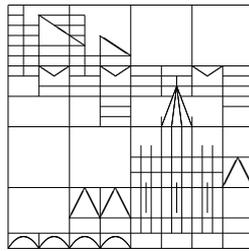


Skript zur Vorlesung

Analysis II

Sommersemester 2016

Robert Denk



Universität Konstanz

Fachbereich Mathematik und Statistik

Stand: 29.08.2016

Inhaltsverzeichnis

1	Die Topologie des \mathbb{R}^n	1
	a) \mathbb{R}^n als normierter Vektorraum	1
	b) Stetigkeit und Kompaktheit	4
2	Funktionen mehrerer Veränderlicher	9
	a) Differenzierbare Abbildungen	9
	b) Der Mittelwertsatz	19
	c) Höhere Ableitungen	21
	d) Extrema unter Nebenbedingungen	29
3	Kurven und Flächen	35
	a) Weglängen	35
	b) Kurven in der Ebene und im Raum	39
	c) m -dimensionale Flächen im \mathbb{R}^n	41
4	Grundlagen der Maß- und Integrationstheorie	46
	a) Maße und messbare Funktionen	46
	b) Das Lebesgue-Integral	54
	c) Iterierte Integrale	60
	d) Der Transformationssatz	64
5	Kurvenintegrale und Flächenintegrale, der Satz von Gauß	67
	a) Kurvenintegrale und 1-Formen	67
	b) Die Integralsätze von Gauß und Stokes	77
6	Lokale Umkehrbarkeit und implizite Funktionen	85
	Literatur	90
	Index	91

1. Die Topologie des \mathbb{R}^n

1.1 Worum geht's? Wir werden in diesem Semester vor allem Funktionen von \mathbb{R}^n nach \mathbb{R}^m betrachten. Dafür müssen die aus dem ersten Semester bekannten Begriffe wie etwa Stetigkeit, Konvergenz und Normen vom \mathbb{R}^1 auf den \mathbb{R}^n übertragen werden. In diesem ersten kurzen Abschnitt werden einige topologische Begriffe und Ergebnisse zusammengefasst, welche zum Teil auch allgemein in normierten oder metrischen Räumen gelten.

a) \mathbb{R}^n als normierter Vektorraum

Im Folgenden sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

1.2 Definition. Sei V ein \mathbb{K} -Vektorraum. Eine Abbildung $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{K}$ heißt ein Skalarprodukt, falls gilt:

- (i) Es gilt $\langle f, f \rangle \geq 0$ ($f \in V$) und $\langle f, f \rangle = 0 \Leftrightarrow f = 0$,
- (ii) Die Abbildung $f \mapsto \langle f, g \rangle$, $V \rightarrow \mathbb{K}$, ist \mathbb{K} -linear, d.h. es gilt

$$\langle \alpha_1 f_1 + \alpha_2 f_2, g \rangle = \alpha_1 \langle f_1, g \rangle + \alpha_2 \langle f_2, g \rangle \quad (\alpha_1, \alpha_2 \in \mathbb{K}, f_1, f_2, g \in V).$$

- (iii) Es gilt $\langle f, g \rangle = \overline{\langle g, f \rangle}$ ($f, g \in V$).

In diesem Fall heißt $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit Skalarprodukt oder Prähilbertraum. Falls V mit der induzierten Norm $\|f\| := \sqrt{\langle f, f \rangle}$ vollständig (d.h. ein Banachraum) ist, so heißt V ein \mathbb{K} -Hilbertraum.

Im Folgenden sei $n \in \mathbb{N}$ fest. Wir schreiben Vektoren des \mathbb{R}^n in der Form

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

(Spaltenvektor), der transponierte Vektor ist definiert durch $x^\top := (x_1, \dots, x_n)$ (Zeilenvektor). Das Standard-Skalarprodukt ist $\langle x, y \rangle := \sum_{i=1}^n x_i y_i$, die zugehörige Norm ist die euklidische Norm $|x| := \langle x, x \rangle^{1/2}$. Damit wird \mathbb{R}^n zu einem Hilbertraum.

Auf dem \mathbb{R}^n existieren noch weitere wichtige Normen, welche im folgenden Beispiel angegeben werden. Der Nachweis der Norm-Eigenschaften ist hier einfach bis auf den Beweis der Dreiecksungleichung, der im nachfolgenden Satz durchgeführt wird.

1.3 Beispiele. a) Für $x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$ und $1 \leq p \leq \infty$ definiert man

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, & \text{falls } p \in [1, \infty), \\ \max_{i=1, \dots, n} |x_i|, & \text{falls } p = \infty. \end{cases}$$

Damit wird $(\mathbb{R}^n, \|\cdot\|_p)$ ein normierter Raum.

b) Man definiert die Folgenräume ℓ^p für $1 \leq p \leq \infty$ durch $\ell^p := \{x \in \mathbb{K}^{\mathbb{N}} : \|x\|_p < \infty\}$, wobei für $x = (x_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ die Norm $\|x\|_p$ definiert wird durch

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}, & \text{falls } 1 \leq p < \infty, \\ \sup_{i \in \mathbb{N}} |x_i|, & \text{falls } p = \infty. \end{cases}$$

Damit wird $(\ell^p, \|\cdot\|_p)$ zu einem normierten Raum.

c) Seien $a, b \in \mathbb{R}$ mit $a < b$. Für $f \in C([a, b]) := \{f: [a, b] \rightarrow \mathbb{K} : f \text{ stetig}\}$ und $1 \leq p \leq \infty$ definiert man

$$\|f\|_p := \begin{cases} \left(\int_a^b |f(x)|^p dx \right)^{1/p}, & \text{falls } p \in [1, \infty), \\ \sup_{x \in [a, b]} |f(x)|, & \text{falls } p = \infty. \end{cases}$$

Damit wird $(C([a, b]), \|\cdot\|_p)$ zu einem normierten Raum.

1.4 Satz. Zu $1 \leq p \leq \infty$ sei der konjugierte Exponent q definiert durch $\frac{1}{p} + \frac{1}{q} = 1$ (d.h. $q = \frac{p}{p-1}$ falls $p > 1$ und $q = \infty$ falls $p = 1$).

a) (*Youngsche Ungleichung*)¹ Falls $p \in (1, \infty)$, so gilt für alle $a, b \geq 0$

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Dabei gilt die Gleichheit nur, wenn $a^p = b^q$.

b) (*Höldersche Ungleichung*)² Für $1 \leq p \leq \infty$ gilt

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \|y\|_q \quad (x \in \ell^p, y \in \ell^q),$$

$$\int_a^b |f(x)g(x)| dx \leq \|f\|_p \|g\|_q \quad (f, g \in C([a, b])).$$

c) (*Minkowskische Ungleichung*)³ Für $1 \leq p \leq \infty$ gilt

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \quad (x, y \in \ell^p),$$

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad (f, g \in C([a, b])).$$

¹William Henry Young, 20.10.1862 – 7.7.1942

²Otto Hölder, 22.12.1859 – 29.8.1937

³Hermann Minkowski, 22.6.1864 – 12.1.1909

Beweis. a) Wir betrachten im positiven Quadranten der (x, y) -Ebene die durch die Gleichung $x^p = y^q$ gegebene Kurve (d.h. $y = x^{p-1}$ bzw. $x = y^{q-1}$). Dann ist die Fläche zwischen der x -Achse (mit $x \in [0, a]$) und dieser Kurve gegeben durch

$$A_1 = \int_0^a x^{p-1} dx = \frac{x^p}{p} \Big|_{x=0}^a = \frac{a^p}{p}.$$

Die Fläche zwischen der y -Achse (mit $y \in [0, b]$) und der Kurve ist analog gegeben durch $A_2 = \frac{b^q}{q}$ (siehe Abbildung 1). Wie die Abbildung zeigt, ist die Summe der beiden Flächen mindestens so groß wie die Fläche des durch die Punkte $(0, 0)$, $(a, 0)$, $(0, b)$ und (a, b) gegebenen Rechtecks. Diese ist ab . Damit erhalten wir

$$A_1 + A_2 = \frac{a^p}{p} + \frac{b^q}{q} \geq ab.$$

Wie die Abbildung ebenfalls zeigt, gilt Gleichheit nur für $a^p = b^q$.

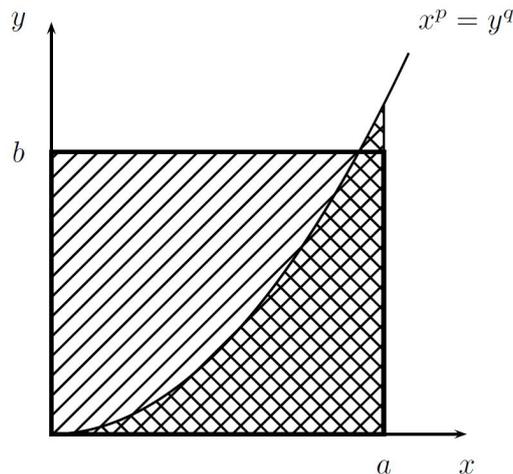


Abbildung 1: Zum Beweis der Youngschen Ungleichung

b) Für $p = 1$ oder $p = \infty$ oder $x = 0$ oder $y = 0$ ist die Behauptung klar. Seien also $p \in (1, \infty)$ und $x \neq 0 \neq y$. Zu $i \in \mathbb{N}$ wenden wir die Youngsche Ungleichung an mit $a := \frac{|x_i|}{\|x\|_p}$ und $b := \frac{|y_i|}{\|y\|_q}$. Wir erhalten

$$\frac{|x_i| \cdot |y_i|}{\|x\|_p \|y\|_q} \leq \frac{|x_i|^p}{p \|x\|_p^p} + \frac{|y_i|^q}{q \|y\|_q^q}.$$

Summation über i und Multiplikation mit $\|x\|_p \cdot \|y\|_q$ liefert die erste Behauptung in b). Die zweite Behauptung wird analog gezeigt, wobei man jetzt

$$a := \frac{|f(x)|}{\|f\|_p} \quad \text{und} \quad b := \frac{|g(x)|}{\|g\|_q}$$

wählt. Die Youngsche Ungleichung ergibt

$$\frac{|f(x)g(x)|}{\|f\|_p\|g\|_q} \leq \frac{|f(x)|^p}{p\|f\|_p^p} + \frac{|g(x)|^q}{q\|g\|_q^q}.$$

Integration über x und Multiplikation mit $\|f\|_p\|g\|_q$ liefert nun die Behauptung.

c) Die Aussage ist wiederum klar für $p = 1$ oder $p = \infty$. Sei also $p \in (1, \infty)$. Dann erhält man unter Verwendung der Hölderschen Ungleichung

$$\begin{aligned} \|x + y\|_p^p &= \sum_{i=1}^{\infty} |x_i + y_i|^p \leq \sum_{i=1}^{\infty} |x_i + y_i|^{p-1} |x_i| + \sum_{i=1}^{\infty} |x_i + y_i|^{p-1} |y_i| \\ &\leq \left(\sum_{i=1}^{\infty} |x_i + y_i|^{(p-1)q} \right)^{1/q} (\|x\|_p + \|y\|_p) \\ &= \|x + y\|_p^{p/q} (\|x\|_p + \|y\|_p). \end{aligned}$$

Da $p - \frac{p}{q} = p - (p - 1) = 1$ ist, folgt hieraus die erste Behauptung in c). Analog beweist man die zweite Behauptung. \square

b) Stetigkeit und Kompaktheit

1.5 Satz (von Bolzano-Weierstraß). Sei $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$ beschränkt. Dann existiert eine konvergente Teilfolge $(x_{k_j})_{j \in \mathbb{N}}$.

Beweis. Im Beweis verwenden wir die Bezeichnung $P_j x := x_j$ für Vektoren $x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$, d.h. $P_j x$ ist die j -te Komponente des Vektors x .

Sei $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$ mit $|x_k| = \|x_k\|_2 \leq M$. Dann gilt $|P_j x_k| \leq M$ für $j = 1, \dots, n, k \in \mathbb{N}$. Da $(P_1 x_k)_{k \in \mathbb{N}}$ eine beschränkte Folge ist, existiert $\limsup_{k \rightarrow \infty} P_1 x_k$ und damit eine Teilfolge $(x_\ell^{(1)})_{\ell \in \mathbb{N}}$ von $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$, für die $(P_1 x_\ell^{(1)})_{\ell \in \mathbb{N}} \subset \mathbb{R}$ konvergiert.

Davon existiert wieder eine Teilfolge $(x_\ell^{(2)})_{\ell \in \mathbb{N}}$, für welche $(P_2 x_\ell^{(2)})_{\ell \in \mathbb{N}} \subset \mathbb{R}$ konvergiert, usw. Insgesamt erhalten wir eine Teilfolge $(x_\ell^{(n)})_{\ell \in \mathbb{N}}$, für die alle Komponenten $(P_j x_\ell^{(j)})_{\ell \in \mathbb{N}} \subset \mathbb{R}$, $j = 1, \dots, n$, konvergent sind, d.h. $(x_\ell^{(n)})_{\ell \in \mathbb{N}} \subset \mathbb{R}^n$ konvergiert. \square

1.6 Satz. Für $K \subset \mathbb{R}^n$ sind äquivalent:

- (i) K ist kompakt,
- (ii) K ist abgeschlossen und beschränkt,

(iii) jede Folge $(x_k)_{k \in \mathbb{N}} \subset K$ besitzt eine konvergente Teilfolge $(x_{k_j})_{j \in \mathbb{N}}$ mit $x := \lim_{j \rightarrow \infty} x_{k_j} \in K$.

Beweis. (i) \Rightarrow (ii). Das gilt allgemein in metrischen Räumen (Analysis I).

(ii) \Rightarrow (iii). Satz von Bolzano-Weierstraß.

(iii) \Rightarrow (i). Wir zeigen folgende Schritte:

(a) K ist abgeschlossen: zu $x \in \overline{K}$ existiert eine Folge $(x_k)_{k \in \mathbb{N}} \subset K$ mit $x_k \rightarrow x$. Nach (iii) existiert eine Teilfolge $(x_{k_j})_{j \in \mathbb{N}}$ mit $\tilde{x} := \lim_{j \rightarrow \infty} x_{k_j} \in K$. Da der Grenzwert eindeutig ist, folgt $x = \tilde{x} \in K$, d.h. $K = \overline{K}$.

(b) K ist totalbeschränkt, d.h.

$$\forall \varepsilon > 0 \exists K_0 \subset K, K_0 \text{ endlich} : K \subset \bigcup_{a \in K_0} B(a, \varepsilon).$$

Sonst existiert ein $\varepsilon > 0$, so dass für alle endlichen $K_0 \subset K$ ein $x \in K$ existiert mit $x \notin \bigcup_{a \in K_0} B(a, \varepsilon)$. Wähle $x_1 \in K$ und (induktiv) $x_k \in K \setminus \bigcup_{\ell=1}^{k-1} B(x_\ell, \varepsilon)$. Dann ist $|x_k - x_\ell| \geq \varepsilon$ ($k \geq \ell$), Widerspruch zu (iii).

(c) K ist kompakt (Prinzip der Intervallschachtelung). Angenommen, $K \subset \bigcup_{i \in I} U_i$ besitzt keine endliche Teilüberdeckung. Zu $\varepsilon_1 := \frac{1}{2}$ existiert nach (b) ein $K_1 \subset K$, K_1 endlich, mit $K \subset \bigcup_{a \in K_1} B(a, \varepsilon_1)$. Dann existiert ein $a_1 \in K_1$ so, dass $K \cap B(a_1, \varepsilon_1)$ keine endliche Teilüberdeckung besitzt.

Zu $\varepsilon_2 := \frac{1}{4}$ existiert $K_2 \subset K$ endlich mit $K \subset \bigcup_{a \in K_2} B(a, \varepsilon_2)$. Es existiert ein $a_2 \in K_2$ so, dass $K \cap B(a_1, \varepsilon_1) \cap B(a_2, \varepsilon_2)$ keine endliche Teilüberdeckung besitzt; usw. Wir erhalten mit $\varepsilon_m := 2^{-m}$ eine Folge $(a_m)_{m \in \mathbb{N}}$ so, dass $A_M := K \cap \bigcap_{m=1}^M B(a_m, \varepsilon_m)$ keine endliche Teilüberdeckung besitzt. Insbesondere ist $A_M \neq \emptyset$. Für $v \in A_M$ folgt $|a_M - a_{M-1}| \leq |a_M - v| + |v - a_{M-1}| \leq 2^{-M} + 2^{-M+1} < 2^{-M+2}$. Somit ist $(a_m)_{m \in \mathbb{N}}$ eine Cauchyfolge, d.h. $z := \lim_{M \rightarrow \infty} a_M \in \mathbb{R}^n$ existiert. Nach (a) ist $z \in K$. Also existiert ein i_0 mit $z \in U_{i_0}$. Da U_{i_0} offen ist, existiert ein $\varepsilon > 0$ mit $B(z, 2\varepsilon) \subset U_{i_0}$. Wähle $M \in \mathbb{N}$ mit $|a_M - z| < \varepsilon$ und $2^{-M} < \varepsilon$. Für $x \in A_M$ gilt dann $|z - x| \leq |z - a_M| + |a_M - x| < \varepsilon + 2^{-M} < 2\varepsilon$, d.h. $A_M \subset B(z, 2\varepsilon) \subset U_{i_0}$. Das ist eine endliche Überdeckung von A_M , Widerspruch. \square

1.7 Korollar (Intervallschachtelung). Seien $A_k \subset \mathbb{R}^n$, $A_k \neq \emptyset$, A_k abgeschlossen, mit $A_k \supset A_{k+1}$ ($k \in \mathbb{N}$). Falls ein A_k beschränkt ist, so ist $\bigcap_{k \in \mathbb{N}} A_k \neq \emptyset$.

Beweis. O.E. sei A_1 beschränkt und damit kompakt. Angenommen $\bigcap_{k \in \mathbb{N}} A_k = \emptyset$. Dann ist $A_1 \subset \bigcup_{k \in \mathbb{N}} (\mathbb{R}^n \setminus A_k) = \mathbb{R}^n$ eine offene Überdeckung. Da A_1 kompakt ist, existiert eine endliche Teilüberdeckung $A_1 \subset \bigcup_{k=1}^K (\mathbb{R}^n \setminus A_k) = \mathbb{R}^n \setminus A_K$, d.h. $A_K = A_1 \cap A_K = \emptyset$. \square

1.8 Bemerkung. a) In diesem Semester werden wir Abbildungen

$$f: X \supset D(f) \rightarrow Y$$

behandeln. Dabei sind $D(f)$ der Definitionsbereich von f und (X, d_X) , (Y, d_Y) metrische Räume. Insbesondere interessieren die Fälle $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$. Man beachte, dass durch die Einschränkung der Metrik d_X bzw. d_Y auch $D(f)$ bzw. $R(f)$ wieder metrische Räume sind.

Wir wiederholen einige aus dem ersten Semester bekannte Aussagen für stetige Funktionen. Seien X und Y metrische Räume. Dann sind für eine Funktion $f: X \rightarrow Y$ die folgenden Aussagen äquivalent:

- (i) f ist stetig.
- (ii) Für alle offenen Teilmengen $V \subset Y$, V ist $f^{-1}(V) \subset X$ offen.
- (iii) Für alle abgeschlossenen Teilmengen $V \subset Y$ ist $f^{-1}(V) \subset X$ abgeschlossen.
- (iv) Für jede Folge $(x_k)_{k \in \mathbb{N}} \subset X$ mit $x_k \rightarrow x \in X$ gilt $f(x_k) \rightarrow f(x)$.

Falls $K \subset X$ kompakt ist und $f: K \rightarrow Y$ stetig ist, so ist der Wertebereich $f(K)$ ebenfalls kompakt und f ist auf K gleichmäßig stetig.

b) Speziell erhalten wir für $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$: Sei $K \subset \mathbb{R}^n$ kompakt und $f: K \rightarrow \mathbb{R}^m$ stetig. Dann ist $f(K)$ kompakt (und damit beschränkt und abgeschlossen), und die Extremwerte von $|f|$ werden angenommen, d.h. $\max_{x \in K} |f(x)|$ und $\min_{x \in K} |f(x)|$ existieren.

1.9 Beispiel. Seien $X = \mathbb{R}^2$ und $Y = \mathbb{R}$; beide Räume seien mit der euklidischen Metrik versehen. Definiere $f: X \rightarrow Y$, $(x, y) \mapsto f(x, y)$ durch

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

Die Funktion f ist in $(x_0, y_0) \neq (0, 0)$ stetig. In $(0, 0)$ ist sie partiell stetig, d.h.

$$\lim_{x \rightarrow 0} f(x, 0) = \lim_{y \rightarrow 0} f(0, y) = 0,$$

aber sie ist nicht stetig, da für $a \neq 0$

$$f(x, ax) = \begin{cases} \frac{a}{1+a^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

ist; die Funktion springt also längs der Geraden $x \mapsto ax$.

Es soll nun ein Ausblick auf ein Analogon zum Zwischenwertsatz gegeben werden. Dazu benötigen wir den Begriff des Zusammenhangs.

1.10 Definition. Ein metrischer Raum X heißt zusammenhängend, falls keine Teilmenge $A \subset X$ existiert mit $\emptyset \subsetneq A \subsetneq X$, welche offen und abgeschlossen ist.

1.11 Bemerkung. Nach Definition ist ein metrischer Raum X genau dann nicht zusammenhängend, wenn zwei nichtleere offene Teilmengen A und B existieren mit $A \cap B = \emptyset$ (d.h. A und B sind disjunkt) und $A \cup B = X$. (Hier kann “offen” durch “abgeschlossen” ersetzt werden.)

1.12 Satz (Zwischenwertsatz). Seien X ein zusammenhängender metrischer Raum, Y ein metrischer Raum und $f : X \rightarrow Y$ eine stetige Abbildung. Dann ist der Wertebereich $R(f)$ zusammenhängend.

Beweis. Angenommen, $R(f)$ ist nicht zusammenhängend. Dann existieren zwei nichtleere offene Mengen A und B mit $R(f) = A \cup B$ und $A \cap B = \emptyset$. (Hier ist die Offenheit im metrischen Raum $R(f)$ zu verstehen.)

Da f stetig ist, sind auch $f^{-1}(A)$ und $f^{-1}(B)$ offen (und nichtleer und disjunkt). Wegen $f^{-1}(A) \cup f^{-1}(B) = f^{-1}(R(f)) = X$, ist X nichtzusammenhängend, Widerspruch. \square

1.13 Beispiele. (i) Der mit der diskreten Metrik versehene Raum $X = \{0, 1\}$ ist nichtzusammenhängend. Wähle $A = \{0\} = B(0, 1/2)$.

(ii) Der mit der euklidischen Metrik versehene Raum $X = \mathbb{R}^n$ ist zusammenhängend.

Eine anschaulichere Definition (die aber nicht äquivalent ist) ist der Wegzusammenhang:

1.14 Definition. Sei X ein metrischer Raum, und seien $a, b \in \mathbb{R}$ mit $a < b$. Dann heißt eine stetige Abbildung $\gamma : [a, b] \rightarrow X$ ein Weg (in X).

Der Raum X heißt wegzusammenhängend, falls je zwei Elemente $x, y \in X$ durch einen Weg in X stetig verbunden werden können, d.h. falls ein Weg $\gamma : [a, b] \rightarrow X$ existiert mit $\gamma(a) = x$ und $\gamma(b) = y$.

Der folgende Satz wird hier nicht bewiesen.

1.15 Satz. Wegzusammenhängende Mengen sind zusammenhängend.

Die Umkehrung dieses Satzes gilt i.a. nicht. Die Menge

$$X := \{(x, y) \in \mathbb{R}^2 \mid 0 < x \leq 1, y = \sin \frac{1}{x}\} \cup \{(0, y) \in \mathbb{R}^2 \mid -1 \leq y \leq 1\},$$

die mit der euklidischen Metrik des \mathbb{R}^2 zu einem metrischen Raum wird, ist ein Beispiel für einen zusammenhängenden, aber nicht wegzusammenhängenden Raum.

Wie auch im ersten Semester, verwenden wir folgende Schreibweise: Seien X und Y normierte \mathbb{K} -Vektorräume, und sei $A: X \rightarrow Y$ eine lineare Abbildung. Dann schreiben wir $Ax := A(x)$. Falls $Y = \mathbb{K}$, heißt eine lineare Abbildung auch ein lineares Funktional.

Eine lineare Abbildung $A: X \rightarrow Y$ ist genau dann stetig, wenn sie beschränkt ist, d.h. falls ein $c > 0$ existiert mit

$$\forall x \in X : \|Ax\|_Y \leq c\|x\|_X.$$

1.16 Definition. Seien X und Y normierte Räume.

Dann definiert man $L(X, Y) := \{A \in L(X, Y) \mid \sup_{\|x\|_X=1} \|Ax\|_Y < \infty\}$, den Raum der beschränkten linearen Abbildungen (Raum der stetigen linearen Operatoren, Raum der beschränkten linearen Operatoren). Für $A \in L(X, Y)$ wird die Operatornorm definiert durch

$$\|A\| := \|A\|_{L(X, Y)} := \sup_{\|x\|_X=1} \|Ax\|_Y = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|_X \leq 1} \|Ax\|_Y.$$

2. Funktionen mehrerer Veränderlicher

2.1 Worum geht's? In diesem Abschnitt werden Funktionen mehrerer reeller Veränderlicher diskutiert; der zentrale Begriff ist hierbei der der „Differenzierbarkeit“. Viele Aussagen aus dem eindimensionalen Fall lassen sich übertragen. Allerdings ist der Begriff der Ableitung selbst jetzt komplizierter. So kann die Ableitung einer Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ an einer Stelle $x \in \mathbb{R}^n$ als Matrix interpretiert werden. Behandelt werden hier unter anderem der Mittelwertsatz, die Taylor-Reihe und die Existenz von Extrema.

a) Differenzierbare Abbildungen

Wir betrachten Funktionen

$$f: \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m.$$

Definiere

$$B(D; \mathbb{R}^m) := \{f: D \rightarrow \mathbb{R}^m \mid \sup_{x \in D} |f(x)| < \infty\}.$$

$B(D; \mathbb{R}^m)$ wird mit der Supremumsnorm $\|f\|_\infty := \sup_{x \in D} |f(x)|$ zu einem normierten Raum.

Genau wie im \mathbb{R}^1 definieren wir

2.2 Definition. Seien $D \subset \mathbb{R}^n$, $f_n: D \rightarrow \mathbb{R}^m$ Funktionen für $n \in \mathbb{N}$. Dann heißt die Funktionenfolge $(f_n)_n$ gleichmäßig konvergent gegen $f: D \rightarrow \mathbb{R}^m$ genau dann, wenn $\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0$ ist.

Komponentenweise zeigt man:

2.3 Satz. *Ist $(f_n)_n \subset C(D; \mathbb{R}^m)$ gleichmäßig konvergent gegen $f: D \rightarrow \mathbb{R}^m$, so ist $f \in C(D; \mathbb{R}^m)$.*

Speziell ergibt sich (wie auch in \mathbb{R}^1) die Vollständigkeit der Räume $B(D; \mathbb{R}^m)$ und $BC(D; \mathbb{R}^m) := C(D; \mathbb{R}^m) \cap B(D; \mathbb{R}^m)$ und $C(K; \mathbb{R}^m)$ für $K \subset \mathbb{R}^n$ kompakt.

Wir werden in diesem Abschnitt die Differentiation von Funktionen

$$f: \mathbb{R}^n \supset U \rightarrow \mathbb{R}^m, \quad U \text{ offen}$$

definieren und diskutieren. Eine direkte Übertragung der Definition für in $U \subset \mathbb{R}^1$ erklärte Funktionen ist nicht möglich, da der Differenzenquotient

$$\frac{f(x+h) - f(x)}{h}$$

für $h \in \mathbb{R}^n$ und $n \geq 2$ nicht erklärt ist.

Einen Hinweis auf eine mögliche Übertragung gibt jedoch die Taylorsche Formel (für $n = m = 1$):

$$f(x+h) = f(x) + f'(x) \cdot h + r(x, h)|h|$$

mit $\lim_{h \rightarrow 0} r(x, h) = 0$.

Mit dieser Darstellung kann man auch die Ableitung f' definieren, und diese Definition werden wir für

$$f: \mathbb{R}^n \supset U \rightarrow \mathbb{R}^m$$

übertragen.

2.4 Definition. (i) Seien $U \subset \mathbb{R}^n$ offen und $V_x := \{h \in \mathbb{R}^n \mid x+h \in U\}$ für $x \in U$. Dann heißt $f: U \rightarrow \mathbb{R}^m$ an der Stelle $x \in U$ differenzierbar, wenn gilt:

Es gibt $A(x) \in L(\mathbb{R}^n, \mathbb{R}^m)$ und eine im Nullpunkt stetige Abbildung $r(x, \cdot): V_x \rightarrow \mathbb{R}^m$ mit $r(x, 0) = 0$ so, dass gilt:

$$f(x+h) = f(x) + A(x)h + r(x, h)|h| \quad \text{für } h \in V_x.$$

(ii) $f: U \rightarrow \mathbb{R}^m$ heißt „in U differenzierbar“, falls f in allen $x \in U$ differenzierbar ist.

(iii) Falls $f: U \rightarrow \mathbb{R}^m$ an der Stelle $x \in U$ differenzierbar ist, heißt $f'(x) := A(x)$ (erste) Ableitung von f an der Stelle x .

Im Falle einer differenzierbaren Funktion heißt die Abbildung

$$f': U \rightarrow L(\mathbb{R}^n, \mathbb{R}^m), \quad x \mapsto f'(x)$$

erste Ableitung von f .

2.5 Bemerkung. Die Ableitung $f'(x) = A(x)$ einer Funktion ist nur dann wohldefiniert, wenn sie eindeutig definiert ist. Wir weisen nach, dass das der Fall ist: Es mögen $A_1(x)$ und $A_2(x)$ die Eigenschaften von $A(x)$ aus der Definition (i) haben. Dann ist

$$(A_1(x) - A_2(x))h + (r_1(x, h) - r_2(x, h))|h| = 0.$$

Hieraus folgt für $y \in \mathbb{R}^n$, $y \neq 0$, und $t \in \mathbb{R} \setminus \{0\}$:

$$\begin{aligned} \left| (A_1(x) - A_2(x)) \frac{y}{|y|} \right| &= \lim_{t \rightarrow 0} \frac{|(A_1(x) - A_2(x))ty|}{|ty|} \\ &= \lim_{t \rightarrow 0} |r_1(x, ty) - r_2(x, ty)| = 0. \end{aligned}$$

Also ist $A_1(x) = A_2(x)$.

Aus der Definition der Ableitung folgt sofort:

2.6 Satz. *Ist f in x differenzierbar, so ist f in x stetig.*

Wir hatten schon bei der Definition der Stetigkeit bemerkt, dass es wesentlich für die Definition ist, dass die Stetigkeit in allen Richtungen gemeint ist.

Für die Stetigkeit in nur einer Richtung hatten wir den Begriff „partielle Stetigkeit“ gebraucht.

Ebenso kann man von „partieller Differenzierbarkeit“ sprechen. Besonders häufig benutzen wir partielle Ableitungen in Richtung der Koordinatenachsen. Wir werden den Zusammenhang zwischen partieller Differenzierbarkeit und Differenzierbarkeit erläutern.

2.7 Definition. Sei $f: \mathbb{R}^n \supset U \rightarrow \mathbb{R}$ eine Funktion und sei $x \in U$. Bezeichne e_i den i -ten Einheitsvektor.

(i) Existiert der Grenzwert

$$\lim_{\mathbb{R} \ni t \rightarrow 0} \frac{f(x + te_i) - f(x)}{h_i},$$

so heißt f an der Stelle x in Richtung des i -ten Einheitsvektors partiell differenzierbar. Wir bezeichnen diese partielle Ableitung mit

$$\frac{\partial f(x)}{\partial x_i} := \lim_{h_i \rightarrow 0} \frac{f(x + h_i e_i) - f(x)}{h_i}.$$

„Partielle Differenzierbarkeit in Richtung des i -ten Einheitsvektors“ wird wie üblich definiert.

(ii) Ist f für $x \in U$ in Richtung aller Einheitsvektoren partiell differenzierbar, so heißt f „an der Stelle x partiell differenzierbar“.

(iii) Der Gradient von $f(x)$ wird für in x partiell differenzierbares f wie folgt definiert:

$$\operatorname{grad} f(x) := \nabla f(x) := \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix} =: \begin{pmatrix} \frac{\partial}{\partial x_1} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{pmatrix} =: \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix}.$$

Das Symbol „ ∇ “ heißt „Nabla“ (von lat. „nablium“, was ein antikes Saiteninstrument bezeichnet).

2.8 Bemerkung. (i) Die partielle Ableitung der Funktion f in Richtung des i -ten Einheitsvektors e_i an der Stelle $x \in U$ entspricht der bereits im ersten Semester

behandelten eindimensionalen Ableitung der Funktion

$$g: \mathbb{R} \supset \tilde{U} \rightarrow \mathbb{R}, \quad t \mapsto f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$$

an der Stelle $t = x_i$. Hierbei ist \tilde{U} das Bild der Projektion der Menge U auf die i -te Komponente. Also ist

$$\frac{\partial f(x)}{\partial x_i} = \frac{d}{dt} g(t)|_{t=x_i}.$$

(ii) Für in x differenzierbares $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ist $f'(x) \in L(\mathbb{R}^n, \mathbb{R})$. Damit existiert ein Zeilenvektor $g = (g_1(x), \dots, g_n(x)) \in \mathbb{R}^{1 \times n}$ mit $f'(x)h = \langle g, h \rangle = \sum_{i=1}^n g_i(x) \cdot h_i$. Setzen wir $h = e_i$, so erhalten wir insbesondere

$$f'(x)e_i = g_i(x).$$

Andererseits ist

$$f'(x)e_i = \frac{f(x + h_i e_i) - f(x)}{h_i} - r(x, h_i e_i) \frac{|h_i e_i|}{h_i} \xrightarrow{h_i \rightarrow 0} \frac{\partial f(x)}{\partial x_i}.$$

Hieraus folgt $g_i(x) = \frac{\partial f(x)}{\partial x_i}$ und damit

$$f'(x) = (\nabla f(x))^\top.$$

Für Funktionen, die in den \mathbb{R}^m abbilden, definiert man analog:

2.9 Definition. Seien $f: \mathbb{R}^n \supset U \rightarrow \mathbb{R}^m$ eine Funktion, $x \in U$ und $e_i \in \mathbb{R}^n$ für $i \in \{1, \dots, n\}$ der i -te Einheitsvektor.

