1 Ancient Egypt

Summary

- Recorded civilization in Nile valley from c. 3150 BC.
- Conquered 322 BC by Alexander the Great (Greece/Macedonia).
- Became a Roman province in 30 BC under Cleopatra.
- Discovery of the Rosetta Stone (AD 1799) containing Greek, hieroglyphic and demotic (post hieratic, cursive) Egyptian script, allowed translation of ancient Egyptian writing.
- Primary mathematical sources: Rhind/Ahmes (A'h-mose) papyrus c. 1650 BC and the Moscow papyrus c. 1700 BC.[1] Part of the Rhind papyrus is shown below. It contained two tables: unit fraction representations of all $\frac{2}{n}$ for $n < 100$ (ish) and expressions for $\frac{3}{10}, \frac{4}{10}, \ldots, \frac{9}{10}$ in terms of unit fractions. Also included were around 100 worked problems. Scribes would learn the method by copying previous problems and changing the numbers.

  The Moscow papyrus is shorter and more focused on geometric problems.

- Few other primary sources. Egyptians wrote on papyrus (plant-based form of paper) which decomposes. Other Egyptian mathematics was likely absorbed uncredited by the Greeks.
- Practical/non-theoretical: worked problems on sums, linear equations, construction and land-measurement (tax-collection). No clear distinction between exact and approximate solutions.

---

[1] Rhind was a Scottish egyptologist. Ahmes was the name of the scribe who wrote/copied the papyrus. The Moscow papyrus is named because it was sold to the Moscow Museum of Fine Art; its author is unknown.

Notation & Egyptian Fractions

The ancient Egyptians had two distinct systems for enumeration: *hieroglyphic* (dating at least to 5000 BC) and *hieratic* (c. 2000 BC). These changed over time, so we give only one version.

Hierogylphic enumeration Essentially decimal symbols for numerals/powers of 10.

- Could be written in any direction: top-to-bottom, right-to-left, etc., or just lumped together: e.g.

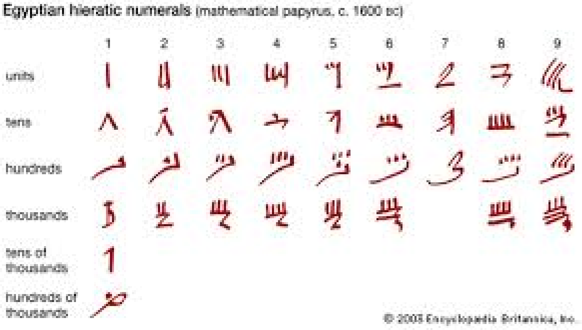
  2349 = |||||||||∩∩∩∩𓍢𓍢𓍢𓆼𓆼

- Slow to write, numbers take up a lot of space.

Numeral	Hieroglyph
1	|
10	∩ (heel bone)
100	𓍢 (snare)
1000	𓆼 (lotus flower)
10000	𓂭 (finger)
100000	𓆐 (fish)
1000000	𓁨 (person)

Hieratic enumeration We will largely ignore this since it is written cursively.

- Different symbols for 1–9, 10–90, 100–900 etc., mapped onto hieratic alphabet.
- System copied later by the Greeks with their own alphabet.
- Pros: less space, easier to write with ink, each number requires fewer symbols.
- Cons: More symbols, slower calculations.


Egyptian hieratic numerals (mathematical papyrus, c. 1600 BC)

For instance, 23 would be written (approximately!) $|||\overline{\wedge}$.

The Eygptians had no numeral for zero, though the hieroglyph *nfr* (beautiful/perfect) was used to denote, for instance, the base floor of a building or to indicate balanced books in accounting.

Fractions Ancient Egyptians worked almost entirely with *reciprocals* of integers ($\frac{1}{n}$ where $n \in \mathbb{N}$). In modern times, any fraction represented as a sum of reciprocals is called an *Egyptian fraction*; their theory is still actively researched. For instance

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{5}$$

is a representation of $\frac{19}{20}$ as an Egyptian fraction.

In hieratic notation, a dot was placed above to denote the reciprocal: e.g., $\dot{\wedge}$ is $\frac{1}{10}$.

In hieroglyphs, a reciprocal was represented by placing an oval over a numeral. We do this with a bar: e.g. $\overline{2} = \frac{1}{2}$. As with integers, combinations of fractions could be written in any order/direction.

The only non-reciprocal fractions with special symbols were $\frac{2}{3}$ and $\frac{3}{4}$, and these only appeared late in Egyptian civilization.[2]

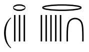
Fraction	Hieroglyph	Modern
$\frac{1}{3}$	$\overline{\|\|\|}$	$\overline{3}$
$\frac{1}{41}$	$\overline{\|\cap\cap\cap\cap}$	$\overline{41}$
$\frac{1}{103204}$	$\overline{\text{\|\|\|\|𓍢𓍢𓆼𓆼𓆼𓆐}}$	$\overline{103204}$
$31+\frac{1}{2}+\frac{1}{25}$	$\overline{\|\|\|\|\|\cap\cap}\ \ \overline{\|\|}\ \ \|\cap\cap\cap$	$\overline{25}\ \overline{2}\ 31$

---

[2] For instance an oval over one-and-a-half sticks for $\frac{2}{3}$, and an oval over three short-long-short sticks for $\frac{3}{4}$.

The Rhind papyrus contains a table, of which we reproduce part, showing how to express $\frac{2}{n}$ as Egyptian fractions for all odd integers $n < 100$. The first column denotes n, and the remaining columns the Egyptian fraction representation. For instance, the first row states

$$\frac{2}{5} = \frac{1}{3} + \frac{1}{15}$$

( $\overline{3}\,\overline{15}$)

5	3	15	
7	4	28	
9	6	18	
11	6	66	
13	8	52	104
15	10	30	

There are several approaches to finding Egyptian fraction representations, and it can be proved that any fraction may be written in such a form. If the denominator is odd, then a simple place to start is with

$$\frac{2}{mn} = \frac{1}{mr} + \frac{1}{nr} \quad \text{where} \quad r = \frac{m+n}{2}$$

Most lines in the Rhind table follow this formula, but not all. Note also that the formula often permits multiple representations.

- The line for $\frac{2}{5}$ has $m = 1$, $n = 5$ and $r = 3$; this is unique up to reordering.
- $\frac{2}{9}$ may be represented

  $$\frac{2}{9} = \begin{cases} \frac{1}{9} + \frac{1}{9} & (m,n,r) = (3,3,3) \\ \frac{1}{5} + \frac{1}{45} & (m,n,r) = (1,9,5) \end{cases}$$

  The first of these is essentially useless and the second is not the expression from the table.
- In the table, $\frac{2}{13}$ is written as a sum of three reciprocals instead of two.

The table reduced the need to divide and sped up the computation of harder fractions: for instance,

$$\frac{5}{13} = 2 \cdot \frac{2}{13} + \frac{1}{13} = 2\left(\frac{1}{8} + \frac{1}{52} + \frac{1}{104}\right) + \frac{1}{13} = \frac{1}{4} + \frac{1}{13} + \frac{1}{26} + \frac{1}{52}$$

Egyptian Calculations & Example Problems

Addition/Subtraction With hieroglyphs this is simple: Write out numbers one above another and count up the symbols! Replace 10 of one by the next symbol. For subtraction, one might need to convert a larger symbol to 10 of a smaller one. Essentially this is 'carrying.' A special symbol was used to denote both addition and subtraction: its meaning changed depending on the direction the text was read.

Multiplication This relied on a base-two algorithm. To compute ab:

1. Write $1, b$
2. Repeatedly double each until the first term is about to exceed a
3. Determine powers of 2 that sum to a
4. Sum the corresponding multiples of b

For example, to compute $13 \cdot 15$, we construct a table where the checked rows are summed.

1	15	✓
2	30	
4	60	✓
8	120	✓

$$\implies 13 \cdot 15 = (1+4+8) \cdot 15 = 15 + 60 + 120 = 195$$

Note how $13 = 1 + 4 + 8 = 2^0 + 2^2 + 2^3$ is essentially the binary representation. We stopped at the fourth row since another doubling would have resulted in the first term (16) exceeding 13. We could instead have reversed the roles of the factors:

1	13	✓
2	26	✓
4	52	✓
8	104	✓

$$\implies 15 \cdot 13 = (1+2+4+8) \cdot 13 = 13 + 26 + 52 + 104 = 195$$

All you need is addition and the ability to multiply by 2!

Division This also relies on doubling/halving, though the answer is non-unique and might require some creativity. To find $\frac{a}{b}$, think about solving the problem $a = bx$ and apply a variant of the multiplication algorithm to find multiples of b summing to a. Here are some examples.

1. To compute $260 \div 13$, we repeatedly double 13 until terms in the right column sum to 260.

1	13	
2	26	
4	52	✓
8	104	
16	208	✓

   Since $260 = 208 + 52$ we conclude that $260 \div 13 = 16 + 4 = 20$

2. To find $5 \div 13$ we start by dividing by 2 with the intent of making terms in the right column sum to 5.

1	13	
$\overline{2}$	$6\,\overline{2}$	
$\overline{4}$	$3\,\overline{4}$	✓
$\overline{8}$	$1\,\overline{2}\,\overline{8}$	✓

   Since $(3\,\overline{4}) + (1\,\overline{2}\,\overline{8}) = 4\,\overline{2}\,\overline{4}\,\overline{8}$ is $\overline{8}$ ($\frac{1}{8}$) short of what we want, we continue the table in a different way. First divide by 13, then continue halving until we obtain the desired $\overline{8}$ in the right column.

$\overline{13}$	1	
$\overline{26}$	$\overline{2}$	
$\overline{52}$	$\overline{4}$	
$\overline{104}$	$\overline{8}$	✓

   We conclude that $5 \div 13 = \overline{4}\,\overline{8}\,\overline{104} = \frac{1}{4} + \frac{1}{8} + \frac{1}{104}$. We could have proceeded differently to obtain the same result as followed from the Rhind table:

$$5 = (3\,\overline{4}) + 1 + \overline{2} + \overline{4} \implies 5 \div 13 = \overline{4}\,\overline{13}\,\overline{26}\,\overline{52}$$

Practical application: Loaf-splitting A typical Egyptian problem might involve determining how to split 5 loaves among 13 people. The previous calculation tells us that we could give each person $\frac{1}{4}+\frac{1}{8}+\frac{1}{104}$ of a loaf. This might seem complicated but it has some advantages over the modern approach where we'd first cut each loaf into 13 equal pieces:

- The large chunks of bread are created by repeatedly cutting in half: this is easy to do accurately, whereas cutting 13th parts is difficult! The remaining 104th parts of a loaf would probably be ignored as crumbs.
- The Egyptian approach only requires 34 cuts, as opposed to 60 in the modern style.

Linear equations Another common type of problem was how to solve linear equations. Solutions were based on the method of *false position.* Essentially one guesses an approximate solution, then modifies it until it works. Here is problem 24 of the Rhind papyrus.

> A heap plus a seventh of a heap is nineteen. What is the heap?

In modern algebra, representing 'heap' by x, we wish to solve $x+\frac{1}{7}x=19$. Here is the Egyptian method.

1. Guess intelligently: $x=14$ is easy to divide by 7 and we obtain $x+\frac{1}{7}x=16$
2. Correct our guess: We want 19, not 16, so we multiply our guess (14) by $\frac{19}{16}=1\,\overline{8}\,\overline{16}$ to obtain the correct answer

1	$1\,\overline{8}\,\overline{16}$	
2	$2\,\overline{4}\,\overline{8}$	✓
4	$4\,\overline{2}\,\overline{4}$	✓
8	$9\,\overline{2}$	✓

$$x=2\,\overline{4}\,\overline{8}+4\,\overline{2}\,\overline{4}+9\,\overline{2}=16\,\overline{2}\,\overline{8}$$

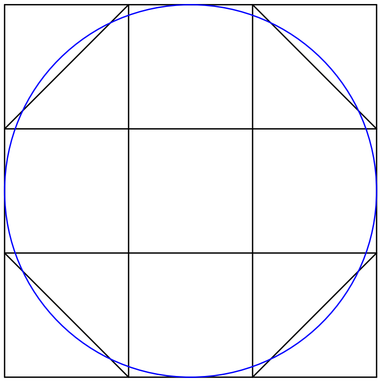
Compare this with the 'modern' method:

$$x+\frac{1}{7}x=19\implies\frac{8}{7}x=19\implies x=\frac{7\cdot 19}{8}=\frac{133}{8}=\frac{128+5}{8}=16\frac{5}{8}$$

Are the any benefits to the Egyptian approach?

Geometry Problem 48 in the Rhind papyrus involves using an octagon to approximated the area of a circle. A square of side 9 is drawn, where each side is split into thirds and the four corner squares are cut in half. The area of the octagon is then

$$81-18=63$$

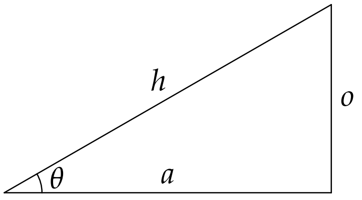


Since the area of the circle is $\frac{81\pi}{4}$, this amounts to the approximation $\pi\approx\frac{28}{9}=3.11111\ldots$

Problem 50 compares the same circle to a square of side 8: essentially $\frac{81\pi}{4}\approx 64$ which corresponds to $\pi\approx\frac{256}{81}=3.16049\ldots$

No explanation is given as to what inspired these methods, nor whether the scribes understood that these were only approximations.

Other problems computed/approximated areas and volumes of triangles, quadrilaterals, boxes, cylinders, and truncated pyramids. Again, these were worked examples without general formulæ. The Egyptians even had a notion of cotangent which they called the *seked*, useful for describing and calculating slopes.



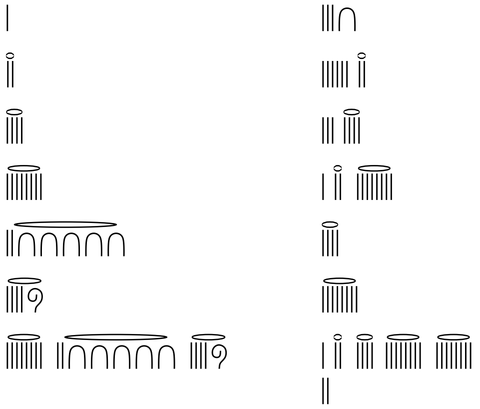
seked $= \frac{a}{o} = \cot\theta$

Exercises Most of these problems are taken from the 'official' textbook.

1. Use Egyptian techniques to multiply 34 by 18 and to divide 93 by 5.
2. Use Egyptian techniques to multiply $\overline{28}$ by $1 + \overline{2} + \overline{4}$ (problem 9 of the Rhind papyrus).
3. A part of the Rhind papyrus table for division by 2 reads as follows:

$$2 \div 11 = \overline{6} + \overline{66}, \qquad 2 \div 13 = \overline{8} + \overline{52} + \overline{104}, \qquad 2 \div 23 = \overline{12} + \overline{276}$$

The calculation of $2 \div 13$ is given below, where the right hand side is a modern rendering, and the terminal '2' indicates that the last entry in the right-hand column is indeed 2:



1	13
$\overline{2}$	$6\ \overline{2}$
$\overline{4}$	$3\ \overline{4}$
$\overline{8}$	$1\ \overline{2}\ \overline{8}$
$\overline{52}$	$\overline{4}$
$\overline{104}$	$\overline{8}$
$\overline{8}\ \overline{52}\ \overline{104}$	$1\ \overline{2}\ \overline{4}\ \overline{8}\ \overline{8}$ 2

Perform similar calculations for $2 \div 11$ and $2 \div 23$.

(*If you want the exact results from the papyrus, you'll need to work with* $\frac{2}{3}$*: denote this by* $\overline{\overline{3}}$)

4. Draw a picture of 5 loaves to help describe how the Egyptians might have divided them between seven people.
5. Use the method of false position to solve problem 28 of the Rhind papyrus:

> A quantity and its 2/3 are added together, and from the sum 1/3 of the sum is subtracted, and 10 remains. What is the quantity?

6. Calculate a quantity such that if it is taken two times along with the quantity itself, the sum comes to 9 (problem 25 of the Moscow papyrus).
7. (a) Find all ways in which $\frac{2}{13} = \frac{1}{a} + \frac{1}{b}$ can be written as a sum of reciprocals ($a \leq b \in \mathbb{N}$).
   (b) Repeat your calculation for $\frac{2}{p}$ whenever p is an odd prime.

2 Babylon/Mesopotamia

Babylon was an ancient city located near modern-day Baghdad, Iraq. The term also serves as a shorthand for the many Mesopotamian[3] empires/civilizations dating back at least to 3000 BC: Sumeria, Akkadia, Babylonia, etc.

Babylonians used *cuneiform* (wedge-shaped) script, typically indentations on clay tablets. Most recovered tablets date from the time of Hammurabi (c. 1800 BC) or the Seleucid dynasty (c. 300 BC) which ruled after the conquests of Alexander the Great.

Mathematical tablets are of two main types: tables of values (multiplication, reciprocals, measures) and worked problems.

Sexagesimal (base-60) Positional Enumeration Our modern decimal system is *base-10 positional*. This means two things, both of which are easy to see with reference, say, to the number 3835.

- The symbol 3 represents both 3000 and 30: the **meaning of a symbol depends on its position**.
- The position of a symbol denotes the **power of 10** by which it should be multiplied. Thus

$$3835 = 3 \cdot 10^3 + 8 \cdot 10^2 + 3 \cdot 10^1 + 5 \cdot 10^0$$

Positional enumeration makes for efficient calculations and easy representation of numbers of vastly different magnitude. Contrast with the difficulty of performing calculations using (non-positional) Egyptian hieroglyphic notation.

One of the key Babylonian contributions to mathematical history is the creation of (arguably) the first positional system of enumeration, dating to at least 2000 BC. Rather than our ten symbols 0–9, Babylonians used only two: roughly $\vee$ for 1 and $\lhd$ for 10, likely made by the same stylus. Any number up to 59 could be written with combinations of these symbols, e.g.,

$$53 = \begin{smallmatrix}\lhd\lhd\lhd\\ \lhd\lhd\end{smallmatrix}\vee\vee\vee$$

The picture shows a typical cuneiform representation. Larger numbers were represented base-60. For instance, the sexagesimal decomposition of 3835 is

$$3835 = 1 \cdot 60^2 + 3 \cdot 60^1 + 55 \cdot 60^0$$

which the Babylonians would have written

$$\vee \qquad \vee\vee\vee \qquad \begin{smallmatrix}\lhd\lhd\lhd\\ \lhd\lhd\end{smallmatrix}\begin{smallmatrix}\vee\vee\vee\\ \vee\vee\end{smallmatrix} \tag{$*$}$$

Rather than using cuneiform, we'll instead write 1, 3, 55;

---

[3] 'Between two rivers,' namely the Tigris and Euphrates. As indicated on the map, these rivers formed the backbone of the *fertile crescent*, a region of early civilization, farming, crop and animal domestication.

Just as, for us, '3' might mean 30, 3000 or even $\frac{3}{1000}$, for the Babylonians ∨ could mean 1, 60, 3600, 216000, or fractions such as $\frac{1}{60}$, $\frac{1}{3600}$ depending on its position. There was no symbol for zero (as a placeholder) until very late in Babylonian history, nor any *sexagesimal point*, so determining position on ancient tablets can be difficult. For instance, rather than 3835, (∗) might instead have represented

$$60+3+\tfrac{55}{60}=63\tfrac{11}{12} \quad \text{or} \quad 60^3+3\cdot 60^2+55\cdot 60=230100$$

To make things easier to read, we use commas to separate terms and, if necessary, a semicolon to denote the sexagesimal point. Thus

$$23,12,0;15=23\cdot 60^2+12\cdot 60+\frac{15}{60}=83520\frac{1}{4}$$

Why base-60? There are many theories, but we cannot be certain. Here are some ideas.

- The Babylonians might have combined two systems (base-10 and base-12) inherited from older cultures.
- Since 60 has many proper divisors (1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30), more numbers have exact representation than with decimal arithmetic: for instance $\frac{1}{3}=;20$, as a terminating sexagesimal, is much simpler than the decimal 0.33333.
- As prolific astronomers and astrologers, the Babylonians might have chosen 60 as a divisor of 360, approximately the number of days in a year. Our modern usage of *degrees-minutes-seconds* for angle, *hours-minutes-seconds* for time, and the standard zodiac are all of Babylonian origin. Indeed Babylonian units of measure often used factors of 60 for magnitude similarly to how modern science uses 1000 (e.g., joules → kilojoules → megajoules).

This sort of historical question is rarely answerable in a satisfying way. Likely no-one 'decided' to use base-60; like most cultural issues, it likely happened slowly and organically, without fanfare.

Basic Sexagesimal Calculations Addition and subtraction would have been as natural to the Babylonians as decimal calculations are to us. For instance, we might write

$$\begin{array}{rr} & 21,49 \\ + & \underset{1}{3},37 \\ \hline & 25,26 \end{array} \qquad \text{(in decimals } 1309+217=1526\text{)}$$

Note how we carry 60 just like we are used to doing with 10 in decimal arithmetic: $49+37=1,26$.

Multiplication is significantly harder. To mimic our familiar long-multiplication process would require memorizing up to the 59 times table! For small factors this might have been fine. For larger factors there is evidence of the Babylonians using two representations of a product in terms of squares

$$xy=\frac{1}{2}\left[(x+y)^2-x^2-y^2\right]=\frac{1}{4}\left[(x+y)^2-(x-y)^2\right]$$

Tablets consisting of tables of squares greatly aided the computation of large products. For instance,

$$31\times 22=\frac{1}{4}\left[53^2-9^2\right]=\frac{1}{4}\left[46,49-1,21\right]=\frac{1}{4}\left[45,28\right]=11,7+15=11,22 \qquad (=682)$$

Fractions & Division As we've already seen, the Babylonians also represented non-integers using sexagesimals. Tables of reciprocals $\frac{1}{n}$ were used to quickly evaluate division using multiplication!

$$m \div n = m \times \frac{1}{n}$$

For example

$$\frac{1}{18} = 0;3,20 \implies \frac{23}{18} = 23(0;3,20) = 1;9 + 0;7,40 = 1;16,40$$

This works nicely provided n has no prime divisors other than 2, 3 or 5, since any such $\frac{1}{n}$ will be an exact terminating sexagesimal.[4] Approximations were used for other reciprocals; a scribe would choose a nearby denominator with an exact sexagesimal and state that the answer was approximate

$$\frac{11}{29} \approx \frac{11}{30} = 11(0;2) = 0;22$$

For more accuracy, one could choose a larger denominator. For instance, if a scribe wanted to divide by 11, they might observe that $11 \cdot 13 = 143 \approx 144$, from which[5]

$$\frac{1}{144} = 0;0,25 \implies \frac{1}{11} \approx \frac{13}{144} = 0;5,25$$

Scribes were explicit in acknowledging the approximation by stating, say, "11 does not divide." Remember that a single digit in the second sexagesimal place means only $\frac{1}{3600}$, so even the most demanding application doesn't require many terms (the above is 99.3% accurate!). The denominators in some of these reciprocal tables were enormous, so far greater accuracy was often possible.

Another table listed all the ways an integer < 10 could be multiplied exactly to get 10.

1	10	5	2
2	5	6	1 40
3	3 20	8	1 15
4	2 30	9	1 6 40

We omit the commas for separation and the sexagesimal point as they did not exist. Moreover 7 is missing since $\frac{1}{7}$ (and thus $\frac{10}{7}$) is not an exact sexagesimal. It should be clear from the table that

$$\frac{10}{6} = 1;40 \quad \text{and} \quad \frac{600}{9} = 1,6;40$$

In the latter case, note that $600 = 10 \cdot 60$ would be written the same as 10, so this amounts to moving the sexagesimal point in $\frac{10}{9} = 1;6,40$.

---

[4] Analogous to the fact that $\frac{1}{n}$ has a terminating decimal if and only if n has no prime divisors other than 2 or 5.

[5] Being rational, $\frac{1}{11} = 0.09090909\ldots = 0;5,27,16,21,49,5,27,16,21,49,\ldots$ has a periodic sexagesimal expansion as can be found using a pocket-calculator:

$$\frac{60}{11} = 5 + \frac{5}{11}, \quad \frac{5 \cdot 60}{11} = 27 + \frac{3}{11}, \quad \frac{3 \cdot 60}{11} = 16 + \frac{4}{11}, \quad \frac{4 \cdot 60}{11} = 21 + \frac{9}{11}, \ldots$$

Linear Systems of Equations These could be solved by a combination of the method of false position (guess and modify as per the Egyptians) and the consideration of homogeneous equations. For instance, here is a (suitably modernized) Babylonian approach to solving the system

$$\begin{cases} 3x + 2y = 11 \\ 2x + y = 7 \end{cases}$$

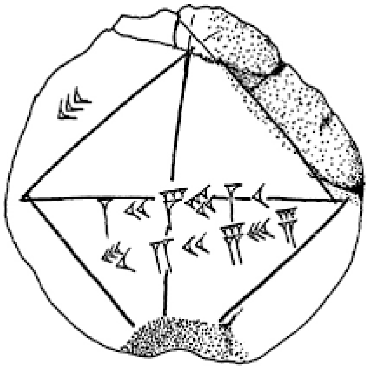
1. Choose one equation, say the second, and set $\hat{x} = \hat{y}$. Solve this (say using false position) to obtain $\hat{x} = \hat{y} = \frac{7}{3} = 2;20$.
2. Since $(d, -2d)$ is the general solution to the homogeneous equation $2x + y = 0$, all solutions to the second equation have the form $x = \hat{x} + d$ and $y = \hat{y} - 2d$. Substitute into the first equation:
$$11 = 3\left(\frac{7}{3} + d\right) + 2\left(\frac{7}{3} - 2d\right) = 11 + \frac{2}{3} - d \implies d = \frac{2}{3}$$
3. Finally compute $x = \hat{x} + d = \frac{7}{3} + \frac{2}{3} = 3$ and $y = \hat{y} - 2d = \frac{7}{3} - \frac{4}{3} = 1$.

Step 2 should should remind you of the 'nullspace' method from modern linear algebra: all solutions to the matrix equation $(2\ 1)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = 7$ have the form

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + \mathbf{n}$$

where $\left(\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right)$ is some particular solution (here $x_0 = y_0 = \frac{7}{3}$) and $\mathbf{n} = \left(\begin{smallmatrix} d \\ -2d \end{smallmatrix}\right)$ lies in the nullspace of the (row) matrix $(2\ 1)$.

The Yale Tablet (YBC 7289) and Square-root Approximations One of the most famous tablets concerns an approximation to $\sqrt{2}$. YBC stands for the *Yale Babylonian Collection* which contains over 45,000 objects. YBC 7289 is shown below alongside an enhanced representation of the numerals.



The tablet depicts a square of side 30 and labels the diagonal in two ways:

- $1;24,51,10$ is an approximation to $\sqrt{2}$, an underestimate by roughly 1 part in 2.5 million!
- $42;25,35$ is an approximation to the diagonal when the side is 30.

The Babylonians more often used the simpler approximation $1;25 = 1.41666\ldots$ which is still very close. Given the impractical accuracy of YBC 7289, it is reasonable to ask how it was found. No-one knows for certain, but two methods are theorized since both were used to solve other problems. It should be stressed that no Babylonian *proofs* of these approaches are known.

1: Square root approximation $\sqrt{a^2 \pm b} \approx a \pm \frac{b}{2a}$. This is essentially the linear approximation from elementary calculus. If one chooses a rational number a whose square is close to 2, then the error will also be small. For instance:

- $\sqrt{2} = \sqrt{1+1} \approx 1 + \dfrac{1}{2} = 1;30$ $\qquad (a = 1)$
- $\sqrt{2} = \sqrt{\left(\dfrac{4}{3}\right)^2 + \dfrac{2}{9}} \approx \dfrac{4}{3} + \dfrac{2/9}{8/3} = \dfrac{4}{3} + \dfrac{1}{12} = \dfrac{17}{12} = 1;25$ $\qquad (a = \frac{4}{3} = 1.3333\ldots)$
- $\sqrt{2} = \sqrt{\left(\dfrac{7}{5}\right)^2 + \dfrac{1}{25}} \approx \dfrac{7}{5} + \dfrac{1/25}{14/5} = \dfrac{99}{70} = 1;24,51,25,42,51,25,42,\ldots$ $\qquad (a = \frac{7}{5} = 1.4)$

2: Method of the Mean It may be checked (Exercise 11) that any sequence defined by the recurrence relation $a_{n+1} = \frac{1}{2}\left(a_n + \frac{2}{a_n}\right)$ converges to $\sqrt{2}$. We apply this when $a_1 = 1$.

$$a_1 = 1, \quad a_2 = \frac{3}{2} = 1;30, \quad a_3 = \frac{17}{12} = 1;25$$

$$a_4 = \frac{577}{408} = 1 + \frac{169}{408} = 1;24,51,10,35,17,\ldots$$

$$a_5 = \frac{665857}{470832} = 1;24,51,10,7,46\ldots$$

It seems incredible that any ancient culture would have bothered to go as far as this to obtain the observed accuracy.

The same approach can be used to approximate other roots. For example, we can approximate $\sqrt{11}$ via $a_{n+1} = \frac{1}{2}(a_n + \frac{11}{a_n})$ and $a_1 = 3$:

$$a_2 = \frac{10}{3} = 3;20, \quad a_3 = \frac{199}{60} = 3;19, \quad a_4 = \frac{79201}{23880} = 3;18,59,50,57,17,\ldots$$

Quadratic Equations The above methods were applied to solve general quadratic equations. A question might be phrased as follows:

> I added twice the side to the square; the result is 2, 51, 40. What is the side?

In modern language, we want the solution to $x^2 + 2x = 2 \cdot 60^2 + 51 \cdot 60 + 40 = 10300$.

Questions such as these were solved using templates, typically as worked examples. The above problem requires the template for solving $x(x+p) = q$ where $p, q > 0$. Since the Babylonians did not recognize negative numbers, the other types of quadratic equation ($x^2 = px + q$, etc.) had different templates.

To make things a little easier, we apply their approach to the simpler equation $x^2 + 4x = 2$:
Set $y = x + p$ $(y = x + 4)$ and decouple the equation:

$$\begin{cases} xy = q \\ y - x = p \end{cases} \qquad\qquad \begin{cases} xy = 2 \\ y - x = 4 \end{cases}$$

Use this to solve for $x + y$:

$$\begin{aligned} 4xy + (y-x)^2 &= p^2 + 4q & \qquad 4xy + (y-x)^2 &= 4^2 + 4 \cdot 2 \\ (y+x)^2 &= p^2 + 4q & (y+x)^2 &= 24 \\ x + y &= \sqrt{p^2 + 4q} & x + y &= \sqrt{24} \approx \ 4;54 \end{aligned}$$

where the square-root was approximated using one of the earlier algorithms, e.g.

$$\sqrt{24} = \sqrt{5^2 - 1} \approx 5 - \frac{1}{10} = 4;54$$

Since $x + y$ and $x - y$ are now known, we have a linear system which is easily solved:

$$x = \frac{\sqrt{p^2 + 4q} - p}{2} \qquad\qquad x \approx 0;27$$

The method of completing the square and the quadratic formula are at least 4000 years old!

While we've written this abstractly, in practice scribes would be copying from a particular example of the same type. There were no abstract formulæ and everything was done without the benefit modern notation. There was moreover typically no written commentary to explain the method; often all historians have to work with is a single column of numbers!

Note also that the template only found the positive solution; the Babylonians had no notion of negative numbers. Amazingly, they were able to address certain cubic equations similarly.

Pythagorean Triples The Plimpton 322 tablet (also at Yale) lists a large number of Pythagorean triples (albeit with some mistakes). Due to the strange manner of encoding, it took scholars a long time to realize what they had.

As an example, line 15 describes the Pythagorean triple $53^2 = 45^2 + 28^2$:

- The first entry 1;23,13,46,40 is the exact sexagesimal for $\left(\frac{53}{45}\right)^2$.
- The second entry is 28.
- The third entry is 53.
- The last two entries indicate line number 15.

The first three (interesting) entries are therefore $\left((\frac{c}{a})^2, b, c\right)$ where $c^2 = a^2 + b^2$. Since the table is broken on the left side it is possible that a missing column explicitly mentioned a.

It is not known how the table was completed, though the first column exhibits a descending pattern that provides clues to its construction. One theory is that a scribe found rational solutions to the equation $v^2 = 1 + u^2$ (equivalently $(v+u)(v-u) = 1$) by starting with a choice of $v+u$ and using a table of reciprocals to calculate $v-u$.

To revisit our example, if $v + u = \frac{9}{5} = 1;48$, then

$$v - u = \frac{1}{v+u} = \frac{5}{9} = 0;33,20$$

We therefore have a linear system of equations in u, v whose solutions are

$$v = 1;10,40 = \frac{53}{45}, \quad u = 0;37,20 = \frac{28}{45}$$

We investigate this further in Exercise 7. The Plimpton tablet has been the source of enormous scholarship; look it up!

Geometry The Babylonians also considered many geometric problems. They used both $\pi \approx 3$ and $\pi \approx 3\frac{1}{8}$ to approximate areas of circles. They had calculations (both correct and erroneous) for the volume of a frustrum (truncated pyramid). They also knew that the altitude of an isosceles triangle bisects its base, and that the angle in semicircle is a right-angle (Thales' Theorem). None of these statements were presented as theorems in a modern sense; we merely have computations and applications that make use of these facts. We simply do not know the depth of Babylonian understanding of such concepts.

Summary

- Sexagesimal positional enumeration. No zero. Fractions also used sexagesimal representation.
- More advanced than Egyptian mathematics but still practical/non-abstract. Perhaps Babylonian mathematics only *appears* more advanced because we have so much more evidence: thousands of tablets versus only a handful of papyri. Like Egypt, we have worked examples without abstraction or statements of general principles.
- Some distinction ('does not divide') between approximate and exact results.
- Some geometry, but algorithmic/numerical methods predominate.

Exercises There is no single correct way to do Babylonian calculations. Play with the ideas and use modern notation to get a feel for things without torturing yourself.

1. Convert the sexagesimal values 0;22,30, 0;8,6, 0;4,10 and 0;5,33,20 into ordinary (modern) fractions in lowest terms.
2. (a) Multiply 25 by 1,4 (b) Multiply 18 by 1,21.
   (*Either compute directly (long multiplication) or use the difference of squares method on page 8*)
3. (a) Use reciprocals to divide 50 by 18. (b) Repeat for 1,21 divided by 32.
4. Use the Babylonian method of false position to solve the linear system $\begin{cases} 3x + 5y = 19 \\ 2x + 3y = 12 \end{cases}$

5. (a) Convert the approximation $\sqrt{2} \approx 1;24,51,10$ to a decimal and verify the accuracy of the approximation on page 10.
   (b) Multiply by 30 to check that the length of the diagonal is as claimed.

6. Babylonian notation is not required for this question.
   (a) Use the square root approximation (pg. 11) with $a = \frac{8}{3}$ to find an approximation to $\sqrt{7}$.
   (b) Taking $a_1 = 3$, apply the method of the mean to find the approximation a_3 to $\sqrt{7}$.

7. Recall that $v^2 = 1 + u^2$ in the construction of the Plimpton tablet.
   (a) If $v + u = \alpha$, show that $u = \frac{1}{2}(\alpha - \alpha^{-1})$ and $v = \frac{1}{2}(\alpha + \alpha^{-1})$.
   (b) Suppose $v + u = 1;30 = \frac{3}{2}$. Find u, v and the corresponding Pythagorean triple.
   (c) Repeat for $v + u = 1;52,30 = \frac{15}{8}$.
   (d) Repeat for $v + u = 2;05 = \frac{25}{12}$. This is line 9 of the tablet.

8. Solve the following problem from tablet YBC 4652. I found a stone, but did not weigh it; after I subtracted one-seventh, added one-eleventh (of the difference), and then subtracted one-thirteenth (of the previous total), it weighed 1 *mina* (= 60 *gin*). What was the stone's weight?

   (*The meaning of the problem isn't completely clear: make your best guess!*)

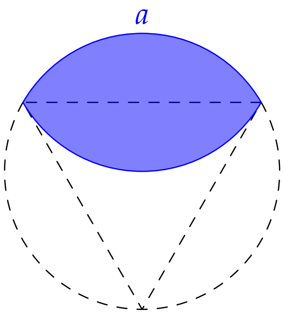
9. Solve the following problem from tablet YBC 6967. A number exceeds its reciprocal by 7. Find the number and the reciprocal.

   (*In this case, two numbers are reciprocals if their product is 60*)

10. For this question it is helpful to think about the corresponding facts for decimals.
    (a) Explain the observation on page 9 regarding which reciprocals n have a terminating sexagesimal. Can you prove this?
    (b) Find the periodic sexagesimal representation of $\frac{1}{7}$ and use geometric series *prove* that you are correct.

11. For this question, look up the AM–GM inequality and remind yourself of some basic Analysis.
    (a) Suppose (a_n) is a sequence satisfying the recurrence $a_{n+1} = \frac{1}{2}(a_n + \frac{2}{a_n})$. Prove that $a_n \geq \sqrt{2}$ whenever $n \geq 2$.
    (b) Prove that $\lim a_n = \sqrt{2}$.

12. (a) The Babylonians used an approximation of the form $A \approx \frac{1}{12}c^2$ for the area of a circle in terms of its circumference. To what approximation for π does this correspond?
    (b) Use a Babylonian method to show that $\sqrt{3} \approx \frac{7}{4}$.
    (c) A 'bulls-eye' (pictured) is constructed using congruent circular arcs built from a circumscribed equilateral triangle. If the arc-length is a, use Babylonian approximations to prove that the area of the bulls-eye is $\approx \frac{9a^2}{32}$. What, approximately, are its dimensions (width/height)?
    (*Use as much modern trigonometry as you like!*)



3 Ancient Greek Mathematics

3.1 Overview of Ancient Greek Civilization & Early Philosophy

Ancient Greek civilization dates from around 800 BC, centered on the peninsula between the Adriatic sea (to the west) and the Aegean (to the east) that forms the mainland of the modern Greek state. Ancient Greek culture was decentralized, consisting of semi-independent city-states connected by trade. Accomplished sea-farers, they extended their reach round the northern and eastern coasts of the Mediterranean, from Iberia (Spain) to the Black Sea, Anatolia (western Turkey), and Egypt, building and capturing a network of city-states and trading outposts.

Philip of Macedon 'unified' the Greek peninsula just before his death in 336 BC. His son, Alexander 'the Great,' launched a massive campaign conquering Persia, Egypt, Babylon, and western India before his own death in Babylon in 323 BC. Alexander left provincial governors to manage captured territory; some of these structures endured for centuries (Egypt's Ptolemaic dynasty), whereas others were overthrown after only a few years (parts of India). While Alexander's conquests did not produce a long-lasting centralized Greek empire, they were effective at expanding the reach of Greek cultural practice and philosophy, and brought external ideas into the Greek tradition.

The core part of Greek territory was absorbed by the Roman empire around 146 BC. In line with typical Roman practice, those Greeks who agreed to accept Roman governors and taxation became Roman citizens.[6] As such, the Greek culture of inquiry and scholarship was left largely intact under Roman rule. Greek culture and learning was central to the later Byzantine (eastern Roman) empire (centered on Byzantium/Constantinople/Istanbul), which lasted until Constantinople was captured by the Islamic Ottomans in 1453. For centuries prior, Islamic scholarship had itself been significantly influenced by the ancient Greeks; the main consequence of the fall of Constantinople for knowledge transfer was to encourage the exodus of scholars to Rome, helping to fuel the nascent European Renaissance.

Greek Territory c. 500 BC

---

[6] While this might sound reasonable, resistance wasn't a realistic option. . .

Greek mathematics is part of a much wider development of science and philosophy encompassing a change of emphasis from practicality to abstraction. One reason for this was the Greek blending of religion/mysticism with natural philosophy: a desire to describe the natural world while preserving the perfection/logic in the gods' design.

Early Greek inquiry into natural phenomena was encouraged through the personification of nature (e.g., sky = man, earth = woman). By 600 BC, philosophers were attempting to describe such phenomena in terms of natural predictable causes and structures. For example, some viewed matter as being comprised of the 'four elements' (fire, earth, water, air) combined in the correct proportions. While the system of the world was seen as divinely-designed, explanations relying on the whims of the gods were discouraged.

While the Greeks certainly used mathematics for practical purposes, philosophers idealized logic and were unhappy with approximations. This led to the development of *axiomatics, theorems* and *proof,* concepts for which there is scant pre-Greek evidence. The ancient Greek language is indeed the source of three words of critical importance:

Mathematics From *mathematos* (μαϑήματος), meaning knowledge or learning; the term covered essentially anything that might be taught in Greek schools.

Geometry Literally *earth-measure*, a combination of two terms:

> *Gi* (γη) Dates from pre-5th century BC, meaning *land, earth* or *soil.* Capitalized (Γη) it could refer to the *Earth* (as a goddess).
>
> *Metron* (μέτρον) A *weight* or *measure*, a *dimension* (length, width, etc.), or the *metre* (rhythm) in music.

