§6.3. Höhere Ableitungen und Taylorformel

§6.1–6.2
Im den Abschnitten werden die Begriffe von Integral (Integrierbarkeit) und Stammfunktionen eingeführt. Wir werden nun eine Anwendung, Taylorformel, besprechen.

Definition 6.3.1 Sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$, (wir betrachten immer $I \neq \emptyset$) und $f: I \to \mathbb{R}$.

Sei $n \in \mathbb{N}_0$.

(i) wir definieren den Begriff der n-maligen Differenzierbarkeit und die n-te Ableitung $f^{(n)}(x_0) = \frac{d^n f}{dx^n}(x_0)$ rekursiv, und zwar wie folgt:

*) jede beliebige Funktion f heißt 0-mal differenzierbar. Die 0-te Ableitung wird durch $f^{(0)} := f$ definiert.

*) Für $n=1$ sagen wir, dass f in $x_0 \in I$ 1-mal differenzierbar ist, falls f in x_0 differenzierbar ist.

In diesem Fall ist die erste Ableitung von f in x_0 als $f^{(1)}(x_0) := f'(x_0)$ definiert.

*) Sei $n \in \mathbb{N}$, $n \geq 2$; wir nehmen an, dass der Begriff der $(n-1)$-maligen Differenzierbarkeit und die $(n-1)$-te Ableitung schon definiert sind.

f heißt n-mal differenzierbar in x_0, falls $\exists\, \varepsilon > 0$ so dass f in jedem Punkt von $(x_0-\varepsilon, x_0+\varepsilon) \cap I$ $(n-1)$-differenzierbar ist, und die Abbildung

$$f^{(n-1)} : (x_0-\varepsilon, x_0+\varepsilon) \cap I \to \mathbb{R}$$
$$x \mapsto f^{(n-1)}(x)$$

in x_0 differenzierbar ist.

Die n-te Ableitung von f in x_0 ist dann durch

$$f^{(n)}(x_0) := \left(f^{(n-1)}\right)'(x_0)$$ definiert.

Wir schreiben auch $f' = f^{(1)}$

$f'' = f^{(2)}$, $f''' = f^{(3)}$, ...

(ii) Fall f n-mal differenzierbar in allen $x \in I$ ist, heißt f n-mal differenzierbar (in I).

Die Funktion $f^{(n)} : I \to \mathbb{R}$ heißt die n-te Ableitung von f.

f heißt unendlich oft differenzierbar, wenn f

n-mal differenzierbar $\forall n \in \mathbb{N}$ ist.

(iii) f heißt n-mal stetig differenzierbar, wenn f n-mal differenzierbar und $f^{(n)}: I \to \mathbb{R}$ stetig ist.

(iv) $\forall n \in \mathbb{N}_0$ setzen wir

$$\mathcal{C}^n(I) := \{ f: I \to \mathbb{R},\ f \text{ n-mal stetig differenzierbar}$$

$$\mathcal{C}^\infty(I) := \{ f: I \to \mathbb{R},\ f \text{ unendlich oft differenzierbar} \}$$

Beachte

$$\mathcal{C}^0(I) = \{ f: I \to \mathbb{R} \text{ stetig} \}$$

$$\mathcal{C}^\infty(I) = \bigcap_{n \in \mathbb{N}} \mathcal{C}^n(I)$$

Beispiel 6.3.2

1) $\exp, \sin, \cos: \mathbb{R} \to \mathbb{R}$ sind unendlich oft differenzierbar

$\log: (0,\infty) \to \mathbb{R}$ ist auch in $\mathcal{C}^\infty((0,\infty))$.

2) Sei $s \in \mathbb{N}$ und $f: (0,\infty) \to \mathbb{R}$, $x \mapsto x^s$

Dann $f \in \mathcal{C}^\infty((0,\infty))$ und

$$f^{(n)}(x) = n! \binom{s}{n} x^{s-n}.$$

3) Sei $P(x) = \sum_{k=0}^{m} a_k x^k$, $a_k \in \mathbb{R}$

Dann $P \in \mathcal{C}^\infty(\mathbb{R})$

$$p^{(n)}(0) = n!\, a_n \quad \text{für } 0 \leq n \leq m$$

$$p^{(n)}(0) = 0 \quad \text{für } n > m.$$

Man bekommt
$$P(x) = \sum_{k=0}^{m} \frac{p^{(k)}(0)}{k!} x^k$$

<u>Bemerkung 6.3.3</u>

Seien $f, g : I \to \mathbb{R}$ n-mal differenzierbar. ($\lambda \in \mathbb{R}$)
Dann sind $f+g$, $f-g$, λf auch n-mal differenzierbar.
Ähnlich ist die Komposition zweier n-mal differenzierbaren Funktionen auch n-mal differenzierbar.
Analog für n-mal stetig differenzierbaren Funktionen. $\Rightarrow$ $\mathcal{C}^n(I)$, $\mathcal{C}^\infty(I)$ sind $\mathbb{R}$-Vektorräume.

<u>Satz 6.3.4</u> (Taylorsche Formel)

Sei $f : I \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion und $x_0 \in I$. Dann gilt $\forall x \in I$:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2$$
$$+ \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_{n+1}(x_0),$$

wobei $R_{n+1}(x) = \frac{1}{n!} \int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)\,dt$.

Beweis Durch Induktion nach n.

Induktionsanfang: für $n=0$ lautet die behauptete Formel:

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\,dt$$

diese ist genau der Fundamentalsatz der Differential- und Integralrechnung (Satz 6.2.4).

Induktionsschritt $\underline{n-1 \to n}$. Wir nehmen die behauptete Formel für $n-1$ an, d.h. es gilt

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + R_n(x_0),$$

wobei

$$R_n(x_0) = \frac{1}{(n-1)!}\int_{x_0}^{x} (x-t)^{n-1} f^{(n)}(t)\,dt$$

$$= \frac{1}{n!}\int_{x_0}^{x} -\left[(x-t)^n\right]' f^{(n)}(t)\,dt$$

(Ableitung bezüglich t)

(nach partielle Integration, Satz 6.2.6)

$$= \frac{1}{n!}\int_{x_0}^{x} \left(f^{(n)}(t)\right)'(x-t)^n\,dt$$

$$+ \left. -\frac{f^{(n)}(t)}{n!}(x-t)^n \right|_{t=x_0}^{t=x}$$

$$= \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_{n+1}(x_0).$$

Daraus folgt die Behauptung ∎

Satz 6.3.5 (Lagrangesche Form des Restglieds)

Seien f, I, x_0 wie im Satz 6.3.4. Dann $\forall x \in I$ $\exists\, \xi$ zwischen x und x_0, so dass

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

($R_{n+1}(x)$ ist wie im Satz 6.3.4)

Beweis Man hat

$$R_{n+1}(x) = \frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(t)\,(x-t)^n\,dt$$

$$= \frac{1}{n!} f^{(n+1)}(\xi)\int_{x_0}^{x}(x-t)^n\,dt$$

für ein ξ zwischen x und x_0, nach Mittelwertsatz der Integralrechnung (Satz 6.1.13).

Es folgt

$$R_{n+1}(x) = \frac{1}{n!} f^{(n+1)}(\xi)\cdot \left.\frac{-(x-t)^{n+1}}{n+1}\right|_{x_0}^{x}$$

$$= \frac{1}{(n+1)!} f^{(n+1)}(\xi)\,(x-x_0)^{n+1}.$$

Folgerung 6.3.6 Sei $f: I \to \mathbb{R}$ eine n-mal stetig differenzierbare Funktion und $x_0 \in I$. Dann gilt

$$\forall x \in I: \quad f(x) = \sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \varphi(x)(x-x_0)^n$$

wobei $\varphi : I \to \mathbb{R}$ eine Funktion mit

$$\lim_{x \to x_0} \varphi(x) = 0.$$

Beweis.
Nach Satz 6.3.5, $\forall x \in I$ $\exists \xi_x$ zwischen x und x_0 so dass

$$R_n(x) = \frac{f^{(n)}(\xi_x)}{n!}(x-x_0)^n$$

$$= \frac{f^{(n)}(\xi_x) - f^{(n)}(x_0)}{n!}(x-x_0)^n$$

$$+ \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

da $\xi_x \to x_0$ wenn $x \to x_0$, nach Stetigkeit von $f^{(n)}$ konvergiert

$$\lim_{x \to x_0} \underbrace{\frac{f^{(n)}(\xi_x) - f^{(n)}(x_0)}{n!}}_{=: \varphi(x)} = 0$$

Daraus folgt

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \varphi(x)(x-x_0)^n$$

mit der behaupten Eigenschaft ∎

Bemerkung: 1) Nach Beispiel 6.3.2 (3) ist die Taylorformel trivial für Polynome: $R_{m+1}(x) = 0$ $\forall x \in \mathbb{R}$, wobei $m =$ Grad von P.

2) Taylorsche Formel gibt uns eine Approximation der Funktionen in $\mathcal{C}^{n+1}(I)$ durch Polynomen, sobald eine gute Abschätzung von Restglied $R_{n+1}(x)$ vorliegt.

§6.4. Uneigentliche Integrale

Seien $-\infty < a < b < \infty$, $f: [a,b] \to \mathbb{R}$ beschränkt.

Falls f integrierbar ist, haben wir im Abschnitt 6.1–6.2 $\int_a^b f(x)\,dx$ definiert.

Wir ~~fort~~setzen nun diesen Begriff fort im Fall: a oder $b \in \{-\infty, \infty\}$

oder f an einer (oder beiden) ~~Grenz~~ Integrationsgrenze nicht definiert ist.

In diesem Fall spricht man von uneigentlichen Integralen.

Fall 1 Eine Integrationsgrenze ist unendlich.

Definition 6.4.1 Sei $a \in \mathbb{R}$, $f: [a, \infty) \to \mathbb{R}$ so dass f über jedem Intervall $[a, R]$, $a < R < \infty$, (Riemann) integrierbar ist. Falls der Grenzwert
$$\lim_{R\to\infty} \int_a^R f(x)\,dx$$
existiert, so heißt das Integral $\int_a^\infty f(x)\,dx$ konvergent und man setzt
$$\int_a^\infty f(x)\,dx := \lim_{R\to\infty} \int_a^R f(x)\,dx.$$

Analog definiert man das Integral $\int_{-\infty}^{a} f(x)\,dx$ für eine Funktion $f: (-\infty, a] \to \mathbb{R}$.

Beispiel 6.4.2 Sei $s \in \mathbb{R}$, $s > 1$.

$$f: [1, \infty) \to \mathbb{R} \qquad \text{stetig}$$
$$x \mapsto \frac{1}{x^s}.$$

Es gilt
$$\int_1^R \frac{dx}{x^s} = \frac{1}{1-s} \cdot \frac{1}{x^{s-1}} \Big|_1^R$$
$$= \frac{1}{s-1}\left(1 - \frac{1}{R^{s-1}}\right)$$
$$\xrightarrow[R\to\infty]{} \frac{1}{s-1} \quad (\text{da } s>1)$$

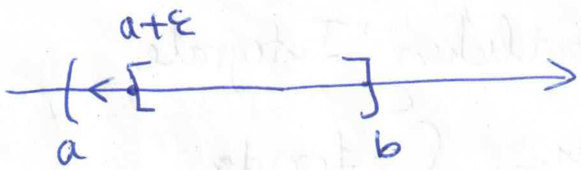
Es folgt: das Integral $\int_1^\infty \frac{dx}{x^s}$ konvergent und $\int_1^\infty \frac{dx}{x^s} = \frac{1}{s-1}$ für $s > 1$.

Andererseits gilt: $\int_1^\infty \frac{dx}{x^s}$ konvergent nicht für $s \leq 1$.

Fall 2 Der Integrand ist an einer Integrationsgrenze nicht definiert. Seien $-\infty < a < b < \infty$

Definition 6.4.3. Sei $f: (a, b] \to \mathbb{R}$ eine Funktion, die über $[a+\varepsilon, b]$ $(\forall\ 0 < \varepsilon < b-a)$ integrierbar ist. Falls

der Grenzwert
$$\lim_{\varepsilon \downarrow 0} \int_{a+\varepsilon}^{b} f(x)\,dx \text{ existiert,}$$
so heißt das Integral $\int_a^b f(x)\,dx$ konvergent und man setzt
$$\int_a^b f(x)\,dx := \lim_{\varepsilon \downarrow 0} \int_{a+\varepsilon}^{b} f(x)\,dx .$$



Man definiert analog $\int_a^b f(x)\,dx$ für $f: [a,b) \to \mathbb{R}$.

<u>Beispiel 6.4.4</u> Sei $s \in \mathbb{R}$, $s < 1$,
$$f : (0,1] \to \mathbb{R}$$
$$x \mapsto \frac{1}{x^s}$$

Man hat
$$\int_\varepsilon^1 \frac{dx}{x^s} = \frac{1}{1-s} \cdot \frac{1}{x^{s-1}} \Bigg|_\varepsilon^1$$
$$= \frac{1}{1-s}\left(1 - \varepsilon^{1-s}\right)$$
$$\xrightarrow[\varepsilon \to 0]{} \frac{1}{1-s}, \text{ da } s<1 .$$

Folglich gilt:
$$\int_0^1 \frac{dx}{x^s} \text{ konvergent und } \int_0^1 \frac{dx}{x^s} = \frac{1}{1-s} .$$

Analog zeigt man: $\int_0^1 \frac{dx}{x^s}$ konvergiert nicht für $s \geq 1$.

Fall 3: Beide Integrationsgrenzen sind kritisch.

Definition 6.4.5 Sei $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$,
$f : (a,b) \to \mathbb{R}$ eine Funktion,
die über jedem kompakten Teilintervall $[\alpha, \beta] \subset (a,b)$ integrierbar ist, und sei $c \in (a,b)$ beliebig.

Falls die beiden uneigentlichen Integrale

$$\int_a^c f(x)\,dx = \lim_{\alpha \downarrow a} \int_\alpha^c f(x)\,dx$$

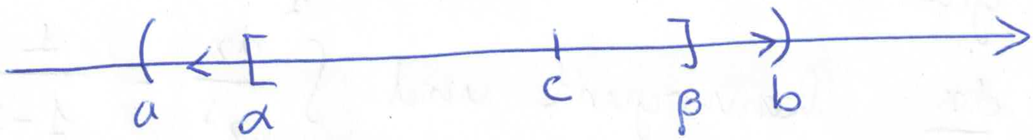
und

$$\int_c^b f(x)\,dx = \lim_{\beta \uparrow b} \int_c^\beta f(x)\,dx$$

konvergieren, heißt das Integral $\int_a^b f(x)\,dx$ konvergent und man setzt

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx .$$

Man bemerkt, diese Definition ist unabhängig von der Auswahl von $c \in (a,b)$.



Beispiel 6.4.6.

1) Das Integral $\int_0^\infty \frac{dx}{x^s}$ divergiert wegen Beispielen 6.4.2 und 6.4.1 $\forall s \in \mathbb{R}$.

2) Das Integral $\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$ konvergiert, weil:

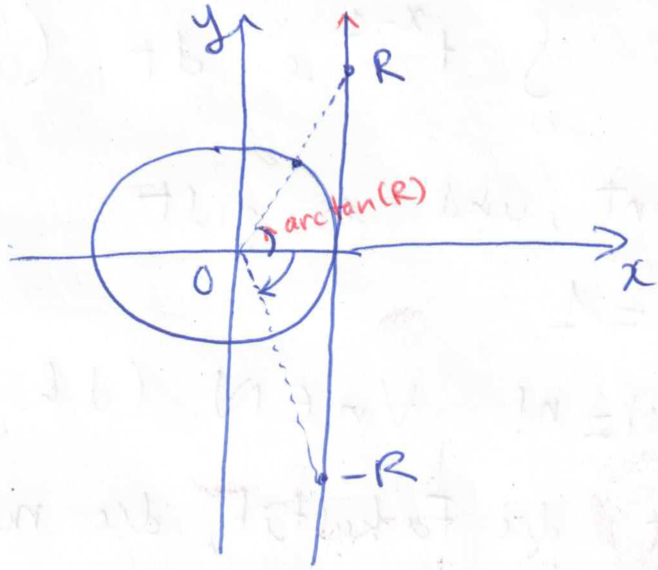
$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \lim_{R\to\infty} \int_{-R}^{0} \frac{dx}{1+x^2} + \lim_{R\to\infty} \int_{0}^{R} \frac{dx}{1+x^2}$$

$$= -\lim_{R\to\infty} \arctan(-R) + \lim_{R\to\infty} \arctan(R)$$

$$= -\left(-\frac{\pi}{2}\right) + \frac{\pi}{2} = \pi .$$

Erinnere

arctan : $\mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$



3) Man kann zeigen $\int_{0}^{\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$

(siehe [Forster, Analysis I] Seite 333)

<u>Satz 6.4.7</u> Sei $f: [1,\infty) \longrightarrow [0,\infty)$ eine monotone fallende Funktion. Dann gilt

$$\sum_{n\geq 1} f(n) \text{ konvergent} \iff \int_{1}^{\infty} f(x)\,dx \text{ konvergent.}$$

(ohne Beweis)

Beispiel: 1) da $\int_1^\infty \frac{dx}{x^s}$ konvergent für $s>1$
divergent für $s \leq 1$,

nach Satz 6.4.7, bekommt man wieder:

$\sum_{n \geq 1} \frac{1}{n^s}$ konvergiert für $s>1$
divergiert für $s \leq 1$.

2) Für $x>0$ kann man zeigen, dass

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} \, dt \quad (\text{Gamma-Funktion})$$

konvergiert, und es gilt

*) $\Gamma(1) = 1$

*) $\Gamma(n+1) = n! \quad \forall n \in \mathbb{N}$ (d.h. Die Gamma-Funktion interpoliert die Fakultät, die nur für natürliche Zahlen definiert ist)

*) $x\,\Gamma(x) = \Gamma(x+1) \quad \forall x > 0$.

7. Gleichmäßig Konvergenz

Erinnerung: Sei $g: D \to \mathbb{C}$, wobei $D \subset \mathbb{C}$.

Die Supremumsnorm von ~~g~~ g ist

$$\|g\|_D := \sup\{ |g(z)| : z \in D \} \subset [0, \infty]$$

g ist beschränkt $\iff \|g\|_D < \infty$
(Definition 4.2.2)

Eine Reihe $\sum_{n \geq 0} g_n$ von Funktionen $: D \to \mathbb{C}$

heißt normal konvergent falls $\sum_{n \geq 0} \|g_n\|_D$

konvergent (d.h. $\sum_{n=0}^{\infty} \|g_n\|_D < \infty$)

Im Satz 4.2.4, 4.2.6 wurde bewiesen: Für Funktionreihe

1) normale Konvergenz $\Rightarrow$ punktweise Konvergenz

2) $\left.\begin{array}{l}\text{normale Konvergenz} \\ g_n \text{ stetig}\end{array}\right\} \Rightarrow \sum_{n=0}^{\infty} g_n$ ist stetig.

Setze $f_n := \sum_{k=0}^{n} g_k$, bekommt man eine

Folge $(f_n)_n$ von Funktionen.

Wir untersuchen nun die Konvergenz der Folgen von Funktionen, und ihre Beziehungen mit der Ableitungen und Integralen.

Definition 7.1 Sei $D \subset \mathbb{C}$, $f_n : D \to \mathbb{C}$, $n \in \mathbb{N}_0$
$f : D \to \mathbb{C}$

wir sagen

1) Die Folge $(f_n)_n$ konvergiert punktweise gegen f, falls $\forall x \in D$ gilt $\lim_{n \to \infty} f_n(x) = f(x)$, d.h.

$$\forall x \in D, \forall \varepsilon > 0 \quad \exists N = N(x, \varepsilon) \text{ so dass}$$
$$|f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N$$

2) Die Folge $(f_n)_n$ konvergiert gleichmäßig gegen f, falls

$$\forall \varepsilon > 0 \quad \exists N = N(\varepsilon) \text{ so dass}$$
$$|f_n(x) - f(x)| < \varepsilon \quad \forall x \in D \text{ und } \forall n \geq N.$$

Der Unterschied ist also der, dass im Fall gleichmäßigen Konvergenz die Zahl N nur von ε, nicht aber von x abhängt.

Beachte

1) $(f_n)_n$ konvergiert gleichmäßig gegen f

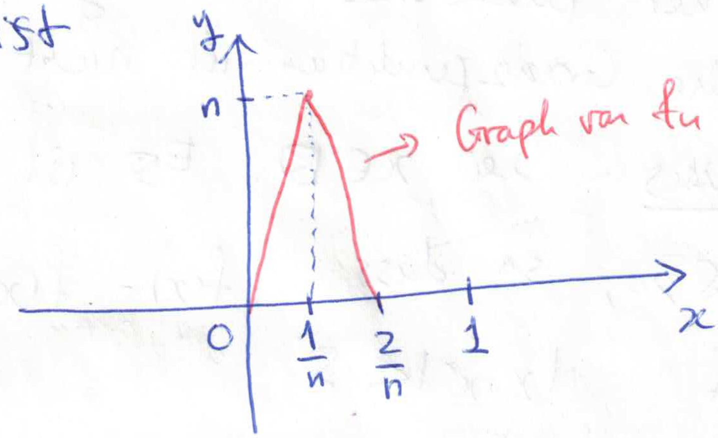
$$\Leftrightarrow \quad \lim_{n \to \infty} \|f_n - f\|_D = 0.$$

2) gleichmäßige Konvergenz $\Rightarrow$ punktweise Konvergenz aber die Umkehrung gilt jedoch nicht.

3) man kann natürlich Folgen $(f_n)_{n \geq n_0}$ $(n_0 \in \mathbb{N}_0$ oder $\mathbb{Z})$ betrachten.
Die Behandlung ist gleich wie Folgen $(f_n)_{n \geq 0}$.

Beispiel 7.2

a) Betrachte $f_n : [0,1] \to \mathbb{R}$ $(n \geq 2)$
dessen Graph ist



$f_n(x) := \max\{n - n^2|x - \frac{1}{n}|, 0\}$. Sei $f : [0,1] \to \mathbb{R}$, $x \mapsto 0$

Man kann zeigen: f_n konvergiert punktweise gegen f, aber nicht gleichmäßig:

$$\|f_n - f\|_{[0,1]} = \sup_{x \in [0,1]} |f_n(x)| = n \to \infty$$

für $n \to \infty$.

$I = [a,b]$

Fragenstellung: 1) Seien $f_n \in \mathcal{C}^0(I)$, $I \subset \mathbb{R}$
so dass f_n konvergiert punktweise oder gleichmäßig gegen f.

1) ist f stetig? Konvergiert $\int_a^b f_n(x)dx$ gegen $\int_a^b f(x)dx$?

2) Falls $f_n \in \mathcal{C}^1(I)$, ist $f \in \mathcal{C}^1(I)$?

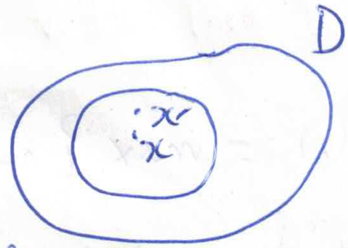
Satz 7.3 Seien $D \subset \mathbb{C}$, $f_n : D \to \mathbb{C}$ stetig

Es gelte: f_n konvergiert gleichmäßig gegen $f : D \to \mathbb{C}$

Dann ist f auch stetig.

(siehe Seite 154 im Skript für ein Beispiel von einer punktweisen konvergenten Folgen und die Grenzfunktion ist nicht stetig)

Beweis: Sei $x \in D$. Es ist zu zeigen, $\forall \varepsilon > 0$ $\exists \delta > 0$, so dass $|f(x) - f(x')| < \varepsilon$ $\forall x' \in D$ mit $|x - x'| < \delta$



Da $(f_n)_n$ gleichmäßig gegen f konvergiert, $\exists$ $n_\varepsilon \in \mathbb{N}$ so dass $|f_{n_\varepsilon}(\xi) - f(\xi)| < \frac{\varepsilon}{3}$ $\forall \xi \in D$.

Da f_{n_ε} im Punkt x stetig ist, $\exists$ $\delta > 0$ so dass $|f_{n_\varepsilon}(x) - f_{n_\varepsilon}(x')| < \frac{\varepsilon}{3}$ $\forall x' \in D$ mit $|x - x'| < \delta$. Daher für $\forall x \in D$ mit $|x - x'| < \delta$ gilt

$$
\begin{aligned}
|f(x) - f(x')| &\leq |f(x) - f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(x')| \\
&\quad + |f_{n_\varepsilon}(x') - f(x')| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon . \quad \blacksquare
\end{aligned}
$$

Satz 7.4 Sei $-\infty < a < b < \infty$

$f_n : [a,b] \to \mathbb{R}$ stetig, $n \in \mathbb{N}_0$

$f : [a,b] \to \mathbb{R}$ stetig.

Konvergiert $(f_n)_n$ gleichmäßig gegen f, so gilt

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

(Man sagt, bei gleichmäßiger Konvergenz man Integration und Limesbildung „verstauschen" darf).

Beweis Es gilt (Satz 6.1.12)

$$\left| \int_a^b f_n(x)\,dx - \int_a^b f(x)\,dx \right| \le \int_a^b |f_n(x) - f(x)|\,dx$$
$$\le \|f_n - f\|_D \int_a^b dx$$
$$= (b-a)\,\|f_n - f\|_D$$
$$\xrightarrow[n\to\infty]{} 0$$

wegen der gleichmäßigen Konvergenz ■

Beispiel: Seien $f_n : [0,1] \to \mathbb{R}$, $f : [0,1] \to \mathbb{R}$ wie im Beispiel 7.2. Erinnere:

f_n konvergiert punktweise gegen f (aber nicht gleichmäßig)

$$\int_0^1 f_n(x)\,dx = 1 \not\longrightarrow \int_0^1 f(x)\,dx = 0$$

d.h. Satz 7.4 gilt im Allgemeinen nicht für nur punktweise konvergente Folgen.

Satz 7.5 Seien $f_n : [a,b] \to \mathbb{R}$ stetig $(-\infty < a < b < \infty)$ differenzierbare Funktionen, die punktweise gegen die Funktion $f : [a,b] \to \mathbb{R}$ konvergieren. Die Folge der Ableitungen $f_n' : [a,b] \to \mathbb{R}$ konvergiere gleichmäßig. Dann ist f differenzierbar und es gilt

$$f'(x) = \lim_{n\to\infty} f_n'(x) \quad \forall x \in [a,b].$$

Beweis Sei $f^* := \lim_{n\to\infty} f_n'$. Nach Satz 7.3 ist f^* stetig. Beachte

$$f_n(x) = f_n(a) + \int_a^x f_n'(t)\,dt \quad \forall x \in [a,b]$$

$$\Big\downarrow{\scriptstyle n\to\infty} \quad \text{nach Satz 7.4}$$

$$\int_a^x f^*(t)\,dx$$

Folglich bekommt man

$$f(x) = f(a) + \int_a^x f^*(t)\,dx$$

Differentiation ergibt $f'(x) = f^*(x)$ (Satz 6.2.1) ∎

Beispiel

wenn $(f_n)_n$ nur gleichmäßig gegen f konvergent, gilt der Satz 7.5 im Allgemeinen nicht:

betrachte
$$f_n : [0,2\pi] \longrightarrow \mathbb{R}$$
$$x \longmapsto \frac{1}{n}\sin nx \quad , n \geq 1.$$

Da $\|f_n\|_{[0,2\pi]} = \frac{1}{n}$, konvergent $(f_n)_n$ gleichmäßig gegen 0. Andererseits gilt
$$f'_n(x) = \cos nx$$
und die Folge $(f'_n)_n$ konvergent punktweise nicht gegen 0 (Aufgabe)

Bemerkung 7.6 Sei $\sum_{n\geq 0} g_n$ eine Reihe von Funktionen $: D \to \mathbb{C}$.

Man sagt, $\sum_{n\geq 0} g_n$ konvergent punktweise (bzw. gleichmäßig) gegen f, wenn die Folge der Partialsummen
$$f_n := \sum_{k=0}^{n} g_k \quad , n \geq 0$$
punktweise (bzw. gleichmäßig) gegen f konvergent.

Bemerkt für Reihen $\sum_{n\geq 0} g_n$:

Normale Konvergenz $\Rightarrow$ Gleichmäßige Konvergenz (x)

$\Rightarrow$ Punktweise Konvergenz.

Die Umkehrung gilt ~~nicht~~. aber nicht.
(Siehe Beispiel 7.3.4 im Marinescus Skript)

Erklärung für (x):

Angenommen, $\sum_{n\geq 0} g_n$ konvergiert normal gegen f :

das bedeutet : $\begin{cases} f(x) = \sum_{n=0}^{\infty} g_n(x) & \forall x \in D \\ \sum_{n=0}^{\infty} \|g_n\|_D < \infty \end{cases}$

Sei $f_n = \sum_{k=0}^{n} g_k$ Betrachte

$$\|f_n - f\|_D = \sup_{x\in D} \{ |f_n(x) - f(x)| \}$$

$$= \sup_{x\in D} \{ | \sum_{k=n+1}^{\infty} g_k(x) | \}$$

$$\leq \sum_{k=n+1}^{\infty} \|g_k\|_D \xrightarrow[n\to\infty]{} 0$$

Folglich f_n konvergiert gleichmäßig gegen f. ∎

Wir betrachten nun Anwendungen auf Potenzreihen.

Sei $\sum_{n\geq 0} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R>0$, wobei $a_n \in \mathbb{R}$, $x \in \mathbb{R}$. Erinnere

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}} \in [0,\infty]. \quad (\text{Satz } 3.4.5)$$

und die Reihe $\sum_{n\geq 0} a_n x^n$ konvergiert normal (also gleichmäßig) auf $[-r,r]$ $\forall\ 0<r<R$. (Satz 4.2.5)

<u>Lemma 7.7</u> Hat $\sum_{n\geq 0} a_n x^n$ den Konvergenzradius R, so haben die durch gliedweise Differentiation bzw. Integration entstehenden Potenzreihen

$$\sum_{n\geq 1} n a_n x^{n-1} \quad \text{und} \quad \sum_{n\geq 0} \frac{1}{n+1} a_n x^{n+1}$$

den Konvergenzradius R.

<u>Beweis</u>. Setze $b_n = n a_n$

$c_n = \frac{a_n}{n+1}$

Da $\sqrt[n]{n} \to 1$ für $n\to\infty$ (Satz 2.1.3 (iii))

folgt $\lim_{n\to\infty} \sqrt[n]{b_n} = \lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \sqrt[n]{c_n}$

Daher erhält man die Behauptung ∎

<u>Satz 7.8</u> Sei $\sum_{n\geq 0} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$, es seien $a, b \in (-R, R)$ $(a<b)$. Dann ist die Summe

$$P: (-R, R) \longrightarrow \mathbb{R}$$
$$x \longmapsto \sum_{n=0}^{\infty} a_n x^n$$

stetig und es gilt

$$\int_a^b P(x)\,dx = \int_a^b \left(\sum_{n=0}^{\infty} a_n x^n\right) dx$$
$$= \sum_{n=0}^{\infty} \int_a^b (a_n x^n)\,dx$$
$$= \sum_{n=0}^{\infty} a_n \frac{b^{n+1} - a^{n+1}}{n+1}$$

Informell darf man bei Potenzreihen die Integration gliedweise berechnen.

<u>Beweis</u> Setze $f_n(x) = \sum_{k=0}^{n} a_k x^k$

Sei $r \in (0, R)$ mit $[a,b] \subset [-r, r]$.
Wir wissen bereits, f_n konvergiert gleichmäßig gegen P auf $[-r, r]$ (daher auf $[a,b]$).
Nach Sätze 7.4, 7.3 ist P stetig und

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b P(x)\,dx$$

Da
$$\int_a^b f_n(x)\,dx = \sum_{k=0}^{n} \int_a^b (a_k x^k)\,dx$$
$$= \sum_{k=0}^{n} a_k \cdot \frac{b^{k+1} - a^{k+1}}{k+1},$$

folgt die Behauptung.

Beispiel 7.9 Erinnere
$$\frac{1}{1+t} = \frac{1}{1-(-t)} = \sum_{n=0}^{\infty} (-t)^n$$

$\forall\ t \in (-1,1)$. Durch Integration auf $[0,x]$ oder $[x,0]$ ($x \in (-1,1)$) bekommt man
$$\int_0^x \frac{dt}{1+t} = \sum_{n=0}^{\infty} \int_0^x (-t)^n\,dt$$
$$= \sum_{n=0}^{\infty} -\frac{(-t)^{n+1}}{n+1}\Big|_0^x$$
$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}$$

Andererseits gilt
$$\int_0^x \frac{dt}{1+t} = \log(1+t)\Big|_0^x = \log(1+x).$$

Daher erhält man
$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}.$$

für $x \in (-1,1)$. Die Formel gilt auch für $x=1$ aber mit zusätzlichen Begründungen.

Satz 7.10 Sei $\sum_{n\geq 0} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R>0$. Dann ist die Funktion

$$p: (-R, R) \to \mathbb{R}$$
$$x \mapsto \sum_{n=0}^{\infty} a_n x^n$$

unendlich oft differenzierbar. Es gilt nämlich

$$p'(x) = \sum_{n=1}^{\infty} n\, a_n x^{n-1} \quad \forall x \in (-R, R)$$

$$p^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)\, a_n x^{n-k}, \quad k \in \mathbb{N}_0$$

Insbesondere $\dfrac{p^{(k)}(0)}{k!} = a_k$.

(Im Beispiel 6.3.2 3) wurde den Fall von Polynomen betrachtet)

Beweis Zuerst nach Lemma 7.7 ist die Reihe $\sum_{n=1}^{\infty} n a_n x^{n-1}$ (und $\sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) a_n x^{n-k}$) konvergent und ihre Konvergenzradius ist gleich R.

Setze $f_n(x) = \sum_{k=0}^{n} a_k x^k$. Dann

$$f_n'(x) = \sum_{k=1}^{n-1} k a_k x^{k-1}$$

Die Folge $(f_n')_n$ konvergiert gleichmäßig gegen $\sum_{n=1}^{\infty} n a_n x^{n-1}$ auf $[-r, r]$ (wobei $0 < r < R$ fest).

Zudem konvergiert f_n gegen P. Daher wegen des Satzes 7.5 ist P differenzierbar und $P'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \quad \forall x \in [-r, r]$

Da r beliebig ist, gilt die Formel für $\forall x \in (-R, R)$.

Wir wenden nun diese Formel auf $\sum_{n=1}^{\infty} n a_n x^{n-1}$ statt $\sum_{n=0}^{\infty} a_n x^n$ an.

und erhalten

$$(P')'(x) = \sum_{n=2}^{\infty} n(n-1) a_n x^{n-2}$$

Durch Induktion nach k bekommt man die zweite Behauptung ∎

<u>Bemerkung 7.11</u> Nach Satz 7.10 gilt

$$P(x) = \sum_{k=0}^{\infty} \frac{P^{(k)}(0)}{k!} x^k$$

Dies veranlässt uns zu

Definition 7.12 Sei $I \subset \mathbb{R}$ Intervall, $x_0 \in I$, $f: I \to \mathbb{R}$ unendlich oft differenzierbar.

Die Potenzreihe

$$T_f(x, x_0) := \sum_{k \geq 0} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

heißt Taylorreihe zu f in x_0.

Die n-te Partialsumme von $T_f(x, x_0)$ heißt das n-te Taylorpolynom $T_{n,f}(x, x_0)$

Die Taylorsche Formel lautet

$$f(x) = T_{n,f}(x, x_0) + \underbrace{R_{n+1}(x)}_{\text{Restglied}}$$

(Satz 6.34)

Bemerkung 7.13

1) Bemerkung 7.11 besagt
$P(x) =$ ihre Taylorreihe in 0.

2) Im Allgemeinen kann die Taylorreihe divergieren.

3) Falls die Taylorreihe von f konvergent ist, konvergiert sie nicht notwendig gegen f.

4) Die Taylorreihe von f konvergiert gegen f genau dann, wenn das Restglieder R_{n+1} gegen 0 konvergiert.

Beispiel 7.14 Sei $f: \mathbb{R} \to \mathbb{R}$ mit

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

Man kann zeigen, $f \in \mathcal{C}^\infty(\mathbb{R})$

$$f^{(n)}(0) = 0 \quad \forall n \in \mathbb{N}_0$$

Daher ist die Taylorreihe trivial konvergent (und konvergiert gegen 0), aber $f(x) \neq 0$ für $x \neq 0$.

Satz 7.15 (Binomische Reihe) Sei $\alpha \in \mathbb{R}$.

Dann gilt $\forall$ $|x| < 1$

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

Bemerkung: Falls $\alpha \in \mathbb{N}$ bricht die Reihe ab da $\binom{\alpha}{n} = 0$ für $n > \alpha$, die Formel folgt aus dem binomischen Lehrsatz (Satz 1.4.12)

Wir beweisen Satz 7.15 nicht

dafür muss man das Restglied in der Taylorformel abschätzen. (siehe Blatt 11 Aufgabe 4 d, Analysis I)

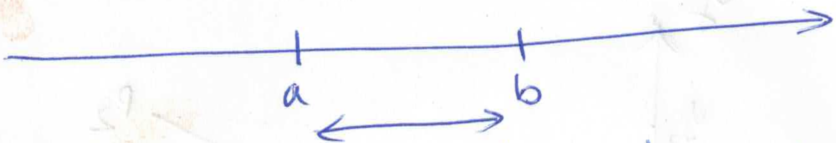
Erinnerung: $\binom{\alpha}{n} = \frac{1}{n!} \prod_{k=0}^{n-1} (\alpha - k)$

Wir verweisen an [Forster, Analysis I, Seite 405-406] für eine Anwendung der binomischen Reihen auf die kinetische Energie eines relativistischen Teilchens.

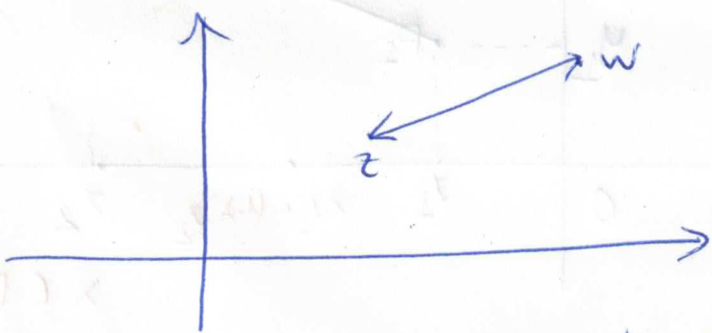
8. Metrische Räume und ihre Topologie

Bislang beschäftigen wir uns mit dem Kalkül der reellen Zahlen oder komplexen Zahlen.

In vielen Situationen genügt es, den Abstand zwischen Punkte zu wissen, z.B. Konvergenz von Folgen



Abstand zwischen a und b : $|b-a|$



$|z-w|$ der Abstand zwischen z und w.

Daher macht es Sinn, Räume mit einem „Abstand“ zwischen ihnen Punkte zu untersuchen.

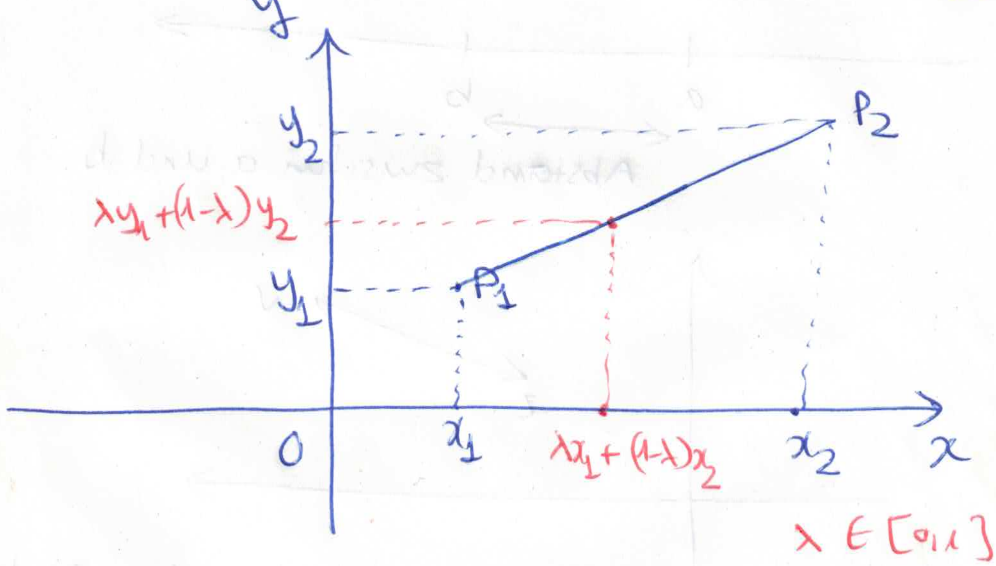
Dazu führen wir nun metrische Räume ein.

