Mehrdimensionale Analysis

Inhalt

I Stetigkeit **1**

1 Definition und grundlegende Eigenschaften 1
2 Kompaktheit 5
3 Stetige Abbildungen kompakter Räume. Zwischenwertsatz . . . 9

II Differentialrechnung im $\mathbb{R}^n$ **11**

4 Partielle Ableitungen 11
5 Differenzierbarkeit 14
6 Vektorfelder, Gradient, Divergenz, Rotation, Laplace 20
7 Die Taylor-Formel 26
8 Der Satz über implizite Funktionen 32
9 Untermannigfaltigkeiten 38
10 Parameterabhängige Integrale 44

III Gewöhnliche Differentialgleichungen **49**

11 Definition und Beispielklassen 49
12 Existenz- und Eindeutigkeitssatz 52
13 Lineare Differentialgleichungen 57
14 Lineare Differentialgleichungen mit konstanten Koeffizienten . . 63
15 Differentialgleichungen 2. Ordnung 69

IV Grundlagen der Funktionentheorie **74**

16 Holomorphe Funktionen 74
17 Die Cauchysche Integralformel 82
18 Der Residuensatz 86

V Das Lebesgue-Integral **92**

19 Treppenfunktionen und Hüllreihen 92
20 Das Lebesgue-Integral 96
21 Der Kleine Satz von Beppo Levi 98
22 Der kleine Satz von Fubini 101
23 Nullmengen und Satz von Riesz-Fischer 105
24 Der Transformationssatz 110
25 Konvergenzsätze. Satz von Fubini 116
26 Beweis des Transformationssatzes 120
27 Integration über Untermannigfaltigkeiten 125
28 Der Gaußsche Integralsatz 130

Literatur

K. Königsberger, "Analysis 2," Springer 2004.
O. Forster, "Analysis 2," Vieweg 2005.

Teil I
Stetigkeit

1 Definition und grundlegende Eigenschaften

Wir definieren Stetigkeit für Abbildungen $f : X \to Y$ zwischen metrischen Räumen (X, d_X) und (Y, d_Y), die wir in Definition 20.13 im letzten Semester definiert hatten. Wichtige Beispiele sind Metriken, die aus einer Norm (Definition 20.8) hervorgehen. Zentral sind die Begriffe Offenheit (Definition 24.1 und 24.3), Konvergenz in metrischen Räumen (Definition 25.1), Vollständigkeit (Definition 25.5).

Definition 1.1 Es seien (X, d_X) und (Y, d_Y) metrische Räume und $a \in X$. Eine Abbildung $f : X \to Y$ heißt *stetig* in a, wenn für alle $\epsilon > 0$ ein $\delta > 0$ existiert, so daß

$$d_Y\big(f(x), f(a)\big) < \epsilon \qquad \text{für alle } x \in X \text{ mit } d_X(x, a) < \delta \ .$$

Die Abbildung $f : X \to Y$ heißt *stetig in* X, wenn f in jedem Punkt $a \in X$ stetig ist.

Eine wichtige Klasse stetiger Abbildungen ist:

Definition 1.2 Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt *Lipschitz-stetig*, wenn es eine Konstante $L \geq 0$ gibt, so daß für alle $x, x' \in X$ gilt

$$d_Y(f(x), f(x')) \leq L \cdot d_X(x, x') \ .$$

Satz 1.3 *Es sei $(V, \| \ \|)$ ein normierter Vektorraum. Dann ist die Norm, aufgefaßt als Abbildung $\| \ \| : V \to \mathbb{R}$ zwischen metrischen Räumen, Lipschitz-stetig mit Lipschitz-Konstante 1.*

Beweis. Zu zeigen ist $\big|\|x\| - \|y\|\big| = d(\|x\|, \|y\|) \leq d(x, y) = \|x - y\|$. Das folgt aus den Dreiecksungleichungen $\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$ und $\|y\| = \|y - x + x\| \leq \|x - y\| + \|x\|$. □

Satz 1.4 *Jede lineare Abbildung $f : V \to W$ zwischen endlich-dimensionalen normierten Vektorräumen $(V, \| \ \|_V)$ und $(W, \| \ \|_W)$ ist Lipschitz-stetig.*

Beweis. Es sei $(v_1, \ldots, v_n)$ Basis von V und $M := \max_{i \in \{1,\ldots,n\}} \big(\|f(v_i)\|_W\big)$. Dann gilt für $x = \sum_{i=1}^n x_i v_i$ und $a = \sum_{i=1}^n a_i v_i$, mit $x_i, a_i \in \mathbb{K}$,

$$\|f(x) - f(a)\|_W = \Big\| \sum_{i=1}^n (x_i - a_i) f(v_i) \Big\|_W \leq M \sum_{i=1}^n |x_i - a_i| =: M \|x - a\|_1 \ .$$

Wegen der Äquivalenz aller Normen in endlich-dimensionalen Vektorräumen gibt es ein $C > 0$, so daß $\|y\|_1 \leq C\|y\|_V$ für alle $y \in V$ und damit $\|f(x) - f(a)\|_W \leq CM\|x - a\|_V$. □

Im unendlich-dimensionalen Fall benötigen wir die Beschränktheit:

Satz 1.5 *Seien $(V, \| \ \|_V)$ und $(W, \| \ \|_W)$ normierte Vektorräume. Eine lineare Abbildung $A : V \to W$ ist genau dann stetig, wenn es eine reelle Zahl $C \geq 0$ gibt mit $\|A(x)\|_W \leq C\|x\|_V$ für alle $x \in V$. In diesem Fall ist A sogar Lipschitz-stetig.*

Beweis. ($\Rightarrow$) Sei A stetig in $0 \in V$. Wegen $A(0) = 0$ gibt es dann zu $\epsilon = 1$ ein $\delta > 0$, so daß

$$d_W\big(A(z), A(0)\big) = \|A(z)\|_W < 1 \qquad \text{für alle } z \in V \text{ mit } d_V(z, 0) = \|z\|_V < \delta \ .$$

Für $x \in V \setminus \{0\}$ setzen wir $z := \frac{\delta}{2\|x\|_V} \cdot x$, mit $\|z\|_V = \frac{\delta}{2} < \delta$, und erhalten

$$\|A(x)\|_W = \big\|\tfrac{2\|x\|_V}{\delta} A(z)\big\|_W = \tfrac{2\|x\|_V}{\delta}\|A(z)\|_W < \tfrac{2}{\delta}\|x\|_V \qquad \text{für alle } x \in V \setminus \{0\} \ .$$

Damit ist $C = \frac{2}{\delta}$. Die Ungleichung $\|A(x)\|_W \leq C\|x\|_V$ gilt auch für $x = 0$.

($\Leftarrow$) Für $x, a \in V$ gilt

$$d_W\big(A(x), A(a)\big) = \big\|A(x) - A(a)\big\|_W = \big\|A(x - a)\big\|_W \leq C\|x - a\|_V \ ,$$

d.h. A ist Lipschitz-stetig und damit stetig. □

Sind V, W normierte Vektorräume, dann bildet in Verallgemeinerung von Satz 25.9 aus dem letzten Semester die Menge

$$\mathcal{B}(V, W) := \{A : V \to W \ : \ A \text{ linear und stetig}\}$$

aller linearen stetigen (d.h. beschränkten) Abbildungen einen Vektorraum, der mit der Operatornorm

$$\|A\|_{op} := \sup_{v \in V \ , \ \|v\| \leq 1} \|A(v)\|_W$$

zu einem normierten Vektorraum wird. Ist W vollständig (also Banach-Raum), so ist auch $(\mathcal{B}(V, W), \| \ \|_{op})$ ein Banach-Raum, was analog zu Satz 25.9 aus dem letzten Semester bewiesen wird.

Satz 1.6 (Folgenkriterium der Stetigkeit) *Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen X, Y ist genau dann stetig im Punkt $a \in X$, wenn für jede Folge $(x_k)_{k \in \mathbb{N}}$ von Punkten aus X mit $\lim_{k \to \infty} x_k = a$ gilt*

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

Beweis. Identisch zum Beweis von Satz 12.6 im ersten Semester, wenn dort der Betrag durch den Abstand ersetzt wird. □

Über das Folgenkriterium läßt sich Satz 12.8 aus dem ersten Semester auf metrische Räume übertragen:

Satz 1.7 *Es seien X, Y, Z metrische Räume und W ein normierter Vektorraum.*

i) *Sind $f_1, f_2 : X \to W$ stetig in $a \in X$, so ist auch $f_1 + f_2 : X \to W$ stetig in a.*

ii) *Sind $f_1, f_2 : X \to \mathbb{K}$ stetig in $a \in X$, so ist auch $f_1 \cdot f_2 : X \to \mathbb{K}$ stetig in a.*

iii) *Es seien $F, g : X \to \mathbb{K}$ stetige Funktionen. Ist $g(a) \neq 0$, so ist $g(x) \neq 0$ in einer Umgebung $U \subset X$ von a, und die Funktion $\frac{f}{g} : U \to \mathbb{K}$ ist stetig in a.*

iv) *Sind $f : X \to Y$ stetig in a und $g : Y \to Z$ stetig in $f(a)$, so ist auch $g \circ f : X \to Z$ stetig in a.*

Satz 1.8 *Es seien X, Y metrische Räume. Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen X, Y ist genau dann stetig im Punkt $a \in X$, wenn zu jeder offenen Umgebung $V \subset Y$ von $f(a) \in V$ eine offene Umgebung $U \subset X$ von $a \in U$ existiert mit $f(U) \subset V$.*

Beweis ($\Rightarrow$) Sei $f : X \to Y$ stetig in a. Wähle eine beliebige offene Umgebung V von $f(a)$. Wegen der Offenheit von V existiert ein $\epsilon > 0$ mit $K_\epsilon(f(a)) \subset V$. Für ein zugehöriges $\delta > 0$ sei $U := K_\delta(a)$. Dann ist $f(U) \subset V$.

($\Leftarrow$) Nach Voraussetzung gibt es zu jeder offenen Kugel $V = K_\epsilon(f(a))$ eine offene Umgebung U von a mit $f(U) \subset V$. Diese enthält eine offene Kugel $K_\delta(a)$, d.h. $f(K_\delta(a)) \subset K_\epsilon(f(a))$. □

Damit läßt sich noch allgemeiner die (globale) Stetigkeit von Abbildungen zwischen topologischen Räumen definieren:

Definition 1.9 Eine Abbildung $f : X \to Y$ zwischen topologischen Räumen heißt *stetig auf ganz X*, wenn das Urbild $f^{-1}(V)$ jeder offenen Menge $V \subset Y$ offen in X ist.

Durch Bildung des Komplements folgt, daß auch das Urbild $f^{-1}(A)$ jeder abgeschlossenen Menge $A \subset Y$ abgeschlossen in X ist: Setze $A : Y \setminus V$ mit V offen, dann ist $f^{-1}(A) \not\subset f^{-1}(V)$. Da f auf ganz X definiert ist, folgt $f^{-1}(A) \subset X \setminus f^{-1}(V)$. Insbesondere gilt:

Satz 1.10 *Für eine stetige Abbildung $f : X \to \mathbb{R}$ gilt:*

i) *$U := \{x \in X \ : \ f(x) < c\}$ ist offen,*

ii) $A := \{x \in X \ : \ f(x) \leq c\}$ *ist abgeschlossen,*

iii) $N := \{x \in X \ : \ f(x) = c\}$ *ist abgeschlossen.*

Beweis. $]-\infty, c[\subset \mathbb{R}$ ist offen und $]-\infty, c] = \mathbb{R} \setminus]c, \infty[$ sowie $\{c\} = \mathbb{R} \setminus (]-\infty, c[\cup]c, \infty[)$ sind abgeschlossen. □

Definition 1.9 besagt *nicht*, daß eine stetige Abbildung offene Mengen auf offene Mengen abbildet. Das wäre falsch, wie das Beispiel $\cos : \mathbb{R} \to [-1, 1]$ zeigt.

Definition 1.11 Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen X, Y heißt *Homöomorphismus*, wenn gilt:

i) f ist bijektiv,

ii) $f : X \to Y$ und $f^{-1} : Y \to X$ sind stetig.

Zwei metrische Räume X, Y heißen *homöomorph*, wenn es einen Homöomorphismus $f : X \to Y$ gibt.

In dieser Definition können die metrischen Räume durch topologische Räume ersetzt werden. Insbesondere bilden Homöomorphismen offene Mengen auf offene Mengen ab, und abgeschlossene auf abgeschlossene.

Beispiel 1.12

i) Die offene Kugel $X = K_1(0) := \{y \in \mathbb{R}^n \ : \ \|y\| < 1\} \subset \mathbb{R}^n$ bezüglich einer beliebigen Norm ist homöomoph zum $\mathbb{R}^n$:
$f : \mathbb{R}^n \to X$ mit $f(x) = \frac{x}{1+\|x\|}$ ist Homöomorphismus mit $f^{-1}(y) = \frac{y}{1-\|y\|}$.

ii) $X = [0, 2\pi[\subset \mathbb{R}, \quad Y = S^1 = \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 = 1\}$
$f : X \to Y$ mit $f(t) = (\cos t, \sin t)$ ist bijektiv und stetig, aber $f^{-1} : Y \to X$ ist nicht stetig in $(1, 0)$. Die Folge von Punkten

$$p_k := \big(\cos(2\pi - \tfrac{1}{k+1}),\ \sin(2\pi - \tfrac{1}{k+1})\big)\,, \qquad k \in \mathbb{N}\,,$$

konvergiert gegen $(1, 0) = f(0)$, aber $f^{-1}(p_k) = 2\pi - \frac{1}{k+1}$ konvergiert nicht gegen 0.

Insbesondere muß die Umkehrung einer bijektiven stetigen Abbildung nicht stetig sein.

Stetige Funktionen lassen sich oft durch Grenzprozesse erhalten.

Definition 1.13 Seien $f : X \to Y$ sowie $f_k : X \to Y$ für $k \in \mathbb{N}$ Abbildungen zwischen metrischen Räumen X, Y. Die Folge der Abbildungen f_k *konvergiert gleichmäßig* gegen die Abbildung f, wenn zu jedem $\epsilon > 0$ ein $n \in \mathbb{N}$ existiert mit $d\big(f_k(x), f(x)\big) < \epsilon$ für alle $x \in X$ und alle $k \geq n$.

Satz 1.14 *Seien X, Y metrische Räume und $f_k : X \to Y$, für $k \in \mathbb{N}$, eine Folge stetiger Abbildungen, die gleichmäßig gegen eine Abbildung $f : X \to Y$ konvergiert. Dann ist f stetig.*

Beweis. Wir zeigen, daß f in jedem Punkt $a \in X$ stetig ist. Sei $\epsilon > 0$ gegeben. Wegen der gleichmäßigen Konvergenz gibt es ein $n \in \mathbb{N}$, so daß $d_Y\big(f_k(x), f(x)\big) < \frac{\epsilon}{3}$ für alle $k \geq n$ und alle $x \in X$. Da f_k in a stetig ist, gibt es ein $\delta > 0$ mit $d_Y\big(f_k(x), f_k(a)\big) < \frac{\epsilon}{3}$ für alle $x \in X$ mit $d_X(x, a) < \delta$. Dann folgt aus der Dreiecksungleichung

$$d_Y\big(f(x), f(a)\big) \leq d_Y\big(f(x), f_k(x)\big) + d_Y\big(f_k(x), f_k(a)\big) + d_Y\big(f_k(a), f(a)\big) < \epsilon$$

für alle $x \in X$ mit $d_X(x, a) < \delta$. □

Satz 1.15 *Es seien X, Y metrische Räume, und eine Folge (f_k) stetiger Funktionen $f_k : X \to Y$ sei eine Cauchy-Folge gleichmäßig auf X, d.h. zu jedem $\epsilon > 0$ gibt es ein $N \in \mathbb{N}$, so daß für alle $k, l \geq N$ und alle $x \in X$ gilt $d_Y\big(f_k(x), f_l(x)\big) < \epsilon$. Ist Y vollständig, so konvergiert (f_k) gleichmäßig gegen eine Grenzfunktion $f : X \to Y$, und diese ist stetig.*

Beweis. Zu jedem Punkt $x \in X$ ist $(f_k(x))_{k \in \mathbb{N}}$ eine Cauchy-Folge, die wegen der Vollständigkeit in Y konvergiert und somit punktweise eine Grenzfunktion $f : X \to Y$ definiert. Dann konvergiert $(f_k)_{k \in \mathbb{N}}$ gleichmäßig gegen f, und nach Satz 1.14 ist f stetig. □

2 Kompaktheit

Wir hatten im ersten Semester gesehen, daß Funktionen über kompakten Intervallen besondere Eigenschaften haben. Dort gilt der Zwischenwertsatz, der Satz vom Maximum/Minumum, und stetige Funktionen sind automatisch gleichmäßig stetig. Zur Verallgemeinerung dieser Eigenschaften auf metrische Räume wird der Begriff der Kompaktheit recht abstrakt definiert.

Definition 2.1 Sei A Teilmenge eines metrischen Raumes X. Unter einer *offenen Überdeckung* von A versteht man eine Familie $(U_i)_{i \in I}$ von offenen Teilmengen $U_i \subset X$ mit $A \subset \bigcup_{i \in I} U_i$, d.h. zu jedem Punkt $x \in A$ gibt es ein $i \in I$ mit $x \in U_i$.

i) Eine Teilmenge A eines metrischen Raumes X heißt *überdeckungskompakt* oder kurz *kompakt*, wenn es zu jeder offenen Überdeckung $(U_i)_{i \in I}$ von A endlich viele Indizes $i_1, \ldots, i_n \in I$ gibt, so daß $A \subset U_{i_1} \cup \cdots \cup U_{i_n}$. (Heine-Borelsche Überdeckungseigenschaft)

ii) Eine Teilmenge K eines metrischen Raums X heißt *folgenkompakt*, wenn jede Folge von Punkten aus K eine konvergente Teilfolge besitzt, deren Grenzwert in K liegt. (Bolzano-Weierstraß-Eigenschaft)

Es wird in i) also nicht gefordert, daß man A durch endlich viele offene Teilmengen von X überdecken kann. Das geht immer, denn A läßt sich durch X selbst überdecken, und X ist offen. Die Forderung ist, daß man jede exotische unendliche Überdeckung von A auf eine endliche Überdeckung reduzieren kann. Insbesondere wird durch Definition 2.1 für $A = X$ die Kompaktheit und Folgenkompaktheit metrischer Räume erklärt.

Satz 2.2 *Es sei X ein metrischer Raum.*

i) *Jede kompakte Teilmenge $A \subset X$ ist auch folgenkompakt.*

ii) *Jede folgenkompakte Teilmenge $K \subset X$ ist beschränkt und abgeschlossen.*

Beweis. i) Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge von Punkten aus A und $K = \{a_k \ : \ k \in \mathbb{N}\}$. Ist K endlich, so hat (a_k) eine konstante Teilfolge. Sei K also unendlich. Angenommen, K hat keinen Häufungspunkt in A. Dann besitzt jeder Punkt $x \in A$ eine offene Umgebung $U(x) \subset X$, die nur endlich viele Punkte aus K enthält. Die offenen Umgebungen $U(x)$ bilden eine offene Überdeckung von A. Da A kompakt ist, genügen bereits endlich viele $U(x_1), \ldots, U(x_n)$ zur Überdeckung, und A enthielte nur endlich viele Punkte aus K, Widerspruch. Sei $a \in A$ Häufungspunkt von K. Dann enthält jede offene Kugel um a mit Radius $\frac{1}{l+1}$ unendlich viele Punkte aus K. Setze $k_l := \min(k \ : d(a_k, a) < \frac{1}{l+1})$. Dann ist $(a_{k_l})_{l\in\mathbb{N}}$ eine Teilfolge von (a_k), die gegen a konvergiert.

ii) Angenommen, K wäre unbeschränkt. Dann gäbe es zu beliebigem $y \in K$ eine Folge $(x_k)_{k\in\mathbb{N}}$ mit $d(x_k, y) > k$, welche keine konvergente Teilfolge besitzt, Widerspruch. Wäre K nicht abgeschlossen, dann gäbe es eine konvergente Folge von Punkten aus K mit Grenzwert außerhalb K, Widerspruch. □

Satz 2.2.i) besagt also:

Satz 2.3 (Bolzano-Weierstraß) *Sei $A \subset X$ kompakte Teilmenge eines metrischen Raumes X und $\{y_k\}_{k\in\mathbb{N}}$ eine Folge von Punkten aus A. Dann gibt es einen Punkt $a \in A$ und eine Teilfolge $\{y_{k_l}\}_{l\in\mathbb{N}}$ von $\{y_k\}_{k\in\mathbb{N}}$, die gegen a konvergiert.* □

Eine wichtige Klasse folgenkompakter Mengen ist durch konvergente Folgen einschließlich ihres Grenzwertes gegeben. Wir zeigen, daß solche Mengen sogar kompakt sind:

Satz 2.4 *Sei X ein metrischer Raum und $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten aus X, die gegen einen Grenzwert $a \in X$ konvergiert. Dann ist die Teilmenge*

$$A := \{x_k \ : \ k \in \mathbb{N}\} \cup \{a\}$$

kompakt in X.

Beweis. Sei $(U_i)_{i\in I}$ eine beliebige offene Überdeckung von A. Dann gibt es einen Index $m \in I$ mit $a \in U_m$. Die Teilmenge $U_m \subset X$ ist offen, enthält also die offene Kugel $K_\epsilon(a)$ für ein $\epsilon > 0$. Da $(x_k)_{k\in\mathbb{N}}$ gegen a konvergiert, gibt es einen Index $n \in \mathbb{N}$, so daß $x_k \in K_\epsilon(a)$ für alle $k \geq n$. Jeder Punkt x_k mit $0 \leq k < n$ liegt in irgendeiner Umgebung U_{i_k} mit $i_k \in I$. Damit gilt

$$A \subset U_{i_0} \cup U_{i_1} \cup \cdots \cup U_{i_{n-1}} \cup U_m \ ,$$

so daß A kompakt ist. □

Ganz entscheidend im Beweis ist die Tatsache, daß der Grenzwert a zu A gehört: Es sei z.B. $A := \{\frac{1}{n+1} \ : \ n \in \mathbb{N}\} \subset \mathbb{R}$. Dann ist A nicht folgenkompakt (der Grenzwert gehört nicht zu A) und damit auch nicht kompakt.

Satz 2.5 *Jede konvergente Folge in einem metrischen Raum ist beschränkt.*

Sei $(x_k)_{k\in\mathbb{N}}$ konvergent mit $\lim_{k\to\infty} x_k = a$. Nach Satz 2.4 ist $\{x_k \ : \ k \in \mathbb{N}\} \cup \{a\}$ kompakt, damit nach Satz 2.2.ii) beschränkt. Die durch Weglassen von $\{a\}$ entstehende Menge bleibt beschränkt. □

Es läßt sich (mit größerem Schreibaufwand) zeigen, daß in beliebigen metrischen Räumen kompakt und folgenkompakt äquivalente Eigenschaften sind. In allgemeinen topologischen Räumen gilt zwar Satz 2.2.i), nicht jedoch die Umkehrung. Wir beschränken uns hier auf den Beweis, daß in endlich-dimensionalen normierten Räumen die Umkehrung gilt, indem wir zeigen:

Satz 2.6 *Für eine Teilmenge K eines endlich-dimensionalen normierten Vektorraums V sind folgende Aussagen äquivalent:*

i) *K ist beschränkt und abgeschlossen.*

ii) *K ist kompakt (Heine-Borel).*

iii) *K ist folgenkompakt (Bolzano-Weierstraß).*

Wegen Satz 2.2 ist nur i)⇒ii) zu zeigen. Wir formulieren zunächst einen Zwischenschritt als eigenen Satz:

Satz 2.7 *Sei X ein metrischer Raum, $Y \subset X$ eine kompakte Teilmenge und $A \subset Y$ eine abgeschlossene Teilmenge. Dann ist A kompakt.*

Beweis. $X \setminus A$ ist offen. Ist $(U_i)_{i\in I}$ eine offene Überdeckung von A, dann ist $Y \subset X = (X \setminus A) \cup \bigcup_{i\in I} U_i$. Da Y kompakt, genügen endlich viele U_i mit $X \setminus A$ zur Überdeckung von Y und damit auch von A. □

Beweis von Satz 2.2. Sei zunächst $V = \mathbb{R}^n$. Wegen Satz 2.7 genügt es zu zeigen, daß der abgeschlossene Quader

$$Q = [a_1, b_1] \times \cdots \times [a_n, b_n] = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \ : \ a_\nu \leq x_\nu \leq b_\nu\} \subset \mathbb{R}^n$$

kompakt ist, denn jede beschränkte abgeschlossene Teilmenge von $\mathbb{R}^n$ liegt in einem abgeschlossenen Quader.

Sei $(U_i)_{i\in I}$ eine unendliche offene Überdeckung von $Q_0 = Q$, die nicht auf eine endliche reduziert werden kann. Durch Halbierung aller Kanten zerlegen wir Q in 2^n gleich große Teilquader der halben Größe. Es gibt dann mindestens einen abgeschlossenen Teilquader, den wir mit Q_1 bezeichnen, der nicht durch endlich viele U_i überdeckt werden kann. Durch Wiederholung des Verfahrens finden wir eine Folge

$$Q = Q_0 \supset Q_1 \supset Q_2 \supset \dots$$

von abgeschlossenen Quadern mit $\operatorname{diam}(Q_k) = \frac{1}{2^k}\operatorname{diam}(Q)$, so daß jeder von ihnen nicht durch endlich viele U_i überdeckt werden kann.

Nach dem Intervallschachtelungsprinzip gibt es einen Punkt $x \in Q_k$ für alle k. Dieser Punkt x liegt in irgendeiner Umgebung U_j mit $j \in I$. Da U_j offen, gibt es ein $\epsilon > 0$, so daß $K_\epsilon(x) \subset U_i$. Dann finden wir aber auch ein $p \in \mathbb{N}$ mit $\operatorname{diam}(Q_p) = \frac{1}{2^p}\operatorname{diam}(Q) < \epsilon$. Somit gilt $Q_p \subset U_j$, d.h. die Quader lassen sich im Widerspruch zur Annahme durch endlich viele U_i überdecken. Also ist Q kompakt.

Sei nun V beliebiger n-dimensionaler reeller[1] Vektorraum und $\phi : \mathbb{R}^n \to V$ ein Isomorphismus (z.B. der kanonische Isomorphismus zu einer Basis). Nach Satz 1.4 ist ϕ ein Homöomorphismus. Jede Überdeckung von $K \subset V$ definiert eine Überdeckung von $\phi^{-1}(K) \subset \mathbb{R}^n$ und umgekehrt, damit ist K genau dann kompakt, wenn $\phi^{-1}(K)$ kompakt ist. Analog ist K genau dann folgenkompakt/abgeschlossen/beschränkt wenn $\phi^{-1}(K)$ folgenkompakt/abgeschlossen/beschränkt ist. (Für die Beschränktheit folgt das aus der Lipschitz-Stetigkeit). □

Daraus ergibt sich folgende Formulierung des Satzes von Bolzano-Weierstraß:

Satz 2.8 *Jede beschränkte Folge in einem endlich-dimensionalen normierten Vektorraum V besitzt eine konvergente Teilfolge.*

Beweis. Die Bildfolge unter dem Isomorphismus $\phi^{-1} : V \to \mathbb{R}^n$ kann in einen abgeschlossenen Quader eingebettet werden, welcher kompakt ist. Die Behauptung folgt dann aus Satz 2.3. □

Satz 2.9 *Es sei $(V^n, \|\ \|)$ ein n-dimensionaler normierter Vektorraum. Dann gilt: Die Vollkugel $B^n := \{x \in V^n \ : \ \|x - a\| \le r\} \subset V^n$ um $a \in V^n$ mit Radius r und die Sphäre $S^{n-1} := \{x \in V^n \ : \ \|x - a\| = r\} \subset V^n$ sind kompakt.*

[1] Für komplexe Vektorräume betrachte $\mathbb{R}^{2n} \xrightarrow{\psi} \mathbb{C}^n \xrightarrow{\phi} V$.

Beweis. Die Norm ist nach Satz 1.3 Lipschitz-stetig mit Lipschitz-Konstante 1, also sind B^n und S^{n-1} abgeschlossen nach Satz 1.10, außerdem beschränkt (klar) und damit kompakt nach Satz 2.6. □

Wir geben nun ein Beispiel dafür an, daß in allgemeinen metrischen Räumen die Umkehrung von Satz 2.2.ii) nicht gilt.

Beispiel 2.10 Wir betrachten den normierten (unendlich-dimensionalen) Vektorraum $\mathcal{C}([0,2\pi])$ der stetigen Funktionen auf $[0,2\pi]$ zusammen mit der Supremumsnorm $\|f\|_\infty = \sup_{x\in[0,2\pi]} |f(x)|$. Die abgeschlossene und beschränkte Einheitsvollkugel

$$B_1(0) := \{f \in \mathcal{C}([0,2\pi]) \ : \ \|f\|_\infty \leq 1\}$$

ist nicht folgenkompakt und damit auch nicht kompakt. Betrachte die Folge $(e_k)_{k\in\mathbb{N}}$ der Funktionen $e_k(x) = e^{\mathrm{i}kx}$. Es gilt $e_k \in B_1(0)$, aber wegen $\|e_k - e_l\|_\infty = \sup_{x\in[0,2\pi]} |2\sin\frac{(k-l)x}{2}| = 2$ für $k \neq l$ besitzt $(e_k)_{k\in\mathbb{N}}$ keine konvergente Teilfolge. ◁

Allgemein läßt sich zeigen, daß die Einheitsvollkugel in einem normierten Vektorraum V genau dann kompakt ist, wenn V endlich-dimensional ist.

3 Stetige Abbildungen kompakter Räume. Zwischenwertsatz

Satz 3.1 *Sei $f : X \to Y$ eine stetige Abbildung zwischen metrischen Räumen X, Y. Ist $A \subset X$ kompakt, dann ist auch $f(A) \subset Y$ kompakt.*

Beweis. Sei $(U_i)_{i\in\mathbb{N}}$ eine offene Überdeckung von $f(A)$. Aus der Stetigkeit von f folgt, daß $V_i := f^{-1}(U_i)$ offen ist. Dann ist $A \subset \bigcup_{i\in I} V_i$, aber tatsächlich genügen endlich viele V_i zur Überdeckung: $A \subset V_{i_1} \cup \cdots \cup V_{i_n}$, also $f(A) \subset U_{i_1} \cup \cdots \cup U_{i_n}$. □

Daraus ergibt sich der Satz vom Minimum/Maximum:

Satz 3.2 *Sei $A \subset X$ eine kompakte Teilmenge eines metrischen Raumes X und $f : X \to \mathbb{R}$ eine stetige Funktion. Dann ist die Einschränkung $f|_A$ auf A beschränkt (d.h. $|f(y)| < \infty$ für alle $y \in A$) und nimmt ihr Supremum und Infimum auf A an, d.h. es gibt $p, q \in A$ mit*

$$f(p) = \sup\{f(y) \ : \ y \in A\} \qquad \textit{und} \qquad f(q) = \inf\{f(y) \ : \ y \in A\} \ .$$

Beweis. Nach Satz 3.1 ist $f(A) \subset \mathbb{R}$ kompakt und nach Satz 2.2 beschränkt (und abgeschlossen), besitzt also ein Supremum M und Infimum m. Wegen der Abgeschlossenheit von $f(A)$ gilt $m, M \in f(A)$. □

Definition 3.3 Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f : X \to Y$ heißt *gleichmäßig stetig*, wenn zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert, so daß

$$d_Y\big(f(x_1), f(x_2)\big) < \epsilon \quad \text{für alle } x_1, x_2 \in X \text{ mit } d_X(x_1, x_2) < \delta \,.$$

Satz 3.4 *Seien X, Y metrische Räume und sei X kompakt. Dann ist jede stetige Abbildung $f : X \to Y$ auch gleichmäßig stetig.*

Beweis. Identisch mit Satz 14.11 aus dem 1. Semester, wenn $|\ |$ durch Abstand ersetzt wird. □

Definition 3.5 Ein metrischer Raum X heißt *zusammenhängend*, wenn es keine Zerlegung $X = U \cup V$ mit U, V offen und $U, V \neq \varnothing$ und $U \cap V = \varnothing$ gibt.

In $\mathbb{R}$ führt diese Definition auf Intervalle:

Satz 3.6 *Eine Teilmenge $X \subset \mathbb{R}$ mit mindestens zwei verschiedenen Punkten ist genau dann zusammenhängend, wenn X ein Intervall ist.*

Beweis. ($\Rightarrow$) Sei $X = I$ ein Intervall. Angenommen, $I = U \cup V$ mit U, V offen und $U, V \neq \varnothing$ und $U \cap V = \varnothing$. Dann gibt es Punkte $u \in U \subset I$ und $v \in V \subset I$ mit $u < v$ oder $u > v$. Sei $u < v$, dann ist $[u, v] \subset I$. Sei $s := \sup\{[u, v] \cap U\}$. Da $U = I \setminus V$ in I abgeschlossen ist, ist $s \in U$, also ist $]s, v] \subset V$. Andererseits ist U offen in I, enthält also auch eine Kugel $K_\epsilon(s)$, Widerspruch.

Sei umgekehrt X kein Intervall. Dann gibt es $u < s < v \in \mathbb{R}$ mit $u, v \in X$ und $s \notin X$. Also sind $U = X \cap]-\infty, s[$ und $V = X \cap]s, \infty[$ offen, disjunkt und nichtleer, außerdem ist $X = U \cup V$. Somit ist X nicht zusammenhängend. □

Satz 3.7 *Es sei $f : X \to Y$ eine stetige Abbildung zwischen metrischen Räumen X, Y. Ist X zusammenhängend, so ist auch $f(X)$ zusammenhängend.*

Beweis. Wäre $f(X)$ nicht zusammenhängend, so gäbe es disjunkte nichtleere offene Mengen U, V mit $f(X) = U \cup V$. Wegen der Stetigkeit von f sind $f^{-1}(U)$ und $f^{-1}(V)$ offen, nichtleer und disjunkt, denn $f(x)$ liegt entweder in U oder in V. Somit wäre $X = f^{-1}(U) \cup f^{-1}(V)$, Widerspruch. □

Satz 3.8 (Zwischenwertsatz) *Sei X ein zusammenhängender metrischer Raum, $a, b \in X$ und $f : X \to \mathbb{R}$ stetig. Dann nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an.*

Beweis. Nach Satz 3.7 ist $f(X)$ zusammenhängend. Ist $f(a) \neq f(b)$, dann ist $f(X)$ nach Satz 3.6 ein Intervall. Für $f(a) = f(b)$ ist nichts zu zeigen. □

Teil II
Differentialrechnung im $\mathbb{R}^n$

4 Partielle Ableitungen

Wir betrachten Funktionen mehrerer Veränderlicher, also Abbildungen $f : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^n$. Sei $x = (x_1, \ldots, x_n) \in U$. Partielle Ableitungen von f sind gewöhnliche Ableitungen, die man erhält, wenn alle Komponenten von x bis auf eine festgehalten werden.

Dazu betrachten wir folgende in U eingebettete Intervalle:

$$I_j = \{(x_1, \ldots, x_{j-1}, t, x_{j+1}, x_n) \in U \ : \ t \in \tilde{I}_j \subset \mathbb{R} \text{ mit } x_j \in \tilde{I}_j\} \ .$$

Die Einschränkung der Funktion $f : U \to \mathbb{R}$ auf I_j wird dann zu einer gewöhnlichen Funktion $f\big|_{I_j} : \tilde{I}_j \to \mathbb{R}$ einer Veränderlicher mit $f\big|_{I_j}(t) = f(x_1, \ldots, x_{j-1}, t, x_{j+1}, x_n)$. Für diese Funktion können wir die Differenzierbarkeit im Punkt x_j betrachten. Das Ergebnis ist die partielle Ableitung von f in der j-ten Koordinatenrichtung:

Definition 4.1 Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge. Eine Funktion $f : U \to \mathbb{R}$ heißt im Punkt $x \in U$ *partiell differenzierbar* in der j-ten Koordinatenrichtung, falls der Grenzwert

$$(\partial_j f)(x) := \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{1}{h}\big(f(x + he_j) - f(x)\big)$$

existiert. Dabei ist $e_j \in \mathbb{R}^n$ der j-te Einheitsvektor und h ist so zu wählen, daß $x + he_j \in U$. Der Grenzwert $(\partial_j f)(x)$ heißt die *j-te partielle Ableitung* von f in x.

Eine Funktion $f : U \to \mathbb{R}$ heißt *partiell differenzierbar*, falls $(\partial_j f)(x)$ für alle $x \in U$ und alle $1 \leq j \leq n$ existiert, und *stetig partiell differenzierbar*, falls alle Funktionen $\partial_j f : U \to \mathbb{R}$ stetig sind.

Oft schreibt man auch $\frac{\partial f}{\partial x_j}$ an Stelle von $\partial_j f$. Zur Existenz des Grenzwertes muß U nicht notwendig offen sein.

Die partielle Ableitung erfüllt die Leibniz-Regel: Seien $f, g : U \to \mathbb{R}$ partiell differenzierbare Funktionen auf U, dann gilt

$$(\partial_j(f \cdot g))(x) = (\partial_j f)(x) \cdot g(x) + f(x) \cdot (\partial_j g)(x) \ , \qquad x \in U \ .$$

Beispiel 4.2 Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x_1, x_2) = e^{x_1^2 + x_2^2} \sin x_1$. Dann ist $(\partial_1 f)(x_1, x_2) = (2x_1 \sin x_1 + \cos x_1)e^{x_1^2+x_2^2}$ und $(\partial_2 f)(x_1, x_2) = 2x_2 \sin x_1 e^{x_1^2+x_2^2}$. $\triangleleft$

Beispiel 4.3 Die partielle Ableitung des Radius $r(x) := \|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$ in der j-ten Koordinatenrichtung ist nach der Kettenregel

$$\frac{\partial r}{\partial x_j}(x_1, \ldots, x_n) = \frac{1}{2\sqrt{x_1^2 + \cdots + x_n^2}} \cdot 2x_j = \frac{x_j}{r} \ .$$

Damit ist $r : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ partiell differenzierbar mit $(\partial_j r)(x) = \frac{x_j}{r}$.

Entsprechend ist nach der Kettenregel jede differenzierbare Funktion $f(r)$ des Radius, aufgefaßt als Funktion $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, partiell differenzierbar mit $(\partial_i f)(x) = x_j \frac{f'(r)}{r}$. Zum Beispiel sind für $f(r) = \mathrm{e}^{-ar^2}$ die partiellen Ableitungen gegeben durch $(\partial_j f)(x) = -2ax_j f(r)$. ◁

Das folgende Beispiel zeigt, daß aus der partiellen Differenzierbarkeit einer Funktion nicht die Stetigkeit folgt.

Beispiel 4.4 Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

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

Dann ist f auf $\mathbb{R}^2 \setminus \{(0,0)\}$ partiell differenzierbar mit

$$(\partial_1 f)(x_1, x_2) = \frac{x_2}{x_1^2 + x_2^2} - \frac{2x_1^2 x_2}{(x_1^2 + x_2^2)^2} = \frac{x_2(x_2^2 - x_1^2)}{(x_1^2 + x_2^2)^2}$$

und analog

$$(\partial_2 f)(x_1, x_2) = \frac{x_1(x_1^2 - x_2^2)}{(x_1^2 + x_2^2)^2} \ .$$

Im Nullpunkt haben wir

$$(\partial_1 f)(0,0) = \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h} = 0 \ , \quad (\partial_1 f)(0,0) = \lim_{h \to 0} \frac{f(0,h) - f(0,0)}{h} = 0 \ ,$$

so daß f auf dem gesamten $\mathbb{R}^2$ partiell differenzierbar ist. Jedoch ist f nicht stetig in $(0,0)$. Die Folge $(y_k)_{k \in \mathbb{N}}$ mit $y_k = (\frac{1}{k+1}, \frac{1}{k+1})$ konvergiert gegen $(0,0)$, aber

$$f(y_k) = \frac{\frac{1}{k+1} \cdot \frac{1}{k+1}}{(\frac{1}{k+1})^2 + (\frac{1}{k+1})^2} = \frac{1}{2}$$

konvergiert nicht gegen $f(0) = 0$. ◁

Eine Verallgemeinerung der partiellen Ableitung ist die Richtungsableitung:

Definition 4.5 Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $v \in \mathbb{R}^n$ ein Vektor. Dann heißt der Differentialquotient

$$(D_v f)(x) := \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(x + tv) - f(x)}{t}$$

die *Richtungsableitung* der Funktion $f : U \to \mathbb{R}$ im Punkt x in Richtung v.

Insbesondere sind die partiellen Ableitungen die Richtungsableitungen in Richtung der Standardbasisvektoren, $(D_{e_i} f)(x) = (\partial_i f)(x)$. Damit folgt aus der Existenz aller Richtungsableitungen die partielle Differenzierbarkeit. Die Umkehrung gilt nicht. Wir werden jedoch im nächsten Abschnitt sehen, daß Stetigkeit der partiellen Ableitungen die Existenz aller Richtungsableitungen impliziert.

Die partielle Ableitung einer partiell differenzierbaren Funktion kann nochmals partiell differenziert werden, usw. Induktiv definieren wir:

Definition 4.6 Sei $U \subset \mathbb{R}^n$ offen und $k \in \mathbb{N} \setminus \{0\}$. Eine Funktion $f : U \to \mathbb{R}$ heißt $(k+1)$-*mal partiell differenzierbar*, wenn sie k-mal partiell differenzierbar ist und alle partiellen Ableitungen $\partial_{i_k} \dots \partial_{i_1} f$ partiell differenzierbar sind. Sind die partiellen Ableitungen $\partial_{i_{k+1}} \partial_{i_k} \dots \partial_{i_1} f$ stetig, so heißt f eine $(k+1)$-*mal stetig partiell differenzierbare Funktion*.

Satz 4.7 (Schwarz) *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Dann vertauschen die zweiten partiellen Ableitungen, d.h. für alle $a \in U$ und alle $i, j = 1, \dots, n$ gilt*

$$(\partial_i \partial_j f)(a) = (\partial_j \partial_i f)(a) \; .$$

Beweis. Der Übersichtlichkeit wegen sei $i = 1, j = 2$ (kann durch Umnumerieren der Koordinaten immer erreicht werden) und dann $n = 2$ (die weiteren Komponenten sind festgehalten und spielen keine Rolle).

Für gegebenes $\delta > 0$ sei $W_\delta(a) \subset U \subset \mathbb{R}^2$ der offene Würfel mit Kantenlänge 2δ und Mittelpunkt $a = (x_0, y_0)$. Es sei (x, y) ein beliebiger Punkt von W_δ, d.h. $|x - x_0| < \delta$ und $|y - y_0| < \delta$. Für festgehaltenes y sei $F_y(x) = f(x, y) - f(x, y_0)$. Nach dem Mittelwersatz der Differentialrechnung gibt es einen Punkt $\xi \in]x_0 - \delta, x_0 + \delta[$, so daß

$$\begin{aligned} F_y(x) - F_y(x_0) &= F_y'(\xi) \cdot (x - x_0) \\ \Rightarrow \quad f(x,y) - f(x,y_0) - f(x_0,y) + f(x_0,y_0) &= \big(\partial_1 f)(\xi, y) - (\partial_1 f)(\xi, y_0)\big) \cdot (x - x_0) \; . \end{aligned}$$

Wir nutzen den Mittelwertsatz nochmals für die Funktion $G_\xi(y) := (\partial_1 f)(\xi, y)$. Es gibt also ein $\eta \in]y_0 - \delta, y_0 + \delta[$, so daß

$$\begin{aligned} G_\xi(y) - G_\xi(y_0) &= G_\xi'(\eta) \cdot (y - y_0) \\ \Rightarrow \quad (\partial_1 f)(\xi, y) - (\partial_1 f)(\xi, y_0) &= (\partial_2 \partial_1 f)(\xi, \eta) \cdot (y - y_0) \; . \end{aligned}$$

Insgesamt gibt es somit ein $(\xi, \eta) \in W_\delta$ mit

$$f(x,y) - f(x,y_0) - f(x_0,y) + f(x_0,y_0) = (\partial_2 \partial_1 f)(\xi, \eta) \cdot (x - x_0)(y - y_0) \; .$$

Wir können aber auch erst x festhalten und den Mittelwertsatz in y anwenden, und als letztes den Mittelwertwsatz in x. Im Ergebnis gibt es einen neuen Punkt $(\tilde{\xi}, \tilde{\eta}) \in W_\delta$ mit

$$f(x,y) - f(x,y_0) - f(x_0,y) + f(x_0,y_0) = (\partial_1 \partial_2 f)(\tilde{\xi}, \tilde{\eta}) \cdot (x - x_0)(y - y_0) \; .$$

Somit gilt $(\partial_2\partial_1 f)(\xi,\eta) = (\partial_1\partial_2 f)(\tilde{\xi},\tilde{\eta})$. Lassen wir δ gegen 0 streben, so konvergieren (ξ,η) und $(\tilde{\xi},\tilde{\eta})$ gegen (x,y), und aus der vorausgesetzten Stetigkeit der zweiten partiellen Ableitungen folgt $(\partial_2\partial_1 f)(x,y) = (\partial_1\partial_2 f)(x,y)$. □

Entsprechend können bei k-mal stetig partiell differenzierbaren Funktionen die partiellen Ableitungen in beliebiger Reihenfolge geschrieben werden:

$$\partial_{i_1} \dots \partial_{i_k} f = \partial_{\pi(i_1)} \dots \partial_{\pi(i_k)} f$$

für eine beliebige Permutation π der Indizes $i_1, \dots, i_k$, denn jede Permutation läßt sich durch Vertauschen benachbarter Elemente darstellen. Es ist deshalb auch üblich, die mehrfachen partiellen Ableitungen zu schreiben als

$$\partial_{i_1} \dots \partial_{i_k} f = \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} \,.$$

5 Differenzierbarkeit

Definition 5.1 Sei $U \subset \mathbb{R}$ eine offene Teilmenge. Eine Abbildung $f : U \to \mathbb{R}^m$ heißt *total differenzierbar* (oder einfach nur *differenzierbar*) im Punkt $x \in U$, falls es eine lineare Abbildung $A(x) : \mathbb{R}^n \to \mathbb{R}^m$ und eine auf einer offenen Umgebung V von $0 \in \mathbb{R}^n$ definierte Abbildung $\phi : V \to \mathbb{R}^m$ gibt, so daß

$$f(x+\xi) = f(x) + A(x) \circ \xi + \phi(\xi) \quad \text{mit} \quad \lim_{\substack{\xi \to 0 \in \mathbb{R}^n \\ \xi \neq 0}} \frac{\phi(\xi)}{\|\xi\|} = 0 \qquad \text{für alle } \xi \in V \,.$$

Dann heißt die lineare Abbildung $(Df)(x) := A(x)$ das *totale Differential* (oder einfach nur das *Differential*) von f im Punkt x.

Einige Bemerkungen:

- Um den Restterm nicht ganz so mühsam zu charakterisieren, schreibt man einfach $o(\|\xi\|)$ und meint eine Abbildung $\phi : V \to \mathbb{R}^m$ mit obigen Eigenschaften.
- Da alle Normen auf endlich-dimensionalen Vektorräumen äquivalent sind, ist die Definition der Differenzierbarkeit unabhängig von der Wahl der Norm.
- Die lineare Abbildung $A(x) : \mathbb{R}^n \to \mathbb{R}^n$ ist bezüglich der Standardbasis durch eine $(m \times n)$-Matrix gegeben, die wir mit dem gleichen Buchstaben bezeichnen, $A(x) = (a_{ij}(x)) \in M(m \times n, \mathbb{R})$. Die Matrix $A = Df$ heißt *Jacobi-Matrix*. Variiert man den Punkt $x \in U$, so ist Df also durch $m \cdot n$ Funktionen $a_{ij} : U \to \mathbb{R}$ bestimmt.
- Oft schreibt man auch df oder f' für das totale Differential. Wir reservieren d für das später einzuführende äußere Differential.

Satz 5.2 *Sei $U \subset \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}^m$ im Punkt $x \in U$ differenzierbar mit $f(x+\xi) = f(x) + A(x) \cdot \xi + o(\|\xi\|)$ und $Df(x) = A(x) = (a_{ij}(x))$. Dann gilt:*

i) *f ist im Punkt $x \in U$ stetig.*

ii) *Alle Komponenten $f_i : U \to \mathbb{R}$ von $f = (f_1, \ldots, f_m)$ sind im Punkt x partiell differenzierbar mit $(\partial_j f_i)(x) = a_{ij}(x)$.*

iii) *f besitzt Richtungsableitungen in jede Richtung, und für $v = \sum_{j=1}^n e_j v_j$ gilt $(Df)(x) \circ v = (D_v f)(x) = \sum_{j=1}^n v_j (\partial_j f)(x)$.*

Beweis. i) Wegen $\lim_{\xi \to 0} A(x) \cdot \xi = 0$ und $\lim_{\xi \to 0} o(\|\xi\|) = 0$ gilt $\lim_{\xi \to 0} f(x+\xi) = f(x)$. Damit ist f stetig.

ii) Ist e_k der k-te Basisvektor der Standardbasis, dann ist $A(x) \circ e_j = \sum_{i=1}^m a_{ij}(x) e_i$, so daß für $\xi = h \cdot e_j$ und $f = \sum_{i=1} f_i e_i$ gilt

$$f_i(x + h \cdot e_j) = f_i(x) + a_{ij}(x) \cdot h + o(h) \qquad \Rightarrow \quad (\partial_j f_i)(x) = a_{ij}(x) \; .$$

iii) Es ist $(Df)(x) \circ v = A(x) \circ v = \sum_{i=1}^m \sum_{j=1}^n v_j a_{ij}(x) e_i$, also

$$f_i(x + h \cdot v) = f_i(x) + \sum_{j=1}^n h v_j a_{ij}(x) + o(|h| \|v\|) \; . \qquad \square$$

Damit bietet sich folgende Strategie zur Überprüfung der Differenzierbarkeit einer Abbildung $f : U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$, an: Man bilde, falls existent, die Matrix $A(x) = (a_{ij}(x))$ der partiellen Ableitungen $a_{ij}(x) = \partial_j f_i$. Die Abbildung f ist genau dann differenzierbar, wenn $f(x+\xi) - f(x) - A(x) \cdot \xi = o(\|\xi\|)$ für alle $\xi \in V$.

Beispiel 5.3 Es sei $f(x,y) = \begin{pmatrix} x \sin y \\ y^2 e^x \end{pmatrix}$. Es ergibt sich

$$\begin{aligned} (\partial_1 f_1)(x,y) &= \sin y \; , & (\partial_2 f_1)(x,y) &= x \cos y \; , \\ (\partial_1 f_2)(x,y) &= y^2 e^x \; , & (\partial_2 f_2)(x,y) &= 2ye^x \; , \end{aligned}$$

also $A(x,y) = \begin{pmatrix} \sin y & x \cos y \\ y^2 e^x & 2ye^x \end{pmatrix}$. Wir betrachten

$$\begin{aligned} & f(x+\xi, y+\eta) - f(x,y) - A \cdot \begin{pmatrix} \xi \\ \eta \end{pmatrix} \\ &= \begin{pmatrix} (x+\xi)\sin(y+\eta) - x \sin y - \xi \sin y - x \cos y \eta \\ (y+\eta)^2 e^{x+\xi} - y^2 e^x - y^2 e^x \xi - 2y\eta e^x \end{pmatrix} \\ &= \begin{pmatrix} -2(x+\xi) \sin y \sin^2 \frac{\eta}{2} + \xi \cos y \sin \eta + x \cos y (\sin \eta - \eta) \\ y^2 e^x (e^\xi - 1 - \xi) + 2y\eta e^x (e^\xi - 1) + \eta^2 e^{x+\xi} \end{pmatrix} = o(\sqrt{\xi^2 + \eta^2}) \; , \end{aligned}$$

da sämtliche Einträge der Matrix für $\sqrt{\xi^2 + \eta^2} \to 0$ mindestens quadratisch gegen 0 gehen. Damit ist f differenzierbar mit $(Df)(x) = A$. $\square$

Ein hinreichendes Kriterium für Differenzierbarkeit ist stetige partielle Differenzierbarkeit:

Satz 5.4 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ partiell differenzierbar auf U. Sind alle partiellen Ableitungen $\partial_j f$ stetig im Punkt $x \in U$, dann ist f im Punkt x total differenzierbar.*

Beweis. Da U offen, gibt es ein $\delta > 0$ mit $K_\delta(x) \subset U$. Wir wählen ein $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ mit $\|\xi\| < \delta$ und betrachten die Punkte $z^{(k)} := x + \sum_{j=1}^{k} \xi_j e_j$. Es gilt $z^{(0)} = x$ und $z^{(n)} = x + \xi$. Da sich benachbarte $z^{(k-1)}$ und $z^{(k)}$ nur in der k-ten Koordinate unterscheiden, können wir den Mittelwertsatz der Differentialrechnung anwenden: Es gibt also ein $\eta^{(k)} \in \mathbb{R}$ mit $|\eta^{(k)}| < \xi_k$, so daß

$$f(z^{(k)}) - f(z^{(k-1)}) = \xi_k \cdot (\partial_k f)(y^{(k)}) \,, \qquad y^{(k)} := z^{(k-1)} + \eta^{(k)} e_k \,.$$

Das bedeutet

$$\begin{aligned} f(x+\xi) &= f(x) + \sum_{k=1}^{n} \xi_k \cdot (\partial_k f)(y^{(k)}) \\ &= f(x) + \sum_{k=1}^{n} (\partial_k f)(x) \cdot \xi_k + \underbrace{\sum_{k=1}^{n} \big((\partial_k f)(y^{(k)}) - (\partial_k f)(x) \big) \cdot \xi_k}_{\phi(\xi)} \,. \end{aligned}$$

Für $\xi \to 0$ strebt y_k gegen x. Aus der Stetigkeit der partiellen Ableitungen folgt $\lim\limits_{\xi \to 0} (\partial_k f)(y^{(k)}) = (\partial_k f)(x)$ und damit $\lim\limits_{\substack{\xi \to 0 \in \mathbb{R}^n \\ \xi \neq 0}} \dfrac{\phi(\xi)}{\|\xi\|} = 0.$ □

Durch Kombination der Sätze 5.2 und 5.4 folgt, daß jede stetig partiell differenzierbare Funktion $f : U \to \mathbb{R}$ auf U auch stetig ist. Außerdem gelten folgende Implikationen:

$$\begin{aligned} & f : U \to \mathbb{R} \text{ stetig partiell differenzierbar} \\ \Rightarrow \quad & f : U \to \mathbb{R} \text{ total differenzierbar} \\ \Rightarrow \quad & f : U \to \mathbb{R} \text{ besitzt Richtungsableitungen in jede Richtung} \\ \Rightarrow \quad & f : U \to \mathbb{R} \text{ partiell differenzierbar} \end{aligned}$$

Die Umkehrungen gelten im allgemeinen nicht.

Es gelten die üblichen Linearitäts- und Produktregeln:

Satz 5.5 *Es sei $U \subset \mathbb{R}^n$ offen und $f_1, f_2 : U \to \mathbb{R}^m$ sowie $f, g : U \to \mathbb{R}$ differenzierbar in $x \in U$. Dann gilt:*

i) *$f_1 + f_2$ ist differenzierbar in x mit $(D(f_1 + f_2))(x) = Df_1(x) + Df_2(x)$,*

ii) *$f \cdot g$ ist differenzierabr in x mit $(D(f \cdot g))(x) = (Df)(x) \cdot g(x) + f(x) \cdot Dg(x)$,*

iii) *Ist $f(x) \neq 0$, dann ist $\frac{1}{f}$ differenzierbar in x mit $(D\frac{1}{f})(x) = -\frac{Df(x)}{(f(x))^2}$.*

Beweis. Ähnlich zu Satz 16.6. aus dem ersten Semester. i) ist klar. Zu ii):

$$\begin{aligned}(f \cdot g)(x+\xi) - (f \cdot g)(x) &\\ &= \big(f(x+\xi) - f(x)\big) \cdot g(x+\xi) + f(x) \cdot \big(g(x+\xi) - g(x)\big) \\ &= \big((Df)(x) \cdot \xi + o(\|\xi)\big) \cdot g(x+\xi) + f(x) \cdot \big(Dg(x) \cdot \xi + o(\|\xi\|)\big) \\ &= (Df)(x) \cdot g(x) + f(x) \cdot Dg(x) + o(\|\xi\|) \; .\end{aligned}$$

Analog ergibt sich iii). □

Satz 5.6 (Kettenregel) *Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offene Mengen sowie $f : U \to \mathbb{R}^m$ und $g : V \to \mathbb{R}^k$ Abbildungen mit $f(U) \subset V$. Die Abbildung f sei im Punkt $x \in U$ differenzierbar, und g sei im Punkt $f(x) \in V$ differenzierbar. Dann ist die Abbildung $g \circ f : U \to \mathbb{R}^k$ im Punkt $x \in U$ differenzierbar, und es gilt*

$$(D(g \circ f))(x) = (Dg)(f(x)) \circ (Df)(x) \; .$$

Beweis. Nach Voraussetzung gilt

$$f(x+\xi) = f(x) + (Df)(x) \cdot \xi + o(\|\xi\|) \; , \qquad g(y+\eta) = g(y) + (Dg)(y) \cdot \eta + o(\|\eta\|) \; .$$

Wir berechnen

$$\begin{aligned}(g \circ f)(x+\xi) = g(f(x+\xi)) &= g\Big(\underbrace{f(x)}_{y} + \underbrace{(Df)(x) \cdot \xi + o(\|\xi\|)}_{\eta} \Big) \\ &= g(f(x)) + (Dg)(f(x)) \cdot \big((Df)(x) \cdot \xi + o(\|\xi\|)\big) \\ &\quad + o\big(\big\|(Df)(x) \cdot \xi + o(\|\xi\|)\big\|\big) \\ \text{Linearität} \Rightarrow \quad &= (g \circ f)(x) + \big((Dg)(f(x)) \cdot (Df)(x)\big) \cdot \xi \\ &\quad + (Dg)(f(x)) \cdot o(\|\xi\|) + o\big(\big\|(Df)(x) \cdot \xi + o(\|\xi\|)\big\|\big) \; .\end{aligned}$$

Da die letzte Zeile wieder $o(\|\xi\|)$ ist, folgt die Behauptung. □

Zu beachten ist, daß $(Dg)(f(x)) \circ (Df)(x)$ die Komposition linearer Abbildungen bzw. die Multiplikation der entspechenden Jacobi-Matrizen ist. Die Reihenfolge von $(Dg)(f(x))$ und $(Df)(x)$ darf nicht geändert werden! Ein wichtiger Spezialfall ist die Ableitung längs einer Kurve.

Definition 5.7 Unter einer stetigen/differenzierbaren/stetig differenzierbaren *Kurve* im $\mathbb{R}^n$ versteht man eine stetige/differenzierbare/stetig differenzierbare Abbildung

$c : I \to \mathbb{R}^n$ (wobei $I \subset \mathbb{R}^n$ aus mehr als einem Punkt besteht). Ist die Kurve $c = (c_1, \ldots, c_n) : I \to \mathbb{R}^n$ differenzierbar in t, dann heißt

$$c'(t) := (Dc)(t) = \big(c_1'(t), \ldots, c_n'(t)\big) \in \mathbb{R}^n$$

der *Tangentialvektor* an c im Punkt $c(t)$.

