Chapter 5

Functions of one variable: Differentiability

(continued)

Theorem 4.11

If $f: U_\rho(0) \to \mathbb{K}$ is a function given by a convergent power series $f(x) = a_0 + a_1 x + \ldots + a_n x^n + \ldots$ which has the radius of convergence $\rho \in]0, \infty]$, then f is differentiable on $U_\rho(0)$ and the derivative is obtained by differentiating term by term:

$$f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1} + \ldots .$$

Proof.

By 3.24 (iii), $\sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$ has the same radius of convergence ρ as $\sum_{n=0}^{\infty} a_n x^n$. Thus $g(z) \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$ defines a function $g: U_\rho(0) \to \mathbb{K}$.

Apply this argument to $\sum_{n=1}^{\infty} n \cdot |a_n| \cdot x^{n-1}$ in place of $\sum_{n=0}^{\infty} a_n x^n$ to get that the power series $\sum_{n=2}^{\infty} (n-1) n \cdot |a_n| \cdot x^{n-2}$ still has the same radius ρ and thus defines a function

$$(*) \qquad R: U_\rho(0) \to \mathbb{K}, \quad R(x) = \sum_{n=2}^{\infty} n(n-1)|a_n| \cdot x^{n-2}.$$

We first prove the following:

Claim. For each $u \in U_\rho(0)$ and $\varepsilon > 0$ there exists $\delta > 0$ s.t.

$$|x-u| < \delta \implies |x| < \rho \wedge |f(x) - f(u) - g(u) \cdot (x-u)| < |x-u| \cdot \varepsilon.$$

Proof of claim.

Let $|u| < \rho$ and $r \in \mathbb{R}$ s.t. $|u| < r < \rho$. Let $|x| \leq r$. Then

$$f(x) - f(u) - (x-u)\cdot g(u) = \sum_{n=0}^{\infty} a_n x^n - \sum_{n=0}^{\infty} a_n u^n - (x-u)\cdot \sum_{n=1}^{\infty} n a_n \cdot u^{n-1}$$

$$= \sum_{n=1}^{\infty} a_n \left(x^n - u^n - n(x-u)\cdot u^{n-1}\right).$$

Thus,

$$|f(x) - f(u) - (x-u)\cdot g(u)| \leq \sum_{n=1}^{\infty} |a_n| \cdot |x^n - u^n - n(x-u)\cdot u^{n-1}|.$$

Now

$$|x^n - u^n - n(x-u)\cdot u^{n-1}| = |(x-u)(x^{n-1} + x^{n-2}\cdot u + \dots + u^{n-1}) - n(x-u)\cdot u^{n-1}|$$

$$\leq |x-u| \cdot \sum_{p+l=n-1} |x^p - u^p| \cdot |u|^l.$$

As $|u|, |x| \leq r$ implies that

$$|x^p - u^p| \leq |x-u| \cdot \sum_{j=0}^{p-1} |x|^{p-1-j} \cdot |u|^j \leq |x-u| \cdot p \cdot r^{p-1}, \quad p \geq 1,$$

we get that

$$|x-u| \cdot \sum_{p+l=n-1} |x^p - u^p| \cdot |u|^l \leq |x-u|^2 \cdot r^{n-2} \cdot \sum_{p=0}^{n-1} p$$

$$= |x-u|^2 \cdot r^{n-2} \cdot \frac{n(n-1)}{2}.$$

Therefore

$$|f(x) - f(u) - (x-u)\cdot g(u)| \leq \sum_{n=1}^{\infty} |a_n| \cdot |x^n - u^n - n(x-u)\cdot u^{n-1}|$$

$$\leq \frac{1}{2} \cdot |x-u|^2 \cdot \sum_{n=1}^{\infty} n(n-1) \cdot |a_n| \cdot r^{n-2}$$

$$= \frac{1}{2} |x-u|^2 \cdot \sum_{n=2}^{\infty} n(n-1) \cdot |a_n| \cdot r^{n-2}$$

$$= |x-u| \cdot |x-u| \cdot \frac{1}{2} R(r), \text{ with } R \text{ as in } (*).$$

For $\varepsilon > 0$ choose $\delta = \min\left\{ \frac{2\varepsilon}{R(r)+1}, r - |u| \right\}$. Then $|x-u| < \delta$ implies $|x| < r < \rho$ and $|f(x) - f(u) - (x-u)\cdot g(u)| < |x-u|\cdot\varepsilon$.

