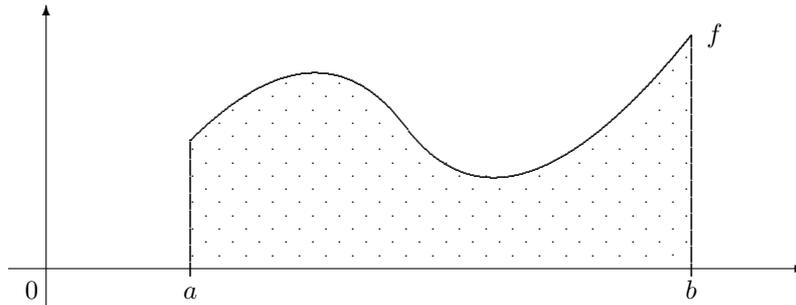


8 Das Riemann-Integral

Es sind wenigstens zwei Probleme, die zur Herausbildung der Integralrechnung geführt haben.

Flächenberechnungen. Gegeben ist eine Funktion $f : [a, b] \rightarrow \mathbb{R}$. Gesucht ist der Inhalt der von den Geraden $x = a$, $x = b$, $y = 0$ und vom Graphen der Funktion f begrenzten Fläche.

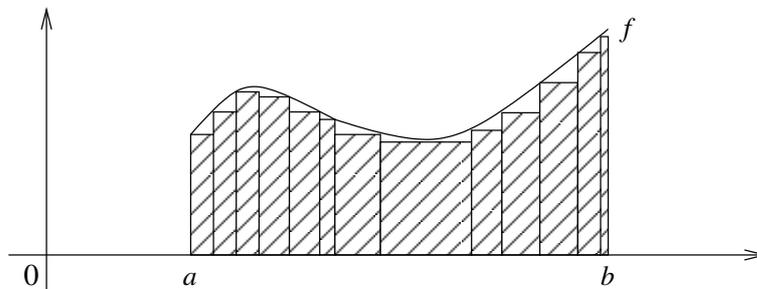


Umkehrung des Differenzierens. Kann man aus der Ableitung einer Funktion die Funktion selbst rekonstruieren? Gibt es für jede Funktion f eine Funktion g mit $g' = f$? Wenn ja, wieviele?

Zur Lösung dieser und anderer Fragen wurden verschiedene Integralbegriffe entwickelt, von denen wir einen der einfachsten – das Riemann-Integral – hier kennenlernen wollen. Eine in vielerlei Hinsicht zufriedenstellendere Integrationstheorie wurde zu Beginn des 20. Jahrhunderts von Lebesgue geschaffen. Wir kommen im 4. Semester darauf zurück.

8.1 Der Begriff des Riemann-Integrals

Eine naheliegende Idee zur Berechnung der Fläche eines komplizierten Gebietes (etwa des oben skizzierten) ist es, das Gebiet durch andere Gebiete “anzunähern”, deren Flächenberechnung einfacher ist, z.B. durch eine endliche Anzahl von Rechtecken. Wir hoffen, dass wir dem “wahren” Flächeninhalt näherkommen, wenn wir die “Approximation” verbessern, indem wir z.B. die Rechtecke schmaler machen. Diese vage Vorstellung wollen wir nun präzisieren.



Für jedes $i = 0, \dots, n$ sei x_i ein Punkt aus $[a, b]$ mit $a = x_0 < x_1 < \dots < x_n = b$. Die Menge $Z := \{x_0, \dots, x_n\}$ heißt eine *Zerlegung* des Intervalls $[a, b]$. Die Länge des i . Teilintervalls $I_i := [x_i, x_{i+1}]$ bezeichnen wir mit $\Delta x_i := x_{i+1} - x_i$ oder $|I_i|$. Die Zahl $|Z| := \max_{0 \leq i \leq n-1} \Delta x_i$ heißt das *Feinheitsmaß* oder die *Maschenweite* von Z . Ist Z eine Zerlegung von $[a, b]$ und geht eine Zerlegung Z' von $[a, b]$ aus Z durch Hinzufügen weiterer Punkte hervor, so heißt Z' eine *Verfeinerung* von Z . Offenbar ist dann $|Z'| \leq |Z|$.

Sei $Z = \{x_0, \dots, x_n\}$ eine Zerlegung von $[a, b]$ und $\xi := (\xi_0, \dots, \xi_{n-1})$ ein zugehöriger *Zwischenvektor*, d.h. es gelte $x_i \leq \xi_i \leq x_{i+1}$ für jedes i . Dann heißt

$$S(Z, \xi, f) := \sum_{i=0}^{n-1} f(\xi_i) \Delta x_i$$

eine *Riemann-Summe* für f .

Sei schließlich $(Z_n)_{n=1}^\infty$ eine Folge von Zerlegungen von $[a, b]$ mit $|Z_n| \rightarrow 0$ für $n \rightarrow \infty$, und für jedes n sei $\xi^{(n)}$ ein Zwischenvektor für Z_n . Dann heißt die Folge $(S(Z_n, \xi^{(n)}, f))_{n=1}^\infty$ der entsprechenden Riemann-Summen eine *Riemann-Folge* für f .

Definition 8.1 Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt Riemann-integrierbar, wenn jede Riemann-Folge $(S(Z_n, \xi^{(n)}, f))_{n=1}^\infty$ für f konvergiert.

Lemma 8.2 Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so haben alle Riemann-Folgen für f den gleichen Grenzwert.

Beweis. Sind $(S_n^{(1)})$ und $(S_n^{(2)})$ Riemann-Folgen für f , so ist auch

$$(S_n) := (S_1^{(1)}, S_1^{(2)}, S_2^{(1)}, S_2^{(2)}, \dots)$$

eine Riemann-Folge für f . Nach Voraussetzung konvergiert die Folge (S_n) . Dann konvergieren auch ihre Teilfolgen $(S_n^{(1)})$ und $(S_n^{(2)})$, und diese haben den gleichen Grenzwert wie (S_n) . ■

Definition 8.3 Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion. Der gemeinsame Grenzwert aller Riemann-Folgen für f heißt das Riemann-Integral von f . Wir bezeichnen diesen Grenzwert mit $\int_a^b f(x) dx$.

Satz 8.4 Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist f beschränkt.

Kurz: Riemann-integrierbare Funktionen sind beschränkt.

Beweis. Wir zeigen die Beschränktheit nach oben. Angenommen, f ist auf $[a, b]$ nach oben unbeschränkt. Dann gibt es eine Folge (x_i) in $[a, b]$ mit $f(x_i) \rightarrow +\infty$. Aus der Folge (x_i) lässt sich eine konvergente Teilfolge auswählen (Bolzano/Weierstraß). Der Einfachheit halber nehmen wir an, dass die Folge (x_i) bereits konvergent ist. Ihr Grenzwert sei \hat{x} .

Sei Z_n die Zerlegung von $[a, b]$ in n gleichlange Teilintervalle. Dann gibt es sicher ein Teilintervall I_m von Z , welches \hat{x} und unendlich viele der Punkte x_i enthält. Für alle Teilintervalle $I_k \neq I_m$ von Z wählen wir beliebige Zwischenpunkte ξ_k und setzen

$$S'_n := \sum_{k \neq m} f(\xi_k) |I_k|.$$

Wegen $\lim_{i \rightarrow \infty} f(x_i) = +\infty$ können wir als Zwischenpunkt $\xi_m \in I_m$ einen der Punkte x_i so wählen, dass $f(\xi_m) > (n - S'_n)/|I_m|$ bzw.

$$S(Z_n, \xi^{(n)}, f) = S'_n + f(\xi_m)|I_m| > n.$$

Es lässt sich also eine Riemann-Folge $(S(Z_n, \xi^{(n)}, f))_{n=1}^\infty$ für f konstruieren, welche bestimmt gegen $+\infty$ divergiert. Dann kann f nicht Riemann-integrierbar sein. ■

8.2 Darbousche Integrale

Wir betrachten einen weiteren Zugang zum Riemann-Integral. Für jede Zerlegung Z von $[a, b]$ und jede *beschränkte* Funktion $f : [a, b] \rightarrow \mathbb{R}$ sei

$$m_i := \inf_{x \in I_i} f(x), \quad M_i := \sup_{x \in I_i} f(x).$$

Wir nennen

$$U(Z, f) := \sum_{i=0}^{n-1} m_i \Delta x_i \quad \text{bzw.} \quad O(Z, f) := \sum_{i=0}^{n-1} M_i \Delta x_i$$

die zugehörige *Unter-* bzw. *Obersumme*.

Lemma 8.5 *Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Dann gilt*

(a) *für jede Zerlegung Z von $[a, b]$ und jede Verfeinerung Z' von Z , dass*

$$U(Z, f) \leq U(Z', f) \quad \text{und} \quad O(Z', f) \leq O(Z, f),$$

(b) *für je zwei Zerlegungen Z_1, Z_2 von $[a, b]$, dass $U(Z_1, f) \leq O(Z_2, f)$.*

Beweis. (a) Wir zeigen nur die erste Behauptung, und diese nur für den Fall, dass Z' genau einen Punkt x^* mehr enthält als Z . Unterscheiden sich Z und Z' um $p > 1$ Punkte, muss diese Überlegung p mal angewandt werden.

Sei $Z = (x_0, \dots, x_n)$ und $x_j < x^* < x_{j+1}$. Für

$$m'_j := \inf_{x \in [x_j, x^*]} f(x) \quad \text{und} \quad m''_j := \inf_{x \in [x^*, x_{j+1}]} f(x)$$

ist $m_j \leq m'_j$ und $m_j \leq m''_j$. Wir können daher abschätzen

$$m_j \Delta x_j = m_j(x^* - x_j) + m_j(x_{j+1} - x^*) \leq m'_j(x^* - x_j) + m''_j(x_{j+1} - x^*).$$

Da beim Übergang von $U(Z, f)$ zu $U(Z', f)$ lediglich der Summand $m_j \Delta x_j$ durch $m'_j(x^* - x_j) + m''_j(x_{j+1} - x^*)$ ersetzt wird, folgt die Behauptung.

(b) Sei Z' eine gemeinsame Verfeinerung von Z_1 und Z_2 . Dann ist $U(Z_1, f) \leq U(Z', f)$ und $O(Z', f) \leq O(Z_2, f)$ nach Teil (a). Die Ungleichung $U(Z', f) \leq O(Z', f)$ gilt offenbar. ■

Aus Lemma 8.5 (b) folgt, dass die Menge der Untersummen einer beschränkten Funktion nach oben und die Menge der Obersummen nach unten beschränkt ist. Damit existieren (als endliche Zahlen)

$$\inf_Z O(Z, f) =: \int_a^{*b} f(x) dx \quad \text{und} \quad \sup_Z U(Z, f) =: \int_{*a}^b f(x) dx.$$

Diese Zahlen heißen *oberes* und *unteres Darboux'sches Integral* von f . Aus Lemma 8.5 (b) folgt sofort

$$\int_{*a}^b f(x) dx \leq \int_a^{*b} f(x) dx.$$

Satz 8.6 Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt und (Z_n) eine Folge von Zerlegungen von $[a, b]$ mit $\lim_{n \rightarrow \infty} |Z_n| = 0$. Dann existieren die Grenzwerte $\lim_{n \rightarrow \infty} U(Z_n, f)$ und $\lim_{n \rightarrow \infty} O(Z_n, f)$, und es gilt:

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

Beweis. Wir führen den Beweis für Untersummen. Sei $\varepsilon > 0$. Dann gibt es eine Zerlegung Z' von $[a, b]$ so, dass

$$0 \leq \int_{*a}^b f(x) dx - U(Z', f) < \varepsilon/2.$$

Diese Zerlegung bestehe aus $r + 2$ Punkten mit $r \geq 1$. Weiter sei

$$M := \sup_{x \in [a, b]} |f(x)|,$$

und $N \in \mathbb{N}$ sei so gewählt, dass

$$|Z_n| \leq \frac{\varepsilon}{4rM} \quad \text{für alle } n \geq N$$

(was wegen $|Z_n| \rightarrow 0$ möglich ist). Für jedes $n \in \mathbb{N}$ sei schließlich $Z'_n := Z_n \cup Z'$ gemeinsame Verfeinerung von Z_n und Z' . Nach Lemma 8.5 (a) gilt

$$\int_{*a}^b f(x) dx - U(Z'_n, f) \leq \int_{*a}^b f(x) dx - U(Z', f) < \varepsilon/2. \quad (8.1)$$

Da beim Übergang von Z_n zu Z'_n höchstens r neue Punkte hinzukommen, gilt

$$U(Z'_n, f) - U(Z_n, f) \leq 2rM|Z_n|. \quad (8.2)$$

Wir überlegen uns (8.2) zunächst für *einen* neu hinzukommenden Punkt t . Liegt t etwa in $[x_i, x_{i+1}]$, wobei die x_i Punkte der Zerlegung Z_n sind, so hat man

$$\begin{aligned} U(Z'_n, f) - U(Z_n, f) &= \inf_{x \in [x_i, t]} f(x)(t - x_i) + \inf_{x \in [t, x_{i+1}]} f(x)(x_{i+1} - t) \\ &\quad - \inf_{x \in [x_i, x_{i+1}]} f(x)(x_{i+1} - x_i) \\ &\leq M(t - x_i) + M(x_{i+1} - t) + M(x_{i+1} - x_i) \\ &= 2M\Delta x_i \leq 2M|Z_n|. \end{aligned}$$

Damit ist (8.2) für einen hinzukommenden Punkt gezeigt. Der allgemeine Fall ergibt sich durch Wiederholung dieser Überlegungen.

Addition von (8.1) und (8.2) ergibt nun

$$0 \leq \int_{*a}^b f(x) dx - U(Z_n, f) < \varepsilon/2 + 2rM|Z_n| \leq \varepsilon/2 + \frac{2rM\varepsilon}{4rM} = \varepsilon$$

für alle $n \geq N$, woraus die behauptete Konvergenz folgt. ■

Definition 8.7 Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt Darboux-integrierbar, wenn

$$\int_{*a}^b f(x) dx = \int_a^{*b} f(x) dx.$$

Satz 8.8 Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Darboux-integrierbar, wenn es für jedes $\varepsilon > 0$ eine Zerlegung Z von $[a, b]$ mit $O(Z, f) - U(Z, f) < \varepsilon$ gibt.

Beweis. (\Leftarrow) Sei $\varepsilon > 0$. Wir wählen eine Zerlegung Z mit $O(Z, f) - U(Z, f) < \varepsilon$. Wegen

$$U(Z, f) \leq \int_{*a}^b f(x) dx \leq \int_a^{*b} f(x) dx \leq O(Z, f)$$

ist dann erst recht

$$0 \leq \int_a^{*b} f(x) dx - \int_{*a}^b f(x) dx < \varepsilon.$$

Dies gilt für jedes $\varepsilon > 0$. Also sind oberes und unteres Darboux-sches Integral gleich.

(\Rightarrow) Sei f Darboux-integrierbar und $J := \int_{*a}^b f(x) dx$. Sei $\varepsilon > 0$ beliebig vorgegeben. Da J das Infimum von Ober- und das Supremum von Untersummen ist, gibt es Zerlegungen Z_1, Z_2 von $[a, b]$ so, dass

$$J - U(Z_1, f) < \varepsilon/2 \quad \text{und} \quad O(Z_2, f) - J < \varepsilon/2.$$

Sei Z eine gemeinsame Verfeinerung von Z_1 und Z_2 . Dann ist nach Lemma 8.5 (a) $U(Z_1, f) \leq U(Z, f)$ sowie $O(Z, f) \leq O(Z_2, f)$ und folglich

$$J - U(Z, f) < \varepsilon/2 \quad \text{und} \quad O(Z, f) - J < \varepsilon/2.$$

Addition dieser Ungleichungen ergibt $O(Z, f) - U(Z, f) < \varepsilon$. ■

Der Zusammenhang zwischen Riemann- und Darboux-Integrierbarkeit wird durch folgenden Satz hergestellt.