Satz 5.8 *Es sei $c = (c_1, \ldots, c_n) : I \to U$ eine in $t \in I$ differenzierbare Kurve und $f : U \to \mathbb{R}^n$ eine in $x = c(t)$ differenzierbare Abbildung. Dann ist $f \circ c : I \to \mathbb{R}^n$ differenzierbar in t mit $D(f \circ c)(t) = (Df)(c(t)) \cdot c'(t)$.* □

Beispiel 5.9 Es sei $c : R^1 \to R^3$, mit $c(t) = (ht; \sin t; \cos t)$ die Schraubenkurve und $f : R^3 \to R^1$ gegeben durch $f(x; y; z) := x^2 + y^2 - z^2$. Gesucht ist $(D(f \circ c))(t)$. Wir drücken die Differentiale durch partielle Ableitungen aus:

$$c'(t) = \begin{pmatrix} h \\ \cos t \\ -\sin t \end{pmatrix} =: A \,,$$

$$Df(x, y, z) = \begin{pmatrix} 2x & 2y & -2z \end{pmatrix} \,, \qquad Df(c(t)) = \begin{pmatrix} 2ht & 2\sin t & -2\cos t \end{pmatrix} =: B \,,$$

also

$$(D(f \circ c))(t) = B \cdot A = 2h^2 t + 4 \sin t \cos t \,. \qquad \triangleleft$$

Über eine differenzierbare Kurve läßt sich der Mittelwertsatz auf den $\mathbb{R}^n$ verallgemeinern:

Satz 5.10 (Mittelwertsatz) *Es seien $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine differenzierbare Funktion. Für zwei Punkte $a, b \in U$ gelte $c(t) := a + (b - a)t \in U$ für alle $t \in [0, 1]$. Dann gibt es einen Zwischenpunkt $\xi = a + (b - a)t_0$, mit $t_0 \in]0, 1[$, so daß $f(b) - f(a) = (Df)(\xi) \circ (b - a)$.*

Beweis. Für $F = f \circ c$ mit der differenzierbaren Kurve $t : [0, 1] \to U$ gilt nach Kettenregel $(DF)(t) = (Df)(c(t)) \cdot c'(t)$. Nach dem Mittelwertsatz der Differentialrechnung gibt es ein $t_0 \in]0, 1[$, so daß für $F(1) = f(b)$, $F(0) = f(a)$ gilt $f(b) - f(a) = F(1) - F(0) = (DF)(t_0) = (Df)(\xi) \cdot (b - a)$, mit $\xi = a + (b - a)t_0$. □

Für $f : U \to \mathbb{R}^m$ gilt für jede Komponente ein Mittelwertsatz, jedoch wird die Ableitung im allgemeinen an verschiedenen Zwischenpunkten $\xi_1, \ldots, \xi_n$ zu nehmen sein. Deshalb wird eine Dastellung als vektorwertiges Integral sinnvoll:

Satz 5.11 (Integralform des Mittelwertsatzes) *Es seien $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ eine differenzierbare Abbildung und $c : [\alpha, \beta] \to U$ eine differenzierbare Kurve mit Randpunkten $c(\alpha) = a$ und $c(\beta) = b$. Dann gilt*

$$f(b) - f(a) = \int_\alpha^\beta dt \; (Df)(c(t)) \circ c'(t) \,.$$

Für die Kurve $c(t) := a + (b-a)t \in U$ gilt unter gleichen Voraussetzungen

$$f(b) - f(a) = \int_0^1 dt\ (Df)(a + (b-a)t) \circ (b-a)\ .$$

Beweis. Nach dem Hauptsatz der Differential- und Integralrechnung ist $F_i(\beta) - F_i(\alpha) = \int_\alpha^\beta dt\ F_i'(t)$. Setzt man $F = f \circ c$, also $F_i = f_i \circ c$, so folgt die Behauptung aus der Kettenregel $(DF)(t) = (Df)(c(t)) \circ c'(t)$. □

Ist $f : U \to \mathbb{R}$ sogar stetig differenzierbar in U, so ist $(Df)(a+(b-a)t)$ stetig. Wir werden später beweisen, daß dann auch das Integral $g(x) = \int_0^1 dt\ (Df)(a + (x-a)t)$ eine stetige Funktion $g : U \to \mathbb{R}$ liefert. Also gilt in Komponentenform

$$f(x) - f(a) = \sum_{i=1}^m g_i(x)(x_i - a_i)$$

für alle $x \in K_r(a) \subset U$ und stetige Funktionen $g : K_r(a) \to \mathbb{R}$. Insbesondere ist $(\partial_i f)(a) = g_i(a)$.

Für den folgenden Schrankensatz benötigen wir die Abschätzung

$$\Big\| \int_a^b dt\ f(t) \Big\| \le \int_a^b dt\ \|f(t)\|$$

für eine vektorwertige Funktion $f : [a,b] \to \mathbb{R}^m$. Dabei ist $\|\ \|$ die aus dem Skalarprodukt erhaltene Norm. Sei $v = \int_a^b dt\ f(t) \in \mathbb{R}^m$, dann gilt

$$\begin{aligned} \|v\|^2 = \langle v, v\rangle = \Big\langle \int_a^b dt\ f(t), v \Big\rangle &= \int_a^b dt\ \langle f(t), v\rangle \\ \text{(Cauchy-Schwarz)} \quad &\le \int_a^b dt\ \|f(t)\|\ \|v\| = \|v\| \int_a^b dt\ \|f(t)\|\ . \end{aligned}$$

Satz 5.12 (Schrankensatz) *Es seien $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Mit $\|\ \|$ werden die aus dem Standardskalarprodukt erhaltenen 2-Normen auf $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ bezeichnet und mit $\|\ \|_{op}$ die Operator-Norm einer linearen Abbildung. Sei $c(t) := a + (b-a)t \in U$ für alle $t \in [0,1]$. Dann gilt $\|f(b) - f(a)\| \le M\|b-a\|$ mit $M = \sup_{t\in[0,1]} \|(Df)(a + (b-a)t)\|_{op}$.*

Beweis. Nach der zweiten Version des Mittelwertsatzes 5.11, obiger Abschätzung und Definition der Operatornorm gilt

$$\begin{aligned} \|f(b) - f(a)\| &= \Big\| \int_0^1 dt\ (Df)(a + (b-a)t) \circ (b-a) \Big\| \\ &\le \int_0^1 dt\ \big\|(Df)(a + (b-a)t) \circ (b-a)\| \\ &\le \int_0^1 dt\ \big\|(Df)(a + (b-a)t)\|_{op} \|b-a\| \le M\|b-a\| \int_0^1 dt\ 1\ . \end{aligned}$$

Die letzte Ungleichung ist Satz 19.11.i) aus dem 1. Semester. □

6 Vektorfelder, Gradient, Divergenz, Rotation, Laplace

Definition 6.1 Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine partiell differenzierbare Funktion. Dann heißt der Vektor

$$(\operatorname{grad} f)(x) = \Big(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\Big) \in \mathbb{R}^n$$

der *Gradient* von f im Punkt $x \in U$.

Satz 6.2 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ differenzierbar. Dann gilt für jeden Punkt $x \in U$ und jeden Vektor $v \in \mathbb{R}^n$*

$$(D_v f)(x) = \langle v, (\operatorname{grad} f)(x)\rangle \ .$$

Beweis. Nach Satz 5.2.iii) ist $(D_v f)(x) = \sum_{j=1}^n v_j(\partial_j f)(x) = \langle v, (\operatorname{grad} f)(x)\rangle$. □

Damit ist der Gradient in Richtung des steilsten Anstiegs der Funktion gerichtet, und die Norm $\|(\operatorname{grad} f)(x)\|$ ist ein Maß für die Stärke des Anstiegs: Ist $\operatorname{grad} f \neq 0$ und ist θ der Winkel zwischen einem Vektor $v \in \mathbb{R}^n$ mit $\|v\| = 1$ und $\operatorname{grad} f$, dann gilt für das Skalarprodukt

$$(D_v f)(x) = \|(\operatorname{grad} f)(x)\| \cos\theta \ .$$

Die Richtungsableitung ist maximal, wenn v in Richtung $\operatorname{grad} f$ zeigt.

Beispiel 6.3 Für $f(x) = r$ gilt $\operatorname{grad} r = \frac{x}{r}$, der steilste Anstieg ist also radial nach außen gerichtet und vom Betrag her überall (außer im Nullpunkt) konstant.

Für $f(x) = \frac{1}{r}$ ist $\operatorname{grad} \frac{1}{r} = -\frac{x}{r^3}$. Der steilste Anstieg ist radial nach innen gerichtet und wächst zum Nullpunkt quadratisch. ◁

Der Gradient ist linear und erfüllt die Leibniz-Regel: Sind $f, g : U \to \mathbb{R}$ partiell differenzierbar, so gilt

$$\begin{aligned}(\operatorname{grad}(f+g))(x) &= (\operatorname{grad} f)(x) + (\operatorname{grad} g)(x) \ , \\ (\operatorname{grad}(f\cdot g))(x) &= (\operatorname{grad} f)(x)\cdot g(x) + f(x)\cdot(\operatorname{grad} g)(x) \ .\end{aligned}$$

Der Gradient $(\operatorname{grad} f)(x)$ ordnet jedem Punkt $x \in U$ einen Vektor zu. So etwas nennt man ein Vektorfeld:

Definition 6.4 Unter einem *Vektorfeld* auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ versteht man eine Abbildung $v : U \to \mathbb{R}^n$. Das Vektorfeld heißt stetig/partiell differenzierbar/..., wenn alle Komponenten von v stetig/partiell differenzierbar/... sind.

Die Vorstellung ist, daß an jedem Punkt $x = (x_1, \ldots, x_n) \in U$ ein Vektor $v(x) = (v_1(x), \ldots, v_n(x))$ angeheftet ist. Eine nützliche Konstruktion besteht darin, diese Vektoren als Tangentialvektoren an Kurven $f : I \to \mathbb{R}^n$ durch $x = f(t_0)$ zu betrachten. Ist $f(t) = (f_1(t), \ldots, f_n(t))$, dann ist also

$$\frac{df_i(t)}{dt} = v_i(f(t)) \ , \quad f_i(t_0) = x_i \qquad \text{für } i = 1, \ldots, n \text{ und } t, t_0 \in I$$

Die Kurve selbst ist dann durch Lösen (Integration) dieser $2n$ Gleichungen zu erhalten und heißt *Integralkurve* des Vektorfeldes durch x. In der Physik werden die Bilder der Integralkurven auch *Feldlinien* bzw. *Stromlinien* des Vektorfelds genannt. Unter recht schwachen Voraussetzungen an das Vektorfeld (Lipschitzstetig) kann man in einer genügend kleinen Umgebung von x die Integralkurven immer finden, aber nicht unbedingt auf ganz U, weil das an Singularitäten des Vektorfeldes scheitern kann. Eine Art von Singularität ist die Divergenz:

Definition 6.5 Sei $v = (v_1, \ldots, v_n) : U \to \mathbb{R}^n$ ein partiell differenzierbares Vektorfeld auf einer offenen Teilmenge $U \subset \mathbb{R}^n$. Dann heißt die Funktion

$$\operatorname{div}(v) := \frac{\partial v_1}{\partial x_1} + \cdots + \frac{\partial v_n}{\partial x_n}$$

die *Divergenz* des Vektorfeldes v.

Die Divergenz ist linear, $\operatorname{div}(v + w) = \operatorname{div}(v) + \operatorname{div}(w)$ und erfüllt folgendes Analogon zur Leibniz-Regel: Sei v ein partiell differenzierbares Vektorfeld und f eine partiell differenzierbare Funktion auf $U \subset \mathbb{R}^n$, dann gilt

$$\operatorname{div}(f \cdot v) = \langle \operatorname{grad}(f), v \rangle + f \cdot \operatorname{div}(v) \ .$$

Die Divergenz eines Vektorfeldes ist ein Maß für die Gesamtbeschleunigung der Integralkurven. Das kann einerseits dadurch erreicht werden, daß die Feldlinien von der Parallelität abweichen (was wahrscheinlich den Namen motiviert hat), oder durch Geschwindigkeitszunahme entlang der Feldlinien.

Beispiel 6.6 Für $v(x) = x \in \mathbb{R}^n$ gilt $\operatorname{div} x = n$. Hier sind die Feldlinien radial vom Nullpunkt nach außen gerichtet und gerade durch den Radius parametrisiert.

Für $v(x) = \frac{x}{\|x\|}$ gilt mit $\|x\| = r$

$$\operatorname{div}\frac{x}{r} = \left\langle \operatorname{grad}\frac{1}{r}, x \right\rangle + \frac{1}{r} \cdot \operatorname{div} x = -\frac{1}{r^3}\langle x, x\rangle + \frac{n}{r} = \frac{n-1}{r} \qquad \text{für } x \neq 0 \ .$$

Die Vektoren $v(x) = \frac{x}{\|x\|}$ haben in jedem Punkt $x \in \mathbb{R}^n \setminus \{0\}$ die Länge 1, aber die Feldlinien sind wieder radial nach außen gerichtet und damit nicht parallel.◁

Gradient und Divergenz sind wichtige Hilfsmittel der Theoretischen Physik. In der Mechanik läßt sich ein konservatives Kraftfeld als (negativer) Gradient eines Potentials schreiben, $F = -\operatorname{grad} V$. In der Elektrodynamik wird die Divergenz von Vektorfeldern zur Formulierung der Maxwellschen Gleichungen gebraucht. Dort gibt es noch eine weitere Konstruktion mit partiellen Ableitungen, die *Rotation*. Die Rotation und das *Vektorprodukt* beruhen auf einer nur im $\mathbb{R}^3$ möglichen Identifikation von sogenannten *Differentialformen*, deren Ideen wir im folgenden vorstellen.

Definition 6.7 Sei V ein reeller Vektorraum und $\{e_1, \ldots, e_n\}$ eine Basis von V. Dann heißt

$$\Lambda^k(V) := \begin{cases} \mathbb{R} & \text{für } k = 0 \\ \operatorname{span}\big(e_{i_1} \wedge \cdots \wedge e_{i_k} \;:\; 1 \le i_1 < i_2 < \cdots < i_k \le n\big) & \text{für } 1 \le k \le n \\ 0 & \text{für } k > n \end{cases}$$

der *Vektorraum der k-fach antisymmetrischen Tensoren über V*. Dabei ist die Verknüpfung $\wedge$ als formelles Symbol anzusehen.

Sei $U \subset \mathbb{R}^n$ offen. Eine stetige/partiell differenzierbare/... Abbildung $\omega : U \to \Lambda^k(\mathbb{R}^n)$ heißt *stetige/partiell differenzierbare/... k-Form* auf U. Der Raum der k-Formen auf U wird mit $\Omega^k(U)$ bezeichnet.

Eine k-Form ist also die Zuordnung eines k-fach antisymmetrischen Tensors zu jedem Punkt von U. Insbesondere können 1-Formen auf U mit Vektorfeldern auf U identifiziert werden, und 0-Formen auf U sind nichts anderes als Funktionen auf U. Im allgemeinen nimmt man an, daß die Formen beliebig oft differenzierbar sind; sie heißen dann auch Differentialformen. Ist $\dim(V) = n$, so gilt $\dim(\Lambda^k(V)) = \binom{n}{k}$.

Definition 6.8 Sei $\{e_1, \ldots, e_n\}$ eine Basis von V. Die durch die Vereinbarung $e_i \wedge e_j = -e_j \wedge e_i$ für alle $i, j = 1, \ldots, n$ und lineare Fortsetzung entstehende assoziative Multiplikation

$$\wedge : \Lambda^k(V) \times \Lambda^l(V) \to \Lambda^{k+l}(V)$$

heißt das *äußere Tensorprodukt* (auch *wedge-Produkt* wegen der englischen Bezeichnung "wedge" (Keil) für $\wedge$). Entsprechend heißt

$$\wedge : \Omega^k(U) \times \Omega^l(U) \to \Omega^{k+l}(U)$$

das *äußere Formenprodukt*.

Beispiel 6.9 Sei $n \ge 3$ und $\omega_1 = f \cdot e_2$ und $\omega_2 = g \cdot e_1 \wedge e_3$ für Funktionen $f, g : U \to \mathbb{R}$, d.h. $\omega_1(x) = f(x)\, e_2$ und $\omega_2(x) = g(x)\, e_1 \wedge e_3$ für $x \in U$. Dann gilt

$$\omega_1 \wedge \omega_2 = (f \cdot g) \cdot (e_2 \wedge e_1 \wedge e_3) = -(f \cdot g) \cdot (e_1 \wedge e_2 \wedge e_3)\,. \qquad \triangleleft$$

Allgemein gilt für $\omega_k \in \Omega_k(U)$ und $\tilde{\omega}_l \in \Omega_l(U)$ die graduierte Kommutativität

$$\omega_k \wedge \tilde{\omega}_l = (-1)^{kl} \tilde{\omega}_l \wedge \omega_k \ .$$

Definition 6.10 Sei $(e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{R}^n$. Die durch

$$d\omega := \sum_{i=1}^{n} e_i \wedge \partial_i(\omega)$$

für eine partiell differenzierbare k-Form $\omega \in \Omega^k(U)$ definierte Abbildung $d : \Omega^k(U) \to \Omega^{k+1}(U)$ heißt *äußere Ableitung* oder *äußeres Differential*.

Beispiel 6.11 Sei $n = 3$ und $\omega = f_1 \cdot (e_2 \wedge e_3) + f_2 \cdot (e_1 \wedge e_3)$, dann ist wegen $e_i \wedge e_i = 0$

$$d\omega = (\partial_1 f_1) \cdot (e_1 \wedge e_2 \wedge e_3) + (\partial_2 f_2) \cdot e_2 \wedge (e_1 \wedge e_3) = (\partial_1 f_1 - \partial_2 f_2) \cdot (e_1 \wedge e_2 \wedge e_3) \ .$$

□

Satz 6.12

i) *Ist $f : U \to \mathbb{R}$ eine partiell differenzierbare Funktion, dann gilt $df = \mathrm{grad}(f)$.*

ii) *Sind $\omega_k \in \Omega^k(U)$ und $\tilde{\omega}_l \in \Omega^l(U)$ partiell differenzierbare Formen, dann gilt die* graduierte Leibniz-Regel

$$d(\omega_k \wedge \tilde{\omega}_l) = (d\omega_k) \wedge \tilde{\omega}_l + (-1)^k \omega_k \wedge d\tilde{\omega}_l \ .$$

iii) *Ist $\omega \in \Omega^k(U)$ zweimal stetig partiell differenzierbar, dann gilt $d(d\omega) = 0$.*

Beweis. i) ist klar. ii) folgt aus der Leibniz-Regel für die partielle Ableitung und der graduierte Kommutativität des $\wedge$-Produkts. iii) folgt aus dem Satz von Schwarz über die Vertauschbarkeit der zweiten partiellen Ableitungen zusammen mit der Antisymmetrie $e_i \wedge e_j = -e_j \wedge e_i$. □

Definition 6.13 Sei $(e_1, \ldots, e_n)$ die Standardbasis von $\mathbb{R}^n$. Die durch

$$*(e_{i_1} \wedge \cdots \wedge e_{i_k}) := \omega \quad \text{falls } \omega \wedge (e_{i_1} \wedge \cdots \wedge e_{i_k}) = e_1 \wedge e_2 \wedge \cdots \wedge e_n$$

und lineare Fortsetzung definierte Abbildung $* : \Omega^k(U) \to \Omega^{n-k}(U)$ heißt die *Hodge*-Abbildung.

Ist z.B. $n = 3$, so gilt $*(e_1 \wedge e_3) = -e_2$, denn $(-e_2) \wedge (e_1 \wedge e_3) = e_1 \wedge e_2 \wedge e_3$. Die Hodge-Abbildung ist ein Isomorphismus von Vektorräumen, denn $\dim(\Lambda^k(\mathbb{R}^n)) = \binom{n}{k} = \binom{n}{n-k} = \dim(\Lambda^{n-k}(\mathbb{R}^n))$. Das Inverse der Hodge-Abbildung ist, bis auf ein mögliches Vorzeichen, die Hodge-Abbildung selbst: $*(*\omega_k) = (-1)^{k(n-k)}\omega_k$ für $\omega_k \in \Omega^k(U)$ und $U \subset \mathbb{R}^n$.

Definition 6.14 Sei $U \subset \mathbb{R}^n$ offen.

i) Die durch $\delta\omega_k := -(-1)^{nk} * (d(*\omega_k))$ für eine partiell differenzierbare k-Form $\omega_k \in \Omega^k(U)$ definierte Abbildung $\delta : \Omega^k(U) \to \Omega^{k-1}(U)$ heißt *Kodifferential*.

ii) Sei $v \in \Omega^1(U)$ ein partiell differenzierbares Vektorfeld. Dann heißt die Funktion
$$\operatorname{div}(v) := *(d(*v)) = (-1)^{n+1}\delta v \in \Omega^0(U)$$
die *Divergenz von* v.

iii) Seien $v, w \in \Omega^1(U)$ Vektorfelder auf U. Dann heißt das Vektorfeld
$$v \times w := *(v \wedge w) \in \Omega^1(U)$$
das *Vektorprodukt* von v, w.

iv) Sei $v \in \Omega^1(U)$ ein partiell differenzierbares Vektorfeld auf U. Dann heißt das Vektorfeld
$$\operatorname{rot}(v) := *(dv) \in \Omega^1(U)$$
die *Rotation* von v.

Wir rechnen nach, das unsere Definition des Vektorprodukts mit der üblichen Vorschrift zusammenfällt. Sei $v = v_1e_1 + v_2e_2 + v_3e_3$ und $w = w_1e_1 + w_2e_2 + w_3e_3$, dann ist

$$\begin{array}{lllllll} v \wedge w & = v_1w_2\, e_1 \wedge e_2 & + v_1w_3\, e_1 \wedge e_3 & - v_2w_1\, e_1 \wedge e_2 & + v_2w_3\, e_2 \wedge e_3 & - v_3w_1\, e_1 \wedge e_3 & - v_3w_2\, e_2 \wedge e_3 \\ *(v \wedge w) & = v_1w_2\, e_3 & - v_1w_3\, e_2 & - v_2w_1\, e_3 & + v_2w_3\, e_1 & + v_3w_1\, e_2 & - v_3w_2\, e_1 \end{array}$$

also tatsächlich

$$(v \times w)_1 = v_2w_3 - v_3w_2\,, \quad (v \times w)_2 = v_3w_1 - v_1w_3\,, \quad (v \times w)_3 = v_1w_2 - v_2w_1\,.$$

Entsprechend rechnen wir die Rotation aus: Sei $v = v_1e_1 + v_2e_2 + v_3e_3$, wobei v_i nun Funktionen sind, dann gilt

$$\begin{array}{lllllll} dv & = \frac{\partial v_2}{\partial x_1}\, e_1 \wedge e_2 & + \frac{\partial v_3}{\partial x_1}\, e_1 \wedge e_3 & - \frac{\partial v_1}{\partial x_2}\, e_1 \wedge e_2 & + \frac{\partial v_3}{\partial x_2}\, e_2 \wedge e_3 & - \frac{\partial v_1}{\partial x_3}\, e_1 \wedge e_3 & - \frac{\partial v_2}{\partial x_3}\, e_2 \wedge e_3 \\ *(dv) & = \frac{\partial v_2}{\partial x_1}\, e_3 & - \frac{\partial v_3}{\partial x_1}\, e_2 & - \frac{\partial v_1}{\partial x_2}\, e_3 & + \frac{\partial v_3}{\partial x_2}\, e_1 & + \frac{\partial v_1}{\partial x_3}\, e_2 & - \frac{\partial v_2}{\partial x_3}\, e_1 \end{array}$$

also tatsächlich

$$(\operatorname{rot} v)_1 = \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}\,, \quad (\operatorname{rot} v)_2 = \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}\,, \quad (\operatorname{rot} v)_3 = \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\,.$$

Satz 6.15 *Sei $U \subset \mathbb{R}^3$ offen, $f : U \to \mathbb{R}$ eine zweifach stetig partiell differenzierbare Funktion und $v : U \to \mathbb{R}^3$ ein zweifach stetig partiell differenzierbares Vektorfeld. Dann gilt*

$$\operatorname{rot}(\operatorname{grad} f) = *(d(df)) = 0\,, \qquad \operatorname{div}(\operatorname{rot} v) = *(d(*(*(dv)))) = 0\,.$$

Ein wichtige Rolle innerhalb der Mathematik spielen Fragestellungen der folgenden Art (übersetzt in physikalische Begriffe):

i) Gegeben sei ein rotationsfreies Vektorfeld $v : U \to \mathbb{R}^3$ mit $U \subset \mathbb{R}^3$ und $\mathrm{rot}(v) = 0$. Unter welchen Voraussetzungen an U existiert eine Funktion $f : U \to \mathbb{R}$, so daß $v = \mathrm{grad}(f)$ gilt?

ii) Gegeben sei ein divergenzfreies Vektorfeld $v : U \to \mathbb{R}^3$ mit $U \subset \mathbb{R}^3$ und $\mathrm{div}(v) = 0$. Unter welchen Voraussetzungen an U existiert ein Vektorfeld $w : U \to \mathbb{R}^n$, so daß $v = \mathrm{rot}(w)$ gilt?

In der Formensprache scheiben sich die Maxwellschen Gleichungen wie folgt:

$$\begin{aligned} dB &= 0 & \delta B + \dot{E} &= -j \\ dE + \dot{B} &= 0 & \delta E &= \rho \end{aligned}$$

Dabei bezeichnet der Punkt über einer Form die Zeitableitung, und die folgenden Formen treten auf:

- $E(t) \in \Omega^1(U)$ ist das elektromagnetische Feld
- $B(t) \in \Omega^2(U)$ ist das Magnetfeld (die Physik arbeitet mit $B_{\mathrm{Physik}} = *B \in \Omega^1(U)$)
- $j(t) \in \Omega^1(U)$ ist die Stromdichte
- $\rho(t) \in \Omega^0(U)$ ist die Ladungsdichte

Wegen $\delta\delta = \pm * d * *d* = 0$ gilt die Kontinuitätsgleichung $\delta j + \dot{\rho} = 0$.

Die links stehenden Gleichungen kann man unter geeigneten Voraussetzungen an U durch $B = dA$ und $E = -d\phi - \dot{A}$ lösen mit $A \in \Omega^1(U)$ und $\phi \in \Omega^0(U)$. Unter der Voraussetzung $\delta A = -\dot{\phi}$ ergibt sich für die rechten Gleichungen

$$\ddot{A} - (\delta d + d\delta)A = j \qquad \ddot{\phi} - \delta d\phi = \rho \; .$$

Definition 6.16 Der durch $\Delta\omega := \delta(d\omega) + d(\delta\omega)$ für $\omega \in \Omega^k(U)$ definierte Operator $\Delta = \delta d + d\delta : \Omega^k(U) \to \Omega^k(U)$ heißt *Laplace-Beltrami-Operator*. Im Spezialfall $k = 0$ heißt $\Delta f = \delta(df) = \mathrm{div}(\mathrm{grad} f)$ der *Laplace-Operator* für eine zweifach partiell differenzierbare Funktion $f : U \to \mathbb{R}$.

Für den Laplace-(Beltrami-)Operator gibt es die Darstellung

$$\Delta\omega = \frac{\partial^2\omega}{\partial x_1^2} + \cdots + \frac{\partial^2\omega}{\partial x_n^2} \; , \qquad \omega \in \Omega^k(U) \; , \; U \subset \mathbb{R}^n \; ,$$

was zumindest für Funktionen $\omega = f \in \Omega^0(U)$ leicht zu überprüfen ist.

Folgende Differentialgleichungen enthalten den Lapace-Operator:

$\ddot{f} - \Delta f = \rho$	inhomogene Wellengleichung
$\ddot{f} - \Delta f = 0$	homogene Wellengleichung
$\Delta f = -\rho$	Poisson-Gleichung
$\Delta f = 0$	Potentialgleichung
$c\dot{f} - \Delta f = 0$	Wärmeleitungsgleichung

Die Maxwellschen Gleichungen im Vakuum ($j = 0$ und $\rho = 0$) führen also auf insgesamt vier skalare Wellengleichungen (elektromagnetische Wellen). Zeitunabhängige elektrische Felder werden durch die Poisson-Gleichung beschrieben. Gibt es zusätzlich keine Ladungsdichte, dann ist die Potentialgleichung zu lösen.

Wichtig ist noch zu zeigen, daß man die Bedingung $\delta A + \dot{\phi} = 0$, die uns die Differentialgleichungen für A und ϕ entkoppelt hat, wirklich lösen kann. Dabei ist entscheidend, daß A und ϕ durch E und B nicht eindeutig festgelegt sind; es gibt eine sogenannte Eichfreiheit: Setzt man $A \mapsto A + df$ und $\phi \mapsto -\dot{f}$ für $f \in \Omega^0(U)$, so bleiben E, B unverändert. Die Separierungsgleichung geht aber über in

$$\delta A + \dot{\phi} \mapsto \delta A + \dot{\phi} - (\ddot{f} - \Delta f) \; .$$

Man bestimmt also zunächst A, ϕ als Lösung der inhomogenen Wellengleichungen und berechnet daraus E und B. Die Eichbedingung $\delta A + \dot{\phi} = 0$ wird zunächst verletzt sein, wir müssen aber nur wissen, daß die Eichung f so gewählt werden kann, daß $\delta A + \dot{\phi} = 0$ zu erfüllen ist. Dann ist die Herleitung der Wellengleichungen für die eichtransformierten Potentiale A, ϕ korrekt. Da aber E und B durch f nicht geändert werden, bestimmen sich die Feldstärken E, B aus Lösungen A, ϕ der Wellengleichungen auch ohne Berücksichtigung der Eichbedingung.

7 Die Taylor-Formel

Um die im folgenden auftretenden vielen Indizes übersichtlicher zu gestalten, hat sich eine abkürzende Schreibweise eingebürgert. Sei $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ ein Multiindex, d.h. ein n-Tupel von Indizes $\alpha_i \in \mathbb{N}$. Dann setzt man

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n \; , \qquad \alpha! := \alpha_1! \, \alpha_2! \cdots \alpha_n! \; .$$

Für eine $|\alpha|$-mal stetig partiell differenzierare Funktion $f : U \to \mathbb{R}$ schreibt sich eine mehrfache partielle Ableitung wie folgt:

$$(\partial^\alpha f)(x) := (\partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n} f)(x) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}(x) \; .$$

Dabei ist $\partial_i^{\alpha_i} f = \underbrace{\partial_i \ldots \partial_i}_{\alpha_i \text{mal}} f$. Ebenso setzt man $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

Satz 7.1 *Sei $U \subset \mathbb{R}$ offen und sei $f : U \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion. Zu $x \in U$ sei ein Vektor $\xi \in \mathbb{R}^n$ so gewählt, daß $x + t\xi \in U$ für alle $t \in [0, 1]$. Dann ist die Funktion einer Veränderlichen*

$$g : [0, 1] \mapsto \mathbb{R} , \qquad g(t) := f(x + t\xi)$$

k-mal stetig differenzierbar auf $[0, 1]$, und es gilt

$$\frac{d^k g}{dt^k}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \xi^\alpha \, (\partial^\alpha f)(x + t\xi) .$$

Dabei läuft die Summe über alle Multiindizes α mit $|\alpha| = k$.

Beweis. Es gilt $g = f \circ c$ mit $c(t) = x + t\xi$. Nach der Kettenregel ist

$$g'(t) = (Df)(c(t)) \circ c'(t) = (Df)(x + t\xi) \circ \xi = \sum_{i=1}^{n} \xi_i (\partial_i f)(x + t\xi) .$$

Durch Wiederholung der Rechnung für jede der Funktionen $(\partial_i f)(x + t\xi) = g_i \circ c(t)$, u.s.w., berechnen wir die höheren Ableitungen, wobei wir zunächst nicht den Satz von Schwarz verwenden, d.h. wir betrachten $\partial_i \partial_j$ und $\partial_j \partial_i$ als verschieden. Es ergibt sich

$$\frac{d^k g}{dt^k}(t) = \sum_{i_1, i_2, \ldots, i_k = 1}^{n} \xi_{i_k} \cdots \xi_{i_2} \xi_{i_1} (\partial_{i_k} \ldots \partial_{i_2} \partial_{i_1} f)(x + t\xi) .$$

Nach dem Satz von Schwarz können wir die partiellen Ableitungen ordnen und zu ∂^α zusammenfassen mit $|\alpha| = k$. Ebenso fassen sich die Produkte der ξ_i zu ξ^α zusammen. Es gibt $\frac{k!}{\alpha_1! \cdots \alpha_n!}$ verschiedene $\partial_{i_k} \ldots \partial_{i_2} \partial_{i_1}$, die nach Umordnung das gleiche ∂^α ergeben. □

Satz 7.2 (Taylor) *Sei $U \subset \mathbb{R}$ offen und sei $f : U \to \mathbb{R}$ eine $(k + 1)$-mal stetig differenzierbare Funktion. Zu $x \in U$ sei ein Vektor $\xi \in \mathbb{R}^n$ so gewählt, daß $x + t\xi \in U$ für alle $t \in [0, 1]$. Dann gibt es ein $\theta \in [0, 1]$, so daß*

$$f(x + \xi) = \sum_{|\alpha| \le k} \frac{\xi^\alpha}{\alpha!} (\partial^\alpha f)(x) + \sum_{|\alpha| = k+1} \frac{\xi^\alpha}{\alpha!} (\partial^\alpha f)(x + \theta\xi) .$$

Beweis. Wir betrachten die Funktion einer Veränderlichen $g : [0, 1] \to \mathbb{R}$, die durch $g[t] := f(x + t\xi)$ gegeben ist. Nach Satz 7.1 ist g eine $(k + 1)$ mal stetig differenzierbare Funktion, auf die wir den eindimensionalen Satz von Taylor (Satz 24.1 aus dem 1. Semester) mit Lagrangescher Form des Restgliedes (Satz 24.2) anwenden können: Es gibt also ein $\theta \in [0, 1]$ mit

$$g(1) = \sum_{j=0}^{k} \frac{1}{j!} \frac{d^j g}{dt^j}(0) + \frac{1}{(k + 1)!} \frac{d^{k+1} g}{dt^{k+1}}(\theta) .$$

Einsetzen von $g(t) = f(x + t\xi)$ und Verwenden von Satz 7.1 liefert die Behauptung. □

Der Satz von Taylor ist wichtig bei Abschätzungen der folgenden Art:

Satz 7.3 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine k-mal stetig differenzierbare Funktion. Dann gilt für jedes $x \in U$*

$$f(x+\xi) = \sum_{|\alpha|\le k} \frac{\xi^\alpha}{\alpha!}(\partial^\alpha f)(x) + o(\|\xi\|^k) \qquad \textit{für } \xi \to 0 \; .$$

Beweis. Da U offen, gibt es ein $\delta > 0$, so daß $K_\delta(x) \subset U$. Dann gibt es zu jedem $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \delta$ ein $\theta \in [0,1]$ mit

$$\begin{aligned} f(x+\xi) &= \sum_{|\alpha|\le k-1} \frac{\xi^\alpha}{\alpha!}(\partial^\alpha f)(x) + \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!}(\partial^\alpha f)(x+\theta\xi) \\ &= \sum_{|\alpha|\le k} \frac{\xi^\alpha}{\alpha!}(\partial^\alpha f)(x) + \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!}\big((\partial^\alpha f)(x+\theta\xi) - (\partial^\alpha f)(x)\big) \; . \end{aligned}$$

Es gilt $|\xi^\alpha| = |\xi_1|^{\alpha_1} \cdots |\xi_n|^{\alpha_n} \le \|\xi\|^{\alpha_1} \cdots \|\xi\|^{\alpha_n} = \|\xi\|^{|\alpha|}$. Da $\partial^\alpha f$ stetig ist, gilt

$$\lim_{\xi\to 0} \frac{1}{\|\xi\|^k} \sum_{|\alpha|=k} \frac{\xi^\alpha}{\alpha!}\big((\partial^\alpha f)(x+\theta\xi) - (\partial^\alpha f)(x)\big) = 0 \; .$$

Das ist genau die Behauptung. □

Für $k = 1$ ist die Menge aller Multiindizes α mit $|\alpha| = 1$ gerade die Menge der Standardbasisvektoren (e_i). Somit erhalten wir die Formel für das totale Differential

$$f(x+\xi) = f(x) + \langle \xi, (\mathrm{grad} f)(x)\rangle + o(\|\xi\|) \; .$$

Für $k = 2$ ergibt sich

$$f(x+\xi) = f(x) + \langle \xi, (\mathrm{grad} f)(x)\rangle + \frac{1}{2}\sum_{i,j=1}^n \xi_i\xi_j(\partial_i\partial_j f)(x) + o(\|\xi\|^2) \; .$$

Dabei gibt es die Möglichkeiten $\alpha = (0,\ldots,0,2,0,\ldots,0)$ mit $\alpha! = 2$ und $\alpha = (0,\ldots,0,1,0,\ldots,0,1,0,\ldots,0)$ mit $\alpha! = 1$, wobei die Einsen an der i-ten und j-ten Stelle stehen mit $i < j$. Da eine in i,j symmetrische Funktion summiert wird, kann die Summe mit $i < j$ durch die halbe Summe mit $i \ne j$ ersetzt werden. Zusammen mit der Summe über $i = j$ von $\alpha = (0,\ldots,0,2,0,\ldots,0)$ ergibt sich die Beziehung.

Definition 7.4 Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine zweifach stetig differenzierbare Funktion. Dann heißt die symmetrische $n \times n$-Matrix

$$(\operatorname{Hess} f)(x) := \big((\partial_i \partial_j f)(x)\big)_{1 \le i,j \le n}$$

die *Hessesche Matrix* von f im Punkt x.

Somit gilt

$$f(x+\xi) = f(x) + \langle a, \xi \rangle + \frac{1}{2}\langle \xi, A\xi \rangle + o(\|\xi\|^2)$$
$$\text{mit } a = (\operatorname{grad} f)(x)\ ,\ A = (\operatorname{Hess} f)(x)\ .$$

Die Hessesche Matrix ist wichtig bei der Untersuchung von lokalen Extrema einer Funktion.

Definition 7.5 Sei $U \subset \mathbb{R}^n$ offen. Ein Punkt $x \in U$ heißt *lokales Maximum* bzw. *lokales Minimum* einer Funktion $f : U \to \mathbb{R}$, wenn es eine Umgebung $V \subset U$ von x gibt, so daß

$$f(x) \ge f(y) \quad \text{bzw.} \quad f(x) \le f(y) \qquad \text{für alle } y \in V\ .$$

Gilt in diesem Fall $f(x) = f(y) \quad \Leftrightarrow \quad x = y$, so heißt das lokale Maximum (bzw. Minimum) *strikt*. Ein *lokales Extremum* ist ein lokales Minimum oder lokales Maximum.

Der folgende Satz liefert eine notwendige Bedingung für ein lokales Extremum:

Satz 7.6 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ partiell differenzierbar. Besitzt f in $x \in U$ ein lokales Extremum, so gilt* $(\operatorname{grad} f)(x) = 0$.

Beweis. Es genügt, die n Funktionen einer Veränderlichen $g_i(t) := f(x + te_i)$ zu betrachten, wobei e_i der i-te Standardbasisvektor ist und $t \in [-\epsilon, \epsilon]$. Hat f ein lokales Extremum in x, so hat g_i ein lokales Extremum in 0. Dann gilt $0 = g_i'(0) = (\partial_i f)(x)$. Da das für jedes $1 \le i \le n$ gilt, folgt die Behauptung. □

Wir beweisen eine hinreichende Bedingung für lokale Extrema unter Verwendung der Hesseschen Matrix. Symmetrische Matrizen bezüglich der Standard-Orthonormalbasis sind diagonalisierbar. Je nach Vorzeichen der Eigenwerte definiert man (Definition 23.13 aus dem 2. Semester):

Definition 7.7 Eine symmetrische Matrix $A = A^t \in M(n, \mathbb{R})$ heißt

- *positiv definit*, falls $\langle \xi, A\xi \rangle > 0 \quad \forall \xi \in \mathbb{R}^n \setminus \{0\}$,
- *positiv semidefinit*, falls $\langle \xi, A\xi \rangle \ge 0 \quad \forall \xi \in \mathbb{R}^n \setminus \{0\}$,

- *negativ (semi)definit*, falls $-A$ positiv (semi)definit ist,
- *indefinit*, falls es $\xi, \eta \in \mathbb{R}^n$ gibt mit $\langle \xi, A\xi \rangle > 0$ und $\langle \eta, A\eta \rangle < 0$.

Für praktische Bestimmungen der Definitheit ist das Determinantenkriterium (Satz 23.15 aus dem 2. Semester)) hilfreich.

Satz 7.8 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig differenzierbar. In einem Punkt $x \in U$ gelte* $(\operatorname{grad} f)(x) = 0$.

i) *Ist* $(\operatorname{Hess} f)(x)$ *positiv definit, so besitzt f in x ein striktes lokales Minimum.*

ii) *Ist* $(\operatorname{Hess} f)(x)$ *negativ definit, so besitzt f in x ein striktes lokales Maximum.*

iii) *Ist* $(\operatorname{Hess} f)(x)$ *indefinit, so besitzt f in x kein lokales Extremum.*

Beweis. Zur Vereinfachung der Schreibweise sei $A := (\operatorname{Hess} f)(x)$. In einer Umgebung V von x gilt

$$f(x+\xi) = f(x) + \frac{1}{2}\langle \xi, A\xi \rangle + \phi(\xi) \qquad \text{mit } \phi(\xi) = o(\|\xi\|^2) \; .$$

Es gibt also zu jedem $\epsilon > 0$ ein $\delta > 0$, so daß $|\phi(\xi)| < \epsilon \|\xi\|^2$ für alle $\xi \in \mathbb{R}^n$ mit $\|\xi\| < \delta$.

i) Die $(n-1)$-Sphäre $S^{n-1} := \{\eta \in \mathbb{R}^n \; : \; \|\eta\| = 1\}$ ist kompakt. Nach Satz 3.2 nimmt die stetige Funktion $g : S^{n-1} \to \mathbb{R}$ mit $g(\eta) := \langle \eta, A\eta \rangle$ auf S^{n-1} ihr Supremum und ihr Infimum an. Es gibt also ein $\mu > 0$ (wegen der positiven Definitheit) mit

$$\mu := \min_{\eta \in S^{n-1}} \{\langle \eta, A\eta \rangle\} \; .$$

Da für beliebiges $\xi \in \mathbb{R}^n \setminus \{0\}$ gilt $\frac{1}{\|\xi\|}\xi \in S^{n-1}$, folgt $\langle \xi, A\xi \rangle \geq \mu \|\xi\|^2$ für alle $\xi \in \mathbb{R}^n$ (für $\xi = 0$ trivialerweise). Wählen wir δ so klein, daß $|\phi(\xi)| \leq \frac{\mu}{4}\|\xi\|^2$, so gilt $\frac{1}{2}\langle \xi, A\xi \rangle + \phi(\xi) \geq \frac{\mu}{4}\|\xi\|^2$. Folglich haben wir

$$f(x+\xi) \geq f(x) + \frac{\mu}{4}\|\xi\|^2 \qquad \text{für alle } \xi \in \mathbb{R}^n \text{ mit } \|\xi\| < \delta \; .$$

Also hat f in x ein striktes lokales Minimum. Analog beweist man ii).

iii) Wir zeigen, daß es $y_1, y_2 \in U$ gibt mit $f(y_1) < f(x) < f(y_2)$. Nach Voraussetzung gibt es ein $\xi \in \mathbb{R}^n$ mit $\langle \xi, A\xi \rangle := \mu > 0$. Dann gilt für hinreichend kleine $t > 0$

$$f(x+t\xi) = f(x) + \frac{\mu}{2}t^2 + \phi(t\xi) \; .$$

Wie zuvor finden wir ein $\delta > 0$, so daß $|\phi(t\xi)| \leq \frac{\mu}{4}t^2$ gilt für alle $0 < t < \delta$. Dann ist $f(x+t\xi) > f(x)$ für alle $0 < t < \delta$. Ebenso folgt aus der Existenz eines $\eta \in \mathbb{R}^n$ mit $\langle \eta, A\eta \rangle := -\mu' < 0$, daß $f(x+t'\eta) < f(x)$ für alle $0 < t' < \delta'$. $\square$

Ist die Hessesche Matrix im Punkt x positiv oder negativ semidefinit, so muß man höhere Ordnungen in der Taylorsche Formel betrachten, um Aussagen über Extrema von f mit $(\operatorname{grad} f)(x) = 0$ zu gewinnen.

Beispiel 7.9 i) $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x, y) = 1 + x^2 + y^2$.
Es gilt $(\operatorname{grad} f)(x, y) = (2x, 2y)$, folglich kann f nur in $(0, 0)$ ein lokales Extremum haben. Wir testen $(\operatorname{Hess} f)(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. Da $(\operatorname{Hess} f)(0, 0)$ positiv definit, hat f im Punkt $(0, 0)$ ein lokales Minimum.

ii) $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x, y) = 1 + x^2 - y^2$.
Es gilt $(\operatorname{grad} f)(x, y) = (2x, -2y)$, folglich kann f nur in $(0, 0)$ ein lokales Extremum haben. Wir testen $(\operatorname{Hess} f)(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. Damit ist $(\operatorname{Hess} f)(0, 0)$ indefinit, so daß f im Punkt $(0, 0)$ kein Extremum hat.

Wir beweisen noch das Maximumprinzip für *harmonische Funktionen* $f : U \to \mathbb{R}$, d.h. $\Delta f = 0$. Dazu benötigen wir

Satz 7.10 (Rotationsinvarianz des Laplace-Operators) *Für jede Orthonormalbasis $(v_1, \ldots, v_n)$ im $\mathbb{R}^n$ gilt $\Delta f = \sum_{i=1}^n D_{v_i} D_{v_i} f$.*

Beweis. Es sei $f : U \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Es sei $v_i = \sum_{j=1}^n v_{ij} e_j$ die Zerlegung nach der Standardbasis $(e_i)_{i=1,\ldots,n}$. Aus der Orthonormalität folgt (siehe Satz 23.12 aus dem 2. Semester) $\sum_{i=1}^n v_{ij} v_{ik} = \delta_{jk}$ und $\sum_{j=1}^n v_{ij} v_{lj} = \delta_{li}$. Nach Satz 5.2.iii) ist $(D_{v_i} f)(x) = \sum_{j=1}^n v_{ij} (\partial_j f)(x)$. Nochmalige Richtungsableitung ergibt $(D_{v_i}(\partial_j f))(x) = \sum_{k=1}^n v_{ik} (\partial_k \partial_j f)(x)$, also

$$\sum_{i=1}^n (D_{v_i} D_{v_i} f)(x) = \sum_{i=1}^n \Big(\sum_{j,k=1}^n v_{ik} v_{ij} (\partial_k \partial_j f)(x) \Big) = \sum_{j,k=1}^n \delta_{jk} (\partial_k \partial_j f)(x) = (\Delta f)(x) \; .$$

□

Satz 7.11 (Maximumprinzip für harmonische Funktionen) *Es sei $U \subset \mathbb{R}^n$ eine beschränkte offene Menge, $\bar{U} := U \cup \partial U$ ihr Abschluß und $f : \bar{U} \to \mathbb{R}$ eine stetige Funktion, die in U harmonisch sei, d.h. $(\Delta f)(x) = 0$ für alle $x \in U$. Dann nimmt f ihr Maximum und Minimum auf dem Rand ∂U an.*

Beweis. Mit f erfüllt auch $-f$ die Voraussetzungen, so daß der Beweis für das Maximum genügt. Die Mengen $\bar{U}$ und ∂U sind kompakt, also nimmt die stetige Funktion f ihr Maximum M auf $\bar{U}$ und m auf ∂U an (Satz 3.2). Angenommen, $m < M$. Dann gibt es ein $\epsilon > 0$, so daß für $f_\epsilon(x) = f(x) + \epsilon \|x\|^2$ gilt $\max_{y \in \partial U} f_\epsilon(y) < M$. Damit wird das Maximum von f_ϵ in $a \in U$ angenommen: $f_\epsilon(a) = \max_{x \in U} f_\epsilon(x) \geq M$. Damit muß $(\operatorname{grad} f)(a) = 0$ und $(\operatorname{Hess} f)(a)$ negativ semidefinit sein (positiv definit und indefinit sind durch Satz 7.8 ausgeschlossen). Nach Satz 23.12 aus dem 2. Semester gibt es eine Orthonormalbasis $(v_i)_{i=1,\ldots,n}$

aus Eigenvektoren von $(\mathrm{Hess}\, f)(a)$. Für die entsprechenden Richtungsableitungen gilt

$$(D_{v_i} D_{v_j} f_\epsilon)(a) = \lambda_i \delta_{ij} \qquad \text{mit } \lambda_i \leq 0 \,.$$

Nach Satz 7.10 ist $(\Delta f_\epsilon)(a) = \sum_{i=1}^n (D_{v_i} D_{v_i} f_\epsilon)(a) = \lambda_1 + \cdots + \lambda_n \leq 0$. Andererseits folgt aus der Harmonizität $\Delta f_\epsilon = \epsilon \,\mathrm{div}(\mathrm{grad}(\|x\|^2)) = 2n\epsilon > 0$, Widerspruch. Also ist $m \geq M$. □

Das Maximumprinzip wird wichtig z.B. beim *Dirichlet-Problem*: Es sei $U \subset \mathbb{R}^n$ offen und beschränkt und $h : \partial U \to \mathbb{R}$ sei stetig. Gesucht ist eine stetige Funktion $f : \bar{U} \to \mathbb{R}$, die in U zweimal stetig differenzierbar ist und $\Delta f = 0$ auf U und $f|_{\partial U} = h$ erfüllt. Nach dem Maximumprinzip kann es höchstens eine Lösung geben, denn die Differenz von zwei Lösungen ist konstant $= 0$ auf dem Rand und damit $= 0$ auf ganz $\bar{U}$.

8 Der Satz über implizite Funktionen

Es geht nun um Funktionen, die implizit definiert sind, z.B. durch Gleichungen der Form $0 = F(x, f(x)) = (f(x))^2 + x^2 - 1$. Wir werden untersuchen, unter welchen Bedingungen sich derartige Gleichungen zumindest im Prinzip nach $f(x)$ auflösen lassen und welche Differenzierbarkeitseigenschaften die Lösungen haben. Im obigen Beispiel ist offenbar $f(x) = \pm\sqrt{1 - x^2}$. Differentiation von $F(x, f(x))$ nach der Kettenregel liefert

$$\begin{aligned} 0 = F'(x, f(x)) &= (\partial_1 F)(x, f(x)) + (\partial_2 F)(x, f(x)) f'(x) \\ \Rightarrow \quad f'(x) &= -\frac{(\partial_1 F)(x, f(x))}{(\partial_2 F)(x, f(x))} = -\frac{x}{f(x)} \,. \end{aligned}$$

Wir werden die Lösung des impliziten Problems iterativ konstruieren. Dazu benötigen wir:

Satz 8.1 (Banachscher Fixpunktsatz) *Sei $A \subset X$ eine abgeschlossene Teilmenge eines Banachraums X (d.h. eines vollständigen normierten Vektorraums $(X, \| \ \|)$). Die Abbildung $\Phi : A \to A$ sei eine* Kontraktion*, d.h. es gibt eine Konstante $\theta \in]0, 1[$, so daß*

$$\|\Phi(f) - \Phi(g)\| \leq \theta \|f - g\| \qquad \textit{für alle } f, g \in A \,.$$

Dann gilt:

i) *Φ besitzt genau einen Fixpunkt f_*, d.h. es gibt ein eindeutig bestimmtes $f_* \in A$ mit $\Phi(f_*) = f_*$.*

ii) *Für einen belieigen Anfangspunkt $g \in A$ konvergiert die durch*

$$f_0 := g \,, \quad f_{k+1} = \Phi(f_k) \textit{ für } k \in \mathbb{N}$$

definierte Folge $(f_k)_{k \in \mathbb{N}}$ gegen f_*, d.h. $\lim_{k \to \infty} f_k = f_*$.

Beweis. 1) Wir beweisen zunächst die Eindeutigkeit des Fixpunktes. Gäbe es zwei Fixpunkte $f_*, g*$, dann ist

$$\|f_* - g_*\| = \|\Phi(f_*) - \Phi(g_*)\| \leq \theta \|f_* - g_*\| \ ,$$

und damit $\|f_* - g_*\| = 0$ wegen $0 < \theta < 1$. Das bedeutet $f_* = g_*$.

2) Sei $(f_k)_{k\in\mathbb{N}}$ wie oben definiert. Dann ist

$$\|f_{k+1} - f_k\| = \|\Phi(f_k - \Phi(f_{k-1})\| \leq \theta \|f_k - f_{k-1}\| \leq \cdots \leq \theta^k \|f_1 - f_0\| \ .$$

Für $m > l$ betrachten wir

$$\begin{aligned} \|f_m - f_l\| &= \Big\| \sum_{k=l}^{m-1} (f_{k+1} - f_k) \Big\| \leq \sum_{k=l}^{m-1} \|f_{k+1} - f_k\| \leq \sum_{k=l}^{m-1} \theta^k \|f_1 - f_0\| \\ &= \frac{(1-\theta^m) - (1-\theta^l)}{1-\theta} \|f_1 - f_0\| \leq \theta^l \frac{\|f_1 - f_0\|}{1-\theta} \ . \end{aligned}$$

Damit ist $(f_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge, die wegen der Vollständigkeit des Banach-Raums gegen einen Punkt f_* konvergiert. Da A abgeschlossen, liegt der Grenzwert f_* sogar in A. Aus $f_* = \lim_{k\to\infty} f_k$ und $f_{k+1} = \Phi(f_k)$ folgt $\Phi(f_*) = f_*$. □

Wir wissen, daß stetige Funktionen über kompakten Mengen vollständig sind. Auf offenen Mengen ist zusätzlich zur Stetigkeit auch die Beschränktheit zu fordern:

Satz 8.2 *Sei $U \subset \mathbb{R}^n$ offen. Der Vektorraum $\mathcal{C}_b(U, \mathbb{R}^m)$ der beschränkten stetigen Abbildungen $f : U \to \mathbb{R}^m$, zusammen mit der Norm $\|f\| := \sup_{x\in U} \|f(x)\|$, ist ein Banachraum.*

Beweis. Zu zeigen ist die Vollständigkeit. Sei $(f_k)_{k\in N}$ bezüglich dieser Norm eine Cauchy-Folge von Abbildungen $f_k \in \mathcal{C}_b(U, \mathbb{R}^m)$, d.h. für jedes $\epsilon > 0$ gibt es ein $k \in \mathbb{N}$ mit $\sup_{x\in U} \|f_i(x) - f_j(x)\| < \epsilon$ für alle $i, j \geq k$. Dann ist auch $\|f_i(x) - f_j(x)\| < \epsilon$ für alle $i, j \geq k$ und alle $x \in U$. Wegen $f_i(x) \in \mathbb{R}^m$ und der Vollständigkeit des $\mathbb{R}^n$ wird durch $f(x) := \lim_{k\to\infty} f_k(x)$ punktweise eine Abbildung $f : U \to \mathbb{R}^m$ definiert.

Zu zeigen ist, daß f stetig ist und daß $(f_k)_{k\in\mathbb{N}}$ in $C_b(U, \mathbb{R}^m)$ gegen f konvergiert. Da $(f_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $C_b(U, \mathbb{R}^m)$ ist, gibt es zu jedem $\epsilon > 0$ ein $k \in \mathbb{N}$, so daß für alle $i, j \geq k$ und einen beliebigen Punkt $x \in U$ gilt $\|f_i(x) - f_j(x)\| < \epsilon$. Das gilt aber auch für den Grenzwert $j \to \infty$:

$$\lim_{j\to\infty} \|f_i(x) - f_j(x)\| = \|f_i(x) - f(x)\| < \epsilon \quad \text{für alle } x \in U \ .$$

Also konergiert $(f_k)_{k\in\mathbb{N}}$ gleichmäßig gegen f, und nach Satz 1.14 ist f stetig. Da $\|f_i(x) - f(x)\| < \epsilon$ für $i \geq k$ und f_i beschränkt ist, ist f ebenfalls beschränkt, und

$$\|f_i - f\| = \sup_{x\in U} \|f_i(x) - f(x)\| < \epsilon \qquad \text{für alle } i \geq k \ .$$

Also gilt $\lim_{k\to\infty} f_k = f \in \mathcal{C}_b(U, \mathbb{R}^m)$. □

Satz 8.3 (über implizite Funktionen) *Seien $U_1 \subset \mathbb{R}^n$ und $U_2 \subset \mathbb{R}^m$ offene Teilmengen und $F = (F_1, \ldots, F_m) : U_1 \times U_2 \to \mathbb{R}^m$ eine stetig differenzierbare Abbildung. Es sei $(a, b) \in U_1 \times U_2$ ein Punkt mit $F(a, b) = 0$, und im Punkt (a, b) sei die Jacobi-Matrix bezüglich der 2. Komponente*

$$\Big(\frac{\partial F}{\partial y}\Big)(a,b) := \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}(a,b) \in GL(m,\mathbb{R})$$

invertierbar. Dann gibt es eine offene Umgebung $V_1 \subset U_1$ von a und eine Umgebung $V_2 \subset U_2$ von b sowie eine stetig differenzierbare Abbildung $g : V_1 \to V_2 \subset \mathbb{R}^m$, so daß

$$F(x, g(x)) = 0 \qquad \textit{für alle } x \in V_1 \ .$$

Mit anderen Worten: Ist eine implizit gegebene Gleichung $F(x, g(x)) = 0$ in einem Punkt a lösbar mit $g(a) = b$ und $F(a, b) = 0$, dann ist sie (unter den gegebenen Voraussetzungen) sogar in einer Umgebung V_1 von a lösbar.

Beweis. Wir setzen $B := \big(\frac{\partial F}{\partial y}\big)(a, b) \in GL(m, \mathbb{R})$. Damit werde eine Abbildung $G : U_1 \times U_2 \to \mathbb{R}^m$ definiert durch

$$G(x, y) := y - B^{-1} \cdot F(x, y) \ , \qquad x \in U_1 \ , \ y \in U_2 \subset \mathbb{R}^m \ .$$

Die Jacobi-Matrix von G im Punkt $(x, y) \in U_1 \times U_2$ ist

$$\Big(\frac{\partial G}{\partial y}\Big)(x,y) = E_m - B^{-1} \cdot \Big(\frac{\partial F}{\partial y}\Big)(x,y) \qquad \Rightarrow \qquad \Big(\frac{\partial G}{\partial y}\Big)(a,b) = 0 \in M(m \times m, \mathbb{R}) \ .$$

Somit ist $\big\|\big(\frac{\partial G}{\partial y}\big)(a,b)\big\|_{op} = 0$. Wegen der Stetigkeit von $\frac{\partial G}{\partial y}$ gibt es Umgebungen $W_1 \subset U_1$ von a und $W_2 \subset U_2$ von b, so daß

$$\Big\|\Big(\frac{\partial G}{\partial y}\Big)(x,y)\Big\|_{op} \leq \frac{1}{2} \qquad \text{für alle } (x, y) \in W_1 \times W_2 \ . \tag{1}$$

Entscheidend ist die Beobachtung

$$F(x, y) = 0 \qquad \Leftrightarrow \qquad y = G(x, y) \tag{2}$$

und insbesondere $b = G(a, b)$. Damit führen wir die Lösung von $F(x, y) = 0$ nach y auf ein Fixpunktproblem zurück.

(1. Konstruktion der Abbildung g) Wir geben zunächst einen vereinfachten Beweis, der die Abbildung g nur punktweise konstruiert und keine Informationen über Stetigkeit liefert. Die Gebiete können dann etwas größer gewählt werden.

Nach dem Satz von Taylor gilt

$$G(x,y) - G(x,\eta) = \Big(\frac{\partial G}{\partial y}\Big)(x,\eta) \cdot (y - \eta) + \phi(y - \eta) \ , \quad \phi(y-\eta) = o(\|y - \eta\|) \ ,$$

für alle (x,y), $(x,\eta) \in W_1 \times W_2$. Es gibt also ein $r > 0$, so daß

$$\|\phi(y-\eta)\| \le \frac{1}{4}\|y-\eta\| \qquad \text{für alle } y \in W_2 \text{ mit } \|y-\eta\| < \frac{5}{2}r \ . \tag{3}$$

Es sei $V_2 := \{y \in W_2 \ : \ \|y-b\| \le r\} \subset W_2$ die abgeschlossene Kugel mit Mittelpunkt b und Radius r. Damit gilt

$$\|G(x,y)-G(x,\eta)\| \le \Big\| \Big(\frac{\partial G}{\partial y}\Big)(x,\eta) \Big\|_{op} \|y-\eta\| + \|\phi(y-\eta)\| \le \frac{3}{4}\|y-\eta\| \tag{4}$$

für alle $x \in W_1$ und alle $y, \eta \in V_2$. Da $G(a,b) = b$, gibt es wegen der Stetigkeit von F eine offene Umgebung $V_1 \subset W_1$ von a, so daß

$$\sup_{x \in V_1} \|G(x,b)-b\| \le \frac{r}{4} \ .$$

Zusammen mit $\eta \mapsto b$ in (4) folgt

$$\sup_{(x,y)\in V_1\times V_2} \|G(x,y)-b\| \le \sup_{(x,y)\in V_1\times V_2} \Big(\|G(x,y)-G(x,b)\| + \|G(x,b)-b\| \Big) \le r \ . \tag{5}$$

Folglich ist für beliebiges, aber festes $x \in V_1$ die Abbildung $\Phi : V_2 \to \mathbb{R}^m$ definiert durch $\Phi(y) := G(x,y)$ eine Kontraktion auf der abgeschlossenen Teilmenge $V_2 \subset \mathbb{R}^m$, so daß es nach dem Banachschen Fixpunktsatz zu jedem $x \in W_1$ genau einen Fixpunkt $y_* \in V_2$ gibt mit $y_* = \Phi(x,y_*) = G(x,y_*)$. Nach (2) löst die so definierte Funktion $g : V_1 \to V_2$ mit $g(x) := y_*$ das Problem $F(x,g(x)) = 0$.

(2. Stetigkeit von g) Wir betrachten jetzt stetige und beschränkte Abbildungen $\gamma \in \mathcal{C}_b(V_1, \mathbb{R}^m)$ mit $\|\gamma\| = \sup_{x\in V_1} \|\gamma(x)\|$. Zu gegebenem $\gamma \in \mathcal{C}_b(V_1,\mathbb{R}^m)$ werde eine Abbildung $\psi : V_1 \to \mathbb{R}^m$ definiert durch

$$\psi(x) := G(x,\gamma(x)) = \gamma(x) - B^{-1} \cdot F(x,\gamma(x)) \ .$$

Wegen der Stetigkeit von F ist ψ ebenfalls stetig. Außerdem ist ψ beschränkt: Falls $\|\gamma - b\| \le r$, dann gilt nach (5) auch $\|\psi - b\| \le r$. Dabei wird b als die konstante Funktion aufgefaßt. Damit wird durch $\Phi(\gamma) := \psi$ eine Abbildung $\Phi : A \to A$ der abgeschlossenen Teilmenge

$$A := \{\gamma \in \mathcal{C}_b(V_1,\mathbb{R}^m) \ : \ \|\gamma - b\| \le r\} \subset \mathcal{C}_b(V_1,\mathbb{R}^m)$$

eines Banach-Raumes auf sich selbst definiert. Nach (4) gilt für alle $\gamma \in A$

$$\begin{aligned} \|\Phi(\gamma_1) - \Phi(\gamma_2)\| &= \sup_{x\in V_1} \|G(x,\gamma_1(x)) - G(x,\gamma_2(x))\| \\ &\le \frac{3}{4} \sup_{x\in V_1} \|\gamma_1(x) - \gamma_2(x)\| = \frac{3}{4}\|\gamma_1 - \gamma_2\| \ . \end{aligned}$$

Folglich ist $\Phi : A \to A$ eine Kontraktion und hat nach dem Banachschen Fixpunktsatz genau einen Fixpunkt $\gamma_* \in \mathcal{C}_b(V_1, \mathbb{R}^m)$. Diese eindeutig bestimmte stetige und beschränkte Funktion $g = \gamma_* : V_1 \to \mathbb{R}^n$ löst die Gleichung $F(x, g(x)) = 0$ für alle $x \in V_1$.

