Theorem 8.3 (Chain Rule): Let $f, g : \mathbb{R} \to \mathbb{R}$. Assume that f is differentiable at x_0 and that g is differentiable at $y_0 = f(x_0)$. Then $h = g \circ f$ is differentiable at x_0 and the derivative is given by:

$$h'(x_0) = g'(y_0)f'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. Recall the definition of the derivative of g at y_0:

$$g'(y_0) = \lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0}$$

which can be rewritten as follows:

$$\begin{aligned} 0 &= \lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} - g'(y_0) \\ &= \lim_{y \to y_0} \frac{g(y) - g(y_0) - g'(y_0)(y - y_0)}{y - y_0}. \end{aligned}$$

Call the numerator in the above limit $\psi(y)$:

$$\psi(y) = g(y) - g(y_0) - g'(y_0)(y - y_0).$$

The function $\psi(y)$ is defined in a small neighborhood of y_0. Then we know that not only $\lim_{y \to y_0} \psi(y) = 0$, but also

$$\lim_{y \to y_0} \frac{\psi(y)}{y - y_0} = 0$$

(this means that $\psi(y) = o(y - y_0)$ as $y \to y_0$). That is, as $y \to y_0$, $\psi(y)$ tends to 0 faster than $y - y_0$. Denote:

$$\varphi(y) = \frac{\psi(y)}{y - y_0} \qquad \text{so that} \qquad \psi(y) = \varphi(y)(y - y_0).$$

Then

$$\lim_{y \to y_0} \varphi(y) = 0.$$

Returning to the definition of ψ, we have:

$$g(y) - g(y_0) - g'(y_0)(y - y_0) = \psi(y) = \varphi(y)(y - y_0).$$

Hence, for all y in a neighborhood of y_0:

$$g(y) - g(y_0) = g'(y_0)(y - y_0) + \varphi(y)(y - y_0).$$

Plugging in $y = f(x)$ and $y_0 = f(x_0)$, and dividing by $x - x_0$ this becomes

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = g'(f(x_0))\frac{f(x) - f(x_0)}{x - x_0} + \varphi(f(x))\frac{f(x) - f(x_0)}{x - x_0}.$$

Recalling that $h = g \circ f$ we have

$$\frac{h(x) - h(x_0)}{x - x_0} = g'(f(x_0))\frac{f(x) - f(x_0)}{x - x_0} + \varphi(f(x))\frac{f(x) - f(x_0)}{x - x_0}.$$

Taking the limit $x \to x_0$, we get

$$\begin{aligned}
\lim_{x\to x_0} \frac{h(x)-h(x_0)}{x-x_0} &= g'(f(x_0)) \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} + \lim_{x\to x_0} \left[\varphi(f(x))\frac{f(x)-f(x_0)}{x-x_0}\right] \\
&= g'(f(x_0)) \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} + \lim_{x\to x_0} \varphi(f(x)) \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \\
&= g'(f(x_0))f'(x_0) + 0 \cdot f'(x_0) \\
&= g'(f(x_0))f'(x_0).
\end{aligned}$$

Hence we find that the derivative of $h = g \circ f$ at x_0 exists and is equal to

$$h'(x_0) = (g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

□

Chain rule: avoiding confusion!

It is easy to get confused with notation when dealing with the chain rule, so it is best to be cautious. If we write $h = g \circ f$ then the chain rule can be expressed as

$$\frac{\mathrm{d}h}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}g}{\mathrm{d}y}(y_0)\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$$

where $y_0 = f(x_0)$. Alternatively, if we write $y = f(x)$ and $z = g(y)$ then we can write the chain rule also as:

$$\frac{\mathrm{d}z}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}z}{\mathrm{d}y}(y_0)\frac{\mathrm{d}y}{\mathrm{d}x}(x_0)$$

where $y_0 = f(x_0)$. This latter expression is easier to remember, because one could imagine that the terms $\mathrm{d}y$ on the right hand side cancel out (though we are not allowed to actually do that!).

Example 8.14: Consider the function $z = h(x) = \sqrt{1-x^2}$ which is the composition of $y = f(x) = 1 - x^2$ with $z = g(y) = \sqrt{y}$. Recalling that

$$f'(x) = -2x \qquad \text{and} \qquad g'(y) = \frac{1}{2\sqrt{y}}$$

we have:

$$\frac{\mathrm{d}h}{\mathrm{d}x}(x) = \frac{\mathrm{d}g}{\mathrm{d}y}(y)\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \frac{1}{2\sqrt{1-x^2}}(-2x) = -\frac{x}{\sqrt{1-x^2}}.$$

