Integrationstheorie

Inhalt

Literatur

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

Teil I
Differenzierbare Abbildungen

Im Gegensatz zur eindimensionalen Differentialrechnung gibt es im Höherdimensionalen verschiedene Differenzierbarkeitsbegriffe.

1 Partielle Ableitungen

Wir betrachten Funktionen mehrerer Veränderlicher, also Abbildungen $f : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^n$. Sei $x = (x_1, \dots, x_n) \in U$. 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:

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

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 $f\big|_{I_j}(t) = f(x_1, \dots, x_{j-1}, t, x_{j+1}, x_n)$. 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 1.1 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 1.2 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 1.3 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 $(\partial_j f)(x) = -2a x_j f(r)$. $\triangleleft$

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

Beispiel 1.4 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 $f(0) = 0$. $\triangleleft$

Eine Verallgemeinerung der partiellen Ableitung ist die Richtungsableitung:

Definition 1.5 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 später zeigen, daß aus der Stetigkeit der partiellen Ableitungen die Existenz aller Richtungsableitungen folgt.

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

Definition 1.6 Sei $U \subset \mathbb{R}^n$ offen und $k \in \mathbb{N}^\times$. 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 1.7 (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

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

2 Differenzierbarkeit

Zentral bei differenzierbaren Abbildungen ist die lineare Approximierbarkeit in Analogie zur Steigung der Tangente an eine eindimensionale Funktion. Diese lineare Approximierbarkeit läßt sich koordinatenfrei formulieren und kann deshalb auch in unendlich-dimensionalen normierten Vektorräumen definiert werden. Für das totale Differential gibt es ein Analogon zur Kettenregel.

Definition 2.1 Sei $U \subset \mathbb{R}^n$ 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.

Eine differenzierbare Abbildung $f : U \to \mathbb{R}^m$ heißt *stetig differenzierbar*, wenn das Differential $Df : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ stetig ist.

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.
- Zur Definition der Stetigkeit von $Df : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ ist $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ als endlich-dimensionaler normierter Vektorraum aufzufassen. Eine mögliche Norm ist die *Operatornorm*.
- Oft schreibt man auch *df* oder f' für das totale Differential. Wir reservieren d für das später einzuführende äußere Differential.

Die Definition läßt sich auf unendlich-dimensionale normierte Vektorräume verallgemeinern:

Definition 2.2 Es seien $(X, \| \ \|_X)$ und $(Y, \| \ \|_Y)$ normierte Vektorräume und $U \subset X$ eine offene Teilmenge. Eine Abbildung $f : U \to Y$ heißt *Fréchet-differenzierbar* im Punkt $x \in U$, falls es einen linearen beschränkten Operator $A(x) \in \mathcal{B}(X, Y)$ und eine auf einer offenen Umgebung V von $0 \in X$ definierte Abbildung $\phi : V \to Y$ gibt, so daß

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

Dabei ist die Konvergenz in Y bezüglich $\| \ \|_Y$ erklärt. Der lineare beschränkte Operator $(Df)(x) = A(x)$ heißt dann die *Fréchet-Ableitung* von f im Punkt x.

Satz 2.3 *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 $f = (f_1, \dots, f_m)$ 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) Die Richtungsableitung ist

$$\begin{aligned}(D_v f)(x) &= \lim_{t\to 0} \frac{1}{t}\big(f(x+tv) - f(x)\big)\\ &= \lim_{t\to 0} \frac{1}{t}\big((Df)(x)\circ(tv) + \phi(tv)\big) = (Df)(x)\circ v\end{aligned}$$

wegen $\phi(tv) = o(\|tv\|) = o(t)$. In der Standardbasis ist $(Df)(x)\circ v = A(x)\circ v = \sum_{i=1}^m \sum_{j=1}^n v_j a_{ij}(x) e_i$. □

Nach Satz 2.3 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 2.4 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 $A(x,y) = \begin{pmatrix} \sin y & x\cos y \\ y^2e^x & 2ye^x \end{pmatrix}$. Wir betrachten

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

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

Aus partieller Differenzierbarkeit folgt nicht die totale Differenzierbarkeit. Zunächst muß die aus den partiellen Ableitungen gebildete Matrix nicht unbedingt linear sein. Aber selbst wenn sie linear ist, kann es noch Probleme mit dem Restglied geben:

Beispiel 2.5 Es sei

$$f(x,y) = \begin{cases} \dfrac{xy^3}{x^2+y^4} & \text{für } (x,y) \neq (0,0)\\ 0 & \text{für } (x,y) = (0,0)\end{cases}$$

Kritisch ist der Punkt $(x, y) = (0, 0)$. Dort existieren alle Richtungsableitungen,

$$\big(D_{(\cos\theta,\sin\theta)}f\big)(0,0) = \lim_{t\to 0}\frac{1}{t}\frac{t^4\cos\theta\sin^3\theta}{t^2\cos^2\theta + t^4\sin^4\theta} = 0\ .$$

Insbesondere ist f partiell differenzierbar in $(0, 0)$ mit $(\partial_x f)(0, 0) = (\partial_y f)(0, 0) = 0$. Die aus den partiellen Ableitungen gebildete Matrix, die Nullmatrix, ist linear. Deshalb ist das Restglied $\phi(\xi, \eta) = f(\xi, \eta)$. Für die gegen $(0, 0)$ konvergente Folge $(\xi_k, \eta_k) = (\frac{1}{2^k}\cos\theta, \frac{1}{2^k}\sin\theta)$ ist

$$\lim_{k\to\infty}\frac{\phi(\xi_k,\eta_k)}{\sqrt{\xi_k^2+\eta_k^2}} = 0\ ,$$

aber für die ebenfalls gegen 0 konvergente Folge $(\xi_k, \eta_k) = (\frac{1}{4^k}\cos\theta, \frac{1}{2^k}\sin\theta)$ haben wir

$$\lim_{k\to\infty}\frac{\phi(\xi_k,\eta_k)}{\sqrt{\xi_k^2+\eta_k^2}} = \lim_{k\to\infty}\frac{\frac{1}{2^{2k+3k}}\cos\theta\sin^3\theta}{\frac{1}{2^{4k}}(\cos^2\theta+\sin^4\theta)}\cdot\frac{2^k}{\sqrt{\frac{1}{2^k}\cos^2\theta+\sin^2\theta}} = \frac{\cos\theta\sin^2\theta}{\cos^2\theta+\sin^4\theta} \neq 0\ .$$

Somit ist $\phi(0, 0) \neq o(\|\xi\|)$, und f ist nicht differenzierbar in $(0, 0)$. $\triangleleft$

Ein hinreichendes Kriterium für (stetige) Differenzierbarkeit ist stetige partielle Differenzierbarkeit:

Satz 2.6 *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, und das Differential Df ist stetig im Punkt x.*

Beweis. Da U offen, gibt es ein $\delta > 0$ mit $K_\delta(x) \subset U$. Wir wählen ein $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ 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 $\lim\limits_{\substack{\xi\to 0\in\mathbb{R}^n\\ \xi\neq 0}} \dfrac{\phi(\xi)}{\|\xi\|} = 0$.

Somit ist f differenzierbar in x mit $(Df)(x) = \big((\partial_1 f)(x), \dots, (\partial_n f)(x)\big)$. Insbesondere ist (Df) stetig in x. □

Durch Kombination der Sätze 2.3 und 2.6 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 2.7 *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 differenzierbar 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 22.6 aus dem ersten Semester. i) ist klar. Zu ii):

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

Analog ergibt sich iii). □

Satz 2.8 (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!

Beispiel 2.9 Es sei

$$\begin{array}{lll}
f:\mathbb{R}\to\mathbb{R}^3 & \text{gegeben durch} & f(t) := (t, t^2+1, \sin t) , \\
g:\mathbb{R}^3\to\mathbb{R} & \text{gegeben durch} & g(x,y,z) := e^{xy} + z .
\end{array}$$

Zu bestimmen sei $D(g\circ f)(\pi)$.

Da f, g stetig partiell differenzierbar sind, gilt

$$(Dg)(x,y,z) = (\partial_x g,\ \partial_y g,\ \partial_z g)(x,y,z) = (ye^{xy},\ xe^{xy},\ 1)$$

und

$$(Df)(t) = \begin{pmatrix} \partial_t f_1 \\ \partial_t f_2 \\ \partial_t f_3 \end{pmatrix}(t) = \begin{pmatrix} 1 \\ 2t \\ \cos t \end{pmatrix}$$

Somit nach Kettenregel

$$\begin{aligned}
D(g\circ f)(\pi) &= (Dg)(f(\pi))\cdot(Df)(\pi) \\
&= \Big((\pi^2+1)e^{\pi(\pi^2+1)},\ \pi e^{\pi(\pi^2+1)},\ 1\Big)\cdot\begin{pmatrix} 1 \\ 2\pi \\ -1 \end{pmatrix} \\
&= (3\pi^2+1)e^{\pi(\pi^2+1)} - 1 .
\end{aligned}$$ ◁

Ein wichtiger Spezialfall ist die Ableitung einer Funktion längs einer Kurve:

Definition 2.10 Unter einer (stetigen) *Kurve* im $\mathbb{R}^n$ versteht man eine (stetige) Abbildung $c : I \to \mathbb{R}^n$ (wobei $I \subset \mathbb{R}^n$ aus mehr als einem Punkt besteht). Eine Kurve $c = (c_1, \dots, c_n) : I \to \mathbb{R}^n$ heißt *(stetig) differenzierbar* in $t \in I$, wenn jede Komponentenfunktion $c_i : I \to \mathbb{R}$ (stetig) differenzierbar in $t \in I$ ist. In diesem Fall heißt

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

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

Satz 2.11 (Ableitung entlang einer Kurve) *Es sei $c = (c_1, \dots, c_n) : I \to U$ 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*

$$D(f \circ c)(t) = (Df)(c(t)) \circ c'(t) = (D_{c'(t)}f)(c(t)) \qquad \square$$

Folglich ist die Ableitung einer stetig partiell differenzierbaren Funktion längs einer Kurve gegeben als die Richtungsableitung der Funktion in Richtung des Tangentialvektors der Kurve.

Beispiel 2.12 Die Ableitung der Funktion $f(x; y; z) := x^2 + y^2 - z^2$ entlang der Schraubenline $c(t) = (ht; \sin t; \cos t)$ ist

$$(f \circ c)'(t) = \left\langle \begin{pmatrix} 2ht \\ 2\sin t \\ -2\cos t \end{pmatrix}, \begin{pmatrix} h \\ \cos t \\ -\sin t \end{pmatrix} \right\rangle = 2h^2 t + 4 \sin t \cos t \; . \qquad \triangleleft$$

Über die Ableitung entlang einer Kurve können wir den Mittelwertsatz verallgemeinern:

Satz 2.13 (Mittelwertsatz) *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*

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

Dabei ist das Integral komponentenweise (vektoriell) zu verstehen. Sind f und c sogar stetig differenzierbar, so gibt es ein $\tau \in [\alpha, \beta]$ mit

$$f(b) - f(a) = (\beta - \alpha)(Df)(c(\tau)) \circ c'(\tau) \; .$$

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

$$\begin{aligned} f(b) - f(a) &= \int_0^1 dt \; (Df)(a + (b - a)t) \circ (b - a) \\ &= (Df)(a + (b - a)\tau) \circ (b - a) \; . \end{aligned}$$

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 $(DF)(t) = (Df)(c(t)) \circ c'(t)$. Die Formel für stetig differenzierbare Funktionen und Kurven folgt aus dem Mittelwertsatz der Integralrechnung. □

Für den folgenden Schrankensatz benötigen wir die Abschätzung

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

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

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

Satz 2.14 (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 $\|f(b) - f(a)\| \leq M\|b - a\|$ mit $M = \sup_{t\in[0,1]} \|(Df)(a + (b - a)t)\|_{op}$.*

Beweis. Nach Mittelwertsatz 2.13, obiger Abschätzung und Definition der Operatornorm (Satz 49.2 im 2. Semester) gilt

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

Aus dem Schrankensatz bzw. Mittelwertsatz folgt in Analogie zu Satz 23.6 aus dem 1. Semester folgender Identitätssatz für differenzierbare Abbildungen, den wir zur Wiederholung der Begriffe der offenen und zusammenhängenden Mengen angeben:

Satz 2.15 *Es sei $G \subset \mathbb{R}^n$ offen und zusammenhängend, und für eine stetig differenzierbare Funktion $f : G \to \mathbb{R}$ gelte $Df = 0$. Dann ist f konstant auf G.*

Beweis. Wir müssen eine stückweise stetig differenzierbare Kurve zwischen $x, y \in G$ konstruieren. Dazu sei

$$U_x := \{y \in G \ : \ \exists \text{stückweise stetig differenzierbare Kurve von } x \text{ nach } y\} \ .$$

Die Menge U_x ist offen: Sei $y \in U_x \subset G$, dann gibt es ein $\epsilon > 0$ mit $K_\epsilon(y) \subset G$. Andererseits läßt sich die Kurve von x nach y fortsetzen zu einer Kurve von x in jeden Punkt von $K_\epsilon(y)$.

Angenommen, $V_x = G \setminus U_x \neq \varnothing$. Sei $v \in V_x \subset G$ und $\epsilon > 0$ mit $K_\epsilon(v) \subset G$. Dann ist auch $K_\epsilon(v) \subset V_x$, denn gäbe es einen Punkt $w \in K_\epsilon(v)$, der durch eine stückweise differenzierbare Kurve mit x verbunden wäre, so wäre auch v mit x verbunden, Widerspruch. Also sind U_x, V_x offen, nichtleer und disjunkt mit $G = U_x \cup V_x$. Damit wäre G nicht zusammenhängend, Widerspruch.

Somit gibt es eine stückweise stetig differenzierbare Kurve zwischen zwei beliebigen Punkten aus G. Ist $Df = 0$ auf G, so ist nach Mittel- bzw. Schrankensatz $f(x) = f(y) = \text{const}$. □

3 Vektorfelder, Gradient und Divergenz

Definition 3.1 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 $x \in U$.

Ist $f : U \to \mathbb{R}$ differenzierbar, dann gilt nach Satz 2.6 für jeden Punkt $x \in U$ und jeden Vektor $v \in \mathbb{R}^n$

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

Wählen wir v als Einheitsvektor, $\|v\| = 1$, dann ist nach der Ungleichung von Cauchy-Schwarz $|D_v f| \leq \|\operatorname{grad} f\|$, mit Gleichheit genau dann, wenn v und $\operatorname{grad} f$ linear abhängig sind. 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.

Beispiel 3.2 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 3.3 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 $x = (x_1, \dots, x_n) \in U$ ein Vektor $v(x) = (v_1(x), \dots, v_n(x))$ angeheftet ist. Eine nützliche Konstruktion besteht darin, diese Vektoren als Tangentialvektoren an Kurven $c : I \to \mathbb{R}^n$ durch $x = c(t_0)$ zu betrachten. Ist $c(t) = (c_1(t), \dots, c_n(t))$, dann ist also

$$c_i'(t) = v_i(c(t)) \;, \quad c_i(t_0) = x_i \qquad \text{für } i = 1, \dots, n \text{ und } t, t_0 \in I$$

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 (Lipschitz-stetig) 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 3.4 Sei $v = (v_1, \dots, v_n) : U \to \mathbb{R}^n$ ein partiell differenzierbares Vektorfeld auf einer offenen Teilmenge $U \subset \mathbb{R}^n$. Dann heißt die Funktion

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

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 3.5 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 $\|x\| = r$

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

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, $F = -\mathrm{grad}\, V$. 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*, auf die wir hier nicht eingehen.

Es sei $U \subset \mathbb{R}^n$ offen und $g : U \to \mathbb{R}$ zweimal stetig differenzierbar. Dann ist $\mathrm{grad}\, f$ ein stetig differenzierbares Vektorfeld auf U, dessen Divergenz die Funktion $\mathrm{div}\mathrm{grad}\, f$ ist. Die Abbildung $\Delta : f \mapsto \mathrm{div}(\mathrm{grad}\, f)$ heißt *Laplace-Operator*, und es gilt

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2} \; .$$

Der Laplace-Operator tritt in wichtigen partiellen Differentialgleichungen auf, z.B. der Poisson-Gleichung

$$\Delta f = -\rho \; , \qquad \rho : U \to \mathbb{R} \; .$$

4 Die mehrdimensionale Taylorsche 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 $\partial_i^{\alpha_i} f = \underbrace{\partial_i \ldots \partial_i}_{\alpha_i \mathrm{mal}} f$. Ebenso setzt man $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

Lemma 4.1 *Sei $U \subset \mathbb{R}$ offen und sei $f : U \to \mathbb{R}$ eine k-mal stetig partiell 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

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

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 Satz 2.11 ist

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

Durch wiederholte Ableitung längs c jeder Funktion $(\partial_i f)(x + t\xi) = (g_i \circ c)(t)$, u.s.w., berechnen wir die höheren Ableitungen, wobei wir zunächst <u>nicht</u> den Satz von Schwarz verwenden, d.h. wir betrachten $\partial_i \partial_j$ und $\partial_j \partial_i$ als verschieden. Es ergibt sich

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

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 $\partial_{i_k} \dots \partial_{i_2} \partial_{i_1}$, die nach Umordnung das gleiche ∂^α ergeben. □

Satz 4.2 (Taylor) *Sei $U \subset \mathbb{R}$ offen und sei $f : U \to \mathbb{R}$ eine $(k+1)$-mal stetig partiell 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*

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

Beweis. Wir betrachten die Funktion einer Veränderlichen $g : [0, 1] \to \mathbb{R}$, die durch $g[t] := f(x + t\xi)$ gegeben ist. Nach Lemma 4.1 ist g eine $(k+1)$ mal stetig differenzierbare Funktion, auf die wir die eindimensionale Taylorsche Formel (Satz 25.1) anwenden können: Es gibt also ein $\theta \in [0, 1]$ mit

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

Einsetzen von $g(t) = f(x + t\xi)$ und Verwenden von Lemma 4.1 liefert die Behauptung. □

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

Satz 4.3 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine k-mal stetig partiell differenzierbare Funktion. Dann gilt für jedes $x \in U$*

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

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

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

Es gilt $|\xi^\alpha| = |\xi_1|^{\alpha_1} \cdots |\xi_n|^{\alpha_n} \leq \|\xi\|^{\alpha_1} \cdots \|\xi\|^{\alpha_n} = \|\xi\|^{|\alpha|}$. 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 aus Satz 2.6

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

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 $\alpha = (0, \dots, 0, 2, 0, \dots, 0)$ mit $\alpha! = 2$ und $\alpha = (0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots, 0)$ 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 $i \neq j$ ersetzt werden. Zusammen mit der Summe über $i = j$ von $\alpha = (0, \dots, 0, 2, 0, \dots, 0)$ ergibt sich die Beziehung.

Definition 4.4 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

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

die *Hessesche Matrix* von f im Punkt x.

Dabei heißt eine quadratische Matrix $A = (a_{ij})$ symmetrisch, wenn $a_{ij} = a_{ji}$ ist. Somit gilt

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

Die Hessesche Matrix ist wichtig bei der Untersuchung von lokalen Extrema einer Funktion. Zunächst gilt in Verallgemeinerung von Satz 23.2

Satz 4.5 *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 $1 \leq i \leq n$ gilt, folgt die Behauptung. □

Wir beweisen eine hinreichende Bedingung für lokale Extrema unter Verwendung der Hesseschen Matrix. Dazu benötigen wir:

Definition 4.6 Eine symmetrische Matrix $A \in M(n, \mathbb{R})$ heißt

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

Ist eine symmetrische Matrix $A \in M(n, \mathbb{R})$ positiv definit, dann definiert $\langle x, y \rangle_A := \langle x, Ax \rangle$ ein neues Skalarprodukt $\langle \ , \ \rangle_A$ auf $\mathbb{R}^n$. Nach Satz 48.14 aus dem 2. Semester folgt positive Definitheit aus der Positivität aller Eigenwerte. Ein besseres Kriterium ist das Determinantenkriterium aus Satz 48.15 des 2. Semesters.

Satz 4.7 *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig partiell 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 Extremwertsatz 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 $0 < t' < \delta'$. □

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 4.8 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. ◁

5 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

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

falls $(\partial_2 F)(x, f(x))$ invertierbar ist. Die Verallgemeinerung dieser Invertierbarkeitsbedingung ist zentral für die lokale Auflösbarkeit.

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

Satz 5.1 (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*

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

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*

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

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 $\Phi(f_*) = f_*$. □

Im Beweis des Satzes über implizite Funktionen werden wir den Banachschen Fixpunktsatz auf den Banach-Raum der stetigen und beschränkten Funktionen anwenden. Wir zitieren Satz 31.4 aus dem 2. Semester:

Satz 31.4 [2. Semester]. *Es sei X ein metrischer Raum und*

$$\mathcal{C}_b(X) := \{f : X \to \mathbb{C} \text{ stetig}\,, \quad \|f\| := \sup_{x\in X} |f(x)| < \infty\}$$

der Vektorraum der stetigen und beschränkten komplexwertigen Funktionen auf X. Dann ist $(\mathcal{C}_b(X), \|\ \|)$ vollständig (d.h. Banach-Raum) bezüglich der Supremums-Norm $\|\ \|$. *Insbesondere ist $(\mathcal{C}(X), \|\ \|)$ Banach-Raum, falls X kompakt ist.*

Satz 5.2 (über implizite Funktionen) *Seien $U_1 \subset \mathbb{R}^n$ und $U_2 \subset \mathbb{R}^m$ offene Teilmengen und $F = (F_1, \dots, F_m) : U_1 \times U_2 \to \mathbb{R}^m$ 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

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

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, und diese Lösung ist sogar differenzierbar.

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 $\left\|\left(\frac{\partial G}{\partial y}\right)(a,b)\right\|_{op} = 0$. 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

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

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.

Wir konstruieren zunächst g als stetige Funktion. Die Gebiete können dann etwas größer gewählt werden. Nach dem Satz von Taylor gilt

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

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

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

Da $G(a,b) = b$, gibt es wegen der Stetigkeit von F eine offene Umgebung $V_1 \subset W_1$ von a, so daß

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

Wir betrachten den Banachraum $\mathcal{C}_b(V_1,\mathbb{R}^m)$. Zu gegebenem $\gamma \in \mathcal{C}_b(V_1,\mathbb{R}^m)$, mit $\|\gamma\| = \sup_{x\in V_1}\|\gamma(x)\|$, 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. Wir zeigen nun: die Zuordnung $\gamma \mapsto \psi = \Phi(\gamma)$ definiert eine Abbildung Φ, welche die abgeschlossene Teilmenge

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

des Banachraums $\mathcal{C}_b(V_1,\mathbb{R}^m)$ auf sich selbst abbildet. Es gilt

$$\begin{aligned}
\|\psi - b\| &= \sup_{x\in V_1}|\psi(x) - b| = \sup_{x\in V_1}|G(x,\gamma(x)) - G(x,b) + G(x,b) - b| \\
&\le \sup_{x\in V_1}\underbrace{|G(x,\gamma(x)) - G(x,b)|}_{\text{Taylor}} + \underbrace{\sup_{x\in V_1}\|G(x,b) - b\|}_{\le\frac{r}{4}\text{ nach (4)}} \\
&\le \frac{r}{4} + \sup_{x\in V_1}\left\|\left(\frac{\partial G}{\partial y}\right)(x,b)\cdot(\gamma(x) - b) + \phi(\gamma(x) - b)\right\| \ .
\end{aligned}$$

Nach (3) gilt $\|\phi(\gamma(x)-b)\| \leq \frac{1}{4}\|\gamma(x)-b\| \leq \frac{r}{4}$ unter der Voraussetzung $\|\gamma-b\| \leq r$. Nach (1) und Definition der Operatornorm folgt

$$\Big\|\Big(\frac{\partial G}{\partial y}\Big)(x,b)\cdot(\gamma(x)-b)\Big\| \leq \Big\|\Big(\frac{\partial G}{\partial y}\Big)(x,b)\Big\|_{op}\|\gamma(x)-b\| \leq \frac{r}{2}$$

für alle $x \in V_1 \subset W_1$ und $|\gamma(x)-b| \leq r$. Damit ist $\|\psi - b\| \leq r$ falls $|\gamma - b\| \leq r$, und $\Psi : A \to A$ bewiesen. Schließlich gilt wieder nach Taylor

$$\begin{aligned}\|\Phi(\gamma_1)-\Phi(\gamma_2)\| &= \sup_{x\in V_1}\|G(x,\gamma_1(x))-G(x,\gamma_2(x))\| \\ &= \sup_{x\in V_1}\Big\|\Big(\frac{\partial G}{\partial y}\Big)(x,\gamma_2(x))\cdot(\gamma_1(x)-\gamma_2(x))+\phi(\gamma_1(x)-\gamma_2(x))\Big\|\end{aligned}$$

Wegen $\|\gamma_1(x)-\gamma_2(x)\| \leq \|\gamma_1(x)-b\|+\|b-\gamma_2(x)\| \leq 2r < \frac{5}{2}r$ gilt $\|\phi(\gamma_1(x)-\gamma_2(x)\| < \frac{1}{4}\|\gamma_1-\gamma_2\|$ nach (3), und nach Definition der Operatornorm und (1) ist

$$\Big\|\Big(\frac{\partial G}{\partial y}\Big)(x,\gamma_2(x))\cdot(\gamma_1(x)-\gamma_2(x))\Big\| \leq \frac{1}{2}\|\gamma_1-\gamma_2\|$$

Damit ist $\|\Phi(\gamma_1)-\Phi(\gamma_2)\| \leq \frac{3}{4}\|\gamma_1-\gamma_2\|$ für alle $\gamma_1,\gamma_2 \in A$ bewiesen, somit Φ eine Kontraktion in A. Nach dem Banachschen Fixpunktsatz gibt es genau einen Fixpunkt $\gamma_* \in \mathcal{C}_b(V_1,\mathbb{R}^m)$, und die eindeutige stetige Abbildung $g=\gamma_* : V_1 \to V_2 \subset \mathbb{R}^n$ löst die Gleichung $F(x,g(x))=0$ für alle $x \in V_1$. (Dabei ist $V_2=\{y\in\mathbb{R}^m \;:\; \|y-b\|\leq r\}$ die Kugel um b mit Radius r).

Nun zur partiellen 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 $g=(g_1,\dots,g_m) : V_1' \to \mathbb{R}^m$ 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 $x=(x_1,\dots,x_n) \in V_1'$, 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 2.6 ist g dann total differenzierbar. $\square$

Wichtig an diesem Beweis ist, daß er nicht nur Existenz und Eindeutigkeit der Lösung von $F(x, y(x)) = 0$ beweist, sondern auch ein konstruktives Verfahren angibt, mit der die Lösung beliebig genau approximiert werden kann. Dieses Verfahren kann insbesondere auf dem Computer implementiert werden.

Beispiel 5.3 Es sei $F(x, y) = x^2 + y^2 - 1$ mit einer Lösung $F(0, 1) = 0$. Natürlich sind diese Gleichung auch exakt zu lösen, soll aber zur Veranschaulichung über das Fixpunktverfahren approximiert werden. Es ist $B = (\partial_y F)(0, 1) = 2$, also $G(x, y) = y - \frac{1}{2}(x^2 + y^2 - 1)$. Ohne nähere Diskussion der Gebiete versuchen wir die Fixpunkt-Konstruktion

$$y_{n+1} = y_n - \frac{1}{2}(x^2 + y_n^2 - 1) , \qquad y_0 = 1 .$$

Die Konvergenz solcher rekursiv definierter Folgen hatten wir bereits im 1. Semester untersucht. Man zeigt, daß für $|x| \leq 1$ die Folge (y_n) monoton fallend und nach unten durch $1 - x^2$ beschränkt ist. Für $x = \frac{1}{2}$ sind die ersten Folgeglieder

$$y_0 = 1 , \quad y_1 = \frac{7}{8} , \quad y_2 = \frac{111}{128} , \quad y_3 = \frac{28383}{32768} = 0.866180\dots .$$

Zum Vergleich: $\sqrt{\frac{3}{4}} = 0.866025\dots.$ $\triangleleft$

Weitere Informationen über die Lösung einer implizit definierten Funktion kann man aus dem Taylor-Polynom gewinnen.

Beispiel 5.4 Durch $F(x, y) = \cos x + \sin y - xy - 1$ werde implizit eine Funktion $y(x)$ mit $y(0) = 0$ erklärt. Dann ist $(\partial_y F)(0, 0) = 1$, das Fixpunkt-Verfahren also anwendbar. Für die Lösung gilt

$$y'(x) = -\frac{(\partial_x F)(x, y(x))}{(\partial_y F)(x, y(x))} = \frac{\sin x + y(x)}{\cos(y(x)) - x}$$

und dann weiter nach Quotienten- und Kettenregel

$$\begin{aligned} y''(x) &= \frac{\cos x + y'(x)}{\cos(y(x)) - x} - \frac{\sin x + y(x)}{(\cos(y(x)) - x)^2}(-y'(x)\sin(y(x)) - 1) \\ &= \frac{\cos x}{\cos(y(x)) - x} + 2\frac{\sin x + y(x)}{(\cos(y(x)) - x)^2} + \sin(y(x))\frac{(\sin x + y(x))^2}{(\cos(y(x)) - x)^3} . \end{aligned}$$

Damit ist das Taylor-Polynom 2. Ordnung von $y(x)$ im Punkt $x = 0$ gegeben durch

$$(T_2 y)(x; 0) = \frac{1}{2}x^2 . \qquad \triangleleft$$

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 5.5 *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 5.6 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 5.5 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. Wir werden das an vielen Stellen benötigen, z.B. in der Theorie der Untermannigfaltigkeiten, die man lokal "geradebiegen" kann, oder bei der höherdimensionalen Verallgemeinerung der Substitutionsregel in Integralen.

Beispiel 5.7 (Kartesischen Koordinaten und Polarkoordinaten) Sei $f : \mathbb{R}_+^\times \times \mathbb{R} \to \mathbb{R}^2$ 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

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

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

6 Untermannigfaltigkeiten

Definition 6.1 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 6.2 (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 $\operatorname{rang}((Df)(x)) = \operatorname{rang}(2(x_1,\dots,x_{n+1})) = 1$ für alle $x \in U$. Somit ist

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

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

Definition 6.3 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.

Beispiel 6.4 Es sei $M = S^2 \subset \mathbb{R}^3$ die Untermannigfaltigkeit aus Beispiel 6.2 zu $f(x) = x_1^2+x_2^2+x_3^2-1$ und $a = (\cos\alpha\sin\beta, \sin\alpha\sin\beta, \cos\beta) \in S^2$, mit $\alpha \in [0, 2\pi]$ und $\beta \in [0, \pi]$. Dann ist $(Df)(a) = 2a \in M(1 \times 3, \mathbb{R})$. Auflösung des linearen Gleichungssystems ergibt nach sinnvoller Wahl der Skalierung

$$T_a S^2 = \operatorname{span}\left(\begin{pmatrix} \sin\alpha \\ -\cos\alpha \\ 0 \end{pmatrix}, \begin{pmatrix} \cos\alpha\cos\beta \\ \sin\alpha\cos\beta \\ -\sin\beta \end{pmatrix}\right) .$$

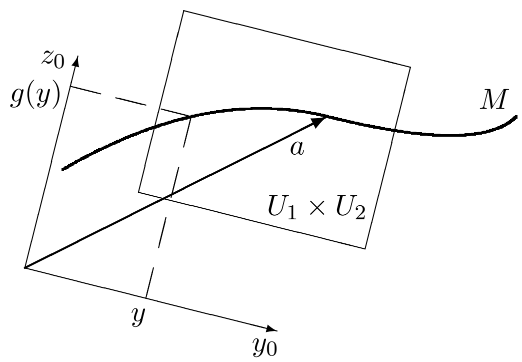
Sinnvollerweise haben wir die aufspannenden Vektoren als Orthonormalsystem gewählt. Schließlich wird der Normalenvektorraum durch den Radiusvektor aufgespannt, $N_a S^2 = \mathbb{R}a^t$. ◁

Definition 6.5 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 6.6 (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 6.7 (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 $1 \leq j \leq n$ und $a_{i,n+j}(y,z) = \frac{\partial f_i}{\partial z_j}(y,z)$ für $1 \leq j \leq k$. 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 $y \in U_1$. $\square$

Beispiel 6.8 Wir sehen uns diese Konstruktion für die S^2 aus Beispiel 6.4 an. Für $a, x \in S^2$ war $T_aS^2 = \operatorname{span}(u,v)$ und $N_aS^2 = \operatorname{span}(a)$ mit

$$a = \begin{pmatrix} \cos\alpha\sin\beta \\ \sin\alpha\sin\beta \\ \cos\beta \end{pmatrix}, \quad x = \begin{pmatrix} \cos\sigma\sin\tau \\ \sin\sigma\sin\tau \\ \cos\tau \end{pmatrix}, \quad u = \begin{pmatrix} \sin\alpha \\ -\cos\alpha \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} \cos\alpha\cos\beta \\ \sin\alpha\cos\beta \\ -\sin\beta \end{pmatrix}.$$

Die Vektoren (a, u, v) bilden ein Orthonormalsystem. Deshalb ist die Zerlegung $x = (y, z)$ entsprechend $\mathbb{R}^{n+k} = T_aS^2 \oplus N_aS^2$ die orthogonale Projektion

$$\begin{aligned} y &= P_{T_aS^2}(x) = \langle u, x\rangle u + \langle v, x\rangle v \\ &= \sin(\alpha - \sigma)\sin\tau \cdot u + \big(\sin\tau\cos\beta\cos(\alpha-\sigma) - \cos\tau\sin\beta\big)v =: y_1u + y_2v \end{aligned}$$

und

$$z = P_{N_aS^2}(x) = \langle a, x\rangle a = \big(\sin\tau\sin\beta\cos(\alpha-\sigma) - \cos\tau\cos\beta\big)a =: z_1a\,.$$

Die Aussage des Lemmas ist nun, daß es eine stetig differenzierbare Abbildung $g : U_1 \to U_2$, mit $U_1 \subset T_aS^2$ und $U_2 \subset N_aS^2$ gibt mit $f(y, g(y)) = 0$. Diese Abbildung ist $g(y) = z_1(y_1, y_2) = \sqrt{1 - y_1^2 - y_2^2}$. $\triangleleft$

Beweis von Satz 6.6. Nach Lemma 6.7 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. □

Beispiel 6.9 Wir diskutieren die Immersion $\phi : T \to V$ für das Beispiel 6.8. Aus dem geometrischen Bild der Orthogonalprojektion der Sphäre auf ihren Tangentialraum ist klar, daß wir $T = K_r(0) = \{(y_1, y_2) \in \mathbb{R}^2 \ : y_1^2 + y_2^2 < r^2\}$ für beliebiges $0 < r < 1$ wählen können. Die Aussage des Satzes ist nun, daß es eine differenzierbare bijektive Umrechnung ϕ mit differenzierbarem Inversen zwischen Punkten $(y_1, y_2) \in T$ und Punkten $x = (\alpha, \beta) \in V \subset S^2$ einer Teilmenge der Sphäre gibt. Diese Umkehrabbildung $\phi^{-1} : V \to T$ ist gegeben durch

$$\phi^{-1}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sin(\alpha-\sigma)\sin\tau \\ \sin\tau\cos\beta\cos(\alpha-\sigma) - \cos\tau\sin\beta \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} .$$

Sie ist nach Konstruktion umkehrbar. ◁

Es sei bemerkt, daß auch die Umkehrung von Satz 6.6 gilt: Ist eine Immersion ϕ mit diesen Eigenschaften gegeben, dann kann man zeigen, daß M Untermannigfaltigkeit ist. Die Immersion ϕ ist nicht eindeutig. Man kann die Konstruktion im Beweisen von Satz 6.6 mit beliebigen Diffeomorphismen von T deformieren, ohne daß sich an der Aussage etwas ändert.

Satz 6.10 *Sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit, lokal definiert durch die Abbildung $f = (f_1, \dots, f_k) : U \to \mathbb{R}^k$, mit $U \subset \mathbb{R}^{n+k}$ offen. Sei $a \in M \cap U$ ein Punkt und $\phi = (\phi_1, \dots, \phi_{n+k}) : T \to \mathbb{R}^{n+k}$ die nach Satz 6.6 existierende Immersion (z.B. konstruiert aus f), die $T \subset \mathbb{R}^n$ homöomorph auf eine Umgebung $V \subset M$ von $\phi(t_a) = a \in M$ abbildet. Dann gilt (unabhängig von der konkreten Wahl von ϕ):*

i) *Zu jedem Tangentialvektor $v \in T_a(M)$ im Sinne von Definition 6.5 gibt es eine Kurve $c :]-\epsilon, \epsilon[\to M$ durch $a = c(0)$, so daß v Tangentialvektor an c im Punkt a ist.*

ii) *Ist umgekehrt v ein Tangentialvektor an eine Kurve $c : I \to M$ 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 $c : I \to M$ eine Kurve auf M durch $c(0) = a$, wobei das Intervall $I =]-\epsilon, \epsilon[$ so gewählt sei, daß die Kurve $\gamma = \phi^{-1} \circ c : I \to T$ durch $t_a = \gamma(0) = \phi^{-1}(a)$ 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 $1 \leq j \leq n$.

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 $v = \sum_{i=1}^n \alpha_i v_i$. Zum entsprechenden Vektor $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ wählen wir die Kurven $\gamma :]-\epsilon, \epsilon[\to T$ mit $\gamma(\tau) = t_a + \tau\alpha$ und $c(\tau) = \phi(\gamma(\tau))$. Dann ist

$$\begin{aligned} c'(0) &= \big(D(\phi \circ \gamma)\big)(0) = (D\phi)(\gamma(0)) \cdot \gamma'(0) = \sum_{i=1}^n (\partial_i \phi)(t_a) \frac{d\gamma_i}{d\tau}\Big|_{\tau=0} \\ &= \sum_{i=1}^n (\partial_i \phi)(t_a) \alpha_i = v . \end{aligned}$$

ii) Sei $c : I \to M$ mit $c(\tau_a) = a$. Dann ist $(f \circ c)(\tau) = 0$ für alle $\tau \in I$, so daß das Differential verschwindet:

$$(D(f \circ c))(\tau_a) = (Df)(a) \cdot c'(\tau_a) = 0 \qquad \Rightarrow \qquad c'(\tau_a) \in T_a(M) .$$

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 6.11 (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

$$F(b) \leq F(a) \qquad \text{bzw. } F(b) \geq F(a) \qquad \text{für alle } b \in V .$$

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

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

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

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

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

Da $c'(0) \in T_a(M)$ als beliebiger Tangentialvektor gewählt werden kann, ist $(DF)(a) = (\operatorname{grad} F)(a) \in \big(T_a(M)\big)^\perp = N_a(M)$. □

Wir geben einige Anwendungen. Der folgende Satz liefert ein numerisch umsetzbares Verfahren, um den größten und kleinsten Eigenwert einer symmetrischen Matrix zu bestimmen:

Satz 6.12 *Für eine symmetrische Matrix $A = A^t \in M(n, \mathbb{R})$ nimmt die Einschränkung der durch $F(x) = \langle x, Ax\rangle$ definierten Funktion $F : \mathbb{R}^n \to \mathbb{R}$ auf die Einheitssphäre S^{n-1} ihr Maximum und Minimum in einem Eigenvektor an, und der Lagrange-Multiplikator ist der Eigenwert.*

Beweis. Die S^{n-1} ist Untermannigfaltigkeit im $\mathbb{R}^n$ definiert durch die Nebenbedingung $f(x) = \langle x, x\rangle - 1$. Als stetige Abbildung auf dem kompakten Raum S^{n-1} nimmt $F : S^{n-1} \to \mathbb{R}$ das Supremum in einem Punkt $v \in S^{n-1}$ an. In diesem Punkt gilt nach Satz 6.11

$$(\operatorname{grad} F)(v) = 2Av = \lambda(\operatorname{grad} f)(v) = 2\lambda v \,,$$

also $Av = \lambda v$ und dann $\langle v, Av\rangle = \lambda\langle v, v\rangle = \lambda$. Analog für das Infimum. □

Die Methode läßt sich fortsetzen, um iterativ sämtliche Eigenwerte und Eigenvektoren einer symmetrischen Matrix zu bestimmen. Sind $v_1, \dots, v_k$ Eigenvektoren zu $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$, dann betrachtet man iterativ die Einschränkung von F auf $M_k := S^{n-1} \cap \big(\operatorname{span}(v_1, \dots, v_k)\big)^\perp$, gegeben durch die Nebenbedingungen $F_{k+1}(x) = \langle x, x\rangle - 1$ und $F_j = 2\langle v_j, x\rangle$ für $1 \leq j \leq k$. Die Differentiale $(DF_{k+1})(x) = 2x^t$ und $(DF_j)(x) = 2v_j^t$ sind linear unabhängig. Der Durchschnitt abgeschlossener Mengen ist abgeschlossen, damit ist M_k kompakt, und F nimmt wieder ihr Supremum in einem Punkt $v_{k+1} \in M_k$ an. Somit gibt es $\mu_1, \dots, \mu_k, \lambda_{k+1}$ mit

$$(\operatorname{grad} F)(v_{k+1}) = 2Av_{k+1} = 2\lambda_{k+1}v_{k+1} + 2\sum_{j=1}^{k} \mu_j v_j \,.$$

Skalarprodukt mit v_{k+1} liefert unter Verwendung von $\langle v_j, v_{k+1}\rangle = \delta_{j,k+1}$ wie zuvor $\lambda_{k+1} = F(v_{k+1})$. Insbesondere ist so bewiesen, daß jede symmetrische Matrix $A = A^t \in M(n, \mathbb{R})$ diagonalisierbar ist und daß es im $\mathbb{R}^n$ eine Orthonormalbasis aus Eigenvektoren von A gibt.

Satz 6.13 (Arithmetisches $\geq$ geometrisches Mittel)
Für beliebige $a_1, \ldots, a_n \geq 0$ gilt

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \cdot a_2 \cdots a_n}$$

mit Gleichheit genau für $a_1 = \cdots = a_n$.