Satz 8.9 *Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie Darboux-integrierbar ist. In diesem Fall gilt*

$$\int_a^b f(x) dx = \int_{*a}^b f(x) dx = \int_a^{*b} f(x) dx. \quad (8.3)$$

Beweis. Sei zunächst f Darboux-integrierbar und $(S(Z_n, \xi^{(n)}, f))$ eine Riemann-Folge für f . Für alle $n \in \mathbb{N}$ ist dann

$$U(Z_n, f) \leq S(Z_n, \xi^{(n)}, f) \leq O(Z_n, f). \quad (8.4)$$

Nach Satz 8.6 konvergieren die linke Seite von (8.4) für $n \rightarrow \infty$ gegen $\int_a^b f(x) dx$ und die rechte Seite gegen $\int_a^{*b} f(x) dx$, und diese Grenzwerte sind gleich. Dann muss aber auch die Riemann-Folge konvergieren, und ihr Grenzwert stimmt mit $\int_{*a}^b f(x) dx$ überein (Satz 4.2). Insbesondere ist f Riemann-integrierbar, und es gilt (8.3).

Sei nun f Riemann-integrierbar und $\varepsilon > 0$ beliebig vorgegeben. Wir konstruieren zwei spezielle Riemann-Folgen für f wie folgt: Z_n sei die Zerlegung von $[a, b]$ in n Intervalle gleicher Länge, und $\xi'_i, \xi''_i \in [x_i, x_{i+1}]$ seien so gewählt, dass

$$\sup_{x \in [x_i, x_{i+1}]} f(x) - f(\xi'_i) < \frac{\varepsilon}{4(b-a)}, \quad f(\xi''_i) - \inf_{x \in [x_i, x_{i+1}]} f(x) < \frac{\varepsilon}{4(b-a)}. \quad (8.5)$$

Wir setzen noch $\xi_{\text{sup}}^{(n)} := (\xi'_0, \dots, \xi'_{n-1})$ und $\xi_{\text{inf}}^{(n)} := (\xi''_0, \dots, \xi''_{n-1})$ und erhalten zwei Riemann-Folgen $(S(Z_n, \xi_{\text{sup}}^{(n)}, f))$ und $(S(Z_n, \xi_{\text{inf}}^{(n)}, f))$, die beide gegen $J := \int_a^b f(x) dx$ konvergieren. Es gibt daher ein N so, dass

$$|S(Z_N, \xi_{\text{sup}}^{(N)}, f) - J| < \varepsilon/4 \quad \text{und} \quad |S(Z_N, \xi_{\text{inf}}^{(N)}, f) - J| < \varepsilon/4. \quad (8.6)$$

Zur Abkürzung sei noch $Z := Z_N$, $\xi_s := \xi_{\text{sup}}^{(N)}$ und $\xi_i := \xi_{\text{inf}}^{(N)}$. Wegen (8.5) ist

$$O(Z, f) \leq S(Z, \xi_s, f) + \frac{\varepsilon}{4} \quad \text{und} \quad U(Z, f) \geq S(Z, \xi_i, f) - \frac{\varepsilon}{4},$$

woraus mit (8.6) schließlich

$$\begin{aligned} O(Z, f) - U(Z, f) &\leq S(Z, \xi_s, f) + \frac{\varepsilon}{4} - S(Z, \xi_i, f) + \frac{\varepsilon}{4} \\ &\leq \frac{\varepsilon}{2} + |S(Z, \xi_s, f) - J| + |J - S(Z, \xi_i, f)| < \varepsilon \end{aligned}$$

folgt. Nach Satz 8.8 ist f Darboux-integrierbar. ■

Eine unmittelbare Folgerung aus Satz 8.8 und Satz 8.9 ist

Folgerung 8.10 (Riemannsches Integritätskriterium) Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Dann ist f genau dann Riemann-integrierbar, wenn für jedes $\varepsilon > 0$ eine Zerlegung Z von $[a, b]$ mit $O(Z, f) - U(Z, f) < \varepsilon$ existiert.

Beispiel. Für die Dirichlet-Funktion $f : [a, b] \rightarrow \mathbb{R}$,

$$f(x) := \begin{cases} 1 & \text{falls } x \text{ rational} \\ 0 & \text{falls } x \text{ irrational} \end{cases}$$

ist jede Obersumme gleich $b - a$ und jede Untersumme gleich 0. Diese Funktion ist also nicht Riemann-integrierbar. ■

8.3 Einige Klassen Riemann-integrierbarer Funktionen

Satz 8.11 Monotone Funktionen auf $[a, b]$ sind Riemann-integrierbar.

Beweis. Sei $f : [a, b] \rightarrow \mathbb{R}$ beispielsweise monoton wachsend. Für jedes $n \in \mathbb{N} \setminus \{0\}$ sei Z_n die Zerlegung von $[a, b]$ in n Teilintervalle gleicher Länge. Dann ist

$$U(Z_n, f) = \sum_{k=0}^{n-1} m_k \Delta x_k = \frac{b-a}{n} \sum_{k=0}^{n-1} f(x_k)$$

sowie

$$O(Z_n, f) = \sum_{k=0}^{n-1} M_k \Delta x_k = \frac{b-a}{n} \sum_{k=0}^{n-1} f(x_{k+1}).$$

Also ist

$$O(Z_n, f) - U(Z_n, f) = \frac{b-a}{n} \left(\sum_{k=1}^n f(x_k) - \sum_{k=0}^{n-1} f(x_k) \right) = \frac{b-a}{n} (f(b) - f(a)).$$

Dies wird kleiner als jedes vorgegebene $\varepsilon > 0$, wenn nur n groß genug ist. ■

Satz 8.12 Stetige Funktionen auf $[a, b]$ sind Riemann-integrierbar.

Beweis. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $\varepsilon > 0$. Da f auf $[a, b]$ gleichmäßig stetig ist (Satz 6.44), gibt es ein $\delta > 0$ so, dass $|f(t_1) - f(t_2)| < \varepsilon/(b - a)$ für alle $t_1, t_2 \in [a, b]$ mit $|t_1 - t_2| < \delta$.

Sei $Z = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$ mit $|Z| < \delta$. Dann ist

$$O(Z, f) - U(Z, f) = \sum_{i=0}^{n-1} (M_i - m_i) \Delta x_i.$$

Da stetige Funktionen auf kompakten Mengen ihr Supremum und Infimum annehmen, gibt es für jedes i Punkte $\xi'_i, \xi''_i \in [x_i, x_{i+1}]$ mit $f(\xi'_i) = M_i$ und $f(\xi''_i) = m_i$. Aus $x_i \leq \xi'_i$ und $\xi''_i \leq x_{i+1}$ folgt weiter $|\xi'_i - \xi''_i| < \delta$ und somit $f(\xi'_i) - f(\xi''_i) < \varepsilon/(b - a)$. Schließlich erhalten wir

$$O(Z, f) - U(Z, f) = \sum_{i=0}^{n-1} (f(\xi'_i) - f(\xi''_i)) \Delta x_i < \frac{\varepsilon}{b - a} \sum_{i=0}^{n-1} \Delta x_i = \varepsilon.$$

Da stetige Funktionen auf kompakten Mengen beschränkt sind, folgt die Behauptung aus dem Riemannsches Integrierbarkeitskriterium (Folgerung 8.10). ■

8.4 Das Lebesguesche Integrierbarkeitskriterium

Dieses Kriterium stellt einen Zusammenhang zwischen der Riemann-Integrierbarkeit einer Funktion und der Menge ihrer Unstetigkeiten her. Wir benötigen dafür den folgenden Begriff. Die Länge eines Intervalls I bezeichnen wir mit $|I|$.

Definition 8.13 Eine Menge $M \subseteq \mathbb{R}$ heißt Nullmenge, wenn es zu jedem $\varepsilon > 0$ höchstens abzählbar viele abgeschlossene Intervalle I_1, I_2, \dots gibt, die M überdecken (d.h. $M \subseteq \cup I_j$) und für die $\sum |I_j| < \varepsilon$ ist.

Man kann in dieser Definition “abgeschlossen” durch “offen” ersetzen.

Lemma 8.14 (a) Jede Teilmenge einer Nullmenge ist eine Nullmenge.

(b) Abzählbare Teilmengen von \mathbb{R} sind Nullmengen.

(c) Die Vereinigung abzählbar vieler Nullmengen ist eine Nullmenge.

(d) Eine kompakte Menge $M \subseteq \mathbb{R}$ ist genau dann eine Nullmenge, wenn es für jedes $\varepsilon > 0$ endlich viele abgeschlossene Intervalle I_j mit $M \subseteq \cup_j I_j$ und $\sum_j |I_j| < \varepsilon$ gibt.

Beweis. (a) ist klar.

(b) Sei $\{x_1, x_2, \dots\}$ eine abzählbare Teilmenge von \mathbb{R} und $\varepsilon > 0$. Offenbar liegt x_j im Intervall $I_j := [x_j - \varepsilon 2^{-j-2}, x_j + \varepsilon 2^{-j-2}]$. Die Vereinigung $\cup_j I_j$ überdeckt daher die Menge $\{x_1, x_2, \dots\}$, und für die Intervalllängen gilt

$$\sum_{j=1}^{\infty} |I_j| = \sum_{j=1}^{\infty} \varepsilon \cdot 2^{-j-1} = \varepsilon/2 < \varepsilon.$$

(c) Seien M_1, M_2, \dots Nullmengen und $\varepsilon > 0$. Für jede Menge M_j gibt es Intervalle I_{j1}, I_{j2}, \dots mit $\sum_k |I_{jk}| < \varepsilon \cdot 2^{-j}$, welche M_j überdecken. Dann überdecken die (nach Satz 2.8 abzählbar vielen) Intervalle $I_{11}, I_{12}, \dots, I_{21}, I_{22}, \dots, I_{31}, I_{32}, \dots$ die Vereinigung $\cup_j M_j$, und für die Intervalllängen gilt

$$\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |I_{jk}| < \sum_{j=1}^{\infty} \varepsilon \cdot 2^{-j} = \varepsilon.$$

(d) Sei $M \subseteq \mathbb{R}$ kompakt und $\varepsilon > 0$. Ist M Nullmenge, so gibt es abgeschlossene Intervalle I_1, I_2, \dots mit $M \subset \cup_j I_j$ und $\sum_j |I_j| =: \delta < \varepsilon$. Ist etwa $I_j = [a_j, b_j]$, so setzen wir $L_j := (a_j - (\varepsilon - \delta)2^{-j-2}, b_j + (\varepsilon - \delta)2^{-j-2})$. Die offenen Intervalle L_j überdecken M . Da M kompakt ist, überdecken bereits endlich viele dieser Intervalle die Menge M , etwa L_1, \dots, L_k . Dann überdecken auch die abgeschlossenen Intervalle $\overline{L}_1, \dots, \overline{L}_k$ die Menge M , und für die Intervalllängen gilt:

$$\sum_{j=1}^k |\overline{L}_j| \leq \sum_{j=1}^{\infty} |\overline{L}_j| = \sum_{j=1}^{\infty} (|I_j| + (\varepsilon - \delta)2^{-j-1}) = \delta + \frac{1}{2}(\varepsilon - \delta) < \varepsilon.$$

Die umgekehrte Aussage ist klar. ■

Sei $M \subseteq \mathbb{R}$. Eine Funktion $f : M \rightarrow \mathbb{R}$ heißt auf M *fast überall stetig*, wenn die Menge ihrer Unstetigkeitsstellen eine Nullmenge ist. Allgemein sagt man, dass eine Eigenschaft *fast überall* erfüllt ist, wenn die ‘Ausnahmepunkte’ eine Nullmenge bilden.

Satz 8.15 (Lebesguesches Integrabilitätskriterium) *Eine Funktion f von $[a, b]$ nach \mathbb{R} ist genau dann Riemann-integrierbar, wenn sie beschränkt und fast überall stetig ist.*

Für den Beweis benötigen wir genauere Aussagen über Unstetigkeitsstellen. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion und T eine nichtleere Teilmenge von $[a, b]$. Die Zahl

$$\Omega_f(T) := \sup_{t \in T} f(t) - \inf_{t \in T} f(t) = \sup\{|f(x) - f(y)| : x, y \in T\}$$

heißt *Oszillation* oder *Schwankung* von f auf T . Für jedes fixierte $x \in [a, b]$ ist die Funktion

$$(0, \infty) \rightarrow \mathbb{R}, \quad \delta \mapsto \Omega_f(U_\delta(x) \cap [a, b])$$

monoton wachsend und nach unten durch 0 beschränkt. Also existiert der Grenzwert

$$\lim_{\delta \rightarrow 0^+} \Omega_f(U_\delta(x) \cap [a, b]) =: \omega_f(x).$$

Die Zahl $\omega_f(x)$ heißt *Oszillation* von f im Punkt x .

Lemma 8.16 *Eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann stetig in $x \in [a, b]$, wenn $\omega_f(x) = 0$.*

Der Beweis benutzt nur die Definition der Stetigkeit und ist HA. ■

Wir bezeichnen die Menge der Unstetigkeitsstellen von f mit $\Delta(f)$, und für jedes $\varepsilon > 0$ setzen wir $\Delta_\varepsilon(f) := \{x \in [a, b] : \omega_f(x) \geq \varepsilon\}$. Eine unmittelbare Konsequenz von Lemma 8.16 ist

Folgerung 8.17 *Für jede beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist*

$$\Delta(f) = \bigcup_{k=1}^{\infty} \Delta_{1/k}(f).$$

Man kann leicht zeigen, dass jede der Mengen $\Delta_\varepsilon(f)$ kompakt ist.

Folgerung 8.18 *$\Delta(f)$ ist eine abzählbare Vereinigung kompakter Mengen.*

Beweis von Satz 8.15. (\Leftarrow) Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt (es gibt also ein C so, dass $|f(x)| \leq C$ für alle $x \in [a, b]$) und $\Delta(f)$ sei eine Nullmenge. Wir zeigen, dass es für jedes $\varepsilon > 0$ eine Zerlegung Z von $[a, b]$ mit $O(Z, f) - U(Z, f) < (2C + b - a)\varepsilon$ gibt. Aus dem Riemannsches Integrabilitätskriterium folgt dann die Riemann-Integrierbarkeit von f .

