Analysis II

Sommersemester 2002

Prof. Dr. E. Kuwert

Mathematisches Institut
Universität Freiburg

Inhaltsverzeichnis

Kapitel 6

Konvergenz im $\mathbb{R}^n$ und in Funktionenräumen

In Analysis 1 wurden die Begriffe ε-Umgebung, offene/abgeschlossene Menge,Häufungspunkt, (Folgen-)Kompaktheit im $\mathbb{R}^n$ bereits definiert. Nach dem griechischen Wort $\tau\acute{o}\pi o\varsigma$ (Ort, Platz, Stelle, Lage) wird diesbezüglich von topologischen Eigenschaften gesprochen. Unsere Übertragung der Definitionen von $\mathbb{R}$ auf $\mathbb{R}^n$ beruhte darauf, dass die Euklidische Norm anstelle des Absolutbetrags benutzt wurde.
Im ersten Abschnitt wiederholen und vertiefen wir die topologischen Begriffe, wobei wir zunächst einen beliebigen Vektorraum mit einer Norm betrachten, aber als Beispiel den $\mathbb{R}^n$ diskutieren. Das geometrische Konzept der Norm kann auf unendlichdimensionale Funktionenräume übertragen werden. Im zweiten Abschnitt werden bekannte Beispiele aus dieser Perspektive erneut betrachtet. Im Anschluß behandeln wir auf dem gegebenen Hintergrund die Konvergenz von Fourierreihen. Diese Reihen spielen sowohl in der Physik als auch in der Informatik (Signalverarbeitung) eine große Rolle.

1 $\mathbb{R}^n$ als normierter Vektorraum

Definition 1.1 (Norm) *Sei X ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Funktion $\|\cdot\| : X \to \mathbb{R}$ heißt Norm, falls gilt:*

$$\begin{aligned} \|x\| &\geq 0, \textit{ Gleichheit } \Rightarrow x = 0. \\ \|\lambda x\| &= |\lambda|\,\|x\| \quad \textit{für alle } \lambda \in \mathbb{K}, x \in X. \\ \|x+y\| &\leq \|x\| + \|y\| \quad \textit{für alle } x, y \in X \textit{ (Dreiecksungleichung).} \end{aligned}$$

Beispiel 1.1 *Die Maximumsnorm auf $X = \mathbb{K}^n$ ist*

$$\|\cdot\|_\infty : \mathbb{K}^n \to \mathbb{R}, \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Beispiel 1.2 *Die 1-Norm auf $X = \mathbb{K}^n$ ist*

$$\|\cdot\|_1 : \mathbb{K}^n \to \mathbb{R}, \quad \|x\|_1 = \sum_{i=1}^n |x_i|.$$

Definition 1.2 (Skalarprodukt) *Sei X ein $\mathbb{K}$-Vektorraum mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung $\langle \cdot, \cdot \rangle : X \times X \longrightarrow \mathbb{K}$ heißt Skalarprodukt, falls gilt:*

$$\begin{aligned} \langle x, y \rangle &= \overline{\langle y, x \rangle} \qquad \textit{(Symmetrie).} \\ \langle \lambda x + \mu y, z \rangle &= \lambda \langle x, z \rangle + \mu \langle y, z \rangle \quad \forall \lambda, \mu \in \mathbb{K}, x, y, z \in X \quad \textit{(Bilinearität).} \\ \langle x, x \rangle &\geq 0, \; \textit{Gleichheit} \Rightarrow x = 0. \qquad \textit{(Positivität).} \end{aligned}$$

Im Fall $\mathbb{K} = \mathbb{C}$ ist das Skalarprodukt im zweiten Eintrag *konjugiert* linear, denn

$$\begin{aligned} \langle x, \lambda y + \mu z \rangle &= \overline{\langle \lambda y + \mu z, x \rangle} \\ &= \overline{\lambda \langle y, x \rangle + \mu \langle z, x \rangle} \\ &= \overline{\lambda} \langle x, y \rangle + \overline{\mu} \langle x, z \rangle. \end{aligned}$$

Beispiel 1.3 *Für $z = (z_1, \ldots, z_n)$, $w = (w_1, \ldots, w_n) \in \mathbb{C}^n$ ist das Standardskalarprodukt gegeben durch*

$$\mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}, \quad \langle z, w \rangle = \sum_{i=1}^{n} z_i \overline{w}_i.$$

Satz 1.1 *Sei $\langle \cdot \cdot \rangle : X \times X \to \mathbb{K}$ ein Skalarprodukt auf dem $\mathbb{K}$-Vektorraum X. Dann ist die Funktion*

$$\| \cdot \| : X \longrightarrow \mathbb{R}, \quad \|x\| = \sqrt{\langle x, x \rangle} \geq 0$$

eine Norm.

Zum Beweis benötigen wir

Satz 1.2 (Ungleichung von Cauchy-Schwarz) *Für ein Skalarprodukt auf einem $\mathbb{K}$-Vektorraum X gilt die Ungleichung*

$$|\langle x, y \rangle| \leq \|x\| \, \|y\| \quad \forall x, y \in X. \tag{1.1}$$

Gleichheit tritt genau dann ein, wenn x, y linear abhängig sind.

Beweis: In Schritt 1 beweisen wir die Ungleichung für $\xi, \eta \in X$, welche die folgenden Normierungsbedingungen erfüllen:

$$\|\xi\| = \|\eta\| = 1 \quad \text{und} \quad \langle \xi, \eta \rangle \in [0, \infty). \tag{1.2}$$

Die allgemeine Aussage folgt dann in Schritt 2 mittels Skalierung (Streckung). In Schritt 3 wird der Gleichheitsfall untersucht.

Schritt 1: Für $\xi, \eta \in X$ mit (1.2) gilt

$$\|\xi\| \, \|\eta\| - |\langle \xi, \eta \rangle| = \frac{1}{2} \left(\|\xi\|^2 + \|\eta\|^2 - \langle \xi, \eta \rangle - \langle \eta, \xi \rangle \right) = \frac{1}{2} \|\xi - \eta\|^2 \geq 0.$$

Schritt 2: Gilt die Ungleichung für $\xi, \eta \in X$, so auch für $x = \lambda \xi, y = \mu \eta$. Denn wir haben

$$|\langle x, y \rangle| = |\lambda \overline{\eta} \langle \xi, \eta \rangle| \leq |\lambda| \, |\eta| \, \|\xi\| \, \|\eta\| = \|x\| \, \|y\|.$$

Es bleibt zu zeigen, dass jedes Paar $x, y \in X$ in der Form $x = \lambda\xi, y = \mu\eta$ geschrieben werden kann, wobei ξ, η die Normierung (1.2) erfüllen. Wir können dazu o.B.d.A. $\langle x, y\rangle \neq 0$ annehmen. Mit $\gamma \in \mathbb{C}$, $|\gamma| = 1$, machen wir den Ansatz

$$\xi = \gamma \frac{x}{\|x\|}, \, \eta = \frac{y}{\|y\|} \quad \Rightarrow \quad \|\xi\| = \|\eta\| = 1,$$

und bestimmen γ aus der Bedingung $\langle \xi, \eta\rangle > 0$:

$$\gamma := \frac{\overline{\langle x, y\rangle}}{|\langle x, y\rangle|} \quad \Rightarrow \quad \langle \xi, \eta\rangle = \gamma \frac{\langle x, y\rangle}{\|x\| \, \|y\|} = \frac{|\langle x, y\rangle|}{\|x\| \, \|y\|} > 0.$$

Die gewünschte Darstellung folgt nun mit $\lambda = \|x\|/\gamma$ und $\mu = \|y\|$.
Schritt 3: Angenommen, für $x, y \in X$ gilt Gleichheit in (1.1). Ist $\langle x, y\rangle = \|x\| \, \|y\| = 0$, so folgt $x = 0$ oder $y = 0$ und x, y sind trivialerweise linear abhängig. Andernfalls können wir ξ, η wie oben definieren. Die Gleichheit in (1.1) gilt dann auch für ξ, η und aus Schritt 1 folgt $\xi = \eta$ bzw.

$$\gamma \frac{x}{\|x\|} = \frac{y}{\|y\|},$$

das heißt die Vektoren x, y sind linear abhängig. □

Beweis: von Satz 1.1
Wir zeigen nur die Dreiecksungleichung:

$$\begin{aligned} \|x + y\|^2 &= \|x\|^2 + 2\,\mathrm{Re}\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \, \|y\| + \|y\|^2 \\ &\leq (\|x\| + \|y\|)^2. \end{aligned}$$

□

Definition 1.3 *Für $p \in [1, \infty)$ ist die p-Norm auf $\mathbb{K}^n$ gegeben durch*

$$\| \cdot \|_p : \mathbb{K}^n \longrightarrow [0, \infty), \quad \|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

Für $p = 1$ erhalten wir die oben definierte 1-Norm, für $p = 2$ die Euklidische Norm. Die Normeigenschaft muss für $\| \cdot \|_p$ allerdings noch gezeigt werden, wobei die Schwierigkeit wieder im Beweis der Dreiecksungleichung liegt.

Lemma 1.1 (Youngsche Ungleichung) *Seien $p, q \in (1, \infty)$ mit $1/p + 1/q = 1$. Dann gilt*

(1.3) $$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad \textit{für alle } x, y \geq 0.$$

Beweis: Wir betrachten für festes $y > 0$ die Funktion

$$f : [0, \infty) \to \mathbb{R}, \quad f(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy.$$

f ist differenzierbar auf $(0,\infty)$ mit $f'(x) = x^{p-1} - y$. Mit $x_0 := y^{\frac{1}{p-1}}$ folgt $f'(x) > 0$ für $x > x_0$, $f'(x) < 0$ für $x < x_0$; somit ist f fallend auf $[0, x_0]$ und wachsend auf $[x_0, \infty)$. Wir erhalten für alle $x > 0$ (beachte $\frac{p}{p-1} = q$)

$$f(x) \geq f(x_0) = \frac{x_0^p}{p} + \frac{y^q}{q} - x_0 y = 0.$$

Gleichheit kann nur eintreten, wenn $x = x_0$ bzw. $x^p = y^q$ ist. $\square$

Das folgende Lemma enthält als Spezialfall $p = q = 2$ die Ungleichung von Cauchy-Schwarz im $\mathbb{K}^n$.

Satz 1.3 (Höldersche Ungleichung) *Seien $p, q \in (1,\infty)$ mit $1/p + 1/q = 1$. Dann gilt*

$$|\langle x, y\rangle| \leq \|x\|_p \|y\|_q \quad \textit{für alle } x, y \in \mathbb{K}^n. \tag{1.4}$$

Beweis: o.B.d.A. $\langle x, y\rangle \neq 0$. Es reicht aus, die Ungleichung für $\|x\|_p = \|y\|_q = 1$ zu zeigen; der allgemeine Fall folgt durch Skalierung. Nun gilt

$$\begin{aligned}
|\langle x, y\rangle| &= \Big|\sum_{i=1}^n x^i y^i\Big| \\
&\leq \sum_{i=1}^n |x^i|\,|y^i| \\
&\leq \sum_{i=1}^n \left(\frac{1}{p}|x^i|^p + \frac{1}{q}|y^i|^q\right) \qquad \text{(Lemma 1.1)} \\
&= \frac{1}{p}\underbrace{\sum_{i=1}^n |x^i|^p}_{=1} + \frac{1}{q}\underbrace{\sum_{i=1}^n |y^i|^q}_{=1} \\
&= 1.
\end{aligned}$$

$\square$

Satz 1.4 (Minkowski-Ungleichung) *Sei $p \in (1,\infty)$. Dann gilt*

$$\|x+y\|_p \leq \|x\|_p + \|y\|_q \quad \textit{für alle } x, y \in \mathbb{K}^n. \tag{1.5}$$

Beweis: oBdA $x + y \neq 0$ (sonst trivial). Durch Skalierung können wir annehmen, dass $\|x+y\|_p = 1$. Es gilt

$$\begin{aligned}
\|x+y\|_p &= \|x+y\|_p^p \\
&= \sum_{i=1}^n |x_i + y_i|^p \\
&\leq \sum_{i=1}^n (|x_i|\,|x_i+y_i|^{p-1} + |y_i|\,|x_i+y_i|^{p-1}) \\
&\leq \left(\Big(\sum_{i=1}^n |x_i|^p\Big)^{\frac{1}{p}} + \Big(\sum_{i=1}^n |y_i|^p\Big)^{\frac{1}{p}} \right) \underbrace{\left(\sum_{i=1}^n |x_i+y_i|^p\right)^{\frac{p-1}{p}}}_{=1} \\
&= \|x\|_p + \|y\|_p.
\end{aligned}$$

□

Bemerkung. Für $p = \infty$, $q = 1$ (also formal $\frac{1}{p} + \frac{1}{q} = 1$) gilt ebenfalls

$$|\langle x, y\rangle| \leq \sum_{i=1}^{n} |x_i|\,|y_i| \leq \max_{1\leq i\leq n} |x_i| \sum_{i=1}^{n} |y_i| = \|x\|_\infty \, \|y\|_1,$$

sowie die jeweiligen Dreiecksungleichungen.

Definition 1.4 *Sei X ein Vektorraum mit Norm $\|\cdot\|$ und $x_0 \in X$, $r > 0$. Die Menge*

$$B_r(x_0) = \{x \in X : \|x - x_0\| < r\}$$

heißt offene Kugel um x_0 mit Radius r.

Beispiel 1.4 *Für die Normen $\|\cdot\|_2$, $\|\cdot\|_1$ und $\|\cdot\|_\infty$ sehen die Einheitskugeln um den Nullpunkt wie folgt aus:*

Definition 1.5 *Sei $\|\cdot\|$ eine Norm auf dem Vektorraum X. Die Menge $\Omega \subset X$ heißt offen bzgl. $\|\cdot\|$, falls zu jedem $x \in \Omega$ ein $\varepsilon > 0$ existiert mit $B_\varepsilon(x) \subset \Omega$.*

Beispiel 1.5 *Die Kugel $B_r(x_0)$ ist offen im Sinn der Definition 1.5. Sei nämlich $x \in B_r(x_0)$ gegeben. Wähle $\varepsilon = r - \|x - x_0\| > 0$ und schließe für $y \in B_\varepsilon(x)$*

$$\|y - x_0\| \leq \|y - x\| + \|x - x_0\| < \varepsilon + \|x - x_0\| = r.$$

Also ist $B_\varepsilon(x) \subset B_r(x_0)$, was zu zeigen war.

Satz 1.5 *Für die offenen Teilmengen eines normierten Raumes X gilt:*

a) *$\emptyset, X$ sind offen.*

b) *Ω_i offen für $i = 1, \ldots, N \quad \Rightarrow \quad \bigcap_{i=1}^{N} \Omega_i$ offen (wobei $N \in \mathbb{N}$).*

c) *Ω_λ offen für $\lambda \in \Lambda \quad \Rightarrow \quad \bigcup_{\lambda\in\Lambda} \Omega_\lambda$ offen (Λ beliebige Indexmenge).*

Beweis:

a) klar.

b) Sei $x \in \bigcap_{i=1}^N \Omega_i$. Zu $i \in \{1, \ldots, N\}$ gibt es dann ein $\varepsilon_i > 0$ mit $B_{\varepsilon_i}(x) \subset \Omega$. Setze $\varepsilon := \min_{1\leq i\leq N} \varepsilon_i > 0$ und erhalte $B_\varepsilon(x) \subset \Omega_i$ für alle i bzw. $B_\varepsilon(x) \subset \bigcap_{i=1}^N \Omega_i$.

c) Sei $x \in \bigcup_{\lambda\in\Lambda} \Omega_\lambda$. Dann ist $x \in \Omega_{\lambda_0}$ für mindestens ein $\lambda_0 \in \Lambda$, also gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \Omega_{\lambda_0} \subset \bigcup_{\lambda\in\Lambda} \Omega_\lambda$ für geeignetes $\varepsilon > 0$.

□

Achtung: Ein abzählbarer Schnitt von offenen Mengen ist nicht notwendig offen, zum Beispiel ist $\bigcap_{n\in\mathbb{N}} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$ nicht offen in $\mathbb{R}$.

Ist $\Omega \subset X$ offen und $x \in \Omega$, so heißt Ω offene Umgebung von x. Insbesondere wird die offene Kugel $B_\varepsilon(x)$ als ε-Umgebung von x bezeichnet.

Lemma 1.2 *In einem normierten Raum X gibt es zu $x, y \in X$ mit $x \neq y$ ein $\varepsilon > 0$ mit $B_\varepsilon(x) \cap B_\varepsilon(y) = \emptyset$.*

BEWEIS: Sei $z \in B_\varepsilon(x) \cap B_\varepsilon(y)$. Dann folgt $\|x - y\| \leq \|x - z\| + \|z - x\| < 2\varepsilon$. Also ist die Behauptung richtig für jedes $\varepsilon \leq \|x - y\|/2$. $\square$

Definition 1.6 *Sei X ein Vektorraum mit Norm $\|\cdot\|$. Die Folge $(x_k)_{k\in\mathbb{N}}$ von Punkten $x_k \in X$ konvergiert gegen $x \in X$ bzgl. $\|\cdot\|$, falls gilt:*

$$\|x_k - x\| \to 0 \text{ mit } k \to \infty \qquad (\text{Notation: } x_k \to x \text{ oder } \lim_{k\to\infty} x_k = x).$$

Äquivalent dazu ist: Für alle $\varepsilon > 0$ gibt es ein $K \in \mathbb{R}$ mit $x_k \in B_\varepsilon(x)$ für alle $k > K$.

Der Grenzwert ist eindeutig bestimmt, denn wäre $y \neq x$ ebenfalls Grenzwert von (x_k), so wählen wir $\varepsilon > 0$ wie in Lemma 1.2 und erhalten

$$x_k \in B_\varepsilon(x) \cap B_\varepsilon(y) = \emptyset \quad \text{für } k \text{ hinreichend groß,}$$

ein Widerspruch.

Definition 1.7 *Eine Teilmenge A eines normierten Raums X heißt abgeschlossen, wenn folgende Implikation stets gilt:*

$$x_k \in A, \quad x_k \to x \quad \Rightarrow \quad x \in A.$$

Satz 1.6 *Sei $\|\cdot\|$ eine Norm auf X. Für $M \subset X$ gilt:*

$$M \text{ offen} \quad \Leftrightarrow \quad X\backslash M \text{ abgeschlossen.}$$

BEWEIS: Kapitel 1, Satz 5.6. $\square$

Folgerung 1.1 *Für die abgeschlossenen Teilmengen eines normierten Raums X gilt:*

a) *$\emptyset, X$ sind abgeschlossen.*

b) *A_i abgeschlossen für $i = 1, \ldots N \quad \Rightarrow \quad \bigcup_{i=1}^N A_i$ abgeschlossen $\quad (N \in \mathbb{N})$.*

c) *A_λ abgeschlossen für $\lambda \in \Lambda \quad \Rightarrow \quad \bigcap_{\lambda\in\Lambda} A_\lambda$ abgeschlossen $\quad (\Lambda$ beliebige Indexmenge).*

BEWEIS: Folgt aus Satz 1.5 und Satz 1.6. $\square$

Definition 1.8 *Sei $\|\cdot\|$ eine Norm auf X und $M \subset X$.*

a) *Die Menge der inneren Punkte von M ist*

$$\operatorname{int} M = \{x \in M : \exists\, \varepsilon > 0 \text{ mit } B_\varepsilon(x) \subset M\}.$$

b) *Der Abschluss von M ist*

$$\overline{M} = \{x \in X : \forall\, \varepsilon > 0 \text{ ist } B_\varepsilon(x) \cap M \neq \emptyset\}.$$

c) *Der Rand von M ist*

$$\partial M = \{x \in X : \forall \varepsilon > 0 \text{ ist } B_\varepsilon(x) \cap M \neq \emptyset \text{ sowie } B_\varepsilon(x) \cap (X \backslash M) \neq \emptyset\}.$$

Beispiel 1.6 *Überlegen Sie, dass für die offene Kugel $B_r(x) = \{y \in X : \|y - x\| < r\}$ gilt:*

$$\begin{aligned} \operatorname{int} B_r(x) &= B_r(x), \\ \overline{B_r(x)} &= \{y \in X : \|y - x\| \leq r\}, \\ \partial B_r(x) &= \{y \in X : \|y - x\| = r\}. \end{aligned}$$

Satz 1.7 *Sei M Teilmenge des normierten Raums X.*

(a) $\operatorname{int} M$ *ist offen und es gilt die Implikation*

$$\Omega \text{ offen }, \Omega \subset M \quad \Rightarrow \quad \Omega \subset \operatorname{int} M.$$

(b) $\overline{M}$ *ist abgeschlossen, und es gilt die Implikation*

$$A \text{ abgeschlossen }, A \supset M \quad \Rightarrow \quad A \supset \overline{M}.$$

(c) ∂M *ist abgeschlossen und es gilt $\partial M = \overline{M} \backslash \operatorname{int} M$.*

Beweis:

a) Sei $x \in \operatorname{int} M$, also $B_r(x) \subset M$ für ein $r > 0$. Für $y \in B_r(x)$ gilt dann $B_\varepsilon(y) \subset B_r(x) \subset M$ mit $\varepsilon = r - \|y - x\| > 0$, vgl. Beispiel 1.5. Es folgt $B_r(x) \subset \operatorname{int} M$, also ist $\operatorname{int} M$ offen. Sei nun $\Omega \subset M$ offen. Zu $x \in \Omega$ gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \Omega \subset M$ und folglich $x \in \operatorname{int} M$. Damit ist a) gezeigt.

b) Nach Definition gilt $X \backslash \overline{M} = \operatorname{int}(X \backslash M)$. Nach a) und Satz 1.6 ist $\overline{M}$ abgeschlossen. Ist andererseits $A \subset X$ beliebige abgeschlossene Menge mit $A \supset M$, so ist $X \backslash A$ offen und $X \backslash A \subset X \backslash M$, also $X \backslash A \subset \operatorname{int}(X \backslash M)$ nach a) und somit $A \supset \overline{M}$. Dies beweist b).

c) Nach Definition gilt $\partial M = \overline{M} \cap (\overline{X \backslash M})$, also ist ∂M abgeschlossen nach b) und Folgerung 1.1. Ferner ist nach Definition $X \backslash \operatorname{int} M = \overline{X \backslash M}$, folglich $\partial M = \overline{M} \cap (X \backslash \operatorname{int} M) = \overline{M} \backslash \operatorname{int} M$.

□

$\overline{M}$ ist die disjunkte Vereinigung der Häufungspunkte und der isolierten Punkte von M.

Definition 1.9 *Die Menge der Häufungspunkte von M ist*

$$\{x \in X : \forall \varepsilon > 0 \quad \text{ist } \#(B_\varepsilon(x) \cap M) = \infty\}.$$

Die Menge der isolierten Punkte von M ist

$$\big\{x \in M : \exists \varepsilon > 0 \quad \text{mit } B_\varepsilon(x) \cap M = \{x\}\big\}.$$

Genau dann ist $x \in X$ Häufungspunkt von M, wenn es eine Folge von Punkten $x_k \in M$ gibt mit $x_k \neq x$ für alle k, so dass $\lim_{k \to \infty} x_k = x$, siehe Kapitel 2, Lemma 5.6.

Definition 1.10 *Seien $(X, \|\cdot\|)$ und $(Y, \|\cdot\|)$ normierte Räume und $D \subset X$. Die Abbildung $f : D \to Y$ heißt stetig in $x_0 \in D$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für $x \in D$ gilt:*

$$\|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \varepsilon.$$

Definition 1.11 *Die Funktion $f : D \to Y$ heißt Lipschitzstetig mit Konstante $L \geq 0$, falls gilt:*

$$\|f(x) - f(x')\| \leq L\|x - x'\| \quad \textit{für alle } x, x' \in D.$$

Beispiel 1.7 *Die Normfunktion $f : X \to \mathbb{R}$, $f(x) = \|x\|$ ist in jedem normierten Raum Lipschitzstetig mit Konstante $L = 1$, denn aus der Dreiecksungleichung folgt*

$$|f(x) - f(x')| = \big|\|x\| - \|x'\|\big| \leq \|x - x'\| \quad \textit{für alle } x, x' \in D.$$

Wir wollen nun zeigen, dass die eingeführten topologischen Begriffe wie Offenheit, Konvergenz, Abgeschlossenheit und Stetigkeit im $\mathbb{R}^n$ bzw. *in endlichdimensionalen Räumen* nicht von der gewählten Norm abhängen. Die Normen sind also möglicherweise geometrisch verschieden - sie haben zum Beispiel andere Isometriegruppen - aber die zugehörigen Topologien sind stets gleich.

Definition 1.12 *Zwei Normen $\|\cdot\|$, $\|\cdot\|'$ auf X heißen äquivalent, wenn es Konstanten $0 < \lambda \leq \Lambda < \infty$ gibt mit*

$$\lambda\|x\|' \leq \|x\| \leq \Lambda\,\|x\|' \quad \textit{für alle } x \in X.$$

Die Normen sind also jeweils gegeneinander abgeschätzt, und es gilt für die zugehörigen Kugeln B, B' bzgl. $\|\cdot\|$ bzw. $\|\cdot\|'$

$$B'_{r/\Lambda}(x) \subset B_r(x) \subset B'_{r/\lambda}(x).$$

Als Konsequenz ist eine Teilmenge von X genau dann offen bzgl. $\|\cdot\|$, wenn sie offen bzgl. $\|\cdot\|'$ ist; analog natürlich für die abgeschlossenen Mengen. Eine Folge ist genau dann konvergent bzgl. $\|\cdot\|$, wenn sie bzgl. der äquivalenten Norm $\|\cdot\|'$ konvergiert.

Satz 1.8 (Äquivalenz von Normen im $\mathbb{R}^n$) *Je zwei Normen $\|\cdot\|, \|\cdot\|'$ auf $\mathbb{R}^n$ sind äquivalent.*

Beweis: *Vorbemerkung.* Durch „äquivalent“ist eine Äquivalenzrelation auf der Menge der Normen gegeben. Insbesondere gilt die Transitivität:

$$\lambda\,\|x\|' \leq \|x\| \leq \Lambda\,\|x\|',\ \mu\,\|x\|'' \leq \|x\|' \leq M\|x\|'' \quad \Rightarrow \quad \lambda\mu\,\|x\|'' \leq \|x\| \leq \Lambda\,M\,\|x\|''.$$

Es reicht deshalb aus zu zeigen, dass jede Norm $\|\cdot\|$ äquivalent zur Euklidischen Norm $|\cdot|$ ist.

Schritt 1: *Abschätzung nach oben*

$$\begin{aligned}
\|x\| &= \|x^1\,e_1 + \ldots + x^n\,e_n\| \\
&\leq |x^1|\,\|e_1\| + \ldots + |x^n|\,\|e_n\| \quad \text{(Definition Norm)} \\
&\leq \left(\sum_{i=1}^n \|e_i\|^2\right)^{1/2} \left(\sum_{i=1}^n |x^i|^2\right)^{1/2} \qquad \text{(Cauchy-Schwarz)} \\
&= \Lambda\,|x| \quad \text{mit} \quad \Lambda := \left(\sum_{i=1}^n \|e_i\|^2\right)^{1/2}.
\end{aligned}$$

Schritt 2 *Abschätzung nach unten*
Die Funktion $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|x\|$ ist Lipschitzstetig *bzgl. der Euklidischen Norm*, denn nach Dreiecksungleichung und Schritt 1 ist

$$|f(x) - f(y)| = \big|\|x\| - \|y\|\big| \le \|x - y\| \le \Lambda|x - y|.$$

Die Sphäre $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ ist abgeschlossen und beschränkt bzgl. $|\cdot|$, also kompakt bzgl. $|\cdot|$ (Heine-Borel, Satz 4.1. in Kapitel 2). Es gibt deshalb ein $\xi \in S^{n-1}$ mit $f(\xi) = \inf_{x \in S^{n-1}} f(x) =: \lambda$ (Kapitel 2, Satz 4.2), und es folgt $\lambda = \|\xi\| > 0$. Also erhalten wir

$$\|x\| = |x| \underbrace{\|\frac{x}{|x|}\|}_{\ge \lambda} \ge \lambda |x| \quad \text{für alle } x \in \mathbb{R}^n.$$

□

Bemerkung. Die Aussage des Satzes gilt in jedem endlichdimensionalen $\mathbb{R}$-Vektorraum V. Wähle dazu eine Basis $\{v_1, \ldots, v_n\}$ mit zugehörigem Koordinatenisomorphismus

$$\varphi : \mathbb{R}^n \to V,\ \varphi(x) = \sum_{i=1}^n x_i v_i.$$

Sind $\|v\|, \|v\|'$ Normen auf V, so erhalten wir auf $\mathbb{R}^n$ die induzierten Normen $\|x\|_{\mathbb{R}^n} = \|\varphi(x)\|$ bzw. $\|x\|'_{\mathbb{R}^n} = \|\varphi(x)\|'$; die Prüfung der Normeigenschaften sei den Hörer/innen überlassen. Jetzt wenden wir Satz 1.8 auf diese Normen $\|\cdot\|_{\mathbb{R}^n}$, $\|\cdot\|'_{\mathbb{R}^n}$ an.

Folgerung 1.2 *Sei $(X, \|\cdot\|)$ endlichdimensional. Dann gilt für eine Folge $(x_k)_{k \in \mathbb{N}}$ und $x \in X$:*

$$\|x_k - x\| \to 0 \quad \Leftrightarrow \quad \lim_{k \to \infty} (x_k)_i = x_i \text{ für alle } i = 1, \ldots, n,$$

wobei $x_1, \ldots, x_n$ die Koordinaten bzgl. einer beliebigen Basis sind.

Definition 1.13 (Vollständigkeit) *Ein normierter Raum $(X, \|\cdot\|)$ heißt vollständig bzw. Banachraum, wenn jede Cauchyfolge $(x_k)_{k \in \mathbb{N}}$ in X gegen ein $x \in X$ konvergiert.*

Folgerung 1.3 *Jeder endlichdimensionale, normierte Raum $(X, \|\cdot\|)$ ist vollständig.*

Definition 1.14 (Folgenkompaktheit) *Sei $\|\cdot\|$ eine Norm auf X. Dann heißt $K \subset X$ kompakt, wenn gilt: jede Folge $(x_k)_{k \in \mathbb{N}}$ mit $x_k \in K$ hat eine Teilfolge $(x_{k_p})_{p \in \mathbb{N}}$, die gegen ein $x \in K$ konvergiert.*

Satz 1.9 (Heine-Borel) *Sei $(X, \|\cdot\|)$ ein endlichdimensionaler, normierter Raum. Dann gilt für $K \subset X$:*

$$K \text{ abgeschlossen und beschränkt} \quad \Rightarrow \quad K \text{ kompakt}.$$

Beweis: Folgt mit Satz 1.8 aus Satz 4.1 in Kapitel 2. □

Bemerkung. Ist umgekehrt $K \subset X$ kompakt, so auch abgeschlossen und beschränkt; dies folgt leicht und gilt sogar ohne die Voraussetzung „X endlichdimensional".

Folgende Aussagen über Stetigkeit/kompakte Mengen sind oft nützlich.

Satz 1.10 *Sei $f : (X, \|\cdot\|) \to (Y, \|\cdot\|)$ stetig und $K \subset X$ kompakt. Dann gilt:*

(1) *$M = f(K)$ ist kompakt.*

(2) *Ist f injektiv, so ist $f^{-1} : M \to X$ stetig.*

Beweis:

(1) Sei (y_k) Folge in M, also $y_k = f(x_k)$. Da K kompakt, gibt es eine Teilfolge mit $x_{k_j} \to x \in K$. Da f stetig, folgt $y_{k_j} = f(x_{k_j}) \to f(x) \in M$.

(2) Angenommen, f^{-1} ist in $y = f(x)$ unstetig. Dann gibt es eine Folge $y_k = f(x_k) \to y$, aber $\|x_k - x\| \geq \varepsilon$ für alle k. Da K kompakt, gibt es eine Teilfolge $x_{k_j} \to x' \in K$ und es gilt $\|x' - x\| \geq \varepsilon$. Da f stetig in x', gilt aber $f(x') = \lim_{j\to\infty} f(x_{k_j}) = y$. Widerspruch zu f injektiv. □

Satz 1.11 (Extrema) *Eine stetige Funktion $f : K \to \mathbb{R}$ auf einer kompakten Menge $K \subset (X, \|\cdot\|)$, $K \neq \emptyset$, ist beschränkt und nimmt ihr Infimum und Supremum an.*

Beweis: (vgl. Kapitel 2, Satz 4.2). □

Beispiel 1.8 *Sei $K \subset X$ kompakt. Dann gibt es zu $x_0 \in X$ ein $x \in K$ mit*

$$\|x - x_0\| = \inf_{y \in K} \|y - x_0\| = \operatorname{dist}(x_0, K).$$

Der Punkt x_0 ist in der Regel nicht eindeutig bestimmt. Sei zum Beispiel $K = \{x \in \mathbb{R}^2 : \|x\|_\infty \leq 1\}$ und $x_0 = (2, 0)$. Dann gilt

$$\operatorname{dist}(x_0, K) = 1 = \|(1, y) - x_0\|_\infty \quad \forall y \in [-1, 1].$$

Satz 1.12 *Sei $K \subset (X, \|\cdot\|)$ kompakt. Dann ist jede Funktion $f \in C^0(K)$ gleichmäßig stetig.*

Beweis: siehe Lemma 1.5, Kapitel 4. □

2 Funktionenräume und Fourierreihen

Wir beginnen damit, Beispiele von Funktionenräumen vorzustellen. Als erstes betrachten wir den Raum

$$\mathcal{B}(D) = \{f : D \to \mathbb{R} : f \text{ ist beschränkt}\} \quad (D \subset \mathbb{R}^n).$$

Auf $\mathcal{B}(D)$ haben wir die Supremumsnorm $\|f\|_D = \sup_{x \in D} |f(x)|$ (Kapitel 4, Lemma 1.3).

