Chapter 14
The Birth of Calculus

He who can digest a second or third fluxion, a second or third difference, need not, methinks, be squeamish about any point in Divinity.
And what are these fluxions? The velocities of evanescent increments. And what are these evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them ghosts of departed quantities?—*Berkeley*

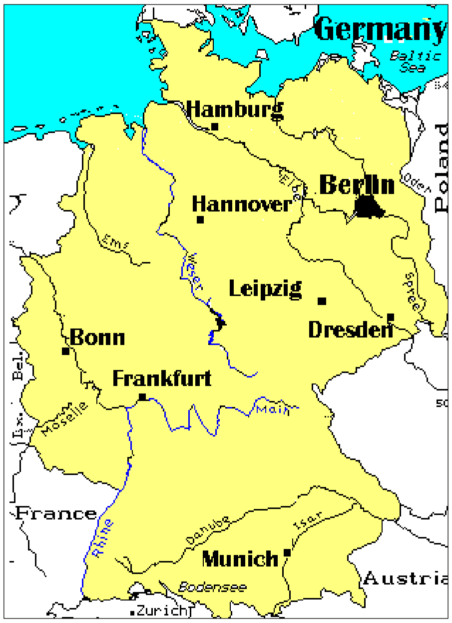
It is almost universally agreed upon that the two characters we encounter in this chapter, **Isaac Newton** (1642-1727) and **Gottfried Leibniz** (1646-1716) are the **discoverers-creators** of calculus. But what do we mean by this? We have seen in the previous chapter that derivatives were already known and so was their connection with tangents and with the extremal values of functions. In addition, the areas under curves of varied complexity had been computed by basically doing Riemann sums integration. Finally, the connection between the two processes of integration and differentiation had been foreseen, and Newton had been exposed to it from Barrow's lectures. One could say that Newton and Leibniz did understand thoroughly the **fundamental theorem of calculus** (as we call it today), and also both appreciated the power and range of the subject. Certainly, Newton used Calculus-type thinking to push the frontiers of mechanics and physics.

European History Highlights

Date	Event
1660	Restoration Age—Charles II becomes king of England.
1662	John Graunt publishes Bills of Mortality.
1663	Seven Year War ends—Canada is ruled by Britain.
1665	Great Plague of London.
1666	National Observatory of Paris is founded. Great Fire of London.
1675	Greenwich Observatory is established.
1682	Peter, the Great begins his reign in Russia.
1685	Huguenots are persecuted in France. J. S. Bach is born.
1686	Fahrenheit is born.
1688	Glorious Revolution in England—William and Mary become monarchs.
1694	Voltaire is born.
1706	Benjamin Franklin is born.

What is the **Fundamental Theorem of Calculus**? Most everyone would agree that essentially it is about the connection between the tangents of the graph of one function and the area under the curve of another. This connection will be clearer after we discuss Leibniz's work.

Leibniz



Like Fermat, **Gottfried Leibniz**, was not a mathematician by trade. He was a diplomat who traveled widely, and as such came to meet and discuss mathematics with all the best-known mathematicians and scientists of his time, including **Huygens** and **Newton**.

It is, in a sense, unfortunate that Leibniz met and corresponded with Newton, since he will, many years after the meeting, be accused of plagiarizing his ideas on calculus from Newton. A long, scandalous dispute followed, and although his name was eventually cleared, the dispute left a bitter taste in the soul of mathematicians on both sides of the English Channel. This led to a partial isolation of English mathematicians from those in the Continent, where calculus, and its consequent disciplines such as differential equations, will explode into a massive and powerful discipline. Leibniz would die unbeknownst to the world and in relative poverty.

Leibniz developed much of our modern notation for calculus such as $\frac{dy}{dx}$ and $\int$ and it is this notation (as opposed to Newton's fluxion notation) that will be adopted in the rest of Europe.

It is fortunate, however, that Leibniz met Huygens, since it is a question posed to him by Huygens that possibly stimulated Leibniz's discovery of the connection between integration and differentiation.

Huygens asked what the sum of the reciprocal of the triangular numbers added to:

$$1+\frac{1}{3}+\frac{1}{6}+\frac{1}{10}+\frac{1}{15}+\cdots=?$$

Recall Oresme. As it turned the answer was already known, but unknown to Leibniz, he plunged ahead into the problem.

He understood that from a given sequence: $\alpha: a_1, a_2, a_3, a_4, \ldots$, one could obtain two other ones, the **difference** and the **sum**. The first one of these: the **difference**, $\Delta(\alpha)$, is defined as follows, $\Delta(\alpha): b_1, b_2, b_3, \ldots$ where: $b_1 = a_2 - a_1$, $b_2 = a_3 - a_2$, $b_3 = a_4 - a_3$, etcetera. Thus, for example, if α is the sequence of triangular numbers, α: $1, 3, 6, 10, 15, \ldots$ then $\Delta(\alpha)$ is $2, 3, 4, 5, \ldots$.

The **sum** (or **series**), $\Sigma(\alpha)$, is defined as follows $\Sigma(\alpha)$: a_1, a_1+a_2, $a_1+a_2+a_3$, $a_1+a_2+a_3+a_4,\ldots$. For example, if we start with the simple sequence 1,1,1,…. and take consecutive sums of it, we obtain the following configuration:

α:	1	1	1	1	1	1	1	1	1	1	1	1
$\Sigma(\alpha)$:	1	2	3	4	5	6	7	8	9	10	11	12
$\Sigma^2(\alpha)$:	1	3	6	10	15	21	28	36	45	55	66	78
$\Sigma^3(\alpha)$:	1	4	10	20	35	56	84	120	165	220	286	364
$\Sigma^4(\alpha)$:	1	5	15	35	70	126	210	330	495	715	1001	1365

If we think of this as a matrix with the rows and columns labeled by 0,1,2,…, then we see that we are dealing with Pascal's triangle since the (i, j) entry is nothing but $\binom{i+j}{i}=\binom{i+j}{j}$ (it is in this form that Pascal originally wrote his triangle).