Wie im Beweis von Lemma 8.14 (d) sieht man, dass $\Delta(f)$ durch abzählbar viele offene Intervalle J_1, J_2, \dots überdeckt werden kann, wobei für die Längen ihrer Abschließungen gilt:

$$\sum_j |\bar{J}_j| < \varepsilon. \tag{8.7}$$

Weiter: in jedem Punkt $x \in [a, b] \setminus \Delta(f)$ ist f stetig. Wir finden daher für jedes solche x ein offenes Intervall U_x mit

$$x \in U_x \quad \text{und} \quad \sup\{f(x) : x \in \bar{U}_x\} - \inf\{f(x) : x \in \bar{U}_x\} < \varepsilon. \tag{8.8}$$

Das System aller Intervalle J_j und U_x bildet eine offene Überdeckung von $[a, b]$. Nach dem Heine-Borelschen Überdeckungssatz lassen sich endlich viele Intervalle J_{j_1}, \dots, J_{j_n} und U_{x_1}, \dots, U_{x_s} auswählen, die ebenfalls $[a, b]$ überdecken. Sei nun Z eine Zerlegung von $[a, b]$ derart, dass jedes ihrer Teilintervalle I_0, \dots, I_ℓ in einem der Intervalle $\bar{J}_{j_1}, \dots, \bar{J}_{j_n}, \bar{U}_{x_1}, \dots, \bar{U}_{x_s}$ liegt (man kann z.B. für Z die Menge aller Endpunkte der Intervalle $\bar{J}_{j_1}, \dots, \bar{U}_{x_s}$ wählen, die in $[a, b]$ liegen). Die Differenz

$$O(Z, f) - U(Z, f) = \sum_{k=0}^{\ell-1} (M_k - m_k) |I_k|$$

schreiben wir als $\Sigma_1 + \Sigma_2$, wobei Σ_1 die Summe aller $(M_k - m_k) |I_k|$ ist, für die I_k in einem der Intervalle \bar{J}_j liegt und Σ_2 die übrigen Summanden enthält. Aus (8.7) bzw. (8.8) folgt dann

$$\Sigma_1 \leq \sum_j 2C |\bar{J}_j| \stackrel{(8.7)}{<} 2C\varepsilon \quad \text{bzw.} \quad \Sigma_2 \stackrel{(8.8)}{\leq} \varepsilon(b - a).$$

Also ist tatsächlich $O(Z, f) - U(Z, f) = \Sigma_1 + \Sigma_2 < (2C + b - a)\varepsilon$.

(\Rightarrow) Sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar. Dann ist f beschränkt nach Satz 8.4. Wir zeigen, dass $\Delta(f)$ eine Nullmenge ist. Nach Folgerung 8.17 ist

$$\Delta(f) = \cup_{k=1}^{\infty} \Delta_{1/k}(f).$$

Da jede abzählbare Vereinigung von Nullmengen eine Nullmenge ist (Lemma 8.9(c)) genügt es somit zu zeigen, dass jede der Mengen $\Delta_{1/k}(f)$ eine Nullmenge ist.

Falls $\Delta_{1/k}(f) = \emptyset$, ist nichts zu beweisen. Sei also $\Delta_{1/k}(f) \neq \emptyset$. Sei $\varepsilon > 0$ beliebig. Nach dem Riemanschen Integrierbarkeitskriterium gibt es eine Zerlegung Z von $[a, b]$ mit

$$O(Z, f) - U(Z, f) < \frac{\varepsilon}{2k}. \quad (8.9)$$

Wir bezeichnen mit \mathcal{M}^* die Menge aller Teilintervalle I_m von Z , die einen Punkt aus $\Delta_{1/k}(f)$ in ihrem Inneren enthalten (beachte: \mathcal{M}^* kann leer sein).

Sei zunächst $I_m \in \mathcal{M}^*$, und $x \in \Delta_{1/k}(f)$ liege im Inneren von I_m . Dann gibt es eine δ -Umgebung $U \subseteq I_m$ von x mit $\Omega_f(U) \geq 1/k$. Erst recht ist dann

$$M_m - m_m = \Omega_f(I_m) \geq \Omega_f(U) \geq 1/k.$$

Hieraus folgt mit (8.9) sofort

$$\frac{\varepsilon}{2k} > O(Z, f) - U(Z, f) \geq \sum_{I_m \in \mathcal{M}^*} (M_m - m_m)|I_m| \geq \frac{1}{k} \sum_{I_m \in \mathcal{M}^*} |I_m|,$$

d.h. es ist

$$\sum_{I_m \in \mathcal{M}^*} |I_m| < \frac{\varepsilon}{2}. \quad (8.10)$$

Schließlich bestimmen wir zu den Teilpunkten x_0, \dots, x_n von Z abgeschlossene Intervalle I'_0, \dots, I'_n mit

$$x_k \in I'_k \quad \text{und} \quad \sum_{k=0}^n |I'_k| < \frac{\varepsilon}{2}. \quad (8.11)$$

Die Menge $\Delta_{1/k}(f)$ wird vom endlichen Intervallsystem $\{I'_0, \dots, I'_n\} \cup \mathcal{M}^*$ überdeckt. Da wegen (8.10) und (8.11) die Summe der entsprechenden Intervalllängen kleiner als $\varepsilon/2 + \varepsilon/2 = \varepsilon$ ist, ist $\Delta_{1/k}(f)$ eine Nullmenge. ■

Folgerung 8.19 *Ist $f : [a, b] \rightarrow \mathbb{R}$ beschränkt und die Menge der Unstetigkeitsstellen von f höchstens abzählbar, so ist f Riemann-integrierbar.*

Man erhält hieraus wieder die Riemann-Integrierbarkeit stetiger und auch die monotoner Funktionen (wie?). Auch ist beispielsweise die Funktion

$$f : x \mapsto \begin{cases} 0 & \text{wenn } x \text{ irrational,} \\ 1/q & \text{wenn } x = p/q \text{ mit } p, q \in \mathbb{Z}, q \neq 0, \text{ g.g.T.}(p, q) = 1 \end{cases}$$

auf jedem Intervall $[a, b]$ Riemann-integrierbar. Es ist nämlich $\Delta(f) = \mathbb{Q} \cap [a, b]$.

8.5 Eigenschaften des Riemann-Integrals

Satz 8.20 Sind $f, g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und stimmen f und g auf einer in $[a, b]$ dichten Menge M überein, so ist

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

Beweis. Sei (Z_n) eine Folge von Zerlegungen von $[a, b]$ mit $|Z_n| \rightarrow 0$. Wegen der Dichtheit von M in $[a, b]$ findet man für jedes Z_n einen Zwischenvektor $\xi^{(n)} = (\xi_0^{(n)}, \dots, \xi_{n-1}^{(n)})$ mit $\xi_i^{(n)} \in M$ für alle i und n . Dann ist aber auch $f(\xi_i^{(n)}) = g(\xi_i^{(n)})$ für alle i und n , und die entsprechenden Riemann-Summen stimmen überein:

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

Hieraus folgt die Behauptung. ■

Folgerung 8.21 Sind $f, g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und ist $N := \{x \in [a, b] : f(x) \neq g(x)\}$ eine Nullmenge, so ist

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

Der Beweis folgt sofort aus Satz 8.20. Da N Nullmenge ist, enthält N keine offenen Intervalle. Also ist $[a, b] \setminus N$ dicht in $[a, b]$, und auf dieser Menge stimmen f und g überein. ■

Satz 8.22 (a) Sei f auf $[a, b]$ Riemann-integrierbar und $a \leq a_1 < b_1 \leq b$. Dann ist f auch auf $[a_1, b_1]$ Riemann-integrierbar.

(b) Sei $a < c < b$, und f sei auf $[a, c]$ und auf $[c, b]$ Riemann-integrierbar. Dann ist f auf $[a, b]$ Riemann-integrierbar, und es gilt

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx. \quad (8.12)$$

Beweis. Beide Aussagen folgen leicht mit dem Lebesgueschen Integrabilitätskriterium. Ist etwa f auf $[a, b]$ Riemann-integrierbar, so ist die Menge der Unstetigkeitsstellen von f auf $[a, b]$ eine Nullmenge. Dann ist erst recht die Menge der Unstetigkeitsstellen auf $[a_1, b_1]$ eine Nullmenge, d.h. f ist auf $[a_1, b_1]$ Riemann-integrierbar. Aussage (b) ist HA. ■

Die Regel (8.12) gilt bei beliebiger Lage der Punkte a, b, c , wenn man für $a < b$ die folgenden Vereinbarungen trifft:

$$\int_b^a f(x) dx := - \int_a^b f(x) dx \quad \text{und} \quad \int_a^a f(x) dx := 0.$$

Satz 8.23 Sind $f, g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und sind $\alpha, \beta \in \mathbb{R}$, so ist auch $\alpha f + \beta g$ Riemann-integrierbar auf $[a, b]$, und es gilt

$$\int_a^b (\alpha f + \beta g)(x) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx.$$

Die Menge der Riemann-integrierbaren Funktionen bildet also einen linearen Raum, und die Abbildung $f \mapsto \int_a^b f(x) dx$ ist linear.

Beweis. Für jede Zerlegung Z und jeden zugehörigen Zwischenvektor ξ gilt

$$S(Z, \xi, \alpha f + \beta g) = \alpha S(Z, \xi, f) + \beta S(Z, \xi, g).$$

Aus der Definition des Riemann-Integrals folgt die Behauptung. ■

8.6 Integralungleichungen und Mittelwertsätze

Satz 8.24 Sind $f, g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und ist $f(x) \geq g(x)$ für alle $x \in [a, b]$, so ist auch

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

Für jede Riemannsumme gilt nämlich $S(Z, \xi, f) \geq S(Z, \xi, g)$. ■

Satz 8.25 (Dreiecksungleichung für Integrale) Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist auch die Funktion $|f|$ Riemann-integrierbar auf $[a, b]$, und es gilt

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx. \quad (8.13)$$

Beachten Sie die Ähnlichkeit zur bekannten Dreiecksungleichung

$$|a_1 + a_2| \leq |a_1| + |a_2| \quad \text{bzw.} \quad \left| \sum_{i=1}^n a_i \right| \leq \sum_{i=1}^n |a_i|.$$

Beweis. Die Riemann-Integrierbarkeit von $|f|$ folgt mit dem Lebesgueschen Integrabilitätskriterium: Ist f in x stetig, so ist auch $|f|$ in x stetig. Also kann $|f|$ nicht mehr Unstetigkeitsstellen als f besitzen. Die Ungleichung (8.13) erhält man aus der Dreiecksungleichung $|S(Z, \xi, f)| \leq S(Z, \xi, |f|)$ für Riemann-Summen oder mit Satz 8.24 aus $f \leq |f|$ und $-f \leq |f|$. ■

Satz 8.26 (Mittelwertsatz der Integralrechnung) Sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und $m := \inf\{f(x) : x \in [a, b]\}$, $M := \sup\{f(x) : x \in [a, b]\}$. Dann gibt es ein $\eta \in [m, M]$ mit

$$\int_a^b f(x) dx = \eta(b - a).$$

Ist f stetig, so gibt es ein $\xi \in [a, b]$ mit $\eta = f(\xi)$.

Die Zahl η ist für $a \neq b$ eindeutig bestimmt und heißt *Mittelwert* von f auf $[a, b]$ (in Analogie zum arithmetischen Mittel von Zahlen).

Beweis. Offenbar ist $m \leq f(x) \leq M$ für alle $x \in [a, b]$. Aus Satz 8.24 folgt daher

$$m(b-a) = \int_a^b m \, dx \leq \int_a^b f(x) \, dx \leq \int_a^b M \, dx = M(b-a).$$

Jede Zahl aus $[m(b-a), M(b-a)]$ lässt sich als $\eta(b-a)$ mit einem $\eta \in [m, M]$ schreiben. Die Behauptung für stetiges f folgt aus dem Zwischenwertsatz und dem Satz von Weierstraß: Mit m und M wird auch jeder Wert $\eta \in [m, M]$ von f angenommen. ■

Satz 8.27 (Erweiterter Mittelwertsatz) *Sind $f, g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist auch $fg : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar. Ist zusätzlich $g \geq 0$ auf $[a, b]$, so gibt es ein $\eta \in [m, M]$ (mit m, M wie in Satz 8.26) so, dass*

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

Ist f stetig auf $[a, b]$, so gibt es ein $\xi \in [a, b]$ mit $f(\xi) = \eta$.

Beweis. Die Riemann-Integrierbarkeit von fg zeigen Sie im Tutorium. Weiter folgt wie im Beweis von Satz 8.26 aus $m \leq f \leq M$ bzw. $mg \leq fg \leq Mg$, dass

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

Ist $\int_a^b g(x) \, dx = 0$, so folgt hieraus auch $\int_a^b f(x)g(x) \, dx = 0$, und für η kann eine beliebige Zahl aus $[m, M]$ genommen werden. Ist dagegen $\int_a^b g(x) \, dx \neq 0$, so ist

$$\eta := \left(\int_a^b f(x)g(x) \, dx \right) / \left(\int_a^b g(x) \, dx \right) \in [m, M]. \quad \blacksquare$$

Ohne Beweis (vgl. Heuser, Nr. 85.7) vermerken wir noch

Satz 8.28 (Zweiter Mittelwertsatz der Integralrechnung) *Sei f monoton und g stetig auf $[a, b]$. Dann existiert ein $\xi \in [a, b]$ mit*

$$\int_a^b f(x)g(x) \, dx = f(a) \int_a^\xi g(x) \, dx + f(b) \int_\xi^b g(x) \, dx.$$

8.7 Hauptsätze der Differential- und Integralrechnung

Diese Sätze stellen einen Zusammenhang zwischen den Begriffen “Ableitung” und “Integral” her, ermöglichen es, eine differenzierbare Funktion bis auf eine Konstante aus ihrer Ableitung zu rekonstruieren und bieten eine bequeme Möglichkeit zur Berechnung zahlreicher Riemann-Integrale.

8.7.1 Stammfunktionen

Definition 8.29 Sind $F, f : [a, b] \rightarrow \mathbb{R}$ Funktionen und ist F differenzierbar auf $[a, b]$ und $F'(x) = f(x)$ für alle $x \in [a, b]$, so heißt F eine Stammfunktion von f .

Mit Folgerung 7.14 erhält man sofort:

Satz 8.30 (a) Ist F Stammfunktion von f und $C \in \mathbb{R}$, so ist auch $F + C$ Stammfunktion von f .

(b) Je zwei Stammfunktionen einer gegebenen Funktion auf einem Intervall unterscheiden sich nur um eine Konstante.

Eine Stammfunktion F von f wird oft als *unbestimmtes Integral* von f bezeichnet, und man schreibt $F = \int f(x) dx$. Dies ist nicht sehr konsequent. Mit f ist ja z.B. auch $F + 1$ Stammfunktion und demzufolge auch $F + 1 = \int f(x) dx$. Wir wollen $\int f(x) dx$ als Bezeichnung für die Menge alle Stammfunktionen betrachten. Anstelle der etwas schwerfälligen Schreibweise

$$\int f(x) dx = \{F + C : C \in \mathbb{R}\}$$

schreibt man meist (jedoch auch nicht sehr exakt) $\int f(x) dx = F + C$.

Aus den uns bekannten Ableitungen spezieller Funktionen erhalten wir die folgenden unbestimmten Integrale (die man zusammen mit einigen anderen als “Grundintegrale” bezeichnet).

- $\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C$ auf $\begin{cases} \mathbb{R} & \text{falls } \alpha \in \mathbb{N}, \\ \mathbb{R} \setminus \{0\} & \text{falls } \alpha = -2, -3, -4, \dots \\ (0, \infty) & \text{falls } \alpha \in \mathbb{R} \setminus \{-1\}. \end{cases}$
- $\int x^{-1} dx = \ln |x| + C$ auf $\mathbb{R} \setminus \{0\}$.
- $\int e^x dx = e^x + C$ auf \mathbb{R} .
- $\int \sin x dx = -\cos x + C, \quad \int \cos x dx = \sin x + C$ auf \mathbb{R} .
- $\int \sinh x dx = \cosh x + C, \quad \int \cosh x dx = \sinh x + C$ auf \mathbb{R} .

8.7.2 Der (erste) Hauptsatz der Differential- und Integralrechnung

Satz 8.31 Die Funktion $F : [a, b] \rightarrow \mathbb{R}$ besitze auf $[a, b]$ eine Riemann-integrierbare Ableitung $f = F'$. Dann ist

$$\int_a^b f(x) dx = \int_a^b F'(x) dx = F(b) - F(a). \quad (8.14)$$

Wir können dieses Resultat auch so formulieren.

