Chapter 4

Differentiation

4.1 The Derivative of a Function

We define the derivative of a function and prove the main properties like product, quotient and chain rule. We relate the derivative of a function with the derivative of its inverse function. We prove the mean value theorem and consider local extrema. Taylor's theorem will be formulated.

Definition 4.1 Let $f\colon (a,b) \to \mathbb{R}$ be a function and $x_0 \in (a,b)$. If the limit

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \tag{4.1}$$

exists, we call f *differentiable* at x_0. The limit is denoted by $f'(x_0)$. We say f *is differentiable* if f is differentiable at every point $x \in (a,b)$. We thus have associated to every function f a function f' whose domain is the set of points x_0 where the limit (4.1) exists; f' is called the *derivative* of f.

Sometimes the Leibniz notation is used to denote the derivative of f

$$f'(x_0) = \frac{\mathrm{d}f(x_0)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} f(x_0).$$

Remarks 4.1 (a) Replacing $h := x - x_0$ we see that $f'(x_0) = \lim\limits_{h\to 0} \dfrac{f(x_0+h)-f(x_0)}{h}$.
(b) The limits

$$\lim_{h\to 0-0} \frac{f(x_0+h)-f(x_0)}{h}, \quad \lim_{h\to 0+0} \frac{f(x_0+h)-f(x_0)}{h}$$

are called *left-hand and right-hand derivatives of f in x_0*, respectively. In particular for $f\colon [a,b] \to \mathbb{R}$, we can consider the right-hand derivative at a and the left-hand derivative at b.

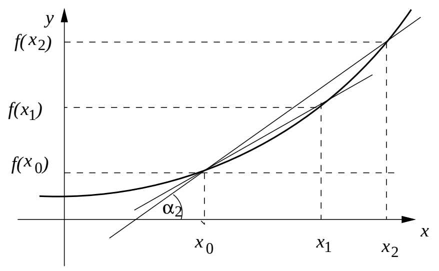
Example 4.1 (a) For $f(x) = c$ the constant function

$$f'(x_0) = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{x\to x_0} \frac{c-c}{x-x_0} = 0.$$

(b) For $f(x) = x$,

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

(c) The slope of the tangent line. Given a function $f\colon (a,b) \to \mathbb{R}$ which is differentiable in x_0. Then $f'(x_0)$ is the slope of the tangent line to the graph of f through the point $(x_0, f(x_0))$.



The slopes of the two secant lines are

$$m_2 = \tan\alpha_2 = \frac{f(x_2) - f(x_0)}{x_2 - x_0}, \quad m_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

One can see: If x approaches x_0, the secant line through $(x_0, f(x_0))$ and $(x, f(x))$ approaches the tangent line through $(x_0, f(x_0))$. Hence, the slope of the tangent line is the limit of the slopes of the secant lines if x approaches x_0:

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

Proposition 4.1 *Let f be defined on (a,b). If f is differentiable at a point $x_0 \in (a,b)$, then f is continuous at x_0.*

Proof. By Proposition 3.2 we have

$$\lim_{x \to x_0} (f(x) - f(x_0)) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) = f'(x_0) \lim_{x \to x_0} (x - x_0) = f'(x_0) \cdot 0 = 0.$$

■

The converse of this proposition is not true. For example $f(x) = |\,x\,|$ is continuous in $\mathbb{R}$ but differentiable in $\mathbb{R} \setminus \{0\}$ since $\lim\limits_{h \to 0+0} \frac{|\,h\,|}{h} = 1$ whereas $\lim\limits_{h \to 0-0} \frac{|\,h\,|}{h} = -1$. Later we will become aquainted with a function which is continuous on the whole line without being differentiable at any point!

Proposition 4.2 *Let $f\colon (r,s) \to \mathbb{R}$ be a function and $a \in (r,s)$. Then f is differentiable at a if and only if there exists a number $c \in \mathbb{R}$ and a function φ defined in a neighborhood of a such that*

$$f(x) = f(a) + (x - a)c + \varphi(x), \tag{4.2}$$

where

$$\lim_{x \to a} \frac{\varphi(x)}{x - a} = 0. \tag{4.3}$$

In this case $f'(a) = c$.

The proposition says that a function f differentiable at a can be approximated by a linear function, in our case by

$$y = f(a) + (x-a)f'(a).$$

The graph of this linear function is the tangent line to the graph of f at the point $(a, f(a))$. Later we will use this point of view to define differentiability of functions $f\colon \mathbb{R}^n \to \mathbb{R}^m$.
Proof. Suppose first f satisfies (4.2) and (4.3). Then

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{x\to a}\left(c + \frac{\varphi(x)}{x-a}\right) = c.$$

Hence, f is differentiable at a with $f'(a) = c$.
Now, let f be differentiable at a with $f'(a) = c$. Put $\varphi(x) = f(x) - f(a) - (x-a)f'(a)$. Then

$$\lim_{x\to a} \frac{\varphi(x)}{x-a} = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} - f'(a) = 0.$$

■

Proposition 4.3 *Suppose f and g are defined on (a,b) and are differentible at a point $x \in (a,b)$. Then $f+g$, fg, and f/g are differentiable at x and*

(a) $(f+g)'(x) = f'(x) + g'(x)$;
(b) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$;
(c) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$.