Dieser Begriff ~~vor~~ ist wesentlich für unsere Untersuchung der Funktionen von mehreren reellen Variablen.

wir beginnen zuerst mit noch einer wichtigen Eigenschaften von Funktionen auf einer reellen Variablen.

§ 8.1 Konvexität und wichtige Ungleichungen

Betrachte die Ebene $\mathbb{R}^2$ und $P_1 = (x_1, y_1)$
$P_2 = (x_2, y_2)$ in $\mathbb{R}^2$



die Strecke zwischen P_1 und P_2 wird parametrisiert durch

$$P(\lambda) = \lambda P_1 + (1-\lambda) P_2 = (\lambda x_1 + (1-\lambda) x_2, \lambda y_1 + (1-\lambda) y_2)$$

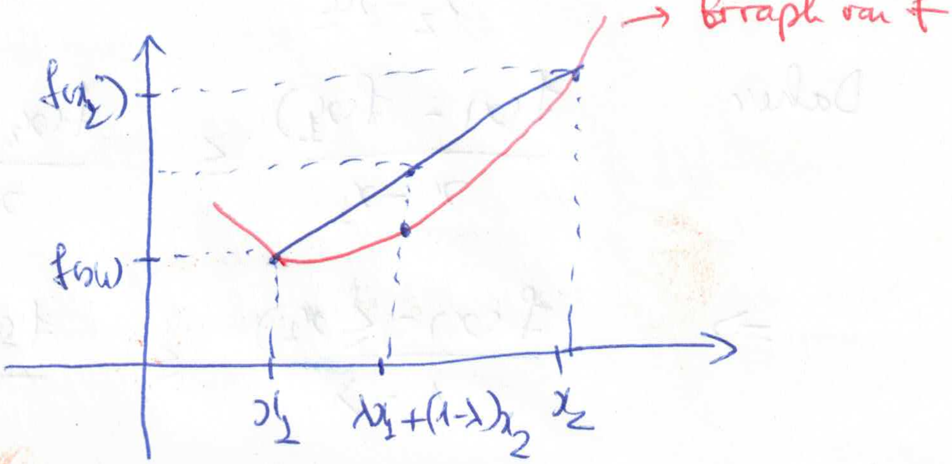
<u>Definition 8.1.1</u> Sei $I \subset \mathbb{R}$ ein Intervall.

Ein Funktion $f: I \to \mathbb{R}$ heißt <u>konvex</u>, wenn $\forall\, x_1, x_2 \in I$, $\lambda \in (0,1)$ gilt

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

f heißt <u>konkav</u> wenn $-f$ konvex ist.

Die Konvexitäts-Bedingung bedeutet (für $x_1 < x_2$), dass der Graph von f im Intervall $[x_1, x_2]$ ~~unterhal~~ unterhalb der Sekante durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$ liegt.



<u>Satz 8.1.2</u> Sei $I \subset \mathbb{R}$ ein offenes Intervall $f: I \to \mathbb{R}$ eine zweimal differenzierbare Funktion. Dann gilt:

f ist genau dann konvex, wenn $f''(x) \geq 0$ $\forall x \in I$.

<u>Beweis</u>

"$\Leftarrow$" Angenommen, $f''(x) \geq 0$ $\forall x \in I$.

Nach Folgerung 5.3.2 ist $f': I \to \mathbb{R}$ monoton wachsend. Seien $x_1, x_2 \in I$, $\lambda \in (0,1)$ $x_1 < x_2$ und $x = \lambda x_1 + (1-\lambda) x_2$. ~~Folgerung 5.3.4 besagt~~

Nach dem Mittelwertsatz (Satz 5.2.6)

$\exists \quad \xi_1 \in (x_1, x), \quad \xi_2 \in (x, x_2)$

mit $\quad \dfrac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1)$

$$\leq$$

$$\dfrac{f(x_2) - f(x)}{x_2 - x} = f'(\xi_2)$$

Daher $\quad \dfrac{f(x) - f(x_1)}{x - x_1} \leq \dfrac{f(x_2) - f(x)}{x_2 - x}$

$$\Rightarrow \quad \dfrac{f(x) - f(x_1)}{1 - \lambda} \leq \dfrac{f(x_2) - f(x)}{\lambda}$$

(weil $\quad x - x_1 = (1-\lambda)(x_2 - x_1)$
$\qquad x_2 - x = \lambda (x_2 - x_1)$)

Es folgt $\quad f(x) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$.

d.h. f konvex.

"$\Rightarrow$" Sei f nun konvex. Wir zeigen $f'' \geq 0$

Angenommen, es gelte nicht $f''(x) \geq 0 \ \forall x \in I$.

Dann $\exists \ x_0 \in I$ mit $f''(x_0) < 0$.

Setze $\quad c := f'(x_0)$, und
$$\varphi(x) := f(x) - c(x - x_0), \quad x \in I.$$

Dann ist φ zweimal differenzierbar und
$$\varphi'(x_0) = 0, \quad \varphi''(x_0) = f''(x_0) < 0.$$

Folgerung 5.3.4 angewendet auf φ impliziert, dass φ ein streng lokales Maximum in x_0 besitzt, d.h. $\exists\ h > 0$ mit $[x_0 - h, x_0 + h] \subset I$ und

$$\varphi(x) < \varphi(x_0) \quad \forall x \in [x_0 - h, x_0 + h] \setminus \{x_0\}.$$

(In der Tat besagt Folgerung 5.3.4 nur, dass φ in x_0 ein lokales Maximum besitzt. Aber aus dem Beweis sieht man klar, dass φ in x_0 eingentlich ein streng lokales Maximum hat).

Daraus folgt

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

$$= \frac{1}{2}\left(f(x_0 - h) + f(x_0 + h)\right)$$

Setzt man $x_1 := x_0 - h$, $x_2 := x_0 + h$, $\lambda := \frac{1}{2}$

So ist $x_0 = \lambda x_1 + (1-\lambda)x_2$, also

$$f(\lambda x_1 + (1-\lambda)x_2) > \lambda f(x_1) + (1-\lambda) f(x_2)$$

Dies steht aber im Widerspruch zur Konvexität von f ∎

<u>Beispiel</u>: 1) $f(x) = ax + b$, $a, b \in \mathbb{R}$, $x \in \mathbb{R}$

$$f''(x) \equiv 0 \quad \forall x \quad \Rightarrow f \text{ konvex.}$$

2) $f(x) = x^2$, $x \in \mathbb{R}$, $f''(x) = 2 > 0 \ \forall x$
$\Rightarrow f$ konvex. Man bekommt also die übliche Ungleichung $\left(\frac{x+y}{2}\right)^2 \leq \frac{x^2 + y^2}{2} \quad \forall x, y \in \mathbb{R}.$

Satz 8.1.3 Seien $p,q \in (1,\infty)$ mit $\frac{1}{p}+\frac{1}{q}=1$. Dann gilt für alle $x,y \in \mathbb{R}_+ \cup \{0\}$ die Ungleichung
$$x^{1/p}\, y^{1/q} \leq \frac{x}{p}+\frac{y}{q}.$$

Beweis Es genügt offenbar, den Satz für $x,y>0$ zu beweisen. Betrachte $\log: \mathbb{R}_+ \to \mathbb{R}$
$$\log' x = \frac{1}{x}, \quad \log'' x = -\frac{1}{x^2} < 0.$$
Daher ist log Funktion konkav. Also
$$\log\left(\frac{1}{p}x+\frac{1}{q}y\right) \geq \frac{1}{p}\log x + \frac{1}{q}\log y$$
$$\left(\lambda := \frac{1}{p}, \quad 1-\lambda = \frac{1}{q}\right)$$
Nimmt man von beiden Seiten die Exponentialfunktion so ergibt sich die Behauptung ∎

Definition 8.1.4 Sei V ein $\mathbb{K}$-Vektorraum ($\mathbb{K}=\mathbb{R}$ oder $\mathbb{C}$). Eine Abbildung $\|\cdot\|: V \to \mathbb{R}$, $v \mapsto \|v\|$ heißt Norm, falls gilt:

(N1) $\|v\| \geq 0 \ \forall v \in V$; und $\|v\|=0 \Leftrightarrow v=0$

(N2) $\|\lambda v\| = |\lambda|\,\|v\| \quad \forall \lambda \in \mathbb{K}, v \in V$.

(N3) $\|v+w\| \leq \|v\|+\|w\| \quad \forall v,w \in V$ (Dreiecksungleichung)

Beispiel: 1) $V = \mathbb{R}$, welcher ein $\mathbb{R}$-Vektorraum ist.

Der Betrag $|\cdot| : \mathbb{R} \to \mathbb{R}$
$x \mapsto |x|$ (Betrag von x)

ist eine Norm auf V.

2) Allgemeiner ist der Betrag komplexer Zahlen eine Norm auf $\mathbb{C} = \mathbb{R}^2$:

$$|\cdot| : \mathbb{R}^2 \to \mathbb{R}$$
$$(x,y) \mapsto \sqrt{|x|^2 + |y|^2}$$

oder $|\cdot| : \mathbb{C} \to \mathbb{R}$ ($\mathbb{C}$ als $\mathbb{C}$-Vektorraum aufgefasst)
$z \mapsto |z|$

Definition 8.1.5

Sei $p \geq 1$ eine reelle Zahl. Die Abbildung

$$\|\cdot\|_p : \mathbb{C}^n \to \mathbb{R}$$
$$(z_1, \dots, z_n) \mapsto (|z_1|^p + \dots + |z_n|^p)^{1/p}$$

heißt die p-Norm auf $\mathbb{C}^n$.

Falls $p = 2$, heißt 2-Norm die euklidische Norm und man schreibt $\|\cdot\|$ statt $\|\cdot\|_2$.

Wenn $n = 1$, findet man wieder den Betrag von z in $\mathbb{C}$.

Wir zeigen nun, dass $\|\cdot\|_p$ genau eine Norm im Sinn der Definition 8.1.4 ist. Dazu brauchen wir Hilfsätzen.

Satz 8.1.6 Seien $p, q \in (1, \infty)$ mit
$$\frac{1}{p} + \frac{1}{q} = 1.$$
Dann gilt für jedes Paar von Vektoren
$$x = (x_1, \dots, x_n) \in \mathbb{C}^n, \quad y = (y_1, \dots, y_n) \in \mathbb{C}^n:$$
$$\sum_{j=1}^{n} |x_j y_j| \leq \|x\|_p \, \|y\|_q$$
(Höldersche Ungleichung)

Beweis. Wir können annehmen, dass $\|x\|_p \neq 0$ und $\|y\|_q \neq 0$, da sonst der Satz trivial ist.

Setze
$$a_j := \frac{|x_j|^p}{\|x\|_p^p}, \qquad b_j := \frac{|y_j|^q}{\|y\|_q^q}$$
$\forall\ 1 \leq j \leq n$.

Dann ist
$$\sum_{j=1}^{n} a_j = 1, \quad \sum_{j=1}^{n} b_j = 1.$$

Nach Satz 8.1.3 angewendet auf a_j, b_j, bekommt man
$$\frac{|x_j y_j|}{\|x\|_p \|y\|_q} \leq \frac{|a_j|}{p} + \frac{|b_j|}{q}$$
Durch Summation über j erhält man
$$\frac{1}{\|x\|_p \|y\|_q} \left(\sum_{j=1}^{n} |x_j y_j| \right) \leq \frac{\sum_{j=1}^{n} |a_j|}{p} + \frac{\sum_{j=1}^{n} |b_j|}{q}$$

$= \frac{1}{p} + \frac{1}{q} = 1$, also die Behauptung.

Satz 8.1.7 (Minkowskische Ungleichung)

Sei $p \in [1, \infty)$. Dann gilt $\forall x, y \in \mathbb{C}^n$

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p .$$

Beweis Für $p=1$ folgt der Satz direkt aus der Dreiecksungleichung für komplexe Zahlen.

Sei nun $p > 1$ und $q \in (1, \infty)$ mit

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \left(q = \frac{p}{p-1} \right)$$

Es sei $z \in \mathbb{C}^n$ der Vektor mit den Komponenten

$$z_j := |x_j + y_j|^{p-1}, \quad j = 1, 2, \dots, n.$$

Dann ist $z_j^q = |x_j + y_j|^{q(p-1)} = |x_j + y_j|^p$.

Es folgt $\|z\|_q^q = \|x+y\|_p^p$, ~~wobei~~

$(z = (z_1, \dots, z_n))$. Nach der Hölderschen Ungleichung gilt

$$\sum_{j=1}^{n} |x_j + y_j| \cdot |z_j| \leq \sum_{j=1}^{n} |x_j z_j| + \sum_{j=1}^{n} |y_j z_j|$$

(Dreiecksungleichung für komplexen Zahlen)

$$\leq (\|x\|_p + \|y\|_p) \|z\|_q$$

also nach der Definition von z

$$\|x+y\|_p^p \le (\|x\|_p + \|y\|_p)\,\|x+y\|_p^{p/q}$$

Da $p - p/q = 1$, folgt daraus die Behauptung ■

<u>Bemerkung 8.1.8</u> Satz 8.1.7 besagt genau, dass die p-Norm $\|\cdot\|_p$ die Dreieckungleichung erfüllt. ((N3)). Die Bedingungen (N1) und (N2) sind trivial für $\|\cdot\|_p$. Daher haben wir gesehen, dass $\|\cdot\|_p$ eine Norm auf $\mathbb{C}^n$ ist.

§ 8.2. Metrische Räume

Definition 8.2.1 Sei X eine Menge. Unter einer Metrik auf X versteht man eine Abbildung

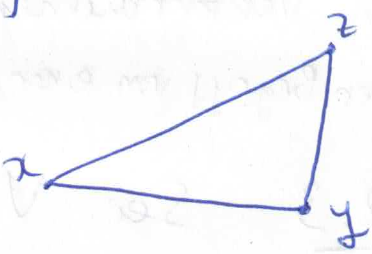
$$d: X \times X \to \mathbb{R}$$
$$(x,y) \mapsto d(x,y)$$

mit folgenden Eigenschaften:

1) $d(x,y) = 0$ genau dann wenn $x = y$;
$d(x,y) \geq 0 \quad \forall x,y \in X$

2) $\forall x,y \in X$ gilt $d(x,y) = d(y,x)$ (Symmetrie)

3) $\forall x,y,z \in X$ gilt
$d(x,z) \leq d(x,y) + d(y,z)$
(Dreiecksungleichung)



Ein metrischer Raum ist ein Paar (X,d), bestehend aus einer Menge X und einer Metrik d auf X. Man nennt $d(x,y)$ den Abstand der Punkte x und y bezüglich der Metrik d.

Beispiel 8.2.2 (1) Die Menge $\mathbb{R}$ und $\mathbb{C}$ sind metrische Räume mit dem Abstand

$$d(x,y) = |x-y| \quad \forall x,y \in \mathbb{R} \text{ (bzw. } x,y \in \mathbb{C})$$

(2) Sei (X,d) ein metrischer Raum und $A \subset X$ eine Teilmenge von X.

Die sogenannte induzierte Metrik auf A ist definiert durch

$$d_A : A \times A \longrightarrow \mathbb{R}$$
$$(x,y) \longmapsto d_A(x,y) := d(x,y)$$

Dann wird A selbst zu einem metrischen Raum.

(3) Auf jeder Menge X kann man eine triviale Metrik einführen durch die Definition

$$d(x,y) := \begin{cases} 0 & \text{für } x = y \\ 1 & \text{für } x \neq y. \end{cases}$$

Wichtigste Beispiele metrischer Räume entstehen aus normierten Vektorräumen, welche wir nun definieren. (Der Begriff von einer Norm wird in der Definition 8.1.4 eingeführt)

Definition 8.2.3 Sei V ein Vektorraum über dem Körper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Unter einer Norm auf V versteht man eine Abbildung

$$\|\cdot\| : V \longrightarrow \mathbb{R}$$
$$x \longmapsto \|x\|$$

mit folgenden Eigenschaften:

(1) $\|x\| = 0 \Leftrightarrow x = 0$; $\quad \|x\| \geq 0 \quad \forall x \in V$.

(2) $\|\lambda x\| = |\lambda| \, \|x\| \quad \forall \lambda \in \mathbb{K}, x \in V$.

(3) $\|x+y\| \leq \|x\| + \|y\| \quad \forall x,y \in V$.

Ein normierter Vektorraum ist ein Paar $(V, \|\cdot\|)$, bestehend aus einem Vektorraum V und einer Norm $\|\cdot\|$ auf V. Man schreibt manchmal nur V statt $(V, \|\cdot\|)$, wenn $\|\cdot\|$ klar aus dem Kontext ist.

Satz 8.2.4 Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann wird durch

$$d(x,y) := \|x-y\| \quad \text{für } x,y \in V$$

eine Metrik auf V definiert.

Beweis offensichtlich. ☒

Für einen normierten Raum $(V, \|\cdot\|)$ betrachten wir immer die induzierte Metrik im Satz 8.2.4. Daher ist $(V, \|\cdot\|)$ eine metrischer Raum.

Beispiel 8.2.5

(1) Sei V ein $\mathbb{K}$-Vektorraum ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$).

Erinnere (aus Lineare Algebra I):

eine Abbildung
$$V \times V \to \mathbb{K}$$
$$(x,y) \mapsto \langle x,y \rangle$$
heißt Skalarprodukt falls gilt:

(i) Die Abbildung $\langle \cdot, y \rangle : V \to \mathbb{K}$ ist $\mathbb{K}$-linear $\forall y \in V$

(ii) $\langle x,y \rangle = \overline{\langle y,x \rangle} \quad \forall x,y \in V$

(iii) $\langle x,x\rangle > 0 \quad \forall x \in V\setminus\{0\}$.

Sei $\langle \cdot,\cdot\rangle$ ein Skalarprodukt auf V.

Setze $\|x\| := \sqrt{\langle x,x\rangle} \quad \forall x \in V$.

Dann ist $\|\cdot\|$ eine Norm auf V und

es gilt $|\langle x,y\rangle| \leq \|x\|\cdot\|y\|$ (*)

$\forall x,y \in V$ (Schwarzsche Ungleichung)

Beweis: Ist $y=0$ so ist (*) klar.

Sei $y\neq 0$ also $\|y\|\neq 0$. Dann gilt

$$\|x\|^2\cdot\|y\|^2 - |\langle x,y\rangle|^2$$

$$= \left\langle x - \frac{\langle x,y\rangle}{\|y\|^2}y,\ x - \frac{\langle x,y\rangle}{\|y\|^2}y \right\rangle \|y\|^2$$

≥ 0.

Die Dreieckungleichung für $\|\cdot\|$ folgt aus Schwarzchen Ungleichung. ∎

(2) Sei $V = \mathbb{R}^n$. Betrachte das kanonische Produkt

$$\langle x,y\rangle := x_1y_1 + \dots + x_ny_n.$$

$\forall\ (x_1,\dots,x_n),\ (y_1,\dots,y_n) \in \mathbb{R}^n$. Man nennt

$$\|x\| := \sqrt{\langle x,x\rangle} = \sqrt{x_1^2+\dots+x_n^2}$$

die euklidische Norm von x (siehe auch Definition 8.1.5)

Der daraus abgeleitete Abstand ist

$$d(x,y) = \|x-y\| = \sqrt{(x_1-y_1)^2+\dots+(x_n-y_n)^2}$$

Im Folgenden verwenden wir immer diesen euklidischen Abstand auf $\mathbb{R}^n$ (und auf Teilmengen von $\mathbb{R}^n$ mit der induzierten Metrik; siehe Beispiel 8.2.2(2))

(3) Sei
$$|\cdot| : \mathbb{R}^n \to \mathbb{R}$$
$$x \mapsto |x| := \max\{|x_1|, \ldots, |x_n|\}$$
wobei $x = (x_1, x_2, \ldots, x_n)$. Diese Abbildung ist eine Norm auf $\mathbb{R}^n$, welche Maximum-Norm heißt.

Es gilt
$$|x| \leq \|x\| \leq \sqrt{n}\, |x|.$$

(4) Nach Bemerkung 8.1.8 ist $\|\cdot\|_p$ (die p-Norm) eine Norm auf $\mathbb{R}^n$. (und auch $\mathbb{C}^n$).

(5) Sei X eine beliebige Menge und $\mathcal{B}(X)$ der Vektorraum aller beschränkten ~~reell~~ reellwertigen Funktionen auf X, d.h. aller Funktionen $f: X \to \mathbb{R}$.

Setze
$$\|f\|_X := \sup\{|f(x)| : x \in X\} < \infty$$
Dann ist $\|\ \|_X$ eine Norm auf $\mathcal{B}(X)$. Diese heißt die Sup-norm (Supremumsnorm). Wir haben diese Norm in der Analysis I und Kapitel 8 für den Fall $X = I$ ein Intervall in $\mathbb{R}$ oder $X = D$ eine Teilmenge von $\mathbb{C}$ eingeführt.

(Aufgabe: Zeigen Sie $\|\cdot\|_X$ ist eine Norm)

(6) Seien $[a,b] \subset \mathbb{R}$

$$\mathcal{C}^0([a,b]) = \{ f : [a,b] \to \mathbb{R} \text{ stetig} \}$$

(siehe §6.3). Bemerkte

$\mathcal{C}^0([a,b])$ ist ein reeller Vektorraum.

Sei $p \geq 1$ eine Konstante. Dann ist

$$\|f\|_p := \left(\int_a^b |f(x)|^p \, dx \right)^{1/p}$$

eine Norm auf $\mathcal{C}^0([a,b])$. Wir nennen $\|\cdot\|_p$ die p-Norm auf $\mathcal{C}^0([a,b])$. Diese ist eine Analoge von der p-Norm auf $\mathbb{R}^n$.

(Dass $\|\cdot\|_p$ eine Norm auf $\mathcal{C}^0([a,b])$ ist folgt aus der Minkowskische Ungleichung (Satz 8.1.7) ; aber wir besprechen den Beweis nicht).

Man bemerkt auch, dass $\mathcal{C}^0([a,b])$ mit der Sup-norm $\|\cdot\|_{[a,b]}$ ein normierter Raum.

Wir führen nun einige topologische Grundbegriffe ein (d.h. offene Menge, abgeschlossene Menge, ...)

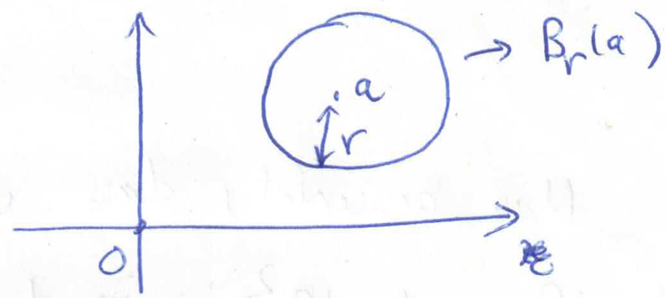
<u>Definition 8.2.6</u> Sei (X, d) ein metrischer Raum, $a \in X$ ein Punkt, $r > 0$. Dann heißt

$$B_r(a) := \{ x \in X : d(a,x) < r \}$$

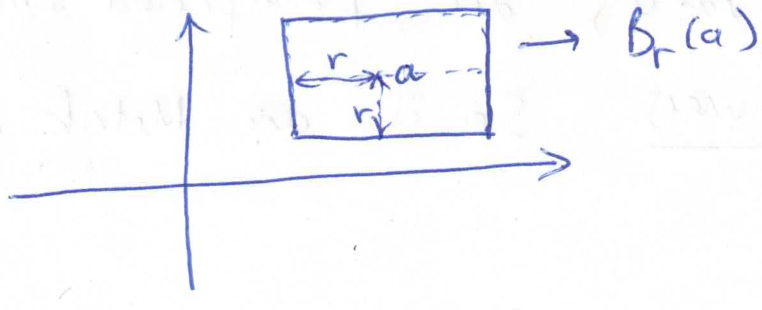
die <u>offene Kugel</u> mit Mittelpunkt a und Radius r (bzgl. der Metrik d).

<u>Beispiel</u>

(1) Betrachte $(\mathbb{R}^2, \|\cdot\|)$ (mit der euklidischen Norm) dann ist $B_r(a)$ die übliche (offene) Kreisscheibe.



(2) Betrachte $(\mathbb{R}^2, |\cdot|)$ (mit der Maximum-Norm) dann $|x-y| = \max\{|x_1-y_1|, |x_2-y_2|\}$ wobei $x = (x_1, x_2)$, $y = (y_1, y_2)$. Was ist die offene Kugel $B_r(a)$ bezüglich dieser Metrik?

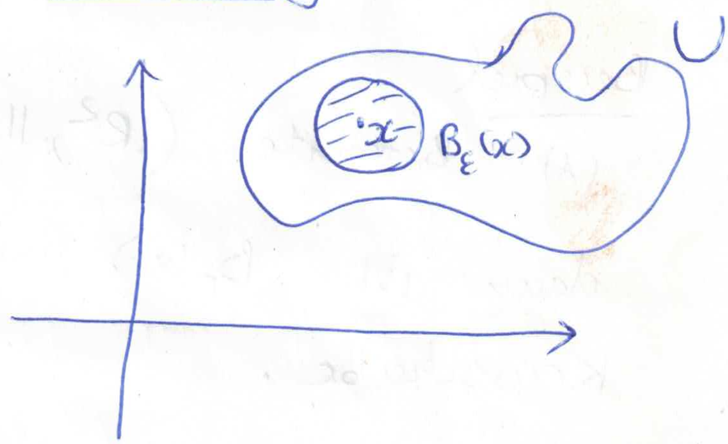


In diesem Fall ist $B_r(a)$ das Rechteck im Bild

Definition 8.2.7 Sei X ein metrischer Raum. Eine Teilmenge $U \subset X$ heißt Umgebung eines Punktes $x \in X$, falls es $\varepsilon > 0$ gibt, so dass

$$B_\varepsilon(x) \subset U.$$

Insbesondere ist $B_\varepsilon(x)$ selbst eine Umgebung von x. Man nennt $B_\varepsilon(x)$ die ε-Umgebung von x

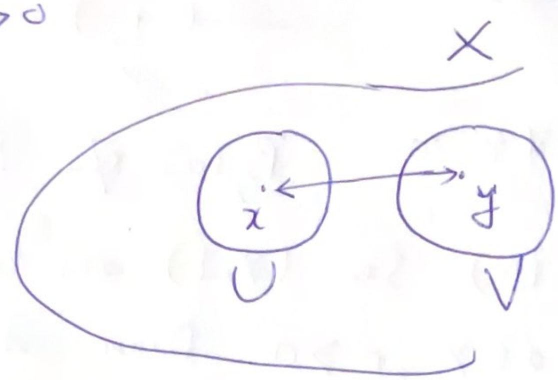


Man bemerkt, dass wir diesen Begriff für $X = \mathbb{R}$ und $\mathbb{R}^2$ in der Analysis I definiert.

Satz 8.2.8 (Hausdorffsches Trennungsaxiom) Sei X ein metrischer Raum. Dann gibt es zu je zwei Punkten $x, y \in X$ mit $x \neq y$ Umgebungen U von x und V von y, die punktfremd sind; d.h. $U \cap V = \emptyset$

Beweis Sei d die Metrik auf X.

Setze $\varepsilon := \frac{1}{3} d(x,y) > 0$



$U := B_\varepsilon(x)$

$V := B_\varepsilon(y)$.

Zeigen wir $U \cap V = \emptyset$.

Gäbe es einen Punkt $z \in U \cap V$, so würde mit der Dreiecksungleichung folgen

$$3\varepsilon = d(x,y) \leq d(x,z) + d(z,y) < \varepsilon + \varepsilon = 2\varepsilon$$

↯ also Widerspruch!.

<u>Definition 8.2.9</u> Eine Teilmenge U eines metrischen Raumes X heißt <u>offen</u>, wenn sie Umgebung jedes ihrer Punkte ist, d.h. $\forall x \in U, \exists \varepsilon > 0$ mit $B_\varepsilon(x) \subset U$.

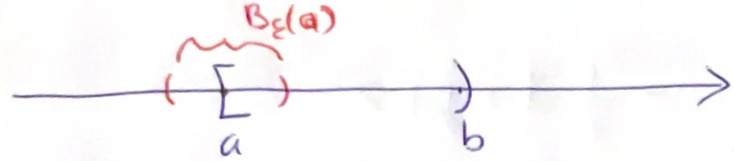
<u>Beispiel 8.2.10</u>

(i) Seien $a, b \in \mathbb{R}$ mit $a < b$. Das Intervall (a,b) ist offen in $\mathbb{R}$, denn ist $x \in (a,b)$, so gilt $B_\varepsilon(x) \subset (a,b)$ für

$$\varepsilon := \min\{|a-x|, |b-x|\}$$

Ähnlich sind (a, ∞), $(-\infty, b)$ auch offen. Andererseits sind $[a,b)$ oder $(a,b]$ $[a,b]$ nicht offen, da:

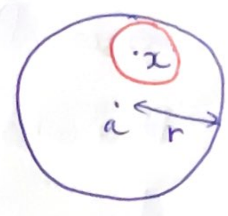
z. B.



$\forall \varepsilon > 0 \qquad B_\varepsilon(a) \not\subset [a,b).$

(ii) Sei (X,d) ein beliebiger metrischer Raum, $a \in X$, $r > 0$. Dann ist $B_r(a)$ offen.

Denn sei $x \in B_r(a)$. Dann für $\varepsilon := r - d(x,a) > 0$ gilt $B_\varepsilon(x) \subset B_r(a)$



weil: $\forall y \in B_\varepsilon(x)$: aus Dreiecksungleichung folgt

$$d(a,y) \leq d(a,x) + d(x,y)$$
$$\leq d(a,x) + \varepsilon < r.$$

(iii) Auf $\mathbb{R}^n$ bezeichnen wir

$$B_\varepsilon(a) := \{x \in \mathbb{R}^n : \|x-a\| < \varepsilon\}$$

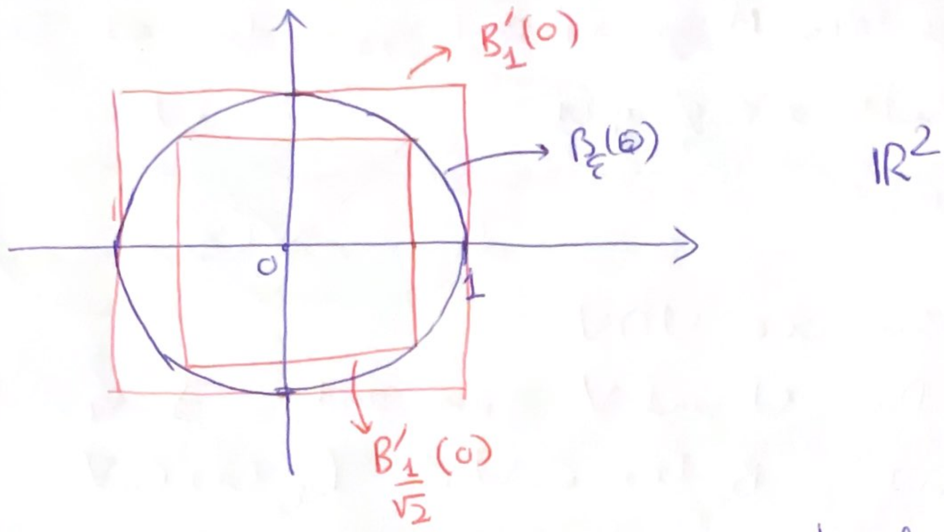
die ε-Umgebung bzgl. der euklidischen Norm und

$$B'_\varepsilon(a) := \{x \in \mathbb{R}^n : |x-a| < \varepsilon\}$$

die ε-Umgebung bzgl. der Maximum-Norm, so gilt

$$B'_{\frac{\varepsilon}{\sqrt{n}}}(a) \subset B_\varepsilon(a) \subset B'_\varepsilon(a).$$

(wegen Beispiel 8.2.5 (3))



Daraus folgt, dass jede offene Menge bzgl. der euklidischen Norm auch offen bzgl. der Maximum Norm ist und umgekehrt.

<u>Satz 8.2.11</u> Sei X ein metrischer Raum.

Dann gilt:

(i) $\emptyset$ und X sind offen.

(ii) Sind U und V offen, so ist auch der Durchschnitt $U \cap V$ offen.

(iii) Sei U_i, $i \in I$ eine Familie offener Teilmengen von X. Dann ist auch die Vereinigung

$\bigcup_{i \in I} U_i$ offen.

<u>Beweis</u>: (i) Der gesamte Raum X ist offen, da X Umgebung jedes Punktes $x \in X$ ist.

Die leere Menge ist auch offen, da es keinen Punkt $x \in \emptyset$ gibt.

(ii)

Sei $x \in U \cap V$.

Da U und V offen sind, $\exists\ \varepsilon_1 > 0,\ \varepsilon_2 > 0$ mit $B_{\varepsilon_1}(x) \subset U$, $B_{\varepsilon_2}(x) \subset V$

Für $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$ gilt dann

$$B_\varepsilon(x) \subset B_{\varepsilon_1}(x) \subset U$$

$$B_\varepsilon(x) \subset B_{\varepsilon_2}(x) \subset V$$

$$\Rightarrow \quad B_\varepsilon(x) \subset U \cap V \rightarrow U \cap V \text{ offen.}$$

(iii) Ist $x \in \bigcup_{i \in I} U_i$, so gibt es einen Index $j \in I$, so dass $x \in U_j$. Da U_j offen ist, existiert ein $\varepsilon > 0$ mit

$$B_\varepsilon(x) \subset U_j \subset \bigcup_{i \in I} U_i$$

<u>Bemerkung</u>: Aus Satz 8.2.11 (ii) gilt: ein Durchschnitt von endlich vielen offenen Mengen wieder offen ist. Dies gilt nicht für unendliche Durchschnitte, z. B.

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right) = [0,1]$$

(wobei $\bigcap_{n=1}^{\infty}$ bedeutet $\bigcap_{n \in \mathbb{N}}$)

Definition 8.2.12 Eine Teilmenge A eines metrischen Raumes X heißt abgeschlossen, wenn ihr Komplement $X \setminus A$ offen ist.

Beispiel 8.2.13

(i) $\forall\ a, b \in \mathbb{R},\ a \leq b$ ist das Intervall $[a,b] \subset \mathbb{R}$ abgeschlossen, denn sein Komplement

$$\mathbb{R} \setminus [a,b] = (-\infty, a) \cup (b, \infty) \quad \text{offen.}$$

(Siehe Beispiel 8.2.10 (i)).

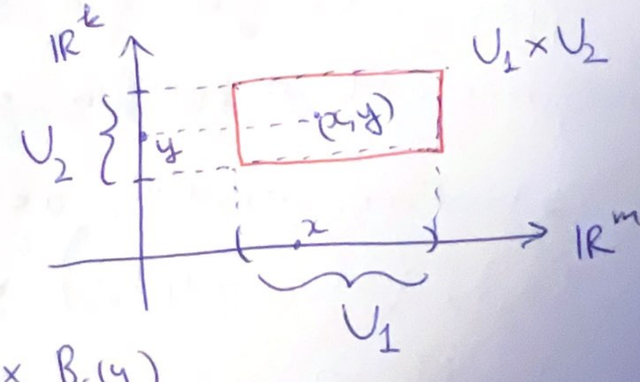
Analog sind $[a, \infty)$, $(-\infty, b]$ auch abgeschlossen, da

$$\mathbb{R} \setminus [a, \infty) = (-\infty, a)$$
$$\mathbb{R} \setminus (-\infty, b] = (b, \infty)$$

offen sind.

(ii) Sind $U_1 \subset \mathbb{R}^k$, $U_2 \subset \mathbb{R}^m$ offen, so ist $U_1 \times U_2$ offen in $\mathbb{R}^{k+m}$: weil für $(x,y) \in U_1 \times U_2$, $\exists\ \varepsilon > 0$ mit

$$\begin{cases} B_\varepsilon(x) \subset U_1 \\ B_\varepsilon(y) \subset U_2 \end{cases}$$



$$\Rightarrow \quad B_\varepsilon((x,y)) \subset B_\varepsilon(x) \times B_\varepsilon(y) \subset U_1 \times U_2.$$

(Beachte $(x',y') \in B_\varepsilon((x,y)) \iff \|x-x'\|^2 + \|y-y'\|^2 < \varepsilon^2$
Es folgt $\|x-x'\|^2 < \varepsilon^2$ und $\|y-y'\|^2 < \varepsilon^2$)

Sind $A_1 \subset \mathbb{R}^k$, $A_2 \subset \mathbb{R}^m$ abgeschlossen, so ist $A_1 \times A_2$ abgeschlossen in $\mathbb{R}^{k+m}$: denn

$$\mathbb{R}^{k+m} \setminus (A_1 \times A_2) = \Big((\underbrace{\mathbb{R}^k \setminus A_1}_{\text{offen}}) \times \mathbb{R}^m \Big) \cup \big(\mathbb{R}^k \times (\underbrace{\mathbb{R}^m \setminus A_2}_{\text{offen}}) \big)$$

Insbesondere folgt daraus, dass jeder Quader

$$Q := \{ (x_1, \dots, x_n) \in \mathbb{R}^n : a_i \leq x \leq b_i \text{ für } i = 1, 2, \dots n \}$$

$(a_i, b_i \in \mathbb{R}, \ a_i \leq b_i)$ abgeschlossen ist.

Analog ist $\tilde{Q} = \{ (x_1, \dots, x_n) \in \mathbb{R}^n : a_i < x < b_i \text{ für } 1 \leq i \leq n \}$

offen in $\mathbb{R}^n$.

(iii) $\forall$ metrischer Raum X sind $\emptyset$ und X abgeschlossen, denn ihre Komplemente X und $\emptyset$ sind offen. Es gibt also Teilmengen, die gleichzeitig offen und abgeschlossen sind.

(iv) Für $a, b \in \mathbb{R}$, $a < b$, ist das Intervall $(a, b]$ (oder $[a, b)$) weder offen noch abgeschlossen.

<u>Bemerkung 8.2.14.</u> Man sagt, die Menge der offenen Teilmengen ~~von~~ eines metrischen Raumes (X,d) bildet eine Topologie auf X. Allgemeiner für eine beliebige Menge X, falls eine Menge $\mathcal{T}$ von Teilmengen von X die folgenden Eigenschaften erfüllt:

(i) $\emptyset, X \in \mathcal{T}$

(ii) Sind $U, V \in \mathcal{T}$, so gilt $U \cap V \in \mathcal{T}$

(iii) Ist I eine beliebige Indexmenge und $U_i \in \mathcal{T}$ für $i \in I$, so folgt

$$\bigcup_{i \in I} U_i \in \mathcal{T}$$

, so heißt $\mathcal{T}$ eine <u>Topologie</u> auf X.

In diesem Fall heißt eine Menge $U \subset X$ <u>offen</u> wenn sie zu $\mathcal{T}$ gehört.

Die Bedeutung dieses allgemeineren Begriff von Topologie ist man betrachtet nur „offene ~~Ts~~ Teilmenge" ohne Bezug ~~zum~~ zu einer Metrik.

Der Begriff von Topologie wird tiefer in einer separaten Vorlesung „Topologie" untersucht

Definition 8.2.15 Sei X ein metrischer Raum und $A \subset X$. Ein Punkt $x \in X$ heißt:

(i) innerer Punkt von A, wenn A eine Umgebung von x ist

(ii) Berührpunkt von A, wenn jede Umgebung von x (mindestens) einen Punkt aus A enthält.

(iii) Randpunkt von A, wenn x Berührpunkt von A und $X \setminus A$ ist

(iv) Häufungspunkt von A, wenn jede Umgebung von x unendlich viele Punkte aus A enthält.