Satz 8.32 Die Funktion $f : [a, b] \rightarrow \mathbb{R}$ sei Riemann-integrierbar und besitze eine Stammfunktion F . Dann gilt (8.14) unabhängig von der Wahl von F .

Anmerkung 1. Es gibt differenzierbare Funktionen, deren Ableitung *nicht* Riemann-integrierbar ist. Beispielsweise ist für

$$F(x) = \begin{cases} x\sqrt{x} \sin \frac{1}{x} & \text{für } x > 0, \\ 0 & \text{für } x = 0 \end{cases}$$

die Ableitung

$$F'(x) = \begin{cases} \frac{3}{2}\sqrt{x} \sin \frac{1}{x} - \frac{1}{\sqrt{x}} \cos \frac{1}{x} & \text{für } x > 0, \\ 0 & \text{für } x = 0 \end{cases}$$

unbeschränkt auf $[0, 1]$. Auch gibt es Riemann-integrierbare Funktionen, die keine Stammfunktion besitzen (z.B. die Funktion $f : [-1, 1] \rightarrow \mathbb{R}$, die auf $[-1, 0)$ gleich -1 und sonst gleich 1 ist).

Anmerkung 2. Statt $F(b) - F(a)$ schreibt man oft $F(x)|_a^b$.

Beweis von Satz 8.31. Sei $F : [a, b] \rightarrow \mathbb{R}$ differenzierbar und $f := F'$ Riemann-integrierbar. Für jede Zerlegung $Z = \{x_0, \dots, x_n\}$ von $[a, b]$ ist

$$F(b) - F(a) = \sum_{i=0}^{n-1} (F(x_{i+1}) - F(x_i)).$$

Nach dem Mittelwertsatz der Differentialrechnung gibt es für jedes $i = 0, \dots, n-1$ ein $\xi_i \in (x_i, x_{i+1})$ so, dass

$$F(x_{i+1}) - F(x_i) = F'(\xi_i)(x_{i+1} - x_i) = f(\xi_i)\Delta x_i.$$

Der Vektor $\xi_Z := (\xi_0, \dots, \xi_{n-1})$ ist ein spezieller Zwischenvektor zur Zerlegung Z , und für diesen gilt

$$F(b) - F(a) = \sum_{i=0}^{n-1} f(\xi_i)\Delta x_i = S(Z, \xi_Z, f). \quad (8.15)$$

Ist nun (Z_n) eine Folge von Zerlegungen von $[a, b]$ mit $|Z_n| \rightarrow 0$, so wählen wir für jedes $n \in \mathbb{N}$ einen Zwischenvektor ξ_{Z_n} wie oben und erhalten aus (8.15)

$$F(b) - F(a) = S(Z_n, \xi_{Z_n}, f) \rightarrow \int_a^b f(x) dx. \quad \blacksquare$$

Beispiel.

$$\int_0^\pi \sin x dx = -\cos x \Big|_0^\pi = -\cos \pi - (-\cos 0) = 2.$$

8.7.3 Der zweite Hauptsatz der Differential- und Integralrechnung

Satz 8.33 Jede auf $[a, b]$ stetige Funktion f besitzt eine Stammfunktion auf $[a, b]$. Eine solche Stammfunktion ist gegeben durch

$$F(x) := \int_a^x f(t) dt, \quad x \in [a, b]. \quad (8.16)$$

Beweis. Jede stetige Funktion ist Riemann-integrierbar. Also existiert für jedes $x \in [a, b]$ das Integral (8.16), und die Funktion F ist wohldefiniert. Wir zeigen, dass F auf $[a, b]$ differenzierbar und $F' = f$ ist.

Seien $x, x+h \in [a, b]$ und $h \neq 0$. Dann ist

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Nach Satz 8.26 (Mittelwertsatz) gibt es ein $\xi \in [x, x+h]$ mit

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} f(\xi)(x+h-x) = f(\xi). \quad (8.17)$$

Halten wir x fest, so hängt ξ nur von h ab, und für $h \rightarrow 0$ strebt ξ gegen x . Da f stetig ist, strebt dann $f(\xi)$ gegen $f(x)$. Also existiert der Grenzwert von (8.17) für $h \rightarrow 0$, und es ist

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = f(x). \quad \blacksquare$$

8.8 Integrationstechniken

Der Hauptsatz der Differential- und Integralrechnung reduziert die Berechnung eines Riemann-Integrals für eine Funktion f auf die Bestimmung einer Stammfunktion für f . Wir lernen nun einige Aussagen kennen, die diese Aufgabe erleichtern. Allerdings bleibt die Bestimmung einer Stammfunktion (im Gegensatz zur „umgekehrten“ Aufgabe, der Bestimmung einer Ableitung) ein schwieriges Problem. Bereits für so einfache Funktionen wie $x \mapsto 1/\ln x$ und $x \mapsto e^{-x^2}$ (die nach Satz 8.33 eine Stammfunktion auf $(0, \infty)$ bzw. \mathbb{R} besitzen) ist es *nicht* möglich,

diese Stammfunktion mit Hilfe endlicher Ausdrücke von „elementaren“ Funktionen (wie Potenzfunktionen oder trigonometrische Funktionen) darzustellen.

Wir gewinnen die Regeln für die Bestimmung von Stammfunktionen durch „Umkehrung“ der uns bekannten Differentiationsregeln.

8.8.1 Linearität

Sind F bzw. G Stammfunktionen von f bzw. g , so ist für beliebige $\alpha, \beta \in \mathbb{R}$

$$(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g.$$

Also besitzt auch $\alpha f + \beta g$ eine Stammfunktion, und es gilt (in Kurzfassung)

$$\int (\alpha f + \beta g) dx = \alpha \int f dx + \beta \int g dx. \quad (8.18)$$

Satz 8.34 *Besitzen $f, g : [a, b] \rightarrow \mathbb{R}$ Stammfunktionen und sind $\alpha, \beta \in \mathbb{R}$, so besitzt auch $\alpha f + \beta g$ eine Stammfunktion auf $[a, b]$, und es gilt (8.18).*

8.8.2 Partielle Integration

Nach der Produktregel $(uv)' = u'v + uv'$ ist uv Stammfunktion von $u'v + uv'$. Besitzt nun eine der Funktionen $u'v$ und uv' eine Stammfunktion, dann auch die andere (Satz 8.34), und wir erhalten

$$uv = \int (u'v + uv') dx = \int u'v dx + \int uv' dx$$

bzw.

$$\int u'v dx = uv - \int uv' dx. \quad (8.19)$$

Satz 8.35 *Seien $u, v : [a, b] \rightarrow \mathbb{R}$ differenzierbar auf $[a, b]$, und uv' besitze eine Stammfunktion. Dann besitzt auch $u'v$ eine Stammfunktion auf $[a, b]$, und es gilt (8.19).*

Beispiel 1.

$$\begin{aligned} \int x \sin x dx &= x(-\cos x) - \int 1 \cdot (-\cos x) dx \\ &= -x \cos x + \int \cos x dx = \sin x - x \cos x + C. \end{aligned}$$

(Hier haben wir (8.19) mit $v(x) = x$ und $u(x) = -\cos x$ benutzt.) ■

Beispiel 2.

$$\begin{aligned}\int \ln x \, dx &= \int 1 \cdot \ln x \, dx = x \ln x - \int x \cdot \frac{1}{x} \, dx \\ &= x \ln x - \int 1 \, dx = x \ln x - x + C.\end{aligned}$$

(Hier benutzten wir (8.19) mit $u(x) = x$ und $v(x) = \ln x$.) ■

Für Riemann-Integrale erhält man mit (8.19):

Satz 8.36 Seien $u, v : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar. Dann ist

$$\int_a^b u'(x)v(x) \, dx = u(x)v(x)|_a^b - \int_a^b u(x)v'(x) \, dx. \quad (8.20)$$

Beweis. Aus der stetigen Differenzierbarkeit von u und v folgt die Stetigkeit von $u'v$ und uv' . Also besitzen beide Funktionen eine Stammfunktion, und die Riemann-Integrale in (8.20) existieren. Die Behauptung folgt nun aus (8.19) und dem Hauptsatz der Differential- und Integralrechnung. ■

8.8.3 Integration durch Substitution

Sei F eine Stammfunktion von f , und g sei differenzierbar. Ist die Verknüpfung $F \circ g$ definiert, so folgt aus der Kettenregel

$$\frac{dF(g(t))}{dt} = F'(g(t)) g'(t) = f(g(t)) g'(t).$$

Also ist $\Phi := F \circ g : t \mapsto F(g(t))$ eine Stammfunktion von $(f \circ g)g' : t \mapsto f(g(t))g'(t)$, d.h.

$$\int f(g(t)) g'(t) \, dt = \int f(x) \, dx \Big|_{x=g(t)}. \quad (8.21)$$

Satz 8.37 Die Funktion f besitze auf dem Intervall I eine Stammfunktion F , die Funktion g sei auf einem Intervall I_1 differenzierbar, und es gelte $g(I_1) \subseteq I$. Dann besitzt die Funktion $(f \circ g)g'$ auf I_1 eine Stammfunktion Φ , und es gilt (8.21) oder kurz $\Phi = F \circ g$.

Besitzt g eine Umkehrfunktion, so ist natürlich $F = \Phi \circ g^{-1}$. Ist g^{-1} darüber hinaus differenzierbar, folgt hieraus: Ist Φ Stammfunktion von $(f \circ g)g'$, so besitzt f eine Stammfunktion F , und es gilt $F = \Phi \circ g^{-1}$.

Beispiel 3. Auf \mathbb{R} suchen wir $\int \cos t \sin^2 t \, dt$. Wählt man $f(x) = x^2$ und $g(t) = \sin t$, so ist $g'(t) = \cos t$ und

$$\int \cos t \sin^2 t \, dt = \int x^2 dx \Big|_{x=\sin t} = \frac{x^3}{3} \Big|_{x=\sin t} + C = \frac{\sin^3 t}{3} + C. \quad \blacksquare$$

Beispiel 4. Für $f(x) = 1/x$ erhält man aus (8.21)

$$\int \frac{g'(t)}{g(t)} dt = \ln |g(t)| + C. \quad \blacksquare$$

Beispiel 5. Ist F Stammfunktion von f und sind $a, b \in \mathbb{R}$ mit $a \neq 0$, so gilt

$$\int f(at + b) dt = \frac{1}{a} \int f(at + b) \cdot a dt = \frac{1}{a} F(at + b) + C. \quad \blacksquare$$

Die Regel (8.21) führt ein Integral der Form $\int f(g(t)) g'(t) dt$ auf ein Integral der Form $\int f(x) dx$ zurück. Häufig möchte man den umgekehrten Weg gehen: Um $\int f(x) dx$ zu bestimmen, versucht man, die Integrationsvariable als $x = g(t)$ mit einer bijektiven differenzierbaren Funktion g zu schreiben und hofft, dass das Integral $\int f(g(t)) g'(t) dt$ ausgewertet werden kann.

Beispiel 6. Wir suchen $\int \sqrt{a^2 - x^2} dx$ auf $(-a, a)$ mit $a > 0$. Dazu substituieren wir $x := a \sin t$ mit $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$ (beachten Sie: auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ verschwindet die Ableitung $t \mapsto a \cos t$ von x nach t nicht), und wir gelangen zu

$$\begin{aligned} & \int \sqrt{a^2 - a^2 \sin^2 t} \cdot a \cos t dt \\ &= a^2 \int \cos^2 t dt = \frac{a^2}{2} \int (1 + \cos(2t)) dt \\ &= \frac{a^2}{2} \left(t + \frac{1}{2} \sin(2t) \right) + C = \frac{a^2}{2} (t + \sin t \cos t) + C \\ &= \frac{a^2}{2} (t + \sin t \sqrt{1 - \sin^2 t}) + C. \end{aligned}$$

Mit der Rücksubstitution $t = \arcsin \frac{x}{a}$ erhalten wir

$$\int \sqrt{a^2 - x^2} dx = \frac{a^2}{2} \left(\arcsin \frac{x}{a} + \frac{x}{a} \sqrt{1 - \left(\frac{x}{a}\right)^2} \right) + C. \quad \blacksquare$$

Beispiel 7. Wir suchen $\int \frac{1}{\sin x} dx$ auf $(0, \pi)$. Die Substitution $x = 2 \arctan t$ führt wegen $\frac{dx}{dt} = \frac{2}{1+t^2}$ und

$$\begin{aligned} \sin x &= 2 \sin \frac{x}{2} \cos \frac{x}{2} = 2 \tan \frac{x}{2} \cos^2 \frac{x}{2} = 2 \tan \frac{x}{2} \frac{\cos^2 \frac{x}{2}}{\sin^2 \frac{x}{2} + \cos^2 \frac{x}{2}} \\ &= 2 \tan \frac{x}{2} \frac{1}{1 + \tan^2 \frac{x}{2}} = \frac{2t}{1 + t^2} \end{aligned}$$

auf das Integral

$$\int \frac{1+t^2}{2t} \cdot \frac{2}{1+t^2} dt = \int \frac{dt}{t} = \ln |t| + C.$$

Die Rücksubstitution $t = \tan \frac{x}{2}$ liefert

$$\int \frac{1}{\sin x} dx = \ln \left| \tan \frac{x}{2} \right| + C. \quad \blacksquare$$

Für Riemann-Integrale kann man die Substitutionsregel wie folgt formulieren.

Satz 8.38 Sei f stetig und g stetig differenzierbar, und die Verkettung $f \circ g$ sei definiert. Dann ist

$$\int_{g(a)}^{g(b)} f(x) dx = \int_a^b f(g(t))g'(t) dt. \quad (8.22)$$

Beweis. Nach Voraussetzung sind f und $(f \circ g)g'$ stetig, so dass diese Funktionen entsprechend Stammfunktionen F und Φ besitzen und die Integrale in (8.22) definiert sind. Nach Satz 8.37 ist $\Phi = F \circ g + C$. Also ist nach dem Hauptsatz der Differential- und Integralrechnung

$$\int_a^b f(g(t))g'(t) dt = (F \circ g)(b) - (F \circ g)(a)$$

und

$$\int_{g(a)}^{g(b)} f(x) dx = F(g(b)) - F(g(a)). \quad \blacksquare$$

Besitzt g eine Umkehrfunktion, so kann man (8.22) schreiben als

$$\int_a^b f(x) dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(t))g'(t) dt.$$

8.9 Stammfunktionen rationaler Funktionen

Eine Funktion $P : x \mapsto a_n x^n + \dots + a_1 x + a_0$ mit $a_i \in \mathbb{R}$ und $a_n \neq 0$ heißt *Polynom* vom Grad n , und n heißt der *Grad* des Polynoms. In Zeichen: $n = \deg P$. Sind P und $Q \neq 0$ Polynome, so ist der Quotient P/Q für alle $x \in \mathbb{R}$ mit $Q(x) \neq 0$ definiert. Funktionen dieser Gestalt heißen *rational*. Für rationale Funktionen lässt sich stets eine Stammfunktion konstruktiv bestimmen. Dazu benötigen wir einige Resultate aus der Algebra, die wir ohne Beweis zitieren.

Seien P, Q Polynome mit $Q \neq 0$. Zur Bestimmung einer Stammfunktion der rationalen Funktion P/Q geht man wie folgt vor.

Schritt 1. Ist $\deg P \geq \deg Q$, so liefert eine Polynomdivision von P durch Q Polynome R und S mit

$$\frac{P}{Q} = R + \frac{S}{Q} \quad \text{mit } \deg S < \deg Q.$$

Für R kann man eine Stammfunktion angeben. Wir betrachten nur noch S/Q .