In the context of sequences, the relation between the difference and the sum is easily understood. If $\alpha: a_1, a_2, a_3, a_4, \ldots$. Then

$$\Delta(\Sigma(\alpha)) = a_2, a_3, a_4, \ldots,$$

which is almost α (all we would need to recover α is attach a_1 at the beginning. Also

$$\Sigma(\Delta(\alpha)) = a_2 - a_1, a_3 - a_1, a_4 - a_1, \ldots,$$

and again α is easily recoverable from this sequence, once a_1 is given.

Leibniz set out to a parallel construction to the triangle above. But instead of consecutive sums he took consecutive differences, but since he was starting with a decreasing sequence, we need to modify the difference to mean:

$$b_1 = a_1 - a_2,\ b_1 = a_1 - a_2,\ b_2 = a_2 - a_3,\ b_3 = a_3 - a_4, \text{ etcetera.}$$

His first (row) sequence was the sequence of harmonic numbers: $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$.

α:	1	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9	1/10	1/11	1/12
$\Delta(\alpha)$:	1/2	1/6	1/12	1/20	1/30	1/42	1/56	1/72	1/90	1/110	1/132	1/156
$\Delta^2(\alpha)$:	1/3	1/12	1/30	1/60	1/105	1/168	1/252	1/360	1/495	1/660	1/858	1/1092
$\Delta^3(\alpha)$:	1/4	1/20	1/60	1/140	1/280	1/504	1/840	1/1320	1/1980	1/2860	1/4004	1/5460
$\Delta^4(\alpha)$:	1/5	1/30	1/105	1/280	1/630	1/1260	1/2310	1/3960	1/6435	1/10010	1/15015	1/21840

Then he easily observed that the second row consisted of the halves of the reciprocals of the triangular numbers. Hence if we let $\alpha: a_1, a_2, a_3, a_4, \ldots$ and $\Delta(\alpha) := b_1, b_2, b_3, \ldots$, then we know $b_1 = a_1 - a_2$, $b_2 = a_2 - a_3$, $b_3 = a_3 - a_2$, …. So, in a similar fashion to the discussion above, $\Sigma(\Delta(\alpha)) := a_1 - a_2$, $a_1 - a_3$, $a_1 - a_4$, …. So the sum of all the reciprocals of the triangular numbers, which is

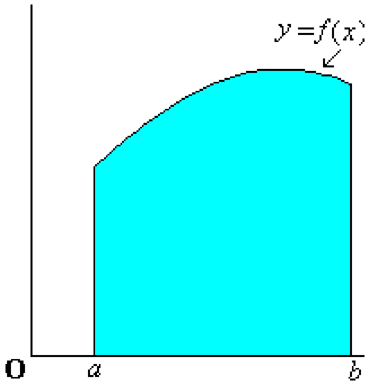
$$2\lim_{n\to\infty}\Sigma(\beta) = 2\lim_{n\to\infty}\Sigma(\Delta(\alpha)) = 2a_1 = 2$$

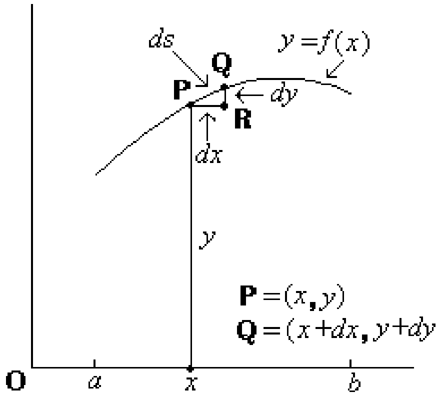
since $a_n \to 0$ as $n \to \infty$.

This was very exciting to him, since he realized he could adequately add one sequence by simply taking differences of another. Although certainly that reminds one of the basic ideas behind the fundamental theorem of calculus, it did not become real calculus until he pushed it further, and this is what we look at next.

In the 1670's, Leibniz discovered a general principle or technique to evaluate areas, which he referred to as **transmutation of areas**. A technique basically equivalent to the **Fundamental Theorem of Calculus**, which we now give in some detail, and in, more or less, modern notation and ideology.

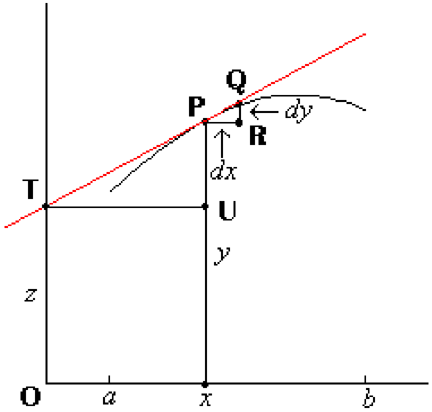
Suppose we have an interval $[a,b]$ and we have a function $y = f(x)$ defined on this interval. We are interested in the area under the curve of this function.





Start by considering two neighboring (**very close**) points **P** and **Q** on the graph of this function, where **P** has coordinates x and $y = f(x)$, and **Q**'s coordinates are $x + dx$ and $y + dy$, where dx is a small change in x, and dy is the corresponding small change in y. We will let **O** denote the origin.

Continuing in the language of indivisibles, we let the length of the curve from point **P** to point **Q** be denoted by ds (recall that s usually denotes length of a curve in Calculus).



Consider the tangent line to the curve at the point **P** and suppose it intersects the $y-$axis at **T** $= (0, z)$.

