Vorlesungsskript zur Vorlesung Analysis 2
(+ Teile der Ergänzungsvorlesung)
Sommersemester 2009

Prof. Dr. M. Schwarz
Mathematisches Institut
Universität Leipzig

Inhaltsverzeichnis

KAPITEL V

Integration

V.1. Das Riemannsche Integral

Voraussetzung: $[a,b] \subset \mathbb{R}$ sei ein kompaktes Integral über das integriert werden soll. $f\colon [a,b] \to \mathbb{R}$ sei eine beschränkte Funktion (nicht notwendigerweise stetig!).

V.1.1. Definition.

Eine **Zerlegung** Z von $[a,b]$ ist eine endliche, geordnete Teilmenge $\{z_0, z_1, \ldots, z_n\} \subset [a,b]$ mit $a = z_0 < z_1 < \ldots < z_n = b$, $n \in \mathbb{N}$ beliebig.
z_i heißt i-ter Zerlegungspunkt,
$L(Z) := \max_{i=1,\ldots,n}(z_i - z_{i-1})$ die **Zerlegungslänge**,
$U(f,Z) := \sum_{i=1}^{n}(z_i - z_{i-1}) \inf_{[z_{i-1},z_i]} f$ die **Untersumme** von f bzgl. Z,
$O(f,Z) := \sum_{i=1}^{n}(z_i - z_{i-1}) \sup_{[z_{i-1},z_i]} f$ die **Obersumme** von f bzgl. Z .
Seien $\mathcal{Z}_{[a,b]} := \{\, Z \,\text{Zerlegung von}\, [a,b] \,\}$ die Menge aller Zerlegungen und $Z, Z' \in \mathcal{Z}_{[a,b]}$. Dann heißt Z' eine **Verfeinerung** von Z genau dann, wenn $Z \subset Z'$.

V.1.2. Bemerkung. (1) $Z \subset Z' \Rightarrow L(Z') \leq L(Z)$,
(2) $Z, Z' \in \mathcal{Z}_{[a,b]} \Rightarrow Z \cup Z' \in \mathcal{Z}_{[a,b]}$ und dies ist eine Verfeinerung von Z und Z', d.h. $\mathcal{Z}_{[a,b]}$ trägt eine Teilordnung.

V.1.3. Hilfssatz. *Es seien $Z, Z' \in \mathcal{Z}_{[a,b]}$, $Z \subset Z'$ und $k = \#(Z' \setminus Z)$. Dann gilt*
(1) $U(f,Z) \leq U(f,Z') \leq U(f,Z) + 2kcL(Z)$,
(2) $O(f,Z) \geq O(f,Z') \geq O(f,Z) - 2kcL(Z)$
mit $c = \sup_{[a,b]} |f|$.

Beweis. *Zu (1)*: Es sei $Z = \{z_0, \ldots, z_n\}$ und $\xi \in [a,b]$ mit $\xi \notin Z$. Dann existiert ein $i \in \{1, \ldots, n\}$ mit $z_{i-1} < \xi < z_i$. Setze dann $Z^* = \{z_0, \ldots, z_{i-1}\xi, z_i, \ldots, z_n\}$, und es folgt $Z \subset Z^*$, und somit

$$U(f,Z^*) - U(f,Z) \;=\; (\xi - z_{i-1}) \inf_{[z_{i-1},\xi]} f + (z_i - \xi) \inf_{[\xi,z_i]} f - (z_i - z_{i-1}) \inf_{[z_{i-1},z_i]} f \,.$$

aus $\inf_{[z_{i-1},z_i]} f \leq \inf_{[z_{i-1},\xi]} f$ und $\inf_{[z_{i-1},z_i]} f \leq \inf_{[\xi,z_i]} f$ folgt dann

$$U(f,Z) \leq U(f,Z^*),$$

und somit

$$\begin{aligned} U(f,Z^*) - U(f,Z) \;&=\; |U(f,Z^*) - U(f,Z)| \\ &\leq\; |\xi - z_{i-1}| \cdot |\inf_{[z_{i-1},\xi]} f| \;+\; |z_i - \xi| \cdot |\inf_{[\xi,z_i]} f| \;+\; |z_i - z_{i-1}| \cdot |\inf_{[z_{i-1},z_i]} f| \\ &\leq\; 2|z_i - z_{i-1}|c \;\leq\; 2cL(Z)\,. \end{aligned}$$

Wir wenden nun die 1-Punkt-Verfeinerung Z^* von Z k-mal an, d.h. Induktion nach k und erhalten die Behauptung.
(2) ergibt sich analog. □

V.1.4. Hilfssatz. *Es gilt $U(f,Z_1) \leq O(f,Z_2)$ für alle $Z_1, Z_2 \in \mathcal{Z}_{[a,b]}$.*

Beweis. Betrachte $Z_1, Z_2 \subset Z_1 \cup Z_2 =: Z'$. Dann ergibt Hilfssatz V.1.3

$$U(f,Z_1) \leq U(f,Z') \leq O(f,Z') \leq O(f,Z_2),$$

mit der mittleren Ungleichung aus der Definition. □

Als Konsequenz erhalten wir:

$$\begin{aligned} &\{\, U(f,Z) \mid Z \in \mathcal{Z}_{[a,b]} \,\} \quad \text{ist oberhalb beschränkt,} \\ &\{\, O(f,Z) \mid Z \in \mathcal{Z}_{[a,b]} \,\} \quad \text{ist unterhalb beschränkt}\,, \end{aligned}$$

und wir definieren

V.1.5. Definition.
das **untere Riemannsche Integral**

$$\underline{\int_a^b} f(x)\,dx := \sup_{Z \in \mathcal{Z}_{[a,b]}} U(f,Z)$$

und das **obere Riemannsche Integral**

$$\overline{\int_a^b} f(x)\,dx := \inf_{Z \in \mathcal{Z}_{[a,b]}} O(f,Z)\,.$$

V.1.6. Beispiel. (1) Betrachte $f\colon [0,1] \to \mathbb{R}$, $f(x) = x^2$ und die Zerlegung $Z_n = \{\, 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n} \,\}$, $n \in \mathbb{N}$. Dann gilt

$$\begin{aligned} U(f,Z_n) &= \sum_{i=1}^{n} \frac{1}{n} f\left(\frac{i-1}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} \frac{(i-1)^2}{n^2} = \frac{1}{n^3}\sum_{i=0}^{n-1} i^2 \\ &= \frac{1}{n^3} \cdot \frac{(n-1)(n-\frac{1}{2})n}{3} = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}\,, \end{aligned}$$

und

$$\begin{aligned} O(f,Z_n) &= \sum_{i=1}^{n} \frac{1}{n} f\left(\frac{i}{n}\right) = \frac{1}{n^3}\sum_{i=1}^{n} i^2 = \frac{1}{3}\frac{(n+\frac{1}{2})(n+1)}{n^2} \\ &= \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}\,. \end{aligned}$$

(2) Sei $f\colon [0,1] \to \mathbb{R}$, $f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \in \mathbb{R}\setminus\mathbb{Q}\,. \end{cases}$. Offenbar ist f beschränkt, und es gilt für alle Zerlegungen Z

- $U(f,Z) = 0$, da $\inf f = 0$ für alle Teilintervalle von $[0,1]$,
- $O(f,Z) = 1$, da $\sup f = 0$ für alle Teilintervalle von $[0,1]$,

also folgt $\underline{\int_0^1} f(x)dx = 0$ und $\overline{\int_0^1} f(x)dx = 1$.

V.1.7. Hilfssatz. *Es gilt für alle Zerlegungen $Z \in \mathcal{Z}_{[a,b]}$*

$$U(f,Z) \le \underline{\int_a^b} f(x)dx \le \overline{\int_a^b} f(x)dx \le O(f,Z)\,.$$

Beweis. folgt direkt aus der Definition. □

V.1.8. Hilfssatz. *Für alle $\epsilon > 0$ existiert ein $\delta = \delta(\epsilon) > 0$ so daß für alle Zerlegungen $Z \in \mathcal{Z}_{[a,b]}$ mit $L(Z) < \delta$ gilt:*

$$0 \le \underline{\int_a^b} f(x)dx - U(f,Z) < \epsilon \qquad \textit{und} \qquad 0 \le O(f,Z) - \overline{\int_a^b} f(x)dx < \epsilon\,.$$

Eine Konsequenz hieraus ist, daß zur Berechnung dieser unteren und oberen Riemannschen Integrale Zerlegungs-Folgen $(Z_n) \subset \mathcal{Z}_{[a,b]}$ mit $L(Z_n) \to 0$ genügen.

Beweis. Es sei $\epsilon > 0$ gegeben und wir wählen eine Zerlegung $Z_o \in \mathcal{Z}_{[a,b]}$ mit $\underline{\int}_a^b f(x)dx - U(f, Z_o) < \frac{\epsilon}{2}$, welche nach der Definition des unteren Integrals als eine Supremum von Untersummen existiert.
Sei nun $m = \# Z_o$ die Anzahl der Zerlegungspunkte und setze $\delta = \frac{\epsilon}{4mc}$ mit $c = \sup_{[a,b]} |f|$. Dann gilt für alle $Z \in \mathcal{Z}_{[a,b]}$ mit $L(Z) < \delta$

$$\underline{\int}_a^b f(x)dx \; - \; U(f, Z) \; < \; \epsilon \,. \tag{1}$$

Dies sieht man mittels $Z' = Z_o \cup Z$, also $\# Z' \setminus Z = k \leq m$, und Hilfssatz V.1.3 liefert dann

$$U(f, Z') - U(f, Z) \; \leq \; 2kcL(Z) \; \leq \; 2mcL(Z) \; < \; \frac{\epsilon}{2} \,. \tag{2}$$

Andererseits folgt aus $U(f, Z') \geq U(f, z_o)$, daß auch

$$\underline{\int}_a^b f(x)dx - U(f, Z') \; < \; \frac{\epsilon}{2} \,. \tag{3}$$

Somit folgt (1) aus der Addition von (2) und (3). $\square$

V.1.9. Definition.
Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ heißt **(Riemann-) integrierbar** auf $[a, b]$ genau dann, wenn ihr unteres und oberes Riemannsches Integral übereinstimmen. Dies ist dann das **Riemannsche Integral von** f, also

$$\underline{\int}_a^b f(x)dx \; = \; \overline{\int}_a^b f(x)dx \; =: \; \int_a^b f(x)dx \,.$$

Wir bezeichnen mit $\mathcal{R}([a, b], \mathbb{R})$ die Menge der auf $[a, b]$ Riemann-integrierbaren Funktionen.

V.1.10. Bemerkung.
Falls f Riemann-integrierbar ist so ist das Integral $\int_a^b f(x)dx$ berechenbar durch eine Wahl einer Zerlegungsfolge $(Z_n)_{n \in \mathbb{N}} \subset \mathcal{Z}_{[a,b]}$ mit $L(Z_n) \to 0$, als

$$\int_a^b f(x)dx \; = \; \lim_{n \to \infty} U(f, Z_n) \; = \; \lim_{n \to \infty} O(f, Z_n) \,.$$

Dies folgt aus Hilfssatz V.1.8.

V.1.11. Beispiel.
Betrachte wieder $f(x) = x^2$ aus Beispiel V.1.6.(1) mit derselben gleichmäßigen Zerlegung Z_n. Dann gilt

$$U(f, Z_n) = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2} \to \frac{1}{3} = \lim_{n \to \infty} O(f, Z_n) = \int_0^1 x^2 dx \,.$$

V.1.12. Satz (**Riemannsches Integrabilitätskriterium**). *Es sei $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion. Dann gilt*

$$f \text{ R.-integrierbar} \Leftrightarrow \forall \epsilon > 0 \, \exists \, Z \in \mathcal{Z}_{[a,b]} \text{ s.d. } O(f, Z) - U(f, Z) < \epsilon \,.$$

Beweis. "$\Rightarrow$": Sei $\epsilon > 0$ gegeben und wähle $\delta(\epsilon/2) > 0$ wie in Hilfssatz V.1.8. Dann finden wir eine Zerlegung $Z \in \mathcal{Z}_{[a,b]}$ mit $L(Z) < \delta$ und es folgt aus der Gleichheit von unterem und oberen Riemannschen Integral $O(f, Z) - U(f, Z) < \epsilon$.
"$\Leftarrow$": Falls f nicht integrierbar ist so ist $\overline{\int}_a^b f(x)dx - \underline{\int}_a^b f(x)dx =: \epsilon_o > 0$ und somit gilt auch $O(f, Z) - U(f, Z) \geq \epsilon_o > 0$ für alle Zerlegungen Z im Widerspruch zur rechten Seite. $\square$

V.1.13. Definition.
Es sei $Z = \{ z_0, \dots, z_n\} \in \mathcal{Z}_{[a,b]}$ und $\xi_1, \dots, x_n \in [a,b]$ mit $z_{i-1} \le \xi_i \le z_i$ f.a. $i = 1, \dots, n$. Dann definieren wir die Zwischen- oder **Riemannsumme** als

$$S(f, Z, \xi) := \sum_{i=1}^{n} (z_i - z_{i-1}) f(\xi_i)\,.$$

Offenbar gilt $U(f,Z) \le S(f,Z,\xi) \le O(f,Z)$, und für alle $\epsilon > 0$ existiert ein $\delta(\epsilon) > 0$, so daß für Riemann-integrierbare f die Abschätzung

$$\left| \int_a^b f(x)dx - S(f, Z\xi) \right| \le O(f,Z) - U(f,Z) < \epsilon$$

für alle Zerlegungen Z mit $L(Z) < \delta$ gilt.

V.1.14. Definition.
Falls $(Z_k)_{k\in\mathbb{N}} \subset \mathcal{Z}_{[a,b]}$ eine Folge von Zerlegungen mit $L(Z_k) \to 0$ ist und jeweils Zwischenpunkt-Tupel $\xi^{(k)}$ zu Z_k gegeben sind mit $z_0 \le \xi_1^{(k)} \le z_1 \le \ldots \le z_{n_k - 1} \le \xi_{n_k}^{(k)} \le z_{n_k}$, so heißt $S_k = S(f, Z_k, \xi^{(k)})$ eine **Riemann-Folge**.

V.1.15. Satz. *Sei $f\colon [a,b] \to \mathbb{R}$ beschränkt. Dann ist f Riemann-integrierbar genau dann, wenn ale Riemannfolgen konvergieren, und in diesem Fall konvergieren die Riemannfolgen gegen das Integral $\int_a^b f(x)dx$.*

Beweis. <u>"$\Rightarrow$"</u>: Wir wissen bereits: Aus f integrierbar folgt $|\int f(x)dx - S(f,Z,\xi)| < \epsilon$ für alle $L(Z) < \delta(\epsilon)$ und somit konvergieren nach Hilfssatz V.1.8 alle Riemannfolgen.
<u>"$\Leftarrow$"</u>: Wir nehmen nun an, daß alle Riemannfolgen konvergieren.
<u>*Schritt 1*</u>: Es seien $S_k = (f, Z_k, \xi^{(k)})$ und $S_k' = (f, Z_k', \xi'^{(k)})$ Riemannfolgen. Dann ist auch $Z_1, Z_1', Z_2, Z_2', \dots$ eine Zerlegungsfolge mit $L(\dots) \to 0$ und $S_1, S_1', S_2, S_2', \dots,$ ist eine Riemannfolge, also ebenfalls konvergent. Daher muss für die Teilfolgen Gleichheit $\lim S_k = \lim S_k' =: \sigma$ gelten. Also ist der Grenzwert $\sigma = \lim S_k$ unabhängig von der Wahl der Riemannfolge.
<u>*Schritt 2*</u>: Wir betrachten $U(f,Z) = \sum_{i=1}^n (z_i - z_{i-1}) \inf_{[z_{i-1}, z_i]} f$ für eine gegebene Zerlegung Z und fixieren $\epsilon > 0$. Wir finden $\xi_i \in [z_{i-1}, z_i]$ für alle $i = 1, \dots, n$, so daß

$$\inf_{z_{i-1}, z_i]} f \;\le\; f(\xi_i) \;\le\; \inf_{z_{i-1}, z_i]} f + \frac{\epsilon}{2(b-a)}\,.$$

Dann gilt $U(f,Z) \le S(f,Z,\xi) \le U(f,Z) + \frac{1}{2}\epsilon$.
<u>*Schritt 3*</u>: Es sei nun $Z_1, Z_2, \dots$ eine Zerlegungsfolge mit $L(Z_k) \to 0$. Nach Schritt 2 finden wir entsprechende Folgen von Zwischenpunkt-Tupeln $\xi^{(k)}$ so daß $0 \le S(f, Z_k, \xi^{(k)}) - U(f, Z_k) \le \frac{\epsilon}{2}$ für alle $k \in \mathbb{N}$. Gemäß Hilfssatz V.1.8 existiert ein $k_o(\epsilon) \in \mathbb{N}$ mit $\underline{\int}_a^b f(x)dx - U(f, Z_k) < \frac{\epsilon}{2}$ für alle $k \ge k_o(\epsilon)$. Also folgt

$$\left| \underline{\int}_a^b f(x)dx \;-\; S(f, Z_k, \xi^{(k)}) \right| < \epsilon \quad \text{f.a. } k \ge k_o\,.$$

Da $S(f, Z_k, \xi^{(k)}) \to \sigma$ folgt $\left| \underline{\int}_a^b f(x)dx - \sigma \right| \le \epsilon$ für alle $\epsilon > 0$.
<u>*Schritt 4*</u>: Mit genau analoger Argumentation für $\overline{\int}_a^b f(x)dx$ folgt

$$\underline{\int}_a^b f(x)dx \;=\; \sigma \;=\; \overline{\int}_a^b f(x)dx\,.$$

Somit ist f Riemann-integrierbar und das Integral gleich σ. □

V.2. Riemann-Integrierbarkeit

V.2.1. Satz. *Jede monotone Funktion $f\colon [a,b] \to \mathbb{R}$ ist Riemann-integrierbar.*

Beweis. Ohne Beschränkung der Allgemeinheit nehmen wir an, daß f monoton wachsend ist, also wegen $f(a) \le f(x) \le f(b)$ f.a. $x \in [a,b]$ auch beschränkt. Sei un $Z_n = \{z_0, \dots, z_n\}$ die gleichmäßige Zerlegung mit $z_i = a + \frac{i}{n}(b-a)$, $i = 0, \dots, n$. Dann gilt wegen der Monotonie

$$U(f, Z_n) = \sum_{i=1}^{n} (z_i - z_{i-1}) \inf_{[z_{i-1}, z_i]} f = \frac{b-a}{n} \sum_{i=1}^{n} f(z_{i-1}),$$
$$O(f, Z_n) = \frac{b-a}{n} \sum_{i=1}^{n} f(z_i)\,.$$

Also gilt $O(f,Z_n) - U(f,Z_n) = \frac{b-a}{n}\big[f(b) - f(a)\big]$. Sei nun $n > \frac{(b-a)[f(b)-f(a)]}{\epsilon}$ so folgt $O(\dots) - U(\dots) < \epsilon$. Also ist gemäß dem Riemannschen Integrabilitätskriterium V.1.12 die Funktion f integrierbar. □

V.2.2. Satz. $C^0([a,b]) \subset \mathcal{R}([a,b], \mathbb{R})$.

V.2.3. Definition.

Es sei $X \subset \mathbb{R}$ eine Teilmenge und $f\colon X \to \mathbb{R}$ eine beliebige Funktion. Dann heißt f **gleichmäßig stetig** genau dann, wenn

$$\forall\, \epsilon > 0\ \exists\, \delta > 0\colon \quad |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \epsilon\,.$$

Beachte, $\delta = \delta(\epsilon, f)$, also δ ist unabhängig von $x \in X$.

V.2.4. Beispiel.

Die Funktion $\ln x$ is auf $(0, \infty)$ stetig aber nicht gleichmäßig stetig. Beweis siehe Übung. Verwende: Für $0 < x < x'$ gilt:

$$\ln x' - \ln x < \epsilon \Leftrightarrow x' - x < x(e^\epsilon - 1)\,.$$

V.2.5. Lemma. *Sei $f \in C^0([a,b], \mathbb{R})$, so ist f gleichmäßig stetig auf $[a,b]$.*

Beweis. Zu zeigen ist: $\forall\, \epsilon\ \exists\, \delta > 0\ \forall\, x, x' \in [a,b]\colon\ |x - x'| < \delta \Rightarrow |f(x) - f(x')| < \epsilon$. Wir führen einen Widerspruchsbeweis und nehme also an, daß

$$\exists\, \epsilon_o > 0\ \forall\, \delta > 0\ \exists x, x' \in [a,b] \quad \text{mit} \quad |x - x'| < \delta \quad \text{und} \quad |f(x) - f(x')| \ge \epsilon\,.$$

Fixiere dieses $\epsilon_o > 0$ und betrachte eine Folge $\delta_n = \frac{1}{n} \to 0$. Somit erhalten wir $x_n, x'_n \in [a,b]$ mit $|x_n - x'_n| < \frac{1}{n} \to 0$ und $|f(x_n) - f(x'_n)| \ge \epsilon_o$.
Da $[a,b]$ ein kompaktes Intervall ist, gilt nach dem Satz von Bolzano-Weierstraß, daß für eine geeignete Teilfolge $(n_k)_{k \in \mathbb{N}}$ $x_{n_k} \to x_o$ und $x'_{n_k} \to x'_o$. Da $|x_n - x'_n| \to 0$, gilt $x_o = x'_o$.
Mit der Stetigkeit von f folgt nun $f(x_{n_k}) \to f(x_o)$ und auch $f(x'_{n_k}) \to f(x_o)$ im Widerspruch zu $|f(x_{n_k}) - f(x'_{n_k})| \ge \epsilon_o > 0$ für alle $k \in \mathbb{N}$. □

Beweis von Satz V.2.2. Es sei $Z_n \in \mathcal{Z}_{[a,b]}$ eine beliebige Zerlegung, also

$$O(f, Z_n) - U(f, Z_n) = \sum_{i=1}^{n} (z_i - z_{i-1}) \big(\sup_{[z_{i-1}, z_i]} f - \inf_{[z_{i-1}, z_i]} f \big)\,.$$

Da f stetig ist, nimmt die Funktion auf jedem Teilintervall $[z_{i-1}, z_i]$ ihr Maximum $f(\xi_i)$und Minimum $f(\eta_i)$ an, an Punkten $\xi_i, \eta_i \in [z_{i-1}, z_i]$.
Nach dem obigen Lemma ist f auch gleichmäßig stetig auf $[a,b]$. Sei nun $\epsilon > 0$ beliebig, dann finden wir ein $\delta = \delta(\frac{\epsilon}{b-a}, f) > 0$ so daß $|f(x) - f(y)| < \frac{\epsilon}{b-a}$ für alle $|x - x'| < \delta$ gilt.

Sei $Z \in \mathcal{Z}_{[a,b]}$ eine Zerlegung mit $L(Z) < \delta$, also $|f(\xi_i) - f(\eta_i)| < \frac{\epsilon}{b-a}$ für alle $i = 1, \dots, n$. Somit gilt

$$0 \leq O(f, Z_n) - U(f, Z_n) = \sum_{i=1}^{n} (z_i - z_{i-1})\big(f(\xi_i) - f(\eta_i)\big) \leq \frac{\epsilon}{b-a} \sum_{i=1}^{n} (z_i - z_{i-1}) = \epsilon .$$

Nach Satz V.1.12 ist f also Riemann-integrierbar. □

V.2.6. Satz. *Es seien $a < b < c$ und $f\colon [a, c] \to \mathbb{R}$ mit $f_{|[a,b]} \in \mathcal{R}([a, b])$ und $f_{|[b,c]} \in \mathcal{R}([b, c])$. Dann gilt auch $f \in \mathcal{R}([a, c])$ mit*

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

Beweis. Sei $\epsilon > 0$. Wähle $Z_1 \in \mathcal{Z}_{[a,b]}$ und $Z_2 \in \mathcal{Z}_{[b,c]}$ mit $O(f_{|[a,b]}, Z_1) - U(f_{|[a,b]}, Z_1) < \frac{\epsilon}{2}$ und $O(f_{|[b,c]}, Z_2) - U(f_{|[b,c]}, Z_2) < \frac{\epsilon}{2}$. Dann bilde die Zerlegung $Z = Z_1 \cup Z_2 \in \mathcal{Z}_{[a,c]}$ und es gilt $O(f, Z) = O(f_{|[a,b]}, Z_1) + O(f_{|[b,c]}, Z_2)$ und $U(f, Z) = U(f_{|[a,b]}, Z_1) + U(f_{|[b,c]}, Z_2)$.
Also erfüllt f wieder das Riemannsche Integrabilitätskriterium V.1.12. □

V.2.7. Satz. *Es sei $f \in \mathcal{R}([a, b])$ und $[c, d] \subset [a, b]$. Dann ist auch $f_{|[c,d]} \in \mathcal{R}([c, d])$.*

Beweis. Es sei $\epsilon > 0$ gegeben. Wir finden eine Zerlegung $Z_o \in \mathcal{Z}_{[a,b]}$ mit $O(f, Z_o) - U(f, Z_o) < \epsilon$. Nehmen wir $\{c, d\}$ hinzu, also $Z' = Z_o \cup \{c, d\}$ so gilt für diese Verfeinerung ebenso $O(f, Z') - U(f, Z') < \epsilon$. Einschränkung auf $[c, d]$ liefert $Z' \cap [c, d] = \bar{Z} \in \mathcal{Z}_{[c,d]}$, also

$$O(f_{|[c,d]}, \bar{Z}) - U(f_{|[c,d]}, \bar{Z}) < \epsilon .$$

□

V.2.8. Definition.
Die Notation für das Integral wird folgendermaßen erweitert: Sei $f \in \mathcal{R}([a, b])$ und $[\alpha, \beta] \subset [a, b]$, dann definieren wir

$$\int_\alpha^\alpha f(x)dx := 0,$$
$$\int_\beta^\alpha f(x)dx := -\int_\alpha^\beta f(x)dx .$$

V.2.9. Satz (**Additivität**). *Für alle $f \in \mathcal{R}([a, b])$ und $\alpha, \beta, \gamma \in [a, b]$ gilt*

$$\int_\alpha^\gamma f(x)dx = \int_\alpha^\beta f(x)dx + \int_\beta^\gamma f(x)dx .$$

Beweis. Folgt unmittelbar aus Satz V.2.6 und Satz V.2.7. □

V.2.10. Satz (**Linearität**). *$\mathcal{R}([a, b])$ ist ein reeller Vektorraum und $f \mapsto \int_a^b f(x)dx$ ist ein lineares Funktional auf $\mathcal{R}([a, b])$.*

Beweis. Betrachte Riemannfolgen, also $(Z_n) \subset \mathcal{Z}_{[a,b]}$ mit $L(Z_n) \to 0$ und Zwischenpunkt-Tupeln $\xi^{(n)}$. Dann folgt die behauptete Linearität aus der für die Riemansummen und der Konvergenz der Riemannfolgen, sowie Satz V.1.15,

$$S(\lambda f, Z_n, \xi^{(n)}) = \lambda S(f, Z_n, \xi^{(n)}),$$
$$S(f + g, Z_n, \xi^{(n)}) = S(f, Z_n, \xi^{(n)}) + S(g, Z_n, \xi^{(n)}),$$

für alle $f, g \in \mathcal{R}([a, b])$, $\lambda \in \mathbb{R}$. □

V.2.11. SATZ. $f \in \mathcal{R}([a,b]) \Rightarrow |f| \in \mathcal{R}([a,b])$.

BEWEIS. Es gilt für $I = [z_{i-1}, z_i] \subset [a,b]$ (Beweis Übung!)

$$\sup_I |f| - \inf_I |f| \leq \sup_I f - \inf_I f\,.$$

Also

$$O(|f|,Z) - U(|f|,Z) = \sum_{i=1}^n (z_i - z_{i-1})\big(\sup_I |f| - \inf_I |f|\big) \leq O(f,Z) - U(f,Z) < \epsilon$$

für Z fein genug. □

V.3. Integral-Eigenschaften

V.3.1. SATZ (**Monotonie**). *Für alle $f, g \in \mathcal{R}([a,b])$ gilt*

$$f(x) \geq g(x) \text{ für alle } x \in [a,b] \quad \Rightarrow \quad \int_a^b f(x)dx \geq \int_a^b g(x)dx\,.$$

BEWEIS. Wie im Beweis der Linearität folgt auch diese Monotonie aus der entsprechenden Eigenschaft für Riemannsummen und den Grenzübergang $L(Z_n) \to 0$. Betrachte also eine Riemannfolge $S(f, Z_n, \xi^{(n)})$, es gilt

$$\begin{aligned} S(f, Z_n, \xi^{(n)}) &= \sum_i (z_i^{(n)} - z_{i-1}^{(n)})\, f(\xi_i^{(n)}) \\ &\geq \sum_i (z_i^{(n)} - z_{i-1}^{(n)})\, g(\xi_i^{(n)}) = S(g, Z_n, \xi^{(n)}) \end{aligned}$$

Die Behauptung folgt dann für $\lim_{n\to\infty}$. □

V.3.2. SATZ (**Integralabschätzungen**). *Für $f \in \mathcal{R}([a,b])$ gilt:*

(1) $m \leq f(x) \leq M$ *für alle* $x \in [a,b] \Rightarrow m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$,
(2) $\int_a^b f(x)dx \leq (b-a)\sup_{x\in[a,b]} f(x)$.

BEWEIS. Folgt aus der Monotonie, Satz V.3.1 □

V.3.3. SATZ (**Dreiecksungleichung für Integrale**). *Sei $f \in \mathcal{R}([a,b])$. Dann gilt*

$$\Big|\int_a^b f(x)dx\Big| \leq \int_a^b |f(x)|dx\,.$$

BEWEIS. Nach Satz V.2.11 ist $|f|$ integrierbar und die Monotonie und Linearität angewendet auf $\pm f \leq |f|$ liefert $\pm\int_a^b f(x)dx \leq \int_a^b |f(x)|dx$ und somit die Behauptung. □

V.3.4. SATZ. *Seien $f \in \mathcal{R}([a,b])$ und $g\colon [a,b] \to \mathbb{R}$ mit $f(x) = g(x)$ für alle bis auf endlich viele $x \in [a,b]$. Dann ist auch $g \in \mathcal{R}([a,b])$ und $\int_a^b g(x)dx = \int_a^b f(x)dx$.*

Dieser Satz besagt, daß Riemann-integrierbare Funktionen integrierbar bleiben und sich das Integral nicht verändert, wenn die Funktionen an endlich vielen Stellen verändert werden. Endlich viele Punkte fallen für das Riemann-Integral also nicht ins Gewicht.

BEWEIS. Betrachte die Differenz $h(x) = f(x) - g(x)$. Also existieren endlich viele Punkte $\{x_1, \ldots, x_n\} \subset [a,b]$ und Werte $h_1, \ldots, h_n \in \mathbb{R}$ mit

$$h(x) = \begin{cases} h_i, & \text{falls } x = x_i, \\ 0, & \text{sonst}\,. \end{cases}$$

Durch geschickte Wahl der Zerlegungspunkte nahe genug um x_i herum in Abhängigkeit von der Höhe h_i finden wir eine Folge von Zerlegungen $(Z_k) \subset \mathcal{Z}_{[a,b]}$ mit

$$-\frac{1}{k} \le U(h, Z_k) \le O(h, Z_k) \le \frac{1}{k} \quad \text{für alle } k \in \mathbb{N}.$$

Also ist $h \in \mathcal{R}([a, b])$ und $\int_a^b h(x)dx$. Aufgrund der Linearität folgt dann die Behauptung. □

V.3.5. Satz (**Mittelwertsatz**). *Es seien $f, g \in C^0([a, b])$ mit $g(x) \ge 0$ für alle $x \in [a, b]$. Dann existiert ein $\xi \in [a, b]$ mit*

$$\int_a^b f(x)g(x)dx = f(\xi)\int_a^b g(x)dx.$$

Spezialfall: $\int_a^b f(x)dx = f(\xi) \cdot (b - a)$ für $g(x) = 1$ für alle $x \in [a, b]$.

Beweis. Seien $m = \min_{[a,b]} f$ und $M = \max_{[a,b]} f$. Minimum und Maximum werden angenommen, da f stetig ist. Aus der Nichtnegativität von g folgt nach der Monotonie

$$m\int_a^b g(x)dx \le \int_a^b f(x)g(x)dx \le M\int_a^b g(x)dx.$$

Also existiert ein $\mu \in [m, M]$ mit $\mu \cdot \int_a^b g(x)dx = \int_a^b f(x)g(x)dx$. Da f stetig ist, liefert der Zwischenwertsatz mindestens ein $\xi \in [a, b]$ mit $f(\xi) = \mu$ und somit die Behauptung. □

Der folgende Satz besagt, daß für integrierbare Funktionen der Mittelwert sowohl durch das normierte Integral als auch durch einen Grenzwert von Mittelwerten über die gleichmäßig verteilten Zwischenpunkte

$$x_i = a + i \cdot \frac{b-a}{n}, \quad i = 0, \dots, n$$

berechnet werden kann. Definiere für $f: [a, b] \to \mathbb{R}$ den diskreten Mittelwert $\mu_n(f) := \frac{1}{n}\sum_{i=1}^n f(x_i)$.

V.3.6. Satz. *Falls $f \in \mathcal{R}([a, b])$ so existiert der Grenzwert $\mu(f) = \lim_{n\to\infty} \mu_n(f)$ und ist gleich dem Mittelwert*

$$\mu(f) = \frac{1}{b-a}\int_a^b f(x)dx.$$

Beweis. Mit der gleichmäßigen Zerlegungsfolge $z_i = a + \frac{i}{n}(b - a)$ und Zwischenpunkte $\xi_i^{(n)} = z_i$ gilt $\mu_n(f)\cdot(b-a) = S(f, Z_n, \xi^{(n)})$. Also folgt die Behauptung unmittelbar aus der Voraussetzung der Integrierbarkeit. □

V.4. Hauptsatz der Differential- und Integralrechnung und Integrationsmethoden

Wir betrachten eine Riemann-integrierbare Funktion $f: [a, b] \to \mathbb{R}$ und definieren aufgrund von Satz V.2.7 die Integralfunktion $F: [a, b] \to \mathbb{R}$ durch

$$F(x) = \int_a^x f(t)dt.$$

Es gelten

V.4.1. Hilfssatz. $F \in C^0([a, b])$.

Beweis. Seien $x, y \in [a, b]$, dann folgt aus Satz V.3.2 und V.3.3 die Abschätzung

$$|F(x) - F(y)| \;=\; \Big| \int_a^x f(t)dt - \int_a^y f(t)dt \Big| \;=\; \Big| \int_x^y f(t)dt \Big| \;\le\; |x-y| \cdot \sup_{[a,b]} |f| \,.$$

Also ist F sogar Lipschitz-stetig. □

V.4.2. Hilfssatz. *Sei f Riemann-integrierbar und stetig in $x_o \in (a, b)$. Dann ist F differenzierbar in x_o und $F'(x_o) = f(x_o)$.*

Beweis. Betrachte

$$\begin{aligned} R(x, x_o) \;&=\; F(x) - F(x_o) - f(x_o) \cdot (x - x_o) \\ &=\; \int_a^x f(t)dt - \int_a^{x_o} f(t)dt - \int_{x_o}^x f(x_o)dt \\ &=\; \int_{x_o}^x \big(f(t) - f(x_o)\big)dt \,. \end{aligned}$$

Da f in x_o stetig ist, existiert zu jedem $\epsilon > 0$ ein $\delta = \delta(\epsilon, f, x_o)$ mit $|f(t) - f(x_o)| < \epsilon$ für alle $|t - x_o| < \delta$. Seien nun $|x - x_o| < \delta$, so folgt $|F(x) - F(x_o)| < \epsilon \cdot |x - x_o|$. Also gilt

$$\lim_{x \to x_o} \frac{R(x, x_o)}{x - x_o} = 0$$

und die Behauptung folgt. □

V.4.3. Theorem. **(a) (Erster Hauptsatz der Differential- und Integralrechnung)** *Sei $f \in C^0([a, b], \mathbb{R})$ und $F(x) = \int_a^x f(t)dt$. Dann ist $F \in C^1([a, b])$ eine Stammfunktion von f.*

(b) (Zweiter Hauptsatz der Differential- und Integralrechnung) *Sei $F\colon [a, b] \to \mathbb{R}$ differenzierbar auf $[a, b]$ und F' integrierbar auf $[a, b]$. Dann gilt*

$$\int_a^b F'(t)dt \;=\; F(x)\Big|_a^b \;:=\; F(b) - F(a) \,.$$

Beweis. <u>(a)</u>: Hilfssatz V.4.2.

<u>(b)</u>: Sei $Z \in \mathcal{Z}_{[a,b]}$, $Z = (z_0, \ldots, z_n)$, und betrachte

$$F(b) - F(a) \;=\; \sum_{i=1}^n F(z_i) - F(z_{i-1}) \,.$$

Nach dem Mittelwertsatz der Differentialrechnung folgt aus der Differenzierbarkeit von F auf $[z_{i-1}, z_i]$, $i = 1, \ldots, n$ die Existenz von Punkten $\xi_i \in (z_{i-1}, z_i)$ so daß $F(z_i) - F(z_{i-1}) = F'(\xi_i) \cdot (z_i - z_{i-1})$, also

$$F(b) - F(a) \;=\; \sum_{i=1}^n (z_i - z_{i-1}) F'(\xi_i) \;=\; S(F', Z, \xi) \,.$$

Sei nun $(Z_n) \subset \mathcal{Z}_{[a,b]}$ eine Zerlegungsfolge mit $L(Z_n) \to 0$ und $\xi^{(n)}$ gemäß dem Mittelwertsatz gewählte Zwischenpunkte. Dann folgt aus der Riemann-Integrierbarkeit von F' nach Satz V.1.15

$$F(b) - F(a) \;=\; S(F', Z_n, \xi^{(n)}) \;\to\; \int_a^b F'(t)dt \quad \text{für } \; n \to \infty,$$

und somit die Behauptung. □

V.4.4. Beispiel. (1) Beispiele für Stammfunktionen als unbestimmte Integrale:

- $\int x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1} + \text{const}$ für $\alpha \in \mathbb{R} \setminus \{-1\}$,
- $\int \frac{1}{x} dx = \ln |x| + \text{const}$,

- $\int \frac{dx}{1+x^2} = \arctan + \text{const}\,.$

(2) Partialbruch-Zerlegung: $\frac{1}{1-x^2} = \frac{1}{(1-x)(1+x)} = \frac{A}{1-x} + \frac{B}{1+x}$, $A = B = \frac{1}{2}$, also

$$\begin{aligned}\int \frac{dx}{1-x^2} &= \frac{1}{2}\int \frac{dx}{1-x} + \frac{1}{2}\int \frac{dx}{1+x} \\ &= -\frac{1}{2}\ln|1-x| + \frac{1}{2}\ln|1+x| + \text{const} = \ln\sqrt{\left|\frac{1+x}{1-x}\right|} + \text{const}\,.\end{aligned}$$

V.4.5. Satz. **(Partielle Integration)**

$$f, g \in C^1([a,b]) \Rightarrow \int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx\,.$$

Beweis. Produktregel $(fg)' = f'g + fg'$ und Hauptsatz V.4.3. □

V.4.6. Satz. **(Substitutionsregel für unbestimmte Integrale)** *Seien I, J Intervalle in $\mathbb{R}$ und $f\colon I \to \mathbb{R}$ stetig, $g\colon J \to \mathbb{R}$ stetig differenzierbar mit $g(J) \subset I$. Dann gilt für $F_0(x) = \int f(x)dx$ und $F_1(x) = \int f(g(t))g'(t)dt$*

$$F_0 \circ g = F_1 + \text{const}\,.$$

Beweis. Kettenregel für Differentiation

$$\frac{d}{dt}\big(F_0(g(t)\big) = F_o'\big(g(t)\big) \cdot g'(t) = f(g(t)) \cdot g'(t) = \frac{d}{dt}F_1(t)$$

und Hauptsatz V.4.3 □

Anwendung:

(1) Wähle Substitutionsfunktion, $x = g(t)$,
(2) berechne $g'(t) = \frac{dx}{dt}$, "formal": $dx = g'(t)dt$, ersetze (substituiere) $\int f(x)dx = \int f(g(t))g'(t)dt$,
(3) berechne Stammfunktion $F_1(t) = \int f(g(t))g'(t)dt$,
(4) berechne $t = g^{-1}(x)$, geht nur wenn $g'(t) \neq 0$, $F_0(x) = F_1(g^{-1}(x))$.

V.4.7. Beispiel.
$\int \frac{dx}{x\sqrt{1+x}}$, $x > 0$

(1) wähle $x = g(t) = t^2 - 1$, $t > 1$,
(2) $g'(t) = 2t$, $dx = 2tdt$,
(3) $\int \frac{2tdt}{(t^2-1)t} = 2\int \frac{dt}{(t+1)(t-1)} = \int \left(\frac{1}{t-1} - \frac{1}{t+1}\right) dt$
$\overset{t\geq 1}{=} \ln(t-1) - \ln(t+1) + \text{const} = \ln\frac{t-1}{t+1} + \text{const}$,
(4) $t = \sqrt{x+1}$, $x > 0$, also $\int \frac{dx}{x\sqrt{1+x}} = \ln\frac{\sqrt{x+1}-1}{\sqrt{x+1}+1} + \text{const}$, für $x > 0$.