Thus, the claim is proved.

The proof of the theorem follows now easily:

Fix $u \in U_\rho(0)$ and define $\sigma : U_\rho(0) \to \mathbb{K}$ by

$$\sigma(x) = \begin{cases} |x-u|^{-1}\cdot\left(f(x) - f(u) - g(u)\cdot(x-u)\right) & \text{for } x \neq u \\ 0 & \text{for } x = u. \end{cases}$$

By the claim we have

$$(\forall \varepsilon > 0)(\exists \delta > 0)\left(|x-u| < \delta \Rightarrow |\sigma(x)| < \varepsilon\right).$$

Hence σ is continuous at u, $\sigma(u) = 0$ and

$$f(x) = f(u) + g(u)\cdot(x-u) + |x-u|\cdot\sigma(x).$$

So $g(u) = f'(u)$.

$\square$

The preceding theorem says that we are allowed to differentiate convergent power series term by term. For $\mathbb{K} = \mathbb{C}$, we may formulate this matter as follows: Any function defined by a convergent power series is holomorphic on its open disc of convergence. If the series has radius $\rho = \infty$, then the function $f : \mathbb{C} \to \mathbb{C}$ which it defines is entire.

The function $f': U_\rho(0) \to \mathbb{K}$ given by

$$f'(x) = \sum_{n=1}^{\infty} n a_n \cdot x^{n-1} = \sum_{n=0}^{\infty} (n+1) \cdot a_{n+1} x^n$$

is again differentiable according to the preceding theorem.

By induction we obtain the following consequence regarding the k-th derivative $f^{(k)}$ of f ($f^{(0)} = f$ and $f^{(k+1)} = (f^{(k)})'$).

Corollary 4.12

Assume that the function $f: U_\rho(0) \to \mathbb{K}$ satisfies the hypotheses of Theorem 4.11. Then all successive derivates $f^{(k)}: U_\rho(0) \to \mathbb{K}$ exist and

$$f^{(k)}(x) = k!\, a_k + \dots + k! \binom{n}{k} a_n x^{n-k} + \dots$$
$$= k! \sum_{n=k}^{\infty} \binom{n}{k} a_n x^{n-k} = k! \sum_{n=0}^{\infty} \binom{n+k}{k} a_{n+k} x^n.$$

Taylor's Theorem (continued)

Definition.

A function $f: I \to \mathbb{R}$ on an interval I which is arbitrarily often (resp. n-times continuously) differentiable is called a C^∞-function (resp. C^n-function). C^∞-functions are also called smooth functions.

<u>Definition</u>

Let $f: I \to \mathbb{R}$ be a smooth function on the interval I and $a \in I$.

Then

$$T^{\infty}_{f,a}(x) \overset{\text{def}}{=\!=} \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} \cdot (x-a)^k$$

is the <u>Taylor series</u> of f at a.

$$T^{n}_{f,a}(x) \overset{\text{def}}{=\!=} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} (x-a)^k$$

is the n-th partial sum or n-th <u>Taylor polynomial.</u>

<u>Proposition</u>

Let $I \subseteq \mathbb{R}$ be an interval, $a \in I$ and $f: I \to \mathbb{R}$ a function given by a power series $f(x) = \sum_{k=0}^{\infty} c_k (x-a)^k$ for all $x \in I$.

Then f is smooth and the Taylor series of f at a coincides with the power series and hence converges to f.

<u>Proof.</u>

By Corollary 4.12, f is smooth and $f^{(k)}(x) = k! \sum_{n=k}^{\infty} \binom{n}{k} a_n x^{n-k} =$

$$= \sum_{n=k}^{\infty} n(n-1)\cdot\ldots\cdot(n-k+1)\, c_n \cdot (x-a)^{n-k}.$$

In particular, $f^{(k)}(a) = k!\, c_k$, i.e. $c_k = \frac{1}{k!} f^{(k)}(a)$. ∎

<u>Warning.</u> As the following example shows, even if the Taylor series $T^{\infty}_{f,a}$ of a smooth function f converges, it needs not converge towards f.