Existieren für $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, m\}$ die Grenzwerte

$$\lim_{h_{ij} \rightarrow 0} \frac{f_j(x + h_{ij} e_i) - f_j(x)}{h_{ij}}, \quad h_{ij} \in \mathbb{R},$$

wobei $f := (f_1, \dots, f_m)^\top$ sei, so heißt die Funktion f „an der Stelle x partiell differenzierbar“.

Die Matrix

$$J_f(x) := \begin{pmatrix} \partial_1 f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_m(x) & \dots & \partial_n f_m(x) \end{pmatrix} =: (\partial_1 f, \dots, \partial_n f)(x)$$

heißt Jacobi-Matrix⁴ von f .

⁴Carl Gustav Jacob Jacobi, 10.12.1804 - 18.2.1851

2.10 Bemerkung. Ebenso wie bei der Interpretation des Gradienten erhalten wir (bei Betrachtung der i -ten Zeile):

$$f'(x) =: (f_{ij}(x))_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \text{ mit } f_{ij}(x) = \frac{\partial f_i(x)}{\partial x_j}.$$

Insgesamt ergibt sich für in x differenzierbare Funktionen der Zusammenhang

$$\forall h \in \mathbb{R}^n : f'(x)h = J_f(x)h = (h \cdot \nabla)f(x) = (h^\top \nabla)f(x),$$

wobei $(h^\top \nabla)$ die Matrixmultiplikation meint. Wir schreiben nun auch ∇f statt J_f .

2.11 Beispiele. (i) Definiere

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad x \mapsto f(x) := \begin{pmatrix} x_1^2 - x_2^2 \\ x_1 x_2 \end{pmatrix}.$$

Wenn f differenzierbar ist, so ist nach obiger Bemerkung

$$f' : \mathbb{R}^2 \rightarrow L(\mathbb{R}^2, \mathbb{R}^2), \quad x \mapsto f'(x) := \begin{pmatrix} 2x_1 & -2x_2 \\ x_2 & x_1 \end{pmatrix}.$$

Es bleibt nachzuweisen, dass diese lineare Abbildung die Eigenschaften der Abbildung $A(x)$ aus Definition 2.5 (ii) erfüllt, was hier aber nicht vorgeführt wird.

(ii) Sei nun

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad x \mapsto f(x) := \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}.$$

Für $x \neq 0$ kann man zeigen: f ist in x differenzierbar, und es gilt:

$$f'(x) = J_f(x) = \frac{1}{|x|^4} (x_2^2(-x_1^2 + x_2^2), 2x_1^3 x_2).$$

Im Nullpunkt ist f jedoch nicht differenzierbar, denn für die partiellen Ableitungen in Richtung der Koordinatenachsen erhält man $\partial_1 f(0) = \partial_2 f(0) = 0$, d.h. es müsste $f'(0) = 0$ sein, aber die Ableitung in Richtung des Vektors $(1, 1)$ ergibt:

$$\lim_{t \rightarrow 0} \frac{f(t(e_1 + e_2))}{t} = \lim_{t \rightarrow 0} \frac{f(t, t)}{t} = \frac{1}{2} \neq 0.$$

Es gibt noch eine weitere mögliche Interpretation der Ableitung f' einer differenzierbaren Funktion. Für

$$f : \mathbb{R}^n \supset U \rightarrow \mathbb{R}^m, \quad x \mapsto f(x)$$

hatten wir bisher gesagt:

$$f' : U \rightarrow L(\mathbb{R}^n, \mathbb{R}^m), \quad x \mapsto f'(x).$$

Ebenso kann man sagen:

$$f' : U \times \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad (x, h) \mapsto f'(x, h) := f'(x)h,$$

wobei f' linear in der zweiten Komponente ist.

Die Menge der differenzierbaren Abbildungen bildet einen Vektorraum. Mit $C^1(U; \mathbb{R}^m)$ bezeichnen wir den Raum der differenzierbaren Abbildungen $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, deren Ableitung f' wieder stetig ist.

Es gelten die folgenden Rechenregeln

2.12 Lemma. (i) Seien $f, g: \mathbb{R}^n \supset U \rightarrow \mathbb{R}$ differenzierbare Funktionen.

Dann ist auch $f \cdot g: U \rightarrow \mathbb{R}$ eine differenzierbare Funktion, und es gilt die Produktregel

$$\forall x \in U : (f \cdot g)'(x) = g(x)f'(x) + f(x) \cdot g'(x).$$

(ii) Seien $f: \mathbb{R}^n \supset U \rightarrow \mathbb{R}^m$ und $g: \mathbb{R}^m \supset V \rightarrow \mathbb{R}^p$ differenzierbare Funktionen mit $f(U) \subset V$. Dann ist auch die Verkettung $g \circ f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ eine differenzierbare Funktion, und es gilt die Kettenregel

$$\forall x \in U : (g \circ f)'(x) = (g' \circ f)(x) \cdot f'(x) \in L(\mathbb{R}^n, \mathbb{R}^p).$$

Beweis. Die erste Behauptung ist klar.

Zu (ii): Sei für $x \in U$ und $h \in \mathbb{R}^n$ (mit $x + h \in U$)

$$H := H(x, h) := f(x + h) - f(x) = f'(x)h + r_f(x, h)|h|.$$

Dann ist

$$\begin{aligned} g(f(x + h)) - g(f(x)) &= g'(f(x))H + r_g(f(x), H)|H| \\ &= g'(f(x))f'(x)h + g'(f(x))r_f(x, h)|h| + r_g(f(x), H)|H| \\ &=: g'(f(x))f'(x)h + r_{g \circ f}(x, h) \cdot |h| \end{aligned}$$

mit

$$\begin{aligned} r_{g \circ f}(x, h) &:= g'(f(x))r_f(x, h) + r_g(f(x), H) \frac{|H|}{|h|} \\ &\rightarrow 0 \text{ für } h \rightarrow 0. \end{aligned}$$

□

Wir wollen nun eine anschauliche Interpretation des Gradienten geben. Hierzu benötigen wir den Begriff der Richtungsableitung.

2.13 Definition. Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare Funktion, $x_0, h_0 \in \mathbb{R}^n$ mit $|h_0| = 1$ und

$$F : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto F(t) := f(x_0 + th_0).$$

Dann ist F eine differenzierbare Funktion und

$$F'(0) = \lim_{t \rightarrow 0} \frac{f(x_0 + th_0) - f(x_0)}{t} = f'(x_0)h_0 = \langle \nabla f(x_0), h_0 \rangle$$

heißt Richtungsableitung von f in Richtung h_0 an der Stelle x_0 .

In einem Punkt x_0 ist die Richtungsableitung betragsmäßig maximal, falls $\nabla f(x_0)$ in Richtung von h_0 zeigt. Das bedeutet, dass die Richtung des Gradienten die Richtung der stärksten Änderung der Funktion f angibt. Der Betrag der Richtungsableitung in Richtung des Gradienten entspricht dem Betrag des Gradienten.

Wir betrachten nun speziell Funktionen $f : \mathbb{R}^2 \supset U \rightarrow \mathbb{R}$. In diesem Fall kann man sich den Graphen von f als Fläche im dreidimensionalen Raum vorstellen. Die Mengen der Punkte im \mathbb{R}^2 , in denen f einen konstanten Wert $c \in \mathbb{R}$ annimmt,

$$N_f(c) := \{x \in U : f(x) = c\}$$

heißen Höhenlinien. Nun definiere für $c \in \mathbb{R}$ eine Abbildung

$$x : \mathbb{R} \rightarrow N_f(c), \quad t \mapsto x(t) = (x_1(t), x_2(t)).$$

Da $f(x(t)) = \text{konstant}$ für alle $t \in \mathbb{R}$ ist, gilt:

$$\begin{aligned} 0 &= \frac{d}{dt} f(x(t)) = \partial_1 f(x(t))x'_1(t) + \partial_2 f(x(t))x'_2(t) \\ &= \langle \nabla f(x(t)), x'(t) \rangle. \end{aligned}$$

Das bedeutet, dass entweder $\nabla f(x(t)) = 0$ oder $x'(t) = 0$ sind, oder aber, dass die beiden Vektoren senkrecht aufeinander stehen.

Abbildungen

$$\beta : \mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto \beta(t)$$

heißen Falllinien, wenn gilt: $\frac{d}{dt}\beta(t) = \nabla f(\beta(t))$. Wir haben gerade gezeigt, dass Höhenlinien und Falllinien senkrecht aufeinander stehen.

2.14 Beispiel. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $x \mapsto f(x) := |x|^2 = x_1^2 + x_2^2$.

Die Funktion f ist differenzierbar, und es gilt für $x_0 \in \mathbb{R}^2$, $x_0 \neq 0$:

$$\nabla f(x_0) = 2x_0 = 2|x_0| \cdot \tilde{x}_0 \quad \text{mit } \tilde{x}_0 = \frac{x_0}{|x_0|}.$$

f ändert sich somit in radialer Richtung am meisten; die Falllinien sind Ursprungsgeraden, die Höhenlinien konzentrische Kreise.

Einige Ableitungen haben aufgrund ihrer Bedeutung eine besondere Bezeichnung erhalten.

2.15 Definition. (i) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine differenzierbare Funktion. Dann heißt die Abbildung

$$\operatorname{div} f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \operatorname{div} f(x) := \sum_{i=1}^n \partial_i f_i(x)$$

Divergenz von f .

(ii) Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ eine differenzierbare Funktion. Dann heißt die Abbildung

$$\operatorname{rot} f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad x \mapsto \operatorname{rot} f(x) := \begin{pmatrix} \partial_2 f_3(x) - \partial_3 f_2(x) \\ \partial_3 f_1(x) - \partial_1 f_3(x) \\ \partial_1 f_2(x) - \partial_2 f_1(x) \end{pmatrix}$$

Rotation von f .

(iii) Der Operator

$$\begin{aligned} \Delta : C^2(\mathbb{R}^n; \mathbb{R}) &\rightarrow C(\mathbb{R}^n; \mathbb{R}) \\ f &\mapsto \Delta f := \sum_{i=1}^n \partial_i^2 f, \end{aligned}$$

also $\Delta = \operatorname{div} \operatorname{grad}$, heißt Laplace-Operator⁵. Für beliebiges $m \in \mathbb{N}$ definiert man

$$\begin{aligned} \Delta : C^2(\mathbb{R}^n, \mathbb{R}^m) &\rightarrow C(\mathbb{R}^n, \mathbb{R}^m) \\ f = (f_1, \dots, f_m)^\top &\mapsto (\Delta f_1, \dots, \Delta f_m)^\top. \end{aligned}$$

2.16 Beispiel. Für $f : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}^3$, $x \mapsto f(x) := \frac{x}{|x|^3}$ und $g : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}$, $x \mapsto g(x) := -\frac{1}{|x|}$ sind $f = \nabla g$, $\Delta g = \operatorname{div} f = 0$ und $\operatorname{rot} f = \operatorname{rot} \nabla g = 0$.

Zum Abschluss dieses Abschnitts untersuchen wir (im Vorgriff auf Abschnitt 6), wann die Umkehrabbildung einer differenzierbaren Funktion, wenn sie existiert, auch wieder differenzierbar ist.

2.17 Satz. Seien $U \subset \mathbb{R}^n$ offen, $a \in U$ und $f : U \rightarrow \mathbb{R}^n$ eine injektive Abbildung, die in a differenzierbar ist mit $\det f'(a) \neq 0$. Seien $b := f(a)$ ein innerer Punkt von $f(U)$ und

$$g := f^{-1} : f(U) \rightarrow \mathbb{R}^n$$

stetig in b .

⁵Pierre Simon Laplace, 28.3.1749 – 5.3.1827

Dann ist g an der Stelle b differenzierbar, und für die Ableitung gilt:

$$g'(b) = (f'(a))^{-1}.$$

2.18 Bemerkung. Die Bedingungen „ $b = f(a)$ ist innerer Punkt von $f(U)$ “ und „ g ist stetig in b “ folgen aus den sonstigen Voraussetzungen, wie wir später beweisen werden. Auch die Injektivität von f folgt aus der Bedingung $\det f'(a) \neq 0$, wobei die Umgebung U eventuell verkleinert werden muss.

Beweis von Satz 2.17. Für $x \in \mathbb{R}^n$, $|x|$ genügend klein, ist

$$\begin{aligned} f(a+x) &= f(a) + f'(a)x + r(a,x)|x| \\ &= f(a) + F(a,x) \cdot x \\ \text{mit } F(a,x) &:= \begin{cases} f'(a) + r(a,x)x_0^\top & \text{für } x \neq 0 \\ f'(a) & \text{für } x = 0, \end{cases} \end{aligned}$$

wobei $x_0 := \frac{x}{|x|}$. Beachte $|x| = \langle \frac{x}{|x|}, x \rangle = x_0^\top \cdot x$.

Die Abbildung $F(a, \cdot)$ hat die folgenden Eigenschaften:

(i) $F(a, \cdot) : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R}^n)$ ist stetig in $x = 0$.

Beweis: Es gilt

$$\begin{aligned} \|F(a,x) - F(a,0)\| &= \|r(a,x)x_0^\top\| = \sup_{\|h\|=1} |r(a,x)| |x_0 \cdot h| \\ &\leq |r(a,x)| \rightarrow 0 \text{ für } |x| \rightarrow 0. \end{aligned}$$

(ii) $F(a, \cdot)$ ist in einer Umgebung $U(0)$ des Nullpunktes invertierbar und

$$\lim_{x \rightarrow 0} (F(a,x))^{-1} = (f'(a))^{-1}.$$

Beweis: Die Invertierbarkeit von $F(a,x)$ für $x \in U(0)$, wobei $U(0)$ eine geeignete Umgebung des Nullpunktes bezeichnet, ist klar, da wegen $\det f'(a) \neq 0$ für $x \in U(0)$ gilt: $\det F(a,x) \neq 0$. Somit existiert die Abbildung $F(a, \cdot)^{-1}$ in $U(0)$; die Stetigkeit dieser Abbildung folgt aus der Identität

$$F(a, x+h)^{-1} - F(a, x)^{-1} = F(a, x+h)^{-1} (F(a, x) - F(a, x+h)) F(a, x)^{-1}.$$

Wähle nun für $x \in \mathbb{R}^n$, $|x|$ genügend klein, y so, dass

$$f(a+x) = b+y, \text{ also } a+x = g(b+y).$$

Bezeichnet $G(y) := (F(a,x))^{-1}$, so ist

$$G(y)y = G(y)(f(a+x) - f(a)) = G(y)(F(a,x)x) = x,$$

und somit ist

$$g(b+y) = g(b) + G(y)y.$$

Aus der Stetigkeit von g folgt:

$$\lim_{y \rightarrow 0} G(y) = \lim_{y \rightarrow 0} (F(a, g(b+y) - g(b))^{-1} = (F(a, 0))^{-1} = (f'(a))^{-1}.$$

Aus den beiden letzten Gleichungen folgt die Existenz der Ableitung von g in b und auch die Identität

$$g'(b) = (f'(a))^{-1}.$$

□

Eine wichtige Anwendung obigen Satzes sind Koordinatentransformationen.

2.19 Beispiel (Polarkoordinaten). Definiere die Koordinatentransformation

$$f : \mathbb{R}^+ \times (0, 2\pi) \rightarrow \mathbb{R}^2, \quad (r, \varphi) \mapsto f(r, \varphi) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cdot \cos \varphi \\ r \cdot \sin \varphi \end{pmatrix}.$$

Die Abbildung f ist injektiv und für die Umkehrabbildung $f^{-1} = g$ erhalten wir

$$g : \mathbb{R}^2 \setminus (\mathbb{R}_0^+ \times \{0\}) \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto g(x, y) = \begin{pmatrix} r \\ \varphi \end{pmatrix}$$

$$\text{mit } r := \sqrt{x^2 + y^2} \text{ und } \varphi := \begin{cases} \arctan(y/x) & \text{für } x > 0, y > 0 \\ \pi/2 & \text{für } x = 0, y > 0 \\ \pi + \arctan(y/x) & \text{für } x < 0 \\ 3\pi/2 & \text{für } x = 0, y < 0 \\ 2\pi + \arctan(y/x) & \text{für } x > 0, y < 0. \end{cases}$$

Nun ist f im ganzen Definitionsbereich differenzierbar mit

$$\det f'(r, \varphi) = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r > 0 \quad \forall (r, \varphi) \in \mathbb{R}^+ \times (0, 2\pi).$$

Auch g ist überall differenzierbar, und wir erhalten:

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

Differenzierbare und injektive Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit invertierbarer Ableitung spielen z.B. für Koordinatentransformationen eine wichtige Rolle und erhalten einen eigenen Namen.

2.20 Definition. Seien $U \subset \mathbb{R}^n$ offen und $f \in C^1(U, \mathbb{R}^n)$ eine invertierbare Funktion mit offenem Wertebereich $V = f(U)$ und stetig differenzierbarer Umkehrabbildung. Dann heißt f Diffeomorphismus von U auf V .

Die Voraussetzung „ $f(U)$ offen“ ist entbehrlich in dieser Definition, wie wir später sehen werden.

b) Der Mittelwertsatz

Die Übertragung des Mittelwertsatzes (Analysis I) auf reellwertige Funktionen im \mathbb{R}^n ist möglich, wenn die lineare Verbindung der beiden betrachteten Punkte a und b , also die Menge

$$\Gamma_{ab} := \{x \in \mathbb{R}^n \mid \exists t \in (0, 1) : x = \gamma_{ab}(t) := a + t(b - a)\},$$

in der Definitionsmenge liegt.

2.21 Satz (Mittelwertsatz). Seien $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ eine differenzierbare Funktion, $a, b \in U$ und $\Gamma_{ab} \subset U$ (Γ_{ab} wie oben definiert).

Dann gibt es ein $c \in \Gamma_{ab}$ mit

$$f(b) = f(a) + f'(c)(b - a).$$

Beweis. Erkläre die Abbildung $F : [0, 1] \rightarrow \mathbb{R}$, $t \mapsto F(t)$ durch $F(t) := f(\gamma_{ab}(t))$. Aus dem Mittelwertsatz für in \mathbb{R} erklärte Funktionen folgt:

$$\begin{aligned} \exists \theta \in (0, 1) : F(1) &= F(0) + F'(\theta) \text{ und} \\ F'(\theta) &= f'(\gamma_{ab}(\theta)) \cdot \gamma'_{ab}(\theta) = f'(\gamma_{ab}(\theta))(b - a). \end{aligned}$$

Mit $c := \gamma_{ab}(\theta)$ folgt die Behauptung. □

2.22 Bemerkungen. (i) Wie der Beweis zeigt, kann die lineare Verbindung der Punkte durch einen beliebigen differenzierbaren in U verlaufenden Weg ersetzt werden, d.h. durch eine differenzierbare Abbildung

$$\gamma : [0, 1] \rightarrow U \text{ mit } \gamma(0) = a \text{ und } \gamma(1) = b.$$

Dann gilt: $f(b) = f(a) + f'(\gamma(\theta))\gamma'(\theta)$ mit einem $\theta \in (0, 1)$.

(ii) Eine andere Formulierung des Satzes lautet: Für $x \in U$ und $h \in \mathbb{R}^n$ mit $|h|$ so klein, dass $x + th \in U$ für $t \in [0, 1]$ ist, gilt:

$$\exists \theta \in (0, 1) : f(x + h) = f(x) + f'(x + \theta h)h.$$

(iii) Für vektorwertige Funktionen $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ kann der Mittelwertsatz komponentenweise angewandt werden; dann erhält man aber in jeder Komponente einen anderen „Mittelwert“ θ_j .

Wie schon für in \mathbb{R} erklärte Funktionen gilt:

2.23 Satz. Seien $U \subset \mathbb{R}^n$ offen und wegzusammenhängend und $f : U \rightarrow \mathbb{R}$ differenzierbar. Dann gilt:

$$f' = 0 \quad \Leftrightarrow \quad f = \text{const.}$$

Beweis. Zu zeigen ist nur „ \Rightarrow “. Seien $a, b \in U$; da U wegzusammenhängend und offen ist, gibt es $k \in \mathbb{N}$ und $k + 1$ Punkte $a_j \in U$, $a_0 = a$, $a_k = b$, so, dass für $j \in \{0, \dots, k-1\}$ die lineare Verbindung von a_j und a_{j+1} in U liegt. Der Mittelwertsatz zwischen a_j und a_{j+1} liefert: $f(a_j) = f(a_{j+1})$. Also ist $f(a) = f(b)$. \square

Wir haben bereits an einem Beispiel gesehen, dass die Existenz der partiellen Ableitungen allein noch nicht die Differenzierbarkeit impliziert. Es gilt jedoch:

2.24 Satz. Seien $U \subset \mathbb{R}^n$ offen, $a \in U$ und $f : U \rightarrow \mathbb{R}^m$. Wenn die partiellen Ableitungen $\partial_1 f, \dots, \partial_n f$ in einer Umgebung $U(a)$ von a existieren und in a stetig sind, dann ist die Funktion f in a differenzierbar.

Beweis. Der Beweis kann komponentenweise geführt werden, daher können wir o.E. $m = 1$ annehmen.

Seien $a, h = (h_1, \dots, h_n) \in \mathbb{R}^n$, $a \in U$ und $|h|$ so klein, dass die „Zwischenpunkte“:

$$\begin{aligned} a_0 &:= a \\ a_1 &:= a_0 + h_1 \cdot e_1 \\ a_2 &:= a_1 + h_2 \cdot e_2 \\ &\vdots \\ a_n &:= a_{n-1} + h_n \cdot e_n = a + h \end{aligned}$$

und deren lineare Verbindungen in U liegen; hierbei bezeichnet e_i den i -ten kanonischen Einheitsvektor.

Dann erhalten wir mit dem Mittelwertsatz:

$$\begin{aligned} f(a+h) - f(a) &= (f(a_n) - f(a_{n-1})) + (f(a_{n-1}) - f(a_{n-2})) + \dots + (f(a_1) - f(a_0)) \\ &= \sum_{j=1}^n \partial_j f(c_j) \cdot h_j \quad \text{mit } c_j \in B(a, |h|). \end{aligned}$$

Insbesondere erhalten wir:

$$\begin{aligned} & |f(a+h) - f(a) - \sum_{j=1}^n (\partial_j f)(a) h_j| \\ &= \left| \sum_{j=1}^n (\partial_j f(c_j) - \partial_j f(a)) h_j \right| \leq |h| \cdot \sum_{j=1}^n |\partial_j f(c_j) - \partial_j f(a)|; \end{aligned}$$

für $h \rightarrow 0$ gilt: $c_j \rightarrow a$, also: $\sum_{j=1}^n |\partial_j f(c_j) - \partial_j f(a)| \rightarrow 0$.
Somit existiert $f'(a)$, und es ist

$$f'(a)h = \sum_{j=1}^n (\partial_j f)(a) h_j.$$

□

2.25 Bemerkung. Aus obigem Beweis folgt insbesondere: Sind die partiellen Ableitungen $\partial_j f$ in einer Umgebung $U(a)$ von a stetig, so existiert f' in $U(a)$ und ist dort stetig.

c) Höhere Ableitungen

Seien $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung. Dann gibt es zwei Interpretationsmöglichkeiten der Ableitung von f :

$$\begin{aligned} f' : U &\rightarrow L(\mathbb{R}^n, \mathbb{R}^m), \quad x \mapsto f'(x), \text{ oder} \\ f' : U \times \mathbb{R}^n &\rightarrow \mathbb{R}^m, \quad (x, h) \mapsto f'(x, h) := f'(x)h \end{aligned}$$

und f' ist linear in der zweiten Variablen.

Der $(n \cdot m)$ -dimensionale Vektorraum $L(\mathbb{R}^n, \mathbb{R}^m)$ entspricht, bezogen auf Standardbasen, dem Raum der $(n \times m)$ -Matrizen und ist isomorph zum $\mathbb{R}^{n \cdot m}$.

Wir können nun analog der obigen Interpretationen höhere Ableitungen definieren (sei $f'' := (f')'$):

$$\begin{aligned} f'' : U &\rightarrow L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m)), \\ f''' : U &\rightarrow L(\mathbb{R}^n, L(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^m))) \end{aligned}$$

usw., oder eben

$$f'' : U \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad (x, h, k) \mapsto f''(x, h, k) = (f''(x)k)h,$$

wobei die Abbildung linear in den letzten beiden Variablen ist, usw.

Für die p -ten Ableitungen $f^{(p)}$ (einer p -mal differenzierbaren Funktion f) ergibt sich analog:

$$f^{(p)} : U \times (\mathbb{R}^n)^p \rightarrow \mathbb{R}^m, \quad (x, h^1, \dots, h^p) \mapsto f^{(p)}(x, h^1, \dots, h^p),$$

und diese Abbildung ist linear in den letzten p Variablen.

Für die erste Ableitung einer Funktion f hatten wir in Bezug auf die Standardbasis gezeigt:

$$f'(x, h) = \sum_{i=1}^n (\partial_i f)(x) h_i = (h^\top \nabla) f(x).$$

Für die zweite Ableitung erhalten wir entsprechend:

$$\begin{aligned} f''(x, h, k) &= \sum_{j=1}^n \partial_j f'(x, h) k_j = \sum_{i,j=1}^n \partial_i \partial_j f(x) h_i k_j \\ &= (k^\top \nabla) (h^\top \nabla) f(x). \end{aligned}$$

Induktiv ergibt sich (für $h = (h^1, \dots, h^p) \in (\mathbb{R}^n)^p$):

$$f^{(p)}(x, h) = ((h^1)^\top \nabla) \cdot \dots \cdot ((h^p)^\top \nabla) f(x).$$

Mit der zweiten Ableitung f'' existieren also auch die zweiten partiellen Ableitungen, die definiert sind durch

$$\partial_j \partial_i f(x) = f''(x, e_i, e_j)$$

(e_i bezeichne den i -ten Einheitsvektor der Standardbasis).

f'' ist symmetrisch in (h, k) :

2.26 Satz (Satz von Schwarz⁶). Seien $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ zweimal differenzierbar. Dann gilt für alle $x \in U$:

$$a) \quad \forall h, k \in \mathbb{R}^n : f''(x, h, k) = f''(x, k, h).$$

$$b) \quad \forall i, j \in \{1, \dots, n\} : \partial_i \partial_j f(x) = \partial_j \partial_i f(x).$$

Beweis. Es genügt, die erste Behauptung zu beweisen; die zweite Aussage folgt mit $h := e_i$ und $k := e_j$ sofort aus dieser.

Wir zeigen:

$$(*) \quad \lim_{s \searrow 0} \frac{1}{s^2} \{f(x + sh + sk) - f(x + sh) - f(x + sk) + f(x)\} = f''(x, h, k).$$

Die linke Seite der Gleichung (*) ist in h und k symmetrisch, so dass die Aussage a) sofort aus (*) folgt.

Beweis von (*):

Sei o.B.d.A. $m = 1$ (es genügt, die Gleichung komponentenweise zu beweisen).

Definiere

$$F : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto F(t) := f(x + th + k) - f(x + th).$$

⁶Hermann Amandus Schwarz, 1843 - 1921

F ist in $(0, 1)$ differenzierbar mit

$$F'(t) = f'(x + th + k, h) - f'(x + th, h).$$

Für

$$A(x, h, k) := F(1) - F(0) = f(x + h + k) - f(x + h) - f(x + k) + f(x)$$

ergibt sich mit dem Mittelwertsatz:

$$\exists \theta \in (0, 1) : A(x, h, k) = F'(\theta), \quad \text{also}$$

$$A(x, h, k)$$

$$\begin{aligned} &= (f'(x + \theta h + k, h) - f'(x, h)) - (f'(x + \theta h, h) - f'(x, h)) \\ &= f''(x, h, \theta h + k) + R(x, h, \theta h + k)|\theta h + k| - f''(x, h, \theta h) - R(x, h, \theta h)|\theta h| \\ &= f''(x, h, k) + R(x, h, \theta h + k) \cdot |\theta h + k| - R(x, h, \theta h)|\theta h|. \end{aligned}$$

Ersetze nun für $s > 0$ h bzw. k durch $s \cdot h$ bzw. $s \cdot k$.

Dann folgt (wiederum aus der Linearität der 2. Ableitung in den letzten beiden Argumenten):

$$A(x, sh, sk) = s^2 \{ f''(x, h, k) + R(x, h, s(\theta h + k)) \cdot |\theta h + k| - R(x, h, s\theta h)|\theta h| \}.$$

(*) folgt nun aus

$$\lim_{s \searrow 0} \frac{A(x, sh, sk)}{s^2} = f''(x, h, k).$$

□

2.27 Bemerkungen. (i) Die Symmetrie von $f''(x, \cdot, \cdot)$ an einer festen Stelle $x \in U$ in den beiden letzten Argumenten gilt auch dann, wenn f in einer Umgebung von x einmal differenzierbar ist und die zweite Ableitung nur an der Stelle x existiert; dies folgt sofort aus dem Beweis des Satzes.

(ii) Wiederum (vgl. Bemerkung 2.25) folgt aus der Stetigkeit aller zweiten partiellen Ableitungen die Existenz von f'' , und f'' ist wieder stetig in x , falls die partiellen Ableitungen in einer Umgebung von x existieren und stetig sind.

Wie schon im \mathbb{R}^1 gilt auch für $U \subset \mathbb{R}^n$ die Taylorsche Formel:

2.28 Satz (von Taylor). Seien $U \subset \mathbb{R}^n$ offen, $x \in U$, $h \in \mathbb{R}^n$ mit $|h|$ so klein, dass die Verbindungsstrecke von x nach $x + h$ in U liegt, und $f \in C^p(U, \mathbb{R})$, wobei $f^{(p)}$ differenzierbar sei.

Dann gilt:

$$f(x + h) = f(x) + f'(x, h) + \cdots + \frac{1}{p!} f^{(p)}(x, \underbrace{h, \dots, h}_{p\text{-mal}}) + R_p(x, h)$$

$$\text{mit } R_p(x, h) = \frac{1}{(p+1)!} f^{(p+1)}(x + \theta h, \underbrace{h, \dots, h}_{(p+1)\text{-mal}}) \quad \text{für ein } \theta \in (0, 1).$$

Beweis. Seien

$$\begin{aligned} \gamma : [0, 1] &\rightarrow \mathbb{R}^n, \quad t \mapsto \gamma(t) := x + th \quad \text{und} \\ F : [0, 1] &\rightarrow \mathbb{R}, \quad t \mapsto F(t) := f(\gamma(t)). \end{aligned}$$

$$F(1) = F(0) + F'(0) + \dots + \frac{1}{p!} F^{(p)}(0) + \frac{1}{(p+1)!} F^{(p+1)}(\theta)$$

für ein $\theta \in (0, 1)$.

Aus

$$\begin{aligned} F'(t) &= f'(\gamma(t), \gamma'(t)) = f'(\gamma(t), h), \\ F''(t) &= f''(\gamma(t), h, \gamma'(t)) = f''(\gamma(t), h, h), \\ &\vdots \\ F^{(p)}(t) &= f^{(p)}(\gamma(t), \underbrace{h, \dots, h}_{p\text{-mal}}) \quad \text{und} \\ F(1) &= f(x+h) \quad \text{und} \quad F(0) = f(x) \end{aligned}$$

folgt die Behauptung. □

Für stetiges $f^{(p+1)}$ erhalten wir (mit demselben Argument) für das Restglied die Darstellung

$$R_p(x, h) = \frac{1}{p!} \int_0^1 (1-t)^p F^{(p+1)}(t) dt = \frac{1}{p!} \int_0^1 (1-t)^p f^{(p+1)}(x + th, \underbrace{h, \dots, h}_{(p+1)\text{-mal}}) dt.$$

Vektorwertige Funktionen f können wir komponentenweise mit der Taylor-Entwicklung darstellen; die Zwischenstelle in der Restglieddarstellung ist dann jedoch von Komponente zu Komponente verschieden ($\theta = \theta_j$, $j = 1, \dots, m$).

Der Begriff der „Taylorreihe“ überträgt sich für beliebig oft differenzierbare Funktionen analog.

Wir führen nun noch eine bequeme Schreibweise für die Ableitung ein.

Es ist

$$f^{(k)}(x, \underbrace{h, \dots, h}_{k\text{-mal}}) = \sum_{i_1, \dots, i_k=1}^n \partial_{i_1} \dots \partial_{i_k} f(x) h_{i_1} \dots h_{i_k}.$$

Wegen der Vertauschbarkeit der partiellen Ableitungen kann man von den n^k Summanden einige zusammenfassen; die „Multiindizes“ ermöglichen eine elegante Schreibweise, die an den \mathbb{R}^1 erinnert. Für $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N}_0)^n$ sei

$$\begin{aligned}\partial^\alpha &:= \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}, \\ h^\alpha &:= h_1^{\alpha_1} \cdots h_n^{\alpha_n}, \\ \alpha! &:= (\alpha_1!) \cdots (\alpha_n!) \text{ und} \\ |\alpha| &:= \alpha_1 + \cdots + \alpha_n.\end{aligned}$$

Bei vorgegebenem α , $k := |\alpha|$, gibt es insgesamt $\frac{k!}{(\alpha_1!) \cdots (\alpha_n!)} = \frac{k!}{\alpha!}$ Summanden, die den Faktor

$$\partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f(x) h_1^{\alpha_1} \cdots h_n^{\alpha_n} = \partial^\alpha f(x) h^\alpha$$

enthalten. In dieser Schreibweise ist also

$$\frac{1}{k!} f^{(k)}(x, \underbrace{h, \dots, h}_{k\text{-mal}}) = \sum_{|\alpha|=k} \frac{1}{\alpha!} \partial^\alpha f(x) h^\alpha,$$

und die Taylorformel wird zu

$$\begin{aligned}f(x+h) &= \sum_{k=0}^p \sum_{|\alpha|=k} \frac{1}{\alpha!} \partial^\alpha f(x) h^\alpha + R_p(x, h) \\ &= \sum_{\substack{|\alpha|=0 \\ \alpha \in \mathbb{N}_0^n}}^p \frac{1}{\alpha!} \partial^\alpha f(x) h^\alpha + R_p(x, h)\end{aligned}$$

Wie im Falle einer Veränderlichen lassen sich nun Kriterien für die Existenz lokaler Extrema angeben.