In (c), *we assume that* $g(x) \neq 0$.

Proof. (a) Since

$$\frac{(f+g)(x+h)-(f+g)(x)}{h} = \frac{f(x+h)-f(x)}{h} + \frac{g(x+h)-g(x)}{h},$$

the claim follows from Proposition 3.2.
Let $h = fg$ and t be variable. Then

$$h(t) - h(x) = f(t)(g(t)-g(x)) + g(x)(f(t)-f(x))$$
$$\frac{h(t)-h(x)}{t-x} = f(t)\frac{g(t)-g(x)}{t-x} + g(x)\frac{f(t)-f(x)}{t-x}.$$

Noting that $f(t) \to f(x)$ as $t \to x$, (b) follows.
Next let $h = f/g$. Then

$$\begin{aligned}\frac{h(t)-h(x)}{t-x} &= \frac{\frac{f(t)}{g(t)} - \frac{f(x)}{g(x)}}{t-x} = \frac{f(t)g(x)-f(x)g(t)}{g(x)g(t)(t-x)}\\ &= \frac{1}{g(t)g(x)}\frac{f(t)g(x)-f(x)g(x)+f(x)g(x)-f(x)g(t)}{t-x}\\ &= \frac{1}{g(t)g(x)}\left(g(x)\frac{f(t)-f(x)}{t-x} - f(x)\frac{g(t)-g(x)}{t-x}\right).\end{aligned}$$

Letting $t \to x$, and applying Propositions 3.2 and 4.1, we obtain (c). ■

Example 4.2 (a) $f(x) = x^n$, $n \in \mathbb{Z}$. We will prove $f'(x) = nx^{n-1}$ by induction on $n \in \mathbb{N}$. The cases $n = 0, 1$ are OK by Example 4.1. Suppose the statement is true for some fixed n. We will show that $(x^{n+1})' = (n+1)x^n$.
By the product rule and the induction hypothesis

$$(x^{n+1})' = (x^n \cdot x)' = (x^n)'x + x^n(x') = nx^{n-1}\,x + x^n = (n+1)x^n.$$

This proves the claim for positive integers n. For negative n consider $f(x) = 1/x^{-n}$ and use the quotient rule.
(b) $(\mathrm{e}^x)' = \mathrm{e}^x$.

$$(\mathrm{e}^x)' = \lim_{h\to 0} \frac{\mathrm{e}^{x+h} - \mathrm{e}^x}{h} = \lim_{h\to 0} \frac{\mathrm{e}^x\mathrm{e}^h - \mathrm{e}^x}{h} = \mathrm{e}^x \lim_{h\to 0} \frac{\mathrm{e}^h - 1}{h} = \mathrm{e}^x; \tag{4.4}$$

the last equation simply follows from Homework 11.4 (c).
(c) $(\sin x)' = \cos x$, $(\cos x)' = -\sin x$. Using $\sin(x+y) - \sin(x-y) = 2\cos(x)\sin(y)$ we have

$$\begin{aligned}(\sin x)' &= \lim_{h\to 0} \frac{\sin(x+h) - \sin x}{h} = \lim_{h\to 0} \frac{2\cos\frac{2x+h}{2}\sin\frac{h}{2}}{h}\\ &= \lim_{h\to 0} \cos\left(x + \frac{h}{2}\right) \lim_{h\to 0} \frac{\sin\frac{h}{2}}{\frac{h}{2}}.\end{aligned}$$

Since $\cos x$ is continuous and $\lim\limits_{h\to 0} \dfrac{\sin h}{h} = 1$ by Proposition 3.25 (b), we obtain $(\sin x)' = \cos x$. The proof for $\cos x$ is analogous.
(d) $(\tan x)' = \dfrac{1}{\cos^2 x}$. Using the quotiont rule for the function $\tan x = \sin x / \cos x$ we have

$$(\tan x)' = \frac{(\sin x)'\cos x - \sin x(\cos x)'}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

The next proposition deals with composite functions and is probably the most important statement about derivatives.

Proposition 4.4 (Chain rule) *Let $g\colon (\alpha,\beta) \to \mathbb{R}$ be differentiable at $x_0 \in (\alpha,\beta)$ and let $f\colon (a,b) \to \mathbb{R}$ be differentiable at $y_0 = g(x_0) \in (a,b)$. Then $h = f \circ g$ is differentiable at x_0, and*

$$h'(x_0) = f'(y_0)g'(x_0). \tag{4.5}$$

Proof. We have

$$\begin{aligned}\frac{f(g(x)) - f(g(x_0))}{x - x_0} &= \frac{f(g(x)) - f(g(x_0))}{g(x) - g(x_0)}\,\frac{g(x) - g(x_0)}{x - x_0}\\ &\underset{x\to x_0}{\longrightarrow} \lim_{y\to y_0} \frac{f(y) - f(y_0)}{y - y_0} \cdot g'(x_0) = f'(y_0)g'(x_0).\end{aligned}$$

Here we used that $y = g(x)$ tends to $y_0 = g(x_0)$ as $x \to x_0$, since g is continuous at x_0. ∎

Example 4.3 (a) Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable; define $F\colon \mathbb{R} \to \mathbb{R}$ by $F(x) := f(ax+b)$ with some $a, b \in \mathbb{R}$. Then

$$F'(x) = af'(ax+b).$$

(b) $x^\alpha = \mathrm{e}^{\alpha \log x}$. Hence, $(x^\alpha)' = (\mathrm{e}^{\alpha \log x})' = \mathrm{e}^{\alpha \log x} \alpha \dfrac{1}{x} = \alpha x^{\alpha - 1}$.

