4. DIFFERENTIABILITY

4.1 Definitions

In the mid-17th century, Fermat made initial advances in determining tangents to 'algebraic' curves defined by polynomials. His arguments were geometric, involving triangles which became similar in the limit. Later, Newton and Leibniz would extend these ideas, and present them more in terms of functions. However, none of these mathematicians had a rigorous definition of what a limit was.

Given a function $f\colon (a,b) \to \mathbb{R}$ and $p \in (a,b)$, we can create a potential approximation to the gradient at p namely

$$\frac{f(p+h)-f(p)}{h},$$

where h is a small value. For a well-behaved function, we expect a situation like the one below:

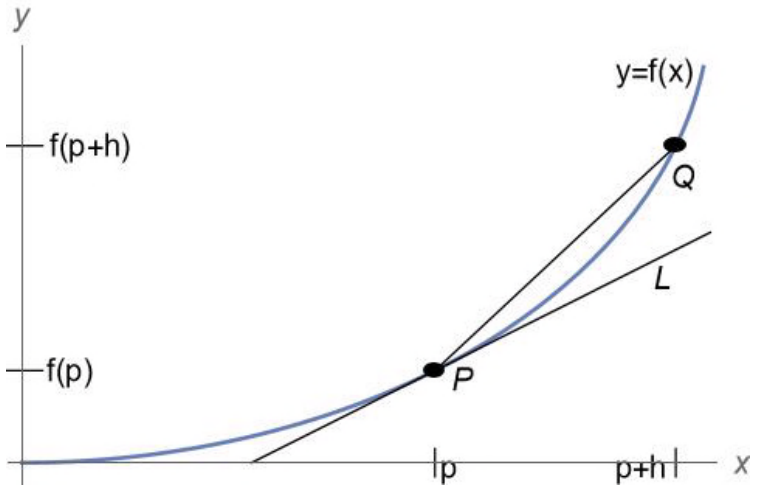


Fig. 4.1. – Newton quotient

We are seeking to determine the gradient to the graph $y = f(x)$ at a point $P = (p, f(p))$ or equivalently the tangent line L at P. A nearby point $Q = (p+h, f(p+h))$ is depicted and the above fraction, called a **Newton quotient**, is the gradient of the chord PQ. As h 'becomes smaller and smaller' the chord's gradient should better approximate the gradient at P.

Newton and Leibniz, however seminal their work, could do little better than the above phrasing. We are now in a position to determine this gradient as a limit. It should be clear this is necessary as we cannot simply set $h = 0$ in the above fraction, as this would yield the indeterminate $0/0$.

Definition 4.1 *We say that a function $f\colon (a,b) \to \mathbb{R}$ is **differentiable** at $p \in (a,b)$ if the limit*

$$\lim_{x\to p} \frac{f(x)-f(p)}{x-p} = \lim_{h\to 0} \frac{f(p+h)-f(p)}{h}$$

exists. (Note these are just different expressions for the same limit.) We denote this limit, if it exists, as

$$f\,'(p) \qquad or \qquad \frac{\mathrm{d}f}{\mathrm{d}x}(p),$$

which is known as the ***derivative*** *of f at p.*

Note that f being differentiable at p is equivalent to writing

$$f(p+h) = f(p) + f\,'(p)h + o(h).$$

Proposition 4.2 *A differentiable function is continuous.*

Proof. Let $f\colon (a,b) \to \mathbb{R}$ be differentiable at $p \in (a,b)$. By AOL

$$\begin{aligned}\lim_{x\to p}\left(f\left(x\right)-f\left(p\right)\right) &= \lim_{x\to p}\left\{(x-p)\left(\frac{f\left(x\right)-f\left(p\right)}{x-p}\right)\right\}\\ &= \lim_{x\to p}(x-p)\times\lim_{x\to p}\left(\frac{f\left(x\right)-f\left(p\right)}{x-p}\right)\\ &= 0\times f\,'(p)=0.\end{aligned}$$

Hence f is continuous at p. ■

Example 4.3 *Let $n \in \mathbb{N}$. The derivative of $f(x) = x^n$ is $f'(x) = nx^{n-1}$.*

Solution. By the binomial theorem

$$(x+h)^n = \sum_{k=0}^{n}\binom{n}{k}h^k x^{n-k} = x^n + nx^{n-1}h + o(h),$$

so that $f\,'(x) = nx^{n-1}$. ■

Example 4.4 *The derivative of* $\exp x$ *is* $\exp x$.

Solution. Note

$$\frac{\exp(x+h)-\exp x}{h} = \frac{e^{x+h}-e^x}{h} = e^x\left(\frac{e^h-1}{h}\right).$$

From the exponential series

$$e^h = 1 + h + h^2\left(\frac{1}{2!}+\frac{h}{3!}+\frac{h^2}{4!}+\cdots\right).$$

Using the ratio test, we can show that the series in parentheses converges for all h. By Corollary 3.12, this means that it defines a continuous function. Hence

$$\lim_{h\to 0}\frac{e^h-1}{h} = 1+\lim_{h\to 0}h\left(\frac{1}{2!}+\frac{h}{3!}+\frac{h^2}{4!}+\cdots\right)\overset{\text{AOL}}{=}1+0\times\frac{1}{2}=1.$$

The result follows. ■

Example 4.5 *The derivative of* $\sin x$ *is* $\cos x$.