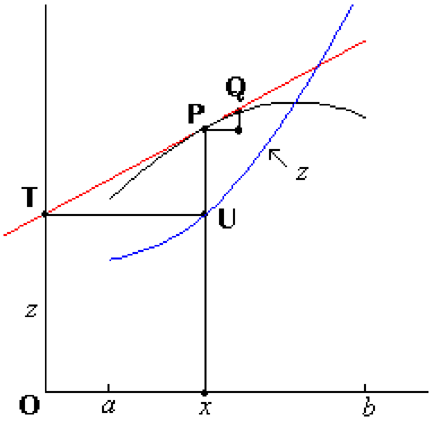
Since the right triangle Δ**TUP** is similar to the right triangle Δ**PRQ**, we have that

$$\frac{y - z}{x} = \frac{dy}{dx},$$

and solving for z we get,

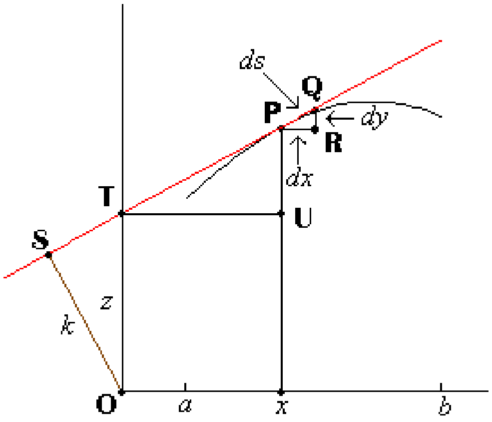
$$z = y - x\frac{dy}{dx}.$$

We can use this expression to define a new function z of x, whose graph is given by:



At the origin **O**, draw the perpendicular to the tangent line

TP and let it intersect this line at point **S**, which has hypotenuse z, and let k be the distance from **S** to the origin.



Since

$$\angle\mathbf{STO}+\angle\mathbf{PTU}=90^\circ,$$

and

$$\angle\mathbf{PTU}=\angle\mathbf{QPR},$$

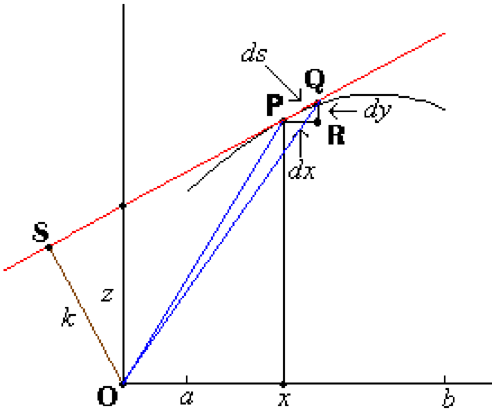
we have

$$\angle\mathbf{STO}=\angle\mathbf{PQR}.$$

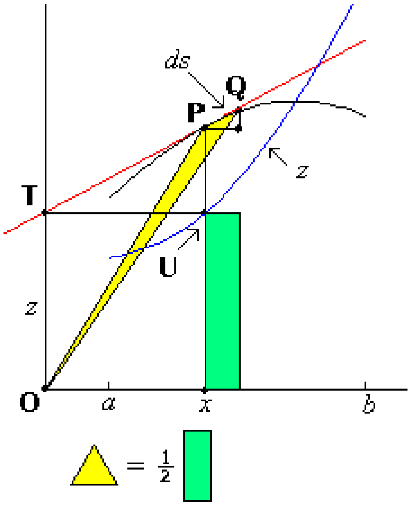
Hence right triangle $\Delta\mathbf{OST}$ is similar to right triangle $\Delta\mathbf{PRQ}$, and we have

$$\frac{dx}{k}=\frac{ds}{z}.$$

Consider now the infinitesimal triangle $\Delta\mathbf{OPQ}$:



Its base is ds and its height is k, hence its area is $\frac{kds}{2}$, which by the similarity above equals $\frac{zdx}{2}$.



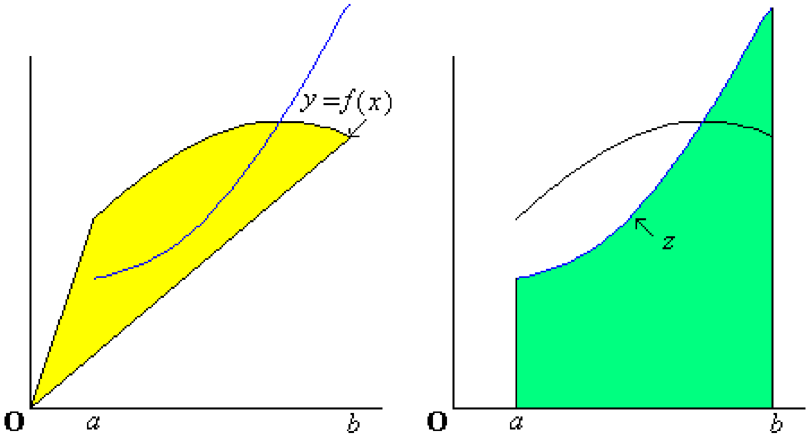
But as the picture on the left illustrates, $\frac{zdx}{2}$ equals half of the area in the indicated infinitesimal rectangle under the function z.

$$\text{area}\left(\Delta\mathbf{OPQ}\right)=\frac{zdx}{2}.$$

Hence we have that the area under the graph of z, which is given by

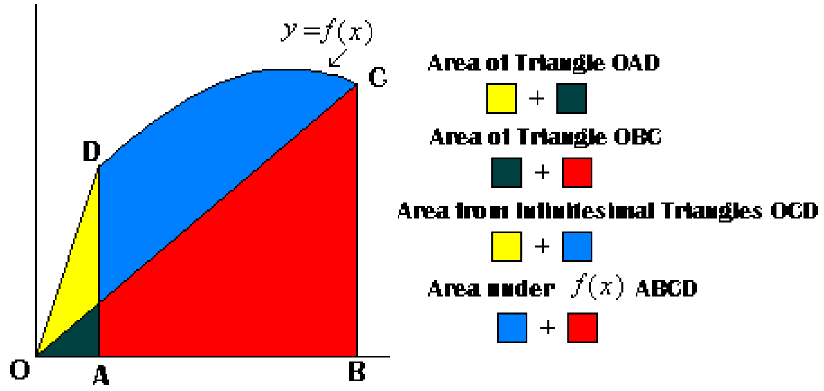
$$\int_a^b zdx$$

is twice the area of the shape made from all the infinitesimal triangles. Hence, in the picture on the right, the area on the left is half of the area on the right.