(c) Suppose $f > 0$ and $g = \log f$. Then $g' = f' \dfrac{1}{f}$; hence $f' = f\, g'$.

Proposition 4.5 *Let $f\colon (a,b) \to \mathbb{R}$ be strictly monotonic and continuous. Suppose f is differentiable at x. Then the inverse function $g = f^{-1}\colon f((a,b)) \to \mathbb{R}$ is differentiable at $y = f(x)$ with*

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{f'(g(y))}. \tag{4.6}$$

Proof. Let $(y_n) \subset f((a,b))$ be a sequence with $y_n \to y$ and $y_n \neq y$ for all n. Put $x_n = g(y_n)$. Since g is continuous (by Corollary 3.16), $\lim_{n\to\infty} x_n = x$. Since g is injective, $x_n \neq x$ for all n. We have

$$\lim_{n\to\infty} \frac{g(y_n) - g(y)}{y_n - y} = \lim_{n\to\infty} \frac{x_n - x}{f(x_n) - f(x)} = \lim_{n\to\infty} \frac{1}{\frac{f(x_n) - f(x)}{x_n - x}} = \frac{1}{f'(x)}.$$

Hence $g'(y) = 1/f'(x) = 1/f'(g(y))$. ∎

We give some applications of this very useful proposition.

Example 4.4 In what follows f is the original function (with known derivative) and g is the inverse function to f. We fix the notion $y = f(x)$ and $x = g(y)$.
(c) $\log\colon \mathbb{R}_+ \setminus \{0\} \to \mathbb{R}$ is the inverse function to $f(x) = \mathrm{e}^x$. By the above proposition

$$(\log y)' = \frac{1}{(\mathrm{e}^x)'} = \frac{1}{\mathrm{e}^x} = \frac{1}{y}.$$

(d) $\arcsin\colon [-1,1] \to \mathbb{R}$ is the inverse function to $y = f(x) = \sin x$. If $x \in (-1,1)$ then

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

Since $y \in [-1,1]$ implies $x = \arcsin y \in [-\pi/2, \pi/2]$, $\cos x \geq 0$. Therefore, $\cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2}$. Hence

$$(\arcsin y)' = \frac{1}{\sqrt{1 - y^2}}, \quad -1 < y < 1.$$

Note that the derivative is not defined at the endpoints $y = -1$ and $y = 1$.
(e)

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

Since $y = \tan x$ we have

$$\begin{aligned} y^2 = \tan^2 x &= \frac{\sin^2 x}{\cos^2 x} = \frac{1 - \cos^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} - 1 \\ \cos^2 x &= \frac{1}{1 + y^2} \\ (\arctan y)' &= \frac{1}{1 + y^2}. \end{aligned}$$

For $a > 0$, $a \neq 1$ and $x > 0$ put

$$\log_a x = \frac{\log x}{\log a}.$$

Then

$$a^{\log_a x} = x = (\log_a a^x).$$

4.2 The Derivatives of Elementary Functions

function	derivative
const.	0
$x^n \ (n \in \mathbb{N})$	nx^{n-1}
$x^\alpha \ (\alpha \in \mathbb{R}, x > 0)$	$\alpha x^{\alpha-1}$
e^x	e^x
$a^x, \ (a > 0)$	$a^x \log a$
$\log x$	$\frac{1}{x}$
$\log_a x$	$\frac{1}{x \log a}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\frac{1}{\cos^2 x}$
$\cot x$	$-\frac{1}{\sin^2 x}$
$\sinh x$	$\cosh x$
$\cosh x$	$\sinh x$
$\tanh x$	$\frac{1}{\cosh^2 x}$
$\coth x$	$-\frac{1}{\sinh^2 x}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
$\arctan x$	$\frac{1}{1+x^2}$
$\operatorname{arccot} x$	$-\frac{1}{1+x^2}$
$\operatorname{arsinh} x$	$\frac{1}{\sqrt{x^2+1}}$
$\operatorname{arcosh} x$	$\frac{1}{\sqrt{x^2-1}}$
$\operatorname{artanh} x$	$\frac{1}{1-x^2}$
$\operatorname{arcoth} x$	$\frac{1}{1-x^2}$

4.2.1 Derivatives of Higher Order

Let $f\colon D \to \mathbb{R}$ be differentiable. If the derivative $f'\colon D \to \mathbb{R}$ is differentiable at $x \in D$, then

$$\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} = f''(x) = (f')'(x)$$

is called the *second derivative* of f at x. Similarly, one defines inductively higher order derivatives. Continuing in this manner, we obtain functions

$$f,\ f',\ f'',\ f^{(3)},\ \dots,\ f^{(k)}$$

each of which is the derivative of the preceding one. $f^{(n)}$ is called the *nth derivative* of f or the derivative of order n of f. We also use the Leibniz notation

$$f^{(k)}(x) = \frac{\mathrm{d}^k f(x)}{\mathrm{d}x^k} = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k f(x).$$