Theorem 8.4 (Derivative of the inverse function)**:** Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and invertible in a neighborhood of $x_0 \in \mathbb{R}$. Moreover, suppose that f is differentiable at x_0 and that $f'(x_0) \neq 0$. Then the inverse function $f^{-1}(y)$ is differentiable at $y_0 = f(x_0)$ and the derivative there is given by

$$\frac{\mathrm{d}}{\mathrm{d}y} f^{-1}(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

Proof. By Theorem 7.9, f^{-1} is continuous on a neighborhood of $y_0 = f(x_0)$. For y in the neighborhood we have

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}}.$$

The Substitution Theorem (Theorem 5.12) implies that taking $y \to y_0$ on the left hand side, is the same as taking $x \to x_0$ on the right hand side:

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{x \to x_0} \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}} = \frac{1}{f'(x_0)}.$$

□

Example 8.15: Let $y = f(x) = \tan x$. We saw that $y'(x) = 1 + \tan^2 x = 1 + y^2(x)$. Denote $x = f^{-1}(y) = \arctan(y)$. Then:

$$\frac{\mathrm{d}}{\mathrm{d}y} \arctan(y) = \frac{1}{1 + y^2}.$$

Example 8.16: Let $y = f(x) = \sin x$. We saw that $y'(x) = \cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2(x)}$. Then:

$$\frac{\mathrm{d}}{\mathrm{d}y} \arcsin(y) = \frac{1}{\sqrt{1 - y^2}}.$$

Similarly, it can be shown that

$$\frac{\mathrm{d}}{\mathrm{d}y} \arccos(y) = -\frac{1}{\sqrt{1 - y^2}}.$$

Example 8.17: Let $y = f(x) = a^x$. Then $y'(x) = (\ln a)a^x = (\ln a)y$. The inverse function is $x = f^{-1}(y) = \log_a y$, and we have:

$$\frac{\mathrm{d}}{\mathrm{d}y} \log_a y = \frac{1}{(\ln a)y}$$

One can check that the same result is obtained for the derivative of $\log_a(-y)$, so that we have

$$\frac{\mathrm{d}}{\mathrm{d}y} \log_a |y| = \frac{1}{(\ln a)y} \qquad y \neq 0.$$

In particular:

$$\frac{\mathrm{d}}{\mathrm{d}y} \ln |y| = \frac{1}{y} \qquad y \neq 0.$$

Logarithmic derivative

If $f : \mathbb{R} \to \mathbb{R}_+$ attains positive values, then $h(x) = \ln(f(x))$ is well-defined. Moreover, if f is differentiable, then so is h and we have $h'(x) = \ln'(f(x))f'(x)$. Using the formula for the derivative of the logarithm, we have:

$$\frac{\mathrm{d}}{\mathrm{d}x} \ln(f(x)) = \frac{f'(x)}{f(x)}.$$

Proposition 8.5: Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable. Then:

$$\begin{aligned} f \text{ odd} \quad &\Rightarrow \quad f' \text{ even} \\ f \text{ even} \quad &\Rightarrow \quad f' \text{ odd} \end{aligned}$$

Proof. Let us consider the first case. The second one follows a similar line of reasoning. If f is odd, then $f(-x) = -f(x)$. Taking derivatives of both sides we have $-f'(-x) = -f'(x)$ so that $f'(-x) = f'(x)$ and the derivative function is even. □

Important derivatives

$$\frac{\mathrm{d}}{\mathrm{d}x} x^\alpha = \alpha x^{\alpha-1} \qquad (\forall \alpha \in \mathbb{R})$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \tan x = 1 + \tan^2 x = \frac{1}{\cos^2 x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arcsin x = \frac{1}{\sqrt{1-x^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arccos x = -\frac{1}{\sqrt{1-x^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arctan x = \frac{1}{1+x^2}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} a^x = (\ln a)\, a^x \qquad (\forall a > 0) \qquad \text{in particular, } \frac{\mathrm{d}}{\mathrm{d}x} e^x = e^x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \log_a |x| = \frac{1}{(\ln a)\, x} \qquad (\forall a > 0,\ a \neq 1) \quad \text{in particular, } \frac{\mathrm{d}}{\mathrm{d}x} \ln |x| = \frac{1}{x}$$

8.3 Where differentiability fails

We have already seen that the function $f(x) = |x|$ is not differentiable at $x_0 = 0$, since the right- and left-limits of the expression $\frac{\Delta f}{\Delta x}$ are different (they are ± 1). This suggests that we can define differentiability on both sides:

Definition 8.6: Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined on an interval $[x_0, x_0 + r)$ for some $r > 0$. Then f is **right-differentiable** at x_0 if

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{exists and is finite. It is denoted } f'_+(x_0)$$

Similarly, if f is defined on $(x_0 - r, x_0]$ then we say that f is **left-differentiable** at x_0 if