Solution. Note

$$\frac{\sin(x+h)-\sin x}{h} = \sin x\left(\frac{\cos h - 1}{h}\right) + \cos h\left(\frac{\sin h}{h}\right).$$

We already saw in Example 1.4(c) that $\frac{1}{h}\sin h \to 1$ as $h \to 0$. Further

$$\frac{\cos h - 1}{h} = \frac{1}{h}\left(-\frac{h^2}{2!} + \frac{h^4}{4!} - \right) = h\left(-\frac{1}{2!} + \frac{h^2}{4!} - \right).$$

By the ratio test the power series in parentheses converges for all h and so defines a continuous function by Corollary 3.12. By AOL

$$\lim_{h\to 0}\frac{\cos h - 1}{h} = 0 \times \left(-\frac{1}{2}\right) = 0.$$

Hence, by AOL,

$$\sin' x = \lim_{h\to 0}\frac{\sin(x+h)-\sin x}{h} = \sin x \times 0 + \cos x \times 1 = \cos x.$$

■

Example 4.6 *The derivative of* $\log x$ *is* x^{-1} *for* $x > 0$.

Solution. For $h, x > 0$, note

$$\frac{\log(x+h)-\log x}{h} = \frac{1}{h}\log\left(1+\frac{h}{x}\right).$$

So, for $0 < h < x$, we can use the logarithmic series to write

$$\frac{\log(x+h)-\log x}{h} = \frac{1}{h}\left(\frac{h}{x} - \frac{h^2}{2x^2} + \frac{h^3}{3x^3} - \cdots\right) = \frac{1}{x} + h\left\{-\frac{1}{2x^2} + \frac{h}{3x^3} - \cdots\right\}.$$

By the ratio test, the series in braces converges for $|h| < |x|$ and so defines a continuous function by Corollary 3.12. By AOL

$$\lim_{h\to 0}\frac{\log(x+h)-\log x}{h} = \frac{1}{x} + 0 \times \left(\frac{-1}{2x^2}\right) = \frac{1}{x}.$$

The result follows. ■

Example 4.7 *The function* $|x|$ *is not differentiable at* 0.

Solution. Note that

$$\lim_{h\to 0^+}\frac{|h|-0}{h} = \lim_{h\to 0^+}\frac{h}{h} = 1; \qquad \lim_{h\to 0^-}\frac{|h|-0}{h} = \lim_{h\to 0^+}\frac{-h}{h} = -1.$$

Hence the derivative of $|x|$ at 0 does not exist. ■

Definition 4.8 *We say that* $f\colon (a,b) \to \mathbb{R}$ *has a* ***left-sided derivative*** *at* $p \in (a,b)$ *if*

$$f_-'(p) = \lim_{x \to p,\, x<p} \frac{f(x)-f(p)}{x-p} = \lim_{h \to 0^-} \frac{f(p+h)-f(p)}{h}$$

exists and has a ***right-sided derivative*** *at* $p \in (a,b)$ *if*

$$f_+'(p) = \lim_{x \to p,\, x>p} \frac{f(x)-f(p)}{x-p} = \lim_{h \to 0^+} \frac{f(p+h)-f(p)}{h}$$

eixsts.

Proposition 4.9 $f\colon (a,b) \to \mathbb{R}$ *is differentiable at* $p \in (a,b)$ *if and only if* $f_-'(p)$ *and* $f_+'(p)$ *exist and are equal.*

Remark 4.10 *In the previous example, we see that* $|x|$ *has left- and right-sided derivatives at* 0 *of* -1 *and* 1*, respectively. As these are not equal,* $|x|$ *is not differentiable at* 0*.*

Proof. If f is differentiable at p, then $f_-'(p)$ and $f_+'(p)$ both exist by definition and equal $f'(p)$.

Conversely, say that $f_-'(p)$ and $f_+'(p)$ exist and equal L. Given any $\varepsilon > 0$, there exist $\delta_1, \delta_2 > 0$ such that

$$\begin{aligned} \left| \frac{f(x)-f(p)}{x-p} - L \right| &< \varepsilon \qquad \text{for} \qquad p-\delta_1 < x < p; \\ \left| \frac{f(x)-f(p)}{x-p} - L \right| &< \varepsilon \qquad \text{for} \qquad p < x < p+\delta_2. \end{aligned}$$

Set $\delta = \min\{\delta_1, \delta_2\} > 0$. Then

$$\left| \frac{f(x)-f(p)}{x-p} - L \right| < \varepsilon \qquad \text{for} \qquad 0 < |x-p| < \delta$$

and hence $f'(p) = L$. ■

Remark 4.11 *Weierstrass, in 1872, gave the first example of a function* $f\colon \mathbb{R} \to \mathbb{R}$ *which is (uniformly) continuous, but nowhere differentiable. His example was*

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x),$$

where $0 < a < 1$ *and* b *is a positive, odd integer, such that* $ab > 1 + 3\pi/2$*. On Sheet 5, Exercise 10, we demonstrate that a similar function is continuous and nowhere differentiable.*

Whilst these functions are sometimes described as pathological, they should be seen as typical of continuous functions. In a sense, which can be made rigorous, a randomly chosen continuous function is nowhere differentiable with probability 1*.*

Definition 4.12 *If f is differentiable for all $p \in (a,b)$, then $f' \colon (a,b) \to \mathbb{R}$ is a well-defined function. If f' is itself differentiable, then we say f is* ***twice differentiable*** *and so on. The second derivative is denoted as*

$$\dot{f}''(p) \qquad or \qquad \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(p).$$

More generally, if it exists, the n***th derivative*** *at p is denoted*

$$f^{(n)}(p) \qquad or \qquad \frac{\mathrm{d}^n f}{\mathrm{d}x^n}(p).$$

Such derivatives are referred to as ***higher derivatives.*** *The first notation is due to Lagrange (1749) and the second to Leibniz (1684).*

If f has derivatives of all orders, then f is said to be ***smooth.***

Example 4.13 *$f(x) = x|x|$ is differentiable everywhere, but is not twice differentiable at 0.*

Solution. For $x > 0$, we have $f(x) = x^2$ and so $f'(x) = 2x$.