Definition 4.2 Let $D \subset \mathbb{R}$ and $k \in \mathbb{N}$ a positive integer. We denote by $\mathrm{C}^k(D)$ the set of all functions $f\colon D \to \mathbb{R}$ such that $f^{(k)}(x)$ exists for all $x \in D$ and $f^{(k)}(x)$ is continuous. Obviously $\mathrm{C}(D) \supset \mathrm{C}^1(D) \supset \mathrm{C}^2(D) \supset \cdots$. Further, we set

$$\mathrm{C}^\infty(D) = \bigcap_{k\in\mathbb{N}} \mathrm{C}^k(D). \tag{4.7}$$

$f \in \mathrm{C}^k(D)$ is called *k times continuously differentiable.*

Using induction over n, one proves the following proposition.

Proposition 4.6 (Leibniz formula) *Let f and g be n times differentiable. Then fg is n times differentiable with*

$$(f(x)g(x))^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x). \tag{4.8}$$

4.3 Local Extrema and the Mean Value Theorem

Many properties of a function f like monotony, convexity, and existence of local extrema can be studied using the derivative f'. From estimates for f' we obtain estimates for the growth of f.

Definition 4.3 Let $f\colon [a,b] \to \mathbb{R}$ be a function. We say that f has a *local maximum* at the point ξ, $\xi \in (a,b)$, if there exists $\delta > 0$ such that $f(x) \le f(\xi)$ for all $x \in [a,b]$ with $|\,x - \xi\,| < \delta$. *Local minima* are defined likewise.

We say that ξ is a *local extremum* if it is either a local maximum or a local minimum.

Proposition 4.7 *Let f be defined on $[a,b]$. If f has a local extremum at a point $\xi \in (a,b)$, and if $f'(\xi)$ exists, then $f'(\xi) = 0$.*

Proof. Suppose f has a local maximum at ξ. According with the definition choose $\delta > 0$ such that

$$a < \xi - \delta < \xi < \xi + \delta < b.$$

If $\xi - \delta < x < \xi$, then

$$\frac{f(x) - f(\xi)}{x - \xi} \geq 0.$$

Letting $x \to \xi$, we see that $f'(\xi) \geq 0$.
If $\xi < x < \xi + \delta$, then

$$\frac{f(x) - f(\xi)}{x - \xi} \leq 0.$$

Letting $x \to \xi$, we see that $f'(\xi) \leq 0$. Hence, $f'(\xi) = 0$. ∎

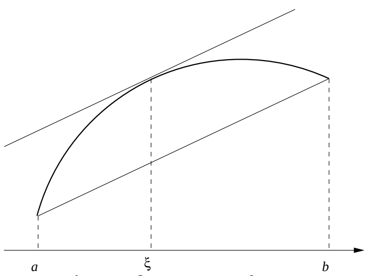
Remarks 4.2 (a) $f'(x) = 0$ is a necessary but not a sufficient condition for a local extremum in x. For example $f(x) = x^3$ has $f'(x) = 0$, but x^3 has no local extremum.
(b) If f attains its local extrema at the boundary, like $f(x) = x$ on $[0, 1]$, we do not have $f'(\xi) = 0$.

Theorem 4.8 (Rolle's Theorem) *Let $f\colon [a, b] \to \mathbb{R}$ be continuous with $f(a) = f(b)$ and let f be differentiable in (a, b). Then there exists a point $\xi \in (a, b)$ with $f'(\xi) = 0$.*

In particular, between two zeros of a differentiable function there is a zero of its derivative.
Proof. If f is the constant function, the theorem is trivial since $f'(x) \equiv 0$ on (a, b). Otherwise, there exists $x_0 \in (a, b)$ such that $f(x_0) > f(a)$ or $f(x_0) < f(a)$. Then f attains its maximum or minimum, respectively, at a point $\xi \in (a, b)$. By Proposition 4.7, $f'(\xi) = 0$. ∎

Theorem 4.9 (Mean Value Theorem) *Let $f\colon [a, b] \to \mathbb{R}$ be continuous and differentiable in (a, b). Then there exists a point $\xi \in (a, b)$ such that*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \tag{4.9}$$



Geometrically, the mean value theorem states that there exists a tangent line through some point $(\xi, f(\xi))$ which is parallel to the secant line AB, $A = (a, f(a))$, $B = (b, f(b))$.

Proof. Define the function $F\colon [a, b] \to \mathbb{R}$ via

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

Then F is continuous in $[a,b]$ and differentiable in (a,b) and $F(a) = f(a) = F(b)$. By Rolle's theorem there exists a point $\xi \in (a,b)$ such that $F'(\xi) = 0$. Since

$$F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a},$$

the claim follows. ∎

Theorem 4.10 (Generalized Mean Value Theorem) *Let f and g be continuous functions on $[a,b]$ which are differentiable on (a,b). Then there exists a point $\xi \in (a,b)$ such that*

$$(f(b) - f(a))g'(\xi) = (g(b) - g(a))f'(\xi).$$

Proof. Put

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

Then h is continuous in $[a,b]$ and differentiable in (a,b) and

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

Again, Rolle's theorem shows that there exists $\xi \in (a,b)$ such that $h'(\xi) = 0$. The theorem follows. ∎

Corollary 4.11 *Suppose f is differentiable on (a,b).*

If $f'(x) \geq 0$ for all $x \in (a,b)$, then f in monotonically increasing.
If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant.
If $f'(x) \leq 0$ for all x in (a,b), then f is monotonically decreasing.