Lemma 2.1 $\big(\mathcal{B}(D), \|\cdot\|_D\big)$ *ist ein Banachraum.*

Beweis: Sei $(f_k)_{k \in \mathbb{N}}$ Cauchyfolge in $\mathcal{B}(D)$. Zu $\varepsilon > 0$ gibt es ein $K(\varepsilon) \in \mathbb{R}$ mit

$$k, l > K(\varepsilon) \Rightarrow \|f_k - f_l\|_D < \varepsilon.$$

Das bedeutet

$$k, l > K(\varepsilon) \Rightarrow |f_k(x) - f_l(x)| < \varepsilon \quad \text{für jedes } x \in D. \tag{2.6}$$

Die Folge $\big(f_k(x)\big)_{k\in\mathbb{N}}$ ist also Cauchyfolge in $\mathbb{R}$ und konvergiert. Wir definieren

$$f : D \to \mathbb{R}, \quad f(x) = \lim_{k\to\infty} f_k(x).$$

Indem wir in (2.6) $l \to \infty$ gehen lassen, folgt

$$k > K(\varepsilon) \Rightarrow |f_k(x) - f(x)| \le \varepsilon \quad \text{für jedes } x \in D. \tag{2.7}$$

Wählen wir insbesondere $\varepsilon := 1$, so folgt für $k > K(1)$ und alle $x \in D$

$$|f(x)| \le |f_k(x)| + |f_k(x) - f(x)| \le \|f_k\|_D + 1.$$

Damit ist $f \in \mathcal{B}(D)$ und nach (2.7) gilt $\|f_k - f\|_D \to 0$ mit $k \to \infty$. □

Bemerkung. Lemma 2.1 gilt einschließlich Beweis analog für Funktionen mit Werten in $\mathbb{R}^m$ oder in $\mathbb{C}$.

Lemma 2.2 *Sei X ein Untervektorraum des Banachraums $(Y, \|\cdot\|)$. Ist $X \subset Y$ abgeschlossen, so ist $(X, \|\cdot\|)$ ebenfalls ein Banachraum.*

Beweis: Sei $(x_k)_{k\in\mathbb{N}}$ Cauchyfolge in X. Dann ist (x_k) Cauchyfolge in Y und nach Voraussetzung gilt $x_k \to y \in Y$. Da X abgeschlossen, ist $y \in X$. □

Satz 2.1 (Vollständigkeit von $C^0(D)$) *Sei $D \subset \mathbb{R}^n$ kompakt. Dann ist $C^0(D)$, versehen mit der Supremumsnorm, ein Banachraum.*

Beweis: Nach Satz 4.2 in Kapitel 2 ist jedes $f \in C^0(D)$ beschränkt und somit $C^0(D) \subset \mathcal{B}(D)$ ein Untervektorraum. Aber nach Kapitel 5, Satz 2.2 ist $C^0(D)$ abgeschlossen in $\mathcal{B}(D)$, das heißt es gilt

$$f_k \in C^0(D), f \in \mathcal{B}(D), \|f_k - f\|_D \to 0 \quad \Rightarrow \quad f \in C^0(D).$$

Die Behauptung folgt aus Lemma 2.1 und Lemma 2.2. □

Bemerkung. Analog wieder für $f \in C^0(D, \mathbb{R}^m)$ bzw. $f \in C^0(D, \mathbb{C})$.

Beispiel 2.1 *Sei $D = [0,1]$. Die Menge $\overline{B_1(0)} = \{f \in C^0(D) : \|f\|_D \le 1\} \subset C^0(D)$ ist abgeschlossen und beschränkt, aber nicht kompakt. Betrachte dazu $f_k \in \overline{B_1(0)}$ mit*

$$f_k(x) = \begin{cases} 1 - kx & 0 \le x \le 1/k, \\ 0 & \textit{sonst.} \end{cases}$$

Angenommen, es gibt eine Teilfolge $(f_{k_j})_{j\in\mathbb{N}}$ und $f \in C^0(D)$ mit $\|f_{k_j} - f\|_D \to 0$. Dann gilt

$$f(x) = \lim_{j\to\infty} f_{k_j}(x) = \begin{cases} 1 & x = 0, \\ 0 & 0 < x \le 1, \end{cases}$$

ein Widerspruch.

In Verallgemeinerung der C^0-Norm können die C^m-Normen definiert werden, die höhere Ableitungen der Funktion mit einbeziehen.

Satz 2.2 *Sei $I \subset \mathbb{R}$ ein kompaktes Intervall. Dann ist für $m \in \mathbb{N}_0$ der Raum $C^m(I)$, versehen mit der Norm*

$$\|f\|_{C^m(I)} = \sum_{j=0}^{m} \|f^{(j)}\|_I,$$

ein Banachraum.

Beweis: Wir zeigen die Behauptung für $m = 1$, der allgemeine Fall ist ähnlich. Sei $(f_k)_{k\in\mathbb{N}}$ eine Cauchyfolge bzgl. $\|\cdot\|_{C^1(I)}$. Dann gilt nach Satz 2.1

$$f_k \to f \in C^0(I) \text{ bzgl. } \|\cdot\|_I, \quad f_k' \to g \in C^0(I) \text{ bzgl. } \|\cdot\|_I.$$

Nach Kapitel 5, Satz 2.4, folgt $f \in C^1(I)$ und $f' = g$, also sogar $f_k \to f$ bzgl. $\|\cdot\|_{C^1(I)}$. □

Beispiel 2.2 *Auf $C^1(I)$ sind die Normen $\|\cdot\|_{C^0(I)} = \|\cdot\|_I$ und $\|\cdot\|_{C^1(I)}$ nicht äquivalent. Betrachte zum Beispiel*

$$f_k : [0,\pi] \to \mathbb{R},\ f_k(x) = \frac{1}{k}\sin kx \qquad \textit{und } f : [0,\pi] \to \mathbb{R},\ f(x) = 0 \textit{ für alle } x.$$

Dann gilt $\|f_k - f\|_{C^0(I)} = \frac{1}{k} \to 0$ mit $k \to \infty$, aber

$$\|f_k - f\|_{C^1(I)} = \|f_k - f\|_I + \|f_k' - f'\|_I = \frac{1}{k} + 1.$$

Wir kommen jetzt zu Normen, die durch Integrale definiert sind.

Satz 2.3 *Sei $I = [a,b] \subset \mathbb{R}$ kompaktes Intervall. Dann ist*

$$\|f\|_p = \Big(\int_I |f|^p\Big)^{\frac{1}{p}} \qquad \textit{für } p \in [1,\infty)$$

eine Norm auf $C^0(I)$.

Beweis: Klar ist, dass $\|f\|_p \geq 0$ und $\|\lambda f\|_p = |\lambda|\,\|f\|_p$.

a) $\|f\|_p = 0 \quad \Rightarrow \quad f = 0.$
   Sei $f \in C^0(I)$ mit $|f(x_0)| \neq 0$. Dann gibt es ein $\varepsilon > 0$ und ein $\delta > 0$ mit

   $$|f(x)| \geq \varepsilon \quad \text{für } x \in I \cap (x_0 - \delta, x_0 + \delta),$$

   siehe Lemma 2.1 in Kapitel 2. Aus der Monotonie des Integrals (Satz 1.6 in Kapitel 4) folgt

   $$\Big(\int_I |f|^p\Big)^{1/p} \geq \Big(\int_{I\cap(x_0-\delta,x_0+\delta)} \varepsilon^p\Big)^{1/p} \geq \varepsilon\,|I \cap (x_0 - \delta, x_0 + \delta)|^{1/p} > 0.$$

b) $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ (Dreiecksungleichung)
   Diese soll in Übungsaufgabe 2, Serie 2 bewiesen werden.

□

Beispiel 2.3 *Sei wie in Beispiel 2.1*

$$f_k(x) = \begin{cases} 1 - kx & 0 \le x \le \frac{1}{k} \\ 0 & \textit{sonst,} \end{cases}$$

und sei $f(x) := 0$ für alle $x \in I$. Dann gilt

$$\|f_k - f\|_1 = \int_0^{\frac{1}{k}} \underbrace{|f_k|}_{\le 1} \le \frac{1}{k} \to 0.$$

Konvergenz bzgl. $\|\cdot\|_I$ impliziert Konvergenz bzgl. $\|\cdot\|_1$, siehe Satz 1.4 in Kapitel 4. Aber nicht umgekehrt, wie das Beispiel zeigt.

Beispiel 2.4 *Der Raum $(C^0(I), \|\cdot\|_1)$ ist kein Banachraum. Sei z. B. $I = [-1, 1]$ und*

$$f_k(x) = \begin{cases} -1 & \textit{für } -1 \le x \le \frac{1}{k} \\ 1 & \textit{für } \frac{1}{k} \le x \le 1, \\ kx & \textit{sonst.} \end{cases}$$

$$f(x) = \begin{cases} -1 & \textit{für } x < 0 \\ 1 & \textit{für } x \ge 0. \end{cases}$$

Es gilt $f \in \mathcal{R}(I)$ sowie

$$\int_I |f - f_k| = \int_{-1/k}^{1/k} \underbrace{|f - f_k|}_{\le 1/k} \le \frac{2}{k} \to 0.$$

Es folgt für $k, l > 4/\varepsilon$

$$\|f_k - f_l\|_1 \le \|f_k - f\|_1 + \|f - f_l\|_1 \le \frac{2}{k} + \frac{2}{l} < \varepsilon,$$

d. h. (f_k) ist $\|\cdot\|_1$ – Cauchyfolge. Angenommen, es gibt $\widetilde{f} \in C^0(I)$ mit $f_k \to \widetilde{f}$ bzgl. $\|\cdot\|_1$. Dann folgt

$$\|f - \widetilde{f}\|_1 \le \|f - f_k\|_1 + \|f_k - \widetilde{f}\|_1 \to 0.$$

Dies ist nur möglich, wenn $f(x) = \widetilde{f}(x)$ für alle $x \neq 0$. Widerspruch, da f im Nullpunkt nicht stetig ergänzt werden kann.

Auch Satz 2.2 und Satz 2.3 gelten analog für Funktionen mit Werten in $\mathbb{R}^m$ oder in $\mathbb{C}$.

Wir kommen jetzt zu den Fourierreihen, d. h. zu Reihen der Form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \quad \text{mit } a_n, b_n \in \mathbb{C}.$$

Ist x ein Zeitparameter, so entspricht jeder Summand einer reinen Schwingung mit Periode $2\pi/n$. Die Partialsummen $a_0/2 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx)$ mit $N \in \mathbb{N}_0$ heißen trigonometrische Polynome und können als (endliche) Überlagerungen dieser reinen Schwingungen betrachtet werden.

Fourier 1822: "Jede willkürlich gegebene, 2π-periodische Funktion lässt sich durch eine trigonometrische Reihe ausdrücken."

Definition 2.1 *Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ heißt 2π-periodisch, falls gilt:*

$$f(x + 2\pi) = f(x) \quad \textit{für alle } x \in \mathbb{R}.$$

Jede 2π-periodische Funktion entspricht einer Funktion auf dem Einheitskreis:

Der Begriff „willkürlich“in der Behauptung von Fourier ist historisch zu verstehen. Eine Funktion im damaligen Sinn ist durch einen analytischen Ausdruck definiert. Demnach bezeichnet Fourier Funktionen, deren Definition zum Beispiel eine Fallunterscheidung im Definitionsbereich erfordert – also bzgl. der x-Variablen – als willkürlich. Wir werden das Problem der Entwicklung in eine trigonometrische Reihe, wie oben von Fourier formuliert, in der folgenden Klasse von Funktionen behandeln:

$$X = \{f : \mathbb{R} \to \mathbb{C} : f \text{ ist } 2\pi - \text{periodisch und stückweise stetig}\}. \tag{2.8}$$

Das Problem weist eine enge Analogie zu der geometrischen/algebraischen Aufgabe auf, einen symmetrischen (selbstadjungierten) Endomorphismus eines unitären oder Euklidischen Vektorraums zu diagonalisieren. Dies wollen wir nun herausarbeiten.

Lemma 2.3 *Seien $f, g, h \in X$ und*

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f\overline{g}. \tag{2.9}$$

Dann gelten folgende Aussagen:

(a) $\langle f, g \rangle = \overline{\langle g, f \rangle}$,

(b) $\langle \lambda f + \mu g, h \rangle = \lambda \langle f, h \rangle + \mu \langle g, h \rangle \quad \forall \lambda, \mu \in \mathbb{C}$,

(c) $\langle f, f \rangle \geq 0$.

Gleichheit in (c) gilt genau dann, wenn die Menge $\{x \in [-\pi, \pi] : f(x) \neq 0\}$ endlich ist.

Beweis: Die Aussagen (a), (b) und (c) folgen aus der Linearität und Monotonie des Integrals. Die Diskussion der Gleichheit in (c) ergibt sich aus Satz 2.3. $\square$

Man überprüft leicht, dass die Ungleichung von Cauchy-Schwarz und die Dreiecksungleichung für die zugehörige Seminorm

$$\|f\|_2 = \left(\int_{-\pi}^{\pi} f\overline{f} \right)^{1/2} \tag{2.10}$$

gelten. In den meisten der folgenden Überlegungen spielt es keine Rolle, dass $\| \cdot \|_2$ nicht strikt positiv definit ist. Wir interessieren uns nun für die zweite Ableitung, die wir als eine lineare Abbildung

$$\Delta : \widetilde{X} \longrightarrow \widetilde{X}, \quad \Delta f = f'', \tag{2.11}$$

auffassen können, wobei $\widetilde{X} = X \cap C^\infty(\mathbb{R}, \mathbb{C})$ der Unterraum der glatten, 2π-periodischen Funktionen ist. Die zweite Ableitung ist zwar schon für $f \in C^2(\mathbb{R}, \mathbb{C})$ definiert, aber dieser Raum wird durch $f \mapsto f''$ nicht in sich abgebildet, sondern nur in $C^0(\mathbb{R}, \mathbb{C})$. Durch Einschränkung auf die C^∞-Funktionen erhalten wir dagegen den Endomorphismus Δ.

Lemma 2.4 *Für $f, g \in \widetilde{X}$ gilt*

$$\langle \Delta f, g\rangle = \langle f, \Delta g\rangle \quad (\Delta \textit{ symmetrisch}), \tag{2.12}$$

$$\langle \Delta f, f\rangle \leq 0 \quad (\Delta \textit{ negativ semidefinit}). \tag{2.13}$$

Beweis: Wir berechnen mit partieller Integration

$$\begin{aligned}
\langle \Delta f, g\rangle &= \int_{-\pi}^{\pi} f''\,\overline{g} \\
&= [f'\,\overline{g}]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'\overline{g}' \\
&= [f'\,\overline{g} - f\,\overline{g}']_{-\pi}^{\pi} + \int_{-\pi}^{\pi} f\,\overline{g}'' \\
&= [f'\,\overline{g} - f\,\overline{g}']_{-\pi}^{\pi} + \langle f, \Delta g\rangle.
\end{aligned}$$

Da die beteiligten Funktionen und ihre Ableitungen 2π-periodisch sind, heben sich die Randwerte bei $x = \pi$ und bei $x = -\pi$ gegenseitig weg, und (2.12) folgt. Außerdem ergibt sich mit $g := f$

$$\langle \Delta f, f\rangle = \underbrace{[f'\,\overline{f}]_{-\pi}^{\pi}}_{=0} - \int_{-\pi}^{\pi} |f'|^2 \leq 0.$$

$\square$

Es folgen nun Konsequenzen für die Eigenwerte und zugehörigen Eigenvektoren – üblicherweise spricht man von Eigenfunktionen – der linearen Abbildung Δ.

Lemma 2.5 *Die Eigenwerte von Δ sind genau die Zahlen $\{-n^2 : n \in \mathbb{N}_0\}$. Die zugehörigen Eigenräume besitzen die Basis $\{u_n, v_n\}$ für $n \geq 1$ und $\{u_0\}$ für $n = 0$, wobei*

$$\begin{aligned}
u_n(x) &= \frac{1}{\sqrt{\pi}}\cos nx, \quad v_n(x) = \frac{1}{\sqrt{\pi}}\sin nx \quad (n \in \mathbb{N}), \\
u_0(x) &= \frac{1}{\sqrt{2\pi}}.
\end{aligned}$$

In der gewählten Normierung bilden $\{u_0, u_1, v_1, u_2, v_2, \dots\}$ ein Orthonormalsystem, das heißt es gilt

$$\langle u_n, u_m\rangle = \langle v_n, v_m\rangle = \delta_{nm},\ \langle u_n, v_m\rangle = 0.$$

Beweis: Sei $\Delta f = \lambda f$. Aus Lemma 2.4 folgt

$$\lambda \underbrace{\|f\|_2^2}_{\neq 0} = \langle \lambda f, f\rangle = \langle \Delta f, f\rangle \leq 0,$$

also $\lambda \in (-\infty, 0]$. Wir setzen $\omega := \sqrt{-\lambda} \geq 0 \Rightarrow \lambda = -\omega^2$. Dann gilt $f'' + \omega^2 f = 0$ und nach Kapitel 3.3 folgt $f(x) = a\cos\omega x + b\sin\omega x$ für geeignete $a, b \in \mathbb{C}$. Da f 2π-periodisch ist, muss $\omega = n \in \mathbb{N}_0$ sein und somit gilt $\lambda = -n^2$.

Eigenfunktionen zu verschiedenen Eigenwerten λ_1, λ_2 sind orthogonal, denn

$$\begin{aligned}
\underbrace{(\lambda_1 - \lambda_2)}_{\neq 0}\langle f_1, f_2\rangle &= \langle \lambda_1 f_1, f_2\rangle - \langle f_1, \lambda_2 f_2\rangle \quad (\text{da } \lambda_2 \in \mathbb{R}) \\
&= \langle \Delta f_1, f_2\rangle - \langle f_1, \Delta f_2\rangle \\
&= 0 \quad (\text{nach } (2.12)).
\end{aligned}$$

Schließlich ist leicht nachzurechnen, dass $\langle u_n, v_n\rangle = 0$ und $\|u_n\|_2 = \|v_n\|_2 = 1$. □

Alternativ kann folgendes ON-System gewählt werden, das sich durch einen Basiswechsel in den zweidimensionalen Eugenräumen, die von u_n, v_n aufgespannt werden, ergibt:

(2.14)
$$\begin{aligned}
w_n &= \frac{1}{\sqrt{2}}(u_n + i\, v_n) = \frac{1}{\sqrt{2\pi}}\, e^{inx} \quad (n \in \mathbb{N}), \\
w_{-n} &= \frac{1}{\sqrt{2}}(u_n - i\, v_n) = \frac{1}{\sqrt{2\pi}}\, e^{-inx} \quad (n \in \mathbb{N}), \\
w_0 &= u_0 = \frac{1}{\sqrt{2\pi}}.
\end{aligned}$$

Man rechnet leicht nach, dass $\langle w_k, w_l\rangle = \delta_{kl}$ für $k, l \in \mathbb{Z}$. Der Vorteil dieses ON-Sytems ist, dass die Funktionen schon Eigenfunktionen der ersten Ableitung

$$D : \widetilde{X} \to \widetilde{X},\ Df = f'$$

sind, und zwar gilt $Dw_k = ikw_k$. Dagegen sind die $\{u_n, v_n\}$ nur Eigenfunktionen von $D^2 = \Delta$. Das System der $\{u_n, v_n\}$ hat andererseits den Vorteil, dass die Funktionen eine Parität besitzen: u_n ist gerade, $u_n(-x) = u_n(x)$, und v_n ist ungerade, $v_n(-x) = -v_n(x)$.

Sei V der von den Eigenfunktionen $\{u_0, u_1, v_1, \dots\}$ erzeugte Raum, also der Raum der trigonometrischen Polynome. Dann ist $V \subset \widetilde{X}$ ein echter Unterraum von $\widetilde{X}$ (Übungsaufgabe 5, Serie 3), das heißt die Eigenfunktionen von Δ bilden keine Basis von $\widetilde{X}$. Wir können aber hoffen, dass sich Funktionen in $\widetilde{X}$ oder sogar in dem größeren Raum X durch trigonometrische Polynome *approximieren* lassen. Die einleitende Frage von Fourier lautet also nun präziser:

> Läßt sich jede 2π-periodische, stückweise stetige Funktion f durch eine Folge von trigonometrischen Polynomen approximieren, und zwar bzgl. $\|\cdot\|_2$? Mit anderen Worten: ist V dicht in X bzgl. $\|\cdot\|_2$, das heißt gilt $\overline{V} = X$?

Zur Approximation einer gegebenen Funktion f ist eine geeignete Folge von trigonometrischen Polynomen zu konstruieren. Hierzu gibt es einen natürlichen Ansatz.

Definition 2.2 *Die Fourierkoeffizienten einer 2π-periodischen, stückweise stetigen Funktion f sind die Zahlen*

(2.15)
$$\begin{aligned}
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\, dx \quad (n \in \mathbb{N}_0), \\
b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\, dx \quad (n \in \mathbb{N}).
\end{aligned}$$

Die Reihe $a_0/2 + \sum_{n=1}(a_n \cos nx + b_n \sin nx)$ mit a_n, b_n aus (2.15) heißt Fourierreihe von f. Die Partialsumme $P_N f = \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx)$ heißt N-tes Fourierpolynom von f.

Beispiel 2.5 *Wir betrachten als Rechenbeispiel die 2π-periodische Funktion f mit*

$$f(x) = \begin{cases} 1 & \textit{für } 0 < x < \pi \\ -1 & \textit{für } -\pi < x < 0. \\ 0 & \textit{für } x = -\pi, 0, \pi. \end{cases}$$

Die Funktion ist ungerade, deshalb gilt $a_n = 0$ für $n \in \mathbb{N}_0$. Wir berechnen für $n \in \mathbb{N}$

$$\begin{aligned} b_n &= \frac{1}{\pi}\left(\int_{-\pi}^{0}(-1)\sin nx\,dx + \int_0^{\pi} 1\,\sin nx\,dx\right) \\ &= \frac{1}{\pi}\left(\left[\frac{1}{n}\cos nx\right]_{-\pi}^{0} + \left[-\frac{1}{n}\cos nx\right]_0^{\pi}\right) \\ &= \frac{2}{n\pi}\,(1-(-1)^n). \end{aligned}$$

Die Fourierreihe der betrachteten Rechteckfunktion lautet also

$$f \sim \frac{4}{\pi}\sum_{j=0}^{\infty}\frac{\sin(2j+1)x}{2j+1}.$$

Dabei bedeutet $\sim$ nur, dass f die berechneten Fourierkoeffizienten hat; von der Konvergenz der Reihe ist noch gar nicht die Rede. Die berechnete Reihe ist jedenfalls nicht absolut konvergent, denn es gilt $\sum_{j=0}^{\infty} 1/(2j+1) = +\infty$. Sie kann auch nicht gleichmäßig konvergieren, denn sonst müßte die Grenzfunktion stetig sein nach Satz 2.2 in Kapitel 5; das ist aber nicht der Fall.

Oft ist es einfacher, mit der Basis aus (2.14) zu rechnen. Wir definieren die zugehörigen Fourierkoeffizienten durch

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,e^{-ikx}\,dx \quad (k \in \mathbb{Z}). \tag{2.16}$$

Aus der Eulerschen Formel $e^{ix} = \cos x + i\sin x$ ergibt sich folgende Umrechnungstabelle:

$$\begin{aligned} a_n &= c_n + c_{-n}, \quad b_n = i(c_n - c_{-n}), \\ c_n &= \frac{1}{2}(a_n - ib_n), \quad c_{-n} = \frac{1}{2}(a_n + ib_n), \quad c_0 = \frac{a_0}{2}, \\ |c_n|^2 + |c_{-n}|^2 &= \frac{1}{2}\,(|a_n|^2 + |b_n|^2), \quad 2\,|c_0|^2 = \frac{1}{2}\,|a_0|^2. \end{aligned} \tag{2.17}$$

Daraus folgt

$$\begin{aligned} P_N f &= c_0 + \sum_{n=1}^{N}\left((c_n + c_{-n})\frac{e^{inx}+e^{-inx}}{2} + i(c_n - c_{-n})\frac{e^{inx}-e^{-inx}}{2i}\right) \\ &= c_0 + \sum_{n=1}^{N}(c_n e^{inx} + c_{-n}e^{-inx}), \qquad \text{bzw.} \end{aligned}$$

$$P_N f = \sum_{k=-N}^{N} c_k e^{ikx} \qquad \text{mit } c_k \text{ aus (2.16)}. \tag{2.18}$$

Welche Bewandtnis hat nun die Definition der Fourierkoeffizienten, d. h. warum ist gerade die Fourierreihe der richtige Ansatz zur Approximation? Betrachte dazu die aufsteigende Sequenz

$$V = \bigcup_{N=0}^{\infty} V_N \text{ mit } V_0 \subset V_1 \subset V_2 \subset \dots, \text{ wobei } V_N = \{ \sum_{k=-N}^{N} c_k e^{ikx} : c_k \in \mathbb{C} \}. \tag{2.19}$$

Lemma 2.6 (Orthogonalitätsrelationen) *Sei f 2π-periodisch und stückweise stetig mit Fourierpolynom $P_N f = \sum_{k=-N}^{N} c_k\, e^{ikx}$, vgl. (2.18). Dann gelten folgende Aussagen:*

(1) *$P_N f$ ist die Orthogonalprojektion von f auf V_N: $f - P_N f \perp V_N$ und $P_N f$ ist der einzige Punkt $q \in V_N$ mit $f - q \perp V_N$.*

(2) $\|P_N f\|_2^2 = 2\pi \sum_{k=-N}^{N} |c_k|^2$.

(3) $\|f - P_N f\|_2^2 = \|f\|_2^2 - \|P_N f\|_2^2$.

(4) $\|f - P_N f\|_2 \leq \|f - q\|_2$ *für alle $q \in V_N$, mit Gleichheit nur für $q = P_N f$.*

Beweis: Sei $\{w_k : k \in \mathbb{Z}\}$ das Orthonormalsystem aus (2.14). Wir berechnen mit (2.16)

$$\langle f, w_k \rangle w_k = \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} f(x)\, e^{-ikx}\, dx \right) e^{ikx} = c_k e^{ikx}.$$

Nach Definition des Fourierpolynoms folgt

$$P_N f = \sum_{k=-N}^{N} \langle f, w_k \rangle w_k, \tag{2.20}$$

Hieraus ergibt sich (1), denn für jedes $l \in \{-N, \dots, N\}$ gilt

$$\langle f - P_N f, w_l \rangle = \langle f, w_l \rangle - \sum_{k=-N}^{N} \langle f, w_k \rangle \underbrace{\langle w_k, w_l \rangle}_{=\delta_{kl}} = 0.$$

Für einen beliebigen Punkt $q \in V_N$ mit $f - q \perp V_N$ ist $P_N f - q = (P_N f - f) + (f - q) \perp V_N$, und andererseits $P_N f - q \in V_N$, also $P_N f - q = 0$. Als nächstes berechnen wir mit (2.20) und (2.14)

$$\begin{aligned}
\|P_N f\|^2 &= \left\langle \sum_{k=-N}^{N} \langle f, w_k \rangle w_k, \sum_{k=-N}^{N} \langle f, w_l \rangle w_l \right\rangle \\
&= \sum_{k,l=-N}^{N} \langle f, w_k \rangle \overline{\langle f, w_l \rangle} \underbrace{\langle w_k, w_l \rangle}_{=\delta_{kl}} \\
&= \sum_{k=-N}^{N} |\langle f, w_k \rangle|^2 \\
&= 2\pi \sum_{k=-N}^{N} |c_k|^2.
\end{aligned}$$

Nun ist $\langle f - P_N f, P_N f - q\rangle = 0$ für $q \in V_N$ nach (1) und somit

$$\|f - q\|_2^2 = \|f - P_N f + P_N f - q\|_2^2 = \|f - P_N f\|_2^2 + \|P_N f - q\|_2^2.$$

Hieraus folgt (4), und mit $q = 0$ ergibt sich auch (3). □

Satz 2.4 (Besselsche Ungleichung) *Für die Fourierkoeffizienten c_k einer 2π-periodischen, stückweise stetigen Funktion f gilt*

$$2\pi \sum_{k=-\infty}^{\infty} |c_k|^2 \leq \|f\|_2^2,$$

mit Gleichheit genau dann, wenn $\lim_{N\to\infty} \|f - P_N f\|_2 = 0$.

Beweis: Aus (1) und (2) in Lemma 2.6 folgt mit $N \to \infty$

$$\|f\|_2^2 - 2\pi \sum_{k=-\infty}^{\infty} |c_k|^2 = \lim_{N\to\infty} \|f - P_N f\|_2^2 \geq 0.$$

□

Proposition 2.1 *Sei $\overline{V}$ der Abschluss von V im Raum X der 2π-periodischen, stückweise stetigen Funktionen bzgl. $\|\cdot\|_2$, das heißt*

$$\overline{V} = \{f \in X : \exists\, q_k \in V \text{ mit } \|f - q_k\|_2 \to 0\}. \tag{2.21}$$

Dann ist $\overline{V}$ abgeschlossen bzgl. $\|\cdot\|_2$ und es gilt

$$f \in \overline{V} \quad \Leftrightarrow \quad \lim_{N\to\infty} \|f - P_N f\|_2 = 0.$$

Beweis: Die Abgeschlossenheit von $\overline{V}$ gilt nach Satz 1.7. Sei nun $f \in \overline{V}$, also $\|f - q_k\|_2 \to 0$ für eine Folge $q_k \in V$. Dann ist $q_k \in V_{N_k}$ für ein $N_k \in \mathbb{N}_0$. Nach (2) und (3) in Lemma 2.6 ist $\|f - P_N f\|_2$ monoton fallend, und Aussage (4) liefert

$$\lim_{N\to\infty} \|f - P_N f\|_2 \leq \|f - P_{N_k} f\|_2 \leq \|f - q_k\|_2 \to 0 \quad \text{mit } k \to \infty.$$

Die umgekehrte Implikation ist trivial. □

Satz 2.5 (Entwicklungssatz: Vollständigkeit der Eigenfunktionen) *Ist f stückweise stetig und 2π-periodisch, so konvergiert die Fourierreihe von f im quadratischen Mittel gegen f, das heißt es gilt $\|f - P_N f\|_2 \to 0$ mit $N \to \infty$.*

Mit anderen Worten: V ist dicht in X bzgl. $\|\cdot\|_2$, das heißt $\overline{V} = X$. Wir stellen den Beweis von Satz 2.5 noch etwas zurück.

Folgerung 2.1 (Parsevalsche Gleichung) *Seien $f, g : \mathbb{R} \to \mathbb{C}$ stückweise stetig und 2π-periodisch, mit Fourierkoeffizienten $c_k, d_k \in \mathbb{C}$ für $k \in \mathbb{Z}$. Dann gilt:*

(1) $\|f\|_2^2 = 2\pi \sum_{k=-\infty}^{\infty} |c_k|^2.$

(2) $\langle f, g\rangle = 2\pi \sum_{k=-\infty}^{\infty} c_k \overline{d_k}$.

Beweis: Aussage (1) ergibt sich aus Satz 2.5 und Satz 2.4. Mit Cauchy-Schwarz schließen wir weiter

$$\begin{aligned} |\langle P_N f, P_N g\rangle - \langle f, g\rangle| &= |\langle P_N f - f, P_N g\rangle + \langle f, P_N g - g\rangle| \\ &\leq \underbrace{\|P_N f - f\|_2}_{\to 0} \underbrace{\|P_N g\|_2}_{\leq \|g\|_2} + \|f\|_2 \underbrace{\|P_N g - g\|_2}_{\to 0} \\ &\to 0 \quad \text{mit } N \to \infty. \end{aligned}$$

Da $\langle e^{ikx}, e^{ilx}\rangle = 2\pi\langle w_k, w_l\rangle = 2\pi\delta_{kl}$, folgt

$$2\pi \sum_{k=-N}^{N} c_k \overline{d_k} = \left\langle \sum_{k=-N}^{N} c_k e^{ikx}, \sum_{l=-N}^{N} c_l e^{ilx} \right\rangle = \langle P_N f, P_N g\rangle \to \langle f, g\rangle.$$

□

Beispiel 2.6 *Sei $f(x) = x$ für $-\pi \leq x < \pi$. Die Funktion ist ungerade (wenn wir $f(\pi) = f(\pi) = 0$ setzen, was die Fourierkoeffizienten nicht berührt), folglich ist $a_n = 0$ für alle $n \in \mathbb{N}_0$. Mit partieller Integration erhalten wir für $n \in \mathbb{N}$*

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin x \, dx \\ &= \frac{1}{\pi} \left[-\frac{x \cos nx}{n} \right]_{-\pi}^{\pi} + \frac{1}{\pi} \underbrace{\int_{-\pi}^{\pi} \frac{\cos nx}{n} \, dx}_{=0} \\ &= \frac{2}{n}(-1)^{n+1}. \end{aligned}$$

Die Fourierreihe ist demnach $2\left(\sin x + \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x + \dots\right)$. Mit den Umrechnungsformeln (2.17) folgt hier

$$2\pi \sum_{k=-\infty}^{\infty} |c_k|^2 = \pi \left(\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) \right) = 4\pi \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Andererseits gilt

$$\|f\|_2^2 = \int_{-\pi}^{\pi} x^2 \, dx = \frac{2}{3}\pi^3.$$

Die Funktion f ist stückweise stetig. Somit ist die Parsevalsche Gleichung (Folgerung 2.1) anwendbar und es ergibt sich die hübsche Formel

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Der Beweis des Satzes 2.5 erfordert etwas Vorarbeit. Wir beginnen mit der Diskussion der *punktweisen Konvergenz* der Fourierreihe, welche auch für sich betrachtet interessant ist. Dazu schreiben wir das Fourierpolynom an der Stelle $x = 0$ wie folgt:

$$P_N f(0) = \sum_{k=-N}^{N} c_k = \int_{-\pi}^{\pi} f(x) \underbrace{\frac{1}{2\pi} \sum_{k=-N}^{N} e^{-ikx}}_{=:D_N(x)} dx. \tag{2.22}$$

Lemma 2.7 *Für die Funktion* $D_N(x) = \dfrac{1}{2\pi} \displaystyle\sum_{k=-N}^{N} e^{-ikx}$ *gilt:*

(1) $\int_{-\pi}^{\pi} D_N(x)\, dx = 1.$

(2) $D_N(x) = \dfrac{1}{2\pi}(\cos Nx + \cot x/2 \, \sin Nx)$ *für* $x \notin 2\pi\mathbb{Z}$.

