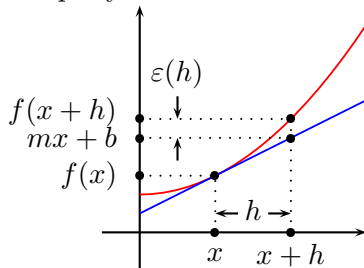


## Linear approximation and the rules of differentiation

The whole body of calculus rest on the idea that if you break things up into little pieces, you can approximate them linearly. Geometrically speaking, we approximate an arc of a curve by a tangent line segment or, in higher dimensions, a piece of a surface by a tangent plane and so on. Note that this approximation is *local* — valid only in small neighbourhood of a point. The derivative is the “coefficient” (or “slope”) of such an approximation.

**Displacement:** If we wish to approximate a function  $f$  near a point  $x$ , we need to know how  $f(x)$  varies with  $x$  for small changes of  $x$ . We denote the change in  $x$  by  $h$  (Leibniz denoted it by  $\Delta x$ ). Thus, if we start by fixing  $x$ , and slightly move away from  $x$ , then our new point is  $x + h$ .

**Linear approximation:** Our hope is to approximate  $f(x + h)$  by a linear function of  $h$  — something of the form  $mh + b$ . In the expression  $mh + b$ , the linear term is  $mh$ , while  $b$  represents a vertical shift (in geometry a linear function plus a shift is called an affine map). The error of our approximation is the difference  $\varepsilon(h) = f(x + h) - (mh + b)$ . In order for the approximation to be a success, the error must be small. It must be a lot smaller than the approximation itself, which means that it must decrease to 0 faster than linearly. The way to express this precisely is with limits. We will require  $\lim_{h \rightarrow 0} \varepsilon(h)/h = 0$ . You can see in the figure below that for a smooth function  $f$ , the error decreases rapidly with  $h$ .



**Differentiability:** A function  $f$  is said to be *differentiable* (or smooth) at  $x$ , if there exists a successful linear approximation to  $f$  at  $x$ , i.e. if there exist  $m$  and  $b$  such that  $\lim_{h \rightarrow 0} [f(x + h) - (mh + b)]/h = 0$ .

**The derivative:** If  $f$  is differentiable at  $x$ , we can find expressions for the coefficients  $m$  and  $b$  in terms of  $f$ . To find  $b$ , take the limit as  $h \rightarrow 0$  of the formula  $f(x + h) = b + mh + \varepsilon(h)$ . Since  $\varepsilon(h) \rightarrow 0$ , we see that  $b = f(x)$ . Thus,  $f(x + h) - f(x) = mh + \varepsilon(h)$ . To find the slope  $m$ , divide by  $h$  and again take the limit as  $h \rightarrow 0$ . Since  $\varepsilon(h)/h \rightarrow 0$ , we see that  $m = \lim_{h \rightarrow 0} [f(x + h) - f(x)]/h$ , i.e.  $m$  is the limit of slopes of secant lines. This is called the derivative of  $f$  at  $x$  and denoted by  $f'(x)$ . Note that, for single variable functions the mere existence of this limit is sufficient to guarantee differentiability. To summarize,  $f$  is differentiable at  $x$  means that

$$f(x + h) = f(x) + f'(x)h + \varepsilon(h), \text{ where } \lim_{h \rightarrow 0} \frac{\varepsilon(h)}{h} = 0 \text{ and } f'(x) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$$

**Example:** Let  $f(x) = x^3$ . Then  $f(x + h) = (x + h)^2 = x^3 + 3x^2h + 3xh^2 + h^3$ . Comparing this expression to  $f(x + h) = f(x) + f'(x)h + \varepsilon(h)$  tells us that  $f'(x) = 3x^2$  and  $\varepsilon(h) = 3xh^2 + h^3$ . Since  $\varepsilon(h)/h = 3xh + h^2 \rightarrow 0$  as  $h \rightarrow 0$  we see that  $f$  is differentiable at any  $x$ , its derivative is  $3x^2$  and it's linear approximation at a point  $x$  is  $f(x + h) = (x + h)^3 \approx x^3 + 3x^2h$ . For example, letting  $x = 2$  we obtain the linear approximation  $8 + 12h$ . If  $h = -0.1$ , then we can conclude that  $1.9^3 \approx 8 + 12 \cdot (-0.1) = 6.8$ . The equation of the tangent line at  $x = 2$  can be obtained by replacing  $h$  in the linear approximation by  $x - 2$ , which gives  $y = 8 + 12(x - 2)$  (point-slope form:  $y - 8 = 12(x - 2)$ , slope-intercept form:  $y = 12x - 16$ ).

**Rules of differentiation:** Most formulas are built up from  $x$  and constants via algebraic operations and composition of algebraic and special functions (e.g. exponentials, logarithms, and trigonometric functions). We can deduce how to handle all such formulas by developing rules of differentiation which deal with each operation and special function.

**The constant rule:** If  $f$  is constant, then  $f'$  is identically zero. This is obvious both geometrically and algebraically, since  $f(x + h) - f(x) = 0$ .