2.29 Definition. Seien $U \subset \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ differenzierbar. Dann heißt $a \in U$ kritische Stelle von f , wenn $f'(a, h) = 0$ ($h \in \mathbb{R}^n$) gilt (d.h. wenn $\nabla f(a) = 0$).

2.30 Satz. Sei a lokale Extremalstelle der differenzierbaren Funktion f . Dann ist a kritische Stelle von f .

Beweis. Sei o.B.d.A. a eine lokale Maximalstelle, d.h.: $\forall x \in V : f(x) \leq f(a)$, wobei $V = V(a)$ eine Umgebung des Punktes a bezeichnet. Sei $\varepsilon > 0$ so klein, dass $B(a, \varepsilon) \subset V$.

Für $h \in B(0, \varepsilon) \subset \mathbb{R}^n$ definiere

$$F : [-1, 1] \rightarrow \mathbb{R}, \quad t \mapsto F(t) := f(a + th).$$

Dann ist für alle $t \in [-1, 1] : F(t) \leq f(a) = F(0)$, also $0 = F'(0) = f'(a, h)$.

Aus der Linearität der Ableitung in der zweiten Komponente folgt hieraus die Behauptung. \square

Im Folgenden wollen wir die Entscheidung, ob ein Extremum vorliegt, mit Hilfe der zweiten Ableitungen treffen. Dazu folgende Definition:

2.31 Definition. Sei $U \subset \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ zweimal differenzierbar. Dann heißt

$$\text{Hess}_f(x) := (\partial_i \partial_j f(x))_{i,j=1,\dots,n}$$

die Hessematrix⁷ von f an der Stelle $x \in U$.

Nach dem Satz von Schwarz 2.26 ist diese Matrix symmetrisch. Für eine skalarwertige Funktion $f: U \rightarrow \mathbb{R}$ erhalten wir damit folgende Darstellung der Taylorformel:

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle \text{Hess}_f(x)h, h \rangle + R_3(x, h).$$

Die Rolle der zweiten Ableitungen spielen nun „quadratische Formen“

$$Q(h, h) := \sum_{i=1}^n \sum_{j=1}^n q_{ij} h_i h_j,$$

mit $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, $q_{ij} \in \mathbb{R}$.

Die quadratische Form $Q(\cdot, \cdot)$ heißt

„positiv (semi)definit“, falls gilt: $\forall h \neq 0: Q(h, h) >_{(=)} 0$,

„negativ (semi)definit“, falls gilt: $\forall h \neq 0: Q(h, h) <_{(=)} 0$ und

„indefinit“, falls gilt: $\exists h^1, h^2: Q(h^1, h^1) < 0 \wedge Q(h^2, h^2) > 0$.

Hierzu zwei Bemerkungen:

- a) Aus der Definition erhält man sofort folgende Aussage: Falls Q positiv definit ist, so existiert ein $p > 0$ mit

$$Q(h, h) \geq p|h|^2 \quad (h \in \mathbb{R}^n).$$

- b) Ist die Matrix $Q = (q_{ij})_{i,j=1,\dots,n}$ symmetrisch — was für $q_{ij} = \partial_i \partial_j f$ der Fall ist — so ist die Matrix diagonalisierbar; es gibt also eine Basis des \mathbb{R}^n aus Eigenvektoren von Q zu reellen Eigenwerten. Die Matrix Q ist genau dann positiv definit, falls alle Eigenwerte positiv sind, und genau dann positiv semidefinit, falls alle Eigenwerte nichtnegativ sind. Analoge Aussagen gelten für negativ definit.

Die uns interessierende quadratische Form ist gegeben durch

$$f''(x, h, h) = \langle \text{Hess}_f(x)h, h \rangle.$$

⁷Ludwig Otto Hesse, 22.4.1811 – 4.8.1874

2.32 Satz. Seien $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine zweimal stetig differenzierbare Funktion und $a \in U$ kritische Stelle von f . Dann gilt:

Ist $\text{Hess}_f(a)$ positiv / negativ definit, so ist a Minimal- / Maximalstelle von f . Falls $\text{Hess}_f(a)$ indefinit ist, ist a keine Extremalstelle von f .

Beweis. Aus der Taylorschen Formel und der zweimaligen stetigen Differenzierbarkeit von f ergibt sich für $h \in \mathbb{R}^n$, $|h|$ genügend klein, und $\theta \in (0, 1)$:

$$\begin{aligned} f(a+h) &= f(a) + f'(a) + \frac{1}{2}f''(a+\theta h, h, h) \\ &= f(a) + \frac{1}{2}f''(a, h, h) + r(a, \theta h)|h|^2 \end{aligned}$$

mit $|r(a, \theta h)| \rightarrow 0$ für $|h| \rightarrow 0$.

Die positive Definitheit von $f''(a, \cdot, \cdot)$ bedeutet:

$$\exists p > 0 : f''(a, h, h) \geq p|h|^2 \quad (h \in \mathbb{R}^n).$$

Wähle $\varepsilon > 0$ so klein, dass für alle $h \in \mathbb{R}^n$ mit $|h| < \varepsilon$ gilt: $|r(a, h)| < p/4$.
Dann folgt:

$$\forall h \in \mathbb{R}^n \setminus \{0\}, \quad |h| < \varepsilon : f(a+h) - f(a) \geq \frac{p}{4}|h|^2 > 0;$$

also liegt ein lokales Minimum in a vor.

Im Falle der negativen Definitheit von $f''(a, \cdot, \cdot)$ geht man analog vor.

Sei nun $f''(a, \cdot, \cdot)$ indefinit, d.h.:

$$\exists h_0, k_0 \in \mathbb{R}^n, |h_0| = |k_0| = 1 : f''(a, h_0, h_0) =: p_0 > 0, \quad f''(a, k_0, k_0) =: n_0 < 0.$$

Mit derselben Argumentation wie oben zeigt man:

$$\exists \varepsilon > 0 \quad \forall t \in \mathbb{R} \setminus \{0\}, \quad |t| < \varepsilon : f(a+th_0) > f(a) \wedge f(a+tk_0) < f(a).$$

In a kann somit kein lokales Extremum der Funktion f vorliegen. □

2.33 Beispiel. Es seien α, β, γ die wahren Werte der Winkel eines Dreiecks. Die gemessenen Werte seien $\alpha+x, \beta+y$ und $\gamma+z = \gamma+(\delta-x-y)$, wobei $\delta := \pi - (\alpha + \beta + \gamma)$ der gegebene Gesamtfehler sei. Die Fehler x und y sollen so bestimmt werden, dass die Summe über die Fehlerquadrate bei gegebenem Gesamtfehler minimal wird. Gesucht ist also ein Minimum der Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2 + (\delta - x - y)^2.$$

Die Ableitung

$$\nabla f(x, y) = \begin{pmatrix} 4x + 2y - 2\delta \\ 2x + 4y - 2\delta \end{pmatrix}$$

verschwindet für $a = \frac{\delta}{3} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Die Hesse-Matrix

$$\text{Hess}_f(x, y) = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$$

hat die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = 6$ und ist somit positiv definit. Folglich liegt in a ein Minimum vor, d.h.: Die Summe über die Fehlerquadrate wird minimiert bei Drittelung des Gesamtfehlers.

2.34 Beispiel (Methode der kleinsten Fehlerquadrate). Im \mathbb{R}^2 seien $n \geq 2$ Punkte $(x_i, y_i), i = 1, \dots, n$, gegeben, wobei mindestens zwei der x_i sich unterscheiden mögen. Gesucht ist eine Ausgleichsgerade g ,

$$g: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto g(x) := \alpha x + \beta,$$

mit der Eigenschaft, dass

$$\sum_{i=1}^n (g(x_i) - y_i)^2 =: \sum_{i=1}^n r_i^2 =: |r|^2 \text{ mit } r := (r_1, \dots, r_n)^\top$$

minimal wird („kleinste Fehlerquadrate“). Hierbei wurde $r := Az - y$ definiert, wobei $z := (\alpha, \beta)^\top$, $y := (y_1, \dots, y_n)^\top$ und

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad z \mapsto f(z) := |r|^2 = \sum_{i=1}^n (y_i - \alpha x_i - \beta)^2.$$

Da $|r|^2 = (Az - y)^\top (Az - y) = z^\top A^\top Az - 2z^\top A^\top y + y^\top y$ für $A := \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$ ist,

ergibt sich für die Ableitungen:

$$\nabla f(z) = 2A^\top Az - 2A^\top y \text{ und } f''(z) = 2A^\top A.$$

Für eine kritische Stelle a muss gelten:

$$A^\top Aa = A^\top y \quad (\text{Normalgleichung}).$$

Man sieht sofort aus $\langle A^\top Ah, h \rangle = |Ah|^2$, dass die Matrix $A^\top A$ positiv definit ist, falls $\ker A = \{0\}$, d.h. falls $\text{rg } A$ maximal ist. Dies ist hier der Fall, da sich zwei der x_i voneinander unterscheiden.

Als letztes Beispiel besprechen wir die konvexen Funktionen.

2.35 Definition. Eine Menge $M \subset \mathbb{R}^n$ heißt konvex, wenn mit x_1 und x_2 auch die Verbindungsstrecke $x_1 + t(x_2 - x_1)$, $t \in [0, 1]$, zu M gehört.

Beliebige Durchschnitte konvexer Mengen sind offenbar konvex.

2.36 Definition. Sei $U \subset \mathbb{R}^n$ offen und konvex.

(i) Dann heißt $f : U \rightarrow \mathbb{R}$ konvex: \Leftrightarrow
 $\forall x_1, x_2 \in U, x_1 \neq x_2, \forall t \in (0, 1) :$

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2).$$

Gilt sogar „<“, so heißt f „streng konvex“.

(ii) f heißt „(streng) konkav“ : $\Leftrightarrow -f$ ist (streng) konvex.

2.37 Bemerkung. (i) Man kann zeigen, dass konvexe Funktionen stetig sind und dass die einseitigen Ableitungen existieren.

(ii) Seien $U \subset \mathbb{R}^n$ offen und konvex und $f : U \rightarrow \mathbb{R}$ zweimal differenzierbar. Man kann zeigen, dass dann gilt:

Wenn f'' in U positiv definit ist, so ist f in U streng konvex.

Wenn f'' in U positiv semidefinit ist, dann ist f in U konvex.

d) Extrema unter Nebenbedingungen

Ebenfalls häufig in der Praxis auftretende Probleme sind die „Extremwertaufgaben unter Nebenbedingungen“:

Es soll das Extremum $x_0 \in \mathbb{R}^n$ der Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ unter den m ($m < n$) „Nebenbedingungen“

$$g_1(x_0) = \dots = g_m(x_0) = 0$$

gefunden werden. Hierbei sind die $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ Funktionen.

Eine erste Idee zum Lösen dieser Aufgabe ist, aus den Nebenbedingungen durch Auflösen Gleichungen für die letzten m x_{0j} ($x_0 =: (x_{0i})_{i=1}^n$) zu erhalten, diese Ausdrücke in die Funktion f einzusetzen und dann die Extremstellen zu berechnen.

2.38 Beispiel. Seien $f : \mathbb{R}^2 \rightarrow \mathbb{R}, x \mapsto x_1^2 + x_2^2$ und $g : \mathbb{R}^2 \rightarrow \mathbb{R}, x \mapsto x_1 + x_2 - 1$. Die Funktion f hat nur im Punkt $(0, 0)$ eine Extremalstelle (ein Minimum), aber die Nebenbedingung ist nicht erfüllt, da $g(0, 0) = -1 \neq 0$ ist. Nun gilt:

$$g(x) = 0 \Leftrightarrow x_1 = 1 - x_2.$$

Definiere $h : \mathbb{R} \rightarrow \mathbb{R}, x_2 \mapsto h(x_2) := f(1 - x_2, x_2)$. Die Funktion h hat ein Minimum in $x_2 = \frac{1}{2}$, d.h. $(\frac{1}{2}, \frac{1}{2})$ ist die Lösung unserer Extremwertaufgabe.

2.39 Bemerkung. Im allgemeinen ist eine Elimination - wie im obigen Beispiel vorgenommen - jedoch zu aufwändig. Man kann das Minimum aber auch mit einem alternativen Ansatz ohne explizite Auflösung berechnen.

Gegeben seien $f: \mathbb{R}^2 \supset U \rightarrow \mathbb{R}$ und $g: U \rightarrow \mathbb{R}$. Gesucht ist das Minimum von $f(x)$ unter der Nebenbedingung $g(x) = 0$. Dabei sei $\nabla g(x) \neq 0$ für alle $x \in U$, o.E. sei $\partial_1 g(x) \neq 0$.

Wir nehmen an, dass man $g(x) = 0$ nach x_1 auflösen kann, d.h. es existiert eine Funktion $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ mit $g(x) = 0 \Leftrightarrow x_1 = \varphi(x_2)$. Für spätere Zwecke setzen wir hier $z := x_2$, d.h. $x = (\varphi(z), z)$. Dann gilt

$$g(\varphi(z), z) = 0 \quad (z \in U). \quad (2-1)$$

Falls f an der Stelle $x_0 = (\varphi(z_0), z_0)$ ein Minimum besitzt, besitzt die Funktion $h: \mathbb{R} \rightarrow \mathbb{R}$, $z \mapsto f(\varphi(z), z)$ an der Stelle z_0 ein Minimum, d.h. es gilt $h'(z_0) = 0$. Mit der Kettenregel erhält man

$$\partial_1 f(x_0) \varphi'(z_0) + \partial_2 f(x_0) = 0. \quad (2-2)$$

Andererseits folgt durch Ableiten von (2-1) nach z :

$$\partial_1 g(x_0) \varphi'(z_0) + \partial_2 g(x_0) = 0. \quad (2-3)$$

Wegen $\partial_1 g(x_0) \neq 0$ kann man (2-3) nach $\varphi'(z_0)$ auflösen:

$$\varphi'(z_0) = -(\partial_1 g(x_0))^{-1} \partial_2 g(x_0).$$

Eingesetzt in (2-2) erhält man

$$\partial_2 f(x_0) - \partial_1 f(x_0) (\partial_1 g(x_0))^{-1} \partial_2 g(x_0) = 0.$$

Wir definieren

$$\lambda_0 := -\partial_1 f(x_0) (\partial_1 g(x_0))^{-1}$$

und erhalten

$$\partial_2 f(x_0) + \lambda_0 \partial_2 g(x_0) = 0. \quad (2-4)$$

Andererseits gilt nach Definition von λ_0 auch

$$\partial_1 f(x_0) + \lambda_0 \partial_1 g(x_0) = 0. \quad (2-5)$$

Aus (2-4) und (2-5) erhält man

$$\nabla f(x_0) + \lambda_0 \nabla g(x_0) = 0. \quad (2-6)$$

Andererseits gilt auch $g(x_0) = 0$. Definiert man also die Funktion

$$F: U \times \mathbb{R} \rightarrow \mathbb{R}, (x, \lambda) \mapsto f(x) + \lambda g(x),$$

so erhält man:

Falls x_0 eine lokale Extremalstelle der Funktion f unter der Nebenbedingung $g(x) = 0$ ist, so existiert ein $\lambda_0 \in \mathbb{R}$ so, dass

$$\nabla F(x_0, \lambda_0) = 0 \quad (2-7)$$

gilt. Dieses λ_0 heißt Lagrange-Multiplikator. Ausgeschrieben bedeutet (2-7):

$$\begin{aligned} \nabla f(x_0) + \lambda_0 \nabla g(x_0) &= 0, \\ g(x_0) &= 0. \end{aligned}$$

2.40 Beispiel. Wir wollen mit obigem Ansatz die Extremwertaufgabe unter Nebenbedingung lösen, die im Beispiel 2.38 bereits behandelt wurde. Es ist

$$F : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}, \quad (x, \lambda) \mapsto F(x, \lambda) := x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 1),$$

also

$$\nabla F(x, \lambda) = \begin{pmatrix} 2x_1 + \lambda \\ 2x_2 + \lambda \\ x_1 + x_2 - 1 \end{pmatrix}.$$

$\nabla F(x_0, \lambda_0) = 0$ liefert: $x_0 = (\frac{1}{2}, \frac{1}{2})^\top$, $\lambda_0 = -1$.

Im allgemeinen Fall betrachtet man $U \subset \mathbb{R}^n$, $f : U \rightarrow \mathbb{R}$ und $g : U \rightarrow \mathbb{R}^m$ mit $m < n$, d.h. man hat m Nebenbedingungen $g_1(x) = \dots = g_m(x) = 0$.

2.41 Satz. Seien $U \subset \mathbb{R}^n$ offen, und seien $f : U \rightarrow \mathbb{R}$ und $g : U \rightarrow \mathbb{R}^m$ mit $m < n$ stetig differenzierbare Funktionen. Der Rang der Matrix $g'(x) \in \mathbb{R}^{m \times n}$ sei für alle $x \in U$ maximal, also gleich m . Sei

$$M := \{x \in U : g_1(x) = \dots = g_m(x) = 0\}.$$

Falls x_0 eine lokale Extremalstelle von $f(x)$ unter der Nebenbedingung $g(x) = 0$ (d.h. von $f|_M$) ist, so existiert ein Lagrange-Multiplikator $\lambda_0 = (\lambda_{0,1}, \dots, \lambda_{0,m}) \in \mathbb{R}^m$ so, dass für die Funktion

$$F : U \times \mathbb{R}^m \rightarrow \mathbb{R}, \quad (x, \lambda) \mapsto F(x, \lambda) := f(x) + \sum_{j=1}^m \lambda_j g_j(x)$$

gilt

$$\nabla F(x_0, \lambda_0) = 0.$$

Insbesondere gilt

$$\nabla f(x_0) + \sum_{j=1}^m \lambda_{0,j} \nabla g_j(x_0) = 0.$$

Wir verwenden im Beweis eine Aussage über die Auflösbarkeit der Gleichung $g(x) = 0$, welche später bewiesen werden wird (Satz über implizite Funktionen).

Beweis. Der Beweis verläuft analog zu den Ausführungen in Bemerkung 2.39. Nach dem Satz über implizite Funktionen existiert (nach eventueller Umnummerierung der Koordinaten) eine Funktion $\varphi: \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m$ so, dass gilt: $g(x) = 0 \Leftrightarrow x = (\varphi(z), z)$. Jetzt kann der Beweis aus Bemerkung 2.39 wörtlich übernommen werden, wobei alle auftretenden Größen Matrizen sind. Dabei ist $\partial_1 f(x_0)$ bzw. $\partial_2 f(x_0)$ zu interpretieren als

$$\begin{aligned}\partial_1 f(x_0) &= \left(\frac{\partial}{\partial x_1} f(x_0), \dots, \frac{\partial}{\partial x_m} f(x_0) \right) \in \mathbb{R}^{1 \times m}, \\ \partial_2 f(x_0) &= \left(\frac{\partial}{\partial x_{m+1}} f(x_0), \dots, \frac{\partial}{\partial x_n} f(x_0) \right) \in \mathbb{R}^{1 \times (n-m)}.\end{aligned}$$

Analog für $\partial_1 g(x_0) \in \mathbb{R}^{m \times m}$, $\partial_2 g(x_0) \in \mathbb{R}^{m \times (n-m)}$. Der Lagrange-Multiplikator λ_0 ist wie oben gegeben durch

$$\lambda_0 := -\partial_1 f(x_0) (\partial_1 g(x_0))^{-1} \in \mathbb{R}^{1 \times m}.$$

□

2.42 Bemerkung. Eine hinreichende Bedingung dafür, dass $x_0 \in \mathbb{R}^n$ eine Extremalstelle von $f|_M$ ist, ist die Definitheit der Hessematrix $\text{Hess}_F(x_0, \lambda_0)$, hierbei seien F und λ_0 wie in der Formulierung des Satzes definiert. Denn dann gilt z.B.

$$F(x_0, \lambda_0) \leq F(x, \lambda_0) \quad (x \in U),$$

das bedeutet aber insbesondere:

$$f(x_0) \leq f(x) \quad (x \in M).$$

An dieser Argumentation sieht man, dass bereits die Definitheit der „linken oberen Ecke“, d.h. der Matrix $(\partial_i \partial_j F(x_0, \lambda_0))_{i,j=1}^n$ genügt

Abschließend wollen wir mit Hilfe des Satzes 2.41 quadratische Formen untersuchen.

2.43 Beispiel. Sei $A = (a_{ij})_{i,j=1}^n$ eine symmetrische Matrix. Wir betrachten die quadratische Form

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto f(x) := \langle x, Ax \rangle = \sum_{i,j=1}^n x_i a_{ij} x_j.$$

Wir suchen die Extremalstellen von f auf der Sphäre $S^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\} \subset \mathbb{R}^n$. Unsere Nebenbedingung wird somit bestimmt durch

$$g: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto g(x) := 1 - |x|^2.$$

Da $S^{n-1} \subset \mathbb{R}^n$ abgeschlossen und beschränkt, also kompakt ist, nimmt die stetige Funktion f das Maximum auf S^{n-1} an, d.h. es gibt ein $x_0 \in S^{n-1}$ mit $f(x_0) = \max_{x \in S^{n-1}} f(x)$. Nach Satz 2.41 existiert ein $\lambda_0 \in \mathbb{R}$ mit $\nabla f(x_0) + \lambda_0 \nabla g(x_0) = 0$. Wegen $\nabla f(x) = 2Ax$ und $\nabla g(x) = -2x$ folgt

$$Ax_0 - \lambda_0 x_0 = 0,$$

d.h. x_0 ist ein Eigenwert von A . Der Funktionswert ist dabei $f(x_0) = \langle Ax_0, x_0 \rangle = \lambda_0 \langle x_0, x_0 \rangle = \lambda_0$.

Aus der Theorie symmetrischer Matrizen ist bekannt, dass A genau n reelle Eigenwerte $\lambda_1, \dots, \lambda_n$ besitzt (inklusive Vielfachheit). Damit ist das Maximum von $f(x)$ auf S^{n-1} gleich dem größten Eigenwert von A (und das Minimum gleich dem kleinsten Eigenwert).

2.44 Beispiel (Portfolio-Optimierung). Ein Anleger habe die Wahl, sein Kapital in einer festverzinslichen Anleihe mit 2% Zinssatz oder in Aktien zu investieren. Für die Aktie seien nach einem Jahr zwei Zustände möglich: Boom mit 10% Ertrag und Rezession mit 5% Verlust. Beide Zustände seien gleichwahrscheinlich. Das heutige Vermögen sei 10 (Rechnungseinheiten), der Besitz eines Kapitals der Größe z bringe dem Anleger einen Nutzen von $\ln z$ (Nutzenfunktion, Modell!). Wieviel Kapital soll der Anleger in Aktien investieren? Ansatz:

$$\begin{aligned} x_1 &:= \text{in festverzinsliche Anleihe investiert,} \\ x_2 &:= \text{in Aktien investiert,} \\ x &:= (x_1, x_2). \end{aligned}$$

Nutzen:

$$\begin{aligned} \text{Boom:} & \quad \ln((1+r)x_1 + ux_2), \\ \text{Rezession:} & \quad \ln((1+r)x_1 + dx_2) \end{aligned}$$

mit

$$r := 0,02, \quad u := 1,1, \quad d := 0,95.$$

Durchschnittlich zu erwartender Nutzen:

$$f(x) := \frac{1}{2} \ln((1+r)x_1 + ux_2) + \frac{1}{2} \ln((1+r)x_1 + dx_2).$$

Nebenbedingung, mit $K := 10$:

$$g(x) := x_1 + x_2 - K = 0.$$

Mit

$$\nabla f(x) = \frac{1}{2} \left(\frac{1+r}{(1+r)x_1 + ux_2} + \frac{1+r}{(1+r)x_1 + dx_2}, \frac{u}{(1+r)x_1 + ux_2} + \frac{d}{(1+r)x_1 + dx_2} \right),$$

$$\nabla g(x) = (1, 1)$$

liefert der Lagrange-Ansatz

$$\frac{1+r}{(1+r)x_1 + ux_2} + \frac{1+r}{(1+r)x_1 + dx_2} = \frac{u}{(1+r)x_1 + ux_2} + \frac{d}{(1+r)x_1 + dx_2}$$

und daraus

$$\alpha_1 x_1 + \alpha_2 x_2 = 0$$

mit

$$\alpha_1 := \frac{1+r}{1+r-u} + \frac{1+r}{1+r-d}, \quad \alpha_2 := \frac{u}{1+r-u} + \frac{d}{1+r-d}.$$

Die Nebenbedingung liefert

$$x_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1} K, \quad x_2 = \frac{\alpha_1}{\alpha_1 - \alpha_2} K,$$

und damit

$$x_1 = -\frac{1}{2} \left(\frac{u}{1+r-u} + \frac{d}{1+r-d} \right).$$

Mit eingesetzten Zahlen, auf zwei Stellen gerundet: $x_1 = 0,89$, $x_2 = 9,11$. (Für eine Maximaleigenschaft wäre natürlich noch die zweite Ableitung zu untersuchen.)

3. Kurven und Flächen

3.1 Worum geht's? Eine Kurve ist eine Äquivalenzklasse von (glatten) Wegen im Raum \mathbb{R}^n . Wesentliche Begriffe sind hier die Weglänge, Tangentialvektor und Krümmung. Wir werden vor allem Kurven im zwei- und dreidimensionalen Raum genauer betrachten. Eine Fläche, etwa eine zweidimensionale Fläche in \mathbb{R}^3 , kann analog zu Kurven durch eine Parametrisierung beschrieben werden. Beispiele von Flächen sind Rotationsflächen oder die stereographische Projektion der Ebene auf die Kugeloberfläche.

a) Weglängen

Wir haben schon in einigen Beispielen Wege, d.h. stetige Abbildungen $\gamma : [a, b] \rightarrow \mathbb{R}^n$ kennengelernt. Wege können einen recht willkürlichen Verlauf nehmen, z.B. können sie im \mathbb{R}^2 eine Dreiecksfläche ganz ausfüllen (Peanokurven u.a.). Wir betrachten nun „glattere“ als stetige Wege.

3.2 Definition. (i) Ein „Weg“ ist eine stetige Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^n$. Ist γ an der Stelle $s_0 \in [a, b]$ (stetig) differenzierbar, so heißt der Weg „an der Stelle s_0 (stetig) differenzierbar“.

(ii) Ist γ überall stetig differenzierbar und ist $\gamma'(s) \neq 0$ für alle $s \in (a, b)$, so heißt γ „glatt“.

(iii) Der Weg γ heißt „stückweise glatt“, falls er aus endlich vielen glatten Wegen zusammengesetzt ist.

(iv) Für glatte Wege (oder die glatten Teilstücke eines stückweise glatten Weges) definieren wir den „Tangenteneinheitsvektor“ durch

$$t : (a, b) \rightarrow \mathbb{R}^n, \quad s \mapsto t(s) = \gamma'(s)/|\gamma'(s)|.$$

3.3 Beispiel. Sei $\gamma : [-1, 1] \rightarrow \mathbb{R}^2$, $t \mapsto \gamma(t)$ definiert durch

$$\gamma(t) := \begin{cases} (-t^2, 0) & \text{für } -1 \leq t \leq 0 \\ (0, t^2) & \text{für } 0 \leq t \leq 1. \end{cases}$$

Dann ist γ stetig differenzierbar mit

$$\gamma'(t) = \begin{cases} (-2t, 0) & \text{für } -1 \leq t \leq 0 \\ (0, 2t) & \text{für } 0 \leq t \leq 1, \end{cases}$$

aber γ ist nicht glatt, da $\gamma'(0) = 0$ ist; γ ist jedoch stückweise glatt, da $\gamma|_{[-1,0]}$ und $\gamma|_{[0,1]}$ glatt sind.

Für stückweise glatte Wege lässt sich die euklidische Länge erklären:

3.4 Definition. Für einen stückweise glatten Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ sei „die Länge $L(\gamma)$ des Weges γ “ erklärt durch

$$L(\gamma) := \int_a^b |\gamma'(s)| \, ds.$$

3.5 Beispiele. (i) Sei $\gamma : [0, 1] \rightarrow \mathbb{R}^n$, $s \mapsto \gamma(s) := a + s(b - a)$ die Strecke von $a \in \mathbb{R}^n$ nach $b \in \mathbb{R}^n$. Dann ist

$$L(\gamma) = \int_0^1 |\gamma'(s)| \, ds = \int_0^1 |b - a| \, ds = |b - a|.$$

Dies entspricht dem euklidischen Abstand der Punkte a und b .

(ii) Sei $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, $s \mapsto \gamma(s) := \begin{pmatrix} \cos s \\ \sin s \end{pmatrix}$ die Kreislinie. Dann ist

$$L(\gamma) = \int_0^{2\pi} |\gamma'(s)| \, ds = \int_0^{2\pi} \left| \begin{pmatrix} -\sin s \\ \cos s \end{pmatrix} \right| \, ds = \int_0^{2\pi} 1 \, ds = 2\pi,$$

also die „übliche“ Länge.

Die Definition der Länge eines Weges motiviert sich aus den folgenden Eigenschaften:

- (1) Für Strecken erhält man den euklidischen Abstand.
- (2) L ist additiv, d.h. für $a < b < c$ und stückweise glatte Wege $\gamma_1 : [a, b] \rightarrow \mathbb{R}^n$, $\gamma_2 : [b, c] \rightarrow \mathbb{R}^n$ mit $\gamma_1(b) = \gamma_2(b)$ und $\gamma : [a, c] \rightarrow \mathbb{R}^n$ mit $\gamma|_{[a,b]} = \gamma_1$, $\gamma|_{[b,c]} = \gamma_2$ gilt: $L(\gamma) = L(\gamma_1) + L(\gamma_2)$.
- (3) Die Länge $L(\gamma)$ eines Weges $\gamma : [a, b] \rightarrow \mathbb{R}^n$ hängt in dem folgenden Sinne stetig von γ ab:
Sei $(\gamma_n)_n$ eine Folge stückweise glatter Wege, $\gamma_n, \gamma : [a, b] \rightarrow \mathbb{R}^n$, mit

$$\lim_{n \rightarrow \infty} \|\gamma_n - \gamma\|_{C^1([a,b])} := \lim_{n \rightarrow \infty} \left(\sup_{s \in [a,b]} (|\gamma_n(s) - \gamma(s)|) + \sup_{s \in [a,b]} (|\gamma_n'(s) - \gamma'(s)|) \right) = 0,$$

dann gilt: $\lim_{n \rightarrow \infty} L(\gamma_n) = L(\gamma)$.

Man kann zeigen (durch Approximation von γ' durch Treppenfunktionen), dass die obengenannten Eigenschaften die Abbildung L , die einem stückweise glatten Weg eine Länge zuordnet, eindeutig bestimmen. Hierbei ist von großer Bedeutung, dass die Approximation in C^1 , d.h. im Raum der stetig differenzierbaren Funktionen, und nicht etwa nur in C^0 , vorgenommen wird. Siehe hierzu das folgende

3.6 Beispiel. Der Weg $\gamma : [0, 1] \rightarrow \mathbb{R}^2$, $s \mapsto \gamma(s) := \begin{pmatrix} s \\ 0 \end{pmatrix}$ wird in C^0 , d.h. bezüglich der Supremumsnorm, approximiert durch die „Sägezähne“ $\gamma_n : [0, 1] \rightarrow \mathbb{R}^2$, $s \mapsto \gamma_n(s) := \begin{pmatrix} s \\ a_n(s) \end{pmatrix}$ mit

$$a_n(s) := \begin{cases} s & \text{für } 0 \leq s \leq \frac{1}{2^n} \\ 2 \cdot \frac{1}{2^n} - s & \text{für } \frac{1}{2^n} \leq s \leq 2 \cdot \frac{1}{2^n} \\ \vdots & \vdots \\ (2^n - 1) \cdot \frac{1}{2^n} - s & \text{für } \frac{2^n - 1}{2^n} \leq s \leq 1. \end{cases}$$

Dann ist $L(\gamma) = 1$ und $L(\gamma_n) = \sqrt{2}$.

Die Länge L hat ferner die folgenden Eigenschaften:

- (4) Seien $a \in \mathbb{R}^n$ und $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung mit $\langle Ax, Ax \rangle = \langle x, x \rangle$ für alle $x \in \mathbb{R}^n$ (d.h. A ist eine Isometrie). Dann gilt für alle stückweise glatten Wege γ :

$$L(A\gamma + a) = L(\gamma).$$

- (5) Die Länge eines stückweise glatten Weges γ ist unabhängig von der Parametrisierung des Weges in folgendem Sinne:

Sei $\varphi : [c, d] \rightarrow [a, b]$ eine stetig differenzierbare Funktion mit $\varphi'(\tilde{s}) > 0$ ($\forall \tilde{s} \in [c, d]$). Dann gilt für $\gamma : [a, b] \rightarrow \mathbb{R}^n$, $s \mapsto \gamma(s)$ und $\tilde{\gamma} : [c, d] \rightarrow \mathbb{R}^n$, $\tilde{s} \mapsto \tilde{\gamma}(\tilde{s}) := \gamma(\varphi(\tilde{s}))$ die Gleichheit $L(\tilde{\gamma}) = L(\gamma)$, denn

$$\begin{aligned} L(\tilde{\gamma}) &= \int_c^d |\tilde{\gamma}'(\tilde{s})| \, d\tilde{s} = \int_c^d |\gamma'(\varphi(\tilde{s}))| \cdot |\varphi'(\tilde{s})| \, d\tilde{s} \\ &= \int_c^d |\gamma'(\varphi(\tilde{s}))| \cdot \varphi'(\tilde{s}) \, d\tilde{s} = \int_{\varphi(c)}^{\varphi(d)} |\gamma'(s)| \, ds = L(\gamma). \end{aligned}$$

Die Möglichkeit, verschiedene Wege durch Umparametrisierung ineinander zu überführen, motiviert die folgende Definition von Äquivalenzklassen:

3.7 Definition. Zwei stückweise glatte Wege $\gamma : [a, b] \rightarrow \mathbb{R}^n$ und $\tilde{\gamma} : [c, d] \rightarrow \mathbb{R}^n$ heißen „äquivalent“, wenn eine Parametertransformation $\varphi : [c, d] \rightarrow [a, b]$ existiert, d.h. eine stetig differenzierbare Funktion mit $\varphi'(\tilde{s}) > 0$ für alle $\tilde{s} \in [c, d]$ mit $\tilde{\gamma}(\tilde{s}) = \gamma(\varphi(\tilde{s})) \, \forall \tilde{s} \in [c, d]$.