Beweis: Für die konstante Funktion $f(x) \equiv 1$ gilt $P_N f = f$ wegen $f \in V_N$. Aus (2.22) folgt dann

$$1 = f(0) = P_N f(0) = \int_{-\pi}^{\pi} D_N(x)\, dx.$$

Für (2) verwenden wir die geometrische Summenformel und im letzten Schritt das Additionstheorem der Sinusfunktion mit $\alpha = Nx$ und $\beta = x/2$:

$$\begin{aligned}
\sum_{k=-N}^{N} e^{-ikx} &= e^{-iNx} \sum_{k=0}^{2N} e^{ikx} \\
&= e^{-iNx} \frac{e^{i(2N+1)x} - 1}{e^{ix} - 1} \\
&= \frac{e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}}{e^{ix/2} - e^{-ix/2}} \\
&= \frac{\sin(N+\frac{1}{2})x}{\sin\frac{x}{2}} \\
&= \cos Nx + \cot\frac{x}{2} \sin Nx.
\end{aligned}$$

□

Nach Definition existieren für Funktionen $f \in X$, also stückweise stetige Funktionen, die rechts- und linksseitigen Grenzwerte

$$f_+(x_0) = \lim_{x \searrow x_0} f(x), \quad f_-(x_0) = \lim_{x \nearrow x_0} f(x).$$

Definition 2.3 *Die stückweise stetige Funktion* $f : \mathbb{R} \to \mathbb{C}$ *heißt an der Stelle* $x_0 \in \mathbb{R}$ *einseitig differenzierbar, wenn die folgenden links- und rechtsseitigen Ableitungen existieren:*

$$f'_+(x_0) = \lim_{x \searrow x_0} \frac{f(x) - f_+(x_0)}{x - x_0}, \quad f'_-(x_0) = \lim_{x \nearrow x_0} \frac{f(x) - f_-(x_0)}{x - x_0}.$$

Satz 2.6 (Dirichlet) *Sei $P_N f$ das Fourierpolynom der 2π-periodischen, stückweise stetigen Funktion $f : \mathbb{R} \to \mathbb{C}$. Falls f an der Stelle $x_0 \in \mathbb{R}$ einseitig differenzierbar ist, so gilt*

$$\lim_{N\to\infty} P_N f(x_0) = \frac{f_+(x_0) + f_-(x_0)}{2}.$$

Beweis: Wir behandeln zunächst den Fall $x_0 = 0$. Jede Funktion kann als Summe einer geraden und einer ungeraden Funktion geschrieben werden, und zwar gilt

$$f(x) = \underbrace{\frac{f(x) + f(-x)}{2}}_{\text{gerade}} + \underbrace{\frac{f(x) - f(-x)}{2}}_{\text{ungerade}}.$$

Ist f ungerade, so gilt $a_n = 0$ für alle $n \in \mathbb{N}_0$, also ist das Fourierpolynom auch eine ungerade Funktion. Damit schließen wir

$$P_N f(0) = 0 = \frac{f_+(0) + f_-(0)}{2} \quad \text{für alle } N \in \mathbb{N}_0,$$

und die Behauptung ist an der Stelle $x_0 = 0$ bewiesen.

Sei nun f gerade. Dann existiert sogar der Grenzwert $\lim_{x\to 0} f(x)$ und wir können nach evtl. Abänderung des Funktionswerts bei $x_0 = 0$ annehmen, dass $f(0) = \lim_{x\to 0} f(x)$; beachten Sie, dass der Wert $f(0)$ in die Behauptung des Satzes nicht eingeht. Mit Lemma 2.7 berechnen wir

$$\begin{aligned} P_N f(0) - f(0) &= \int_{-\pi}^{\pi} (f(x) - f(0))\, D_N(x)\, dx \\ &=: \underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi} g(x) \cos Nx\, dx}_{=:A_N} + \underbrace{\frac{1}{\pi}\int_{-\pi}^{\pi} h(x) \sin Nx\, dx}_{=:B_N}, \end{aligned}$$

wobei

$$\begin{aligned} g(x) : &= \frac{1}{2}\,(f(x) - f(0)) \\ h(x) : &= \begin{cases} \frac{1}{2}\,(f(x) - f(0)) \cot \frac{x}{2} & x \notin 2\pi\mathbb{Z} \\ 0 & x \in 2\pi\mathbb{Z}. \end{cases} \end{aligned}$$

Die Funktionen g und h sind stückweise stetig und 2π-periodisch. Für g ist das klar, für h folgt es aus

$$\frac{1}{2}(f(x) - f(0)) \cot \frac{x}{2} = \frac{f(x) - f(0)}{x}\,\frac{x/2\, \cos x/2}{\sin x/2} \quad \to f'_\pm(0) \text{ für } x \searrow 0 \text{ bzw. } x \nearrow 0.$$

Damit sind die Zahlen A_N, B_N Fourierkoeffizienten der stückweise stetigen Funktionen $g \in X$ bzw. $h \in X$. Mit der Besselschen Ungleichung, Satz 2.4, folgt nun

$$\sum_{N=0}^{\infty} |A_N|^2 < \infty, \quad \text{sowie} \quad \sum_{N=0}^{\infty} |B_N|^2 < \infty,$$

und das Nullfolgenkriterium liefert $\lim_{N\to\infty} A_N = \lim_{N\to\infty} B_N = 0$. Damit ist auch für gerade Funktionen f die Aussage an der Stelle $x_0 = 0$ bewiesen.

Sei schließlich $x_0 \in \mathbb{R}$ beliebig. Die Funktion $\widetilde{f}(x) = f(x_0 + x)$ hat die Fourierkoeffizienten

$$\begin{aligned}
\widetilde{c_k} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+x_0)e^{-ikx}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi+x_0}^{\pi+x_0} f(t)e^{-ik(t-x_0)}\,dt \quad \text{(Substitution } x_0 + x = t\text{)} \\
&= \frac{e^{ikx_0}}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\,dt \quad \text{(Linearität des Integrals, Periodizität)} \\
&= e^{ikx_0}c_k \quad (c_k = \text{Fourierkoeffizient von } f).
\end{aligned}$$

Es folgt

$$P_N f(x_0) = \sum_{k=-N}^{N} c_k e^{ikx_0} = \sum_{k=-N}^{N} \widetilde{c_k} = P_N\widetilde{f}(0).$$

Damit erhalten wir die Behauptung auch an der Stelle x_0, und der Satz ist bewiesen. □

Beispiel 2.7 *Wir berechnen die Fourierreihe der Sägezahnfunktion $f(x) = |x|$ für $|x| \leq \pi$. Da f gerade ist, folgt $b_n = 0$ für alle $n \in \mathbb{N}$. Weiter gilt $a_0 = \frac{2}{\pi}\int_0^{\pi} |x|\,dx = \pi$, und für $n \geq 1$*

$$\begin{aligned}
a_n &= \frac{2}{\pi}\int_0^{\pi} x\,\cos nx\,dx \\
&= \frac{2}{\pi}\left(\underbrace{\left[\frac{x\,\sin x}{n}\right]_0^{\pi}}_{=0} - \frac{1}{n}\int_0^{\pi}\sin x\,dx\right) \\
&= \frac{2}{n^2\pi}\left[\cos nx\right]_0^{\pi} \\
&= \frac{2}{n^2\pi}((-1)^n - 1).
\end{aligned}$$

Die Sägezahnfunktion ist überall einseitig differenzierbar und stetig. Der Satz von Dirichlet ergibt somit für alle $x \in [-\pi, \pi]$ die Reihendarstellung

$$|x| = \frac{\pi}{2} - \frac{4}{\pi}\sum_{j=0}^{\infty}\frac{\cos(2j+1)x}{(2j+1)^2}.$$

Zum Beispiel folgt an der Stelle $x = 0$

$$\frac{\pi^2}{8} = \sum_{j=0}^{\infty}\frac{1}{(2j+1)^2} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots .$$

Damit die Fourierreihe von f nicht nur punktweise, sondern gleichmäßig konvergiert, sind an die Regularität der Funktion f zusätzliche Bedingungen zu stellen, vgl. auch Beispiel 2.5.

Satz 2.7 *Die 2π-periodische Funktion $f : \mathbb{R} \to \mathbb{C}$ sei stetig und stückweise C^1, d. h. es gibt eine Unterteilung $-\pi = x_0 < x_1 < \ldots < x_J = \pi$, so dass $f|_{[x_{j-1},x_j]}$ stetig differenzierbar ist. Dann konvergiert die Fourierreihe von f gleichmäßig gegen f.*

Beweis: Es reicht aus zu zeigen, *dass* die Fourierreihe gleichmäßig konvergiert. Als Grenzfunktion kommt dann nämlich nur f in Frage, denn für jedes $x \in \mathbb{R}$ gilt $\lim_{N\to\infty} P_N f(x) = f(x)$ nach Satz 2.6. Nach dem Majorantenkriterium (Übungsaufgabe 2, Serie 2) folgt die gleichmäßige Konvergenz einer Funktionenreihe $\sum_{k=-\infty}^{\infty} f_k$ von 2π-periodischen Funktionen aus der Bedingung $\sum_{k=-\infty}^{\infty} \|f_k\|_{[-\pi,\pi]} < \infty$. Im Fall der Fourierreihe ist $f_k(x) := c_k e^{ikx}$ und folglich $\|f_k\|_{[-\pi,\pi]} = |c_k|$. Es ist also nur zu zeigen, dass $\sum_{k=-\infty}^{\infty} |c_k| < \infty$.

Sei g die 2π-periodische, stückweise stetige Funktion mit

$$g(x) = \begin{cases} f'(x) & \text{falls } x \in (x_{j-1}, x_j) \text{ für } j = 1, \ldots, J \\ 0 & \text{sonst.} \end{cases}$$

Um die Fourierkoeffizienten von f abzuschätzen, integrieren wir auf den Teilintervallen partiell und erhalten

$$\begin{aligned} c_k &= \frac{1}{2\pi} \sum_{j=1}^{J} \int_{x_{j-1}}^{x_j} f(x) e^{-ikx}\, dx \\ &= \underbrace{\frac{1}{2\pi} \sum_{j=1}^{J} \left[f(x) \frac{i}{k} e^{-ikx} \right]_{x_j-1}^{x_j}}_{=0} - \frac{1}{2\pi} \sum_{j=1}^{J} \int_{x_j-1}^{x_j} f'(x) \frac{i}{k} e^{-ikx}\, dx \\ &= -\frac{i}{2\pi k} \int_{-\pi}^{\pi} g(x)\, dx \\ &=: -\frac{i}{k}\, d_k, \end{aligned}$$

wobei d_k die Fourierkoeffizienten von $g \in X$ bezeichnet. Es folgt für $k \neq 0$

$$|c_k| = \frac{1}{|k|}\, |d_k| \leq \frac{1}{2} \left(\frac{1}{k^2} + |d_k|^2 \right).$$

Die linke Reihe $\sum_{k\neq 0} 1/k^2$ konvergiert nach einer Anwendung des Integralvergleichskriteriums, siehe Kapitel 5, Beispiel 1.5. Die rechte Reihe $\sum_{k=-\infty}^{\infty} |d_k|^2$ konvergiert wegen der Besselschen Ungleichung, Satz 2.4. Damit ist $\sum_{k=-\infty}^{\infty} |c_k| < \infty$, was zu zeigen war. □

Zum Beweis des Entwicklungssatzes fehlt jetzt nur noch das folgende Approximationsresultat:

Lemma 2.8 *Jede 2π-periodische, stückweise stetige Funktion f kann bzgl. $\|\cdot\|_2$ durch stetige, stückweise lineare, 2π-periodische Funktionen g_k approximiert werden, d. h. $\|f - g_k\|_2 \to 0$ mit $k \to \infty$.*

Beweis: Sei zunächst f eine Treppenfunktion, d. h. es gibt eine Unterteilung $-\pi = x_0 < \ldots < x_J = \pi$ mit $f|_{(x_{j-1},x_j)} \equiv c_j \in \mathbb{C}$ für $j = 1, \ldots, J$. Sei $\delta > 0$. Wir ersetzen die Funktion f auf der Umgebung $(x_j - \delta, x_j + \delta)$ jeder Sprungstelle durch die lineare Funktion, die zwischen den Werten von f an den Stellen $x_j \pm \delta$ interpoliert. Dazu definieren wir die linearen Hilfsfunktionen

$$\lambda_j : [x_j - \delta, x_j + \delta] \to \mathbb{R},\ \lambda_j(x) = \frac{x - (x_j - \delta)}{2\delta}.$$

Es gilt $\lambda_j(x_j - \delta) = 0$, $\lambda_j(x_j + \delta) = 1$ und $0 \leq \lambda_j(x) \leq 1$ für $x \in (x_j - \delta, x_j + \delta)$. Nun setzen wir für $x \in [-\pi, \pi]$

$$f_\delta(x) = \begin{cases} (1 - \lambda(x))\, c_j + \lambda_j(x)\, c_{j+1} & \text{für } x \in (x_j - \delta, x_j + \delta) \\ f(x) & \text{sonst,} \end{cases}$$

wobei $c_0 := c_J$ und $c_{J+1} := c_1$, so dass $f_\delta(\pi) = f_\delta(-\pi)$. Für $\delta > 0$ hinreichend klein haben wir

$$(x_i - \delta, x_i + \delta) \cap (x_j - \delta, x_j + \delta) = \emptyset \quad \text{für } i, j \in \{0, \ldots, J\} \text{ mit } i \neq j,$$

so dass f_δ wohldefiniert ist und 2π-periodisch auf $\mathbb{R}$ fortgesetzt werden kann. Nun gilt für $x \in (x_j - \delta, x_j + \delta)$ nach Konstruktion

$$\begin{aligned} |f(x) - f_\delta(x)| &\leq \max\{|c_j - f_\delta(x)|, |c_{j+1} - f_\delta(x)|\} \\ &= \max\{|\lambda_j(x)(c_j - c_{j+1})|, |(1 - \lambda_j(x))(c_j - c_{j+1})|\} \\ &\leq |c_j - c_{j+1}|. \end{aligned}$$

Daraus folgt

$$\int_{-\pi}^{\pi} |f - f_\delta|^2 = \sum_{j=1}^{n} \int_{x_j - \delta}^{x_j + \delta} |f - f_\delta|^2 \leq 2\,\delta \sum_{j=1}^{n} |c_j - c_{j+1}|^2 \to 0 \text{ mit } \delta \to 0.$$

Sei schließlich $f \in X$ beliebig, das heißt es gibt eine Unterteilung $-\pi = x_0 < \ldots < x_J = \pi$ und stetige Funktionen $f_j : I_j = [x_{j-1}, x_j] \to \mathbb{C}$ mit $f = f_j$ auf (x_{j-1}, x_j). Nach Satz 1.5 in Kapitel 4 existieren Treppenfunktionen $g_j : I_j \to \mathbb{C}$ mit $\|f_j - g_j\|_{I_j} < \varepsilon$. Definiere nun auf $[-\pi, \pi]$ die Treppenfunktion

$$g(x) = \begin{cases} g_j(x) & \text{für } x \in (x_{j-1}, x_j) \\ f(x) & \text{für } x = x_j,\ 0 \leq j \leq J. \end{cases}$$

Es gilt dann $\|f - g\|_{[-\pi,\pi]} < \varepsilon$ und somit

$$\int_{-\pi}^{\pi} |f - g|^2 \leq 2\pi\varepsilon^2 \to 0 \quad \text{mit } \varepsilon \to 0.$$

$\square$

Der *Beweis von Satz 2.5* ergibt sich aus Satz 2.7 und Lemma 2.8 wie folgt: Sei V der Raum der trigonometrischen Polynome, und $\overline{V}$ sein Abschluss in X bzgl. $\|\cdot\|_2$, vgl. Proposition 2.1. Nach Satz 2.7 werden die 2π-periodischen, stetigen und stückweise linearen Funktionen bzgl. $\|\cdot\|_{[-\pi,\pi]}$, also erst recht bzgl. $\|\cdot\|_2$, durch ihre Fourierreihe approximiert, liegen also in $\overline{V}$. Aber nach Lemma 2.8 ist *jede* Funktion in X durch solche stückweise linearen Funktionen approximierbar. Da $\overline{V}$ abgeschlossen bzgl. $\|\cdot\|_2$, folgt $\overline{V} = X$.

Ausblick: Sei L^2 der Raum der Folgen $c = (c_k)_{k\in\mathbb{Z}}$ mit

$$\|c\|_2 := \left(2\pi \sum_{k\in\mathbb{Z}} |c_k|^2\right)^{1/2} < \infty.$$

Dieser Raum ist ein Banachraum, also vollständig (Übungsaufgabe 5, Serie 2). Durch die Fourierreihe wird jeder 2π-periodischen, stückweise stetigen Funktion f ein Element von L^2 zugeordnet, nämlich die Folge der Fourierkoeffizienten. Die Fourierreihe liefert also eine Abbildung

$$\mathcal{F} : X \to L^2,\ f \mapsto \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx\right)_{k\in\mathbb{Z}}.$$

Nach der Parsevalschen Gleichung, Satz 2.1, gilt

$$\|\mathcal{F}(f)\|_2 = \|f\|_2,$$

das heißt die Abbildung $\mathcal{F}$ ist isometrisch bzgl. der jeweiligen Normen $\|\cdot\|_2$. Im wesentlichen ist die Funktion $f \in X$ durch die Folge $\mathcal{F}(f)$ vollständig kodiert: für zwei Funktionen $f, g \in X$ mit $\mathcal{F}(f) = \mathcal{F}(g)$ folgt $\|\mathcal{F}(f-g)\|_2 = 0$, also stimmen sie an allen mit Ausnahme von endlich vielen Punkten überein. Offen bleibt die Frage, ob jeder Folge $(c_k)_{k\in\mathbb{Z}}$ umgekehrt auch eine Funktion entspricht. Der zugehörige Funktionenraum müßte bzgl. $\|\cdot\|_2$ vollständig sein; dies trifft auf X nicht zu. Oft wird der Entwicklungssatz im etwas größeren Raum der 2π-periodischen, Riemann-integrierbaren Funktionen gezeigt, was sich aus unseren Resultaten ohne große Mühe folgern lässt. Aber auch dieser Raum erweist sich als nicht vollständig bzgl. $\|\cdot\|_2$. Der Raum, für den die obige Abbildung ein Isomorphismus wird, ist der Raum der der quadratintegrierbaren Funktionen nach Lebesgue. Ihn werden wir in Analysis III kennenlernen.

Kapitel 7

Differentiation von Funktionen auf dem $\mathbb{R}^n$

1 Die Ableitungsbegriffe für Funktionen von mehreren Variablen

Bislang haben wir Funktionen betrachtet, die von einer reellen Variablen abhängen. Der Fall von Funktionen mit *Werten* in $\mathbb{R}^n$ konnte dabei gut mitbehandelt werden; nur gelegentlich – etwa beim Zwischenwertsatz und beim Schrankensatz - war Vorsicht geboten. Wir wollen nun Funktionen betrachten und differenzieren, die auf einer offenen Menge $\Omega \subset \mathbb{R}^n$ definiert sind und somit von n Variablen $x_1, \ldots, x_n$ abhängen. Hierfür gibt es diverse Beispiele.

Beispiel 1.1 *Ein Flächenstück kann durch seine Höhenfunktion $f : \Omega \to \mathbb{R}$, $x_3 = f(x_1, x_2)$, über einem ebenen Grundgebiet $\Omega \subset \mathbb{R}^2$ definiert werden.*

Beispiel 1.2 *Das elektrische Feld einer Kugelladung $q \in \mathbb{R}$ im Nullpunkt wird beschrieben durch*

$$f : \Omega = \mathbb{R}^3 \backslash \{0\} \to \mathbb{R}^3,\ f(x) = q\frac{x}{|x|^3},\ \text{wobei } |x| = (x_1^2 + x_2^2 + x_3^2)^{1/2}.$$

Beispiel 1.3 *Betrachte die komplexe Funktion*

$$f : \mathbb{C} \to \mathbb{C},\ f(z) = e^z.$$

Mit $z = x_1 + i\,x_2$, $f = f_1 + i\,f_2$ folgt die reelle Darstellung

$$f : \mathbb{R}^2 \to \mathbb{R}^2,\ f(x_1, x_2) = (e^{x_1} \cos x_2, e^{x_1} \sin x_2).$$

Beispiel 1.4 *Eine Beobachtungsgröße y möge von einem Parameter x abhängen. Es seien $N \geq 2$ Messdaten (x_k, y_k), $1 \leq k \leq N$, gegeben. Es werde die Hypothese zugrunde gelegt, dass die Abhängigkeit affin linear ist, d. h. $y = ax + b$ mit zu bestimmenden a, b. Um bestmögliche Werte zu bestimmen, wird oft die Methode der kleinsten Quadrate benutzt: a und b werden so bestimmt, dass die Funktion*

$$f(a, b) = \sum_{k=1}^{N} (ax_k + b - y_k)^2$$

minimal wird (sogenannte Ausgleichsgerade). Analog können Ausgleichspolynome höherer Ordnung betrachtet werden.

Beispiel 1.5 *Eine Parametrisierung der Sphäre $S^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$ ist wie folgt gegeben:*

$$f : \underbrace{(0, \pi) \times (0, 2\pi)}_{=\Omega\subset\mathbb{R}^2} \to S^2 \subset \mathbb{R}^3,\ f(\vartheta, \varphi) = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta) \in \mathbb{R}^3.$$

Eine Funktion mehrerer Variablen kann zu einer Funktion nur einer der Variablen gemacht werden, indem die übrigen Veränderlichen als feste Parameter aufgefasst werden. Wir beginnen damit, diesen Ansatz zu untersuchen.

Definition 1.1 *Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$. Die partielle Ableitung von f nach x_i (der i-ten Koordinate) an der Stelle $x \in \Omega$ ist der Grenzwert (falls existent)*

$$\partial_i f(x) = \lim_{t\to 0} \frac{f(x + te_i) - f(x)}{t}.$$

Andere Bezeichnungen sind $\frac{\partial f}{\partial x_i}(x)$ und $D_i f(x_0)$.

Die partielle Ableitung $\partial_i f(x)$ ist die gewöhnliche Ableitung an der Stelle $t = 0 \in \mathbb{R}$ der reellen Funktion

$$(-\delta, \delta) \to \mathbb{R}^m, \quad t \mapsto f(x + te_i),$$

bzw. alternativ die gewöhnliche Ableitung an der Stelle $x_i \in \mathbb{R}$ der reellen Funktion

$$(x_i - \delta, x_i + \delta) \to \mathbb{R}^m, \quad \xi \mapsto f(x_1, \ldots, x_{i-1}, \xi, x_{i+1}, \ldots, x_n).$$

Die partielle Ableitung kann also mittels der Differentiationsregeln für Funktionen einer reellen Variablen berechnet werden, indem die übrigen Variablen als Konstante aufgefasst werden.

Beispiel 1.6 *Sei $g : (0, \infty) \to \mathbb{R}$ stetig differenzierbar und*

$$f : \mathbb{R}^n\backslash\{0\} \to \mathbb{R}, \quad f(x) = g(r) \quad \textit{mit } r(x) = |x|.$$

Dann gilt für $x \neq 0$

$$\begin{aligned} \partial_i f(x) &= g'(r)\, \partial_i\big((x_1^2 + \ldots + x_n^2)^{\frac{1}{2}}\big) \\ &= g'(r)\, \frac{1}{2}(x_1^2 + \ldots + x_n^2)^{-\frac{1}{2}}\, 2x_i, \ \textit{also} \\ \partial_i f(x) &= g'(r)\, \frac{x_i}{r}. \qquad (1.1) \end{aligned}$$

Die Funktion $\partial_j f : \Omega \to \mathbb{R}^m$, $x \mapsto \partial_j f(x)$ kann erneut nach einer Variablen x_i partiell differenziert werden, vgl Definition 1.2. In diesem Beispiel erhalten wir, wenn g zweimal *stetig differenzierbar ist,*

$$
\begin{aligned}
\partial_i(\partial_j f)(x) &= \partial_i\left(\frac{g'(r)}{r}\right) x_j + \frac{g'(r)}{r}\,\partial_i\, x_j \\
&= \left(\frac{g'(r)}{r}\right)' \frac{x_i}{r} x_j + \frac{g'(r)}{r}\,\delta_{ij} \\
&= \frac{g''(r)r - g'(r)}{r^2}\frac{x_i\, x_j}{r} + \frac{g'(r)}{r}\,\delta_{ij}, \quad \textit{also}
\end{aligned}
$$

$$
\partial_i\partial_j f(x) = g''(r)\frac{x_i\, x_j}{r^2} + \frac{g'(r)}{r}\big(\delta_{ij} - \frac{x_i\, x_j}{r^2}\big). \tag{1.2}
$$

Der Operator $\Delta = \sum_{i=1}^n \partial_i^2$ heißt Laplaceoperator, die Lösungen von $\Delta f = 0$ harmonische Funktionen. Die rotationssymmetrischen harmonischen Funktionen können wir ausrechnen:

$$
0 \overset{!}{=} \Delta f(x) = g''(r) + \frac{n-1}{r} g'(r) = r^{1-n}\big(r^{n-1} g'(r)\big)'.
$$

Diese Gleichung hat die Lösungen, mit Integrationskonstanten $a, b \in \mathbb{R}$,

$$
g(r) = \begin{cases} a\, r^{2-n} + b & \text{für } n \geq 3 \\ a\, \log r + b & \text{für } n = 2. \end{cases}
$$

Als Anwendung der partiellen Ableitung behandeln wir parameterabhängige Integrale.

Lemma 1.1 *Seien $I = [\alpha, \beta]$, $J = [a, b]$ kompakte Intervalle und $f \in C^0(I \times J)$. Dann ist die Funktion*

$$
\varphi : [\alpha, \beta] \to \mathbb{R}, \quad \varphi(t) = \int_a^b f(t, x)\, dx
$$

wohldefiniert und stetig.

Beweis: Für jedes $t \in I$ ist $f(t, \cdot\,) \in C^0(J)$, also integrierbar nach Satz 1.5 in Kapitel 4. Nach Satz 1.11 in Kapitel 6 ist f gleichmäßig stetig auf $I \times J$, das heißt es gilt

$$
\omega_f(\delta) = \sup_{|(t,x)-(t',x')| \leq \delta} |f(t, x) - f(t', x')| \to 0 \quad \text{mit } \delta \to 0.
$$

Es folgt

$$
\begin{aligned}
|\varphi(t) - \varphi(t_0)| &= \Big|\int_a^b \big(f(t, x) - f(t_0, x)\big)\, dx\Big| \\
&\leq |J| \sup_{x \in J} |f(t, x) - f(t_0, x)| \\
&\leq |J|\ \omega_f(|t - t_0|) \to 0 \quad \text{mit } t \to t_0.
\end{aligned}
$$

$\square$

Satz 1.1 *Sei $f \in C^0(I \times J)$ nach $t \in I$ partiell differenzierbar und $\partial_t f : I \times J \to \mathbb{R}$ stetig. Dann ist die Funktion $\varphi(t) = \int_a^b f(t, x)\, dx$ stetig differenzierbar und es gilt*

$$
\varphi'(t) = \frac{d}{dt}\int_a^b f(t, x)\, dx = \int_a^b \partial_t f(t, x)\, dx.
$$

Beweis: Da $\partial_t f$ gleichmäßig stetig, gilt (siehe oben)

$$\omega_{\partial_t f}(\delta) = \sup_{|(t,x)-(t',x')|\leq\delta} |\partial_t f(t,x) - \partial_t f(t',x')| \to 0 \quad \text{mit } \delta \searrow 0.$$

Wir berechnen

$$\begin{aligned}
& \left|\frac{\varphi(t+h)-\varphi(t)}{h} - \int_a^b \partial_t f(t,x)\,dx\right| \\
= & \left|\frac{1}{h}\int_a^b \big(f(t+h,x) - f(t,x) - h\,\partial_t f(t,x)\big)\,dx\right| \\
= & \frac{1}{|h|}\left|\int_a^b \int_t^{t+h} \big(\partial_t f(\tau,x) - \partial_t f(t,x)\big)\,d\tau\,dx\right| \\
\leq & |J|\ \omega_{\partial_t f}(|h|) \to 0 \quad \text{mit } h \to 0.
\end{aligned}$$

□

Beispiel 1.7 *Sei* $\phi(t) = (\int_0^t e^{-x^2}dx)^2$ *für* $t \geq 0$. *Wir wollen das uneigentliche Integral* $\lim_{t\to\infty}\phi(t)$ *berechnen. Es gilt*

$$\begin{aligned}
\phi'(t) &= 2\,e^{-t^2}\int_0^t e^{-x^2}\,dx \\
&= 2\,te^{-t^2}\int_0^1 e^{-t^2y^2}\,dy \qquad \textit{(Substitution } x = ty) \\
&= \int_0^1 2te^{-(1+y^2)t^2}\,dy \\
&= -\int_0^1 \partial_t \frac{e^{-(1+y^2)t^2}}{1+y^2}\,dy \\
\textit{(Satz 1.1)} \quad &= \frac{d}{dt}\underbrace{\Big(-\int_0^1 \frac{e^{-(1+y^2)t^2}}{1+y^2}\,dy\Big)}_{=:\psi(t)}.
\end{aligned}$$

Es folgt $\phi(t) = \psi(t) + c$ *mit einer Konstanten* c. *Einsetzen von* $t = 0$ *liefert*

$$c = \phi(0) - \psi(0) = \int_0^1 \frac{dy}{1+y^2} = \arctan 1 = \frac{\pi}{4},$$

also $\phi(t) = \psi(t) + \frac{\pi}{4}$. *Aber es gilt*

$$|\psi(t)| \leq e^{-t^2}\int_0^1 \frac{dy}{1+y^2} \to 0 \quad \textit{mit } t \to \infty$$

und damit ergibt sich das Resultat

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}.$$

Eine zweite nützliche Anwendung ist die Vertauschbarkeit der Integrationsreihenfolge in Mehrfachintegralen.

Satz 1.2 (Kleiner Fubini) *Seien $I = [\alpha, \beta]$, $J = [a, b]$ kompakte Intervalle. Dann gilt für $f \in C^0(I \times J)$*

$$\int_\alpha^\beta \left(\int_a^b f(t,x)\,dx \right) dt = \int_a^b \left(\int_\alpha^\beta f(t,x)\,dt \right) dx.$$

Beweis: Wir setzen

$$\begin{aligned} \phi(t) &= \int_\alpha^t \left(\int_a^b f(\tau,x)\,dx \right) d\tau, \\ \psi(t) &= \int_a^b \left(\int_\alpha^t f(\tau,x)\,d\tau \right) dx. \end{aligned}$$

Nach Hauptsatz ist $\phi'(t) = \int_a^b f(t,x)\,dx$, während nach Satz 1.1 und Hauptsatz auch

$$\psi'(t) = \int_a^b \frac{\partial}{\partial t} \left(\int_\alpha^t f(\tau,x)\,d\tau \right) dx = \int_a^b f(t,x)\,dx.$$

Da ferner $\phi(\alpha) = 0 = \psi(\alpha)$, folgt $\phi(t) = \psi(t)$ für alle t und mit $t = \beta$ die Behauptung. $\square$

Wie schon festgestellt, können partielle Ableitungen iteriert werden.

Definition 1.2 *Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$. Die partiellen Ableitungen der Ordnung $k \geq 1$ sind induktiv definiert durch*

$$\partial_{i_1}\partial_{i_2}\ldots\partial_{i_k} f = \partial_{i_1}(\partial_{i_2}\ldots\partial_{i_k} f) \quad (i_j \in \{1,\ldots,n\}).$$

Wir bezeichnen mit $C^k(\Omega, \mathbb{R}^m)$ den $\mathbb{R}$-Vektorraum der k-mal stetig differenzierbaren Funktionen auf Ω mit Werten in $\mathbb{R}^m$, das heißt

$$C^k(\Omega,\mathbb{R}^m) = \{f : \text{alle partiellen Ableitungen } \partial_{i_1}\ldots\partial_{i_j} f \text{ für } 0 \leq j \leq k \text{ sind definiert und stetig auf } \Omega\}.$$

Satz 1.3 (Symmetrie der 2. Ableitung, H. A. Schwarz) *Sei $\Omega \subset \mathbb{R}^n$ offen. Für $f \in C^2(\Omega)$ und $1 \leq i, j \leq n$ gilt $\partial_i\,\partial_j\,f = \partial_j\,\partial_i\,f$.*

Beweis: Wir betrachten die Differenzenquotienten

$$\partial_j^t f(x) = \frac{1}{t}\big(f(x + te_j) - f(x)\big) \quad \text{wobei } t \neq 0.$$

Nach Definition gilt

$$\partial_i \partial_j f(x) = \lim_{s\to 0} \big(\lim_{t \to 0} \partial_i^s \partial_j^t f(x) \big). \tag{1.3}$$

Das Problem besteht darin, die beiden Grenzwerte zu vertauschen. Der Differenzenquotient vertauscht mit der partiellen Ableitung:

$$\partial_i \partial_j^t f(x) = \frac{1}{t}\big(\partial_i f(x + te_j) - \partial_i f(x)\big) = \partial_j^t \partial_i f(x).$$

Wir verwenden nun den Mittelwertsatz der Differentialrechnung. Ist g nach x_i partiell differenzierbar, so gilt für ein $\alpha \in [0,1]$:

$$\partial_i^s g(x) = \partial_i g(x + \alpha s e_i).$$

Wir wenden das nacheinander an auf $g = \partial_j^t f$ und auf $g = \partial_i f$:

$$\begin{aligned}
\partial_i^s \partial_j^t f(x) &= \partial_i(\partial_j^t f)(x + \alpha s e_i) \quad \big(\alpha = \alpha(s,t) \in [0,1]\big) \\
&= \partial_j^t(\partial_i f)(x + \alpha s e_i) \\
&= \partial_j(\partial_i f)(x + \alpha s e_i + \beta t e_j) \quad (\beta = \beta(s,t) \in [0,1]).
\end{aligned} \tag{1.4}$$

Da $\partial_j(\partial_i f)$ nach Voraussetzung stetig in x, folgt aus (1.3) und (1.4) die Behauptung.