Beweis. i) Es sei M die $n-1$-dimensionale Untermannigfaltigkeit

$$M = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \;:\; x_i > 0\;,\; f(x) := x_1 + \cdots + x_n - 1 = 0\}\;,$$

betrachtet als Durchschnitt von $(\mathbb{R}_+^\times)^n$ mit $f^{-1}(0)$. Dann hat $(Df)(x) = (1, \ldots, 1)$ maximalen Rang. Wir betrachten die Einschränkung der Funktion $F(x) = x_1 x_2 \cdots x_n$ auf den kompakten Abschluß $\overline{M}$ von M. Nach dem Satz vom Maximum nimmt F ihr Supremum in einem Punkt $y \in \overline{M}$ an, und da $F(x) = 0$ für $x \in \overline{M} \setminus M$, gilt sogar $y \in M$. Somit gibt es ein λ mit

$$(\partial_i F)(y) = \frac{y_1 \cdots y_n}{y_i} = \lambda (\partial_i f)(y) = \lambda \qquad \underline{\text{für alle}}\ i = 1, \ldots n\;,$$

also $y_1 = y_2 = \cdots = y_n = \frac{1}{n}$ und damit $x_1 \cdot x_2 \cdots x_n \leq \frac{1}{n^n}$ für alle $x \in M$ mit Gleichheit genau für $x = y$.

ii) Seien nun $a_1, \ldots, a_n > 0$ und $a := a_1 + \cdots + a_n$. Wir setzen $x_i = \frac{a_i}{a}$, dann gilt nach i)

$$x_1 \cdot x_2 \cdots x_n = \frac{a_1 \cdot a_2 \cdots a_n}{a^n} \leq \frac{1}{n^n}\;.$$

Elementare Umformungen liefern die Behauptung. Für Randpunkte mit $a_i = 0$ ist die Ungleichung offensichtlich erfüllt. □

Durch ähnliche Techniken erhält man einen weiteren Beweis der Hölderschen 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, $\mathrm{rang}(a_{ij}(x)) = k$ 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 6.6 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 6.14 Eine eindimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^2$ werde durch $f(x_1, x_2) = \frac{1}{2}(l^2 - x_1^2 - x_2^2) = 0$ definiert. Das ist relevant für die Dynamik eines ebenen Pendels, bestehend aus einem Massenpunkt aufgehängt an einem masselosen Seil der Länge l. Es gilt

$$(Df)(x) = (-x_1,\ -x_2)\ , \qquad \operatorname{rang}((Df)(x)) = 1 \qquad \text{für } x_1^2 + x_2^2 \neq 0\ .$$

Damit bildet $n(x) := (\operatorname{grad} f)(x) = (-x_1, -x_2)$ eine Basis von $N_{(x_1,x_2)}(M)$ und $v(x) = (-x_2, x_1)$ eine Basis von $T_{(x_1,x_2)}(M)$. Auf den Massepunkt wirke die Kraft $(K_1, K_2) = (0, -mg)$. Das d'Alembertsche Prinzip liefert

$$\Big(m\ddot{x}_1,\ m\ddot{x}_2 + mg\Big) = \lambda(x)\big(-x_1, -x_2\big)\ . \tag{*}$$

Zusammen mit der Gleichung $f = 0$ haben wir 3 Gleichungen zur Bestimmung der 3 Funktionen x_1, x_2, λ. Skalarprodukt von (*) mit $(\dot{x}_1, \dot{x}_2)$ und Verwendung von $x_1^2 + x_3^2 = l^2$, also $x_1\dot{x}_1 + x_2\dot{x}_2 = 0$, liefert den Energieerhaltungssatz

$$E = \frac{m}{2}((\dot{x}_1)^2 + (\dot{x}_2)^2) + mgx_2 = \text{const}\ .$$

Skalarprodukt von (*) mit (x_1, x_2) und Verwendung von $x_1\ddot{x}_1 + x_2\ddot{x}_2 + (\dot{x}_1)^2 + (\dot{x}_2)^2 = 0$ liefert

$$\lambda = \frac{m}{l^2}\big((\dot{x}_1)^2 + (\dot{x}_2)^2 - gx_2\big) = \frac{1}{l^2}(2E - 3mgx_2)\ .$$

Die Zwangskraft $Z = (-\lambda x_1, -\lambda x_2)$ ist dann die Seilspannung. Die Gleichungen lassen sich in Polarkoordinaten $x_1 = l\sin\phi$ und $x_2 = -l\cos\phi$ 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 $v^2 := \|\dot{x}\|^2 = l^2\dot{\phi}^2$. ◁

In der Statistischen Physik spielt die Boltzmann-Verteilung eine wichtige Rolle, die sich ebenfalls aus einem Extremalproblem mit Nebenbedingung ergibt:

Beispiel 6.15 (Boltzmann-Verteilung) Ein System besitze N Energieniveaus $E_1, \ldots, E_N$, wobei das Niveau E_i mit Wahrscheinlichkeit p_i angenommen wird. Die Gesamtenergie ist dann $U = \sum_{i=1}^N E_i p_i$ unter der Nebenbedingung $f(p_i) = \sum_{i=1}^N p_i - 1 = 0$. Der Wahrscheinlichkeitsverteilung wird eine Entropie $S = -k_B \sum_{i=1}^N p_i \ln p_i$ zugeordnet, wobei k_B die Boltzmann-Konstante ist. Es wird nun jene Verteilung realisiert, die die freie Energie $F = U - TS$ minimiert, wobei T die Temperatur ist. Somit gilt

$$\frac{\partial F}{\partial p_i} = E_i + k_B T(1 + \ln p_i) = \lambda \frac{\partial f}{\partial p_i} = \lambda 1 \;,$$

oder $p_i = \exp\left(\frac{\lambda}{k_B T} - 1 - \frac{E_i}{k_B T}\right)$. Die Nebenbedingung $\sum_{i=1}^N p_i = 1$ erzwingt

$$\exp\left(\frac{\lambda}{k_B T} - 1\right) = \frac{1}{Z} \qquad \text{mit} \qquad Z = \sum_{i=1}^N e^{-\frac{E_i}{k_B T}}$$

Dabei heißt Z die Zustandssumme, und insgesamt folgt $p_i = \frac{1}{Z} e^{-\frac{E_i}{k_B T}}$. $\triangleleft$

7 Differenzierbare Kurven

Eine eindimensionale Untermannigfaltigkeit im $\mathbb{R}^n$ ist eine differenzierbare Kurve mit nicht verschwindendem Tangentialvektor, denn die zugehörige Immersion $\phi = c : I \to \mathbb{R}^n$ bildet ein offenes Intervall $I \subset \mathbb{R}$ homöomorph auf die Teilmenge $M = c(I) \subset \mathbb{R}^n$ ab. Diese Teilmenge heißt die *Spur* der Kurve c. Die Umkehrung gilt nicht immer: Wenn die Spur der Kurve sich selbst schneidet, ist c nicht bijektiv. Die Rangbedingung ist verletzt für $c'(t) = 0$.

Definition 7.1 Eine differenzierbare Kurve $c : I \to \mathbb{R}^n$ heißt *regulär*, falls $c'(t) \neq 0$.

Im folgenden beschränken wir uns auf reguläre Kurven. Selbstschnittpunkte der Spur sind zugelassen.

Satz 7.2 *Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $c : I \to \mathbb{R}^n$ eine reguläre stetig differenzierbare Kurve mit $c(t) = (c_1(t), \ldots, c_n(t))$ und $c_1' \neq 0$ auf I. Dann gilt:*

i) *$c_1 : I \to J$ ist eine Diffeomorphismus zwischen offenen Intervallen.*

ii) *Es sei $M = c(I)$ die Spur der Kurve. Auf der Teilmenge $U = J \times \mathbb{R}^{n-1} \subset \mathbb{R}^n$ existiert eine differenzierbare Abbildung $f : U \to \mathbb{R}^{n-1}$ mit $\operatorname{rang}((Df)(x)) = n - 1$ und $M \cap U = f^{-1}(0)$. Insbesondere ist M eine Untermannigfaltigkeit.*

Beweis. Nach Voraussetzung ist $c_1 : I \to J$ ein Diffeomorphismus mit Umkehrung $c_1^{-1} : J \to I$. Setze $f = (f_2, \ldots, f_n)$ mit $f_j = x_j - c_j(c_1^{-1}(x_1))$. Dann ist das Differential nach Kettenregel gegeben in Blockdarstellung durch $(Df)(x) = (v(x_1), E_{n-1}) \in M((n-1) \times n, \mathbb{R})$ mit $v_i(x_1) = c'_{i+1}(c_1^{-1}(x_1)) \cdot \big(c'_1(c_1^{-1}(x_1))\big)^{-1}$, hat somit maximalen Rang. Zusammen mit $t = c_1^{-1}(x_1)$ definieren die Nullstellen $x_j = c_j(t)$ die Spur. $\square$

Offenbar verhindert die Bedingung $c'_1 \neq 0$ Selbstschnittpunkte, denn c_1 bleibt monoton wachsend bzw. fallend.

Es sei $c : I \to \mathbb{R}^n$ eine stetige Kurve und $T = \{t_0, \ldots, t_m\}$ eine endliche Teilmenge von I mit $t_0 < t_1 < \ldots, t_m$. Dann definieren die zugehörigen Kurvenpunkte $c(t_1), \ldots, c(t_m)$ ein der Kurve einbeschriebendes *Sehnenpolygon* der Länge

$$L(c(t_1), \ldots, c(t_m)) = \sum_{i=1}^{m} \|c(t_i) - c(t_{i-1})\| \,.$$

Nach Dreiecksungleichung wächst die Länge des Sehnenpolygons, wenn weitere Kurvenpunkte hinzugefügt werden.

Definition 7.3 Es sei $c : I \to \mathbb{R}^n$ eine stetige Kurve. Falls die Menge der Längen von einbeschriebenen Sehnenpolynomen beschränkt ist, dann heißt das Supremum dieser Menge die *Bogenlänge der Kurve* c.

Satz 7.4 *Eine Kurve $c : [\alpha, \beta] \to \mathbb{R}$, die auf $]\alpha, \beta[$ stetig differenzierbar ist, hat die Bogenlänge*

$$L(c) = \int_\alpha^\beta dt\, \|c'(t)\| = \int_\alpha^\beta dt\, \sqrt{\langle c'(t), c'(t)\rangle} \,.$$

Beweis. i) Nach dem Mittelwertsatz 2.13 gilt mit $f = \mathrm{id} : \mathbb{R}^n \to \mathbb{R}^n$

$$c(t_i) - c(t_{i-1}) = \int_{t_{i-1}}^{t_i} dt\, c'(t)$$

und dann mit der an diesen Satz dort anschließenden Abschätzung

$$\|c(t_i) - c(t_{i-1})\| \leq \int_{t_{i-1}}^{t_i} dt\, \|c'(t)\| \,.$$

Somit gilt für jede Unterteilung des Intervalls $[\alpha, \beta]$ die Abschätzung $L(c(t_0), \ldots, c(t_m)) \leq L(c)$.

ii) Wir zeigen: Zu jedem $\epsilon > 0$ gibt es eine Unterteilung mit $L(c) - L(c(t_0), \ldots, c(t_m)) < \epsilon$. In der Standardbasis (e_j) des $\mathbb{R}^n$ sei $c(t) = \sum_{j=1}^n c_j(t) e_j$. Da die $c'_j(t)$ als stetige Funktionen Riemann-integrierbar sind, gibt es Treppenfunktionen $\phi_j : [\alpha, \beta] \to \mathbb{R}$ mit $|c'_j(t) - \phi_j(t)| < \frac{\epsilon}{2n|\beta-\alpha|}$ für alle $t \in [\alpha, \beta]$. Mit

$\phi(t) = \sum_{j=1}^n \phi_j(t)e_j$ folgt nach Dreiecksungleichung $\|c'(t) - \phi(t)\| \leq \frac{\epsilon}{2|\beta-\alpha|}$ für alle $t \in [\alpha, \beta]$. Wir wählen eine Unterteilung $\alpha = t_0 < t_1 < \cdots < t_m = \beta$, so daß ϕ auf jedem offenen Teilintervall $]t_{i-1}, t_i[$ konstant ist. Nach Dreiecksungleichung gilt

$$\begin{aligned}\Big\| \int_{t_{i-1}}^{t_i} dt\ \phi(t) \Big\| &\leq \Big\| \int_{t_{i-1}}^{t_i} dt\ (\phi(t) - c'(t)) \Big\| + \Big\| \int_{t_{i-1}}^{t_i} dt\ c'(t) \Big\| \\ &\leq \int_{t_{i-1}}^{t_i} dt\ \|\phi(t) - c'(t)\| + \Big\| \int_{t_{i-1}}^{t_i} dt\ c'(t) \Big\| \\ &\leq \frac{\epsilon}{2|\beta-\alpha|}(t_i - t_{i-1}) + \big\|c(t_i) - c(t_{i-1})\big\| .\end{aligned}$$

Da ϕ auf $]t_{i-1}, t_i[$ konstant ist, gilt $\Big\| \int_{t_{i-1}}^{t_i} dt\ \phi(t) \Big\| = \int_{t_{i-1}}^{t_i} dt\ \|\phi(t)\|$ und somit nach Summieren über i

$$\sum_{i=1}^m \big\|c(t_i) - c(t_{i-1})\big\| \geq \int_\alpha^\beta dt\ \|\phi(t)\| - \frac{\epsilon}{2} .$$

Andererseits ist $\|c'(t)\| \leq \|c'(t) - \phi(t)\| + \|\phi(t)\|$, also $\|\phi(t)\| \geq \|c'(t)\| - \frac{\epsilon}{2|\beta-\alpha|}$. Integrieren über $[\alpha, \beta]$ liefert dann die Behauptung $\sum_{i=1}^m \big\|c(t_i) - c(t_{i-1})\big\| \geq \int_\alpha^\beta dt\ \|c'(t)\| - \epsilon$. □

Beispiel 7.5 Der ebene Kreisbogen $c : [\phi_1, \phi_2] \to \mathbb{R}^2$ mit $c(t) = (r\cos t, r\sin t)$ hat die Bogenlänge

$$L(c) = \int_{\phi_1}^{\phi_2} dt\ \sqrt{r^2\sin^2 t + r^2\cos^2 t} = r(\phi_2 - \phi_1) .$$

Insbesondere hat der Kreis vom Radius r, realisiert durch die Kurve $c : [0, 2\pi[\to \mathbb{R}^2$ mit $c(t) = (r\cos t, r\sin t)$ den Umfang $L(c) = 2\pi r$. ◁

Beispiel 7.6 Für $0 < b < a$ werde die Ellipsenkurve $c : [0, 2\pi[\to \mathbb{R}^2$ definiert durch $c(t) = (a\cos t, b\sin t)$. Ausgedrückt durch $\epsilon^2 := \frac{a^2-b^2}{a^2}$ hat sie die Bogenlänge

$$L(c) = \int_0^{2\pi} dt\ \sqrt{a^2\cos^2 t + b^2\sin^2 t} = a\int_0^{2\pi} dt\ \sqrt{1 - \epsilon^2\cos^2 t} = a \cdot E(2\pi; \epsilon) .$$

Das verbleibende sogenannte *elliptische Integral* kann nicht mehr elementar gelöst werden. Für den Viertelbogen findet man durch Potenzreihenentwicklung und

Beispiel 29.3 aus dem 2. Semester

$$
\begin{aligned}
E(2\pi;\epsilon) &= 4\sum_{k=0}^{\infty}\binom{\frac{1}{2}}{k}(-\epsilon^2)^k\int_0^{\frac{\pi}{2}} dt\ \sin^{2k} t \\
&= 2\pi\Big(1+\sum_{k=1}^{\infty}\epsilon^{2k}\frac{(-1)\cdot 1\cdot 3\cdots(2k-3)}{2^k k!}\cdot\frac{1\cdot 3\cdots(2k-3)(2k-1)}{2^k k!}\Big) \\
&= 2\pi\Big(1-\sum_{k=1}^{\infty}\frac{\epsilon^{2k}}{2k-1}\Big(\frac{\prod_{j=1}^{k}(2j-1)}{2^k k!}\Big)^2\Big)\ . \qquad \triangleleft
\end{aligned}
$$

Für ebene Kurven $c :]\alpha,\beta[\to \mathbb{R}^2$ ist es oft zweckmäßig, sie in Polarkoordinaten $c(t) = (r(t)\cos\phi(t), r(t)\sin\phi(t))$ darzustellen. Dann ist

$$c'(t) = \big(r'(t)\cos\phi(t) - r(t)\phi'(t)\sin\phi(t),\ r'(t)\sin\phi(t) + r(t)\phi'(t)\cos\phi(t)\big)$$

und folglich

$$L(c) = \int_{\alpha}^{\beta} dt\ \sqrt{\big(r'(t)\big)^2 + \big(r(t)\big)^2\big(\phi'(t)\big)^2}\ .$$

Entscheidend bei der Interpretation der Bogenlänge ist die Invarianz unter Reparametrisierung:

Satz 7.7 *Es sei $I \subset \mathbb{R}$ ein offenes Intervall, $c : I \to \mathbb{R}^n$ eine reguläre differenzierbare Kurve und $\phi^{-1} : I \to J$ ein Diffeomorphismus. Dann ist auch $\tilde{c} := c\circ\phi : J \to \mathbb{R}$ eine reguläre differenzierbare Kurve, und es gilt $L(c) = L(\tilde{c})$.*

Beweis. Sei $I =]\alpha,\beta[$. Der Diffeomorphismus ϕ ist insbesondere monoton. Sei zunächst $(\phi^{-1})'(t) > 0$ für alle $t \in I$, dann ist $J =]\phi^{-1}(\alpha),\phi^{-1}(\beta)[$. Differenzierbarkeit von $c\circ\phi(\tau)$ folgt aus der Kettenregel, mit $\tilde{c}'(\tau) = c'(\phi(\tau))\cdot\phi'(\tau)$. Damit gilt für die Bogenlänge unter Verwendung der Substitutionsregel für das Riemann-Integral

$$
\begin{aligned}
L(\tilde{c}) &= \int_{\phi^{-1}(\alpha)}^{\phi^{-1}(\beta)} d\tau\ \big\|\tilde{c}'(\tau)\big\| = \int_{\phi^{-1}(\alpha)}^{\phi^{-1}(\beta)} d\tau\ \big\|c'(\phi(\tau))\big\|\cdot|\phi(\tau)| \\
&= \int_{\phi^{-1}(\alpha)}^{\phi^{-1}(\beta)} d\tau\ \big\|c'(\phi(\tau))\big\|\cdot\phi(\tau) = \int_{\phi(\phi^{-1}(\alpha))}^{\phi(\phi^{-1}(\beta))} d\phi\ \big\|c'(\phi)\big\| = \int_{\alpha}^{\beta} d\phi\ \big\|c'(\phi)\big\|\ .
\end{aligned}
$$

Ist $(\phi^{-1})' < 0$ auf I, dann ist $J =]\phi^{-1}(\beta),\phi^{-1}(\alpha)[$. Das dann entstehende Vorzeichen beim Vertauschen der Integrationsgrenzen wird kompensiert durch das Vorzeichen in $|\phi(\tau)| = -\phi(\tau)$, so daß die Bogenlänge unabhängig vom Vorzeichen von ϕ' ist. □

Folglich ist die Bogenlänge einer Kurve allein eine Eigenschaft der Spur der Kurve. Die konkrete Parametrisierung ist unerheblich und kann möglichst sinnvoll gewählt werden. Oft bringt die *Parametrisierung nach Bogenlänge* Vorteile:

Definition 7.8 Eine reguläre differenzierbare Kurve $c : I \to \mathbb{R}^n$ heißt *nach Bogenlänge parametrisiert*, wenn $\|c'(t)\| = 1$ für alle $t \in I$.

Satz 7.9 *Zu jeder regulären stetig differenzierbaren Kurve $c : I \to \mathbb{R}^n$ gibt es eine diffeomorphe Reparametrisierung $\phi^{-1} : I \to J$, so daß $\tilde{c} := c \circ \phi : J \to \mathbb{R}^n$ nach Bogenlänge parametrisiert ist.*

Beweis. Es sei $I =]\alpha, \beta[$. Wir definieren

$$\phi^{-1}(\tau) := \int_\alpha^\tau dt\ \|c'(t)\| \ .$$

Dann ist ϕ^{-1} Stammfunktion zu $\|c'(t)\|$, d.h. $(\phi^{-1})'(\tau) = \|c'(t)\|$ und somit

$$\tilde{c}'(\phi^{-1}(\tau)) = (D(c \circ \phi))(\phi^{-1}(\tau)) = c'(\tau) \cdot \phi'(\phi^{-1}(\tau)) = c'(\tau) \cdot \frac{1}{(\phi^{-1})'(\tau)} = \frac{c'(\tau)}{\|c'(\tau)\|} \ .$$

Folglich ist $\|\tilde{c}'(\phi^{-1}(\tau))\| = 1$. □

8 Variationsrechnung

In der Variationsrechnung sucht man Extremwerte (genauer: stationäre Werte) von Funktionen (genauer: Funktionalen) auf unendlich-dimensionalen normierten Vektorräumen. Es ist die wichtigste Methode zur Gewinnung von Bewegungsgleichungen in Mechanik und Feldtheorie. Wir nehmen vereinfachend Fréchet-Differenzierbarkeit an, tatsächlich genügt das unendlich-dimensionale Analogon der Richtungsableitung.

Abkürzend bezeichnen wir mit $X^k = \mathcal{C}^k([\alpha, \beta], \mathbb{R}^n)$ den (unendlich-dimensionalen) Vektorraum der k-mal stetig differenzierbaren Kurven $c : [\alpha, \beta] \to \mathbb{R}^n$. Man kann zeigen, daß X^k zusammen mit der Norm

$$\|c\|_{(k)} = \sum_{l=0}^{k} \sup_{t \in [\alpha,\beta]} \|c^{(l)}(t)\|$$

ein Banach-Raum ist. Das *Wirkungsfunktional* ist eine Abbildung $S : X^2 \to \mathbb{R}$. Wir interessieren uns für *stationäre* Punkte von S, d.h. Kurven $c \in X^2$ mit $(DS)(c) = 0$, wobei wir die Randpunkten der Kurve $a = c(\alpha)$ und $c = c(\beta)$ festgehalten.

In wichtigen Fällen läßt sich die Wirkung S schreiben als Riemann-Integral über die *Lagrange-Funktion* $\mathcal{L}$ der Kurve,

$$S(c) = \int_\alpha^\beta dt\ \mathcal{L}(t, c(t), c'(t)) \ .$$

Dabei wird $\mathcal{L} : [\alpha,\beta] \times X^2 \times X^1 \to \mathbb{R}$ als zweimal stetig differenzierbar vorausgesetzt. Betrachten wir $(t,q,v) \in [\alpha,\beta] \times X^2 \times X^1 \to \mathbb{R}$ als Variablen, dann ist

$$\begin{aligned}
&\mathcal{L}(t+\tau, q+r, v+w) \\
&= \mathcal{L}(t,q,v) + \Big(\frac{\partial \mathcal{L}}{\partial t}\Big)(t,q,v) \circ \tau + \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t,q,v) \circ r + \Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t,q,v) \circ w \\
&+ o(\|(t,r,w)\|)
\end{aligned}$$

mit $\tau \in \mathbb{R}$, $r \in X^2$ und $w \in X^1$ derart, daß $(t+\tau, q+r, v+w)$ in der gewählten Umgebung von (t,q,v) bleibt. Zweimalige Differenzierbarkeit von $\mathcal{L}$ bedeutet, daß die Abbildungen

$$\begin{aligned}
\Big(\frac{\partial \mathcal{L}}{\partial t}\Big) &: \mathbb{R} \times X^2 \times X^1 \to \mathrm{Hom}(\mathbb{R},\mathbb{R}) \;, \\
\Big(\frac{\partial \mathcal{L}}{\partial q}\Big) &: \mathbb{R} \times X^2 \times X^1 \to \mathrm{Hom}(X^2,\mathbb{R}) \;, \\
\Big(\frac{\partial \mathcal{L}}{\partial v}\Big) &: \mathbb{R} \times X^2 \times X^1 \to \mathrm{Hom}(X^1,\mathbb{R})
\end{aligned}$$

selbst wieder differenzierbar sind in (t,q,v) .

Es sei $U \subset X^2$ die Teilmenge der zweimal stetig differenzierbaren Kurven $c : [\alpha,\beta] \to \mathbb{R}^n$ mit festgehaltenen Randpunkten $a = c(\alpha)$ und $c = c(\beta)$. Um das Differential zu berechnen, ist $S(c+\xi)$ mit $S(c)$ zu vergleichen. Dabei ist $\xi \in X^2$ so zu wählen, daß $c+\xi \in U$ liegt, d.h. ξ verschwindet an den Randpunkten. Wir bezeichnen mit

$$X_0^2 := \{\xi \in \mathcal{C}^k([\alpha,\beta],\mathbb{R}^n)\;,\;\xi(\alpha) = \xi(\beta) = 0\}$$

den Untervektorraum der zweimal stetig differenzierbarten Kurven, die am Rand verschwinden. Dann gilt für $\xi \in X_0^2$ für eine zweimal stetig differenzierbare Lagrange-Funktion

$$\begin{aligned}
S(c+\xi) &= \int_\alpha^\beta dt\; \mathcal{L}(t, c(t)+\xi(t), c'(t)+\xi'(t)) \\
&= \int_\alpha^\beta dt\; \Big\{\mathcal{L}(t,c(t),c'(t)) + \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t,c(t),c'(t)) \circ \xi(t) \\
&\qquad + \Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t,c(t),c'(t)) \circ \xi'(t) + o(\|\xi\|) + o(\|\xi'\|)\Big\} \;.
\end{aligned}$$

Es gilt $\int_a^b dt\; (o(\|\xi\|) + o(\|\xi'\|)) = o(\|\xi\|_{(2)})$. Für die festgehaltene Kurve $c \in X^2$ können wir das Differential $\big(\frac{\partial \mathcal{L}}{\partial v}\big)(t,c(t),c'(t))$ auffassen als Abbildung

$$\mathbb{R} \ni t \mapsto \Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t,c(t),c'(t)) \in \mathrm{Hom}\big(X^1,\mathbb{R}\big)$$

Nach Kettenregel und allen Voraussetzungen ist diese Funktion stetig differenzierbar auf $]\alpha, \beta[$ mit stetiger Fortsetzung auf $[\alpha, \beta]$, so daß wir partiell integrieren dürfen:

$$\begin{aligned}(DS)(c) \circ \xi &= \int_\alpha^\beta dt \,\Big\{ \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t, c(t), c'(t)) \circ \xi(t) + \Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t, c(t), c'(t)) \circ \xi'(t) \Big\} \\ &= \int_\alpha^\beta dt \,\Big\{ \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t, c(t), c'(t)) - \Big(\Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t, c(t), c'(t))\Big)' \Big\} \circ \xi(t) \\ &\quad + \Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t, c(t), c'(t)) \circ \xi(t) \Big|_\alpha^\beta .\end{aligned}$$

Der Randbeitrag in der letzten Zeile verschwindet wegen $\xi(\alpha) = \xi(\beta) = 0$. Unter Verwendung des weiter unten angegebenen *Fundamentallemmas der Variationsrechnung* haben wir somit bewiesen:

Satz 8.1 (Euler-Lagrange-Gleichungen) *Über eine Lagrange-Funktion* $\mathcal{L} : [\alpha, \beta] \times X^2 \times X^1 \to \mathbb{R}$ *werde durch*

$$S(c) := \int_\alpha^\beta dt \; \mathcal{L}(t, c(t), c'(t))$$

ein Wirkungsfunktional $S : X^2 \to \mathbb{R}$ *auf dem Banach-Raum* X^2 *der zweimal stetig differenzierbaren Kurven* $c : [a, b] \to \mathbb{R}^n$ *definiert. Dann gilt: Das Wirkungsfunktional ist auf der Teilmenge* $U \subset X^2$ *der zweimal stetig differenzierbaren Kurven mit festgehaltenen Randpunkten genau dann stationär in einem Punkt* $c \in U$, *d.h.* $(DS)(c) = 0$, *wenn für die Kurve* c *die Euler-Lagrange-Gleichungen*

$$\frac{d}{dt}\Big(\Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t, c(t), c'(t))\Big) - \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t, c(t), c'(t)) = 0$$

in jedem Punkt $t \in [\alpha, \beta]$ *gelten.*

Nach obiger Diskussion ist nur die Richtung $(DS)(c) = 0 \quad \Rightarrow$ Euler-Lagrange zu zeigen. Sie ergibt sich aus

Lemma 8.2 (Fundamentallemma der Variationsrechnung) *Für eine stetige Funktion* $f \in \mathcal{C}([\alpha, \beta])$ *gelte* $\displaystyle\int_\alpha^\beta f(t)g(t) = 0$ *für alle zweimal stetig differenzierbaren Funktionen* $g \in \mathcal{C}^2_0([\alpha, \beta], \mathbb{R})$, *welche am Rand verschwinden. Dann ist* $f(t) = 0$ *für alle* $t \in [\alpha, \beta]$.

Beweis. Wegen der Stetigkeit von f genügt es, $f(t) = 0$ für alle $t \in]\alpha, \beta[$ zu zeigen. Angenommen, es gibt ein $t_0 \in]\alpha, \beta[$ 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 $t \in K_\delta(t_0) \subset \;]\alpha, \beta[$. Dann gibt es (sogar beliebig oft) differenzierbare nichtnegative Funktionen g, die außerhalb $K_\delta(t_0)$ identisch verschwinden. Eine Wahl ist z.B.

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

Damit ist

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

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

In den Euler-Lagrange-Gleichungen ist zunächst

$$\frac{d}{dt}\Big(\Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t, c(t), c'(t))\Big) - \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t, c(t), c'(t)) \in \mathrm{Hom}(X^2, \mathbb{R})$$

eine lineare Abbildung. Evaluiert im Punkt $t \in\,]\alpha, \beta[$ geht aber nur der Vektor $\xi(t) \in \mathbb{R}^n$ ein, so daß es genügt, die Gleichungen in den Vektoren der Standardbasis des $\mathbb{R}^n$ auszuwerten. Somit sind die Euler-Lagrange-Gleichungen äquivalent zum System

$$\frac{d}{dt}\Big(\Big(\frac{\partial \mathcal{L}}{\partial v}\Big)(t, c(t), c'(t)) \circ e_k\Big) - \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t, c(t), c'(t)) \circ e_k = 0$$
$$\text{für alle } k = 1, \dots, n \text{ und } t \in\,]\alpha, \beta[\ .$$

Beispiel 8.3 Wir betrachten noch einmal das Pendel der Fadenlänge l aus Beispiel 6.14. Die Nebenbedingung definiert eine Untermannigfaltigkeit $M = S^1 \subset \mathbb{R}^2$, auf der die Bewegung stattfindet. Die Lagrange-Funktion ist dann auf einer lokalen Karte (T, ϕ) dieser Untermannigfaltigkeit definiert. Sinnvollerweise wählt man $T \subset \mathbb{R}$ als abgerollten Kreisbogen und dann die Kurve $[\alpha, \beta] \ni t \mapsto c(t) \in T \subset \mathbb{R}$ als Zuodnung des Auslenkungswinkels zum Zeitpunkt t. Dann sei die Lagrange-Funktion gegeben als

$$\mathcal{L}(t, c(t), c'(t)) = \frac{ml^2}{2}\big(c'(t)\big)^2 + mgl\cos\big(c(t)\big)\ .$$

Nach Definition der Einschränkungen der Jacobi-Matrix ist

$$\begin{aligned}\mathcal{L}(t, c(t) + r, c'(t)) &= \mathcal{L}(t, c(t), c'(t)) + mgl\big(\cos(c(t) + r) - \cos(c(t))\big) \\ &= \mathcal{L}((t, c(t), c'(t)) + mgl\big(\cos(c(t))(\cos r - 1) - \sin(c(t))\sin r\big) \\ &= \mathcal{L}(t, c(t), c'(t)) + \Big(\frac{\partial \mathcal{L}}{\partial q}\Big)(t, c(t), c'(t)) \circ r + o(\|r\|)\end{aligned}$$

mit $\left(\frac{\partial\mathcal{L}}{\partial q}\right)(t,c(t),c'(t)) = -mgl\sin(c(t))$ und

$$\begin{aligned}
\mathcal{L}(t,c(t),c'(t)+w) &= \frac{ml^2}{2}\big(c'(t)+w\big)^2 + mgl\cos(c(t)) \\
&= \mathcal{L}(t,c(t),c'(t)) + ml^2c'(t)w + \frac{m}{2}w^2 \\
&= \mathcal{L}(t,c(t),c'(t)) + \Big(\frac{\partial\mathcal{L}}{\partial v}\Big)(t,c(t),c'(t))\circ w + o(\|w\|)
\end{aligned}$$

mit $\left(\frac{\partial\mathcal{L}}{\partial v}\right)(t,c(t),c'(t)) = mc'(t)$. Somit lautet die Euler-Lagrange-Gleichung

$$\big(ml^2c'(t)\big)' - (-mgl\sin(c(t))) = 0 \qquad \Rightarrow \qquad c''(t) + \omega^2\sin(c(t)) = 0$$

mit $\omega = \sqrt{\frac{g}{l}}$. Das ist eine Differentialgleichung 2. Ordnung zur Lösung der Integralkurve $c(t)$. Wir werden im nächsten Kapitel sehen, daß zu ihrer eindeutigen Lösung Anfangsbedingungen $c(t_0)$ und $c'(t_0)$ vorzugeben sind. $\triangleleft$

Diese Methode, sich die Bahnkurven eines mechanischen Systems zu verschaffen, heißt *Hamiltonsches Prinzip*: Von allen möglichen Bahnkurven eines mechanischen Systems zwischen festgehaltenen Anfangs- und Endpunkten ist jene Kurve realisiert, für die die Wirkung stationär ist. In vielen Fällen sind die stationären Punkte Minima der Wirkung.

Symmetrien der Lagrange-Funktion führen auf Erhaltungsgrößen (Noether-Theorem). Wenn $\mathcal{L}$ nicht *explizit* vom Kurvenparameter abhängt, sondern nur in der Form $\mathcal{L}(c(t),c'(t))$, dann führt diese Translationssymmetrie auf den *Energieerhaltungssatz*

$$\frac{d}{dt}E(t) = 0\,; \qquad E(t) = \Big(\frac{\partial\mathcal{L}}{\partial v}(c(t),c'(t))\Big)\circ c'(t) - \mathcal{L}(c(t),c'(t))$$

für die Kurve, in der die Wirkung stationär ist:

$$\begin{aligned}
&\frac{d}{dt}\Big\{\Big(\frac{\partial\mathcal{L}}{\partial v}(c(t),c'(t))\Big)\circ c'(t)\Big\} \\
&= \Big(\frac{\partial\mathcal{L}}{\partial q}(c(t),c'(t))\Big)\circ c'(t) + \Big(\frac{\partial\mathcal{L}}{\partial v}(c(t),c'(t))\Big)\circ c''(t) = \frac{d}{dt}\mathcal{L}(c(t),c'(t))\,.
\end{aligned}$$

Beispiel 8.4 (Brachistochrone, Johann Bernoulli 1696) Die Brachistochrone ist jene Kurve $c(t) = (x(t),z(t))$ in der x-z-Ebene zwischen den festgehaltenen Punkten $(0,h)$ und $(a,0)$, mit $h,a>0$, die ein darauf reibungsfrei unter dem Einfluß der Gravitationskraft gleitender Massepunkt in kürzester Zeit durchläuft, wobei der Massepunkt in $(0,h)$ ruht. Die zeitunabhängige Lagrange-Funktion (e_z ist der Einheitsvektor in z-Richtung)

$$\mathcal{L}(c(t),c'(t)) = \frac{m}{2}\langle c'(t),c'(t)\rangle - mg\langle c(t),e_z\rangle$$

führt auf den Energieerhaltungssatz

$$\frac{m}{2}\langle c'(t), c'(t)\rangle + mg\langle c(t), e_z\rangle = mgh = \text{const.}$$

Wir können $x' > 0$ annehmen, so daß wir die Spur der Kurve nach dem Kurvenparameter auflösen können: $z(t) = z(x(t))$. Dann gilt nach Kettenregel

$$\begin{aligned} mgh &= \frac{m}{2}\big(x'(t)\big)^2\Big(1 + \Big(\frac{dz(x)}{dx}\Big)^2\Big) + mgz(x) \\ \Rightarrow \qquad \frac{dt(x)}{dx} &= \sqrt{\frac{1 + \Big(\frac{d(h-z(x))}{dx}\Big)^2}{2g(h - z(x))}} \;. \end{aligned}$$

Somit gilt für die Gesamtzeit T, in der der Massepunkt die Kurve durchläuft,

$$T = \int_0^a dx\,\Big(\frac{dt(x)}{dx}\Big) = \int_0^a dx\,\sqrt{\frac{1 + \Big(\frac{dy(x)}{dx}\Big)^2}{2gy(x)}}\;, \qquad y(x) = h - z(x)\;.$$

Damit ist nach Identifikation von x mit einem neuen Kurvenparameter die Bestimmung der schnellsten Kurve $y(x)$ äquivalent zur Lösung der Euler-Lagrange-Gleichung für die Lagrange-Funktion

$$\mathcal{L}(y(x), y'(x)) = \sqrt{\frac{1 + \langle y'(x), y'(x)\rangle}{y(x)}}\;, \quad y : [0, a] \to \mathbb{R}\;.$$

Da die neue Lagrange-Funktion zeitunabhängig ist, gilt wieder der Energieerhaltungssatz

$$\begin{aligned} E &= \frac{\langle y'(x), y'(x)\rangle}{\sqrt{y(x)}\sqrt{1 + \langle y'(x), y'(x)\rangle}} - \sqrt{\frac{1 + \langle y'(x), y'(x)\rangle}{y(x)}} \\ &= -\frac{1}{\sqrt{y(x)}\sqrt{1 + \langle y'(x), y'(x)\rangle}}\;, \\ \Rightarrow \qquad & \big(y'\big)^2 - \frac{1}{E^2 y} = -1 \end{aligned}$$

für die Kurve, für die die Wirkung stationär wird. Solche Differentialgleichungen betrachten wir im nächsten Kapitel. Wir beschränken uns hier auf Angabe der Lösung in Parameterform:

$$y(\tau) = \frac{1}{2E^2}\big(1 - \cos\tau\big)\;, \qquad x(\tau) = \frac{1}{2E^2}\big(\tau - \sin\tau\big)\;.$$

Dann ist nach Kettenregel

$$y'(x) = \frac{\frac{dy}{d\tau}}{\frac{dx}{d\tau}} = \frac{\sin\tau}{1 - \cos\tau} \qquad \Rightarrow \qquad \big(y'\big)^2 - \frac{1}{E^2 y} = \frac{\sin^2\tau - 2(1 - \cos\tau)}{(1 - \cos\tau)^2} = -1\;.$$

Somit ist die Brachistochrone in Parameterform gegeben durch

$$\begin{pmatrix} x(\tau) \\ z(\tau) \end{pmatrix} = \frac{1}{2E^2} \begin{pmatrix} \tau \\ 2E^2h - 1 \end{pmatrix} + \begin{pmatrix} \cos\tau & -\sin\tau \\ \sin\tau & \cos\tau \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{2E^2} \end{pmatrix}$$

Es handelt sich um die Spiegelung der Kurve des Randpunktes eines rollendes Rades vom Radius $R = \frac{1}{2E^2}$. ◁

Wiederholung

- Definitionen von partieller Ableitung, Richtungsableitung und (totalem) Differential und ihre gegenseitigen Beziehungen
- Satz von Schwarz
- Kettenregel
- differenzierbare Kurve, reguläre Kurve, Ableitung entlang Kurve, Bogenlänge
- Definition von Vektorfeldern, Gradient, Divergenz
- mehrdimensioanle Taylorsche Formel, Extremwertdiskussion über Gradient und Hessesche Matrix
- Banachscher Fixpunktsatz, Satz über implizite Funktionen
- Untermannigfaltigkeiten definiert über Nebenbedingungen sowie über Kartenabbildungen (d.h. lokales Koordinatensystem)
- Tangentialraum und Normalenvektorraum. Extrema mit Nebenbedingungen
- Euler-Lagrange-Gleichungen, Energieerhaltungssatz

Teil II
Gewöhnliche Differentialgleichungen

9 Definition und Interpretation

Unter einer gewönlichen Differentialgleichung versteht man eine Gleichung zwischen einer Funktion $x(t)$, ihren Ableitungen $x', x'', \ldots x^{(n)}$ und der Variablen t, genauer:

Definition 9.1 Es sei $I \subset \mathbb{R}$ ein Intervall, $G \subset \mathbb{R}^{n+2}$ eine Teilmenge mit $n \geq 1$ und $x : I \to \mathbb{R}$ eine n-mal differenzierbare Funktion, so daß für $t \in I$ gilt $(t, x(t), x'(t), \ldots, x^{(n)}(t)) \in G$. Die Funktion x erfüllt eine *gewöhnliche Differentialgleichung n-ter Ordnung*, wenn es eine Funktion $F : G \to \mathbb{R}$ gibt, so daß

$$F(t, x(t), x'(t), \ldots, x^{(n)}(t)) = 0 \qquad \forall t \in I \; .$$

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

$$x^{(n)}(t) = f(t, x(t), x'(t), \ldots, x^{(n-1)}(t)) \; , \tag{*}$$

dann heißt die Differentialgleichung *explizit*.

Eine explizite Differentialgleichung n-ter Ordnung (*) läßt sich identifizieren mit einer Differentialgleichung erster Ordnung für eine Bahnkurve $c : I \to \mathbb{R}^n$. Setzt man

$$c(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_n(t) \end{pmatrix} = \begin{pmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix} ,$$

so ergibt die Ableitung unter Verwendung von (*) das äquivalente *System von Differentialgleichungen 1. Ordnung*

$$c'(t) = \begin{pmatrix} x'(t) \\ x''(t) \\ \vdots \\ f(t, x(t), x'(t), \ldots, x^{(n-1)}(t)) \end{pmatrix} =: v(t, c(t)) \; .$$

Dieses können wir als Differentialgleichung $c'(t) = v(t, c(t))$ einer Bahnkurve interpretieren. Solche Bewegungsgleichungen entstehen z.B. aus den Euler-Lagrange-Gleichungen nach Übergang in die Hamiltonsche Formulierung (Legendre-Transformation).

Ist allgemeiner $U \subset \mathbb{R}^n$ eine offene Teilmenge und $v : I \times U \to \mathbb{R}^n$ ein *zeitabhängiges Vektorfeld*, dann definiert die Differentialgleichung $c'(t) = v(t, c(t))$

zu jedem Zeitpunkt $t \in I$ eine *Schar von Kurven* durch jeden Punkt $y \in \Omega$ derart, daß der Tangentialvektor $c'(t)$ im Punkt $y = c(t)$ gegeben ist durch den Vektor $v(t, y)$. Diese Kurvenschar heißt die *Integralkurve* des Vektorfeldes. Anschaulich interpretiert man die Integralkurven als Stromlinien des Feldes v. Zur Bestimmung einer konkreten Integralkurven ist der Startpunkt $y_0 = y(t_0)$ vorzugeben. Wir werden sehen, daß dann unter geeigneten Stetigkeitsbedingungen an das Vektorfeld die Integralkurve in einer offenen Umgebung des Startpunktes eindeutig bestimmt ist. Andere Startpunkte liefern entweder verschiedene Kurven oder setzen eine bisherige Kurve fort, so daß U schließlich durch eine Schar von Kurven überdeckt wird.

10 Elementare Lösungsmethoden

Wir beginnen mit einigen einfachen Klassen gewöhnlicher Differentialgleichungen 1. Ordnung.

Definition 10.1 (DGL 1. Ordnung mit getrennten Variablen) Seien $I, J \subset \mathbb{R}$ offene Intervalle und $f : I \to \mathbb{R}$ und $g : J \to \mathbb{R}$ stetige Funktionen, wobei $g(x) \neq 0$ für alle $x \in J$. Dann heißt

$$x'(t) = f(t)\, g(x(t))$$

Differentialgleichung mit getrennten Variablen.

Satz 10.2 *Zu gegebenem Anfangspunkt* $(t_0, x_0) \in I \times J$ *definieren wir Funktionen* $F : I \to \mathbb{R}$ *und* $G : J \to \mathbb{R}$ *durch*

$$F(t) := \int_{t_0}^{t} ds\, f(s)\ , \qquad G(x) := \int_{x_0}^{x} \frac{dy}{g(y)}\ .$$

Ist $I' \subset I$ *ein offenes Intervall mit* $F(I') \subset G(J)$, *dann gibt es genau eine Lösung* $x : I' \to \mathbb{R}$ *der Differentialgleichung* $x'(t) = f(t)\, g(x(t))$ *mit der Anfangsbedingung* $x(t_0) = x_0$. *Diese Lösung erfüllt die Gleichung*

$$G(x(t)) = F(t) \qquad \text{für alle } t \in I'\ .$$

Beweis. i) Wir nehmen die Gültigkeit von $G(x(t)) = F(t)$ an. Wegen $G'(x) = \frac{1}{g(x)} \neq 0$ für alle $x \in J$ ist G streng monoton (und stetig differenzierbar), so daß nach dem Satz über implizite Funktionen eine eindeutige stetig differenzierbare Umkehrfunktion $G^{-1} : G(J) \to \mathbb{R}$ auf dem offenen Intervall $G(J)$ existiert.

ii) Nach dem Satz über implizite Funktionen ist $H := G^{-1} \circ F : I' \to \mathbb{R}$ stetig differenzierbar mit

$$H'(t) = (G^{-1})'(F(t)) \cdot F'(t) = \frac{1}{G'(G^{-1}(F(t)))} \cdot F'(t) = g(H(t)) \cdot f(t)\ ,$$

d.h. H erfüllt die Differentialgleichung mit der Anfangsbedingung $H(t_0) = G^{-1}(F(t_0)) = G^{-1}(0) = x_0$. Also existiert eine Lösung.

iii) Sei $x(t)$ eine beliebige Lösung der Differentialgleichung $x'(t) = f(t)\, g(x(t))$ mit $x(t_0) = x_0$. Wir setzen $t \mapsto s$ und integrieren über s von t_0 nach t:

$$\int_{t_0}^{t} ds\ \frac{x'(s)}{g(x(s))} = \int_{t_0}^{t} ds\ f(s)\ .$$

Nach Substitution $y = x(s)$ ergibt sich

$$\int_{x_0}^{x(t)} \frac{dy}{g(y)} = \int_{t_0}^{t} ds\ f(s)\ .$$

Damit erfüllt *jede* Lösung von $x'(t) = f(t)\, g(x(t))$ mit Anfangsbedingung $x(t_0) = x_0$ die Gleichung $G(x(t)) = F(t)$. Nach dem Satz über implizite Funktionen besitzt $G(x(t)) = F(t)$ in einer Umgebung von t_0 genau eine Lösung $x(t)$. $\square$

Zu beachten ist, daß die Umkehrfunktion G^{-1} im allgemeinen nicht durch elementare Funktionen auszudrücken ist.