Proof. All conclusions can be read off from the equality

$$f(x) - f(t) = (x - t)f'(\xi)$$

which is valid for each pair x, t, $a < t < x < b$ and for some $\xi \in (t,x)$. ∎

4.3.1 Local Extrema and Convexity

Proposition 4.12 *Let $f\colon (a,b) \to \mathbb{R}$ be differentiable and suppose $f''(\xi)$ exists at a point $\xi \in (a,b)$. If*

$$f'(\xi) = 0 \quad \textit{and} \quad f''(\xi) > 0,$$

then f has a local minimum at ξ. Similarly, if

$$f'(\xi) = 0 \quad \textit{and} \quad f''(\xi) < 0,$$

f has a local maximum at ξ.

Remark 4.3 The condition of Proposition 4.12 is sufficient but not necessary for the existence of a local extremum. For example, $f(x) = x^4$ has a local minimum at $x = 0$, but $f''(0) = 0$.

Proof. We consider the case $f''(\xi) > 0$; the proof of the other case is analogous. Since

$$f''(\xi) = \lim_{x\to\xi} \frac{f'(x) - f'(\xi)}{x - \xi} > 0.$$

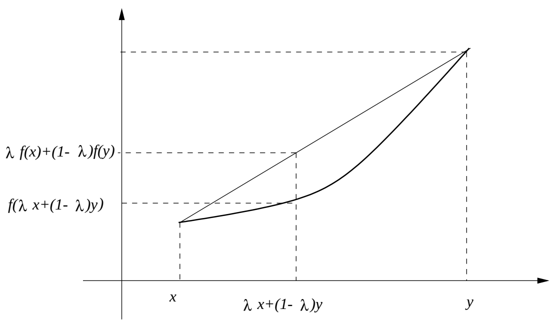
By Homework 10.4 there exists $\delta > 0$ such that

$$\frac{f'(x) - f'(\xi)}{x - \xi} > 0, \quad \text{for all } x \text{ with } \quad 0 < |\, x - \xi \,| < \delta.$$

Since $f'(\xi) = 0$ it follows that

$$\begin{aligned} f'(x) < 0 \quad &\text{if} \quad \xi - \delta < x < \xi, \\ f'(x) > 0 \quad &\text{if} \quad \xi < x < \xi + \delta. \end{aligned}$$

Hence, by Corollary 4.11, f is decreasing at $(\xi - \delta, \xi)$ and increasing at $(\xi, \xi + \delta)$. Therefore, f has a local minimum at ξ. ∎



Definition 4.4 A function $f\colon (a, b) \to \mathbb{R}$ is said to be *convex* if for all $x, y \in (a, b)$ and all $\lambda \in [0, 1]$

$$f(\lambda x + (1 - \lambda))y) \le \lambda f(x) + (1 - \lambda) f(y). \tag{4.10}$$

A function f is said to be *concave* if $-f$ is convex.

Proposition 4.13 *Suppose $f\colon (a, b) \to \mathbb{R}$ is twice differentiable. Then f is convex if and only if $f''(x) \ge 0$ for all $x \in (a, b)$.*

Proof. The proof is in Appendix C to this chapter. ∎

4.4 L'Hospital's Rule

Theorem 4.14 (L'Hospital's Rule) *Suppose f and g are differentiable in (a, b) and $g(x) \neq 0$ for all $x \in (a, b)$, where $-\infty \le a < b \le +\infty$. Suppose*

$$\lim_{x\to a+0} \frac{f'(x)}{g'(x)} = A. \tag{4.11}$$

If

$$\text{(a)} \quad \lim_{x\to a+0} f(x) = \lim_{x\to a+0} g(x) = 0 \quad \textit{or} \tag{4.12}$$

$$\text{(b)} \quad \lim_{x\to a+0} f(x) = \lim_{x\to a+0} g(x) = +\infty, \tag{4.13}$$

then

$$\lim_{x\to a+0} \frac{f(x)}{g(x)} = A. \tag{4.14}$$

The analogous statements are of course also true if $x \to b - 0$, *or if* $g(x) \to -\infty$.