For $x < 0$, we have $f(x) = -x^2$ and so $f'(x) = -2x$.

Finally, for $x = 0$, we have

$$\frac{f(h) - f(0)}{h} = \frac{h|h| - 0}{h} = |h| \to 0.$$

Hence $f'(0) = 0$ and altogether we see $f'(x) = 2|x|$, and we already know that $f'(x)$ is not differentiable at 0. ■

Example 4.14 *Let $f \colon \mathbb{R} \to \mathbb{R}$ be differentiable at p. For each of the following statements, either prove the statement or give a counter-example.*

(a) If $f'(p) \geqslant 0$ then there exists $\delta > 0$ such that $f(x) \geqslant f(p)$ for all $x \in (p, p+\delta)$.

(b) If $f'(p) > 0$ then there exists $\delta > 0$ such that $f(x) > f(p)$ for all $x \in (p, p+\delta)$.

(c) If $f'(p) > 0$ then there exists $\delta > 0$ such that f is increasing on $(p, p+\delta)$.

Solution. (a) False. Consider $f(x) = -x^2$ and $p = 0$. Then, for any $\delta > 0$, we have $f(x) < f(0)$ for all $x \in (p, p+\delta)$.

(b) True. There exists $\delta > 0$ such that for all $x \in (p, p+\delta)$ we have

$$-\frac{f'(p)}{2} < \frac{f(x) - f(p)}{x - p} - f'(p),$$

which rearranges to

$$\frac{f(x) - f(p)}{x - p} > \frac{f'(p)}{2} > 0,$$

and so $f(x) > f(p)$.

(c) False. Consider f defined by $f(0) = 0$ and for $x \neq 0$ we define

$$f(x) = \frac{x}{2} + x^2 \sin\left(\frac{1}{x}\right).$$

Take $p = 0$. By Sheet Exercise 1(ii), f is differentiable at 0 and $f'(0) = 1/2 > 0$. However for $x > 0$ we have

$$f'(x) = \frac{1}{2} + 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

This derivative takes negative values in any $(p, p + \delta)$. ■

Remark 4.15 *(Generalizations) In due course, you will meet differentiation for functions of more than one (real) variable and for functions of a complex variable.*

Recall that the derivative $f'(p)$ of a function $f\colon \mathbb{R} \to \mathbb{R}$ at p satisfies

$$f(p + h) = f(p) + f'(p)h + o(h).$$

This generalizes to $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ being differentiable at $\mathbf{p}$ with derivative $D\mathbf{f}_\mathbf{p}\colon \mathbb{R}^2 \to \mathbb{R}^2$ if

$$\mathbf{f}(\mathbf{p} + \mathbf{h}) = \mathbf{f}(\mathbf{p}) + D\mathbf{f}_\mathbf{p}\mathbf{h} + o(|\mathbf{h}|).$$

So $D\mathbf{f}_\mathbf{p}$ is actually a linear map $D\mathbf{f}_\mathbf{p}\colon \mathbb{R}^2 \to \mathbb{R}^2$ which is represented by the Jacobian matrix wrt the standard basis, as we see below. If $\mathbf{f} = (f, g)$ and $\mathbf{h} = (h.k)$ then

$$\begin{aligned} f(p_1 + h, p_2 + k) &= f(p_1, p_2) + \left[h\frac{\partial f}{\partial x}(p_1, p_2) + k\frac{\partial f}{\partial y}(p_1, p_2)\right] + \text{ higher order terms;} \\ g(p_1 + h, p_2 + k) &= g(p_1, p_2) + \left[h\frac{\partial g}{\partial x}(p_1, p_2) + k\frac{\partial g}{\partial y}(p_1, p_2)\right] + \text{ higher order terms.} \end{aligned}$$

Hence

$$\mathbf{f}(\mathbf{p} + \mathbf{h}) = \mathbf{f}(\mathbf{p}) + \begin{pmatrix} \partial f/\partial x & \partial f/\partial y \\ \partial g/\partial x & \partial g/\partial y \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} + \text{ higher order terms.}$$

So, for the function $\mathbf{f}(x, y) = (x, -y)$, we have

$$\mathbf{f}(p_1 + h, p_2 + k) = (p_1 + h, -p_2 - k) = (p_1, -p_2) + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

and we see, for any $\mathbf{p}$, that

$$D\mathbf{f}_\mathbf{p} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

By comparison, a function $f\colon \mathbb{C} \to \mathbb{C}$ is said to be (complexly) differentiable at $p \in \mathbb{C}$ if the limit

$$\lim_{h \to 0} \frac{f(p + h) - f(p)}{h}$$

exists. Crucially $\mathbb{C}$ is different from $\mathbb{R}^2$ in that it is 'oriented'. $i = \sqrt{-1}$ is an integral part of $\mathbb{C}$ that has no counterpart in $\mathbb{R}^2$, and multiplication by i introduces a natural sense of anti-clockwise to $\mathbb{C}$ that $\mathbb{R}^2$ does not have. The function $\mathbf{f}(x, y) = (x, -y)$ above is the equivalent of complex conjugation, $f(z) = \overline{z}$, which is NOT differentiable. Note

$$\frac{f(p + h) - f(p)}{h} = \frac{\overline{p + h} - \overline{p}}{h} = \frac{\overline{h}}{h}.$$

This has no limit as $h \to 0$ in the complex plane. If h is real, we get 1 as the limit and, if h is purely imaginary, we get a limit of -1. Because complex conjugation reverses the natural orientation of $\mathbb{C}$, the function is not (complex) differentiable. In terms of the previous notation, a further requirement on the Jacobian for the function $\mathbf{f}$ to be complex differentiable is that

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial y} \qquad \text{and} \qquad \frac{\partial f}{\partial y} = -\frac{\partial g}{\partial x}.$$

These are known as the Cauchy-Riemann equations.