Schritt 2. Man zerlege das Nennerpolynom Q in Faktoren 1. und 2. Grades. Dass dies möglich ist, folgt aus nachstehendem Satz.

Satz 8.39 (Fundamentalsatz der Algebra) Sei $n \geq 1$. Für jedes Polynom $Q(x) = \sum_{i=0}^n q_i x^i$ mit $q_n \neq 0$ und $q_i \in \mathbb{R}$ gibt es reelle Zahlen b_i, c_i, d_i mit $b_i \neq b_j$ und $(c_i, d_i) \neq (c_j, d_j)$ für $i \neq j$ und positive natürliche Zahlen k_i, m_j, r und s so, dass

$$Q(x) = q_n \prod_{i=1}^r (x - b_i)^{k_i} \prod_{j=1}^s (x^2 + 2c_j x + d_j)^{m_j} \quad \text{für alle } x \in \mathbb{R} \quad (8.23)$$

mit $k_1 + \dots + k_r + 2(m_1 + \dots + m_s) = n$ und $d_j - c_j^2 > 0$ für alle j .

Zur Bestimmung der b_i, c_i, d_i ermittelt man die komplexen Nullstellen $\lambda_1, \dots, \lambda_n$ von Q . Dann ist $Q(x) = q_n(x - \lambda_1) \dots (x - \lambda_n)$. Die Terme $(x - \lambda)(x - \bar{\lambda})$ mit $\lambda \notin \mathbb{R}$ werden zu $(x - \lambda)(x - \bar{\lambda}) = x^2 - (\lambda + \bar{\lambda})x + |\lambda|^2$ zusammengefasst.

Einen Beweis von Satz 8.39 finden Sie in Heuser, Analysis I, Satz 69.3 und in der Vorlesung zur Funktionentheorie. Die Bestimmung der Nullstellen von Q ist oft schwierig.

Schritt 3. Ist die Zerlegung (8.23) gefunden, wählt man den Ansatz

$$\frac{S(x)}{Q(x)} = \sum_{i=1}^r \sum_{k=1}^{k_i} \frac{A_{ik}}{(x - b_i)^k} + \sum_{j=1}^s \sum_{m=1}^{m_j} \frac{B_{jm}x + C_{jm}}{(x^2 + 2c_j x + d_j)^m} \quad (8.24)$$

mit zu bestimmenden reellen Zahlen A_{ik}, B_{jm} und C_{jm} .

Satz 8.40 (Partialbruchzerlegung) Sei Q wie in (8.23) und S ein Polynom mit $\deg S < \deg Q$. Dann existieren Zahlen A_{ik}, B_{jm} und C_{jm} , so dass (8.24) gilt, und diese Zahlen sind eindeutig bestimmt.

Der Beweis ist in Heuser, Analysis I, Satz 69.5. ■

Die Zahlen A_{ik}, B_{jm} und C_{jm} können ermittelt werden, indem man (8.24) mit Q multipliziert und durch Koeffizientenvergleich ein lineares Gleichungssystem für die gesuchten Größen aufstellt.

Schritt 4. Zu allen in (8.24) vorkommenden Brüchen lassen sich Stammfunktionen durch partielle Integration und Substitution effektiv bestimmen. Einige der folgenden Regeln müssen dazu wiederholt angewandt werden. Man beachte, dass $d > c^2$. Außerdem sei $m \geq 2$.

$$\int \frac{dx}{(x - b)^k} = \begin{cases} \frac{1}{1-k}(x - b)^{1-k} & \text{falls } k > 1, \\ \ln |x - b| & \text{falls } k = 1, \end{cases}$$

$$\int \frac{dx}{x^2 + 2cx + d} = \frac{1}{\sqrt{d - c^2}} \arctan \frac{x + c}{\sqrt{d - c^2}},$$

$$\int \frac{dx}{(x^2 + 2cx + d)^m} = \frac{x + c}{2(m - 1)(d - c^2)(x^2 + 2cx + d)^{m-1}}$$

$$\begin{aligned}
& + \frac{(2m-3)}{2(m-1)(d-c^2)} \int \frac{dx}{(x^2+2cx+d)^{m-1}}, \\
\int \frac{\alpha x + \beta}{x^2+2cx+d} dx &= \frac{\alpha}{2} \ln(x^2+2cx+d) + (\beta - \alpha c) \int \frac{dx}{x^2+2cx+d}, \\
\int \frac{\alpha x + \beta}{(x^2+2cx+d)^m} dx &= \frac{-\alpha}{2(m-1)(x^2+2cx+d)^{m-1}} \\
& + (\beta - \alpha c) \int \frac{dx}{(x^2+2cx+d)^{m-1}}.
\end{aligned}$$

Beispiel. Man bestimme $\int \frac{x^4+1}{x^4-x^3-x+1} dx$.

Schritt 1: Polynomdivision.

$$\frac{x^4+1}{x^4-x^3-x+1} = 1 + \frac{x^3+x}{x^4-x^3-x+1}.$$

Schritt 2: Faktorisierung des Nennerpolynoms.

$$x^4 - x^3 - x + 1 = (x-1)(x^3-1) = (x-1)^2(x^2+x+1).$$

Schritt 3: Partialbruchzerlegung. Der Ansatz

$$\frac{x^3+x}{x^4-x^3-x+1} = \frac{A_1}{x-1} + \frac{A_2}{(x-1)^2} + \frac{Bx+C}{x^2+x+1}$$

liefert nach Multiplikation mit x^4-x^3-x+1

$$x^3+x = A_1(x-1)(x^2+x+1) + A_2(x^2+x+1) + (Bx+C)(x-1)^2$$

bzw.

$$x^3+x = (A_1+B)x^3 + (A_2-2B+C)x^2 + (A_2+B-2C)x + (A_2-A_1+C).$$

Ein Vergleich der Koeffizienten auf der linken bzw. rechten Seite ergibt das lineare Gleichungssystem

$$\begin{aligned}
\text{bei } x^3 : & \quad A_1 + B = 1 \\
\text{bei } x^2 : & \quad A_2 - 2B + C = 0 \\
\text{bei } x^1 : & \quad A_2 + B - 2C = 1 \\
\text{bei } x^0 : & \quad A_2 - A_1 + C = 0.
\end{aligned}$$

Die Lösung dieses Gleichungssystems ist $A_1 = \frac{2}{3}$, $A_2 = \frac{2}{3}$, $B = \frac{1}{3}$, $C = 0$. Die zu integrierende Funktion lautet also

$$\frac{x^4+1}{x^4-x^3-x+1} = 1 + \frac{2}{3} \frac{1}{x-1} + \frac{2}{3} \frac{1}{(x-1)^2} + \frac{1}{3} \frac{x}{x^2+x+1}.$$

Schritt 4: Integration

$$\begin{aligned}\int \frac{x^4 + 1}{x^4 - x^3 - x + 1} dx &= \int 1 dx + \frac{2}{3} \int \frac{dx}{x-1} + \frac{2}{3} \int \frac{dx}{(x-1)^2} \\ &\quad + \frac{1}{3} \int \frac{x dx}{x^2 + x + 1} \\ &= x + \frac{2}{3} \ln|x-1| - \frac{2}{3} \frac{1}{x-1} + \frac{1}{6} \ln(x^2 + x + 1) \\ &\quad - \frac{1}{3\sqrt{3}} \arctan \frac{2x+1}{\sqrt{3}} + C. \quad \blacksquare\end{aligned}$$

8.10 Uneigentliche Integrale

Bisher haben wir ausschließlich beschränkte Funktionen auf kompakten Intervallen integriert. Durch naheliegende Grenzprozesse erweitern wir nun die Definition des Riemann-Integrals zu sogenannten uneigentlichen Integralen.

8.10.1 Integrale mit unbeschränktem Integrationsintervall

Die Funktion $f : [a, \infty) \rightarrow \mathbb{R}$ sei auf jedem Intervall $[a, t]$ mit $t > a$ Riemann-integrierbar. Wenn der Grenzwert

$$\lim_{t \rightarrow \infty} \int_a^t f(x) dx \quad (8.25)$$

existiert und endlich ist, so bezeichnen wir ihn mit $\int_a^\infty f(x) dx$ und nennen ihn *uneigentliches Riemann-Integral* von f . Man sagt auch, dass f auf $[a, \infty)$ *Riemann-integrierbar* ist oder dass $\int_a^\infty f(x) dx$ *konvergiert*. Ist der Grenzwert (8.25) unendlich oder existiert er nicht, so heißt $\int_a^\infty f(x) dx$ *divergent*. Schließlich heißt $\int_a^\infty f(x) dx$ *absolut konvergent*, wenn $\int_a^\infty |f(x)| dx$ konvergiert. Wie bei Reihen folgt aus der absoluten Konvergenz die gewöhnliche (Cauchy-Kriterium).

Analoge Definitionen trifft man für

$$\int_{-\infty}^a f(x) dx = \lim_{s \rightarrow -\infty} \int_s^a f(x) dx$$

und für

$$\begin{aligned}\int_{-\infty}^\infty f(x) dx &= \int_{-\infty}^a f(x) dx + \int_a^\infty f(x) dx \\ &= \lim_{s \rightarrow -\infty} \int_s^a f(x) dx + \lim_{t \rightarrow +\infty} \int_a^t f(x) dx.\end{aligned}$$

Beispiel 1. Es ist

$$\begin{aligned} \int_1^\infty x^\alpha dx &= \lim_{t \rightarrow \infty} \int_1^t x^\alpha dx = \lim_{t \rightarrow \infty} \begin{cases} \frac{t^{\alpha+1}}{\alpha+1} - \frac{1}{\alpha+1} & \text{falls } \alpha \neq -1 \\ \ln t & \text{falls } \alpha = -1 \end{cases} \\ &= \begin{cases} \infty & \text{falls } \alpha \geq -1 \quad (\text{Divergenz}), \\ -\frac{1}{\alpha+1} & \text{falls } \alpha < -1 \quad (\text{Konvergenz}). \end{cases} \quad \blacksquare \end{aligned}$$

Beispiel 2. Wir berechnen $\int_0^\infty x^n e^{-x} dx$. Eine Stammfunktion des Integranden ist

$$F(x) = -e^{-x} \sum_{k=0}^n \frac{n!}{k!} x^k.$$

Dies kann man einfach durch Differenzieren bestätigen (HA). Wir überlegen uns, dass $\lim_{x \rightarrow \infty} F(x) = 0$. Hierfür genügt es zu zeigen, dass

$$\lim_{x \rightarrow \infty} \frac{x^k}{e^x} = 0 \quad \text{für jedes } k \geq 0. \quad (8.26)$$

Aus der Definition der Exponentialfunktion durch eine Potenzreihe folgt für jedes $k \in \mathbb{N}$ und jedes $x > 0$

$$e^x \geq \frac{x^{k+1}}{(k+1)!} \quad \text{bzw.} \quad 0 \leq \frac{x^k}{e^x} \leq \frac{(k+1)!}{x}.$$

Der Grenzübergang $x \rightarrow \infty$ in der rechten Ungleichung liefert (8.26). Zusammengefasst erhalten wir

$$\int_0^\infty x^n e^{-x} dx = \lim_{t \rightarrow \infty} \int_0^t x^n e^{-x} dx = \lim_{t \rightarrow \infty} F(t) - F(0) = -F(0) = n!. \quad \blacksquare$$

Beispiel 3. Wir zeigen, dass $\int_0^\infty \frac{\sin x}{x} dx$ konvergiert. An der Stelle 0 ist der Integrand nicht definiert. Wegen $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$ lässt sich die Funktion $x \mapsto \frac{\sin x}{x}$ aber zu einer auf $[0, \infty)$ stetigen Funktion fortsetzen, wenn man ihren Wert an der Stelle 0 durch 1 festlegt. Insbesondere existiert $\int_0^1 \frac{\sin x}{x} dx$ als gewöhnliches Riemann-Integral. Wir müssen also nur noch zeigen, dass $\int_1^\infty \frac{\sin x}{x} dx$ konvergiert.

Partielle Integration liefert für $t > 1$

$$\int_1^t \frac{\sin x}{x} dx = -\frac{\cos x}{x} \Big|_1^t - \int_1^t \frac{\cos x}{x^2} dx.$$

Offenbar existiert der Grenzwert

$$\lim_{t \rightarrow \infty} -\frac{\cos x}{x} \Big|_1^t = \lim_{t \rightarrow \infty} \left(-\frac{\cos t}{t} + \cos 1 \right) = \cos 1.$$

Es verbleibt, die Existenz des Grenzwertes $\lim_{t \rightarrow \infty} \int_1^t \frac{\cos x}{x^2} dx$ zu beweisen. Wir benutzen das Cauchy-Kriterium und schätzen für $1 \leq t_1 < t_2$ ab:

$$\left| \int_{t_1}^{t_2} \frac{\cos x}{x^2} dx \right| \leq \int_{t_1}^{t_2} \frac{|\cos x|}{x^2} dx \leq \int_{t_1}^{t_2} \frac{1}{x^2} dx = \frac{1}{t_1} - \frac{1}{t_2} < \frac{1}{t_1}.$$

Ist nun $\varepsilon > 0$ beliebig vorgegeben, so gilt für alle $t_2 > t_1 > 1/\varepsilon$

$$\left| \int_{t_1}^{t_2} \frac{\cos x}{x^2} dx \right| < \varepsilon.$$

Also existiert $\int_1^\infty \frac{\cos x}{x^2} dx$ und damit auch $\int_0^\infty \frac{\sin x}{x} dx$. Was man auf diesem Wege nicht erhält ist, dass dieses Integrals den Wert $\pi/2$ hat. ■

Wir haben oben $\int_{-\infty}^\infty f(x) dx$ definiert als $\lim_{s \rightarrow -\infty, t \rightarrow +\infty} \int_s^t f(x) dx$. In diesem Sinn existiert z.B. $\int_{-\infty}^\infty x dx$ *nicht* als uneigentliches Riemann-Integral. Es ist aber

$$\lim_{t \rightarrow \infty} \int_{-t}^t x dx = \lim_{t \rightarrow \infty} \frac{x^2}{2} \Big|_{-t}^t = 0.$$

Wir nehmen dies als Anlass zu folgender Definition.

Definition 8.41 Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ auf jedem Intervall $[-t, t]$ Riemann-integrierbar und existiert der Grenzwert $\lim_{t \rightarrow \infty} \int_{-t}^t f(x) dx$ im eigentlichen Sinn, so heißt dieser Grenzwert der Cauchyscher Hauptwert. Wir bezeichnen ihn mit

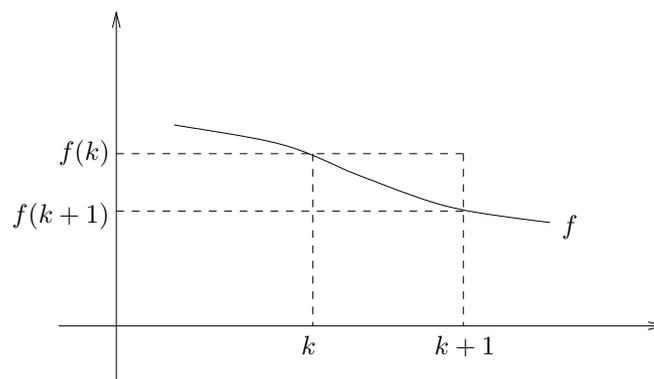
$$\text{V.P.} \int_{-\infty}^{\infty} f(x) dx.$$

Beispielsweise ist $\text{V.P.} \int_{-\infty}^{\infty} x dx = 0$.

