A Brief Introduction to Infinitesimal Calculus

Section 1: Intuitive Proofs with "Small" Quantities

Abraham Robinson discovered a rigorous approach to calculus with infinitesimals in 1960 and published it in

Non-standard Analysis, *Proceedings of the Royal Academy of Sciences*, Amsterdam, ser A, 64, 1961, p.432-440

This solved a 300 year old problem dating to Leibniz and Newton. Extending the ordered field of (Dedekind) "real" numbers to include infinitesimals is not difficult algebraically, but calculus depends on approximations with transcendental functions. Robinson used mathematical logic to show how to extend all real functions in a way that preserves their properties in a precise sense. These properties can be used to develop calculus with infinitesimals.

Infinitesimal numbers have always fit basic intuitive approximation when certain quantities are "small enough," but Leibniz, Euler, and many others could not make the approach free of contradiction. Robinson's discovery offers the possibility of making rigorous foudations of calculus more accessible.

Section 1 of this article uses some intuitive approximations to derive a few fundamental results of analysis. We use approximate equality, $x \approx y$, only in an intuitive sense that "x is sufficiently close to y". Intutive approximation gives compelling arguments for the results, but not technically complete proofs.

H. Jerome Keisler developed simpler approaches to Robinson's logic and began using infinitesimals in beginning U. S. calculus courses in 1969. The experimental and first edition of his book were used widely in the 1970's. Section 2 of this article completes the proofs of Section 1 using Keisler's approach to the logic of infinitesimals from

Elementary Calculus: An Infinitesimal Approach*, 2nd Edition*, PWS Publishers, 1986, now available free at http://www.math.wisc.edu/~keisler/calc.html

Continuity & Extreme Values

A foundation of real analysis is:

Theorem: The Extreme Value Theorem

Suppose a function $f[x]$ is continuous on a compact interval $[a, b]$. Then $f[x]$ attains both a maximum and minimum, that is, there are points x_{MAX} and $x_{\min}$ in $[a, b]$, so that for every other x in $[a, b]$, $f[x_{\min}] \le f[x] \le f[x_{\text{MAX}}]$.

Formulating the meaning of "continuous" is a large part of making this result precise. We will take the intuitive "definition" that $f[x]$ is continuous means that if an input value x_1 is close to another, x_2, then the output values are close. We summarize this as: $f[x]$ is continuous if and only if

$$\boxed{a \le x_1 \approx x_2 \le b \Longrightarrow f[x_1] \approx f[x_2]}$$

Given this property of $f[x]$, if we partition $[a, b]$ into tiny increments,

$$a < a + \frac{1\,(b-a)}{H} < a + \frac{2\,(b-a)}{H} < \cdots < a + \frac{k(b-a)}{H} < \cdots < b$$

the maximum of the finite partition occurs at one (or more) of the points $x_M = a + \frac{k(b-a)}{H}$. This means that for any other partition point $x_1 = a + \frac{j(b-a)}{H}$, $f[x_M] \geq f[x_1]$.

Any point $a \leq x \leq b$ is within $\frac{1}{H}$ of a partition point $x_1 = a + \frac{j(b-a)}{H}$, so if H is very large, $x \approx x_1$ and

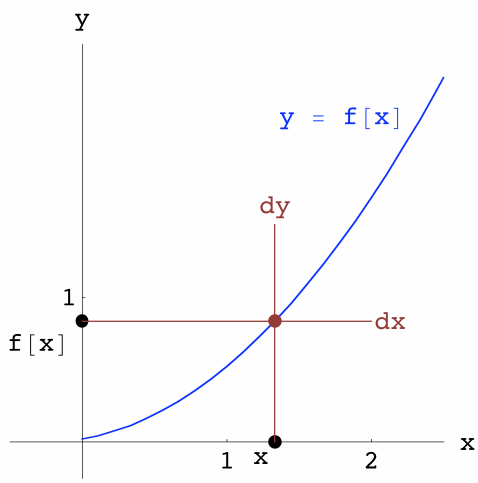
$$f[x_M] \geq f[x_1] \approx f[x]$$

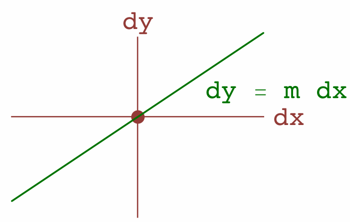
so we have found the approximate maximum.

It is not hard to make this idea into a sequential argument where $x_{M[H]}$ depends on H, but there is quite some trouble to make the sequence $x_{M[H]}$ converge (using some form of compactness of $[a, b]$.) Robinson's theory simply shows that the hyperreal x_M chosen when $1/H$ is infinitesimal, is infinitely near an ordinary real number where the maximum occurs. (A very general and simple re-formulation of compactness.) We complete this proof as a simple example of Keisler's Axioms in Section 2.