Jede solche Äquivalenzklasse heißt „orientierte stückweise glatte Kurve“ (die Orientierung wird festgelegt durch das Vorzeichen der Ableitung der Parametertransformation).

Wir sind insbesondere an einer ausgezeichneten Parameterdarstellung interessiert, nämlich an dem „nach Bogenlänge parametrisierten Repräsentanten γ “ einer Kurve, für den gilt: $|\gamma'(s)| = 1$ ($\tilde{s} \in [c, d]$). Sei dazu

$$\psi : [a, b] \rightarrow [0, L], \quad \psi(s) := \int_a^s |\gamma'(s)| \, ds$$

(„Bogenlänge von γ “). Dann gilt $\psi(a) = 0$ und $\psi(b) = L(\gamma)$. Für die Umkehrfunktion $\varphi := \psi^{-1}$, $\tilde{s} \mapsto \varphi(\tilde{s})$, erhalten wir $\varphi'(\tilde{s}) = \frac{1}{|\gamma'(\varphi(\tilde{s}))|}$. Mit $\psi'(s) = |\gamma'(s)|$ ergibt sich für die Umparametrisierung $\tilde{\gamma}(\tilde{s}) := \gamma(\varphi(\tilde{s}))$

$$|\tilde{\gamma}'(\tilde{s})| = |\gamma'(\varphi(\tilde{s}))\varphi'(\tilde{s})| = 1.$$

3.8 Beispiel. Für $a, b \in \mathbb{R}^+$ definiere

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad s \mapsto \gamma(s) := \begin{pmatrix} a \cos s \\ b \sin s \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix};$$

(γ beschreibt eine Ellipse); dann ist

$$|\gamma'(s)| = \left| \begin{pmatrix} -a \sin s \\ b \cos s \end{pmatrix} \right| = \sqrt{a^2 \sin^2 s + b^2 \cos^2 s},$$

d.h. die Bogenlänge ist

$$\sigma = \psi(s) = \int_0^s \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} \, dt.$$

Dies ist ein elliptisches Integral und i.a. nicht elementar berechenbar. Für den Spezialfall des Kreises, d.h. $a = b =: r$, ergibt sich: $\sigma = r \cdot s$ bzw. $\varphi(\sigma) = \frac{\sigma}{r}$.

Für

$$\tilde{\gamma} : [0, 2\pi r] \rightarrow \mathbb{R}^2, \quad \sigma \mapsto \tilde{\gamma}(\sigma) := \gamma(\sigma/r) = r \cdot \begin{pmatrix} \cos \sigma/r \\ \sin \sigma/r \end{pmatrix}$$

gilt: $|\tilde{\gamma}'(\sigma)| = 1$; das ist die gesuchte ausgezeichnete Darstellung.

Sei im folgenden $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ein differenzierbarer Weg mit $|\gamma'(s)| = 1$ für $s \in [a, b]$. Dann ist Tangenteneinheitsvektor $t(s)$ gegeben durch $t(s) = \gamma'(s) \quad \forall s \in [a, b]$. ($\gamma'(s)$ ist der „Geschwindigkeitsvektor“.) Ist γ zweimal differenzierbar, so heißt $\gamma''(s)$ „Krümmungsvektor“ oder „Beschleunigungsvektor“ zu γ im Punkt $\gamma(s)$. Tangenten- und Krümmungsvektor stehen senkrecht aufeinander, wie die Differentiation des Ausdrucks $\langle \gamma'(s), \gamma'(s) \rangle = 1$ beweist. Ist $\gamma''(s) \neq 0$, so sind $\gamma'(s)$ und $\gamma''(s)$ linear unabhängig. Die von $\gamma'(s)$ und $\gamma''(s)$ aufgespannte Ebene heißt „Schmiegebene“ in $\gamma(s)$. Der Ausdruck $\kappa(s) := |\gamma''(s)|$ heißt „Krümmung“ von γ in $\gamma(s)$.

3.9 Beispiele. (i) Für $a, b \in \mathbb{R}^n$ mit $|b| = 1$ und $\gamma : (-\infty, \infty) \rightarrow \mathbb{R}^n$, $s \mapsto \gamma(s) := a + sb$ gilt: $\forall s : \kappa(s) := |\gamma''(s)| = 0$.

(ii) Seien $a, b \in \mathbb{R}^n$ orthonormiert und $m \in \mathbb{R}^n$. Definiere $\gamma : [0, 2\pi r] \rightarrow \mathbb{R}^n$, $s \mapsto \gamma(s) := m + r(a \cos \frac{s}{r} + b \sin \frac{s}{r})$ (γ beschreibt einen Kreis mit Mittelpunkt m und Radius r). Dann sind

$$|\gamma'(s)| = \sqrt{|a|^2 \cdot \sin^2(s/r) + |b|^2 \cos^2(s/r)} = 1 \text{ und}$$

$$\begin{aligned}\gamma''(s) &= -\frac{1}{r}\left(a\cos\frac{s}{r} + b\sin\frac{s}{r}\right), \text{ also ist} \\ \kappa(s) &= \frac{1}{r}.\end{aligned}$$

(iii) Beliebigen zweimal differenzierbaren Kurven γ kann man einen „Krümmungskreis“ zuordnen, der die Kurven von zweiter Ordnung berührt.

b) Kurven in der Ebene und im Raum

Wir behandeln zunächst ebene Kurven: Sei $\gamma : [a, b] \rightarrow \mathbb{R}^2$, $s \mapsto \gamma(s) := \begin{pmatrix} x(s) \\ y(s) \end{pmatrix}$ eine ebene Kurve mit $|\gamma'(s)| = 1$ für $s \in [a, b]$. Dann ist

$$t(s) = \begin{pmatrix} x'(s) \\ y'(s) \end{pmatrix}$$

die Tangente, und es sei

$$n(s) := \begin{pmatrix} -y'(s) \\ x'(s) \end{pmatrix}$$

die „Normale“.

Tangente und Normale stehen senkrecht aufeinander für alle $s \in [a, b]$ und sind ungleich 0. Es existieren somit Abbildungen $\alpha, \beta, \delta, \varepsilon : [a, b] \rightarrow \mathbb{R}$ mit

$$\begin{aligned}t'(s) &= \alpha(s)t(s) + \beta(s)n(s) \\ \text{und } n'(s) &= \delta(s)t(s) + \varepsilon(s)n(s).\end{aligned}$$

Diese Funktionen lassen sich leicht bestimmen:

$$\begin{aligned}\alpha(s) &= \langle t(s), t'(s) \rangle = 0, \text{ da } |t(s)| = 1, \\ \beta(s) &= \langle t'(s), n(s) \rangle = -\langle n'(s), t(s) \rangle = -\delta(s), \text{ da } \langle n(s), t(s) \rangle = 0 \text{ und} \\ \varepsilon(s) &= \langle n'(s), n(s) \rangle = 0, \text{ da } |n(s)| = 1.\end{aligned}$$

Also lauten unsere Gleichungen:

$$\begin{pmatrix} t' \\ n' \end{pmatrix} = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} \begin{pmatrix} t \\ n \end{pmatrix}$$

und $|\beta(s)| = \kappa(s)$ ($\kappa(s) := |t'(s)|$ bezeichnet wieder die Krümmung).

Diese Gleichungen heißen „Frenetsche Gleichungen“⁸ im \mathbb{R}^2 .

⁸Jean Frédéric Frenet, 7.2.1816 – 12.6.1900

Schreiben wir

$$t(s) = \begin{pmatrix} \cos \alpha(s) \\ \sin \alpha(s) \end{pmatrix},$$

wobei $\alpha(s)$ den „Kontingenzwinkel“ bezeichnet (also den Winkel, den die Tangente mit dem Vektor $(1, 0)$ einschließt), so ist

$$t'(s) = \alpha'(s) \begin{pmatrix} -\sin \alpha(s) \\ \cos \alpha(s) \end{pmatrix} = \alpha'(s) \cdot n(s).$$

Folglich ist $\beta(s) = \alpha'(s)$.

Wir betrachten nun Raumkurven

$$\gamma : [a, b] \rightarrow \mathbb{R}^3, \quad s \mapsto \gamma(s) := \begin{pmatrix} x(s) \\ y(s) \\ z(s) \end{pmatrix}$$

mit $|\gamma'(s)| = 1$. Wiederum bezeichnen

$$\begin{aligned} t(s) &:= \gamma'(s) \text{ die Tangente und} \\ n(s) &:= \frac{t'(s)}{|t'(s)|} \text{ die Normale.} \end{aligned}$$

(Es gilt $\langle t(s), n(s) \rangle = 0$, da $|t(s)| = 1$.) Um eine Basis des \mathbb{R}^3 zu erhalten, benötigen wir eine dritte vektorwertige Funktion b , die die folgenden Eigenschaften haben soll: $\forall s \in [a, b]$:

1. $\langle t(s), b(s) \rangle = 0 \wedge \langle n(s), b(s) \rangle = 0$
2. $\det(t(s), n(s), b(s)) = 1$
3. $|b(s)| = 1$.

b heißt „Binormale“ und ist definiert durch

$$b(s) := t(s) \times n(s).$$

Hierbei ist für zwei Vektoren $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3) \in \mathbb{R}^3$ das „Kreuzprodukt $a \times b$ “ erklärt durch:

$$a \times b := \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Wie zuvor kann man nun die Vektoren $t'(s)$, $n'(s)$ und $b'(s)$ bzgl. der Basis $\{t(s), n(s), b(s)\}$ darstellen, und wir erhalten durch ähnliche Überlegungen wie für ebene Kurven die Frenetschen Gleichungen im \mathbb{R}^3 :

$$\begin{pmatrix} t'(s) \\ n'(s) \\ b'(s) \end{pmatrix} = \begin{pmatrix} 0 & \kappa(s) & 0 \\ -\kappa(s) & 0 & \tau(s) \\ 0 & -\tau(s) & 0 \end{pmatrix} \begin{pmatrix} t(s) \\ n(s) \\ b(s) \end{pmatrix}.$$

Hierbei ist

$$\tau(s) := \langle n'(s), b(s) \rangle = \langle -b'(s), n(s) \rangle$$

die „Windung“ (oder „Torsion“) von γ . Es gilt (zur kürzeren Schreibweise verzichten wir auf das Ausschreiben der Variablen s):

$$\begin{aligned} \tau &= \langle b, n' \rangle = \langle t \times n, n' \rangle = \langle t, n \times n' \rangle \\ &= \langle t, n \times \left(\frac{t''}{\kappa} + \left(\frac{1}{\kappa} \right)' t' \right) \rangle = \frac{1}{\kappa^2} \langle t, t' \times t'' \rangle \\ &= \frac{1}{\kappa^2} \langle \gamma', \gamma'' \times \gamma''' \rangle = \frac{1}{\kappa^2} \det(\gamma', \gamma'', \gamma'''). \end{aligned}$$

3.10 Beispiel. Betrachten wir die Spirale

$$\gamma : (-\infty, \infty) \rightarrow \mathbb{R}^3, \quad s \mapsto \begin{pmatrix} r \cdot \cos ks \\ r \cdot \sin ks \\ c \cdot ks \end{pmatrix}$$

mit $k := \frac{1}{\sqrt{r^2 + c^2}}$. Dann sind

$$\gamma'(s) = k \begin{pmatrix} -r \sin ks \\ r \cos ks \\ c \end{pmatrix}, \quad \gamma''(s) = k^2 \begin{pmatrix} -r \cos ks \\ -r \sin ks \\ 0 \end{pmatrix}, \quad \gamma'''(s) = k^3 \begin{pmatrix} r \sin ks \\ -r \cos ks \\ 0 \end{pmatrix}.$$

Wir erhalten somit:

$$\begin{aligned} |\gamma'(s)| &= 1, \\ \kappa(s) &= |\gamma''(s)| = k^2 \cdot r = \frac{r}{r^2 + c^2} \text{ und} \\ \tau(s) &= \frac{1}{\kappa^2(s)} \det(\gamma'(s), \gamma''(s), \gamma'''(s)) = \frac{c}{r^2 + c^2}. \end{aligned}$$

c) m -dimensionale Flächen im \mathbb{R}^n

3.11 Definition. Sei $U \subset \mathbb{R}^m$ offen, $m < n$. Die Abbildung $\gamma : U \rightarrow \mathbb{R}^n$ sei stetig differenzierbar und $\gamma' : U \rightarrow L(\mathbb{R}^m, \mathbb{R}^n)$ habe für alle $u \in U$ den Rang m . Dann heißt γ (Parametrisierung oder) „Parameterdarstellung einer m -dimensionalen Fläche“ in \mathbb{R}^n („ m -Fläche“).

Im Falle $m = 1$, also für eine 1-Fläche, erhält man gerade einen glatten Weg. Aufgrund der Rangbedingung ist für jedes feste $u \in U$

$$\tau(h) := \gamma(u) + \gamma'(u, h), \quad h \in \mathbb{R}^m,$$

ebenfalls Parameterdarstellung einer m -Fläche. τ ist in h linear, d.h. $\mathcal{R}(\tau)$ ist ein affiner Unterraum des \mathbb{R}^n , durch den $\mathcal{R}(\gamma)$ in der Nähe von $\gamma(u)$ approximiert wird. τ ist „die“ Parameterdarstellung des m -dimensionalen Tangentialraums für γ im Punkte $\gamma(u)$.

3.12 Beispiel (Rotationsflächen im \mathbb{R}^3). Anschaulich ist eine Rotationsfläche die Menge von Punkten, die man erhält, wenn man eine ebene Kurve um eine Achse, die in der Kurvenebene liegt und die Kurve nicht schneidet, dreht. Als 2-Fläche lässt sich eine Rotationsfläche durch

$$\gamma(u, v) = \begin{pmatrix} f(u) \cos v \\ f(u) \sin v \\ g(u) \end{pmatrix}$$

mit $f(u) > 0$ und $f'(u)^2 + g'(u)^2 > 0$ beschreiben. Nehmen wir also beispielsweise $f(u) = u$ als Radius und v als Winkel im \mathbb{R}^2 , so variiert $\gamma(u, v)$ für festes u nur mit v und beschreibt eine Rotation. Zur Überprüfung, dass durch $\gamma(u, v)$ tatsächlich eine 2-Fläche beschrieben wird, berechnen wir

$$\gamma'(u, v) = \begin{pmatrix} f'(u) \cos v & -f(u) \sin v \\ f'(u) \sin v & f(u) \cos v \\ g'(u) & 0 \end{pmatrix}$$

und sehen, dass $\text{rg}(\gamma') = 2$ ist.

3.13 Definition. Zwei Parameterdarstellungen $\gamma : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$, $\tilde{\gamma} : \tilde{U} \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ von m -Flächen heissen äquivalent: \Leftrightarrow

$$\exists \Phi : \tilde{U} \rightarrow U \quad \text{Diffeomorphismus mit} \quad \tilde{\gamma} = \gamma \circ \Phi.$$

Man schreibt $\gamma \sim \tilde{\gamma}$. Die durch die Parameterdarstellungen γ und $\tilde{\gamma}$ beschriebenen m -Flächen haben die „gleiche Orientierung“, wenn $\det \Phi' > 0$ ist.

Eine spezielle Parameterdarstellung $\gamma = (\gamma_1, \dots, \gamma_n)$ einer m -Fläche ist durch:

$$\begin{aligned} \gamma_1(u) &= u_1 \\ &\vdots \\ \gamma_m(u) &= u_m \\ \gamma_{m+1}(u) &= f_1(u) \\ &\vdots \\ \gamma_n(u) &= f_{n-m}(u) \end{aligned}$$

mit $u \in U \subset \mathbb{R}^m$ und $f_i \in \mathcal{C}^1(U, \mathbb{R})$ ($i = 1, \dots, n - m$) gegeben, denn es gilt:

$$\gamma'(u) = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 1 \\ \partial_1 f_1 & \dots & \partial_m f_1 \\ \vdots & & \vdots \\ \partial_1 f_{n-m} & \dots & \partial_m f_{n-m} \end{pmatrix}$$

und damit $\text{rg}(\gamma'(u)) = m$. Wir werden nun zeigen, dass jede m -Fläche (bzw. die zugehörige Parameterdarstellung) lokal äquivalent ist zu einem γ dieses speziellen und einfachen Typs. Dabei wenden wir wieder den Satz von der lokalen Umkehrbarkeit an, der später bewiesen wird.

3.14 Satz. Sei $\gamma : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine Parameterdarstellung einer m -Fläche. Für ein $u_0 \in U$ gelte

$$\det \begin{pmatrix} \partial_1 \gamma_1 & \dots & \partial_m \gamma_1 \\ \vdots & & \vdots \\ \partial_1 \gamma_m & \dots & \partial_m \gamma_m \end{pmatrix} (u_0) \neq 0.$$

Dann gibt es eine Umgebung W von u_0 mit

$$\gamma|_W \sim \tilde{\gamma} : V \rightarrow \mathbb{R}^n,$$

wobei $V \subset \mathbb{R}^m$ offen geeignet gewählt werden kann und $\tilde{\gamma}_1(v) = v_1, \dots, \tilde{\gamma}_m(v) = v_m$ sowie $\tilde{\gamma}_{m+1}(v) = f_1(v), \dots, \tilde{\gamma}_n(v) = f_{n-m}(v)$ ist.

Beweis. Man betrachte die Abbildung

$$\Psi : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^m, \quad \Psi(u) := \begin{pmatrix} \gamma_1(u) \\ \vdots \\ \gamma_m(u) \end{pmatrix}.$$

Nach dem Satz von der lokalen Umkehrbarkeit, der sich aufgrund der Voraussetzungen des Satzes anwenden lässt, gibt es eine Umgebung W von u_0 und eine offene Menge $V \subset \mathbb{R}^m$, so dass $\Psi|_W : W \rightarrow V$ ein Diffeomorphismus ist. Sei $\Phi := (\Psi|_W)^{-1} : V \rightarrow W$ und $\tilde{\gamma} := \gamma \circ \Phi : V \rightarrow \mathbb{R}^n$. Dann sind γ und $\tilde{\gamma}$ äquivalent, und es gilt nach Konstruktion:

$$\tilde{\gamma}_j(v) = v_j, \quad j = 1, \dots, m.$$

□

Lokal lässt sich also eine m -Fläche in der Form

$$\begin{aligned} x_{m+1} &= f_1(x_1, \dots, x_m) \\ &\vdots \\ x_n &= f_{n-m}(x_1, \dots, x_m) \end{aligned}$$

schreiben oder implizit als Nullstellenmenge der Funktionen g_1, \dots, g_{n-m} , die durch

$$\begin{aligned} 0 &= g_1(x) = x_{m+1} - f_1(x_1, \dots, x_m) \\ &\vdots \\ 0 &= g_{n-m}(x) = x_n - f_{n-m}(x_1, \dots, x_m) \end{aligned}$$

gegeben ist. Man sieht leicht ein, dass für $g = (g_1, \dots, g_{n-m})$ gilt: $\text{rg}(g') = n - m$.

3.15 Beispiel (Stereographische Projektion). Es sei $\gamma : \mathbb{R}^2 \rightarrow S^2 \setminus (0, 0, -1)^\top$ gegeben durch

$$\gamma(u, v) = \frac{1}{1 + u^2 + v^2} \begin{pmatrix} 2u \\ 2v \\ 1 - u^2 - v^2 \end{pmatrix},$$

dann wird der \mathbb{R}^2 auf die „gelochte Kugeloberfläche (ohne Südpol)“ abgebildet. Es gilt:

$$\gamma(0, 0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad \lim_{\alpha \rightarrow \infty} \gamma(\alpha, \alpha) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Anschaulich ist die stereographische Projektion die Abbildung, die jedem Punkt P der x - y -Ebene den Schnittpunkt der Einheitssphäre (mit Mittelpunkt im Koordinatenursprung) und der Geraden, die durch den Punkt P und den „Südpol“ $S = (0, 0, -1)^\top$ verläuft, zuordnet. Es gilt:

$$\gamma'(u, v) = \frac{2}{(1 + u^2 + v^2)^2} \begin{pmatrix} 1 - u^2 + v^2 & -2uv \\ -2uv & 1 + u^2 - v^2 \\ -2u & -2v \end{pmatrix},$$

also $\text{Rang}(\gamma') = 2$.

Für die Anwendung von Satz 3.14 berechnen wir:

$$\det \begin{pmatrix} \partial_1 \gamma_1 & \partial_2 \gamma_1 \\ \partial_1 \gamma_2 & \partial_2 \gamma_2 \end{pmatrix} = \frac{4}{(1 + u^2 + v^2)^4} (1 - (u^2 + v^2)^2) \neq 0,$$

falls (u, v) nicht auf dem Einheitskreis liegt. In einer Umgebung U von $u_0 := (0, 0)$, finden wir also eine zu γ äquivalente „einfache“ Parameterdarstellung $\tilde{\gamma}$. Es seien

$$x = \frac{2u}{1 + w^2}, \quad y = \frac{2v}{1 + w^2} \quad \text{und} \quad z = \frac{1 - w^2}{1 + w^2},$$

mit $w^2 = u^2 + v^2$. Dann gilt

$$x^2 + y^2 = \frac{4w^2}{(1+w^2)^2} \quad \text{und} \quad z^2 = \frac{(1-w^2)^2}{(1+w^2)^2} = 1 - x^2 - y^2,$$

d.h. $z = \pm\sqrt{1-x^2-y^2}$,

wobei die negative Wurzel nicht gewählt werden kann, wenn wir für U eine Umgebung des Nullpunktes betrachten, da der Südpol nicht zum Wertebereich gehört. Es ergibt sich also:

$$\tilde{\gamma}(x, y) = \begin{pmatrix} x \\ y \\ \sqrt{1-x^2-y^2} \end{pmatrix}$$

für die „einfache“ Parameterdarstellung der Kugeloberfläche, genauer: der oberen Halbkugel

$$0 \leq x^2 + y^2 \leq 1.$$

4. Grundlagen der Maß- und Integrationstheorie

4.1 Worum geht's? Eine der fundamentalen Fragen in der Mathematik ist das Messen von Flächen (oder das Berechnen eines Volumens). Aus mathematischer Sicht ist dies ein Teil der Maßtheorie. Ein Maß ist eine Abbildung, welche einer Teilmenge des \mathbb{R}^n einen Wert zuordnet. Dabei stellt sich heraus, dass man nicht beliebige Teilmengen messen kann, was auf den Begriff der σ -Algebra führt. Neben dem Maßbegriff ist auch der Begriff des zugehörigen Integrals wichtig, bei welchem z.B. Funktionen integriert werden können, welche auf einer Fläche definiert sind. Wir werden in diesem Abschnitt einen kurzen Abriss der Lebesgueschen Maß- und Integrationstheorie kennenlernen.

a) Maße und messbare Funktionen

Um mehrdimensionale Integrale definieren zu können, benötigen wir einen mathematischen Begriff, der die Fläche bzw. das Volumen einer Teilmenge des \mathbb{R}^2 bzw. \mathbb{R}^3 angibt. Anschaulich sieht man (z.B. auch für die Länge im \mathbb{R}^1), dass folgende Bedingungen für das n -dimensionale Volumen λ_n sinnvoll sind:

- (i) Das Volumen eines n -dimensionalen „Intervalls“ $\prod_{j=1}^n [a_j, b_j]$ ist das Produkt der Seitenlängen $\prod_{j=1}^n (b_j - a_j)$.
- (ii) Die leere Menge hat Volumen 0, die Werte von λ_n sind ≥ 0 .
- (iii) Das Volumen einer disjunkten Vereinigung von zwei Mengen ist die Summe der einzelnen Volumen.

Betrachtet man z.B. $(0, 1] = \bigcup_{n \in \mathbb{N}} ((\frac{1}{2})^n, (\frac{1}{2})^{n-1}]$, so scheint auch folgende „Stetigkeitsbedingung“ plausibel:

- (iv) Das Volumen einer abzählbaren disjunkten Vereinigung von Mengen ist gleich der abzählbaren Summe der einzelnen Volumina.

Da z.B. \mathbb{R}^n sicher unendliches Volumen besitzt, muss λ_n auch den Wert $+\infty$ annehmen können. Während die Bedingung (i) speziell das n -dimensionale Volumen auszeichnet, wurden die Bedingungen (ii)-(iv) verwendet, um einen allgemeinen Maßbegriff zu definieren. Dies hat den Vorteil, dass z.B. auch die Wahrscheinlichkeit von Zufallsereignissen als Maß gesehen werden kann.

Welchen Mengen kann ein Volumen zugeordnet werden? Neben den Intervallen aus Bedingung (i) sollten wegen (iv) auch alle (disjunkten) Vereinigungen von Intervallen „messbar“ sein. Allerdings lässt sich zeigen, dass nicht allen Teilmengen des \mathbb{R}^n in sinnvoller Weise ein Volumen zugeordnet werden kann. Daher wird das n -dimensionale Volumen als eine Abbildung $\lambda_n: \mathcal{A} \rightarrow [0, \infty]$ mit $\mathcal{A} \subset \mathcal{P}(X)$ definiert. Dabei ist der Definitionsbereich ein System von Mengen, welche Axiome erfüllt, die

zu den Bedingungen (i)-(iv) passen.

4.2 Definition (σ -Algebra). Sei X eine Menge, $\mathcal{P}(X) := \{A : A \subset X\}$ die Potenzmenge von X und $\mathcal{A} \subset \mathcal{P}(X)$.

Dann heißt \mathcal{A} eine σ -Algebra über X , falls gilt:

- (i) $\emptyset \in \mathcal{A}$,
- (ii) Für jedes $A \in \mathcal{A}$ gilt $A^c := \{x \in X : x \notin A\} \in \mathcal{A}$.
- (iii) Für $A_n \in \mathcal{A}$ ($n \in \mathbb{N}$) gilt $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

In diesem Fall heißt (X, \mathcal{A}) Messraum, und die Mengen $A \in \mathcal{A}$ heißen (\mathcal{A} -)messbar.

4.3 Definition-Bemerkung. (i) Die (bezüglich Mengeninklusion) größte σ -Algebra ist $\mathcal{P}(X)$, die kleinste ist $\{\emptyset, X\}$. Falls \mathcal{A}_i eine σ -Algebra ist für $i \in I$, wobei I eine nichtleere Indexmenge ist, dann ist $\bigcap_{i \in I} \mathcal{A}_i$ wieder eine σ -Algebra.

(ii) Sei $\mathcal{E} \subset \mathcal{P}(X)$ beliebig. Dann ist

$$\sigma(\mathcal{E}) := \bigcap \{ \mathcal{A} \supset \mathcal{E} : \mathcal{A} \text{ ist } \sigma\text{-Algebra über } X \}$$

die kleinste σ -Algebra, die \mathcal{E} enthält (von \mathcal{E} erzeugte σ -Algebra). In diesem Fall heißt \mathcal{E} ein Erzeugendensystem der σ -Algebra $\sigma(\mathcal{E})$.

4.4 Definition (Maß). Sei (X, \mathcal{A}) ein Messraum.

(i) Eine Abbildung $\mu : \mathcal{A} \rightarrow [0, \infty]$ ein Maß auf \mathcal{A} , falls gilt:

- (1) $\mu(\emptyset) = 0$,
- (2) σ -Additivität: Für $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ mit $A_n \cap A_m = \emptyset$ ($n \neq m$) gilt

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n). \quad (4-1)$$

In diesem Fall heißt (X, \mathcal{A}, μ) ein Maßraum.

(ii) Ein Maß μ auf einer σ -Algebra \mathcal{A} heißt

- σ -endlich (oder normal), falls es eine Folge $(A_n)_n \subset \mathcal{A}$ gibt mit $\bigcup_{n \in \mathbb{N}} A_n = X$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$,
- endlich, falls $\mu(X) < \infty$ (und damit $\mu(A) < \infty$ für alle $A \in \mathcal{A}$).
- ein Wahrscheinlichkeitsmaß, falls $\mu(X) = 1$.

(iii) Sei μ ein Maß auf \mathcal{A} . Dann heißt eine Menge $A \subset X$ eine μ -Nullmenge, falls $A \in \mathcal{A}$ und $\mu(A) = 0$ gilt. Falls für eine Aussage $M(x)$ gilt $\{x \in X : M(x) \text{ gilt nicht}\}$ ist eine μ -Nullmenge, so sagt man, die Aussage $M(x)$ gilt μ -fast überall.

4.5 Bemerkung. In obiger Definition und auch im folgenden tritt der Wert ∞ auf. Dabei sind folgende Rechenregeln zu beachten:

- $\infty \cdot 0 = 0 \cdot \infty = 0$,
- $\infty \cdot a = a \cdot \infty = \infty \quad (0 < a \leq \infty)$,
- $\infty + a = a + \infty = \infty \quad (-\infty < a \leq \infty)$.
- Der Ausdruck $\infty - \infty$ ist nicht definiert.

4.6 Beispiele. (i) Dirac-Maß: Zu $x \in X$ definiere

$$\delta_x(A) := \chi_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

Dann ist δ_x ein Maß auf $\mathcal{P}(X)$ und damit auf jeder σ -Algebra. Das Maß δ_x wird als Dirac-Maß oder auch Punktmaß bezeichnet.

(ii) Zählmaß: Definiere

$$\zeta(A) := \begin{cases} |A|, & \text{falls } A \text{ endlich,} \\ \infty, & \text{falls } A \text{ unendlich.} \end{cases}$$

Dann ist ζ ein Maß auf $\mathcal{P}(X)$, welches genau dann σ -endlich ist, falls X abzählbar ist.

4.7 Bemerkung. (i) Sei (X, \mathcal{A}, μ) ein Maßraum und $E \in \mathcal{A}$. Dann ist $\mathcal{A} \cap E := \{A \cap E : A \in \mathcal{A}\}$ eine σ -Algebra, und die Einschränkung $\mu|_E := \mu|_{\mathcal{A} \cap E}$ ist wieder ein Maß. Man erhält einen neuen Maßraum $(E, \mathcal{A} \cap E, \mu|_E)$. Man spricht von der Spur- σ -Algebra und dem Spurmaß.

(ii) Sei $X \neq \emptyset$ eine Menge, (Y, \mathcal{B}) ein Messraum und $f: X \rightarrow Y$ eine Funktion. Dann ist $\sigma(f) := f^{-1}(\mathcal{B}) := \{f^{-1}(B) : B \in \mathcal{B}\}$ eine σ -Algebra auf X , die von f erzeugte σ -Algebra.

4.8 Satz. Seien (X, \mathcal{A}) ein Messraum und $\mu: \mathcal{A} \rightarrow [0, \infty]$ eine Abbildung mit $\mu(\emptyset) = 0$ und $\mu(A \dot{\cup} B) = \mu(A) + \mu(B)$ für $A, B \in \mathcal{A}$ disjunkt. Betrachte die folgenden Aussagen:

- (a) μ ist σ -additiv.
- (b) Für alle $A_n \in \mathcal{A}$ mit $A_1 \subset A_2 \subset \dots$ und $\bigcup_{n \in \mathbb{N}} A_n =: A \in \mathcal{A}$ gilt

$$\lim_{n \rightarrow \infty} \mu(A_n) = \mu(A)$$

(d.h. μ ist stetig von unten).

(c) Für alle $A_n \in \mathcal{A}$ mit $A_1 \supset A_2 \supset \dots$, $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$ und $\mu(A_1) < \infty$ gilt

$$\lim_{n \rightarrow \infty} \mu(A_n) = 0$$

(d.h. μ ist stetig von oben).

Dann gilt (a) \iff (b) \implies (c). Falls μ endlich ist, sind alle drei Aussagen äquivalent.

Beweis. (a) \implies (b). Mit $A_0 := \emptyset$ und $\tilde{A}_n := A_n \setminus A_{n-1}$ ist $A = \bigcup_{n \in \mathbb{N}} \tilde{A}_n$ und $A_n = \bigcup_{k=1}^n \tilde{A}_k$. Also ist

$$\mu(A) = \sum_{n \in \mathbb{N}} \mu(\tilde{A}_n) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mu(\tilde{A}_k) = \lim_{n \rightarrow \infty} \mu(A_n).$$

(b) \implies (a). Sei $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ paarweise disjunkt, $A := \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$. Setze $\tilde{A}_n := A_1 \dot{\cup} \dots \dot{\cup} A_n$. Dann gilt $\tilde{A}_n \nearrow A$ (d.h. $\tilde{A}_1 \subset \tilde{A}_2 \subset \dots$ und $\bigcup \tilde{A}_n = A$), und nach (b) gilt $\mu(\tilde{A}_n) \rightarrow \mu(A)$. Wegen $\mu(\tilde{A}_n) = \sum_{k=1}^n \mu(A_k)$ gilt also $\sum_{k=1}^{\infty} \mu(A_k) = \mu(A)$.

(b) \implies (c). Wegen $\mu(A_1 \setminus A_n) = \mu(A_1) - \mu(A_n)$ und $A_1 \setminus A_n \nearrow A_1$ gilt nach (b)

$$\mu(A_1) = \lim_{n \rightarrow \infty} \mu(A_1 \setminus A_n) = \mu(A_1) - \lim_{n \rightarrow \infty} \mu(A_n)$$

und damit $\mu(A_n) \rightarrow 0$.

Sei nun μ endlich.

(c) \implies (b). Falls $A_n \nearrow A$, gilt $A \setminus A_n \searrow \emptyset$ und damit gilt $\mu(A \setminus A_n) \rightarrow 0$ nach (c). Somit folgt $\mu(A_n) \rightarrow \mu(A)$. \square

4.9 Definition (Borel- σ -Algebra). (i) Zu $a, b \in \mathbb{R}^n$ schreibe

$$a < b : \iff \forall j = 1, \dots, n : a_j < b_j,$$

analog $a \leq b$. Für $a, b \in \mathbb{R}^n$ mit $a < b$ sei $(a, b] := \{x \in \mathbb{R}^n : a_j < x_j \leq b_j (j = 1, \dots, n)\} = \prod_{j=1}^n (a_j, b_j]$ das n -dimensionale (halboffene) Intervall (Quader). Analog $[a, b]$, (a, b) , (a, b) .