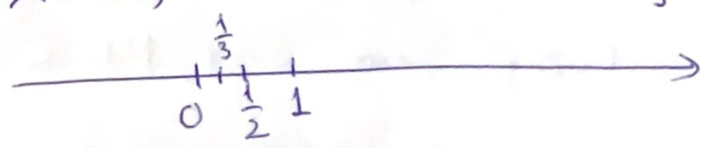
(v) isolierter Punkt von A, wenn $\exists$ Umgebung U von x mit $U \cap A = \{x\}$.

Bemerkung

$\{$Häufungspunkt$\} \cup \{$isolierter Punkte$\}$

$\subset \{$Berührpunkte$\}$

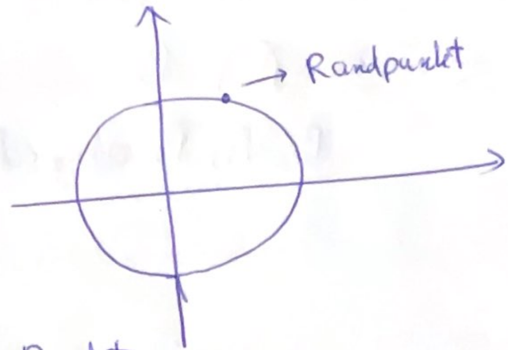
Beispiel: (i) $X = \mathbb{R}$, $A = \{ \frac{1}{n} : n \in \mathbb{N} \}$



A hat keinen inneren Punkt.
0 ist die einzige Häufungspunkt von A
$A \cup \{0\}$ ist die Menge der Berührpunkte von A.

alle Punkte von A sind isolierte Punkte.

(ii) $X = \mathbb{R}^2$, $A = B_1(0)$ (der offene Kreisscheibe)



A hat keinen isolierten Punkt

alle Punkte in A sind innere Punkte

$S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ ist die Menge der Randpunkte von A.

<u>Satz 8.2.16</u> Sei X ein metrischer Raum und $A \subset X$. Dann gilt:

(i) $A \setminus \partial A$ ist offen

(ii) $A \cup \partial A$ ist abgeschlossen.

(iii) ∂A ist abgeschlossen, wobei ∂A die Menge der Randpunkte von A bezeichnet.

<u>Beweis</u> (Aufgabe)

<u>Definition 8.2.17</u> Sei X ein metrischer Raum $A \subset X$, so heißt

$\mathring{A} := A \setminus \partial A$ das ~~Inne~~ Innere von A

$\overline{A} := A \cup \partial A$ die abgeschlossene Hülle von A.

Man kann zeigen, dass

$$\mathring{A} = \bigcup_{U \subset A,\, U \text{ offen}} U$$

$$\overline{A} = \bigcap_{B \supset A,\, B \text{ abgeschlossen}} B$$

.

§ 8.3. Grenzwerte. Stetigkeit

Wir führen nun den Begriff der Konvergenz von Punkt-Folgen und Stetigkeit von Abbildungen zwischen metrischen Räumen ein.

Dies verallgemeinert entsprechende Begriffsbildungen für Folgen in $\mathbb{R}$ oder $\mathbb{C}$ und Funktionen einer reellen Variablen.

Definition 8.3.1 Sei (X, d_X) ein metrischer Raum und $(x_k)_{k \in \mathbb{N}} \subset X$ (eine Folge von Punkten aus X).

Die Folge (x_k) heißt *konvergent* gegen den Punkt $a \in X$, in Zeichen

$$\lim_{k \to \infty} x_k = a,$$

wenn gilt: Zu ~~jeder Umgebung~~ jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, so dass

$$d_X(x_k, a) < \varepsilon \quad \forall\ k \geq N.$$

Man bemerkt, wenn $X = \mathbb{R}$ oder $(\mathbb{R}^2 \cong \mathbb{C})$, finden wir wieder den Konvergenzbegriff reeller Folgen in der Analysis I.

~~Definition 8.3~~

Satz 8.3.2 Sei $(x_k)_{k\in\mathbb{N}}$ eine Falge von Punkten im $\mathbb{R}^n$, $x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$, $k \in \mathbb{N}$.
Genau dann konvergiert die Falge $(x_k)_k$ gegen den Punkt $a = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$, wenn für $1 \leq l \leq n$ gilt

$$\lim_{k\to\infty} x_{kl} = a_l .$$

Beweis "$\Rightarrow$" Es gelte $\lim\limits_{k\to\infty} x_k = a$.
Dann gibt es zu vorgegebenem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass $\| x_k - a \| < \varepsilon$ $\forall k \geq N$. Daraus folgt für $l = 1, 2, \dots, n$,

$$| x_{kl} - a_l | \leq \| x_k - a \| < \varepsilon \quad \forall k \geq N .$$

(Erinnere: $\| x_k - a \| = \sqrt{(x_{k1} - a_1)^2 + \dots + (x_{kn} - a_n)^2}$)

Also ist $\lim\limits_{k\to\infty} x_{kl} = a_l$.

"$\Leftarrow$". angenommen, dass $\lim\limits_{k\to\infty} x_{kl} = a_l$ für $l = 1, 2, \dots n$. Zu vorgegebenem $\varepsilon > 0$ gibt es dann ein $N_l \in \mathbb{N}$ so dass

$$| x_{kl} - a_l | < \varepsilon' := \frac{\varepsilon}{\sqrt{n}} \quad \forall k \geq N_l .$$

Setze $N := \max\{N_1, N_2, \dots, N_n\}$

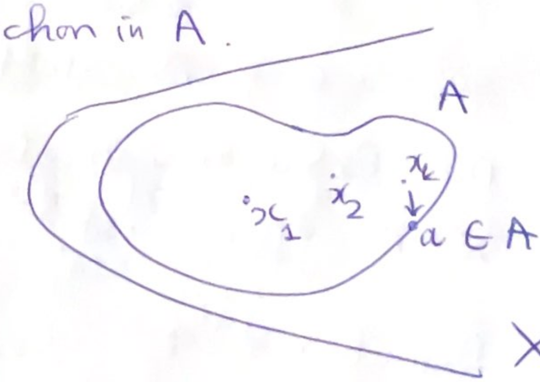
Es gilt dann
$$\|x_k - a\| = \left(\sum_{l=1}^{n} |x_{kl} - a_l|^2\right)^{1/2} < \sqrt{n}\,\varepsilon' = \varepsilon.$$

Also gilt $\lim_{k\to\infty} x_k = a$ ∎

Der Satz bedeutet: Die Konvergenz von Punktfolgen im $\mathbb{R}^n$ kann einfach auf die Konvergenz von Folgen reeller Zahlen zurückgeführt werden.

Bemerkung: Der Grenz~~punkt~~wert a einer konvergenten Folge $(x_k)_{k\in\mathbb{N}} \subset X$ ist eindeutig bestimmt.

Satz 8.3.3 Sei (X, d_X) ein metrischer Raum. Eine Teilmenge $A \subset X$ ist genau dann abgeschlossen, wenn gilt: $\forall\ (x_k)_{k\in\mathbb{N}} \subset A$ mit $\lim_{k\to\infty} x_k = a \in X$, liegt der Grenzwert a schon in A.



Beweis. "$\Rightarrow$" angenommen, A ist abgeschlossen, und $\exists\ (x_k)_{k\in\mathbb{N}} \subset A$ mit $\lim_{k\to\infty} x_k = a \in X$ aber $a \notin A$.

Da $X \setminus A$ offen ist, ~~ist~~ dann gibt es eine ε-Umgebung

$B_\varepsilon(a) \subset X \setminus A$.

Es folgt

$d(x,a) \geq \varepsilon \quad \forall\, x \in A$

Insbesondere $d(x_k, a) \geq \varepsilon$

$\forall\, k \in \mathbb{N}$

Widerspruch!

"$\Leftarrow$" Das Folgenkriterium sei erfüllt; wir zeigen, dass A abgeschlossen ist, d.h. $X \setminus A$ offen ist.

Sei $a \in X \setminus A$ ein beliebiger Punkt.

<u>Behauptung</u>: Es gibt ein $\varepsilon > 0$ mit $B_\varepsilon(a) \subset X \setminus A$.

Wäre dies nicht der Fall, könnten wir zu jedem $k > 0$ ein $x_k \in A$ finden mit $d_X(x_k, a) < \frac{1}{k}$.

Dann gilt $\lim\limits_{k \to \infty} x_k = a$; dies im Widerspruch zu $a \in X \setminus A$ steht.

Die Behauptung ist also richtig, was zeigt, dass $X \setminus A$ offen ist. $\blacksquare$

<u>Definition 8.3.4</u> Sei (X,d) ein metrischer Raum. Eine Folge $(x_k)_{k \in \mathbb{N}}$ von Punkten aus X heißt <u>Cauchy-Folge</u>, wenn gilt: $\forall\, \varepsilon > 0 \;\exists\, N \in \mathbb{N}$,

so dass $d(x_k, x_m) < \varepsilon \quad \forall \; k, m \geq N$.

Bemerkung: Jede konvergente Folge in einem metrischen Raum ist eine Cauchy-Folge. Dies ist genau wie im Fall von reellen Folgen.
(siehe Satz 2.3.6)

Definition 8.3.5: Ein metrischer Raum heißt vollständig, wenn in ihm jede Cauchy-Folge konvergiert.
Ein vollständiger normierter Vektorraum heißt Banach-Raum.

Erinnere (Satz 2.3.6) in $\mathbb{R}$ jede Cauchy-Folge konvergiert. Daher ist $\mathbb{R}$ ein ~~metrischer~~ vollständiger metrischer Raum. (bzgl. der euklischen Metrik).
In der Tat ist $\mathbb{R}$ ein Banach-Raum bzgl. der euklischen Norm (siehe Beispiel 8.2.5(2)).

Satz 8.3.6 Im $\mathbb{R}^n$ konvergiert jede Cauchy-Folge, d.h. $\mathbb{R}^n$ ist ein Banach-Raum.

Beweis Sei $x_k = (x_{k1}, x_{k2}, \cdots, x_{kn})$, $k \in \mathbb{N}$, so dass ~~und~~ $(x_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge ist. Da

$$|x_{k\ell} - x_{m\ell}| \leq \| x_k - x_m \| \quad \forall \; 1 \leq \ell \leq n,$$

ist die Folge $(x_{k\ell})_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$.

Daher ist $(x_{k\ell})_{k\in\mathbb{N}}$ konvergent. $\forall\ 1\leq \ell \leq n$.

Nach Satz 8.3.2 konvergiert $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$. ■

<u>Bemerkung ~~8.3.6~~</u> Seien $[a,b]\subset\mathbb{R}$

$\mathcal{C}^0([a,b])$ wie im

Beispiel 8.2.5 (6). Erinnere $\mathcal{C}^0([a,b])$ ist ein normierter Raum mit der Sup-norm. Wir zeigen nun $\left(\mathcal{C}^0([a,b]), \|\cdot\|_{[a,b]}\right)$ ist ein Banach-Raum.

In der Tat sei $(f_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}^0([a,b])$, d.h. $\forall \varepsilon>0\ \exists N\in\mathbb{N}$ mit

$\|f_k - f_m\|_{[a,b]} < \varepsilon \quad \forall k,m \geq N$. Es folgt

$(f_k(x))_{k\in\mathbb{N}}$ ist eine Cauchy-Folge in $\mathbb{R}$

Daraus folgt: $(f_k(x))_k$ konvergent und

Setze $f(x) := \lim\limits_{k\to\infty} f_k(x) \in\mathbb{R}$.

Man bekommt $f: [a,b]\to\mathbb{R}$ und

f_k konvergiert gleichmäßig gegen f für $k\to\infty$

(<u>weil</u>: $|f_k(x) - f_m(x)| < \varepsilon \quad \forall x\in[a,b]\ \forall k,m\geq N$

In Betracht $m\to\infty$ bekommt man

$|f_k(x) - f(x)| < \varepsilon\ \forall x\in[a,b]\ \forall k\geq N$

$\Rightarrow \quad \|f_k - f\|_{[a,b]} < \varepsilon \quad \forall k\geq N$)

Nach Satz 7.3 ist f dann stetig, anders gesagt $f \in \mathcal{C}^0([a,b])$. Daher ist $(\mathcal{C}^0([a,b]), \|\cdot\|_{[a,b]})$ ein Banach-Raum.

<u>Definition 8.3.7</u> Eine Teilmenge A eines metrischen Raumes (X,d) heißt <u>beschränkt</u> wenn $\exists M \in \mathbb{N}$ so dass $d(x,y) \leq M \quad \forall x,y \in A$.

A

<u>Beispiel</u>: i) jede offene Kugel $B_r(x) \subset (X,d)$ ist beschränkt.

ii) $A \subset (X,d)$ beschränkt $\Leftrightarrow \exists \, x \in X, \ r \in \mathbb{R}_+$ so dass $A \subset B_r(x)$.

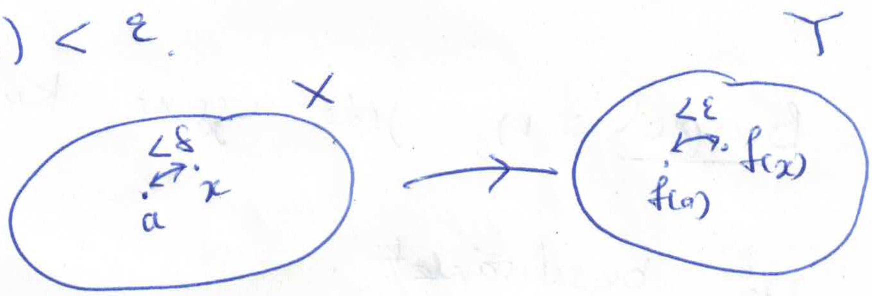
iii) jede konvergente Folge in (X,d) ist beschränkt, allgemeiner ist jede Cauchy-Folge beschränkt.

(iv) Wenn $X = \mathbb{R}$ oder $\mathbb{C}$ (mit der euklidischen Metrik), findet man den üblichen Begriff von Beschränktheit.

Wir besprechen nun die Stetigkeit von Abbildungen zwischen metrischen Räume. Dies verallgemeinert den Abschnitt 4.1, in dem wir diesen Begriff für Abbildungen zwischen Teilmengen von $\mathbb{C} \cong \mathbb{R}$ betrachtet.

<u>Definition 8.3.8</u> Seien (X, d_X), (Y, d_Y) metrische Räume und $f: X \to Y$ eine Abbildung. f heißt <u>stetig im Punkt $a \in X$</u>, falls $\forall \varepsilon > 0$ $\exists$ $\delta = \delta(\varepsilon, a) > 0$ so dass $\forall x \in X$ mit $d_X(x,a) < \delta$ gilt

$$d_Y(f(x), f(a)) < \varepsilon.$$



f heißt <u>stetig auf X</u>, wenn f in jedem Punkt $a \in X$ stetig.

Man beachte, dass wenn $X \subset \mathbb{C}$, $Y \subset \mathbb{C}$, ist die Definition 8.3.8 die übliche Definition für die Stetigkeit (Definition 4.1.1)

<u>Beispiel</u>: i)
$$f: X \to Y, \quad x \mapsto y_0 \quad \forall x \in X$$
ist stetig (konstante Abbildungen)

ii) $f: X \to X$ (identische Abbildung)
$x \mapsto x$

ist stetig.

iii) Sei $f: X \to Y$ stetig, $A \subset X$. Dann ist $f|_A : A \to Y$ (die Einschränkung von f auf A) auch stetig. (Erinnere, dass A selbst ein metrischer Raum ist mit der von X induzierten Metrik).

Definition 8.3.9 Sei X, Y metrische Räume, $A \subset X$, $a \in X$ einer Häufungspunkt von A. Eine Funktion $f: A \to Y$ hat in a den Grenzwert $b \in Y$, falls die Abbildung

$$F: A \cup \{a\} \to Y$$
$$x \mapsto f(x), \; x \in A \setminus \{a\}$$
$$a \mapsto b$$

in a stetig ist. In diesem Fall schreibt man

$$\lim_{x \to a} f(x) = b.$$

Bemerkung: $f: X \to Y$ stetig iff $\forall a \in X$ gilt $\lim_{x \to a} f(x) = f(a)$.

Satz 8.3.10 (Im Vergleich zu den Satz 4.1.4)
Sei (X, d_X), (Y, d_Y) metrische Räume und $a \in X$,

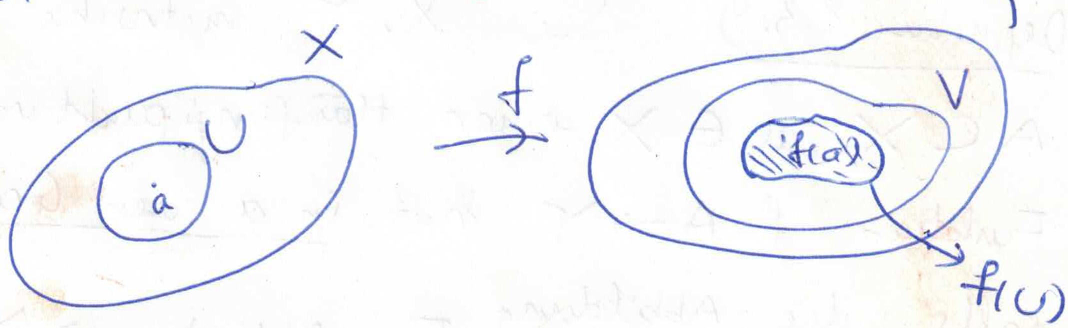
$f: X \to Y$ eine Abbildung.

Die folgende Aussagen sind äquivalent:

(i) f ist stetig in a.

(ii) $\forall$ Umgebung V von $f(a)$ in Y gibt es eine Umgebung U von a in X mit $f(U) \subset V$

(iii) $\forall$ $(x_k)_{k\in\mathbb{N}} \subset X$ mit $\lim\limits_{k\to\infty} x_k = a$ gilt $\lim\limits_{k\to\infty} f(x_k) = f(a)$. (Folgenkriterium für Stetigkeit)



Beweis Der Beweis von Satz 4.1.4 verallgemeinert sich formell auf den Fall von metrischen Räume.

(i) $\Rightarrow$ (ii): Sei V eine Umgebung von $f(a)$ in Y. Dann $\exists\, \varepsilon > 0$ mit $B_\varepsilon(f(a)) \subset V$. Nach Stetigkeit in a $\exists\, \delta > 0$ mit: $\forall x \in B_\delta(a)$ gilt $f(x) \in B_\varepsilon(f(a)) \subset V$. Daher $f(U) \subset V$.

(ii) $\Rightarrow$ (iii) Sei $\varepsilon > 0$. Nach (ii) $\exists\, \varepsilon_1 > 0$ so dass $f(B_{\varepsilon_1}(a)) \subset B_\varepsilon(f(a))$.

Da $\lim_{k\to\infty} x_k = a$, $\exists\, N \in \mathbb{N}$ mit $x_k \in B_{\varepsilon_1}(a)$ $\forall k \geq N$. Es folgt $f(x_k) \in B_\varepsilon(f(a))$ $\forall k \geq N$

anders gesagt, $d_Y(f(x_k), f(a)) \leq \varepsilon$ $\forall k \geq N$.

Deshalb gilt $\lim_{k\to\infty} f(x_k) = f(a)$.

<u>(iii) $\Rightarrow$ (i)</u> Angenommen, f ist nicht in a stetig.

Dann $\exists\, \varepsilon > 0$ $\forall \delta > 0$ $\exists\, x \in X$ mit $d_X(x,a) < \delta$ und $d_Y(f(x), f(a)) \geq \varepsilon$. Wähle $\delta = \frac{1}{k}$ und $x_k \in B_{\frac{1}{k}}(a)$, bekommt man

$$d_Y(f(x_k), f(a)) \geq \varepsilon \quad \forall k \in \mathbb{N}.$$

Andererseits $\lim_{k\to\infty} x_k = a$, Widerspruch zu (iii)! ∎

<u>Folgerung 8.3.11</u> Seien X, Y metrische Räume und $f: X \to Y$ eine Abbildung. Dann gilt:

f ist stetig $\Leftrightarrow$ $\forall V \subset Y$ offen ist das Urbild $f^{-1}(V)$ auch offen (in X).

<u>Beweis</u>. "$\Leftarrow$" ist klar nach Satz 8.3.10 (i)$\Leftrightarrow$(ii)

"$\Rightarrow$" ist auch offensichtlich nach dem gleichen Grund.

Bemerkung 8.3.12 Sei X ein metrischer Raum. $f: X \to \mathbb{R}$ stetig, $c \in \mathbb{R}$. Dann sind

$$U_1 := f^{-1}((-\infty, c)) = \{x \in X : f(x) < c\}$$

$$U_2 := f^{-1}((c, \infty)) = \{x \in X : f(x) > c\}$$

offen. Daher ist

$$A := \{x \in X : f(x) = c\} = X \setminus (U_1 \cup U_2)$$

abgeschlossen.

Satz 8.3.13 Seien X, Y, Z metrische Räume und $f: X \to Y$, $g: Y \to Z$ Abbildungen. Ist f stetig im Punkt $a \in X$ und g stetig in $b := f(a) \in Y$, so ist $g \circ f : X \to Z$ stetig in a.

Beweis. Ist $\lim_{k\to\infty} x_k = a$, so folgt $\lim_{k\to\infty} f(x_k) = b$ wegen der Stetigkeit von f in a. Da g stetig in b ist, erhält man $\lim_{k\to\infty} g(f(x_k)) = g(b)$, also

$$\lim_{k\to\infty} (g \circ f)(x_k) = g \circ f(a).$$ Nach Folgenkriterium ist $g \circ f$ stetig in a. ∎

Betrachte nun $f: X \to \mathbb{R}^n$, $i_l : \mathbb{R}^n \to \mathbb{R}$ mit $i_l(x_1, \dots, x_n) = x_l$ für $1 \le l \le n$.

Setze $f_\ell := i_\ell \circ f$. Man bekommt

$$f(x) = (f_1(x), \dots, f_n(x)) \quad \forall x \in X.$$

Man sagt, f_ℓ die ℓ-te Komponenten-Funktion von f ist.

Folgerung 8.3.14 Sei X ein metrischer Raum.
Eine Abbildung $f = (f_1, \dots, f_n) : X \to \mathbb{R}^n$ ist genau dann stetig, wenn alle Komponenten $f_\ell : X \to \mathbb{R}$, $\ell = 1, 2, \dots, n$ stetig sind.

Beweis "$\Rightarrow$" ist offensichtlich nach Satz 8.3.13

Wir zeigen nun "$\Leftarrow$". Sei $(x_k)_{k\in\mathbb{N}} \subset X$ eine Folge mit $\lim\limits_{k\to\infty} x_k = a \in X$.

Da f_ℓ stetig ist, gilt $\lim\limits_{k\to\infty} f_\ell(x_k) = f_\ell(a)$.

Daraus folgt $\lim\limits_{k\to\infty} f(x_k) = f(a)$

(siehe Satz 8.3.2). ∎

Folgerung 8.3.15 Folgende Abbildungen sind stetig:

a) add : $\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ $\quad (x,y) \mapsto x+y$
b) mult : $\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ $\quad (x,y) \mapsto xy$
c) quot : $\mathbb{R}\times \mathbb{R}^* \to \mathbb{R}$ $\quad (x,y) \mapsto xy^{-1}$

wobei $\mathbb{R}^* := \mathbb{R}\setminus\{0\}$

<u>Beweis</u> Direkt aus Folgenkriterium der Stetigkeit.

<u>Satz 8.3.16</u> Sei X ein metrischer Raum und seien $f, g : X \to \mathbb{R}$ stetige Abbildungen. Dann sind auch die Funktionen $f+g : X \to \mathbb{R}$, $f \cdot g : X \to \mathbb{R}$, $\lambda f : X \to \mathbb{R}$ $(\lambda \in \mathbb{R})$ stetig. Gilt außerdem $g(x) \neq 0 \; \forall x \in X$, so ist auch $\frac{f}{g} : X \to \mathbb{R}$ stetig.

<u>Beweis</u>: Nach Folgerung 8.3.14 ist $(f, g) : X \to \mathbb{R}^2$ stetig. Nun gilt

$$f+g = \mathrm{add} \circ (f, g)$$
$$fg = \mathrm{mult} \circ (f, g)$$
$$f/g = \mathrm{quot} \circ (f, g)$$

Aus Folgerung 8.3.15 und Satz 8.3.13 sind ~~all~~ alle dreie Funktionen stetig. Die Stetigkeit von λf ist offenbar. ■

<u>Beispiel 8.3.17</u> Sei $r \in \mathbb{N}_0$.

Sei $I_r = \{ (k_1, \dots, k_n) : k_j \in \mathbb{N}_0, \; k_1 + \dots + k_n \le r \}$

Eine Polynomfunktion $F : \mathbb{R}^n \to \mathbb{R}$ vom Grad $\le r$ ist von der Gestalt:

$$F(x_1, \dots, x_n) = \sum_{(k_1, \dots, k_n) \in I_r} a_{(k_1 \dots k_n)} \, x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

wobei $c_{(k_1,\dots,k_n)} \in \mathbb{R}$. Nach wiederholte Anwendungen von Satz 8.3.16 ist F stetig.

<u>Definition 8.3.18</u> Seien X eine beliebige Menge, (Y, d_Y) ein metrischer Raum, sowie

$$f_n : X \to Y, n \in \mathbb{N} \quad \text{und} \quad f : X \to Y$$

Abbildungen. Man sagt, die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen f, falls zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass

$$d_Y(f_n(x), f(x)) < \varepsilon \quad \forall x \in X, \forall n \geq N.$$

(In Vergleich zu der Definition 7.1 (2))

Diese verallgemeinert Definition 7.1 (2), in der der Fall $X \subset \mathbb{C} \cong \mathbb{R}^2$, $Y = \mathbb{C}$ betrachtet wird.

<u>Satz 8.3.19</u> Seien X, Y metrischer Räume

$$f_n : X \to Y \quad n \in \mathbb{N}$$

eine Folge stetiger Funktionen, die gleichmäßig gegen die Funktion $f : X \to Y$ konvergiere. Dann ist f auch stetig.

<u>Beweis</u>. Im Satz 7.3 haben ~~da~~ wir die Aussage für $X \subset \mathbb{C}$, $Y \subset \mathbb{C}$ bewiesen. Der Beweis überträgt sich formell auf den Fall

mit metrischen Räumen. ∎

<u>Beispiel 8.3.20</u> Sei $U \subset \mathbb{R}^n$ eine Teilmenge und $\mathcal{C}_b(U, \mathbb{R}^m) := \{ f : U \to \mathbb{R}^m \text{ stetig und beschränkt} \}$

wobei f beschränkt bedeutet: das Bild $f(U)$ ist beschränkt in $\mathbb{R}^m$.

Man beachtet, dass $\mathcal{C}_b(U, \mathbb{R}^m)$ ein $\mathbb{R}$-Vektorraum ist. Wir versehen $\mathcal{C}_b(U, \mathbb{R}^m)$ mit der Sup-norm $\|\cdot\|_U$.

Wie im Bemerkung ~~8.3.6~~ in der Seite 309, nach Satz 8.3.19, ist $(\mathcal{C}_b(U, \mathbb{R}^m), \|\cdot\|_U)$ ein Banach-Raum.

($\|f\|_U := \sup\{ \|f(x)\| : x \in U \} < \infty$)

Wir betrachten nun einen speziellen Fall von stetigen Abbildungen.

<u>Satz 8.3.21</u> (Lineare Abbildungen) Seien V und W normierte Vektorräume (über $\mathbb{R}$ oder $\mathbb{C}$) und sei

$A : V \to W$ eine lineare Abbildung.

A ist genau dann stetig, wenn es eine reelle Konstante $C \geq 0$ gibt, so dass

$$\|A(x)\| \leq C\|x\| \quad \forall x \in V.$$

<u>Beweis</u> "$\Rightarrow$" Angenommen, A ist stetig.

Insbesondere ist A im Nullpunkt stetig. Dann gibt es zu $\varepsilon = 1$ ein $\delta > 0$, so dass

$$\|A(z)\| < 1 \quad \forall\, z \in V \text{ mit } \|z\| < \delta.$$

Wir setzen $C := \frac{2}{\delta}$. Sei $x \in V \setminus \{0\}$ beliebig, $\lambda := \frac{1}{C\|x\|}$ und $z := \lambda x$. Dann gilt

$$\|z\| = |\lambda| \, \|x\| = \frac{\delta}{2} < \delta, \text{ also } \|A(z)\| < 1.$$

Nun ist

$$A(z) = A(\lambda x) = \lambda A(x) = \frac{1}{C\|x\|} A(x),$$

also folgt

$$\|A(x)\| \leq C\|x\|.$$

"$\Longleftarrow$" Es gebe eine Konstante $C \geq 0$ mit $\|A(x)\| \leq C\|x\| \quad \forall x \in V$. Dann gilt

$$\|A(x) - A(x_0)\| = \|A(x - x_0)\| \leq C \underbrace{\|x - x_0\|}_{\xrightarrow{x \to x_0}\ 0}$$

Daher $\lim\limits_{x \to x_0} A(x) = A(x_0)$, d.h. A in x_0 stetig $\forall x_0 \in V$ ∎

<u>Definition 8.3.22</u> Seien V und W normierte ~~Räume~~ Vektorräume und $A : V \to W$ eine stetige lineare Abbildung. Dann wird ihre Norm definiert als

$$\|A\| := \sup\{\|A(x)\| : x \in V \text{ mit } \|x\| \leq 1\}$$

<u>Bemerkung 8.3.23</u> (i) Nach Satz 8.3.21 gilt

$$\|A\| < \infty$$

und $\|A(x)\| \leq \|A\| \, \|x\| \quad \forall x \in V.$

Setze

$$\mathcal{C}(V,W) := \{ A : V \to W \text{ stetige lineare Abbildungen} \}$$

Dann ist $\mathcal{C}(V,W)$ Vektorraum und mit der Norm $\|\cdot\|$ in der Definition 8.3.22 ist $\mathcal{C}(V,W)$ ein normierter Vektorraum.

(ii) Betrachte nun einen speziellen Fall:

$V = \mathbb{R}^n$ und $W = \mathbb{R}^m$. Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Bezüglich der kanonischen Basen von $\mathbb{R}^n, \mathbb{R}^m$ wird A durch eine $m \times n$-Matrix

$$(a_{ik})_{1 \leq i \leq m,\ 1 \leq k \leq n} \in M(m \times n, \mathbb{R})$$

gegeben:

$$A : \mathbb{R}^n \to \mathbb{R}^m$$

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \to \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Nach Beispiel 8.3.17 und ~~Satz~~ Folgerung 8.3.14 ist A stetig (da Komponenten-Funktionen von A Polynomfunktionen sind).

Man kann zeigen: $\max\limits_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}} |a_{ik}| \leq \|A\| \leq \sqrt{nm} \max\limits_{\substack{1 \leq i \leq m \\ 1 \leq k \leq n}} |a_{ik}|$

(Aufgabe)

§ 8.4. Kompaktheit

Nach Satz von Balzano-Weierstraß (Satz 2.3.3), besitzt jede beschränkte Folge in $\mathbb{R}$ eine Teilfolge, die konvergent ist. Dies gilt auch für $\mathbb{R}^n$.

Satz 8.4.1 (Balzano-Weierstraß in $\mathbb{R}^n$) Jede beschränkte Folge in $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.

Beweis Sei $(x_k)_{k\in\mathbb{N}} \subset \mathbb{R}^n$ eine beschränkte Folge, d.h. $\exists\ M>0$ so dass $\|x_k\| \leq M \ \forall k \in \mathbb{N}$

Man schreibt $x_k = (x_{k1}, \dots, x_{kn}) \in \mathbb{R}^n$.

Dann gilt $|x_{kj}| \leq M \quad \forall\, k \in \mathbb{N},\ \forall\ 1 \leq j \leq n$

(da $\|x_k\| = \sqrt{x_{k1}^2 + \dots + x_{kn}^2} \geq x_{kj}$)

Anders gesagt, die Folge $(x_{kj})_{k\in\mathbb{N}}$ ist eine beschränkte Folge in $\mathbb{R}$ für $1 \leq j \leq n$.

Daher für $j=1$ $\exists\ (s_k^{(1)})_{k\in\mathbb{N}} \subset \mathbb{N}$ mit

$$\lim_{k\to\infty} x_{s_k^{(1)}1} = a_1 \in \mathbb{R}$$

für $j=2$ $\exists\ (s_k^{(2)})_{k\in\mathbb{N}} \subset (s_k^{(1)})_{k\in\mathbb{N}}$ mit

$$\lim_{k\to\infty} x_{s_k^{(2)}2} = a_2 \in \mathbb{R}$$

rekursiv bekommt man Folgen

$$(s_k^n)_k \subset (s_k^{n-1})_k \subset \cdots \subset (s_k^2)_k \subset (s_k^1)_k$$

und $\lim_{k\to\infty} x_{s_k^j j} = a_j \in \mathbb{R}$

Daraus folgt $\lim_{k\to\infty} x_{s_k^n j} = a_j \in \mathbb{R} \quad \forall j$

(da $(x_{s_k^n j})_k \subset (x_{s_k^j j})_k$)

Daher gilt
$$\lim_{k\to\infty} x_{s_k^n} = (a_1, \dots, a_n) \in \mathbb{R}^n.$$ $\blacksquare$

<u>Definition 8.4.2</u> Sei X ein metrischer Raum und $A \subset X$ eine Teilmenge. A heißt <u>kompakt</u>, wenn jede Folge in A eine gegen einen Punkt aus A konvergierende Teilfolge enthält.

<u>Beispiel</u> (i) Jedes kompakte Intervall in $\mathbb{R}$ ist kompakt. d.h. $\forall\ a, b \in \mathbb{R}\ (a<b)$ ist $[a,b]$ kompakt.

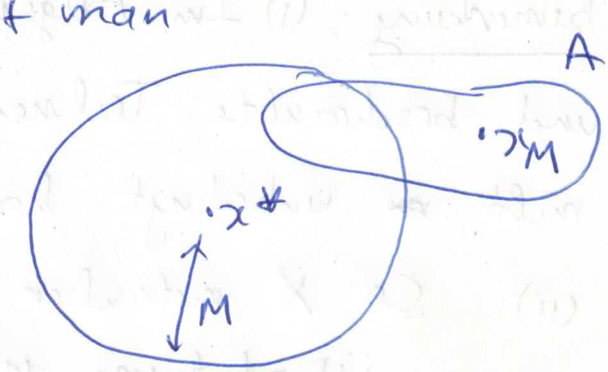
(ii) Sei $(x_k)_k \subset X$ eine konvergente Folge mit $\lim_{k\to\infty} x_k = a$. Dann ist die Menge $A := \{x_k : k\in\mathbb{N}\} \cup \{a\}$ kompakt.

<u>Satz 8.4.3</u> Sei (X,d) ein metrischer Raum, $A \subset X$ eine kompakte Teilmenge. Dann ist A abgeschlossen und beschränkt.

Beweis Sei $x^* \in X$. Wäre A nicht beschränkt, gelte: $\forall M>0 \quad \exists x_M \in A \setminus B_M(x^*)$.

Wähle $M = k \in \mathbb{N}$ bekommt man $(x_k)_{k\in\mathbb{N}} \subset A$ mit

$d(x_k, x^*) \geq k \quad \forall k \in \mathbb{N}$



Nach Kompaktheit von A $\exists$ $(x_{n_k})_k \subset (x_k)$ mit $\lim\limits_{k\to\infty} x_{n_k} = a \in A$. Es folgt

$d(x_{n_k}, x^*) \geq k \quad \forall k \in \mathbb{N}$. Nimmt $k \to \infty$ in der Ungleichung erhält man $d(a, x^*) = \infty$ Widerspruch! Daher ist A beschränkt.

Die Abgeschlossenheit von A folgt direkt aus Satz 8.3.3 ∎

Satz 8.4.4 (Heine-Borel) Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis "$\Rightarrow$" ist direkt aus Satz 8.4.3.

Nun sei A beschränkte und abgeschlossene Teilmenge in $\mathbb{R}^n$. Sei $(x_k)_k \subset A$ eine Folge. Daher ist $(x_k)_k$ beschränkt. Satz 8.4.1 $\Rightarrow$ $(x_k)_k$ besitzt eine konvergente Teilfolge. Der Grenzpunkt dieser Teilfolge muss notwendig in A liegen wegen der Abgeschlossenheit von A (Satz 8.3.3). ∎

<u>Bemerkung</u> 8.4.5: (i) Im Allgemeinen ist eine abgeschlossene und beschränkte Teilmenge eines metrischen Raumes nicht unbedingt kompakt.

(ii) Sei X metrischer Raum, $A \subset X$. Unter einer <u>offenen Überdeckung</u> von A versteht man eine Familie $(U_i)_{i \in I}$ von offenen Teilmengen $U_i \subset X$ mit

$$A \subset \bigcup_{i \in I} U_i$$

Dabei ist I eine beliebige Indexmenge.

Man kann zeigen, dass die folgende Aussagen äquivalent sind:

(1) A ist kompakt

(2) A erfüllt die Überdeckungseigenschaft: $\forall$ offene Überdeckung $(U_i)_{i \in I}$ von A $\exists$ endlich viele Indizes $i_1, \dots, i_k \in I$ mit

$$A \subset U_{i_1} \cup \dots \cup U_{i_k}.$$

(iii) Sei $A \subset \mathbb{R}$ kompakt. Da A beschränkt ist, sind $\sup A$, $\inf A$ endlich. Es existieren Folgen $(x_k)_k \subset A$, $(y_k)_k \subset A$ mit $\lim_{k \to \infty} x_k = \sup A$ und $\lim_{k \to \infty} y_k = \inf A$. Da A abgeschlossen ist, gilt

$$\sup A \in A, \quad \inf A \in A.$$

Zusatzblatt zur Bemerkung 8.4.5(r)

Beispiel i) Sei X eine unendliche Menge versehen mit der trivialen Metrik ($d(x,y) = 1$ falls $x \neq y$)

Dann ist X beschränkt und abgeschlossen. aber X ist nicht kompakt.

ii) Man weiß schon, dass $\mathcal{C}^0([a,b])$ oder $\mathcal{C}_b([a,b], \mathbb{R})$ Banach-Räume mit der Supremum-norm sind (Seit 309 in Skript oder Beispiel 8.3.20).

Analog ~~sind~~ ist $\mathcal{C}_b([a,b], \mathbb{C}) = \{ f : [a,b] \to \mathbb{C}$ beschränkt stetig $\}$ auch Banach-Raum mit der Sup-norm.

Betrachte

$$\overline{B_1(0)} := \{ f \in \mathcal{C}_b([a,b], \mathbb{C}) : \|f\|_{[a,b]} \leq 1 \}$$

welche beschränkt und abgeschlossen ist.

Seien nun $a = 0$, $b = \pi$

$$f_j : [0,\pi] \to \mathbb{C}$$
$$x \longmapsto e^{2jix}$$

Beachte, $f_j \in \overline{B_1(0)}$ und $\|f_j - f_\ell\|_{[0,\pi]} = 2$ für $\ell \neq j$. Daher erhält $(f_j)_j \subset \overline{B_1(0)}$ keine konvergente Teilfolge.

$\Rightarrow \overline{B_1(0)}$ ist nicht kompakt.

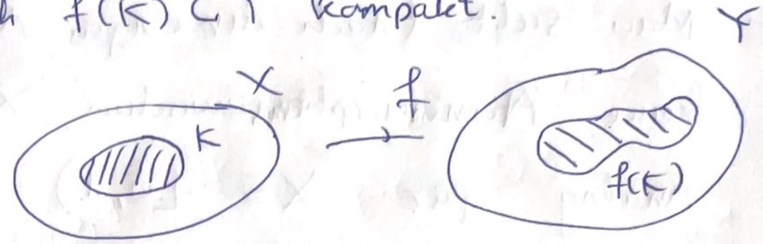
(iii) Seien $[a_j, b_j] \subset \mathbb{R}$ kompakte Intervalle für $1 \leq j \leq n$.

Setze $A := [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$

Da A beschränkt und abgeschlossen ist, so ist A kompakt in $\mathbb{R}^n$.

Satz 8.4.6 Seien X, Y metrische Räume und $f: X \to Y$ eine stetige Abbildung. Ist $K \subset X$ kompakt, so ist auch $f(K) \subset Y$ kompakt.

Beweis.