(3. partielle Differenzierbarkeit der Lösung) Die Jacobi-Matrix $(\frac{\partial F}{\partial y})(x, y) \in M(m, \mathbb{R})$ ist genau dann invertierbar, wenn ihre Determinante ungleich Null ist. Da die Determinante als Polynom der Matrixelemente eine stetige Funktion der Matrixelemente ist und $\det B = \det((\frac{\partial F}{\partial y})(a, b)) \neq 0$ ist, gibt es eine Umgebung $V_1' \subset V_1$ von a, so daß $(\frac{\partial F}{\partial y})(x, g(x))$ invertierbar ist für alle $x \in V_1'$.

Über die Lösung $g = (g_1, \ldots, g_m) : V_1' \to \mathbb{R}^m$ definieren wir eine Abbildung $\tilde{F} : V_1' \to \mathbb{R}^m$ durch $\tilde{F}(x) := F(x, g(x))$. Da $\tilde{F}(x) = 0$ für alle $x = (x_1, \ldots, x_n) \in V_1'$, verschwinden auch alle partiellen Ableitungen von $\tilde{F}$:

$$0 = \Big(\frac{\partial \tilde{F}}{\partial x_j}\Big)(x) = \Big(\frac{\partial F}{\partial x_j}\Big)(x, y)\Big|_{y=g(x)} + \sum_{i=1}^{m} \Big(\frac{\partial F}{\partial y_i}\Big)(x, y)\Big|_{y=g(x)} \cdot \Big(\frac{\partial g_i}{\partial x_j}\Big)(x) \; .$$

Sind $K_{il}(x, y)$ die Matrixelemente der inversen Matrix $J^{-1} = (K_{il}(x, y))$ der Jacobi-Matrix

$$J = (J_{li}(x, y)) \in M(m, \mathbb{R}) \; , \qquad J_{li}(x, y) = \Big(\frac{\partial F_l}{\partial y_i}\Big)(x, y) \; ,$$

von F im Punkt (x, y), dann erhalten wir

$$\Big(\frac{\partial g_i}{\partial x_j}\Big)(x) = -\sum_{l=1}^{m} K_{il}(x, y) \Big(\frac{\partial F_l}{\partial x_j}\Big)(x, y)\Big|_{y=g(x)} \; .$$

Insbesondere ist $g : V_1' \to \mathbb{R}^m$ partiell differenzierbar. Aus der stetigen Differenzierbarkeit von F und der Stetigkeit der Bildung der inversen Matrix folgt, daß die $\frac{\partial g_i}{\partial x_j}$ stetig sind. Nach Satz 5.4 ist g dann total differenzierbar. □

Eine erste nützliche Anwendung des Satzes über implizite Funktionen sind die Höhenlinien von Funktionen zweier Veränderlicher $f : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^2$ offen. Eine *Niveaumenge* von f ist die Menge

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

für $c \in \mathbb{R}$. Dabei kann $N_f(c)$ leer sein, aus einem Punkt bestehen oder eine eigentliche Höhenlinie darstellen. Im letzten Fall kann der Satz über implizite Funktionen genutzt werden, um die Höhenline lokal zu konstruieren.

Es werde vorausgesetzt, daß eine Lösung $(a, b) \in U$ mit $f(a, b) = c$ bekannt ist. Wir können $(\operatorname{grad} f)(a, b) \neq 0$ voraussetzen (sonst gibt es ein lokales Extremum oder einen Sattelpunkt, was wir schon vorher untersucht hatten). Ist

$(\frac{\partial f}{\partial x})(a,b) \neq 0$, dann konstruiert für $F(x,y) := f(x,y) - c$ der Satz über implizite Funktionen ein offenes Intervall I mit $a \in I$ und eine eindeutig bestimmte (partiell) differenzierbare Lösung $g : I \to \mathbb{R}$, so daß $f(x, g(x)) = c$ für alle $x \in I$ gilt. Ist $(\frac{\partial f}{\partial y})(a,b) \neq 0$, so kann/muß man die Rolle von x und y vertauschen: Für $F(y,x) := f(x,y) - c$ konstruiert der Satz über implizite Funktionen für ein offenes Intervall J mit $b \in J$ eine eindeutig bestimmte (partiell) differenzierbare Lösung $h : J \to \mathbb{R}$, so daß $f(h(y), y) = c$ für alle $y \in J$.

Dadurch wird eine offene Teilmenge V der Höhenlinie konstruiert. Man kann einen so gewonnenen Punkt der Teilmenge als neuen Startpunkt für den Satz über implizite Funktionen nehmen und so die Höhenlinie über die Teilmenge V hinaus ausdehnen. Das Konstruktionsverfahren läßt sich auf dem Computer implementieren und kann zur dreidimensionalen Visualisierung von Funktionen $f : U \to \mathbb{R}$ verwandt werden.

Eine weitere wichtige Anwendung des Satzes über implizite Funktionen ist das lokale Invertieren einer Abbildung $f : V \to U$ mit $U, V \subset \mathbb{R}^n$.

Satz 8.4 *Sei $V \subset \mathbb{R}^n$ offen und $f : V \to \mathbb{R}^n$ eine stetig differenzierbare Abbildung in einem Punkt $b \in V$. Das Differential von f sei invertierbar in $b \in V$, d.h. $(Df)(b) \in GL(n, \mathbb{R})$. Dann gibt es eine offene Umgebung $V_0 \subset V$ von b und eine offene Umgebung $U_0 \subset \mathbb{R}^n$ von $a := f(b)$ so daß $f : V_0 \to U_0$ bijektiv ist und die Umkehrabbildung $g = f^{-1} : U_0 \to V_0$ stetig differenzierbar ist. Außerdem gilt $(Dg)(a) = ((Df)(b))^{-1}$.*

Beweis. Wir verwenden den Satz über implizite Funktionen mit $F : \mathbb{R}^n \times V \to \mathbb{R}^n$ gegeben durch $F(x,y) := x - f(y)$. Ziel ist lokale Auflösung nach $y = g(x) = f^{-1}(x)$.

Es gilt $F(a,b) = 0$ und $(\frac{\partial F}{\partial y})(a,b) = -(Df)(b) \in GL(n, \mathbb{R})$. Nach dem Satz über implizite Funktionen existiert eine offene Umgebung U' von a, eine Umgebung $V' \subset V$ von b und eine eindeutig bestimmte stetig differenzierbare Abbildung $g : U' \to V'$, so daß

$$F(x, g(x)) = x - f(g(x)) = 0 \qquad \text{für alle } x \in U' \ .$$

Wegen der Stetigkeit von f gibt es eine offene Umgebung $V_0 \subset V'$ von b mit $f(V_0) := U_0 \subset U'$. Aus der Stetigkeit von g folgt, daß U_0 offen ist. Also ist $f : V_0 \to U_0$ bijektiv mit $f^{-1} = g : U_0 \to V_0$.

Aus $y = g(f(y))$ und der Kettenregel folgt

$$(D(g \circ f))(y) = E_n = (Dg)(f(y)) \cdot (Df)(y) \qquad \text{für alle } y \in V_0$$

und insbesondere $(Dg)(a) = ((Df)(b))^{-1}$. □

Definition 8.5 Unter einem *Diffeomorphismus* zwischen offenen Teilmengen $U, V \subset \mathbb{R}^n$ versteht man eine bijektive stetig differenzierbare Abbildung $f : V \to U$, so daß f^{-1} ebenfalls stetig differenzierbar ist.

Der Satz 8.4 über die lokale Invertierbarkeit besagt also, daß sich jede differenzierbare Abbildung $f : V \to U$, deren Differential in einem Punkt $b \in V$ invertierbar ist, zu einem lokalen Diffeomorphismus fortsetzen läßt.

Als Anwendung der lokalen Invertierbarkeit betrachten wir die Umrechnung zwischen kartesischen Koordinaten und Polarkoordinaten. Dazu sei $f : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{R}^2$ gegeben durch $f(r, \phi) := (r \cos \phi, r \sin \phi)$. Das totale Differential (Jacobi-Matrix) ist

$$(Df)(r,\phi) = \begin{pmatrix} \frac{\partial f_1}{\partial r} & \frac{\partial f_1}{\partial \phi} \\ \frac{\partial f_2}{\partial r} & \frac{\partial f_2}{\partial \phi} \end{pmatrix} (r,\phi) = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}$$

Damit gilt $\det((Df)(r, \phi)) = r > 0$, die Jacobi-Matrix ist also in jedem Punkt $(r, \phi) \in \mathbb{R}_+^* \times \mathbb{R}$ invertierbar. Folglich ist f in jedem Punkt von $\mathbb{R}^2 \setminus \{0\}$ lokal invertierbar. Das ist in diesem Fall auch direkt zu erhalten: Ist $f(r, \phi) = (x, y)$, dann ist $r = \sqrt{x^2 + y^2}$ und $\cos \phi = \frac{x}{r}$, $\sin \phi = \frac{y}{r}$. Damit erhalten wir die Jacobi-Matrix für *eine* lokale Umkehrung $g : U \to V$ mit $f(g(x, y)) = (x, y)$ zu

$$(Dg)(x,y) = \big((Df)(r,\phi)\big)^{-1} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\frac{\sin\phi}{r} & \frac{\cos\phi}{r} \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} .$$

Eine Bijektion läßt sich z.B. finden zwischen

$$f : \mathbb{R}_+^* \times \,]-\tfrac{\pi}{2}, \tfrac{\pi}{2}[\,\to \mathbb{R}_+^* \times \mathbb{R} \, .$$

Für $(x, y) \in \mathbb{R}_+^* \times \mathbb{R}$ setzen wir dann $g(x, y) = \Big(\sqrt{x^2 + y^2}, \arctan \frac{y}{x}\Big)$. Es existiert aber keine globale Bijektion $f : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{R}^2$, da f periodisch im Winkel ϕ ist.

9 Untermannigfaltigkeiten

Definition 9.1 Eine Teilmenge $M \subset \mathbb{R}^{n+k}$ heißt *n-dimensionale Untermannigfaltigkeit*, wenn zu jedem Punkt $a \in M$ eine offene Umgebung $U \subset \mathbb{R}^{n+k}$ von a und eine stetig differenzierbare Abbildung $f : U \to \mathbb{R}^k$ existieren, so daß

i) $U \cap M = f^{-1}(0)$,

ii) für alle $x \in U$ mit $f(x) = 0 \in \mathbb{R}^k$ hat das Differential $(Df)(x) \in M(k \times (n + k), \mathbb{R})$ den maximalen Rang k.

Mit diesen Bezeichnungen heißt k die *Kodimension* von M.

Beispiel 9.2 (Sphäre S^n) Dazu sei $U := \mathbb{R}^{n+1} \setminus \{0\}$ und $f : U \to \mathbb{R}^1$ gegeben durch $f(x) = x_1^2 + \cdots + x_{n+1}^2 - 1$. Dann ist $\operatorname{rang}((Df)(x)) = \operatorname{rang}(2(x_1, \ldots, x_{n+1})) = 1$ für alle $x \in U$. Somit ist

$$U \cap S^n = S^n = f^{-1}(0) = \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \ : \ x_1^2 + \cdots + x_{n+1}^2 = 1\}$$

eine n-dimensionale Untermannigfaltigkeit. Die Kodimension ist 1. ◁

Definition 9.3 Sei $M \subset \mathbb{R}^{n+k}$ Untermannigfaltigkeit, $a \in M$ und $f : U \to \mathbb{R}^k$ die die Untermannigfaltigkeit M definierende differenzierbare Abbildung mit $a \in U \subset \mathbb{R}^{n+k}$. Dann heißt der Untervektorraum

$$T_a(M) := \ker((Df)(a)) = \{v \in \mathbb{R}^{n+k} \ : \ (Df)(a) \cdot v = 0\} \subset \mathbb{R}^{n+k}$$

der *Tangentialraum* von M im Punkt $a \in M$. Sei orthogonales Komplement bezüglich des kanonischen Skalarprodukts $\langle \ , \ \rangle$ im $\mathbb{R}^{n+k}$,

$$N_a(M) := T_a(M)^\perp := \{w \in \mathbb{R}^{n+k} \ : \ \langle v, w\rangle = 0 \text{ für alle } v \in T_a(M)\}$$

heißt der *Normalenvektorraum* von M im Punkt a. Elemente $v \in T_a(M)$ bzw. $w \in N_a(M)$ heißen *Tangentialvektoren* bzw. *Normalenvektoren* an M im Punkt a.

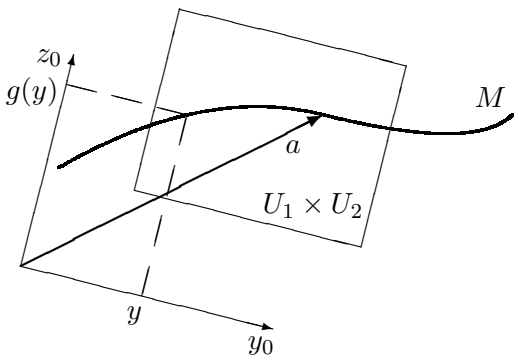
Als Kern einer linearen Abbildung ist $T_a(M) \subset \mathbb{R}^{n+k}$ ein Untervektorraum. Wegen $\dim(\mathbb{R}^{n+k}) = \dim(\ker(Df)) + \operatorname{rang}(Df)$ ist $\dim(T_a(M)) = n$ für alle $a \in M$. Nach dem Schmidtschen Orthonormalisierungsverfahren ist $N_a(M)$ ein k-dimensionaler Vektorraum.

Definition 9.4 Sei $T \subset \mathbb{R}^n$ offen. Eine stetig differenzierbare Abbildung $\phi : T \to \mathbb{R}^{n+k}$ heißt *Immersion*, wenn $\operatorname{rang}((D\phi)(t)) = n$ für alle $t \in T$.

Satz 9.5 (lokales Koordinatensystem) *Sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit. Dann gibt es zu jedem Punkt $a \in M$ eine offene Umgebung $V \subset M$, eine offene Teilmenge $T \subset \mathbb{R}^n$ und eine Immersion $\phi : T \to \mathbb{R}^{n+k}$, die T homöomorph auf V abbildet.*

Ein Homöomorphismus war eine bijektive stetige Abbildung ϕ mit stetigem Inversen ϕ^{-1}.

Lemma 9.6 (angepaßte Koordinaten) *Sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit. Dann gibt es zu jedem $a \in M$ eine Zerlegung $\mathbb{R}^{n+k} = T_a(M) \oplus N_a(M) \simeq \mathbb{R}^n \times \mathbb{R}^k$ mit offenen Mengen $U_1 \subset T_a(M) \simeq \mathbb{R}^n$ und $U_2 \subset N_a(M) \simeq \mathbb{R}^k$ sowie eine differenzierbare Abbildung $g : U_1 \to U_2$, so daß $M \cap (U_1 \times U_2) = \big\{(y, g(y)) \ : \ y \in U_1\big\}$.*



Beweis: Im $\mathbb{R}^{n+k}$ werde ein Koordinatensystem so gewählt, daß die ersten n Koordinatenrichtungen den Tangentialraum $T_a(M)$ aufspannen und die letzten k Koordinatenrichtungen den Normalenvektorraum $N_a(M)$. In diesen Koordinaten habe $x \in M$ die Darstellung $x = (y, z)$ mit $y \in \mathbb{R}^n$ und $z \in \mathbb{R}^k$. Speziell ist $a = (y_0, z_0)$. Seien $V_1 \subset \mathbb{R}^n$ eine offene Umgebung von y_0 und $V_2 \subset \mathbb{R}^k$ eine offene Umgebung von z_0, so daß $\operatorname{rang}((Df)((y,z))) = k$ für alle $(y,z) \in M \cap (V_1 \times V_2)$. Sei $(Df)(x,y) =: (a_{ij}(y,z))$ mit $a_{ij}(y,z) = \frac{\partial f_i}{\partial y_j}(y,z)$ für $1 \le j \le n$ und $a_{i,n+j}(y,z) = \frac{\partial f_i}{\partial z_j}(y,z)$ für $1 \le j \le k$. Die Jacobi-Matrix bezüglich der 2. Kompenente ist invertierbar in (y_0, z_0), d.h. $\frac{\partial f}{\partial z}(y_0, z_0) \in GL(k, \mathbb{R})$: Für $w \neq 0$ ist $\binom{0}{w} \in \big(\ker((Df)(y_0,z_0))\big)^\perp$, also $0 \neq (Df(y_0,z_0)) \cdot \binom{0}{w} = \frac{\partial f}{\partial z}(y_0,z_0) \cdot w$, d.h. $\frac{\partial f}{\partial z}(y_0,z_0)$ ist injektiv, wegen der Gleichheit der Dimensionen dann bijektiv. Nach dem Satz über implizite Funktionen existiert eine offene Umgebung $U_1 \subset V_1$ von y_0 und eine offene Umgebung $U_2 \subset V_2$ von z_0 sowie eine eindeutig bestimmte stetig differenzierbare Abbildung $g : U_1 \to U_2$ mit $f(y, g(y)) = 0$ für alle $y \in U_1$. □

Beweis von Satz 9.5. Nach Lemma 9.6 existieren offene Teilmengen $U_1 \subset \mathbb{R}^n$ und $U_2 \subset \mathbb{R}^k$ sowie eine stetig differenzierbare Abbildung $g : U_1 \to U_2$, so daß $M \cap (U_1 \times U_2) = \big\{(y, g(y)) \,:\, y \in U_1\big\}$. Wir setzen $V = M \cap (U_1 \times U_2)$ und $T = U_1$ sowie $\phi(y) = (y, g(y))$. Surjektivität von $\phi : T \to V$ folgt aus der Konstruktion und Injektivität aus der Eindeutigkeit von g. Dann ist $\phi^{-1} : (y, g(y)) \mapsto y$ die Projektion auf die erste Komponente, und damit stetig.

Schließlich gilt $(D\phi)(y) = \begin{pmatrix} E_n \\ (Dg)(y) \end{pmatrix}$. Da g differenzierbar auf T ist, ist auch ϕ differenzierbar in y, und es gilt $\operatorname{rang}((D\phi)(y)) = n$. Damit ist $\phi : T \to V$ eine Immersion. □

Es sei bemerkt, daß auch die Umkehrung von Satz 9.5 gilt: Ist eine Immersion ϕ mit diesen Eigenschaften gegeben, dann kann man zeigen, daß M Untermannigfaltigkeit ist.

Satz 9.7 *Sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit, lokal definiert durch die Abbildung $f = (f_1, \ldots, f_k) : U \to \mathbb{R}^k$, mit $U \subset \mathbb{R}^{n+k}$ offen. Sei $a \in M \cap U$ ein Punkt und $\phi = (\phi_1, \ldots, \phi_{n+k}) : T \to \mathbb{R}^{n+k}$ die aus f konstruierte Immersion, die $T \subset \mathbb{R}^n$ homöomorph auf eine Umgebung $V \subset M$ von $\phi(t_a) = a \in M$ abbildet. Dann gilt:*

i) *Zu jedem Tangentialvektor $v \in T_a(M)$ im Sinne von Definition 9.4 gibt es eine Kurve $\gamma :]-\epsilon, \epsilon[\to M$ durch $a = \gamma(0)$, so daß v Tangentialvektor an γ im Punkt a ist.*

ii) *Ist umgekehrt v ein Tangentialvektor an eine Kurve $\gamma : I \to M$ durch $a \in M$, dann gilt $v \in \ker(Df)(a) = T_a(M)$.*

iii) *Die Familie* $\big((\partial_j\phi)(t_a)\big)_{j=1,\dots,n}$ *der Vektoren* $(\partial_j\phi)(t_a) \in \mathbb{R}^{n+k}$ *ist eine Basis von* $T_a(M)$.

iv) *Die Familie* $\big((\operatorname{grad} f_i)(a)\big)_{i=1,\dots,k}$ *der Vektoren* $(\operatorname{grad} f_i)(a) \in \mathbb{R}^{n+k}$ *ist eine Basis von* $N_a(M)$.

Beweis. Sei $\gamma : I \to M$ eine Kurve auf M durch $\gamma(0) = a$, wobei das Intervall $I =]-\epsilon, \epsilon[$ so gewählt sei, daß die Kurve $s = \phi^{-1} \circ \gamma : I \to T$ durch $t_a = s(0) = \phi^{-1}(a)$ vollständig in T liegt. Wegen $f(\phi(t)) = 0$ für alle $t \in T$ gilt

$$0 = (D(f \circ \phi))(t_a) = (Df)(a) \cdot (D\phi)(t_a) \in M(k \times n, \mathbb{R}) \; . \qquad (*)$$

Schreiben wir $(D\phi)(t_a) = (v_1, \dots, v_n) \in M((n+k) \times n, \mathbb{R})$, mit Spaltenvektoren $v_j = (\partial_j\phi)(t_a) \in \mathbb{R}^{n+k}$, so folgt $v_j \in \ker((Df)(a)) = T_a(M)$ für alle $1 \le j \le n$.

iii) Wegen $\operatorname{rang}((D\phi)(t)) = \operatorname{rang}((\partial_j\phi_i)(t)) = n$ ist die Familie $(v_j)_{j=1,\dots,n}$ linear unabhängig und spannt somit einen n-dimensionalen Untervektorraum von $\mathbb{R}^{n+k}$ auf, der wegen der Gleichheit der Dimensionen identisch mit $T_a(M)$ ist. Also ist $\big((\partial_j\phi)(t_a)\big)_{j=1,\dots,n}$ eine Basis von $T_a(M)$.

i) Damit läßt sich jeder Vektor $v \in T_a(M)$ schreiben als $v = \sum_{i=1}^n c_i v_i$. Zum entsprechenden Vektor $c = (c_1, \dots, c_n) \in \mathbb{R}^n$ wählen wir die Kurven $s :]-\epsilon, \epsilon[\to T$ mit $s(\tau) = t_a + c\tau$ und $\gamma(\tau) = \phi(s(\tau))$. Dann ist

$$\gamma'(0) = (\frac{d}{d\tau}(\phi \circ s))(0) = \sum_{i=1}^n (\partial_i\phi)(t_a) s_i'(0) = \sum_{i=1}^n (\partial_i\phi)(t_a) c_i = v \; .$$

ii) Nach der Kettenregel gilt für den Tangentialvektor v and die Kurve γ

$$v = \gamma'(0) = \Big(\frac{d}{\tau}(\phi \circ s)\Big)(0) = \sum_{i=1}^n (\partial_i\phi)(t_a)\, s_i'(0) \; .$$

Damit ist $v \in T_a(M)$.

iv) Wegen $(Df)(a) = (a_{ij}) \in M(k \times (n+k), \mathbb{R})$, $a_{ij} = (\partial_j f_i)(a)$ sind die Zeilen von $(Df)(a)$ gegeben durch die Vektoren $(\operatorname{grad} f_i)(a)$, $i = 1, \dots, k$. Nach $(*)$ gilt $(Df)(a) \cdot v = 0$ für alle $v \in T_a(M)$ und somit $(\operatorname{grad} f_i)(a) \in N_a(M)$. Wegen $\operatorname{rang}((Df)(a)) = k$ ist die Familie $\big((\operatorname{grad} f_i)(a)\big)_{i=1,\dots,n}$ linear unabhängig und damit wegen der Gleichheit der Dimensionen eine Basis von $N_a(M)$. □

Satz 9.8 (Extrema mit Nebenbedingungen) *Sei* $M \subset \mathbb{R}^{n+k}$ *eine* n*-dimensionale Untermannigfaltigkeit, lokal definiert durch die Abbildung* $f = (f_1, \dots, f_k) : U \to \mathbb{R}^k$, *für* $U \subset \mathbb{R}^{n+k}$ *offen, mit* $M \cap U = f^{-1}(0)$ *und* $(\operatorname{rang}(Df)(x)) = k$ *für alle* $x \in M \cap U$.

Gegeben sei eine stetig differenzierbare Funktion $F : U \to \mathbb{R}$, *so daß die Einschränkung* $F\big|_M : M \cap U \to \mathbb{R}$ *im Punkt* $a \in M$ *ein lokales Maximum (bzw. Minimum) besitzt, d.h. es gibt eine Umgebung* $V \subset M \cap U$ *von* a, *so daß*

$$F(b) \le F(a) \qquad \text{bzw. } F(b) \ge F(a) \qquad \text{für alle } b \in V \; .$$

Dann gilt $(\operatorname{grad} F)(a) \in N_a(M)$*, es gibt also Konstanten* $\lambda_1, \ldots, \lambda_k$*, so daß*

$$(\operatorname{grad} F)(a) = \sum_{i=1}^{k} \lambda_i (\operatorname{grad} f_i)(a) \; .$$

Diese Konstanten $\lambda_1, \ldots, \lambda_k$ *heißen* Lagrange-Multiplikatoren.

Beweis. Da $F\big|_M$ in $a \in M$ ein lokales Extremum hat, hat die nochmalige Einschränkung auf eine Kurve $\gamma :]-\epsilon, \epsilon[\to M$ durch $\gamma(0) = a$ ein lokales Extremum in $t = 0$:

$$0 = \Big(\frac{d}{d\tau}(F \circ \gamma)\Big)(0) = (DF)(a) \cdot \gamma'(0) \; .$$

Wegen $\gamma'(0) \in T_a(M)$ gilt $(DF)(a) = (\operatorname{grad} F)(a) \in N_a(M)$. □

Einige wichtige Beweise mittels Lagrange-Multiplikatoren stellen wir als Übungsaufgaben (Blatt 6):

1. Für eine symmetrische Matrix $A = A^t \in M(n, \mathbb{R})$ nimmt die Einschränkung der Funktion $F : \mathbb{R}^n \to \mathbb{R}$, $F(x) = \langle x, Ax \rangle$, auf die Einheitssphäre ihr Maximum und Minimum in einem Eigenvektor an, und der Lagrange-Multiplikator ist der Eigenwert.
2. Die Ungleichung zwischen geometrischem und arithmetischem Mittel: $\sqrt[n]{x_1 \cdots x_n} \leq \frac{1}{n}(x_1 + \cdots + x_n)$ für $x_i \geq 0$.
3. Die Höldersche Ungleichung $\Big|\sum_{i=1}^{n} x_i y_i\Big| \leq \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |y_i|^q\Big)^{\frac{1}{q}}$ für $x_i, y_i \in \mathbb{R}$ und $\frac{1}{p} + \frac{1}{q} = 1$.

Untermannigfaltigkeiten spielen eine wichtige Rolle in der Mechanik. Gegeben sei ein mechanisches System aus N Teilchen (gleicher Masse m). Eine Konfiguration des Systems wird beschrieben durch einen Punkt $x = (x_1, \ldots, x_{3N}) \in \mathbb{R}^{3N}$ (Angabe aller Koordinaten zu gegebenem Zeitpunkt). Dem System werden k holonome Zwangsbedingungen auferlegt, beschrieben durch k Gleichungen $f_1(x) = 0$, $\ldots$, $f_k(x) = 0$. Wir setzen voraus, daß die aus den partiellen Ableitungen $a_{ij}(x) := (\frac{\partial f_i}{\partial x_j})(x)$ gebildete Matrix der partiellen Ableitungen maximalen Rang hat, $\operatorname{rang}(a_{ij}(x)) = k$ für alle $x \in M$. Dann definieren die Gleichungen eine $n = (3N - k)$-dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^{n+k}$. Konfigurationen des Systems mit Zwangsbedingungen sind dann durch Punkte aus M zu beschreiben. Die Dimension von M entspricht der Zahl der Freiheitsgrade.

Nach Satz 9.5 gibt es zu jedem Punkt $a \in M$ eine offene Umgebung $V \subset M$ sowie eine Umgebung $T \subset \mathbb{R}^n$ und eine Immersion $\phi : T \to M$, die T homöomorph auf V abbildet. Die Koordinaten $(q_1, \ldots, q_n)$ eines Punktes $q \in T$ heißen verallgemeinerte Koordinaten.

Sei $(K(a)) \in \mathbb{R}^{3N}$ eine Familie von Kräften, die auf die Teilchen wirken. Dann wird die Beschleunigung der Teilchen beschrieben durch das d'Alembertsche Prinzip

$$\langle m\ddot{x}(a) - K(a), v\rangle = 0 \quad \text{für beliebige } v \in T_aM \ .$$

Die Forderung besagt, daß die durch $Z(a) := m\ddot{x}(a) - K$ definierte Zwangskraft keine virtuelle Arbeit verrichtet bzw. ein Normalenvektor ist. Somit gibt es Lagrange-Multiplikatoren λ_l, $l = 1, \ldots, k$ mit $Z(a) = \sum_{l=1}^{k} \lambda_l(a)\,(\operatorname{grad} f_l)(a)$. Es liegt dann nahe, die Zwangskraft mit dem Gradienten einer Funktion $W : U \to \mathbb{R}$ in Verbindung zu setzen, deren Einschränkung auf M im Punkt a ein lokales Extremum hat. Diese Funktion ist die Wirkung und die Extremalitätsforderung heißt *Hamiltonsches Prinzip*.

Beispiel 9.9 Eine eindimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^3$ werde durch $f(x_1, x_2, x_3) = \begin{pmatrix} x_2 \\ \frac{1}{2}(l^2 - x_1^2 - x_3^2) \end{pmatrix} = 0$ definiert. Das ist relevant für die Dynamik eines Pendels, bestehend aus einem Massenpunkt aufgehängt an einem masselosen Seil der Länge l, und zusätzlicher Beschränkung der Bewegung auf die Ebene $x_2 = 0$. Es gilt

$$(Df)(x) = \begin{pmatrix} 0 & 1 & 0 \\ -x_1 & 0 & -x_3 \end{pmatrix} , \qquad \operatorname{rang}((Df)(x)) = 2 \qquad \text{für } x_1^2 + x_3^2 \neq 0 \ .$$

Damit bilden $n_1(x) := (\operatorname{grad} f_1)(x) = (0, 1, 0)$ und $n_2(x) := (\operatorname{grad} f_2)(x) = (-x_1, 0, -x_3)$ eine Basis von $N_{(x_1,x_2,x_3)}(M)$ und $v(x) = (-x_3, 0, x_1)$ eine Basis von $T_{(x_1,x_2,x_3)}(M)$. Auf den Massepunkt wirke die Kraft $(K_1, K_2, K_3) = (0, 0, -mg)$. Das d'Alembertsche Prinzip liefert

$$\Big(m\ddot{x}_1,\ m\ddot{x}_2,\ m\ddot{x}_3 + mg\Big) = \lambda_1(x)\big(0, 1, 0\big) + \lambda_2(x)\big(-x_1, 0, -x_3\big) \ .$$

Zusammen mit den beiden Gleichungen $f = 0$ haben wir 5 Gleichungen zur Bestimmung der 5 Funktionen $x_1, x_2, x_3, \lambda_1, \lambda_2$. Wegen $x_2 = 0$ ist auch $\lambda_1 = 0$. Multiplikation mit $(\dot{x}_1, 0, \dot{x}_3)^t$ und Verwendung von $x_1^2 + x_3^2 = l^2$, also $x_1\dot{x}_1 + x_3\dot{x}_3 = 0$, liefert den Energieerhaltungssatz

$$E = \frac{m}{2}((\dot{x}_1)^2 + (\dot{x}_3)^2) + mgx_3 = \text{const}$$

Multiplikation mit $(x_1, 0, x_3)^t$ und Verwendung von $x_1\ddot{x}_1 + x_3\ddot{x}_3 + (\dot{x}_1)^2 + (\dot{x}_3)^2 = 0$ liefert

$$\lambda_2 = \frac{m}{l^2}\big((\dot{x}_1)^2 + (\dot{x}_3)^2 - gx_3\big) = \frac{1}{l^2}(2E - 3mgx_3) \ .$$

Die Zwangskraft $Z = (-\lambda_2 x_1, 0, -\lambda_2 x_3)$ ist dann die Seilspannung. Die Gleichungen lassen sich in Polarkoordinaten $x_1 = l\sin\phi$ und $x_3 = -l\cos\phi$ entkoppeln ($\phi = 0$ ist die Ruhelage). Der Betrag der Seilspannung ist dann $\|Z\| = ml\dot{\phi}^2 + mg\cos\phi = \frac{mv^2}{l} + mg\cos\phi$, mit $v^2 := \|\dot{x}\|^2 = l^2\dot{\phi}^2$. ◁

10 Parameterabhängige Integrale

Wir zeigen, daß man bei geeigneten Voraussetzungen Integration und partielle Ableitungen vertauschen kann.

Satz 10.1 *Sei $I = [a,b] \subset \mathbb{R}$ ein kompaktes Intervall, $U \subset \mathbb{R}^n$ offen und $f : I \times U \to \mathbb{R}$ eine stetige Funktion. Dann ist die durch $\phi(x) := \int_a^b dt\ f(t,x)$ definierte Funktion $\phi : U \to \mathbb{R}$ stetig.*

Beweis. Es sei $(x_k)_{k\in N}$ eine Folge von Punkten $x_k \in U$, die gegen $x \in U$ konvergiert. Wir betrachten die Folge der Funktionen $F_k : I \to \mathbb{R}$, $F_k(t) := f(t,x_k)$. Da die Menge $X := I \times (\{x_k\ :\ k \in \mathbb{N}\} \cup \{x\})$ kompakt ist, ist nach Satz 3.4 die Abbildung $f : X \to \mathbb{R}$ gleichmäßig stetig. Zu gegebenem $\epsilon > 0$ gibt es also ein $\delta > 0$, so daß für alle $(t,x),(t',x') \in X$ mit $\|(t,x)-(t',x')\| < \delta$ gilt $\|f(t,x)-f(t',x')\| < \epsilon$. Insbesondere gibt es ein $l \in \mathbb{N}$, so daß $\|f(t,x) - f(t,x_k)\| < \epsilon$ für alle $t \in I$ und alle $k \geq l$. Folglich ist $(F_k)_{k\in\mathbb{N}}$ gleichmäßig stetig, so daß nach Satz 22.5 aus dem 1. Semester Integration und Grenzwertbildung vertauschen:

$$\lim_{k\to\infty} \int_a^b dt\ f(t,x_k) = \int_a^b dt \big(\lim_{k\to\infty} f(t,x_k) \big) = \int_a^b dx\ f(t,x)\ .$$

Damit ist $\phi : U \to \mathbb{R}$ stetig. □

Satz 10.2 *Seien $I, J \subset \mathbb{R}$ kompakte Intervalle und $f : I \times J \to \mathbb{R}$ eine stetige Funktion, so daß $f(t,x)$ nach x stetig partiell differenzierbar ist. Dann ist die stetige Funktion $\phi : J \to \mathbb{R}$ mit $\phi(x) := \int_I dt\ f(t,x)$ stetig differenzierbar, und es gilt*

$$\phi'(x) = \int_I dt\ \frac{\partial f}{\partial x}(t,x)\ .$$

Beweis. Es sei $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten $x_k \in J$, die gegen $x \in J$ konvergiert, wobei $x_k \neq x$ für alle $k \in \mathbb{N}$. Nach dem Mittelwertsatz gibt es zu jedem $t \in I$ und $k \in \mathbb{N}$ ein $y_k \in J$ zwischen x und x_k, so daß

$$\frac{f(t,x_k) - f(t,x)}{x_k - x} = (\partial_2 f)(t,y_k)\ .$$

Da $I \times J$ kompakt ist, ist $(\partial_2 f)(t,x)$ auf $I \times J$ gleichmäßig stetig. Also gibt es zu jedem $\epsilon > 0$ ein $\delta > 0$, so daß

$$\|(\partial_2 f)(t,x) - (\partial_2 f)(t',x')\| < \epsilon \text{ für alle } (t,x),(t',x') \in I \times J \text{ mit } \|(t,x)-(t',x')\| < \delta\ .$$

Insbesondere gibt es ein $l \in \mathbb{N}$, so daß

$$\|(\partial_2 f)(t,x) - (\partial_2 f)(t,y_k)\| < \epsilon \qquad \text{für alle } t \in I \text{ und alle } k \geq l\ .$$

Also ist die Folge $((\partial_2 f)(t,y_k))_{k\in\mathbb{N}}$ gleichmäßig konvergent auf I, so daß Integration und Grenzwertbildung vertauschen:

$$\begin{aligned}\phi'(x) &= \lim_{k\to\infty}\frac{\phi(x_k)-\phi(x)}{x_k-x} = \lim_{k\to\infty}\frac{1}{x_k-x}\Big(\int_I dt\ f(t,x_k) - \int_I dt\ f(t,x)\Big)\\ &= \int_I dt\Big(\lim_{k\to\infty}\frac{f(t,x_k)-f(t,x)}{x_k-x}\Big) = \int_I dt\ (\partial_2 f)(t,x)\ . \qquad \square\end{aligned}$$

Das Ergebnis überträgt sich sofort auf mehrere Variablen, da diese bei partiellen Ableitungen festgehalten werden: Ist $I\subset\mathbb{R}$ ein kompaktes Intervall, $U\subset\mathbb{R}^n$ offen und $f: I\times[a,b]\to\mathbb{R}$ eine stetige Funktion, die stetig partiell differenzierbar nach $x_1,\dots,x_k$ ist, dann gilt

$$\frac{\partial}{\partial x_i}\int_a^b dt\ f(t,x_1,\dots,x_n) = \int_a^b dt\ \Big(\frac{\partial f}{\partial x_i}\Big)(t,x_1,\dots,x_n)\ .$$

Eine wichtige Anwendung sind Doppelintegrale:

Satz 10.3 *Ist $f:[a,b]\times[c,d]\to\mathbb{R}$ stetig, so gilt*

$$\int_c^d dx\int_a^b dt\ f(t,x) = \int_a^b dt\int_c^d dx\ f(t,x)\ .$$

Beweis. Sei $F(\xi) = \displaystyle\int_c^\xi dx\int_a^b dt\ f(t,x)$ und $G(\xi) = \displaystyle\int_a^b dt\int_c^\xi dx\ f(t,x)$. Nach Satz 10.1 ist $\displaystyle\int_a^b dt\ f(t,x)$ stetig, so daß nach Satz 20.3 aus dem 1. Semester (Stammfunktionen) gilt $F'(\xi) = \displaystyle\int_a^b dt\ f(t,\xi)$. Andererseits gilt nach Satz 10.2 $G'(\xi) = \displaystyle\int_a^b dt\ \frac{d}{d\xi}\Big(\int_c^\xi dx\ f(t,x)\Big) = \int_a^b dt\ f(t,\xi)$. Nach Satz 17.6 aus dem 1. Semester ist $F-G=\text{const}$, und $F(c)=G(c)=0$ liefert die Behauptung. $\square$

Eine weitere wichtige Anwendung sind die Euler-Lagrange-Gleichungen der Variationsrechnung. Mit $\mathcal{C}^2[a,b]$ werde der Vektorraum der zweimal stetig differenzierbaren Funktionen über dem kompakten Intervall $[a,b]\subset\mathbb{R}$ bezeichnet. Zu $\alpha,\beta\in\mathbb{R}$ betrachten wir die Teilmenge

$$\mathcal{C}^2_{\alpha,\beta}[a,b] := \{f\in\mathcal{C}^2[a,b]\ :\ f(a)=\alpha\ ,\ f(b)=\beta\}\subset\mathcal{C}^2[a,b]$$

der zweimal stetig differenzierbaren Funktionen mit vorgegebenen Werten am Rand des Intervalls. Für eine zweimal stetig differenzierbare Abbildung $L:[a,b]\times\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ betrachten wir die Abbildung

$$\begin{aligned}&S:(\mathcal{C}^2_{\alpha,\beta}[a,b])^n\to\mathbb{R}\ ,\\ &S(\phi_1,\dots,\phi_n) := \int_a^b dt\ L\big(t,\phi_1(t),\dots,\phi_n(t),\phi_1'(t),\dots,\phi_n'(t)\big)\ ,\end{aligned}$$

mit $(\mathcal{C}^2_{\alpha,\beta}[a,b])^n := \mathcal{C}^2_{\alpha_1,\beta_1}[a,b] \times \cdots \times \mathcal{C}^2_{\alpha_n,\beta_n}[a,b]$. In physikalischen Anwendungen heißt L die *Lagrange-Funktion* und S die Wirkung. Zur Fixierung der partiellen Ableitungen setzen wir $L(t,q_1,\dots,q_n,v_1,\dots,v_n)$ und entsprechend $(\partial_{1+i}L)(t,q,v) = \frac{\partial L}{\partial q_i}(t,q,v)$ und $(\partial_{1+n+i}L)(t,q,v) = \frac{\partial L}{\partial v_i}(t,q,v)$. Eine typische Problemstellung der Variationsrechnung ist es, das Infimum von S auf $(\mathcal{C}^2_{\alpha,\beta}[a,b])^n$ zu suchen:

Satz 10.4 *Mit diesen Bezeichnungen gilt: Eine notwendige Bedingung für die Realisierung $S(\phi) = \inf_{\psi\in(\mathcal{C}^2_{\alpha,\beta}[a,b])^n} S(\psi)$ des Infimums durch die Funktion $\phi \in (\mathcal{C}^2_{\alpha,\beta}[a,b])^n$ ist die Erfüllung der Euler-Lagrange-Gleichungen*

$$\Big(\frac{d}{dt}\frac{\partial L}{\partial v_i}\Big)(t,\phi(t),\phi'(t)) - \Big(\frac{\partial L}{\partial q_i}\Big)(t,\phi(t),\phi'(t)) = 0\,, \qquad 1 \le i \le n\,.$$

Beweis. Es sei $\phi \in (\mathcal{C}^2_{\alpha,\beta}[a,b])^n$ eine Funktion mit

$$S(\phi) \le S(\psi) \qquad \text{für alle } \psi \in (\mathcal{C}^2_{\alpha,\beta}[a,b])^n\,.$$

Zu $g \in \mathcal{C}^2_{0,0}[a,b]$ sei $\delta_i = (\underbrace{0,\dots,0}_{i-1}, g, \underbrace{0,\dots,0}_{n-i}) \in (\mathcal{C}^2_{0,0}[a,b])^n$. Dann ist $\psi + \epsilon\delta_i \in (\mathcal{C}^2_{\alpha,\beta}[a,b])^n$ mit $S(\phi) \le S(\phi+\epsilon\delta_i)$. Dann besitzt $F(\epsilon) := S(\phi+\epsilon\delta_i)$ ein Minimum bei $\epsilon = 0$, und es gilt notwendigerweise $(\frac{d}{d\epsilon}F)(0) = 0$. Die Voraussetzungen von Satz 10.2 für Differentiation unter dem Integral sind erfüllt, so daß wir mit der Kettenregel erhalten:

$$\begin{aligned}(\frac{d}{d\epsilon}F)(\epsilon) &= \int_a^b dt\ \frac{\partial}{\partial\epsilon} L\big(t,\phi(t)+\epsilon\delta_i(t),\phi'(t)+\epsilon\delta_i'(t)\big)\\ &= \int_a^b dt\ \Big(\frac{\partial L}{\partial q_i}(\dots)\cdot g(t) + \frac{\partial L}{\partial v_i}(\dots)\cdot g'(t)\Big)\,.\end{aligned}$$

Dabei steht $(\dots)$ für $\big(t,\phi(t)+\epsilon\delta_i(t),\phi'(t)+\epsilon\delta_i'(t)\big)$. Im letzten Term integrieren wir partiell unter Verwendung von $g(a) = g(b) = 0$:

$$\begin{aligned}&\int_a^b dt\ \frac{\partial L}{\partial v_i}(\dots)\cdot g'(t)\\ &= \frac{\partial L}{\partial v_i}(a,\phi(a)+\epsilon\delta_i(a),\phi'(a)+\epsilon\delta_i'(a))\cdot \underbrace{g(a)}_{=0} - \frac{\partial L}{\partial v_i}(b,\phi(b)+\epsilon\delta_i(b),\phi'(b)+\epsilon\delta_i'(b))\cdot \underbrace{g(b)}_{=0}\\ &- \int_a^b dt\ \Big(\frac{d}{dt}\frac{\partial L}{\partial v_i}\Big)(\dots)\cdot g(t)\,.\end{aligned}$$

Somit folgt

$$0 = (\frac{d}{d\epsilon}F)(0) = \int_a^b dt\ \Big(\frac{\partial L}{\partial q_i}(t,\phi(t),\phi'(t)) - \Big(\frac{d}{dt}\frac{\partial L}{\partial v_i}\Big)(t,\phi(t),\phi'(t))\Big)\cdot g(t)\,.$$

Das ergibt die Euler-Lagrange-Gleichungen, falls wir zeigen können, daß aus $\int_a^b f(t)g(t)$ für alle $g \in \mathcal{C}^2_{0,0}[a,b]$ für die stetige Funktion $f : [a,b] \to \mathbb{R}$ folgt $f(t) = 0$ für alle $t \in [a,b]$. Wegen der Stetigkeit von f genügt es, $f(t) = 0$ für alle $t \in]a,b[$ zu zeigen.

Angenommen, es gibt ein $t_0 \in]a,b[$ mit $f(t_0) = \epsilon > 0$. (Der Beweis für $f(t_0) = -\epsilon < 0$ ist analog.) Wegen der Stetigkeit von f gibt es dann ein $\delta > 0$, so daß $f(t) > \frac{\epsilon}{2}$ für alle $t \in K_\delta(t_0) \subset]a,b[$. Dann gibt es (sogar beliebig oft) differenzierbare nichtnegative Funktionen g, die außerhalb $K_\delta(t_0)$ identisch verschwinden. Eine Wahl ist z.B.

$$g(t) = \begin{cases} 0 & \text{für } t \le t_0 - \delta \text{ oder } t \ge t_0 + \delta \\ e^{-\frac{1}{t-(t_0-\delta)} - \frac{1}{(t_0+\delta)-t}} & \text{für } t_0 - \delta < t < t_0 + \delta \end{cases}$$

Damit ist

$$\int_a^b dt\ f(t)g(t) = \int_{t_0-\delta}^{t_0+\delta} dt\ f(t)g(t) > \frac{\epsilon}{2} \int_{t_0-\delta}^{t_0+\delta} dt\ g(t) > 0\ ,$$

im Widerspruch zur Annahme. Daraus folgt $f(t) = 0$ für alle t und schließlich die Gültigkeit der Euler-Lagrange-Gleichungen. □

Das Supremum $S(\phi) = \sup_{\psi \in (\mathcal{C}^2_{\alpha,\beta}[a,b])^n} S(\psi)$ führt offenbar auf dieselben Euler-Lagrange-Gleichungen. Die Erfahrung zeigt, daß Bewegungsgleichungen für ein physikalisches System sich oft als Euler-Lagrange-Gleichungen zu einer Wirkung erhalten lassen. Entsprechend fordert man das *Hamiltonsche Prinzip*: Von allen möglichen Bahnkurven eines physikalischen Systems zwischen festgehaltenen Anfangs- und Endkonfigurationen ist jene Bahn realisiert, für die die Wirkung extremal ist.

In der einfachsten Situation ist die Lagrange-Funktion gegeben durch die Differenz aus kinetischer Energie und potentieller Energie. Wir diskutieren den Fall ohne potentielle Energie, aber mit allgemeinem Skalarprodukt in der kinetischen Energie.

Beispiel 10.5 (Geodätengleichung) Durch eine in jedem Punkt $x \in \mathbb{R}^n$ positiv definite symmetrische Matrix $G(x) = G^t(x) = (g_{ij}(x)) \in GL(n,\mathbb{R})$ werde punktweise ein allgemeines Skalarprodukt $\langle\ ,\ \rangle_{G(x)} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mit $\langle v(x), w(x)\rangle_{G(x)} := \langle v(x), G(x) \cdot w(x)\rangle$ definiert. Für Bahnkurven $\gamma : \mathbb{R} \to \mathbb{R}^n$ mit $t \mapsto x(t) = (x_1(t), \ldots, x_n(t))$ betrachten wir die Lagrange-Dichte

$$L(t,q,v) := \frac{m}{2}\langle \dot{x}(t), \dot{x}(t)\rangle_{G(x)} = \frac{m}{2} \sum_{i,j=1}^n g_{ij}(x(t))\, \dot{x}_i(t)\dot{x}_j(t)\ .$$

Die Euler-Lagrange-Gleichungen sind dann

$$\begin{aligned}
0 &= \frac{d}{dt}\Big(m\sum_{j=1}^{n} g_{kj}(x(t))\dot{x}_j(t)\Big) - \frac{m}{2}\sum_{i,j=1}^{n}\frac{\partial g_{ij}}{\partial x_k}(x(t))\dot{x}_i(t)\dot{x}_j(t) \\
&= m\sum_{j=1}^{n} g_{kj}(x(t))\,\ddot{x}_j(t) + m\sum_{i,j=1}^{n}\frac{\partial g_{kj}}{\partial x_i}(x(t))\dot{x}_i(t)\dot{x}_j(t) - \frac{m}{2}\sum_{i,j=1}^{n}\frac{\partial g_{ij}}{\partial x_k}(x(t))\dot{x}_i(t)\dot{x}_j(t) \\
&= m\sum_{j=1}^{n} g_{kj}(x(t))\ddot{x}_j(t) + \frac{m}{2}\sum_{i,j=1}^{n}\Big(\frac{\partial g_{kj}}{\partial x_i}(x(t)) + \frac{\partial g_{ki}}{\partial x_j}(x(t)) - \frac{\partial g_{ij}}{\partial x_k}(x(t))\Big)\dot{x}_i(t)\dot{x}_j(t)\ .
\end{aligned}$$

Durch Multiplikation mit der inversen Matrix $G^{-1}(x) = (h_{lk}(x)) \in GL(n,\mathbb{R})$ entsteht die *Geodätengleichung*

$$\begin{aligned}
0 &= \ddot{x}_l(t) + \sum_{i,j=1}^{n}\Gamma_{ij}^{l}(x(t))\dot{x}_i(t)\dot{x}_j(t)\ , \\
\Gamma_{ij}^{l}(x(t)) &:= \frac{1}{2}\sum_{k=1}^{n} h_{lk}(x(t))\Big(\frac{\partial g_{kj}}{\partial x_i}(x(t)) + \frac{\partial g_{ki}}{\partial x_j}(x(t)) - \frac{\partial g_{ij}}{\partial x_k}(x(t))\Big)\ .
\end{aligned}$$

Im Fall des kanonischen Skalarprodukts $G = E_n$ ergibt sich das Trägheitsgesetz $\ddot{x}_l = 0$, da auf das Teilchen dann keine Kraft wirkt. In der allgemeinen Relativitätstheorie beschreibt die *Metrik* $G \neq E_n$ das Gravitationsfeld. Dann beschreiben die Geodätengleichungen die Bahnen von Testteilchen im Gravitationsfeld. Für geeignete Wahl von G, welche dem Gravitationsfeld einer kugelsymmetrischen Masseverteilung entspricht, ergeben sich aus den Bahngleichungen berühmte Effekte wie die Lichtablenkung an der Sonne, beobachtbar bei einer totalen Sonnenfinsternis, und die Periheldrehung des Merkur.

Teil III
Gewöhnliche Differentialgleichungen

11 Definition und Beispielklassen

Unter einer gewönlichen Differentialgleichung versteht man eine Gleichung zwischen einer Funktion y, ihren Ableitungen $y', y'', \dots y^{(n)}$ und der Variablen x, genauer:

Definition 11.1 Es sei $I \subset \mathbb{R}$ ein Intervall, $G \subset \mathbb{R}^{n+2}$ eine Teilmenge mit $n \geq 1$ und $y : I \to \mathbb{R}$ eine n-mal differenzierbare Funktion, so daß für $x \in I$ gilt $(x, y(x), y'(x), \dots, y^{(n)}(x)) \in G$. Die Funktion y erfüllt eine *gewönliche Differentialgleichung n-ter Ordnung*, wenn es eine Funktion $F : G \to \mathbb{R}$ gibt, so daß

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0 \qquad \forall x \in I \ .$$

Läßt sich diese Gleichung nach $y^{(n)}(x)$ auflösen zu

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)) \ ,$$

dann heißt die Differentialgleichung *explizit*.

Wir hatten in Abschnitt 25 des 1. Semesters bereits einige elementare Lösungsmethoden vorgestellt. Wir beschränken uns auf eine kurze Wiederholung der Definitionen und Lösungen.

Definition 11.2 (DGL 1. Ordnung mit getrennten Variablen) Seien $I, J \subset \mathbb{R}$ offene Intervalle und $f : I \to \mathbb{R}$ und $g : I \to \mathbb{R}$ stetige Funktionen, wobei $g(y) \neq 0$ für alle $y \in J$. Dann heißt

$$y'(x) = f(x)\, g(y)$$

Differentialgleichung mit getrennten Variablen.

Satz 11.3 *Zu gegebenem Anfangspunkt $(x_0, y_0) \in I \times J$ definieren wir Funktionen $F : I \to \mathbb{R}$ und $G : J \to \mathbb{R}$ durch*

$$F(x) := \int_{x_0}^{x} dt\ f(t) \ , \qquad G(y) := \int_{y_0}^{y} \frac{ds}{g(s)} \ .$$

Ist $I' \subset I$ ein Intervall mit $F(I') \subset G(J)$, dann gibt es genau eine Lösung $y : I' \to \mathbb{R}$ der Differentialgleichung $y' = f(x)\, g(y)$ mit der Anfangsbedingung $y(x_0) = y_0$. Diese Lösung erfüllt die Gleichung

$$G(y(x)) = F(x) \qquad \textit{für alle } x \in I' \ .$$

Definition 11.4 (lineare Differentialgleichung 1. Ordnung) Es seien $I \subset \mathbb{R}$ und $a, b : I \to \mathbb{R}$ stetige Funktionen. Dann heißt

$$y'(x) = a(x)\, y(x) + b(x)$$

eine *lineare Differentialgleichung 1. Ordnung*. Diese heißt für $b = 0$ *homogen*, sonst *inhomogen*.

Satz 11.5 *Es sei $x_0 \in I$ und $y_0 \in \mathbb{R}$. Dann gibt es genau eine Lösung $y : I \to \mathbb{R}$ der Differentialgleichung $y' = a(x)y$ mit der Anfangsbedingung $y(x_0) = y_0$, nämlich*

$$y(x) = y_0 \exp\Big(\int_{x_0}^{x} dt\; a(t)\Big)\,.$$

Satz 11.6 (Variation der Konstanten) *Es seien $I \subset \mathbb{R}$ ein Intervall, $a, b : I \to \mathbb{R}$ stetige Funktionen sowie $x_0 \in I$ und $y_0 \in \mathbb{R}$. Dann gibt es genau eine Lösung $y : I \to \mathbb{R}$ der Differentialgleichung $y' = a(x)y + b(x)$ mit der Anfangsbedingung $y(x_0) = y_0$, nämlich*

$$y(x) = \tilde{y}(x)\Big(y_0 + \int_{x_0}^{x} dt\; b(t)\tilde{y}^{-1}(t)\Big) \qquad \textit{mit} \qquad \tilde{y}(s) = \exp\Big(\int_{x_0}^{s} dt\; a(t)\Big)\,.$$

Eine weitere wichtige elementar lösbare Klasse sind die *exakten Differentialgleichungen*, deren Lösungen eine Kurvenschar $F(x, y) = C$ in der x-y-Ebene bilden.

Definition 11.7 (Exakte Differentialgleichung) Sei $G \subset \mathbb{R}^2$ offen und zusammenhängend und $g, h : G \to \mathbb{R}$ stetig. Die Differentialgleichung

$$g(x, y) + h(x, y)y' = 0$$

heißt *exakt*, wenn es eine stetig differenzierbare Abbildung $F : G \to \mathbb{R}$ gibt, so daß $g = \frac{\partial F}{\partial x}$ und $h = \frac{\partial F}{\partial y}$. In diesem Fall heißt F die *Stammfunktion* der Differentialgleichung.

Die Stammfunktion einer exakten Differentialgleichung ist bis auf eine Konstante eindeutig: Sei $\tilde{F}$ eine zweite Stammfunktion, dann ist $u := F - \tilde{F}$ stetig differenzierbar mit $\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = 0$. Wegen der stetigen Differenzierbarkeit verschwinden auch alle Richtungsableitungen von u, also ist u konstant (Übungsaufgabe 4 von Blatt 4). Ist $\tilde{F}(x) = F(x, y(x)) = C$, dann ist

$$0 = \frac{d\tilde{F}}{dx}(x) = \frac{\partial F}{\partial x}(x, y) + \frac{\partial F}{\partial y}(x, y)\; y' = g(x, y) + h(x, y)\, y'\,,$$

d.h. aus $F(x, y(x)) = C$ wird die Differentialgleichung zurückerhalten. Durch lokale Auflösung der Gleichung $F(x, y(x)) = C$ mittels des Satzes über implizite Funktionen kann die Kurvenschar explizit konstruiert werden.

Beispiel 11.8 Die Differentialgleichung $2y\,y' + 2x = 0$ ist exakt, denn $F(x,y) = x^2 + y^2$ ist Stammfunktion. Die Niveaukurven $F(x,y) = R^2$ sind konzentrische Kreise um den Nullpunkt. ◁

Satz 11.9 (Notwendige Bedingung für Exaktheit) *Auf einer offenen und zusammenhängenden Teilmenge $G \subset \mathbb{R}^2$ sei die Differentialgleichung $g(x,y) + h(x,y)\,y' = 0$ gegeben mit stetig differenzierbaren Abbildungen $g, h : G \to \mathbb{R}$. Ist die Differentialgleichung exakt, so gilt*

$$\frac{\partial g}{\partial y}(x,y) = \frac{\partial h}{\partial x}(x,y) \qquad \text{für alle } (x,y) \in G\ .$$

Beweis. Nach Definition gibt es eine stetig differenzierbare Stammfunktion $F : G \to \mathbb{R}$ mit $g = \frac{\partial F}{\partial x}$ und $h = \frac{\partial F}{\partial y}$. Da g, h stetig differenzierbar sind, ist F sogar zweimal stetig differenzierbar. Nach dem Satz von Schwarz gilt dann

$$\frac{\partial}{\partial x} h = \frac{\partial}{\partial x}\frac{\partial F}{\partial y} = \frac{\partial}{\partial y}\frac{\partial F}{\partial x} = \frac{\partial}{\partial y} g\ . \qquad \square$$

Wir kommen nun zur Berechnung der Stammfunktion.

Satz 11.10 *Es seien $I, J \subset \mathbb{R}$ Intervalle und $G = I \times J$. Die Abbildungen $g, h : I \times J \to \mathbb{R}$ seien stetig differenzierbar, und es gelte $\frac{\partial g}{\partial y}(x,y) = \frac{\partial h}{\partial x}(x,y)$ für alle $(x,y) \in I \times J$. Sei $(x_0, y_0) \in I \times J$ ein beliebiger Anfangspunkt. Dann ist*

$$F(x,y) := \int_{x_0}^{x} dt\ g(t,y) + \int_{y_0}^{y} ds\ h(x_0,s)$$

eine Stammfunktion der Differentialgleichung $g(x,y) + h(x,y)y' = 0$.

Beweis. Nach dem Hauptsatz der Differential- und Integralrechnung gilt $\frac{\partial F}{\partial x} = g(x,y)$. Wegen Stetigkeit von g können wir unter dem Integral differenzieren, so daß außerdem gilt

$$\frac{\partial F}{\partial y} = \int_{x_0}^{x} dt\ \frac{\partial g}{\partial y}(t,y) + h(x_0,y) = \int_{x_0}^{x} dt\ \frac{\partial h}{\partial t}(t,y) + h(x_0,y) = h(x,y)\ . \qquad \square$$

Ohne Beweis bemerken wir, daß man die Forderung an G stark abschwächen kann: Es genügt, daß $G \subset \mathbb{R}^2$ *einfach zusammenhängend* ist, d.h. daß sich ein beliebiger geschlossener Weg in G auf einen Punkt zusammenziehen läßt. Ist dann $(x_0, y_0) \in G$ ein beliebiger Anfangspunkt und $\gamma : [a,b] \to G$ eine beliebige (stückweise) differenzierbare Kurve mit $\gamma(a) = (x_0, y_0)$ und $\gamma(b) = (x,y)$, so ist

$$F(x,y) = \int_a^b dt \Big\langle \big(g(\gamma(t)), h(\gamma(t))\big), \dot\gamma(t) \Big\rangle$$

eine Stammfunktion der exakten Differentialgleichung $g(x,y) + h(x,y)\,y' = 0$ mit $\frac{\partial g}{\partial y}(x,y) = \frac{\partial h}{\partial x}(x,y)$.