Proof. First we consider the case of finite $a \in \mathbb{R}$. (a) One can extend the definition of f and g via $f(a) = g(a) = 0$. Then f and g are continuous at a. By the generalized mean value theorem, for every $x \in (a, b)$ there exists a $\xi \in (a, x)$ such that

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)}.$$

If x approaches a then ξ also approaches a, and (a) follows.
(b) Given $\varepsilon > 0$ choose $\delta > 0$ such that

$$\left| \frac{f'(t)}{g'(t)} - A \right| < \varepsilon$$

if $t \in (a, a + \delta)$. By the generalized mean value theorem for any $x, y \in (a, a + \delta)$ with $x \neq y$,

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - A \right| < \varepsilon.$$

We have

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(y)}{g(x) - g(y)} \frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}}.$$

The right factor tends to 1 as x approaches a, in particular there exists $\delta_1 > 0$ with $\delta_1 < \delta$ such that $x \in (a, a + \delta_1)$ implies

$$\left| \frac{f(x)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)} \right| < \varepsilon.$$

Further, the triangle inequality gives

$$\left| \frac{f(x)}{g(x)} - A \right| < 2\varepsilon.$$

This proves (b).
The case $x \to +\infty$ can be reduced to the limit process $y \to 0 + 0$ using the substitution $y = 1/x$. ∎

L'Hospital's rule aplies also applies in the cases $A = +\infty$ and $A = -\infty$.

Example 4.5 (a) $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = \lim\limits_{x\to 0} \dfrac{\cos x}{1} = 1.$

(b) $\lim\limits_{x\to 0+0} \dfrac{\sqrt{x}}{1-\cos x} = \lim\limits_{x\to 0+0} \dfrac{\frac{1}{2\sqrt{x}}}{\sin x} = \lim\limits_{x\to 0+0} \dfrac{1}{2\sqrt{x}\sin x} = +\infty.$

(c)

$$\lim_{x\to 0+0} x\log x = \lim_{x\to 0+0} \frac{\log x}{\frac{1}{x}} = \lim_{x\to 0+0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x\to 0+0} -x = 0.$$

Remark 4.4 It is easy to transform other indefinite expressions to $\dfrac{0}{0}$ or $\dfrac{\infty}{\infty}$ of l'Hospital's rule.

$$0\cdot\infty: \quad f\cdot g = \frac{f}{\frac{1}{g}}$$

$$\infty-\infty: \quad f-g = \frac{\frac{1}{g}-\frac{1}{f}}{\frac{1}{fg}};$$

$$0^0: \quad f^g = \mathrm{e}^{g\log f}.$$

Similarly, expressions of the form 1^∞ and ∞^0 can be transformed.

4.5 Taylor's Theorem

The aim of this section is to show how n times differentiable functions can be approximated by polynomials of degree n.

First consider a polynomial $p(x) = a_n x^n + \cdots + a_1 x + a_0$. We compute

$$\begin{aligned} p'(x) &= na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1, \\ p''(x) &= n(n-1)a_n x^{n-2} + (n-1)(n-2)a_{n-1}x^{n-2} + \cdots + 2a_2, \\ &\vdots \\ p^{(n)}(x) &= n!\, a_n. \end{aligned}$$

Inserting $x=0$ gives $p(0)=a_0$, $p'(0)=a_1$, $p''(0)=2a_2$, ... , $p^{(n)}(0)=n!a_n$. Hence,

$$p(x) = p(0) + \frac{p''(0)}{1!}x + \frac{p''(0)}{2!}x^2 + \cdots + \frac{p^{(n)}(0)}{n!}x^n. \tag{4.15}$$

Now, fix $a\in\mathbb{R}$ and let $q(x)=p(x+a)$. Since $q^{(k)}(0)=p^{(k)}(a)$, (4.15) gives

$$\begin{aligned} p(x+a) = q(x) &= \sum_{k=0}^{n} \frac{q^{(k)}(0)}{k!}\, x^k, \\ p(x+a) &= \sum_{k=0}^{n} \frac{p^{(k)}(a)}{k!}\, x^k. \end{aligned}$$

Replacing in the above equation $x+a$ by x yields

$$p(x) = \sum_{k=0}^{n} \frac{p^{(k)}(a)}{k!}\,(x-a)^k. \tag{4.16}$$

Theorem 4.15 (Taylor's Theorem) *Suppose f is a real function on $[r,s]$, $n \in \mathbb{N}$, $f^{(n)}$ is continuous on $[r,s]$, $f^{(n+1)}(t)$ exists for all $t \in (r,s)$. Let a and x be distinct points of $[r,s]$ and define*

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k. \tag{4.17}$$

Then there exists a point ξ between x and a such that

$$f(x) = P_n(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}\,(x-a)^{n+1}. \tag{4.18}$$

For $n=0$, this is just the mean value theorem. $P_n(x)$ is called the *nth Taylor polynomial of f at $x=a$*, and the second summand of (4.18)

$$R_{n+1}(x,a) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\,(x-a)^{n+1}$$

is called the *Lagrange remainder term.*
In general, the theorem shows that f can be approximated by a polynomial of degree n, and that (4.18) allows to estimate the error, if we know the bounds of $\left| f^{(n+1)}(x) \right|$.
Proof. Consider a and x to be fixed; let M be the number defined by

$$f(x) = P_n(x) + M(x-a)^{n+1}$$

and put

$$g(t) = f(t) - P_n(t) - M(t-a)^{n+1}, \quad \text{for} \quad r \le t \le s. \tag{4.19}$$

We have to show that $(n+1)!M = f^{(n+1)}(\xi)$ for some ξ between a and x. By (4.17) and (4.19),

$$g^{(n+1)}(t) = f^{(n+1)}(t) - (n+1)!M, \quad \text{for} \quad r < t < s. \tag{4.20}$$