Theorem From *theoreo* (ϑεώρέω), meaning 'I contemplate/consider.' In a mathematical context this become *theorema* (ϑεώρημα): a proposition to be proved.

Ancient Greece had several schools, mostly private and open only to men. Typically arithmetic was taught until age 14, followed by geometry and astronomy until age 18. The most famous scholars of ancient Greece were the Athenian trio of Socrates, Plato and Aristotle,[7] whose writings became central to the Western/Islamic philosophical tradition. Plato's *Academy* in Athens was a model for centuries of schooling; the centrality of geometry to the curriculum was evidenced by the famous inscription above its entryway: "Let none ignorant of geometry enter here."

Ancient Greek Enumeration

The Greeks had two primary forms of enumeration, both dating from around 800–500 BC.

In *Attic Greek* (Attica = Athens) strokes were used for 1–4, with larger numerals using the first letter of the words for 5, 10, 100, 1000 and 10000. For example,

- Πεντε (pente) is Greek for five, whence Π denoted the number 5.
- Δεϰα (deca) means ten, so Δ represented 10.
- H (hekaton), X (khilias) and M (myrion/myriad) denoted 100, 1000 and 10000 respectively.
- Larger numbers were written using combinations of these symbols, similarly to both Egyptian hieroglyphs and (the later) Roman numerals: e.g., $\mathrm{XHH}\Pi|| = 1207$.

---

[7] Each taught his successor, with the birth of Socrates to the death of Aristotle covering 470–322 BC.

Ionic Greek (Ionia = mid Anatolian coast) numerals used the Greek alphabet, an approach possibly copied from Egyptian hieratic enumeration. Larger numbers used a left subscript mark (like a comma) to denote thousands and/or M (with superscripts) for 10000 as in Attic Greek. For example,

$$35298 = \text{,λ,εσϟη} = \overset{\gamma}{\text{M}}\text{,εσϟη}$$

The ancient Greek alphabet included three archaic symbols ϛ ϟ ϡ (stigma, qoppa, sampi), with which you're likely unfamiliar.

1	α	10	ι	100	ρ
2	β	20	κ	200	σ
3	γ	30	λ	300	τ
4	δ	40	μ	400	υ
5	ε	50	ν	500	ϕ
6	ϛ	60	ξ	600	χ
7	ζ	70	o	700	ψ
8	η	80	π	800	ω
9	θ	90	ϟ	900	ϡ

The Greeks also used Egyptian fractions, denoting reciprocals with an accent over the symbol: e.g., $\acute{\vartheta} = \frac{1}{9}$. The use of Egyptian fractions dominated in Europe well into the middle ages.

Ionic Greek enumeration has persisted, with few changes, into modern times, although Hindu–Arabic numerals are also in common usage. Eventually a bar was placed over numbers to distinguish them from words (e.g., $\overline{\xi\vartheta} = 89$), while modern practice is to insert a *keraia* (similar to an apostrophe) at the end of a number: thus 35298 = ͵λ͵εσϟη′.

Both the Attic and Ionic systems were suitable for record-keeping but terrible for calculations! Later Greek mathematicians adapted the Babylonian sexagesimal system for calculation purposes, helping cement its modern use of in astronomy, navigation and time-keeping.

Exercises 3.1. 1. State the number 1789 in both Attic and Ionic notation.

2. Represent $\frac{8}{9}$ as a sum of distinct unit fractions (Egyptian style). Express the result in (Ionic) Greek notation.

   (*The answer to this problem is non-unique*)

3. For tax purposes, the ancient Greeks would approximate the area of a quadrilateral field by multiplying the averages of the two pairs of opposite sides. In one example, the two pairs of opposite sides were given as

   $$\begin{aligned} a &= \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} \quad \text{opposite} \quad c = \frac{1}{8} + \frac{1}{16}, \quad \text{and,} \\ b &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \quad \text{opposite} \quad d = 1 \end{aligned}$$

   where the lengths are in fractions of of a *schonion*, a measure of approximately 150 feet. Find the average of a and c, the average of b and d, and thus the approximate area of the field in square *schonion*. The taxman then rounds up the answer to collect a little more tax!

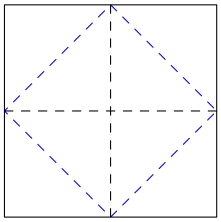
3.2 Pre-Euclidean Greek Mathematics

The publication of Euclid's *Elements* (c. 300 BC) forms a natural breakpoint in Greek mathematics, since is subsumed much of what came before it. In this section, we consider the contributions of several pre-Euclidean mathematicians. There are very few sources for Greek mathematics & philosophy before 400 BC, so almost everything is inferred from the writings and commentaries of others.[8]

Thales of Miletus (c. 624–546 BC) One of the first people known to state abstract general principles, Thales was a trader based in Miletus, a city-state in Anatolia. He travelled widely and was likely exposed to mathematical ideas from all round the Mediterranean. Here are some statements at least partly attributed to Thales:

- The angles at the base of an isosceles triangle are equal.
- Any circle is bisected by its diameter.
- A triangle inscribed in a semi-circle is right-angled (still known as Thales' Theorem).

Thales' major development is *generality*: his propositions concern *all* triangles, circles, etc. The Babylonians & Egyptians had examples of such results, but we have little indication that they viewed these abstractly. In accordance with the Greek idea of a theorem ('to look at'), Thales' reasoning was almost certainly pictorial. As an example of contemporary geometric reasoning, by 425 BC Socrates could describe how to halve/double the area of a square by joining the midpoints of edges.



Pythagoras of Samos (c. 572–497 BC) Like Thales, Pythagoras travelled widely, eventually settling in Croton (southeast Italy) where he founded a school lasting 100 years after his death. Plato is believed to have learned much of his mathematics from a Pythagorean named Archytas.

The Pythagoreans practiced a mini-religion whose ideas lay outside the mainstream of Greek society.[9] One of their mottos, "All is number," emphasised their belief in the centrality of pattern and proportion. The following quote[10] gives some flavor of the Pythagorean way of life.

> After a testing period and after rigorous selection, the initiates of this order were allowed to hear the voice of the Master [Pythagoras] behind a curtain; but only after some years, when their souls had been further purified by music and by living in purity in accordance with the regulations, were they allowed to see him. This purification and the initiation into the mysteries of harmony and of numbers would enable the soul to approach [become] the Divine and thus escape the circular chain of re-births.

The Pythagoreans were particularly interested in musical harmony and its relationship to number. They are credited with relating musical intervals to the ratios of lengths of vibrating strings:

- Identical strings whose lengths are in the ratio 2:1 vibrate an *octave* apart.
- A *perfect fifth* corresponds to the ratio 3:2.
- A *perfect fourth* corresponds to the ratio 4:3.

The use of these ratios to tune musical instruments is still known as *Pythagorean tuning.*

---

[8] For instance, most of our knowledge of Socrates comes from the voluminous writings of Plato and Aristotle. The earliest known Greek textbook/compilation (*Elements of Geometry*) was written around 430 BC by Hippocrates of Chios; no copy survives, though most of its material probably made it into Book I of Euclid.

[9] They were vegetarians, believed in the transmigration of souls, and accepted women as students; controversial indeed!

[10] Van der Waerden, *Science Awakening* pp 92–93

Theorems 21–34 in Book IX of Euclid's *Elements* are Pythagorean in origin. For instance:

Theorem. **(IX .21)** *A sum of even numbers is even.*

(IX .27) *Odd less odd is even.*

The Pythagoreans also studied perfect numbers, those which equal the sum of their proper divisors (e.g., $6 = 1 + 2 + 3$), and they seem to have observed the following famous result.

Theorem (IX. 36). *If $2^n - 1$ is prime then $2^{n-1}(2^n - 1)$ is perfect.*

They moreover considered square and triangular numbers ($\frac{1}{2}m(m+1)$) and tried to express geometric shapes as numbers, in service of their belief that all matter could be formed from basic shapes.

Incommensurability and the Pythagorean Theorem As with other ancient cultures, the only numbers in Greek mathematics were *positive integers.* These were used to *compare* lengths/sizes of objects.

Definition. Lengths are in the ratio $m : n$ if some sub-length divides exactly m times into the first and n times into the second.
Two lengths are *commensurable* if some sub-length divides exactly into both.

Ratio 3 : 2

While modern mathematics has no problem with *irrational ratios* (e.g., the diagonal of a square to its side is $\sqrt{2} : 1$), this conflicted with the core Pythagorean belief that *any two lengths were commensurable.* Identifying lengths with real numbers (underlined), we restate their assertion in modern language:

$$\forall \underline{m}, \underline{n} \in \mathbb{R}^+, \ \exists \underline{\ell} \in \mathbb{R}^+, \ \exists a, b \in \mathbb{N}, \text{ such that } \underline{m} = a\underline{\ell} \text{ and } \underline{n} = b\underline{\ell}$$

This is complete nonsense, for it insists that every ratio of real numbers $\underline{m} : \underline{n} = a : b$ is *rational*!

The Pythagorean commensurability supposition stems from their basic tenets: all is number (including length ratios), and the design of the gods is perfect (number means positive integer). The discovery of incommensurable ratios produced something of a crisis; a possibly apocryphal story states that a disciple named Hippasus (c. 500 BC) was set adrift at sea as punishment for its revelation.

By 340 BC, however, the Greeks were happy to state that incommensurable lengths exist.

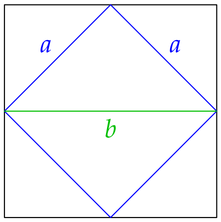
Theorem (Aristotle). *If the diagonal and side of a square are commensurable, then odd numbers equal even numbers.*

Inferred proof. In Socrates' doubled-square, suppose that side : diagonal $= a : b$; these are integers!

Assume at least one of a or b is odd, else the common sub-length may be doubled. The larger square is twice the smaller, whence the square numbers have ratio

$$b^2 : a^2 = 2 : 1$$

It follows that b^2 is even and thus divisible by 4. But then a^2 is also even, whence *both* a, b are even. Whichever of a, b was odd is also even: contradiction!



■

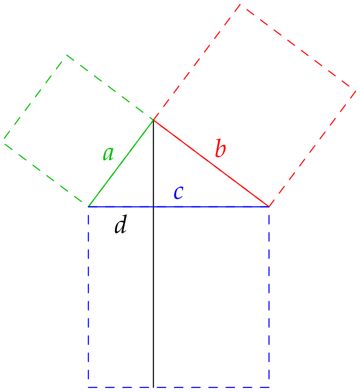
Note the similarity of this argument to the modern proof of the irrationality of $\sqrt{2}$.

While there is no evidence that the Pythagoreans ever provided a correct proof of their famous Theorem, one argument possibly attributable to them relied on the flawed notion of commensurability.

'Proof' of the Pythagorean Theorem. Label the right-triangle a, b, c and drop the altitude to the hypotenuse c as shown. Let d be the part of the hypotenuse under a. Similar triangles tell us that

$$a : d = c : a \implies a^2 : ad = cd : ad \implies a^2 = cd$$

Thus the square on a has the same area as the rectangle below the d-side of the hypotenuse. Repeat the calculation on the other side to obtain $b^2 = c(c - d)$, and sum to complete the proof. $\blacksquare$



Since the only *numbers* were integers, a, b, c, d denote *integer multiples* of an assumed common sub-length. This restriction destroys the generality of the argument, since most triangles cannot be so labelled.
Book I of the *Elements* is plainly organized so as to provide a rigorous repair of the above argument which did not depend on the flawed notion of commensurability. With our modern understanding of real numbers and continuity, there is nothing wrong with the above.

Theaetetus of Athens (417–369 BC) Theaetetus is likely the source for much of the most difficult parts (Books VII & X) of the *Elements*. His essential definition of (in)commensurability comes from applying what is now known as the Euclidean algorithm to line segments.

Definition. Let $\underline{a} > \underline{b}$ be lengths/segments.[11] Repeatedly apply the division algorithm:

$$\begin{array}{lll} \underline{a} = q_1\underline{b} + \underline{r}_1 & \underline{r}_1 < \underline{b} & (\exists q_1 \in \mathbb{N}_0 \text{ and a length } \underline{r}_1 < \underline{b}) \\ \underline{b} = q_2\underline{r}_1 + \underline{r}_2 & \underline{r}_2 < \underline{r}_1 & \\ \underline{r}_1 = q_3\underline{r}_2 + \underline{r}_3 & \underline{r}_3 < \underline{r}_2, \quad \text{etc.} & \end{array}$$

We say that $\underline{a}$ and $\underline{b}$ are *commensurable* if the algorithm terminates: some remainder $\underline{r}_n$ divides exactly into $\underline{r}_{n-1}$. Otherwise $\underline{a}$ and $\underline{b}$ are *incommensurable.*
Ratios are *equal* $\underline{a} : \underline{b} = \underline{c} : \underline{d}$ precisely when the sequences of quotients in the algorithm are equal.

If $\underline{a}$ and $\underline{b}$ are commensurable, then $\underline{r}_n$ is their *greatest common sub-length*. If we write $\underline{a} = a\underline{r}_n$ and $\underline{b} = b\underline{r}_n$ for some integers a, b and rewrite the algorithm in the modern fashion, the result is the standard Euclidean algorithm computation of $\gcd(a, b) = 1$.

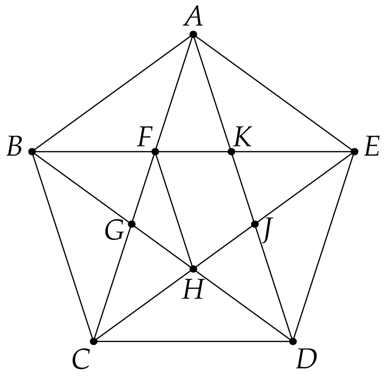
Example 1 $37 : 13 = 148 : 52$ since we obtain the same sequence of quotients $(2, 1, 5, 2)$:

$$\begin{array}{ll} \underline{37} = 2 \cdot \underline{13} + \underline{11} & \underline{148} = 2 \cdot \underline{52} + \underline{44} \\ \underline{13} = 1 \cdot \underline{11} + \underline{2} & \underline{52} = 1 \cdot \underline{44} + \underline{8} \\ \underline{11} = 5 \cdot \underline{2} + \underline{1} & \underline{44} = 5 \cdot \underline{8} + \underline{4} \\ \underline{2} = 2 \cdot \underline{1} & \underline{8} = 2 \cdot \underline{4} \end{array}$$

---

[11] Thus '$\underline{a}$ is longer than $\underline{b}$' is a statement about lengths, not numbers. *Only* the quotients q_k need be integers.

Example 2 We sketch a proof that the side $\overline{AB}$ and diagonal $\overline{AC}$ of a regular pentagon $ABCDE$ are incommensurable.



1. Prove that $\triangle BAG$ is isosceles (just count angles!).
2. Take $a = |AC|$ and $b = |AB| = |AG|$. The first line of the algorithm reads

   $$|AC| = |AG| + |GC| = |AB| + |GC|$$

   so we write $\underline{a} = q_1\underline{b} + \underline{r}_1$ where $r_1 = |GC|$ and $q_1 = 1$.
3. Since $|GC| = |AF|$, the second line of the algorithm reads

   $$|AG| = |AF| + |FG| = |GC| + |FG|$$

   so that we again have a quotient of $q_2 = 1$.
4. Appealing to congruent isosceles triangles $\triangle DCG \cong \triangle EHF$ we see that $|GC| = |FH|$ is the diagonal of the interior regular pentagon. The third line of the algorithm is therefore the same as the first: we are back to considering the ratio of the diagonal to the side of a regular pentagon. The algorithm therefore continues forever with all quotients being 1.

Example 3 In modern language, the diagonal to the side of a square is the incommensurable ratio $\sqrt{2} : 1$. We apply Theaetetus' algorithm:

$$\underline{\sqrt{2}} = 1 \cdot \underline{1} + (\underline{\sqrt{2} - 1})$$
$$\underline{1} = 2 \cdot (\underline{\sqrt{2} - 1}) + (\underline{3 - 2\sqrt{2}})$$

Observe that $3 - 2\sqrt{2} = (\sqrt{2} - 1)^2$, whence the second line reads $1 = 2x + x^2$. The following lines in the algorithm may therefore be obtained by repeatedly multiplying by x:

$$x = 2x^2 + x^3, \quad x^2 = 2x^3 + x^4, \quad \text{etc.},$$

resulting in a never-ending sequence[12] of quotients: $1, 2, 2, 2, 2, 2, \ldots$

Eudoxus of Knidos (c. 390–337 BC) Eudoxus was arguably the most prolific pre-Euclidean mathematician. Apart from attending and perhaps teaching at Plato's academy, he is famous for explaining how to calculate with ratios of lengths (segments). For example:

Definition. $A : B > C : D$ if there exist positive integers m, n such that $mA > nB$ and $mC \leq nD$.

At first glance it appears as if Eudoxus is telling us how to compare *rational* numbers; if A, B, C, D are integers, we see that

$$\frac{A}{B} > \frac{n}{m} \geq \frac{C}{D}$$

which is trivially satisfied by taking $m = D$ and $n = C$. To Eudoxus however, A, B, C, D could also be interpreted as *segments*. Building on the work of Theaetetus, his mathematics told the Greeks how to approximate incommensurable ratios with rational ratios.

---

[12] If you're interested in number theory, investigate the relationship of Theaetetus' algorithm to continued fractions. . .

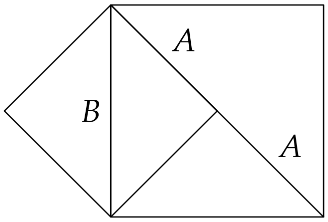
Examples 1. To see that $13 : 3 > 17 : 4$, simply choose $m = 4$ and $n = 17$ to obtain

$$4 \cdot 13 = 52 > 51 = 3 \cdot 17$$

2. We show that the side : diagonal of a square is greater than 1 : 2; equivalently, the diagonal is less than twice the side.
   In the picture, side : diagonal $= A : B$. Choosing $m = D = 2$ and $n = C = 1$, we see that

   $$mA = 2A > B = nB \qquad \text{(diag > side of large square)}$$

   In modern language, this is merely $\frac{1}{\sqrt{2}} > \frac{1}{2}$.



Zeno of Elea (c. 450 BC) Zeno's arguments have provided philosophical fodder for millennia. Several illustrate the essential difficulties of the infinite and the infinitesimal that lie at the heart of the controversy around the development of calculus. Here are two of the most famous:

Achilles and Tortoise Achilles chases a Tortoise. After time t_0, Achilles reaches the Tortoise's starting position, but the Tortoise has moved on. After another time t_1, Achilles reaches the Tortoise's second position; again the Tortoise has moved. In this manner Achilles spends $t_0 + t_1 + t_2 + \cdots$ in the chase. Zeno's paradoxical conclusion is that Achilles never catches the Tortoise.

This paradox may be resolved (see Exercise 8), at least assuming both Achilles and the Tortoise travel at constant speeds: even though it be split into infinitely many subintervals of time, the total duration of the chase can expressed as the (finite) value of a convergent *infinite series.*

Arrow paradox An arrow is shot from a bow. At any given instant the arrow doesn't move. If time is made up of instants, then the arrow never moves.

This time Zeno debates the the idea that a finite time period can be considered as a sum of infinitesimal instants. The same difficulty is central to *integration.*

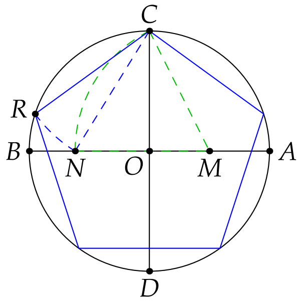
Constructions and Geometry By the mid 5th century BC, Greek mathematicians were solving geometric problems using *ruler-and-compass* (peg-and-cord) constructions. This approach could have come to Greece from India, or might have arisen organically. Constructions were based on three rules, which became the first three postulates (axioms) of Book I of Euclid's *Elements.*

1. Two points may be joined by a straight line segment.
2. Any segment may be extended indefinitely.
3. Given a center and radius, one may draw a circle.

Theorems were often stated as *problems*: e.g., *to bisect a given angle.* A proof provided first a *construction,* then an argument justifying that the construction really had solved the problem.

By the time of Euclid, the Greeks knew how to construct an equilateral triangle, a square and a regular pentagon in a given circle.

Example We construct a regular pentagon in a given circle.[13]



1. Draw perpendicular diameters $\overline{AB}$ and $\overline{CD}$ and bisect $\overline{OA}$ at M.
2. Draw an arc centered at M with radius $|CM|$. Let N be the intersection of this arc with $\overline{OB}$.
3. Draw an arc centered at C with radius $|CN|$. Let R be the intersection of this arc with the original circle.
4. Move $|CR|$ around the circle to create a regular pentagon.

A purely geometric proof of the validity of this construction is too difficult for us, but you can easily check it using a calculator, Pythagoras' and the cosine rule: if the circle has radius 2, then

$$|CR|^2 = |CN|^2 = |ON|^2 + |OC|^2 = (\sqrt{5}-1)^2 + 2^2 = 10 - 2\sqrt{5} = 2^2 + 2^2 - 2 \cdot 2 \cdot 2 \cos 72^\circ$$

Construction problems have motivated mathematicians ever since. In 1796, Gauss (then 19) constructed a regular 17-gon. A classification of constructable regular polygons took until 1837.

Theorem. *A regular n-gon is constructable if and only if $n = 2^k F_1 \cdots F_r$ where $F_1, \ldots, F_r$ are distinct primes of the form $2^{(2^n)} + 1$.*

After the 17-gon, the next prime-sided constructable n-gon has $257 = 2^{2^3} + 1$ sides!

By 400 BC, the Greeks were referencing the second and third *impossible constructions of antiquity*:

1. Trisecting a general angle.
2. Doubling (the volume of) a given cube.
3. Squaring a circle (construct a square with the same area as a given circle).[14]

It wasn't until the advent of field theory in the 1800s that such were proved to be impossible using ruler-and-compass constructions.

Summary Several of the mathematical techniques in this section are difficult, and the results technical. It isn't important to become proficient with all of these ideas! Instead play with them to help develop an appreciation of two overarching points:

1. Even before Euclid, the focus of Greek mathematics was more abstract and less practical than other ancient cultures (Egyptians, Babylonians, Chinese, etc.), largely due to the influence of wider Greek philosophy and religion. The modern liberal arts ideal of learning for its own sake—to celebrate the beauty of knowledge and to expand the mind—is, to a large extent, a Greek inheritance.

2. The ancient Greeks pondered fundamental mathematical questions and concepts—*number* versus *length*, continuity, irrationality, infinitesimals, constructions—ideas that have stimulated mathematical research ever since. These particular issues would not rigorously be resolved until the 1800s when luminaries such as Gauss, Cauchy and Riemann developed modern analysis and algebra.

---

[13] Theorem IV. 11 of the *Elements* presents a less practical construction. Ours follows from Theorem XIII. 10: if a regular pentagon, hexagon and decagon are inscribed in a circle, then their sides form a right-triangle.

[14] "You can't square that circle" is now a metaphor for something that can't be done.

Exercises 3.2. 1. Construct five Pythagorean triples using the formula $\left(n, \frac{n^2-1}{2}, \frac{n^2+1}{2}\right)$ where n is odd. Construct five more using the formula $\left(m, \left(\frac{m}{2}\right)^2 - 1, \left(\frac{m}{2}\right)^2 + 1\right)$ where m is even.

2. Suppose $2^n - 1 = p$ is prime (its only positive divisors are itself and 1). List the positive divisors of $2^{n-1}(2^n - 1)$ and hence prove Theorem IX. 36.

3. Draw a *picture* with dots to show that eight times any triangular number plus 1 makes a square, and that any odd square diminished by 1 becomes eight times a triangular number. That is:
   (a) $8 \cdot \frac{1}{2}m(m+1) + 1$ is a perfect square.
   (b) If n is odd, then $n^2 - 1 = 8 \cdot \frac{1}{2}m(m+1)$ for some m.

4. Find a construction (using the *ruler-and-compass* constructions) to bisect a given angle, and show that it is correct.

5. Sketch a construction inscribing a regular hexagon in a circle.

   (*Assume you can construct an equilateral triangle on a given segment—Thm I. 1 of Euclid, pg. 27*)

6. (A line-doubling paradox) One line has twice the length of another and so has *more* points. However, there is a bijective correspondence between the points on these lines; the two lines therefore have the *same number* of points.

   Explain the second observation. How can you resolve the paradox?

7. The *cycle of fifths* is a musical concept stating that twelve perfect fifths equals seven octaves (pg. 18). State this claim *numerically*, and show that it is a contradiction.

   (*Hint: two strings are seven octaves apart if their lengths are in the ratio* $2^7 : 1$)

8. We use modern language to resolve Zeno's paradox of Achilles and the Tortoise. Suppose Achilles travels at speed v_A, the tortoise at speed $v_B < v_A$, and that the tortoise starts a distance d ahead of Achilles.
   (a) Prove that $t_n = \frac{d}{v_A}\left(\frac{v_T}{v_A}\right)^n$ for each positive integer n.
   (b) Compute $\sum_{n=0}^{\infty} t_n$ using the geometric series formula from calculus.
   (c) Verify the time-value computed in (b) as would a modern Physicist; by considering the motion of Achilles relative to the tortoise.

9. Use Theaetetus' definition of equal ratios to prove that $46 : 6 = 23 : 3$.

10. (Hard) A line of length 1 is divided at x so that $\frac{1}{x} = \frac{x}{1-x}$. Prove that 1 and x are incommensurable. Indeed, show that $1 : x$ is *the same* as diagonal : side of a regular pentagon.

    (*Hint: the first line of the algorithm is* $1 = x + x^2 \ldots$)

11. (Hard) Let $a > b$ and c be *positive lengths.* Use Eudoxus' definition to *prove* that $c : b > c : a$.

    (*Hint: let* n *be the smallest integer such that* $n(a-b) \geq c$*; its existence is the "archimedean property"*)

3.3 Euclid and the *Elements*

Euclid worked in the Library of Alexandria, named for the Greek general Alexander the Great who conquered Egypt in 323 BC. The Library was constructed around 320 BC as a means of organizing the knowledge of the world and for the demonstration of Greek power. Although it was seriously damaged on several occasions, the Library remained a center of scholarship until around AD 500. Below is a map of the city around AD 400: note the size and centrality of the Library.

It is hard to argue against Euclid's *Elements* (c. 300 BC) as the most influential mathematics text ever produced. Likely a compilation of earlier mathematical work rather than a pure original, it was edited and added to over the centuries, eclipsing and subsuming other works. Particular import were the edits of Theon of Alexandria (c. AD 400) and his daughter Hypatia, both prolific scholars in their own right. Due to edits such as these, the precise contents of the original are unknown.

Extant fragments date to around AD 100. The earliest (almost) complete copy is from the the 9th century; written in Greek and held at the Vatican, it is missing some of the edits of Theon & Hypatia, thus demonstrating that multiple versions were in circulation.

Earliest Fragment c. AD 100

Full copy 9th C

Until the mid 20th century, some version of the *Elements* would have been used as a high-school textbook in most western and middle-eastern countries. Many editions and variations have been produced, four of which are shown below:

EVCLIDIS
ELEMENTORVM
LIBER TERTIVS
CVM SCHOLIIS ANTIQVIS,
ET COMMENTARIIS
Federici Commandini Vrbinatis.

SCHOLIVM.

Propoſitum Euclidi eſt hoc loco tractare de ijs, quæ circulis accidunt cum ad rectas lineas, & ad angulos comparantur.

DIFFINITIONES.

I.

AEQVALES circuli ſunt, quorum diametri ſunt æquales, vel quorum quæ ex centris ſunt æquales.

II.

Recta linea circulum contingere dicitur, quæ contingens circulum, et producta ipſum non ſecat.

III.

Circuli contingere ſe ſe dicũtur, qui contingentes ſe ipſos non ſecant.

Incirculo

LIBER III. 37

IIII.

In circulo æqualiter diſtare à centro rectæ lineæ dicuntur, quando à centro ad ipſas perpẽdiculares ductę sũt ęquales.

V.

Magis autem diſtare à centro dicitur ea, ad quam maior perpendicularis cadit.

VI.

Portio circuli eſt figura, quę recta linea, et circuli circumferẽtia continetur.

VII.

Portionis autẽ angulus eſt, qui recta linea, et circuli circũferentia cõprehẽditur.

VIII.

In portione angulus eſt, quando in circũferentia portionis ſumatur aliquod punctum, atque ab ipſo ad terminos lineæ eius, quæ baſis eſt portionis, rectę lineæ ducantur, angulus vero ductis lineis ſit cõtentus.

VIIII.

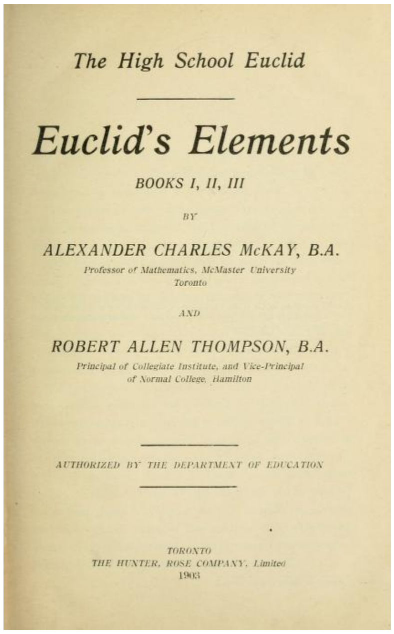
Quando autem continentes angulum rectæ lineæ aſſumunt circumferentiam, in illa conſiſtere angulus dicitur.

X.

Sector circuli eſt, quando angulus ad centrum cõſtiterit, figura contenta rectis lineis angulum comprehendentibus, et circumferentia ab ipſis aſſumpta.

K Similes

Latin translation, 1572



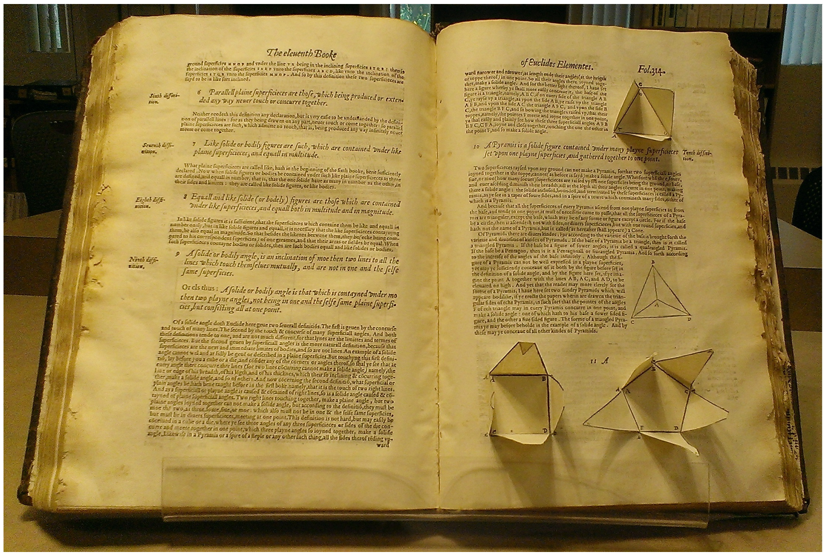
The High School Euclid

Euclid's Elements
BOOKS I, II, III
BY
ALEXANDER CHARLES McKAY, B.A.
Professor of Mathematics, McMaster University
Toronto
AND
ROBERT ALLEN THOMPSON, B.A.
Principal of Collegiate Institute, and Vice-Principal
of Normal College, Hamilton

AUTHORIZED BY THE DEPARTMENT OF EDUCATION

TORONTO
THE HUNTER, ROSE COMPANY, Limited
1903

High School textbook, 1903

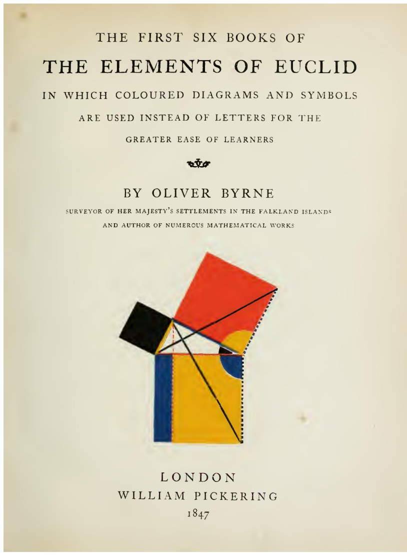


The eleuenth Booke

of Euclides Elementes.

Fol.314.

Pop-up edition, 1500s



THE FIRST SIX BOOKS OF
THE ELEMENTS OF EUCLID
IN WHICH COLOURED DIAGRAMS AND SYMBOLS
ARE USED INSTEAD OF LETTERS FOR THE
GREATER EASE OF LEARNERS

BY OLIVER BYRNE
SURVEYOR OF HER MAJESTY'S SETTLEMENTS IN THE FALKLAND ISLANDS
AND AUTHOR OF NUMEROUS MATHEMATICAL WORKS

LONDON
WILLIAM PICKERING
1847

Color edition, 1847

You can download Byrne's color edition here (very large file!). It is notably different from many earlier editions: it contains a much longer list of definitions, inserts many more axioms, and relabels propositions 4 and 5 as axioms (pages xviii–xxiii). The picture on the cover page is Euclid's proof of Pythagoras' Theorem (Book I, Thm. 47).

A Brief Overview of the *Elements*

The *Elements* consists of thirteen books covering two- and three-dimensional geometry, computations and number theory. Whether some version of every book was written by Euclid himself is unknown.[15] The key feature of the *Elements* is its *axiomatic* presentation. Each book begins with a list of axioms/postulates and definitions and proceeds to prove theorems deduced from these. This *axiomatic method* is essentially universal in modern mathematics, and its advent is fundamentally what sets Greek mathematics apart from everything that came before.

We briefly discuss Book I, then give some flavor of the remainder of the text with a few example results. Several examples of material from later books were mentioned in the previous section.

Book I Consists of 48 theorems, culminating with Pythagoras' and its converse. It seems likely that Euclid organized Book I with the goal of proving this important result in a rigorous manner: recall (pg. 20) how the Pythagorean 'proof' relied on the erroneous notion of commensurability. Here are the postulates from Book I: the first three are what define ruler-and-compass constructions (pg. 22).

P1 Given any two points, a straight line can be drawn between them

P2 Any line may be indefinitely extended

P3 Given a center and a radius, a circle may be drawn

P4 All right angles are equal to each other

P5 If a straight line crosses two others so that the angles on the same side make less than two right angles, then the two lines meet on that side of the original.

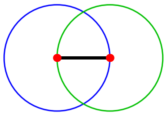
The fifth postulate is awkwardly phrased. An equivalent modern statement is *Playfair's axiom*:

> There is at most one parallel through a given point not on line.

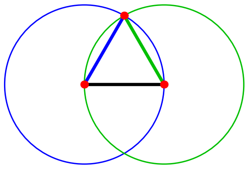
For centuries, mathematicians tried to prove that this postulate was a theorem of the others until; it was eventually shown to be necessary with the advent of hyperbolic geometry in the 1800s. Euclid's refusal to use the parallel postulate until Theorem 29 suggests he understood this awkwardness.

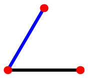
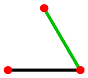
Theorems are generally presented as *problems*: a pictorial construction is provided, then Euclid proves that the construction really solves the problem. Here is Euclid's first theorem.

Theorem (I. 1). *Problem: To construct an equilateral triangle on a given segment.*

Proof. Given  construct two circles (P3) 

Join one of the circle intersections to the endpoints of the original segment (P1)



The result is an equilateral triangle; indeed the three sides are congruent, for  are radii of a common circle, as are  ■

---

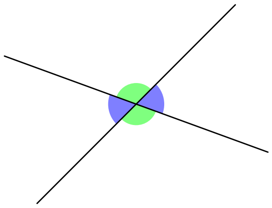
[15] It is not even certain that Euclid was a single person as opposed to a figurehead or a name for a collective. It is hardly surprising that we know so little about someone who lived 2300 years ago. The same questions are sometimes raised about William Shakespeare who lived only 400 years ago!

After this Euclid proceeds to establish several well-known results. Since this isn't a geometry class, we'll omit most of the details. You can find more of these here, in Byrne, or elsewhere.

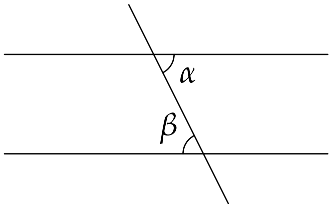
Thm I. 4 Side angle side: if triangles have two pairs of congruent sides and the angles between them are also congruent, then the remaining sides and angles are congruent.

Thm I. 15 Vertical angles: if two lines meet, then the opposite angles made are congruent.

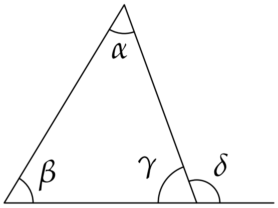
Thms I. 27 & 29 Angles and parallels: if a line falls on two other lines, then the two lines are parallel if and only if the alternate angles are congruent ($\alpha \cong \beta$ in the picture).



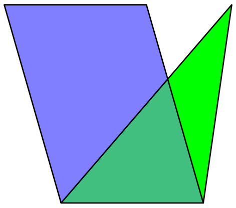
Thm I. 15



Thms I. 27/29



Thm I. 32



Thm I. 41

Thm I. 32 Angle sums in a triangle: if one side of a triangle is protruded, the exterior angle equals the sum of the opposite interior angles.

In the picture $\alpha + \beta = \delta$; in modern language $\alpha + \beta + \gamma = 180^\circ$.

Thm I. 41 A parallelogram and triangle on the same base and with the same height have area in the ratio 2:1.

The last two results of Book I are Pythagoras and its converse.

Theorem (I. 47). *The square on the hypotenuse of a right-triangle has area equal to the sum of the areas of the squares on the remaining sides.*

Proof. Given a right-angle at A, drop the perpendicular from A across $|BC|$ to L.

$\triangle FBC$ and $\square ABFG$ share the same base $\overline{BF}$ and height $\overline{AB}$. By Thm I. 41,

$$\text{area}(\square ABFG) = 2\,\text{area}(\triangle BCF)$$

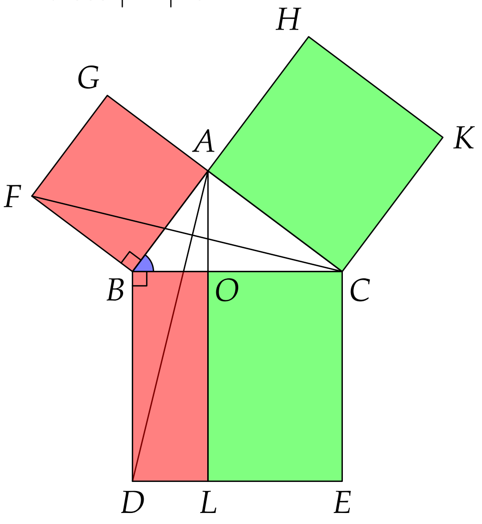
Similarly (base $\overline{BD}$, height $\overline{BO}$)

$$\text{area}(\square BOLD) = 2\,\text{area}(\triangle ABD)$$

Side-Angle-Side (Thm I. 4) $\implies$ $\triangle ABD \cong \triangle FBC$; the triangles have the same area, and so

$$\text{area}(\square ABFG) = \text{area}(\square BOLD)$$

Similarly $\text{area}(\square ACKH) = \text{area}(\square OCEL)$.



■

The converse (Thm I. 48) is Exercise 3.

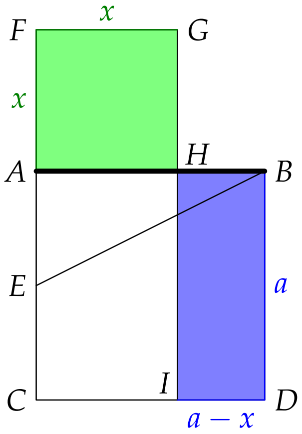
Book II Geometric solutions to problems that would now be treated using algebra. Much of this is attributable to the Pythagoreans.

(Thm II. 11) A segment may be divided so that the rectangle contained by the whole and one of the sub-segments is equal to the square on the remaining sub-segment.

We rephrase this in modern language. Suppose the given segment $\overline{AB}$ has length a, our goal is to find H on $\overline{AB}$ such that $|AH| = x$ and $x^2 = a(a-x)$. Euclid is providing a *geometric solution* to a quadratic equation!

1. Construct $\square ABDC$ on $\overline{AB}$
2. Let E be the midpoint of $\overline{AC}$ and connect $\overline{EB}$
3. Extend $\overline{AC}$ and lay off $\overline{EF} \cong \overline{EB}$ on $\overline{AC}$ extended
4. Construct $\square AFGH$ and drop the perpendicular to I
5. To finish: prove that area$(\square AFGH)$ = area$(\square BDIH)$

The solution $x = |AH| = \frac{\sqrt{5}-1}{2}a$ means that $\overline{AH} : \overline{HB}$ is the *golden ratio.*



Book III Theorems regarding circles and tangency. Most of this material likely came from Hippocrates.
(Thm III. 31) Thales' Thm: a triangle in a semi-circle is right-angled.

Book IV Construction of regular 3, 4, 5, 6 and 15-sided polygons inscribing and exscribing a circle. Also likely the work of Hippocrates.

