8 Analytic Geometry and Calculus

8.1 Axes and Co-ordinates

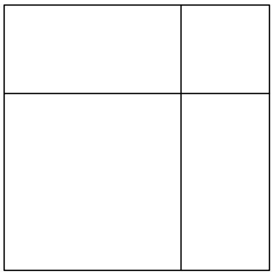
Modern mathematics is dominated by *algebra*. We trust equations and the rules of algebra more than pictorial, geometric reasoning. For example, when modern mathematicians are likely to interpret the expression $(x+y)^2 = x^2 + 2xy + y^2$ within the laws (axioms) of algebra:

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y) && \text{(definition of 'square')}\\ &= x(x+y) + y(x+y) && \text{(distributive law)}\\ &= x^2 + xy + yx + y^2 && \text{(distributive law twice more)}\\ &= x^2 + 2xy + y^2 && \text{(commutativity)}\end{aligned}$$

For most of mathematical history, however, this result would have been stated *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 the purposes of *proof*; practitioners still felt the need to justify calculations geometrically. In modern times this seems backwards: if a college student today were asked to prove Euclid's result,[59] they'd likely start by labeling the 'parts' x and y, before appealing to the algebraic formula at the top of the page! Of course, each piece of algebra has a geometric basis:

- Distributivity is the fact that a rectangle on a side and two parts equals the sum of the rectangles on the side and each of the parts respectively.
- Commutativity is the fact that a rectangle has the same area after a 90° rotation.

Modern mathematics has converted geometric rules into algebra and largely forgotten their origin. The slow triumph of algebra over geometry is one of the great revolutions of mathematical history, completely changing the way mathematicians *think.* More practically, algebra permits easy generalization: how would one geometrically justify an expression such as

$$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4\ ?$$

Euclidean Geometry is *synthetic:* based on purely geometric axioms without formulæ or co-ordinates. *Analytic geometry* married algebra, geometry, and number, by introducing *axes* and *co-ordinates.* Modern geometry is primarily analytic and it is now rare to find a mathematician working synthetically—algebra's domination of Euclidean geometry is total. The critical contribution to this revolution was made almost simultaneously by two Frenchmen...

---

[59] But importantly *not* a 5th-grader, who'd likely consider algebra to be alien and pointless: *isn't the picture obvious*? Keep this in mind if you ever have the privilege of teaching at grade-school: it is *your* ideas that are strange...

Pierre de Fermat (1601–1665)

Fermat made great strides in several areas (probability, analytic geometry, early calculus, number theory, optics), even though mathematics was a pastime rather than his profession.[60] You've likely encountered his name attached to two famous results in number theory:

Fermat's Little Theorem If p is prime and x an integer, then $x^p \equiv x \pmod{p}$.

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 ($n = 4$). This problem has captivated mathematicians ever since, motivating several developments in algebra and number theory; a complete proof finally came in the 1990s.

Some of Fermat's fame comes from his enigma. Much of what we know comes from letters to friends in which he claims much but rarely offers complete proofs—it is often unclear whether he merely suspected a general statement. When Fermat died, his notes and letters contained many unproven claims. Leonhard Euler (1707–83) in particular expended much effort on these.

Fermat's work on analytic geometry was largely developed by 1629, though it wasn't published during his lifetime. He approach was similar to that of Descartes which we describe below. The two corresponded, not always cordially, but both acknowledged the other's contributions. We'll return to Fermat when we discuss the beginnings of calculus in the next section.

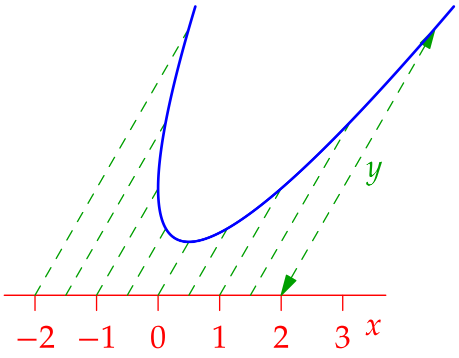
René Descartes (1596–1650)

The chalk to Fermat's cheese, Descartes rigorously recorded everything. His defining work is 1637's *Discours de la méthode*...[61] 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 his fear of suffering the same persecution as Galileo.[62] 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.



---

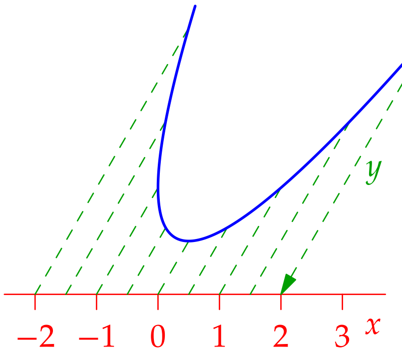
[60] Fermat was wealthy but not aristocratic, attending the University of Orléans for three years where he trained as a lawyer.

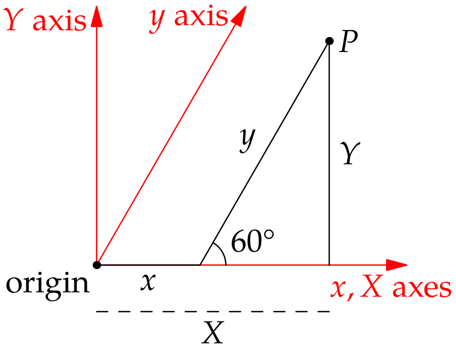
[61] *...of rightly conducting one's reason and of seeking truth in the sciences.* The primary part of this work is philosophical and contains his famous phrase *cogito egro sum* (*I think therefore I am*).

[62] 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.

While neither Descartes nor Fermat had a second axis, their approach implicitly imagines one, namely 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 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.

Example 1 The previous picture shows some of the flexibility inherent in Descartes' approach. The curve $y = x^2 + 1$ is drawn, where the 'y-axis' is inclined 60° to the horizontal. To represent the curve in a more standard fashion, we perform a change of co-ordinates. Suppose a point P has co-ordinates (x, y) with respect to the original slanted axes and (X, Y) with respect to the familiar orthogonal axes of van Schooten.





The second picture shows the relationship between the two co-ordinate systems

$$\begin{cases} X = x + y\cos 60^\circ = x + \frac{1}{2}y \\ Y = y\sin 60^\circ = \frac{\sqrt{3}}{2}y \end{cases}$$

A little algebra quickly shows that $\sqrt{3}X - Y = \sqrt{3}x$. It follows that for any point on the curve, we have

$$(\sqrt{3}X - Y)^2 = 3x^2 = 3(y-1) = 3\left(\frac{2}{\sqrt{3}}Y - 1\right)$$
$$\implies 3X^2 - 2\sqrt{3}XY + Y^2 - 2\sqrt{3}Y + 3 = 0$$

This is an implicit equation for a parabola in rectangular co-ordinates. Descartes has essentially chosen a second axis in such a way as to make the equation of this parabola very simple.[63]

---

[63] That the implicit equation is a parabola may be confirmed in modern language by computing its *discriminant*: a non-degenerate quadratic curve $aX^2 + 2bXY + cY^2 + \cdots = 0$ is a parabola if and only if $b^2 - ac = 0$. You've possibly encountered this in a linear algebra course, where the standard approach is to *rotate* both axes instead by diagonalizing a symmetric matrix: indeed the discriminant is negative the determinant of the matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ and $(x\ y)A\binom{x}{y} = aX^2 + 2bXY + cY^2 \ldots$

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.*

> The three medians[64] of a triangle meet at a common point (the *centroid*), which moreover lies two-thirds of the way along each median.

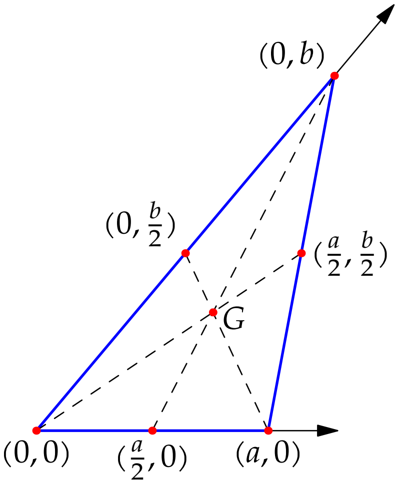
Choose axes pointing along two sides of a triangle with the origin at one vertex. If the two sides lying on the axes have lengths a and b, then the third side necessarily has equation

$$bx + ay = ab$$

Moreover, the midpoints of the three sides plainly have co-ordinates as pictured.

To complete the proof, simply compute the co-ordinates of the point two-thirds of the way along each median: for the median through the upper vertex, this is

$$\frac{2}{3}\left(\frac{a}{2},0\right)+\frac{1}{3}(0,b)=\frac{1}{3}(a,b)$$



One obtains the same co-ordinates for the other two medians, whence *all three meet at a common point G*, the *centroid* of the triangle.

