Vorlesung aus dem Sommersemester 2010

Analysis 2

Prof. Dr. Hans-Dieter Donder

geTEXt von Viktor Kleen & Florian Stecker

Inhaltsverzeichnis

9 Gleichmäßige Konvergenz, Taylorreihen

Definition. Sei $D \subseteq \mathbb{R}$. Für $n \in \mathbb{N}$ sei $f_n \colon D \to \mathbb{R}$. Weiterhin sei $f \colon D \to \mathbb{R}$.

(a) Die Folge $(f_n)_{n\in\mathbb{N}}$ *konvergiert punktweise* gegen f, wenn für alle $x \in D$ die Folge $(f_n(x))_{n\in\mathbb{N}}$ gegen $f(x)$ konvergiert.

$$\forall x \in D \colon \forall \varepsilon > 0 \colon \exists m \in \mathbb{N} \colon \forall n \geq m \colon |f_n(x) - f(x)| < \varepsilon$$

(b) Die Folge $(f_n)_{n\in\mathbb{N}}$ *konvergiert gleichmäßig* gegen f, wenn gilt:

$$\forall \varepsilon > 0 \colon \exists m \in \mathbb{N} \colon \forall x \in D \colon \forall n \geq m \colon |f_n(x) - f(x)| < \varepsilon$$

Bemerkung. Wenn $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f konvergiert, so konvergiert $(f_n)_{n\in\mathbb{N}}$ auch punktweise gegen f.

Beispiel. Sei $I = [0,1]$ und $f_n \colon I \to \mathbb{R}, x \mapsto x/n$. Weiterhin sei $f = 0$. Dann konvergiert $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f. Denn sei $\varepsilon > 0$. Wähle $m \in \mathbb{N}$ mit $1/m < \varepsilon$. Dann gilt für alle $x \in I$ und alle $n \geq m$

$$|f_n(x) - f(x)| = x/n \leq 1/n \leq 1/m < \varepsilon$$

Zur Erinnerung: Für $f \colon D \to \mathbb{R}$ sei $\|f\| = \sup\{|f(x)| \colon x \in D\}$.

Bemerkung. Seien $f_n \colon D \to \mathbb{R}$, $f \colon D \to \mathbb{R}$. $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen f, wenn

$$\lim_{n\to\infty} \|f_n - f\| = 0$$

Bemerkung. Sei $I = [a,b]$. Ein $f \colon I \to \mathbb{R}$ ist Regelfunktion genau dann, wenn es eine Folge $(f_n)_{n\in\mathbb{N}}$ von Treppenfunktionen gibt, die gleichmäßig gegen f konvergieren.

Satz 1. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von stetigen Funktionen auf D, die gleichmäßig gegen f konvergiert. Dann ist auch f stetig.*

Beweis. Sei $a \in D$. Wir müssen zeigen, dass f stetig in a ist. Sei also $\varepsilon > 0$. Da $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f konvergiert, existiert ein $m \in \mathbb{N}$ mit $|f_m(x) - f(x)| < \varepsilon/3$ für alle $x \in D$. Da f_m stetig in a ist, existiert ein $\delta > 0$ mit $|f_m(x) - f_m(a)| < \varepsilon/3$ für alle $x \in D$ mit $|x - a| < \delta$. Dann gilt für alle $x \in D$ mit $|x - a| < \delta$

$$\begin{aligned} |f(x) - f(a)| &= |(f(x) - f_m(x)) + (f_m(x) - f_m(a)) + (f_m(a) - f(a))| \leq \\ &\leq |f(x) - f_m(x)| + |f_m(x) - f_m(a)| + |f_m(a) - f(a)| < \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \end{aligned}$$

□

Bemerkung. Dies gilt nicht bei punktweiser Konvergenz. Sei $I = [0,1]$ und definiere $f_n \colon I \to \mathbb{R}, x \to x^n$. Die f_n sind stetig und es gilt für $x \in I$

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & x \neq 1 \\ 1 & x = 1 \end{cases}$$

Also konvergiert $(f_n)_{n\in\mathbb{N}}$ punktweise gegen $f \colon I \to \mathbb{R}$ mit $f(x) = \delta_{x1}$. f ist aber in 1 nicht stetig.

Satz 2. *Sei $I = [a, b]$ ein kompaktes Intervall. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Regelfunktionen auf I, die gleichmäßig gegen f konvergiert. Dann ist auch f eine Regelfunktion und es gilt*

$$\int_a^b f \,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n \,\mathrm{d}x$$

Beweis. Wir zeigen zuerst, dass f eine Regelfunktion ist. Hierzu müssen wir zeigen, dass für alle $\varepsilon > 0$ eine Treppenfunktion $g\colon I \to \mathbb{R}$ existiert mit $\|f - g\| < \varepsilon$. Sei also $\varepsilon > 0$. Da $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f konvergiert, gilt $\lim_{n\to\infty} \|f_n - f\| = 0$. Also existiert ein $m \in \mathbb{N}$ mit $\|f_m - f\| \leq \varepsilon/2$. Da f_m eine Regelfunktion ist, existiert eine Treppenfunktion $g\colon I \to \mathbb{R}$ mit $\|f_m - g\| \leq \varepsilon/2$. Dann gilt aber $\|f - g\| \leq \|f - f_m\| + \|f_m - g\| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$. Wir zeigen nun $\int_a^b f \,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n \,\mathrm{d}x$. Dies ist trivial für $a = b$. Sei also $a < b$. Sei nun $\varepsilon > 0$. Wegen $\lim_{n\to\infty} \|f_n - f\| = 0$ existiert $m \in \mathbb{N}$ mit $\|f_n - f\| < \varepsilon/(b - a)$ für alle $n \geq m$. Somit gilt für alle $n \geq m$:

$$\left|\int_a^b f \,\mathrm{d}x - \int_a^b f_n \,\mathrm{d}x\right| = \left|\int_a^b f - f_n \,\mathrm{d}x\right| \leq (b - a)\|f_n - f\| < \varepsilon \qquad \square$$

Satz 3. *Sei I ein echtes Intervall. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge stetig differenzierbarer Funktionen auf I, die punktweise gegen f konvergiert. Weiterhin konvergiere die Folge der Ableitungen $(f_n')_{n\in\mathbb{N}}$ gleichmäßig gegen g. Dann ist f auch stetig differenzierbar und $f' = g$.*

Beweis. Nach Satz 1 ist g stetig. Also genügt es zu zeigen, dass f differenzierbar ist und $f' = g$ gilt. Wähle hierzu ein festes $a \in I$. Sei nun $x \in I$. Natürlich konvergiert auch die Folge der Einschränkungen von f_n' auf $[a, x]$ (bzw. $[x, a]$) gleichmäßig gegen die Einschränkung von g auf $[a, x]([x, a])$. Also folgt aus Satz 2

$$\int_a^x g(t) \,\mathrm{d}t = \lim_{n\to\infty} \int_a^x f_n'(t) \,\mathrm{d}t = \lim_{n\to\infty} (f_n(x) - f_n(a)) = f(x) - f(a)$$

d.h. $f(x) = f(a) + \int_a^x g(t) \,\mathrm{d}t$. Somit ist nach Satz 7 aus §7 f differenzierbar und es gilt $f' = g$. $\square$

Bemerkung. Wenn in Satz 3 $I = [a, b]$ kompakt ist, so folgt automatisch, dass $(f_n)_{n\in\mathbb{N}}$ sogar gleichmäßig gegen f konvergiert, denn für all $x \in [a, b]$ gilt:

$$\begin{aligned}
|f_n(x) - f(x)| &= \left|\left(\int_a^x f_n'(t) \,\mathrm{d}t - \int_a^x g(t) \,\mathrm{d}t\right) + (f_n(a) - f(a))\right| \leq \\
&\leq \left|\int_a^x f_n'(t) - g(t) \,\mathrm{d}t\right| + |f_n(a) - f(a)| \leq \\
&\leq (b - a)\|f_n' - g\| + |f_n(a) - f(a)| \underset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

Also ist $\lim_{n\to\infty} \|f_n - f\| = 0$.

Definition. Sei $D \subseteq \mathbb{R}$. Für $k \in \mathbb{N}$ sei $f_k\colon D \to \mathbb{R}$. Die Reihe $\sum_{k=0}^{\infty} f_k$ *konvergiert punktweise* (bzw. *konvergiert gleichmäßig*) gegen f, wenn die Folge der Partialsummen $\left(\sum_{k=0}^{n} f_k\right)_{n\in\mathbb{N}}$ punktweise (bzw. gleichmäßig) gegen f konvergiert. Wir schreiben dann $\sum_{k=0}^{\infty} f_k = f$.

Satz 4 (Konvergenzkriterium von Weierstraß)**.** *Seien $f_k\colon D \to \mathbb{R}$ Funktionen für $k \in \mathbb{N}$. Ist $\sum_{k=0}^{\infty} \|f_k\|$ konvergent, so ist $\sum_{k=0}^{\infty} f_k$ gleichmäßig konvergent.*

Beweis. Wir zeigen zuerst, dass $\sum_{k=0}^{\infty} f_k$ punktweise konvergiert. Sei hierzu $x \in D$. Da $|f_k(x)| \leq \|f_k\|$ für alle $k \in \mathbb{N}$ ist $\sum_{k=0}^{\infty} f_k(x)$ absolut konvergent. Definiere also $f\colon D \to \mathbb{R}, x \mapsto \sum_{k=0}^{\infty} f_k(x)$. Dann konvergiert $\sum_{k=0}^{\infty} f_k$ punktweise gegen f. Wir zeigen nun, dass $\sum_{k=0}^{\infty} f_k$ sogar gleichmäßig gegen f konvergiert. Sei also $\varepsilon > 0$. Da $\sum_{k=0}^{\infty} \|f_k\|$ konvergiert, existiert $m \in \mathbb{N}$ mit $\sum_{k=m}^{\infty} \|f_k\| < \varepsilon$. Dann gilt für alle $x \in D$ und alle $n \geq m$:

$$\left|\sum_{k=0}^{n} f_k(x) - f(x)\right| = \left|\sum_{k=n+1}^{\infty} f_k(x)\right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} \|f_k\| < \varepsilon$$

□

Beispiel. Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius r. Sei $0 \leq c < r$ und setze $I = [-c, c]$. Für $k \in \mathbb{N}$ definiere $f_k\colon I \to \mathbb{R}, x \mapsto a_k x^k$. Dann ist für alle $k \in \mathbb{N}$ und alle $x \in I$

$$|f_k(x)| = |a_k x^k| = |a_k||x|^k \leq |a_k c^k|$$

also $\|f_k\| \leq |a_k c^k|$. Da $\sum_{k=0}^{\infty} a_k c^k$ absolut konvergiert, ist also $\sum_{k=0}^{\infty} \|f_k\|$ absolut konvergent. Somit ist nach Satz 4 $\sum_{k=0}^{\infty} f_k$ gleichmäßig konvergent. Mit Satz 1 folgt also, dass $\sum_{k=0}^{\infty} f_k$ stetig ist.

Definition. Sei $a \in \mathbb{R}$. Eine *Potenzreihe mit Mittelpunkt a* ist eine Reihe der Form $\sum_{k=0}^{\infty} a_k (x-a)^k$ mit $a_k \in \mathbb{R}$. Sie konvergiert (bzw. divergiert) an der Stelle $b \in \mathbb{R}$, wenn $\sum_{k=0}^{\infty} a_k (b-a)^k$ konvergiert (bzw. divergiert).

$\sum_{k=0}^{\infty} a_k (x-a)^k$ konvergiert natürlich an der Stelle b genau dann, wenn $\sum_{k=0}^{\infty} a_k x^k$ an der Stelle $b - a$ konvergiert. Somit können wir alle Ergebnisse von früher auf diesen allgemeineren Begriff übertragen. Setzen wir also

$$I = \left\{ b \in \mathbb{R}\colon \sum_{k=0}^{\infty} a_k (x-a)^k \text{ konvergiert an der Stelle } b \right\}$$

so ist I ein Intervall mit Mittelpunkt a. Den Radius von I nennen wir den Konvergenzradius von $\sum_{k=0}^{\infty} a_k (x-a)^k$. Die zugehörige Funktion wird definiert durch $f(b) = \sum_{k=0}^{\infty} a_k (b-a)^k$. f ist stetig und in $\mathring{I}$ differenzierbar und es gilt für $b \in \mathring{I}$

$$f'(b) = \sum_{k=1}^{\infty} k a_k (b-a)^{k-1}$$

Damit ist f in $\mathring{I}$ natürlich beliebig oft differenzierbar. Man kann nun folgende Frage stellen: Sei $g\colon I \to \mathbb{R}$ beliebig oft differenzierbar und I ein Intervall. Sei $a \in \mathring{I}$. Gibt es dann eine Potenzreihe $\sum_{k=0}^{\infty} a_k (x-a)^k$ mit Mittelpunkt a und Konvergenzradius $r > 0$ und ein $0 \leq \overline{r} \leq r$, so dass $g(b) = \sum_{k=0}^{\infty} a_k (b-a)^k$ für alle $b \in I$ mit $|b-a| < \overline{r}$. Leider ist dies nicht immer der Fall. Nach dem Identitätssatz gibt es aber nur einen Kandidaten. Man kann sogar die Koeffizienten bestimmen. Es muss nämlich gelten

$$a_k = \frac{g^{(k)}(a)}{k!}$$

Definition. Sei I ein Intervall, $g\colon I \to \mathbb{R}$ eine in $\mathring{I}$ beliebig oft differenzierbare Funktion und $a \in \mathring{I}$. Setze dann

$$T(g,a) = \sum_{k=0}^{\infty} \frac{g^{(k)}(a)}{k!}(x-a)^k$$

$T(g,a)$ ist die *Taylorreihe von g mit Entwicklungspunkt a*. Wir bezeichnen mit $T(g,a)$ auch die zugehörige Funktion.

Ist der Konvergenzradius r von $T(g,a)$ echt größer als 0 und gibt es ein $0 < \overline{r} \leq r$ mit $g(b) = T(g,a)(b)$ für alle $b \in I$ mit $|b-a| < \overline{r}$, so sagen wir g besitzt im Punkt a ein Taylorentwicklung. Dies motiviert folgendes:

Definition. Seien $f\colon I \to \mathbb{R}$, I ein Intervall, $a \in I$. Sei f n-mal differenzierbar in a. Setze dann

$$T_n(f,a) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

$T_n(f,a)$ ist das *n-te Taylorpolynom von f mit Entwicklungspunkt a*. Wir betrachten $T_n(f,a)$ als Funktion $I \to \mathbb{R}$.

Satz 5 (Taylorsche Formel)**.** *Sei I ein Intervall und $f\colon I \to \mathbb{R}$ eine $(n+1)$-mal differenzierbare Funktion, und sei $a \in I$. Dann gilt für alle $x \in I$:*

$$f(x) = T_n(f,a)(x) + R_{n+1}(x) \quad \textit{wobei} \quad R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,\mathrm{d}t$$

Beweis. Durch Induktion über n.

- Induktionsanfang: Es ist $f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t = T_0(f,a)(x) + R_1(x)$.
- Induktionsschritt:

$$\begin{aligned}
R_{n+1}(x) &= \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,\mathrm{d}t = \\
&= -\int_a^x f^{(n+1)}(t)\left(\frac{(x-t)^{n+1}}{(n+1)!}\right)'\,\mathrm{d}t = \\
&= -f^{(n+1)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\Bigg|_{t=a}^{x} + \int_a^x f^{(n+2)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\,\mathrm{d}t = \\
&= \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} + \frac{1}{(n+1)!}\int_a^x (x-t)^{n+1} f^{(n+2)}(t)\,\mathrm{d}t = \\
&= \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} + R_{n+2}(x)
\end{aligned}$$

Also ist nach Induktionsvoraussetzung

$$\begin{aligned} f(x) &= T_n(f,a)(x) + R_{n+1}(x) = \\ &= T_n(f,a)(x) + \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1} + R_{n+2}(x) = \\ &= T_{n+1}(f,a)(x) + R_{n+2}(x) \end{aligned}$$

□

Korollar (Langrangesches Restglied)**.** *Sei I ein Intervall und $f\colon I \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion. Weiterhin seien $a, x \in I$. Dann existiert ein ξ zwischen a und x mit*

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

Beweis. Sei $R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,\mathrm{d}t = \int_a^x f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,\mathrm{d}t$ wie in Satz 5. Nach dem Mittelwertsatz der Integralrechnung existiert ein ξ zwischen a und x, so dass gilt

$$R_{n+1}(x) = f^{(n+1)}(\xi)\int_a^x \frac{(x-t)^n}{n!}\,\mathrm{d}t = -f^{(n+1)}(\xi)\frac{(x-t)^{n+1}}{(n+1)!}\bigg|_{t=a}^{x} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

□

Korollar. *Sei I ein Intervall und $f\colon I \to \mathbb{R}$ eine n-mal stetig differenzierbare Funktion. Weiterhin sei $a \in I$. Dann gibt es eine stetige Funktion $r\colon I \to \mathbb{R}$ mit $r(a) = 0$ und*

$$f(x) = T_n(f,a)(x) + r(x)(x-a)^n$$

für alle $x \in I$.

Beweis. Für $n = 0$ setze $r(x) = f(x) - f(a)$. Sei also $n > 0$. Definiere $r\colon I \to \mathbb{R}$ durch

$$r(x) = \begin{cases} \frac{1}{(x-a)^n}(f(x) - T_n(f,a)(x)) & x \neq a \\ 0 & x = a \end{cases}$$

Offenbar ist $f(x) = T_n(f,a)(x) + r(x)(x-a)^n$ für $x \in I$, denn $T_n(f,a)(a) = f(a)$. Weiter ist r stetig in $I \setminus \{a\}$. Wir müssen also nur noch zeigen, dass r stetig in a ist. Sei also $(x_k)_{k\in\mathbb{N}}$ eine folge aus $I \setminus \{a\}$, die gegen a konvergiert. Dann ist

$$\begin{aligned} r(x_k) &= \frac{1}{(x_k-a)^n}\left(f(x_k) - T_n(f,a)(x_k)\right) = \\ &= \frac{1}{(x_k-a)^n}\left(f(x_k) - T_{n-1}(f,a)(x_k) - \frac{f^{(n)}(a)}{n!}(x_k-a)^n\right) = \\ &= \frac{1}{(x_k-a)^n}\left(f(x_k) - T_{n-1}(f,a)(x_k)\right) - \frac{f^{(n)}(a)}{n!} \end{aligned}$$

Nach obigem Korollar existiert ein ξ_k zwischen a und x_k mit

$$f(x_k) - T_{n-1}(f,a)(x_k) = \frac{f^{(n)}(\xi_k)}{n!}(x_k-a)^n$$

Dann ist $r(x_k) = \frac{1}{n!}(f^{(n)}(\xi_k) - f^{(n)}(a))$. Wegen $\lim_{k\to\infty} x_k = a$ ist aber auch $\lim_{k\to\infty} \xi_k = a$. Da $f^{(n)}$ stetig ist, folgt als $\lim_{k\to\infty} r(x_k) = 0$. $\square$

Satz 6. *Sei $f\colon I \to \mathbb{R}$ n-mal stetig differenzierbar, wobei I ein Intervall ist und $n \geq 2$. Sei $a \in \mathring{I}$ und es gelte $f^{(k)}(a) = 0$ für $1 \leq k < n$, aber $f^{(n)}(a) \neq 0$.*

(a) Ist n gerade und $f^{(n)}(a) > 0$, so besitzt f in a ein lokales Minimum.

(b) Ist n gerade und $f^{(n)}(a) < 0$, so besitzt f in a ein lokales Maximum.

(c) Ist n ungerade, so besitzt f in a kein lokales Extremum.

Beweis. Nach Voraussetzung ist $T_{n-1}(f, a) = f(a)$.

zu (a) Sei n gerade und $f^{(n)}(a) > 0$. Wegen $f^{(n)}$ stetig und $a \in \mathring{I}$ existiert ein $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq I$ und $f^{(n)}(b) > 0$ für alle $b \in U_\varepsilon(a)$. Sei nun $x \in U_\varepsilon(a)$, $x \neq a$. Nach dem ersten Korollar zu Satz 5 existiert ein ξ zwischen a und x mit

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

Dann ist aber $f^{(n)}(\xi) > 0$ und wegen n gerade ist auch $(x-a)^n > 0$. Also folgt $f(x) > f(a)$.

zu (b) wende (a) auf $-f$ an.

zu (c) Sei nun n ungerade. Sei o.E. $f^{(n)}(a) > 0$ (sonst betrachte $-f$). Sei $\varepsilon > 0$. Es ist zu zeigen, dass $x_0, x_1 \in U_\varepsilon(a) \cap I$ existieren mit $f(x_0) < f(a) < f(x_1)$. Wegen $f^{(n)}$ stetig in a und $a \in \mathring{I}$ existiert ein $\delta > 0$ mit $\delta \leq \varepsilon$ und $U_\delta(a) \subseteq I$ und $f^{(n)}(b) > 0$ für alle $b \in U_\delta(a)$. Wähle $x_0, x_1 \in U_\delta(a)$ mit $x_0 < a < x_1$. Dann existiert wieder ξ_i zwischen a und x_i mit

$$f(x_i) = f(a) + \frac{f^{(n)}(\xi_i)}{n!}(x_i - a)^n$$

Dann gilt $f^{(n)}(\xi_i) > 0$. Wegen n ungerade ist aber $(x_0 - a)^n < 0$ und $(x_1 - a)^n > 0$. Also $f(x_0) < f(a) < f(x_1)$. $\square$

Beispiel. (a) Definiere $f\colon \mathbb{R} \to \mathbb{R}$ durch

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

f ist beliebig oft differenzierbar und es gilt $f^{(n)}(0) = 0$ für ale n. Die Taylorreihe von f mit Entwicklungspunkt 0 ist also identisch 0. Somit besitzt f keine Taylorentwicklung in 0.

(b) Früher wurde gezeigt

$$\ln(1+b) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{b^k}{k} \quad \forall b \in (-1,1)$$

Also

$$\ln(x) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(x-1)^k}{k} \quad \forall x \in (0,2)$$

Also besitzt ln eine Taylorentwicklung im Punkt 1.

10 Metrische Räume

Wir wollen nun $\mathbb{R}^n$ untersuchen. Für beliebige n ist $\mathbb{R}^n$ nicht mehr ein Körper. Aber $\mathbb{R}^n$ ist noch ein $\mathbb{R}$-Vektorraum, d.h. wir haben Addition und Skalarmultiplikation. Wir haben auch ein Analogon zum Absolutbetrag. Hierzu setzen wir für $a = (a_1, \dots, a_n) \in \mathbb{R}^n$

$$\|a\| = \sqrt{\sum_{k=1}^{n} a_k^2}$$

$\|\cdot\|$ ist die euklidische Norm. Für $a \in \mathbb{R}$ ist $\|a\| = |a|$. Für die euklidische Norm $\|\cdot\|\colon \mathbb{R}^n \to \mathbb{R}$ gilt:

(N1) $\|x\| = 0 \iff x = 0$

(N2) $\forall x \in \mathbb{R}^n \colon \forall \lambda \in \mathbb{R} \colon \|\lambda x\| = |\lambda|\|x\|$

(N3) $\forall x, y \in \mathbb{R}^n \colon \|x+y\| \leq \|x\| + \|y\|$

Beweis. Seien $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$.