12 Existenz- und Eindeutigkeitssatz

Wir betrachten im weiteren die Differentialgleichung $y' = f(x, y)$ mit $y \in \mathbb{R}^n$. Genauer sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $f : G \to \mathbb{R}^n$ eine stetige Abbildung, gesucht ist eine Lösung $y : I \to \mathbb{R}^n$ der Differentialgleichung $y' = f(x, y)$, so daß $(y(x), x) \in G$ für alle $x \in I \subset \mathbb{R}$. Meist sucht man Lösungen mit vorgegebener Anfangsbedingung $y(x_0) = y_0$. Diese Problemstellung heißt *System von n Differentialgleichungen 1. Ordnung.*

Viele interessantere Differentialgleichungen lassen sich auf ein solches System überführen. Sei z.B. eine explizite gewöhnliche Differentialgleichung n-ter Ordnung gegeben, $z^{(n)} = \tilde{f}(x, z, z', \ldots, z^{(n-1)})$, dann setzen wir

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} := \begin{pmatrix} z \\ z' \\ \vdots \\ z^{(n-2)} \\ z^{(n-1)} \end{pmatrix} \quad \Rightarrow \quad y' = \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_n \\ \tilde{f}(x, y_1, \ldots, y_n) \end{pmatrix} = f(x, y) \; .$$

Wir lösen das System von Differentialgleichungen durch Zurückführen auf eine *Integralgleichung.*

Satz 12.1 *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ eine stetige Abbildung sowie $(x_0, y_0) \in G$. Dann gilt: Eine stetige Abbildung $y : I \to \mathbb{R}^n$ eines Intervalls $I \subset \mathbb{R}$ mit $x_0 \in I$ und $(x, y(x)) \in G$ für alle $x \in I$ löst genau dann die Differentialgleichung $y' = f(x, y)$ zur Anfangsbedingung $y(x_0) = y_0$, wenn folgende Integralgleichung gilt:*

$$y(x) = y_0 + \int_{x_0}^{x} dt \; f(t, y(t)) \qquad \text{für alle } x \in I \; .$$

Beweis. ($\Leftarrow$) Für $x = x_0$ folgt $y(x_0) = y_0$. Wegen der Stetigkeit von y und f ist auch die Abbildung $F : I \to \mathbb{R}$ mit $F(t) := f(t, y(t))$ stetig auf I. Damit gilt nach dem Fundamentalsatz der Differential- und Integralrechnung

$$\frac{d}{dx} \int_{x_0}^{x} dt \; f(t, y(t)) = f(x, y(x)) \; ,$$

d.h. y ist sogar differenzierbar mit $y'(x) = f(x, y(x))$.

($\Rightarrow$) Es gilt

$$\int_{x_0}^{x} dt \; f(t, y(t)) = \int_{x_0}^{x} dt \; y'(t) = y(x) - y(x_0) = y(x) - y_0 \; ,$$

d.h. die Integralgleichung ist erfüllt. □

Es wird sich zeigen, daß Existenz und Eindeutigkeit der Lösung bewiesen werden kann, wenn die Funktion f einer *Lipschitz-Bedingung* genügt.

Definition 12.2 Sei $G \subset \mathbb{R} \times \mathbb{R}^n$. Eine Abbildung $f : G \to \mathbb{R}^n$ genügt einer *Lipschitz-Bedingung* (bezüglich der Variablen y), wenn es ein $L \geq 0$ gibt, so daß

$$\|f(x,y) - f(x,\tilde{y})\| \leq L\|y - \tilde{y}\| \qquad \text{für alle } (x,y),(x,\tilde{y}) \in G \; .$$

Die Funktion f genügt einer *lokalen Lipschitz-Bedingung*, wenn jeder Punkt $(x_0, y_0) \in G$ eine Umgebung $U \subset \mathbb{R} \times \mathbb{R}^n$ besitzt, so daß f in $G \cap U$ einer Lipschitz-Bedingung mit möglicherweise von U abhängiger Lipschitz-Konstanten $L(U)$ genügt.

Ein nützliches Kriterium ist stetige partielle Differenzierbarkeit nach y:

Satz 12.3 *Sei $G \subset \mathbb{R}\times\mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n$ stetig partiell differenzierbar in den letzten n Koordinatenrichtungen $y_1, \ldots, y_n$. Dann genügt f in G lokal einer Lipschitz-Bedingung.*

Beweis. Wegen der Offenheit von G gibt es zu beliebigem $(x_0, y_0) \in G$ ein $r > 0$, so daß

$$V := \{(x,y) \in \mathbb{R} \times \mathbb{R}^n \; : \; |x - x_0| \leq r \, , \; \|y - y_0\| \leq r\} \subset G \; .$$

Die Behauptung folgt nun aus dem Schrankensatz 5.12 für V, und die Lipschitz-Konstante ist $L := \sup_{(x,y)\in V} \big\|(Df)(x,y)\big\|$. □

Satz 12.4 (Eindeutigkeitssatz) *Es sei $G \subset \mathbb{R}\times\mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ eine stetige Abbildung, die lokal einer Lipschitz-Bedingung genügt. Sind $y, \tilde{y}$ zwei Lösungen derselben Differentialgleichung $y' = f(x,y)$ und $\tilde{y}' = f(x,\tilde{y})$ über einem Intervall $I \subset \mathbb{R}$, und in einem Punkt $x_0 \in I$ gelte $y(x_0) = \tilde{y}(x_0)$, so folgt*

$$y(x) = \tilde{y}(x) \qquad \textit{für alle } x \in I \; .$$

Beweis. Die Eindeutigkeit wird stückweise vom Punkt x_0 aus ausgedehnt.

i) Sei dazu $y(a) = \tilde{y}(a)$ für ein $a \in I$. Dann gilt für alle $x \in I$ die Integralgleichung

$$y(x) - \tilde{y}(x) = \int_a^x dt \, \big(f(t, y(t)) - f(t, \tilde{y}(t))\big) \; .$$

Da f lokal einer Lipschitz-Bedingung genügt, gibt es reelle Zahlen $L \geq 0$ und $\delta > 0$, so daß

$$\|f(t, y(t)) - f(t, \tilde{y}(t))\| \leq L\|y(t) - \tilde{y}(t)\| \qquad \text{für alle } t \in I \cap K_\delta(a) \; .$$

Wie üblich ist $K_\delta(a) := \{t \in \mathbb{R} \; : \; |t - a| < \delta\}$. Sei nun $\epsilon := \min(\delta, \frac{1}{2L})$ und

$$M := \sup_{t \in I \cap K_\epsilon(a)} \|y(t) - \tilde{y}(t)\| \; .$$

Dann folgt für alle $x \in I \cap K_\epsilon(a)$

$$\|y(x) - \tilde{y}(x)\| \leq L\Big| \int_a^x dt \, \big\|y(t) - \tilde{y}(t)\big\|\Big| \leq LM\epsilon \leq \frac{M}{2} \; .$$

Das bedeutet $M = \sup_{x\in I\cap K_\epsilon} \|y(x) - \tilde{y}(x)\| \leq \frac{M}{2}$ und damit $M = 0$, d.h. die Funktionen $y, \tilde{y}$ stimmen sogar auf $I \cap K_\epsilon(a)$ überein.

ii) Wir zeigen nun $y(x) = \tilde{y}(x)$ für alle $x \in I$ mit $x \geq x_0$. Dazu sei

$$x_1 := \sup\left\{\xi \in I \ : \ y(\xi) = \tilde{y}(\xi)\right\} .$$

Für $x_1 = +\infty$ oder x_1 gleich dem Intervallende ist nichts mehr zu zeigen. Ansonsten gibt es ein $\delta > 0$ mit $]x_1, x_1 + \delta[\subset I$. Da $y, \tilde{y}$ als Lösungen der Differentialgleichung insbesondere stetig sind, gilt $y(x_1) = \tilde{y}(x_1)$. Dann aber gibt es nach i) ein $\epsilon > 0$, so daß $y(x) = \tilde{y}(x)$ für alle $x \in I \cap K_\epsilon(x_1)$, im Widerspruch zur Definition von x_1. Analog wird der Fall $x \leq x_0$ behandelt. □

Satz 12.5 (Picard-Lindelöf) *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n$ eine stetige Abbildung, die lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jedem $(x_0, y_0) \in G$ ein $\epsilon > 0$ und eine Lösung $y : [x_0 - \epsilon, x_0 + \epsilon]$ der Differentialgleichung $y' = f(x, y)$ mit Anfangsbedingung $y(x_0) = y_0$. Nach Satz 12.4 ist diese Lösung eindeutig.*

Beweis. Der Beweis verwendet den Banachschen Fixpunktsatz für die zugehörige Integralgleichung. Dazu sei $\mathcal{C}([x_0 - \epsilon, x_0 + \epsilon], \mathbb{R}^n)$ der Banachraum der stetigen Abbildungen $\phi : [x_0 - \epsilon, x_0 + \epsilon] \to \mathbb{R}^n$ mit der Supremumsnorm

$$\|\phi\| := \sup_{x\in[x_0-\epsilon,x_0+\epsilon]} |\phi(x)| .$$

(Die Vollständigkeit beweist man unter Verwendung von Satz 1.15 wie im Satz 25.7 aus dem letzten Semester.)

Gewählt sei ein Quader

$$Q_{\delta,r} := \{(x, y) \in \mathbb{R} \times \mathbb{R}^n \ : \ |x - x_0| \leq \delta \ , \ \|y - y_0\| \leq r\}$$

mit $\delta, r > 0$, so daß die Abbildung f in $Q_{\delta,r}$ einer Lipschitz-Bedingung mit der Lipschitz-Konstanten L genügt. Da f stetig und $Q_{\delta,r}$ kompakt ist, gibt es eine reelle Zahl $M > 0$, so daß $\|f(x, y)\| \leq M$ für alle $(x, y) \in Q_{\delta,r}$. Mit diesen Daten sei $\epsilon := \min(\delta, \frac{r}{M}, \frac{1}{2L})$. Wir betrachten die abgeschlossene Teilmenge

$$A := \{\phi \in \mathcal{C}([x_0 - \epsilon, x_0 + \epsilon], \mathbb{R}^n) \ , \ \|\phi - y_0\| \leq r\}$$

des Banachraums. Wir zeigen, daß

$$(T(\phi))(x) := y_0 + \int_{x_0}^{x} dt \ f(t, \phi(t))$$

eine Kontraktion $T : A \to A$ definiert.

i) Zunächst ist $T(\phi) \in A$ zu zeigen für alle $\phi \in A$. Nach Konstruktion ist

$$(t, \phi(t)) \in Q_{\delta,r} \subset G \qquad \text{für alle } t \in [x_0 - \epsilon, x_0 + \epsilon] .$$

Damit ist f sinnvoll erklärt und stetig in t. Es gilt

$$\|(T(\phi))(x) - y_0\| = \Big\| \int_{x_0}^{x} dt\ f(t, \phi(t)) \Big\| \leq |x - x_0|\, M \leq \epsilon M \leq r\ ,$$

also $T(\phi) \in A$.

ii) Seien $\phi_1, \phi_2 \in A$, dann gilt

$$\begin{aligned} \|T(\phi_1) - T(\phi_2)\| &= \Big\| \int_{x_0}^{x} dt\ \big(f(t, \phi_1(t)) - f(t, \phi_2(t)) \Big\| \\ &\leq |x - x_0|\, L \sup_{t \in [x_0, x]} |\phi_1(t) - \phi_1(t)| \leq \frac{1}{2} \|\phi_1 - \phi_2\|\ . \end{aligned}$$

Also ist $T : A \to A$ eine Kontraktion auf einer abgeschlossenen Teilmenge eines Banachraums. Nach dem Banachschen Fixpunktsatz gibt es damit einen Fixpunkt $y \in A$ mit $T(y) = y$, also

$$y(x) = y_0 + \int_{x_0}^{x} dt\, f(t, y(t))\ ,$$

d.h. eine (eindeutige) Lösung der Differentialgleichung. □

Es sei betont, daß die Größe des Intervalls 2ϵ durch die Lipschitz-Bedingung an f und die Norm $\|f\|$ bestimmt ist. Wenn $\|f(x, y)\|$ mit y also anwächst, wird die Fortsetzbarkeit der Lösung immer kleiner.

Beispiel 12.6 Betrachtet werde die Differentialgleichung $y' = 2xy^2$. Die Funktion $f(x, y) = 2xy^2$ ist auf ganz $\mathbb{R}^2$ nach y partiell differenzierbar, genügt also auf ganz $\mathbb{R}^2$ einer lokalen Lipschitz-Bedingung. Die eindeutige Lösung zu $y(0) = 0$ ist offenbar $y = 0$. Sei dann $y(0) > 0$. Da $G = \mathbb{R} \times \mathbb{R}_+$ einfach zusammenhängend ist, ist die entsprechende Differentialgleichung $2x - \frac{1}{y^2}\, y' = 0$ auf G exakt mit Stammfunktion

$$F(x, y) = \int_{x_0}^{x} dt\, (2t) - \int_{y_0}^{y} \frac{ds}{s^2} = x^2 - x_0^2 + \frac{1}{y} - \frac{1}{y_0}\ .$$

Die Niveaukurven sind also durch $x^2 + \frac{1}{y} = C$ gegeben, d.h. $y = \frac{1}{C - x^2}$, insbesondere $y(0) = \frac{1}{C} > 0$. Die Lösung läßt sich also nur bis zu $]-\sqrt{C}, \sqrt{C}[$ fortsetzen. Im umgekehrten Fall $y(0) < 0$ ergibt sich $y = -\frac{1}{C + x^2}$ mit $y(0) = -\frac{1}{C} < 0$ so daß die Lösung auf ganz $\mathbb{R}$ fortgesetzt werden kann. ◁

Ohne Beweis erwähnen wir, daß die *Existenz* einer Lösung der Differentialgleichung $y' = f(x, y)$ bereits durch die *Stetigkeit* von f garantiert ist:

Satz 12.7 (Peano) *Es sei $x_0 \in \mathbb{R}$ und $y_0 \in \mathbb{R}^n$ und*

$$Q_{\delta,r} := \{(x,y) \in \mathbb{R} \times \mathbb{R}^n \ : \ |x - y_0| \le \delta \ , \ \|y - y_0\| \le r\}$$

für $\delta, r > 0$. Die Funktion $f : Q_{\delta,r} \to \mathbb{R}^n$ sei stetig (also beschränkt) mit $M := \max_{(x,y)\in Q_{\delta,r}} \|f(x,y)\|$. Dann gibt es mindestens eine Lösung $y : [x_0 - \epsilon, x_0 + \epsilon]$ der Differentialgleichung $y' = f(x,y)$ mit Anfangsbedingung $y(x_0) = y_0$, wobei $\epsilon := \min(\delta, \frac{r}{M})$ ist (mit $\epsilon := \delta$ für $M = 0$).

Der Beweis ohne die Lipschitz-Bedingung ist aber aufwendiger und wird deshalb weggelassen.

Beispiel 12.8 Betrachtet werde die Differentialgleichung $y' = (y^2)^{\frac{1}{3}}$ mit $y(0) = 0$. Die Funktion $f(x,y) = (y^2)^{\frac{1}{3}}$ ist stetig auf ganz $\mathbb{R}^2$, aber nicht stetig partiell differenzierbar in $y = 0$, und tatsächlich genügt f in keiner Umgebung von $(0,0)$ einer lokalen Lipschitz-Bedingung. Nach dem Existenzsatz von Peano gibt es lokal eine Lösung der Differentialgleichung, z.B. $y(x) = 0$ für alle $x \in \mathbb{R}$. Das ist aber nicht die einzige Lösung. Durch Trennen der Variablen findet man, daß $y_c(x) := \frac{1}{27}(x - c)^3$ die Differentialgleichung zur Anfangsbedingung $y(c) = 0$ erfüllt. Durch Verkleben mit der Nullösung konstruieren wir zu $c, c' \ge 0$ die Lösung

$$y_{cc'}(x) := \begin{cases} \frac{1}{27}(x-c)^3 & \text{für } x \ge c \ge 0 \\ 0 & \text{für } -c' \le x \le c \\ \frac{1}{27}(x+c')^3 & \text{für } x \le -c' \le 0 \end{cases}$$

Man rechnet nach, daß $y_{cc'}$ tatsächlich stetig differenzierbar ist. ◁

Das Verfahren von Picard-Lindelöf über den Banachschen Fixpunktsatz ist konstruktiv und kann deshalb leicht numerisch implementiert werden.

Beispiel 12.9 Gegeben sei die Differentialgleichung $y' = 2xy$ auf $G = \mathbb{R} \times \mathbb{R}$ mit Anfangsbedingung $y(0) = y_0$. In einer gewissen Umgebung $[-\epsilon, \epsilon]$ von 0 konvergiert deshalb die Folge $(y_k(x))_{k\in\mathbb{N}}$ von Funktionen mit $y_0(x) = y_0$ und

$$y_{k+1}(x) = y_0 + \int_0^x dt \ f(t, y_k(t))$$

gegen die eindeutige Lösung. Wir finden

$$\begin{aligned} y_1(x) &= y_0 + \int_0^x dt \ 2ty_0 = y_0(1 + x^2) \\ y_2(x) &= y_0 + \int_0^x dt \ 2ty_0(1 + t^2) = y_0\Big(1 + x^2 + \frac{x^4}{2}\Big) \\ y_k(x) &= y_0\Big(1 + x^2 + \frac{x^4}{2} + \cdots + \frac{x^{2k}}{k!}\Big) \end{aligned}$$

Damit ergibt sich die Lösung $y(t) = \lim_{k\to\infty} y_k(x) = y_0\, e^{x^2}$, was man natürlich auch mit elementaren Methoden finden kann. ◁

13 Lineare Differentialgleichungen

Wir spezifizieren nun den Existenz- und Eindeutigkeitssatz auf Systeme linearer Differentialgleichungen $y' = A(x)y + b(x)$. Dabei ist $A(x)$ eine Matrix und $b(x)$ ein Vektor. Im Hinblick auf die Diagonalisierbarkeit von A erweist es sich als nützlich, mit komplexen Matrizen zu arbeiten. Sei dazu $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Die Variable x bleibt aber reell.

Definition 13.1 Es sei $I \subset \mathbb{R}$ ein Intervall, $A = (a_{ij}) : I \to M(n, \mathbb{K})$ eine stetige Abbildung und $b = (b_1, \ldots, b_n)^t : I \to \mathbb{K}^n$ eine stetige vektorwertige Funktion. Dann heißt $y' = A(x)y + b(x)$ ein *inhomogenes lineares Differentialgleichungssystem* und $y' = A(x)y$ ein (bzw. das zugehörige) *homogene(s) lineare(s) Differentialgleichungssystem*.

Gesucht sind Lösungen $y : I \to \mathbb{K}^n$, wobei für $\mathbb{K} = \mathbb{C} \simeq \mathbb{R}^2$ die Differenzierbarkeit komponentenweise betrachtet wird.

Satz 13.2 *Es sei $I \subset \mathbb{R}$ ein Intervall, $A = (a_{ij}) : I \to M(n, \mathbb{K})$ eine stetige Abbildung und $b = (b_1, \ldots, b_n)^t : I \to \mathbb{K}^n$ eine stetige vektorwertige Funktion. Dann gibt es zu jedem $x_0 \in I$ und jedem $y_0 \in \mathbb{K}^n$ genau eine Lösung $y : I \to \mathbb{K}^n$ der linearen Differentialgleichung $y' = A(x)y + b(x)$ mit Anfangsbedingung $y(x_0) = y_0$.*

Beweis. i) Wir setzen $f(x, y) = A(x)y + b(x)$. Sei $J \subset I$ ein kompaktes Intervall, dann folgt aus der Stetigkeit von A

$$L := \sup_{x \in J} \|A(x)\| < \infty$$

Somit gilt für beliebige $x \in J$ und $y, \tilde{y} \in \mathbb{K}^n$

$$\|f(x, y) - f(x, \tilde{y})\| = \|A(x)(y - \tilde{y})\| \leq L\|y - \tilde{y}\| \ .$$

Folglich genügt $f(x, y)$ auf ganz $J \times \mathbb{K}^n$ einer globalen Lipschitz-Bedingung, so daß nach Satz 12.4 die Lösung auf $J \times \mathbb{K}^n$ eindeutig ist.

ii) Nach Picard-Lindelöf existiert die Lösung in einer Umgebung von x_0. Zu zeigen bleibt, daß sie auf ganz I existiert. Dazu definieren wir Folgen von Abbildungen $y_k : I \to \mathbb{K}^n$ durch

$$y_{k+1}(x) := y_0 + \int_{x_0}^{x} dt \ f(t, y_k(t)) \ , \qquad y_0(x) = y_0 \ .$$

Die so konstruierten Abbildungen y_k sind stetig auf I. Also gilt

$$M := \sup_{x \in J} \|y_1(x) - y_0(x)\| < \infty \ .$$

Wir zeigen mit vollständiger Induktion

$$\|y_{k+1}(x) - y_k(x)\| \leq M \frac{L^k \|x - x_0\|^k}{k!} \ .$$

Für $k = 0$ ist das die Definition von M. Der Induktionsschritt ist

$$\begin{aligned}
\|y_{k+2}(x) - y_{k+1}(x)\| &= \Big\| \int_{x_0}^{x} dt \, \big(f(t, y_{k+1}(t)) - f(t, y_k(t)) \big) \Big\| \\
&\leq \Big| \int_{x_0}^{x} dt \, \big\| f(t, y_{k+1}(t)) - f(t, y_k(t)) \big\| \Big| \\
&\leq L \Big| \int_{x_0}^{x} dt \, \big\| y_{k+1}(t) - y_k(t) \big\| \Big| \\
&\leq L \Big| \int_{x_0}^{x} dt \, M \frac{L^k |t - x_0|^k}{k!} \Big| = M \frac{L^{k+1} |x - x_0|^{k+1}}{(k+1)!} \ .
\end{aligned}$$

Setzen wir $r := \sup_{x \in J} |x - x_0|$, dann ist die Reihe $\sum_{k=0}^{\infty} (y_{k+1}(x) - y_k(x))$ absolut konvergent auf J, da sie durch die Reihe $M \sum_{k=0}^{\infty} \frac{L^k r^k}{k!} = M e^{Lr}$ majorisiert wird. Somit konvergiert die Folge (y_k) auf jedem kompakten Teilintervall $J \subset I$ gleichmäßig gegen die Grenzfunktion

$$y := \lim_{k \to \infty} y_k = y_0 + \sum_{k=0}^{\infty} (y_{k+1} - y_k) \ ,$$

welche somit stetig auf J und damit auf I ist. Diese Grenzfunktion erfüllt nach Konstruktion die Integralgleichung

$$y(x) := y_0 + \int_{x_0}^{x} dt \, f(t, y(t))$$

und löst deshalb auf I die Differentialgleichung $y' = A(x)y + b(x)$ mit Anfangsbedingung $y(x_0) = y_0$. □

Ähnlich wie in der linearen Algebra bilden die Lösungen der homogenen Differentialgleichung, *wenn man keine Anfangsbedingung stellt*, einen Vektorraum:

Satz 13.3 *Es sei $I \subset \mathbb{R}$ ein (nicht-triviales) Intervall und $A : I \to M(n, \mathbb{K})$ stetig. Dann bildet die Menge V_A aller Lösungen der homogenen Differentialgleichung $y' = A(x)y$ einen n-dimensionalen Vektorraum über $\mathbb{K}$.*

Für ein k-Tupel von Lösungen $y_{(1)}, \ldots, y_{(k)} \in V_A$ sind folgende Aussagen äquivalent:

i) *Die vektorwertigen Funktionen $(y_{(1)}, \ldots, y_{(k)})$ sind linear unabhängig.*

ii) *Es gibt ein $x_0 \in I$, so daß die Vektoren $y_{(1)}(x_0), \ldots, y_{(k)}(x_0) \in \mathbb{K}^n$ linear unabhängig sind.*

iii) *Für jedes $x_0 \in I$ sind die Vektoren $y_{(1)}(x_0), \ldots, y_{(k)}(x_0) \in \mathbb{K}^n$ linear unabhängig.*

Beweis. Klar ist, daß V_A ein Untervektorraum des (unendlich-dimensionalen) Vektorraums $\mathcal{C}(I, \mathbb{K}^n)$ aller stetigen Abbildungen $f : I \to \mathbb{K}^n$ ist, denn

$$\begin{aligned}(\lambda_1 y_{(1)} + \lambda_2 y_{(2)})' &= \lambda_1 {y_{(1)}}' + \lambda_2 {y_{(2)}}' = \lambda_1(A(x)y_{(1)}) + \lambda_2(A(x)y_{(2)}) \\ &= A(x) \cdot (\lambda_1 y_{(1)} + \lambda_2 y_{(2)}) \,.\end{aligned}$$

Die Äquivalenzen iii)⇒ii)⇒i) sind klar, und i)⇒iii) folgt aus der Eindeutigkeit der Nullösung: Ist $\sum_{i=1}^k \lambda_i y_i(x_0) = 0$, so folgt $\sum_{i=1}^k \lambda_i y_i(x) = 0$ für alle $x \in I$.

Wir zeigen $\dim(V_A) = n$. Seien $e_1, \ldots, e_n \in \mathbb{K}^n$ die Standardbasisvektoren und $x_0 \in I$. Dann gibt nach dem Existenz- und Eindeutigkeitssatz eindeutige Lösungen $y_{(i)} : I \to \mathbb{K}^n$ der Differentialgleichung ${y_{(i)}}' = A(x)y_{(i)}$ mit Anfangsgbedingung $y_{(i)}(x_0) = e_i$. Die Lösungen sind linear unabhängig, da sie an der Stelle x_0 linear unabhängig sind. Jede weitere Lösung $y_{(n+1)}$ ist linear abhängig, da sie an der Stelle x_0 linear abhängig ist. □

Definition 13.4 Unter einem *Lösungs-Fundamentalsystem* der Differentialgleichung $y' = A(x)y$ versteht man eine Basis $(y_{(1)}, \ldots, y_{(n)})$ des Vektorraums V_A der Lösungen von $y' = A(x)y$.

Schreibt man das Lösungs-Fundamentalsystem als Matrix $Y = (y_{ij})$, mit $y_{ij} := y_{(j)i}$ (die j-te Spalte von Y ist $y_{(j)}$), dann gilt offenbar $\det Y(x_0) \neq 0$ für wenigstens einen (und damit für jeden) Punkt $x_0 \in I$. Eine beliebige Lösung schreibt sich damit als $y(x) = Y(x) \cdot \lambda$ mit $\lambda = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{K}^n$.

Beispiel 13.5 Gegeben sei die Schwingungsdifferentialgleichung $z'' + \omega^2 z = 0$. Wir setzen $y_1 := z$ und $y_2 := -\frac{1}{\omega} z'$, was also auf das Differentialgleichungssystem $y_1' = -\omega y_2$ und $y_2' = \omega y_1$ führt. In Matrixschreibweise ergibt sich $y' = Ay$ mit $A = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$. Die Lösungstheorie dieses Problems mit konstanten Koeffizienten behandeln wir später. Man bestätigt durch Nachrechnen, daß folgende Funktionen $y_i : \mathbb{R} \to \mathbb{R}^2$ Lösungen der Differentialgleichung sind:

$$y_{(1)}(x) = \begin{pmatrix} \cos \omega x \\ \sin \omega x \end{pmatrix} \,, \qquad y_{(2)}(x) = \begin{pmatrix} -\sin \omega x \\ \cos \omega x \end{pmatrix} \,.$$

Diese sind in $x_0 = 0$ und damit auf ganz $\mathbb{R}$ linear unabhängig, was man auch durch

$$\det Y(x) = \det \begin{pmatrix} \cos \omega x & -\sin \omega x \\ \sin \omega x & \cos \omega x \end{pmatrix} = 1$$

sieht. ◁

Wir charakterisieren nun den Lösungsraum der inhomogenen Differentialgleichung:

Satz 13.6 *Es sei $I \subset \mathbb{R}$ ein (nicht-triviales) Intervall und $A : I \to M(n, \mathbb{K})$ und $b : I \to \mathbb{K}^n$ stetige Abbildungen. Dann gilt: Die Menge aller Lösungen*

$$L_{A,b} := \{y : I \to \mathbb{K}^n \ : \ y' = A(x)y + b(x)\}$$

der inhomogenen Differentialgleichung ist der affine Raum

$$L_{A,b} = \tilde{y} + V_A \ ,$$

wobei V_A der Vektorraum aller Lösungen der zugehörigen homogenen Differentialgleichung $y' = A(x)y$ und $\tilde{y} \in L_{A,b}$ eine beliebige Lösung der inhomogenen Differentialgleichung ist.

Mit anderen Worten: die allgemeine Lösung der inhomogenen Differentialgleichung ist die Summe aus einer speziellen Lösung und der allgemeinen Lösung der homogenen Differentialgleichung.

Beweis. i) Sei $y \in L_{A,b}$ eine Lösung von $y' = A(x)y + b(x)$, dann erfüllt $u := y - \tilde{y}$ die homogene Differentialgleichung $u' = A(x)u$. Das bedeutet $u \in V_A$, also $y \in \tilde{y} + V_A$ und damit $L_{A,b} \subset \tilde{y} + V_A$.

ii) Sei umgekehrt $y \in \tilde{y} + V_A$ gegeben, also $y = \tilde{y} + u$ mit $u' = A(x)u$. Dann gilt

$$y' = \tilde{y}' + u' = (A(x)\tilde{y} + b(x)) + A(x)u(x) = A(x)y + b(x) \ ,$$

also $\tilde{y} + V_A \subset L_{A,b}$. □

Die spezielle Lösung $\tilde{y}$ bekommt man wie in den elementaren Lösungsmethoden durch "Variation der Konstanten":

Satz 13.7 (Variation der Konstanten) *Sei $Y = (y_{(1)}, \ldots, y_{(n)}) : I \to GL(n, \mathbb{K})$ ein Lösungsfundamentalsystem der homogenen Differentialgleichung $y' = A(x)y$. Dann wird eine Lösung $\tilde{y}$ der inhomogenen Differentialgleichung $\tilde{y}' = A(x)\tilde{y} + b(x)$ erhalten durch den Ansatz $\tilde{y}(x) = Y(x) \cdot v(x)$, wobei die differenzierbare Abbildung $v : I \to \mathbb{K}^n$ der Differentialgleichung $Y(x)v' = b(x)$ genügt. Ihre Lösung ist somit*

$$v(x) = \int_{x_0}^{x} dt \; Y^{-1}(t) \cdot b(t) + const \, .$$

Beweis. Es gilt

$$\tilde{y}'(x) = Y'(x) \cdot v(x) + Y(x)v'(x) = A(x) \cdot Y(x) \cdot v(x) + Y(x)v'(x) \stackrel{!}{=} A(x)\tilde{y}(x) + b(x) \ ,$$

also $v'(x) = Y^{-1}(x) \cdot b(x)$. Umgeschreiben in eine Integralgleichung ergibt sich die Lösung. □

Beispiel 13.8 Gegeben sei die Schwingungsdifferentialgleichung mit periodischer äußerer Kraft $z'' + \omega^2 z = b\cos(\Omega x)$. Mit $y_1 := z$ und $y_2 := -\frac{1}{\omega}z'$ ergibt sich $y_1' = -\omega y_2$ und $y_2' = \omega y_1 - \frac{b}{\omega}\cos(\Omega x)$. Die spezielle Lösung ist für $\Omega \neq \pm\omega$ damit

$$\begin{aligned}
v(x) &= \int_{x_0}^{x} dt \begin{pmatrix} \cos\omega t & \sin\omega t \\ -\sin\omega t & \cos\omega t \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -\frac{b}{\omega}\cos\Omega t \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
&= -\frac{b}{2\omega}\int_{x_0}^{x} dt \begin{pmatrix} \sin(\omega t+\Omega t) + \sin(\omega t-\Omega t) \\ \cos(\omega t+\Omega t) + \cos(\omega t-\Omega t) \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
&= \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} - \frac{b}{2\omega(\omega+\Omega)} \begin{pmatrix} -\cos(\omega x+\Omega x) \\ \sin(\omega x+\Omega x) \end{pmatrix} - \frac{b}{2\omega(\omega-\Omega)} \begin{pmatrix} -\cos(\omega x-\Omega x) \\ \sin(\omega x-\Omega x) \end{pmatrix} \\
&= \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} + \frac{b}{\omega(\omega^2-\Omega^2)} \begin{pmatrix} \omega\cos\omega x\cos\Omega x + \Omega\sin\omega x\sin\Omega x \\ -\omega\sin\omega x\cos\Omega x + \Omega\cos\omega x\sin\Omega x \end{pmatrix}.
\end{aligned}$$

Damit ergibt sich

$$y(x) = Y(x)\cdot v(x) = \begin{pmatrix} c_1'\cos\omega x - c_2'\sin\omega x \\ c_1'\sin\omega x + c_2'\cos\omega x \end{pmatrix} + \frac{b}{(\omega^2-\Omega^2)} \begin{pmatrix} \cos\Omega x \\ \frac{\Omega}{\omega}\sin\Omega x \end{pmatrix}.$$

Für $\Omega^2 \to \omega^2$ wird die Amplitude der speziellen Lösung unendlich (Resonanz). Tatsächlich muß das Integral in diesem Fall anders berechnet werden.

$$\begin{aligned}
v(x) &= -\frac{b}{2\omega}\int_{x_0}^{x} dt \begin{pmatrix} \sin(2\omega t) \\ \cos(2\omega t)+1 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
&= \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} - \frac{b}{4\omega^2} \begin{pmatrix} -\cos(2\omega x) \\ \sin(2\omega x) + 2\omega x \end{pmatrix}
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
y(x) = Y(x)\cdot v(x) &= \begin{pmatrix} c_1'\cos\omega x - c_2'\sin\omega x \\ c_1'\sin\omega x + c_2'\cos\omega x \end{pmatrix} \\
&\quad + \frac{b}{4\omega^2} \begin{pmatrix} \cos\omega x\cos 2\omega x + \sin\omega x\sin 2\omega x + 2\omega x\sin\omega x \\ \sin\omega x\cos 2\omega x - \cos\omega x\sin 2\omega x - 2\omega x\cos\omega x \end{pmatrix} \\
&= \begin{pmatrix} (c_1' + \frac{b}{4\omega^2})\cos\omega x - c_2'\sin\omega x \\ (c_1' - \frac{b}{4\omega^2})\sin\omega x + c_2'\cos\omega x \end{pmatrix} + \frac{b}{2\omega} \begin{pmatrix} x\sin\omega x \\ -x\cos\omega x \end{pmatrix}.
\end{aligned}$$

Die Amplitude wächst also mit der Zeit x an. Wir werden später eine einfachere Berechnungsmethode für diese Differentialgleichung angeben. ◁

Wir übertragen nun die Aussagen zu linearen Differentialgleichungssystemen auf lineare Differentialgleichungen n-ter Ordnung.

Definition 13.9 Sei $I \subset \mathbb{R}$ ein Intervall und $b, a_k : I \to \mathbb{K}$ stetige Funktionen für $k = 0, 1, \ldots, n-1$. Dann heißt

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_0(x)y = b(x)$$

eine *lineare Differentialgleichung n-ter Ordung*. Sie heißt *homogen* für $b = 0$, sonst *inhomogen*.

Satz 13.10 *In den Bezeichnungen von Definition 13.9 gilt:*

i) *Die Menge V_a aller Lösungen der homogenen linearen Differentialgleichung n-ter Ordnung ist ein n-dimensionaler Vektorraum über $\mathbb{K}$.*

ii) *Die Menge $L_{a,b}$ aller Lösungen der inhomogenen linearen Differentialgleichung n-ter Ordnung ist der affine Raum $L_{a,b} = \tilde{y} + V_a$, wobei $\tilde{y}$ eine beliebige Lösung der inhomogenen Differentialgleichung ist.*

iii) *Ein n-Tupel $(y_{(1)}, \dots, y_{(n)})$ von Lösungen der homogenen Differentialgleichung ist genau dann linear unabhängig, wenn in einem (und damit jedem) Punkt $x \in I$ die "Wronski-Determinante"*

$$W(x) := \det \begin{pmatrix} y_{(1)}(x) & y_{(2)}(x) & \dots & y_{(n)}(x) \\ y'_{(1)}(x) & y'_{(2)}(x) & \dots & y'_{(n)}(x) \\ \vdots & \vdots & \ddots & \vdots \\ y^{(n-1)}_{(1)}(x) & y^{(n-1)}_{(2)}(x) & \dots & y^{(n-1)}_{(n)}(x) \end{pmatrix}$$

ungleich Null ist.

Beweis. Alle Aussagen folgen sofort aus der zu Beginn des Abschnitts 12 gegebenen Umschreibung

$$\hat{y} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{pmatrix} := \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$$

in ein System von linearen Differentialgleichungen erster Ordnung und den dafür bewiesenen Sätzen. Insbesondere ist $W(x) = \det \hat{Y}(x)$. □

Beispiel 13.11 Gegeben sei die Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 1 \qquad \text{auf } I := \mathbb{R}^*_+ \,.$$

Der Ansatz $y = x^\alpha$ führt für die zugehörigen homogene Gleichung auf

$$\Big(\alpha(\alpha-1) - \frac{\alpha}{2} + \frac{1}{2}\Big)x^{\alpha-2} = (\alpha-1)(\alpha-\tfrac{1}{2})x^{\alpha-2} = 0$$

und damit auf $y_{(1)}(x) = x$ und $y_{(2)}(x) = \sqrt{x}$. Die Wronski-Determinante ist damit

$$W(x) = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} = -\frac{1}{2}\sqrt{x}\,,$$

d.h. $(x, \sqrt{x})$ ist ein Lösungsfundamentalsystem. Um eine spezielle Lösung der inhomogenen Gleichung zu finden, nutzt man, daß für $y = cx^2$ jeder Summand der linken Seite eine Konstante ist. Damit findet man $c = \frac{2}{3}$ in der speziellen Lösung und

$$y = \frac{2}{3}x^2 + c_1 x + c_2\sqrt{x}$$

als allgemeine Lösung der Differentialgleichung. ◁

14 Lineare Differentialgleichungen mit konstanten Koeffizienten

Lineare Differentialgleichungen mit konstanten Koeffizienten können auf Eigenwertprobleme für lineare Abbildungen zurückgeführt werden, die wir in der Linearen Algebra behandelt haben.

Zur Fixierung der Bezeichnungen sei $\mathbb{C}[T]$ die Menge aller Polynome (endlicher Ordnung) in einer formalen Größe T, d.h. $P_n(T) := a_0 + a_1 T + \cdots + a_n T^n \in \mathbb{C}[T]$ für ein $n \in \mathbb{N}$ und $a_i \in \mathbb{C}$. Die Menge der Polynome $\mathbb{C}[T]$ bildet eine sogenannte Algebra, d.h. einen Vektorraum mit Produkt, wobei alle Distributiv-Gesetze gelten. Wir interessieren uns für die Menge $\mathbb{C}[\frac{d}{dx}]$ der *Differentialoperatoren*.

Ist $\mathcal{C}^k(I)$ der Vektorraum der k-mal stetig differenzierbaren komplexwertigen Funktionen $f : I \to \mathbb{C}$ auf dem Intervall $I \subset \mathbb{R}$ und $P_n(\frac{d}{dx}) = a_0 + a_1\frac{d}{dx} + \cdots + a_n\frac{d^n}{dx^n}$ mit $n \leq k$, dann ist dieser Differentialoperator n-ter Ordnung $P_n(\frac{d}{dx})$ eine lineare Abbildung

$$P_n(\tfrac{d}{dx}) : \mathcal{C}^k(I) \to \mathcal{C}^{k-n}(I) \,, \qquad f(x) \mapsto a_0 + a_1 f'(x) + \cdots + a_n f^{(n)}(x) \,.$$

Wir können $a_n = 1$ annehmen. Insbesondere ist $P_n(\frac{d}{dx})$ ein Endomorphismus des Vektorraums $\mathcal{C}^\infty(I)$ der beliebig oft differenzierbaren Funktionen. Damit läßt sich die Eigenwerttheorie von Endomorphismen eines Vektorraums auf Differentialoperatoren übertragen. Zwar ist $\mathcal{C}^\infty(I)$ ein unendlich-dimensionaler Vektorraum, aber nach Satz 13.10 ist der Vektorraum $\ker(P_n(\frac{d}{dx})) \subset \mathcal{C}^\infty(I)$ der Lösungen y von $P_n(\frac{d}{dx})y = 0$ *endlich-dimensional*.

Die gesamte Theorie der Differentialoperatoren $P_n(\frac{d}{dx})$ beruht auf der Beobachtung, daß

$$P_n(\tfrac{d}{dx})e^{\lambda x} = P_n(\lambda)e^{\lambda x} \,.$$

Folglich gilt:

Satz 14.1 *Ist $\lambda \in \mathbb{C}$ eine Nullstelle des Polynoms P, d.h. $P(\lambda) = 0$, dann ist $y = e^{\lambda x}$ eine Lösung der Differentialgleichung $P(\frac{d}{dx})e^{\lambda x} = 0$.*

Im einfachsten Fall sind alle n Nullstellen von P_n verschieden:

Satz 14.2 *Sei $P_n(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0$ ein Polynom, welches n paarweise voneinander verschiedene Nullstellen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ habe. Dann bilden die Funktionen $y_{(1)}, \ldots, y_{(n)} : \mathbb{R} \to \mathbb{C}$ mit*

$$y_{(k)}(x) := e^{\lambda_k x} , \qquad k = 1, \ldots, n ,$$

ein Fundamentalsystem von Lösungen der Differentialgleichung

$$P_n(\tfrac{d}{dx})y = y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y + a_0 = 0 .$$

Beweis. Nach Satz 14.1 ist jede dieser Funktionen $y_{(k)}$ Lösung der Differentialgleichung. Zur Überprüfung der linearen Unabhängigkeit berechnen wir die Wronski-Determinante an der Stelle $x = 0$. Mit $y_{(k)}^{(j)}(x) = \lambda_k^j e^{\lambda_k x}$ ergibt sich

$$W(x) = \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{pmatrix}$$

Das ist genau die Vandermonde-Determinate (Aufgabe 1 von Blatt 9 aus der Linearen Algebra),

$$W(x) = \prod_{k>l} (\lambda_k - \lambda_l) \neq 0 ,$$

da die Nullstellen paarweise verschieden sind. □

Man sieht aber auch, daß für mehrfache Nullstellen die Lösungen $e^{\lambda_k x}$ nicht linear unabhängig sind.

Beispiel 14.3 Gegeben sei die Differentialgleichung $y''' - 2y'' + y' - 2y = 0$. Sie schreibt sich als $P_3(\frac{d}{dx})y = 0$ mit

$$P_3(T) = T^3 - 2T^2 + T - 2 = (T-2)(T^2+1) = (T-2)(T-i)(T+i) .$$

Alle Nullstellen von P_3 sind paarweise verschieden, so daß

$$y_{(1)} = e^{ix} , \qquad y_{(2)} = e^{-ix} , \qquad y_{(3)} = e^{2x}$$

ein Lösungsfundamentalsystem bildet. Wegen

$$c_1 e^{ix} + c_2 e^{-ix} = (c_1 + c_2)\cos x + (ic_1 - ic_2)\sin x$$

bildet dann

$$y_{(1)} = \cos x , \qquad y_{(2)} = \sin x , \qquad y_{(3)} = e^{2x}$$

ein reelles Fundamentalsystem. ◁

Ganz allgemein gilt: Ist $\lambda = i\mu$ mit $\mu \in \mathbb{R}$ eine rein imaginäre Nullstelle eines rellen Polynoms $P(T) \in \mathbb{R}[T]$, dann ist auch $\bar{\lambda} = -i\mu$ eine Nullstelle, so daß sich die Lösungen $e^{\pm i\mu x}$ der entsprechenden Differentialgleichung äquivalent durch $\cos \mu x$ und $\sin \mu x$ ausdrücken lassen.

Beispiel 14.4 Wir erinnern noch einmal an die Schwingungsdifferentialgleichung $y'' + \omega^2 y = 0$. Mit $P_2(T) = T^2 + \omega^2 = (T - i\omega)(T + i\omega)$ finden wir sofort das Lösungsfundamentalsystem $y_{(1)} = \cos \omega x$ und $y_{(2)} = \sin \omega x$. ◁

Es verbleibt die Diskussion mehrfacher Nullstellen. Nach dem Fundamentalsatz der Algebra können wir jedes Polynom $P_n(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_0 \in \mathbb{C}[T]$ in Linearfaktoren zerlegen,

$$P_n(T) = (T - \lambda_1)^{k_1} \cdots (T - \lambda_r)^{k_r} \ ,$$

mit paarweise verschiedenen Nullstellen $\lambda_i \in \mathbb{C}$ und $k_1 + \cdots + k_r = n$. Wir benötigen zwei Hilfssätze:

Lemma 14.5 *Sei $\lambda \in \mathbb{C}$ und $k \in \mathbb{N}$ sowie $I \subset \mathbb{R}$. Für jede k-mal stetig differenzierbare Funktion $f : I \to \mathbb{C}$ gilt*

$$(\tfrac{d}{dx} - \lambda)^k (f(x)\, e^{\lambda x}) = f^{(k)}(x)\, e^{\lambda x} \ .$$

Beweis. Für $k = 0$ ist nichts zu zeigen, und für $k = 1$ ergibt sich

$$(\tfrac{d}{dx} - \lambda)(f(x)\, e^{\lambda x}) = f'(x)\, e^{\lambda x} + f(x)\, (e^{\lambda x})' - \lambda\, f(x)\, e^{\lambda x} = f'(x)\, e^{\lambda x} \ .$$

Der Induktionsschritt $k \mapsto k + 1$ ist

$$\begin{aligned} (\tfrac{d}{dx} - \lambda)^{k+1} (f(x)\, e^{\lambda x}) &= (\tfrac{d}{dx} - \lambda)(f^{(k)}(x)\, e^{\lambda x}) \\ &= (f^{(k)}(x))'\, e^{\lambda x}) \end{aligned}$$

nach obiger Rechnung für $k = 1$. □

Lemma 14.6 *Es sei $P_n(T) \in \mathbb{C}[T]$ ein Polynom und $\lambda \in \mathbb{C}$, so daß $P_n(\lambda) \neq 0$. Ist $g_k : \mathbb{R} \to \mathbb{C}$ eine Polynomfunktion k-ten Grades, so gilt*

$$P_n(\tfrac{d}{dx})\big(g_k(x)\, e^{\lambda x}\big) = h_k(x)\, e^{\lambda x} \ ,$$

wobei $h_k : \mathbb{R} \to \mathbb{C}$ ebenfalls eine Polynomfunktion k-ten Grades ist.

Beweis. Das Polynom $P_n(T)$ läßt sich (z.B. über den Satz von Taylor) umordnen nach Potenzen von $T - \lambda$:

$$P_n(T) = \sum_{j=0}^{n} c_i (T - \lambda)^j$$

mit $c_j \in \mathbb{C}$ und $c_0 = P_n(\lambda) \neq 0$. Damit gilt nach Lemma 14.5

$$P_n(\tfrac{d}{dx})\big(g_k(x)\,e^{\lambda x}\big) = \sum_{j=0}^{n} c_j(\tfrac{d}{dx}-\lambda)^j\big(g_k(x)\,e^{\lambda x}\big) = \sum_{j=0}^{n} c_j g_k^{(j)}(x)\,e^{\lambda x}\ .$$

Wegen $c_0 \neq 0$ ist $h_k(x) := \sum_{j=0}^{n} c_j g_k^{(j)}(x)$ wieder eine Polynomfunktion vom Grad k. □

Satz 14.7 *Es sei $P_n(T) = (T-\lambda_1)^{k_1}\cdots(T-\lambda_r)^{k_r}$ ein Polynom n-ten Grades mit paarweise verschiedenen Nullstellen $\lambda_j \in \mathbb{C}$ der Vielfachheit k_j. Dann besitzt die Differentialgleichung $P(\frac{d}{dx})y = 0$ ein Lösungs-Fundamentalsystem aus den Funktionen*

$$y_{(jm)}(x) := x^m e^{\lambda_j x}\ , \qquad 1 \le j \le r\ ,\ 0 \le m \le k_j - 1\ .$$

Beweis. i) Für gewähltes j läßt sich das Polynom P_n schreiben als $P_n(T) = Q_j(T)(T-\lambda_j)^{k_j}$ mit $Q_j(\lambda_j) \neq 0$. Dann gilt nach Lemma 14.5

$$P_n(\tfrac{d}{dx})(x^m e^{\lambda_j x}) = Q_j(\tfrac{d}{dx})(\tfrac{d}{dx}-\lambda_j)^{k_j}(x^m e^{\lambda_j x}) = Q_j(\tfrac{d}{dx})\ (x^m)^{(k_j)} e^{\lambda_j x} = 0$$

wegen $m < k_j$. Somit erfüllen alle Funktionen $y_{(jm)}$ die Differentialgleichung.

ii) Zu zeigen bleibt die lineare Unabhängigkeit der $y_{(jm)}$. Eine Linearkombination der $y_{(jm)}$ hat die Form $\tilde{y}(x) = \sum_{j=1}^{r} g_{(j)}(x)e^{\lambda_j x}$, wobei $g_{(j)}(x)$ eine Polynomfunktion vom Grad $\le k_j - 1$ ist (wir lassen den Polynomgrad zur Verbesserung der Lesbarkeit weg). Wir zeigen durch Induktion nach r, daß $\tilde{y} = 0$ genau dann, wenn $g_{(j)} = 0$ für alle j.

Für $r = 1$ folgt aus $\tilde{y}(x) = g_{(1)}(x)e^{\lambda_1 x} = 0$ für alle $x \in \mathbb{R}$, daß $g_{(1)}(x) = 0$ ist.

Im Schritt von $r-1$ nach r sei dann $\sum_{j=1}^{r} g_{(j)}(x)e^{\lambda_j x} = 0$. Ist eines der $g_{(j)}$ gleich Null, so sind wir nach Induktionsannahme fertig. Ansonsten wenden wir $(\frac{d}{dx}-\lambda_r)^{k_r}$ an und benutzen Lemma 14.5 und Lemma 14.6:

$$\begin{aligned} 0 &= (\tfrac{d}{dx}-\lambda_r)^{k_r}\Big(\sum_{j=1}^{r} g_{(j)}(x)e^{\lambda_j x}\Big) = (\tfrac{d}{dx}-\lambda_r)^{k_r}\Big(\sum_{j=1}^{r-1} g_{(j)}(x)e^{\lambda_j x}\Big) \\ &= \sum_{j=1}^{r-1}\sum_{i=0}^{k_r} c_{ji}(\tfrac{d}{dx}-\lambda_j)^i\Big(g_{(j)}(x)e^{\lambda_j x}\Big) = \sum_{j=1}^{r-1} \underbrace{\Big(\sum_{i=0}^{k_r} c_{ji} g_{(j)}^{(i)}(x)\Big)}_{h_{(j)}(x)} e^{\lambda_j x}\ , \end{aligned}$$

wobei $h_{(j)}(x)$ wegen $c_{j0} = (\lambda_j - \lambda_r)^{k_r} \neq 0$ wieder ein Polynom vom Grad $\le k_j - 1$ ist. Nach Induktionsvoraussetzung ist dann $h_{(j)} = 0$ für alle j und weiter wegen $c_{j0} \neq 0$ auch $g_{(j)}^{(0)}(x) = 0$, im Widerspruch zur Annahme. □

Beispiel 14.8 Gegeben sei die Schwingungsdifferentialgleichung mit kritischer Reibung

$$0 = y'' + 2\omega y' + \omega^2 y = P_2(\tfrac{d}{dx})(y)$$

mit $P_2(T) = T^2 + 2\omega T + \omega^2 = (T + \omega)^2$. Die allgemeine Lösung ist deshalb $y = (c_1 + c_2 x)e^{-\omega x}$. ◁

Wir betrachten nun inhomogene lineare Differentialgleichungen $P_n(\frac{d}{dx})y = b(x)$. Es bietet sich an, die Lösung der homogenen Gleichung zu bestimmen, sie in die Matrixform einer linearen Differentialgleichung 1. Ordnung zu überführen und dann durch Variation der Konstanten eine spezielle Lösung des inhomogenen Problems zu berechnen. In manchen Fällen kommt man aber durch einen geeigneten Lösungsansatz schneller ans Ziel.

Zunächst folgende Beobachtung: Ist $b(x) = b_1(x) + \cdots + b_k(x)$ und $y_{(k)}$ Lösung von $P_n(\frac{d}{dx})y_{(k)} = b_k(x)$, so ist $y = y_{(1)} + \cdots + y_{(k)}$ Lösung von $P_n(\frac{d}{dx})y = b(x)$. Der Lösungsansatz funktioniert dann für rechte Seiten der Form $b(x) = g_m(x)e^{\mu x}$, wobei g_m eine Polynomfunktion vom Grad m ist.

Satz 14.9 *Sei $P_n(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_1 T + a_0 \in \mathbb{C}[T]$ ein Polynom und $\mu \in \mathbb{C}$, so daß $P(\mu) \neq 0$ (keine Resonanz). Dann gilt:*

i) *Die Differentialgleichung $P_n(\frac{d}{dx})y = e^{\mu x}$ besitzt die spezielle Lösung $y(x) = \frac{1}{P_n(\mu)}e^{\mu x}$.*

ii) *Ist $g_m : \mathbb{R} \to \mathbb{C}$ eine Polynomfunktion vom Grad m, so besitzt die Differentialgleichung $P_n(\frac{d}{dx})y = g_m(x)e^{\mu x}$ eine spezielle Lösung der Form $y(x) = h_m(x)e^{\mu x}$, wobei $h_m : \mathbb{R} \to \mathbb{C}$ wieder eine Polynomfunktion vom Grad m ist.*

Beweis. i) ist klar wegen $P_n(\frac{d}{dx})e^{\mu x} = P_n(\mu)e^{\mu x}$.

ii) mit Induktion nach m. Der Fall $m = 0$ ist Teil i). Nach Lemma 14.6 ist $P_n(\frac{d}{dx})(x^m e^{\mu x}) = \tilde{h}_m(x)e^{\mu x}$ für eine Polynomfunktion $\tilde{h}_m(x)$ vom Grad m. Wir schreiben $g_m(x) = c\tilde{h}_m(x) + g_{m-1}(x)$ für ein $c \in \mathbb{C}$ und ein Polynom $g_{m-1}(x)$ vom Grad $m - 1$. Nach Induktionsvoraussetzung gibt es eine Polynomfunktion $h_{m-1}(x)$ vom Grad $m - 1$, so daß $P_n(\frac{d}{dx})y = g_{m-1}(x)e^{\mu x}$ die spezielle Lösung $y(x) = h_{m-1}(x)e^{\mu x}$ hat. Dann hat $P_n(\frac{d}{dx})y = g_m(x)e^{\mu x}$ die spezielle Lösung $y(x) = (cx^m + h_{m-1}(x))e^{\mu x}$. □

Beispiel 14.10 Gegeben sei die Differentialgleichung $y''' - y = x$, also $P_3(\frac{d}{dx})y = b(x)$ mit $P_3(T) = T^3 - 1 = (T - 1)(T^2 + T + 1) = (T - 1)(T - \tau_1)(T - \tau_2)$ und $b(x) = xe^{0x}$. Damit ist $P_3(0) \neq 0$, so daß es eine spezielle Lösung $y = c_1 x + c_0$ der Differentialgleichung gibt. Wir testen

$$(\frac{d^3}{dx^3} - 1)(c_1 x + c_0) = -c_1 x - c_0 \stackrel{!}{=} x \ ,$$

was auf $c_0 = 0$ und $c_1 = -1$ führt. Zusammen mit der allgemeinen Lösung ($\tau_{1,2} = -\frac{1}{2} \pm \frac{i}{2}\sqrt{3}$) des homogenen Problem ergibt sich als allgemeinste Lösung

$$y = -x + ae^x + e^{-\frac{1}{2}x}\Big(b\cos\frac{\sqrt{3}}{2}x + c\sin\frac{\sqrt{3}}{2}x\Big)\ . \qquad \triangleleft$$

Satz 14.11 *Sei $P_n(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_1T + a_0 \in \mathbb{C}[T]$ ein Polynom und $g_m : \mathbb{R} \to \mathbb{C}$ eine Polynomfunktion vom Grad m. Die komplexe Zahl $\mu \in \mathbb{C}$ sei eine k-fache Nullstelle von P_n (Resonanzfall). Dann besitzt die Differentialgleichung $P_n(\frac{d}{dx})y = g_m(x)e^{\mu x}$ eine spezielle Lösung der Form $y(x) = h_{m+k}(x)e^{\mu x}$, wobei $h_{k+m}(x) = \sum_{j=k}^{m+k} c_j x^j$ eine Polynomfunktion vom Grad $m+k$ ist, in der die untersten Potenzen x^j mit $j < k$ nicht auftreten.*

Beweis. Nach Voraussetzung gibt es eine Darstellung $P_n(T) = Q_{n-k}(T)(T-\mu)^k$ mit $Q_{n-k}(\mu) \neq 0$. Nach Satz 14.9 gibt es eine Polynomfunktion $\tilde{h}_m(x)$, so daß $Q_{n-k}(\frac{d}{dx})(\tilde{h}_m(x)e^{\mu x}) = g_m(x)e^{\mu x}$. Es gibt dann eine Polynomfunktion $h_{m+k}(x) = \sum_{j=k}^{m+k} c_j x^j$, so daß $h^{(k)}_{m+k}(x) = \tilde{h}_m(x)$. Nach Lemma 14.5 gilt damit

$$\begin{aligned} P_n(\tfrac{d}{dx})\big(h_{m+k}(x)e^{\mu x}\big) &= Q_{n-k}(\tfrac{d}{dx})(\tfrac{d}{dx}-\mu)^k\big(h_{m+k}(x)e^{\mu x}\big) \\ &= Q_{n-k}(\tfrac{d}{dx})\big(h^{(k)}_{m+k}(x)e^{\mu x}\big) = Q_{n-k}(\tfrac{d}{dx})\big(\tilde{h}_m(x)e^{\mu x}\big) \\ &= g_m(x)e^{\mu x}\ . \end{aligned}$$ $\square$

Beispiel 14.12 Wir betrachten noch einmal die Differentialgleichung $y'' + \omega^2 y = b\cos\Omega x = \mathrm{Re}(be^{i\Omega x})$. Wir rechnen im Komplexen und nehmen am Ende den Realteil der Lösung. Wir haben $P_2(T) = (T - i\omega)(T + i\omega)$. Zunächst sei $\Omega^2 \neq \omega^2$. Wegen $P_2(i\Omega) = \omega^2 - \Omega^2$ ergibt sich für $\Omega \neq \omega$ die komplexe Lösung zu

$$y = c_1e^{i\omega x} + c_2e^{-i\omega x} + \frac{b}{\omega^2 - \Omega^2}e^{i\Omega x}\ , \qquad c_1, c_2 \in \mathbb{C}\ .$$

Der Realteil ist

$$\mathrm{Re}(y) = a_1\cos\omega x + a_2\sin\omega x + \frac{b}{\omega^2 - \Omega^2}\cos\Omega x\ , \quad a_1 = \mathrm{Re}(c_1{+}c_2)\ ,\ a_2 = \mathrm{Im}(c_2{-}c_1)\ .$$

Im Resonanzfall $\Omega = \pm\omega$ ist Ω eine einfache Nullstelle, so daß eine spezielle Lösung die Form $cxe^{i\omega x}$ hat. Wegen

$$(\tfrac{d^2}{dx^2} + \omega^2)(cxe^{i\omega x}) = 2ic\omega e^{i\omega x}$$

ergibt sich

$$y = c_1e^{i\omega x} + c_2e^{-i\omega x} + \frac{bx}{2i\omega}e^{i\omega x}\ , \qquad c_1, c_2 \in \mathbb{C}\ .$$

Der Realteil ist

$$\mathrm{Re}(y) = a_1\cos\omega x + a_2\sin\omega x + \frac{b}{2\omega}x\sin\omega x\ , \qquad a_1 = \mathrm{Re}(c_1{+}c_2)\ ,\ a_2 = \mathrm{Im}(c_2{-}c_1)\ .$$

Wir bestätigen damit die zuvor in Beispiel 13.8 erhaltenen Lösungen. $\triangleleft$

15 Differentialgleichungen 2. Ordnung

Viele Probleme der Physik führen auf Differentialgleichungen 2. Ordnung. Das typische Beispiel ist das Newtonsche Gesetz für die eindimensionale Bewegung der Mechanik $y'' = f(y, y')$, dabei ist f die durch die Masse dividierte Kraft. Die übliche Anfangsbedingung ist die Vorgabe von Ort und Geschwindigkeit zur Zeit x_0, d.h. $y(x_0) = y_0$ und $y'(x_0) = v_0$. Diese Differentialgleichung läßt sich in ein System linearer Differentialgleichungen 1. Ordnung überführen und dann durch Picard-Lindelöf integrieren. Ist die Kraft $f(y)$ unabhängig von der Geschwindigkeit, aber z.B. nichtlinear, dann bietet sich noch ein anderer Weg an: Da $\mathbb{R}$ einfach zusammenhängend ist, ist das Potential

$$U(y) := -\int_{y_0}^{y} ds\ f(s)$$

unabhängig vom Integrationsweg. Es gilt $\frac{dU}{dy}(y) = -f(y)$. Durch Multiplikation des Newtonschen Gesetzes mit der Geschwindigkeit y' ergibt sich dann

$$0 = y''y' + \frac{dU}{dy}y' = \Big(\frac{1}{2}y'^2 + U(y)\Big)' \ .$$

Die Lösung ist der Energieerhaltungssatz $\frac{1}{2}y'^2 + U(y) = E = \frac{1}{2}v_0^2 + U(y_0) = \text{const}$. Das ist nun eine Differentialgleichung mit getrennten Variablen mit der Lösung

$$x - x_0 = \pm \int_{y_0}^{y} \frac{ds}{\sqrt{2(E - U(s))}} =: T(y) \ .$$

Der Ort selbst bestimmt sich dann durch die Umkehrfunktion zu $y(x) = T^{-1}(x - x_0)$.