As demonstrated by the centroid proof, Descartes' 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.

Descartes found that his method could solve many problems in a more efficient manner than using synthetically; in particular, finding complicated intersections of curves. As we'll see in the next section, such arguments were necessary for the development of calculus, ans would often rely on his Factor Theorem (pg. 93).

Given the novelty of his approach, Descartes typically gave geometric proofs to back up his algebraic work. However, he also saw the future: once several examples were done, he claimed that it was no longer necessary to draw physical lines and provide a geometric argument; *the algebra was the proof.*

Exercises 8.1. *Key concepts: Basic algebra encodes geometric facts, Analytic geometry helps algebra & number eclipse geometry, Fermat & Descartes' single-axis approach*

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 $b^2 - 2x^2 = 2xy + y^2$.

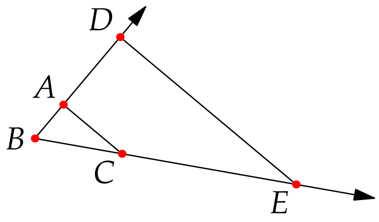
   (*Hint: add* x^2 *to both sides and remember that axes do not have to be orthogonal...*)

---

[64] A segment joining a vertex and the midpoint of the opposite side.

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 $E \in \overrightarrow{BC}$.

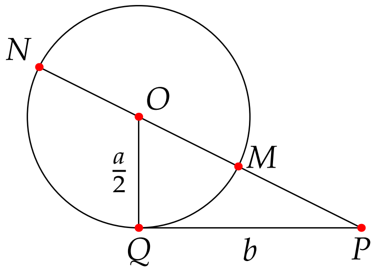
   

   Prove that $|BE|$ is the product of $|BD|$ and $|BC|$.

   Similarly, given lengths $|BE|$ and $|BD|$, explain how to 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 $x^3 + bx^2 = bc$ $(b, c > 0)$, Fermat would introduce a new variable y by setting both sides of the equation equal to bxy. The (positive) solution to the cubic is the x-solution to the system of equations

   $$\begin{cases} x^2 + bx = by \\ xy = c \end{cases}$$

   Sketch these curves explicitly for the cubic $x^3 + 4x^2 = 24$. What is the solution?

8.2 The Beginnings of Calculus

The the word *calculus* (*small stone* in Latin) relates to the use of pebbles and tokens for counting (calculating) in ancient times (consider, for instance, the abacus). In the 1600s, the modern theory became known as the *infinitesimal calculus*: the method of calculating using infinitesimals. At its heart 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.* Otherwise said, the *slope* of the displacement-time graph.

- The *displacement* of a particle is the *net area* under its *velocity-time* graph.

These relationships are essentially geometric. As we've seen, the advent of analytic geometry in the 1620-30s facilitated easier geometric computations. With this powerful tool available, the rapid development of calculus was arguably inevitable.

In the context of the above, the *Fundamental Theorem of Calculus* is intuitive: complete knowledge of displacement (from some initial point) is equivalent to complete knowledge of velocity. A modern statement is more daunting.

Theorem (Fundamental Theorem of Calculus).

1. *If f is integrable on $[a,b]$, then $F(x) := \int_a^x f(u)\,\mathrm{d}u$ is continuous on $[a,b]$. Moreover, if f is continuous at $c \in (a,b)$, then $F'(c) = f(c)$.*

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 modern version is abstract and very widely applicable: we've gone way beyond considerations of velocity (f) and displacement (F). The challenge of *teaching*[65] 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, however, with some older considerations of the velocity and area problems.

The Velocity Problem pre-1600

The modern ideas of *uniform* or *average* velocity are straightforward:

> Measure displacement over a given time interval and divide one by the other.

Several ancient Greek mathematicians considered versions of this, and of uniform acceleration, the challenge of considering a ratio of two unlike quantities (displacement : time) proved difficult to surmount. Around 1200, Gerard of Brussels resurrected the ratio approach as a definition of velocity, though it was not considered a numerical quantity in its own right.

---

[65] 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!

Gerard was credited by the Oxford/Merton Thinkers (1330s) as influencing their investigations of *instantaneous velocity,* a much more difficult issue. They offered the following definition.

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.

While vague, this teases at the idea of inertial motion that Newton would later formalize in his first law of motion. The Merton Thinkers also made the first known statement of the 'mean speed theorem.'

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.*

For centuries, Galileo was thought to have been the first to state such ideas (compare his falling body discussion, pg. 100). The Oxford group had no algebra to back up their claims and essentially only asserted examples. For instance, if bodies subject to the same uniform acceleration are accelerated from rest over time intervals in the ratio 2 : 1, then the distances traveled will be in the ratio 4 : 1. In modern notation:

$$v_1 = at, \quad v_2 = 2at \implies d_1 = \frac{1}{2}v_1 t = \frac{1}{2}at^2, \quad d_2 = \frac{1}{2}v_2 \cdot 2t = 2at^2 = 4d_1$$

In the 1350s, Nicolas Oresme (Paris) considered velocity geometrically by drawing what are essentially velocity-time graphs. This is roughly the approach taken later by Galileo, and another 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.

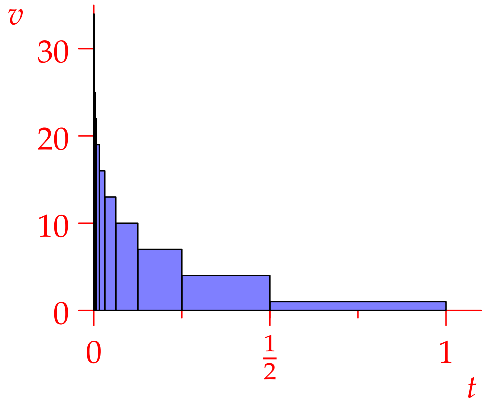
A rigorous definition of *instantaneous velocity* is difficult because it requires *limits*: one measures average velocity over smaller and smaller intervals. You are in good company if you find this challenging: Zeno's arrow paradox is an objection to the idea of instantaneous velocity! Even if one accepts the concept, its *direct* measurement is nigh-impossible: for instance, radar Doppler-shift (as used to catch speeding motorists) requires measuring the wavelength of a radar beam, which amounts to computing an average velocity over a very small time interval.

The Area Problem pre-1600

We've previously seen two situations in which calculus-like methods described areas.

- Archimedes (sec. 3.4) computed the area inside a parabola using infinitely many triangles in what amounts to the evaluation of an infinite series. His 'cross-section' approach to finding areas and volumes also feels modern, though this work was unknown until 1899.
- Kepler (pg. 98) argued for his second law (equal areas in equal times) using infinitesimally small triangles to approximate segments of an ellipse: this is more akin to integration. He also applied this method to other problems, crediting Archimedes with the approach.

The modern integral (via Riemann sums) is a special case of approximating 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 traveled 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. In modern language, the distance traveled is an infinite series

$$d = \sum_{n=0}^{\infty}(1+3n)\left(\frac{1}{2^n} - \frac{1}{2^{n+1}}\right) = \sum_{n=0}^{\infty}\frac{1+3n}{2^{n+1}} = 4$$

Similarly 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 acceleration over intervals. These are not true Riemann sums, nor are they physical, for a particle cannot suddenly change speed.

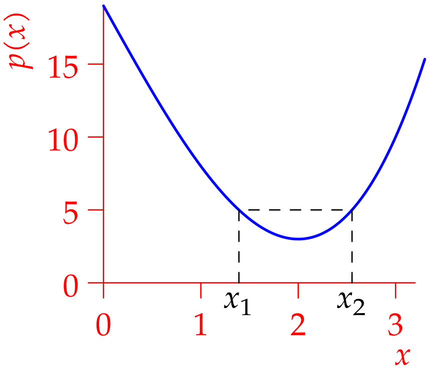
Calculus à la Fermat & Descartes

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 identified with the slope of the displacement-time graph, which may be computed using variations of the modern method of *secant lines.* We discuss their competing methods.

Fermat's method of adequation We illustrate with an example. Our goal is to find the local minimum (located at $m = 2$) of the cubic function $p(x) = x^3 - 12x + 19$.

Fermat argues that if x_1 and 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 algebraic identity holds for any x_1, x_2 with $p(x_1) = p(x_2)$, Fermat claims that it holds also when $x_1 = x_2 = m$: note the assumption of continuity and limit-taking! He concludes

$$3m^2 - 12 = 0 \implies m = 2$$

By considering values of x near m, it is clear to Fermat that he really has found a local minimum. We recognize the idea that the slope of the tangent line is zero at local extrema.