All the above aspects are covered in the Part A Metric Spaces and Complex Analysis courses.

4.2 AOL

We now meet the sum, product and quotient rules for differentiation. These rules were first appreciated by Leibniz in 1675 and first published in 1684.

Proposition 4.16 *(**Sum, product and quotient rules of differentiation**) Let $f, g \colon \mathbb{R} \to \mathbb{R}$ be differentiable at p. Then $f+g$ and fg are differentiable at p and, if $g'(p) \neq 0$, then f/g is differentiable at p. Further*

$$\begin{aligned} (f+g)'(p) &= f'(p) + g'(p). \\ (fg)'(p) &= f(p)\, g'(p) + f'(p)\, g(p). \\ \left(\frac{f}{g}\right)'(p) &= \frac{g(p)\, f'(p) - f(p)\, g'(p)}{g(p)^2}. \end{aligned}$$

Remark 4.17 *Note, consequently, that the differentiable functions form a real vector space, in fact a real algebra – a vector space which is closed under multiplication. We also note that differentiation is a linear transformation (where well-defined).*

Notation 4.18 *We write:*

- *$C^0(\mathbb{R})$ for the algebra of continuous functions from $\mathbb{R}$ to $\mathbb{R}$.*
- *$C^1(\mathbb{R})$ for the algebra of continuously differentiable functions from $\mathbb{R}$ to $\mathbb{R}$.*
- *$C^\infty(\mathbb{R})$ for the algebra of smooth (= infinitely differentiable) functions from $\mathbb{R}$ to $\mathbb{R}$.*

Proof. (a) As $x \to p$ then

$$\begin{aligned} & \frac{[f(x) + g(x)] - [f(p) + g(p)]}{x - p} \\ = \; & \left(\frac{f(x) - f(p)}{x - p}\right) + \left(\frac{g(x) - g(p)}{x - p}\right) \\ \to \; & f'(p) + g'(p) \end{aligned}$$

by AOL.

(b) As $x \to p$ then

$$\begin{aligned} & \frac{f(x)g(x) - f(p)g(p)}{x-p} \\ = \; & f(x)\left(\frac{g(x)-g(p)}{x-p}\right) + \left(\frac{f(x)-f(p)}{x-p}\right)g(p) \\ \to \; & f(p)g'(p) + f'(p)g(p), \end{aligned}$$

by AOL and the continuity of f at p.

(c) Given products of differentiable functions are differentiable, we need only prove that $1/g$ is differentiable at p. Now

$$\frac{1}{x-p}\left(\frac{1}{g(x)} - \frac{1}{g(p)}\right) = \frac{1}{g(x)g(p)}\left(\frac{g(p)-g(x)}{x-p}\right).$$

As $g(p) \neq 0$, then $g(x) \neq 0$ in some neighbourhood around p. So, as $x \to p$, we have

$$\frac{1}{x-p}\left(\frac{1}{g(x)} - \frac{1}{g(p)}\right) \to \frac{-g'(p)}{g(p)^2}.$$

by AOL and the continuity of g at p. Combined with the product rule this gives

$$\left(\frac{f}{g}\right)'(p) = f'(p)\frac{1}{g(p)} - f(p)\frac{g'(p)}{g(p)^2} = \frac{g(p)f'(p) - f(p)g'(p)}{g(p)^2}.$$

■

Example 4.19 *(a) Differentiate x^n where n is a negative integer.*
(b) Differentiate $\csc x$.

Solution. (a) The derivative of x^{-n} is $-nx^{-n-1}$. So, by the quotient rule, with $f(x) = 1$ and $g(x) = x^{-n}$, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^n) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x^{-n}}\right) = \frac{-(-nx^{-n-1})}{(x^{-n})^2} = nx^{n-1}.$$

(b) The derivative of $\sin x$ is $\cos x$. So, by the quotient rule, with $f(x) = 1$ and $g(x) = \sin x$, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}(\csc x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\sin x}\right) = \frac{-\cos x}{\sin^2 x} = -\csc x \cot x.$$

■

Proposition 4.20 ***(General Leibniz product rule)*** *Let $f, g\colon (a,b) \to \mathbb{R}$ be n-times differentiable functions and $p \in (a,b)$. Then*

$$(fg)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(p) g^{(n-k)}(p).$$

Proof. The $n = 1$ case is the standard product rule. Suppose the result holds for n and assume f, g are $(n+1)$-times differentiable. By the product rule

$$\begin{aligned}
& (fg)^{(n+1)}(x) \\
= & \sum_{k=0}^{n} \binom{n}{k} \left[f^{(k+1)}(p) g^{(n-k)}(p) + f^{(k)}(p) g^{(n-k+1)}(p) \right] \\
= & f^{(n+1)}(p)g(p) + \sum_{k=0}^{n-1} \binom{n}{k} f^{(k+1)}(p) g^{(n-k)}(p) + \sum_{k=1}^{n} \binom{n}{k} f^{(k)}(p) g^{(n-k+1)}(p) + f(p) g^{(n+1)}(p) \\
= & f^{(n+1)}(p)g(p) + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] f^{(k)}(p) g^{(n-k+1)}(p) + f(p) g^{(n+1)}(p) \\
= & f^{(n+1)}(p)g(p) + \sum_{k=1}^{n} \binom{n+1}{k} f^{(k)}(p) g^{(n+1-k)}(p) + f(p) g^{(n+1)}(p) \\
= & \sum_{k=0}^{n+1} \binom{n+1}{k} f^{(k)}(p) g^{(n+1-k)}(p).
\end{aligned}$$

The result follows by induction. ■

Proposition 4.21 ***(Chain Rule)*** *Let $f, g \colon \mathbb{R} \to \mathbb{R}$ be differentiable functions, with $p \in \mathbb{R}$. Then $g \circ f$ is differentiable and*

$$(g \circ f)'(p) = f'(p) g'(f(p)).$$

Remark 4.22 *It's quite inviting to think one can simply write*

$$\frac{g(f(x)) - g(f(p))}{x - p} = \left(\frac{g(f(x)) - g(f(p))}{f(x) - f(p)} \right) \left(\frac{f(x) - f(p)}{x - p} \right)$$

and let $x \to p$. The problem with this argument is that $f(x) - f(p)$ might be zero arbitrarily close to p. This explains why we need to take a little care with the following proof.