Sei $(y_k)_{k \in \mathbb{N}}$ eine Folge in $f(K)$. Wir zeigen, dass es eine Teilfolge von $(y_k)_k$ gibt, welche gegen einem Punkt in $f(K)$ konvergiert.

Da $y_k \in f(K)$, $\exists\, x_k \in K$ mit $f(x_k) = y_k$.

Daher $(x_k)_k \subset K$. Wegen der Kompaktheit von K

$\exists\, (x_{k_\ell})_\ell \subset (x_k)_k$ mit $\lim\limits_{\ell \to \infty} x_{k_\ell} = a \in K$.

Stetigkeit von $f \Rightarrow \lim\limits_{\ell \to \infty} f(x_{k_\ell}) = f(a) \in f(K)$

Folglich $\lim\limits_{\ell \to \infty} y_{k_\ell} = f(a) \in f(K)$. ∎

Bemerkung Seien X, Y metrische Räume

$f: X \to Y$ heißt Homöomorphismus, falls f bijektiv ist und f, f^{-1} stetig sind. Man sagt, X, Y sind homöomorph falls $\exists$ Homöomorphismus zwischen X und Y.

Aus Satz 8.4.6 bekommt man: ist einer der beiden X, Y kompakt, so auch der andere.

z.B. $X = \{ x \in \mathbb{R}^n : \|x\| \leq 1 \}$ kompakt
$Y = \mathbb{R}^m$ nicht kompakt

Man sieht (nach der obigen Aussage), dass es keine Homöomorphismus zwischen X und Y gibt.

Analog für $X = [0,1)$ und
$Y = S^1 = \{ (x,y) \in \mathbb{R}^2 : x^2+y^2=1 \}$

<u>Satz 8.4.7</u> Sei X eine kompakte metrische Raum und $f : X \to \mathbb{R}$ eine stetige Funktion. Dann ist f beschränkt und nimmt ihre Maximum und Minimum an, d.h. $\exists$ $a, b \in X$ mit

$$f(a) = \sup \{ f(x) : x \in X \}$$
$$f(b) = \inf \{ f(x) : x \in X \}$$

(Im Vergleich zum Satz 4.4.3 für den Fall $X = [a,b] \subset \mathbb{R}$)

<u>Beweis</u> Nach Satz 8.4.6 ist $f(X)$ kompakt in $\mathbb{R}$. Daher ist $f(X)$ beschränkt und

$$\sup f(X) \in f(X) \quad \text{(Bemerkung 8.4.5(iii))}$$
$$\inf f(X) \in f(X)$$

Daraus folgt $\exists$ $a, b \in X$: $f(a) = \sup f(X)$
$f(b) = \inf f(X)$ ∎

Definition 8.4.8 (siehe auch Definition 4.4.5)

Seien X, Y metrische Räume. Eine Abbildung $f: X \to Y$ heißt **gleichmäßig stetig**, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$d_Y(f(x), f(x')) < \varepsilon \quad \forall x, x' \in X \text{ mit } d_X(x, x') < \delta.$$

Im Satz 4.4.6 wurde es gezeigt, dass jede stetige Abbildung $f: [a,b] \to \mathbb{R}$ gleichmäßig stetig ist. $(-\infty < a < b < \infty)$. Hier ist eine Verallgemeinerung:

Satz 8.4.9 Seien X, Y metrische Räume und sei X kompakt. Dann ist jede stetige Abbildung $f: X \to Y$ gleichmäßig stetig.

Beweis Analog zu dem Beweis von Satz 4.4.6.

Angenommen, f nicht gleichmäßig stetig. Dann gäbe es $\varepsilon > 0$: $\forall n \in \mathbb{N}$ $\exists x_n, y_n \in X$ mit $d_X(x_n, y_n) < \frac{1}{n}$ und $d_Y(f(x_n), f(y_n)) > \varepsilon$.

Nach Kompaktheit von X $\exists (x_{n_\ell})_\ell \subset (x_n)_n$, $(y_{n_\ell}) \subset (y_n)_n$ (man kann dieselbe Index n_ℓ wählen)

so dass

$$\lim_{l\to\infty} x_{n_l} = a \in X$$

$$\lim_{l\to\infty} y_{n_l} = b \in X$$

Da $d_X(x_{n_l}, y_{n_l}) < \frac{1}{n}$, folgt $a = b$.

Stetigkeit von f impliziert

$$\lim_{l\to\infty} f(x_{n_l}) = f(a)$$

$$\lim_{l\to\infty} f(y_{n_l}) = f(b) = f(a)$$

Das widerspricht der Tatsache dass

$$d_Y(f(x_{n_l}), f(y_{n_l})) > \varepsilon \ \forall l$$

(lässt $l\to\infty$ erhält man $d_Y(f(a), f(a)) > \varepsilon$) ∎

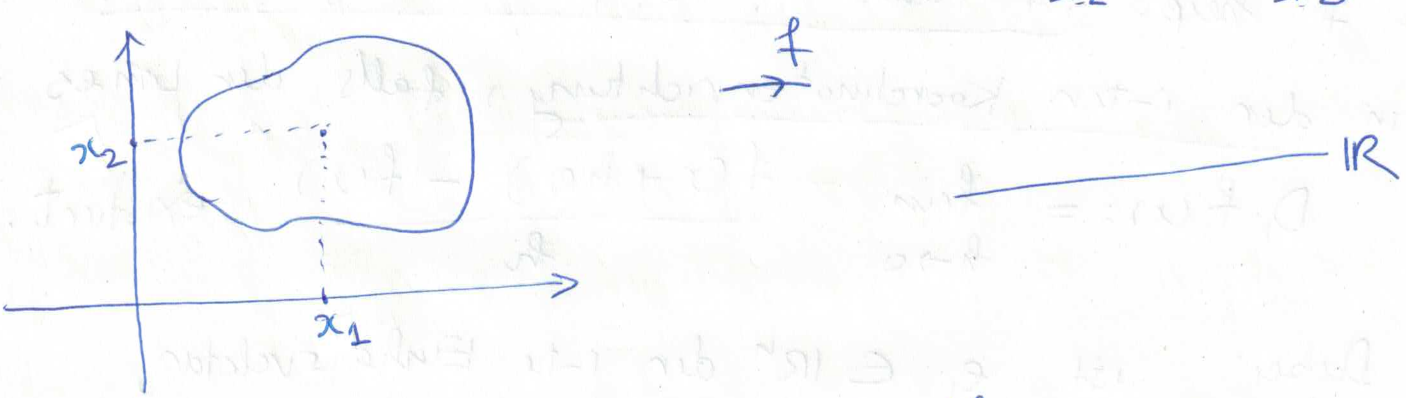
§ 9. Partielle Ableitungen

§ 9.1. Definitionen und Beispiele

Im letzten Kapitel wurde der Begriff von Stetigkeit untersucht. Wir beschäftigen uns nun mit der Differenzierbarkeit.

Diesen Begriff haben wir im Kapitel 5 für Funktionen einer Variable eingeführt. Um die Idee zu erklären, betrachten wir das folgende Beispiel:

Sei $U \subset \mathbb{R}^2$ eine offene Menge, $f: U \to \mathbb{R}$, $(x_1, x_2) \mapsto f(x_1, x_2)$



f ist eine Funktion von zwei Variablen.

Wir fassen x_2 als einen Parameter auf.

Daher $\forall x_2$ fest bekommt man $g_{x_2}(x_1) := f(x_1, x_2)$ eine Funktion von x_1. Man kann die Differenzierbarkeit von g_{x_2} untersuchen. Die Ableitung von g_{x_2}, wenn existiert, heißt die partielle Ableitung von f bzgl. x_1 für x_1 als Parameter.

Analog

Erinnere
$$g'_{x_2}(x_1) = \lim_{h\to 0} \frac{g_{x_2}(x_1+h) - g_{x_2}(x_1)}{h}$$
$$= \lim_{h\to 0} \frac{f(x_1+h, x_2) - f(x_1, x_2)}{h}$$

Schreibt $f(x_1+h, x_2) = f\big((x_1,x_2) + h\underbrace{(1,0)}_{=: e_1}\big)$

$$\Rightarrow \quad g'_{x_2}(x_1) = \lim_{h\to 0} \frac{f((x_1,x_2)+he_1) - f(x_1,x_2)}{h}.$$

Im Allgemeinen haben wir

<u>Definition 9.1.1</u> Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und $f: U \to \mathbb{R}$ eine reelle Funktion.

f heißt <u>im Punkt $x \in U$ partiell differenzierbar in der i-ten Koordinatenrichtung</u>, falls der Limes
$$D_i f(x) := \lim_{h\to 0} \frac{f(x+he_i) - f(x)}{h} \quad \text{existiert}$$

Dabei ist $e_i \in \mathbb{R}^n$ der i-te Einheitsvektor,
$$e_i := (0, \dots, 0, \underbrace{1}_{i\text{-te Stelle}}, 0, \dots, 0)$$

Für den Limes $h \to 0$, betrachte man $h \in \mathbb{R}\setminus\{0\}$ hinreichend klein mit $x + he_i \in U$.

$D_i f(x)$ heißt i-te <u>Partielle Ableitung von f in x</u>.

Wie in der Diskussion am Anfang kann man die partiellen Ableitungen von $f: U \to \mathbb{R}$ als gewöhnliche Ableitungen von Funktionen einer Variablen interpretieren:

Sei $x = (x_1, x_2, \dots, x_n) \in U$ ein fester Punkt.

Für $i = 1, 2, \dots, n$ betrachten wir die Funktionen

$$\xi \longmapsto f_i(\xi) := f(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots x_n)$$

Aus der Definition der partiellen Ableitung folgt

$$D_i f(x) = \lim_{h \to 0} \frac{f_i(x_i + h) - f_i(x_i)}{h} = f_i'(x_i).$$

Deshalb gelten für die partiellen Ableitungen analoge Rechenregeln wie für die gewöhnlichen Ableitungen.

<u>Definition 9.1.2</u> Sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$.

f heißt <u>partiell differenzierbar</u>, falls $D_i f(x)$ für $\forall x \in U$ und $1 \leq i \leq n$ existiert.

f heißt <u>stetig</u> ~~db~~ <u>partiell differenzierbar</u>, falls zusätzlich alle partiellen Ableitungen $D_i f: U \to \mathbb{R}$ stetig.

<u>Schreibweise</u>: statt $D_i f$ schreibt man auch $\frac{\partial f}{\partial x_i}$, und

$$D_i f(x) = \frac{\partial f}{\partial x_i}(x) = \frac{\partial f(x)}{\partial x_i}.$$

Beispiel 9.1.3

i) Sei $f: \mathbb{R}^2 \to \mathbb{R}$
$(x,y) \mapsto e^{xy}$

Dann
$$\frac{\partial f}{\partial x}(x,y) = y\, e^{xy} \quad \left((e^{cx})' = c\, e^{cx} \quad c \text{ Konstante} \right)$$

$$\frac{\partial f}{\partial y}(x,y) = x\, e^{xy}$$

ii) Sei $r: \mathbb{R}^n \to \mathbb{R}$
$x \mapsto \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \|x\|$

$x = (x_1, x_2, \dots x_n)$

Behauptung: Die Funktion r ist in $\mathbb{R}^n \setminus \{0\}$ partiell differenzierbar und es gilt
$$\frac{\partial r}{\partial x_i}(x) = \frac{x_i}{r(x)} \quad \text{für } x = (x_1, x_2, \dots, x_n) \neq 0.$$

Beweis: Betrachte $1 \le i \le n$ und $(x_1, x_2, \dots x_n) \in \mathbb{R}^n \setminus \{0\}$
$$f_i(\xi) = \sqrt{x_1^2 + \dots + x_{i-1}^2 + \xi^2 + x_{i+1}^2 + \dots + x_n^2}$$

wobei ξ in einer kleinen Umgebung von x_i in $\mathbb{R}$ ist.

Daher gilt (Kettenregel)
$$f_i'(\xi) = \frac{1}{2\sqrt{x_1^2 + \dots + x_{i-1}^2 + \xi^2 + x_{i+1}^2 + \dots + x_n^2}} \cdot 2\xi$$
$$= \frac{\xi}{f_i(\xi)}$$

Daraus folgt $\frac{\partial r}{\partial x_i}(x) = \frac{x_i}{r(x)}$ (Durch Substitution $\xi = x_i$)

Allgemeiner betrachte $f: \mathbb{R}_+ \to \mathbb{R}$ eine differenzierbare Funktion ($\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$).
Dann ~~die~~ ist die zusammengesetzte Funktion:

$$\mathbb{R}^n \setminus \{0\} \xrightarrow{f \circ r} \mathbb{R}$$
$$x \longmapsto f(r(x))$$

partiell differenzierbar und es gilt

$$\frac{\partial f(r)}{\partial x_i} = f'(r) \frac{\partial r}{\partial x_i} = f'(r) \frac{x_i}{r}$$

wobei wir $f \circ r$ kurz mit $f(r)$ bezeichnet.

(iii) Sei $n \geq 2$. Sei $F: \mathbb{R}^n \to \mathbb{R}$

$$x \longmapsto \frac{x_1 x_2 \cdots x_n}{(r(x))^n}, \quad x \neq 0$$
$$0 \longmapsto 0$$

wie in (ii) nach Kettenregel bekommt man

$$\frac{\partial F}{\partial x_1}(x) = \frac{x_2 \cdots x_n}{(r(x))^n} + x_1 x_2 \cdots x_n \cdot \frac{\partial}{\partial x_1}(r^{-n})$$

$$= \frac{x_2 \cdots x_n}{(r(x))^n} - n \cdot \frac{x_1^2 x_2 \cdots x_n}{(r(x))^{n+2}}$$

für $x \in \mathbb{R}^n \setminus \{0\}$.

Da F symmetrische von $x_1, \ldots x_n$ abhängt, (d.h. $F(x) = F(\sigma(x))$, $\forall \sigma: \{1, \ldots, n\} \to \{1, \ldots, n\}$ Permutation) erhält man die Formel für die partielle Ableitung in i-ter Koordinatenrichtung durch Vertauschen der Rollen von x_1 und x_i.

Wir berechnen die partielle Ableitungen in 0.

$$\frac{\partial F}{\partial x_i}(0) = \lim_{h \to 0} \frac{F(he_i) - F(0)}{h} = 0$$

da $F(he_i) = 0 \quad \forall h \in \mathbb{R}$. Daraus folgt, dass F in ganz $\mathbb{R}^n$ partiell differenzierbar ist. Aber

Behauptung: $\frac{\partial F}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$ nicht stetig in 0.

$F : \mathbb{R}^n \to \mathbb{R}$ nicht stetig in 0.

Beweis: Angenommen, F stetig in 0.

Dann $\forall \; (a_k)_k \subset \mathbb{R}^n$ mit $\lim_{k \to \infty} a_k = 0$ gelte

$$\lim_{k \to \infty} F(a_k) = F(0) = 0$$

wähle $a_k := \left(\frac{1}{k}, \dots, \frac{1}{k}\right), \; k \geq 1.$

Es gilt $r(a_k) = \frac{\sqrt{n}}{k}$

$$F(a_k) = \frac{k^{-n}}{\left(\frac{\sqrt{n}}{k}\right)^n} = n^{-\frac{n}{2}}$$

(unabhängig von k)

$\Rightarrow \lim_{k \to \infty} F(a_k) = n^{-\frac{n}{2}} \neq 0$ Widerspruch!

Analog für die Unstetigkeit von $\frac{\partial F}{\partial x_i}$ in 0.

Dieses Beispiel zeigt im Allgemeinen

Partielle Differenzierbarkeit $\underset{(n \geq 2)}{\not\Rightarrow}$ Stetigkeit.

Dies ist im Gegensatz zum Fall $n=1$, da jede differenzierbare Funktion einer reellen Variable stetig ist.

<u>Bemerkung</u>: Wir definieren im nächsten Kapitel eine „richtige" Verallgemeinerung des Differenzierbarkeitsbegriffs auf Funktionen mehrerer Variablen, welche die Stetigkeit nach sich zieht. Wir zeigen ~~auch~~ auch später, dass jede stetig partiell differenzierbare Funktion auch stetig ist.

<u>Definition 9.1.4</u> Sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine partiell differenzierbare Funktion. Dann heißt der Vektor

$$\operatorname{grad} f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)$$

der <u>Gradient</u> von f im Punkt $x \in U$.

Wenn $n=1$ ist $\operatorname{grad} f(x)$ nichts anders als die Ableitung von f in x. Wie im Fall $n=1$ misst $\operatorname{grad} f(x)$ das Verhalten von f in der Nähe von x.

§ 9.2. Höhere Ableitungen

Sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine partielle differenzierbare Funktion. Sind alle partiellen Ableitungen $D_i f: U \to \mathbb{R}$ selbst wieder partiell differenzierbar, so heißt f zweimal partiell differenzierbar. Man kann dann die partiellen Ableitungen zweiter Ordnung

$$D_j D_i f := D_j(D_i f)$$ bilden.

$(k \geq 1)$

Allgemeiner definiert man rekursiv: $f: U \to \mathbb{R}$ heißt <u>$(k+1)$-mal partiell differenzierbar</u>, wenn sie k-mal partiell differenzierbar ist und alle partiellen Ableitungen k-ter Ordnung

$$D_{i_k} \cdots D_{i_2} D_{i_1} f: U \to \mathbb{R} \text{ partiell differenzierbar}$$

sind.

(Eine partiell differenzierbare Funktion heißt <u>1-mal</u> partiell differenzierbar)

f heißt <u>k-mal stetig partiell differenzierbar</u> $(k \geq 1)$, wenn sie k-mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $\leq k$ stetig sind.

Im Allgemeinen gilt es nicht dass

$$D_j \, D_i f = D_i D_j f$$

d.h.! Die Reihenfolge der Differentiation ist wichtig.

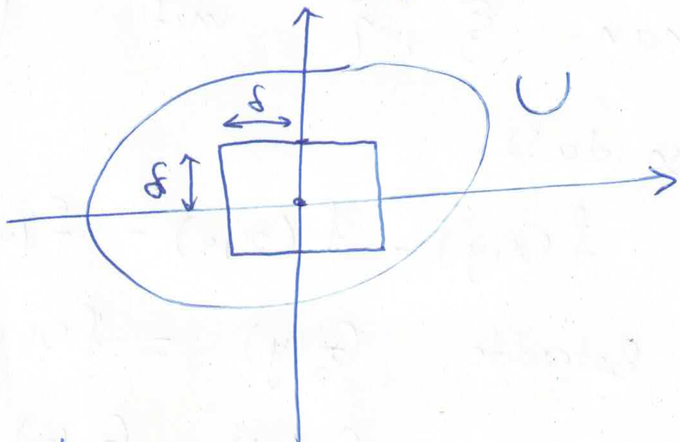
Aber wir haben

Satz 9.2.1 Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig partiell differenzierbare Funktion. Dann gilt $\forall a \in U$, $\forall 1 \leq i \leq n$, $1 \leq j \leq n$

$$D_j \, D_i f(a) = D_i D_j f(a).$$

Beweis Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $i = 1$, $j = 2$, $n = 2$, $a = 0$. Statt (x_1, x_2) schreiben wir zur Vereinfachung (x, y). Sei $\delta > 0$ mit

$$\{ (x,y) \in \mathbb{R}^2 : |x| < \delta, \, |y| < \delta \} \subset U$$



Für festes $|y| < \delta$ sei

$F_y : (-\delta, \delta) \to \mathbb{R}$ gegeben

durch $F_y(x) := f(x,y) - f(x,0)$

Nach dem Mittelwertsatz (Satz 5.2.6)

$\exists\ \xi \in (-|x|, |x|)$ mit

$$F_y(x) - F_y(0) = F_y'(\xi)\, x$$

Beachte $\quad F_y'(\xi) = D_1 f(\xi, y) - D_1 f(\xi, 0).$

Der Mittelwertsatz, angewendet auf die Funktion $y \longmapsto D_1 f(\xi, y)$ liefert ein η mit $|\eta| \leq |y|$, so dass

$$D_1 f(\xi, y) - D_1 f(\xi, 0) = D_2 D_1 f(\xi, \eta)\, y.$$

Folglich

$$f(x,y) - f(x,0) - f(0,y) + f(0,0)$$

$$= F_y(x) - F_y(0)$$

$$= F_y'(\xi)\, x = D_2 D_1 f(\xi, \eta)\, xy \qquad (1)$$

Unter Vertauschung der Rollen von x und y bekommt man $\tilde{\xi}, \tilde{\eta}$ mit $|\tilde{\xi}| \leq |x|$, $|\tilde{\eta}| \leq |y|$

so dass

$$f(x,y) - f(x,0) - f(0,y) + f(0,0) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta})\, xy \qquad (2)$$

[Betrachte $G_x(y) := f(x,y) - f(0,y)$

$$G_x(y) - G_x(0) = G_x'(\tilde{\eta})\, y \quad \text{mit } |\tilde{\eta}| \leq |y|$$

$$G_x'(\tilde{\eta}) = D_2 f(x, \tilde{\eta}) - D_2 f(0, \tilde{\eta})$$

$$= D_1 D_2 f(\tilde{\xi}, \tilde{\eta})\, x \quad \text{mit } |\tilde{\xi}| \leq |x| \ ]$$

Aus (1) und (2) folgt für $xy \neq 0$

$$D_2 D_1 f(\xi, \eta) = D_1 D_2 (\tilde{\xi}, \tilde{\eta})$$

wobei (ξ, η), $(\tilde{\xi}, \tilde{\eta})$ natürlich von (x, y) abhängen mit

$$|\xi| \leq |x|, \quad |\tilde{\xi}| \leq x$$
$$|\eta| \leq |y|, \quad |\tilde{\eta}| \leq y$$

Lässt man nun (x, y) gegen $(0,0)$ streben, so gilt auch $(\xi, \eta) \to 0$ und $(\tilde{\xi}, \tilde{\eta}) \to 0$. Aus der Stetigkeit von $D_1 D_2 f$ und $D_2 D_1 f$ folgt

$$D_2 D_1 f(0,0) = D_1 D_2 f(0,0). \quad \blacksquare$$

<u>Satz 9.2.2</u> Sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine k-mal stetig partiell differenzierbare Funktion $(k \geq 1)$. Dann gilt

$$D_{i_k} \cdots D_{i_2} D_{i_1} f = D_{i_{\sigma(k)}} \cdots D_{i_{\sigma(2)}} D_{i_{\sigma(1)}} f$$

für alle $i_1, i_2, \ldots i_k \in \{1, 2, \ldots, n\}$ und jede Permutation σ der Zahlen $1, 2, \ldots, k$.

z.B. f 3-mal stetig partiell differenzierbare Funktion. Dann gilt

$$D_1 D_2 D_3 f = D_2 D_3 D_1 f$$
$$= D_1 D_3 D_2 f = \text{usw.}$$

<u>Beweis</u>: wir präsentieren keinen vollständigen Beweis. Der Beweis erfolgt durch vollständige Induktion über k mithilfe der Tatsache, dass sich jede Permutation

Zusatzblatt zur Seite 332

Rechenregeln für partielle Ableitungen.

Sei $U \subset \mathbb{R}^n$ offen, seien $1 \leq i \leq n$, und

$f, g: U \to \mathbb{R}$ partiell differenzierbar in der i-ten Koordinatenrichtung in jedem $x \in U$. Dann gilt:

1) $$D_i(f+g) = D_i f + D_i g$$

2) $$D_i(f \cdot g) = (D_i f)\, g + f\,(D_i g)$$

3) $$D_i(\lambda f) = \lambda\, D_i f \quad \forall \lambda \in \mathbb{R}.$$

4) $$D_i\left(\frac{f}{g}\right) = \frac{(D_i f)\, g - f\, D_i g}{g^2}$$

falls $g(x) \neq 0 \;\; \forall x \in U$.

5) Sei $W \subset \mathbb{R}$ offen mit $f(U) \subset W$
$h: W \to \mathbb{R}$ differenzierbar.

Dann ist $h \circ f$ partiell differenzierbar in der i-ten Koordinatenrichtung in jedem $x \in U$ und es gilt

$$D_i(h \circ f)(x) = h'(f(x))\, D_i f(x) \quad \forall x \in U.$$

$$\left(\begin{tikzcd} U \arrow[r, "f"] \arrow[rr, bend right, "h \circ f"'] & f(U) \subset W \arrow[r, "h"] & \mathbb{R} \end{tikzcd} \right)$$

aus Vertauschungen benachbarter Glieder zusammensetzen lässt. ■

<u>Schreibweise</u> Man verwendet auch die Schreibweisen

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = D_j D_i f, \quad \frac{\partial^2 f}{\partial x_i^2} = D_i^2 f = D_i D_i f$$

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}} = D_{i_k} D_{i_{k-1}} \cdots D_{i_1} f \quad \text{usw.}$$

<u>Beispiel 9.2.3</u> Sei $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ zweimal stetig partiell differenzierbare Funktion. Man setzt

$$\Delta f := \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2}$$

Man nennt $\Delta := \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$ den Laplace-Operator, welcher von großen ~~Bedeu~~ Bedeutung in der Mathematik sowie Physik ist: z.B. die Funktionen f, die die Gleichung $\Delta f = 0$ erfüllt, heißen harmonische Funktionen.

Man untersucht den Operator Δ tiefer in der Vorlesung „Differentialgleichungen".

§ 10. Totale Differenzierbarkeit

§ 10.1. Definition und Kettenregel

Wir beginnen mit einer Bemerkung über die Differenzierbarkeit in einer Variable.

<u>Satz 10.1.1</u> Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f: I \to \mathbb{R}$, $x_0 \in I$. Die folgende Aussage sind äquivalent:

(i) f ist differenzierbar in x_0.

(ii) $\exists\, \lambda \in \mathbb{R}$, $\varepsilon > 0$ und eine Funktion $\rho: (-\varepsilon, \varepsilon) \to \mathbb{R}$ mit $\lim\limits_{h \to 0} \frac{\rho(h)}{h} = 0$ und es gilt

$$f(x_0+h) = f(x_0) + \lambda h + \rho(h) \;\; \forall h \in (-\varepsilon, \varepsilon).$$

<u>Beweis</u>:

(i) $\Rightarrow$ (ii): Setze $\lambda := f'(x_0)$ und $\rho(h) := f(x_0+h) - f(x_0) - \lambda h$

für $h \in (-\varepsilon, \varepsilon)$, wobei $\varepsilon > 0$ hinreichend klein so dass $(x_0 - \varepsilon, x_0 + \varepsilon) \subset I$. Beachte

$$\lim_{h\to 0} \frac{\rho(h)}{h} = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - \lambda h}{h} = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} - f'(x_0) = 0.$$

(ii) $\Rightarrow$ (i): Nach Voraussetzung auf $\rho(h)$ gilt

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = \lambda + \lim_{h\to 0}\frac{\rho(h)}{h} = \lambda$$

Daher $f'(x_0)=\lambda$. ■

Bemerkung 10.1.2

(i) es gilt

$$\lim_{h\to 0}\frac{\rho(h)}{h}=0 \iff \lim_{h\to 0}\frac{\rho(h)}{|h|}=0$$

(ii) Sei $A: \mathbb{R}\to\mathbb{R}$, $h\mapsto \lambda h$ lineare Abbildung.

Dann kann die Aussage (ii) des Satze 10.1.1 sich umformulieren wie:

(ii)' $\exists$ eine lineare Abbildung $A:\mathbb{R}\to\mathbb{R}$ und $\varepsilon>0$, $\rho: (-\varepsilon,\varepsilon)\to\mathbb{R}$ mit $\lim_{h\to 0}\frac{\rho(h)}{|h|}=0$

und $f(x_0+h) = f(x_0) + Ah + \rho(h)$

(man schreibt Ah für den Wert $A(h)$)

Definition 10.1.3 Sei $U \subset \mathbb{R}^n$ offen

$f: U \to \mathbb{R}^m$ eine Abbildung.

f heißt im Punkt $x \in U$ total differenzierbar (oder differenzierbar schlechthin) falls es eine lineare Abbildung $A: \mathbb{R}^n \to \mathbb{R}^m$ gibt, so dass

$$f(x+\xi) = f(x) + A\xi + \rho(\xi), \quad (*)$$

wobei ρ eine in eine Umgebung von $0 \in \mathbb{R}^n$ definierte Funktion mit Werten in $\mathbb{R}^m$ ist mit

$$\lim_{\xi \to 0} \frac{\rho(\xi)}{\|\xi\|} = 0 \quad (**)$$

(Statt $(**)$ kann man auch schreiben

$$\rho(\xi) = o(\|\xi\|)$$
)

Bemerkung: (i) Wenn $n = m = 1$, liefert dies die übliche Definition der Differenzierbarkeit (Bemerkung 10.1.2)

(ii) Wir sehen später dass A eindeutig bestimmt ist. Man nennt A das Differential von f im x. Schreibweise: $Df(x) := A$.

Wenn $n = m = 1$, ist das Differential $Df(x)$ die lineare Abbildung
$$\mathbb{R} \to \mathbb{R}$$
$$\xi \mapsto f'(x) \cdot \xi .$$

iii) Die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ kann bzgl. der kanonischen Basen von $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ durch eine $m \times n$-Matrix
$$(a_{ij}) \in M(m \times n, \mathbb{R}), \quad 1 \le i \le m, \quad 1 \le j \le n$$
beschrieben werden. Nämlich hat man
$$A : \mathbb{R}^n \to \mathbb{R}^m$$
$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Im Folgenden identifizieren wir die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ mit der sie beschreibenden Matrix.

Seien
$$f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} \quad \text{und} \quad \varphi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \\ \varphi_m \end{bmatrix}$$
die Komponenten-Darstellungen von f und φ. Dann schreibt sich die Gleichung $(*)$ ausführlich als
$$f_i(x + \xi) = f_i(x) + \sum_{j=1}^{n} a_{ij} \xi_j + \varphi_i(\xi)$$
für $i = 1, 2, \ldots, m$, $\xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$. Daraus folgt:

$f : U \to \mathbb{R}^m$ ist im x differenzierbar $\Leftrightarrow$ $f_i : U \to \mathbb{R}$ ist im x differenzierbar $\forall\, 1 \le i \le m$.

(Man beachtet, dass die Abbildung $: \mathbb{R}^n \to \mathbb{R}$

$$\xi \mapsto \sum_{j=1}^{n} a_{ij} \cdot \xi_j$$

lineare ist)

IV) Sei $A_0 : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung.

Da
$$A_0(x+\xi) = A_0(x) + A_0(\xi) \quad \forall x \in \mathbb{R}^n,\ \xi \in \mathbb{R}^n,$$

ist A_0 differenzierbar in jedem $x \in \mathbb{R}^n$ und

$$DA_0(x) = A_0 \quad \forall x \in \mathbb{R}^n.$$

(Die Funktion ρ in der Gleichung (*) ist identisch gleich 0 in ~~diese~~ diesem Fall).

Satz 10.1.4 Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Abbildung, die im $x \in U$ differenzierbar ist, und zwar gelte

$$f(x+\xi) = f(x) + A\xi + o(\|\xi\|)$$

(wobei $o(\|\xi\|)$ eine Funktion ρ von einer Umgebung von $0 \in \mathbb{R}^n$ nach $\mathbb{R}^m$ bezeichnet mit

$$\lim_{h \to 0} \frac{\rho(\xi)}{\|\xi\|} = 0 \quad)$$

mit der Matrix $A = (a_{ij}) \in M(m \times n, \mathbb{R})$. Dann gilt:

i) f ist im x stetig.

ii) Alle Komponenten $f_i : U \to \mathbb{R}$ $(1 \le i \le m)$ von f ~~ist~~

sind in x partiell differenzierbar mit

$$\frac{\partial f_i}{\partial x_j}(x) = a_{ij}.$$

Bemerkung 10.1.5: Aus ii) sieht man, dass die Matrix A (das Differential $Df(x)$) eindeutig bestimmt ist. Die Matrix von $Df(x)$ nennt man Funktional-Matrix oder auch Jacobi-Matrix von f im Punkt x. Schreibweise

$$J_f(x) := \frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)}(x) := \left(\frac{\partial f_i}{\partial x_j}(x) \right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

(wie vorher identifizieren wir oft $Df(x)$ mit $J_f(x)$).

Die i-te Zeile der Jacobi-Matrix ist der Gradient der Funktion f_i:

$$\left(\frac{\partial f_i}{\partial x_1}(x), \frac{\partial f_i}{\partial x_2}(x), \dots, \frac{\partial f_i}{\partial x_n}(x) \right) = \operatorname{grad} f_i(x)$$

Im Fall $m = 1$ (d.h. $f: U \to \mathbb{R}$) ist also $J_f(x)$ einfach der Gradient von f in x.

Beweis vom Satz 10.1.4

i) da $\lim\limits_{\xi \to 0} A\xi = 0$ und $\lim\limits_{\xi \to 0} o(\|\xi\|) = 0$
(Bemerkung 8.3.23 (iii))
ist f im x stetig.

ii) $\forall\ 1 \leq i \leq m$ gilt

$$f_i(x+\xi) = f_i(x) + \sum_{j=1}^{m} a_{ij}\xi_j + \rho_i(\xi),$$

wobei $\lim\limits_{\xi\to 0} \frac{\rho_i(\xi)}{\|\xi\|} = 0$.

Folglich

$$f_i(x+he_j) = f_i(x) + h\,a_{ij} + \rho_i(he_j)$$

$$\left(\xi := he_j \Rightarrow \xi_{j'} = 0 \ \forall\, j' \neq j,\ \xi_j = 1 \right)$$

Daher erhält man

$$\frac{\partial f_i}{\partial x_j}(x) = \lim_{h\to 0} \frac{f_i(x+he_j) - f_i(x)}{h} = a_{ij} + \lim_{h\to 0} \frac{\rho_i(he_j)}{h}$$

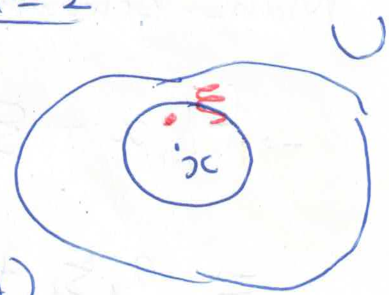
$$= a_{ij}. \quad \blacksquare$$

<u>Satz 10.1.6</u> Sei $U \subset \mathbb{R}^n$ offen

$f: U \to \mathbb{R}$ partiell differenzierbar.

Alle partiellen Ableitungen $D_k f$ seien im Punkt $x \in U$ stetig. Dann ist f in x total differenzierbar.

<u>Beweis</u> wir zeigen die Aussage für <u>$n=2$</u>



Da U offen ist, $\exists\ \delta > 0$ so dass

$B_\delta(x) \subset U$. Sei $\xi \in B_\delta(x)$

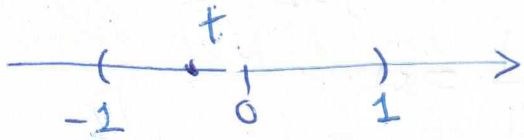
Definiere $z^{(0)} := x$ (x fest)

$z^{(1)} := x + \xi_1 e_1$

$z^{(2)} := x + \xi_1 e_1 + \xi_2 e_2 = x + \xi$

(bemerkte, $\xi_1 e_1 + \xi_2 e_2 = (\xi_1, \xi_2)$)

Setze $g(t) := f(x + t\xi_1 e_1)$, $t \in (-1, 1)$.



Nach dem Mittelwertsatz für differenzierbare Funktionen einer ~~Vab~~ Variable $\exists\ \theta_1 \in [0,1)$ so dass

$$g(1) - g(0) = g'(\theta_1)$$
$$= D_1 f(y_1)\, \xi_1$$

wobei $y_1 := x + \theta_1 \xi_1 e_1$

$$\Rightarrow \quad f(x + \xi_1 e_1) - f(x) = D_1 f(y_1)\, \xi_1$$

Analog $\exists\ \theta_2 \in [0,1]$ mit

$$f(x + \xi_1 e_1 + \xi_2 e_2) - f(x + \xi_1 e_1)$$
$$= D_2 f(y_2)\, \xi_2$$

wobei $y_2 = x + \xi_1 e_1 + \theta_2 \xi_2 e_2$ für ein $\theta_2 \in [0,1]$.

Man bekommt

$$f(x + \xi) - f(x) = D_1 f(y_1)\xi_1 + D_2 f(y_2)\xi_2$$
$$= a_1 \xi_1 + a_2 \xi_2 + (D_1 f(y_1) - a_1)\xi_1 + (D_2 f(y_2) - a_2)\xi_2$$

wobei $a_1 := D_1 f(x)$, $a_2 := D_2 f(x)$.

Setze $\rho(\xi) := \left(D_1 f(y_1) - a_1\right)\xi_1 + \left(D_2 f(y_2) - a_2\right)\xi_2$

Wenn $\xi \to 0$, streben y_1, y_2 gegen x.

Daraus folgt, nach Stetigkeit der $D_1 f$, $D_2 f$ in x,

dass $$\lim_{\xi \to 0} \frac{\rho(\xi)}{\|\xi\|} = 0$$

(Man schätzt ab:

$$\left| \frac{\rho(\xi)}{\|\xi\|} \right| \leq |D_1 f(y_1) - a_1| \frac{|\xi_1|}{\|\xi\|} + |D_2 f(y_2) - a_2| \frac{|\xi_2|}{\|\xi\|}$$

$$\leq \underbrace{|D_1 f(y_1) - a_1| + |D_2 f(y_2) - a_2|}_{\downarrow\ \xi \to 0 \atop 0} \quad)$$

Zusätzlich gilt

$$f(x+\xi) = f(x) + a_1 \xi_1 + a_2 \xi_2 + \rho(\xi)$$

$\Rightarrow$ f differenzierbar in x; und

$$Df(x)(\xi) = a_1 \xi_1 + a_2 \xi_2 .$$

Der allgemeine Fall wird ähnlich bewiesen. Anstelle von $z^{(0)}, z^{(1)}, z^{(2)}$ betrachtet man

$$z^{(k)} = x + \sum_{j=1}^{k} \xi_j e_j \quad , \quad k = 0, 1, \dots, n.$$

∎

Folgerung 10.1.7 Sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$ stetig partiell differenzierbar. Dann ist f in U stetig.

Beweis: Aus Satz 10.1.6 und Satz 10.1.4 (i).

Bemerkung 10.1.8 $\forall f: U \to \mathbb{R}$ gilt:

$$f \text{ stetig partiell differenzierbar in } U$$
$$\Downarrow$$
$$f \text{ total differenzierbar in jedem } x \in U$$
$$\Downarrow$$
$$f \text{ partiell differenzierbar in } U.$$

Die Umkehrungen gelten im Allgemeinen nicht.

Wegen dieser Zusammenhänge nennt man eine stetige partiell differenzierbare Funktion kurz stetig differenzierbar.

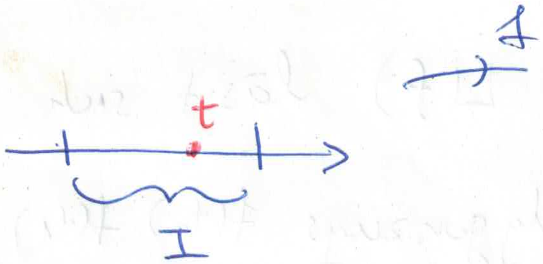
Definition 10.1.9 Sei $I \subset \mathbb{R}$ ein Intervall

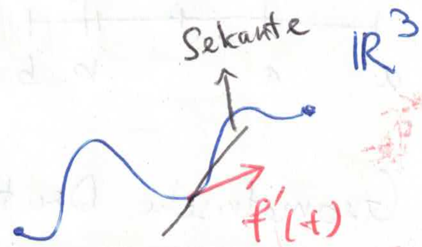
$f : I \to \mathbb{R}^m$ eine Abbildung.

Wir sagen, f ist eine Kurve.

Schreiben wir $f = (f_1, f_2, \ldots, f_m)$. Dann $f_i : I \to \mathbb{R}$ für $1 \le i \le m$.

f heißt differenzierbar (bzw. stetig differenzierbar) wenn alle f_i $(1 \le i \le n)$ differenzierbar (bzw. stetig differenzierbar) sind.