Hence the proof will be complete if we can show that $g^{(n+1)}(\xi) = 0$ for some ξ between a and x.
Since $P_n^{(k)}(a) = f^{(k)}(a)$ for $k = 0, 1, \dots, n$, we have

$$g(a) = g'(a) = \cdots = g^{(n)}(a) = 0.$$

Our choice of M shows that $g(x) = 0$, so that $g'(\xi_1) = 0$ for some ξ_1 between a and x, by Rolle's theorem. Since $g'(a) = 0$ we conclude similarly that $g''(\xi_2) = 0$ for some ξ_2 between a and ξ_1. After $n+1$ steps we arrive at the conclusion that $g^{(n+1)}(\xi_{n+1}) = 0$ for some ξ_{n+1} between a and ξ_n, that is, between a and x. ∎

Definition 4.5 Suppose that f is a real function defined on $[r,s]$ such that $f^{(n)}(t)$ exists for all $t \in (r,s)$ and all $n \in \mathbb{N}$. Let x and a points of $[r,s]$. Then

$$T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k \tag{4.21}$$

is called the *Taylor series* of f at a.

Remarks 4.5 (a) The radius r of convergence of a Taylor series can be 0.
(b) If T_f converges, it may happen that $T_f(x) \neq f(x)$. If $T_f(x)$ at a point a converges to $f(x)$ in a certain neighborhood $U_r(a)$, $r > 0$, f is called to be *analytic* at a.

Example 4.6 We give an example for (b). Define $f \colon \mathbb{R} \to \mathbb{R}$ via

$$f(x) = \begin{cases} \mathrm{e}^{-1/x^2}, & \text{if} \quad x \neq 0, \\ 0, & \text{if} \quad x = 0. \end{cases}$$

We will show that $f \in \mathrm{C}^\infty(\mathbb{R})$ with $f^{(k)}(0) = 0$. For we will prove by induction on n that there exists a polynomial p_n such that

$$f^{(n)}(x) = p_n\left(\frac{1}{x}\right) \mathrm{e}^{-1/x^2}, \quad x \neq 0$$

and $f^{(n)}(0) = 0$. For $n = 0$ the statement is clear taking $p_0(x) \equiv 1$. Suppose the statement is true for n. First, let $x \neq 0$ then

$$f^{(n+1)}(x) = \left(p_n\left(\frac{1}{x}\right) \mathrm{e}^{-1/x^2}\right)' = \left(-\frac{1}{x^2} p_n'\left(\frac{1}{x}\right) + \frac{2}{x^3} p_n\left(\frac{1}{x}\right)\right) \mathrm{e}^{-1/x^2}.$$

Choose $p_{n+1}(t) = -p_n'(t)t^2 + 2p_n(t)t^3$.
Secondly,

$$f^{(n+1)}(0) = \lim_{h\to 0} \frac{f^{(n)}(h) - f^{(n)}(0)}{h} = \lim_{h\to 0} \frac{p_n\left(\frac{1}{h}\right) \mathrm{e}^{-1/h^2}}{h} = \lim_{x\to\pm\infty} x p_n(x) \mathrm{e}^{-x^2} = 0,$$

where we used Proposition 2.5 in the last equality.
Hence $T_f \equiv 0$ at 0—the Taylor series is identically 0—and $T_f(x)$ does not converge to $f(x)$ in a neigborhood of 0.

4.5.1 Examples of Taylor Series

(a) Power series coincide with their Taylor series.

$$\mathrm{e}^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}, \quad \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad x \in (-1,1).$$

(b) $f(x) = \log(1+x)$, see Homework 13.5.

(c) $f(x) = (1+x)^\alpha$, $\alpha \in \mathbb{R}$, $a = 0$. We have

$$f^{(k)}(x) = \alpha(\alpha-1)\cdots(\alpha-k+1)(1+x)^{\alpha-k}, \quad \text{in particular} \quad f^{(k)}(0) = \alpha(\alpha-1)\cdots(\alpha-k+1).$$

Therefore,

$$(1+x)^\alpha = \sum_{k=1}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}\, x^k + R_n(x) \tag{4.22}$$

The quotient test shows that the corresponding power series converges for $|\,x\,| < 1$. Consider the Lagrange remainder term with $0 < \xi < x < 1$ and $n+1 > \alpha$. Then

$$|\,R_{n+1}(x)\,| = \left| \binom{\alpha}{n+1}(1+\xi)^{\alpha-n-1}x^{n+1} \right| \le \left| \binom{\alpha}{n+1} x^{n+1} \right| \le \left| \binom{\alpha}{n+1} \right| \longrightarrow 0$$

as $n \to \infty$. Hence,

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n, \quad 0 < x < 1. \tag{4.23}$$