Beispiel 10.3 Wir betrachten die Differentialgleichung $x'(t) = e^{x(t)} \sin t$ zu beliebigem Anfangspunkt $(t_0, x_0) \in \mathbb{R}^2$. Damit ist $F(t) = \cos t_0 - \cos t$ und $G(x) = e^{-x_0} - e^{-x}$, also

$$e^{-x(t)} = \cos t - \cos t_0 + e^{-x_0}\ .$$

Wegen $e^{-x_0} > 0$ gibt es ein offenes Intervall $I' \subset \mathbb{R}$ mit $t_0 \in I'$, so daß $\cos t - \cos t_0 + e^{-x_0} > 0$ für alle $t \in I'$. Dann ergibt sich die Lösung zu

$$x(t) = -\ln\big(\cos t - \cos t_0 + e^{-x_0}\big)$$

für alle $t \in I'$. $\triangleleft$

Definition 10.4 (lineare Differentialgleichung 1. Ordnung) Es seien $I \subset \mathbb{R}$ und $a, b : I \to \mathbb{R}$ stetige Funktionen. Dann heißt

$$x'(t) = a(t)\, x(t) + b(t)$$

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

Satz 10.5 *Es sei $t_0 \in I$ und $x_0 \in \mathbb{R}$. Dann gibt es genau eine Lösung $x : I \to \mathbb{R}$ der Differentialgleichung $x'(t) = a(t)x(t)$ mit der Anfangsbedingung $x(t_0) = x_0$, nämlich*

$$x(t) = x_0 \exp\Big(\int_{t_0}^{t} ds\ a(s)\Big)\ .$$

Beweis. Es handelts sich um einen Spezialfall einer Differentialgleichung mit getrennten Variablen. □

Der inhomogene Fall wird durch *Variation der Konstanten* gelöst. Darunter versteht man den Ansatz $x(t) = \tilde{x}(t)u(t)$, wobei $\tilde{x}(t)$ die homogene Gleichung $\tilde{x}'(t) = a(t)\tilde{x}(t)$ mit $\tilde{x}(t_0) = 1$ löst. Damit ergibt sich

$$\begin{aligned} x'(t) &= \tilde{x}'(t)\,u(t) + \tilde{x}(t)\,u'(t) = \big(a(t)\,u(t) + u'(t)\big)\tilde{x}(t) \\ &= a(t)\,u(t)\,\tilde{x}(t) + b(t)\;, \end{aligned}$$

also $u'(t)\tilde{x}(t) = b(t)$ mit Anfangsbedingung $u(t_0) = x_0$. Das ist wieder eine spezielle Differentialgleichung mit getrennten Variablen mit der eindeutigen Lösung

$$u(t) = x_0 + \int_{t_0}^{t} ds\, \frac{b(s)}{\tilde{x}(s)}\;.$$

Somit ist bewiesen:

Satz 10.6 (Variation der Konstanten) *Es seien $I \subset \mathbb{R}$ ein Intervall, $a, b : I \to \mathbb{R}$ stetige Funktionen sowie $t_0 \in I$ und $x_0 \in \mathbb{R}$. Dann gibt es genau eine Lösung $x : I \to \mathbb{R}$ der Differentialgleichung $x'(t) = a(t)y(t) + b(t)$ mit der Anfangsbedingung $x(t_0) = x_0$, nämlich*

$$x(t) = \tilde{x}(t)\Big(x_0 + \int_{t_0}^{t} ds\, \frac{b(s)}{\tilde{x}(s)}\Big) \qquad \textit{mit} \qquad \tilde{x}(s) = \exp\Big(\int_{t_0}^{s} dr\; a(r)\Big)\;.$$

Beispiel 10.7 Betrachtet werde die Differentialgleichung $x'(t) = 2tx(t) + t$ mit $x(0) = c$. Dann hat die homogene Gleichung $\tilde{x}'(t) = 2t\tilde{x}(t)$ mit $\tilde{x}(0) = 1$ die Lösung

$$\tilde{x}(t) = \exp\Big(\int_0^t ds\; 2s\Big) = e^{t^2}\;.$$

Also erhalten wir

$$\begin{aligned} x(t) &= e^{t^2}\Big(c + \int_0^t ds\; se^{-s^2}\Big) = e^{t^2}\Big(c + \frac{1}{2}\Big(1 - e^{-t^2}\Big)\Big) \\ &= \frac{2c+1}{2}e^{t^2} - \frac{1}{2}\;. \end{aligned}$$ ◁

Beispiel 10.8 (Freier Fall mit Reibung) Der freie Fall eines Massenpunktes (in x-z-Ebene) wird bei geschwindigkeitsproportionaler Reibungskraft beschrieben durch die Differentialgleichungen

$$mz''(t) = -mg - rz'(t)\;, \qquad mx''(t) = -rx'(t)\;.$$

Als Anfangsbedingung sei $z'(0) = v_z$ und $x'(0) = v_x$ gegeben. Damit ergeben sich die Geschwindigkeiten zu

$$\begin{aligned}
x'(t) &= v_x \exp\Big(\int_0^t ds\ (-\frac{r}{m})\Big) = v_x e^{-\frac{r}{m}t} , \\
z'(t) &= e^{-\frac{r}{m}t}\Big(v_z - \int_0^t ds\ g e^{\frac{r}{m}s}\Big) = v_z e^{-\frac{r}{m}t} - \frac{mg}{r}(1 - e^{-\frac{r}{m}t}) ,
\end{aligned}$$

Für $t \to \infty$ fällt der Massenpunkt also mit konstanter Geschwindigkeit $\frac{mg}{r}$ in negative z-Richtung. Die Anfangsgeschwindigkeiten sind exponentiell gedämpft.

Beide Lösungen sind selbst Differentialgleichungen mit getrennten Variablen. Sind die Anfangsbedingungen $z(0) = h$ und $x(0) = 0$, dann ergibt sich

$$\begin{aligned}
x(t) &= \int_0^t ds\ v_x e^{-\frac{r}{m}s} = \frac{v_x m}{r}\Big(1 - e^{-\frac{r}{m}t}\Big) , \\
z(t) &= h + \int_0^t ds \Big(v_z e^{-\frac{r}{m}s} - \frac{mg}{r}(1 - e^{-\frac{r}{m}s})\Big) \\
&= h + \frac{v_z m}{r}\Big(1 - e^{-\frac{r}{m}t}\Big) - \frac{mg}{r}\Big(t - \frac{m}{r}(1 - e^{-\frac{r}{m}t})\Big) . \qquad \triangleleft
\end{aligned}$$

Manchmal findet man Substitutionen, durch die zunächst kompliziertere Differentialgleichungen auf obige Situationen zurückgeführt werden:

Beispiel 10.9 (Bernoullische Differentialgleichung) Die Bernoullische Differentialgleichung lautet

$$y'(t) = a(t)y(t) + b(t)\big(y(t)\big)^{\alpha} , \quad \alpha \neq 0, 1 .$$

Zu unterscheiden ist $\alpha \in \mathbb{Z}$ und $\alpha \in \mathbb{R} \setminus \mathbb{Z} > 0$. Für $\alpha > 0$ gibt es zumindest die triviale Lösung $y = 0$. Für $(\alpha < 0, \alpha \notin \mathbb{Z})$ ist nur $y(t) > 0$ sinnvoll.

Sei $\alpha \in \mathbb{R} \setminus \mathbb{Z}$ und $y(t_0) > 0$. Wir substituieren $x(t) = \big(y(t)\big)^{1-\alpha}$ und erhalten nach Kettenregel in einer Umgebung von $(x(t_0), t_0)$

$$x'(t) = (1 - \alpha)\big(y(t)\big)^{-\alpha} y'(t) = (1 - \alpha)\Big(a(t)x(t) + b(t)\Big) . \qquad \triangleleft$$

Viele Probleme der Physik führen auf Differentialgleichungen 2. Ordnung. Das typische Beispiel ist das Newtonsche Gesetz für die eindimensionale Bewegung der Mechanik $x'' = f(x, x')$, dabei ist f die durch die Masse dividierte Kraft. Die übliche Anfangsbedingung ist die Vorgabe von Ort und Geschwindigkeit zur Zeit t_0, d.h. $x(t_0) = x_0$ und $x'(t_0) = v_0$. Diese Differentialgleichung läßt sich in ein System linearer Differentialgleichungen 1. Ordnung überführen und dann mit den vorzustellenden Methoden lösen. Ist die Kraft $f(y)$ unabhängig von der Geschwindigkeit, aber z.B. nichtlinear, dann bietet sich noch ein anderer Weg an: Wir werden zeigen, daß das Potential

$$U(x) := -\int_{x_0}^{x} ds\ f(s)$$

unabhängig vom Integrationsweg definiert ist und $\frac{dU}{dx}(x) = -f(x)$ erfüllt. Durch Skalarprodukt des Newtonschen Gesetzes mit der Geschwindigkeit x' ergibt sich dann

$$0 = x''x' + \frac{dU}{dx}x' = \Big(\frac{1}{2}x'^2 + U(x)\Big)' .$$

Die Lösung ist der Energieerhaltungssatz $\frac{1}{2}x'^2+U(x) = E = \frac{1}{2}v_0^2+U(y_0) = \text{const.}$ Das ist nun eine Differentialgleichung mit getrennten Variablen mit der Lösung

$$t - t_0 = \pm \int_{x_0}^{x} \frac{ds}{\sqrt{2(E - U(s))}} =: T(x) .$$

Der Ort selbst bestimmt sich dann durch die Umkehrfunktion zu $x(t) = T^{-1}(t - t_0)$.

Beispiel 10.10 Die Differentialgleichung des mathematischen Pendels ist $x'' = -\omega^2 \sin x$. Dabei ist x der Auslenkungswinkel und $\omega^2 = \frac{g}{L}$, mit g der Schwerebeschleunigung und L der Fadenlänge. Für kleine Winkel ist $\sin x \approx x$, und es wird die einfachere lineare Schwingungsdifferentialgleichung erhalten. Wir stellen aber nicht die Bedingung kleiner Winkel. Das Potential ergibt sich zu

$$U(x) = -\int_0^x ds\ (-\omega^2 \sin s) = \omega^2(1 - \cos x) = 2\omega^2 \sin^2 \tfrac{x}{2} .$$

Dieses entspricht der Gesamtenergie im maximalen Auslenkungswinkel $a = x(0)$, so daß sich folgende Lösung ergibt:

$$t = \frac{1}{2\omega} \int_x^a \frac{ds}{\sqrt{\sin^2 \frac{a}{2} - \sin^2 \frac{s}{2}}}$$

Es interessiert die Zeit $\frac{T}{2}$ einer halben Schwingung zum anderen Umkehrpunkt $x = -a$. 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

$$T = \frac{2\pi}{\omega}\Big(1 + \sum_{n=1}^{\infty} \Big(\frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)}\Big)^2 \sin^{2n} \tfrac{a}{2}\Big)$$

gezeigt. Insbesondere ist $\lim_{a\to 0} T = \frac{2\pi}{\omega}$. ◁

11 Existenz- und Eindeutigkeitssatz

Wir betrachten im weiteren die Differentialgleichung $x'(t) = v(t, x(t))$ mit $x(t) \in \mathbb{R}^n$. Genauer sei $G \subset \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $v : G \to \mathbb{R}^n$ eine stetige Abbildung, gesucht ist eine Lösung $x : I \to \mathbb{R}^n$ der Differentialgleichung $x'(t) = v(t, x(t))$, so daß $(t, x(t)) \in G$ für alle $t \in I \subset \mathbb{R}$. Meist sucht man Lösungen mit vorgegebener Anfangsbedingung $x(t_0) = x_0$. Diese Problemstellung heißt *System von n Differentialgleichungen 1. Ordnung.*

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

Satz 11.1 *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $v : G \to \mathbb{R}^n$ eine stetige Abbildung sowie $(t_0, x_0) \in G$. Dann gilt: Eine stetige Abbildung $x : I \to \mathbb{R}^n$ eines Intervalls $I \subset \mathbb{R}$ mit $t_0 \in I$ und $(t, x(t)) \in G$ für alle $t \in I$ löst genau dann die Differentialgleichung $x'(t) = v(t, x(t))$ zur Anfangsbedingung $x(t_0) = x_0$, wenn folgende Integralgleichung gilt:*

$$x(t) = x_0 + \int_{t_0}^{t} ds\, v(s, x(s)) \qquad \textit{für alle } t \in I\ .$$

Beweis. ($\Leftarrow$) Für $t = t_0$ folgt $x(t_0) = x_0$. Wegen der Stetigkeit von x und v ist auch die Abbildung $F : I \to \mathbb{R}$ mit $F(s) := v(s, x(s))$ stetig auf I. Damit gilt nach dem Fundamentalsatz der Differential- und Integralrechnung

$$\frac{d}{dt} \int_{t_0}^{t} ds\, v(s, x(s)) = v(t, x(t))\ ,$$

d.h. x ist sogar differenzierbar mit $x'(t) = v(t, x(t))$.

($\Rightarrow$) Es gilt

$$\int_{t_0}^{t} ds\, v(s, x(s)) = \int_{t_0}^{t} ds\, x'(s) = x(t) - x(t_0) = x(t) - x_0\ ,$$

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

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

Die Funktion v genügt einer *lokalen Lipschitz-Bedingung*, wenn jeder Punkt $(t_0, x_0) \in G$ eine Umgebung $U \subset \mathbb{R} \times \mathbb{R}^n$ besitzt, so daß v in $G \cap U$ einer Lipschitz-Bedingung mit möglicherweise von U abhängiger Lipschitz-Konstanten $L(U)$ genügt.

Hinreichend ist stetige Differenzierbarkeit:

Satz 11.3 *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $v : G \to \mathbb{R}^n$ stetig partiell differenzierbar in den letzten n Koordinatenrichtungen $x_1, \ldots, x_n$. Dann genügt f in G lokal einer Lipschitz-Bedingung.*

Beweis. Wegen der Offenheit von G gibt es zu beliebigem $(t_0, x_0) \in G$ ein $r > 0$, so daß

$$V := \{(t, x) \in \mathbb{R} \times \mathbb{R}^n \;:\; |t - t_0| \leq r\,,\; \|x - x_0\| \leq r\} \subset G\,.$$

Die Behauptung folgt nun aus dem Schrankensatz 2.14 für V, und die Lipschitz-Konstante ist $L := \sup_{(t,x)\in V} \big\|(Dv)(t, x)\big\|$. □

Satz 11.4 (Eindeutigkeitssatz) *Es sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $v : G \to \mathbb{R}^n$ eine stetige Abbildung, die lokal einer Lipschitz-Bedingung genügt. Sind $x, \tilde{x}$ zwei Lösungen derselben Differentialgleichung $x'(t) = v(t, x(t))$ und $\tilde{x}'(t) = v(t, \tilde{x}(t))$ über einem Intervall $I \subset \mathbb{R}$, und in einem Punkt $t_0 \in I$ gelte $x(t_0) = \tilde{x}(t_0)$, so folgt*

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

Beweis. Die Eindeutigkeit wird stückweise vom Punkt t_0 aus ausgedehnt.

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

$$x(t) - \tilde{x}(t) = \int_a^t ds\, \big(v(s, x(s)) - v(s, \tilde{x}(s))\big)\,.$$

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

$$\|v(s, x(s)) - v(s, \tilde{x}(s))\| \leq L\|x(s) - \tilde{x}(s)\| \qquad \text{für alle } s \in I \cap K_\delta(a)\,.$$

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

$$M := \sup_{s \in I \cap K_\epsilon(a)} \|x(s) - \tilde{x}(s)\|\,.$$

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

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

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

ii) Wir zeigen nun $x(t) = \tilde{x}(t)$ für alle $t \in I$ mit $t \geq t_0$. Dazu sei

$$t_1 := \sup\big\{s \in I \;:\; x(s) = \tilde{x}(s)\big\}\,.$$

Für $t_1 = +\infty$ oder t_1 gleich dem Intervallende ist nichts mehr zu zeigen. Ansonsten gibt es ein $\delta > 0$ mit $]t_1, t_1 + \delta[\subset I$. Da $x, \tilde{x}$ als Lösungen der Differentialgleichung insbesondere stetig sind, gilt $x(t_1) = \tilde{x}(t_1)$. Dann aber gibt es nach i) ein $\epsilon > 0$, so daß $x(t) = \tilde{x}(t)$ für alle $x \in I \cap K_\epsilon(t_1)$, im Widerspruch zur Definition von t_1. Analog wird der Fall $t \leq t_0$ behandelt. □

Satz 11.5 (Picard-Lindelöf) *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $v : G \to \mathbb{R}^n$ eine stetige Abbildung, die lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jedem $(t_0, x_0) \in G$ ein $\epsilon > 0$ und eine Lösung $x : [t_0 - \epsilon, t_0 + \epsilon] :\to \mathbb{R}^n$ der Differentialgleichung $x'(t) = v(t, x(t))$ mit Anfangsbedingung $x(t_0) = x_0$. Nach Satz 11.4 ist diese Lösung eindeutig.*

Beweis. Der Beweis verwendet den Banachschen Fixpunktsatz für die zugehörige Integralgleichung auf dem Banach-Raum $\mathcal{C}([t_0 - \epsilon, t_0 + \epsilon], \mathbb{R}^n)$ der stetigen Abbildungen $\phi : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}^n$ mit der Supremumsnorm

$$\|\phi\|_{\sup} := \sup_{t \in [t_0 - \epsilon, t_0 + \epsilon]} \|\phi(t)\| .$$

Gewählt sei ein Quader

$$Q_{\delta,r} := \{(t, x) \in \mathbb{R} \times \mathbb{R}^n \ : \ |t - t_0| \leq \delta \ , \ \|x - x_0\| \leq r\}$$

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

$$A := \{\phi \in \mathcal{C}([t_0 - \epsilon, t_0 + \epsilon], \mathbb{R}^n) \ , \ \|\phi - x_0\|_{\sup} \leq r\}$$

des Banachraums. Wir zeigen, daß

$$(T(\phi))(t) := x_0 + \int_{t_0}^{t} ds \ v(t, \phi(s))$$

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

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

$$(s, \phi(s)) \in Q_{\delta,r} \subset G \qquad \text{für alle } s \in [t_0 - \epsilon, t_0 + \epsilon] .$$

Damit ist v sinnvoll erklärt und stetig in s. Es gilt

$$\|(T(\phi))(t) - x_0\| = \Big\| \int_{t_0}^{t} ds \ v(s, \phi(s)) \Big\| \leq |t - t_0| \, M \leq \epsilon M \leq r ,$$

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

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

$$\begin{aligned}\|(T(\phi_1) - T(\phi_2))(t)\| &= \Big\| \int_{t_0}^t ds\, \big(v(s,\phi_1(s)) - v(s,\phi_2(s))\big)\Big\| \\ &\leq |t-t_0|\, L \sup_{s\in[t_0,t]} \|\phi_1(s) - \phi_2(s)\| \leq \frac{1}{2}\|\phi_1 - \phi_2\|_{\sup} \;.\end{aligned}$$

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

$$x(t) = x_0 + \int_{t_0}^t ds\, v(s, x(s)) \;,$$

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

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

Beispiel 11.6 Betrachtet werde die Differentialgleichung $x'(t) = 2t\,(x(t))^2$. Die Funktion $v(t,x) = 2tx^2$ ist auf ganz $\mathbb{R}^2$ nach x partiell differenzierbar, genügt also auf ganz $\mathbb{R}^2$ einer lokalen Lipschitz-Bedingung. Die eindeutige Lösung zu $x(0) = 0$ ist offenbar $x = 0$. Sei dann $x(0) = \frac{1}{C} > 0$, so daß die Methode der Trennung der Variablen anwendbar ist. Es gilt

$$G(x(t)) = F(t) \;, \quad G(x(t)) = \int_{\frac{1}{C}}^{x(t)} dy \frac{1}{y^2} = C - \frac{1}{x(t)} \;, \quad F(t) = \int_0^t ds\; 2s = t^2 \;,$$

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

Ohne Beweis erwähnen wir, daß die *Existenz* einer Lösung der Differentialgleichung $x'(t) = v(t, x(t))$ bereits durch die *Stetigkeit* von v garantiert ist:

Satz 11.7 (Peano) *Es sei $t_0 \in \mathbb{R}$ und $x_0 \in \mathbb{R}^n$ und*

$$Q_{\delta,r} := \{(t,x) \in \mathbb{R} \times \mathbb{R}^n \;:\; |t - t_0| \leq \delta \;,\; \|x - x_0\| \leq r\}$$

für $\delta, r > 0$. Die Funktion $v : Q_{\delta,r} \to \mathbb{R}^n$ sei stetig (also beschränkt) mit $M := \max_{(t,x)\in Q_{\delta,r}} \|v(t,x)\|$. Dann gibt es mindestens eine Lösung $x : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}$ der Differentialgleichung $x'(t) = v(t, x(t))$ mit Anfangsbedingung $x(t_0) = x_0$, 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 11.8 Betrachtet werde die Differentialgleichung $x' = (x^2)^{\frac{1}{3}}$ mit $x(0) = 0$. Die Funktion $v(t,x) = (x^2)^{\frac{1}{3}}$ ist stetig auf ganz $\mathbb{R}^2$, aber nicht stetig partiell differenzierbar in $x = 0$, und tatsächlich genügt v in keiner Umgebung von $(0,0)$ einer lokalen Lipschitz-Bedingung. Nach dem Existenzsatz von Peano gibt es lokal eine Lösung der Differentialgleichung, z.B. $x(t) = 0$ für alle $t \in \mathbb{R}$. Das ist aber nicht die einzige Lösung. Durch Trennen der Variablen findet man, daß $x_c(t) := \frac{1}{27}(t-c)^3$ die Differentialgleichung zur Anfangsbedingung $x(c) = 0$ erfüllt. Durch Verkleben mit der Nullösung konstruieren wir zu $c, c' \geq 0$ die Lösung

$$x_{cc'}(t) := \begin{cases} \frac{1}{27}(t-c)^3 & \text{für } t \geq c \geq 0 \\ 0 & \text{für } -c' \leq t \leq c \\ \frac{1}{27}(t+c')^3 & \text{für } t \leq -c' \leq 0 \end{cases}$$

Man rechnet nach, daß $x_{cc'}$ tatsächlich stetig differenzierbar ist. $\triangleleft$

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

Beispiel 11.9 Gegeben sei die Differentialgleichung $x'(t) = 2t\,x(t)$ auf $G = \mathbb{R}\times\mathbb{R}$ mit Anfangsbedingung $x(0) = x_0$. In einer gewissen Umgebung $[-\epsilon, \epsilon]$ von 0 konvergiert deshalb die Folge $(x_k(t))_{k\in\mathbb{N}}$ von Funktionen mit $x_0(t) = y_0$ und

$$x_{k+1}(t) = x_0 + \int_0^t ds\; v(s, x_k(s))$$

gegen die eindeutige Lösung. Wir finden

$$\begin{aligned} x_1(t) &= x_0 + \int_0^t ds\; 2sx_0 = x_0(1+t^2) \\ x_2(t) &= x_0 + \int_0^t ds\; 2sx_0(1+s^2) = x_0\Big(1 + t^2 + \frac{t^4}{2}\Big) \\ x_k(t) &= y_0\Big(1 + t^2 + \frac{t^4}{2} + \dots + \frac{t^{2k}}{k!}\Big)\,. \end{aligned}$$

Damit ergibt sich die Lösung $x(t) = \lim_{k\to\infty} x_k(t) = x_0\, e^{t^2}$, was man natürlich auch mit elementaren Methoden finden kann. $\triangleleft$

12 Lineare Differentialgleichungen

Wir spezifizieren nun den Existenz- und Eindeutigkeitssatz auf Systeme linearer Differentialgleichungen $x'(t) = A(t)x(t) + b(t)$. Dabei ist $A(t)$ eine Matrix und $b(t)$ 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 t bleibt aber reell.

Definition 12.1 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 $x'(t) = A(t)x(t) + b(t)$ ein *inhomogenes lineares Differentialgleichungssystem* und $x'(t) = A(t)x(t)$ ein (bzw. das zugehörige) *homogene(s) lineare(s) Differentialgleichungssystem*.

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

Satz 12.2 *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 $t_0 \in I$ und jedem $x_0 \in \mathbb{K}^n$ genau eine Lösung $x : I \to \mathbb{K}^n$ der linearen Differentialgleichung $x'(t) = A(t)x(t) + b(t)$ mit Anfangsbedingung $x(t_0) = x_0$.*

Beweis. Für jedes kompakte Intervall $J \subset I$ ist die durch $v(t, x) = A(t)x + b(t)$ definierte Abbildung $v : J \times \mathbb{R}^n \to \mathbb{R}^n$ Lipschitz-stetig mit der Lipschitz-Konstanten $L = \sup_{t \in J} \|A(t)\|_{op}$. Damit folgt Existenz und Eindeutigkeit der Lösung nach Picard-Lindelöf. Insbesondere konvergiert die Folge $(x_k)_{k \in \mathbb{N}}$ mit

$$x_{k+1}(t) := x_0 + \int_{t_0}^{t} ds\; v(s, x_k(s)) \;, \qquad x(t_0) = x_0 \;,$$

auf J gleichmäßig gegen die Lösung x. □

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

Satz 12.3 *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 $x'(t) = A(t)x(t)$ einen n-dimensionalen Vektorraum über $\mathbb{K}$.*

Ein n-Tupel von Lösungen $x_{(1)}, \ldots, x_{(n)} \in V_A$ bildet genau dann eine Basis von V_A, wenn die Vektoren $x_{(1)}(t), \ldots, x_{(k)}(t) \in \mathbb{K}^n$ für wenigstens ein $t \in I$, und damit für jedes $t \in I$, eine Basis von $\mathbb{K}^n$ bilden.

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

$$\begin{aligned}(\lambda_1 x_{(1)} + \lambda_2 x_{(2)})' &= \lambda_1 x_{(1)}{}' + \lambda_2 x_{(2)}{}' = \lambda_1 (A(t)x_{(1)}) + \lambda_2 (A(t)x_{(2)}) \\ &= A(t) \cdot (\lambda_1 x_{(1)} + \lambda_2 x_{(2)}) \;.\end{aligned}$$

Sei $x \in V_A$ und $t_0 \in I$ beliebig, dann definiert $V_A \ni x \mapsto F_{t_0}(x) := x(t_0) \in \mathbb{K}^n$ eine lineare Abbildung $F_{t_0} : V_A \to \mathbb{K}^n$. Diese lineare Abbildung ist surjektiv, denn zu jedem $x_0 \in \mathbb{K}^n$ gibt es ein $x \in V_A$ mit $x(t_0) = x_0$. Außerdem ist F_{t_0}

wegen der Eindeutigkeit der Lösung des Anfangswertproblems injektiv. Nach dem Dimensionssatz für Vektorräume gilt dann $\dim(V_A) = n$. Da der Anfangspunkt $t_0 \in I$ beliebig ist, ist F_{t_0} für alle $t_0 \in I$ ein Isomorphismus, überführt also Basen wieder in Basen. □

Definition 12.4 Unter einem *Lösungs-Fundamentalsystem* der Differentialgleichung $x'(t) = A(t)x(t)$ versteht man eine Basis $(x_{(1)}, \ldots, x_{(n)})$ des Vektorraums V_A der Lösungen von $x'(t) = A(t)x(t)$.

Schreibt man das Lösungs-Fundamentalsystem als Matrix $\Phi = (x_{ij})$, mit $x_{ij} := x_{(j)i}$ (die j-te Spalte von Φ ist $x_{(j)}$), dann gilt offenbar $\det \Phi(t_0) \neq 0$ für wenigstens einen (und damit für jeden) Punkt $t_0 \in I$. Eine beliebige Lösung schreibt sich damit als $x(t) = \Phi(t) \cdot \lambda$ mit $\lambda = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{K}^n$.

Beispiel 12.5 Gegeben sei die Schwingungsdifferentialgleichung $z'' + \omega^2 z = 0$. Wir setzen $x_1 := z$ und $x_2 := -\frac{1}{\omega} z'$, was also auf das Differentialgleichungssystem $x_1' = -\omega x_2$ und $x_2' = \omega x_1$ führt. In Matrixschreibweise ergibt sich $x' = Ax$ 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 $x_i : \mathbb{R} \to \mathbb{R}^2$ Lösungen der Differentialgleichung sind:

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

Diese sind in $t_0 = 0$ und damit auf ganz $\mathbb{R}$ linear unabhängig, was man auch durch

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

sieht. Da A spurfrei ist, ist $\det \Phi$ konstant. ◁

Wir charakterisieren nun den Lösungsraum der inhomogenen Differentialgleichung:

Satz 12.6 *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*

$$L_{A,b} := \{x : I \to \mathbb{K}^n \ : \ x'(t) = A(t)x(t) + b(t)\}$$

der inhomogenen Differentialgleichung ist der affine Raum

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

wobei V_A der Vektorraum aller Lösungen der zugehörigen homogenen Differentialgleichung $x'(t) = A(t)x(t)$ und $\tilde{x} \in L_{A,b}$ 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 $x \in L_{A,b}$ eine Lösung von $x'(t) = A(t)x(t) + b(t)$, dann erfüllt $u := x - \tilde{x}$ die homogene Differentialgleichung $u'(t) = A(t)u(t)$. Das bedeutet $u \in V_A$, also $x \in \tilde{x} + V_A$ und damit $L_{A,b} \subset \tilde{x} + V_A$.

ii) Sei umgekehrt $x \in \tilde{x} + V_A$ gegeben, also $x = \tilde{x} + u$ mit $u'(t) = A(t)u(t)$. Dann gilt

$$x'(t) = \tilde{x}'(t) + u'(t) = (A(t)\tilde{x}(t) + b(t)) + A(t)u(t) = A(t)x(t) + b(t) \;,$$

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

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

Satz 12.7 (Variation der Konstanten) *Sei* $\Phi = (x_{(1)}, \dots, x_{(n)}) : I \to GL(n, \mathbb{K})$ *ein Lösungsfundamentalsystem der homogenen Differentialgleichung* $x'(t) = A(t)x(t)$. *Dann wird eine Lösung* $\tilde{x}$ *der inhomogenen Differentialgleichung* $\tilde{x}' = A(t)\tilde{x}(t) + b(t)$ *erhalten durch den Ansatz* $\tilde{x}(t) = \Phi(t) \cdot y(t)$, *wobei die differenzierbare Abbildung* $y : I \to \mathbb{K}^n$ *der Differentialgleichung* $\Phi(t)y'(t) = b(t)$ *genügt. Ihre Lösung ist somit*

$$y(t) = \int_{t_0}^{t} ds\; \Phi^{-1}(s) \cdot b(s) + const\,.$$

Beweis. Es gilt

$$\tilde{x}'(t) = \Phi'(t) \cdot y(t) + \Phi(t)y'(t) = A(t) \cdot \Phi(t) \cdot y(t) + \Phi(t)y'(t) \stackrel{!}{=} A(t)\tilde{x}(t) + b(t) \;,$$

also $y'(t) = \Phi^{-1}(t) \cdot b(t)$. Umgeschreiben in eine Integralgleichung ergibt sich die Lösung. □

Beispiel 12.8 Gegeben sei die Schwingungsdifferentialgleichung mit periodischer äußerer Kraft $z'' + \omega^2 z = b\cos(\Omega t)$. Mit $x_1 := z$ und $x_2 := -\frac{1}{\omega}z'$ ergibt sich $x_1' = -\omega x_2$ und $x_2' = \omega x_1 - \frac{b}{\omega}\cos(\Omega t)$. Die spezielle Lösung ist für $\Omega \neq \pm\omega$ damit

$$\begin{aligned}
y(t) &= \int_{t_0}^{t} ds \begin{pmatrix} \cos\omega s & \sin\omega s \\ -\sin\omega s & \cos\omega s \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -\frac{b}{\omega}\cos\Omega s \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
&= -\frac{b}{2\omega} \int_{t_0}^{t} ds \begin{pmatrix} \sin(\omega s + \Omega s) + \sin(\omega s - \Omega s) \\ \cos(\omega s + \Omega s) + \cos(\omega s - \Omega s) \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\
&= \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} - \frac{b}{2\omega(\omega+\Omega)} \begin{pmatrix} -\cos(\omega t+\Omega t) \\ \sin(\omega t+\Omega t) \end{pmatrix} - \frac{b}{2\omega(\omega-\Omega)} \begin{pmatrix} -\cos(\omega t-\Omega t) \\ \sin(\omega t-\Omega t) \end{pmatrix} \\
&= \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} + \frac{b}{\omega(\omega^2 - \Omega^2)} \begin{pmatrix} \omega\cos\omega t\cos\Omega t + \Omega\sin\omega t\sin\Omega t \\ -\omega\sin\omega t\cos\Omega t + \Omega\cos\omega t\sin\Omega t \end{pmatrix} .
\end{aligned}$$

Damit ergibt sich

$$x(t) = \Phi(t) \cdot y(t) = \begin{pmatrix} c_1' \cos \omega t - c_2' \sin \omega t \\ c_1' \sin \omega t + c_2' \cos \omega t \end{pmatrix} + \frac{b}{(\omega^2 - \Omega^2)} \begin{pmatrix} \cos \Omega t \\ \frac{\Omega}{\omega} \sin \Omega t \end{pmatrix} .$$

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

$$\begin{aligned} y(t) &= -\frac{b}{2\omega} \int_{t_0}^{t} ds \begin{pmatrix} \sin(2\omega s) \\ \cos(2\omega s) + 1 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= \begin{pmatrix} c_1' \\ c_2' \end{pmatrix} - \frac{b}{4\omega^2} \begin{pmatrix} -\cos(2\omega t) \\ \sin(2\omega t) + 2\omega t \end{pmatrix} \end{aligned}$$

Daraus folgt

$$\begin{aligned} x(t) = \Phi(t) \cdot y(t) &= \begin{pmatrix} c_1' \cos \omega t - c_2' \sin \omega t \\ c_1' \sin \omega t + c_2' \cos \omega t \end{pmatrix} \\ &\quad + \frac{b}{4\omega^2} \begin{pmatrix} \cos \omega t \cos 2\omega t + \sin \omega t \sin 2\omega t + 2\omega t \sin \omega t \\ \sin \omega t \cos 2\omega t - \cos \omega t \sin 2\omega t - 2\omega t \cos \omega t \end{pmatrix} \\ &= \begin{pmatrix} (c_1' + \frac{b}{4\omega^2}) \cos \omega t - c_2' \sin \omega t \\ (c_1' - \frac{b}{4\omega^2}) \sin \omega t + c_2' \cos \omega t \end{pmatrix} + \frac{b}{2\omega} \begin{pmatrix} t \sin \omega t \\ -t \cos \omega t \end{pmatrix} . \end{aligned}$$

Die Amplitude wächst also mit der Zeit t an. Wir werden später eine einfachere Berechnungsmethode für diese Differentialgleichung angeben. $\triangleleft$

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

Definition 12.9 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

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

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

Satz 12.10 *In den Bezeichnungen von Definition 12.9 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 $L_{a,b} = \tilde{x} + V_a$, wobei $\tilde{x}$ eine beliebige Lösung der inhomogenen Differentialgleichung ist.*

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

$$W(t) := \det \begin{pmatrix} x_{(1)}(t) & x_{(2)}(t) & \dots & x_{(n)}(t) \\ x'_{(1)}(t) & x'_{(2)}(t) & \dots & x'_{(n)}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x^{(n-1)}_{(1)}(t) & x^{(n-1)}_{(2)}(t) & \dots & x^{(n-1)}_{(n)}(t) \end{pmatrix}$$

ungleich Null ist.

Beweis. Alle Aussagen folgen sofort aus der in Abschnitt 9 gegebenen Umschreibung

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

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

Beispiel 12.11 Gegeben sei die Differentialgleichung

$$x''(t) - \frac{1}{2t}x'(t) + \frac{1}{2t^2}x(t) = 1 \qquad \text{auf } I := \mathbb{R}^\times_+ \,.$$

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

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

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

$$W(t) = \det \begin{pmatrix} t & \sqrt{t} \\ 1 & \frac{1}{2\sqrt{t}} \end{pmatrix} = -\frac{1}{2}\sqrt{t} \neq 0 \,,$$

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

$$x(t) = \frac{2}{3}t^2 + c_1 t + c_2\sqrt{t}$$

als allgemeine Lösung der Differentialgleichung. ◁

13 Einige spezielle Funktionen

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

$$(1-t^2)x''(t) - 2t\,x(t) + n(n+1)x(t) = 0\;, \qquad n \in \mathbb{N}\;.$$

Satz 13.1 *Das Legendresche Polynom n-ter Ordnung*

$$P_n(t) := \frac{1}{2^n n!}\Big(\frac{d}{dt}\Big)^n (t^2-1)^n$$

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

$$\Big(\frac{d}{dt}\Big)^n (fg)(t) = \sum_{j=0}^{n} \binom{n}{j} f^{(j)}(t)\, g^{(n-j)}(t)\;.$$

Wir betrachten die $(n+1)$-ste Ableitung von $z(t) := (t^2-1)\frac{d}{dt}(t^2-1)^n = 2nt(t^2-1)^n$:

$$\begin{aligned}
&\frac{d^{n+1}}{dt^{n+1}}\Big((t^2-1)\frac{d}{dt}(t^2-1)^n\Big)\\
&= \binom{n+1}{0}(t^2-1)\frac{d^{n+2}}{dt^{n+2}}(t^2-1)^n + \binom{n+1}{1}2t\frac{d^{n+1}}{dt^{n+1}}(t^2-1)^n + \binom{n+1}{2}2\frac{d^n}{dt^n}(t^2-1)^n\\
&= (t^2-1)\frac{d^{n+2}}{dt^{n+2}}(t^2-1)^n + 2(n+1)t\frac{d^{n+1}}{dt^{n+1}}(t^2-1)^n + n(n+1)\frac{d^n}{dt^n}(t^2-1)^n\\
&\equiv \binom{n+1}{0}2nt\frac{d^{n+1}}{dt^{n+1}}(t^2-1)^n + \binom{n+1}{1}2n\frac{d^n}{dt^n}(t^2-1)^n\\
&= 2nt\frac{d^{n+1}}{dt^{n+1}}(t^2-1)^n + 2n(n+1)\frac{d^n}{dt^n}(t^2-1)^n\;.
\end{aligned}$$

Zusammenfassung der Terme liefert

$$0 = (1-t^2)P_n''(t) - 2tP_n'(t) + n(n+1)P_n(t)\;. \qquad \square$$

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 13.2 *Sei $I \subset \mathbb{R}$ ein Intervall und $a,b : I \to \mathbb{K}$ stetige Funktionen. Es sei $x_{(1)} : I \to \mathbb{K}$ eine Lösung der Differentialgleichung $x''(t)+a(t)x'(t)+b(t)x(t) = 0$ und in einem Intervall $J \subset I$ gelte $x_{(1)}(t) \neq 0$ für alle $t \in J$. Dann erhält man auf J eine zweite linear unabhängige Lösung $x_{(2)} : J \to \mathbb{K}$ der Differentialgleichung durch den Ansatz $x_{(2)}(t) = u(t)x_{(1)}(t)$, wobei u dann eine nichtkonstante Lösung der Differentialgleichung*

$$u''(t) + \Big(2\frac{x'_{(1)}(t)}{x_{(1)}(t)} + a(t)\Big)u'(t) = 0$$

ist, die im ersten Schritt zu

$$u'(t) = \frac{1}{x^2_{(1)}(t)} \exp\Big(-\int_{t_0}^t ds\; a(s)\Big)$$

integriert werden kann.

Sehr ähnlich werden folgende Differentialgleichungen behandelt:

Beispiel 13.3 Die *Hermitesche Differentialgleichung* zum Parameter $n \in \mathbb{N}$ ist

$$x''(t) - 2t\,x'(t) + 2nx(t) = 0\,, \qquad t \in \mathbb{R}\,,\; n \in \mathbb{N}\,.$$

Eine Lösung ist das *Hermitesche Polynom* n-ter Ordnung

$$H_n(t) := (-1)^n e^{t^2}\Big(\frac{d}{dt}\Big)^n e^{-t^2}\,.$$

Es tritt auf als Lösung der eindimensionalen Schrödinger-Gleichung für das Potential $U(t) = \frac{1}{2}\omega^2 t^2$ eines harmonischen Oszillators.

Beispiel 13.4 Die *Laguerresche Differentialgleichung* zum Parameter $n \in \mathbb{N}$ ist (mit r statt t als "Zeit")

$$rx''(r) + (1-r)x'(r) + nx(r) = 0\,, \qquad r \in \mathbb{R}^*_+\,,\; n \in \mathbb{N}\,.$$

Eine Lösung ist das *Laguerresche Polynom* n-ter Ordnung

$$L_n(r) := \frac{1}{n!} e^r \Big(\frac{d}{dr}\Big)^n \big(r^n e^{-r}\big) \; .$$

Es tritt auf als Lösung der zweidimensionalen Schrödinger-Gleichung in Radialkoordinaten für das Potential $U(r) = \frac{1}{2}\omega^2 r^2$ eines harmonischen Oszillators.

Beispiel 13.5 Die *hypergeometrische Differentialgleichung* zu den 3 Parametern $\alpha, \beta, \gamma \in \mathbb{R}$ ist

$$t(1-t)x'' + (\gamma - (\alpha+\beta+1)t)x' - \alpha\beta x = 0 \; , \qquad t \in \mathbb{C} \setminus \{0, 1, \infty\} \; .$$

Eine Lösung ist gegeben durch die *hypergeometrische Funktion*

$$F\Big(\begin{matrix}\alpha\ \beta \\ \gamma\end{matrix}\Big| t\Big) := 1 + \frac{\alpha\beta}{\gamma} \cdot \frac{t^1}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)} \cdot \frac{t^2}{2!} + \dots \; .$$

Ist α oder β eine negative ganze Zahl $-n$ (und γ geeignet), dann bricht die Reihe ab, und $F\big(\begin{smallmatrix}\alpha\ \beta \\ \gamma\end{smallmatrix}\big| t\big)$ wird ein Polynom n-ter Ordnung in t.