Book V Ratios and magnitudes à la Eudoxus.

(Thm V. 11) If $a : b = c : d$ and $c : d = e : f$ then $a : b = e : f$

Book VI Ratios of magnitudes applied to geometry (similarity results).

(Thm VI. 4) Triangles with equal angles have corresponding sides proportional.

(Thm VI. 8) The altitude from the right angle of a right triangle divides it into two triangles similar to each other and the the original.

(Thm VI. 31) Corrected Pythagorean proof (pg. 20) of Pythagoras' using the proportions of Eudoxus and Thm VI. 8.

Book VII Divisibility and the Euclidean algorithm. Probably due to the Pythagoreans and Theaetetus.

Book VIII Number progressions, geometric sequences. Possibly due to studies in music by Archytas (a Pythagorean who taught Plato mathematics).

Book IX Number Theory: even/odd + perfect numbers.

(Thm IX. 20) There are infinitely many primes.

Book X Discussion of commensurable and incommensurable ratios. Long and difficult, possibly derived from Theaetetus.

Book XI Solid geometry (lines/planes in 3D).

(Thm XI. 28) A parallelepiped is bisected by its diagonal plane.

Book XII Ratios of areas and volumes (Eudoxus).

(Thm XII. 2) The areas of circles are in the same ratio as the squares on their diameters.

Book XIII Construction of regular polyhedra inside a sphere and their classification.

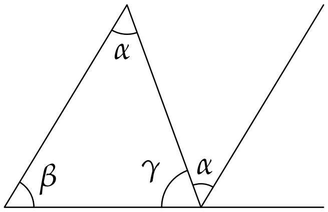
(Thm XIII. 10) If a regular pentagon, hexagon and decagon are inscribed in the same circle, then their sides form a right-triangle.

One could study the *Elements* and its influence for a lifetime and not be done! Hopefully this very brief overview convinces you why the book had such a profound impact on mathematics.

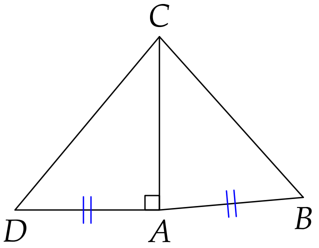
Exercises 3.3. 1. Prove Thales' Theorem (III. 31) (pg. 29).

(*Hint: start by joining the center of the circle to the apex of the triangle…*)

2. Use the picture to provide a proof of Thm I. 32: the sum of the three interior angles of a triangle is equal to two right angles.
   Show that the proof depends on Thm I. 29, and therefore on the parallel postulate.
   (*This isn't quite the same as Euclid's argument*)



3. Suppose that the square on side $\overline{BC}$ of $\triangle ABC$ has the same area as the sum of the squares on the other sides $\overline{AB}$, $\overline{AC}$. As in the picture, draw a perpendicular $\overline{AD} \cong \overline{AB}$.
   (a) Explain why $\overline{DC} \cong \overline{BC}$.
   (b) Hence conclude that $\triangle ABC$ is right-angled at A.



4. Prove Thm III. 3: A diameter of a circle bisects a chord if and only if it is perpendicular to the chord.

5. Verify that Euclid's construction for Thm II. 11 really does solve the given problem.
   (*You can use modern algebra!*)

6. Draw a semi-circle with diameter $9 + 5 = 14$. Solve the equation $\frac{9}{x} = \frac{x}{5}$ *geometrically*, by constructing a vertical line whose length is x.

7. Show that areas of similar segments of circles are proportional to the squares of the length of their chords.
   (*You may assume that areas of circles are proportional to the squares on their diameters and can use modern algebra/trigonometry if you wish*)

3.4 Archimedes of Syracuse

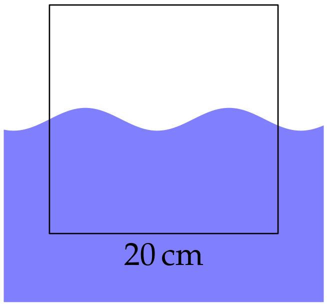
Archimedes (287–212 BC) is arguably the greatest ancient mathematician. Syracuse is on the island of Sicily at the foot of the Italian peninsula; at the time of Archimedes' birth this was a Greek city-state, though under threat from the expanding Roman Empire. Archimedes famously helped defend Syracuse against the Romans using catapults, though he ultimately died at their hands after the city fell. He is believed to have travelled to Alexandria in his youth and perhaps studied with scholars at the library, including Eratosthenes (pg. 34).

Archimedes' genius was practical not just mathematical. Beyond his anti-Roman catapults, he is credited with a large number of inventions and technical innovations, including *Archimedes' screw,* still used in modern irrigation systems to elevate water. He is acknowledged as the founder of *hydrostatics,* where *Archimedes' principle* states that an object immersed in water loses weight equal to that of the displaced water. A famous story recounts Archimedes using this to detect whether a smith had used all the gold he had been given in the manufacture of a crown.

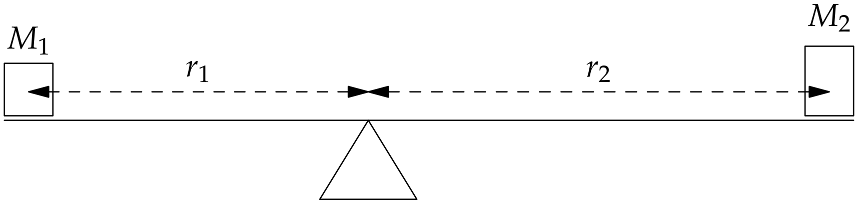
Example A cube with side length 20 cm floats such that the waterline is halfway up the cube. By Archimedes' principle, the weight of the cube is the same as that of the volume of displaced water:

$$20 \times 20 \times 10 = 4000\,\text{cm}^3$$

which has a weight (mass) of roughly 4 kg.



Levers Archimedes made great study of levers, both for practical purposes and as a method of calculation.



Given masses M_1, M_2 located distances r_1, r_2 from a pivot, Archimedes states:

- The lever balances $\iff M_1 : M_2 = r_2 : r_1$
- The lever rotates clockwise $\iff M_1 : M_2 < r_2 : r_1$
- The lever rotates counter-clockwise $\iff M_1 : M_2 > r_2 : r_1$

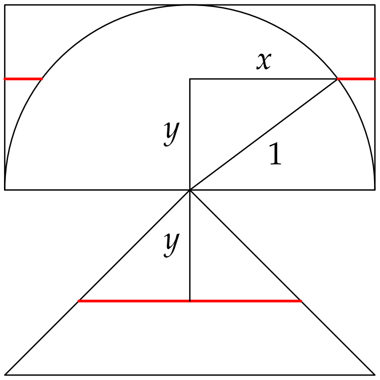
In modern terms we'd compare the *torques* $\tau_1 = M_1 r_1$ and $\tau_2 = M_2 r_2$. Since torque requires the multiplication of non-numerical quantities, Archimedes would instead have considered this using Eudoxus' theory of proportions.

For example, to find the mass M_2 required to balance a lever given $M_1 = 12\,\text{lb}$, $r_1 = 4\,\text{ft}$ and $r_2 = 3\,\text{ft}$, Archimedes would have observed that

$$M_2 : M_1 = 4 : 3 \implies M_2 = 16\,\text{lb}$$

***The Method*: is Archimedes the founder of calculus?** A previously unknown work of Archimedes was discovered in 1899. As an amazing application of the lever principle, Archimedes makes an argument that looks remarkably like modern calculus; he could be claimed to be its earliest practitioner by 1800 years! The method was outlined in a letter to Eratosthenes and includes part of an argument for proving Archimedes' favorite theorem, a picture of this result was engraved on his tomb.

Theorem. *A cone, hemisphere and cylinder with the same base and height have volumes in the ratio* 1 : 2 : 3*. Using modern formulæ, if the height is r, then the volumes are* $\frac{1}{3}\pi r^3 : \frac{2}{3}\pi r^3 : \pi r^3$.



Here is a modernized version illustrating Archimedes' approach. Suppose the 'base' is a disk with radius 1, remove the hemisphere from the cylinder and place the cone beneath. Compare the cross-sections the same distance y from the apex of the cone.

- The circular cross-section of the cone has radius y whence its area is proportional to the square on the radius: πy^2.
- The upper annular cross-section has area proportional to the difference of the squares on the radius of the cylinder and on the distance x. By Pythagoras' the cross-sectional area is

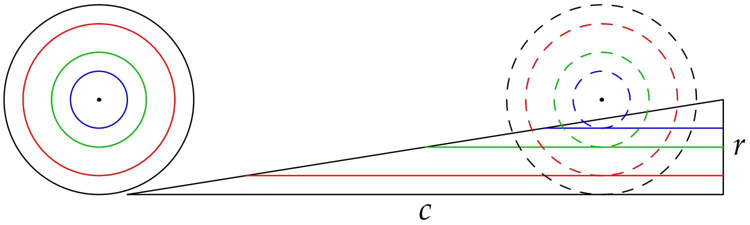
$$\pi(1^2 - x^2) = \pi(1 - (1 - y^2)) = \pi y^2$$

The cross-sections are therefore in balance with respect to a vertical lever whose pivot lies at the center of the picture. Archimedes concludes that the cone and the upper-figure are in balance: that is

$$V_{\text{cone}} = V_{\text{cylinder}} - V_{\text{hemisphere}}$$

Combining this with the fact that $V_{\text{cylinder}} = 3V_{\text{cone}}$ (e.g., Exercise 7d) gives the desired result.

Here is another argument of Archimedes' with a suggestion of calculus. A disk comprises infinitely many concentric circles, the circumference of each being proportional to its radius. 'Unwind' these circles to obtain a triangle; one side is the radius of the disk, the other its circumference. The area of a circle is therefore that of a triangle with sides the radius and circumference of the circle: $A = \frac{1}{2}rc$.



The Method includes several of these calculus-like discussions. While efficient, Archimedes felt that his approach didn't constitute a proof and provided alternative arguments elsewhere in his writings. The essential problem is this;

> Can we really say that an area *equals* its cross-sectional lines? Or that a volume *equals* its cross-sectional areas? Lines have no width so if we add them up we have no area. If they have width, then infinitely many of them have infinite area.

These are really variations of Zeno's paradoxes (pg. 22) regarding infinitesimals and indivisibles!

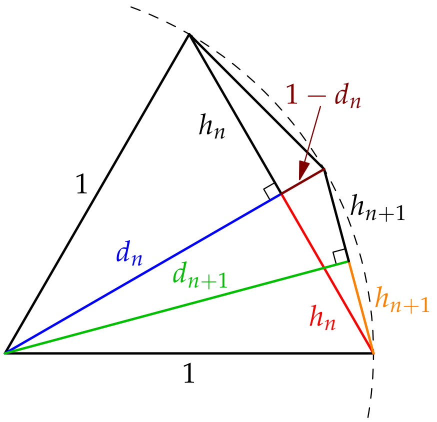
Archimedes' arguments would be resurrected in the early 1600s by Cavalieri and Galileo as the development of calculus gathered pace. The same duality of presentation characterised this later development: Newton and others found the infinitesimal approach efficient, but felt the need to present *geometric* proofs to convince readers that their results weren't mere trickery.

It is tempting to imagine what might have happened if Archimedes' *method* had been accepted and preserved as part of the Greek canon; if calculus had been developed 1800 years earlier, how might this have affected technological development? Would the space-race have happened in AD 500?!

Quadratures Archimedes also approximated areas and arc-lengths of various figures using limit-like argumentation. Here is how he approached the area/circumference of a circle.

1. Inscribe a regular hexagon in a circle (of radius 1 say) and compute its perimeter (6).
2. Halve each angle to obtain a regular dodecagon: compute its perimeter ($12\sqrt{2-\sqrt{3}}$).
3. Repeat the angle-halving process: Archimedes did this with 24-, 48- and 96-gons to obtain an increasing sequence of perimeters bounded above by the circumference of the circle (2π).
4. Repeat the same calculation with circumscribed polygons to obtain a decreasing sequence of over-estimates.
5. Using 96-sided polygons allowed Archimedes to obtain the estimate $3\frac{10}{71} < \pi < 3\frac{1}{7}$.

Archimedes' halving process relied on an induction step, an approximation of which we mimic here. Suppose we have an isosceles triangle with equal legs 1, altitude d_n, and chord $2h_n$. We halve the angle to find the new altitude d_{n+1} and chord $2h_{n+1}$. Everything follows from three applications of Pythagoras':



$$1 = d_n^2 + h_n^2$$
$$(2h_{n+1})^2 = h_n^2 + (1-d_n)^2$$
$$1 = d_{n+1}^2 + h_{n+1}^2$$

Expanding and cancelling, we obtain

$$d_{n+1}^2 = \frac{1}{2}(1+d_n), \qquad h_{n+1}^2 = 1 - d_{n+1}^2 = \frac{1}{2}(1-d_n)$$

Since $d_0 = \frac{\sqrt{3}}{2}$ and $h_0 = \frac{1}{2}$, we may compute the entirety of both sequences:

$$d_1 = \frac{1}{2}\sqrt{2+\sqrt{3}}, \quad d_2 = \frac{1}{2}\sqrt{2+\sqrt{2+\sqrt{3}}}, \quad \ldots \quad d_n = \frac{1}{2}\sqrt{2+\sqrt{2+\cdots+\sqrt{2+\sqrt{3}}}}$$
$$h_1 = \frac{1}{2}\sqrt{2-\sqrt{3}}, \quad h_2 = \frac{1}{2}\sqrt{2-\sqrt{2+\sqrt{3}}}, \quad \ldots \quad h_n = \frac{1}{2}\sqrt{2-\sqrt{2+\cdots+\sqrt{2+\sqrt{3}}}}$$

where the n^{th} terms have n copies of the digit 2 under the square-root. The circumference and area of the $6\cdot 2^n$-sided polygon inscribed in the circle are therefore

$$C_n = 12\cdot 2^n h_n, \qquad A_n = 6\cdot 2^n d_n h_n = 6\cdot 2^{n-1} h_{n-1} = \frac{1}{2}C_{n-1}$$

These sequences increase to 2π and π respectively. For a 96-sided polygon, Archimedes would have had to approximate

$$C_4 = 12 \cdot 2^4 h_4 = 96\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{3}}}}} > 6.282 \implies \pi > 3.141 > 3\frac{10}{71}$$

Other Highlights of the later Greek Period: 300 BC–AD 500

We'll consider ancient astronomy, including Greek contributions, in the next chapter. Here are a few of the other developments of the late Greek period and some historical context.

- Eratosthenes (276–194 BC) grew up in Cyrene (c. 500 miles west of Alexandria in modern-day Libya) and moved to Alexandria in adulthood to become its librarian. He is credited with a simple algorithm for finding primes: the *Sieve of Eratosthenes.*
  - List the integers $n \geq 2$.
  - Leave 2 and delete all its multiples.
  - Leave 3 and delete its multiples.
  - Repeat ad infinitum: each time one reaches a number, leave it and delete its multiples.
  - The remaining list contains all the primes.
- Apollonius (225 BC) writes an eight-volume book on conic sections building on earlier work of Menaechmusus (350 BC).
- By 146 BC the Greek empire had fallen under Roman rule. Alexandria remained important. Educated Greeks still spoke and wrote in Greek rather than (Roman) Latin. For context, Julius Caesar ruled Rome around this time (died 44 BC).
- Heron (AD 75) proves the formula $\sqrt{s(s-a)(s-b)(s-c)}$ for the area of triangle, where $s = \frac{1}{2}(a+b+c)$ is the semi-perimeter. This was likely known to Archimedes; Heron's work was a compilation of earlier mathematics.
- Around AD 100 the Neopythagorean's worked in Alexandria, studying music, philosophy, and number, with the intent of reviving the teachings of Pythagoras.
- Around AD 400, Theon and Hypatia produce the most widely-read edition of Euclid's *Elements* as well as improving upon several earlier mathematical topics.
- In AD 395 the Roman empire split into eastern and western parts centered on Rome and Byzantium/Constantinople. The western empire rapidly declined under the pressures of corruption and barbarian attacks, collapsing completely by AD 500. Alexandria experienced riots and a bloody power-struggle (Hypatia was murdered by a mob in 415) and the library of Alexandria was severely damaged and possibly destroyed at this time. In 642, Alexandria was captured by the new Islamic caliphate. Much of the material in the library survived by being copied and transported to various places of learning; particularly Constantinople and Baghdad. For the next 600 years, the knowledge of Alexandria was largely a mystery to (western) Europe.

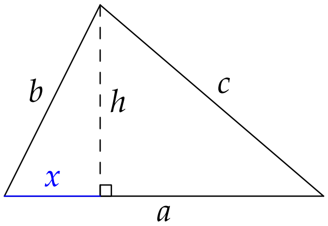
Exercises 3.4. 1. If a weight of 8 kg is placed 10 m from the pivot of a lever and a weight of 12 kg is placed 8 m from the pivot in the opposite direction, toward which weight will the lever incline? Answer using Archimedes' language.

2. Use Eratosthenes' Sieve to find all the primes < 100.

3. (a) Prove Heron's formula as follows.

   i. Let h be the altitude and x the base of the left-hand right-triangle. Apply Pythagoras' to the two right-triangles to show that

$$x = \frac{a^2 + b^2 - c^2}{2a}$$



   ii. Substitute in $h^2 = b^2 - x^2$ to find h in terms of a, b, c and thus deduce Heron's formula.

   (b) Find the area of a triangle with sides 4, 7 and 10.

4. Suppose C_n is the circumference of a $6 \cdot 2^n$-sided inscribed polygon in a unit circle. Show that the circumference of the corresponding *circumscribed* polygon is $\mathrm{C}_n^{\mathrm{ex}} = \frac{1}{d_n} C_n$.

5. Use the modern formula $A = \frac{1}{2} ab \sin C$ to prove that, for any $k \in \mathbb{N}$

$$\frac{1}{2} k \sin \frac{2\pi}{k} < \pi < k \tan \frac{\pi}{k}$$

   Moreover, explain why both sides converge to π.

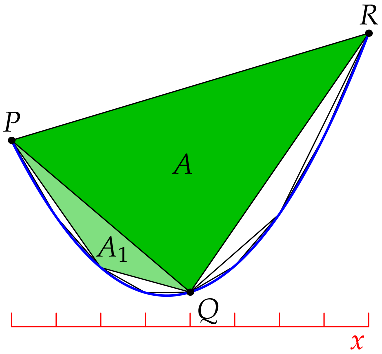
6. Instead of modern algebra, Archimedes used several geometric lemmas to help find the areas of polygons inscribed in and circumscribing circles. Here is one; prove it!

   Let $\overline{OA}$ be the radius of a circle and $\overline{AC}$ be tangent to the circle at A. Let D lie on $\overline{AC}$ such that $\overline{OD}$ bisects $\angle COA$. Then

$$\frac{|DA|}{|OA|} = \frac{|CA|}{|CO| + |OA|} \quad \text{and} \quad |DO|^2 = |OA|^2 + |DA|^2$$

   (*Hint: draw a picture and let T be the intersection of the circle and $\overline{OC}$*)

7. Archimedes used a geometric series approach to evaluate the area inside a parabola. Use modern algebra for this question.



   (a) Suppose $y = a + bx + cx^2$ is the equation of a parabola. If P, Q, R have x co-ordinates in an arithmetic sequence $x - \epsilon, x, x + \epsilon$, show that the area of $\triangle PQR$ is $A = |c|\epsilon^3$; *independent of x!*

   (b) With reference to the picture, explain why the areas of the labelled triangles satisfy $A_1 = \frac{1}{8} A$.

   (c) Use a geometric series to prove that the area inside the parabola bounded by $\overline{PR}$ is $\frac{4}{3} A$

   (d) How might this result have been applied to show that the volume of a cone equals $\frac{1}{3}$ that of a cylinder with the same base and height?

4 Ancient Astronomy & Trigonometry

Throughout history, astronomy has been (and remains) an important driver of mathematical development. In particular, early *trigonometry* (*triangle-measure*) was largely developed to facilitate astronomical computations, for which there are many practical benefits: for instance,

Calendars The phases of the moon (whence *month*), the seasons, and the solar year are paramount. Without an accurate calendar, food production, gathering and hunting are more difficult: *When* will the rains come? *When* should we plant/harvest? *When* will the buffalo return?

Navigation The simplest navigational observation in the northern hemisphere is that the stars appear to orbit *Polaris* (the pole star), thus providing a fixed reference point/direction in the night sky. As humans travelled further, accurate computations became increasingly important.

Religion and Astrology In modern times, we distinguish *astronomy* (the science) from *astrology* (how the heavens influence our lives). However, for most of human history the two were inseparable. In light-polluted modern cities, it is hard to imagine the significance the night sky held for our ancestors, even a couple of centuries ago. Almost all religions imbue the heavens with meaning; understanding and predicting heavenly movements provided a massive historically imperative for mathematical and technological development. Here are just a few examples of the relationship between astronomy, astrology and culture.

- The concept of *heaven* as the domain of the gods, whether explicitly in the sky or simply atop a high mountain (e.g., Olympus in Greek mythology, Moses ascending Mt. Sinai, etc.).
- Many ancient structures were constructed in alignment with heavenly objects:
  - Ancient Egyptians viewed the region around Polaris as their heaven; pyramids included shafts emanating from the burial chamber so that the deceased could 'ascend to the stars.'
  - Several Mayan temples and observatories appear to be oriented to the solstices (page 37). Such alignments are also found elsewhere in the Americas and throughout the world.
  - Venus and Sirius—respectively the brightest planet and star in the night sky—were also important objects of alignment.
- The modern (western) zodiac comes from pre-1000 BCBabylon. A tablet dated to 686 BC describes 60–70 constellations and stars with aspects familiar to modern astrologers, including Taurus, Leo, Scorpio and Capricorn. During the same period Chinese and Indian astronomers developed different systems of constellations.[16]
- Calendars mark religious festivals, practices and even the age of the world.
  - The traditional Hebrew calendar dates the beginning of the world to 3760 BC.
  - The Mayan long count calendar dates the creation of the world to 3114 BC.
  - The modern Gregorian calendar arose to facilitate an accurate determination of Easter.
- The *star in the east* is associated to the birth of Jesus in Christianity.
- Muslims orient themselves towards Mecca when at prayer; we'll see later how this direction (the *qibla*) may be computed, but the required data is astronomical.

---

[16] Chinese astronomy has 28 constellations (or *mansions*). As a point of comparison, Taurus corresponds roughly to the Chinese 'White Tiger of the West' (*Baihu*, and similar terms in various East-Asian languages).

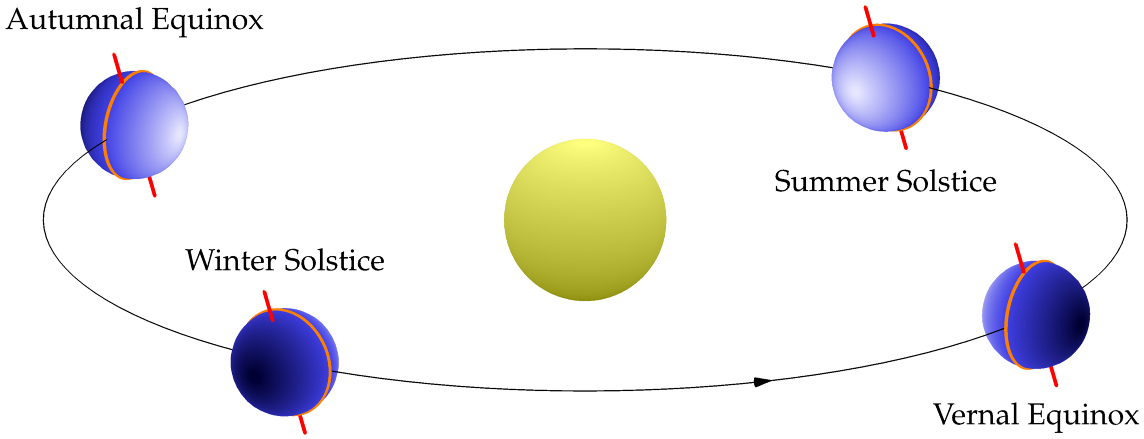
4.1 Astronomical Terminology and Early Measurement

Seasonal variation exists because the earth's axis is tilted approximately 23.5° with respect to the *ecliptic* (sun-earth orbital plane). Summer, in a given hemisphere, is when the earth's axis is tilted towards the sun, resulting in more sunlight and longer days. Astronomically, the seasons are determined by four dates:

Summer/Winter Solstices (≈ 21st June/December) The north pole is maximally tilted towards/away from the sun. *Solstice* comes from Latin meaning 'sun stationary.' The location of the rising/setting sun and its maximal (noon) elevation changes throughout the year, with extremes on the solstices, the summer solstice being when the setting sun is most northerly and (north of the tropics) the noon sun is highest in the sky. Indeed the tropics (of Cancer/Capricorn) are the lines of latitude where the sun is directly overhead at noon on one solstice.

Vernal/Autumnal Equinoxes (≈ 21st March/September) Earth's axis is perpendicular to the Sun-Earth orbital radius. *Equinox* means *equal night*: day and night both last approximately 12 hours everywhere since Earth's axis passes through the day-night boundary.

The picture shows the orientation of the ecliptic, the earth's axis and the day-night boundary.



Astronomical measurements must be be conducted relative to this set-up.

Fixed stars These form the background with respect to which everything else is measured. The *ecliptic* is the sun's apparent path over the year set against the fixed stars. Planets (*wandering stars*) are also seen to move relative to this background.

Celestial longitude Measured from zero to 360° around the ecliptic with 0° at the vernal equinox. One degree corresponds approximately to the sun's apparent daily motion. The ecliptic is divided into twelve equal segments: Aries is 0–30° (March to April 21st); Taurus is 30–60°, etc.

Celestial latitude Measured in degrees north or south of the ecliptic; the sun has latitude zero.

This formulation was largely co-opted by the Greeks from Babylon. The Greeks kept the Babylonian base-60 degrees-minutes-seconds system, which, with minor modifications, persists to this day.[17]

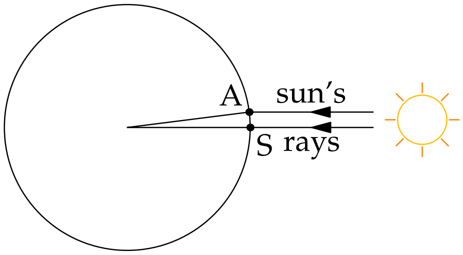
---

[17] Modern astronomers typically measure latitude and longitude (declination/right-ascension) with respect to the earth's equatorial plane rather than the ecliptic. Such *equatorial co-ordinates* are first known to have been introduced by Hipparchus of Nicaea (page 40). Right-ascension is measured in hours-minutes-seconds rather than degrees, where 24 hours = 360°, though modern scientific practice is to use decimals rather than sexagesimal minutes and seconds.

The Circumference of the Earth

One of the earliest problems in astronomy is to find the size of the earth. Eratosthenes of Cyrene (c. 200 BC, page 34) performed one of the first accurate estimations by measuring the sun's rays at noon in two different places.

- Syene (modern-day Aswan, Egypt) is approximately 5,000 *stadia* south of Alexandria.
- When the sun is directly overhead at Syene, it is inclined $7°12' = \frac{1}{50} \cdot 360°$ at Alexandria.
- The circumference of the earth is therefore approximately $50 \cdot 5000 = 250,000$ stadia.



Eratosthenes' original calculation is lost, though it was a little more complicated than the above. From other (shorter) distances, historians have inferred that Eratosthenes' *stadion* was $\approx$ 172 yards, making his approximation for the circumference of the earth $\approx 24,500$ miles, astonishingly accurate in comparison to the modern value of $\approx 25,000$ miles. Later mathematicians provided other estimates based on other locations, but the basic method was the same.

Modelling the Heavens

Early Greek analysis reflects several assumptions.

- Spheres and circles are perfect, matching the 'perfect design' of the universe. The earth is a sphere and the fixed stars (constellations) lie on a larger 'celestial sphere.' Models relied on spheres and circles rotating at constant rates.
- The earth is stationary, and the celestial sphere rotates around it once per day.
- The planets lie on concentric spherical shells centered on the earth.

When such assumptions are tested by observation, two major contradictions appear:

Variable brightness The apparent brightness of heavenly bodies, particularly planets, is non-constant.

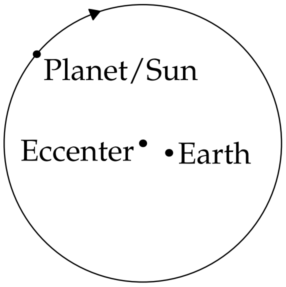
Retrograde motion Planets mostly follow the east-west motion of the heavens, though sometimes they are seen to slow down and reverse course.

If planets move at constant speed around circles centered on the earth, how can these observations be explained? Attempts to produce accurate models while preserving spherical/circular motion led to the development of new mathematics.

One early approach is due to Eudoxus of Knidos (c. 370 BC, page 21), who developed a concentric-sphere model where planets and the sun are attached to separate spheres, each of which has its poles attached to the sphere outside it, with the outermost sphere being that of the fixed stars. The motion generated by such a model[18] is highly complex. Eudoxus' approach is capable of producing retrograde motion, but not the variable brightness of stars and planets.

---

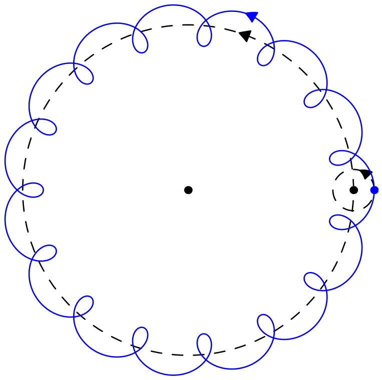
[18] The link is to a very nice flash animation of Eudoxus' model that would have been far beyond Eudoxus' ability to visualize and measure.



Epicycles & Eccentric Orbits Apollonius of Perga ($2^{\text{nd}}/3^{\text{rd}}$ C. BC) is most famous for his study of conic sections, but is relevant here for developing two models of solar/planetary motion.

In the *eccenter* model, a planetary/solar orbit is a circle (the deferent) whose center is *not* the earth. This straightforwardly addresses the problem of variable brightness since the planet is not a fixed distance from the earth.

The obvious criticism is *why?* What philosophical justification could there be for the eccenter? Eudoxus' model might have been complex and impractical, but was more in line with the assumptions of spherical/circular motion.



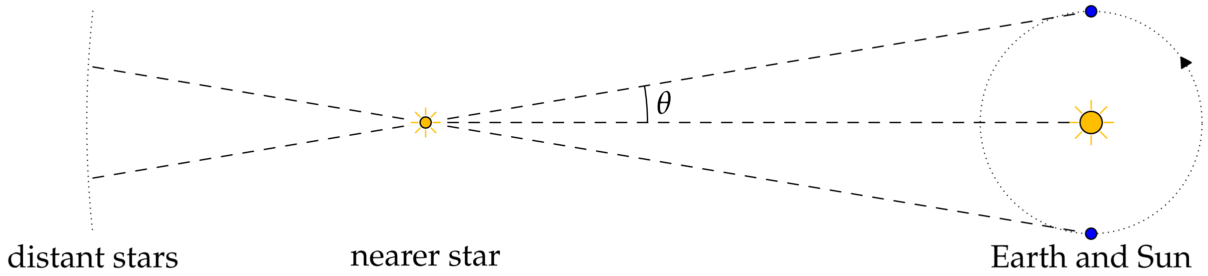
Apollonius' second approach used *epicycles*: small circles attached to a larger circle—you'll be familiar with these if you've played with the toy *Spirograph.* An observer at the center sees the apparent brightness change, and potentially observes retrograde motion. In modern language, the motion is parametrized by the vector-valued function

$$\mathbf{x}(t) = R\begin{pmatrix}\cos\omega t\\ \sin\omega t\end{pmatrix} + r\begin{pmatrix}\cos\psi t\\ \sin\psi t\end{pmatrix}$$

where R, r, ω, ψ are the radii and frequencies (rad/s) of the circles.

Combining these models allowed Apollonius to describe very complex motion. Calculation was difficult however, requiring finding lengths of chords of various circles from a given angle, and vice versa. It is from this requirement that some of the earliest notions of trigonometry arose.

One might ask why the Greeks didn't make the 'obvious' fix and place the sun at the center of the cosmos. In fact Aristarchus of Samos (c. 310–230 BC) did precisely this, suggesting that the fixed stars were really just other suns at exceptional distance! However, the great thinkers of the time (Plato, Aristotle, etc.) had a very strong objection to Aristarchus' proposal: *parallax.*



If the earth moves around the sun and the fixed stars are really independent objects, then the position of a nearer star should appear to change throughout the year. The angle θ in the picture is the *parallax* of the nearer star. Unfortunately for Aristarchus, the Greeks were incapable of observing any parallax.[19] It took 2000 years before the work of Copernicus and Kepler in the 15-1600s forced astronomers to take *heliocentric* models seriously (*Helios* is the Greek sun-god).

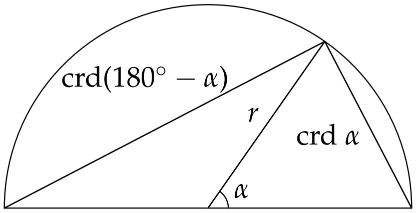
---

[19] The astronomical unit of one *parsec* is the distance to a star exhibiting one arc-second ($\frac{1}{3600}°$) of parallax: roughly 3.3 light-years or 3×10^{13} km, an unimaginable distance to anyone before the scientific revolution. Since the nearest star to our sun, *Proxima Centauri*, lies 4.2 light years $= 0.77$ parsecs away, the rejection of Aristarchus' hypothesis is understandable!

Hipparchus of Nicaea/Rhodes (c. 190–120 BC)

Born in Nicaea (northern Turkey) but doing much of his work on the Mediterranean island of Rhodes, Hipparchus was one of the pre-eminent Greek astronomers. He made use of Babylonian eclipse data to fit Apollonius' eccenter and epicycle models to the observed motion of the moon. As part of this work, he needed to accurately compute chords of circles. His *chord tables* are acknowledged as the earliest lists of trigonometric values.

In an imitation of Hipparchus' approach, we define a function crd which returns the length of the chord in a given circle subtended by a given angle. In modern language



$$\operatorname{crd}\alpha = 2r\sin\frac{\alpha}{2}$$

Hipparchus chose a circle with circumference 360° (in fact he used $60 \cdot 360 = 21600$ arc-minutes), whence $r = \frac{21600}{2\pi} \approx 57,18;$ (base-60). Note that this is sixty times the number of *degrees per radian.*[20] His chord table was constructed starting with two obvious values:

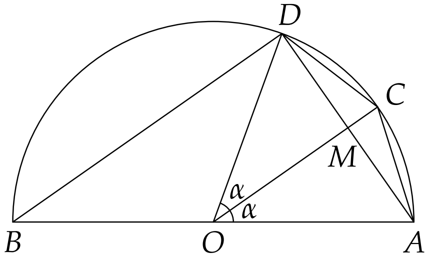
$$\operatorname{crd} 60^\circ = r = 57,18; \qquad \operatorname{crd} 90^\circ = \sqrt{2}r = 81,2;$$

Since (Thales Theorem) the large triangle is right-angled, the Pythagorean theorem can be used to obtain chords for angles $180^\circ - \alpha$. In modern language

$$\operatorname{crd}(180^\circ - \alpha) = \sqrt{(2r)^2 - (\operatorname{crd}\alpha)^2} = 2r\sqrt{1 - \sin^2(\alpha/2)} = 2r\cos\frac{\alpha}{2}$$

Pythagoras was again used to halve and double angles in an approach analogous to Archimedes' quadrature of the circle (page 33). We rewrite the argument in this language.

In the picture, we double the angle α; plainly M is the midpoint of $\overline{AD}$ and $|DB| = \operatorname{crd}(180^\circ - 2\alpha)$. Since $\angle BDA = 90^\circ$, it follows that $\overline{BD}$ is parallel to $\overline{OM}$ and so



$$|OM| = \frac{1}{2}|BD| = \frac{1}{2}\operatorname{crd}(180^\circ - 2\alpha)$$

Now apply Pythagoras to $\triangle CMD$:

$$\begin{aligned}(\operatorname{crd}\alpha)^2 &= \left(\frac{1}{2}\operatorname{crd}2\alpha\right)^2 + \left(r - \frac{1}{2}\operatorname{crd}(180^\circ - 2\alpha)\right)^2 \qquad (|CD|^2 = |DM|^2 + |CM|^2)\\ &= \frac{1}{4}(\operatorname{crd}2\alpha)^2 + r^2 - r\operatorname{crd}(180^\circ - 2\alpha) + \frac{1}{4}\operatorname{crd}(180^\circ - 2\alpha)^2\\ &= \frac{1}{4}(\operatorname{crd}2\alpha)^2 + r^2 - r\operatorname{crd}(180^\circ - 2\alpha) + \frac{1}{4}(4r^2 - (\operatorname{crd}2\alpha)^2)\\ &= 2r^2 - r\operatorname{crd}(180^\circ - 2\alpha) = 2r^2 - r\sqrt{4r^2 - (\operatorname{crd}2\alpha)^2}\end{aligned}$$

In modern notation this is one of the double-angle trigonometric identities!

$$4r^2\sin^2\frac{\alpha}{2} = 2r^2 - 2r^2\cos\alpha \iff \cos\alpha = 1 - 2\sin^2\frac{\alpha}{2}$$

---

[20] One radian is the angle subtended by an arc equal in length to the radius of the circle. Hipparchus essentially does this in reverse: the circumference is fixed so that *degree* now measures both subtended angle *and* circumferential distance.

Example To calculate crd 30°, we start with crd 60° $= r$. Then

$$\begin{aligned} &\operatorname{crd} 120^\circ = \sqrt{4r^2 - r^2} = \sqrt{3}\,r \\ \implies &\operatorname{crd} 30^\circ = \sqrt{2r^2 - r\operatorname{crd}(180^\circ - 60^\circ)} = \sqrt{2r^2 - \sqrt{3}r^2} = \sqrt{2-\sqrt{3}}\,r \end{aligned}$$

In modern language this yields an exact value for sin 15°:

$$\operatorname{crd} 30^\circ = 2r\sin 15^\circ \implies \sin 15^\circ = \frac{1}{2}\sqrt{2-\sqrt{3}}$$

Continuing this process, we obtain crd 150° $= \sqrt{2+\sqrt{3}}\,r$, whence

$$(\operatorname{crd} 15^\circ)^2 = 2r^2 - r\operatorname{crd} 150^\circ = \left(2 - \sqrt{2+\sqrt{3}}\right) r^2 \implies \operatorname{crd} 15^\circ = \sqrt{2-\sqrt{2+\sqrt{3}}}\,r$$

Again translating: $\sin 7.5^\circ = \frac{1}{2}\sqrt{2-\sqrt{2+\sqrt{3}}}$.
In similar fashion, Hipparchus computed the chords of 7.5°, 15°, . . . , 172.5°, in steps of 7.5°. Of course everything was an estimate since he had to rely on repeated approximations for square-roots. All Hipparchus' original work is lost, so we know of his approach only by reference. In particular, while the above method is probably due to Hipparchus, we see it first in the work of Ptolemy, as we'll consider next.

Exercises 4.1. 1. Calculate crd 150°, crd 165°, and crd $172\frac{1}{2}$° using the method of Hipparchus.
(*Leave your answers as a multiple of* $r =$ crd 60°)

2. *Sirius*, the brightest star in the sky, is 2.64 parsecs (8.6 light-years) from the sun. Use modern trigonometry to find its parallax.

3. The tropic of cancer is the line of latitude (approximately) 23.5° north of the equator marking the locations where the sun is directly overhead at noon on the summer solstice.[21] At the arctic circle on the *winter* solstice, the sun is precisely on the horizon.

   (a) Explain why the latitude of the arctic circle is 66.5° north.

   (b) Find the angle the sun makes *above* the horizon at the arctic circle at noon on the summer solstice.

4. Consider the epicycle model where the position vector of a planet is given by

$$\mathbf{x}(t) = R\begin{pmatrix}\cos\omega t\\ \sin\omega t\end{pmatrix} + r\begin{pmatrix}\cos\psi t\\ \sin\psi t\end{pmatrix}$$

   (a) Suppose $R = 4$ and $r = 1$, $\omega = 1$ and $\psi = 2$, so that the epicycle rotates twice every orbit. Sketch a picture of the full orbit.

   (b) Suppose that ω, ψ are positive constants. Prove that an observer will see retrograde motion if and only if $r\psi > R\omega$.
   (*Hint: differentiate* $\mathbf{x}'(t)$ *and think about its direction*)

---

[21] Syene (page 38) is almost exactly on the Tropic of Cancer.

4.2 Ptolemy's *Almagest*

Born in Egypt and living much of his life in Alexandria, Claudius Ptolemy (c. AD 100–170) was a Greek/Egyptian/Roman[22] astronomer and mathematician. Around AD 150, he produced the *Mathematica Syntaxis*, better known as the *Almagest*. The latter term is derived from the Arabic *al-mageisti* (great work), reflecting its importance to later Islamic learning.