Fermat's approach dates from the 1620s and is similar to earlier work of Viète. He later alters his method by considering values $p(x)$ and $p(x+e)$ for small e (x is 'adequated by e'), with the difference more easily divided by e without factorizing. Compare with the previous calculation:

$$\begin{aligned} 0 &= \frac{p(x+e)-p(x)}{e} = \frac{x^3+3x^2e+3xe^2+e^3-12x-12e+19-x^3+12x-19}{e} \\ &= \frac{3x^2e+3xe^2+e^3-12e}{e} = 3x^2-12+3xe+e^2 \end{aligned}$$

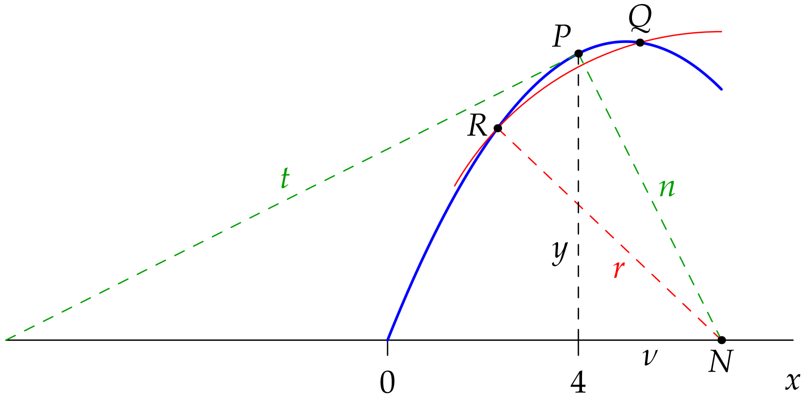
Finish by setting $e=0$ and solving for x as before. Observe the derivative $p'(x)=3x^2-12$ and how e plays the role of h in the modern definition

$$p'(x) = \lim_{h\to 0} \frac{p(x+h)-p(x)}{h}$$

Fermat's method works for any polynomial: in such cases, the limit definition of derivative requires no more than simple evaluation at $h=0$. He moreover 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 parabola $y=\frac{1}{4}(10x-x^2)$ at the point $P=(4,6)$.



Let $N=(4+\nu,0)$ be where the normal to the curve intersects the x-axis.[66] Draw a circular arc with radius r centered at N. If r is close to n, this 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 using algebra: 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,

$$\begin{cases} \big(x-(4+\nu)\big)^2+y^2=r^2 \\ y=\frac{1}{4}(10x-x^2) \end{cases} \implies (x-Q_x)(x-R_x)f(x)=0$$

---

[66] At the time, ν was known as the *subnormal* and t the *tangent*.

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 $Q = R = P$ and the above equation has a *double-root*:

$$\begin{cases}\left(x-(4+\nu)\right)^2+y^2=n^2\\ y=\frac{1}{4}(10x-x^2)\end{cases} \implies (x-P_x)^2f(x)=(x-4)^2f(x)=0$$

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

$$\frac{y}{\sqrt{t^2-y^2}}=\frac{\nu}{y}=\frac{1}{2}$$

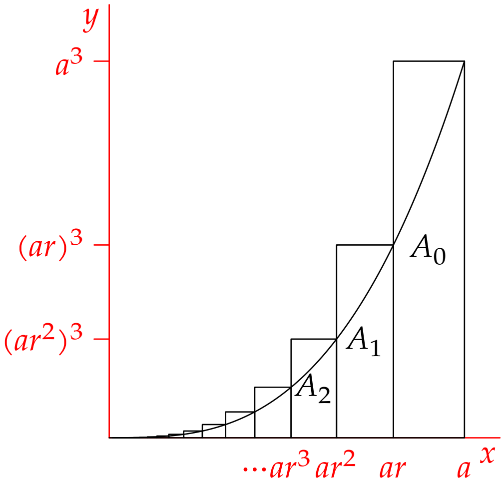
Fermat and Area The previous methods permit *differentiation*, albeit inefficiently. Fermat also approached the area problem in a manner similar to Oresme. Here is an example of this approach, where we find the area under the curve $y = x^3$ between $x = 0$ and $x = a$.

Let $0 < r < 1$ be constant. The rectangle on the interval $[ar^{k+1}, ar^k]$ touching the curve at its upper right-corner has area

$$A_k=(ar^k-ar^{k+1})\cdot(ar^k)^3=a^4(1-r)r^{4k}$$

The sum of the areas is therefore

$$\begin{aligned}\sum_{n=0}^{\infty}A_k=a^4(1-r)\sum_{n=0}^{\infty}r^{4k}&=\frac{a^4(1-r)}{1-r^4}\\ &=\frac{a^4}{1+r+r^2+r^3}\end{aligned}$$



Setting $r = 1$ recovers the area under the curve: $\frac{1}{4}a^4$.

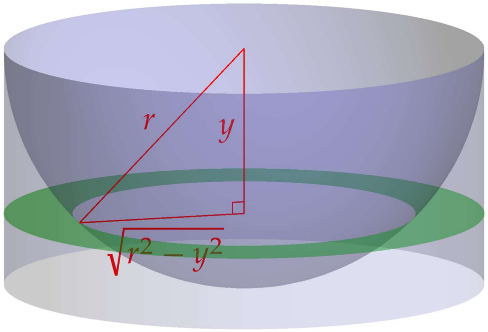
This is non-rigorous by modern standards and again implicitly invokes limits[67] by setting $r = 1$. By similar calculations, Fermat is able to establish the power law $\int_0^a x^n \, dx = \frac{1}{n+1}a^{n+1}$ for any positive integer n.

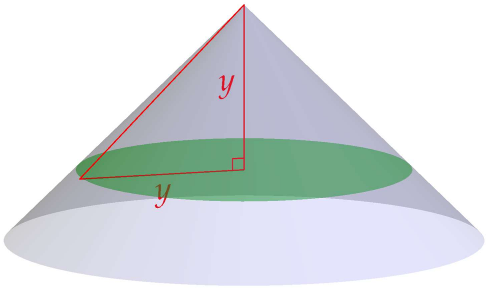
---

[67] 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.

Italian Calculus in the 17th Century: Area and Volume Problems

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[68] 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 those at the top of each solid assert 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 discussion and advocacy on these matters that first gained him notoriety with the Church. His later evangelism for the Copernican heliocentrism merely rekindled old animosities.

Bonaventura Cavalieri (1598–1647) & Indivisibles Cavalieri, a student of Galileo and a Jesuat scholar, gave a more thorough discussion in his 1635 text *Geometria Indivisibilis*.

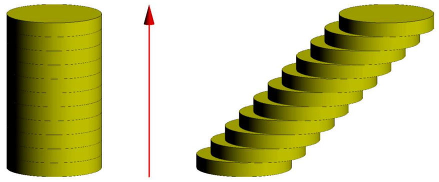
To Cavalieri, a solid (say Galileo's cone) consists of infinitely many, infinitely small, but *indivisible* pieces. Loosely speaking, indivisibles and infinitesimals are the same, though the former nods more towards the debated Greek idea of atomism: an indivisible cannot be further divided...

Cavalieri's principle remains a foundational idea in the application of calculus:

> 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.

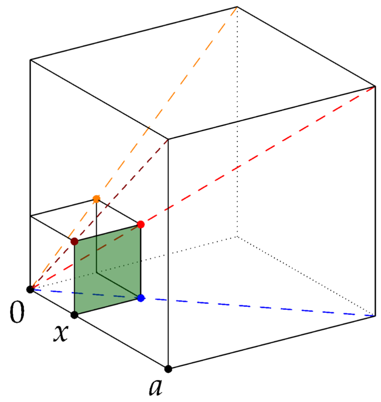
Another classic example involves sliding a stack of coins or a deck of cards; the volume of the slanted coin stack equals that of the cylinder.



---

[68] As did Archimedes 1900 years earlier (pg. 41), though Galileo was unaware of it.

Extending his principle, Cavalieri inferred the power law $\int_0^a x^n \mathrm{d}x = \frac{1}{n+1}a^{n+1}$, giving reasonable arguments for $n = 1$ and 2. Here is a sketch of his approach for $n = 2$.



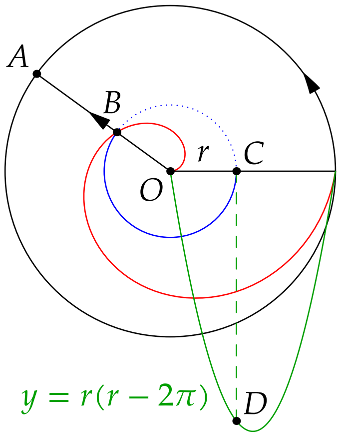
Draw a cube of side x inside a cube of side a. This consists of three congruent pyramids with base a^2 and eight a.