(N1) $\|x\| = \sqrt{\sum_{k=1}^{n} a_k^2} = 0 \iff x_1 = \dots = x_n = 0 \iff x = 0$

(N2) $\|\lambda x\| = \sqrt{\sum_{k=1}^{n} (\lambda x_k)^2} = |\lambda| \sqrt{\sum_{k=1}^{n} x_k^2} = |\lambda|\|x\|$

(N3) Setze $\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k$. Wir zeigen zuerst die *Cauchy-Schwarz-Ungleichung*

$$\langle x, y \rangle \leq \|x\|^2 \|y\|^2$$

Dies ist trivial für $y = 0$. Sei also $y \neq 0$. Setze $\lambda = \langle x, y\rangle / \|y\|^2$. Es ist:

$$\begin{aligned} 0 \leq \|x - \lambda y\|^2 &= \sum_{k=1}^{n} (x_k - \lambda y_k)^2 = \\ &= \sum_{k=1}^{n} (x_k^2 - 2\lambda x_k y_k + \lambda^2 y_k^2) = \\ &= \sum_{k=1}^{n} x_k^2 - 2\lambda \sum_{k=1}^{n} x_k y_k + \lambda^2 \sum_{k=1}^{n} y_k^2 = \\ &= \|x\|^2 - 2\lambda\langle x, y\rangle + \lambda^2 \|y\|^2 \end{aligned}$$

Durch Multiplikation mit $\|y\|^2$ also

$$0 \leq \|x\|^2\|y\|^2 - 2\langle x, y\rangle^2 + \langle x, y\rangle^2$$

also $\langle x, y\rangle^2 \leq \|x\|^2\|y\|^2$ Dann gilt aber

$$\begin{aligned} \|x + y\|^2 &= \sum_{k=1}^{n} (x_k + y_k)^2 = \sum_{k=1}^{n} (x_k^2 + 2x_k y_k + y_k^2) = \\ &= \sum_{k=1}^{n} x_k^2 + 2\sum_{k=1}^{n} x_k y_k + \sum_{k=1}^{n} y_k^2 = \\ &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leq \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = \\ &= (\|x\| + \|y\|)^2 \end{aligned}$$

Also $\|x + y\| \leq \|x\| + \|y\|$ □

Definition. Sei V ein $\mathbb{R}$-Vektorraum und $\|\cdot\|\colon V \to \mathbb{R}^+$. Dann ist $\|\cdot\|$ eine *Norm*, wenn gilt:

(N1) $\|x\| = 0 \iff x = 0$

(N2) $\forall x \in V\colon \forall \lambda \in \mathbb{R}\colon \|\lambda x\| = |\lambda|\|x\|$

(N3) $\forall x, y \in V\colon \|x + y\| \leq \|x\| + \|y\|$

V mit $\|\cdot\|$ heißt dann *normierter Vektorraum.*

Sei nun V mit $\|\cdot\|$ ein normierter Vektorraum. Wir können dann den Abstand zwischen zwei Punkten $x, y \in V$ definieren durch $d(x, y) = \|x - y\|$. Hierfür gilt dann $d\colon V^2 \to \mathbb{R}^+$

(M1) $d(x, y) = 0 \iff x = y$

(M2) $\forall x, y \in V\colon d(x, y) = d(y, x)$ (Symmetrie)

(M3) $\forall x, y, z \in V\colon d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

Beweis. (M1) $d(x,y) = 0 \iff \|x-y\| = 0 \iff x-y = 0 \iff x = y$

(M2) $d(x,y) = \|x-y\| = \|-(y-x)\| = |-1|\|y-x\| = \|y-x\| = d(y,x)$

(M3) $d(x,z) = \|x-z\| = \|(x-y)+(y-z)\| \leq \|x-y\| + \|y-z\| = d(x,y) + d(y,z)$ □

Definition. Sei X eine Menge. Eine *Metrik* auf X ist eine Funktion $d\colon X^2 \to \mathbb{R}^+$ mit

(M1) $d(x,y) = 0 \iff x = y$

(M2) $\forall x,y \in X\colon d(x,y) = d(y,x)$ (Symmetrie)

(M3) $\forall x,y,z \in X\colon d(x,z) \leq d(x,y) + d(y,z)$ (Dreiecksungleichung)

X mit d ist ein *metrischer Raum.*

Wir sagen auch einfach „X ist ein metrischer Raum" und bezeichnen dann die Metrik mit d. Dies machen wir auch für verschiedene Räume. Speziell ist also jeder $\mathbb{R}^n$ ein metrischer Raum mit der euklidischen Metrix $d(x,y) = \|x-y\|$. Dies benutzen wir immer auf dem $\mathbb{R}^n$. Sei nun X ein metrischer Raum.

Definition. Für $x \in X$ und $\varepsilon > 0$ setzen wir

$$U_\varepsilon(x) = \{y \in X\colon d(x,y) < \varepsilon\}$$

$U_\varepsilon(x)$ ist die *ε-Umgebung von X.*

Definition. Sei $x \in X$ und $U \subseteq X$. U ist eine *Umgebung von x*, wenn ein $\varepsilon > 0$ existiert mit $U_\varepsilon(x) \subseteq U$. Speziell ist also für $\varepsilon > 0$ $U_\varepsilon(x)$ eine Umgebung von X.

Satz 1 (Trennungseigenschaft von Hausdorff)**.** *Zu je 2 verschiedenen Punkten $x,y \in X$ gibt es Umgebungen U von x und V von y mit $U \cap V = \varnothing$.*

Beweis. Seien $x,y \in X$, $x \neq y$. Setze $\varepsilon = d(x,y)$. Dann ist $\varepsilon > 0$ nach (M1). Setze $\delta = \varepsilon/2$ und $U = U_\delta(x)$ und $V = U_\delta(y)$. Dann ist U Umgebung von x und V Umgebung von y. Es ist $U \cap V \neq \varnothing$, denn nehme an, dass $z \in U \cap V$. Dann ist

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

was ein Widerspruch ist. Ist X ein metrischer Raum und $A \subseteq X$, so ist A mit der auf A eingeschränkten Metrix selbst ein metrischer Raum. Auf diese Weise wird insbesondere jede Teilmenge von $\mathbb{R}^n$ zu einem metrischen Raum. □

Im folgenden seien X,Y,Z stets metrische Räume.

Definition. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge von X und $a \in X$. Die Folge $(a_n)_{n\in\mathbb{N}}$ *konvergiert* gegen a, wenn gilt: für alle $\varepsilon > 0$ existiert ein $m \in \mathbb{N}$ mit $d(a_n, a) < \varepsilon$ für alle $n \geq m$. a ist dann *Grenzwert* (oder Limes) von $(a_n)_{n\in\mathbb{N}}$.

Wir können dies auf verschiedene Weise umformulieren.

Bemerkung. $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a genau dann wenn $\lim_{n\to\infty} d(a_n, a) = 0$

Bemerkung. Es ist äquivalent:

1. $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a.
2. für alle $\varepsilon > 0$ existiert ein $m \in \mathbb{N}$ mit $a_n \in U_\varepsilon(a)$ für alle $n \geq m$.
3. für alle Umgebungen U von a existiert ein $m \in \mathbb{N}$ mit $a_n \in U$ für alle $n \geq m$.

Aus Satz 1 folgt also, dass jede Folge höchstens einen Grenzwert besitzt. Wir können also schreiben:

$$\lim_{n\to\infty} a_n = a \qquad \text{für „}(a_n)_{n\in\mathbb{N}} \text{ konvergiert gegen } a\text{“}$$

Definition. Eine Folge $(a_n)_{n\in\mathbb{N}}$ aus X ist *beschränkt*, wenn ein $b \in X$ und $r > 0$ existieren mit $a_n \in U_r(b)$ für alle $n \in \mathbb{N}$.

Bemerkung. Jede konvergente Folge ist beschränkt.

Beweis. Sei $\lim_{n\to\infty} a_n = a$. Dann existiert ein $m \in \mathbb{N}$ mit $a_n \in U_1(a)$ für alle $n \geq m$. Setze $r = \max\{d(a_k, a) \colon k < m\} + 1$. Dann gilt $a_n \in U_r(a)$ für alle n. □

Satz 2. *Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge aus $\mathbb{R}^n$ und $b \in \mathbb{R}^n$. Sei $b = (b_1, \ldots, b_n)$ und $a_k = (a_{k1}, \ldots, a_{kn})$. Dann gilt $\lim_{k\to\infty} a_k = b$ genau dann, wenn $\lim\limits_{k\to\infty} a_{kl} = b_l$ für alle $1 \leq l \leq n$.*

Beweis. Sei $\lim_{k\to\infty} a_k = b$. Dann gilt $\lim_{k\to\infty} d(a_k, b) = 0$. Sei $1 \leq l \leq n$. Dann ist für $k \in \mathbb{N}$

$$0 \leq |a_{kl} - b_l| \leq \sqrt{\sum_{j=1}^{n} (a_{kj} - b_j)^2} = d(a_k, b)$$

Also $\lim_{k\to\infty} |a_{kl} - b_l| = 0$, d.h. $\lim_{k\to\infty} a_{kl} = b_l$.
Sei umgekehrt $\lim_{k\to\infty} a_{kl} = b_l$ für alle $1 \leq l \leq n$. Dann ist für alle $1 \leq l \leq n$ $\lim_{k\to\infty}(a_{kl} - b_l) = 0$. Also ist

$$\lim_{k\to\infty} d(a_k, b) = \lim_{k\to\infty} \sqrt{\sum_{j=1}^{n} (a_{kj} - b_j)^2} = 0$$

Damit folgt sofort für Folgen $(a_k)_{k\in\mathbb{N}}$, $(b_k)_{k\in\mathbb{N}}$ und $c \in \mathbb{R}$.

$$\lim_{k\to\infty} (a_k + b_k) = \lim_{k\to\infty} a_k + \lim_{k\to\infty} b_k$$

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

□

Definition. Sei $(a_n)_{n\in\mathbb{N}}$ Folge aus X und $a \in X$. a ist *Häufungspunkt* von $(a_n)_{n\in\mathbb{N}}$, wenn es für jede Umgebung U von a unendlich viele n gibt mit $a_n \in U$.

Bemerkung. a ist ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$, wenn es eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$ gibt, die gegen a konvergiert.

Beweis. Wie in Analysis I.

Satz 3 (Bolzano-Weierstraß)**.** *Im $\mathbb{R}^n$ besitzt jede beschränkte Folge einen Häufungspunkt.*

Beweis. Durch Induktion über n.

- Induktionsanfang ($n = 1$): nach Analysis I
- Induktionsschritt: Sei $(a_k)_{k\in\mathbb{N}}$ beschränkte Folge in $\mathbb{R}^{n+1}$. Sei $a_k = (a_{k1}, \dots, a_{k,n+1})$ für $k \in \mathbb{N}$. Setze $b_k = (a_{k1}, \dots, a_{kn})$. Dann ist $(b_k)_{k\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}^n$. Nach Induktionsvoraussetzung besitzt also $(b_k)_{k\in\mathbb{N}}$ eine konvergente Teilfolge $(b_{k(l)})_{l\in\mathbb{N}}$. Nach Analysis I besitzt die beschränkte Folge $(a_{k(l),n+1})_{l\in\mathbb{N}}$ eine konvergente Teilfolge $(a_{k(l)(j),n+1})_{j\in\mathbb{N}}$. Nach Satz 2 ist dann $(a_{k(l)(j)})_{j\in\mathbb{N}}$ eine konvergente Teilfolge von $(a_k)_{k\in\mathbb{N}}$. □

Definition. Eine Folge $(a_n)_{n\in\mathbb{N}}$ aus X ist eine *Cauchy-Folge*, wenn gilt:

$$\forall \varepsilon > 0 \colon \exists m \in \mathbb{N} \colon \forall n \geq m \colon d(a_n, a_m) < \varepsilon$$

Bemerkung. Ist $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge, so gilt stärker:

$$\forall \varepsilon > 0 \colon \exists m \in \mathbb{N} \colon \forall k, n \geq m \colon d(a_n, a_k) < \varepsilon$$

Beweis. Wie in Analysis I

Bemerkung. Jede konvergente Folge ist eine Cauchy-Folge. Die Umkehrung gilt nicht in beliebigen metrischen Räumen.

Beweis. Sei $\lim_{n\to\infty} a_n = a$. Sei $\varepsilon > 0$. Dann existiert $m \in \mathbb{N}$ mit $d(a_n, a) < \varepsilon/2$ für alle $n \geq m$. Dann ist aber für $n \geq m$ $d(a_n, a_m) \leq d(a_n, a) + d(a_m, a) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Beispiel. Sei $X = (0, 1)$ ein offenes Intervall. In X ist $(1/n)_{n\in\mathbb{N}}$ eine Cauchy-Folge, die in X (!) nicht konvergiert.

Definition. Ein metrischer Raum ist *vollständig*, wenn jede Cauchy-Folge aus X konvergiert (in X).

Satz 4. *$\mathbb{R}^n$ ist vollständig.*

Beweis. Sei $(a_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge aus $\mathbb{R}^n$. Sei $a_k = (a_{k1}, \dots, a_{kn})$. Dann ist für jedes $1 \leq l \leq n$ $(a_{kl})_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$, denn $0 \leq |a_{kl} - a_{jl}| \leq d(a_k, a_j)$ für alle k, j. Also ist $(a_{kl})_{k\in\mathbb{N}}$ konvergent. Somit ist nach Satz 2 $(a_k)_{k\in\mathbb{N}}$ konvergent. □

Definition. Sei $f \colon X \to Y$, $a \in X$.

1. f ist *stetig in a*, wenn für alle $\varepsilon > 0$ ein $\delta > 0$ existiert mit $d(f(x), f(a)) < \varepsilon$ für alle $x \in X$ mit $d(x, a) < \delta$.
2. f ist *stetig*, wenn f in jedem Punkt $a \in X$ stetig ist.

Bemerkung. Somit sind also folgende Aussagen äquivalent:

1. f ist stetig in a
2. Für alle $\varepsilon > 0$ existiert ein $\delta > 0$ mit $f[U_\delta(a)] \subseteq U_\varepsilon(f(a))$
3. Für alle Umgebungen U von $f(a)$ existiert eine Umgebung V von a mit $f[V] \subseteq U$

Bemerkung. Eine Funktion $f\colon X \to Y$ ist genau dann stetig im Punkt $a \in X$, wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$, die gegen a konvergiert, die Bildfolge $(f(x_n))_{n\in\mathbb{N}}$ gegen $f(a)$ konvergiert.

Beweis. Wie in Analysis I

Beispiel.

- Sei $1 \leq i \leq n$. Definiere $p_i\colon \mathbb{R}^n \to \mathbb{R}$ durch $p_i(x_1, \dots, x_n) = x_i$. p_i ist die i-te Projektion. p_i ist stetig. Das folgt sofort aus Satz 2.
- Die Funktion $+\colon \mathbb{R}^2 \to \mathbb{R}$ ist stetig.
- Die Funktion $\cdot\colon \mathbb{R}^2 \to \mathbb{R}$ ist stetig.
- Sind $f, g\colon X \to \mathbb{R}$ stetig, $c \in \mathbb{R}$, so sind $f + g$, $f \cdot g$, cf stetig (entsprechend abgeändert gilt das auch für $1/g$).

Satz 5. *Seien $f\colon X \to Y$ und $g\colon Y \to Z$. Die Funktion f sei in $a \in X$ stetig und g sei in $f(a) \in Y$ stetig. Dann ist $g \circ f$ in a stetig.*

Beweis. Sei U eine Umgebung von $(g \circ f)(a) = g(f(a))$. Wegen g in $f(a)$ stetig existiert dann eine Umgebung V von $f(a)$ mit $g[V] \subseteq U$. Wegen f in a stetig existiert eine Umgebung W von a mit $f[W] \subseteq V$. Dann ist $(g \circ f)[W] = g[f[W]] \subseteq g[V] \subseteq U$. □

Beispiel. Die Funktion $\|\cdot\|\colon \mathbb{R}^n \to \mathbb{R}$ ist stetig. Denn

$$\|(x_1, \dots, x_n)\| = \sqrt{\sum_{k=1}^{n} x_k^2}$$

Definition. Sei $f\colon X \to \mathbb{R}^n$. Dann existieren eindeutig bestimmte Funktionen $f_1, \dots, f_n$ von $X \to \mathbb{R}$ mit $f(x) = (f_1(x), \dots, f_n(x))$. f_i ist die i-te *Komponentenfunktion* von f. Wir schreiben hierfür etwas ungenau $f = (f_1, \dots, f_n)$.

Satz 6. *Sei $f\colon X \to \mathbb{R}^n$, $f = (f_1, \dots, f_n)$, und sei $a \in X$. Dann ist f genau dann in a stetig, wenn jedes f_i mit $1 \leq i \leq n$ in a stetig ist.*

Beweis. Sei f in a stetig. Es ist $f_i = p_i \circ f$. Da p_i stetig ist, ist f_i in a stetig. Sei umgekehrt jedes f_i in a stetig. Sei $\lim_{k\to\infty} x_k = a$. Dann ist $\lim_{k\to\infty} f_i(x_k) = f_i(a)$ für $1 \leq i \leq n$. Also $\lim_{k\to\infty} f(x_k) = (f_1(a), \dots, f_n(a)) = f(a)$. □

Definition. Sei $A \subseteq X$ und $b \in X$. b ist *Häufungspunkt* von A, wenn in jeder Umgebung von b ein Element von A liegt, das von b verschieden ist.

Bemerkung. Seien A,b wie oben. Es sind äquivalent:

1. b ist ein Häufungspunkt von A
2. Es gibt eine Folge aus $A \setminus \{b\}$, die gegen b konvergiert
3. In jeder Umgebung von b liegen unendlich viele Elemente von A.

Definition. Seien $A \subseteq X$ und $f\colon A \to Y$ und $b \in Y$. Weiterhin sei $a \in A$ oder a ein Häufungspunkt von A. Dann ist $\lim_{x \to a} f(x) = b$, falls für jede Folge $(a_n)_{n\in\mathbb{N}}$ aus A, die gegen a konvergiert, die Folge $(f(a_n))_{n\in\mathbb{N}}$ gegen b konvergiert.

Definition.

1. Eine Teilmenge A von X ist *abgeschlossen*, wenn sie jeden ihrer Häufungspunkte enthält.
2. Eine Teilmenge U von X ist *offen*, wenn für alle $x \in U$ U eine Umgebung von x ist.

Beispiel. Seien $a \in X$ und $\varepsilon > 0$. Setze $U = U_\varepsilon(a)$. Dann ist U offen.

Beweis. Sei $x \in U$. Setze $\delta = \varepsilon - d(x,a) > 0$. Dann ist $U_\delta(x) \subseteq U$, denn sei $z \in U_\delta(x)$. Dan ist $d(z,a) \leq d(z,x) + d(x,a) < \delta + d(x,a) = \varepsilon$. Somit ist U Umgebung von X. □

Bemerkung. Sei $U \subseteq X$. Dann ist U offen genau dann, wenn $X \setminus U$ abgeschlossen ist.

Beweis. Sei also U offen. Dann ist kein $x \in U$ ein Häufungspunkt von $X \setminus U$, denn für $x \in U$ ist U Umgebung von x mit $U \cap (X \setminus U) = \varnothing$. Also ist $X \setminus U$ abgeschlossen. Sei umgekehrt $X \setminus U$ abgeschlossen und $x \in U$. Dann ist x kein Häufungspunkt von $X \setminus U$. Somit existiert eine Umgebung W von x mit $W \cap (X \setminus U) \subseteq \{x\}$. Wegen $x \in U$ ist dann $W \subseteq U$. Also ist U Umgebung von x. □

Satz 7. *Sei $f\colon X \to Y$, Dann sind folgende Aussagen äquivalent:*

1. *f ist stetig*
2. *Für alle offenen $U \subseteq Y$ ist $f^{-1}[U]$ offen.*
3. *Für alle abgeschlossenen $A \subseteq Y$ ist $f^{-1}[A]$ abgeschlossen.*

Beweis.

1. Sie f stetig und $U \subseteq Y$ offen. Sei $a \in f^{-1}[U]$. Wir müssen zeigen, dass $f^{-1}[U]$ eine Umgebung von a ist. Nun ist $f(a) \in U$. Wegen U offen ist also U eine Umgebung von $f(a)$. Wegen f stetig existiert also eine Umgebung V von a mit $f[V] \subseteq U$. Dann ist aber $V \subseteq f^{-1}[U]$. Also ist $f^{-1}[U]$ Umgebung von a.
2. Sei $A \subseteq Y$ abgeschlossen. Dann ist nach obiger Bemerkung $Y \setminus A$ offen. Also $f^{-1}[Y \setminus A]$ offen. Aber $f^{-1}[Y \setminus A] = X \setminus f^{-1}[A]$. Also ist $f^{-1}[A]$ abgeschlossen

3. Sei $a \in X$. Wir müssen zeigen, dass f in a stetig ist. Sei also $\varepsilon > 0$. Setze $U = U_\varepsilon(f(a))$. Wegen U offen ist dann $Y \setminus U$ abgeschlossen. Also $f^{-1}[Y \setminus U] = X \setminus f^{-1}[U]$ abgeschlossen. Somit ist $V = f^{-1}[U]$ offen. Wegen $a \in V$ ist also V eine Umgebung von a mit $f[V] \subseteq U$. $\square$

Beispiel. Sei $f\colon X \to \mathbb{R}$ stetig, $b \in \mathbb{R}$. Dann ist $A = \{x \in X\colon f(x) = b\}$ abgeschlossen und $B = \{x \in X\colon f(x) < b\}$ offen. Denn $A = f^{-1}[\{b\}]$, $B = f^{-1}[(-\infty, b)]$.

Definition. Eine Teilmenge A von X ist *kompakt*, wenn jede Folge aus A einen Häufungspunkt besitzt, der in A liegt.

Definition. Eine Teilmenge A von X ist *beschränkt*, wenn es ein $b \in X$ und $r > 0$ gibt mit $A \subseteq U_r(b)$ oder wenn $A = \varnothing$.

Satz 8. *Jede kompakte Teilmenge von X ist abgeschlossen und beschränkt.*

Beweis. Sei $A \subseteq X$ kompakt. Wir zeigen zuerst, dass A abgeschlossen ist. Sei hierzu a ein Häufungspunkt von A. Dann existiert eine Folge $(a_n)_{n\in\mathbb{N}}$ aus $A \setminus \{a\}$, die gegen a konvergiert. Also ist a der einzige Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$. Wegen Kompaktheit von A ist also $a \in A$. Weiterhin kann A nicht unbeschränkt sein. Denn nehme an, dass dies der Fall ist. Dann ist natürlich $X \neq \varnothing$. Sei also $x \in X$ Wenn A unbeschränkt ist, dann ist für alle $n \geq 1$ A keine Teilmenge von $U_n(x)$. Wähle also für jedes $n \geq 1$ ein $a_n \in A$ mit $a_n \notin U_n(x)$ und setze $a_0 = x$. Dann besitzt offenbar die Folge $(a_n)_{n\in\mathbb{N}}$ keinen Häufungspunkt, denn jede Teilfolge von $(a_n)_{n\in\mathbb{N}}$ ist unbeschränkt. Also ist A nicht kompakt. $\square$

Bemerkung. Die Umkehrung von Satz 8 gilt aber nicht in jedem metrischen Raum.