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

$$F\Big(\begin{matrix}\alpha\ \beta \\ \gamma\end{matrix}\Big| t\Big) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \int_0^1 ds\; s^{\beta-1}(1-s)^{\gamma-\beta-1}(1-ts)^{-\alpha} \; .$$

Beispiel 13.6 Die *Besselsche Differentialgleichung* zum Parameter $p \in \mathbb{R}$ ist (mit r statt t als "Zeit")

$$x''(r) + \frac{1}{r}x'(r) + \Big(1 - \frac{p^2}{r^2}\Big)x(r) = 0 \; , \qquad p \in \mathbb{R} \; .$$

Dabei ist zunächst $r \in \mathbb{R}^*_+$, jedoch kann man die Gleichung auf $r \in \mathbb{C} \setminus \{0\}$ 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 $u(r,\phi) = x(r)e^{ip\phi}$ die Besselsche Differentialgleichung. 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 $p \in \mathbb{R}$)

$$J_p(r) = \frac{r^p}{2^p} \sum_{k=0}^{\infty} (-1)^k \frac{r^{2k}}{2^{2k} k! \Gamma(p+k+1)} \; .$$

Eine zweite linear unabhängige Lösung ist die Neumannsche Funktion $N_p(r)$. Für $r \in \mathbb{C}$ sind die komplexen Linearkombinationen $H_p(r) = J_p(r) \pm iN_p(r)$ (Hankel-Funktionen) nützlich.

Es gibt nützliche Integraldarstellungen der Besselfunktion, z.B.

$$J_n(r) = \frac{1}{\pi} \int_0^{\pi} d\theta \; \cos(n\theta - r \sin\theta) , \qquad n \in \mathbb{N} .$$

Daraus folgt $|J_n(r)| \leq 1$ für alle $r \in \mathbb{R}$, 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.

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 $\mathbb{C}[\frac{d}{dt}]$ 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 $P_n(\frac{d}{dt}) = a_0 + a_1 \frac{d}{dt} + \cdots + a_n \frac{d^n}{dt^n}$ mit $n \leq k$, dann ist dieser Differentialoperator n-ter Ordnung $P_n(\frac{d}{dt})$ eine lineare Abbildung

$$P_n(\tfrac{d}{dt}) : \mathcal{C}^k(I) \to \mathcal{C}^{k-n}(I) , \qquad f(t) \mapsto a_0 f(t) + a_1 f'(t) + \cdots + a_n f^{(n)}(t) .$$

Wir können $a_n = 1$ annehmen. Insbesondere ist $P_n(\frac{d}{dt})$ 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 12.10 ist der Vektorraum $\ker(P_n(\frac{d}{dt})) \subset \mathcal{C}^\infty(I)$ der Lösungen x von $P_n(\frac{d}{dt})x = 0$ *endlich-dimensional.*

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

$$P_n(\tfrac{d}{dt}) e^{\lambda t} = P_n(\lambda) e^{\lambda t} .$$

Folglich gilt:

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

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 $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ habe. Dann bilden die Funktionen $x_{(1)}, \dots, x_{(n)} : \mathbb{R} \to \mathbb{C}$ mit*

$$x_{(k)}(t) := e^{\lambda_k t} , \qquad k = 1, \dots, n ,$$

ein Fundamentalsystem von Lösungen der Differentialgleichung

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

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

$$W(0) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix}$$

Das ist genau die Vandermonde-Determinate (Aufgabe 2a von Blatt 10 aus dem 2. Semester),

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

da die Nullstellen paarweise verschieden sind. □

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

Beispiel 14.3 Gegeben sei die Differentialgleichung $x'''(t) - 2x''(t) + x'(t) - 2x(t) = 0$. Sie schreibt sich als $P_3(\frac{d}{dx})y = 0$ mit

$$P_3(T) = T^3 - 2T^2 + T - 2 = (T-2)(T^2+1) = (T-2)(T-\mathrm{i})(T+\mathrm{i}) .$$

Alle Nullstellen von P_3 sind paarweise verschieden, so daß

$$x_{(1)}(t) = e^{\mathrm{i}t} , \qquad x_{(2)}(t) = e^{-\mathrm{i}t} , \qquad x_{(3)}(t) = e^{2t}$$

ein Lösungsfundamentalsystem bildet. Wegen

$$c_1 e^{\mathrm{i}t} + c_2 e^{-\mathrm{i}t} = (c_1 + c_2)\cos t + (ic_1 - ic_2)\sin t$$

bildet dann

$$x_{(1)}(t) = \cos t , \qquad x_{(2)}(t) = \sin t , \qquad x_{(3)}(t) = e^{2t}$$

ein reelles Fundamentalsystem. ◁

Ganz allgemein gilt: Ist $\lambda = \mathrm{i}\mu$ mit $\mu \in \mathbb{R}$ eine rein imaginäre Nullstelle eines rellen Polynoms $P(T) \in \mathbb{R}[T]$, dann ist auch $\bar{\lambda} = -\mathrm{i}\mu$ eine Nullstelle, so daß sich die Lösungen $e^{\pm \mathrm{i}\mu t}$ der entsprechenden Differentialgleichung äquivalent durch $\cos(\mu t)$ und $\sin(\mu t)$ ausdrücken lassen.

Beispiel 14.4 Wir erinnern noch einmal an die Schwingungsdifferentialgleichung $x''(t) + \omega^2 x(t) = 0$. Mit $P_2(T) = T^2 + \omega^2 = (T - \mathrm{i}\omega)(T + \mathrm{i}\omega)$ finden wir sofort das Lösungsfundamentalsystem $x_{(1)}(t) = \cos(\omega t)$ und $x_{(2)}(t) = \sin(\omega t)$. ◁

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*

$$(\tfrac{d}{dt} - \lambda)^k (f(t)\, e^{\lambda t}) = f^{(k)}(t)\, e^{\lambda t} \;.$$

Beweis. Für $k = 0$ ist nichts zu zeigen, und für $k = 1$ ergibt sich

$$(\tfrac{d}{dt} - \lambda)(f(t)\, e^{\lambda t}) = f'(t)\, e^{\lambda t} + f(t)\, (e^{\lambda t})' - \lambda\, f(t)\, e^{\lambda t} = f'(t)\, e^{\lambda t} \;.$$

Der Induktionsschritt $k \mapsto k + 1$ ist

$$\begin{aligned}(\tfrac{d}{dt} - \lambda)^{k+1}(f(t)\, e^{\lambda t}) &= (\tfrac{d}{dt} - \lambda)(f^{(k)}(t)\, e^{\lambda t}) \\ &= (f^{(k)}(t))'\, e^{\lambda t})\end{aligned}$$

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*

$$P_n(\tfrac{d}{dt})\big(g_k(t)\, e^{\lambda t}\big) = h_k(t)\, e^{\lambda t} \;,$$

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

$$P_n(\tfrac{d}{dt})\big(g_k(t)\,e^{\lambda t}\big) = \sum_{j=0}^{n} c_j(\tfrac{d}{dt} - \lambda)^j\big(g_k(t)\,e^{\lambda t}\big) = \sum_{j=0}^{n} c_j g_k^{(j)}(t)\,e^{\lambda t}\ .$$

Wegen $c_0 \neq 0$ ist $h_k(t) := \sum_{j=0}^{n} c_j g_k^{(j)}(t)$ 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 $P(\frac{d}{dt})x = 0$ ein Lösungs-Fundamentalsystem aus den Funktionen*

$$x_{(jm)}(t) := t^m e^{\lambda_j t}\ , \qquad 1 \leq j \leq r\ ,\ 0 \leq m \leq k_j - 1\ .$$

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

$$P_n(\tfrac{d}{dt})(t^m e^{\lambda_j t}) = Q_j(\tfrac{d}{dt})(\tfrac{d}{dt} - \lambda_j)^{k_j}(t^m e^{\lambda_j t}) = Q_j(\tfrac{d}{dt})\,(t^m)^{(k_j)} e^{\lambda_j t} = 0$$

wegen $m < k_j$. Somit erfüllen alle Funktionen $x_{(jm)}$ die Differentialgleichung.

ii) Zu zeigen bleibt die lineare Unabhängigkeit der $x_{(jm)}$. Eine Linearkombination der $x_{(jm)}$ hat die Form $\tilde{x}(t) = \sum_{j=1}^{r} g_{(j)}(t)e^{\lambda_j t}$, wobei $g_{(j)}(t)$ eine Polynomfunktion vom Grad $\leq k_j - 1$ ist (wir lassen den Polynomgrad zur Verbesserung der Lesbarkeit weg). Wir zeigen durch Induktion nach r, daß $\tilde{x} = 0$ genau dann, wenn $g_{(j)} = 0$ für alle j.

Für $r = 1$ folgt aus $\tilde{x}(t) = g_{(1)}(t)e^{\lambda_1 t} = 0$ für alle $t \in \mathbb{R}$, daß $g_{(1)}(t) = 0$ ist.

Im Schritt von $r-1$ nach r sei dann $\sum_{j=1}^{r} g_{(j)}(t)e^{\lambda_j t} = 0$. Ist eines der $g_{(j)}$ gleich Null, so sind wir nach Induktionsannahme fertig. Ansonsten wenden wir $(\frac{d}{dt} - \lambda_r)^{k_r}$ an und benutzen Lemma 14.5 und Lemma 14.6:

$$\begin{aligned} 0 &= (\tfrac{d}{dt} - \lambda_r)^{k_r}\Big(\sum_{j=1}^{r} g_{(j)}(t)e^{\lambda_j t}\Big) = (\tfrac{d}{dt} - \lambda_r)^{k_r}\Big(\sum_{j=1}^{r-1} g_{(j)}(t)e^{\lambda_j t}\Big) \\ &= \sum_{j=1}^{r-1}\sum_{i=0}^{k_r} c_{ji}(\tfrac{d}{dt} - \lambda_j)^i\Big(g_{(j)}(t)e^{\lambda_j t}\Big) = \sum_{j=1}^{r-1} \underbrace{\Big(\sum_{i=0}^{k_r} c_{ji} g_{(j)}^{(i)}(t)\Big)}_{h_{(j)}(t)} e^{\lambda_j t}\ , \end{aligned}$$

wobei $h_{(j)}(t)$ wegen $c_{j0} = (\lambda_j - \lambda_r)^{k_r} \neq 0$ wieder ein Polynom vom Grad $\leq k_j - 1$ ist. Nach Induktionsvoraussetzung ist dann $h_{(j)} = 0$ für alle j und weiter wegen $c_{j0} \neq 0$ auch $g_{(j)}^{(0)}(t) = 0$, im Widerspruch zur Annahme. □

Beispiel 14.8 Gegeben sei die Schwingungsdifferentialgleichung mit kritischer Reibung

$$0 = x''(t) + 2\omega x'(t) + \omega^2 x(t) = P_2(\tfrac{d}{dt})(x)$$

mit $P_2(T) = T^2 + 2\omega T + \omega^2 = (T+\omega)^2$. Die allgemeine Lösung ist deshalb $x(t) = (c_1 + c_2 t)e^{-\omega t}$. $\triangleleft$

Wir betrachten nun inhomogene lineare Differentialgleichungen $P_n(\frac{d}{dt})x = b(t)$. 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 $b(t) = b_1(t) + \cdots + b_k(t)$ und $x_{(k)}$ Lösung von $P_n(\frac{d}{dt})x_{(k)} = b_k(t)$, so ist $x = x_{(1)} + \cdots + x_{(k)}$ Lösung von $P_n(\frac{d}{dt})x = b(t)$. Der Lösungsansatz funktioniert dann für rechte Seiten der Form $b(t) = g_m(t)e^{\mu t}$, 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 $P_n(\frac{d}{dt})x = e^{\mu t}$ besitzt die spezielle Lösung $x(t) = \frac{1}{P_n(\mu)}e^{\mu t}$.*

ii) *Ist $g_m : \mathbb{R} \to \mathbb{C}$ eine Polynomfunktion vom Grad m, so besitzt die Differentialgleichung $P_n(\frac{d}{dt})x = g_m(t)e^{\mu t}$ eine spezielle Lösung der Form $x(t) = h_m(t)e^{\mu t}$, wobei $h_m : \mathbb{R} \to \mathbb{C}$ wieder eine Polynomfunktion vom Grad m ist.*

Beweis. i) ist klar wegen $P_n(\frac{d}{dt})e^{\mu t} = P_n(\mu)e^{\mu t}$.

ii) mit Induktion nach m. Der Fall $m = 0$ ist Teil i). Nach Lemma 14.6 ist $P_n(\frac{d}{dt})(t^m e^{\mu t}) = \tilde{h}_m(t)e^{\mu t}$ für eine Polynomfunktion $\tilde{h}_m(t)$ vom Grad m. Wir schreiben $g_m(t) = c\tilde{h}_m(t) + g_{m-1}(t)$ für ein $c \in \mathbb{C}$ und ein Polynom $g_{m-1}(t)$ vom Grad $m-1$. Nach Induktionsvoraussetzung gibt es eine Polynomfunktion $h_{m-1}(t)$ vom Grad $m-1$, so daß $P_n(\frac{d}{dt})x = g_{m-1}(t)e^{\mu t}$ die spezielle Lösung $x(t) = h_{m-1}(t)e^{\mu t}$ hat. Dann hat $P_n(\frac{d}{dt})x = g_m(t)e^{\mu t}$ die spezielle Lösung $x(t) = (ct^m + h_{m-1}(t))e^{\mu t}$. $\square$

Beispiel 14.10 Gegeben sei die Differentialgleichung $x'''(t) - x(t) = t$, also $P_3(\frac{d}{dt})x = b(t)$ mit $P_3(T) = T^3 - 1 = (T-1)(T^2+T+1) = (T-1)(T-\tau_1)(T-\tau_2)$ und $b(t) = te^{0x}$. Damit ist $P_3(0) \neq 0$, so daß es eine spezielle Lösung $x(t) = c_1 t + c_0$ der Differentialgleichung gibt. Wir testen

$$(\frac{d^3}{dt^3} - 1)(c_1 t + c_0) = -c_1 t - c_0 \stackrel{!}{=} t\ ,$$

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

$$x(t) = -t + ae^t + e^{-\frac{1}{2}t}\Big(b\cos\frac{\sqrt{3}}{2}t + c\sin\frac{\sqrt{3}}{2}t\Big) . \qquad \triangleleft$$

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 $P_n(\frac{d}{dt})y = g_m(t)e^{\mu t}$ eine spezielle Lösung der Form $x(t) = h_{m+k}(t)e^{\mu t}$, wobei $h_{k+m}(t) = \sum_{j=k}^{m+k} c_jt^j$ eine Polynomfunktion vom Grad $m+k$ ist, in der die untersten Potenzen t^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 $\tilde{h}_m(t)$, so daß $Q_{n-k}(\frac{d}{dt})(\tilde{h}_m(t)e^{\mu t}) = g_m(t)e^{\mu t}$. Es gibt dann eine Polynomfunktion $h_{m+k}(t) = \sum_{j=k}^{m+k} c_jt^j$, so daß $h^{(k)}_{m+k}(t) = \tilde{h}_m(t)$. Nach Lemma 14.5 gilt damit

$$\begin{aligned}
P_n(\tfrac{d}{dt})\big(h_{m+k}(t)e^{\mu t}\big) &= Q_{n-k}(\tfrac{d}{dt})(\tfrac{d}{dt}-\mu)^k\big(h_{m+k}(t)e^{\mu t}\big) \\
&= Q_{n-k}(\tfrac{d}{dt})\big(h^{(k)}_{m+k}(t)e^{\mu t}\big) = Q_{n-k}(\tfrac{d}{dt})\big(\tilde{h}_m(t)e^{\mu t}\big) \\
&= g_m(t)e^{\mu t} .
\end{aligned}$$

□

Beispiel 14.12 Wir betrachten noch einmal die Differentialgleichung $x''(t) + \omega^2x(t) = b\cos(\Omega t) = \mathrm{Re}(be^{i\Omega t})$. 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

$$x(t) = c_1e^{i\omega t} + c_2e^{-i\omega t} + \frac{b}{\omega^2-\Omega^2}e^{i\Omega t} , \qquad c_1, c_2 \in \mathbb{C} .$$

Der Realteil ist

$$\mathrm{Re}(x(t)) = a_1\cos\omega t + a_2\sin\omega t + \frac{b}{\omega^2-\Omega^2}\cos\Omega t , \quad a_1 = \mathrm{Re}(c_1+c_2) , \; a_2 = \mathrm{Im}(c_2-c_1) .$$

Im Resonanzfall $\Omega = \pm\omega$ ist Ω eine einfache Nullstelle, so daß eine spezielle Lösung die Form $cte^{i\omega t}$ hat. Wegen

$$(\tfrac{d^2}{dt^2} + \omega^2)(cte^{i\omega t}) = 2ic\omega e^{i\omega t}$$

ergibt sich

$$x(t) = c_1e^{i\omega t} + c_2e^{-i\omega t} + \frac{bt}{2i\omega}e^{i\omega t} , \qquad c_1, c_2 \in \mathbb{C} .$$

Der Realteil ist

$$\mathrm{Re}(x(t)) = a_1\cos\omega t + a_2\sin\omega t + \frac{b}{2\omega}t\sin\omega t , \qquad a_1 = \mathrm{Re}(c_1+c_2) , \; a_2 = \mathrm{Im}(c_2-c_1) .$$

Wir bestätigen damit die zuvor in Beispiel 12.8 erhaltenen Lösungen. ◁

Wiederholung

- Umformungen zwischen expliziter Differentialgleichung und vektorieller DGL $x'(t) = v(t, x(t))$; Interpretation von v als Geschwindigkeitsfeld und Lösungen $x(t)$ als Bahnkurven, Stromlinien, Feldlinien
- DGL mit getrennten Variablen, lineare DGL, Variation der Konstanten
- Bezeihung zwischen Anfangswertproblem $x'(t) = v(t, x(t))$ mit $x(t_0) = x_0$ und Integralgleichung; lokale Existenz und Eindeutigkeit unter Lipschitz-Bedingung
- Lineare DGL: Lösungsfundamentalsystem, Variation der Konstanten im inhomogenen Fall
- Lineare DGL mit konstanten Koeffizienten: Ansatz für Lösungsfundamentalsystem und spezielle Inhomogenitäten mit Fallunterscheidungen nach Vielfachheit der Nullstellen

Teil III
Grundlagen der Funktionentheorie

15 Differentialformen und Kurvenintegrale

Definition 15.1 Sei $U \subset \mathbb{R}^n$ offen. Unter einer *Differentialform 1. Grades* bzw. einer *1-Form* auf U versteht man eine Abbildung

$$\omega : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{K}) , \qquad \mathbb{K} = \mathbb{R} \text{ oder } \mathbb{K} = \mathbb{C} .$$

Beispiel 15.2 Sei $f : U \to \mathbb{R}$ stetig differenzierbar. Dann ist $\omega = Df$ eine 1-Form, $\omega(\xi) = (Df)(\xi) : \mathbb{R}^n \to \mathbb{R}$ linear. ◁

Wir bezeichnen hier mit $\xi = (\xi_1, \dots, \xi_n)$ die Koordinaten auf U, da die Bezeichnung $x_i \in \mathcal{C}(U)$ für die *Koordinatenfunktionen* reserviert ist:

$$x_i(\xi) = \xi_i .$$

Dann ist $x_i(\xi + h) = \xi_i + h_i = x_i(\xi) + (Dx_i)(\xi) \circ h$. Für diese Differentiale Dx_i der Koordinatenfunktionen führt man die besondere Bezeichnung dx_i ein,

$$dx_i := Dx_i \quad \Rightarrow \quad dx_i(\xi) \circ h = h_i \quad \forall \xi \in U .$$

Beispiel 15.3 Sei $v : U \to \mathbb{R}^n$ ein Vektorfeld. Dann wird durch $\omega = \langle v, \ . \ \rangle$, d.h $\omega(\xi) \circ h = \langle v(\xi), h \rangle$ für $h \in \mathbb{R}^n$, eine 1-Form definiert. ◁

Wegen der Linearität ist eine 1-Form ω durch ihre Werte auf den Basisvektoren $(e_i)_{i=1,\dots,n}$ des $\mathbb{R}^n$ bestimmt, also durch die Funktionen $\omega \circ e_i = \omega_i : U \to \mathbb{K}$. Dann ist

$$\omega(\xi) \circ \Big(\sum_{i=1}^n h_i e_i \Big) = \sum_{i=1}^n h_i \omega(\xi) \circ e_i = \sum_{i=1}^n h_i \omega_i(\xi) = \sum_{i=1}^n \omega_i(\xi) \ (dx_i)(\xi) \circ h$$

$$\Rightarrow \quad \omega(\xi) = \sum_{i=1}^n \omega_i(\xi) dx_i(\xi) , \qquad dx_i \circ e_j = \delta_{ij} .$$

Beispiel 15.4

i) Ist $f : U \to \mathbb{R}$ stetig differenzierbar, dann gilt $Df = \sum_{i=1}^n (\partial_i f) dx_i$

ii) Ist $v = \sum_{i=1}^n v_i e_i$ ein stetiges Vektorfeld, dann gilt für die durch $\omega = \langle v, \ . \ \rangle$ definierte 1-Form $\omega = \sum_{i=1}^n v_i dx_i$.

1-Formen im $\mathbb{R}^n$ können über Kurven integriert werden. Dadurch läßt sich z.B. die Arbeit definieren, welche ein Kraftfeld an einem Massepunkt verrichtet, der sich auf einer Bahnkurve c bewegt. Ist die Bahn geradlinig mit $c(t) = c_0 + vt$, wobei $c_0, v \in \mathbb{R}^n$ und $t \in [\alpha, \beta]$, und ist das Kraftfeld $F(c(t)) = F \in \mathbb{R}^n$ konstant, dann ist die Arbeit erklärt als

$$W = \langle F, v(\beta - \alpha)\rangle = \int_\alpha^\beta dt \langle F, c'(t)\rangle .$$

Es ist deshalb sinnvoller, das Kraftfeld als eine 1-Form aufzufassen, welches eine lineare Abbildung des Tangentialraums nach $\mathbb{R}$ implementiert. Als Definition der Arbeit, die auch für nichtkonstante Kräfte sinnvoll bleibt, bietet sich an:

$$W(F, c) = \int_\alpha^\beta dt\; F(c(t)) \circ c'(t) , \qquad F : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}) .$$

Definition 15.5 Es sei $c : [\alpha, \beta] \to U \subset \mathbb{R}^n$ eine Kurve im Definitionsbereich einer 1-Form $\omega : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{K})$. Die 1-Form ω heißt *längs c integrierbar*, wenn es eine Zahl $I \in \mathbb{K}$ gibt mit folgender Eigenschaft: Zu jedem $\epsilon > 0$ existiert ein $\delta > 0$, so daß für eine beliebige Zerlegung $\alpha \leq t_0 < t_1 < \cdots < t_m \leq \beta$ mit $|t_k - t_{k-1}| < \delta$ und beliebige Wahl von $\tau_k \in [t_{k-1}, t_k]$ gilt

$$\Big| I - \sum_{k=1}^m \omega(c(\tau_k)) \circ (c(t_k) - c(t_{k-1}))\Big| < \epsilon .$$

In diesem Fall heißt $I =: \int_c \omega$ das Integral der 1-Form ω längs c.

Es stellt sich heraus, daß selbst für stetige 1-Formen ω mehr als Stetigkeit der Kurve c gefordert werden muß, damit ω längs c integriert werden kann. Es genügt (stückweise) Differenzierbarkeit der Kurve:

Satz 15.6 *Ist $\omega = \sum_{i=1}^n \omega_i dx_i$ stetig und $c = (c_1, \ldots, c_n)$ stetig differenzierbar, dann ist ω längs c integrierbar, und es gilt*

$$\int_c \omega = \int_\alpha^\beta dt\; \omega(c(t)) \circ c'(t) = \sum_{i=1}^n \int_\alpha^\beta dt\; \omega_i(c(t)) c_i'(t) .$$

Beweis. Nach Satz 7.4 besitzt c eine Bogenlänge $L(c)$. Die stetigen Funktionen $\omega_i \circ c : [\alpha, \beta] \to \mathbb{K}$ sind gleichmäßig stetig. Wähle $\delta > 0$ derart, daß für alle $t, t' \in [\alpha, \beta]$ mit $|t - t'| < \delta$ gilt

$$\big|\omega_i(c(t)) - \omega_i(c(t'))\big| < \frac{\epsilon}{nL(c)} , \qquad \forall i = 1, \ldots, n .$$

Wähle eine Unterteilung $\alpha \leq t_0 < t_1 < \cdots < t_m \leq \beta$ mit $|t_k - t_{k-1}| < \delta$ und $\tau_k \in [t_{k-1}, t_k]$. Dann gilt

$$\begin{aligned}
&\Big|\sum_{i=1}^{n}\Big(\int_\alpha^\beta dt\;\omega_i(c(t))c_i'(t) - \sum_{k=1}^{m}\omega_i(c(\tau_k))(c_i(t_k) - c_i(t_{k-1}))\Big)\Big| \\
&\Big|\sum_{i=1}^{n}\sum_{k=1}^{m}\int_{t_{k-1}}^{t_k} dt\;\big(\omega_i(c(t)) - \omega_i(c(\tau_k))\big)c_i'(t)\Big| \\
&\leq \sum_{i=1}^{n}\sum_{k=1}^{m}\int_{t_{k-1}}^{t_k} dt\;\big|\omega_i(c(t)) - \omega_i(c(\tau_k))\big||c_i'(t)| \\
&< \sum_{i=1}^{n}\sum_{k=1}^{m}\frac{\epsilon}{nL(c)}\int_{t_{k-1}}^{t_k} dt\;|c_i'(t)| \leq \sum_{i=1}^{n}\sum_{k=1}^{m}\frac{\epsilon}{nL(c)}\int_{t_{k-1}}^{t_k} dt\;\|c'(t)\| \\
&\leq \frac{\epsilon}{L(c)}\int_\alpha^\beta dt\;\|c'(t)\| = \epsilon\;.
\end{aligned}$$

Dabei haben wir Cauchy-Schwarz $|c_i'| = |\langle c', e_i\rangle \leq \|c'\|\|e_i\|$ benutzt sowie die Formel für die Bogenlänge aus Satz 7.4. □

Beispiel 15.7 Das Integral der durch ein homogenes Gravitationsfeld gegebenen 1-Form $\kappa = -mg\,dx_3$ soll längs der Kurven $c, \gamma : [0,1] \to \mathbb{R}^3$ mit $c(t) = (1 - \cos(\pi t))e_1 + \sin(\pi t)e_2 + h(1-t)e_3$ und $\gamma(t) = 2te_1 + h(1-t^2)e_3$ berechnet werden. Es gilt $c(0) = \gamma(0) = (0,0,h)$ und $c(1) = \gamma(1) = (2,0,0)$ sowie

$$\int_c \kappa = -mg\int_0^1 dt\;c_3'(t) = mgh\;, \qquad \int_\gamma \kappa = -mg\int_0^1 dt\;\gamma_3'(t) = mgh\;. \tag{5}$$

Für dieses Beispiel ist die Arbeit unabhängig vom Weg. ◁

Beispiel 15.8 Auf $U = \mathbb{R}^2 \setminus \{0\}$ heißt $\omega(\xi) := -\frac{\xi_2}{\xi_1^2+\xi_2^2}dx_1 + \frac{\xi_1}{\xi_1^2+\xi_2^2}dx_2$ die *Windungsform*. Wir integrieren ω längs des geschlossenen Weges $c : [0, 2\pi] \to U$, mit $c(t) = e_1\cos t + e_2 \sin t$. Dann ist $c'(t) = -e_1 \sin t + e_2 \cos t$, somit

$$\omega(c(t)) \circ c'(t) = \sin^2 t + \cos^2 t = 1 \qquad \Rightarrow \qquad \int_c \omega = 2\pi\;.$$ ◁

16 Exakte 1-Formen

Definition 16.1 Sei $U \subset \mathbb{R}^n$ offen. Unter einer *Stammfunktion (bzw. Potential) einer 1-Form* $\omega = \sum_{i=1}^n \omega_i dx_i : U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{K})$ versteht man eine differenzierbare Funktion $F : U \to \mathbb{K}$ mit $\omega = DF(= dF)$, d.h. $(\partial_i F)(\xi) = \omega_i(\xi)$ für alle $\xi \in U$ und $i = 1, \ldots, n$. Eine 1-Form ω, die auf U eine Stammfunktion besitzt, heißt *exakt*.

Nach dem Hauptsatz der Differential- und Integralrechnung besitzt jede stetige 1-Form fdx auf einem *Intervall* $I \subset \mathbb{R}$ eine Stammfunktion F mit $F' = f$. Im Höherdimensionalen gibt es nicht-exakte 1-Formen. Es wird sich zeigen, daß die Exaktheit sogar vom Definitionsbereich U abhängt. Wenn U zusammenhängend ist, dann unterscheiden sich zwei Stammfunktionen (falls es sie gibt) nach Satz 2.15 nur um eine Konstante. Für exakte 1-Formen gilt das Analogon des Hauptsatzes der Differential- und Integralrechnung:

Satz 16.2 *Ist F Stammfunktion einer stetigen 1-Form ω auf U, so gilt für jede (stückweise) stetig differenzierbare Kurve $c : [\alpha, \beta] \to U$*

$$\int_c \omega = F(c(\beta)) - F(c(\alpha)) .$$

Insbesondere ist $\int_c \omega = 0$ für jede geschlossene Kurve $c : I \to U$.

Beweis. Nach Voraussetzung ist $F \circ c : [\alpha, \beta] \to \mathbb{R}$ differenzierbar auf $[\alpha, \beta]$ mit $(F \circ c)'(t) = (DF)(c(t)) \circ c'(t) = \omega(c(t)) \circ c'(t)$. Somit gilt

$$F(c(\beta)) - F(c(\alpha)) = \int_\alpha^\beta dt\ (F \circ c)'(t) = \int_\alpha^\beta dt\ \omega(c(t)) \circ c'(t) = \int_c \omega . \qquad \square$$

Beispiel 16.3 Die Newtonsche Gravitationskraft einer Punktmasse M in $0 \in \mathbb{R}^3$ wirkend auf eine Punktmasse m in $\xi \in U = \mathbb{R}^3 \setminus \{0\}$ ist gegeben durch die 1-Form $\omega(\xi) = -\frac{\gamma Mm}{\|\xi\|^3} \sum_{i=1}^3 \xi_i dx_i$. Diese 1-Form ist exakt mit Potential $\Phi = \frac{\gamma Mm}{\|\xi\|}$. Somit gilt für die Arbeit der Gravitationskraft längs einer beliebigen Bahnkurve c zwischen $\xi, \eta \in U$ die Formel $\int_c \omega = \Phi(\eta) - \Phi(\xi) = \gamma Mm\big(\frac{1}{\|\eta\|} - \frac{1}{\|\xi\|}\big)$. Insbesondere verrichtet die Gravitationskraft keine Arbeit auf einer geschlossenen Bahnkurve. ◁

Beispiel 16.4 Die Windungsform aus Beispiel 15.8 kann nicht exakt sein, da das Integral über S^1 nicht verschwindet. ◁

Für eine exakte 1-Form ω auf U hängt das Integral $\int_c \omega$ nur von Anfangs- und Endpunkt der Kurve ab und nicht vom Wegverlauf dazwischen. Wie bei eindimensionalen Integralen ist es deshalb für exakte ω sinnvoll zu schreiben $\int_a^b \omega := \int_c \omega$, wobei für $c : [\alpha, \beta] \to U$ eine beliebige Kurve mit $c(\alpha) = a$ und $c(\beta) = b$ gewählt werden kann.

In Umkehrung von Satz 16.2 garantiert Wegunabhängigkeit der Integration die Existenz der Stammfunktion.

Satz 16.5 *Eine stetige 1-Form ω auf einer zusammenhängenden offenen Menge $U \subset \mathbb{R}^n$, die in U wegunabhängig integriert werden kann, besitzt in U die (bis auf Addition einer Konstanten eindeutige) Stammfunktion $F(\xi) = \int_a^\xi \omega$*

Beweis. Es gibt eine offene Kugel $K_r(\xi) \subset U$. Dann liegt für jedes $h \in \mathbb{R}^n$ mit $\|h\| < r$ die Kurve $c(t) = \xi + ht$ für $t \in [0,1]$ in U, und es gilt

$$\begin{aligned} F(\xi+h) - F(\xi) - \omega(\xi) \circ h &= \Big(\int_0^1 dt\; \omega(\xi+ht) \circ h \Big) - \omega(\xi) \circ h \\ &= \int_0^1 dt\; \big(\omega(\xi+ht) - \omega(\xi)\big) \circ h\;. \end{aligned}$$

In der Standardbasis ist $\big(\omega(\xi+ht) - \omega(\xi)\big) \circ h = \sum_{i=1}^n \big(\omega_i(\xi+ht) - \omega_i(\xi)\big) h_i$. Wegen der Stetigkeit der ω_i folgt $F(\xi+h) - F(\xi) - \omega(\xi) \circ h = o(\|h\|)$. □

Auch wenn zunächst unklar ist, ob eine 1-Form eine Stammfunktion besitzt, kann man zunächst das Integral $F(\xi) = \int_a^\xi \omega$ berechnen und nachträglich überprüfen, ob $dF = \omega$ gilt.

Beispiel 16.6 Zur Berechnung der Stammfunktion des homogenen Gravitationsfeldes $\omega = -mg dx_3$ im Punkt $\xi = (\xi_1, \xi_2, \xi_3)$ integrieren wir ω entlang der Kurve $c(t) = \xi_1 t e_1 + \xi_2 t x_2 + \xi_3 t e_3$, $t \in [0,1]$. Es gilt mit $c'(t) = \sum_{i=1}^n \xi_i e_i$

$$\Phi(\xi) := \int_0^\xi \omega = \int_0^1 dt\; (-mg)\xi_3 = -mg\xi_3\;.$$

Wie erwartet ist $D\Phi = \omega$. ◁

Eine 1-Form $\omega = \sum_{i=1}^n \omega_i dx^i$ sei stetig differenzierbar. Falls $\omega = DF$ das Differential einer Stammfunktion F ist, d.h. $\omega_i = \partial_i F$, dann muß nach dem Satz von Schwarz gelten

$$\partial_j \partial_k F = \partial_k \partial_j F \qquad \Rightarrow \qquad \partial_j \omega_k = \partial_k \omega_j \qquad \text{für alle } j,k = 1, \dots n\;.$$

Diese $\frac{n(n-1)}{2}$ unabhängigen Gleichungen heißen *Integrationsbedingungen*. Wir führen folgende Bezeichnung ein:

Definition 16.7 Eine differenzierbare 1-Form $\omega = \sum_{i=1}^n \omega_i dx^i$ heißt *geschlossen*, wenn für alle $j,k = 1, \dots, n$ gilt $(d\omega)_{jk} := \partial_j \omega_k - \partial_k \omega_j = 0$.

Damit haben wir gezeigt:

Satz 16.8 *Notwendig für die Exaktheit einer differenzierbaren 1-Form* $\omega = \sum_{i=1}^n \omega_i dx_i$ *ist, daß* ω *geschlossen ist,* $(d\omega)_{jk} := \partial_j \omega_k - \partial_k \omega_j = 0$ *für alle* $j,k = 1, \dots, n$. □

Auf Teilmengen $U \subset \mathbb{R}^2$ ist nur die Bedingung $\partial_1 \omega_2 - \partial_2 \omega_1 = 0$ zu überprüfen.

Beispiel 16.9 Die Windungsform aus Beispiel 15.8 ist in jedem Punkt $\xi \in \mathbb{R}^2 \setminus \{0\}$ geschlossen:

$$\begin{aligned}
\partial_2\omega_1 &= \partial_2\Big(-\frac{\xi_2}{\xi_1^2+\xi_2^2}\Big) = -\frac{1}{\xi_1^2+\xi_2^2} + \frac{\xi_2\cdot 2\xi_2}{(\xi_1^2+\xi_2^2)^2} = \frac{\xi_2^2-\xi_1^2}{(\xi_1^2+\xi_2^2)^2}\;,\\
\partial_1\omega_2 &= \partial_1\Big(+\frac{\xi_1}{\xi_1^2+\xi_2^2}\Big) = \frac{1}{\xi_1^2+\xi_2^2} - \frac{\xi_1\cdot 2\xi_1}{(\xi_1^2+\xi_2^2)^2} = \frac{\xi_2^2-\xi_1^2}{(\xi_1^2+\xi_2^2)^2}\;.
\end{aligned}$$

Da die Windungsform nicht exakt ist, ist Geschlossenheit (im allgemeinen) nicht hinreichend für Exaktheit! ◁

Für Teilmengen $U \subset \mathbb{R}^3$ sind drei Bedingungen zu überprüfen:

$$\begin{aligned}
(\operatorname{rot} v)_1 &:= (d\omega)_{23} = \partial_2\omega_3 - \partial_3\omega_2 = 0\;,\\
(\operatorname{rot} v)_2 &:= (d\omega)_{31} = \partial_3\omega_1 - \partial_1\omega_3 = 0\;,\\
(\operatorname{rot} v)_3 &:= (d\omega)_{12} = \partial_1\omega_2 - \partial_2\omega_1 = 0\;.
\end{aligned}$$

Die dabei auftretende Konstruktion $\operatorname{rot} v$ heißt die *Rotation* des Vektorfeldes $v = \sum_{i=1}^3 \omega_i e_i$ zur Differentialform $\omega = \sum_{i=1}^3 \omega_i dx_i = \langle v,\ .\rangle$. Damit also das Vektorfeld v Gradient eines Potentials sein kann, muß zumindest die Rotation von v verschwinden. Das Beispiel der Windungsform, welches dreidimensionale Analoga besitzt, zeigt jedoch, daß $\operatorname{rot} v = 0$ noch nicht hinreichend ist, damit sich die zugehörige Differentialform $\omega = \langle v,\ .\rangle$ wegunabhängig integrieren läßt. Die Integrabilitätsbedingungen sind jedoch hinreichend für Exaktheit im Falle von *Sterngebieten*:

Satz 16.10 (Lemma von Poincaré) *Es sei $U \subset \mathbb{R}^n$ ein Sterngebiet, d.h. es gibt einen Punkt $a \in U$ (das Zentrum), so daß für jeden Punkt $\xi \in U$ die Verbindungsstrecke $c_\xi(t) := a + (\xi - a)t$ für $t \in [0,1]$ vollständig in U liegt. Ist eine differenzierbare 1-Form ω geschlossen auf dem offenen Sterngebiet U, so ist sie auf U auch exakt.*

Beweis. Für $\xi \in U$ setzen wir

$$F(\xi) := \int_{c_\xi} \omega = \int_0^1 dt\ \omega(a + (\xi - a)t) \circ (\xi - a)$$

Wegen der Offenheit gibt es ein $r > 0$, so daß $\xi + \eta, c_{(\xi+\eta)(t)} \in U$ für $\|\eta\| < r$ und $t \in [0,1]$. Dann gilt für die partiellen Ableitungen (mit $|h| < r$)

$$\begin{aligned}
(\partial_j F)(\xi) &= \lim_{h\to 0} \frac{1}{h}\big(F(\xi + he_j) - F(\xi)\big)\\
&= \lim_{h\to 0} \frac{1}{h} \int_0^1 dt\ \Big(\omega(a + (\xi{+}he_j{-}a)t) \circ (\xi{+}he_j{-}a) - \omega(a + (\xi{-}a)t) \circ (\xi{-}a)\Big)
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{h\to 0} \int_0^1 dt \sum_{i=1}^n \frac{\omega_i(a+(\xi+he_j-a)t) - \omega_i(a+(\xi-a)t)}{h}(\xi_i+h\delta_{ij}-a_i) \\
&+ \lim_{h\to 0} \int_0^1 dt\, \omega_j(a+(\xi+he_j-a)t) \\
&= \int_0^1 dt \sum_{i=1}^n t(\partial_j\omega_i)(a+(\xi-a)t)\,(\xi_i-a_i) + \int_0^1 dt\, \omega_j(a+(\xi-a)t)\,.
\end{aligned}
$$

Gelten die Integrationsbedingungen, so folgt mit Kettenregel $(\omega_j \circ c)'(t) = (D\omega_j)(c(t)) \circ c'(t)$

$$
\begin{aligned}
(\partial_j F)(\xi) &= \int_0^1 dt \sum_{i=1}^n t(\partial_i\omega_j)(a+(\xi-a)t)\,(\xi_i-a_i) + \int_0^1 dt\, \omega_j(a+(\xi-a)t) \\
&= \int_0^1 dt\, t\frac{d}{dt}\big(\omega_j(a+(\xi-a)t)\big) + \int_0^1 dt\, \omega_j(a+(\xi-a)t) \\
&= \int_0^1 dt\, \frac{d}{dt}\big(t\omega_j(a+(\xi-a)t)\big) \\
&= \big(t\omega_j(a+(\xi-a)t)\big)\Big|_0^1 = \omega_j(\xi)\,.
\end{aligned}
$$

□

Beispiel 16.11 Da die Windungsform ω geschlossen ist, ist sie nach dem Lemma von Poincaré auf Sterngebieten exakt. Ein solches ist z.B. gegeben durch die geschlitzte Ebene $U = \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\})$ (die Punkte $(b,0)$ mit $b \leq 0$ fehlen). Jeder Punkt $(a,0)$ mit $a \geq 0$ ist mögliches Zentrum des Sterngebietes U. Legen wir das Potential in $(1,0)$ fest zu $F(1,0) = 0$, dann ist das Potential in $(\xi_1,\xi_2) = re^{\mathrm{i}\phi}$, mit $|\phi| < \pi$ und $r > 0$ gegeben durch Integration von ω z.B. längs der reellen Achse zum Punkt $(r,0)$ und dann längs des Kreisbogens von $re^{\mathrm{i}0}$ nach $re^{\mathrm{i}\phi}$:

$$
\begin{aligned}
F(re^{\mathrm{i}\phi}) &= \int_0^1 dt\, \omega_1((1+(r-1)t,0))\cdot(r-1) \\
&+ \int_0^1 dt\, \Big(\omega_1((r\cos(\phi t), r\sin(\phi t)))\cdot(-r\phi\sin(\phi t)) \\
&\qquad\qquad + \omega_2((r\cos(\phi t), r\sin(\phi t)))\cdot(r\phi\cos(\phi t))\Big) \\
&= \int_0^1 dt\, \Big(\phi\sin^2(\phi t) + \phi\cos^2(\phi t)\Big) = \phi\,.
\end{aligned}
$$

Zurückübersetzt in kartesische Koordinaten ist $F(\xi_1,\xi_2) = \arccos\frac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}}$. Dann ist für $\xi_2 \neq 0$

$$
\begin{aligned}
(\partial_1 F)(\xi_1,\xi_2) &= \frac{1}{-\sin\big(\arccos\frac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}}\big)}\cdot\Big(\frac{1}{\sqrt{\xi_1^2+\xi_2^2}} - \frac{\xi_1^2}{(\sqrt{\xi_1^2+\xi_2^2})^3}\Big) \\
&= -\frac{\sqrt{\xi_1^2+\xi_2^2}}{\xi_2}\cdot\frac{\xi_2^2}{(\sqrt{\xi_1^2+\xi_2^2})^3} = -\frac{\xi_2}{\xi_1^2+\xi_2^2}\,,
\end{aligned}
$$

$$
\begin{aligned}
(\partial_2 F)(\xi_1,\xi_2) &= \frac{1}{-\sin\big(\arccos\frac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}}\big)}\cdot\Big(-\frac{\xi_1\xi_2}{(\sqrt{\xi_1^2+\xi_2^2})^3}\Big)\\
&= \frac{\sqrt{\xi_1^2+\xi_2^2}}{\xi_2}\cdot\frac{\xi_1\xi_2}{(\sqrt{\xi_1^2+\xi_2^2})^3} = \frac{\xi_1}{\xi_1^2+\xi_2^2}\ .
\end{aligned}
$$

Somit ist, nach stetiger Fortsetzung auch zu Punkten ($\xi_2 = 0$, $\xi_1 > 0$), die Funktion $F(\xi_1,\xi_2) = \arccos\frac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}}$ auf der geschlitzten Ebene eine Stammfunktion zur Windungsform. ◁

Eine weiteres wichtiges Anwendungfeld dieser Methoden ist die Klasse der *exakten Differentialgleichungen*. Dabei geht es ausschließlich um Differentialgleichungen 1. Ordnung für eine Funktion $y(x)$.