V.4.8. Satz. **(Substitutionsregel für bestimmte Integrale)** *Sei $f\colon [a,b] \to \mathbb{R}$ stetig, $g\colon [\alpha,\beta] \to \mathbb{R}$ stetig diff'bar mit $g([\alpha,\beta]) = [a,b]$, $g(\alpha) = a$, $g(\beta) = b$. Dann gilt*

$$\int_a^b f(x)dx = \int_\alpha^\beta f\big(g(t)\big)g'(t)dt\,.$$

Beweis. Betrachte die Stammfunktion F von f, dann gilt nach der Kettenregel $F\big(g(t)\big)' = f\big(g(t)\big)g'(t)$ und somit nach dem Hauptsatz V.4.3

$$F(b) - F(a) = F\big(g(\beta)\big) - F\big(g(\alpha)\big)\,.$$

□

V.5. Unbeschränkte Intervalle und uneigentliche Integrale

V.5.1. Definition.
Sei $f\colon [a,\infty) \to \mathbb{R}$ mit $f_{|[a,b]} \in \mathcal{R}([a,b])$ für alle $a < b < \infty$. Dann definieren wir das **uneigentliche Integral**

$$\int_a^\infty f(x)dx := \lim_{b\to\infty} \int_a^b f(x)dx\,, \quad \text{falls der Grenzwert existiert,}$$

und analog $$\int_{-\infty}^b f(x)dx := \lim_{a\to-\infty} \int_a^b f(x)dx\,, \quad \text{falls der Grenzwert existiert,}$$

in dem Fall von $f\colon (\infty, b] \to \mathbb{R}$. Für $f\colon (\infty,\infty) \to \mathbb{R}$ definieren wir

$$\int_{-\infty}^\infty f(x)dx := \int_{-\infty}^a f(x)dx + \int_a^\infty f(x)dx\,, \quad \text{falls beide Grenzwerte existieren}$$

mit $a \in \mathbb{R}$ beliebig.

V.5.2. Beispiel. (1) $\int_{-\infty}^0 e^x dx = \lim_{t\to-\infty} \int_t^0 e^x dx = \lim_{t\to-\infty} e^x\big|_t^0$
$= 1 - \lim_{t\to-\infty} e^t = 1.$

(2)

$$\int_1^\infty x^\alpha dx = \begin{cases} \lim_{t\to\infty} \ln x\big|_1^t & , \alpha = -1 \\ \lim_{t\to\infty} \frac{x^{\alpha+1}}{\alpha+1}\big|_1^t & , \alpha \neq -1 \end{cases}$$
$$= \begin{cases} +\infty & , \alpha = -1, \\ -\frac{1}{\alpha+1} & , \alpha < -1, , \\ +\infty & , \alpha > -1 \end{cases}$$

d.h. $\int_1^\infty x^\alpha dx$ existiert/konvergiert genau für $\alpha < -1$.

V.5.3. Definition.
Sei $f\colon [a,b) \to \mathbb{R}$ unbeschränkt und $f_{|[a,t]} \in \mathcal{R}([a,t])$ für alle $a < t < b$. Dann definieren wir

$$\int_a^b f(x)dx := \lim_{t\uparrow b} \int_a^t f(x)dx\,, \quad \text{falls der Grenzwert existiert}\,,$$

und analog für $f\colon (a,b] \to \mathbb{R}$, $f\colon (a,b) \to \mathbb{R}$.

V.5.4. Beispiel. (1)

$$\int_0^1 x^\alpha dx = \begin{cases} \lim_{t\downarrow 0} \ln x\big|_t^1 & , \alpha = -1 \\ \lim_{t\downarrow 0} \frac{x^{\alpha+1}}{\alpha+1}\big|_t^1 & , \alpha \neq -1 \end{cases}$$
$$= \begin{cases} +\infty & , \alpha = -1, \\ \frac{1}{\alpha+1} & , \alpha > -1, , \\ +\infty & , \alpha < -1 \end{cases}$$

d.h. $\int_0^1 x^\alpha dx$ existiert/konvergiert genau für $\alpha > -1$.

(2) $\int_{-\infty}^\infty \frac{dx}{1+x^2} = \lim_{t\to-\infty} \arctan x\big|_t^0 + \lim_{t\to\infty} \arctan x\big|_0^t = \pi.$

V.5.5. Satz. **(Cauchy-Konvergenzkriterium)** *Sei $f\colon [a,\infty) \to \mathbb{R}$ Riemann-integrierbar auf allen kompakten Teilintervallen. Dann gilt*

$$\int_a^\infty f(x)dx \textit{ existiert} \Leftrightarrow \forall\, \epsilon > 0\; \exists\, x_o > a \textit{ s.d. } \left|\int_t^{t'} f(x)dx\right| < \epsilon \textit{ f.a. } t, t' > x_o\,.$$

Beweis. "$\Rightarrow$": Sei $\gamma = \int_a^\infty f(x)dx$ und $\epsilon > 0$. Dann existiert ein $x_o > a$ in Abhängigkeit von ϵ, so daß

$$\begin{aligned}\left|\int_t^{t'} f(x)dx\right| &= \left|\int_a^{t'} f(x)dx - \int_a^t f(x)dx\right| \\ &\leq \left|\int_a^{t'} f(x)dx - \gamma\right| + \left|\gamma - \int_a^t f(x)dx\right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

für alle $t, t' \geq x_o$.
"$\Leftarrow$": Es sei $(c_n) \subset \mathbb{R}$ eine Folge mit $\lim_{n\to\infty} c_n = \infty$ und $\gamma_n = \int_a^{c_n} f(x)dx$. Da $|\gamma_n - \gamma_m| = \left|\int_{c_m}^{c_n} f(x)dx\right| < \epsilon$ für alle $m, n \geq n_o$ mit $n_o \in \mathbb{N}$ groß genug so daß $c_n \geq x_o$ für alle $n \geq n_o$, folgt daß $(\gamma_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist und somit existiert der Grenzwert $\lim_{n\to\infty} \gamma_n$.
Wir behaupten nun daß für jede solche Folge mit $c_n \to \infty$ sich derselbe Grenzwert $\lim \gamma_n$ ergibt. Dies sieht man wie folgt. Seien $(c_n), (c'_n)$ zwei solche Folgen und wir bilden dann die Reißverschlußfolge $(\bar{c}_n) = (c_1, c'_1, c_2, c'_2, \ldots)$. Mit demselben Argument wie vorher für (c_n) ergibt sich wieder daß $(\bar{\gamma}_n)$ konvergiert. Allerdings sind nun (γ_n) und (γ'_n) darin als Teilfolgen enthalten und müssen folglich denselben Grenzwert haben. Also existiert $\lim_{t\to\infty} \int_a^t f(x)dx$. □

V.5.6. Satz. *Sei $f\colon [a,\infty) \to \mathbb{R}$ mit $f \in \mathcal{R}([a,b])$ für alle $a < b$ und $\int_a^\infty |f(x)|dx$ existiere. Dann existiert auch $\int_a^\infty f(x)dx$ und*

$$\left|\int_a^\infty f(x)dx\right| \leq \int_a^\infty |f(x)|dx\,.$$

Beweis. Zu jedem $\epsilon > 0$ existiert nach Voraussetzung wegen Satz V.5.5 ein $x_o(\epsilon) > a$ so daß

$$\left|\int_t^{t'} f(x)dx\right| \leq \int_t^{t'} |f(x)|dx < \epsilon \quad \text{für alle } t' > t \geq x_o\,.$$

Also folgt die Behauptung wiederum mittels Satz V.5.5. □

V.5.7. Satz. **(Majoranten- und Minorantenkriterium)** *Es seien $f, g\colon [a,\infty) \to \mathbb{R}$. Dann gilt: Falls*

(a) *$|f(x)| \leq g(x)$ für alle $x \geq a$ und $\int_a^\infty g(x)dx$ existiert, dann konvergiert auch $\int_a^\infty |f(x)|dx$, und falls*
(b) *$0 \leq g(x) \leq f(x)$ für alle $x \geq a$ und $\int_a^\infty g(x)dx$ divergiert, so divergiert auch $\int_a^\infty f(x)dx$.*

Beweis. Es genügt der Beweis von (a), (b) folgt dannunmittelbar aus (a). Nach Voraussetzung an g und mittels Satz V.5.5 existiert für alle $\epsilon > 0$ ein $x_o > a$ so daß für alle $t' > t \geq x_o$

$$\int_t^{t'} |f(x)|dx \leq \int_t^{t'} g(x)dx < \epsilon\,.$$

Also ist wiederum mittels Satz V.5.5 $|f|$ auch endlich integrabel über $[a,\infty)$. □

V.5.8. Beispiel.
Wir behaupten, daß das Integral $\int_{-\infty}^\infty e^{-x^2}dx$ konvergiert. Es genügt offenbar zu

zeigen, daß $\int_1^\infty e^{-x^2} dx$ konvergiert. Wir betrachten folgende Abschätzung.

$$\begin{aligned} y \geq 0 \ &\Rightarrow \ e^y = 1 + y + \frac{1}{2}y^2 + \ldots \geq y \\ &\Rightarrow \ y \geq \ln y \quad \text{f.a. } \ y > 0 \quad \Rightarrow \ x^2 \geq \ln x^2 \text{ f.a. } \ x \neq 0 \\ &\Rightarrow \ e^{-x^2} \ \leq \ e^{-\ln x^2} \ = \ \frac{1}{x^2} \, . \end{aligned}$$

Nun existiert das Integral $\int_1^\infty \frac{dx}{x^2}$ und somit liefert Satz V.5.7 die Konvergenz von $\int_1^\infty e^{-x^2}$, ohne daß wir bislang den Wert des Integrals berechnen können! Dieses wird erst mit mehrdimensionalen Integrations-Methoden möglich sein. Wir werden zeigen, daß

$$\int_{-\infty}^{\infty} e^{-x^2} dx \ = \ \sqrt{\pi} \, .$$

V.6. Ergänzungen und Anwendungen

V.6.1. Funktionenfolgen.

V.6.1. Definition.

Es sei $X \subset \mathbb{C}$ ein gegebener Definitionsbereich für Funktionen $f_n, f\colon X \to \mathbb{C}$, $n \in \mathbb{N}$. Dann definieren wir die **Supremumsnorm auf** X durch

$$\|f\|_X := \sup_{z \in X} \big|f(z)\big| \, .$$

Wir sagen, die Folge (f_n) **konvergiert punktweise gegen** f genau dann, wenn

$$f_n(z) \to f(z) \text{ für } \ n \to \infty, \quad \text{f.a. } \ z \in X \, ,$$

und wir sagen, die Folge (f_n) **konvergiert gleichmäßig gegen** f genau dann, wenn

$$\|f_n - f\|_X \to 0 \quad \text{für } \ n \to \infty \, .$$

Es ist leicht zu sehen, daß gleichmäßige Konvergenz immer auch punktweise Konvergenz impliziert. Dagegen ist die Umkehrung falsch wie folgendes Beispiel zeigt:

V.6.2. Beispiel.

Es sei $X = (-1, 1) \subset \mathbb{R}$ und $f_n(x) = x^n$. Da $|x| < 1$, konvergiert $f_n \to 0$ punktweise. Allerdings gilt

$$\|f_n\|_{(-1,1)} \ = \ \sup_{-1<x<1} |x|^n \ = \ 1 \quad \text{f.a. } \ n \in \mathbb{N}$$

und somit konvergiert (f_n) nicht gleichmäßig. Allerdings, wenn wir f_n auf ein beliebiges kompaktes Intervall $[-1+\epsilon, 1-\epsilon]$, $0 < \epsilon < 1$, einschränken, so konvergiert $(f_{n|[-1+\epsilon,1-\epsilon]})$ gleichmäßig gegen 0.

V.6.3. Beispiel.

Betrache eine Potenzreihe $f(z) = \sum_{k=0}^\infty a_k z^k$ also Funktionenfolge von Polynomfunktionen $f_n(z) = \sum_{k=0}^n a_k z^k$ auf einem Kreisscheibengebiet $X = B_R(0) \subset \mathbb{C}$. Angenommen, (f_n) ist normal konvergent, d.h. $\sum_{k=0}^\infty \|a_k z^k\|_{B_R(0)} = \sum_{k=0}^\infty |a_k| R^k$ konvergiert. Dann gilt $f_n \to f$ gleichmäßig auf $B_R(0)$.

V.6.4. Satz. *Es seien $(f_n), f\colon X \to \mathbb{C}$ wie oben und $(f_n) \subset C^0(X, \mathbb{C})$. Dann ist auch f stetig auf X.*

Beweis. Vgl. III.3.1.
Sei $z_o \in X$. Es gilt f.a. $z \in X$

$$\begin{aligned} |f(z) - f(z_o)| \ &\leq \ |f(z) - f_n(z)| \ + \ |f_n(z) - f_n(z_o)| \ + \ |f_n(z_o) - f(z_o)| \\ &\leq \ 2\|f - f_n\|_X \ + \ |f_n(z) - f_n(z_o)| \, . \end{aligned}$$

Sei $\epsilon > 0$. Wenn $f_n \to f$ gleichmäßig, so existiert ein $n_o(\epsilon/3)$ so daß

$$\|f - f_n\| < \frac{\epsilon}{3} \quad \text{f.a.} \;\; n \geq n_o\,.$$

Da f_{n_o} in z_o stetig ist, existiert ein $\delta = \delta(f_{n_o}, z_o)$ so daß

$$|f_{n_o}(z) - f_{n_o}(z_o)| < \frac{\epsilon}{3} \quad \text{f.a.} \;\; |z - z_o| < \delta\,.$$

Für dieses δ gilt dann $|f(z) - f(z_o)| < \epsilon$ für alle $|z - z_o| < \delta$. □

V.6.5. Satz. *Seien nun $(f_n) \subset \mathcal{R}([a,b])$, $f\colon [a,b] \to \mathbb{R}$ mit $f_n \to f$ gleichmäßig. Dann gilt auch $f \in \mathcal{R}([a,b])$ und*

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

Beweis. Es gilt für jedes Teilintervall $I \subset [a,b]$,

$$\begin{aligned} \sup_I f \;&=\; \sup_I (f - f_n + f_n) \;\leq\; \sup_I (f - f_n) + \sup_I f_n \\ &\leq\; \sup_{[a,b]} |f - f_n| + \sup_I f_n, \end{aligned}$$

und wegen $-\inf f = \sup(-f)$ haben wir

$$\begin{aligned} -\inf_I f \;&\leq\; \sup_I(-f + f_n) \;+\; \sup_I(-f_n) \\ &\leq\; \sup_{[a,b]} |f - f_n| \;-\; \inf_I f_n\,. \end{aligned}$$

Also gilt für alle $Z \in \mathcal{Z}_{[a,b]}$

$$\begin{aligned} O(f,Z) - U(f,Z) \;&=\; \sum_{i=1}^k \Big(\sup_{[z_{i-1},z_i]} f - \inf_{[z_{i-1},z_i]} f\Big) \cdot (z_i - z_{i-1}) \\ &\leq\; 2\|f - f_n\|_{[a,b]}(b - a) \;+\; O(f_n, Z) \;-\; U(f_n, Z)\,. \end{aligned}$$

Sei nun $\epsilon > 0$ und $n_o(\epsilon) \in \mathbb{N}$ so daß $\|f - f_{n_o}\| < \frac{\epsilon}{4(b-a)}$. Dann folgt aus der Annahme $f_{n_o} \in \mathcal{R}([a,b])$, daß eine Zerlegung $Z \in \mathcal{Z}_{[a,b]}$ existiert mit $O(f_{n_o}, Z) - U(f_{n_o}, Z) < \frac{\epsilon}{2}$. Also folgt $O(f,Z) - U(f,Z) < \epsilon$. Somit ist auch f auf $[a,b]$ Riemann-integrierbar. Außerdem folgt aus $\big|O(f,Z) - O(f_n,Z)\big| \leq \|f - f_n\| \cdot (b - a)$ der Grenzwert

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

□

V.6.6. Korollar. *Seien $(f_n) \subset C^1([a,b])$ mit $f_n \to f$ punktweise auf $[a,b]$ und $f_n' \to g$ gleichmäßig auf $[a,b]$. Dann gilt: $f \in C^1([a,b])$, $f' = g \in C^0([a,b])$ und*

$$\frac{d}{dx} \lim_{n\to\infty} f_n \;=\; \lim_{n\to\infty} \frac{d}{dx} f_n\,.$$

Beweis. Aus $(f_n) \subset C^1([a,b])$ folgt $(f_n') \subset C^0([a,b])$. Satz V.6.4 impliziert dann $g \in C^0([a,b])$ und Satz V.6.5 den Grenzwert für alle $x \in [a,b]$:

$$\int_a^x f_n'(t)dt \;\to\; \int_a^x g(t)dt \quad \text{für} \;\; n \to \infty\,.$$

Aus dem Hauptsatz der Differential- und Integralrechnung folgt dann

$$\lim_{n\to\infty} \big(f_n(x) - f_n(a)\big) = \int_a^x g(t)dt\,.$$

Andererseits gilt $f_n(x) \to f(x)$ und $f_n(a) \to f(a)$. Außerdem besagt der Hauptsatz ebenfalls, daß $h(x) = \int_a^x g(t)dt$ differenzierbar ist mit $h'(x) = g(x)$. Also folgt

$$f(x) - f(a) \;=\; h(x) \in C^1([a,b])$$

und somit $f \in C^1([a,b])$, $f' = g = \lim f_n$. □

V.6.7. Beispiel.
Wir betrachten die Funktion $f(x) = \arctan x$, $x \in \mathbb{R}$. Hierfür gilt

$$f'(x) = \frac{1}{1+x^2}\,.$$

Mittels der geometrischen Reihe folgt

$$f'(x) \;=\; \sum_{k=0}^{\infty}(-x^2)^k \quad \text{f.a. } |x| < 1\,.$$

Wir betrachten die Potenzreihe $f_n(x) = \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} x^{2k+1}$ mit Konvergenzradius 1 und erhalten gleichmäßige Konvergenz der Funktionenfolge (f_n) eingeschränkt auf jedes abgeschlossene Intervall $[-1+\epsilon, 1-\epsilon]$ mit $0 < \epsilon < 1$. Aus Korollar V.6.6 folgt somit

$$\arctan x \;=\; \sum_{n=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} \;+\; \text{const} \quad \text{f.a. } -1 < x < 1\,.$$

Da $\arctan 0 = 0$, folgt const $= 0$. Der Abelsche Grenzwertsatz zusammen mit der Konvergenz von $\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}$ nach dem Leibniz-Kriterium liefern dann

$$\arctan x \;=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \quad \text{f.a. } |x| \le 1,$$

insbesondere

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \;=\; 1 - \frac{1}{3} + \frac{1}{5} \mp \ldots \;=\; \frac{\pi}{4}\,.$$

V.6.2. Integraltest für Reihen.

V.6.8. Satz. *Es sei $f\colon [m,\infty) \to \mathbb{R}$ eine monoton fallende Funktion, $m \in \mathbb{N}$, mit $f(x) \ge 0$ für alle $x \ge m$. Dann gilt*

$$\int_m^{\infty} f(x)dx \text{ existiert} \;\Leftrightarrow\; \sum_{n=m}^{\infty} f(n) \text{ konvergiert}.$$

Beweis. ”⇒”: Aufgrund der Monotonie können wir auf jedem ganzen Teilintervall

$$\int_n^{n+1} f(x)dx \;\ge\; \inf_{[n,n+1]} f(x) \;=\; f(n+1)$$

abschätzen und somit gilt

$$\sum_{n=m+1}^{M+1} f(n) \;\le\; \sum_{n=m}^{M} \int_n^{n+1} f(x)dx \;\le\; \int_m^{M+1} f(x)dx\,.$$

Da $f(x) \ge 0$ für alle $x \ge m$ und das uneigentliche Integral nach Voraussetzung existiert, folgt, daß die Reihe $\sum_{n=m}^{\infty} f(n)$ konvergiert und der Wert nach oben beschränkt ist durch $f(m) + \int_m^{\infty} f(x)dx$.

"$\Leftarrow$": Es sei nun $b > m$. Dann gilt aufgrund der Additivität des Integrals

$$\begin{aligned}\int_m^b f(x)dx &= \int_m^{m+1} f(x)dx + \ldots + \int_{[b]-1}^{[b]} f(x)dx + \int_{[b]}^b f(x)dx \\ &\leq f(m) + f(m+1) + \ldots + f([b]-1) + (b-[b]) \cdot f([b]) \\ &\leq \sum_{n=m}^{\infty} f(n) < \infty .\end{aligned}$$

Die erste Abschätzung folgt aus dem monotonen Abfall von f. Also ist insgesamt $F(b) = \int_m^b f(x)dx$ eine monoton wachsende und oberhalb beschränkte Funktion. Daher existiert der Grenzwert $\lim_{b\to\infty} F(b)$ und dieser ist oberhalb beschränkt durch $\sum_{n=m}^{\infty} f(n)$. □

V.6.9. Beispiel.
Wir betrachten $f(x) = \frac{1}{x^2}$ und schätzen ab:

$$\sum_{n=2}^{\infty} \frac{1}{n^2} \leq \int_1^{\infty} \frac{dx}{x^2} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} .$$

Aus der Integralberechnung zu 1 folgt somit

$$1 \leq \sum_{n=1}^{\infty} \frac{1}{n^2} \leq 2$$

Mit nicht-trivialen Methoden, die uns hier noch nicht zur Verfügung stehen, lässt sich dieser Wert der Zeta-Funktion berechnen zu

$$\zeta(2) = \frac{\pi^2}{6} .$$

V.7. ∗ Fourier-Reihen

Wir betrachte komplexwertige Funktionen in einer reellen Variable.

V.7.1. Definition.
Eine Funktion $f\colon \mathbb{R} \to \mathbb{C}$ heißt **periodisch** mit Periode τ, wenn

$$f(x+\tau) = f(x) \quad \text{f.a. } x \in \mathbb{R}, .$$

Falls f τ-periodisch ist s ist f auch $k\tau$-periodisch für alle $k \in \mathbb{Z}$.

V.7.2. Hilfssatz. *Ist $f\colon \mathbb{R} \to \mathbb{C}$ stetig, periodisch und nicht konstant, so besitzt f eine kleinste positive Periode,*

Beweis. Angenommen, das ist falsch. Dann finden wir eine Folge $\tau_n \to 0$, $\tau_n > 0$, mit

$$f(x+\tau_n) = f(x) \quad \text{f.a. } x \in \mathbb{R},\, n \in \mathbb{N} .$$

Da $\tau_n \to 0$, existiert für jedes $y \in \mathbb{R}$ eine Folge $(m_n) \in \mathbb{Z}^{\mathbb{N}}$ mit $m_n\tau_n \to y$. Also folgt aus der Stetigkeit

$$f(y) = \lim_{n\to\infty} f(m_n\tau_n) = f(0) .$$

Somit muss f konstant sein. □

V.7.3. Beispiel.
$e_k(x) := e^{ikx}$ ist 2π-periodisch mit minimaler Periode $\frac{2\pi}{k}$, falls $k \in \mathbb{Z} \setminus \{0\}$. Wir haben

$$e_k(x) = \cos kx + i \sin kx,$$

$\cos kx$ und $\sin kx$ sind ebenfalls $\frac{2\pi}{k}$-periodisch.

V.7.4. DEFINITION.
Seien $n \in \mathbb{N}$ und $c_k = \alpha_k + i\beta_k \in \mathbb{C}$, $k \in \{-n \ldots, 0, \ldots, n\} \subset \mathbb{Z}$. Dann heißt

$$T(x) \;=\; \sum_{k=-n}^{n} c_k e^{ikx}$$

ein **trigonometrisches Polynom**.
Nach der Umformung

$$\begin{aligned}
\sum_{|k|\le n} c_k e^{ikx} \;&=\; \sum_{k=-n}^{n} (\alpha_k + i\beta_k)(\cos kx + i \sin kx) \\
&=\; (\alpha_0 + i\beta_0) \;+\; \sum_{k=1}^{n} \big[(\alpha_k + i\beta_k)(\cos kx + i\sin kx) \\
&\qquad\qquad + (\alpha_{-k} + i\beta_{-k})(\cos kx - i \sin kx)\big] \\
&=\; \sum_{k=1}^{n} (\alpha_k + \alpha_{-k}) \cos kx + (\beta_{-k} - \beta_k) \sin kx \\
&\quad + i\big[\beta_0 + \sum_{k=1}^{n} (\beta_k + \beta_{-k}) \cos kx + (\alpha_{-k} - \alpha_k) \sin kx\big] \,,
\end{aligned}$$

sehen wir, daß wir trigonometrische Polynome auch allgemein schreiben können in der Form

$$T(x) \;=\; a_0 + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx, \quad a_0, a_k, b_k \in \mathbb{C},\; k \in \mathbb{N}\,.$$

Da Integration eine lineare Operation ist, können wir diese problemlos auch auf komplexwertige Funktionen erweitern. Sei $f\colon [a,b] \to \mathbb{C}$ eine beschränkte Funktion, d.h. $|f(x)| \le c$ f.a. $x \in [a,b]$. Dann heißt f **Riemann-integrierbar** , wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ Riemann-integrierbar sind und

$$\int_a^b f(x)dx \;:=\; \int_a^b \operatorname{Re} f(x)dx \;+ i \int_a^b \operatorname{Im} f(x)dx\,.$$

Wir haben für die trigonometrischen Funktionen e_k, $\cos kx$ und $\sin kx$ folgende **Orthogonalitätsrelationen**:

$$\begin{aligned}
\langle e_k, e_l\rangle \;:=&\; \frac{1}{2\pi}\int_0^{2\pi} e_k(x) e_{-l}(x) dx \;=\; \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-l)x} dx \\
=&\; \begin{cases} \frac{1}{2\pi(k-l)} e^{i(k-l)x}\big|_0^{2\pi} & , k \ne l, \\ 1 & , k = l, \end{cases} \\
=&\; \begin{cases} 0, & k \ne l, \\ 1, & k = l \end{cases} \;=:\; \delta_{k,l}\,.
\end{aligned}$$

(4)

$$\begin{aligned}
\langle \sin nx, \sin mx\rangle \;&=\; \frac{1}{2\pi}\int_0^{2\pi} \sin nx \sin mx \, dx \;=\; \frac{1}{2}\delta_{n,m} \\
&=\; \frac{1}{2\pi}\int_0^{2\pi} \cos nx \cos mx \, dx \;=\; \langle \cos nx, \cos mx\rangle, \\
\langle \sin nx, \cos mx\rangle \;&=\; \frac{1}{2\pi}\int_0^{2\pi} \sin nx \cos mx \, dx \;=\; 0 \quad \text{f.a. } \; n, m, \in \mathbb{Z}\,.
\end{aligned}$$

Sei nun $f(x) = \sum_{k=-n}^{n} c_k e_k(x)$, dann folgt aus den Orthogonalitätsrelationen für e_k,

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e_{-k}(x)\,dx\,.$$

Entsprechend erhalten wir für

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{k=0}^{n} a_k \cos kx + b_k \sin kx, \\ a_0 &= \frac{1}{\pi}\int_0^{2\pi} f(x)\,dx, \\ a_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,dx, \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx\,dx\,. \end{aligned} \tag{5}$$

V.7.5. Definition.

Sei $f\colon \mathbb{R} \to \mathbb{C}$ eine 2π-periodische Funktion, so daß das Produkt $f \cdot e_k \in \mathcal{R}([0, 2\pi])$ für alle $k \in \mathbb{Z}$ Riemann-integrierbar ist. Dann heißt

$$\hat{f}(k) := \frac{1}{2\pi}\int_o^{2\pi} f(x)e_{-k}(x)\,dx$$

der k-te **Fourier-Koeffizient**, $s_n(f) = \sum_{k=-n}^{n} \hat{f}(k)e_k$ heißt die k-te **Fourier-(Partial-)Summe** von f, sowie $\sum_{k=-\infty}^{\infty} \hat{f}(k)e_k := (s_n)_{n\in\mathbb{N}_0}$ die **Fourier-Reihe** von f.

Betrachten wir eine solche Fourier-Reihe $(s_n)_{n\in\mathbb{N}_0}$ einer Funktion f, so stellt sich die Frage, für welche f diese Funktionenreihe gleichmäßig auf $[0, 2\pi]$ konvergiert. Man beachte, daß s_n für alle $n \in \mathbb{N}_0$ 2π-periodisch ist, selbst wenn f nicht die Periode 2π hat. Dies schränkt also bereits die Klasse der Funktionen mit gleichmäßig konvergenter Fourier-Reihe ein.

V.7.6. Definition.

Das trigonometrishce Polynom

$$\sigma_n(f) = \frac{1}{n}\sum_{k=0}^{n-1} s_k(f)$$

als Mittelwert der ersten n Fourier-Polynomen heißt das n-te **Fejér-Polynom** von f.

Es gilt offenbar: Falls $s_n(f)(x) \to f(x)$ für $n \to \infty$ so auch $\sigma_n(f)(x) \to f(x)$. Ohne Beweis zitieren wir folgenden

V.7.7. Satz (von Fejér). *Sei $f\colon \mathbb{R} \to \mathbb{C}$ 2π-periodisch und stückweise stetig, so gilt:*

(a) *Für alle $x \in \mathbb{R}$ gilt $\sigma_n(f)(x) \to \frac{f(x-)+f(x+)}{2}$, insbesondere*
(b) *falls f in x stetig ist, so gilt $\sigma_n(f)(x) \to f(x)$.*
(c) *Ist f stetig auf ganz $\mathbb{R}$, so konvergiert $\sigma_n(f) \to f$ gleichmäßig auf $\mathbb{R}$.*

Aber man kann auch folgendes zeigen:

V.7.8. Satz. *Für alle $x \in \mathbb{R}$, $\epsilon > 0$ und $f\colon \mathbb{R} \to \mathbb{C}$ existiert eine stetige Funktion $\tilde{f}\colon \mathbb{R} \to \mathbb{C}$ mit $\|f - \tilde{f}\|_\infty < \epsilon$ und $\sup_{n\in\mathbb{N}} \left|s_n(\tilde{f})(x)\right| = \infty$.*

Das bedeutet, beliebig nahe an f finden wir immer stetige Funktionen, deren Fourier-Reihen in x divergieren. Also brauchen wir im Gegensatz zu Fejér-Polynomen für Fourier-Polynome stärkere Voraussetzungen, um die Reihe konvergieren zu lassen.

Wir wiederholen folgende

V.7.9. Definition.

Eine Funktion $f\colon [a,b] \to \mathbb{C}$ heißt **stückweise stetig**, wenn es eine Zerlegung $Z = (a = z_0 < z_1 < \ldots < z_l = b) \in \mathcal{Z}_{[a,b]}$ gibt, so daß $f_{|[z_{i-1},z_i]}$ für alle $i = 1,\ldots,l$ stetig ist. Wir definieren dann die Notation

$$f(z_i+) := \lim_{t\downarrow z_i} f(t), \quad f(z_i-) := \lim_{t\uparrow z_i} f(t)\,.$$

f heißt außerdem **stückweise differenzierbar**, wenn f stückweise stetig ist bzgl. einer Zerlegung $Z \in \mathcal{Z}_{[a,b]}$ so daß $f_{|(z_{i-1},z_i)}$ differenzierbar ist, so daß

$$f'(z_i+) = \lim_{t\downarrow 0} \frac{f(t+z_i) - f(z_i+)}{t} \quad \text{existiert und ebenso auch}$$

$$f'(z_i-) = \lim_{t\downarrow 0} \frac{f(z_i-) - f(z_i-t)}{t}\,.$$

Das Hauptresultat ist

V.7.10. Theorem (**Dirichlet**). *Sei $f\colon \mathbb{R} \to \mathbb{C}$ 2π-periodisch und stückweise stetig und $f'(x+)$ und $f'(x-)$ existieren in $x \in \mathbb{R}$. Dann konvergiert die Fourier-Reihe $s_n(f)$ in x gegen $\frac{f(x+)+f(x-)}{2}$.*

V.7.11. Beispiel.

Wir betrachten die stückweise konstante Funktion

$$f(x) = \begin{cases} 1, & 0 \le x \le \pi, \\ 0, & \text{sonst.} \end{cases}$$

Wir setzen dann f 2πperiodisch fort zu

$$\tilde{f}(x) = \begin{cases} 1, & 2k\pi \le x \le (2k+1)\pi, \\ 0, & (2k+1)\pi < x < (2k+2)\pi, \ k \in \mathbb{Z}\,. \end{cases}$$

Also berechnen wir gemäß (5)

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_{\pi}^{\pi} f(x)\,dx = 1, \\ a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx = 0, \text{ wegen ungeradem Integrand,} \\ b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx = \frac{1}{\pi}\int_0^{\pi} \sin kx\,dx \\ &= \frac{1}{k\pi}(-\cos kx)\Big|_0^{\pi} = \begin{cases} 0, & k \text{ gerade,} \\ \frac{2}{k\pi}, & k \text{ ungerade}\,. \end{cases} \end{aligned}$$

Also erhalten wir die Fourier-Reihe von f als

$$F(x) = \frac{1}{2} + \sum_{l=0}^{\infty} \frac{2}{(2l+1)\pi} \sin(2l+1)x\,.$$

Gemäß Theorem V.7.10 konvergiert diese Fourier-Reihe gegen

$$F(x) = \begin{cases} 0, & -\pi < x < 0, \\ 1, & 0 < x < \pi, \\ \frac{1}{2}, & x = 0, \pi\,. \end{cases}$$

Wenn wir zum Beispiel $x = \frac{1}{2}$ einsetzen erhalten wir wegen $\sin(2l+1)\frac{\pi}{2} = (-1)^l$

$$\frac{1}{2} + \sum_{l=0}^{\infty} \frac{2}{(2l+1)\pi}(-1)^l \;=\; 1,$$

$$\text{also} \quad \sum_{l=0}^{\infty} \frac{(-1)^l}{2l+1} \;=\; \frac{\pi}{4}\,.$$

Man vergleiche dies mit der Taylorreihe von $\arctan x$.

Wir wolen nun das Theorem von Dirichlet beweisen.

V.7.12. Definition.

Das trigonometrische Polynom

$$D_n(x) := \sum_{k=-n}^{n} e_k(x) \;=\; \sum_{k=-n}^{n} e^{ikx}$$

heißt der **Dirichlet-Kern n-ten Grades**.

Es gilt

$$\begin{aligned}
\sum_{k=-n}^{n} e^{ikx} &= e^{-inx}\sum_{k=0}^{2n} e^{ikx} \;=\; e^{-inx}\sum_{k=0}^{2n}\left(e^{ix}\right)^k \\
&= e^{-inx}\frac{(1-e^{ix})(1+\ldots+e^{i2nx})}{1-e^{ix}} \\
&= e^{-inx}\frac{1-e^{i(2n+1)x}}{1-e^{ix}} \;=\; \frac{e^{-inx}-e^{i(n+1)x}}{1-e^{ix}} \\
&= \frac{e^{-i(n+\frac{1}{2})x}-e^{i(n+\frac{1}{2})x}}{e^{-i\frac{x}{2}}-e^{i\frac{x}{2}}} \\
&= \frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}}\,.
\end{aligned} \tag{6}$$

Außerdem haben wir

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{\pi} D_n(t)\,dt \;&=\; \frac{1}{2} + \frac{1}{2\pi}\sum_{k=1}^{n}\int_0^{\pi} e^{ikx} + e^{-ikx}\,dx \\
&= \frac{1}{2} + \frac{1}{\pi}\sum_{k=1}^{n}\int_0^{\pi} \cos kx\,dx \\
&= \frac{1}{2}\,.
\end{aligned} \tag{7}$$

Die zentralen Schritte liegen in folgenden Lemmata:

V.7.13. Lemma (**Riemannsches Lemma**). *Sei $F\colon [a,b] \to \mathbb{C}$ stückweise stetig. Dann gilt*

$$\lim_{k\to\infty}\int_a^b F(x)\sin kx\,dx \;=\; 0\,.$$

Beweis. Sei $\epsilon > 0$. Ohne Beschränkung der Allgemeinheit gelte $F(x+) = F(x)$ für alle $x \in [a,b]$. Sei $Z = (a = z_0 < z_1 < \ldots < z_l = b) \in \mathcal{Z}_{[a,b]}$ und $\xi = (\xi_1,\ldots,\xi_l)$, $\xi_i \in (z_{i-1}, z_i)$, so daß

$$\sup_{(z_{i-1},z_i)} F - \inf_{(z_{i-1},z_i)} F \;<\; \epsilon \quad \text{f.a. } i = 1,\ldots,l\,.$$

Also mit $\varphi = \sum_{i=1}^{l} F(\xi_i) \cdot 1_{[z_{i-1}, z_i)}$ gilt $\|F - \varphi\|_\infty < \epsilon$. Wir berechnen

$$\int_a^b \varphi(x) \sin kx \, dx \;=\; \sum_{i=1}^{l} F(\xi_i) \frac{1}{k} \big(\cos k z_{i-1} - \cos k z_i \big) \;\to\; 0$$

für $k \to \infty$, da $|\cos x| \leq 1$. Daraus folgt

$$\left| \int_a^b F(x) \sin kx \, dx \right| \;\leq\; \epsilon \cdot (b-a) \;+\; \int_a^b \varphi(x) \sin kx \, dx,$$

und somit folgt die Behauptung. □

V.7.14. Lemma (**Dirichlet'sches Lemma**). *Sei $f\colon [-\pi, \pi] \to \mathbb{C}$ stückweise stetig und links- und rechtsseitig differenzierbar in* 0. *Dann gilt:*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_n(t) \, dt \;\to\; \frac{f(0-) + f(0+)}{2} \quad \textit{für } n \to \infty \,.$$

Beweis. Wir haben mit (7)

$$\frac{1}{2\pi} \int_0^{\pi} f(0+) D_n(t) \, dt \;=\; \frac{1}{2} f(0+),$$

und erhalten somit

$$\begin{aligned} A \;:=&\; \frac{1}{2\pi} \int_0^{\pi} \frac{f(t) - f(0+)}{t} \cdot \frac{t}{\sin \frac{t}{2}} \sin(n + \tfrac{1}{2}) t \, dt \\ =&\; \frac{1}{2\pi} \int_0^{\pi} f(t) D_n(t) \, dt \;-\; \frac{1}{2} f(0+) \,. \end{aligned}$$

Aus der rechtsseitigen Differenzierbarkeit folgt

$$\lim_{t \downarrow 0} \frac{f(t) - f(0+)}{t} \cdot \frac{t}{\sin \frac{t}{2}} \;=\; f'(0+) \cdot 2,$$

also ist $\frac{f(t) - f(0+)}{t} \cdot \frac{t}{\sin \frac{t}{2}}$ stückweise stetig und wir können das Riemannsche Lemma V.7.13 anwenden. Daraus folgt $\lim_{n \to \infty} A = 0$, also

$$\lim_{n \to \infty} \frac{1}{2\pi} \int_0^{\pi} f(t) D_n(t) \, dt \;=\; \frac{1}{2} f(0+) \,.$$

Analog zeigt man

$$\lim_{n \to \infty} \frac{1}{2\pi} \int_{-\pi}^{0} f(t) D_n(t) \, dt \;=\; \frac{1}{2} f(0-) \,.$$

Zusammensetzen beider Identitäten ergibt die Behauptung. □

Beweis des Theorems V.7.10. Wir berechnen

$$\begin{aligned} s_n(f)(x) \;=&\; \sum_{k=-n}^{n} \hat{f}(k) e^{ikx} \\ =&\; \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k=-n}^{n} f(t) e^{ik(x-t)} \, dt, \quad \text{substitutiere } \; x - t = y \\ =&\; \frac{1}{2\pi} \int_{x+\pi}^{x-\pi} f(x-y) \sum_{k=-n}^{n} e^{iky} (-dy) \;=\; \frac{1}{2\pi} \int_{x-\pi}^{x+\pi} f(x-y) D_n(y) \, dy \\ =&\; \frac{1}{2\pi} \int_{-\pi}^{\pi} F(y) D_n(y) \, dy, \quad \text{da Integrand } 2\pi\text{-periodisch}, \end{aligned}$$

wobei $F(y) := f(x-y)$. Da f nach Voraussetzung überall links- und rechtsseitig differenzierbar ist, ist das Dirichlet'sche Lemma V.7.14 anwendbar, und es folgt

$$\lim_{n\to\infty} s_n(f)(x) = \frac{F(0-)+F(0+)}{2} = \frac{f(x+)+f(x-)}{2}.$$

Dies ist die behauptete Konvergenz der Fourier-Reihe von f. $\square$

V.7.15. Beispiel.