Beispiel 15.1 Die Differentialgleichung des mathematischen Pendels ist $y'' = -\omega^2 \sin y$. Dabei ist y der Auslenkungswinkel und $\omega^2 = \frac{g}{L}$, mit g der Schwerebeschleunigung und L der Fadenlänge. Für kleine Winkel ist $\sin y \approx y$, und es wird die einfachere lineare Schwingungsdifferentialgleichung (harmonischer Oszillator) erhalten. Wir stellen aber nicht die Bedingung kleiner Winkel. Das Potential ergibt sich zu

$$U(y) = -\int_{0}^{y} ds\ (-\omega^2 \sin s) = \omega^2(1 - \cos y) = 2\omega^2 \sin^2 \tfrac{y}{2} \ .$$

Dieses entspricht der Gesamtenergie im maximalen Auslenkungswinkel $a = y(0)$, so daß sich folgende Lösung ergibt:

$$x = \frac{1}{2\omega} \int_{y}^{a} \frac{ds}{\sqrt{\sin^2 \frac{a}{2} - \sin^2 \frac{s}{2}}}$$

Es interessiert die Zeit $\frac{T}{2}$ einer halben Schwingung zum anderen Umkehrpunkt $y = -a$. Wir substituieren $\sin\frac{s}{2} = \sin\frac{a}{2}\sin u$, also

$$\begin{aligned}\frac{1}{2}\cos\tfrac{s}{2}\,ds = \sin\tfrac{a}{2}\cos u\,du \quad &\Rightarrow \quad \frac{1}{2}\sqrt{1-\sin^2\tfrac{s}{2}}\,ds = \sin\tfrac{a}{2}\sqrt{1-\sin^2 u}\,du \\ &\Rightarrow \quad \frac{1}{2}\sqrt{1-\sin^2\tfrac{a}{2}\sin^2 u}\,ds = \sqrt{\sin^2\tfrac{a}{2}-\sin^2\tfrac{s}{2}}\,du\ .\end{aligned}$$

Das ergibt

$$\frac{T}{2} = \frac{1}{\omega}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{du}{\sqrt{1-\sin^2\frac{a}{2}\sin^2 u}} = \frac{2}{\omega}E(\sin\tfrac{a}{2})\ ,$$

wobei $E(k) := \displaystyle\int_0^{\frac{\pi}{2}} \frac{du}{\sqrt{1-k^2\sin^2 u}}$ ein elliptisches Integral ist. In Beispiel 22.9 aus dem 1. Semester hatten wir

$$T = \frac{2\pi}{\omega}\left(1+\sum_{n=1}^{\infty}\Big(\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots(2n)}\Big)^2\sin^{2n}\tfrac{a}{2}\right)$$

gezeigt. Insbesondere ist $\lim_{a\to 0} T = \frac{2\pi}{\omega}$. ◁

Wir betrachten nun einige Beispiele für Differentialgleichungen zweiter Ordnung, welche sich aus wichtigen partiellen Differentialgleichungen ergeben, wenn gewisse Symmetrien vorliegen. Ihre Lösungen sind spezielle Polynome.

Die *Legendresche Differentialgleichung* auf dem Intervall $I =]-1,1[$ ist gegeben durch

$$(1-x^2)y'' - 2xy + n(n+1)y = 0\ , \qquad n\in\mathbb{N}\ .$$

Satz 15.2 *Das Legendresche Polynom n-ter Ordnung*

$$P_n(x) := \frac{1}{2^n n!}\Big(\frac{d}{dx}\Big)^n (x^2-1)^n$$

ist Lösung der Legendreschen Differentialgleichung zum Parameter n.

Beweis. Der Vorfaktor ist hier irrelavant, er kommt von der Orthonormalitätseigenschaft der Legendreschen Polynome. Aus dem Pascalschen Dreieck folgt die Leibniz-Regel n-ter Ordnung

$$\Big(\frac{d}{dx}\Big)^n (fg)(x) = \sum_{j=0}^{n}\binom{n}{j} f^{(j)}(x)\,g^{(n-j)}(x)\ .$$

Wir betrachten die $(n+1)$-ste Ableitung von $z(x) := (x^2-1)\frac{d}{dx}(x^2-1)^n = 2nx(x^2-1)^n$:

$$
\begin{aligned}
&\frac{d^{n+1}}{dx^{n+1}}\Big((x^2-1)\frac{d}{dx}(x^2-1)^n\Big)\\
&= \binom{n+1}{0}(x^2-1)\frac{d^{n+2}}{dx^{n+2}}(x^2-1)^n + \binom{n+1}{1}2x\frac{d^{n+1}}{dx^{n+1}}(x^2-1)^n + \binom{n+1}{2}2\frac{d^n}{dx^n}(x^2-1)^n\\
&= (x^2-1)\frac{d^{n+2}}{dx^{n+2}}(x^2-1)^n + 2(n+1)x\frac{d^{n+1}}{dx^{n+1}}(x^2-1)^n + n(n+1)\frac{d^n}{dx^n}(x^2-1)^n\\
&\equiv \binom{n+1}{0}2nx\frac{d^{n+1}}{dx^{n+1}}(x^2-1)^n + \binom{n+1}{1}2n\frac{d^n}{dx^n}(x^2-1)^n\\
&= 2nx\frac{d^{n+1}}{dx^{n+1}}(x^2-1)^n + 2n(n+1)\frac{d^n}{dx^n}(x^2-1)^n\ .
\end{aligned}
$$

Zusammenfassung der Terme liefert

$$0 = (1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x)\ . \qquad \square$$

Die Legendreschen Polynome treten bei der Lösung der sehr wichtigen partiellen Differentialgleichung

$$\Delta\psi + U(r)\psi = 0$$

in Kugelkoordinaten (r,θ,ϕ) auf (Δ ist der Laplace-Operator). Beispiele sind die (zeitunabhängige) Schrödinger-Gleichung mit kugelsymmetrischen Potential (z.B. Wasserstoffatom). Der Ansatz $\psi(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi)$ führt mit der Normierungsbedingung für die Wahrscheinlichkeit $\int_{\mathbb{R}^3} dx\,|\psi(r,\theta,\phi)|^2 = 1$ auf Lösungen, die im allgemeinen durch ganze Zahlen parametrisiert werden (Quantisierung der Energieniveaus). Man findet (bis auf Vorfaktoren) $\Theta_{l0}(\theta) = P_l(\cos\theta)$.

Die Legendresche Differentialgleichung ist durch die Legendreschen Polynome noch nicht vollständig gelöst, da gemäß der allgemeinen Theorie linearer Differentialgleichungen n-ter Ordnung eine zweite linear unabhängige Lösung benötigt wird. Man kann zeigen, daß diese zweite Lösung durch das folgende Reduktionsverfahren erhalten werden kann:

Satz 15.3 *Sei $I \subset \mathbb{R}$ ein Intervall und $a, b : I \to \mathbb{K}$ stetige Funktionen. Es sei $y_{(1)} : I \to \mathbb{K}$ eine Lösung der Differentialgleichung $y'' + a(x)y' + b(x)y = 0$ und in einem Intervall $J \subset I$ gelte $y_{(1)}(x) \neq 0$ für alle $x \in J$. Dann erhält man auf J eine zweite linear unabhängige Lösung $y_{(2)} : J \to \mathbb{K}$ der Differentialgleichung durch den Ansatz $y_{(2)}(x) = u(x)y_{(1)}(x)$, wobei u dann eine nichtkonstante Lösung der Differentialgleichung*

$$u'' + \Big(2\frac{y_{(1)}'(x)}{y_{(1)}(x)} + a(x)\Big)u'(x) = 0$$

ist, die im ersten Schritt zu

$$u'(x) = \frac{1}{y_{(1)}^2(x)} \exp\Big(- \int_{x_0}^{x} dt\; a(t) \Big)$$

integriert werden kann.

Sehr ähnlich werden folgende Differentialgleichungen behandelt:

Beispiel 15.4 Die *Hermitesche Differentialgleichung* zum Parameter $n \in \mathbb{N}$ ist

$$y'' - 2xy' + 2ny = 0 \;, \qquad x \in \mathbb{R} \;,\; n \in \mathbb{N} \;.$$

Eine Lösung ist das *Hermitesche Polynom* n-ter Ordnung

$$H_n(x) := (-1)^n e^{x^2} \Big(\frac{d}{dx} \Big)^n e^{-x^2} \;.$$

Es tritt auf als Lösung der eindimensionalen Schrödinger-Gleichung für das Potential $U(x) = \frac{1}{2}\omega^2 x^2$ eines harmonischen Oszillators.

Beispiel 15.5 Die *Laguerresche Differentialgleichung* zum Parameter $n \in \mathbb{N}$ ist

$$xy'' + (1-x)y' + ny = 0 \;, \qquad x \in \mathbb{R}_+^* \;,\; n \in \mathbb{N} \;.$$

Eine Lösung ist das *Laguerresche Polynom* n-ter Ordnung

$$L_n(x) := \frac{1}{n!} e^x \Big(\frac{d}{dx} \Big)^n \big(x^n e^{-x} \big) \;.$$

Es tritt auf als Lösung der zweidimensionalen Schrödinger-Gleichung in Radialkoordinaten für das Potential $U(x) = \frac{1}{2}\omega^2 r^2$ eines harmonischen Oszillators.

Beispiel 15.6 Die *hypergeometrische Differentialgleichung* zu den 3 Parametern $\alpha, \beta, \gamma \in \mathbb{R}$ ist

$$x(1-x)y'' + (\gamma - (\alpha + \beta + 1)x)y' - \alpha\beta y = 0 \;, \qquad x \in \mathbb{C} \setminus \{0, 1, \infty\} \;.$$

Eine Lösung ist gegeben durch die *hypergeometrische Funktion*

$$F\Big(\begin{matrix} \alpha & \beta \\ & \gamma \end{matrix} \Big| x \Big) := 1 + \frac{\alpha\beta}{\gamma} \cdot \frac{x^1}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)} \cdot \frac{x^2}{2!} + \dots \;.$$

Ist α oder β eine negative ganze Zahl $-n$ (und γ geeignet), dann bricht die Reihe ab, und $F\Big(\begin{matrix} \alpha & \beta \\ & \gamma \end{matrix} \Big| x \Big)$ wird ein Polynom n-ter Ordnung in x.

Durch die sehr allgemeine 3-parametrige Form lassen sich viele andere spezielle Funktionen durch hypergeometrische ausdrücken. Viele Integrale berechnen sich zu hypergeometrischen Funktionen, wichtig ist dabei die Identität

$$F\Big(\begin{matrix} \alpha & \beta \\ & \gamma \end{matrix} \Big| x \Big) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \int_0^1 dt\; t^{\beta-1}(1-t)^{\gamma-\beta-1}(1-tx)^{-\alpha} \;.$$

Beispiel 15.7 Die *Besselsche Differentialgleichung* zum Parameter $p \in \mathbb{R}$ ist

$$y'' + \frac{1}{x}y' + \Big(1 - \frac{p}{x^2}\Big)y = 0 \,, \qquad p \in \mathbb{R} \,.$$

Dabei ist zunächst $x \in \mathbb{R}_+^*$, jedoch kann man die Gleichung auf $x \in \mathbb{C} \setminus \{0\}$ ausdehnen. Die Besselsche Differentialgleichung tritt auf bei zweidimensionalen Schwingungen und Wellen in Radialkoordinaten. Die Gleichung $\Delta u = -u$ in 2 Dimensionen liefert mit dem Ansatz $u(r,\phi) = f(r)e^{ip\phi}$ die Besselsche Differentialgleichung für $y(x) \mapsto f(r)$. Die Gleichung $\Delta u = -u$ entsteht z.B. aus der Wellengleichung $\big(\Delta - \frac{\partial^2}{\partial t^2}\big)\psi$ mit dem Ansatz $\psi(r,\phi,t) = u(r,\psi)e^{it}$.

Die Lösungen der Besselschen Differentialgleichung heißen *Zylinderfunktionen*. Die einfachste ist die *Besselfunktion* (zum Parameter $p \in \mathbb{R}$)

$$J_p(x) = \frac{x^p}{2^p} \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2^{2k} k! \Gamma(p+k+1)} \,.$$

Eine zweite linear unabhängige Lösung ist die Neumannsche Funktion $N_p(x)$. Für $x \in \mathbb{C}$ sind die komplexen Linearkombinationen $H_p(x) = J_p(x) \pm iN_p(x)$ (Hankel-Funktionen) nützlich.

Es gibt nützliche Integraldarstellungen der Besselfunktion, z.B.

$$J_n(x) = \frac{1}{\pi} \int_0^{\pi} d\theta \;\, \cos(n\theta - x \sin\theta) \,, \qquad n \in \mathbb{N} \,.$$

Daraus folgt $|J_n(x)| \leq 1$ für alle $x \in \mathbb{R}$, was aus der Reihenformel nicht offensichtlich ist. Bestimmte Integrale mit Zylinderfunktionen sind elementar berechenbar (zumindest tabelliert). Ihre Bedeutung ist vergleichbar mit Sinus und Cosinus bzw. Exponentialfunktion. Die Nullstellen der Besselfunktion sind wichtig bei zweidimensionalen *Randwertproblemen*, z.B. bei Schwingungen einer am Rand eingespannten Membran.

Teil IV
Grundlagen der Funktionentheorie

16 Holomorphe Funktionen

Wir behandeln nun die komplexe Differenzierbarkeit von komplexwertigen Funktionen $f : U \to \mathbb{C}$ auf offenen Teilmengen $U \subset \mathbb{C}$.

Definition 16.1 Es sei $U \subset \mathbb{C}$ offen. Eine Funktion $f : U \to \mathbb{C}$ heißt *komplex differenzierbar* im Punkt $z \in U$, falls der Grenzwert

$$f'(z) = \lim_{w \to z, w \neq z} \frac{f(w) - f(z)}{w - z}$$

existiert. Die Funktion $f : U \to \mathbb{C}$ heißt *holomorph* im Punkt $z \in U$, wenn f in einer offenen Umgebung $V \subset U$ von z komplex differenzierbar ist, und *holomorph auf* U, falls f in jedem Punkt von U holomorph ist.

Wie üblich wird die Konvergenz bezüglich des Abstands $|\ |$ auf $\mathbb{C}$ definiert. Wie im Reellen folgt aus der komplexen Differenzierbarkeit die Stetigkeit, außerdem die lineare Approximierbarkeit nach Satz 16.4 aus dem 1. Semester: Ist $f : U \to \mathbb{C}$ differenzierbar in $z_0 \in U$, dann gibt ein $\phi : U \to \mathbb{C}$ mit

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \phi(z) \qquad \text{mit } \lim_{z \to z_0} \frac{\phi(z)}{z - z_0} = 0 \ .$$

Für komplex-differenzierbare Funktionen f, g gelten die üblichen Rechenregeln

$$(f + g)'(z) = f'(z) + g'(z) \ , \qquad (f \cdot g)'(z) = f'(z)g(z) + f(z)g'(z) \ ,$$
$$(f \circ g)'(z) = f'(g(z)) \cdot g'(z) \ .$$

Analog zu Satz 16.12 aus dem 1. Semster gilt:

Satz 16.2 *Jede Potenzreihe $f(z) = \sum_{k=0}^{\infty} a_k z^k$ mit $a_k \in \mathbb{C}$ ist im Inneren ihres Konvergenzkreises $K_R(0)$ komplex differenzierbar und damit (für $R > 0$) holomorph in $K_R(0)$. Ihre Ableitung ist $f'(z) = \sum_{k=1}^{\infty} k a_k z^{k-1}$.* □

Wir können $f : U \to \mathbb{C}$ auch auffassen als $f = u + \mathrm{i}v : U \to \mathbb{R}^2$ mit $U \subset \mathbb{R}^2$ und komplexe und reelle Differenzierbarkeit vergleichen. Wir werden oft $(x, y) \in \mathbb{R}^2$ mit $x + \mathrm{i}y \in \mathbb{C}$ identifizieren.

Satz 16.3 *Es sei $U \subset \mathbb{C} \simeq \mathbb{R}^2$ offen.*

i) *Eine Funktion $f : U \to \mathbb{C}$ mit $f = u + \mathrm{i}v$ sei in $z = x + \mathrm{i}y \in U$ komplex differenzierbar. Dann sind $u, v : U \to \mathbb{R}$ partiell nach x, y differenzierbar, und es gelten die Cauchy-Riemannschen Differentialgleichungen*

$$\partial_x u = \partial_y v \ , \qquad \partial_y u = -\partial_x v \ .$$

ii) *Die Funktionen $u, v : U \to \mathbb{R}$ seien stetig partiell differenzierbar auf U, und es gelten die Cauchy-Riemannschen Differentialgleichungen. Dann ist $f = u + \mathrm{i}v$ holomorph auf U.*

Beweis. i) Ist f komplex differenzierbar in z, dann gilt

$$\partial_x f(x+\mathrm{i}y) = \lim_{h\to 0\ ,\ h\in\mathbb{R}} \frac{1}{h}\Big(f(x+h+\mathrm{i}y) - f(x+\mathrm{i}y)\Big) = f'(z)\ ,$$
$$\partial_y f(x+\mathrm{i}y) = \lim_{h\to 0\ ,\ h\in\mathbb{R}} \frac{\mathrm{i}}{\mathrm{i}h}\Big(f(x+\mathrm{i}(y+h)) - f(x+\mathrm{i}y)\Big) = \mathrm{i}f'(z)\ ,$$

also $\partial_x f = -\mathrm{i}\partial_y f$ und nach Zerlegung in Real- und Imaginärteil

$$\partial_x u + \mathrm{i}\partial_y v = -\mathrm{i}\partial_y u + \partial_y v\ .$$

ii) Nach Voraussetzung sowie Satz 5.4 ist $f = u + \mathrm{i}v$ total differenzierbar, und das Differential ist gegeben durch die Jacobi-Matrix der partiellen Ableitungen:

$$f(x+h_1+\mathrm{i}(y+h_2)) = f(x+\mathrm{i}y) + h_1(\partial_x f)(x+\mathrm{i}y) + h_2(\partial_y f)(x+\mathrm{i}y) + \phi(h_1+\mathrm{i}h_2)$$

mit $\lim_{h_1,h_2\to 0} \frac{|\phi(h_1+\mathrm{i}h_2)|}{\sqrt{h_1^2+h_2^2}} = 0$. Unter Verwendung der Cauchy-Riemannschen Differentialgleichungen gilt

$$\begin{aligned}
&\frac{f(x+h_1+\mathrm{i}(y+h_2)) - f(x+\mathrm{i}y)}{h_1+\mathrm{i}h_2} \\
&= \frac{1}{h_1+\mathrm{i}h_2}\Big(h_1(\partial_x u + \mathrm{i}\partial_x v)(x+\mathrm{i}y) + h_2(\partial_y u + \mathrm{i}\partial_y v)(x+\mathrm{i}y)\Big) + \frac{\phi(h_1+\mathrm{i}h_2)}{h_1+\mathrm{i}h_2} \\
&= \partial_x u + \mathrm{i}\partial_x v + \frac{\phi(h_1+\mathrm{i}h_2)}{h_1+\mathrm{i}h_2}\ .
\end{aligned}$$

Wegen $\lim_{h_1,h_2\to 0} \frac{|\phi(h_1+\mathrm{i}h_2)|}{\sqrt{h_1^2+h_2^2}} = 0$ existiert der Limes

$$f'(x+\mathrm{i}y) = \lim_{h_1+\mathrm{i}h_2\to 0} \frac{f(x+h_1+\mathrm{i}(y+h_2)) - f(x+\mathrm{i}y)}{h_1+\mathrm{i}h_2} = \partial_x u + \mathrm{i}\partial_x v\ .$$

Die Rechnung gilt für beliebige $x+\mathrm{i}y \in U$, also ist f holomorph auf U. □

Beispiel 16.4 Es sei $f(z) = z\bar{z} = x^2 + y^2$. Dann ist f komplex differenzierbar in 0, aber nicht in $z \neq 0$ und damit nirgends holomorph. Zwar gilt $f'(0) = \lim_{z\to 0} \frac{z\bar{z}}{z} = \lim_{z\to 0} \bar{z} = 0$, aber mit $u = x^2 + y^2$ und $v = 0$ ergibt sich

$$\partial_x u = 2x\ ,\quad \partial_y u = 2y\ ,\quad \partial_x v = \partial_y v = 0\ .$$

Die Cauchy-Riemannschen Differentialgleichungen gelten also nur in $x = y = 0$, somit ist f in keiner Umgebung von 0 komplex differenzierbar. ◁

Definition 16.5 Es sei $U \subset \mathbb{C}$ offen und $\gamma : [a,b] \to U$, mit $\gamma(t) = x(t) + \mathrm{i}y(t)$, eine differenzierbare Kurve. Sei $f : U \to \mathbb{C}$ stetig. Dann wird das *Kurvenintegral* von f entlang γ erklärt als

$$\int_\gamma dz\ f(z) := \int_a^b dt\ f(\gamma(t))\dot\gamma(t) = \int_a^b dt\ \big(f(\gamma(t))\dot x(t) + \mathrm{i}f(\gamma(t))\dot y(t)\big)\ .$$

Ist $\gamma : [a,b] \to U$ stückweise differenzierbar, d.h. γ ist stetig und es existieren $a = t_0 < t_1 < \cdots < t_n = b$, so daß die Einschränkung von γ auf $]t_{i-1}, t_i[$ differenzierbar ist, dann setzt man $\displaystyle\int_\gamma dz\ f(z) := \sum_{i=1}^n \int_{\gamma|_{[t_{i-1},t_i]}} dz\ f(z)$.

Satz 16.6 *Sei $\gamma : [a,b] \to U$ eine differenzierbare Kurve und $f : U \to \mathbb{C}$ stetig. Dann gilt*

$$\Big|\int_\gamma dz\ f(z)\Big| \le \sup_{t\in[a,b]} |f(\gamma(t))|\, L(\gamma)\ ,$$

wobei $L(\gamma) := \displaystyle\int_a^b dt\ \sqrt{(\dot x(t))^2 + (\dot y(t))^2}$ *die* Kurvenlänge *von γ ist.*

Beweis. Das folgt aus $\Big|\displaystyle\int_a^b dt\ g(t)\Big| \le \int_a^b dt\ |g(t)|$ für $g(t) = f(\gamma(t))\big(\dot x(t) + \mathrm{i}\dot y(t)\big)$ mit $|g(t)| = |f(\gamma(t))|\sqrt{(\dot x(t))^2 + (\dot y(t))^2}$. □

Wäre f sogar stetig differenzierbar, dann sind die Cauchy-Riemannschen Differentialgleichungen nach Satz 11.9 notwendig für die Existenz von Stammfunktionen zu $u(x,y) + \mathrm{i}v(x,y)\mathrm{i}y'(x) = 0$ und $\mathrm{i}v(x,y) + u(x,y)\mathrm{i}y'(x) = 0$. Wir hatten im Anschluß an Satz 11.10 behauptet, daß für einfach zusammenhängende Gebiete U die Stammfunktionen mittels einer beliebigen Kurve $\gamma : [a,b] \to U$ mit $\gamma(a) = x_0 + \mathrm{i}y_0$ und $\gamma(b) = x + \mathrm{i}y$ gegeben sind durch

$$F(x,y) = \int_a^b dt\Big(u(x(t),y(t))\dot x(t) - v(x(t),y(t))\dot y(t)\Big)\ ,$$
$$G(x,y) = \int_a^b dt\Big(v(x(t),y(t))\dot x(t) + u(x(t),y(t))\dot y(t)\Big)\ ,$$

also

$$(F + \mathrm{i}G)(x + \mathrm{i}y) = \int_a^b dt\Big(f(\gamma(t))\dot x(t) + \mathrm{i}f(\gamma(t))\dot y(t)\Big) = \int_\gamma dz\ f(z)\ .$$

Insbesondere ergibt sich für eine auf U stetig differenzierbare Funktion und eine geschlossene Kurve $\gamma : [a,b] \to U$ mit $\gamma(a) = \gamma(b)$ durch Unterteilung in $c \in]a,b[$ die Identität $\displaystyle\oint_\gamma dz\ f(z) = 0$. Wir zeigen, daß das sogar ohne die Voraussetzung der stetigen Differenzierbarkeit von f und mit Abschwächung an U gilt. Zunächst sei $\gamma = \partial\Delta$ der Rand eines Dreiecks Δ.

Satz 16.7 (Lemma von Goursat) *Es sei $U \subset \mathbb{C}$ offen, $f : U \to \mathbb{C}$ holomorph und Δ ein offenes Dreieck mit $\bar{\Delta} = \Delta \cup \partial\Delta \subset U$. (Sind $a, b, c \in U$ die Eckpunkte von Δ, dann ist $\partial\Delta = \vec{ab} \cup \vec{bc} \cup \vec{ca}$ mit positivem Umlaufsinn, d.h. entgegen dem Uhrzeigersinn.) Dann gilt*

$$\int_{\partial\Delta} dz\ f(z) = 0\ .$$

Beweis. Durch Verbinden der Seitenmittelpunkte entstehen aus Δ vier kongruente Dreiecke Δ_a, Δ_b, Δ_c, Δ_m. Werden diese Dreiecke positiv umlaufen, dann gilt

$$\int_{\partial\Delta} dz\ f(z) = \int_{\partial\Delta_a} dz\ f(z) + \int_{\partial\Delta_b} dz\ f(z) + \int_{\partial\Delta_c} dz\ f(z) + \int_{\partial\Delta_m} dz\ f(z)\ ,$$

da in der Summe die Kanten von $\partial\Delta_m$ zweimal in entgegengesetzte Richtung durchlaufen werden. Ist Δ_1 jenes Teildreieck, für das das Kurvenintegral den betragsmäßig größten Wert hat, dann gilt

$$\Big|\int_{\partial\Delta} dz\ f(z)\Big| \le 4\Big|\int_{\partial\Delta_1} dz\ f(z)\Big|\ .$$

Das Dreieck Δ_1 werde erneut in 4 Teildreiecke zerlegt, Δ_2 sei jenes mit betragsmäßig größtem Kurvenintegral. Durch Wiederholung des Verfahrens entsteht eine Folge $\Delta = \Delta_0 \supset \Delta_1 \supset \cdots \supset \Delta_n$ von Dreiecken mit

$$\Big|\int_{\partial\Delta} dz\ f(z)\Big| \le 4^n\Big|\int_{\partial\Delta_n} dz\ f(z)\Big|\ .$$

Wegen der Vollständigkeit von $\mathbb{C}$ gibt es ein $z_0 \in \Delta_n$ für alle n. Nach Voraussetzung ist f in z_0 komplex differenzierbar, d.h. es gibt eine Funktion $r : U \to \mathbb{C}$ mit

$$f(z) = f(z_0) + (z - z_0)f'(z_0) + |z - z_0|r(z) \qquad \text{mit } \lim_{z\to z_0} r(z) = 0\ .$$

Für ein Kurvenintegral über die lineare Funktion $f(z_0) + (z - z_0)f'(z_0)$ gilt

$$\begin{aligned}
&\int_\gamma dz\ \big(f(z_0) + (z - z_0)f'(z_0)\big)\\
&= \int_a^b dt\ \Big(f(z_0) + (x(t) + \mathrm{i}y(t) - z_0)f'(z_0)\Big)\big(\dot{x}(t) + \mathrm{i}\dot{y}(t)\big)\\
&= \int_a^b dt\ \frac{d}{dt}\Big(\big(f(z_0) - z_0f'(z_0)\big)\big(x(t) + \mathrm{i}y(t)\big) + \frac{1}{2}f'(z_0)\big(x(t) + \mathrm{i}y(t)\big)^2\Big)\\
&= \Big(\big(f(z_0) - z_0f'(z_0)\big)\big(\gamma(b) - \gamma(a)\big) + \frac{1}{2}f'(z_0)\big(\gamma(b)^2 - \gamma(a)^2\big)\Big)\ .
\end{aligned}$$

Für eine geschlossene Kurve γ wie z.B. $\gamma = \partial\Delta_n$ gilt somit

$$\int_{\partial\Delta_n} dz\ \big(f(z_0) + (z - z_0)f'(z_0)\big) = 0\ .$$

Es sei $L := L(\Delta)$ der Umfang von Δ. Dann ist $L(\Delta_n) = 2^{-n}L$. Wegen $z_0 \in \Delta_n$ ist $|z - z_0| \leq 2^{-n}L$ für alle $z \in \partial\Delta_n$. Somit gilt nach Satz 16.6

$$\Big| \int_{\partial\Delta_n} dz\ f(z) \Big| = \Big| \int_{\partial\Delta_n} dz\ |z-z_0| r(z) \Big| \leq 2^{-n} L \cdot \sup_{z\in\partial\Delta_n} |z-z_0||r(z)| \leq 4^{-n} L^2 \sup_{z\in\overline{\Delta_n}} |r(z)| \ .$$

Somit gilt

$$\Big| \int_{\partial\Delta} dz\ f(z) \Big| \leq L^2 \sup_{z\in\overline{\Delta_n}} |r(z)| \ .$$

Für $n \to \infty$ geht $z \to z_0$ und damit $\lim_{n\to\infty} \sup_{z\in\overline{\Delta_n}} |r(z)| = 0$. Das ist die Behauptung. □

Wir zeigen nun, daß wir ein Kurvenintegral durch ein Integral über ein eingeschriebes Polygon approximieren können.

Satz 16.8 *Es sei $U \subset \mathbb{C}$ offen, $\gamma : [a,b] \to U$ eine (stückweise) differenzierbare Kurve und $f : U \to \mathbb{C}$ stetig. Dann gibt es zu jedem $\epsilon > 0$ eine Unterteilung $a = t_0 < t_1 < \cdots < t_n = b$, so daß für die zugehörigen Sekanten $\gamma_i(t) := \gamma(t_{i-1}) + \frac{t-t_{i-1}}{t_i - t_{i-1}}\big(\gamma(t_i) - \gamma(t_{i-1})\big)$, mit $t \in [t_{i-1}, t_i]$, von γ gilt $\gamma_i \subset U$ und*

$$\Big| \int_\gamma dz\ f(z) - \sum_{i=1}^n \int_{\gamma_i} dz\ f(z) \Big| < \epsilon \ .$$

Beweis. Es sei $L < \infty$ die Länge von γ und $\delta > 0$ vorgegeben. Wegen der Kompaktheit von γ und der stückweisen Differenzierbarkeit (lineare Approximierbarkeit) gibt es eine endliche Überdeckung von γ durch offene Kugeln $K_1, \ldots, K_n \subset \mathbb{C}$ vom Radius $\leq \delta$ und eine Unterteilung $a = t_0 < t_1 < \cdots < t_n = b$ derart, daß

$$\Gamma_i := \{ s\gamma(t) + (1-s)\gamma_i(t) \ :\ t \in [t_{i-1}, t_i],\ s \in [0,1] \} \subset K_i \cap U \ . \qquad (*)$$

Die Menge Γ_i der Punkte (*) ist als Bild von $[t_{i-1}, t_i] \times [0,1]$ unter einer stetigen Abbildung kompakt, so daß f auf $\Gamma_1 \cup \cdots \cup \Gamma_n$ gleichmäßig stetig ist. Zum gegebenen $\epsilon > 0$ gibt es also ein $\delta > 0$, so zu jeder (*) erfüllenden Überdeckung (K_i) und Unterteilung (t_i) für beliebige $z, z' \in \Gamma_i \subset K_i \cap U$ gilt $|f(z) - f(z')| < \frac{\epsilon}{2L}$. Ferner gibt es nach Satz 19.12 (Riemannsche Summen) aus dem 1. Semester zu $\epsilon > 0$ eine (zunächst andere) Unterteilung $a = t_0 < t_1 < \cdots < t_n = b$ mit $|t_i - t_{i-1}| < \delta'$, so daß für beliebige Wahl der Stützstellen $\xi_i \in [t_{i-1}, t_i]$ gilt

$$\Big| \int_a^b dt\ f(\gamma(t))\dot\gamma(t) - \sum_{i=1}^n f(\gamma(\xi_i))\dot\gamma(\xi_i)(t_i - t_{i-1}) \Big| < \frac{\epsilon}{2} \ .$$

Wegen der Lipschitz-Stetigkeit von γ gibt es ein $M > 0$ mit $|\gamma(t_{i-1}) - \gamma(t_i)| \leq M\delta'$. Nach möglicher Verkleinerung von δ oder δ' können wir diese Zerlegung als

identisch mit (*) annehmen. Wir wählen ξ_i als Zwischenpunkt im Mittelwertsatz der Differentialrechnung für γ:

$$\dot\gamma(\xi_i) = \frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} = \dot\gamma_i(t) \qquad \text{für beliebiges } t \in [t_i, t_{i-1}] \; .$$

Nach Konstruktion der Γ_i gilt $|f(\gamma_i(t)) - f(\gamma(\xi_i))| < \frac{\epsilon}{2L}$ für beliebige $t \in [t_{i-1}, t_i]$. Somit ist

$$\begin{aligned} &\Big| \sum_{i=1}^n f(\gamma(\xi_i))\dot\gamma(\xi_i)(t_i - t_{i-1}) - \sum_{i=1}^n \int_{\gamma_i} dz\; f(z) \Big| \\ &\leq \sum_{i=1}^n \Big| \int_{t_{i-1}}^{t_{i-1}} dt\; \Big(f(\gamma(\xi_i))\dot\gamma_i(t) - f(\gamma_i(t))\dot\gamma_i(t) \Big) \Big| \leq \sum_{i=1}^n \frac{\epsilon}{2L} L(\gamma_i) \; . \end{aligned}$$

Die letzte Ungleichung ist Satz 16.6. Nun ist die Länge $L(\gamma_i) = \displaystyle\int_{t_{i-1}}^{t_{i-1}} dt\; |\dot\gamma_i(t)|$ der Sekante γ_i kleiner gleich der Länge des Kurvensegments zwischen $\gamma(t_{i-1})$ und $\gamma(t_i)$ und somit $\sum_{i=1}^n L(\gamma_i) \leq L$. Damit ergibt sich die Behauptung aus der Dreiecksungleichung. □

Definition 16.9 Es sei $U \subset \mathbb{C}$ offen.

i) Zwei stetige Kurven $\gamma_0, \gamma_1 : [a,b] \to U$ mit $\gamma_0(a) = \gamma_1(a) = A$ und $\gamma_0(b) = \gamma_1(b) = B$ heißen *homotop* in U, wenn sie in U stetig ineinander deformiert werden können, d.h. wenn es eine stetige Abbildung $H : [a,b] \times [0,1] \mapsto H(t,s) \in U$ gibt mit $H(t,0) = \gamma_0(t)$ und $H(t,1) = \gamma_1(t)$ für alle $t \in [a,b]$ sowie $H(a,s) = A$ und $H(b,s) = B$ für alle $s \in [0,1]$.

ii) Eine geschlossene stetige Kurve $\gamma : [a,b] \to U$ mit $\gamma(a) = \gamma(b) = A$ heißt *kontrahierbar* in U (oder *nullhomotop*), wenn sie homotop zum Punkt $\gamma_0(t) = A$ für alle $t \in [a,b]$ ist.

iii) U heißt *einfach zusammenhängend*, wenn U zusammenhängend ist und jede geschlossene Kurve in U kontrahierbar ist.

Satz 16.10 (Cauchyscher Integralsatz) *Es sei $U \subset \mathbb{C}$ offen und $f : U \to \mathbb{C}$ eine holomorphe Funktion. Dann gilt:*

i) *Für jede in U kontrahierbare stückweise differenzierbare geschlossene Kurve $\gamma : [a,b] \to U$ gilt $\displaystyle\int_\gamma dz\; f(z) = 0$.*

ii) *Sind γ_0, γ_1 homotope stückweise differenzierbare Kurven in U mit gemeinsamem Anfangs- und Endpunkt, dann gilt*

$$\int_{\gamma_0} dz\; f(z) = \int_{\gamma_1} dz\; f(z) \; .$$

iii) *Ist U einfach zusammenhängend und $\gamma : [a,b] \to U$ eine beliebige stückweise differenzierbare Kurve von $\gamma(a) = z_0$ nach $\gamma(b) = z$, dann ist das Integral*

$$F(z) := \int_{z_0}^{z} dw\ f(w) = \int_{\gamma} dw\ f(w)$$

unabhängig von γ und eine holomorphe Stammfunktion zu f, d.h. es gilt $F'(z) = f(z)$.

Beweis. i) Nach Satz 16.8 gibt es zu beliebigem $\epsilon > 0$ eine Unterteilung $a = t_0 < t_1 < \cdots < t_n = b$ und zugehörige Sekanten $\gamma_i : [t_{i-1}, t_i] \to U$ von γ mit $\Big| \int_\gamma dz\ f(z) - \sum_{i=1}^{n} \int_{\gamma_i} dz\ f(z) \Big| < \epsilon$. Nach Konstruktion ist mit γ auch das geschlossene Polygon $P(\gamma) := \gamma_1 \cup \cdots \cup \gamma_n$ in U kontrahierbar. Es sei $Q(\gamma)$ das Innere von $P(\gamma)$, welches sich triangulieren läßt, d.h. es gibt endlich viele abgeschlossene Dreiecke $\Delta_1, \ldots, \Delta_m \subset U$, welche sich höchstens in Ecken oder Kanten schneiden, mit $\Delta_1 \cup \cdots \cup \Delta_m = Q(\gamma)$ und $\partial(\Delta_1 \cup \cdots \cup \Delta_m) = P(\gamma)$. Nach dem Lemma von Goursat ist $\int_{\partial\Delta_i} dz\ f(z) = 0$ für jedes Dreieck Δ_i. In der Summe heben sich die Beiträge aller inneren Kanten weg, so daß

$$0 = \sum_{i=1}^{m} \int_{\partial\Delta_i} dz\ f(z) = \int_{P(\gamma)} dz\ f(z)\ .$$

Somit ist $\Big| \int_\gamma dz\ f(z) \Big| < \epsilon$ für beliebiges $\epsilon > 0$, also $\int_\gamma dz\ f(z) = 0$.

ii) folgt aus i) für die kontrahierbare Kurve $\gamma = \gamma_0 \cup (-\gamma_1)$, wobei $-\gamma_1$ in umgekehrte Richtung wie γ_1 durchlaufen wird, was $\int_{-\gamma_1} dz\ f(z) = -\int_{\gamma_1} dz\ f(z)$ ergibt.

iii) Nach Aufteilung in Polygonzüge können wir annehmen, daß für beliebige $z+h$ aus einer Umgebung von z das Dreieck Δ mit Eckpunkten z_0,z, $z+h$ in U liegt. Sei $F(z) := \int_{\overline{z_0,z}} dw\ f(w)$ und analog für $F(z+h)$. Nach dem Lemma von Goursat ist

$$0 = \int_{\partial\Delta} dw\ f(w) = \int_{\overline{z_0,z}} dw\ f(w) + \int_{\overline{z,z+h}} dw\ f(w) - \int_{\overline{z_0,z+h}} dw\ f(w)\ .$$

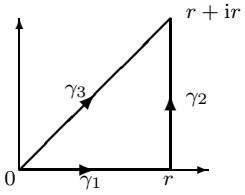
Damit ist

$$\begin{aligned} \Big| \frac{F(z+h) - F(z)}{h} - f(z) \Big| &= \Big| \frac{1}{h} \int_{\overline{z,z+h}} dw\ f(w) - \frac{1}{h} \int_{\overline{z,z+h}} dw\ f(z) \Big| \\ &\leq \sup_{w \in \overline{z,z+h}} |f(w) - f(z)|\ . \end{aligned}$$

Wegen der Stetigkeit von f ist $\lim_{h\to 0} \sup_{w\in\overline{z,z+h}} |f(w)-f(z)| = 0$. $\square$

Der Cauchysche Integralsatz ist die Grundlage der gesamten Funktionentheorie. Seine Konsequenzen sind grundlegend verschieden zur reellen Dfferentialrechnung.

Beispiel 16.11 (Fresnel-Integral) Gesucht ist $\int_0^\infty dt\ \ e^{-it^2}$. Dazu integrieren wir die holomorphe Funktion $f(z) = e^{-z^2}$ über die geschlossene Kurve $(\gamma_1, \gamma_2, -\gamma_3)$:



$$\int_{\gamma_3} dz f(z) = \int_{\gamma_1} dz f(z) + \int_{\gamma_2} dz f(z)\ .$$

Mit $\gamma_2(t) = r + it$ ist $|f(\gamma_2(t))| = e^{-r^2+t^2} \le e^{-r^2+rt}$ und damit

$$\Big|\int_{\gamma_2} dz\ f(z)\Big| \le \int_0^r dt\ |f(\gamma_2(t))\dot{\gamma_2}(t)| = e^{-r^2}\int_0^r dt\ e^{rt} = \frac{1-e^{-r^2}}{r}\ .$$

Für $r\to\infty$ konvergiert das Integral gegen 0, also gilt

$$\lim_{r\to\infty}\int_{\gamma_3} f(z)dz = \lim_{r\to\infty}\int_{\gamma_1} f(z)dz = \lim_{r\to\infty}\int_0^r dt\ e^{-t^2} = \frac{\sqrt{\pi}}{2}\ .$$

Andererseits ist $\gamma_3(t) = (1+\mathrm{i})t$ mit $t\in[0,r]$, also $(\gamma_3(t))^2 = 2it^2$ und

$$\lim_{r\to\infty}\int_0^r dt\ e^{-2\mathrm{i}t^2}(1+\mathrm{i}) = \frac{(1+\mathrm{i})}{\sqrt{2}}\int_0^\infty d\tau\ e^{-\mathrm{i}\tau^2} = \sqrt{\mathrm{i}}\int_0^\infty d\tau\ e^{-\mathrm{i}\tau^2}\ .$$

Somit gilt $\displaystyle\int_0^\infty dt\ e^{-\mathrm{i}t^2} = \frac{1}{2}\sqrt{\frac{\pi}{\mathrm{i}}}$ und nach Zerlegung in Real- und Imaginärteil mit $e^{-\mathrm{i}t^2} = \cos t^2 - \mathrm{i}\sin t^2$ schließlich

$$\int_0^\infty dt\ \cos(t^2) = \int_0^\infty dt\ \sin(t^2) = \sqrt{\frac{\pi}{8}} \quad\Rightarrow\quad \int_0^\infty dx\ \frac{\cos x}{\sqrt{x}} = \int_0^\infty dx\ \frac{\sin x}{\sqrt{x}} = \sqrt{\frac{\pi}{2}}\ .$$

$\triangleleft$

Beispiel 16.12 (Hauptzweig des komplexen Logarithmus) Es sei $\mathbb{C}^- := \mathbb{C}\setminus\mathbb{R}^-$ die geschlitzte komplexe Ebene (die reellen Zahlen ≤ 0 fehlen). Jede geschlossene Kurve $\gamma : [a,b]\to\mathbb{C}^-$ ist kontrahierbar. Die Funktion $f(z) = \frac{1}{z}$ ist wegen $f'(z) = -\frac{1}{z^2}$ holomorph auf $\mathbb{C}^-$ und kann deshalb wegunabhängig integriert werden. Es sei

$$L(z) = \int_1^z \frac{dz}{z} := \int_\gamma \frac{dz}{z} \qquad (\gamma \text{ ist beliebige Kurve zwischen 1 und } z \text{ in } \mathbb{C}^-)$$

Dann ist $L(z)$ holomorph auf $\mathbb{C}^-$ mit $L'(z) = \frac{1}{z}$ und $L(1) = 0$. Es gilt $(z\mathrm{e}^{-L(z)})' = \mathrm{e}^{-L(z)}(1 - zL'(z)) = 0$, also ist $z\mathrm{e}^{-L(z)} = \mathrm{const} = 1 \cdot \mathrm{e}^{-L(1)} = 1$, d.h.

$$\mathrm{e}^{L(z)} = z \qquad \text{für alle } z \in \mathbb{C}^- \ .$$

Man nennt $L(z)$ den Hauptzweig des komplexen Logarithmus. Mit der Kurve

$$\gamma = \gamma_1 \cup \gamma_2 \ , \quad \gamma_1(t) = t : [1, r] \to \mathbb{C} \ , \quad \gamma_2(t) = r\mathrm{e}^{\mathrm{i}t} : [0, \varphi] \to \mathbb{C}$$

von 1 über r nach $z = r\mathrm{e}^{\mathrm{i}\varphi}$ ergibt sich mit $\dot\gamma_1(t) = 1$ und $\dot\gamma_2(t) = \mathrm{i}r\mathrm{e}^{\mathrm{i}t}$

$$L(z) = \int_1^r dt \frac{1}{t} \cdot 1 + \int_0^\varphi dt \frac{1}{re^{\mathrm{i}\varphi}} \cdot \mathrm{i}re^{\mathrm{i}t} = r + \mathrm{i}\phi \ . \qquad \triangleleft$$

Ist $f : U \to \mathbb{C}$ eine nullstellenfreie Funktion auf einer einfach zusammenhängenden offenen Teilmenge $U \subset \mathbb{C}$, dann wird durch $e^F = f$ der *holomorphe Logarithmus* F von f definiert. Wie in Beispiel 16.12 ist

$$F(z) = \int_{z_0}^z dw \ \frac{f'(w)}{f(w)}$$

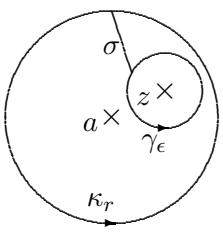
eine holomorphe Stammfunktion zu $\frac{f'}{f}$, und es gilt $fe^{-F} = \mathrm{const}$. Wegen $e^z = 1 \quad \Leftrightarrow \quad z \in 2\pi\mathrm{i}\mathbb{Z}$ unterscheiden sich zwei holomorphe Logarithmen nur um Vielfache von $2\pi\mathrm{i}$ und sind somit durch Angabe ihres Wertes in einem Punkt $z_0 \in U$ eindeutig bestimmt. Über den holomorphen Logarithmus können komplexe Potenzen nullstellenfreier Funktionen definiert werden als $f^\alpha = e^{\alpha F}$ für $\alpha \in \mathbb{C}$.

17 Die Cauchysche Integralformel

Satz 17.1 (Cauchysche Integralformel) *Es sei f holomorph in einer offenen Teilmenge $U \subset \mathbb{C}$, welche die abgeschlossene Kreisscheibe $\overline{K_r(a)}$ mit Mittelpunkt $a \in U$ und Radius r enthält. Der Umfang der Kreisscheibe ist dann die Kurve $\kappa_r(t) = a + re^{\mathrm{i}t}$ mit $t \in [0, 2\pi]$. Dann gilt für jeden Punkt $z \in K_r(a)$ im Inneren des Kreises*

$$f(z) = \frac{1}{2\pi\mathrm{i}} \int_{\kappa_r} d\zeta \ \frac{f(\zeta)}{\zeta - z} \ .$$

Beweis. Zu $z \in K_r(a)$ gibt es ein $\epsilon > 0$, so daß die abgeschlossene Kreisscheibe $\overline{K_\epsilon(z)}$ um z mit Radius ϵ im Inneren von $K_r(a)$ liegt. Sei $\gamma_\epsilon(\tau) = z + \epsilon e^{\mathrm{i}\tau}$ mit $\tau \in [0, 2\pi]$ der Umfang. Der entstehende (asymmetrische) Kreisring $KR_{r,\epsilon} := \overline{K_r(a)} \setminus K_\epsilon(z)$ werde aufgeschnitten entlang einer beliebigen (stückweise differenzierbaren) Kurve $\sigma \in KR_{r,\epsilon}$.



Dann ist $\frac{f(\zeta)}{\zeta - z}$ bezüglich ζ holomorph in dem so entstehenden einfach zusammenhängenden Gebiet Γ, das von den Kurven $\kappa_r, \sigma, -\gamma_\epsilon, -\sigma$ berandet wird. Dabei werden die beiden Kurven σ in verschiedene Richtungen durchlaufen, so daß sich die Randintegrale wegheben. Außerdem wird γ_ϵ in negative Richtung durchlaufen. Somit gilt nach Satz 16.10

$$\int_{\kappa_r} d\zeta\, \frac{f(\zeta)}{\zeta - z} = \int_{\gamma_\epsilon} d\zeta\, \frac{f(\zeta)}{\zeta - z} = \int_{\gamma_\epsilon} d\zeta\, \frac{f(\zeta) - f(z)}{\zeta - z} + f(z) \int_{\gamma_\epsilon} d\zeta\, \frac{1}{\zeta - z}\,.$$

Insbesondere ist die rechte Seite unabhängig von ϵ, also können wir den Limes $\epsilon \to 0$ betrachten. Da f in einer Umgebung von z komplex differenzierbar ist, ist $\frac{f(\zeta)-f(z)}{\zeta - z}$ auf $\overline{K_\epsilon(z)}$ beschränkt. Da der Umfang $L(\gamma_\epsilon)$ mit ϵ gegen 0 geht, ist $\lim\limits_{\epsilon\to 0} \int\limits_{\gamma_\epsilon} d\zeta\, \dfrac{f(\zeta) - f(z)}{\zeta - z} = 0$. Schließlich gilt auf dem inneren Kreis $\frac{1}{\zeta(\gamma_\epsilon(\tau)) - z} = \frac{1}{\epsilon e^{\mathrm{i}\tau}}$ und $\dot{\gamma_\epsilon}(\tau) = \mathrm{i}\epsilon\, e^{\mathrm{i}t}$, damit

$$\int_{\gamma_\epsilon} d\zeta\, \frac{1}{\zeta - z} = \int_0^{2\pi} dt \frac{\mathrm{i}\epsilon\, e^{\mathrm{i}t}}{\epsilon e^{\mathrm{i}t}} = 2\pi\mathrm{i}\,,$$

was die Cauchysche Integralformel beweist. □

Entscheidend für die gesamte Funktionentheorie ist die Tatsache, daß man den Wert $f(z)$ durch ein Kurvenintegral berechnen kann, wobei die Kurve *außerhalb* von problematischen Punkten der Funktion gewählt werden kann. Außerdem geht der Punkt z im Kurvenintegral gar nicht in die Funktion f ein, sondern tritt nur im Faktor $\frac{1}{\zeta - z}$ auf. Dadurch lassen sich bemerkenswerte Aussagen gewinnen.

Satz 17.2 (Potenzreihenentwicklung) *Eine holomorphe Funktion f auf einer offenen Teilmenge $U \subset \mathbb{C}$ kann in jeder offenen Kreisscheibe $K_\rho(a) \subset U$ in eine Potenzreihe $f(z) = \sum_{n=0}^{\infty} a_n (z-a)^n$ entwickelt werden. Der Konvergenzradius ist mindestens so groß wie der Abstand des Mittelpunktes a zum Rand von U. Die Entwicklungskoeffizienten sind gegeben durch die Integrale*

$$a_n = \frac{1}{2\pi\mathrm{i}} \int_{\partial K_r(a)} d\zeta\, \frac{f(\zeta)}{(\zeta - a)^{n+1}}$$

für einen beliebigen Radius $0 < r < \rho$. Ist $|f(\zeta)| < M$ für alle $\zeta \in \partial K_r(a)$, dann können die Koeffizienten abgeschätzt werden durch $|a_n| \le \frac{M}{r^n}$.

Beweis. Für $z \in K_r(a)$ mit $r < \rho$ und $\zeta \in \partial K_r(a)$ gibt es eine reelle Zahl $0 < q < 1$, so daß $|\frac{z-a}{\zeta - a}| \le 1 - q$. Dann gilt

$$\frac{f(\zeta)}{\zeta - z} = \frac{f(\zeta)}{\zeta - a} \cdot \frac{1}{1 - \frac{z-a}{\zeta - a}} = \frac{f(\zeta)}{\zeta - a} \sum_{n=0}^{\infty} \Big(\frac{z-a}{\zeta - a}\Big)^n\,,$$

wobei die Reihe, die durch $\sum_{n=0}^{\infty}(1-q)^n = \frac{1}{q}$ majorisiert wird, gleichmäßig konvergent ist. Damit gilt

$$f(z) = \int_{\partial K_r(a)} d\zeta\, \frac{f(\zeta)}{\zeta - z} = \sum_{n=0}^{\infty} \Big(\frac{1}{2\pi \mathrm{i}} \int_{\partial K_r(a)} d\zeta\, \frac{f(\zeta)}{(\zeta-a)^{n+1}}\Big)(z-a)^n \ .$$

Die Abschätzung ergibt sich aus $|\frac{f(\zeta)}{(\zeta-a)^{n+1}}| \le \frac{M}{r^{n+1}}$ für alle $\zeta \in \partial K_r(a)$ und der Länge $2\pi r$ des Randes. □

Als wichtige Konsequenz ergibt sich:

Satz 17.3 *Jede holomorphe Funktion $f : U \to \mathbb{C}$ ist beliebig oft komplex differenzierbar, alle Ableitungen $f^{(k)}$ sind holomorph und gegeben durch*

$$f^{(k)}(z) = \frac{k!}{2\pi \mathrm{i}} \int_{\partial K_r(a)} d\zeta\, \frac{f(\zeta)}{(\zeta - z)^{k+1}} \ , \qquad z \in K_r(a) \ . \qquad \square$$

Diese Aussage ist grundlegend verschieden von der reellen Differentialrechnung: Für eine reell differenzierbare Funktion muß die Ableitung nicht einmal stetig sein (z.B. $f(x) = x^2 \sin \frac{1}{x}$). Selbst wenn eine Funktion beliebig oft reell differenzierbar ist, muß sie nicht in eine Potenzreihe entwickelbar sein (z.B. $f(x) = e^{-\frac{1}{x^2}}$ in $x = 0$).

Satz 17.4 *Ist $f(z) = \sum_{k=0}^{\infty} a_k (z-c)^k$ eine in $K_R(c)$ konvergente Potenzreihe, dann kann f um jeden Punkt $w \in K_R(c)$ in eine Potenzreihe $f(z) = \sum_{n=0}^{\infty} b_n (z-w)^n$ entwickelt werden. Der Konvergenzradius dieser Reihe ist mindestens $R - |w-c|$, und es gilt $b_n = \sum_{k=n}^{\infty} \binom{k}{n} a_k (w-c)^{k-n}$.*

Beweis. f ist holomorph in $w \in K_R(c)$ und deshalb in eine Potenzreihe entwickelbar. Der Konvergenzradius folgt aus Satz 17.2, sowie

$$b_n = \frac{1}{n!} f^{(n)}(w) = \frac{1}{n!} \sum_{k=n}^{\infty} a_k k(k-1)\cdots(k-n+1)(w-c)^{k-n} \ . \qquad \square$$

Oft ist der Konvergenzradius der umentwickelten Reihe größer als $R - |w-c|$, so daß die umentwicklete Reihe die Funktion f über $K_R(c)$ hinaus fortsetzt. Man spricht dann von einer *analytischen Fortsetzung* von f.

Beispiel 17.5 Es sei $f(z) = \sum_{k=0}^{\infty} z^k$ mit Konvergenzradius $R = 1$. Wir entwickeln f um $w = -\frac{1}{2}$. Innerhalb des Konvergenzkreises ist

$$f(z) = \frac{1}{1-z} = \frac{1}{\frac{3}{2} - (z + \frac{1}{2})} = \frac{2}{3} \sum_{n=0}^{\infty} \Big(\frac{2}{3}\Big(z + \frac{1}{2}\Big)\Big)^n \ .$$

Die Reihe konvergiert für $|z + \frac{1}{2}| < \frac{3}{2}$ und wird damit auf einen größeren Kreis analytisch fortgesetzt. □

Wir zeigen, daß eine mögliche analytische Fortsetzung eindeutig ist:

Satz 17.6 *Es sei $U \subset \mathbb{C}$ offen und zusammehängend. Für zwei holomorphe Funktionen $f, g : U \to \mathbb{C}$ sind äquivalent:*

i) $f = g$

ii) *Die Identitätsmenge $\{w \in U \ : \ f(w) = g(w)\}$ hat einen Häufungspunkt in U.*

iii) *Es gibt ein $c \in U$ und ein $n \in \mathbb{N}$, so daß $f^{(k)}(c) = g^{(k)}(c)$ für alle $k \geq n$.*

Beweis. i)⇒ii) ist klar.

ii)⇒iii) Für $h = f - g$ hat die Nullstellenmenge von h einen Häufungspunkt $c \in U$. Angenommen, es gäbe ein $k \in \mathbb{N}$ mit $h^{(k)}(c) \neq 0$, und sei n das Minimum dieser k. Wegen der Potenzreihenentwicklung ist $h(z) = (z-c)^n h_n(z)$ mit $h_n(z) = \sum_{k=n}^{\infty} a_k (z-c)^k$ und $h_n(c) \neq 0$. Wegen der Stetigkeit von h_n gilt dann auch $h_n(z) \neq 0$ für alle z aus einer ϵ-Umgebung von c, im Widerspruch zur Voraussetzung, daß c Häufungspunkt der Nullstellenmenge ist.

iii)⇒i) Es sei $h = f - g$ und $S_k := \{w \in U \ : \ h^{(k)}(w) = 0\}$. Nach Satz 1.10.iii) ist S_k abgeschlossen in U. Da der Durchschnitt beliebig vieler abgeschlossener Teilmengen wieder abgeschlossen ist, ist $S := \bigcap_{k=0}^{\infty} S_k$ abgeschlossen in U. Andererseits ist S auch offen in U, denn für $z_1 \in S$ ist die Potenzreihenentwicklung von h in einer beliebigen offenen Kreisscheibe $K \subset U$ mit Mittelpunkt z_1 die Nullreihe. Damit verschwinden sämtliche Ableitungen $h^{(k)}(z)$ für alle $z \in K$, also ist $K \subset S$. Da U zusammenhängend ist, folgt $S = U$. □

Bemerkenswert ist, daß $f = g$ in ganz U aus zwei entgegengesetzten Bedingungen folgt: Aus der Gleichheit aller Ableitungen an *nur einem* Punkt sowie aus der Gleichheit an genügend vielen Punkten in U. Das ist grundlegend verschieden vom reellen Fall. Für die Funktionen $f(X) = 0$ und $g(x) = e^{-\frac{1}{x^2}}$ sind in $x = 0$ alle Ableitungen gleich, aber offenbar ist $f \neq g$. Als wichtige Konsequenz ergibt sich, daß holomorphe Funktionen $f, g : U \to \mathbb{C}$, die auf $I \subset \mathbb{R}$ (also als reelle Funktionen) übereinstimmen, bereits auf ganz U identisch sind.

Beispiel 17.7 Für den komplexen Logarithmus gilt $L(1+z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} z^{k+1}$ für alle $z \in K_1(0)$, denn $1 + z \in \mathbb{C}^-$, und die Gleichheit gilt auf dem reellen Intervall $]-1, 1[$.

Eine komplexe Funktion f, die überall auf $\mathbb{C}$ definiert und holomorph ist, heißt *ganze Funktion*. Nach Satz 17.2 gibt es für eine ganze Funktion f die Darstellung $f(z) = \sum_{n=0}^{\infty} a_n z^n$ mit Konvergenzradius ∞.

Satz 17.8 (Liouville) *Jede beschränkte ganze Funktion ist konstant.*

Beweis. Ist $|f| \leq M$ auf $\mathbb{C}$, dann erfüllen die Entwicklungskoeffizienten nach Satz 17.2 die Abschätzung $|a_n| \leq \frac{M}{r^n}$ für *beliebiges* $r > 0$. Also ist $a_n = 0$ für alle $n \geq 1$ und $f(z) = a_0$. □

Der Satz von Liouville hat kein Analogon in der reellen Differentialrechnung. Z.B. ist $f(x) = \sin x$ beliebig oft differenzierbar auf $\mathbb{R}$, beschränkt, und nichtkonstant.

Satz 17.9 (Fundamentalsatz der Algebra) *Jedes Polynom vom Grad ≥ 1 mit komplexen Koeffizienten besitzt in $\mathbb{C}$ eine Nullstelle.*

Beweis. Angenommen, das Polynom P habe keine Nullstelle, dann ist $\frac{1}{P}$ holomorph auf ganz $\mathbb{C}$. Außerdem ist $\frac{1}{P(z)} \to 0$ für $|z| \to \infty$, d.h. $\frac{1}{P(z)}$ ist beschränkt. Nach dem Satz von Liouville ist $\frac{1}{P}$ dann konstant, also wäre auch P konstant. Widerspruch. □

Nach Abdividieren der Nullstellen läßt sich somit jedes komplexe Polynom $P(z) = \sum_{k=0}^{n} a_k z^k$ vom Grad n, normiert auf $a_n = 1$, faktorisieren in $P(z) = \prod_{i=1}^{n}(z - b_i)$.

18 Der Residuensatz

Wichtig für die Ausnutzung der Cauchyschen Integralformel zur Berechnung von Integralen ist eine genauere Diskussion möglicher Singularitäten von komplexen Funktionen.