In diesem Fall heißt $f'(t) := (f_1'(t), \ldots, f_m'(t)) \in \mathbb{R}$ den Tangentialvektor der Kurve f in t. Wie im Fall $m = 1$ lässt ~~sich~~ $f'(t)$ sich als Limes von Sekanten auffassen, denn

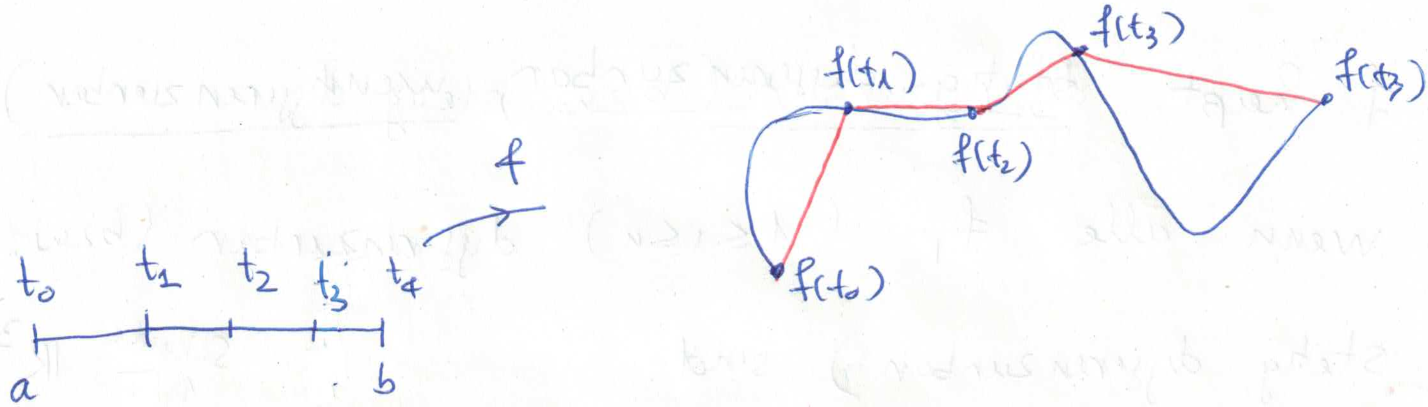
$$f'(t) = \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}.$$

Wenn I offen, sieht man dass differenzierbare Kurven total differenzierbar im Sinn von Def. 10.1.3 sind, und gilt $Df(t) = f'(t)$. (wie Spaltenvektor)

Definition 10.110 Sei $[a,b] \subset \mathbb{R}$, $a<b$.

Sei $f: [a,b] \to \mathbb{R}^m$ eine stetige differenzierbare Kurve. Die (Bogen) Länge von f ist definiert durch

$$L(f) := \int_a^b \| f'(t) \| dt .$$



Geometrische Deutung: $L(f)$ lässt sich als Limes von Länge L_k des Polygonzugs $f(t_0) f(t_1) \dots f(t_k)$

$$L_k := \sum_{i=1}^{k} \| f(t_i) - f(t_{i-1}) \|$$ auffassen

für ~~eine~~ Unterteilung $a = t_0 < t_1 < \dots < t_k = b$ von $[a,b]$ mit $t_1 - t_0 = t_2 - t_1 = \dots = t_k - t_{k-1} =: \delta$ wenn die Feinheit $\delta \to 0$.

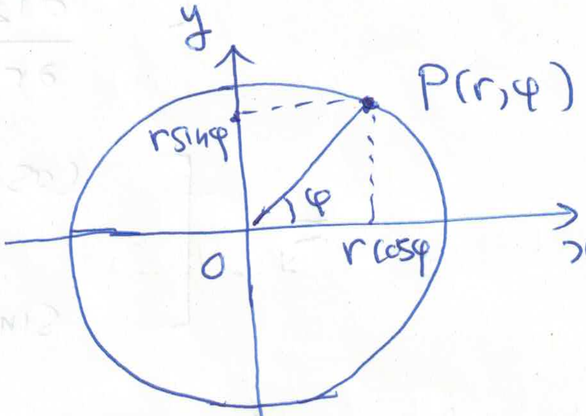
(Für weitere Informationen verweisen wir auf das Kapitel 4 im Forsters Analysis II)

Beispiel 10.1.11 (Polarkoordinaten in $\mathbb{R}^2$)

Erinnere (Satz 5.4.4) dass die Abbildung

$$P : \mathbb{R}_+ \times [0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus \{0\}$$
$$(r, \varphi) \longmapsto (r\cos\varphi, r\sin\varphi)$$

ein Bijektion ist.



Man sagt, dass (r, φ) die <u>Polarkoordinaten</u> auf $\mathbb{R}^2$ sind.

Setze $U := \mathbb{R}_+ \times (0, 2\pi) \subset \mathbb{R}^2$, welche offen in $\mathbb{R}^2$ ist. Von nun an betrachten wir die Einschränkung von P auf U:

$$P : \mathbb{R}_+ \times (0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus \{0\}$$

Man bemerkt,

$$P(U) = \left\{ (r\cos\varphi, r\sin\varphi) : \begin{array}{l} r > 0 \\ \varphi \in (0, 2\pi) \end{array} \right\}$$

$$= \mathbb{R}^2 \setminus \{ (x, y) : y = 0, x \geq 0 \}$$

Schreibe $P = (P_1, P_2)$ Komponenten-Funktionen

$P_1(r, \varphi) = r\cos\varphi$, $P_2(r, \varphi) = r\sin\varphi$, für $(r, \varphi) \in U$. Da $\cos\varphi$, $\sin\varphi$ stetig differenzierbar sind, so sind P_1, P_2 auch stetig differenzierbar.

Daraus folgt ist P stetig differenzierbar (Bemerkung 10.1.8) und die Jacobi-Matrix von P im (r,φ) lautet

$$J_P(r,\varphi) = \begin{bmatrix} \frac{\partial P_1}{\partial r}(r,\varphi) & \frac{\partial P_1}{\partial \varphi}(r,\varphi) \\ \frac{\partial P_2}{\partial r}(r,\varphi) & \frac{\partial P_2}{\partial \varphi}(r,\varphi) \end{bmatrix}$$

$$= \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}$$

Wir berechnen die Determinante der Matrix $J_P(r,\varphi)$:

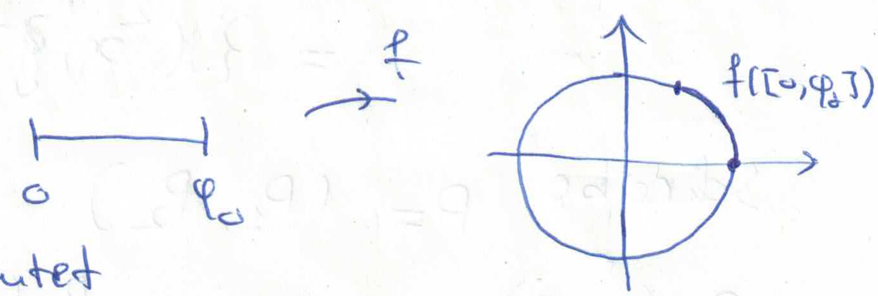
$$\det J_P(r,\varphi) = \cos\varphi\,(r\cos\varphi) - (\sin\varphi)(-r\sin\varphi)$$
$$= r\,(\cos^2\varphi + \sin^2\varphi) = r \neq 0$$

für $(r,\varphi) \in U$. Sei $\varphi_0 \in (0,2\pi)$.

Betrachte die Kurve $f: [0,\varphi_0] \to \mathbb{R}^2$

$$\varphi \longmapsto (\cos\varphi, \sin\varphi)$$

$f(\varphi) = P(1,\varphi)$



Die Länge von f lautet

$$L(f) = \int_0^{\varphi_0} \|f'(\varphi)\|\,d\varphi = \int_0^{\varphi_0} \sqrt{(-\sin\varphi)^2 + \cos^2\varphi}\,d\varphi$$
$$= \int_0^{\varphi_0} d\varphi = \varphi_0 .$$

Folglich ist die Länge der Bogen $f([0, \varphi_0])$ gleich genau φ_0. Die Tatsache haben wir in der Analysis I (Seite 218 im Skript) ohne Beweis akzeptiert.

Wir untersuchen nun allgemeine Eigenschaften von differenzierbaren Funktionen.

Satz 10.1.12 (Kettenregel)

Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offene Mengen und

$$f : U \to \mathbb{R}^m, \quad g : V \to \mathbb{R}^k$$

Abbildungen mit $f(U) \subset V$. Die Abbildung f sei im Punkt $x \in U$ differenzierbar und die Abbildung g im Punkt $y := f(x) \in V$ differenzierbar. Dann ist die zusammengesetzte Abbildung $g \circ f : U \to \mathbb{R}^k$ im x differenzierbar und es gilt

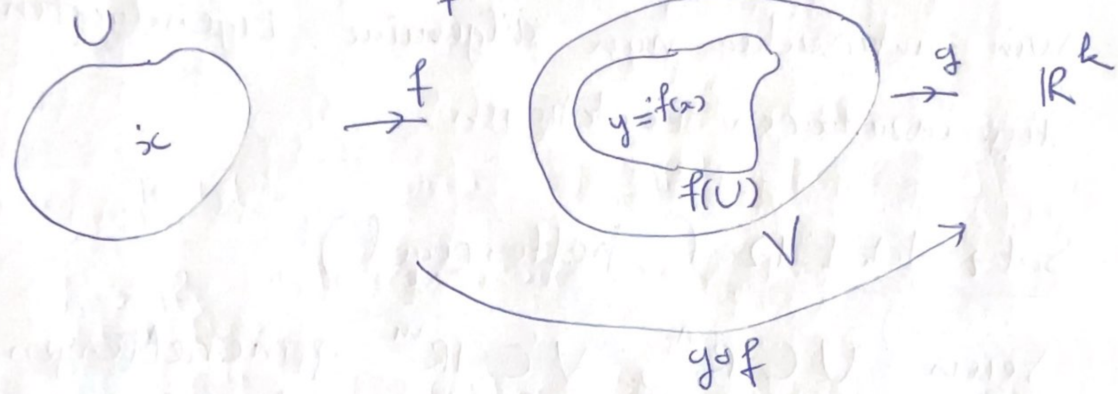
$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$$

$$J_{g \circ f}(x) = J_g(f(x)) \cdot J_f(x)$$

(Multiplikation von Matrizen)

Wenn $n = m = k = 1$ findet man wieder die übliche Kettenregel für Funktionen einer Variable (in diesem Fall sind Jacobi-Matrizen nichts anderes als reelle Zahlen)

<u>Beweis</u> Sei $A := J_f(x)$, $B := J_g(y)$



Wir zeigen, $J_{g \circ f}(x) = BA$.

Nach Definition der totalen Differenzierbarkeit gilt

$$f(x+\xi) = f(x) + A\xi + \varphi(\xi)$$

$$g(y+\eta) = g(y) + B\eta + \psi(\eta)$$

mit $\varphi(\xi) = o(\|\xi\|)$ $\left(\text{d.h. } \lim_{\xi \to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0\right.$
$\psi(\eta) = o(\|\eta\|)$ $\left. \lim_{\eta \to 0} \frac{\psi(\eta)}{\|\eta\|} = 0 \right)$

für $\forall \xi$ in einer Umgebung von $0 \in \mathbb{R}^n$
$\forall \eta$ in einer Umgebung von 0 in $\mathbb{R}^m$

Insbesondere $\exists\ \delta > 0\ \ \forall \xi$ mit $\|\xi\| \le \delta$ gilt

$$\|\varphi(\xi)\| \le \|\xi\|$$

Durch Ersetzen $\eta = f(x+\xi) - f(x)$
$= A\xi + \varphi(\xi)$

bekommt man

$$g\circ f(x+\xi) = g(f(x+\xi)) = g(f(x)) + B(A\xi + \varphi(\xi)) + \psi(A\xi + \varphi(\xi))$$

$$= g(f(x)) + BA\xi + B\varphi(\xi) + \psi(A\xi + \varphi(\xi))$$

$$= g\circ f(x) + BA\xi + \chi(\xi) \qquad (*)$$

mit $\chi(\xi) := B\varphi(\xi) + \psi(A\xi + \varphi(\xi))$

Zusätzlich gilt

$$\|B\varphi(\xi)\| \leq \|B\|\,\|\varphi(\xi)\|$$ ($\|B\|$ Norm von B: Bemerkung 83.23)

und für $\|\xi\| < \delta$:

$$\|A\xi + \varphi(\xi)\| \leq \|A\xi\| + \|\varphi(\xi)\|$$
$$\leq \|A\|\,\|\xi\| + \|\varphi(\xi)\| \leq (\|A\|+1)\,\|\xi\|$$

Nach Voraussetzung auf φ, ψ erhält man

$$\lim_{\xi \to 0} \frac{\|B\varphi(\xi)\|}{\|\xi\|} = 0 \quad \text{und} \quad \psi(\eta) = \|\eta\|\,\psi_1(\eta)$$

mit $\lim\limits_{\eta \to 0} \psi_1(\eta) = 0$. Daher gilt

$$\|\psi(A\xi + \varphi(\xi))\| = \|A\xi + \varphi(\xi)\|\,\|\psi_1(A\xi + \varphi(\xi))\|$$
$$\leq (\|A\|\,\|\xi\| + \|\varphi(\xi)\|)\,\|\psi_1(A\xi + \varphi(\xi))\|$$
$$\leq (\|A\|+1)\,\|\xi\| \cdot \|\psi_1(A\xi + \varphi(\xi))\|$$

$$\Rightarrow \lim_{\xi \to 0} \frac{\| \psi(A\xi + \varphi(\xi)) \|}{\|\xi\|} = 0 .$$

Also $$\lim_{\xi \to 0} \frac{\|\chi(\xi)\|}{\|\xi\|} = 0 .$$

Dies und (*) implizieren: $g \circ f$ ist differenzierbar in x

und $J_{g\circ f(x)} = BA$. ∎

<u>Folgerung 10.1.13</u> Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen,

$f: V \to \mathbb{R}$, $x \mapsto f(x)$ differenzierbar und

$\varphi: U \to \mathbb{R}^m$, $t \mapsto \varphi(t)$ eine differenzierbare Abbildung mit $\varphi(U) \subset V$.

Dann ist die Funktion

$$F := f\circ\varphi : U \to \mathbb{R}$$
$$t \mapsto f(\varphi(t))$$

partiell differenzierbar und es gilt für $i = 1, 2, \dots, n$

$$\frac{\partial F}{\partial t_i}(t_1, \dots, t_n) = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(\varphi_1(t), \dots, \varphi_m(t)) \frac{\partial \varphi_j}{\partial t_i}(t_1, \dots, t_n)$$

wobei $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_m)$, $t = (t_1, \dots, t_m)$
$x = (x_1, \dots, x_m)$.

<u>Beweis</u>: Nach Satz 10.1.12 ist F differenzierbar und $J_F(t) = J_f(\varphi(t))\, J_\varphi(t)$. (*)

Das impliziert: F ist partiell differenzierbar.

Wir berechnen

$$J_F(t) = \left(\frac{\partial F}{\partial t_1}(t), \ldots, \frac{\partial F}{\partial t_n}(t) \right) \in \mathbb{R}^n$$

$$J_f(\varphi(t)) = \left(\frac{\partial f}{\partial x_1}(\varphi(t)), \ldots, \frac{\partial f}{\partial x_m}(\varphi(t)) \right) \in \mathbb{R}^m$$

(wir stellen $J_F(t)$, $J_f(\varphi(t))$ als Vektoren dar)

$$J_\varphi(t) = \begin{bmatrix} \frac{\partial \varphi_1}{\partial t_1}(t) & \cdots & \frac{\partial \varphi_1}{\partial t_n}(t) \\ \vdots & & \\ \frac{\partial \varphi_m}{\partial t_1}(t) & \cdots & \frac{\partial \varphi_m}{\partial t_n}(t) \end{bmatrix} \quad (m \times n)\text{-Matrix}$$

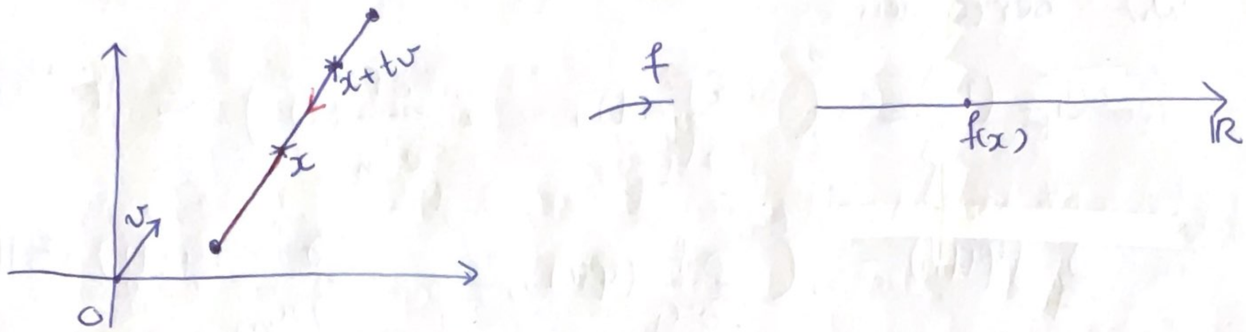
Die Behauptung ergibt sich deshalb durch Matrix-Multiplikation aus (*). ∎

<u>Definition 10.1.14</u> (Richtungsableitung) Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine Funktion. Weiter sei $x \in U$ und $v \in \mathbb{R}^n$ mit $\|v\| = 1$. Unter der <u>Richtungsableitung von f in x in Richtung v</u> versteht man (im Fall der Existenz) den Limes

$$D_v f(x) := \frac{d}{dt} f(x+tv) \Big|_{t=0} = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

Für $v = e_i$ ist also $D_{e_i} f$ gleich der i-ten partiellen

Ableitung $D_v f$.



Sei $\varphi : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$

$t \mapsto x + tv = (x_1 + tv_1, \ldots, x_n + tv_n)$

wobei $\varepsilon > 0$ hinreichend klein mit $\varphi((-\varepsilon, \varepsilon)) \subset U$.

Betrachte $F := f \circ \varphi : (-\varepsilon, \varepsilon) \to \mathbb{R}$.

Dann
$$D_v f(x) = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$
$$= \lim_{t \to 0} \frac{f(\varphi(t)) - f(\varphi(0))}{t} = F'(0)$$

(Im Fall der Existenz) Das bedeutet:
die Richtungsableitung lässt sich beschreiben als die gewöhnliche Ableitung einer Funktion einer Variable.

<u>Satz 10.1.15</u> Sei $U \subset \mathbb{R}^n$ offen
$f : U \to \mathbb{R}$ differenzierbar.

Dann $\forall\ v \in \mathbb{R}^n$ mit $\|v\| = 1$, $x \in U$ gilt

$$D_v f(x) = \langle v, \operatorname{grad} f(x) \rangle.$$

Beweis: Sei φ, F wie oben.

Dann bekommt man

$$D_v f(x) = F'(0)$$

Aus Folgerung 10.1.13 (angewendet auf $n=1$) gilt

$$F'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\varphi(t)) \cdot \frac{\partial \varphi_i}{\partial t}(t)$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\varphi(t)) \cdot v_i = \langle v, \operatorname{grad} f(\varphi(t)) \rangle$$

(da $\varphi_i(t) = x_i + t v_i \Rightarrow \varphi_i'(t) = v_i$)

Ersetzen $t=0$, erhalten wir

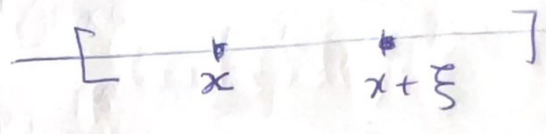
$$F'(0) = \langle v, \operatorname{grad} f(x) \rangle \qquad \blacksquare$$

§ 10.2. Mittelwertsatz

Diskussion in Dimension 1

Sei $f: I \to \mathbb{R}$ eine stetige differenzierbare Funktion, wobei $I \subset \mathbb{R}$ ein Intervall $\neq \emptyset$.

Seien $x, x+\xi \in I$



Aus dem Fundamentalsatz der Differential- und Integralrechnung folgt:

$$f(x+\xi) - f(x) = \int_x^{x+\xi} f'(u)\,du$$

$$= \int_0^1 f'(x+t\xi)\,\xi\,dt$$

(Durch Substitution (Satz 6.2.5) $\varphi(t) := x + t\xi$, $\varphi'(t) = \xi$, $t \in [0,1]$

$$\int_0^1 f'(\varphi(t))\,\varphi'(t)\,dt = \int_x^{x+\xi} f'(u)\,du$$)

$$= \left(\int_0^1 f'(x+t\xi)\,dt\right)\xi. \quad (*)$$

Diese Formel impliziert den Mittelwertsatz für Differentialrechnung (Satz 5.2.6). Denn nach

Satz 6.1.13 (Mittelwertsatz ~~zu~~ der Integralrechnung)

$\exists$ $t_0 \in [0,1]$ mit

$$f'(x+t_0\xi) = \int_0^1 f'(x+t\xi)\,dt$$

Daher erhält man

$$f(x+\xi) - f(x) = f'(x+t_0\xi)\,\xi.$$

Dies ist die übliche Fassung des Mittelwertsatzes.

Wir betrachten nun eine Verallgemeinerung des Mittelwertsatzes, welche ähnlich wie (*) ist.

<u>Definition 10.2.1</u> Sei $A = [a_{ij}]_{\substack{1 \le i \le m \\ 1 \le j \le n}}$

eine Matrix, deren Koeffizienten a_{ij} stetige Funktionen auf dem Intervall $[t_0, t_1] \subset \mathbb{R}$ seien. Dann versteht man unter dem Integral

$$\int_{t_0}^{t_1} A(t)\,dt$$

die Matrix mit den Koeffizienten $\int_{t_0}^{t_1} a_{ij}(t)\,dt$

$1 \le i \le m$, $1 \le j \le n$.

<u>Satz 10.2.2</u> (Mittelwertsatz)

Sei $U \subset \mathbb{R}^n$ offen und

$f: U \to \mathbb{R}^m$ stetig differenzierbar. Sei $x \in U$

und $\xi \in \mathbb{R}^n$ ein Vektor derart, dass die ganze Strecke $x + t\xi$, $0 \leq t \leq 1$, in U liegt. Dann gilt

$$f(x+\xi) - f(x) = \underbrace{\left(\int_0^1 J_f(x+t\xi)\,dt\right)}_{(m \times n)\text{- Matrix}} \cdot \underbrace{\xi}_{\text{Vektor in } \mathbb{R}^n}$$

<u>Beweis</u> Es seien $f_i : U \to \mathbb{R}$ die Komponenten von f $1 \leq i \leq m$. Wir betrachten

$$g_i : [0,1] \to \mathbb{R}$$
$$t \longmapsto g_i(t) = f_i(x + t\xi)$$

Es gilt

$$f_i(x+\xi) - f_i(x) = g_i(1) - g_i(0) = \int_0^1 g_i'(t)\,dt$$
$$= \int_0^1 \left(\sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x+t\xi)\xi_j \right) dt$$

(Kettenregel)

$$= \sum_{j=1}^n \left(\int_0^1 \frac{\partial f_i}{\partial x_j}(x+t\xi)\,dt \right) \xi_j$$

Außerdem da
$$J_f(x+t\xi) = \left[\frac{\partial f_i}{\partial x_j}(x+t\xi) \right]_{\substack{1\le i\le m \\ 1\le j\le n}}$$

gilt
$$\int_0^1 J_f(x+t\xi)\,dt = \left[\int_0^1 \frac{\partial f_i}{\partial x_j}(x+t\xi)\,dt \right]_{\substack{1\le i\le m \\ 1\le j\le n}}$$

Schreibt
$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix} = \left(\int_0^1 J_f(x+t\xi)\,dt \right) \xi$$

Man sieht
$$v_i = \sum_{j=1}^{n} \left(\int_0^1 \frac{\partial f_i}{\partial x_j}(x+t\xi)\,dt \right) \xi_j$$

Daraus folgt $v_i = f_i(x+\xi) - f_i(x)$

Man bekommt daher die Behauptung. ■

<u>Lemma 10.2.3</u>. Sei $v : [a,b] \to \mathbb{R}^m$ eine stetige vektorwertige Funktion auf $[a,b] \subset \mathbb{R}$.

Dann gilt
$$\Big\| \underbrace{\int_a^b v(t)\,dt}_{\text{Vektor in } \mathbb{R}^m} \Big\| \le \int_a^b \|v(t)\|\,dt$$

Beweis: $v = (v_1, v_2, \dots, v_m)$

$$\Rightarrow u := \int_a^b v(t)\,dt = \left(\int_a^b v_1(t)\,dt, \int_a^b v_2(t)\,dt, \dots, \int_a^b v_m(t)\,dt \right) \in \mathbb{R}^m$$

Setze $K := \|u\| = \sqrt{\langle u, u \rangle}$.

Dann gilt

$$K^2 = \langle u, u \rangle = \left\langle \int_a^b v(t)\,dt, u \right\rangle$$

$$= \left(\int_a^b v_1(t)\,dt\right) u_1 + \dots + \left(\int_a^b v_m(t)\,dt\right) u_m$$

$$= \int_a^b (v_1(t)\, u_1)\,dt + \dots + \int_a^b (v_m(t) \cdot u_m)\,dt$$

$$= \int_a^b \langle v(t), u \rangle\,dt$$

$$\leq \int_a^b \|v(t)\|\,\|u\|\,dt \quad \text{(wegen der Cauchy-Schwarz-Ungleichung)}$$

$$= \|u\| \int_a^b \|v(t)\|\,dt$$

$$= K \int_a^b \|v(t)\|\,dt$$

Daher folgt die Behauptung ∎

Folgerung 10.2.4: Seien U, $f: U \to \mathbb{R}^m$ stetig differenzierbar, $x \in U$, $\xi \in \mathbb{R}^m$ so dass die ganze Strecke $x + t\xi$, $0 \leq t \leq 1$, in U liegt. Es sei

$$M := \sup_{0 \leq t \leq 1} \| J_f(x + t\xi) \|$$ Dann gilt

$$\| f(x+\xi) - f(x) \| \leq M \| \xi \| .$$

(Wir identifizieren $J_f(x+t\xi)$ mit $Df(x+t\xi)$ die Norm von $J_f(x+t\xi)$ ist dann die Norm der linearen Abbildung $Df(x+t\xi)$, siehe Definition 8.3.22)

<u>Beweis</u> Satz 10.2.2 impliziert

$$\| f(x+\xi) - f(x) \| = \left\| \left(\int_0^1 J_f(x+t\xi)\, dt \right) \xi \right\|$$

$$= \left\| \int_0^1 \left(J_f(x+t\xi) \cdot \xi \right) dt \right\|$$

$$\leq \int_0^1 \| J_f(x+t\xi) \cdot \xi \| \, dt$$

(Nach Lemma 10.2.3 angewendet auf
$v(\xi) := J_f(x+t\xi).\xi$)

$$\leq \int_0^1 \| J_f(x+t\xi) \| \, \| \xi \| \, dt$$

$$\leq \int_0^1 M \| \xi \| dt = M \| \xi \|$$

∎

§ 11. Taylor-Formel, lokale Extrema

§ 11.1. Taylor-Formel

Ziel: wir verallgemeinern die Taylor-Formel im Kapitel 6 auf den Fall von mehreren Variablen.

Bezeichnungen: Für $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}_0^n$ (man sagt, α ist ein n-Tupel) setzen wir

$$|\alpha| := \alpha_1 + \alpha_2 + \dots + \alpha_n \quad, (\text{dann } |\alpha| \in \mathbb{N}_0)$$

$$\alpha! := \alpha_1!\, \alpha_2! \cdots \alpha_n!$$

Sei $k \in \mathbb{N}_0$. Wenn $k \geq 1$, haben wir k-mal stetig differenzierbare Funktionen definiert.

Wenn $k = 0$: wir sagen, jede Funktion $f: U \to \mathbb{R}$ ($U \subset \mathbb{R}^n$ offen) ist 0-mal differenzierbar und setzen

$$\frac{\partial^0 f}{\partial x_j^0} := f \quad \text{für } 1 \leq j \leq n$$

$$x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

jede stetige Funktion $f: U \to \mathbb{R}$ heißt 0-mal stetig differenzierbar.

Ist f nun eine $|\alpha|$-mal stetig differenzierbare Funktion, so setzt man

$$D^{\alpha} f := D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

wobei $D_i^{\alpha_i} = \underbrace{D_i D_i \cdots D_i}_{\alpha_i\text{-mal}}$

z.B. $n=2$, $\alpha = (1,3)$, $|\alpha| = 4$

$$D^{(1,3)} f = \frac{\partial^4 f}{\partial x_1^1 \partial x_2^3} = \frac{\partial}{\partial x_1}\left(\frac{\partial}{\partial x_2}\left(\frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_2}\right)\right)\right)$$

(Nach Satz 9.2.1 ist die Reihenfolge der Differentiation nicht wichtig in diesem Fall)

$\alpha = (0,1)$

$$D^{\alpha} f = \frac{\partial^1 f}{\partial x_1^0 \partial x_2^1} = \frac{\partial f}{\partial x_2}.$$

Für $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ sei

$$x^{\alpha} := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$

Satz 11.1.1 Sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion ($k \geq 0$). Sei $x \in U$ und $\xi \in \mathbb{R}^n$ ein Vektor derart, dass die Strecke $x + t\xi$, $0 \leq t \leq 1$, ganz in U liegt.

Dann ist die Funktion

$$g : [0,1] \to \mathbb{R}$$
$$t \longmapsto f(x+t\xi)$$

k-mal stetig differenzierbar und es gilt

$$g^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D^\alpha f)(x+t\xi)\xi^\alpha .$$

(unter $\sum_{|\alpha|=k}$ versteht man $\sum_{\alpha \in I}$, wobei

$$I = \{ \alpha \in \mathbb{N}_0^n : |\alpha| = k \} \;)$$

Beweis wir zeigen zunächst durch Induktion über k, dass

$$g^{(k)}(t) = \sum_{i_1, i_2, \dots i_k = 1}^{n} D_{i_k} \cdots D_{i_1} f(x+t\xi)\xi_{i_1}\xi_{i_2} \cdots \xi_{i_k} \quad (*)$$

(wobei: man schreibt $\sum_{i_1, i_2, \dots i_k = 1}^{n}$ kurz für

$\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n}$). Sei $\varphi : [0,1] \to \mathbb{R}^n$, $t \mapsto x+t\xi$

(die Reihenfolge der Summen nach $i_1 \dots i_k$ ist nicht wichtig)

Für $k=1$ ergibt sich aus der Kettenregel (Folgerung 10.1.13)

$$g'(t) = (f \circ \varphi)'(t)$$

$$= \frac{\partial f}{\partial x_1}(\varphi(t)) \frac{\partial \varphi_1}{\partial t}(t) + \dots + \frac{\partial f}{\partial x_n}(\varphi(t)) \frac{\partial \varphi_n}{\partial t}(t)$$

$$= D_1 f(x+t\xi) \cdot \xi_1 + \dots + D_n f(x+t\xi)\,\xi_n$$

(da $\varphi_j(t) = x_j + t\xi_j$, $1 \le j \le n$

$\dfrac{\partial \varphi_j}{\partial t}(t) = \xi_j$)

Induktionsschritt $k-1 \to k$: wir nehmen an, dass

$$g^{(k-1)}(t) = \sum_{i_1,\dots,i_{k-1}=1}^{n} D_{i_{k-1}} \cdots D_{i_1} f(x+t\xi)\, \xi_{i_1} \cdots \xi_{i_{k-1}}$$

Betrachte

$$g^{(k)}(t) = \frac{d}{dt}\left(g^{(k-1)}(t)\right)$$

$$= \frac{d}{dt}\left(\sum_{i_1,\dots,i_{k-1}=1}^{n} D_{i_{k-1}} \cdots D_{i_1} f(x+t\xi)\, \xi_{i_1} \cdots \xi_{i_{k-1}}\right)$$

(nach t)

$$= \sum_{j=1}^{n} D_j \left(\sum_{i_1,\dots,i_{k-1}=1}^{n} D_{i_{k-1}} \cdots D_{i_1} f(x+t\xi)\, \xi_{i_1} \cdots \xi_{i_{k-1}}\right) \xi_j$$

$$= \sum_{i_k=1}^{n} \sum_{i_1,\dots,i_{k-1}=1}^{n} D_{i_k} D_{i_{k-1}} \cdots D_{i_1} f(x+t\xi)\, \xi_{i_1} \cdots \xi_{i_k}$$

(wir haben den Index i_k statt j geschrieben)

$$= \sum_{i_1,\dots,i_k=1}^{n} D_{i_k} D_{i_{k-1}} \cdots D_{i_1} f(x+t\xi)\, \xi_{i_1} \cdots \xi_{i_k}.$$

Daher ist $(*)$ bewiesen.

Kommt unter den Indizes $(i_1,\dots,i_k)$ der Index 1 genau α_1-mal, der Index 2 genau α_2-mal, ..., den Index n genau α_n-mal vor, so nach Satz 9.21

(Verstauschung der Differentiation) gilt

$$D_{i_k} \cdots D_{i_1} f(x+t\xi)\xi_{i_1} \cdots \xi_{i_k} = D_1^{\alpha_1} \cdots D_n^{\alpha_n} f(x+t\xi)\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$$

$$= D^\alpha f(x+t\xi)\xi^\alpha$$

Beachte: $\alpha_1 + \dots + \alpha_n = |\alpha| = k$.

es gilt genau $\dfrac{k!}{\alpha_1! \dots \alpha_n!}$ k-Tupel $(i_1, \dots, i_k)$ von Zahlen

$1 \leq i_l \leq n$, $(1 \leq l \leq k)$, bei denen der Index j genau α_j-mal vorkommt ($j = 1, 2, \dots, n$). (kombinatorische Aufgabe)

Daraus folgt

$$g^{(k)}(t) = \sum_{i_1, \dots, i_k = 1}^{n} D_{i_k} \cdots D_{i_1} f(x+t\xi)\xi_{i_1} \cdots \xi_{i_k}$$

$$= \sum_{|\alpha| = k} \frac{k!}{\alpha_1! \dots \alpha_n!} D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n} f(x+t\xi)\xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$$

(wir gruppieren die gleichen Terme zusammen)

$$= \sum_{|\alpha| = k} \frac{k!}{\alpha!} D^\alpha f(x+t\xi)\xi^\alpha$$

Satz 11.1.2 (Taylorsche Formel)

Sei $U \subset \mathbb{R}^n$ offen, $x \in U$, $\xi \in \mathbb{R}^n$ Vektor derart, dass die Strecke $x + t\xi$, $0 \leq t \leq 1$, ganz in U liegt. Weiter sei $f: U \to \mathbb{R}$ eine $(k+1)$-mal stetig differenzierbare Funktion. Dann $\exists\ \theta \in [0,1]$ so dass

$$f(x+\xi) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha + \sum_{|\alpha| = k+1} \frac{D^\alpha f(x+\theta\xi)}{\alpha!}\xi^\alpha$$

- Wenn $\underline{n=1}$, bekommt man wieder

$$f(x+\xi) = \sum_{\alpha=0}^{k} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha + \frac{D^{k+1} f(x+\theta\xi)}{(k+1)!}\xi^{k+1}$$

(Satz 6.3.5) ($\alpha \in \mathbb{N}_0$, $0 \leq \alpha \leq k$).

- $n=2$, $k=1$ erhält man

$$f(x+\xi) = \sum_{|\alpha|\leq 1} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha + \sum_{|\alpha|=2} \frac{D^\alpha f(x+\theta\xi)}{\alpha!}\xi^\alpha$$

$$= \frac{D^{(0,0)} f(x)}{0!\,0!}\xi_1^0\xi_2^0 + \frac{D^{(1,0)} f(x)}{1!\,0!}\xi_1^1\xi_2^0$$

$$+ \frac{D^{(0,1)}}{0!\,1!} f(x)\,\xi_1^0\xi_2^1 + \sum_{|\alpha|=2} \frac{D^\alpha f(x+\theta\xi)}{\alpha!}\xi^\alpha$$

$$= f(x) + \left(\frac{\partial f(x)}{\partial x_1}\xi_1 + \frac{\partial f}{\partial x_2}(x)\,\xi_2\right)$$

$$+ \sum_{|\alpha|=2} \frac{D^\alpha f(x+\theta\xi)}{\alpha!}\xi^\alpha .$$

($\sum_{|\alpha|\leq k}$ bedeutet die Summe über alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \leq k$)

<u>Beweis von Satz 11.1.2</u>

Sei $g : [0,1] \to \mathbb{R}$

$$t \mapsto g(t) := f(x+t\xi)$$

Nach Satz 11.1.1 ist g $(k+1)$-mal stetig differenzierbar. Taylorsche Formel (Satz 6.3.5) (angewendet auf g) liefert

$\exists\ \theta \in [0,1]$ mit

$$g(1) = \sum_{m=0}^{k} \frac{g^{(m)}(0)}{m!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}.$$

$\parallel$

$f(x+\xi)$

Satz ~~11.1~~. 11.1.1 besagt:

$$\frac{g^{(m)}(0)}{m!} = \sum_{|\alpha|=m} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha$$

und

$$\frac{g^{(k+1)}(\theta)}{(k+1)!} = \sum_{|\alpha|=k+1} \frac{D^\alpha f(x+\theta\xi)}{\alpha!}\xi^\alpha$$

woraus die Behauptung folgt. ∎

<u>Folgerung 11.1.3</u> Sei $U \subset \mathbb{R}^n$ offen und $f: U \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion. Dann gilt $\forall\ x \in U$, $\forall\ \xi$ in einer hinreichend kleinen Umgebung von $0 \in \mathbb{R}^n$;

$$f(x+\xi) = \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha + o(\|\xi\|^k)$$

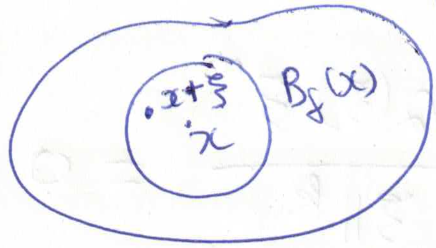
(für $\xi \to 0$)

d.h. $$\lim_{\xi\to 0} \frac{f(x+\xi) - \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!}\xi^\alpha}{\|\xi\|^k} = 0$$

($\sum_{|\alpha|\le k}$ bedeutet die Summe über alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \le k$)

Beweis

$U \subset \mathbb{R}^n$



Da U offen ist, $\exists\ \delta > 0$ so dass $B_\delta(x) \subset U$.

$\Rightarrow \quad x + \xi \in U \quad \forall\ \xi \in \mathbb{R}^n$ mit $\|\xi\| < \delta$.

Nach Satz 11.1.2 gibt es zu jedem $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \delta$ ein $\theta \in [0,1)$ mit

$$f(x+\xi) = \sum_{|\alpha| \le k-1} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha + \sum_{|\alpha| = k} \frac{D^\alpha f(x+\theta\xi)}{\alpha!} \xi^\alpha$$

$$= \sum_{|\alpha| \le k} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha + \sum_{|\alpha| = k} r_\alpha(\xi)\, \xi^\alpha$$

wobei $\quad r_\alpha(\xi) := \dfrac{D^\alpha f(x+\theta\xi) - D^\alpha f(x)}{\alpha!}$.

Wegen der Stetigkeit von $D^\alpha f$ in x gilt

$$\lim_{\xi \to 0} r_\alpha(\xi) = 0.$$

Außerdem gilt

$$|\xi^\alpha| = |\xi_1^{\alpha_1}| \cdots |\xi_n^{\alpha_n}|$$

$$= |\xi_1|^{\alpha_1} \cdots |\xi_n|^{\alpha_n}$$

$$\le \|\xi\|^{\alpha_1} \cdots \|\xi\|^{\alpha_n} = \|\xi\|^{\alpha_1 + \dots + \alpha_n}$$

$$= \|\xi\|^{|\alpha|} = \|\xi\|^k$$

Folglich

$$\lim_{\xi \to 0} \frac{\sum r_\alpha(\xi)\, \xi^\alpha}{\|\xi\|^k} = 0$$

$\blacksquare$

<u>Bemerkung 11.1.4</u>. Seien U, f wie in der Folgerung 11.1.3. Sei $0 \leq m \leq k$ ganze Zahl. Definiere

$$P_m(\xi) := \sum_{|\alpha| = m} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha$$

Die Formel in der Folgerung 11.1.3 lässt sich umschreiben als

$$f(x+\xi) = \sum_{m=0}^{k} P_m(\xi) + o(\|\xi\|^k)$$

Wir berechnen P_m explizit für $m = 0, 1, 2$.

i) <u>$m = 0$</u>:

$$P_0(\xi) = \frac{D^0 f(x)}{0!} \xi^0 = f(x) \quad (\forall \xi)$$

ii) <u>$m = 1$</u>: Die einzigen n-tupel $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| = 1$ sind die n Einheitsvektoren

$$e_i = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \dots, 0)$$

Es gilt $D^{e_i} f = D_i f$, $e_i! = 1$ und $\xi^{e_i} = \xi_i$.