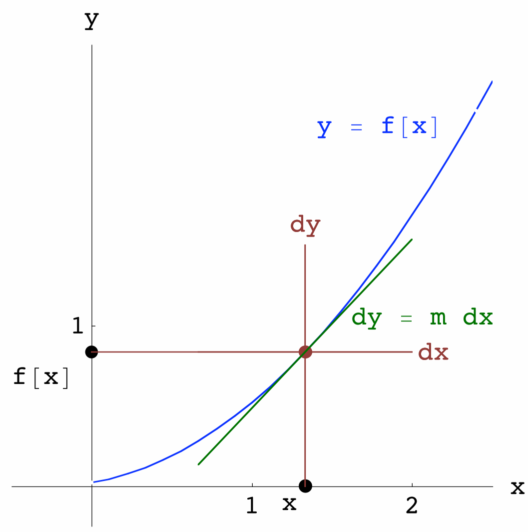
Microscopic tangency in one variable

In begining calculus you learned that the derivative measures the slope of the line tangent to a curve $y = f[x]$ at a particular point, $(x, f[x])$. We begin by setting up convenient "local variables" to use to discuss this problem. If we fix a particular $(x, f[x])$ in the x-y-coordinates, we can define new parallel coordinates (dx, dy) through this point. The (dx, dy)-origin is the point of tangency to the curve.





A line in the local coordinates through the local origin has equation dy $= m$ dx for some slope m. Of course we seek the proper value of m to make dy $= m$ dx tangent to $y = f[x]$.



The Tangent as a Limit

You probably learned the derivative from the approximation

$$\lim_{\Delta x \to 0} \frac{f[x+\Delta x]-f[x]}{\Delta x} = f'[x]$$

If we write the error in this limit explicitly, the approximation can be expressed as

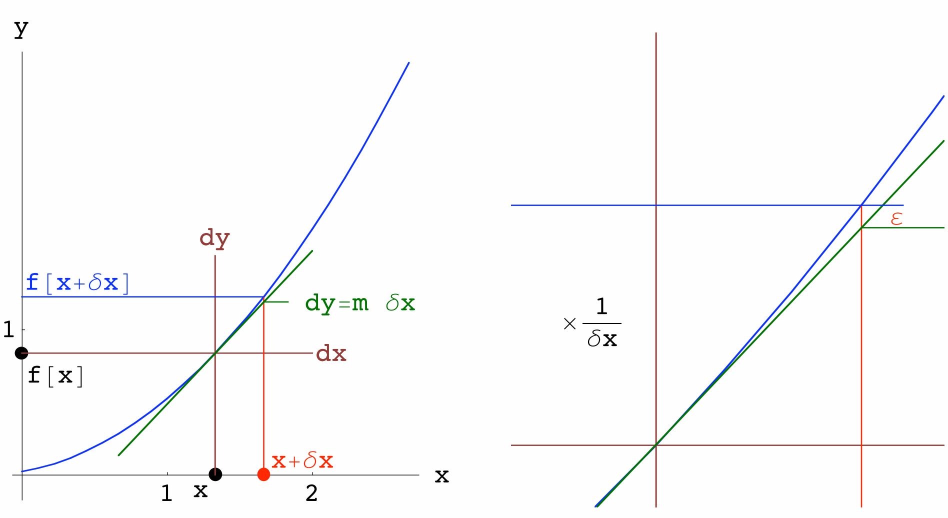
$$\frac{f[x+\Delta x]-f[x]}{\Delta x} = f'[x]+\varepsilon \text{ or } f[x+\Delta x]-f[x] = f'[x]\cdot\Delta x+\varepsilon\cdot\Delta x$$

where $\varepsilon \to 0$ as $\Delta x \to 0$. Intuitively we may say the error is small, $\varepsilon \approx 0$, in the formula

$$\boxed{f[x+\delta x]-f[x] = f'[x]\cdot\delta x+\varepsilon\cdot\delta x} \qquad (1.1.1)$$

when the change in input is small, $\delta x \approx 0$. The nonlinear change on the left side equals a linear change plus a term that is small compared with the input change.

The error ε has a direct graphical interpretation as the error measured above $x+\delta x$ after magnification by $1/\delta x$. This magnification makes the small change δx appear unit size and the term $\varepsilon\cdot\delta x$ measures ε after magnification.



When we focus a powerful microscope at the point $(x,\ f[x])$ we only see the linear curve $\mathrm{dy} = m \cdot \mathrm{dx}$, because $\varepsilon \approx 0$ is smaller than the thickness of the line. The figure below shows a small box magnified on the right.

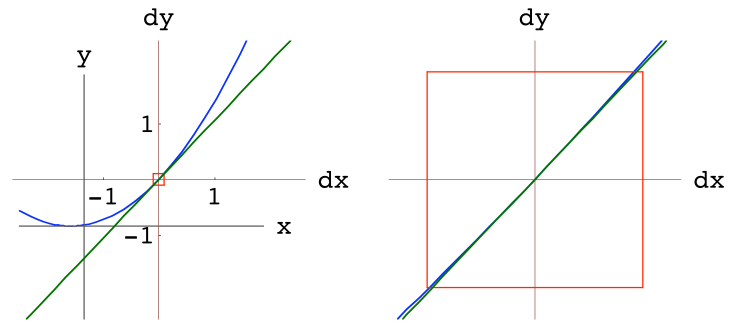


Figure 1.1.1: A Magnified Tangent

The Fundamental Theorem of Integral Calculus

Now we use the intuitive microcsope approximation (1.1.1) to prove:

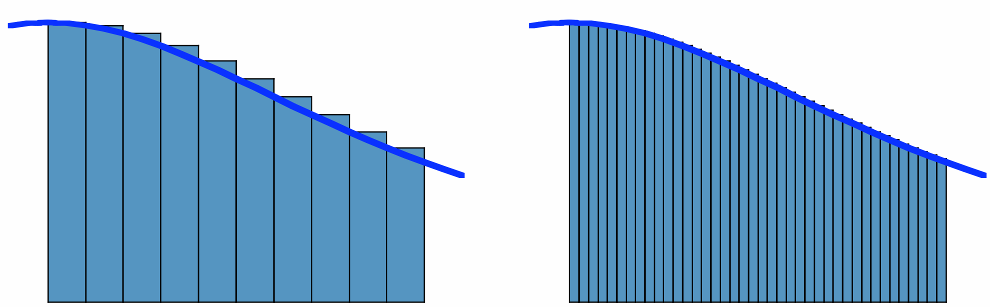
Theorem: The Fundamental Theorem of Integral Calculus: Part 1