Satz 18.1 (Riemannscher Hebbarkeitssatz) *Es sei f eine auf $U \setminus \{a\}$ holomorphe Funktion, und es existiere eine Umgebung $V \subset U$ von $a \in U$, so daß f auf $V \setminus \{a\}$ beschränkt ist. Dann gibt es eine Fortsetzung $\tilde{f}$ von f, die holomorph auf ganz U ist.*

Beweis. Wir definieren eine Funktion $\phi : U \to \mathbb{C}$ durch

$$\phi(z) := \begin{cases} (z-a)^2 f(z) & \text{für } z \neq a \\ 0 & \text{für } z = a \end{cases}$$

Wegen der Beschränktheit von f auf $V \setminus \{a\}$ ist ϕ holomorph auf $V \setminus \{a\}$ und dann auf $U \setminus \{a\}$, und es gilt

$$\phi'(a) = \lim_{z \to a} \frac{\phi(z) - \phi(a)}{z - a} = 0 \, .$$

Damit besitzt ϕ die Potenzreihenentwicklung $\phi(z) = \sum_{n=2}^{\infty} a_n (z-a)^n$, und die Fortsetzung von f kann definiert werden als

$$\tilde{f}(z) := \sum_{n=0}^{\infty} a_{n+2} (z-a)^n \, . \qquad \square$$

Definition 18.2 Ist f holomorph in einer Umgebung $U \setminus \{a\}$ eines Punktes $a \in U$, so heißt a eine *isolierte Singularität von* f, und zwar:

i) Eine *hebbare Singularität*, wenn f holomorph in den Punkt a fortgesetzt werden kann.

ii) Ein *Pol*, wenn keine holomorphe Fortsetzung in a existiert, aber ein $k \in \mathbb{N} \setminus \{0\}$ derart, daß $(z-a)^k f$ holomorph in den Punkt a fortgesetzt werden kann. Die kleinste derartige Zahl k heißt die *Vielfachheit des Pols*. Der Punkt a ist genau dann ein k-facher Pol von f, wenn es in U eine Darstellung $f(z) = \frac{g(z)}{(z-a)^k}$ gibt, wobei g holomorph in U ist (insbesondere auch in a) und $g(a) \neq 0$ gilt.

iii) Eine *wesentliche Singularität*, wenn sie weder hebbar noch Pol ist.

Eine Funktion $f : U \to \mathbb{C}$ auf einer offenen Teilmenge $U \subset \mathbb{C}$ heißt *meromorph*, wenn sie bis auf Pole in U holomorph ist.

Jede rationale Funktion ist meromorph.

Ist f holomorph in $U \setminus \{a\}$ und liegt in a ein Pol der Ordnung k vor, dann hat die eindeutige Fortsetzung von f auf U die *Laurent-Reihenentwicklung*

$$f(z) = \sum_{n=-k}^{\infty} a_n (z-a)^n \ .$$

Die Laurent-Reihe ist konvergent in einem Kreisring

$$K_{R,r}(a) = \{z \in U \ , \ 0 < r < |z-a| < R\} \ .$$

Dabei ist R der Konvergenzradius des *Nebenteils der Laurent-Reihe* $\displaystyle\sum_{n=0}^{\infty} a_n (z-a)^n$.

Der *Hauptteil der Laurent-Reihe* $\displaystyle\sum_{n=-k}^{-1} a_n (z-a)^n = \sum_{n=1}^{k} \frac{a_{-n}}{(z-a)^n}$ ist als endliche Summe beschränkt in $K_{R,r}(a)$. Für die Entwicklungskoeffizienten gilt:

Satz 18.3 *Eine auf $U \setminus \{a\}$ holomorphe Funktion habe in a einen k-fachen Pol. Dann sind die Entwicklungskoeffizienten der Laurent-Reihe $f(z) = \displaystyle\sum_{n=-k}^{\infty} a_n (z-a)^n$ gegeben durch*

$$a_n = \frac{1}{2\pi \mathrm{i}} \int_{\partial K_r(a)} d\zeta \frac{f(\zeta)}{(\zeta - a)^{n+1}}$$

für einen beliebigen Kreis $K_r(a)$ um a mit Radius $r > 0$, so daß $\overline{K_r(a)} \subset U$.

Beweis. Die Funktion $h(z) = (z-a)^k f(z)$ läßt sich holomorph auf U fortsetzen und besitzt eine Potenzreihenentwicklung $h(z) = (z-a)^k f(z) = \sum_{n=0}^{\infty} a_n (z-a)^n$ mit

$$a_n = \frac{1}{2\pi \mathrm{i}} \int_{\partial K_r(a)} d\zeta \, \frac{h(\zeta)}{(\zeta - a)^{n+1}} = \frac{1}{2\pi \mathrm{i}} \int_{\partial K_r(a)} d\zeta \, \frac{f(\zeta)}{(\zeta - a)^{n-k+1}} \, .$$

Also ist

$$f(z) = \sum_{n=0}^{\infty} a_n (z-a)^{n-k} = \sum_{n=0}^{\infty} \Big(\frac{1}{2\pi \mathrm{i}} \int_{\partial K_r(a)} d\zeta \, \frac{f(\zeta)}{(\zeta - a)^{n-k+1}} \Big) (z-a)^{n-k}$$

Substitution $n - k \mapsto n$ liefert die Behauptung. □

Offenbar hat eine auf $U \setminus \{a\}$ holomorphe Funktion genau dann eine wesentliche Singularität in a, wenn der Hauptteil der Laurent-Entwicklung $f(z) = \sum_{n=-\infty}^{\infty} a_n (z-a)^n$ nicht abbricht. Ein Beispiel einer wesentlichen Singularität ist $e^{\frac{1}{z}} = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{1}{z^k}$ in $z = 0$. Eine Funktion f ist in jeder Umgebung einer wesentlichen Singularität stark oszillierend; ohne Beweis erwähnen wir:

Satz 18.4 (Picard) *Mit höchstens einer Ausnahme nimmt f in jeder Umgebung einer wesentlichen Singularität jede komplexe Zahl unendlich oft an.*

Für $e^{\frac{1}{z}}$ wird die Null nicht angenommen.

Definition 18.5 Es sei f eine auf $U \setminus \{a\}$ holomorphe Funktion mit Laurent-Reihenentwicklung $f(z) = \sum_{n=-\infty}^{\infty} a_n (z-a)^n$. Dann heißt der Koeffizient $a_{-1} = \mathrm{res}_a f$ das *Residuum* von f in a.

Offenbar ist $\mathrm{res}_a f = 0$, wenn f in a holomorph ist oder (wegen der Eindeutigkeit der Laurent-Reihe) in a eine hebbare Singularität besitzt. Nach Satz 18.3 gilt:

Satz 18.6 *Es sei f eine auf $U \setminus \{a\}$ holomorphe Funktion. Dann gilt*

$$\mathrm{res}_a f = \frac{1}{2\pi \mathrm{i}} \int_{\partial K_r(a)} f(z) dz$$

für einen beliebigen Kreis $K_r(a)$ um a mit Radius $r > 0$, so daß $\overline{K_r(a)} \subset U$. □

Wir geben Berechnungsvorschriften für das Residuum in einigen wichtigen Spezialfällen an:

- Ist $f(z) = \frac{g(z)}{z-a}$, und ist g holomorph in einer Umgebung von a, so folgt aus der Cauchyschen Integralformel $\mathrm{Res}_a f = g(a)$.

- Ist allgemeiner $f = \frac{g}{h}$ Quotient von in a holomorphen Funktionen g, h mit $h(a) = 0$ und $h'(a) \neq 0$, dann ist wegen der stetigen Differenzierbarkeit $h(z) = (z-a)(h'(a) + \phi(z))$ mit $\lim_{z\to a} \phi(z) = 0$. Es gibt also ein $r > 0$, so daß $|\phi(z)| < h'(a)$, so daß $\frac{1}{h'(a)+\phi(z)}$ holomorph auf $K_r(a)$ ist. Damit gilt
$$\operatorname{res}_a \frac{g}{h} = \frac{g(a)}{h'(a) + \phi(a)} = \frac{g(a)}{h'(a)} \ .$$

- Hat f in a einen k-fachen Pol, d.h. die Laurent-Reihe ist $f(z) = \sum_{n=-k}^{\infty} a_n (z-a)^n$, dann folgt
$$\operatorname{res}_a f = a_{-1} = \frac{1}{(k-1)!} \frac{d^{(k-1)}}{dz^{k-1}} \big((z-a)^k f(z)\big)\Big|_{z=a} \ .$$

Satz 18.7 (Residuensatz) *Es sei $U \subset \mathbb{C}$ offen, $S \subset U$ eine Teilmenge ohne Häufungspunkt in U und f holomorph auf $U \setminus S$. Sei $A \subset U$ eine Teilmenge, die in U einfach zusammenhängend ist und deren Rand $\gamma := \partial A$ in U liegt und keinen Punkt aus S trifft ($S \cap \gamma = \varnothing$). Dann gilt*
$$\int_\gamma f(z)dz = 2\pi\mathrm{i} \sum_{a\in S\cap A} \operatorname{res}_a f \ .$$

Beweis. Analog zur Cauchyschen Integralformel werden um jeden Punkt $a \in S \cap A$ Kreise $K_{\epsilon_a}(a)$ gelegt, die im Inneren von A liegen und sich nicht schneiden (S hat keinen Häufungspunkt). Dann ist f holomorph auf einer Umgebung von $\Gamma := A \setminus \bigcup_{a\in S\cap A} K_{\epsilon_a}(a)$. Durch Aufschneiden von Γ zwischen γ und jedem Kreis $K_{\epsilon_a}(a)$ entsteht ein einfach zusammenhängendes Gebiet, so daß das Integral von f über dessen Rand verschwindet. Die Schnitte werden zweimal in umgekehrter Richtung durchlaufen, so daß sich die Integrale gegenseitig aufheben. Die Integrale über die $\partial K_{\epsilon_a}(a)$ ergeben bis auf einen Faktor $-2\pi\mathrm{i}$ (Durchlauf in umgekehrter Richtung) das jeweilige Residuum $\operatorname{Res}_a f$. □

Der Residuensatz ist ein mächtiges Werkzeug zur Berechnung von Integralen.

Beispiel 18.8 Gesucht ist $I(a) := \displaystyle\int_0^{2\pi} \frac{dt}{a + \cos t}$ für $a > 1$. Da auf dem Einheitskreis $z(t) = e^{it}$ gilt $\cos t = \frac{1}{2}(z(t) + \frac{1}{z(t)})\big|_{\partial K_1(0)}$. Dann ist $\dot z(t) = \mathrm{i}z(t)$, und wir erhalten
$$I(a) = \int_0^{2\pi} dt\ \frac{1}{a + \frac{1}{2}(z(t) + \frac{1}{z(t)})} \frac{\dot z(t)}{\mathrm{i}z(t)} = \frac{2}{\mathrm{i}} \int_{\partial K_1(0)} \frac{dz}{z^2 + 2az + 1} = \frac{2}{\mathrm{i}} \int_{\partial K_1(0)} dz\ f(z)$$

mit $f(z) := \dfrac{1}{(z-(-a-\sqrt{a^2-1}))(z-(-a+\sqrt{a^2-1}))}$. Nur die Polstelle bei $z=\sqrt{a^2-1}-a$ liegt im Inneren des Einheitskreises, also folgt

$$\int_0^{2\pi} \frac{dt}{a+\cos t} = 4\pi \mathrm{res}_{\sqrt{a^2-1}-a} f = \frac{2\pi}{\sqrt{a^2-1}} \,. \qquad \triangleleft$$

Allgemein gilt für solche Art von Integralen:

Satz 18.9 *Es sei $R(x,y)$ eine rationale Funktion in zwei Variablen und $R(\cos t, \sin t)$ sei für alle $t \in [0,2\pi]$ erklärt. Dann gilt*

$$\int_0^{2\pi} dt\; R(\cos t, \sin t) = 2\pi \sum_{a \in K_1(0)} \mathrm{res}_a \tilde{R} \,, \qquad \tilde{R}(z) := \frac{1}{z} R\Big(\frac{1}{2}\Big(z+\frac{1}{z}\Big), \frac{1}{2\mathrm{i}}\Big(z-\frac{1}{z}\Big)\Big) \,.$$

Eine andere wichtige Klasse von reellen Integralen, die mit dem Residuensatz berechnet werden können, ist die folgende:

Satz 18.10 *Es sei R eine rationale Funktion (einer Variablen), die auf der reellen Achse keinen Pol habe und in ∞ eine mindestens zweifache Nullstelle, d.h. wenn $R(x) = P(x)/Q(x)$ mit Polynomen P, Q, dann ist $\deg(Q) - \deg(P) \geq 2$. In diesem Fall gilt*

$$\int_{-\infty}^{\infty} dx\; R(x) = 2\pi\mathrm{i} \sum_{a \in H} \mathrm{res}_a R \,,$$

wobei $H = \{z \in \mathbb{C} \,:\, \mathrm{Im}(z) \geq 0\}$ die obere Halbebene ist.

Beweis. Man integriert über den Halbkreis bestehend aus dem Durchmesser $[-r, r]$ auf der reellen Achse und dem halben Umfang $z = re^{it}$ mit $t \in [0, \pi]$. Dabei wird r so groß gewählt, daß alle Pole in H von $R(z)$ im Inneren des Halbkreises liegen. Nach Voraussetzung verschwindet dann für $r \to \infty$ das Integral über den Halbkreisbogen. $\square$

Beispiel 18.11 Gesucht ist $I_n := \displaystyle\int_{-\infty}^{\infty} \frac{dx}{1+x^{2n}}$ mit $n \in \mathbb{N} \setminus \{0\}$. Aufgefaßt als komplexe Funktion sind die (einfachen) Pole $z^{2n} = -1$ der oberen Halbebene bei $a_k = e^{\frac{i\pi(2k+1)}{2n}}$, $k = 0, 1, \ldots, n-1$. Also gilt mit $\mathrm{res}_a(\frac{1}{h}) = \frac{1}{h'(a)}$

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^{2n}} = \sum_{k=0}^{n-1} \frac{2\pi\mathrm{i}}{2n e^{\frac{\mathrm{i}\pi(2k+1)(2n-1)}{2n}}} = -\frac{\pi\mathrm{i}}{n} e^{\frac{\mathrm{i}\pi}{2n}} \sum_{k=0}^{n-1} e^{\frac{\mathrm{i}\pi k}{n}} = \frac{\pi}{\mathrm{i}n} e^{\frac{\mathrm{i}\pi}{2n}} \frac{1-e^{\frac{\mathrm{i}\pi n}{n}}}{1-e^{\frac{\mathrm{i}\pi}{n}}} = \frac{\frac{\pi}{n}}{\sin\frac{\pi}{2n}} \,.$$

$\triangleleft$

Integrale der Form $I = \int_0^\infty dx\; R(x)$, wobei R eine gerade Funktion ist, werden über $I = \frac{1}{2}\int_{-\infty}^{\infty} dx\; R(x)$ ausgerechnet.

Satz 18.12 *Es sei R eine rationale Funktion ohne Pol auf der reellen Achse und mit mindestens einfacher Nullstelle in ∞. Dann ist für jedes $\alpha > 0$ das folgende Integral existent und durch den Residuensatz berechenbar zu*

$$\int_{-\infty}^{\infty} dx\ R(x)e^{\mathrm{i}\alpha x} = 2\pi\mathrm{i}\sum_{a\in H} \mathrm{res}_a\Big(R(z)e^{\mathrm{i}\alpha z}\Big)\ .$$

Beweis. Das zugehörige bestimmte Integral über $[-r,r]$ wird durch ein Quadrat in der oberen Halbebene geschlossen. Die entsprechenden Kurvenstücke sind $\gamma_1(t) = r+\mathrm{i}t$ mit $t\in[0,r]$, $\gamma_2(t) = t+\mathrm{i}r$ mit $t\in[-r,r]$ und $\gamma_3(t) = -r+\mathrm{i}t$ mit $t\in[0,r]$. Unter Beachtung des Umlaufsinns liefert der Residuensatz

$$\begin{aligned}2\pi\mathrm{i}\sum_{a\in H} \mathrm{res}_a\Big(R(z)e^{\mathrm{i}\alpha z}\Big) &= \int_{-r}^{r} dx\ R(x)e^{\mathrm{i}\alpha x} + \mathrm{i}\int_0^r dt\ R(r{+}\mathrm{i}t)e^{\mathrm{i}\alpha(r+\mathrm{i}t)}\\ &\quad - \int_{-r}^{r} dt\ R(t{+}\mathrm{i}r)e^{\mathrm{i}\alpha(t+\mathrm{i}r)} - \mathrm{i}\int_0^r dt\ R(-r{+}\mathrm{i}t)e^{\mathrm{i}\alpha(-r+\mathrm{i}t)}\ .\end{aligned}$$

Da R eine mindestens einfache Nullstelle in ∞ hat, gilt $|R(z)| = R(|z|) \le \frac{M}{1+|z|}$ und damit

$$\begin{aligned}\Big|\int_0^r dt\ R(\pm r{+}\mathrm{i}t)e^{\mathrm{i}\alpha(\pm r+\mathrm{i}t)}\Big| &\le \frac{M}{1+r}\int_0^r dt\ e^{-\alpha t} \le \frac{M}{\alpha(1+r)}\ ,\\ \Big|\int_{-r}^r dt\ R(t{+}\mathrm{i}r)e^{\mathrm{i}\alpha(t+\mathrm{i}r)}\Big| &\le \frac{M}{1+r}e^{-\alpha r}\cdot 2r\ .\end{aligned}$$

Folglich verschwinden die Integrale über γ_i im Limes $r\to\infty$. Die Existenz des Limes $r\to\infty$ für das reelle Integral ergibt sich nach partieller Integration und Verwendung von $|R'(x)| \le \frac{M}{(1+|x|)^2}$. □

Nach Zerlegung in Real- und Imaginärteil können die reellen Integrale

$$\begin{aligned}\int_{-\infty}^{\infty} dx\ R(x)\cos(\alpha x) &= \mathrm{Re}\Big(2\pi\mathrm{i}\sum_{a\in H}\mathrm{res}_a\Big(R(z)e^{\mathrm{i}\alpha z}\Big)\Big)\ ,\\ \int_{-\infty}^{\infty} dx\ R(x)\sin(\alpha x) &= \mathrm{Im}\Big(2\pi\mathrm{i}\sum_{a\in H}\mathrm{res}_a\Big(R(z)e^{\mathrm{i}\alpha z}\Big)\Big)\end{aligned}$$

berechnet werden.

Beispiel 18.13 Gesucht ist $\displaystyle\int_{-\infty}^{\infty}\frac{x\sin x}{1+x^2} = \mathrm{Im}\Big(\int_{-\infty}^{\infty}\frac{xe^{\mathrm{i}x}}{1+x^2}\Big)$. Die Voraussetzungen von Satz 18.12 sind erfüllt, es gibt einen Pol bei $x=\mathrm{i}$ in der oberen Halbebene, so daß gilt

$$\int_{-\infty}^{\infty}\frac{xe^{\mathrm{i}x}}{1+x^2} = 2\pi\mathrm{i}\ \mathrm{res}_{\mathrm{i}}\Big(\frac{ze^{\mathrm{i}z}}{(z+\mathrm{i})(z-\mathrm{i})}\Big) = 2\pi\mathrm{i}\frac{\mathrm{i}e^{\mathrm{i}\cdot\mathrm{i}}}{2\mathrm{i}} = \frac{\pi\mathrm{i}}{e}\ .$$

Somit gilt $\displaystyle\int_{-\infty}^{\infty}\frac{x\sin x}{1+x^2} = \frac{\pi}{e}$. ◁

Teil V
Das Lebesgue-Integral

19 Treppenfunktionen und Hüllreihen

Ein n-dimensionaler Quader $Q \subset \mathbb{R}^n$ ist das direkte Produkt $Q = I_1 \times \cdots \times I_n$ zusammenhängender beschränkter nichtleerer Intervalle $I_1, \ldots, I_n \subset \mathbb{R}^n$. Dabei dürfen die Intervalle, offen, abgeschlossen, halboffen und zu einem Punkt entartet sein. Die Länge eines Intervalls I ist definiert als

$$|I| := \operatorname{diam}(I) = \sup_{t_1,t_2 \in I} |t_1 - t_2| \, .$$

Dann definieren wir das Volumen eines Quaders zu

$$v(I_1 \times \cdots \times I_n) := |I_1| \cdots |I_n| \, .$$

Diese Definition ist bereits nichttrivial, denn z.B. haben das offene Intervall $]a, b[$ und das abgeschlossene Intervall $[a, b]$ das gleiche Volumen $|a - b|$. Wichtig ist, daß die Volumina sich additiv unter Zerlegungen verhalten. Sind I, I_1, I_2 zusammenhängende Intervalle mit $I = I_1 \cup I_2$, dann ist $|I| = |I_1| + |I_2|$. Damit gilt für die Quader $v(Q_1 \cup Q_2) = v(Q_1) + v(Q_2)$. Ausgeartete Quader, für die mindestens eine Kante aus einem zu einem Punkt entarteten Intervall $I = \{t\}$, $t \in \mathbb{R}$, besteht, haben das Volumen Null.

Definition 19.1 Eine Funktion $\phi : \mathbb{R}^n \to \mathbb{C}$ heißt *Treppenfunktion*, falls es endlich viele paarweise disjunkte Quader $Q_1, \ldots, Q_k$ gibt, so daß

i) Für jedes $1 \leq i \leq k$ ist die Einschränkung $\phi|_{Q_i}$ von ϕ auf Q_i eine konstante Funktion.

ii) $\phi(x) = 0$ für alle $x \in \mathbb{R}^n \setminus (\bigcup_{i=1}^n Q_i)$.

Durch Teilungen der Quader kann man stets erreichen, daß zwei beliebige Treppenfunktion ϕ_1, ϕ_2 bezüglich des gleichen Satzes paarweise disjunkter Quader $\{Q_1, \ldots, Q_k\}$ definiert sind. Daraus folgt, daß eine endliche Linearkombinationen von Treppenfunktionen wieder eine Treppenfunktion ist. Somit bildet die Menge aller Treppenfunktionen einen Vektorraum. Sind ϕ_1, ϕ_2 Treppenfunktionen, so auch $|\phi_1|$ sowie $\max(\phi_1, \phi_2)$ und $\min(\phi_1, \phi_2)$ Treppenfunktionen.

Es ist bequem, die *charakteristische Funktion* einer Teilmenge $A \subset \mathbb{R}^n$ einzuführen als

$$\mathbf{1}_A : \mathbb{R}^n \to \mathbb{R} \, , \qquad \mathbf{1}_A(a) := \begin{cases} 1 & \text{für } x \in A \, , \\ 0 & \text{für } x \notin A \, . \end{cases}$$

Dann kann jede Treppenfunktion ϕ als Linearkombination $\phi = \sum_{i=1}^k c_i \mathbf{1}_{Q_i}$ mit $c_i \in \mathbb{C}$ geschrieben werden. Wegen der Möglichkeit der Zerteilung von Quadern ist eine solche Darstellung aber nicht eindeutig.

Definition 19.2 Das *Integral einer Treppenfunktion* $\phi = \sum_{i=1}^{k} c_i \mathbf{1}_{Q_i} : \mathbb{R}^n \to \mathbb{C}$ ist definiert als

$$\int_{\mathbb{R}^n} dx\ \phi(x) := \sum_{i=1}^{k} c_i\, v(Q_i)\ .$$

Satz 19.3 *Das so definierte Integral einer Treppenfunktion ist unabhängig von der Darstellung der Treppenfunktion. Außerdem gilt für Treppenfunktionen* ϕ, ψ

i) $\displaystyle\int_{\mathbb{R}^n} dx\ (\alpha\phi + \beta\psi)(x) = \alpha \int_{\mathbb{R}^n} dx\ \phi(x) + \beta \int_{\mathbb{R}^n} dx\ \psi(x)\ , \qquad \alpha, \beta \in \mathbb{C}\ .$

ii) $\displaystyle\Big| \int_{\mathbb{R}^n} dx\ \phi(x) \Big| \le \int_{\mathbb{R}^n} dx\ |\phi(x)|\ .$

iii) *Sind* ϕ, ψ *reellwertig mit* $\phi \le \psi$, *so gilt* $\displaystyle\int_{\mathbb{R}^n} dx\ \phi(x) \le \int_{\mathbb{R}^n} dx\ \psi(x)\ .$

Der Beweis der Unabhängigkeit des Integrals von der Darstellung ist länglich (siehe Königsberger: Analysis II, §7.1, Satz 7.1) und wird hier nicht wiedergegeben. Die Eigenschaften i)–iii) sind dann klar.

Wir verallgemeinern Treppenfunktionen nun zu *Hüllreihen* durch folgende Schritte:

- Die Quader dürfen sich überlappen.
- Wir lassen abzählbare (unendliche) Linearkombinationen von $\mathbf{1}_{Q_i}$ zu, $\Phi = \sum_{i\in\mathbb{N}} c_i \mathbf{1}_{Q_i}$ (das sind dann im allgemeinen keine Treppenfunktionen mehr).
- "Unendlich" wird als Funktionswert zugelassen. Dabei setzt man $c+\infty = \infty$ für alle $c \in \mathbb{C}\cup\{\infty\}$ und $0\cdot\infty = 0$ sowie $c\cdot\infty = \infty$ für $c \in \mathbb{C}^*\cup\{\infty\}$.

Definition 19.4 Eine Reihe $\Phi = \sum_{k\in\mathbb{N}} c_k \mathbf{1}_{Q_k}$ heißt *Hüllreihe* einer Funktion $f : \mathbb{R}^n \to \mathbb{C}\cup\{\infty\}$, wenn gilt:

i) Die Quader $Q_k \in \mathbb{R}^n$ sind offen und die $c_k \in \mathbb{R}_+$ sind nichtnegative reelle Zahlen.

ii) Es gilt $|f(x)| \le \Phi(x) = \sum_{k\in\mathbb{N}} c_k \mathbf{1}_{Q_k}(x)$ für alle $x \in \mathbb{R}^n$.

Der *Inhalt* einer Hüllreihe $\Phi = \sum_{k\in\mathbb{N}} c_k \mathbf{1}_{Q_k}$ ist definiert als $I(\Phi) := \sum_{k\in\mathbb{N}} c_k\, v(Q_k) \in \mathbb{R}^+\cup\{\infty\}$, mit $I(\Phi) := \infty$, falls die Reihe nicht konvergiert.

Jede Funktion $f : \mathbb{R}^n \to \mathbb{C}\cup\{\infty\}$ besitzt eine Hüllreihe, z.B. $\Phi = \sum_{k=1}^{\infty} \mathbf{1}_{\hat{Q}_k}$, wobei $\hat{Q}_k \subset \mathbb{R}^n$ der offene Würfel mit Mittelpunkt 0 und Kantenlänge k ist.

Definition 19.5 Sei $f : \mathbb{R}^n \to \mathbb{C}\cup\{\infty\}$ eine beliebige Funktion. Dann ist die L^1-*Halbnorm* von f erklärt als das Infimum der Inhalte von Hüllreihen zu f,

$$\|f\|_1 := \inf\{I(\Phi)\ :\ \Phi \text{ ist Hüllreihe zu } f\}\ .$$

Entscheidend für Besonderheiten des Lebesgue-Integrals ist folgendes Beispiel:

Beispiel 19.6 Es sei $f = \mathbf{1}_Q$ die charakteristische Funktion eines ausgearteten Quaders Q (eine Kante $I_j = \{a\}$ hat die Länge Null). Sei $Q^{(k)}$ ein offener Quader, der Q enthält, wobei die j-te Kante von $Q^{(k)}$ gegeben ist durch $I_j^{(k)} =]a - \frac{1}{k+1}, a + \frac{1}{k+1}[$. Dann ist $\Phi_k = \mathbf{1}_{Q^{(k)}}$ eine Hüllreihe für jedes $k \in \mathbb{N}$. Bleiben alle anderen Kanten festgehalten, dann gilt $v(Q^{(k)}) = \frac{1}{k+1} v(Q^{(0)})$. Nach Definition des Infimums ist $0 \leq \|\mathbf{1}_Q\|_1 \leq v(Q^{(k)}) = \frac{1}{k+1} v(Q^{(0)})$ für alle $k \in \mathbb{N}$, also $\|\mathbf{1}_Q\|_1 = 0$. ◁

Daraus folgt, daß die L^1-Halbnorm *keine Norm* ist, denn $\|f\|_1 = 0 \not\Rightarrow f = 0$. Es gelten aber die anderen Normeigenschaften:

Satz 19.7 *Für $f, g : \mathbb{R}^n \to \mathbb{C} \cup \{\infty\}$ und $c \in \mathbb{C}$ gilt:*

i) $\|cf\|_1 = |c| \|f\|_1$

ii) $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$

iii) *Aus $|f(x)| \leq |g(x)|$ für alle $x \in \mathbb{R}^n$ folgt $\|f\|_1 \leq \|g\|_1$.*

iv) *Für nichtnegative reellwertige Funktionen $f_k : \mathbb{R}^n \to \mathbb{R}_+ \cup \{\infty\}$ gilt* $\Big\| \sum_{k=0}^{\infty} f_i \Big\|_1 \leq \sum_{k=0}^{\infty} \|f_i\|_1$.

Beweis. i) und iii) sind klar, und ii) folgt wegen $|f(x) + g(x)| \leq |f(x)| + |g(x)|$ aus iii) und iv).

iv) Zu gegebenem $\epsilon > 0$ existiert nach Definition des Infimums zu f_k eine Hüllreihe $\Phi_k = \sum_i c_{ki} \mathbf{1}_{Q_{ki}}$ mit $I(\Phi_k) = \sum_i c_{ki} v(Q_{ki}) \leq \|f_k\|_1 + \frac{\epsilon}{2^k}$. Dann ist $\Phi := \sum_{k=0}^{\infty} \Phi_k = \sum_{k=0}^{\infty} \Big(\sum_i c_{ki} \mathbf{1}_{Q_{ki}} \Big)$ Hüllreihe zu $\sum_{k=0}^{\infty} f_k$, mit

$$I(\Phi) = \sum_{k=0}^{\infty} \Big(\sum_i c_{ki} v(Q_{ki}) \Big) \leq \sum_{k=0}^{\infty} \Big(\|f_k\|_1 + \frac{\epsilon}{2^k} \Big) = \sum_{k=0}^{\infty} \|f_k\|_1 + 2\epsilon \, .$$

Somit gilt $\Big\| \sum_{k=0}^{\infty} f_k \|_1 \leq \sum_{k=0}^{\infty} \|f_k\|_1 + 2\epsilon$ für alle $\epsilon > 0$, und daraus folgt iv). □

Satz 19.8 *Für einen abgeschlossenen Quader $A \subset \mathbb{R}^n$ gilt*
$\|\mathbf{1}_A\|_1 = v(A) = \int_{\mathbb{R}^n} dx \, \mathbf{1}_A(x)$.

Beweis. i) Zu jedem $\epsilon > 0$ gibt es einen offenen Quader $Q_\epsilon \supset A$ mit $v(Q_\epsilon) \leq v(A) + \epsilon$. Damit ist $\|\mathbf{1}_A\|_1 \leq v(Q_\epsilon) \leq v(A) + \epsilon$ für alle $\epsilon > 0$, also $\|\mathbf{1}_A\|_1 \leq v(A)$.

ii) Zu zeigen ist, daß auch durch eine unendliche Zerteilung von A die entsprechende Hüllreihe $\Phi = \sum_{k=0}^{\infty} c_k \mathbf{1}_{Q_k}$ zu $\mathbf{1}_A$, mit Q_k offen, keinen kleineren Inhalt haben kann. Wegen $\Phi(x) \geq 1$ für alle $x \in A$ gibt es zu gegebenem $\epsilon > 0$ einen

Index $N(x)$ mit $\sum_{k=0}^{N(x)} c_k \mathbf{1}_{Q_k}(x) \geq 1-\epsilon$. Da die Q_k offen sind und der Durchschnitt endlich vieler offener Mengen offen ist, gibt es eine Umgebung $U(x) \subset A$ von x, so daß wegen der Konstantheit von $\mathbf{1}_{Q_k}$ auf Q_k gilt

$$\sum_{k=0}^{N(x)} c_k \mathbf{1}_{Q_k}(x) \geq 1-\epsilon \qquad \text{für alle } x \in U(x) \; .$$

Da A kompakt ist, wird A durch endlich viele $U(x_1), \ldots, U(x_p)$ überdeckt. Also gibt es ein $N = \max(N(x_1), \ldots, N(x_p))$, so daß $\sum_{k=0}^{N} c_k \mathbf{1}_{Q_k}(x) \geq 1-\epsilon = (1-\epsilon)\mathbf{1}_A(x)$ für alle $x \in A$. Da $\sum_{k=0}^{N} c_k \mathbf{1}_{Q_k}(x)$ eine Treppenfunktion ist, folgt mit Satz 19.3.iii)

$$I(\Phi) = \sum_{k=0}^{\infty} c_k v(Q_k) \geq \sum_{k=0}^{N} c_k v(Q_k) \geq (1-\epsilon) v(A) \; .$$

Also gilt $\|\mathbf{1}_A\|_1 = v(A) = \int_{\mathbb{R}^n} dx \; \mathbf{1}_A(x)$. □

Satz 19.9 *Für jede Treppenfunktion $\phi = \sum_{i=1}^{k} c_i \mathbf{1}_{Q_i}$ gilt*

$$\|\phi\|_1 = \int_{\mathbb{R}^n} dx \; |\phi(x)| \; .$$

Beweis. Da $\|\phi\|_1 = \||\phi|\|_1$, können wir $\phi(x) \geq 0$ annehmen.

i) Wir zerlegen die beliebigen Quader Q_i, deren Kanten offen, abgeschlossen, halboffen und entartet sein dürfen, in disjunkte offene Quader und disjunkte Ränder, so daß

$$\phi = \sum_{i=1}^{k'} c_i' \mathbf{1}_{Q_i^o} + \sum_{j=1}^{l} d_j \mathbf{1}_{R_j} \; , \qquad c_i', d_j > 0 \; .$$

Zu jedem der entarteten Quader R_j mit Volumen 0 wählen wir einen nichtentarteten offenen Quader $R_j^o \supset R_j$ mit $v(R_j^0) < \epsilon$. Dann ist $\Phi_\epsilon := \sum_{i=0}^{k'} c_i' \mathbf{1}_{Q_i^o} + \sum_{j=0}^{l} d_j \mathbf{1}_{R_j^0}$ eine Hüllreihe zu ϕ mit

$$\|\phi\|_1 \leq \inf_\epsilon I(\Phi_\epsilon) = \inf_\epsilon \Big(\sum_{i=1}^{k'} c_i' v(Q_i^o) + \epsilon \sum_{j=1}^{l} d_j \Big) = \int_{\mathbb{R}^n} dx \; \phi(x) \; .$$

ii) Sei A ein abgeschlossener Quader derart, daß $\phi(x) = 0$ für alle $x \notin A$, und sei $m = \max_{x \in A} \phi(x)$. Dann ist $\psi := m\mathbf{1}_A - \phi$ wieder eine nichtnegative Treppenfunktion, so daß $-\int_{\mathbb{R}^n} dx \; \psi(x) \leq -\|\psi\|_1$ und daraus

$$\begin{aligned} \int_{\mathbb{R}^n} dx \; \phi(x) &= \int_{\mathbb{R}^n} dx \; (m\mathbf{1}_A(x) - \psi(x)) = m \int_{\mathbb{R}^n} dx \; \mathbf{1}_A(x) - \int_{\mathbb{R}^n} dx \; \psi(x) \\ &\leq \|m1_A\|_1 - \|\psi\|_1 = \|\phi + \psi\|_1 - \|\psi\|_1 \leq \|\phi\|_1 \; . \end{aligned}$$

□

20 Das Lebesgue-Integral

Definition 20.1 Eine Funktion $f : \mathbb{R}^n \to \mathbb{C} \cup \{\infty\}$ heißt *(Lebesgue-)integrierbar*, wenn es eine Folge von Treppenfunktionen $(\phi_k)_{k\in\mathbb{N}}$ gibt mit $\lim_{k\to\infty} \|f - \phi_k\|_1 = 0$. In diesem Fall heißt

$$\int_{\mathbb{R}^n} dx\ f(x) := \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ \phi_k(x)$$

das *Lebesgue-Integral* von f.

Zur Definition ist zu bemerken, daß die Folge der Integrale $\Big(\int_{\mathbb{R}^n} dx\ \phi_k(x)\Big)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{C}$ ist: Wegen der L^1-Konvergenz gegen f gibt es zu jedem $\epsilon > 0$ ein $k \in \mathbb{N}$, so daß $\|f - \phi_l\|_1 < \frac{\epsilon}{2}$ für alle $l \geq k$. Dann gilt für alle $m, l \geq k$

$$\begin{aligned}\Big|\int_{\mathbb{R}^n} dx\ \phi_l(x) - \int_{\mathbb{R}^n} dx\ \phi_m(x)\Big| &= \Big|\int_{\mathbb{R}^n} dx\ (\phi_l - \phi_m)(x)\Big| \leq \int_{\mathbb{R}^n} dx\ |(\phi_l - \phi_m)(x)| \\ &= \|\phi_l - \phi_m\|_1 \leq \|\phi_l - f\|_1 + \|f - \phi_m\|_1 < \epsilon\ .\end{aligned}$$

Wegen der Vollständigkeit der komplexen Zahlen existiert der Grenzwert und damit das Lebesgue-Integral von f, und das Integral ist endlich.

Offenbar ist jede Treppenfunktion integrierbar. Die folgende Eigeschaft des Lebesgue-Integrals ist grundlegend verschieden vom Riemann-Integral:

Satz 20.2 *Ist $f : \mathbb{R}^n \to \mathbb{C} \cup \{\infty\}$ integrierbar, dann ist auch $|f|$ integrierbar, und es gilt*

$$\Big|\int_{\mathbb{R}^n} dx\ f(x)\Big| \leq \int_{\mathbb{R}^n} dx\ |f(x)| = \|f\|_1\ .$$

Beweis. Sei $(\phi_k)_{k\in\mathbb{N}}$ eine Folge von Treppenfunktionen mit $\lim_{k\to\infty} \|f - \phi_k\|_1 = 0$. Wegen $|(|f| - |\phi_k|)| \leq |f - \phi_k|$ gilt $\|(|f| - |\phi_k|)\|_1 \leq \|f - \phi_k\|_1$ und insbesondere $\lim_{k\to\infty} \|(|f| - |\phi_k|)\|_1 = 0$. Also ist $|f|$ integrierbar. Für die Integrale gilt die Abschätzung

$$\begin{aligned}\Big|\int_{\mathbb{R}^n} dx\ f(x)\Big| &= \Big|\lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ \phi_k(x)\Big| = \lim_{k\to\infty} \Big|\int_{\mathbb{R}^n} dx\ \phi_k(x)\Big| \leq \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ |\phi_k(x)| \\ &= \int_{\mathbb{R}^n} dx\ |f|(x)\ .\end{aligned}$$

Wegen $\|\phi_k\|_1 \leq \|\phi_k - f\|_1 + \|f\|_1 \leq 2\|\phi_k - f\|_1 + \|\phi_k\|_1$ gilt mit Satz 19.9

$$\Big(\int_{\mathbb{R}^n} dx\ |\phi_k|(x)\Big) - \|\phi_k - f\|_1 \leq \|f\|_1 \leq \Big(\int_{\mathbb{R}^n} dx\ |\phi_k|(x)\Big) + \|\phi_k - f\|_1\ .$$

Für $k \to \infty$ erhalten wir $\|f\|_1 = \int_{\mathbb{R}^n} dx\ |f(x)|$. □

Bemerkung: Aus der Integrierbarkeit von $|f|$ folgt *nicht* die Integrierbarkeit von f, denn z.B. könnte f in einem beschränkten Intervall I unkontrolliert zwischen $+1$ und -1 springen, während $|f| = 1$ als Treppenfunktion über I integrierbar ist.

Satz 20.3 *Das Lebesgue-Integral erfüllt die folgenden Rechenregeln für integrierbare Funktionen $f, g : \mathbb{R}^n \to \mathbb{C} \cup \{\infty\}$ und $\alpha, \beta \in \mathbb{C}$:*

i) $\displaystyle\int_{\mathbb{R}^n} dx\ (\alpha f + \beta g)(x) = \alpha \int_{\mathbb{R}^n} dx\ f(x) + \beta \int_{\mathbb{R}^n} dx\ g(x)\ .$

ii) *Sind f, g reellwertig mit $f(x) \leq g(x)$ für alle $x \in \mathbb{R}^n$, so folgt* $\displaystyle\int_{\mathbb{R}^n} dx\ f(x) \leq \int_{\mathbb{R}^n} dx\ g(x)$.

iii) *Ist g beschränkt, so ist $f \cdot g$ integrierbar.*

Beweis. i) ergibt sich durch approximierende Treppenfunktionen. ii) folgt aus $g - f = |g - f|$ und $\displaystyle\int_{\mathbb{R}^n} dx\ (g - f) = \int_{\mathbb{R}^n} dx\ |g - f| = \|g - f\|_1 \geq 0$ nach Satz 20.2.

iii) Sei $|g(x)| \leq M < \infty$ und seien ϕ_k, ψ_k Treppenfunktionen, die in der L^1-Halbnorm gegen f, g konvergieren. Für $\epsilon > 0$ gibt es ein $k \in \mathbb{N}$, so daß $\|f - \phi_k\|_1 < \frac{\epsilon}{2M}$. Sei dann $\mu := \max_{x \in \mathbb{R}^n} |\phi_k(x)|$. Dann gibt es ein $l \in \mathbb{N}$, so daß $\|g - \psi_l\|_1 < \frac{\epsilon}{2\mu}$. Es gilt

$$\begin{aligned} |(fg - \phi_k\psi_l)(x)| &= |(f - \phi_k)(x)g(x) + \phi_k(x)(g - \psi_l)(x)| \\ &\leq M|(f - \phi_k)(x)| + \mu|(g - \psi_l)(x)| \end{aligned}$$

und deshalb

$$\|fg - \phi_k\psi_l\|_1 \leq \|(M|f - \phi_k| + \mu|g - \psi_l|)\| \leq M\|f - \phi_k\|_1 + \mu\|g - \psi_l\|_1 < \epsilon\ .$$

□

Insbesondere gilt:

- $f : \mathbb{R}^n \to \mathbb{C} \cup \{\infty\}$ ist genau dann integrierbar, wenn $\mathrm{Re}(f)$ und $\mathrm{Im}(f)$ integrierbar sind, und es gilt

$$\int_{\mathbb{R}^n} dx\ f(x) = \int_{\mathbb{R}^n} dx\ (\mathrm{Re}\, f)(x) + \mathrm{i} \int_{\mathbb{R}^n} dx\ (\mathrm{Im}\, f)(x)$$

- Sind f, g integrierbare reellwertige Funktionen, dann sind auch $\max(f, g) = \frac{1}{2}(f + g + |f - g|)$ und $\min(f, g) = \frac{1}{2}(f + g - |f - g|)$ integrierbar sowie $f^+ := \max(f, 0)$ und $f^- := \max(-f, 0)$. Es gilt $f^+ \geq 0$, $f^- \geq 0$ sowie $f = f^+ - f^-$ und $|f| = f^+ + f^-$. Somit gilt auch: $f = (\mathrm{Re}\, f)^+ - (\mathrm{Re}\, f)^- + \mathrm{i}(\mathrm{Im}\, f)^+ - \mathrm{i}(\mathrm{Im}\, f)^-$ ist genau dann integrierbar, wenn alle vier nichtnegativen Anteile $(\mathrm{Re}\, f)^\pm, (\mathrm{Im}\, f)^\pm$ integrierbar sind.

Die Integration einer Funktion $f : A \to \mathbb{C} \cup \{\infty\}$ über eine Teilmenge $A \subset \mathbb{R}^n$ wird zurückgeführt auf die Integration im $\mathbb{R}^n$ durch

$$f_A(x) := \begin{cases} f(x) & \text{für } x \in A\,, \\ 0 & \text{für } x \notin A\,. \end{cases}$$

Dann heißt f integrierbar über $A \subset \mathbb{R}^n$, wenn f_A integrierbar ist, und wir setzen $\int_A dx\ f(x) := \int_{\mathbb{R}^n} dx\ f_A(x)$.

Riemann-integrierbare Funktionen auf kompakten Intervallen sind auch Lebesgue-integrierbar:

Satz 20.4 *Es sei $f : [a,b] \to \mathbb{C}$ eine über $[a,b]$ Riemann-integrierbare Funktion. Dann ist f über $[a,b]$ auch Lebesgue-integrierbar, und Riemann- und Lebesgue-Integral stimmen überein, $\int_{[a,b]} dx\ f(x) = \int_a^b dx\ f(x)$.*

Beweis. Wir können f als reellwertig annehmen. Nach Satz 19.6 aus dem 1. Semester gibt es zu $\epsilon = \frac{1}{k+1}$ Treppenfunktionen ϕ_k, ψ_k mit $\phi_k \le f \le \psi_k$ und $\int_a^b dx(\psi_k - \phi_k)(x) \le \dfrac{1}{k+1}$. Da für Treppenfunktionen Lebesgue- und Riemann-Integral übereinstimmen, gilt $\|\psi_k - \phi_k\|_1 \le \frac{1}{k+1}$ nach Satz 19.9. Aus $0 \le \psi_k - f \le \psi_k - \phi_k$ folgt $\|\psi_k - f\|_1 \le \|\psi_k - \phi_k\|_1 \le \frac{1}{k+1}$ nach Satz 19.7. Somit ist f Lebesgue-integrierbar, und

$$\int_{[a,b]} dx\ f(x) = \lim_{k\to\infty} \int_a^b dx\ \psi_k(x) = \int_a^b dx\ f(x)\,,$$

denn das Infimum der Folge der Integrale über ψ_k ist Oberintegral zu f. □

21 Der Kleine Satz von Beppo Levi

Wir geben ein erstes Approximationsverfahren für das Lebesgue-Integral an:

Satz 21.1 (Kleiner Satz von Beppo Levi) *Zu $f : \mathbb{R}^n \to \mathbb{R}^n \cup \{\infty\}$ gebe es eine monoton wachsende oder monoton fallende Folge $(\phi_k)_{k\in\mathbb{N}}$ von Treppenfunktionen, so daß (ϕ_k) punktweise gegen f konvergiert und die Folge der Integrale $\int_{\mathbb{R}^n} dx\ \phi_k(x)$ beschränkt ist. Dann ist f integrierbar, und es gilt $\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ \phi_k(x)$.*

Beweis. Sei $(\phi_k)_{k\in\mathbb{N}}$ monoton wachsend. Wegen der punktweisen Konvergenz gilt $(f - \phi_k)(x) = \sum_{i=k}^{\infty} (\phi_{i+1} - \phi_i)(x)$. Nach Satz 19.7.iv) und Satz 19.9 gilt

$$\|f - \phi_k\|_1 = \Big\| \sum_{i=k}^{\infty} (\phi_{i+1} - \phi_i) \Big\|_1 \le \sum_{i=k}^{\infty} \|\phi_{i+1} - \phi_i\|_1 = \sum_{i=k}^{\infty} \int_{\mathbb{R}^n} dx\ (\phi_{i+1} - \phi_i)(x)\,.$$

Da die Folge der Integrale $\Big(\int_{\mathbb{R}^n} dx\ \phi_i(x)\Big)_{i\in\mathbb{N}}$ monoton und beschränkt ist, konvergiert sie gegen einen Grenzwert F, d.h. es gilt $\|f-\phi_k\|_1 \le F - \int_{\mathbb{R}^n} dx\ \phi_k(x)$. Damit ist f integrierbar mit $\int_{\mathbb{R}^n} dx\ f(x) = F$. □

Wir wenden dieses Approximationsverfahren auf uneigentliche Riemann-Integrale von Regelfunktionen an. Das waren nach Definition 20.1 aus dem 1. Semester Funktionen, die sich gleichmäßig durch Treppenfunktionen approximieren lassen, d.h. zu einer Regelfunktion $f : [a,b] \to \mathbb{R}$ gibt es zu jedem $\epsilon > 0$ Treppenfunktionen ϕ, ψ mit $f(x) - \epsilon \le \phi(x) \le f(x)$ und $f(x) \le \psi(x) \le f(x) + \epsilon$. Regelfunktionen sind Riemann-integrierbar, und für sie gilt der Hauptsatz der Differential- und Integralrechnung. Jede stetige Funktion ist Regelfunktion.

Satz 21.2 *Eine Regelfunktion $f :]a,b[\to \mathbb{R}$, wobei $a = \infty$ und/oder $b = \infty$ zugelassen ist, ist genau dann über $]a,b[$ Lebesgue-integrierbar, wenn $|f|$ uneigentlich Riemann-integrierbar ist, d.h. für Folgen $[a_k, b_k] \subset]a,b[$ kompakter Intervalle mit $\lim_{k\to\infty} a_k = a$ und $\lim_{k\to\infty} b_k = b$ gilt $\lim\limits_{k\to\infty} \int_{a_k}^{b_k} dx\ |f(x)| < \infty$. In diesem Fall gilt $\int_{]a,b[} dx\ f(x) = \int_a^b dx\, f(x)$.*

Beweis. Wir wählen kompakte Intervalle $I_k = [a_k, b_k] \subset]a,b[$ mit $I_k \subset I_{k+1}$ und $\bigcup_{k=0}^\infty I_k =]a,b[$.

($\Rightarrow$) Sei f über $]a,b[$ Lebesgue-integrierbar, dann sind auch f^+ und f^- Lebesgue-integrierbar, und wegen $f^\pm_{[a_k,b_k]} \le f^\pm_{]a,b]}$ gilt mit Satz 20.4

$$\int_{a_k}^{b_k} dx\ f^\pm(x) = \int_{[a_k,b_k]} dx\ f^\pm(x) \le \int_{]a,b[} dx\ f^\pm(x)\ .$$

Damit existiert $\int_a^b dx\ f^\pm(x) := \lim\limits_{k\to\infty} \int_{a_k}^{b_k} dx\ f^\pm(x)$ und ist durch $\int_{]a,b[} dx\ f^\pm(x)$ beschränkt.

($\Leftarrow$) Sei $|f| :]a,b[\to \mathbb{R}$ uneigentlich Riemann-integrierbar. Nach dem Majorantenkriterium Satz 21.7 aus dem 1. Semester ist dann auch $f :]a,b[\to \mathbb{R}$ uneigentlich Riemann-integrierbar, und somit auch $f^\pm = \frac{1}{2}(|f| \pm f)$. Da auch $f^\pm$ Regelfunktionen sind, gibt es eine monoton wachsende Folge von Treppenfunktionen $\phi_k^\pm :]a,b[\to \mathbb{R}_+$, die auf $]a,b[$ punktweise gegen $f^\pm$ konvergiert. Die Folge der Integrale $\int_a^b dx\ \phi_k^\pm(x)$ ist nach dem Majorantenkriterium beschränkt durch $\int_a^b dx\ f^\pm(x)$. Nach dem kleinen Satz von Beppo Levi ist dann auch die Grenzfunktion $f^\pm_{]a,b[} = \lim_{k\to\infty} \phi_k^\pm$ Lebesgue-integrierbar mit

$$\int_{]a,b[} dx\ f^\pm(x) = \lim_{k\to\infty} \int_a^b dx\ \phi_k^\pm(x) \le \int_a^b dx\ f^\pm(x)\ .$$

Zusammen mit der umgekehrten Abschätzung aus ($\Rightarrow$) folgt dann $\int_{]a,b[} dx\ f^{\pm}(x) = \int_a^b dx\ f^{\pm}(x)$. $\square$

Die absolute Konvergenz des Riemann-Integrals kann nicht auf einfache Konvergenz reduziert werden. Z.B. gilt für das uneigentliche Riemann-Integral $\int_{-\infty}^{\infty} dx\ \frac{\sin x}{x} = \pi$, aber $\frac{\sin x}{x}$ ist nicht Lebesgue-integrierbar über $\mathbb{R}$. Denn mit $\frac{\sin x}{x}$ wäre auch $\left|\frac{\sin x}{x}\right|$ Lebesgue-integrierbar, was nicht der Fall ist.

Wir untersuchen nun die Integrierbarkeit von stetigen Funktionen über offenen und kompakten Teilmengen. Im eindimensionalen werden stetige Funktionen als Regelfunktionen beliebig genau durch Treppenfunktionen approximiert. Das gilt auch im $\mathbb{R}^n$:

Lemma 21.3 *Sei $U \subset \mathbb{R}^n$ offen und $K \subset U$ kompakt. Sei $f : K \to \mathbb{R}$ stetig mit $f \geq 0$. Dann gibt es zu jedem $\epsilon > 0$ Treppenfunktionen $\phi_\epsilon, \psi_\epsilon \geq 0$ mit*

$$f(x) - \epsilon \leq \phi_\epsilon(x) \leq f(x) \leq \psi_\epsilon(x) \leq f(x) + \epsilon \qquad \text{für alle } x \in K$$

und $\phi_\epsilon(x) = \psi_\epsilon(x) = 0$ für alle $x \in \mathbb{R} \setminus U$.

Beweis. Nach Satz 3.4 ist $f : K \to \mathbb{R}$ gleichmäßig stetig, d.h. für jedes $\epsilon > 0$ gibt es ein $\delta > 0$, so daß für die Supremumsnorm gilt

$$\|f(x) - f(x')\| < \epsilon \qquad \text{für alle } x, x' \in K \text{ mit } \|x - x'\| < \delta\ .$$

Für jeden Punkt $x \in U$ werde ein abgeschlossener Würfel $W_x \subset U$ mit Mittelpunkt x und Kantenlänge $0 < l \leq \delta$ gewählt. Sei $W_i^o := W_i \setminus \partial W_i$ das (offene) Innere eines solchen Würfels. Dann ist $K \subset \bigcup_{x \in U} W_i^o$. Da K kompakt ist, gibt es endlich viele Würfel $W_1^0, \ldots, W_p^0$, die K überdecken, und insbesondere ist $K \subset W_1 \cup \cdots \cup W_p \subset U$. Nun ist auch $W_i \cap K$ kompakt, so daß f auf $W_i \cap K$ ein Maximum M_i und Minimum m_i annimmt. Dann sind

$$\psi_k := \max(M_1 \mathbf{1}_{W_1}, \ldots M_p \mathbf{1}_{W_p})\ , \qquad \phi_\epsilon := \min(m_1 \mathbf{1}_{W_1}, \ldots m_p \mathbf{1}_{W_p})$$

Treppenfunktionen mit den geforderten Eigenschaften. $\square$

Satz 21.4 *Sei $U \subset \mathbb{R}^n$ eine beschränkte offene Teilmenge und $K \subset \mathbb{R}^n$ eine kompakte Teilmenge. Dann ist jede beschränkte stetige Funktion $f : U \to \mathbb{R}$ über U integrierbar und jede stetige Funktion $g : K \to \mathbb{R}$ über K integrierbar.*

Beweis. Wegen $f = f^+ - f^-$ mit $f^+(x) := \max(f(x), 0) \geq 0$ und $f^-(x) := \max(-f(x), 0) \geq 0$ genügt es, die Integrierbarkeit von nichtnegativen Funktion f zu zeigen.

i) Jede offene Teilmenge $U \subset \mathbb{R}^n$ ist Vereinigung abzählbar vieler kompakter Teilmengen $K_j \subset U$, d.h. $U = \bigcup_{j=0}^{\infty} K_j$. (Siehe Forster, Analysis 3, §7, Hilfssatz 1). Die Einschränkung von f auf K_j wird nach Lemma 21.3 von unten gleichmäßig durch eine Treppenfunktion ϕ_j zu $\epsilon_i = \frac{1}{2^j}$ approximiert. Dann ist $(\phi_j)_{j\in\mathbb{N}}$ eine monoton wachsende Folge von Treppenfunktionen, die punktweise gegen f konvergiert. Da U beschränkt ist, gibt es einen abgeschlossenen Quader $Q \subset \mathbb{R}^n$ mit $U \subset Q$. Da f beschränkt ist, gibt es ein $M \in \mathbb{R}$ mit $0 \le \phi_j(x) \le f_U(x) \le M\mathbf{1}_Q(x)$ für alle $x \in \mathbb{R}^n$. Also gilt

$$\int_{\mathbb{R}^n} dx\ \phi_j(x) \le M v_Q < \infty\ ,$$

so daß f_U nach Satz 21.1 von Beppo Levi integrierbar ist.

ii) Umgekehrt ist jede kompakte Menge K abzählbarer Durchschnitt von offenen Teilmengen $U_j \supset K$, d.h. $K = \bigcap_{j=0}^{\infty} U_j$. Damit wird f_K nach Lemma 21.3 von oben durch eine monoton fallende Folge ψ_j von Treppenfunktionen zu $\epsilon_j = \frac{1}{2^j}$, die außerhalb U_j verschschwinden, punktweise approximiert. Die Folge der Integrale $\Big(\int_{\mathbb{R}^n} dx\ \psi_j(x)\Big)_{j\in\mathbb{N}}$ ist nach unten durch 0 beschränkt, so daß f_K nach Satz 21.1 integrierbar ist. □

22 Der kleine Satz von Fubini

Es geht nun darum, die Integration einer Funktion $f : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ auf getrennte Integrationen über $\mathbb{R}^p$ und $\mathbb{R}^{n-p}$ zurückzuführen mit $0 < p < n$. Letztendlich genügt es also, eindimensionale Integrale ausrechnen zu können, und diese reduzieren sich unter geeigneten Bedingungen (Satz 20.4 und Satz 21.2) auf die bekannten Riemann-Integrale. Zunächst diskutieren wir Treppenfunktionen.

Satz 22.1 *Es sei $\phi = \sum_{i=1}^{k} c_i \mathbf{1}_{Q_i} : \mathbb{R}^p \times \mathbb{R}^{n-p} \to \mathbb{R}$ eine Treppenfunktion, geschrieben $\phi(x,y)$ für $x \in \mathbb{R}^p$ und $y \in \mathbb{R}^{n-p}$. Dann ist auch $\Phi : \mathbb{R}^{n-p} \to \mathbb{R}$ mit $\Phi(y) := \int_{\mathbb{R}^p} dx\ \phi(x,y)$ eine Treppenfunktion, und es gilt*

$$\int_{\mathbb{R}^p\times\mathbb{R}^{n-p}} d(x,y)\ \phi(x,y) = \int_{\mathbb{R}^{n-p}} dy \Big(\int_{\mathbb{R}^p} dx\ \phi(x,y)\Big)\ .$$

Beweis. Die entsprechende Zerlegung der Quader sei $Q_i = Q_i' \times Q_i''$ mit $Q_i' \subset \mathbb{R}^p$ und $Q_i'' \subset \mathbb{R}^{n-p}$. Dann gilt $v(Q_i) = v(Q_i') \cdot v(Q_i'')$. Für festes $y \in \mathbb{R}^{n-p}$ definieren wir

$$\phi_y(x) := \sum_{i=1}^{k} \big(c_i \mathbf{1}_{Q_i''}(y)\big)\mathbf{1}_{Q_i'}(x)$$

$$\Rightarrow \quad \Phi(y) := \int_{\mathbb{R}^p} dx\ \phi_y(x) = \sum_{i=1}^{k} \big(c_i \mathbf{1}_{Q_i''}(y)\big) v(Q_i') = \sum_{i=1}^{k} \big(c_i v(Q_i')\big)\mathbf{1}_{Q_i''}(y)\ .$$

Also ist $\Phi = \sum_{i=1}^k \big(c_i v(Q_i')\big)\mathbf{1}_{Q_i''} : \mathbb{R}^{n-p} \to \mathbb{R}$ eine Treppenfunktion. Ihr Integral ist

$$\int_{\mathbb{R}^{n-p}} dy\ \Phi(y) = \sum_{i=1}^k \big(c_i v(Q_i')\big) v(Q_i'') = \sum_{i=1}^k c_i v(Q_i) = \int_{\mathbb{R}^p \times \mathbb{R}^{n-p}} d(x,y)\ \phi(x,y)\ .$$

□

Offenbar gilt für diese Situation

$$\int_{\mathbb{R}^p \times \mathbb{R}^{n-p}} d(x,y)\ \phi(x,y) = \int_{\mathbb{R}^{n-p}} dy \Big(\int_{\mathbb{R}^p} dx\ \phi(x,y) \Big) = \int_{\mathbb{R}^p} dx \Big(\int_{\mathbb{R}^{n-p}} dy\ \phi(x,y) \Big)\ .$$

Satz 22.2 (Fubini) *Es sei $A \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$ eine kompakte Teilmenge oder eine beschränkte offene Teilmenge. Für festes $y \in \mathbb{R}^{n-p}$ sei $A_y := \{x \in \mathbb{R}^p \ :\ (x,y) \in A\} \subset \mathbb{R}^p$ und für festes $x \in \mathbb{R}^p$ sei $A_x := \{y \in \mathbb{R}^{n-p} \ :\ (x,y) \in A\} \subset \mathbb{R}^{n-p}$. Eine Funktion $f : A \to \mathbb{R}$ sei stetig und beschränkt. Dann gilt:*

i) *Ist $A_y \neq \varnothing$, dann ist die durch $f_y(x) := f(x,y)$ definierte Funktion $f_y : A_y \to \mathbb{R}$ über A_y integrierbar, und die durch*

$$F(y) := \begin{cases} \displaystyle\int_{A_y} dx\ f(x,y) & \text{für } A_y \neq \varnothing \\ 0 & \text{für } A_y = \varnothing \end{cases}$$

definierte Funktion $F : \mathbb{R}^{n-p} \to \mathbb{R}$ ist über $\mathbb{R}^{n-p}$ integrierbar.

ii) *Ist $A_x \neq \varnothing$, dann ist die durch $f_x(y) := f(x,y)$ definierte Funktion $f_x : A_x \to \mathbb{R}$ über A_x integrierbar, und die durch*

$$G(x) := \begin{cases} \displaystyle\int_{A_x} dy\ f(x,y) & \text{für } A_x \neq \varnothing \\ 0 & \text{für } A_x = \varnothing \end{cases}$$

definierte Funktion $G : \mathbb{R}^p \to \mathbb{R}$ ist über $\mathbb{R}^p$ integrierbar.

iii) *Es gilt*

$$\int_A d(x,y)\ f(x,y) = \int_{\mathbb{R}^{n-p}} dy\ F(y) = \int_{\mathbb{R}^n} dx\ G(x)\ .$$

Beweis. Wir beweisen den Satz für A offen und beschränkt und können wie üblich $f \geq 0$ annehmen. Nach Lemma 21.3 gibt es eine monoton wachsende Folge $(\phi_k)_{k\in\mathbb{N}}$ von nichtnegativen Treppenfunktionen $\phi_k : \mathbb{R}^p \times \mathbb{R}^{n-p} \to \mathbb{R}$, die gleichmäßig gegen f_A konvergiert. Dann ist für festgehaltenes y die durch $\phi_{ky} := \phi_k(x,y)$ definierte Folge von nichtnegativen Treppenfunktionen $\phi_{ky} : \mathbb{R}^p \to \mathbb{R}$ monoton wachsend und gegen f_y konvergent. Durch Einbettung in einen genügend großen abgeschlossenen Quader folgt analog zum Beweis von Satz 21.4.i), daß die Folge

der Integrale $\Phi_k(y) := \int_{\mathbb{R}^p} dx\ \phi_{yk}(x)$ auch beschränkt ist. Nach Satz 21.1 gilt somit

$$F(y) = \lim_{k\to\infty} \Phi_k(y) = \lim_{k\to\infty} \int_{\mathbb{R}^p} dx\ \phi_k(x,y)\ .$$

Die Integrale $\Phi_k(y) : \mathbb{R}^{n-p} \to \mathbb{R}$ sind selbst wieder nichtnegative Treppenfunktionen, die monoton wachsend gegen $F(y)$ konvergieren. Wie zuvor ist die Folge der Integrale $\int_{\mathbb{R}^{n-p}} dy\ \Phi_k(y)$ beschränkt, so daß nach Satz 21.1 gilt

$$\int_{\mathbb{R}^{n-p}} dy\ F(y) = \lim_{k\to\infty} \int_{\mathbb{R}^{n-p}} dy\ \Phi_k(y) = \lim_{k\to\infty} \int_{\mathbb{R}^{n-p}} dy\ \Big(\int_{\mathbb{R}^p} dx\ \phi_k(x,y) \Big)\ .$$

Mit Satz 22.1 und $\lim_{k\to\infty} \int_{\mathbb{R}^n} d(x,y)\ \phi_k(x,y) = \int_{\mathbb{R}^n} d(x,y)\ f(x,y)$ folgt die Behauptung.