**The power rule (positive integer):** Let  $f(x) = x^n$ , where  $n$  is a positive integer. Then  $f(x + h) = (x + h)^n = x^n + nx^{n-1}h + C_n^2 x^{n-2}h^2 + \dots h^n$ . Comparing this expression to  $f(x + h) = f(x) + f'(x)h + \varepsilon(h)$  tells us that  $f'(x) = nx^{n-1}$  and  $\varepsilon(h) = C_n^2 x^{n-2}h^2 + \dots h^n$ . Thus,  $f$  is differentiable and  $f'(x) = nx^{n-1}$ . In particular,  $x' = 1$ , which is obvious anyway. Also, it's nice to see agreement with the above example:  $(x^3)' = 3x^2$ .

**The sum rule:** If  $f$  and  $g$  are differentiable at  $x$  and  $u(x) = f(x) + g(x)$ , then  $u(x+h) = f(x+h) + g(x+h) = f(x) + f'(x)h + \varepsilon_f(h) + g(x) + g'(x)h + \varepsilon_g(h) = u(x) + [f'(x) + g'(x)]h + [\varepsilon_f(h) + \varepsilon_g(h)]$ . Thus  $u$  is differentiable at  $x$  and  $u'(x) = f'(x) + g'(x)$ .

**The product rule:** If  $f$  and  $g$  are differentiable at  $x$  and  $u(x) = f(x)g(x)$ , then  $u(x+h) = f(x+h)g(x+h) = [f(x) + f'(x)h + \varepsilon_f(h)][g(x) + g'(x)h + \varepsilon_g(h)] = u(x) + [f'(x)g(x) + f(x)g'(x)]h + [f'(x)g'(x)h^2 + f(x)\varepsilon_g(h) + f'(x)h\varepsilon_g(h) + \varepsilon_f(h)g(x) + \varepsilon_f(h)g'(x)h + \varepsilon_f(h)\varepsilon_g(h)]$ . Thus  $u$  is differentiable at  $x$  and  $u'(x) = f'(x)g(x) + f(x)g'(x)$ .

**The constant factor rule:** If  $f$  is differentiable at  $x$ ,  $c$  is a constant, and  $u(x) = cf(x)$ , then  $u$  is differentiable at  $x$  and  $u'(x) = c'f(x) + cf'(x) = cf'(x)$ .

**The quotient rule:** If  $f$  and  $g$  are differentiable at  $x$ ,  $g(x) \neq 0$  and  $u(x) = f(x)/g(x)$ , then  $f(x) = u(x)g(x)$ , so  $f'(x) = u'(x)g(x) + u(x)g'(x)$ . Solving for  $u'(x)$  we obtain  $u'(x) = [f'(x) - u(x)g'(x)]/g(x) = [f'(x) - f(x)g'(x)/g(x)]/g(x) = [f'(x)g(x) - f(x)g'(x)]/[g(x)]^2$ .

**Example:** Let  $f(x) = 1/x$ . Then  $f'(x) = [1'x - 1x']/x^2 = -1/x^2$ .

**The chain rule:** If  $g$  is differentiable at  $x$ ,  $f$  is differentiable at  $g(x)$  and  $u(x) = f(g(x))$ , then letting  $s = g(x+h) - g(x)$  we see that  $f(g(x+h)) = f(g(x) + s) = f(g(x)) + f'(g(x))s + \varepsilon_f(s) = f(g(x)) + f'(g(x))[g(x+h) - g(x)] + \varepsilon_f(s) = f(g(x)) + f'(g(x))[g(x) + g'(x)h + \varepsilon_g(h) - g(x)] + \varepsilon_f(s) = f(g(x)) + f'(g(x))g'(x)h + [f'(g(x))\varepsilon_g(h) + \varepsilon_f(s)]$ . Thus  $u$  is differentiable and  $u'(x) = f'(g(x))g'(x)$ .

**The inverse function rule:** Suppose  $u$  is the compositional inverse of  $f$  and  $f$  is differentiable at  $u(x)$ . Since  $f(u(x)) = x$ , we have  $[f(u(x))]' = 1$ . On the other hand, by the chain rule  $[f(u(x))]' = f'(u(x))u'(x)$ . Thus,  $u'(x) = 1/f'(u(x))$ .

**Example:** Let  $f(x) = \sqrt{x}$ , where  $x \geq 0$ . Then  $f$  is the compositional inverse of  $g$ , where  $g(x) = x^2$ , so  $f'(x) = 1/g'(f(x)) = 1/[2f(x)] = 1/[2\sqrt{x}]$ .

**The power rule (positive rational):** Let  $f(x) = x^r$ , where  $r = n/m$  and  $n, m$  are positive integers. Then  $f(x)^m = x^{rm} = x^n$ , so  $[f(x)^m]' = nx^{n-1}$ . On the other hand, by the power rule and the chain rule  $[f(x)^m]' = mf(x)^{m-1}f'(x)$ . Thus,  $f'(x) = nx^{n-1}/[mf(x)^{m-1}] = nx^{n-1}/[mx^{r(m-1)}] = rx^{n-1-rm+r} = rx^{r-1}$ .