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

$$\int_0^a x^2 \,\mathrm{d}x = \frac{1}{3}a^3$$

Cavalieri also used his method (Book IV, Prop 19 of *Geometria Indivisibilis*) to calculate the area enclosed in an Archimedean spiral.
Suppose the line $\overline{OA}$ rotates at a constant speed about O, and that the point B moves at a constant speed (from O) along this line. In polar co-ordinates, B traces a curve with equation

$$r = k\theta, \qquad 0 \le \theta \le 2\pi$$



We take $k = 1$ for simplicity.
If $|OB| = r$, then the arc *inside* the spiral has length

$$\ell(r) = 2\pi r \cdot \frac{2\pi - \theta}{2\pi} = r(2\pi - r)$$

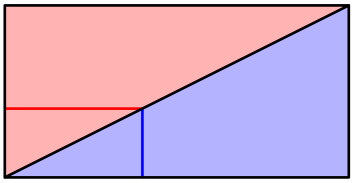
Imagine this arc as a noodle which, when cut at B and allowed to fall, forms the line $\overline{CD}$. The area within the spiral therefore equals that inside the parabola. Thanks to Archimedes (and Cavalieri himself), this is $\frac{4}{3}$ the area of the largest triangle that can fit inside. Otherwise said, the area inside the spiral is

$$\frac{4}{3} \cdot \frac{1}{2}(2\pi)\ell(\pi) = \frac{4}{3}\pi^3$$

Cavalieri's calculation was slightly different, with his parabola being drawn in a rectangle and the difference subtracted from a triangle. The above is a little easier to visualize.

Unlike Galileo, Cavalieri did not court controversy. He was well-aware of the contentious nature of indivisibles and took great pains to distinguish 'all the lines/squares' of a figure from the figure itself. This sleight-of-hand and the dense, difficult, nature of *Geometria Indivisibilis* helped keep Cavalieri out of trouble. But challenges certainly arose: his political rivals (the Jesuit order) worked hard to stamp out the 'dangerous' study of indivisibles. Cavalieri's later work tried to refute some of these challenges, not least the criticism that he lacked a rigorous definition of *indivisible*.

Evangelista Torricelli (1608–1647) A contemporary of Galileo and Cavalieri, Torricelli made great use of Cavalieri's principle and advocated for its careful application. Here is an example.

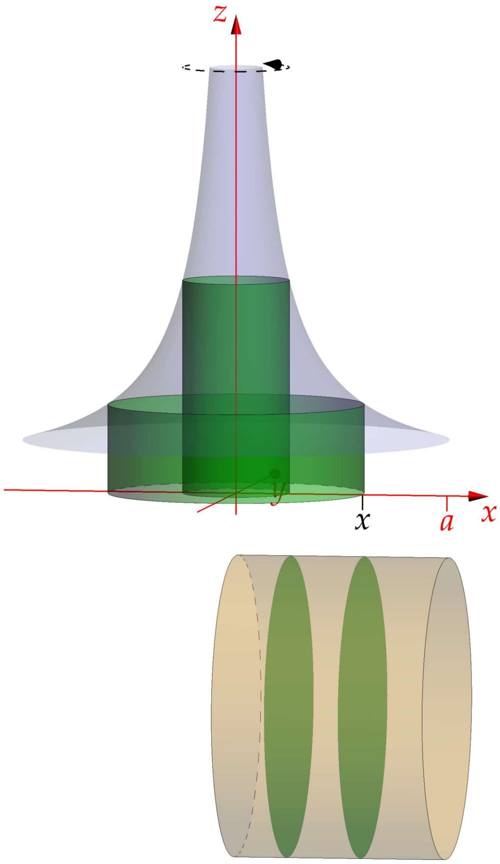


Suppose a rectangle has sides in the ratio 2 : 1, and split it diagonally into two congruent triangles (of necessarily equal area).

The indicated segments are also in the ratio 2 : 1. In Cavalieri's language, 'the lines' of the upper triangle are twice 'the lines' of the lower. So is the area of the upper triangle twice that of the lower?! Obviously not, since the triangles are congruent. Torricelli observes Cavalieri's principle has been misapplied: the cross-sectional lines were not measured with reference to a common line. In modern language, we could view the diagonal as the line $x = 2y$ and observe that this substitution result in and equality of integrals

$$\int_0^2 \frac{1}{2}x\,\mathrm{d}x = \int_0^1 2y\,\mathrm{d}y$$

The issue 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 is it possible for such to lie above a *finite* volume? Examples like this abound in calculus, such as $\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = 2$: a finite area underneath an unbounded curve. This isn't trickery, but it does illustrate how easy it is to work with the infinite and the infinitesimal *without* stopping to ask if their *meaning* are well understood...

Scientific Power Moves Northwards

Galileo, Cavalieri and Torricelli mark the end of 400 years of Italian dominance in science and mathematics, dating back to Fibonacci. Their ideas were too controversial to be allowed to thrive so close to Rome at a time of great threat to the Church's power.

The various protestant reformations and wars of the 15-1600s brought great changes to northern Europe. While deeply destructive at the time, they also proved to be a catalyst for new ideas, leading directly to the Enlightenment period of the 16-1700s. Ideas of reformed government and individual liberty[69] helped to unleash greater productivity and scientific output. The net result was that northern Europe became a more fertile place for philosophical and scientific development.

Exercises 8.2. *Key concepts: The Area and Velocity problems as motivators of calculus, Analytic geometry provides a key computational tool, Differentiation à la Descartes & Fermat, Cross-sections & Cavalieri's Principle*

1. Find the maximum of $p(x) = 5 + x - 2x^2$ using Fermat's first method.
2. Use Fermat's second "adequation" $(+e)$ method to find the local maximum of the function $bx - x^3$, where $b > 0$. How might Fermat decide that the positive solution to this problem really does provide a maximum?
3. Consider Fermat's first method of determining maxima and minima. Justify the method by showing that if $p(a) = M$ is a local maximum for $p(x)$, then the polynomial $p(x) - M$ is divisible by $(x - a)^2$.
4. Use Descartes' method of normals to compute the slope of the curve $y = x^2$ at the point (a, a^2).
5. Suppose that the surface of a sphere of radius r is subdivided into infinitesimal regions of equal (infinitesimal) area.

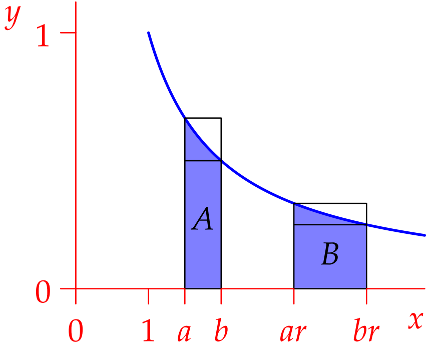
   Following Kepler, use the formula for the volume of a cone ($\frac{1}{3}$base·height) to find the relationship between the volume V of the sphere and its surface area A.
6. (A problem of Kepler) Show that the largest circular cylinder that can be inscribed in a sphere is one in which the ratio of diameter 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. Mimicking Fermat, establish the power law $\int_0^a x^n \, dx = \frac{1}{n+1}a^{n+1}$ for any positive integer n.

---

[69] For instance, Hobbes' *Leviathan*, written during the English Civil War (1642–51), is partly a plea to constrain absolute monarchical power. The War itself indeed proved decapitatingly effective at reining in a King...

8. 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 pictured, 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

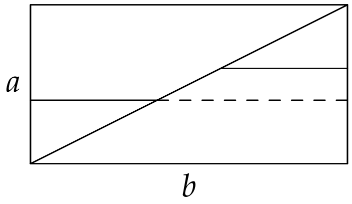
$$A(x_1 x_2) = A(x_1) + A(x_2)$$

   Why are you not surprised by this?

   (*For simplicity, assume* $1 < x_1 < x_1 x_2$)

9. Consider two copies of a 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,* $\int_0^a b \, \mathrm{d}y = 2 \int_0^a \frac{b}{a} y \, \mathrm{d}y$)



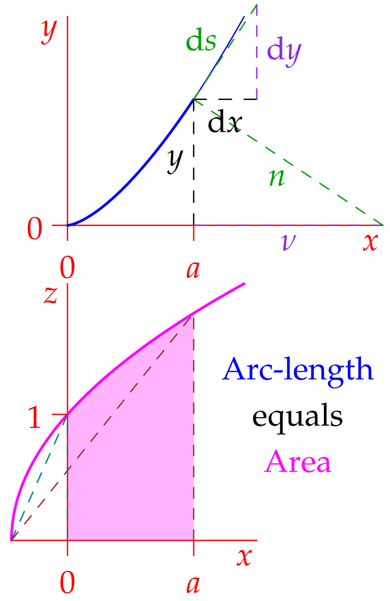
10. Repeat Cavalieri's analysis of the spiral (pg. 114) to find the area inside one revolution of the curve $r = k\theta$ for any $k > 0$.

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 the northern European work on calculus, culminating in the efforts of Newton and Leibniz. The overarching story is of how algebraic notation helped turn calculus from a clunky *geometric* activity into an efficient problem-solving tool.