Suppose we want to find $\int_a^b f[x]\, dx$. If we can find another function $F[x]$ so that the differential gives $dF[x] = F'[x]\,\mathrm{dx} = f[x]\,\mathrm{dx}$ for every x, $a \le x \le b$, then

$$\int_a^b f[x]\, dx = F[b] - F[a]$$

The definition of the integral we use is the real number approximated by a sum of small slices,

$$\int_a^b f[x]\, dx \approx \sum_{\substack{x=a \\ \text{step } \delta\mathrm{x}}}^{b-\delta\mathrm{x}} f[x] \cdot \delta\mathrm{x} \quad \text{,when } \delta\mathrm{x} \approx 0$$



Telescoping sums & derivatives

We know that if $F[x]$ has derivative $F'[x] = f[x]$, the differential approximation above says,

$$F[x + \delta\mathrm{x}] - F[x] = f[x] \cdot \delta\mathrm{x} + \varepsilon \cdot \delta\mathrm{x}$$

so we can sum both sides

$$\sum_{\substack{x=a \\ \text{step } \delta\mathrm{x}}}^{b-\delta\mathrm{x}} F[x + \delta\mathrm{x}] - F[x] = \sum_{\substack{x=a \\ \text{step } \delta\mathrm{x}}}^{b-\delta\mathrm{x}} f[x] \cdot \delta\mathrm{x} + \sum_{\substack{x=a \\ \text{step } \delta\mathrm{x}}}^{b-\delta\mathrm{x}} \varepsilon \cdot \delta\mathrm{x}$$

The telescoping sum satisfies,

$$\sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} F[x+\delta x]-F[x]=F[b']-F[a]$$

so we obtain the approximation,

$$\int_a^b f[x]\,dx \approx \sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} f[x]\cdot\delta x \;=\; F[b']-F[a]-\;\sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} \varepsilon\cdot\delta x$$

This gives,

$$\left|\sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} f[x]\cdot\delta x \;-(F[b']-F[a])\right| \leq \left|\sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} \varepsilon\cdot\delta x\right| \leq \sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} |\varepsilon|\cdot\delta x$$

$$\leq \text{Max}[\,|\varepsilon|]\cdot \sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} \delta x = \text{Max}[\,|\varepsilon|]\cdot(b'-a)\approx 0$$

or $\int_a^b f[x]\,dx \approx \sum_{\substack{x=a\\ \text{step } \delta x}}^{b-\delta x} f[x]\cdot\delta x \approx F[b']-F[a]$. Since $F[x]$ is continuous, $F[b']\approx F[b]$, so $\int_a^b f[x]\,dx = F[b]-F[a]$.

We need to know that all the epsilons above are small when the step size is small, $\varepsilon\approx 0$, when $\delta x\approx 0$ for all $x=a,\ a+\delta x,\ a+2\,\delta x,\ \cdots$. This is a uniform condition that has a simple appearance in Robinson's theory. There is something to explain here because the theorem stated above is false if we take the usual pointwise notion of derivative and the Reimann integral. (There are pointwise differentiable functions whose derivative is not Riemann integrable.)

The condition needed to make this proof complete is natural geometrically and plays a role in the intuitive proof of the inverse function theorem in the next section.

Continuity of the Derivative

We show now that the differential approximation

$$\boxed{f[x+\delta x]-f[x]=f'[x]\cdot\delta x+\varepsilon\cdot\delta x}$$

forces the derivative function $f'[x]$ to be continuous,

$$\boxed{x_1\approx x_2 \Longrightarrow f'[x_1]\approx f'[x_2]}$$

Let $x_1\approx x_2$, but $x_1\neq x_2$. Use the differential approximation with $x=x_1$ and $\delta x=x_2-x_1$ and also with $x=x_2$ and $\delta x=x_1-x_2$, geometrically, looking at the tangent approximation from both endpoints.

$$f[x_2]-f[x_1]=f'[x_1]\cdot(x_2-x_1)+\varepsilon_1\cdot(x_2-x_1)$$

$$f[x_1]-f[x_2]=f'[x_2]\cdot(x_1-x_2)+\varepsilon_2\cdot(x_1-x_2)$$

Adding, we obtain

$$0=((f'[x_1]-f'[x_2])+(\varepsilon_1-\varepsilon_2))\cdot(x_2-x_1)$$

Dividing by the nonzero term (x_2-x_1) and adding $f'[x_2]$ to both sides, we obtain,

$f'[x_2]=f'[x_1]+(\varepsilon_1-\varepsilon_2)$ or $f'[x_2]\approx f'[x_1]$, since the difference between two small errors is small.