□

Folgerung 1.1 *Für eine Funktion $f \in C^k(\Omega, \mathbb{R}^m)$ vertauschen die partiellen Ableitungen bis zur Ordnung k, d. h.*

$$\partial_{i_{\sigma(1)}} \dots \partial_{i_{\sigma(k)}} f = \partial_{i_1} \dots \partial_{i_k} f \quad \forall \sigma \in S(k).$$

BEWEIS: Nach Satz 1.3 können benachbarte Operatoren ∂_i, ∂_j vertauscht werden. Durch diese Vertauschungen wird die symmetrische Gruppe erzeugt. □

Eine geringfügige Verallgemeinerung der partiellen Ableitung ist die Richtungsableitung.

Definition 1.3 *Sei $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^m$. Für $x \in \Omega$ und $v \in \mathbb{R}^n$ heißt der Grenzwert*

$$\partial_v f(x) = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

Richtungsableitung von f an der Stelle x in Richtung v. Dies ist die gewöhnliche Ableitung der Funktion $t \mapsto f(x+tv)$ an der Stelle $t = 0$.

Beispiel 1.8 *Die Richtungsableitung von $r(x) = |x|$ im Punkt $x \in \mathbb{R}^n \backslash \{0\}$ in Richtung $v \in \mathbb{R}^n$ ist*

$$\begin{aligned}
\partial_v r(x) &= \frac{d}{dt}|x + tv| \,|_{t=0} \\
&= \frac{d}{dt}(|x+tv|^2)^{1/2} \,|_{t=0} \\
&= \frac{d}{dt}(|x|^2 + 2t\langle x, v\rangle + |v|^2)^{1/2} \,|_{t=0} \\
&= \frac{1}{2}(|x|^2)^{-1/2}\, 2\langle x, v\rangle \\
&= \langle \frac{x}{|x|}, v\rangle.
\end{aligned}$$

Für die mehrdimensionale Differentialrechnung ist der Begriff der partiellen Ableitung allein nicht ausreichend. Zwei entscheidende Mängel sind:

- Die Existenz der partiellen Ableitungen $\partial_1 f, \dots, \partial_n f$ im Punkt $x \in \mathbb{R}^n$ impliziert *nicht* die Stetigkeit von f im Punkt x.

- Es gilt keine Version der Kettenregel. Sei z. B. $c : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ mit $c(0) = x \in \mathbb{R}^n$. Die Existenz der partiellen Ableitungen $\partial_1 f, \dots, \partial_n f$ im Punkt x impliziert *nicht* die Differenzierbarkeit der Verkettung $f \circ c$ an der Stelle $t = 0$.

Beispiel 1.9 *Sei $\Omega = \mathbb{R}^2$ und*

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & (x,y) \neq 0 \\ 0 & (x,y) = (0,0). \end{cases}$$

Dann gilt

$$\partial_1 f(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = 0 = \partial_2 f(0,0).$$

Aber für $v = (\xi, \eta)$ mit $\xi \neq 0$, $\eta \neq 0$ gilt

$$f(tv) = \frac{\xi\eta}{\xi^2 + \eta^2} \neq 0 \quad \textit{für } t \neq 0.$$

Die Funktion f ist nicht stetig im Nullpunkt.

Die Definition der partiellen Ableitungen hängt von der Wahl der Koordinaten (für uns meist die Standardbasis) ab. Es wäre denkbar, dass dies Ursache des Problems ist, und dass sich die Missstände beheben lassen, indem man die Existenz aller Richtungsableitungen fordert. Das ist aber nicht so:

Beispiel 1.10 *Sei wieder auf $\Omega = \mathbb{R}^2$*

$$f(x,y) = \begin{cases} \frac{2xy^2}{x^2+y^4} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

Dann existieren im Punkt (0,0) alle Richtungsableitungen, denn für $v = (\xi, \eta) \neq 0$ ist

$$\partial_v f(0,0) = \lim_{t \to 0} \frac{2\xi\eta^2}{\xi^2 + t^2\eta^4} = \begin{cases} 2\eta^2/\xi & \xi \neq 0 \\ 0 & \xi = 0. \end{cases}$$

Dennoch ist die Funktion im Nullpunkt unstetig: für $c(t) = (\alpha t^2, t)$ mit $\alpha \in \mathbb{R}\backslash\{0\}$ gilt

$$f\big(c(t)\big) = \begin{cases} \frac{2\alpha}{\alpha^2+1} & \textit{für } t \neq 0 \\ 0 & \textit{für } t = 0. \end{cases}$$

Definition 1.4 *Sei $\Omega \subset \mathbb{R}^n$ offen, und $f : \Omega \to \mathbb{R}^m$. Die lineare Abbildung $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ heißt Ableitung von f im Punkt x_0 (Notation: $Df(x_0) = A$), falls gilt:*

$$\lim_{x \to x_0} \frac{f(x) - \big(f(x_0) + A(x - x_0)\big)}{|x - x_0|} = 0. \tag{1.5}$$

In Satz 1.4 unten wird gezeigt, dass es höchstens ein $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ mit (1.5) gibt. Wenn es ein solches A gibt, so heißt f differenzierbar im Punkt x_0.

Die Aussage „$Df(x_0) = A$“ bedeutet, dass sich die Funktion $f(x)$ für $x \to x_0$ in erster Ordnung wie die affinlineare Funktion $x \mapsto f(x_0) + A(x - x_0)$ verhält. Mit der Substitution $h = x - x_0$ läßt sich das äquivalent auch so formulieren:

$$R(h) := f(x_0 + h) - \big(f(x_0) + A\,h\big) \quad \Rightarrow \quad \lim_{h \to 0} \frac{R(h)}{|h|} = 0. \tag{1.6}$$

Beispiel 1.11 *Sei* $b : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ *symmetrische Bilinearform und*

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad f(x) = b(x,x).$$

Für $x_0 \in \mathbb{R}^n$ *ist dann* $Df(x_0) = 2b(x_0, \cdot) \in L(\mathbb{R}^n, \mathbb{R})$*, das heißt*

$$Df(x_0)v = 2\,b(x_0, v) \quad (v \in \mathbb{R}^n).$$

Denn es gilt

$$\begin{aligned} \frac{f(x_0+h) - \big(f(x_0) + 2b(x_0,h)\big)}{|h|} &= \frac{1}{|h|}\big(b(x_0+h, x_0+h) - b(x_0,x_0) - b(x_0,h) - b(h,x_0)\big) \\ &= \frac{1}{|h|} b(h,h) \longrightarrow 0 \quad \text{mit } h \to 0, \end{aligned}$$

da $|b(h,h)| = |\sum_{i,j=1}^n b_{ij}h_ih_j| \leq |h|^2 \sum_{i,j=1}^n |b_{ij}|$.

Beispiel 1.12 *Die Funktion* $f : I = (a,b) \to \mathbb{R}^m$ *habe in* $x_0 \in I$ *die Ableitung* $f'(x_0) \in \mathbb{R}^m$ *(im Sinne von Analysis 1). Wir setzen*

$$A : \mathbb{R} \to \mathbb{R}^m, \quad A\,h = f'(x_0)h.$$

Dann folgt für $h \neq 0$

$$\frac{\big|f(x_0+h) - \big(f(x_0) + Ah\big)\big|}{|h|} = \left|\frac{f(x_0+h) - f(x_0)}{h} - f'(x_0)\right| \to 0 \quad \text{mit } h \to 0,$$

Also ist f *differenzierbar im Sinne von Definition 1.4 mit Ableitung* $Df(x_0) = A$.

Beispiel 1.13 *Ist* $f : \Omega \to \mathbb{R}^m$ *Einschränkung der linearen Abbildung* $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ *auf* $\Omega \subset \mathbb{R}^n$,

$$f : \Omega \to \mathbb{R}^m, f(x) = Ax \quad (\Omega \subset \mathbb{R}^n),$$

so ist f *in allen* $x \in \Omega$ *differenzierbar mit Ableitung* $Df(x_0) = A$.

Satz 1.4 (Berechnung der Ableitung) *Die Funktion* $f : \Omega \to \mathbb{R}^m$ *habe im Punkt* $x_0 \in \Omega$ *die Ableitung* $A \in L(\mathbb{R}^m, \mathbb{R}^n)$*. Dann gilt:*

(1) $Av = \partial_v f(x_0)$ *(Richtungsableitung)* *für alle* $v \in \mathbb{R}^n$.

(2) *Bezüglich der Standardbasen von* $\mathbb{R}^n$ *und* $\mathbb{R}^m$ *hat* A *die Matrixdarstellung (Jacobimatrix)*

$$(\partial_j f_i(x_0)) = \begin{pmatrix} \partial_1 f_1(x_0) & \dots & \dots & \partial_n f_1(x_0) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \partial_1 f_m(x_0) & \dots & \dots & \partial_n f_m(x_0) \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

Insbesondere ist A *durch (1.5) eindeutig bestimmt.*

Beweis: Sei $v \in \mathbb{R}^n$, oBdA $v \neq 0$. Mit $h = tv$ folgt aus (1.6)

$$\begin{aligned}
\frac{f(x_0+tv)-f(x_0)}{t} &= Av + \frac{f(x_0+tv)-\big(f(x_0)+A(tv)\big)}{t} \\
&= Av + \frac{R(tv)}{t} \\
&\to Av \quad \text{mit } t \to 0.
\end{aligned}$$

Damit ist (1) gezeigt. Für (2) wähle $v := e_j$ und erhalte aus (1) und Lemma 1.1 in Kapitel 3

$$Df(x_0)e_j = \partial_j f(x_0) = \sum_{i=1}^{m} \partial_j f_i(x_0) e_i.$$

□

Definition 1.5 *Sei $\Omega \subset \mathbb{R}^n$ offen, und $f : \Omega \to \mathbb{R}$ in $x_0 \subset \Omega$ differenzierbar. Der Gradient von f in x_0 (bzgl. des Standardskalarprodukts $\langle \cdot , \cdot \rangle$) ist der Vektor*

$$\operatorname{grad} f(x_0) = \big(\partial_1 f(x_0), \ldots, \partial_n f(x_0)\big) \in \mathbb{R}^n.$$

Der Gradient ist charakterisiert durch die Eigenschaft

$$Df(x_0)v = \langle \operatorname{grad} f(x_0), v\rangle \quad \textit{für alle } v \in \mathbb{R}^n.$$

Beispiel 1.14 *Der Gradient von $f(x) = g(r)$ mit $r(x) = |x|$ ist*

$$\operatorname{grad} f(x) = g'(r)\frac{x}{r} \quad \textit{(vgl. Beispiel 1.6).}$$

Beispiel 1.15 *Sei $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ symmetrische Bilinearform, und $B \in L(\mathbb{R}^n, \mathbb{R}^n)$ der darstellende Endomorphismus bzgl. des Standardskalarprodukts*

$$b(x,y) = \langle Bx, y\rangle \quad \textit{für alle } x, y \in \mathbb{R}^n.$$

Sei $f : \mathbb{R}^n \to \mathbb{R}$, $\quad f(x) = b(x,x)$. Dann gilt

$$\operatorname{grad} f(x) = 2Bx \quad \textit{für alle } x \in \mathbb{R}^n.$$

Denn nach Beispiel 1.11 gilt für alle $v \in \mathbb{R}^n$

$$\langle \operatorname{grad} f(x), v\rangle = Df(x)v = 2b(x,v) = \langle 2Bx, v\rangle.$$

Satz 1.5 (Komponentenweise Differentiation) *Die Funktion $f : \Omega \to \mathbb{R}^m$ ist genau dann in $x_0 \in \Omega$ differenzierbar, wenn alle Komponentenfunktionen $f_i : \Omega \to \mathbb{R}^m$, $i = 1, \ldots, m$, differenzierbar sind.*

Beweis: Wir können von vornherein annehmen, dass die partiellen Ableitungen $a_{ij} = \partial_j f_i(x_0)$ für $i = 1, \ldots, m$, $j = 1, \ldots, n$ existieren, denn das ist nach Satz 1.4 in beiden Richtungen gegeben. Dann ist mit Satz 1.4 die Funktion f genau dann differenzierbar in x_0, wenn gilt

$$\lim_{h \to 0} \frac{1}{|h|} \left(f(x_0+h) - \big(f(x_0) + \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} h_j e_i\big) \right) = 0.$$

Entsprechend sind alle f_i genau dann in x_0 differenzierbar, wenn

$$\lim_{h\to 0}\frac{1}{|h|}\left(f_i(x_0+h)-\big(f_i(x_0)+\sum_{j=1}^n a_{ij}h_j\big)\right)=0 \text{ für alle } i=1,\dots,m.$$

Die Behauptung folgt nun aus der Äquivalenz von Konvergenz und komponentenweiser Konvergenz im $\mathbb{R}^m$. □

Es wäre mühselig, die Differenzierbarkeit von Funktionen stets anhand der Definition zu prüfen. Wie oben erklärt, reicht die Existenz der partiellen Ableitungen nicht aus. Der folgende Satz liefert das fehlende, hinreichende Kriterium.

Satz 1.6 (stetig partiell differenzierbar ⇒ differenzierbar) *Sei $\Omega\subset\mathbb{R}^n$ offen. Die Funktion $f:\Omega\to\mathbb{R}^m$ sei in allen $x\in\Omega$ nach $x_1,\dots,x_n$ partiell differenzierbar. Sind die Ableitungen $\partial_i f:\Omega\to\mathbb{R}^m$ in $x_0\in\Omega$ stetig, so ist f in x_0 differenzierbar.*

Beweis: Wir können $m=1$ annehmen (Folgerung 1.5). Kandidatin für die Ableitung ist (vgl. Satz 1.4)

$$A:\mathbb{R}^n\to\mathbb{R},\quad Ah=\sum_{k=1}^n\partial_k f(x_0)h_k.$$

Sei $x_k=x_0+\sum_{i=1}^k h_ie_i$ für $1\le k\le n$. Nach dem Mittelwertsatz der Differentialrechnung gilt

$$f(x_k)-f(x_{k-1})=f(x_{k-1}+h_ke_k)-f(x_{k-1})=\partial_k f(x_{k-1}+s_kh_ke_k)h_k$$

mit $s_k\in[0,1]$ geeignet. Es folgt

$$\begin{aligned}\frac{1}{|h|}\left|f(x_0+h)-\big(f(x_0)+Ah\big)\right| &= \frac{1}{|h|}\Big|\sum_{k=1}^n\big(f(x_k)-f(x_{k-1})-\partial_k f(x_0)h_k\big)\Big|\\ &= \frac{1}{|h|}\Big|\sum_{k=1}^n\big(\partial_k f(x_0+s_kh_ke_k)-\partial_k f(x_0)\big)h_k\Big|\\ &\to 0\quad\text{mit } h\to 0.\end{aligned}$$

□

Wir wollen nun zeigen, dass der eingeführte Begriff der Differenzierbarkeit die dargestellten Nachteile der partiellen Ableitung nicht hat.

Satz 1.7 (Differenzierbarkeit ⇒ Stetigkeit) *Ist $f:\Omega\to\mathbb{R}^m$ differenzierbar in $x_0\in\Omega\subset\mathbb{R}^n$, so ist f auch stetig in x_0.*

Beweis: Sei $Df(x_0)=A\in L(\mathbb{R}^n,\mathbb{R}^m)$ und $R(h)=f(x_0+h)-\big(f(x_0)+Ah\big)$. Für $x\neq x_0$ gilt

$$\begin{aligned}|f(x)-f(x_0)| &= |f(x_0+x-x_0)-f(x_0)|\\ &= |A(x-x_0)+R(x-x_0)|\\ \text{(Kap. 2, (2.2))}\quad &\le |A||x-x_0|+\frac{|R(x-x_0)|}{|x-x_0|}|x-x_0|\\ &\to 0\quad\text{mit } x\to x_0.\end{aligned}$$

$\square$

Wir kommen jetzt zu den Differentiationsregeln.

Satz 1.8 (Kettenregel) *Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen und $U \xrightarrow{f} V \xrightarrow{g} \mathbb{R}^p$. Es sei f differenzierbar in $x_0 \in U$, g differenzierbar in $y_0 = f(x_0) \in V$. Dann ist $g \circ f$ differenzierbar in $x_0 \in U$ und es gilt*

$$D(g \circ f)(x_0) = Dg\big(f(x_0)\big)Df(x_0).$$

Bemerkung. Für die zugehörigen Jacobimatrizen bedeutet das

$$\underbrace{\left(\frac{\partial (g \circ f)_i}{\partial x_j}(x_0)\right)}_{p\times n} = \underbrace{\left(\frac{\partial g_i}{\partial y_k}(f(x_0))\right)}_{p\times m} \cdot \underbrace{\left(\frac{\partial f_k}{\partial x_j}(x_0)\right)}_{m\times n}$$

bzw. elementweise

$$\frac{\partial (g \circ f)_i}{\partial x_j}(x_0) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}\big(f(x_0)\big)\frac{\partial f_k}{\partial x_j}(x_0).$$

Beweis: Wir setzen $Df(x_0) = A$, $Dg(y_0) = B$ und

$$\begin{aligned} R_f(\xi) &= f(x_0 + \xi) - \big(f(x_0) + A\xi\big), \\ R_g(\eta) &= g(y_0 + \eta) - \big(g(y_0) + B\eta\big). \end{aligned}$$

Kandidatin für die Ableitung von $g \circ f$ bei x_0 ist BA. Wir bilden die entsprechende Restfunktion für $\xi \neq 0$

$$\begin{aligned} \frac{R_{g\circ f}(\xi)}{|\xi|} &:= \frac{1}{|\xi|}\Big((g \circ f)(x_0 + \xi) - \big((g \circ f)(x_0) + BA\xi\big)\Big) \\ (\text{Def. } R_f, y_0 = f(x_0)) &= \frac{1}{|\xi|}\Big(g(y_0 + \underbrace{A\xi + R_f(\xi)}_{=:\eta(\xi)}) - g(y_0) - BA\xi\Big) \\ (\text{Def. } R_g(\eta)) &= \frac{1}{|\xi|}\Big(g(y_0) + B\eta(\xi) + R_g\big(\eta(\xi)\big) - g(y_0) - BA\xi\Big) \\ &= \underbrace{B\frac{R_f(\xi)}{|\xi|}}_{\mathrm{I}} + \underbrace{\frac{R_g\big(\eta(\xi)\big)}{|\xi|}}_{\mathrm{II}}. \end{aligned}$$

Zu zeigen ist, dass mit $\xi \to 0$ die rechte Seite gegen Null geht. Es gilt

$$|\mathrm{I}| \leq |B|\frac{|R_f(\xi)|}{|\xi|} \to 0 \quad \text{nach Vor. und Def. von } R_f.$$

Wir wollen II mit $|\eta(\xi)|$ erweitern. Um den Fall $\eta(\xi) = 0$ mitzubehandeln, definieren wir

$$\varepsilon_g(\eta) = \begin{cases} |R_g(\eta)|/|\eta| & \text{für } \eta \neq 0 \\ 0 & \text{für } \eta = 0. \end{cases}$$

Da $Dg(y_0) = B$, ist ε_g nach Definition von R_g stetig im Punkt $\eta = 0$. Beachte nun

$$|\eta(\xi)| = |A\xi + R_f(\xi)| \leq \left(|A| + \frac{|R_f(\xi)|}{|\xi|}\right)|\xi|.$$

Es folgt aus der Definition von ε_g

$$|\mathrm{II}| = \varepsilon_g\big(\eta(\xi)\big)\frac{|\eta(\xi)|}{|\xi|} \leq \underbrace{\varepsilon_g\big(\eta(\xi)\big)}_{\to 0}\underbrace{\left(|A| + \frac{|R_f(\xi)|}{|\xi|}\right)}_{\leq \mathrm{const}} \to 0 \text{ mit } \xi \to 0.$$

$\square$

Beispiel 1.16 (Kettenregel für Kurven) *Sei $c : I = (a,b) \longrightarrow \Omega \subset \mathbb{R}^n$ differenzierbar in $t_0 \in I$, und $f : \Omega \longrightarrow \mathbb{R}$ differenzierbar in $x_0 = c(t_0)$. Dann ist $f \circ c : I \longrightarrow \mathbb{R}$ in t_0 differenzierbar und es gilt*

$$(f \circ c)'(t_0) = Df(x_0)\, c'(t_0) = \langle \operatorname{grad} f(x_0), c'(t_0)\rangle.$$

Falls $f\big(c(t)\big)$ konstant (Höhenlinie), so gilt $\operatorname{grad} f(x_0) \perp c'(t_0)$.

Beispiel 1.17 (Parameterdifferentiation) *Sei $f \in C^0(I \times J)$, $I \times J = [\alpha,\beta] \times [a,b]$, nach $t \in I$ stetig partiell differenzierbar, also $\frac{\partial f}{\partial t} \in C^0(I \times J)$. Sind $x_1, x_2 : [\alpha,\beta] \to (a,b)$ differenzierbar, so gilt*

$$\frac{d}{dt}\int_{x_1(t)}^{x_2(t)} f(t,x)\,dx = \big[f\big(t, x_i(t)\big)\, x_i'(t)\big]_{i=1}^{i=2} - \int_{x_1(t)}^{x_2(t)} \frac{\partial f}{\partial t}(t,x)\,dx.$$

Denn die Funktion

$$\phi : I \times (a,b) \times (a,b) \to \mathbb{R}, \quad \phi(t,x,y) = \int_x^y f(t,\xi)\,d\xi$$

ist stetig partiell differenzierbar nach Hauptsatz der Differential- und Integralrechnung (bzgl. x und y) und nach Satz 1.1 zur Parameterableitung (bzgl. t). Nach Satz 1.6 ist ϕ differenzierbar und die Kettenregel ergibt

$$\begin{aligned}
\frac{d}{dt}\int_{x_1(t)}^{x_2(t)} f(t,x)\,dx &= \frac{d}{dt}\phi\big(t, x_1(t), x_2(t)\big) \\
&= \frac{\partial \phi}{\partial t}(t, x_1, x_2) + \frac{\partial \phi}{\partial x}(\ldots)x_1' + \frac{\partial \psi}{\partial y}(\ldots)x_2' \\
&= \big[f\big(t, x_i(t)\big)x_i'(t)\big]_{i=1}^{i=2} + \int_{x_1(t)}^{x_2(t)} \frac{\partial f}{\partial t}(t,x)\,dx.
\end{aligned}$$

Außer der Kettenregel sind natürlich auch die anderen bekannten Ableitungsregeln erneut zu beweisen.

Satz 1.9 (Ableitungsregeln) *Sei $\Omega \subset \mathbb{R}^n$ offen und $x_0 \in \Omega$. Die Funktionen $f, g : \Omega \to \mathbb{R}^m$ seien in x_0 differenzierbar. Dann gelten folgende Aussagen:*

(1) *Für $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g : \Omega \longrightarrow \mathbb{R}^m$ in x_0 differenzierbar und es gilt*

$$D(\alpha f + \mu g)(x_0) = \alpha Df(x_0) + \beta Dg(x_0).$$

(2) *Sind f, g reellwertig (also $m = 1$), so ist $fg : \Omega \to \mathbb{R}$ in $x_0 \in \Omega$ differenzierbar und*

$$\begin{aligned} D(fg)\,(x_0) &= Df(x_0)\,g(x_0) + f(x_0)\,Dg(x_0), \quad \textit{bzw.} \\ \operatorname{grad}(fg)(x_0) &= \operatorname{grad} f(x_0)\,g(x_0) + f(x_0)\operatorname{grad} g(x_0). \end{aligned}$$

(3) *Sind f, g reellwertig und ist $g(x_0) \neq 0$, so ist f/g in x_0 differenzierbar und es gilt*

$$D\left(\frac{f}{g}\right)(x_0) = \frac{Df(x_0)\,g(x_0) - f(x_0)\,Dg(x_0)}{g(x_0)^2}.$$

Beweis: Setze $Df(x_0) = A$, $Dg(x_0) = B$, sowie

$$R_f(h) = f(x_0 + h) - \big(f(x_0) + A(h)\big), \quad R_g(h) = g(x_0 + h) - (g(x_0) + Bh).$$

Mit den entsprechend definierten Restfunktionen prüfen wir jeweils die Definition von Ableitung.

(1)

$$\begin{aligned} \frac{R_{\alpha f+\beta g}(h)}{|h|} &:= \frac{(\alpha f + \beta g)(x_0 + h) - \big((\alpha f + \beta g)(x_0) + (\alpha A + \beta B)h\big)}{|h|} \\ &= \alpha\frac{f(x_0 + h) - \big(f(x_0) + Ah\big)}{|h|} + \beta\frac{g(x_0 + h) - \big(g(x_0) + Bh\big)}{|h|} \\ &\quad \to 0 \text{ mit } h \to 0. \end{aligned}$$

(2) Übungsaufgaben

(3) oBdA $f \equiv 1$ (sonst schreibe $\frac{f}{g} = f \cdot \frac{1}{g}$ und verwende die Produktregel (2)).

$$\begin{aligned} R_{1/g}(h) &:= \frac{1}{g}(x_0 + h) - \left(\frac{1}{g}(x_0) - \frac{1}{g(x_0)^2}Bh\right) \\ &= \frac{1}{g(x_0 + h)g(x_0)}\Big(g(x_0) - g(x_0 + h) + \frac{g(x_0 + h)}{g(x_0)}Bh\Big) \\ &= \frac{1}{g(x_0 + h)g(x_0)}\left(\left(\frac{g(x_0 + h)}{g(x_0)} - 1\right)Bh - R_g(h)\right) \end{aligned}$$

Es folgt

$$\frac{|R_{1/g}(h)|}{|h|} \leq \frac{1}{|g(x_0 + h)\,g(x_0)|}\left(\underbrace{\left|\frac{g(x_0 + h)}{g(x_0)} - 1\right|}_{\to 0}|B| + \underbrace{\frac{|R_g(h)|}{|h|}}_{\to 0}\right).$$

Die rechte Seite geht gegen Null mit $h \to 0$, da g stetig in x_0 (Satz 1.7) und $g(x_0) \neq 0$. $\square$

2 Erste Anwendungen der Ableitung

Ein sehr häufig auftretendes Problem in der Analysis ist es, Informationen über die Ableitung in Eigenschaften der Funktion zu übersetzen. Für Funktionen einer Variablen stehen hierzu zwei Argumente zur Verfügung.

a) der Mittelwertsatz (siehe Kapitel 3.2):

$$f(x_1) - f(x_0) = f'(\xi)(x_1 - x_0) \quad ([x_0, x_1] \subset \mathbb{R}),$$

b) der Hauptsatz der Differential- und Integralrechnung (siehe Kapitel 4.2):

$$f(x_1) - f(x_0) = \int_{x_0}^{x_1} f'(x)\,dx \quad ([x_0, x_1] \subset \mathbb{R}).$$

Vergleich der Vor-/Nachteile:

- Der Mittelwertsatz verlangt nur, dass f differenzierbar ist. Der Hauptsatz braucht $f \in C^1$ bzw. f stückweise C^1.
- Der Mittelwertsatz gilt so nicht für vektorwertige Funktionen (die komponentenweise Anwendung ergibt unterschiedliche Zwischenstellen ξ). Beim Hauptsatz ist f vektorwertig zulässig.
- Die Zwischenstelle ξ im Mittelwertsatz hängt in unkontrollierbarer Weise von x_0, x_1 und f ab.

Wie kann man diese Werkzeuge nun für Funktionen mehrerer Variabler $f : \Omega \to \mathbb{R}$ mit $\Omega \subset \mathbb{R}^n$ einsetzen? Die einfache Antwort heißt: indem man f längs Kurven

$$\gamma : [a, b] \to \Omega, \quad \gamma = \gamma(t)$$

auswertet.

Lemma 2.1 *Sei $\gamma : I = [a, b] \to \Omega$ stetig und stückweise C^1, d. h. es gibt eine Unterteilung $a = t_0 < t_1 < \ldots < t_N = b$, so dass $\gamma|_{[t_{j-1}, t_j]}$ stetig differenzierbar ist für $j = 1, \ldots, N$. Dann gilt für jede Funktion $f \in C^1(\Omega, \mathbb{R}^m)$*

$$f\big(\gamma(b)\big) - f\big(\gamma(a)\big) = \int_a^b Df\big(\gamma(t)\big)\gamma'(t)\,dt. \tag{2.1}$$

Beweis: f ist differenzierbar nach Satz 1.6. Ist $\gamma \in C^1([a, b], \mathbb{R}^n)$, so folgt aus Kettenregel und Hauptsatz

$$\int_a^b Df\big(\gamma(t)\big)\,\gamma'(t)\,dt = \int_a^b \frac{d}{dt} f\big(\gamma(t)\big)\,dt = f\big(\gamma(b)\big) - f\big(\gamma(a)\big).$$

Ist f nur stückweise C^1, so folgt durch Anwendung auf jedem Teilintervall

$$\begin{aligned} f\big(\gamma(b)\big) - f\big(\gamma(a)\big) &= \sum_{j=1}^N f\big(\gamma(t_j)\big) - f\big(\gamma(t_{j-1})\big) \\ &= \sum_{j=1}^N \int_{t_{j-1}}^{t_j} Df\big(\gamma(t)\big)\,\gamma'(t)\,dt \\ &= \int_a^b Df\big(\gamma(t)\big)\,\gamma'(t)\,dt. \end{aligned}$$

Dabei ist γ' stückweise stetig (an den Unterteilungspunkten nicht notwendig definiert). $\square$

Definition 2.1 *Eine Menge $M \subset \mathbb{R}^n$ heißt wegweise zusammenhängend, falls es zu je zwei Punkten $x_0,\ x_1 \in M$ eine stetige Kurve $c : [0,1] \to M \subset \mathbb{R}^n$ gibt mit $c(0) = x_0,\ c(1) = x_1$.*

Lemma 2.2 *Sei $\Omega \subset \mathbb{R}^n$ offen und wegweise zusammenhängend. Dann gibt es eine stetige, stückweise affin-lineare Abbildung $\gamma : [0,1] \to \Omega$, d.h. einen Polygonzug, mit $\gamma(0) = x_0,\ \gamma(1) = x_1$.*

Beweis: Sei $c \in C^0([0,1],\Omega)$ mit $c(0) = x_0$ und $c(1) = x_1$ gegeben. Betrachte die Zerlegung $t_i = i/N$ für $i = 0,\ldots,N$ des Intervalls $[0,1]$, und definiere einen Polygonzug $\gamma : [0,1] \to \mathbb{R}^n$ mit $\gamma(t_i) = c(t_i)$ für alle i durch

$$\gamma(t) = \frac{t_i - t}{t_i - t_{i-1}}\, c(t_{i-1}) + \frac{t - t_{i-1}}{t_i - t_{i-1}}\, c(t_i) \qquad \text{für } t \in [t_{i-1}, t_i].$$

Dann folgt für $t \in [t_{i-1}, t_i]$ die Abschätzung

$$\begin{aligned} |\gamma(t) - c(t_{i-1})| &= \frac{t - t_{i-1}}{t_i - t_{i-1}}\, |c(t_i) - c(t_{i-1})| \\ &\le \sup_{|t-s| \le 1/N} |c(t) - c(s)| \to 0 \text{ mit } N \to \infty, \end{aligned}$$

da c gleichmäßig stetig auf $[0,1]$. Nun ist $c([0,1])$ kompakt und somit $\operatorname{dist}(c([0,1]), \mathbb{R}^n \backslash \Omega) > 0$. Daraus folgt für N hinreichend groß $\gamma(t) \in \Omega$ für alle $t \in [0,1]$. $\square$

Satz 2.1 (Konstanzsatz) *Sei $\Omega \subset \mathbb{R}^n$ offen und wegweise zusammenhängend. Dann gilt:*

$$Df(x) = 0 \quad \textit{für alle } x \in \Omega \quad \Rightarrow \quad f \textit{ ist konstant.}$$

Beweis: Seien x_0, x_1 beliebige Punkte in Ω. Nach Lemma 2.2 gibt es einen Polygonzug $\gamma : [0,1] \to \Omega$ mit $\gamma(0) = x_0$ und $\gamma(1) = x_1$. Da γ stetig und stückweise C^1, folgt aus Lemma 2.1

$$f(x_1) - f(x_0) = \int_0^1 Df(\gamma(t))\gamma'(t)\, dt = 0.$$

$\square$

Definition 2.2 *$M \subset \mathbb{R}^n$ heißt konvex, falls gilt:*

$$x_0, x_1 \in M \;\Rightarrow\; (1-t)\, x_0 + t\, x_1 \in M \quad \textit{für alle } t \in [0,1].$$

Satz 2.2 (Schrankensatz) *Sei $\Omega \subset \mathbb{R}^n$ offen und konvex. Für $f \in C^1(\Omega, \mathbb{R}^m)$ gelte $\sup_{x \in \Omega} |Df(x)| \le L < \infty$. Dann folgt*

$$|f(x_1) - f(x_0)| \le L\, |x_1 - x_0|.$$

Beweis: *Erinnerung:* Ist $\varphi : I = [a,b] \to \mathbb{R}^m$ stetig, so gilt $|\int_a^b \varphi| \le \int_a^b |\varphi|$. Dies folgt durch Betrachtung der Riemannschen Summen mit der Dreiecksungleichung, siehe Kapitel 4.1, Satz 1.6.