Deshalb erhält man

$$P_1(\xi) = \sum_{i=1}^{n} D_i f(x)\,\xi_i = \langle \operatorname{grad} f(x), \xi \rangle$$

Folglich gilt (Approximation erster Ordnung: durch lineare Abbildung)

$$f(x+\xi) = f(x) + \langle \operatorname{grad} f(x), \xi \rangle + o(\|\xi\|)$$

wenn f einmal stetig differenzierbar ist.

Dies ist eine Umformulierung von der Differenzierbarkeit von f da $Df(x)\xi = \langle \operatorname{grad} f(x), \xi \rangle$ (siehe Bemerkung 10.1.5)

iii) $m=2$: wir bestimmen alle $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| = 2$, d.h. $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ mit $\alpha_j \in \mathbb{N}_0\ \forall j$

und $\alpha_1 + \alpha_2 + \dots + \alpha_n = 2$

Es gibt zwei Fälle:

Erster Fall: $\exists\ 1 \leq j \leq n$ mit $\alpha_j = 2$

Dann $\forall\ j' \neq j$, $\alpha_{j'} = 0$, und

$$\alpha = 2e_j = (0, 0, \dots, 0, \underset{\substack{\uparrow \\ \text{i-te Stelle}}}{2}, 0, \dots, 0)$$

Zweiter Fall: $0 \leq \alpha_j \leq 1\ \forall\ 1 \leq j \leq n$

$\Rightarrow \alpha_j \in \{0, 1\}$. Daher $\exists\ 1 \leq i < j \leq n$

mit $\alpha_i = \alpha_j = 1$ und $\alpha_{j'} = 0\ \forall\ j' \notin \{i, j\}$.

Man bekommt $\alpha = e_i + e_j = (0, \dots, 0, \underset{\substack{\uparrow \\ \text{i-te}}}{1}, 0, \dots, 0, \underset{\substack{\uparrow \\ \text{j-te Stelle}}}{1}, 0, \dots, 0)$

Es gilt
$$D^{2e_i} f = D_i^2 f, \quad (2e_i)! = 2! = 2$$
$$D^{e_i+e_j} f = D_i D_j f, \quad (e_i+e_j)! = 1!\,1! = 1$$
für $i \neq j$.

Daraus folgt
$$P_2(\xi) = \sum_{i=1}^{n} \frac{D^{2e_i} f(x)\, \xi^{2e_i}}{(2e_i)!} + \sum_{1\le i<j\le n} \frac{D^{e_i+e_j} f(x)}{(e_i+e_j)!} \xi^{e_i+e_j}$$

($\sum\limits_{1\le i<j\le n}$ bedeutet die Summe über alle $(i,j) \in \mathbb{N}^2$ mit $1 \le i < j \le n$)

$$= \frac{1}{2} \sum_{i=1}^{n} D_i^2 f(x)\, \xi_i^2 + \sum_{1\le i<j\le n} D_i D_j f(x)\, \xi_i \xi_j$$

Da $D_i D_j f = D_j D_i f$, gilt
$$\sum_{1\le i<j\le n} D_i D_j f(x)\, \xi_i \xi_j = \frac{1}{2} \sum_{\substack{i\neq j \\ 1\le i,j\le n}} D_i D_j f(x)\, \xi_i \xi_j$$

man erhält daher
$$P_2(\xi) = \frac{1}{2} \sum_{i,j=1}^{n} D_i D_j f(x)\, \xi_i \xi_j ,$$

d.h. P_2 ist eine quadratische Form mit der Matrix $\left(\frac{1}{2} D_i D_j f(x)\right)_{\substack{1\le i\le n \\ 1\le j\le n}}$

Dies motiviert uns die folgende Definition.

Definition 11.1.5 (Hesse-Matrix) Sei $U \subset \mathbb{R}^n$ offen $f: U \to \mathbb{R}$ zweimal stetig differenzierbar. Unter der Hesse-Matrix von f im Punkt $x \in U$ versteht man die $n \times n$-Matrix

$$(\mathrm{Hess}\, f)(x) := (D_i D_j f(x))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$$

Die Hesse-Matrix ist symmetrisch, da $D_i D_j f = D_j D_i f$

für $n=2$ ist $\mathrm{Hess}\, f(x)$

$$(\mathrm{Hess}\, f)(x) = \begin{bmatrix} D_1 D_1 f(x) & D_1 D_2 f(x) \\ D_2 D_1 f(x) & D_2 D_2 f(x) \end{bmatrix}$$

für $n=1$ ist $\mathrm{Hess} f(x)$ die zweite Ableitung $f''(x)$ von f.

Nach der Berechnungen von $P_1(\xi)$, $P_2(\xi)$ und Folgerung 11.1.3 erhält man sofort

(Approximation zweiter Ordnung)

Folgerung 11.1.4 Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $f: U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Dann gilt

$$f(x+\xi) = f(x) + \langle \mathrm{grad} f(x), \xi \rangle + \frac{1}{2}\langle \xi, A\xi \rangle + o(\|\xi\|^2)$$

wobei $A = (\mathrm{Hess}\, f)(x)$.

§ 11.2. Lokale Extrema

Ziel:
Wir untersuchen lokale Extrema von Funktionen mehrerer Variablen. Verallgemeinerungen von Sätzen von § 5.3 werden aufgestellt.

Erinnerung: Sei $I \subset \mathbb{R}$ offenes Intervall
$f: I \to \mathbb{R}$ differenzierbar, $x \in I$.

Dann gilt
(i) hat f in x lokales Extremum, dann ist $f'(x) = 0$

(ii) ist f zweimal differenzierbar und es gelte $f'(x) = 0$, $f''(x) > 0$ (bzw. $f''(x) < 0$), dann besitzt f in x ein strenges lokales Minimum (bzw. Maximum) (Folgerung 5.3.4)

Im Fall von mehreren Variablen spielt $\operatorname{grad} f$ (bzw. $\operatorname{Hess}(f)$) die Rolle von $f'(x)$ (bzw. $f''(x)$).

Definition 11.2.1 Sei $U \subset \mathbb{R}^n$ offen
$f: U \to \mathbb{R}$, $x \in U$.

Man sagt, f hat in x ein lokales Maximum (bzw. lokales Minimum) wenn $\exists$ eine Umgebung $V \subset U$ von x so dass $f(x) \geq f(y)$ (bzw. $f(x) \leq f(y)$)

für $\forall y \in V$.

Trifft in dieser Definition der Fall $f(x) = f(y)$ nur für $x = y$ ein, so spricht man von einem strikten lokalen Maximum oder Minimum.

Ein lokales Extremum ist ein lokaler Maximum oder Minimum.

Der Punkt x heißt einen kritischen Punkt von f wenn $\operatorname{grad} f(x) = 0$

Satz 11.2.2 (notwendige Bedingung für lokales Extremum) Sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$ partiell differenzierbar. Besitzt f in $x \in U$ ein lokales Extremum, so gilt

$$\operatorname{grad} f(x) = 0.$$

Beweis Sei $\varepsilon > 0$ mit $B_\varepsilon(x) \subset U$.

Betrachte

$$g_i(t) := f(x + te_i), \quad 1 \leq i \leq n, \quad -\varepsilon < t < \varepsilon.$$

Nach Voraussetzung ist g_i differenzierbar.

Hat f in x ein lokales Extremum, so hat g_i in 0 ein lokales Extremum. Daher gilt

$g_i'(0) = 0$. Da $D_i f(x) = g_i'(0)$,

erhält man $D_i f(x) = 0 \quad \forall\, 1 \leq i \leq n$

anders gesagt, $\operatorname{grad} f(x) = (D_1 f(x), \dots, D_n f(x))$
$= 0$. ∎

Wir suchen nun eine hinreichende Bedingung für die Existenz von lokalen Extrema. Dazu braucht man die Definitheit von der Hesse-Matrix.

Definition 11.2.3 Sei $A \in M(n \times n, \mathbb{R})$ eine symmetrische $n \times n$-Matrix.

i) A heißt positiv definit, falls
$$\langle \xi, A\xi \rangle > 0 \quad \forall\, \xi \in \mathbb{R}^n \setminus \{0\}.$$

ii) A heißt positiv semidefinit, falls
$$\langle \xi, A\xi \rangle \geq 0 \quad \forall\, \xi \in \mathbb{R}^n.$$

iii) A heißt negativ definit (bzw. negativ semidefinit), falls die Matrix $-A$ positiv definit (bzw. positiv semidefinit) ist.

iv) A heißt indefinit, falls $\exists\, \xi, \eta \in \mathbb{R}^n$ so dass $\langle \xi, A\xi \rangle > 0$ und $\langle \eta, A\eta \rangle < 0$.

<u>Bemerkung</u>: Aus linearen Algebra weißt man:

$\forall$ symmetrische $n \times n$-Matrix $A \in M(n \times n, \mathbb{R})$ gibt es eine Orthonormalbasis $v_1, \dots, v_n \in \mathbb{R}^n$ von Eigenvektoren:

$$\begin{cases} A v_i = \alpha_i \cdot v_i \\ \langle v_i, v_j \rangle = \delta_{ij} \end{cases} \quad \forall 1 \le i, j \le n.$$

wobei $\alpha_1, \dots, \alpha_n$ Eigenwerte von A (reelle) sind.

Sei $\xi \in \mathbb{R}^n$. Schreibt

$$\xi = \sum_{i=1}^{n} \lambda_i \cdot v_i$$

Dann gilt $$\langle \xi, A\xi \rangle = \sum_{i=1}^{n} \alpha_i \lambda_i^2$$

Damit bekommt man:

i) A ist genau dann positiv (bzw. negativ) definit, falls alle Eigenwerte von A positiv (>0) (bzw. negativ (<0)) sind.

ii) A ist genau dann positiv (bzw. negativ) semidefinit, wenn alle Eigenwerte ≥ 0 (bzw. ≤ 0) sind.

iii) A ist genau dann indefinit, wenn A mindestens einen positiven und einen negativen Eigenwert besitzt.

Die Berechnung der Eigenwerte einer Matrix kann recht kompliziert sein. Der folgenden Satz liefert ein Kriterium, welches praktischer ist.

Satz 11.2.4 (ohne Beweis: aus linearer Algebra)

Sei
$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in M(n\times n, \mathbb{R})$$

eine reelle symmetrische $n\times n$-Matrix. Dann gilt:

A ist positiv definit, genau dann, wenn ~~alle~~ für $k = 1, 2, \ldots, n$ gilt
$$\det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \\ a_{k1} & \cdots & a_{kk} \end{bmatrix} > 0$$

(die Determinanten der oberen linken quadratischen Teilmatrizen)

Beispiel

1) Sei $A \in M(2\times 2, \mathbb{R})$, symmetrische Matrix.
$$A = \begin{bmatrix} a & u \\ u & a \end{bmatrix}$$

Nach Satz 11.2.4 gilt

A ist positiv definit $\Leftrightarrow$ $a > 0$, $\det A > 0$

A ist negativ definit $\Leftrightarrow$ $a < 0$, $\det A > 0$

(weil: A negativ definit $\Leftrightarrow$ $-A$ positiv definit
$$-A = \begin{bmatrix} -a & -u \\ -u & -a \end{bmatrix}$$

$\Leftrightarrow \quad -a > 0, \quad \det(-A) > 0$

$\Leftrightarrow \quad a < 0, \quad \det A = \det(-A) > 0$)

A ist indefinit $\Leftrightarrow$ $\det A < 0$.

(denn: Seien α_1, α_2 Eigenwerte von A

Dann $\det A = \alpha_1 \alpha_2$

A ist indefinit $\Leftrightarrow$ A besitzt einen positiven Eigenwert und einen negativen Eigenwert

$\Leftrightarrow \alpha_1 \alpha_2 < 0$.)

ii) Sei
$$A = \begin{bmatrix} a & u & w \\ u & b & v \\ w & v & c \end{bmatrix} \in M(3\times 3, \mathbb{R})$$

A ist positiv definit $\Leftrightarrow$ $a > 0$, $\det \begin{bmatrix} a & u \\ u & b \end{bmatrix} > 0$

$\det A > 0$

$\Leftrightarrow a > 0, \quad ab - u^2 > 0, \det A > 0.$

<u>Satz 11.2.5</u> (hinreichende Bedingung für lokales Extremum). Sei $U \subset \mathbb{R}^n$ offen, $f: U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $x \in U$ ein Punkt mit $\operatorname{grad} f(x) = 0$.

i) Ist $(\operatorname{Hess} f)(x)$ positiv definit, so besitzt f in x ein striktes lokales Minimum.

ii) Ist $(\operatorname{Hess} f)(x)$ negativ definit, so besitzt f in x ein striktes lokales Maximum.

ii) Ist $(\mathrm{Hess} f)(x)$ indefinit, so besitzt f in x kein lokales Extremum.

<u>Beweis</u> Sei $A := (\mathrm{Hess} f)(x)$. Nach Folgerung 11.1.4 gilt

$$f(x+\xi) = f(x) + \overbrace{\langle \mathrm{grad} f(x), \xi \rangle}^{=0 \text{ nach Voraussetzung}} + \frac{1}{2}\langle \xi, A\xi \rangle + \varphi(\xi), \qquad (1)$$

für ξ in einer Umgebung von $0 \in \mathbb{R}^n$

$\varphi(\xi) = o(\|\xi\|^2)$, d.h. $\lim\limits_{\xi \to 0} \dfrac{\varphi(\xi)}{\|\xi\|^2} = 0$.

Deshalb gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass

$$|\varphi(\xi)| \leq \varepsilon \|\xi\|^2 \quad \text{für } \|\xi\| < \delta.$$

i) Wir nehmen nun an, dass A positiv definit ist.

Sei $S = \{\xi \in \mathbb{R}^n ; \|\xi\| = 1\}$ die Sphäre vom Radius 1.

Da S kompakt ist, nimmt die stetige Funktion

$$S \longrightarrow \mathbb{R} \qquad (\text{Satz } 8.4.7)$$
$$\xi \longmapsto \langle \xi, A\xi \rangle$$

auf S ihr Minimum an. Da $\langle \xi, A\xi \rangle > 0 \;\forall \xi \in S$ gilt

$$\alpha := \min\{\langle \xi, A\xi \rangle : \xi \in S\} > 0$$

Behauptung: Es gilt $\langle \xi, A\xi \rangle \geq \alpha \|\xi\|^2$ (2)

$\forall \xi \in \mathbb{R}^n$.

Beweis zur Behauptung: oder $\xi = 0$ für $\xi \in S$ ist das trivial.

Für einen beliebigen Vektor $\xi \in \mathbb{R}^n \setminus \{0\}$ setze

$\xi^* := \dfrac{\xi}{\|\xi\|} \in S$. Daher $\langle \xi^*, A\xi^* \rangle \geq \alpha$.

Folglich gilt

$$\langle \xi, A\xi \rangle = \|\xi\|^2 \langle \xi^*, A\xi^* \rangle \geq \alpha \|\xi\|^2.$$

Wir zeigen nun: f hat in x ein striktes lokales Minimum.

Wir wählen $\delta > 0$ so klein, dass

$$|\varphi(\xi)| \leq \frac{\alpha}{4} \|\xi\|^2 \text{ für } \|\xi\| < \delta$$

Nach (1) und (2) gilt

$$\begin{aligned} f(x+\xi) &\geq f(x) + \frac{\alpha}{2}\|\xi\|^2 + \varphi(\xi) \\ &\geq f(x) + \frac{\alpha}{2}\|\xi\|^2 - \frac{\alpha}{4}\|\xi\|^2 \\ &\geq f(x) + \frac{\alpha}{4}\|\xi\|^2 > f(x) \end{aligned}$$

für ξ mit $\|\xi\| < \delta$. Das bedeutet: f hat in x ein striktes lokales Minimum.

ii) Ist A negativ definit, betrachten wir $-f$ anstelle von f, und die Behauptung folgt aus (i).

iii) Sei A indefinit. Wir müssen zeigen dass in jeder Umgebung von x $\exists$ y', y'' mit $f(y'') < f(x) < f(y')$.

Da A indefinit ist, $\exists$ $\xi \in \mathbb{R}^n \setminus \{0\}$ mit $\alpha := \langle \xi, A\xi \rangle > 0$. Daher nach (£) für kleine reelle Zahlen t gilt

$$f(x+t\xi) = f(x) + \frac{1}{2}\langle t\xi, A(t\xi)\rangle + \varphi(t\xi)$$

$$= f(x) + \frac{t^2\alpha}{2} + \varphi(t\xi)$$

da $\varphi(t\xi) = o(\|t\xi\|^2) = o(t^2\|\xi\|^2)$

$\exists$ $\delta > 0$ mit $|\varphi(t\xi)| \leq \frac{\alpha}{4}t^2$ für $|t| < \delta$.

Daraus folgt $\underbrace{f(x+t\xi)}_{y' :=} \geq f(x) + \frac{t^2\alpha}{2} - \frac{t^2\alpha}{4}$

$> f(x)$

für $0 < |t| < \delta$.

Analog $\exists$ $\eta \in \mathbb{R}^n \setminus \{0\}$ mit $\langle \eta, A\eta \rangle < 0$. Man zeigt wie oben dass für $t \in \mathbb{R}$ mit $|t|$ genügend klein $t \neq 0$ gilt

$$f(\underbrace{x+t\eta}_{y'' :=}) < f(x)$$ ∎

<u>Bemerkung 11.2.6</u> Seien f, x wie im Satz 11.2.5. Wenn $(\text{Hess} f)(x)$ nur semidefinit ist, kann man im Allgemeinen keine Aussage (über lokale Extrema) machen. z.B.

$$f_1(x,y) = x^2 + y^3 \quad , (x,y) \in \mathbb{R}^2$$

$$\text{grad} f_1(x,y) = (2x, 3y^2)$$

$\Rightarrow$ $(0,0)$ ist ein kritischer Punkt von f_1.

$$\text{Hess} f_1(x,y) = \begin{bmatrix} \partial^2_{x^2} f_1(x,y) & \partial_x \partial_y f_1(x,y) \\ \partial_x \partial_y f_1(x,y) & \partial^2_{y^2} f_1(x,y) \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 0 & 6y \end{bmatrix}$$

$$\text{Hess} f_1(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{positiv semidefinit.}$$

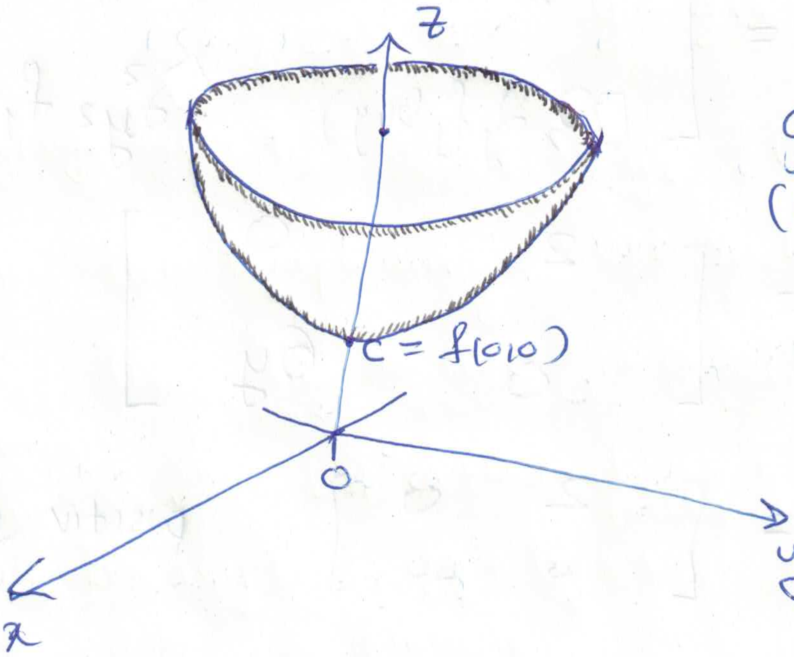
Man beachte: f_1 hat in $(0,0)$ weder ein lokales Minimum noch ein lokales Maximum.

Beispiel i) Sei $f(x,y) := c + x^2 + y^2$ (wobei $c \in \mathbb{R}$ eine Konstante), $(x,y) \in \mathbb{R}^2$.

$\text{grad} f(x,y) = (2x, 2y)$

$\text{Hess} f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ positiv definit

Daher ist $(0,0)$ ein kritischer Punkt von f und nach Satz 11.2.5 (i) besitzt f im $(0,0)$ ein lokales Minimum



graph von f (Paraboloid)

ii) Sei $f(x,y) = c - x^2 - y^2$, $(x,y) \in \mathbb{R}^2$

$\text{grad} f(x,y) = (-2x, -2y)$

$\text{Hess} f(x,y) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$ negativ definit

$\Rightarrow$ f besitzt im $(0,0)$ ein lokales Maximum.

iii) Sei $f(x,y) = c + x^2 - y^2$, $(x,y) \in \mathbb{R}^2$.

$\operatorname{grad} f(x,y) = (2x, -2y)$

$\operatorname{Hess} f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$ indefinit.

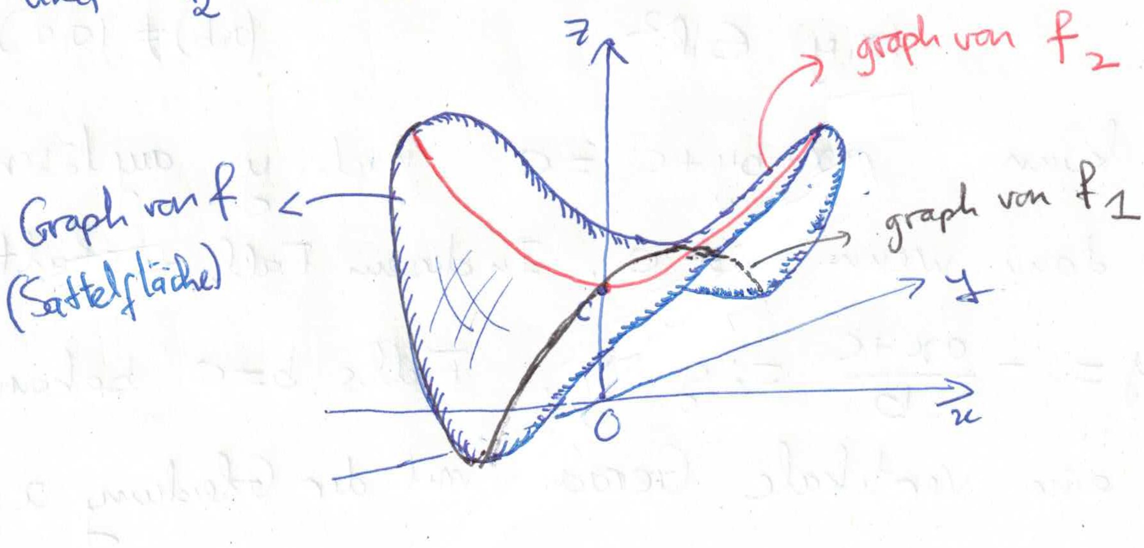
Dann ist $(0,0)$ ein kritischen Punkt von f, in dem f weder ein lokales Minimum noch ein lokales Maximum besitzt: wir untersuchen f genauer.

Betrachten wir die Einschränkungen von f auf $\{(x,0) : x \in \mathbb{R}\}$ und $\{(0,y) : y \in \mathbb{R}\}$, d.h.

$$f_1 : \mathbb{R} \to \mathbb{R}$$
$$y \mapsto f(0,y) = c - y^2$$

$$f_2 : \mathbb{R} \to \mathbb{R}$$
$$x \mapsto f(x,0) = c + x^2$$

Man beachte, dass f_1 in 0 ~~eine~~ ein lokales Maximum hat, und f_2 in 0 ein lokales Minimum besitzt



§ 12. Implizite Funktionen

Motivation. Seien $U_1 \subset \mathbb{R}$, $U_2 \subset \mathbb{R}$ offene Menge.

Sei $F: U_1 \times U_2 \to \mathbb{R}$ eine Funktion.
$(x, y) \mapsto F(x,y)$

In vielen Situationen muss man die Gleichung $F(x,y)=0$ nach y auflösen, d.h. man muss eine Funktion g finden so dass

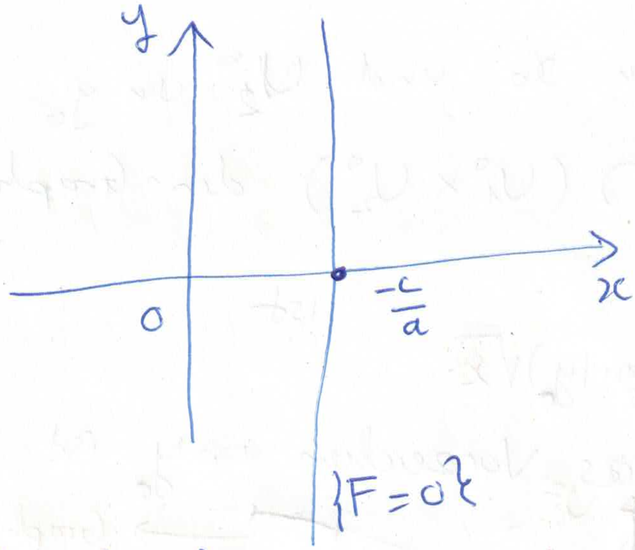
$$F(x,y)=0 \iff y = g(x).$$

anders ~~gesag~~ gesagt, die Menge $\{(x,y) : F(x,y)=0\}$ $= F^{-1}(0)$ ist genau der Graph von g. Wir suchen eine natürliche Bedingung, die die Aussage impliziert.

Wir betrachten nun Beispiele:

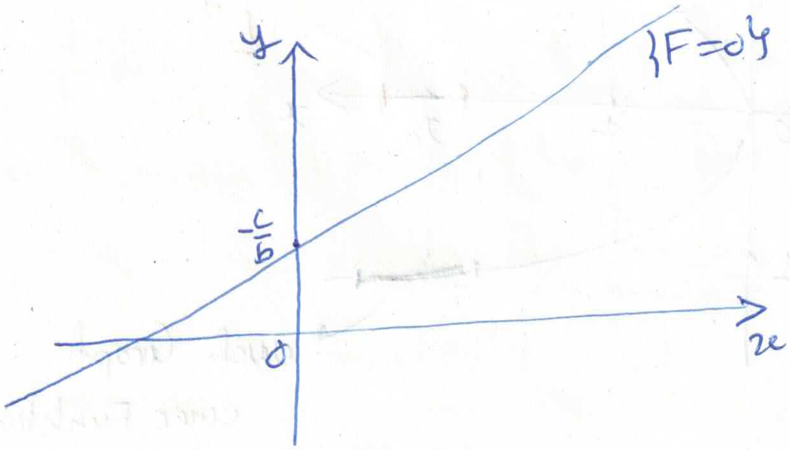
(1) Sei $F(x,y) = ax + by + c$, $a, b, c \in \mathbb{R}$
$(x,y) \in \mathbb{R}^2$ $(a,b) \neq (0,0)$.

Man kann $ax+by+c=0$ nach y auflösen genau dann wenn $b \neq 0$. In diesem Fall erhält man $y = -\frac{ax+c}{b} =: g(x)$. Falls $b=0$ bekommen wir eine vertikale Gerade mit der Gleichung $x = \frac{-c}{a}$



(kein Graph über die x-Achse)

$\underline{b=0}$



$\underline{b\neq 0}$

Graph von g

Beachte, $\dfrac{\partial F}{\partial x}(x,y) = a$, $\dfrac{\partial F}{\partial y}(x,y) = b$

Daher wenn $\dfrac{\partial F}{\partial y} \neq 0$, lässt $F(x,y)=0$ sich nach y auflösen.

(ii) Sei $F: \mathbb{R}^2 \to \mathbb{R}$
$(x,y) \mapsto x-y^2$

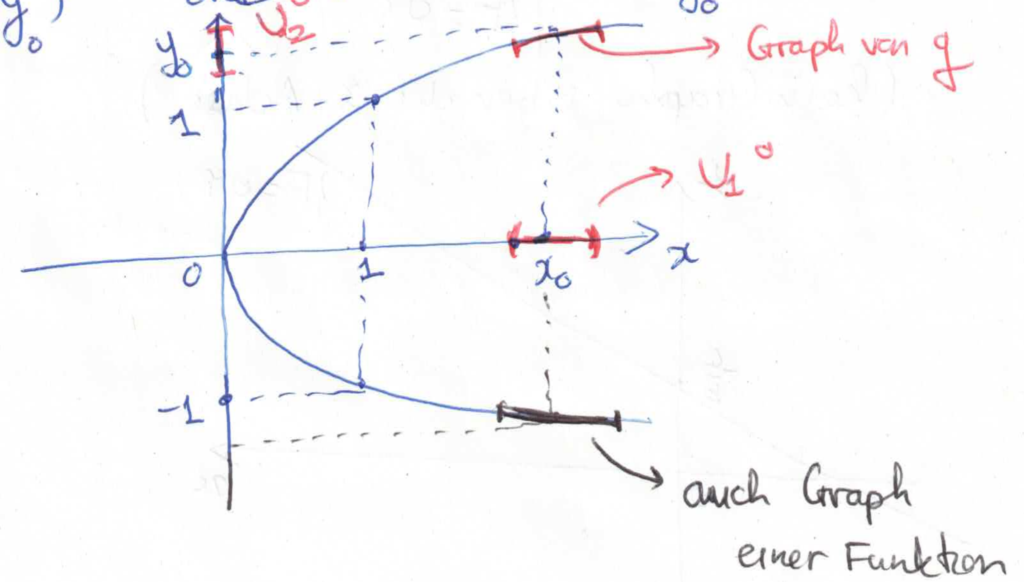
Dann ist $F^{-1}\{0\} = \{(x,y) : x = y^2\}$ kein Graph über die x-Achse. In der Tat, für $x=1$ gilt $(1,1), (1,-1) \in F^{-1}(0)$.

Aber für jeder $(x_0, y_0) \in F^{-1}(0)$ mit $y_0 \neq 0$ $\exists$

eine Umgebung U_1° von x_0 und U_2° von y_0 sodass $F^{-1}(0) \cap (U_1^\circ \times U_2^\circ)$ der Graph von

$$g : U_1^\circ \to \mathbb{R} \qquad \text{ist}$$
$$x \mapsto \operatorname{sgn}(y_0)\sqrt{x}$$

wobei $\operatorname{sgn}(y_0)$ das Vorzeichen von y_0 ist

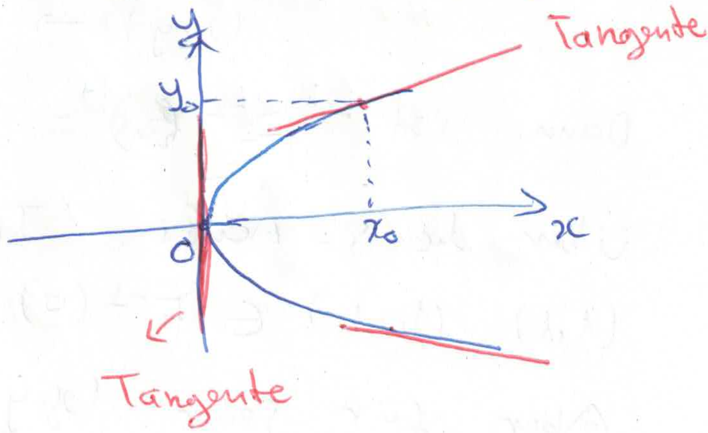


Man sagen, $F^{-1}(0)$ ist ein lokaler Graph nahe (x_0, y_0).

Andererseits ist $F^{-1}(0)$ kein lokaler Graph nahe (über x-Achse) $(0,0)$.

Was ist die Unterschied zwischen (x_0, y_0) $(y_0 \neq 0)$ und $(0,0)$? Diese liegt bei der Tangente ~~in~~ im diesen Punkten:

Die Tangente an $F^{-1}(0)$ im $(0,0)$ ist <u>vertikal</u>. und diese für (x_0, y_0) ist <u>nicht</u> vertikal.



d.h. $\frac{\partial f}{\partial y}(0) = 0$ wobei $f(y) = y^2$

Hier man löst $F(x,y)=0 \Leftrightarrow x=y^2=f(y)$.

Beachte $\frac{\partial F}{\partial y}(x,y) = \frac{\partial f}{\partial y}(y)$

Zum Schluss wenn $\frac{\partial F}{\partial y}(x_0,y_0) \neq 0 \Rightarrow F^{-1}(0)$

ist ein lokaler Graph nahe (x_0,y_0).

Beide Beispiele zeigen, dass die Bedingung $\frac{\partial F}{\partial y}(x_0,y_0)\neq 0$

eine Rolle spielt in der Existenz einer Funktion g

so dass

$$F(x,y)=0 \Leftrightarrow y=g(x) \quad (\text{nahe } (x_0,y_0))$$

(Man sagt, in diesem Fall, g werde durch die Gleichung $F(x,y)=0$ ~~dg~~ implizit definiert).

Der Satz über implizite Funktionen besagt, dass in der Tat:

$$\frac{\partial F}{\partial y}(x_0,y_0) \neq 0 \Rightarrow F^{-1}(0) \text{ ist ein lokaler Graph nahe } (x_0,y_0).$$

Dies gilt auch für den Fall, dass x,y mehr-dimensional sind.

Um den Satz über implizite Funktionen genauer zu besprechen, brauchen wir den Banachschen Fixpunktsatz

Satz 12.1 (Banachscher Fixpunktsatz)

Sei $(V, \| \|)$ ein Banachraum, $A \subset V$ abgeschlossen.

Die Abbildung $\Phi: A \to A$ sei eine Kontraktion,

d.h. es gebe eine Konstante θ mit $0 < \theta < 1$

so dass

$$\| \Phi(f) - \Phi(g) \| \leq \theta \| f - g \| \quad \forall f, g \in A$$

Dann besitzt Φ genau einen Fixpunkt, d.h.

$\exists!\ f_* \in A$ mit $\Phi(f_*) = f_*$.

Für einen beliebigen Anfangswert $f_0 \in A$ konvergiert die durch $f_k := \Phi(f_{k-1})$ rekursiv definierte Folge $(f_k)_{k \in \mathbb{N}}$ gegen den Fixpunkt f_*.

Beweis

i) Zur Eindeutigkeit. Seien f_*, g_* zwei Fixpunkte von Φ. Dann gilt

$$\| f_* - g_* \| = \| \Phi(f_*) - \Phi(g_*) \| \leq \theta \| f_* - g_* \| .$$

Daraus folgt: $f_* = g_*$ da $\theta \in (0, 1)$.

ii) Zur Existenz: Sei $f_0 \in A$ beliebig. Setze

$$f_k := \Phi(f_{k-1}) \qquad k \in \mathbb{N}.$$

Es gilt

$$\| f_{k+1} - f_k \| = \| \Phi(f_k) - \Phi(f_{k-1}) \|$$
$$\leq \theta \| f_k - f_{k-1} \| \leq \theta^2 \| f_{k-1} - f_{k-2} \| \leq$$
$$\dots$$
$$\leq \theta^k \| f_1 - f_0 \|$$

Deshalb erhält man

$$\| f_m - f_k \| = \| \sum_{i=k}^{m-1} (f_{i+1} - f_i) \|$$
$$\leq \sum_{i=k}^{m-1} \| f_{i+1} - f_i \|$$
$$\leq \left(\sum_{i=k}^{m-1} \theta^i \right) \| f_1 - f_0 \|$$
$$= \theta^k \left(\sum_{j=0}^{m-k-1} \theta^j \right) \| f_1 - f_0 \|$$
$$\leq \frac{1}{1-\theta} \cdot \theta^k \cdot \| f_1 - f_0),$$ welche

gegen 0 konvergiert, wenn $k \to \infty$. Das zeigt, $(f_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in V. Da V Banach ist, $\exists \; f_* \in V$ mit $f_* = \lim_{k \to \infty} f_k$

Da $\left. \begin{matrix} f_k \in A \quad \forall k \\ A \text{ abgeschlossen} \end{matrix} \right\}$, gilt $f_* \in A$.

Zeitlem nach Stetigkeit von Φ auf A (folgt aus dass Φ eine Kontraktion ist), bekomme man

$$\Phi(f_*) = \lim_{k\to\infty} \Phi(f_k)$$
$$= \lim_{k\to\infty} f_{k+1} = f_* \qquad \blacksquare$$

Definition 12.2. Sei $F: U_1 \times U_2 \to \mathbb{R}^m$, $(x,y) \mapsto F(x,y)$ stetig differenzierbar, wobei $U_1 \subset \mathbb{R}^k$ offen, $U_2 \subset \mathbb{R}^n$ offen.

Setze

$$\frac{\partial F}{\partial x} := \frac{\partial(F_1,\dots,F_m)}{\partial(x_1,\dots,x_k)} := \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_k} \\ \vdots & & \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_k} \end{bmatrix}$$

$m\times k$-Matrix.

$$\frac{\partial F}{\partial y} := \frac{\partial(F_1,\dots,F_m)}{\partial(y_1,\dots,y_n)} := \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_n} \end{bmatrix}$$

$m\times n$-Matrix, wobei $F = (F_1,\dots,F_m)$.

Bemerkung: Erinnere, dass ~~die~~ für $U \subset \mathbb{R}^n$ ~~offen~~

$$\mathcal{C}_b(U,\mathbb{R}^m) = \{ f: U \to \mathbb{R}^m \text{ stetig und beschränkt} \}$$

ein Banachraum mit der Sup-norm (Beispiel 8.3.2d)

$$\|f\|_U := \{ \|f(x)\| : x \in U \}.$$

Wir werden später den Banachschen Fixpunktsatz auf $\mathcal{C}_b(U, \mathbb{R}^m)$ anwenden.

Lemma 12.3 Seien $U_1 \subset \mathbb{R}^k$ und $U_2 \subset \mathbb{R}^m$ offen und

$$F: U_1 \times U_2 \to \mathbb{R}^m, \quad (x,y) \mapsto F(x,y)$$

stetig differenzierbar sodass die Matrix $\frac{\partial F}{\partial y}(x,y)$ invertierbar ist. $\forall (x,y)$

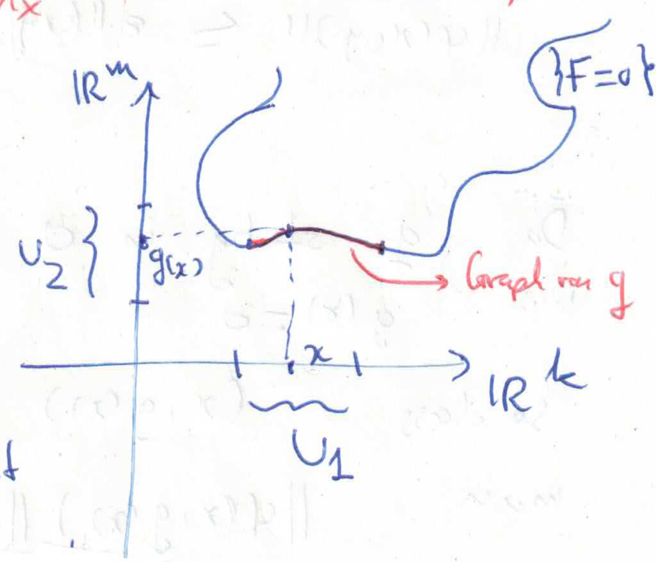
Seien $(a,b) \in U_1 \times U_2$ mit $F(a,b) = 0$

$g: U_1 \to U_2$ eine stetige Abbildung mit

$$\begin{cases} F(x, g(x)) = 0 & \forall x \in U_1 \\ g(a) = b \end{cases}$$

Dann ist g stetig differenzierbar und es gilt

$$\underbrace{\frac{\partial g}{\partial x}(x)}_{m\times k\text{-Matrix}} = -\underbrace{\left(\frac{\partial F}{\partial y}(x, g(x))\right)^{-1}}_{m\times m\text{-Matrix}} \cdot \underbrace{\frac{\partial F}{\partial x}(x, g(x))}_{m\times k\text{-Matrix}}$$



Beweis:

Wir zeigen zuerst dass g differenzierbar in x_0 ist $\forall x_0 \in U_1$.