This fact can be used to prove:

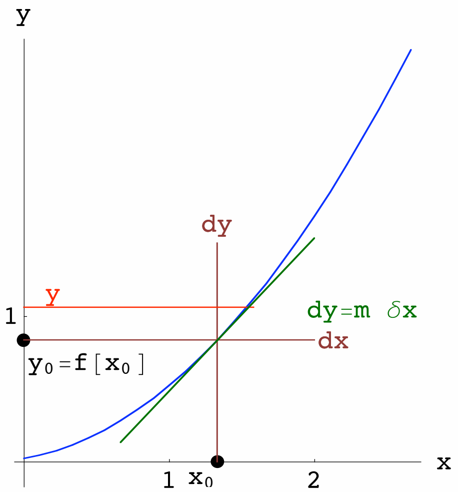
Theorem: The Inverse Function Theorem

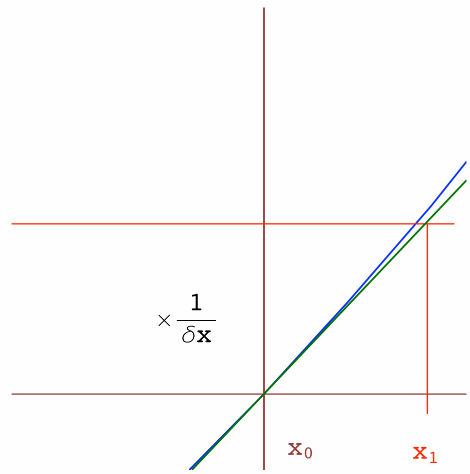
If $f'[x_0]\neq 0$ then $f[x]$ has an inverse function in a small neighborhood of x_0, that is, if $y\approx y_0=f[x_0]$, then there is a unique $x\approx x_0$ so that $y=f[x]$.

We saw above that the differential approximation makes a microscopic view of the graph look linear. If $y \approx y_1$ the linear equation $\mathrm{dy} = m \cdot \mathrm{dx}$ with $m = f'[x_1]$ can be inverted to find a first approximation to the inverse,

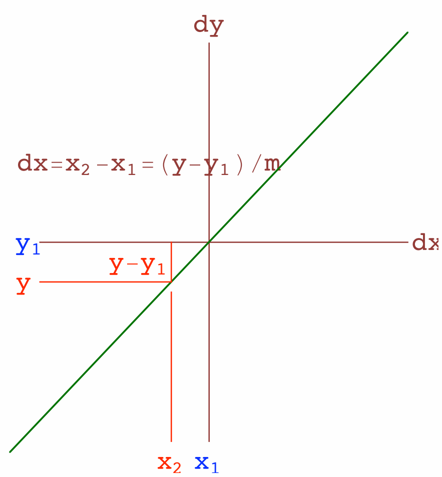
$$y - y_0 = m \cdot (x_1 - x_0)$$

$$x_1 = x_0 + \frac{1}{m}(y - y_0)$$





We test to see if $f[x_1] = y$. If not, examine the graph microscopically at $(x_1, y_1) = (x_1, f[x_1])$. Since the graph appears the same as it's tangent to within ε and since $m = f'[x_1] \approx f'[x_2]$, the local coordinates at (x_1, y_1) look like a line of slope m:



Solving for the x-value which gives output y, we get

$$y - y_1 = m \cdot (x_2 - x_1)$$

$$x_2 = x_1 + \frac{1}{m}(y - f[x_1])$$

Continue in this way generating a sequence of approximations, $x_1 = x_0 + \frac{1}{m}(y - y_0)$, $x_{n+1} = G[x_n]$, where the recursion function $G[\xi] = x + \frac{1}{m}(y - f[\xi])$. The distance between successive approximations is

$$|x_2 - x_1| = |G[x_1] - G[x_0]| \leq |G'[\xi_1]| \cdot |x_1 - x_0|$$

$$|x_3 - x_2| = |G[x_2] - G[x_1]| \leq |G'[\xi_2]| \cdot |x_2 - x_1| \leq |G'[\xi_2]| \cdot |G'[\xi_1]| \cdot |x_1 - x_0|$$

by the Mean Value Theorem for Derivatives. Notice that $G'[\xi] = 1 - f'[\xi]/m \approx 0$, for $\xi \approx x_0$, so $|G'[\xi]| < 1/2$ in particular, and

$$|x_2 - x_1| = |G[x_1] - G[x_0]| \leq \frac{1}{2} \cdot |x_1 - x_0|$$

$$|x_3 - x_2| = |G[x_2] - G[x_1]| \leq \frac{1}{2} \cdot |x_2 - x_1| \leq \frac{1}{2^2} \cdot |x_1 - x_0|$$

$$\vdots$$

$$|x_{n+1} - x_n| \leq \frac{1}{2^n} \cdot |x_1 - x_0|$$

$$|x_{n+1} - x_0| \leq |x_{n+1} - x_n| + |x_n - x_{n-1}| + \cdots + |x_1 - x_0| \leq (1 + \frac{1}{2} + \cdots + \frac{1}{2^n}) \cdot |x_1 - x_0|$$

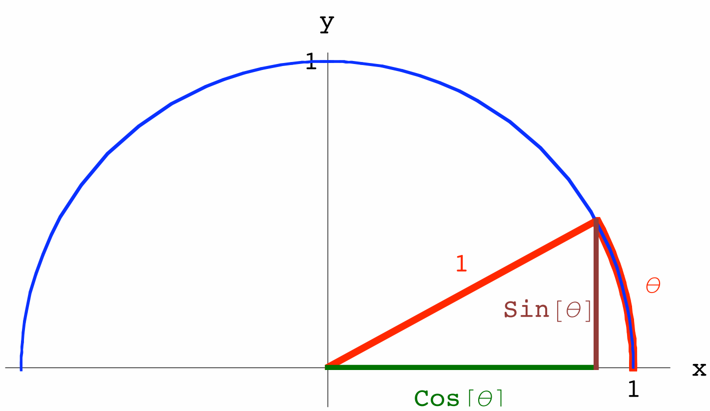
A geometric series estimate shows that the series converges, $x_n \to x \approx x_0$ and $f[x] = y$.