Example 4.64

Define a function $f: \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{for } x \le 0 \\ e^{-\frac{1}{x}} & \text{for } x > 0. \end{cases}$$

Then f is smooth and $f^{(n)}(x) \overset{(*)}{=\!=} \begin{cases} 0 & \text{for } x \le 0 \\ P_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}} & \text{for } x > 0, \end{cases}$

for a suitable polynomial P_n of degree $\le 2n$.

The claim trivially holds for all $x < 0$ and all n.

We prove by induction on n that $(*)$ holds for all $x \geq 0$ and all $n \geq 0$.

The case $n = 0$ is obvious.

Let $x > 0$. Then $(*)$ is true for a whole interval neighborhood of x

and

$$f^{(n+1)}(x) = \left(f^{(n)}(x)\right)' = \left(P_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}}\right)' = P_n'\left(\frac{1}{x}\right) \cdot \left(-\frac{1}{x^2}\right) e^{-\frac{1}{x}} +$$

$$+ P_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}} \cdot \frac{1}{x^2} \overset{\text{def}}{=\!=} P_{n+1}\left(\frac{1}{x}\right) e^{-\frac{1}{x}}, \text{ where}$$

$P_{n+1}(y) \overset{\text{def}}{=\!=} P_n'(y) \cdot (-y)^2 + P_n(y) \cdot y^2$, so P_{n+1} is a polynomial of degree $\le 2(n+1)$.

It remains to prove that $f^{(n+1)}(0) = 0$. This follows from the fact that the left derivative of $f^{(n)}$ at 0 and the right derivative of $f^{(n)}$ at 0 equal 0:

$$D_0^- f^{(n)} = \lim_{\substack{x \to 0 \\ x < 0}} \frac{f^{(n)}(x) - f^{(n)}(0)}{x - 0} = 0, \text{ and } D_0^+ f^{(n)} = \lim_{\substack{x \to 0 \\ x > 0}} \frac{f^{(n)}(x) - f^{(n)}(0)}{x - 0} =$$

$$= \lim_{\substack{x \to 0 \\ x > 0}} \frac{1}{x} P_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}} = \lim_{y \to \infty} \frac{y P_n(y)}{e^y} = 0.$$

Now $(*)$ implies that $T^{\infty}_{f,0}(x)$ is the constant 0 function, but $f(x) > 0$ for all $x > 0$.

$\square$

Monotonicity of differentiable functions

Theorem 4.33

Let $f: I \longrightarrow \mathbb{R}$ be a differentiable function on an interval I.

Then

(i) f is isotone $\iff f' \geq 0$ on I;

(ii) f is antitone $\iff f' \leq 0$ on I;

(iii) f is constant $\iff f' = 0$ on I.

Proof

(i) „$\Rightarrow$" Let $a \in I$. Then, since f is isotone, for any $x \in I$ we have that $\dfrac{f(x) - f(a)}{x - a} \geq 0$. It follows that $f'(a) = \lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a} \geq 0$.

„$\Leftarrow$" Let $a, b \in I$ be such that $a < b$. Since f is differentiable on $[a, b]$, we can apply the Mean Value theorem to get a point $u \in I$ s.t. $u \in]a, b[$ and $\dfrac{f(b) - f(a)}{b - a} = f'(u) \geq 0$. Since $b - a > 0$, we get that $f(b) - f(a) \geq 0$, i.e. $f(b) \geq f(a)$. Hence, f is isotone.

(ii) f is antitone $\iff -f$ is isotone $\overset{(i)}{\iff} (-f)' \geq 0$ on $I \iff -f' \geq 0$ on I
$\iff f' \leq 0$ on I.

(iii) is an immediate consequence of (i) and (ii).

$\square$

Theorem 4.45

Let $f: I \to \mathbb{R}$ be a differentiable function on an interval I.

The following statements are equivalent:

(i) f is strictly isotone

(ii) $f' \geq 0$ on I and for each pair of numbers $a < b$ in I there exists a number $w \in]a,b[$ with $f'(w) > 0$.

Proof

(i) $\Rightarrow$ (ii) Assume that f is strictly isotone. Then $f' \geq 0$ on I by Theorem 4.33. Now let $a < b$ in I. Applying the Mean Value Theorem, we get a $w \in]a,b[$ s.t. $f'(w) = \dfrac{f(b) - f(a)}{b-a} > 0$, since $f(b) > f(a)$.