Definition 16.12 (Exakte Differentialgleichung) Sei $U \subset \mathbb{R}^2$ offen und zusammenhängend und $g, h : U \to \mathbb{R}$ stetig. Die Differentialgleichung

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

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

Die Stammfunktion einer exakten Differentialgleichung ist bis auf eine Konstante eindeutig. Die Lösungen $y(x)$ einer exakten Differentialgleichung ergeben sich wie folgt: Ist $F(x,y)$ eine Stammfunktion, dann löse man die Gleichung $F(x,y(x)) = C$, für $C = \text{const}$, lokal mittels des Satzes über implizite Funktionen auf. Diese Lösung ist dann eine durch C parametrisierte Kurvenschar. Denn ist $h(x,y) = (\partial_y F)(x_0,y_0) \neq 0$, dann gilt in einer Umgebung von (x_0,y_0)

$$y'(x) = -\frac{(\partial_x F)(x,y)}{(\partial_y F)(x,y)} = -\frac{g(x,y)}{h(x,y)}\ .$$

Somit wird die exakte Differentialgleichung zurückerhalten.

Beispiel 16.13 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. ◁

Über die Verbindung $g(x,y) + h(x,y)y'(x) = 0 \quad\Leftrightarrow\quad \omega = g(x,y)dx + h(x,y)dy$ und die Identifizierung der Stammfunktionen erhalten wir:

Satz 16.14 (Notwendige Bedingung für Exaktheit) *Auf einer offenen und zusammenhängenden Teilmenge $U \subset \mathbb{R}^2$ sei die Differentialgleichung $g(x,y) +$*

$h(x,y)\,y' = 0$ gegeben mit stetig differenzierbaren Abbildungen $g, h : U \to \mathbb{R}$. Ist die Differentialgleichung exakt, so gilt

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

Für Sterngebiete läßt sich die Stammfunktion über das Kurvenintegral berechnen:

Satz 16.15 *Es sei $U \subset \mathbb{R}^2$ ein offenes Sterngebiet. Die Abbildungen $g, h : U \to \mathbb{R}$ seien stetig differenzierbar, und es gelte $\frac{\partial g}{\partial y}(x,y) = \frac{\partial h}{\partial x}(x,y)$ für alle $(x,y) \in U$. Sei $(x_0, y_0) \in U$ ein beliebiger Anfangspunkt und $c(t) = (x(t), y(t))$ für $t \in [0,1]$ eine beliebige stückweise stetig differenzierbare Kurve in U mit $(x(0), y(0)) = (x_0, y_0)$ und $(x(1), y(1)) = (x, y)$. Dann ist*

$$\begin{aligned} F(x,y) &:= \int_{(x_0,y_0)}^{(x,y)} \big(g(x,y)dx + h(x,y)dy\big) \\ &= \int_0^1 dt \, \big(g(x(t),y(t))x'(t) + h(x(t),y(t))y'(t)\big) \end{aligned}$$

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

Beispiel 16.16 Gesucht wird für geeignete Intervalle $I, J \subset \mathbb{R}$ eine Lösung $(x,y) \in I \times J$ der Differentialgleichung

$$\underbrace{\Big(1 - \frac{2x}{y(x)}\Big)}_{g(x,y)} + \underbrace{\Big(\frac{2}{(y(x))^3} + \frac{x^2-1}{(y(x))^2}\Big)}_{h(x,y)} y'(x) = 0$$

Wegen $\partial_y\big(1 - \frac{2x}{y}\big) = \frac{2x}{y^2} = \partial_x\big(\frac{2}{y^3} + \frac{x^2-1}{y^2}\big)$ ist die zugehörige 1-Form $\omega = gdx + hdy$ geschlossen. Sei $J \subset \mathbb{R}^\times_+$. Da $I \times J$ ein Sterngebiet ist, ist eine Stammfunktion gegeben z.B. durch Integration von ω längs eines horizontalen Weges $c(t) = (x_0 + t(x - x_0), y_0)$ und dann längs eines vertikalen Weges $c(t) = (x, y_0 + t(y - y_0))$, für $(x_0, y_0) \in I \times J$ und jeweils $t \in [0,1]$:

$$\begin{aligned} F(x,y) &= \int_0^1 dt \; g(x_0 + t(x-x_0), y_0) \cdot (x - x_0) + \int_0^1 dt \; h(x, y_0 + t(y - y_0)) \cdot (y - y_0) \\ &= \int_0^1 dt \, \Big(1 - \frac{2(x_0 + t(x-x_0))}{y_0}\Big)(x - x_0) \\ &\quad + \int_0^1 dt \, \Big(\frac{2}{(y_0 + t(y-y_0))^3} + \frac{x^2-1}{(y_0+t(y-y_0))^2}\Big) \cdot (y - y_0) \\ &= \Big(1 - \frac{x + x_0}{y_0}\Big)(x - x_0) - \frac{1}{y^2} + \frac{1}{y_0^2} - \frac{x^2-1}{y} + \frac{x^2-1}{y_0} \\ &= x - \frac{x^2-1}{y} - \frac{1}{y^2} - C_0 \ , \qquad C_0 = x_0 - \frac{x_0^2 - 1}{y_0} - \frac{1}{y_0^2} \ . \end{aligned}$$

Damit ist die Lösung der Differentialgleichung gegeben durch die implizit definierten Kurven $F(x,y) = C - C_0$, also $(x-C)y^2 - (x^2-1)y - 1 = 0$. Die Lösung ist

$$y(x) = \begin{cases} \frac{x^2-1\pm\sqrt{(x^2-1)^2+4(x-C)}}{2(x-C)} & \text{für } x > C\ , \\ -\frac{1}{C^2-1} & \text{für } x = C \neq 1\ , \\ \frac{1-x^2\pm\sqrt{(1-x^2)^2-4(C-x)}}{2(C-x)} & \text{für } x < C\ ,\ (1-x^2)^2 > 4(C-x)\ . \end{cases}$$

Je nach Anfangsbedingung ist nur eines der Vorzeichen realisiert. Nehmen wir z.B. $C = 0$, dann ist in der Kurve durch $(x_0, y_0) = (1,1)$ für $x > 0$ das positive Vorzeichen realisiert. Diese Lösung setzt sich stetig fort zu $-0.22527 \leq x \leq 0$, wobei das negative Vorzeichen realisiert ist. Der andere Zweig würde durch $(x_0, y_0) = (1,-1)$ gehen und sich zu $x > 0$ fortsetzen, ist aber durch die Bedingung $y > 0$ ausgeschlossen. ◁

Manchmal ist eine Differentialgleichung $g(x,y)\,dx + h(x,y)\,dy = 0$ nicht exakt, aber durch Multiplikation mit einem *integrierenden Faktor* $m(x,y)$ exakt zu machen, d.h. $m(x,y)\,g(x,y)\,dx + m(x,y)\,h(x,y)\,dy = 0$ ist exakt. Ein integrierender Faktor kann relativ leicht gefunden werden, wenn er nur von x oder nur von y abhängt. Offenbar ist die zugehörige 1-Form $m(x)g(x,y)dx + m(x)h(x,y)dy$ genau dann geschlossen, wenn gilt

$$\frac{(\partial_y g)(x,y) - (\partial_x h)(x,y)}{h(x,y)} = \frac{m'(x)}{m(x)} = (\ln m)'(x)\ .$$

Das erfordert, daß die linke Seite unabhängig von y ist. Analog ist $m(y)g(x,y)dx + m(y)h(x,y)dy$ genau dann geschlossen, wenn

$$\frac{(\partial_x h)(x,y) - (\partial_y g)(x,y)}{g(x,y)} = \frac{m'(y)}{m(y)} = (\ln m)'(y)\ .$$

Beispiel 16.17 Gegeben sei die Differentialgleichung

$$\underbrace{(1 - x^2\,y(x))}_{g} + \underbrace{(x^2\,y(x) - x^3)}_{h}\,y'(x) = 0\ .$$

Sie ist nicht exakt, aber

$$\frac{(\partial_y g)(x,y) - (\partial_x h)(x,y)}{h(x,y)} = \frac{-x^2 - (2xy - 3x^2)}{x^2y - x^3} = -\frac{2}{x}$$

ist unabhängig von y. Somit führt $(\ln m)'(x) = -2(\ln x)'$ auf den integrierenden Faktor $m(x) = \frac{C}{x^2}$, d.h.

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

ist eine exakte Differentialgleichung. Eine Stammfunktion auf einem Sterngebiet ist gegeben durch

$$\begin{aligned} F(x,y) &= \int_0^1 dt\,\Big(\frac{1}{(x_0+(x-x_0)t)^2}-y_0\Big)(x-x_0) \\ &\qquad + \int_0^1 dt\,((y_0+(y-y_0)t)-x)(y-y_0) \\ &= -\frac{1}{x}+\frac{1}{x_0}-y_0(x-x_0)+(y_0-x)(y-y_0)+\frac{1}{2}(y-y_0)^2 \\ &= -\frac{1}{x}-xy+\frac{1}{2}y^2-C_0\;, \qquad C_0=-\frac{1}{x_0}-x_0y_0+\frac{1}{2}y_0^2\;. \end{aligned}$$

Die Lösungen der ursprünglichen Differentialgleichung sind somit

$$y(x)=x\pm\sqrt{x^2+\frac{2}{x}+C}\;. \qquad \triangleleft$$

17 Holomorphe Funktionen

Wir behandeln nun die komplexe Differenzierbarkeit von komplexwertigen Funktionen $f:U\to\mathbb{C}$ auf offenen Teilmengen $U\subset\mathbb{C}$.

Definition 17.1 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 22.4 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

$$\begin{aligned} &(f+g)'(z)=f'(z)+g'(z)\;, \qquad (f\cdot g)'(z)=f'(z)g(z)+f(z)g'(z)\;, \\ &(f\circ g)'(z)=f'(g(z))\cdot g'(z)\;. \end{aligned}$$

Analog zu Satz 22.12 aus dem 1. Semster gilt:

Satz 17.2 *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 17.3 *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

$$\begin{aligned}
\partial_x f(x+\mathrm{i}y) &= \lim_{h\to 0\,,\; h\in\mathbb{R}} \frac{1}{h}\Big(f(x+h+\mathrm{i}y) - f(x+\mathrm{i}y)\Big) = f'(z)\;,\\
\partial_y f(x+\mathrm{i}y) &= \lim_{h\to 0\,,\; h\in\mathbb{R}} \frac{\mathrm{i}}{\mathrm{i}h}\Big(f(x+\mathrm{i}(y+h)) - f(x+\mathrm{i}y)\Big) = \mathrm{i}f'(z)\;,
\end{aligned}$$

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 2.6 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. $\square$

Beispiel 17.4 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. $\triangleleft$

Holomorphe Funktionen (also komplexwertige Funktionen, die in einer Umgebung komplex differenzierbar sind) haben die bemerkenswerte Eigenschaft, daß sie sogar beliebig oft differenzierbar sind. Wir werden das in mehreren Teilschritten beweisen. Der erste Schritt ist die Betrachtung von Kurvenintegralen.

Sei $U \subset \mathbb{C}$ offen. Jede stetige komplexwertige Funktion $f : U \to \mathbb{C}$ definiert eine 1-Form $\omega = f dz = f dx + \mathrm{i} f dy : U \to \mathrm{Hom}(\mathbb{R}^2, \mathbb{C})$ durch die Identifikation $\mathbb{C} = \mathbb{R}^2$. Ist $c(t) = x(t) + iy(t) \equiv \big(x(t), y(t)\big)$ (stückweise) stetig differenzierbar, so folgt für das sich ergebende Kurvenintegral der 1-Form $f dz$ längs c

$$\int_c f(z)dz = \int_\alpha^\beta dt\; f(c(t))x'(t) + if(c(t))y'(t) \equiv \int_\alpha^\beta dt\; f(c(t))c'(t) \;.$$

Die Standardabschätzungen liefern:

Satz 17.5 *Sei $c : [\alpha, \beta] \to U$ eine differenzierbare Kurve und $f : U \to \mathbb{C}$ stetig. Dann gilt*

$$\Big|\int_c dz\; f(z)\Big| \leq \sup_{t\in[\alpha,\beta]} |f(c(t))|\, L(c) \;,$$

wobei $L(c) := \int_a^b dt\; \sqrt{(x'(t))^2 + (y'(t))^2}$ die Bogenlänge von C ist. $\square$

Ist f *stetig* komplex differenzierbar auf U, so liefern die Cauchy-Riemannschen Differentialgleichungen $\partial_y f = \mathrm{i}\partial_x f$ und damit die *Geschlossenheit* von $f dx + \mathrm{i} f dy$. Nach dem Lemma von Poincaré ist dann $f dz$ exakt auf Sterngebieten und läßt sich dort wegunabhängig integrieren. Wir zeigen nun, daß das sogar *ohne* die Voraussetzung der *stetigen* Differenzierbarkeit gilt. Zunächst sei $c = \partial\Delta$ der Rand eines Dreiecks Δ.

Satz 17.6 (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

$$\Big|\int_{\partial\Delta} dz\ f(z)\Big| \leq 4\Big|\int_{\partial\Delta_1} dz\ f(z)\Big|\ .$$

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

$$\Big|\int_{\partial\Delta} dz\ f(z)\Big| \leq 4^n\Big|\int_{\partial\Delta_n} dz\ f(z)\Big|\ .$$

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

$$f(z) = f(z_0) + (z - z_0)f'(z_0) + |z - z_0|r(z) \qquad \text{mit } \lim_{z\to z_0} r(z) = 0\ .$$

Da jede polynomiale 1-Form eine Stammfunktion besitzt, gilt

$$\int_{\partial\Delta_n} dz\ \big(f(z_0) + (z - z_0)f'(z_0)\big) = 0\ .$$

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 17.5

$$\Big|\int_{\partial\Delta_n} dz\ f(z)\Big| = \Big|\int_{\partial\Delta_n} dz\ |z-z_0|r(z)\Big| \leq 2^{-n}L\cdot \sup_{z\in\partial\Delta_n} |z-z_0||r(z)| \leq 4^{-n}L^2 \sup_{z\in\overline{\Delta_n}} |r(z)|\ .$$

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. □

Lemma 17.7 *Es sei $U \subset \mathbb{C}$ offen, $c : [\alpha, \beta] \to U$ eine (stückweise) stetig differenzierbare Kurve und ω eine stetige 1-Form auf U. Dann gibt es zu jedem $\epsilon > 0$ eine Unterteilung $\alpha = t_0 < t_1 < \cdots < t_m = \beta$, so daß für die zugehörigen Sekanten $\tilde{c}_j(t) := c(t_{j-1}) + \frac{t-t_{j-1}}{t_j-t_{j-1}}\big(c(t_j) - c(t_{j-1})\big)$, mit $t \in [t_{j-1}, t_j]$, von c gilt*

$$\Big|\int_c \omega - \sum_{j=1}^m \int_{\tilde{c}_j} \omega\Big| < \epsilon\,.$$

Beweis. Durch leichte Abwandlung des Beweises von Satz 15.6. Wegen der Stetigkeit von ω_i und der Offenheit von U gibt es zu jedem $t \in [\alpha, \beta]$ eine offene Kugel $K_{\delta_t}(c(t)) \subset U$, so daß $|\omega_i(c(t)) - \omega_i(a)| < \frac{\epsilon}{4nL(c)}$ für alle $a \in K_{\delta_t}(c(t))$. Dann ist $|\omega_i(a) - \omega_i(b)| < \frac{\epsilon}{2nL(c)}$ für alle $a, b \in K_{\delta_t}(c(t))$. Die Menge $\bigcup_{t\in[\alpha,\beta]} K_{\delta_t}(c(t))$ ist eine Überdeckung der Spur von c. Da die Spur von c kompakt ist als Bild von $[\alpha, \beta]$ unter einer stetigen Abbildung, gibt es endlich viele offene Kugeln $K_1, \dots, K_m \subset U$ mit Mittelpunkten $c(\tau_1), \dots, c(\tau_m)$, welche die Spur von c überdecken, und für die gilt $|\omega_i(a_j) - \omega_i(b_j)| < \frac{\epsilon}{2nL(c)}$ für alle $a_j, b_j \in K_j$. Sei $t_0 = \alpha$ und $t_m = \beta$. Nach aufsteigender Anordnung entsprechend des Kurvenparameters sei $c(t_0) \in K_1$ und $c(t_m) \in K_m$. Weiterhin gibt es zu jedem $1 \leq j < m$ einen Punkt $c(t_j) \in K_j \cap K_{j+1}$ (mit $t_j > t_k$ für $j > k$). Nach Konstruktion ist $c([t_{j-1}, t_j]) \subset K_j$ und $\tilde{c}_j([t_{j-1}, t_j]) \subset K_j$ sowie $\tau_j \in [t_{j-1}, t_j]$. Wegen $\int_{t_{j-1}}^{t_j} dt\, c_i'(t) = \int_{t_{j-1}}^{t_j} dt\, \tilde{c}_{ji}'(t) = c_i(t_j) - c_i(t_{j-1})$ gilt

$$\begin{aligned}
&\Big|\int_c \omega - \sum_{j=1}^m \int_{\tilde{c}_j} \omega\Big| \\
&= \Big|\sum_{i=1}^n \sum_{j=1}^m \int_{t_{j-1}}^{t_j} dt\, \big(\omega_i(c(t))c_i'(t) - \omega_i(c(\tau_j))c_i'(t) + \omega_i(c(\tau_j))\tilde{c}_{ji}'(t) - \omega_i(\tilde{c}_j(t))\tilde{c}_{ji}'(t)\big)\Big| \\
&\leq \sum_{i=1}^n \sum_{j=1}^m \int_{t_{j-1}}^{t_j} dt\, \Big(\big|\omega_i(c(t)) - \omega_i(c(\tau_j))\big|\big|c_i'(t)\big| + \big|\omega_i(\tilde{c}_j(t)) - \omega_i(c(\tau_j))\big|\big|\tilde{c}_{ji}'(t)\big|\Big) < \epsilon
\end{aligned}$$

nach ähnlichen Schritten wie im Beweis von Satz 15.6 sowie der Supremums-Eigenschaft der Bogenlänge $\displaystyle\sum_{j=1}^m \int_{t_{j-1}}^{t_j} dt\, \|\tilde{c}_j\| \leq L(c)$. □

Definition 17.8 Es sei $U \subset \mathbb{C}$ offen.

i) Zwei stetige Kurven $c_0, c_1 : [\alpha, \beta] \to U$ mit $c_0(\alpha) = c_1(\alpha) = a$ und $c_0(\beta) = c_1(\beta) = b$ heißen *homotop* in U, wenn sie in U stetig ineinander deformiert werden können, d.h. wenn es eine stetige Abbildung $H : [\alpha, \beta] \times [0, 1] \mapsto H(t, s) \in U$ gibt mit $H(t, 0) = c_0(t)$ und $H(t, 1) = c_1(t)$ für alle $t \in [\alpha, \beta]$ sowie $H(\alpha, s) = a$ und $H(\beta, s) = b$ für alle $s \in [0, 1]$.

ii) Eine geschlossene stetige Kurve $c : [\alpha, \beta] \to U$ mit $c(\alpha) = c(\beta) = a$ heißt *kontrahierbar* in U (oder *nullhomotop*), wenn sie homotop zum Punkt $c_0(t) = a$ für alle $t \in [\alpha, \beta]$ ist.

iii) U heißt *einfach zusammenhängend*, wenn U zusammenhängend ist und jede geschlossene Kurve in U kontrahierbar ist.

Satz 17.9 (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 stetig differenzierbare geschlossene Kurve $c : [\alpha, \beta] \to U$ gilt $\displaystyle\int_c dz\ f(z) = 0$.*

ii) *Sind c_0, c_1 homotope stückweise stetig differenzierbare Kurven in U mit gemeinsamem Anfangs- und Endpunkt, dann gilt*

$$\int_{c_0} dz\ f(z) = \int_{c_1} dz\ f(z)\ .$$

iii) *Ist U einfach zusammenhängend und $c : [\alpha, \beta] \to U$ eine beliebige stückweise stetig differenzierbare Kurve von $c(\alpha) = z_0$ nach $c(\beta) = z$, dann ist das Integral*

$$F(z) := \int_{z_0}^{z} dw\ f(w) = \int_c dw\ f(w)$$

unabhängig von c und eine holomorphe Stammfunktion zu f, d.h. es gilt $F'(z) = f(z)$.

Beweis. i) Sollte sich die Kurve c selbst schneiden, so überlegt man sich, daß das Kurvenintegral zerfällt in eine (nicht eindeutige) Summe von Integralen längs Kurven ohne Selbstschnittpunkte. Es genügt dann, i) nur für Kurven ohne Selbstschnittpunkte zu beweisen. Nach Lemma 17.7 gibt es zu jedem $\epsilon > 0$ einen geschlossenen Polygonzug $P(c)$ in U, der wegen der Offenheit von U auch ohne Selbstschnittpunkte gewählt werden kann, so daß $\Big|\displaystyle\int_c \omega - \int_{P(c)} \omega\Big| < \epsilon$. Nach Konstruktion ist mit c auch $P(c)$ in U kontrahierbar. Es sei $Q(c)$ das Innere von $P(c)$, welches sich triangulieren läßt, d.h. es gibt endlich viele abgeschlossene Dreiecke $\Delta_1, \dots, \Delta_k \subset U$, welche sich höchstens in Ecken oder Kanten schneiden, mit $\Delta_1 \cup \dots \cup \Delta_k = Q(c)$ und $\partial(\Delta_1 \cup \dots \cup \Delta_k) = P(c)$. Nach dem Lemma von Goursat ist $\displaystyle\int_{\partial\Delta_i} dz\ f(z) = 0$ für jedes Dreieck Δ_i. In der Summe heben sich die Beiträge aller inneren Kanten weg, so daß

$$0 = \sum_{i=1}^{k} \int_{\partial\Delta_i} dz\ f(z) = \int_{P(c)} dz\ f(z)\ .$$

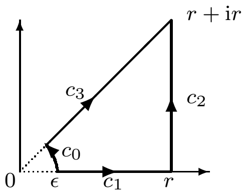
Somit ist $\Big|\displaystyle\int_c dz\ f(z)\Big| < \epsilon$ für beliebiges $\epsilon > 0$, also $\displaystyle\int_\gamma dz\ f(z) = 0$.

ii) folgt aus i) für die kontrahierbare Kurve $c = c_0 \cup (-c_1)$, wobei $-c_1$ in umgekehrte Richtung wie c_1 durchlaufen wird, was $\displaystyle\int_{-c_1} dz\ f(z) = -\int_{c_1} dz\ f(z)$ ergibt.

iii) folgt aus ii) zusammen mit Satz 16.5. □

Der Cauchysche Integralsatz ist die Grundlage der gesamten Funktionentheorie. Seine Konsequenzen sind grundlegend verschieden zur reellen Dfferentialrechnung. Außerdem erlaubt er bereits die Berechnung komplizierterer reeller Integrale durch Schließen des Integrationsweges im Komplexen.

Beispiel 17.10 Gesucht ist (für noch festzulegende $s \in \mathbb{R}$) das Riemann-Integral $\int_0^\infty dt\ t^{2s-1}e^{-it^2}$. Dazu integrieren wir die holomorphe Funktion $f(z) = z^{2s-1}e^{-z^2}$ über die geschlossene Kurve $(c_1, c_2, -c_3, -c_0)$:



$$\int_{c_0} dzf(z) + \int_{c_3} dzf(z) = \int_{c_1} dzf(z) + \int_{c_2} dzf(z)\ .$$

Mit $c_2(t) = r + it$ ist $|f(c_2(t))| = \sqrt{r^2+t^2}^{2s-1} e^{-r^2+t^2} \leq \max(r^{s-1}, (\sqrt{2}r)^{2s-1})e^{-r^2+rt}$ und damit

$$\Big|\int_{c_2} dz\ f(z)\Big| \leq \int_0^r dt\ |f(c_2(t))c_2'(t)| = c_s r^{2s-1}e^{-r^2}\int_0^r dt\ e^{rt} = c_s\frac{1-e^{-r^2}}{r^{2-2s}}\ .$$

Für $s < 1$ konvergiert das Integral im Limes $r \to \infty$ gegen 0.

Mit $c_0(t) = \epsilon e^{\mathrm{i}t}$, $0 \leq t \leq \frac{\pi}{4}$, ist $|f(c_0(t))| = \epsilon^{2s-1}e^{-\epsilon^2(2s-1)\sin(2t)} \leq \epsilon^{2s-1}\max(1, e^{(1-2s)\epsilon^2})$ und damit

$$\Big|\int_{c_0} dz\ f(z)\Big| \leq \int_0^{\frac{\pi}{4}} dt\ |f(c_0(t))c_0'(t)| \leq \frac{\pi}{4}\epsilon^s\tilde{c}_s\ .$$

Für $s > 0$ konvergiert das Integral im Limes $\epsilon \to 0$, also gilt für $0 < s < 1$

$$\begin{aligned}
\lim_{r\to\infty,\epsilon\to 0}\int_{c_3} f(z)dz &= \lim_{r\to\infty,\epsilon\to 0}\int_{c_1} f(z)dz + \lim_{r\to\infty}\int_{c_2} f(z)dz - \lim_{\epsilon\to 0}\int_{c_0} f(z)dz \\
&= \lim_{r\to\infty,\epsilon\to 0}\int_\epsilon^r dt\ t^{2s-1}e^{-t^2} = \frac{1}{2}\lim_{r\to\infty,\epsilon\to 0}\int_\epsilon^r (2tdt)\ (t^2)^{s-1}e^{-t^2} \\
&= \frac{1}{2}\Gamma(s)\ .
\end{aligned}$$

Andererseits ist $c_3(t) = e^{\mathrm{i}\frac{\pi}{4}}t$ mit $t \in [\epsilon, \sqrt{2}r]$, also $(c_3(t))^2 = \mathrm{i}t^2$ und

$$\lim_{r\to\infty,\epsilon\to 0} \int_\epsilon^{\sqrt{2}r} dt\; e^{\mathrm{i}\frac{\pi}{2}s} t^{2s-1} e^{-\mathrm{i}t^2} = e^{\frac{\mathrm{i}\pi s}{2}} \int_0^\infty dt\; t^{2s-1} e^{-\mathrm{i}t^2} \;.$$

Somit gilt $\displaystyle\int_0^\infty dt\; t^{2s-1} e^{-\mathrm{i}t^2} = \frac{1}{2}\Gamma(s) e^{-\frac{\mathrm{i}\pi s}{2}}$. Für $s = \frac{1}{2}$ ist das das Fresnel-Integral $\displaystyle\int_0^\infty dt\; e^{-\mathrm{i}t^2} = \frac{1}{2}\sqrt{\frac{\pi}{\mathrm{i}}} = \sqrt{\frac{\pi}{8}}(1-\mathrm{i})$. Allgemein erhalten wir nach Zerlegung in Real- und Imaginärteil mit $e^{-\mathrm{i}t^2} = \cos t^2 - \mathrm{i}\sin t^2$ schließlich für $0 < s < 1$

$$\int_0^\infty dt\; t^{2s-1} \cos(t^2) = \frac{1}{2}\Gamma(s) \cos\frac{\pi s}{2} \;, \qquad \int_0^\infty dt\; t^{2s-1} \sin(t^2) = \frac{1}{2}\Gamma(s) \sin\frac{\pi s}{2} \;,$$

$$\int_0^\infty dx\; x^{s-1} \cos x = \Gamma(s) \cos\frac{\pi s}{2} \;, \qquad \int_0^\infty dx\; x^{s-1} \sin x = \Gamma(s) \sin\frac{\pi s}{2} \;.$$

Im letzten Integral existiert der Limes $s \to 0$ mit $\displaystyle\int_0^\infty dx\, \frac{\sin x}{x} = \frac{\pi}{2}$. Der Limes $s \to 1$ ist nicht möglich. ◁

Beispiel 17.11 (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 $c : [\alpha, \beta] \to \mathbb{C}^-$ 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

$$L(z) = \int_1^z \frac{dz}{z} := \int_c \frac{dz}{z} \qquad (c \text{ ist beliebige Kurve zwischen } 1 \text{ und } z \text{ in } \mathbb{C}^-)$$

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 $z\mathrm{e}^{-L(z)} = \text{const} = 1 \cdot \mathrm{e}^{-L(1)} = 1$, 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

$$c = c_1 \cup c_2 \;, \quad c_1(t) = t : [1, r] \to \mathbb{C} \;, \quad c_2(t) = r\mathrm{e}^{\mathrm{i}t} : [0, \varphi] \to \mathbb{C}$$

von 1 über r nach $z = r\mathrm{e}^{\mathrm{i}\varphi}$ ergibt sich mit $c_1'(t) = 1$ und $c_2'(t) = \mathrm{i}r\mathrm{e}^{\mathrm{i}t}$

$$L(z) = \int_1^r dt \frac{1}{t} \cdot 1 + \int_0^\varphi dt \frac{1}{r\mathrm{e}^{\mathrm{i}\varphi}} \cdot \mathrm{i}r\mathrm{e}^{\mathrm{i}t} = r + \mathrm{i}\phi \;.$$ ◁

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 17.11 ist

$$F(z) = \int_{z_0}^z dw\, \frac{f'(w)}{f(w)}$$

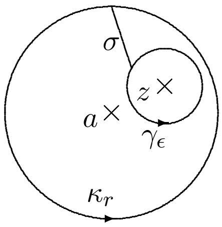
eine holomorphe Stammfunktion zu $\frac{f'}{f}$, und es gilt $fe^{-F} = \text{const.}$ Wegen $e^z = 1 \quad\Leftrightarrow\quad z \in 2\pi\mathrm{i}\mathbb{Z}$ unterscheiden sich zwei holomorphe Logarithmen nur um Vielfache von $2\pi\mathrm{i}$ und sind somit durch Angabe ihres Wertes in einem Punkt $z_0 \in U$ eindeutig bestimmt. Über den holomorphen Logarithmus können komplexe Potenzen nullstellenfreier Funktionen definiert werden als $f^\alpha = e^{\alpha F}$ für $\alpha \in \mathbb{C}$. $\triangleleft$

18 Die Cauchysche Integralformel

Satz 18.1 (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 17.9

$$\int_{\kappa_r} d\zeta\, \frac{f(\zeta)}{\zeta - z} = \int_{\gamma_\epsilon} d\zeta\, \frac{f(\zeta)}{\zeta - z} = \int_{\gamma_\epsilon} d\zeta\, \frac{f(\zeta) - f(z)}{\zeta - z} + f(z) \int_{\gamma_\epsilon} d\zeta\, \frac{1}{\zeta - z} \;.$$

Insbesondere ist die linke 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 $\lim\limits_{\epsilon \to 0} \int_{\gamma_\epsilon} d\zeta\, \frac{f(\zeta) - f(z)}{\zeta - z} = 0$. Schließlich gilt auf dem inneren Kreis $\frac{1}{\zeta(\gamma_\epsilon(\tau)) - z} = \frac{1}{\epsilon e^{\mathrm{i}\tau}}$ und $\gamma'_\epsilon(\tau) = \mathrm{i}\epsilon\, e^{\mathrm{i}t}$, 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. $\square$

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 18.2 (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 $|a_n| \leq \frac{M}{r^n}$.

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 $|\frac{z-a}{\zeta - a}| \leq 1 - q$. 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 vertauschen Summe und Integral, und es 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 $|\frac{f(\zeta)}{(\zeta - a)^{n+1}}| \leq \frac{M}{r^{n+1}}$ für alle $\zeta \in \partial K_r(a)$ und der Länge $2\pi r$ des Randes. □

Als wichtige Konsequenz ergibt sich:

Satz 18.3 *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 $x = 0$).

Beispiel 18.4 Für $c(t) = 1 + e^{2\pi \mathrm{i} t}$, $t \in [0,1]$ ist

$$\int_c \frac{d\zeta}{(\zeta+1)(\zeta-1)^2} = \int_{\partial K_1(1)} \frac{d\zeta}{2}\Big(\frac{1}{\zeta-1} - \frac{1}{\zeta+1}\Big)\frac{1}{(\zeta-1)}$$
$$= \int_{\partial K_1(1)} \frac{d\zeta}{4}\Big(\frac{2}{(\zeta-1)^2} - \frac{1}{\zeta-1} + \frac{1}{\zeta+1}\Big) .$$

Mit $f_1(\zeta) = 1$ und $z = 1 \in K_1(1)$ gilt

$$\int_{\partial K_1(1)} d\zeta \frac{1}{(\zeta-1)^2} = \int_{\partial K_1(1)} d\zeta \frac{f_1(\zeta)}{(\zeta-z)^2}\Big|_{z=1\in K_1(1)} = \frac{2\pi \mathrm{i}}{1!} f_1'(z) = 0 .$$

Im dritten Integral wird wegen $z = -1 \notin K_1(1)$ über eine auf $K_1(1)$ holomorphe Funktion integriert, deshalb ist nach Cauchyschem Integralsatz

$$\int_{\partial K_1(1)} d\zeta \frac{1}{\zeta+1} = \int_{\partial K_1(1)} d\zeta \frac{1}{\zeta-z}\Big|_{z=-1\notin K_1(1)} = 0 .$$

Schließlich liefert die Cauchysche Integralformel

$$\int_c \frac{d\zeta}{(\zeta+1)(\zeta-1)^2} = -\frac{1}{4}\int_{\partial K_1(1)} \frac{1}{\zeta-z}\Big|_{z=1\in K_1(1)} = -\frac{1}{4}\cdot 2\pi \mathrm{i} = -\frac{\mathrm{i}\pi}{2} . \qquad \triangleleft$$

Satz 18.5 *Ist $f(z) = \sum_{k=0}^\infty a_k(z-z_0)^k$ eine in $K_R(z_0)$ konvergente Potenzreihe, dann kann f um jeden Punkt $w \in K_R(z_0)$ in eine Potenzreihe $f(z) = \sum_{n=0}^\infty b_n(z-w)^n$ entwickelt werden. Der Konvergenzradius dieser Reihe ist mindestens $R - |w-z_0|$, und es gilt $b_n = \displaystyle\sum_{k=n}^\infty \binom{k}{n} a_k(w-z_0)^{k-n}$.*

Beweis. f ist holomorph in $w \in K_R(z_0)$ und deshalb in eine Potenzreihe entwickelbar. Der Konvergenzradius folgt aus Satz 18.2, sowie

$$b_n = \frac{1}{n!}f^{(n)}(w) = \frac{1}{n!}\sum_{k=n}^\infty a_k k(k-1)\cdots(k-n+1)(w-z_0)^{k-n} . \qquad \square$$

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 18.6 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. $\triangleleft$

Wir zeigen, daß eine mögliche analytische Fortsetzung eindeutig ist:

Satz 18.7 *Es sei $U \subset \mathbb{C}$ offen und zusammenhängend. 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 $z_0 \in U$, so daß $f^{(k)}(z_0) = g^{(k)}(z_0)$ für alle $k \in \mathbb{N}$.*

Beweis. i)⇒ii) ist klar.

ii)⇒iii) Für $h = f - g$ hat die Nullstellenmenge von h einen Häufungspunkt $z_0 \in U$. Angenommen, es gäbe ein $k \in \mathbb{N}$ mit $h^{(k)}(z_0) \neq 0$, und sei n das Minimum dieser k. Wegen der Potenzreihenentwicklung ist $h(z) = (z - z_0)^n h_n(z)$ mit $h_n(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$ und $h_n(z_0) \neq 0$. 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 $S_k := \{w \in U \ : \ h^{(k)}(w) = 0\}$. Als Urbild einer abgeschlossenen Menge unter einer stetigen Abbildung ist S_k abgeschlossen in U (Satz 19.2.iii) aus dem 1. Semester). 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 18.8 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 18.2 gibt es für eine ganze Funktion f die Darstellung $f(z) = \sum_{n=0}^{\infty} a_n z^n$ mit Konvergenzradius ∞.

Satz 18.9 (Liouville) *Jede beschränkte ganze Funktion ist konstant.*

Beweis. Ist $|f| \leq M$ auf $\mathbb{C}$, dann erfüllen die Entwicklungskoeffizienten nach Satz 18.2 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 18.10 (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 $P(z) = \prod_{i=1}^{n}(z - b_i)$.

19 Der Residuensatz

Wichtig für die Ausnutzung der Cauchyschen Integralformel zur Berechnung von Integralen ist eine genauere Diskussion möglicher Singularitäten von komplexen Funktionen.

Satz 19.1 (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

$$\tilde{f}(z) := \sum_{n=0}^{\infty} a_{n+2}(z-a)^n \ .$$

□

Definition 19.2 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 $k \in \mathbb{N}^\times$ 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* $\sum_{n=0}^{\infty} a_n(z-a)^n$. Der *Hauptteil der Laurent-Reihe* $\sum_{n=-k}^{-1} a_n(z-a)^n = \sum_{n=1}^{k} \frac{a_{-n}}{(z-a)^n}$ ist als endliche Summe beschränkt in $K_{R,r}(a)$. Für die Entwicklungskoeffizienten gilt:

Satz 19.3 *Eine auf $U \setminus \{a\}$ holomorphe Funktion habe in a einen k-fachen Pol. Dann sind die Entwicklungskoeffizienten der Laurent-Reihe $f(z) = \sum_{n=-k}^{\infty} a_n(z-a)^n$ 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 $h(z) = (z-a)^k f(z) = \sum\limits_{n=0}^{\infty} a_n (z-a)^n$ 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 $f(z) = \sum\limits_{n=-\infty}^{\infty} a_n (z-a)^n$ 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 19.4 (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 19.5 Es sei f eine auf $U \setminus \{a\}$ holomorphe Funktion mit Laurent-Reihenentwicklung $f(z) = \sum\limits_{n=-\infty}^{\infty} a_n (z-a)^n$. 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 19.3 gilt:

Satz 19.6 *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
$$\mathrm{res}_a \frac{g}{h} = \frac{g(a)}{h'(a) + \phi(a)} = \frac{g(a)}{h'(a)} \;.$$
- 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
$$\mathrm{res}_a f = a_{-1} = \frac{1}{(k-1)!} \frac{d^{(k-1)}}{dz^{k-1}} \big((z-a)^k f(z)\big)\Big|_{z=a} \;.$$

Satz 19.7 (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 mit folgenden Eigenschaften:*

i) *A ist einfach zusammenhängend in U,*

ii) *der Rand $\gamma := \partial A$ liegt in U und ist stückweise stetig differenzierbar,*

iii) *$S \cap \gamma = \varnothing$, d.h. der Rand ∂A trifft keinen Punkt aus S.*

Dann gilt
$$\int_\gamma f(z)dz = 2\pi \mathrm{i} \sum_{a \in S\cap A} \mathrm{res}_a f \;.$$

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 $\mathrm{Res}_a f$. □

Der Residuensatz ist ein mächtiges Werkzeug zur Berechnung von Integralen.

Beispiel 19.8 Gesucht ist $I(a) := \displaystyle\int_0^{2\pi} \frac{dt}{a + \cos t}$ für $a > 1$. 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 $z'(t) = \mathrm{i}z(t)$, und wir erhalten
$$I(a) = \int_0^{2\pi} dt\, \frac{1}{a + \frac{1}{2}(z(t) + \frac{1}{z(t)})} \frac{z'(t)}{\mathrm{i}z(t)} = \frac{2}{\mathrm{i}} \int_{\partial K_1(0)} \frac{dz}{z^2 + 2az + 1} = \frac{2}{\mathrm{i}} \int_{\partial K_1(0)} dz\; f(z)$$

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 19.9 *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 19.10 *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. □

Beispiel 19.11 Gesucht ist $I_n := \displaystyle\int_{-\infty}^{\infty} \frac{dx}{1+x^{2n}}$ mit $n \in \mathbb{N}^\times$. Aufgefaßt als komplexe Funktion sind die (einfachen) Pole $z^{2n} = -1$ der oberen Halbebene bei $a_k = e^{\frac{i\pi(2k+1)}{2n}}$, $k = 0, 1, \dots, n-1$. Also gilt mit $\mathrm{res}_a(\frac{1}{h}) = \frac{1}{h'(a)}$

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^{2n}} = \sum_{k=0}^{n-1} \frac{2\pi\mathrm{i}}{2ne^{\frac{\mathrm{i}\pi(2k+1)(2n-1)}{2n}}} = -\frac{\pi\mathrm{i}}{n} e^{\frac{\mathrm{i}\pi}{2n}} \sum_{k=0}^{n-1} e^{\frac{\mathrm{i}\pi k}{n}} = \frac{\pi}{\mathrm{i}n} e^{\frac{\mathrm{i}\pi}{2n}} \frac{1 - e^{\frac{\mathrm{i}\pi n}{n}}}{1 - e^{\frac{\mathrm{i}\pi}{n}}} = \frac{\frac{\pi}{n}}{\sin\frac{\pi}{2n}} \,.$$

◁

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 19.12 *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 zusätzlichen Kurvenstücke sind $c_1(t) = r + \mathrm{i}t$ mit $t \in [0, r]$, $c_2(t) = t + \mathrm{i}r$ mit $t \in [-r, r]$ und $c_3(t) = -r + \mathrm{i}t$ 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 $|R(z)| \leq \frac{M}{1+|z|}$ und damit

$$\begin{aligned} \Big|\int_0^r dt\ R(\pm r{+}\mathrm{i}t)e^{\mathrm{i}\alpha(\pm r+\mathrm{i}t)}\Big| &\leq \frac{M}{1+r}\int_0^r dt\ e^{-\alpha t} \leq \frac{M}{\alpha(1+r)} , \\ \Big|\int_{-r}^{r} dt\ R(t{+}\mathrm{i}r)e^{\mathrm{i}\alpha(t+\mathrm{i}r)}\Big| &\leq \frac{M}{1+r}e^{-\alpha r}\cdot 2r . \end{aligned}$$

Folglich verschwinden die Integrale über c_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 $|R'(x)| \leq \frac{M}{(1+|x|)^2}$. □

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 19.13 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 19.12 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 $\displaystyle\int_{-\infty}^{\infty} \frac{x\sin x}{1+x^2} = \frac{\pi}{e}$. ◁

Eine andere wichtige Anwendung des Residuensatzes ist die Lokalisierung von Nullstellen und Polstellen.