To complete this proof we need to show that $G[\xi]$ is a contraction on some nonzero interval. The function $G[\xi]$ must map the interval into itself and have derivative less than $1/2$ on the interval. The precise definiton of the derivative matters because the result is false if $f'[x]$ is defined by a pointwise limit. The function $f[x] = x + x^2 \text{Sin}[\pi/x]$ with $f[0] = 0$ has pointwise derivative 1 at zero, but is not increasing in any neighborhood of zero. (If you "move the microscope an infinitesimal amount" when looking at $y = x^2 \text{Sin}[\pi/x] + x$, the graph will look nonlinear.)

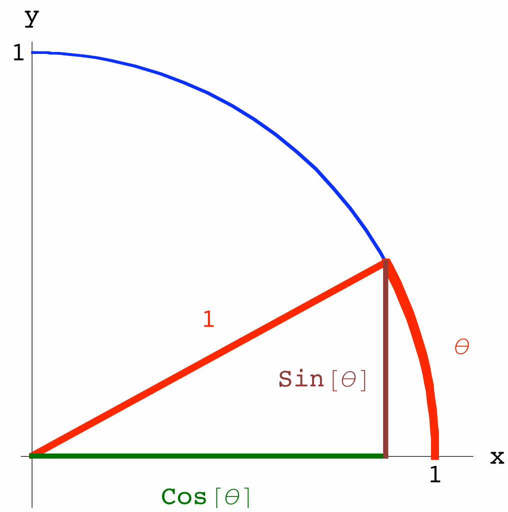
Trig, Polar Coordinates & Holditch's Formula

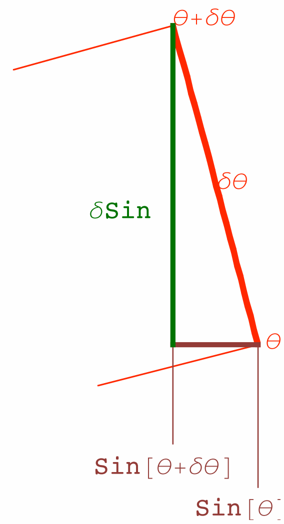
With hindsight, Robinson's solution to the old problem of rigorous use of infinitesimal numbers in calculus boils down to Keisler's Funcrtion Extension Axiom that we discuss in Section 2. Extending the (Dedekind) "real" numbers to include infinitesimals algebraically is not at all difficult, but calculus depends on small approximations with transcendental functions like sine, cosine, and the natural logarithm. Following are some intuitive approximations with non-algebraic graphs.

Sine and cosine in radian measure give the x and y location of a point on the unit circle measured a distance θ along the circle as shown next.



Now make a small change in the angle and magnify by $1/\delta\theta$ to make the change appear unit size.





Since magnification does not change lines, the radial segments from the origin to the tiny increment of the circle meet the circle at right angles and appear under magnification to be parallel lines. Smoothness of the circle means that under powerful magnification, the circular increment appears to be a line. The difference in values of the sine is the long vertical leg of the increment "triangle" above on the right. The apparant hypotenuse with length $\delta\theta$ is the circular increment.

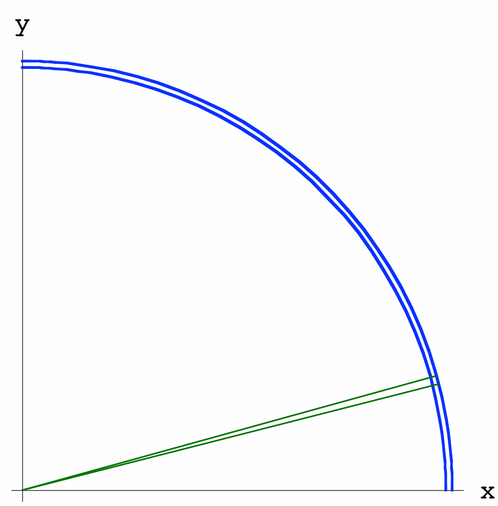
Since the radial lines meet the circle at right angles the large triangle on the unit circle at the left with long leg Cos[θ] and hypotenuse 1 is similar to the increment triangle, giving

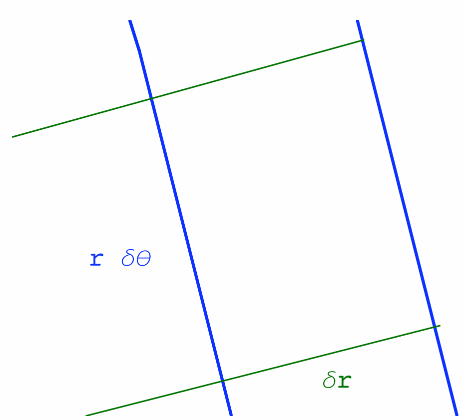
$$\frac{\mathrm{Cos}[\theta]}{1} \approx \frac{\delta\mathrm{Sin}}{\delta\theta}$$

We write approximate similarity because the increment "triangle" actually has one circular side that is ≈-straight. In any case, this is a convincing argument that $\frac{d\mathrm{Sin}}{d\theta} = \mathrm{Cos}[\theta]$. A similar geometric argument on the increment triangle shows that $\frac{d\mathrm{Cos}}{d\theta} = -\mathrm{Sin}[\theta]$.