**The power rule (rational):** Let  $f(x) = x^r$ , where  $r$  is rational. If  $r > 0$ , we already know the answer, so assume  $r = -s$ , where  $s$  is a positive rational. Then  $f(x) = x^{-s} = (1/x)^s$ , so  $f'(x) = s(1/x)^{s-1}(1/x)' = s(1/x)^{s-1}(-1/x^2) = -sx^{-s+1-2} = -sx^{-s-1} = rx^{r-1}$ .

**Exponentials and logarithms:** Suppose  $b > 0$  and  $f(x) = b^x$ . Then  $f(x+h) - f(x) = b^{x+h} - b^x = b^x b^h - b^x = b^x(b^h - 1)$ . Thus,  $f'(x) = cb^x$ , where  $c = \lim_{h \rightarrow 0} (b^h - 1)/h$ . In other words,  $f'(x) = cf(x)$  and we would like to figure out the proportionality constant. A bit of estimation shows that for  $b = 2$ ,  $c < 1$  and for  $b = 3$ ,  $c > 1$ . This suggests the existence of a number  $e \approx 2.71828$  (Napier's number) between 2 and 3 such that  $\lim_{h \rightarrow 0} (e^h - 1)/h = 1$ . Then, in particular,  $(e^x)' = e^x$ . The compositional inverse function to  $e^x$  is called the natural logarithm  $\ln(x) = \log_e(x)$ . By the inverse function rule  $[\ln(x)]' = 1/e^{\ln(x)} = 1/x$ . Returning to the issue of computing  $c$ , write  $b^h = e^{\ln(b^h)} = e^{h \ln(b)}$ . Then  $c = \lim_{h \rightarrow 0} (b^h - 1)/h = \lim_{h \rightarrow 0} (e^{h \ln(b)} - 1)/h = \ln b \lim_{h \rightarrow 0} (e^{h \ln(b)} - 1)/[h \ln(b)] = \ln b$ . Thus,  $f'(x) = \ln(b)b^x$ . Again, by the inverse function rule  $[\log_b(x)]' = 1/[\ln(b)b^{\log_b(x)}] = 1/[\ln(b)x]$ .

**Example:** Let  $f(x) = x^x$ . Then  $\ln f(x) = \ln(x^x) = x \ln(x)$ , so  $[\ln f(x)]' = 1 \ln(x) + x(1/x) = \ln(x) + 1$ . On the other hand  $[\ln f(x)]' = [1/f(x)]f'(x)$ . Thus,  $f'(x) = f(x)[\ln(x) + 1] = x^x[\ln(x) + 1]$ .

**Trigonometric functions:** First, the sine:  $\sin(x+h) - \sin(x) = \sin(x)\cos(h) + \cos(x)\sin(h) - \sin(x) = \sin(x)[\cos(h) - 1] + \cos(x)\sin(h)$ . Since  $\lim_{h \rightarrow 0} \sin(h)/h = 1$  and  $\lim_{h \rightarrow 0} [\cos(h) - 1]/h = 0$ , we have  $[\sin(x)]' = \cos(x)$ . Now the cosine:  $[\cos(x)]' = [\sin(\pi/2 - x)]' = \cos(\pi/2 - x)(-1) = -\sin(x)$ . The remaining trigonometric functions are defined in terms of sine and cosine, so can be differentiated using the usual rules. The inverse trigonometric functions are differentiated using the inverse function rule. For example,  $[\arcsin(x)]' = 1/\cos(\arcsin(x)) = 1/\sqrt{1 - \sin(\arcsin(x))^2} = 1/\sqrt{1 - x^2}$ .

**Summary:**  $c' = 0$ ,  $(x^r)' = rx^{r-1}$ ,  $(f+g)' = f' + g'$ ,  $(fg)' = f'g + fg'$ ,  $(cf)' = cf'$ ,  $(f/g)' = [f'g - fg']/g^2$ ,  $[f(g)]' = f'(g)g'$ ,  $(f^{-1})' = 1/f'(f^{-1})$ ,  $(e^x)' = e^x$ ,  $(b^x)' = \ln(b)b^x$ ,  $[\ln(x)]' = 1/x$ ,  $[\log_b(x)]' = 1/[\ln(b)x]$ ,  $[\sin(x)]' = \cos(x)$ ,  $[\cos(x)]' = -\sin(x)$ ,  $[\tan(x)]' = [\sec(x)]^2$ ,  $[\cot(x)]' = -[\csc(x)]^2$ ,  $[\sec(x)]' = \tan(x)\sec(x)$ ,  $[\csc(x)]' = -\cot(x)\csc(x)$ ,  $[\arcsin(x)]' = [\arccos(x)]' = 1/\sqrt{1 - x^2}$ ,  $[\arctan(x)]' = 1/[1 + x^2]$ .