But by cutting and pasting in the picture below, we get that the area under $f(x)$ equals the area

$$\mathbf{OCD}+\mathbf{OBC}-\mathbf{OAD}.$$



But by the result above, **OCD** equals one half the area under z. Easily, **OBC** equals $\frac{1}{2}bf(b)$ and **OAD** is $\frac{1}{2}af(a)$, so symbolically we have:

$$\int_a^b ydx=\frac{1}{2}\left(\int_a^b zdx+bf(b)-af(a)\right)$$

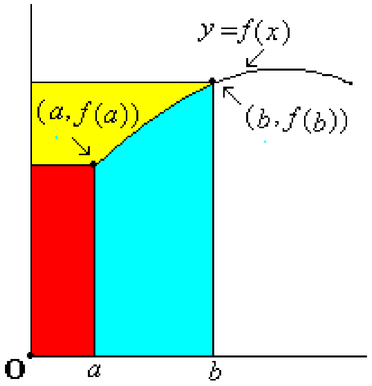
❶

and we have *exchanged one computation of areas for another* that may turn out to be simpler than the original as we will exemplify below.

But before we do that let's make a couple of observations.

First, $bf(b)-af(a)$ can be simplified using standard evaluation notation to $[xy]_a^b$. Second, we can use the fact that $z=y-x\frac{dy}{dx}$ to substitute in the equation above to obtain

$$\begin{aligned}\int_a^b ydx &= \frac{1}{2}\left(\int_a^b zdx+bf(b)-af(a)\right)\\ &= \frac{1}{2}\left(\int_a^b\left(y-x\frac{dy}{dx}\right)dx+[xy]_a^b\right)\\ &= \frac{1}{2}\int_a^b ydx-\frac{1}{2}\int_a^b xdy+\frac{1}{2}[xy]_a^b\end{aligned}$$

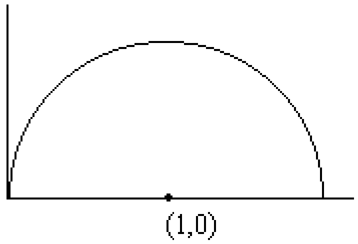


which after clearing and multiplying by 2 yields the *integration by parts* formula due to Leibniz:

$$\int_a^b ydx=[xy]_a^b-\int_{f(a)}^{f(b)} xdy,$$

which is tantamount—in the picture—to the three shaded areas filling in the rectangle—confirming the geometric reasoning he had employed in the more sophisticated equation ❶ above.

Leibniz himself was pleased with the following application of his ideas. Consider a circle of radius 1 centered at the point (1,0) so that it is tangent to the $y-$ axis.



Then its upper semicircle has the equation $y=\sqrt{2x-x^2}$. Since $y^2=2x-x^2$, differentiating, we get

$$2ydy=2dx-2xdx$$

and so

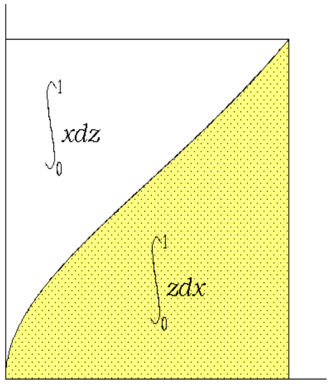
$$\frac{dy}{dx}=\frac{1-x}{y},$$

hence

$$z=y-x\frac{1-x}{y}=\sqrt{\frac{x}{2-x}}.$$

Solving for x, we have $$x=\frac{2z^2}{1+z^2}.$$

But, then we have



$$\begin{aligned}
\frac{\pi}{4} &= \int_0^1 ydx && \text{classical fact}\\
&= \frac{1}{2}\left(\int_0^1 zdx+\left[x\sqrt{2x-x^2}\right]_0^1\right) && \text{by ❶ above}\\
&= \frac{1}{2}\left(\left(1-\int_0^1 xdz\right)+1\right) && \text{by picture:}\\
&= 1-\int_0^1 \frac{z^2}{1+z^2}dz && \text{by substitution}\\
&= 1-\int_0^1 z^2\left(1-z^2+z^4-z^6+\cdots\right)dz && \text{by geometric series}\\
&= 1-\left[\frac{z^3}{3}-\frac{z^5}{5}+\frac{z^7}{7}-\frac{z^9}{9}+\cdots\right]_0^1 && \text{by term-wise integration}\\
&= 1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\cdots && \text{by evaluation}
\end{aligned}$$

However, although one can find the expression

$$\frac{\pi}{4}=1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots$$

quite beautiful, it is not very practical for doing computation—it converges too slowly. For example, if we use the fact that any alternating series toggles between exceeding and being less than the true sum, we have the not very good estimate $3.108269 < \pi < 3.173842$ even after 30 terms.

Newton

Born on Christmas Day, 1642, to a relatively poor widow, **Isaac Newton** showed promise as a student, and thus a brother of his mother agreed to support him in college. He attended Cambridge University. In 1665, during an outbreak of the plague, he was sent home, and it was during that period that he developed some of his best ideas. Soon after that, his teacher, Isaac Barrow resigned his position so that Newton can be appointed to follow him. For the next 30 years Newton was a professor at Cambridge—alas, a terrible lecturer, hardly any one would attend his lectures, but a widely known scholar. In 1693, he suffered a nervous breakdown, partially caused by the stress suffered during the dispute with Leibniz. After he recovered, he was appointed in charge of the Royal Mint where he spent the remainder of his life. When he died, he was the most famous scientist in the world, and was buried with all the glory and ceremony at Westminster Abbey.

Sir Isaac Newton is one of the most distinguished names in the history of mathematics and science. He can be considered one of the founders of modern science, and his book **Philosophiae Naturalis Principia Mathematica** (1687) (often referred simply as the **Principia**) is a major book in Western civilization.

We will be far from doing justice to Newton since we could—without much effort—spend a whole semester with him, just as with Euler or Gauss.

Newton had major impact on both mechanics and optics. We briefly glance at one of his three laws of motion.