The *Almagest* is essentially a textbook on geocentric cosmology. It shows how to compute the motions of the moon, sun and planets, describing lunar parallax, eclipses, the constellations, and elementary spherical trigonometry (this last probably courtesy of Menelaus c. AD 100). It contains our best evidence as to the accomplishments of Hipparchus and describes his calculations. The *Almagest* formed the basis of Western/Islamic astronomical theory well into the 1600s.

Ptolemy's Calculations Ptolemy used several innovations to compute more chords at greater accuracy than Hipparchus.

Initial Data Ptolemy took $r = 60$ so that $\operatorname{crd} 60^\circ = 60$. He also had more initial data:

$$\operatorname{crd} 90^\circ = 60\sqrt{2}, \quad \operatorname{crd} 36^\circ = 30(\sqrt{5}-1), \quad \operatorname{crd} 72^\circ = 30\sqrt{10-2\sqrt{5}}$$

Halving/Doubling Angles Ptolemy used what was probably Hipparchus' method:

$$\operatorname{crd}^2 \alpha = 2r^2 - r \operatorname{crd}(180^\circ - 2\alpha) = 60\big(120 - \operatorname{crd}(180^\circ - 2\alpha)\big)$$

$$\operatorname{crd}(180^\circ - \alpha) = \sqrt{(2r)^2 - \operatorname{crd}^2 \alpha} = \sqrt{120^2 - \operatorname{crd}^2 \alpha}$$

approximating square-roots to the desired accuracy. For example,

$$\operatorname{crd} 30^\circ = \sqrt{60(120 - \operatorname{crd} 120^\circ)} = \sqrt{60\left(120 - 60\sqrt{3}\right)} = 60\sqrt{2-\sqrt{3}} \approx 31;3,30$$

Multiple-Angle Formula Ptolemy computed $\operatorname{crd} 12^\circ = \operatorname{crd}(72^\circ - 60^\circ)$, then halved this for angles of 6°, 3°, 1.5°, and 0.75°. Chords for all integer multiples of 1.5° were computed using multiple-angle/addition formulæ.

Interpolation The observation $\alpha < \beta \implies \frac{\operatorname{crd}\beta}{\operatorname{crd}\alpha} < \frac{\beta}{\alpha}$ allowed Ptolemy to compute chords for every half-degree to the incredible accuracy of two sexagesimal places. For approximating between half-degrees, his table indicated how much should be added for each arc-minute ($\frac{1}{60}{}^\circ$). For example, the second line of Ptolemy's table reads

$$1^\circ \qquad 1;2,50 \qquad ;1,2,50$$

The first two columns state that $\operatorname{crd} 1^\circ = 1;2,50$ to two sexagesimal places.[23] The third entry says, for example, that

$$\operatorname{crd} 1^\circ 5' \approx 1;2,50 + 5(;1,2,50) = 1;8,4,10 \approx 1;8,4$$

To obtain these arc-minute approximations, it is believed Ptolemy computed half-angle chords to an accuracy of *five sexagesimal places* (1 part in over 750 million!). The construction of the chord-table must have been a gargantuan task, one for which Ptolemy likely had much assistance.

---

[22] Ptolemy (Ptolemaeus) is a Greek name, while Claudius is Roman, reflecting the changing cultural situation in Egypt.

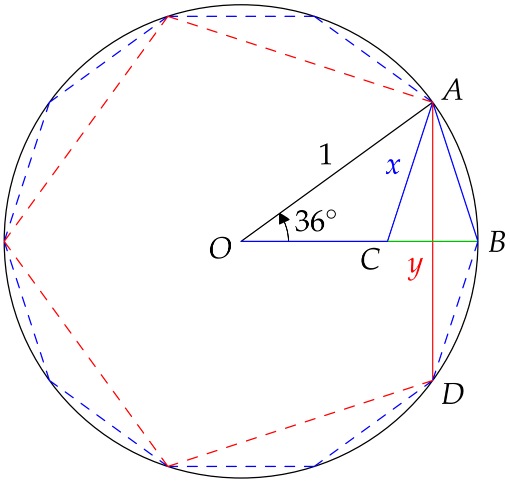
[23] This is $1 + \frac{2}{60} + \frac{50}{60^2} = 1.0472222\ldots = 120 \sin \frac{1.00003625\ldots}{2}{}^\circ$, an already phenomenal level of accuracy.

How did Ptolemy know exact values for crd 36° **and** crd 72°**?** Everything is in Euclid's *Elements*!

Theorem.

1. *(Thm XIII. 9) In a circle, the sides of a regular inscribed hexagon and decagon are in the golden ratio (this ratio is* 60 : crd 36° *in Ptolemy).*
2. *(Thm XIII. 10) In a circle, the square on an inscribed pentagon equals the sum of the squares on an inscribed hexagon and decagon.*

Purely Euclidean proofs are too difficult for us, so here is a way to see things in modern notation.



1. Let $\overline{AB} = x$ be the side of a regular decagon inscribed in a unit circle with center O.

   $\triangle OAB$ is isosceles with angles 36°, 72°, 72°.

   Let C lie on $\overline{OB}$ such that $\overline{AC} = x$.

   Count angles to see that $\triangle OAB$ and $\triangle ABC$ are similar, that $\angle OAC = 36°$ and so $\overline{OC} = x$.

   Similarity now tells us that

   $$x = \frac{1-x}{x} \implies x = \frac{\sqrt{5}-1}{2}$$

   In a circle of radius 60, this gives the exact value

   $$\text{crd}\,36° = 60x = 30(\sqrt{5}-1)$$

2. Now let $\overline{AD} = y$ be the side of a regular pentagon inscribed in the same circle. Applying Pythagoras, we see that

   $$\left(\frac{y}{2}\right)^2 + \left(\frac{1-x}{2}\right)^2 = x^2$$

   Since $x^2 = 1 - x$, this multiplies out to give Euclid's result

   $$y^2 = 1^2 + x^2$$

   from which we obtain the exact value

   $$\text{crd}\,72° = 60y = 30\sqrt{10-2\sqrt{5}}$$

While these values were geometrically precise, Ptolemy used sexagesimal approximations to square-roots to obtain the values stated in his tables:

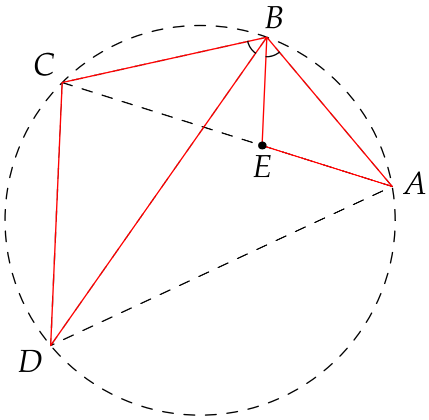
$$\text{crd}\,36° = 37;4,55 \qquad \text{crd}\,72° = 70;32,3$$

He must have used a far higher degree of accuracy in order to obtain similarly accurate values for other chords.

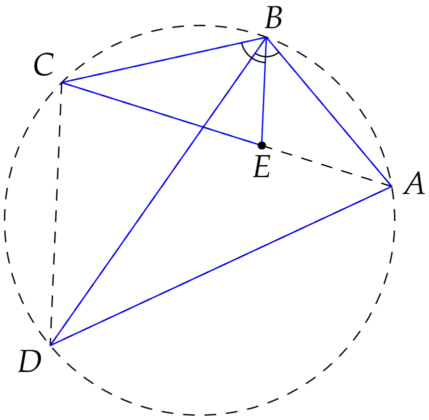
Angle-addition & the Multiple-angle Formula Computation of $\operatorname{crd}(\alpha \pm \beta)$ was facilitated by versions of the multiple-angle formulæ of modern trigonometry.

Theorem (Ptolemy's Theorem). *Suppose a quadrilateral is inscribed in a circle. Then the product of the diagonals equals the sum of the products of the opposite sides.*[24]

Proof. Choose E on $\overline{AC}$ such that $\angle ABE \cong \angle DBC$. Then $\angle ABD \cong \angle EBC$. Since $\angle BAE \cong \angle BDC$ are inscribed angles of the same arc $\overline{BC}$, we obtain two pairs of similar triangles:



$\triangle ABE \sim \triangle DBC$



$\triangle ABD \sim \triangle EBC$

The proof follows immediately: since $\frac{|AE|}{|CD|} = \frac{|AB|}{|BD|}$ and $\frac{|CE|}{|AD|} = \frac{|BC|}{|BD|}$, we have

$$|AC|\,|BD| = (|AE| + |CE|)\,|BD| = |AB|\,|CD| + |AD|\,|BC|$$

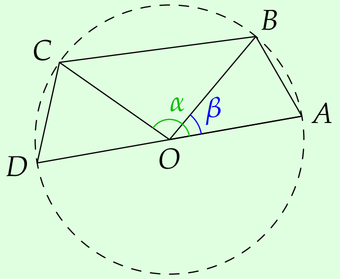
■

Corollary. *If $\alpha > \beta$, then*

$$120 \operatorname{crd}(\alpha - \beta) = \operatorname{crd}\alpha \operatorname{crd}(180^\circ - \beta) - \operatorname{crd}\beta \operatorname{crd}(180^\circ - \alpha)$$

In modern language, divide through by 120^2 to obtain

$$\sin\frac{\alpha - \beta}{2} = \sin\frac{\alpha}{2}\cos\frac{\beta}{2} - \sin\frac{\beta}{2}\cos\frac{\alpha}{2}$$



Proof. If $|AD| = 120$ is a diameter of the pictured circle, then Ptolemy's Theorem says

$$\operatorname{crd}\alpha \operatorname{crd}(180^\circ - \beta) = \operatorname{crd}\beta \operatorname{crd}(180^\circ - \alpha) + 120 \operatorname{crd}(\alpha - \beta)$$

■

Similar expressions for $\operatorname{crd}(\alpha + \beta)$ and $\operatorname{crd}(180^\circ - (\alpha \pm \beta))$ were also obtained, essentially recovering all versions of the modern multiple-angle formulæ for $\sin(\alpha \pm \beta)$ and $\cos(\alpha \pm \beta)$.

---

[24] It is generally considered that this result predates Ptolemy, though there is some debate as to whether it belongs in the *Elements*. Book VI traditionally contains 33 propositions, however some editions append four corollaries of which Ptolemy's Theorem is the last (Thm VI. D).

Examples 1. Here is how Ptolemy might have calculated $\operatorname{crd} 42°$. Let $\alpha = 72°$ and $\beta = 30°$, then

$$120 \operatorname{crd} 42° = \operatorname{crd} 72° \operatorname{crd} 150° - \operatorname{crd} 30° \operatorname{crd} 108°$$

Since $\operatorname{crd} 72° = 30\sqrt{10 - 2\sqrt{5}}$ is known, and

$$\operatorname{crd} 108° = \operatorname{crd}(180° - 2 \cdot 36°) = 120 - \frac{1}{60} \operatorname{crd}^2 36° = 30(1 + \sqrt{5})$$

we see that

$$\begin{aligned} \operatorname{crd} 42° &= \frac{1}{120}\left(30\sqrt{10 - 2\sqrt{5}} \cdot 60\sqrt{2 + \sqrt{3}} - 60\sqrt{2 - \sqrt{3}} \cdot 30(1 + \sqrt{5})\right) \\ &= 15\left(\sqrt{10 - 2\sqrt{5}} \cdot \sqrt{2 + \sqrt{3}} - (1 + \sqrt{5})\sqrt{2 - \sqrt{3}}\right) \approx 43;0,15 \approx 43.0042 \end{aligned}$$

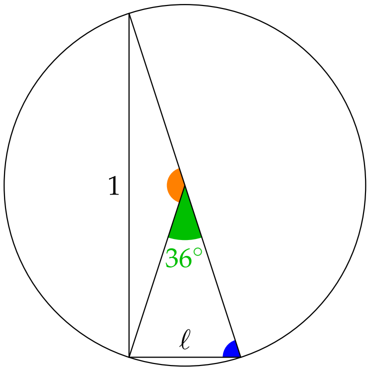
Note all the square-roots which had to be approximated!

2. The *Almagest* also contained many practical examples. Here is one such.

A stick of length 1 is placed in the ground. The angle of elevation of the sun is 72°. What is the length of its shadow?

Ptolemy tells us to draw a picture. The lower isosceles triangle has base angles 72° and the length of the shadow is ℓ. The ratio of the chords is then computed:

$$\begin{aligned} 1 : \ell &= \operatorname{crd} 144° : \operatorname{crd} 36° \\ \implies \ell &= \frac{\operatorname{crd} 36°}{\operatorname{crd} 144°} = \frac{30(\sqrt{5} - 1)}{30\sqrt{10 + 2\sqrt{5}}} \approx 0.32491 \end{aligned}$$



This is precisely $\cot 72°$, though Ptolemy had no such notion.

Exercises 4.2. 1. What are the exact values of $\sin 36°$ and $\sin 18°$?

2. (a) Restate the interpolation formula $\alpha < \beta \implies \frac{\operatorname{crd} \beta}{\operatorname{crd} \alpha} < \frac{\beta}{\alpha}$ in terms of the sine function. What facts about $\frac{\sin x}{x}$ does this reflect?
   (b) Find $\operatorname{crd} 57'$ (arc-minutes!) to two sexagesimal places.

3. Find the exact value of $\operatorname{crd} 54°$

4. Prove the following using Ptolemy's Theorem. What is this in modern language?

$$120 \operatorname{crd}\big(180° - (\alpha + \beta)\big) = \operatorname{crd}(180° - \alpha) \operatorname{crd}(180° - \beta) - \operatorname{crd} \alpha \operatorname{crd} \beta$$

5. Use Ptolemy's Theorem to establish a version of the double-angle formula: $\sin \alpha = 2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2}$.
(*Hint: draw a symmetric quadrilateral one of whose diagonals is a diameter*)

6. Calculate the length of a noon shadow of a pole of length 60 using Ptolemy's methods:
   (a) On the vernal equinox at latitude 40°.
   (b) At latitude 36° north on both the summer and winter solstices.
   (*Hint: recall Exercise 4.1.3*)

5 Ancient Chinese Mathematics

Documented civilization Dates from c. 3000 BC. Until around AD 200, China refers roughly to the area in the map: a small fraction of the modern state's territory north of the Yangtze and centered around the Yellow river.

Earliest mathematics *Oracle Bone* enumeration dates from the Shang dynasty (c. 1600–1046 BC) commensurate with the earliest known Chinese character script. Most information on the Shang comes from later commentaries, though original oracle bones have been excavated, particularly from the ancient capital Anyang. Astronomy, the calendar and trade were dominant drivers of mathematical development.

Zhou dynasty (1046–256 BC) and the Warring States period (475–221 BC) Many mathematical texts were written, though most have been lost; their content must be inferred from later commentaries. Rapid change created pressure for new systems of thought and spurred technological development. Feudal lords employed philosophers, of whom the most famous was Confucius (c. 500 BC).[25] Technological developments included the compass (for navigation) and the use of iron in warfare.[26]

Later history and expansion Between the 221 BC victory of the Qin Emperor Shi Huang Di[27] and the forced abdication (at the age of 6) of the last Qing Emperor Puyi in 1912, China was ruled by a succession of dynasties. By the end of the Qing, Chinese territory had expanded to roughly its modern borders. The Chinese Civil War (1927–1949) resulted in victory for the communists under Mao Zedong and the foundation of the modern Chinese state. While this simple description might suggest a long calm in which culture and technology could develop in comfort, in reality the empire experienced many rebellions, schisms and flux, often exacerbated by the changing whims of emperors and later leaders.

Transmission of knowledge East Asia (modern China, Korea, Japan, etc.) is geographically separated from other areas of early civilization by tundra, desert, mountains and jungle. During the Han dynasty (c. 200 BC–AD 220) a network of trading routes known as the *silk road* was established, connecting China, India, Persia and Eastern Europe; the Great Wall was in part constructed to protect these trade routes. Geographical separation meant that trade was limited, and there is little evidence of mathematical and philosophical ideas making the journey until many centuries later. For instance, there is no evidence of sexagesimal notation being used in China, suggesting that Babylonian and Greek astronomy did not travel eastwards beyond India. Similarly, certain eastern mathematical ideas such as matrix-style calculations saw no analogue in the west until many centuries later. There are, however, indications that early decimal calculations in India may have been inspired by the Chinese counting board approach. On balance, it seems reasonable to conclude that Chinese and Mediterranean mathematics developed essentially independently.

---

[25] Confucius was an adviser to Lu, a vassal state of the Zhou. *Confucianism* emphasises stability and unity as a counter to turmoil. *Taoism,* the competing contemporary philosophical system, is more comfortable with change and adaptation. Very loosely these were the conservative and progressive political philosophies of their day.

[26] Sun Tzu's military classic *The Art of War* dates from this time.

[27] Famous for book-burning, rebuilding the great walls, and for the Terracotta Army of Xi'an.

Early Mathematical Texts

Zhou Bi Suan Jing (*The Mathematical Classic of the Zhou Gnomon*[28] *and the Circular Paths of Heaven*) The oldest suspected Chinese mathematical work was likely compiled some time in the period 500–200 BC. Largely concerned with astronomical calculations, it was presented in the form of a dialogue between the 11th century Duke of Zhou (of *I Ching* fame) and Shang Gao (one of his ministers, and a skilled mathematician). It contained perhaps the earliest statement of Pythagoras' Theorem as well as simple rules for fractions and arithmetic.

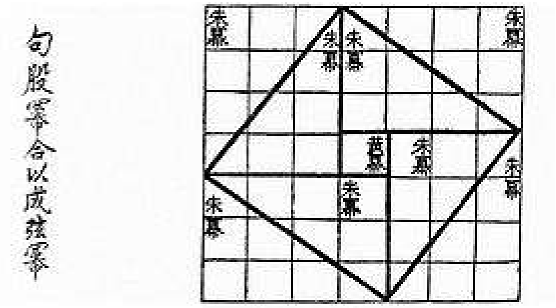
Suanshu Shu (*A Book on Arithmetic*) Compiled around 300–150 BC, it covered topics such as fractions, the areas of rectangular fields, and the computation of fair taxes.

Jiu Zhang Suan Shu (*Nine Chapters on the Mathematical Arts*) Written between 300 BC and AD 200, this the most famous ancient Chinese mathematical text. Many topics are covered, including square roots, ratios (false position and the rule of three[29]), simultaneous linear equations, areas and volumes, right-angled triangles, etc. The *Nine Chapters* was hugely influential, in part due to the detailed commentary and solution manual to its 246 problems written by Liu Hui in AD 263. Several of our examples below come from Liu's work.

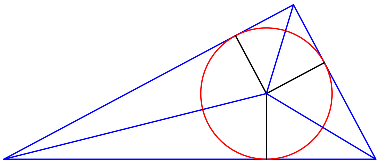
These texts typically involved worked examples with wide application. There is no notion of axiomatics on which one could construct a modern-style proof.

The *Gao Gu* Unsurprisingly, the Chinese do not attribute Pythagoras' Theorem to the Greeks: its name instead refers to the shorter and the longer of the two non-hypotenuse sides of the triangle.

Here is an early mention of the *gao gu*. Is this a 'proof'? Is it a claim about *all* right-triangles, or merely an observation of the triple $(3, 4, 5)$? It can be made rigorous (see Exercise 4), but it is unclear whether this was the intention of the author.



Another example describes how to find the diameter of the circle inscribed in a right-triangle with *gao* 8 and *gu* 15. A picture was drawn and the answer stated:



$$d = \frac{2 \cdot 8 \cdot 15}{8 + 15 + 17} = 6$$

Here is a modern explanation. Given $a = 8$ and $b = 5$, the hypotenuse is $c = \sqrt{8^2 + 15^2} = 17$, and the area of the large triangle is the sum of three smaller triangles, each having height $\frac{1}{2}d$:

$$\frac{1}{2}ab = \frac{1}{2}a \cdot \frac{1}{2}d + \frac{1}{2}b \cdot \frac{1}{2}d + \frac{1}{2}c \cdot \frac{1}{2}d \implies d = \frac{2ab}{a+b+c}$$

Again we ask: is this a general method or an example?

---

[28] Gnomon: "One that knows or examines." Also refers to the elevated piece of a sun/moondial.

[29] Given equal ratios $a : b = c : d$, where a, b, c are known, then $d = \frac{bc}{a}$.

The Bamboo Problem Here is another problem from the *Nine Chapters*, as depicted in Yang Hui's 1262 *Analysis of the Nine Chapters*.

折抵地爲弦以句及股弦并求股故先令句自乘見矩冪令如高而一凡爲高一丈爲股弦并之以除此冪得差所得以減竹高而半其餘卽折者之高也此率與係索之類更相返覆也亦可如上術令高自乘爲股弦并冪去本自乘爲矩冪減之餘爲實倍高爲法則得折之高數也

去根如勾折處

如股折梢如弦

通長如股弦和

股弦和與勾求股法曰勾自乘爲實幷股弦較乘股弦和如股弦和而一乙除得股弦較以減股弦和餘二段

A bamboo has height 10 *chi*. It breaks and the top touches the ground 3 *chi* from the base of the stem. What is the height of the break?

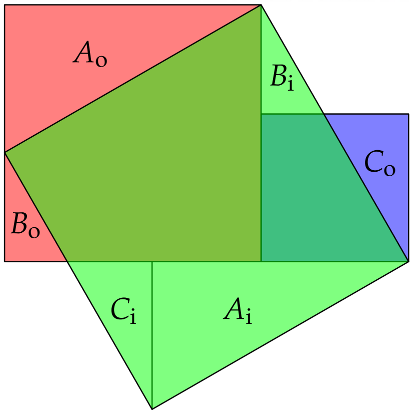
In modern language: if a, b, c are the sides of the triangle with hypotenuse c, we know $b + c = 10$ and $a = 3$; we want b. The solution given is

$$b = \frac{1}{2}\left(10 - \frac{3^2}{10}\right) = \frac{91}{20}\,chi$$

Think about why!

The Out-In Principle Liu made many other contributions to mathematics, including estimating π in a manner similar to Archimedes. He made particular use of the *out-in principle* for comparing area and volume:

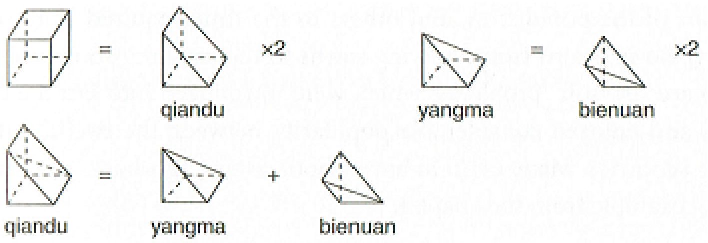
1. Area and volume are invariant under translations.
2. If a figure is subdivided, the sum of the areas/volumes of the parts equals that of the whole.



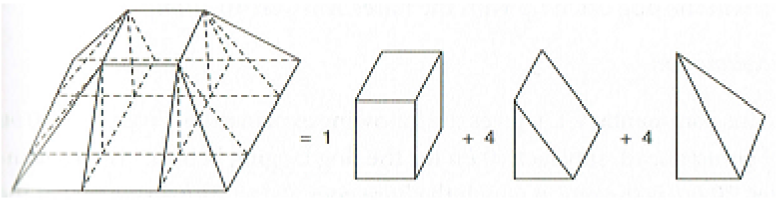
These are essentially axioms for area/volume in Euclidean geometry. For instance, Liu gave the argument shown in the picture in justification of the *gao gu*: the large square is subdivided and the *in* pieces A_i, B_i, C_i translated to new *out* pieces A_o, B_o, C_o to assemble the required squares.

Liu extended the out-in principle to analyze solids, comparing the volumes of four basic solids:

- Cube (*lifang*)
- Right triangular prism (*qiandu*)
- Rectangular pyramid (*yangma*)
- Tetrahedron (*bienuan*)



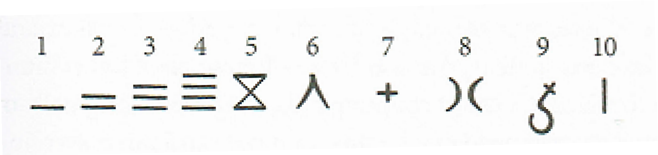
These could be assembled to calculate the volume of, say, a truncated pyramid:

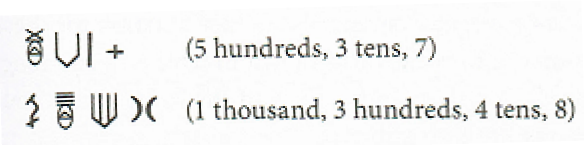


Chinese Enumeration

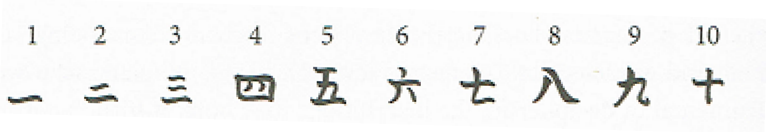
The ancient Chinese had two parallel systems of enumeration. Both are essentially decimal.

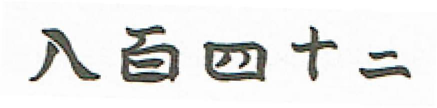
Oracle Bone Script and Modern Numerals The earliest Chinese writing, *oracle bone* script, dates from around 1600 BC. The numbers 1–10 had distinct symbols, as did 20, 100, 1000 and 10000. These were decorated to denote various multiples. Some examples are shown below.





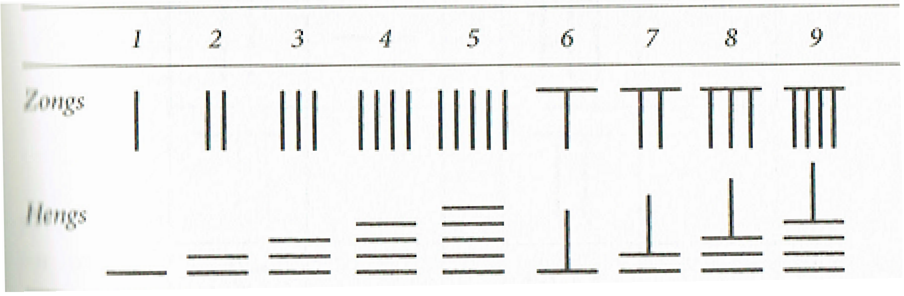
Given all the possibilities for decoration, the system is complex and more advanced than other contemporary systems. Modern Chinese numerals are a direct descendant of this script:





Observe the similarity between the expressions for the first 10 digits. The second image denotes 842, where the second and fourth symbols represent 100's and 10's respectively (literally *eight hundred four ten two*). A zero symbol is not required as a separator: one could not confuse 205 (*two hundred five*) with 250 (*two hundred five ten*). The system is partly positional: for instance the symbol for 8 can also mean 800 if placed correctly, but only if followed by the symbol for 100.

Rod Numerals and the Counting Board The second dominant form of enumeration dates from around 300 BC and was in wide use by AD 300. Numbers were denoted by patterns known as *zongs* and *hengs*: *zongs* represent units, 100's, 10000's, etc., while *hengs* were for 10's, 1000's, 100000's, etc.



Rod numerals were immensely practical—in extremis they could easily be scratched in the dirt! More commonly, short bamboo sticks or *counting rods*—of which any merchant would carry a bundle—would be used in conjunction with a *counting board*: a grid of squares on which sticks could be placed for ease of calculation. This technology facilitated easy trade and gave rise to several methods of calculation which will seem familiar. There was no need for a zero in this system as an empty space did the job. Variations of the rod numeral system persisted in China morphing into the *Suzhou* system which can still be found in some traditional settings.

Basic Counting Board Calculations Addition and subtraction are straightforward by carrying and borrowing in the usual way. The smallest number was typically placed on the right. Multiplication is a little more fun. Here we multiply 387 by 147.

3	8	7
1	4	7

Arrange rods: we use modern numerals for clarity

		3	8	7
4	4	1		
1	4	7		

$3 \times 147 = 441$, note the position of 147

			8	7
4	4	1		
1	1	7	6	
	1	4	7	

Delete 3, move 147 and multiply: $8 \times 147 = 1176$

			8	7
5	5	8	6	
	1	4	7	

Sum rows

				7
5	5	8	6	
	1	0	2	9
		1	4	7

Delete 8, move 147 and multiply: $7 \times 147 = 1029$

5	6	8	8	9
		1	4	7

Sum rows: in conclusion, $387 \times 147 = 56889$

The algorithm is just long-multiplication, starting with the largest digit (3) instead of the units as is more typical in Western education.

Division is similar to long-division. To divide 56889 by 147 one might have the following sequence of boards

5	6	8	8	9
		1	4	7

⟶

		3		
5	6	8	8	9
1	4	7		

⟶

		3	8	
1	2	7	8	9
	1	4	7	

⟶

		3	8	7
	1	0	2	9
		1	4	7

In the first two boards, 147 goes 3 times into 568.
In board 3, we subtract 3×147 from 568 to leave 127, shift 147 one place to the right, and observe that 147 goes 8 times into 1278.
In the final step we have subtracted 8×147 from 1278 to leave 102, before shifting 147 to its final position on the right. Since 147 divides exactly seven times into 1029, we are done.

There is nothing stopping us from dividing when the result is not an integer; one simply continues as in long-division, with fractions represented as decimals.

Simultaneous linear equations The coefficients of a linear system were placed in adjacent *columns* and then *column operations* were performed. The method is identical to what you learn in a linear algebra class, but with columns rather than rows. Here is an example.

$$\begin{cases}3x+2y=7\\2x+y=4\end{cases} \longrightarrow \begin{array}{|c|c|}\hline 3&2\\\hline 2&1\\\hline 7&4\\\hline\end{array} \longrightarrow \begin{array}{|c|c|}\hline 1&2\\\hline 1&1\\\hline 3&4\\\hline\end{array} \longrightarrow \begin{array}{|c|c|}\hline 1&1\\\hline 1&0\\\hline 3&1\\\hline\end{array} \longrightarrow \begin{array}{|c|c|}\hline 0&1\\\hline 1&0\\\hline 2&1\\\hline\end{array} \longrightarrow x=1,\ y=2$$

This matrix method was essentially unique to East Asia until the 1800s.

Euclidean algorithm The counting board lent also itself to the computation of greatest common divisors, which were used for simplifying fractions. Here is the process applied to $\frac{35}{91}$:

$$\begin{array}{llllll} 35 & 35 & 35 & 14 & 14 & 7\\ 91 & 56 & 21 & 21 & 7 & 7 \end{array}$$

At each stage, one subtracts the smaller number from the larger. Once the same number is in each row you stop. You should recognize the division algorithm at work! Since $\gcd(35,91)=7$, both could be divided by 7 to obtain $\frac{35}{91}=\frac{5}{13}$ in lowest terms.

Negative numbers There is a strong case that the Chinese are the oldest adopters of negative numbers, though these were not originally thought of as such. Different colored rods were used to denote a deficiency in a quantity, commonly when balancing accounts. The *Nine Chapters* describes using red and black rods in this manner. This practice was known by AD 1, roughly 500 years before negative numbers were used in calculations in India. It is possible that there was some transference of this idea from China to India.

Music, Mysticism and Approximations Like the Pythagoreans, the ancient Chinese were interested in music and pattern for mystical reasons. While the Pythagoreans delighted in the pentagram, the Chinese created *magic squares* (grids whose rows, columns and diagonals sum to the same total) as symbols of perfection.

8	3	4
1	5	9
6	7	2

3×3 magic square

The notion of equal temperament in musical tuning was first 'solved' in China by Zhu Zaiyu (1536–1611), some 30 years before Mersenne & Stevin established the same idea in Europe. This required the computation of the twelfth-root of 2 which Zhu found using approximations for square and cube roots:

$$\sqrt[12]{2}=\sqrt[3]{\sqrt{\sqrt{2}}}$$

Zhu's approximation was correct to 24 decimal places! Indeed the Chinese emphasis on practicality meant that they often had the most accurate mathematical approximations of their time:

- Approximations to π including $\frac{22}{7}$, $\sqrt{10}$, $\frac{355}{113}$, $\frac{377}{120}$. Most accurate in the world from 400–1400.
- Methods for approximating square and cube roots were found earlier than in Europe. Approximations to solutions of higher-order equations similar to the Horner–Ruffini/Newton–Raphson method were also discovered earlier.
- Pascal's triangle first appeared in China around 1100. It later appeared in Islamic mathematics before making its way to Europe.

Two Famous Problems

We finish with two famous Chinese problems. The first is known as the *Hundred Fowl Problem* and dates from the 5th century AD. It was copied later in India and then by Leonardo da Pisa (Fibonacci) in Europe, thus demonstrating how some Chinese mathematics travelled westwards.

> If cockerels cost 5 *qian* (a copper coin), hens 3 *qian*, and 3 chicks cost 1 *qian*, and if 100 fowl are bought for 100 *qian*, how many cockerels, hens and chicks are there?

In modern language, we want non-negative integers x, y, z satisfying

$$\begin{cases} 5x + 3y + \frac{1}{3}z = 100 \\ x + y + z = 100 \end{cases}$$

The stated answers are $(4, 18, 78)$, $(8, 11, 81)$, $(12, 4, 84)$ while the solution $(0, 25, 75)$ was ignored.

Finally we consider the *Chinese Remainder Theorem* for solving simultaneous congruence equations. This result dates from the 4th century AD, after which it travelled to India where it was described by Bhramagupta, and thence to Europe. This example comes from Qin Jiushao's *Shu Shu Jiu Zhang* (Nine Sections of Mathematics, 1247).

> Three thieves stole three identical vessels filled with rice, but whose exact capacity was unknown. The thieves were caught and their vessels examined: the quantities left in each vessel were 1 *ge*, 14 *ge* and 1 *ge* respectively. The thieves did not know the exact quantities they'd stolen. The first used a horse ladle (capacity 19 *ge*) to take rice from the first vessel. The second used a wooden shoe (17 *ge*) to take rice from his vessel. The third used a bowl (12 *ge*). What was the total amount of rice stolen?

In modern language, the capacity x of each vessel satisfies

$$x \equiv 1 \pmod{19}, \qquad x \equiv 14 \pmod{17}, \qquad x \equiv 1 \pmod{12}$$

The given answer, $x = 3193$ *ge*, represents the smallest possible capacity of each vessel, with all other solutions being congruent modulo $19 \cdot 17 \cdot 12 = 3876$, as you should be able to confirm if you've studied number theory! The total amount of rice stolen is then

$$(x - 1) + (x - 14) + (x - 1) = 3x - 16 = 9563$$

Since congruence equations are simply underdetermined linear equations

$$x \equiv 1 \pmod{19} \iff \exists y \in \mathbb{Z} \text{ such that } x = 1 + 19y$$

solutions to both of these problems can be effected using counting board methods.

Exercises 5. 1. Verify the result of the *Bamboo problem.*

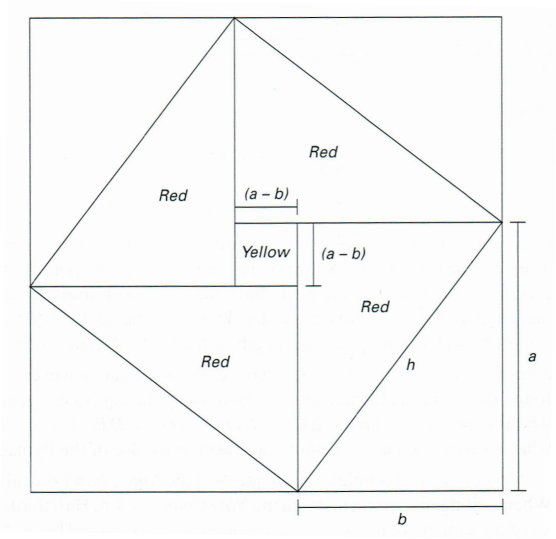
2. Solve the *Hundred Fowl Problem* by substituting $z = 100 - x - y$ in the first equation and observing that x must be divisible by 4.

3. Use a counting board method to:

   (a) Solve the linear system $\begin{cases} 8x + y = 28 \\ 3x + 2y = 17 \end{cases}$

   (b) Multiply 218×191.

4. Turn Zhao Shuang's pictorial argument from the *Arithmetical Classic of the Gnomon* into a proof of Pythagoras' Theorem.



5. Solve problem 24 of chapter 9 of the *Nine Chapters.*

   > A deep well 5 ft in diameter is of unknown depth (to the water level). If a 5 ft post is erected at the edge of the well, the line of sight from the top of the post to the edge of the water passes through a point 0.4 ft from the lip of the well below the post. What is the depth of the well?

6. Solve problem 26 of chapter 6 of the *Nine Chapters.*

   > Five channels bring water into a reservoir. If only the first channel is open, the reservoir fills in $\frac{1}{3}$ of a day. The second channel by itself fills the reservoir in 1 day, the third channel in $2\frac{1}{2}$ days, the fourth in 3 days, and the fifth in 5 days. If all the channels are opened together, how long will the reservoir take to fill?

6 Indian and Islamic Mathematics

6.1 India, the Hindu–Arabic Numerals & Zero

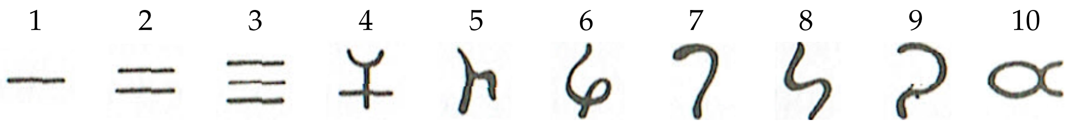
The Indian/South Asian subcontinent is bordered to the north by the Himalayan mountains and to the east by dense jungle. Its primary historical frontier comprised the fertile Indus valley to the west, now the central corridor of Pakistan, where recorded civilization dates to at least 2500 BC. During the first millennium BC, Hinduism developed as an amalgamation of previous practices and beliefs; Buddhism and Jainism began to spread in the later part of this period, particularly in the Ganges valley further east.

Alexander the Great's conquests reached the Indus in 326 BC, bringing Greek, Babylonian and Egyptian knowledge in his wake. The Greek overlords he left behind were rapidly overthrown and the subcontinent became largely unified under the Mauryan Empire for the next 150 years. After this came 1000 years of shifting control with several invasions from the west by the Persians. Islam conquered the Indus around AD 1000, with most of India becoming part of the Islamic Mughal Empire by the 1500s; after the Mughal decline and fragmentation, the British became dominant in 1857.

The modern political situation reflects this complicated history. India gained independence from Britain in 1947 after World War II and was shortly thereafter partitioned according to religion: the greater Indus valley and the lower Ganges/Brahmaputra comprise the modern Islamic states of Pakistan and Bangladesh, with the majority of the landmass becoming the nominally secular but majority Hindu *country* of India. The upper Indus valley (Kashmir) remains contested and has been the site of several military conflicts between India, Pakistan and China.

Ancient India's contributions to world knowledge and development are significant; it is estimated that India accounted for 25–30% of the world's economy during the 1st millenium AD! It was moreover a technological and cultural crossroads between East (China) and West (Greece, Persia, Rome, etc.); while some trade and knowledge passed north of the Himalayas directly between China and the Middle East/Europe, far more percolated slowly through India, being improved upon and given back in turn.

Brahmi Numerals & Numerical Naming Our primary focus is on possibly the most important practical mathematical development in history: the decimal positional system of enumeration, complete with fully-functional zero. The Brahmi numerals, one of the earliest antecedents of modern numerals, first appeared around the 3rd century BC.



The example dates from around 100 BC and was used in Mumbai/Bombay. Additional symbols denoted multiples of 10, 100, 1000, 10000, etc. As with Chinese characters, the system was partly positional (800 would be written by prefixing the symbol for 100 by that for 8) and there was no symbol or placeholder for zero.

Symbols are only part of the story. The modern approach to naming numbers and constructing large numbers can also be linked to the same period. The table below gives old Sanskrit names.

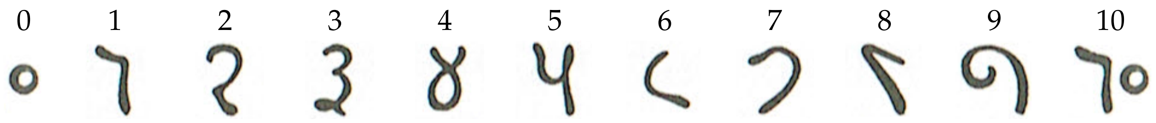
1	2	3	4	5	6	7	8	9
eka	dvi	tri	catur	pancha	sat	sapta	asta	nava
10	20	30	40	50	60	70	80	90
dasa	vimsati	trimsati	catvarimsat	panchasat	sasti	saptati	asiti	navati
100	1000	10000	100000	1000000	10^7	10^8	10^9	10^{10}
sata	sahasra	ayuta	niyuta	prayuta	arbuda	nyarbuda	samudra	madhya

Many European languages have Sanskrit roots; it should be no surprise that several ancient Sanskrit numbers are similar (e.g., *dva* in Russian, *quatre* in French). The construction of larger numbers should also seem familiar: for example *tri sahasra sat sata panchasat nava* is precisely how we *read* 3659.

Such familiarity has its limits, for old Sanskrit verbiage doesn't map perfectly onto modern English. For instance, old Sanskrit had distinct words for powers of 10 up to (at least!) 10^{62}, and employed a version of pre-subtraction: e.g., *ekanna-niyuta* meant 'one less than 100000,' or 99999.