Als eine Anwendung uneigentlicher Integrale vermerken wir das *Integralkriterium* für die Konvergenz von Reihen.

Satz 8.42 Sei $f : [1, \infty) \rightarrow [0, \infty)$ monoton fallend. Dann konvergiert die Reihe $\sum_{n=1}^\infty f(n)$ genau dann, wenn das Integral $\int_1^\infty f(x) dx$ konvergiert.

Beweis. Für jedes $k \geq 1$ ist $f(k+1) \leq \int_k^{k+1} f(x) dx \leq f(k)$.



Aufsummieren von $k = 1, \dots, n-1$ ergibt für jedes $n \geq 2$

$$f(2) + \dots + f(n) \leq \int_1^n f(x) dx \leq f(1) + \dots + f(n-1).$$

Für die Partialsumme $s_n := \sum_{k=1}^n f(k)$ gilt somit

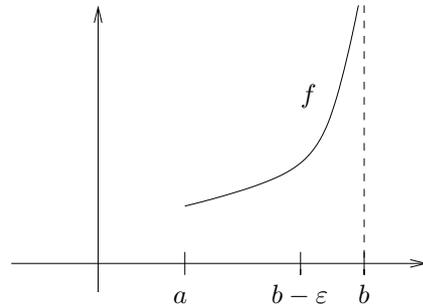
$$s_n - f(1) \leq \int_1^n f(x) dx \leq s_{n-1}.$$

Aus der linken Ungleichung folgt: Ist $\int_1^\infty f(x) dx$ konvergent, so bleiben die s_n beschränkt, also (da alle Reihenglieder nichtnegativ sind) konvergiert $\sum_{n=1}^\infty f(n)$. Analog liefert die rechte Ungleichung die umgekehrte Behauptung. ■

Beispiel 4. Aus Beispiel 1 wissen wir, daß $\int_1^\infty x^{-\alpha} dx$ für alle $\alpha > 1$ konvergiert. Nach dem Integralkriterium konvergiert dann auch $\sum_{n=1}^\infty \frac{1}{n^\alpha}$ für alle $\alpha > 1$. ■

8.10.2 Integrale mit offenem Integrationsintervall

Sei $f : [a, b) \rightarrow \mathbb{R}$ für jedes $\varepsilon \in (0, b - a)$ auf $[a, b - \varepsilon]$ Riemann-integrierbar. Wenn der Grenzwert $\lim_{\varepsilon \searrow 0} \int_a^{b-\varepsilon} f(x) dx$ existiert, so bezeichnen wir ihn mit $\int_a^b f(x) dx$ und sagen, f sei auf $[a, b)$ uneigentlich integrierbar. Eine analoge Definition trifft man für links halb-offene Intervalle.



Ist $a < c < b$ und f auf $[a, b] \setminus \{c\}$ definiert, und existieren die uneigentlichen Integrale $\int_a^c f(x) dx$ und $\int_c^b f(x) dx$, so definiert man

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx = \lim_{\varepsilon \searrow 0} \int_a^{c-\varepsilon} f(x) dx + \lim_{\delta \searrow 0} \int_{c+\delta}^b f(x) dx.$$

Schließlich definiert man in diesem Fall den Cauchyschen Hauptwert durch

$$\text{V.P.} \int_a^b f(x) dx := \lim_{\varepsilon \searrow 0} \left(\int_a^{c-\varepsilon} f(x) dx + \int_{c+\varepsilon}^b f(x) dx \right)$$

(falls dieser Grenzwert existiert und endlich ist).

Beispiel 5. $\int_0^1 \frac{1}{x^\alpha} dx$ ist für $\alpha \geq 1$ divergent und für $\alpha < 1$ konvergent. Im konvergenten Fall ist dieses Integral gleich $\frac{1}{1-\alpha}$.

Beispiel 6. $\int_0^1 \ln x dx$ konvergiert und hat den Wert -1 .

Beispiel 7. $\text{V.P.} \int_{-1}^1 \frac{1}{x} dx = \lim_{\varepsilon \searrow 0} \left(\int_{-1}^{-\varepsilon} \frac{1}{x} dx + \int_{\varepsilon}^1 \frac{1}{x} dx \right) = 0$.

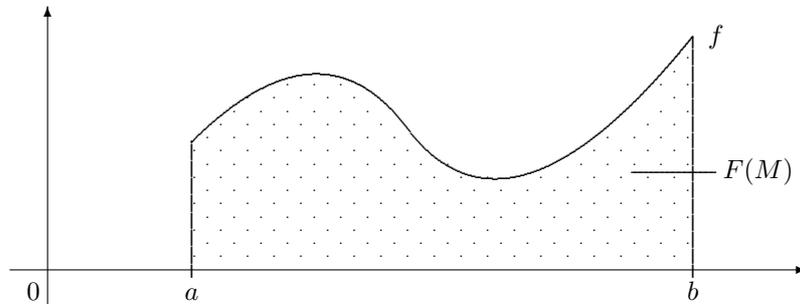
8.11 Flächeninhalte

Einer der Gründe zur Einführung des bestimmten Integrals war der Wunsch, Flächeninhalte zu definieren und zu berechnen. Ist $f : [a, b] \rightarrow [0, \infty)$ Riemann-

integrierbar, so *definieren* wir als Flächeninhalt der Menge

$$M := \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq f(x), a \leq x \leq b\}$$

die Zahl $F(M) := \int_a^b f(x) dx$.

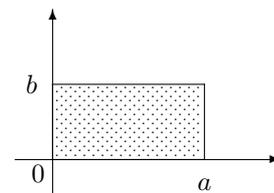


Mit dieser Definition lassen sich auch die Inhalte komplizierter Mengen definieren und berechnen, wenn man akzeptiert, dass der Flächeninhalt die folgenden (aus unserer Erfahrung heraus plausiblen) Eigenschaften aufweist:

- (a) Geht M' aus M durch Verschiebung, Drehung oder Spiegelung an einer Geraden hervor, so ist $F(M') = F(M)$.
- (b) Kann man M in zwei sich nicht überlappende Teilmengen A, B zerlegen, von denen jede einen Flächeninhalt besitzt, so ist $F(M) = F(A) + F(B)$.

Die zweite Forderung ist sehr ungenau (man sollte etwas über $A \cap B$ voraussetzen). Darauf kommen wir später zurück.

Beispiel 1. Für $f : [0, a] \rightarrow \mathbb{R}$, $x \mapsto b$ findet man $F(M) = \int_0^a b dx = ab$. Der von uns definierte Flächeninhalt stimmt also für Rechtecke mit dem „bekannten“ Flächeninhalt überein.



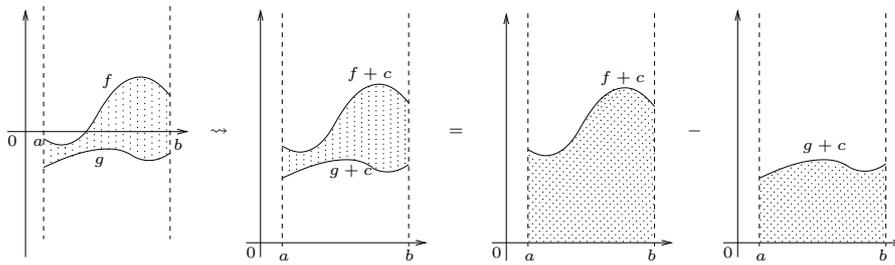
Beispiel 2. Die Dirichletfunktion

$$f(x) = \begin{cases} 1 & \text{falls } x \text{ rational,} \\ 0 & \text{falls } x \text{ irrational,} \end{cases}$$

ist auf keinem Intervall $[a, b]$ mit $a < b$ Riemann-integrierbar. Unsere Definition erlaubt es daher nicht, der Menge $M = \{(x, y) : a \leq x \leq b, 0 \leq y \leq f(x)\}$ einen Flächeninhalt zuzuschreiben. ■

Beispiel 3. Die Funktionen f, g seien auf $[a, b]$ Riemann-integrierbar, und es sei $f(x) \geq g(x)$ für alle $x \in [a, b]$. Gesucht ist der Flächeninhalt der Menge

$$M = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, g(x) \leq y \leq f(x)\}.$$



Wir verschieben M um $c > 0$ in Richtung der positiven y -Achse, bis das Bild von M komplett oberhalb der x -Achse liegt. Mit den Eigenschaften (a), (b) folgt:

$$F(M) = \int_a^b (f(x) + c) dx - \int_a^b (g(x) + c) dx = \int_a^b (f(x) - g(x)) dx. \quad \blacksquare$$

Beispiel 4. Oft ist der Graph von f in Parameterdarstellung gegeben, etwa

$$\{(x, y) \in \mathbb{R}^2 : x = x(t), y = y(t), t \in [\alpha, \beta]\}$$

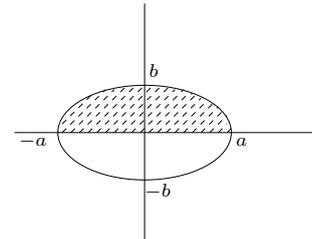
mit $x(\alpha) = a$ und $x(\beta) = b$. Unter entsprechenden Voraussetzungen an x und y (vgl. Satz 8.38) gilt dann

$$\int_a^b f(x) dx = \int_\alpha^\beta f(x(t)) \dot{x}(t) dt = \int_\alpha^\beta y(t) \dot{x}(t) dt,$$

wobei $\dot{x}(t) = \frac{dx}{dt}(t)$. Beispielsweise wird durch

$$x = a \cos t, y = b \sin t \quad \text{mit } t \in [0, 2\pi]$$

eine Ellipse beschrieben. Für ihren Flächeninhalt findet man

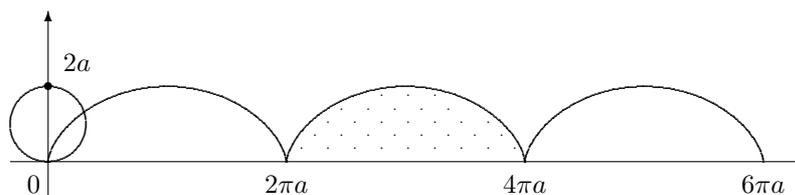


$$\begin{aligned} F(M) &= 2 \int_{-a}^a f(x) dx = 2 \int_\pi^0 y(t) \dot{x}(t) dt = 2ab \int_\pi^0 \sin t (-\sin t) dt \\ &= 2ab \int_0^\pi \sin^2 t dt = 2ab \left(-\frac{\sin t \cos t}{2} + \frac{t}{2} \right) \Big|_0^\pi = \pi ab. \quad \blacksquare \end{aligned}$$

Beispiel 5. Durch

$$x = a(t - \sin t), \quad y = a(1 - \cos t) \quad \text{mit } t \in \mathbb{R}$$

wird eine *Zykloide* definiert. Diese Kurve beschreibt den Weg eines Punktes auf der Kreisperipherie beim Abrollen des Kreises.



Für die Fläche unter einem Zykloidenbogen findet man

$$\begin{aligned} F(M) &= \int_0^{2\pi} y(t) \dot{x}(t) dt = a^2 \int_0^{2\pi} (1 - \cos t)(1 - \cos t) dt \\ &= a^2 \int_0^{2\pi} (1 - 2 \cos t + \cos^2 t) dt = a^2 \left(t - 2 \sin t + \frac{\cos t \sin t}{2} + \frac{t}{2} \right) \Big|_0^{2\pi} \\ &= 3a^2\pi. \quad \blacksquare \end{aligned}$$

9 Folgen und Reihen von Funktionen

In diesem Abschnitt betrachten wir verschiedene Arten der Konvergenz einer Funktionenfolge. Besonders interessiert uns die Frage, ob sich Eigenschaften der einzelnen Glieder einer konvergenten Funktionenfolge (f_n) auf die Grenzfunktion f vererben, etwa

- ist f stetig, wenn alle f_n stetig sind?
- ist f differenzierbar (integrierbar), wenn alle f_n differenzierbar (integrierbar) sind, und gilt in diesem Fall

$$f' = (\lim f_n)' = \lim f_n' \quad \text{bzw.} \quad \int_a^b f(x) dx = \lim \int_a^b f_n(x) dx?$$

In diesen Fragen geht es letztlich darum, ob zwei *Grenzprozesse vertauscht* werden können. Ein Beispiel, in dem dieses Vertauschen erlaubt ist, haben wir in Abschnitt 6.3 kennengelernt: Ist $\sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$, so gilt für jedes $z_0 \in \mathbb{C}$ mit $|z_0| < R$

$$\lim_{z \rightarrow z_0} \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} a_n z_0^n = \sum_{n=0}^{\infty} a_n (\lim_{z \rightarrow z_0} z)^n.$$

Wir gehen diese Probleme nun systematisch an und betrachten anschließend zwei spezielle Klassen von Funktionenreihen: Potenzreihen und Fourierreihen.

9.1 Punktweise Konvergenz

Sei X eine nichtleere Menge, N ein metrischer Raum mit einer Metrik d und für jedes $n \in \mathbb{N}$ sei eine Funktion $f_n : X \rightarrow N$ gegeben. Dann heißt $(f_n)_{n \in \mathbb{N}}$ eine *Funktionenfolge* auf X (beachten Sie: alle Glieder einer Funktionenfolge sind auf der gleichen Menge definiert und bilden in eine gleiche Menge hinein ab).

Definition 9.1 Die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ konvergiert auf X punktweise gegen eine Funktion $f : X \rightarrow N$, wenn für jedes $x \in X$ gilt

$$d(f_n(x), f(x)) \rightarrow 0 \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} f_n(x) = f(x).$$

Der punktweise Grenzwert einer Funktionenfolge (f_n) ist eindeutig bestimmt.

Beispiel 1. Die Funktionen $f_n : [0, 1] \rightarrow \mathbb{R}$ seien durch $f_n(x) = x^n$ gegeben. Für $0 \leq x < 1$ ist $\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} x^n = 0$, während $\lim_{n \rightarrow \infty} f_n(1) = 1$ ist. Die stetigen (sogar differenzierbaren) Funktionen f_n konvergieren also punktweise gegen die Funktion

$$f(x) = \begin{cases} 0 & \text{für } 0 \leq x < 1, \\ 1 & \text{für } x = 1, \end{cases}$$

die im Punkt 1 *nicht* stetig ist. ■

Beispiel 2. Für $n \in \mathbb{N}$ und $x \in \mathbb{R}$ sei

$$f_n(x) := \lim_{k \rightarrow \infty} (\cos(n! \pi x))^{2k}$$

(der Grenzwert existiert, da $0 \leq (\cos(n! \pi x))^2 \leq 1$). Wir zeigen, dass die Folge (f_n) punktweise gegen die Dirichletfunktion

$$f(x) = \begin{cases} 1 & \text{falls } x \text{ rational,} \\ 0 & \text{falls } x \text{ irrational} \end{cases}$$

konvergiert. Sei zunächst x rational. Wir schreiben x als p/q mit $p, q \in \mathbb{Z}$ und $q > 0$. Für beliebiges $n \geq q$ ist dann

$$(\cos(n! \pi p/q))^2 = 1.$$

Es ist also $f_n(x) = 1$ für alle $n \geq q$, woraus $\lim_{n \rightarrow \infty} f_n(x) = 1$ für jedes rationale x folgt. Ist dagegen x irrational, so ist $n!x$ niemals ganzzahlig. Es ist daher in diesem Fall

$$0 \leq (\cos(n! \pi x))^2 < 1,$$

woraus für jedes n folgt

$$f_n(x) = \lim_{k \rightarrow \infty} (\cos(n! \pi x))^{2k} = 0.$$

Es ist daher $\lim_{n \rightarrow \infty} f_n(x) = 0$ für jedes irrationale x .