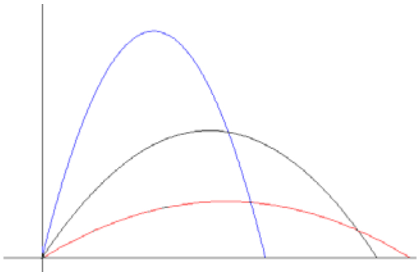
❶ **The Second Law of Motion:** $F = m \times a$

Force = mass × **acceleration**,

but actually Newton was even more accurate than that:

Force = mass $\overset{\bullet}{\times}$ **velocity**, where the dot stood for **fluxion** which is the word he used for derivative, or equivalently, in Leibniz's notation $F = \frac{d(m \times v)}{dt}$, so if mass is a constant, we get the more common version of the second law. However, if mass is not constant, we get a different law—one that is valid at the very high speeds of atomic particles of modern physics.

A very easy application, when this law is linked with Galileo's conclusion that gravity is constant, is the calculation of the path followed by a projectile such as a cannon ball: **Suppose a projectile is shot from the origin at angle θ with speed v_0. What is the path of the projectile?** If we separate the force into its two components, one in the x-direction and one in the y-direction (this idea is much older than Newton), we get that $F_x = 0$ while $F_y = -g$, a constant. But then $a_x = 0$ while $a_y = \frac{-g}{m}$, a constant. But, by integrating accelerations, we get velocities, $v_x = v_0 \cos(\theta)$ and $v_y = \frac{-g}{m} t + v_0 \sin(\theta)$, if we assume time is measured so $t = 0$ is when we shot the projectile. Hence, $s_x = v_0 \cos(\theta) t$ and $s_y = \frac{-g}{2m} t^2 + v_0 \sin(\theta) t$, and if we graph this path, we get a parabola.



❷ The Law of Gravitation

Newton may have independently arrived to this law:

> **two objects attract each other with a force proportional to their masses and inversely proportional to the square of their distance,**

but others had also proposed it (**Hooke** for one). But he was definitely the first one to have done something with it. He used mathematics (calculus ideas) to infer Kepler's first two laws of planetary motion from the law of gravitation, and naturally this served as a major piece of evidence of support for this law.

❸ The Binomial Theorem

One of his first successes, and a definite step toward calculus, was **his extension of the binomial theorem to other exponents besides positive integers.** What started as a technique to improve the computation of squared roots, and other roots, became a broader weapon, and made him a superb manipulator of series—, which was critical to his whole view of calculus. We recall that

$$(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i ,$$

where n is an arbitrary **positive** integer. One of the advantages of our present day notation is that we can easily write Newton's binomial theorem, where α **is an arbitrary**

number now:

$$(1+x)^{\alpha}=\sum_{i=0}^{\infty}\binom{\alpha}{i}x^{i},$$

where $\binom{\alpha}{i}$ is defined by $\dfrac{\alpha\times(\alpha-1)\times(\alpha-2)\times\cdots\times(\alpha-(i-1))}{i\times(i-1)\times(i-2)\times\cdots\times 1}$.

Make the following observations:

- there are i factors in the numerator and i factors in the denominator.
- if n is a positive integer, then for $i>n$, $\binom{n}{i}=0$, and thus Newton's version is a true extension of the finite case.
- Note the recursion, $\binom{\alpha}{i+1}=\dfrac{\alpha-i}{i+1}\times\binom{\alpha}{i}$.
- $\binom{\alpha}{i}$ is a polynomial in α of degree i—with roots $0,1,\ldots,i-1$.

We compute a few of these polynomials (using mainly the recursion given above.

$\binom{\alpha}{0}=1$, by definition or agreement;

$\binom{\alpha}{1}=\alpha$;

$\binom{\alpha}{2}=\dfrac{\alpha\times(\alpha-1)}{2\times 1}=\dfrac{\alpha^{2}-\alpha}{2}$;

$\binom{\alpha}{3}=\dfrac{\alpha\times(\alpha-1)\times(\alpha-2)}{3\times 2\times 1}=\dfrac{\alpha^{3}}{6}-\dfrac{\alpha^{2}}{2}+\dfrac{\alpha}{3}$;

$\binom{\alpha}{4}=\dfrac{\alpha-3}{4}\times\binom{\alpha}{3}=\dfrac{\alpha^{4}}{24}-\dfrac{\alpha^{3}}{4}+\dfrac{11\alpha^{2}}{24}-\dfrac{\alpha}{4}$;

$\binom{\alpha}{5}=\dfrac{\alpha-4}{5}\times\binom{\alpha}{4}=\dfrac{\alpha^{5}}{120}-\dfrac{\alpha^{4}}{12}+\dfrac{7\alpha^{3}}{24}-\dfrac{5\alpha^{2}}{12}+\dfrac{\alpha}{5}$;

$\binom{\alpha}{6}=\dfrac{\alpha^{6}}{720}-\dfrac{\alpha^{5}}{48}+\dfrac{17\alpha^{4}}{144}-\dfrac{5\alpha^{3}}{16}+\dfrac{137\alpha^{2}}{360}-\dfrac{\alpha}{6}$;

$\binom{\alpha}{7}=\dfrac{\alpha^{7}}{5040}-\dfrac{\alpha^{6}}{240}+\dfrac{5\alpha^{5}}{144}-\dfrac{7\alpha^{4}}{48}+\dfrac{29\alpha^{3}}{90}-\dfrac{7\alpha^{2}}{20}+\dfrac{\alpha}{7}$.

So, for example, if we are interested in taking square roots, we let $\alpha=\dfrac{1}{2}$ and we get the following coefficients,

$$\binom{\frac{1}{2}}{0}=1 \qquad \binom{\frac{1}{2}}{1}=\frac{1}{2} \qquad \binom{\frac{1}{2}}{2}=-\frac{1}{8} \qquad \binom{\frac{1}{2}}{3}=\frac{1}{16}$$

$$\binom{\frac{1}{2}}{4}=-\frac{5}{128} \qquad \binom{\frac{1}{2}}{5}=\frac{7}{256} \qquad \binom{\frac{1}{2}}{6}=-\frac{21}{1024} \qquad \binom{\frac{1}{2}}{7}=\frac{33}{2048}$$

So if we put them together with the binomial theorem, we get that

$$\sqrt{1+x}\approx 1+\frac{x}{2}-\frac{x^2}{8}+\frac{x^3}{16}-\frac{5x^4}{128}+\frac{7x^5}{256}-\frac{21x^6}{1024}+\frac{33x^7}{2048}+\textbf{higher order terms},$$

so if x is small, we should have a reasonable approximation.