Hendrick van Heuraet (1634–1660)

Working in Holland, van Heuraet studied Descartes' analytic geometry and method of normals.



He 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' of Descartes. His method appeared in Frans van Schooten's 1659 version of Descartes' *La Geometrie.*

To see why this claim makes sense, recall Descartes' method of normals (page 111) and observe that the ratio $n : y$ equals that of $\mathrm{d}s : \mathrm{d}x$ in a (modern) differential triangle.

This is easiest to see in an example: here we calculate the arc-length of the curve $y^2 = x^3$ in van Heureat's style. By Descartes' method of normals, we substitute the curve into the equation for the circle with radius n tangent to the curve:

$$\begin{cases}(x-a-v)^2 + y^2 = n^2 \\ y^2 = x^3\end{cases}$$

$$\begin{aligned}\implies 0 &= x^3 + x^2 - 2(x+v)x + (a+v)^2 - n^2 \\ &= (x-a)^2(x+2a+1) + (3a^2 - 2v)x + v^2 + 2av - 2a^3 - n^2\end{aligned}$$

Since this last must have a double-root at $x = a$, the remainder is zero and we conclude that

$$3a^2 - 2v = 0 \implies v = \frac{3}{2}a^2$$

But then

$$n = \sqrt{v^2 + y^2} = \sqrt{\frac{9}{4}a^4 + a^3} \implies z = \frac{n}{y} = \sqrt{\frac{9}{4}a + 1}$$

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. Recalling Archimedes, van Heuraet knew that the area under a parabola is $\frac{4}{3}$ that of the corresponding triangle, whence

$$\text{Arc-length} = \frac{4}{3}\left[\frac{1}{2}\left(\frac{4}{9} + a\right)z(a) - \frac{1}{2}\cdot\frac{4}{9}z(0)\right] = \left[\frac{4}{9} + a\right]^{3/2} - \frac{8}{27}$$

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 University, 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, we need to find y satisfying the *integral equation*

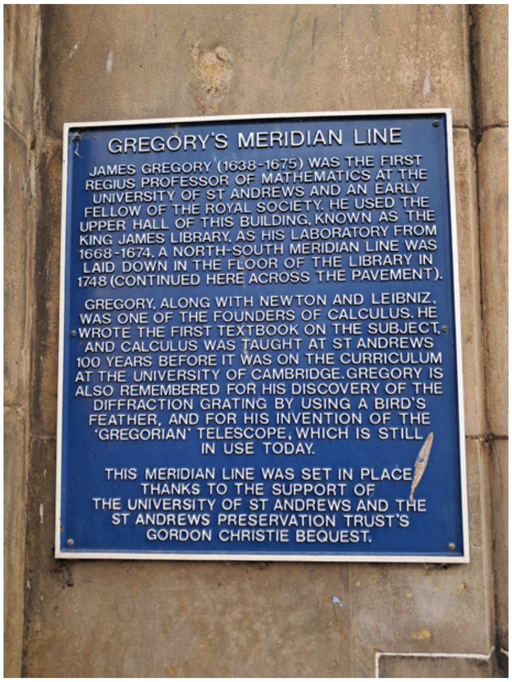
$$\int_0^a \sqrt{1+y'^2}\,\mathrm{d}x = \int_0^a z\,\mathrm{d}x$$

which is tantamount to 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 part 1 of the Fundamental Theorem: if you want something whose slope (derivative) is given, define it to be the area under the curve!

Obtaining *power series* expressions was a large part of the early development of calculus. Gregory's contributions here include his (re)discovery[70] of the arctangent series

$$\arctan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \cdots$$

Gregory was also a prolific astronomer. The picture shows part of a meridian line (zero-line of longitude) he had drawn through his office in St. Andrews to assist with his astronomical work



GREGORY'S MERIDIAN LINE

JAMES GREGORY (1638-1675) WAS THE FIRST REGIUS PROFESSOR OF MATHEMATICS AT THE UNIVERSITY OF ST ANDREWS AND AN EARLY FELLOW OF THE ROYAL SOCIETY. HE USED THE UPPER HALL OF THIS BUILDING, KNOWN AS THE KING JAMES LIBRARY, AS HIS LABORATORY FROM 1668-1674. A NORTH-SOUTH MERIDIAN LINE WAS LAID DOWN IN THE FLOOR OF THE LIBRARY IN 1748 (CONTINUED HERE ACROSS THE PAVEMENT).

GREGORY, ALONG WITH NEWTON AND LEIBNIZ, WAS ONE OF THE FOUNDERS OF CALCULUS. HE WROTE THE FIRST TEXTBOOK ON THE SUBJECT, AND CALCULUS WAS TAUGHT AT ST ANDREWS 100 YEARS BEFORE IT WAS ON THE CURRICULUM AT THE UNIVERSITY OF CAMBRIDGE. GREGORY IS ALSO REMEMBERED FOR HIS DISCOVERY OF THE DIFFRACTION GRATING BY USING A BIRD'S FEATHER, AND FOR HIS INVENTION OF THE 'GREGORIAN' TELESCOPE, WHICH IS STILL IN USE TODAY.

THIS MERIDIAN LINE WAS SET IN PLACE THANKS TO THE SUPPORT OF THE UNIVERSITY OF ST ANDREWS AND THE ST ANDREWS PRESERVATION TRUST'S GORDON CHRISTIE BEQUEST.

---

[70] A version of this series was known to the 14th Century Indian mathematician Mādhava, as were the Maclaurin series for sine and cosine, though this material was unknown in Europe. Gregory independently rediscovered it in 1671, shortly before Leibniz did the same thing.

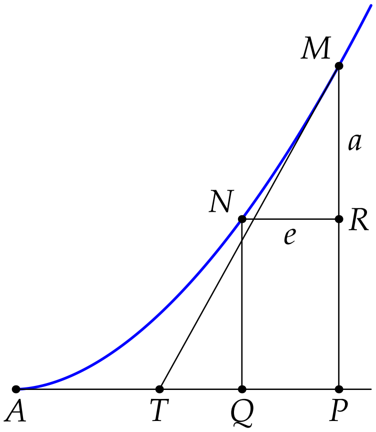
Isaac Barrow (1630–1677)

Like Gregory, Barrow also studied mathematics in Italy (and France), before returning to England where he become the inaugural Lucasian Professor of Mathematics[71] at Trinity College, Cambridge.

Barrow's work on calculus was predominantly geometric: he stated geometric versions of both parts of the Fundamental Theorem, crediting Gregory with part of the argument.

Newton encouraged Barrow to develop more algebraic methods for computing tangents and subtangents. One result was a modification of Fermat's algorithm for differentiation.

Unlike Fermat, Barrow is explicit in referring to "infinitely small" arcs and segments. Indeed Barrow is arguably the originator of the term *differential*: he referred to the *differential triangle* $\triangle MNR$ (see also page 118), since its sides are the *differences* between the co-ordinates of the infinitesimally separated points M, N.



For instance, here is how Barrow found the slope of the so-called kappa curve $x^2(x^2+y^2) = r^2y^2$ at a point (x, y):

- Replace x and y with $x+e$ and $y+a$ respectively, and expand to obtain

$$(x^2+2xe+e^2)(x^2+2xe+e^2+y^2+2ya+a^2) = r^2(y^2+2ya+a^2)$$

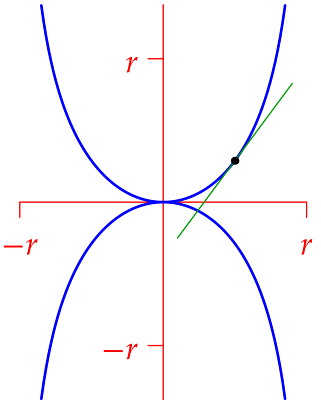
- Delete everything from the original equation $x^2(x^2+y^2) = r^2y^2$ and every expression containing two or more of the terms e, a, obtaining

$$x^2(2xe+2ya)+2xe(x^2+y^2) = 2r^2ya \implies x(2x^2+y^2)e = (r^2-x^2)ya$$

- The slope is therefore the ratio $a : e = x(2x^2+y^2) : y(r^2-x^2)$.

If we re-write the infinitesimals $e = \mathrm{d}x$ and $a = \mathrm{d}y$, and use modern fractional notation, this is precisely what we'd get using implicit differentiation:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{x(2x^2+y^2)}{y(r^2-x^2)}$$



Note again the essential difficulty with these algorithmic approaches to calculus: the infinitesimal quantities e, a are necessary for the calculation, but most are discarded when no longer useful! Is it really legitimate to calculate with such objects?[72]

---

[71] The Lucasian Chair is one of the world's most prestigious academic positions. While named as a position in *Mathematics*, it has become more associated with theoretical *Physics*, in part due to the fame of its second incumbent, Issac Newton, and later chairs including George Stokes, Paul Dirac and Stephen Hawking.