Satz 19.14 (Nullstellen und Polstellen zählendes Integral) *Es sei f eine nichtkonstante meromorphe Funktion auf $U \subset \mathbb{C}$, ferner $S \subset U$ die Menge der Nullstellen und Polstellen von f und $A \subset U$ eine Teilmenge, deren Rand ∂A die Voraussetzungen des Residuensatzes erfüllt. Dann gilt für die Anzahl der Nullstellen N_A von f in A und die Anzahl der Polstellen P_A von f in A, jeweils mit Vielfachhheit gezählt, die Formel*

$$N_A - P_A = \frac{1}{2\pi \mathrm{i}} \int_{\partial A} dz \frac{f'(z)}{f(z)}$$

Beweis. Ist f meromorph, so ist $\frac{f'}{f}$ holomorph außerhalb der Null- und Polstellen von f. In der Umgebung V einer Null- oder Polstelle a gilt $f(z) = (z-a)^k g(z)$ für ein $k \in \mathbb{Z}$ und eine auf V holomorphe nullstellenfreie Funktion g. Dabei ist $k > 0$ für eine k-fache Nullstelle und $k < 0$ für einen Pol der Ordnung $|k|$. Somit ist $\frac{f'}{f}(z) = \frac{k}{z-a} + \frac{g'(z)}{g(z)}$ und dann $\mathrm{res}_a \frac{f'}{f} = k$. Der Residuensatz liefert die Behauptung. □

Satz 19.15 (Rouché) *Es seien f, g holomorphe Funktionen auf $U \subset \mathbb{C}$ und $A \subset U$ eine Teilmenge mit stückweise stetig differenzierbarem Rand $\partial A \subset U$. Es gelte $|g(z)| < |f(z)|$ für alle $z \in \partial A$. Dann haben f und $f+g$ die gleiche Anzahl von Nullstellen in A.*

Beweis. Es gibt eine Umgebung $V \subset U$ von ∂A mit $|\frac{g}{f}| < 1$. Dann ist $h := 1 + \frac{g}{f}$ auf V holomorph mit Bild $h(V) \subset K_1(1)$. Der komplexe Logarithmus L ist holomorph auf $K_1(1)$, so daß nach Kettenregel gilt $(L \circ h)'(z) = \frac{h'(z)}{h(z)}$. Somit verschwindet das Kurvenintegral $\displaystyle\int_c dz \frac{h'(z)}{h(z)} = 0$ für jede geschlossene Kurve c mit $h \circ c \in K_1(1)$, insbesondere für das Bild von ∂A unter $\partial A \ni z \mapsto h(z) = 1 + \frac{g(z)}{f(z)}$. Es gilt $\frac{h'}{h} = \frac{(f+g)'(z)}{(f+g)(z)} - \frac{f'(z)}{f(z)}$, so daß nach Satz 19.14 die holomorphen Funktionen f und $f+g$ die gleiche Zahl von Nullstellen in A haben. □

Beispiel 19.16 Wieviele Nullstellen von $h(z) = z^8 - 5z^3 + z - 2$ liegen in $K_1(0)$? Setze $f(z) = -5z^3$ und $g(z) = z^8 - z - 2$. Dann gilt für $z \in \partial K_1(0)$, d.h. $|z| = 1$ nach Dreiecksungleichung $|g(z)| \leq 4 < |f(z)| = 5$. Die Gleichung $f(z) = -5z^3 = 0$ hat eine dreifache Nullstelle in $0 \in K_1(0)$. Somit hat nach dem Satz von Rouché die Funktion h ebenfalls 3 Nullstellen in $K_1(0)$. ◁

Wiederholung

- Definition von 1-Formen, mit Differential als Beispiel

- Integral von 1-Formen längs Kurven
- Exakte 1-Formen: Wegunabhängigkeit des Kurvenintegrals
- Geschlossene 1-Formen als notwendige Bedingung für Exaktheit, hinreichend auf Sterngebieten
- Anwendung: exakte Differentialgleichungen; auf Sterngebieten Lösung durch Kurvenintegral
- Definitionen komplexe Differenzierbarkeit und Holomorphie, Cauchy-Riemannsche Differentialgleichungen
- Definitionen kontrahierbare Kurve und einfach zusammenhängende Gebiete
- komplexe Kurvenintegrale, Cauchyscher Integralsatz
- Cauchysche Integralformel und Folgerungen für Potenzreihenentwicklung und holomorphe Fortsetzbarkeit. Satz von Liouville
- Diskussion isolierter Singularitäten, Laurent-Entwicklung
- Residuum mit Berechnungsmöglichkeiten, Residuensatz
- Berechnung reeller Integrale über Residuensatz oder Cauchyschem Integralsatz

Teil IV
Das Lebesgue-Integral

20 Treppenfunktionen und L^1-Halbnorm

Grundlage jedes Integralbegriffs ist das geometrisch definierte Integral von *Treppenfunktionen*. Für $A \subset \mathbb{R}^n$ definiert man die *charakteristische Funktion* $\mathbf{1}_A : \mathbb{R}^n \to \mathbb{R}$ *von* A als

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{für } x \in A \,, \\ 0 & \text{für } x \notin A \,. \end{cases}$$

Sei $Q \subset \mathbb{R}^n$ ein achsenparalleler n-dimensionaler Quader, wobei die Zugehörigkeit von Teilmengen der $(n-1)$-dimensionalen Randflächen zu Q keine Rolle spielt. *Treppenfunktionen* $\phi : \mathbb{R}^n \to \mathbb{C}$ sind (nach Unterteilung in disjunkte Teilquader) konstant auf den Teilquadern: $\phi(x) = \sum_{l=0}^N c_l \mathbf{1}_{Q_l}(x)$ mit $c_l \in \mathbb{C}$. Dadurch 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.

Definition 20.1 Das *Integral einer Treppenfunktion* $\phi = \sum_{l=0}^N c_l \mathbf{1}_{Q_l} : \mathbb{R}^n \to \mathbb{C}$ ist definiert als

$$\int_{\mathbb{R}^n} dx\ \phi(x) := \sum_{l=0}^N c_l\, v(Q_l) \,,$$

wobei $v(Q_l)$ das Volumen des Quaders Q_l ist.

Satz 20.2 *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) $\displaystyle\Big|\int_{\mathbb{R}^n} dx\ \phi(x)\Big| \leq \int_{\mathbb{R}^n} dx\ |\phi(x)| \,.$

iii) *Sind* ϕ, ψ *reellwertig mit* $\phi \leq \psi$, *so gilt* $\displaystyle\int_{\mathbb{R}^n} dx\ \phi(x) \leq \int_{\mathbb{R}^n} dx\ \psi(x) \,.$

Beweis. Durch Zurückführen auf eindimensionale Integrale für Treppenfunktionen mittels des folgenden Satzes. Die Eigenschaften i)–iii) sind dann klar. □

Satz 20.3 (Fubini für Treppenfunktionen) *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* $\Phi(y) := \displaystyle\int_{\mathbb{R}^p} dx\ \phi(x,y)$ *eine Treppenfunktion, und es gilt*

$$\int_{\mathbb{R}^p \times \mathbb{R}^{n-p}} d(x,y)\ \phi(x,y) = \int_{\mathbb{R}^{n-p}} dy \Big(\int_{\mathbb{R}^p} dx\ \phi(x,y) \Big) = \int_{\mathbb{R}^p} dx \Big(\int_{\mathbb{R}^{n-p}} dy\ \phi(x,y) \Big) \,.$$

Beweis. Die entsprechende Zerlegung der Quader sei $Q_l = Q'_l \times Q''_l$ mit $Q'_l \subset \mathbb{R}^p$ und $Q''_l \subset \mathbb{R}^{n-p}$. Dann gilt $v(Q_l) = v(Q'_l) \cdot v(Q''_l)$. Für festes $y \in \mathbb{R}^{n-p}$ definieren wir

$$\phi_y(x) := \sum_{l=0}^{N} \big(c_l \mathbf{1}_{Q''_l}(y)\big)\mathbf{1}_{Q'_l}(x)$$

$$\Rightarrow \quad \Phi(y) := \int_{\mathbb{R}^p} dx\ \phi_y(x) = \sum_{l=0}^{N} \big(c_l \mathbf{1}_{Q''_l}(y)\big)v(Q'_l) = \sum_{l=0}^{N} \big(c_l v(Q'_l)\big)\mathbf{1}_{Q''_l}(y)\ .$$

Also ist $\Phi = \sum_{l=0}^{N} \big(c_l v(Q'_l)\big)\mathbf{1}_{Q''_l} : \mathbb{R}^{n-p} \to \mathbb{R}$ eine Treppenfunktion. Ihr Integral ist

$$\int_{\mathbb{R}^{n-p}} dy\ \Phi(y) = \sum_{l=0}^{N} \big(c_l v(Q'_l)\big)v(Q''_l) = \sum_{l=0}^{N} c_l v(Q_l) = \int_{\mathbb{R}^p \times \mathbb{R}^{n-p}} d(x,y)\ \phi(x,y)\ .$$

Durch Vertauschen der Rollen von x, y entsteht der andere Teil der Gleichung. □

Das Integral beliebiger Funktionen $f : A \to \mathbb{C}$ mit $A \subset \mathbb{R}^n$ wird dann erklärt als Grenzwert der Integrale einer Folge $(\phi_k)_{k\in\mathbb{N}}$ von Treppenfunktionen, die f in geeignter Weise approximiert. Da Funktionen über A einen *unendlich-dimensionalen* Vektorraum bilden, gibt es verschiedene Konvergenzbegriffe, die dann auf verschiedene Integrale führen. In Analogie zum eindimensionalen Fall läßt sich auch ein mehrdimensionales Riemann-Integral einführen über die monotone beschränkte Konvergenz reellwertiger Treppenfunktionen:

Definition 20.4 Eine Funktion $f : A \to \mathbb{R}$ heißt *Riemann-integrierbar* über $A \subset \mathbb{R}^n$, falls es Folgen $(\phi_k)_{k\in\mathbb{N}}$ und $(\psi_k)_{k\in\mathbb{N}}$ von Treppenfunktionen gibt mit

- (ϕ_k) ist monoton wachsend mit $\phi_k(x) \leq f(x)$,
- (ψ_k) ist monoton fallend mit $\psi_k(x) \geq f(x)$,
- $\lim\limits_{k\to\infty} \int_A dx\ (\psi_k(x) - \phi_k(x)) = 0$

In diesem Fall definiert man das *Riemann-Integral* von f als $\int_A dx\ \ f(x) = \lim\limits_{k\to\infty} \int_A dx\ \psi_k(x) \equiv \lim\limits_{n\to\infty} \int_A dx\ \phi_k(x)$.

Dieses Riemann-Integral hat folgende Schwächen:

- Über $\langle f, g\rangle = \int_A dx\ \overline{f(x)}g(x)$ läßt sich ein rudimentäres Skalarprodukt erklären. Der Vektorraum der Riemann-integrierbaren Funktionen ist jedoch nicht vollständig in der induzierten (Halb-)Norm.

- Für das Riemann-Integral gilt folgender Vertauschungssatz: Ist (f_k) Folge Riemann-integrierbarer Funktionen, die *gleichmäßig gegen f konvergiert*, so gilt $\displaystyle\int_A dx \lim_{k\to\infty} f_k(x) = \lim_{k\to\infty}\int_A dx\, f_k(x)$. Es zeigt sich aber, daß gleichmäßige Konvergenz der Funktionsfolgen oft eine zu starke Forderung ist.

Das Lebesgue-Integral beruht auf einer anderen Konvergenz von Treppenfunktionen, in denen die L^1-Halbnorm eine zentrale Rolle spielt. Dazu werden die endlichen Summen in Treppenfunktion zu Reihen verallgemeinert, die den Wert ∞ annehmen dürfen. Wir vereinbaren $c+\infty = \infty$ für alle $c \in \mathbb{C}\cup\{\infty\}$ und $0\cdot\infty = 0$ sowie $c\cdot\infty = \infty$ für $c \in \mathbb{C}^\times \cup \{\infty\}$.

Definition 20.5 Eine Reihe $\Phi = \sum_{l=0}^{\infty} c_l \mathbf{1}_{Q_l}$ heißt *Hüllreihe* einer Funktion $f : \mathbb{R}^n \to \mathbb{C}\cup\{\infty\}$, wenn gilt:

i) Alle Quader $Q_l \in \mathbb{R}^n$ sind offen, und $c_l \geq 0$.

ii) Es gilt $|f(x)| \leq \Phi(x) = \sum_{l=0}^{\infty} c_l \mathbf{1}_{Q_l}(x)$ für alle $x \in \mathbb{R}^n$.

Der *Inhalt* einer Hüllreihe $\Phi = \sum_{l=0}^{\infty} c_l \mathbf{1}_{Q_l}$ ist definiert als $I(\Phi) := \sum_{l=0}^{\infty} c_l\, v(Q_l) \in \mathbb{R}^+ \cup \{\infty\}$, 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. $\Phi = \sum_{l=0}^{\infty} \mathbf{1}_{\hat{Q}_l}$, wobei $\hat{Q}_l \subset \mathbb{R}^n$ der offene Würfel mit Mittelpunkt 0 und Kantenlänge $2l+1$ ist.

Definition 20.6 Sei $f : \mathbb{R}^n \to \mathbb{C}\cup\{\infty\}$ eine beliebige Funktion. Dann ist die L^1-*Halbnorm* von f erklärt als

$$\|f\|_1 := \inf\{I(\Phi) \;:\; \Phi \text{ ist Hüllreihe zu } f\}\;.$$

Es folgt $\|f\|_1 = \|\,|f|\,\|_1$. *Achtung:* Entgegen der üblichen Bezeichnung ist die L^1-Halbnorm $\|\;\|_1$ *keine Norm*! Nach folgendem Beispiel gilt: $\|f\|_1 = 0 \;\not\Rightarrow\; f = 0$.

Beispiel 20.7 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 $\|\mathbf{1}_Q\|_1 = 0$. ◁

Es gelten aber die anderen Normeigenschaften:

Satz 20.8 *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* $\Big\| \sum\limits_{k=0}^{\infty} f_i \Big\|_1 \leq \sum\limits_{k=0}^{\infty} \|f_i\|_1$.

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 $\Phi_k = \sum_{l=0}^{\infty} c_{kl} \mathbf{1}_{Q_{kl}}$ mit $I(\Phi_k) = \sum_{l=0}^{\infty} c_{kl} v(Q_{kl}) \leq \|f_k\|_1 + \frac{\epsilon}{2^k}$. Dann ist $\Phi := \sum_{k=0}^{\infty} \Phi_k = \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{\infty} c_{kl} \mathbf{1}_{Q_{kl}} \Big)$ Hüllreihe zu $\sum_{k=0}^{\infty} f_k$, mit

$$I(\Phi) = \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{\infty} c_{kl} v(Q_{kl}) \Big) \leq \sum_{k=0}^{\infty} \Big(\|f_k\|_1 + \frac{\epsilon}{2^k} \Big) = \sum_{k=0}^{\infty} \|f_k\|_1 + 2\epsilon \;.$$

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). $\square$

Sei V ein Vektorraum. Eine Abbildung $\| \; \| : V \to \mathbb{R}_+$, die i) und ii) wie in Satz 20.8 erfüllt, nennt man Halbnorm. Eine allgemeine Halbnorm *genügt nicht* zur Definition einer Konvergenz (im üblichen Sinn), da der Grenzwert nicht eindeutig ist. Die zusätzlichen Bedingungen iii) und iv) sowie der folgende Satz 20.10 liefern jedoch einen Konvergenzbegriff für die spezielle L^1-Halbnorm.

Satz 20.9 *Für einen abgeschlossenen Quader* $A \subset \mathbb{R}^n$ *gilt* $\|\mathbf{1}_A\|_1 = v(A) = \displaystyle\int_{\mathbb{R}^n} dx \; \mathbf{1}_A(x)$.

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 $\Phi = \sum_{l=0}^{\infty} c_l \mathbf{1}_{Q_l}$ zu $\mathbf{1}_A$, mit Q_l 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 $\sum_{l=0}^{N(x)} c_l \mathbf{1}_{Q_l}(x) \geq 1 - \epsilon$. Da die Q_l 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 $\mathbf{1}_{Q_l}$ auf Q_l gilt

$$\sum_{l=0}^{N(x)} c_l \mathbf{1}_{Q_l}(x) \geq 1 - \epsilon \qquad \text{für alle } x \in U(x) \;.$$

Da A kompakt ist, wird A durch endlich viele $U(x_1), \dots, U(x_p)$ überdeckt. Also gibt es ein $N = \max(N(x_1), \dots, N(x_p))$, so daß $\sum_{l=0}^{N} c_l \mathbf{1}_{Q_l}(x) \geq 1 - \epsilon = (1 - \epsilon)\mathbf{1}_A(x)$ für alle $x \in A$. Da $\sum_{l=0}^{N} c_l \mathbf{1}_{Q_l}(x)$ eine Treppenfunktion ist, folgt mit Satz 20.2.iii):

$$I(\Phi) = \sum_{l=0}^{\infty} c_l v(Q_l) \geq \sum_{l=0}^{N} c_l v(Q_l) \geq (1-\epsilon) v(A) \;.$$

Also gilt $\|\mathbf{1}_A\|_1 = v(A) = \int_{\mathbb{R}^n} dx\; \mathbf{1}_A(x)$. $\square$

Satz 20.10 *Für jede Treppenfunktion $\phi = \sum_{l=1}^{N} c_l \mathbf{1}_{Q_l}$ 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_l, deren Kanten offen, abgeschlossen, halboffen und entartet sein dürfen, in disjunkte offene Quader Q_i^o und disjunkte Ränder R_j, so daß

$$\phi = \sum_{i=1}^{N'} c'_i \mathbf{1}_{Q_i^o} + \sum_{j=1}^{M} d_j \mathbf{1}_{R_j} \;, \qquad c'_i, d_j > 0 \;.$$

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 $\Phi_\epsilon := \sum_{i=0}^{N'} c'_i \mathbf{1}_{Q_i^o} + \sum_{j=0}^{M} d_j \mathbf{1}_{R_j^0}$ eine Hüllreihe zu ϕ mit

$$\|\phi\|_1 \leq \inf_\epsilon I(\Phi_\epsilon) = \inf_\epsilon \Big(\sum_{i=0}^{N'} c'_i v(Q_i^o) + \epsilon \sum_{j=0}^{M} d_j \Big) = \int_{\mathbb{R}^n} dx\; \phi(x) \;.$$

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

$$\begin{aligned}\int_{\mathbb{R}^n} dx\; \phi(x) &= \int_{\mathbb{R}^n} dx\; (m\mathbf{1}_A(x) - \psi(x)) = m \int_{\mathbb{R}^n} dx\; \mathbf{1}_A(x) - \int_{\mathbb{R}^n} dx\; \psi(x) \\ &\leq \|m1_A\|_1 - \|\psi\|_1 = \|\phi + \psi\|_1 - \|\psi\|_1 \leq \|\phi\|_1 \;.\end{aligned}$$ $\square$

21 Das Lebesgue-Integral

Definition 21.1 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.

Eine Funktion $f : A \to \mathbb{C} \cup \{\infty\}$ heißt *integrierbar über eine Teilmenge* $A \subset \mathbb{R}^n$, wenn die Funktion $f_A := f \cdot \mathbf{1}_A$ integrierbar ist (über $\mathbb{R}^n$), und man setzt $\displaystyle\int_A dx\; f(x) := \int_{\mathbb{R}^n} dx\; f_A(x)$.

Zur Definition ist zu bemerken, daß die Folge der Integrale $\Big(\displaystyle\int_{\mathbb{R}^n} dx\; \phi_k(x)\Big)_{k\in\mathbb{N}}$ 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 $m, l \geq k$

$$\begin{aligned} \Big|\int_{\mathbb{R}^n} dx\; \phi_l(x) - \int_{\mathbb{R}^n} dx\; \phi_m(x)\Big| &= \Big|\int_{\mathbb{R}^n} dx\; (\phi_l - \phi_m)(x)\Big| \leq \int_{\mathbb{R}^n} dx\; |(\phi_l - \phi_m)(x)| \\ \text{Satz 20.10} \quad \Rightarrow \qquad &= \|\phi_l - \phi_m\|_1 \leq \|\phi_l - f\|_1 + \|f - \phi_m\|_1 < \epsilon\;. \end{aligned}$$

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. Riemann-integrierbare Funktionen auf kompakten Intervallen sind auch Lebesgue-integrierbar:

Satz 21.2 *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,* $\displaystyle\int_{[a,b]} dx\; f(x) = \int_a^b dx\; f(x)$.

Beweis. Wir können f als reellwertig annehmen. Zu $\epsilon = \frac{1}{k+1}$ gilt es Treppenfunktionen ϕ_k, ψ_k mit $\phi_k \leq f \leq \psi_k$ und $\displaystyle\int_a^b dx(\psi_k - \phi_k)(x) \leq \frac{1}{k+1}$. Da für Treppenfunktionen Lebesgue- und Riemann-Integral übereinstimmen, gilt $\|\psi_k - \phi_k\|_1 \leq \frac{1}{k+1}$ nach Satz 20.10. Aus $0 \leq \psi_k - f \leq \psi_k - \phi_k$ folgt $\|\psi_k - f\|_1 \leq \|\psi_k - \phi_k\|_1 \leq \frac{1}{k+1}$ nach Satz 20.8. 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. □

Die folgende Eigeschaft des Lebesgue-Integrals ist grundlegend verschieden vom Riemann-Integral:

Satz 21.3 *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 20.10

$$\Big(\int_{\mathbb{R}^n} dx\; |\phi_k|(x)\Big) - \|\phi_k - f\|_1 \leq \|f\|_1 \leq \Big(\int_{\mathbb{R}^n} dx\; |\phi_k|(x)\Big) + \|\phi_k - f\|_1\;.$$

Für $k\to\infty$ erhalten wir $\|f\|_1 = \displaystyle\int_{\mathbb{R}^n} dx\; |f(x)|$. □

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. Dagegen folgt (Majorantenkriterium Satz 30.8 aus dem 2. Semester) für *Regelfunktionen* aus der uneigentlichen Riemann-Integrierbarkeit von $|f|$ die uneigentliche Riemann-Integrierbarkeit von f.

Satz 21.4 *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 21.3.

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.

Definition 21.5 Eine Menge $A \subset \mathbb{R}^n$ heißt *(Lebesgue-) meßbar*, wenn die Funktion 1 über A integrierbar ist. Das n-dimensionale Volumen $v_n(A) := \int_A dx$ heißt dann auch das *Lebesgue-Maß* von A. Die leere Menge hat das Lebesgue-Maß Null.

Satz 21.6 *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 $v_n(A) \leq v_n(B)$.*

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

$$\Big|\int_{A\cap B} dx\; f(x)\Big| \leq \int_A dx\; \big|f(x)\big|\,.$$

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 $\mathbf{1}_A \leq \mathbf{1}_B$.

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 21.4.iii). Die Abschätzung folgt aus $|f_{A\cap B}| \leq |f_A|$. □

Wir zeigen nun, daß Singularitäten, die nur auf einer Teilmenge vom Maß Null auftreten, im Lebesgue-Integral keine Rolle spielen.

Definition 21.7 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 21.8 i) *Jede Teilmenge einer Nullmenge ist eine Nullmenge.*

ii) *Die Vereinigung abzählbar vieler Nullmengen ist eine Nullmenge.*

Beweis. i) $M \subset N \quad \Rightarrow \quad \mathbf{1}_M \leq \mathbf{1}_N \quad \Rightarrow \quad \|\mathbf{1}_M\|_1 \leq \|\mathbf{1}_N\| = 0$.

ii) Ist $N = \bigcup_{k=1}^\infty N_k$ mit $v_n(N_k) = 0$, dann ist $0 \leq \mathbf{1}_N(x) \leq \sum_{k=1}^\infty \mathbf{1}_{N_k}(x)$ und somit $\|\mathbf{1}_N\|_1 \leq \sum_{k=1}^\infty \|\mathbf{1}_{N_k}\|_1 = 0$. □

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 $\int_{\mathbb{R}} dx\; \mathbf{1}_{\mathbb{Q}}(x) = 0$.

Definition 21.9 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 21.10 *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 21.11 (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 $\displaystyle\int_{\mathbb{R}^n} dx\; f(x) = \int_{\mathbb{R}^n} dx\; g(x)$. 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 $\lim\limits_{k\to\infty} \|f - \phi_k\|_1 = 0$. 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 $\displaystyle\int_{\mathbb{R}^n} dx\ g(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ \phi_k(x) = \int_{\mathbb{R}^n} dx\ f(x)$. □

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 $\displaystyle\int_{A\cup B} dx\ f(x) = \int_A dx\ f(x) + \int_B dx\ f(x)$.

Satz 21.12 *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. □

22 Vollständigkeit des Lebesgue-Integrals

Definition 22.1 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 22.2 (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 21.10 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 22.3 *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 21.12

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. Die Menge aller Äquivalenzklassen

$$L^1(\mathbb{R}^n) := \mathcal{L}^1(\mathbb{R}^n)/\mathcal{N}(\mathbb{R}^n) = \{[f] \;:\; f \in \mathcal{L}^1(\mathbb{R}^n)\}$$

wird zu einem Vektorraum mit $[0] = \mathcal{N}(\mathbb{R}^n)$ und $c_1[f_1] + c_2[f_2] := [c_1 f_1 + c_2 f_2]$.

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

$$\|f\|_1 = \|g + (f - g)\|_1 \leq \|g\|_1 + \|f - g\|_1 = \|g\|_1 \;,$$

und umgekehrt $\|g\|_1 \leq \|f\|_1$. 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 21.12 $\|[f]\|_1 = 0$ 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.

23 Berechnung von Lebesgue-Integralen über Fubini

Wir geben zunächst ohne Beweis die wichtigste Methode an, um *stetige Funktionen über kompakte Teilmengen* zu integrieren.

Satz 23.1 (kleiner Satz von Fubini) *Es sei $A \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$ eine kompakte Teilmenge und $f : A \to \mathbb{R}$ stetig. 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 $A_x := \{y \in \mathbb{R}^{n-p} \;:\; (x, y) \in A\} \subset \mathbb{R}^{n-p}$. Dann gilt:*

i) *$f : A \to \mathbb{R}$ ist integrierbar.*

ii) *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.

iii) *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.

iv) *Es gilt*

$$\int_A d(x,y)\ f(x,y) = \int_{\mathbb{R}^{n-p}} dy\ F(y) = \int_{\mathbb{R}^n} dx\ G(x)\ .$$

Der Beweis wird später nachgeholt.

In Kombination mit Satz 21.2 können wir das Lebesgue-Integral schrittweise auf eindimensionale Riemann-Integrale zurückführen ($p = 1$). Für $p = 1$ und A kompakt ist $A_y = \bigcup_{k=1}^N [x_{1k}(y), x_{2k}(y)] \subset \mathbb{R}$ Vereinigung von Intervallen. Mit $B := \{y \in \mathbb{R}^{n-1} \ :\ A_y \neq \varnothing\} \subset \mathbb{R}^{n-1}$ gilt dann

$$\int_A d(x,y)\ f(x,y) = \int_B dy \Big(\sum_{k=1}^N \int_{x_{1k}(y)}^{x_{2k}(y)} dx\ f(x,y) \Big)\ .$$

Insbesondere gilt für die Integration stetiger Funktionen über das Rechteck $[a,b] \times [c,d] \subset \mathbb{R}^2$

$$\int_{[a,b]\times[c,d]} d(x,y)\ f(x,y) = \int_c^d dy \Big(\int_a^b dx\ f(x,y) \Big) = \int_a^b dx \Big(\int_c^d dy\ f(x,y) \Big)\ .$$

Definition 23.2 Das *Volumen* einer kompakte Menge $A \subset \mathbb{R}^n$ ist gegeben durch das Integral $v_n(A) = \displaystyle\int_{\mathbb{R}^n} dx\ \mathbf{1}_A(x) = \int_A dx$. Ist $\mu : A \to \mathbb{R}$ die *Dichte* von A, so ist die (n-dimensionale) *Masse* von A erklärt als $m_n(A) = \displaystyle\int_A dx\ \mu(x)$.

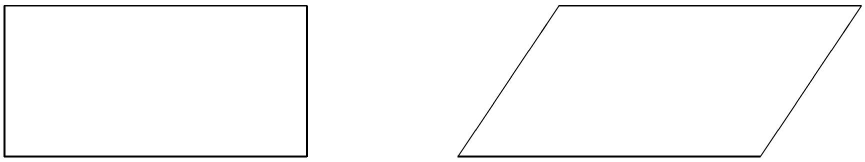
Der kleine Satz von Fubini ermöglicht (da $\mathbf{1}_A$ stetig auf A) die schrittweise Berechnung von Volumina. Ist $A \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$ und $A_y := \{x \in \mathbb{R}^p \ :\ (x,y) \in A\} \subset \mathbb{R}^p$ die in Satz 27.1 eingeführte Schnittmenge, dann gilt

$$v_n(A) = \int_{\mathbb{R}^{n-p}} dy\ v_p(A_y)\ .$$

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):



Beispiel 23.4 Es sei $A = \{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \leq r\}$ der abgeschlossene Vollkreis vom Radius r. In obigen Bezeichnungen ist $A_y = [-\sqrt{r^2-y^2}, \sqrt{r^2-y^2}]$ und $B = [-r,r]$. 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

$$\begin{aligned} v_2(A) &= \int_A d(x,y)\ 1 = \int_{-r}^{r} dy\ \Big(\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} dx\ 1 \Big) = \int_{-r}^{r} dy\ 2\sqrt{r^2-y^2} \\ &= 4\int_0^r dy\ \sqrt{r^2-y^2} \overset{y=r\cos t}{=} 4r^2 \int_0^{\frac{\pi}{2}} dy\ \sin^2 t = 4r^2 \cdot \frac{1}{2} \cdot \frac{\pi}{2} = \pi r^2 \end{aligned}$$

nach Beispiel 29.3 aus dem 2. Semster. ◁

Satz 23.5 (Volumen der n-dimensionalen Vollkugel)
Das Volumen $v_n(A) =: \kappa_n(r)$ der n-dimensionalen Vollkugel $A = \{(x_1,\dots,x_n) \in \mathbb{R}^n \ : \ x_1^2 + \dots + x_n^2 \leq r^2\}$ vom Radius r ist

$$\kappa_n(r) = \frac{\pi^{\frac{n}{2}} r^n}{\Gamma(\frac{n+2}{2})} = \begin{cases} \dfrac{\pi^{\frac{n}{2}}}{\frac{n}{2}!}\ r^n & \text{für } n \text{ gerade} \\[2ex] \dfrac{2^n \pi^{\frac{n-1}{2}}\, \frac{n-1}{2}!}{n!}\ r^n & \text{für } n \text{ ungerade} \end{cases}$$

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 $A_{x_n} = \{(x_1,\dots,x_{n-1}) \in \mathbb{R}^{n-1} \ : \ x_1^2 + \dots + x_{n-1}^2 \leq r^2 - x_n^2\}$ 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 29.3 aus dem 2. 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

$$\begin{aligned}\kappa_n(r) &= \int_{-r}^{r} dx_n \int_{A_{x_n}} dy\ 1 = \int_{-r}^{r} dx_n \kappa_{n-1}(\sqrt{r^2-x_n^2}) = 2\int_0^r dx_n \kappa_{n-1}(\sqrt{r^2-x_n^2}) \\ &= \frac{2\pi^{\frac{n-1}{2}}}{\frac{n-1}{2}!} \int_0^r dx_n\ (r^2-x_n^2)^{\frac{n-1}{2}} \stackrel{x_n=r\cos t}{=} \frac{2\pi^{\frac{n-1}{2}}}{\frac{n-1}{2}!} r^n \int_0^{\frac{\pi}{2}} dt\ \sin^n t\ .\end{aligned}$$

Nach Beispiel 29.3 aus dem 2. 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. □

Definition 23.6 Das Trägheitsmoment einer kompakten Menge $A \subset \mathbb{R}^n$ bezüglich einer Achse L ist definiert als

$$\Theta(A) := \int_A dx\ \mu(x) \big(\mathrm{d}(x,L)\big)^2\ ,$$

wobei μ die Dichte und $\mathrm{d}(x,L)$ der Abstand eines Punktes $x \in A$ zur Drehachse L ist.

Beispiel 23.7 (Trägheitsmoment eines homogenen Kreiszylinders) Es sei

$$A = \{(x,y,z) \in \mathbb{R}^3\ :\ x^2+y^2 \leq r^2\ ,\ 0 \leq z \leq h\}$$

ein homogener (konstante Diche μ) gerader Kreiszylinder, die z-Achse sei die Drehachse L. Dann gilt $\mathrm{d}((x,y,z),L) = \sqrt{x^2+y^2}$, so daß wir für das Trägheitsmoment nach Fubini erhalten

$$\begin{aligned}\Theta(A) &= \int_0^h dz \int_{-r}^{r} dy \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} dx\ \mu\cdot(x^2+y^2) \\ &= \mu h \int_{-r}^{r} dy\ \Big(\frac{1}{3}x^3+y^2x\Big)\Big|_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \\ &= \mu h \int_{-r}^{r} dy\ \Big(\frac{2}{3}\big(\sqrt{r^2-y^2}\big)^3 + 2y^2\sqrt{r^2-y^2}\Big) \\ &= 4\mu h \int_0^r dy\ \Big(\frac{1}{3}\big(\sqrt{r^2-y^2}\big)^3 + y^2\sqrt{r^2-y^2}\Big)\end{aligned}$$

$$
\begin{aligned}
&\overset{y=r\cos t}{=} 4\mu h \int_{\frac{\pi}{2}}^{0} dt\,(-r\sin t)\Big(\frac{1}{3}r^3\sin^3 t + r^3\cos^2 t\sin t\Big)\\
&= 4\mu h r^4 \int_0^{\frac{\pi}{2}} dt\,\Big(\sin^2 t - \frac{2}{3}\sin^4 t\Big)\\
&= 4\mu h r^4\Big(\frac{1}{2}\cdot\frac{\pi}{2} - \frac{2}{3}\cdot\frac{3\cdot 1}{4\cdot 2}\cdot\frac{\pi}{2}\Big) && \leftarrow \text{Beispiel 29.3 (2. Semester)}\\
&= \frac{\pi}{2}\mu h r^4 = \frac{1}{2}mr^2\;,
\end{aligned}
$$

wobei $m = \mu v_3(A) = \mu\pi r^2 h$ die Gesamtmasse des Zylinders ist. $\triangleleft$

Definition 23.8 Der *Schwerpunkt* einer kompakten Menge $A \subset \mathbb{R}^n$ der Dichte μ ist der Vektor $s = (s_1, \dots, s_n) \subset \mathbb{R}^n$ mit

$$
s_i = \frac{\displaystyle\int_A d(x_1,\dots,x_n)\; x_i\cdot\mu(x_1,\dots,x_n)}{\displaystyle\int_A d(x_1,\dots,x_n)\;\mu(x_1,\dots,x_n)}\;.
$$

Beispiel 23.9 Es sei $A = \{(x,y,z) \in \mathbb{R}^3\;,\; x^2+y^2+z^2 \leq r^3\;,\; z \geq 0$ die halbe homogene Vollkugel vom Radius r und Dichte μ. Der Nenner ist die Masse $m = \mu v_3(A) = \frac{2\pi\mu}{3}r^3$. Für den Schwerpunkt $s = (s_x, s_y, s_z)$ gilt $s_x = s_y = 0$ aus Symmetriegründen. Es verbleibt

$$
\begin{aligned}
s_z &= \frac{1}{\mu v_3(K)}\int_0^r dz \int_{-\sqrt{r^2-z^2}}^{\sqrt{r^2-z^2}} dy \int_{-\sqrt{r^2-z^2-y^2}}^{\sqrt{r^2-z^2-y^2}} dx\;\mu z\\
&= \frac{3}{2\pi r^3}\int_0^r dz \int_{-\sqrt{r^2-z^2}}^{\sqrt{r^2-z^2}} dy\; 2z\sqrt{r^2-z^2-y^2}\\
&= \frac{6}{\pi r^3}\int_0^r dz \int_0^{\sqrt{r^2-z^2}} dy\; z\sqrt{r^2-z^2-y^2}\\
&\overset{y=\sqrt{r^2-z^2}\cos t}{=} \frac{6}{\pi r^3}\int_0^r dz \int_{\frac{\pi}{2}}^{0} dt\;\big(-\sqrt{r^2-z^2}\sin t\big) z\sqrt{r^2-z^2}\sin t\\
&= \frac{6}{\pi r^3}\int_0^r dz\; z(r^2-z^2)\cdot\frac{1}{2}\cdot\frac{\pi}{2} && \leftarrow \text{Beispiel 29.3 (2. Semester)}\\
&= \frac{3}{2r^3}\Big(\frac{r^2}{2}\cdot r^2 - \frac{r^4}{4}\Big) = \frac{3}{8}r\;.
\end{aligned}
$$

$\triangleleft$

24 Konvergenzsätze

Wir zeigen nun, daß im Gegensatz zum Riemann-Integral das Lebesgue-Integral sogar mit Grenzwerten von monotonen beschränkten Funktionsfolgen vertauscht.

Satz 24.1 (von der monotonen Konvergenz bzw. Satz von Beppo Levi) *Es sei* $(f_k)_{k\in\mathbb{N}}$ *eine monoton wachsende (oder monoton fallende) Folge integrierbarer reellwertiger 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* $\displaystyle\int_{\mathbb{R}^n} dx\ f_k(x)$ *beschränkt ist. In diesem Fall gilt* $\displaystyle\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty}\int_{\mathbb{R}^n} dx\ f_k(x)$*.*

Beweis. ($\Leftarrow$) Wegen $f_k \leq f$ ist $\displaystyle\int_{\mathbb{R}^n} dx\ f_k(x) \leq \int_{\mathbb{R}^n} dx\ f(x)$, so daß die Beschränktheit der Integrale notwendig ist.

($\Rightarrow$) Die Folge $\displaystyle\Big(\int_{\mathbb{R}^n} dx\ f_k(x)\Big)_{k\in\mathbb{N}}$ 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

$$\begin{aligned}\|f_k - f_l\|_1 &= \int_{\mathbb{R}^n} dx\ \big|f_k(x) - f_l(x)\big| = \int_{\mathbb{R}^n} dx\ \big(f_k(x) - f_l(x)\big) \\ &= \int_{\mathbb{R}^n} dx\ f_k(x) - \int_{\mathbb{R}^n} dx\ f_l(x) = \Big|\int_{\mathbb{R}^n} dx\ f_k(x) - \int_{\mathbb{R}^n} dx\ f_l(x)\Big| < \epsilon\ .\end{aligned}$$

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 wenden dieses Approximationsverfahren auf uneigentliche Riemann-Integrale von Regelfunktionen an. Das waren nach Definition 28.1 aus dem 2. 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 $f(x) - \epsilon \leq \phi(x) \leq f(x)$ und $f(x) \leq \psi(x) \leq f(x) + \epsilon$. Regelfunktionen sind Riemann-integrierbar, und für sie gilt der Hauptsatz der Differential- und Integralrechnung. Jede stetige Funktion ist Regelfunktion.

Satz 24.2 *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* $\displaystyle\lim_{k\to\infty}\int_{a_k}^{b_k} dx\ |f(x)| < \infty$*. In diesem Fall gilt* $\displaystyle\int_{]a,b[} dx\ f(x) = \int_a^b dx\, f(x)$*.*

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 $f^{\pm}_{[a_k,b_k]} \leq f^{\pm}_{]a,b]}$ gilt mit Satz 21.2

$$\int_{a_k}^{b_k} dx\ f^{\pm}(x) = \int_{[a_k,b_k]} dx\ f^{\pm}(x) \leq \int_{]a,b[} dx\ f^{\pm}(x) \ .$$

Damit existiert $\displaystyle\int_a^b dx\ f^{\pm}(x) := \lim_{k\to\infty} \int_{a_k}^{b_k} dx\ f^{\pm}(x)$ und ist durch $\displaystyle\int_{]a,b[} dx\ f^{\pm}(x)$ beschränkt.

($\Leftarrow$) Sei $|f| :]a, b[\to \mathbb{R}$ uneigentlich Riemann-integrierbar. Nach dem Majorantenkriterium Satz 30.8 aus dem 2. 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 Satz von Beppo Levi ist dann auch die Grenzfunktion $f^{\pm}_{]a,b[} = \lim_{k\to\infty} \phi_k^{\pm}$ Lebesgue-integrierbar mit

$$\int_{]a,b[} dx\ f^{\pm}(x) = \lim_{k\to\infty} \int_a^b dx\ \phi_k^{\pm}(x) \leq \int_a^b dx\ f^{\pm}(x) \ .$$

Zusammen mit der umgekehrten Abschätzung aus ($\Rightarrow$) folgt dann $\displaystyle\int_{]a,b[} dx\ f^{\pm}(x) = \int_a^b dx\ f^{\pm}(x)$. □

Die absolute Konvergenz des Riemann-Integrals kann nicht auf einfache Konvergenz reduziert werden. Z.B. gilt für das uneigentliche Riemann-Integral $\displaystyle\int_{-\infty}^{\infty} dx\ \frac{\sin x}{x} = \pi$, 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 holen nun den Beweis nach, daß stetige beschränkte Funktionen über kompakte und beschränkte offene Teilmengen integrierbar sind. Dazu konstruieren wir geeignete Approximationen durch Treppenfunktionen:

Lemma 24.3 *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*

$$f(x) - \epsilon \leq \phi_\epsilon(x) \leq f(x) \leq \psi_\epsilon(x) \leq f(x) + \epsilon \qquad \textit{für alle } x \in K$$

und $\phi_\epsilon(x) = \psi_\epsilon(x) = 0$ für alle $x \in \mathbb{R}^n \setminus U$.

Beweis. Stetige Funktionen auf kompakten Mengen sind gleichmäßig stetig (Satz 21.14 aus dem 1. Semester), 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 $W_1^0, \dots, W_p^0$, die K überdecken, und insbesondere ist $K \subset W_1 \cup \dots \cup W_p \subset U$. 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

$$\psi_k := \max(M_1 \mathbf{1}_{W_1}, \dots M_p \mathbf{1}_{W_p}) \,, \qquad \phi_\epsilon := \min(m_1 \mathbf{1}_{W_1}, \dots m_p \mathbf{1}_{W_p})$$

Treppenfunktionen mit den geforderten Eigenschaften. □

Lemma 24.4 *Jede offene Teilmenge $U \subset \mathbb{R}^n$ ist Vereinigung abzählbar vieler Würfel, die höchstens Randpunkte gemeinsam haben.*

Beweis. Für $k \in \mathbb{N}$ sei $\mathcal{W}_k$ die Menge aller Würfel $W = I_1 \times \dots \times I_n$, so daß für jedes j eine ganze Zahl $m_j \in \mathbb{Z}$ existiert mit $I_j = [\frac{m_j}{2^k}, \frac{m_j+1}{2^k}]$. Dann ist $\mathcal{W}_k$ abzählbar durch die ganzen Zahlen $m_1, \dots, m_n$. Nach Konstruktion schneiden sich beliebige Würfel aus $\mathcal{W}_k, \mathcal{W}_l$ höchstens in Randpunkten. Dann ist $U = \bigcup_{k=0}^\infty W_k$ mit $W_0 := \mathcal{W}_0 \cap U$ und dann rekursiv W_k als jene Teilmenge von $\mathcal{W}_k \cap U$, deren Inneres nicht in W_l liegt für $l < k$. □

Satz 24.5
i) *Sei $U \subset \mathbb{R}^n$ eine beschränkte offene Teilmenge. Dann ist jede beschränkte stetige Funktion $f : U \to \mathbb{R}$ über U integrierbar. Insbesondere ist U meßbar.*

ii) *Sei $K \subset \mathbb{R}^n$ eine kompakte Teilmenge. Dann ist jede stetige Funktion $g : K \to \mathbb{R}$ über K integrierbar. Insbesondere ist K meßbar.*

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) Nach Lemma 24.4 ist U Vereinigung abzählbar vieler kompakter Teilmengen $K_j \subset U$, d.h. $U = \bigcup_{j=0}^\infty K_j$. Die Einschränkung von f auf $\bigcup_{j=0}^l K_j$ wird nach Lemma 24.3 von unten gleichmäßig durch eine Treppenfunktion ϕ_l zu $\epsilon_l = \frac{1}{2^l}$ approximiert. Dann ist $(\phi_l)_{l \in \mathbb{N}}$ 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 $0 \leq \phi_j(x) \leq f_U(x) \leq M \mathbf{1}_Q(x)$ für alle $x \in \mathbb{R}^n$. Also gilt

$$\int_{\mathbb{R}^n} dx \; \phi_l(x) \leq M v_Q < \infty \,,$$

so daß f_U nach Satz 24.1 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 24.3 von oben durch eine monoton fallende Folge ψ_l von Treppenfunktionen zu $\epsilon_l = \frac{1}{2^l}$, die außerhalb $\bigcap_{j=0}^{l} U_j$ verschschwinden, punktweise approximiert. Die Folge der Integrale $\Big(\displaystyle\int_{\mathbb{R}^n} dx\ \psi_j(x)\Big)_{j\in\mathbb{N}}$ ist nach unten durch 0 beschränkt, so daß f_K nach Satz 24.1 integrierbar ist. □

Wir benötigen ein weiteres Integrierbarkeitskriterium:

Satz 24.6 (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 $|f_k| \leq F$ für alle $k \in \mathbb{N}$. Dann ist f integrierbar, und es gilt $\displaystyle\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ f_k(x)$.*

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 $g_{k,\nu} = \max(f_k, f_{k+1}, \dots, f_{k+\nu})$. 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 $g_k := \lim_{\nu\to\infty} g_{k,\nu} = \sup_{j\geq k} f_j$ integrierbar, mit

$$\Big|\int_{\mathbb{R}^n} dx\ g_k(x)\Big| \leq \int_{\mathbb{R}^n} dx\ F(x)\ .$$

Andererseits konvergiert $(g_k)_{k\in\mathbb{N}}$ monoton fallend punktweise 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 $\displaystyle\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ g_k(x)$.