O.B.d.A nehmen wir an, $x_0 = 0$, $g(x_0) = 0$

Setze $A := \frac{\partial F}{\partial x}(0,0)$, $B := \frac{\partial F}{\partial y}(0,0)$

Aus der Differenzierbarkeit von F im $(0,0)$ folgt

$F(x,y) = Ax + By + \varphi(x,y)$ mit $\varphi(x,y) = o(\|(x,y)\|)$

Da $F(x, g(x)) = 0 \quad \forall x \in U_1$, gilt

$$Ax + Bg(x) + \varphi(x, g(x)) = 0 \qquad (1)$$

Folglich bekommt man $g(x) = -B^{-1}Ax - \underbrace{B^{-1}\varphi(x,g(x))}_{\psi(x):=}$

(da B invertierbar ist).

Um die Differenzierbarkeit in 0 von g zu zeigen, genügt es zu überprüfen dass $\psi(x) = o(\|x\|)$.

Setze $c_1 := \|B^{-1}A\|$, $c_2 := \|B^{-1}\|$

wegen $\varphi(x,y) = o(\|(x,y)\|)$ gibt es zu $\varepsilon := \frac{1}{2c_2}$ eine Umgebung $U' \subset U_1 \times U_2$ von $(0,0)$ so dass

$$\|\varphi(x,y)\| \leq \varepsilon\|(x,y)\| \leq \frac{1}{2c_2}(\|x\| + \|y\|)$$
$$\forall (x,y) \in U'.$$

Da $\left.\begin{matrix} g \text{ stetig in } 0 \text{ ist} \\ g(0) = 0 \end{matrix}\right\}$, $\exists$ eine Umgebung U_1' von 0 $\in \mathbb{R}^k$

so dass $(x, g(x)) \in U' \ \forall x \in U_1'$. Daher erhält man

$$\|\varphi(x, g(x))\| \leq \frac{1}{2c_2}(\|x\| + \|g(x)\|)$$

Die Gleichung (1) liefert

$$\begin{aligned}\|g(x)\| &\le c_1\|x\| + c_2 \cdot \|\varphi(x,g(x))\| \\ &\le c_1\|x\| + \frac{1}{2}(\|x\| + \|g(x)\|) \\ &= \left(c_1 + \frac{1}{2}\right)\|x\| + \frac{1}{2}\|g(x)\|\end{aligned}$$

Daraus folgt $\|g(x)\| \le (2c_1+1)\|x\|$.

und
$$\begin{aligned}\psi(x) &= B^{-1}\varphi(x,g(x)) = o(\|(x,g(x))\|) \\ &= o(\|x\| + \|g(x)\|) = o(\|x\|).\end{aligned}$$

Man bekommt
$$g(x) = -B^{-1}A\,x + o(\|x\|)$$
$$\frac{\partial g}{\partial x}(0) = -B^{-1}A$$

Anders gesagt
$$\frac{\partial g}{\partial x}(x) = -\underbrace{\left(\frac{\partial F}{\partial x}(x,g(x))\right)^{-1}}_{\text{stetig in } x}\cdot\underbrace{\frac{\partial F}{\partial y}(x,g(x))}_{\text{stetig in } x}$$

wegen der stetigen Differenzierbarkeit von F.
Daher ist g stetig differenzierbar. ∎

Satz 12.4 (über implizite Funktionen)

Seien $U_1 \subset \mathbb{R}^k$, $U_2 \subset \mathbb{R}^m$ offene Teilmengen
und
$$F : U_1 \times U_2 \to \mathbb{R}^m$$
$$(x,y) \mapsto F(x,y)$$
stetig differenzierbar

Sei $(a,b) \in U_1 \times U_2$ mit $F(a,b) = 0$.

Die Matrix $\frac{\partial F}{\partial y}(a,b)$ sei invertierbar.

Dann $\exists$ eine offene ~~Umgebungen~~ Umgebung $V_1 \subset U_1$ von a

$\exists$ eine Umgebung $V_2 \subset U_2$ von b

sowie eine stetig differenzierbare Abbildung

$g : V_1 \to V_2$ mit $g(a) = b$, so dass

$F(x, g(x)) = 0 \quad \forall x \in V_1$.

Ist $(x,y) \in V_1 \times V_2$ mit $F(x,y) = 0$ so folgt

$y = g(x)$.

<u>Bemerkung:</u>

i) Ist $\frac{\partial F}{\partial y}$ invertierbar im Punkt (a,b), so ist es auch invertierbar in einer gewissen Umgebung von (a,b): <u>weil</u>

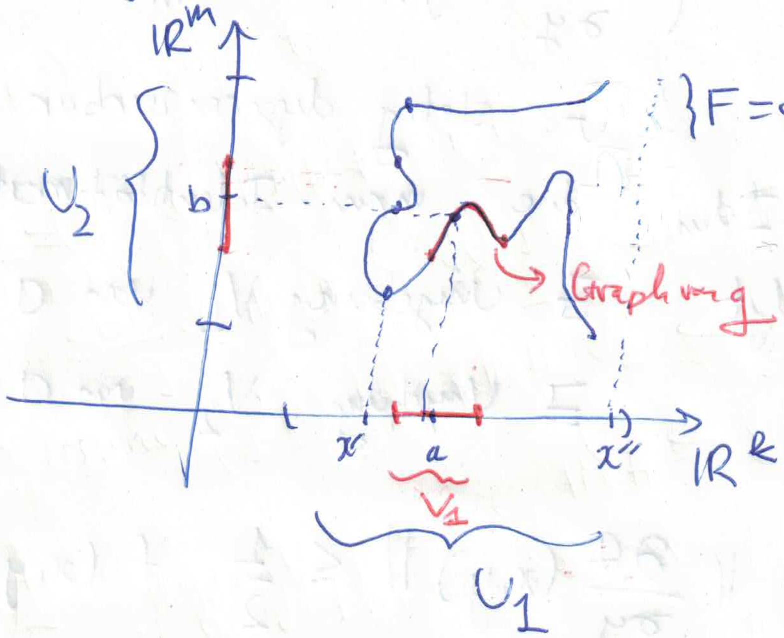
die Funktion
$$\tau : U_1 \times U_2 \to \mathbb{R}$$
$$(x,y) \mapsto \det\left(\frac{\partial F}{\partial y}(x,y)\right)$$

stetig ist und es gilt

$\frac{\partial F}{\partial y}(x,y)$ ist invertierbar $\Leftrightarrow$ $\tau(x,y) \neq 0$.

ii) Im Allgemeinen ist V_1 (bzw. V_2) eine echte Teilmenge von U_1 (bzw. U_2). In ganz $U_1 \times U_2$

könnte es zu einem gegebenen x mehreren y-Werte (oder auch gar kein) geben, die die Gleichung $F(x,y)=0$ genügen.



Im Bild
für x' gibt es mehreren y mit $F(x',y)=0$
für x'' gibt es kein y mit $F(x'',y)=0$.

<u>Beweis von Satz 12.4</u> O.B.d.A sei $(a,b)=(0,0)$.

Setze $B := \frac{\partial F}{\partial y}(0,0)$.

Definiere
$$G : U_1 \times U_2 \to \mathbb{R}^m$$
$$(x,y) \mapsto y - B^{-1} F(x,y).$$

Beachte
$$F(x,y)=0 \iff G(x,y)=y$$

d.h.
Die Gleichung lösen $F(x,y)=0$ (x fest) $\iff$ einen Fixpunkt von der Abbildung $G(x,\cdot): U_2 \to \mathbb{R}^m$

finden.

Beachte
$$\begin{cases} G(0,0) = 0 \\ \frac{\partial G}{\partial y}(x,y) = Id_m - B^{-1}\frac{\partial F}{\partial y}(x,y) \\ G \text{ stetig differenzierbar} \end{cases}$$

wobei Id_m die $m\times m$-Identitätsmatrix bezeichnet.

Es folgt $\exists$ Umgebung V_1 von $0 \in \mathbb{R}^k$
$\exists$ Umgebung V_2 von $0 \in \mathbb{R}^m$

so dass

1) $$\left\| \frac{\partial G}{\partial y}(x,y) \right\| \leq \frac{1}{2} \quad \forall (x,y) \in V_1 \times V_2$$
$$\left(\text{da } \frac{\partial G}{\partial y}(0,0) = 0 \right)$$

2) $$\varepsilon := \sup_{x\in V_1} \| G(x,0)\| \leq \frac{r}{2}, \text{ wobei} \quad (1)$$
$$V_2 = \overline{\mathbb{B}_r(0)} \subset \mathbb{R}^m.$$

Es folgt
$$\| G(x,y) - G(x,y')\| \leq \frac{1}{2}\|y-y'\| \quad (2)$$
$\forall\, y, y' \in V_2 = \overline{\mathbb{B}_r(0)}$.

Setzt man $y' = 0$, so ergibt sich zusammen mit (1):
$$\|y\| \leq r \Rightarrow \| G(x,y)\| \leq r \quad (3)$$

$\forall x \in V_1$. Deshalb $\forall x$ festes in V_1

ist die Abbildung $G_x : V_2 \to V_2$

$y \mapsto G(x, y)$

eine Kontraktion auf die abgeschlossene Menge V_2 im Banachraum $\mathbb{R}^m$.

Nach dem Banachem Fixpunktsatz $\exists !\, y_x \in V_2$ mit $G(x, y_x) = y_x$, d.h. $\exists !\, y_x$ mit $F(x, y_x) = 0$.

Setzen $g(x) := y_x$, bekommen wir

$g : V_1 \to V_2$ eindeutig bestimmt mit:

$x \mapsto g(x)$

$$\left\{\begin{array}{l} F(x,y) = 0 \\ (x,y) \in V_1 \times V_2 \end{array}\right. \iff \left\{\begin{array}{l} y = g(x) \\ x \in V_1 . \end{array}\right. \qquad (4)$$

Wir zeigen nun dass g stetig ist.

Wenn wir diese schaffen nach Lemma 123, ist g automatisch stetig differenzierbar und der Satz ist bewiesen.

Erinnere $\mathcal{C}_b(V_1, \mathbb{R}^m) = \{ \varphi : V_1 \to \mathbb{R}^m$ stetig beschränkt$\}$

$\|\varphi\|_{V_1} = \sup\{ \|\varphi(x)\| : x \in V_1 \}$.

Nach (3) sieht man dass

wenn $\|\varphi\|_{V_1} \le r \Rightarrow \|\psi\|_{V_1} \le r$, wobei

$$\psi : V_1 \longrightarrow \mathbb{R}^m$$
$$x \longmapsto G(x, \varphi(x)).$$

Es folgt, die Abbildung $\Phi : A \longrightarrow A$
$$\varphi \longmapsto \psi$$

wobei $A := \{ \varphi \in \mathscr{C}_b(V_1, \mathbb{R}^m) : \|\varphi\| \le r \}$.

ist wohl-definiert. Aus (2) folgt

$$\|\Phi(\varphi_1) - \Phi(\varphi_2)\|_{V_1} = \sup_{x \in V_1} \|G(x, \varphi_1(x)) - G(x, \varphi_2(x))\|$$
$$\le \frac{1}{2} \sup_{x \in V_1} \|\varphi_1(x) - \varphi_2(x)\|$$
$$= \frac{1}{2} \|\varphi_1 - \varphi_2\|_{V_1}$$

das bedeutet, Φ ist eine Kontraktion.

Nach dem Banachen Fixpunktsatz besitzt Φ einen eindeutigen Fixpunkt $\tilde{g} \in A \subset \mathscr{C}_b(V_1, \mathbb{R}^m)$.

d.h. $G(x, \tilde{g}(x)) = \tilde{g}(x)$, $x \in V_1$.

Nach (4) gilt $\tilde{g}(x) = g(x)$ $\forall x \in V_1$ und dass g stetig beschränkt ist. ∎

Wir besprechen nun eine Anwendung von Satz über impliziten Funktionen.

Umkehrabbildungen:

Frage: Seien $U_1, U_2 \subset \mathbb{R}^n$ offene Mengen und $f: U_1 \to U_2$ stetig differenzierbar.

Wir interessieren uns für die Frage, wann die Abbildung bijektiv ist und die Eigenschaften von Umkehrabbildung.

Dazu erinnern wir daran:

Erinnerung: (nach Kapitel 4 + 5: Satz 4.3.5 und Satz 5.1.7)

Sei $I, J \subset \mathbb{R}$ Intervalle
$f: I \to J$ stetig differenzierbar und bijektiv

Sei $f'(x) \neq 0 \quad \forall x \in I$.
Dann erfüllt $g := f^{-1}: J \to I$

$$g'(y) = \frac{1}{f'(g(y))} \quad \forall y \in J \qquad (*)$$

und g ist stetig differenzierbar. Nach $(*)$ kann man durch Induktion zeigen, dass wenn f k-mal stetig differenzierbar ist, dann ist g auch k-mal stetig differenzierbar $\forall k \in \mathbb{N}_0$.

Man hat ein Analog in der höheren Dimensionen.

Satz 12.5 (Satz über die Umkehrabbildung)

Sei $U \subset \mathbb{R}^n$ offen und

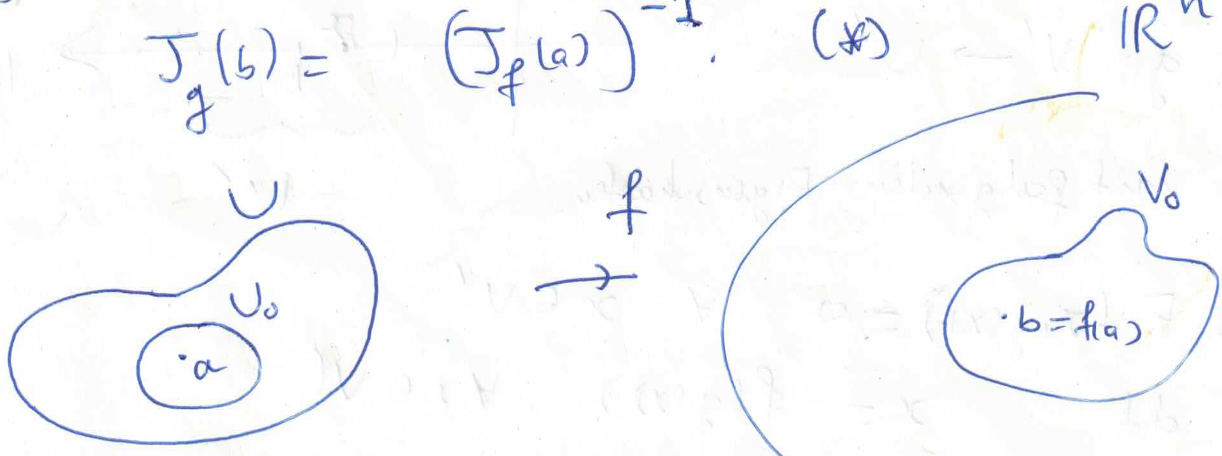
$f: U \to \mathbb{R}^n$ stetig differenzierbar.

Sei $a \in U$ und $b := f(a)$. Die Jacobi-Matrix $J_f(a)$ sei invertierbar. Dann gibt es eine offene Umgebung $U_0 \subset U$ von a und eine offene Umgebung V_0 von b, so dass f die Menge U_0 bijektiv auf V_0 abbildet und die Umkehrabbildung

$$g := f^{-1} : V_0 \to U_0$$

stetig differenzierbar ist. Es gilt

$$J_g(b) = \left(J_f(a)\right)^{-1}. \quad (*)$$



$$\mathbb{R}^n \xrightarrow{Df(a)} \mathbb{R}^n$$

invertierbar

$$v \longmapsto J_f(a) \cdot v$$

Beweis Sei $F: \mathbb{R}^n \times U \to \mathbb{R}^n$

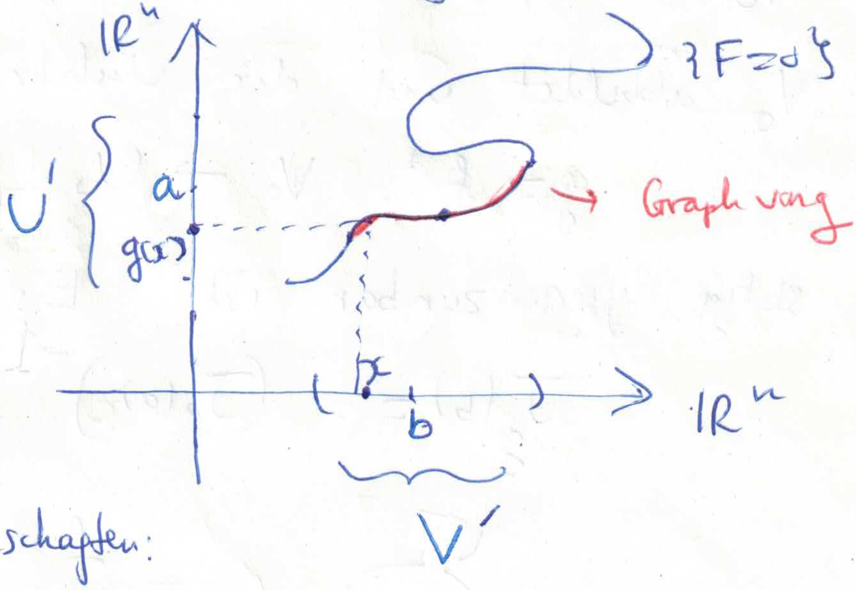
$$(x,y) \mapsto F(x,y) := x - f(y)$$

Beachte $F(b,a) = 0$, $\frac{\partial F}{\partial y}(x,y) = -J_f(y)$

Daher ist $\frac{\partial F}{\partial y}(b,a)$ invertierbar. Deshalb kann man den Satz über impliziten Funktionen auf F anwenden und bekommen eine offene Umgebung V' von b, eine Umgebung $U' \subset U$ von a

und eine stetig differenzierbare Abbildung

$g: V' \to U'$



mit folgenden Eigenschaften:

(1) $F(x, g(x)) = 0 \quad \forall\ x \in V'$

d.h. $x = f(g(x)) \quad \forall x \in V'$

(2) Ist $(x,y) \in V' \times U'$ mit $F(x,y) = 0$ (d.h. $x = f(y)$) so folgt $y = g(x)$.

Man erhält: $V' \xrightarrow{g} g(V') \xrightarrow{f} f(g(V')) = V'$

$$x \longmapsto g(x) \longmapsto x$$

Beide Abbildungen sind bijektiv und $g \circ f(y) = y \ \forall y \in g(V')$.

411

Wegen der Stetigkeit von f gibt es eine offene Umgebung $U_0 \subset U'$ von a mit $f(U_0) \subset V'$.

Daher gilt $V_0 := f(U_0) = g^{-1}(U_0)$

(wegen (2) gilt $U_0 \subset g(V') \Rightarrow g \circ f(y) = y \ \forall y \in U_0$)

Da g stetig ist, ist V_0 eine offene Umgebung von b. Man bekommt: $f: U_0 \to V_0$ bijektiv und $g = f^{-1}: V_0 \to U_0$ ihre Umkehrabbildung

mit $J_g(f(y))\, J_f(y) = J_{g \circ f}(y) = Id_{\mathbb{R}^n}$

(Kettenregel)

d.h. $J_g(f(y)) = (J_f(y))^{-1} \ \forall y \in U_0$.

Insbesondere gilt $J_g(b) = (J_f(a))^{-1}$ ∎

Bemerkung 12.6

1) Die Existenz von U_0, V_0 ist ~~neu~~ sogar neu wenn $n = 1$.

2) Wenn $n=1$, ist $(*)$ die bekannte Formel

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

3) Sei $f: U \to \mathbb{R}^m$ ($U \subset \mathbb{R}^n$ offen)
$k \in \mathbb{N}_0$.

Wir sagen, f ist k-mal stetig differenzierbar, wenn alle Komponenten von f k-mal stetig differenzierbar sind.

f heißt unendlich oft ~~de~~ stetig differenzierbar (oder glatt) wenn f k-mal stetig differenzierbar ist $\forall k \in \mathbb{N}_0$.

Für $k \in \mathbb{N}_0 \cup \{\infty\}$ bezeichnen wir mit $\mathcal{C}^k(U, \mathbb{R}^m)$ die Menge aller k-mal stetig differenzierbaren Abbildungen von U nach $\mathbb{R}^m$ (wobei ∞-mal ~~heißt~~ bedeutet unendlich oft)

Nun betrachte $f: U \to V$ ($U, V \subset \mathbb{R}^n$ offen) bijektiv und $f \in \mathcal{C}^k(U, \mathbb{R}^n)$, $k \in \mathbb{N} \cup \{\infty\}$.
Es gelte auch $J_f(a)$ sei invertierbar. $\forall a \in U$.
Nach Satz 12.5 ist $g = f^{-1}: V \to U$ stetig differenzierbar und $J_g(b) = (J_f(a))^{-1}$, $b = f(a)$
Es folgt $g \in \mathcal{C}^k(V, \mathbb{R}^n)$. Man sagt, f $\mathcal{C}^k$-invertierbar ist;

und man nennt f <u>Diffeomorphismus der Klasse C^k.</u>

<u>Beispiel 12.7</u> (wieder Polarkoordinaten in $\mathbb{R}^2$)

Wir wissen bereits dass

$$P: \mathbb{R}_+ \times (0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus \{(x,y) : y = 0, x \geq 0\}$$
$$(r, \varphi) \longmapsto (r\cos\varphi, r\sin\varphi)$$

(Beispiel 10.1.11)

eine Bijektion ist und P stetig differenzierbar (P ist sogar glatt) ist. Zudem gilt

$$\det J_P(r,\varphi) = r \neq 0 \quad \text{d.h.}$$

$J_P(r,\varphi)$ ist invertierbar für $(r,\varphi) \in \mathbb{R}_+ \times (0, 2\pi)$.

Nach Satz über die Umkehrabbildung ist die Umkehrabbildung P^{-1} stetig differenzierbar. wegen der ~~Aus~~ Bemerkung 12.6 ist P^{-1} eingentlich glatt. Man kann diese Aussage wieder bestätigen durch einer expliziten Berechnung von P^{-1} wie folgt:

Setze $\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi \end{cases}$, $(x,y) \in \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\}$

Nach Satz 7.10 sind die Funktionen

$$\exp : \mathbb{R} \to \mathbb{R}, \quad t \mapsto e^t$$

$$\cos : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \cos t, \qquad \sin : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \sin t$$

glatt. Es folgt: die Umkehrfunktionen

$$\log : \mathbb{R}_+ \to \mathbb{R}$$

$$\arcsin : (-1,1) \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

$$\arccos : (-1,1) \to (0,\pi)$$

(wir betrachten arcsin, arccos nur auf $(-1,1)$)

sind auch glatt.

Sei $U_1 := \{(x,y) \in \mathbb{R}^2 : y > 0\}$

$U_2 := \{(x,y) \in \mathbb{R}^2 : y < 0\}$

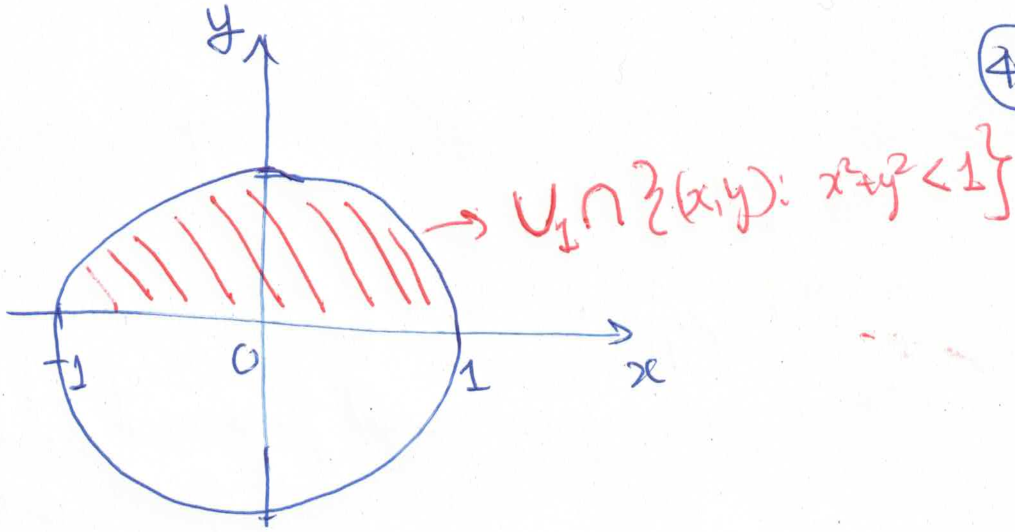
$U_3 := \{(x,y) \in \mathbb{R}^2 : x < 0\}$

Beachte $U_1 \cup U_2 \cup U_3 = \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\}$

Man zeigt, P^{-1} ist glatt auf U_1, U_2, U_3.

$$P^{-1} : U_1 \to \mathbb{R}_+ \times (0, 2\pi)$$

$$(x,y) \mapsto \left(\sqrt{x^2+y^2},\ \arccos \frac{x}{\sqrt{x^2+y^2}}\right)$$



Es folgt: $P^{-1}|_{U_1}$ ist glatt. Analog für $P^{-1}|_{U_2}$, $P^{-1}|_{U_3}$. Man erhält daher die Glattheit von P^{-1}.

In der obigen Begründung hat man die folgende Aussage benutzt:

<u>Behauptung</u>: Seien $W_1 \subset \mathbb{R}^m$, $W_2 \subset \mathbb{R}^k$, $W_3 \subset \mathbb{R}^n$ offene Mengen.
Sei $s \in \mathbb{N}_0 \cup \{\infty\}$
Seien $f: W_1 \to W_2$ s-mal stetig differenzierbar
$g: W_2 \to W_3$ s-mal stetig differenzierbar
Dann ist $g \circ f: W_1 \to W_3$ auch s-mal stetig differenzierbar.

<u>Beweisidee</u>: man benutzt die Kettenregel und eine Induktion über s.

§ 13. Gewöhnliche Differentialgleichungen

§ 13.1. Definitionen und Beispiele

In vielen Bereichen aus Mathematik, Physik sowie Biologie, Wirtschaft, ... muss man sehr oft eine Funktion suchen, die eine Gleichung über ihre (partiellen) Ableitungen erfüllt.

Definition 13.1.1

(i) Eine Differentialgleichung (DGL) ist eine Gleichung, in der Ableitungen einiger Funktionen von einer oder mehreren Variablen auftreten. Die Unbekannten sind dabei die Funktionen.

(ii) Hängen die Funktionen von nur einer reellen Variablen ab, so heißt die DGL gewöhnlich, sonst partielle. In diesem Kapitel betrachten wir nur gewöhnlichen Differentialgleichungen (GDGL)

(iii) Werden mehrere Funktionen gesucht, so spricht man von einem Differentialgleichungenssystem sonst von einer einzel DGL.

(iv) Die allgemeine Form einer GDGL lautet

$$F\,(t, x, x', \dots, x^{(n)}) = 0 \qquad (1)$$

wobei

$$F : I \times \underbrace{\mathbb{R}^m \times \dots \times \mathbb{R}^m}_{(n+1)\text{-mal}} \to \mathbb{R}^m, \quad n \in \mathbb{N}$$

und $I \subset \mathbb{R}$ ein Intervall ist. Eine Lösung von (1) ist eine n-mal differenzierbare Funktion $x : I \to \mathbb{R}^m$ mit $F\,(t, x(t), x'(t), \dots, x^{(n)}(t)) = 0 \quad \forall\, t \in I$.

Die Gleichung (1) nennt man implizit. Sie ist explizit, wenn sie die Form annimmt

$$x^{(n)} = f\,(t, x, x', \dots, x^{(n-1)})$$

wobei $f : I \times (\mathbb{R}^m)^n \to \mathbb{R}^m$

(In diesem Fall ist $F\,(t, x, x', \dots, x^{(n)})$

$$= x^{(n)} - f\,(t, x, x', \dots, x^{(n-1)})\,)$$

Erinnere, wenn $x = (x_1, \dots, x_m) : I \to \mathbb{R}^m$, dann gilt

$$x^{(n)} := (x_1^{(n)}, \dots, x_m^{(n)}), \quad \forall\, n \in \mathbb{N}_0.$$

(v) Die höchste Ordnung eines in der DGL auftretenden Differentials wird **Ordnung** der DGL genannt.

Beispiel:

1) $x'' - 2x' + x = t, \quad t \in \mathbb{R}$

diese DGL ist von Ordnung 2 und sie ist explizit

denn sie ist äquivalent zu $\quad x'' = 2x' - x + t.$

$(m = 1)$
$n = 2$

2) Betrachte die DGL
$$x'(t) = f(t), \quad t \in I$$

wobei $f : I \to \mathbb{R}$ eine stetige Funktion auf einem Intervall $I \subset \mathbb{R}$ ist.

Nach dem Fundamentalsatz der Differential- und Integralrechnung sind alle Lösungen von der Form:
$$x(t) = x_0 + \int_{t_0}^{t} f(s)\,ds, \quad t \in I$$

Dabei ist $x(t_0) = x_0$.

3) Betrachte die DGL $\quad x^2 + x'^2 = -1$

Diese Gleichung hat keine Lösung $x: \mathbb{R} \to \mathbb{R}$

da $x^2 + x'^2 \geq 0$.

4) die DGL $$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

$(x,y) \in U \subset \mathbb{R}^2$, $u: U \to \mathbb{R}$ unbekannt

ist eine partielle DGL, da die gesuchte Funktion von 2 Variablen x, y abhängt.

5) Betrachte ~~das~~ DGLsystem:

$$\begin{cases} x_1' = x_2 \\ x_2' = -x_1 \end{cases}$$

wobei $x_1, x_2: I \to \mathbb{R}$ unbekannte differenzierbare Funktionen. Diese ist ein Differentialgleichungssystem

6) (Fall eines Massenpunktes)

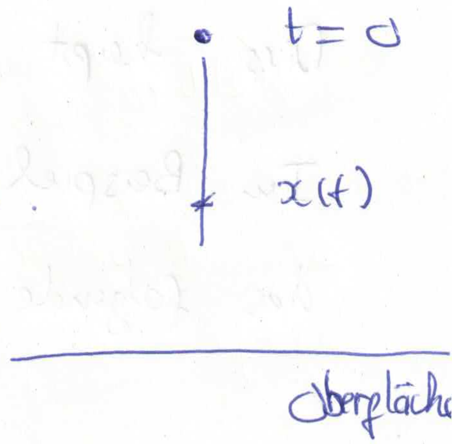
Sei $x(t)$ die Position eines Massenpunktes der Masse m über dem Erdboden zur Zeit t.

Zur Zeit $t=0$ ~~sei~~ sei $\begin{cases} x(0) = x_0 \\ x'(0) = v_0 \end{cases}$

(Man sagt, diese Bedingung ist die Anfangsbedingung

Angenommen, der Massenpunkt befinde sich in geringer Höhe über dem Erdboden, so dass die ~~Erd beschleuder~~ Erdbeschleunigung g an der Oberfläche maßgebend ist. Der Luftwiderstand werde vernachlässigt. Dann gilt

$$x''(t) = -g$$



welche eine GDGL von Ordnung 2 ist.

Die Lösung der GDGL ist

$$x(t) = -\frac{1}{2}t^2 g + c_1 + c_2 t$$

mit Konstanten c_1, c_2. Wegen der Anfangsbedingung ist $c_1 = x_0$ und $c_2 = v_0$, also

$$x(t) = x_0 + v_0 t - \frac{1}{2}t^2 g$$

(Gesetz des freien Falls)

Aus Beispiele sehen wir dass für DGL man oft die Anfangsbedingungen betrachtet.

<u>Definition 13.1.2</u> (Anfangswertprobleme)

Seien ~~t_0~~ $\in I$, $x_0, x_0', \dots, x_0^{(n-1)} \in \mathbb{R}^m$.

Zu bestimmen sind alle Funktionen $x : I \to \mathbb{R}^m$

mit
$$\begin{cases} F(t, x(t), x'(t), \ldots, x^{(n)}(t)) = 0 \\ \left.\begin{aligned} x(t_0) &= x_0 \\ x'(t_0) &= x_0' \\ &\vdots \\ x^{(n-1)}(t_0) &= x_0^{(n-1)} \end{aligned}\right\} \text{Anfangsbedingungen} \end{cases}$$

Dies heißt eine Anfangswertproblem (AWP).

Im Beispiel des freien Falls haben wir das folgende AWP besprochen:
$$\begin{cases} x'' = -g \\ x(0) = x_0 \\ x'(0) = v_0 \end{cases}$$

<u>Bemerkung 13.1.3</u> (Äquivalenz zwischen der expliziten DGL n-ter Ordnung und DGL-Systeme ~~Erster~~ erster Ordnung)

Betrachte die explizite DGL n-ter Ordnung
$$x^{(n)} = f(t, x, x', \ldots, x^{(n-1)}) \qquad (2)$$
wobei $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R} \times (\mathbb{R}^m)^n$ offen.

Wir führen neue Unbekannte Funktionen $u_1, u_2, \ldots, u_n$ ein durch
$$u_1 := x, \; u_2 := x', \; \ldots, \; u_n := x^{(n-1)}$$
und bezeichnen $u := (u_1, \ldots, u_n) \in (\mathbb{R}^m)^n$.

Dann ist (2) äquivalent zu

$$\begin{cases} u_1' = u_2 \\ u_2' = u_3 \\ \vdots \\ u_{n-1}' = u_n \\ u_n' = f(t, u_1, \dots, u_n) \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} u_1' \\ u_2' \\ \vdots \\ u_{n-1}' \\ u_n' \end{pmatrix} = \begin{pmatrix} u_2 \\ u_3 \\ \vdots \\ u_n \\ f(t,u) \end{pmatrix} \qquad (3)$$

In der Tat wenn $x : I \to \mathbb{R}^m$ (2) löst, so löst $u = (x, x', \dots, x^{(n-1)})$ das System (3). Umgekehrt löst u (3), so löst $x := u_1$ die DGL (2).

Die Anfangsbedingungen für (2)

$$x(t_0) = x_0,\ x'(t_0) = x_0', \dots, x^{(n-1)}(t_0) = x_0^{(n-1)}$$

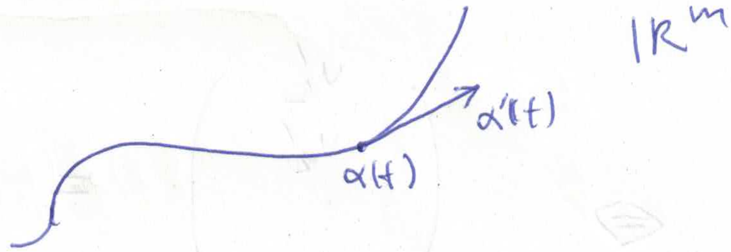
übertragen sich in Anfangsbedingungen für (3):

$$u(t_0) = (x_0, x_0', \dots, x_0^{(n-1)})$$

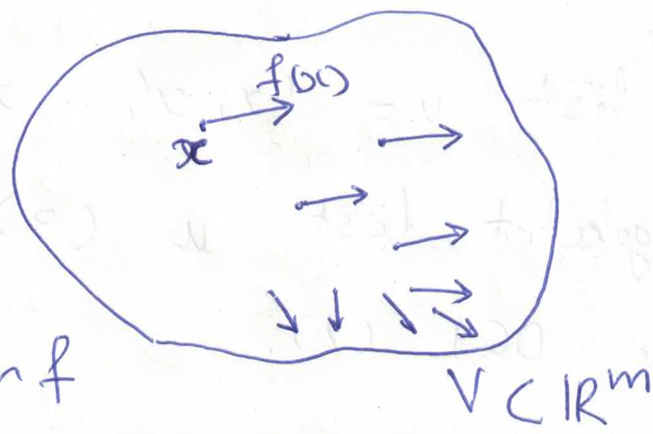
Geometrische Interpretation einer DGL erster Ordnung.

Sei $\alpha: I \to \mathbb{R}^m$ eine differenzierbare Kurve
$t \mapsto (\alpha_1(t), \dots, \alpha_m(t))$

Der Tangentialvektor $\alpha'(t) = (\alpha_1'(t), \dots, \alpha_m'(t))$ heißt auch Geschwindigkeitsvektor zur Zeit t



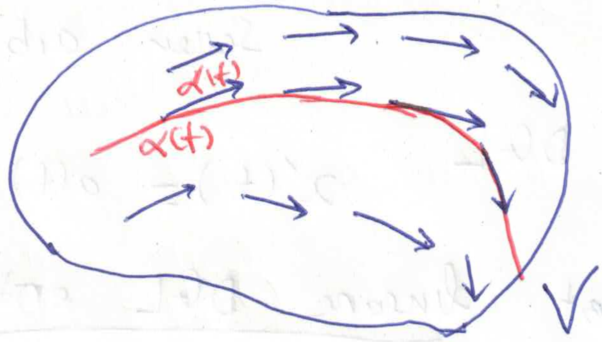
Sei $V \subset \mathbb{R}^m$ eine offene Menge. Ein Vektorfeld auf V ist eine Abbildung $f: V \to \mathbb{R}^m$.



Eine Integralkurve von f ist eine differenzierbare Kurve $\alpha: I \to V$ mit $\alpha'(t) = f(\alpha(t))$ $\forall t \in I$, d.h. der Geschwindigkeitsvektor zur Zeit t ist derjenige Vektor, der dem Punkt $\alpha(t)$ von dem Vektorfeld zugeordnet wird.

das bedeutet, $\alpha: I \to V$ ist genau eine Lösung der DGL $x'(t) = f(x(t))$.

Sie ist eine explizite DGL, deren rechte Seite unabhängig von t ist. Eine solche DGL heißt **autonom**.



Allgemeiner betrachten wir die DGL

$$x' = f(t,x) \quad \text{wobei}$$

$$f : U \to \mathbb{R}^n, \quad U \subset \mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n \text{ offen ist}$$
$$(t,x) \mapsto f(t,x) \qquad\qquad (t,x)$$

Definiere

$$F : U \to \mathbb{R}^{n+1}$$
$$(t,x) \mapsto (1, f(t,x))$$

$$\tilde{x} : I \to U$$
$$t \mapsto (t, x(t))$$

Das Bild $\tilde{x}(I)$ ist genau der Graph von $x : I \to \mathbb{R}^n$

Es gilt $\tilde{x}'(t) = (1, x'(t)) = (1, f(t,x))$
$$= F(t,x)$$

Damit ist $\tilde{x}$ eine Integralkurve von Vektorfeldes F.

§ 13.2. Elementare Lösungsmethoden

Definition 13.2.1 Sei $I \subset \mathbb{R}$ ein Intervall und seien $a, b : I \to \mathbb{R}$ stetig.

Die DGL
$$x'(t) = a(t)\,x(t) + b(t) \qquad (*)$$
heißt lineare DGL erster Ordnung.

Die DGL
$$x'(t) = a(t)\,x(t) \qquad (**)$$
heißt homogene lineare DGL erster Ordnung

Bemerkung 13.2.2 Sei $x : I \to \mathbb{R}$ eine Lösung von $x'(t) = a(t)x(t) + b(t)$ und $x_p : I \to \mathbb{R}$ eine Lösung von $x_p'(t) = a(t)\,x_p(t)$.

Dann gilt: $x + x_p$ ist Lösung von $(*)$.

Umgekehrt für jede Lösung $\tilde{x}$ von $(*)$ ist $\tilde{x} - x_p$ eine Lösung von $(**)$.

Wir lösen zuerst $(**)$.

Satz 13.2.3 Seien $t_0 \in I$, $x_0 \in \mathbb{R}$. Dann gibt es genau eine Lösung von $(**)$ mit $x(t_0) = x_0$, nämlich $x: I \to \mathbb{R}$

$$x(t) = x_0 \exp\left(\int_{t_0}^{t} a(s)\,ds \right) \quad \forall t \in I$$

<u>Beweis</u>: Sei $A: I \to \mathbb{R}$

$$t \mapsto \int_{t_0}^{t} a(s)\,ds$$

Dann ist $A'(t) = a(t) \quad \forall t$, und $A(t_0) = 0$.