Bildet man also von den (beliebig oft differenzierbaren!) Funktionen

$$x \mapsto (\cos(x! \pi x))^{2k}$$

erst den punktweisen Grenzwert bzgl. k und dann bzgl. n , so erhält man eine Funktion, die in keinem Punkt stetig ist! ■

Der durch die punktweise Konvergenz hergestellte Zusammenhang zwischen der Folge (f_n) und ihrer Grenzfunktion f ist offenbar zu schwach, um z.B. das Vererben der Stetigkeit zu garantieren. Im nächsten Abschnitt betrachten wir einen wesentlich stärkeren Konvergenzbegriff, der dieses Vererben garantiert.

9.2 Gleichmäßige Konvergenz

Sei wieder X eine nichtleere Menge, (N, d) ein metrischer Raum und $(f_n)_{n \in \mathbb{N}}$ eine Folge von Funktionen $f_n : X \rightarrow N$.

Definition 9.2 Die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ konvergiert auf X gleichmäßig gegen die Funktion $f : X \rightarrow N$, wenn für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass

$$d(f_n(x), f(x)) < \varepsilon \quad \text{für alle } n \geq n_0 \text{ und alle } x \in X.$$

Beachten Sie: Bei punktweiser Konvergenz darf n_0 von x abhängen, während bei gleichmäßiger Konvergenz n_0 *unabhängig* von x gefunden werden kann. Aus der gleichmäßigen Konvergenz folgt offenbar die punktweise Konvergenz.

Beispiel 1. Die Funktionenfolgen in den Beispielen 1 und 2 aus Abschnitt 9.1 konvergieren *nicht* gleichmäßig. Wir überlegen uns dies für Beispiel 1. Sei $\varepsilon > 0$ und $x \in (0, 1)$. Wegen

$$|x^n - 0| < \varepsilon \Leftrightarrow x^n < \varepsilon \Leftrightarrow n \ln x < \ln \varepsilon \Leftrightarrow n > \ln \varepsilon / \ln x$$

kann es kein n_0 so geben, dass $|f_n(x) - f(x)| < \varepsilon$ für alle $n \geq n_0$ und $x \in M$. ■

Beispiel 2. Auf \mathbb{R} sei $f_n(x) := \frac{1}{n}[nx]$ ($[y]$ ist die größte ganze Zahl, die kleiner als oder gleich y ist.) Aus $[nx] \leq nx < [nx] + 1$ folgt

$$\frac{1}{n}[nx] \leq x \leq \frac{1}{n}[nx] + \frac{1}{n} \quad \text{bzw.} \quad f_n(x) \leq x \leq f_n(x) + \frac{1}{n},$$

d.h. es ist

$$0 \leq x - f_n(x) < \frac{1}{n} \quad \text{für alle } x \in \mathbb{R}.$$

Hieraus folgt sofort die gleichmäßige Konvergenz der Funktionen f_n gegen die Funktion $f(x) = x$. ■

Im Weiteren sei $N = \mathbb{R}$, versehen mit dem üblichen Abstand. Alle Überlegungen dieses Abschnittes bleiben aber auch für $N = \mathbb{R}^k$ (z.B. mit der Euklidischen Norm) und insbesondere für $N = \mathbb{C}$ (mit dem üblichen Betragsabstand) richtig. Wir zeigen, dass man die gleichmäßige Konvergenz als Konvergenz in einem geeigneten metrischen Raum (dessen Elemente Funktionen sind) auffassen kann.

Eine Funktion $f : X \rightarrow \mathbb{R}$ heißt *beschränkt*, wenn

$$\|f\|_\infty := \sup_{x \in X} |f(x)| < \infty. \tag{9.1}$$

Die Zahl $\|f\|_\infty$ heißt die *Supremumsnorm* von f . Es ist klar, dass die Menge $M(X)$ aller beschränkten reellwertigen Funktionen auf X einen reellen linearen Raum bildet und dass (9.1) ein Norm auf $M(X)$ in folgendem Sinn definiert:

Definition 9.3 Sei L ein reeller linearer Raum. Eine Abbildung $\|\cdot\| : L \rightarrow \mathbb{R}$ heißt Norm auf L , wenn

- (a) $\|x\| \geq 0$ für alle $x \in L$, und $\|x\| = 0$ genau dann, wenn $x = 0$.
- (b) $\|\alpha x\| = |\alpha| \|x\|$ für alle $\alpha \in \mathbb{R}$ und $x \in L$.
- (c) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in L$ (Dreiecksungleichung).

Für die Norm (9.1) folgt die Eigenschaft (c) aus

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty,$$

indem man auf der linken Seite das Supremum über alle $x \in X$ bildet.

Lemma 9.4 *Ist L ein linearer Raum und $\|\cdot\|$ eine Norm auf L , so wird durch $d(x, y) := \|x - y\|$ eine Metrik auf L definiert.*

Beweis. Aus dem ersten Normaxiom folgt $d(x, y) \geq 0$ und dass $d(x, y) = 0$ genau dann, wenn $x = y$. Aus dem zweiten Normaxiom erhalten wir die Symmetrie von d :

$$d(x, y) = \|x - y\| = \|-(y - x)\| = |-1| \|y - x\| = d(y, x).$$

Die Dreiecksungleichung für d folgt unmittelbar aus dem dritten Normaxiom. ■

Man nennt $d(x, y) := \|x - y\|$ die durch die Norm $\|\cdot\|$ *induzierte Metrik*. Solange nichts anderes gesagt ist, verstehen wir normierte Räume immer mit den induzierten Metriken und machen sie so zu metrischen Räumen. Man überlegt sich leicht, dass die gleichmäßige Konvergenz einer Folge (f_n) beschränkter reellwertiger Funktionen nichts anderes ist als die Konvergenz dieser Folge im metrischen Raum $M(X)$ mit der durch die Supremumsnorm induzierten Metrik, d.h. $f_n \rightarrow f$ bedeutet

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \|f - f_n\|_\infty = \sup_{x \in X} |f(x) - f_n(x)| < \varepsilon.$$

Definition 9.5 *Ein normierter linearer Raum heißt ein Banachraum, wenn er bezüglich der durch die Norm induzierten Metrik vollständig ist.*

Einige Beispiele für Banachräume kennen wir bereits: \mathbb{R} und \mathbb{C} mit den üblichen Beträgen als Norm und \mathbb{R}^k mit der $\|\cdot\|_1$, $\|\cdot\|_2$ oder $\|\cdot\|_\infty$ -Norm.

Satz 9.6 *Der lineare Raum $M(X)$ der beschränkten reellwertigen Funktionen auf X , versehen mit der Supremumsnorm, ist ein Banachraum.*

Beweis. Sei (f_n) eine Cauchyfolge in $M(X)$, d.h.

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall m, n \geq n_0 : d(f_n, f_m) = \sup_{x \in X} |f_n(x) - f_m(x)| < \varepsilon. \quad (9.2)$$

Offenbar ist für jedes feste $x \in \mathbb{R}$ die Folge $(f_n(x))$ eine Cauchyfolge in \mathbb{R} . Da \mathbb{R} vollständig ist, konvergiert die Folge $(f_n(x))$ gegen eine Zahl, die wir $f(x)$ nennen. Hierdurch wird eine Funktion $f : X \rightarrow \mathbb{R}$ festgelegt. Wir zeigen: f ist beschränkt (d.h. $f \in M(X)$) und $d(f, f_n) = \|f - f_n\|_\infty \rightarrow 0$.

Aus (9.2) wissen wir:

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall m, n \geq n_0 \quad \forall x \in X : |f_n(x) - f_m(x)| < \varepsilon.$$

Vollziehen wir hierin den Grenzübergang $m \rightarrow \infty$, folgt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 \quad \forall x \in X : |f_n(x) - f(x)| \leq \varepsilon. \quad (9.3)$$

Wir wählen z.B. $\varepsilon = 1$ und das zugehörige n_0 und erhalten

$$|f(x)| \leq |f_{n_0}(x)| + |f_{n_0}(x) - f(x)| \leq \|f_{n_0}\|_\infty + 1 \quad \text{für alle } x \in X,$$

d.h. f ist beschränkt. Die Konvergenz von f_n gegen f bzgl. der Supremumsnorm folgt ebenfalls sofort aus (9.3). ■

Wir vermerken noch einige wichtige Konsequenzen der Vollständigkeit des Raumes $M(X)$. Diese gelten entsprechend für beliebige Banachräume.

Seien f_n Funktionen aus $M(X)$. Die *Funktionenreihe* $\sum_{k=0}^{\infty} f_k$ heißt *punktweise* bzw. *gleichmäßig konvergent*, wenn die Folge ihrer Partialsummen $s_n := \sum_{k=0}^n f_k$ punktweise bzw. gleichmäßig konvergiert. Aus der Vollständigkeit des Raumes $(M(X), \|\cdot\|_{\infty})$ folgt sofort das Cauchysche Konvergenzkriterium.

Satz 9.7 (Cauchy-Kriterium) Sei (f_n) eine Folge in $M(X)$.

(a) Die Folge (f_n) konvergiert genau dann gleichmäßig, wenn sie eine Cauchyfolge in $M(X)$ ist, d.h. wenn

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall m, n \geq n_0 : \|f_n - f_m\|_{\infty} < \varepsilon.$$

(b) Die Reihe $\sum_{k=0}^{\infty} f_k$ konvergiert genau dann gleichmäßig, wenn es für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq n_0$ und alle $r \in \mathbb{N}$ gilt

$$\left\| \sum_{k=n}^{n+r} f_k \right\|_{\infty} < \varepsilon.$$

In Analogie zu Zahlenfolgen definieren wir

Definition 9.8 Die Reihe $\sum_{n=0}^{\infty} f_n$ heißt *absolut konvergent*, wenn die Reihe $\sum_{n=0}^{\infty} \|f_n\|_{\infty}$ konvergiert.

Wie im Beweis von Satz 5.9 zeigt man:

Satz 9.9 Jede absolut konvergente Reihe in $M(X)$ konvergiert gleichmäßig.

Dieser Satz ist bemerkenswert, da man z.B. statt der gleichmäßigen Konvergenz einer *Funktionenreihe* nur die absolute Konvergenz einer *Zahlenreihe* untersuchen muß, für die wir zahlreiche Kriterien kennen.

Beispiel 3. Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^x}$ konvergiert für jedes $x > 1$ (ist also auf $(1, \infty)$ punktweise konvergent), und sie ist gleichmäßig konvergent auf jedem Intervall $[c, \infty)$ mit $c > 1$. Die Funktionen $f_n(x) = n^{-x}$ sind nämlich auf $[c, \infty)$ streng monoton fallend. Also ist $\|f_n\|_{\infty} = n^{-c}$, und für $c > 1$ konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^c}$. Die Funktion $\zeta(x) := \sum_{n=1}^{\infty} \frac{1}{n^x}$ heißt *Riemannsche Zetafunktion*. ■

Beispiel 4. Jede Potenzreihe $\sum_{n=0}^{\infty} a_n z^n$ mit Konvergenzradius $R > 0$ ist auf der Kreisscheibe $\{z \in \mathbb{C} : |z| < R\}$ punktweise konvergent. Für $0 \leq r < R$ konvergiert sie auf jeder Kreisscheibe $\{z \in \mathbb{C} : |z| \leq r\}$ sogar gleichmäßig. Für die Funktionen $f_n(z) = a_n z^n$ ist nämlich

$$\|f_n\|_{\infty} = \sup_{|z| \leq r} |a_n z^n| = |a_n| r^n,$$

und die Reihe $\sum_{n=0}^{\infty} |a_n| r^n$ konvergiert, da jede Potenzreihe im Inneren ihres Konvergenzbereichs absolut konvergiert (Satz 6.16). ■

9.3 Gleichmäßige Konvergenz und Stetigkeit

In diesem Abschnitt sei (X, d) ein metrischer Raum, und wir betrachten Funktionen $X \rightarrow \mathbb{R}$. Die Sätze 9.10 und 9.11 lassen sich problemlos auf Funktionen $X \rightarrow \mathbb{R}^k$ mit $k > 1$ übertragen.

Satz 9.10 *Die Funktionen $f_n : X \rightarrow \mathbb{R}$ sollen gleichmäßig gegen $f : X \rightarrow \mathbb{R}$ konvergieren. Ist jede der Funktionen f_n in einem Punkt $x_0 \in X$ stetig, so ist auch die Grenzfunktion f in x_0 stetig.*

Beweis. Sei $\varepsilon > 0$ beliebig. Wir wählen $N \in \mathbb{N}$ so, dass $\|f - f_N\|_\infty < \varepsilon/3$. Dann ist für alle $x \in X$

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &\leq 2\|f - f_N\|_\infty + |f_N(x) - f_N(x_0)| \\ &< 2\varepsilon/3 + |f_N(x) - f_N(x_0)|. \end{aligned}$$

Da f_N in x_0 stetig ist, finden wir eine Umgebung U von x_0 so, dass $|f_N(x) - f_N(x_0)| < \varepsilon/3$ für alle $x \in U$. Für alle $x \in U$ ist dann

$$|f(x) - f(x_0)| < 2\varepsilon/3 + |f_N(x) - f_N(x_0)| < \varepsilon.$$

Also ist f in x_0 stetig. ■

Sind insbesondere alle Funktionen f_n auf ganz X stetig, so ist auch f auf ganz X stetig. Wir interpretieren dies wieder als eine Vollständigkeitsaussage. Sei $C_b(X)$ die Menge aller beschränkten stetigen Funktionen $f : X \rightarrow \mathbb{R}$. Es ist $C_b(X) \subseteq M(X)$, und $C_b(X)$ ist ein normierter linearer Raum bzgl. der Supremumsnorm. Ist X kompakt, so ist *jede* stetige Funktion auf X beschränkt (Satz 6.40). Die Menge aller *beschränkten* stetigen Funktionen stimmt dann also überein mit der Menge $C(X)$ aller stetigen Funktionen auf X .

Satz 9.11 *Der lineare Raum $C_b(X)$, versehen mit der Supremumsnorm, ist ein Banachraum.*

Beweis. Ist (f_n) eine Cauchyfolge in $C_b(X)$, so ist (f_n) erst recht eine Cauchyfolge in $M(X)$. Da $M(X)$ vollständig ist, konvergiert die Folge (f_n) gleichmäßig gegen eine beschränkte Grenzfunktion f . Aus Satz 9.10 wissen wir, dass f stetig ist, also zu $C_b(X)$ gehört. ■

Wir können das Bewiesene auch so formulieren: $C_b(X)$ ist abgeschlossen in $M(X)$ bzgl. $\|\cdot\|_\infty$. Als eine Anwendung geben wir einen weiteren Beweis von Satz 6.19.

Beispiel. Sei $f(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$, und sei $|z_0| < R$. Dann ist f in z_0 stetig. Um dies einzusehen, wählen wir ein r in $(|z_0|, R)$. Nach Beispiel 4 aus 9.2 konvergiert die Reihe $\sum_{n=0}^{\infty} a_n z^n$ auf $X :=$

$\{z \in \mathbb{C} : |z| \leq r\}$ gleichmäßig. Da alle Partialsummen auf X stetig sind, ist f auf X und insbesondere in $z_0 \in X$ stetig. ■

Aus der Stetigkeit der Grenzfunktion folgt aber im Allgemeinen *nicht*, dass die Funktionenfolge gleichmäßig konvergiert. Es gilt jedoch:

Satz 9.12 (Dini) *Seien $f_n : [a, b] \rightarrow \mathbb{R}$ stetige Funktionen, die punktweise und monoton (d.h. für jedes $x \in [a, b]$ ist die Folge $(f_n(x))$ monoton) gegen eine stetige Funktion f konvergieren. Dann ist die Konvergenz sogar gleichmäßig.*

Einen Beweis finden Sie in Barner/Flohr, Analysis I, S. 321.