Gwalior Numerals During the first few centuries AD, a fully positional decimal place system came into being. The earliest evidence comes from a manuscript found in Bakhshālī (Pakistan) in 1881, which has been carbon-dated to the 3rd or 4th century. The manuscript contains the earliest known version of the modern symbol for zero, a circular dot. It is conjectured that the decimal place system was inspired by the Chinese counting-board method, though convincing proof has yet to be uncovered. Regardless of attribution, Chinese mathematicians were copying the method by the 8th century.

The examples below are better understood than the Bakhshālī manuscript and come from Gwalior (northern India) around 876.



The similarity with modern numerals is clear; 0, 1, 2, 3, 4, 7, 9, 10 are very familiar. Zero has evolved from the Bakhshālī dot to a hollow circle. The symbols for 2 and 3 are conjectured to have developed in an attempt to write earlier versions (e.g. the Brahmi numerals) cursively; try writing three horizontal strokes quickly…

The system is fully positional. Below are the numbers 270 and 30984:

Sanskrit is written left-to-right, with the leftmost digits representing the largest powers of 10. Note how zero is used as a placeholder to clarify position so that, e.g., 27, 207, and 270 are clearly distinguishable.

Zero On the right is a table of modern Sanskrit names and numerals; the digits and names are certainly similar to their Gwalior counterparts.

०	१	२	३	४
0	1	2	3	4
shuunyá	ekaḥ	dvau	tryaḥ	catvāraḥ
५	६	७	८	९
5	6	7	8	9
pañca	ṣaṭ	sapta	aṣṭa	nava

The Sanskrit *shuunyá* means *void* or *emptiness.* It is related to *svi* (hollow), which in turn derives from an ancient word meaning *to grow.* This reflects a major idea within religions of the area, with the void being the source of all things, of creation and creativity. Contemplation of the void (the doctrine of Shunyata) is recommended before composing music, creating art, etc. This contrasts with the Abrahamic religions where the void is something to be feared; an early conception of hell was the eternal absence of God.

The Gwalior numerals travelled westwards, with Europe eventually inheriting the system via Islam; as such they are today known the *Hindu–Arabic* numerals. Here is a short version of the etymological journey of zero into European languages.

- *Shunya* was transliterated to *sifr* in Arabic where the double-meaning persisted: *al-sifr* was the number zero, while *safira* meant *it was empty.*
- The term came to Europe in the 12th-13thcenturies courtesy of Fibonacci where it became *cifra.* This was blended with *zephyrum* (*west wind/zephyr*) providing an alternate spelling.
- Cifra ultimately became the words *cipher* (English), *chiffre* (French) and *ziffer* (German), meaning a figure, digit, or code.
- Zephyrum became *zefiro* in Italian and *zero* in Venetian.

Zero and the Hindu–Arabic numerals also travelled eastwards, with Qin Jiushao introducing the zero symbol into China in the 13th century.

Our modern understanding of zero is a fusion of several concepts:

Numerical positioning For instance, to distinguish 101 from 11.

Absence of a quantity 101 contains no 10's.

Symbol First a dot (*bindu*), then a circle (*chidra/randhra* meaning *hole*). The relationship between *shunya* and a symbol was established by AD 2-300, as this quote from AD 400 (Vasavadatta) illustrates

> The stars shone forth, like zero dots [shunya-bindu] scattered as if on a blue rug. The Creator reckoned the total with a bit of the moon for chalk.

Mathematical operations By the time of Brahmagupta (7th C.), a mathematical text might contain a section called *shunya-gania,* with computations involving zero, including addition, multiplication, subtraction, effects on $\pm$-signs, division and the relationship with ∞ (*ananta*). In the 12th C., Bhaskaracharya stated:

> If you were to divide by zero you would get a number that was "as infinite as the god Vishnu."

Other ancient cultures had one or more of these aspects of zero, but the Indians were the first to put them all together.

- The Egyptian hieroglyph *nfr* (beautiful/complete) indicated zero remainder in calculations as early as 1700 BC and was also used as a reference point/level in buildings.
- Very late in Babylonian times, a placeholder symbol was used to separate powers of 60. It was not used as a number.
- With the Chinese counting board, an empty space served as a placeholder.
- Various Mesoamerican cultures, such as the Maya, had a zero symbol that was used as a placeholder, particularly when writing dates.

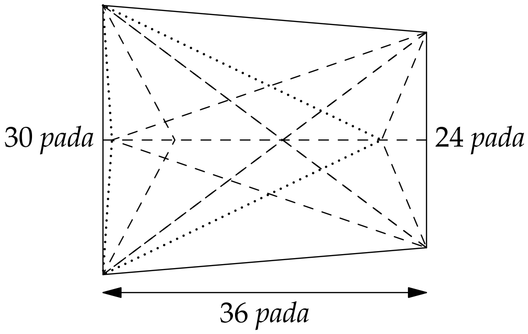
'Real' Indian Mathematics

Indian mathematicians made great progress on several fronts, not merely the decimal place system.

Much ancient work was influenced by religion. For instance, the pre-Hindu *sulbasutras* contained instructions for laying out altars using ruler-and-compass constructions. These could be quite complex, as the construction of the base of the *Mahavedi* (great altar) shows: The center line is divided left-to-right in the ratio

1 : 7 : 12 : 11 : 5

and the altar contains five distinct Pythagorean triples!



Of particular importance to our continuing narrative is Indian work on trigonometry. Here are some highlights:

- The early 5th C. text *Paitāmahasiddhānta* is assumed to be an extension of Hipparchus' work, since it contains a table of chords based on a circle of radius 57,18; rather than Ptolemy's 60.
- Indian mathematicians instituted the use of *half-chords*, in line with our modern understanding of sine. Indeed the word *sine* is the result of a long sequence of (mis)translations and transliterations via Arabic and Latin from the Sanskrit *jyā-ardha* (*chord-half*).[30] The Indians also began to distinguish 'base sine' and 'perpendicular sine' (cosine).
- Created tables of sines/half-chords from 0° to 90° in steps of $3\frac{3}{4}°$, using linear interpolation to approximate values in between. By 650, Bhramagupta had much better approximations, using quadratic polynomials to interpolate. By 1530, Indian mathematicians had discovered cubic and higher approximations (essentially Taylor polynomials 130 years before Newton) for even greater accuracy of sine, cosine and arctangent.

Navigation was one of the drivers of this development. While Mediterranean sailors rarely strayed long out of sight of land, the Indians sailed the ocean and required accurate measurements to find their latitude.

---

[30] This is also the root of the word *sinus* meaning *bay* or *gulf* (e.g., in your nose).

Exercises 6.1. 1. The *Mahavedi* (pg. 57) contains five Pythagorean triples; find them.

2. To simplify square root expressions, Bhaskara used the formula

$$\sqrt{a+\sqrt{b}} = \sqrt{\frac{1}{2}\left(a+\sqrt{a^2-b}\right)} + \sqrt{\frac{1}{2}\left(a-\sqrt{a^2-b}\right)}$$

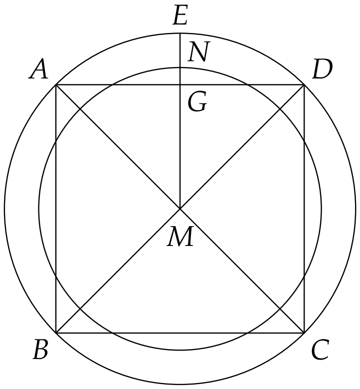
Prove Bhaskara's formula and use it to simplify $\sqrt{2+\sqrt{3}}$.

3. Here is an Indian method for 'finding' a circle whose area is equal to a given square.

In a square $ABCD$, let M be the intersection of the diagonals. Draw a circle with M as the center and MA the radius; let ME be the radius of the circle perpendicular to the side AD and cutting AD at G. Let $GN = \frac{1}{3}GE$. Then MN is the radius of the desired circle.
Show that if $AB = s$ and $MN = r$, then

$$\frac{r}{s} = \frac{2+\sqrt{2}}{6}$$

Show that this implies a value for π equal to 3.088311755.



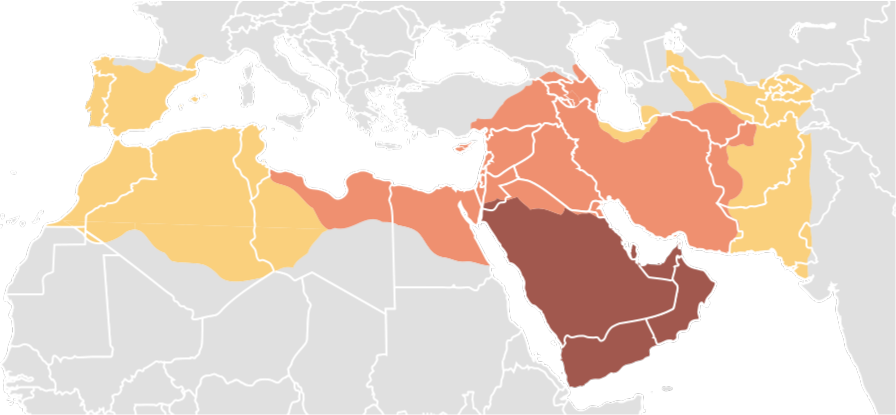
4. Solve the following problem of Mahāvīra.

> Of a collection of mango fruits, the king took 1/6; the queen took 1/5 of the remainder, and the three chief princes took 1/4, 1/3, 1/2 of what remained at each step. The youngest child took the remaining three mangoes. O you, who are clever in working miscellaneous problems on fractions, give out the measure of that collection of mangoes.

6.2 Islamic Mathematics I: Algebra

Muhammad ibn Abdullah was born in Mecca (modern Saudi Arabia) in 570. Around 610 he began preaching *Islam* (*submission to the will of God*)—the third of the major Abrahamic religions, (chronologically) following Judaism and Christianity. After several years of exile, he returned with an army, conquering Mecca a few years before his death in 632.

Through military conquest, Muhammad's successors expanded the caliphate (empire) at a truly remarkable speed. At the time of his death, the Arabian peninsula was Islamic. By 660 Islam had reached Libya and most of Persia, and by 750 extended from Iberia & Morocco to Afganistan & Pakistan. Serious schisms eventually arose[31] and several successor empires emerged, the longest-lasting of which was the Ottoman Empire (c. 1300–1922). Even though centralized political control ended long ago, Islam remains dominant in the region pictured below (with the notable exceptions of Spain and Portugal) and over a greater region of Africa and south-east Asia (e.g., Indonesia).



As with the Romans, early Muslims permitted conquered peoples—including Jews and Christians (*people of the book*)—to maintain their culture, provided they acknowledged their overlords and paid taxes. Those who converted to Islam were welcomed as full citizens, though deconversion (apostasy) was not tolerated. Many of the great Islamic thinkers were born on the periphery and travelled to the great centers of learning, particularly Baghdad during the Islamic golden age (8th–13th centuries). Knowledge was also absorbed from Alexandria and western India (Pakistan). In the mid-700s papermaking came from China, greatly facilitating the dissemination and consolidation of knowledge. Schools (*madrassas*) reflected a strong cultural and religious focus on learning.

The Islamic golden age overlapped the European *dark ages* (c. 500–1200) following the fall of Rome, during which European philosophical development stagnated. By 1200, the crusades[32] were well underway and Islam had come to be seen as the enemy of Christian Europe. The infusion of knowledge that came to Europe from Islam around this time helped spur the European renaissance & later scientific revolution. Among European scholars almost to the present day, it was fashionable to credit Islam merely with the *preservation* of ancient 'European' knowledge; a claim both fanciful and chauvinistic, but plainly stemming from medieval animosity.

---

[31] In particular between the Sunni and Shia branches of the faith. Much of the modern-day tension between Saudi Arabia and Iran stems from this rupture.

[32] A series of religious–military campaigns 1096–1291 with the goal of wresting control of the Holy Land, particularly Jerusalem, from Islam.

Algebra & Algorithms

Proof and axiomatics were learned from Greek texts such as the *Elements.* Like the Greeks, Islamic scholars gave primacy to geometry and proved algebraic relations in a geometric manner.[33] Practical and accurate calculation was more important than to the Greeks, and great advances were made in this area. This included completing the development of the Indian decimal place system (hence the dual credit *Hindu–Arabic* numerals).

The second most obvious legacy of Islamic mathematics is encountered daily in every mathematics classroom. *Algebra*[34] comes from the Arabic *al-ğabr*, meaning *restoring*. It originally referred to moving a deficient (negative) quantity from one side of an equation to another. A second term *al-muqabala* (*comparing/balancing*) meant to subtract the same positive quantity from both sides of an equation.

$$\begin{aligned} &\textit{Al-ğabr}: && x^2+7x=4-2x^2 \implies 3x^2+7x=4 \\ &\textit{Al-muqābala}: && x^2+7x=4+5x \implies x^2+2x=4 \end{aligned}$$

Islamic scholars did not use symbols or equations in a modern sense; statements were instead written out in sentences.

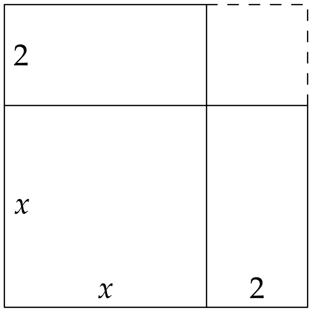
Muhammad ibn Mūsā al-Khwārizmī (780–850) Born near the Aral Sea in modern Uzbekistan, al-Khwārizmī eventually became chief librarian at the great school of learning, the *House of Wisdom,* in Baghdad. His *Compendious book on the calculation by restoring and balancing*[35] (820) is a synthesis of Babylonian methods and Euclidean axiomatics; an algorithm demonstrated a solution, followed by a geometric proof. After being translated into Latin in the 1100s it became a standard textbook of European mathematics, displacing Euclid in places due to its greater emphasis on practical calculation. The word *algorithm* reflects its importance: the Latin *dixit algorismi* literally means *al-Kwārizmī says.*

Example. Here is al-Khwārizmī's approach to the equation $x^2+4x=60$, or, more properly:

> What must be the square which, when increased by four of its roots, amounts to sixty?

The algorithm may be applied to *any* equation of the form $x^2+ax=b$ where $a,b>0$: here a is the number of 'roots,' and b the total 'amount.'

- Halve the number of roots $\left(2=\frac{1}{2}a\right)$
- Multiply by itself $\left(4=\frac{1}{4}a^2\right)$
- Add to the total amount $\left(64=\frac{1}{4}a^2+b\right)$
- Take the root of this $\left(8=\sqrt{\frac{1}{4}a^2+b}\right)$
- Subtract half the number of roots $\left(6=\sqrt{\frac{1}{4}a^2+b}-\frac{a}{2}\right)$



Al-Kwārizmī essentially constructs the quadratic formula $=\frac{-a+\sqrt{a^2+4b}}{2}$, while the pictorial justification is Euclid's (*Elements*, Thm II. 4). The geometry should be obvious: the original square (x^2) has been increased by four of its roots; the algorithm is simply 'completing the square' $(x+2)^2=64$.

---

[33] Like Book II of the *Elements.* Such Greek texts were venerated by Islamic scholars; recognizing the depth of Ptolemy's work on astronomy and trigonometry, they bestowed the name by which it is now known, the *Almagest* (*Great Work*).

[34] Many words beginning *al-* are of Arabic origin (alkali, albatross, etc.), as are others that have been latinized (elixir).

[35] *Al-kitāb al-mukhtasar fī hisāb al-ğabr wa'l-muqābala.*

Other algorithms were supplied to solve every type of quadratic.

It is hard to notice from our example, but the crucial development from a math-history point of view is the abstraction, in a modern sense the *algebra*; al-Khwārizmī's approach applies equally to numbers as it does to geometric objects, a very different approach to the geometry-focused Greeks.

As an example of the power of this idea, consider how Abū Kāmil (Egypt 850–930) generalized Euclid's Book II geometric-algebra arguments to permit substitution, provided the resulting equation was quadratic.

$$\text{If } y = \frac{1+x}{3+x} \text{ and } y^2 + y = 1 \text{ then } x = \sqrt{5}$$

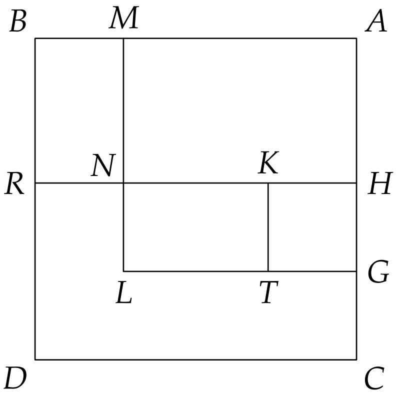
Abū Kāmil essentially substitutes $y = \frac{1+x}{3+x}$ into the quadratic (with solution $y = \frac{\sqrt{5}-1}{2}$). While al-Khwārizmī's methods were geometrically justified, when combined in this fashion the entire process no-longer admits a straightforward geometric interpretation. This method of substitution was an early step towards establishing the modern primacy of algebra and number over geometry and length.

Over the following centuries, this algebraic approach was further improved. In particular, Omar Khayyam (1048–1131) produced ground-breaking work on cubic equations, astronomy, the binomial theorem, and irrational numbers.

Exercises 6.2. 1. Solve the equations $\frac{1}{2}x^2 + 5x = 28$ and $2x^2 + 10x = 48$ using al-Khwārizmī's methods (first multiply or divide by 2).

2. Al-Khwārizmī gives the following algorithm for solving the equation $bx + c = x^2$.

   - Halve the number of roots.
   - Multiply this by itself.
   - Add this square to the number.
   - Extract the square root.
   - Add this to half the roots.

   Translate this into a formula. Give a geometric argument for the validity of the approach using the picture: HC has length b where G is the midpoint; rectangle $ABRH$ has area c; $KHGT$ and $AMLG$ are squares; and the large square $ABDC$ has side-length x.

3. Solve the following problems by Abū Kāmil (use modern algebra!).

   (a) Suppose 10 is divided into two parts and the product of one part by itself equals the product of the other part by the square root of 10. Find the parts.

   (b) Suppose 10 is divided into two parts, each of which is divided by the other, and the sum of the quotients equals the square-root of 5. Find the parts.

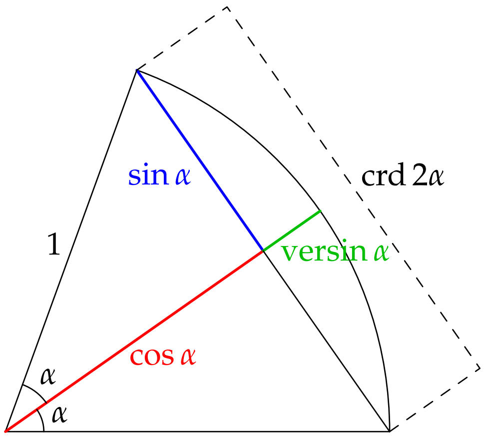
6.3 Islamic Mathematics II: Spherical Trigonometry and the *Qibla*

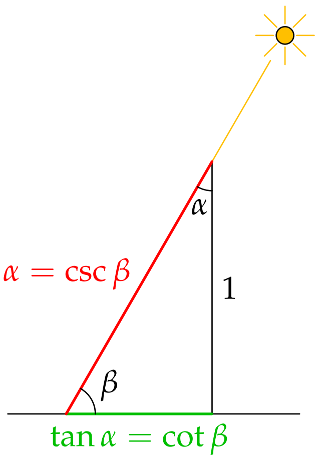
Late 8th century Indian work on trigonometry, linking back to Hipparchus, was known in Baghdad, as was the work of Ptolemy. Islamic scholars were interested in trigonometry for reasons beyond mere astronomy. A primary requirement in Islam is to face the Ka'aba in the Great Mosque at Mecca when at prayer: this is the *qibla* (*direction* in Arabic). A mosque is typically built so that one wall faces Mecca for convenience; if not possible, an arrow indicating the *qibla* might be placed in an alcove. In Muhammad's time (when Muslims faced Jerusalem not Mecca), determining the *qibla* was relatively easy, though as Islam spread the curvature of the earth made determination more difficult. The religious impetus behind this problem motivated Islamic mathematics for centuries, and the methods developed (with minor modifications) are still used today, though in modern times the mathematics is very much hidden behind GPS technology!

Terminology and Trigonometric Tables Scholars worked with the Indian *half-chord* (sine), and with circles of various radii. Al-Battānī (c. 858–929) introduced an early version of *cosine* as the *complementary half-chord* for angles less than 90°, and an analogue of the modern function *versine*:[36]

$$\operatorname{versin}\theta = 1 - \cos\theta$$

Al-Bīrūnī (973–1048) defined versions of tangent, cotangent, secant and cosecant by projecting (e.g., a sundial) onto either a horizontal or a vertical plane. In the second picture below, the gnomon is the vertical stick of length 1. With this definition, al-Bīrūnī moves towards the modern consideration of trigonometry in terms of *triangles* rather than circles.





Trigonometric tables with improved accuracy over Ptolemy were created for all these 'functions.' Abū al-Wafā (940–998) and his descendants computed sine & tangent values for every *minute* of arc accurate to *five sexagesimal places* (one part in 777 million!) via repeated applications of the half-angle formula and interpolating using the downwards concavity of the sine function (draw a picture!):

$$\sin(\alpha+\beta) - \sin\alpha < \sin\alpha - \sin(\alpha-\beta) \qquad \text{whenever} \qquad 0° < \alpha-\beta < \alpha+\beta < 90°$$

---

[36] *Versed sine* refers to the measurement of a length in a re*versed* direction (perpendicular) to that of sine.

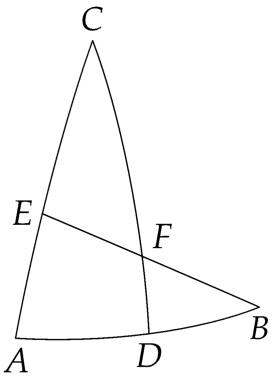
Calculating the *Qibla* In what follows we observe several conventions:

- A single letter A refers to a *point* or to the *angle measure* in a triangle with vertex A.
- $\overset{\frown}{AB}$ means the *great-circle arc* joining points A, B or its *arc-length.* A *spherical triangle* $\triangle ABC$ comprises three points on a sphere joined by great-circle arcs.
- $\overline{AB}$ means the *straight line* joining A, B with *length* $|AB|$.
- All results are modernized and applied to a unit sphere (center O). The arc-length along a great-circle therefore equals the central angle subtended by that arc in radians: $\overset{\frown}{AB} = \measuredangle AOB$. 3D movable versions of all pictures are online—click them!

Ptolemy and the Indians had already done some relevant work, though Ptolemy's approach relies heavily on Menelaus' Theorem (c. AD 100).

Theorem (Menelaus). *For the pictured configuration of spherical triangles on a sphere of radius 1,*

$$\frac{\sin \overset{\frown}{CE}}{\sin \overset{\frown}{AE}} = \frac{\sin \overset{\frown}{CF}}{\sin \overset{\frown}{DF}} \cdot \frac{\sin \overset{\frown}{BD}}{\sin \overset{\frown}{AB}}$$



Applying Menelaus is difficult since one typically needs to create many new spherical triangles. Al-Wafā simplified matters with an alternative result.

Theorem (Al-Wafā). *If $\triangle ABC$ and $\triangle ADE$ are spherical triangles with right angles at B, D and a common acute angle at A, then*

$$\frac{\sin \overset{\frown}{BC}}{\sin \overset{\frown}{AC}} = \frac{\sin \overset{\frown}{DE}}{\sin \overset{\frown}{AE}}$$

In fact these ratios equal $\sin \alpha$ where α is the acute angle, though al-Wafā didn't say this.

Proof. Let O be the center of the sphere. Project C orthogonally to the plane containing O, A, B to produce K, then project K to $\overline{OA}$ to get L.

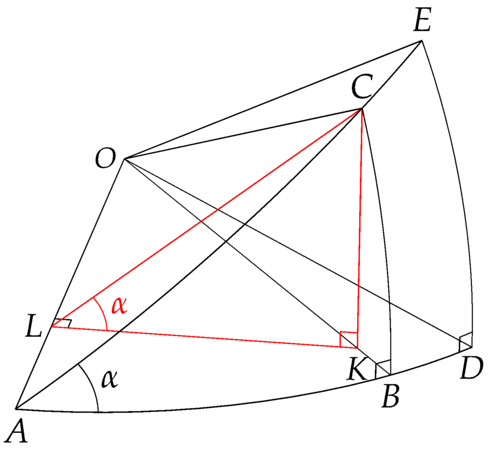
Consider the right-angled planar triangle CKL. Since α is the angle between two planes, we have $\alpha = \measuredangle CLK$. Moreover

$$|CK| = \sin \measuredangle COK = \sin \measuredangle COB = \sin \overset{\frown}{BC}$$
$$|CL| = \sin \measuredangle COL = \sin \measuredangle COA = \sin \overset{\frown}{AC}$$

The usual sine formula for plane triangles says

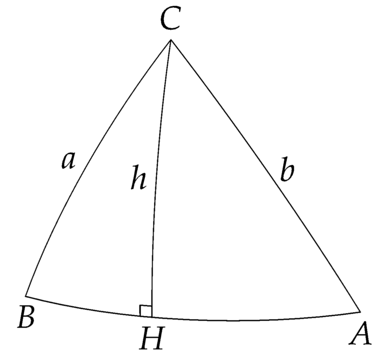
$$\sin \alpha = \frac{|CK|}{|CL|} = \frac{\sin \overset{\frown}{BC}}{\sin \overset{\frown}{AC}}$$

The same ratio is obtained for $\triangle ADE$.



$\blacksquare$

By dropping a perpendicular in a spherical triangle, Al-Wafā's result quickly yields the spherical sine rule. For the pictured triangle, drop the perpendicular to $H \in \widehat{AB}$ from C. Al-Wafā says

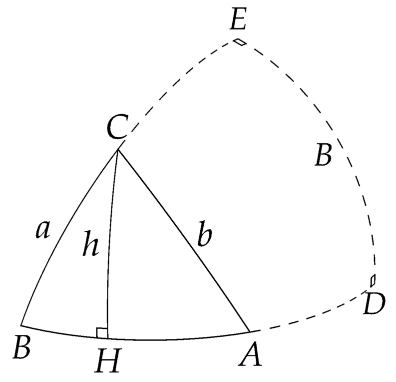


$$\sin B = \frac{\sin h}{\sin a} \quad \text{and} \quad \sin A = \frac{\sin h}{\sin b}$$

By equating the $\sin h$ terms and permuting symmetrically, we've proved:

Corollary (Sine rule). *If a, b, c are the side-lengths of a spherical triangle with angles A, B, C, then*

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

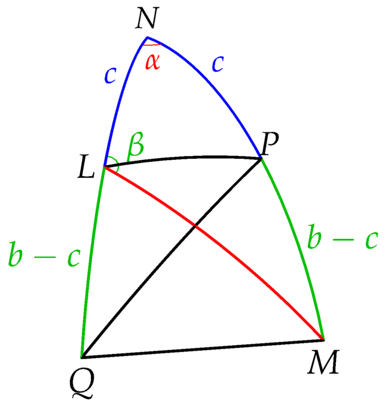


Al-Wafā's proof was a little more complicated. He extended $\widehat{AB}$ and $\widehat{BC}$ to quarter-circles resulting in a spherical triangle with right-angles at D and E. Since $\widehat{DE}$ is an arc with central angle B, we have $\widehat{DE} = B$. Since $\widehat{BD} = 90°$, Al-Wafā's theorem implies

$$\frac{\sin h}{\sin a} = \frac{\sin B}{\sin 90°} \implies \sin h = \sin a \sin B$$

Mirroring this by extending $\widehat{AB}$ past B and equating the $\sin h$ terms yields the result.

Using this approach, al-Wafā could solve spherical triangles and thus compute the *qibla*. As with his sine rule argument, his method required several auxiliary triangles and is difficult to follow.



Al-Bīrūnī further simplified matters by developing what is essentially the cosine rule. We apply his method to find the *qibla* from a location L (remember: our sphere (Earth!) has radius 1).

Let M be Mecca and N the north pole. The *qibla* is β, the initial bearing from L to M. Our (known) initial data are the latitudes and longitudes of L, M, specifically:

- α is the difference in the longitudes.
- b, c are the *colatitudes*[37] of M, L respectively.

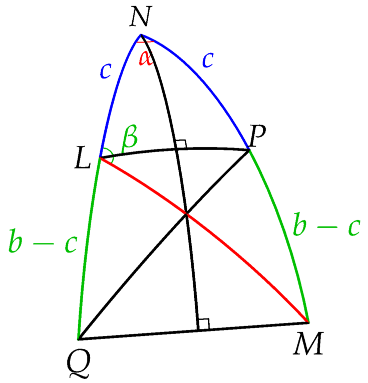
The cosine rule follows from Ptolemy's Theorem (pg. 44). Extend $\widehat{NL}$ to Q with the same latitude as M. Similarly let $P \in \widehat{NM}$ have the same latitude as L. By symmetry, L, P, Q, M are *coplanar,* whence the quadrilateral $\square LPQM$ lies on the intersection of a plane and a sphere: a circle! Measured as straight lines (chords) and using symmetry ($|PQ| = |LM|$ and $|LQ| = |PM|$), Ptolemy says

$$|LM|\,|PQ| = |LQ|\,|PM| + |LP|\,|QM| \implies |LM|^2 = |LQ|^2 + |LP|\,|QM|$$

---

[37] *Colatitude* (equals 90° minus latitude) is measured southwards from the north pole. Since our model sphere has radius 1, the arc-lengths b, c equal the colatitudes in radians.

The great-circle arc-lengths on the sphere may be found from straight-line distances via the usual chord relations: e.g.,



$$|LM| = \text{crd}\,\widehat{LM} = 2\sin\frac{\widehat{LM}}{2}$$

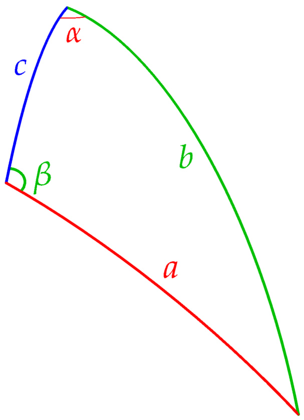
Ptolemy's theorem now becomes a relation between *arc-lengths*

$$\sin^2\frac{\widehat{LM}}{2} = \sin^2\frac{b-c}{2} + \sin\frac{\widehat{LP}}{2}\sin\frac{\widehat{QM}}{2}$$

By bisecting α we obtain two pairs of right-triangles; al-Wafā's theorem tells us that

$$\sin\frac{\alpha}{2} = \frac{\sin\frac{\widehat{LP}}{2}}{\sin c} = \frac{\sin\frac{\widehat{QM}}{2}}{\sin b}$$

$$\implies \sin^2\frac{\widehat{LM}}{2} = \sin^2\frac{b-c}{2} + \sin^2\frac{\alpha}{2}\sin c\sin b \tag{$*$}$$



To complete the proof we apply the multiple-angle formulæ

$$\sin^2\frac{x}{2} = \frac{1}{2}(1-\cos x) \qquad \cos(b-c) = \cos b\cos c + \sin b\sin c$$

Corollary (Cosine rule). *In a spherical triangle with sides a, b, c and angle α opposite a, we have*

$$\cos a = \cos b\cos c + \sin b\sin c\cos\alpha$$

For our triangle of interest $a = \widehat{LM}$. Given points L, M (and thus b, c, α), one uses the cosine rule to compute a and then the sine rule to find the *qibla* β (whew!):

$$\frac{\sin b}{\sin\beta} = \frac{\sin a}{\sin\alpha} \implies \sin\beta = \frac{\sin\alpha\sin b}{\sin a}$$

Example. For fun, here is some real-world data. Mecca and London have, respectively, co-ordinates 21°25′ N, 39°49′ E and 51°30′ N, 8′ W. This corresponds to

$$\alpha = 39°57', \qquad b = 68°35', \qquad c = 38°30'$$

By al-Bīrūnī's cosine rule,

$$\cos a = \cos 68°35' \cos 38°30' + \sin 68°35' \sin 38°30' \cos 39°57' \implies a = 43.110°$$

Since Earth's circumference is 24900 miles, the distance London → Mecca is $\frac{43.110\times 24900}{360} = 2981$ miles. Al-Wafā's sine rule computes the *qibla*

$$\beta = 180° - \sin^{-1}\frac{\sin\alpha\sin b}{\sin a} = 118°59'$$

where we subtracted from 180° since the relevant angle is plainly obtuse. Check it yourself at the Great Circle Mapper (the website uses airports for slightly different initial data).

Spherical Trigonometry Cheat Sheet

Let $\triangle ABC$ be a spherical triangle with side-lengths a, b, c on a sphere of radius 1.

Basic trigonometry. If $\triangle ABC$ is right-angled at C

$$\sin A = \frac{\sin a}{\sin c} \qquad \cos A = \frac{\tan b}{\tan c} \qquad \tan A = \frac{\tan a}{\sin b}$$

Al-Wafā essentially proved the first; the others follow from trig identities ($\cos^2 A = 1 - \sin^2 A \ldots$)

Sine rule (Al-Wafā)

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

Cosine rule (Al-Bīrūnī)

$$\cos c = \cos a \cos b + \sin a \sin b \cos C$$

The spherical Pythagorean Theorem is the special case $\cos c = \cos a \cos b$ $(C = 90°)$.

If the sphere has radius r, simply divide all lengths by r before applying the results; e.g.,

$$\sin A = \frac{\sin(a/r)}{\sin(c/r)}$$

As $r \to \infty$, we have $\sin \frac{a}{r} \approx \frac{a}{r}$ and $\cos \frac{a}{r} \approx 1 - \frac{a^2}{2r^2}$, which recover the flat (Euclidean geometry) versions of these statements.

Examples. 1. On a sphere of radius 1, an equilateral triangle has side length $\frac{\pi}{3}$. Splitting it in half creates two right-triangles with adjacent $\frac{\pi}{6}$ and hypotenuse $\frac{\pi}{3}$. The angles in the triangle are therefore

$$\alpha = \cos^{-1} \frac{\tan \frac{\pi}{6}}{\tan \frac{\pi}{3}} = \cos^{-1} \frac{1}{3} \approx 70.53°$$

The angle sum in the triangle is $3\alpha \approx 211.59°$!

2. An airfield is at C and two planes are at A and B. The bearings and distances to the aircraft are 45°, 2000 miles, and 90°, 4000 miles respectively. Find the distance between the aircraft.

This is just the cosine rule! We have a spherical triangle with sides 2000 and 4000 with angle 45° between them. If $r = 4000$ miles is Earth's radius, then

$$\begin{aligned} \cos \frac{c}{r} &= \cos \frac{2000}{r} \cos \frac{4000}{r} + \sin \frac{2000}{r} \sin \frac{4000}{r} \cos 45° \\ &= \cos \frac{1}{2} \cos 1 + \frac{1}{\sqrt{2}} \sin \frac{1}{2} \sin 1 \\ \Longrightarrow \; c &= 2833 \text{ miles} \end{aligned}$$

This is a little closer than the value (2947 miles) one would obtain from assuming a flat Earth!

Modern navigators typically use a slightly different, though equivalent, approach to minimize the error inherent in estimating cosine for small values: look up the *haversine formula* if you're interested.

Exercises 6.3. 1. A right-isosceles triangle on the surface of a unit sphere has equal legs of length $\frac{\pi}{4}$. Find the length of the hypotenuse and the sum of the angles in the triangle.

2. Explain the observation on page 62 that

$$0^\circ < \alpha - \beta < \alpha + \beta < 90^\circ \implies \sin(\alpha+\beta) - \sin\alpha < \sin\alpha - \sin(\alpha - \beta)$$

is the downwards concavity of the sine function.

3. Suppose we have a spherical triangle (sphere radius 1) as on page 65 with data

$$c = 30^\circ, \quad b = 60^\circ, \quad \alpha = 60^\circ$$

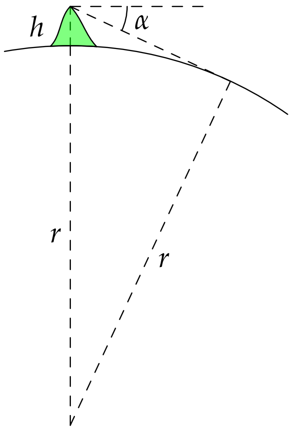
   (a) Use the cosine rule to find a.
   (b) Compute the remaining angles in the triangle. What do you observe about the sum of the angles $\alpha + \beta + \gamma$?

4. Determine the *qibla* for Rome (latitude 41°53′ N, longitude 12°30′ E).
   Repeat for the UCI campus (33°39′ N, 117°51′ W).

5. Al-Bīrūnī devised a method for determining the radius r of the earth by sighting the horizon from the top of a mountain of known height h. He would measure α, the angle of depression from the horizontal to which one sights the apparent horizon. Show that

$$r = \frac{h\cos\alpha}{1-\cos\alpha}$$

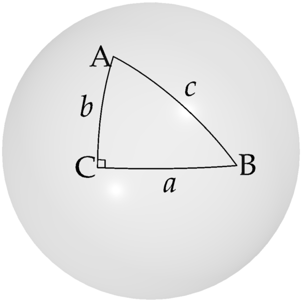
   In a particular case, al-Bīrūnī measures $\alpha = 34'$ from a mountain of height $652;3,18$ cubits. Assuming that a cubit equals $18''$, convert your answer to miles and compare with the modern value. Discuss the efficacy of this method.



6. On a sphere of radius r, Pythagoras' Theorem may be stated

$$\cos\frac{c}{r} = \cos\frac{a}{r}\cos\frac{b}{r} \qquad (*)$$

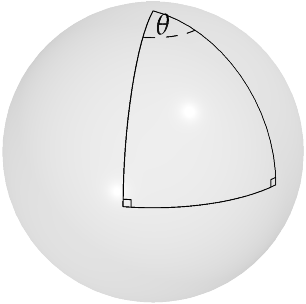
   where c is the hypotenuse and a, b the other side-lengths. Use the Maclaurin series $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$ to expand $(*)$ to degree 4.
   Suppose a, b are constant so that c is a function of r. Prove that $\lim_{r\to\infty} c^2 = a^2 + b^2$. Why does this make sense?



7. Construct a triangle on the surface of a sphere of radius r by taking two lines of longitude making an angle θ from the north pole to the equator. Prove that the area of the triangle is

$$A = r^2\theta$$

   What does Pythagoras $(*)$ say for this triangle?
   (*Hint: What fraction of the sphere is covered by the triangle?*)

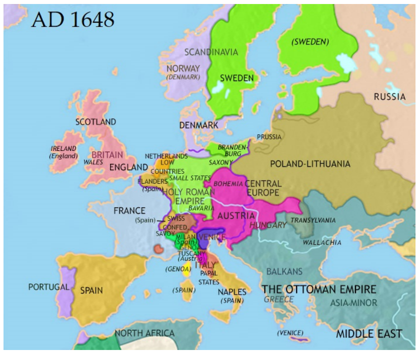


7 The Renaissance in Europe

7.1 Fibonacci and Notational Development

Between the fall of Rome in 476 and the early renaissance[38] (c. 1100) came Europe's dark ages. The one-dimensional view is that this was a time of limited learning and technological progress; as ever, the reality was more complex. Learning persisted in monasteries, though new researches took second place to a conservative focus on preserving the wisdom of the ancients.

By 1100, the small shifting kingdoms that had characterised Europe were starting to come together as more stable feudal states.[39] The maps below show the major feudal societies of medieval Europe (France, England, Holy Roman Empire, Poland, Austria) consolidating. While it would have meant little to the average peasant, a large stable nation is able to produce and support a larger elite population with the time and money to pursue education, and fund libraries and universities. By the 1700s, the borders of western Europe are largely recognizable; political and social organization had expanded so that a much larger proportion of the population—though small in comparison to today—were able to take advantage of and contribute to the growth of knowledge. The European renaissance is often contrasted with a decline in Islamic power, but again the story is more complex: Islam retreats from Spain, while the Ottoman Empire becomes dominant in south-east Europe.



Between 900 and 1300, the population of Europe roughly tripled to around 100 million. Trade increased, along with which came knowledge.[40] Learning (and land) came via wars with Islam (the Crusades, Spain, etc.). It helped that Islamic scholars so venerated the Greeks; Europeans could tell themselves that they were merely 'reclaiming' ancient knowledge which had been 'stolen' by their cultural and religious enemies. The fall of Constantinople to Mehmet in 1453 marks both the high point of Islamic conquest and the start of the decline of Islamic scientific dominance. Many intellectuals fled Constantinople—where the Byzantines still preserved much of Alexandria's learning—for Rome, helping to further fuel developments. With powerful enemies to the east, Europeans began travelling greater distances by sea,[41] beginning the colonial era of global European empire.

---

[38] Literally *rebirth*. Dates vary by location and discipline (Italy vs. France, art vs. philosophy, etc.) but a wide net would encompass the 12th to 17th centuries.

[39] An arrangement where powerful landowners could demand service (rent, food, warriors) from their tenants.

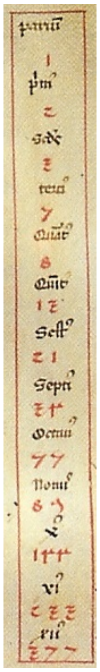
[40] A particularly important trading hub was Venice, from where Marco Polo (1254–1324), perhaps the most famous trader of the period, travelled the silk road to China.