Proof. We first introduce $h \colon \mathbb{R} \to \mathbb{R}$ by

$$h(y) = \begin{cases} \frac{g(y) - g(f(p))}{y - f(p)} & \text{when } y \neq f(p), \\ g'(f(p)) & \text{when } y = f(p). \end{cases}$$

As g is differentiable at $f(p)$, then h is continuous, so that

$$\lim_{x \to p} h(f(p)) = g'(f(p)),$$

as f is also continuous. Further, when $x \neq p$, even if $f(x) = f(p)$, we have

$$\frac{g(f(x)) - g(f(p))}{x - p} = h(f(x)) \left(\frac{f(x) - f(p)}{x - p} \right).$$

Let $x \to p$ and the limit on the RHS equals $g'(f(p)) f'(p)$. The result follows. ■

Remark 4.23 *Note that the proof only relies on g being differentiable at $f(p)$ and f being differentiable at p, so the chain rule is true with less general hypotheses than those presented here.*

You may have met the chain rule before as

$$\frac{\mathrm{d}g}{\mathrm{d}x} = \frac{\mathrm{d}g}{\mathrm{d}f}\frac{\mathrm{d}f}{\mathrm{d}x}.$$

As stated in the previous proposition, it is clearer where these derivatives are being evaluated.

Example 4.24 *(a) Differentiate x^a where $a \in \mathbb{R}$ and $x > 0$.*

(b) Differentiate a^x where $a > 0$ and $x \in \mathbb{R}$.

Solution. We showed earlier that the derivatives of e^x and $\log x$ are respectively e^x and x^{-1}.
(a) By definition $x^a = \exp(a \log x)$. So, by the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}x^a = \frac{\mathrm{d}}{\mathrm{d}x}e^{a\log x} = \frac{a}{x}e^{a\log x} = \frac{a}{x}x^a = ax^{a-1}.$$

(b) By definition $a^x = \exp(x \log a)$. So, by the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = \frac{\mathrm{d}}{\mathrm{d}x}e^{x\log a} = \log a \times e^{x\log a} = \log a \times a^x.$$

■

Example 4.25 *(a) Differentiate $\cos x$.*

(b) Differentiate $\tan x$.

(c) Differentiate $\sinh x$.

(d) Differentiate $\cos(x^2)$.

(e) Find an antiderivative of $x^n \log x$.

Solution. (a) We have $\cos x = \sin(\pi/2 - x)$ and know that $\sin' x = \cos x$. By the chain rule,

$$\cos' x = \frac{\mathrm{d}}{\mathrm{d}x}\sin\left(\frac{\pi}{2} - x\right) = -\cos\left(\frac{\pi}{2} - x\right) = -\sin x.$$

(b) Recall that $\tan x = \sin x/\cos x$ so that, by the quotient rule,

$$\tan' x = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\sin x}{\cos x}\right) = \frac{\cos x(\cos x) - \sin x(-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.$$

(c) By the sum and chain rules

$$\sinh' x = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{e^x - e^{-x}}{2}\right) = \frac{e^x - (-e^{-x})}{2} = \cosh x.$$

(d) By the chain rule $\cos(x^2)$ differentiates to $-2x\sin(x^2)$.
(e) We can write

$$x^n \log x = \frac{1}{n+1}x^n \log\left(x^{n+1}\right) = \frac{1}{(n+1)^2}u'(x)\log u(x)$$

where $u(x) = x^{n+1}$. Knowing an antiderivative of $\log x$ to be $x\log x - x$, we see from the chain rule that an antiderivative of $x^n \log x$ is

$$\frac{u\log u - u}{(n+1)^2} = \frac{x^{n+1}\log x^{n+1} - x^{n+1}}{(n+1)^2} = \frac{x^{n+1}\log x}{n+1} - \frac{x^{n+1}}{(n+1)^2}.$$

■

4.3 Extrema and Mean-values

Definition 4.26 *Let $S \subseteq \mathbb{R}$, $p \in I$ and $f : S \to \mathbb{R}$. We say p is a*
(a) ***global maximum*** *if $f(x) \leqslant f(p)$ for all $x \in S$.*
(b) ***global minimum*** *if $f(x) \geqslant f(p)$ for all $x \in S$.*
(c) ***local maximum*** *if there exists $\delta > 0$ such that $f(x) \leqslant f(p)$ for all $x \in S$ with $|x - p| < \delta$.*
(d) ***local minimum*** *if there exists $\delta > 0$ such that $f(x) \geqslant f(p)$ for all $x \in S$ with $|x - p| < \delta$.*
Note that a global extremum is automatically a local one.

Example 4.27 *(a) The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = \sin x$ has infinitely many global maxima, where $\sin x = 1$, and infinitely many global minima, where $\sin x = -1$.*

(b) The function $g : [0, 2\pi] \to \mathbb{R}$ given by $g(x) = \sin x$ has a single global maximum at $\pi/2$, a single global minimum at $3\pi/2$, a local minimum at 0 and a local maximum at 2π.