Als Anwendung betrachten wir die Funktion $\frac{\sin x}{x}$. Wir berechnen das Integral

$$\begin{aligned} I_n &= \int_{-(n+\frac{1}{2})\pi}^{(n+\frac{1}{2})\pi} \frac{\sin x}{x}\,dx, \quad \text{substituieren } x=(n+\frac{1}{2})t, \\ &= \int_{-\pi}^{\pi} \frac{\sin(n+\frac{1}{2})t}{t}\,dt \\ &= \frac{1}{2}\int_{-\pi}^{\pi} f(t)D_n(t)\,dt, \end{aligned}$$

mit $f(t) = \frac{\sin\frac{1}{2}t}{\frac{1}{2}t}$. f ist differenzierbar in 0. Also ist wiederum das Dirichlet'sche Lemma anwendbar und wir erhalten $\lim_{n\to\infty} I_n = \pi f(0)$. Also folgt

$$\int_{-\infty}^{\infty} \frac{\sin x}{x}\,dx = \pi.$$

Daß das Integral existiert, ist bereits früher in einer Übungsaufgabe gezeigt worden.

KAPITEL VI

Topologische Grundbegriffe metrischer Räume

VI.1. Metrische und normierte Räume

VI.1.1. Definition.
Es sei X eine Menge und $d\colon X \times X \to \mathbb{R}$ eine Abbildung. (X, d) heißt **metrischer Raum** genau dann, wenn

(i) $d(p, q) \geq 0$ für alle $p, q \in X$ und $d(p, q) = 0$ genau dann, wenn $p = q$ (*Nicht-Degeneriertheit*),
(ii) $d(p, q) = d(q, p)$ für alle $p, q \in X$ (*Symmetrie*),
(iii) $d(p, r) \leq d(p, q) + d(q, r)$ für alle $p, q, r \in X$ (*Dreiecks-Ungleichung*).

VI.1.2. Beispiel. (1) Sei X beliebig und d_{diskr} definiert durch

$$d_{\mathsf{diskr}}(p, q) := \begin{cases} 0, & p = q, \\ 1, & p \neq q\,. \end{cases}$$

d_{diskr} heißt die **diskrete** Metrik.
(2) $X = \mathbb{R}$, $d(p, q) := |p - q|$.

VI.1.3. Definition.
Sei V ein reeller (nicht notwendig endlich-dimensionaler) Vektorraum und $\|\cdot\|\colon V \to \mathbb{R}$ eine reelle Funktion. Dann heißt $(V, \|\cdot\|)$ ein **normierter Vektorraum** genau dann, wenn

(i) $\|v\| \geq 0$ für alle $v \in V$ und $\|v\| = 0$ genau dann, wenn $v = 0$ (*Nicht-Degeneriertheit*),
(ii) $\|\lambda v\| = |\lambda| \cdot \|v\|$ für alle $\lambda \in \mathbb{R}$, $v \in V$,
(iii) $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$ (*Dreiecks-Ungleichung*).

Offenbar gilt

VI.1.4. Satz. *Jeder normierte Vektorraum $(V, \|\cdot\|)$ trägt eine induzierte Metrik $(V, d_{\|\cdot\|})$, $d_{\|\cdot\|}(v, w) := \|v - w\|$.*

VI.1.5. Beispiel. (1) $V = \mathbb{R}^n$, $p \geq 1$, $\|v\|_p := \|(v_1, \ldots, v_n)\|_p := \Big(\sum_{i=1}^n |v_i|^p\Big)^{\frac{1}{p}}$,
(2) $V = \mathbb{R}^n$, $\|v\|_\infty := \max_{i=1,\ldots,n} |v_i|$.
(3) Sei $V = \mathbb{R}^2$ und $\|(x, y)\|_{\frac{1}{2}} := \big(\sqrt{|x|} + \sqrt{|y|}\big)^2$. Dann ist $\|\cdot\| = \|\cdot\|_{\frac{1}{2}}$ keine Norm. Betrachte zum Beispiel $v = (1, 1) = (1, 0) + (0, 1)$. Es gilt $\|(1, 0)\| = 1 = \|(0, 1)\|$, aber $\|(1, 1)\| = 4 > \|(1, 0)\| + \|(0, 1)\|$. Also ist die Dreiecksungleichung nicht gültig.
(4) Für alle $v \in \mathbb{R}^n$ gilt:

$$\begin{aligned} \|v\|_\infty \;&\leq\; \|v\|_2 \;\leq\; \sqrt{n}\|v\|_\infty, \\ \|v\|_\infty \;&\leq\; \|v\|_1 \;\leq\; n\|v\|_\infty, \\ \text{allgemein}\quad n^{-\frac{1}{q}}\|v\|_q \;&\leq\; \|v\|_p \;\leq\; n^{\frac{1}{p}}\|v\|_q \quad \text{f.a. } p, q \in \mathbb{N} \cup \{\infty\}, \\ \text{insbesondere}\quad \|v\|_q \;&\leq\; \|v\|_p \quad \text{f.a. } p \leq q\,. \end{aligned}$$

(5) $V = C^0\big([a,b],\mathbb{R}\big)$, $\|f\|_\infty := \|f\|_{[a,b]} := \sup_{x\in[a,b]} \big|f(x)\big|$,

Diese Beispiele reihen sich in eine umfassende Theorie von Funktionen- und Folgenräume ein, deren allgemeinste Form wir später noch in der Maß- und Integrations-Theorie kennenlernen werden. Als Einführung sei nachfolgender Abschnitt VI.1.1 mit den wesentlichen Eigenschaften eingefügt.

VI.1.6. DEFINITION.
Seien V ein reeller Vektorraum und $\|\cdot\|_i$, $i = 1, 2$ Normen auf V. Dann heißen $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalent genau dann, wenn Konstanten $c_1, c_2 > 0$ existieren, so daß

$$c_1\|v\|_1 \;\le\; \|v\|_2 \;\le\; c_2\|v\|_1 \quad \text{f.a. } \; v \in V\,.$$

VI.1.7. SATZ. *Je zwei Normen auf $\mathbb{R}^n$ sind äquivalent.*

Den Beweis können wir erst später in Abschnitt VI.7 führen. Als unmittelbare Verallgemeinerung ergibt sich, daß je zwei Normen auf einem beliebigen endlichdimensionalen Vektorraum äquivalent sind.
Dagegen ist diese Aussage im unendlich-dimensionalen Fall falsch, wie folgendes Gegenbeispiel zeigt.

VI.1.8. SATZ. *Die Normen $\|\cdot\|_p$, $\|\cdot\|_q$ auf $C^0([a,b],\mathbb{R})$ sind für $p \neq q$ nicht äquivalent.*

BEWEIS. Wir führen hier den Beweis nur für $p = 1$ und $q = 2$ auf $[0,1]$. Der Beweis für den allgemeinen Fall lässt sich leicht daraus ableiten.
Wir betrachten die Funktionenfolge

$$f_n(x) = \begin{cases} 2n - 2n^2x, & 0 \le x \le \frac{1}{n},\\ 0, & \frac{1}{n} < x \le 1\,. \end{cases}$$

Wie man leicht sieht, gilt $\|f_n\|_1 = 1$ für alle $n \in \mathbb{N}$, jedoch $\|f_n\|_2 = \frac{4}{3}n \to \infty$. □

VI.1.9. DEFINITION.
Sei X eine Menge mit zwei Metriken d_i, $i = 1, 2$. Dann heißen diese Metriken **äquivalent** genau dann, wenn

$$\text{ex. } \; c > 0 \text{ s.d. } \; cd_1(p,q) \;\le\; d_2(p,q) \;\le\; \frac{1}{c}d_1(p,q) \quad \text{f.a. } \; p,q \in X\,.$$

Seien (X_i, d_i), $i = 1, 2$ zwei metrische Räume. Dann heißt eine surjektive Abbildung $f\colon (X_1, d_1) \to (X_2, d_2)$ eine **Isometrie** genau dann, wenn

$$d_2\big(f(p), f(q)\big) \;=\; d_1(p,q) \quad \text{f.a. } \; p,q \in X_1\,.$$

Offenbar ist eine Isometrie automatisch injektiv, also eine Bijektion. Verzichtet man auf die Forderung der Surjektivität so nennt man f eine **isometrische Einbettung**.

VI.1.10. DEFINITION.
Sei (X,d) ein metrischer Raum. Dann heißt die Teilmenge

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

der **(offene) Ball** vom Radius $r > 0$ und Mittelpunkt $x \in X$. Der Ball $B_1(x_o)$ heißt auch **Einheitsball** zum Mittelpunkt x_o.
Sei des weiteren $A \subseteq X$ eine beliebige Teilmenge. Dann wird auf A durch

$$d_A\colon A \times A \to \mathbb{R}, \quad d_A(p,q) = d(p,q) \quad \text{f.a. } \; p,q \in A$$

eine Metrik auf A induziert. Also (A, d_A) mit der **induzierten Metrik** wird selbst als ein metrische Raum aufgefasst.

VI.1.11. BEISPIEL.
Die n-dimensionale Einheitssphäre

$$S^n := \{ (x_0, \ldots, x_n) \in \mathbb{R}^{n+1} \,|\, x_0^2 + \ldots + x_n^2 = 1 \} \subseteq \mathbb{R}^{n+1}$$

trägt die durch die euklidische Metrik aus $\|\cdot\|_2$ induzierte Metrik. Der Durchmesser in der induzierten Metrik beträgt 2. (Dagegen beträgt die Länge der kürzesten Verbindungskurve zwischen zwei Antipoden auf S^n π. Kurvenlängen werden wir später noch genauer betrachten.)

VI.1.1. L^p-Räume.

VI.1.12. DEFINITION.
Es seien $p, q \in (1, \infty)$ sogenannte **konjugierte Exponenten** mit $\frac{1}{p} + \frac{1}{q} = 1$. Wir definieren $l_p(\mathbb{R}) = \{ x = (x_n) \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n\in\mathbb{N}} |x_n|^p < \infty \}$ und

$$\|x\|_p := \Big(\sum_{n=1}^{\infty} |x_n|^p\Big)^{\frac{1}{p}} \quad \text{für } \; x \in l_p(\mathbb{R})\,.$$

Sei $f\colon [a, b] \to \mathbb{R}$ stückweise stetig, dann definieren wir

$$\|f\|_p := \Big(\int_a^b |f(x)|^p dx \Big)^{\frac{1}{p}} .$$

Wir wissen nun aus Aufgabe 1 a), Übungsserie 5, daß für alle $u, v \geq 0$ und konjugierte Exponenten

$$uv \,\leq\, \frac{u^p}{p} + \frac{v^q}{q} \tag{8}$$

gilt.
Seien nun $x = (x_n) \in l_p$, $y = (y_n) \in l_q$ und $N \in \mathbb{N}$. Wir definieren $\langle x, y\rangle_N := \sum_{n=1}^N x_n y_n$, und falls diese Reihe für $N \to \infty$ absolut konvergiert, $\langle x, y\rangle := \lim_{N\to\infty} \langle x, y\rangle_N$.
Es seien $\tilde{x} = \frac{x}{\|x\|_p}$ und $\tilde{y} = \frac{y}{\|y\|_q}$, falls $x \neq 0$ und $y \neq 0$. Dann gilt mit (8)

$$\langle \tilde{x}, \tilde{y}\rangle_N \leq \textstyle\sum_{n=1}^N \big(\frac{\tilde{x}_n^p}{p} + \frac{\tilde{y}_n^q}{q}\big) \,\leq\, \frac{1}{p}\|\tilde{x}\|_p^p + \frac{1}{q}\|\tilde{y}\|_q^q \,=\, \frac{1}{p} + \frac{1}{q} = 1\,.$$

Daraus folgt $\langle x, y\rangle_N \leq \|x\|_p \cdot \|y\|_q$ für alle $N \in \mathbb{N}$ und somit liegt auch absolute Konvergenz vor und es gilt $\langle x, y\rangle \leq \|x\|_p \cdot \|y\|_q$ für alle $x \in l_p$, $y \in l_q$.
Betrachte nun $f, g\colon [a, b] \to \mathbb{R}$ stückweise stetig und $\langle f, g\rangle = \int_a^b f(x)g(x)dx$. Ohne Beschränkung der Allgemeinheit seien f und g nicht die Nullfunktionen, also $\|f\|_p \neq 0$ und $\|g\|_q \neq 0$. Dann normieren wir wieder $\tilde{f} = \frac{f}{\|f\|_p}$ und $\tilde{g} = \frac{g}{\|g\|_q}$, und es folgt analog mit (8)

$$\big|\langle \tilde{f}, \tilde{g}\rangle\big| \,\leq\, \int_a^b \big|\tilde{f}(x)\tilde{g}(x)\big|dx \,\leq\, \int_a^b \Big(\frac{|\tilde{f}|^p}{p} + \frac{|\tilde{g}|^q}{q}\Big)dx \,=\, 1\,.$$

Also folgt wieder $\langle f, g\rangle \leq \|f\|_p \cdot \|g\|_q$. Wir haben somit für $p, q > 1$ folgenden Satz bewiesen

VI.1.13. SATZ. **(Hölder-Ungleichung)** *Seien $p, q \in (1, \infty)$ konjugierte Exponenten bzw. $p = 1$ und $q = \infty$. Dann gilt für alle $x \in l_p$ und $y \in l_q$ bzw. $f, g\colon [a, b] \to \mathbb{R}$ stückweise stetig*

$$\langle x, y\rangle \,\leq\, \|x\|_p \cdot \|y\|_q \quad \textit{bzw.} \quad \langle f, g\rangle \,\leq\, \|f\|_p \cdot \|g\|_q\,.$$

Der Fall $p = 1$ und $q = \infty$ führt auf die einfache Abschätzung gegen die Supremums-Norm $\|\cdot\|_\infty$ und ist unmittelbar direkt zu beweisen.
Mit Hilfe der Hölder-Ungleichung lässt sich nun leicht die wesentliche Bedingung der Norm-Eigenschaft beweisen:

VI.1.14. SATZ. **(Minkowski-Ungleichung)** *Es seien $p \in [1, \infty]$ und $x, y \in l_p$ bzw. $f, g \colon [a, b] \to \mathbb{R}$ stückweise stetig. Dann gilt $x + y \in l_p$ und*

$$\|x+y\|_p \;\le\; \|x\|_p + \|y\|_p \quad \textit{sowie}$$
$$\Big(\int_a^b \big(f(x)+g(x)\big)^p dx\Big)^{\frac{1}{p}} \;\le\; \Big(\int_a^b f(x)^p dx\Big)^{\frac{1}{p}} + \Big(\int_a^b g(x)^p dx\Big)^{\frac{1}{p}} .$$

BEWEIS. Betrachte $x, y \in l_p$, dann folgt aus der Hölder-Ungleichung

$$\begin{aligned} \sum_{n=1}^{N} x_n (x_n + y_n)^{p-1} \;&\le\; \|x\|_p \Big(\sum_{n=1}^{N} (x_n + y_n)^{(p-1)q} \Big)^{\frac{1}{q}} \\ &=\; \|x\|_p \Big(\sum_{n=1}^{N} (x_n + y_n)^{p} \Big)^{1-\frac{1}{p}}, \end{aligned}$$

da $p + q = pq$ für konjugierte Exponenten. Daraus folgt

$$\begin{aligned} \sum_{n=1}^{N} (x_n + y_n)^{p} \;&=\; \sum_{n=1}^{N} (x_n + y_n)(x_n + y_n)^{p-1} \\ &\le\; \big(\|x\|_p + \|y\|_p\big) \Big(\sum_{n=1}^{N} (x_n + y_n)^{p} \Big)^{1-\frac{1}{p}} \end{aligned}$$

und somit

$$\Big(\sum_{n=1}^{N} (x_n + y_n)^{p} \Big)^{\frac{1}{p}} \;\le\; \|x\|_p + \|y\|_p \quad \text{f.a. } \; N \in \mathbb{N} .$$

Also ist diese Reihe wiederum konvergent und $x + y \in l_p$, sowie

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

Der Beweis für die Minkowski-Ungleichung für die Integral-Normen für f, g erfolgt vollkommen analog. Ebenso leicht lässt sich die Behauptung für $p = 1$ und $p = \infty$ mit Hilfe der bekannten Betragsungleichung und Supremums-Eigenschaften zeigen. □

VI.2. Grundbegriffe

VI.2.1. DEFINITION.
Eine Teilmenge $U \subseteq X$ eines metrischen Raumes (X, d) heißt **Umgebung** eines Punktes $x \in X$, wenn es einen offenen Ball um x gibt, der ganz zu U gehört, d.h. $x \in B_r(x) \subseteq U$ f.e. $r > 0$.

VI.2.2. BEISPIEL.
(1) $(-\epsilon, \epsilon)$ ist eine Umgebung von $0 \in \mathbb{R}$.
(2) $\{ (x, y) \,|\, x \cdot y \ge 0 \}$ ist keine Umgebung von $(0, 0)$ in $\mathbb{R}^2$.
(3) Der Schnitt von endlich vielen Umgebungen eines Punktes $x \in X$ ist wieder eine Umgebung.
(4) $\bigcap_{n \in \mathbb{N}} (-\frac{1}{n}, \frac{1}{n})$ ist keine Umgebung von 0 in $\mathbb{R}$.

VI.2.3. SATZ (**Hausdorff-Eigenschaft von metrischen Räumen**). *Je zwei verschiedene Punkte $p, q \in X$ eines metrischen Raumes (X, d) besitzen disjunkte Umgebungen, $p \in U \subseteq X$, $q \in V \subseteq X$, $U \cap V = \emptyset$.*

BEWEIS. Es sei $r = d(p, q)$ und $\epsilon = \frac{r}{3}$. Dann sind die ϵ-Bälle um p und q disjunkt. Andernfalls existiert ein $x \in B_\epsilon(p) \cap B_\epsilon(q)$ und die Dreiecksungleichung der Metrik d impliziert mit

$$d(p, q) \le d(p, x) + d(x, q) < 2\epsilon = \frac{2}{3} r < r$$

einen Widerspruch. □

VI.2.4. Definition.
Sei (X,d) ein metrischer Raum und $A \subseteq X$ eine Teilmenge und $x \in X$ ein Punkt. Dann heißt x ein

- **innerer Punkt** von $A \overset{def}{\Leftrightarrow}$ ex. Umgebung U von x s.d. $U \subseteq A$,
- **äußerer Punkt** von $A \overset{def}{\Leftrightarrow}$ ex. Umgebung U von x s.d. $U \cap A = \emptyset$, also genau dann, wenn x ein innerer Punkt von $X \setminus A$ ist,
- **Randpunkt** von A $\overset{def}{\Leftrightarrow}$ f.a. Umgebungen U von x gilt $U \cap A \neq \emptyset$ und $U \cap (X \setminus A) \neq \emptyset$, d.h. jede Umgebung von x trifft sowohl A als auch das Komplement.

Wir definieren die folgende Mengen,

$$\begin{aligned} \overset{\circ}{A} &:= \{\, x \in A \mid x \text{ ist innerer Punkt von } A \,\}, \\ \partial A &:= \{\, x \in A \mid x \text{ ist Randpunkt von } A \,\}, \\ \overline{A} &:= A \cup \partial A\,. \end{aligned}$$

Wir bezeichnen mit $\overset{\circ}{A}$ das **Innere** von A, ∂A den **Rand** von A und $\overline{A}$ den **Abschluss** oder die **abgeschlossene Hülle** von A in X.

VI.2.5. Beispiel. (1) Sei $A = \{\,(x,y) \mid x,y \geq 0,\ \max(x,y) < 1\,\} = [0,1) \times [0,1)$ als Teilmenge von $\mathbb{R}^2$. Dann sind $(\frac{1}{2}, \frac{1}{2})$ ein innerer Punkt von A, $(2,1)$ ein äußerer Punkt von A und $(1,1)$, $(0,0)$, $(\frac{1}{2},0)$ Randpunkte von A. Es gilt

$$\begin{aligned} \overset{\circ}{A} &= (0,1) \times (0,1), \\ \overline{A} &= [0,1] \times [0,1], \\ \partial A = \overline{A} \setminus \overset{\circ}{A} &= \{0,1\} \times [0,1] \cup [0,1] \times \{0,1\}\,. \end{aligned}$$

(2) $X = \mathbb{R}$, $A = \mathbb{Q} \subseteq \mathbb{R}$. Dann gilt $\overline{A} = \mathbb{R}$, $\overset{\circ}{A} = \emptyset$, $\partial A = \mathbb{R}$.

VI.2.6. Satz. *Für Teilmengen $A, B \subseteq X$ eines metrischen Raumes gilt:*

(a) $\overline{A} = \{\, x \in X \mid$ *jede Umgebung* U *von* x *trifft* A*, d.h.* $U(x) \cap A \neq \emptyset \,\}$*,*
(b) $\overset{\circ}{A} = A \setminus \partial A$*,*
(c) $\overset{\circ}{A} = \{\, x \in X \mid A$ *ist Umgebung von* $x \,\}$*,*
(d) $\overset{\circ}{A} \subseteq A \subseteq \overline{A}$*,*
(e) $A \subseteq B \Rightarrow \overset{\circ}{A} \subseteq \overset{\circ}{B}$ *und* $\overline{A} \subseteq \overline{B}$*,*
(f) $X \setminus \overset{\circ}{A} = \overline{X \setminus A}$*,* $X \setminus \overline{A} = (X \setminus A)^\circ$*,*
(g) $\partial A = \partial(X \setminus A)$
(h) $\overset{\circ}{\emptyset} = \emptyset$*,* $\overset{\circ}{X} = X$*,* $\overline{X} = X$*.*

von (f). Die anderen Aussagen sind ebenso leichte Übungen.

$$\begin{aligned} x \in X \setminus \overset{\circ}{A} &\Leftrightarrow x \text{ ist kein innerer Punkt von } A \\ &\Leftrightarrow \text{ex. keine Umgebung } U(x) \subseteq A \\ &\Leftrightarrow \text{jede Umgebung von } x \text{ trifft das Komplement},; U(x) \cap X \setminus A \neq \emptyset \\ &\overset{(a)}{\Leftrightarrow} x \in \overline{X \setminus A}\,. \end{aligned}$$

Der 2. Teil erfolgt analog. □

VI.2.7. Definition.
Sei $A \subseteq X$ Teilmenge eines metrischen Raumes.

$$\begin{aligned} A \text{ heißt } \textbf{offen in } X &\overset{def}{\iff} \partial A \cap A = \emptyset \\ &\Leftrightarrow A = \mathring{A} \\ &\Leftrightarrow \forall x \in A \, \exists \, \text{Umgebung } U(x) \subseteq A \\ &\Leftrightarrow A \text{ ist Umgebung für alle } x \in A\,. \\ A \text{ heißt } \textbf{abgeschlossen in } X &\overset{def}{\iff} \partial A \subseteq A \\ &\Leftrightarrow A = \overline{A} \\ &\Leftrightarrow X \setminus A \text{ ist offen in } X\,. \end{aligned}$$

Beachten Sie, daß diese Eigenschaften "offen" und "abgeschlossen" einer Teilmenge nur relative Eigenschaften sind, relativ zum umgebenden metrischen Raum!

VI.2.8. Beispiel.
(1) $(0,1)$ ist offen in $\mathbb{R}$,
(2) S^n ist abgeschlossen in $\mathbb{R}^{n+1}$.
(3) $\{\,(x,y) \,|\, x \cdot y = 1\,\}$ ist abgeschlossen in $\mathbb{R}^2$.

VI.2.9. Satz (**Haupteigenschaften offener und abgeschlossener Mengen**). *Sei (X,d) ein metrischer Raum.*

(1) *Der Durchschnitt endlich vieler in X offener Mengen ist wieder offen in X.*
(2) *Die Vereinigung beliebig vieler in X offener Mengen ist wieder offen in X.*
(3) *Der Durchschnitt beliebig vieler in X abgeschlossener Mengen ist wieder abgeschlossen in X.*
(4) *Die Vereinigung endlich vieler in X abgeschlossener Mengen ist wieder abgeschlossen in X.*
(5) *$\mathring{A}$ ist die größte in X offene Menge, die in A enthalten ist,*

$$\mathring{A} = \bigcup \{\, B \subseteq X \,|\, B \textit{ offen}\,,\ B \subseteq A\,\}\,.$$

(6) *$\overline{A}$ ist die kleinste in X abgeschlossene Menge, welche A enthält,*

$$\bigcap \{\, B \subseteq X \,|\, B \textit{ abgeschlossen}\,,\ A \subseteq B\,\}\,.$$

(7) *$\emptyset$ und X sind abgeschlossen und offen in X.*

Beweis. *zu (1)*: Es seien A und B offen in X und $x \in A \cap B$. Dann finden wir $\epsilon_1 > 0$ mit $B_{\epsilon_1}(x) \subseteq A$ und $\epsilon_2 > 0$ mit $B_{\epsilon_2}(x) \subseteq B$. Sei $\epsilon = \min(\epsilon_1, \epsilon_2)$. Dann ist $B_\epsilon(x) \subseteq A \cap B$, also $A \cap B$ offen.
zu (2): Es sei $\{A_i \,|\, i \in I\,\}$ eine beliebige Familie von in X offenen Mengen. Dann gilt $x \in \cup_{i \in I} A_i$ genau dann, wenn $x \in A_{i_o}$ für ein $i_o \in I$. Da A_{i_o} offen ist, existiert ein $B_\epsilon(x) \subseteq A_{i_o}$ und somit auch $\subseteq \cup_{i \in I} A_i$. Also ist die Vereinigung ebenfalls offen.
zu (3): Dies folgt direkt aus (2) mit der Eigenschaft, daß abgeschlossene Mengen genau die Komplemente offener Mengen sind und Komplemente von Durchschnitten die Vereinigungen der Einzelkomplemente sind.
zu (4): analog aus (1).
zu (5): Sei $B \subseteq A$ und B offen. Dann gilt $B = \mathring{B} \subseteq \mathring{A}$, also gilt

$$\bigcup \{\, B \subseteq X \,|\, B \text{ offen}\,,\ B \subseteq A\,\} \subseteq \mathring{A}\,.$$

Die Rückrichtung $\supset$ ist trivial, da $\mathring{A}$ in der Menge der B selbst vorkommt.

<u>*zu (6)*</u>: folgt mittels $X \setminus \overline{A} = (X \setminus A)^\circ$ aus (5).
<u>*(7)*</u> ist trivial.

□

VI.2.10. Beispiel.
Sei (X, d) ein metrischer Raum und $A \subseteq X$ eine Teilmenge mit der induzierten Metrik d_A. Dann gilt: Eine Menge $U \subseteq A$ ist offen in (A, d_A) genau dann, wenn eine in X offene Menge V existiert mit $U = V \cap A$. Zum Beweis betrachte

$$U = \bigcup \{ B_{d_A}(x, r) \,|\, x \in U,\, r > 0 \text{ s.d. } B_{d_A}(x, r) \subseteq U \},$$

wobei $B_{d_A}(x, r)$ die Bälle in (A, d_A) bzgl. der induzierten Metrik sind und nach Definition

$$B_{d_A}(x, r) \;=\; B_d(x, r) \cap A$$

für alle $x \in A$. Sei also $V = \bigcup \{ B_d(x, r) \,|\, x \in U,\, r > 0 \text{ s.d. } B_{d_A}(x, r) \subseteq U \}$. Somit ist V als Vereinigung von in X offenen Bällen in X offen und es gilt $U = V \cap A$. Das Offenheit immer eine relative Eigenschaft einer Teilmenge ist, kann man sprachlich noch damit unterstreichen, daß man sagt, U ist **relativ offen** in A.

VI.2.11. Definition.
Sei $A \subseteq (X, d)$ eine Teilmenge eines metrischen Raums und $x \in \overline{A}$. Dann definieren wir in den folgenden beiden Fällen:

Fall 1: Falls eine Umgebung U von x in X existiert mit $U \cap A = \{x\}$, so heißt x ein **isolierter Punkt** von A.

Fall 2: Falls für jede Umgebung U von x in X gilt $(U \cap A) \setminus \{x\} \neq \emptyset$, d.h. jede Umgebung von x enthält einen Punkt von A verschieden von x, so heißt x ein **Häufungspunkt** von A.

VI.2.12. Satz. *x ist ein Häufungspunkt von A genau dann, wenn jede Umgebung von x unendlich viele Punkte von A enthält.*

Beweis. <u>"$\Leftarrow$"</u>: ist direkt nach Definition klar.
<u>"$\Rightarrow$"</u>: Angenommen, U ist eine Umgebung von x, die nur endlich viele Punkte aus A enthält, $A \cap U = \{a_1, \ldots, a_n\}$. Aufgrund der Hausdorff-Eigenschaft von (X, d), siehe Satz VI.2.3, existiert zu jedem a_i, $i = 1, \ldots, n$ eine Umgebung $a_i \in V_i$ und $U_i(x)$ von x, so daß $V_i \cap U_i = \emptyset$ für alle $i = 1, \ldots, n$. Also ist $U \cap U_1 \cap \ldots \cap U_n$ eine Umgebung von x, welche keine Punkt aus A verschieden von x enthält. Also ist x kein Häufungspunkt von A. □

VI.3. Konvergenz

VI.3.1. Definition.
Sei (X, d) ein metrischer Raum und $(x_n) \in X^{\mathbb{N}}$ eine gegebene Folge. Wir definieren:

> (x_n) **konvergiert in X gegen** $x \in X$ genau dann, wenn zu jeder Umgebung U von x ein $n_o \in \mathbb{N}$ existiert mit $x_n \in U$ für alle $n \geq n_o$. Wir schreiben in diesem Fall $\lim_{n\to\infty} x_n = x$.

Unmittelbar aus Definition einer Umgebung folgen die äquivalenten Formulierungen für den Konvergenzfall:

VI.3.2. Satz. *Folgende Aussagen sind für $(x_n) \in X^{\mathbb{N}}$, $x \in X$ äquivalent:*

- $\lim_{n\to\infty} x_n = x$,
- $\forall\, \epsilon > 0\ \exists\, n_o(\epsilon) \in \mathbb{N}$ *mit* $d(x_n, x) < \epsilon$ *für alle* $n \geq n_o$,
- $\forall\, \epsilon > 0\ \exists\, n_o(\epsilon) \in \mathbb{N}$ *mit* $x_n \in B_d(x, \epsilon)$ *für alle* $n \geq n_o$,
- $\big(d(x, x_n)\big)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ *ist eine Nullfolge.*

Ebenso folgt direkt aus der Hausdorff-Eigenschaft VI.2.3

VI.3.3. SATZ. *Jede konvergente Folge in einem metrischen Raum hat einen eindeutigen Grenzwert.*

Mittels Folgenkonvergenz ergibt sich eine alternative Charakterisierung abgeschlossener Mengen:

VI.3.4. SATZ. *Sei (X, d) ein metrischer Raum und $A \subseteq X$. Dann gilt*

$$\overline{A} = \{ x \in X \mid \textit{ex. } (x_n) \in A^{\mathbb{N}} \textit{ mit } \lim_{n\to\infty} x_n = x \}.$$

BEWEIS. $\underline{"\subseteq"}$: Sei $x \in \overline{A}$ und ObdA $x \notin A$. Also gilt für alle Umgebung U von x: $u \cap A \neq \emptyset$. Betrachte eine Folge von Umgebungen durch Bälle $B_{\frac{1}{n}}(x)$. Also existiert für alle $n \in \mathbb{N}$ ein $x_n \in B_{\frac{1}{n}}(x) \cap A$. Offenbar konvergiert x_n gegen x.
$\underline{"\supseteq"}$: Sei $x = \lim x_n$ mit $(x_n) \subset A$. Sei U eine beliebige Umgebung von x. Dann existiert ein $n_o \in \mathbb{N}$ mit $x_n \in U$ f.a.$n \geq n_o$. Also insbesondere $U \cap A \neq \emptyset$. □

Eine direkte Folgerung ist die oft verwendete Charakterisierung abgeschlossener Teilräume:

> $A \subseteq X$ ist abgeschlossen genau dann, wenn der Grenzwert jeder Folge $(x_n) \in A^{\mathbb{N}}$ aus A, welche in X konvergiert, ebenfalls zu A gehört.

Eine unmittelbare Verallgemeinerung der Grenzwertsätze für reelle Folgen ist

VI.3.5. SATZ. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann gilt:*
- $\lim x_n = x$, $\lim y_n = y \Rightarrow \lim(x_n + y_n) = x + y$,
- $\lim x_n = x$, $\lambda \in \mathbb{R} \Rightarrow \lim(\lambda x_n) = \lambda x$,
- (x_n) *konvergiert* $\Rightarrow$ (x_n) *ist beschränkt, d.h. ex.* $c > 0$ *mit* $\|x_n\| \leq c$ *f.a.* $n \in \mathbb{N}$.

VI.3.6. SATZ. *Sei $\|\cdot\|$ eine gegebene Norm auf $\mathbb{R}^n$ und $(x_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}^n$, $x_k = (x_k^1, \dots, x_k^n)$, $k \in \mathbb{N}$. Dann gilt: $x_k \to \xi \in \mathbb{R}^n$ konvergiert bzgl. $\|\cdot\|$ genau dann, wenn alle Koordinatenfolgen konvergieren, $x_k^i \to \xi^i$ für $k \to \infty$, f.a. $i = 1, \dots, n$.*

BEWEIS. Nach Satz VI.1.7 ist $\|\cdot\|$ äquivalent zur Maximumsnorm $\|\cdot\|_\infty$. Also gilt

$$\lim_{k\to\infty} \|x_k - \xi\| = 0 \Leftrightarrow \lim_{k\to\infty} \|x_k - \xi\|_\infty = 0 \Leftrightarrow \lim_{k\to\infty} |x_k^i - \xi^i| = 0 \text{ f.a. } i = 1, \dots, n.$$

□

VI.3.7. BEISPIEL. (1) Sei $X \neq \emptyset$ eine beliebige Menge und $BX = \{ f : X \to \mathbb{R} \mid f \text{ ist beschränkt} \}$. Dann ist BX mit der Supremumsnorm $\|\cdot\|_\infty$ ein normierter Vektorraum, und es gilt: $(f_n) \in (BX)^{\mathbb{N}}$ konvergiert gegen f bzgl. $\|\cdot\|_\infty$ genau dann, wenn $f_n \to f$ gleichmäßig auf X.
(2) Seien $X = C^0([a, b], \mathbb{R})$, $(f_n), f \in X$, also stetige Funktionen auf $[a, b]$. Dann gilt

$$\begin{aligned} & f_n \to f \quad \text{kvg. im quadratischen Mittel} \\ \overset{\text{def}}{\Longleftrightarrow} \; & \lim_{n\to\infty} \|f_n - f\|_2 = 0 \\ \Longleftrightarrow \; & \lim_{n\to\infty} \int_a^b |f_n(x) - f(x)|^2 \, dx \\ \overset{\text{def}}{\Longleftrightarrow} \; & f_n \xrightarrow{L^2} f \, . \end{aligned}$$

(3) Betrachte due Funktionenfolge $f_n(x) = n^{\frac{1}{5}} e^{-nx^2} \in C^0(\mathbb{R}, \mathbb{R})$. Es gilt für die Quadratintegralnorm

$$\begin{aligned} \|f_n\|_2 &= n^{\frac{1}{5}} \Big(\int_{-\infty}^{\infty} e^{-2nx^2}\, dx \Big)^{\frac{1}{2}}, \quad \text{substituiere } y = \sqrt{2n}x, \\ &= n^{\frac{1}{5}} \Big(\frac{1}{\sqrt{2n}} \int_{-\infty}^{\infty} e^{-y^2} dy \Big)^{\frac{1}{2}} \\ \frac{n^{\frac{1}{5}} \pi^{\frac{1}{4}}}{(2n)^{\frac{1}{4}}} &= \big(\frac{\pi}{2}\big)^{\frac{1}{4}} \cdot n^{\frac{1}{5} - \frac{1}{4}} \to 0\,. \end{aligned}$$

Also gilt $f_n \to 0$ in $L^2(\mathbb{R})$, aber $\sup_{x \in \mathbb{R}} |f_n(x)| = n^{\frac{1}{5}} \to \infty$, und $f_n(x) \to 0$ punktweise für alle $x \neq 0$.

Vollkommen analog zum Fall reeller Folgen gilt

VI.3.8. Satz. *Die Häufungspunkte einer Folge $(x_n) \in X^{\mathbb{N}}$ sind genau die Grenzwerte der konvergenten Teilfolgen von (x_n).*

Beweis. Übung. □

VI.4. Vollständigkeit

VI.4.1. Definition.
Sei (X, d) ein metrischer Raum und $(x_n) \in X^{\mathbb{N}}$. (x_n) heißt **Cauchy-Folge** in (X, d) genau dann, wenn

$$\forall\, \epsilon > 0\, \exists\, n_o \in \mathbb{N}\colon \quad d(x_n, x_m) < \epsilon \quad \text{f.a. } m, n \geq n_o\,.$$

VI.4.2. Beispiel.
Jede konvergente Folge ist eine Cauchy-Folge, da $d(x_n, x_m) < d(x, x_n) + d(x, x_m)$.

VI.4.3. Definition.
Ein metrischer Raum (X, d) heißt vollständig, wenn jede Cauchy-Folge auch in X konvergiert.

VI.4.4. Beispiel. (1) $(\mathbb{Q}, d)$ mit der durch den Betrag induzierten Metrik ist nicht vollständig.

(2) Sei BX wieder der Raum der beschränkten Funktionen auf einer beliebigen Menge X mit der Supremums-Norm wie in VI.3.7. Dann ist $(BX, \|\cdot\|_\infty)$ vollständig.
Beweis: Sei $(f_n) \in (BX)^{\mathbb{N}}$ eine Cauchy-Folge. Also

$$\forall\, \epsilon > 0\, \exists\, n_o \in \mathbb{N}\colon \ \sup_{x \in X} |f_n(x) - f_m(x)| < \epsilon \quad \text{f.a. } n, m \geq n_o,$$

insbesondere ist für jedes feste $x \in X$ die reelle Folge $(f_n(x)) \in \mathbb{R}^{\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Da $\mathbb{R}$ qua Axiom vollständig ist existiert eine Funktion $f\colon X \to \mathbb{R}$ so daß $\lim_{n \to \infty} f_n(x) = f(x)$, also liegt zumindest punktweise Konvergenz vor. Da $|f_n(x) - f_m(x)| < \epsilon$ für alle $m, n \geq n_o$ gilt, so folgt nach $\lim_{m \to \infty}$ ebenfalls $|f_n(x) - f(x)| \leq \epsilon$ für alle $n \geq n_o$ und $x \in X$. Inbesondere ist f beschränkt und $\lim_{n \to \infty} \|f_n - f\|_\infty = 0$, d.h. (f_n) konvergiert in $(BX, \|\cdot\|_\infty)$.

(3) $\big(C^0([a, b], \mathbb{R}), \|\cdot\|_\infty\big)$ ist vollständig. *Beweis*:
Da $C^0([a, b], \mathbb{R}) \subset B([a, b])$, ist jede Cauchy-Folge in $(C^0([a, b]), \|\cdot\|_\infty)$ konvergent gegen ein $f \in B([a, b])$. Es bleibt also zu zeigen, daß dieser Grenzwert f wieder stetig ist. Dies zu zeigen ist äquivalent zur Abgeschlossenheit von $C^0([a, b])$ in $B([a, b])$. *Beweis hiervon*:
Sei $x_o \in [a, b]$ und $\epsilon > 0$. Wähle ein $n \in \mathbb{N}$ so daß $\|f_n - f\|_\infty < \frac{\epsilon}{3}$, und

ein $\delta > 0$ so daß $|f_n(x) - f_n(x_o)| < \frac{\epsilon}{3}$ für alle $x \in [a,b]$ mit $|x - x_o| < \delta$. Dann folgt

$$\begin{aligned} |f(x) - f(x_o)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(x_o)| + |f_n(x_o) - f(x_o)| \\ &< \epsilon \quad \text{f.a. } |x - x_o| < \delta \,. \end{aligned}$$

Also ist f stetig.

(4) Der normierte Raum $\big(C^0([0,2],\mathbb{R}), \|\cdot\|_2\big)$ ist nicht vollständig. Als Beweis genügt es eine konkrete nicht-konvergente Cauchy-Folge zu finden. Betrachte

$$f_n(x) = \begin{cases} x^n, & 0 \le x \le 1, \\ 1, & 1 \le x \le 2, \end{cases} \quad \text{und} \quad f(x) = \begin{cases} 0, & 0 \le x \le 1, \\ 1, & x \ge 1 \,. \end{cases}$$

Es gilt $\|f_n - f\|_2^2 = \left(\int_0^1 x^{2n} dx\right)^{\frac{1}{2}} = \frac{1}{\sqrt{2n-1}} \to 0$. Also ist (f_n) eine Cauchy-Folge bzgl. $\|\cdot\|_2$. Außerdem ist $(f_n) \subset C^0([0,2])$, aber $f \notin C^0$. Da je zwei stetige Funktionen ϕ, ψ mit $\|\phi - \psi\|_2 = 0$ bereits identisch sind, muss jede stetige Funktion g mit $\|f_n - g\|_2 \to 0$ auf $[0,2] \setminus \{1\}$ mit f übereinstimmen, was aufgrund der Sprungstelle von f in 1 nicht möglich ist. Also kann (f_n) nicht bzgl. $\|\cdot\|_2$ konvergent sein.