[41] Christopher Columbus (born Genoa 1451) famously 'discovered' America in 1492 while looking for sea routes to Asia.

Scientific and philosophical progress was spurred by the translation of works from Arabic and ancient Greek into Latin, with the first universities being formed to teach this canon: Bologna 1088, Paris 1150, and Oxford 1167. The typical student was a young man of wealth who had been privately tutored in grammar, logic & rhetoric (the *trivium*). At university he would study the Greek-influenced *quadrivium* (geometry, astronomy, arithmetic & music). While Islamic improvements were incorporated, scholars gave pre-eminence to the Greeks: Euclid for geometry, Aristotle for logic/physics, Hippocrates/Galen for medicine, Ptolemy for astronomy. Early universities were often funded by the Church and 'research' was more likely to involve the justification of biblical passages using Aristotle than the conduct of experiments.

Leonardo da Pisa (Fibonacci c. 1175–1250) Fibonacci[42] likely first encountered the Hindu–Arabic numerals while trading with his father in North Africa. He was impressed by the ease of calculation they afforded and is the first European known to use them (contemporary Europeans used Roman numerals and Egyptian fractions). Fibonacci's 1202 text *Liber Abaci* was written to instruct traders in their use. The first page below explains how to compute with decimal fractions, with the two columns at the bottom of the page showing how to repeatedly multiply 100 (and then 10) by the fraction $\frac{9}{10}$. Thus:

$$100, \quad 90, \quad 81, \quad \frac{9}{10}72\ (=72.9), \quad \frac{1}{10}\frac{6}{10}65\ (=65.61), \quad \frac{9}{10}\frac{4}{10}\frac{0}{10}59\ (=59.049), \quad \text{etc.}$$



Note how the fractional part is written backwards on the left, using a bar to separate numerator and denominator; the Indians wrote fractions without a bar and it is thought Islamic scholars inserted it for clarity in the 1100s. The second picture is of Fibonacci's famous sequence: read top-to-bottom 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377. Amongst other inheritances from the Hindu–Arabic tradition, Fibonacci is the first known European to work with negative numbers, provided these represented deficiencies or debts in accounting.

---

[42] The name was given to Leonardo by French scholars of the 1800s: *filius Bonacci* means *son of Bonacci.*

Algebraic Notation and Development

To a modern reader, the most obvious mathematical development of the renaissance is notational. Fibonacci's fractional notation was cutting-edge for the 1200s but essentially everything else except numbers and fractions was written in sentences. The next 500 years slowly brought notational improvements which eventually allowed algebra to eclipse geometry as the primary language of reasoning. Here is a brief summary.

Italian Abacists, 14th C. This group continued Fibonacci's advocacy for the Hindu–Arabic system against the traditional use of Roman numerals, and also for the use of accompanying algorithms. Their approach was highly practical and largely for conducting trade. Here is a typical problem described by the group:

> The *lira* earns three *denarii* a month in interest. How much will sixty *lire* earn in eight months?

The Abacists introduced shorthands and symbols for unknowns and certain mathematical operations. *Cosa* ('thing') was used for an unknown. *Censo, cubo* and *radice* meant, respectively, square, cube and (square-)root. These expressions could be combined, for example 'ce cu' (read *censo di cubo*) referred to the sixth power of an unknown $(x^3)^2 = x^6$.

Luca Pacioli, Italy late 1400s. Introduced $\overline{p}, \overline{m}$ (*piu, meno*) for plus and minus. For example $8\overline{m}2$ denoted eight minus two.

Nicolas Chuquet, France 1484. His text *Triparty en la science des nombres* borrowed Pacioli's $\overline{p}$ and $\overline{m}$, and introduced an R-notation for roots. For instance $R^4 7$ meant $\sqrt[4]{7}$, while

$$\sqrt[5]{4-\sqrt{2}} \quad \text{would be written} \quad R^5\underline{4\overline{m}R2}$$

The underline indicates grouping (modern parentheses).

Christoff Rudolff, Vienna 1520s. Introduced symbols similar to x and ζ for an unknown and its square. He had other symbols for odd powers and produced tables showing how to multiply these. The words he used show Italian and French influence: algebra in the German-speaking world was known as the art of the *coss* (German for *thing*). Rudolff also introduced ± symbols as *algebraic operations,* though these had been used for around 30 years as a prefix denoting an excess or deficiency in a quantity (profit/loss in accounting). A period denoted equals and he is also credited with the first use of the square-root sign $\surd$, which is nothing more than a stylized r. This would be written *in front* of a number to denote a root, e.g. $\surd 23$ rather than $\sqrt{23}$.

Robert Recorde, England 1557. Introduced the equals sign in *The Whetstone of Witt,* asserting, 'No two things are more equal than a pair of parallel lines.'

Francois Viète, France 1540–1603. Before Viète, mathematicians typically described how to solve particular equations algorithmically via examples (e.g. $x^3 + 3x = 14$ rather than $x^3 + bx = c$), expecting readers to change numbers to fit a required situation. Viète pioneered the modern use of abstract constants, using letters to represent both unknowns and constants.

Simon Stevin, Holland 1548–1620. *De Thiende* (The Tenths) demonstrated how to calculate using decimals rather than fractions. Stevin arguably completed the journey whereby the concept of *number* subsumed that of *magnitude,* asserting that every ratio is a number. This increased the application of algebra by permitting the description of any geometric magnitude.

William Oughtred, England 1575–1660. Introduced × for multiplication, though he often simply used juxtaposition. Oughtred combined Viète's general approach (abstract constants) with symbolic algebra. For instance, to solve a quadratic equation $A_q + BA + C = 0$, where A_q means 'A-squared' (A-*quadratum* in Latin) and B, C are constants, he'd write the quadratic formula as

$$A = \surd : \frac{1}{4}B_q - C : -\frac{1}{2}B$$

In Oughtred's notation, colons were parentheses.

Thomas Harriot, England 1560–1621. Made several steps towards modern notation including juxtaposition for multiplication and a modern *encompassing* root-sign. For example,

$$\sqrt[4]{cccc + 27aa\sqrt[3]{2+b}} \quad \text{meant} \quad \sqrt[4]{c^4 + 27a^2\sqrt[3]{2+b}}$$

René Descartes, France/Holland 1596–1650. Used exponents for powers (a^2, a^3) and solidified the convention of using letters at the end of the alphabet (x, y, z) for unknowns and those at the beginning (a, b, c) for constants.

While modern mathematics uses many specialized symbols ($\emptyset$, $\Rightarrow$, e, π, etc.), basic notation is essentially unchanged from the mid 1600s. This is not to say that all mathematicians uniformly used the most modern notation: for instance, papers of Leonhard Euler (1700s) used Harriot's juxtaposition notation for exponents, and some published works from the late 1800s still wrote equations in words.

It is also worth mentioning in this context Gutenberg's 1439 invention of the printing press, which naturally aided the dissemination of all learning. The relative ease of production meant a great increase in the availability of written material, but also in the rejection/abandonment of some older texts when money could not be found to make an updated printed edition.

Exercises 7.1. 1. Repeatedly divide 10 by 5 a total of five times (to $\frac{10}{5^5}$), expressing the results using Fibonacci's notation.

2. What would Nicolas Chuquet have meant by the expression $4\overline{p}R^3\underline{7\,\overline{m}R5}$?

3. How would William Oughtred have expressed the solution to the quadratic equation $A_q = BA + C$? What about Thomas Harriot?

4. I am owed 3240 *florins.* The debtor pays me 1 *florin* the first day, 2 the second day, 3 the third say, etc. How many days does it take to pay off the debt?

5. (A problem of Antonio de Mazzinghi) Find two numbers such that multiplying one by the other makes 8 and the sum of their squares is 27.

   (*Hint: let the numbers be* $x \pm \sqrt{y}$)

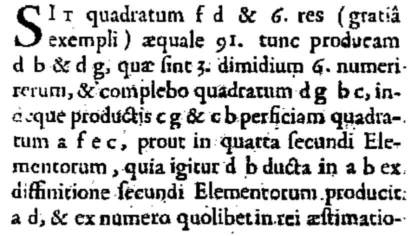
7.2 Polynomials: Cardano, Factorization & the Fundamental Theorem

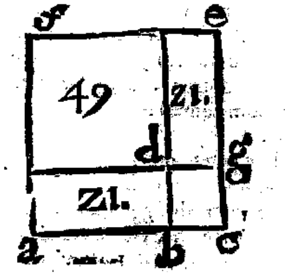
As an example of contemporary algebraic notation, we consider Girolamo Cardano's 1545 *Ars Magna* (*Great Art or the Rules of Algebra*), in which he describes how to solve quadratic and cubic equations.

The example on the right (start of *Caput V*, page 9 of the linked pdf), is the beginning of Cardano's description of how to solve $x^2 + 6x = 91$, *quadratum & 6 res aequalie 91* (square and 6 things equals 91), by completing the square. He employs several single-letter abbreviations but still writes in sentences and provides a pictorial justification. In modern algebra, the argument is nothing more than completing the square:

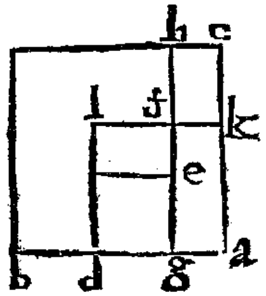
$$\begin{aligned} x^2 + 6x = 91 &\implies x^2 + 2 \cdot 3x + 3^2 = 91 + 3^2 \\ &\implies (x+3)^2 = 100 \\ &\implies x + 3 = 10 \\ &\implies x = 7 \end{aligned}$$

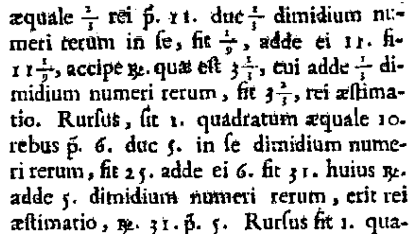
where the picture justifies $7^2 + 2 \cdot 21 + 3^2 = 10^2$.


SIT quadratum f d & 6. res (gratiâ exempli) æquale 91. tunc producam d b & d g, quæ ſint 3. dimidium 6. numeri rerum, & complebo quadratum d g b c, indeque productis c g & c b perficiam quadratum a f e c, prout in quarta ſecundi Elementorum, quia igitur d b ducta in a b ex diffinitione ſecundi Elementorum producit a d, & ex numero quolibet in rei æſtimatio-



The quadratic algorithms were well-established by this time; on the following page of his text, we find a picture from the work of al-Khwārizmī (Exercise 6.2.2), after which comes more obvious mathematical notation. Even though Rudolff's ± and $\surd$ were in use, Cardano wrote almost everything in words augmented with fractional notation and Pacioli's $\overline{p}$ and $\overline{m}$.




æquale $\frac{2}{3}$ rei p̄. 11. duc $\frac{1}{3}$ dimidium numeri rerum in ſe, fit $\frac{1}{9}$, adde ei 11. fit 11$\frac{1}{9}$, accipe ℞. quæ eſt 3$\frac{1}{3}$, cui adde $\frac{1}{3}$ dimidium numeri rerum, fit 3$\frac{2}{3}$, rei æſtimatio. Rurſus, ſit 1. quadratum æquale 10. rebus p̄. 6. duc 5. in ſe dimidium numeri rerum, fit 25. adde ei 6. fit 31. huius ℞. adde 5. dimidium numeri rerum, erit rei æſtimatio, ℞. 31. p̄. 5. Rurſus ſit 1. qua-

As was typical for the time, Cardano describes negative solutions as fictitious; he even writes the square root of −15 at one point, though only to mention its absurdity. He also follows the Islamic approach of solving a concrete problem of each type rather than proceeding abstractly, though observe the *gratia exempli* ("for the sake of an example") as acknowledgement that the general problem $x^2 + ax = b$ may be solved identically.

It is for the solution of cubic (and quartic) equations that Cardano is most famous. Below we describe, in modern notation, Cardano's method for solving the cubic equation $x^3 + bx = c$ where $b, c > 0$, though we stress (again) that Cardano only gave *examples* not a general formula.

Let u, v satisfy $u^3 - v^3 = c$ and $uv = \frac{b}{3}$. Then $x = u - v$ is seen to solve the cubic.

$$\begin{aligned} x^3 + bx &= (u-v)^3 + b(u-v) = u^3 - 3u^2v + 3uv^2 - v^3 + b(u-v) \\ &= (u^3 - v^3) + (u-v)(b - 3uv) = c \end{aligned}$$

However u and v also satisfy

$$(u^3 + v^3)^2 = (u^3 - v^3)^2 + 4(uv)^3 = c^2 + 4\left(\frac{b}{3}\right)^3$$

so we obtain a system of linear equations in the unknowns u^3, v^3 which are easily solved:

$$\begin{cases} u^3 + v^3 = \sqrt{c^2 + 4\left(\frac{b}{3}\right)^3} \\ u^3 - v^3 = c \end{cases} \implies \begin{cases} u = \sqrt[3]{\sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3} + \frac{c}{2}} \\ v = \sqrt[3]{\sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3} - \frac{c}{2}} \end{cases}$$

$$\implies x = u - v = \sqrt[3]{\sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3} + \frac{c}{2}} - \sqrt[3]{\sqrt{\left(\frac{c}{2}\right)^2 + \left(\frac{b}{3}\right)^3} - \frac{c}{2}}$$

Examples 1. We apply Cardano's method to $x^3 + 6x = 7$, which has the obvious solution $x = 1$.

$$\begin{cases} u^3 - v^3 = 7 \\ uv = \frac{6}{3} = 2 \end{cases} \implies (u^3 + v^3)^2 = 7^2 + 4 \cdot 2^3 = 81 \implies u^3 + v^3 = 9$$

$$\implies u^3 = \frac{1}{2}(9 + 7) = 8, \qquad v^3 = \frac{1}{2}(9 - 7) = 1$$

$$\implies x = u - v = \sqrt[3]{8} - \sqrt[3]{1} = 2 - 1 = 1$$

2. This time we solve $x^3 + 3x = 14$.

$$\begin{cases} u^3 - v^3 = 14 \\ uv = \frac{3}{3} = 1 \end{cases} \implies (u^3 + v^3)^2 = 14^2 + 4 \cdot 1^3 = 200 \implies u^3 + v^3 = 10\sqrt{2}$$

$$\implies \begin{cases} u^3 = \frac{1}{2}(10\sqrt{2} + 14) = 5\sqrt{2} + 7 = (\sqrt{2} + 1)^3 \\ v^3 = 5\sqrt{2} - 7 = (\sqrt{2} - 1)^3 \end{cases}$$

$$\implies x = u - v = (\sqrt{2} + 1) - (\sqrt{2} - 1) = 2$$

As this shows, Cardano's formula might produce an ugly expression for a simple answer—of course, the cube root of $5\sqrt{2} \pm 7$ is the tip of everyone's tongue!

3. The equation $x^3 + 3x = 10$ may be solved using Cardano's method, though in this case the answer is just ugly.

$$\begin{cases} u^3 - v^3 = 10 \\ uv = \frac{3}{3} = 1 \end{cases} \implies (u^3 + v^3)^2 = 10^2 + 4 \cdot 1^3 = 104 \implies u^3 + v^3 = 2\sqrt{26}$$

$$\implies u^3 = \sqrt{26} + 5, \qquad v^3 = \sqrt{26} - 5$$

$$\implies x = u - v = \sqrt[3]{\sqrt{26} + 5} - \sqrt[3]{\sqrt{26} - 5} \approx 1.6989$$

Being unable or unwilling to work directly with negative numbers, Cardano modified his method to solve other cubics such as $x^3 + c = bx$, and moreover described how to remove a quadratic term from a cubic using what is now known as the *Tschirnhaus substitution* ($x = y - \frac{a}{3}$):

$$x^3 + ax^2 + bx + c = \left(y - \frac{a}{3}\right)^3 + a\left(y - \frac{a}{3}\right)^2 + b\left(y - \frac{a}{3}\right) + c = y^3 - \frac{a^2}{3}y + \cdots \tag{$*$}$$

Cardano's student, Lodovico Ferrari, extended the method to solve quartic equations in terms of the solution of a resultant cubic.

Negative Solutions and Complex Numbers

By the late 1500s, mathematicians were mentioning negative solutions to equations—these were usually described as *fictitious,* or *false roots,* but this didn't stop them from being investigated. Rafael Bombelli (1526–1572, Rome) even introduced a notation for complex numbers, described their algebra, and showed how they could be used to find solutions to any quadratic or cubic equation. For instance, he wrote $4+3i$ and $4-3i$ as follows:

> 4 *p di m* 3, read '*quattro piu di meno tre*' (four plus of minus three), and,
> 4 *m di m* 3, '*quattro meno di meno tre.*'

Given Bombelli's belief in the fictitiousness of complex numbers, the effort he expended in their honor is extraordinary: this is a prime example of pure abstraction, math for the sake of math!

In modern language, the three roots of Cardano's cubic $x^3+bx=c$ are

$$u-v, \quad \zeta u-\zeta^2 v, \quad \zeta^2 u-\zeta v$$

where $\zeta = e^{2\pi i/3} = \frac{-1+\sqrt{3}i}{2}$ is a primitive cube root of unity. Together with the Tschirnhaus substitution $(*)$, Cardano's formula therefore solves all cubic equations.

Examples 1. Returning to one of our previous examples, if $x^3+3x=14$, then $u=\sqrt{2}+1$ and $v=\sqrt{2}-1$, from which the three solutions are

$$\begin{aligned} &u-v=2 \\ &\zeta u-\zeta^2 v = (\sqrt{2}+1)\frac{-1+\sqrt{3}i}{2} - (\sqrt{2}-1)\frac{-1-\sqrt{3}i}{2} = -1+\sqrt{6}i \\ &\zeta^2 u-\zeta v = (\sqrt{2}+1)\frac{-1-\sqrt{3}i}{2} - (\sqrt{2}-1)\frac{-1+\sqrt{3}i}{2} = -1-\sqrt{6}i \end{aligned}$$

2. To find a solution to $x^3+3x^2=3$, we perform the Tschirnhaus substitution $x=y-\frac{3}{3}=y-1$ before applying Cardano's method:

$$\begin{aligned} &y^3-3y^2+3y-1+3(y^2-2y+1)=3 \implies y^3-3y=1 \qquad (b=-3,\ c=1) \\ &\implies \begin{cases} u^3-v^3=1 \\ uv=-\frac{3}{3}=-1 \end{cases} \implies (u^3+v^3)^2 = 1^2+4(-1)^3=-3 \\ &\implies u^3+v^3=\sqrt{3}i \implies u^3=\frac{1}{2}(\sqrt{3}i+1), \quad v^3=\frac{1}{2}(\sqrt{3}i-1) \\ &\implies x=u-v-1=\sqrt[3]{\frac{\sqrt{3}i+1}{2}}-\sqrt[3]{\frac{\sqrt{3}i-1}{2}}-1 \end{aligned}$$

This is ugly! If you know Euler's formula and you choose a compatible pair of cube roots (you need $uv=-1$), this will evaluate to one of three real numbers: the positive solution is in fact $x=2\cos 20^\circ - 1 \approx 0.8794$.

Factorization & the Fundamental Theorem of Algebra

By the late 1500s, Viète's abstraction allowed him to streamline Cardano's methods. He also investigated the relationship between the coefficients of a polynomial and its roots. For Viète the roots had to be positive, but later improvements by Thomas Harriot and Albert Girard (1629) applied this to all polynomials with any roots. For instance, if $ax^2+bx+c=0$ has roots r_1, r_2, then

$$(r_1-r_2)\big(a(r_1+r_2)+b\big) = a(r_1^2-r_2^2)+b(r_1-r_2) = \big(ar_1^2+br_1+c\big)-\big(ar_2^2+br_2+c\big) = 0$$

Provided the roots are distinct,[43] we conclude that

$$\frac{b}{a} = -(r_1+r_2), \quad \frac{c}{a} = -r_1^2-\frac{b}{a}r_1 = -r_1^2+(r_1+r_2)r_1 = r_1r_2$$

These are the quadratic version of what are known as Viète's formulas; other versions exist for every degree. Their use amounts to an early form of factorization.

Example To solve $3x^2-2x-1=0$, first spot that $r_1=1$ is a root. By Viète's formulas,

$$r_1+r_2 = -\frac{b}{a} = \frac{2}{3} \implies r_2 = -\frac{1}{3} \quad \text{or alternatively} \quad r_1r_2 = \frac{c}{a} = -\frac{1}{3} \implies r_2 = -\frac{1}{3}$$

Think about the relationship between this approach and factorization!

A nice side-effect is a method for obtaining the quadratic formula by solving a pair of simultaneous equations analogous to Cardano's cubic approach:

$$\begin{cases} r_1+r_2 = -\frac{b}{a} \\ r_1-r_2 = \sqrt{(r_1+r_2)^2-4r_1r_2} = \sqrt{\left(\frac{b}{a}\right)^2-4\frac{c}{a}} \end{cases} \implies r_1, r_2 = -\frac{b}{2a} \pm \frac{1}{2}\sqrt{\left(\frac{b}{a}\right)^2-4\frac{c}{a}}$$

Viète's formulas were central to the later development of Galois Theory (1830) and the Abel–Ruffini theorem regarding the insolubility of quintic and higher-degree polynomials.

The remaining key results regarding solutions of polynomials also appeared around this time:

Fundamental Theorem of Algebra Including multiplicity, a degree n polynomial has n complex roots. Girard offered the first version in 1629, though a complete proof didn't appear until the work of Argand, Cauchy and Gauss in the early 1800s.

Factor Theorem In 1637, Descartes proved that $p(r)=0 \iff p(x)$ is divisible by $x-r$, essentially via long-division.

For instance, to find the roots of $p(x)=x^3+2x^2-13x+10$, Descartes would observe:

- $p(1)=0 \implies x-1$ is a factor; use long-division to obtain $p(x)=(x-1)(x^2+3x-10)$.
- $q(x)=x^2+3x-10$ has $q(2)=0$; divide to get $q(x)=(x-2)(x+5)$.
- The roots of $p(x)$ are therefore 1, 2 and -5 (Descartes called this last a *false root*).

We'll return to Descartes in the context of Analytic Geometry in the next chapter.

---

[43] In fact the formulas hold even when $r_1=r_2$.

Exercises 7.2. 1. (a) Find Viète's formulas for the polynomial $p(x) = x^3 + ax^2 + bx + c$ with roots r_1, r_2, r_3; that is, find the coefficients a, b, c in terms of the roots.

   (*Hint: Multiply out* $p(x) = (x - r_1)(x - r_2)(x - r_3) \ldots$)

   (b) Solve $x^3 - 6x^2 + 9x - 4 = 0$ using Girard's method: first, determine one solution by inspection, then use Viète's formulas for the cubic to investigate the relationship between the remaining roots.

2. (a) Apply Cardano's method to the equation $x^3 + 6x = 20$.

   (*Hint: to finish, compute* $(1 + \sqrt{3})^3$)

   (b) If $b, c > 0$, Cardano's method finds a single positive solution to $x^3 + bx = c$. Explain why such an equation always has exactly one real solution which is moreover positive.

3. Prove that if t is a root of $x^3 = cx + d$, then

$$r_1 = \frac{t}{2} + \sqrt{c - \frac{3t^2}{4}} \quad \text{and} \quad r_2 = \frac{t}{2} - \sqrt{c - \frac{3t^2}{4}}$$

   are both roots of $x^3 + d = cx$. Use this to solve $x^3 + 3 = 8x$.

4. Consider the cubic equation $aaa - 3raa + ppa = 2xxx$ (as written by Harriot). Show that the substitution $a = e + r$ reduces this to an equation without a square term.

   As an example, reduce the equation $aaa - 18aa + 87a = 110$ to a cubic in e without a square term. Find all three solutions in e and therefore find the solutions to the original equation in a.

5. Find all roots of the cubic $p(x) = 2x^3 - 3x^2 - 3x + 2$ using Descartes' factor theorem.

6. Use Cardano's method (with Tschirnhaus substitution and complex numbers!) to find the solutions to the equation

$$x^3 + 4 = 3x^2$$

   Verify that you get the same solutions using Descartes' factor theorem.

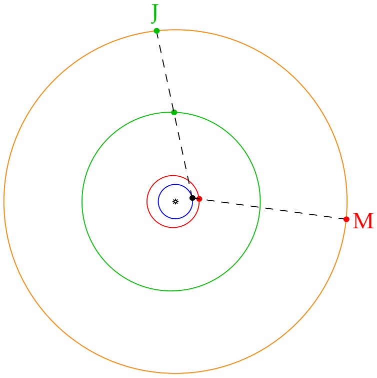
   (*Hint: all solutions are integers!*)

7. (If you are very comfortable with complex numbers) Use Euler's formula to verify that the equation $x^3 + 3x^2 = 3$ has the positive solution $x = 2\cos 20° - 1$. It also has two negative real solutions: find them!

7.3 Astronomy and Trigonometry in the Renaissance

Distinctly European progress in trigonometry came courtesy of Johannes Müller (a.k.a. Regiomontanus,[44] 1436–1476). *De Triangulis Omnimodius* (*Of all kinds of triangles,* 1463) provided an axiomatic update of Ptolemy's *Almagest* and its Islamic improvements. Though the title refers to triangles, his approach remains circle-based (chords and half-chords). Regiomontanus was a renowned astronomer; his tracking of a comet from late 1471 to spring 1472 provided controversial evidence that objects could move between the, supposedly fixed, heavenly spheres of ancient Greek theory.

Domenico Novara (1454–1504), Regiomontanus' student, inherited much of his unpublished work. He became a student of Pacioli in Florence and an astronomer at the University of Bologna, though he is now perhaps best known as adviser to a young Pole, Nicolaus Copernicus (1473–1543), who studied in Bologna from 1496 with the ostensible intent of joining the priesthood…



Copernicus concluded that Ptolemy's geocentric model could not be reconciled with astronomical observation. *De revolutionibus orbium celestium* (*On the revolutions of the heavenly spheres*), published a year after his death, describes how to compute within a *heliocentric* (*sun-centered*) model. This, Copernicus believed, was the obvious solution to the problem of *retrograde motion* that had plagued the ancient Greeks.[45]

The animation demonstrates Copernicus' solution: with the sun at the center, the retrograde motion of Mars and Jupiter are easily explained. The outer circle represents the 'fixed stars' against which the motion of the planets are observed.

Copernicus' work is now described as a revolution, though it was not perceived so at the time. *De revolutionibus* was dedicated to the Pope, welcomed by the Church, and used by Vatican astronomers to aid in calculation. The difficulty and narrow readership of his work made it unthreatening to contemporary dogma. Copernicus did not present heliocentrism as reality nor did he advocate for overturning long-held beliefs. Within a century, however, the Copernican theory had found its bulldog in Galileo, and conflict between science and the Church became unavoidable.

Trigonometry is finally about triangles!

Georg Rheticus (1514–1574) defined trigonometric functions purely in terms of triangles, referring to the *perpendiculum* (sine) and *basis* (cosine) of a right-triangle with fixed hypotenuse. Rheticus was a student of Copernicus and helped posthumously to publish his work.

In 1595, Bartholomew Pitiscus finally introduced the modern term with his book *Trigonometriæ*, in which he purposefully sets out to solve problems related to triangles. The picture is the title page from the second edition (1600—MDC in Roman numerals). Both Rheticus and Pitiscus had problems which look very familiar to modern readers, such as solving for unknown sides of triangles.

Bartholomæi
Pitiſci Grunbergenſis
Sileſij.
TRIGONOMETRIÆ
Siue
De dimenſione Trianguloꝝ
LIBRI QVINQVE. Item
PROBLEMATVM VARIORV̄
nempe
Geodæticorum,
Altimetricorum,
Geographicorum,
Gnomonicorum, et
Aſtronomicorum:
LIBRI DECEM.
TRIGONOMETRIÆ SVBIVNCTI,
AD VSVM EIVS DEMON-
STRANDVM.

Sancti Vrbani Bibliothecæ,

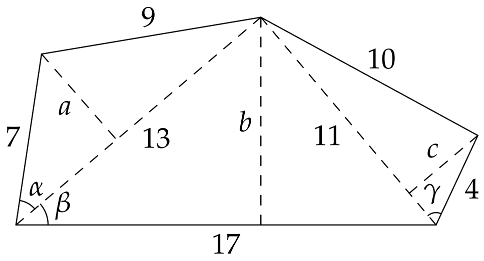
Auguſtæ Vindelic.

Sumptib. et impens.
Dominici. Cuſtodis
chalcographi.
M.DC..

---

[44] His grand-sounding name is a latinization of his birthplace Königsberg (*King's Mountain*), Bavaria, Germany.

[45] Copernicus wasn't the first to suggest such a model. Several ancient Greek scholars embraced heliocentrism, with Aristarchus of Samos (c. 310–230 BC) credited with its first presentation. However, Aristarchus' views were rejected by the mainstream Greeks, and it is likely Copernicus never encountered his work.



Example (Pitiscus) A field has five straight edges of lengths 7, 9, 10, 4 and 17 in order. The distance from the first to third vertex is 13 and from the third to fifth is 11. What is the area of the field?

The problem can be solved easily using Heron's formula, but Pitiscus opts for trigonometry. We give a modernized version that depends on applying the law of cosines to the three large triangles.

$$\cos\alpha = \frac{7^2+13^2-9^2}{2\cdot 7\cdot 13} = \frac{137}{182} \qquad \cos\beta = \frac{17^2+13^2-11^2}{2\cdot 17\cdot 13} = \frac{337}{442}$$
$$\cos\gamma = \frac{4^2+11^2-10^2}{2\cdot 4\cdot 11} = \frac{37}{88}$$

The values of α,β,γ and therefore the altitudes a,b,c of the three major triangles could be read off a table, or found exactly using Pythagoras':

$$a = 7\sin\alpha = 7\sqrt{1-\cos^2\alpha} = \frac{7}{182}\sqrt{182^2-137^2} = \frac{3}{26}\sqrt{1595},$$
$$b = 13\sin\beta = \frac{1}{34}\sqrt{81795}, \qquad c = 4\sin\gamma = \frac{5}{22}\sqrt{255}$$

The total area is easily computed:

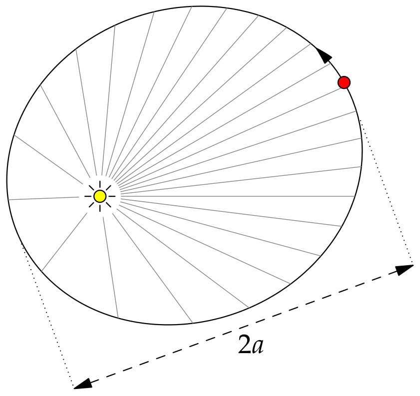
$$A = \frac{1}{2}(13a+17b+11c) = \frac{1}{4}\left(3\sqrt{1595}+\sqrt{81795}+5\sqrt{255}\right) \approx 121.4$$

Kepler's Laws

Johannes Kepler (1571–1630) was a student of Tycho Brahe[46] for the last two years of Brahe's life. He inherited Brahe's position, decades-worth of astronomical data, and his philosophy on the importance of observation-based theory. Kepler also embraced the mystical Pythagorean view that nature reflects harmony, a belief that partly drove his scientific pursuits. To Kepler, any observation of a natural, simple ratio was something of great import. For example, in observing that the daily movement of Saturn at its furthest point from the sun was roughly 4/5 of that at its nearest point, his temptation was to assume that 'roughly' must be 'exactly.'

Thanks to Brahe, Kepler had data on roughly thirteen orbits of Mars and two of Jupiter. From these data, he posited three laws.

1. Planets move in ellipses with the sun at one focus.
2. The orbital radius sweeps out equal areas in equal times. In the picture, the sectors all have the same area and the planet moves more slowly the further it is from the sun.
3. The square of the orbital period is proportional to the cube of the semi-major axis of the ellipse: $T^2 \propto a^3$.

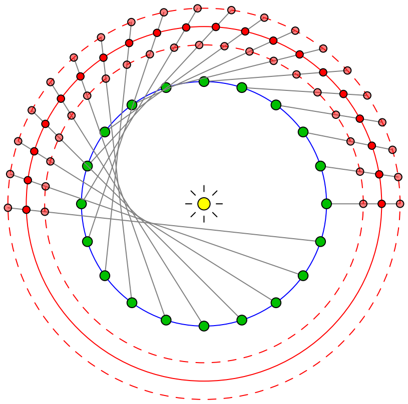


---

[46] Tycho Brahe (1546–1601) was a Danish astronomer who worked for Austro-Hungarian Emperor Rudolph II in Prague for 25 years, producing a wealth of accurate astronomical measurements. While these helped burnish the Copernican theory, he is also known for his 1572 observation of a *nova* (a *new* star that later disappeared, now understood to be the death of a distant star), and then a comet in 1577; both provided yet more evidence for the changeability of the heavens.

Kepler's laws are empirical observations rather than the result of mathematical proof. However, the process of their discovery demonstrates Kepler's tremendous mathematical abilities.

Starting point Kepler began by assuming the essential correctness of the Copernican model in that all planets exhibit uniform circular motion round the sun.



Orbital Estimation Kepler's data told him the *direction* to each planet, but not the *distance,* though his Copernican assumption allowed him to estimate the *relative distance*.[47] For example, using the direction from Earth to Mars at equally spaced times and by drawing circles of different radii for possible orbits of Mars, he was able estimate an orbit where Mars' motion would also be uniform. This required an enormous number of trigonometric calculations: each measurement of planetary longitude/latitude *relative to Prague* had to be converted to measurements relative to the sun. Everything was subject to errors of estimation.

Modifying the model Kepler altered his model to reflect Earth's slightly non-circular orbit. He first tried an equant model (in the style of the ancient Greeks), offsetting the center of the orbit slightly from the sun. Despite this, he failed to fit his data for Mars to pure circular motion.

The First Law Abandoning circles, Kepler now permitted planets to move in ovals. He decided to approximate Mars' orbit with an ellipse and set out to calculate its parameters, stumbling on an almost perfect match when the sun was placed at a focus. The *geometric* significance of the focus was exactly the natural beauty sought by Kepler. Having now established the first law for Mars, he repeated the exercise for the other known planets (Mercury, Venus, Earth, Jupiter, Saturn) as well as he could given his inferior data.

The Second Law Kepler's second law followed an infinitesimal argument based on inspired guesswork. To fit his elliptical orbits, planetary velocity was non-constant, appearing inversely proportional to the distance from the sun ($v = \frac{k}{r}$). Kepler used this to approximate the area of a sector swept out by the radius vector: over a small time-interval Δt, a planet travels a distance $v\Delta t$ and thus sweeps out an approximate triangle of area $\frac{1}{2}rv\Delta t = \frac{1}{2}k\Delta t$. In modern language, this is the conservation of angular momentum. Kepler had no justification beyond the fact that it seemed to fit the data. In particular, he did not know why a planet should move more slowly when further from the sun.

The Third Law This was stated with very limited analysis. Given the relative sizes of each orbit and its period, some inspired guesswork allowed him to observe that $\frac{T^2}{a^3}$ is approximately the same value for each.

Kepler's discoveries were revealed over many years in several texts, and his magnum opus *Epitome astronomiæ Copernicanæ* was published in 1621. Within a century Issac Newton had provided a mathematical justification of Kepler's laws based on the theory of calculus and his own axioms: an inverse square law for gravitational acceleration and his own three laws of motion.

---

[47] Following Ptolemy, Brahe thought the Earth-Sun distance was around 1/10 of its true value. Kepler thought this an underestimate by at least a factor of three. In 1659, Christiaan Huygens found the distance to an accuracy of 3%.

A Religious Interlude: Protestantism, the Counter-Reformation and Calendar Reform

In 1563, Pope Gregory began the Catholic Church's push-back against the spread of Protestantism,[48] the *counter-reformation.* Of particular interest to science and mathematics was the newly created *Index Librorum Prohibitorum,* a list of books contradicting Church doctrine. This was also a response to the new technology of mass-printing, which made disseminating controversial new ideas easier. Heliocentrism came quickly under attack: Kepler's book was banned immediately upon publication in 1621. However, his location far from Rome meant that Kepler and his ideas were relatively safe. The ultimate result of Gregory's crackdown was the slow ceding of scientific power to northern (Protestant) Europe where papal diktat had no effect.

In contrast to the anti-scientific book-banning fervor of the counter-reformation, Pope Gregory is also famous for shepherding an astounding scientific achievement: calendar reform. By 1500, astronomers knew the solar year to be roughly $11\frac{1}{4}$ minutes shorter than the $365\frac{1}{4}$ days of the Julian calendar.[49] For 1200 years, Easter had been decreed to be the Sunday after the first full moon after the vernal equinox (March 20th/21st), but by 1500 the equinox was happening 10–11 days earlier. The impetus to correct the date of Easter meant that calendar reform and astronomical modelling were now an important Church project.

A century of effort[50] resulted in the *Gregorian calendar* (designed by Aloysius Lilius and Christopher Clavius). Gregory imposed the new calendar in all Catholic countries in 1582. The 10 day deficit was corrected by eliminating October 5th–14th 1582. To prevent the error re-occurring, the computation of leap-years was also changed: centuries are now leap-years only if divisible by 400, thus 1600 was a leap year, but 1900 was not. The Gregorian calendar is astonishingly accurate, losing only one day every 3000 years. Since it emanated from Rome, many Protestant parts of Europe took decades if not centuries to adopt the new calendar. The Eastern Orthodox Church still computes Easter using the Julian calendar, which is now 13 days behind the Gregorian.

Galileo Galilei (1564–1642)

Based in northern Italy, Galileo was close to the center of Church power; unlike Copernicus and Kepler, he openly challenged its orthodoxy. While undoubtedly a great mathematician, he is more importantly considered the father of the scientific revolution for his reliance on experiment and observation. He famously observed Jupiter's moons with a telescope of his own invention, noting that objects orbiting an alien body was counter to Ptolemaic theory. Skeptics, when shown this image, preferred to assert that it must be somewhere *inside* the telescope!

In 1632 Galileo published *Dialogue Concerning the Two Chief World Systems,* a Socratic discussion between three characters: Salviati argued for Copernicus, Simplicio was for Ptolemy, and Sagredo was an independent questioner. The character of Simplicio was provocatively modeled on conservative philosophers who refused to consider experiments and bore a notable resemblance to the Pope. Salviati almost always came out on top and Simplicio was made to appear foolish. The text resulted in Galileo's conviction for heresy; all his publications, past and future, were banned, and he spent the remainder of his life under house arrest.

---

[48] Martin Luther's *Ninety-five Theses* (1517) is generally considered the start of the Protestant Reformation. Europe saw several gruesome religious wars over the next 150 years as various countries broke from Catholicism and Rome.

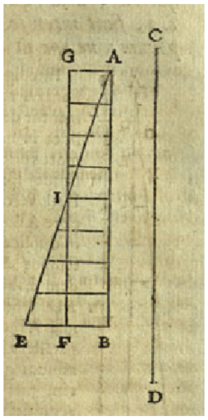
[49] Named for Julius Caesar, the Julian year has 365 days with a leap-day added every four years.

[50] Pope Sixtus IV tried to recruit Regiomontanus to the cause in 1475, though the mathematician died first. Copernicus was among those invited to consider proposals in the early 1500s, though he distanced himself, perhaps because he knew that his developing heliocentric ideas would not be accepted.

Despite Church efforts, Galileo's works continued to be distributed by his supporters and he continued working. His most important scientific text, *Discourses Concerning Two New Sciences* (materials science and kinematics) was smuggled out of Italy to be published in Holland in 1638. In this book he resurrects his characters from *Two Chief World Systems* and famously refutes Aristotle's claim that heavier objects fall more rapidly than lighter ones.[51] Here are two results from this text.

Theorem. *If acceleration is uniform, then the average speed is the average of the initial and terminal speeds.*

Proof. Galileo argues pictorially.



In essence, $\overline{CD}$ is the time-axis, increasing downward. Velocity is measured horizontally from the time-axis to the uniformly sloped line $\overline{AE}$, of which I is the midpoint.

The distance travelled is the area between $\overline{AE}$ and the time-axis, which plainly equals the area of the rectangle between $\overline{GF}$ and the time-axis. The velocity corresponding to $\overline{GF}$ is the average of those corresponding to A and E.

Here is the calculation using modern algebra. Suppose the object has velocity v_A as it passes A and v_B as it passes v_B, and that $t = |AB|$. Then the distance travelled is

$$v_{\text{av}}t = v_At + \text{area}(\triangle ABE) = v_At + \frac{1}{2}(v_B - v_A)t = \frac{1}{2}(v_A + v_B)t$$

whence v_{av} is the average of the initial and final speeds.

■

Corollary. *A falling object dropped from rest will traverse distance in proportion to time-squared,*

$$d_1 : d_2 = t_1^2 : t_2^2$$

This is Galileo's version of the well-known kinematics formula $d = \frac{1}{2}gt^2$.