For example, if we are interested in $\sqrt{7}$, then we can handle it this way, $\sqrt{7}=\sqrt{9-2}=3\sqrt{1-\frac{2}{9}}$, so we let $x=-\frac{2}{9}$, and we get $\frac{11248487}{12754584}$ for $\sqrt{1-\frac{2}{9}}$, which approximates to 0.8819171993, which when multiplied by 3, gives 2.645751598, a good estimate for $\sqrt{7}$.

Sometimes, a closed expression for the coefficients is desired (and can be found), although it may be difficult to find the pattern at first. Let us revisit the coefficients we just have computed: 1, $\frac{1}{2}$, $-\frac{1}{8}$, $\frac{1}{16}$, $-\frac{5}{128}$, $\frac{7}{256}$, $-\frac{21}{1024}$, $\frac{33}{2048}$ and at first the pattern does elude us. But let us go back to the definition:

$$\binom{\frac{1}{2}}{i}=\frac{\frac{1}{2}\times\left(\frac{1}{2}-1\right)\times\left(\frac{1}{2}-2\right)\times\cdots\times\left(\frac{1}{2}-(i-1)\right)}{i\times(i-1)\times(i-2)\times\cdots\times 1}=$$

$$\frac{\frac{1}{2}\times\frac{-1}{2}\times\frac{-3}{2}\times\frac{-5}{2}\times\cdots\times\frac{-(2i-3)}{2}}{i!}=\frac{(-1)^{i-1}1\times 3\times 5\times\cdots\times(2i-3)}{2^i i!}.$$

Can we simplify this further? Perhaps simplify is the wrong word, and what we are trying to do is reduce the expression to more familiar functions. If we could reduce everything to the **factorial**, we would be at peace. We need to understand then the product of consecutive odds:

$$1\times 3\times 5\times\cdots\times(2i-3)=\frac{(2i-3)!}{2\times 4\times 6\times\cdots\times(2i-4)}=$$

$$\frac{(2i-3)!}{(2\times 1)\times(2\times 2)\times(2\times 3)\times\cdots\times(2\times(i-2))}=\frac{(2i-3)!}{2^{i-2}(i-2)!},$$

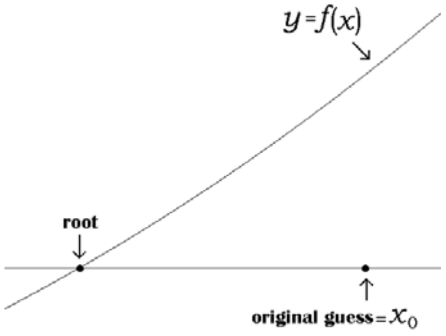
and so we conclude that

$$\binom{\frac{1}{2}}{i}=\frac{(-1)^{i-1}(2i-3)!}{2^{2i-2}(i-2)!\,i!}$$

for any *i*. Now whether this expression is computationally acceptable depends very much on our control of the factorial. But we do have what is called a closed expression.

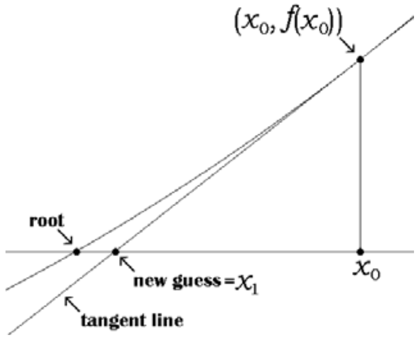
❹ Newton's Method

Very often, early in our mathematical career, perhaps as early as the first course in calculus, we get exposed to a powerful procedure for finding roots of equations called **Newton's Method**. It is more effective than the other two methods we discussed earlier in the Descartes section—the bisection method and the method of false position. Its ideal name would be the **tangent method** (in contrast to the secant method, a variation of the method of false position) since it finds its new guess by following the tangent line to the function at a guess.



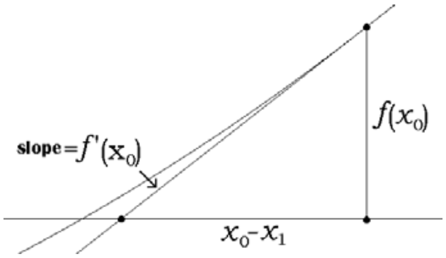
More formally, we review the procedure:

Start with one guess x_0 (as opposed to two guesses necessary in the other methods), and then **follow one's nose** by using the tangent line at the point of the original guess. How do we simply find the new guess, x_1?



We follow the tangent line at the point of the graph corresponding to our initial guess:

Geometrically then, we have an idea of where to locate our new approximation. But how do we find the point efficiently? Easily—use the slope:



slope of the tangent line = $\frac{\textbf{rise}}{\textbf{run}}$,

$$f'(x_0) = \frac{f(x_0) - 0}{x_0 - x_1},$$

and solving for x_1, we get

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Once we get x_1, we can use the same expression to get x_2, and then x_3, etceteras, continuing the iteration. When do we stop? **If there is not much change from one *x* to the next *x*, most probably we are close to a root, and we can stop.**

We give an application of the three methods to an original example of Newton's:

$$x^3 - 2x - 5 = 0.$$

Newton's	
x	$f(x)$
2	-1
2.1	0.061
2.094568121	0.000185723
2.094551482	0.0000000017

Method of False Position			
-	+	New	New Value
2	3	2.058823529	-0.390799919
2.058823529	3	2.08126366	-0.147204060
2.081263659	3	2.08963921	-0.054676503
2.089639210	3	2.092739574	-0.020202866
2.092739574	3	2.093883708	-0.007450506
2.093883708	3	2.094305451	-0.002745673
2.094305451	3	2.094460846	-0.001011574
2.094460846	3	2.094518093	-0.000372653

And we can easily see that at least in this example, Newton's is an easy winner. We note, again, that Newton's has the advantage of requiring only one guess.