Das Verfahren wird wiederholt für $h_{k,\nu} = \min(f_k, f_{k+1}, \dots, f_{k+\nu})$ mit Limes und $h_k := \lim_{k\to\infty} h_{k,\nu} = \inf_{j\geq k} f_j$ und $\displaystyle\int_{\mathbb{R}^n} dx\ f(x) = \lim_{k\to\infty} \int_{\mathbb{R}^n} dx\ h_k(x)$. Aus $h_k \leq f_k \leq g_k$ folgt die Behauptung. □

Der Satz von Lebesgue wird insbesondere beim Vertauschen von Integral und Reihenentwicklung benutzt:

Beispiel 24.7 Wir beweisen $\displaystyle\int_0^{\infty} dx\ \frac{x^{s-1}}{e^x - 1} = \Gamma(s)\zeta(s)$ 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

$$|f_k| \leq F = \sum_{n=1}^{\infty} x^{\mathrm{Re}(s)-1}e^{-nx} = \frac{x^{\mathrm{Re}(s)-1}}{e^x - 1} ,$$

und für $\mathrm{Re}(s) > 1$ ist F wieder Lebesgue-integrierbar. Somit gilt nach dem Satz von Lebesgue

$$\int_0^\infty dx\, \frac{x^{s-1}}{e^x - 1} = \lim_{k\to\infty} \sum_{n=1}^{k} \int_0^\infty dx\; x^{s-1}e^{-nx} = \lim_{k\to\infty} \sum_{n=1}^{k} \frac{1}{n^s}\Gamma(s) = \zeta(s)\Gamma(s) . \qquad \square$$

25 Der Transformationssatz

Der Transformationssatz ist eine mächtige Methode zur Berechnung von Integralen.

Satz 25.1 *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 5.5 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.

25.1 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 $\varphi, \vartheta_1, \dots, \vartheta_{n-2}$. Dann ist $T : (r, \varphi, \vartheta_1, \dots, \vartheta_{n-1}) \mapsto (y_1, \dots, y_n) = T_n(r, \varphi, \vartheta_1, \dots, \vartheta_{n-2})$ definiert durch $T_2(r,\varphi) = (r\cos\varphi, r\sin\varphi)$ und dann rekursiv $y_n = r\cos\vartheta_{n-2}$ und $(y_1, \dots, y_{n-1}) = T_{n-1}(r, \varphi, \vartheta_1, \dots, \vartheta_{n-3}) \cdot \sin\vartheta_{n-2}$. Konkret heißt das

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} r\cos\varphi\sin\vartheta_1\cdots\sin\vartheta_{n-2} \\ r\sin\varphi\sin\vartheta_1\cdots\sin\vartheta_{n-2} \\ r\cos\vartheta_1\sin\vartheta_2\cdots\sin\vartheta_{n-2} \\ \vdots \\ r\cos\vartheta_{n-3}\sin\vartheta_{n-2} \\ r\cos\vartheta_{n-2} \end{pmatrix}.$$

Damit die Transformation T bijektiv wird, ist (z.B.) $\vartheta_1, \dots, \vartheta_{n-2} \in \,]0,\pi[$ und $\varphi \in \,]0, 2\pi[$ zu wählen.

Satz 25.2 *Es gilt* $|\det(DT)(r, \varphi, \vartheta_1, \dots, \vartheta_{n-2})| = r^{n-1}(\sin\vartheta_1)^1 \cdots (\sin\vartheta_{n-2})^{n-2}$.

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 25.3 *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 $f(T(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2})) \cdot r^{n-1} C(\varphi, \vartheta_1, \ldots, \vartheta_{n-2})$ über $I \times \Pi$ integrierbar ist. In diesem Fall gilt (unter Verwendung des später bewiesenen Satzes von Fubini)*

$$\begin{aligned} &\int_{K(I)} dy\; f(y) \\ &= \int_I dr\; r^{n-1} \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta_1\; \sin \vartheta_1 \cdots \int_0^\pi d\vartheta_{n-2}\; \sin^{n-2} \vartheta_{n-2}\; f(T(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2}))\,. \end{aligned}$$

□

Beispiel 25.4 Das Trägheitsmoment einer homogenen (dreidimensionalen) Vollkugel K vom Radius R mit Dichte μ ist bezüglich einer durch den Mittelpunkt gehenden Achse gegeben durch

$$
\begin{aligned}
\Theta(K) &= \mu \int_K d(x,y,z)\ (x^2+y^2) \\
&= \mu \int_0^R dr\ r^2 \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta\ \sin\vartheta\ (r^2 \sin^2\vartheta) \\
&= \mu \frac{R^5}{5} \cdot 2\pi \int_0^\pi d\vartheta\ \sin^3\vartheta = \mu \frac{R^5}{5} \cdot 2\pi \int_0^\pi d\vartheta\ (\sin\vartheta - \sin\vartheta\cos^2\vartheta) \\
&= \mu \frac{R^5}{5} \cdot 2\pi \Big(-\cos\vartheta + \frac{1}{3}\cos^3\vartheta \Big)\Big|_0^\pi \\
&= \mu \frac{R^5}{5} \cdot 2\pi \Big(2 - \frac{2}{3} \Big) = \frac{2R^2}{5} \mu \cdot \frac{4\pi}{3} R^3 = \frac{2}{5} m R^2 \ . \qquad \triangleleft
\end{aligned}
$$

Wenn in Satz (25.3) die Funktion f nicht von den Winkeln abhängt, also rotationssymmetrisch ist, dann erhalten wir:

Satz 25.5 *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 25.3 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 $I = [0,1]$:

$$
\begin{aligned}
\kappa_n &= \int_{[0,1]} dr\ r^{n-1} \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta_1\ \sin\vartheta_1 \cdots \int_0^\pi d\vartheta_{n-2}\ \sin^{n-2}\vartheta_{n-2} \\
&= \frac{1}{n} \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta_1\ \sin\vartheta_1 \cdots \int_0^\pi d\vartheta_{n-2}\ \sin^{n-2}\vartheta_{n-2}\ . \qquad \square
\end{aligned}
$$

25.2 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 25.6 *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 25.7 (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:

$$\begin{aligned}
\int_{(\mathbb{R}^+)^2} d(x,y)\, x^{p-1}y^{q-1}e^{-x-y} &= \int_0^\infty du_1\; u_1^{p+q-1}e^{-u_1} \int_0^1 du_2\; (1-u_2)^{p-1}u_2^{q-1} \\
&= \Gamma(p+q) \int_0^1 du_2\; (1-u_2)^{p-1}u_2^{q-1} \\
\text{Fubini } &= \int_0^\infty dx\; x^{p-1}e^{-x} \int_0^\infty dy\; y^{q-1}e^{-y} = \Gamma(p)\Gamma(q)\;.
\end{aligned}$$

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 25.8 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. $\triangleleft$

Beispiel 25.9 Das Trägheitsmoment des Standardsimplex Δ^3 bei Rotation um die z-Achse ist

$$\begin{aligned}
\Theta &= \int_{\Delta^3} d(x,y,z)\; \mu \cdot (x^2+y^2) \\
&= \int_0^1 du_1\; u_1^2 \int_0^1 du_2 \int_0^1 du_3 (1-u_3)\; \mu \cdot \Big((u_1(1-u_2)(1-u_3))^2 + (u_1u_2(1-u_3))^2 \Big) \\
&= \mu \int_0^1 du_1\; u_1^4 \int_0^1 du_2\; (1-2u_2+2u_2^2) \int_0^1 du_3 (1-u_3)^3 \\
&= \frac{\mu}{5\cdot 4}\Big(1 - \frac{2}{2} + \frac{2}{3}\Big) = \frac{\mu}{30} = \frac{1}{5} m(\Delta^3)\;. \qquad \triangleleft
\end{aligned}$$

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

$$\Delta^{\alpha_1,\dots,\alpha_n}_{a_1,\dots,a_n} := \left\{(y_1,\dots,y_n) \in \mathbb{R}^n \ : \ y_i \geq 0 \ , \ \left(\frac{y_1}{a_1}\right)^{\alpha_1} + \dots \left(\frac{y_n}{a_n}\right)^{\alpha_n} \leq 1\right\}$$

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

$$|\det(DT)(x_1,\dots,x_n)| = \frac{a_1\cdots a_n}{\alpha_1\cdots\alpha_n} x_1^{\frac{1}{\alpha_1}-1}\cdots x_n^{\frac{1}{\alpha_n}-1} \ .$$

Satz 25.10 *Eine auf auf dem verallgemeinerten Simplex* ($\Delta^{\alpha_1,\dots,\alpha_n}_{a_1,\dots,a_n}$) *definierte Funktion f ist genau dann über dieses verallgemeinerte Simplex integrierbar, wenn die Funktion $f(a_1x_1^{\frac{1}{\alpha_1}},\dots,a_nx_n^{\frac{1}{\alpha_n}})x_1^{\frac{1}{\alpha_1}-1}\cdots x_n^{\frac{1}{\alpha_n}-1}$ über das Standardsimplex integrierbar ist. In diesem Fall gilt*

$$\int_{\Delta^{\alpha_1,\dots,\alpha_n}_{a_1,\dots,a_n}} dy\ f(y) = \frac{a_1\cdots a_n}{\alpha_1\cdots\alpha_n}\int_{\Delta^n} d(x_1,\dots,x_n)\ x_1^{\frac{1}{\alpha_1}-1}\cdots x_n^{\frac{1}{\alpha_n}-1} f(a_1x_1^{\frac{1}{\alpha_1}},\dots,a_nx_n^{\frac{1}{\alpha_n}}) \ .$$

Durch Kombination mit der Jacobi-Transformation entsteht so ein Diffeomorphismus $W^n \xrightarrow{J_n} \Delta^n \xrightarrow{T} \Delta^{\alpha_1,\dots,\alpha_n}_{a_1,\dots,a_n}$, mit dem wir Integrationen über ein verallgemeinertes Simplex auf Integrationen über den Würfel zurückführen können.

Beispiel 25.11 Wir berechnen das Volumen eines Ellipsoiden-Oktanten $EO := \Delta^{2,2,2}_{a,b,c}$ über die Jacobi-Transformation:

$$\begin{aligned}
v_3(EO) &= \int_{\Delta^{2,2,2}_{a,b,c}} dy \\
&= \frac{abc}{8}\int_{\Delta^3} d(x_1,x_2,x_3)\ x_1^{-\frac{1}{2}}x_2^{-\frac{1}{2}}x_3^{-\frac{1}{2}} \\
&= \frac{abc}{8}\int_0^1 du_1 u_1^2 \int_0^1 du_2 \int_0^1 du_3 (1-u_3)\cdot u_1^{-\frac{3}{2}}u_2^{-\frac{1}{2}}(1-u_2)^{-\frac{1}{2}}u_3^{-\frac{1}{2}}(1-u_3)^{-1} \\
&= \frac{abc}{8}\int_0^1 du_1\ u_1^{\frac{1}{2}}\int_0^1 du_2\ u_2^{-\frac{1}{2}}(1-u_2)^{-\frac{1}{2}}\int_0^1 du_3\ u_3^{-\frac{1}{2}} \\
&= \frac{abc}{8}\cdot\frac{2}{3}\cdot B(\tfrac{1}{2},\tfrac{1}{2})\cdot 2 = \frac{1}{8}\cdot\frac{4abc^3}{3}\frac{(\Gamma(\frac{1}{2}))^2}{\Gamma(1)} \ .
\end{aligned}$$

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 $E = \left\{(x,y,z)\in\mathbb{R}^3 \ : \ \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\leq 1\right\}$ hat das Volumen $v_3(E) = \frac{4\pi}{3}abc$. $\triangleleft$

26 Beweis des Transformationssatzes

Wir unterteilen den Beweis des Transformationssatzes (Satz 25.1) 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* $\|(x_1, \dots, x_i)\|_\infty := \max(x_1, \dots, x_n)$ 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 $a_1, \dots, a_n \in \mathbb{R}^n$ und $P(a_1, \dots, a_n) := \{x = t_1 a_1 + \cdots + t_n a_n \ : \ t_i \in [0, 1]\} \subset \mathbb{R}^n$ das durch diese Vektoren aufgespannte Parallelotop. Dann gilt*

$$v_n(P(a_1, \dots, a_n)) = |\det(a_1, \dots, a_n)| ,$$

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) $|\det(\dots, \lambda a_i, \dots)| = |\lambda| |\det(\dots, a_i, \dots)|$

(D2) $|\det(\dots, a_i, \dots, a_j, \dots)| = |\det(\dots, a_i + a_j, \dots, a_j, \dots)|$

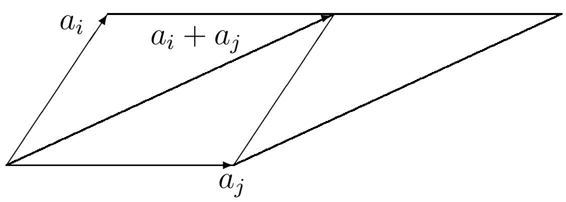
(D3) $|\det(e_1, \dots, e_n)| = 1$

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 $P_\lambda := P(a_1, \dots, a_{i-1}, \lambda a_i, a_{i+1}, \dots, a_n)$. 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 $\lambda = l \in \mathbb{N}^\times$ 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 $r_1 \le \lambda \le r_2$ mit $|r_1 - r_2| \le \frac{\epsilon}{v_n(P_1)}$. Das ergibt $v_n(P_{r_1}) \le v_n(P_\lambda) \le v_n(P_{r_2})$ und damit $|v_n(P_\lambda) - \lambda v_n(P_1)| \le \epsilon$. 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. $\square$

Sei nun $W = P(e_1, \dots, e_n) \subset \mathbb{R}^n$ 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 $T(W) = P(a_1, \dots, a_n)$. 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*
$v_n(T(W)) \le \max_{x \in W} |\det(DT)(x)| \cdot v_n(W)$.

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 $v_n(T(W_1)) \ge \alpha v_n(W_1)$. Durch Wiederholung dieser Zerlegung gewinnt man eine Folge $W_1 \supset W_2 \supset \dots$ von Würfeln mit $v_n(T(W_i)) \geq \alpha v_n(W_i)$. Nach Intervallschachtelungsprinzip (z.B. Satz 21.4) 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 $\displaystyle\lim_{x \to 0, x \neq 0} \frac{\phi(x)}{\|x\|_\infty} = 0$. 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 kompakte Mengen $A_0 \subset A_1 \subset \dots$ 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 $1, \frac{1}{2}, \dots, \frac{1}{2^k}$ 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* $\mathrm{supp}(\phi) := \overline{\{y \in \mathbb{R}^n \;:\; \phi(y) \neq 0\}}$ *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 21.14 aus dem 1. Semester), gibt es zu jedem $\epsilon > 0$ eine Zerlegung $Q = Q_1 \cup \dots \cup Q_p$ in kompakte Quader, die nur Randpunkte gemeinsam haben und so klein sind, daß $\max_{y\in Q_i} |\det(DT^{-1})(y)|^{-1} - \min_{y\in Q_i} |\det(DT^{-1})(y)|^{-1} \leq \epsilon$. Dann gilt im Urbild $T^{-1}(Q_i)$

$$\begin{aligned}\max_{x\in T^{-1}(Q_i)} |\det(DT)(x)| v_n(T^{-1}(Q_i)) - \min_{x\in T^{-1}(Q_i)} |\det(DT)(x)| v_n(T^{-1}(Q_i))&\\ \leq \epsilon v_n(T^{-1}(Q_i)) \;.&\end{aligned}$$

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

$$\Big|\int_{T^{-1}(Q_i)} dx\; |\det(DT)(x)| - v_n(Q_i)\Big| \leq \epsilon v_n(T^{-1}(Q_i)) \;.$$

Summation über alle Teilquader liefert

$$\begin{aligned}\Big|\int_{T^{-1}(Q)} dx\; |\det(DT)(x)| - v_n(Q)\Big| &\leq \sum_{i=1}^p \Big|\int_{T^{-1}(Q_i)} dx\; |\det(DT)(x)| - v_n(Q_i)\Big|\\ &\leq \epsilon v_n(T^{-1}(Q)) \;.\end{aligned}$$

Für $\epsilon \to 0$ ergibt sich die Behauptung. $\square$

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 $\|f - \phi_\epsilon\|_1 \leq \frac{\epsilon}{2}$. Wegen $|f_V - \phi_\epsilon \mathbf{1}_V| \leq |f_V - \phi_\epsilon|$ gilt dann auch $\|f_V - \phi_\epsilon \mathbf{1}_V\|_1 \leq \frac{\epsilon}{2}$.

Sei $B \subset \mathbb{R}^n$ eine beschränkte offene Teilmenge mit $\mathrm{supp}(\phi_\epsilon) \subset B$ und $M = \max_{x \in \mathrm{supp}(\phi_\epsilon)} |f_\epsilon(x)|$. Dann gibt es zu der beschränkten offenen Teilmenge $V \cap B$ eine Vereinigung $A = Q_0 \cup \dots \cup Q_k \subset V \cap B$ 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 $\mathrm{supp}(\phi_\epsilon \mathbf{1}_A) \subset V$, für die gilt

$$\|\phi_\epsilon \mathbf{1}_A - \phi_\epsilon \mathbf{1}_V\|_1 = \|\phi_\epsilon \mathbf{1}_A - \phi_\epsilon \mathbf{1}_{V \cap B}\|_1 \leq M|v_n(A) - v_n(V \cap B)| < \frac{\epsilon}{2} .$$

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 Auwahl einer Teilfolge gemäß Satz 22.3 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. $\square$

27 Der Satz von Fubini

Es geht nun um die Verallgemeinerung des Satzes von Fubini für stetigen Funktionen (bisher nicht bewiesen) 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 27.1 (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* $F(y) := \begin{cases} \displaystyle\int_X dx\ f(x,y) & \textit{für } y \in Y \setminus N \\ 0 & \textit{für } y \in N \end{cases}$
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 22.3 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 $\sum_{i=0}^{\infty} c_i \mathbf{1}_{Q_i}$ mit $\sum_{i=0}^{\infty} c_i v(Q_i) < \epsilon$. 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) $\mathbf{1}_{A_y}(x) \leq \sum_{i=0}^{\infty} c_i \mathbf{1}_{Q_i'}(x) \mathbf{1}_{Q_i''}(y)$ und damit $a(y) := \displaystyle\int_X dx \mathbf{1}_{A_y}(x) \leq \sum_{i=0}^{\infty} c_i v(Q_i') \mathbf{1}_{Q_i''}(y)$. Integration der so definierten Funktion a über Y liefert $\|a\|_{1,Y} \leq \epsilon$ für alle $\epsilon > 0$, somit $\|a\|_{1,Y} = 0$. Damit existiert eine 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 f_y.

Sei $H_k(y) := \displaystyle\int_X dx\ \ |\phi_{k+1,y}(x) - \phi_{k,y}(x)| = \|\phi_{k+1,y} - \phi_{k,y}\|_{1,X}$. Nach dem Satz von Fubini für Treppenfunktionen (Satz 20.3) gilt $\displaystyle\int_Y dy\ \ H_k(y) = \int_{X\times Y} d(x,y)\ |\phi_{k+1}(x,y) - \phi_k(x,y)| = \|\phi_{k+1} - \phi_k\|_1$ und damit

$$\sum_{k=0}^{\infty} \int_Y dy\ H_k(y) < \infty\ . \tag{*}$$

Die Folge $(G_s)_{s\in\mathbb{N}}$ der Funktionen $G_s = \sum_{k=0}^{s} H_s$ ist monoton wachsend. Nach (*) gilt

$$I_s := \int_Y dy\ G_s(y) = \sum_{k=0}^{s} \int_Y dy\ H_k(y) \leq \sum_{k=0}^{\infty} \int_Y dy\ H_k(y) < \infty\ ,$$

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 21.10 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 $N = N' \cup N''$,

$$\sum_{k=0}^{\infty} \|\phi_{k+1,y} - \phi_{k,y}\|_{1,X} < \infty \ . \tag{**}$$

Nach dem Cauchy-Kriterium für Reihen gibt es zu $\epsilon > 0$ ein $N \in \mathbb{N}$, so daß $\sum_{k=l}^{l'} \|\phi_{k+1,y} - \phi_{k,y}\|_{1,X} < \epsilon$ für alle $l' > l \geq N$. Insbesondere 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 f_y integrierbar, und für $y \in Y \setminus N$ gilt

$$\lim_{k\to\infty} \int_X dx\ \phi_k(x,y) = \int_X dx\ f(x,y) = F(y) \ .$$

Somit ist i) bewiesen.

Wir betrachten nun die Treppenfunktionen $\Phi_k(y) := \int_X dx\ \phi_{k,y}(x)$. Die Folge $(\Phi_k)_{k\in\mathbb{N}}$ konvergiert fast überall punktweise gegen F. Außerdem ist

$$\begin{aligned}
\sum_{k=0}^{\infty} \|\Phi_{k+1} - \Phi_k\|_{1,Y} &= \sum_{k=0}^{\infty} \int_Y dy\ \Big| \int_X dx\ (\phi_{k+1,y}(x) - \phi_{k,y}(x)) \Big| \\
&\leq \sum_{k=0}^{\infty} \int_Y dy \int_X dx\ \big|\phi_{k+1,y}(x) - \phi_{k,y}(x)\big| = \sum_{k=0}^{\infty} \|\phi_{k+1} - \phi_k\|_1 < \infty \ .
\end{aligned}$$

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

$$\int_{X\times Y} d(x,y)\ f(x,y) = \int_X dx \int_Y dy\ f(x,y) = \int_Y dy \int_X dx\ f(x,y) \ .$$

Eine wichtige Anwendung des Satzes von Fubini ist die Faltung:

Satz 27.2 *Seien $f, g : \mathbb{R}^n \to \mathbb{C}$ integrierbar, dann existiert für fast alle $x \in \mathbb{R}^n$ die Faltung $f * g : \mathbb{R}^n \to \mathbb{C}$ definiert durch*

$$(f * g)(x) := \int_{\mathbb{R}^n} dy\ f(x-y)g(y)\ .$$

Es gilt

i) *$f * g$ ist integrierbar mit* $\displaystyle\int_{\mathbb{R}^n} dx\ (f * g)(x) = \int_{\mathbb{R}^n} dx\ f(x) \int_{\mathbb{R}^n} dy\ g(y)$

ii) $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$

iii) $f * g = g * f$

Beweis. Die Funktion $(x, y) \mapsto f(x)g(y)$ ist über $\mathbb{R}^{2n}$ integrierbar. Die Abbildung $T : \mathbb{R}^{2n} \ni (x, y) \mapsto (x-y, y) \in \mathbb{R}^{2n}$ ist ein Diffeomorphismus mit $\det(DT) = 1$ Nach Transformationssatz ist $(x, y) \mapsto f(x-y)g(y)$ integrierbar mit

$$\int_{\mathbb{R}^{2n}} d(x,y)\ f(x)g(y) = \int_{\mathbb{R}^{2n}} d(x,y)\ f(x-y)g(y)\ .$$

Nach Satz 27.1.i) existiert für fast alle x das Integral $(f*g)(x)$, nach Satz 27.1.ii) ist diese Faltung integrierbar und erfüllt i).

Für festes x gilt $|(f * g)(x)| \leq \int_{\mathbb{R}^n} dy\ |f(x-y)|\,|g(y)|$ außerhalb einer Nullmenge und damit

$$\|f * g\|_1 = \int_{\mathbb{R}^n} dx\ |(f * g)(x)| \leq \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy\ |f(x-y)|\,|g(y)| = \|f\|_1\,\|g\|_1$$

nach Anwendung der Transformation $T(x, y) = (x + y, y)$. Somit ist ii) gezeigt.

Für festes x ist $z = T(y) = x - y$ ein Diffeomorphismus mit $|\det DT| = 1$. Nach Transformationssatz gilt $\displaystyle\int_{\mathbb{R}^n} dy\ f(x-T(y))g(T(y)) = \int_{\mathbb{R}^n} dz\ f(x-z)g(z).$ □

Die Voraussetzung der Integrierbarkeit über $X \times Y$ im Satz von Fubini ist entscheidend. Es gibt Beispiele für Funktionen, für die die Integrale $\displaystyle\int_X dx \int_Y dy\ \ f(x,y)$ und $\displaystyle\int_Y dy \int_X dx\ \ f(x,y)$ existieren, ohne daß $f(x, y)$ integrierbar ist. Die Umkehrung ist der Satz von Tonelli. Dazu benötigen wir:

Definition 27.3 Eine Teilmenge $A \subset \mathbb{R}^n$ heißt σ-*kompakt*, falls sie Vereinigung abzählbar vieler kompakter Mengen ist.

Eine Funktion $f : A \to \mathbb{R} \cup \{\infty\}$ auf einer σ-kompakten Menge $A \subset \mathbb{R}^n$ heißt *lokal-integrierbar*, wenn sie über jede kompakte Teilmenge $K \subset A$ integrierbar ist.

Jede kompakte und jede offene Menge ist σ-kompakt.

Satz 27.4 (Tonelli) *Eine lokal-integrierbare 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*

$$\int_{X\times Y} d(x,y)\; f(x,y) = \int_X dx \int_Y dy\; f(x,y) = \int_Y dy \int_X dx\; f(x,y)\;.$$

($\Leftarrow$) Mit f ist auch $|f|$ über $X \times Y$ integrierbar, und der Satz von Fubini liefert die Existenz und Eigenschaften der iterierten Integrale.

($\Rightarrow$) Sei $X \times Y = \bigcup_{k=0}^{\infty} A_k$, mit A_k kompakt, sowie $B_l := \bigcup_{k=0}^{l} A_k$ und $f_l := f\mathbf{1}_{B_l}$. Die Folge $\{|f_l|\}$ konvergiert punktweise monoton wachsend gegen $|f|$. Mit f_l ist auch $|f_l|$ integrierbar, somit gilt für $|f_l|$ der Satz von Fubini,

$$\int_{X\times Y} d(x,y)\; |f_l(x,y)| = \int_Y dy \int_X dx |f_l(x,y)| \leq \int_Y dy \int_X dx |f(x,y)|\;.$$

Nach Voraussetzung ist das Integral beschränkt, so daß nach dem Satz von Beppo Levi die Grenzfunktion $|f|$ integierbar ist. Schließlich folgt die Integrierbarkeit von f aus dem Satz von Lebesgue. □

Satz 27.5 (Majorantenkriterium) *Es sei $A \subset \mathbb{R}^n$ eine σ-kompakte Menge, $f : A \to \mathbb{C} \cup \{\infty\}$ eine lokal-integrierbare Funktion, für die es eine integrierbarer Majorante $F : A \to \mathbb{R}$ gibt mit $|f| \leq F$. Dann ist auch f integrierbar.*

Beweis. Es sei $A = \bigcup_{l\in\mathbb{N}} A_l = A$ mit A_l kompakt und $B_k := \bigcup_{l=0}^{l} A_l$. Dann ist $f_k := f\mathbf{1}_{B_k}$ integrierbar mit $\lim_{k\to\infty} f_k = f$ und $|f_k| \leq |f| \leq F$. Nach dem Satz von Lebesgue ist f integrierbar. □

28 L^p-Räume und Fourier-Transformation

Satz 28.1 ($\mathcal{L}^p$-Räume) *Es sei $A \subset \mathbb{R}^n$ eine σ-kompakte Teilmenge. Für $1 < p < \infty$ sei*

$$\begin{aligned}\mathcal{L}^p(A) &:= \{f : A \to \mathbb{C} \cup \{\infty\}\;:\\ &\qquad\qquad f \text{ ist lokal-integrierbar und } |f|^p \text{ ist über } A \text{ integrierbar}\}\;,\\ \mathcal{L}^\infty(A) &:= \{f : A \to \mathbb{C} \cup \{\infty\}\;:\\ &\qquad\qquad f \text{ ist lokal-integrierbar und fast überall beschränkt}\}\;.\end{aligned}$$

Dann ist $\mathcal{L}^p$ für $1 \leq p \leq \infty$ ein Vektorraum.

Beweis. Klar für $p = \infty$, ansonsten sind für $f, g \in \mathcal{L}^p(A)$ auch $f + g$ und $|f + g|^p$ lokal-integrierbar. Es gilt

$$|f + g|^p \leq (|f| + |g|)^p \leq 2^p \max(|f|^p, |g|^p)\;.$$ □

Aus dem Majorantenkriterium folgt, daß jede lokal-integrierbare Funktion, für die $|f|$ integrierbar ist, auch selbst integrierbar ist.

Satz 28.2 *Für* $f \in \mathcal{L}^p(A)$ *mit* $1 \leq p < \infty$ *sei* $\|f\|_p := \Big(\displaystyle\int_A dx\; |f(x)|^p\Big)^{\frac{1}{p}}$. *Für* $f \in \mathcal{L}^\infty(A)$ *sei*

$$\|f\|_\infty := \inf\{M \in \mathbb{R} \;:\; \text{außerhalb einer Nullmenge gilt } |f(x)| < M\}\;.$$

Dann ist $L^p(A) := \mathcal{L}^p(A)/\mathcal{N}^p(A)$ *mit* $\mathcal{N}^p(A) := \big\{f \in \mathcal{L}^p(A) \;:\; \|f\|_p = 0\big\}$ *ein Banach-Raum bezüglich* $\|\;\|_p$*, d.h. ein vollständiger normierter Vektorraum. Insbesondere ist* $L^2(A)$ *ein Hilbert-Raum mit Skalarprodukt*

$$\langle f,g\rangle := \int_A dx\; \overline{f(x)}g(x)\,, \qquad f,g \in L^2(A)\;.$$

Offenbar ist $\|f\|_p = 0$ genau dann, wenn $f = 0$ fast überall, d.h. $L^p(A)$ besteht wieder aus den Äquivalenzklassen fast überall gleicher $\mathcal{L}^p$-Funktionen.

Beweis. Zu zeigen ist die Dreiecksungleichung in $\mathcal{L}^p(A)$ und die Vollständigkeit. Wie im Fall der Folgenräume $\ell^p(\mathbb{N})$ benötigen wir zunächst:

Satz 28.3 (Höldersche Ungleichung) *Es seien* $1 \leq p,q \leq \infty$ *mit* $\frac{1}{p}+\frac{1}{q} = 1$ *sowie* $f \in \mathcal{L}^p(A)$ *und* $g \in \mathcal{L}^q(A)$*. Dann ist* $fg \in \mathcal{L}^1(A)$*, und es gilt die Höldersche Ungleichung*

$$\|fg\|_1 \leq \|f\|_p\|g\|_q\;.$$

Beweis. Für $p = 1$, $q = \infty$ ist außerhalb einer Nullmenge $|g| < M$. Dann ist auch fg lokal-integrierbar, und $M|f|$ ist integrierbare Majorante für fg.

Für $1 < p,q < \infty$ können wir $\|f\|_p, \|g\|_q \neq 0$ annehmen; sonst wäre $f = 0$ oder $g = 0$ fast überall, somit nichts zu zeigen. In Analogie zum Beweis der Hölderschen Ungleichung für Summen in Satz 24.7 aus dem 1. Semester gilt (außerhalb möglicher Nullmengen N) in jedem Punkt $x \in A \setminus N$

$$\frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_q} \leq \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|g(x)|^q}{\|g\|_q^q}$$

Nach Integration über A folgt $\|fg\|_1 = \displaystyle\int_A dx|fg|(x) \leq \|f\|_p\|g\|_q < \infty$. Außerhalb von Nullmengen sind f,g als lokal-integrierbare Funktionen beschränkt, somit ist auch fg lokal-integrierbar. Aus dem Majorantenkriterium folgt dann $fg \in \mathcal{L}^1(A)$. □

Der Beweis der *Minkowskischen Ungleichung* $\|f+g\|_p \leq \|f\|_p+\|g\|_p$ für $f,g \in \mathcal{L}^p$ ist dann analog zum Beweis von Satz 24.8 aus dem 1. Semester.

(Vollständigkeit für $1 < p < \infty$) Es sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge von Funktionen $f_n \in \mathcal{L}^p(A)$. Wegen $|\|f_n\|_p - \|f_m\|_p| \leq \|f_n - f_m\|_p$ ist dann auch $(\|f_n\|_p)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Insbesondere ist $\|f_n\|_p \leq M$ beschränkt für alle n, und damit ist $|f_n(x)| < \infty$ fast überall.

Es sei $(f_{n_k})_{k=1,2,\dots}$ eine Teilfolge mit $n_1 < n_2 < \dots$ und $\|f_n - f_{n_k}\|_p < \frac{1}{2^k}$ für alle $n \geq n_k$. Dann gilt $\sum_{k=1}^\infty \|f_{n_k} - f_{n_{k+1}}\|_p \leq 1$. Sei $g_k(x) := f_{n_k}(x) - f_{n_{k+1}}(x)$ und $G_n(x) := \sum_{k=1}^n |g_k(x)|$ sowie $G(x) := \sum_{k=1}^\infty |g_k(x)|$. Da $G_n^p \in \mathcal{L}^1(A)$, gilt

$$\int_A dx\; (G_n(x))^p = \|G_n\|_p^p \leq \Big(\sum_{k=1}^n \|g_k\|_p\Big)^p \leq 1\;.$$

Nach dem Satz von Beppo Levi ist die punktweise gebildete Grenzfunktion $H(x) := \lim_{n\to\infty}(G_n(x))^p$ integrierbar, d.h. $\|H\|_1 = \|G\|_p^p < \infty$. Insbesondere ist $G(x)$ fast überall endlich. Also existiert fast überall punktweise der Limes $g(x) := \sum_{k=1}^\infty g_k(x)$ mit absoluter Konvergenz der Reihe.

Wir zeigen $g \in \mathcal{L}^p(A)$. Die Folge der Partialsummen $s_n := \Big|\sum_{k=1}^n g_k\Big|^p$ konvergiert fast überall punktweise gegen $|g|^p$ mit $|s_n| \leq |g|^p$ fast überall. Wegen

$$\int_A dx \Big|\sum_{k=1}^n g_k(x)\Big|^p = \Big(\Big\|\sum_{k=1}^n g_k\Big\|_p\Big)^p \leq \Big(\sum_{k=1}^n \|g_k\|_p\Big)^p \leq 1$$

sind die Integrale der Partialsummen s_n beschränkt. Nach dem Satz von Lebesgue ist dann $|g|^p$ integrierbar, also $g \in \mathcal{L}^p(A)$. Damit ist $f := f_{n_1} - g = \lim_{k\to\infty} f_{n_k} \in \mathcal{L}^p(A)$.

Die Folge $|f - f_{n_k}|^p$ konvergiert für $k \to \infty$ fast überall gegen Null. Außerdem ist nach Abänderung auf einer möglichen Nullmenge $|f - f_{n_k}|^p \leq |g|^p$ für alle k, und $|g|^p$ ist integrierbar. Nach dem Satz von Lebesgue gilt dann

$$\lim_{k\to\infty} \int_A dx |f(x) - f_{n_k}(x)|^p = \int_A dx \Big(\lim_{k\to\infty} |f(x) - f_{n_k}(x)|^p\Big) = 0\;,$$

d.h. (f_n) konvergiert in der L^p-Norm gegen die punktweise gebildete Grenzfunktion $f \in \mathcal{L}^p(A)$.

(Vollständigkeit für $p = \infty$) Wie zuvor folgt $\|f_n\|_\infty \leq M$. Die abzählbare Vereinigung von Nullmengen ist Nullmenge, also gibt es eine gemeinsame Nullmenge N, so daß $|f_n(x)| < M$ für alle $x \in A \setminus N$. Aus $\|f_n - f_k\|_\infty < \epsilon$ folgt dann $|f_n(x) - f_k(x)| < \epsilon$, somit aus der Vollständigkeit von $\mathbb{C}$ die Existenz der Grenzfunktion $f = \lim_{n\to\infty} f_n \in \mathcal{L}^\infty(A)$. □

Für $k \in \mathbb{R}^n$ sei $e_k \in \mathcal{L}^\infty(\mathbb{R}^n)$ definiert durch $e_k(x) := e^{-i\langle k,x\rangle}$. Es gilt $\|e_k\|_\infty = 1$. Ist $f \in \mathcal{L}^1(\mathbb{R}^n)$, dann gilt nach der Hölderschen Ungleichung $\|fe_k\|_1 \leq \|f\|_1\|e_k\|_\infty = \|f\|_1 < \infty$. Also existiert das Integral

$$\hat{f}(k) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dx\; f(x)\, e^{-i\langle k,x\rangle}\;, \qquad f \in \mathcal{L}^1(\mathbb{R}^n)\;,$$

mit $\sup_{k\in\mathbb{R}^n} |\hat{f}(k)| \leq \frac{1}{(2\pi)^{\frac{n}{2}}}\|f\|_1 < \infty$.

Satz 28.4 *Für $f \in \mathcal{L}^1(\mathbb{R}^n)$ ist die Abbildung $\mathbb{R}^n \ni k \mapsto \hat{f}(k)$ stetig.*

Beweis. Sei (k_l) konvergent gegen k. Dann konvergiert $f_l(x) := f(x)e^{-\mathrm{i}\langle k_l,x\rangle}$ punktweise (bezüglich x) gegen $f(x)e^{-\mathrm{i}\langle k_l,x\rangle}$. Da $|f_l(x)| \leq |f(x)|$ durch eine integrierbare Funktion beschränkt ist, vertauschen nach dem Satz von Lebesgue Grenzwert und Integral, d.h. das Integral ist stetig in k. □

Definition 28.5 Die Abbildung $\mathcal{F} : \mathcal{L}^1(\mathbb{R}^n) \to \mathcal{C}_b(\mathbb{R}^n)$ mit $(\mathcal{F}(f))(k) := \hat{f}(k)$ heißt die *(kontinuierliche) Fourier-Transformation*.

Beispiel 28.6 Für $x \in \mathbb{R}$ sei $f(x) = e^{-\frac{\lambda^2}{2}x^2}$ mit $\lambda > 0$. Dann gilt

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dx\; e^{-\frac{\lambda^2}{2}x^2 + \mathrm{i}kx} = \frac{1}{\sqrt{2\pi}} e^{-\frac{k^2}{2\lambda^2}} \int_{\mathbb{R}} dx\; e^{-\frac{\lambda^2}{2}(x + \frac{\mathrm{i}k}{\lambda^2})^2} .$$

Das verbleibende Integral kann mit der Cauchyschen Integralformel für $e^{-\frac{\lambda^2}{2}z^2}$ berechnet werden zu $\int_{\mathbb{R}} dx\; e^{-\frac{\lambda^2}{2}(x+\frac{\mathrm{i}k}{\lambda^2})^2} = \sqrt{\frac{2\pi}{\lambda^2}}$. Somit ist $\hat{f}(k) = \frac{1}{\lambda}e^{-\frac{k^2}{2\lambda^2}}$. ◁

Satz 28.7 (Faltungssatz) *Seien f, g integrierbar, dann gilt $\widehat{(f * g)}(k) = (2\pi)^{\frac{n}{2}}\hat{f}(k)\hat{g}(k)$.*

Beweis. Die Definitionen ergeben

$$\begin{aligned}\widehat{(f * g)}(k) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dx\; (f \star g)(x)e^{-\mathrm{i}\langle k,x\rangle} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy\; f(x-y)e^{-\mathrm{i}\langle k,x-y\rangle} g(y)e^{-\mathrm{i}\langle k,y\rangle} .\end{aligned}$$

Die Transformation $T(x, y) = (x + y, y)$ mit $\det DT = 1$ liefert die Behauptung. □

Im allgemeinen muß eine $\mathcal{C}_b$-Funktion nicht wieder integrierbar sein.

Satz 28.8 (Umkehrformel) *Sei $f \in \mathcal{L}^1(\mathbb{R}^n)$ derart, daß $\hat{f} \in \mathcal{L}^1(\mathbb{R}^n)$. Dann gilt*

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dk\; \hat{f}(k)e^{\mathrm{i}\langle k,x\rangle} \qquad \text{für fast alle } x \in \mathbb{R}^n .$$

Zu beachten ist das andere Vorzeichen in $e^{\mathrm{i}\langle k,x\rangle}$ gegenüber der Fourier-Transformation!

Beweis. Für $\lambda > 0$ sei $F_\lambda(x) := \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dk\; \hat{f}(k)e^{\mathrm{i}\langle k,x\rangle}e^{-\frac{\lambda^2}{2}\langle k,k\rangle}$. Da $\hat{f}$ beschränkt ist, existiert das Integral für $\lambda > 0$. Dann gilt mit dem Transformationssatz

$$\begin{aligned} F_\lambda(x) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dk \Big(\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dy\; f(y) e^{-\mathrm{i}\langle k,y\rangle} \Big) e^{\mathrm{i}\langle k,x\rangle} e^{-\frac{\lambda^2}{2}\langle k,k\rangle} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dk\; \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dy\; f(y+x) e^{-\mathrm{i}\langle k,y\rangle} e^{-\frac{\lambda^2}{2}\langle k,k\rangle} . \end{aligned}$$

Die Funktion $(y,k) \mapsto f(y)e^{-\mathrm{i}\langle k,y\rangle}e^{-\frac{\lambda^2}{2}\langle k,k\rangle}$ ist über $\mathbb{R}^{2n}$ integrierbar, ebenso $(y,k) \mapsto f(x+y)e^{-\mathrm{i}\langle k,y\rangle}e^{-\frac{\lambda^2}{2}\langle k,k\rangle}$ nach Transformationssatz, so daß wir nach Fubini die Integrale vertauschen dürfen. Mit Beispiel 28.6 (mit $x \mapsto -k$ und $k \mapsto y$) und Fubini gilt $\displaystyle\int_{\mathbb{R}^n} dx\; e^{-\frac{\lambda^2}{2}\langle k,k\rangle} e^{\mathrm{i}\langle y,k\rangle} = \frac{(2\pi)^{\frac{n}{2}}}{\lambda^n} e^{-\frac{\langle y,y\rangle}{2\lambda^2}}$, also

$$F_\lambda(x) = \int_{\mathbb{R}^n} dy\; f(y+x)\delta_\lambda(y) = (f * \delta_\lambda)(x) , \qquad \delta_\lambda(y) := \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{\lambda^n} e^{-\frac{\langle y,y\rangle}{2\lambda^2}} .$$

Wir werden zeigen, daß für $\lambda \to 0$ die rechte Seite in der L^1-Norm gegen $f(x)$ konvergiert. In diesem Sinn geht $\delta_\lambda(y)$ für $\lambda \to 0$ in die sogenannte *Diracsche δ-Distribution* über. Andererseits konvergiert $\hat{f}(k)e^{\mathrm{i}\langle k,x\rangle}e^{-\frac{\lambda^2}{2}\langle k,k\rangle}$ für $\lambda \to 0$ punktweise gegen $\hat{f}(k)e^{\mathrm{i}\langle k,x\rangle}$, und $|\hat{f}(k)e^{\mathrm{i}\langle k,x\rangle}e^{-\frac{\lambda^2}{2}\langle k,k\rangle}|$ ist beschränkt durch die integrierbare Funktion $|\hat{f}(k)|$. Nach dem Satz von Lebesgue gilt dann $\displaystyle\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dk\; \hat{f}(k)e^{\mathrm{i}\langle k,x\rangle} = \lim_{\lambda\to 0} F_\lambda(x)$. Beides zusammen liefert die Behauptung.