[72] This is the objection at the heart of Bishop George Berkeley's (after whom the Californian city and university are named) famous 1734 critique of calculus, that infinitesimals are merely the "ghosts of departed quantities."

Isaac Newton (1642–1727)

Newton enrolled as an undergraduate at Trinity College, Cambridge in 1661, at a time when mathematics was not generally taught to undergraduates. Newton, however, had been introduced to the subject in grade-school and, taking advantage of the freedom Cambridge afforded him to study essentially whatever he liked, threw himself into mathematics, particularly Euclid, Viète, Oughtred, Wallis and Descartes. He discussed mathematics with Barrow and, on the former's retirement in 1669, assumed the Lucasian Professorship, a position he held until 1702.

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 him lecturing to an empty room even after no-one had turned up!

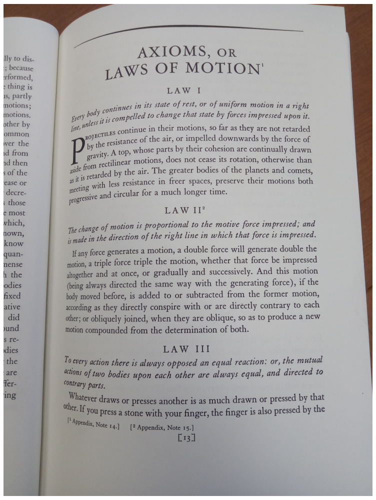
We are mostly interested in Newton's mathematics, though his fame comes more from its wide application, not least to gravitation. His magnum opus, one of the central works of the Scientific Revolution, was published in 1686: *Philosophiæ Naturalis Principia Mathematica*. The text begins with Newton's familiar laws of motion:

I. Inertial motion. A body unaffected by an external force continues moving in a straight line at uniform speed.

II. $F = ma$. A force acting on a body causes a proportional acceleration.

III. Equal-and-opposite forces. If one body exerts a force on a second, then the second exerts on the first a force of the same magnitude and opposite direction.

In Euclidean fashion, these are the *axioms* of Newtonian mechanics.



AXIOMS, OR
LAWS OF MOTION[1]

LAW I

Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed upon it.

PROJECTILES continue in their motions, so far as they are not retarded by the resistance of the air, or impelled downwards by the force of gravity. A top, whose parts by their cohesion are continually drawn aside from rectilinear motions, does not cease its rotation, otherwise than as it is retarded by the air. The greater bodies of the planets and comets, meeting with less resistance in freer spaces, preserve their motions both progressive and circular for a much longer time.

LAW II[2]

The change of motion is proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed.

If any force generates a motion, a double force will generate double the motion, a triple force triple the motion, whether that force be impressed altogether and at once, or gradually and successively. And this motion (being always directed the same way with the generating force), if the body moved before, is added to or subtracted from the former motion, according as they directly conspire with or are directly contrary to each other; or obliquely joined, when they are oblique, so as to produce a new motion compounded from the determination of both.

LAW III

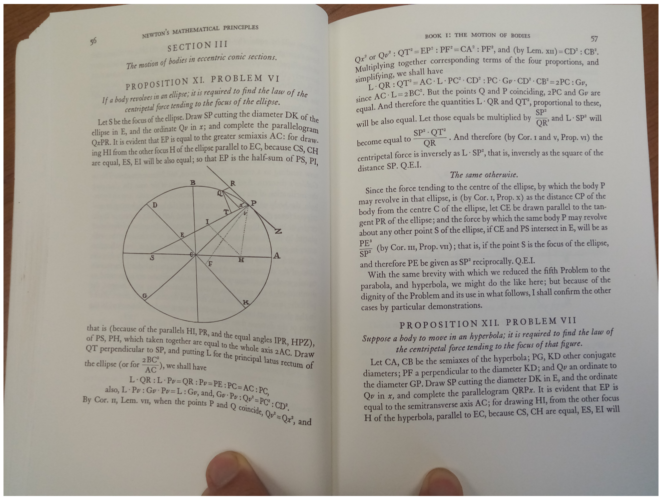
To every action there is always opposed an equal reaction: or, the mutual actions of two bodies upon each other are always equal, and directed to contrary parts.

Whatever draws or presses another is as much drawn or pressed by that other. If you press a stone with your finger, the finger is also pressed by the

[[1] Appendix, Note 14.] [[2] Appendix, Note 15.]

[13]

Newton's Laws of Motion



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 Qv in x; and complete the parallelogram $QxPR$. 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 $\frac{2BC^2}{AC}$), we shall have

$$L\cdot QR : L\cdot Pv = QR : Pv = PE : PC = AC : PC,$$

$$\text{also, } L\cdot Pv : Gv\cdot Pv = L : Gv, \text{ and, } Gv\cdot Pv : Qv^2 = PC^2 : CD^2.$$

By Cor. II, Lem. VII, when the points P and Q coincide, $Qv^2 = Qx^2$, and

BOOK I: THE MOTION OF BODIES 57

Qx^2 or $Qv^2 : QT^2 = EP^2 : PF^2 = CA^2 : PF^2$, and (by Lem. XII) $= CD^2 : CB^2$. Multiplying together corresponding terms of the four proportions, and simplifying, we shall have

$$L\cdot QR : QT^2 = AC\cdot L\cdot PC^2\cdot CD^2 : PC\cdot Gv\cdot CD^2\cdot CB^2 = 2PC : Gv,$$

since $AC\cdot L = 2BC^2$. But the points Q and P coinciding, 2PC and Gv are equal. And therefore the quantities $L\cdot QR$ and QT^2, proportional to these, will be also equal. Let those equals be multiplied by $\frac{SP^2}{QR}$, and $L\cdot SP^2$ will become equal to $\frac{SP^2\cdot QT^2}{QR}$. And therefore (by Cor. I and V, Prop. VI) the centripetal force is inversely as $L\cdot SP^2$, 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 $\frac{PE^3}{SP^2}$ (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 Qv an ordinate to the diameter GP. Draw SP cutting the diameter DK in E, and the ordinate Qv in x, and complete the parallelogram $QRPx$. 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

Kepler's 1st law $\Longrightarrow$ inverse-square force

Newton proceeds to apply calculus to prove the relationship between Kepler's laws of planetary motion (pg. 97) and an inverse-square gravitational force. At first glance it is hard to see where calculus appears in the pictured discussion. Observe the small parallelogram $QxPR$ and how Newton notes at the bottom of the page, "when the points P and Q coincide,…" The elliptical arc $\overset{\frown}{PQ}$ is *infinitesimal*.

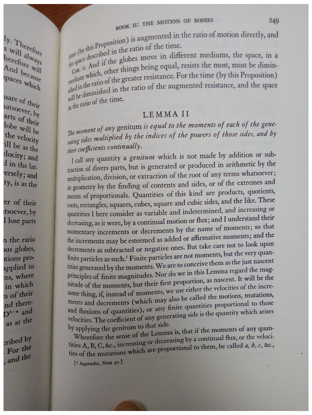
The *Principia* is Newton's first published work involving calculus, though many of its results and methods were worked out 20 years previously, just after completing his undergraduate studies and while Cambridge was closed due to the 1665–6 plague epidemic.

Newton's geometric presentation was typical for the time. He comments on how calculations are more efficient using indivisible/infinitesimal 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 versions of many familiar 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 pages below, 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.

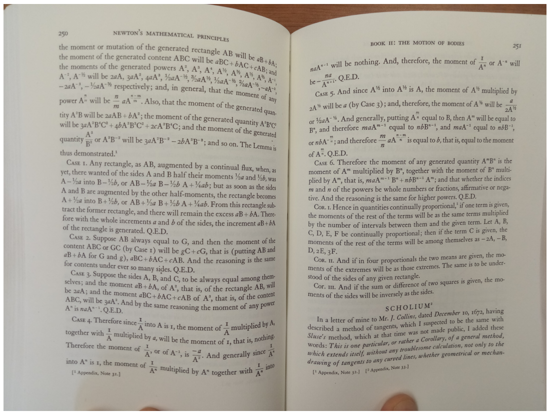
LEMMA II

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.

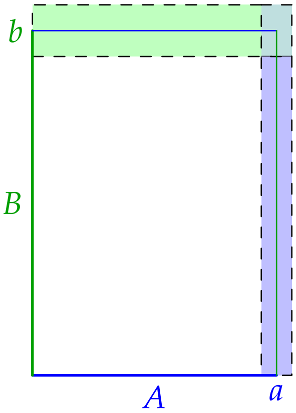
Cor. ii. And if in four proportionals the two means are given, the moments of the extremes will be as those extremes. The same is to be understood of the sides of any given rectangle.

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 and 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,

$$(A + a)(B + b) - AB \approx aB + bA$$

ignoring the double-infinitesimal quantity ab, 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 $\dfrac{n}{m}aA^{\frac{n-m}{m}}$.

in 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 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*. A fluent is a 'flowing quantity,' to us a *smooth function*, though such a concept was not defined in any modern sense.

