The University of Sydney

MOOC Introduction to Calculus

Notes for 'Finding derivatives from first principles'

Important Ideas and Useful Facts:

(i) The derivative of a linear function: Let $f(x) = mx + k$ where m and k are constants be a linear function, whose graph is a line. Then

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{m(x+h) + k - (mx+k)}{h} \\ &= \lim_{h\to 0} \frac{mx + mh + k - mx - k}{h} = \lim_{h\to 0} \frac{mh}{h} = \lim_{h\to 0} m = m , \end{aligned}$$

which is just the slope m of the associated line. This is to be expected, because a line coincides with its tangent line at every point.

(ii) The derivative is additive: The derivative is *additive* and *respects addition* in the following sense: if $y = g(x)$ and $y = h(x)$ are functions with derivative $g'(x)$ and $h'(x)$ respectively, and $f(x) = g(x) + h(x)$ then

$$f'(x) = g'(x) + h'(x) .$$

This follows quickly from either of the limit definitions and the fact that limits respect addition. For example,

$$\begin{aligned} f'(a) &= \lim_{b\to a} \frac{f(b) - f(a)}{b-a} = \lim_{b\to a} \frac{g(b) + h(b) - \big(g(a) + h(a)\big)}{b-a} \\ &= \lim_{b\to a} \frac{g(b) - g(a) + h(b) - h(a)}{b-a} = \left(\lim_{b\to a} \frac{g(b) - g(a)}{b-a} \right) + \left(\lim_{b\to a} \frac{h(b) - h(a)}{b-a} \right) \\ &= g'(a) + h'(a) , \end{aligned}$$

and the result follows, taking $x = a$.

For example, if $f(x) = x^3 + x^2$ then $f'(x) = 3x^2 + 2x$.

(iii) Constants come out the front: The derivative *respects constant multiples* in the following sense: if $y = g(x)$ with derivative $g'(x)$ and $f(x) = kg(x)$ for some constant k then

$$f'(x) = kg'(x) .$$

This follows quickly from either of the limit definitions and the fact that limits respect constant multiples. For example,

$$f'(a) = \lim_{b\to a} \frac{f(b) - f(a)}{b-a} = \lim_{b\to a} \frac{kg(b) - kg(a)}{b-a} = k\left(\lim_{b\to a} \frac{g(b) - g(a)}{b-a} \right) = kg'(a) ,$$

and the result follows, taking $x = a$.

For example, if $f(x) = 10x^3$ then $f'(x) = (10)(3x^2) = 30x^2$.

(iv) The derivative of a polynomial: From the derivatives of power functions (discussed previously) and the fact that derivatives respect addition and constant multiples, we are able to quickly obtain the derivative of any polynomial function

$$p(x) \;=\; a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_{n-1} x^{n-1} + a_n x^n$$

of degree $n \geq 1$ where $a_0, \ldots, a_n$ are constants, namely,

$$p'(x) \;=\; a_1 + 2a_2 x + 3a_3 x^2 + \ldots (n-1)a_{n-1} x^{n-2} + n a_n x^{n-1} \;,$$

which is a polynomial of degree $n-1$, that is, of degree one less than that of $p(x)$.

(v) Second, third and higher order derivatives: Let $y = f(x)$. The derivative of f is denoted by $y' = f'(x)$, and is also called the *first derivative*. The *second derivative* is the derivative $(y')'$ of the (first) derivative and denoted more simply just by

$$y'' \;=\; f''(x) \;.$$

The *third derivative* is the derivative $(y'')'$ of the second derivative and denoted by

$$y''' \;=\; f'''(x) \;.$$

We can continue differentiating to get *higher-order derivatives*. Because of the proliferation of dashes, it is conventional to denote the result of differentiating n times by

$$y^{(n)} \;=\; f^{(n)}(x) \;,$$

which is an abbreviation for using n dashes. For example, $y''' = y^{(3)}$ and $y'''' = y^{(4)}$, $y''''' = y^{(5)}$, and so on.