Beispiel. Sei $X = \{1/(n+1)\colon n \in \mathbb{N}\}$. Dann ist X abgeschlossen (in X!) und beschränkt. Aber X ist nicht kompakt in X, da $(1/(n+1))_{n\in\mathbb{N}}$ in X keinen Häufungspunkt besitzt.

Satz 9. *Sei $A \subseteq \mathbb{R}^n$. Dann ist A genau dann kompakt, wenn A abgeschlossen und beschränkt ist.*

Beweis. Ist A kompakt, so ist A abgeschlossen und beschränkt nach Satz 8. Sei also umgekehrt A abgeschlossen und beschränkt. Wir müssen zeigen, dass A kompakt ist. Sei hierzu $(a_n)_{n\in\mathbb{N}}$ eine Folge aus A. Wegen A beschränkt ist dann auch $(a_n)_{n\in\mathbb{N}}$ beschränkt. Nach Satz 3 besitzt dann $(a_n)_{n\in\mathbb{N}}$ einen Häufungspunkt a. Wegen A abgeschlossen ist $a \in A$. $\square$

Definition. Sei $\mathcal{U} \subseteq \mathcal{P}(X)$ und $A \subseteq X$. $\mathcal{U}$ ist *offene Überdeckung* von A, wenn gilt:

1. Jedes $U \in \mathcal{U}$ ist offen
2. Für alle $a \in A$ existiert $U \in \mathcal{U}$ mit $a \in U$.

Satz 10 (Heine-Borelscher Überdeckungssatz)**.** *Eine Teilmenge A von X ist genau dann kompakt, wenn es für jede offene Überdeckung $\mathcal{U}$ von A endlich viele $U_1, \ldots, U_n \in \mathcal{U}$ gibt mit $A \subseteq U_1 \cup \ldots \cup U_n$.*

Beweis. Von links nach rechts: Sei $A \subseteq X$ kompakt. Sei $\mathcal{U}$ eine offene Überdeckung von A. Für $\delta > 0$ setze $A_\delta = \{x \in A \colon \exists U \in \mathcal{U} \colon U_\delta(x) \subseteq U\}$. Wir zeigen zuerst $(*)$ es existiert $\delta > 0$ mit $A_\delta = A$.

Beweis von $(*)$: Nehme an, dass dies nicht der Fall ist. Dann existiert für alle $n \geq 1$ ein $x_n \in A$ mit $x_n \notin A_{1/n}$. Wegen A kompakt besitzt die Folge $(x_n)_{n\geq 1}$ einen Häufungspunkt $x \in A$. Dann existiert $U \in \mathcal{U}$ mit $x \in U$. Wegen U offen gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq U$. Sei $m \in \mathbb{N}$ mit $1/m \leq \varepsilon/2$. Da x ein Häufungspunkt von $(x_n)_{n\geq 1}$ ist, gibt es ein $k \geq m$ mit $d(x_k, x) < 1/m$. Dann ist aber für $z \in U_{1/k}(x_k)$

$$d(z, x) \leq d(z, x_k) + d(x_k, x) < 1/k + 1/m \leq 2/m \leq \varepsilon$$

Somit ist $U_{1/k}(x_k) \subseteq U_\varepsilon(x) \subseteq U$. Also $x_k \in A_{1/k}$, was ein Widerspruch zur Wahl von x_k ist. Wähle nun $\delta > 0$ wie in $(*)$. Es genügt natürlich nun zu zeigen, dass es $x_1, \ldots, x_n \in A$ gibt mit $A \subseteq U_\delta(x_1) \cup \ldots \cup U_\delta(x_k)$. Nehme an, dass dies nicht der Fall ist. Dann können wir rekursiv eine Folge $(x_n)_{n\in\mathbb{N}}$ definieren mit $x_n \in A \setminus (U_\delta(x_0) \cup \ldots \cup U_\delta(x_{n-1}))$. Dann gilt aber für alle $m, n \in \mathbb{N}$ mit $m < n$, dass $d(x_m, x_n) \geq \delta$. Offenbar hat aber dann $(x_n)_{n\in\mathbb{N}}$ keinen Häufungspunkt. Dies ist aber ein Widerspruch zur Kompaktheit von A.

Von rechts nach links: Die andere Richtung des Satzes zeigen wir indirekt. Sei also $A \subseteq X$ nicht kompakt. Dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ aus A, die keinen Häufungspunkt in A besitzt. Somit existiert für jedes $x \in A$ eine offene Umgebung V_x von x mit $\{n \in \mathbb{N} \colon x_n \in V_x\}$ ist endlich. Setze nun $\mathcal{U} = \{V_x \colon x \in A\}$. Dann ist $\mathcal{U}$ eine offene Überdeckung von A. Sind aber $V_{z_1}, \ldots, V_{z_n}$ endlich viele Elemente von $\mathcal{U}$, so ist $\{n \in \mathbb{N} \colon x_n \in V_{z_1} \cup \ldots \cup V_{z_n}\}$ endlich. Also ist $A \not\subseteq V_{z_1} \cup \ldots \cup V_{z_n}$. □

Satz 11. *Sei $f \colon X \to Y$ stetig und $A \subseteq X$ kompakt. Dann ist $f[A]$ kompakt.*

Beweis. Sei $(y_n)_{n\in\mathbb{N}}$ eine Folge aus $f[A]$. Dann existiert für jedes $n \in \mathbb{N}$ ein $x_n \in A$ mit $y_n = f(x_n)$. Da A kompakt ist, besitzt $(x_n)_{n\in\mathbb{N}}$ eine Teilfolge $(x_{n(k)})_{k\in\mathbb{N}}$, die gegen ein $x \in A$ konvergiert. Wegen f stetig gilt aber dann

$$\lim_{k\to\infty} y_{n(k)} = \lim_{k\to\infty} f(x_{n(k)}) = f(x) \in f[A]$$ □

Definition. Sei $f \colon X \to Y$. f ist *gleichmäßig stetig*, wenn gilt:

$$\forall \varepsilon > 0 \colon \exists \delta > 0 \colon \forall x, x' \in X \colon \bigl(d(x, x') < \delta \Longrightarrow d\bigl(f(x), f(x')\bigr) < \varepsilon\bigr)$$

Satz 12. *Ist $f \colon X \to Y$ stetig und X kompakt, so ist f sogar gleichmäßig stetig.*

Beweis. Sei $\varepsilon > 0$. Wegen f stetig gibt es zu jedem $x \in X$ ein $\delta(x) > 0$ mit $f\left[U_{\delta(x)}(x)\right] \subseteq U_{\varepsilon/2}\left(f(x)\right)$. Setze $\mathcal{U} = \left\{U_{\delta(x)/2}(x) \colon x \in X\right\}$. Dann ist $\mathcal{U}$ eine offene Überdeckung von X. Wegen X kompakt gibt es also nach Satz 10 $x_1, \ldots, x_n \in X$ mit $X = U_{\delta(x_1)/2}(x_1) \cup \ldots \cup U_{\delta(x_n)/2}(x_n)$. Setze $\delta = \min\{\delta(x_1)/2, \ldots, \delta(x_n)/2\} > 0$. Seien $a, a' \in X$ mit $d(a, a') < \delta$. Dann gibt es ein $1 \leq j \leq n$ mit $a \in U_{\delta(x_j)/2}(x_j)$. Wegen $d(a, a') < \delta(x_j)/2$ ist dann $a' \in U_{\delta(x_j)}(x_j)$, denn $d(a', x_j) \leq d(a', a) + d(a, x_j) < \delta(x_j)$. Somit

$$d\left(\left(f(a), f(a')\right)\right) \leq d\left(f(a), f(x_j)\right) + d\left(f(x_j), f(a')\right) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$ □

Definition. Zwei Mengen A, B heißen *disjunkt*, wenn $A \cap B = \varnothing$ gilt.

Definition. Eine Teilmenge A von X ist *zusammenhängend*, wenn es keine disjunkten offenen Mengen U, V gibt mit $A \subseteq U \cup V$, $U \cap A \neq \varnothing$ und $V \cap A \neq \varnothing$.

Satz 13. *Eine Teilmenge von $\mathbb{R}$ ist genau dann zusammenhängend, wenn sie ein Intervall ist.*

Beweis. Die Richtung von links nach rechts zeigen wir indirekt. Sei also $A \subseteq \mathbb{R}$ kein Intervall. Dann gibt es Punkt $u, v \in A$ mit $u < v$ und ein $x \in X$ mit $u < x < v$ und $x \notin A$. Setze $U = (-\infty, x)$ und $V = (x, \infty)$. Dann sind U, V disjunkt, beide offen, und es gilt $A \subseteq U \cup V$, $U \cap A \neq \varnothing$ und $V \cap A \neq \varnothing$. Also ist A nicht zusammenhängend.
Sei umgekehrt I ein Intervall. Sien U, V offen mit $I \in U \cup V$, $U \cap I \neq \varnothing$ und $V \cap I \neq \varnothing$. Wir müssen zeigen, dass U,V nicht disjunkt sind. Wähle hierzu $u \in U \cap I$ und $v \in V \cap I$. Wenn $u = v$ sind wir fertig. Wir können also o.E. annehmen, dass $u < v$. Da I ein Intervall ist, gilt $[u, v] \subseteq I$. Sei $s = \sup(U \cap [u, v])$. Ist $s \in U$, so wegen U offen $s = v$. Also ist in diesem Fall $U \cap V \neq \varnothing$. Sei also $s \notin U$. Wegen $s \in I$ ist dann $s \in V$. Wegen V offen existiert dann $\varepsilon > 0$ mit $(s - \varepsilon, s + \varepsilon) \subseteq V$. Somit aber nach Definition von s auch $U \cap V \neq \varnothing$. □

Satz 14. *Sei $f \colon X \to Y$ stetig und $A \subseteq X$ zusammenhängend. Dann ist $f[A]$ zusammenhängend.*

Beweis. Seien U, V offene Teilmengen von Y mit $f[A] \subseteq U \cup V$, $U \cap f[A] \neq \varnothing$ und $V \cap f[A] \neq \varnothing$. Wegen f stetig sind $f^{-1}[U]$ und $f^{-1}[V]$ offen. Außerdem gilt $A \subseteq f^{-1}[U] \cup f^{-1}[V]$, $A \cap f^{-1}[U] \neq \varnothing$ und $A \cap f^{-1}[V] \neq \varnothing$. Wegen A zusammenhängend ist also $f^{-1}[U] \cap f^{-1}[V] \neq \varnothing$. Also $U \cap V \neq \varnothing$. □

Definition. Für $n \in \mathbb{N}$ seien $f_n \colon X \to Y$ Funktionen. Weiterhin sei $f \colon X \to Y$ Funktion.

1. Die Folge $(f_n)_{n \in \mathbb{N}}$ *konvergiert punktweise* gegen f, wenn für alle $x \in X$ gilt, dass $\lim_{n \to \infty} f_n(x) = f(x)$.

2. Die Folge $(f_n)_{n \in \mathbb{N}}$ *konvergiert gleichmäßig* gegen f, wenn gilt:

$$\forall \varepsilon > 0 \colon \exists m \in \mathbb{N} \colon \forall x \in X \colon \forall n \geq m \colon d(f_n(x), f(x)) < \varepsilon$$

Satz 15. *Konvergiert eine Folge $(f_n)_{n \in \mathbb{N}}$ von stetigen Funktionen $f_n \colon X \to Y$ gleichmäßig gegen f, so ist auch f stetig.*

Beweis. Wie in §9.

11 Kurven

Definition. Eine *Kurve* im $\mathbb{R}^n$ ist eine stetige Funktion $f \colon I \to \mathbb{R}^n$, wobei I ein echtes Intervall ist. Sei $f = (f_1, \ldots, f_n)$. f ist (stetig) *differenzierbar*, falls $f_1, \ldots, f_n$ (stetig) differenzierbar sind. Man definiert dann $f' \colon I \to \mathbb{R}^n$ durch $f'(t) = (f_1'(t), \ldots, f_n'(t))$ für $t \in I$.

Beispiel.

1. $f\colon [0, 2\pi] \to \mathbb{R}^2$ definiert durch $f(t) = (r\cos t, r\sin t)$ (Kreis mit Radius $r > 0$)
2. Sei $g\colon I \to \mathbb{R}$ stetig, I echtes Intervall. Dann ist $f\colon I \to \mathbb{R}^2$ definiert durch $f(t) = (t, g(t))$ eine Kurve.

Definition. Sei $I = [a, b]$ ein kompaktes Intervall mit $a < b$.

1. Eine *Zerlegung* von I ist eine endliche Folge $(t_0, \dots, t_m)$ mit
$$a = t_0 < t_1 < t_2 < \ldots < t_m = b$$
2. Sei $f\colon I \to \mathbb{R}^n$ eine Kurve. Für eine Zerlegung $Z = (t_0, \dots, t_m)$ von I setze
$$p_f(Z) = \sum_{k=1}^{m} \|f(t_k) - f(t_{k-1})\|$$

Definition. Sei $f\colon I \to \mathbb{R}^n$ eine Kurve, wobei $I = [a, b]$ mit $a < b$. f ist *rektifizierbar*, wenn $P_f := \{p_f(Z)\colon Z \text{ Zerlegung von } I\}$ beschränkt ist. Ist dies der Fall, so heißt $L := \sup P_f$ die Länge von f.

Notation. Sei $f\colon [a, b] \to \mathbb{R}^n$ eine Kurve, und sei $f = (f_1, \dots, f_n)$. Setze dann
$$\int_a^b f(t)\,\mathrm{d}t = \left(\int_a^b f_1(t)\,\mathrm{d}t \quad, \dots, \quad \int_a^b f_n(t)\,\mathrm{d}t\right)$$

zur Erinnerung: Für $u = (u_1, \dots, u_n) \in \mathbb{R}^n$ und $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ hatten wir gesetzt $\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i$ und gezeigt $|\langle u, v\rangle| \leq \|u\|\|v\|$.

Lemma 1. *Sei $g\colon [a, b] \to \mathbb{R}^n$ eine Kurve. Dann gilt*
$$\left\|\int_a^b g(t)\,dt\right\| \leq \int_a^b \|g(t)\|\,\mathrm{d}t$$

Beweis. Sei $g = (g_1, \dots, g_n)$ und setze für $1 \leq k \leq n$ $u_k = \int_a^b g_k(t)\,\mathrm{d}t$, sowie $u = (u_1, \dots, u_n)$. Dann gilt
$$\|u\|^2 = \sum_{k=1}^{n} u_k \cdot \int_a^b g_k(t)\,\mathrm{d}t = \int_a^b \sum_{k=1}^{n} u_k g_k(t)\,\mathrm{d}t = \int_a^b \langle u, g(t)\rangle\,\mathrm{d}t \leq \|u\| \int_a^b \|g(t)\|\,\mathrm{d}t$$

Hieraus folgt die Behauptung, da sie für $u = 0$ trivial ist. □

Satz 2. *Jede stetig differenzierbare Kurve $f\colon [a, b] \to \mathbb{R}^n$ ist rektifizierbar und hat die Länge*
$$L = \int_a^b \|f'(t)\|\,\mathrm{d}t$$

Beweis. Setze $L = \int_a^b \|f'(t)\|\,\mathrm{d}t$ und $I = [a, b]$ und sei $f = (f_1, \ldots, f_n)$. Wir zeigen zuerst, dass f rektifizierbar ist. Hierzu zeigen wir, dass $p_f(Z) \leq L$ für jede Zerlegung Z von I. Sei also $Z = (t_0, \ldots, t_m)$ eine Zerlegung von I. Dann gilt

$$\begin{aligned}
p_f(Z) &= \sum_{k=1}^{m} \|f(t_k) - f(t_{k-1})\| = \\
&= \sum_{k=1}^{m} \|(f_1(t_k) - f_1(t_{k-1}), \ldots, f_n(t_k) - f_n(t_{k-1}))\| = \\
&= \sum_{k=1}^{m} \left\| \left(\int_{t_{k-1}}^{t_k} f_1'(t)\,\mathrm{d}t \;, \;\ldots\;, \int_{t_{k-1}}^{t_k} f_n'(t)\,\mathrm{d}t \right) \right\| = \\
&= \sum_{k=1}^{m} \left\| \int_{t_{k-1}}^{t_k} f(t)\,\mathrm{d}t \right\| \leq \sum_{k=1}^{m} \int_{t_{k-1}}^{t_k} \|f'(t)\|\,\mathrm{d}t = \int_a^b \|f'(t)\|\,\mathrm{d}t
\end{aligned}$$

Wir zeigen nun, dass L die Länge von f ist. Hierzu müssen wir noch zeigen, dass für jedes $K < L$ eine Zerlegung Z von I existiert mit $p_f(Z) > K$. Sei also $K < L$ und setze $\varepsilon = L - K > 0$. Wegen I kompakt sind $f_1', \ldots, f_k'$ gleichmäßig stetig. Also existiert ein $\delta > 0$ mit

$$|f_k'(c) - f_k'(d)| < \frac{\varepsilon}{\sqrt{n}(b-a)}$$

für alle $1 \leq k \leq n$ und alle $c, d \in I$ mit $|c - d| < \delta$. Wähle nun eine Zerlegung $Z = (t_0, \ldots, t_n)$ von I so fein, dass für alle $1 \leq k \leq n$ $t_k - t_{k-1} < \delta$ gilt. Nach dem Mittelwertsatz der Integralrechnung gibt es $\eta_k \in [t_{k-1}, t_k]$ mit $L = \sum_{k=1}^{m} \|f'(\eta_k)\|(t_k - t_{k-1})$. Nach dem Mittelwertsatz der Differentialrechnung gibt es $\xi_{k1}, \ldots, \xi_{kn} \in [t_{k-1}, t_k]$ mit $f_j(t_k) - f_j(t_{k-1}) = f_j'(\xi_{kj})(t_k - t_{k-1})$ für $j \in \{1, \ldots, n\}$. Dann gilt $|\eta_k - \xi_{kj}| < \delta$. Setzt man also $v_k = (f_1'(\xi_{k1}), \ldots, f_n'(\xi_{kn}))$, so gilt

$$\begin{aligned}
\|f'(\eta_k) - v_k\| &= \sqrt{(f_1'(\eta_k) - f_1'(\xi_{k1}))^2 + \cdots + (f_n'(\eta_k) - f_n'(\xi_{kn}))^2} < \\
&< \sqrt{\frac{\varepsilon^2}{n(b-a)^2} + \cdots + \frac{\varepsilon^2}{n(b-a)^2}} = \frac{\varepsilon}{b-a}
\end{aligned}$$

Somit erhalten wir

$$\begin{aligned}
L - p_f(Z) &= \sum_{k=1}^{m} \|f'(\eta_k)\|(t_k - t_{k-1}) - \sum_{k=1}^{m} \|f(t_k) - f(t_{k-1})\| = \\
&= \sum_{k=1}^{m} \big(\|f'(\eta_k)(t_k - t_{k-1})\| - \|f(t_k) - f(t_{k-1})\|\big) \leq \\
&\leq \sum_{k=1}^{m} \|f'(\eta_k)(t_k - t_{k-1}) - (f(t_k) - f(t_{k-1}))\| = \\
&= \sum_{k=1}^{m} \|f'(\eta_k)(t_k - t_{k-1}) - v_k(t_k - t_{k-1})\| = \\
&= \sum_{k=1}^{m} \|f'(\eta_k) - v_k\|(t_k - t_{k-1}) < \frac{\varepsilon}{b-a} \cdot \sum_{k=1}^{m}(t_k - t_{k-1}) = \varepsilon
\end{aligned}$$

Somit ist $p_f(Z) > K$. $\square$

Beispiel. Sei $f\colon [0, 2\pi] \to \mathbb{R}^2, t \mapsto (r\cos t, r\sin t)$ der Kreis mit Radius r. Dann

$$L = \int_0^{2\pi} \sqrt{r^2\cos^2 t + r^2 \sin^2 t}\,\mathrm{d}t = \int_0^{2\pi} r\,\mathrm{d}t = 2\pi r$$

12 Mehrdimensionale Differenzierbarkeit

Wir benötigen hier elementare Kenntnisse und Notationen aus der linearen Algebra.

Konvention.

- Matrix = reelle Matrix
- Im Zusammenhang mit Matrizenrechnungen betrachten wir Elemente von $\mathbb{R}^n$ als Spaltenvektoren, d.h. für eine $m \times n$-Matrix A und $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ist:

$$Ax = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

- Außerdem unterscheiden wir nicht zwischen $m \times n$-Matrizen und linearen Abbildungen $\mathbb{R}^n \to \mathbb{R}^m$. Ist also A eine $m \times n$-Matrix, so ist $A\colon \mathbb{R}^n \to \mathbb{R}^m, x \mapsto Ax$. Beachte, dass A stetig ist.

Frage. Sei $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m, a \in U$. Wann ist f differenzierbar in $a \in U$? $\lim\limits_{x \to a} \frac{f(x)-f(a)}{x-a}$ gibt keinen Sinn.

Wir hatten im Fall $m = n = 1$ aber schon umformuliert. Es gilt (falls a Häufungspunkt von U): f ist differenzierbar in a genau dann, wenn gilt:

Es gibt ein $c \in \mathbb{R}$ und eine in a stetige Funktion $\tilde{r}\colon U \to \mathbb{R}$ mit $\tilde{r}(a) = 0$ und $f(x) = f(a) + c(x-a) + \tilde{r}(x)(x-a)$ für alle $x \in U$. Dabei ist c eindeutig bestimmt und es gilt $c = f'(a)$.

Wir können (1) noch geringfügig äquivalent umformulieren zu:

Es gibt ein $c \in \mathbb{R}$ und eine a stetige Funktion $r\colon U \to \mathbb{R}$ mit $r(a) = 0$ und $f(x) = f(a) + c(x-a) + r(x)|x-a|$ für alle $x \in U$. Dies geht einfach durch die Festsetzung

$$r(x) = \begin{cases} \tilde{r}(x) & \text{falls } a \le x \\ -\tilde{r}(x) & \text{falls } x < a \end{cases}$$

Die Bedingung lässt sich kanonisch auf den mehrdimensionalen Fall übertragen.

Definition. Sei $f\colon U \to \mathbb{R}^m$, wobei $U \subseteq R^n$ offen sei. Weiterhin sei $a \in U$.

- f ist *differenzierbar* in a, wenn es eine $m \times n$-Matrix C und eine in a stetige Funktion $r\colon U \to \mathbb{R}^n$ gibt mit $r(a) = 0$ und $f(x) = f(a) + C(x-a) + r(x)\|x-a\|$ für alle $x \in U$.
- f ist differenzierbar in einer Teilmenge $D \subseteq U$, wenn f in jedem $a \in D$ differenzierbar ist.
- f ist differenzierbar, wenn f in jedem $a \in U$ differenzierbar ist.

Für $m = n = 1$ ist dies also äquivalent zur früheren Definition. Wir wollen zeigen, dass die Matrix C durch f und a eindeutig bestimmt ist.