Sei nun $\gamma(t) = (1-t)x_0 + tx_1$ für $0 \le t \le 1$. Aus Lemma 2.1 folgt wegen $\gamma'(t) = (x_1 - x_0)$ für alle t

$$\begin{aligned} |f(x_1) - f(x_0)| &= \left| \int_0^1 Df\big(\gamma(t)\big)(x_1 - x_0)\, dt \right| \\ &\le \int_0^1 |Df\big(\gamma(t)\big)(x_1 - x_0)|\, dt \\ &\le \int_0^1 \underbrace{|Df\big(\gamma(t)\big)|}_{\le L} |x_1 - x_0|\, dt \\ &\le L\,|x_1 - x_0|. \end{aligned}$$

□

Bemerkung. Die Voraussetzung in Satz 2.2 kann abgeschwächt werden zu $\sup_{x\in\Omega} \|Df(x)\| \le L < \infty$, wobei $\|A\| = \sup_{|v|\le 1} |Av|$ die Operatornorm bezeichnet.

Folgerung 2.1 *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^m)$. Dann gibt es zu jeder kompakten Menge $K \subset \Omega$ eine Konstante $L < \infty$ mit*

$$|f(x) - f(y)| \le L\,|x - y| \quad \textit{für alle } x, y \in K.$$

Beweis: (*indirekt*) Da $f : K \to \mathbb{R}^m$ stetig, ist $M := \sup_{x\in K} |f(x)| < \infty$ (Satz 4.2 in Kapitel 2). Wäre die Behauptung falsch, so gibt es zu $k \in \mathbb{N}$ Punkte x_k, $y_k \in K$ mit

$$|f(x_k) - f(y_k)| > k\,|x_k - y_k| \quad (k \in \mathbb{N}).$$

Da K kompakt, gilt nach Wahl einer Teilfolge und Umnumerierung $x_k \to x \in K$ mit $k \to \infty$. Aber

$$|x_k - y_k| < \frac{1}{k}\,|f(x_k) - f(y_k)| \le \frac{2M}{k} \to 0 \quad \text{mit } k \to \infty.$$

Daraus folgt auch $y_k \to x$ mit $k \to \infty$. Wähle nun $r > 0$ mit $\overline{B_r(x)} \subset \Omega$. Da Df stetig, ist

$$L := \sup_{y\in\overline{B_r(x)}} |Df(y)| < \infty.$$

Da $x_k, y_k \in B_r(x)$ für k hinreichend groß, folgt aus Satz 2.2 für k hinreichend groß

$$k|x_k - y_k| < |f(x_k) - f(y_k)| \le L|x_k - y_k|,$$

ein Widerspruch. □

Zu den wichtigen Anwendungen der ersten und zweiten Ableitung gehört natürlich die Diskussion von Extremwerten. Wir wollen dazu die Taylorentwicklung von Funktionen mehrerer Variabler herleiten. Wir müssen dies zunächst für Funktionen einer Variablen durchführen - gewissermaßen ein Ausflug zurück nach Analysis 1.
Die Idee der Taylorentwicklung ist es, eine gegebene Funktion f mit einem Polynom zu vergleichen, das an einer festen Stelle x_0 mit f „von höherer Ordnung“ übereinstimmt, das heißt einschließlich einer Reihe von Ableitungen.

Lemma 2.3 *Sei $I = (a,b) \subset \mathbb{R}$, $x_0 \in I$ und $n \in \mathbb{N}_0$. Zu $f \in C^n(I)$ gibt es genau ein Polynom $P : \mathbb{R} \to \mathbb{R}$ vom Grad höchstens n mit $P^{(k)}(x_0) = f^{(k)}(x_0)$ für $k = 0, 1, \ldots, n$, und zwar*

$$P(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j. \tag{2.2}$$

Das Polynom P heißt Taylorpolynom n-ten Grades von f mit Entwicklungspunkt x_0 und wird mit $P^n_{x_0} f$ bezeichnet.

Beweis: Es gilt $\left(\frac{d}{dx}\right)^k (x - x_0)^j \,|_{x=x_0} = k!\, \delta_{jk}$. Daraus ergibt sich allgemein für $k \leq n$

$$P(x) = \sum_{j=0}^{n} a_j (x - x_0)^j \Rightarrow P^{(k)}(x_0) = k!\, a_k. \tag{2.3}$$

Für P wie in (2.2) gilt $a_k = \frac{f^{(k)}(x_0)}{k!}$, also folgt $P^{(k)}(x_0) = f^{(k)}(x_0)$ und P leistet das Verlangte.

Für die Eindeutigkeit reicht es aus zu zeigen, dass ein Polynom P vom Grad höchstens n mit $P^{(k)}(x_0) = 0$ für $k = 0, 1, \ldots, n$ das Nullpolynom ist. Nun gilt

$$x^k = (x_0 + x - x_0)^k = \sum_{i=0}^{k} \binom{k}{i} x_0^{k-i} (x - x_0)^i.$$

Also gilt $P(x) = \sum_{j=0}^{n} a_j (x - x_0)^j$ mit geeigneten $a_j \in \mathbb{R}$, und aus (2.3) folgt $a_j = \frac{1}{j!} P^{(j)}(x_0) = 0$. Somit ist P das Nullpolynom (vgl. Kap. 2, Lemma 1.2). $\square$

Folgerung 2.2 *Ist f ein Polynom vom Grad höchstens n, so gilt $P^n_{x_0} f = f$, das heißt das n-te Taylorpolynom von f ist f selbst.*

Definition 2.3 *Für $f \in C^n(I)$ und $x_0 \in I$ heißt die Funktion*

$$R^n_{x_0} f : (a,b) \to \mathbb{R}, \quad R^n_{x_0} f(x) = f(x) - P^n_{x_0} f(x)$$

Restglied der Taylorentwicklung n-ter Ordnung mit Entwicklungspunkt x_0.

Knackpunkt bei der Taylorentwicklung ist die Abschätzung des Restglieds; diese besagt, wie gut die Funktion f durch das Taylorpolynom $P^n_{x_0} f$ approximiert wird. Entscheidend für die Abschätzung sind die in den folgenden beiden Sätzen gelieferten Darstellungen.

Satz 2.3 (Integraldarstellung des Restglieds) *Sei $f \in C^{n+1}(I)$ für ein $n \in \mathbb{N}_0$ und $x_0 \in I$. Dann hat das Restglied $R^n_{x_0} f(x) = f(x) - P^n_{x_0} f(x)$ die Darstellung*

$$R^n_{x_0} f(x) = \frac{1}{n!} \int_{x_0}^{x} (x - y)^n f^{(n+1)}(y)\, dy. \tag{2.4}$$

Beweis: Durch Induktion über $n \in \mathbb{N}_0$. Der Fall $n = 0$ folgt aus dem Hauptsatz der Differential- und Integralrechnung:

$$R^0_{x_0} f(x) = f(x) - f(x_0) = \int_{x_0}^{x} f'(y)\, dy.$$

Der Induktionsschritt $n-1 \Rightarrow n \geq 1$ beruht auf partieller Integration:

$$\begin{aligned}
R^n_{x_0} f(x) &= R^{n-1}_{x_0} f(x) - \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \quad \text{(Definition } R^n_{x_0}\text{)}\\
\text{(Induktion)} &= \frac{1}{(n-1)!}\int_{x_0}^x (x-y)^{n-1} f^{(n)}(y)\,dy - \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\\
\text{(part. Int.)} &= \frac{1}{(n-1)!}\left[-\frac{(x-y)^n}{n}\,f^{(n)}(y)\right]_{y=x_0}^{y=x} + \frac{1}{n!}\int_{x_0}^x (x-y)^n\, f^{(n+1)}(y)\,dy\\
&\quad -\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\\
&= \frac{1}{n!}\int_{x_0}^x (x-y)^n\, f^{(n+1)}(y)\,dy.
\end{aligned}$$

□

Lemma 2.4 (Kapitel 4, Folgerung 1.3) *Seien $f, \varphi : I = [a,b] \to \mathbb{R}$ stetig und $\varphi \geq 0$ auf I (oder $\varphi \leq 0$ auf I). Dann gibt es ein $\xi \in I$ mit*

$$\int_a^b f\varphi = f(\xi)\int_a^b \varphi.$$

Beweis: oBdA sei $\varphi \geq 0$ (sonst betrachte $-\varphi$). Wir definieren die stetige Funktion

$$F : I \to \mathbb{R}, \quad F(x) = \int_a^b \big(f(y) - f(x)\big)\varphi(y)\,dy.$$

Ist $x \in I$ Minimalstelle (Maximalstelle) von f, so folgt $F(x) \geq 0$ (bzw. $F(x) \leq 0$). Nach dem Zwischenwertsatz existiert ein $\xi \in I$ mit $F(\xi) = 0$. □

Satz 2.4 (Lagrangedarstellung des Restglieds) *Unter den Voraussetzungen von Satz 2.3 existiert ein $\xi \in [x_0, x]$ (bzw. $\xi \in [x, x_0]$ für $x \leq x_0$), so dass gilt:*

$$R^n_{x_0} f(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}. \tag{2.5}$$

Beweis: Sei $x \geq x_0$ (sonst analog). Dann ist $(x-y)^n \geq 0$ auf $[x_0, x]$. Nach dem MWS der Integralrechnung existiert ein $\xi \in [x_0, x]$ mit

$$R^n_{x_0} f(x) = f^{(n+1)}(\xi)\int_{x_0}^x \frac{(x-y)^n}{n!}\,dy = f^{(n+1)}(\xi)\,\frac{(x-x_0)^{n+1}}{(n+1)!}$$

□

Beispiel 2.1 *Betrachte für $x \in (-1,1)$ die Funktion*

$$\begin{aligned}
f(x) &= (1-x)^{-1/2}, & f(0) &= 1\\
f'(x) &= \tfrac{1}{2}(1-x)^{-3/2}, & f'(0) &= 1/2\\
f''(x) &= \tfrac{3}{4}(1-x)^{-5/2}, & f''(0) &= 3/4.
\end{aligned}$$

Das Taylorpolynom erster Ordnung bei $x_0 = 0$ lautet

$$P_0^1 f(x) = \sum_{k=0}^{1} \frac{f^{(k)}(0)}{k!} x^k = 1 + \frac{1}{2}x,$$

mit der Lagrange-Restglieddarstellung

$$R_0^1 f(x) = \frac{f''(\xi)}{2} x^2 = \frac{3}{8}(1-\xi)^{-5/2} x^2 \quad \textit{mit } \xi \in [0, x].$$

Als Anwendung erhalten wir für die relativistische Energie die Entwicklung

$$\begin{aligned}
E &= \frac{m_0 c^2}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \\
&= m_0 c^2 \, f(x) \quad (\textit{ wobei } x = \left(\frac{v}{c}\right)^2) \\
&= m_0 c^2 \, (1 + \frac{1}{2} x + \frac{f''(\xi)}{2} x^2) \\
&= m_0 c^2 + \frac{1}{2} m_0 v^2 + \Delta E.
\end{aligned}$$

Dabei ist der Term nullter Ordnung die Ruheenergie, der Term erster Ordnung die klassische kinetische Energie E_{kin} und für den relativistischen Korrekturterm ergibt sich aus der Restglieddarstellung die Abschätzung

$$\frac{\Delta E}{E_{\text{kin}}} = f''(\xi) x \leq f''(x) x < 0{,}008 \quad \textit{für } x \leq 0{,}01.$$

Bei Geschwindigkeiten $v \leq \frac{1}{10} c$ beträgt die relativistische Korrektur weniger als ein Prozent der klassischen kinetischen Energie.

Definition 2.4 *Für $f \in C^\infty(I)$ und $x_0 \in I$ heißt die Reihe*

$$P_{x_0} f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

Taylorreihe von f mit Entwicklungspunkt x_0.

Die Taylorreihe ist eine Potenzreihe, mit Entwicklungspunkt $x_0 \in \mathbb{R}$. Insbesondere hat sie einen Konvergenzradius $R \in [0, +\infty]$, vgl. Kap. 5, Satz 2.1:

$$\begin{array}{lcll}
|x - x_0| & < & R & \Rightarrow \text{ absolute Konvergenz} \\
|x - x_0| & > & R & \Rightarrow \text{ Divergenz}
\end{array}$$

Selbst wenn die Taylorreihe einen positiven Konvergenzradius hat, muss sie nicht notwendig gegen die gegebene Funktion konvergieren.

Beispiel 2.2 *Betrachte $f : \mathbb{R} \to \mathbb{R}$ mit*

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \textit{für } x > 0 \\ 0 & \textit{für } x \leq 0 \end{cases}$$

Dann gilt $f \in C^\infty(\mathbb{R})$ und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$, siehe Folgerung 2.6 in Kapitel 3. Damit sind alle Koeffizienten der Taylorreihe Null und diese konvergiert gegen die Nullfunktion, nicht gegen f.

Definition 2.5 *Eine Funktion $f : (a,b) \to \mathbb{R}$ heißt analytisch, wenn es zu jedem $x_0 \in (a,b)$ eine Umgebung $U_\delta(x_0) \subset (a,b)$ gibt, auf der die Funktion f als Potenzreihe mit Entwicklungspunkt x_0 darstellbar ist.*

Satz 2.5 *Seien $a_k \in \mathbb{R}$ für $k \in \mathbb{N}_0$. Für $f : I = (a,b) \to \mathbb{R}$ und $x_0 \in I$ sind folgende Aussagen äquivalent:*

(1) *f besitzt auf I die Potenzreihendarstellung*

$$f(x) = \sum_{k=0}^{\infty} a_k (x-x_0)^k \quad \textit{für alle } x \in I.$$

(2) *$f \in C^\infty(I)$, $a_k = \frac{f^{(k)}(x_0)}{k!}$ und für alle $x \in I$ gilt $R^n_{x_0} f(x) \to 0$ mit $n \to \infty$.*

Beweis: $(1) \Rightarrow (2)$: Nach Voraussetzung gilt für den Konvergenzradius der Potenzreihe (vgl. Satz 2.1, Kap. 5)

$$R \geq \max(|x_0 - a|, |x_0 - b|).$$

Nach Satz 2.5 in Kapitel 5 ist $f \in C^\infty(I)$ und die Ableitungen können durch gliedweise Differentiation berechnet werden. Also folgt

$$f^{(k)}(x_0) = \sum_{j=0}^{\infty} a_j \, k! \, \delta_{jk} = k! a_k.$$

Folglich ist die Potenzreihe die Taylorreihe von f mit Entwicklungspunkt x_0, und es folgt

$$R^n_{x_0} f(x) = f(x) - \sum_{k=0}^{n} a_k \, x^k \to 0 \quad \text{nach Voraussetzung.}$$

$(2) \Rightarrow (1)$: Nach Voraussetzung gilt für $n \in \mathbb{N}_0$

$$\begin{aligned} f(x) &= \sum_{k=0}^{n} a_k (x-x_0)^k + \underbrace{R^n_{x_0} f(x)}_{\to 0 \text{ mit } n \to \infty} \\ \Rightarrow \quad f(x) &= \sum_{k=0}^{\infty} a_k (x-x_0)^k. \end{aligned}$$

□

Beispiel 2.3 *Mit Hilfe der Restgliedabschätzung können wir erneut die Potenzreihendastellung der Funktionen* exp, cos *und* sin *beweisen. Wir betrachten zum Beispiel $f(x) = \cos x$. Es gilt dann*

$$f^{(n)}(x) = \begin{cases} (-1)^k \cos x & n = 2k \\ (-1)^k \sin x & n = 2k-1 \end{cases}.$$

Insbesondere

$$f^{(n)}(0) = \begin{cases} (-1)^k & n = 2k \\ 0 & n = 2k-1 \end{cases}.$$

Die Taylorreihe mit Entwicklungspunkt Null lautet also

$$P_0 f(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\, x^{2k}.$$

Für das Restglied folgt mit Lagrange

$$R_0^{2k} f(x) = \frac{(-1)^{k+1}\, \sin\xi}{(2k+1)!}\, x^{2k+1} \to 0 \quad \textit{mit } k \to \infty.$$

Wir kommen jetzt zur mehrdimensionalen Analysis und wollen die Taylorentwicklung verallgemeinern.

Satz 2.6 (Taylorentwicklung im $\mathbb{R}^n$) *Sei $\Omega \subset \mathbb{R}^n$ offen und konvex, und sei $f \in C^{k+1}(\Omega)$ für ein $k \in \mathbb{N}_0$. Dann gibt es zu x_0, $x \in \Omega$ ein $\xi = (1-\tau)x_0 + \tau x$ (für $\tau \in (0,1)$ geeignet), so dass mit $h = x - x_0$ gilt:*

$$f(x) = \sum_{j=0}^{k} \frac{(D^j f)(x_0)\,(h,\dots,h)}{j!} + \frac{D^{k+1} f(\xi)(x_0)\,(h,\dots,h)}{(k+1)!}. \tag{2.6}$$

Dabei setzen wir

$$D^j f(x)(\underbrace{h,\dots,h}_{j \text{ Einträge}}) := \sum_{i_1,\dots,i_j=1}^{n} (\partial_{i_1} \dots \partial_{i_j} f)(x)\, h_{i_1} \dots h_{i_j}.$$

Beweis: Betrachte die Funktion

$$\varphi : [0,1] \to \mathbb{R}, \quad \varphi(t) = f(x_0 + t\,h).$$

Wir zeigen durch Induktion für $j = 0, 1, \dots, k+1$

$$\varphi^{(j)} t = D^j f(x_0 + th)(h,\dots,h) \tag{$*$}$$

Für $j = 0$ gilt das per Definition, und für $j = 1$ folgt aus der Kettenregel

$$\varphi'(t) = Df(x_0 + th)h = \sum_{i=1}^{n} \partial_i f(x_0 + th)\, h_i.$$

Ist die Aussage für $j \geq 1$ gezeigt, so folgt

$$\begin{aligned}
\varphi^{(j+1)}(t) &= \frac{d}{dt} \sum_{i_1,\dots,i_j=1}^{n} (\partial_{i_1} \dots \partial_{i_j} f)(x_0 + th)\, h_{i_1} \dots h_{i_j} \\
\text{(Fall } i = 1) &= \sum_{i_1,\dots,i_j=1}^{n} \sum_{i=1}^{n} (\partial_i\, \partial_{i_1} \dots \partial_{i_j} f)(x_0 + th)\, h_i\, h_{i_1} \dots h_{i_j} \\
&= D^{j+1} f(x_0 + th)(h,\dots,h).
\end{aligned}$$

Es folgt insbesondere $\varphi \in C^{k+1}([0,1])$. Nach Satz 2.4 gibt es ein $\tau \in (0,1)$ mit

$$\begin{aligned}\varphi(1) &= \sum_{j=0}^{k} \frac{\varphi^{(j)}(0)}{j!} + \frac{\varphi^{(k+1)}(\tau)}{(k+1)!} \\ &\overset{(*)}{=} \sum_{j=0}^{k} \frac{D^j f(x_0)(h,\dots,h)}{j!} + \frac{D^{k+1} f(\xi)(h,\dots,h)}{(k+1)!}.\end{aligned}$$

□

Beispiel 2.4 *Wir berechnen die Taylorentwicklung erster Ordnung im Punkt $(x_0, y_0) = (1,1)$ der Funktion*

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

und zwar lautet das Taylorpolynom erster Ordnung

$$P^1_{(1,1)} f(x,y) = \frac{1}{2}(x-1) - \frac{1}{2}(y-1) = \frac{1}{2}(x-y).$$

Das Restglied lautet in Lagrangedarstellung mit Zwischenpunkt (ξ, η)

$$\begin{aligned}R^1_{(1,1)} f(x,y) &= \frac{1}{2!}\Big(-\frac{4\eta}{(\xi+\eta)^3}(x-1)^2 + \frac{2(\xi-\eta)}{(\xi+\eta)^3}(x-1)(y-1) \\ &\quad + \frac{2(\xi-\eta)}{(\xi+\eta)^3}(y-1)(x-1) + \frac{4\xi}{(\xi+\eta)^3}(y-1)^2\Big) \\ &= \frac{2}{(\xi+\eta)^3}\big(-\eta(x-1)^2 + (\xi-\eta)(x-1)(y-1) + \xi(y-1)^2\big).\end{aligned}$$

Abschätzung des Restglieds, falls $\max(|x-1|, |y-1|) \le \frac{1}{2}$): dann folgt $\xi, \eta \in \left[\frac{1}{2}, \frac{3}{2}\right]$, $\xi + \eta \ge 1$ und $|\xi - \eta| \le 1$. Einsetzen dieser Abschätzungen ergibt

$$|R^1_{(1,1)} f(x,y)| \le 4\big((x-1)^2 + (y-1)^2\big).$$

Folgerung 2.3 *Unter den Voraussetzungen von Satz 2.6 gilt in Multiindexnotation*

$$f(x) = \sum_{j=0}^{k} \sum_{|\alpha|=j} \frac{\partial^\alpha f(x_0)}{\alpha!} h^\alpha + \sum_{|\alpha|=k+1} \frac{\partial^\alpha f(\xi)}{\alpha!} h^\alpha.$$

BEWEIS: Sei $\alpha = (\alpha_1, \dots, \alpha_n)$ ein Multiindex mit totaler Differentiationsordnung $|\alpha| = j$. Die Zahl der Tupel $(i_1, \dots, i_j)$, in denen jede Nummer $\nu \in \{1, \dots, n\}$ genau α_ν-mal vorkommt, beträgt $\frac{j!}{(\alpha_1!)\cdot\ \cdot(\alpha_n!)}$. Deshalb folgt

$$\frac{1}{j!} \sum_{i_1,\dots,i_j=1}^{n} (\partial_{i_1} \dots \partial_{i_j} f)(x)\, h_{i_1} \dots h_{i_j} = \frac{1}{j!} \sum_{|\alpha|=j} \frac{j!}{(\alpha_1!)\cdot\ \cdot(\alpha_n!)} \partial^\alpha f(x)\, h^\alpha.$$

Die Behauptung folgt nun aus Satz 2.6. □

Definition 2.6 *Sei $f \in C^1(\Omega)$. Ein Punkt $x_0 \in \Omega$ mit $Df(x_0) = 0$ heißt kritischer Punkt von f.*

Definition 2.7 *Sei $f \in C^2(\Omega)$ mit $\Omega \subset \mathbb{R}^n$ offen und $x_0 \in \Omega$. Die zweite Ableitung von f im Punkt x_0 ist die Bilinearform*

$$D^2 f(x_0) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad D^2 f(x_0)(v,w) = \sum_{i,j=1}^{n} \partial_i \partial_j f(x_0)\, v_i\, w_j.$$

Bezüglich der Standardbasis besitzt $D^2 f(x_0)$ die $n \times n$-Matrix $\big(\partial_i \partial_j f(x_0)\big)_{1 \le i,j \le n}$; diese heißt Hessematrix von f an der Stelle x_0. Die zugehörige quadratische Form

$$q : \mathbb{R}^n \to \mathbb{R}, \quad q(v) = D^2 f(x_0)(v,v) = \sum_{i,j=1}^{n} \partial_i \partial_j f(x_0)\, v_i\, v_j$$

heißt Hesse-Form von f in x_0.

Bemerkung. Für $f \in C^2(\Omega)$ ist $D^2 f(x_0)$ eine symmetrische Bilinearform, denn nach Satz 1.4 (Schwarz) gilt $\partial_i \partial_j f(x_0) = \partial_j \partial_i f(x_0)$ und folglich

$$\begin{aligned} D^2 f(x_0)(v,w) &= \sum_{i,j=1}^{n} \partial_i \partial_j f(x_0) v_i w_j \\ &= \sum_{i,j=1}^{n} \partial_j \partial_i f(x_0) w_j v_i \\ &= D^2 f(x_0)(w,v). \end{aligned}$$

Anders ausgedrückt: *die Hessematrix ist symmetrisch.*

In einem kritischen Punkt bestimmt die Hesseform das lokale Verhalten der Funktion:

Lemma 2.5 *Sei $f \in C^2(\Omega)$ mit $\Omega \subset \mathbb{R}^n$ offen und $Df(x_0) = 0$. Dann gilt*

$$f(x_0 + h) = f(x_0) + \frac{1}{2} D^2 f(x_0)(h,h) + R(h), \qquad \textit{wobei } \lim_{h \to 0} \frac{R(h)}{|h|^2} = 0.$$

Beweis: Nach Satz 2.6 mit $k = 1$ gilt mit $\tau \in [0,1]$ wegen $Df(x_0) = 0$

$$\begin{aligned} f(x_0 + h) &= f(x_0) + \frac{1}{2} D^2 f(x_0 + \tau h)(h,h) \\ &= f(x_0) + \frac{1}{2} D^2 f(x_0)(h,h) + R(h). \end{aligned}$$

Dabei gilt

$$\frac{|R(h)|}{|h|^2} \le \frac{1}{2} |D^2 f(x_0 + th) - D^2 f(x_0)| \longrightarrow 0 \text{ mit } h \to 0.$$

Dabei wurde benutzt, dass mit Cauchy-Schwarz

$$\left| \sum_{i,j=1}^{n} B_{ij} h_i h_j \right| \le |B||h|^2.$$

□

Um also das lokale Verhalten einer Funktion bei einem kritischen Punkt zu verstehen, ist die Hesse-Form zu analysieren.

Definition 2.8 *Eine symmetrische Bilinearform* $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *(bzw. die zugehörige quadratische Form* $q(v) = b(v,v)$*) heißt*

$$\begin{array}{lll} \textit{positiv definit} & \Leftrightarrow & b(v,v) > 0 \quad \forall\, v \in \mathbb{R}^n \backslash \{0\}, \\ \textit{positiv semidefinit} & \Leftrightarrow & b(v,v) \geq 0 \quad \forall\, v \in \mathbb{R}^n \backslash \{0\}, \\ \textit{indefinit} & \Leftrightarrow & \exists\, v, w \in \mathbb{R}^n \textit{ mit } b(v,v) > 0,\, b(w,w) < 0. \end{array}$$

Analog für negativ (semi-)definit.

Satz 2.7 (Lokale Extrema) *Sei* $f \in C^2(\Omega)$*,* $\Omega \subset \mathbb{R}^n$ *offen und* $x_0 \in \Omega$*.*

(a) *Wenn* f *in* x_0 *ein lokales Minimum hat, so folgt* $Df(x_0) = 0$ *und* $D^2 f(x_0)$ *ist positiv semidefinit.*

(b) *Ist* $Df(x_0) = 0$ *und* $D^2 f(x_0)$ *(strikt) positiv definit, so hat* f *in* x_0 *ein lokales Minimum.*

Hinweis. Der Test (b) reicht *nicht* aus, um eine globale Extremaleigenschaft nachzuweisen.
Beweis:
(a) Für $h \in \mathbb{R}^n$ beliebig hat die Funktion $\varphi(t) = f(x_0 + th)$ bei $t = 0$ ein lokales Minimum.

$$\begin{array}{rclcll} \Rightarrow 0 & = & \varphi'(0) & = & Df(x_0)h & \forall h \in \mathbb{R}^n, \\ 0 & \leq & \varphi''(0) & = & D^2 f(x_0)(h,h) & \forall h \in \mathbb{R}^n. \end{array}$$

Zum Beweis von (b) benötigen wir folgendes

Lemma 2.6 *Sei* $(V, \langle\cdot\,,\cdot\rangle)$ *ein Euklidischer Vektorraum mit* $\dim V < \infty$*. Für eine symmetrische Bilinearform* $b : V \times V \to \mathbb{R}$ *setze* $\lambda := \inf_{|v|=1} b(v,v)$*, wobei* $|v| = \sqrt{\langle v, v\rangle}$*. Dann gibt es ein* $\xi \in V$ *mit* $|\xi| = 1$ *und* $b(\xi,\xi) = \lambda$*.*

Beweis: Durch Wahl einer Orthonormalbasis (Gram-Schmidt-Verfahren) können wir annehmen, dass $X = \mathbb{R}^n$ und $\langle\cdot\,,\cdot\rangle$ das Standardskalarprodukt ist. Die Funktion

$$\mathbb{R}^n \to \mathbb{R}, \quad x \mapsto b(x,x) = \sum_{i,j=1}^{n} b_{ij}\, x_i\, x_j$$

ist stetig, und $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ ist abgeschlossen und beschränkt. Also wird $\inf_{|x|=1} b(x,x)$ in einem Punkt $\xi \in S^{n-1}$ angenommen. $\square$

Beweis von Satz 2.7 (b) Sei $\lambda = \inf_{v=1} D^2 f(x_0)(v,v)$. Nach Lemma 2.6 gibt es ein $\xi \in \mathbb{R}^n$, $|\xi| = 1$, mit

$$\lambda = D^2 f(x_0)(\xi,\xi) > 0 \quad \text{(da } D^2 f(x_0) \text{ positiv definit).}$$

Für $h \in \mathbb{R}^n \backslash \{0\}$ folgt

$$D^2 f(x_0)(h,h) = |h|^2 D^2 f(x_0) \left(\frac{h}{|h|}, \frac{h}{|h|} \right) \geq \lambda\, |h|^2.$$

Also folgt aus Lemma 2.5

$$\begin{array}{rcl} f(x_0 + h) - f(x_0) & = & \frac{1}{2} D^2 f(x_0)(h,h) + R(h) \\ & \geq & \frac{\lambda}{2} |h|^2 \big(1 - \frac{2R(h)}{\lambda\, |h|^2}\big) \\ \text{(Lemma 2.5)} & \geq & \frac{\lambda}{4} |h|^2 \quad \text{für } h \text{ hinreichend klein.} \end{array}$$

□

Um die Hesseform (und damit die Funktion f in der Nähe eines kritischen Punkts) zu verstehen, ist folgendes Resultat aus der Linearen Algebra entscheidend:

Satz 2.8 (Hauptachsentransformation) *Sei $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ symmetrische Bilinearform. Dann existieren Zahlen $\lambda_1 \leq \ldots \leq \lambda_n$ und eine Orthonormalbasis $\{v_1, \ldots, v_n\}$ mit $b(v_k, v_1) = \lambda_k\, \delta_{kl}$.*

Beweis: Wir konstruieren die Basis $\{v_1, \ldots, v_n\}$ und die Zahlen $\lambda_1, \ldots, \lambda_n$ induktiv wie folgt: seien $v_1, \ldots, v_{k-1}$ und $\lambda_1, \ldots, \lambda_{k-1}$ schon gefunden, wobei $1 \leq k \leq n$. Für $k = 1$ ist also noch gar nichts gefunden. Sei V_{k-1} der von $v_1, \ldots, v_{k-1}$ aufgespannte Raum, bzw. $V_0 = \{0\}$. Definiere nun

$$\lambda_k = \inf\{b(v, v) : v \in V_{k-1}^\perp,\ |v| = 1\}. \tag{2.7}$$

Nach Lemma 2.6 (angewandt auf $V_{k-1}^\perp$) gibt es ein $v_k \in V_{k-1}^\perp$ mit $|v_k| = 1$ und

$$b(v_k, v_k) = \lambda_k. \tag{2.8}$$

Per Konstruktion gelten (2.7) und (2.8) für alle $k = 1, \ldots, n$. Insbesondere

$$V_{k-1} \subset V_k \quad \Rightarrow \quad V_{k-1}^\perp \supset V_k^\perp \quad \Rightarrow \quad \lambda_{k-1} \leq \lambda_k. \tag{2.9}$$

Betrachte nun für $k < l$ und $t \in \mathbb{R}$ den Vektor

$$\begin{aligned} v(t) &:= (\cos t)\, v_k + (\sin t)\, v_l \in V_{k-1}^\perp \quad \text{für alle } t, \\ |v(t)|^2 &= \cos^2 t\, |v_k|^2 + 2 \sin t \cos t\, \langle v_k, v_l \rangle + \sin^2 t\, |v_l|^2 = 1. \end{aligned}$$

Also ist $v(t)$ in der Infimumsbildung in (2.7) zugelassen. Da $v(0) = v_k$, hat die Funktion $t \mapsto b(v(t), v(t))$ bei $t = 0$ ein Minimum und es folgt

$$0 = \frac{d}{dt} b(v(t), v(t))|_{t=0} = 2b(v(0), v'(0)) = 2b(v_k, v_l).$$

Damit gilt $b(v_k, v_l) = \lambda_k\, \delta_{kl}$. □

Sei $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ symmetrische Bilinearform mit Eigenwerten $\lambda_1 \leq \ldots \leq \lambda_n$ und Eigenvektorbasis $\{v_1, \ldots, v_n\}$. Für $x = \xi_1 v_1 + \ldots + \xi_n v_n$ folgt:

$$b(x, x) = \lambda_1 \xi_1^2 + \ldots + \lambda_n \xi_n^2.$$

3 Variationsformeln und Kurvenintegrale

Sei $I = [a, b] \to \mathbb{R}$. Wir betrachten hier ein Variationsintegral (oder Funktional) $\mathcal{F}$, dass jeder Abbildung $c \in C^1(I, \Omega)$ (anschaulich: ein Weg) eine reelle Zahl $\mathcal{F}(c) \in \mathbb{R}$ zuordnet, und zwar wie folgt:

$$\mathcal{F} : C^1(I, \Omega) \to \mathbb{R}, \quad \mathcal{F}(c) = \int_a^b f\big(t, c(t), c'(t)\big)\, dt$$

Dabei ist $f : I \times \Omega \times \mathbb{R}^n \longrightarrow \mathbb{R}$, $f = f(t, x, v)$. In der Physik heißt die Funktion f Lagrangefunktion.