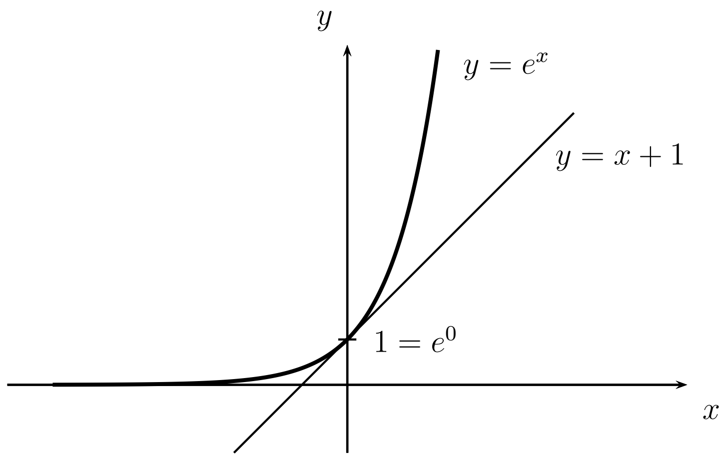
If $p(x)$ is a polynomial of degree $n \geq 1$ then, as noted above, $p'(x)$ is a polynomial of degree $n-1$. It follows that the higher order derivative $p^{(k)}$, the result of differentiating the polynomial $k \leq n$ times, is a polynomial of degree $n-k$. In particular $p^{(n)}$ is a constant polynomial (of degree 0), so one further differentiation will produce zero. This shows that, by differentiating a polynomial sufficiently often, eventually the higher-order derivatives can be made to "disappear" in the sense of becoming zero.

For example, if $p(x) = x^2 + 6x + 5$, then $p'(x) = 2x + 6$, $p''(x) = 2$ and $p'''(x) = 0$. Differentiating any quadratic (of degree two) three times produces zero.

(vi) The derivative of the exponential function is itself: Consider the natural exponential function $y = f(x) = e^x$. We claim that

$$y' = f'(x) = e^x \;.$$

Recall that Euler's number e was chosen so that the slope of the tangent line to the curve at the y-intercept has slope 1.



But, with our derivative notation, this slope is just $f'(0)$. Hence, we have that

$$1 \;=\; f'(0) \;=\; \lim_{h\to 0}\frac{f(0+h)-f(0)}{h} \;=\; \lim_{h\to 0}\frac{e^{0+h}-e^0}{h} \;=\; \lim_{h\to 0}\frac{e^h-1}{h}\ .$$

Remarkably, we can use this to prove that $f'(x) = e^x$ for all x. Using an exponential law at the third step, and the above limit at the second last step, we get

$$f'(x) \;=\; \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \;=\; \lim_{h\to 0}\frac{e^{x+h}-e^x}{h} \;=\; \lim_{h\to 0}\frac{e^xe^h-e^x}{h} \;=\; \lim_{h\to 0}\frac{e^x(e^h-1)}{h}$$

$$=\; e^x\left(\lim_{h\to 0}\frac{e^h-1}{h}\right) \;=\; e^x(1) \;=\; e^x\ ,$$

as required. Thus the natural exponential function reproduces itself by differentiation, and coincides with all of its higher derivatives, by contrast with polynomial functions, which eventually become zero by repeated differentiation.

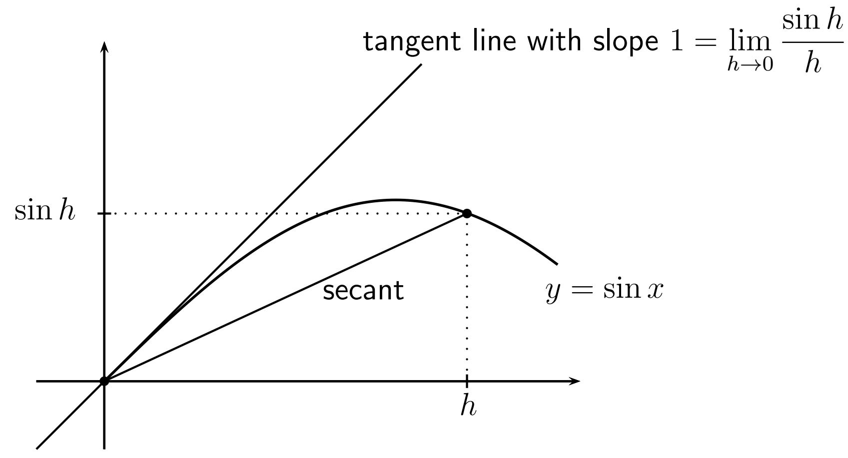
(vii) The derivative of the sine function is the cosine function: Consider $y = f(x) = \sin x$. We claim that

$$y' \;=\; f'(x) \;=\; \cos x\ .$$

Before proving this, first observe that the slope of the tangent line to the sine curve at the origin is 1 (which coincides with $\cos(0)$). This follows because of the fact