Lemma 1. *Sei $f\colon U \to \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$ offen. Sei f differenzierbar in $a \in U$ und seien daher $C \in \mathbb{R}^{m\times n}$ und $r\colon U \to \mathbb{R}^m$ eine in a stetige Funktion mit $r(a) = 0$ und $f(x) = f(a) + C(x-a) + r(x)\|x-a\|$ für alle $x \in U$.*
Dann gilt für alle $x \in \mathbb{R}^n$:

$$Cx = \lim_{t\to 0} \frac{f(a+tx) - f(a)}{t}$$

Beweis. Für $t \neq 0$ mit $a + tx \in U$ gilt:

$$\frac{f(a+tx) - f(a)}{t} = \frac{f(a) + Ctx + r(a+tx)\|tx\| - f(a)}{t} = $$
$$= Cx + r(a+tx)\frac{|t|\|x\|}{t}$$

Wegen r stetig in a und $r(a) = 0$ ist aber $\lim_{t\to 0} r(a+tx) = 0$. Insgesamt folgt also die Behauptung. □

Ist also f differenzierbar in a und C $m \times n$-Matrix wie in Definition, so ist C eindeutig bestimmt durch f, a. Wir nennen also C die *Ableitung* (Differential, Jacobimatrix) von f im Punkt a und setzen:

$$C = f'(a) = \mathrm{D}f(a)$$

Beispiel. 1. Sei $f\colon \mathbb{R}^n \to \mathbb{R}^m$ eine konstante Abbildung, d.h. etwa $f(x) = c$ für alle $x \in \mathbb{R}^n$. Dann ist für alle $a \in \mathbb{R}^n$ $f'(a) = 0$, denn:

$$f(x) = f(a) + 0(x-a) + 0\|x-a\|$$

2. Sei $f\colon \mathbb{R}^n \to \mathbb{R}^m$ linear, d.h. etwa $f(x) = Bx$ für alle $x \in \mathbb{R}^n$. Dann ist für alle $a \in \mathbb{R}^n$ $f'(a) = B$, denn:

$$f(x) = f(a) + B(x-a) + 0\|x-a\| \quad \text{für alle } x \in \mathbb{R}^n$$

Satz 2. *Sei $f\colon U \to \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$ offen, und sei $a \in U$. Ist f differenzierbar in a, so ist f stetig in a.*

Beweis. Sei f differenzierbar. Dann existiert $r\colon U \to \mathbb{R}^m$ mit r stetig in a und $r(a) = 0$ mit $f(x) = f(a) + \mathrm{D}f(a)(x-a) + r(x)\|x-a\|$. Da $\mathrm{D}f(a)$ als lineare Abbildung stetig ist, ist also auch f in a stetig. □

Satz 3. *Sei $f\colon U \to \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^m$ offen, und sei $a \in U$. Weiterhin sei $f = (f_1, \ldots, f_m)$. Dann ist f differenzierbar in a genau dann, wenn $f_1, \ldots, f_m$ differenzierbar in a sind.*
Ist dies der Fall, so gilt:

$$f'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix}$$

Beweis. Sei f differenzierbar in $a \in U$. Dann existiert eine in a stetige Funktion $r\colon U \to \mathbb{R}^m$ mit $r(a) = 0$ und $f(x) = f(a) + f'(a)(x-a) + r(x)\|x-a\|$ für alle $x \in U$. Seien $r = (r_1, \ldots, r_m)$ und $v_1, \ldots, v_m$ die Zeilen von $f'(a)$. Ist dann $1 \leq k \leq m$, so ist r_k stetig in a, und es gilt $r_k(a) = 0$. Außerdem ist $f_k(x) = f_k(a) + v_k(x-a) + r_k(x)\|x-a\|$ für alle $x \in U$. Also ist f_k differenzierbar in a, und es gilt $f_k'(a) = v_k$.
Die andere Richtung folgt analog. □

Satz 4 (Linearität)**.** *Seien $f, g\colon U \to \mathbb{R}^m$ differenzierbar in a, $U \subseteq \mathbb{R}^n$ offen, und sei $c \in \mathbb{R}$. Dann sind auch $f + g$ und cf differenzierbar in a und es gilt:*

$$(f+g)'(a) = f'(a) + g'(a)$$
$$(cf)'(a) = cf'(a)$$

Beweis. Seien $r, s\colon U \to \mathbb{R}^m$ in a stetig und $r(a) = s(a) = 0$ und für alle $x \in U$:

$$f(x) = f(a) + f'(a)(x-a) + r(x)\|x-a\|$$
$$g(x) = g(a) + g'(a)(x-a) + s(x)\|x-a\|$$

Dann ist für alle $x \in U$

$$(f+g)(x) = f(x) + g(x) = (f+g)(a) + \big(f'(a) + g'(a)\big)(x-a) + (r(x) + s(x))\|x-a\|$$

und

$$(cf)(x) = cf(x) + cf'(a)(x-a) + cr(x)\|x-a\|$$

Weiterhin sind $r+s$ und cr in a stetig und $(r+s)(a) = 0 = cr(a)$. Hieraus folgt die Behauptung. □

Definition (Die Abbildungsnorm)**.** Sei A eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$. Wir setzen dann:

$$\|A\| = \sup\{\|Ax\|\colon x \in \mathbb{R}^n \text{ und } \|x\| \leq 1\}$$

$\|A\|$ ist die Norm von A. Es ist $\|A\| < \infty$, denn $S = \{x \in \mathbb{R}^n\colon \|x\| \leq 1\}$ ist abgeschlossen und beschränkt in $\mathbb{R}^n$, also kompakt. Weiterhin ist $g\colon \mathbb{R}^n \to \mathbb{R}, x \mapsto \|Ax\|$ stetig. Also ist $g(S)$ kompakt und daher beschränkt, woraus $\|A\| = \sup g(S) < \infty$ folgt.

Man rechnet leicht nach, dass die Abbildungsnorm eine Norm auf dem Vektorraum $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ ist. Weiterhin gilt für alle $x \in \mathbb{R}^n$:

$$\|Ax\| \leq \|A\|\|x\|$$

denn sei $x \in \mathbb{R}^n$. Für $x = 0$ ist die Behauptung trivial. Sei also $x \neq 0$. Setze dann $z = \frac{1}{\|x\|}x$. Also ist $\|Ax\| = \|A(\|x\|z)\| = \|\|x\|Az\| = \|x\|\|Az\| \leq \|A\|\|x\|$.

Satz 5 (Kettenregel)**.** *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen, und sei $f\colon U \to \mathbb{R}^m$ und $g\colon V \to \mathbb{R}^k$ mit $f(U) \subseteq V$. Sei f in a differenzierbar und g in $f(a)$ differenzierbar. Dann ist $g \circ f$ in a differenzierbar, und es gilt:*

$$(g \circ f)'(a) = g'\left(f(a)\right) f'(a)$$

Beweis. Seien $C := f'(a)$, $D := g'(f(a))$ und setze $b := f(a)$. Also ist zu zeigen, dass $(g \circ f)'(a) = DC$. Es existieren Funktionen $r\colon U \to \mathbb{R}^m$ und $s\colon V \to \mathbb{R}^k$ mit r stetig in a, s stetig in b und $r(a) = 0 = s(b)$, sowie für alle $x \in U$ und $y \in V$:

$$\begin{aligned} f(x) &= f(a) + C(x-a) + r(x)\|x-a\| \\ g(y) &= g(b) + D(y-b) + s(y)\|y-b\| \end{aligned}$$

Dann ist für alle $x \in U$

$$\begin{aligned} (g \circ f)(x) = g(f(x)) &= g(b) + D(f(x)-b) + s(f(x))\|f(x)-b\| \\ &= g(b) + D(C(x-a) + r(x)\|x-a\|) + s(f(x))\|f(x)-b\| \\ &= g(b) + DC(x-a) + \eta(x)\|x-a\| \end{aligned}$$

mit

$$\eta(x) = \begin{cases} Dr(x) + s(f(x))\frac{\|f(x)-b\|}{\|x-a\|} & \text{für } x \neq a \\ 0 & \text{für } x = a \end{cases}$$

Somit genüg es zu zeigen, dass η stetig in a ist, d.h. dass $\lim_{x\to a} \eta(x) = 0$. Nun ist aber $\lim_{x\to a} r(x) = 0$, also $\lim_{x\to a} Dr(x) = 0$. Weiterhin ist f stetig in a, also ist wegen s stetig in $f(a)$ $\lim_{x\to a} s(f(x)) = s(f(a)) = 0$.

Also genügt es zu zeigen, dass eine Umgebung W von a existiert und ein $K \in \mathbb{R}$ mit $\frac{\|f(x)-b\|}{\|x-a\|} \leq K$ für alle $x \in U \cap W$ mit $x \neq a$. Nun ist aber für $x \in U, x \neq a$

$$\begin{aligned}\frac{\|f(x)-b\|}{\|x-a\|} &= \frac{\|C(x-a)+r(x)\|x-a\|\|}{\|x-a\|}\\ &\leq \frac{\|C(x-a)\|+\|r(x)\|x-a\|\|}{\|x-a\|}\\ &\leq \frac{\|C\|\|x-a\|+\|r(x)\|\|x-a\|}{\|x-a\|}\\ &\leq \|C\|+\|r(x)\|\end{aligned}$$

Da r stetig in a ist, folgt also das Gewünschte. □

Sei $f\colon U \to \mathbb{R}^m$ differenzierbar in $a \in U$, wobei $U \subseteq \mathbb{R}^n$ offen. Wie berechnet man $f'(a)$?
Ist $f = (f_1, \ldots, f_m)$, so gilt nach Satz 3

$$f'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix}$$

Es genügt also den Fall $m = 1$ zu betrachten. Sei also $f'(a) = v = (c_1, \ldots, c_n)$. Nach Lemma 1 gilt dann für alle $x \in \mathbb{R}^n$

$$vx = \lim_{t\to 0} \frac{f(a+tx)-f(a)}{t}$$

Speziell also für die Einheitsvektoren $e_1, \ldots, e_n$

$$c_i = ve_i = \lim_{t\to 0} \frac{f(a+te_i)-f(a)}{t}$$

Setzt man also für $1 \leq i \leq n$ $E_i = \{x \in \mathbb{R}\colon (a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n) \in U\}$ und definiert $g_i\colon E_i \to \mathbb{R}, x \mapsto f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n)$, so ist

$$c_i = \lim_{t\to 0} \frac{f(a+te_i)+f(a)}{t} = \lim_{t\to 0} \frac{g_i(a_i+t)-g_i(a_i)}{t} = g_i'(a_i)$$

Damit ist alles auf den eindimensionalen Fall zurückgeführt, wenn wir wissen, dass f in a differenzierbar ist.

Dies motiviert folgende Definitionen:

Definition. Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \to \mathbb{R}, a \in U$.

1. Sei $v \in \mathbb{R}^n$. Existiert der Grenzwert $\lim_{t\to 0} \frac{f(a+tv)-f(a)}{t}$, so nennt man ihn die *Ableitung von f im Punkt a in Richtung v* und bezeichnet ihn mit $D_v f(a)$.

2. ist e_i der i-te Einheitsvektor im $\mathbb{R}^n$, so setzen wir $D_i f(a) := D_{e_i} f(a)$. $D_i f(a)$ heißt die *i-te partielle Ableitung von f in a*. Statt $D_i f(a)$ schreibt man auch $\frac{\partial f}{\partial x_i}$.

3. f ist *partiell differenzierbar* in a, falls $D_i f(a)$ für alle $1 \le i \le n$ existiert. Ist dies der Fall, so setzt man

$$\nabla f(a) = (D_1 f(a), \dots, D_n f(a))$$

   $\nabla f(a)$ ist der *Gradient von f im Punkt a.*

4. f ist *partiell differenzierbar*, falls f in jedem $a \in U$ partiell differenzierbar ist.

Bemerkung. f, U, a wie oben, $1 \le i \le n$, $E_i = \{x \in \mathbb{R}\colon (a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n) \in U\}$. Definiere $g_i\colon E_i \to \mathbb{R}, x \mapsto f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, n)$. Dann ist $D_i f(a) = g_i'(a_i)$.

Beweis.

$$D_i f(a) = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{t \to 0} \frac{g_i(a_i + t) - g_i(a_i)}{t} = g_i'(a_i)$$

□

Beispiel. Sei $f\colon \mathbb{R}^3 \to \mathbb{R}, (x_1, x_2, x_3) \mapsto x_1 x_2 x_3 + x_1^2 x_2 + x_2 x_3^2$. Dann ist für $x = (x_1, x_2, x_3) \in \mathbb{R}^3$:

$$\begin{aligned} D_1 f(x) &= x_2 x_3 + 2x_1 x_2 \\ D_2 f(x) &= x_1 x_3 + x_1^2 + x_3^2 \\ D_3 f(x) &= x_1 x_2 + 2x_2 x_3 \end{aligned}$$

Der folgende Satz ist nur eine Umformulierung eines Spezialfalls von Lemma 1.

Satz 6. *Sei $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ sei differenzierbar in a. Dann existiert für alle $v \in \mathbb{R}^n$ die Ableitung von f im Punkt a in Richtung v, und es gilt $D_v f(a) = f'(a)v$. Insbesondere gilt also $f'(a) = \nabla f(a)$.*

Beweis. Nach Lemma 1 ist

$$f'(a)v = \lim_{t \to 0} \frac{f(a + tv) - f(a)}{t} = D_v f(a)$$

Sei $f'(a) = (c_1, \dots, c_n)$. Dann gilt für $1 \le i \le n$

$$c_i = f'(a)e_i = D_i f(a)$$

Also ist $f'(a) = (c_1, \dots, c_n) = \nabla f(a)$. □

Korollar. *Sei $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ differenzierbar in a. Sei $f = (f_1, \dots, f_m)$. Dann gilt*

$$f'(a) = \begin{pmatrix} \nabla f_1(a) \\ \vdots \\ \nabla f_m(a) \end{pmatrix} = \begin{pmatrix} D_1 f_1(a) & \cdots & D_n f_1(a) \\ \vdots & & \vdots \\ D_1 f_m(a) & \cdots & D_n f_m(a) \end{pmatrix}$$

Beweis. Nach Satz 3 sind $f_1, \ldots, f_m$ differenzierbar in a, und es gilt

$$f'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix}$$

Somit folgt aus Satz 6 die Behauptung. □

Aber aus der partiellen Differenzierbarkeit von einer Funktion $f\colon U \to \mathbb{R}$ in a folgt nicht die Differenzierbarkeit von f in a. Es folgt nicht einmal, dass f stetig in a ist.

Beispiel. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch:

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & \text{für } (x_1, x_2) \neq (0,0) \\ 0 & \text{sonst} \end{cases}$$

f ist partiell differenzierbar, denn:
Sei $(x_1, x_2) \in \mathbb{R}^2$. Ist $(x_1, x_2) \neq (0,0)$, so ist f offenbar partiell differenzierbar in (x_1, x_2). Weiterhin:

$$D_1 f(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = 0$$
$$D_2 f(0,0) = \lim_{t \to 0} \frac{f(0,t) - f(0,0)}{t} = 0$$

f ist jedoch in 0 nicht stetig, denn sei $a_k = (1/k, 1/k)$. Dann ist $\lim_{k\to\infty} a_k = 0$, aber $\lim_{k\to\infty} f(1/k, 1/k) = \lim_{k\to\infty} 1/2 = 1/2 \neq 0$.

Satz 7. *Sei $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ partiell differenzierbar. Sind alle partiellen Ableitungen von f in $a \in U$ stetig, so ist f differenzierbar in a.*

Beweis. Sei $v = \nabla f(a) = (D_1 f(a), \ldots, D_n f(a))$. Definiere dann den Rest $r\colon U \to \mathbb{R}$ durch

$$r(x) = \begin{cases} \frac{f(x) - f(a) - v(x-a)}{\|x-a\|} & \text{falls } x \in U \setminus \{a\} \\ 0 & \text{falls } x = a \end{cases}$$

Dann ist für alle $x \in U$:

$$f(x) = f(a) + v(x-a) + r(x)\|x-a\|$$

Also genügt es zu zeigen, dass r in a stetig ist, d.h. dass gilt $\lim_{x\to a} r(x) = 0$.

Sei hierzu $\varepsilon > 0$. Da $D_1 f, \ldots, D_n f$ in a stetig sind, existiert $\delta > 0$ mit $U_\delta(a) \subseteq U$ und

$$|D_i f(x) - D_i f(a)| < \frac{\varepsilon}{n} \quad \text{für alle } x \in U_\delta(a) \text{ und alle } 1 \leq i \leq n$$

Wir zeigen, dass $|r(x)| < \varepsilon$ für alle $x \in U_\delta(a)$. Sei also $x \in U_\delta(a)$, $x = (x_1, \ldots, x_n)$. Weiterhin sei $a = (a_1, \ldots, a_n)$. Ohne Einschränkung sei $x \neq a$. Für $i \leq n$ setze $v_i = (x_1, \ldots, x_i, a_{i+1}, \ldots, a_n)$. Dann ist $d(v_i, a) = \sqrt{\sum_{k=1}^{i} (x_k - a_k)^2} \leq \sqrt{\sum_{k=1}^{n} (x_k - a_k)^2} =$

$d(x,a) < \delta$. Also $v_i \in U_\delta(a)$. Beachte, dass $v_0 = a$ und $v_n = x$, also $f(x) - f(a) = \sum_{i=1}^n f(v_i) - f(v_{i-1})$. Nun existiert für alle $1 \leq i \leq n$ eine $w_i \in U_\delta(a)$ mit

$$f(v_i) - f(v_{i-1}) = D_i f(w_i)(x_i - a_i)$$

Ist $x_i = a_i$, so ist die Behauptung trivial. Sei also $x_i \neq a_i$ und etwa $a_i < x_i$. Dann ist $(x_1, \ldots, x_{i-1}, t, a_{i+1}, \ldots, a_n) \in U_\delta(a)$ für alle $t \in [a_i, x_i]$. Wir können also $g_i \colon [a_i, x_i] \to \mathbb{R}$ definieren durch $g_i(t) = f(x_1, \ldots, x_{i-1}, t, a_{i+1}, \ldots, a_n)$. Nach Voraussetzung ist dann g_i differenzierbar, denn es ist $g_i'(t) = D_i f(x_1, \ldots, x_{i-1}, t, a_{i+1}, \ldots, a_n)$. Nach dem Mittelwertsatz der Differentialrechnung existiert also ein $\xi_i \in (a_i, x_i)$ mit $g_i(x_i) - g_i(a_i) = g_i'(\xi_i)(x_i - a_i)$. Setze dann $w_i = (x_1, \ldots, x_{i-1}, \xi_i, a_{i+1}, \ldots, a_n)$. Dann ist

$$f(v_i) - f(v_{i-1}) = g_i(x_i) - g_i(a_i) = g_i'(\xi_i)(x_i - a_i) = D_i f(w_i)(x_i - a_i)$$

und $w_i \in U_\delta(a)$. Falls $x_i < a_i$ betrachte analog $[x_i, a_i]$. Insgesamt erhalten wir schließlich

$$\begin{aligned}
|r(x)| &= \left| \frac{f(x) - f(a) - v(x-a)}{\|x-a\|} \right| = \\
&= \frac{\left|\sum_{i=1}^n f(v_i) - f(v_{i-1}) - \sum_{i=1}^n D_i f(a)(x_i - a_i)\right|}{\|x-a\|} = \\
&= \frac{\left|\sum_{i=1}^n (D_i f(w_i) - D_i f(a))(x_i - a_i)\right|}{\|x-a\|} \leq \\
&\leq \sum_{i=1}^n |D_i f(w_i) - D_i f(a)| \underbrace{\frac{|x_i - a_i|}{\|x-a\|}}_{\leq 1} < n \frac{\varepsilon}{n} = \varepsilon
\end{aligned}$$

□

Korollar. *Sei $U \subseteq \mathbb{R}^n$ offen und $f \colon U \to \mathbb{R}^m$ und sei $f = (f_1, \ldots, f_m)$. Seien $f_1, \ldots, f_m$ partiell differenzierbar. Sind alle partiellen Ableitungen von $f_1, \ldots, f_m$ in $a \in U$ stetig, so ist f in a differenzierbar.*

Beweis. Nach Satz 7 sind $f_1, \ldots, f_m$ differenzierbar in a. Also ist f differenzierbar in a. □

Definition. Sei $U \subseteq \mathbb{R}^n$ offen, $f \colon U \to \mathbb{R}^m$ und sei $f = (f_1, \ldots, f_m)$. f ist *stetig differenzierbar*, wenn f differenzierbar ist und alle partiellen Ableitungen $D_i f_j$ für $1 \leq i \leq n, 1 \leq j \leq m$ stetig sind.

Bemerkung. Sei f wie oben differenzierbar. Wir können eine $m \times n$-Matrix $A = (a_{ij})_{i,j}$ mit dem Vektor

$$(a_{11}, \ldots, a_{1n}, \ldots, a_{m1}, \ldots, a_{mn}) \in \mathbb{R}^{mn}$$

identifizieren. Dann ist f stetig differenzierbar genau dann, wenn die Abbildung $f' \colon U \to \mathbb{R}^{mn}$ stetig ist. (Dies gilt auch für die Abbildungsnorm!)

Definition. Für $a, b \in \mathbb{R}^n$ sei $[a,b] = \{a + t(b-a) \colon 0 \leq t \leq 1\}$. $[a,b]$ ist die *Strecke* von a nach b.

Bemerkung. Für $a, b \in \mathbb{R}, a \leq b$ stimmt diese Notation mit der Intervallnotation überein.

Satz 8 (Mittelwertsatz)**.** *Sei $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ stetig differenzierbar. Weiterhin seien $a, b \in U$ und $[a, b] \subseteq U$. Dann gilt:*

$$f(b) - f(a) = \int_0^1 f'(a + t(b-a))(b-a)\,\mathrm{d}t$$

Beweis. Sei $V = \{t \in \mathbb{R}\colon a + t(b-a) \in U\}$. Dann ist V offen und nach Voraussetzung $[0,1] \subseteq V$. Definiere $g\colon V \to \mathbb{R}^m$ durch $g(t) = f(a + t(b-a))$. Also ist $g(0) = f(a)$ und $g(1) = f(b)$. Nach Kettenregel ist g differenzierbar, und es gilt:

$$g'(t) = f'(a + t(b-a))(b-a)$$

Da f' stetig ist, ist also auch g' stetig. Sei jetzt $g = (g_1, \ldots, g_m)$. Dann gilt:

$$\begin{aligned} f(b) - f(a) &= g(1) - g(0) \\ &= (g_1(1) - g_1(0), \ldots, g_m(1) - g_m(0)) \\ &= \left(\int_0^1 g_1'(t)\,\mathrm{d}t, \ldots, \int_0^1 g_m'(t)\,\mathrm{d}t \right) \\ &= \int_0^1 g'(t)\,\mathrm{d}t \end{aligned}$$

□

Bemerkung. Sei oben $m = n = 1$. Dann hat man

$$f(b) - f(a) = \int_0^1 f'(a + t(b-a))(b-a)\,\mathrm{d}t = (b-a)\int_0^1 f'(a + t(b-a))\,\mathrm{d}t$$

Aber nach dem Mittelwertsatz der Integralrechnung existiert ein $0 \le t_0 \le 1$ mit

$$\int_0^1 f'(a + t(b-a))\,\mathrm{d}t = f'(a + t_0(b-a))$$

Setze also $\xi_0 = a + t_0(b-a)$. Dann gilt $f(b) - f(a) = f'(\xi_0)(b-a)$.