Proof. Let d_1, d_2, v_1, v_2 represent the distances travelled and the speeds of the dropped body at times t_1 and t_2. Since acceleration is uniform,

$$v_1 : v_2 = t_1 : t_2$$

By the Theorem,

$$d_1 = \frac{0 + v_1}{2}t_1 = \frac{1}{2}v_1t_1, \quad \text{and} \quad d_2 = \frac{1}{2}v_2t_2$$

whence

$$d_1 : d_2 = v_1t_1 : v_2t_2 = t_1^2 : t_2^2$$

■

---

[51] Supposedly by dropping weights off the leaning tower of Pisa, though take this story with a pinch of salt...

Galileo follows this by decomposing the motion of a projectile into horizontal (uniform speed) and vertical (uniform acceleration) components, thereby proving that projectiles follow parabolic paths.

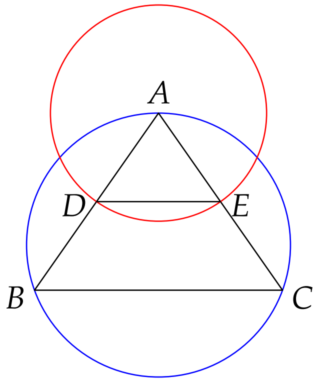
Galileo covered several other important mathematical topics, some of which we'll mention when we discuss calculus. While his mathematical ideas were cutting-edge for the time, it is his insistence on testing theory against data that makes him a true revolutionary. By 1600 very few of Aristotle's easy-to-refute claims had been rejected due to experimental testing; the hostility Galileo provoked by doing so perhaps explains why. This is the core of the scientific revolution: primacy is given to experiment and observation over ancient 'wisdom,' whatever the source.

Galileo was finally cleared of heresy by the Catholic Church in 1992.

Exercises 7.3. 1. *Compute* today's date in the Julian calendar and explain your calculation.

2. (A problem of Copernicus) Given the three sides of an isosceles triangle, to find the angles.

   Suppose $\overline{AB}$ and $\overline{AC}$ are the equal legs of the triangle. Circumscribe a circle around the triangle and draw another with center A and radius $\overline{AD} = \frac{1}{2}\overline{AB}$.

   (a) Why is Copernicus introducing the second circle?

   (b) Explain why the ratio of each of the equal sides to the base of $\triangle ABC$ equals that of the radius $\overline{AD}$ to the chord $\overline{DE}$.

   (c) If $|AB| = |AC| = 10$ and $|BC| = 6$, use modern trigonometry to find the three angles of the original triangle.

3. Verify that Heron's formula gives the same solution to Pitiscus' problem (pg. 78).

4. Given that Earth's orbital period is 1 year and that the mean distance of Mars from the sun is 1.524 times that of Earth, use Kepler's third law to determine the orbital period of Mars.

5. According to Kepler's second law, at what point in a planet's orbit will it be moving fastest?

6. Galileo states the following.

   > A projectile fired at an angle $\alpha = 45°$ above the horizontal at a given initial speed reaches a distance of 20,000. Then, with the *same* initial speed it will reach a distance of 17,318 when $\alpha = 60°$, or $\alpha = 30°$.

   Check this statement: if you want a challenge, try to do without the standard Physics formulæ and instead use ratios!

8 Analytic Geometry and Calculus

8.1 Axes and Co-ordinates

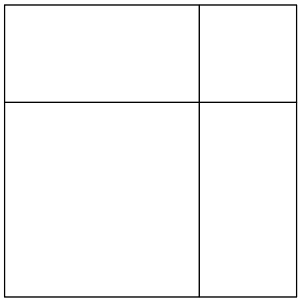
Modern mathematics is almost entirely *algebraic*: we trust equations and the rules of algebra more than pictures. For example, modern mathematics considers the expression $(x+y)^2 = x^2 + 2xy + y^2$ to follow from the laws (axioms) of algebra:

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y) && \text{(definition of `square')}\\ &= x(x+y) + y(x+y) && \text{(distributive law)}\\ &= x^2 + xy + yx + y^2 && \text{(distributive law twice more)}\\ &= x^2 + 2xy + y^2 && \text{(commutativity)}\end{aligned}$$

For most of mathematical history, this result would have been viewed *geometrically* as in Euclid's *Elements* (Thm II. 4):

> The square on two parts equals the squares on each part plus twice the rectangle on the parts.

The proof was a simple picture.



We've seen how algebra and algebraic notation were slowly adopted in renaissance Europe. While its utility for efficient calculation was noted, algebra was not initially considered acceptable for *proof* and calculations would be justified geometrically. From our modern viewpoint this seems backwards: if a student today were asked to prove Euclid's result, they'd likely label the 'parts' x and y, before using the algebraic formula at the top of the page to justify the result! Of course each line in the algebraic proof has a geometric basis:

- Distributivity says that the rectangle on a side and two parts equals the sum of the rectangles on the side and each of the parts respectively.
- Commutativity says that a rectangle has the same area if rotated 90°.

Modern mathematics has converted geometric rules into algebra and largely forgotten their geometric origin! This slow movement from geometry to algebra is one of the great revolutions of mathematical history, completely changing the way mathematicians *think.* More practically, the conversion to algebra allows easier generalization: how would one geometrically justify an expression such as

$$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 \text{ ?}$$

Euclidean Geometry is *synthetic:* based on purely geometric axioms without formulæ or co-ordinates. The revolution of *analytic geometry* was to marry algebra and geometry using *axes* and *co-ordinates.* Modern geometry is primarily analytic and it is now rare to find a mathematician working solely in synthetic geometry—algebra's domination of Euclidean geometry is total! The critical step in this revolution was made almost simultaneously by two Frenchmen…

Pierre de Fermat (1601–1665) Mathematics was Fermat's pastime rather than his profession, though this didn't prevent him making great strides in several areas such as probability, analytic geometry, early calculus, number theory and optics.[52] Some of Fermat's fame comes from his enigma, with most of what we know of his work coming in letters to friends in which he rarely offers proofs. He would regularly challenge friends to prove results, and it is often unknown whether he had proofs himself or merely suspected a general statement. Being outside the mainstream, his ideas were often ignored or downplayed. When he died, his notes and letters contained many unproven claims. Leonhard Euler (1707–1783) in particular expended much effort proving several of these.

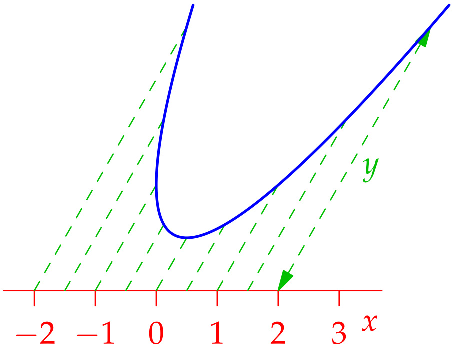
Fermat's approach to analytic geometry was not dissimilar to that of Descartes which we shall describe below: he introduced a single axis which allowed the conversion of curves into algebraic equations. We'll return to Fermat when we discuss the beginnings of calculus in the next section.

René Descartes (1596–1650) In his approach to mathematics, Descartes is the chalk to Fermat's cheese, rigorously recording everything. His defining work is 1637's *Discours de la méthode...*[53] While enormously influential in philosophy, *Discours* was intended to lay the groundwork for investigation within mathematics and the sciences—Descartes finishes *Discours* by commenting on the necessity of experimentation in science and on his reluctance to publish due to the environment of hostility surrounding Galileo's prosecution.[54] The copious appendices to *Discours* contain Descartes' scientific work. It is in one of these, *La Géométrie,* that Descartes introduces axes and co-ordinates.

We now think of Cartesian axes and co-ordinates as *plural*, but both Fermat and Descartes used only one axis. Here is a sketch of their approach.

Draw a straight line (the axis) containing two fixed points labelled 0 (the *origin*) and 1. All points on the axis are identified with numbers x (originally only positive).

A curve is described as an algebraic relationship between x and the distance y from the axis to the curve measured using a family of parallel lines intersecting both.



Neither Descartes nor Fermat had a second axes, though their approach implicitly imagines one, the measuring line through the origin. It therefore makes sense for us to speak of the *co-ordinates* (x, y); the modern terms *abscissa* (x) and *ordinate* (y) date from shortly after the time of Descartes. It wasn't long before a second axis orthogonal to the first was instituted (Frans van Schooten, 1649), an approach that quickly became standard.

---

[52] Fermat was wealthy but not aristocratic, attending the University of Orléans for three years where he trained as a lawyer. You've likely encountered his name in relation to two famous results in number theory:

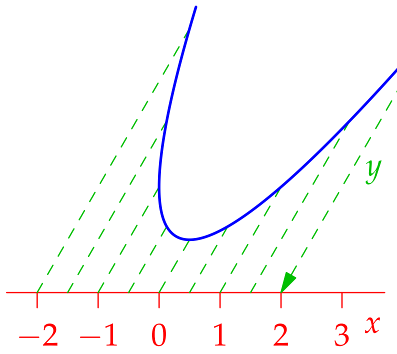
Fermat's Little Theorem p prime $\Longrightarrow x^p \equiv x \pmod{p}$ for all integers x.

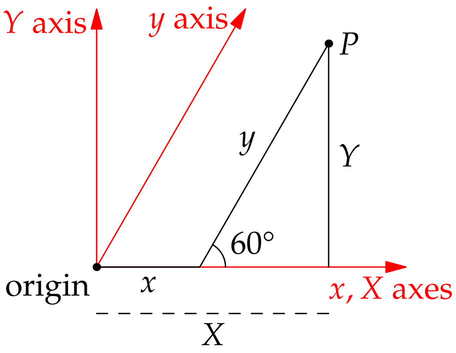
Fermat's Last Theorem If $n \in \mathbb{N}_{\geq 3}$, then $x^n + y^n = z^n$ has no integer solutions with $xyz \neq 0$. Fermat is not believed to have proved this beyond a special case $(n = 4)$, with a complete proof not appearing until the 1990s.

[53] *...of rightly conducting one's reason and of seeking truth in the sciences.* The primary part of this work is philosophical and contains his famous phrase *cogito egro sum* (*I think therefore I am*).

[54] At this time, France was still Catholic. Descartes had moved thence to Holland in part to pursue his work more freely. In 1649 Descartes moved to Sweden where he died the following year.

Example 1 The previous picture shows some of the flexibility in Descartes' approach. The curve $y = x^2 + 1$ is drawn, where the 'y-axis' is inclined 60° to the horizontal. To recognize the curve in a more standard fashion, we can perform a change of co-ordinates. Suppose $P = (x, y)$ with respect to the slanted axes and (X, Y) with respect to the usual orthogonal axes.





The second picture shows that

$$\begin{cases} X = x + y\cos 60^\circ = x + \frac{1}{2}y \\ Y = y\sin 60^\circ = \frac{\sqrt{3}}{2}y \end{cases}$$

For any point on the curve,

$$\begin{aligned} \sqrt{3}X - Y = \sqrt{3}x &\implies (\sqrt{3}X - Y)^2 = 3x^2 = 3(y-1) = 3\left(\frac{2}{\sqrt{3}}Y - 1\right) \\ &\implies 3X^2 - 2\sqrt{3}XY + Y^2 - 2\sqrt{3}Y + 3 = 0 \end{aligned}$$

which is an implicit equation for a parabola.[55]

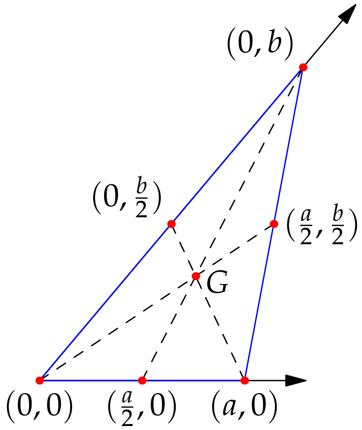
Example 2 Analytic geometry affords easy proofs of many results that are significantly harder in Euclidean geometry. For instance, here is the famous *centroid theorem.*

Choose axes pointing along two sides of a triangle with the origin at one vertex.

If the two axes-side lengths are a and b, then the third side has equation $bx + ay = ab$ or $y = b - \frac{b}{a}x$ and the midpoints of the sides have co-ordinates as in the picture. Now compute the point 1/3 of the way along each median: for instance

$$\frac{2}{3}\left(\frac{a}{2}, 0\right) + \frac{1}{3}(0, b) = \frac{1}{3}(a, b)$$

One obtains the same result with the other medians, whence *all three meet at a common point* G, the *centroid* of the triangle.



This ability to *choose axes to fit the problem* is a critical advantage of analytic geometry, largely dispensing with the tedious consideration of *congruence* in synthetic geometry.

---

[55] This really is a parabola, just rotated! This may be confirmed by computing its *discriminant*: a non-degenerate quadratic curve $aX^2 + bXY + cY^2 +$ linear terms is a parabola if and only if $b^2 - 4ac = 0$.

Descartes used his method to solve problems that had proved more difficult synthetically, such as finding complicated intersections of curves. As we'll see in the next section, such arguments would often rely on his Factor Theorem (pg. 75). Given the novelty of his approach, Descartes typically gave geometric proofs of all assertions to back up his algebraic work. However he also saw the future, stating that once several examples were done it was no longer necessary to draw physical lines and provide a geometric argument, *the algebra was the proof.*

Exercises 8.1.

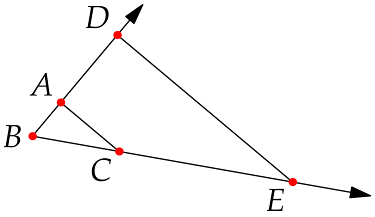
1. Assume that $xy = c$ represents a hyperbola with asymptotes the x- and y-axes. Show that $xy + c = rx + sy$ also represents a hyperbola, and find its asymptotes.
2. Determine the locus of the equation $b^2 - 2x^2 = 2xy + y^2$.

   (*Hint: add x^2 to both sides and remember that 'axes' do not have to be orthogonal…*)
3. We describe a method whereby Descartes constructed the product of two lengths.

   Let $\overrightarrow{BC}$ and $\overrightarrow{BD}$ be rays forming an acute angle at B. Our goal is to multiply $|BD|$ by $|BC|$.

   Suppose $|AB| = 1$, where A lies on $\overrightarrow{BD}$.

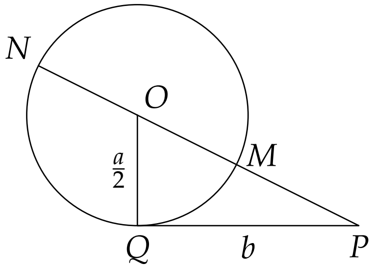
   Join $\overline{AC}$ and draw $\overline{DE}$ parallel to $\overline{AC}$ so that $E \in \overrightarrow{BC}$.

   Prove that $|BE|$ is the product of $|BD|$ and $|BC|$.

   Similarly, given lengths $|BE|$ and $|BD|$, construct a segment whose length is the quotient $\frac{|BE|}{|BD|}$.
4. Here is a geometric justification of Descartes for the solution of the equation $x^2 = ax + b^2$, where $a, b > 0$.

   - Let $\overline{OQ}$ and $\overline{PQ}$ be perpendicular with lengths $\frac{a}{2}$ and b respectively.
   - Draw the circle centered at O with radius $\frac{a}{2}$.
   - Draw and extend the line $\overleftrightarrow{OP}$.
   - The solution is $x = |NP|$.

   (a) Prove that Descartes' construction indeed recovers a solution.

   (b) Show that the other solution to the equation $x^2 = ax + b^2$ is negative (*false* to Descartes). How is it visible in the picture?

   (c) Explain how the same picture could be used to solve the equation $x^2 + ax = b^2$.
5. Following the work of Islamic mathematicians such as Omar Khayyam, Fermat could describe the solutions to certain cubic equations as the intersection of two conics. For example, to solve $x^3 + bx^2 = bc$ $(b, c > 0)$, he would introduce a new variable y by setting both sides equal to bxy. The (positive) solution is then the x-solution to the system of equations

   $$\begin{cases} x^2 + bx = by \\ xy = c \end{cases}$$

   Sketch these curves explicitly for the example $x^3 + 4x^2 = 24$.

8.2 The Beginnings of Calculus

At the heart of calculus is the relationship between *velocity, displacement, rate of change* and *area.*

- The *instantaneous velocity* of a particle is the *rate of change* of its *displacement.*
- The *displacement* of a particle is the *net area* under its *velocity-time* graph.

To state these principles requires graphs. Analytic geometry makes computation much easier (*rate of change* means *slope*). Once it appeared in the early 1600s, the rapid development of calculus was arguably inevitable. However, many of the basic ideas long predate Descartes and Fermat.

In the context of the above principles, the *Fundamental Theorem of Calculus* is intuitive: complete knowledge of displacement (from some starting point) is equivalent to complete knowledge of velocity. The modern statement is more daunting, though familiar:

Theorem. 1. *If f is continuous on $[a, b]$, then $F(x) := \int_a^x f(u)\,\mathrm{d}u$ is continuous on $[a, b]$, differentiable on (a, b), and $F'(x) = f(x)$.*

2. *If F is continuous on $[a, b]$ with integrable derivative on (a, b), then $\int_a^b F'(x)\,\mathrm{d}x = F(b) - F(a)$.*

The triumph of the modern version is its abstraction and wide applicability: we've gone way beyond considerations of velocity (f) and displacement (F). The challenge of *teaching*[56] and *proving* the Fundamental Theorem lies in developing and understanding what is meant by *continuous, differentiable* and *integrable.* The quest for good definitions of these concepts is the story of analysis in the 17-1800s. We begin with some older considerations of the velocity and area problems.

The Velocity Problem pre-1600

The modern ideas of *uniform/average* velocity are straightforward:

> Measure how far an object travels in a given time interval and divide one by the other.

While several ancient Greek mathematicians considered this (and uniform acceleration), the challenge of considering a ratio of two unlike quantities (distance : time) proved difficult to surmount. Around 1200, Gerard of Brussels resurrected this approach as a definition of velocity, though it was not considered a numerical quantity in its own right.

Gerard was credited in the 1330s by the Oxford/Merton Thinkers as influencing their investigations of *instantaneous velocity,* a much more difficult issue. They offered the following definition and made the first statement of the 'mean speed theorem,' though both are vague and logically dubious.

Definition. The *velocity* of a particle at an instant will be measured as the uniform velocity along the path that would have been taken by the particle if it continued with that velocity.

This is really a convoluted idea of inertial motion.

Theorem. *If a particle is uniformly accelerated from rest to some velocity, it will travel half the distance it would have traveled over the same interval with the final velocity.*

---

[56] Calculus students can easily be taught the mechanics of calculus (the power law, chain rule, etc.) without having any idea of its meaning; witness the power and curse of analytic geometry and algebra!

For centuries it was thought that Galileo was the first to state such ideas (compare his falling body discussion, pg. 81), but the Oxford group beat him by 250 years. They had no algebra with which to prove their assertions and essentially only asserted examples.

In the 1350's, Nicolas Oresme (Paris) considered velocity geometrically by (essentially) drawing velocity-time graphs. As we saw previously, this is essentially the approach taken by Galileo; it is also an early version of *axes.* A major difference is that Galileo married mathematics to *observation*; uniform acceleration for Galileo was precisely the motion of a falling body.

A proper definition of *instantaneous velocity* is difficult because it requires *limits,* measuring average velocity over smaller and smaller intervals. You are in good company if you find this challenging: Zeno's arrow paradox is essentially an objection to the very idea of instantaneous velocity! Even if one accepts the concept, its direct measurement, even in modern times, is essentially impossible.[57]

The Area Problem pre-1600

We've previously seen two situations in which calculus-like methods were used to describe areas.

- Archimedes (sec. 3.4) computed/approximated the area of a circle and inside a parabola using infinitely many triangles. His 'cross-section' approach to finding area and volume also seems modern, though this work remained unknown until 1899.
- Kepler (pg. 79) argued for his second law (equal areas in equal times) using infinitesimally small triangles to approximate segments of an ellipse. He also applied this method to several other problems, crediting Archimedes with the approach.

The modern notion of Riemann sums is just a special case of approximating an area using small rectangles: the philosophical challenge is again the notion of *limits* and infinitesimals.

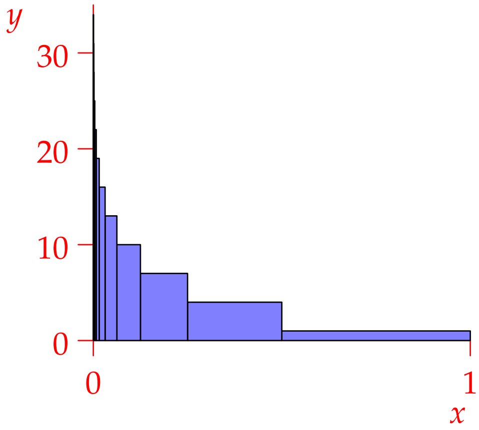
In an early antecedent, Oresme describes how to compute the distance travelled by a particle whose speed is constant on a sequence of intervals. For example:

Over the time interval $\left[\frac{1}{2^{n+1}}, \frac{1}{2^n}\right)$ a particle travels at speed $1+3n$. How far does it travel in 1 second?

Oresme drew boxes to compute areas and obtain

$$d = \sum_{n=0}^{\infty}(1+3n)\left(\frac{1}{2^n}-\frac{1}{2^{n+1}}\right) = \sum_{n=0}^{\infty}\frac{1+3n}{2^{n+1}} = 4$$

Similar to Archimedes, the infinite sum was evaluated by spotting two patterns:



$$\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}\cdots+\frac{1}{2^{n+1}} = 1-\frac{1}{2^{n+1}} \qquad \frac{0}{2}+\frac{1}{2^2}+\frac{2}{2^3}+\frac{3}{2^4}+\cdots+\frac{n}{2^{n+1}} = 1-\frac{n+2}{2^{n+1}}$$

Oresme had neither our notation nor our (limit-dependent) concept of an infinite series! He also worked with similar problems for uniform accelerations over intervals. These are not true Riemann sums, nor are they physical, for a particle cannot suddenly change speed!

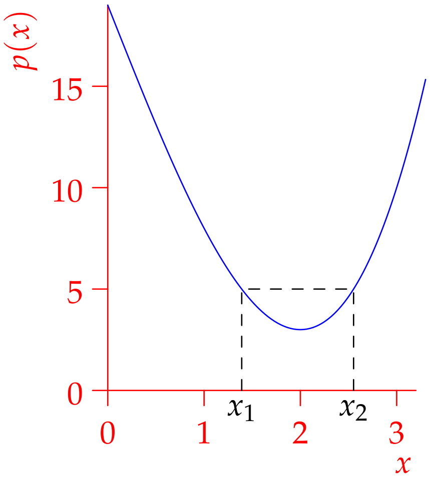
---

[57] For instance, radar Doppler-shift (as used to catch speeding motorists) requires measuring the wavelength of a radar beam, which essentially compute the average velocity over a very small time interval.

Calculus à la Fermat & Descartes

The advent of analytic geometry allowed Fermat and Descartes to turn the computation of instantaneous velocity and related differentiation problems into algebraic processes. The velocity of an object is now identified with the slope of the displacement-time graph, which can be computed using variations on the modern method of *secant lines.* We discuss their competing methods.

Fermat's method of adequation Consider the function $p(x) = x^3 - 12x + 19$; the goal is to find the minimum, which we know to be located at the x-value $m = 2$.
Fermat argues that if x_1, x_2 are located near m such that $p(x_1) = p(x_2)$, then the polynomial $p(x_2) - p(x_1)$ (which equals zero!) is divisible by $x_2 - x_1$. Indeed



$$
\begin{aligned}
0 &= \frac{p(x_2) - p(x_1)}{x_2 - x_1} \\
&= \frac{x_2^3 - 12x_2 + 19 - x_1^3 + 12x_1 - 19}{x_2 - x_1} \\
&= \frac{(x_2 - x_1)(x_2^2 + x_1x_2 + x_1^2 - 12)}{x_2 - x_1} \\
&= x_2^2 + x_1x_2 + x_1^2 - 12
\end{aligned}
$$

Since this holds for any x_1, x_2 with $p(x_1) = p(x_2)$, Fermat claims it also holds when $x_1 = x_2 = m$ (note the assumption of continuity!), and he concludes

$$3m^2 - 12 = 0 \implies m = 2$$

By considering values of x near m, it is clear that Fermat really has found a local minimum. We recognize the idea that the slope of the tangent line is zero at local extrema.

This approach dates from the 1620s and is similar to earlier work of Viète. Fermat later alters his method by considering values $p(x)$ and $p(x + e)$ for small e (x is 'adequated by e'). The difference $p(x + e) - p(x)$ is more easily divided by e without factorizing. Compared with the above, we obtain

$$
\begin{aligned}
0 &= \frac{p(x+e) - p(x)}{e} = \frac{x^3 + 3x^2e + 3xe^2 + e^3 - 12x - 12e + 19 - x^3 + 12x - 19}{e} \\
&= \frac{3x^2e + 3xe^2 + e^3 - 12e}{e} = 3x^2 - 12 + 3xe + e^2
\end{aligned}
$$

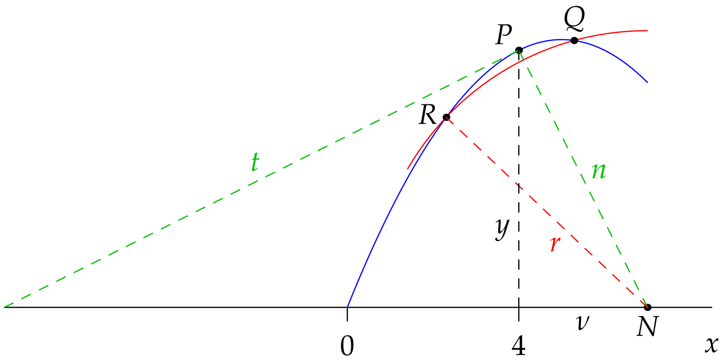
Fermat finished by setting e to zero and solving for x as before. Observe the derivative $p'(x) = 3x^2 - 12$ and how e plays the role of h in the modern expression

$$p'(x) = \lim_{h \to 0} \frac{p(x+h) - p(x)}{h}$$

Fermat's method works for any polynomial, where the limit definition of derivative requires no more than simple evaluation at $h = 0$. Fermat also extended his method to cover implicit curves and their tangents.

Descartes' method of normals Descartes and Fermat are known to have corresponded regarding their methods. Descartes indeed seems to have felt somewhat challenged by Fermat, and engaged in some criticism of his approach. Descartes' method (in *La Géométrie*) relies on circles and repeated roots of polynomials in order to compute tangents.

In this example, we compute the slope of the curve $y = \frac{1}{4}(10x - x^2)$ at the point $P = (4, 6)$.



Let $N = (4 + \nu, 0)$ be where the normal to the curve intersects the x-axis.[58] Draw a circular arc radius r centered at N. If r is close to n, the circle intersects the curve in two points Q, R near to P. The line $\overleftrightarrow{QR}$ plainly approximates the tangent at P.

The co-ordinates of Q, R may be found by solving algebraic equations: substituting $y = \frac{1}{4}(10x - x^2)$ into the equation for the circle results in an equation with two known roots, namely the x-values of Q and R. By the factor theorem,

$$\begin{cases} \big(x - (4+\nu)\big)^2 + y^2 = r^2 \\ y = \frac{1}{4}(10x - x^2) \end{cases} \implies (x - Q_x)(x - R_x)f(x) = 0$$

where $f(x)$ is some polynomial (in this case quadratic). Rather than doing this explicitly, Descartes observes that if the radius r is adjusted until it *equals* n, then Q and R coincide with P and the above equation has a *double-root*:

$$\begin{cases} \big(x - (4+\nu)\big)^2 + y^2 = n^2 \\ y = \frac{1}{4}(10x - x^2) \end{cases} \implies (x - P_x)^2 f(x) = (x-4)^2 f(x) = 0$$

Factorization can be done by hand using long-division (note that ν and n are currently unknown!): substituting as above, we obtain

$$\begin{aligned} 0 &= x^4 - 20x^3 + 116x^2 - 32(4+\nu)x + 16(4+\nu)^2 - 16n^2 \\ &= (x-4)^2(x^2 - 12x + 4) + 32(3-\nu)x + 16(12 + 8\nu + \nu^2 - n^2) \end{aligned}$$

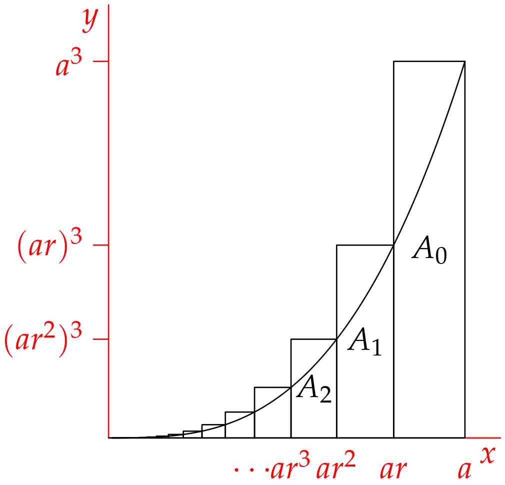
Since the remainder $32(3-\nu)x + 16(12 + 8\nu + \nu^2 - n^2)$ must be the zero polynomial, we conclude that $\nu = 3$. By similar triangles, the slope of the curve at P is therefore

$$\frac{y}{\sqrt{t^2 - y^2}} = \frac{\nu}{y} = \frac{1}{2}$$

---

[58] At the time, ν was known as the *subnormal* and t the *tangent*.

Fermat and Area The previous methods permit *differentiation*, albeit inefficiently. Fermat also approached the area problem, in a manner similar to Oresme. Here is an example where we find the area under the curve $y = x^3$ between $x = 0$ and $x = a$.



Let $0 < r < 1$ be constant. The rectangle on the interval $[ar^{n+1}, ar^n]$ touching the curve at its upper right-corner has area

$$A_n = (ar^n - ar^{n+1}) \cdot (ar^n)^3 = a^4(1-r)r^{4n}$$

The sum of the areas is therefore

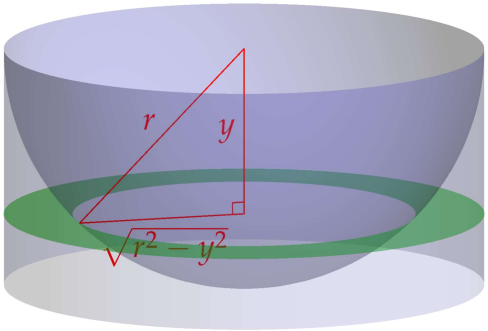
$$\sum_{n=0}^{\infty} A_n = a^4(1-r) \sum_{n=0}^{\infty} r^{4n} = \frac{a^4(1-r)}{1-r^4}$$
$$= \frac{a^4}{1+r+r^2+r^3}$$

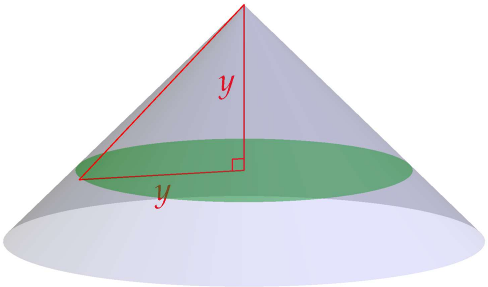
Setting $r = 1$ recovers the area under the curve: $\frac{1}{4}a^4$.

This is non-rigorous by modern standards and again implicitly invokes limits[59] by setting $r = 1$. Regardless, Fermat is able to establish the power law $\int_0^a x^n \, dx = \frac{1}{n+1}a^{n+1}$ for any positive integer n.

Italian Calculus in the 17th Century: the Area and Volume Problems

In contrast to the work of Fermat and Descartes, contemporary Italian scholars were more focused on integration problems. Here is Galileo's classic 'soup bowl' problem, where he compares the volume between a hemisphere and a cylinder to that of a cone.





Galileo observed[60] that the cross-sectional areas on both sides are equal (to πy^2). Since all cross-sections are equal, so must be the volumes. Unfortunately for Galileo, he couldn't sufficiently address two philosophical objections:

The zero-measure problem If cross-sections are 'equal,' then the top cross-sections state that a circle 'equals' a point.

Infinitesimals sum to the whole? Can we claim that equal cross-sections imply equal volumes?

It was Galileo's advocacy on these points that first gained him notoriety with the Church. His later evangelism for the Copernican theory merely rekindled old animosities.

---

[59] For Fermat, $r = \frac{n}{m}$ would have been rational, and he'd have set $m = n$ at the end as in his adequation method

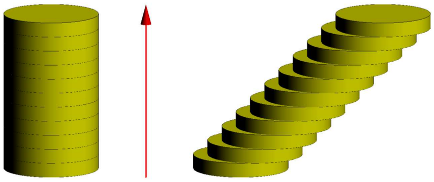
[60] As did Archimedes 1900 years earlier (pg. 32), though Galileo was unaware of it.

Bonaventura Cavalieri (1598–1647) Cavalieri, a student of Galileo and a Jesuat scholar, gave a more thorough discussion of indivisibles in 1635. He is particularly remembered for *Cavalieri's principle*:

> If geometric figures have proportional cross-sectional measure at every point relative to a line, then the figures have measure in the same proportion.

Galileo's soup bowl is an example of this reasoning, where the 'line' is any vertical.

Another classic example involves sliding a stack of coins or a deck of cards; the volume of the slanted coin stack equals that of the cylinder.

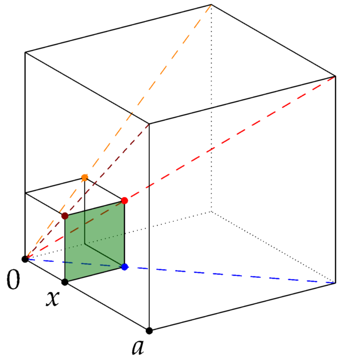


Extending his principle, Cavalieri inferred the power law $\int_0^a x^n \, dx = \frac{1}{n+1}a^{n+1}$, giving reasonable arguments for $n = 1$ and 2. Here is a sketch of his approach for $n = 2$.

Draw a cube of side x inside a cube of side a. This consists of three congruent pyramids with base a^2 and eight a.

Consider the pyramid apex O and whose base is the square face nearest the viewer. The cross-section of this pyramid at position x has area x^2. In Cavalieri's language, the pyramid is 'all the squares;' in modern language

$$\int_0^a x^2 \, dx = \frac{1}{3}a^3$$



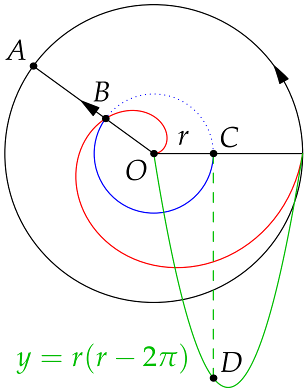
Cavalieri also used his method (Book IV, Prop 19 of *Geometria Indivisibilis*) to calculate the area enclosed in an Archimedean spiral. Suppose B moves at constant speed along a line $\overline{OA}$ rotating at constant speed about O. In polar co-ordinates B traces a curve $r = k\theta$ for $0 \le \theta \le 2\pi$ (we take $k = 1$).

If $|OB| = r$, then the arc *inside* the spiral has length

$$\ell(r) = 2\pi r \cdot \frac{2\pi - \theta}{2\pi} = r(2\pi - r)$$

Imagine the arc as a noodle which, when cut at B and allowed to fall, forms the line $\overline{CD}$. The area in the spiral therefore equals that within the parabola which (thanks to Archimedes and Cavalieri himself) is $\frac{4}{3}$ that of the largest triangle that can fit inside, namely

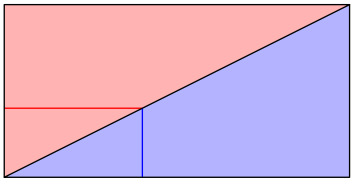
$$\frac{4}{3} \cdot \frac{1}{2}(2\pi)\ell(\pi) = \frac{4}{3}\pi^3$$



Cavalieri did this slightly differently, his parabola being drawn in a rectangle and the difference subtracted from a triangle, but the above is easier to visualize.

Cavalieri did not court controversy like Galileo, being well-aware of the contentious nature of indivisibles and taking great pains to distinguish 'all the lines/squares' of a figure from the figure itself. *Geometria Indivisibilis* was dense and difficult, so it wasn't easy to challenge. Cavalieri therefore remained relatively safe, even as his political rivals (the Jesuit order) worked hard to stamp out the 'dangerous' study of indivisibles.

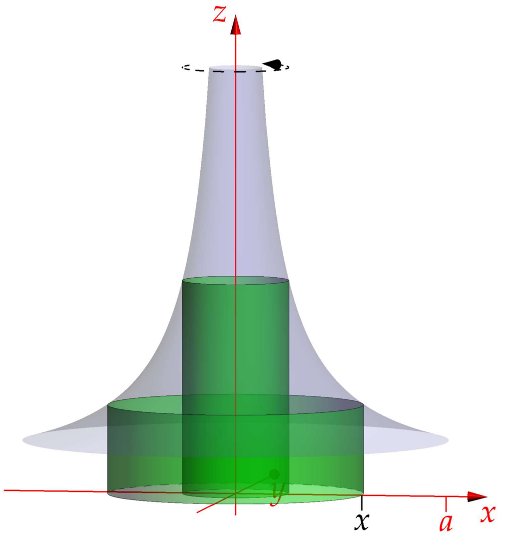
Evangelista Torricelli (1608–1647) A contemporary of Galileo and Cavalieri, Torricelli made several applications of Cavalieri's principle and advocated for its careful use.

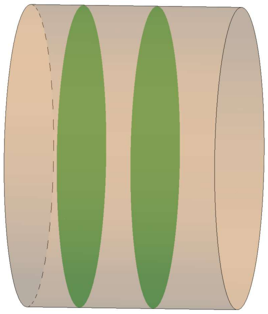


If the sides of a rectangle are in the ratio 2 : 1, so also are indicated red and blue segments. In Cavalieri's language, 'all the lines' of the red triangle are twice 'all the lines' of the blue triangle! Torricelli observes that the red triangle cannot have twice the area of the blue, since they are congruent and points out that Cavalieri's principle has been misapplied: the cross-sections were not measured with reference to a common line.

In modern language, $\int_0^2 \frac{1}{2}x\,\mathrm{d}x = \int_0^1 2y\,\mathrm{d}y$ are equal via the substitution $x = 2y$. The point is that the infinitesimals are also in the *same ratio* $\mathrm{d}x : \mathrm{d}y = 2 : 1$.

Another of Torricelli's examples offers a seeming paradox.





The hyperbola with equation $z = \frac{1}{x}$ is rotated around the z-axis. A cylinder centered on the z-axis with radius x lying under the surface has surface area

$$A = \text{circumference} \cdot \text{height} = 2\pi xz = 2\pi$$

Underneath the graph at x, Torricelli draws a circular disk with area 2π. Since the area of this disk is independent of x, he argues that the volume under the original surface out to radius a equals the volume of the solid cylinder:

$$V = 2\pi a$$

Torricelli argues that this is a correct use of Cavalieri's principle since the cylindrical 'cross-sections' and the circular cross-sections are both measured with respect to the same line (the x-axis).

This is precisely the method of volume by cylindrical shells that we learn in modern calculus:

$$V = \int_0^a 2\pi x \cdot \frac{1}{x}\,\mathrm{d}x = 2\pi a$$

The conundrum is that the surface is *infinitely tall*! How can we justify the idea that it lies above a *finite* volume?

Galileo, Cavalieri and Torricelli mark the end of 400 years of Italian dominance of in science and mathematics dating back to Fibonacci. Their ideas were too controversial to thrive so close to the center of Church power. The center of European science and philosophy therefore moved northward. The English and French (protestant) reformations of the 1500s together with developing ideas of reformed government[61] meant that Northern Europe provided more fertile ground for new ideas.

---

[61] For instance Hobbes' *Leviathan* written during the English Civil War (1642–1651) was a plea for the constraint of absolute monarchical power. The War itself indeed proved decapitatingly effective at reining in a King...

Exercises 8.2. 1. Find the maximum of $p(x) = 5 + x - 2x^2$ using Fermat's first method.

2. Use Fermat's second method ("$+e$") to find the maximum of $bx - x^3$. How might Fermat decide which of the two solutions to choose as his maximum?

3. Justify Fermat's first method of determining maxima and minima by showing that if M is a maximum of $p(x)$, then the polynomial $p(x) - M$ always has a factor $(x-a)^2$, where a is the value of x giving the maximum.

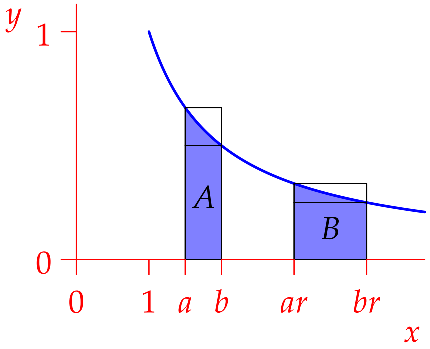
4. Use Descartes' method of normals to compute the slope of the curve $y = x^2$ at (a, a^2).

5. Suppose the surface of a sphere of radius r is subdivided into infinitesimal regions of equal area. Following Kepler, use the formula for the volume of a cone ($\frac{1}{3}$base$\cdot$height) to find the relationship between the volume V of the sphere and its surface area A.

6. (A problem of Kepler) Show that the largest circular cylinder that can be inscribed in a sphere is one in which the ratio of diameter to altitude is $\sqrt{2} : 1$.