(ii) Die von allen Intervallen erzeugte σ -Algebra

$$\mathcal{B}(\mathbb{R}^n) := \sigma(\{(a, b] : a, b \in \mathbb{R}^n, a < b\})$$

heißt die „Borel- σ -Algebra“ im \mathbb{R}^n . Die Mengen in $\mathcal{B}(\mathbb{R}^n)$ heißen Borel-messbar.

Es gilt sogar:

4.10 Lemma. Sei τ das Mengensystem aller offenen Teilmengen des \mathbb{R}^n . Dann gilt $\mathcal{B}(\mathbb{R}^n) = \sigma(\tau)$. Insbesondere ist jede offene Teilmenge und damit jede abgeschlossene Teilmenge Borel-messbar.

Beweis. Um für $\mathcal{E}_1, \mathcal{E}_2 \subset \mathcal{P}(X)$ die Gleichheit $\sigma(\mathcal{E}_1) = \sigma(\mathcal{E}_2)$ zu zeigen, reicht es offensichtlich, die Inklusionen $\mathcal{E}_1 \subset \sigma(\mathcal{E}_2)$ und $\mathcal{E}_2 \subset \sigma(\mathcal{E}_1)$ zu zeigen.

(i) Aus $(a, b] = \bigcup_{k \in \mathbb{N}} \prod_{j=1}^n (a_j, b_j + \frac{1}{k})$ sieht man sofort $(a, b] \in \sigma(\tau)$ für alle $a, b \in \mathbb{R}^n$ mit $a < b$.

(ii) Sei $U \subset \mathbb{R}^n$ offen. Zu jedem Punkt $x \in U$ existiert ein $\varepsilon_x > 0$ mit

$$I_x := \{y \in \mathbb{R}^n : x_i - \varepsilon_x < y_i < x_i + \varepsilon_x \ (i = 1, \dots, n)\} \subset U.$$

Wie oben sieht man $I_x \in \mathcal{B}(\mathbb{R}^n)$. Weiter gilt

$$U = \bigcup_{x \in U} I_x.$$

Diese offene Überdeckung besitzt eine abzählbare Teilüberdeckung (Übung). Dies ist aber eine abzählbare Vereinigung, und somit ist $U \in \mathcal{B}(\mathbb{R}^n)$. Also ist $\tau \subset \mathcal{B}(\mathbb{R}^n)$. \square

Der folgende Satz wird im Abschnitt Maßtheorie bewiesen.

4.11 Definition und Satz (Lebesgue-Maß). Es existiert genau ein Maß $\lambda = \lambda_n: \mathcal{B}(\mathbb{R}^n) \rightarrow [0, \infty]$ mit der Eigenschaft

$$\lambda\left(\prod_{j=1}^n (a_j, b_j]\right) = \prod_{j=1}^n (b_j - a_j)$$

für alle $a, b \in \mathbb{R}^n$ mit $a < b$. Das Maß λ_n heißt das n -dimensionale Lebesgue-Maß.

4.12 Bemerkung. Man sieht sofort folgende Eigenschaften des Lebesgue-Maßes:

(i) Die Menge $\{x\}$ ist für jeden Punkt $x \in \mathbb{R}^n$ eine λ -Nullmenge, denn $\{x\} = \bigcap_{k \in \mathbb{N}} \prod_{j=1}^n (x_j - \frac{1}{k}, x_j]$ und damit nach Satz 4.8: $\lambda(\{x\}) = \lim_{k \rightarrow \infty} (\frac{1}{k})^n = 0$.

(ii) Analog ist $\lambda(Q) = 0$ für entartete Quader $Q = \prod_{j=1}^n [a_j, b_j]$ mit mindestens einem $a_j = b_j$. Falls $U \subset \mathbb{R}^n$ ein linearer Unterraum der Form $U = \{x \in \mathbb{R}^n : x_1 = \dots = x_k = 0\}$ ist, folgt aus $U = \bigcup_{N \in \mathbb{N}} (U \cap [-N, N]^n)$ somit $\lambda(U) = 0$.

(iii) $\lambda((a, b)) = \lambda([a, b)) = \lambda((a, b]) = \lambda([a, b]) = \prod_{j=1}^n (b_j - a_j)$ für alle $a, b \in \mathbb{R}^n$ mit $a \leq b$.

(iv) $\lambda(\mathbb{Q}) = 0$, da \mathbb{Q} abzählbare Vereinigung von Nullmengen ist.

(v) λ ist σ -endlich, da $\mathbb{R}^n = \bigcup_{N \in \mathbb{N}} [-N, N]^n$ gilt.

4.13 Beispiel (Cantor-Menge). Definiere iterativ die Mengen $(C_n)_{n \in \mathbb{N}}$ durch

$$\begin{aligned} C_1 &:= [0, 1], \\ C_2 &:= C_1 \setminus \left(\frac{1}{3}, \frac{2}{3}\right), \\ C_3 &:= C_2 \setminus \left[\left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)\right], \\ &\vdots \end{aligned}$$

(d.h. man nimmt jeweils in den verbleibenden Intervallen das mittlere Drittel weg). Die Cantormenge C ist nun definiert als $C := \bigcap_{n \in \mathbb{N}} C_n$. Als abzählbarer Durchschnitt abgeschlossener Mengen ist $C \in \mathcal{B}(\mathbb{R})$. Andererseits gilt $\lambda(C_1) = 1$, $\lambda(C_2) = 1 - \frac{1}{3}$, $\lambda(C_3) = 1 - \frac{1}{3} - \frac{2}{9}$ etc. Man erhält

$$\lambda(C) = 1 - \frac{1}{3} \cdot \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = 0.$$

Die Elemente in C sind gerade die Zahlen $x \in [0, 1]$, welche eine 3-adische Entwicklung der Form

$$x = 0.a_1a_2a_3 \dots$$

mit $a_n \in \{0, 2\}$ besitzen.

Die Abbildung $x = 0.a_1a_2a_3 \dots \mapsto \sum_{n=1}^{\infty} a_n 2^{-n}$ ist surjektiv von C nach $[0, 1]$. Also besitzt die Menge C die gleiche Mächtigkeit (Kardinalität) wie das Intervall $[0, 1]$, nämlich $|[0, 1]| = |\mathbb{R}| = \mathfrak{c}$. Insbesondere ist C überabzählbar, d.h. die Cantormenge ist eine überabzählbare Lebesgue-Nullmenge.

Um das Integral bezüglich des Lebesgue-Maßes zu definieren, gehen wir ähnlich wie in Analysis I schrittweise vor: Zunächst wird das Integral für Treppenfunktionen definiert (die hier allerdings allgemeiner sind), dann wird die Definition auf allgemeinere Funktionen durch einen Grenzprozess ausgeweitet. Wir beginnen mit dem Begriff der messbaren Funktionen.

4.14 Definition (messbare Funktionen). Seien (X, \mathcal{A}) und (S, \mathcal{S}) Messräume. Für eine Abbildung $f: X \rightarrow S$ setze $f^{-1}(B) := \{x \in X : f(x) \in B\}$ und $f^{-1}(\mathcal{S}) := \{f^{-1}(B) : B \in \mathcal{S}\} \subset \mathcal{P}(X)$. Dann heißt f messbar (genauer \mathcal{A} - \mathcal{S} -messbar), falls $f^{-1}(\mathcal{S}) \subset \mathcal{A}$, d.h. falls für alle $B \in \mathcal{S}$ gilt $f^{-1}(B) \in \mathcal{A}$.

Falls $(S, \mathcal{S}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, so heißt eine \mathcal{A} - \mathcal{S} -messbare Funktion f auch \mathcal{A} -messbar. Falls auch $(X, \mathcal{A}) = (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, so heißt f Borel-messbar oder auch nur messbar.

4.15 Bemerkung. (i) Jede konstante Funktion ist messbar bezüglich jeder σ -Algebra.

b) Sind $f: (X, \mathcal{A}) \rightarrow (S_1, \mathcal{S}_1)$ und $g: (S_1, \mathcal{S}_1) \rightarrow (S_2, \mathcal{S}_2)$ messbar, so auch $g \circ f: (X, \mathcal{A}) \rightarrow (S_2, \mathcal{S}_2)$. Denn es gilt

$$(g \circ f)^{-1}(\mathcal{S}_2) = f^{-1}(g^{-1}(\mathcal{S}_2)) \subset f^{-1}(\mathcal{S}_1) \subset \mathcal{A}.$$

4.16 Lemma. Seien (X, \mathcal{A}) und (S, \mathcal{S}) Messräume und $\mathcal{S} = \sigma(\mathcal{E})$ (d.h. \mathcal{E} ist ein Erzeugendensystem von \mathcal{S}). Dann ist $f: X \rightarrow S$ genau dann \mathcal{A} - \mathcal{S} -messbar, wenn $f^{-1}(\mathcal{E}) \subset \mathcal{A}$.

Beweis. Das Mengensystem $\mathcal{S}' := \{B \subset S : f^{-1}(B) \in \mathcal{A}\}$ ist eine σ -Algebra über S . Nach Definition ist f genau dann \mathcal{A} - \mathcal{S} -messbar, wenn $\mathcal{S} = \sigma(\mathcal{E}) \subset \mathcal{S}'$. Dies ist aber äquivalent zu $\mathcal{E} \subset \mathcal{S}'$, d.h. zu $f^{-1}(\mathcal{E}) \subset \mathcal{A}$. \square

4.17 Korollar. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Dann sind äquivalent:

- (i) f ist Borel-messbar.
- (ii) Für jedes $a \in \mathbb{R}$ gilt $\{x \in X : f(x) > a\} \in \mathcal{B}(\mathbb{R}^n)$.
- (iii) Für jedes $a \in \mathbb{R}$ gilt $\{x \in X : f(x) \geq a\} \in \mathcal{B}(\mathbb{R}^n)$.
- (iv) Für jedes $a \in \mathbb{R}$ gilt $\{x \in X : f(x) < a\} \in \mathcal{B}(\mathbb{R}^n)$.
- (v) Für jedes $a \in \mathbb{R}$ gilt $\{x \in X : f(x) \leq a\} \in \mathcal{B}(\mathbb{R}^n)$.

Beweis. (i) \implies (ii) ist klar wegen $(a, \infty) \in \mathcal{B}(\mathbb{R})$.

(ii) \implies (i). Nach Lemma 4.16 ist nur zu zeigen, dass $\mathcal{E} := \{(a, \infty) : a \in \mathbb{R}\}$ ein Erzeugendensystem von $\mathcal{B}(\mathbb{R})$ ist. Dies folgt aber aus $(a, b] = (a, \infty) \setminus (b, \infty)$ für $a < b$, da damit $(a, b] \in \sigma(\mathcal{E})$ gilt.

Die Äquivalenz von (ii)–(v) folgt aus den Darstellungen

$$\begin{aligned} \{f \geq a\} &= \bigcap_{k \in \mathbb{N}} \left\{ f > a - \frac{1}{k} \right\}, \\ \{f < a\} &= \{f \geq a\}^c, \\ \{f \leq a\} &= \bigcap_{k \in \mathbb{N}} \left\{ f < a + \frac{1}{k} \right\}, \\ \{f > a\} &= \{f \leq a\}^c. \end{aligned}$$

\square

4.18 Bemerkung. Im Folgenden treten auch die Werte $\pm\infty$ als Funktionswerte auf. Dazu setze $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ und die Borel- σ -Algebra auf $\overline{\mathbb{R}}$ durch $\mathcal{B}(\overline{\mathbb{R}}) := \sigma(\{(a, \infty] \subset \overline{\mathbb{R}} : a \in \mathbb{R}\})$. Die obigen Aussagen gelten analog auch für messbare Funktionen $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$.

4.19 Satz. (i) Alle stetigen Funktionen $f: \mathbb{R}^n \rightarrow \mathbb{R}$ sind Borel-messbar.

(ii) Sei $(f_k)_{k \in \mathbb{N}}$ eine Folge (Borel-)messbarer Funktionen $f_k: \mathbb{R}^n \rightarrow \mathbb{R}$. Dann sind auch die Funktionen $\inf_{k \in \mathbb{N}} f_k$, $\sup_{k \in \mathbb{N}} f_k$, $\liminf_{n \in \mathbb{N}} f_k$ und $\limsup_{k \in \mathbb{N}} f_k$ messbar.

(iii) Der Grenzwert einer punktweise konvergenten Folge $(f_k)_{k \in \mathbb{N}}$ messbarer Funktionen ist messbar.

(iv) Seien $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ messbar und $F: \mathbb{R}^2 \rightarrow \mathbb{R}$ stetig, so ist auch $h: \mathbb{R}^n \rightarrow \mathbb{R}$, $h(x) := F(f(x), g(x))$ messbar. Insbesondere sind mit f und g auch $\max\{f, g\}$, $\min\{f, g\}$, $f \pm g$ und $f \cdot g$ messbar, ebenso $|f|^r$ für $r > 0$ und f^r für $r \in \mathbb{N}$.

Beweis. (i) Die Urbilder $\{f > a\} = f^{-1}((a, \infty))$ sind offen, also messbar.

(ii) Es gilt

$$\{\sup_k f_k \leq a\} = \bigcap_k \{f_k \leq a\} \in \mathcal{B}(\mathbb{R}^n)$$

für alle $a \in \mathbb{R}$. Mit Korollar 4.17 folgt die Behauptung für das Supremum. Wegen $\inf f_k = -\sup(-f_k)$ und $\limsup f_k = \inf_k \sup_{m \geq k} f_m$ folgt der Rest daraus, da mit f offensichtlich auch $-f$ messbar ist.

(iii) Für eine konvergente Folge gilt $\lim f_k = \limsup f_k$ und damit folgt die Behauptung aus (ii).

(iv) Da F stetig ist, ist zu $a \in \mathbb{R}$ die Menge $G_a := \{(x, y) \in \mathbb{R}^2 : F(x, y) > a\}$ offen. Wie in Beweis von Lemma 4.10 können wir schreiben $G_a = \bigcup_{k \in \mathbb{N}} A_k$ mit $A_k = (a_k, b_k) \times (c_k, d_k)$. Da f und g messbar sind, gilt

$$\{a_k < f < b_k\} = f^{-1}((a_k, b_k)) \in \mathcal{B}(\mathbb{R}^n),$$

und ebenso $\{c_k < g < d_k\} \in \mathcal{B}(\mathbb{R}^n)$. Also ist auch

$$\{(f, g) \in A_k\} = \{a_k < f < b_k\} \cap \{c_k < g < d_k\} \in \mathcal{B}(\mathbb{R}^n)$$

und damit

$$\{x \in \mathbb{R}^n : F(f(x), g(x)) > a\} = \{(f, g) \in G_a\} = \bigcup_{k \in \mathbb{N}} \{(f, g) \in A_k\} \in \mathcal{B}(\mathbb{R}^n).$$

Die restlichen Aussagen von (iv) folgen aus der Stetigkeit der Abbildungen $(x, y) \mapsto x \pm y$, $(x, y) \mapsto x \cdot y$, $(x, y) \mapsto \max\{x, y\}$, $(x, y) \mapsto \min\{x, y\}$, $(x, y) \mapsto |x|^r$ und $(x, y) \mapsto x^r$. \square

b) Das Lebesgue-Integral

Im folgenden sei (X, \mathcal{A}, μ) ein Maßraum.

4.20 Definition. (i) Eine Stufenfunktion (oder Treppenfunktion oder einfache Funktion) ist eine Funktion $f: X \rightarrow \mathbb{R}$ der Form $f = \sum_{i=1}^k c_i \chi_{A_i}$ mit $c_i \in \mathbb{R}$ und $A_i \in \mathcal{A}$. Dabei ist

$$\chi_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

die charakteristische Funktion von $A \subset X$.

(ii) $B(X, \mathcal{A}; \mathbb{R})$ bezeichne den Raum aller beschränkten \mathcal{A} -messbaren Funktionen $f: X \rightarrow \mathbb{R}$.

4.21 Satz. (i) Zu jeder messbaren Funktion $f: X \rightarrow \overline{\mathbb{R}}$ existiert eine Folge $(s_k)_{k \in \mathbb{N}}$ von Stufenfunktionen mit $s_k(x) \rightarrow f(x)$ ($x \in X$).

Falls zusätzlich $f \geq 0$ gilt, so kann man eine monoton wachsende Folge $(s_j)_j$ finden.

(ii) Falls $f: X \rightarrow \mathbb{R}$ beschränkt und messbar ist, so Folge von Stufenfunktionen, welche gleichmäßig gegen f konvergiert. (Der Raum der Stufenfunktionen liegt somit dicht in $B(X, \mathcal{A}; \mathbb{R})$ bzgl. $\|\cdot\|_\infty$ -Norm.)

Beweis. (i) Sei zunächst $f: X \rightarrow \mathbb{R}$ messbar mit $f \geq 0$. Für $k, j \in \mathbb{N}$ mit $1 \leq j \leq k \cdot 2^k$ definiere

$$A_{kj} := \left\{ x \in X : \frac{j-1}{2^k} \leq f(x) < \frac{j}{2^k} \right\} = f^{-1} \left(\left[\frac{j-1}{2^k}, \frac{j}{2^k} \right) \right)$$

und $A_k := f^{-1}([k, \infty))$. Da f messbar ist, sind A_{kj} und A_k messbar. Definiere nun

$$s_k := \sum_{j=1}^{k \cdot 2^k} \frac{j-1}{2^k} \chi_{A_{kj}} + k \cdot \chi_{A_k}.$$

Dann konvergiert die Folge $(s_k)_{k \in \mathbb{N}}$ monoton wachsend punktweise gegen f .

Im allgemeinen Fall zerlege man $f = f_+ - f_-$ mit $f_+ := \max\{f, 0\}$ und $f_- := -\min\{0, f\}$ und wende obige Konstruktion auf f_+ und f_- an.

(ii) Die obige Konstruktion zeigt, dass die Folge $(s_k)_{k \in \mathbb{N}}$ bei beschränktem messbaren f sogar gleichmäßig konvergiert. \square

4.22 Definition (allgemeines Lebesgue-Integral). (i) Sei $s = \sum_{j=1}^k c_j \chi_{A_j}$ mit $c_j \in \mathbb{R}$ und $A_j \in \mathcal{A}$ eine Stufenfunktion mit $s \geq 0$. Definiere das Integral von s bzgl. μ

durch

$$\int s d\mu := \sum_{j=1}^k c_j \mu(A_j) \in [0, \infty].$$

(ii) Sei nun $f: X \rightarrow \mathbb{R}$ messbar mit $f \geq 0$. Definiere

$$\int f d\mu := \sup \left\{ \int s d\mu : s \text{ Stufenfunktion mit } 0 \leq s \leq f \right\} \in [0, \infty].$$

(iii) Falls $f: X \rightarrow \mathbb{R}$ messbar, definiert man

$$\int f d\mu := \int f_+ d\mu - \int f_- d\mu, \quad (4-2)$$

falls nicht beide Integrale den Wert $+\infty$ haben.

(iv) Eine Funktion $f: X \rightarrow \mathbb{R}$ heißt (Lebesgue-)integrierbar, falls f messbar ist und beide Integrale in (4-2) endlich sind. Die Menge aller integrierbaren Funktionen wird mit $\mathcal{L}^1(\mu) = \mathcal{L}^1(X, \mu) = \mathcal{L}^1(X)$ bezeichnet. Andere Schreibweisen sind etwa $\int f(x) d\mu(x) := \int f \mu(dx) := \int f d\mu$. Der Index „1“ wird manchmal unten geschrieben: $\mathcal{L}_1(\mu)$.

Falls $\mu = \lambda$ das Lebesgue-Maß im \mathbb{R}^n ist, so schreibt man $\int f(x) dx$.

(v) Für $A \in \mathcal{A}$ definiert man

$$\int_A f d\mu := \int \chi_A f d\mu.$$

Für $A \in \mathcal{B}(\mathbb{R}^n)$ schreiben wir $\mathcal{L}^1(A) := \mathcal{L}^1(\lambda|_A)$, wobei λ das n -dimensionale Lebesgue-Maß bezeichne.

4.23 Satz (Elementare Eigenschaften des Lebesgue-Integrals). Sei $f: X \rightarrow \mathbb{R}$ messbar und $A \in \mathcal{A}$.

(i) Sei zusätzlich f beschränkt und $\mu(X) < \infty$. Dann ist $f \in \mathcal{L}^1(\mu)$.

(ii) Monotonie: Sind $f, g \in \mathcal{L}^1(\mu)$ mit $f \leq g$, so ist

$$\int f d\mu \leq \int g d\mu.$$

Speziell gilt: Ist $a \leq f(x) \leq b$ ($x \in A$) und $\mu(A) < \infty$, so gilt

$$a\mu(A) \leq \int_A f d\mu \leq b\mu(A).$$

(iii) Ist A eine μ -Nullmenge, so gilt $\int_A f d\mu = 0$.

(iv) Sind $A, B \in \mathcal{A}$ disjunkt und $f \in \mathcal{L}^1(X, \mu)$, so gilt $\int_{A \cup B} f d\mu = \int_A f d\mu + \int_B f d\mu$.

(v) Ist $f \in \mathcal{L}^1(X, \mu)$, so ist $f \in \mathcal{L}^1(A, \mu|_A)$.

Beweis. (i) Zu $f \in B(X, \mathcal{A}, \mathbb{R})$ existiert ein $K > 0$ mit $0 \leq f_{\pm} \leq K$. Für jede Stufenfunktion s mit $0 \leq f_+$ gilt somit $\int s d\mu \leq K\mu(X)$. Somit gilt $\int f_+ d\mu \leq K\mu(X) < \infty$. Analog sieht man $\int f_- d\mu < \infty$.

(ii) Ist $f \leq g$, so folgt $f_+ \leq g_+$ und $f_- \geq g_-$. Für jede Stufenfunktion s mit $0 \leq s \leq f_+$ gilt daher auch $s \leq g_+$. Damit folgt $\int f_+ d\mu \leq \int g_+ d\mu$. Genauso folgt $\int f_- d\mu \geq \int g_- d\mu$ und damit die erste Behauptung.

Die zweite Behauptung folgt mit $a\chi_A \leq f\chi_A \leq b\chi_A$.

(iii) Für jede Stufenfunktion $s = \sum_{j=1}^k c_j \chi_{A_j}$ mit $0 \leq s \leq f_+$ gilt

$$\int_A s d\mu = \sum_{j=1}^k c_j \mu(A \cap A_j) = 0$$

wegen $\mu(A \cap A_j) \leq \mu(A) = 0$. Damit ist $\int_A f_+ d\mu = 0$. Analog folgt $\int_A f_- d\mu = 0$.

(iv) Die Gleichheit gilt nach Definition des Integrals für Stufenfunktionen. Für integrierbare Funktionen $f \geq 0$ und Stufenfunktionen $0 \leq s \leq f$ gilt $0 \leq s\chi_A \leq f\chi_A$ und $0 \leq s\chi_B \leq f\chi_B$, andererseits ist für zwei Stufenfunktionen s_1, s_2 mit $s_1 \leq f\chi_A$ und $s_2 \leq f\chi_B$ auch $s := s_1\chi_A + s_2\chi_B$ eine Stufenfunktion mit $s \leq f$. Geht man zum Supremum über, folgt die Behauptung für integrierbares $f \geq 0$ und damit für $f \in \mathcal{L}^1(\mu)$.

(v) Für jede Stufenfunktion $s = \sum_{j=1}^k c_j \chi_{A_j}$ mit $0 \leq s \leq f_+$ gilt

$$\int_A s d\mu = \sum_{j=1}^k c_j \mu(A \cap A_j) \leq \sum_{j=1}^k c_j \mu(A_j) = \int s d\mu.$$

Damit folgt $\int_A f_+ d\mu \leq \int f_+ d\mu < \infty$, analog für f_- . □

4.24 Lemma. Sei $f: X \rightarrow \mathbb{R}$ messbar.

(i) Ist $f \geq 0$ mit $\int f d\mu = 0$, so ist $f = 0$ μ -f.ü..

(ii) Es gilt $f \in \mathcal{L}^1(\mu)$ genau dann, wenn $|f| \in \mathcal{L}^1(\mu)$ gilt. In diesem Fall ist

$$\left| \int f d\mu \right| \leq \int |f| d\mu.$$

(iii) (Majorantenkriterium) Sei $g \in \mathcal{L}^1(\mu)$ mit $|f| \leq g$ μ -f.ü.. Dann ist $f \in \mathcal{L}^1(\mu)$.

Beweis. (i) Wir schreiben $\{f \neq 0\} = \bigcup_{j=1}^{\infty} \{\frac{1}{j+1} \leq f < \frac{1}{j}\} \cup \{f \geq 1\}$. Falls $\mu(\{f \neq 0\}) > 0$, so hat wegen der σ -Additivität von μ eine der Mengen auf der rechten Seite positives Maß und damit nach Satz 4.23 (ii) $\int f d\mu > 0$.

(ii) Die Mengen $A := \{f \geq 0\}$ und $B := \{f < 0\}$ sind messbar, ebenso die Funktion $|f|$. Nach Satz 4.23 (iv) gilt

$$\int |f| d\mu = \int_A |f| d\mu + \int_B |f| d\mu = \int f_+ d\mu + \int f_- d\mu < \infty.$$

Die Abschätzung folgt aus $-|f| \leq f \leq |f|$ und der Monotonie des Integrals (Satz 4.23 (ii)).

(iii) Nach Änderung auf einer Nullmenge können wir o.E. $|f| \leq g$ annehmen. Dann ist $f_- \leq g$ und $f_+ \leq g$ und somit $\int f_{\pm} d\mu < \infty$. \square

Der Vorteil des soeben definierten Lebesgue-Integrals liegt zum einen an der großen Allgemeinheit (beliebige Maße), zum anderen an starken Konvergenzaussagen. Die folgenden Aussagen werden hier nur zitiert und im Abschnitt Maßtheorie bewiesen.

Im folgenden sei (X, \mathcal{A}, μ) ein Maßraum.

4.25 Satz (Linearität des Integrals). Seien $f_1, f_2 \in \mathcal{L}^1(\mu)$, $\alpha_1, \alpha_2 \in \mathbb{R}$. Dann ist $\alpha_1 f_1 + \alpha_2 f_2 \in \mathcal{L}^1(\mu)$ und

$$\int (\alpha_1 f_1 + \alpha_2 f_2) d\mu = \alpha_1 \int f_1 d\mu + \alpha_2 \int f_2 d\mu.$$

4.26 Satz (von Lebesgue über monotone Konvergenz). Sei $(f_k)_{k \in \mathbb{N}}$ eine Folge messbarer Funktionen mit $0 \leq f_1 \leq f_2 \leq \dots$, und sei $f: X \rightarrow [0, \infty]$ definiert durch $f(x) := \lim_{k \rightarrow \infty} f_k(x)$ ($x \in X$). Dann gilt

$$\lim_{k \rightarrow \infty} \int f_k d\mu = \int f d\mu = \int \lim_{k \rightarrow \infty} f_k d\mu.$$

4.27 Satz. Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge nichtnegativer messbarer Funktionen und $f: X \rightarrow [0, \infty]$ definiert durch $f := \sum_{n=1}^{\infty} f_n$. Dann gilt

$$\int f d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

Falls $f_n \in \mathcal{L}^1(\mu)$ ($n \in \mathbb{N}$) und die Summe auf der rechten Seite konvergiert, so gilt $f \in \mathcal{L}^1(\mu)$, und die Reihe $\sum_{n=1}^{\infty} f_n(x)$ konvergiert für μ -fast alle $x \in X$.

4.28 Satz (Lemma von Fatou). Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge nichtnegativer messbarer Funktionen und $f: X \rightarrow [0, \infty]$ definiert durch $f := \liminf_{n \rightarrow \infty} f_n$. Dann gilt

$$\int f d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu.$$

Der folgende Satz ist einer der wichtigsten Sätze der Integrationstheorie und heißt auch Satz von Lebesgue über dominierte Konvergenz.

4.29 Satz (von Lebesgue über majorisierte Konvergenz). Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge messbarer Funktionen, und der Grenzwert $f(x) := \lim_{n \rightarrow \infty} f_n(x) \in \overline{\mathbb{R}}$ existiere μ -fast überall. Weiter existiere eine Funktion $g \in \mathcal{L}^1(\mu)$ mit $|f_n(x)| \leq g(x)$ für μ -fast alle $x \in X$ und alle $n \in \mathbb{N}$. Dann ist $f \in \mathcal{L}^1(\mu)$ und

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int \lim_{n \rightarrow \infty} f_n d\mu.$$

Im folgenden werden wir stets den Maßraum $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$ mit $\lambda = \lambda_n$ betrachten. In diesem Fall schreibt man auch

$$\int_A f(x) dx := \int_A f(x) d\lambda(x) := \int_A f d\lambda.$$

Ein weiterer wichtiger Integralbegriff wurde schon im ersten Semester erwähnt: Das Riemann-Integral.

4.30 Definition (Riemann-Integral). (i) Eine Stufenfunktion $s: \mathbb{R}^n \rightarrow \mathbb{R}$, $s(x) = \sum_{i=1}^k c_i \chi_{A_i}$ heißt eine Riemann-Stufenfunktion, falls A_i n -dimensionale Intervalle sind, d.h. $A_i = (a_i, b_i]$ mit $a_i, b_i \in \mathbb{R}^n$, $a_i < b_i$.

(ii) Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ eine beschränkte Funktion mit kompaktem Träger

$$\text{supp } f := \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}.$$

Dann heißt

$$\overline{\int} f(x) dx := \inf \left\{ \int s(x) dx : s \text{ Riemann-Stufenfunktion, } s \geq f \right\}$$

das Oberintegral von f und

$$\underline{\int} f(x) dx := \sup \left\{ \int s(x) dx : s \text{ Riemann-Stufenfunktion, } s \leq f \right\}$$

das Unterintegral von f .

f heißt Riemann-integrierbar, wenn $\overline{\int} f(x) dx = \underline{\int} f(x) dx =: \int f(x) dx$ (Riemann-Integral).

4.31 Satz. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ eine beschränkte Funktion mit kompaktem Träger.

(i) Falls f gleichmäßig stetig ist, so ist f Riemann-integrierbar.

(ii) Falls f Riemann-integrierbar ist, so ist f (nach eventueller Abänderung auf einer Nullmenge) Lebesgue-integrierbar, und

$$\int f(x)dx = \int f d\lambda,$$

wobei links das Riemann-Integral und rechts das Lebesgue-Integral steht.

Beweis. (i) Wir wählen ein Intervall $(-N, N]^n$ mit $\text{supp } f \subset (-N, N]^n$ und dann eine Folge von Partitionen

$$(-N, N]^n = \bigcup_{k=1}^{m_j} I_{kj} \quad (j = 1, 2, \dots)$$

mit Intervallen I_{kj} , deren Feinheit gegen 0 konvergiert. Setze für $j \in \mathbb{N}$

$$s_j := \sum_{k=1}^{m_j} c_{kj} \chi_{I_{kj}}, \quad S_j := \sum_{k=1}^{m_j} C_{kj} \chi_{I_{kj}}$$

mit

$$c_{kj} := \inf_{y \in I_{kj}} f(y), \quad C_{kj} := \sup_{y \in I_{kj}} f(y).$$

Da f gleichmäßig stetig ist, gilt $\int S_j(x)dx - \int s_j(x)dx \rightarrow 0$ ($j \rightarrow \infty$). Damit ist f Riemann-integrierbar und $\int f(x)dx = \lim_{j \rightarrow \infty} \int S_j(x)dx = \lim_{j \rightarrow \infty} \int s_j(x)dx$.

(ii) Sei f Riemann-integrierbar. Definiert man s_j und S_j wie oben, so gilt

$$s_1 \leq s_2 \leq \dots \leq f \leq \dots \leq S_2 \leq S_1.$$

Als Grenzwerte monotoner Folgen existieren punktweise $s(x) := \lim_{j \rightarrow \infty} s_j(x)$ und $S(x) := \lim_{j \rightarrow \infty} S_j(x)$ und sind wieder messbar. Es gilt $s \leq f \leq S$.

Nach dem Satz von der monotonen Konvergenz gilt

$$\int s d\lambda = \lim_{j \rightarrow \infty} \int s_j d\lambda = \lim_{j \rightarrow \infty} \int s_j(x) dx,$$

letzteres, da auf Riemann-Stufenfunktionen Riemann- und Lebesgue-Integral nach Definition übereinstimmen. Genauso gilt

$$\int S d\lambda = \lim_{j \rightarrow \infty} \int S_j(x) dx.$$

Da f Riemann-integrierbar ist, gilt außerdem

$$\lim_{j \rightarrow \infty} \int s_j(x) dx = \lim_{j \rightarrow \infty} \int S_j(x) dx$$

und damit $\int (S - s)d\lambda = 0$. Nach Lemma 4.24 folgt $S = s$ λ -fast überall und somit wegen $s \leq f \leq S$ auch $s = S = f$ fast überall. Damit ist f nach eventueller Änderung auf einer Nullmenge messbar mit

$$\int f d\lambda = \int s d\lambda = \int f(x) dx.$$

Beachte, dass wegen Beschränktheit von f und Kompaktheit des Trägers von f automatisch $f \in \mathcal{L}^1(\lambda)$ gilt. \square

c) Iterierte Integrale

4.32 Beispiel. Sei $M = \{x \in [0, 1]^2 : x_2 \leq x_1^2\}$. Dann ist

$$\begin{aligned} \int_M |x|^2 dx &= \int_M (x_1^2 + x_2^2) dx \stackrel{(*)}{=} \int_0^1 \int_0^{x_1^2} (x_1^2 + x_2^2) dx_2 dx_1 \\ &= \int_0^1 (x_1^4 + \frac{x_1^6}{3}) dx_1 = (\frac{x_1^5}{5} + \frac{x_1^7}{7 \cdot 3}) \Big|_0^1 = \frac{1}{5} + \frac{1}{21} = \frac{26}{105}. \end{aligned}$$

Frage: Ist (*) erlaubt? Die Antwort (ja) liefert der Satz von Fubini.