Satz 9 (Schrankensatz)**.** *Seien $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^m$ stetig differenzierbar. Weiterhin seien $a, b \in U$ mit $[a, b] \subseteq U$. Setze $M = \sup\{\|f'(x)\|\colon x \in [a,b]\} < \infty$. Dann gilt $\|f(b) - f(a)\| \le M\|b-a\|$.*

Beweis. Beachte, dass $M = \sup\{\|f'(a + t(b-a))\|\colon t \in [0,1]\}$. Setze zur Abkürzung $h(t) = f'(a + t(b-a))(b-a)$. Nach Mittelwertsatz gilt:

$$\|f(b) - f(a)\| = \left\| \int_0^1 h(t)\,\mathrm{d}t \right\|$$

Da h stetig ist, gilt nach Lemma 1 aus Paragraph 11

$$\left\| \int_0^1 h(t)\,\mathrm{d}t \right\| \le \int_0^1 \|h(t)\|\,\mathrm{d}t$$

Setze nun $M^* = \sup\{\|h(t)\|\colon t \in [0,1]\}$. Dann gilt $\int_0^1 \|h(t)\|\,\mathrm{d}t \le M^*$. Aber für $t \in [0,1]$ ist $\|h(t)\| = \|f'(a + t(b-a))(b-a)\| \le \|f'(a + t(b-a))\|\|b-a\| \le M\|b-a\|$. Also ist $M^* \le M\|b-a\|$. Insgesamt $\|f(b) - f(a)\| \le M\|b-a\|$. □

Bemerkung. Beachte, dass in Satz 9 $M < \infty$, da die Abbildungsnorm stetig ist und $[a, b]$ kompakt ist.

Bemerkung. Im Schrankensatz wird die Schranke M mit Hilfe der Abbildungsnorm definiert. Für $m = 1$ oder $n = 1$ stimmt diese aber mit der euklidischen Norm überein. Sei zuerst $m = 1$. Sei also $v = (v_1, \ldots, v_n), v \neq 0$ und $\|v\|'$ die Abbildungsnorm, $\|v\|$ die euklidische Norm. Für $(x_1, \ldots, x_n) = x \in \mathbb{R}^n$ mit $\|x\| \leq 1$ gilt dann $\|vx\| = \|\langle v, x\rangle\| \leq \|v\|\|x\| \leq \|v\|$. Also $\|v\|' \leq \|v\|$. Andererseits gilt für $x = \frac{v}{\|v\|}$, dass $\|x\| = 1$ und $vx = \frac{\langle v,v\rangle}{\|v\|} = \|v\|$. Also $\|v\| \leq \|v\|'$.

Sei nun $n = 1$, $v = (v_1, \ldots, v_n) \in \mathbb{R}^m$ und $\|v\|'$ die Abbildungsnorm, $\|v\|$ die euklidische Norm. Dann gilt

$$\|v\|' = \sup\{\|vx\| \colon x \in \mathbb{R}, |x| \leq 1\} = \sup\{|x|\|v\| \colon x \in \mathbb{R}, |x| \leq 1\} = \|v\|$$

Definition. Für $a_0, \ldots, a_k \in \mathbb{R}^n$ setze

$$[a_0, a_1, \ldots, a_k] := \bigcup_{i=1}^{k} [a_{i-1}, a_i]$$

$[a_0, a_1, \ldots, a_k]$ ist ein *Streckenzug von a_0 nach a_k*.

Lemma 10. *Sei $U \subseteq \mathbb{R}^n$ offen. Dann ist U genau dann zusammenhängend, wenn es für alle $a, b \in U$ einen Streckenzug S von a nach b gibt mit $S \subseteq U$.*

Beweis. Sei U zusammenhängend und seien $a, b \in U$. Setze

$$V_0 = \{x \in U \colon \exists S \text{ Streckenzug von } a \text{ nach } x. S \subseteq U\}.$$

1. V_0 ist offen, denn sei $x \in V_0$. Wegen U offen existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq U$. Dann ist aber auch $U_\varepsilon(x) \subseteq V_o$, denn sei $w \in U_\varepsilon(x)$. Wegen $x \in V_0$ existiert ein Streckenzug $S = [a_0, \ldots, a_k]$ von a nach x mit $S \subseteq U$. Wegen $w \in U_\varepsilon(x)$ ist aber $[x, w] \subseteq U_\varepsilon(x) \subseteq U$. Also ist $T = [a_0, \ldots, a_k, w]$ ein Streckenzug von a nach w mit $T \subseteq U$. Somit ist $w \in V_0$. Also ist $U_\varepsilon(x) \subseteq V_0$.

2. $V_1 = U \setminus V_0$ ist offen, denn sei $x \in V_1$. Wegen U offen existiert ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq U$. Dann ist aber auch $U_\varepsilon(x) \subseteq V_1$, denn andernfalls existiert ein $w \in U_\varepsilon(x)$ mit $w \notin V_1$. Also ist wegen $w \in U$ $w \in V_0$. Also existiert ein Streckenzug $S = [a_0, \ldots, a_k]$ von a nach w mit $S \subseteq U$. Nun ist aber $[w, x] \subseteq U_\varepsilon(x) \subseteq U$. Also ist $T = [a_0, \ldots, a_k, x]$ ein Streckenzug von a nach x mit $T \subseteq U$. Somit wäre $x \in V_0$, was ein Widerspruch ist.

Somit ist $U = V_0 \cup V_1$, V_0 und V_1 sind offen, $V_0 \cap V_1 = \emptyset$. Wegen $a \in V_0$ ist $V_0 \neq \emptyset$. Da U zusammenhängend ist, ist also $V_1 = \emptyset$, d.h. $V_0 = U$. Somit ist $b \in V_0$, woraus das Gewünschte folgt.

Sei umgekehrt die rechte Seite von Lemma 10 erfüllt. Wir müssen zeigen, dass U zusammenhängend ist. Seien V_0, V_1 offen mit $U \subseteq V_0 \cup V_1$, $U \cap V_0 \neq \emptyset \neq U \cap V_1$. Wir müssen zeigen, dass $V_0 \cap V_1 \neq \emptyset$. Wähle hierzu $a \in U \cap V_0$ und $b \in U \cap V_1$. Sei also

$S = [a_0, \dots, a_k]$ ein Streckenzug von a nach b mit $S \subseteq U$. Definiere $f\colon [0,k] \to \mathbb{R}^n$ durch $f(j+t) = a_j + t(a_{j+1} - a_j)$ für $j \in \{0, \dots, k-1\}$ und $0 \leq t \leq 1$. Dann ist f stetig und $f([0,k]) = S$. Also sind $f^{-1}(V_0)$ und $f^{-1}(V_1)$ offen (in $[0,k]$), $f^{-1}(V_0) \cap [0,k] \neq \emptyset \neq f^{-1}(V_1) \cap [0,k]$ und $[0,k] \subseteq f^{-1}(V_0) \cup f^{-1}(V_1)$. Wegen $[0,k]$ zusammenhängend ist also $f^{-1}(V_0) \cap f^{-1}(V_1) \neq \emptyset$ und somit $V_0 \cap V_1 \neq \emptyset$. □

Satz 11. *Seien $U \subseteq \mathbb{R}^n$ offen und zusammenhängend und $f\colon U \to \mathbb{R}^n$ differenzierbar. Ist $f' = 0$, so ist f konstant.*

Beweis. Seien $a, b \in U$. Wir müssen zeigen, dass $f(a) = f(b)$. Nach Lemma 10 existiert ein Streckenzug $S = [a_0, \dots, a_k]$ von a nach b mit $S \subseteq U$. Nach dem Schrankensatz gilt dann, dass $f(a_{j+1}) - f(a_j) = 0$ für alle $j < k$. Also $f(a) = f(a_0) = f(a_1) = \cdots = f(a_k) = f(b)$. □

Definition. Seien $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ differenzierbar.

1. f ist *zweimal differenzierbar*, wenn f' auch differenzierbar ist.

2. f ist *zweimal stetig differenzierbar*, wenn f' sogar stetig differenzierbar ist.

Bemerkung. Sei f zweimal differenzierbar. Dann ist nach Satz 6 $f' = (D_1 f, \dots, D_n f)$. Wegen f' differenzierbar ist also nach Satz 3 für $1 \leq i \leq n$ $D_i f$ differenzierbar. Nach Satz 6 ist aber dann $D_i f$ partiell differenzierbar. Für $1 \leq j \leq n$ existieren also $D_j D_i f$. Ist f sogar zweimal stetig differenzierbar, so ist $D_j D_i f$ stetig.

Satz 12 (Satz von Schwarz). *Seien $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt für alle $1 \leq i, j \leq n$*

$$D_j D_i f = D_i D_j f$$

Beweis. Da ja bei der partiellen Differentiation nach x_i die anderen Variablen fest gehalten werden, können wir o.E. annehmen, dass $n = 2$, $i = 1$, $j = 2$. Sei nun $a = (a_1, a_2) \in \mathbb{R}^2$. Wir wollen zeigen, dass $D_2 D_1 f(a) = D_1 D_2 f(a)$. Wegen U offen existiert ein $\rho > 0$ mit $U_\rho(a) \subseteq U$. Setze $\varepsilon = \rho/\sqrt{2}$ und $V = \{(x_1, x_2) \in \mathbb{R}^2 \colon |x_1 - a_1| < \varepsilon \wedge |x_2 - a_2| < \varepsilon\}$. Dann ist $V \subseteq U_\rho(a) \subseteq U$. Für $0 < \delta < \varepsilon$ setze $V_\delta = \{(x_1, x_2) \in \mathbb{R}^2 \colon |x_1 - a_1| < \delta \wedge |x_2 - a_2| < \delta\}$. Wir zeigen zuerst für alle $0 < \delta < \varepsilon$, dass $w, x \in V_\delta$ existieren mit $D_2 D_1 f(w) = D_1 D_2 f(x)$. Sei dafür $0 < \delta < \varepsilon$ und $b_k = a_k + \delta$ für $k = 1, 2$. Definiere $g\colon [a_1, b_1] \to \mathbb{R}$ durch $g(t) = f(t, b_2) - f(t, a_2)$. Da g differenzierbar ist, existiert ein $\xi \in [a_1, b_1]$ mit $g(b_1) - g(a_1) = g'(\xi)(b_1 - a_1) = (D_1 f(\xi, b_2) - D_1 f(\xi, a_2)) \cdot \delta$. Definiere nun $h\colon [a_2, b_2] \to \mathbb{R}$ durch $h(t) = D_1 f(\xi, t)$. Da auch h differenzierbar ist, existiert nach dem Mittelwertsatz ein $\eta \in [a_2, b_2]$ mit $h(b_2) - h(a_2) = h'(\eta)(b_2 - a_2) = D_2 D_1 f(\xi, \eta) \cdot \delta$. Insgesamt erhalten wir also $(\xi, \eta) \in V_\delta$ und

$$\begin{aligned} f(b_1, b_2) - f(b_1, a_2) - f(a_1, b_2) + f(a_1, a_2) &= g(b_1) - g(a_1) = \\ &= (D_1 f(\xi, b_2) - D_1 f(\xi, a_2)) \cdot \delta = \\ &= (h(b_2) - h(a_2)) \cdot \delta = D_2 D_1 f(\xi, \eta) \cdot \delta^2 \end{aligned}$$

Indem wir analog die Funktion $\tilde{g}\colon [a_2, b_2] \to \mathbb{R}$ definiert durch $\tilde{g}(t) = f(b_1, t) - f(a_1, t)$ betrachten, finden wir $(\tilde{\xi}, \tilde{\eta}) \in V_\delta$ mit

$$f(b_1, b_2) - f(a_1, b_2) - f(b_1, a_2) + f(a_1, a_2) = \tilde{g}(b_2) - \tilde{g}(a_2) = D_1 D_2 f(\tilde{\xi}, \tilde{\eta}) \cdot \delta^2$$

Setze nun $w = (\xi, \eta)$, $x = (\tilde{\xi}, \tilde{\eta})$. Dann sind $w, x \in V_\delta$ und damit $D_2 D_1 f(w) = D_1 D_2 f(x)$. Also können wir nun zwei Folgen $(w_k)_{k\in\mathbb{N}}$ und $(x_k)_{k\in\mathbb{N}}$ aus U definieren mit $\lim_{k\to\infty} w_k = a = \lim_{k\to\infty} x_k$ und $D_2 D_1 f(w_k) = D_1 D_2 f(x_k)$. Da $D_2 D_1 f$ und $D_1 D_2 f$ stetig sind, folgt dann

$$D_2 D_1 f(a) = \lim_{k\to\infty} D_2 D_1 f(w_k) = \lim_{k\to\infty} D_1 D_2 f(x_k) = D_1 D_2 f(a) \qquad \square$$

Definition. Seien $U \subseteq R^n$ offen und $f\colon U \to \mathbb{R}$ zweimal differenzierbar. Für $a \in U$ setzen wir dann

$$\operatorname{Hess} f(a) = (D_i D_j f(a))_{1\le i,j\le n}$$

$\operatorname{Hess} f(a)$ ist die *Hessesche Matrix* von f im Punkt a. $\operatorname{Hess} f(a)$ ist also eine $n \times n$-Matrix. Ist f zweimal stetig differenzierbar, so ist nach Satz 12 $\operatorname{Hess} f(a)$ symmetrisch, d.h. falls $\operatorname{Hess} f(a) = (a_{ij})_{1\le i,j\le n}$, so $a_{ij} = a_{ji}$ für alle $1 \le i, j \le n$.

Bemerkung. Es ist einfach $\operatorname{Hess} f(a) = f''(a)$.

Satz 13 (Taylorformel vom Grad 2)**.** *Seien $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ zweimal stetig differenzierbar. Weiterhin sei $a \in U$ und sei $A = \operatorname{Hess} f(a)$. Dann existiert eine in a stetige Funktion $r\colon U \to \mathbb{R}$ mit $r(a) = 0$ und $f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}(x-a)^T A(x-a) + r(x)\|x-a\|^2$*

Beweis. Definiere $r\colon U \to \mathbb{R}$ durch

$$r(x) = \begin{cases} \dfrac{f(x) - f(a) - f'(a)(x-a) - \frac{1}{2}(x-a)^T A(x-a)}{\|x-a\|^2} & \text{falls } x \ne a \\ 0 & \text{falls } x = a \end{cases}$$

Wir müssen zeigen, dass r in a stetig ist. Hierzu zeigen wir: Sei $\delta > 0$ mit $U_\delta(a) \subseteq U$. Dann existiert für alle $x = (x_1, \ldots, x_n) \in U_\delta(a) \setminus \{a\}$ ein $\xi \in [a, x]$ mit

$$r(x) = \frac{1}{2} \sum_{i,j=1}^{n} (D_i D_j f(\xi) - D_i D_j f(x)) \cdot \frac{x_i - a_i}{\|x-a\|} \cdot \frac{x_j - a_j}{\|x-a\|}$$

Hieraus folgt das Gewünschte. Denn sei $\varepsilon > 0$. Wegen $D_i D_j f$ stetig für $1 \le i, j \le n$ existiert dann $\delta > 0$ mit $U_\delta(a) \subseteq U$ und $|D_i D_j f(w) - D_i D_j f(a)| < \varepsilon/n^2$ für alle $1 \le i, j \le n$ und alle w mit $\|w - a\| < \delta$. Sei nun $x = (x_1, \ldots, x_n) \in U_\delta(a) \setminus \{a\}$. Wähle hierzu $\xi \in [a, x]$ wie oben. Dann ist $\xi \in U_\delta(a)$, also $\|\xi - a\| < \delta$. Somit gilt

$$\begin{aligned} |r(x)| &\le \left| \sum_{i,j=1}^{n} (D_i D_j f(\xi) - D_i D_j f(a)) \cdot \frac{x_i - a_i}{\|x-a\|} \cdot \frac{x_j - a_j}{\|x-a\|} \right| \le \\ &\le \sum_{i,j=1}^{n} |D_i D_j f(\xi) - D_i D_j f(a)| \cdot \frac{|x_i - a_i|}{\|x-a\|} \cdot \frac{|x_j - a_j|}{\|x-a\|} < n^2 \cdot \frac{\varepsilon}{n^2} = \varepsilon \end{aligned}$$

Jetzt bleibt noch die obige Behauptung zu zeigen. Sei hierzu $\delta > 0$ mit $U_\delta(a) \subseteq U$ und $x = (x_1, \dots, x_n) \in U_\delta(a) \setminus \{a\}$. Setze $z = x - a$, und sei $z = (z_1, \dots, z_n)$. Sei $V = \{t \in \mathbb{R} \colon a + tz \in U\}$ und definiere $g \colon V \to \mathbb{R}$ durch $g(t) = f(a + tz)$. Dann ist V offen und nach Kettenregel ist

$$g'(t) = f'(a + tz) \cdot z = \sum_{j=1}^{n} D_j f(a + tz) \cdot z_j$$

Somit ist auch g' differenzierbar und es gilt

$$g''(t) = \sum_{j=1}^{n} ((D_j f)'(a + tz) \cdot z) \cdot z_j = \sum_{i=1}^{n} \sum_{j=1}^{n} D_i D_j f(a + tz) z_i z_j$$

Nach der Taylorformel aus §9 (Lagrangesches Restglied) angewandt auf g im Punkt 0 existiert ein $t_0 \in [0, 1]$ mit $g(1) = g(0) + g'(0) + \frac{1}{2} g''(t_0) = g(0) + g'(0) + \frac{1}{2} g''(0) + \frac{1}{2}(g''(t_0) - g''(0))$. Setze nun $\xi = a + t_0 z \in [a, x]$. Dann gilt

$$f(x) = g(1) = f(a) + f'(a)z + \frac{1}{2} z^T A z + \frac{1}{2} \sum_{i,j=1}^{n} (D_i D_j f(\xi) - D_i D_j f(a)) z_i z_j$$

also

$$r(x) = \frac{1}{2} \sum_{i,j=1}^{n} (D_i D_j f(\xi) - D_i D_j f(a)) \frac{z_i}{\|z\|} \frac{z_j}{\|z\|}$$

□

Definition. Seien $U \subseteq \mathbb{R}^n$ offen, $f \colon U \to \mathbb{R}$ und $a \in U$.

- f hat in a ein *lokales Maximum*, wenn eine Umgebung W von a existiert mit $f(x) \leq f(a)$ für alle $x \in W \cap U$.
- f hat in a ein *lokales Minimum*, wenn eine Umgebung W von a existiert mit $f(x) \geq f(a)$ für alle $x \in W \cap U$.
- f hat in a ein *lokales Extremum*, wenn f in a ein lokales Maximum oder ein lokales Minimum hat.

Satz 14. *Seien $U \subseteq \mathbb{R}^n$ offen und $f \colon U \to \mathbb{R}$ partiell differenzierbar in $a \in U$. Besitzt f in a ein lokales Extremum so gilt $\nabla f(a) = 0$.*

Beweis. Sei o.E $a \in U$ lokales Maximum von f (sonst betrachte $-f$). Dann existiert $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq U$ und $f(x) \leq f(a)$ für alle $x \in U_\varepsilon(a)$. Sei nun $1 \leq j \leq n$. Wir müssen zeigen, dass $D_j f(a) = 0$. Nun ist für alle $t \in (-\varepsilon, \varepsilon)$ $a + te_j \in U_\varepsilon(a) \subseteq U$. Wir können also $g_j \colon (-\varepsilon, \varepsilon) \to \mathbb{R}$ definieren durch $g_j(t) = f(a + te_j)$. g_j besitzt in 0 ein lokales Maximum und ist in 0 differenzierbar, da $g'(0) = D_j f(a)$. Also ist $0 = g_j'(0) = D_j f(a)$. □

Definition. Sei A eine symmetrische $n \times n$-Matrix.

- A ist *positiv definit*, falls $x^T A x > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.

- A ist *negativ definit*, falls $x^TAx < 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.
- A ist *indefinit*, falls es $x, y \in \mathbb{R}^n$ gibt mit $x^TAx > 0$ und $y^TAy < 0$.

Satz 15. *Seien $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ zweimal stetig differenzierbar. Weiterhin sei $a \in U$ mit $f'(a) = 0$. Dann gilt*

(a) Ist $\operatorname{Hess} f(a)$ *positiv definit, so hat f in a ein lokales Minimum*

(b) Ist $\operatorname{Hess} f(a)$ *negativ definit, so hat f in a ein lokales Maximum*

(c) Ist $\operatorname{Hess} f(a)$ *indefinit, so hat f in a kein lokales Extremum*

Beweis. Sei $A = \operatorname{Hess} f(a)$. Nach der Taylorformel existiert wegen $f'(a) = 0$ eine in a stetige Funktion $r\colon U \to \mathbb{R}$ mit $r(a) = 0$ und $f(x) = f(a) + \frac{1}{2}(x-a)^T A(x-a) + r(x)\|x-a\|^2$ für alle $x \in U$.

zu (a): Sei A positiv definit. Setze $K = \{x \in \mathbb{R}^n \colon \|x\| = 1\}$. Da K kompakt und $x \mapsto x^TAx$ stetig ist, existiert ein $z \in K$ mit $z^TAz = \min\{x^TAx \colon x \in K\} =: \alpha$. Da A positiv definit ist, ist $\alpha > 0$. Wegen $\lim_{x\to a} r(x) = 0$ existiert also ein $\delta > 0$ mit $U_\delta(a) \subseteq U$ und $|r(x)| < \alpha/2$ für alle $x \in U_\delta(a)$. Nun gilt aber für alle $y \in \mathbb{R}^n$ $y^TAy \geq \alpha\|y\|^2$, denn für $y \neq 0$ ist $\frac{y}{\|y\|} \in K$, also $\alpha \leq \frac{y^T}{\|y\|} A \frac{y}{\|y\|}$. Somit erhalten wir für alle $x \in U_\delta(a)$

$$
\begin{aligned}
f(x) &= f(a) + \tfrac{1}{2}(x-a)^T A(x-a) + r(x)\|x-a\|^2 \\
&\leq f(a) + \tfrac{\alpha}{2}\|x-a\|^2 + r(x)\|x-a\|^2 \\
&= f(a) + (\tfrac{\alpha}{2} + r(x))\|x-a\|^2 \geq f(a)
\end{aligned}
$$

Also hat f in a ein lokales Minimum.

zu (b): Wegen $\operatorname{Hess}(-f)(a) = -\operatorname{Hess} f(a)$ und ($-A$ positiv definit gdw. A negativ definit) können wir einfach (a) auf $-f$ anwenden.

zu (c): Sei A indefinit. Wähle also $u_0, u_1 \in \mathbb{R}^n$ mit $u_0^TAu_0 > 0$ und $u_1^TAu_1 < 0$. Sei W Umgebung von a. Wir müssen zeigen, dass es $v_0, v_1 \in U \cap W$ gibt mit $f(v_0) < f(a) < f(v_1)$. Setze hierzu $\alpha := u_0^TAu_0$. Dann existiert $t > 0$ mit $a + tu_0 \in U \cap W$ und $|r(a+tu_0)| < \alpha/(2\|u_0\|^2)$. Somit gilt

$$
\begin{aligned}
f(a+tu_0) &= f(a) + \frac{1}{2} tu_0 Atu_0 + r(a+tu_0)\|tu_0\|^2 = \\
&= f(a) + \frac{1}{2} t^2\alpha + r(a+tu_0)t^2\|u_0\|^2 = \\
&= f(a) + t^2\|u_0\|^2 \left(\frac{\alpha}{2\|u_0\|^2} + r(a+tu_0) \right) > f(a)
\end{aligned}
$$

Analog findet man $v_0 \in U \cap W$ mit $f(v_0) < f(a)$, indem man u_1 betrachtet. □

Bemerkung. Sei A eine symmetrische $n \times n$-Matrix. Ist $n = 1$ und $A = (a)$, so gilt natürlich:

(1) A positiv definit gdw. $\alpha > 0$

(2) A negativ definit gdw. $\alpha < 0$

(3) A ist niemals indefinit

Sei nun $n = 2$ und etwa $A = \left(\begin{smallmatrix} \alpha & \beta \\ \beta & \gamma \end{smallmatrix}\right)$. Setze $\Delta = \det A = \alpha\gamma - \beta^2$. Dann gilt:

(1) Ist $\Delta > 0$ und $\alpha > 0$, so ist A positiv definit

(2) Ist $\Delta > 0$ und $\alpha < 0$, so ist A negativ definit

(3) Ist $\Delta < 0$, so ist A indefinit

Beweis. Für $w = (x, y) \in \mathbb{R}^2$ gilt

$$\begin{aligned} \alpha w^T A w &= \alpha(\alpha x^2 + 2\beta xy + \gamma y^2) = \\ &= \alpha^2 x^2 + 2\alpha\beta xy + \alpha\gamma y^2 = \\ &= (\alpha^2 x^2 + 2\alpha\beta xy + \beta^2 y^2) + (\alpha\gamma y^2 - \beta^2 y^2) = \\ &= (\alpha x + \beta y)^2 + \Delta y^2 \geq 0 \end{aligned}$$

Hieraus folgt (1) und (2). Sei nun $\Delta < 0$. Ist $\alpha = 0$, so $\beta \neq 0$ und $(x, 1)^T A(x, 1) = 2\beta x + \gamma$. Also A indefinit. Ist $\alpha \neq 0$, so gilt für $x = -\beta/\alpha$

$$\alpha(x, 1)^T A(x, 1) = \Delta < 0 \quad \text{und} \quad \alpha(1, 0)^T A(1, 0) = \alpha^2 > 0$$

Also ist A indefinit. $\square$

Beispiel. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x, y) = x^3 + xy^2 + x^2 - y^2$. Es ist

$$f'(x, y) = (3x^2 + y^2 + 2x, 2(x - 1)y)$$

Also sind die Nullstellen von f' $(0, 0)$ und $(-2/3, 0)$. Weiterhin ist

$$\operatorname{Hess} f(x, y) = \begin{pmatrix} 6x + 2 & 2y \\ 2y & 2x - 2 \end{pmatrix}$$

Es gilt $\det \operatorname{Hess} f(0, 0) = -4 < 0$ und $\det \operatorname{Hess} f(-2/3, 0) = 20/3 > 0$ und die erste Komponente von $\operatorname{Hess} f(-\frac{2}{3}, 0)$: $6(-2/3) + 2 = -2 < 0$, also ist $(-2/3, 0)$ ein lokales Maximum.