Für A kompakt ist f durch eine monoton fallende Folge von nichtnegativen Treppenfunktionen zu approximieren, ansonsten ist der Beweis identisch. □

Auf diese Weise können wir die Integration von stetigen und beschränkten Funktionen $f : A \to \mathbb{R}$ über beschränkte offene oder kompakte Teilmengen $A \subset \mathbb{R}^n$ schrittweise auf eindimensionale Integrationen zurückführen ($p = 1$). Für $p = 1$ sei $B := \{y \in \mathbb{R}^{n-1} \ : \ A_y \neq \varnothing\} \subset \mathbb{R}^{n-1}$ und (im einfachsten Fall) $A_y = \bigcup_{k=1}^N [x_{1k}(y), x_{2k}(y)] \subset \mathbb{R}$. Dann gilt

$$\int_A d(x,y)\ f(x,y) = \int_B dy \Big(\sum_{k=1}^N \int_{x_{1k}(y)}^{x_{2k}(y)} dx\ f(x,y) \Big)\ .$$

Wenn A_y aus offenen Intervallen besteht, ist Satz 21.2 zu beachten.

Insbesondere gilt für die Integration stetiger Funktionen über das Rechteck $[a,b] \times [c,d] \subset \mathbb{R}^2$

$$\int_{[a,b]\times[c,d]} d(x,y)\ f(x,y) = \int_c^d dy \Big(\int_a^b dx\ f(x,y) \Big) = \int_a^b dx \Big(\int_c^d dy\ f(x,y) \Big)\ .$$

Beispiel 22.3 Es sei $A = \{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \leq r\}$ der abgeschlossene Vollkreis vom Radius r sowie $B = [-r, r]$ und $A_y = [-\sqrt{r^2 - y^2}, \sqrt{r^2 - y^2}]$. Somit gilt für eine stetige Funktion $f : A \to \mathbb{R}$

$$\int_A d(x,y)\ f(x,y) = \int_{-r}^r dy\ \Big(\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} dx\ f(x,y) \Big)\ .$$

Insbesondere ist das Volumen (also der Flächeninhalt) von A gegeben durch

$$\begin{aligned} v(A) &= \int_A d(x,y)\ 1 = \int_{-r}^r dy\ \Big(\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} dx\ 1 \Big) = \int_{-r}^r dy\ 2\sqrt{r^2 - y^2} \\ &= 4 \int_0^r dy\ \sqrt{r^2 - y^2} \overset{y = r\cos t}{=} 4r^2 \int_0^{\frac{\pi}{2}} dy\ \sin^2 t \end{aligned}$$

Die Stammfunktion zu $f(t) = \sin^2 t$ ist $F(t) = \frac{1}{2}(t - \cos t \sin t)$. Damit ist $\int_0^{\frac{\pi}{2}} dy\ \sin^2 t = \frac{\pi}{4}$ und $v(A) = \pi r^2$. ◁

Satz 22.4 (Volumen der n-dimensionalen Vollkugel) *Das Volumen $\kappa_n(r)$ der n-dimensionalen Vollkugel $A = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \ : \ x_1^2 + \cdots + x_n^2 \leq r^2\}$ vom Radius r ist*

$$\kappa_n(r) = \frac{\pi^{\frac{n}{2}} r^n}{\Gamma(\frac{n+2}{2})} = \begin{cases} \dfrac{\pi^{\frac{n}{2}}}{\frac{n}{2}!}\, r^n & \text{für } n \text{ gerade} \\ \dfrac{2^n \pi^{\frac{n-1}{2}} \frac{n-1}{2}!}{n!}\, r^n & \text{für } n \text{ ungerade} \end{cases}$$

Beweis. Durch Induktion nach n. Die Aussage gilt offenbar für $n = 1$ mit $\kappa_1 = 2r$ und $n = 2$ mit $\kappa_2 = \pi r^2$.

i) Sei n gerade und die Behauptung bewiesen bis $n-1$. Dann ist $A_{x_n} = \{(x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1} \ : \ x_1^2 + \cdots + x_{n-1}^2 \leq r^2 - x_n^2\}$ und somit

$$\begin{aligned} \kappa_n(r) &= \int_{-r}^{r} dx_n \int_{A_{x_n}} dy\ 1 = \int_{-r}^{r} dx_n \kappa_{n-1}(\sqrt{r^2 - x_n^2}) = 2\int_0^r dx_n \kappa_{n-1}(\sqrt{r^2 - x_n^2}) \\ &= \frac{2^n \pi^{\frac{n-2}{2}} \frac{n-2}{2}!}{(n-1)!} \int_0^r dx_n\ (r^2 - x_n^2)^{\frac{n-1}{2}} \\ &\overset{x_n = r\cos t}{=} \frac{2^n \pi^{\frac{n-2}{2}} \frac{n-2}{2}!}{(n-1)!} \cdot r^n \int_0^{\frac{\pi}{2}} dt\ \sin^n t\ . \end{aligned}$$

Nach Beispiel 20.8 aus dem 1. Semster ist

$$\int_0^{\frac{\pi}{2}} dt\ \sin^{2k} t = \frac{(2k-1)(2k-3)\cdots 1}{2k(2k-2)\cdots 2} \cdot \frac{\pi}{2} = \frac{(2k-1)!\pi}{2^{2k} k!(k-1)!}\ ,$$

und $2k \mapsto n$ liefert die Behauptung.

ii) Sei n ungerade und die Behauptung bewisen bis $n-1$. Dann ist

$$\begin{aligned} \kappa_n(r) &= \int_{-r}^{r} dx_n \int_{A_{x_n}} dy\ 1 = \int_{-r}^{r} dx_n \kappa_{n-1}(\sqrt{r^2 - x_n^2}) = 2\int_0^r dx_n \kappa_{n-1}(\sqrt{r^2 - x_n^2}) \\ &= \frac{2\pi^{\frac{n-1}{2}}}{\frac{n-1}{2}!} \int_0^r dx_n\ (r^2 - x_n^2)^{\frac{n-1}{2}} \overset{x_n = r\cos t}{=} \frac{2\pi^{\frac{n-1}{2}}}{\frac{n-1}{2}!} r^n \int_0^{\frac{\pi}{2}} dt\ \sin^n t\ . \end{aligned}$$

Nach Beispiel 20.8 aus dem 1. Semster ist

$$\int_0^{\frac{\pi}{2}} dt\ \sin^{2k+1} t = \frac{2k(2k-2)\cdots 2}{(2k+1)(2k-1)\cdots 3 \cdot 1} = \frac{2^{2k} k!k!}{(2k+1)!}\ ,$$

und $2k+1 \mapsto n$ liefert die Behauptung. □

23 Nullmengen und Satz von Riesz-Fischer

Definition 23.1 Eine Menge $A \subset \mathbb{R}^n$ heißt *(Lebesgue-) meßbar*, wenn die Funktion 1 über A integrierbar ist. In diesem Fall heißt

$$v_n(A) := \int_A dx\ 1(x) = \int_A dx = \int_{\mathbb{R}^n} dx\ \mathbf{1}_A(x)$$

das *n-dimensionale Volumen* bzw. das *Lebesgue-Maß* von A. Die leere Menge hat das Lebesgue-Maß Null.

Aus Satz 21.4 folgt, daß jede kompakte Teilmenge $K \subset \mathbb{R}^n$ und jede offene und beschränkte Teilmenge $U \subset \mathbb{R}^n$ meßbar ist.

Für Quader $Q \subset \mathbb{R}^n$ ist offenbar $v_n(Q) = v(Q)$ das zu Beginn eingeführte Volumen. Für den Vollkreis $A = \{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \le r\}$ gilt $v_2(A) = \pi r^2$ nach Beispiel 22.3.

Satz 23.2 *Sind $A, B \subset \mathbb{R}^n$ meßbar, so gilt*

i) *$A \cup B$ und $A \cap B$ sind meßbar, und es gilt $v_n(A \cup B) = v_n(A) + v_n(B) - v_n(A \cap B)$.*

ii) *Aus $A \subset B$ folgt $v_n(A) \le v_n(B)$.*

Ist eine Funktion f über $A \subset \mathbb{R}^n$ integrierbar und ist $B \subset \mathbb{R}^n$ meßbar, dann ist f auch über $A \cap B$ integrierbar, und es gilt

$$\Big| \int_{A\cap B} dx\ f(x) \Big| \le \int_A dx\ \big|f(x)\big| \ .$$

Beweis. i) folgt aus $\mathbf{1}_{A\cap B} = \mathbf{1}_A \cdot \mathbf{1}_B$ und $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \cdot \mathbf{1}_B$ und ii) aus $\mathbf{1}_A \le \mathbf{1}_B$.

Die letzte Behauptung folgt aus $f_{A\cap B} = f_A \cdot \mathbf{1}_B$. Da f_A und $\mathbf{1}_B$ integrierbar sind und $\mathbf{1}_B$ beschränkt ist, ist $f_{A\cap B}$ integrierbar nach Satz 20.3.iii). Die Abschätzung folgt aus $|f_{A\cap B}| \le |f_A|$. □

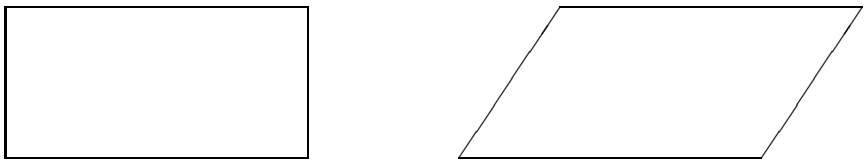
Der kleine Satz von Fubini ermöglicht wieder (da $\mathbf{1}_A$ stetig) die schrittweise Berechnung von Volumina. Ist $A \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$ und $A_y := \{x \in \mathbb{R}^p \ : \ (x,y) \in A\} \subset \mathbb{R}^p$ die in Satz 22.2 eingeführte Schnittmenge, dann gilt

$$v_n(A) = \int_{\mathbb{R}^{n-p}} dy\ v_p(A_y) \ .$$

Insbesondere folgt

Satz 23.3 (Prinzip von Cavalieri) *Seien $A, B \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$ zwei kompakte Mengen, und es gelte $v_p(A_y) = v_p(B_y)$ für alle $y \in \mathbb{R}^{n-p}$. Dann gilt $v_n(A) = v_n(B)$.* □

Zum Beispiel haben ein Rechteck und ein Parallelogramm mit gleicher Grundlänge und gleicher Höhe das gleiche zweidimensionale Volumen (Flächeninhalt):



Wir haben bisher vor allem stetige Funktionen integriert. Wir zeigen nun, daß Singularitäten, die nur auf eine Teilmenge vom Maß Null auftreten, keine Rolle spielen.

Definition 23.4 Eine Menge $N \subset \mathbb{R}^n$ heißt *(Lebesgue-) Nullmenge*, wenn sie eine der folgenden äquivalenten Eigenschaften hat:

i) N ist meßbar mit $v_n(N) = 0$.

ii) Die charakteristische Funktion von N hat die L^1-Halbnorm Null: $\|\mathbf{1}_N\|_1 = 0$.

Beweis: i)⇒ii) $0 = v_n(N) = \int_{\mathbb{R}^n} dx\ \mathbf{1}_N(x) = \int_{\mathbb{R}^n} dx\ |\mathbf{1}_N(x)| = \|\mathbf{1}_N\|_1$
(⇐) Für die Treppenfunktionen $\phi_k = 0$ gilt $\|\mathbf{1}_N - \phi_k\|_1 = 0$ und damit $v_n(N) = \int_{\mathbb{R}^n} dx\ \mathbf{1}_N(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ \phi_k(x) = 0$.

Satz 23.5 i) *Jede Teilmenge einer Nullmenge ist eine Nullmenge.*

ii) *Die Vereinigung abzählbar vieler Nullmengen ist eine Nullmenge.*

Beweis. i) $M \subset N \quad \Rightarrow \quad \mathbf{1}_M \le \mathbf{1}_N \quad \Rightarrow \quad \|\mathbf{1}_M\|_1 \le \|\mathbf{1}_N\| = 0$.

ii) Ist $N = \bigcup_{k=1}^{\infty} N_k$ mit $v_n(N_k) = 0$, dann ist $0 \le \mathbf{1}_N(x) \le \sum_{k=1}^{\infty} \mathbf{1}_{N_k}(x)$ und somit $\|\mathbf{1}_N\|_1 \le \sum_{k=1}^{\infty} \|\mathbf{1}_{N_k}\|_1 = 0$. □

Zum Beispiel ist die Menge der rationalen Zahlen $\mathbb{Q} \subset \mathbb{R}$ eine Nullmenge, so daß $\mathbf{1}_{\mathbb{Q}}$ über $\mathbb{R}$ integrierbar ist mit $\int_{\mathbb{R}} dx\ \mathbf{1}_{\mathbb{Q}}(x) = 0$.

Satz 23.6 *Es sei $A \subset \mathbb{R}^{n-1}$ eine offene oder abgeschlossene Teilmenge und $\Gamma = \{(x, g(x)) \ :\ x \in A\ ,\ g : A \to \mathbb{R}$ stetig$\} \subset \mathbb{R}^n$ der Graph einer stetigen Funktion. Dann ist Γ eine Nullmenge.*

Beweis. Jede offene oder abgeschlossene Menge ist Vereinigung abzählbar vieler kompakter Teilmengen, so daß es genügt, den Satz für A kompakt zu beweisen. Das Bild einer kompakten Menge unter einer stetigen Funktion ist ein kompaktes Intervall. Also ist Γ kompakt, und das Volumen ist nach dem Satz von Fubini

$$v_n(\Gamma) = \int_A dx \int_{g(x)}^{g(x)} dy\ 1(x) = 0.$$ □

Definition 23.7 Es sei E eine Eigenschaft, die einer Teilmenge $A \subset \mathbb{R}^n$ zukommt. Wir sagen, die Eigenschaft E gilt *fast überall* auf A, wenn die Menge der Punkte, für die E nicht gilt, eine Nullmenge ist.

Satz 23.8 *Eine Funktion $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ mit $\|f\|_1 < \infty$ ist fast überall endlich, d.h. $N := \{x \in \mathbb{R}^n \ : \ f(x) = \infty\}$ ist eine Nullmenge. Insbesondere ist jede integrierbare Funktion fast überall endlich.*

Beweis. Für alle $\epsilon > 0$ gilt $\mathbf{1}_N \leq \epsilon |f|$, also $\|\mathbf{1}_N\|_1 \leq \epsilon \|f\|_1$ und damit $\|\mathbf{1}_N\|_1 = 0$, da $\|f\|_1 < \infty$. □

Satz 23.9 (Modifikationssatz) *Seien $f, g : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ Funktionen, die fast überall gleich sind, und f sei integrierbar. Dann ist auch g integrierbar, und es gilt $\int_{\mathbb{R}^n} dx \ f(x) = \int_{\mathbb{R}^n} dx \ g(x)$. Insbesondere gibt es zu jeder integrierbaren Funktion f eine integrierbare Funktion g mit gleichem Lebesgue-Integral, die fast überall mit f übereinstimmt und nur Werte $\neq \infty$ annimmt.*

Beweis. Zu f gibt es eine Folge $\{\phi_k\}$ von Treppenfunktionen mit $\lim_{k\to\infty} \|f - \phi_k\|_1 = 0$. Es sei $N := \{x \in \mathbb{R}^n \ : \ f(x) \neq g(x)\}$. Dann gilt $|g - f| \leq \sum_{j=1}^{\infty} \mathbf{1}_N$ und somit

$$0 \leq \|g - \phi_k\|_1 \leq \|g - f\|_1 + \|f - \phi_k\|_1 \leq \Big(\sum_{j=1}^{\infty} \|\mathbf{1}_N\|_1 \Big) + \|f - \phi_k\|_1 = \|f - \phi_k\|_1 \ .$$

Also gilt $\lim_{k\to\infty} \|g - \phi_k\|_1 = 0$, d.h. g ist integrierbar, mit $\int_{\mathbb{R}^n} dx \ g(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx \ \phi_k(x) = \int_{\mathbb{R}^n} dx \ f(x)$. □

Der Modifikationssatz ist auch hilfreich bei Zusammensetzungen von Gebieten: Sei f über A und über B integrierbar und sei $A \cap B$ eine Nullmenge, dann ist f auch über $A \cup B$ integrierbar, da $f_{A\cup B}$ fast überall mit der Funktion $f_A + f_B$ übereinstimmt. Also gilt $\int_{A\cup B} dx \ f(x) = \int_A dx \ f(x) + \int_B dx \ f(x)$. Insbesondere können wir Funktionen auf einer beschränkten offenen oder kompakten Teilmenge, für die die Menge der Unstetigkeitsstellen eine Nullmenge ist, integrieren. Das betrifft z.B. stückweise stetige und beschränkte Funktionen.

Satz 23.10 *Für eine Funktion $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ gilt $\|f\|_1 = 0$ genau dann, wenn f fast überall gleich Null ist. Insbesondere verschwindet jede nichtnegative integrierbare Funktion f mit $\int_{\mathbb{R}^n} dx \ f(x) = 0$ fast überall.*

Beweis. Die Richtung ($\Leftarrow$) ist klar. Sei also $\|f\|_1 = 0$. Wir betrachten $N := \{x \in \mathbb{R}^n \ : \ f(x) \neq 0\}$. Dann gilt $N = \bigcup_{k=1}^{\infty} N_k$ mit $N_k := \{x \in \mathbb{R}^n \ : \ |f(x)| \geq \frac{1}{k}\}$. Dann ist $\mathbf{1}_{N_k} \leq k \cdot |f|$ und somit $0 \leq \|\mathbf{1}_{N_k}\|_1 \leq k\|f\|_1 = 0$. Also ist jedes N_k und damit N eine Nullmenge. □

Mit Hilfe der Eigenschaften von Nullmengen zeigen wir jetzt, daß das Lebesgue-Integral in gewisser Weise vollständig ist.

Definition 23.11 Eine Folge $(f_k)_{k\in\mathbb{N}}$ von Funktionen $f_k : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ heißt *L^1-konvergent* gegen eine Funktion $f : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$, wenn $\lim_{k\to\infty} \|f - f_k\|_1 = 0$ gilt. Die Funktion f heißt dann der *L^1-Grenzwert* von $(f_k)_{k\in\mathbb{N}}$.

Die Folge $(f_k)_{k\in\mathbb{N}}$ heißt *L^1-Cauchyfolge*, wenn es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß $\|f_k - f_l\|_1 < \epsilon$ für alle $k, l \geq N$.

Der L^1-Grenzwert kann nicht eindeutig sein, denn aus $\lim_{k\to\infty} \|f - f_k\|_1 = \lim_{k\to\infty} \|g - f_k\|_1 = 0$ folgt $\|f - g\|_1 = 0$, d.h. f und g sind nur fast überall gleich. Wie üblich ist jede L^1-konvergente Folge eine L^1-Cauchyfolge. Für integrierbare Funktionen gilt aber auch die Umkehrung:

Satz 23.12 (Riesz-Fischer) *Es sei $\mathcal{L}^1(\mathbb{R}^n)$ der Vektorraum der über $\mathbb{R}^n$ integrierbaren Funktion. Jede L^1-Cauchyfolge $(f_k)_{k\in\mathbb{N}}$ integrierbarer Funktionen $f_k \in \mathcal{L}^1(\mathbb{R}^n)$ besitzt einen Grenzwert $f \in \mathcal{L}^1(\mathbb{R}^n)$, und es gilt*

i) $\displaystyle\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ f_k(x)$

ii) *Es gibt eine Teilfolge von $(f_k)_{k\in\mathbb{N}}$, die fast überall punktweise gegen f konvergiert.*

Beweis. Es werden Indizes $k_1 < k_2 < \ldots$ so gewählt, daß $\|f_k - f_{k_\nu}\|_1 < \frac{1}{2^\nu}$ für alle $k \geq k_\nu$. Dann gilt $\sum_{\nu=1}^\infty \|f_{k_\nu} - f_{k_{\nu+1}}\|_1 \leq 1$. Abkürzend sei $g_\nu := f_{k_\nu} - f_{k_{\nu+1}}$ und $g = \sum_{\nu=1}^\infty |g_\nu|$. Nach Satz 23.8 ist wegen $\|g\|_1 \leq 1$ die Menge $N = \{x \in \mathbb{R}^n \ : \ g(x) = \infty\}$ eine Nullmenge. Außerdem gibt es eine Nullmenge N_1 mit $f_{k_1}(x) \neq \infty$ für alle $x \notin N_1$. Dann setzen wir

$$f(x) := \begin{cases} \lim\limits_{\nu\to\infty} f_{k_\nu} = f_{k_1} + \sum\limits_{n=1}^\infty g_\nu & \text{für } x \in \mathbb{R}^n \setminus (N \cup N_1) \\ 0 & \text{für } x \in N \cup N_1 \end{cases}$$

Damit ist $f(x) \neq \infty$, und die Teilfolge (f_{k_ν}) konvergiert fast überall gegen f, d.h. ii) ist gezeigt.

Die Teilfolge ist so gewählt, daß es zu jedem $\epsilon > 0$ ein $\rho \in \mathbb{N}$ gibt, so daß $\sum_{\nu=\rho}^\infty \|g_\nu\|_1 < \epsilon$ und $\|f_k - f_{k_\rho}\|_1 < \epsilon$ für alle $k \geq k_\rho$. Da f_{k_ρ} integrierbar ist, gibt es eine Treppenfunktion ϕ mit $\|f_{k_\rho} - \phi\| < \epsilon$. Somit gilt für $k \geq k_\rho$

$$\|f - \phi\|_1 \leq \|f - f_{k_\rho}\|_1 + \|f_{k_\rho} - \phi\|_1 < \Big\| \sum_{\nu=\rho}^\infty g_\nu \Big\|_1 + \epsilon < 2\epsilon\ ,$$

d.h. f ist integrierbar. Für $k \geq k_\rho$ gilt

$$\|f - f_k\|_1 \leq \|f - f_{k_\rho}\|_1 + \|f_{k_\rho} - f_k\|_1 < 2\epsilon\ ,$$

d.h. (f_k) ist L^1-konvergent gegen f. Schließlich gilt

$$\Big| \int_{\mathbb{R}^n} dx\ f(x) - \int_{\mathbb{R}^n} dx\ f_k(x) \Big| \leq \int_{\mathbb{R}^n} dx\ \big| f(x) - f_k(x) \big| = \|f - f_k\|_1 < 2\epsilon\ .$$

also die in i) behauptete Konvergenz der Integrale. □

Eine integrierbare Funktion ist L^1-Grenzwert einer Folge von Treppenfunktionen. Nach dem Satz von Riesz-Fischer kann man erwarten, daß fast überall auch punktweise Konvergenz zu erreichen ist:

Satz 23.13 *Jede integrierbare Funktion $f \in \mathcal{L}^1(\mathbb{R})$ ist L^1-Grenzwert einer Folge $(\phi_k)_{k\in\mathbb{N}}$ von Treppenfunktionen mit*

i) $\displaystyle\sum_{k=0}^{\infty} \|\phi_{k+1} - \phi_k\|_1 < \infty$

ii) (ϕ_k) *konvergiert fast überall punktweise gegen f.*

Beweis. Es gibt eine Folge (ψ_l) von Treppenfunktionen mit $\lim_{l\to\infty} \|f - \psi_l\|_1 = 0$. Nach dem Satz von Riez-Fischer gibt es eine Teilfolge (ϕ_k) mit Eigenschaft i), die fast überall punktweise gegen eine integrierbare Funktion $\tilde{f}$ konvergiert. Nach Konstruktion sind beide L^1-Grenzwerte $f, \tilde{f}$ fast überall gleich. □

Wir wissen, daß die L^1-Halbnorm keine Norm auf dem Vektorraum $\mathcal{L}^1(\mathbb{R}^n)$ aller über $\mathbb{R}^n$ integrierbaren Funktionen ist: aus $\|f\|_1 = 0$ folgt nach Satz 23.10 nur, daß f fast überall Null ist. Es bietet sich deshalb an, fast überall gleiche Funktionen zu *Äquivalenzklassen* zusammenzufassen. Sei dazu $\mathcal{N}(\mathbb{R}^n) = \{f \in \mathcal{L}^1(\mathbb{R}^n) \ : \ \|f\|_1 = 0\}$. Offenbar ist $\mathcal{N}(\mathbb{R}^n)$ ein Untervektorraum von $\mathcal{L}^1(\mathbb{R}^n)$. Zwei Funktionen $f, g \in \mathcal{L}^1(\mathbb{R}^n)$ heißen *äquivalent* ($f \sim g$), wenn $f - g \in \mathcal{N}(\mathbb{R}^n)$, d.h. wenn f und g fast überall gleich sind. Die entsprechende Äquivalenzklasse einer Funktion $f \in \mathcal{L}^1(\mathbb{R}^n)$ wird mit $[f]$ oder $f + \mathcal{N}(\mathbb{R}^n)$ bezeichnet. Dann ist die Menge aller Äquivalenzklassen

$$L^1(\mathbb{R}^n) := \mathcal{L}^1(\mathbb{R}^n)/\mathcal{N}(\mathbb{R}^n) = \{[f] \ : \ f \in \mathcal{L}^1(\mathbb{R}^n)\}$$

ein Untervektorraum von $\mathcal{L}^1(\mathbb{R}^n)$. Dabei ist die Linearkombination von Klassen definiert als Klasse der Linearkombination: $c_1[f_1] + c_2[f_2] := [c_1 f_1 + c_2 f_2]$.

Durch die Äquivalenzklassenbildung wird das Problem mit der Normeigenschaft von $\|\ \|_1$ behoben: Dazu definieren wir $\|[f]\|_1 := \|f\|_1$. Diese Definition ist sinnvoll, denn aus $f \sim g$, also $[f] = [g]$, folgt

$$\|f\|_1 = \|g + (f - g)\|_1 \leq \|g\|_1 + \|f - g\|_1 = \|g\|_1\ ,$$

die Norm ist also unabhängig von der Wahl des Repräsentanten. Insbesondere gilt die Dreiecksungleichung sowie $\|c[f]\|_1 = |c|\|f\|_1$. Schließlich gilt nach Satz 23.10 $\|[f]\|_1 = 0$ genau dann, wenn $[f] = [0]$. Damit ist $\big(L^1(\mathbb{R}^n), \|\ \|_1\big)$ ein normierter Raum, nach dem Satz von Riesz-Fischer sogar ein Banach-Raum.

24 Der Transformationssatz

Der Transformationssatz ist eine mächtige Methode zur Berechnung von Integralen.

Satz 24.1 *Seien $U, V \subset \mathbb{R}^n$ offene Teilmengen und sei $T : U \to V$ ein Diffeomorphismus. Eine Funktion $f : V \to \mathbb{R} \cup \{\infty\}$ ist genau dann über V integrierbar, wenn die Funktion $|\det(DT)| \cdot (f \circ T) : U \to \mathbb{R} \cup \{\infty\}$ über $U = T^{-1}(V)$ integrierbar ist. In diesem Fall gilt*

$$\int_U dx \, |\det(DT)(x)| \, f(T(x)) = \int_V dy \, f(y) \, .$$

Zur Erinnerung: Ein Diffeomorphismus T ist eine differenzierbare bijektive Abbildung mit differenzierbarem Inversen. Das Differential DT ist dann eine lineare Abbildung $DT : \mathbb{R}^n \to \mathbb{R}^n$, so daß die Determinante korrekt definiert ist. Für genügend kleine U sichert der Satz über die inverse Funktion Satz 8.4 die Invertierbarkeit von DT. Der Transformationssatz läßt sich aber auch verwenden, wenn T nur auf einer Nullmenge N kein Diffeomorphismus ist, da Nullmengen im Lebesgue-Integral keine Rolle spielen. In diesem Fall genügt es, über $U \setminus N$ bzw. $T^{-1}(U \setminus N)$ zu integrieren.

Der Beweis des Transformationssatzes erfordert Methoden, die wir erst später bereitstellen. Dieser Abschnitt stellt typische Folgerungen und Anwendungen vor.

- Im $\mathbb{R}^1$ reduziert sich der Transformationssatz auf die Substitutionsregel: Sei $T : [a, b] \to [\alpha, \beta]$ eine bijektive stetig differenzierbare Abbildung, dann gilt $\displaystyle\int_{[a,b]} dx \, |T'(x)| \, f(T(x)) = \int_{[\alpha,\beta]} dy \, f(y)$.
- Im Spezialfall einer nichtausgearteten affinen Transformation $y = T(x) = Ax + b \in \mathbb{R}^n$ mit $\det A \neq 0$ ist $f : K \to \mathbb{R}^n \cup \{\infty\}$ genau dann über $K \subset \mathbb{R}^n$ integrierbar, wenn $f \circ T$ über $T^{-1}(K)$ integrierbar ist, und es gilt
$$\int_{T^{-1}(K)} dx \, f(Ax + b) = \frac{1}{|\det A|} \int_K dy \, f(y) \, .$$
- Beschreibt A eine Rotation oder Spiegelung (dann ist $|\det A| = 1$) und wählen wir für f die konstante Funktion $f = 1$, so folgt, daß K genau dann meßbar ist, wenn $T^{-1}(K)$ meßbar ist, und es gilt $v_n(K) = v_n(T^{-1}(K))$. Volumina bleiben also bei Kombinationen aus Verschiebung, Drehung und Spiegelung erhalten.

24.1 Integrale über Kugelschalen

Sehr häufig treten Integrale über n-dimensionale Kugeln oder Kugelschalen auf. Solche Integrale lassen sich durch eine Transformation T zu Polarkoordinaten

vereinfachen (und mit dem Satz von Fubini oft auch lösen). Polarkoordianten im $\mathbb{R}^n$ bestehen aus dem Radius r und $n-1$ Winkeln $\varphi, \vartheta_1, \ldots, \vartheta_{n-2}$. Dann ist $T : (r, \varphi, \vartheta_1, \ldots, \vartheta_{n-1}) \mapsto (y_1, \ldots, y_n) = T_n(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2})$ definiert durch $T_2(r, \varphi) = (r\cos\varphi, r\sin\varphi)$ und dann rekursiv $y_n = r\cos\vartheta_{n-2}$ und $(y_1, \ldots, y_{n-1}) = T_{n-1}(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-3}) \cdot \sin\vartheta_{n-2}$. Konkret heißt das

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} r\cos\varphi\sin\vartheta_1\cdots\sin\vartheta_{n-2} \\ r\sin\varphi\sin\vartheta_1\cdots\sin\vartheta_{n-2} \\ r\cos\vartheta_1\sin\vartheta_2\cdots\sin\vartheta_{n-2} \\ \vdots \\ r\cos\varphi_{n-3}\sin\vartheta_{n-2} \\ r\cos\vartheta_{n-2} \end{pmatrix} .$$

Damit die Transformation T bijektiv wird, ist (z.B.) $\vartheta_1, \ldots, \vartheta_{n-2} \in]0, \pi[$ und $\varphi \in]0, 2\pi[$ zu wählen.

Satz 24.2 *Es gilt* $|\det(DT)(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2})| = r^{n-1}(\sin\vartheta_1)^1 \cdots (\sin\vartheta_{n-2})^{n-2}$.

Beweis. Das Differential der Transformation ist

$$DT = \begin{pmatrix} \frac{\partial y_1}{\partial r} & \frac{\partial y_1}{\partial \varphi} & \frac{\partial y_1}{\partial \vartheta_1} & \cdots & \frac{\partial y_1}{\partial \vartheta_{n-2}} \\ \frac{\partial y_2}{\partial r} & \frac{\partial y_2}{\partial \varphi} & \frac{\partial y_2}{\partial \vartheta_1} & \cdots & \frac{\partial y_2}{\partial \vartheta_{n-2}} \\ \frac{\partial y_3}{\partial r} & \frac{\partial y_3}{\partial \varphi} & \frac{\partial y_3}{\partial \vartheta_1} & \cdots & \frac{\partial y_3}{\partial \vartheta_{n-2}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial r} & \frac{\partial y_n}{\partial \varphi} & \frac{\partial y_n}{\partial \vartheta_1} & \cdots & \frac{\partial y_n}{\partial \vartheta_{n-2}} \end{pmatrix}$$

Für $n = 2$ ist $\partial_r y = (\cos\varphi, \sin\varphi)$ und $\partial_\varphi y = (-r\sin\varphi, -r\cos\varphi)$, so daß die Determinantenformel gilt. Im Schritt von n auf $n+1$ für $n \geq 2$ haben wir mit $\tilde{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ und $y = (\tilde{y}\sin\vartheta_{n-1}, r\cos\vartheta_{n-1}) \in \mathbb{R}^{n+1}$

$$\partial_r y = \begin{pmatrix} \partial_r\tilde{y}\cdot\sin\vartheta_{n-1} \\ \cos\vartheta_{n-1} \end{pmatrix} , \qquad \partial_\varphi y = \begin{pmatrix} \partial_\varphi\tilde{y}\cdot\sin\vartheta_{n-1} \\ 0 \end{pmatrix} ,$$
$$\partial_{\vartheta_i} y = \begin{pmatrix} \partial_{\vartheta_i}\tilde{y}\cdot\sin\vartheta_{n-1} \\ 0 \end{pmatrix} \text{ für } 1 \leq i \leq n-2 , \qquad \partial_{\vartheta_{n-1}} y = \begin{pmatrix} \tilde{y}\cos\vartheta_{n-1} \\ -r\sin\vartheta_{n-1} \end{pmatrix} .$$

Nach Induktionsannahme gelte die Determinatenformel für $n \geq 2$. Im Schritt von n auf $n+1$ betrachten wir zunächst $\sin\vartheta_{n-1} = 0$. Dann ist gilt $\operatorname{rang}(dT) = 2$ und damit $\det dT = 0$. Sei also $\sin\vartheta_{n-1} \neq 0$. Dann addieren wir das $(-r\frac{\cos\vartheta_{n-1}}{\sin\vartheta_{n-1}})$-fache der ersten Spalte zur letzten. Wegen $r\partial_r\tilde{y} = \tilde{y}$ wird die neue letzte Spalte zu

$$\begin{pmatrix} -r\partial_r\tilde{y}\cdot\cos\vartheta_{n-1} \\ -r\frac{\cos^2\vartheta_{n+1}}{\sin\vartheta_{n-1}} \end{pmatrix} + \begin{pmatrix} \tilde{y}\cos\vartheta_{n-1} \\ -r\sin\vartheta_{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{-r}{\sin\vartheta_{n-1}} \end{pmatrix} .$$

Entwicklung nach der neuen letzten Spalte und Herausziehen des Faktors $\sin\vartheta_{n-1}$ aus jeder der ersten n Spalten der Unterdeterminante bestätigt die Determinatenformel. □

Sei nun $\Pi :=]0, 2\pi[\times(]0,\pi[)^{n-2}$ und $I \subset]0,\infty[$ ein offenes Intervall. Dann ist das Bild von $I \times \Pi$ unter T die offene Teilmenge $K(I) \setminus N \subset \mathbb{R}^n$, wobei $K[I] := \{x \in \mathbb{R}^n \,:\, \|x\| \in I\}$ die offene Kugelschale der Radien im Intervall I ist und N eine Nullmenge, die durch Aufschneiden der Kugel bei $y_2 = 0$ entlang der positiven y_1-Achse erhalten wird. Da für das Lebesgue-Integral Nullmengen (N und Ränder von I) keine Rolle spielen, erhalten wir

Satz 24.3 *Sei $I \subset [0,\infty[$ ein beliebiges Intervall und $K(I) = \{x \in \mathbb{R}^n \,:\, \|x\|_2 \in I\}$ die entsprechende Kugelschale. Eine auf $K(I)$ definierte Funktion f ist genau dann über die Kugelschale $K(I)$ integrierbar, wenn die Funktion $f(T(r,\varphi,\vartheta_1,\dots,\vartheta_{n-2}))\cdot r^{n-1}C(\varphi,\vartheta_1,\dots,\vartheta_{n-2})$ über $I\times\Pi$ integrierbar ist. In diesem Fall gilt (unter Verwendung des später bewiesenen Satzes von Fubini)*

$$\int_{K(I)} dy\; f(y)$$
$$= \int_I dr\; r^{n-1} \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1\; \sin\vartheta_1 \cdots \int_0^{\pi} d\vartheta_{n-2}\; \sin^{n-2}\vartheta_{n-2}\; f(T(r,\varphi,\vartheta_1,\dots,\vartheta_{n-2}))\;.$$

□

Als Beispiel berechnen wir nochmals das Volumen der dreidimensionalen Kugel vom Radius R, d.h. $I = [0,R]$. Die Funktion 1 ist integrierbar, so daß gilt

$$v(K([0,R])) = \int_0^R dr\; r^2 \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1\; \sin\vartheta_1 = \frac{R^3}{3}\cdot 2\pi\cdot\big(-\cos\vartheta_1\big|_0^{\pi}\big) = \frac{4\pi}{3}R^3\;.$$

Wenn in Satz (24.3) die Funktion f nicht von den Winkeln abhängt, also rotationssymmetrisch ist, dann erhalten wir:

Satz 24.4 *Es sei f eine Funktion auf dem Intervall $]a,b[$. Die Funktion $\tilde f$ auf $\mathbb{R}^n$ mit $\tilde f(x) = f(\|x\|)$ ist genau dann über die Kugelschale K_I integrierbar, wenn die Funktion $f(r)r^{n-1}$ über I integrierbar ist. In diesem Fall gilt mit $\kappa_n := \kappa_n(1)$*

$$\int_{K(I)} dx\; f(\|x\|) = n\kappa_n \int_I dr\; r^{n-1}\, f(r)\;.$$

Beweis. Unter Verwendung von Satz 24.3 ist nur zu zeigen, daß das Winkelintegral $\int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1\; \sin\vartheta_1 \cdots \int_0^{\pi} d\vartheta_{n-2}\; \sin^{n-2}\vartheta_{n-2} = n\kappa_n$ liefert. Das folgt aber sofort

für das Volumen der Einheitsvollkugel mit $f = 1$ und $I = [0, 1]$:

$$\begin{aligned}\kappa_n &= \int_{[0,1]} dr\; r^{n-1} \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1\; \sin\vartheta_1 \cdots \int_0^{\pi} d\vartheta_{n-2}\; \sin^{n-2}\vartheta_{n-2} \\ &= \frac{1}{n} \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta_1\; \sin\vartheta_1 \cdots \int_0^{\pi} d\vartheta_{n-2}\; \sin^{n-2}\vartheta_{n-2}\;. \end{aligned}$$ □

24.2 Integration über Teilmengen von $(\mathbb{R}_+)^n$

Häufig treten Integrationen auf, die auf das Standardsimplex

$$\Delta^n := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \;:\; x_i \geq 0\;,\; x_1 + \cdots + x_n \leq 1\} \subset \mathbb{R}^n$$

zurückgeführt werden können, z.B. bei Funktionen auf $(\mathbb{R}_+)^n$, die entscheidend von der Summe $x_1 + \cdots + x_n$ abhängen. In diesem Fall ist eine auf Jacobi zurückgehende Transformation hilfreich. Dazu definiert man

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = J_2(u_1, u_2) := \begin{pmatrix} u_1(1-u_2) \\ u_1 u_2 \end{pmatrix}$$

und dann rekursiv für $\tilde{x} = (x_1, \ldots, x_n)^t$

$$\begin{pmatrix} \tilde{x} \\ x_{n+1} \end{pmatrix} = J_{n+1}(u_1, \ldots, u_n, u_{n+1}) := \begin{pmatrix} J_n(u_1, \ldots, u_n)\,(1-u_{n+1}) \\ u_1 u_{n+1} \end{pmatrix}\;.$$

Wir zeigen, daß J_n einen Diffeomorphismus implementiert zwischen

- $\mathbb{R}_+ \times (]0,1[)^{n-1}$ und $(\mathbb{R}_+)^n$,
- bzw. $(]0,1[)^n$ und $(\Delta^n)^o := \Delta^n \setminus \partial\Delta^n$.

Zunächst zur Bijektivität. Klar ist, daß das Bild Teilmenge von $(\mathbb{R}_+)^n$ ist. Es gilt $x_1 + \cdots + x_n = u_1$ zunächst für $n = 2$ und dann rekursiv für alle n. Sei also $u_1 > 0$. Damit gilt $0 < x_n < u_1$, es gibt also eine bijektive Zuordnung zwischen $u_n \in\,]0,1[$ und $x_n = u_1 u_n$. Sei dann zusätzlich u_n fixiert, dann ist $x_1 + \cdots + x_{n-1} = u_1 - x_n = u_1(1-u_n)$. Insbesondere folgt $0 < x_{n-1} < u_1(1-u_n)$, damit eine bijektive Zuordnung zwischen $u_{n-1} \in\,]0,1[$ und $x_n = u_1 u_{n-1}(1-u_n)$, usw.

Das Differential von J ist

$$\begin{aligned}(DJ_2)(u_1, u_2) &= \begin{pmatrix} 1-u_2 & -u_1 \\ u_2 & u_1 \end{pmatrix}\;, \\ (DJ_{n+1})(\tilde{u}, u_{n+1}) &= \begin{pmatrix} (1-u_{n+1})(DJ_n)(\tilde{u}) & -J_n(\tilde{u}) \\ u_{n+1} \cdot e_1 & u_1 \end{pmatrix}\;,\end{aligned}$$

wobei $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^n$ der erste Einheitsvektor ist. Also ist J differenzierbar. Es gilt $\det(DJ_2)(u_1, u_2) = u_1$. In der Rekursionsformel ist die erste Spalte

gegeben durch $\big((1-u_{n+1})\frac{\partial J_n}{\partial u_1}, u_{n-1}\big)^t = \big(\frac{(1-u_{n+1})}{u_1}J_n, u_{n-1}\big)^t$, so daß Addition der $\frac{u_1}{1-u_{n+1}}$-fachen ersten Spalte zur letzten ergibt:

$$\begin{aligned}\det(DJ_{n+1})(\tilde{u}, u_{n+1}) &= \det\begin{pmatrix} (1-u_{n+1})(DJ_n)(\tilde{u}) & 0 \\ u_{n+1}\cdot e_1 & u_1 + \frac{u_1 u_{n+1}}{1-u_{n+1}} \end{pmatrix} \\ &= u_1(1-u_{n+1})^{n-1}\det\big(\ (DJ_n)(\tilde{u})\ \big)\ .\end{aligned}$$

Somit gilt $|\det(DJ_n)(u_1,\ldots,u_n)| = u_1^{n-1}(1-u_3)(1-u_4)^2\cdots(1-u_n)^{n-2}$, und J_n ist ein Diffeomorphismus. Damit haben wir den folgenden Satz bewiesen:

Satz 24.5 *Eine auf $(\mathbb{R}_+)^n$ bzw. auf $(\Delta^n)^o$ definierte Funktion f ist genau dann über $(\mathbb{R}_+)^n$ bzw. $(\Delta^n)^o$ integrierbar, wenn die Funktion $|\det(DJ_n)|f\circ J_n$ über $\mathbb{R}_+\times W^{n-1}$ bzw. über W^n integrierbar ist, wobei $W^k := (]0,1[)^k$ der offene Würfel ist. In diesem Fall gilt*

$$\begin{aligned}&\int_{\substack{(\mathbb{R}_+)^n \\ bzw.\ \Delta^n}} dx\ f(x) \\ &= \int_{\substack{\mathbb{R}^+ \\ bzw.\ ]0,1[}} du_1\ u_1^{n-1}\int_{]0,1[} du_2 \int_{]0,1[} du_3\ (1-u_3)\ldots\int_{]0,1[} du_n\ (1-u_n)^{n-2}\ f(J_n(u_1,\ldots,u_n)).\end{aligned}$$

Speziell erhalten wir $v_n(\Delta^n) = \frac{1}{n!}$.

Beispiel 24.6 (Beta-Funktion) Wir integrieren die für $p,q>0$ stetige und beschränkte Funktion $f(x,y) = x^{p-1}y^{q-1}e^{-x-y}$ über $(\mathbb{R}^+)^2$ mit der Jacobi-Formel und mit dem Satz von Fubini:

$$\begin{aligned}\int_{(\mathbb{R}^+)^2} d(x,y)\ x^{p-1}y^{q-1}e^{-x-y} &= \int_0^\infty du_1\ u_1^{p+q-1}\int_0^1 du_2\ (1-u_2)^{p-1}u_2^{q-1} \\ &= \Gamma(p+q)\int_0^1 du_2\ (1-u_2)^{p-1}u_2^{q-1} \\ \text{Fubini}\ &= \int_0^\infty dx\ x^{p-1}e^{-x}\int_0^\infty dy\ y^{q-1}e^{-y} = \Gamma(p)\Gamma(q)\ .\end{aligned}$$

Somit gilt für die als Beta-Funktion bezeichnete Funktion $B(p,q)$

$$B(p,q) := \int_0^1 dt\ (1-t)^{p-1}t^{q-1} = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\ . \qquad \triangleleft$$

Beispiel 24.7 Mit der Jacobi-Abbildung lassen sich z.B. zweidimensionale Integrale des folgenden Typs lösen (dabei ist $p,q>0$):

$$\begin{aligned}\int_{\Delta^2} d(x,y)\,x^{p-1}y^{q-1}f(x+y) &= \int_{]0,1[} du_1\ u_1^{p+q-1}\ f(u_1)\int_{]0,1[} du_2\ (1-u_2)^{p-1}u_2^{q-1} \\ &= B(p,q)\int_{]0,1[} du_1\ u_1^{p+q-1}\ f(u_1)\ .\end{aligned}$$

Die obige Gleichung gilt, wenn eines der Integrale existiert. Statt über Δ^2 und $]0,1[$ kann auch über $(\mathbb{R}_+)^2$ und $\mathbb{R}_+$ integriert werden. $\triangleleft$

Die Integration über das Standardsimplex ist deshalb so wichtig, weil sich durch Potenzabbildungen viele Integrationsgebiete darauf zurückführen lassen. Dazu wird für $\alpha_i, a_i > 0$ folgende Transformation betrachtet:

$$(y_1, \ldots, y_n) = T(x_1, \ldots, x_n) := (a_1 x_1^{\frac{1}{\alpha_1}}, \ldots, a_n x_n^{\frac{1}{\alpha_n}}) \; .$$

Die Transformation T bildet $(\mathbb{R}_+)^n$ diffeomorph auf sich selbst ab. Sie bildet andererseits das Innere des Standardsimplex Δ^n diffeomorph auf das Innere des verallgemeinerten Simplex

$$\Delta^{\alpha_1,\ldots,\alpha_n}_{a_1,\ldots,a_n} := \left\{ (y_1, \ldots, y_n) \in \mathbb{R}^n \; : \; y_i \geq 0 \; , \; \left(\frac{y_1}{a_1}\right)^{\alpha_1} + \ldots \left(\frac{y_n}{a_n}\right)^{\alpha_n} \leq 1 \right\}$$

ab. Das sind dann z.B. Viertelkreise ($n = 2$, $\alpha_1 = \alpha_2 = 2$, $a_1 = a_2 = r$) oder Kugeloktanten, ...

Die Determinante des Differentials ist offenbar

$$|\det(DT)(x_1, \ldots, x_n)| = \frac{a_1 \cdots a_n}{\alpha_1 \cdots \alpha_n} x_1^{\frac{1}{\alpha_1}-1} \cdots x_n^{\frac{1}{\alpha_n}-1} \; .$$

Satz 24.8 *Eine auf auf dem verallgemeinerten Simplex $(\Delta^{\alpha_1,\ldots,\alpha_n}_{a_1,\ldots,a_n})$ definierte Funktion f ist genau dann über dieses verallgemeinerte Simplex integrierbar, wenn die Funktion $f(a_1 x_1^{\frac{1}{\alpha_1}}, \ldots, a_n x_n^{\frac{1}{\alpha_n}}) x_1^{\frac{1}{\alpha_1}-1} \cdots x_n^{\frac{1}{\alpha_n}-1}$ über das Standardsimplex integrierbar ist. In diesem Fall gilt*

$$\int_{\Delta^{\alpha_1,\ldots,\alpha_n}_{a_1,\ldots,a_n}} dy \; f(y) = \frac{a_1 \cdots a_n}{\alpha_1 \cdots \alpha_n} \int_{\Delta^n} d(x_1, \ldots, x_n) \; x_1^{\frac{1}{\alpha_1}-1} \cdots x_n^{\frac{1}{\alpha_n}-1} f(a_1 x_1^{\frac{1}{\alpha_1}}, \ldots, a_n x_n^{\frac{1}{\alpha_n}}) \; .$$

Durch Kombination mit der Jacobi-Transformation entsteht so ein Diffeomorphismus $W^n \xrightarrow{J_n} \Delta^n \xrightarrow{T} \Delta^{\alpha_1,\ldots,\alpha_n}_{a_1,\ldots,a_n}$, mit dem wir Integrationen über ein verallgemeinertes Simplex auf Integrationen über den Würfel zurückführen können.

Beispiel 24.9 Wir berechnen das Volumen eines Ellipsoiden-Oktanten $EO := \Delta^{2,2,2}_{a,b,c}$ über die Jacobi-Transformation:

$$\begin{aligned} v_3(EO) &= \int_{\Delta^{2,2,2}_{a,b,c}} dy \\ &= \frac{abc}{8} \int_{\Delta^3} d(x_1, x_2, x_3) \; x_1^{-\frac{1}{2}} x_2^{-\frac{1}{2}} x_3^{-\frac{1}{2}} \end{aligned}$$

$$= \frac{abc}{8} \int_0^1 du_1 u_1^2 \int_0^1 du_2 \int_0^1 du_3 (1-u_3) \cdot u_1^{-\frac{3}{2}} u_2^{-\frac{1}{2}} (1-u_2)^{-\frac{1}{2}} u_3^{-\frac{1}{2}} (1-u_3)^{-1}$$
$$= \frac{abc}{8} \int_0^1 du_1 \, u_1^{\frac{1}{2}} \int_0^1 du_2 \, u_2^{-\frac{1}{2}} (1-u_2)^{-\frac{1}{2}} \int_0^1 du_3 \, u_3^{-\frac{1}{2}}$$
$$= \frac{abc}{8} \cdot \frac{2}{3} \cdot B(\tfrac{1}{2}, \tfrac{1}{2}) \cdot 2 = \frac{1}{8} \cdot \frac{4abc^3}{3} \frac{(\Gamma(\frac{1}{2}))^2}{\Gamma(1)} \, .$$

Für $a = b = c = r$ entsteht das Volumen des Kugeloktanten $KO = \Delta^{2,2,2}_{r,r,r}$ mit $v_n(KO) = \frac{1}{8}\kappa_3(r) = \frac{1}{8} \cdot \frac{4\pi r^3}{3}$, so daß wir $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ erhalten. Somit ist $v_2(EO) = \frac{\pi abc}{6}$, und das Ellipsoid $E = \{(x,y,z) \in \mathbb{R}^3 \ : \ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1\}$ hat das Volumen $v_3(E) = \frac{4\pi}{3} abc$. ◁

25 Konvergenzsätze. Satz von Fubini

Wir verallgemeinern nun den Satz 21.1, der uns ein Integrierbarkeitskriterium für monotone Folgen von Treppenfunktionen mit beschränktem Integral geliefert hat, auf monotone Folgen integrierbarer Funktionen mit beschränktem Integral.

Satz 25.1 (von der monotonen Konvergenz bzw. Satz von Beppo Levi) *Es sei $(f_k)_{k\in\mathbb{N}}$ eine monoton wachsende Folge integrierbarer Funktionen $f_k \in \mathcal{L}^1(\mathbb{R}^n)$. Die punktweise gebildete Grenzfunktion $f = \lim_{k\to\infty} f_k$ ist genau dann integrierbar, wenn die Folge der Integrale $\int_{\mathbb{R}^n} dx \ f_k(x)$ beschränkt ist. In diesem Fall gilt $\int_{\mathbb{R}^n} dx \ f(x) = \lim\limits_{k\to\infty} \int_{\mathbb{R}^n} dx \ f_k(x)$.*

Beweis. ($\Leftarrow$) Wegen $f_k \leq f$ ist $\int_{\mathbb{R}^n} dx \ f_k(x) \leq \int_{\mathbb{R}^n} dx \ f(x)$, so daß die Beschränktheit der Integrale notwendig ist.

($\Rightarrow$) Die Folge $\Big(\int_{\mathbb{R}^n} dx \ f_k(x) \Big)_{k\in\mathbb{N}}$ sei beschränkt (und monoton wachsend). Also konvergiert sie gegen einen Grenzwert in $\mathbb{R}$. Jede konvergente Folge ist eine Cauchy-Folge, so daß es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß für alle $k \geq l \geq N$ gilt

$$\|f_k - f_l\|_1 = \int_{\mathbb{R}^n} dx \, \big| f_k(x) - f_l(x) \big| = \int_{\mathbb{R}^n} dx \, \big(f_k(x) - f_l(x) \big)$$
$$= \int_{\mathbb{R}^n} dx \ f_k(x) - \int_{\mathbb{R}^n} dx \ f_l(x) = \Big| \int_{\mathbb{R}^n} dx \ f_k(x) - \int_{\mathbb{R}^n} dx \ f_l(x) \Big| < \epsilon \, .$$

Folglich ist $(f_k)_{k\in\mathbb{N}}$ eine L^1-Cauchyfolge, die nach dem Satz von Riesz-Fischer einen L^1-Grenzwert $\tilde{f}$ hat, so daß eine Teilfolge (f_{k_ν}) fast überall punktweise gegen $\tilde{f}$ konvergiert. Damit gilt $f = \lim_{\nu\to\infty} f_{k_\nu} = \tilde{f}$ fast überall punktweise.

Nach dem Modifikationssatz ist dann auch f integrierbar mit

$$\int_{\mathbb{R}^n} dx\ f(x) = \int_{\mathbb{R}^n} dx\ \tilde{f}(x) = \lim_{\nu\to\infty} \int_{\mathbb{R}^n} dx\ f_{k_\nu}(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ f_k(x)\ . \qquad \square$$

Wir benötigen ein weiteres Integrierbarkeitskriterium:

Satz 25.2 (von der majorisierten Konvergenz bzw. Satz von Lebesgue) *Es sei $(f_k)_{k\in\mathbb{N}}$ eine Folge integrierbarer Funktionen auf $\mathbb{R}^n$, die fast überall punktweise gegen eine Funktion f konvergiert. Es gebe eine integrierbare Funktion F mit $|f_k| \le F$ für alle $k \in \mathbb{N}$. Dann ist f integrierbar, und es gilt* $\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ f_k(x)$.

Beweis. Es genügt, reellwertige Funktionen zu betrachten. Es gibt eine Nullmenge N, so daß $F(x) < \infty$ und $\lim_{k\to\infty} f_k(x) \to f(x)$ für alle $x \notin N$. Für $x \in N$ sei $F(x) = f(x) = f_k(x) = 0$ gesetzt, so daß (f_k) punktweise gegen f konvergiert. Sei $g_{k,\nu} = \max(f_k, f_{k+1}, f_{k+\nu})$. Dann ist für festes k die Folge $(g_{k,\nu})_{\nu\in\mathbb{N}}$ monoton wachsend, und die Folge der Integrale ist beschränkt durch das Integral von F. Nach dem Satz von Beppo Levi ist $g_k := \lim_{k\to\infty} g_{k,\nu} = \sup_{j\ge k} f_j$ integrierbar, mit

$$\Big| \int_{\mathbb{R}^n} dx\ g_k(x) \Big| \le \int_{\mathbb{R}^n} dx\ F(x)\ .$$

Andererseits konvergiert $(g_k)_{k\in\mathbb{N}}$ monoton fallend gegen f. Analog zum Beweis des Satzes von Beppo Levi ist damit (g_k) eine L^1-Cauchy-Folge, die eine L^1-Grenzwert $\tilde{g}$ hat, gegen den eine Teilfolge g_{k_l} fast überall punktweise konvergiert. Somit ist $f = \tilde{g} = \lim_{l\to\infty} g_{k_l}$ fast überall, und es gilt $\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ g_k(x)$.

Das Verfahren wird wiederholt für $h_{k,\nu} = \min(f_k, f_{k+1}, f_{k+\nu})$ mit Limes und $h_k := \lim_{k\to\infty} h_{k,\nu} = \inf_{j\ge k} f_j$ und $\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ h_k(x)$. Aus $h_k \le f_k \le g_k$ folgt die Behauptung. $\square$

Beispiel 25.3 Wir beweisen $\int_0^\infty dx\ \dfrac{x^{s-1}}{e^x - 1} = \Gamma(s)\zeta(s)$ für $s \in \mathbb{C}$ mit $\mathrm{Re}(s) > 1$. Diese Formel wird beim Planckschen Strahlungsgesetz benötigt.

Der Integrand ist punktweiser Limes der Regelfunktionen $f_k = \sum_{n=1}^k x^{s-1}e^{-nx}$. Es gilt $|x^{s-1}| = |e^{\ln x(s-1)}| = |e^{\ln x(\mathrm{Re}(s)-1)}| = x^{\mathrm{Re}(s)-1}$. Damit ist $|f_k|$ über $\mathbb{R}_+$ uneigentlich Riemann-integrierbar, folglich f_k Lebesgue-integrierbar. Es gilt

$$|f_k| \le F = \sum_{n=1}^\infty x^{\mathrm{Re}(s)-1}e^{-nx} = \frac{x^{\mathrm{Re}(s)-1}}{e^x - 1}\ ,$$

und für $\mathrm{Re}(s) > 1$ ist F wieder Lebesgue-integrierbar. Somit gilt nach dem Satz von Lebesgue

$$\int_0^\infty dx\, \frac{x^{s-1}}{e^x - 1} = \lim_{k\to\infty} \sum_{n=1}^{k} \int_0^\infty dx\, x^{s-1} e^{-nx} = \lim_{k\to\infty} \sum_{n=1}^{k} \frac{1}{n^s} \Gamma(s) = \zeta(s)\Gamma(s)\,. \quad \square$$

Es geht nun um die Verallgemeinerung des Satzes von Fubini von stetigen Funktionen auf beliebige integrierbare Funktionen. Zur Vereinfachung der Schreibweise sei $X = \mathbb{R}^p$ und $Y = \mathbb{R}^{n-p}$. Wie üblich entsteht aus $f(x, y)$ durch Festhalten von $y \in Y$ die Funktion f_y auf X und durch Festhalten von $x \in X$ die Funktion f_x auf Y.

Satz 25.4 (Fubini (allgemeinster Fall)) *Es sei $f : X \times Y \to \mathbb{R} \cup \{\infty\}$ eine integrierbare Funktion. Dann gilt:*

i) *Abgesehen von einer möglichen Nullmenge $N \subset Y$ ist für festes $y \in Y \setminus N$ die Funktion f_y über X integrierbar.*

ii) *Die durch* $F(y) := \begin{cases} \int_X dx\, f(x,y) & \text{für } y \in Y \setminus N \\ 0 & \text{für } y \in N \end{cases}$
definierte Funktion $F : Y \to \mathbb{R}$ ist über Y integrierbar, und es gilt

$$\int_{x\times Y} d(x,y)\, f(x,y) = \int_Y dy\, F(y) \equiv \int_Y dy \int_X dx\, f(x,y)\,.$$

Beweis. Zu f gibt es nach Satz 23.13 eine Folge $(\phi_k)_{k\in\mathbb{N}}$ von Treppenfunktionen und eine Nullmenge $A \subset X \times Y$, so daß für $(x, y) \notin A$ gilt $\lim_{k\to\infty} \phi_k(x, y) = f(x, y)$ und außerdem $\sum_{k=0}^{\infty} \|\phi_{k+1} - \phi_k\|_1 < \infty$. Im folgenden bezeichnen $\| \ \|_{1,X}$ und $\| \ \|_{1,Y}$ die L^1-Halbnormen auf X und Y.