Damit gilt $x(t) = x_0 e^{A(t)}$.

Man beachte $x(t_0) = x_0$

$$x'(t) = x_0 A'(t) e^{A(t)}$$

$$= x(t)\, a(t)$$

Wir zeigen die Eindeutigkeit der Lösung.

Sei $u: I \to \mathbb{R}$ eine Lösung von $(**)$ mit $u(t_0) = x_0$.

Betrachte die Funktion $g: I \to \mathbb{R}$ mit

$g(t) = u(t) e^{-A(t)}$. Es gilt

$$g'(t) = \frac{d}{dt}\left(u(t) e^{-A(t)} \right)$$

$$= u'(t) e^{-A(t)} - u(t) A'(t) e^{-A(t)}$$

$$= a(t) u(t) e^{-A(t)} - u(t) a(t) e^{-A(t)}$$

$$= 0$$

Es folgt: g ist eine konstante Funktion, d.h.

$\exists\, c \in \mathbb{R}$ mit $u(t)\, e^{-A(t)} = c \quad \forall t \in I$

Ersetzen $t = t_0$ bekommen wir $c = u(t_0) = x_0$.

Daher $u(t) = x_0\, e^{A(t)} = x(t)$ ∎

Wir möchten nun die inhomogene DGL (*) auflösen.
Nach Bemerkung 13.2.2 genügt es eine partikuläre Lösung von (*) zu finden. Dazu benutzt man die Methode von „Variation der Konstante". Die Idee ist wie folgt:

die Lösung von (**) ist von der Form

$x(t) = c\, e^{A(t)}$ wobei c eine Konstante

ist. Wir suchen eine Lösung $u(t) = c(t)\, e^{A(t)}$ von (*) wobei $c(t)$ eine Funktion von t ist.

Es gilt
$$\begin{aligned} u'(t) &= c'(t)\, e^{A(t)} + c(t)\, A'(t)\, e^{A(t)} \\ &= c'(t)\, e^{A(t)} + c(t)\, a(t)\, e^{A(t)} \\ &= c'(t)\, e^{A(t)} + u(t)\, a(t) \end{aligned}$$

Da $u'(t) = a(t)u(t) + b(t)$, bekommt man

$$b(t) = c'(t)\, e^{A(t)} \iff c'(t) = e^{-A(t)}\, b(t)$$

Wir wählen $c(t) = \int_{t_0}^{t} b(s)\, e^{-A(s)}\, ds$

und erhalten eine partikuläre Lösung x_p von (*) mit $x_p(t_0) = 0$, $x_p(t) = \left(\int_{t_0}^{t} b(s) e^{-A(s)} ds\right) e^{A(t)}$

Daraus folgt

Satz 13.2.4 Das AWP $\left\{ \begin{array}{l} x'(t) = a(t)x(t) + b(t) \\ x(t_0) = x_0 \end{array} \right.$

hat eine eindeutige Lösung

$$(1) \quad x(t) = \left(x_0 + \int_{t_0}^{t} b(s)\, e^{-A(s)} ds \right) e^{A(t)}$$

wobei $A(t) = \int_{t_0}^{t} a(s)ds$.

Beweis Die Formel (1) zeigt

$$x(t) = x_0 e^{A(t)} + x_p(t)$$

daher gilt $x(t_0) = x_0$ und da $x_0 e^{A(t)}$ eine Lösung von (**), nach Bemerkung 13.2.2 ist $x(t)$ eine Lösung von dem AWP.

Wir zeigen die Eindeutigkeit von der Lösung.

Sei $\tilde{x}(t)$ eine Lösung des AWPs
$\tilde{x}(t_0) = x_0$ und $\tilde{x}'(t) = a(t)\tilde{x}(t) + b(t)$.

Also gilt: $x - \tilde{x}$ ist eine Lösung von (**) mit $(x - \tilde{x})(t_0) = 0$. Nach Satz 13.2.3 ist $x = \tilde{x}$ auf I. ■

Beispiel: wir betrachten die DGL

$$x' = 2xt + t, \quad t \in \mathbb{R}$$

$(a(t) = 2t,\ b(t) = t)$

Die homogene Gleichung $\begin{cases} x' = 2xt \\ x(0) = x_0 \end{cases}$ besitzt die Lösung

$$\varphi(t) = x_0 \exp\left(\int_0^t 2s\,ds\right) = x_0 e^{t^2}$$

Deshalb hat die DGL $x' = 2xt + t$ mit der Anfangsbedingung $x(0) = x_0$ die eindeutige Lösung

$$x(t) = \left(x_0 + \int_0^t s e^{-s^2} ds\right) e^{t^2}$$

$$= \left(x_0 + \frac{-e^{-s^2}}{2}\Big|_0^t\right) e^{t^2}$$

$$= \left(x_0 + \frac{1}{2} - \frac{e^{t^2}}{2}\right) e^{t^2}$$

<u>Definition 13.2.5</u> Seien $I, J \subset \mathbb{R}$ nicht-leere offene Intervalle. Seien $g: I \to \mathbb{R}$, $h: J \to \mathbb{R}$ stetige Funktionen. Sei $f: I \times J \to \mathbb{R}$

$$(t,x) \mapsto g(t)\,h(x).$$

Es werde vorausgesetzt, dass $h(x) \neq 0 \quad \forall x \in J$.

Die DGL $\quad x' = f(t,x) = g(t)\,h(x)$

heißt eine <u>DGL mit getrennten Variablen</u>.

<u>Satz 13.2.6</u> Seien die Notationen wie in der Definition 13.2.5. Sei $(t_0, x_0) \in I \times J$. Definiere $G: I \to \mathbb{R}$, $H: J \to \mathbb{R}$ durch

$$G(t) := \int_{t_0}^{t} g(s)\,ds, \quad H(x) := \int_{x_0}^{x} \frac{du}{h(u)}$$

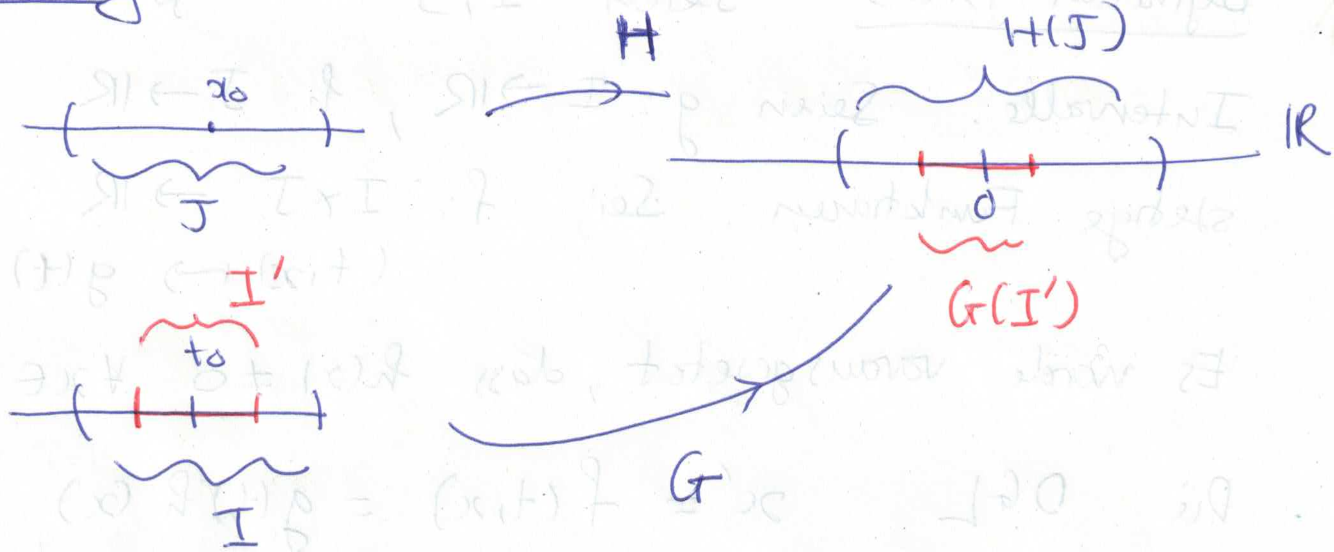
Es sei $I' \subset I$ ein Intervall mit $t_0 \in I'$ und $G(I') \subset H(J)$. Dann besitzt das AWP

$$\begin{cases} x' = f(t,x) \\ x(t_0) = x_0 \end{cases}$$

genau eine eindeutige Lösung $x: I' \to \mathbb{R}$ und x genügt

$$(1) \quad G(t) = H(x(t)) \quad \forall t \in I'.$$

Bemerkung:



Da $H'(x) = \frac{1}{h(x)} \neq 0 \quad \forall x \in J$, ist H streng monoton auf J. Deshalb ist $H(J)$ eine ~~offen~~ Umgebung von $H(x_0) = 0$. Denn $G(t_0) = 0$, man sieht dass es immer eine Intervall I' gibt mit $\begin{cases} t_0 \in I' \\ G(I') \subset H(J) \end{cases}$ wegen der Stetigkeit von G.

Beweis von Satz 13.2.6

Wir zeigen zunächst: ist $\varphi : I' \to \mathbb{R}$ eine Lösung des AWPs $\begin{cases} x' = f(t,x) \\ x(t_0) = x_0 \end{cases}$, so gilt (*).

Aus $\varphi'(t) = f(t, \varphi(t)) = g(t)\, h(\varphi(t))$

folgt
$$\frac{\varphi'(t)}{h(\varphi(t))} = g(t) \quad \forall t \in I'$$

Es folgt

$$H(\varphi(t)) = \int_{t_0}^{t} \frac{\varphi'(s)\, ds}{h(\varphi(s))} = \int_{t_0}^{t} g(s)\,ds = G(t)$$

(Durch Substitution $x = \varphi(s)$)

d.h. (1) gilt.

Erinnere, dass H streng monoton ist. Sei $H^{-1} : H(J) \to J$ die Umkehrabbildung von H.

Da H stetig differenzierbar ist, so gilt: H^{-1} ist auch stetig differenzierbar. Aus (1) folgt

(2) $\varphi(t) = H^{-1}(G(t))$ da $G(t) \in H(J)$ $\forall t \in I'$.

das zeigt, φ eindeutig bestimmt ist falls φ (eine Lösung) existiert.

Existenz der Lösung:

Wir nehmen (2) als Definition einer Funktion $\varphi : I' \to \mathbb{R}$. Sie ist stetig differenzierbar da H^{-1}, G stetig differenzierbar sind. Zudem gilt

$\varphi(t_0) = H^{-1}(G(t_0)) = H^{-1}(0) = x_0$.

d.h. φ löst das AWA $\begin{cases} x' = f(t,x) \\ x(t_0) = x_0 \end{cases}$

$\blacksquare$

Beispiel: $x' = tx \quad (g(t) = t,\ h(x) = x$
$I = \mathbb{R},\ J = (0,\infty))$

Sei $t_0 \in \mathbb{R}$, $x_0 \in (0,\infty)$. Beachte, $h(x) \neq 0$ $\forall x \in J$. Aus $x' = tx$ folgt

$$\frac{x'}{x} = t \Rightarrow \int_{t_0}^{t} \frac{x'(t)\,dt}{x(t)} = \int_{t_0}^{t} s\,ds = \frac{t^2 - t_0^2}{2}.$$

Man erhält $\displaystyle\int_{x_0}^{x(t)} \frac{du}{u} = \frac{t^2 - t_0^2}{2}$

die linke Seite ist gleich $\log u \Big|_{x_0}^{x} = \log x - \log x_0$.

Deshalb gilt $\displaystyle\log x(t) = \frac{t^2 - t_0^2}{2} + \log x_0$.

$$x(t) = x_0\, e^{\frac{t^2 - t_0^2}{2}}.$$

Man kann natürlich die Funktion G, H explizit berechnen und ~~wen~~ den Satz 13.2.6 direkt anwenden. Aber die obige Präsentation ist einprägsamer.

Definition 13.2.7 Sei $I \subset \mathbb{R}\setminus\{0\}$, $J \subset \mathbb{R}$ Intervalle, $\varphi: J \to \mathbb{R}$ stetig. Sei $U = \{(t,x) \in I \times \mathbb{R} : \frac{x}{t} \in J\}$. Sei $f: U \to \mathbb{R}$ mit $f(t,x) = \varphi(\frac{x}{t})$.

Die DGL
$$x'(t) = f(t, x(t)) = \varphi\left(\frac{x(t)}{t}\right)$$
heißt homogene DGL (nicht zu verwechseln mit den homogenen linearen DGL). Der Name kommt aus der Tatsache, dass die Funktion f homogen (von Grad 1) ist, d.h.
$$f(\lambda t, \lambda x) = \lambda f(t, x) \quad \forall \lambda \in \mathbb{R}\setminus\{0\}$$
mit $(t, x) \in U$, $(\lambda t, \lambda x) \in U$.

Satz 13.2.8 Sei $(t_0, x_0) \in U$. Dann gilt:
eine Funktion $x: I \to \mathbb{R}$ eine Lösung von
$$\begin{cases} x' = f(t, x) \\ x(t_0) = x_0 \end{cases}$$
genau dann wenn $u: I \to \mathbb{R}$
$$t \mapsto \frac{x(t)}{t}$$
eine Lösung von
$$(*) \begin{cases} u'(t) = \frac{1}{t}\left(\varphi(u(t)) - u(t)\right) \\ u(t_0) = \frac{x_0}{t_0} \end{cases}$$

Beweis: "$\Rightarrow$" da $u(t) = \frac{x(t)}{t}$, gilt
$$u'(t) = \frac{x'(t) \cdot t - x(t)}{t^2} = \frac{1}{t}\left(x'(t) - \frac{x(t)}{t^2}\right)$$
$$= \frac{1}{t}\left(\varphi(u(t)) - u(t)\right)$$

Sei umgekehrt vorausgesetzt, dass u eine Lösung von $(*)$ ist. Dann gilt

$$(x(t))' = \frac{d}{dt}(t\,u(t)) = u(t) + t\,u'(t)$$

$$= \varphi(u(t)) = f(t, x(t)).$$

<u>Beispiel</u>: Betrachte

$$x'(t) = 1 + \frac{x(t)}{t} + \frac{x^2(t)}{t^2}$$

$\varphi(u) = 1 + u + u^2$

$(I = \mathbb{R}_+, J = \mathbb{R})$. Setze $u(t) := \frac{x(t)}{t}$.

Dann gilt

$$u'(t) = \frac{1}{t}\left(\varphi(u(t)) - u(t)\right) = \frac{1}{t}\left(1 + u^2(t)\right)$$

Es folgt $\frac{u'}{1+u^2} = \frac{1}{t}$ (Trennung der Variablen)

$$\int_{u_0}^{u(t)} \frac{dv}{1+v^2} = \int_{t_0}^{t} \frac{u'(s)}{1+u^2(s)}\,ds = \int_{t_0}^{t} \frac{ds}{s}, \text{ wobei } u_0 := \frac{x_0}{t_0}$$

Man bekommt $\arctan u(t) - \arctan u_0$

$$= \log t - \log t_0 \quad \left(u_0 = \frac{x_0}{t_0}\right)$$

$$\Rightarrow \quad u(t) = \tan\left(\log t - \log t_0 + \arctan u_0\right)$$

Folglich gilt

$$x(t) = t u(t)$$
$$= t \cdot \tan\left(\log t - \log t_0 + \arctan \frac{x_0}{t_0} \right)$$

Da $\arctan u(t) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, gilt

$$\log t - \log t_0 + \arctan \frac{x_0}{t_0} \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Das impliziert, die Lösung x nur definiert in der Menge $\left\{ t \in \mathbb{R}_+ : \log t - \log t_0 + \arctan \frac{x_0}{t_0} \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \right\}$

(die Menge ist eine offene Umgebung von t_0 in $\mathbb{R}_+$)

§ 13.3. Existenz- und Eindeutigkeitssatz

Im Allgemeinen kann man DGLs nicht explizit auflösen. Man kann aber die Existenz und Eigenschaften der Lösungen ~~unter~~ untersuchen. Ziel ist es, wir besprechen grundlegende Sätze in der Theorie der DGL über Existenz und Eindeutigkeit der Lösungen der DGLs.

Satz 13.3.1 Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ und $f: U \to \mathbb{R}^n$ stetig. Weiter sei $(t_0, x_0) \in U$ gegeben. Dann gilt:
Eine stetige Funktion $\varphi: I \to \mathbb{R}^n$, welche auf einem Intervall $I \subset \mathbb{R}$ mit $t_0 \in I$ definiert ist, und deren Graph in U enthalten ist (d.h. $(t, \varphi(t)) \in U \;\; \forall t \in I$), ist genau dann Lösung der DGL $\begin{cases} x' = f(t,x) \\ x(t_0) = x_0 \end{cases}$

wenn die folgende Integralgleichung gilt:

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds \quad \forall t \in I$$

<u>Beweis</u> Sei die Integralgleichung erfüllt.

Setzt man darin $t = t_0$, ergibt sich $\varphi(t_0) = x_0$.

Da f stetig ist, nach dem Fundamentalsatz der Differential- und Integralrechnung gilt

$$\varphi'(t) = \frac{d}{dt} \int_{t_0}^{t} f(s, \varphi(s))\, ds$$

$$= f(t, \varphi(t))$$

Daher erfüllt φ die DGL $\varphi'(t) = f(t, \varphi(t))$ mit der Anfangsbedingung $\varphi(t_0) = x_0$.

Sei umgekehrt φ Lösung der DGL mit $\varphi(t_0) = x_0$. Insbesondere ist φ differenzierbar, mit $\varphi'(t) = f(t, \varphi(t))$. Es folgt

$$\int_{t_0}^{t} \varphi'(s)\, ds = \int_{t_0}^{t} f(s, \varphi(s))\, ds$$

$$\| $$

$$\varphi(t) - \varphi(t_0).$$

Dann gilt $\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds$ ∎

Definition 13.3.2 Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ und

$$f: U \to \mathbb{R}^n$$
$$(t,x) \mapsto f(t,x)$$

Man sagt, f genüge in U einer Lipschitz-Bedingung (bzgl. der Variable x) mit der Lipschitz-Konstanten $L \geq 0$, wenn

$$\| f(t,x) - f(t,\tilde{x}) \| \leq L \| x - \tilde{x} \|$$

$\forall (t,x), (t,\tilde{x}) \in U$.

Man sagt, f genüge in U lokal einer Lipschitz-Bedingung (bzgl. der Variable x) falls jeder Punkt $(t,x) \in U$ eine Umgebung V besitzt, so dass f in $U \cap V$ einer Lipschitz-Bedingung mit einer gewissen (von V abhängigen) Konstanten $L \geq 0$ genügt.

Der folgende Satz liefert ein nützliches Kriterium für die Lipschitz-Bedingung.

Satz 13.3.3 Sei $U \subset \mathbb{R} \times \mathbb{R}^n$

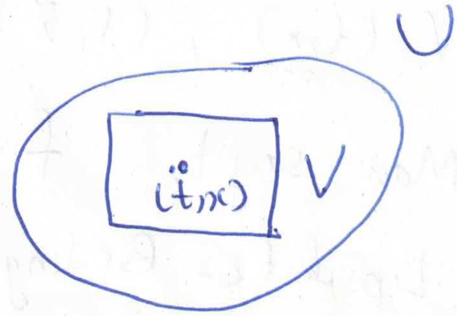
$$f : U \to \mathbb{R}^n$$
$$(t,x) \mapsto f(t,x)$$

Ist f stetig partiell differenzierbar bezüglich der Variable $x = (x_1, x_2, \dots, x_n)$, so genügt f in U lokal einer Lipschitz-Bedingung (bzgl. x)

Beweis Sei $(t_0, x_0) \in U$. Da U offen ist, $\exists\, r > 0$ mit:

$$V := \left\{ (t,x) \in \mathbb{R} \times \mathbb{R}^n : \begin{array}{l} |t - t_0| \leq r \\ \|x - x_0\| \leq r \end{array} \right\}$$

liegt ganz in U.



Beachte, V ist kompakt.

Zudem sind $\frac{\partial f}{\partial x_j}(t,x)$ (nach Voraussetzung) $\forall\, j = 1, 2, \dots, n$ stetig auf V

Es folgt:

$$L := \sup_{(t,x) \in V} \| \underbrace{\frac{\partial f}{\partial x}}_{\text{eine } n\times n\text{-Matrix}}(t,x) \| < \infty$$

Folgerung 10.2.4 impliziert

$$\|f(t,x) - f(t,\tilde{x})\| \leq L\|x - \tilde{x}\|$$

$\forall\, (t,x), (t,\tilde{x}) \in V$. ∎

Satz 13.3.4 (Eindeutigkeitssatz)

Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ und $f: U \to \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung (bzgl. x) genügt. Seien

$\varphi, \psi : I \to \mathbb{R}^n$ zwei Lösungen der DGL

$$x' = f(t,x)$$

auf einem Intervall $I \subset \mathbb{R}$. Gilt dann

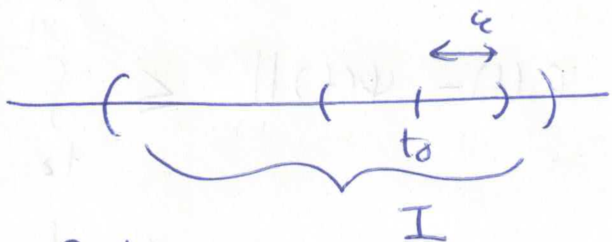
$$\varphi(t_0) = \psi(t_0) \quad \text{für ein } t_0 \in I$$

so folgt $\quad \varphi(t) = \psi(t) \quad \forall\, t \in I$.

Beweis: Sei $\varphi(t_0) = \psi(t_0)$ für ein $t_0 \in I$.

Schritt 1 Wir zeigen dass es ein $\varepsilon > 0$ gibt mit

$$\varphi(t) = \psi(t) \quad \forall\, t \in (t_0 - \varepsilon, t_0 + \varepsilon) \cap I$$



Da f lokal einer Lipschitz-Bedingung (bzgl. x) genügt, $\exists$ eine Konstante $L \geq 0$, eine $\delta > 0$, so dass

$$\| f(t,\varphi(t)) - f(t,\psi(t)) \| \leq L \| \varphi(t) - \psi(t) \| \quad (1)$$

$\forall\, t \in I \cap B_\delta(t_0) = \{ t \in I : |t - t_0| < \delta \}$.

Wir können δ nehmen so dass $I \cap \overline{B_\delta(t_0)}$ kompakt ist.

Da φ, ψ stetig auf I sind, sind φ, ψ beschränkt auf $I \cap \overline{B_\delta(t_0)}$. Daher ist

$$M_\varepsilon := \sup\{ \| \varphi(t) - \psi(t) \| : t \in [t_0-\varepsilon, t_0+\varepsilon] \cap I \}$$

endlich $\forall \varepsilon \in (0, \delta]$.

Sei nun $\varepsilon := \min\{\delta, \frac{1}{2L}\}$.

Nach Satz 13.3.1 gilt

$$\varphi(t) = \varphi(t_0) + \int_{t_0}^{t} f(s, \varphi(s))\, ds$$

$$\psi(t) = \psi(t_0) + \int_{t_0}^{t} f(s, \psi(s))\, ds$$

wegen $\varphi(t_0) = \psi(t_0)$ erhält man

$$\varphi(t) - \psi(t) = \int_{t_0}^{t} \left(f(s, \varphi(s)) - f(s, \psi(s)) \right) ds$$

Aus diese und (1) folgt $\forall t \in I \cap \overline{B_\varepsilon(t_0)}$

$$\| \varphi(t) - \psi(t) \| \leq \int_{t_0}^{t} \| f(s, \varphi(s)) - f(s, \psi(s)) \|\, ds$$

$$\leq L \int_{t_0}^{t} \| \varphi(s) - \psi(s) \|\, ds$$

$$\leq L |t - t_0| \cdot M_\varepsilon \leq L \cdot \varepsilon \cdot M_\varepsilon \leq \frac{M_\varepsilon}{2}$$

denn $\varepsilon \leq \frac{1}{2L}$, $\varepsilon \leq \delta$.

Daher gilt

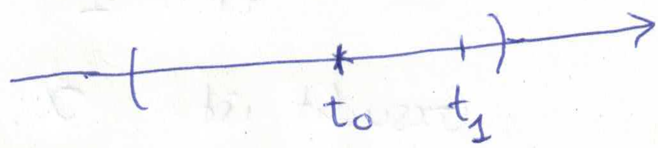
$$M_\varepsilon \leq \frac{M_\varepsilon}{2} \Rightarrow \frac{M_\varepsilon}{2} \leq 0$$

Es folgt $M_\varepsilon = 0$ (da $M_\varepsilon \geq 0$). Anders gesagt

$\varphi(t) = \psi(t) \quad \forall\, t \in I \cap \overline{B_\varepsilon}(t_0)$.

Schritt 2 Wir zeigen jetzt $\varphi(t) = \psi(t) \quad \forall\, t \in I, \; t \geq t_0$

Sei $t_1 := \sup\{ t \in I,\ t \geq t_0 : \varphi|_{[t_0, t]} \equiv \psi|_{[t_0, t]} \}$



Falls $t_1 = \infty$ oder t_1 gleich dem rechten Intervallende von I ist, sind wir fertig. Andernfalls gibt es $\delta > 0$, sodass $[t_1, t_1 + \delta] \subset I$.

Da φ, ψ stetig sind, gilt $\varphi(t_1) = \psi(t_1)$
(Nehmen wir nur eine Folge $(\tilde{t}_k)_{k \in \mathbb{N}}$, welche gegen t_1 konvergiert, mit $\varphi|_{[t_0, \tilde{t}_k]} \equiv \psi|_{[t_0, \tilde{t}_k]}$)

Nach dem Schritt 1 $\exists\, \varepsilon > 0$ so dass

$\varphi(t) = \psi(t) \quad \forall\, t \in \overline{B_\varepsilon}(t_1) \cap I$

Dies steht im Widerspruch zur Definition von t_1.

Daher $\varphi(t) = \psi(t) \; \forall\, t \geq t_0, \; t \in I$.

analog bekommen wir $\varphi(t) = \psi(t) \; \forall t \leq t_0, \; t \in I$.

Der Beweis ist daher fertig. ■

Beispiel 13.3.5

1) Betrachte $\begin{cases} x' = \sqrt{x} \\ x(t_0) = x_0 \end{cases}$, $\quad t_0 \in \mathbb{R}, \; x_0 \in (0,\infty)$

Diese DGL ist eine DGL mit getrennten Variablen

$$g: \mathbb{R} \to \mathbb{R}, \; t \mapsto 1 \qquad , \qquad h: (0,\infty) \to \mathbb{R}, \; x \mapsto \sqrt{x}$$

Gesucht ist $x: I' \to (0,\infty)$ mit $t_0 \in I'$

und $\begin{cases} x'(t) = \sqrt{x(t)} \quad (1) \\ x(t_0) = x_0 \end{cases}$

Aus (1) folgt $\frac{x'}{\sqrt{x}} = 1 \Rightarrow$

$$\int_{t_0}^{t} 1 \, ds = \int_{x_0}^{x(t)} \frac{du}{\sqrt{u}} = 2\sqrt{u}\Big|_{x_0}^{x(t)} = 2\left(\sqrt{x(t)} - \sqrt{x_0}\right)$$

Daher

$$t - t_0 = 2\left(\sqrt{x(t)} - \sqrt{x_0}\right)$$

$$\Leftrightarrow \quad \sqrt{x(t)} = \frac{1}{2}(t - t_0) + \sqrt{x_0}$$

Da $\sqrt{x(t)} > 0$, gilt $t - t_0 + 2\sqrt{x_0} > 0$

äquivalent $t > t_0 - 2\sqrt{x_0}$

Man bekommt $x: (t_0 - 2\sqrt{x_0}, \infty) \to \infty$
(2) $\qquad t \mapsto \frac{1}{4}(t - t_0 + 2\sqrt{x_0})^2$

$(I' = (t_0 - 2\sqrt{x_0}, \infty))$

wir können diese Funktion x mit Null links von $t_0 - 2\sqrt{x_0}$ fortsetzen und eine Lösung auf $\mathbb{R}$ erhalten:

$$x : \mathbb{R} \to \mathbb{R}$$

$$t \longmapsto \frac{1}{4}\left(t - t_0 + 2\sqrt{x_0}\right)^2, \quad t \geq t_0 - 2\sqrt{x_0}$$

(2)

$$t \longmapsto 0 \quad \text{wenn } t < t_0 - 2\sqrt{x_0}$$

Was passiert wenn $x_0 = 0$?

Man beachte: $x(t) \equiv 0 \quad (t \in \mathbb{R})$ ist eine Lösung von AWPs $\begin{cases} x' = \sqrt{x} \\ x(t_0) = 0 \end{cases}$

Andererseits ist die durch (2) definierte (wenn $x_0 = 0$) ~~definierte~~ Funktion auch eine Lösung des AWPs.

Daher bekommt man zwei Lösungen für das AWP $\begin{cases} x' = \sqrt{x} \\ x(t_0) = 0 \end{cases}$

Dies ~~steht~~ widerspricht dem Satz 13.3.4 nicht da die Funktion $f(t,x) = \sqrt{x}$ nicht die Lipschitz-Bedingung (bzgl. x) in einer Umgebung von $x_0 = 0$ erfüllt.

2) Betrachte das AWP

$$\begin{cases} x' = x^2 \sin x + xt - tx^2 \\ x(0) = 0 \end{cases}$$

Man hat $f(t,x) = x^2 \sin x + xt - tx^2$ für $(t,x) \in \mathbb{R}^2$

Beachte dass $\varphi: \mathbb{R} \to \mathbb{R}$, $t \mapsto 0$ $(\varphi(t) \equiv 0 \ \forall t \in \mathbb{R})$

eine Lösung des AWPs. Zudem da f partiell differenzierbar bzgl. x ist, genügt f lokal einer Lipschitz-Bedingung (bzgl. x). Nach Satz 13.3.4 wenn $\psi: \mathbb{R} \to \mathbb{R}$ eine Lösung des AWPs ist, dann gilt $\psi = \varphi$. d.h. φ ist die ~~einzige~~ einzige Lösung des AWPs.

Satz 13.3.6 (Existenzsatz von Picard-Lindelöf)

Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f: U \to \mathbb{R}^n$, $(t,x) \mapsto f(t,x)$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung bzgl. x besitzt. Sei $(t_0, x_0) \in U$.

Dann gibt es eine $\varepsilon > 0$ und eine Lösung

...

$$\varphi : [t_0-\varepsilon, t_0+\varepsilon] \to \mathbb{R}, \quad t \mapsto \varphi(t)$$

von AWPs

$$(*) \begin{cases} x' = f(t,x) \\ x(t_0) = x_0 \end{cases}$$

<u>Bemerkung:</u> Nach Satz 13.3.4 ist die Lösung φ eindeutig bestimmt.

<u>Beweis:</u> Wir wenden den Banachschen Fixpunktsatzes an. Zu diesem Zwecke, beachte, nach Satz 13.3.1, dass: φ ist eine Lösung von $(*)$ $\Leftrightarrow$ gilt:

$$\varphi(t) = x_0 + \int_t^{t_0} f(s, \varphi(s))\,ds$$

$\forall t \in [t_0-\varepsilon, t_0+\varepsilon]$.

Sei

$$\mathcal{C}([t_0-\varepsilon, t_0+\varepsilon], \mathbb{R}^n) = \{\text{stetige Funktionen } \psi : [t_0-\varepsilon, t_0+\varepsilon] \to \mathbb{R}^n\}$$

(für ein noch zu bestimmendes $\varepsilon > 0$) mit der Supremumsnorm $\|\psi\| := \sup\{\|\psi(t)\| : |t-t_0| \le \varepsilon\}$

(da $[t_0-\varepsilon, t_0+\varepsilon]$ kompakt ist, ist ψ beschränkt).

Nach Voraussetzung über f, $\exists\, \delta > 0, r > 0$ so dass

$$Q_{\delta,r} := \{ (t,x) \in \mathbb{R}\times\mathbb{R}^n : \; |t-t_0| \le \delta, \; |x-x_0| \le r \} \subset U$$

und f in $Q_{\delta,r}$ einer Lipschitz-Bedingung mit einer gewissen Konstanten $L>0$ genügt.

Da f stetig und $Q_{\delta,r}$ kompakt ist, $\exists\, M>0$ mit $\|f(t,x)\| \le M \;\; \forall\; (t,x) \in Q_{\delta,r}$.

Wir setzen jetzt

$$\varepsilon := \min\{\delta, \frac{r}{M}, \frac{1}{2L}\}$$

und $A := \{ \varphi \in \mathcal{C}([t_0-\varepsilon, t_0+\varepsilon], \mathbb{R}^n) : \|\varphi - x_0\| \le r \}$

welche eine abgeschlossene Teilmenge von $\mathcal{C}([t_0-\varepsilon, t_0+\varepsilon], \mathbb{R}^n)$ ist.

<u>Behauptung</u>: die Abbildung $T: A \to A$, die einer Funktion $\varphi \in A$ die Funktion $\eta := T(\varphi)$,

$$\eta(t) := x_0 + \int_{t_0}^{t} f(s,\varphi(s))\, ds$$

$\forall\, t \in [t_0-\varepsilon, t_0+\varepsilon]$ zuordnet, ist wohldefiniert und eine Kontraktion.

<u>Beweis</u>: Sei $\varphi \in A$, da $\|\varphi - x_0\| \le r$, liegt der Graph von φ in $U \Rightarrow f(t,\varphi(t))$ wohldefiniert ist und stetig als Funktion von t ist.

Da $\eta = T(\varphi)$, erhält man

$$\|\eta - x_0\| = \sup_{|t-t_0|\le\varepsilon} \|\eta(t) - x_0\|$$

$$= \sup_{|t-t_0|\le\varepsilon} \left\| \int_{t_0}^{t} f(s,\varphi(s))\,ds \right\|$$

$$\le \sup_{|t-t_0|\le\varepsilon} |t-t_0|\, M \le \varepsilon . M \le r$$

(Lemma 10.2.3)

d.h. $\eta \in A$, und T ist wohl definiert.

Wir zeigen: T ist eine Kontraktion.

Seien $\varphi_1, \varphi_2 \in A$. Dann gilt

$$T(\varphi_1)(t) - T(\varphi_2)(t) = \int_{t_0}^{t} \left(f(s,\varphi_1(s)) - f(s,\varphi_2(s)) \right) ds$$

Nach der Lipschitz-Bedingung von f (bzgl. x) gilt

$$\| T(\varphi_1)(t) - T(\varphi_2)(t) \| \le \int_{t_0}^{t} L \|\varphi_1(s) - \varphi_2(s)\|\, ds$$

$$\le L\, |t-t_0|\, \|\varphi_1 - \varphi_2\|$$

$$\le \frac{1}{2} \|\varphi_1 - \varphi_2\|$$

Es folgt die Behauptung. Nach Banachschen Fixpunktsatz $\exists\ \varphi \in A$ mit $T(\varphi) = \varphi$, d.h.

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\, ds \quad \forall t \in [t_0-\varepsilon, t_0+\varepsilon]$$

Dies ist die gesuchte Lösung des AWP ∎

<u>Bemerkung 13.3.7</u> (Picard-Lindelöfsches ~~Inte~~ Iterationsverfahren)

Nach dem Beweis des Banachschen Fixpunktsatz für eine Kontraktion $T: A \to A$ haben wir gesehen, dass $\forall$ jeden Anfangswert $\varphi_0 \in A$ die Folge $\varphi_k := T(\varphi_{k-1})$, $k \geq 0$ gegen den Fixpunkt konvergiert.

Angewendet auf den Satz 13.3.6 bedeutet dies: für $\varphi_0 \equiv x_0$ (konstante Funktion gleich x_0)

und $$\varphi_k := x_0 + \int_{t_0}^{t} f(s, \varphi_{k-1}(s))\, ds \quad \forall k \geq 1$$

dann konvergiert $(\varphi_k)_{k \in \mathbb{N}}$ gleichmäßig auf $[t_0-\varepsilon, t_0+\varepsilon]$ gegen ~~den~~ die Lösung φ von $\begin{cases} \varphi' = f(t,\varphi) \\ \varphi(t_0) = x_0 \end{cases}$

$\varphi: [t_0-\varepsilon, t_0+\varepsilon] \to \mathbb{R}$

Beispiel 13.3.8 Betrachte

$$(*) \begin{cases} x' = 2tx \\ x(0) = x_0 \end{cases} \qquad (t_0 = 0)$$

Nach Satz 13.3.6 (Picard-Lindelöf)

$\exists\, \varepsilon > 0$ so dass für $\varphi_0 : [-\varepsilon, \varepsilon] \to \mathbb{R}$, $t \mapsto x_0$

$$\varphi_k(t) := x_0 + \int_{t_0}^{t} 2s\, \varphi_{k-1}(s)\, ds \qquad \forall\, k \geq 1$$

Dann gilt: φ_k konvergiert ~~gegen~~ gleichmäßig auf $[-\varepsilon, \varepsilon]$ gegen die Lösung $\varphi : [-\varepsilon, \varepsilon] \to \mathbb{R}$ von $(*)$. Man berechnet

$$\varphi_1(t) = x_0 + 2x_0 \int_0^t s\, ds = x_0(1+t^2)$$

$$\varphi_2(t) = x_0 + 2x_0 \int_0^t s(1+s^2)\, ds$$

$$= x_0\left(1 + t^2 + \frac{t^4}{2}\right)$$

allgemein beweist man durch vollständige Induktion

$$\varphi_k(t) = x_0\left(1 + \frac{t^2}{1!} + \frac{t^4}{2!} + \frac{t^6}{3!} + \dots + \frac{t^{2k}}{k!}\right)$$

Daraus folgt $\varphi(t) = \lim_{k \to \infty} \varphi_k(t) = x_0 e^{t^2}$

$\forall\, t \in [-\varepsilon, \varepsilon]$

Man überprüft direkt, dass diese Funktion φ eingentlich die Lösung von (*) auf ganz $\mathbb{R}$ ist.

Das AWP (*) ist in der Tat eine homogene lineare DGL. Daher findet man wieder φ mit der Methoden aus § 13.2.

<u>Satz 13.3.9</u> (Eindeutigkeit- und Existenzsatz für DGL höherer Ordnung)

Sei $U \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt.

Sei $t_0 \in \mathbb{R}$, $x_0, x_0', \dots, x_0^{(n-1)} \in \mathbb{R}$ mit

$(t_0, x_0, x_0', \dots, x_0^{(n-1)}) \in U$.

Betrachte das AWP

$$(*) \quad \begin{cases} x^{(n)} = f(t, x, x', \dots, x^{(n-1)}) \\ x(t_0) = x_0 \\ x'(t_0) = x_0' \\ \vdots \\ x^{(n-1)}(t_0) = x_0^{(n-1)} \end{cases}$$

i) Sei $I \subset \mathbb{R}$ Intervall mit $t_0 \in I$, seien $\varphi, \psi : I \to \mathbb{R}$ zwei Lösungen von (*)

Dann gilt $\varphi(t) = \psi(t) \;\forall\, t \in I$.

(Eindeutigkeit)

ii) $\exists\, \varepsilon > 0$ so dass $\exists$ eine Lösung $\varphi : [t_0 - \varepsilon, t_0 + \varepsilon] \to \mathbb{R}$ von (*) (Existenz) (die Lösung ist eindeutig bestimmt nach (1'))

Beweis: Bemerkung 13.1.3 besagt:

für $u = (u_1, \dots, u_n) := (x, x', \dots, x^{(n-1)})$

gilt:

$x : I \to \mathbb{R}$ eine Lösung von $(*)$ $\Leftrightarrow$

$u : I \to \mathbb{R}^n$ eine Lösung von $\begin{cases} u' = F(t,u) \\ u(t_0) = u_0 \end{cases}$

wobei $u_0 = (x_0, x_0', \dots, x_0^{(n-1)}) \in U$

$F(t,u) := (u_2, u_3, \dots, u_n, f(t,u))$

Da f einer Lipschitz-Bedingung bzgl x genügt, genügt F auch einer Lipschitz-Bedingung bzgl. u

Nach Sätzen 13.3.5, 13.3.4 gilt die Behauptung ∎