13 Umkehrsatz, Implizite Funktionen

Sei $\mathbb{R}^{m \times n}$ der Vektorraum der $m \times n$-Matrizen. Wir können natürlich $\mathbb{R}^{m \times n}$ kanonisch mit $\mathbb{R}^{m \cdot n}$ identifizieren, aber wir nehmen auf $\mathbb{R}^{m \times n}$ die Abbildungsnorm, die definiert war durch

$$\|A\| = \sup\{\|Ax\| \colon x \in \mathbb{R}^n \wedge \|x\| \leq 1\}$$

Nun hatten wir viele Sätze nur für den $\mathbb{R}^n$ mit der euklidischen Norm bewiesen. Es gilt aber

Satz 1 (Normäquivalenzsatz). *Sei $\|\cdot\|$ die euklidische Norm auf $\mathbb{R}^n$ und $\|\cdot\|^*$ eine weitere Norm auf $\mathbb{R}^n$. Dann existieren $c_0, c_1 > 0$ mit $c_0\|x\| \leq \|x\|^* \leq c_1\|x\|$ für alle $x \in \mathbb{R}^n$.*

Beweis. Sei $e_1, \dots, e_n$ die kanonische Basis des $\mathbb{R}^n$. Für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ gilt dann

$$\begin{aligned}\|x\|^* &= \left\|\sum_{k=1}^n x_k e_k\right\|^* \leq \sum_{k=1}^n \|x_k e_k\|^* = \sum_{k=1}^n |x_k|\|e_k\|^* = \\ &= \langle(|x_1|, \dots, |x_n|), (\|e_1\|^*, \dots, \|e_n\|^*)\rangle \leq \\ &\leq \|(|x_1, \dots, |x_n|)\| \cdot \|(\|e_1\|^*, \dots, \|e_n\|^*)\| = \|(\|e_1\|^*, \dots, \|e_n\|^*)\| \cdot \|x\|\end{aligned}$$

Wir können also $c_1 = \|(\|e_1\|^*, \dots, \|e_n\|^*)\|$ wählen. Somit gilt also auch $\|x-y\|^* \leq c_1\|x-y\|$ für alle $x, y \in \mathbb{R}^n$. Also ist $\|\cdot\|^*$ stetig (bzgl. $\|\cdot\|$). Sei nun $S = \{x \in \mathbb{R}^n \colon \|x\| = 1\}$. Wegen S kompakt und $\|\cdot\|^*$ stetig nimmt also $\|\cdot\|^*$ auf S ihr Minimum. Setze also $c_0 = \min\{\|x\|^* \colon x \in S\}$. Dann ist $c_0 > 0$. Sei nun $x \in \mathbb{R}^n, x \neq 0$. Dann ist $x/\|x\| \in S$. Somit gilt $c_0 \leq \|x/\|x\|\|^* = \|x\|^*/\|x\|$ und daher $c_0\|x\| \leq \|x\|^*$. Für $x = 0$ gilt dies natürlich auch. □

Damit hängt aber z.B. der Konvergenzbegriff auf $\mathbb{R}^n$ nicht von der speziell gewählten euklidischen Norm auf $\mathbb{R}^n$ ab. Ist $\|\cdot\|^*$ eine weitere Norm auf $\mathbb{R}^n$, so gilt für eine Folge $(x_k)_{k\in\mathbb{N}}$: $(\|x_k\|)_{k\in\mathbb{N}}$ ist Nullfolge genau dann wenn $(\|x_k\|^*)_{k\in\mathbb{N}}$ eine Nullfolge ist.

Definition. Sei $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Setze dann $\|x\|_\infty = \max\{|x_i| \colon 1 \leq i \leq n\}$. $\|\cdot\|_\infty$ ist die *Maximumsnorm.*

Bemerkung. Sei $U \in \mathbb{R}^n$ offen und $f \colon U \to \mathbb{R}^m$ differenzierbar. Dann ist f stetig differenzierbar genau dann, wenn $f' \colon U \to \mathbb{R}^{m\times n}$ stetig ist. Dies erhält man durch Anwendung des Äquivalenzsatzes auf die Abbildungsnorm.

Bemerkung. Seien $A \in \mathbb{R}^{k\times m}$ und $B \in \mathbb{R}^{m\times n}$. Dann gilt $\|AB\| \leq \|A\|\|B\|$.

Beweis. Sei $x \in \mathbb{R}^n$. Dann gilt $\|ABx\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\|$. Also

$$\begin{aligned}\|AB\| &= \sup\{\|ABx\| \colon x \in \mathbb{R}^n \wedge \|x\| \leq 1\} \\ &\leq \sup\{\|A\|\|B\|\|x\| \colon x \in \mathbb{R}^n \wedge \|x\| \leq 1\} \\ &= \|A\|\|B\|\end{aligned}$$

□

Definition. Für $n \in \mathbb{N}$ sei $\mathrm{GL}(n) = \{A \in \mathbb{R}^{n\times n} \colon A \text{ invertierbar}\}$.

Lemma 2. $\mathrm{GL}(n)$ *ist offen und die Abbildung* $\mathrm{Inv} \colon \mathrm{GL}(n) \to \mathrm{GL}(n)$ *definiert durch* $\mathrm{Inv}(A) = A^{-1}$ *ist stetig.*

Beweis. Die Determinantenabbildung $\det \colon \mathbb{R}^{n\times n} \to \mathbb{R}$ ist stetig und $\mathrm{GL}(n) = \{A \in \mathbb{R}^{n\times n} \colon \det(A) \neq 0\} = \det^{-1}(\mathbb{R} \setminus \{0\})$. Also ist $\mathrm{GL}(n)$ offen. Nach der Cramerschen Regel sind die Komponentenfunktionen von Inv rationale Funktionen, also stetig. □

Wir untersuchen nun die Differenzierbarkeit der Umkehrfunktion.

Satz 3. *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen und sei $f\colon U \to V$ bijektiv. Weiterhin sei $a \in U$ und seien f in a differenzierbar und f^{-1} in $f(a)$ differenzierbar. Dann gilt $m = n$, $Df(a)$ ist invertierbar und es ist $Df^{-1}(f(a)) = (Df(a))^{-1}$.*

Beweis. Setze $A = Df(a)$ und $B = Df^{-1}(f(a))$. Weiterhin seien id_U die identische Abbildung auf U und id_V die identische Abbildung auf V. Dann gilt natürlich $D\,\mathrm{id}_U(x) = E_n$, die n-te Einheitsmatrix, für alle $x \in U$ und $D\,\mathrm{id}_V(y) = E_m$, die m-te Einheitsmatrix, für alle $y \in V$. Wegen $f^{-1} \circ f = \mathrm{id}_U$ und $f \circ f^{-1} = \mathrm{id}_V$ gilt somit nach Kettenregel $BA = E_n$ und $AB = E_m$. Hieraus folgt nach linearer Algebra die Behauptung. □

Frage. Wann ist f^{-1} differenzierbar? Das folgende Kriterium werden wir später noch wesentlich verstärken.

Lemma 4. *Seien $U, V \in \mathbb{R}^n$ offen und sei $f\colon U \to V$ bijektiv und stetig differenzierbar. Weiterhin sei f^{-1} stetig und für alle $x \in U$ sei $Df(x)$ invertierbar. Dann ist f^{-1} stetig differenzierbar.*

Beweis. Wir zeigen zuerst, dass f^{-1} in jedem Punkt $b = f(a) \in V$ differenzierbar ist. Dabei können wir o.E. $a = 0$ und $f(a) = 0$ annehmen, denn sonst betrachte die Funktion $g(x) = f(x + a) - f(a)$. Außerdem genügt es die Differenzierbarkeit von $(C \circ f)^{-1}$ in 0 zu zeigen, wobei $C = (Df(0))^{-1}$. Es ist aber $D(C \circ f)(0) = CDf(0) = E_n$, wobei E_n die n-te Einheitsmatrix ist. Also können wir auch o.E. annehmen, dass $Df(0) = E_n$. Wegen f differenzierbar in 0 existiert eine in 0 stetige Funktion $r\colon U \to \mathbb{R}^n$ mit $r(0) = 0$ und $f(x) = f(0) + E_n x + r(x)\|x\| = x + r(x)\|x\|$ für alle $x \in U$. Definiere nun $s\colon V \to \mathbb{R}^n$ durch

$$s(y) = \begin{cases} -\frac{r(x)\|x\|}{\|y\|} & \text{falls } y \neq 0,\ y = f(x) \\ 0 & \text{falls } y = 0 \end{cases}$$

Dann gilt für alle $y = f(x) \in V$ mit $y \neq 0$

$$\begin{aligned} f^{-1}(y) = x &= f(x) - r(x)\|x\| = \\ &= y - \frac{r(x)\|x\|}{\|y\|}\|y\| = \\ &= y + s(y)\|y\| = \\ &= f^{-1}(0) + E_n y + s(y)\|y\| \end{aligned}$$

Natürlich gilt dies auch für $y = 0$. Also genügt es zu zeigen, dass s in 0 stetig ist. Sei also $\varepsilon > 0$, wobei o.E. $\varepsilon \leq 1$. Da r stetig ist in 0 und $r(0) = 0$ gibt es ein $\eta > 0$ mit $\|r(x)\| < \varepsilon/2$ für alle x mit $\|x\| < y$. Da f^{-1} in 0 stetig ist und $f^{-1}(0) = 0$ sowie V offen gibt es ein $\delta > 0$ mit $\|f^{-1}(y)\| < \eta$ für alle y mit $\|y\| < \delta$. Dann gilt für alle $y = f(x)$ mit $\|y\| < \delta$ und $y \neq 0$

$$\|x\| = \|y - r(x)\|x\|\| \leq \|y\| + \|r(x)\|\|x\| \leq \|y\| + \|x\|/2$$

also $\|x\| \leq 2\|y\|$ und somit auch

$$\|s(y)\| = \frac{\|r(x)\|\|x\|}{\|y\|} \leq 2\|r(x)\| < \varepsilon$$

Somit ist f^{-1} differenzierbar. Wir zeigen nun noch, dass Df^{-1} stetig ist. Nach Satz 2 ist $Df^{-1} \circ f = \mathrm{Inv} \circ Df$, also $Df^{-1} = Inv \circ Df \circ f^{-1}$. Nach Lemma 2 ist Inv stetig. Df und f^{-1} sind auch stetig. Also ist Df^{-1} stetig. □

Wir werden später zeigen, dass man in Lemma 4 die Voraussetzungen, dass V offen ist und f^{-1} stetig ist, weglassen können.

Definition. Sei X ein metrischer Raum und $f\colon X \to X$. f ist eine *Kontraktion*, wenn ein $c < 1$ existiert mit $d(f(x), f(y)) \leq c\, d(x, y)$ für alle $x, y \in X$.

Bemerkung. Jede Kontraktion ist stetig.

Satz 5 (Banachscher Fixpunktsatz)**.** *Sei $X \neq \varnothing$ ein vollständiger metrischer Raum. Weiterhin sei $f\colon X \to X$ eine Kontraktion. Dann besitzt f genau einen Fixpunkt, d.h. ein $x \in X$ mit $f(x) = x$.*

Beweis. Sei $c < 1$ mit $d(f(x), f(y)) \leq c\, d(x, y)$ für alle $x, y \in X$. Wir zeigen zuerst die Eindeutigkeit. Seien dazu $x, y \in X$ Fixpunkte von f. Dann gilt $d(x, y) = d(f(x), f(y)) \leq c\, d(x, y)$. Wegen $c < 1$ also $d(x, y) = 0$, d.h. $x = y$. Sei für die Existenz $a \in X$ beliebig. Definiere dann eine Folge $(x_n)_{n\in\mathbb{N}}$ rekursiv durch $x_0 = a$ und $x_{n+1} = f(x_n)$. Wir sagen zuerst durch Induktion, dass für alle $n \in \mathbb{N}$ gilt

$$d(x_n, x_{n+1}) \leq c^n d(x_0, x_1)$$

Der Fall $n = 0$ ist trivial. Außerdem gilt nach Induktionsvoraussetzung

$$d(x_{n+1}, x_{n+2}) = d(f(x_n), f(x_{n+1})) \leq c\, d(x_n, x_{n+1}) \leq c^{n+1}\, d(x_0, x_1)$$

Somit gilt für $k, n \in \mathbb{N}$

$$d(x_n, x_{n+k}) \leq \sum_{j=0}^{k-1} d(x_{n+j}, x_{n+j+1}) \leq \sum_{j=0}^{k-1} c^{n+j} d(x_0, x_1) = \frac{c^n}{1-c} d(x_0, x_1)$$

Wegen $c < 1$ ist aber $(\frac{c^n}{1-c} d(x_0, x_1)_{n\in\mathbb{N}}$ eine Nullfolge. Somit ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge, denn sei $\varepsilon > 0$. Dann existiert $n \in \mathbb{N}$ mit $\frac{c^n}{1-c} d(x_0, x_1) < \varepsilon$. Also ist $d(x_n, x_{n+k}) \leq \frac{c^n}{1-c} d(x_0, x_1) < \varepsilon$ für alle $k \in \mathbb{N}$. Wegen der Vollständigkeit von X ist somit $(x_n)_{n\in\mathbb{N}}$ konvergent. Sei also $x = \lim_{n\to\infty} x_n$. Da f stetig ist, gilt $f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x$. □

Um den folgenden Satz zu motivieren, betrachten wir folgende Situation im eindimensionalen Fall. Seien $U \subseteq \mathbb{R}$ offen und $f\colon U \to \mathbb{R}$ stetig differenzierbar. Weiterhin sei $a \in U$ mit $f'(a) \neq 0$. Sei etwa $f'(a) > 0$. Da f' stetig ist, existiert ein offenes Intervall $I \subseteq U$ mit $a \in I$ und $f'(x) > 0$ für alle $x \in I$. Somit ist $h = f|_I$ streng monoton wachsend. Also ist $W := f(I)$ offen und $h^{-1}\colon W \to \mathbb{R}$ ist stetig differenzierbar, denn $(h^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$.

Wir verallgemeinern dies auf den mehrdimensionalen Fall.

Satz 6 (Lokaler Umkehrsatz)**.** *Seien $U \subseteq \mathbb{R}^n$ offen und $f: U \to \mathbb{R}^n$ stetig differenzierbar. Weiterhin sei $a \in U$ mit $Df(a)$ invertierbar. Dann gibt es eine offene Umgebung $V \subseteq U$ von a derart, dass $W := f(V)$ offen ist, die Einschränkung $h = f|_V$ injektiv ist und $h^{-1}: W \to \mathbb{R}^n$ stetig differenzierbar ist.*

Beweis. Sei $C = Df(a)^{-1}$. Es genügt den Satz für $C \circ f$ zu beweisen. Nun ist aber $D(C \circ f)(a) = C\,Df(a) = E_n$. Weiterhin können wir ohne Einschränkung annehmen, dass $a = 0 = f(a)$, denn sonst betrachte die Funktion $g(x) = f(x+a) - f(a)$. Somit nehmen, wir also insgesamt an, dass $a = f(a) = 0$ und $Df(0) = E_n$.

Wegen Df stetig in 0 und $Df(0) = E_n$ sowie U offen existiert eine $\varepsilon > 0$ mit $K = \{x \in \mathbb{R}^n : \|x\| \leq \varepsilon\} \subseteq U$ und $\|E_n - Df(x)\| \leq 1/2$ für alle $x \in K$. Beachte, dass K abgeschlossen ist.

Für $y \in \mathbb{R}^n$ definiere $g_y: U \to \mathbb{R}^n, x \mapsto y + x - f(x)$. Es gilt also

$$g_y(x) = x \iff f(x) = y$$

Außerdem ist $g_y(0) = y - f(0) = y$. Weiterhin ist $Dg_y = E_n - Df$. Mit dem Schrankensatz erhalten wir

(1) $\|g_y(x_2) - g_y(x_1)\| \leq \dfrac{1}{2}\|x_2 - x_1\|$ für alle $x_1, x_2 \in K$.

denn sind $x_1, x_2 \in K$, so ist

$$\|g_y(x_1) - g_y(x_2)\| \leq \sup\{\|Dg_y(x)\| : x \in [x_1, x_2]\}\|x_2 - x_2\| \leq \frac{1}{2}\|x_2 - x_1\|$$

Somit gilt

(2) $\|g_y(x)\| < \varepsilon$ für alle x, y mit $\|x\| \leq \varepsilon$ und $\|y\| < \frac{\varepsilon}{2}$.

denn für solche x, y ist

$$\|g_y(x)\| = \|(g_y(x) - g_y(0)) + y\| \leq \|g_y(x) - g_y(0)\| + \|y\| \underset{(1)}{\leq} \frac{\|x\|}{2} + \|y\| < \varepsilon$$

Setze nun $W = U_{\varepsilon/2}(0)$. W ist offen. Für $y \in W$ gilt nach (2), dass $g_y: K \to K$. Da K abgeschlossen ist und $\mathbb{R}^n$ vollständig, ist K ein vollständiger metrischer Raum. Weiterhin ist nach (1) für $y \in W$ g_y (eingeschränkt auf K) eine Kontraktion. Somit existiert nach dem Fixpunktsatz für jedes $y \in W$ genau ein $x \in K$ mit $g(x) = x$, d.h. mit $f(x) = y$. Setze nun $V = f^{-1}[W]$. Dann ist wegen f stetig V eine offene Umgebung ovn O mit $V \subseteq K \subseteq U$ und für die Einschränkung h von f auf V ist $h: V \to W$ bijektiv. Außerdem ist W offen. Wir müssen noch zeigen, dass h^{-1} stetig differenzierbar ist. Hierzu wollen wir Lemma 4 auf h anwenden. Wir müssen also noch zeigen, dass h^{-1} stetig ist und für alle $x \in V$ $Dh(x)$ invertierbar ist.

Wir zeigen zuerst die Stetigkeit von h^{-1}. Seien $y_1, y_2 \in W$ und setze $x_1 = h^{-1}(y_1)$, $x_2 = h^{-1}(y_2)$. Dann ist $x_2 - x_1 = (g_0(x_2) - g_0(x_1)) + (f(x_2) - f(x_1))$. Also nach (1) $\|x_2 - x_1\| \leq \frac{1}{2}\|x_2 - x_1\| + \|f(x_2) - f(x_1)\|$ d.h. $\|h^{-1}(y_2) - h^{-1}(y_1)\| \leq 2\|y_2 - y_1\|$. Hieraus folgt die Stetigkeit von h^{-1} sofort.

Sei schließlich $x \in V$. Wir müssen zeigen, dass $Dh(x) = Df(x)$ invertierbar ist. Wegen $V \subseteq K$ gilt $\|E_n - Df(x)\| \leq \frac{1}{2}$. Also gilt für alle $v \in \mathbb{R}^n$ $\|v - Df(x)v\| \leq \frac{1}{2}\|v\|$ und somit $Df(x)v \neq 0$ für alle $v \neq 0$. Somit ist $Df(x)$ injektiv, d.h. invertierbar. □

Korollar. *Seien $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^n$ stetig differenzierbar und $Df(x)$ invertierbar für alle $x \in U$. Dann ist $f[U]$ offen.*

Beweis. Sei $b \in f[U]$. Wähle $a \in U$ mit $b = f(a)$. Nach Satz 6 existiert eine offene Umgebung $V \subseteq U$ von a mit $W = f[V]$ offen. Wegen $b \in W$ ist W eine offene Umgebung von b und $W = f[V] \subseteq f[U]$. Also ist $f[U]$ offen. □

Korollar. *Seien $U \subseteq \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}^n$ eine injektive stetig differenzierbare Funktion mit $Df(x)$ invertierbar für alle $x \in U$. Dann ist $f[U]$ offen und $f^{-1}\colon f[U] \to \mathbb{R}^n$ ist stetig differenzierbar.*

Beweis. Nach obigem Korollar ist $f[U]$ offen. Sei nun $b \in f[U]$ und sei $a = f^{-1}(b)$. Nach Satz 6 existiert eine offene Umgebung $W \subseteq f[U]$ von b derart, dass f^{-1} eingeschränkt auf W stetig differenzierbar ist. Hieraus folgt die Behauptung. □

Beispiel (Polarkoordinaten)**.** Sei $f\colon (0,\infty) \times (0, 2\pi) \to \mathbb{R}^2$ definieret durch $f(r,\varphi) = (r\cos\varphi, r\sin\varphi)$. f ist injektiv und stetig differenzierabr. Es ist

$$f'(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

Somit $\det f'(r,\varphi) = r\cos^2\varphi + r\sin^2\varphi = r > 0$. Also ist $f'(r,\varphi)$ invertierbar. Sei $g = f^{-1}$. Für $(x,y) = f(r,\varphi)$ gilt

$$g'(x,y) = (f'(r,\varphi))^{-1} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{1}{r}\sin\varphi & \frac{1}{r}\cos\varphi \end{pmatrix} = \begin{pmatrix} \frac{x}{\|(x,y\|} & \frac{y}{\|(x,y\|} \\ -\frac{y}{\|(x,y\|^2} & \frac{x}{\|(x,y\|^2} \end{pmatrix}$$

Sei nun $F\colon \mathbb{R}^2 \to \mathbb{R}$ und sei $F(a,b) = 0$. Wir suchen Bedingungen für die eindeutige Lösbarkeit der Gleichung $F(x,y) = 0$ in Abhängigkeit von x in einer Umgebung von (a,b).