(c) The function $h : [-1, 2] \to \mathbb{R}$ given by $h(x) = x^3 - x$ has a global maximum at $x = 2$, a global minimum at $x = 1/\sqrt{3}$, a local minimum at $x = -1$ and a local maximum at $x = -1/\sqrt{3}$. (See the explanation following the next theorem.)

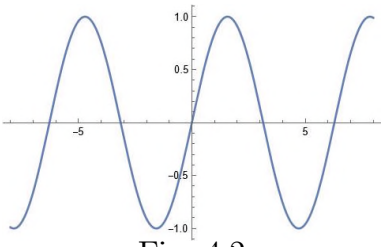


Fig. 4.2

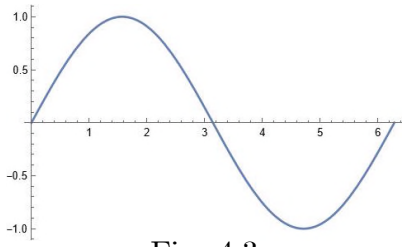


Fig. 4.3

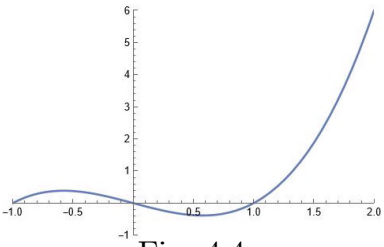


Fig. 4.4

Theorem 4.28 *(**Fermat's theorem**, 1629-30) Let $f : (a, b) \to \mathbb{R}$ be differentiable with p a local extremum. Then $f'(p) = 0$.*

Remark 4.29 *Note it's important that p is an internal local extremum. In Fig. 4.4 we see that -1 and 2 are local extrema where the derivative is non-zero.*

Proof. WLOG suppose that p is a local maximum, so that $f(p) \geqslant f(x)$ on some $(p - \delta, p + \delta)$. So for $p - \delta < x < p$ we have

$$\frac{f(x) - f(p)}{x - p} \geqslant 0,$$

as the numerator is non-positive and denominator is negative. And for $p < x < p + \delta$ we have

$$\frac{f(x) - f(p)}{x - p} \leqslant 0,$$

as the numerator is non-positive and denominator is positive. Recalling that limits respect weak inequalities, we have both $f'(p) \geqslant 0$ and $f'(p) \leqslant 0$. The result follows. ■

Theorem 4.30 *(**Rolle's theorem**, 1691) Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$, differentiable on (a,b), and such that $f(a) = f(b)$. Then there exists $c \in (a,b)$ such that $f'(c) = 0$.*

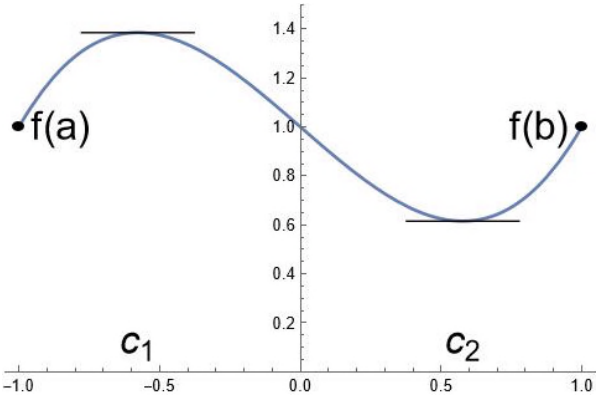


Fig. 4.5. – Rolle's theorem

Proof. By the boundedness theorem, f achieves its extrema. If f is constant, we may take any $c \in (a,b)$. Otherwise, f has a global extremum, distinct from $f(a) = f(b)$, and which is achieved internally at some $c \in (a,b)$. By Fermat's theorem, $f'(c) = 0$. ■

Example 4.31 *Let $f\colon \mathbb{R} \to \mathbb{R}$ be twice differentiable such that $f''(x) > 0$ for all x.. Show that f has at most two roots. Give examples to show that f may have $0, 1$ or 2 roots.*

Solution. SFAC that f has three roots $x_1 < x_2 < x_3$. By Rolle's theorem, f' has roots $c_1 \in (x_1, x_2)$ and $c_2 \in (x_2, x_3)$. But then, by Rolle again, f'' has a root in (c_1, c_2), contradicting the hypotheses.

The functions $x^2 + 1$, x^2, $x^2 - 1$ all have a positive second derivative of 2, and respectively have $0, 1, 2$ roots. ■

The above argument can be readily generalized to demonstrate the following.

Corollary 4.32 *Let $f\colon \mathbb{R} \to \mathbb{R}$ be n-times differentiable and suppose that $f^{(n)}$ has k roots. Then f has at most $n + k$ roots.*

Combined with the IVT, this corollary can sometimes be used to determine the exact number of real roots of a polynomial.

Example 4.33 *Determine the number of real solutions to the equation $x^5 + x^3 - 7x - 2 = 0$.*

Solution. Let $f(x) = x^5 + x^3 - 7x - 2 = 0$, so that

$$f'(x) = 5x^4 + 3x^2 - 7.$$

This is a quadratic in x^2 with discriminant $3^2 - 4 \times 5 \times (-7) > 0$. Hence there are two possible values of x^2 that solve $f'(x) = 0$. However their product is -7, so one of these values is negative yielding no real x. That is, $f'(x) = 0$ has two real roots.

By the above, $f(x) = 0$ has at most three real roots. We can see that

$$f(-2) = -28, \quad f(-1) = 3, \quad f(0) = -2, \quad f(2) = 24.$$

As $f(x)$ is continuous, f has three roots in the intervals $(-2,-1)$, $(-1,0)$ and $(0,2)$ by the IVT.