(5) Sei nun $X = C^1([a,b])$ und definiere für $f \in X$,

$$\|f\|_{C^1} := \|f\|_\infty + \|f'\|_\infty \,.$$

Dann ist $(C^1([a,b]), \|\cdot_{C^1})$ ein vollständiger normierter Vektorraum. *Beweis*: Wie man leicht aus $\|\cdot\|_\infty$ folgert, ist X ein normierter Vektorraum. Sei nun (f_n) eine Cauchy-Folge in X. Dann ist sowohl (f_n) eine Cauchy-Folge in $(C^0, \|\cdot\|_\infty)$ als auch (f_n') eine Cauchy-Folge in $(C^0, \|\cdot\|_\infty)$. Also gilt $f_n \to f$ und $f_n' \to g$ gleichmäßig auf $[a,b]$ für stetige Funktionen f und g. Aus Korollar V.6.6 folgt dann $f \in C^1$, $f' = g$ und somit $f_n \to f$ bzgl. $\|\cdot\|_{C^1}$.

VI.4.5. Satz. *Jede Cauchy-Folge $(x_n) \subset (X,d)$ mit Häufungspunkt ist konvergent.*

Beweis. Sei (x_n) eine Cauchy-Folge mit Häufungspunkt x. Also existiert eine Teilfolge (n_k) mit $x_{n_k} \to x$. Sei $\epsilon > 0$ und finde $n_o \in \mathbb{N}$ mit $d(x_n, x_m) < \frac{\epsilon}{2}$ für alle $m, n \ge n_o$. Finde des weiteren $k_o \in \mathbb{N}$ so daß $d(x_{n_k}, x) < \frac{\epsilon}{2}$ für alle $k \ge k_o$. Sei nun $k \in \mathbb{N}$ groß genug, daß sowohl $k \ge k_o$ als auch $n_k \ge n_o$ gilt. Dann folgt für alle $n \ge n_o$

$$d(x_n, x) < d(x_n, x_{n_k}) + d(x_{n_k}, x) < \epsilon \,.$$

□

VI.4.6. Satz. *Eine abgeschlossene Teilmenge eines vollständigen Raumes ist wieder vollständig.*

Beweis. Sei (X,d) vollständig, $A \subseteq X$ abgeschlossen und $(x_n) \in A^{\mathbb{N}}$ eine Cauchy-Folge in A. Dann ist (x_n) ebenfalls Cauchy-Folge in X also dort konvergent gegen $x \in X$. Da A abgeschlossen ist, folgt gemäß der Folgerung aus Satz VI.3.4, daß $x \in A$. □

Hierzu ist auch folgende Umkehrung gültig.

VI.4.7. Satz. *Sei (X,d) ein beliebiger metrischer Raum, $A \subseteq X$, d_A die induzierte Metrik auf A und sei (A, d_A) vollständig. Dann ist A ein abgeschlossener Teilraum von X.*

Beweis. Sei $(x_n) \subset A$ mit $x_n \to x \in X$. Dann ist (x_n) eine Cauchy-Folge in A bzgl. d_A und aufgrund der Vollständigkeit von A auch in A konvergent. Aus der

Eindeutigkeit von Grenzwerten folgt somit $x \in A$. Also ist $A \subset X$ abgeschlossen nach Satz VI.3.4. □

Es gilt folgende Verallgemeinerung des Intervallschachtelungsprinzips:

VI.4.8. SATZ. *Sei (X, d) vollständig und $A_1 \supset A_2 \supset A_3 \supset \ldots$ eine absteigende Folge von nichtleeren abgeschlossenen Teilmengen $A_n \subset X$, $A_n \neq \emptyset$ f.a. $n \in \mathbb{N}$. Sei außerdem die Folge der Durchmesser von A_n eine Nullfolge,*

$$\operatorname{diam}(A_n) \stackrel{\mathsf{def}}{=} \sup \big\{ d(x, y) \,|\, x, y \in A_n \big\} \to 0 \quad \text{für } n \to \infty .$$

Dann gilt: $\bigcap_{n \in \mathbb{N}} A_n = \{ x_\infty \}$ für ein $x_o \in X$, das heißt es gibt genau ein x_o welches in allen A_n enthalten ist.

BEWEIS. *Eindeutigkeit*: Falls $x, y \in \bigcap A_n$, so gilt $d(x, y) \leq \operatorname{diam} A_n \to 0$, also $d(x, y) = 0$ und somit $x = y$.
Existenz: Seien $x_n \in A_n$ f.a. $n \in \mathbb{N}$. Dann ist (x_n) eine Cauchy-Folge, da $\operatorname{diam} A_n \to 0$. Da X vollständig ist, folgt $x_n \to x_\infty$, und da alle A_n abgeschlossen sind, gilt $x_\infty \in A_n$ f.a. $n \in \mathbb{N}$. □

VI.5. Stetigkeit

VI.5.1. DEFINITION.
Seien (X, d_X), (Y, d_Y) metrische Räume und $f\colon X \to Y$ eine Abbildung. Dann heißt f **stetig in** $x_o \in X$ genau dann, wenn zu jeder Umgebung V von $f(x_o)$ in Y eine Umgebung U von x_o in X existiert mit $f(U) \subseteq V$.

VI.5.2. SATZ. *Sei $f\colon X \to Y$, $x_o \in X$, dann sind folgende Aussagen äquivalent:*

(a) *f ist stetig in x_o.*
(b) *$\forall \epsilon > 0\ \exists \delta = \delta(f, \epsilon, x_o) > 0$ so daß $d_X(x, x_o) < \delta \Rightarrow d_Y(f(x), f(x_o)) < \epsilon$,*
(c) *für jede Folge $(x_n) \in X^{\mathbb{N}}$ mit $\lim x_n = x_o$ gilt $\lim f(x_n) = f(x_o)$.*

BEWEIS. "(a)⇒(b)": $B_\epsilon\big(f(x_o)\big)$ ist eine Umgebung von $f(x_o)$ in Y. Also existiert gemäß (a) eine Umgebung U von x_o in X mit $f(U) \subseteq B_\epsilon\big(f(x_o)\big)$. Nach Definition einer Umgebung existiert ein $\delta > 0$ so daß $B_\delta(x_o) \subseteq U$. Also folgt (b).
"(b)⇒(c)": Sei $\lim x_n = x_o$ in X und $\epsilon > 0$. Gemäß (b) existiert ein $\delta > 0$ mit $f\big(B_\delta(x_o)\big) \subseteq B_\epsilon\big(f(x_o)\big)$. Sei $n_o \in \mathbb{N}$ mit $x_n \in B_\delta(x_o)$ für alle $n \geq n_o$. Also folgt $d_Y\big(f(x_n), f(x_o)\big) < \epsilon$ f.a. $n \geq n_o$. Also gilt $\lim f(x_n) = f(x_o)$.
"(c)⇒(a)": Beweis durch Widerspruch. Angenommen es existiert eine Umgebung V von $f(x_o)$ in Y so daß für jede Umgebung U von x_o in X ein $y \in f(U) \setminus V$ existiert. Wähle als U die Folge der Bälle $B_{\frac{1}{n}}(x_o)$ bzgl. d_X. Also existiert eine Folge $(x_n) \in X^{\mathbb{N}}$ mit $f(x_n) \notin V$. Da nun $x_n \to x_o$, folgt gemäß (c), daß $f(x_n) \to f(x_o)$. Also muss auch $f(x_n) \in V$ f.a. $n \geq n_o$ gelten für ein $n_o \in \mathbb{N}$ groß genug. Widerspruch zur Annahme. □

VI.5.3. SATZ. *Seien X, Y, Z metrische Räume, $f\colon X \to Y$ stetig in $x_o \in X$ und $g\colon Y \to Z$ stetig in $y_o = f(x_o)$. Dann ist $g \circ f\colon X \to Z$ stetig in x_o.*

BEWEIS. folgt z.B. mit VI.5.2.(b). □

VI.5.4. DEFINITION.
Seien X, Y metrische Räume, dann heißt $f\colon X \to Y$ **stetig (auf X)**, falls f in jedem Punkt $x_o \in X$ stetig ist. Wir schreiben für die Menge aller stetigen Abbildungen von X nach Y, $C^0(X, Y)$, oder auch $C(X, Y)$.

VI.5.5. SATZ. *Äquivalent sind die Aussagen:*

(a) $f \in C^0(X,Y)$,
(b) *Urbilder* $f^{-1}(A)$ *abgeschlossener Teilmengen* $A \subseteq Y$ *sind wieder abgeschlossen in* X.
(c) *Urbilder* $f^{-1}(U)$ *offener Teilmengen* $U \subseteq Y$ *sind wieder offen in* X.

Beweis. Offenbar sind (b) und (c) äquivalent, da offene Teilmengen genau die Komplemente abgeschlossener Teilmengen sind.
"(a)⇒ (b)": Sei $A \subset Y$ abgeschlossen und $(x_n) \in X^{\mathbb{N}}$ mit $x_n \in f^{-1}(A)$ f.a. $n \in \mathbb{N}$ und $x_n \to x$ in X. Dann gilt wegen der Folgenstetigkeit von f in x, $f(x_n) \to f(x)$. Also gilt $x \in f^{-1}(A)$. Somit ist $f^{-1}(A)$ gemäß Satz VI.3.4 ebenfalls abgeschlossen.
"(c)⇒ (a)": Sei V eine Umgebung von $f(x_o)$ in Y. Dann existiert ein offener Ball $B_\epsilon(f(x_o)) \subseteq V$. Nach (c) ist $f^{-1}\big(B_\epsilon(f(x_o))\big)$ offen und somit eine Umgebung von x_o, welche durch f ganz in V abgebildet wird. Also ist f stetig in x_o. □

VI.5.6. Beispiel. (1) Umgekehrt müssen stetige Abbildungen offene Teilmengen nicht unbedingt wieder auf offene Mengen abbilden: Betrachte $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = \sin x$. Dann ist $\mathbb{R} \subseteq \mathbb{R}$ offen, aber $f(\mathbb{R}) = [-1,1]$ ist nicht offen in $\mathbb{R}$. Ein analoges Beispiel für abgeschlossene Mengen ist $f(x) = \arctan x$ mit $f(\mathbb{R}) = (-\pi/2, \pi/2)$ nicht abgeschlossen in $\mathbb{R}$.

(2) Sei $A \subseteq X$ eine nichtleere Teilmenge eines metrischen Raums und definiere die Abstandsfunktion zu A durch
$$d(x,A) := \inf_{a \in A} d(x,a) \quad \text{für } x \in X\,.$$
Dann ist $f(x) = d(x,A)$ eine stetige Funktion auf X, denn $d(x,A) \le d(x,a) \le d(x,y) + d(y,a)$ f.a. $a \in A$, $x,y \in X$. Hieraus folgt $d(x,A) - d(x,y) \le d(y,a)$ f.a. $a \in A$ und somit $d(x,A) - d(x,y) \le d(y,A)$ und $|d(x,A) - d(y,A)| \le d(x,y)$. f ist also Lipschitz-stetig mit Lipschitz-Konstante 1.

(3) Eine Folgerung aus (2) ist, daß $U_\epsilon(A) := \{\, x \in X \mid d(x,A) < \epsilon \,\}$ für alle $\epsilon > 0$ eine offene Umgebung von A ist, denn $U_\epsilon(A) = f^{-1}\big((-\infty,\epsilon)\big) \subset \mathbb{R}$ offen.

(4) <u>Funktionen in den $\mathbb{R}^n$</u>: Sei (X,d) ein metrischer Raum und betrachte $f\colon X \to \mathbb{R}^n$ mit $f(x) = \big(f^1(x), \dots, f^n(x)\big)$, also $f^i\colon X \to \mathbb{R}$ die i-te Koordinatenfunktion. Dann gilt: f ist stetig in $x_o \in X$ genau dann, wenn alle n Koordinatenfunktionen f^i in x_o stetig sind. Dies folgt mittels der Folgenstetigkeit aus Satz VI.3.6.

(5) <u>Abbildungen auf dem $\mathbb{R}^n$</u>: Sei $U \subset \mathbb{R}^n$, (X,d) ein metrischer Raum. Betrachte die Abbildung $f\colon U \to X$ in n Variablen $(x^1,\dots,x^n) \mapsto f(x^1,\dots,x^n) \in X$. Sei f stetig in $(x^1,\dots,x^n)$ und betrachte für ein $i \in \{1,\dots,n\}$ die Einschränkung auf die i-te Variable
$$\begin{aligned} &g_i\colon \{\, t \in \mathbb{R} \mid (x^1,\dots,x^{i-1},t,x^{i+1},\dots,x^n) \in U \,\} \to X,\\ &g_i(t) = f(x^1,\dots,t,\dots,x^n)\,. \end{aligned}$$
Dann ist g_i stetig in x^i. Sei nämlich $t_k \to x^i$. Dann gilt
$$(x^1,\dots,x^{i-1},t_k,x^{i+1},\dots,x^n) \to (x^1,\dots,x^n)$$
und aus der Stetigkeit von f folgt $g(t_k) = f(x^1,\dots,t_k,\dots,x^n) \to f(x^1,\dots,x^n) = g(x^i)$.

(6) Die Umkehrung von (5) ist im Allgemeinen falsch. Betrachte z.B. $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch
$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{falls } (x,y) \ne (0,0),\\ 0, & \text{falls } x = y = 0\,. \end{cases}$$

Es gilt $t \mapsto f(t,y)$ ist für jedes feste y eine stetige Funktion, denn

$$f(t,y) = \begin{cases} 0, & \text{falls } y = 0, \\ \frac{1}{y} \cdot \frac{t}{1+t^2/y^2}, & \text{falls } y \neq 0\,. \end{cases}$$

Analog ist auch $t \mapsto f(x,t)$ für alle festen x stetig in der Variablen t. Aber für die Folge $(x_n, y_n) = (\frac{1}{n}, \frac{1}{n}) \to (0,0)$ gilt $f(x_n, y_n) = \frac{1}{2}$ für alle $n \in \mathbb{N}$. Also ist $\lim f(x_n, y_n) = \frac{1}{2} \neq 0 = f(0,0)$, und f ist somit nicht stetig in $(0,0)$.

VI.5.7. Satz (Stetigkeit von algebraischen Operationen).

(a) *Folgende Funktionen sind stetig:*
$\mathrm{add}\colon \mathbb{R}^2 \to \mathbb{R}$, $\mathrm{add}(x,y) = x + y$,
$\mathrm{mult}\colon \mathbb{R}^2 \to \mathbb{R}$, $\mathrm{mult}(x,y) = x \cdot y$,
$\mathrm{quot}\colon \mathbb{R} \times \mathbb{R}^\times \to \mathbb{R}$, $\mathrm{quot}(x,y) = \frac{x}{y}$.

(b) *Seien $f, g \in C^0(X, \mathbb{R})$ dann gilt*
$f + g, f \cdot g \in C^0(X, \mathbb{R})$ *und*
$\frac{f}{g} \in C^0(X, \mathbb{R})$, *falls $g(x) \neq 0$ f.a. $x \in X$.*

Beweis. Die Stetigkeit von add und mult ist lediglich eine Umformulierung der Grenzwertsätze über die Additivität von Konvergenz von reellen Zahlenfolgen, und entsprechend für die Multiplikation. Die Stetigkeit von quot folgt aus der Stetigkeit von $x \mapsto \frac{1}{x}$ auf $\mathbb{R}^\times$ und der Stetigkeit von mult.
(b) Folgt aus der Stetigkeit bei Kompositionen gemäß Satz VI.5.3, da $f + g = \mathrm{add} \circ (f,g)$ mit $(f,g)\colon X \to \mathbb{R}^2$ stetig nach Voraussetzung, entsprechend für $f \cdot g$ und f/g.

□

VI.5.8. Beispiel.
Betrachten wir den endlich-dimensionalen Vektorraum

$$M(n \times n, \mathbb{R}) \;=\; \big\{\, A = (a_{ij}) \mid a_{ij} \in \mathbb{R} \,\big\} \;\cong\; \mathbb{R}^{n^2},$$

und die darauf definierte Determinanten-Funktion $\det\colon M(n \times n, \mathbb{R}) \to \mathbb{R}$,

$$\det A \;=\; \sum_{\sigma \in S_n} (-1)^{|\sigma|} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)},$$

mit S_n die Gruppe der Permutationen von n Elementen und $|\sigma|$ das Vorzeichen der Permutation σ (Anzahl der enthaltenen Transpositionen modulo 2). Dann ist det eine Verknüpfung von Multiplikationen und Additionen der n^2 vielen Koordinatenfunktionen. Somit ist nach dem vorigen Satz det stetig.
Als Anwendung erhalten wir mittels Satz VI.5.5, daß der topologische Teilraum

$$\begin{aligned} \mathsf{GL}(n, \mathbb{R}) \;&=\; \big\{\, A \in M(n \times n, \mathbb{R}) \mid A \text{ ist invertierbar} \,\big\} \\ &=\; \big\{ A \in M(n \times n, \mathbb{R}) \mid \det A \neq 0 \,\big\} \\ &=\; \det{}^{-1}(\mathbb{R}^\times) \end{aligned}$$

als Urbild der offenen Teilmenge $\mathbb{R}^\times \subset \mathbb{R}$ unter der stetigen Abbildung det wieder offen ist. Das ist äquivalent zu

$$\begin{aligned} &\forall\, A \in \mathsf{GL}(n, \mathbb{R}) \quad \exists\, \delta(A) > 0 \text{ s.d. } \forall B = (b_{ij}) \in M(n \times n, \mathbb{R})\colon \\ &\max_{1 \leq i,j \leq n} |a_{ij} - b_{ij}| < \delta \quad \Rightarrow \quad B \text{ ist invertierbar.} \end{aligned}$$

Der nachfolgende Satz von Tietze-Urysohn zeigt wie allgemein sich neue stetige Funktionen auf beliebigen metrischen Räumen bilden lassen und wie groß somit der Vektorraum der stetigen Funktionen ist. Sowie X keine endliche Menge ist, ist $C^0(X,\mathbb{R})$ bereits unendlich-dimensional. (Beweis!)

VI.5.9. Satz (Fortsetzungssatz von Tietze-Urysohn). *Seien (X,d) ein metrischer Raum, $A\subset X$ eine abgeschlossene Teilmenge und $f\in C^0(A,\mathbb{R})$ eine beschränkte stetige Funktion. Dann existiert eine stetige Fortsetzung $F\in C^0(X,\mathbb{R})$, also $F_{|A}=f$, so daß $\sup_{x\in X}F(x)=\sup_{x\in A}f(x)$ und $\inf_X F=\inf_A f$.*

Beweis. <u>1. Schritt</u>: Falls f konstant ist, so ist der Satz trivialerweise richtig. Sei f also nicht-konstant. Dann können wir nach geeigneter Transformation $\tilde f(x)=\alpha f(x)+\beta$, $\alpha,\beta\in\mathbb{R}$ annehmen, daß $\sup_A f=2$ und $\inf_A f=1$ gilt.
<u>2. Schritt</u>: Wir definieren $F\colon X\to\mathbb{R}$ nun durch

$$F(x):=\begin{cases}\dfrac{\inf_{y\in A}\big(f(y)\cdot d(x,y)\big)}{d(x,A)}\,, & x\notin A,\\ f(x)\,, & x\in A\,.\end{cases}$$

Hierbei ist $d(x,A)=\inf_{y\in A}d(x,y)$ die stetige Abstandsfunktion zur Teilmenge A. Zu zeigen ist nun, daß F stetig ist und daß $\sup_X F=2$ und $\inf_X F=1$.
<u>3. Schritt</u>: Sei $x\notin A$, dann gilt $\frac{d(x,y)}{d(x,A)}\geq 1$ f.a. $y\in A$, und somit $\frac{f(y)\cdot d(x,y)}{d(x,A)}\geq f(y)\geq 1$ und daher $F(x)\geq 1$. Ebenso folgt aus $f(y)\cdot d(x,y)\leq 2d(x,y)$, daß $\inf_{y\in A}f(y)\cdot d(x,y)\leq 2d(x,A)$ und somit $F(x)\leq 2$.
Nach Definition ist F eine Fortsetzung von f. Daher gilt auch umgekehrt $\inf_X F\leq\inf_A f$ und $\sup_X F\geq\sup_A f$. Es bleibt also nur noch die Stetigkeit von F zu zeigen.
<u>4. Schritt</u>: Wir schreiben $h\colon X\to\mathbb{R}$ für $h(x)=\inf_{y\in A}\big(f(y)\cdot d(x,y)\big)$ und machen folgende Fallunterscheidung für $x_o\in X$.
<u>*Fall 1*</u>: Falls $x_o\in\overset{\circ}{A}$, so ist F stetig in x_o nach Definition, da es in einer ganzen Umgebung von x_o mit f übereinstimmt.
<u>*Fall 2*</u>: Nach Voraussetzung ist A abgeschlossen in X. Also ist $X\setminus A$ offen. Sei $x_o\in X\setminus A$, $\epsilon>0$ gegeben und setze $r=d(x_o,A)>0$. Ohne Beschränkung der Allgemeinheit können wir annehmen, daß $\epsilon/2<r$. Dann gilt für alle $d(x,x')<\epsilon/2$ nach der Dreiecksgleichung $d(x,y)<d(x',y)+\epsilon/2$ und somit $h(x)\leq h(x')+\epsilon$, da $f(y)\leq 2$. Analog gilt $h(x')\leq h(x)+\epsilon$ und somit

$$\big|h(x)-h(x')\big|\leq\epsilon\quad\text{f.a. }d(x,x')<\epsilon/2\,.$$

Also ist h stetig auf $X\setminus A$ und somit mittels Satz VI.5.7 (b) auch F.
<u>*Fall 3*</u>: Sei nun $x_o\in\partial A$ und $\epsilon>0$ gegeben.
Wir wählen nun $r>0$ klein genug, so daß $|f(y)-f(x_o)|<\epsilon$ für alle $y\in C:=A\cap B_r(x_o)$. Betrachte $D=A\setminus C=\{\,y\in A\,|\,d(y,x_o)\geq r\,\}$ und $x\in X\setminus A$ mit $d(x,x_o)\leq\frac{r}{4}$. Für alle $y\in D$ gilt $d(x,y)\geq d(x_o,y)-d(x,x_o)\geq\frac{3}{4}r$. Also folgt

$$\inf_{y\in D}\big(f(y)\cdot d(x,y)\big)\geq\tfrac{3}{4}r, \tag{9}$$

da $f\geq 1$ nach Schritt 1.
Andererseits gilt für $x\in X\setminus A$ mit $d(x,x_o)\leq\frac{r}{4}$

$$f(x)\cdot d(x,x_o)\leq 2\cdot\tfrac{r}{4}<\tfrac{3}{4}r\,.$$

Das bedeutet, daß der Punkt $y=x_o$ in $A\cap B_r(x_o)$ liegt und $f(y)\cdot d(x,y)<\frac{3}{4}r$ erfüllt. Die Abschätzung (9) impliziert somit, daß das Infimum für $y\in A$ bereits gleich dem Infimum für $y\in C$ ist, also

$$h(x)=\inf_{y\in A}\big(f(y)\cdot d(y,x)\big)=\inf_{y\in A\cap B_r(x_o)}\big(f(y)\cdot d(y,x)\big)\,.$$

Um mit h zu rechnen, können wir uns also bereits mit dem Infimum auf $A \cap B_r(x_o)$ einschränken. Dies erlaubt nun folgende Abschätzungen für $y \in A \cap B_r(x_o)$: Nach Eigenschaft von r gilt $f(x_o) - \epsilon \leq f(y) \leq f(x_o) + \epsilon$ und somit

$$\begin{aligned}\big(f(x_o) - \epsilon\big) d(x, A) &= \inf_{y \in A \cap B_r(x_o)} \left[\big(f(x_o) - \epsilon\big) d(x, y)\right] \leq h(x) \\ &\leq \inf_{y \in A \cap B_r(x_o)} \left[\big(f(x_o) + \epsilon\big) d(x, y)\right] = \big(f(x_o) + \epsilon\big) d(x, A),\end{aligned}$$

denn $\inf_{y \in A \cap B_r(x_o)} d(x, y) = d(x, A)$, wegen $x \in B_{\frac{r}{4}}(x_o)$. Also folgt

$$\left|\frac{h(x)}{d(x, A)} - f(x_o)\right| \leq \epsilon \quad \text{f.a. } x \in B_{\frac{r}{4}}(x_o) \cap X \setminus A\,.$$

Dies beweist die Stetigkeit von F in $x_o \in \partial A$. □

Das folgende Korollar zeigt, daß im Satz von Tietze-Urysohn sogar die Beschränktheitsannahme weggelassen werden kann.

VI.5.10. KOROLLAR. *Sei (X, d) ein metrischer Raum und $A \subset X$ eine abgeschlossene Teilmenge und $f \in C^0(A, \mathbb{R})$. Dann existiert eine stetige Fortsetzung $F \in C^0(X, \mathbb{R})$, $F_{|A} = f$ mit $\sup_X F = \sup_A f$ und $\inf_X F = \inf_A f$.*

BEWEIS. Nach Tietze-Urysohn hat die beschränkte Funktion $\arctan f(x)$ eine entsprechende Fortsetzung g. Sei dann $F = \tan \circ g$. □

Eine weitere Erweiterung besteht in der Beobachtung, daß aus dem Satz von Tietze-Urysohn auch die Existenz von mengentrennenden stetigen Funktionen folgt

VI.5.11. KOROLLAR. *Für alle disjunkten abgeschlossenen Teilmengen eines metrischen Raums $A, B \subset X$, $A \cap B = \emptyset$ existiert eine stetige Funktion $F\colon X \to [0, 1]$ mit $F_{|A} = 0$ und $F_{|B} = 1$.*

Ein weiterer oft verwendeter Sachverhalt ist die Einschränkung stetiger Abbildungen auf Teilmenge:
Es sei (X, d) ein metrischer Raum, $A \subset X$ ein metrischer Raum mit der induzierten Metrik d_A. Dann ist die Eischränkung $f_{|A}$ einer stetigen Funktion $f \in C^0(X, \mathbb{R})$ wieder stetig, $f_{|A} \in C^0(A, \mathbb{R})$. Der Beweis folgt unmittelbar aus der Folgendefinition der Stetigkeit.

VI.5.12. BEISPIEL.
Betrachte $S^1 = \{\, (x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1 \,\}$ als Teilraum von $\mathbb{R}^2$. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$. Dann ist $g\colon S^1 \to \mathbb{R}$, $g(x, y) = xy$ als Einschränkung von f auf S^1 ebenfalls stetig.

Betrachte nun eine stetige Abbildung $f\colon X \to Y$ zwischen metrischen Räumen, deren Bildmenge in einer Teilmenge $B \subset Y$ enthalten ist, $f(X) \subset B$. Dann ist $f\colon X \to B$ bezüglich des Raumes (B, d_B) mit der induzierten Metrik ebenfalls stetig.

VI.5.13. BEISPIEL.
Die Abbildung $g\colon [0, 2\pi] \to S^1$, $g(\varphi) = (\cos\varphi, \sin\varphi)$ ist stetig.

VI.5.14. BEMERKUNG.
Beachte, daß für eine stetige, bijektive Abbildung $f\colon X \to Y$ die Umkehrabbildung $f^{-1}\colon Y \to X$ nicht stetig sein muss. Als Beispiel dient wieder die stetige Abbildung $g\colon [0, 2\pi) \to S^1$, $g(\varphi) = (\cos\varphi, \sin\varphi)$, welche durch Entfernen des rechten Intervallrandes bijektiv wird. Zum Beispiel ist für $x, y > 0$ die Umkehrabbildung $g^{-1}(x, y) = \arctan(\frac{y}{x})$. Aber die Abbildung $g^{-1}\colon S^1 \to [0, 1]$ ist nicht stetig, wie man sich leicht klarmacht.

Mit stärkeren, hier nicht zur Verfügung stehenden Methoden der Topologie lässt sich sogar zeigen:

VI.5.15. SATZ. *Es gibt keine stetige bijektive Abbildung* $f\colon [0,1] \to S^1$ *mit stetiger Umkehrfunktion.*

VI.6. $*$ Banach-Algebren

Sei $(V, \|\cdot\|)$ ein vollständiger, normierter Vektorraum. Einen solchen bezeichnen wir als auch **Banach-Raum**.
Eine **Reihe** in V ist eine Folge von Partialsummen

$$\sum_{k=1}^{\infty} x_k := \Big(\sum_{k=1}^{n} x_k\Big)_{n\in\mathbb{N}}, \quad \text{bei gegebenem } (x_k)_{k\in\mathbb{N}} \in V^{\mathbb{N}}.$$

Eine solche Reihe heißt **absolut (oder normal) konvergent**, wenn die reelle Zahlenreihe $\sum_{k=1}^{\infty} \|x_k\|$ in $\mathbb{R}$ absolut konvergiert.

VI.6.1. LEMMA. *Wenn die Reihe* $\sum x_k$ *absolut konvergent ist, existiert der Grenzwert* $\lim_{n\to\infty} \sum_{k=1}^{n} x_k$ *in* V.

BEWEIS. Wenn $\sum x_k$ absolut konvergiert, so ist $(\sum_{k=1}^{n} x_k)_{n\in\mathbb{N}}$ aufgrund der Dreiecksungleichung eine Cauchy-Folge in V und somit aufgrund der vorausgesetzten Vollständigkeit von V konvergent. □

VI.6.2. DEFINITION.
Eine **Banach-Algebra** ist ein Banach-Raum $\mathcal{A} = (V, \|\cdot\|)$, der außerdem mit einer Multiplikation $V \times V \to V$, $(x, y) \mapsto xy$, versehen ist, welche der Assoziativität genügt und

$$\|xy\| \le \|x\| \cdot \|y\| \quad \text{f.a. } x, y \in \mathcal{A}.$$

Wir sprechen von einer Banach-Algebra **mit Eins**, wenn es ein multiplikativ neutrales Element $e \in \mathcal{A}$ gibt. $\mathcal{A}$ heißt eine komplexe Banach-Algebra, wenn V ein Vektorraum über dem Körper der komplexen Zahlen ist.

Die Multiplikation muss nicht notwendig kommutativ sein!
Ein wichtiges Beispiel ist die endlich-dimensionale Algebra der $n \times n$-Matrizen $A \in M(n \times n, \mathbb{R})$ mit der Operatornorm $\|A\|_{\mathsf{Op}} = \max_{\|v\|=1} \|Av\|$, bei einer beliebigen Norm $\|\cdot\|$ auf $\mathbb{R}^n$. Aufgrund der endlichen Dimension der Algebra, ist die Operatornorm äquivalent zu jeder anderen Norm auf $\mathbb{R}^{n^2}$ und somit vollständig.

VI.6.3. SATZ. *Sei* $\mathcal{A}$ *eine komplexe Banach-Algebra mit Eins* $1 = e$ *, und*

$$P(z) = \sum_{k=0}^{\infty} a_k z^k, \quad (a_k) \in \mathbb{C}^{\mathbb{N}}$$

sei eine gegebene komplexe Potenzreihe mit Konvergenzradius $R > 0$. *Wir erhalten dann eine wohldefinierte Abbildung*

$$P_{\mathcal{A}}\colon \mathcal{A} \supset B_R(0) \to \mathcal{A}, \quad x \mapsto \sum_{k=0}^{\infty} a_k x^k,$$

(hier ist $x^0 = e = 1$*) und* $P_{\mathcal{A}}$ *ist Lipschitz-stetig auf* $B_{R'}(0)$ *f.a.* $0 < R' < R$.

BEWEIS. Seien $x, y \in B_R(0) \subset \mathcal{A}$ mit $\|x\|, \|y\| \le \rho < R$. Dann gilt

$$\left\| \sum_{k=n}^{n+m} a_k x^k \right\| \le \sum_{k=n}^{n+m} |a_k| \rho^k \to 0 \quad \text{für } n \to \infty,$$

aufgrund des Konvergenzradius' der Potenzreihe $P(z)$. Somit ist $P_{\mathcal{A}}(x)$ als absolut konvergente Reihe in $\mathcal{A}$ wohldefiniert. Ähnlich sehen wir

$$\begin{aligned}\|x^k - y^k\| &= \big\|(x-y)(x^{k-1} + yx^{k-2} + y^2x^{k-3} + \ldots + xy^{k-2} + y^{k-1}\big\| \\ &\le \|x-y\| \cdot k\rho^{k-1}\,.\end{aligned}$$

Daraus folgt

$$\|P_{\mathcal{A}}(x) - P_{\mathcal{A}}(y)\| \le \sum_{k=1}^{\infty} k|a_k|\rho^{k-1} \cdot \|x-y\|\,.$$

Das bedeutet, daß $P_{\mathcal{A}}$ auf $B_{R'}(0)$ Lipschitz-stetig ist, da auch $P'(z)$ eine Potenzreihe mit Konvergenzradius R ist. □

VI.6.4. Beispiel. (a) Sei $P(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} = e^z$. Daraus erhalten wir eine Abbildung auf der Matrix-Algebra, die wir ebenfalls mit $\exp\colon M(n\times n) \to M(n\times n)$ bezeichnen.

Falls A, B zwei kommutierende Matrizen sind, $[A,B] := AB - BA = 0$, so gilt mit Hilfe des Cayley-Produkts

$$\begin{aligned}\exp(A+B) &= \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} \overset{!}{=} \sum_{k=0}^{\infty} \frac{A^k + kAB^{k-1} + \ldots + B^k}{k!} \\ &= \exp(A)\cdot\exp(B)\,, \quad \text{falls } \; AB = BA\,.\end{aligned}$$

Hieraus folgt wegen $\mathrm{id} = \exp(0) = \exp(A-A) = \exp(A)\cdot\exp(-A)$

$$\exp\colon M(n\times n, \mathbb{R}) \to \mathsf{GL}(n,\mathbb{R}), \quad \exp(A)^{-1} = \exp(-A)\,.$$

(b) Betrachte die geometrische Reihe $G(z)$ mit Konvergenzradius $R = 1$. Daraus erhalten wir die Abbildung

$$P_G\colon B_1(0) \to \mathcal{A}, \quad P_G(x) = \sum_{k=0}^{\infty} x^k\,.$$

Es gilt dann $P_G(x)\cdot(1-x) = 1$. Also nimmt P_G seine Werte in der multiplikativen Gruppe der invertierbaren Elemente $\mathcal{A}^\times = \{\, x \in \mathcal{A} \,|\, \text{ex. } x^{-1} \,\}$ an.

Als Konsequenz erhalten wir folgende Beschreibung der Inversion:

$$\mathrm{Inv}\colon \mathcal{A}^\times \to \mathcal{A}^\times, \quad x \mapsto x^{-1} \quad \text{ist stetig.}$$

Beweis: Sei $a \in \mathcal{A}^\times$ und $r := \|a^{-1}\|^{-1}$. Sei $x \in B_r(a)$, d.h. $\|x-a\| < \frac{1}{\|a^{-1}\|}$. Daraus folgt

$$\|1 - a^{-1}x\| \le \|a^{-1}\| \cdot \|a-x\| < 1\,.$$

Also ist P_G anwendbar auf $1 - a^{-1}x$ und wir erhalten aus

$$P_G(1-a^{-1}x)\cdot(a^{-1}x) = 1$$

daß $a^{-1}x \in \mathcal{A}^\times$. Daher ist auch $x \in \mathcal{A}^\times$ und somit ist $\mathcal{A}^\times$ eine offene Teilmenge der Banachalgebra $\mathcal{A}$ und

$$x^{-1} = P_G(1-a^{-1}x)\cdot a^{-1} \quad \text{ist stetig in } \; x\,.$$

VI.7. Kompaktheit

VI.7.1. Definition.
Es sei (X,d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Des weiteren sei $\mathfrak{U} = \{\, U_j \mid j \in J \,\}$ ein System von Teilmengen, $U_j \subset X$, mit einer Indexmenge J. $\mathfrak{U}$ heißt eine

- **Überdeckung von** A, wenn $A \subset \bigcup_{i \in J} U_j$, und als solche
- eine **offene Überdeckung**, wenn alle U_j offene Teilmenge sind,
- eine **endliche Überdeckung**, wenn J endlich ist,
- eine **lokal-endliche Überdeckung** von A, wenn für alle $a \in A$ die Menge $\{\, j \in J \mid a \in U_j \,\}$ endlich ist.

VI.7.2. Beispiel.
- $U_n = [n, n+1]$, $n \in \mathbb{Z}$ ist eine lokal-endliche Überdeckung von $\mathbb{R}$.
- $V_n = (-n, n)$, $n \in \mathbb{N}$ ist eine offene, nicht lokal-endliche Überdeckung von $\mathbb{R}$.

VI.7.3. Definition.
Es sei $\{\, U_j \mid j \in J \,\}$ eine Überdeckung von $A \subset X$ und $J_o \subset J$. Dann heißt $\{\, U_j \mid j \in J_o \,\}$ eine **Teilüberdeckung** von A, falls es immer noch eine Überdeckung von A ist.

VI.7.4. Beispiel.
- $V_n = (-n, n)$, $n \in \mathbb{N}$ ist eine offene Überdeckung von $\mathbb{R}$ ohne eine endliche Teilüberdeckung.
- Sei $A = \{\, \frac{1}{n} \mid n \in \mathbb{N} \,\} \cup \{0\} \subset \mathbb{R}$ und $\mathfrak{U} = \{\, U_j \mid j \in J \,\}$ eine beliebige offene Überdeckung von A, bestehend aus in $\mathbb{R}$ offenen Mengen. Dann besitzt $\mathfrak{U}$ eine endliche Teilüberdeckung, wie man folgendermaßen sieht:

  Es existiert ein $j_o \in J$ mit $0 \in U_{j_o}$. Da $\frac{1}{n} \to 0$ für $n \in \mathbb{N}$, existiert ein $n_o \in \mathbb{N}$ so daß $\frac{1}{n} \in U_{j_o}$ für alle $n \geq n_o$, denn U_{j_o} ist offen in $\mathbb{R}$. Betrachte also die endlich vielen Punkte $1, \ldots, \frac{1}{n_o} \in A$. Für alle $k \in \{1, \ldots, n_o\}$ existiert ein $j_k \in J$ mit $\frac{1}{k} \in U_{j_k}$. Also ist $U_{j_1} \cup \ldots \cup U_{j_{n_o}} \cup U_{j_o}$ eine endliche Teilüberdeckung.

Die zentrale Definition ist nun

VI.7.5. Definition.
Es sei (X,d) ein metrischer Raum und $A \subset X$. Dann heiŝt A **kompakt**, wenn jede offene Überdeckung von A eine endliche Teilüberdeckung besitzt.

Beachte, daß Kompaktheit ein topologisch absoluter Begriff ist. Also bereits von dem metrischer Raum (X,d) selbst kann gefragt werden, ob er kompakt ist oder nicht. Der Standardraum $\mathbb{R}$ ist nicht kompakt, wie wir soeben gesehen haben, denn es gibt mindestens eine offene Überdeckung ohne endliche Teilüberdeckung. Umgekehrt jedoch ist die Kompakheit eines Raumes direkt mittels der Definition meist schwer zu zeigen. Hierfür müssen im Folgenden noch bessere Kriterien entwickelt werden.