Beispiel. $F(x,y) = x - y^2$. Es ist $F(1,1) = 0$. Definiere $g\colon (0,\infty) \to \mathbb{R}$ durch $g(x) = \sqrt{x}$ ist die Lösung in Umgebung von $(1,1)$.

Wir machen dies etwas allgemeiner und führen hierzu einige Notationen ein.

Notation. Für die Funktion $g\colon X \to Y$ sei $\mathrm{Graph}(g) = \{(x, g(x))\colon x \in X\}$ der Graph von g.

Bemerkung. Seien $(u,v) \in \mathbb{R}^k \times R^m$ und W eine Umgebung von (u,v). Dann existieren Umgebungen W_0 von u und W_1 von v mit $W_0 \times W_1 \subseteq W$.

Beweis. Für $x = (x_1,\dots,x_k) \in \mathbb{R}^k$ und $y = (y_1,\dots,y_m) \in \mathbb{R}^m$ ist $\|(x,y)\|^2 = \|x\|^2 + \|y\|^2 \le (\|x\| + \|y\|)^2$ also $\|(x,y)\| \le \|x\| + \|y\|$. Wähle also $\varepsilon > 0$ mit $U_\varepsilon((u,v)) \subseteq W$. Dann ist $U_{\varepsilon/2}(u) \times U_{\varepsilon/2}(v) \subseteq U_\varepsilon((u,v)) \subseteq W$. □

Notation. Seien nun $U_1 \subseteq \mathbb{R}^k$ und $U_2 \subseteq \mathbb{R}^m$ offen und sei $F\colon U_1 \times U_2 \to \mathbb{R}^m$ differenzierbar. Weiterhin sei $F = (F_1, \dots, F_m)$. Für $(a, b) \in U_1 \times U_2$ setze

$$\frac{\partial F}{\partial x}(a,b) = \begin{pmatrix} D_1 F_1(a,b) & \dots & D_k F_1(a,b) \\ \vdots & \ddots & \vdots \\ D_1 F_m(a,b) & \dots & D_k F_m(a,b) \end{pmatrix}$$

$$\frac{\partial F}{\partial y}(a,b) = \begin{pmatrix} D_{k+1} F_1(a,b) & \dots & D_{k+m} F_1(a,b) \\ \vdots & \ddots & \vdots \\ D_{k+1} F_m(a,b) & \dots & D_{k+m} F_m(a,b) \end{pmatrix}$$

Es gilt also

$$DF(a,b) = \left(\frac{\partial F}{\partial x}(a,b) \;\middle|\; \frac{\partial F}{\partial y}(a,b) \right)$$

Somit ist für $(u, v) \in \mathbb{R}^k \times \mathbb{R}^m$

$$DF(a,b) \cdot (u,v) = \frac{\partial F}{\partial x}(a,b) \cdot u + \frac{\partial F}{\partial y}(a,b) \cdot v$$

Satz 7 (Implizite Funktionen). *Seien $U_1 \subseteq \mathbb{R}^k$ und $U_2 \subseteq \mathbb{R}^m$ offen, und sei $F\colon U_1 \times U_2 \to \mathbb{R}^m$ stetig differenzierbar. Weiterhin sei $(a, b) \in U_1 \times U_2$ mit $F(a, b) = 0$ und $\frac{\partial F}{\partial y}(a, b)$ invertierbar. Dann gibt es offene Umgebungen $V_1 \subseteq U_1$ von a und $V_2 \subseteq U_2$ von b sowie eine stetig differenzierbare Funktion $g\colon V_1 \to V_2$ mit*

$$\operatorname{Graph}(g) = \{(x, y) \in V_1 \times V_2 \colon F(x, y) = 0\}$$

Weiterhin gilt

$$g'(a) = -\left(\frac{\partial F}{\partial y}(a,b)\right)^{-1} \frac{\partial F}{\partial x}(a,b)$$

Beweis. Definiere $H\colon U_1 \times U_2 \to \mathbb{R}^{k+m}$ durch $H(x, y) = (x, F(x, y))$. Wegen F stetig differenzierbar ist auch H stetig differenzierbar. Es ist

$$DH(a,b) = \begin{pmatrix} E_k & 0 \\ \frac{\partial F}{\partial x}(a,b) & \frac{\partial F}{\partial y}(a,b) \end{pmatrix}$$

wobei E_k die k-te Einheitsmatrix ist. Es ist $\det(DH(a, b)) \cdot \det(\frac{\partial F}{\partial y}(a, b)) = \det(\frac{\partial F}{\partial y}(a, b))$. Wegen $\frac{\partial F}{\partial y}(a, b)$ invertierbar ist also auch $DH(a, b)$ invertierbar. Wir können also auf H den lokalen Umkehrsatz in (a, b) anwenden. Somit gibt es eine offene Umgebung $V \subseteq U_1 \times U_2$ von (a, b) und $W \subseteq \mathbb{R}^{k+m}$ von $H(a, b)$ derart, dass die Einschränkung h von H auf V eine Bijektion von V auf W ist, für die h^{-1} stetig differenzierbar ist. Natürlich existiert eine Funktion $h^*\colon W \to \mathbb{R}^m$ mit $h^{-1}(u, v) = (u, h^*(u, v))$ für alle $(u, v) \in W$. Dann gilt

$$\forall (x,y) \in V \colon \Big(F(x,y) = 0 \iff H(x,y) = (x,0) \iff y = h^*(x,0) \Big) \tag{$*$}$$

Nach Bemerkung existieren offene Umgebungen W_1 von A und V_2 von b mit $W_1 \times V_2 \subseteq V$. Nun ist $h^*(a,0) = b$, da $F(a,b) = 0$, und wegen h^{-1} stetig ist die Abbildung $x \mapsto h^*(x,0)$ stetig. Also existiert eine offene Umgebung $V_1 \subseteq W_1$ von a mit $h^*(x,0) \in V_2$ für alle $x \in V_1$. Definiere nun $g\colon V_1 \to V_2$ durch $g(x) = h^*(x,0)$. Wegen h^{-1} stetig differenzierbar ist auch g stetig differenzierbar. Weiterhin gilt

$$\mathrm{Graph}(g) = \{(x,y) \in V_1 \times V_2 \colon F(x,y) = 0\}$$

„$\subseteq$“ Sei $x \in V_1$. Dann $(x, g(x)) \in V_1 \times V_2 \subseteq V$ und $g(x) = h^*(x,0)$. Also nach $(*)$ $F(x,(g(x)) = 0$.

„$\supseteq$“ Sei $(x,y) \in V_1 \times V_2$ mit $F(x,y) = 0$. Dann $(x,y) \in V$. Also nach $(*)$ $y = h^*(x,0) = g(x)$, d.h. $(x,y) = (x, g(x)) \in \mathrm{Graph}(g)$.

Nun noch zum Zusatz. Es ist $F(a, g(a)) = 0$, da $g(a) = b$, und F differenzierbar in $(a, g(a))$ und g differenzierbar in a. Also gilt nach Kettenregel für $G(x) = F(x, g(x))$, da G in der Umgebung von a konstant null ist

$$0 = DG(a) = DF(a,b)\begin{pmatrix} E_k \\ Dg(a) \end{pmatrix}$$

Nun ist $DF(a,b) = \left(\frac{\partial F}{\partial x}(a,b) \middle| \frac{\partial F}{\partial y}(a,b)\right)$. Also erhält man

$$0 = \frac{\partial F}{\partial x}(a,b) + \frac{\partial F}{\partial y}(a,b) \cdot Dg(a)$$

und daher

$$Dg(a) = -\left(\frac{\partial F}{\partial y}(a,b)\right)^{-1} \frac{\partial F}{\partial x}(a,b)$$

□

Beispiel. $F(x,y) = x^2 + y^2 - 1$. Definiere $g_0\colon (-1,1) \to \mathbb{R}$ durch $g_0(x) = \sqrt{1-x^2}$ und $g_1\colon (-1,1) \to \mathbb{R}$ durch $g_1(x) = -\sqrt{1-x^2}$.

- g_0 ist Lösung an den Stellen (a,b) mit $b > 0$.
- g_1 ist Lösung an den Stellen (a,b) mit $b < 0$.
- An den Stellen $(-1,0)$ und $(1,0)$ gibt es keine Lösung, da $\frac{\partial F}{\partial y}(-1,0) = 0 = \frac{\partial F}{\partial y}(1,0)$.

Satz 8 (Lokale Extrema mit Nebenbedingung)**.** *Seien $U \subseteq \mathbb{R}^n$ offen und sei $f\colon U \to \mathbb{R}$ stetig differenzierbar. Setze $M = \{x \in U \colon f(x) = 0\}$. Sei $a \in M$ mit $f'(a) \neq 0$. Weiterhin sei $h\colon U \to \mathbb{R}$ eine stetig differenzierbare Funktion, die in a ein lokales Maximum (bzw. Minimum) unter der Nebenbedingung $f = 0$ besitzt, d.h. es gebe eine Umgebung $V \subseteq U$ von a mit $h(a) \geq (x)$ (bzw. $h(a) \leq h(x)$) für alle $x \in V \cap M$. Dann gibt es ein $\lambda \in \mathbb{R}$ mit $h'(a) = \lambda f'(a)$. (Solch ein λ heißt* Lagrangescher Multiplikator*)*

Beweis. Wir nehmen o.E. an, dass $D_n f(a) \neq 0$. Sonst nummeriere die Koordinaten um. Sei o.E. $n \geq 2$, da sonst die Behauptung trivial ist. Sei $a = (a_1, \ldots, a_n)$ und setze $c = (a_1, \ldots, a_{n-1})$. Nach Satz 7 gibt es offene Umgebungen V von c und W von a_n mit $V \times W \subseteq U$ und eine stetig differenzierbare Funktion $g \colon V \to W$ mit $\operatorname{Graph}(g) = M \cap (V \times W)$. Außerdem gilt

$$D_i g(c) = -\frac{D_i f(a)}{D_n f(a)} \quad \text{für } 1 \leq i \leq n-1 \tag{$*$}$$

Sei nun h^* definiert durch $h^*(x) = h(x, g(x))$. Da für alle $x \in V$ $(x, g(x)) \in M$, besitzt nach Voraussetzung h^* in c ein lokales Extremum. Weil h^* differenzierbar ist, ist also $Dh^*(c) = 0$. Somit gilt nach Kettenregel

$$0 = D_i h^*(c) = D_i h(a) + D_n h(a) \cdot D_i g(c) \quad \text{für } 1 \leq i \leq n-1$$

Setze nun $\lambda = D_n h(a)/D_n f(a)$. Für $1 \leq i \leq n-1$ gilt dann

$$D_i h(a) = -D_n h(a) \cdot D_i g(c) \overset{(*)}{=} \lambda D_i f(a)$$

Diese Gleichung gilt auch für $i = n$. Also ist $h'(a) = \lambda f'(a)$. $\square$

Beispiel. Man bestimme diejenigen Punkte auf der Kugeloberfläche $S^2 = \{(x, y, z) \colon x^2 + y^2 + z^2 = 1\}$, die von dem Punkt $a = (1, 1, 1)$ den kleinsten bzw. größten Abstand haben. Definiere hierzu $f, h \colon \mathbb{R}^3 \to \mathbb{R}$ durch

$$\begin{aligned} f(x, y, z) &= x^2 + y^2 + z^2 - 1 \\ h(x, y, z) &= \|(x, y, z) - a\|^2 = (x-1)^2 + (y-1)^2 + (z-1)^2 \end{aligned}$$

Gesucht sind globale Extrema von h mit der Nebenbedingung $f = 0$. Es ist $Df(x, y, z) = (2x, 2y, 2z)$ und $Dh(x, y, z) = \big(2(x-1), 2(y-1), 2(z-1)\big)$. Ist $Dh(x, y, z) = \lambda Df(x, y, z)$ und $(x, y, z) \in S^2$, so ist $(x-1) = \lambda x$, $y - 1 = \lambda y$ und $z - 1 = \lambda z$ und $x^2 + y^2 + z^2 = 1$. Also $x = y = z = 1/(1-\lambda)$. Wegen $(x, y, z) \in S^2$ also $3/(1-\lambda)^2 = 1$, also $\lambda = 1 \pm \sqrt{3}$. Somit ist $(x, y, z) = \pm\frac{1}{\sqrt{3}}(1, 1, 1)$. Setze $a_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$ und $a_2 = -\frac{1}{\sqrt{3}}(1, 1, 1)$. Da h stetig ist und S^2 kompakt ist, existieren $b_1, b_2 \in S^2$ mit $h(b_1) \leq h(x) \leq h(b_2)$ für alle $x \in S^2$. Nun ist aber $Df(b_1) \neq 0$ und $Df(b_2) \neq 0$, da $b_1, b_2 \in S^2$. Nach Satz 7 existieren also $\lambda_1, \lambda_2 \in \mathbb{R}$ mit $Dh(b_1) = \lambda_1 Df(b_1)$ und $Dh(b_2) = \lambda_2 Df(b_2)$. Wegen $h(a_1) < h(a_2)$ ist also nach obiger Rechnung $b_1 = a_1$ und $b_2 = a_2$. Somit gilt

- $\frac{1}{\sqrt{3}}(1, 1, 1)$ ist der Punkt aus S^2 mit kleinstem Abstand von $(1, 1, 1)$
- $-\frac{1}{\sqrt{3}}(1, 1, 1)$ ist der Punkt aus S^2 mit größtem Abstand von $(1, 1, 1)$

14 Parameterabhängige Integrale

Satz 1. *Seien $a, b \in \mathbb{R}$ mit $a < b$, $D \subseteq \mathbb{R}^n$ und $\varphi, \psi\colon D \to \mathbb{R}$ stetig mit $a \leq \varphi(y) \leq \psi(y) \leq b$ für alle $y \in D$. Weiterhin sei $f\colon [a,b] \times D \to \mathbb{R}$ stetig. Dann ist die Funktion $g\colon D \to \mathbb{R}$ stetig, welche definiert ist durch*

$$g(y) = \int\limits_{\varphi(y)}^{\psi(y)} f(x,y) \,\mathrm{d}x$$

Beweis. Sei $c \in D$ und $(y_k)_{k\in\mathbb{N}}$ eine Folge in D mit $\lim_{k\to\infty} y_k = c$. Wir müssen zeigen, dass $\lim_{k\to\infty} g(y_k) = g(c)$. Sei hierzu $\varepsilon > 0$. Setze nun $Q = \{y_k : k \in \mathbb{N}\} \cup \{c\}$. Dann ist Q kompakt. Also ist auch $K = [a,b] \times Q$ kompakt. Somit ist die Einschränkung von f auf K gleichmäßig stetig. Daher gibt es ein $\delta > 0$ mit

(1) $|f(x,y) - f(x,z)| < \frac{\varepsilon}{3(b-a)}$ für alle $x \in [a,b]$ und alle $y, z \in Q$ mit $\|y - z\| < \delta$.

Weiterhin existiert M mit

(2) $|f(x,y)| < M$ für alle $(x,y) \in K$.

Wegen $\lim_{k\to\infty} y_k = c$ und φ,ψ stetig existiert $m \in \mathbb{N}$ mit für alle $k \geq m$

(3) $\|y_k - c\| < \delta$ und $|\varphi(y_k) - \varphi(c)|, |\psi(y_k) - \psi(c)| < \frac{\varepsilon}{3M}$.

Dann gilt für alle $k \geq m$

$$\begin{aligned}
|g(y_k) - g(c)| &= \left| \int_{\varphi(y_k)}^{\psi(y_k)} f(x,y_k) \,\mathrm{d}x - \int_{\varphi(c)}^{\psi(c)} f(x,c) \,\mathrm{d}x \right| \leq \\
&\leq \left| \int_{\varphi(y_k)}^{\psi(y_k)} f(x,y_k) \,\mathrm{d}x - \int_{\varphi(c)}^{\psi(c)} f(x,y_k) \,\mathrm{d}x \right| + \left| \int_{\varphi(c)}^{\psi(c)} f(x,y_k) - f(x,c) \,\mathrm{d}x \right| \leq \\
&\leq \left| \int_{\varphi(c)}^{\psi(y_k)} f(x,y_k) \,\mathrm{d}x + \int_{\varphi(y_k)}^{\psi(c)} f(x,y_k) \,\mathrm{d}x \right| + \frac{\varepsilon}{3(b-a)} |\psi(c) - \varphi(c)| \leq \\
&\leq M \cdot |\varphi(y_k) - \varphi(c)| + M \cdot |\psi(c) - \psi(y_k)| + \frac{\varepsilon}{3} < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\end{aligned}$$

□

Satz 2. *Seien $a, b \in \mathbb{R}$ mit $a < b$, $U \subseteq \mathbb{R}^n$ offen und $f\colon [a,b] \times U \to \mathbb{R}, (x, y_1, \ldots, y_n) \mapsto f(x, y_1, \ldots, y_n)$ stetig. Ist f bezüglich y_i stetig partiell differenzierbar, so ist auch die Funktion $g\colon U \to \mathbb{R}$ definiert durch $g(y) = \int_a^b f(x,y) \,\mathrm{d}x$ bezüglich y_i stetig partiell differenzierbar und es gilt*

$$\frac{\partial g}{\partial y_i}(y) = \int_a^b \frac{\partial f}{\partial y_i}(x,y) \,\mathrm{d}x$$

Beweis. o.E. sei $n = 1$. Sei $c \in U$ und $(c_k)_{k\in\mathbb{N}}$ eine Folge aus $U \setminus \{c\}$, die gegen c konvergiert. Wir müssen zeigen

$$\lim_{k\to 0} \frac{g(c_k) - g(c)}{c_k - c} = \int_a^b D_2 f(x,c) \,\mathrm{d}x$$

Sei also $\varepsilon > 0$. Setze $f_k(x) = \big(f(x,c_k) - f(x,c)\big)/\big(c_k - c\big)$. Dann gilt

$$\frac{g(c_k)-g(c)}{c_k-c} - \int_a^b D_2 f(x,c)\,\mathrm{d}x = \int_a^b \big(f_k(x) - D_2 f(x,c)\big)\,\mathrm{d}x \tag{1}$$

Wähle $\rho > 0$ mit $Q = [c-\rho, c+\rho] \subseteq U$. Nach dem Mittelwertsatz gilt

$$\text{für alle } k \text{ und alle } x \in [a,b] \text{ existiert } y \in [c_k, c] \text{ mit } f_k(x) = D_2 f(x,y) \tag{2}$$

Setze $K = [a,b] \times Q$. K ist kompakt und daher $D_2 f$ eingeschränkt auf K gleichmäßig stetig. Somit gibt es ein $0 < \delta \leq \rho$ mit

$$|D_2 f(x,y) - D_2 f(x,c)| < \frac{\varepsilon}{b-a} \quad \text{für alle } x \in [a,b] \text{ und alle } y \text{ mit } |y-c| < \delta \tag{3}$$

Wegen $\lim_{k\to\infty} c_k = c$ existiert ein m mit

$$|c_k - c| < \delta \quad \text{für alle } k \geq m \tag{4}$$

Für $k \geq m$ gilt dann

$$\begin{aligned}
&\left|\frac{g(c_k)-g(c)}{c_k-c} - \int_a^b D_2 f(x,c)\right| \leq \\
&\overset{(1)}{\leq} \sup\{|f_k(x) - D_2 f(x,c)|\colon x \in [a,b]\} \cdot (b-a) \leq \\
&\overset{(2)}{\leq} \sup\{|D_2 f(x,y) - D_2 f(x,c)|\colon x \in [a,b] \wedge y \in [c_k,c]\} \cdot (b-a) \leq \\
&\overset{(3,4)}{\leq} \frac{\varepsilon}{b-a} \cdot (b-a) = \varepsilon
\end{aligned}$$

□

15 Approximationssätze

Sei X ein kompakter metrischer Raum. Sei dann $C(X)$ die Menge aller stetigen Funktionen von X nach $\mathbb{R}$. Für $f,g \in C(X)$ und $\alpha \in \mathbb{R}$ definiere $f+g$, αf durch

$$\begin{gathered}
(f+g)(x) = f(x) + g(x) \\
(\alpha f)(x) = \alpha f(x)
\end{gathered}$$

für alle $x \in X$. Mit dieser Operation ist $C(X)$ ein $\mathbb{R}$-Vektorraum. Wir können auf $C(X)$ eine Norm definieren durch (für $f \in C(X)$)

$$\|f\| = \sup\{|f(x)|\colon x \in X\}$$

Weil X kompakt ist, ist auch $f[X]$ kompakt und daher beschränkt. Also $\|f\| < \infty$.(Wir benutzen die obige Notation auch für beliebige Funktionen von X nach $\mathbb{R}$.) Man rechnet leicht nach, dass $\|\cdot\|$ eine Norm ist (die Supremumsnorm). Somit ist $C(X)$ mit $\|\cdot\|$ ein normierter Vektorraum. Insbesondere ist es also ein metrischer Raum mit der Metrik $d(f,g) = \|f-g\|$.

Satz 1. *$C(X)$ ist vollständig.*

Beweis. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $C(X)$. Da ja für alle $m, n \in \mathbb{N}$ und $x \in X$ $|f_m(x) - f_n(x)| \leq \|f_m - f_n\|$ gilt, ist für jedes $x \in X$ $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$ und daher konvergent. Sei für $x \in X$ $f(x) = \lim_{n\to\infty} f_n(x)$. Wir zeigen, dass $f\colon X \to \mathbb{R}$ stetig ist und $(f_n)_{n\in\mathbb{N}}$ gegen f konvergiert (in $C(X)$). Hierzu zeigen wir

$$\text{für alle } \varepsilon > 0 \text{ existiert ein } m \in \mathbb{N} \text{ mit } \|f_n - f\| \leq 2\varepsilon \text{ für alle } n \geq m \tag{$*$}$$

Dann konvergiert also f_n „gleichmäßig“ gegen f. Wie in §9 folgt dann, dass f stetig ist, da alle f_n stetig sind. Somit konvergiert $(f_n)_{n\in\mathbb{N}}$ gegen f (in $C(X)$).

Nun zum Beweis von $(*)$. Sei $m \in \mathbb{N}$ mit $\|f_n - f_k\| < \varepsilon$ für alle $k, n \geq m$. Sei nun $n \geq m$ und $x \in X$. Wähle ein $k \geq m$ mit $|f_k(x) - f(x)| < \varepsilon$. Dann ist $|f_n(x) - f(x)| \leq |f_k(x) - f(x)| + |f_n(x) - f_k(x)| < \varepsilon + \|f_n - f_k\| < \varepsilon + \varepsilon = 2\varepsilon$. Also $\|f_n - f\| \leq 2\varepsilon$. □

Definition. Sei Y ein metrischer Raum und $A \subseteq Y$. A ist *dicht* in Y, wenn für alle $y \in Y$ und alle $\varepsilon > 0$ ein $z \in A$ existiert mit $d(y, z) < \varepsilon$.