(ii) $\Rightarrow$ (i) Since $f' \geq 0$ on I, by Theorem 4.33, f is isotone. Let $a < b$ in I. We have to prove that $f(a) < f(b)$. Applying (ii), there is a $w \in]a,b[$ with $f'(w) > 0$. By Theorem 4.25 there is at least one $w' \in]a,b[$ s.t. $f(w) < f(w')$. It follows that

$$f(a) \leq f(w) < f(w') \leq f(b),$$ and hence $f(a) < f(b)$.

$\square$

Remark

A similar statement holds for strictly antitone functions.

Definition 4.22

Let $X \subseteq \mathbb{K}$ and $f: X \longrightarrow \mathbb{K}$ be a function. A function $F: X \to \mathbb{K}$ is called an antiderivative or a primitive of f if F is differentiable and $F' = f$.

We draw an important conclusion from Theorem 4.33 which clarifies the issue to which extent the derivative of a differentiable function determines this function. We cannot expect uniqueness, because all constant functions have the same derivative, namely the constant function with value zero.

Theorem 4.41

Let $f, g: I \to \mathbb{R}$ be two differentiable functions on an interval I. Then the following statements are equivalent:

(i) $f' = g'$ on I;

(ii) $f - g$ is a constant function;

(iii) there is a number $c \in \mathbb{K}$ s.t. for all $x \in I$ we have $f(x) = g(x) + c$.

Proof

(ii) $\Leftrightarrow$ (iii) Obviously.

(i) $\Leftrightarrow$ (ii) Applying Theorem 4.33, we get

$$f' = g' \Leftrightarrow f' - g' = 0 \Leftrightarrow (f-g)' = 0 \Leftrightarrow f - g \text{ is a constant function.}$$

$\square$

Theorem 4.53 (The Generalized Mean Value Theorem)

Let $a<b$ and $f,g : [a,b] \to \mathbb{R}$ be continuous functions which are differentiable on $]a,b[$. Then there exists a $u \in \mathbb{R}$ with $a<u<b$ and

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

If $g(a) \neq g(b)$ and $g'(x) \neq 0$ on $]a,b[$, then

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

Proof

Consider the function $F : [a,b] \to \mathbb{R}$ defined by

$$F(x) = (g(b)-g(a))\, f(x) - (f(b)-f(a))\, g(x).$$

F is continuous and differentiable on $]a,b[$, and, furthermore,

$$F(a) = (g(b)-g(a))\, f(a) - (f(b)-f(a))\, g(a) = g(b) f(a) - f(b) g(a) =$$
$$= (g(b)-g(a))\, f(b) - (f(b)-f(a))\, g(b) = F(b).$$

Hence, by Rolle's Theorem 4.28[1] there exists a $u \in \mathbb{R}$ with $a<u<b$ s.t. $0 = F'(u) = (g(b)-g(a))\, f'(u) - (f(b)-f(a))\, g'(u).$ $\square$

Remark

The ordinary Mean Value theorem is the special case with $g(x) \overset{\text{def}}{=} x$.

1) Recall Rolle's Theorem 4.28: If $f : [a,b] \to \mathbb{R}$ $(a<b)$ is a continuous function which is differentiable on $]a,b[$, then $f(a)=f(b)$ implies the existence of a number $u \in]a,b[$ s.t. $f'(u)=0$.

Definition

Let $X \subseteq \mathbb{R}$, Y be a metric space and consider an accumulation point a of $X \cap]a, \infty[$. Furthermore, let $f: X \to Y$ be a function. We say that an element $y \in Y$ is the limit of f as x approaches a from above (or from the right), written

$$y = \lim_{\substack{x \to a \\ x > a}} f(x)$$

if

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in X)\left(a < x < a + \delta \implies d_Y(f(x), y) < \varepsilon \right).$$

Remark

In an analogous way we define

$$y = \lim_{\substack{x \to a \\ x < a}} f(x),$$

the limit of f as x approaches a from below (or from the left).