Beispiel 3.1 *Die Bogenlänge von $c \in C^1(I,\Omega)$ ist*

$$\mathcal{F}(c) = \int_a^b |c'(t)|\,dt, \quad f(t,x,v) = |v|.$$

Heuristische Begründung: Wähle eine Zerlegung $a = t_0 < \ldots < t_N = b$ von I. Dann sollte näherungsweise gelten

$$\begin{aligned} |c(t_i) - c(t_{i-1})| &= |\int_{t_{i-1}}^{t_i} c'(t)\,dt| \approx |c'(\tau_i)|\,\Delta t_i, \text{ für } \tau \in [t_{i-1}, t_i], \text{ und} \\ \sum_{i=1}^N |c(t_i) - c(t_{i-1})| &\approx \sum_{i=1}^N |c'(\tau_i)|\Delta t_i \approx \int_a^b |c'(t)|\,dt. \end{aligned}$$

Wir wollen die obige Formel für die Bogenlänge als Definition hinnehmen.

Beispiel 3.2 *Gewichtete Bogenlänge (Querfeldeinlauf)*

$$\mathcal{F}(c) = \int_a^b \omega\big(c(t)\big)\,|c'(t)|\,dt, \quad f(t,x,v) = \omega(x)|v|.$$

Beispiel 3.3 *Wirkungsintegral der klassischen Mechanik*
Hier beschreibt $c \in C^1(I,\Omega)$ die Bahn eines Teilchens der Masse m in einem Kraftfeld mit Potential $V : \Omega \to \mathbb{R}$. Es ist dann $\frac{m}{2}|\dot{c}(t)|^2$ die kinetische Energie und $V(c(t))$ die potentielle Energie des Teichens zur Zeit t. Das Wirkungsintegral lautet

$$\mathcal{F}(c) = \int_a^b \Big(\frac{m}{2}|\dot{c}(t)|^2 - V\big(c(t)\big)\Big)\,dt, \text{ also } f(t,x,v) = \frac{m}{2}|v|^2 - V(x).$$

Um allgemein zu untersuchen, wie $\mathcal{F}(c)$ vom Weg c abhängt, betrachten wir Variationen

$$c : (-\varepsilon_0, \varepsilon_0) \times I \to \mathbb{R}^n, \quad c = c(\varepsilon, t)$$

wobei $c(0,\,\cdot\,) = c$. Wir bezeichnen die Ableitung

$$\varphi : I \to \mathbb{R}^n,\ \varphi(t) = \frac{\partial c}{\partial \varepsilon}(0,t)$$

als zugehöriges Vektorfeld der Variation.

Lemma 3.1 (Erste Variation) *Sei $\Omega \subset \mathbb{R}^n$ offen, $I = [a,b] \subset \mathbb{R}^n$. $f : I \times \Omega \times \mathbb{R}^n \to \mathbb{R}$. Die Funktionen $f : I \times \Omega \times \mathbb{R}^n \to \mathbb{R}$ sowie $D_v f = (\frac{\partial f}{\partial v_1}, \ldots, \frac{\partial f}{\partial v_n}) : I \times \Omega \times \mathbb{R}^n \to \mathbb{R}^n$ seien stetig differenzierbar. Für $c \in C^2\big((-\varepsilon_0, \varepsilon_0) \times I, \Omega\big)$ ist dann die Funktion*

$$\phi(\varepsilon) = \mathcal{F}(c(\varepsilon,\cdot)) = \int_a^b f\Big(t, c(\varepsilon,t), \frac{\partial c}{\partial t}(\varepsilon,t)\Big)\,dt$$

differenzierbar, und zwar gilt mit $\varphi = \frac{\partial c}{\partial \varepsilon}(0,\cdot) : I \to \mathbb{R}^n$

$$\phi'(0) = \int_a^b \langle L_f(c), \varphi\rangle\,dt + \big[\langle D_v f(t,c,c'), \varphi\rangle\big]_{t=a}^{t=b}. \tag{3.1}$$

Dabei ist $\langle\cdot,\cdot\rangle$ das Standardskalarprodukt im $\mathbb{R}^n$ und $L_f(c) : I \to \mathbb{R}^n$ ist gegeben durch

$$L_f(c) = D_x f(t,c,c') - \frac{d}{dt}\big[D_v f\,(t,c,c')\big].$$

Beweis: Die Verkettung $(\varepsilon, t) \mapsto f\big(t, c(\varepsilon, t), \frac{\partial c}{\partial t}(\varepsilon, t)\big)$ ist nach Voraussetzung von der Klasse C^1. Nach Satz 1.1 kann unter dem Integralzeichen differenziert werden. Es folgt, wobei wir $(\dots) = (t, c(t), c'(t))$ abkürzen,

$$\phi'(0) = \int_a^b \left(\langle D_x f(\dots), \frac{\partial c}{\partial \varepsilon}(0,t) \rangle + \langle D_v f(\dots), \frac{\partial}{\partial \varepsilon}\frac{\partial c}{\partial t}(0,t) \rangle \right) dt.$$

Nun gilt nach Schwarz $\frac{\partial}{\partial \varepsilon}\frac{\partial c}{\partial t} = \frac{\partial}{\partial t}\frac{\partial c}{\partial \varepsilon}$. Also folgt

$$\phi'(0) = \int_a^b \big(\langle D_x f(t,c,c'), \varphi \rangle + \langle D_v f(t,c,c'), \varphi' \rangle \big)\, dt. \tag{3.2}$$

Durch partielle Integration des hinteren Terms ergibt sich

$$\phi'(0) = \int_a^b \langle \underbrace{D_x f(t,c,c') - \frac{d}{dt}\big[D_v f(t,c,c')\big]}_{= L_f(c)}, \varphi \rangle \, dt + \big[\langle D_v f(t,c,c'), \varphi \rangle\big]_{t=a}^{t=b}.$$

□

Lemma 3.2 *Sei $I = (a,b)$. Für $f \in C^0(I, \mathbb{R}^n)$ gelte*

$$\int_a^b \langle f, \varphi \rangle = 0 \quad \text{für alle } \varphi \in C_c^\infty(I, \mathbb{R}^n). \tag{3.3}$$

Dann ist f die Nullfunktion.

Beweis: Es gibt eine Funktion $\eta \in C^\infty(\mathbb{R})$ mit $\eta(s) = 0$ für $|s| \geq 1$, $\eta \geq 0$ und $\int_{\mathbb{R}} \eta = 1$, z. B.

$$\eta(s) = \begin{cases} a \exp \frac{1}{s^2 - 1} & \text{für } |s| < 1 \\ 0 & \text{sonst.} \end{cases}$$

Dabei sei $a > 0$ so gewählt, dass $\int_{\mathbb{R}} \eta = 1$ ist.

Wir zeigen die Aussage des Lemmas zunächst für $n = 1$. Angenommen, $\varepsilon := f(t_0) > 0$ für ein $t_0 \in I$. Dann gibt es ein $\delta > 0$ mit

$$f(t) \geq \frac{\varepsilon}{2} \quad \text{für } t \in [t_0 - \delta, t_0 + \delta] \subset I.$$

Betrachte die reskalierte Funktion

$$\eta_{t_0,\delta}(t) = \frac{1}{\delta}\, \eta \left(\frac{t - t_0}{\delta} \right).$$

Dann gilt $\eta_{t_0,\delta}(t) = 0$ für $|t - t_0| \geq \delta$ und somit folgt aus der Voraussetzung

$$0 = \int_a^b f(t)\, \eta_{t_0,\delta}(t)\, dt \geq \varepsilon \int_a^b \eta_{t_0,\delta}(t)\, dt = \varepsilon > 0,$$

ein Widerspruch. Sei nun $f : I \to \mathbb{R}^n$. Für $\varphi \in C_c^\infty(I)$ liefert die Voraussetzung

$$0 = \int_a^b \langle f, \varphi\, e_i \rangle = \int_a^b f_i \varphi.$$

Aus obigem folgt $f_i = 0$, und das Lemma ist bewiesen. □

Satz 3.1 (Euler-Lagrange-Gleichungen) *Sei $\mathcal{F}(c) = \int_a^b f(t,c,c')$ ein Variationsintegral mit $f \in C^2(I \times \Omega \times \mathbb{R}^n)$, $f = f(t,x,v)$. Sei $c \in C^2(I,\Omega)$ stationärer Punkt von $\mathcal{F}$, d. h.*

$$\frac{d}{d\varepsilon}\mathcal{F}(c+\varepsilon\varphi)|_{\varepsilon=0} = 0 \quad \text{für alle } \varphi \in C_c^\infty(I,\mathbb{R}^n).$$

Dann gelten die Euler-Lagrange-Gleichungen

$$L_f(c) = 0 \Leftrightarrow \frac{\partial f}{\partial x_i}(t,c,c') - \frac{d}{dt}\frac{\partial f}{\partial v_i}(t,c,c') = 0 \text{ für } i = 1,\ldots,n.$$

Bemerkung. In den Euler-Lagrange-Gleichungen kommen Ableitungen der gesuchten Funktion c vor, deshalb spricht man von Differentialgleichungen (zweiter Ordnung, da zweite Ableitungen von c auftreten).

Beweis: Die Randterme in 3.1 verschwinden, da $\varphi(a) = \varphi(b) = 0$ nach Voraussetzung. Die Aussage folgt dann aus Lemma 3.1 und 3.2. □

Beispiel 3.4 (Gewichtete Bogenlänge eines Graphen) *Sei $\omega : \mathbb{R}^2 \to \mathbb{R}$ eine Gewichtsfunktion wie in Beispiel 3.2. Wir betrachten Wege $c(t) = \big(t, u(t)\big)$, die als Graph einer Funktion $u : I = [a,b] \longrightarrow \mathbb{R}$ mit $u(a) = \alpha$, $u(b) = \beta$ gegeben sind. Damit erhalten wir das Funktional*

$$\mathcal{F}(u) := \int_a^b \omega\big(c(t)\big)\,|c'(t)|\,dt = \int_a^b \omega(t,u(t))\,\big(1+|u'(t)|^2\big)^{1/2}\,dt.$$

Wir berechnen

$$\begin{aligned} f(t,x,v) &= \omega(t,x)\,(1+|v|^2)^{1/2} \\ \frac{\partial f}{\partial x}(t,x,v) &= \frac{\partial \omega}{\partial x}(t,x)\,(1+|v|^2)^{1/2} \\ \frac{\partial f}{\partial v}(t,x,v) &= \omega(t,x)\,\frac{v}{(1+|v|^2)^{1/2}} \end{aligned}$$

unf erhalten die Euler-Lagrange-Gleichung $L_f(u) = 0$

$$\frac{\partial \omega}{\partial x}(t,u)\sqrt{1+|u'|^2} - \frac{d}{dt}\left(\omega(t,u)\,\frac{u'}{\sqrt{1+(u')^2}}\right) = 0.$$

Für $\omega \equiv 1$ ergibt sich insbesondere

$$\frac{u'}{\sqrt{1+(u')^2}} = c \quad \Rightarrow \quad u' = \frac{c}{\sqrt{1-c^2}} \quad (c = \textit{Konstante}).$$

Also ist u linear, die Verbindung eine Gerade.

Beispiel 3.5 *Bewegung eines Massenpunkts in einem konservativen Kraftfeld:*

$$\mathcal{F}(c) = \int_a^b \Big(\frac{m}{2}|\dot{c}^2| - V\big(c(t)\big)\Big)\,dt, \quad f(t,x,v) = \frac{m}{2}\,|v|^2 - V(x)$$

Sei $F(x) = -\operatorname{grad} V(x)$ das zugehörige Kraftfeld. Dann ergibt sich

$$\frac{\partial f}{\partial x_i}(t,x,v) = F_i(x), \quad \frac{\partial f}{\partial v_i}(t,x,v) = m\,v_i.$$

Die Euler-Lagrange-Gleichungen lauten somit.

$$m\ddot{c} = F(c).$$

Dies sind die Newtonschen Bewegungsgleichungen.

Definition 3.1 *Sei $\Omega \subset \mathbb{R}^n$ offen und $F \in C^0(\Omega, \mathbb{R}^n)$ ein Vektorfeld auf Ω. Ist $c : I = [a, b] \to \Omega$ stückweise C^1, so heißt*

$$\int_c F \cdot ds = \int_a^b \langle F\big(c(t)\big), \dot{c}(t)\rangle\, dt$$

das Kurvenintegral von F längs c.

Beispiel 3.6 *Ist $\Omega \subset \mathbb{R}^3$ und $F : \Omega \to \mathbb{R}^3$ ein Kraftfeld, so ist $\int_c F \cdot ds$ die bei Transport einer „Probemasse " längs c verrichtete Arbeit.*

Beispiel 3.7 (Winkelfeld) *Sei $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ die Drehung um $\pi/2$ im $\mathbb{R}^2$. Betrachte das Vektorfeld*

$$W : \mathbb{R}^2 \backslash \{0\} \to \mathbb{R}^2, \quad W(z) = \frac{Jz}{|z|^2} = \begin{pmatrix} -\frac{y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \end{pmatrix}.$$

Ein Weg $c : I = [a, b] \to \mathbb{R}^2 \backslash \{0\}$ sei in Polardarstellung gegeben, d. h.

$$c(t) = r(t)\begin{pmatrix} \cos\vartheta(t) \\ \sin\vartheta(t) \end{pmatrix} \quad \text{mit } r, \vartheta \in C^1(I), r > 0.$$

Wir berechnen das Kurvenintegral

$$\begin{aligned} \int_c W \cdot ds &= \int_a^b \left\langle \begin{pmatrix} -\frac{1}{r}\sin\vartheta \\ \frac{1}{r}\cos\vartheta \end{pmatrix}, r' \begin{pmatrix} \cos\vartheta \\ \sin\vartheta \end{pmatrix} + r\,\vartheta' \begin{pmatrix} -\sin\vartheta \\ \cos\vartheta \end{pmatrix} \right\rangle dt \\ &= \int_a^b \vartheta' \,(\underbrace{\sin^2\vartheta + \cos^2\vartheta}_{=1})\, dt \end{aligned}$$

Es folgt

$$\vartheta(b) - \vartheta(a) = \int_c W \cdot ds \quad (\textit{Winkeldifferenz} = \textit{Kurvenintegral von } W). \tag{3.4}$$

Lemma 3.3 (Eigenschaften des Kurvenintegrals)

(1) Linearität: $\int_c (\lambda_1 F_1 + \lambda_2 F_2) \cdot ds = \lambda_1 \int_c F_1 \cdot ds + \lambda_2 \int_c F_2 \cdot ds$

(2) Additivität gegenüber Zerlegungen: *Sei $c : I = [a, b] \to \Omega \subset \mathbb{R}^n$ stückweise C^1 und $c_i = c\,|_{[t_{i-1}, t_i]}$, wobei $a = t_0 < \ldots < t_N = b$ Zerlegung von I ist. Dann gilt*

$$\int_c F \cdot ds = \sum_{i=1}^N \int_{c_i} F \cdot ds$$

(3) Invarianz bei Umparametrisierungen: *Sei* $c : J \to \Omega \subset \mathbb{R}^n$ *stückweise* C^1. *Ist* $\varphi \in C^1(I, J)$ *bijektiv und* $\varphi' > 0$, *so gilt*

$$\int_{c\circ\varphi} F \cdot ds = \int_c F \cdot ds.$$

(4) Abschätzung:

$$\Big| \int_c F \cdot ds \Big| \leq (\sup_{t\in I} |F \circ c|) \cdot \text{Länge}(c).$$

Beweis: (1) und (2) folgen aus der Definition und Standardeigenschaften (Kap. 4, Satz 2.3).

$$\begin{aligned}
\int_c F \cdot ds &= \int_{\alpha=\varphi(a)}^{\beta=\varphi(b)} \langle F\big(c(t)\big), c'(t)\rangle \, dt \\
&= \int_a^b \langle F\big(c(\varphi(\tau))\big), c'\big(\varphi(\tau)\big)\rangle \varphi'(\tau) \, d\tau \\
&= \int_a^b \langle F\big((c\circ\varphi)(\tau)\big), (c\circ\varphi)'(\tau)\rangle \, dt \\
&= \int_{c\circ\varphi} F \cdot ds.
\end{aligned}$$

Schließlich folgt (4) aus

$$\begin{aligned}
\left| \int_c F \cdot ds \right| &= \left| \int_a^b \langle F\big(c(t)\big), c'(t)\rangle \, dt \right| \\
&\leq \int_a^b |\langle F\big(c(t)\big), c'\big(t)\big)\rangle| dt \quad \text{(Kap. 4, Gl. (1.5))} \\
&\leq \int_a^b \big|F\big(c(t)\big)\big| \, |c'(t)| \, dt \quad \text{(Cauchy-Schwarz)} \\
&\leq \big(\sup_{t\in I} \big|F\big(c(t)\big)\big|\big) \underbrace{\int_a^b |c'(t)| \, dt}_{=\text{Länge}(c)} \quad \text{(Kap. 4, Gl. (1.4))},
\end{aligned}$$

vgl. Beispiel 3.1. □

Definition 3.2 *Sei* $\Omega \subset \mathbb{R}^n$ *offen. Ein Vektorfeld* $F \in C^0(\Omega, \mathbb{R}^n)$ *heißt Gradientenfeld (Physik: konservativ), wenn es eine Funktion* $\varphi \in C^1(\Omega)$ *gibt mit* $\operatorname{grad} \varphi = F$. *Die Funktion* φ *heißt Stammfunktion oder Potential von* F.

Bemerkung. Ist Ω zusammenhängend, so ist eine Stammfunktion bis auf eine additive Konstante eindeutig bestimmt:

$$\begin{aligned}
\operatorname{grad} \varphi_1 = \operatorname{grad} \varphi_2 \quad &\Rightarrow \quad \operatorname{grad}(\varphi_1 - \varphi_2) = 0 \\
\text{(Satz 2.1)} \quad &\Rightarrow \quad \varphi_1 - \varphi_2 \quad \text{ist konstant.}
\end{aligned}$$

Der Begriff „konservativ" bezieht sich auf den Energieerhaltungssatz: bei Transport einer Probemasse in einem konservativen Kraftfeld entspricht der Gewinn an Lagenergie der aufgewendeten Arbeit, unabhängig vom gewählten Weg.

Satz 3.2 (Wegunabhängigkeit des Kurvenintegrals) *Sei $\Omega \subset \mathbb{R}^n$ offen und zusammenhängend. Für ein Vektorfeld $F \in C^0(\Omega, \mathbb{R}^n)$ sind folgende Aussagen äquivalent:*

(1) *F ist ein Gradientenfeld.*

(2) *Für jeden geschlossenen Weg $c : I \to \Omega$, c stückweise C^1, ist $\int_c F \cdot ds = 0$.*

(3) *Für je zwei stückweise C^1-Wege $c_0, c_1 : I \to \Omega$ mit $c_0(a) = c_1(a)$, $c_0(b) = c_1(b)$ gilt*

$$\int_{c_0} F \cdot ds = \int_{c_1} F \cdot ds.$$

Beweis: (1) $\Rightarrow$ (2) : Sei $F = \operatorname{grad} \varphi$ mit $\varphi \in C^1(\Omega)$. Dann folgt

$$\begin{aligned} 0 &= \varphi\big(c(b)\big) - \varphi\big(c(a)\big) \\ &= \int_a^b \langle \operatorname{grad} \varphi\big(c(t)\big), c'(t) \rangle \, dt \quad \text{(Lemma 2.1)} \\ &= \int_c F \cdot ds. \end{aligned}$$

(2) $\Rightarrow$ (3) : Definiere den geschlossenen, stückweise C^1-Weg

$$c(t) = \begin{cases} c_0(t) & a \le t \le b \\ c_1(2b - t) & b \le t \le 2b - a \end{cases}.$$

Es folgt

$$0 = \int_c F \cdot ds = \int_{c_0} F \cdot ds - \int_{c_1} F \cdot ds \quad \big((2), (3) \text{ in Lemma 3.3}\big)$$

(3) $\Rightarrow$ (1) : Sei $x_0 \in \Omega$ fest. Für $x \in \Omega$ definiere

$$\varphi(x) := \int_{\gamma_x} F \cdot ds,$$

wobei $\gamma_x : [0,1] \to \Omega$ stetig und stückweise C^1 ist mit $\gamma_x(0) = x_0$ und $\gamma_x(1) = 1$. Zum Beispiel können wir nach Lemma 2.2 für γ_x einen Polygonzug wählen. Nach Voraussetzung hängt $\varphi(x)$ nicht von der Wahl des Wegs γ_x ab. Also ist die Funktion $\varphi : \Omega \to \mathbb{R}$ wohldefiniert.

Sei nun $x \in \Omega$ und $r > 0$ mit $B_r(x) \subset \Omega$. Für $h \in \mathbb{R}^n$ mit $|h| < r$ ist $c(t) = x + th \in B_r(x)$ für $t \in [0,1]$. Es folgt aus Lemma 3.3 (2)

$$\begin{aligned} \frac{\big|\varphi(x+h) - \big(\varphi(x) + \langle F(x), h \rangle\big)\big|}{|h|} &= \frac{1}{|h|} \Big| \int_c F \cdot ds - \langle F(x), h \rangle \Big| \\ &= \frac{1}{|h|} \Big| \int_0^1 \langle F(x+th) - F(x), h \rangle \, dt \Big| \\ &\le \sup_{|y-x| \le |h|} |F(y) - F(x)| \to 0 \quad \text{mit } h \to 0. \end{aligned}$$

Also gilt $D\varphi(x)\, h = \langle F(x), h \rangle$ bzw. $\operatorname{grad} \varphi(x) = F(x)$. □

Frage. Wie können wir entscheiden, ob ein gegebenes Vektorfeld $F : \Omega \to \mathbb{R}^n$ ein Gradientenfeld ist?

Satz 3.3 (notwendige Bedingung für Gradientenfelder) *Sei $\Omega \subset \mathbb{R}^n$ offen. Falls das Vektorfeld $F \in C^1(\Omega, \mathbb{R}^n)$ ein Gradientenfeld ist, so gilt:*

$$\partial_i F_j = \partial_j F_i \quad \text{auf } \Omega \quad \text{für } 1 \leq i, j \leq n.$$

Bemerkung. Für $n = 3$ lässt sich das schreiben als $\operatorname{rot} F = 0$, wobei

$$\operatorname{rot} F = (\partial_2 F_3 - \partial_3 F_2, \partial_3 F_1 - \partial_1 F_3, \partial_1 F_2 - \partial_2 F_1).$$

BEWEIS: Ist $F = \operatorname{grad} \varphi$, so gilt $\varphi \in C^2(\Omega)$ und aus dem Satz von Schwarz (Satz 1.4) folgt

$$\partial_i F_j = \partial_i \partial_j \varphi = \partial_j \partial_i \varphi = \partial_j F_i.$$

□

Beispiel 3.8 $F(x, y) = \frac{1}{2}(-y, x)$ *auf $\mathbb{R}^2$. Es gibt keine Stammfunktion, denn*

$$\partial_1 F_2 = 1, \ \text{aber } \partial_2 F_1 = -1.$$

Beispiel 3.9 *Betrachte erneut das Winkelfeld $W(z) = (-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2})$, mit den Ableitungen*

$$\begin{aligned}
\partial_1 W_2 &= \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}, \\
\partial_2 W_1 &= -\frac{x^2+y^2-2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}.
\end{aligned}$$

Die notwendige Bedingung gilt, aber dennoch ist das Kurvenintegral nicht wegunabhängig. betrachte zum Beispiel den geschlossenen Weg

$$c(t) = r(\cos kt, \sin kt), \quad t \in [0, 2\pi], k \in \mathbb{Z}.$$

Der Weg c ist in Polardarstellung gegeben, und zwar ist $r(t) \equiv r > 0$ konstant und $\vartheta(t) = kt$ für $0 \leq t \leq 2\pi$. Aus Gleichung 3.4 folgt

$$\int_c W \cdot ds = \vartheta(2\pi) - \vartheta(0) = 2\pi k \quad (\neq 0 \text{ für } k \in \mathbb{Z}\backslash\{0\}). \tag{3.5}$$

Das Integral ist also tatsächlich nicht wegunabhängig. Die notwendige Bedingung ist nicht hinreichend.

Definition 3.3 *Eine Variation $c \in C^0([0,1] \times I, \Omega)$ von Wegen $c(\varepsilon, \cdot) : I = [a, b] \to \Omega$ heißt*

a) *Homotopie mit festen Endpunkten, falls $c(\varepsilon, a) = p$ und $c(\varepsilon, b) = q$ für alle $\varepsilon \in [0, 1]$.*

b) *Homotopie von geschlossenen Wegen, falls $c(\varepsilon, a) = c(\varepsilon, b)$ für alle $\varepsilon \in [0, 1]$.*

Satz 3.4 (Homotopieinvarianz des Kurvenintegrals) *Sei $\Omega \subset \mathbb{R}^n$ offen und $F \in C^1(\Omega, \mathbb{R}^n)$. Dann sind folgende Aussagen äquivalent:*

(1) $\partial_i F_j = \partial_j F_i$ *auf Ω für alle $1 \leq i, j \leq n$.*

(2) *Das Kurvenintegral ist homotopieinvariant, d. h. sind $c_0, c_1 : I \to \Omega$ stückweise C^1 und gibt es eine Homotopie $c \in C^0([0,1] \times I, \Omega)$ zwischen $c_0 = c(0, \cdot\,)$ und $c_1 = c(1, \cdot\,)$ mit festen Endpunkten (bzw. geschlossen), so gilt*

$$\int_{c_0} F \cdot ds = \int_{c_1} F \cdot ds.$$

Beweis: (1) $\Rightarrow$ (2): Wir zeigen die Aussage nur unter der stärkeren Voraussetzung $c \in C^2([0,1] \times I, \Omega)$, ansonsten siehe Skript Analysis II, SS 97. Es gilt

$$\int_{c_\varepsilon} F \cdot ds = \int_a^b f(t, c_\varepsilon(t), c'_\varepsilon(t))\, dt \qquad (c_\varepsilon = c(\varepsilon, \cdot\,), \quad ' = \frac{\partial}{\partial t})$$

mit $f(t,x,v) = \langle F(x), v\rangle$. Für einen beliebigen Weg γ gilt, wenn $L_f(\gamma)$ den zugehörigen Euler-Lagrange-Operator bezeichnet,

$$\begin{aligned}
\big(L_f(\gamma)\big)_i &= \frac{\partial f}{\partial x_i}(t,\gamma,\gamma') - \frac{d}{dt}\frac{\partial f}{\partial v_i}(t,\gamma,\gamma') \\
&= \langle \partial_i F(\gamma), \gamma'\rangle - \frac{d}{dt} F_i(\gamma) \\
&= \sum_{j=1}^n \big(\partial_i F_j(\gamma) - \partial_j F_i(\gamma)\big)\, \gamma'_j \\
&= 0 \text{ nach Voraussetzung.}
\end{aligned}$$

Aus Lemma 3.1 folgt mit $\varphi = \frac{\partial c}{\partial \varepsilon}(\varepsilon, \cdot\,)$

$$\frac{\partial}{\partial \varepsilon} \int_{c_\varepsilon} F \cdot ds = \int_a^b \langle \underbrace{L_f(c_\varepsilon)}_{=0}, \varphi\rangle + \Big[\sum_{i=1}^n \underbrace{\frac{\partial f}{\partial v_i}(t, c_\varepsilon, c'_\varepsilon)}_{F_i(c_\varepsilon)}\ \varphi_i\Big]_{t=a}^{t=b} = 0.$$

Der Randterm verschwindet, da für eine Variation mit festen Endpunkten $\varphi(a) = \varphi(b) = 0$ und für eine Variation von geschlossenen Wegen $c(\varepsilon, a) = c(\varepsilon, b)$ und $\varphi(a) = \varphi(b)$.
Um die Rückrichtung zu beweisen, formulieren wir ein allgemeineres Resultat.

Definition 3.4 *Die offene Menge $\Omega \subset \mathbb{R}^n$ heißt sternförmig (bzgl. $x_0 \in \Omega$), falls gilt:*

$$(1-\varepsilon)x_0 + \varepsilon x \in \Omega \quad \textit{für alle } \varepsilon \in [0,1], x \in \Omega.$$

Beispiel 3.10 $\mathbb{R}^2 \backslash \mathbb{R}_0^+$ *ist sternförmig bzgl.* $(-1, 0)$.

Satz 3.5 (Lemma von Poincaré) *Sei $\Omega \subset \mathbb{R}^n$ offen und sternförmig. Für $F \in C^1(\Omega, \mathbb{R}^n)$ ist äquivalent:*

(1) $\partial_i F_j = \partial_j F_i$ *auf Ω für alle* $1 \le i, j \le n$,
(2) *F ist ein Gradientenfeld.*

Beweis: (2) $\Rightarrow$ (1) folgt aus Satz 3.3. Sei $c : I \to \Omega$ beliebiger geschlossener, stückweise C^1-Weg. Ist Ω sternförmig bzgl. x_0, so ist

$$c(\varepsilon, t) = (1-\varepsilon)\, x_0 + \varepsilon\, c(t) \quad (0 \le \varepsilon \le 1)$$

geschlossene Homotopie zwischen c und $c(0,t) \equiv x_0$. Es folgt aus Satz 3.4, (1) $\Rightarrow$ (2)

$$\int_c F \cdot ds = \int_{c(0,\,\cdot\,)} F \cdot ds = 0 \quad (\text{da } c(0,\,\cdot\,) \text{ konstant}).$$

Aus Satz 3.2 folgt die Behauptung. □

Beweis von (2) $\Rightarrow$ (1) in Satz 3.4:
Zu $x \in \Omega$ wähle $B_r(x) \subset \Omega$. Da $B_r(x)$ sternförmig, ist $\int_c F \cdot ds = 0$ für jeden geschlossenen, stückweise C^1-Weg $c : I \to B_r(x)$. Nach Satz 3.2 ist $F|_{B_r(x)}$ ein Gradientenfeld, also $\partial_i F_j = \partial_j F_i$ auf $B_r(x)$ nach Satz 3.3. □

Zusammenfassung:

$$\begin{array}{ccc}
F \text{ Gradientenfeld} & \Leftrightarrow & \int F \cdot ds \text{ wegunabhängig} \\
\Downarrow & \Uparrow \text{ falls } \Omega \text{ sternförmig } \Uparrow & \Downarrow \\
\partial_i F_j = \partial_j F_i & \Leftrightarrow & \int F \cdot ds \text{ homotopie invariant}
\end{array}$$

Fundamentalsatz der Algebra (Kap. 2, Satz 1.2)
Jedes komplexe Polynom vom Grad $n \geq 1$ hat mindestens eine Nullstelle $z_0 \in \mathbb{C}$.
Beweis: (indirekt) Angenommen, das Polynom $p(z) = z^n + a_{n-1}\, z^{n-1} + \ldots + a_0$ hat keine Nullstelle auf $\mathbb{C}$. Setze $p_n(z) = z^n$ und $q(z) = p(z) - p_n(z)$. Da $\deg q \leq n-1$, gilt $q(z)/z^n \to 0$ mit $|z| \to \infty$. Also gibt es ein $R > 0$ mit

$$|q(z)| \leq \frac{1}{2}\, |z|^n \quad \text{für } |z| = R.$$

Betrachte nun den geschlossenen Weg

$$c : I = [0, 2\pi] \longrightarrow \mathbb{R}^2 \backslash \{0\}, \quad c(t) = p(R\, e^{it}).$$

Wir berechnen die Windungszahl von c auf zwei Weisen:

a) Sei $c(\varepsilon, t) = p_n(R\, e^{it}) + \varepsilon\, q(R\, e^{it})$, $0 \leq \varepsilon \leq 1$.

$$\begin{aligned}
\Rightarrow c(0,t) &= R^n\, e^{int} = R^n(\cos nt, \sin nt) \\
c(1,t) &= p(R\, e^{it}) = c(t) \\
|c(\varepsilon, t)| &\geq R^n - |q(R\, e^{it})| \geq \frac{1}{2} R^n > 0.
\end{aligned}$$

Aus Satz 3.4 folgt, da $\partial_1 W_2 = \partial_2 W_1$ nach Beispiel 3.9

$$\int_c W \cdot ds = \int_{c(1,\,\cdot\,)} W \cdot ds = 2\pi n \quad (\text{Gleichung 3.5}).$$

b) Sei $\gamma(\varrho, t) = p(\varrho\, e^{it})$ für $0 \leq \varrho \leq R$. Da p nach Annahme keine Nullstelle hat, folgt wieder mit Satz 3.4

$$\int_c W \cdot ds = \int_{\gamma(R,\,\cdot\,)} W \cdot ds = \int_{\gamma(0,\,\cdot\,)} W \cdot ds = 0, \quad \text{da } \gamma(0,\,\cdot\,) \text{ konstant}.$$

Kombination von a) und b) liefert $2\pi n = 0$, im Widerspruch zur Annahme. □

Kapitel 8

Lokale Auflösung von Gleichungen

Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^m)$. Wir beschäftigen uns hier mit der Lösbarkeit einer nichtlinearen Gleichung $f(x) = y$ zu gegebenem $y \in \mathbb{R}^m$. Dabei nehmen wir an, dass eine Lösung der Gleichung $f(x_0) = y_0$ gegeben ist, und stellen uns folgende Fragen:

(1) Gibt es zu jedem y nahe bei y_0 eine Lösung x nahe bei x_0 von $f(x) = y$?