VI.7.6. Satz. *Es sei (X,d) ein metrischer Raum. Dann gilt:*
(a) *Ist $A \subset X$ kompakt, so ist A beschränkt, also* $\operatorname{diam}(A) = \sup\{\, d(x,y) \mid x, y \in A \,\} < \infty$.
(b) *Ist $A \subset X$ kompakt, so ist A abgeschlossen in X.*
(c) *Ist (X,d) kompakt und $A \subset X$ abgeschlossen, so ist A ebenfalls kompakt.*
(d) *$A_1, \ldots, A_n \subset X$ kompakt $\Rightarrow A_1 \cup \ldots \cup A_n$ kompakt.*

Beweis. (a): Sei $x_o \in X$ ein beliebiger, fester Punkt in X. Da jeder Punkt in X zu x_o einen endlichen Abstand hat, ist $\{\, B_n(x_o) \mid n \in \mathbb{N} \,\}$ eine offene Überdeckung von X, insbesondere von A. Wenn A kompakt ist, gibt es also eine endliche Teilüberdeckung, d.h. es existieren endlich viele $n_1, \ldots, n_k \in \mathbb{N}$ mit $A \subset B_{n_1}(x_o) \cup \ldots \cup B_{n_k}(x_o)$. Da

alle diese Bälle denselben Mittelpunkt haben, ist ihre Vereinigung gleich dem jeweils größten unter ihnen, $A \subset B_m(x_o)$, $m = \max(n_1, \ldots, n_k)$. Somit ist $\operatorname{diam}(A) < m$.
(b): Sei also A kompakt. Wir zeigen nun, daß dann $X \setminus A$ in X offen ist. Ohne Einschränkung nehmen wir an $A \neq X$. Sei also $x_o \in X \setminus A$. Aufgrund der Hausdorff-Eigenschaft des metrischen Raumes X (siehe VI.2.3) existiert für jedes $a \in A$ disjunkte offene Umgebungen $U(a)$ von a in X und $V_a(x_o)$ von x_o , $U(a) \cap V_a(x_o) = \emptyset$. Insbesondere ist dann $\{\, U(a) \,|\, a \in A \,\}$ eine offene Überdeckung von A. Nach Voraussetzung der Kompaktheit von A existiert hierzu eine endliche Teilüberdeckung $A \subset U(a_1) \cup \ldots \cup U(a_n)$. Sei nun zu diesen Indices der Teilüberdeckung $V = V_{a_1} \cap \ldots \cap V_{a_n}$. Da dies ein endlicher Schnitt von offenen Mengen ist, ist V immer noch eine offene Umgebung von x_o und zwar disjunkt zu jedem $U(a_i)$, $i = 1, \ldots, n$. Also gilt $V \cap \big(U(a_1) \cup \ldots \cup U(a_n)\big) = \emptyset$ und somit $V \cap A = \emptyset$. Also ist $V \subset X \setminus A$, was zu zeigen war.
(c): Sei $\{\, U_j \,|\, j \in J \,\}$ eine beliebige offene Überdeckung von A. Da A nach Voraussetzung abgeschlossen ist, ist $X \setminus A$ offen und

$$\{\, X \setminus A \,\} \cup \{\, U_j \,|\, j \in J \,\} \quad \text{eine offene Überdeckung von } X\,.$$

Wenn X kompakt ist, existiert also eine endliche Teilüberdeckung, welche ohne $X \setminus A$ auch eine endliche Teilüberdeckung von A liefert. Also ist A kompakt.
(d): Dies zeigt man leicht durch Induktion nach n. Hier skizzieren wir den Beweis für $n = 2$. Sei $\{\, U_j \,|\, j \in J \,\}$ eine offene Überdeckung von $A_1 \cup A_2$. Dann gibt es zwei Teilmengen $J_1, J_2 \subset J$ so daß $J_1 \cup J_2 = J$ und $\{\, U_j \,|\, j \in J_l \,\}$ jeweils eine offene Überdeckung von A_l, $l = 1, 2$ liefert. Da beide Mengen kompakt sind, gibt es jeweils eine endliche Teilüberdeckung $A_l \subset U_{j_{l,1}} \cup \ldots \cup U_{j_{l,n_l}}$, $l = 1, 2$. Dann ist $U_{j_{1,1}} \cup \ldots \cup U_{j_{1,n_1}} \cup U_{j_{2,1}} \cup \ldots \cup U_{j_{2,n_2}}$ eine endliche Teilüberdeckung von $A_1 \cup A_2$. □

VI.7.7. Definition.
Es sei (X, d) ein metrischer Raum.
(X, d) heißt **folgenkompakt**, wenn jede Folge $(x_n) \in X^{\mathbb{N}}$ eine konvergente Teilfolge hat.
Eine Teilmenge $D \subset X$ heißt **dicht** in X, wenn ihr Abschluß ganz X ergibt, $\overline{D} = X$, also für alle $x \in X$ und $\epsilon > 0$ existiert ein $y \in D$ mit $d(x, y) < \epsilon$.
Der metrische Raum (X, d) heißt **separabel**, wenn er eine abzählbare dichte Teilmenge besitzt.

VI.7.8. Satz. *Jeder kompakte metrische Raum (X, d) ist folgenkompakt.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine beliebige Folge in X. Zu zeigen ist, daß (x_n) eine konvergente Teilfolge also einen Häufungspunkt in X hat. Angenommen das ist falsch.
Dann existiert zu jedem $x \in X$ eine offene Umgebung $U(x)$, die nur endlich viele Elemente von (x_n) enthält. Wir betrachten die somit gewonnene offene Überdeckung $\bigcup_{x\in X} U(x)$ von X. Aufgrund der Kompaktheit von X existieren endlich viele Punkte $p_1, \ldots, p_n \in X$ mit $U(p_1) \cup \ldots \cup U(p_n) = X$. Diese endlich vielen Umgebungen enthalten nach Konstruktion insgesamt nur endlich viele Folgenglieder von (x_n). Also kann diese Folge nur aus endlich viel verschiedenen Punkte bestehen und hat somit eine konvergente Teilfolge im Widerspruch zur Annahme. □

Etwas aufwendiger ist es zu zeigen, daß auch die Umkehrung gilt. Die Folgenkompaktheit ist also für metrische Räume äquivalent zur Kompaktheit im Sinne der Überdeckungseigenschaft.

VI.7.9. Satz. *Jeder folgenkompakte metrische Raum (X, d) ist kompakt.*

Beweis. <u>*1. Schritt: Beh.*</u>: Zu jedem $\epsilon > 0$ gibt es eine endliche Überdeckung durch Bälle des Radius ϵ, also $\forall\ \epsilon > 0\ \exists\ x_1, \ldots, x_n \in X$, $n = n(\epsilon)$ mit $X = B_\epsilon(x_1) \cup \ldots \cup B_\epsilon(x_n)$.
Bew.: Sei $\epsilon > 0$ und $x_1 \in X$ beliebig. Dann wählen wir rekursiv x_n, $n \in \mathbb{N}$ mit $x_n \in X \setminus \big(\bigcup_{i=1}^{n-1} B_\epsilon(x_i)\big)$. Also gilt $x_j \notin B_\epsilon(x_i)$ für alle $j \neq i$. Das bedeutet $d(x_i, x_j) \geq \epsilon$ für alle $i \neq j$ und somit kann $(x_n)_{n\in\mathbb{N}}$ keine konvergente Teilfolge haben, was der vorausgesetzten Folgenkompaktheit widerspricht. Also muss die rekursive Konstruktion von x_n nach endlich vielen Schritten abbrechen, und die Behauptung folgt.
<u>*2. Schritt: Beh.*</u>: (X, d) ist separabel, also es existiert eine abzählbare dichte Teilmenge $D \subset X$.
Gemäß Schritt 1 existiert zu jedem $n \in \mathbb{N}$ eine endliche Menge $D_n \subset X$ mit $\bigcup\{\, B_{\frac{1}{n}}(x) \mid x \in D_n \,\} = X$. Sei $D = \bigcup_{n\in\mathbb{N}} D_n$, also als abzählbare Vereinigung wieder eine abzählbare Menge. Sei nun $y \in X$ beliebig und $\epsilon > 0$. Dann existiert ein $n \in \mathbb{N}$ und ein $x \in D_n$ mit $\frac{1}{n} < \epsilon$ und $y \in B_{\frac{1}{n}}(x)$, also $d(y, x) < \epsilon$, $x \in D$. Somit ist D dicht in X.
<u>*3. Schritt*</u>: Betrachte die folgende Menge spezieller Bälle $S = \big\{\, B_r(x) \mid x \in D,\ r \in \mathbb{Q} \,\big\}$. Also ist S eine abzählbare Menge von Bällen.
Beh.: Jede offene Menge in $\emptyset \neq U \subset X$ ist eine Vereinigung spezieller Bälle aus S.
Bew.: Sei $y \in U$. Dann existiert ein $r_y \in \mathbb{Q}$ mit $B_{2r_y}(y) \subset U$. Da D eine dichte Menge ist, existiert ein $x_y \in B_{r_y}(y) \cap D$. Somit gilt $y \in B_{r_y}(x_y) \subset B_{2r_y}(y) \subset U$. Also folgt $U = \bigcup \big\{\, B_{r_y}(x_y) \mid y \in U \,\big\}$.
Der zentrale Teil des Beweises ist nun folgender
<u>*4. Schritt: Beh.*</u>: Jede offene Überdeckung von X durch spezielle Bälle enthält eine endliche Teilüberdeckung.
Bew.: Angenommen das ist falsch. Betrachte eine Abzählung $S = (B_n)_{n\in\mathbb{N}}$. Nach Annahme kann keine endliche Menge von speziellen Bällen X überdecken, also

$$B_1 \cup \ldots \cup B_n \subsetneqq X \quad \text{f.a. } n \in \mathbb{N}\,.$$

Sei dann $C_n = X \setminus (B_1 \cup \ldots \cup B_n)$. Wir wissen also, daß für alle $n \in \mathbb{N}$ die Mengen C_n nichtleer und abgeschlossen sind und $C_{n+1} \subset C_n$ f.a. $n \in \mathbb{N}$. Somit finden wir eine Folge $(c_n) \in X^{\mathbb{N}}$ mit $c_n \in C_n$ f.a. $n \in \mathbb{N}$. Nach Voraussetzung hat nun (c_n) eine konvergente Teilfolge, $z_{n_k} \to z_\infty$ für $k \to \infty$. Da alle C_n abgeschlossen sind und alle bis auf endliche viele Glieder der Teilfolge in C_n liegen, gilt $c_\infty \in \bigcap_{n\in\mathbb{N}} C_n \neq \emptyset$. Aber das bedeutet $c_\infty \notin B_n$ für alle $n \in \mathbb{N}$, so daß S keine Überdeckung von X sein kann, im Widerspruch zur Konstruktion von S.
<u>5. Schritt</u>: Sei $\{\, U_j \mid j \in J \,\}$ eine gegeben offene Überdeckung von X. Nach Schritt 3 ist jede offene Menge U_j eine abzählbare Vereinigung von speziellen Bällen,

$$U_j = \bigcup_{k=1}^{n_j} B_{k,j}, \quad B_{k,j} \in S\,.$$

Wir erhalten hieraus eine bestimmte abzählbare Überdeckung von X durch spezielle Bälle $\{\, B_{1,j_1}, \ldots, B_{n_1,j_1}, B_{1,j_2}, \ldots\}$. Nach Schritt 4 hat diese eine endliche Teilüberdeckung $\{\, B_{i,j'_1}, \ldots, B_{i',j'_r} \,\}$, woraus folgt daß auch die entsprechenden $U_{j'_1}, \ldots, U_{j'_r}$ eine endliche Überdeckung von X darstellen.
Also ist X kompakt □

VI.7.10. Bemerkung.
- X kompakt $\Rightarrow$ X separabel
- $\mathbb{R}^n$ ist nicht kompakt aber separabel, da $\mathbb{Q}^n \subset \mathbb{R}^n$ dicht und abzählbar, $\{\, B_{\frac{1}{n}}(q) \mid q \in \mathbb{Q}^n \,\}$ ist eine Familie spezieller Bälle.
- $(0, 1]$ ist nicht kompakt, da z.B. $(\frac{1}{n})_{n\in\mathbb{N}}$ keinen Häufungspunkt in $(0, 1])$ hat.

VI.7.11. DEFINITION.
Ein metrischer Raum (X, d) hat die **Heine-Borel-Eigenschaft**, wenn für alle Teilmengen $A \subset X$ gilt: A ist genau dann kompakt, wenn A in X abgeschlossen und beschränkt ist.

VI.7.12. SATZ. *Der metrische Raum $(\mathbb{R}^n, d_\infty)$ hat die Heine-Borel-Eigenschaft.*

BEWEIS. Wir wissen bereits, daß "abgeschlossen und beschränkt" aus der Kompaktheit folgt. Sei nun umgekehrt $A \subset \mathbb{R}^n$ abgeschlossen und beschränkt und $(x_k) \in A^{\mathbb{N}}$ eine beliebige Folge in A. Wir müssen nun die Existenz einer konvergenten Teilfolge zeigen.
Da A beschränkt ist, existiert ein $M > 0$ so daß für alle Koordinatenfolgen $(x_k^i)_{k\in\mathbb{N}}$, $i = 1, \dots, n$ gilt: $\max_{i=1,\dots,n} |x_k^i| \le M$ f.a. $k \in \mathbb{N}$. Anwendung des Auswahlsatzes von Bolzano liefert eine konvergente Teilfolge der ersten Koordinatenfolge

$$x_{k_l}^1 \to x_\infty^1 \quad \text{für } \ l \to \infty\,.$$

Wir ersetzen die ursprügliche Punktefolge $(x_k)_{k\in\mathbb{N}}$ durch die entsprechende Teilfolge $(x_{k_l})_{l\in\mathbb{N}}$ und betrachten nun daraus die Folge der zweiten Koordinaten, (x_k^2). Auch diese hat nach Bolzano eine in $\mathbb{R}$ konvergente Teilfolge. Insgesamt n-mal wählen wir eine Teilfolge der jeweils vorigen (Teil-)folge. Nach diesen endlich vielen Teilfolgenwahlen, bleibt eine Teilfolge übrig, die hinsichtlich aller n Koordinaten konvergiert. Also konvergiert dann auch die entsprechende Punktfolge in $\mathbb{R}^n$ und somit auch in A, da dieses als abgeschlossen vorausgesetzt ist. $\square$

VI.7.13. SATZ. *Bilder kompakter Mengen unter stetigen Abbildungen zwischen metrischen Räumen sind wieder kompakt.*

BEWEIS. Sei $f \colon X \to Y$ stetig und $A \subset X$ kompakt. Betrachte $f(A) \subset Y$ und $\{\, V_j \mid j \in J\}$ eine offene Überdeckung von $f(A)$. Gemäß Satz VI.5.5 sind die Urbilder $U_j = f^{-1}(V_j) \subset X$ offen in X, also ist $\{\, U_j \mid j \in J\,\}$ eine offene Überdeckung von A. Da A kompakt ist, existiert eine endliche Teilüberdeckung $\{\, U_{j_1}, \dots, U_{j_r}\,\}$. Dann ist auch $\{\, V_{j_1} = f(U_{j_1}), \dots, V_{j_r} = f(U_{j_r})\,\}$ eine endliche Teilüberdeckung von $f(A)$. $\square$

Eine wichtige Konsequenz hieraus ist

VI.7.14. KOROLLAR. *$f \in C^0(X, \mathbb{R})$, X kompakt $\Rightarrow$ $\max_{x\in X} f$ und $\min_{x\in X} f$ existieren, also jede stetige Funktion auf einem kompakten Raum hat ein Maximum und ein Minimum.*

BEWEIS. $f(X) \subset \mathbb{R}$ ist eine kompakte Menge von reellen Zahlen, also inbesondere beschränkt und abgeschlossen. Daher existieren hiervon Supremum und Infimum und gehören selbst zu dem Wertebereich. $\square$

VI.7.1. Anwendungen. Eine schöne Anwendung der Existenz von Maximum und Minimum auf kompakten Mangen ist nun endlich der Beweis des Satzes VI.1.7: Je zwei Normen auf $\mathbb{R}^n$ sind äquivalent.

BEWEIS VON SATZ VI.1.7. Sei $\|\cdot\|$ eine beliebige Norm auf $\mathbb{R}^n$ und $e_1, \dots, e_n \in \mathbb{R}^n$ die kanonische Basis. Wir setzen $c = \max_{i=1,\dots,n} \|e_i\|$. Dann folgt für alle $x = (x^1, \dots, x^n) \in \mathbb{R}^n$

$$\|x\| \le \sum_{i=1}^n |x^i| \|e_i\| \le cn\|x\|_\infty\,. \tag{10}$$

Sei nun S die Einheitssphäre in $\|\cdot\|_\infty$, $S = \{\, x \in \mathbb{R}^n \mid \|x\|_\infty \,\}$. Nach Satz VI.7.12 ist S kompakt in $(\mathbb{R}^n, \|\cdot\|_\infty)$. Die Funktion $f \colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|$ ist aufgrund

(10) stetig und besitzt dann nach Korollar VI.7.14 ein Maximum und ein Minimum, $M = \max_S f$, $m = \min_S f$. Dies bedeutet

$$m \leq \left\| \tfrac{x}{\|x\|_\infty} \right\| \leq M \quad \text{f.a. } x \in \mathbb{R}^n \setminus \{0\},$$

was äquivalent ist zu $m\,\|x\|_\infty \leq \|x\| \leq M\,\|x\|_\infty$ f.a. $x \in \mathbb{R}^n$. Somit ist die gegebene Norm $\|\cdot\|$ äquivalent zu $\|\cdot\|_\infty$. □

Eine weitere Anwendung ist der Satz von Ascoli-Arzela, welcher in der Theorie der Lösungen von Differentialgleichungen eine wichtige Rolle spielt.
In Vorbereitung hierauf haben wir zunächst

VI.7.15. SATZ. *Jede auf einem kompakten Raum X definierte stetige Abbildung $f\colon X \to Y$ ist gleichmäßig stetig, d.h. für alle $\epsilon > 0$ existiert ein $\delta = \delta(\epsilon) > 0$ so daß für alle $x, x' \in X$ mit $d_X(x, x') < \delta$ gilt: $d_Y(f(x), f(x')) < \epsilon$.*

Hier ist wichtig, daß δ nur von ϵ abhängt und nicht von $x \in X$.

BEWEIS. Sei $\epsilon > 0$ beliebig. Da f in jedem $x \in X$ stetig ist, existiert zunächst für jedes $x \in X$ ein δ_x in Abhängigkeit von ϵ und x so daß für alle $x' \in B_{\delta_x}(x)$ gilt: $d_Y(f(x), f(x')) < \frac{\epsilon}{2}$. Wir betrachten nun die offene Überdeckung $\{\, B_{\frac{1}{2}\delta_x}(x) \mid x \in X \,\}$ von X. Da X kompakt ist, existiert eine endliche Teilüberdeckung

$$X = B_{\frac{1}{2}\delta_{x_1}}(x_1) \cup \ldots \cup B_{\frac{1}{2}\delta_{x_r}}(x_r)\,.$$

Sei $\delta = \frac{1}{2}\min(\delta_{x_1}, \ldots, \delta_{x_r})$. Dann existiert für alle $x, x' \in X$ mit $d_X(x, x') < \delta$ ein $i \in \{1, \ldots, r\}$ mit $x, x' \in B_{\delta_{x_i}}(x_i)$ und somit gilt $d_Y(f(x), f(x_i)) < \frac{\epsilon}{2}$ und $d_Y(f(x'), f(x_i)) < \frac{\epsilon}{2}$, also nach Dreiecksungleichung $d_Y(f(x), f(x')) < \epsilon$. □

VI.7.16. DEFINITION.
Eine Teilmenge $A \subset X$ eines metrischen Raumes (X, d) heißt **relativ kompakt** in X, wenn ihr Abschluß $\overline{A}$ kompakt ist.

Aus den Sätzen VI.7.8 und VI.7.9 folgt, daß eine Teilmenge $A \subset X$ genau dann relativ kompakt ist, wenn jede Folge in A eine in X konvergente Teilfolge hat.

VI.7.17. DEFINITION.
Es seien X, Y zwei metrische Räume und $H \subset C^0(X, Y)$ eine Menge von stetigen Abbildungen $f\colon X \to Y$. Die Abbildungsmenge H heißt **gleichgradig stetig** in einem Punkt $x_o \in X$, wenn es zu jedem $\epsilon > 0$ ein $\delta = \delta(\epsilon, x_o)$ gibt, so daß aus $d_X(x, x_o) < \delta$ die Abschätzung $d_Y(f(x), f(x_o)) < \epsilon$ für alle $f \in H$ folgt. Das bedeutet, daß δ zwar von ϵ und dem Punkt x_o abhängt, aber gleich gültig für alle $f \in H$ ist. H heißt einfach nur gleichgradig stetig, wenn es in jedem Punkt $x \in X$ gleichgradig stetig ist.

VI.7.18. BEISPIEL.
Es sei $X = [0, 1]$, $Y = \mathbb{R}$ und $H \subset C^1([0, 1], \mathbb{R})$ mit

$$\|f'\|_\infty = \max_{x \in [0,1]} |f'(x)| \leq m \quad \text{für alle } \ f \in H\,.$$

Dann folgt aus dem Mittelwertsatz die Abschätzung

$$|f(x) - f(y)| \leq |f'(\xi)| \cdot |x - y| \quad \text{für ein } \xi \in (x, y) \text{ f.a. } 0 \leq x < y \leq 1\,.$$

Also sind alle $f \in H$ Lipschitz-stetig mit uniformer Lipschitz-Konstante m, insbesondere ist H gleichgradig stetig.

Der zentrale Satz ist nun

VI.7.19. THEOREM (Satz von Ascoli-Arzela). *Es seien (X, d_X) ein kompakter metrischer Raum, (Y, d_Y) ein vollständiger metrischer Raum und $H \subset C^0(X, Y)$ eine Teilmenge des metrischen Raumes stetiger Abbildungen mit der Supremums-Metrik*

$d_\infty(f,g) = \max_{x\in X} d_Y(f(x), g(x))$. *Dann gilt: H ist genau dann relativ kompakt in $\big(C^0(X,Y), d_\infty\big)$, wenn H gleichgradig stetig ist und für alle $x \in X$ die Wertemenge $\{\, f(x) \mid f \in H\,\}$ in Y relativ kompakt ist.*

VI.7.20. LEMMA. *X präkompakt und vollständig $\Rightarrow$ X kompakt.*

Beachte: *präkompakt* heißt, daß zu jedem $\epsilon > 0$ eine endliche Überdeckung durch ϵ-Bälle existiert.

BEWEIS. Sei $\{U_j\}_{j\in J}$ eine offene Überdeckung von X und sei angenommen, daß $\{U_j\}$ keine endliche Teilüberdeckung besitzt. Wir konstruieren wie folgt rekursiv eine Folge von offenen Bällen $(B_n)_{n\in\mathbb{N}}$:
Sei $\epsilon_1 > 0$ beliebig. Da X präkompakt ist, existieren $x_1^{(1)}, \ldots, x_{N_1}^{(1)}$ so daß

$$X \subset \bigcup_{i=1}^{N_1} B(x_i^{(1)}, \epsilon_1)\,.$$

Nach Annahme läßt sich mindestens einer dieser Bälle nicht durch endlich viele der vorgegebenen offenen Mengen U_j überdecken, d.h. ex. $i \in \{1, \ldots, N_1\}$ mit: Die Überdeckung $\{\, B(x_i^{(1)}, \epsilon_1) \cap U_j\}_{j\in J}$ von $B(x_i^{(1)}, \epsilon_1)$ besitzt keine endliche Teilüberdeckung. Wir setzen $B_1 := B(x_i^{(1)}, \epsilon_1)$ als den ersten Ball der gesuchten Folge fest.
Zur Rekursion: Sei nun B_{n-1} bereits gegeben. Dann betrachte $\epsilon_n = \frac{1}{2}\epsilon_{n-1}$ und wegen X präkompakt $x_1^{(n)}, \ldots, x_{N_n}^{(n)}$ mit $X \subset \bigcup_{i=1}^{N_n} B(x_i^{(n)}, \epsilon_n)$. Betrachte unter diesen Bällen nur diejenigen, welche B_{n-1} schneiden: $B_{n-1} \cap B(x_i^{(n)}, \epsilon_n) \neq \emptyset$. Für einen von ihnen, $B(x_k^{(n)}, \epsilon)$, muss wiederum gelten, daß keine endlich viele der U_j diesen überdecken: Die Überdeckung $\{\, B(x_k^{(n)}, \epsilon_n) \cap U_j\}_{j\in J}$ von $B(x_k^{(n)}, \epsilon_n)$ besitzt keine endliche Teilüberdeckung. Dann sei dies $B_n := B(x_k^{(n)}, \epsilon_n)$.
Wir erhalten also eine Folge von Bällen $(B_n)_{n\in\mathbb{N}}$ mit:

(a) Für die Radien ϵ_n gilt: $\epsilon_n = \frac{1}{2^{n-1}}\epsilon_1$.
(b) Je zwei aufeinanderfolgende Bälle haben nichtleeren Schnitt, $B_{n+1} \cap B_n \neq \emptyset$.
(c) Keiner der Bälle B_n besitzt eine endliche Teilüberdeckung aus $\{U_j\}_{j\in J}$.

Nach der Dreiecksungleichung angewandt auf (b) und wegen (a) folgt, daß die Folge der Mittelpunkte der Bälle B_n eine Cauchyfolge bildet, also wegen der Vollständigkeit von X konvergiert. Sei x_∞ der Grenzwert und U_j eine der offenen Mengen, welche x_∞ enthält. Dann enthält U_j auch fast alle (alle bis auf endlich viele) der Bälle B_n, im Widerspruch zu (c). □

BEWEIS DES SATZES VON ASCOLI-ARZELA.
Hinreichende Bedingung: Da $(C^0(X,Y), d_\infty)$ vollständig ist und somit auch der Abschluß von H darin, genügt für die hinreichende Bedingung nach Lemma VI.7.20 zu zeigen, daß H präkompakt ist. Sei also $\epsilon > 0$.
Da H gleichgradig stetig ist, existiert für alle $x \in X$ eine Umgebung $V(x) = B_{d_X}(x, \delta)$ in X, so daß

$$d_Y(f(x), f(y)) < \frac{\epsilon}{4} \quad \text{f.a. } f \in H,\ y \in V(x)\,.$$

Da X kompakt ist, läßt es sich durch endlich viele dieser Umgebungen überdecken,

$$X \subset \bigcup_{i=1}^{m} V(x_i)$$

Andererseits ist jedes $H(x_i) = \{ f(x_i) \,|\, f \in H \}$ in Y relativ kompakt und somit auch die endliche Vereinigung dieser

$$K = \bigcup_{i=1}^{m} H(x_i) \quad \text{rel. kp. in } Y\,.$$

Somit ist K ebenfalls präkompakt und es gibt eine endliche Menge $\{ y_j \,|\, 1 \le j \le n \} \in K$ so daß

$$K \subset \bigcup_{j=1}^{n} B_{d_Y}(y_j, \epsilon/4)\,.$$

Sei nun $S = \{1, \ldots, n\}^{\{1,\ldots,m\}}$ die endliche Menge aller Abbildung von $\{1, \ldots, m\}$ nach $\{1, \ldots, m\}$ und für jedes $\sigma \in S$ sei

$$L_\sigma := \{ f \in H \,|\, d_Y(f(x_i), y_{\sigma(i)}) \le \epsilon/4 \text{ f.a. } 1 \le i \le m \}\,.$$

Aus der Konstruktion der $\{y_j\}$ folgt, daß $\{ L_\sigma \}_{\sigma\in S}$ eine endliche Überdeckung von H bildet. Es bleibt also zu zeigen, daß jedes L_σ in einem Ball vom Radius ϵ liegt: Seien $f, g \in L_\sigma$. Für alle $x \in X$ ex. x_i mit $x \in V(x_i)$, also

$$d_Y(f(x), f(x_i)) < \epsilon/4 \text{ und } d_Y(g(x), g(x_i)) < \epsilon/4\,.$$

Da wegen $f, g \in H_\sigma$, $d_Y(f(x_i), g(x_i)) \le \epsilon/2$, so gilt auch $d_Y(f, g) < \epsilon$.
Notwendige Bedingung: Sei H relativ kompakt. Dann existieren für jedes $\epsilon > 0$ endlich viele $f_1, \ldots, f_n \in H$, so daß für alle $f \in H$ ein $i \in \{1, \ldots, n\}$ existiert mit $d_\infty(f, f_i) < \epsilon/3$, d.h. $d_Y(f(x), f_i(x)) < \epsilon/3$ f.a. $x \in X$. Somit ist jedes $H(x)$ in Y relativ kompakt.
Es existiert außerdem f.a. $x \in X$ eine Umgebung $V(x)$ so daß $d_Y(f_i(y), f_i(x)) < \epsilon/3$ f.a. $y \in V(x)$ und f.a. $i = 1, \ldots, n$. Dann gilt nach Konstruktion der f_i auch $d_Y(f(y), f(x)) < \epsilon$ f.a. $f \in H$ und $y \in V(x)$, d.h. H ist gleichgradig stetig. □

Eine weitere Anwendung der Eigenschaft kompakter Mengen ist das folgende sogenannte Tuben-Lemma:

VI.7.21. Lemma (Tuben-Lemma). *Sei X ein metrischer Raum, K ein kompakter metrischer Raum, $x_o \in X$ und $W \subset X \times K$ offen, so daß*

$$\{x_o\} \times K \subset W\,.$$

Dann existiert eine offene Umgebung U von x_o in X, so daß

$$U \times K \subset W\,.$$

Beweis. Da W offen ist, existiert zu jedem $k \in K$ eine Umgebung von (x_o, k) innerhalb W von Produktform, d.h. es existiert $V(k)$ eine offene Umgebung von k in K und ein Ball $B_{\epsilon(k)}(x_o)$ so daß

$$(x_o, k) \in B_{\epsilon(k)}(x_o) \times V(k) \subset W\,.$$

Dann ist $\bigcup_{k\in K} V(k) = K$ eine offene Überdeckung, welche aufgrund der Kompaktheit von K eine endliche Teilüberdeckung $V(k_1) \cup \ldots \cup V(k_n) = K$ besitzt. Sei dann $\epsilon = \min\big(\epsilon(k_1), \ldots, \epsilon(k_n)\big)$. Dann ist $U = B_\epsilon(x_o)$ die gesuchte offene Menge,

$$B_\epsilon(x_o) \times K \subset W\,.$$

□

Hieraus ergibt sich folgende wichtige Anwendung für parameter-abhängige Integrale.

VI.7.22. Satz. *Es sei X ein metrischer Raum, $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall und $f\colon X \times [a, b] \to \mathbb{R}$ stetig. Dann ist*

$$\varphi\colon X \to \mathbb{R}, \quad \varphi(x) = \int_a^b f(x,t)dt$$

stetig.

Beweis. Es sei $x_o \in X$ und $h\colon X \times [a, b] \to \mathbb{R}$ gegeben durch $h(x, t) = f(x, t) - f(x_o, t)$. Sei $\epsilon > 0$. Dann ist

$$W = \big\{ (x, t) \in X \times [a, b] \,|\, |h(x, t)| < \epsilon \big\}$$

eine offene Umgebung von $\{x_o\} \times [a, b]$. Nach dem Tubenlemma existiert eine Umgebung $U(x_o) \subset X$ so daß $U(x_o) \times [a, b] \subset W$ und somit

$$\big|f(x, t) - f(x_o, t)\big| < \epsilon \quad \text{für alle } \; x \in U(x_o),\, t \in [a, b]\,.$$

Hieraus folgt

$$\big|\varphi(x) - \varphi(x_o)\big| \; < \; \int_a^b \epsilon\, dt \; = \; \epsilon \cdot (b - a) \quad \text{f.a. } \; x \in U(x_o)\,.$$

Also ist φ stetig in x_o. □

VI.8. ∗ Vervollständigung metrischer Räume

Wir haben in Satz VI.4.6 gesehen, daß abgeschlossene Teilmengen vollständiger Räume wieder vollständig sind. Das heißt, daß sich beliebige Teilmengen vollständiger Räume immer mittels des Abschlusses vervollständigen lassen. Durch den Abschluss werden alle möglicherweise noch nicht enthaltenen Grenzwerte von nicht-konvergenten Cauchy-Folgen hinzugefügt.

Handelt es sich nun aber um einen beliebigen, möglicherweise nicht vollständigen metrischen Raum, ist die Frage nach einer Vervollständigung konzeptionell schwieriger. Nun fehlt ein umgebender größerer Raum, in dem die hinzuzufügenden Grenzwerte nicht-konvergenter Cauchy-Folgen lokalisiert werden können.

Für die Vervollständigung metrischer Räume brauchen wir ein neues Konzept, wie wir durch Daten aus dem gegebenen Raum heraus neue Punkte hinzufügen und den Raum erweitern. Dieses Problem ist vergleichbar mit dem Schritt der Erweiterung des Zahlenraums der rationalen Zahlen zu den reellen Zahlen, hierbei handelt es sich genau um eine solche metrische Vervollständigung.

Sei (X, d) ein metrischer Raum. Wir wiederholen folgende wichtige Begriffe:

- (X, d) heißt *vollständig*, wenn für jede Folge $(x_n) \in X^{\mathbb{N}}$ gilt: Ist (x_n) eine Cauchy-Folge, so ist sie in X konvergent.
- Eine Teilmenge $D \subset X$ heißt *dicht*, wenn ihr Abschluss ganz X ergibt.

VI.8.1. Definition.
Es seien (x, d_X) und (Y, d_Y) zwei metrische Räume. Dann heißt eine Abbildung $f\colon X \to Y$ eine **isometrische Einbettung**, wenn für alle $x_1, x_2 \in X$ gilt:

$$d_Y\big(f(x_1), f(x_2)\big) \; = \; d_X(x_1, x_2)\,.$$

Eine surjektive isometrische Einbettung heißt **Isometrie**.

VI.8.2. Satz. *Zu jedem metrischen Raum X gibt es einen vollständigen metrischen Raum $\hat{X}$ und eine isometrische Einbettung $i\colon X \hookrightarrow \hat{X}$, so daß $i(X) \subset \hat{X}$ dicht ist.*

VI.8.3. Definition.
Eine solche isometrische, dichte Einbettung $i\colon X \hookrightarrow \hat{X}$ in einen vollständigen metrischen Raum $\hat{X}$ heißt **Vervollständigung** von X.

VI.8.4. SATZ. *Je zwei Vervollständigungen von X sind isometrisch.*

Dies bedeutet, daß zu je zwei Vervollständigungen $i_1 \colon X \hookrightarrow \hat{X}_1$ und $i_2 \colon X \hookrightarrow \hat{X}_2$ eine Isometrie $j \colon \hat{X}_1 \to \hat{X}_2$ existiert, so daß $i_2 = j \circ i_1$. Diese Isometrie ist dann auch eindeutig bestimmt. Also ist eine Vervollständigung bis auf Isometrie eindeutig.

VI.8.5. BEISPIEL.
(1) $(\mathbb{Q}, |\cdot|) \subset (\mathbb{R}, |\cdot|)$.
(2) $\big(\{(x,y) \,|\, x^2+y^2<1\}, d_2\big) \subset \big(\{(x,y) \,|\, x^2+y^2 \leq 1\}, d_2\big)$.
(3) $X = \big(C^0([0,1],\mathbb{R}), \|\cdot\|_p\big)$ mit $\|f\|_p = \big(\int_0^1 |f(x)|^p dx\big)^{1/p}$ für $p \geq 1$. Dann wird die Vervollständigung $\hat{X}$ mit $L^p([0,1],\mathbb{R})$ bezeichnet. $\|\cdot\|_p$ liefert auf L^p eine vollständige Norm. In der Vorlesung Maß- und Integrationstheorie werden diese L^p-Räume eingehender untersucht.

Sei nun (X,d) ein beliebiger metrischer Raum.

VI.8.6. HILFSSATZ. *Für je zwei Cauchyfolgen (x_n) und (y_n) in X existiert der Grenzwert $\lim_{n\to\infty} d(x_n,y_n)$.*

BEWEIS. Aus der Cauchyfolgen-Eigenschaft folgt daß zu jedem $\epsilon > 0$ $n_1, n_2 \in \mathbb{N}$ existieren so daß

$$d(x_n,x_m) < \tfrac{\epsilon}{2} \text{ f.a. } m,n \geq n_1 \text{ und } d(y_n,y_m) < \tfrac{\epsilon}{2} \text{ f.a. } m,n \geq n_2 \,.$$

Sei nun $n_o = \max(n_1,n_2)$, dann folgt aus der Dreiecksungleichung

$$\big|d(x_n,y_n) - d(x_m,y_m)\big| \;\leq\; d(x_n,x_m) + d(y_n,y_m) < \epsilon \quad \text{f.a. } \; m,n \geq n_o \,.$$

Also ist $\big(d(x_n,y_n)\big)n \in \mathbb{N}$ eine Cauchy-Folge in $\mathbb{R}$ und die Konverenz folgt aus der Vollständigkeit von $\mathbb{R}$. □

Wir benutzen diesen Sachverhalt für folgende

VI.8.7. DEFINITION.
Betrachte $CF(X) := \big\{ (x_n) \in X^{\mathbb{N}} \,|\, (x_n) \text{ ist C.F.} \big\}$. Wir definieren auf $CF(X)$ die Relation

$$(x_n) \sim (y_n) \overset{\text{def}}{\Leftrightarrow} \lim_{n\to\infty} d(x_n,y_n) = 0 \,.$$

Gemäß Hilfssatz VI.8.6 ist die Relation $\sim$ auf $CF(X)$ wohldefiniert. Der Beweis des nächsten Hilfssatzes ist eine leichte Übung.

VI.8.8. HILFSSATZ. *$\sim$ ist eine Äquivalenzrelation.*

VI.8.9. DEFINITION.
Sei $(x_n) \in CF(X)$, dann bezeichnen wir die Äquivalenzklasse bzgl. $\sim$ mit

$$[x_n] := \big\{ (y_n) \in CF(X) \,|\, (y_n) \sim (x_n) \big\},$$

und die Menge aller Äquivalenzklassen mit $\hat{X} := \big\{ [x_n] \,\big|\, (x_n) \in CF(X) \big\}$.

Wir stellen hierzu folgendes fest:

(1) Zu jedem $x \in X$ erhalten wir die Äquivalenzklasse $[x] \in \hat{X}$ der konstanten Folge $x_n = x$ f.a. $n \in \mathbb{N}$. Die Äquivalenzklasse der konstanten Folge $(x) \in X^{\mathbb{N}}$ ist genau die Menge alle der Folgen $(x_n) \in X^{\mathbb{N}}$, die gegen x konvergieren. Wir erhalten also eine wohldefinierte Abbildung

$$i \colon X \to \hat{X}, \quad x \mapsto [x]\,,$$

und das Bild $i(X)$ entspricht genau der Menge der Äquivalenzklassen der konvergenten Cauchy-Folgen.

(2) Die Äquivalenzklassen zu verschiedenen Punkten $x \neq y$ sind verschieden, $[x] \neq [y]$. Also ist i injektiv. Wir können also X mit $i(X)$ als Menge identifizieren.

(3) Wenn (X, d) bereits vollständig ist, ist jede Cauchy-Folge konvergent und nach Definition gilt $i(X) = \hat{X}$.
(4) Falls (X, d) nicht vollständig ist, enthält $\hat{X} \backslash \{i(X)\}$ genau die Äquivalenzklassen der nicht-konvergenten Cauchy-Folgen

Das Ziel ist es nun, auf $\hat{X}$ eine Metrik zu finden, bezüglich der $\hat{X}$ vollständig ist und so daß i eine dichte, isometrische Einbettung wird.

VI.8.10. Hilfssatz. *Seien $\xi, \eta \in \hat{X}$ repräsentiert durch Cauchy-Folgen (x_n) und (y_n). Dann ist $\hat{d}(\xi, \eta) := \lim_{n\to\infty} d(x_n, y_n)$ wohldefiniert, d.h. es hängt nur von ξ und η aber nicht von der Wahl der Repräsentanten (x_n), (y_n) ab.*

Beweis. Seien $\xi = [x_n] = [x'_n]$. Dann folgt aus der Dreiecksungleichung

$$\lim d(x_n, y_n) \leq \lim d(x_n, x'_n) + \lim d(x'_n, y_n) = 0 + \lim d(x'_n, y_n)\,,$$

und umgekehrt auch $\lim d(x'_n, y_n) \leq \lim d(x_n, y_n)$, also $\lim d(x'_n, y_n) = \lim d(x_n, y_n)$. Da die Metrik d symmetrisch ist, folgt daraus auch die Unabhängigkeit von der Wahl des Repräsentanten (y_n). □

VI.8.11. Hilfssatz. *$\hat{d}\colon \hat{X} \times \hat{X} \to [0, \infty)$ ist eine Metrik auf $\hat{X}$.*

Beweis. Dies ist eine leichte Übung. Man verwendet im wesentlichen die Dreiecksungleichung und die Definition der Äquivalenzrelation $\sim$. □

Der wesentliche Punkt der Konstruktion von $\hat{X}$ liegt in dem folgenden

VI.8.12. Hilfssatz. *$(\hat{X}, \hat{d})$ ist ein vollständiger metrischer Raum.*

Beweis. Sei $(\xi_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\hat{X}$. Jedes einzelne Folgenglied $\xi_n \in \hat{X}$ ist also selbst eine Äquivalenzklasse $\xi_n = [x_{n1}, x_{n2}, \ldots]$ einer Cauchy-Folge $(x_{nk})_{k\in\mathbb{N}}$ in X.
Wir bemerken dabei:

(a) Da $(x_{nk})_{k\in\mathbb{N}}$ eine Cauchy-Folge in X ist, existiert zu jedem $\epsilon = \frac{1}{n}$ ein $k(n) \in \mathbb{N}$ so daß $d(x_{ni}, x_{nj}) < \frac{1}{n}$ für alle $i, j \geq k(n)$ gilt.
(b) Da $(\xi_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\hat{X}$ ist, existiert zu jedem $\epsilon > 0$ ein $n(\epsilon) \in \mathbb{N}$ derart, daß $\hat{d}(\xi_n, \xi_m) < \frac{\epsilon}{4}$ für alle $m, n \geq n(\epsilon)$ gilt.
(c) Aus (2) folgt somit, daß für alle $m, n \geq n(\epsilon)$ ein $N(m, n) \in \mathbb{N}$ existiert, so daß $d(x_{nk}, x_{mk}) < \frac{\epsilon}{2}$ für alle $k \geq N(m, n)$ gilt, denn $\hat{d}(\xi_m, \xi_n) = \lim_{k\to\infty} d(x_{nk}, x_{mk})$.