Proposition 4.54 (Rule of Bernoulli and de l'Hôpital)

Let $f, g: [a, b] \to \mathbb{R}$ $(a < b)$ be continuous functions which are differentiable on $]a, b[$. Assume further that

(i) $(\exists \delta > 0)(\forall t \in]a, a+\delta[)\left(g(t) \neq 0 \wedge g'(t) \neq 0 \right)$;

(ii) the limit $\ell = \lim\limits_{\substack{x \to a \\ x > a}} \dfrac{f(x)}{g(x)}$ exists;

(iii) $f(a) = g(a) = 0$.

Then the limit $\lim\limits_{\substack{x \to a \\ x > a}} \dfrac{f(x)}{g(x)}$ exists as well and coincides with ℓ.

Proof

Let $x \in]a, a+\delta[$. Because of (i) we can apply the Generalized Mean Value Theorem 4.53 and obtain a number $u(x) \in]a, x[$ s.t.

$$\frac{f'(u(x))}{g'(u(x))} = \frac{f(x)-f(a)}{g(x)-g(a)} \overset{(iii)}{=} \frac{f(x)}{g(x)}.$$

Since $\lim\limits_{\substack{x \to a \\ x > a}} u(x) = a$, we obtain by (ii) that

$$\lim_{\substack{x\to a \\ x>a}} \frac{f(x)}{g(x)} = \lim_{\substack{x\to a \\ x>a}} \frac{f'(u(x))}{g'(u(x))} = \lim_{\substack{y\to a \\ y>a}} \frac{f'(y)}{g'(y)} = \ell.$$

□

There are numerous variations of this rule. The following is one of them.

Proposition (Rule of Bernoulli and de l'Hôpital - second form)

Let $f, g : [a,b] \to \mathbb{R}$ $(a<b)$ be continuous functions, which are differentiable on $]a,b[$. Assume further that

(i) $(\exists \delta > 0)\ (\forall t \in]b-\delta, b[)\ (g(t) \neq 0 \wedge g'(t) \neq 0)$;

(ii) the limit $\ell = \lim\limits_{\substack{x\to b \\ x<b}} \frac{f'(x)}{g'(x)}$ exists;

(iii) $f(b) = g(b) = 0$.

Then the limit $\lim\limits_{\substack{x\to b \\ x<b}} \frac{f(x)}{g(x)}$ exists as well and coincides with ℓ.

Proof.

Analogous to the proof of Proposition 4.54.

□

Convexity of functions

In the preceding chapter we discussed the convexity of a function. The question is now whether the tools of differential calculus allow us a new approach to this property. This is indeed the case since we reduced convexity to monotonicity of suitable functions in Theorem 3.63.

Theorem 4.55

Let $f: I \to \mathbb{R}$ be a differentiable on an interval I. Then the following two statements are equivalent.

(i) f is convex (resp. strictly convex).

(ii) f' is isotone (resp. strictly isotone).

Proof

(i) $\Rightarrow$ (ii) Let $a < b$ in I. We have to show that $f'(a) \leq f'(b)$ (resp. $f'(a) < f'(b)$). Choose a point $a < u < b$. Then

$$\mu_a(u) = \frac{f(u)-f(a)}{u-a} = \frac{f(a)-f(u)}{a-u} = \mu_u(a) \underset{(\text{resp.} <)}{\leq} \mu_u(b) = \frac{f(b)-f(u)}{b-u} =$$

$$= \frac{f(u)-f(b)}{u-b} = \mu_b(u),$$ since by Theorem 3.63, μ_u is (strictly) isotone.

Since $f'(a) = \lim\limits_{\substack{t \to a \\ t > a}} \mu_a(t)$, $f'(b) = \lim\limits_{\substack{t \to b \\ t < b}} \mu_b(t)$, and μ_a and μ_b are (strictly) isotone, we get that $f'(a) \leq \mu_a(u) \underset{(\text{resp.} <)}{\leq} \mu_b(u) \leq f'(b)$, which proves (ii).

(ii) $\Rightarrow$ (i) Assume that f' is isotone (resp. strictly isotone). We use now Theorem 3.63 to prove that f is convex (resp. strictly convex). Thus, we must show that for all $a \in I$ the function μ_a is isotone (resp. strictly isotone), that is for all $a \in I$ and for all $u, v \in I \setminus \{a\}$,

$$\mu_a(u) \underset{(\text{resp. } <)}{\leq} \mu_a(v)$$

There are three cases:

(1) $u < a < v$

(2) $a < u < v$

(3) $u < v < a$.