4.33 Satz (von Tonelli). Sei $f: \mathbb{R}^2 \rightarrow [0, \infty]$ messbar. Dann sind $y \mapsto \int_{\mathbb{R}} f(x, y) dx, \mathbb{R} \rightarrow [0, \infty]$ für jedes feste x und $x \mapsto \int_{\mathbb{R}} f(x, y) dy, \mathbb{R} \rightarrow [0, \infty]$ messbar (analog für vertauschtes x und y), und es gilt

$$\int_{\mathbb{R}^2} f(x, y) d(x, y) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) dx \right) dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) dy \right) dx \in [0, \infty].$$

Beweis. Wir zeigen den Satz nur für den Spezialfall, dass f Riemann-integrierbar ist. Nach dem Beweis von Satz 4.31 (ii) existiert dann eine Folge $(s_k)_{k \in \mathbb{N}}$ von Riemann-Stufenfunktionen mit $s_k \nearrow f$ punktweise.

Sei $A = (a, b], a, b \in \mathbb{R}^2, a < b$. Dann sind folgende Funktionen offensichtlich messbar:

$$\begin{aligned} y &\mapsto \chi_A(x, y), \mathbb{R} \rightarrow \mathbb{R}, \quad \text{für festes } x, \\ x &\mapsto \chi_A(x, y), \mathbb{R} \rightarrow \mathbb{R}, \quad \text{für festes } y, \\ y &\mapsto \int_{\mathbb{R}} \chi_A(x, y) dx, \mathbb{R} \rightarrow \mathbb{R}, \\ x &\mapsto \int_{\mathbb{R}} \chi_A(x, y) dy, \mathbb{R} \rightarrow \mathbb{R}. \end{aligned}$$

Weiter ist

$$\int_{\mathbb{R}^2} \chi_A d(x, y) = \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_A(x, y) dx dy = \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_A(x, y) dy dx = (b_1 - a_1)(b_2 - a_2).$$

Wegen der Linearität des Integrals folgen die Behauptungen für alle Riemann-Stufenfunktionen s_k . Als punktwiser Limes messbarer Funktionen sind somit $y \mapsto f(x, y) = \lim_k s_k(x, y)$ und $x \mapsto f(x, y)$ messbar. Mit dem Satz über monotone Konvergenz folgt die Messbarkeit von $y \mapsto \int f(x, y) dx = \lim_k \int s_k(x, y) dx$ und $x \mapsto \int f(x, y) dy$. Wieder mit monotoner Konvergenz folgt

$$\begin{aligned} \int_{\mathbb{R}^2} f(x, y) d(x, y) &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^2} s_k(x, y) d(x, y) \\ &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}} \int_{\mathbb{R}} s_k(x, y) dx dy = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) dx dy. \end{aligned}$$

□

4.34 Satz (von Fubini). Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ (Lebesgue-)integrierbar. Dann sind (eventuell nach Änderung auf einer Nullmenge) die Funktionen $y \mapsto \int_{\mathbb{R}} f(x, y) dx$ und $x \mapsto \int_{\mathbb{R}} f(x, y) dy$ integrierbar, und es gilt

$$\int_{\mathbb{R}^2} f(x, y) d(x, y) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) dx \right) dy = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x, y) dy \right) dx.$$

Beweis. Der Satz von Tonelli, angewendet auf $|f|$, liefert die Messbarkeit von $x \mapsto \int |f(x, y)| dy$. Damit ist

$$N := \left\{ x \in \mathbb{R} : \int_{\mathbb{R}} |f(x, y)| dy = \infty \right\}$$

Borel-messbar, und wegen

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x, y)| dy \right) dx = \int_{\mathbb{R}^2} |f(x, y)| d(x, y) < \infty$$

ist $\lambda_1(N) = 0$. Definiert man

$$\tilde{f}(x, y) := \begin{cases} f(x, y), & \text{falls } x \notin N, \\ 0, & \text{sonst,} \end{cases}$$

und $g(x) := \int_{\mathbb{R}} \tilde{f}(x, y) dy$, so folgt $\tilde{f} = f$ fast überall und $g \in \mathcal{L}^1(\mathbb{R})$.

Falls $f \geq 0$ (und damit auch $\tilde{f} \geq 0, g \geq 0$), so folgt wieder mit dem Satz von Tonelli

$$\int_{\mathbb{R}} g(x) dx = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \tilde{f}(x, y) dy \right) dx = \int_{\mathbb{R}^2} \tilde{f}(x, y) d(x, y) = \int_{\mathbb{R}^2} f(x, y) d(x, y).$$

Für beliebiges $f \in \mathcal{L}^1(\mathbb{R}^2)$ folgt die Behauptung durch die Zerlegung $f = f_+ - f_-$. □

4.35 Korollar. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ integrierbar. Dann gilt

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_n) dx_1 \dots dx_n,$$

wobei die Integrationsreihenfolge auf der rechten Seite frei wählbar ist.

Beweis. Iteriertes Anwenden des Satzes von Fubini. □

4.36 Korollar (Prinzip von Cavalieri⁹). Für $A \in \mathcal{B}(\mathbb{R}^n)$ und $t \in \mathbb{R}$ sei

$$A_t := \{(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1} : (x_1, \dots, x_{n-1}, t) \in A\}$$

der Schnitt zum Wert t . Dann ist $A_t \in \mathcal{B}(\mathbb{R}^{n-1})$, und

$$\lambda_n(A) = \int_{\mathbb{R}} \lambda_{n-1}(A_t) dt.$$

Beweis. Dies folgt durch die Anwendung des Satzes von Tonelli auf χ_A . □

4.37 Korollar. Sei $D \in \mathcal{B}(\mathbb{R}^{n-1})$ und $f: D \rightarrow \mathbb{R}$ Borel-messbar. Dann gilt $\text{graph } f \in \mathcal{B}(\mathbb{R}^n)$ und $\lambda_n(\text{graph } f) = 0$. Insbesondere hat jede Hyperebene im \mathbb{R}^n Lebesgue-Maß 0.

Beweis. O.E. sei $D = \mathbb{R}^{n-1}$, denn sonst setze f durch 0 zu einer messbaren Funktion \tilde{f} auf \mathbb{R}^n fort. Die Behauptung für f folgt dann aus $\text{graph } f = (D \times \mathbb{R}) \cap \text{graph } \tilde{f}$.

Die Koordinatenprojektionen $\text{pr}_i: x \mapsto x_i$, $i = 1, \dots, n$, sind stetig und damit Borel-messbar. Also ist auch

$$g: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto f(x_1, \dots, x_{n-1}) - x_n$$

Borel-messbar. Damit ist $\text{graph } f = \{g = 0\} \in \mathcal{B}(\mathbb{R}^n)$. Wegen $\chi_{\text{graph } f}(x_1, \dots, x_{n-1}, t) = 0$ für $t \neq f(x_1, \dots, x_{n-1})$ folgt

$$\int_{\mathbb{R}} \chi_{\text{graph } f}(x', t) d\lambda(t) = 0$$

für alle $x' := (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$. Nach dem Satz von Tonelli folgt

$$\lambda_n(\text{graph } f) = \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} \chi_{\text{graph } f}(x', t) d\lambda_1(t) \right) d\lambda_{n-1}(x') = \int_{\mathbb{R}^{n-1}} 0 d\lambda(x') = 0.$$

□

⁹Bonaventura Cavalieri 1598-30.11.1647

4.38 Beispiel (Kreisfläche, Kugelvolumen). (i) Sei $M := \{(x, y)^\top \in \mathbb{R}^2 : |(x, y)^\top| \leq r\}$ der Kreis mit Radius r . Nach dem Prinzip von Cavalieri erhält man für die Kreisfläche

$$\lambda_2(M) = \int_{-r}^r \lambda_1([- \sqrt{r^2 - x^2}, \sqrt{r^2 - x^2}]) dx = 4 \int_0^r \sqrt{r^2 - x^2} dx.$$

Mit der Substitution $x = r \sin t$, $dx = r \cos t dt$, $t \in [0, \frac{\pi}{2}]$, erhält man

$$\lambda_2(M) = 4 \int_0^{\pi/2} r^2 \cos^2 t dt = 4r^2 \frac{\pi}{4} = \pi r^2.$$

(ii) Sei nun $K := \{(x, y, z)^\top \in \mathbb{R}^3 : |(x, y, z)^\top| \leq r\}$ die Kugel im \mathbb{R}^3 mit Radius r . Für $x \in [-r, r]$ ist der Schnitt zum Wert x gegeben durch $K_x = \{(y, z) \in \mathbb{R}^2 : y^2 + z^2 \leq r^2 - x^2\}$, d.h. nach (i) ist $\lambda_2(K_x) = \pi(r^2 - x^2)$. Damit ist

$$\lambda_3(K) = \int_{-r}^r \pi(r^2 - x^2) dx = 2\pi(r^3 - \frac{r^3}{3}) = \frac{4}{3} \pi r^3.$$

Somit besitzt ein Kreis im \mathbb{R}^2 mit Radius r die Fläche πr^2 und eine Kugel im \mathbb{R}^3 mit Radius r das Volumen $\frac{4}{3} \pi r^3$.

4.39 Beispiel (Trägheitsmoment einer Kugel). Sei $K := \{(x, y, z)^\top \in \mathbb{R}^3 : |(x, y, z)^\top| \leq r\}$ (Kugel mit Radius $r > 0$). Das Trägheitsmoment von K bzgl. der z -Achse ist gegeben durch $J := \int_K (x^2 + y^2) d(x, y, z)$. Nach dem Satz von Fubini ist

$$\begin{aligned} J &= 2 \int_K x^2 d(x, y, z) = \int_{-r}^r \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \int_{-\sqrt{r^2-x^2-y^2}}^{\sqrt{r^2-x^2-y^2}} x^2 dz dy dx \\ &= 16 \int_0^r \int_0^{\sqrt{r^2-x^2}} \int_0^{\sqrt{r^2-x^2-y^2}} x^2 dz dy dx \\ &= 16 \int_0^r \int_0^{\sqrt{r^2-x^2}} x^2 \sqrt{r^2 - x^2 - y^2} dy dx. \end{aligned}$$

Substituiert man $y = \sqrt{r^2 - x^2} \sin t$, $dy = \sqrt{r^2 - x^2} \cos t dt$, so erhält man

$$\begin{aligned} J &= 16 \int_0^r x^2 \int_0^{\pi/2} \sqrt{r^2 - x^2} \sqrt{1 - \sin^2 t} \sqrt{r^2 - x^2} \cos t dt dx \\ &= 16 \int_0^r x^2 (r^2 - x^2) dx \int_0^{\pi/2} \cos^2 t dt \\ &= 16 \left(\frac{r^5}{3} - \frac{r^5}{5} \right) \cdot \frac{\pi}{4} = \frac{8}{15} \pi r^5. \end{aligned}$$

d) Der Transformationssatz

Im folgenden sei $\lambda = \lambda_n$ das Lebesgue-Maß im \mathbb{R}^n .

4.40 Definition und Satz. (i) Sei $U \subset \mathbb{R}^n$ offen, μ ein Maß auf $(U, \mathcal{B}(U))$, und $\Phi: U \rightarrow \mathbb{R}^n$ messbar. Dann ist $\mu \circ \Phi^{-1}: \mathcal{B}(\mathbb{R}^n) \rightarrow [0, \infty]$, $A \mapsto \mu(\Phi^{-1}(A))$ ein Maß, das Bildmaß von μ unter Φ .

(ii) Ein Maß μ auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ heißt translationsinvariant, falls $\mu \circ \Phi^{-1} = \mu$ für alle Translationen $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto x + c$ mit $c \in \mathbb{R}^n$ gilt. Das Maß μ heißt bewegungsinvariant, falls $\mu \circ \Phi^{-1} = \mu$ gilt für alle Bewegungen $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Sx + c$ mit $S \in \mathbb{R}^{n \times n}$ orthogonale Matrix und $c \in \mathbb{R}^n$.

Beweis. Die σ -Additivität folgt sofort aus

$$\lambda\left(\Phi^{-1}\left(\bigcup_{k \in \mathbb{N}} A_k\right)\right) = \lambda\left(\bigcup_{k \in \mathbb{N}} \Phi^{-1}(A_k)\right) = \sum_{k \in \mathbb{N}} \lambda(\Phi^{-1}(A_k)),$$

die anderen Eigenschaften eines Maßes sind klar. □

4.41 Satz. (i) λ ist translationsinvariant.

(ii) Sei $W := ((0, \dots, 0), (1, \dots, 1]) = (0, 1]^n$ der n -dimensionale Einheitswürfel. Falls μ ein translationsinvariantes Maß auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ mit $\mu(W) =: \alpha < \infty$ ist, so ist $\mu = \alpha\lambda$.

Beweis. (i) Für $a, b \in \mathbb{R}^n$ mit $a < b$ und $\Phi(x) = x + c$, $c \in \mathbb{R}^n$, gilt $\lambda \circ \Phi^{-1}((a, b]) = \lambda((a - c, b - c]) = \prod_{j=1}^n (b_j - a_j) = \lambda((a, b])$. Nach Satz 4.11 ist $\lambda \circ \Phi^{-1} = \lambda$.

(ii) Sei $a_k := (\frac{1}{k}, \dots, \frac{1}{k})$, $W_k := (0, a_k]$ und $G_k := \{\frac{1}{k}, \frac{2}{k}, \dots, 1\}^n$. Dann ist $W = W_1 = \bigcup_{r \in G_k} (r - a_k, r]$ und damit $\mu(W) = \alpha = k^n \mu(W_k)$. Wegen $\alpha < \infty$ folgt $\mu(W_k) = \frac{\alpha}{k^n} = \alpha \lambda(W_k)$.

Betrachte nun $(a, b]$ mit $a, b \in \mathbb{Q}^n$, wobei o.E. $a = 0$, $b = (\frac{m_1}{k}, \dots, \frac{m_n}{k})$. Dann ist $(0, b] = \bigcup_{r \in H_k} (r - a_k, r]$ mit der Indexmenge $H_k := \prod_{i=1}^n \{\frac{1}{k}, \dots, \frac{m_i}{k}\}$. Damit

$$\mu((0, b]) = \sum_{r \in H_k} \mu((r - a_k, r]) = \alpha \cdot \frac{m_1}{k} \cdot \dots \cdot \frac{m_n}{k} = \alpha \lambda((0, b]).$$

Somit gilt $\mu((a, b]) = \alpha \lambda((a, b])$ für alle $a, b \in \mathbb{Q}^n$ mit $a < b$.

Zu $a, b \in \mathbb{R}^n$, $a < b$, wähle nun $a_k, b_k \in \mathbb{Q}^n$ mit $(a_k, b_k] \nearrow (a, b]$ und erhalte $\mu((a, b]) = \alpha \lambda((a, b])$. Nach Satz 4.11 folgt $\mu = \alpha\lambda$. □

4.42 Satz. λ ist bewegungsinvariant.

Beweis. Sei $\Phi(x) = Sx + c$, d.h. $\Phi = T_c \circ S$ mit $S: \mathbb{R}^n \rightarrow \mathbb{R}^n$ orthogonal und $T_c: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto x + c$ mit $c \in \mathbb{R}^n$. Dann gilt $\Phi = T_c \circ S = S \circ T_d$ mit $d := S^{-1}c$ und damit

$$\lambda \circ \Phi^{-1} = (\lambda \circ S^{-1}) \circ T_c^{-1} = (\lambda \circ T_d^{-1}) \circ S^{-1} = \lambda \circ S^{-1} \quad (4-3)$$

nach Satz 4.41 (i). Zu zeigen ist also noch $\lambda \circ S^{-1} = \lambda$.

Da $S^{-1}(W) \subset \mathbb{R}^n$ beschränkt ist, folgt $\alpha := (\lambda \circ S^{-1})(W) < \infty$. Nach (4-3) ist $\lambda \circ S^{-1}$ translationsinvariant, und nach Satz 4.41 (ii) folgt $\lambda \circ S^{-1} = \alpha \cdot \lambda$. Für $K := \overline{B(0, 1)} = \{x \in \mathbb{R}^n : |x| \leq 1\}$ ist $S^{-1}(K) = K$, d.h. $(\lambda \circ S^{-1})(K) = \lambda(S^{-1}(K)) = \lambda(K) > 0$ und damit $\alpha = 1$. \square

4.43 Satz. Sei $T \in GL(n, \mathbb{R})$. Dann gilt

$$(\lambda \circ T^{-1})(B) = \frac{1}{|\det T|} \lambda(B) \quad (B \in \mathcal{B}(\mathbb{R}^n))$$

bzw.

$$\lambda(T(B)) = |\det T| \cdot \lambda(B) \quad (B \in \mathcal{B}(\mathbb{R}^n)).$$

Beweis. Wie im Beweis von Satz 4.41 sieht man, dass $\lambda \circ T^{-1}$ translationsinvariant ist. Da $\alpha := (\lambda \circ T^{-1})(W) = \lambda(T^{-1}(W)) < \infty$, folgt aus Satz 4.41 (ii) $\lambda \circ T^{-1} = \alpha \cdot \lambda$.

Nach einem Satz der linearen Algebra existieren orthogonale Matrizen S_1, S_2 und eine Diagonalmatrix $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ mit $T = S_1 D S_2$. Für $B := S_1(W)$ folgt

$$\begin{aligned} \alpha &= \alpha \lambda(W) = \alpha \cdot (\lambda \circ S_1)(W) = \alpha \lambda(B) \\ &= (\lambda \circ T^{-1})(B) = \lambda(S_2^{-1} \circ D^{-1} \circ S_1^{-1}(B)) \\ &= \lambda \circ S_2^{-1}(D^{-1}(W)) = \lambda(D^{-1}(W)) \\ &= \lambda\left(\prod_{j=1}^n (0, |\lambda_j|^{-1}]\right) = \prod_{j=1}^n |\lambda_j|^{-1} = |\det D|^{-1} = |\det T|^{-1}. \end{aligned}$$

\square

Für lineare bijektive Abbildungen haben wir also die Gleichheit $\lambda(T(B)) = |\det T| \cdot \lambda(B)$ ($B \in \mathcal{B}(\mathbb{R}^n)$). Seien nun $U, V \subset \mathbb{R}^n$ offen und $\Phi: U \rightarrow V$, $z \mapsto \Phi(z) =: x$, ein C^1 -Diffeomorphismus. Dann gilt $x_2 - x_1 = \Phi(z_2) - \Phi(z_1) = \Phi'(z_1)(z_2 - z_1) + \dots$. Somit liegt es nahe, dass Würfelinhalte lokal um den Faktor $|\det \Phi'(z)|$ verstärkt werden, d.h. dass gilt " $dx = |\det \Phi'(z)| dz$ ". Dies ist tatsächlich der Fall, wie der folgende Satz zeigt, der erst in Analysis III vollständig bewiesen wird.

4.44 Satz (Transformationsatz). Seien $U, V \subset \mathbb{R}^n$ offen und $\Phi: U \rightarrow V$ ein C^1 -Diffeomorphismus. Sei $f: V \rightarrow \mathbb{R}$ messbar. Dann ist f über $V = \Phi(U)$ genau dann integrierbar, wenn $(f \circ \Phi) \cdot |\det \Phi'|: U \rightarrow \mathbb{R}$ integrierbar über U ist, und dann gilt

$$\int_{\Phi(U)} f(x) dx = \int_U f(\Phi(z)) \cdot |\det \Phi'(z)| dz.$$

4.45 Beispiel (Polarkoordinaten im \mathbb{R}^2). Bereits in Beispiel 2.19 hatten wir die Koordinatentransformation von kartesischen Koordinaten auf Polarkoordinaten in \mathbb{R}^2 angegeben:

$$\mathbb{R}^2 \setminus \{(x, 0) \mid x \geq 0\} = \Phi(r, \varphi) = (r \cos \varphi, r \sin \varphi),$$

wobei $\Phi: Q \rightarrow \mathbb{R}^2 \setminus \{(x, 0) \mid x \geq 0\}$ mit $Q := \{(r, \varphi) \mid r > 0, 0 < \varphi < 2\pi\}$. Φ ist bijektiv, und es gilt

$$\det \Phi'(r, \varphi) = \begin{vmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{vmatrix} = r > 0 \text{ in } Q,$$

also ist Φ ein Diffeomorphismus. ($\Phi: \overline{Q} \rightarrow \mathbb{R}^2$ ist dagegen nicht bijektiv, deshalb die Abbildung auf die „geschlitzte Ebene“ $\mathbb{R}^2 \setminus \{(x, 0) \mid x \geq 0\}$). Die unbeschränkte Menge $N := \{(x, 0) \mid x \geq 0\}$ ist eine Nullmenge, so dass wir schreiben können:

$$\int_{B(0,R)} f(x, y) d(x, y) = \int_{B(0,R) \setminus N} f(x, y) d(x, y) = \int_0^R \int_0^{2\pi} f(r \cos \varphi, r \sin \varphi) r \, d\varphi dr,$$

wobei die letzte Gleichung gerade aus der Transformationsformel resultiert.

Für $f = 1$ erhält man wieder die Fläche des Kreises:

$$\lambda(B(0, R)) = \int_{B(0,R)} 1 d(x, y) = \int_0^R \int_0^{2\pi} r \, d\varphi dr = 2\pi \frac{R^2}{2} = \pi R^2.$$

Für $f(x, y) = e^{-(x^2+y^2)}$ erhält man

$$\int_{\mathbb{R}^2} f(x, y) d(x, y) = \int_0^\infty \int_0^{2\pi} e^{-r^2} r \, dr \, d\varphi = \frac{2\pi}{2} (-e^{-r^2}) \Big|_0^\infty = \pi.$$

Wegen $\int_{\mathbb{R}^2} e^{-(x^2+y^2)} d(x, y) = \left(\int_{\mathbb{R}} e^{-t^2} dt \right)^2$ folgt damit $\int_{\mathbb{R}} e^{-t^2} dt = \sqrt{\pi}$ und somit für die Wahrscheinlichkeitsdichte der Standard-Normalverteilung

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-t^2/2} dt = 1.$$

5. Kurvenintegrale und Flächenintegrale, der Satz von Gauß

5.1 Worum geht's? Nachdem die maß- und integrationstheoretischen Grundlagen bereitgestellt wurden, können wir nun Funktionen integrieren, welche auf einer Kurve oder einer Fläche definiert sind. Für das Flächenintegral ist die Gramsche Determinante zentral, welche eine der Fläche angepasste Skalierung des Integranden bewirkt. Einer der wichtigsten Sätze für Flächenintegrale ist der Satz von Gauß, welcher das Analogon zur partiellen Integration in \mathbb{R} darstellt.

a) Kurvenintegrale und 1-Formen

Unser Ziel ist es nun, Wegintegrale zu erklären. Wir erinnern uns an die Definition eines Weges als stetige Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^n$. Diese können noch recht exotisch sein, wenn man an das Beispiel der Peanokurve $\gamma : [0, 1] \rightarrow [0, 1]^2$ (surjektiv) denkt. Für stückweise differenzierbare Wege haben wir deren Länge über

$$L(\gamma) = \int_a^b |\gamma'(t)| dt$$

festgelegt. Zunächst behandeln wir Kurvenintegrale, also Integrale über ein Integrationsgebiet $\Gamma = \gamma([a, b])$ und wollen zu einer Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung: $\int_\alpha^\beta f'(x) dx = f(\beta) - f(\alpha)$ gelangen. Später beweisen wir dann eine entsprechende Aussage über Flächenintegrale.

5.2 Definition. Unter einem Vektorfeld auf $D \subset \mathbb{R}^n$ verstehen wir eine Abbildung

$$v : D \rightarrow \mathbb{R}^n,$$

die jedem Punkt $x \in D$ einen Vektor $v(x) \in \mathbb{R}^n$ zuordnet. Meist denkt man sich den Vektor $v(x)$ im Fußpunkt x angetragen.

Wir führen nun Integrale über Γ ein und denken dabei beispielsweise an die folgende Situation: Es sei $\gamma : [a, b] \rightarrow \Gamma \subset \mathbb{R}^3$ ein glatter Weg, $v : \Gamma \rightarrow \mathbb{R}^3$ ein Vektorfeld. Ist v etwa ein Kraftfeld, so berechnet sich die physikalische Arbeit als „Kraft mal Weg“ und man möchte das Integral

$$A := \int_\Gamma \langle v(x), dx \rangle := \int_a^b \langle v(\gamma(t)), \gamma'(t) \rangle dt, \quad \langle \cdot, \cdot \rangle : \text{Skalarprodukt in } \mathbb{R}^3$$

berechnen, also v „längs Γ “ integrieren.

Einen Kalkül zur Berechnung solcher Integrale liefern die „Pfaffschen Differentialformen“¹⁰. In diesem Zusammenhang sind auch die Namen Henri Poincaré¹¹ und Cartan¹² zu nennen.

5.3 Definition. Sei $\Gamma \subset \mathbb{R}^n$ Bild eines glatten Weges und $\omega : \Gamma \times \mathbb{R}^n \rightarrow \mathbb{R}$ stetig und linear im zweiten Argument. Dann heißt ω eine 1-Form in Γ .

Auch für eine beliebige Teilmenge $U \subset \mathbb{R}^n$ nennt man ein entsprechendes $\omega : U \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine 1-Form in U . Eine Funktion $f : U \rightarrow \mathbb{R}$ heißt auch eine 0-Form in U .

Die 1-Form ω läßt sich explizit beschreiben, wenn wir die zweite Variable h bezüglich der Standardbasis des \mathbb{R}^n schreiben: $h = \sum_{i=1}^n h_i e_i$, e_i i -ter Einheitsvektor des \mathbb{R}^n . Damit ist $\omega(x, h) = \sum_{i=1}^n \omega(x, e_i) h_i$, d.h. mit $\omega_i(x) := \omega(x, e_i)$ gilt:

$$\omega(x, h) = \langle \omega(x), h \rangle, \quad \omega(x) = \begin{pmatrix} \omega_1(x) \\ \vdots \\ \omega_n(x) \end{pmatrix}.$$

Dem klassischen „ dx “ geben wir hiermit folgenden Sinn:

$$\begin{aligned} dx_i : \mathbb{R}^n &\rightarrow \mathbb{R}, \quad h \mapsto h_i = dx_i(h) = \langle h, e_i \rangle \\ \Rightarrow \omega(x, \cdot) &= \sum_{i=1}^n \omega_i(x) dx_i \equiv \langle \omega(x), dx \rangle. \end{aligned}$$

5.4 Definition. Sei γ eine Darstellung von Γ . Dann ist

$$\int_{\Gamma} \omega := \int_a^b \omega(\gamma(t), \gamma'(t)) dt$$

das Kurvenintegral der 1-Form ω längs Γ .

Erinnern wir uns an das obige Beispiel der physikalischen Arbeit, so ist $\omega(x, \cdot) = \langle v(x), dx \rangle$, d.h. $A = \int_{\Gamma} \omega$.

Die Definition des Kurvenintegrals ist von der speziellen Darstellung unabhängig, solange nur dieselbe Orientierung gewählt wird: Denn seien $\gamma_j : [a_j, b_j] \rightarrow \mathbb{R}^n$ ($j = 1, 2$) und $\varphi : [a_2, b_2] \rightarrow [a_1, b_1]$ mit $\varphi' > 0$, so dass $\gamma_2 = \gamma_1 \circ \varphi$. Dann gilt

$$\int_{a_1}^{b_1} \omega(\gamma_1(t), \gamma_1'(t)) dt = \int_{a_2}^{b_2} \omega(\gamma_1 \circ \varphi(\tau), \gamma_1' \circ \varphi(\tau)) \varphi'(\tau) d\tau = \int_{a_2}^{b_2} \omega(\gamma_2, \gamma_2')(\tau) d\tau.$$

¹⁰Johann Friedrich Pfaff, 22.12.1765 – 21.4.1825

¹¹Henri Poincaré, 29.4.1854 – 17.7.1912

¹²Élie Joseph Cartan, 9.4.1869 – 6.5.1951

Weiterhin gelten offensichtlich die folgenden Rechenregeln:

$$\begin{aligned}\int_{\Gamma} \omega_1 + \omega_2 &= \int_{\Gamma} \omega_1 + \int_{\Gamma} \omega_2 \\ \forall c \in \mathbb{R} : \int_{\Gamma} c\omega &= c \int_{\Gamma} \omega \\ \int_{\Gamma_1} \omega + \int_{\Gamma_2} \omega &= \int_{\Gamma} \omega, \quad \text{für } \Gamma = \Gamma_1 + \Gamma_2.\end{aligned}$$

5.5 Beispiele. (i) Es sei das Kraftfeld $v(x) = \frac{x}{|x|^2}$ in \mathbb{R}^2 gegeben und $\Gamma = \Gamma_1 + \Gamma_2 \subset \mathbb{R}^2$ durch

$$\Gamma_1 = \left\{ x = \gamma_1(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, 0 \leq t \leq \frac{\pi}{2} \right\}, \quad \Gamma_2 = \left\{ x = \gamma_2(t) = \begin{pmatrix} 0 \\ t \end{pmatrix}, 1 \leq t \leq 2 \right\}$$

Dann berechnet sich die Arbeit wie folgt:

$$A = \int_{\Gamma} \omega = \int_{\Gamma_1} \omega + \int_{\Gamma_2} \omega, \quad \text{wobei } \omega(x, \cdot) = \langle v(x), dx \rangle.$$

Wegen $\omega(\gamma_1, \gamma_1') = \langle v(\gamma_1(t)), \gamma_1'(t) \rangle = 0$ ist $\int_{\Gamma_1} \omega = 0$

$$\Rightarrow A = \int_{\Gamma_2} \omega = \int_1^2 \langle v(\gamma_2(t)), \gamma_2'(t) \rangle dt = \int_1^2 \left\langle \frac{1}{t^2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle dt = \int_1^2 \frac{1}{t} dt = \ln 2.$$

(ii) Sei $\Gamma := \Gamma_1 + \Gamma_2 + \Gamma_3$ gegeben durch

$$\gamma_1(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, 0 \leq t \leq \pi; \quad \gamma_2(t) = \begin{pmatrix} t-1 \\ -t \end{pmatrix}, 0 \leq t \leq 1; \quad \gamma_3(t) = \begin{pmatrix} t \\ t-1 \end{pmatrix}, 0 \leq t \leq 1 \text{ und}$$

$$v(x) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \cdot \frac{1}{|x|^2}, \quad \text{d.h.: } \omega(x, \cdot) = \langle v(x), dx \rangle = \frac{-x_2}{x_1^2 + x_2^2} dx_1 + \frac{x_1}{x_1^2 + x_2^2} dx_2, \quad \text{für } x \neq 0.$$

$$\begin{aligned}\Rightarrow \int_{\Gamma_1} \omega &= \int_0^{\pi} \left\langle \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \right\rangle dt = \pi \\ \int_{\Gamma_2} \omega &= \int_0^1 \left\langle \begin{pmatrix} t \\ t-1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle \frac{1}{t^2 + (t-1)^2} dt \\ &= \int_0^1 \frac{1}{t^2 + (1-t)^2} dt = \int_{-1/2}^{1/2} \frac{1}{(\tau + \frac{1}{2})^2 + (\tau - \frac{1}{2})^2} d\tau \\ &= 2 \int_{-1/2}^{1/2} \frac{1}{4\tau^2 + 1} d\tau = \int_{-1}^1 \frac{1}{s^2 + 1} ds = \arctan s \Big|_{s=-1}^{s=1} = \frac{\pi}{2}\end{aligned}$$

$$\text{Analog: } \int_{\Gamma_3} \omega = \frac{\pi}{2}$$

$$\Rightarrow \int_{\Gamma} \omega = 2\pi.$$

Wenn man Γ_1 zu einer vollen Kreislinie $\tilde{\Gamma}_1$ verlängert, so ergibt sich: $\int_{\tilde{\Gamma}_1} \omega = 2\pi$.

In den letzten Beispielen haben wir gesehen, dass für das erste $\int_{\text{Kreislinie}} \omega = 0$ gilt, während man für das zweite und den geschlossenen Weg Γ $\int_{\Gamma} \omega = 2\pi$ erhält. Außerdem gilt für das zweite, um vom Punkt $(1, 0)$ zum Punkt $(-1, 0)$ zu kommen im Fall von Γ_1 : $\int_{\Gamma_1} \omega = \pi$, während man für den Weg $\tilde{\Gamma} = \Gamma_2 + \Gamma_3$ in umgekehrter Richtung durchlaufen $\int_{\tilde{\Gamma}} \omega = -\pi$ erhält.

Wir sehen also, dass die „Arbeit“, also das Integral, vom Integrationsweg abhängen kann (aber nicht muss). Es drängen sich nun folgende Fragen auf:

- Wann ist $\int_{\Gamma} \omega$ nur von den Endpunkten von Γ abhängig?
- Wann sind die Integrale über geschlossene Kurven Null?

5.6 Definition-Bemerkung. (i) Eine Funktion wird auch Nullform genannt. Sei $U \subset \mathbb{R}^n$ offen, $f \in C^1(U; \mathbb{R})$. Dann ist $df: U \times \mathbb{R}^n \rightarrow \mathbb{R}$, $(x, h) \mapsto f'(x, h) = \langle \nabla f(x), h \rangle$ eine 1-Form in U . Sie heißt das totale Differential von f . Es gilt

$$\begin{aligned} \int_{\Gamma} df &= \int_a^b f'(\gamma(t), \gamma'(t)) dt = \int_a^b (f \circ \gamma)'(t) dt = f(\gamma(b)) - f(\gamma(a)) \\ &= f(B) - f(A) \end{aligned}$$

mit $B := \gamma(b)$ und $A := \gamma(a)$, d.h. das Integral hängt nur von den Endpunkten ab.

(ii) Existiert zur 1-Form $\omega: U \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine 0-Form $f: U \rightarrow \mathbb{R}$ mit $\omega(x, h) = f'(x, h) = (df)(x, h)$, so heißt ω exakt und f Stammfunktion zu ω .

Ist $\omega(x, h) = \langle \omega(x), h \rangle$, so ist also für exaktes $\omega: \omega(x) = \nabla f$.

(iii) Eine 1-Form $\omega: U \times \mathbb{R}^n \rightarrow \mathbb{R}$ heißt wegunabhängig, falls $\int_{\Gamma_1} \omega = \int_{\Gamma_2} \omega$ für alle in U verlaufenden Wege Γ_1, Γ_2 mit denselben Anfangs- und Endpunkten gilt.