(2) Ist die Lösung eindeutig bestimmt?

(3) Falls nein, wie sieht die Lösungsmenge $f^{-1}\{y_0\}$ nahe bei x_0 aus?

Betrachte zunächst den Fall, dass f affin-linear ist, das heißt

$$f(x) = y_0 + A(x - x_0) \quad \text{mit } A \in L(\mathbb{R}^n, \mathbb{R}^m).$$

Es gilt dann

$$f(x) = y \quad \Leftrightarrow \quad A(x - x_0) = y - y_0.$$

Hierzu macht die Lineare Algebra folgende Aussagen:

(1) Es gibt eine Lösung für alle $y \quad \Leftrightarrow \quad \operatorname{rang} A = m$.

(2) Es gibt höchstens eine Lösung $\quad \Leftrightarrow \quad \ker A = \{0\} \quad \Leftrightarrow \quad \operatorname{rang} A = n$.

(3) $f^{-1}\{y_0\} = x_0 + \ker A$ ist ein affiner Unterraum der Dimension $n - \operatorname{rang} A$.

Sei nun $f \in C^1(\Omega, \mathbb{R}^m)$ beliebig. Dann definiere R_f durch die Entwicklung

$$f(x) = \underbrace{f(x_0)}_{=y_0} + \underbrace{Df(x_0)}_{=A}(x - x_0) + R_f(x - x_0).$$

Es folgt

$$f(x) = y \quad \Leftrightarrow \quad A(x - x_0) + R_f(x - x_0) = y - y_0.$$

Wir hoffen, dass sich die Aussagen (1), (2) und (3) geeignet übertragen lassen.

1 Diffeomorphismen

Definition 1.1 *Eine Abbildung* $f : U \to V$ *zwischen offenen Mengen* $U, V \subset \mathbb{R}^n$ *heißt Diffeomorphismus der Klasse* C^r *(wobei* $r \in \mathbb{N} \cup \{\infty\}$*), falls gilt:*

- f *ist bijektiv*
- f, f^{-1} *sind* r*-mal stetig differenzierbar.*

Beispiel 1.1 *Sei* $I = (a, b)$ *und* $f \in C^1(I)$ *mit* $f' > 0$ *(oder* $f' < 0$*). Dann ist* $f(I) =: J$ *ein offenes Intervall und* $f : I \to J$ *ein Diffeomorphismus der Klasse* C^1*.*

Dies ergibt sich aus Folgendem:

- Da f streng monoton wachsend und stetig, gilt $f(I) = (\alpha, \beta)$ mit $\alpha = \lim_{x \searrow a} f(x)$ und $\beta = \lim_{x \nearrow b} f(x)$ nach Kap.2, Satz 3.2. Außerdem ist $f : I \to J$ bijektiv.
- $g = f^{-1}$ ist differenzierbar mit $g' = \frac{1}{f' \circ g}$, insbesondere ist g von der Klasse C^1 (Kap. 3, Satz 1.4).

Umgekehrt: Ist $f : I \to J$ ein C^1-Diffeomorphismus zwischen offenen Intervallen mit Umkehrfunktion g, so folgt

$$g(f(x)) = x \quad \Rightarrow \quad g'(f(x))\, f'(x) = 1 \quad \Rightarrow \quad f'(x) \neq 0.$$

Nach dem Zwischenwertsatz ist also entweder $f' > 0$ oder $f' < 0$ (vgl. Bemerkung zu Satz 1.4. in Kap. 3). Zum Beispiel ist die Abbildung

$$f : (-1, 1) \to (-1, 1), \quad f(x) = x^3$$

zwar bijektiv (streng monoton wachsend) und von der Klasse C^1, aber sie ist kein C^1-Diffeomorphismus, denn $f'(0) = 0$. Die Umkehrabbildung

$$g : (-1, 1) \to (-1, 1), \quad g(y) = \begin{cases} \sqrt[3]{y} & \text{für } y \geq 0 \\ -\sqrt[3]{-y} & \text{für } y < 0 \end{cases}$$

ist im Punkt $y = 0$ nicht differenzierbar.

Beispiel 1.2 *Die Polarkoordinatenabbildung*
Sei $U = \{(r, \vartheta) \in \mathbb{R}^2 : r > 0,\, 0 < \vartheta < 2\pi\}$ *und* $V = \mathbb{R}^2 \backslash \{(x, 0) : x \geq 0\}$*. Dann ist die Abbildung*

$$f : U \to V, \quad f(r, \vartheta) = (r \cos \vartheta, r \sin \vartheta)$$

ein Diffeomorphismus der Klasse C^∞*. Überlegen Sie, dass die Umkehrabbildung* $g : V \to U$ *wie folgt gegeben ist:*

$$g(x, y) = \begin{cases} (\sqrt{x^2 + y^2},\, \arccos \frac{x}{\sqrt{x^2+y^2}}) & \textit{für } y \geq 0 \\ (\sqrt{x^2 + y^2},\, 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}}) & \textit{für } y < 0. \end{cases}$$

Beachte, dass für $x < 0$ *auch folgende Darstellung gilt:*

$$g(x, y) = (\sqrt{x^2 + y^2},\, \frac{\pi}{2} + \arccos \frac{y}{\sqrt{x^2 + y^2}}).$$

Somit ist in der Tat $g \in C^\infty(V, U)$*.*

Beispiel 1.3 (Inversion) *Die Inversion an der Standardsphäre* $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ *ist die Abbildung*

$$f : \mathbb{R}^2 \backslash \{0\} \longrightarrow \mathbb{R}^2 \backslash \{0\}, \quad f(x) = \frac{x}{|x|^2}.$$

Es gilt $f^{-1} = f \in C^\infty(\mathbb{R}^2\backslash\{0\}, \mathbb{R}^2\backslash\{0\})$, *und* f *bildet die beschränkte Menge* $U = \{x : 0 < |x| < 1\}$ *diffeomorph auf die unbeschränkte Menge* $V = \{y : |y| > 1\}$ *ab.*

Lemma 1.1 (Ableitung der Umkehrfunktion) *Seien* $U, V \subset \mathbb{R}^n$ *offen und* $f : U \to V$ *bijektiv mit Umkehrabbildung* $g : V \to U$. *Ist* f *in* $x_0 \in U$ *sowie* g *in* $y_0 = f(x_0) \in V$ *differenzierbar, so ist* $Df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^n)$ *invertierbar und* $Dg(y_0) = Df(x_0)^{-1}$.

Beweis: Aus $g\big(f(x)\big) = x$ und $f\big(g(y)\big) = y$ folgt mit der Kettenregel

$$Dg(y_0)\, Df(x_0) = \mathbb{1}, \quad Df(x_0)\, Dg(y_0) = \mathbb{1}.$$

□

Bemerkung. Bekanntlich heißt $\det Df(x_0)$ Jacobideterminante von f in x_0. In der Situation von Lemma 1.1 folgt aus dem Determinantenmultiplikationssatz

$$\det Dg(y_0)\, \det Df(x_0) = 1.$$

Lemma 1.2 (Höhere Ableitungen der Umkehrfunktion) *Seien* $U, V \subset \mathbb{R}^n$ *offen und* $f : U \to V$ *bijektiv. Ist* $f \in C^r(U, V)$ *für ein* $r \in \mathbb{N} \cup \{\infty\}$ *und ist die Umkehrabbildung* $g : V \to U$ *differenzierbar, so ist auch* $g \in C^r(V, U)$.

Beweis: Nach Lemma 1.1 gilt $Dg = (Df \circ g)^{-1}$, also nach der Cramerschen Regel

$$\partial_i g_j = (-1)^{i+j} \frac{M_{ij}\,(Df)}{\det(Df)} \circ g. \tag{1.1}$$

Dabei bezeichnet $M_{ij}(Df)$ die Determinante der Matrix, die aus Df durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Sei nun $f \in C^r(U, V)$ und induktiv $g \in C^{r-1}(U, V)$ schon gezeigt. Dabei gilt für $r = 0$ bereits $g \in C^0(U, V)$ wegen Satz 1.7. Nach Produktregel, Quotientenregel und Kettenregel ist dann die rechte Seite in (1.1) von der Klasse C^{r-1}, also $g \in C^r(V, U)$. □

Zum Nachweis der Existenz einer Lösung der Gleichung $f(x) = y$ verwenden wir folgendes grundlegende Resultat.

Satz 1.1 (Fixpunktsatz von Banach) *Sei* $(X, \|\cdot\|)$ *ein Banachraum,* $A \subset X$ *abgeschlossen und* $F : A \to A$ *eine Kontraktion, das heißt es gibt ein* $\theta \in [0, 1)$ *mit*

$$\|F(x) - F(y)\| \leq \theta\, \|x - y\| \quad \textit{für alle } x, y \in A. \tag{1.2}$$

Dann gibt es genau ein $x \in A$ *mit* $F(x) = x$.

Beweis: Betrachte die Rekursion $x_{n+1} = F(x_n)$ mit beliebigem Startwert $x_0 \in A$. Es folgt aus (1.2) für $n \geq 1$

$$\|x_{n+1} - x_n\| = \|F(x_n) - F(x_{n-1})\| \leq \theta \|x_n - x_{n-1}\|. \tag{1.3}$$

Wir können uns einen müder werdenden Frosch vorstellen, dessen Sprünge jedes Mal um ein Faktor $\theta \in [0,1)$ kürzer werden. Wie weit kann der Frosch insgesamt kommen? Es folgt per Induktion aus (1.3)

$$\|x_{n+1} - x_n\| \leq \theta^n \|x_1 - x_0\| \quad \text{für } n \in \mathbb{N}_0 \tag{1.4}$$

und hieraus weiter

$$\begin{aligned}
\|x_n - x_0\| &\leq \sum_{j=0}^{n-1} \|x_{j+1} - x_j\| \\
&\leq \sum_{j=0}^{n-1} \theta^j \|x_1 - x_0\| \\
&\leq \|x_1 - x_0\| \sum_{j=0}^{\infty} \theta^j \\
&= \frac{1}{1-\theta} \|x_1 - x_0\|.
\end{aligned}$$

Indem wir x_n als Startwert auffassen, folgt für $m > n$

$$\|x_m - x_n\| \leq \frac{1}{1-\theta} \|x_{n+1} - x_n\| \leq \frac{\theta^n}{1-\theta} \|x_1 - x_0\|. \tag{1.5}$$

Also ist $(x_n)_{n\in\mathbb{N}_0}$ eine Cauchyfolge. Da $(X, \|\cdot\|)$ vollständig, existiert $x = \lim_{n\to\infty} x_n$. Wegen $x_n \in A$ für alle n und A abgeschlossen folgt $x \in A$. Da F stetig (Lipschitz), folgt schließlich

$$F(x) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

Damit ist die Existenz des Fixpunkts bewiesen. Ist $y \in A$ ein weiterer Fixpunkt, so folgt

$$\|x - y\| = \|F(x) - F(y)\| \leq \theta \|x - y\|,$$

also $y = x$. Dies zeigt die Eindeutigkeit. $\square$

Bemerkung. Aus (1.5) folgt mit $m \to \infty$ die Fehlerabschätzung

$$\|x - x_n\| \leq \frac{\theta^n}{1-\theta} \|x_1 - x_0\|.$$

Satz 1.2 (über inverse Funktionen) *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^n)$. Ist $Df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^n)$ invertierbar, so gibt es eine offene Umgebung U von x_0, so dass gilt:*

- *$V = f(U)$ ist offene Umgebung von $y_0 = f(x_0)$*
- *$f : U \to V$ ist Diffeomorphismus der Klasse C^1.*

Zusatz. Ist $f \in C^r(\Omega, \mathbb{R}^n)$ für ein $r \in \mathbb{N} \cup \{\infty\}$, so ist $g = (f|_U)^{-1} \in C^r(V, \mathbb{R}^n)$.

Beweis: oBdA $x_0 = 0$, $y_0 = 0$, sonst betrachte

$$\widetilde{f} : \widetilde{\Omega} = \{\xi = x - x_0 : x \in \Omega\} \to \mathbb{R}^n, \quad \widetilde{f}(\xi) = f(x_0 + \xi) - f(x_0).$$

Schritt 1 *Formulierung als Fixpunktproblem*
Setze $A = Df(0)$ und $R_f(x) = f(x) - Ax$. Es gilt

$$f(x) = y \quad \Leftrightarrow \quad Ax + R_f(x) = y \quad \Leftrightarrow \quad x = A^{-1}\big(y - R_f(x)\big).$$

Für $y \in \mathbb{R}^n$ definiere

$$\phi_y : \Omega \to \mathbb{R}^n, \quad \phi_y(x) = A^{-1}\big(y - R_f(x)\big).$$

Damit gilt

$$f(x) = y \quad \Leftrightarrow \quad \phi_y(x) = x. \tag{1.6}$$

Schritt 2 *Konstruktion der Lösung (Inversen)*
Wir bestimmen $\varepsilon > 0$ und $\delta > 0$, so dass für jedes $y \in B_\varepsilon(0)$ die Abbildung $\phi_y : \overline{B_\delta(0)} \to \overline{B_\delta(0)}$ eine Kontraktion ist.

Wir setzen $\Lambda = |A^{-1}| \in (0, \infty)$. Da nach Voraussetzung $DR_f(x) = Df(x) - A$ stetig mit $DR_f(0) = 0$, gibt es ein $\delta_0 > 0$ mit

$$\overline{B_{\delta_0}(0)} \subset \Omega \quad \text{und } |DR_f(x)| \le \frac{1}{2\Lambda} \quad \text{für } |x| \le \delta_0.$$

Aus dem Schrankensatz (Satz 2.2 in Kap. 4) folgt

$$|x_1|, |x_2| \le \delta_0 \quad \Rightarrow \quad |R_f(x_1) - R_f(x_2)| \le \frac{1}{2\Lambda}\,|x_1 - x_2|. \tag{1.7}$$

Wir berechnen

$$\begin{aligned}
|\phi_y(x_1) - \phi_y(x_2)| &= \big|A^{-1}\big(y - R_f(x_1)\big) - A^{-1}\big(y - R_f(x_2)\big)\big| \\
&= \big|A^{-1}\big(R_f(x_1) - R_f(x_2)\big)\big| \\
&\le \Lambda\,|R_f(x_1) - R_f(x_2)|.
\end{aligned}$$

Also folgt aus (1.7) die Kontraktionseigenschaft

$$|x_1|, |x_2| \le \delta_0 \quad \Rightarrow \quad |\phi_y(x_1) - \phi_y(x_2)| \le \frac{1}{2}\,|x_1 - x_2|. \tag{1.8}$$

Weiter schätzen wir ab:

$$\begin{aligned}
|\phi_y(x)| &= \big|A^{-1}\big(y - R_f(x)\big)\big| \\
&\le |A^{-1}|\,\big(|y| + |R_f(x)|\big) \\
&= \Lambda\big(|y| + |R_f(x) - R_f(0)|\big) \quad (\text{da } R_f(0) = 0) \\
&\le \Lambda\,|y| + \frac{1}{2}\,|x| \quad \text{für } |x| \le \delta_0 \text{ nach (1.7).}
\end{aligned}$$

Also folgt für $\delta \le \delta_0$ und $\varepsilon = \delta/2\Lambda$

$$|x| \le \delta, \quad |y| < \varepsilon \quad \Rightarrow \quad |\phi_y(x)| < \frac{\delta}{2} + \frac{\delta}{2} = \delta. \tag{1.9}$$

Nach (1.9) und (1.7) ist $\phi_y : \overline{B_\delta(0)} \longrightarrow \overline{B_\delta(0)}$ eine Kontraktion.

Definiere die offenen Mengen $V = B_\varepsilon(0)$ und $U = f^{-1}(V) \cap B_\delta(0)$. Nach dem Banachschen Fixpunktsatz und (1.6) gibt es zu jedem $y \in V$ genau ein $x \in \overline{B_\delta(0)}$ mit $f(x) = y$. Tatsächlich ist $x \in B_\delta(0)$, denn nach (1.9) gilt $|x| = |\phi_y(x)| < \delta$. Somit ist $f : U \to V$ injektiv und surjektiv, also bijektiv.

Schritt 3 *Differenzierbarkeit der Inversen*
Sei $g : V \to U$ die Umkehrabbildung von $f : U \to V$.

a) *g ist stetig in* $y = 0$, denn $g(0) = 0$ und wie oben gezeigt gilt $|g(y)| = |\phi_y(g(y))| \leq \Lambda|y| + \frac{1}{2}|g(y)|$, also $|g(y)| \leq 2\Lambda|y| \to 0$ mit $y \to 0$.

b) Es gilt $Dg(0) = A^{-1}$, insbesondere ist g in $y = 0$ differenzierbar. Dazu brechnen wir

$$\begin{aligned} \frac{|g(y) - A^{-1}y|}{|y|} &= \frac{|\phi_y\big(g(y)\big) - A^{-1}y|}{|y|} \\ &= \frac{\big|A^{-1}\big(R_f\big(g(y)\big)\big)\big|}{|y|} \\ &\leq \Lambda \underbrace{\frac{\big|R_f\big(g(y)\big)\big|}{|g(y)|}}_{\to 0} \underbrace{\frac{|g(y)|}{|y|}}_{\leq 2\Lambda \text{ nach a)}} \\ &\to 0 \quad \text{mit } y \to 0. \end{aligned}$$

c) Für $\delta > 0$ hinreichend klein ist schliesslich g auf ganz V differenzierbar. Da $Df : \Omega \to L(\mathbb{R}^n, \mathbb{R}^n)$ und $\det : L(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$ stetig sind, können wir nämlich $\delta \leq \delta_0$ so klein wählen, dass $\det Df(x) \neq 0$ für alle $x \in B_\delta(0)$, also $Df(x)$ invertierbar. Ist dann $y \in V$, und $x = g(y) \in U$, so sind in x die Voraussetzungen des Satzes erfüllt. Nach dem schon Bewiesenen ist dann g in y differenzierbar.

Schritt 4 *Höhere Differenzierbarkeit*
Nach Lemma 1.2 ist $g \in C^1(V, U)$. Ist $f \in C^r(U, V)$ für ein $r \in \mathbb{N} \cup \{\infty\}$, so ist $g \in C^r(V, U)$ ebenfalls nach Lemma 1.2. $\square$

Als wichtige Konsequenz des Satzes halten wir fest:

Folgerung 1.1 *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^1(\Omega, \mathbb{R}^n)$. Ist $Df(x)$ invertierbar für alle $x \in \Omega$, so ist $f(\Omega) \subset \mathbb{R}^n$ offen.*

Beweis: Nach Satz 1.2 hat $y_0 = f(x_0)$ die offene Umgebung $V \subset f(\Omega)$. $\square$

2 Implizite Funktionen

Wir betrachten jetzt den Fall eines unterbestimmten Systems, wenn es also weniger Gleichungen gibt als Unbekannte. Wir können die Funktion dann wie folgt schreiben, indem wir die Variablen in zwei Gruppen einteilen:

$$f : \Omega \to \mathbb{R}^m, \quad f = f(x, y), \quad \text{wobei } (x, y) \in \Omega \subset \mathbb{R}^k \times \mathbb{R}^m = \mathbb{R}^n.$$

Gegeben sei eine Lösung (x_0, y_0) der Gleichung $f(x, y) = z_0$. Wir interessieren uns dafür, wie die Lösungsmenge dieser Gleichung nahe bei (x_0, y_0) aussieht. Können wir nach y auflösen, d. h. die Lösungsmenge als Graph einer Funktion $y = g(x)$ darstellen?

Beispiel 2.1 *Betrachte die Gleichung*

$$f(x, y) = x^2 + y^2 = 1 \quad \textit{für } (x, y) \in \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$$

Sei eine Lösung $(x_0, y_0) \in \mathbb{R}^2$ gegeben, also ein Punkt auf dem Einheitskreis. Ist $y_0 > 0$, so ist die Lösungsmenge nahe bei (x_0, y_0) Graph der Funktion $y = \sqrt{1 - x^2}$. Analog haben wir im Fall $y_0 < 0$ die lokale Graphendarstellung $y = -\sqrt{1 - x^2}$. Dagegen ist die Lösungsmenge in keiner Umgebung von $(1, 0)$ (und ebenso in keiner Umgebung von $(-1, 0)$) als Graph über der x-Achse darstellbar, denn es gibt die zwei Lösungen $y = \pm\sqrt{1 - x^2}$.

Es kann auch vorkommen, dass (x_0, y_0) der einzige Punkt mit $f(x_0, y_0) = z_0$ ist, z.B. bei der Gleichung $x^2 + y^2 = 0$.

Beispiel 2.2 *Sei $f : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$ linear. Wir unterteilen die $m \times (k + m)$-Matrix von f in eine $m \times k$-Matrix A und eine $m \times m$-Matrix B, d. h.*

$$f(x, y) = Ax + By \quad \textit{mit } A \in L(\mathbb{R}^k, \mathbb{R}^m), \quad B \in L(\mathbb{R}^m, \mathbb{R}^m).$$

Die Gleichung $Ax + By = z_0$ hat zu $x \in \mathbb{R}^k$ eine eindeutige Auflösung nach y dann und nur dann, wenn B invertierbar ist. Ist das der Fall, so lautet die Auflösung $y = B^{-1}(z_0 - Ax)$.

Allgemein schreiben wir die Jacobimatrix von $f = f(x, y)$ in der Form

$$Df(x, y) = (D_x f,\, D_y f) \in (\mathbb{R}^{m \times k},\, \mathbb{R}^{m \times m}).$$

Wenn wir nach $y = g(x)$ auflösen wollen, so sollte nach Beispiel 2.2 die Ableitung $D_y f(x_0, y_0) \in \mathbb{R}^{m \times m}$ invertierbar sein. In den Anwendungen ist die Einteilung in die beiden Variablengruppen nicht immer vorgegeben, das heißt es könnte nach verschiedenen Gruppen von je m Variablen aufgelöst werden. So kann der Einheitskreis in einer Umgebung von $(1, 0)$ zwar nicht als Graph $y = g(x)$ geschrieben werden, wohl aber als Graph $x = g(y)$.

Merkregel. Die Ableitung nach den Variablen, nach denen aufgelöst werden soll, muss invertierbar sein. Im (häufigen) Fall $m = 1$ bedeutet das $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$.

Satz 2.1 (über implizite Funktionen) *Sei $\Omega \subset \mathbb{R}^k \times \mathbb{R}^m$ offen, $f \in C^1(\Omega, \mathbb{R}^m)$ und sei $z_0 \in \mathbb{R}^m$. Für den gegebenen Punkt $(x_0, y_0) \in \Omega$ gelte:*

(a) $f(x_0, y_0) = z_0$

(b) $D_y f(x_0, y_0) \in \mathbb{R}^{m \times m}$ *ist invertierbar.*

Dann gibt es offene Umgebungen U von x_0, V von y_0 sowie eine Funktion $g \in C^1(U, V)$ mit

$$\{(x, y) \in U \times V : f(x, y) = z_0\} = \{\big(x, g(x)\big) : x \in U\}, \tag{2.10}$$

insbesondere ist $g(x_0) = y_0$. Die Funktion g hat die Ableitung

$$Dg(x_0) = -\big(D_y f(x_0, y_0)\big)^{-1} D_x f(x_0, y_0). \tag{2.11}$$

Zusatz. Für jedes $r \in \mathbb{N} \cup \{\infty\}$ gilt die Implikation

$$f \in C^r(\Omega, \mathbb{R}^m) \quad \Rightarrow \quad g \in C^r(U, \mathbb{R}^m).$$

BEWEIS: Wir verwenden einen Trick, um den Satz über inverse Funktionen anwenden zu können. Und zwar betrachten wir

$$F : \Omega \to \mathbb{R}^k \times \mathbb{R}^m, \quad F(x,y) = \big(x, f(x,y)\big).$$

Es gilt

$$DF = \begin{pmatrix} E_k & 0 \\ D_x f & D_y f \end{pmatrix} \in \begin{pmatrix} \mathbb{R}^{k\times k} & \mathbb{R}^{k\times m} \\ \mathbb{R}^{m\times k} & \mathbb{R}^{m\times m} \end{pmatrix}.$$

Nach Voraussetzung gilt

$$\det DF(x_0, y_0) = \det D_y f(x_0, y_0) \neq 0.$$

Nach dem Umkehrsatz existieren offene Umgebungen $U_0 \times V$ von (x_0, y_0) sowie W von (x_0, z_0), so dass $F : U_0 \times V \to W$ diffeomorph ist. Wir bezeichnen die zugehörige Umkehrabbildung mit $G \in C^1(W, U_0 \times V)$. Ist $(x,z) \in W$, so gilt also $(x,z) = (x, f(x,y))$ mit $(x,y) \in U_0 \times V$ nach Konstruktion, und es folgt

$$G(x,z) = G(x, f(x,y) = G(F(x,y)) = (x,y).$$

Also ist G von der Form $G(x,z) = (x, g_0(x,z))$ mit $g_0 \in C^1(W, \mathbb{R}^m)$.

Sei nun $U = \{x \in U_0 : (x, z_0) \in W\}$. Da W offen in $\mathbb{R}^k \times \mathbb{R}^m$, ist U offen in $\mathbb{R}^k$ und für $(x,y) \in U \times V$ gilt

$$\begin{aligned} f(x,y) = z_0 \quad \Leftrightarrow \quad F(x,y) &= (x, z_0) \\ \Leftrightarrow \quad (x,y) &= G(x, z_0) \quad (\text{da } (x, z_0) \in W) \\ \Leftrightarrow \quad y &= g_0(x, z_0). \end{aligned}$$

Also gilt die Aussage des Satzes mit $g(x) = g(x, z_0)$. Die Formel für die Ableitung folgt aus der Kettenregel:

$$f\big(x, g(x)\big) = z_0 \quad \Rightarrow \quad D_x f(x_0, y_0) + D_y f(x_0, y_0)\, Dg(x_0) = 0.$$

$\square$

Beispiel 2.3 *Wir wollen als triviales Beispiel untersuchen, wie die Nullstelle eines quadratischen Polynoms von seinen Koeffizienten abhängt. Sei*

$$f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, \quad f(p,q,\lambda) = \lambda^2 + 2p\,\lambda + q = (\lambda + p)^2 - (p^2 - q).$$

Die Menge $N = \{(p,q,\lambda) \in \mathbb{R}^2 \times \mathbb{R} : f(p,q,\lambda) = 0\}$ *ist die Vereinigung der beiden disjunkten Graphen* $G^\pm = \{(p,q,\lambda^\pm(p,q)) : p^2 > q\}$, *wobei* $\lambda^\pm(p,q) = -p \pm \sqrt{p^2 - q}$, *mit der Menge* $\overline{G}^+ \cap \overline{G}^- = \{(p,q,\lambda) : p^2 = q,\ \lambda = -p\}$.

Fall 1 $\frac{\partial f}{\partial \lambda}(p_0, q_0, \lambda_0) = 2(\lambda_0 + p_0) \neq 0 \quad \Leftrightarrow \quad p_0^2 - q_0 > 0.$
Dann liegt (p_0, q_0, λ_0) *in einem der beiden Graphen* G^+ *oder* G^-, *und* N *ist in einer*

Umgebung $U \times V$ als Graph von λ^+ oder λ^- darstellbar.

Fall 2 $\frac{\partial f}{\partial \lambda}(p_0, q_0, \lambda_0) = 2(\lambda_0 + p_0) \quad \Leftrightarrow \quad p_0^2 - q_0 = 0.$
Hier macht der implizite Funktionensatz keine Aussage. Tatsächlich lässt sich die Menge N in keiner Umgebung von (p_0, q_0, λ_0) als Graph $\lambda = \lambda(p, q)$ darstellen: für $p^2 < q$ hat die Gleichung überhaupt keine Lösung, für $p^2 = q$ genau eine und für $p^2 > q$ die zwei verschiedenen Lösungen $\lambda^{\pm}(p, q)$.

Beispiel 2.4 *Betrachte allgemeiner das Polynom*

$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(b, \lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_0.$$

Sei λ_0 einfache Nullstelle von $f(a, \lambda)$ für $a \in \mathbb{R}^n$, d.h. $f(a, \lambda) = (\lambda - \lambda_0)\, q(\lambda)$ für ein Polynom $q(\lambda)$ mit $q(\lambda_0) \neq 0$. Es folgt $\frac{\partial f}{\partial \lambda}(a, \lambda_0) = q(\lambda_0) \neq 0$. Nach dem Satz über implizite Funktionen existiert eine Umgebung $U \times V$ von (a, λ_0), so dass zu jedem $b \in U$ genau eine Nullstelle $\lambda(b) \in V$ von $f(b, \cdot)$ existiert. Diese hängt unendlich oft differenzierbar von b ab, und es gilt für $0 \leq i \leq n-1$

$$\frac{\partial \lambda}{\partial b_i}(a, \lambda_0) = -\frac{\lambda_0^i}{n\, \lambda_0^{n-1} + (n-1)\, a_{n-1}\, \lambda_0^{n-2} + \ldots + a_1}.$$

Wir kommen nun zu einer geometrischen Anwendung des Satzes über implizite Funktionen. Ist $f \in C^1(\mathbb{R}^2)$ und $z_0 \in \mathbb{R}$, so kann im allgemeinen die Niveaumenge

$$M = \{(x, y) \in \mathbb{R}^2 : f(x, y) = z_0\}$$

Punkte enthalten, in denen M nicht lokal wie eine Linie aussieht, z.B. Kreuzungspunkte von Linien, isolierte Punkte usw. Ist aber $Df(x, y) \neq 0$ für alle $(x, y) \in M$, so ist M nach dem impliziten Funktionensatz in der Nähe jedes Punkts als Graph über der x-Achse oder der y-Achse darstellbar und damit wirklich eine Höhenlinie im strengen Sinn des Worts. Die folgende Definition verallgemeinert diesen Sachverhalt auf beliebige Dimensionen.

Definition 2.1 *Sei $0 \leq k \leq n$. Eine Menge $M \subset \mathbb{R}^n$ heisst k-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ der Klasse C^r (wobei $r \in \mathbb{N} \cup \{\infty\}$), falls Folgendes gilt: zu jedem $p \in M$ gibt es eine offene Umgebung Ω und eine Funktion $f \in C^r(\Omega, \mathbb{R}^{n-k})$ (wobei wir $\mathbb{R}^0 := \{0\}$ setzen) mit $M \cap \Omega = f^{-1}\{0\}$ und $\operatorname{rang} Df(q) = n - k$ für alle $q \in \Omega$.*

Beispiel 2.5 *Die Ellipsoidoberfläche*

$$M = \{(x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\} \quad (a, b, c > 0).$$

Es gilt $M = f^{-1}\{0\}$ mit

$$f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 \quad \textit{und} \quad \operatorname{grad} f(x, y, z) = 2\left(\frac{x}{a^2}, \frac{y}{b^2}, \frac{z}{c^2}\right) \neq 0$$

für alle $(x, y, z) \in \Omega = \mathbb{R}^3 \backslash \{0\}$.