Combining all the above, we see that $f(x)$ has precisely three real roots. ■

Theorem 4.34 ***(Mean-value theorem)*** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$, differentiable on (a,b). Then there exists $c \in (a,b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b-a}.$$

The theorem gets its name as the expression on the RHS of the equation is the mean value of the gradient.

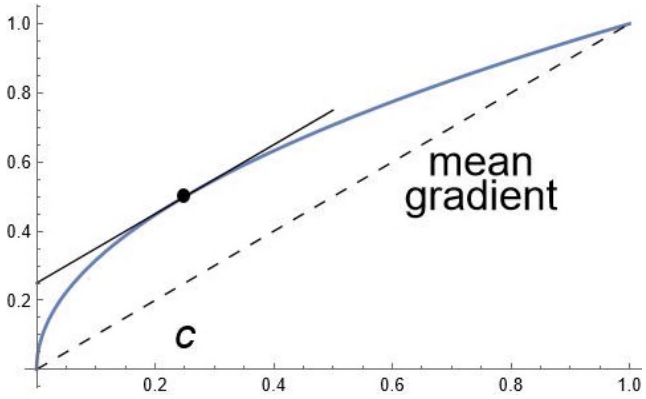


Fig. 4.6. – MVT

Solution. Set $g(x) = f(x) - K(x-a)$ where

$$K = \frac{f(b) - f(a)}{b-a}.$$

Then g is continuous on $[a,b]$, differentiable on (a,b) and

$$g(b) = f(b) - (f(b) - f(a)) = f(a) = g(a).$$

By Rolle's theorem, there exists $c \in (a,b)$ such that

$$0 = g'(c) = f'(c) - K.$$

Once rearranged, the result follows. ■

Corollary 4.35 ***(Constancy theorem)*** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$, differentiable on (a,b) and such that $f'(x) = 0$ for all $x \in (a,b)$. Then f is constant.*

Proof. Take $a \leqslant x < y \leqslant b$. By the MVT, there is $c \in (x,y)$ such that

$$0 = f'(c) = \frac{f(y) - f(x)}{y-x}.$$

It follows that $f(x) = f(y)$ for all $x, y \in [a,b]$; that is, f is constant. ■

Example 4.36 *Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable and such that*

$$f'(x) = f(x) \qquad f(0) = 1.$$

Show that $f(x) = e^x$.

Solution. We introduce $g(x) = f(x)e^{-x}$. By the product and chain rules we have

$$g'(x) = f'(x)e^{-x} - f(x)e^{-x} = (f'(x) - f(x))\, e^{-x} = 0.$$

By the constancy theorem, $g(x)$ is constant; we note $g(0) = f(0)e^{-0} = 1$ and hence $g(x) = f(x)e^{-x} = 1$. The result follows. ■

Proposition 4.37 *A differentiable function $f\colon I \to \mathbb{R}$ with bounded derivative is Lipschitz continuous (and so uniformly continuous).*

Proof. Say that $|f'(x)| < K$ for all x. By the MVT for $x < y$ there exists $c \in (x, y)$ such that

$$|f(x) - f(y)| = |f'(c)|\,|x - y| \leqslant K\,|x - y|\,,$$

showing f is Lipschitz. ■

4.4 Inverse functions – a reprise

There is a differential equivalent of the previous CIFT. The graph of the inverse, guaranteed by the CIFT, is the reflection of the original graph in the $y = x$ line. So we would expect that a curve will gradient m would reflect to a curve with gradient m^{-1} at the correspondent point. This explains the extra hypothesis in the theorem below .

Theorem 4.38 ***(DIFT)*** *Let $f\colon [a, b] \to [c, d]$ be a strictly monotonic, continuous function, so that f is invertible and $g = f^{-1}$ is continuous by CIFT. If f is differentiable at $p \in (a, b)$ and $f'(p) \neq 0$, then g is differentiable at $f(p)$ and*

$$g'(f(p)) = \frac{1}{f'(p)}.$$

Proof. We will write $y = f(x)$ for $x \in [a, b]$, so that $x - g(y)$. Should the last limit below exist we have

$$g'(f(p)) = \lim_{y \to f(p)} \frac{g(y) - g(f(p)}{y - f(p)} = \lim_{y \to f(p)} \frac{x - p}{f(x) - f(p)}.$$

By CIFT g is continuous and so $x \to p$ as $y \to f(p)$ and hence

$$\lim_{y \to f(p)} \frac{x - p}{f(x) - f(p)} = \left(\lim_{y \to f(p)} \frac{f(x) - f(p)}{x - p} \right)^{-1} = \frac{1}{f'(p)}$$

by AOL. Thus the limit exists and equals $g'(f(p))$. ■

Corollary 4.39 *With the hypotheses of DIFT, but further assuming that f is twice differentiable at p, then g is twice differentiable at $f(p)$ and*

$$g'(f(p)) = \frac{-f''(p)}{f'(p)^2}.$$

Remark 4.40 *Note that we cannot simply use the chain rule to differentiate $g(f(x)) = x$ and find $g'(f(x))f'(x) = 1$ as this presumes that g is differentiable.*