Eine Antwort auf die obigen Fragen und eine Verbindung zur vorangegangenen Definition liefert der

5.7 Satz. Seien $U \subset \mathbb{R}^n$ offen und zusammenhängend und $\omega: U \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine 1-Form. Dann gilt: ω ist genau dann exakt, wenn ω wegunabhängig ist.

Beweis. „ \Rightarrow “: Ist bereits in Bemerkung 5.6 (i) bewiesen worden.

„ \Leftarrow “: Sei $x_0 \in U$ fest gewählt und $\Gamma(x_0, x)$ verbinde x_0 mit $x \in U$, dann ist

$$f(x) := \int_{\Gamma(x_0, x)} \omega$$

wohldefiniert, da das Integral nach Voraussetzung wegunabhängig ist. Das so definierte f ist die gesuchte Funktion, denn $f \in C^1(U, \mathbb{R})$ mit folgendem Argument: Verbindet man x mit $x+h$ geradlinig (für kleine h) und $\Gamma_{x,x+h}$ sei die Verbindungsstrecke, dann ist

$$\begin{aligned} f(x+h) - f(x) &= \int_{\Gamma_{x,x+h}} \omega = \int_0^1 \omega(x+th, h) dt \\ &= \omega(x+\tau h, h) \quad \text{für ein } \tau \in (0, 1), \tau = \tau(x, h) \\ &= \omega(x, h) + \omega(x+\tau h, h) - \omega(x, h) \\ &= \omega(x, h) + r(x, h)|h| \\ \text{mit } r(x, h) &:= \begin{cases} \frac{\omega(x+\tau h, h) - \omega(x, h)}{|h|} & , h \neq 0 \\ 0 & , h = 0 \end{cases} \\ &= \omega\left(x+\tau h, \frac{h}{|h|}\right) - \omega\left(x, \frac{h}{|h|}\right) \quad \text{für } h \neq 0, \end{aligned}$$

d.h., da ω stetig: $|r(x, h)| < \varepsilon$, falls $|\tau h| < \delta(\varepsilon, x)$. Somit ist f an der Stelle x differenzierbar mit $f'(x) = \omega(x, \cdot)$. \square

5.8 Beispiele. Wenn wir die Beispiele 5.5 betrachten, so gilt (i) $\omega(x) = \frac{x}{|x|^2} = \nabla \ln|x|$ in $U = \mathbb{R}^2 \setminus \{0\}$, d.h. ω ist exakt.

(ii) $\omega(x) = \frac{1}{|x|^2} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$ ist nicht exakt, da für den geschlossenen Weg Γ gilt $\int_{\Gamma} \omega = 2\pi \neq 0$.

Als *notwendige Bedingung* für die Exaktheit haben wir wegen

$$\omega(x, h) = f'(x, h) \Rightarrow \sum_{i=1}^n \omega_i(x) h_i = \sum_{i=1}^n \partial_i f(x) h_i \Rightarrow \omega_i = \partial_i f$$

für $\omega \in \mathcal{C}^1$:

$$\partial_j \omega_i = \partial_j \partial_i f = \partial_i \partial_j f = \partial_i \omega_j.$$

Die Bedingung ist jedoch *nicht hinreichend*, wie Beispiel 5.5 (ii) zeigt:

$$\partial_2 \omega_1 = \frac{-1 \cdot |x|^2 + 2x_2^2}{|x|^4} = \frac{x_2^2 - x_1^2}{|x|^4}, \quad \text{sowie} \quad \partial_1 \omega_2 = \frac{|x|^2 - 2x_1^2}{|x|^4} = \frac{x_2^2 - x_1^2}{|x|^4}.$$

Eine hinreichende Bedingung erhält man unter zusätzlichen Voraussetzungen an U :

5.9 Definition. Eine Teilmenge $U \subset \mathbb{R}^n$ heißt *sternförmig* bezüglich eines Punktes $p \in U$: \Leftrightarrow für alle $x \in U$ ist die Verbindungsstrecke \overline{px} in U enthalten.

5.10 Satz. Sei $U \subset \mathbb{R}^n$ offen und sternförmig und ω eine C^1 -1-Form mit $\partial_j \omega_i = \partial_i \omega_j$, $i, j = 1, \dots, n$. Dann ist ω exakt.

Beweis. O.B.d.A. sei U sternförmig bezüglich des Nullpunktes. Durch geradliniges Integrieren längs der Verbindungsstrecke von 0 nach x geben wir eine Stammfunktion an: $\Gamma_{0,x} = \{\gamma(t) = tx \mid 0 \leq t \leq 1\}$:

$$f(x) := \int_{\Gamma_{0,x}} \omega = \int_0^1 \omega(tx, x) dt = \int_0^1 \sum_{i=1}^n \omega_i(tx) x_i dt$$

Da nach Voraussetzung ω_i stetig differenzierbar ist, folgt

$$\partial_j f(x) = \int_0^1 \omega_j(tx) dt + \int_0^1 \sum_{i=1}^n x_i t \partial_j \omega_i(tx) dt = \int_0^1 \omega_j(tx) dt + \int_0^1 \sum_{i=1}^n x_i t \partial_i \omega_j(tx) dt.$$

Mit partieller Integration ist $\int_0^1 t \sum_{i=1}^n x_i \partial_i \omega_j(tx) dt = t \omega_j(tx) \Big|_0^1 - \int_0^1 \omega_j(tx) dt = \omega_j(x) - \int_0^1 \omega_j(tx) dt$

$$\Rightarrow \partial_j f(x) = \omega_j(x).$$

□

5.11 Beispiele. Wiederum betrachten wir die Beispiele 5.5:

(i) Dieses Beispiel zeigt, dass in $U = \mathbb{R}^2 \setminus \{0\}$ nicht beide Bedingungen nötig sind, denn U ist nicht sternförmig, doch es gilt $\omega(x) = \nabla \ln |x|$ dort.

(ii) Dieses Beispiel wollen wir nun auf der sternförmigen Menge $U := \{x \mid x_1 > 0\}$ betrachten und eine Stammfunktion von

$$\omega(x) = \frac{1}{|x|^2} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

berechnen. Wir wählen $p = (1, 0)$ und nach $x = (x_1, x_2)$ den Weg $\Gamma = \Gamma_1 + \Gamma_2$ mit

$$\Gamma_1 = \{\gamma_1(t) = (t, 0) \mid 1 \leq t \leq x_1\} \text{ und } \Gamma_2 = \{\gamma_2(t) = (x_1, t) \mid 0 \leq t \leq x_2\},$$

welcher eine einfachere Berechnung des Integrals über ω zuläßt als die direkte Verbindung \overline{px} . Diese freie Wahl des Weges ist jetzt wegen der Wegunabhängigkeit (nach Satz 5.7) möglich.

$$\begin{aligned} \Rightarrow f(x) &= \int_{\Gamma} \omega = \int_1^{x_1} \frac{1}{t^2} \left\langle \begin{pmatrix} 0 \\ t \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle dt + \int_0^{x_2} \frac{1}{x_1^2 + t^2} \left\langle \begin{pmatrix} -t \\ x_1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle dt \\ &= \int_0^{x_2} \frac{x_1}{x_1^2 + t^2} dt = \int_0^{x_2/x_1} \frac{1}{1 + \tau^2} d\tau = \arctan \frac{x_2}{x_1}. \end{aligned}$$

Man darf aber nicht beliebig um den Nullpunkt herum integrieren, weil man sonst womöglich einen einmal gewählten Ast der arctan-Funktion verlassen wird.

Im folgenden werden wir analog zur Bogenlänge von Kurven ein Flächenmaß einführen. Sei also $\gamma : T \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine Parameterdarstellung einer m -dimensionalen Fläche, $\Gamma = \gamma(U)$ mit $U \subset \mathbb{R}^m$ offen und $\gamma'(u, \cdot)$ habe Rang p .

Was soll nun das Maß $A(\Gamma)$ von Γ sein? Betrachte dazu ein m -dimensionales Parallelepiped (Spat) $P := \{x = \sum_{j=1}^m \lambda_j a_j : \lambda_j \in [0, 1]\}$ mit Kantenvektoren $a_1, \dots, a_m \in \mathbb{R}^n$. Es gilt $P = S([0, 1]^m)$ mit der Matrix $S := (a_1, \dots, a_m)$.

Im Falle $n = m$ erhält man (vgl. Satz 4.43) $A(P) = \lambda_m(S([0, 1]^m)) = |\det S|$, d.h. $A(P)^2 = (\det S)^2 = (\det S^\top) \det S = \det(S^\top S) = \det G(a_1, \dots, a_m)$

$$G(a_1, \dots, a_m) := \left(\langle a_i, a_j \rangle \right)_{i,j=1,\dots,m} = \begin{pmatrix} \langle a_1, a_1 \rangle & \dots & \langle a_1, a_m \rangle \\ \vdots & & \vdots \\ \langle a_m, a_1 \rangle & \dots & \langle a_m, a_m \rangle \end{pmatrix}.$$

Diese Formel ist aber auch für $m < n$ sinnvoll und zeigt, dass das m -dimensionale Volumen eines Würfels unter der Abbildung S mit dem Faktor $\sqrt{G(a_1, \dots, a_m)}$ multipliziert wird.

Sei nun $\gamma : U \rightarrow \mathbb{R}^n$ eine Parametrisierung einer m -dimensionalen Fläche Γ . Lokal an der Stelle $x = \gamma(u)$ wird die Fläche (approximativ) durch die Vektoren $\gamma(u + he_i) - \gamma(u)$, $i = 1, \dots, m$, aufgespannt. Im Limes erhält man die partiellen Ableitungen $\partial_i \gamma(u)$. Mit $a_i := \partial_i \gamma(u) \in \mathbb{R}^n$ gilt $\gamma'(u) = (a_1, \dots, a_m) \in \mathbb{R}^{n \times m}$. Es ist daher plausibel, dass das m -dimensionale Volumen eines Würfels lokal mit dem Faktor $\sqrt{(\langle a_i, a_j \rangle)_{i,j=1,\dots,m}} = \sqrt{\det(\gamma'(u)^\top \gamma'(u))}$ verstärkt wird.

5.12 Definition. Sei $U \subset \mathbb{R}^n$ offen und $\gamma : U \rightarrow \mathbb{R}^n$ eine Parametrisierung einer m -dimensionalen Fläche Γ .

(i) Dann heißt

$$g : U \rightarrow \mathbb{R}^{m \times m}, u \mapsto g(u) = \gamma'(u)^\top \gamma'(u) = (\langle \partial_i \gamma(u), \partial_j \gamma(u) \rangle)_{i,j=1,\dots,m}$$

die ‘Gramsche Matrix’¹³ (oder Maßtensor) zu γ und $\det g(u)$ die ‘Gramsche Determinante’ zu γ .

(ii) Der Flächeninhalt $A(\Gamma)$ (‘Areal’) von Γ ist definiert durch

$$A(\Gamma) := \int_U \sqrt{\det g(u)} du = \int_U \sqrt{\det(\gamma'(u)^\top \gamma'(u))} du.$$

5.13 Lemma. (i) Der Inhalt einer Fläche ist invariant gegenüber Bewegungen im \mathbb{R}^n .

(ii) Der Inhalt einer Fläche ist unabhängig von der Parametrisierung der Fläche.

¹³Jorgen Pedersen Gram, 27.6.1850 – 29.4.1916

Beweis. (i) Sei $T(x) = Sx + c$ eine Bewegung. Dann ist $\tilde{\gamma} := T \circ \gamma$ eine Parametrisierung der Fläche $\tilde{\Gamma} := T(\Gamma)$. Es folgt $\tilde{\gamma}'(u) = S\gamma'(u)$ und damit $\tilde{\gamma}'(u)^\top \tilde{\gamma}'(u) = \gamma'(u)^\top S^\top S \gamma'(u) = \gamma'(u)^\top \gamma'(u)$, da S orthogonal ist.

(ii) Seien $\gamma_j: U_j \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$, $j = 1, 2$ äquivalente Parametrisierungen von Γ , also $\gamma_2 = \gamma_1 \circ \Phi$ mit $\Phi: U_2 \rightarrow U_1$ diffeomorph und $\det \Phi' > 0$. Dann folgt $\gamma_2'(u) = \gamma_1'(\Phi(u))\Phi'(u)$ und damit

$$g_{\gamma_2}(u) = \gamma_2'(u)^\top \gamma_2'(u) = \Phi'(u)^\top \gamma_1'(\Phi(u)) \gamma_1'(\Phi(u)) \Phi'(u),$$

also

$$\det g_{\gamma_2}(u) = |\det \Phi'(u)|^2 \det g_{\gamma_1}(u).$$

Somit erhält man für die zugehörigen Flächeninhalte nach dem Transformationssatz

$$\int_{U_2} \sqrt{\det g_{\gamma_2}(u)} du = \int_{U_2} \sqrt{\det g_{\gamma_1}(u)} |\det \Phi'(u)| du = \int_{U_1} \sqrt{\det g_{\gamma_1}(u)} du.$$

□

Wir hatten für das Flächenmaß $A(\Gamma) = \int_U \sqrt{\det g(u)} du$ mit $g_{ij}(u) = \langle \partial_i \gamma(u), \partial_j \gamma(u) \rangle$ hergeleitet, wobei $\gamma: U \rightarrow \Gamma$ eine Parametrisierung der m -Fläche Γ war. Es gilt somit “ $dA = \sqrt{\det g(u)} du$ ”. Dies ist die Grundlage des Integrals über m -dimensionale Flächen.

5.14 Definition. Sei $\gamma: U \rightarrow \mathbb{R}^n$ eine Parametrisierung einer m -Fläche Γ mit Gramscher Matrix $g(u)$ und $f: \Gamma \rightarrow \mathbb{R}$ eine Funktion. Dann heißt f integrierbar über Γ , falls $u \mapsto f(\gamma(u))\sqrt{\det g(u)}$ integrierbar über U ist. In diesem Fall setzt man

$$\int_{\Gamma} f dA := \int_{\Gamma} f(x) dA(x) := \int_U f(\gamma(u)) \sqrt{\det g(u)} du.$$

Andere Schreibweise ist $\int_{\Gamma} f dS$ (“surface”). Im Fall $m = 1$ schreibt man auch $\int_{\Gamma} f ds$ (s : Bogenlänge).

Eine Teilmenge $M \subset \Gamma$ heißt integrierbar, falls χ_M integrierbar ist, und in diesem Fall heißt $A(M) := \int_{\Gamma} \chi_M dA$ der m -dimensionale Flächeninhalt von $M \subset \mathbb{R}^n$.

5.15 Bemerkung. (i) Wie in Lemma 5.13 (ii) sieht man mit dem Transformationssatz, dass das Integral wohldefiniert, d.h. unabhängig von der Parametrisierung ist.

(ii) Für $f = 1$ erhält man wieder $A(\Gamma) = \int_{\Gamma} 1 dA$. Für $m = 1$ ist $g(u) = \langle \gamma'(u), \gamma'(u) \rangle$ für $u \in U = (a, b)$ und damit

$$A(\Gamma) = \int_{\Gamma} 1 dA = \int_a^b |\gamma'(u)| du$$

die Bogenlänge der Kurve Γ . In diesem Fall ist

$$\int_{\Gamma} f(x) ds(x) = \int_a^b f(\gamma(u)) |\gamma'(u)| du.$$

(iii) Als Beziehung zum Kurvenintegral für 1-Formen ergibt sich für $n = 2$, $m = 1$, $\omega(x, h) = \langle v(x), h \rangle$ mit $v(x) = \begin{pmatrix} v_1(x) \\ v_2(x) \end{pmatrix}$:

$$\int_{\Gamma} \omega = \int_a^b \omega(\gamma(t), \gamma'(t)) dt = \int_a^b \langle v(\gamma(t)), \vec{t}(\gamma(t)) \rangle |\gamma'(t)| dt = \int_{\Gamma} \langle v, \vec{t} \rangle ds,$$

wobei $\vec{t}(\gamma(t)) = \frac{\gamma'(t)}{|\gamma'(t)|}$ den Tangenteneinheitsvektor bezeichnet.

(iv) Für $m = n$ erhält man

$$\begin{aligned} A(\Gamma) &= \int_{\Gamma} 1 dA = \int_U 1 \cdot \sqrt{\det(\gamma'(u)^\top \gamma'(u))} du \\ &= \int_U 1 \cdot |\det \gamma'(u)| du = \int_{\Gamma} 1 dx = \lambda_n(\Gamma), \end{aligned}$$

der n -dimensionale Flächeninhalt im \mathbb{R}^n ist also wieder das Lebesgue-Maß (Volumen).

5.16 Beispiel (Rotationsflächen). Sei $I \subset \mathbb{R}$ ein offenes Intervall, $f \in C^1(I; \mathbb{R})$ mit $f(t) \geq 0$, $t \in I$. Setze $\Gamma := \{(x, y, z)^\top \in \mathbb{R}^3 : z \in I, x^2 + y^2 = f(z)^2\}$. Dann ist

$$\gamma: I \times (0, 2\pi) \rightarrow \mathbb{R}^3, t \mapsto \begin{pmatrix} f(t) \cos \varphi \\ f(t) \sin \varphi \\ t \end{pmatrix}$$

bis auf eine Nullmenge eine Parametrisierung von Γ . Es gilt

$$\gamma'(t, \varphi) = \begin{pmatrix} f'(t) \cos \varphi & -f(t) \sin \varphi \\ f'(t) \sin \varphi & f(t) \cos \varphi \\ 1 & 0 \end{pmatrix}$$

und damit

$$\gamma'(t)^\top \gamma'(t) = \begin{pmatrix} f'(t)^2 + 1 & 0 \\ 0 & f(t)^2 \end{pmatrix}.$$

Also ist $\sqrt{\det g(t, \varphi)} = f(t) \sqrt{f'(t)^2 + 1}$ und damit

$$\begin{aligned} A(\Gamma) &= \int_{\Gamma} 1 dA = \int_{I \times (0, 2\pi)} \sqrt{\det g(t, \varphi)} d(t, \varphi) = \int_0^{2\pi} \int_I f(t) \sqrt{f'(t)^2 + 1} dt d\varphi \\ &= 2\pi \int_I f(t) \sqrt{f'(t)^2 + 1} dt. \end{aligned}$$

Speziell sei $f(t) = r$ und $I = (0, L)$. Dann ist Γ die Oberfläche eines Zylinders der Länge L mit Radius r , und man erhält $A(\Gamma) = 2\pi \int_0^L r dt = 2\pi Lr$.

5.17 Beispiel (Oberfläche eines Graphen). Sei $U \subset \mathbb{R}^n$ offen, $F \in C^1(U; \mathbb{R})$. Dann ist $\Gamma := \text{graph } F \subset \mathbb{R}^n$ eine $(n-1)$ -dimensionale Fläche mit Parametrisierung $\gamma: U \rightarrow \mathbb{R}^n, u \mapsto (u, F(u))^\top$. Man kann nachrechnen, dass $\det g(u) = \det(\gamma'(u)^\top \gamma'(u)) = 1 + |\nabla F(u)|^2$ gilt. Somit ist für integrierbares $f: \Gamma \rightarrow \mathbb{R}$

$$\int_{\Gamma} f dA = \int_U f(u, F(u)) \sqrt{1 + |\nabla F(u)|^2} du.$$

Insbesondere erhält man $A(\text{graph } F) = \int_U \sqrt{1 + |\nabla F(u)|^2} du$.

Sei nun speziell $U := \{u \in \mathbb{R}^{n-1} : |u| < r\}$ und $F(u) := \sqrt{r^2 - |u|^2}$ für $r > 0$. Dann ist $\Gamma := \text{graph } F = \{x \in \mathbb{R}^n : |x| = r, x_n > 0\}$ die Nordhalbkugel mit Radius r . Wegen $\nabla F(u) = -\frac{u}{\sqrt{r^2 - |u|^2}}$ erhält man $\det g(u) = 1 + |\nabla F(u)|^2 = 1 + \frac{|u|^2}{r^2 - |u|^2} = \frac{r^2}{r^2 - |u|^2}$. Somit ist

$$\int_{\Gamma} f dA = \int_{|u| < r} f(u, \sqrt{r^2 - |u|^2}) \frac{r}{\sqrt{r^2 - |u|^2}} du = r^{n-1} \int_{|v| < 1} \frac{f(rv, r\sqrt{1 - |v|^2})}{\sqrt{1 - |v|^2}} dv,$$

wobei die Substitution $u = rv, du = r^{n-1} dv$ verwendet wurde.

5.18 Satz. Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ integrierbar. Dann gilt

$$\int_{\mathbb{R}^n} f(x) dx = \int_0^\infty \left(\int_{|x|=r} f(x) dA(x) \right) dr.$$

Beweis. Sei $U := \{u \in \mathbb{R}^{n-1} : |u| < 1\}$ und $\Phi: U \times (0, \infty) \rightarrow \mathbb{R}^n, (u, r)^\top \mapsto (ru, r\sqrt{1 - |u|^2})^\top$. Dann ist $\Phi: U \times (0, \infty) \rightarrow \mathbb{R}^n \setminus \{0\}$ ein Diffeomorphismus, und nach dem Transformationssatz folgt

$$\begin{aligned} \int_{\mathbb{R}^n} f(x) dx &= \int_{\mathbb{R}^n \setminus \{0\}} f(x) dx = \int_{U \times (0, \infty)} f(\Phi(u, r)) |\det \Phi'(u, r)| d(u, r) \\ &= \int_0^\infty \int_U f(ru, r\sqrt{1 - |u|^2}) \frac{r^{n-1}}{\sqrt{1 - |u|^2}} du dr = \int_0^\infty \int_{|u|=r} f(u) dA(u) dr. \end{aligned}$$

□

5.19 Korollar (rotationssymmetrische Funktionen). (i) Sei $f: (0, \infty) \rightarrow \mathbb{R}$ mit $x \mapsto f(|x|)$ integrierbar. Dann gilt

$$\int_{\mathbb{R}^n} f(|x|) dx = A(S^{n-1}) \int_0^\infty f(r) r^{n-1} dr.$$

(ii) Sei $B_n := \{x \in \mathbb{R}^n : |x| \leq 1\}$ und $S^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$. Dann gilt $A(S^{n-1}) = n\lambda_n(B_n)$. Speziell ist $A(S^1) = 2\pi, A(S^2) = 4\pi$.

Beweis. (i) Nach Satz 5.18 ist

$$\begin{aligned}\int_{\mathbb{R}^n} f(|x|) dx &= \int_0^\infty \int_{|x|=r} f(|x|) dA(x) dr = \int_0^\infty f(r) \int_{|x|=r} 1 dA(x) dr \\ &= A(S^{n-1}) \int_0^\infty f(r) r^{n-1} dr.\end{aligned}$$

(ii) folgt aus (i) mit $f = \chi_{(0,1)}$ wegen $\int_0^1 r^{n-1} dr = \frac{1}{n}$. \square

b) Die Integralsätze von Gauß und Stokes

In \mathbb{R} lautet der Hauptsatz der Differential- und Integralrechnung $\int_a^b f' = f(b) - f(a)$ für stetig differenzierbares f . Entsprechendes haben wir für Kurvenintegrale und 1-Formen in Bemerkung 5.6 bewiesen:

$$\int_{\Gamma} df = f(B) - f(A),$$

wobei A, B die Randpunkte von Γ sind.

In diesem Abschnitt werden wir Analoga für glatte bzw. stückweise glatte m -Flächen, $m \geq 1$ beliebig, erhalten.

Wir benötigen dafür noch einige geometrische Begriffe.

5.20 Definition. Sei $\Gamma \subset \mathbb{R}^n$ eine m -Fläche mit Parametrisierung $\gamma: U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$. Sei $u \in U$ und $a := \gamma(u)$. Der von den Vektoren $\partial_1 \gamma(u), \dots, \partial_m \gamma(u)$ aufgespannte (m -dimensionale) Vektorraum $T_a \Gamma := \{\gamma'(u)h : h \in \mathbb{R}^m\} \subset \mathbb{R}^n$ heißt der “Tangentenraum” an Γ im Punkt a . Die Elemente $v \in T_a \Gamma$ heißen “Tangentialvektoren” an Γ im Punkt a .

5.21 Definition. a) Sei $G \subset \mathbb{R}^n$. Dann heißt $\partial G := \overline{G} \cap \overline{\mathbb{R}^n \setminus G}$ der Rand von G .

b) Eine Teilmenge $G \subset \mathbb{R}^n$ heißt ein Gebiet, falls G offen und zusammenhängend ist.

c) Sei $G \subset \mathbb{R}^n$ beschränktes Gebiet (d.h. G ist offen, zusammenhängend und beschränkt). Dann besitzt G einen “glatten Rand”, falls zu jedem $a \in \partial G$ eine offene Umgebung $V \subset \mathbb{R}^n$ und ein $\psi \in C^1(V; \mathbb{R})$ gibt mit $\nabla \psi(v) \neq 0 (v \in V)$ und $G \cap V = \{x \in V : \psi(x) < 0\}$.

5.22 Bemerkung. (i) Für jede Teilmenge $G \subset \mathbb{R}^n$ ist der Rand von G beschrieben durch

$$\partial G = \{x \in \mathbb{R}^n : \forall \varepsilon > 0 : B(x, \varepsilon) \cap G \neq \emptyset, B(x, \varepsilon) \cap (\mathbb{R}^n \setminus G) \neq \emptyset\}.$$

(ii) In der Situation von Definition 5.21 c) gilt $\partial G \cap V = \{x \in V : \psi(x) = 0\}$: Sei $x \in \partial G \cap V$. Falls $\psi(x) < 0$, dann existiert eine Umgebung $W \subset V$ mit $\psi < 0$ in W . Insbesondere folgt $W \subset G$ und damit ist x kein Randpunkt von G . (Analog für den Fall $\psi(x) > 0$.)

Sei andererseits $x \in V$ mit $\psi(x) = 0$. Für $v := \nabla\psi(x) \neq 0$ gilt

$$\psi(x+h) = \psi(x) + \langle \nabla\psi(x), h \rangle + o(|h|) = \langle v, h \rangle + o(|h|), \quad (h \in \mathbb{R}^n, h \rightarrow 0).$$

Für $h := tv$, $t \in \mathbb{R}$, erhält man $\psi(x+tv) = t|v|^2 + o(t|v|)$, $t \rightarrow 0$. Also existiert ein $\varepsilon > 0$ mit $\psi(x-tv) < 0 < \psi(x+tv)$ für $0 < t < \varepsilon$. Somit enthält jede Umgebung von x Punkte in G und Punkte in $\mathbb{R}^n \setminus G$, d.h. $x \in \partial G$.

(iii) Sei G beschränktes Gebiet mit glattem Rand, und $a \in \partial G$. Nach dem Satz über implizite Funktionen existiert nach eventueller Umnummerierung der Koordinaten eine offene Menge $U \subset \mathbb{R}^{n-1}$, ein Intervall $(\alpha, \beta) \subset \mathbb{R}$ und $g \in C^1(U; \mathbb{R})$ mit

$$G \cap (U \times I) = \{x = (x', x_n) \in U \times I : x_n < g(x')\}$$

und

$$\partial G \cap (U \times I) = \{x = (x', x_n) \in U \times I : x_n = g(x')\}.$$

D.h. man kann in Definition 5.21 $\psi(x) = x_n - g(x')$ wählen. Insbesondere ist durch

$$\gamma: U \rightarrow \mathbb{R}^n, u \mapsto \begin{pmatrix} u \\ g(u) \end{pmatrix}$$

eine Parametrisierung von $\partial G \cap (U \times I)$ gegeben, d.h. ∂G ist (lokal) eine $n-1$ -dimensionale Fläche.

5.23 Definition-Bemerkung. Sei G beschränktes Gebiet mit glattem Rand, und $a \in \partial G$. Dann heißt

$$N_a(\partial G) := (T_a(\partial G))^\perp := \{x \in \mathbb{R}^n : \forall v \in T_a(\partial G) : \langle x, v \rangle = 0\}$$

der ‘‘Raum aller Normalenvektoren’’. Es gilt $\dim T_a(\partial G) = n-1$ (siehe Bemerkung 5.22 (ii)) und damit $\dim N_a(\partial G) = 1$. Damit existiert genau ein Vektor $\nu(a) \in N_a(\partial G)$ mit $|\nu(a)| = 1$ und der Eigenschaft $a + t\nu(a) \notin G$ für hinreichend kleine t . Der Vektor $\nu(a)$ heißt ‘‘äußerer Normalenvektor’’ von ∂G im Punkt a . Andere Schreibweisen $n(a), \vec{n}(a)$.

Falls G lokal durch $\psi(x) < 0$ dargestellt wird, so ist $\nu(a) = \frac{\nabla\psi(a)}{|\nabla\psi(a)|}$ (denn der Gradient ist orthogonal zur Höhenlinie und zeigt nach außen). Wird speziell ∂G lokal durch $x_n = g(x')$, d.h. mit $\psi(x) = x_n - g(x')$ dargestellt, so ist

$$\nu(a) = \frac{\nabla\psi(a)}{|\nabla\psi(a)|} = \frac{1}{\sqrt{1 + |\nabla g(a')|^2}} \begin{pmatrix} -\nabla g(a') \\ 1 \end{pmatrix} \quad (a \in \partial A, a' := (a_1, \dots, a_{n-1})).$$

Damit ist $\nu: \partial A \rightarrow \mathbb{R}^n$, $a \mapsto \nu(a)$, stetig.

Im folgenden sei $G \subset \mathbb{R}^n$ ein beschränktes Gebiet mit glattem Rand. Der Satz von Gauß besagt, dass für glatte Vektorfelder $V: \overline{G} \rightarrow \mathbb{R}^n$ die Gleichheit $\int_G \operatorname{div} V(x) dx = \int_{\partial G} \langle V(x), \nu(x) \rangle dA(x)$ gilt. Beachte dabei $\operatorname{div} V = \partial_1 V_1 + \dots + \partial_n V_n$. Wir betrachten zunächst V mit $\operatorname{supp} V \subset G$. Dann ist $V = 0$ auf ∂G und das rechte Integral ist gleich 0. Wir zeigen, dass $\int \partial_j V_j(x) dx = 0$ für alle $j = 1, \dots, n$ gilt:

5.24 Lemma. Sei $h \in C^1(G; \mathbb{R})$ mit $\operatorname{supp} h \subset G$. Dann gilt $\int_G \partial_j h(x) dx = 0$ für $j = 1, \dots, n$.

Beweis. O.E. sei $j = 1$. Setze h durch 0 zu $\tilde{h} \in C^1(\mathbb{R}^n; \mathbb{R})$ fort. Dann ist $\int_G \partial_j h(x) dx = \int_{\mathbb{R}^n} \partial_j h(x) dx$. Wähle $R > 0$ mit $\operatorname{supp} h \subset [-R, R]^n$. Für festes $(x_2, \dots, x_n) \in \mathbb{R}^{n-1}$ gilt nach dem Hauptsatz der Differential- und Integralrechnung

$$\int_{\mathbb{R}} \partial_1 h(x) dx_1 = \int_{-R}^R \partial_1 h(x) dx_1 = \tilde{h}(x_1, \dots, x_n) \Big|_{x_1=-R}^R = 0.$$

Integration bzgl. (x_2, \dots, x_n) liefert die Behauptung. \square

In einem zweiten Schritt betrachten wir ein Vektorfeld $V: \overline{G} \rightarrow \mathbb{R}^n$, dass nur in der Nähe eines Punktes $a \in \partial G$ nicht verschwindet. Lokal kann man den Rand durch $x_n = g(x')$ darstellen (siehe Bem. 5.23). Die Behauptung wird wieder für jede Komponente von V formuliert:

5.25 Satz. Sei $U \subset \mathbb{R}^{n-1}$ offen, zusammenhängend und beschränkt, $I = (\alpha, \beta) \subset \mathbb{R}$, $g \in C^1(U; \mathbb{R})$ mit $g(U) \subset I$. Sei $W := \{x = (x', x_n) \in U \times I : x_n < g(x')\}$ und $M := \{x \in U \times I : x_n = g(x')\}$. Sei weiter $\nu(x) := \frac{1}{\sqrt{1+|\nabla g(x')|^2}} \begin{pmatrix} -\nabla g(x') \\ 1 \end{pmatrix}$ der äußere Normalenvektor. Für $f \in C^1(U \times I; \mathbb{R})$ mit $\operatorname{supp} f \subset U \times I$ gilt

$$\int_W \partial_j f(x) dx = \int_M f(x) \nu_j(x) dA(x) \quad (j = 1, \dots, n).$$

Beweis. Wegen $M = \operatorname{graph} g$ gilt nach Beispiel 5.17 für integrierbares $h: M \rightarrow \mathbb{R}$

$$\int_M h dA = \int_U h(u, g(u)) \sqrt{1 + |\nabla g(u)|^2} du.$$

(i) Sei $j \in \{1, \dots, n-1\}$. Definiere $F: U \times I \rightarrow \mathbb{R}$, $F(x', z) := \int_{\alpha}^z f(x', x_n) dx_n$. Dann ist $\frac{\partial}{\partial z} F(x', z) = f(x', z)$ und $\partial_j F(x', z) = \int_{\alpha}^z \partial_j f(x', z) dx_n$. Setzt man $h(x') := F(x', g(x'))$, so ist $\operatorname{supp} h \subset U$ kompakt, und nach der Kettenregel folgt

$$\partial_j h(x') = \partial_j F(x', g(x')) = \int_{\alpha}^{g(x')} \partial_j f(x', x_n) dx_n + f(x', g(x')) \partial_j g(x').$$

Wende Lemma 5.24 (mit $n-1$ statt n und U statt G) an und erhalte $\int_U \partial_j h(x') dx' = 0$. Damit

$$\begin{aligned} \int_W \partial_j f(x) dx &= \int_U \int_{\alpha}^{g(x')} \partial_j f(x', x_n) dx_n dx' \\ &= - \int_U f(x', g(x')) \partial_j g(x') dx' \\ &= \int_U f(x', g(x')) \nu_j(x', g(x')) \sqrt{1 + |\nabla g(x')|^2} dx' \\ &= \int_M f(x) \nu_j(x) dA(x). \end{aligned}$$

(ii) Sei nun $j = n$. Für festes $x' \in U$ hat $x_n \mapsto f(x', x_n)$ kompakten Träger in $I = (\alpha, \beta)$. Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$\int_{\alpha}^{g(x')} \partial_n f(x', x_n) dx_n = f(x', g(x')).$$

Damit

$$\begin{aligned} \int_W \partial_n f(x) dx &= \int_U \int_{\alpha}^{g(x')} \partial_n f(x', x_n) dx_n dx' \\ &= \int_U f(x', g(x')) = \int_U f(x', g(x')) \nu_n(x', g(x')) \sqrt{1 + |\nabla g(x')|^2} dx' \\ &= \int_M f(x) \nu_n(x) dA(x). \end{aligned}$$

□

5.26 Satz (Gaußscher Integralsatz). Sei $G \subset \mathbb{R}^n$ beschränktes Gebiet mit glattem Rand ∂G , und $\nu: \partial G \rightarrow \mathbb{R}^n$ die äußere Normale. Sei $V: \overline{G} \rightarrow \mathbb{R}^n$ ein glattes Vektorfeld. Dann gilt

$$\int_G \operatorname{div} V(x) dx = \int_{\partial G} \langle V(x), \nu(x) \rangle dA(x).$$

Beweisskizze. Wir haben die “lokalen Versionen” bereits bewiesen: Falls $\operatorname{supp} V \subset G$, so folgt die Behauptung aus Lemma 5.24, falls $\operatorname{supp} V$ hinreichend klein mit $\operatorname{supp} V \cap \partial G \neq \emptyset$ ist, so kann man ∂G in $\operatorname{supp} V$ darstellen in der Form $x_n = g(x')$, und die Behauptung folgt aus Satz 5.25.