Wir benötigen folgende Notation. Sei $X \subset \mathbb{R}^n$ ein Unterraum mit orthogonalem Komplement $X^\perp$. Für $U \subset X$ und $V \subset X^\perp$ schreiben wir dann

$$U + V = \{x + y : x \in U,\, y \in V\}.$$

Im Fall $X = \mathbb{R}^k \times \{0\}$ ist also $U + V = U \times V$. Ferner setzen wir für $M \subset \mathbb{R}^n$ und $p \in \mathbb{R}^n$

$$M - p = \{q - p : q \in M\}.$$

Satz 2.2 (Untermannigfaltigkeiten = lokale Graphen) *Für $M \subset \mathbb{R}^n$ sind die folgenden Aussagen äquivalent:*

(1) *M ist eine k-dimensionale Untermannigfaltigkeit der Klasse C^r, das heißt zu $p \in M$ gibt es eine offene Umgebung $\Omega \subset \mathbb{R}^n$ und eine Funktion $f \in C^r(\Omega, \mathbb{R}^{n-k})$ mit $M \cap \Omega = f^{-1}\{0\}$ und $\operatorname{rang} Df = n - k$ auf Ω.*

(2) *Zu jedem $p \in M$ gibt es einen k-dimensionalen Unterraum $X \subset \mathbb{R}^n$, offene Teilmengen $U \subset X$, $V \subset X^\perp$ und eine Funktion $g \in C^r(U, V)$ mit $g(0) = 0$, so dass*

$$(M - p) \cap (U + V) = \{x + g(x) : x \in U\}.$$

BEWEIS:
Wir überlegen zunächst, dass sowohl (1) als auch (2) invariant unter Translationen und Drehungen des $\mathbb{R}^n$ sind. Betrachte dazu

$$\Phi : \mathbb{R}^n \to \mathbb{R}^n,\ \Phi(q) = Sq + a, \qquad \text{wobei } S \in \mathbb{O}(n) \text{ und } a \in \mathbb{R}^n.$$

Es gilt $\Phi^{-1}(q') = S^{-1}(q' - a)$ und $D\Phi^{-1}(q') = S^{-1}$ für alle $q' \in \mathbb{R}^n$. Sind die Daten unter (1) gegeben, so ist $\widetilde{\Omega} := \Phi(\Omega)$ offene Umgebung von $\widetilde{p} := \Phi(p) \in \Phi(M) =: \widetilde{M}$, und mit $\widetilde{f} := f \circ \Phi^{-1} : \widetilde{\Omega} \to \mathbb{R}^{n-k}$ gilt

$$\begin{aligned}
\widetilde{M} \cap \widetilde{\Omega} &= \Phi(M \cap \Omega) = \Phi(f^{-1}\{0\}) = \widetilde{f}^{-1}\{0\},\\
\operatorname{rang} D\widetilde{f} &= \operatorname{rang}\big(Df(\Phi^{-1})\, S^{-1}\big) = n - k \quad \text{auf } \widetilde{\Omega},\\
\ker D\widetilde{f}(\widetilde{p}) &= \ker\big(Df(p)\, S^{-1}\big) = S(\ker Df(p)).
\end{aligned}$$

Sind andererseits die Daten wie in (2) gegeben, so setzen wir $\widetilde{p} = \Phi(p) \in \Phi(M) =: \widetilde{M}$, $\widetilde{X} = SX$, $\widetilde{U} = SU$, $\widetilde{V} = SV$ und $\widetilde{g} = S \circ g \circ S^{-1}|_{\widetilde{U}}$. Dann ist $\widetilde{X}$ ein k-dimensionaler Unterraum, $\widetilde{U}$ offen in $\widetilde{X}$, $\widetilde{V}$ offen in $S(X^\perp) = \widetilde{X}^\perp$, $\widetilde{g} \in C^r(\widetilde{U}, \widetilde{V})$, $\widetilde{g}(0) = 0$ und schliesslich

$$\begin{aligned}
(\widetilde{M} - \widetilde{p}) \cap (\widetilde{U} + \widetilde{V}) &= S\big((M - p) \cap (U + V)\big)\\
&= S(\{x + g(x) : x \in U\})\\
&= \{\widetilde{x} + S \circ g \circ S^{-1}(\widetilde{x}) : \widetilde{x} \in \widetilde{U}\}\\
&= \{\widetilde{x} + \widetilde{g}(\widetilde{x}) : \widetilde{x} \in \widetilde{U}\}.
\end{aligned}$$

Nach diesen Vorüberlegungen ist der eigentliche Beweis einfach. Für (1) $\Rightarrow$ (2) können wir nämlich nach Translation um $-p$ und einer geeigneten Drehung S annehmen, dass $p = 0$ und $D_y f(0, 0)$ invertierbar, wobei $(x, y) \in \mathbb{R}^k \times \mathbb{R}^{n-k} = \mathbb{R}^n$. Aussage (2) folgt dann sofort aus dem Satz 2.1 über implizite Funktionen. Umgekehrt können wir für (2) $\Rightarrow$ (1) annehmen, dass wieder $p = 0$ und dass $X = \mathbb{R}^k \times 0$. Dann definieren wir $\Omega = U \times V$ und $f \in C^r(\Omega, \mathbb{R}^{n-k})$ durch $f(x,y) = y - g(x)$. Es folgt $M \cap \Omega = f^{-1}\{0\}$ sowie

$$Df(x,y) = (-Dg(x), E_{n-k}) \quad \Rightarrow \quad \operatorname{rang} Df = n-k.$$

$\square$

Definition 2.2 *Ein Vektor $v \in \mathbb{R}^n$ heisst Tangentialvektor der Untermannigfaltigkeit $M \subset \mathbb{R}^n$ im Punkt $p \in M$, falls es eine Abbildung $\gamma : (-\varepsilon, \varepsilon) \to M$ gibt mit $\gamma(0) = p$, $\gamma'(0) = v$. Die Menge der Tangentialvektoren von M im Punkt p wird mit T_pM bezeichnet.*

Folgerung 2.1 *Sei M eine k-dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ der Klasse C^1. Dann gilt $T_pM = \ker Df(p)$, falls $p \in M \cap \Omega = f^{-1}\{0\}$ mit $\operatorname{rang}(Df) = n-k$ auf Ω. Insbesondere ist T_pM ein k-dimensionaler Unterraum des $\mathbb{R}^n$.*

Beweis: Für $\gamma : (-\varepsilon, \varepsilon) \to M$ mit $\gamma(0) = p$ und $\gamma'(0) = v$ gilt

$$0 = \frac{d}{dt} f\big(\gamma(t)\big)|_{t=0} = Df\big(\gamma(0)\big)\gamma'(0) = Df(p)\,v \quad \Rightarrow \quad T_pM \subset \ker Df(p).$$

Nach Satz 2.2 gibt es andererseits einen k-dimensionalen Unterraum X, offene Teilmengen $U \subset X$, $V \subset X^\perp$ und eine Funktion $g \in C^1(U,V)$ mit $g(0) = 0$, so dass $(M-p) \cap (U+V) = \{x + g(x) : x \in U\}$. Sei $G : U \to \mathbb{R}^n$, $G(x) = x + g(x)$ und $\xi \in X$ beliebig. Dann gilt $p + G(t\xi) \in M$ für $t \in (-\varepsilon, \varepsilon)$ und folglich

$$DG(0)\,\xi = \frac{d}{dt}(p + G(t\xi))|_{t=0} \in T_pM \quad \Rightarrow \quad \operatorname{Bild} DG(0) \subset T_pM.$$

Aber $\dim \operatorname{Bild} DG(0) = k = \dim \ker Df(p)$, denn mit der Orthogonalprojektion $P_X : \mathbb{R}^n \to X$ ist $P_X \cdot DG(0) = D(P_X \circ G)(0) = \mathrm{Id}_X$. Daraus folgt die Behauptung. $\square$

Folgerung 2.2 (Extrema mit Nebenbedingungen) *Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega, \mathbb{R}^m)$ und $\varphi \in C^1(\Omega)$. Gilt dann für $p \in f^{-1}\{0\}$*

(1) *$\varphi(q) \geq \varphi(p)$ für alle $q \in \Omega$ mit $f(q) = 0$,*

(2) *$\operatorname{rang} Df(p) = m$,*

so gibt es $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ mit $\operatorname{grad} \varphi(p) = \sum_{i=1}^m \lambda_i \operatorname{grad} f_i(p)$.

Beweis: Nach evtl. Verkleinerung von Ω ist $\operatorname{rang} Df(p) = m$ auf Ω und $M := f^{-1}\{0\}$ ist k-dimensionale Untermannigfaltigkeit. Ist $\gamma : (-\varepsilon, \varepsilon) \to M$ mit $\gamma(0) = p$ und $\gamma'(0) = v$, so hat hat $\varphi \circ \gamma$ in $t = 0$ ein lokales Maximum und folglich

$$0 = \frac{d}{dt}\, \varphi\big(\gamma(t)\big)|_{t=0} = \langle \operatorname{grad} \varphi(p), v \rangle.$$

Also gilt $\operatorname{grad} \varphi(p) \in (T_pM)^\perp$. Da $f_j|_M \equiv 0$, gilt analog $\operatorname{grad} f_j(p) \in (T_pM)^\perp$ für $1 \leq j \leq m$. Aber $\dim (T_pM)^\perp = n - k = m$ nach Folgerung 2.1, und die Vektoren $\operatorname{grad} f_j(p)$ sind die Zeilenvektoren der Matrix $Df(p)$ mit Rang m. Wegen der Gleichheit von Zeilenrang und Spaltenrang ist $\{\operatorname{grad} f_j(p) : 1 \leq j \leq m\}$ eine Basis von $(T_{p_0}M)^\perp$, und die Behauptung folgt. $\square$

Beispiel 2.6 *Sei* $\mu = \inf\{\langle Ax, x\rangle : |x| = 1\}$ *für eine symmetrische Matrix* $A \in \mathbb{R}^{n\times n}$*. Zu minimieren ist also die Funktion* $\varphi(x) = \langle Ax, x\rangle$ *unter der Nebenbedingung* $f(x) = |x|^2 - 1 = 0$*. Da* $S^{n-1} = f^{-1}\{0\}$ *kompakt und* φ *stetig, wird das Infimum in einem Punkt* $x_0 \in S^{n-1}$ *angenommen. Nach Folgerung 2.2 gibt es ein* $\lambda \in \mathbb{R}$ *mit* $\operatorname{grad}\varphi(x_0) = \lambda \operatorname{grad} f(x_0)$*, also* $Ax_0 = \lambda x_0$*. Somit hat jede symmetrische Matrix mindestens einen Eigenvektor, vgl. Kapitel 7, Satz 2.8.*

Kapitel 9

Anfangswertprobleme für gewöhnliche Differentialgleichungen

Aus Zeitgründen konnte in der Vorlesung nur kurz auf gewöhnliche Differentialgleichungen eingegangen werden. Es wurden allerdings zentrale Aussagen (Picard-Lindelöf) behandelt, so dass Sie den Stoff ohne allzu große Mühe ergänzen können, zum Beispiel aus meinem Skript vom Sommersemester '97 oder aus dem Buch Analysis II von O. Forster.

1 Eindeutigkeit und Existenz der Lösung

Als Einstieg betrachten wir das Problem, die zeitliche Entwicklung einer Population (Bakterien, Bevölkerung, Kontostand, Atome, ...) vorherzusagen oder rückwärtig zu bestimmen. Dabei ist zur Zeit $t_0 \in \mathbb{R}$ ein Wert $x(t_0) = x_0$ gegeben – wir sprechen von einem Anfangswert – und wir interessieren uns für die Größe $x(t)$ der Population zur Zeit t. Je nach Problemstellung sind unterschiedliche Wachstumsgesetze denkbar:

- *Natürliches Wachstum* (Kontostand, radioaktiver Zerfall)
  Hier ist die Zuwachs-/Zerfallsrate konstant,
  $$x' = \alpha\, x \quad \text{mit } \alpha \in \mathbb{R}.$$
  Eindeutige Lösung des Anfangswertproblems ist, siehe Kapitel 3, Folgerung 2.5,
  $$x(t) = x_0\, e^{\alpha(t-t_0)}.$$

- *Logistisches Wachstum*
  Ab einem gewissen Schwellenwert soll Abnahme statt Zunahme eintreten (z. B. Schafe auf einer Weide)
  $$x' = (\alpha - \beta x)x = \alpha x - \beta x^2 \quad \text{mit } \alpha, \beta > 0.$$
  Man kann sich überlegen, dass im Fall $x_0 \neq 0$ die Lösung für $t \geq t_0$ wie folgt lautet:
  $$x(t) = \frac{1}{\frac{\beta}{\alpha} + \left(\frac{1}{x_0} - \frac{\beta}{\alpha}\right) e^{-\alpha(t-t_0)}}.$$

- *Räuber-Beute-Modell*
  Gegeben sind zwei Populationen $x(t)$ =Beute und $y(t)$ =Räuber, zum Beispiel Gänse und Füchse.

$$\begin{aligned} x' &= (\alpha - \beta y)x \\ y' &= (-\gamma + \delta x)y, \qquad (\alpha, \beta, \gamma, \delta > 0). \end{aligned}$$

  Es ergeben sich zwei gekoppelte Gleichungen, deren Lösung ohne weiteres nicht so leicht zu ermitteln ist.

Allgemein interessieren wir uns für folgende Situation.

Definition 1.1 *Sei G offen in $\mathbb{R} \times \mathbb{R}^n$ und $f \in C^0(G, \mathbb{R}^n)$, $f = f(t,x)$. Eine Funktion $x \in C^1(I, \mathbb{R}^n)$, $I \subset \mathbb{R}$ Intervall, heißt Lösung der Differentialgleichung $x' = f(\cdot\,, x)$, falls gilt:*

$$\{(t, x(t)) : t \in I\} \subset G \tag{1.1}$$

$$x'(t) = f(t, x(t)) \qquad \text{für alle } t \in I. \tag{1.2}$$

Gilt außerdem

$$x(t_0) = x_0 \text{ für gegebenes } (t_0, x_0) \in G, \tag{1.3}$$

so heißt x Lösung des zugehörigen Anfangswertproblems.

Wir können uns $x(t)$ als Fahrt eines Fahrzeugs vorstellen, das zur Zeit t_0 in x_0 startet und durch Vorgabe der Geschwindigkeit zur Zeit t am Ort $x(t)$ gesteuert wird. Es stellen sich die folgenden Fragen:

> **Eindeutigkeit:** Ist eine Lösung $x \in C^1(I, \mathbb{R}^n)$ des Anfangswertproblems eindeutig bestimmt?
>
> **Existenz:** Gibt es eine Lösung $x \in C^1(I, \mathbb{R}^n)$ des Anfangswertproblems?
>
> **Abhängigkeit von den Daten**: Wie hängt die Lösung von (t_0, x_0) und f ab?

In der Vorlesung wird nur auf die ersten beiden Fragen Bezug genommen.

Beispiel 1.1 *Sei $G = \mathbb{R} \times \mathbb{R}$ und $f(t,x) = 2\sqrt{|x|}$. Dann hat das Anfangswertproblem*

$$x' = f(\cdot\,, x) \qquad x(0) = 0$$

unendlich viele verschiedene Lösungen $x \in C^1(\mathbb{R})$, nämlich die Funktionen

$$x(t) = \begin{cases} 0 & \text{für } t \le s \\ (t-s)^2 & \text{für } t > s \end{cases}$$

mit $s \geq 0$, sowie außerdem die Nullfunktion.

Das Beispiel zeigt, dass für die Eindeutigkeit eine Regularitätsvoraussetzung erforderlich ist.

Satz 1.1 (Eindeutigkeit der Lösung) *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen. Ist $f \in C^1(G, \mathbb{R}^n)$, so hat das Anfangswertproblem*

$$x' = f(\,\cdot\,, x), \qquad x(t_0) = x_0$$

höchstens eine Lösung $x \in C^1(I, \mathbb{R}^n)$.

Für den Beweis brauchen wir folgende Aussage, die direkt aus dem Schrankensatz, Satz 2.2 in Kapitel 7, folgt.

Lemma 1.1 *Sei $f \in C^0(G, \mathbb{R}^n)$, $f = f(t,x)$, mit $D_x f \in C^0(G, \mathbb{R}^{n \times n})$. Dann ist f lokal Lipschitz-stetig bzgl. x: zu $(t_0, x_0) \in G$ gibt es ein $\varepsilon > 0$ und ein $L < \infty$, so dass gilt:*

$$\begin{aligned} Z_\varepsilon(t_0, x_0) : &= \{(t,x) : |t - t_0| \leq \varepsilon, |x - x_0| \leq \varepsilon\} \subset G, && (1.4)\\ |f(t, x_1) - f(t, x_2)| &\leq L\,|x_1 - x_2| \quad \text{für } (t, x_1), (t, x_2) \in Z_\varepsilon(t_0, x_0). && (1.5) \end{aligned}$$

Beweis von Satz 1.1 Seien $x_1, x_2 \in C^1(I, \mathbb{R}^n)$ zwei Lösungen des Anfangswertproblems.

Schritt 1: Es gibt ein $\delta > 0$ mit $x_1 = x_2$ auf $(t_0 - \delta, t_0 + \delta)$.
Seien $\varepsilon > 0$, $L < \infty$ wie in Lemma 1.1. Wähle $\delta > 0$ mit $\big(t, x_1(t)\big)$, $\big(t, x_2(t)\big) \in Z_\varepsilon(t_0, x_0)$ für $|t - t_0| \leq \delta$. Für $u(t) = |x_1(t) - x_2(t)|^2$ folgt

$$\begin{aligned} |u'| &= 2|\langle x_1 - x_2, x_1' - x_2'\rangle| \\ &= 2|\langle x_1 - x_2, f(\,\cdot\,, x_1) - f(\,\cdot\,, x_2)\rangle| \\ &\leq 2|x_1 - x_2|\,|f(\,\cdot\,, x_1) - f(\,\cdot\,, x_2)| \quad \text{(Cauchy-Schwarz)} \\ &\leq 2\,L\,u. \end{aligned}$$

Mit Folgerung 2.5 in Kapitel 2 erhalten wir $u(t) \leq u(t_0)\, e^{2L|t - t_0|}$ bzw.

$$|x_1(t) - x_2(t)| \leq |x_1(t_0) - x_2(t_0)|\, e^{L|t - t_0|}. \tag{1.6}$$

Die Behauptung ergibt sich aus $x_1(t_0) = x_2(t_0)$.

Schritt 2: $x_1 = x_2$ auf ganz I.
Betrachte die Menge $M = \{t \in I : x_1(t) = x_2(t)\}$. Dann gilt $t_0 \in M$ nach Voraussetzung und M ist abgeschlossen in I, da x_1, x_2 stetig. Nach Schritt 1 gibt es zu $t \in M$ aber ein $\delta > 0$ mit $(t - \delta, t + \delta) \subset M$, d. h. M ist offen. Daraus folgt $M = I$.

$\square$

Wir kommen nun zur Frage der Existenz. Das folgende Beispiel zeigt, dass wir nur eine zeitlich lokale Lösung erwarten können.

Beispiel 1.2 *Betrachte $f \in C^\infty(\mathbb{R} \times \mathbb{R})$, $f(t,x) = x^2$. Das zugehörige Anfangswertproblem*

$$x' = x^2, \quad x(0) = 1$$

hat die eindeutige Lösung

$$x : (-\infty, 1) \to \mathbb{R}, \quad x(t) = \frac{1}{1 - t}.$$

Die Lösung ist nach rechts nicht fortsetzbar, denn es gilt $\lim_{t \nearrow 1} x(t) = +\infty$.

Eine entscheidende Beobachtung zur Konstruktion der lokalen Lösung ist, dass das Anfangswertproblem als Integralgleichung geschrieben werden kann.

Lemma 1.2 *Für $x : I \to \mathbb{R}^n$ und $(t_0, x_0) \in G$ sind folgende Aussagen äquivalent:*

(1) *$x \in C^1(I, \mathbb{R}^n)$ ist Lösung des Anfangswertproblems*

$$x'(t) = f\big(t, x(t)\big) \text{ für alle } t \in I, \quad x(t_0) = x_0.$$

(2) *$x \in C^0(I, \mathbb{R}^n)$ erfüllt die Gleichung*

$$x(t) = x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\, ds \quad \text{für alle } t \in I.$$

Beweis:
(1)⇒(2): folgt aus dem Hauptsatz der Differential- und Integralrechung durch Integration von t_0 bis t.
(2)⇒(1): Da $t \mapsto f(t, x(t))$ stetig, folgt $x \in C^1(I, \mathbb{R}^n)$ und die Differentialgleichung ergibt sich durch Differentiation, wieder mit dem Hauptsatz. □

Satz 1.2 (Kurzzeitexistenzsatz von Picard-Lindelöf) *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen, $f \in C^1(G, \mathbb{R}^n)$ und $(t_0, x_0) \in G$. Dann gibt es ein $\delta > 0$, so dass das Anfangswertproblem*

$$x' = f(\,\cdot\,, x), \quad x(t_0) = x_0 \quad \text{auf } I = [t_0 - \delta, t_0 + \delta]$$

eine Lösung $x \in C^1(I, \mathbb{R}^n)$ besitzt.

Beweis:
Schritt 1 *Formulierung als Fixpunktproblem*
Seien $\varepsilon > 0$ und $L < \infty$ wie in Lemma 1.1 gewählt, sowie $\delta \in (0, \varepsilon]$ zunächst beliebig.
Der Raum $X = C^0(I, \mathbb{R}^n)$ ist, versehen mit der Supremumsnorm $\|\cdot\|_I$, ein Banachraum (Satz 2.1 in Kapitel 6). Auf der abgeschlossenen Teilmenge

$$A = \{x \in X : |x(t) - x_0| \le \varepsilon \quad \text{für alle } t\} \subset X$$

betrachten wir die Abbildung

$$F : A \to X, \quad [F(x)](t) = x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\, ds.$$

Nach Lemma 1.2 ist ein Fixpunkt von F, also $F(x) = x$, eine Lösung des Anfangswertproblems.

Schritt 2 Wir zeigen nun, dass für hinreichend kleines $\delta \in (0, \varepsilon]$ Folgendes gilt:

a) $F(A) \subset A$

b) $\|F(x) - F(y)\|_I \le \frac{1}{2}\, \|x - y\|_I$ für alle $x, y \in A$.

Aus dem Banachschen Fixpunktsatz, Satz 1.1 in Kapitel 8, folgt dann die Existenz eines Fixpunkts $x \in A$, also der gewünschten Lösung der Anfangswertproblems.

Da $f : G \to \mathbb{R}^n$ stetig und $Z_\varepsilon(t_0, x_0) \subset G$ kompakt, ist $M = \sup\{|f(t,x)| : |t - t_0| \leq \varepsilon, |x - x_0| \leq \varepsilon\} < \infty$ und für $x \in A$ folgt

$$|[F(x)](t) - x_0| = \Big| \int_{t_0}^{t} f\big(s, x(s)\big)\, ds \Big| \leq M(t - t_0) \leq M\,\delta. \tag{1.7}$$

Weiter schätzen wir für $x, y \in A$ ab

$$\begin{aligned}
|[F(x)](t) - [F(y)](t)| &= \Big| \int_{t_0}^{t} \big(f(s, x(s)) - f(s, y(s))\big)\, ds \Big| \\
&\leq |t - t_0| \sup_{s \in I} \big| f\big(s, x(s)\big) - f\big(s, y(s)\big) \big| \\
&\leq L\,\delta \sup_{s \in I} |x(s) - y(s)|,
\end{aligned}$$

also

$$\|F(x) - F(y)\|_I \leq L\,\delta\, \|x - y\|_I. \tag{1.8}$$

Nach (1.7) und (1.8) sind die Bedingungen a) und b) für die Anwendung des Banachschen Fixpunktsatzes mit $\delta = \min\{\frac{\varepsilon}{M}, \frac{1}{2L}\} > 0$ erfüllt. □

Wir verweisen auf das Skript Analysis II vom SS '97 für eine Diskussion der Frage, unter welchen Voraussetzungen die Existenz einer globalen Lösung bewiesen werden kann.

2 Separation der Variablen

Als Ausgangspunkt betrachten wir nochmals die Euler-Lagrange-Gleichungen.

Satz 2.1 (Energieerhaltungssatz) *Sei $\Omega \subset \mathbb{R}^n$ offen und $f \in C^2(\Omega \times \mathbb{R}^n)$, $f = f(x, v)$ (also f unabhängig von t). Ist $x \in C^2(I, \Omega)$ eine Lösung der Euler-Lagrange-Gleichungen, so gilt der „Energieerhaltungssatz“*

$$\frac{d}{dt}[\langle D_v f(x, x'), x' \rangle - f(x, x')] = 0.$$

Beweis:

$$\frac{d}{dt}\left[\sum_{j=1}^{n} \frac{\partial f}{\partial v_j}(x, x') x'_j - f(x, x')\right] = \sum_{j=1}^{n} \left(\underbrace{\frac{d}{dt}\frac{\partial f}{\partial v_j}(x, x') - \frac{\partial f}{\partial x_j}(x, x')}_{=0}\right) x'_j = 0$$

□

Beispiel 2.1 *Sei $f(x,v) = \frac{m}{2}|v|^2 - V(x)$ die Lagrangefunktion eines Teilchens der Masse $m > 0$ im Potential V. Dann gilt*

$$\begin{aligned}\langle D_v f(x,x'), x'\rangle - f(x,x') &= \langle mx', x'\rangle - \Big(\frac{m}{2}\,|x'|^2 - V(x)\Big)\\ &= \frac{m}{2}\,|x'|^2 + V(x).\end{aligned}$$

Es folgt der Energieerhaltungssatz

$$\frac{m}{2}\,|x'|^2 + V(x) = E \quad (= \text{Konstante}).$$

Bei Systemen mit einem Freiheitsgrad ($n = 1$) lautet der Energieerhaltungssatz

$$x' = \pm\sqrt{\frac{2}{m}\big(E - V(x)\big)}.$$

Anfangswertprobleme dieses Typs können nun „explizit“ gelöst werden mit dem nachstehendem Verfahren.

Satz 2.2 (Separation der Variablen) *Seien $f \in C^0(I)$, $g \in C^0(J)$ für Intervalle $I, J \subset \mathbb{R}$, und $g(x) \neq 0$ für $x \in J$. Betrachte das Anfangswertproblem*

$$x'(t) = \frac{f(t)}{g\big(x(t)\big)} \quad \text{für } t \in I, \quad x(t_0) = x_0. \tag{2.1}$$

Definiere $F \in C^1(I)$, $G \in C^1(J)$ durch

$$F(t) = \int_{t_0}^t f(s)\,ds, \quad G(x) = \int_{x_0}^x g(y)\,dy.$$

Ist I so klein gewählt dass $F(I) \subset G(J)$, so folgt:

a) Es gibt genau ein $x \in C^1(I)$ mit $G\big(x(t)\big) = F(t)$ für alle $t \in I$.

b) Diese Funktion ist die eindeutig bestimmte Lösung des Anfangswertproblems (2.1).

Beweis: Nach Analysis 1 ist $J^* = G(J)$ ein offenes Intervall und es existiert die Umkehrfunktion $H \in C^1(J^*)$ von G. Für $x \in C^1(I)$ mit $x(I) \subset J$ gilt:

$$\begin{array}{cccll} & x \text{ ist Lösung} & \text{von} & (2.1) & \\ \Leftrightarrow & \frac{d}{dt}\,G\big(x(t)\big) & = & \frac{d}{dt}\,F(t) & \text{und } x(t_0) = x_0 \\ \Leftrightarrow & G\big(x(t)\big) & = & F(t) & \text{(beachte } G(x_0) = F(t_0) = 0) \\ \Leftrightarrow & x(t) & = & H\big(F(t)\big). & \end{array}$$

Insbesondere ist eine Lösung von (2.1) eindeutig bestimmt. Definieren wir $x = H \circ F$, so ist $x \in C^1(I)$ mit $x(I) = H\big(F(I)\big) \subset H\big(G(J)\big) = J$ nach Voraussetzung und x ist Lösung von (2.1). $\square$

Lösungsrezept.

- $\dfrac{dx}{dt} = \dfrac{f(t)}{g(x)} \quad \longrightarrow \quad g(x)\,dx = f(t)\,dt$

- Integration von t_0 bis t und x_0 bis x:

$$G(x) = \int_{x_0}^{x} g(y)\, dy = \int_{t_0}^{t} f(s)\, ds = F(t)$$

- Auflösung nach x ergibt $x = x(t)$ oder Auflösung nach t ergibt $t = t(x)$.

Beispiel 2.2 (Homogene, lineare Gleichung) *Betrachte für $a \in C^0(I)$ das Anfangswertproblem*

$$x'(t) = a(t)x(t), \quad x(t_0) = x_0.$$

Nach dem Lösungsrezept ergibt sich (unter welchen Voraussetzungen?)

- $\dfrac{dx}{x} = a(t)\, dt$

- *Integration von t_0 bis t bzw. x_0 bis x:*

$$\log \frac{x}{x_0} = \int_{x_0}^{x} \frac{dy}{y} = \int_{t_0}^{t} a(s)\, ds$$

- $x(t) = x_0 \exp \int_{t_0}^{t} a(s)\, ds.$

Beispiel 2.3 (Rotationsminimalflächen) *Wird der Graph einer Funktion $r : I \to (0, \infty)$, $r = r(x)$, um die x-Achse rotiert, so hat die entstehende Fläche (wie man zeigen kann) den Flächeninhalt*

$$A(r) = 2\pi \int_I r(x)\sqrt{1 + r'(x)^2}\, dx.$$

Falls r diesen Flächeninhalt relativ zu allen Graphen $u : I \to (0, \infty))$ mit $u|_{\partial I} = r|_{\partial I}$ minimiert, so ist r Lösung der zugehörigen Euler-Lagrange-Gleichung. Aus Satz 2.1 folgt

$$\begin{aligned} \frac{r\, r'}{\sqrt{1 + (r')^2}}\, r' - r\sqrt{1 + (r')^2} &= -a \quad (= Konst.) \\ \Leftrightarrow \quad (r')^2 &= \left(\frac{r}{a}\right)^2 - 1. \end{aligned}$$

Wir lösen mit Separation unter der Annahme $r' > 0$:

$$\frac{dr}{dt} = \sqrt{\left(\frac{r}{a}\right)^2 - 1} \quad \longrightarrow \quad \left(\left(\frac{r}{a}\right)^2 - 1\right)^{-1/2} dr = dx.$$

Integration von r_0 bis r, x_0 bis x liefert Arcosh $\frac{r}{a}$ − Arcosh $\frac{r_0}{a} = x - x_0$ *beziehungweise*

$$r(x) = a \cosh \frac{x - x_1}{a} \qquad \textit{mit } x_1 = x_0 - \operatorname{Arcosh} \frac{r_0}{a}.$$

3 Lineare Differentialgleichungen

Für lineare Differentialgleichungen erhält man eine sehr präzise Theorie. Wir betrachten hier allgemein ein System

$$x'(t) = A(t)\,x(t) + b(t), \quad x(t_0) = x_0. \tag{3.1}$$

Dabei sind $A : I \to \mathbb{R}^{n\times n} \simeq L(\mathbb{R}^n, \mathbb{R}^n)$ und $b \in I \to \mathbb{R}^n$ stetig und $x_0 \in \mathbb{R}$. Gesucht ist eine Lösung $x \in C^1(I, \mathbb{R}^n)$ von (3.1). Es erweist sich aber als nützlich, im Folgenden auch komplexwertige Funktionen zuzulassen, also

$$A(t) \in \mathbb{C}^{n\times n}, \quad b(t) \in \mathbb{C}^n, \quad x_0 \in \mathbb{C}^n \quad \text{und } x(t) \in \mathbb{C}^n.$$

Tatsache. Das lineare Anfangwertproblem (3.1) ist auf ganz I (und nicht nur lokal) lösbar. Dies ergibt sich durch eine genauere Analyse des Satzes von Picard-Lindelöf, siehe Skript Analysis II, SS '97.

Satz 3.1 *Sei $A \in C^0(I, \mathbb{K}^{n\times n})$ mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Dann ist die Lösungsmenge des homogenen Systems*

$$L_A = \{x \in C^1(I, \mathbb{K}^n) : x' = A\,x \text{ auf } I\}$$

ein n-dimensionaler $\mathbb{K}$-Vektorraum. Und zwar ist für jedes $t_0 \in I$ die Abbildung

$$\delta_{t_0} : L_A \to \mathbb{K}^n, \quad \delta_{t_0}(x) = x(t_0)$$

ein Vektorraumisomorphismus.

Beweis: Mit $x, y \in L_A$ ist auch $\lambda\,x + \mu\,y \in L_A$ für $\lambda, \mu \in \mathbb{R}$. Deshalb ist L_A ein Untervektorraum von $C^1(I, \mathbb{K}^n)$. Die Abbildung δ_{t_0} ist surjektiv, weil das Anfangswertproblem zu jedem $x_0 \in \mathbb{K}^n$ eine Lösung auf ganz I hat, siehe obige Tatsache. δ_{t_0} ist injektiv, weil die Lösung des Anfangswertproblems nach Satz 2.1 eindeutig bestimmt ist. □

Definition 3.1 *Eine Basis von L_A heißt (Lösungs-)Fundamentalsystem auf I.*

Als Beispiel betrachten wir die Differentialgleichung zweiter Ordnung (Schwingung mit Reibung)

$$x'' + 2\beta\,x' + \omega_0^2 x = 0 \quad (\beta_0, \omega_0 \geq 0).$$

Wir können die Gleichung äquivalent in ein System erster Ordnung umschreiben:

$$\begin{aligned} x' &= y \\ y' &= -\omega_0^2 x - 2\beta\,y \end{aligned}$$

Ein Fundamentalsystem ergibt sich aus dem Ansatz $x(t) = e^{\lambda t}$ mit $\lambda \in \mathbb{C}$, also

$$0 = e^{\lambda t}(\lambda^2 + 2\,\beta\,\lambda + \omega_0^2).$$

Fall 1 $0 < \beta^2 - \omega_0^2 =: \omega^2$ (wobei $\omega > 0$)
Dann hat das Polynom die beiden reellen Nullstellen $\lambda_\pm = -\beta \pm \omega < 0$. Die beiden Lösungen

$$\begin{pmatrix} x_\pm(t) \\ y_\pm(t) \end{pmatrix} = \begin{pmatrix} e^{\lambda_\pm t} \\ \lambda_\pm\, e^{\lambda_\pm t} \end{pmatrix}$$

bilden ein Fundamentalsystem, denn

$$\det \begin{pmatrix} x_+(0) & x_-(0) \\ y_+(0) & y_-(0) \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ \lambda_+ & \lambda_- \end{pmatrix} = \lambda_- - \lambda_+ = -2\omega < 0.$$

Fall 2 $0 > \beta^2 - \omega_0^2 =: (i\omega)^2$ (wobei $w > 0$). Dann hat das Polynom die beiden nichtreellen Nullstellen $\lambda_\pm = -\beta \pm i\omega$. Analog zu oben erhalten wir nun ein komplexes Fundamentalsystem

$$\begin{pmatrix} x_\pm(t) \\ y_\pm(t) \end{pmatrix} = \begin{pmatrix} e^{\lambda_\pm t} \\ \lambda_\pm \, e^{\lambda_\pm t} \end{pmatrix}.$$

Über $\mathbb{R}$ erhalten wir durch Linearkombination

$$\begin{array}{rclcrcl} x_1(t) & = & e^{-\beta t} \cos \omega t & \quad & x_2(t) & = & e^{-\beta t} \sin \omega t \\ y_1(t) & = & x_1'(t) & & y_2(t) & = & x_2'(t) \end{array}$$

Fall 3 $0 = \beta^2 - \omega_0^2$
In diesem Fall hat das Polynom die eine reelle Nullstelle $-\beta$. Ein Fundamentalsystem lautet

$$\begin{array}{rclcrcl} x_1(t) & = & e^{-\beta t} & \quad & x_2(t) & = & t\, e^{-\beta t} \\ y_1(t) & = & -e^{-\beta t} & & y_2(t) & = & e^{-\beta t} - \beta\, t\, e^{-\beta t} \end{array}$$

An der Stelle $t = 0$

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$