The Polar Area Differential

The derivation of sine and cosine is related to the area differential in polar coordinates. If we take an angle increment of $\delta\theta$ and magnify a view of the circular arc on a circle of radius r, the length of the circular increment is $r \cdot \delta\theta$, by similarity.





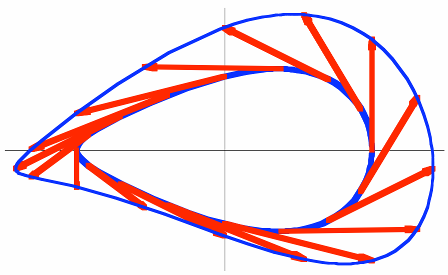
A magnified view of circles of radii r and $r + \delta r$ between the rays at angles θ and $\theta + \delta\theta$ appears to be a rectangle with sides of lengths δr and $r \cdot \delta\theta$. If this were a true rectangle, its area would be $r \cdot \delta\theta \cdot \delta \mathrm{r}$, but it is only an approximate rectangle. Technically, we can show that the area of this region is $r \cdot \delta\theta \cdot \delta \mathrm{r}$ plus a term that is small compared with this infinitesimal,

$$\delta \mathrm{A} = r\, \delta\theta\, \delta \mathrm{r} + \varepsilon \cdot \delta\theta \delta \mathrm{r}$$

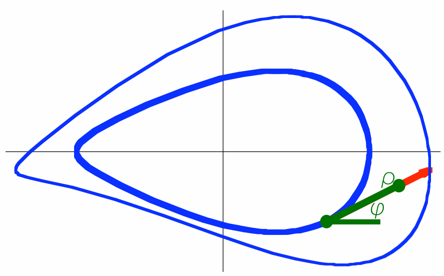
Keisler's Infinite Sum Theorem assures us that we can neglect this size error and integrate with respect to $r\, d\theta\, \mathrm{dr}$.

Holditch's Formula

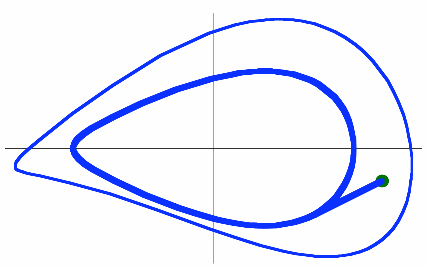
The area swept out by a tangent of length R as it traverses an arbitrary convex curve in the plane is $A = \pi R^2$.

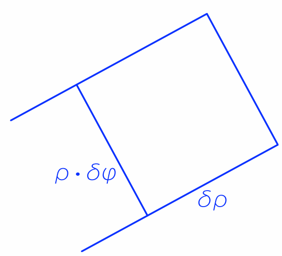


We can see this interesting result by using a variation of polar coordinates and the infinitesimal polar area increment above. Since the curve is convex, each tangent meets the curve at a unique angle, φ, and each point in the region swept out by the tangents is a distance ρ along that tangent.

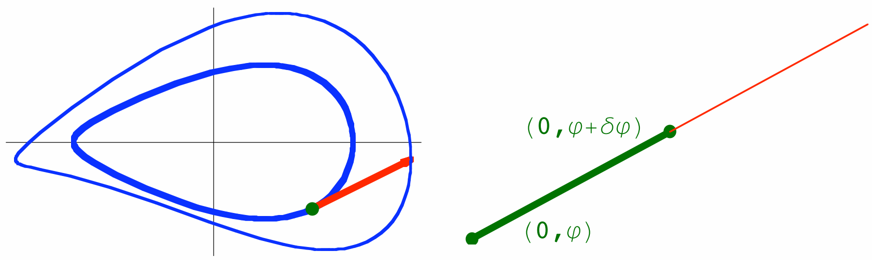


We look at an infinitesimal increment of the region in ρ-φ-coordinates, first holding the the φ-base point on the curve and changing φ. Microscopically this produces an increment like the polar increment:

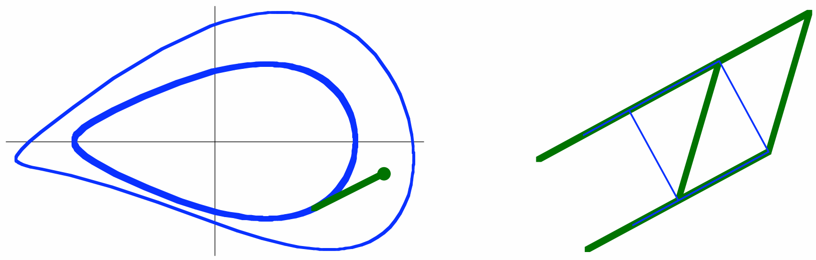




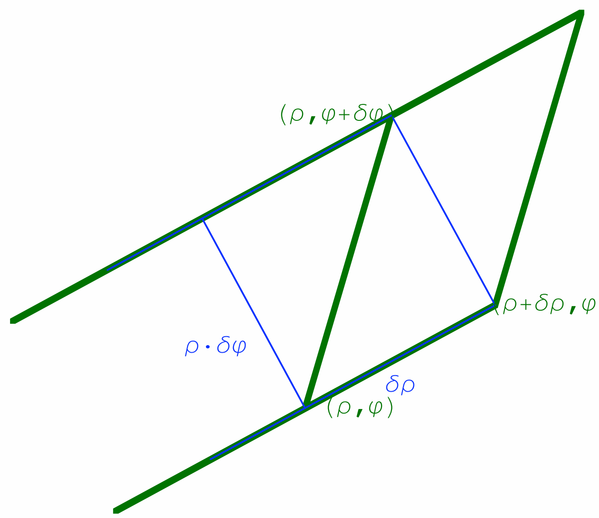
Next, looking at the base point of the tangent on the curve, moving to the correct $\varphi + \delta\varphi$-base point, moves along the curve. Microscopically this looks like translation along the tangent line (by smoothness):