But we also should remark that depending on the nature of the equation, the method can be unstable, and not lead to a root at all—the curious reader may attempt to find a root of $x^{\frac{1}{3}}$ by Newton's method.

Bisection Method			
-	+	New	New Value
2	3	2.5	5.625
2	2.5	2.25	1.890625
2	2.25	2.125	0.345703125
2	2.125	2.0625	-0.351318359
2.0625	2.125	2.09375	-0.008941650
2.09375	2.125	2.109375	0.166835785
2.09375	2.109375	2.1015625	0.078562260
2.09375	2.1015625	2.09765625	0.034714282
2.09375	2.09765625	2.095703125	0.012862332
2.09375	2.095703125	2.094726563	0.001954348
2.09375	2.094726563	2.094238281	-0.003495149
2.094238281	2.094726563	2.094482422	-0.000770775

Nevertheless, the method is so useful it is programmed in most hand-held calculators.

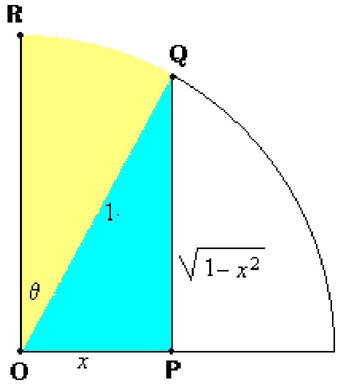
Needless to say, the method just described has been polished through time, and we now spend some time describing what Newton originally did. He used one of the original ideas behind calculus. That idea is **ignoring terms of higher order than 1**, in other words, ignoring everything except linear terms—actually, that is what approximating a curve by the tangent line is all about. Newton was an expert at that technique.

We illustrate, analytically, his original thinking behind his method with the example above: Solve $x^3 - 2x - 5 = 0$.

We start with one approximation, 2, the same as above, and so we let $x = 2 + p$, and obtain $p^3 + 6p^2 + 10p - 1 = 0$—remember x is supposed to be a root. Ignoring all but the linear term, we get $10p - 1 = 0$, $p = 0.1$ and we have a better approximation $x = 2.1$.

Again, we let $x = 2.1 + p$, and substituting, we get $p^3 + 6.3p^2 + 11.23p + .061 = 0$, and thus, by again considering only the linear term, we have that $p = -\frac{61}{11230} \approx -0.00543187$, which quickly gives us an estimate for $x = 2.1 - 0.00543187 = 2.09456813$—just as before.

❺ The Series for the Arcsine.



Consider the circle $x^2+y^2=1$ on the first quadrant. And take an arbitrary point **P** on the $x-$axis, at distance x from the origin **O**. Then we know the height of the circle at that point is $\sqrt{1-x^2}=(1-x^2)^{\frac{1}{2}}$, which by his binomial theorem (see ❸ above, and substitute $-x^2$ for x) Newton could expand into

$$\sqrt{1-x^2}=1-\frac{x^2}{2}-\frac{x^4}{8}-\frac{x^6}{16}-\frac{5x^8}{128}-\frac{7x^{10}}{256}-\frac{21x^{12}}{1024}-\frac{33x^{14}}{2048}+\cdots.$$

He could then integrate this series to obtain the total shaded area **OPQR**, consisting of the triangle Δ**OPQ** and the section of the circle **OQR**:

$$\Delta+\angle=x-\frac{x^3}{6}-\frac{x^5}{40}-\frac{x^7}{112}-\frac{5x^9}{1152}-\frac{7x^{11}}{2816}-\frac{21x^{13}}{13312}-\frac{33x^{15}}{30720}+\cdots.$$

But the area of the triangle Δ**OPQ** is known,

$$\Delta=\frac{x\sqrt{1-x^2}}{2}=\frac{x}{2}-\frac{x^3}{4}-\frac{x^5}{16}-\frac{x^7}{32}-\frac{5x^9}{256}-\frac{7x^{11}}{512}-\frac{21x^{13}}{2048}-\frac{33x^{15}}{4096}+\cdots.$$

And thus the section of the circle $\angle$**OQR** is given by the difference of the two series:

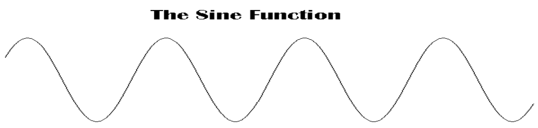
$$\angle=\frac{x}{2}+\frac{x^3}{12}+\frac{3x^5}{80}+\frac{5x^7}{224}+\frac{35x^9}{2304}+\frac{63x^{11}}{5632}+\frac{231x^{13}}{26624}+\frac{143x^{15}}{20480}+\cdots.$$

But $\angle=\frac{\theta}{2}$, and $x=\cos\left(\frac{\pi}{2}-\theta\right)=\sin(\theta)$, and so $\theta=\arcsin(x)$, and we have that

$$\theta=\arcsin(x)=x+\frac{x^3}{6}+\frac{3x^5}{40}+\frac{5x^7}{112}+\frac{35x^9}{1152}+\frac{63x^{11}}{2816}+\frac{231x^{13}}{13312}+\frac{143x^{15}}{10240}+\cdots.$$

❻ The Series for the Sine (and the Cosine).