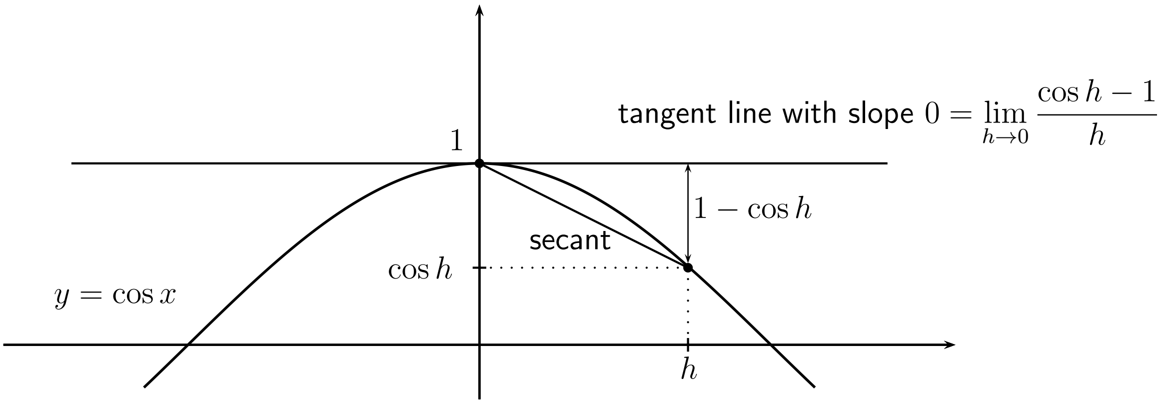
$$\lim_{h\to 0}\frac{\sin h}{h} \;=\; 1\ ,$$

which we carefully proved using the squeeze law, but stated here using h instead of x, and because the ratio $\frac{\sin h}{h}$ is the slope of the secant joining the origin to the point on the curve for $x = h$.



Next observe that the slope of the cosine curve at $x = 0$ is zero. This is the case, because the smooth cosine curve turns around and has a peak at the point $(0, 1)$, so that the tangent line must be sitting flat, that is, horizontal.

If one joins the point $(0, 1)$ to some nearby point $(h, \cos h)$ to form a secant, then it has (negative) slope $\frac{\cos h - 1}{h}$, which must therefore approach 0 as h approaches 0.



We therefore get the following limit:

$$\lim_{h\to 0} \frac{\cos h - 1}{h} \;=\; 0 \;.$$

We can now find the derivative of $\sin x$ as follows, using these two special limits at the second last step, after expanding the numerator at the third step using the formula for the sine of a sum of angles (explained later in these notes), and using our usual tricks for rearranging fractions and manipulating limits:

$$\begin{aligned}
f'(x) \;&=\; \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \;=\; \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} \\
&=\; \lim_{h\to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \;=\; \lim_{h\to 0} \frac{\sin x(\cos h - 1) + \cos x \sin h}{h} \\
&=\; \sin x\left(\lim_{h\to 0} \frac{\cos h - 1}{h}\right) + \cos x\left(\lim_{h\to 0} \frac{\sin h}{h}\right) \\
&=\; (\sin x)(0) + (\cos x)(1) \;=\; \cos x \;,
\end{aligned}$$

as required.

(viii) The derivative of the cosine function is the negative sine function: Consider $y = f(x) = \cos x$. We claim that

$$y' \;=\; f'(x) \;=\; -\sin x \;.$$

We again use the two special limits that were used in the previous proof and the same tricks for manipulating fractions and limits, but now, at the third step, use the formula for the cosine of a sum of angles (explained later in these notes):

$$\begin{aligned}
f'(x) \;&=\; \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \;=\; \lim_{h\to 0} \frac{\cos(x+h) - \cos x}{h} \\
&=\; \lim_{h\to 0} \frac{\cos x \cos h - \sin x \sin h - \cos x}{h} \;=\; \lim_{h\to 0} \frac{\cos x(\cos h - 1) - \sin x \sin h}{h} \\
&=\; \cos x\left(\lim_{h\to 0} \frac{\cos h - 1}{h}\right) - \sin x\left(\lim_{h\to 0} \frac{\sin h}{h}\right) \\
&=\; (\cos x)(0) - (\sin x)(1) \;=\; -\sin x \;,
\end{aligned}$$

as required.

(ix) The formula for the sine of a sum of angles: We sketch a proof of the following formula:

$$\sin(\alpha+\beta) \;=\; \sin\alpha\cos\beta + \sin\beta\cos\alpha\ .$$

Sketch of proof: This is only a sketch, because it relies on the particular diagram below that assumes both α and β are first quadrant angles. (To get a complete proof, one has to also explain how to reduce the problem for arbitrary angles to this case.)

Consider a right-angled triangle with vertices P, Q and R and angle α at P. We add a point T to produce another right-angled triangle with vertices P, R and T and angle β at P. Hence the angle at P of the larger triangle with vertices P, Q and T is $\alpha+\beta$.

We add a point S, extending the line through Q and R such that the line segment joining S to T is parallel to the line segment joining P and Q. Thus the triangle with vertices R, S and T is right-angled with angle α at R.