Including this near-translation in the infinitesimal area increment produces a parallelogram:



of height $\rho \cdot \delta\varphi$ and base $\delta\rho$, or area $\delta\text{A} = \rho \cdot \delta\varphi \cdot \delta\rho$:



Integrating gives the total area of the region

$$\int_0^R \int_0^{2\pi} \rho \, d\varphi \, d\rho = \pi R^2$$

Leibniz's Formula for Radius of Curvature

The radius r of a circle drawn through three infinitely nearby points on a curve in the (x, y)-plane satisfies

$$\frac{1}{r} = -\frac{d}{dx}\left(\frac{dy}{ds}\right)$$

where s denotes the arclength. For example, if $y = f[x]$, so $ds = \sqrt{1 + (f'[x])^2}\; dx$, then

$$\frac{1}{r} = -\frac{d}{dx}\left(\frac{f'[x]}{\sqrt{1+(f'[x])^2}}\right) = -\frac{f''[x]}{(1+f'[x]^2)^{3/2}}$$

If the curve is given parametrically, $y = y[t]$ and $x = x[t]$, so $ds = \sqrt{x'[t]^2 + y'[t]^2}\; dt$, then

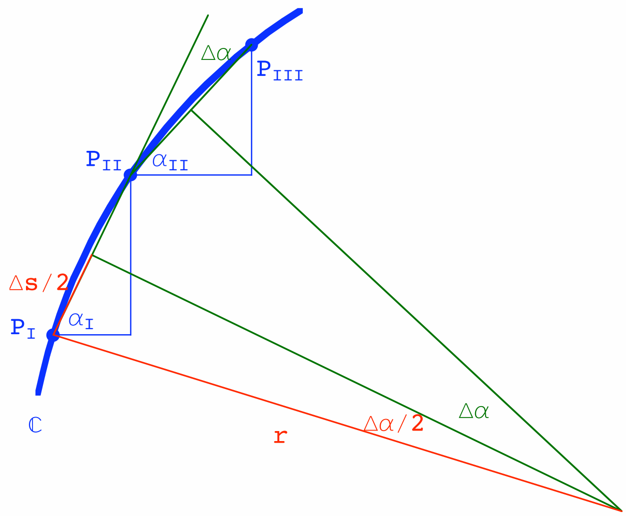
$$\frac{1}{r} = -\frac{d\left(\frac{dy}{ds}\right)}{dx} = \frac{y'[t]\,x''[t] - x'[t]\,y''[t]}{\left(x'[t]^2 + y'[t]^2\right)^{3/2}}$$

Changes

Consider three points on a curve $\mathbb{C}$ with equal distances Δs between the points. Let α_{I} and α_{II} denote the angles between the horizontal and the segments connecting the points as shown. We have the relation between the changes in y and α:

$$\mathrm{Sin}[\alpha] = \frac{\Delta \mathrm{y}}{\Delta \mathrm{s}} \tag{1.1.2}$$

The difference between these angles, $\Delta\alpha$, is shown near p_{III}.



The angle between the perpendicular bisectors of the connecting segments is also $\Delta\alpha$, because they meet the connecting segments at right angles.

These bisectors meet at the center of a circle through the three points on the curve whose radius we denote r. The small triangle with hypotenuse r gives

$$\mathrm{Sin}\left[\frac{\Delta\alpha}{2}\right] = \frac{\Delta \mathrm{s}/2}{r} \tag{1.1.3}$$

Small Changes

Now we apply these relations when the distance between the successive points is an infinitesimal δs. The change

$$-\delta\,\mathrm{Sin}[\alpha] = -\delta\left(\frac{\delta \mathrm{y}}{\delta \mathrm{s}}\right) = \mathrm{Sin}[\alpha] - \mathrm{Sin}[\alpha - \delta\alpha] = \mathrm{Cos}[\alpha]\,\cdot\delta\alpha + \vartheta\cdot\delta\alpha, \tag{1.1.4}$$

with $\vartheta \approx 0$, by smoothness of sine (see above). Smoothness of sine also gives,

$$\mathrm{Sin}\left[\frac{\delta\alpha}{2}\right] = \frac{\delta\alpha}{2} + \eta\cdot\delta\alpha, \text{ with } \eta \approx 0$$

Combining this with formula (1.1.3) for the infinitesimal case (assuming $r \neq 0$), we get

$$\delta\alpha = \frac{\delta \mathrm{s}}{r} + \iota\cdot\delta\alpha, \text{ with } \iota \approx 0$$

Now substitute this in (1.1.4) to obtain

$$-\delta\left(\frac{\delta \mathrm{y}}{\delta \mathrm{s}}\right) = \mathrm{Cos}[\alpha]\,\frac{\delta \mathrm{s}}{r} + \zeta\cdot\delta \mathrm{s}, \text{ with } \zeta \approx 0$$

By trigonometry, $\mathrm{Cos}[\alpha] = \delta\mathrm{x}/\delta\mathrm{s}$, so

$$-\frac{\delta\left(\frac{\delta \mathrm{y}}{\delta \mathrm{s}}\right)}{\delta \mathrm{x}} = \frac{1}{r} + \zeta\cdot\frac{\delta \mathrm{s}}{\delta \mathrm{x}} \approx \frac{1}{r}, \text{ as long as } \frac{\delta \mathrm{s}}{\delta \mathrm{x}} \text{ is not infinitely large.}$$

Keisler's Function Extension Axiom allows us to apply formulas (1.1.3) and (1.1.4) when the change is infinitesimal, as we shall see. We still have a gap to fill in order to know that we may replace infinitesimal differences with differentials (or derivatives), especially because we have a difference of a quotient of differences.