Derivatives of fluents were *fluxions*, denoted by placing a dot above the corresponding fluent: $\dot{x}, \dot{y}$. The modern short-hand f' comes from this notation.[73] To be in *flux* is to be changing, hence 'rate of change.'

Anti-derivatives were denoted with an accent. Thus $\acute{x}$ is a *fluent of which x is the fluxion.*

Newton had several algorithms for computing fluxions, often variants of those of earlier mathematicians such as Barrow. For instance, here he finds the relationship between the fluxions of fluents x, y satisfying $x^2 + 3xy^3 + y = 5$.

1. Rearrange the expression as a polynomial in x:

$$x^2 + (3y^3)x + (y - 5)$$

2. Multiply terms by a decreasing arithmetic sequence with constant differences -1 (for instance, $(2, 1, 0)$), and the entire expression by $\frac{\dot{x}}{x}$ to obtain

$$2x^2 + 1 \cdot 3y^3 x + 0(y - 5) \rightsquigarrow (2x + 3y^3)\dot{x}$$

3. Repeat for y, multiplying through using the *same arithmetic sequence* and then by $\frac{\dot{y}}{y}$. In this example, 2 corresponds to x^2, so we start with 3 for y^3:

$$\begin{aligned}(3x)y^3 + 0y^2 + y + (x^2 - 5) &\rightsquigarrow 3(3x)y^3 + 2 \cdot 0y^2 + 1 \cdot y + 0 \cdot (x^2 - 5) \\ &\rightsquigarrow (9xy^2 + 1)\dot{y}\end{aligned}$$

4. Sum these expressions, set equal to zero and rearrange for the required ratio:

$$(2x + 3y^2)\dot{x} + (9xy^2 + y)\dot{y} = 0 \implies \frac{\dot{y}}{\dot{x}} = -\frac{2x + 3y^3}{9xy^2 + 1}$$

The arithmetic sequence and subsequent multiplication by $\frac{\dot{x}}{x}$ simply encodes the power law for derivatives. The result is exactly what you'd obtain using modern implicit differentiation. Perhaps surprisingly, it doesn't matter what arithmetic sequence is used in Newton's algorithm, the result is the same...

---

[73] Newton's dot-notation ('pricked letters') persists separately in modern *dynamics* to denote derivatives with respect to time. For instance, $\ddot{\mathbf{r}} = -\frac{GM}{r^3}\mathbf{r}$ is the standard differential equation for gravitation arising from Newton's second law and his inverse-square force law (note the double dot for the second derivative).

Power Series & the Binomial Series Newton made tremendous use of power series expressions. A simplistic and overly dismissive reading of his calculus would suggest that he merely applied the power law to integrate and differentiate series term-by-term.

The binomial series was discovered during the plague years, though it first appeared in a private letter of 1676:

$$(1+x)^\alpha = 1+\sum_{k=1}^{\infty}\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}x^k$$

This permitted expansion of 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:

$$P+PQ\,\frac{m}{n} = P\frac{m}{n}+\frac{m}{n}AQ+\frac{m-n}{2n}BQ+\frac{m-2n}{3n}CQ+\frac{m-3n}{4n}DQ+\cdots \tag{$*$}$$

Exponents were written 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 modern language, $(*)$ reads

$$(P+PQ)^{m/n} = P^{m/n}+\frac{m}{n}P^{m/n}Q+\frac{m(m-n)}{2n^2}P^{m/n}Q^2+\frac{m(m-n)(m-2n)}{6n^3}P^{m/n}Q^3+\cdots$$

Newton's 'proof' wouldn't pass modern muster, with his discoveries largely the result of inspired pattern-spotting and an incredible willingness to grind out calculations. Several examples were explicitly verified by multiplying out or using long-division. For instance, the series for $\sqrt{1+x}$ may be obtained by comparing coefficients:

$$\begin{aligned}
&1+x = (1+ax+bx^2+cx^3+\cdots)^2 = 1+2ax+(a^2+2b)x^2+(2c+2ab)x^3+\cdots\\
&\implies a=\frac{1}{2}, \quad b=-\frac{1}{2}a^2=-\frac{1}{8}, \quad c=-ab=\frac{1}{16}, \quad \text{etc.}\\
&\implies \sqrt{1+x} = 1+\frac{1}{2}x-\frac{1}{8}x^2+\frac{1}{16}x^3+\cdots
\end{aligned}$$

Of particular note are Newton's (re-)discoveries of the sine and cosine series of Mādhava in 1669 (similarly to when Gregory established the arctangent series[74]). Such were differentiated as if they were polynomials: for instance

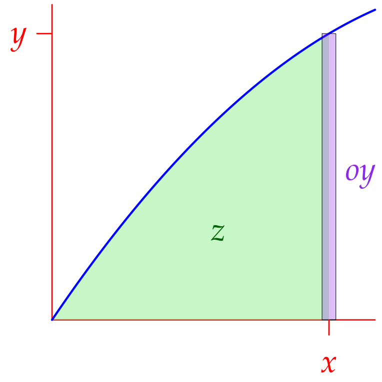
$$\sin x = x-\frac{1}{3!}x^3+\frac{1}{5!}x^5-\frac{1}{7!}x^7+\cdots \implies \cos x = 1-\frac{1}{2!}x^2+\frac{1}{4!}x^4-\frac{1}{6!}x^6+\cdots$$

Newton did not establish the full theory of power series as you would encounter in a modern analysis course, skirting over issues of convergence. The major question is whether term-by-term integration and differentiation of series is legitimate. The answer is (mostly) yes, though it was 150 years before Cauchy, Weierstraß and others rigorously confirmed this.

---

[74] Indeed Newton found the sine series by inverting that for arcsine, which can itself be found from the binomial series: see Exercise 6.

The Power Law & the Fundamental Theorem Here is one of Newton's arguments establishing the power law from fractional exponents. Observe the importance of the binomial series.



1. Suppose 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 infinitesimally by 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

$$\mathrm{d}z = y\,\mathrm{d}x \leftrightsquigarrow \frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\int^x y(t)\,\mathrm{d}t = y(x)$$

Using similar approaches, Newton produced one of the first tables of integrals, listing much of what you'd find inside the covers of a modern calculus textbook. By combining power series and the power law, he was able to efficiently integrate and differentiate an enormous variety of functions. His methods are a vast improvement over the predominantly *geometric* emphasis of earlier expositions.

Later Life & Legacy As mentioned previously, Newton's fame really came from his applications of mathematics to science. To some extent, Newton is where the subjects of mathematics and physics split: physicists & engineers took the easy-to-use tools of calculus and applied them to model the real world, while mathematicians tried to create new computational tools and fretted over whether already-existing methods were logically rigorous.

Newton slowly retired from science and mathematics in the early 1700s. Developing health problems, lapses of memory and depression probably contributed, though he continued to work on mathematical problems and to supervise new editions of his works. For much of his later years a battle raged—largely between surrogates—for supremacy with Leibniz over the invention of the calculus. Until his death in 1727, he lived off his income as Warden and Master of the Royal Mint. This required relatively little work on Newton's part, though he is said to have taken his responsibilities seriously.

Gottfried Wilhelm Leibniz (1646–1716)

Leibniz hailed from Leipzig, southwest of Berlin, in what was then part of the Holy Roman Empire. He eventually became a diplomat and then counsellor to the Duke of Hanover.

His initial academic studies were in philosophy, following his professor father. His taste for advanced mathematics was fueled during a 1672–6 sojourn in Paris during which he learned of van Shooten's expansion of Descartes' analytic geometry and of the *differential triangle* (page 118), which was already in use by others such as Pascal and Barrow.

Notation & the Fundamental Theorem The familiar modern notations for derivatives $\frac{\mathrm{d}y}{\mathrm{d}x}$ and integrals $\int y\,\mathrm{d}x$ come from Leibniz. Very loosely, here are their origins, and how they relate to the fundamental theorem. Suppose

$$(x_0, z_0), (x_1, z_1), \dots, (x_n, z_n), \qquad x_0 < x_1 < \cdots < x_n$$

describe a sequence of points along a curve $z(x)$ defined on an interval $[x_0, x_n]$. Form the sequences of *differences* and *sums* of the ordinates z_i:

$$\big(\delta z_i\big) = \big(z_1 - z_0,\ z_2 - z_1,\ \dots,\ z_n - z_{n-1}\big), \qquad \big(\sum z_i\big) = \big(z_0,\ z_0 + z_1,\ \dots,\ z_0 + \cdots + z_n\big)$$

Two relationships between differences and sums are immediate:

1. The difference sequence of the sums returns the original sequence:

$$\big(\delta \sum z_i\big) = \big(z_0, z_1, \dots, z_n\big)$$

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 by 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: the integral symbol is simply an elongated letter "S"!