(1) $u < a < v$.

The Mean Value Theorem yields $s \in]u, a[$, $t \in]a, v[$ s.t.

$$f'(s) = \frac{f(a) - f(u)}{a - u} = \frac{f(u) - f(a)}{u - a} = \mu_a(u),$$

$$f'(t) = \frac{f(v) - f(a)}{v - a} = \mu_a(v).$$

Since f' is isotone (resp. strictly isotone) we get that

$$\mu_a(u) = f'(s) \underset{(\text{resp. } <)}{\leq} f'(t) = \mu_a(v).$$

(2) $a < u < v$.

Applying again the Mean Value Theorem, we obtain $s \in]a, u[$, $t \in]u, v[$ s.t. $f'(s) = \dfrac{f(u) - f(a)}{u - a} = \mu_a(u)$ and $f'(t) = \dfrac{f(v) - f(u)}{v - u} = \mu_u(v)$.

Since $s<t$ and f' is isotone (resp. strictly isotone), we have that

$f'(s) \underset{(\text{resp.} <)}{\leq} f'(t)$, that is $\mu_a(u) \underset{(<)}{\leq} \mu_u(v)$.

But $\mu_a(u) \underset{(<)}{\leq} \mu_u(v) \iff \mu_a(u) \underset{(<)}{\leq} \dfrac{f(v)-f(u)}{v-u} \iff$

$$\iff \mu_a(u) \underset{(<)}{\leq} \frac{\big(\mu_a(v)\cdot(v-a)+f(a)\big)-\big(\mu_a(u)\cdot(u-a)+f(a)\big)}{v-u}$$

$$\iff \mu_a(u)\cdot(u-v) \underset{(<)}{\leq} \mu_a(v)\cdot(v-a)-\mu_a(u)\cdot(u-a) \iff \mu_a(u)(v-a) \underset{(<)}{\leq} \mu_a(v)(v-a)$$

$$\iff \mu_a(u) \underset{(<)}{\leq} \mu_a(v).$$

Thus μ_a is isotone (resp. strictly isotone).

(3) $u<v<a$.

The proof is similar with (2).

$\square$

<u>Theorem 4.56</u>

Let $f: I \to \mathbb{R}$ be twice differentiable on the interval $I \subseteq \mathbb{R}$.

Then the following statements are equivalent.

(i) f is convex (resp. strictly convex)

(ii) $f'' \geq 0$ on I (resp. $f'' \geq 0$ on I and for each pair of numbers $a<b$ in I there exists $u \in]a,b[$ with $f''(u)>0$).

<u>Proof.</u>

By Theorems 4.55 and 4.33, f is convex $\iff$ f' is isotone $\iff$ $f'' \geq 0$ on I.

Similarly, applying Theorems 4.55 and 4.45 we obtain that f is strictly convex $\iff$ f' is strictly isotone $\iff$ $f'' \geq 0$ on I and for any $a<b$ in I there exists $u \in]a,b[$ with $f''(u)>0$.

$\square$

Remark

Analogous statements hold for (strict) concavity (then with „$\le$" instead of „$\ge$" and „$<$" instead of „$>$").

As was proved in Corollary 4.27, if a function $f: I \to \mathbb{R}$ attains a local extremum in an interior point $a \in I$, then $f'(a) = 0$ (that is f is stationary at a), provided that f is differentiable at a.

The converse does not hold as the example $f(x) = x^3$, $f: \mathbb{R} \to \mathbb{R}$ shows: $f'(0) = 0$ but f has no local extremum at 0.

Thus in looking for local maxima or minima of a differentiable function, it is necessary to determine the points at which f is stationary, but this will not suffice in general.

However, we have the following criterion.

Proposition 4.57

Assume that $f: I \to \mathbb{R}$ is a function which is differentiable at the point $a \in I$ of the interval I. Then f attains a local minimum in a, provided that the following conditions are satisfied.

(i) If $a = \min I$, then $f'(a) > 0$.

(ii) If $a = \max I$, then $f'(a) < 0$.

(iii) If a is an inner point of I, then f is differentiable on a neighborhood of a, $f''(a)$ exists and both $f'(a) = 0$ and $f''(a) > 0$ hold.