Most of us take **Taylor's series** expansions very much for granted. Yet this most wonderful and powerful theorem **could not have even been thought of without enough examples to hint at its existence.**



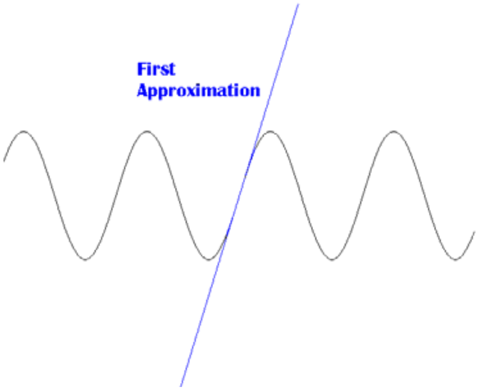
One of the best early examples is provided by the series expansion of the **sine** (and also the **cosine**)—which Newton developed. Start with the series of the Arcsine that we have above:

$$\theta=\arcsin(x)=x+\frac{x^3}{6}+\frac{3x^5}{40}+\frac{5x^7}{112}+\frac{35x^9}{1152}+\frac{63x^{11}}{2816}+\frac{231x^{13}}{13312}+\frac{143x^{15}}{10240}+\cdots.$$

The idea is to **solve for** x in terms of θ, $x=\sin(\theta)$. At all times, **we will eliminate all nonlinear terms**, continuing with the basic idea behind his method.

$$x+\frac{x^3}{6}+\frac{3x^5}{40}+\frac{5x^7}{112}+\frac{35x^9}{1152}+\frac{63x^{11}}{2816}+\frac{231x^{13}}{13312}+\frac{143x^{15}}{10240}+\cdots-\theta=0\,.$$

Dropping all nonlinear terms, we have as a first approximation, $x\approx\theta$, which as the picture indicates fits the sine function for small x's.



We are going to illustrate the method by using the first four terms of the series for the Arcsine:

$$x+\frac{x^3}{6}+\frac{3x^5}{40}+\frac{5x^7}{112}-\theta=0. \qquad ◎$$

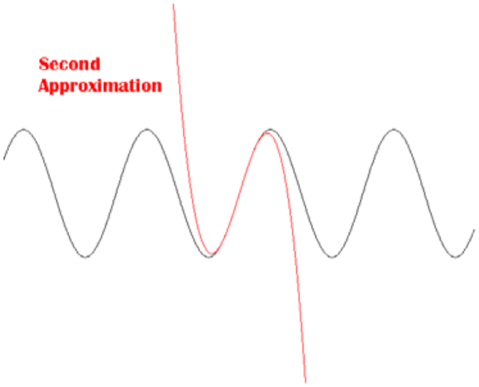
We let now $x=\theta+p$, and we get:

$$(\theta+p)+\frac{(\theta+p)^3}{6}+\frac{3(\theta+p)^5}{40}+\frac{5(\theta+p)^7}{112}-\theta=0\,,$$

Thus, $\dfrac{280\theta^3+126\theta^5+75\theta^7}{1680}+\left(\dfrac{16+8\theta^2+6\theta^4+5\theta^6}{16}\right)p+$ **higher order terms** = 0.

Ignoring the higher order terms and solving for p, we get that:

$$p=-\frac{\frac{280\theta^3+126\theta^5+75\theta^7}{1680}}{\frac{16+8\theta^2+6\theta^4+5\theta^6}{16}}=-\frac{\theta^3}{6}+\text{higher order terms,}$$



and so

$$x\approx\theta-\frac{\theta^3}{6}\,.$$

Now we let $x=\theta-\dfrac{\theta^3}{6}+p$, and substitute in ◎, and we obtain:

$$\left(\theta-\frac{\theta^3}{6}+p\right)+\frac{\left(\theta-\frac{\theta^3}{6}+p\right)^3}{6}+\frac{3\left(\theta-\frac{\theta^3}{6}+p\right)^5}{40}+\frac{5\left(\theta-\frac{\theta^3}{6}+p\right)^7}{112}-\theta=0.$$

And so when we expand,

$$\frac{1{,}306{,}368\theta^5+\textbf{higher order terms in }\theta}{156{,}764{,}160}{}^{1}+\left(\frac{756{,}496+\textbf{higher order terms in }\theta}{746{,}496}\right)p$$

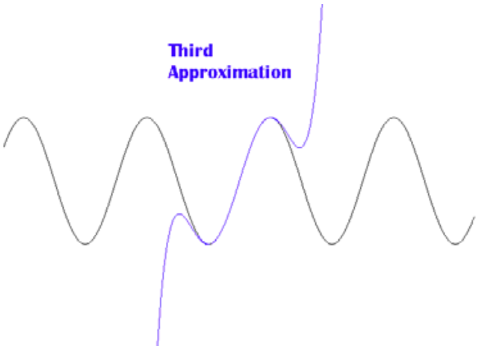
$$+\ \text{higher order terms in } p=0.$$

---

[1] In fact, the higher order terms in θ are $622{,}080\theta^7+5{,}019{,}840\theta^9-3{,}538{,}080\theta^{11}+1{,}088{,}640\theta^{13}-187{,}488\theta^{15}+18{,}900\theta^{17}-1{,}050\theta^{19}+25\theta^{21}$. Similarly, for the other fraction.

Simplifying, we have

$$\frac{\theta^5 + \textbf{higher order terms in } \theta}{120} + (1 + \textbf{higher order terms in } \theta)\,p + \cdots = 0.$$



When solving for p and ignoring anything but the lowest term order we get $p = \frac{\theta^5}{120}$, and so we have

$$x \approx \theta - \frac{\theta^3}{6} + \frac{\theta^5}{120}.$$

Continuing in this way, he eventually predicted the correct pattern:

$$\sin(\theta) = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \frac{\theta^9}{9!} - \frac{\theta^{11}}{11!} + \cdots.$$

From this series he used the fact that

$$\cos(\theta) = \sqrt{1 - \sin^2(\theta)},$$

and substituted the series for the sine into the series for $\sqrt{1 - x^2}$ —an absolutely formidable calculation.

Newton did this in order to establish

$$\cos(\theta) = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \frac{\theta^8}{8!} - \frac{\theta^{10}}{10!} + \cdots.$$

Did he know that the derivative of the sine function is the cosine? Perhaps not, since he would most probably have used the derivative to find the series for the cosine instead. To be strictly accurate historically, we have given a modern representation of Newton's actual calculations. At that time it was more customary to view the sine as corresponding to an arc more than an angle—so in fact radians were not needed.

Next we give the pictures for the next eight approximations of the sine series. And we see a steady improvement in our approximations.