Wir betrachten nun $\zeta = \big(x_{1k(1)}, x_{2k(2)}, \ldots\big)$ mit $k(n)$ aus (a). Wir behaupten, daß ζ wiederum eine Cauchy-Folge in X ist. Sei nämlich $\epsilon > 0$ und $n_o(\epsilon) = \max\big(n(\epsilon), \frac{4}{\epsilon}\big)$ mit $n(\epsilon)$ aus (b). Dann gilt für $n, m \geq n_o(\epsilon)$ die Abschätzung

$$\begin{aligned} d\big(x_{nk(n)}, x_{mk(m)}\big) &\leq d\big(x_{nk(n)}, x_{ni}\big) + d\big(x_{ni}, x_{mi}\big) + d\big(x_{mi}, x_{mk(m)}\big) \\ &< \frac{1}{n} + \frac{\epsilon}{2} + \frac{1}{m} \leq \frac{\epsilon}{4} + \frac{\epsilon}{2} + \frac{\epsilon}{4} < \epsilon\,. \end{aligned} \tag{11}$$

Die Abschätzung für den ersten Summanden gilt aufgrund von (a), sofern $i \geq k(n)$, die für den zweiten Summanden aufgrund von (c), sofern $i \geq N(n, m)$, und die für den dritten Summanden wieder aufgrund von (a), sofern $i \geq k(m)$. Da aber die linke Seite nicht von i abhängt können wir i ahängig von m und n groß genug wählen, so daß ϵ eine obere Schranke ist und somit ζ eine Cauchy-Folge.
Also stellt $[\zeta]$ ein Element von $\hat{X}$ dar, und wir behaupten des weiteren, daß

$$\lim_{n\to\infty} \xi_n = \zeta\,.$$

Dies sieht man folgendermaßen: Sei $\epsilon > 0$ gegeben und wähle $n_o(\epsilon) = \max\big(n(\epsilon), \frac{4}{\epsilon}\big)$ mit $n(\epsilon)$ aus obigem (b). Seien nun $n \geq n_o(\epsilon)$ und $l \geq \max\big(k(n), n_o(\epsilon)\big)$. Dann folgt

$$d\big(x_{nl}, x_{lk(l)}\big) \;\leq\; d\big(x_{nl}, x_{nk(n)}\big) \;+\; d\big(x_{nk(n)}, x_{lk(l)}\big) \;<\; \frac{1}{n} + \epsilon,$$

wegen $l \geq k(n)$ und (11). Also gilt

$$\hat{d}\big(\xi_n, [\zeta]\big) = \lim_{l\to\infty} d(x_{nl}, x_{lk(l)}) \;\leq\; \frac{1}{n} + \epsilon\,.$$

Das impliziert $\lim_{n\to\infty} \hat{d}(\xi_n, [\zeta]) \leq \epsilon$ für alle $\epsilon > 0$ und daher $\xi_n \to \zeta$.
Somit haben wir gezeigt, daß jede beliebige Cauchy-Folge $(\xi_n) \in \hat{X}^{\mathbb{N}}$ konvergiert, also $\hat{X}$ vollständig ist. □

Seien $x, y \in X$ zwei beliebige Punkte und $[x], [y] \in \hat{X}$ die zugehörigen Äquivalenzklassen der konstanten Folgen. Dann gilt nach Definition

$$\hat{d}\big([x], [y]\big) \;=\; \lim_{n\to\infty} d(x, y) = d(x, y)\,.$$

Das bedeutet, daß $i\colon X \hookrightarrow \hat{X}$ eine isometrische Einbettung ist.

VI.8.13. Hilfssatz. *Das Bild $i(X)$ liegt dicht in $\hat{X}$.*

Beweis. Sei $\xi\hat{X}$ eine beliebige Äquivalenzklasse $\xi = [(x_n)]$. Sei außerdem $\epsilon > 0$ gegeben. Dann existiert ein $n_o(\epsilon)$ mit $d(x_n, x_m) < \frac{\epsilon}{2}$ für alle $m, n \geq n_o(\epsilon)$, da (x_n) nach Voraussetzung eine Cauchy-Folge ist.
Sei $x = x_{n_o} \in X$. Dann gilt also $d(x, x_n) < \frac{\epsilon}{2}$ für alle $n \geq n_o(\epsilon)$ und daher $\hat{d}\big([x], \xi\big) = \lim_{n\to\infty} d(x, x_n) < \epsilon$, das heißt $[x] \in B_{\hat{d}}(\xi, \epsilon) \cap i(X)$. □

Ingesamt haben wir also Satz VI.8.2 bewiesen. □
Satz VI.8.4 folgt unmittelbar aus folgendem allgemeineren

VI.8.14. Satz. *Sei $A \subset (X, d)$ eine dichte Teilmenge eines metrischen Raumes und $f\colon A \to Y$ eine gleichmäßig stetige Abbildung in einen vollständigen metrischen Raum Y. Dann existiert eine eindeutige stetige Fortsetzung $F\colon X \to Y$ $F_{|A} = f$.*

Diesen Satz haben wir in konkreten Fällen bereits in der Analysis in einer reellen Variablen mehrmals angewandt, nämlich bei der Definition der Exponentialfunktion $a \mapsto a^x$ und des Logarithmus $x \mapsto \ln x$ für reelle bzw. positive reelle Zahlen. Wir hatten die Funktion zuerst für natürliche, dann mittels der jeweiligen Funktionalgleichung für rationale Zahlen fortgesetzt und schließlich eine eindeutige Fortsetzung auf $\mathbb{R}$ gefunden, da die Funktionen in den rationalen Variablen stetig waren und somit gleichmäßig stetig auf allen Schnittmengen mit kompakten Intervallen.

Beweis. Sei $x \in X$. Da $A \subset X$ dicht liegt, finden wir eine Folge $(a_n) \in A^{\mathbb{N}}$ mit $a_n \to x$. Da f gleichmäßig stetig ist, bildet f Cauchy-Folgen in A auf Cauchy-Folgen in Y ab. Da Y vollständig ist, konvergiert $(f(a_n))$. Für je zwei Folgen $a_n \to x$ und $a'_n \to x$ gilt $d(a_n, a'_n) \to 0$ und somit wegen der gleichmäßigen Stetigkeit wiederum $\lim f(a_n) = \lim f(a_n)$. Also ist $F(x) = \lim f(a_n)$ wohldefiniert.
Sei nun $\epsilon > 0$ und $\delta > 0$, so daß $d_Y(f(a), f(a')) < \frac{\epsilon}{2}$ für alle $a, a' \in A$ mit $d_X(a, a') < \delta$. Dann gilt für alle $x, x' \in X$ mit $d_X(x, x') < \frac{\delta}{2}$, daß wir zum einen $a, a' \in A$ finden mit $d_X(x, a) < \frac{\delta}{4}$ und $d_X(x', a') < \frac{\delta}{4}$ und außerdem auch $d_Y(f(x), f(a)) < \frac{\epsilon}{4}$ und $d_Y(f(x'), f(a')) < \frac{\epsilon}{4}$. Dann gilt wegen der Dreiecksungleichung $d_X(a, a') < \delta$ und somit auch $d_Y(f(x), f(x')) < \epsilon$. Also ist auch die Fortsetzung F gleichmäßig stetig.
Da je zwei stetige Abbildungen, die auf einer dichten Teilmenge übereinstimmen, identisch sind, ist F als stetige Fortsetzung somit eindeutig. □

VI.9. Zusammenhang

VI.9.1. Definition.
Ein metrischer Raum (X, d) heißt **zusammenhängend**, falls er sich nicht in zwei nicht-leere disjunkte offene Mengen zerlegen lässt, d.h. falls $X = U \cup V$ mit $U \cap V = \emptyset$, U, V offen in X so ist $U = \emptyset$ oder $V = \emptyset$.
Ein metrischer Raum heißt **wegzusammenhängend**, falls es zu je zwei Punkten $p, q \in X$ eine stetige Abbildung *(Weg)* $c\colon [0,1] \to X$ mit $c(0) = p$ und $c(1) = q$ gibt.

VI.9.2. Satz. *Bilder (weg-)zusammenhängender Räume unter stetigen Abbildungen sind wieder (weg-)zusammenhängend.*

Beweis. Die Aussage für Wegzusammenhang folgt unmittelbar daraus, daß Kompositionen stetiger Abbildungen wieder stetig sind.
Sei also nun X zusammenhängend und $f\colon X \to Y$ stetig. Um zu zeigen, daß auch $f(X)$ zusammenhängend ist, nehmen wir an, daß es zwei offene Mengen $U, V \subset Y$ gibt mit $U \cap V \cap f(X) = \emptyset$ und $f(X) \subset U \cup V$. Dann ist $f^{-1}(U) \cup f^{-1}(V) = X$ eine offene Überdeckung und $f^{-1}(U) \cap f^{-1}(V) = \emptyset$. Also ist entweder $f^{-1}(U) = \emptyset$ oder $f^{-1}(V) = \emptyset$. Das bedeutet, daß $f(X) \cap U = \emptyset$ oder $f(X) \cap V = \emptyset$. Somit ist $f(X)$ zusammenhängend. □

VI.9.3. Satz. *X wegzusammenhängend $\Rightarrow$ X zusammenhängend.*

Beweis. Sei $X = U \cup V$, $U \cap V = \emptyset$ eine disjunkte offene Überdeckung. Falls sowohl $U \neq \emptyset$ als auch $V \neq \emptyset$, so existieren $p \in U$ und $q \in V$ und nach Annahme des Wegzusammenhangs ein stetiger Weg $c\colon [0,1] \to X$ von p nach q. Nach dem vorigen Satz VI.9.2 ist $I = c([0,1])$ zusammenhängend. Also kann nicht sowohl $U \cap I \neq \emptyset$ als auch $V \cap I \neq \emptyset$ sein. Daher muss U oder V leer sein. Also ist X zusammenhängend. □

VI.9.4. Satz. *Jede zusammenhängende Teilmenge von $\mathbb{R}$ ist entweder 1-punktig oder ein Intervall, und jedes Intervall ist zusammenhängend.*

Beweis. Nach dem vorigen Satz ist jedes Intervall zusammenhängend, denn wegzusammenhängend ist es offenbar.
Sei umgekehrt $A \subset \mathbb{R}$ eine zusammenhängende Menge mit mindestens zwei Elementen. Dann existieren $s, t \in A$ mit $s < t$. Es gilt $(s, t) \subset A$, da sonst ein $x \in (s, t) \cap \mathbb{R} \setminus A$ existiert und $U = (-\infty, x) \cap A$ und $V = (x, \infty) \cap A$ würden eine disjunkte Zerlegung von A in zwei nicht leere in A offene Mengen liefern, im Widerspruch zum vorausgesetzten Zusammenhang von A. Aber nun sind die Teilmengen A von $\mathbb{R}$ mit der Eigenschaft, daß $s, t \in A$ und $s < t$ auch $(s, t) \subset A$ impliziert, genau die Intervalle in $\mathbb{R}$. □

Das folgende Korollar stellt eine Verallgemeinerung des Zwischenwertsatzes dar.

VI.9.5. Korollar. *Sei X ein zusammenhängender metrischer Raum und $f\colon X \to \mathbb{R}$ eine stetige Funktion. Dann gilt für alle $a, b \in X$: f nimmt jeden Wert zwischen $f(a)$ und $f(b)$ an.*

Beweis. Nach VI.9.2 ist $f(X) \subset \mathbb{R}$ zusammenhängend und somit nach VI.9.4 ein Intervall. □

Das folgende Beispiel zeigt, daß Wegzusammenhang eine stärkere Eigenschaft als Zusammenhang ist.

VI.9.6. Beispiel.
Die Menge $X = I \cup J \subset \mathbb{R}^2$ mit

$$I = \{0\} \times [-1,1] \qquad \text{und} \qquad J = \left\{ (x, \sin \tfrac{1}{x}) \in \mathbb{R}^2 \,|\, 0 < x \leq 1 \right\}$$

ist zusammenhängend aber nicht wegzusammenhängend.
Offenbar ist $I \cup J$ der Abschluß von J. Hieraus und dem Wegzusammenhang von J folgt, daß $I \cup J$ zusammenhängend ist.
Um zu zeigen, daß $I \cup J$ nicht wegzusammenhängend ist, nehmen wir an, daß es einen stetigen Weg $c\colon \ \to I \cup J$ gibt mit $c(0) \in I$ und $c(1) \in J$. Sei $\tau \sup\{ t \in [0,1] \,|\, c(t) \in I \}$. Offenbar ist $\{ t \in (0,1] \,|\, c(t) \in J \}$ offen in $[0,1]$. Daraus folgt $c(\tau) \in I$ und $\tau < 1$. Betrachte die stetige Projektion $p\colon I \cup J \to [0,1]$, $(x,y) \mapsto y$. Dann ist $f = p \circ c\colon [0,1] \to [0,1]$ stetig, aber $\lim_{t \downarrow \tau} f(t)$ kann nicht wohldefiniert sein.

Den Begriff des Zusammenhang wollen wir nun auf das Homöomorphie-Problem anwenden:

VI.9.7. Definition.
Eine stetige Abbildung $f\colon X \to Y$ heißt ein **Homöomorphismus**, wenn f bijektiv und f^{-1} ebenfalls stetig ist. Falls es zu X und Y einen Homöomorphismus gibt, so heißen X und Y **homöomorph**. Dies ist eine Äquivalenzrelation und die Äquivalenzklasse homöomorpher Räume heißt **Homöomorphietyp**.

VI.9.8. Beispiel.

- $\mathbb{R}^2$ und $B_1(0) = \{ (x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 < 1 \}$ sind homöomorph. Betrachte die Abbildung

$$f\colon B_1(0) \to \mathbb{R}^2, \quad (x,y) \mapsto \left(\frac{x}{\sqrt{1-x^2-y^2}}, \frac{y}{\sqrt{1-x^2-y^2}} \right).$$

  f ist offenbar stetig für $x^2+y^2 < 1$. Wie man leicht verifiziert, ist $g(x,y) = (\frac{x}{\sqrt{1+x^2+y^2}}, \frac{y}{\sqrt{1+x^2+y^2}})$ die Umkehrabbildung zu f und offenbar ebenfalls stetig.
- $S^1 = \{ (x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 = 1 \}$ und $[0,1]$ sind nicht homöomorph. Angenommen, $f\colon S^1 \to [0,1]$ ist ein Homöomorphismus. Dann wären $[0,1/2) \cup (1/2,1]$ und $S^1 \setminus \{f^{-1}(1/2)\}$ vermöge f immer noch homöomorph. Aber $S^1 \setminus \{f^{-1}(1/2)\}$ ist noch zusammenhängend, $[0,1/2) \cup (1/2,1]$ dagegen nicht, was im Widerspruch zu der stetigen und surjektiven Abbildung f steht.

KAPITEL VII

Differentiation in nomierten Vektorräumen

VII.1. Totale und partielle Differentiation

VII.1.1. Definition.
Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Vektorräume. Dann bezeichnen wir mit

$$\mathcal{L}(X, Y) := \big\{ L \in \mathrm{Hom}(X, Y) \,|\, L \text{ stetig} \big\}$$

die Menge aller linearen Abbildungen, welche bezüglich $\|\cdot\|_X$ und $\|\cdot\|_Y$ stetig sind.
Das folgende Lemma erklärt, warum die stetig linearen Abbildungen auch beschränkte Operatoren heißen.

VII.1.2. Lemma. *Eine lineare Abbildung $L \in \mathrm{Hom}(X, Y)$ ist stetig genau dann, wenn*

$$\sup_{x \neq 0} \frac{\|Lx\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|Lx\|_Y < \infty .$$

Beweis. Sei zunächst $c = \sup_{x \neq 0} \frac{\|Lx\|_Y}{\|x\|_X} < \infty$. Dann folgt

$$\|Lx\|_Y \le c\,\|x\|_X \quad \text{f.a. } x \in X,$$

also ist L c-Lipschitz-stetig.
Umgekehrt angenommen, es gilt

$$\sup_{x \neq 0} \frac{\|Lx\|_Y}{\|x\|_X} = \infty,$$

dann existiert eine Folge $(x_n) \in X^{\mathbb{N}}$ mit $\frac{\|Lx_n\|_Y}{\|x_n\|_X} \to \infty$. Wir setzen $x_n' = \frac{x_n}{\|Lx_n\|_Y} \in X$ und erhalten $\|x_n'\|_X \to 0$, aber $\|Lx_n'\|_Y = 1$ für alle $n \in \mathbb{N}$, das heißt $Lx_n' \not\to 0$ konvergiert nicht. Also ist L dann nicht stetig in 0. □

Wir haben insbesondere gesehen, daß eine lineare Abbildung genau dann stetig ist, wenn sie bereits in 0 stetig ist.
Das nachfolgende Lemma zeigt, daß wir auf die explizite Voraussetzung der Stetigkeit linearer Abbildung im endlich-dimensionalen Fall verzichten können.

VII.1.3. Lemma. *Sei $X = \mathbb{R}^n$. Dann ist jede lineare Abbildung $f \in \mathcal{L}(\mathbb{R}^n, Y)$ in einen normierten Vektorraum Y automatisch stetig.*

Beweis. Wir betrachten $x = (x_1, \ldots, x^n) \in X = \mathbb{R}^n$. Dann haben wir

$$f(x_1, \ldots, x_n) = \sum_{i=1}^n x^i f(e_i),$$

wobei $e_1, \ldots, e_n \in \mathbb{R}^n$ die Standard-Basis von $\mathbb{R}^n$ ist. Sei nun $c = \max_{i=1,\ldots,n} \|f(e_i)\|_Y$ und $\|\cdot\|_\infty$ die Maximumsnorm auf $\mathbb{R}^n$. Dann folgt

$$\|f(x)\|_Y \le c\|x\|_\infty, \quad \text{f.a. } x \in \mathbb{R}^n .$$

Da je zwei Normen auf $\mathbb{R}^n$ äquivalent sind, folgt somit die Stetigkeit von f bzgl. jeder Norm auf $\mathbb{R}^n$. □

Dieses Lemma besagt also

$$\operatorname{Hom}(\mathbb{R}^n, Y) = \mathcal{L}(\mathbb{R}^n, Y)\,.$$

VII.1.4. Definition.
Für stetige lineare Abbildungen $L \in \mathcal{L}(X, Y)$ heißt der Ausdruck

$$\|L\|_{\mathsf{Op}} := \sup_{x \neq 0} \frac{\|Lx\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|Lx\|_Y < \infty$$

die sogenannte **Operator-Norm** von L. Es ist eine einfache Übungsaufgabe zu zeigen, daß es sich dabei um eine Norm auf $\mathcal{L}(X, Y)$ handelt.
Betrachte nun zu den normierten Vektorräumen $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ eine Umgebung $U \subset X$ eines gegebenen Punktes x_o. Dann heißt eine Abbildung $f\colon U \to Y$ **differenzierbar in** x_o, wenn es eine stetig lineare Abbildung $L \in \mathcal{L}(X, Y)$ gibt so daß

$$\lim_{x \to x_o} \frac{f(x) - f(x_o) - L[x - x_o]}{\|x - x_o\|_X} = 0\,.$$

Wir verwenden von nun ab die Konvention, die Argumentvariable bei linearen Abbildungen durch eckige Klammern kenntlich zu machen, $L[h] := L(h)$, $h \in X$.
Eine alternative Bezeichnung zur Differenzierbarkeit von f in x_o ist auch f in x_o als **linearisierbar** zu bezeichnen.

Das nächste Lemma zeigt nun, daß die Linearisierung von f in x_o eindeutig ist, sofern sie existiert.

VII.1.5. Lemma. *Ist f differenzierbar in x_o wie in der Definition, so ist $L \in \mathcal{L}(X, Y)$ eindeutig bestimmt und wird mit*

$$Df(x_o) := L \quad \textit{bezeichnet}$$

und heißt das **Differential von f in** x_o.

Beweis. Sei

$$\lim_{x \to x_o} \frac{f(x) - f(x_o) - L[x - x_o]}{\|x - x_o\|_X} = 0\,.$$

Dann betrachten wir ein beliebiges $h \in X \setminus \{0\}$ und erhalten mit $x = x_o + th$, $t \in \mathbb{R}$

$$\begin{aligned}
0 &= \lim_{t \to 0} \left| \frac{f(x_o + th) - f(x_o) - L[th]}{\|th\|_X} \right| \\
&= \frac{1}{\|h\|_X} \lim_{t \to 0} \left| \frac{f(x_o + th) - f(x_o)}{|t|} - \frac{t}{|t|} L[h] \right| \\
&= \frac{1}{\|h\|_X} \lim_{t \to 0} \left| \frac{f(x_o + th) - f(x_o)}{t} - L[h] \right|\,.
\end{aligned}$$

Also ist $L \in \mathcal{L}(X, Y)$ eindeutig durch f bestimmt als

$$L[h] = \lim_{t \to 0} \frac{f(x_o + th) - f(x_o)}{t}\,.$$

□

VII.1.6. Definition.
Sei $f\colon U \to Y$ wie oben. Falls für ein $h \in X \setminus \{0\}$ der Grenzwert

$$\lim_{t \to 0} \frac{f(x_o + th) - f(x_o)}{t} \quad \text{existiert,}$$

so wird er mit $\partial_h f(x_o) \in Y$ bezeichnet und heißt die **Richtungsableitung** von f in x_o in Richtung h.

VII.1.7. Bemerkung.

Beachte, daß streng genommen $h \in X$ nur ein Richtungsvektor ist, wenn $\|h\|_X = 1$ gilt. Seien $h, k \in X \setminus \{0\}$ mit $k = \lambda h$, $\lambda \in \mathbb{R}$. So gilt

$$\begin{aligned} \partial_k f(x_o) &= \lim_{t\to 0} \frac{f(x_o + tk) - f(x_o)}{t} \\ &= \lim_{t\to 0} \frac{f(x_o + t\lambda h) - f(x_o)}{\lambda t} \cdot \lambda \\ &= \lambda \cdot \lim_{\tau\to 0} \frac{f(x_o + \tau h) - f(x_o)}{\tau} = \lambda \cdot \partial_h f(x_o)\,. \end{aligned}$$

In komplizierteren Fällen sind oft der Beweis der Differenzierbarkeit einer Abbildung f in einem Punkt x_o und die Berechnung des Differentials zwei unterschiedliche Probleme. Grundsätzlich jedoch erfordert der direkte Beweis der Differenzierbarkeit erst einmal die Bestimmung der möglichen Linearisierung, also die Bestimmung der Richtungsableitung. Wie wir in den Übungsaufgaben sehen, gibt es jedoch Fälle, in denen Richtungsableitungen in einem Punkt x_o existieren, obwohl die Abbildung f in x_o nicht differenzierbar ist!

VII.1.8. Definition.

In dem Spezialfall der Standardvektorräume $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$ betrachten wir die kanonische Basis $e_1, \dots, e_n \in \mathbb{R}^n$. Sei $U \subset \mathbb{R}^n$ eine Umgebung eines Punktes $x_o \in \mathbb{R}^n$ und $f\colon U \to \mathbb{R}^m$. Dann heißt die Richtungsableitung von f in Richtung e_i die **partielle Ableitung von f nach x^i**, die gleichwertig verwendeten Notationen sind

$$\begin{aligned} \frac{\partial f}{\partial x^i}(x_o) &:= \partial_{x^i} f(x_o) := \partial_i f(x_o) := \partial_{e_i} f(x_o) \\ &= \lim_{t\to 0} \frac{f(x_o^1, \dots, x_o^i + t, \dots, x_o^n) - f(x_o^1, \dots, x_o^n)}{t} \in \mathbb{R}^m\,. \end{aligned}$$

Wenn alle partielle Ableitungen $\partial_{x^i} f(x_o)$, $i = 1, \dots, n$, existieren, dann heißt f in x_o **partiell differenzierbar**.

VII.1.9. Beispiel.

Betrachte $f\colon \mathbb{R}^3 \setminus \{z \neq 0\} \to \mathbb{R}^2$, $f(x, y, z) = (xy, x^2 - \frac{y}{z})$. Dann berechnen sich die partiellen Ableitungen zu

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y, z) &= (y, 2x), \\ \frac{\partial f}{\partial y}(x, y, z) &= (x, -\frac{1}{z}), \\ \frac{\partial f}{\partial z}(x, y, z) &= (0, \frac{y}{z^2})\,. \end{aligned}$$

VII.1.10. Definition.

Wir betrachten nun eine Abbildung $f\colon \mathbb{R}^n \supset U \to \mathbb{R}^m$,

$$f(x^1, \dots, x^n) = \begin{pmatrix} f^1(x^1, \dots, x^n) \\ \vdots \\ f^m(x^1, \dots, x^n) \end{pmatrix}.$$

E sei f differenzierbar in x_o, dann ist f in alle Richtungen partiell differenzierbar, und wir haben

$$\begin{aligned} Df(x_o)[e_i] &= \frac{\partial f}{\partial x^i}(x_o) = \partial_{e_i} f(x_o) \\ &= \left(\frac{\partial f^1}{\partial x^i}(x_o), \dots, \frac{\partial f^m}{\partial x^i}(x_o)\right)^T \\ &= \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x_o) & \cdots & \frac{\partial f^1}{\partial x^n}(x_o) \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1}(x_o) & \cdots & \frac{\partial f^m}{\partial x^n}(x_o) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}. \end{aligned}$$

Die $m \times n$-Matrix

$$Jf(x_o) := \left(\frac{\partial f^i}{\partial x^j}(x_o)\right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

heißt die **Jakobi-Matrix** von f im Punkt x_o und stellt die lineare Abbildung $Df(x_o) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ bezüglich der kanonische Basen auf $\mathbb{R}^n$ und $\mathbb{R}^m$ dar.
Die Abbildung $f\colon \mathbb{R}^n \supset U \to Y$ heißt in x_o stetig partiell differenzierbar nach x^i, wenn die partielle Ableitung $\frac{\partial f}{\partial x^i}\colon U \to Y$ in einer Umgebung von x_o existiert und in x_o stetig ist.
f heißt in x_o **zweimal partiell differenzierbar**, wenn die zweifachen partiellen Ableitungen

$$\frac{\partial^2 f}{\partial x^j \partial x^i}(x) = \frac{\partial}{\partial x^j}\left(\frac{\partial f}{\partial x^i}\right)(x), \quad 1 \le i, j \le n,$$

in $x = x_o$ existieren. Sind diese auch stetig in x_o, so sprechen wir von stetiger partieller Differenzierbarkeit. Analog definieren wir rekursiv die k-fach (stetige) partielle Differenzierbarkeit. Wir bezeichnen

$$\frac{\partial^r f}{\partial x^{i_1} \dots \partial x^{i_r}} = \frac{\partial}{\partial x^{i_1}} \cdots \frac{\partial}{\partial x^{i_r}} f$$

als die **partielle Ableitung r-ter Ordnung** zu dem Multi-Index $(i_1, \dots, i_r) \in \{1, \dots, n\}^r$.

VII.1.11. Satz. *Sei $f\colon \mathbb{R}^n \supset U \to Y$ eine gegeben Abbildung in einen normierten Vektorraum Y. Falls die zweifachen partiellen Ableitungen $\frac{\partial^2 f}{\partial x^i \partial x^j}$ und $\frac{\partial^2 f}{\partial x^j \partial x^i}$ in $U \subset X$ existieren und stetig sind, so sind sie gleich.*

Dieser Satz besagt, daß im Falle der zweifachen stetigen partiellen Differenzierbarkeit die Reihenfolge der partiellen Differentiation keine Rolle spielt. Ohne Voraussetzung der Stetigkeit der partiellen Ableitungen ist die Vertauschbarkeit im Allgemeinen falsch (siehe Übungsaufgaben).

Beweis. Ohne Beschränkung der Allgemeinheit können wir $n = 2$ annehmen, und $Y = \mathbb{R}$. Wir schreiben $(x^1, x^2) = (x, y) \in \mathbb{R}^2$.
Seien $(x_o, y_o) \in U$ und $\epsilon > 0$ so daß

$$Q := \left\{ (x, y) \,\middle|\, |x - x_o| < \epsilon,\ |y - y_o| < \epsilon \right\} \subset U\,.$$

Seien nun $h, k \in \mathbb{R} \setminus \{0\}$ mit $|h|, |k| < \epsilon$.
Betrachte zunächst $\varphi\colon (x_o - \epsilon, x_o + \epsilon) \to \mathbb{R}$,

$$\varphi(x) := f(x, y_o + k) - f(x, y_o)\,.$$

Da f partiell nach x differenzierbar ist, ist auch φ in einer reellen Variablen differenzierbar und es gilt

$$\varphi'(x) = \frac{\partial f}{\partial x}(x, y_o + k) - \frac{\partial f}{\partial x}(x, y_o)\,.$$

Aus dem Mittelwertsatz folgt, daß es für alle $x_o + h \in (x_o - \epsilon, x_o + \epsilon)$ ein x_1 zwischen x_o und $x_o + h$ gibt , so daß

$$\varphi(x_o + h) - \varphi(x_o) \;=\; h \cdot \varphi'(x_1) \;=\; h\big(\partial_x f(x_1, y_o + k) - \partial_x f(x_1, y_o)\big)\,.$$

Des weiteren ist auch $\partial_x f$ differenzierbar nach y, so daß eine erneute Anwendung des Mittelwertsatzes ein y_1 zwischen y_o und $y_o + k$ ergibt mit

$$\varphi(x_o + h) - \varphi(x_o) \;=\; hk \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)\,. \tag{12}$$

Analog existieren für $\psi\colon (y_o - \epsilon, y_o + \epsilon) \to \mathbb{R}$ mit

$$\psi(y) = f(x_o + h, y_o) - f(x_o, y_o)$$

Zahlen x_2 zwischen x_o und $x_o + h$ und y_2 zwischen y_o und $y_o + k$, so daß

$$\psi(y_o + k) - \psi(y_o) \;=\; hk \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2)\,. \tag{13}$$

Andererseits gilt

$$\begin{aligned} \varphi(x_o + h) - \varphi(x_o) \;&=\; f(x_o + h, y_o + k) - f(x_o + h, y_o) - f(x_o, y_o + k) + f(x_o, y_o) \\ &=\; \psi(y_o + k) - \psi(y_o)\,. \end{aligned}$$

Hieraus folgt mit (12) und (13)

$$\frac{\partial^2 f}{\partial x \partial y}(x_2, y_2) \;=\; \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)\,. \tag{14}$$

Wir lassen nun $(h, k) \to (0, 0)$, dann konvergieren auch $(x_1, y_1) \to (x_o, y_o)$ und $(x_2, y_2) \to (x_o, y_o)$.
Da nun die zweifachen partiellen Ableitungen $\frac{\partial^2 f}{\partial x \partial y}$ und $\frac{\partial^2 f}{\partial y \partial x}$ als stetig vorausgesetzt waren, folgt aus (14) die Identität

$$\frac{\partial^2 f}{\partial x \partial y}(x_o, y_o) \;=\; \frac{\partial^2 f}{\partial y \partial x}(x_o, y_o)\,.$$

□

Eine einfaches Beispiel für partielle Differenzierbarkeit, bei der die partiellen Ableitungen nicht stetig sind, ist

$$f\colon \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) = \begin{cases} 1, \text{ falls } \; xy = 0, \\ 0, \text{ falls } \; xy \neq 0\,. \end{cases}$$

VII.1.12. KOROLLAR (**Satz von H.A. Schwarz**). *Sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$. Falls alle partiellen Ableitungen der Ordnung $\le k$ existieren und auf U stetig sind, $\frac{\partial^r f}{\partial x^{i_1} \ldots \partial x^{i_r}} \in C^0(U, \mathbb{R})$, f.a. $i_1, \ldots, i_r \in \{1, \ldots, n\}$, $1 \le r \le k$, so sind sie unabhängig von der Reihenfolge des partiellen Differenzierens,*

$$\frac{\partial^r f}{\partial x^{i_1} \ldots \partial x^{i_r}} \;=\; \frac{\partial^r f}{\partial x^{\sigma(i_1)} \ldots \partial x^{\sigma(i_r)}} \quad \textit{f.a. } \; \sigma \in S_r\,.$$

Der Beweis folgt durch Induktion nach k aus Satz VII.1.11
Ein wichtiges, praktikables Kriterium für die Differenzierbarkeit einfach vektorwertiger Funktionen in endlich vielen Variablen ergibt sich aus der stetigen partiellen Differenzierbarkeit.

VII.1.13. Satz (**hinreichende Bedingung für Differenzierbarkeit**). *Sei $U \subset \mathbb{R}^n$ offen, $x_o \in U$, $f\colon U \to \mathbb{R}^m$, $f = (f^1, \ldots, f^m)$. Falls alle partiellen Ableitungen aller Koordinatenfunktionen $\frac{\partial f^i}{\partial x^j}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$, in U existieren und in x_o stetig sind, dann ist f auch differenzierbar in x_o.*

Beweis. Da eine vektorwertige Funktion $f\colon U \to \mathbb{R}^m$ genau dann differenzierbar ist, wenn alle ihre Koordinatenfunktion differenzierbar sind (das folgt aus der Definition zusammen mit dem Satz VI.3.6), können wir ohne Beschränkung der Allgemeinheit annehmen, daß $m = 1$. Außerdem führen wir der einfacheren Notation halber den Beweis nur für $n = 2$, der allgemeine Fall verläuft vollkommen analog. Somit betrachten wir $U \subset \mathbb{R}^2$ offen, $(x_o, y_o) \in U$ und nehmen an, daß die partiellen Ableitungen $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ in U existieren und bei (x_o, y_o) stetig sind. Aus der Offenheit von U folgt daß wir uns auf eine Rechteckumgebung der Kantenlänge 2ϵ einschränken können,

$$Q = \left\{ (x, y) \in \mathbb{R}^2 \,|\, |x - x_o| < \epsilon,\, |y - y_o| < \epsilon \right\} \subset U\,.$$

Seien nun $h, k \in \mathbb{R}$ fest mit $|h|, |k| < \epsilon$, also

$$(x_,o, y_o + k),\ (x_o + h, y_o + k) \in Q\,.$$

Da $y \mapsto f(x_o, y)$ nach Voraussetzung zwischen y_o und $y_o + k$ differenzierbar ist, folgt aus dem Mittelwertsatz die Existenz von $s, t \in (0, 1)$ so daß

$$\begin{aligned}
f(x_o, y_o + k) - f(x_o, y_o) &= k \cdot \tfrac{\partial f}{\partial y}(x_o, y_o + tk) \\
&= k \cdot \tfrac{\partial f}{\partial y}(x_o, y_o) \\
&\quad + k \cdot \big(\tfrac{\partial f}{\partial y}(x_o, y_o + tk) - \tfrac{\partial f}{\partial y}(x_o, y_o)\big), \text{ und} \\
f(x_o + h, y_o + k) - f(x_o, y_o + k) &= h \cdot \tfrac{\partial f}{\partial x}(x_o + sh, y_o + k) \\
&= h \cdot \tfrac{\partial f}{\partial x}(x_o, y_o) \\
&\quad + h \cdot \big(\tfrac{\partial f}{\partial x}(x_o + sh, y_o + k) - \tfrac{\partial f}{\partial x}(x_o, y_o)\big)\,.
\end{aligned}$$

Somit erhalten wir

$$f(x_o + h, y_o + k) - f(x_o, y_o) = Jf(x_o, y_o) \cdot \begin{pmatrix} h \\ k \end{pmatrix} + g(h, k)\,,$$

mit

$$g(h, k) = k \cdot \big(\tfrac{\partial f}{\partial y}(x_o, y_o + tk) - \tfrac{\partial f}{\partial y}(x_o, y_o)\big) + h \cdot \big(\tfrac{\partial f}{\partial x}(x_o + sh, y_o + k) - \tfrac{\partial f}{\partial x}(x_o, y_o)\big)\,.$$

Es bleibt zu zeigen, daß $\lim_{(h,k)\to(0,0)} \frac{g(h,k)}{\|(h,k)\|} = 0$. Ohne Beschränkung der Allgemeinheit können wir annehmen, daß die Einheitsvektoren in der Norm $\|\cdot\|$ auf $\mathbb{R}^2$ Einheitslänge haben, also $|h| = \|(h, 0)\| \le \|(h, k)\|$, $|k| \le \|(h, k)\|$. Daher gilt

$$\begin{aligned}
\left|\frac{g(h, k)}{\|(h, k)\|}\right| &\le \big|\tfrac{\partial f}{\partial y}(x_o, y_o + tk) - \tfrac{\partial f}{\partial y}(x_o, y_o)\big| + \big|\tfrac{\partial f}{\partial x}(x_o + sh, y_o + k) - \tfrac{\partial f}{\partial x}(x_o, y_o)\big| \\
&\to 0 \quad \text{für } h, k \to 0\,.
\end{aligned}$$

Daraus folgt die Differenzierbarkeit von f in (x_o, y_o). □

Zusammenfassend sehen wir also für Abbildungen $\mathbb{R}^n \subset U \xrightarrow{f} \mathbb{R}^m$

$$f \text{ stetig partiell diff'bar} \quad \underset{\not\Leftarrow}{\Rightarrow} \quad f \text{ diff'bar} \quad \underset{\not\Leftarrow}{\Rightarrow} \quad f \text{ partiell diff'bar}\,.$$

VII.1.14. Definition.
Eine Abbildung $f\colon X \supset U \to Y$ zwischen normierten Vektorräumen heißt **differenzierbar**, wenn sie in jedem Punkt $x_o \in U$ differenzierbar ist.

Wenn f differenzierbar ist, so erhalten wir durch das Differential eine Abbildung

$$Df\colon U \to \mathcal{L}(X,Y), \quad x \mapsto Df(x),$$

also in den normierten Vektorraum $(\mathcal{L}(X,Y), \|\cdot\|_{\mathsf{Op}})$.

VII.1.15. Definition.
Die Abbildung $f\colon X \supset U \to Y$ heißt **stetig differenzierbar**, wenn f differenzierbar ist, und $Df\colon U \to \mathcal{L}(X,Y)$ stetig. f heißt **zweimal differenzierbar**, wenn f stetig differenzierbar ist und $Df\colon U \to \mathcal{L}(X,Y)$ wiederum differenzierbar. Somit definierten wir rekursiv, was es bedeutet, daß f k-mal stetig differenzierbar ist. Die Menge der k-mal stetig differenzierbaren Abbildungen $f\colon X \supset U \to Y$ wird mit $C^k(U,Y)$ bezeichnet.

Seien X und Y zwei normierte Vektorräume, dann gibt es eine kanonische Identifizierung von $\mathcal{L}(X,\mathcal{L}(X,Y))$ mit den bilinearen Abbildungen $\mathcal{L}^{(2)}(X,Y) := \mathcal{L}(X \times X, Y)$,

$$\Phi\colon \mathcal{L}(X,\mathcal{L}(X,Y)) \xrightarrow{\cong} \mathcal{L}^{(2)}(X,Y), \quad \Phi(L)[x,x'] = L[x][x'] .$$

Sei nun $f\colon X \supset U \to Y$ eine k-mal differenzierbare Abbildung. Dann können wir das Differential k-ter Ordnung

$$D^{(k)}(x_o) = D(\ldots D(Df)\ldots)(x_o) \in \mathcal{L}(X,\mathcal{L}(X,\ldots,\mathcal{L}(X,Y))\ldots)$$

mit einer k-fach multilinearen Abbildung $D^{(k)}(x_o) \in \mathcal{L}^{(k)}(X,Y) = \mathcal{L}(X\times\ldots\times X, Y)$ identifizieren.

In dem Spezialfall $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$ werden k-fach multilineare Abbildungen $L \in \mathcal{L}^{(k)}(\mathbb{R}^n,\mathbb{R}^m) = \mathcal{L}(\mathbb{R}^n \times \ldots \times R^n, \mathbb{R}^m)$ auch als Tensoren bezeichnet und bezüglich den kanonischen Basen durch mehrdimensionale Matrizen dargestellt,

$$L = \left(L^j_{i_1\ldots i_k}\right)_{\substack{1 \le j \le m \\ 1 \le i_1,\ldots,i_k \le n}} .$$

Diese Matrizen sind im Falle des Differentials k-ter Ordnung genau durch die partiellen Ableitungen k-ter Ordnung gegeben. $D^{(k)}f(x_o) \in \mathcal{L}^{(k)}(\mathbb{R}^n,\mathbb{R}^m)$ wird dargestellt durch die $k+1$-dimensionale $m \times n \times \ldots \times n$-Matrix

$$\left(\frac{\partial^k f^j}{\partial x^{i_1}\ldots\partial x^{i_k}}(x_o)\right)_{\substack{1 \le j \le m \\ 1 \le i_1,\ldots,i_k \le n}} .$$

Aufgrund von Satz VII.1.13 können wir also die k-mal stetig differenzierbaren Abbildungen $f\colon \mathbb{R}^n \supset U \to \mathbb{R}^m$ folgendermaßen identifizieren:

VII.1.16. Korollar. *f ist k-mal stetig differenzierbar, wenn alle partielle Ableitungen bis zur k-ten Ordnung $\frac{\partial^k f^j}{\partial x^{i_1}\ldots\partial x^{i_r}}$, $j = 1,\ldots,m$, $i_1,\ldots,i_r \in \{1,\ldots,n\}$, $1 \le r \le k$, existieren und in U stetig sind.*

VII.1.17. Definition.
$C^\infty(U,Y)$ für $U \subset \mathbb{R}^n$ ist der Vektorraum aller Abbildungen $f\colon U \to Y$ deren sämtliche partiellen Ableitungen $\frac{\partial^k f}{\partial x^{i_1}\ldots\partial x^{i_k}}$ existieren und stetig sind, $1 \le i_1,\ldots,i_k \le n$, $k \in \mathbb{N}$. Diese Abbildungen heißen **glatt** auf U.