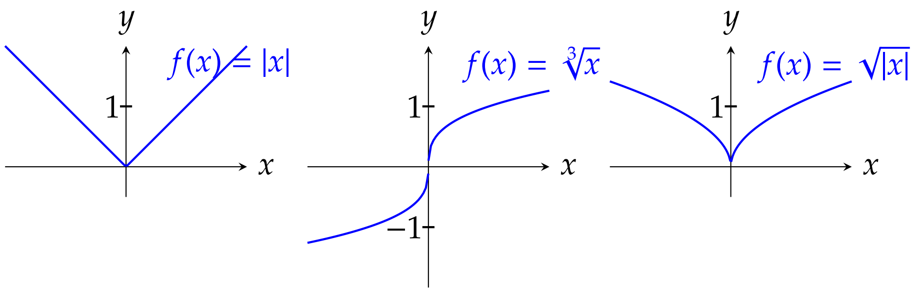
$$\lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{exists and is finite. It is denoted } f'_-(x_0).$$

Proposition 8.7: A function f is differentiable at x_0 if and only if it is both left- and right-differentiable at x_0, and $f'_+(x_0) = f'_-(x_0)$. In this case $f'(x_0) = f'_+(x_0) = f'_-(x_0)$.

Proof. The proof is almost immediate, and left as an exercise (see Proposition 4.4). □

Remark:

1. If $f'_+(x_0) \neq f'_-(x_0)$ are both finite, then x_0 is a **corner point** of f (e.g. 0 is a corner point of $|x|$).
2. If one or both of $f'_+(x_0)$ and $f'_-(x_0)$ are $+\infty$ or $-\infty$, then x_0 is a point with **vertical tangent** of f. For example for $f(x) = \sqrt{x}$
$$f'_+(0) = +\infty.$$
Another example is $f(x) = \sqrt[3]{x}$, for which $f'(x) = \frac{1}{3}x^{-\frac{2}{3}}$ is defined for $x \neq 0$. However,
$$f'_\pm(0) = \lim_{x \to 0^\pm} \frac{1}{3}x^{-\frac{2}{3}} = +\infty.$$
3. If $f'_+(x_0)$ and $f'_-(x_0)$ are infinite of opposite signs, then x_0 is a **cusp** point of f. For example, consider $f(x) = \sqrt{|x|}$. Then by definition
$$f'_\pm(0) = \lim_{x \to 0^\pm} \frac{\sqrt{|x|}}{x} = \lim_{x \to 0^\pm} \frac{\sqrt{|x|}}{\operatorname{sign}(x)|x|} = \lim_{x \to 0^\pm} \frac{1}{\operatorname{sign}(x)\sqrt{|x|}} = \pm\infty.$$



Theorem 8.8: If a function f is continuous at x_0 and differentiable in a neighborhood of x_0 (excluding x_0 itself) then f is differentiable at x_0 if $\lim_{x \to x_0} f'(x)$ exists and is finite. If so, then $f'(x_0) = \lim_{x \to x_0} f'(x)$.

Proof. We don't have (yet) the tools to prove this theorem. We'll come back to it later. □

Example 8.18: Are there numbers $a, b \in \mathbb{R}$ such that the function

$$f(x) = \begin{cases} a\sin(2x) - 4 & x < 0 \\ b(x-1) + e^x & x \geq 0 \end{cases}$$

is differentiable at $x_0 = 0$?

First, f needs to be continuous at $x_0 = 0$. So we need:

$$\lim_{x\to 0^-} a\sin(2x) - 4 = -4 \qquad \text{to be equal to} \qquad -b+1$$

Hence $b = 5$. Now we need the left- and right-derivatives to agree.

$$\lim_{x\to 0^+} f'(x) = \lim_{x\to 0^+} (b + e^x) = 6,$$
$$\lim_{x\to 0^-} f'(x) = \lim_{x\to 0^-} 2a\cos(2x) = 2a.$$

For these to be equal, we impose $a = 3$.

8.4 Extrema and critical points

We can now dig deeper into our previous definitions of the supremum, infimum, maximum and minimum of sets.

Local maximum

A point x_0 is a **local maximum point** for f if there exists a neighborhood $I_r(x_0)$ such that

$$f(x) \le f(x_0), \qquad \forall x \in \mathrm{dom}(f) \cap I_r(x_0).$$

Then $f(x_0)$ is a **local maximum** of f.

Global maximum

A point x_0 is a **global maximum point** for f if

$$f(x) \le f(x_0), \qquad \forall x \in \mathrm{dom}(f).$$

Then $f(x_0)$ is the **global maximum** of f. The maximum is **strict** if $f(x) < f(x_0)$ for all $x \neq x_0$.

Local minimum

A point x_0 is a **local minimum point** for f if there exists a neighborhood $I_r(x_0)$ such that

$$f(x) \ge f(x_0), \qquad \forall x \in \mathrm{dom}(f) \cap I_r(x_0).$$

Then $f(x_0)$ is a **local minimum** of f.