(4.23) is called the *binomial series*. Its radius of convergence is $R = 1$. Looking at other forms of the remainder term gives that (4.23) holds for $-1 < x < 1$.
(d) $y = f(x) = \arctan x$. Since $y' = 1/(1+x^2)$ and $y'' = -2x/(1+x^2)^2$ we see that

$$y'(1+x^2) = 1.$$

Differentiating this n times and using Leibniz's formula, Proposition 4.6 we have

$$\sum_{k=0}^{n} (y')^{(k)}(1+x^2)^{(n-k)} \binom{n}{k} = 0.$$

$$\Longrightarrow \binom{n}{n} y^{(n+1)}\,(1+x^2) + \binom{n}{n-1} y^{(n)}\,2x + \binom{n}{n-2} y^{(n-1)}\,2 = 0;$$

$$x = 0: \quad y^{(n+1)} + n(n-1)y^{(n-1)} = 0.$$

This yields

$$y^{(n)}(0) = \begin{cases} 0, & \text{if} \quad n = 2k, \\ (-1)^k(2k)!, & \text{if} \quad n = 2k+1. \end{cases}$$

Therefore,

$$\arctan x = \sum_{k=0}^{n} \frac{(-1)^k}{2k+1}\, x^{2k+1} + R_{2n+2}(x). \tag{4.24}$$

One can prove that $-1 < x \le 1$ implies $R_{2n+2}(x) \to 0$ as $n \to \infty$. In particular, $x = 1$ gives

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - + \cdots.$$

4.6 Appendix C

Corollary 4.16 (to the mean value theorem) *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a differentiable function with*

$$f'(x) = cf(x) \quad \textit{for all } x \in \mathbb{R}, \tag{4.25}$$

where $c \in \mathbb{R}$ is a fixed number. Let $A = f(0)$. Then

$$f(x) = A\mathrm{e}^{cx} \quad \textit{for all } x \in \mathbb{R}. \tag{4.26}$$

Proof. Consider $F(x) = f(x)\mathrm{e}^{-cx}$. Using the product rule for derivatives and (4.25) we obtain

$$F'(x) = f'(x)\mathrm{e}^{-cx} + f(x)(-c)\mathrm{e}^{-cx} = (f'(x) - cf'(x))\,\mathrm{e}^{-cx} = 0.$$

By Corollary 4.11, $F(x)$ is constant. Since $F(0) = f(0) = A$, $F(x) = A$ for all $x \in \mathbb{R}$; the statement follows. ∎

The Continuity of derivatives

We have seen that there exist derivatives f' which are not continuous at some point. However, not every function is a derivative. In particular, derivatives which exist at every point of an interval have one important property: The intermediate value theorem holds. The precise statement follows.

Proposition 4.17 *Suppose f is differentiable on $[a, b]$ and suppose $f'(a) < \lambda < f'(b)$. Then there is a point $x \in (a, b)$ such that $f'(x) = \lambda$.*

Proof. Put $g(t) = f(t) - \lambda t$. Then g is differentiable and $g'(a) < 0$. Therefore, $g(t_1) < g(a)$ for some $t_1 \in (a, b)$. Similarly, $g'(b) > 0$, so that $g(t_2) < g(b)$ for some $t_2 \in (a, b)$. Hence, g attains its minimum in the *open* interval (a, b) in some point $x \in (a, b)$. By Proposition 4.7, $g'(x) = 0$. Hence, $f'(x) = \lambda$. ∎

Corollary 4.18 *If f is differentiable on $[a, b]$, then f' cannot have discontinuities of the first kind.*

Proof of Proposition 4.13. (a) Suppose first that $f'' \geq 0$ for all x. By Corollary 4.11, f' is increasing. Let $a < x < y < b$ and $\lambda \in [0, 1]$. Put $t = \lambda x + (1 - \lambda)y$. Then $x < t < y$ and by the mean value theorem there exist $\xi_1 \in (x, t)$ and $\xi_2 \in (t, y)$ such that

$$\frac{f(t) - f(x)}{t - x} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(y) - f(t)}{y - t}.$$

Since $t - x = (1 - \lambda)(y - x)$ and $y - t = \lambda(y - x)$ it follows that

$$\frac{f(t) - f(x)}{1 - \lambda} \leq \frac{f(y) - f(t)}{\lambda}$$
$$\Longrightarrow f(t) \leq \lambda f(x) + (1 - \lambda) f(y).$$

Hence, f is convex.
(b) Let $f\colon (a,b) \to \mathbb{R}$ be convex and twice differentiable. Suppose to the contrary $f''(x_0) < 0$ for some $x_0 \in (a,b)$. Let $c = f'(x_0)$; put

$$\varphi(x) = f(x) - (x - x_0)c.$$

Then $\varphi\colon (a,b) \to \mathbb{R}$ is twice differentiable with $\varphi'(x_0) = 0$ and $\varphi''(x_0) < 0$. Hence, by Proposition 4.12, φ has a local maximum in x_0. By definition, there is a $\delta > 0$ such that $U_\delta(x_0) \subset (a,b)$ and

$$\varphi(x_0 - \delta) < \varphi(x_0), \quad \varphi(x_0 + \delta) < \varphi(x_0).$$

It follows that

$$f(x_0) = \varphi(x_0) > \frac{1}{2}\left(\varphi(x_0 - \delta) + \varphi(x_0 + \delta)\right) = \frac{1}{2}\left(f(x_0 - \delta) + f(x_0 + \delta)\right).$$

This contradicts the convexity of f if we set $x = x_0 - \delta$, $y = x_0 + \delta$, and $\lambda = 1/2$. ∎