Example 4.41 *(a)* ***Logarithm.*** *Recall that* $\exp\colon \mathbb{R} \to (0,\infty)$ *has inverse* $\log\colon (0,\infty) \to \mathbb{R}$. *Noting* $\exp' x = \exp x \neq 0$, *by DIFT,*

$$\log'(y) = \frac{1}{\exp x} = \frac{1}{y}.$$

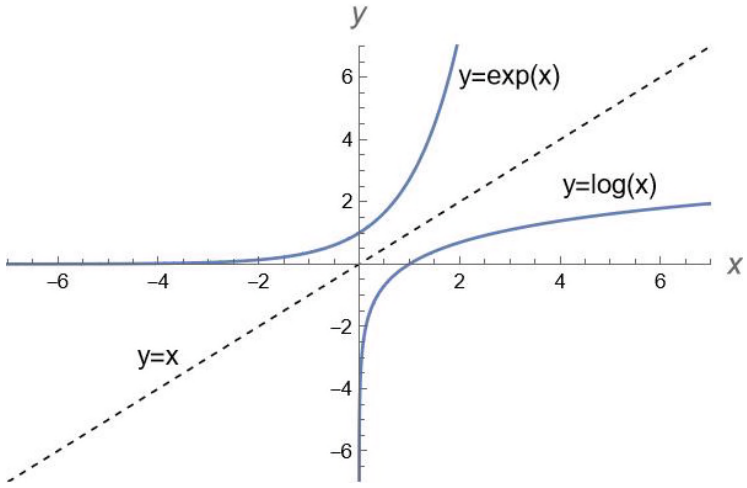


Fig 4.7. exp *and* log *graphs*

(b) ***Inverse trigonometric functions.*** *Sine, restricted as* $\sin\colon [-\pi/2, \pi/2] \to [-1,1]$, *satisfies the hypotheses of DIFT and we find* $\arcsin\colon [-1,1] \to [-\pi/2, \pi/2]$ *satisfies*

$$\arcsin'(y) = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}}.$$

Similarly cosine, restricted as $\cos\colon [0,\pi] \to [-1,1]$, *satisfies the hypotheses of DIFT and we find* $\arccos\colon [-1,1] \to [0,\pi]$ *satisfies*

$$\arccos'(y) = \frac{1}{-\sin x} = \frac{-1}{\sqrt{1-\cos^2 x}} = \frac{-1}{\sqrt{1-y^2}}.$$

And tangent, restricted as $\tan\colon (-\pi/2, \pi/2) \to (-\infty,\infty)$, *satisfies the hypotheses of DIFT and we find* $\arctan\colon (-\infty,\infty) \to (-\pi/2, \pi/2)$ *satisfies*

$$\arctan'(y) = \frac{1}{\sec^2 x} = \frac{1}{1+\tan^2 x} = \frac{1}{1+y^2}.$$

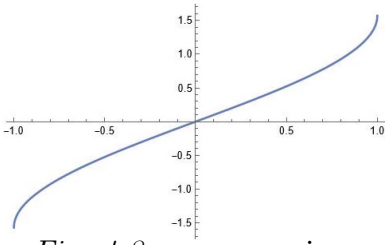


Fig. 4.8. – $y = \arcsin x$

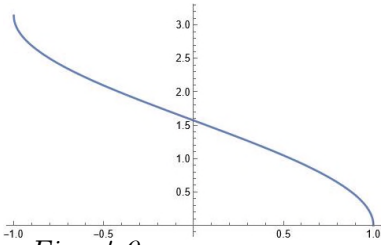


Fig. 4.9. – $y = \arccos x$

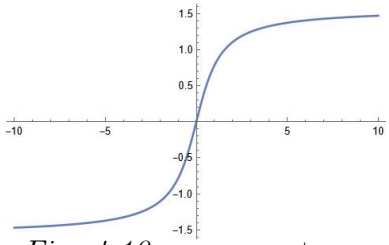


Fig. 4.10. – $y = \arctan x$

Fig. 5.5 – the inverse trigonometric functions

Example 4.42 *(c)* ***Inverse hyperbolic functions.*** $\sinh\colon \mathbb{R} \to \mathbb{R}$ *satisfies the hypotheses of DIFT and we find* $\operatorname{arcsinh}\colon \mathbb{R} \to \mathbb{R}$ *satisfies*

$$\operatorname{arcsinh}'(y) = \frac{1}{\cosh x} = \frac{1}{\sqrt{1+\sinh^2 x}} = \frac{1}{\sqrt{1+y^2}}.$$

$\cosh\colon [0,\infty) \to [1,\infty)$ *satisfies the hypotheses of DIFT and we find* $\operatorname{arccosh}\colon [1,\infty) \to [0,\infty)$ *satisfies*

$$\operatorname{arccosh}'(y) = \frac{1}{\sinh x} = \frac{1}{\sqrt{\cosh^2 x - 1}} = \frac{1}{\sqrt{y^2-1}}.$$

$\tanh\colon \mathbb{R} \to (-1,1)$ *satisfies the hypotheses of DIFT and we find* $\operatorname{arctanh}\colon (-1,1) \to \mathbb{R}$ *satisfies*

$$\operatorname{arctanh}'(y) = \frac{1}{\operatorname{sech}^2 x} = \frac{1}{1-\tanh^2 x} = \frac{1}{1-y^2}.$$

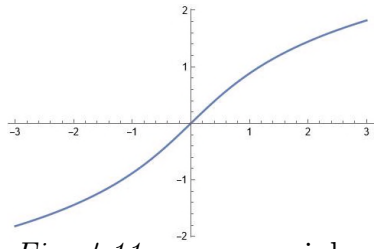


Fig. 4.11. – $y = \operatorname{arcsinh} x$

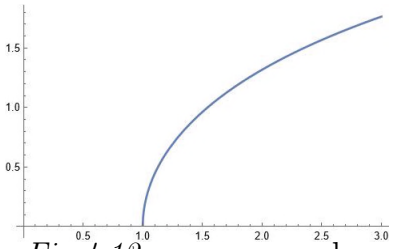


Fig 4.12. – $y = \operatorname{arccosh} x$

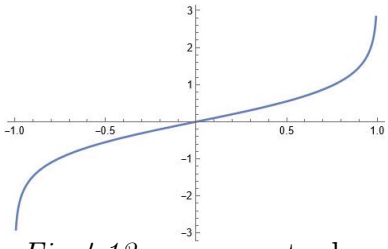


Fig 4.13. – $y = \operatorname{arctanh} x$