Proof. (i), (ii) By Theorem 4.25.

(iii) Since f' is defined on a neighborhood U of a, f' is differentiable at a, and $f''(a) > 0$, we can apply Theorem 4.25 to f'. We get a $\delta > 0$ s.t. $]a-\delta, a+\delta[\subseteq U$ and $f'(x) < f'(a) = 0$ for $a-\delta < x < a$, while $f'(x) > f'(a) = 0$ for $a < x < a+\delta$.

We prove in the sequel that $f(x) \geq f(a)$ for all $x \in]a-\delta, a+\delta[$ and, hence, f attains a local minimum at a.

The proof is by contradiction. Assume that there would exist an $u \in]a-\delta, a[$ with $f(u) < f(a)$. Then the Mean Value Theorem yields a $v \in]u, a[$ s.t.

$$f'(v) = \frac{f(a) - f(u)}{a-u} > 0$$

in contradiction with what we have just derived.

One shows analogously that there is no $u \in]a, a+\delta[$ with $f(u) < f(a)$.

$\square$

<u>Remark</u>

Similar statements hold for local maxima with $f'(a) < 0$ in (i), $f'(a) > 0$ in (ii), respectively $f''(a) < 0$ in (iii).

If f is twice differentiable and $f'(a) = f''(a) = 0$, nothing can be concluded:

- for $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ one has $f'(0) = f''(0) = 0$ and f has NO local extremum at 0.
- for $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$ one has $f'(0) = f''(0) = 0$ and f has a global minimum at 0.

Let us consider $f: X \to \mathbb{R}$. If $f(a) = \max f(X)$, we say that f attains a (global) maximum at a. Similarly, if $f(a) = \min f(X)$, f attains a (global) minimum at a.

It must be stressed that the above methods do not tell us whether f attains a global maximum (minimum) at a point at which it takes a local maximum (minimum).

Such a global piece of information can only be gained from additional information on f. Thus, even if all local extrema are found, one has not yet solved the problem of finding the global extrema.

One particularly "nice" situation is when our function f is defined on a compact interval $[a, b] \subseteq \mathbb{R}$, and is differentiable. Then, by Theorem of the Minimum and the Maximum 3.52, f attains a global maximum and a global minimum. And, since f is differentiable, the points at which these global extrema are attained are between the endpoints a, b and the points at which f is stationary.

Example

Let $f: [-1, 1] \to \mathbb{R}$, $f(x) = x - x^3$.

f is differentiable and $f'(x) = 1 - 3x^2$. Thus, $f'(x) = 0 \iff x = \pm \frac{1}{\sqrt{3}}$.

Since $f\left(\frac{1}{\sqrt{3}}\right) = \frac{2}{3\sqrt{3}}$, $f\left(-\frac{1}{\sqrt{3}}\right) = -\frac{2}{3\sqrt{3}}$, $f(-1) = f(1) = 0$, it follows that f attains a global maximum at $\frac{1}{\sqrt{3}}$ and a global minimum at $-\frac{1}{\sqrt{3}}$.

Points of inflection

Definition 4.58

Let $I \subseteq \mathbb{R}$ be an open interval. A function $f: I \longrightarrow \mathbb{R}$ is said to have a <u>point of inflection</u> at $a \in I$ if there exists a $\delta > 0$ s.t. f is convex (resp. concave) on $]a-\delta, a[$ but convex (resp. concave) on $]a, a+\delta[$.

For example, the function $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ has a point of inflection at 0.

Remark

Assume that f is twice differentiable on the open interval I and $a \in I$.

Then

(i) if f has a point of inflection at a, then $f''(a) = 0$

(ii) if $f''(a) = 0$ and there exists $\delta > 0$ s.t.

$(f'' \leq 0$ on $]a-\delta, a[$ and $f'' \geq 0$ on $]a, a+\delta[)$

or

$(f'' \geq 0$ on $]a-\delta, a[$ and $f'' \leq 0$ on $]a, a+\delta[)$

then f has a point of inflection at a.

Thus, the vanishing of the second derivative is a necessary condition for the presence of a point of inflection, but does not suffice. Let us consider the function $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$. Then $f''(0) = 0$, but f has no points of inflection, since f is a convex function.