We now scale the diagram so that the length of the line segment joining P to T is 1 unit.

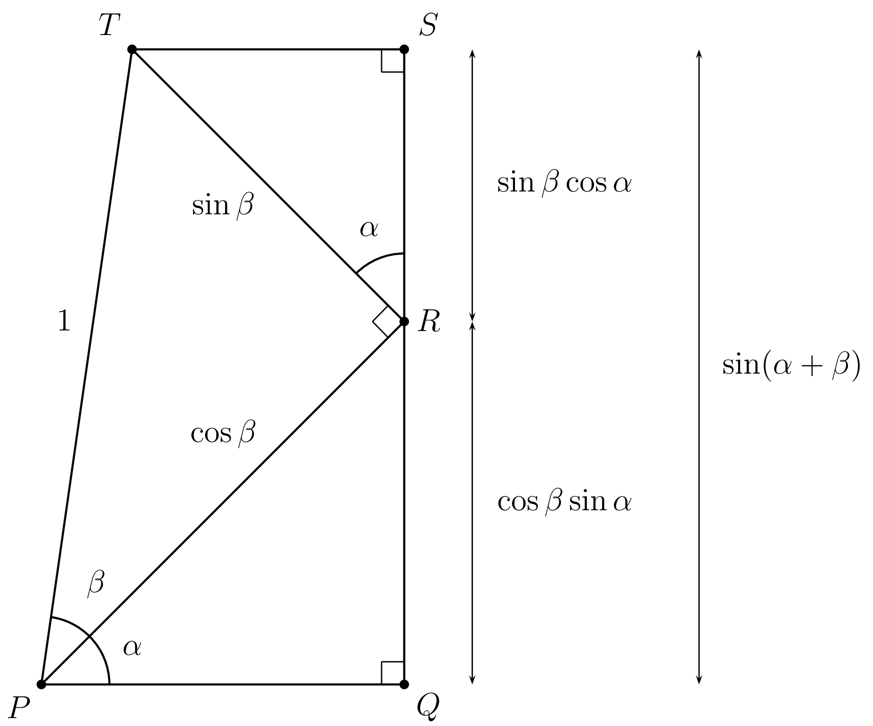
Note that the altitude of T above the horizontal line segment joining P to Q is the sine of the combined angle $\alpha+\beta$, and this is the length of the line segment joining Q to S.

Further, the lengths of the line segments joining P to R and R to T are $\cos\beta$ and $\sin\beta$ respectively.

Hence the lengths of the line segments joining Q to R and R to S are $\cos\beta\sin\alpha$ and $\sin\beta\cos\alpha$, respectively. Adding these up gives $\sin(\alpha+\beta)$, which is

$$\cos\beta\sin\alpha + \sin\beta\cos\alpha \;=\; \sin\alpha\cos\beta + \sin\beta\cos\alpha\ ,$$

completing the sketch of the proof.



(x) The formula for the cosine of a sum of angles: We use the formula for the sine of a sum of angles to deduce the following formula:

$$\cos(\alpha+\beta) \;=\; \cos\alpha\cos\beta - \sin\alpha\sin\beta\ .$$

Proof: Recall, from when we introduced the graphs of $y = \sin x$ and $y = \cos x$, that the curves are both sinusoidal and match up perfectly by a horizontal translation. Algebraically, this translation is captured by the following equation, for all angles x:

$$\cos x \;=\; \sin\left(x+\frac{\pi}{2}\right)\ .$$

Thus, for any angles α and β, using this equation and the formula for the sine of a sum of angles repeatedly,

$$\begin{aligned}
\cos(\alpha+\beta) \;&=\; \sin\left(\alpha+\beta+\frac{\pi}{2}\right) \;=\; \sin\alpha\cos\left(\beta+\frac{\pi}{2}\right)+\sin\left(\beta+\frac{\pi}{2}\right)\cos\alpha \\
&=\; \sin\alpha\sin(\beta+\pi)+\left(\sin\beta\cos\left(\frac{\pi}{2}\right)+\sin\left(\frac{\pi}{2}\right)\cos\beta\right)\cos\alpha \\
&=\; \sin\alpha\big(\sin\beta\cos\pi+\sin\pi\cos\beta\big)+\big((\sin\beta)(0)+(1)(\cos\beta)\big)\cos\alpha \\
&=\; \sin\alpha\big((\sin\beta)(-1)+(0)(\cos\beta)\big)+\cos\beta\cos\alpha \\
&=\; -\sin\alpha\sin\beta+\cos\beta\cos\alpha \;=\; \cos\alpha\cos\beta-\sin\alpha\sin\beta\ ,
\end{aligned}$$

completing the proof.