Wir beweisen zunächst, daß es eine Nullmenge $N' \subset Y$ gibt, so daß für $y \in Y \setminus N'$ gilt, daß $A_y := \{x \in X \ : \ (x, y) \in A\} \subset X$ eine Nullmenge ist. Wegen $\|\mathbf{1}_A\|_1 = 0$ gibt es für jedes $\epsilon > 0$ zu $\mathbf{1}_A$ eine Hüllreihe $\sum_{i=0}^{\infty} \mathbf{1}_{Q_i}$ mit $\sum_{i=0}^{\infty} v(Q_i) < \epsilon$. Die Quader zerlegen sich in $Q_i = Q_i' \times Q_i''$ mit $Q_i' \in X$ und $Q_i'' \in Y$. Für festes y ist (ohne Ausschluß einer Nullmenge) $\mathbf{1}_{A_y} \le \sum_{i=1}^{\infty} \mathbf{1}_{Q_i'} \mathbf{1}_{Q_i''}(y)$ und damit $a(y) := \int_X dx \mathbf{1}_{A_y}(x) \le \sum_{i=1}^{\infty} v(Q_i') \mathbf{1}_{Q_i''}(y)$. Integration der so definierten Funktion a über Y liefert $\|a\|_{1,Y} = 0$, d.h. die Existenz einer Nullmenge $N' \subset Y$, so daß $0 = a(y) = \|\mathbf{1}_{A_y}\|_{1,X} = 0$ für $y \notin N$. Das war zu zeigen.

Folglich konvergiert für festes $y \in Y \setminus N'$ die Folge $(\phi_{k,y})_{k\in\mathbb{N}}$ fast überall auf X (nämlich für $x \notin A_y$) gegen ϕ_y.

Sei $H_k(y) := \int_X dx\, |\phi_{k+1,y}(x) - \phi_{k,y}(x)|$. Nach dem Kleinen Satz von Fubini (Satz 22.1) gilt $\int_Y dy\, H_k(y) = \int_{X\times Y} d(x,y)\, |\phi_{k+1}(x,y) - \phi_k(x,y)| = \|\phi_{k+1} - \phi_k\|_1$

und damit

$$\sum_{k=0}^{\infty} \int_Y dy\ H_k(y) < \infty\ . \qquad (*)$$

Die Folge $(G_s)_{s\in\mathbb{N}}$ der Funktionen $G_s = \sum_{k=0}^{s} H_s$ ist monoton wachsend. Nach (*) gilt

$$I_s := \int_Y dy\ G_s(y) = \sum_{k=0}^{s} \int_Y dy\ H_k(y) \le \sum_{k=0}^{\infty} \int_Y dy\ H_k(y) < \infty\ ,$$

d.h. die Folge $(I_s)_{s\in\mathbb{N}}$ der Integrale ist beschränkt. Nach dem Satz von Beppo Levi ist damit die Grenzfunktion $\sum_{k=0}^{\infty} H_k$ integrierbar. Nach Satz 23.8 gibt es höchstens eine Nullmenge N'', so daß $H_k(y) < \infty$ für alle $y \in Y \setminus N''$. Somit gilt für alle $y \in Y \setminus N$, mit $N = N' \cup N''$,

$$\sum_{k=0}^{\infty} \|\phi_{k+1,y} - \phi_{k,y}\|_{1,X} < \infty\ . \qquad (**)$$

Damit ist $(\phi_{k,y})_{k\in\mathbb{N}}$ eine L^1-Cauchyfolge auf X, die nach dem Satz von Riesz-Fischer fast überall punktweise gegen eine über X integrierbare Funktion $\tilde{f}_y$ konvergiert. Damit ist nach dem Modifikationssatz auch ϕ_y integrierbar, und für $y \in Y \setminus N$ gilt

$$\lim_{k\to\infty} \int_X dx\ \phi_k(x,y) = \int_X dx\ f(x,y) = F(y)\ .$$

Wir betrachten nun die Treppenfunktionen $\Phi_k(y) := \int_X dx\ \phi_k(x,y)$. Die Folge $(\Phi_k)_{k\in\mathbb{N}}$ konvergiert fast überall punktweise gegen F. Außerdem ist

$$\begin{aligned}\sum_{k=0}^{\infty} \|\Phi_{k+1} - \Phi_k\|_{1,Y} &= \sum_{k=0}^{\infty} \int_Y dy\ \Big| \int_X dx\ (\phi_{k+1,y}(x) - \phi_{k,y}(x))\Big| \\ &\le \sum_{k=0}^{\infty} \int_Y dy \int_X dx\ \big|\phi_{k+1,y}(x) - \phi_{k,y}(x)\big| = \sum_{k=0}^{\infty} \|\phi_{k+1} - \phi_k\|_1 < \infty\ .\end{aligned}$$

Somit ist $(\Phi_k)_{k\in\mathbb{N}}$ eine L^1-Cauchyfolge, die nach dem Satz von Riesz-Fischer fast überall punktweise gegen eine integrierbare Funktion $\tilde{F}$ konvergiert. Damit ist auch F integrierbar, und es gilt

$$\begin{aligned}\int_Y dy\ F(y) &= \lim_{k\to\infty} \int_y dy\ \Phi_k(y) = \lim_{k\to\infty} \int_y dy\ \int_X dx\ \phi_k(x,y) \\ &= \lim_{k\to\infty} \int_{X\times Y} d(x,y)\ \phi_k(x,y) = \int_{X\times Y} d(x,y)\ f(x,y)\ . \qquad \square\end{aligned}$$

Im Beweis können die Rollen von X, Y vertauscht werden, so daß für eine über $X \times Y$ integrierbare Funktion gilt

$$\int_{X\times Y} d(x,y)\ f(x,y) = \int_X dx \int_Y dy\ f(x,y) = \int_Y dy \int_X dx\ f(x,y)\ .$$

Die Voraussetzung der Integrierbarkeit über $X \times Y$ ist entscheidend. Es gibt Beispiele für Funktionen, für die die Integrale $\int_X dx \int_Y dy\ \ f(x,y)$ und $\int_Y dy \int_X dx\ f(x,y)$ existieren, ohne daß $f(x,y)$ integrierbar ist. Die Umkehrung ist der Satz von Tonelli, den wir ohne Beweis angeben.

Definition 25.5 Es sei A eine Vereinigung abzählbar vieler kompakter Mengen. Eine Funktion $f : A \to \mathbb{R} \cup \{\infty\}$ heißt *lokal-integrierbar*, wenn sie über jede kompakte Teilmenge $K \subset A$ integrierbar ist.

Satz 25.6 (Tonelli) *Eine lokal-integrierbare oder fast überall stetige Funktion $f : X \times Y \to \mathbb{R}$ ist genau dann über $X \times Y$ integrierbar, wenn wenigstens eines der iterierten Integrale $\int_X dx \int_Y dy\ |f(x,y)|$ oder $\int_Y dy \int_X dx\ |f(x,y)|$ existiert. Ist das der Fall, so gilt*

$$\int_{X\times Y} d(x,y)\ f(x,y) = \int_X dx \int_Y dy\ f(x,y) = \int_Y dy \int_X dx\ f(x,y)\ .$$

26 Beweis des Transformationssatzes

Wir unterteilen den Beweis in folgende Schritte:

i) Diskussion der Nullmengen

ii) Volumen eines affin transformierten Würfels

iii) Volumen eines diffeomorph transformierten Würfels

iv) Beweis für Treppenfunktionen

v) Beweis im allgemeinen Fall

In den Beweisen ist es vorteilhaft, auf $\mathbb{R}^n$ die *Maximumsnorm* $\|(x_1, \ldots, x_i)\|_\infty := \max(x_1, \ldots, x_n)$ einzuführen.

Lemma 26.1 *Es seien $U, V \subset \mathbb{R}^n$ offen, $T : U \to V$ ein Diffeomorphismus und $N \subset U$ eine Nullmenge. Dann ist auch $T(N) \subset V$ eine Nullmenge.*

Beweis. Für $x, y \in U$ und $t \in [0,1]$ sei $g(t) := T(x + t(y - x))$. Dann gibt es nach dem Mittelwertsatz ein $\theta \in [0,1]$ mit

$$T(y) - T(x) = g(1) - g(0) = g'(\theta) = (DT)(x + \theta(y - x)) \cdot (y - x)\ ,$$

wobei im letzten Schritt die Kettenregel benutzt ist. Damit gilt nach Definition der Norm einer linearen Abbildung

$$\|T(y) - T(x)\|_\infty \leq \sup_{\theta\in[0,1]} \|(DT)(x + \theta(y - x))\| \cdot \|y - x\|_\infty \ .$$

Zu N gibt es für jedes $\epsilon > 0$ eine Überdeckung durch abzählbar viele kompakte Würfel $W_k \subset U$ mit $\sum_{k=0}^\infty v_n(W_k) < \epsilon$. Entsprechend ist $T(N) \subset \bigcup_{k=0}^\infty T(W_k)$. Sei jetzt $x, y \in N \cap W_k$. Auf W_k ist $\|(DT)(x+\theta(y-x))\|$ als stetige Funktion auf einer kompakten Menge beschränkt, d.h. $\|T(y) - T(x)\|_\infty \leq L \cdot \|y - x\|_\infty$ für alle $x, y \in W_k$ und somit auch für alle $x, y \in N \cap W_k$. Also gilt $v_n(T(W_k)) \leq L^n v_n(W_k)$, d.h. $T(N \cap W_k)$ ist eine Nullmenge. Dann ist auch $T(N)$ als Vereinigung von abzählbar vielen Nullmengen eine Nullmenge. □

Lemma 26.2 *Es seien $a_1, \ldots, a_n \in \mathbb{R}^n$ und $P(a_1, \ldots, a_n) := \{x = t_1 a_1 + \cdots + t_n a_n \ : \ t_i \in [0, 1]\} \subset \mathbb{R}^n$ das durch diese Vektoren aufgespannte Parallelotop. Dann gilt*

$$v_n(P(a_1, \ldots, a_n)) = |\det(a_1, \ldots, a_n)| \ ,$$

wobei a_i auf der rechten Seite die i-te Zeile einer $(n \times n)$-Matrix ist.

Beweis. Aus der Definition und dem Beweis der Eindeutigkeit der Determinante im letzten Semester folgt, daß der Betrag der Determinante eindeutig definiert ist durch

(D1) $|\det(\ldots, \lambda a_i, \ldots)| = |\lambda||\det(\ldots, a_i, \ldots)|$

(D2) $|\det(\ldots, a_i, \ldots, a_j, \ldots)| = |\det(\ldots, a_i + a_j, \ldots, a_j, \ldots)|$

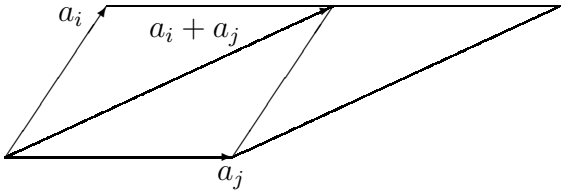
(D3) $|\det(e_1, \ldots, e_n)| = 1$

Die Punkte in (D1), (D2) bedeuten, daß die jeweiligen Zeilen der rechten und linken Seite identisch sind.

Wir beweisen, daß auch das Volumen diese Eigenschaften hat. (D3) ist klar.

(D1) Sei $P_\lambda := P(a_1, \ldots, a_{i-1}, \lambda a_i, a_{i+1}, \ldots, a_n)$. Die Parallelotope P_1 und P_{-1} sind nur gegeneinander verschoben und haben nach Cavalieri das gleiche Volumen. Wir können uns also auf $\lambda > 0$ beschränken. Für natürliche Zahlen $\lambda = l \in \mathbb{N} \setminus \{0\}$ gilt offenbar $v_n(P_l) = l v_n(P_1)$ nach Aneinandereihung von l Parallelotopen P_1 in i-ter Richtung. Sei $\lambda = \frac{p}{q}$ eine rationale Zahl mit $p, q \in \mathbb{N} \setminus \{0\}$. Dann gilt $v_n(P_{q\cdot\lambda}) = q v_n(P_\lambda) = v_n(P_p) = p v_n(P_1) = \frac{p}{q} v_n(P_q)$. Schließlich finden wir für $\lambda \in \mathbb{R}_+$ zu jedem $\epsilon > 0$ rationale Zahlen $r_1 \leq \lambda \leq r_2$ mit $|r_1 - r_2| \leq \frac{\epsilon}{v_n(P_1)}$. Das ergibt $v_n(P_{r_1}) \leq v_n(P_\lambda) \leq v_n(P_{r_2})$ und damit $|v_n(P_\lambda) - \lambda v_n(P_1)| \leq \epsilon$. Somit gilt $(D1)$ für alle $\lambda \in \mathbb{R}$.

(D2) Nach dem Prinzip von Cavalieri genügt es, die jeweiligen Flächen in der $\{i, j\}$-Ebene zu vergleichen:



Wieder nach Cavalieri haben die durch $\{a_i, a_j\}$ bzw. $\{a_i + a_j, a_j\}$ aufgespannten Parallelogramme die gleiche Fläche. Das beendet den Beweis. □

Sei nun $W = P(e_1, \ldots, e_n) \subset \mathbb{R}^n$ der Einheitswürfel und $T : \mathbb{R}^n \to \mathbb{R}^n$ mit $T : x \mapsto A \cdot x$ eine lineare Abbildung. Dann ist $A \cdot e_i = a_i$ die i-te Spalte von A bzw. die i-te Zeile von A^t. Aus der Linearität von T folgt somit $T(W) = P(a_1, \ldots, a_n)$. Aus Lemma 26.2 und $\det A^t = \det A$ ergibt sich schließlich $v_n(T(W)) = |\det A| \cdot v_n(W)$.

Lemma 26.3 *Für jeden kompakten Würfel $W \subset U$ gilt*
$v_n(T(W)) \leq \max\limits_{x \in W} |\det(DT)(x)| \cdot v_n(W)$.

Beweis. Da jede kompakte Teilmenge meßbar ist und das Bild einer kompakten Teilmenge im $\mathbb{R}^n$ unter einer stetigen Abbildung wieder kompakt ist, sind $v_n(W)$ und $v_n(T(W))$ definiert. Wegen Lemma 26.1 genügt es, den Fall $v_n(W) > 0$ zu beweisen.

i) Wir setzen $\alpha := \frac{v_n(T(W))}{v_n(W)}$. Durch Halbierung sämtlicher Kanten zerlegen wir W in 2^n achsenparallele gleich große Teilwürfel. Dann gibt es einen Teilwürfel W_1 mit $v_n(T(W_1)) \geq \alpha v_n(W_1)$. Durch Wiederholung dieser Zerlegung gewinnt man eine Folge $W_1 \supset W_2 \supset \ldots$ von Würfeln mit $v_n(T(W_i)) \geq \alpha v_n(W_i)$. Nach Intervallschachtelungsprinzip (z.B. Satz 2.6) gibt es einen Punkt $a \in W$, der in allen W_i liegt. Sei $b := T(a)$ der Bildpunkt. Wir können das Koordinatensystem so verschieben, daß $a = b = 0$ gilt.

Ist m_k der Mittelpunkt des k-ten Würfels und hat der Ausgangswürfel W die Kantenlänge $2L$, dann ist $W_k = \{x \in U \ : \ \|x - m_k\|_\infty \leq \frac{L}{2^k}\}$. Nach Definition der Differenzierbarkeit von T im Nullpunkt gilt $T(x) = T(0) + (DT)(0) \cdot x + \phi(x)$ mit $\lim\limits_{x \to 0, x \neq 0} \dfrac{\phi(x)}{\|x\|_\infty} = 0$. Wir setzen $A := (DT)(0) \in GL(n, \mathbb{R})$. Wegen $T(0) = 0$ gilt dann $T(x) = A \cdot (x + \|x\|_\infty \cdot r(x))$, wobei $r(x) := \frac{1}{\|x\|_\infty} A^{-1} \cdot \phi(x)$ gegen 0 konvergiert für $x \neq 0$. Also gibt es zu jedem $\epsilon > 0$ ein $\delta > 0$, so daß $\|r(x)\|_\infty < \frac{\epsilon}{2}$ für alle $x \in \mathbb{R}^n$ mit $\|x\|_\infty < \delta$. Sei l ein Index, so daß für alle $x \in W_l$ gilt $\|x\|_\infty \leq 2 \cdot \frac{L}{2^l} < \delta$. Dann gilt

$$\|(x + \|x\|_\infty \cdot r(x)) - m_l\|_\infty \leq \|x - m_l\|_\infty + \|x\|_\infty \cdot \|r(x)\|_\infty \leq \frac{L}{2^l} + 2 \cdot \frac{L}{2^l} \cdot \frac{\epsilon}{2} = \frac{L}{2^l}(1 + \epsilon)$$

für alle $x \in W_l$. Somit ist die Menge $V_l := \{x + \|x\| \cdot r(x) \ : \ x \in W_l\}$ enthalten im Würfel $W_l^\epsilon := \{x \in \mathbb{R}^n \ : \ \|x - m_l\|_\infty \leq \frac{L}{2^k} \cdot (1 + \epsilon)\}$.

Also gilt $T(W_l) = A \cdot V_l \subset A \cdot W_l^\epsilon$ und weiter

$$v_n(T(W_l)) \leq v_n(A \cdot W_l^\epsilon) = |\det A|(1+\epsilon)^n v_n(W_l) \ .$$

ii) Angenommen, es gelte $\alpha > \max_{x\in W} |\det(DT)(x)| \geq |\det A|$. Dann finden wir ein $\epsilon > 0$, für das auch $\alpha > (1+\epsilon)^n |\det A|$ gilt. Das bedeutet $v_n(T(W_l)) < \alpha v_n(W_l)$ im Widerspruch zur Konstruktion von W_l. □

Lemma 26.4 *Sei $K \subset U$ eine kompakte Teilmenge, so daß der Rand ∂K eine Nullmenge ist. Dann gilt*

$$\min_{x\in K} |\det(DT)(x)| \cdot v_n(K) \leq v_n(T(K)) \leq \max_{x\in K} |\det(DT)(x)| \cdot v_n(K) \ .$$

Beweis. Die kompakte Menge K und damit ihr offenes Innere $K \setminus \partial K$ ist meßbar mit $v_n(K) = v_n(K \setminus \partial K)$. Zu $K \setminus \partial K$ gibt es eine Ausschöpfung $A_0 \subset A_1 \subset \ldots$ mit $K \setminus \partial K = \bigcup_{k=0}^\infty A_k$, wobei die kompakten Teilmengen $A_k = \cup_{i_k=0}^{p_k} W_{i_k}$ durch Zusammenkleben von kompakten Würfeln W_{i_k} der Kantenlängen $1, \frac{1}{2}, \ldots, \frac{1}{2^k}$ entlang ihrer Ränder gebildet werden. Dann ist $v_n(K \setminus \partial K) = \sum_{k=0}^\infty \sum_{i_k=0}^{p_k} v_n(W_{i_k})$. Wegen der Stetigkeit und Bijektivität von T gilt $T(K) \setminus \partial T(K) = T(K \setminus \partial K) = T(\bigcup_{k=0}^\infty \bigcup_{i_k=0}^{p_k} W_{i_k})$ und dann mit Lemma 26.1 und Lemma 26.3

$$\begin{aligned} v_n(T(K)) = v_n(T(K \setminus \partial K)) &= \sum_{k=0}^\infty \sum_{i_k=0}^{p_k} v_n(T(W_{i_k})) \\ &\leq \max_{x\in K} |\det(DT)(x)| \cdot \sum_{k=0}^\infty \sum_{i_k=0}^{p_k} v_n(W_{i_k}) = \max_{x\in K} |\det(DT)(x)| \cdot v_n(K) \ . \end{aligned}$$

Andererseits folgt daraus durch Vertauschung der Rollen von K und $T(K)$

$$v_n(K) = v_n(T^{-1}(T(K))) \leq \max_{y\in T(K)} |\det(DT^{-1})(y)| \cdot v_n(T(K)) \ .$$

Nach dem Satz über implizite Funktionen gilt $(DT^{-1})(y) = ((DT)(x))^{-1}$ mit $x := T^{-1}(y)$, also $|\det(DT^{-1})(y)| = \frac{1}{|\det(DT)(x)|}$. Nun ist $|\det(DT^{-1})(y)|$ dort maximal, wo $|\det(DT)(x)|$ minimal ist. Das bedeutet

$$v_n(T(K)) \geq \min_{x\in K} |\det(DT)(x)| \cdot v_n(K) \ . \qquad \square$$

Satz 26.5 *Der Transformationssatz gilt für jede Treppenfunktionen ϕ auf V, deren Träger* $\operatorname{supp}(\phi) := \overline{\{y \in \mathbb{R}^n \ : \ \phi(y) \neq 0\}}$ *Teilmenge von V ist.*

Beweis. Wegen der Linearität des Integrals genügt es, den Transformationssatz für die charakteristische Funktion eines Quaders zu beweisen. Weiter brauchen wir nach Lemma 26.1 nur kompakte Quader $Q \in V$ zu betrachten, da der Rand

eines Quaders eine Nullmenge ist. Die Integrierbarkeit von $|\det DT|\mathbf{1}_Q \circ T$ ist klar, denn $\mathbf{1}_Q \circ T$ verschwindet außerhalb der kompakten Menge $T^{-1}(Q) \subset U$, und $|\det DT|$ ist stetig auf $T^{-1}(Q)$. Zu zeigen bleibt

$$\int_Q dy = v_n(Q) = \int_{T^{-1}(Q)} dx\, |\det(DT)(x)| \,.$$

Da die stetige Funktion $|\det(DT^{-1})|^{-1}|$ auf der kompakten Menge Q gleichmäßig stetig ist (Satz 3.4), gibt es zu jedem $\epsilon > 0$ eine Zerlegung $Q = Q_1 \cup \cdots \cup Q_p$ in kompakte Quader, die nur Randpunkte gemeinsam haben und so klein sind, daß $\max_{y\in Q_i} |\det(DT^{-1})(y)|^{-1} - \min_{y\in Q_i} |\det(DT^{-1})(y)|^{-1} \le \epsilon$. Dann gilt im Urbild $T^{-1}(Q_i)$

$$\max_{x\in T^{-1}(Q_i)} |\det(DT)(x)| v_n(T^{-1}(Q_i)) - \min_{x\in T^{-1}(Q_i)} |\det(DT)(x)| v_n(T^{-1}(Q_i)) \le \epsilon v_n(T^{-1}(Q_i)) \,.$$

Sowohl $\displaystyle\int_{T^{-1}(Q_i)} dx\, |\det(DT)(x)|$ als auch $v_n(Q_i)$ nach Lemma 26.4 sind enthalten im Intervall

$$\Big[\min_{x\in T^{-1}(Q_i)} |\det(DT)(x)| v_n(T^{-1}(Q_i)), \max_{x\in T^{-1}(Q_i)} |\det(DT)(x)| v_n(T^{-1}(Q_i))\Big] \,.$$

Also gilt

$$\Big| \int_{T^{-1}(Q_i)} dx\, |\det(DT)(x)| - v_n(Q_i)\Big| \le \epsilon v_n(T^{-1}(Q_i)) \,.$$

Summation über alle Teilquader liefert

$$\Big| \int_{T^{-1}(Q)} dx\, |\det(DT)(x)| - v_n(Q)\Big| \le \sum_{i=1}^{p} \Big| \int_{T^{-1}(Q_i)} dx\, |\det(DT)(x)| - v_n(Q_i)\Big| \le \epsilon v_n(T^{-1}(Q)) \,.$$

Für $\epsilon \to 0$ ergibt sich die Behauptung. □

Beweis des Transformationssatzes. i) Nach Definition der Integrierbarkeit gibt es zu jeder über $V \subset \mathbb{R}^n$ integrierbaren Funktion f und jedem $\epsilon > 0$ eine auf einer beschränkten Teilmenge des $\mathbb{R}^n$ definierte Treppenfunktion ϕ_ϵ mit $\|f - \phi_\epsilon\|_1 \le \frac{\epsilon}{2}$. Wegen $|f_V - \phi_\epsilon \mathbf{1}_V| \le |f_V - \phi_\epsilon|$ gilt dann auch $\|f_V - \phi_\epsilon \mathbf{1}_V\|_1 \le \frac{\epsilon}{2}$.

Sei $B \subset \mathbb{R}^n$ eine beschränkte offene Teilmenge mit $\operatorname{supp}(\phi_\epsilon) \subset B$ und $M = \max_{x\in\operatorname{supp}(\phi_\epsilon)} |f_\epsilon(x)|$. Dann gibt es zu der beschränkten offenen Teilmenge $V \cap B$ eine Vereinigung $A = Q_0 \cup \cdots \cup Q_k \subset V \cap B$ von endlich vielen kompakten Quadern Q_i mit $\big|v_n(V\cap B) - v_n(A)\big| < \frac{\epsilon}{2M}$. Damit ist $\phi_\epsilon \mathbf{1}_A$ eine Treppenfunktion mit $\operatorname{supp}(\phi_\epsilon \mathbf{1}_A) \subset V$, für die gilt

$$\|\phi_\epsilon \mathbf{1}_A - \phi_\epsilon \mathbf{1}_V\|_1 = \|\phi_\epsilon \mathbf{1}_A - \phi_\epsilon \mathbf{1}_{V\cap B}\|_1 \le M|v_n(A) - v_n(V\cap B)| < \frac{\epsilon}{2} \,.$$

Somit gilt $\|f_V - \phi_\epsilon \mathbf{1}_A\|_1 \leq \|f_V - \phi_\epsilon \mathbf{1}_V\|_1 + \|\phi_\epsilon \mathbf{1}_V - \phi_\epsilon \mathbf{1}_A\|_1 < \epsilon$, d.h. wir können annehmen, daß die approximierenden Treppenfunktionen zu f_V ihren Träger in V haben. Nach Auswahl einer Teilfolge gemäß Satz 23.13 gibt es also zu f_V eine Familie $(\phi_k)_{k\in\mathbb{N}}$ von Treppenfunktionen mit Träger in V, die fast überall punktweise gegen f konvergieren und außerdem L^1-konvergent gegen f sind.

ii) Wir betrachten die Folge der Funktionen $\tilde{\phi}_k := |\det(DT)|(\phi_k \circ T)$. Nach Satz 26.5 ist $\tilde{\phi}_k$ über U integrierbar, und es gilt

$$\|\tilde{\phi}_k - \tilde{\phi}_l\|_{1,U} = \int_U dx\, |\tilde{\phi}_k(x) - \tilde{\phi}_l(x)| = \int_V |\phi_k(y) - \phi_l(y)| = \|\phi_k - \phi_l\|_{1,V}\ .$$

Damit ist $(\tilde{\phi}_k)_{k\in\mathbb{R}^n}$ eine L^1-Cauchyfolge auf U, so daß eine Teilfolge fast überall punktweise gegen eine über U integrierbare Funktion $\tilde{f}$ konvergiert mit $\displaystyle\int_U dx\ \tilde{f}(x) = \lim_{k\to\infty} \int_U dx\ \tilde{\phi}_k(x)$. Andererseits konvergiert $\tilde{\phi}_k$ auch fast überall punktweise gegen die Funktion $|\det(DT)|(f\circ T)$. Nach dem Modifikationssatz ist dann auch $|\det(DT)|(f \circ T)$ über U integrierbar, und es gilt

$$\int_U dx\, |\det(DT)(x)|\ f(T(x)) = \lim_{k\to\infty} \int_U dx\ \tilde{\phi}_k(x) = \lim_{k\to\infty} \int_V dy\ \phi_k(y) = \int_V dy\ f(y)\ .$$

Ist umgekehrt $|\det(DT)|(f\circ T)$ über U integrierbar, dann folgt durch Vertauschen der Rollen von T und T^{-1}, daß $|\det(DT^{-1})|\big(|\det(DT)|(f\circ T)\big)\circ T^{-1} = f$ über V integrierbar ist. Damit ist der Transformationssatz bewiesen. □

27 Integration über Untermannigfaltigkeiten

Wir haben bisher die Methoden entwickelt, um Funktionen über Teilmengen $A \subset \mathbb{R}^n$ zu integrieren und z.B. Volumina solcher Teilmengen zu berechnen. Wir können damit aber noch nicht die Oberfläche des Randes von A berechnen. Die dazu notwendigen Ideen sollen nun kurz vorgestellt werden, wobei wir aus Zeitgründen keine Beweise angeben können. Wir erinnern an die folgende Charakterisierung von Untermannigfaltigkeiten, die wir in Satz 9.5 für die Richtung ($\Rightarrow$) bewiesen hatten:

Satz 27.1 *Eine Teilmenge $M \subset \mathbb{R}^{n+k}$ ist genau dann eine n-dimensionale differenzierbare Untermannigfaltigkeit, wenn es zu jedem Punkt $a \in M$ eine offene Umgebung $V \subset M$, eine offene Umgebung $T \subset \mathbb{R}^n$ und eine Immersion $\phi : T \to \mathbb{R}^{n+k}$ gibt, so daß T durch ϕ homöomorph auf V abgebildet wird.*

Bemerkungen. Zur Erinnerung: Immersion bedeutet, daß ϕ differenzierbar ist mit $\operatorname{rang}(D\phi)(t) = n$ für alle $t \in T$.

Insbesondere gibt es eine Überdeckung einer Untermannigfaltigkeit durch offene Mengen V_i. Dann heißt (V_i, ϕ_i) mit $\phi_i : T_i \to V_i$ eine *lokale Karte* von M. Für

$V_{ij} := V_i \cap V_j \neq \varnothing$ gibt es zwei Homöomorphismen $\phi_i^{-1} : V_{ij} \to \phi_i^{-1}(V_{ij}) \subset T_i \subset \mathbb{R}^n$ und $\phi_j^{-1} : V_{ij} \to \phi_j^{-1}(V_ij) \subset T_j \subset \mathbb{R}^n$. Über die Konstruktion von ϕ im Beweis von Satz 9.5 zeigt man, daß $\tau_{ij} := \phi_j^{-1} \circ \phi_i : \phi_i^{-1}(V_{ij}) \to \phi_j^{-1}(V_{ij})$ sogar ein Diffeomorphismus ist zwischen Teilmengen des $\mathbb{R}^n$. Man sagt, die Kartenwechsel sind Diffeomorphismen.

Die Integration einer Funktion f über die Teilmenge $V \subset M$ wird nun über einen analogen Transformationssatz durch Integration der Funktion "$|\det D\phi|$"$(f \circ \phi)$ über T erklärt. Das Problem dabei ist, daß die Determinante der rechteckigen Matrix $D\phi$ so nicht existiert. Man zeigt, daß

$$\text{"}|\det D\phi|\text{"} := \sqrt{\det((D\phi)^t \cdot (D\phi))}$$

die richtigen Eigenschaften hat. Dabei ist $(D\phi)^t(D\phi)$ punktweise eine $n \times n$-Matrix. Entsprechend definiert man das Integral einer Funktion f über eine Karte (V, ϕ) von M mit $\phi(T) = V$ zu

$$\int_{(V,\phi)} dS\ f(x) := \int_T du\ \sqrt{\det((D\phi)^t(u) \cdot (D\phi)(u))}\ f(\phi(u))\ . \qquad (*)$$

Die Idee ist wieder zu beweisen, daß das durch die n Vektoren $a_1, \ldots, a_n \subset \mathbb{R}^{n+k}$ aufgespannte Parallelotop das Volumen $\det(A^t \cdot A)$ hat, wobei a_i die Spalten von $A \in M((n+k) \times n, \mathbb{R})$ sind. Dann identifiziert man das Parallelotop mit dem Bild des n-dimensionalen Einheitswürfels im $\mathbb{R}^{n+k}$, dessen letzte k Komponenten identisch Null sind, unter einer affinen Transformation. Durch analoge Konvergenzbetrachtungen wie im Transformationssatz beweist man, daß durch (*) das Integral einer Funktion über $V \subset M$ sinnvoll definiert ist.

Beispiel 27.2 (Oberfläche der dreidimensionalen Kugel) Es sei

$$M := \{(x_1, x_2, x_3) \in \mathbb{R}^3\ :\ x_1^2 + x_2^2 + x_3^2 = R^2\}$$

die Oberfläche der dreidimensionalen Kugel vom Radius R. Mittels Polarkoordinaten gewinnen wir die folgende Abbildung $\phi :]0, 2\pi[\times]0, \pi[\to V \subset M$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \phi(\varphi, \vartheta) := \begin{pmatrix} R\cos\varphi\sin\vartheta \\ R\sin\varphi\sin\vartheta \\ R\cos\vartheta \end{pmatrix}\ .$$

Das offene Rechteck $T :=]0, 2\pi[\ \times\ ]0, \pi[$ wird durch ϕ homöomorph auf die Teilmenge $V := M \setminus HK$ abgebildet, d.h. aus der Kugeloberfläche wird der Halbkreis $HK := \{(x_1, x_2, x_3) \in \mathbb{R}^3\ :\ x_2 = 0\ ,\ x_1 \geq 0\ ,\ x_1^2 + x_3^3 = R^2\}$ herausgeschnitten. Dann ist

$$(D\phi)(\varphi, \vartheta) = \begin{pmatrix} -R\sin\varphi\sin\vartheta & R\cos\varphi\cos\vartheta \\ R\cos\varphi\sin\vartheta & R\sin\varphi\cos\vartheta \\ 0 & -R\sin\vartheta \end{pmatrix}\ ,$$

$$(D\phi)^t(\varphi, \vartheta) \cdot (D\phi)(\varphi, \vartheta) = \begin{pmatrix} R^2\sin^2\vartheta & 0 \\ 0 & R^2 \end{pmatrix}\ ,$$

so daß wir erhalten:

$$\int_{(V,\phi)} dS\ f(x) = R^2 \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta\ \sin\vartheta\ f(\phi(\varphi,\vartheta))\ .$$

Der Halbkreis HK ist eine Nullmenge. Man kann wieder zeigen, daß Nullmengen für die Integrationstheorie ignoriert werden können. Also stimmt das Integral mit dem Integral über ganz M überein. Insbesondere erhalten wir für $f = 1$ die Oberfläche der zweidimensionalen Sphäre vom Radius R zu $\int_M dS = 4\pi R^2$. $\triangleleft$

Eine wichige Konsequenz des Transformationssatzes ist, daß (*) unabhängig von der Wahl der Karte ist. Gibt es zu V zwei Karten (V,ϕ_1) und (V,ϕ_2) mit Immersionen $\phi_i : T_i \to \mathbb{R}^{n+k}$, so daß $\phi_i : T_i \to V_i$ Homöomorphismen sind, so gilt $\int_{(V,\phi_1)} dS\ f(x) = \int_{(V,\phi_2)} dS\ f(x)$. Zum Beweis verwendet man, daß $\phi_2 \circ \phi_1^{-1} : T_1 \to T_2$ ein Diffeomorphismus ist und den entsprechenden Transformationssatz, der $|\det D(\phi_2 \circ \phi_1^{-1})|$ beinhaltet.

Das nutzt man aus, um Integrationen über Untermannigfaltigkeiten zu definieren, die aus mehreren Karten zusammengesetzt werden müssen. Wir betrachten nur den einfachsten Fall, daß es endlich viele Karten $(V_1,\phi_1),\ldots,(V_p,\phi_p)$ gibt, die $M = V_1 \cup \cdots \cup V_p$ überdecken. Dann kann man immer eine Familie von Funktionen $f_i : M \to \mathbb{R}$ konstruieren mit

- $\mathrm{supp}(\alpha_i) \subset V_i$
- $\sum_{i=1}^p \alpha_i(x) = 1$ für alle $x \in M$.

Eine solche Familie heißt *Zerlegung der Eins*. Mittels Zerlegung der Eins erhalten wir:

$$\begin{aligned}\int_M dS\ f(x) &= \sum_{i=1}^p \int_M dS\ f(x)\alpha_i(x) = \sum_{i=1}^p \int_{V_i} dS\ (f\alpha_i)(x) \\ &= \sum_{i=1}^p \int_{T_i} du_i\ \sqrt{\det((D\phi_i)^t(u_i)\cdot(D\phi_i)(u_i))}\ (f\alpha_i)(\phi_i(u_i))\ .\end{aligned}$$

Die Eigenschaften der Zerlegung der Eins garantieren, daß diese Definition unabhängig von der Wahl der Überdeckung und der α_i ist. Die Konstruktion verallgemeinert sich sogar auf abzählbar viele Karten, wenn sich jeweils nur endlich viele schneiden und $|f\alpha_i|$ integrierbar ist:

Definition 27.3 Es sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit, ausgestattet mit einem Atlas lokaler Karten (V_i,ϕ_i) entsprechend Satz 27.1, so daß $M = \bigcup_{i=0}^\infty V_i$ und jeder Punkt $x \in M$ nur in endlich vielen V_i enthalten ist.

Eine auf M definierte Funktion f heißt *über M integrierbar*, wenn es eine dem Atlas $(V_i,\phi_i)_{i\in\mathbb{N}}$ untergeordnete Zerlegung der Eins $(\alpha_i)_{i\in\mathbb{N}}$ gibt, so daß

i) Jede Funktion $f\alpha_i$ ist über V_i (damit über M) integrierbar

ii) $\sum_{i=0}^{\infty} \int_{(V_i,\phi_i)} dS\ |(f\alpha_i)(x)| < \infty.$

Dann ist das Integral von f über M (unabhängig von der Zerlegung der Eins) definiert durch

$$\int_M dS\ f(x) := \sum_{i=0}^{\infty} \int_{T_i} du_i\ \sqrt{\det((D\phi_i)^t(u_i) \cdot (D\phi_i)(u_i))}\ (f\alpha_i)(\phi_i(u_i))\ .$$

Wir sehen uns noch einige interessante Integrale über Karten an:

Beispiel 27.4 (Integrale entlang Kurven) Es sei $I \subset \mathbb{R}^n$ ein offenes Intervall und $\gamma : I \to \mathbb{R}^n$ eine differenzierbare Kurve. Unter der Annahme $(D\gamma)(t) = \gamma'(t) \neq 0$ handelt es sich um eine Immersion, d.h. γ spielt die Rolle von ϕ in Definition 27.3. Sei f jetzt eine Funktion auf der Kurve $\Gamma := \gamma(I)$, dann ist das Kurvenintegral gegeben durch

$$\int_\Gamma dS\ f(x) = \int_I dt\ \|\gamma'(t)\|\ f(t)\ .$$

Das motiviert auch die Definition 16.5 des komplexen Kurvenintegrals. Insbesondere ist das eindimensionale Volumen der Kurve $\Gamma := \gamma(I)$ ihre Bogenlänge L,

$$L := v_1(\Gamma) = \int_\Gamma dS = \int_I dt\ \|\gamma'(t)\|\ . \qquad \triangleleft$$

Die Berechnung von Determinanten des Typs $\det(A^t \cdot A)$ mit $A \in M((n+k) \times n, \mathbb{R})$ kann für große n, k sehr umständlich werden. Hier hilft das Determinanten-Multiplikationstheorem (Binet-Cauchy-Theorem) entscheidend weiter:

Satz 27.5 (Binet-Cauchy) *Es seien $A = (a_1, \ldots, a_{n+k}) \in M((n+k) \times n, \mathbb{R})$ und $B = (b_1, \ldots, b_{n+k}) \in M((n+k) \times n, \mathbb{R})$ zwei rechteckige Matrizen, gebildet aus den Zeilenvektoren $a_i, b_i \in \mathbb{R}^n$. Für $1 \leq m_1 < m_2 < \cdots < m_n \leq n+k$ seien quadratische Matrizen $A^{m_1 m_2 \ldots m_n} := (a_{m_1}, a_{m_2}, \ldots, a_{m_n}) \in M(n \times n, \mathbb{R})$ und $B^{m_1 m_2 \ldots m_n} := (b_{m_1}, b_{m_2}, \ldots, b_{m_n}) \in M(n \times n, \mathbb{R})$ definiert. Dann gilt*

$$\det(A^t \cdot B) = \sum_{1 \leq m_1 < m_2 \cdots < m_n \leq n+k} (\det A^{m_1 m_2 \ldots m_n}) \cdot (\det B^{m_1 m_2 \ldots m_n})$$

Die Summe läuft über die $\binom{n+k}{n} = \frac{(n+k)!}{n!k!}$ verschiedenen Möglichkeiten, n der $n+k$ Zeilen der Matrizen auszuwählen.

Ein Beweis findet sich z.B. in G. Fischer: Lineare Algebra, Kapitel 3.3.

Beispiel 27.6 Es sei $T \subset \mathbb{R}^n$ offen und die Höhenfunktion $h : T \to \mathbb{R}$ differenzierbar. Dann ist die Abbildung $\phi : T \to \mathbb{R}^{n+1}$ mit $\phi(u) := (u, h(u))$ eine Immersion. Zur Berechnung von Integralen über den Graphen $\Gamma := \phi(T) \subset \mathbb{R}^{n+1}$ benötigen wir die Determinante der Matrix $G(u) = (D\phi)^t(u) \cdot (D\phi)(u)$. Dabei ist $(D\phi)(u) = \begin{pmatrix} E_n \\ (\operatorname{grad} h)(u) \end{pmatrix}$, wenn $(\operatorname{grad} h)(u) \in \mathbb{R}^n$ als Zeilenvektor betrachtet wird. Dann gilt mit den Bezeichnungen aus Satz 27.5

$$\det((D\phi)^{1,2,\dots,n}) = 1 \ , \quad \det((D\phi)^{1,\dots,i-1,i+1,\dots,n+1}) = \pm\partial_i h$$

und damit $\det(D\phi)^t(u) \cdot (D\phi)(u) = 1 + \|(\operatorname{grad} h)(u)\|^2$. Somit erhalten wir das Integral einer Funktion f über den Graph $\Gamma := \phi(T) \subset \mathbb{R}^{n+1}$ (Höhenfäche) zu

$$\int_\Gamma dS \ f(x) = \int_T du \ \sqrt{1 + \|(\operatorname{grad} h)(u)\|^2} \ f(u, h(u)) \ .$$

Wir berechnen auf diese Weise noch einmal die Oberfläche der Halbkugel HK. Dazu sei $T := \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 < R^2\}$ und $h(x, y) := \sqrt{R^2 - x^2 - y^2}$. Dann ist $HK := (x, y, h(x, y)) \subset \mathbb{R}^3$, und wir erhalten

$$v_2(HK) = \int_{HK} dS = \int_T d(x, y) \frac{R}{\sqrt{R^2 - x^2 - y^2}} \ .$$

Das Integral lösen wir in Polarkoordinaten $x = r\cos\varphi$, $y = r\sin\varphi$. Mit Satz 24.3 erhalten wir

$$v_2(HK) = \int_0^R dr \, r \int_0^{2\pi} d\varphi \ \frac{1}{\sqrt{1 - \frac{r^2}{R^2}}} \overset{r=R\sin t}{=} 2\pi R^2 \int_0^{\frac{\pi}{2}} dt \ \sin t = 2\pi R^2 \ . \qquad \triangleleft$$

Beispiel 27.7 (Rotationsflächen im $\mathbb{R}^3$) Sei $I \subset \mathbb{R}$ offen und die Radiusfunktion $r : I \to \mathbb{R}_+$ differenzierbar. Sei $M := \{(x, y, z) \in \mathbb{R}^3 \ : \ z \in I \ , \ x^2 + y^2 = (r(z))^2\}$ die Rotationsfläche. Dann ist die Abbildung

$$\phi : \ I \times]0, 2\pi[\to M \setminus N \ , \qquad \phi(\varphi, z) := \big(r(z)\cos\varphi, \ r(z)\sin\varphi, \ z\big)$$

eine Immersion, wobei der Nullmeridian $N = \{(x, y, z) \in \mathbb{R}^3 \ : \ z \in I \ , \ y = 0 \ , \ x^2 = (r(z))^2\}$ herausgeschnitten ist. Wir erhalten

$$(D\phi)(z, \varphi) = \begin{pmatrix} r'(z)\cos\varphi & -r(z)\sin\varphi \\ r'(z)\sin\varphi & r(z)\cos\varphi \\ 1 & 0 \end{pmatrix}$$

und damit $(\det(D\phi)^t \cdot (D\phi))(z, \varphi) = r(z)^2(1 + (r'(z))^2)$. Da N eine Nullmenge ist, erhalten wir das Integral einer Funktion f über die Rotationsfläche zu

$$\int_M dS \ f(x) = \int_I dz \int_0^{2\pi} d\varphi \ r(z)\sqrt{1 + (r(z))^2} \, f\big(r(z)\cos\varphi, \ r(z)\sin\varphi, \ z\big) \ .$$

Für $I =]-R, R[$ und $r(z) = \sqrt{R^2 - z^2}$ erhalten wir die Oberfläche der dreidimensionalen Kugel $M = \{(x, y, y) \in \mathbb{R}^3 \ : \ x^2 + y^2 + z^2 = R^2\}$ zu

$$v_2(M) = \int_M dS = 2\pi \int_{-R}^{R} dz \, \sqrt{R^2 - z^2} \sqrt{1 + \Big(\frac{z}{\sqrt{R^2 - z^2}}\Big)^2} = 2\pi R \int_{-R}^{R} dz = 4\pi R^2.$$

◁

28 Der Gaußsche Integralsatz

Wir betrachten jetzt (differenzierbare) Hyperflächen im $\mathbb{R}^n$, d.h. $(n-1)$-dimensionale Untermannigfaltigkeiten $M \subset \mathbb{R}^n$. Lokal auf einer Teilmenge $U \subset M$ ist der Normalenvektorraum $N_a(U)$ ein eindimensionaler Untervektorraum des $\mathbb{R}^n$, gegeben durch Vielfache des Gradienten der Funktion $f : U \to \mathbb{R}^n$, die die Untermannigfaltigkeit beschreibt.

Definition 28.1 Ein *Einheitsnormalenfeld* auf einer Hyperfläche $M \subset \mathbb{R}^n$ ist ein stetiges Vektorfeld $\nu : M \to \mathbb{R}^n$, so daß in jedem Punkt $x \in M$ gilt

i) $\nu(x)$ steht senkrecht auf dem Tangentialraum $T_x(M)$

ii) $\|\nu(x)\| = 1$

Eine differenzierbare Hyperfläche heißt *orientierbar*, wenn es auf ihr ein Einheitsnormalenfeld gibt. Ein Paar (M, ν) mit festgelegtem Einheitsnormalenfeld ν heißt *orientierte Hyperfläche*.

Entweder es existieren zwei Einheitsnormalenfelder ν und $-\nu$, oder gar keines. Lokal in jeder Karte (V, ϕ) von M existiert immer ein Einheitsnormalenfeld $\nu = \frac{\operatorname{grad} f}{\|\operatorname{grad} f\|}$. Beim Zusammensetzen der Karten zu einer Überdeckung von M kann es aber das Problem geben, daß auf dem Durchschnitt $V_i \cap V_j$ sich die Einheitsnormalenfelder der Karten um das Vorzeichen unterscheiden. Bekanntestes Beispiel einer nichtorientierbaren Hyperfläche ist das Möbiusband.

Definition 28.2 Es sei (M, ν) eine orientierte differenzierbare Hyperfläche im $\mathbb{R}^n$. Ein Vektorfeld $F : M \to \mathbb{R}^n$ heißt *integrierbar über* M, wenn die Funktion $\langle F, \nu \rangle$ über M integrierbar ist. In diesem Fall setzt man

$$\int_{(M,\nu)} \vec{dS} \ F(x) := \int_M dS \ \langle F(x), \nu(x) \rangle \ .$$

Zur Formulierung des Gaußschen Integralsatzes benötigen wir den Begriff des $\mathcal{C}^1$-Polyeders:

Definition 28.3 Es sei $G \subset \mathbb{R}^n$ eine kompakte Teilmenge und ∂G der Rand von G. Ein Randpunkt $x \in \partial G$ heißt *regulärer Randpunkt*, wenn es eine Umgebung $U \subset \mathbb{R}^n$ von x und eine stetig differenzierbare Funktion $f : U \to \mathbb{R}$ gibt mit $(\operatorname{grad} f)(y) \neq 0$

für alle $y \in U$, so daß $G \cap U = \{y \in U \ : \ f(u) \leq 0\}$. Jeder nicht reguläre Randpunkt von ∂G heißt singulär. Die Menge der regulären Randpunkte heißt *regulärer Rand* $\partial_r G$. Die Menge der singulären Randpunkte heißt *singulärer Rand* $\partial_s G$. Die Menge G heißt $\mathcal{C}^1$-*Polyeder*, wenn $\partial_s G$ eine $n-1$-dimensionale Nullmenge ist.

Die Definition besagt, daß der reguläre Rand eine $(n-1)$-dimensionale Untermannigfaltigkeit (Hyperfläche) ist. Singuläre Randpunkte sind z.B. die Ecken und Kanten eines Quaders. Diese dürfen wir nicht ausschließen, da der Beweis des Gaußschen Integralsatzes auf den Fall der Quader zurückgeführt wird. Da es für $\partial_r G$ "innen" und "außen" gibt, ist $\partial_r G$ orientierbar. Das äußere Einheitsnormalenfeld ist dann dadurch ausgezeichnet, daß es für jeden Punkt $x \in \partial G \subset \mathbb{R}^n$ ein $\epsilon > 0$ gibt, so daß $x + t\nu(x) \notin G$ für alle $t \in]0, \epsilon[$.

In Vorbereitung des Gaußschen Integralsatzes sei an die Divergenz eines Vektorfeldes F auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ erinnert: Ist $F = (F_1, \ldots, F_n)$ mit differenzierbaren Funktionen $F_i : U \to \mathbb{R}$, dann ist die Divergenz des Vektorfeldes die Funktion $(\mathrm{div} F) = \partial_1 F_1 + \cdots + \partial_n F_n$.

Theorem 28.4 (Gauß) *Es sei $G \subset \mathbb{R}^n$ ein $\mathcal{C}^1$-Polyeder, und $\partial_r G$ sei durch das äußere Einheitsnormalenfeld orientiert. Sei $F : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld auf einer offenen Teilmenge $U \subset \mathbb{R}^n$ mit $G \subset U$. Ist die Divergenz des Vektorfeldes* div F *über $G \subset U$ integrierbar und das Vektorfeld F über den regulären Rand $\partial_r G$ integrierbar, dann gilt*

$$\int_G dy \ (\mathrm{div}\, F)(y) = \int_{\partial_r G} d\vec{S} \ F(x) \ .$$

Der entscheidende Schritt im Beweis ist die Betrachtung der Situation für einen kompakten achsenparallelen Quader, der offenbar ein $\mathcal{C}^1$-Polyeder ist.

Lemma 28.5 *Es sei $Q \subset \mathbb{R}^n$ ein offener Quader und $F = (F_1, \ldots, F_n)$ ein stetig differenzierbares Vektorfeld auf einer Umgebung U von Q. Dann gilt*

$$\int_Q dy \ (\mathrm{div}\, F)(y) = \int_{\partial Q} d\vec{S} \ F(x)$$

Beweis. Es sei $\nu = (\nu_1, \ldots, \nu_n)$ das äußere Einheitsnormalenfeld auf ∂Q. Wegen Linearität der Integrale genügt es zu zeigen, daß für jede auf U stetig differenzierbare Funktion f und jede Komponente $i = 1, \ldots, n$ gilt

$$\int_Q dy \ (\partial_i f)(y) = \int_{\partial_r Q} (\nu_i f)(x) \ .$$

Durch Umnumerierung der Richtungen können wir $i = n$ annehmen. Dann ist $Q = Q' \times]a, b[$, wobei $Q' \subset \mathbb{R}^{n-1}$ wieder ein offener Quader ist. Entsprechend sei

$y = (y', z)$ die Parametrisierung mit $y' \in \mathbb{R}^{n-1}$ und $z \in \mathbb{R}$. Der reguläre Rand von Q ist

$$\partial_r Q = \big((Q')^o \times \{a\}\big) \cup \big((Q')^o \times \{b\}\big) \cup \big(\partial_r Q' \times]a, b[\,\big) \ ,$$

wobei $(Q')^o$ das offene Innere von Q' ist. Für die n-te Komponente ν_n des äußeren Einheitsnormalenfeldes auf dem regulären Rand gilt dann

$$\nu_n(x) = \begin{cases} 1 & \text{auf } (Q')^o \times \{b\} \\ -1 & \text{auf } (Q')^o \times \{b\} \\ 0 & \text{auf } \partial Q' \times]a, b[\end{cases}$$

Also ist die Funktion $f\nu_n$ nur über die Randflächen $(Q')^o \times \{b\}$ und $(Q')^o \times \{a\}$ zu integrieren. Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$\begin{aligned} \int_{\partial_r Q} dx\ (\nu_n f)(x) &= \int_{(Q')^o} dy'\ f(y', b) - \int_{(Q')^o} dy'\ f(y', a) \\ &= \int_{(Q')^o} dy' \int_a^b dz\ (\partial_n f)(y', z) = \int_Q d(y', z)\ (\partial_n f)(y', z) \ . \quad \square \end{aligned}$$

Wir sehen uns einige Anwendungen des Gaußschen Integralsatzes an.

Beispiel 28.6 (Oberfläche der Einheitssphäre)
Es sei $G = K_n := \{x \in \mathbb{R}^n \ : \ \|x\| \leq 1\}$ die n-dimensionale Einheitskugel und $S^{n-1} := \partial G = \{x \in \mathbb{R}^n \ : \ \|x\| = 1\}$ die $(n-1)$-dimensionale Sphäre. Wir betrachten das Vektorfeld $F = x$ mit $(\operatorname{div} F)(x) = n$. Das äußere Einheitsnormalenfeld auf S^{n-1} ist $\nu(x) = x$. Dann gilt

$$\int_{K_n} dx\ (\operatorname{div} F)(x) = n \int_{K_n} dx\ = n\kappa_n = \int_{S^{n-1}} dS\ \langle x, x\rangle = \int_{S^{n-1}} dS =: \omega_n \ .$$

Die Oberfläche der S^{n-1} ist also $v_{n-1}(S^{n-1}) =: \omega_n = n\kappa_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$. $\triangleleft$

Beispiel 28.7 Es sei $G \subset \mathbb{R}^n$ ein $\mathcal{C}^1$-Polyeder, $a \in \mathbb{R}^n \setminus \partial G$ ein Punkt und $F(x) := \dfrac{x-a}{\|x-a\|^n}$. Wir beweisen

$$\int_{\partial_r G} \vec{dS}\ F(x) = \begin{cases} 0 & \text{für } a \notin G \\ \omega_n & \text{für } a \in G \end{cases} \qquad (*)$$

Zunächst gilt für $x \neq a$

$$\begin{aligned} (\operatorname{div} F)(x) &= \big\langle x - a, \operatorname{grad} \tfrac{1}{\|x-a\|^n} \big\rangle + \tfrac{1}{\|x-a\|^n} \operatorname{div}(x-a) \\ &= \big\langle x - a, \tfrac{-n}{\|x-a\|^{n+1}} \tfrac{x-a}{\|x-a\|} \big\rangle + \tfrac{n}{\|x-a\|^n} = 0 \ . \end{aligned}$$

Ist $a \notin G$, dann liefert der Gaußsche Integralsatz sofort die Behauptung (*).

Ist andererseits $a \in G$, dann gibt es wegen $a \notin \partial G$ eine offene Kugel $K_r(a) \subset G$. Dann ist $G_a := G \setminus K_r(a)$ wieder ein $\mathcal{C}^1$-Polyeder, und $(\operatorname{div} F)(y) = 0$ für alle $y \in G_a$. Es gilt $\partial_r G_a = \partial_r G \cup \partial K_r(a)$. Das äußere Einheitsnormalenfeld ν auf $\partial K_r(a)$ aus Sicht von G_a ist das innere Einheitsnormalenfeld aus Sicht von $K_r(a)$, so daß gilt $\nu(x) = -\frac{x-a}{\|x-a\|} = -\frac{1}{r}(x-a)$. Das ergibt $\langle \nu(x), F(x) \rangle = -\frac{1}{r^{n-1}}$ für alle $x \in K_a(r)$ und damit

$$\int_{\partial_r G} d\vec{S}\, F(x) = -\int_{\partial K_r(a)} d\vec{S}\, F(x) = \frac{1}{r^{n-1}} \int_{\partial K_r(a)} dS = \omega_n \; .$$

Die Gleichung (*) verallgemeinert sich auf Linearkombinationen von Vektorfeldern F. Sei $G \subset \mathbb{R}^3$ ein $\mathcal{C}^1$-Polyeder und seien $q_1, \ldots, q_k$ die Punktladungen in den Punkten $a_1, \ldots, a_k \in \mathbb{R}^3 \setminus \partial G$, dann ist nach dem Coulombschen Gesetz die elektrische Feldstärke in einem Punkt $x \neq a_i$ gegeben durch

$$E = \sum_{i=1}^{k} \frac{q_i}{4\pi} \frac{x - a_i}{\|x - a_i\|^3} \; .$$

Der Fluß des elektrischen Feldstärke durch die Oberfläche ∂G ist dann gleich der Gesamtladung in G:

$$\int_{\partial G} d\vec{S}\, E(x) = \sum_{\{i\,:\; a_i \in G\}} q_i \; . \qquad \triangleleft$$

Satz 28.8 (Greensche Formeln) *Sei $G \subset \mathbb{R}^n$ ein $\mathcal{C}^1$-Polyeder und f, g zweimal stetig differenzierbare Funktionen auf einer offenen Umgebung von G. Dann gilt*

$$\int_G dy\, \big\langle \operatorname{grad} f, \operatorname{grad} g \big\rangle(y) = \int_{\partial G} dS\, (f\, D_\nu g)(x) - \int_G dy\, (f\, \Delta g)(y) \; ,$$
$$\int_G dy\, (f\, \Delta g - g\, \Delta f)(y) = \int_{\partial G} dS\, (f\, D_\nu g - g\, D_\nu f)(x) \; ,$$

wobei $D_\nu f = \langle \nu, \operatorname{grad} f \rangle$ die Richtungsableitung in die äußere Normalenrichtung ist.

Beweis. Man wendet den Gaußschen Integralsatz auf das Vektorfeld $F = f \operatorname{grad} g$ an und benutzt die Leibnizregel. $\square$

Die Greenschen Formeln spielen eine wichtige Rolle bei der Lösung wichtiger partieller Differentialgleichungen.

Definition 28.9 Sei $U \subset \mathbb{R}^n$ offen. Eine zweimal stetig differenzierbare Funktion f auf U heißt *harmonisch*, wenn $(\Delta f)(x) = 0$ für alle $x \in U$.

Die Newtonschen Potentiale $N_a : (\mathbb{R}^n \setminus \{a\}) \to \mathbb{R}$ mit

$$N_a(x) := \begin{cases} -\dfrac{1}{(n-2)\omega_n} \dfrac{1}{\|x-a\|^{n-2}} & \text{für } n > 2 \\ \dfrac{1}{2\pi} \ln \|x-a\| & \text{für } n = 2 \end{cases}$$

sind auf $\mathbb{R}^n \setminus \{a\}$ harmonisch.

Satz 28.10 (Mittelwertsatz harmonischer Funktionen) *Es sei $h : U \to \mathbb{R}$ eine harmonische Funktion auf einer offenen Menge $U \subset \mathbb{R}^n$. Dann gilt für jede Kugel $K_r(a) \subset U$*

$$h(a) = \frac{1}{\omega_n r^{n-1}} \int_{\partial K_r(a)} dS\ h(x)\ .$$

Beweis. (für $n \geq 2$) Sei $G := \overline{K_r(a)} \setminus K_\rho(a)$ die Kugelschale mit innerem Radius ρ und äußerem Radius $r > \rho$, und sei $S_r := \partial K_r(a)$ und $S_\rho := \partial K_\rho(a)$. Dann sind h, N_a harmonisch auf G, so daß nach der 2. Greenschen Formel gilt

$$\int_{S_r} dS\ (hD_\nu N_a - N_a D_\nu h)(x) = \int_{S_\rho} dS\ (hD_\nu N_a - N_a D_\nu h)(x)\ .$$

Dabei ist ν jeweils das äußere Einheitsnormalenfeld auf den Sphären. Die 1. Greensche Formel für $G = \overline{K_R(a)}$ sowie $f \mapsto 1$ und $g \mapsto h$ liefert $\int_{S_R} dS\ (D_\nu h)(x) = 0$ für $R = \rho$ und $R = r$. Da N_a auf S_R konstant ist, folgt $\displaystyle\int_{S_r} dS\ (hD_\nu N_a)(x) = \int_{S_\rho} dS\ (hD_\nu N_a)(x)$. Für alle $x \in S_R$ gilt $(D_\nu N_a)(x) = \frac{1}{\omega_n} \frac{1}{\|x-a\|^{n-1}} = \frac{1}{\omega_n R^{n-1}}$ und damit

$$\frac{1}{\omega_n r^{n-1}} \int_{S_r} dS\ h(x) = \frac{1}{\omega_n \rho^{n-1}} \int_{S_\rho} dS\ h(x)\ .$$

Für $\rho \to 0$ folgt aus der Stetigkeit von h die Behauptung. □

Mit den Greenschen Formeln beweist man auch den folgenden Satz über eine Lösung der Potentialgleichung:

Satz 28.11 *Sei $\rho : \mathbb{R}^n \to \mathbb{R}$ eine stetig differenzierbare Funktion mit kompaktem Träger. Für $x \in \mathbb{R}^n$ sei*

$$\phi(x) := \int_{\mathbb{R}^n} dy\ N_y(x)\, \rho(y)\ .$$

Dann ist $\phi : \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar, und es gilt $\Delta\phi = \rho$.

Dabei kann man sich ρ als Ladungsdichte vorstellen und ϕ als elektrisches Potential. Auf diese Weise findet man das Coulombsche Gesetz als Lösung der statischen Maxwellschen Gleichungen.