Im allgemeinen Fall wird der Satz durch eine Partition der Eins bewiesen. Es existiert eine offene Überdeckung $\overline{G} \subset \bigcup_{i \in I} W_i$ mit folgender Eigenschaft: Entweder ist $W_i \subset$

G oder (nach Ummummerierung) $W_i = U_i \times (\alpha_i, \beta_i)$ mit $U_i \subset \mathbb{R}^{n-1}$ und $\overline{G} \cap W_i = \{x \in U_i \times (\alpha_i, \beta_i) : x_n \leq g_i(x')\}$ mit $g_i \in C^1(U_i; \mathbb{R})$.

Da \overline{G} kompakt ist, kann man eine endliche Teilüberdeckung $\{W_1, \dots, W_N\}$ auswählen. Eine zugehörige Partition der Eins besteht aus Funktionen $\varphi_i \in C^\infty(\mathbb{R}^n; \mathbb{R})$ mit $\text{supp } \varphi \subset W_i$ und $\sum_{i=1}^N \varphi_i = 1$ in \overline{G} . (Eine solche Partition der Eins kann unter Verwendung der Funktion $t \mapsto \exp(-\frac{1}{1-t^2})$ konstruiert werden.) Wegen $\sum_{i=1}^N \varphi_i = 1$ ist

$$\begin{aligned} \int_G \text{div } V(x) dx &= \sum_{i=1}^N \int_G \text{div } V(x) \varphi_i(x) dx \\ &= \sum_{i=1}^N \int_{\partial G} \langle V(x) \varphi_i(x), \nu(x) \rangle dA(x) = \int_{\partial G} \langle V(x), \nu(x) \rangle dA(x). \end{aligned}$$

Hierbei wurde für die mittlere Gleichheit Lemma 5.24 bzw. Satz 5.25 verwendet. \square

5.27 Korollar. Sei G beschränktes Gebiet mit glattem Rand.

(i) Für $f \in C^1(\overline{G}; \mathbb{R})$ gilt

$$\int_G \partial_j f(x) dx = \int_{\partial G} f(x) \nu_j(x) dA(x).$$

(ii) Seien $f, g \in C^2(\overline{G}; \mathbb{R})$ und $\Delta := \sum_{j=1}^n \partial_j^2$ der Laplace-Operator. Dann gelten die Greenschen Formeln¹⁴

$$\begin{aligned} \int_G f \Delta g dx &= - \int_G \langle \nabla f, \nabla g \rangle dx + \int_{\partial G} f \frac{\partial g}{\partial \nu} dA(x), \\ \int_G (f \Delta g - g \Delta f) dx &= \int_{\partial G} \left(f \frac{\partial g}{\partial \nu} - g \frac{\partial f}{\partial \nu} \right) dA(x). \end{aligned}$$

Dabei ist $\frac{\partial f}{\partial \nu} = \langle \nabla f, \nu \rangle$ die Richtungsableitung von f in Richtung ν .

(iii) Für $f, g \in C^1(\overline{G}; \mathbb{R})$ gilt die Formel der partiellen Integration:

$$\int_G f(x) (\partial_j g)(x) dx = - \int_G (\partial_j f)(x) g(x) dx + \int_{\partial G} f(x) g(x) n_j(x) dA(x).$$

Beweis. (i) Setze im Satz von Gauß $V = (0, \dots, 0, f, 0, \dots, 0)^\top$ (j -te Stelle) und erhalte $\text{div } V = \partial_j f$ und $\langle V, \nu \rangle = f \nu_j$.

(ii) Setze $V := f \nabla g$. Dann ist $\text{div } V = f \Delta g + \langle \nabla f, \nabla g \rangle$ und $\langle V, \nu \rangle = f \langle \nabla g, \nu \rangle = f \frac{\partial g}{\partial \nu}$. Dies gibt die erste Formel, die zweite folgt, wenn man die Rollen von f und g vertauscht und die Gleichungen subtrahiert.

(iii) Wende (i) auf fg anstelle von f an. \square

¹⁴George Green, 14.7.1793–31.3.1841

5.28 Beispiel. Sei $V(x) := x$ ($x \in \mathbb{R}^n$). Dann ist $\operatorname{div} V(x) = \sum_{j=1}^n \partial_j x_j = n$. Somit folgt

$$\lambda_n(G) = \int_G 1 dx = \frac{1}{n} \int_G \operatorname{div} V(x) dx = \frac{1}{n} \int_{\partial G} \langle x, \nu(x) \rangle dA(x).$$

Speziell für $G = B_n = B(0, 1)$ (n -dimensionale Einheitskugel) folgt wegen $\nu(x) = x$ die Gleichheit

$$\lambda_n(B_n) = \frac{1}{n} \int_{S^{n-1}} \langle x, x \rangle dA(x) = \frac{1}{n} \int_{S^{n-1}} 1 dA(x) = \frac{1}{n} A(S^{n-1}).$$

In \mathbb{R}^2 lautet der Satz von Gauß

$$\int_G (\partial_1 W_1 + \partial_2 W_2) dx = \int_{\partial G} (W_1 \nu_1 + W_2 \nu_2) dA(x).$$

Setzt man $V := \begin{pmatrix} -W_2 \\ W_1 \end{pmatrix}$, erhält man

$$\int_G (\partial_1 V_2 - \partial_2 V_1) dx = \int_{\partial G} (V_2 \nu_1 - V_1 \nu_2) dA(x).$$

Falls $\gamma: (\alpha, \beta) \rightarrow \mathbb{R}^2$ eine Parametrisierung von ∂G ist, gilt

$$\nu(\gamma(\tau)) = \frac{1}{|\gamma'(\tau)|} \begin{pmatrix} \gamma_2'(\tau) \\ -\gamma_1'(\tau) \end{pmatrix}.$$

Das Kurvenintegral auf der rechten Seite wird dann zu

$$\begin{aligned} \int_{\partial G} (V_2 \nu_1 - V_1 \nu_2) dA(x) &= \int_{\alpha}^{\beta} \frac{1}{|\gamma'(\tau)|} (V_2(\gamma(\tau)) \gamma_2'(\tau) + V_1(\gamma(\tau)) \gamma_1'(\tau)) |\gamma'(\tau)| d\tau \\ &= \int_{\alpha}^{\beta} \langle V(\gamma(\tau)), \gamma'(\tau) \rangle d\tau = \int_{\alpha}^{\beta} \omega \end{aligned}$$

mit der 1-Form $\omega(x, h) := \langle V(x), h \rangle$.

Man definiert die Rotation in \mathbb{R}^2 durch

$$\operatorname{rot} V: \mathbb{R}^2 \rightarrow \mathbb{R}, \operatorname{rot} V := \partial_1 V_2 - \partial_2 V_1.$$

Somit erhält man:

5.29 Satz (Satz von Stokes¹⁵ im \mathbb{R}^2). Sei $G \subset \mathbb{R}^2$ beschränktes Gebiet mit glattem Rand und $V \in C^1(\overline{G}; \mathbb{R}^2)$. Dann gilt

$$\int_G \operatorname{rot} V(x) dx = \int_{\partial G} \langle V(x), dx \rangle.$$

¹⁵George Gabriel Stokes, 13.8.1819 – 1.2.1903

Im \mathbb{R}^3 definiert man für $V \in C^1(G; \mathbb{R}^3)$ die klassische Rotation durch

$$\operatorname{rot} V := \nabla \times V := \begin{pmatrix} \partial_2 V_3 - \partial_3 V_2 \\ \partial_3 V_1 - \partial_1 V_3 \\ \partial_1 V_2 - \partial_2 V_1 \end{pmatrix}.$$

Der Satz von Stokes im \mathbb{R}^3 handelt von zweidimensionalen Flächen im \mathbb{R}^3 . Sei $G \subset M$ für eine zweidimensionale Fläche $M \subset \mathbb{R}^3$, und sei $\nu(x)$ der Normalenvektor zu M , $\nu_{\partial G}(x)$ der Normalenvektor zu ∂G „auf M “ und $t(x)$ der Tangenteneinheitsvektor zu ∂G . Dann bilden für $x \in \partial G$ die Vektoren $\nu(x), \nu_{\partial G}(x), t(x)$ eine Orthonormalbasis des \mathbb{R}^3 . Der folgende Satz wird nicht bewiesen:

5.30 Satz (Satz von Stokes im \mathbb{R}^3). Sei $M \subset \mathbb{R}^3$ eine 2-Fläche, $G \subset M$ beschränkt mit glattem Rand. Für $V \in C^1(\overline{G}; \mathbb{R}^3)$ gilt

$$\int_G \langle \operatorname{rot} V(x), \nu(x) \rangle dA(x) = \int_{\partial G} \langle V(x), dx \rangle.$$

5.31 Bemerkung. (i) Man beachte, dass auf der linken Seite ein Flächenintegral, auf der rechten Seite ein Kurvenintegral steht. Definiert man die 1-Form $\omega(x, h) := \langle V(x), h \rangle$, so steht auf der rechten Seite $\int_{\partial G} \omega$.

(ii) Es gelten die folgenden Rechenregeln: Seien f eine reellwertige Funktion und V wie oben, dann ist

$$\begin{aligned} \operatorname{rot} \nabla f &= 0, \\ \operatorname{div} \operatorname{rot} V &= 0, \\ \operatorname{div}(fV) &= \langle \nabla f, V \rangle + f \operatorname{div} V, \\ \operatorname{rot}(fV) &= (\nabla f) \times V + f \operatorname{rot} V, \\ \operatorname{div} \nabla f &= \Delta f, \\ \Delta V &= \nabla \operatorname{div} V - \operatorname{rot} \operatorname{rot} V, \end{aligned}$$

wobei der Laplaceoperator in der letzten Zeile komponentenweise zu verstehen ist.

(iii) In die Definition von $\int_G, \int_{\partial G}$ geht eine gewisse „Orientierung“ mit ein, denn die Integrale sind nur unabhängig von Parametrisierungen in äquivalenten Klassen, d.h. unter Transformationen φ mit $\det \varphi' > 0$. Allgemein heißt eine Basis des \mathbb{R}^n „positiv orientiert“, falls $\det B > 0$, wobei $B = (b_1, \dots, b_n)$ und $\{b_1, \dots, b_n\}$ Basis des \mathbb{R}^n . Zwei Basen heißen „gleich orientiert“, falls $A = \Phi B$ mit $\det \Phi > 0$. Beispielsweise ist die Standardbasis positiv orientiert und über $\gamma : \mathbb{R}^p \rightarrow \mathbb{R}^n$ erhält man eine Orientierung in den Tangentialräumen des von γ beschriebenen p -Flächenstücks Γ aus der Orientierung von $\{e_1, \dots, e_p\}$, d.h. wir haben eine natürliche Orientierung im Tangentialraum in $\gamma(u) : \operatorname{Or}(u) = \{\partial_1 \gamma(u), \dots, \partial_p \gamma(u)\}$. Eine p -Fläche heißt „orientierbar“, falls alle so erhaltenen Orientierungen $\operatorname{Or}(u)$ gleichorientiert sind.

Auf ∂G wird eine „positive Orientierung induziert“, falls gilt: Ist $p \in \partial G$ und $\{w_1, \dots, w_{n-1}\}$ eine Basis des Tangentialraums von ∂G in p , so ist für jeden nach „außen“ weisenden Vektor $w = \{w, w_1, \dots, w_{n-1}\}$ positiv orientiert (d.h. gleich orientiert zur gegebenen Orientierung des \mathbb{R}^n).

(iv) Führt man „ p -Differentialformen“ ω ein für $1 \leq p \leq n$ – der Fall $p = 1$ ist bekannt – sowie einen Ableitungsoperator d , der aus $(p-1)$ -Formen geeignete p -Formen macht – Beispiel $p = 1$, $p-1 = 0$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ Funktion, Nullform, so

$$df := \sum_{k=1}^n \partial_k f dx_k,$$

also df die zugehörige 1-Form –, so lassen sich die Sätze von Gauß und von Stokes vereinheitlichen zu

$$\boxed{\int_M d\omega = \int_{\partial M} \omega}$$

wobei M eine orientierte p -Fläche und ω eine $(p-1)$ -Form ist. Für $p = n$ erhält man den Gaußschen Satz, für $1 \leq p \leq n-1$ den Stokesschen Satz.

Wir haben in diesem Abschnitt gesehen, dass Flächen in verschiedener Weise beschrieben werden können. Dies wird in folgendem Satz zusammengefasst, dessen Beweis sich im Wesentlichen aus dem Satz über lokale Umkehrbarkeit ergibt:

5.32 Satz. Seien $M \subset \mathbb{R}^n$ eine Menge und $m < n$. Dann sind äquivalent:

- (i) Für alle $x \in M$ existiert eine offene Umgebung $V \subset \mathbb{R}^n$ von x und eine Parametrisierung $\gamma : \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$ einer m -dimensionalen Fläche mit $M \cap V = \gamma(U)$.
- (ii) Für alle $x \in M$ existieren offene Umgebungen $V \subset \mathbb{R}^n$, $U \subset \mathbb{R}^m$ und $f \in C^1(U; \mathbb{R}^{n-m})$ mit $M \cap V = \text{graph } f = \{(u, f(u)) : u \in U\}$.
- (iii) Für alle $x \in M$ existieren offene Umgebungen $V \subset \mathbb{R}^n$, $W \subset \mathbb{R}^n$ und $g \in C^1(W; \mathbb{R}^{n-m})$ so, dass der Rang der Matrix $g'(y)$ für alle $y \in W$ gleich $n-m$ ist und $M \cap V = \{x \in W : g(x) = 0\}$.

5.33 Definition. Eine Menge $M \subset \mathbb{R}^n$ heißt eine m -dimensionale (C^1 -)Untermannigfaltigkeit des \mathbb{R}^n , falls M die äquivalenten Bedingungen aus Satz 5.32 erfüllt. Eine Parametrisierung wie in (i) heißt eine Karte von M . Ersetzt man überall C^1 durch C^k mit $k \in \mathbb{N} \cup \{\infty\}$, heißt M eine C^k -Untermannigfaltigkeit.

6. Lokale Umkehrbarkeit und implizite Funktionen

6.1 Worum geht's? In diesem kurzen Abschnitt werden drei wichtige Sätze vorgestellt, welche für Beweise der Analysis in vielen Fällen nützlich sind: Der Banachsche Fixpunktsatz, der Satz über lokale Umkehrbarkeit und der damit verwandte Satz über implizite Funktionen. Der Satz über implizite Funktionen wurde in den letzten Abschnitten bereits mehrmals verwendet, z.B. um eine Nullstellenmenge als Graph einer Funktion zu schreiben.

Wir beginnen mit einem Fixpunktsatz, der an vielen Stellen der Analysis eine zentrale Rolle spielt.

6.2 Definition. Sei (M, d) ein metrischer Raum. Dann heißt eine Abbildung $\Phi: M \rightarrow M$ kontrahierend oder eine Kontraktion, falls es ein $c \in [0, 1)$ gibt mit $d(\Phi(x), \Phi(y)) \leq cd(x, y)$ für alle $x, y \in M$.

6.3 Satz (Banachscher Fixpunktsatz). Sei (M, d) vollständiger metrischer Raum und $\Phi: M \rightarrow M$ eine Kontraktion. Dann besitzt Φ genau einen Fixpunkt $z \in M$, d.h. $\Phi(z) = z$.

Definiert zu $x_0 \in M$ die Folge $(x_n)_{n \in \mathbb{N}_0} \subset M$ durch $x_{n+1} := \Phi(x_n)$ ($n \in \mathbb{N}_0$), so gilt $x_n \rightarrow z$ ($n \rightarrow \infty$) und $d(x_n, z) \leq \frac{c^n}{1-c} d(x_1, x_0)$ (a priori-Abschätzung).

Beweis. Es gilt $d(x_n, x_{n+1}) \leq cd(x_{n-1}, x_n) \leq \dots \leq c^n d(x_0, x_1)$ und damit für $m \geq 1$

$$\begin{aligned} d(x_n, x_{n+m}) &\leq d(x_n, x_{n+1}) + \dots + d(x_{n+m-1}, x_{n+m}) \\ &\leq (1 + c + c^2 + \dots + c^{m-1})d(x_n, x_{n+1}) \leq \frac{c^n}{1-c} d(x_0, x_1). \end{aligned}$$

Wegen $c^n \rightarrow 0$ ($n \rightarrow \infty$) ist $(x_n)_n \subset M$ eine Cauchyfolge. Da M vollständig ist, existiert der Grenzwert $z := \lim_{n \rightarrow \infty} x_n$. Als Kontraktion ist Φ Lipschitz-stetig und damit insbesondere stetig, also gilt

$$\Phi(z) = \Phi\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} \Phi(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = z.$$

Da die Metrik $d: M \times M \rightarrow \mathbb{R}$ stetig ist, folgt

$$d(x_n, z) = d(x_n, \lim_{m \rightarrow \infty} x_{n+m}) = \lim_{m \rightarrow \infty} d(x_n, x_{n+m}) \leq \frac{c^n}{1-c} d(x_0, x_1).$$

Somit existiert ein Fixpunkt z und erfüllt die a priori-Abschätzung. Zu zeigen ist noch die Eindeutigkeit. Seien dazu z, z' Fixpunkte von Φ . Dann ist $d(z, z') = d(\Phi(z), \Phi(z')) \leq cd(z, z')$, und wegen $c < 1$ folgt $z = z'$. \square

Für Funktionen $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ hatten wir gezeigt, dass strenge Monotonie ein notwendiges und hinreichendes Kriterium für Injektivität ist. Für Funktionen $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, n \geq 2$, existiert kein Analogon. Wir wollen hier zunächst für den Fall $n = m$ ein hinreichendes Kriterium für die lokale Invertierbarkeit von Funktionen beweisen.

6.4 Satz (Satz über lokale Umkehrbarkeit). Seien $U \subset \mathbb{R}^n$ offen, $f \in C^1(U, \mathbb{R}^n)$, $a \in U$ und $\det f'(a) \neq 0$. Dann gibt es eine offene Umgebung $V \subset U$ von a mit:

- (i) $f|_V$ ist injektiv.
- (ii) $f(V)$ ist offen.
- (iii) $g := (f|_V)^{-1}$ ist stetig differenzierbar.

Das heißt: $f|_V : V \rightarrow f(V)$ ist Diffeomorphismus.

Beweis. Seien o.B.d.A. $a = 0$, $f(a) = 0$ und $f'(0) = I_n$ ($n \times n$ -Einheitsmatrix).

(a) Festlegung von V : Für $y \in \mathbb{R}^n$ definiere

$$\Phi_y : U \rightarrow \mathbb{R}^n, \quad x \mapsto \Phi_y(x) := x - f(x) + y.$$

Ein Fixpunkt der Funktion Φ ist Urbild von y . Im zweiten Schritt dieses Beweises werden wir $r > 0$ so bestimmen, dass

$$\Phi_y : \overline{B(0, 2r)} \rightarrow \overline{B(0, 2r)}$$

für alle $y \in B(0, r)$ kontrahierend ist.

Aus dem Banachschen Fixpunktsatz folgt dann: Für alle $y \in B(0, r)$ existiert genau ein $x \in B(0, 2r)$ mit $f(x) = y$. Es wird in Teil (b) des Beweises folgen: $x \in B(0, 2r)$. Definiere nun ($r > 0$ sei bereits bestimmt)

$$V := f^{-1}(B(0, r)) \cap B(0, 2r).$$

V ist offen, da f stetig ist, und

$$f : V \rightarrow B(0, r) = f(V)$$

ist bijektiv. Die ersten beiden Behauptungen des Satzes sind mit dieser Wahl des V somit erfüllt.

(b) Wahl von r : Aus $f'(0) = I_n$ und der Stetigkeit der ersten Ableitung folgt:

$$\exists r > 0 \quad \forall x, |x| \leq 2r : \|I_n - f'(x)\|_{L(\mathbb{R}^n, \mathbb{R}^n)} \leq \frac{1}{2}.$$

Aus dem Mittelwertsatz folgt für $\Phi_0 = \text{id} - f$ mit $0 < \theta_j < 1$, $j = 1, \dots, n$, und $x \in B(0, 2r)$:

$$|x - f(x)| = |\Phi_0(x) - \Phi_0(0)| = \left| \left((\Phi_0^j)'(\theta_j x, x) \right)_{j=1}^n \right| \leq \frac{1}{2} |x|$$

(hierbei sei $\Phi_0 =: (\Phi_0^j)_{j=1}^n$.)

Für $|y| < r$ und $|x| \leq 2r$ ergibt sich hieraus

$$|\Phi_y(x)| \leq |x - f(x)| + |y| < 2r,$$

d.h.: $\forall y \in \mathbb{R}^n, |y| < r : \Phi_y : \overline{B(0, 2r)} \rightarrow \overline{B(0, 2r)}$.

Insbesondere: $\Phi_y(x) = x \Rightarrow |x| < 2r$.

(c) Zur Kontraktion von Φ_y in $\overline{B(0, 2r)}$: Seien $x_1, x_2 \in \overline{B(0, 2r)}$. Dann gilt:

$$\Phi_y^j(x_2) - \Phi_y^j(x_1) = (\Phi_y^j)'(x_1 + \theta_j(x_2 - x_1), x_2 - x_1)$$

mit $0 < \theta_j < 1, j = 1, \dots, n$. Hieraus folgt:

$$|\Phi_y(x_2) - \Phi_y(x_1)| \leq \frac{1}{2}|x_2 - x_1|.$$

(d) Zur Behauptung (iii): Es genügt, die Stetigkeit der Umkehrabbildung g zu zeigen. Die Differenzierbarkeit von g folgt dann aus Satz 2.17; die Stetigkeit der Ableitung ist sofort aus der in Satz 2.17 bewiesenen Darstellung der Ableitung und der Stetigkeit der Inversenbildung $GL(n, \mathbb{R}) \rightarrow \mathbb{R}^{n \times n}, A \mapsto A^{-1}$ ersichtlich.

Seien nun $x_i \in \overline{B(0, 2r)}$ und $y_i := f(x_i)$ für $i = 1, 2$. Dann ist

$$\begin{aligned} \Phi_0(x_i) &= x_i - f(x_i) \quad , \text{ also} \\ x_1 - x_2 &= \Phi_0(x_1) - \Phi_0(x_2) + f(x_1) - f(x_2) \quad \text{und somit} \\ |x_1 - x_2| &\leq \frac{1}{2}|x_2 - x_1| + |f(x_2) - f(x_1)|. \end{aligned}$$

Hieraus folgt sofort die Lipschitz-Stetigkeit von g . □

In den Anwendungen tritt häufig folgendes Problem auf: Gegeben sei eine Abbildung

$$f : D \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n, (x, u) \mapsto f(x, u).$$

Dabei seien $x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n, u = (u_1, \dots, u_m)^\top \in \mathbb{R}^m$ und $f(x, u) = (f_1(x, u), \dots, f_n(x, u))^\top \in \mathbb{R}^n$.

Gesucht ist nun eine Abbildung

$$\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^n, u \mapsto \varphi(u)$$

mit $f(\varphi(u), u) = 0$; φ heißt „Auflösung“.

Im allgemeinen ist eine solche Auflösung nicht möglich, wie für $n = m = 1$ das Beispiel

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, u) \mapsto f(x, u) := 1 + x^2 + u^2$$

zeigt; hier ist keine reelle Auflösung möglich. Es sind also Voraussetzungen an f erforderlich, mindestens z.B.:

$$\exists(x_1, u_1) \in \mathbb{R}^{n+m} : f(x_1, u_1) = 0.$$

Wir werden folgende Schreibweise benutzen: Für $h \in \mathbb{R}^n$ und $k \in \mathbb{R}^m$ definiere

$$f' = f'(x, u; h, k) =: \tilde{f}'_1(x, u; h) + \tilde{f}'_2(x, u; k)$$

mittels

$$\begin{aligned} & \begin{pmatrix} \partial_1 f_1 & \dots & \partial_n f_1 & \partial_{n+1} f_1 & \dots & \partial_{n+m} f_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial_1 f_n & \dots & \partial_n f_n & \partial_{n+1} f_n & \dots & \partial_{n+m} f_n \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \\ k_1 \\ \vdots \\ k_m \end{pmatrix} \\ &= \begin{pmatrix} \partial_1 f_1 & \dots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_n & \dots & \partial_n f_n \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + \begin{pmatrix} \partial_{n+1} f_1 & \dots & \partial_{n+m} f_1 \\ \vdots & & \vdots \\ \partial_{n+1} f_n & \dots & \partial_{n+m} f_n \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix}. \end{aligned}$$

6.5 Satz (Satz über implizite Funktionen). Seien $D \subset \mathbb{R}^n \times \mathbb{R}^m$ offen und $f \in C^1(D, \mathbb{R}^n)$. Es gebe $a \in \mathbb{R}^n, b \in \mathbb{R}^m$ mit $f(a, b) = 0$ und $\det \tilde{f}'_1(a, b) \neq 0$. Dann existieren eine Umgebung $U = U(b) \subset \mathbb{R}^m$ von b und eine eindeutig bestimmte Abbildung $\varphi \in C^1(U, \mathbb{R}^n)$ mit:

$$\forall u \in U : f(\varphi(u), u) = 0.$$

Beweis. Idee:

- Anwendung des Satzes über die lokale Umkehrbarkeit (Satz 6.4) auf die Abbildung $x \mapsto f(x, u)$, wobei u fest sei.
- Wähle $\varphi(u) := g(0, u)$, wobei $g(\cdot, u)$ die Umkehrabbildung zu $f(\cdot, u)$ bezeichne.
- Untersuchung der Eigenschaften von φ .

Definiere

$$F : D \rightarrow \mathbb{R}^{n+m}, (x, u) \mapsto F(x, u) := (f(x, u), u).$$

F ist stetig differenzierbar mit der Ableitung

$$F'(x, u) = \begin{pmatrix} \tilde{f}'_1(x, u) & \tilde{f}'_2(x, u) \\ 0 & I_m \end{pmatrix}.$$

Es ist $\det F'(x, u) = \det \tilde{f}'_1(x, u)$, also ist insbesondere $\det F'(a, b) \neq 0$.

Aus Satz 6.4 folgt nun, dass zu $(a, b) \in D$ eine Umgebung $W \subset D$ existiert, so dass $F|_W$ eine C^1 -invertierbare Abbildung mit offenem Wertebereich ist. $G : F(W) \rightarrow W$ sei die Umkehrabbildung von $F|_W$. Aus $F(x, u) = (y, v)$ erhält man für $G: G(y, v) = (x, u)$, wobei $v = u$ ist. Setze $g(y, v) := x$. Dann ist $g \in C^1(F(W), \mathbb{R}^n)$. g liefert die gesuchte Abbildung φ mittels $\varphi(v) := g(0, v)$, denn:

Da $D(G) = F(W)$ offen ist, existiert eine Umgebung $U = U(b) \subset \mathbb{R}^m$ mit: $\forall v \in U(b) : (0, v) \in F(W)$. Dort ist φ erklärt. Aus $F \circ G = \text{id}_{F(W)}$ folgt für $(y, v) \in F(W)$: $f(g(y, v), v) = y$. Mit $y = 0$ (und $v = u$) erhalten wir wie gewünscht:

$$f(\varphi(u), u) = 0 \quad \forall u \in U.$$

Zu zeigen bleibt die Eindeutigkeit dieser Auflösung.: Nach Definition der Abbildung g ist $g(f(x, u), u) = x$. Aus $f(x, u) = 0$ folgt somit $g(0, u) = x$, was äquivalent ist zu $x = \varphi(u)$. \square

6.6 Beispiele. (i) Seien $n = m = 1$ und $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $(x, u) \mapsto f(x, u) := x + x^2 - u^2$. Dann sind $f(0, 0) = 0$, $\tilde{f}'_1(x, u) = 1 + 2x$, also $\tilde{f}'_1(0, 0) = 1$. Die Voraussetzungen des Satzes über implizit gegebene Funktionen sind somit für $(a, b) = (0, 0)$ erfüllt.

Durch formales Auflösen erhalten wir: $\varphi(u) = -\frac{1}{2} + \sqrt{\frac{1}{4} + u^2}$ für $u \in \mathbb{R}$ (da $\varphi(0) \stackrel{!}{=} 0$).

(ii) Seien $n = 2$, $m = 1$ und

$$f : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}^2, \quad (x, y, u) \mapsto f(x, y, u) = \begin{pmatrix} 1 + x + y - u \\ -1 + x + y^2 + u \end{pmatrix}.$$

Für $(a, b) = ((0, 0), 1)$ erhalten wir $f(a, b) = (0, 0)$ und aus

$$\tilde{f}'_1(x, y, u) = \begin{pmatrix} 1 & 1 \\ 1 & 2y \end{pmatrix}$$

$\det \tilde{f}'_1(a, b) = 2y - 1 \Big|_{y=0} = -1$. Die Voraussetzungen des Satzes sind somit erfüllt.

Formales Auflösen und die Bedingung $\varphi(1) = (0, 0)$ liefert:

$$\varphi(u) = \frac{1}{2} \begin{pmatrix} 2u - 3 + \sqrt{9 - 8u} \\ 1 - \sqrt{9 - 8u} \end{pmatrix}.$$

Literatur

- [AE99] AMANN, Herbert ; ESCHER, Joachim: *Analysis. II*. Birkhäuser Verlag, Basel, 1999 (Grundstudium Mathematik. [Basic Study of Mathematics]). – xii+412 S.
- [BF89] BARNER, Martin ; FLOHR, Friedrich: *Analysis. II*. Second. Walter de Gruyter & Co., Berlin, 1989 (de Gruyter Lehrbuch. [de Gruyter Textbook]). – 449 S.
- [DR11] DENK, Robert ; RACKE, Reinhard: *Kompendium der Analysis. Ein kompakter Bachelor-Kurs von reellen Zahlen zu partiellen Differentialgleichungen. Band 1: Differential- und Integralrechnung, gewöhnliche Differentialgleichungen*. Wiesbaden: Vieweg+Teubner, 2011. – xii + 317 S.
- [For77] FORSTER, Otto: *Analysis 2. Differentialrechnung im \mathbf{R}^n* . Verlag Vieweg, Braunschweig, 1977. – viii+163 S. – Gewöhnliche Differentialgleichungen, Rororo Vieweg: Mathematik Grundkurs, No. 31
- [Heu91] HEUSER, Harro: *Lehrbuch der Analysis. Teil 2*. Sixth. B. G. Teubner, Stuttgart, 1991 (Mathematische Leitfäden. [Mathematical Textbooks]). – 737 S.
- [Rud99] RUDIN, Walter: *Reelle und komplexe Analysis*. R. Oldenbourg Verlag, Munich, 1999. – xiv+499 S. – Translated from the third English (1987) edition by Uwe Krieg
- [Wal07] WALTER, Rolf: *Einführung in die Analysis. 2*. Walter de Gruyter & Co., Berlin, 2007 (de Gruyter Lehrbuch. [de Gruyter Textbook]). – x+465 S.

Index

- Ableitung, 10
- Ausgleichsgerade, 28
- Diffeomorphismus, 19
- differenzierbar, 10
- Divergenz, 16
- Extrema unter Nebenbedingungen, 29
- Falllinien, 15
- gleichmäßig konvergent, 9
- Gradient, 11
- Höhenlinien, 15
- Höldersche Ungleichung, 2
- Hessematrix, 26
- Hilbertraum, 1
- Intervallschachtelung, 5
- Jacobi-Matrix, 12
- Kettenregel, 14
- konjugierte Exponent, 2
- konvex, 28, 29
- kritische Stelle, 25
- $L(X, Y)$, 8
- Laplace-Operator, 16
- lineares Funktional, 8
- ℓ^p , 2
- Methode der kleinsten Fehlerquadrate, 28
- Minkowskische Ungleichung, 2
- Mittelwertsatz, 19
- Multiindex, 25
- Normalgleichung, 28
- Operator, 8
- Operatornorm, 8
- partiell differenzierbar, 11
- Polarkoordinaten, 18
- Portfolio-Optimierung, 33
- Prähilbertraum, 1
- Produktregel, 14
- quadratische Form, 32
- Richtungsableitung, 15
- Rotation, 16
- Satz
 - Mittelwert-, 19
 - von Bolzano-Weierstraß, 4
 - von Schwarz, 22
 - von Taylor, 23
 - Zwischenwert-, 7
- Skalarprodukt, 1
- Standard-Skalarprodukt, 1
- stetige lineare Operatoren, 8
- Supremumsnorm, 9
- totalbeschränkt, 5
- Weg, 7
- wegzusammenhängend, 7
- Youngsche Ungleichung, 2
- zusammenhängend, 7
- Zwischenwertsatz, 7