   (*Hint: Relate the problem to finding the maximum of the function* $f(x) = x - \frac{1}{4}x^3$, *for which you can use modern calculus*)

7. We consider a version of Gregory St. Vincent's 1647 approach to the area under the hyperbola $xy = 1$.

   (a) If $1 < a < b$ and $r > 0$ as in the picture, explain why the areas A and B under the curve satisfy the same inequalities

   $$\frac{b-a}{b} < A, B < \frac{b-a}{a}$$

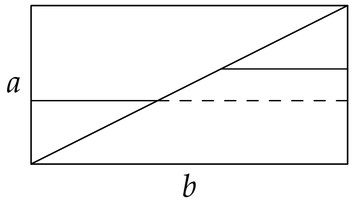
   (*Since* $[a, b]$ *may be subdivided into arbitrarily many subintervals, the areas* A *and* B *are therefore equal*)

   (b) If $A(x)$ is the area under the hyperbola between 1 and x, explain why A satisfies the *logarithmic* identity

   $$A(x_1 x_2) = A(x_1) + A(x_2)$$

   Why are you not surprised by this?
   (*For simplicity, assume* $1 < x_1 < x_1 x_2$)

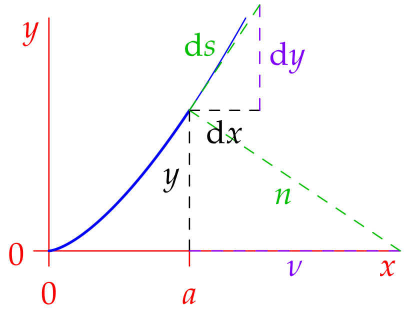
8. Consider two copies of triangle with sides a, b arranged into a rectangle. Argue that 'all the lines' of the rectangle are twice 'all the lines' of the triangle.
   (*In modern language,* $\int_0^a b \, dy = 2 \int_0^a \frac{b}{a} y \, dy$)

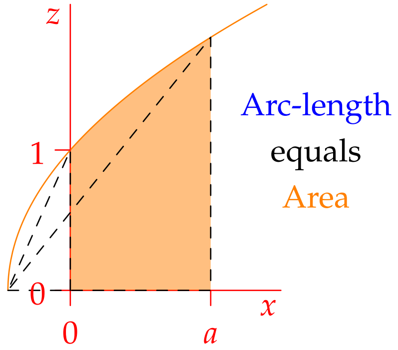
   

9. Repeat Cavalieri's analysis of the spiral (pg. 92) to find the area inside one revolution of the curve $r = k\theta$ for any $k > 0$.

8.3 Calculus in the late 1600s

By the second half of the 17th century, the mathematical center of Europe had moved northwards, to France, Germany, Holland and Britain. In this section we present some of this work, culminating in the efforts of Newton and Leibniz.





Hendrick van Heuraet (1634–1660) Working in Holland, van Heuraet studied Descartes' and argued that the arc-length of a curve described by a function y equals the area under the curve $z = \frac{n}{y}$ where n is the 'normal' curve. His method appeared in Frans van Schooten's 1659 version of Descartes' *La Geometrie.* To see why this should make sense, recall Descartes' method of normals and observe that the ratio $n : y$ equals that of $\mathrm{d}s : \mathrm{d}x$ in a differential triangle.

His most famous example involved calculating the arc-length of the curve $y^2 = x^3$. By Descartes' method,

$$\begin{cases} (x-a-v)^2 + y^2 = n^2 \\ y^2 = x^3 \end{cases}$$

$$\begin{aligned} \implies 0 &= x^3 + x^2 - 2(x+v)x + (a+v)^2 - n^2 \\ &= (x-a)^2(x+2a+1) + (3a^2-2v)x \\ &\qquad + v^2 + 2av - 2a^3 - n^2 \end{aligned}$$

Since the remainder must be zero, we conclude that $v = \frac{3}{2}a^2$, from which

$$n = \sqrt{v^2 + y^2} = \sqrt{\frac{9}{4}a^4 + a^3} \implies z = \frac{n}{y} = \sqrt{\frac{9}{4}a + 1}$$

The arc-length from $x = 0$ to a is therefore the area under the parabola $z^2 = \frac{9}{4}x + 1$ between the same limits. By the usual Archimedean $\frac{4}{3}$-triangle approach, we see that

$$\text{Arc-length} = \frac{4}{3}\left[\frac{1}{2}\cdot\left(\frac{4}{9}+a\right)z(a) - \frac{1}{2}\cdot\frac{4}{9}\cdot 1\right] = \left[a + \frac{4}{9}\right]^{3/2} - \frac{8}{27}$$

James Gregory (1638–1675) Hailing from Aberdeen, Scotland, Gregory studied in Italy with Stefano Angeli, a pupil of Torricelli, before returning to Scotland where he became chair of mathematics at St. Andrews, and then Edinburgh.

Gregory repeats van Heuraet's work relating the length of a curve to the area under another, before considering whether the process can be reversed: given a curve $z(x)$, can we find a curve $y(x)$ such that the arc-length of y is given by the area under z? In modern language, given z, find y such that

$$\int_0^a \sqrt{1 + y'^2}\,\mathrm{d}x = \int_0^a z\,\mathrm{d}x$$

which we'd view as solving the ODE $\frac{\mathrm{d}y}{\mathrm{d}x} = \sqrt{z^2 - 1}$. Gregory's solution was to *define* y to be the area under the curve $\sqrt{z^2 - 1}$ from $x = 0$ to a. This is essentially part 1 of the fundamental theorem: if you want something whose slope (derivative) is given, define it to be the area under the curve!

Isaac Barrow (1630–1677) Like Gregory, Barrow also studied mathematics in Italy (and France), before returning to England to become the inaugural Lucasian Professor of Mathematics[62] at Trinity College Cambridge. Barrow's work remained predominantly geometric; he stated proving geometric versions of both parts of the fundamental theorem, though credited Gregory with part of the argument. In a precursor of Newton's work, he also modified Fermat's algorithm for differentiation. For example, here is how Barrow would have found the slope of the curve $x^2 + 2xy^2 = c$ at a point (x, y):

- Replace x and y with $x + e$ and $y + a$ respectively and expand:

$$x^2 + 2ex + e^2 + 2xy^2 + 4axy + 2a^2x + 2ey^2 + 4eay + 2ea^2 = c$$

- Delete everything from the original equation $x^2 + 2xy^2 = c$ and every expression containing two or more of the terms e, a:

$$2ex + 4axy + 2ey^2 = 0$$

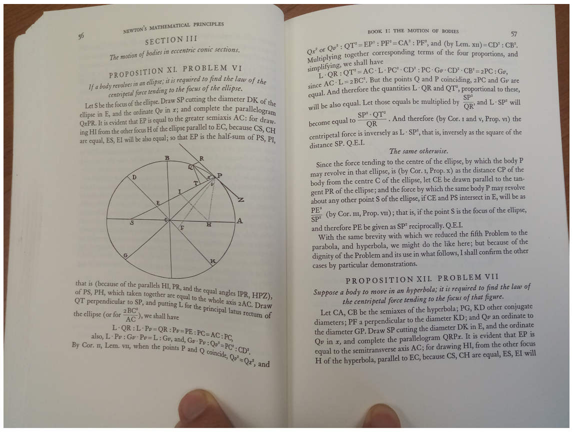
- Rearrange: the slope is the ratio $a : e = -x - y^2 : 2xy$

This is implicit differentiation, where $\mathrm{d}x = e$ and $\mathrm{d}y = a$. Note again the essential difficulty with these *algorithmic* approaches to calculus: the infinitesimal quantities e, a are necessary for the calculation, but most of them are discarded when no longer useful! Can we really calculate with objects that must simultaneously exist and be zero?[63]

Isaac Newton (1642–1727)

The caricature of Newton is of an obsessive genius—difficult to get along with, but with a phenomenal ability to concentrate on problems. One possibly apocryphal story describes how he continued lecturing to an empty room even after no-one had turned up!

We are mostly interested in Newton's mathematics, though his wider fame comes from its wide application, particularly to gravitation. In *Philosophiæ Naturalis Principia Mathematica* (1686), Newton applied his three laws of motion[64] and the machinery of calculus to prove the relationship between Kepler's laws of planetary motion (pg. 78) and an inverse-square gravitational force. The *Principia* is Newton's first published work involving calculus, though many of its results and his method appear to have been worked out 20 years previously, just after completing his undergraduate studies and while Cambridge was closed due to the 1665–6 plague epidemic.



Kepler's 1st law $\Longrightarrow$ inverse-square force

56 NEWTON'S MATHEMATICAL PRINCIPLES

SECTION III

The motion of bodies in eccentric conic sections.

PROPOSITION XI. PROBLEM VI

If a body revolves in an ellipse; it is required to find the law of the centripetal force tending to the focus of the ellipse.

Let S be the focus of the ellipse. Draw SP cutting the diameter DK of the ellipse in E, and the ordinate Q*v* in *x*; and complete the parallelogram Q*x*PR. It is evident that EP is equal to the greater semiaxis AC: for drawing HI from the other focus H of the ellipse parallel to EC, because CS, CH are equal, ES, EI will be also equal; so that EP is the half-sum of PS, PI, that is (because of the parallels HI, PR, and the equal angles IPR, HPZ), of PS, PH, which taken together are equal to the whole axis 2AC. Draw QT perpendicular to SP, and putting L for the principal latus rectum of the ellipse (or for $\frac{2\mathrm{BC}^2}{\mathrm{AC}}$), we shall have

$$\mathrm{L\cdot QR : L\cdot P}v\mathrm{ = QR : P}v\mathrm{ = PE : PC = AC : PC},$$

$$\text{also, } \mathrm{L\cdot P}v : \mathrm{G}v\cdot\mathrm{P}v = \mathrm{L} : \mathrm{G}v, \text{ and, } \mathrm{G}v\cdot\mathrm{P}v : \mathrm{Q}v^2 = \mathrm{PC}^2 : \mathrm{CD}^2.$$

By Cor. II, Lem. VII, when the points P and Q coincide, $\mathrm{Q}v^2 = \mathrm{Q}x^2$, and

BOOK I: THE MOTION OF BODIES 57

$\mathrm{Q}x^2$ or $\mathrm{Q}v^2 : \mathrm{QT}^2 = \mathrm{EP}^2 : \mathrm{PF}^2 = \mathrm{CA}^2 : \mathrm{PF}^2$, and (by Lem. XII) $= \mathrm{CD}^2 : \mathrm{CB}^2$. Multiplying together corresponding terms of the four proportions, and simplifying, we shall have

$$\mathrm{L\cdot QR : QT^2 = AC\cdot L\cdot PC^2\cdot CD^2 : PC\cdot G}v\mathrm{\cdot CD^2\cdot CB^2 = 2PC : G}v,$$

since $\mathrm{AC\cdot L = 2BC^2}$. But the points Q and P coinciding, 2PC and G*v* are equal. And therefore the quantities L · QR and QT^2, proportional to these, will be also equal. Let those equals be multiplied by $\frac{\mathrm{SP}^2}{\mathrm{QR}}$, and $\mathrm{L\cdot SP^2}$ will become equal to $\frac{\mathrm{SP^2\cdot QT^2}}{\mathrm{QR}}$. And therefore (by Cor. I and V, Prop. VI) the centripetal force is inversely as $\mathrm{L\cdot SP^2}$, that is, inversely as the square of the distance SP. Q.E.I.

The same otherwise.

Since the force tending to the centre of the ellipse, by which the body P may revolve in that ellipse, is (by Cor. I, Prop. X) as the distance CP of the body from the centre C of the ellipse, let CE be drawn parallel to the tangent PR of the ellipse; and the force by which the same body P may revolve about any other point S of the ellipse, if CE and PS intersect in E, will be as $\frac{\mathrm{PE}^3}{\mathrm{SP}^2}$ (by Cor. III, Prop. VII); that is, if the point S is the focus of the ellipse, and therefore PE be given as SP^2 reciprocally. Q.E.I.

With the same brevity with which we reduced the fifth Problem to the parabola, and hyperbola, we might do the like here; but because of the dignity of the Problem and its use in what follows, I shall confirm the other cases by particular demonstrations.

PROPOSITION XII. PROBLEM VII

Suppose a body to move in an hyperbola; it is required to find the law of the centripetal force tending to the focus of that figure.

Let CA, CB be the semiaxes of the hyperbola; PG, KD other conjugate diameters; PF a perpendicular to the diameter KD; and Q*v* an ordinate to the diameter GP. Draw SP cutting the diameter DK in E, and the ordinate Q*v* in *x*, and complete the parallelogram QRP*x*. It is evident that EP is equal to the semitransverse axis AC; for drawing HI, from the other focus H of the hyperbola, parallel to EC, because CS, CH are equal, ES, EI will

---

[62] One of the most prestigious world-wide academic positions in *theoretical Physics*, in large part due to the fame of its second incumbent: Issac Newton. Later chairs include Paul Dirac and Stephen Hawking.

[63] This is at the heart of Bishop George Berkeley's (after whom the Californian city and university are named) famous 1734 objection to calculus; that infinitesimals are merely the "ghosts of departed quantities."

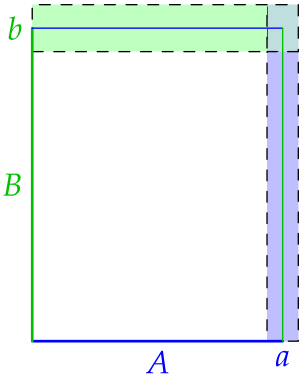
[64] I. Inertial motion; II. $F = ma$; III. Equal-and-opposite forces. Consider these as the *axioms* of Newtonian mechanics.

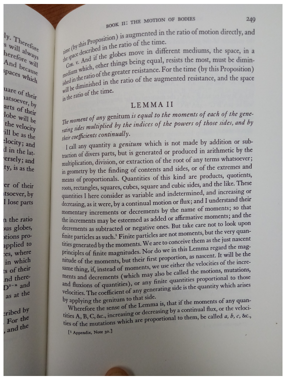
Newton's geometric presentation was typical for the time. He comments on how calculations are more efficient using indivisible methods, but that the 'hypothesis of indivisibles seems somewhat harsh' (he likely wants to counter the impression that his method is philosophically shaky). Newton's approach makes it hard for modern readers to extract calculus algorithms; indeed the *Principia* is not a calculus textbook and it is not really possible to learn calculus directly from it. Nevertheless, it contains notions of many standard concepts, for instance:

Limits/continuity Book I, Lemma I: "Quantities, …which in any finite time converge continually to equality, and before the end of that time approach nearer to each other than by any given difference, become ultimately equal."

One can see modern ideas appearing (e.g., $\forall\epsilon > 0, |a-b| < \epsilon \implies a = b$) though there is a long way to go! What, for instance does *approach* mean?

Product Rule In the following pages, Newton argues for the product rule by augmenting the sides of a rectangle.





BOOK II: THE MOTION OF BODIES 249

time (by this Proposition) is augmented in the ratio of motion directly, and the space described in the ratio of the time.

Cor. v. And if the globes move in different mediums, the space, in a medium which, other things being equal, resists the most, must be diminished in the ratio of the greater resistance. For the time (by this Proposition) will be diminished in the ratio of the augmented resistance, and the space in the ratio of the time.

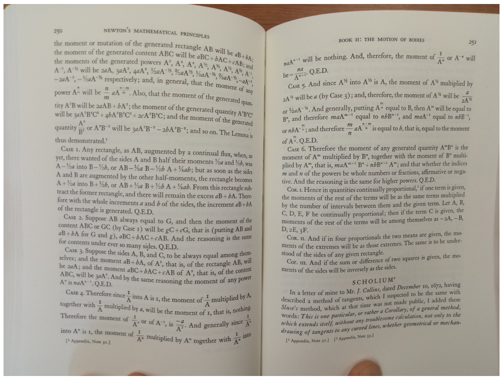
LEMMA II

The moment of any genitum *is equal to the moments of each of the generating sides multiplied by the indices of the powers of those sides, and by their coefficients continually.*

I call any quantity a *genitum* which is not made by addition or subtraction of divers parts, but is generated or produced in arithmetic by the multiplication, division, or extraction of the root of any terms whatsoever; in geometry by the finding of contents and sides, or of the extremes and means of proportionals. Quantities of this kind are products, quotients, roots, rectangles, squares, cubes, square and cubic sides, and the like. These quantities I here consider as variable and indetermined, and increasing or decreasing, as it were, by a continual motion or flux; and I understand their momentary increments or decrements by the name of moments; so that the increments may be esteemed as added or affirmative moments; and the decrements as subtracted or negative ones. But take care not to look upon finite particles as such.[1] Finite particles are not moments, but the very quantities generated by the moments. We are to conceive them as the just nascent principles of finite magnitudes. Nor do we in this Lemma regard the magnitude of the moments, but their first proportion, as nascent. It will be the same thing, if, instead of moments, we use either the velocities of the increments and decrements (which may also be called the motions, mutations, and fluxions of quantities), or any finite quantities proportional to those velocities. The coefficient of any generating side is the quantity which arises by applying the genitum to that side.

Wherefore the sense of the Lemma is, that if the moments of any quantities A, B, C, &c., increasing or decreasing by a continual flux, or the velocities of the mutations which are proportional to them, be called *a*, *b*, *c*, &c.,

[1 Appendix, Note 30.]



250 NEWTON'S MATHEMATICAL PRINCIPLES

the moment or mutation of the generated rectangle AB will be $aB+bA$; the moment of the generated content ABC will be $aBC+bAC+cAB$; and the moments of the generated powers A^2, A^3, A^4, $A^{\frac{1}{2}}$, $A^{\frac{3}{2}}$, $A^{\frac{1}{3}}$, $A^{\frac{2}{3}}$, A^{-1}, A^{-2}, $A^{-\frac{1}{2}}$ will be $2aA$, $3aA^2$, $4aA^3$, $\frac{1}{2}aA^{-\frac{1}{2}}$, $\frac{3}{2}aA^{\frac{1}{2}}$, $\frac{1}{3}aA^{-\frac{2}{3}}$, $\frac{2}{3}aA^{-\frac{1}{3}}$, $-aA^{-2}$, $-2aA^{-3}$, $-\frac{1}{2}aA^{-\frac{3}{2}}$ respectively; and, in general, that the moment of any power $A^{\frac{n}{m}}$ will be $\frac{n}{m}aA^{\frac{n-m}{m}}$. Also, that the moment of the generated quantity A^2B will be $2aAB+bA^2$; the moment of the generated quantity $A^3B^4C^2$ will be $3aA^2B^4C^2+4bA^3B^3C^2+2cA^3B^4C$; and the moment of the generated quantity $\frac{A^3}{B^2}$ or A^3B^{-2} will be $3aA^2B^{-2}-2bA^3B^{-3}$; and so on. The Lemma is thus demonstrated.[1]

Case 1. Any rectangle, as AB, augmented by a continual flux, when, as yet, there wanted of the sides A and B half their moments $\frac{1}{2}a$ and $\frac{1}{2}b$, was $A-\frac{1}{2}a$ into $B-\frac{1}{2}b$, or $AB-\frac{1}{2}a\,B-\frac{1}{2}b\,A+\frac{1}{4}ab$; but as soon as the sides A and B are augmented by the other half-moments, the rectangle becomes $A+\frac{1}{2}a$ into $B+\frac{1}{2}b$, or $AB+\frac{1}{2}a\,B+\frac{1}{2}b\,A+\frac{1}{4}ab$. From this rectangle subtract the former rectangle, and there will remain the excess $aB+bA$. Therefore with the whole increments *a* and *b* of the sides, the increment $aB+bA$ of the rectangle is generated. Q.E.D.

Case 2. Suppose AB always equal to G, and then the moment of the content ABC or GC (by Case 1) will be $gC+cG$, that is (putting AB and $aB+bA$ for G and g), $aBC+bAC+cAB$. And the reasoning is the same for contents under ever so many sides. Q.E.D.

Case 3. Suppose the sides A, B, and C, to be always equal among themselves; and the moment $aB+bA$, of A^2, that is, of the rectangle AB, will be $2aA$; and the moment $aBC+bAC+cAB$ of A^3, that is, of the content ABC, will be $3aA^2$. And by the same reasoning the moment of any power A^n is naA^{n-1}. Q.E.D.

Case 4. Therefore since $\frac{1}{A}$ into A is 1, the moment of $\frac{1}{A}$ multiplied by A, together with $\frac{1}{A}$ multiplied by a, will be the moment of 1, that is, nothing. Therefore the moment of $\frac{1}{A}$, or of A^{-1}, is $\frac{-a}{A^2}$. And generally since $\frac{1}{A^n}$ into A^n is 1, the moment of $\frac{1}{A^n}$ multiplied by A^n together with $\frac{1}{A^n}$ into

[1 Appendix, Note 31.]

BOOK II: THE MOTION OF BODIES 251

naA^{n-1} will be nothing. And, therefore, the moment of $\frac{1}{A^n}$ or A^{-n} will be $-\frac{na}{A^{n+1}}$. Q.E.D.

Case 5. And since $A^{\frac{1}{2}}$ into $A^{\frac{1}{2}}$ is A, the moment of $A^{\frac{1}{2}}$ multiplied by $2A^{\frac{1}{2}}$ will be a (by Case 3); and, therefore, the moment of $A^{\frac{1}{2}}$ will be $\frac{a}{2A^{\frac{1}{2}}}$ or $\frac{1}{2}aA^{-\frac{1}{2}}$. And generally, putting $A^{\frac{m}{n}}$ equal to B, then A^m will be equal to B^n, and therefore maA^{m-1} equal to nbB^{n-1}, and maA^{-1} equal to nbB^{-1}, or $nbA^{-\frac{m}{n}}$; and therefore $\frac{m}{n}aA^{\frac{m-n}{n}}$ is equal to b, that is, equal to the moment of $A^{\frac{m}{n}}$. Q.E.D.

Case 6. Therefore the moment of any generated quantity A^mB^n is the moment of A^m multiplied by B^n, together with the moment of B^n multiplied by A^m, that is, $maA^{m-1}B^n+nbB^{n-1}A^m$; and that whether the indices m and n of the powers be whole numbers or fractions, affirmative or negative. And the reasoning is the same for higher powers. Q.E.D.

Cor. i. Hence in quantities continually proportional,[1] if one term is given, the moments of the rest of the terms will be as the same terms multiplied by the number of intervals between them and the given term. Let A, B, C, D, E, F be continually proportional; then if the term C is given, the moments of the rest of the terms will be among themselves as −2A, −B, D, 2E, 3F.

Cor. ii. And if in four proportionals the two means are given, the moments of the extremes will be as those extremes. The same is to be understood of the sides of any given rectangle.

Cor. iii. And if the sum or difference of two squares is given, the moments of the sides will be inversely as the sides.

SCHOLIUM[2]

In a letter of mine to Mr. *J. Collins,* dated *December* 10, 1672, having described a method of tangents, which I suspected to be the same with *Sluse's* method, which at that time was not made public, I added these words: *This is one particular, or rather a Corollary, of a general method, which extends itself, without any troublesome calculation, not only to the drawing of tangents to any curved lines, whether geometrical or mechan-*

[1 Appendix, Note 32.] [2 Appendix, Note 33.]

If a, b are infinitesimal changes in the sides of a rectangle with sides A, B, then the *moment of mutation* (infinitesimal change) of the generated rectangle AB is the quantity $aB + bA$. In more familiar language,

$$(A+a)(B+b) - AB \approx aB + bA$$

where Newton ignores the double-infinitesimal quantity ab, exactly as did Barrow and others before him.

Power Law In the same pages, Newton asserts the general power law for rational exponents

> …the moment of any power $A^{\frac{n}{m}}$ will be $\frac{n}{m}aA^{\frac{n-m}{m}}$.

in what seems like a non-rigorous appeal to induction based on the product rule. In reality, Newton established this using infinitesimal arguments; we'll see one of his methods for this shortly.

In contrast to the mostly synthetic presentation in the *Principia,* Newton made great use of algebra in his private calculations and correspondence with friends. Some of these private works were published many years later. We discuss some of his methods in what follows.

Fluxions and Fluents Newton's main language for calculus (he had several!) referred to time-dependent quantities x, y as *fluents* and their derivatives as *fluxions,* denoted using dots $\dot{x}, \dot{y}$; the modern notation y' comes from this.[65] Anti-derivatives were denoted by placing an accent directly over a quantity: $\acute{x}$ is a *fluent of which x is the fluxion.* Newton had several algorithms for computing fluxions, often variants of those of previous mathematicians: for instance, here he finds the relationship between the fluxions of fluents x, y satisfying $x^2 + 3xy^3 + y = 5$.

- Rearrange as a polynomial in x: thus $x^2 + (3y^3)x + (y - 5) = 0$.
- Multiply terms by a decreasing arithmetic sequence (e.g. $2, 1, 0$) and the entire expression by $\frac{\dot{x}}{x}$:

$$(2x + 3y^3)\dot{x}$$

- Repeat for y, using the *same arithmetic sequence*; in this example, 2 corresponds to x^2, so we start with 3 for y^3:

$$(3x)y^3 + 0y^2 + y + (x^2 - 5) = 0 \rightsquigarrow (9xy^2 + y)\dot{y}$$

- Sum these expressions, set equal to zero and rearrange for the required ratio:

$$(2x + 3y^2)\dot{x} + (9xy^2 + y)\dot{y} = 0 \implies \frac{\dot{y}}{\dot{x}} = -\frac{2x + 3y^2}{9xy^2 + y}$$

The arithmetic sequence encodes the power law for derivatives, and the result is exactly what you'd expect from modern implicit differentiation.

The Binomial Series Also discovered during the plague years, but first appearing in a private letter of 1676, is Newton's discovery of the binomial series

$$(1 + x)^\alpha = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha - 1)(\alpha - 2) \cdots (\alpha - k + 1)}{k!} x^k$$

which allowed Newton to expand expressions such as

$$(1 + x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \cdots$$

Newton's version was only for fractional exponents and is more difficult to read:

$$P + PQ\,\frac{m}{n} = P\frac{m}{n} + \frac{m}{n}AQ + \frac{m - n}{2n}BQ + \frac{m - 2n}{3n}CQ + \frac{m - 3n}{4n}DQ + \cdots \qquad (*)$$

Newton wrote exponents using juxtaposition, and $A, B, C, D, \ldots$, meant 'the previous term': thus $P\frac{m}{n} = P^{m/n}$, $A = P^{m/n}$ and $B = \frac{m}{n}AQ = \frac{m}{n}P^{m/n}Q$. In more modern language, $(*)$ reads

$$(P + PQ)^{m/n} = P^{m/n} + \frac{m}{n}P^{m/n}Q + \frac{m(m - n)}{2n^2}P^{m/n}Q^2 + \frac{m(m - n)(m - 2n)}{6n^3}P^{m/n}Q^3 + \cdots$$

---

[65] A fluent is a 'flowing' quantity: to us, a *smooth function,* though this was not formally defined. To be in *flux* is to be changing, hence 'rate of change.' Newton's dot-notation ('pricked letters') persists in the modern field of *dynamics*: for instance $\ddot{\mathbf{r}} = -\frac{GM}{r^3}\mathbf{r}$ is the differential equation arising from Newton's second law together with the inverse-square law for gravitation (note the double dot for the second derivative).

Newton's 'proof' wouldn't pass modern muster. His discoveries were largely the result of some inspired pattern-spotting! Several examples were explicitly verified by multiplying out or using long-division. For instance the series for $\sqrt{1+x}$ may be obtained

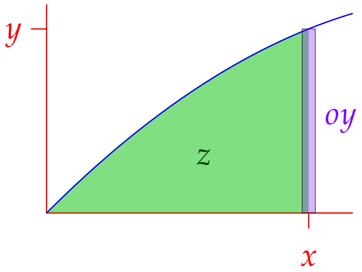
$$\begin{aligned}
&1+x=(1+ax+bx^2+cx^3+\cdots)^2=1+2ax+(a^2+2b)x^2+(2c+2ab)x^3+\cdots\\
\implies &a=\frac{1}{2}, \quad b=-\frac{1}{2}a^2=-\frac{1}{8}, \quad c=-ab=\frac{1}{16}, \quad \text{etc.}\\
\implies &\sqrt{1+x}=1+\frac{1}{2}x-\frac{1}{8}x^2+\frac{1}{16}x^3+\cdots
\end{aligned}$$

Newton did not work through the full theory of infinite series as you would encounter in a typical undergraduate analysis course. For instance:

- $\sqrt{1+x}$ is well-defined for $x \geq -1$, yet the resulting series only converges when $-1 \leq x \leq 1$. What happens in general: must a series converge to the original/generating function?

- Is it legitimate to differentiate and integrate power series term-by-term as if they are polynomials? Answer: (mostly) yes, though it was 150 years before Cauchy, Weierstraß and others could rigorously confirm this.

The Power Law & the Fundamental Theorem Newton's ability to expand expressions as infinite series was essential to his arguments. Here is one of his arguments to prove the power law.



1. Assume that the area under a curve y is given by a function

   $$z=\frac{x^{\frac{m}{n}+1}}{\frac{m}{n}+1}=\frac{n}{m+n}x^{\frac{m+n}{n}}$$

2. If x is increased by an infinitesimal amount o, then the new area under the curve is found using the binomial series

   $$\begin{aligned}
   z+oy&=\frac{n}{m+n}(x+o)^{\frac{m+n}{n}}=\frac{n}{m+n}x^{\frac{m+n}{n}}\left(1+\frac{o}{x}\right)^{\frac{m+n}{n}}\\
   &=\frac{n}{m+n}x^{\frac{m+n}{n}}\left(1+\frac{m+n}{n}\cdot\frac{o}{x}+\frac{(m+n)m}{2n^2}\cdot\frac{o^2}{x^2}+\cdots\right)
   \end{aligned}$$

3. Following Barrow, Newton cancels the terms in the original equation, divides by o, and throws out all remaining o-terms. The result is

   $$y=\frac{n}{m+n}x^{\frac{m+n}{n}}\frac{m+n}{n}\cdot\frac{1}{x}=x^{m/n}$$

   Note the link-up with the fundamental theorem, which is intuitively obvious when y is a 'flowing' (continuous) quantity: the additional area is approximately an infinitesimal rectangle $oy=\mathrm{d}z$ with base $o=\mathrm{d}x$ and height y, whence

   $$\mathrm{d}z=y\,\mathrm{d}x \rightsquigarrow \frac{\mathrm{d}z}{\mathrm{d}x}=\frac{\mathrm{d}}{\mathrm{d}x}\int^x y(t)\,\mathrm{d}t=y(x)$$

Using similar approaches, Newton produced one of the first tables of integrals, listing much of what you'd find inside the covers of an undergraduate calculus textbook! He also obtained power series for trigonometric and logarithmic functions, partly following the work of Gregory and others. By combining his approach with the power law, he was able to efficiently integrate and differentiate an enormous variety of functions.

Gottfried Wilhelm Leibniz (1646–1716)

Leibniz hailed from Leipzig, southwest of Berlin, in what was then part of the Holy Roman Empire. His initial studies were in philosophy, following his professor father. Though he eventually became a diplomat and then counsellor to the Duke of Hanover, his taste for advanced mathematics was fueled during his 1672–76 sojourn in Paris, where he was introduced to van Shooten's expansion of Descartes' geometric ideas, in particular the concept of the *differential triangle* (page 95) which was already in use by others such as Pascal and Barrow.

The familiar notations for derivatives $\frac{\mathrm{d}y}{\mathrm{d}x}$ and integrals $\int y\,\mathrm{d}x$ come from Leibniz. Very loosely, here is its origin and how it relates to the fundamental theorem. Suppose that

$$(x_0, z_0), (x_1, z_1), \ldots, (x_n, z_n), \qquad x_0 < x_1 < \cdots < x_n$$

describe a sequence of points along a curve $z(x)$ defined on an interval $[x_0, x_n]$. One may then form the sequences of *differences* and *sums* of the ordinates z_i:

$$(\delta z_i) = (z_1 - z_0,\ z_2 - z_1,\ \ldots,\ z_n - z_{n-1}), \qquad \left(\sum z_i\right) = (z_0,\ z_0 + z_1,\ \ldots,\ z_0 + \cdots + z_n)$$

Two relationships between sums and differences are immediate:

1. The difference sequence of the sums returns the original sequence:

$$\left(\delta \sum z_i\right) = (z_0, z_1, \ldots, z_n)$$

2. The sum of the difference sequence is the net change in the ordinate:

$$\sum(\delta z_i) = (z_1 - z_0) + (z_2 - z_1) + \cdots + (z_n - z_{n-1}) = z_n - z_0$$

Leibniz's notation arises from viewing a curve as an *infinite sequence* of points: he writes $\mathrm{d}z$ for the infinitesimal *differences* and $\int$ for the *sum* of infinitely many infinitesimal objects. Observation 2 becomes $\int \mathrm{d}z = z(x_n) - z(x_0)$: the sum of the infinitesimal changes in z is its net change. Indeed, if we assume that $z(x)$ describes the *area*[66] under a curve $y(x)$, the fact that $\mathrm{d}z = y\,\mathrm{d}x$ recovers part 2 of the fundamental theorem of calculus:

2. $\int y\,\mathrm{d}x = z$: the area under the curve is the sum of its infinitesimal increments.

Observation 1 makes sense once we apply it to a 'sequence' of infinitesimals $z_i \rightsquigarrow y\,\mathrm{d}x$, whence it becomes part 1 of the fundamental theorem:

1. $\mathrm{d}\left(\int y\,\mathrm{d}x\right) = y\,\mathrm{d}x$, or alternatively, $\frac{\mathrm{d}}{\mathrm{d}x}\left(\int y\,\mathrm{d}x\right) = y$: the rate of change of the area function is the ordinate.

---

[66] Notation is chosen to match Newton's discussion of the power law (page 99).

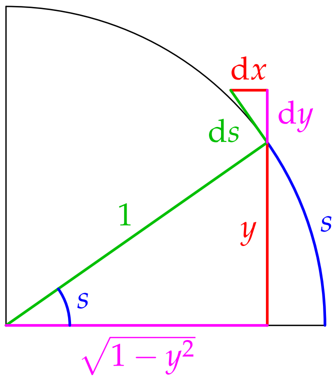
While we are happy to refer to infinitesimals and infinite sums, Leibniz was more cagey in his publications out of fear of criticism. He referred to each $\mathrm{d}x$ as an arbitrary (if small) *finite* line segment, and therefore—like every other contemporary practitioner of calculus—fails to get to grips with the essential paradoxes involved. Regardless, he and his followers became adept at manipulating differential expressions. For instance, if $y = z^3 + 2z$, Leibniz might compute

$$\begin{aligned} \mathrm{d}y &= (z+\mathrm{d}z)^3 + 2(z+\mathrm{d}z) - z^3 - 2z = 3z^2\,\mathrm{d}z + 3z(\mathrm{d}z)^2 + (\mathrm{d}z)^3 + 2\,\mathrm{d}z \\ &= (3z^2+2z)\,\mathrm{d}z \end{aligned}$$

where the $(\mathrm{d}z)^2$ and $(\mathrm{d}z)^3$ terms are discarded due to their (relatively) infinitesimal size. Using such approaches Leibniz justified general formulas such as the linearity of derivatives, the product, quotient and power rules. Even the chain rule is easy in Leibniz's notation: given $y(x) = \sqrt{1+x^3}$, Leibniz might perform a substitution $u = 1+x^3$ and observe that

$$\mathrm{d}y = \mathrm{d}\sqrt{u} = \sqrt{u+\mathrm{d}u} - \sqrt{u} = \frac{u+\mathrm{d}u-u}{\sqrt{u+\mathrm{d}u}+\sqrt{u}} = \frac{\mathrm{d}u}{2\sqrt{u}} = \frac{3x^2\,\mathrm{d}x}{2\sqrt{1+x^3}}$$

Like Newton, Leibniz worked extensively with power series. As a final example here is his computation of that for sine. In contrast to Newton's approach,[67] Leibniz derived the well-known second-order ODE satisfied by sine.



In a circle of radius 1, let s be the polar angle and $y = \sin s$ the ordinate. By considering the differential triangle, we see that

$$\frac{\mathrm{d}y}{\mathrm{d}s} = \sqrt{1-y^2} \implies (1-y^2)(\mathrm{d}s)^2 = (\mathrm{d}y)^2 \qquad (*)$$

Leibniz supposes that infinitesimals $\mathrm{d}s$ describing the arc are constant and applies d again (with the product rule):[68]

$$-2y\,(\mathrm{d}y)(\mathrm{d}s)^2 = 2(\mathrm{d}y)\mathrm{d}(\mathrm{d}y) \implies \frac{\mathrm{d}(\mathrm{d}y)}{(\mathrm{d}s)^2} = -y \qquad \left(= \frac{\mathrm{d}^2y}{\mathrm{d}s^2}\right)$$

He then writes $y(s) = \sin s = c_0 + c_1 s + c_2 s^2 + c_3 s^3 \cdots$ as a power series. Since at $s = 0$, $y = \sin s = 0$ and $\mathrm{d}s = \mathrm{d}y$ (by $*$), Leibniz sees that $c_0 = 0$ and $c_1 = 1$. He now differentiates twice and equates coefficients:

$$\begin{aligned} &\frac{\mathrm{d}(\mathrm{d}y)}{(\mathrm{d}s)^2} = 2c_2 + 6c_3 s + 12c_4 s^2 + 20c_5 s^3 + 30c_6 s^4 + \cdots = -y = -s - c_2 s^2 - c_3 s^3 - c_4^4 - \cdots \\ &\implies 0 = c_2 = c_4 = c_6 = \cdots, \quad c_3 = -\frac{1}{6}, \quad c_5 = -\frac{1}{20}c_3 = \frac{1}{120}, \ \ldots \end{aligned}$$

to obtain the familiar series

$$y = \sin s = s - \frac{1}{6}s^3 + \frac{1}{120}s^5 + \cdots = s - \frac{1}{3!}s^3 + \frac{1}{5!}s^5 - \frac{1}{7!}s^7 + \cdots$$

While not precisely the same as modern calculus—in particular, the differentials $\mathrm{d}x, \mathrm{d}y, \mathrm{d}s$ are separate quantities—this should all seem very familiar! The computational efficiency of such notation super-charged mathematics and its applicability to real-world problems.

---

[67] Newton first found a series for arcsine before computing its inverse.

[68] We've bracketed everything. When these are removed, we obtain the standard Leibniz notation for second derivatives!

Exercises 8.3. 1. Show that to find the length of the arc of the parabola $y = x^2$ one needs to determine the area under the hyperbola $y^2 - 4x^2 = 1$.

2. Use Barrow's a, e method to determine the slope of the tangent line to the curve $x^3 + y^3 = c^3$.

3. Calculate a power series for $\frac{1}{1-x^2}$ by using long-division.

4. Use the binomial series to obtain a power series expression for $\ln(1+x)$, which Newton knew to describe the area under the curve $\frac{1}{1+x}$.

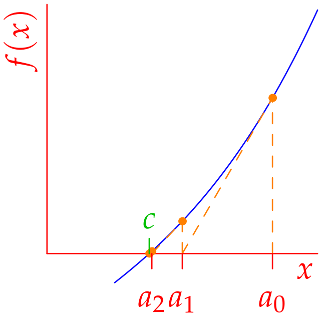
5. Using his calculus, Newton was able to extend older methods of approximation.

   - Suppose $f(c) = 0$ and that a_0 is an initial approximation to c.
   - The tangent line at $\big(a_0, f(a_0)\big)$ has equation

     $$y = f(a_0) + f'(a_0)(x - a_0)$$

     which intersects the x-axis at $a_1 := a_0 - \frac{f(a_0)}{f'(a_0)}$.
   - Iterate to obtain a sequence (a_n) that (typically) converges to c:

     $$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)} \qquad \lim_{n\to\infty} a_n = c$$



   (a) If $f(x) = x^2 - c$, show that Newton's method is the Babylonian method of the mean.

   (b) Use Newton's method to solve the equation $x^2 - 2 = 0$ to a result accurate to eight decimal places. How many steps does this take?

6. Calculate the relationship of the fluxions in the equation $x^3 - ax^2 + axy - y^3 = 0$ using multiplication by the progression 4, 3, 2, 1. Compare to what happens if you use the progression 3, 2, 1, 0. What do you notice?

7. Use Leibniz's differential triangle to argue that

   $$x = \cos s \quad \text{and} \quad \mathrm{d}x = -y\,\mathrm{d}s$$

   Where does the negative sign come from? Hence find the standard power series representation for cosine and conclude that the rate of change (derivative) of sine is cosine.

8. Given a curve $y(x)$ with $y(0) = 0$, Leibniz's *transmutation theorem* relates the area under y and the curve $z(x) = y - x\frac{\mathrm{d}y}{\mathrm{d}x}$ obtained by considering the y-intercepts of the tangent lines to the original:

   $$\int_0^a y\,\mathrm{d}x = \frac{1}{2}\left[ay(a) + \int_0^a z\,\mathrm{d}x\right]$$

   If $x^p = y^q$, find z and use the transmutation theorem to find the area $\int_0^a y\,\mathrm{d}x$.

   How does the transmutation theorem relate to *integration by parts*?