Observation 2 now asserts that the sum of the infinitesimal changes in z is its net change

$$\int \mathrm{d}z = z(x_n) - z(x_0)$$

If $z(x)$ describes the *area*[75] under a curve $y(x)$, the fact that $\mathrm{d}z = y\,\mathrm{d}x$ recovers part 2 of the fundamental theorem: the area equals the sum of its infinitesimal increments

$$\int y\,\mathrm{d}x = z$$

---

[75] Notation is chosen to match Newton's discussion of the power law (page 125).

Observation 1 becomes part 1 of the fundamental theorem once applied to a 'sequence' of infinitesimals ($z_i \rightsquigarrow y\,\mathrm{d}x$): the rate of change of the area function is the ordinate,

$$\mathrm{d}\left(\int y\,\mathrm{d}x\right) = y\,\mathrm{d}x, \qquad \text{or alternatively,} \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left(\int y\,\mathrm{d}x\right) = y$$

While modern mathematicians are happy to refer to infinitesimals and infinite sums, Leibniz was more cagey in his publications, likely out of fear of criticism. He referred to each $\mathrm{d}x$ as an arbitrary, minuscule, but *finite* line segment. Like every other contemporary practitioner, he therefore fails to fully grapple with the essential paradoxes of calculus.

This very issue lies behind a common practice when teaching modern calculus. Instructors often insist that introductory students write differentials only in *combination*: $\frac{\mathrm{d}z}{\mathrm{d}x} = y$ is preferred to $\mathrm{d}z = y\,\mathrm{d}x$, where the former is the limit

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x}$$

visualized as a rate of change (slope). The helps instructors dodge a discussion of infinitesimals by punting the conceptual difficulty to that of limit, something that is typically easier to motivate, even though a formal definition is beyond the grasp of beginning students.

Example Regardless of his logical sleights-of-hand, Leibniz 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.

General Calculus Laws Such approaches enabled Leibniz to justify general formulas such as the linearity of derivatives, the product, quotient and power rules

$$\mathrm{d}(f+g) = \mathrm{d}f + \mathrm{d}g, \qquad \mathrm{d}(fg) = g\,\mathrm{d}f + f\,\mathrm{d}g, \qquad \mathrm{d}\left(\frac{f}{g}\right) = \frac{g\,\mathrm{d}f - f\,\mathrm{d}g}{g^2}$$

Even the chain rule is straightforwardly familiar in Leibniz's notation. For instance, to differentiate $y(x) = \sqrt{1+x^3}$, Leibniz would perform the substitution $u = 1 + x^3$ and observe that

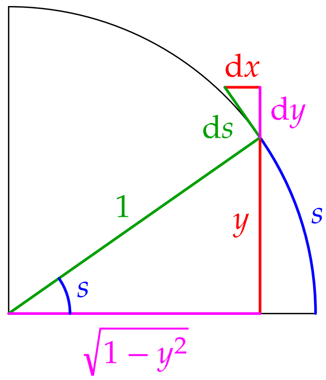
$$\mathrm{d}y = \mathrm{d}\sqrt{u} = \sqrt{u+\mathrm{d}u} - \sqrt{u} = \frac{u+\mathrm{d}u-u}{\sqrt{u+\mathrm{d}u}+\sqrt{u}} = \frac{\mathrm{d}u}{2\sqrt{u}} = \frac{3x^2\,\mathrm{d}x}{2\sqrt{1+x^3}}$$

Leibniz's notations and approaches are likely much more familiar to the modern reader than those of Newton.

The Sine Series Like Newton, Leibniz worked extensively with power series. As a final example, here is his computation of the Maclaurin series for the sine function. In contrast to Newton's approach (footnote 74), Leibniz derived the well-known second-order ODE for the sine function.

In a circle of radius 1, let s be the polar angle and $y = \sin s$ the ordinate. Observing that the differential triangle is similar to the large triangle with hypotenuse 1, we conclude

$$\frac{\mathrm{d}y}{\mathrm{d}s} = \frac{\sqrt{1-y^2}}{1} \implies (1-y^2)(\mathrm{d}s)^2 = (\mathrm{d}y)^2 \qquad (*)$$



Leibniz now supposes that the infinitesimals ds describing the arc are constant and applies d again: by the product rule, we obtain

$$-2y\,(\mathrm{d}y)(\mathrm{d}s)^2 = 2(\mathrm{d}y)\mathrm{d}(\mathrm{d}y) \implies \frac{\mathrm{d}(\mathrm{d}y)}{(\mathrm{d}s)^2} = -y \qquad \left(= \frac{\mathrm{d}^2y}{\mathrm{d}s^2}\right)$$

We bracketed everything to make clear that Leibniz is applying d (to compute differences). When the brackets are removed, we obtain the modern notation for second derivatives!
Leibniz now writes

$$y(s) = \sin s = c_0 + c_1 s + c_2 s^2 + c_3 s^3 \cdots$$

as a power series, whose coefficients are, as yet, unknown. Since at $s = 0$, $y = \sin s = 0$ and $\mathrm{d}s = \mathrm{d}y$ (by $*$), Leibniz concludes that $c_0 = 0$ and $c_1 = 1$. He now differentiates the series twice and compares coefficients:

$$\begin{aligned}
\frac{\mathrm{d}(\mathrm{d}y)}{(\mathrm{d}s)^2} &= 2c_2 + 6c_3 s + 12c_4 s^2 + 20c_5 s^3 + 30c_6 s^4 + \cdots \\
-y &= -s - c_2 s^2 - c_3 s^3 - c_4 s^4 - \cdots \\
\implies 0 = c_2 = c_4 = c_6 &= \cdots, \quad c_3 = -\frac{1}{6}, \quad c_5 = -\frac{1}{20}c_3 = \frac{1}{120}, \ \cdots
\end{aligned}$$

The result is the familiar series

$$y = \sin s = s - \frac{1}{6}s^3 + \frac{1}{120}s^5 + \cdots = s - \frac{1}{3!}s^3 + \frac{1}{5!}s^5 - \frac{1}{7!}s^7 + \cdots$$

Legacy Leibniz's notation and approach are not precisely the same as modern calculus, though they are very familiar, more so than Newton's. While the battle between their supporters for primacy of discovery was fought to a grudging draw, Leibniz's notation very much won the day. The development of calculus (analysis) over the next 150 years was largely a continental European affair with the British, stubbornly wedded to Newton's approach, falling behind. Regardless, the computational efficiency afforded by both men's efforts ended up super-charging mathematics and its applicability to real-world problems.

Exercises 8.3. *Key concepts: Differential triangle, Calculus becomes algebraic/algorithmic, Power series, Notations of Newton & Leibniz*

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 $x^3 + y^3 = c^3$ at a given point $P = (x, y)$.

3. Suppose $y > 0$ and that the point $P = (\frac{r}{2}, y)$ lies on Barrow's *kappa curve*. Find y. Compute the slope of the tangent line at P, and show that it is independent of r.

4. Calculate a power series for $\frac{1}{1-x^2}$ by using long-division.

5. Use the binomial series to obtain a power series expression for $\ln(1+x)$, which Newton knew to describe the area under the curve $\frac{1}{1+x}$.

6. (a) Use the binomial series to show that

$$(1-y^2)^{-1/2} = 1 + \frac{1}{2}y^2 + \frac{3}{8}y^4 + \frac{5}{16}y^6 + \cdots$$

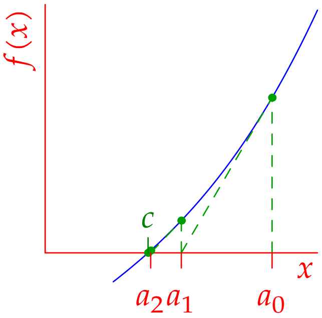
   (b) Use part (a) to find a power series expression for $\arcsin x$ up to an x^5 term.

   (c) By writing $\sin z = c_0 + c_1 z + c_2 z^2 + c_3 z^3 + \cdots$ as a power series, evaluate enough of

$$x = \sin(\arcsin x)$$

   to obtain the coefficients $c_0, \ldots, c_5$.

7. 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 $(a_0, f(a_0))$ 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?

8. 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?

9. Use Leibniz's differential triangle (page 128) 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.

10. Suppose $y(x)$ describes a non-negative curve with $y(0) = 0$. Leibniz's *transmutation theorem* relates the area under this curve 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:

$$\int_0^a y\,\mathrm{d}x = \frac{1}{2}\left[ay(a) + \int_0^a z\,\mathrm{d}x\right]$$

   (a) Justify the claim that $z(x)$ is the y-intercept of the tangent line to $y(x)$.

   (b) Use integration by parts to prove the transmutation theorem.

   (c) If $x^p = y^q$, find z and use the transmutation theorem to compute the area $\int_0^a y\,\mathrm{d}x$.