VII.1.1. Laplace-Operator. Als Beispiel zum Rechnen mit mehrfachen partiellen Ableitungen betrachten wir den Laplace-Operator in kartesischen Koordinaten auf dem $\mathbb{R}^n$.

Sei $f \in C^2(U)$, $U \subset \mathbb{R}^n$ offen. Dann heißt

$$\triangle f := \frac{\partial^2 f}{\partial x_1^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}$$

der Laplace-Operator (in kartesischen Koordinaten). Hier ist $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$. Betrachten wir zum Beispiel

- $U = \mathbb{R}^n \setminus \{0\}$, $f(x_1, \ldots, x_n) = r$, $r = \sqrt{x_1^2 + \ldots + x_n^2}$. Dann ist f beliebig oft partiell differenzierbar auf U, also glatt. Wir berechnen

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= \frac{x_i}{r}, \\ \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} &= \frac{1}{r^2}\Big(\frac{\partial x_i}{\partial x_j} \cdot r - x_i \cdot \frac{x_j}{r}\Big) \\ \frac{\partial x_i}{\partial x_j} &= \delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \neq j\,. \end{cases}, \\ \frac{\partial^2 f}{\partial x_i^2} &= \frac{1}{r^2}\big(r - \frac{x_i^2}{r}\big) = \frac{r^2 - x_i^2}{r^3}, \\ \triangle f &= (n-1)\frac{1}{r}\,. \end{aligned}$$

- U wie vorher, $f(x_1, \ldots, x_n) = \frac{1}{r}$. Dann folgt

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= -\frac{x_i}{r^3}, \\ \frac{\partial^2 f}{\partial x_i \partial x_j} &= -\frac{1}{r^5}(\delta ij r^2 - 3x_i x_j), \\ \frac{\partial^2 f}{\partial x_i^2} &= \frac{3x_i^2 - r^2}{r^5}, \\ \triangle f &= (3-n)\frac{1}{r^3}\,. \end{aligned}$$

VII.2. Regeln für das Differential

VII.2.1. Satz. (a) *Sei $f\colon X \to Y$ eine konstante Abbildung. Dann ist f differenzierbar und $Df \equiv 0$.*

(b) *Sei $f \in \mathcal{L}(X,Y)$ eine lineare Abbildung, dann ist f differenzierbar und das Differential $Df(x) = Df(0)$ ist unabhängig von $x \in X$ und gegeben durch*

$$Df(0)[v] = f(v) \quad \text{f.a. } v \in X\,.$$

(c) *Seien $f, g\colon X \supset U \to Y$ differenzierbar in x_o. Dann ist auch $f+g$ differenzierbar in x_o und $D(f+g)(x_o)[v] = Df(x_o)[v] + Dg(x_o)[v]$. Für $\lambda \in \mathbb{R}$ ist dann auch λf differenzierbar in x_o mit $D(\lambda f)(x_o)[v] = \lambda Df(x_o)[v]$, $v \in X$.*

Der Beweis ist eine leichte Übung. Aussage (c) bedeutet, daß die Differentiation für $f\colon X \supset U \to Y$,

$$D\colon C^k(U,Y) \to C^{k-1}\big(U, \mathcal{L}(X,Y)\big)$$

ein lineare Operator ist.

VII.2.2. Satz (**Kettenregel**). *Seien $U \subset X$ und $V \subset Y$ jeweils offene Teilmengen normierter Vektorräume, Z ein normierter Vektorraum und $f\colon U \to Y$, $g\colon V \to Z$ Abbildungen mit $f(U) \subset V$. Falls f in $x_o \in U$ differenzierbar ist und g in $y_o = f(x_o)$, so ist die Komposition $g \circ f$ differenzierbar in x_o und*

$$D(g \circ f)(x_o) = Dg(y_o) \circ Df(x_o) \in \mathcal{L}(X,Z)\,.$$

In den endlich-dimensionalen Fällen $X = \mathbb{R}^p$, $Y = \mathbb{R}^n$, $Z = \mathbb{R}^m$, $(x^1, \ldots, x^p) \mapsto (f^1(x), \ldots, f^n(x))$, $(y^1, \ldots, y^n) \mapsto (g^1(y), \ldots, g^m(y))$, berechnet sich die Jakobi-Matrix $J(g \circ f)(x_o) \in M(m \times p, \mathbb{R})$ aus der Matrix-Multiplikation von $J(g)(y_o) \in$

$M(m \times n, \mathbb{R})$ *und* $Jf(x_o) \in M(n \times p, \mathbb{R})$,

$$\begin{pmatrix} \frac{\partial(g^1\circ f)}{\partial x^1} & \dots & \frac{\partial(g^1\circ f)}{\partial x^p} \\ \vdots & \dots & \vdots \\ \frac{\partial(g^m\circ f)}{\partial x^1} & \dots & \frac{\partial(g^m\circ f)}{\partial x^p} \end{pmatrix} = \begin{pmatrix} \frac{\partial g^1}{\partial y^1} & \dots & \frac{\partial g^1}{\partial y^n} \\ \vdots & \dots & \vdots \\ \frac{\partial g^m}{\partial y^1} & \dots & \frac{\partial g^m}{\partial y^n} \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \dots & \frac{\partial f^1}{\partial x^p} \\ \vdots & \dots & \vdots \\ \frac{\partial f^n}{\partial x^1} & \dots & \frac{\partial f^n}{\partial x^p} \end{pmatrix}.$$

VII.2.3. Bemerkung.

- Der Spezialfall $m = 1$ drückt sich folgendermaßen aus:

$$\begin{aligned} \frac{\partial}{\partial x^i}(g \circ f) &= \frac{\partial}{\partial x^i}\Big(g\big(f^1(x^1 \dots, x^p), \dots, f^n(x^1, \dots, x^p)\big)\Big) \\ &= \sum_{j=1}^{n} \frac{\partial g}{\partial y^j}(f^1, \dots, f^n) \cdot \frac{\partial f^j}{\partial x^i}(x^1, \dots, x^p). \end{aligned}$$

- In dem Spezialfall $p = 1$ und $m = 1$ bezeichnen wir $f\colon \mathbb{R} \supset U \to \mathbb{R}^n$ als einen Weg und wir erhalten die Kettenregel in der Form

$$\frac{d}{dx}(g \circ f) = \frac{d}{dx}\big(g(f^1(x), \dots, f^n(x))\big) = \sum_{j=1}^{n} \frac{\partial g}{\partial y^j} \cdot \frac{\partial f^j}{\partial x}.$$

- Es empfiehlt sich in allen Fällen die selbe Notation für die Koordinatenfunktionen von f wir für die Variablen von g zu verwenden also

$$f(x^1, \dots, x^p) = (y^1(x^1, \dots, x^n), \dots, y^n(x^1, \dots, x^p)).$$

Dadurch erscheinen die Formeldarstellungen der Kettenregel suggestiv in Form einer "Variablen-Kürzung",

$$\frac{\partial(g^k \circ f)}{\partial x^i} = \sum_{j=1}^{n} \frac{\partial g^k}{\partial y^j} \frac{\partial y^j}{\partial x^i}.$$

VII.2.4. Definition.

Sei $\gamma\colon I = (a, b) \to \mathbb{R}^n$ ein differenzierbarer Weg, dann identifizieren wir das Differential $D\gamma(t)$ im Punkt $t \in (a, b)$ mit dem Ableitungsvektor

$$\dot\gamma(t) := \frac{d}{dt}\gamma(t) = \big(\gamma_1'(t), \dots, \gamma_n'(t)\big) \in \mathbb{R}^n$$

aus den Ableitungen der n Koordinatenfunktionen. Der Vektor $\dot\gamma(t) \in \mathbb{R}^n$ heißt **Tangentialvektor** an γ in $\gamma(t)$. Das Differential $D\gamma(t) \in \mathcal{L}(\mathbb{R}, \mathbb{R}^n)$ wird ausgedrückt durch

$$D\gamma(t)[h] = h \cdot \dot\gamma(t), \quad h \in \mathbb{R}.$$

VII.2.5. Beispiel.

Sei $f\colon (a, b) \to \mathbb{R}$ gegeben durch

$$f(x) = \int_a^{x^2} h(x, t)dt,$$

mit einer stetig differenzierbaren Funktion $h\colon (a, b) \times (a, b^2) \to \mathbb{R}$. Dann gilt

$$f'(x) = 2x \cdot h(x, x^2) + \int_a^x \frac{\partial h}{\partial x^1}(x, t)dt.$$

Dies sieht man folgendermaßen: Sei g die Funktion in zwei Variablen, $g(y_1, y_2) = \int_a^{y_2} h(y_1, t)dt$. Dann berechnen wir die partiellen Ableitungen zu

$$\frac{\partial g}{\partial y_1} = \int_a^{y_2} \frac{\partial h}{\partial y_1}(y_1, t)dt, \quad \frac{\partial g}{\partial y_2} = h(y_1, y_2).$$

Mit $\gamma(x) = (x, x^2)$ erhalten wir $f(x) = (g \circ \gamma(x))$ und nach der Kettenregel

$$\begin{aligned} f'(x) &= \frac{\partial g}{\partial y_1} \cdot \gamma_1' + \frac{\partial g}{\partial y_2} \cdot \gamma_2' \\ &= \int_a^{x^2} \frac{\partial h}{\partial x_1}(x,t)dt + h(x, x^2) \cdot 2x \,. \end{aligned}$$

Beweis der Kettenregel VII.2.2. Da f bei x_o differenzierbar ist, gilt

(15) $$f(x) = f(x_o) + Df(x_o)[x - x_o] + R_f(x_o, x) \quad \text{mit} \quad \lim_{x \to x_o} \frac{R_f(x_o, x)}{\|x - x_o\|} = 0,$$

und analog für g differenzierbar bei y_o,

(16) $$g(y) = g(y_o) + Dg(y_o)[y - y_o] + R_g(y_o, y) \quad \text{mit} \quad \lim_{y \to y_o} \frac{R_g(y_o, y)}{\|y - y_o\|} = 0\,.$$

Somit ergibt sich für $x \in U$,

$$\begin{aligned} g\big(f(x)\big) &= g(y_o) + Dg(y_o)\big[f(x) - f(x_o)\big] + R_g\big(y_o, f(x)\big) \\ &= g(f(x_o)) + Dg(y_o)\big[Df(x_o)[x - x_o] + R_f(x_o, x)\big] + R_g(y_o, f(x)) \\ &= (g \circ f)(x_o) + \big(Dg(y_o) \circ Df(x_o)\big)[x - x_o] + R_1(x_o, x) + R_2(x_o, x)\,, \end{aligned}$$

wobei $R_1(x_o, x) = Dg(y_o)\big[R_f(x_o, x)\big]$ und $R_2(x_o, x) = R_g(y_o, f(x))$. Wir wissen, daß wegen (15)

$$\frac{R_1(x_o, x)}{\|x - x_o\|} = Dg(y_o)\left[\frac{R_f(x_o, x)}{\|x - x_o\|}\right] \to 0 \quad \text{für} \quad x \to x_o\,.$$

Es bleibt also die Grenzwertuntersuchung von

$$\frac{R_2(x_o, x)}{\|x - x_o\|} = \begin{cases} \frac{R_g(y_o, f(x))}{\|f(x) - f(x_o)\|} \cdot \frac{\|f(x) - f(x_o)\|}{\|x - x_o\|}\,, & \text{falls } f(x) \neq f(x_o), \\ 0\,, & \text{sonst.} \end{cases}$$

Da f bei x_o stetig ist, folgt $\lim_{x \to x_o} \frac{R_g(y_o, f(x))}{\|f(x) - f(x_o)\|} = 0$ aus (16). Also genügt es zu zeigen, daß $\lim_{x \to x_o} \frac{\|f(x) - f(x_o)\|}{\|x - x_o\|}$ existiert. Dies folgt wegen (15) aus

$$\frac{\|f(x) - f(x_o)\|}{\|x - x_o\|} = \frac{\|Df(x_o)[x - x_o] + R_f(x_o, x)\|}{\|x - x_o\|} \leq \|Df(x_o)\|_{\mathsf{Op}} + \frac{R_f(x_o, x)}{\|x - x_o\|}\,.$$

□

VII.3. Gradienten, Graphen und Tangentialräume

VII.3.1. Definition.
Sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ reellwertig und differenzierbar in U. Dann ist der (kartesische) **Gradient von f in x_o** definiert durch

$$\nabla f(x_o) := \operatorname{grad} f(x_o) := \left(\frac{\partial f}{\partial x^1}(x_o), \ldots, \frac{\partial f}{\partial x^n}(x_o)\right) \in \mathbb{R}^n\,.$$

Sei $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf $\mathbb{R}^n$, dann gilt

VII.3.2. Lemma. *Ist f differenzierbar in x_o, so erfüllt das Differential $Df(x_o) \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R})$ die Identität*

$$Df(x_o)[v] = \partial_v f(x_o) = \langle \nabla f(x_o), v \rangle \quad \textit{f.a.} \;\; v \in \mathbb{R}^n \setminus \{0\}\,.$$

Beweis. Folgt direkt aus der Definition. □

VII.3.3. Satz. *Sei f differenzierbar in x_o, dann gilt:*

(a) $\nabla f(x_o) = 0$ *genau dann wenn* $\frac{\partial f}{\partial x^i}(x_o) = 0$ *f.a.* $i = 1, \ldots, n$*, genau dann wenn* $Df(x_o) = 0$.
(b) $\big|\partial_v f(x_o)\big| \leq \|\nabla f(x_o)\|_2$ *für alle Einheitsvektoren* $v \in S^{n-1}$*, d.h.* $\|v\|_2 = 1$.
(c) *Falls* $\nabla f(x_o) \neq 0$ *und* $v = \frac{\nabla f(x_o)}{\|\nabla f(x_o)\|_2}$*, so ist die Richtungsableitung* $\partial_v f(x_o) = \|\nabla f(x_o)\|_2$.

Die Aussagen (b) und (c) bedeuten, daß der Gradient in die Richtung des stärksten Anstiegs von f zeigt.

Beweis. (a) ist trivial und (b) und (c) folgen aus der Cauchy-Schwarz-Ungleichung

$$\big|\partial_v f(x_o)\big| = \big|\langle \nabla f(x_o), v\rangle\big| \leq \|\nabla f(x_o)\|_2 \cdot \|v\|_2 \,.$$

□

Sei $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ differenzierbar. Sei $\gamma\colon \mathbb{R} \supset I \to U$ eine differenzierbare Kurve, welche die Differentialgleichung

$$\dot\gamma(t) = \nabla f\big(\gamma(t)\big) \quad \text{f.a. } t \in I$$

erfüllt. Eine solche Kurve heißt **Gradientenflußkurve**. Es gilt für den infinitesimalen Anstieg von f entlang einer solchen Kurve

$$\frac{d}{dt}\Big(f\big(\gamma(t)\big)\Big) = df\big(\gamma(t)\big) \cdot \dot\gamma(t) = \big\langle \nabla f(\gamma(t)), \nabla f(\gamma(t))\big\rangle = \big\|\nabla f(\gamma(t))\big\|_2^2 \,.$$

Also ist f entlang der Gradientenflußkurve monoton steigend und genau dann konstant, wenn $\gamma(t)$ konstant in einem **kritischen Punkt** von f liegt, d.h. $\nabla f(\gamma(t)) = 0$.

VII.3.4. Definition.
Sei wieder $U \subset \mathbb{R}^n$ offen und $f\colon U \to \mathbb{R}$ reellwertig und differenzierbar. Dann heißt

$$N_c(f) = \big\{\, x \in U \mid f(x) = c \,\big\}, \quad c \in \mathbb{R}$$

die **Niveaumenge** von f zum Niveau c.

VII.3.5. Beispiel.
- $f\colon \mathbb{R}^3 \to \mathbb{R}$, $f(x,y,z) = x^2 + y^2 - z^2$, dann ist

$$\begin{aligned} N_0(f) &= \big\{\,(x,y,z) \mid z^2 = x^2 + y^2 \,\big\} = \text{Kegel}, \\ N_1(f) &= \big\{\,(x,y,z) \mid x^2 + y^2 = z^2 + 1 \,\big\} = \text{einschaliger Hyperboloid}, \\ N_{-1}(f) &= \big\{\,(x,y,z) \mid z^2 = x^2 + y^2 + 1 \,\big\} = \text{zweischaliger Hyperboloid}\,. \end{aligned}$$

- $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = xy$. $N_c(f)$ sind für alle $c \neq 0$ Hyperbeln, für $c = 0$ das Achsenkreuz.

Betrachte nun wieder eine reellwertige, differenzierbare Funktion $f\colon \mathbb{R}^n \supset U \to \mathbb{R}$, einen Punkt in der Niveaumenge $x_o \in N_c(f)$ so daß $\nabla f(x_o) \neq 0$. Sei des weiteren $\gamma\colon (-\epsilon, \epsilon) \to U$ ein differenzierbarer Weg mit $\gamma(0) = x_o$ und $f\big(\gamma(t)\big) = c$ f.a. $|t| < \epsilon$, also γ verläuft innerhalb der Niveaumenge $N_c(f)$.

VII.3.6. Definition.
Falls $\nabla f(x_o) \neq 0$, so heißt

$$T_{x_o}N_c(f) := \big\{\, \dot\gamma(0) \in \mathbb{R}^n \mid \gamma\colon (-\epsilon,\epsilon) \to U \text{ diff'bar }, \ f(\gamma(t)) = c \text{ f.a. } |t| < \epsilon \,\big\}$$

der **Tangentialraum an die Niveaumenge** $N_c(f)$ **in** x_o**.** Ein Element $v \in T_{x_o}N_c(f)$ heißt **Tangentialvektor**.

VII.3.7. Satz. *Seien* $f\colon U \to \mathbb{R}$ *und* $x_o \in U$ *wie oben mit* $\nabla f(x_o) \neq 0$*. Dann ist* $T_{x_o}N_c(f)$ *ein* $(n-1)$*-dimensionaler Untervektorraum von* $\mathbb{R}^n$ *und der Gradient* $\nabla f(x_o)$ *steht senkrecht zu* $T_{x_o}N_c(f)$*,*

$$\big\langle \nabla f(x_o), v\big\rangle = 0 \quad \textit{f.a. } v \in T_{x_o}N_c(f)\,.$$

Umgekehrt gilt:

$$T_{x_o}N_c(f) \;=\; \big(\mathbb{R}\cdot\nabla f(x_o)\big)^{\perp} \;=\; \big\{\, v\in\mathbb{R}^n \,|\, \langle v,\nabla f(x_o)\rangle = 0 \,\big\}\,.$$

Beweis. "$\underline{T_{x_o}N_c(f)\subseteq\big(\mathbb{R}\cdot\nabla f(x_o)\big)^{\perp}}$": Sei $\gamma\colon(-\epsilon,\epsilon)\to U$ ein differenzierbarer Weg innerhalb $N_c(f)$ mit $\gamma(0)=x_o$¿ Dann gilt

$$0=\frac{d}{dt}\big(f(\gamma(t))\big) \;=\; \langle\nabla f(x_o)(\gamma(t)),\dot\gamma(t)\rangle,$$

also $\dot\gamma(0)\perp\nabla f(x_o)$.
"$\underline{\big(\mathbb{R}\cdot\nabla f(x_o)\big)^{\perp}\subseteq T_{x_o}N_c(f)}$": Diese Richtung ist etwas subtiler zu beweisen und benötigt das Argument des Theorems über Implizite Funktionen, welches wir erst etwas später kennenlernen werden, siehe VII.5.2. Der Beweis wird in VII.5 nachgeliefert. □

VII.3.8. Bemerkung.
Es gilt immer, daß $\big(\mathbb{R}\cdot\nabla f(x_o)\big)^{\perp} = \ker Df(x_o)\subset\mathbb{R}^n$ ein $(n-1)$-dimensionaler Untervektorraum ist, falls $\nabla f(x_o)\neq 0$. Der vorige Satz identifiziert diesen Raum mit dem Tangentialraum.

VII.3.9. Beispiel.
Sei $f\colon M(n\times n,\mathbb{R})\to\mathbb{R}$ gegeben durch $f(A)=\det A$ und $N_1(f)=\mathsf{SL}(n,\mathbb{R})$, die Gruppe der speziellen linearen Transformationen. Die Linearisierung der Determinante in der Identität ist durch den Spur-Operator gegeben, $Df(\mathbf{1})[B]=\operatorname{spur}B$, also folgt aus $Df(\mathbf{1})[\mathbf{1}]=n\neq 0$, daß $\nabla f(\mathbf{1})\neq 0$ mit $\mathbf{1}\in N_1(f)$. Somit erhalten wir die Darstellung des Tangentialraumes an der Identität $\mathbf{1}\in\mathsf{SL}(n,\mathbb{R})$ durch

$$\mathfrak{sl}(n,\mathbb{R}) := T_{\mathbf{1}}\mathsf{SL}(n,\mathbb{R}) \;=\; kerDf(\mathbf{1}) \;=\; \big\{\,\text{spurfreie } (n\times n)-\text{Matrizen}\,\big\}\,.$$

Neben den Niveaumenge ist eine weitere wichtige Klasse von nichtlinearen Teilmengen von $\mathbb{R}^n$ die der Graphen von Funktionen. Sei wieder $U\subset\mathbb{R}^n$ und $f\colon U\to\mathbb{R}$ differenzierbar. Dann definieren wir den **Graph von** f durch

$$\operatorname{graph} f \;=\; \big\{\,(x^1,\dots,x^n,x^{n+1})\in\mathbb{R}^{n+1}\,|\,(x^1,\dots,x^n)\in U,\; x^{n+1}=f(x^1,\dots,x^n)\,\big\}\,.$$

Diese Menge $\operatorname{graph} f$ ist die Niveaumenge zu $F\colon U\times\mathbb{R}\to\mathbb{R}$, $F(x^1,\dots,x^{n+1}) = x^{n+1}-f(x^1,\dots,x^n)$, also $\operatorname{graph} = N_0(F)$. Es gilt

$$\nabla F(x^1,\dots,x^{n+1}) \;=\; \big(-\nabla f(x^1,\dots,x^n),1\big) \;\neq\; 0 \quad \text{f.a. } \; \xi=(x^1,\dots,x^{n+1})\in N_0(F)\,.$$

Somit erhalten wir den Tangentialraum an den Graphen im Punkt ξ in der Darstellung

$$\begin{aligned} T_\xi \operatorname{graph} f \;&=\; \big\{\,(v^1,\dots,v^{n+1})\in\mathbb{R}^{n+1}\,|\,v^{n+1}-\langle\nabla f(x^1,\dots,x^n),(v^1,\dots,v^n)\rangle = 0\,\big\}\\ &= \operatorname{span}\big\{\,(1,0,\dots,0,\frac{\partial f}{\partial x^1}),\,(0,1,0,\dots,0,\frac{\partial f}{\partial x^2},\dots,(0,\dots,0,1,\frac{\partial f}{\partial x^n})\,\big\}\,.\end{aligned}$$

Im Spezialfall $\nabla f(x^1\dots,x^n)=0$ ergibt sich dann

$$T_{(x^1,\dots,x^n,f(\dots))}\operatorname{graph} f \;=\; \mathbb{R}^n\times\{0\}\,.$$

VII.4. Umkehrsatz

VII.4.1. Satz (**Banachscher Fixpunktsatz**). *Sei (X,d) ein vollständiger, metrischer Raum und $F\colon X\to X$ eine Kontraktionsabbildung, d.h.*

$$d\big(F(x),F(y)\big) \;\le\; L\cdot d(x,y) \quad \textit{f.a. } \; x,y\in X, \textit{mit } \; L<1\,.$$

Dann gilt: Es existiert ein eindeutiger Fixpunkt $\xi\in X$, d.h. $F(\xi)=\xi$.

Beweis. Eindeutigkeit: Angenommen, es existieren $\xi, \xi' \in X$ mit $F(\xi) = \xi$ und $F(\xi') = \xi'$. Dann gilt

$$d(\xi, \xi') = d\big(F(\xi), F(\xi')\big) \le L \cdot d(\xi, \xi')\,.$$

Daraus folgt wegen $0 \le L < 1$

$$(1-L) \cdot d(\xi, \xi') \le 0 \;\Rightarrow\; \xi = \xi'\,.$$

Existenz: Die Idee besteht darin, durch eine rekursive Anwendung von F eine Folge in X zu konstruieren, die gegen einen Fixpunkt konvergiert. Sei $x_o \in X$ beliebig. Dann definieren wir rekursiv

$$(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}} \quad \text{durch} \;\; x_n = F(x_{n-1}),\; n \in \mathbb{N}\,.$$

Schritt 1: (x_n) ist eine Cauchy-Folge.
Bew.:

$$\begin{aligned} d(x_{n+k}, x_n) &\le d(x_{n+k}, x_{n+k-1}) + \ldots + d(x_{n+1}, x_n) \\ &\le \big(L^{k-1} + L^{k-2} + \ldots + 1\big) d(x_{n+1}, x_n) \\ &= \frac{1-L^k}{1-L} d(x_{n+1}, x_n) \le \frac{L^n}{1-L} d(x_1, x_o) \to 0\,. \end{aligned}$$

Schritt 2: Da (X,d) vollständig ist, konvergiert die Folge (x_n). Sei $\xi = \lim x_n$. Dann folgt aus der Stetigkeit von F

$$F(\xi) = \lim F(x_n) = \lim x_{n+1} = \xi\,.$$

Also ist ξ ein Fixpunkt von F. □

VII.4.2. Definition.
Es seien X und Y normierte Vektorräume und $U \subset X$ und $V \subset Y$ offene Teilmengen. Dann heißt eine Abbildung $f\colon U \to V$ ein **Diffeomorphismus**, wenn f bijektiv ist, f differenzierbar in U und f^{-1} differenzierbar in V ist.

Offenbar ist jeder Diffeomorphismus ein Homöomorphismus, aber die Umkehrung ist falsch. Betrachte zum Beispiel den Homöomorphismus

$$f\colon \mathbb{R} \to \mathbb{R}, \quad f(x) = x^3\,.$$

Die Umkehrabbildung $f^{-1}(x) = \operatorname{sgn}(x)\sqrt[3]{|x|}$ ist in 0 nicht differenzierbar. Daß dieses Beispiel charakteristisch für das einzige Hindernis zur Umkehrung ist, sieht man im folgenden

VII.4.3. Satz. *Seien $U \subset X$, $V \subset Y$ wie oben und $f\colon U \to V$ ein Homöomorphismus und differenzierbar in $x_o \in U$. Dann ist f^{-1} genau dann in $y_o = f(x_o) \in V$ differenzierbar, wenn das Differential $Df(x_o) \in \mathcal{L}(X,Y)$ invertierbar ist und die Inverse ebenfalls stetig. In diesem Fall gilt dann*

$$D(f^{-1})(y_o) = \big(Df(x_o)\big)^{-1} \in \mathcal{L}(Y,X)\,.$$

Beweis. Daß die stetige Invertierbarkeit eine notwendige Bedingung ist, folgt aus der Kettenregel,

$$\mathrm{id} = D(\mathrm{id})(y_o) = Df(x_o) \circ D(f^{-1})(y_o)\,.$$

Wir nehmen nun umgekehrt an, dass $Df(x_o) \in \mathcal{L}(X,Y)$ stetig invertierbar ist. Durch die Transformation

$$\tilde{f}(x) = Df(x_o)^{-1}\big(f(x_o + x) - f(x_o)\big)$$

erreichen wir, daß wir $x_o = y_o = 0$ und $Df(x_o) = \mathrm{id}$ annehmen können.

Also bedeutet die Differenzierbarkeit von f in 0

$$f(x) \;=\; x + R(x) \quad \text{mit} \quad \lim_{x\to 0} \frac{R(x)}{\|x\|} = 0\,. \tag{17}$$

Das bedeutet für die als stetig angenommene Umkehrabbildung $y = f^{-1}(y) + R\big(f^{-1}(y)\big)$, also

$$f^{-1}(y) \;=\; y + \tilde{R}(y) \quad \text{mit} \;\; \tilde{R}(y) = -R\big(f^{-1}(y)\big)\,.$$

Zu zeigen ist demnach $\lim_{y\to 0} \frac{\tilde{R}(y)}{\|y\|} = 0$. Wegen (17) und aufgrund der Homöomorphie-Eigenschaft von f finden wir $r, \delta > 0$ mit

$$\|R(x)\| \;\le\; \frac{1}{2}\|x\| \quad \text{f.a. } \|x\| \le r \quad \text{und} \qquad \|f^{-1}(y)\| \le r \quad \text{f.a. } \|y\| \le \delta\,,.$$

Das liefert $\|\tilde{R}(y)\| \le \frac{1}{2}\|f^{-1}(y)\|$ für alle $\|y\| \le \delta$, und wir erhalten

$$\|f^{-1}(y)\| \;=\; \|y + \tilde{R}(y)\| \;\le\; \|y\| + \frac{1}{2}\|f^{-1}(y)\| \quad \text{f.a. } \|y\| \le \delta\,.$$

Durch Umstellung ergibt sich daraus

$$\|f^{-1}(y)\| \;\le\; 2\|y\| \quad \text{f.a. } \|y\| \le \delta$$

und hiermit

$$\frac{\|\tilde{R}(y)\|}{\|y\|} \;\le\; 2\frac{\|R(f^{-1}(y))\|}{\|f^{-1}(y)\|} \;=\; 2\frac{\|R(x)\|}{\|x\|} \;\to 0 \quad \text{für} \;\; y \to 0,$$

denn das ist äquivalent zu $x \to 0$. □

VII.4.4. Korollar. *Sei $f\colon X \supset U \to V \subset Y$ ein stetig differenzierbarer Homöomorphismus wie oben, und sei $Df(x)$ für alle $x \in U$ ein Isomorphismus. Dann ist auch f^{-1} stetig differenzierbar.*

Beweis. Nach Satz VII.4.3 ist f^{-1} in allen $y \in V$ differenzierbar. Bleibt zu zeigen, daß $D(f^{-1})(y)$ auch stetig von y abhängt. Dies folgt aus der Formel

$$D(f^{-1})(y) \;=\; \Big(Df\big(f^{-1}(y)\big)\Big)^{-1}$$

Zusammen mit dem Sachverhalt, daß die Inversions-Operation $A \mapsto A^{-1}$ auf der offenen Teilmenge der Banach-Algebra $\mathcal{L}(X, X)$ der stetig linearen und invertierbaren Operatoren gemäß VI.6.4 (b) eine stetige Abbildung ist, folgt die Stetigkeit von $y \mapsto D(f^{-1})(y)$. □

Sei nun $U \subset X$ und $f \in C^1(U, Y)$ wie bisher und $K \subset U$ eine kompakte und konvexe Teilmenge, d.h. für alle $x, y \in K$ gilt auch $x + t(y - x) \in K$ für alle $t \in [0, 1]$. Dann gilt

VII.4.5. Lemma. *Für alle $x, y \in K$ ist*

$$\|f(x) - f(y)\| \;\le\; \|Df\|_K \,\cdot \|x - y\|\,,$$

mit $\|Df\|_K = \max_{x\in K} \|Df(x)\|_{\mathsf{Op}}$.

Beachte, daß das Supremum von $\|Df(x)\|$ für $x \in K$ wegen der Kompaktheit und der Stetigkeit von $Df(x)$ als Maximum angenommen wird.

Beweis. Sei $\gamma(t) = y + t(x - y)$, $t \in [0, 1]$, die Strecke von y nach x innerhalb K. Dann betrachte wir zu $\epsilon > 0$ die Funktion

$$F_\epsilon\colon [0, 1] \to \mathbb{R}, \quad F_\epsilon(t) \;=\; \big\|f(\gamma(t)) - f(y)\big\| - t\big(\|Df\|_K + \epsilon\big)\|x - y\|\,.$$

Wir behaupten, daß $F_\epsilon(1) \le 0$ für alle $\epsilon > 0$ gilt. Angenommen wir haben $F_\epsilon > 0$, dann folgt wegen $F_\epsilon(0) = 0$ aus der Stetigkeit von F_ϵ

$$\forall\, c \in \big(0, F_\epsilon(1)\big) \quad \exists\, t_o \in (0,1] \text{ mit } F_\epsilon(t_o) = c \text{ und } F_\epsilon(t) > c \quad \text{f.a. } t_o < t \le 1\,.$$

Also gilt

$$\varphi(t) := \frac{F_\epsilon(t) - F_\epsilon(t_o)}{t - t_o} > 0 \quad \text{f.a. } t_o < t \le 1\,.$$

Andererseits haben wir

$$\begin{aligned} F_\epsilon(t) - F_\epsilon(t_o) &= \big\|f(\gamma(t)) - f(y)\big\| - \big\|f(\gamma(t_o)) - f(y)\big\| - (t-t_o)\big(\|Df\|_K + \epsilon\big)\|x-y\| \\ &\le \big\|f(\gamma(t)) - f(\gamma(t_o))\big\| - (t-t_o)\big(\|Df\|_K + \epsilon\big)\|x-y\|\,. \end{aligned}$$

Daraus folgt

$$\varphi(t) \le \left\|\frac{f(\gamma(t)) - f(\gamma(t_o))}{t - t_o}\right\| - \big(\|Df\|_K + \epsilon\big)\|x-y\|\,. \tag{18}$$

Aus

$$\lim_{t \downarrow t_o} \left\|\frac{f(\gamma(t)) - f(\gamma(t_o))}{t - t_o}\right\| = \big\|Df(\gamma(t_o))[x-y]\big\| \le \|Df\|_K \|x-y\|$$

und (18) folgt dann

$$\varphi(t) < 0 \quad \text{für } t - t_o > 0 \quad \text{klein genug}\,,$$

im Widerspruch zur Annahme. Also gilt

$$\|f(x) - f(y)\| \le \big(\|Df\|_K + \epsilon\big)\|x-y\| \quad \text{f.a. } \epsilon > 0\,.$$

Daraus folgt die Behauptung. □

Nun haben wir alle Vorbereitungen abgeschlossen, um den zentralen Umkehrsatz beweisen zu können.

VII.4.6. Theorem (**Umkehrsatz**). *Seien X und Y normierte Vektorräume und X vollständig, also ein Banachraum. Sei $U \subset X$ offen und $\Phi \in C^1(U,Y)$ eine stetig differenzierbare Abbildung. Wenn das Differential $D\Phi(x_o) \in \mathcal{L}(X,Y)$ in einem Punkt $x_o \in U$ stetig invertierbar ist, so existieren Umgebungen $U_o(x_o) \subset U$ und $V(\Phi(x_o)) \subset Y$, so daß die Restriktion*

$$\Phi_{|U_o} \colon U_o(x_o) \xrightarrow{\sim} V(\Phi(x_o))$$

ein Diffeomorphismus ist, d.h. Φ ist in x_o ein **lokaler Diffeomorphismus**.

Beweis. Ohne Einschränkung können wir annehmen, daß $X = Y$, $x_o = 0$, $\Phi(x_o) = 0$ und $D\Phi(x_o) = \mathrm{id}$, denn wir können Φ durch

$$\tilde\Phi(x) = D\Phi(x_o)^{-1}\big(\Phi(x_o + x) - \Phi(x_o)\big)$$

ersetzen und die Aussage des Theorems für den Spezialfall $\tilde\Phi$ impliziert dann den allgemeinen Fall.

<u>Schritt 1</u>: Zunächst zeigen wir die lokale Invertierbarkeit der Abbildung Φ, d.h. wir zeigen: Für alle y nahe bei 0 existiert ein eindeutiges x nahe bei 0, so daß $\Phi(x) = y$. Hierbei verwenden wir die *Methode des Fixpunktprinzips.* Sei

$$\varphi_y \colon U \to X, \quad \varphi_y(x) = y + x - \Phi(x) \quad \text{für } y \in X\,.$$

Also gilt

$$\Phi(x) = y \quad \Leftrightarrow \quad \varphi_y(x) = x\,.$$

Wir zeigen, daß für y in einer hinreichend klein gewählten Umgebung von 0 eine abgeschlossene Teilmenge $B \subset X$ existiert, so daß $\varphi_y(B) \subset B$ und

$$\big\|\varphi_y(x) - \varphi_y(x')\big\| \le \frac{1}{2}\|x - x'\|,$$

d.h. $\varphi_{y|B}$ ist eine Kontraktionsabbildung, und B ist ein vollständiger metrischer Raum, so daß der Banachsche Fixpunktsatz VII.4.1 anwendbar ist.
Da $x \mapsto D\Phi(x)$ stetig ist und $D\Phi(0) = \mathrm{id}$, existiert ein $r > 0$ so daß $\overline{B_{2r}(0)} \subset U$ und

$$\big\| \mathrm{id} - D\Phi(x) \big\|_{\mathsf{Op}} \leq \frac{1}{2} \quad \text{f.a. } x \in \overline{B_{2r}(0)}\,.$$

Da $D\varphi_y(x) = \mathrm{id} - D\Phi(x)$ flgt $\|D\varphi_y(x)\|_{\mathsf{Op}} \leq \frac{1}{2}$ für alle $\|x\| \leq 2r$. Somit folgt aus Lemma VII.4.5

$$\big\|\varphi_y(x) - \varphi_y(x')\big\| \leq \frac{1}{2}\|x - x'\| \quad \text{f.a. } x, x' \in B_{2r}(0)\,. \tag{19}$$

Außerdem gilt

$$\|\varphi_y(x)\| \leq \big\|\varphi_y(x) - \varphi_y(0)\big\| + \|y\| \leq \frac{1}{2}\|x\| + \|y\| < 2r \quad \text{f.a. } \|x\| \leq 2r,\ \|y\| < r\,.$$

Also gilt: $\varphi_y(\overline{B_{2r}(0)}) \subset \overline{B_{2r}(0)}$ für alle $\|y\| < r$ und $\varphi_{y|\overline{B_{2r}(0)}}$ ist eine Kontraktionsabbildung mit Kontraktionsfaktor $\frac{1}{2}$.
Aus dem Banachschen Fixpunktsatz folgt nun die Existenz einer Abbildung

$$\Psi\colon B_r(0) \to \overline{B_{2r}(0)}, \quad \text{mit } \Phi \circ \Psi(y) = y\,.$$

<u>*Schritt 2*</u>: Wir behaupten, daß Ψ stetig ist.
Bew.: Seien $x_i = \Psi(y_i)$, $i = 1, 2$. Also haben wir

$$\Psi(y_1) - \Psi(y_2) = x_1 - x_2 = \varphi_0(x_1) - \varphi_0(x_2) + \Phi(x_1) - \Phi(x_2)\,.$$

Hieraus folgt mittels (19)

$$\|x_2 - x_1\| \leq \frac{1}{2}\|x_2 - x_1\| + \|\Phi(x_2) - \Phi(x_1)\|$$

und somit

$$\|\Psi(y_2) - \Psi(y_1)\| = \|x_2 - x_1\| \leq 2\|y_2 - y_1\|\,.$$

<u>*Schritt 3*</u>: Sei nun $U_o = \Phi^{-1}\big(B_r(0)\big) \cap B_{2r}(0)$. Da aufgrund von VI.6.4 (b) die Menge der invertierbaren stetig linearen Operatoren offen in $\mathcal{L}(X, X)$ ist und $x \mapsto D\Phi(x)$ stetig, gibt es ein $r > 0$ hinreichend klein, so daß $D\Phi(x)$ für alle $x \in U_o$ ein Isomorphismus ist. Damit können wir Satz VII.4.3 auf $\Phi_{|U_o}\colon U_o \to \Phi(U_o)$ anwenden und erhalten den behaupteten Diffeomorphismus.

□

VII.4.7. Beispiel.
Betrachte die Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (e^x \cos y, e^x \sin y)$. Daraus ergibt sich für das Differential

$$Jf(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix},$$

also

$$\det Jf(x, y) = e^x \neq 0 \quad \text{f.a. } (x, y) \in \mathbb{R}^2\,.$$

Nach dem Umkehrsatz ist also f in allen Punkten ein lokaler Diffeomorphismus. Beachte jedoch, daß f selbst nicht injektiv ist, also kein globaler Diffeomorphismus!