First differences and derivatives have a fairly simple rigorous version in Robinson's theory, just using the differential approximation (1.1.1). This can be used to derive many classical differential equations like the tractrix, catenary, and isochrone, see: Chapter 5 Differenital Equations from Increment Geometry in Projects for Calculus: The Language of Change on my website at http://www.math.uiowa.edu/%7Estroyan/ProjectsCD/estroyan/indexok.htm

Second differences and second derivatives have a complicated history. See

H. J. M. Bos, **Differentials, Higher-Order Differentials and the Derivative in the Leibnizian Calculus**, *Archive for History of Exact Sciences*, vol. 14, nr. 1, 1974.

This is a very interesting paper that begins with a course in calculus as Leibniz might have presented it.

The natural exponential

The natural exponential function satisfies

$$y[0] = 1$$

$$\frac{dy}{dx} = y$$

We can use (1.1.1) to find an approximate solution,

$$y[\delta\text{x}] = y[0] + y'[0] \cdot \delta\text{x} = 1 + \delta\text{x}$$

Recursively,

$$y[2\,\delta\text{x}] = y[\delta\text{x}] + y'[\delta\text{x}] \cdot \delta\text{x} = y[\delta\text{x}] \cdot (1 + \delta\text{x}) = (1 + \delta\text{x})^2$$

$$y[3\,\delta\text{x}] = y[2\,\delta\text{x}] + y'[2\,\delta\text{x}] \cdot \delta\text{x} = y[2\,\delta\text{x}] \cdot (1 + \delta\text{x}) = (1 + \delta\text{x})^3$$

$$\vdots$$

$$y[x] = (1 + \delta\text{x})^{x/\delta\text{x}}, \text{ for } x = 0,\ \delta\text{x},\ 2\,\delta\text{x},\ 3\,\delta\text{x},\ \cdots$$

This is the product expansion $e \approx (1 + \delta\text{x})^{1/\delta\text{x}}$, for $\delta\text{x} \approx 0$.

No introduction to calculus is complete without mention of this sort of "infinite algebra" as championed by Euler as in

L. Euler, *Introductio in Analysin Infinitorum*, Tomus Primus, Lausanne, 1748. Reprinted as L. Euler, Opera Omnia, ser. 1, vol. 8. Translated from the Latin by J. D. Blanton, *Introduction to Analysis of the Infinite*, Book I, Springer–Verlag, New York, 1988.

A wonderful modern interpretaion of these sorts of computations is in

Mark McKinzie and Curtis Tuckey, Higher Trigonometry, Hyperreal Numbers and Euler's Analysis of Infinities, *Math Magazine*, vol. 74, nr. 5, Dec. 2001, p. 339-368

W. A. J. Luxemburg's reformulation of the proof of one of Euler's central formulas

$$\text{Sin}[z] = z \prod_{k=1}^{\infty} \left(1 - \left(\frac{z}{k\,\pi}\right)^2\right)$$

appears in our monograph, *Introduction to the Theory of Infinitesimals,* Academic Press Series on Pure and Applied Math. vol 72, 1976, Academic Press, New York.

Concerning the History of the Calculus

Chapter X of Robinson's monograph

Non-standard Analysis, North-Holland Publishing Co., Amsterdam, 1966. Revised edition by Princeton University Press, Princeton, 1996, begins:

> *The history of a subject is usually written in the light of later developments. For over half a century now, accounts of the history of the Differential and Integral Calculus have been based on the belief that even though the idea of a number system containing infinitely small and infinitely large elements might be consistent, it is useless for the development of Mathematical Analysis. In consequence, there is in the writings of this period an noticable contrast between the severity with which the ideas of Leibniz and his successors are treated and the leniency accorded to the lapses of the early proponents of the doctrine of limits. We do not propose here to subject any of these works to a detailed criticism. However, it will serve as a starting point for our discussion to try to give a fair summary of the contemporary impression of the history of the Calculus...*

I recomend that you read Robinson's Chapter X. I have often wondered if mathematicians in the time of Weierstrass said things like, 'Karl's epsilon-delta stuff isn't very interesting. All he does is re-prove old formulas of Euler.'

I have a non-standard interest in the history of infinitesimal calculus. It really is not historical. Some of the old derivations like Bernoulli's derivation of Leibniz' formula for the radius of curvature seem to me to have a compelling clarity. Robinson's theory of infiitesimals offers me an opportunity to see what is needed to complete these arguments with a contemporary standard of rigor.

Working on such problems has led me to believe that the best theory to use to underly calculus when we present it to beginners is one based on the kind of derivatives described in Section 2 and not the pointwise approach that is the current custom in the U.S. I believe we want a theory that supports intuitive reasoning like the examples above and pointwise defined derivatives do not.