Bemerkung. Seien A, Y wie oben. Dann gilt: A ist dicht in Y genau dann, wenn für alle $y \in Y$ eine Folge $(a_n)_{n\in\mathbb{N}}$ aus A existiert, die gegen y konvergiert.

Sei weiterhin X ein kompakter metrischer Raum. Wir interessieren uns für dichte Teilmengen von $C(X)$. Für $f, g \in C(X)$ definiere $fg\colon X \to \mathbb{R}$ durch $(fg)(x) = f(x)g(x)$ für $x \in X$. Also $fg \in C(X)$.

Definition. Sei $A \subseteq C(X)$. A ist eine *Algebra*, wenn $A \neq \varnothing$ und für alle $f, g \in A$ und $c \in \mathbb{R}$ gilt, dass $f + g, fg, cf \in A$.

Beispiel. Sei $X = [a, b]$ ein kompaktes Intervall. Dann ist die Menge aller Polynomfunktionen auf X eine Algebra.

Definition. Sei $A \subseteq C(X)$. A *trennt die Punkte* von X, wenn für alle $x, y \in X$ mit $x \neq y$ ein $f \in A$ mit $f(x) \neq f(y)$ existiert.

Bemerkung. Ist A wie im obigen Beispiel, so trennt A die Punkte von X, denn wir können einfach als f die Identität wählen.

Bemerkung. Sei $A \subseteq C(X)$ eine Algebra. Setze

$$\overline{A} = \{f \in C(X)\colon \text{es existiert eine Folge } (f_n)_{n\in\mathbb{N}} \text{ aus } A\text{, die gegen } f \text{ konvergiert}\}$$

Dann ist $\overline{A}$ eine Algebra.

Beweis. Um z.B. zu zeigen, dass $f + g \in \overline{A}$ für $f, g \in \overline{A}$, benutze die Tatsache, dass wenn $(f_n)_{n\in\mathbb{N}}$ gegen f und $(g_n)_{n\in\mathbb{N}}$ gegen g konvergiert, folgt, dass $(f_n + g_n)$ gegen $f + g$ konvergiert. □

Lemma 2. *Sei $X = [a, b]$ ein kompaktes Intervall. Definiere $f\colon X \to \mathbb{R}$ durch $f(x) = |x|$. Dann existiert eine Folge von Polynomfunktionenen auf X, die gleichmäßig gegen f konvergiert.*

Beweis. Der Beweis folgt am Ende der Vorlesung. □

Definition. Seien $f, g\colon X \to \mathbb{R}$. Definiere die Funktionen $\max(f, g)$ und $\min(f, g)$ von X nach $\mathbb{R}$ durch (für $x \in \mathbb{R}$)

$$\max(f, g)(x) = \max(f(x), g(x))$$
$$\min(f, g)(x) = \min(f(x), g(x))$$

(entsprechend für endlich viele Funktionen)

Bemerkung. Es gelten

$$\max(f, g) = \frac{f+g}{2} + \frac{|f-g|}{2}$$
$$\min(f, g) = \frac{f+g}{2} - \frac{|f-g|}{2}$$

Beweis. Nur die erste Gleichung. Aus Symmetriegründen gilt o.E. $\max(f(x), g(x)) = f(x)$. Aber $f(x) = \frac{1}{2}(f(x) + g(x)) + \frac{1}{2}(f(x) - g(x))$. □

Satz 3 (Satz von Stone-Weierstraß)**.** *Sei X ein kompakter metrischer Raum, und sei $A \subseteq C(X)$ eine Algebra, welche die konstanten Funktionen enthält und die Punkte von X trennt. Dann ist A dicht in $C(X)$.*

Beweis. Wir setzen zuerst voraus, dass für A zusätzlich gilt:

$$\forall f, g \in A\colon \max(f, g), \min(f, g) \in A \tag{$*$}$$

Nun gilt

$$\forall x, y \in X, \alpha, \beta \in \mathbb{R}\colon x \neq y \Longrightarrow \exists h \in A\colon h(x) = \alpha \wedge h(y) = \beta \tag{$+$}$$

Denn seien x, y, α, β wie in (+). Da A die Punkte von X trennt, gibt es ein $g \in A$ mit $g(x) \neq g(y)$. Definiere nun $h\colon X \to \mathbb{R}$ durch

$$h(z) = \alpha + (\beta - \alpha)\frac{g(z) - g(x)}{g(y) - g(x)}$$

Da A eine Algebra ist und alle konstanten Funktionen enthält, ist $h \in A$. Außerdem $h(x) = \alpha$, $h(y) = \beta$.
Sei nun $f \in C(X)$ und $\varepsilon > 0$. Wir suchen ein $g \in A$ mit $f(x) - \varepsilon < g(x) < f(x) + \varepsilon$ für alle $x \in X$. Solch ein g finden wir in zwei Schritten. Gemäß (+) wähle für je zwei $x, y \in X$ ein $h_{x,y} \in A$ mit $h_{x,y}(x) = f(x)$ und $h_{x,y}(y) = f(y)$. (für $x = y$ wähle die konstante Funktion $f(x)$). Sei nun $x \in X$. Für $y \in X$ ist $(h_{x,y} - f)(y) = 0$ und $h_{x,y} - f$ ist stetig. Also existiert eine offene Umgebung U_y von y mit

$$\forall z \in U_y\colon h_{x,y} < f(z) + \varepsilon \tag{Δ}$$

Setze $\mathcal{U} = \{U_y\colon y \in X\}$. Dann ist $\mathcal{U}$ eine offene Überdeckung von X. Wegen X kompakt existieren also endlich viele $y_1, \ldots, y_n \in X$ mit $X = U_{y_1} \cup \cdots \cup U_{y_n}$. Setze nun $h_x = \min(h_{x,y_1}, \ldots, h_{x,y_n})$. Wegen $(*)$ ist dann $h_x \in A$ und nach (Δ) gilt

$$\forall z \in X\colon h_x(z) < f(z) + \varepsilon \tag{$\circ$}$$

Weiterhin ist $h_x(x) = f(x)$, also $(h_x - f)(x) = 0$. Wegen Stetigkeit existiert also für alle $x \in X$ eine offene Umgebung V_x von x mit

$$\forall z \in U_x \colon f(z) - \varepsilon < h_x(z) \tag{$\heartsuit$}$$

Setze nun $\mathcal{V} = \{v_x \colon x \in X\}$. Dann ist $\mathcal{V}$ eine offene Überdeckung von X. Wegen X kompakt existieren also endlich viele $x_1, \dots, x_n \in X$ mit $X = V_{x_1} \cup \dots \cup V_{x_n}$. Setze nun $g = \max(h_{x_1}, \dots, h_{x_n})$. Wegen $(*)$ ist $g \in A$ und nach $(\circ)$,$(\heartsuit)$

$$\forall z \in X \colon f(z) - \varepsilon < g(z) < f(z) + \varepsilon$$

Sei nun A beliebig wie im Satz und sei $\overline{A}$ der Abschluss von A (wie in früherer Bemerkung). Es genügt zu zeigen, dass $\overline{A}$ dicht in $C(X)$ liegt. Nach früherer Bemerkung ist $\overline{A}$ eine Algebra. Nach dem ersten Teil des Beweises genügt es also zu zeigen, dass für alle $f, g \in \overline{A}$ $\max(f,g), \min(f,g) \in \overline{A}$ gilt. Hierzu reicht es aber aus, dass für alle $f, g \in A$ gilt, dass $\max(f,g), \min(f,g) \in \overline{A}$. Nach obiger Bemerkung ist aber

$$\begin{aligned} \max(f,g) &= \frac{f+g}{2} + \frac{|f-g|}{2} \\ \min(f,g) &= \frac{f+g}{2} - \frac{|f-g|}{2} \end{aligned}$$

Da $\overline{A}$ eine Algebra ist, genügt es also zu zeigen, dass für $f \in A$ $|f| \in \overline{A}$ gilt. Sie hierzu $f \in A$ und setze $c = \|f\|$. Es genügt zu zeigen, dass für alle $\varepsilon > 0$ ein $g \in A$ existiert mit $\|g - |f|\| < \varepsilon$. Sei hierzu $\varepsilon > 0$. Nach Lemma 2 existiert ein Polynom p mit $|p(t) - |t|| \leq \varepsilon$ für alle $t \in [-c, c]$. Dann ist aber $|p(f(x)) - |f(x)|| \leq \varepsilon$ für alle $x \in X$. Aber $p \circ f \in A$, da A eine Algebra ist. □

Korollar. *Sei $X \subseteq \mathbb{R}^n$ kompakt und sei P die Menge aller Polynomfunktionen auf X. Dann ist P dicht in $C(X)$.*

Beweis. Offenbar ist P eine Algebra, die alle konstanten Funktionen enthält. Um zu zeigen, dass P die Punkte von X trennt, betrachte für $1 \leq i \leq n$ die Polynomfunktionen $p_i(x_1, \dots, x_n) = x_i$. Sind $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in X$ mit $x \neq y$, so existiert i mit $x_i \neq y_i$. Dann ist $p_i(x) = x_i \neq y_i = p_i(y)$. Also folgt die Behauptung aus Satz 3. □

Es gibt noch eine komplexwertige Version von Satz 3. Sei dann für einen kompakten metrischen Raum X $C(X, \mathbb{C})$ die Menge aller stetigen Funktionen von X nach $\mathbb{C}$. Wir können auf $C(X, \mathbb{C})$ die Supremumsnorm analog definieren durch

$$\|f\| = \sup\{|f(x)| \colon x \in X\}$$

$C(X, \mathbb{C})$ ist dann ein normierter $\mathbb{C}$-Vektorraum. Wir sagen, dass $A \subseteq C(X, \mathbb{C})$ eine $\mathbb{C}$-Algebra ist, wenn für alle $f, g \in A$ und $c \in \mathbb{C}$ $f + g, fg, cf \in A$ gilt. Für $f \colon X \to \mathbb{C}$ definiere die Konjugierte $\overline{f}$ durch $\overline{f}(x) = \overline{f(x)}$ für $x \in X$. Wir sagne, dass A selbstkonjugiert ist, wenn für alle $f \in A$ gilt $\overline{f} \in A$.

Satz 4. *Sei X ein kompakter metrischer Raum und sei $A \subseteq C(X,\mathbb{C})$ eine selbstkonjugierte $\mathbb{C}$-Algebra, welche die konstanten Funktionen enthält und die Punkte von X trennt. Dann ist A dicht in $C(X,\mathbb{C})$.*

Beweis. Sei $A(\mathbb{R}) = A \cap C(X)$. Dann ist A eine Algebra.

$$A(\mathbb{R}) \text{ trennt die Punkte von } X \tag{1}$$

Seien hierzu $x, y \in X$ mit $x \neq y$. Wähle $f \in A$ mit $f(x) \neq f(y)$. Definiere $g\colon X \to \mathbb{C}$ durch $g(z) = (f(z) - f(x))/(f(y) - f(x))$. Dann ist $g \in A$ und $g(x) = 0$, $g(y) = 1$. Setze nun $h = g + \overline{g}$. Wegen A selbstkonjugiert ist dann $h \in A(\mathbb{R})$ und $h(x) = 0 \neq 2 = h(y)$ □(1) Somit folgt aus Satz 3

$$A(\mathbb{R}) \text{ ist dicht in } C(X)$$

Sei nun $\varphi \in C(X,\mathbb{C})$ und $\varepsilon > 0$. Dann existieren $\psi_0, \psi_1 \in C(X)$ mit $\varphi = \psi_0 + i\psi_1$. Gemäß (2) wähle $f_0, f_1 \in A(\mathbb{R})$ mit $\|\psi_0 - f_0\| \leq \varepsilon$ und $\|\psi_1 - f_1\| \leq \varepsilon$. Setze $f = f_0 + if_1$. Dann ist $f \in A$ und $\|\varphi - f\| \leq 2\varepsilon$. □

Definition. Sei $S = \{z \in \mathbb{C}\colon |z| = 1\}$. Eine Funktion $f\colon S \to \mathbb{C}$ ist ein *trigonometrisches Polynom* (auf S), wenn es ein $n \in \mathbb{N}$ gibt und $c_k \in \mathbb{C}$ für $k \in \mathbb{Z}$ mit $-n \leq k \leq n$, sodass für alle $z \in S$ gilt

$$f(z) = \sum_{k=-n}^{n} c_k z^k$$

Satz 5. *Sei $S = \{z \in \mathbb{C}\colon |z| = 1\}$, und sei T die Menge aller trigonometrischen Polynome auf S. Dann ist T dicht in $C(S,\mathbb{C})$.*

Beweis. Offenbar ist T eine $\mathbb{C}$-Algebra, welche alle konstanten Funktionen enthält und die die Punkte von S trennt. Wegen Satz 4 brauchen wir also nur noch zu zeigen, dass T selbstkonjugiert ist. Sei hierzu $f \in T$ und etwa $f(z) = \sum_{k=-n}^{n} c_k z^k$ für $z \in \mathbb{C}$. Nun ist aber für $z \in S$ $\overline{z} = z^{-1}$. Also gilt für $z \in S$

$$\overline{f}(z) = \sum_{k=-n}^{n} c_k z^{-k} = \sum_{k=-n}^{n} c_{-k} z^k$$

Somit ist $\overline{f} \in T$. □

Der Kreis S wird ja parametrisiert durch die Funktion $g\colon [0, 2\pi] \to S$ definiert durch $g(t) = e^{it} = \cos t + i \sin t$. Jedem Element von $C(S,\mathbb{C})$ entspricht dann kanonisch eine stetige Funktion $f^*\colon [0, 2\pi] \to \mathbb{C}$ mit $f^*(0) = f^*(2\pi)$ definiert durch $f^*(t) = f(e^{it})$. f^* kann man dann zu einer 2π-periodischen Funktion f^{**} fortsetzen, die auf ganz $\mathbb{R}$ definiert ist, indem man $f^{**}(2\pi n + t) = f^*(t)$ für $n \in \mathbb{Z}$ und $t \in [0, 2\pi]$ setzt. Auf diese Weise erhält man alle 2π-periodischen stetigen Funktionen $h\colon \mathbb{R} \to \mathbb{C}$. Beim Übergang von f zu f^* zu f^{**} verändert sich die Supremumsnorm nicht.

Wir sagen nun $f\colon \mathbb{R} \to \mathbb{C}$ ist ein trigonometrisches Polynom auf $\mathbb{R}$, wenn ein $n \in \mathbb{N}$ existiert und $c_k \in \mathbb{C}$ für $k \in \mathbb{Z}$ mit mit $-n \leq k \leq n$, sodass gilt

$$f(t) = \sum_{k=-n}^{n} c_k e^{ikt}$$

Korollar. *Sei B die Menge aller stetigen 2π-periodischen Funktionen von $\mathbb{R}$ nach $\mathbb{C}$ (mit Supremumsnorm) und sei T die Menge aller trigonometrischen Polynome (auf $\mathbb{R}$). Dann ist T dicht in B.*

Beachte noch, dass sich wegen $e^{ix} = \cos x + i \sin x$ jedes trigonometrische Polynom schreiben lässt als

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kt) + i \sum_{k=1}^{n} b_k \sin(kt)$$

Definition. Sei $(K_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $\mathbb{R} \to \mathbb{R}$. Dann ist $(K_n)_{n\in\mathbb{N}}$ eine *Diracfolge*, wenn gilt

(DIR1) Für alle $x \in \mathbb{R}$ und alle $n \in \mathbb{N}$ ist $K_n(x) \geq 0$.

(DIR2) Für alle $n \in \mathbb{N}$ ist K_n stetig und $\int_{-\infty}^{\infty} K_n(t)\,\mathrm{d}t = 1$.

(DIR3) Für alle $\varepsilon, \delta > 0$ existiert ein $m \in \mathbb{N}$, sodass für alle $n \geq m$ gilt:

$$\int_{-\infty}^{-\delta} K_n(t)\,\mathrm{d}t + \int_{\delta}^{\infty} K_n(t)\,\mathrm{d}t < \varepsilon$$

Sei $(K_n)_{n\in\mathbb{N}}$ eine Diracfolge. Für eine stetige und beschränkte Funktion $f\colon \mathbb{R} \to \mathbb{R}$ definiere die *Faltung* $f * K_n$ von f und K_n durch

$$(f * K_n)(x) = \int_{-\infty}^{\infty} f(t) K_n(x-t)\,\mathrm{d}t$$

Satz 6. *Sei $(K_n)_{n\in\mathbb{N}}$ eine Diracfolge. Weiterhin sei $S \subseteq \mathbb{R}$ beschränkt und sei $f\colon \mathbb{R} \to \mathbb{R}$ stetig und beschränkt. Dann konvergiert die Folge $(f * K_n)_{n\in\mathbb{N}}$ auf S gleichmäßig gegen f.*

Beweis. Setze $f_n = f * K_n$. Durch einfache lineare Substitution erhält man

$$f_n(x) = \int_{-\infty}^{\infty} f(x-t) K_n(t)\,\mathrm{d}t$$

Wegen (DIR2) gilt aber

$$f(x) = f(x) \int_{-\infty}^{\infty} K_n(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} f(x) K_n(t)\,\mathrm{d}t$$

Also ist

$$f_n(x) - f(x) = \int_{-\infty}^{\infty} \big(f(x-t) - f(x)\big) K_n(t)\,\mathrm{d}t$$

Sei nun $\varepsilon > 0$. Wegen S beschränkt existiert ein kompaktes Intervall $I = [c, d]$ mit $S \subseteq [c+\mu, d-\mu]$ für ein $\mu > 0$. Dann ist f auf I gleichmäßig stetig. Also existiert ein $\delta > 0$ mit $\big|f(x-t) - f(x)\big| < \frac{\varepsilon}{2}$ für alle $x \in S$ und $t \in \mathbb{R}$ mit $|t| < \delta$. Sei nun $M > 0$ eine

Schranke für f, d.h. $|f(x)| \leq M$ für alle $x \in \mathbb{R}$. Gemäß (DIR3) wähle m so, dass für alle $n \geq m$ gilt

$$\int_{-\infty}^{-\delta} K_n(t)\,\mathrm{d}t + \int_{\delta}^{\infty} K_n(t)\,\mathrm{d}t < \frac{\varepsilon}{4M}$$

Wir erhalten dann für $x \in S$

$$\begin{aligned}
\big|f_n(x) - f(x)\big| &= \left|\int_{-\infty}^{\infty} \big(f(x-t) - f(x)\big) K_n(t)\,\mathrm{d}t\right| \leq \\
&\leq \int_{-\infty}^{\infty} \big|f(x-t) - f(x)\big| K_n(t)\,\mathrm{d}t = \\
&= \int_{-\infty}^{-\delta} \big|f(x-t) - f(x)\big| K_n(t)\,\mathrm{d}t + \int_{-\delta}^{\delta} \big|f(x-t) - f(x)\big| K_n(t)\,\mathrm{d}t + \\
&\quad + \int_{\delta}^{\infty} \big|f(x-t) - f(x)\big| K_n(t)\,\mathrm{d}t \leq \\
&\leq 2M \left(\int_{-\infty}^{-\delta} K_n(t)\,\mathrm{d}t + \int_{\delta}^{\infty} K_n(t)\,\mathrm{d}t\right) + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} K_n(t)\,\mathrm{d}t < \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

□

Beweis von Lemma 2. Wir zeigen, dass für ein kompaktes Intervall $[a, b]$ und eine stetige Funktion $f\colon [a, b] \to \mathbb{R}$ eine Folge von Polynomfunktionen auf $[a, b]$ existiert, die gleichmäßig gegen f konvergiert.

Sei o.E. $a \neq b$. Wir reduzieren die Behauptung zuerst auf den Fall $[a, b] = [0, 1]$ und $f(a) = f(b) = 0$. Wir müssen ja zeigen:

($*$) für alle $\varepsilon > 0$ existiert ein Polynom P mit $\|f - P\| \leq \varepsilon$.

Definiere $g\colon [0, 1] \to \mathbb{R}, t \mapsto f\big((b-a)t+a\big)$. Falls ein Polynom P existiert mit $\|g-P\| \leq \varepsilon$, so gilt $\left|f(x) - P\left(\frac{x-a}{b-a}\right)\right| \leq \varepsilon$ für alle $x \in [a, b]$. Aber $P\left(\frac{x-a}{b-a}\right)$ ist ein Polynom. Somit können wir voraussetzen, dass $[a, b] = [0, 1]$. Setze nun $h(x) = f(x) - f(0) - x\big(f(1) - f(0)\big)$ für alle $x \in [0, 1]$. Wenn wir ($*$) für h beweisen, so gilt es auch für f. Aber $h(0) = h(1) = 0$. Wir können also auch $f(a) = f(b) = 0$ voraussetzen.

Wir setzen f auf ganz $\mathbb{R}$ fort, indem wir es außerhalb von $[0, 1]$ gleich 0 setzen. Wir nennen dies immer noch f. Dann ist f stetig und beschränkt. Für $n \in \mathbb{N}$ setze

$$c_n = \int_{-1}^{1} \left(1 - t^2\right)^n \,\mathrm{d}t$$

Definiere $K_n\colon \mathbb{R} \to \mathbb{R}$ durch

$$K_n(t) = \begin{cases} \frac{(1-t^2)^n}{c_n} & \text{falls } -1 \leq t \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Beachte, dass $K_n(t) = K_n(-t)$ gilt. Wir zeigen, dass $(K_n)_{n\in\mathbb{N}}$ eine Dirac-Folge ist. (DIR1) und (DIR2) sind klar. Es ist außerdem

$$\frac{c_n}{2} = \int_0^1 (1-t^2)^n \,\mathrm{d}t = \int_0^1 (1+t)^n (1-t)^n \,\mathrm{d}t \geq \int_0^1 (1-t)^n \,\mathrm{d}t = \frac{1}{n+1}$$

d.h. $c_n \geq \frac{2}{n+1}$. Für $1 \geq \delta > 0$ folgt

$$\begin{aligned}\int_\delta^1 K_n(t)\,\mathrm{d}t &= \int_\delta^1 \frac{(1-t^2)^n}{c_n}\,\mathrm{d}t \leq \int_\delta^1 \frac{n+1}{2}(1-\delta^2)^n\,\mathrm{d}t = \\ &= \frac{n+1}{2}(1-\delta^2)^n(1-\delta) \underset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

Also gilt auch (DIR3). Schließlich zeigen wir, dass $f * K_n|_{[0,1]}$ ein Polynom ist. Sei also $f * K_n = f_n$. Es ist

$$f_n(x) = \int_{-\infty}^{\infty} f(t)K_n(x-t)\,\mathrm{d}t$$

Da $f(t)K_n(x-t)$ für $t \in \mathbb{R} \setminus [0,1]$ gleich 0 ist, ist

$$f_n(x) = \int_0^1 f(t)K_n(x-t)\,\mathrm{d}t$$

Nun ist aber K_n in dem Intervall $[-1,1]$ ein Polynom. Somit ist $K_n(x-t)$ ein Polynom in x und t und daher darstellbar als

$$K_n(x-t) = \sum_{k=0}^{2n} g_k(t)x^k$$

wobei $g_0, \ldots, g_{2n}$ Polynome in t sind. Daher gilt

$$f_n(x) = \int_0^1 f(t)\sum_{k=0}^{2n} g_k(t)x^k\,\mathrm{d}t = \sum_{k=0}^{2n}\left(\int_0^1 f(t)g_k(t)\,\mathrm{d}t\right)x^k$$

Also ist f_n in $[0,1]$ ein Polynom. Somit folgt $(*)$ aus Satz 6. $\square$