Ausgangspunkt ist $\displaystyle\frac{1}{(2\pi\lambda^2)^{\frac{n}{2}}} \int_{\mathbb{R}} dx\; e^{-\langle y,y\rangle/(2\lambda^2)} = 1$. Damit gilt

$$f(x) - F_\lambda(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dy\; (f(x) - f(x+y)) \frac{e^{-\langle y,y\rangle/(2\lambda^2)}}{\lambda^n} . \qquad (*)$$

Zunächst sei f stetig mit kompaktem Träger $\mathrm{supp}(f) \subset \overline{K_R(0)}$. Dann ist f auch gleichmäßig stetig, d.h. zu $\epsilon > 0$ gibt es ein $1 > \delta > 0$, so daß $|f(x) - f(x+y)| < \frac{\epsilon}{2v_n(\overline{K_{R+1}(0)})}$ für alle $x \in \mathbb{R}^n$ und $y \in K_\delta(0)$. Das y-Integral wird zerlegt

- in ein Integral über $K_\delta(0)$: Dort ist $\mathrm{supp}_x(f(x) - f(x+y)) \subset \overline{K_{R+1}(0)}$, so daß das x-Integral durch $|f(x) - f(x+y)|$ und das Volumen von $\overline{K_{R+1}(0)}$ abgeschätzt werden kann. Das y-Integral wird anschließend auf $\mathbb{R}^n$ ausgedehnt und wird unabhängig von λ.
- in ein Integral über $\mathbb{R}^n \setminus K_\delta(0)$. Dann ist $0 < e^{-\langle y,y\rangle/(2\lambda^2)} \leq e^{-\frac{\delta^2}{4\lambda^2}} e^{-\langle y,y\rangle/(4\lambda^2)}$.

Integration des Betrages von (*) über $x \in \mathbb{R}^n$ liefert dann $\|f - F_\lambda\|_1 = I_1 + I_2$ mit

$$
\begin{aligned}
I_1 &= \int_{\overline{K_{R+1}(0)}} dx \int_{K_\delta(0)} dy\, |f(x) - f(x+y)| \frac{e^{-\langle y,y\rangle/(2\lambda^2)}}{(2\pi\lambda^2)^{\frac{n}{2}}} \\
&\leq v_n(K_{R+1}(0)) \sup_{x,x+y \in \overline{K_{R+1}(0)}} |f(x) - f(x+y)| \int_{\mathbb{R}^n} dy \frac{e^{-\langle y,y\rangle/(2\lambda^2)}}{(2\pi\lambda^2)^{\frac{n}{2}}} < \frac{\epsilon}{2}\,, \\
I_2 &= \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n \setminus K_\delta(0)} dy\, |f(x) - f(x+y)| \frac{e^{-\langle y,y\rangle/(2\lambda^2)}}{(2\pi\lambda^2)^{\frac{n}{2}}} \\
&\leq e^{-\frac{\delta^2}{4\lambda^2}} \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy\, (|f(x)| + |f(x+y)|) \frac{e^{-\langle y,y\rangle/(4\lambda^2)}}{(2\pi\lambda^2)^{\frac{n}{2}}} \\
&\leq e^{-\frac{\delta^2}{4\lambda^2}} 2\|f\|_1 2^{\frac{n}{2}}\,.
\end{aligned}
$$

Wir können nun $\lambda(\delta)$ so klein wählen, daß $e^{-\frac{\delta^2}{4\lambda^2}} 2\|f\|_1 2^{\frac{n}{2}} < \frac{\epsilon}{2}$. Also gibt es zu jedem $\epsilon > 0$ ein $\lambda(\epsilon) > 0$, so daß $\|f - F_{\lambda(\epsilon)}\|_1 < \epsilon$ für stetige Funktionen f mit kompaktem Träger, d.h. es gilt fast überall $f(x) = \lim_{\lambda \to 0} F_\lambda(x)$. Da integrierbare Funktionen bezüglich $\| \ \|_1$ durch stetige Funktionen mit kompaktem Träger approximiert werden können, gilt die Aussage für alle $f \in \mathcal{L}^1(\mathbb{R}^n)$ mit $\hat{f} \in \mathcal{L}^1(\mathbb{R}^n)$. □

Satz 28.9 *Sei $k \in \mathbb{N}$ und $f \in \mathcal{L}^1(\mathbb{R}^n)$ derart, daß für alle Multi-Indizes α mit $|\alpha| < k$ sogar $x^\alpha f(x)$ integrierbar ist. Dann existieren die mehrfachen partiellen Ableitungen $\partial_\alpha \hat{f}$ der Fourier-Transformierten, und es gilt $\widehat{(x^\alpha f)}(k) = \dfrac{\mathrm{i}^{|\alpha|}\partial^{|\alpha|}}{\partial k^\alpha} \hat{f}(k)$.*

Beweis. Es genügt, den Fall einer partiellen Ableitung in Richtung e_i zu beweisen. Sei $\epsilon_l = \frac{1}{l}$. Dann ist

$$
\frac{\hat{f}_\alpha(k + \epsilon_l e_i) - \hat{f}_\alpha(k)}{\epsilon_l} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} dx\, f(x) e^{-\mathrm{i}\langle k,x\rangle} \frac{(e^{-\mathrm{i}x_i\epsilon_l} - 1)}{\epsilon_l}
$$

Die Funktion $\frac{(e^{-\mathrm{i}x_i\epsilon_l}-1)}{\epsilon_l}$ ist nach dem Schrankensatz beschränkt durch $|x_i|$, somit vertauschen nach dem Satz von Lebesgue Integral und Grenzwert $l \to \infty$. □

Aus der Umkehrformel ergibt sich:

Satz 28.10 *Sei $f \in \mathcal{C}^k(\mathbb{R}^n)$ derart, daß f und $\partial^\alpha f$ für alle Multi-Indizes mit $|\alpha| \leq k$ integrierbar sind. Dann gilt $\widehat{(\partial^\alpha f)}(k) = (\mathrm{i}k)^\alpha \hat{f}(k)$.* □

Daraus gewinnt man eine oft nützliche Lösungsmethode für lineare (sogar partielle) Differentialgleichungen:

Beispiel 28.11 (Wärmeleitungsgleichung) Der Ausgleich von Temperaturdifferenzen in homogenen Medien wird durch die Wärmeleitungsgleichung

$$\partial_t T(t,x) = c(\Delta T)(t,x)$$

beschrieben. Dabei ist Δ der Laplace-Operator, $T(t,x)$ die Temperatur zur Zeit $t \in \mathbb{R}_+$ am Ort $x \in \mathbb{R}^n$ und c die Wärmeleitfähigkeit. Fourier-Transformation in x führt mit Satz 28.10 auf die gewöhnliche Differentialgleichung $\frac{d}{dt}\hat{T}(t,k) = -c\|k\|^2\hat{T}(t,k)$ mit Lösung $\hat{T}(t,k) = \hat{T}(0,k)e^{-c\|k\|^2 t}$. Nach Beispiel 28.6 gilt $e^{-c\|k\|^2 t} = \hat{f}(k)$ für $f(x) = \frac{1}{(2ct)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4ct}}$. Aus dem Faltungssatz folgt dann $\hat{T}(t,k) = \frac{1}{(2\pi)^{\frac{n}{2}}}(\widehat{T(0,\,.\,)*f})(k)$, so daß die Rücktransformation auf folgende Lösung der Wärmeleitungsgleichung führt:

$$T(t,x) = \frac{1}{(2\pi)^{\frac{n}{2}}}(T(0,\,.\,)*f)(x) = \frac{1}{(4\pi ct)^{\frac{n}{2}}}\int_{\mathbb{R}^n} dy\; T(0,y)e^{-\frac{\|x-y\|^2}{4ct}}\;. \qquad \triangleleft$$

Eine weitere Folgerung der Umkehrformel ist:

Satz 28.12 (Plancherel) *Sei $f \in \mathcal{L}^2(\mathbb{R}^n) \cap \mathcal{L}^1(\mathbb{R}^n)$, dann ist $\mathcal{F}(f) \in \mathcal{L}^2(\mathbb{R}^n)$, und es gilt $\|f\|_2 = \|\mathcal{F}(f)\|_2$. Insbesondere folgt $\langle f,g\rangle_2 = \langle \hat{f},\hat{g}\rangle_2$.*

Beweis. Sei f zunächst so, daß $\hat{f}$ integrierbar ist und bezeichne $g := \overline{\hat{f}}$. Dann ist $(x,k) \mapsto f(x)g(k)e^{-\mathrm{i}\langle k,x\rangle}$ über $\mathbb{R}^{2n}$ integrierbar, und nach Fubini und Umkehrformel gilt

$$\begin{aligned}
&\frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^{2n}} d(x,k)\; f(x)g(k)e^{-\mathrm{i}\langle k,x\rangle}\\
&\qquad = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} dx\; f(x)\overline{\int_{\mathbb{R}^n} dk\; \hat{f}(k)e^{\mathrm{i}\langle k,x\rangle}} = \int_{\mathbb{R}^n} dx\; \big|f(x)\big|^2\\
&\qquad = \frac{1}{(2\pi)^{\frac{n}{2}}}\int_{\mathbb{R}^n} dk\; \overline{\hat{f}(k)}\int_{\mathbb{R}^n} dx\; f(x)e^{-\mathrm{i}\langle k,x\rangle} = \int_{\mathbb{R}^n} dk\; \big|\hat{f}(k)\big|^2\;.
\end{aligned}$$

Nach Approximation folgt die Aussage für alle f. Die Identität für das Skalarprodukt folgt aus den Polarisationsformeln und der Linearität von $\mathcal{F}$. □

29 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 6.6 für die Richtung $(\Rightarrow)$ bewiesen hatten:

Satz 29.1 *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 6.6 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

$$\int_{(V,\phi)} dS\; f(x) := \int_T du\; \sqrt{\det((D\phi)^t(u) \cdot (D\phi)(u))}\; f(\phi(u))\,. \tag{*}$$

Die Idee ist wieder zu beweisen, daß das durch die n Vektoren $a_1, \dots, a_n \subset \mathbb{R}^{n+k}$ 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 29.2 (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

$$\begin{aligned}
(D\phi)(\varphi, \vartheta) &= \begin{pmatrix} -R\sin\varphi\sin\vartheta & R\cos\varphi\cos\vartheta \\ R\cos\varphi\sin\vartheta & R\sin\varphi\cos\vartheta \\ 0 & -R\sin\vartheta \end{pmatrix} , \\
(D\phi)^t(\varphi, \vartheta) \cdot (D\phi)(\varphi, \vartheta) &= \begin{pmatrix} R^2\sin^2\vartheta & 0 \\ 0 & R^2 \end{pmatrix} ,
\end{aligned}$$

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 $\displaystyle\int_M dS = 4\pi R^2$. $\triangleleft$

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 $\displaystyle\int_{(V,\phi_1)} dS \ f(x) = \int_{(V,\phi_2)} dS \ f(x)$. 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 $(V_1, \phi_1), \dots, (V_p, \phi_p)$ gibt, die $M = V_1 \cup \dots \cup V_p$ überdecken. Dann kann man immer eine Familie von Funktionen $\alpha_i : M \to \mathbb{R}$ 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 29.3 Es sei $M \subset \mathbb{R}^{n+k}$ eine n-dimensionale Untermannigfaltigkeit, ausgestattet mit einem Atlas lokaler Karten (V_i, ϕ_i) entsprechend Satz 29.1, 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) $\displaystyle\sum_{i=0}^{\infty} \int_{(V_i,\phi_i)} dS\, |(f\alpha_i)(x)| < \infty.$

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 29.4 (Integrale längs Kurven) Es sei $I \subset \mathbb{R}^n$ ein offenes Intervall und $c : I \to \mathbb{R}^n$ eine differenzierbare Kurve. Unter der Annahme $(Dc)(t) = c'(t) \neq 0$ handelt es sich um eine Immersion, d.h. c spielt die Rolle von ϕ in Definition 29.3. Sei f jetzt eine Funktion auf der Spur $\Gamma := c(I)$, dann ist das Kurvenintegral gegeben durch

$$\int_\Gamma dS\; f(x) = \int_I dt\; \|c'(t)\|\; f(c(t))\;.$$

Insbesondere ist $L(c) = \displaystyle\int_\Gamma dS = \int_I dt\; \|c'(t)\|$ die Bogenlänge. ◁

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 29.5 (Binet-Cauchy) *Es seien $A = (a_1, \dots, a_{n+k}) \in M((n+k) \times n, \mathbb{R})$ und $B = (b_1, \dots, b_{n+k}) \in M((n+k) \times n, \mathbb{R})$ zwei rechteckige Matrizen, gebildet aus den Zeilenvektoren $a_i, b_i \in \mathbb{R}^n$. Für $1 \leq m_1 < m_2 < \dots < m_n \leq n+k$ seien quadratische Matrizen $A^{m_1 m_2 \dots m_n} := (a_{m_1}, a_{m_2}, \dots, a_{m_n}) \in M(n \times n, \mathbb{R})$ und $B^{m_1 m_2 \dots m_n} := (b_{m_1}, b_{m_2}, \dots, b_{m_n}) \in M(n \times n, \mathbb{R})$ definiert. Dann gilt*

$$\det(A^t \cdot B) = \sum_{1 \leq m_1 < m_2 \dots < m_n \leq n+k} (\det A^{m_1 m_2 \dots m_n}) \cdot (\det B^{m_1 m_2 \dots m_n})$$

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 29.6 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 $(D\phi)(u) = \begin{pmatrix} E_n \\ (\mathrm{grad}\, h)(u) \end{pmatrix}$, wenn $(\mathrm{grad}\, h)(u) \in \mathbb{R}^n$ als Zeilenvektor betrachtet wird. Dann gilt mit den Bezeichnungen aus Satz 29.5

$$\det((D\phi)^{1,2,\dots,n}) = 1 , \quad \det((D\phi)^{1,\dots,i-1,i+1,\dots,n+1}) = \pm\partial_i h$$

und damit $\det(D\phi)^t(u) \cdot (D\phi)(u) = 1 + \|(\mathrm{grad}\, h)(u)\|^2$. Somit erhalten wir das Integral einer Funktion f über den Graph $\Gamma := \phi(T) \subset \mathbb{R}^{n+1}$ (Höhenfäche) zu

$$\int_\Gamma dS\ f(x) = \int_T du\ \sqrt{1 + \|(\mathrm{grad}\, h)(u)\|^2}\ f(u, h(u)) \ .$$

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 25.3 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 29.7 (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

$$v_2(M) = \int_M dS = 2\pi \int_{-R}^{R} dz\ \sqrt{R^2-z^2}\sqrt{1+\Big(\frac{z}{\sqrt{R^2-z^2}}\Big)^2} = 2\pi R \int_{-R}^{R} dz = 4\pi R^2.$$

$\triangleleft$

30 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 30.1 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 30.2 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

$$\int_{(M,\nu)} d\vec{S}\ F(x) := \int_M dS\ \langle F(x),\nu(x)\rangle\ .$$

Zur Formulierung des Gaußschen Integralsatzes benötigen wir den Begriff des $\mathcal{C}^1$-Polyeders:

Definition 30.3 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 $y \in U$ und $G \cap U = \{y \in U \ : \ f(u) \leq 0\}$. 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 $F = (F_1, \dots, F_n)$ mit differenzierbaren Funktionen $F_i : U \to \mathbb{R}$, dann ist die Divergenz des Vektorfeldes die Funktion $(\operatorname{div} F) = \partial_1 F_1 + \dots + \partial_n F_n$.

Theorem 30.4 (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*

$$\int_G dy\, (\operatorname{div} F)(y) = \int_{\partial_r G} d\vec{S}\; F(x) = \int_{\partial_r G} dS\; \langle \nu, F(x) \rangle \;.$$

Der entscheidende Schritt im Beweis ist die Betrachtung der Situation für einen kompakten achsenparallelen Quader, der offenbar ein $\mathcal{C}^1$-Polyeder ist.

Lemma 30.5 *Es sei $Q \subset \mathbb{R}^n$ ein offener Quader und $F = (F_1, \dots, F_n)$ ein stetig differenzierbares Vektorfeld auf einer Umgebung U von Q. Dann gilt*

$$\int_Q dy\, (\operatorname{div} F)(y) = \int_{\partial Q} d\vec{S}\; F(x)$$

Beweis. Es sei $\nu = (\nu_1, \dots, \nu_n)$ 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 $i = 1, \dots, n$ 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

$$\nu_n(x) = \begin{cases} 1 & \text{auf } (Q')^o \times \{b\} \\ -1 & \text{auf } (Q')^o \times \{a\} \\ 0 & \text{auf } \partial Q' \times]a,b[\end{cases}$$

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

$$\begin{aligned} \int_{\partial_r Q} dx\; (\nu_n f)(x) &= \int_{(Q')^o} dy'\; f(y', b) - \int_{(Q')^o} dy'\; f(y', a) \\ &= \int_{(Q')^o} dy' \int_a^b dz\; (\partial_n f)(y', z) = \int_Q d(y', z)\; (\partial_n f)(y', z) \;. \qquad \square \end{aligned}$$

Wir sehen uns einige Anwendungen des Gaußschen Integralsatzes an.

Beispiel 30.6 (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 $v_{n-1}(S^{n-1}) =: \omega_n = n\kappa_n = \frac{2\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$. ◁

Beispiel 30.7 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

$$\int_{\partial_r G} d\vec{S}\; F(x) = \begin{cases} 0 & \text{für } a \notin G \\ \omega_n & \text{für } a \in G \end{cases} \tag{*}$$

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 $(\mathrm{div}\ F)(y) = 0$ 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 $q_1, \dots, q_k$ die Punktladungen in den Punkten $a_1, \dots, a_k \in \mathbb{R}^3 \setminus \partial G$, dann ist nach dem Coulombschen Gesetz die elektrische Feldstärke in einem Punkt $x \neq a_i$ gegeben durch

$$E(x) = \sum_{i=1}^{k} \frac{q_i}{4\pi} \frac{x - a_i}{\|x - a_i\|^3} \ .$$

Der Fluß des elektrischen Feldstärke durch die Oberfläche ∂G ist dann gleich der Gesamtladung in G:

$$\int_{\partial G} d\vec{S}\ E(x) = \sum_{\{i\,:\,a_i \in G\}} q_i \ . \qquad \triangleleft$$

Definition 30.8 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, denn mit $r := \|x - a\| > 0$ gilt $(n \geq 3)$

$$\Delta \tfrac{1}{r^{n-2}} = \mathrm{div}\big(\mathrm{grad} \tfrac{1}{r^{n-2}}\big) = \mathrm{div}\big(\tfrac{(2-n)}{r^{n-1}} \tfrac{x-a}{r}\big) = (2-n)\mathrm{div}\big(\tfrac{x-a}{\|x-a\|^2}\big) = 0$$

nach Beispiel 30.7. Wir zeigen nun, daß die Verallgemeinerung des Coulombschen Gesetzes auf kontinuierliche Ladungsdichten die Poisson-Gleichung löst:

Satz 30.9 *Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt und $\rho \in \mathcal{C}^1(\Omega) \cap \mathcal{C}^1(\bar{\Omega})$ mit $\rho(x) = 0$ für $x \in \mathbb{R}^n \setminus \bar{\Omega}$. Für $x \in \mathbb{R}^n$ sei*

$$\phi(x) := \int_\Omega dy\ N_y(x)\, \rho(y) \ .$$

Dann ist $\phi : \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar, und ϕ löst die Poisson-Gleichung $\Delta\phi = \rho$.

Beweis. (für $n \geq 3$) Formal gilt

$$(\partial_i \phi)(x) = \int_\Omega dy\ (\partial_i N_y)(x)\,\rho(y) = \frac{1}{\omega_n}\int_\Omega dy\ \frac{(x-y)_i}{\|x-y\|^n}\,\rho(y)$$

Für $x \notin \Omega$ existiert das Integral ohne Probleme. Für $x \in \Omega$ teilen wir Ω auf in die abgeschlossene Kugel $\overline{K_\epsilon(x)}$ und den Rest. Nach Voraussetzung ist ρ stetig und beschränkt, und nach Übergang zu Polarkoordinaten in $\overline{K_\epsilon(x)}$ folgt die Integrierbarkeit von $\frac{(x-y)_i}{\|x-y\|^n}\,\rho(y)$. Der Satz von Lebesgue sichert dann die Korrektheit der formalen Lösung.

Da ρ stetig differenzierbar ist, gilt nach Vertauschen von x, y

$$\begin{aligned}(\partial_i\, \phi)(x) &= -\int_\Omega dy\ (\partial_i N_x)(y)\,\rho(y)\\ &= -\int_\Omega dy\ (\partial_i(N_x(\,.\,)\rho))(y) + \int_\Omega dy\ N_x(y)(\partial_i\rho)(y)\end{aligned}$$

Wir identifizieren $\partial_i(N_x(\,.\,)\rho) = \mathrm{div}(N_x(\,.\,)\rho e_i)$ und erhalten über den Gaußschen Integralsatz

$$(\partial_i\, \phi)(x) = -\int_{\partial_r\Omega} dS\ \nu_i N_y(x)\rho(y) + \int_\Omega dy\ N_y(x)\frac{\partial}{\partial y_i}\big(\rho(y)-\rho(x)\big)\ .$$

Nochmalige partielle Ableitung ergibt

$$\begin{aligned}(\partial_i\partial_i\, \phi)(x) &= -\int_{\partial_r\Omega} dS\ \nu_i(\partial_i N_y)(x)\rho(y) + \int_\Omega dy\ (\partial_i N_y)(x)\frac{\partial}{\partial y_i}\big(\rho(y)-\rho(x)\big)\\ &= \int_{\partial_r\Omega} dS\ \nu_i(\partial_i N_x)(y)\rho(y)\\ &\quad - \int_\Omega dy\ \frac{\partial}{\partial y_i}\Big(\frac{\partial N_x(y)}{\partial y_i}\big(\rho(y)-\rho(x)\big)\Big) + \int_\Omega dy\ \frac{\partial^2 N_x(y)}{\partial y_i\partial y_i}\big(\rho(y)-\rho(x)\big)\\ &= \rho(x)\int_{\partial_r\Omega} dS\ \nu_i(\partial_i N_x)(y) + \int_\Omega dy\ \frac{\partial^2 N_x(y)}{\partial y_i\partial y_i}\big(\rho(y)-\rho(x)\big)\ ,\end{aligned}$$

wobei im letzten Schritt wieder der Gaußsche Integralsatz benutzt wurde. Nach Summation über i entsteht

$$(\Delta\phi)(x) = \rho(x)\int_{\partial_r\Omega} dS\ \langle \nu, \mathrm{grad}_y N_x(y)\rangle + \int_\Omega dy\ \Delta_y N_x(y)\big(\rho(y)-\rho(x)\big)\ .$$

Nach Beispiel 30.7 gilt $\displaystyle\int_{\partial_r\Omega} dS\ \langle \nu, \mathrm{grad}_y N_x(y)\rangle = \begin{cases} 1 & \text{für } x \in \Omega \\ 0 & \text{für } x \notin \Omega \end{cases}$. Im Ω-Integral ist $\Delta_y N_x(y) = 0$ für $x \neq y$, so daß wir das Integral auf $K_\epsilon(x)$ beschränken dürfen.

Dann gilt mit $\big\|\rho(y)-\rho(x)\big| \le M\|x-y\|$ und $\|x-y\|=r$

$$\begin{aligned}
&\Big|\int_{K_\epsilon(x)} dy\;(\Delta_y N_x(y))\big(\rho(y)-\rho(x)\big)\Big| \\
&\le \frac{M}{(n-2)\omega_n}\sum_{i=1}^n \int_{K_\epsilon(0)} dy\;\Big|\frac{\partial^2}{\partial y_i^2}\frac{1}{r^{n-2}}\Big| r \le \frac{M}{\omega_n}\sum_{i=1}^n \int_{K_\epsilon(0)} dy\;\Big|\frac{\partial}{\partial y_i}\frac{y_i}{r^n}\Big| r \\
&\le \frac{M}{\omega_n}\sum_{i=1}^n \int_{K_\epsilon(0)} dy\;\Big(\frac{1}{r^n}+\frac{n y_i y_i}{r^{n+2}}\Big) r = (n+1)M\epsilon\;.
\end{aligned}$$

Für $\epsilon \to 0$ folgt die Behauptung. □

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. Die Lösung ϕ der Poisson-Gleichung kann nicht eindeutig sein: $\phi + h$ mit h harmonisch ist ebenfalls eine Lösung. Die Eindeutigkeit kann durch Randbedingungen erzwungen werden.

Satz 30.10 (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*

$$\begin{aligned}
\int_G dy\;\big\langle \operatorname{grad} f, \operatorname{grad} g\big\rangle(y) &= \int_{\partial G} dS\;(f\,D_\nu g)(x) - \int_G dy\;(f\,\Delta g)(y)\;,\\
\int_G dy\;(f\,\Delta g - g\,\Delta f)(y) &= \int_{\partial G} dS\;(f\,D_\nu g - g\,D_\nu f)(x)\;,
\end{aligned}$$

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 $\operatorname{div} F = \langle \operatorname{grad} f, \operatorname{grad} g\rangle + f\Delta g$. □

Satz 30.11 *Es sei $\Omega \subset \mathbb{R}^n$ eine offene und beschränkte Teilmenge, für die der Gaußsche Integralsatz gilt, und $\nu : \partial\Omega \to \mathbb{R}^n$ sei das äußere Einheitsnormalenfeld. Eine Funktion $h \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ sei harmonisch, $\Delta h = 0$ in Ω. Dann gilt:*

i) $\displaystyle\int_{\partial\Omega} dS\;\langle \operatorname{grad} h, \nu\rangle = 0$ *(Gaußscher Integralsatz)*

ii) *Ist zu $a \in \Omega$ die Kugel $\overline{K_r(a)} \subset \Omega$, dann gilt*
$\displaystyle h(a) = \frac{1}{\omega_n r^{n-1}}\int_{\partial K_r(a)} dS\;h(x)$ *(Mittelwertsatz)*

iii) *Ist Ω zusammenhängend und nimmt h ihr Maximum oder Minimum auf $\overline{\Omega}$ im Inneren Ω an, so ist h auf $\overline{\Omega}$ konstant (Maximumprinzip).*

iv) *Erfüllen h_1, h_2 die Voraussetzungen und ist $h_1 = h_2$ auf $\partial\Omega$ mit Ω zusammenhängend, so gilt $h_1 = h_2$ auf $\overline{\Omega}$.*

Beweis. ii) (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 $(\Delta\phi)(x) = \rho(x)$.

iii) Sei $M := \max_{x\in\overline{\Omega}} h(x)$ und $F := \{x \in \Omega\ :\ h(x) = M\} \subset \Omega$. Wir zeigen, F ist offen und abgeschlossen. Das heißt $F = \Omega$ (dann ist h konstant) oder $F = \varnothing$ (dann nimmt h das Maximum auf dem Rand an).

Sei $F \neq \varnothing$ und $(x_k)_{k\in\mathbb{N}}$ eine Folge von Punkten $x_k \in F$, die gegen $x \in \Omega$ konvergiert. Dann ist $h(x_k) = M$ für alle k und wegen der Stetigkeit von h folgt $h(x) = \lim_{k\to\infty} h(x_k) = M$, also ist F abgeschlossen.

Sei $x \in F$ und $r > 0$ derart, daß $\overline{K_r(x)} \subset \Omega$. Dann folgt aus dem Mittelwertsatz $h(y) = M$ für alle $y \in \partial K_r(x)$, denn wäre $h(y) < M - \epsilon$ für ein $\epsilon > 0$, dann folgt aus der stetigen Differenzierbarkeit von h die Existenz einer Teilmenge $V \subset \partial K_r(x)$ mit Volumen $\geq \delta > 0$, auf der $h(y) < M - \epsilon/2$ gilt. Nach ii) wäre dann auch $h(x) < M$, Widerspruch. Also gibt es zu jedem $x \in F$ auch eine Kugel $K_r(x) \subset F$, damit ist F offen. Die Aussagen für das Minimum ergeben sich mit $h \mapsto -h$.

iv) Mit h_1, h_2 ist auch $h := h_1 - h_2$ harmonisch mit $h = 0$ auf $\partial\Omega$. Dann ist $h = 0$ auf $\overline{\Omega}$ nach iii). □

Eine wichtige Anwendung betrifft das Dirichlet-Problem

$$(\Delta\phi)(x) = \rho(x) \text{ für alle } x \in \Omega\ , \qquad \phi(y) = f \text{ für alle } y \in \partial\Omega$$

mit vorgegebenen Ladungsdichten ρ und Randwerten f. Nach Satz 30.11 gibt es höchstens eine Lösung des Dirichlet-Problems.

Satz 30.12 (Sektorformel von Leibniz) *Es sei $c : [\alpha, \beta] \to \mathbb{R}^2$ eine differenzierbare reguläre Kurve und $S_c := \{s \cdot c(t) \ : \ s \in [0,1] \ , \ t \in [\alpha, \beta]\}$ der von c gebildete Sektor. Dann ist mit $c(t) = (x(t), y(t))$ der Flächeninhalt des Sektors gegeben durch*

$$v_2(S_c) = \frac{1}{2} \int_c \big(- y dx + x dy \big) = \frac{1}{2} \int_\alpha^\beta dt \ \big(- y(t)x'(t) + x(t)y'(t) \big) \ .$$

Beweis. Der Tangentialvektor an c in t ist $(x'(t), y'(t))$. Damit ist $\nu := \Big(\frac{y'(t)}{\|c(t)\|}, - \frac{x'(t)}{\|c(t)\|} \Big)$ mit $\|c(t)\| = \sqrt{(y'(t))^2 + (x'(t))^2}$ ein Einheitsnormalenvektor an den von c gebildeten Rand von S_c. Wird c im positiven Sinn durchlaufen, dann ist ν der äußere Einheitsnormalenvektor. Betrachte das Vektorfeld $F = \mathrm{id} : \mathbb{R}^2 \to \mathbb{R}^2$, d.h. $F(x,y) = (x,y)$. Nach Beispiel 29.4 gilt

$$\int_{c([\alpha,\beta])} dS \ \langle F, \nu \rangle = \int_\alpha^\beta dt \ \|c'(t)\| \langle F(c(t)), \nu \rangle = \int_\alpha^\beta dt \ \big(x(t)y'(t) - y(t)x'(t) \big) \ .$$

Auf den Radien $sc(\alpha)$ und $sc(\beta)$, mit $s \in [0,1]$ ist F tangential gerichtet, somit $\langle F, \nu \rangle = 0$. Andererseits ist $\mathrm{div}(F) = 2$ und damit $\int_G dy \ (\mathrm{div} \ F)(y) = 2 v_2(G)$. Der Gaußsche Integralsatz liefert die Behauptung. $\square$

Beispiel 30.13 Sei $c(t) = (x(t), y(t))$ mit $x = r \cos t$ und $y = r \sin t$ ein Kreisbogen, dann ist

$$v_2(S_c) = \frac{r^2}{2} \int_\alpha^\beta dt \ (\sin^2 t + \cos^2 t) = \frac{r^2(\beta - \alpha)}{2} \ .$$

Der Vollkreis $\beta - \alpha = 2\pi$ hat die Fläche πr^2 $\triangleleft$

Als weitere Anwendung leiten wir die Keplerschen Gesetze aus dem Newtonschen Gravitationsgesetz

$$x''(t) = - \frac{M x(t)}{\|x(t)\|^3}$$

her. Dabei ist M für alle Planeten der gleiche Wert. Wir benötigen das Vektorprodukt von Vektoren $a = (a_1, a_2, a_3)$ und $b = (b_1, b_2, b_3)$:

$$(a \times b)_1 := a_2 b_3 - a_3 b_2 \ , \qquad (a \times b)_2 := a_3 b_1 - a_1 b_3 \ , \qquad (a \times b)_3 := a_1 b_2 - a_2 b_1 \ .$$

Dann gilt $a \times b = -b \times a$, insbesondere $a \times a = 0$. Weiter gilt die Graßmann-Identität $a \times (b \times c) = b \langle a, c \rangle - c \langle a, b \rangle$ sowie $\langle a, b \times c \rangle = \det(a,b,c) = \langle b, c \times a \rangle = \langle c, a \times b \rangle$ (Spatprodukt).

Der Drehimpuls $J = x \times x'$ ist eine Erhaltungsgröße:

$$J' = x \times x'' + x' \times x' = - \frac{G}{\|x\|^3} x \times x = 0 \ .$$

Weiter sei $A := \frac{1}{M} J \times x' + \frac{x}{\|x\|}$ der Lenz-Runge-Vektor. Er ist ebenfalls eine Erhaltungsgröße:

$$A' = \frac{1}{M} J \times x'' + \frac{x'}{\|x\|} - \frac{x\langle x, x'\rangle}{\|x\|^3} = \frac{1}{\|x\|^3}\Big(-(x \times x') \times x + x'\langle x, x\rangle - x\langle x, x'\rangle\Big) = 0$$

wegen der Graßmann-Identität. Die Gleichung $\langle x, J\rangle = 0$ (Spatprodukt) besagt, daß der Bahnvektor x in einer Ebene E senkrecht zum konstanten Drehimpulsvektor J liegt. Dann ist auch $A \perp J$, d.h. A, x liegen in E. Somit gilt

$$\langle A, x\rangle = \|A\|\|x\| \cos\phi = \frac{1}{M}\langle J \times x', x\rangle + \|x\| = -\frac{1}{M}\langle x \times x', J\rangle + \|x\| = -\frac{\|J\|^2}{M} + \|x\| \;.$$

Ist $A = 0$, so ist $r := \|x\| = \frac{\|J\|^2}{M} = \text{const}$, die Bahn also ein Kreis. Ansonsten folgt

$$r = \frac{p}{1 - \epsilon\cos\phi}\;, \qquad \epsilon = \|A\|\;,\; p = \frac{\|J\|^2}{M}\;.$$

Das ist die Darstellung eines Kegelschnittes in Polarkoordinaten $(r(\phi), \phi)$ mit Brennpunkt $r = 0$. Somit folgt:

Die Bahnen der Planeten sind ebene Kegelschnitte, in deren einen Brennpunkt die Sonne steht.

Sei $J = \|J\|e_3$ und $x = x_1e_1 + x_2e_2$, dann ist $\|J\| = x_1x_2' - x_2x_1'$. Nach der Sektorformel von Leibniz gilt

$$\|J\|(t_2 - t_1) = \int_{t_1}^{t_2} dt\; \|J\| = \int_{t_1}^{t_2} dt\; (x_1x_2' - x_2x_1') = 2v_2(S_{t_1,t_2})\;,$$

wobei S_{t_1,t_2} der Sektor unter der Bahnkurve zwischen t_1, t_2 ist. Somit gilt:

Der Radiusvektor Sonne-Planet überstreicht in gleichen Zeiten gleiche Flächen.

Für $0 \leq \epsilon < 1$ ist die Bahn eine geschlossene Ellipse. Ist T die Umlaufzeit, dann ist $F = \frac{\|J\|T}{2}$ die Fläche der Ellipse. Andererseits ist

$$F = \int_0^{2\pi} d\phi \int_0^{r(\phi)} d\rho\; \rho = \int_0^{2\pi} d\phi \frac{1}{2}\Big(\frac{p}{1 - \epsilon\cos\phi}\Big)^2 = \frac{\pi p^2}{(1 - \epsilon^2)^{\frac{3}{2}}}\;.$$

Das Integral kann über den Residuensatz berechnet werden:

$$\begin{aligned}
\int_0^{2\pi} d\phi \frac{1}{(1 - \epsilon\cos\phi)^2} &= 2\pi\, \mathrm{res}_{z_0 \in K_1(0)}\Big(\frac{1}{z}\frac{1}{(1 - \frac{\epsilon}{2}(z + \frac{1}{z}))^2}\Big) \\
&= 2\pi\, \mathrm{res}_{z_0 = \frac{1 - \sqrt{1-\epsilon^2}}{\epsilon}}\Big(\frac{4z}{(2z - \epsilon(z^2 + 1))^2}\Big) \\
&= 2\pi\Big(\frac{4z(z - z_0)^2}{(2z - \epsilon(z^2 + 1))^2}\Big)'\Big|_{z = z_0}
\end{aligned}$$

$$
\begin{aligned}
&= 2\pi\Big(\frac{4z}{(\epsilon z - 1 - \sqrt{1-\epsilon^2})^2}\Big)'\Big|_{z=z_0} \\
&= 2\pi\Big(\frac{4}{(\epsilon z - 1 - \sqrt{1-\epsilon^2})^2} - \frac{8z\epsilon}{(\epsilon z - 1 - \sqrt{1-\epsilon^2})^3}\Big)\Big|_{z=z_0} \\
&= \frac{2\pi}{(\sqrt{1-\epsilon^2})^3} \;.
\end{aligned}
$$

Die große Halbachse a der Ellipse ist definiert durch $2a := r(0) + r(\pi) = \frac{p}{1-\epsilon} + \frac{p}{1+\epsilon^2} = \frac{2p}{1-\epsilon^2} = \frac{2\|J\|^2}{M(1-\epsilon^2)}$. Somit gilt: $\|J\|T = 2\pi\frac{\|J\|^4}{M^2(1-\epsilon^2)^{\frac{3}{2}}} = \frac{2\pi}{\sqrt{M}}\|J\|a^{\frac{3}{2}}$.

Die Quadrate der Umlaufzeiten zweier Planeten verhalten sich zueinander wie die dritten Potenzen der großen Halbachsen ihrer Bahnen.

Der Gaußsche Integralsatz ist Spezialfall des *Integralsatzes von Stokes*. Dieser wird in der Sprache der Differentialformen formuliert, von denen wir bisher nur 1-Formen eingeführt haben. Eine 1-Form im $\mathbb{R}^n$ ließ sich schreiben als $\omega = \sum_{i=1}^n \omega_i dx_i$. Man führt formal ein *äußeres Produkt* $\wedge$ ein, welches assoziativ, distributiv und *antikommutativ* auf den Differentialen dx_i ist:

$$
dx_i \wedge (dx_j \wedge dx_k) = (dx_i \wedge dx_j) \wedge dx_k \;, \qquad dx_i \wedge dx_j = -dx_j \wedge dx_i \;.
$$

Insbesondere ist $dx_i \wedge dx_i = 0$. Dann hat eine k-Form $\omega^{(k)}$ die Darstellung $\omega^{(k)} = \sum_{1\leq i_1<i_2<\dots<i_k\leq n} \omega^{(k)}_{i_1 i_2\dots i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}$, mit differenzierbaren Funktionen $\omega^{(k)}_{i_1 i_2\dots i_k} : \mathbb{R}^n \to \mathbb{R}$. Wir bezeichen mit $\Lambda^k(\mathbb{R}^n)$ den Vektorraum der k-Formen im $\mathbb{R}^n$. Eine k-Form im $\mathbb{R}^n$ hat also $\binom{n}{k}$ Komponenten. Weiter wird das äußere Differential $d : \Lambda^n(\mathbb{R}^n) \to \Lambda^{k+1}(\mathbb{R}^n)$ eingeführt durch

$$
d\omega^{(k)} = \sum_{1\leq i_1<i_2<\dots<i_k\leq n} \sum_{j=1}^n \partial_j \omega^{(k)}_{i_1\dots i_k} dx_j \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \;.
$$

Wegen Satz von Schwarz und der Antisymmetrie von $\wedge$ gilt $d^2 = 0$.

Eine k-Form $\omega^{(k)} \in \Lambda^k(\mathbb{R}^n)$ läßt sich über k-dimensionale Untermannigfaltigkeiten im $\mathbb{R}^n$ integrieren. Zunächst definiert man für $\omega^{(n)} \in \Lambda^n(\mathbb{R}^n)$ das Integral über $G \subset \mathbb{R}^n$ als

$$
\int_G \omega_{12\dots n} dx_1 \wedge dx_2 \wedge \dots \wedge dx_n := \int_G \omega_{12\dots n} d(x_1, x_2, \dots, x_n) \;.
$$

Über die Kartenabbildungen $\phi : T \to V \subset M$ lassen sich k-Formen $\omega^{(k)} \in \Lambda^k(\mathbb{R}^n)$ über k-dimensionale Untermannigfaltigkeiten $M \subset \mathbb{R}^n$ integrieren. Dann gilt:

Satz 30.14 (Stokes) *Sei $M \subset \mathbb{R}^n$ eine k-dimensionale orientierte Untermannigfaltigkeit, deren Rand ∂M eine $(k-1)$-dimensionale orientierte Untermannigfaltigkeit ist. Dann gilt für $\omega^{(k-1)} \in \Lambda^k(\mathbb{R}^n)$*

$$
\int_M d\omega^{(k-1)} = \int_{\partial M} \omega^{(k)} \;.
$$

Symbolisch schreibt man $\langle M, d\omega\rangle = \langle \partial M, \omega\rangle$, d.h. Differential d und Rand ∂ sind zueinander adjungiert bezüglich der durch das Integral gegebenen Linearform.

Aus Dimensionsgründen lassen sich k-Formen mit $(n-k)$-Formen identifizieren, indem man $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ auf das $\wedge$-Produkt der $(n-k)$ fehlenden dx_j abbildet (mit Vorzeichen). Diese Abbildung wird mit $*$ bezeichnet, $* : \Lambda^k(\mathbb{R}^n) \to \Lambda^{n-k}(\mathbb{R}^n)$. Es gilt $*^2 = \pm 1$. Insbesondere identifizieren wir n-Formen mit Funktionen (0-Formen) und 1-Formen mit $(n-1)$-Formen. Außerdem können 1-Formen mit Vektorfeldern $v \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ identifiziert werden durch eine Abbildung $I : \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n) \to \Lambda^1(\mathbb{R}^n)$ definiert durch $I(e_i) = dx_i$. Damit erhalten wir eine neue Interpretation der Divergenz eines Vektorfeldes F als die Funktion

$$\operatorname{div} F = *d * (I(F)) \ .$$

Denn ist $F = fe_1$, dann $I(F) = f dx_1$, weiter $*(I(F)) = f dx_2 \wedge \cdots \wedge dx_n$ und damit $d * (I(F)) = (\partial_1 f) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$, denn alle anderen partiellen Ableitungen $(\partial_k f)$ mit $k \neq 1$ führen auf einen Faktor dx_k, der in $dx_2 \wedge \cdots \wedge dx_n$ bereits vorkommt. Nochmaliges Anwenden von $*$ nimmt alle dx_i wieder weg, $*d * (I(F)) = (\partial_1 f)$. Beachtet man noch die Beziehung $\int_G *f = \int_G dx\ f(x)$ für das Volumenintegral einer Funktion f, so liefert der Stokessche Integralsatz

$$\int_G d * (I(F)) = \int_G *(*d * (I(F))) = \int_G dx \operatorname{div}(F) = \int_{\partial G} *(I(F)) =: \int_{\partial G} d\vec{S}\ F$$

den Gaußschen.

Im $\mathbb{R}^3$ wird die *Rotation* eines Vektorfeldes F erklärt als das Vektorfeld

$$\operatorname{rot} F = I^{-1}(*d(I(F)) \ .$$

Nehmen wir $F = fe_3$ also $I(F) = f dx_3$, dann $d(I(F)) = (\partial_1 f) dx_1 \wedge dx_3 + (\partial_2 f) dx_2 \wedge dx_3 = -(\partial_1 f) dx_3 \wedge dx_1 + (\partial_2 f) dx_2 \wedge dx_3$ und $*d(I(F)) = -(\partial_1 f) dx_2 + (\partial_2 f) dx_1$ sowie $\operatorname{rot} F = -(\partial_1 f) e_2 + (\partial_2 f) e_1$. Ist $M \subset \mathbb{R}^3$ eine zweidimensionale Untermannigfaltigkeit mit eindimensionalem Rand ∂M, dann liefert Stokes

$$\int_M d(I(F)) = \int_M *I(\operatorname{rot} F) = \int_M d\vec{S} \operatorname{rot} F = \int_{\partial M} I(F) \ .$$

Das dritte Integral ist $\int_M d\vec{S} \operatorname{rot} F = \int_M dS\ \langle \nu, \operatorname{rot} F\rangle$, wobei ν wieder ein Einheitsnormalenvektor von M ist. Die rechte Seite ist das Kurvenintegral der Differentialform $I(F)$ längs der Randkurve, deren Umlaufsinn durch die Wahl von ν entsprechend der rechten Hand zu wählen ist.