Eine typische Anwendung des Umkehrsatzes ist

VII.4.8. Satz (**Offenheitssatz**). *Seien X, Y vollständige normierte Vektorräume, $U \subset X$ offen und $\Phi \in C^1(U, Y)$ mit $D\Phi(x)$ stetig invertierbar in allen $x \in U$. Dann ist auch das Bild $\Phi(U)$ offen in Y.*

Beweis. Gemäß dem Umkehrsatz ist Φ in allen $x_o \in U$ ein lokaler Diffeomorphismus. Daraus folgt insbesondere, daß jeder Punkt $\Phi(x_o)$ eine Umgebung in Y besitzt, die noch ganz zu $\Phi(U)$ gehört. Also ist das Bild offen. □

VII.4.9. SATZ (**Diffeomorphiesatz**). *Seien* $\Phi\colon X \supset U \to Y$ *wie in Satz VII.4.8 und zudem* Φ *auch injektiv. Dann ist*

$$\Phi\colon U \to V := \Phi(U) \subset Y$$

ein Diffeomorphismus.

BEWEIS. Nach Konstruktion und Voraussetzung ist $\Phi\colon U \to V$ bijektiv und für jede offene Menge $U' \subset U$ ist $\Phi(U') \subset V$ gemäß Satz VII.4.8 ebenfalls offen. Das bedeutet aber gerade, daß $\Psi := \Phi^{-1}\colon V \to U$ stetig ist. Also ist Φ ein Homöomorphismus. Darauf ist aber nun wiederum Satz VII.4.3 anwendbar und die Behauptung folgt. □

Eine der wichtigsten Anwendungen in der mehrdimensionalen Analysis ist der Satz über

VII.5. Implizite Funktionen

VII.5.1. BEISPIEL. (1) Betrachte die Gleichung in 3 Variablen

$$x^2 + y^2 + z^2 = 1\,.$$

Die geometrische Sicht hierauf bedeutet, die Lösungsmenge als eine lokal parametrisierbare zwei-dimensionale Fläche im $\mathbb{R}^3$ zu beschreiben. Eine algebraische Sicht hierzu besteht darin, die Gleichung nach 1 Variablen aufzulösen, z.B.

$$z = z(x, y) = \pm\sqrt{1 - x^2 - y^2}\,.$$

Das $\pm$-Zeichen verdeutlicht die fehlende Eindeutigkeit einer solchen Auflösung.

(2) Betrachte nun die Menge von orthogonalen $n \times n$-Matrizen

$$\mathsf{O}(n, \mathbb{R}) = \left\{ A \in M(n \times n, R) \,|\, A \cdot A^T = \mathbf{1} \right\},$$

als Lösungsmenge $\{ A \,|\, F(A) = \mathbf{1} \}$ bzgl. der Abbildung

$$\begin{aligned} F\colon M(n \times n, \mathbb{R}) \to S(n, \mathbb{R}) &= \{ A \in M(n \times n, \mathbb{R}) \,|\, A = A^T \}, \\ A &\mapsto A \cdot A^T\,. \end{aligned}$$

Hier ist nun das Ziel, $\mathsf{O}(n, \mathbb{R})$ als lokal parametrisierbares Objekt durch $\dim \mathsf{O}(n)$-viele Variablen zu beschreiben, mit $\dim \mathsf{O}(n) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$.

(3) Wir betrachten nun ein System aus q Gleichungen in $n = p + q$ reellen Variablen. Wir erwarten hier also p Freiheitsgrade.

$$\begin{cases} F_1(x^1, \ldots, x^p, y^1, \ldots, y^q) &= 0, \\ \quad\vdots & \\ F_q(x^1, \ldots, x^p, y^1, \ldots, y^q) &= 0\,. \end{cases} \tag{20}$$

Das Problem besteht nun darin, (20) so nach $(y^1, \ldots, y^q)$ aufzulösen daß $y = y(x)$ Funktionen in $x^1, \ldots, x^p$ sind,

$$y^i = f_i(x^1, \ldots, x^p),\ i = 1, \ldots, q,$$

d.h. wir suchen Funktionen $f_1, \ldots, f_q$ in $x^1, \ldots, x^p$, so daß (20) erfüllt ist,

$$F_j(x, f_1(x), \ldots, f_q(x)) = 0 \quad \text{f.a. } j = 1, \ldots, q,\ x = (x^1, \ldots, x^p)\,.$$

Die gesuchten Funktionen $f_1, \ldots, f_q$ heißen dann die **impliziten Funktionen**.

Diese Theorie lässt sich wieder in einem recht allgemeinen Kontext von Banachräumen behandeln, d.h. also auch für unendlich-dimensionale Räume.
Es seien X, Y, Z drei Banachřaume,

$$U \subset X \times Y \text{ offen, und } \quad F\colon U \to Z$$

mindestens einmal stetig differenzierbar.
Wir definieren die **partiellen Differentiale** in $(x, y) \in U$,

$$\begin{aligned}
&D_X F(x,y) \in \mathcal{L}(X,Z) \quad \text{definiert durch} \\
&D_X F(x,y)[h] = DF(x,y)\big[(h,0)\big] \quad \text{f.a. } h \in X, \\
&D_Y F(x,y) \in \mathcal{L}(Y,Z) \quad \text{definiert durch} \\
&D_Y F(x,y)[k] = DF(x,y)\big[(0,k)\big] \quad \text{f.a. } k \in Y\,.
\end{aligned}$$

Wir haben nun folgendes

VII.5.2. Theorem (**Satz über Implizite Funktionen**). *Sei*

$$(x_o, y_o) \in U \quad \textit{mit } \; F(x_o, y_o) = 0$$

und $D_Y F(x_o, y_o) \in \mathcal{L}(Y, Z)$ sei ein stetig linearer Isomorphismus (also auch stetige Inverse). Dann existieren offene Umgebungen $U'(x_o) \subset X$, $U''(y_o) \subset Y$ und $g\colon U' \to U''$ stetig differenzierbar mit:

(a) $U'(x_o) \times U''(y_o) \subset U$,
(b) $g(x_o) = y_o$.
(c) $F(x, g(x)) = 0$ *für alle* $x \in U'(x_o)$.

Also ist g eine lokal in x_o definierte implizite Funktion bezüglich der Bedingung $g(x_o) = y_o$.

Wir rufen uns hierbei in Erinnerung, daß im Falle endlich-dimensionaler Vektorräume X, Y, Z die konkreten Normen auf diesen keine Rolle spielen und alle linearen Abbildungen automatisch auch stetig sind.
Der Beweis von VII.5.2 wird nun auf den Umkehrsatz VII.4.6 zurückgeführt. In der Tat kann man sogar die Äquivalenz dieses Impliziten-Funktionen-Theorems mit dem Umkehrsatz beweisen. (Übungsaufgabe!)

Beweis. Wir betrachten die Abbildung

$$\begin{aligned}
\Phi\colon U \;&\to\; X \times Z, \\
(x,y) \;&\mapsto\; \big(x, F(x,y)\big)\,.
\end{aligned}$$

Das Differential berechnet sich zu

$$D\Phi(x,y)\big[(h,k)\big] \;=\; \big(h, D_X F(x,y)[h] + D_Y F(x,y)[k]\big)\,.$$

Zu gegebenem $(a, b) \in X \times Z$ lässt sich die Gleichung

$$\big(h, D_X F[h] + D_Y F[k]\big) \;=\; (a,b)$$

eindeutig auflösen zu $h = a$ und $k = (D_Y F)^{-1}\big(b - D_X F[a]\big)$. Das bedeutet, daß $D\Phi(x_o, y_o)$ ein stetig linearer Isomorphismus ist. Also ist der Umkehrsatz auf Φ anwendbar und es existiert eine Umgebung $U_o(x_o, y_o) \subset U$ sowie eine Umgebung $V(x_o, 0) \subset X \times Z$ so daß die Einschränkung von Φ ein Diffeomorphismus ist,

$$\Phi_{|U_o}\colon U_o \xrightarrow{\approx} V \quad \text{Diffeom.}$$

Die dazu Inverse schreiben wir als

$$\begin{aligned}
&\Psi = (\Phi_{|U_o})^{-1}\colon V \to U_o, \\
&\Psi(\xi,\eta) = \big(\xi, h(\xi,\eta)\big)\,,
\end{aligned}$$

wodurch die Abbildung $h\colon X \times Z \supset V \to Y$ definiert wird. Für $(x,y) \in U_o$ gilt nun

$$\begin{aligned} F(x,y) = 0 &\Leftrightarrow \Phi(x,y) = (x,0) \\ &\Leftrightarrow (x,y) = \big(x, h(x,0)\big) \\ &\Leftrightarrow y = h(x,0)\,. \end{aligned} \tag{21}$$

Seien nun $U'(x_o)$ und $U''(y_o)$ Umgebungen in X bzw. Y so daß

$$U' \times U'' \subset U_o \quad \text{und} \quad h(x,0) \in U'' \text{ f.a. } x \in U'\,.$$

Dann definieren wir $g\colon U' \to U''$ durch

$$g(x) := h(x,0)$$

und erhalten damit die gesuchte implizite Funktion. □

VII.5.3. Bemerkung. (a) Aus (21) folgt, daß für U' klein genug die implizite Funktion $g\colon U' \to U''$ durch die Bedingung $g(x_o) = y_o$ **eindeutig bestimmt** ist.

(b) Falls die vorgegebene Abbildung F k-mal stetig differenzierbar ist, so ist auch die implizite Funktion $g \in C^k$, **k-mal stetig differenzierbar**.

VII.5.4. Beispiel. (1) Sei $F(x,y,z) = x^2 + y^2 - z^2 - 1$. Dann ist die Lösungsmenge

$$\{\, F(x,y,z) = 0 \,\} \;=\; \{\, x^2 + y^2 = 1 + z^2 \,\}$$

ein einschaliger Hyperboloid. Wir betrachten nun die spezielle Lösung $(1,0,0)$ und berechnen in diesem Punkt die partielle Ableitung $\partial_x F = 2x \neq 0$. Also ist nach dem Satz über implizite Funktionen die Gleichung lokal bei $(1,0,0)$ nach x auflösbar, $x = x(y,z)$. Dies lässt sich in diesem Fall natürlich auch explizit durchführen,

$$x = \sqrt{1 + y^2 - y^2} \quad \text{f.a. } (y,z) \in B_r(0), \text{ z.B. } 0 < r < 1\,.$$

(2) Sei nun

$$F(x,y) = e^{\sin(xy)} + x^2 - 2y - 1 \overset{!}{=} 0\,.$$

Wir wollen nun lokal in $(0,0)$, welches offenbar eine Lösung ist, nach y auflösen. De Berechnung

$$\partial_y F = x\cos(xy)e^{\sin(xy)} - 2, \quad \partial_y F(0,0) = -2 \neq 0$$

ergibt, daß dies möglich ist, also die implizite Funktion $y = y(x)$ existiert nahe bei $x = 0$. Dies ist nun ein Beispiel, in dem das Auflösen nicht explizit mit bekannten Formeln möglich ist. Wie bestimmt man aber nun die implizite Funktion konkret? Zumindest die Taylorreihen-Entwicklung lässt sich wie folgt bestimmen: Das Ableiten der Gleichung $F\big(x, y(x)\big) = 0$ führt zu

$$\partial_x F + \partial_y F \cdot y' = 0, \tag{22}$$

was sich nach y' auflösen lässt, wenn $\partial_y F \neq 0$ gilt, was genau der hinreichenden Bedingung für das Implizite-Funktionen-Theorem entspricht. Wir erhalten in dem Beispiel $y(0) = 0$, $y'(0) = \frac{1}{2} \cdot 0 = 0$. Wenn man (22) weiter ableitet erhält man

$$\partial_{xx} F + 2\partial_{xy} F \cdot y' + \partial_{yy} F \cdot (y')^2 + \partial_y F \cdot y'' = 0,$$

welches sich in $x = 0$ wieder nach y'' auflösen lässt, usw.

(3) Wir betrachten nun das obige Beispiel der orthogonalen Gruppe $\mathsf{O}(n,\mathbb{R})$. Sei

$$F\colon M(n \times n, \mathbb{R}) \to S(n,\mathbb{R}), \quad A \mapsto A \cdot A^T\,.$$

Sei außerdem

$$A(n,\mathbb{R}) \;=\; \{\, A \in M(n \times n, \mathbb{R}) \mid A = -A^T \,\}$$

der $\frac{n(n-1)}{2}$-dimensionale Vektorraum der antisymmetrischen Matrizen. Wir haben die kanonische Zerlegung in symmetrische und antisymmetrische Matrizen

$$\begin{aligned} M(n \times n, \mathbb{R}) &\cong A(n) \oplus S(n), \\ A &\mapsto \left(\frac{A - A^T}{2}, \frac{A + A^T}{2} \right). \end{aligned}$$

Wir setzen $X = A(n)$ und $Y = Z = S(n)$ und erhalten bzgl. der Zerlegung

$$F\colon X \times Y \to Z, \quad F(A, S) = (A + S) \cdot (A + S)^T .$$

Das ergibt

$$F(A, S) = -A^2 + A \cdot S - S \cdot A + S^2 .$$

Man beachte, daß die Matrizenmultiplikation $A \cdot S$ nicht kommutativ ist!
Es gilt nun $F(0, \mathbf{1}) = \mathbf{1}$, und die Linearisierung ist

$$\begin{aligned} DF(A, S)[h, k] = &- h \cdot A - A \cdot h \\ &+ h \cdot S - S \cdot h \\ &+ A \cdot k - k \cdot A \\ &+ k \cdot S + S \cdot k . \end{aligned}$$

Also ist das partielle Differential

$$D_Y F(0, \mathbf{1})[k] = 2k, \quad D_Y F(0, \mathbf{1}) \in \mathcal{L}(S(n), S(n))$$

ein Isomorphismus. Somit ist das Implizite-Funktionen-Theorem anwendbar, und es existiert eine Abbildung, definiert auf einer Umgebung von $0 \in A(n)$,

$$g\colon A(n) \supset U(0) \to S(n), \quad g(0) = \mathbf{1} ,$$

so daß $g\colon U(0) \to \mathsf{O}(n, \mathbb{R})$ ein Homöomorphismus auf eine Umgebung von $\mathbf{1} \in \mathsf{O}(n, \mathbb{R})$ ist. Die Gruppe der orthogonalen $n \times n$-Matrizen wird also zumindest in einer Umgebung der Einheitsmatrix durch eine offene Menge von antisymmetrischen Matrizen parametrisiert. In diesem oberflächlichen, vorläufigen Sinne können wir also von der Dimension $\frac{n(n-1)}{2}$ sprechen. Eine Präzisierung wird erst später mit Hilfe der Theorie der sogenannten **Mannigfaltigkeiten** erfolgen.

Mit Hilfe des Satzes über implizite Funktionen können wir nun auch den Beweis der Identifizierung des Tangentialraumes einer Niveaumenge abschließen.

Beweis. *der zweiten Hälfte von Satz VII.3.7.* Sei $f\colon \mathbb{R}^n \supset U \to \mathbb{R}$ eine mindestens einmal stetig differenzierbare Funktion, und wir betrachten die Niveaumenge $N_c(f)$ zu dem Wert $c \in \mathbb{R}$. Sei $x_o \in N_c(f)$ mit $\nabla f(x_o) \neq 0$.
Das bedeutet, daß für mindestens einen Koordinatenindex $i \in \{1, \dots, n\}$ die partielle Ableitung $\frac{\partial f}{\partial x^i}(x_o) \neq 0$. Ohne Beschränkung der Allgemeinheit (d.h. nach geeigneter Vertauschung der Koordinatenindices) sei $i = n$. Damit erfüllt die Funktion

$$f\colon \mathbb{R}^{n-1} \times \mathbb{R} \supset U \to \mathbb{R}$$

die Voraussetzung für die Anwendung des Impliziten-Funktionen-Theorems und wir erhalten eine offene Umgebung $U'\big((x_o^1, \dots, x_o^{n-1})\big) \subset \mathbb{R}^{n-1}$ und eine mindestens einmal stetig differenzierbare Funktion $g\colon U' \to \mathbb{R}$ mit

$$g(x_o^1, \dots, x_o^{n-1}) = x_o^n \quad \text{und} \quad f\big(y^1, \dots, y^{n-1}, g(y^1, \dots, g^{n-1})\big) = c \text{ f.a. } y \in U' .$$

Für U' klein genug existiert somit eine Umgebung $U''(x_o^n) \subset \mathbb{R}$ so daß

$$N_c(f) \cap \big(U' \times U''\big) = \operatorname{graph} g .$$

Aus der Darstellung des Tangentialraums von $T_{x_o}\operatorname{graph} g$ in VII.3 folgt nun, daß dieser identisch ist mit dem Untervektorraum $\big(\mathbb{R}\cdot\nabla f(x_o)\big)^{\perp}\subset\mathbb{R}^n$. □

VII.6. Taylorformel und lokale Extremwerte

In diesem Abschnitt verallgemeinern wir die Theorie der Taylorentwicklung auf Funktionen in mehreren Variablen. Wir schränken uns allerdings auf den endlichdimensionalen Fall ein. Auch hierzu gilt wieder eine entsprechend allgemeinere Version auch in dem Banachraum-Kontext.
Sei $f\colon \mathbb{R}^n\supset U\to\mathbb{R}^m$ eine mindestens k-mal stetig differenzierbare Abbildung. Dann ist das Differential k-ter Ordnung,

$$D^{(k)}f(x)\in\mathcal{L}^{(k)}(\mathbb{R}^n\times\ldots\times\mathbb{R}^n,\mathbb{R}^n)$$

gegeben durch

$$D^{(k)}f(x)[v_1,\ldots,v_k] \;=\; \sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n}\frac{\partial^k f}{\partial x^{i_1}\ldots\partial x^{i_k}}(x)\cdot v_1^{i_1}\cdot\ldots\cdot v_k^{i_k},$$

für $v_j=(v_j^1,\ldots,v_j^n)\in\mathbb{R}^n$, $j=1,\ldots,k$. Nach Satz VII.1.12 ist $D^{(k)}f(x)$ nicht nur k-multilinear, sondern auch symmetrisch, d.h.

$$D^{(k)}f(x)[v_1,\ldots,v_k] \;=\; D^{(k)}f(x)[v_{\sigma(1)},\ldots,v_{\sigma(k)}]\quad\text{für alle Permutationen }\;\sigma\in S_k\,.$$

Wir nehmen nun der Einfachheit halber den Spezialfall $m=1$ an, und daß $f\in C^{k+1}(U,\mathbb{R})$. Sei $x_o\in U$ fest und $r>0$ mit $B_r(x_o)\subset U$. Sei $h\in\mathbb{R}^n$ mit $\|h\|<r$, also $x_o+h\in B_r(x_o)$. Das Ziel ist nun eine Taylorentwicklung für $f(x_o+h)=f(x_o)+\ldots$ zu finden, indem wir dies auf die bekannte Theorie in einer reellen Variablen zurückführen.
Wir betrachten $g\colon[0,1]\to\mathbb{R}$, $g(t)=f(x_o+th)=f(x_o^1+th^1,\ldots,x_o^n+th^n)$. Also ist g nach der Kettenregel ebenfalls $(k+1)$-mal stetig differenzierbar und die bekannte Taylorformel ergibt

$$g(1)=g(0)+g'(0)+\frac{1}{2}g''(0)+\ldots+\frac{g^{(k+1)}(\tau)}{(k+1)!}\quad\text{für ein }\;\tau\in(0,1)\,.$$

Also haben wir

$$f(x_o+h)=f(x_o)+\sum_{i=1}^{k}\frac{g^{(i)}(0)}{i!}+\frac{1}{(k+1)!}g^{(k+1)}(\tau)\,.$$

Wir berechnen für $g(t)=f(x_o+th)$,

$$\begin{aligned}
g'(t) \;&=\; Df(x_o+th)[h]=\big\langle\nabla f(x_o+th),h\big\rangle=\sum_{i=1}^{n}\frac{\partial f}{\partial x^i}(x_o+th)\cdot h^i,\\
g''(t) \;&=\; D^{(2)}f(x_o+th)[h,h]=\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x^i\partial x^j}(x_o+th)h^ih^j,\\
&\;\vdots\\
g^{(k)}(t) \;&=\; \sum_{i_1,\ldots,i_k=1}^{n}\frac{\partial^k f}{\partial x^{i_1}\ldots\partial x^{i_k}}(x_o+th)h^{i_1}\cdot\ldots\cdot h^{i_k},
\end{aligned}$$

wobei in den mehrfachen partiellen Ableitungen die Reihenfolge keine Rolle spielt, also

$$\begin{aligned} f(x_o+h) \;&=\; f(x_o) + \sum_{r=1}^{k} \frac{1}{r!} \sum_{i_1,\ldots,i_r=1}^{n} \frac{\partial^r f}{\partial x^{i_1}\ldots\partial x^{i_r}}(x_o) h^{i_1}\cdot\ldots\cdot h^{i_r} \\ &\quad + \frac{1}{(k+1)!} \sum_{i_1,\ldots,i_{k+1}=1}^{n} \frac{\partial^{k+1} f}{\partial x^{i_1}\ldots\partial x^{i_{k+1}}}(x_o+\tau h)) h^{i_1}\cdot\ldots\cdot h^{i_{k+1}} \end{aligned} \tag{23}$$

für ein $\tau \in (0,1)$. In koordinatenfreier Notation haben wir also

$$\begin{aligned} f(x_o+h) \;&=\; f(x_o) + Df(x_o)[h] + \frac{1}{2}D^{(2)}f(x_o)[h,h] + \ldots + \\ &\quad + \frac{1}{k!}D^{(k)}f(x_o)[h,\ldots,h] + \frac{1}{(k+1)!}D^{(k+1)}f(x_o+\tau h)[h,\ldots,h]\,. \end{aligned}$$

In dem Spezialfall $k=2$ können wir dies auch folgendermaßen schreiben,

$$\begin{aligned} f(x_o+h) \;&=\; f(x_o) + \langle \nabla f(x_o), h\rangle + \frac{1}{2}h^T\cdot Hf(x_o)\cdot h + \mathrm{o}(\|h\|^2), \\ \text{mit }\; Hf(x_o) &= \begin{pmatrix} \frac{\partial^2 f}{(\partial x^1)^2} & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x^n \partial x^1} & \cdots & \frac{\partial^2 f}{(\partial x^n)^2} \end{pmatrix}, \quad \lim_{h\to 0}\frac{\mathrm{o}(\|h\|^2)}{\|h\|^2} = 0\,. \end{aligned} \tag{24}$$

VII.6.1. Definition.
Die symmetrische (da f mind. zweimal stetig diff'bar) Matrix $Hf(x_o)$ heißt die **Hessesche Matrix von f bei x_o**.
Wir nennen $x_o \in U$ ein **lokales Maximum bzw. Minimum** von f, wenn es ein $r>0$ gibt mit $B_r(x_o)\subset U$ und

$$f(x) \;\leq\; \text{bzw.}\; \geq\; f(x_o) \quad \text{f.a. } \; x\in B_r(x_o)\,.$$

In beiden Fällen sprechen wir auch von einer **lokalen Extremstelle ode Extremum**. Das lokale Extremum heißt **isoliert** oder **strikt**, wenn $<$ bzw. $>$ für alle $x \in B_r(x_o)\setminus\{x_o\}$ gilt. Der Punkt $x_o\in U$ heißt **kritischer Punkt** von f, wenn $\nabla f(x_o)$ existiert und $\nabla f(x_o)=0$.

VII.6.2. Satz. *Sei f in $x_o \in U$ partiell differenzierbar, und x_o sei eine lokale Extremstelle von f. Dann ist x_o auch ein kritischer Punkt von f, also $\nabla f(x_o)=0$.*

Beweis. Der Beweis lässt sich leicht auf die bekannte Differentialrechnung in einer Variablen zurückführen. Wir betrachten nämlich $g(t)=f(x_o+te_i)$ jeweils für $i=1,\ldots,n$, $t\in(-r,r)$. Dann ist g in einer reellen Variablen bei $t=0$ differenzierbar und hat dort eine lokale Extremstelle. Daraus folgt $g'(0)=0$ und dies ist äquivalent zu $\frac{\partial f}{\partial x^i}(x_o)=0$. Da dies für alle $i=1,\ldots,n$ gilt, folgt $\nabla f(x_o)=0$. □

Wir wiederholen hier noch einmal folgende Begriffe für quadratische Formen aus der Linearen Algebra:

VII.6.3. Definition.
Sei $A\in M(n\times n,\mathbb{R})$ eine symmetrische Matrix, $A=A^T$ und sei $Q\colon \mathbb{R}^n\to\mathbb{R}$ die durch A definierte quadratische Form

$$Q(v) \;=\; v^T\cdot A\cdot v \;=\; \sum_{i,j=1}^{n} a_{ij}v^iv^j\,.$$

Dann heißt Q bzw. A

positiv definit, wenn $Q(v)>0$ für alle $v\neq 0$,
negativ definit, wenn $Q(v)<0$ für alle $v\neq 0$,

indefinit, wenn es $v, w \in \mathbb{R}^n$ gibt mit $Q(v) > 0$ und $Q(v) < 0$,
positiv semidefinit, wenn $Q(v) \geq 0$ für alle $v \in \mathbb{R}^n$,
negativ semidefinit, wenn $Q(v) \leq 0$ für alle $v \in \mathbb{R}^n$.

VII.6.4. LEMMA. *Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch und positiv definit. Dann hat die zugeordnete quadratische Form $Q(v) = v^T A v$ in 0 ein striktes Minimum und es existiert ein $m > 0$ mit*

$$Q(v) \geq m\|v\|^2 \quad \textit{f.a.} \quad v \in \mathbb{R}^n .$$

Die Abschätzung gilt für jede Norm $\|\cdot\|$ auf $\mathbb{R}^n$, allerdings hängt die Konstante m von dieser Norm ab. Der Beweis dieses Lemma ist eine einfache Übungsaufgabe aus der linearen Algebra und benutzt die Diagonalisierbarkeit von A aufgrund der Symmetrie.

Der nachfolgende Satz besagt nun, daß aus dem notwendigen Kriterium für ein lokales Extremum auch ein hinreichendes gemacht werden kann, wenn zudem die Hessesche als Darstellung der zweiten Ableitung von f entsprechend definit ist.

VII.6.5. SATZ. *Sei $U \subset \mathbb{R}^n$ offen, $f \in C^2(U, \mathbb{R})$ eine reellwertige Funktion mit einem kritischen Punkt $x_o \in U$. Dann gilt:*

(a) *Ist $Hf(x_o)$ positiv definit, so ist x_o ein striktes lokales Minimum.*
(b) *Ist $Hf(x_o)$ negativ definit, so ist x_o ein striktes lokales Maximum.*
(c) *Ist $Hf(x_o)$ indefinit, so ist x_o keine lokale Extremstelle.*

BEWEIS. (a): Nach Lemma VII.6.4 existiert ein $m > 0$ so daß $h^T Hf(x_o)h \geq m\|h\|^2$ für alle $h \in \mathbb{R}^n$. Die Taylor-Entwicklung (24) liefert

$$f(x_o + h) = f(x_o) + \frac{1}{2}h^T Hf(x_o)h + R(h) \quad \text{mit} \quad \lim_{h \to 0} \frac{R(h)}{\|h\|^2} = 0 .$$

Demnach wählen wir $\delta > 0$, so daß $|R(h)| \leq \frac{1}{4}m\|h\|^2$ für alle $\|h\| < \delta$ gilt. Wir erhalten daraus

$$\begin{aligned} f(x_o + h) &\geq f(x_o) + \frac{1}{2}m\|h\|^2 - \big|R(h)\big| \\ &\geq f(x_o) + \frac{1}{4}m\|h\|^2 \quad \text{f.a.} \quad \|h\| < \delta . \end{aligned}$$

Also hat f in x_o ein striktes lokales Minimum.
(b): Dies folgt unmittelbar aus (a) durch Ersetzen von f durch $-f$.
(c): Sei $v \in \mathbb{R}^n \setminus \{0\}$ mit $v^T Hf(x_o)v \geq m_1\|v\|^2$. Dann ist 0 ein striktes lokales Minimum von $t \mapsto f(x_o + tv)$. Entsprechend finden wir aufgrund der Indefinitheit auch eine Richtung $w \in \mathbb{R}^n \setminus \{0\}$, auf die eingeschränkt x_o ein striktes lokales Maximum ist. Also kann insgesamt x_o keine Extremstelle sein. □

VII.6.1. Extremwerte unter Nebenbedingungen. Ein wichtiges Problem bei praktischen Anwendungen ist die Suche nach Extremwerten von Funktionen in mehreren Veränderlichen, wobei diese Variablen noch durch i.A. nichtlineare Nebenbedingungen untereinander gekoppelt sind.

Wir betrachten zunächst folgende Beispiele für einen solchen Problemtyp:

VII.6.6. BEISPIEL. (1) Hat die Funktion $f(x, y) = xy$ auf der Geraden $x + y = 1$ ein Maximum? Hier ist die Lösung durch Elimination einer Variablen mittels der Nebenbedingung sehr einfach: $y = 1 - x$ aus der Nebenbedingung wird in f eingesetzt und ergibt die Funktion

$$\varphi(x) = f(x, 1 - x) = x - x^2$$

in einer Variablen. Die notwendige Bedingung für eine Extremstelle ist $\varphi'(x) = 0$, also $1 - 2x = 0$. Man sieht sofort, daß wegen $\varphi''(\frac{1}{2}) = -2$ die

Funktion f in $(\frac{1}{2}, \frac{1}{2})$ ein Maximum unter der Nebenbedingung $x+y=1$ hat.

(2) Wir betrachten nun das Problem, die Oberfläche eines Quaders der Kantenlängen $x, y, z > 0$ zu Minimieren, und zwar unter der Nebenbedingung eines konstanten Volumens $V = xyz = \text{const}$. Die Oberfläche ist durch $A = 2(xy + yx + xz)$ gegeben. Wieder erlaubt uns die Nebenbedingung, explizit eine der drei Variablen durch Auflösen zu eliminieren, z.B. $z = \frac{V}{xy}$. Demnach erhalten wir die Funktion in den verbliebenen freien Variablen

$$\varphi(x,y) = 2\big(xy + (x+y)\frac{V}{xy}\big) = 2\big(xy + V(\frac{1}{x} + \frac{1}{y})\big)\,.$$

Dies ist nun nach Satz VII.6.2 und Satz VII.6.5 leicht behandelbar: Die notwendige Bedingung

$$\nabla\varphi = 2\big(y - \frac{V}{x^2}, x - \frac{V}{y^2}\big) \stackrel{!}{=} 0 \quad \text{ergibt} \;\; x = y = \sqrt[3]{V}\,.$$

Das Kriterium der positiv definiten Hesseschen im kritischen Punkt $(x,y) = (\sqrt[3]{V}, \sqrt[3]{V})$

$$H^2 f = \begin{pmatrix} 4\frac{V}{x^3} & 2 \\ 2 & 4\frac{V}{y^3} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$$

ergibt ein striktes (lokales) Minimum, welches auch ein globales, also absolutes Minimum ist. Also ist der Quader mit minimaler Oberfläche bei vorgegebenem Volumen ein Würfel der Kantenlänge $\sqrt[3]{V}$.

(3) Welchen Abstand hat der Punkt $(2,0)$ in der Ebene von der Kurve C, die durch die kubische Gleichung $x^3 + y^3 - 3xy = 0$ gegeben ist? (Man mache sich selbst eine grafische Veranschaulichung hiervon!) Analytisch bedeutet dieses Problem das Minimieren des Abstandes $f(x,y) = \sqrt{(x-2)^2 + y^2}$ unter der Nebenbedingung $x^3 + y^3 - 3xy = 0$. Hier haben wir nun das praktische Problem, daß die Nebenbedingung kein einfaches Auflösen nach einer der beiden Variablen erlaubt.

Dies ist ein typisches Problem und verlangt nach einer anderen Methode.

Die Antwort nach einer geeigneten Methode zu Auffinden von Extrema unter nichtlinearen Nebenbedingungen liegt wieder im Prinzip der Impliziten Funktion. Wir betrachten nun das Problem allgemein:
Gegeben sei eine offene Teilmenge $U \subset \mathbb{R}^n$ und Funktionen $f, g_1, \ldots, g_k \in C^1(U, \mathbb{R})$. Gesucht seien lokale Extreme von f unter den k Nebenbedingungen $g_1 = 0, \ldots\, g_k = 0$, d.h. wir suchen Punkte $a \in U$ mit

(a) $g_1(a) = g_2(a) = \ldots = g_k(a) = 0$, also $a \in N_0(g_1) \cap \ldots \cap N_0(g_k)$, und
(b) es existiert eine Umgebung $V(a)$ in U mit $f(x) \leq f(a)$ (bzw. $\geq$) für alle $x \in V(a) \cap N_0(g_1) \cap \ldots \cap N_0(g_k)$.

VII.6.7. Satz (**Lagrange-Multiplikatoren-Regel**). *Sei $a \in U$ ein Extremum von f unter den Nebenbedingungen $g_1(a) = \ldots = g_k(a) = 0$, und seien $\nabla g_1(a), \ldots, \nabla g_k(a) \in \mathbb{R}^n \setminus \{0\}$ linear unabhängig. Dann existieren $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ mit*

$$\nabla f(a) = \sum_{i=1}^{k} \lambda_i \nabla g_i(a)\,.$$

Dieser Satz liefert ein *notwendiges* Kriterium für die Existenz von Extrema unter Nebenbedingungen, vorausgesetzt, daß die Nebenbedingungen einer Nichtdegeneriertheitsbedingung genügen. Die reellen Zahlen $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ nennt man hierbei die **Lagrange-Multiplikatoren**.

Dieses notwendige Kriterium lässt sich wie folgt in einem Algorithmus zusammenfassen: Wir betrachten die kombinierte Funktion

$$F\colon U \times \mathbb{R}^k \to \mathbb{R},$$

$$F(x^1, \dots, x^n, \lambda_1, \dots, \lambda_k) \;=\; f(x^1, \dots, x^n) - \sum_{i=1}^{k} \lambda_i \cdot g_i(x^1, \dots, x^n),$$

in $n+k$ vielen Variablen und berechnen hierzu den Gradienten in $\mathbb{R}^{n+k}$,

$$\nabla F(x^1, \dots, x^n, \lambda_1, \dots, \lambda_k) = \Big(\frac{\partial f}{\partial x^1} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial x^1}, \dots, \frac{\partial f}{\partial x^n} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial x^n}, -g_1, \dots, -g_k\Big)\,.$$

Also gilt

$$\begin{aligned} &\nabla F(x^1, \dots, x^n, \lambda_1, \dots, \lambda_k) = 0 \\ \Leftrightarrow\ &\nabla f(x^1, \dots, x^n) \;=\; \sum_{i=1}^{k} \lambda_i \nabla g_i(x^1, \dots, x^n) \quad \text{und} \\ &g_1(x^1, \dots, x^n) = \ldots = g_k(x^1, \dots, x^n) = 0\,. \end{aligned}$$

Also: Falls a ein Extremum von f unter den Nebenbedingungen $g_1 = \ldots = g_k = 0$ ist und außerdem $\nabla g_1, \ldots, \nabla g_k$ linear unabhängig, so folgt als notwendige Bedingung

$$\nabla F(a, \lambda_1, \dots, \lambda_k) = 0\,.$$

VII.6.8. Beispiel.
Betrachte $f(x,y,z) = 5x + y - 3z$. Gesucht seien Extrema von f und den Nebenbedingungen, daß diese sowohl auf der Ebene $E = \{\, x + y + z = 0 \,\}$ als auch auf der Sphäre $S = \{\, x^2 + y^2 + z^2 = 1 \,\}$ liegen. E schneidet S in einer kompakten Kreiskurve $K = E \cap S$, also nimmt f auf K Maximum und Minimum an. Wir betrachten

$$F(x,y,z,\lambda_1,\lambda_2) = 5x + y - 3z - \lambda_1(x+y+z) - \lambda_2(x^2+y^2+z^2-1),$$

und wir berechnen $\nabla F = 0$ zu

$$\begin{aligned} \frac{\partial F}{\partial x} &= 5 - \lambda_1 - 2\lambda_2 x \overset{!}{=} 0, \\ \frac{\partial F}{\partial y} &= 1 - \lambda_1 - 2\lambda_2 y \overset{!}{=} 0, \\ \frac{\partial F}{\partial x} &= -3 - \lambda_1 - 2\lambda_2 z \overset{!}{=} 0, \\ \frac{\partial F}{\partial x} &= -(x+y+z) \overset{!}{=} 0, \\ \frac{\partial F}{\partial x} &= -(x^2+y^2+z^2-1) \overset{!}{=} 0\,. \end{aligned}$$

Dieses Gleichungssystem von 5 Gleichungen in 5 Variablen führt auf 2 Lösungen für (x,y,z), nämlich

$$p_1 = \Big(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\Big) \quad \text{und} \quad p_2 = \Big(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\Big)\,.$$

Notwendigerweise muss einer der beiden Punkte das Maximum sein, der andere das Minimum. Wir setzen in f ein und erhalten $f(p_1) = 4\sqrt{2}$ und $f(p_2) = -4\sqrt{2}$.

Beweis der Lagrange-Multiplikatoren-Regel VII.6.7. Wir betrachten die Abbildung $G\colon U \to \mathbb{R}^k$, gebildet aus den k Nebenbedingungen,

$$G(x^1, \dots, x^n) \;=\; \big(g_1(x), \dots, g_k(x)\big),$$

und erhalten die Menge der Punkte, welche alle k Nebenbedingungen erfüllen als das Nullstellengebilde

$$N := G^{-1}(0) = N_0(g_1) \cap \ldots \cap N_0(g_k)\,.$$

Seien nun in einem Punkt $a \in N$ die Gradienten $\nabla g_1(a)$, ..., $\nabla g_k(a)$ linear unabhängig. Dann hat das Differential $DG(a)$ vollen Rang k, d.h. $DG(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^k)$ ist surjektiv. Wir finden demnach einen Untervektorraum $V \subset \mathbb{R}^n$ komplementär zu $\ker DG(a)$, also

$$DG_{|V} : V \xrightarrow{\cong} \mathbb{R}^k\,.$$

Mit $X = \ker DG(a)$, $Y = V$ und $Z = \mathbb{R}^k$ sind für

$$F : X \times Y \subset U \to Z, \quad F(x, y) := G(a + x + y),$$

die Voraussetzungen für den Satz VII.5.2 über Implizite Funktionen erfüllt, mit $F(0, 0) = 0$, und wir finden eine offene Umgebung $U_o(0) \subset X = \ker DG(a)$ und eine mindestens einmal stetig differenzierbare Funktion

$$\varphi : U_o \to V, \quad \text{mit} \;\; G\big(a + x + \varphi(x)\big) = 0 \text{ f.a. } x \in U_o\,.$$

In der Tat erhalten wir erhalten somit eine Parametrisierung einer Umgebung von a innerhalb N in der Form $x \mapsto a + x + \varphi(x)$.
Sei nun $a \in N$ außerdem auch ein Extremum von f unter der Nebenbedingung N. Das bedeutet, daß $0 \in U_o$ ein Extremum der Funktion

$$h : U_o \to \mathbb{R}, \quad h(x) = f(a + x + \varphi(x)),$$

ist. Also muss die gemäß Satz VII.6.2 notwendige Bedingung $Dh(0) = 0$ für die Linearisierung aus $\mathcal{L}(\mathbb{R}^{n-k}, \mathbb{R})$ erfüllt sein. Diese notwendige Bedingung drückt sich nach der Kettenregel aus als

$$Df(a)\big[v + D\varphi(0)[v]\big] = 0 \quad \text{f.a. } v \in \ker DG(a)\,.$$

Andererseits gilt wegen $G\big(a + x + \varphi(x)\big) = 0$ f.a. $x \in U_o$, daß auch

$$DG(a)\big[v + D\varphi(0)[v]\big] = 0 \quad \text{f.a. } v \in \ker DG(a)\,.$$

Also ist $D\varphi(0) = 0$ auf $X = \ker DG(a)$. Daraus folgt $Df(a)[v] = 0$ für alle $v \in \ker DG(a)$, d.h.

$$\ker DG(a) \subset \ker Df(a)\,. \tag{25}$$

Wir schreiben die Jakobi-Matrix $JG(a)$ als Matrix, die aus den Gradienten $\nabla g_1(a)^T$, ..., $\nabla g_k(a)^T$ als Zeilenvektoren besteht. Die lineare Unabhängigkeit dieser Gradienten bedeutet, daß die Matrix Zeilenrang k hat. Wenn wir diese Matrix zu der $(k+1) \times n$-Matrix

$$A = \begin{pmatrix} \nabla g_1(a)^T \\ \vdots \\ \nabla g_k(a)^T \\ \nabla f(a)^T \end{pmatrix}$$

ergänzen, sehen wir, daß (25) äquivalent dazu ist, daß der Zeilenrang immer noch k beträgt. Also ist (25) äquivalent zu

$$\nabla f(a) \in \operatorname{span}\big(\nabla g_1(a), \ldots, \nabla g_k(a)\big)\,.$$

□

VII.6.1