9.4 Gleichmäßige Konvergenz und Integrierbarkeit/Differenzierbarkeit

Sei $[a, b]$ ein endliches Intervall. Wir bezeichnen mit $R([a, b])$ die Menge der Riemann-integrierbaren Funktionen auf $[a, b]$. Da Riemann-integrierbare Funktionen beschränkt sind (Satz 8.4), können wir $R([a, b])$ als Teilraum von $M([a, b])$ auffassen.

Satz 9.13 (a) *Die Funktionen $f_n \in R([a, b])$ sollen gleichmäßig gegen eine Funktion f konvergieren. Dann ist $f \in R([a, b])$, und es gilt*

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

(b) *Der Raum $R([a, b])$, versehen mit der Supremumsnorm, ist ein Banachraum.*

Beweis. (a) Wir zeigen zuerst, dass f Riemann-integrierbar ist. Sei $\Delta(f_n)$ die Menge aller Unstetigkeitsstellen von f_n und $U := \cup_{n \in \mathbb{N}} \Delta(f_n)$. In jedem Punkt x aus $[a, b] \setminus U$ ist jede der Funktionen f_n stetig. Nach Satz 9.10 ist dann auch f in allen diesen Punkten stetig. Also ist $\Delta(f) \subseteq U$. Nach dem Lebesgueschen Integrabilitätskriterium (Satz 8.15) ist jedes $\Delta(f_n)$ eine Nullmenge. Nach Lemma 8.14 sind auch U und $\Delta(f)$ Nullmengen. Wieder nach Satz 8.15 ist $f \in R([a, b])$. Damit ist klar, dass $\int_a^b f(x) dx$ existiert, und wir haben die Abschätzung

$$\left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| \leq \int_a^b |f_n(x) - f(x)| dx \leq (b-a) \|f_n - f\|_\infty.$$

Aus der gleichmäßigen Konvergenz von f_n gegen f folgt nun Behauptung (a).

(b) Sei (f_n) eine Cauchyfolge Riemann-integrierbarer Funktionen. Dann ist (f_n) eine Cauchyfolge in $M([a, b])$ und folglich konvergent mit einer Grenzfunktion f . Aus Teil (a) wissen wir, dass $f \in R([a, b])$. Schließlich folgt aus Satz 8.23, dass $R([a, b])$ ein linearer Raum ist. ■

Es ist bemerkenswert, dass die gleichmäßige Konvergenz einer Folge differenzierbarer Funktionen *nicht* genügt, um die Differenzierbarkeit der Grenzfunktion und die Vertauschbarkeit zu erzwingen.

Beispiel. Die Folge der Funktionen $f_n(x) := \frac{1}{n} \sin nx$ konvergiert auf \mathbb{R} gleichmäßig gegen die Funktion $f = 0$. Die Folge der Ableitungen $f'_n(x) = \cos nx$ konvergiert an der Stelle $x = 0$ gegen 1. Es ist aber

$$1 = \lim f'_n(0) \neq (\lim f_n)'(0) = f'(0) = 0. \quad \blacksquare$$

Satz 9.14 Für die Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$ gelte

- (a) sie sind auf $[a, b]$ differenzierbar,
- (b) die Folge (f'_n) ihrer Ableitungen konvergiert gleichmäßig auf $[a, b]$,
- (c) es gibt ein $x_0 \in [a, b]$, für das die Folge $(f_n(x_0))$ konvergiert.

Dann konvergiert die Folge (f_n) gleichmäßig gegen eine differenzierbare Funktion f , und die Folge (f'_n) konvergiert gleichmäßig gegen f' .

Unter den getroffenen Annahmen dürfen Funktionenfolgen also gliedweise differenziert werden, und es gilt (im Sinne der gleichmäßigen Konvergenz)

$$\lim_{n \rightarrow \infty} f'_n = (\lim_{n \rightarrow \infty} f_n)'.$$

Beweis. Wir zeigen, dass die Folge (f_n) gleichmäßig gegen eine Funktion f konvergiert. Sei $\varepsilon > 0$. Wir wählen $n_0 \in \mathbb{N}$ so, dass

$$|f_m(x_0) - f_n(x_0)| < \varepsilon/2 \quad \text{für alle } m, n \geq n_0 \quad (9.4)$$

und

$$\|f'_m - f'_n\|_\infty < \frac{\varepsilon}{2(b-a)} \quad \text{für alle } m, n \geq n_0. \quad (9.5)$$

Der Mittelwertsatz der Differentialrechnung, angewandt auf die Funktion $f_m - f_n$, liefert die für alle $x, y \in [a, b]$ und $m, n \geq n_0$ gültige Abschätzung

$$\begin{aligned} |(f_m(x) - f_n(x)) - (f_m(y) - f_n(y))| &= |f'_m(\xi) - f'_n(\xi)| |x - y| \\ &\leq \frac{\varepsilon}{2(b-a)} |x - y| \end{aligned} \quad (9.6)$$

mit einem $\xi \in [x, y]$. Wir benutzen (9.6) mit $y = x_0$ und erhalten

$$\begin{aligned} |f_m(x) - f_n(x)| &\leq |(f_m(x) - f_n(x)) - (f_m(x_0) - f_n(x_0))| + |f_m(x_0) - f_n(x_0)| \\ &\leq \frac{\varepsilon}{2(b-a)} |x - x_0| + \frac{\varepsilon}{2} < \varepsilon \end{aligned}$$

für alle $x \in [a, b]$ und $m, n \geq n_0$. Also ist (f_n) eine Cauchyfolge in $C([a, b])$ und nach Satz 9.6 gleichmäßig konvergent. Sei f ihr Grenzwert. Wir zeigen die

Differenzierbarkeit von f in jedem Punkt $y \in [a, b]$. Dazu betrachten wir auf $[a, b]$ die Funktionen

$$F_n(x) := \begin{cases} \frac{f_n(x) - f_n(y)}{x - y}, & \text{falls } x \neq y, \\ f'_n(y), & \text{falls } x = y. \end{cases} \quad F(x) := \begin{cases} \frac{f(x) - f(y)}{x - y} & \text{falls } x \neq y, \\ \lim_{n \rightarrow \infty} f'_n(y) & \text{falls } x = y. \end{cases}$$

Im Beweis von Satz 7.2 haben wir gesehen, dass die Funktionen F_n stetig auf $[a, b]$ sind. Weiter ist klar, dass die Funktionen F_n auf $[a, b]$ *punktweise* gegen F streben. Wir zeigen, dass diese Konvergenz sogar gleichmäßig ist. Für $x \neq y$ erhalten wir aus (9.6) nach Division durch $|x - y|$

$$|F_m(x) - F_n(x)| < \frac{\varepsilon}{2(b-a)} \quad \forall m, n \geq n_0$$

und somit

$$\sup_{x \in [a, b] \setminus \{y\}} |F_m(x) - F_n(x)| \leq \frac{\varepsilon}{2(b-a)} \quad \forall m, n \geq n_0.$$

Wir wählen n_0 so groß, dass auch noch

$$|F_m(y) - F_n(y)| = |f'_m(y) - f'_n(y)| \leq \frac{\varepsilon}{2(b-a)}$$

für alle $m, n \geq n_0$. Dann ist

$$\sup_{x \in [a, b]} |F_m(x) - F_n(x)| \leq \frac{\varepsilon}{2(b-a)} \quad \forall m, n \geq n_0.$$

Also ist (F_n) eine Cauchyfolge in $M([a, b])$. Nach Satz 9.6 konvergiert diese Folge auf $[a, b]$ gleichmäßig. Da sie auf $[a, b]$ punktweise gegen F konvergiert, folgt, dass (F_n) auf ganz $[a, b]$ gleichmäßig gegen F konvergiert. Nach Satz 9.10 ist F auf $[a, b]$ und insbesondere in y stetig. Es existiert also der Grenzwert $\lim_{x \rightarrow y} F(x)$ und stimmt mit $F(y)$ überein. Dann ist die Funktion f an der Stelle y differenzierbar, und es gilt

$$f'(y) = \lim_{n \rightarrow \infty} f'_n(y). \quad \blacksquare$$

9.5 Ergänzungen zu Potenzreihen

Wir wenden zunächst das Resultat aus Satz 9.14 auf Potenzreihen an. Sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Die Partialsummen $s_n(x) = \sum_{k=0}^n a_k x^k$ sind offenbar differenzierbar auf $(-R, R)$, und ihre Ableitung ist

$$s'_n(x) = \sum_{k=1}^n k a_k x^{k-1}.$$

Die s'_n sind die Partialsummen der Potenzreihe $\sum_{n=1}^{\infty} n a_n x^{n-1}$. Aus Folgerung 6.18 wissen wir, dass diese Potenzreihe den gleichen Konvergenzradius wie die

Ausgangsreihe besitzt. Also ist (Beispiel 4 aus 9.2) die Folge (s'_n) auf jeder kompakten Teilmenge von $\{x \in \mathbb{R} : |x| < R\}$ gleichmäßig konvergent. Aus Satz 9.14 folgt nun:

Satz 9.15 *Die Potenzreihe $f(x) = \sum_{n=0}^{\infty} a_n x^n$ ist in jedem Punkt $x \in (-R, R)$ differenzierbar, und ihre Ableitung kann gliedweise bestimmt werden:*

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

Im dritten Semester übertragen wir dieses Resultat auf Potenzreihen im Komplexen. Eine wiederholte Anwendung von Satz 9.15 zeigt, dass Potenzreihen unendlich oft differenzierbar sind. Damit haben wir auch Satz 7.11 bewiesen.

Wir vermerken noch, dass man wegen Satz 9.13 Potenzreihen auch gliedweise integrieren darf. Zum Beispiel ist

$$\int_0^y \left(\sum_{n=0}^{\infty} a_n x^n \right) dx = \sum_{n=0}^{\infty} \int_0^y a_n x^n dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} y^{n+1}$$

für alle y im Konvergenzintervall $(-R, R)$.

Das nächste Resultat besagt, dass die Werte einer Potenzreihe im Inneren ihres Konvergenzkreises eindeutig festgelegt sind, wenn man ihre Werte nur in einer Folge von Punkten z_n mit $\lim_{n \rightarrow \infty} z_n = 0$ kennt. Genauer:

Satz 9.16 (Identitätssatz für Potenzreihen) *Seien*

$$f(z) = \sum_{n=0}^{\infty} f_n (z - z_0)^n \quad \text{und} \quad g(z) = \sum_{n=0}^{\infty} g_n (z - z_0)^n$$

Potenzreihen, die auf einer Kreisscheibe $K := \{z \in \mathbb{C} : |z - z_0| < R\}$ mit $R > 0$ konvergieren, und sei $(z_n)_{n=1}^{\infty} \subseteq K \setminus \{z_0\}$ eine Folge mit Grenzwert z_0 . Ist $f(z_n) = g(z_n)$ für alle $n \geq 1$, so stimmen beide Funktionen bzw. beide Potenzreihen auf K überein, d.h. es ist

$$f(z) = g(z) \quad \forall z \in K \quad \text{und} \quad f_n = g_n \quad \forall n \in \mathbb{N}.$$

Beweis. Nach Satz 6.19 sind f und g stetig auf K . Also gilt

$$f_0 = f(z_0) = \lim_{n \rightarrow \infty} f(z_n) = \lim_{n \rightarrow \infty} g(z_n) = g(z_0) = g_0.$$

Wir nehmen dies als Induktionsanfang. In der Induktionsvoraussetzung nehmen wir an, wir hätten für ein gewisses $n \in \mathbb{N}$ bereits gezeigt, dass

$$f_0 = g_0, \quad f_1 = g_1, \quad \dots, \quad f_n = g_n$$

und wollen zeigen, dass dann auch $f_{n+1} = g_{n+1}$. Dazu genügt es, die oben gemachten Überlegungen auf die Potenzreihen

$$\tilde{f}(z) := \frac{f(z) - \sum_{k=0}^n f_k(z - z_0)^k}{(z - z_0)^{n+1}} = f_{n+1} + f_{n+2}(z - z_0) + f_{n+3}(z - z_0)^2 + \dots$$

und

$$\tilde{g}(z) := \frac{g(z) - \sum_{k=0}^n g_k(z - z_0)^k}{(z - z_0)^{n+1}} = g_{n+1} + g_{n+2}(z - z_0) + g_{n+3}(z - z_0)^2 + \dots$$

anzuwenden. (Diese sind durch die Brüche nur für $z \neq z_0$, durch die rechten Seiten aber auch für $z = z_0$ definiert.) ■

Der Identitätssatz ist Grundlage des *Koeffizientenvergleichs*: Hat man ein und dieselbe Funktion in zwei Potenzreihen $\sum a_n(z - z_0)^n$ und $\sum b_n(z - z_0)^n$ mit positivem Konvergenzradius entwickelt, so folgt $a_n = b_n$ für alle $n \in \mathbb{N}$.

Beispiel. Ist

$$f(z) := \sum_{n=0}^{\infty} a_n z^n$$

eine gerade Funktion, d.h. $f(z) = f(-z)$ für alle z , so folgt aus

$$f(z) = f(-z) = \sum_{n=0}^{\infty} a_n (-z)^n = \sum_{n=0}^{\infty} (-1)^n a_n z^n,$$

dass $a_n = (-1)^n a_n$ für alle n und daher $a_n = 0$ für alle ungeraden n ist. ■

Das dritte Resultat dieses Abschnittes betrifft die Division von Potenzreihen.

Satz 9.17 Die Potenzreihe $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ konvergiere für $|z - z_0| < R$, und es sei $a_0 \neq 0$. Dann lässt sich $1/f$ in einer gewissen Umgebung von z_0 in eine konvergente Potenzreihe entwickeln.

Sehen wir uns die Potenzreihenentwicklung von $1/f$ zunächst formal und für $z_0 = 0$ an. Aus $f \cdot f^{-1} = 1$ bzw. $(\sum_{n=0}^{\infty} a_n z^n)(\sum_{n=0}^{\infty} b_n z^n) = 1$ folgt gemäß dem Cauchy-Produkt

$$\sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k b_{n-k} \right) z^n = 1 = 1 + 0z^1 + 0z^2 + \dots$$

Durch Koeffizientenvergleich folgen hieraus die Gleichungen

$$\begin{aligned} a_0 b_0 &= 1 && \text{(bei } z^0) \\ a_1 b_0 + a_0 b_1 &= 0 && \text{(bei } z^1) \\ a_2 b_0 + a_1 b_1 + a_0 b_2 &= 0 && \text{(bei } z^2) \quad \text{u.s.w.} \end{aligned}$$

Wegen $a_0 \neq 0$ kann man aus der ersten Gleichung b_0 ermitteln, dann aus der zweiten b_1 , aus der dritten b_2 usw. Hierdurch wird eine Potenzreihe $\sum_{n=0}^{\infty} b_n z^n$ eindeutig festgelegt. Satz 9.17 sagt aus, dass diese Reihe einen positiven Konvergenzradius besitzt.

Beispiel 1. Nach Satz 9.17 läßt sich die Tangensfunktion in einer Umgebung des Punktes 0 in eine Potenzreihe entwickeln. Wir bestimmen die ersten Koeffizienten der Potenzreihe $\tan x := \sum_{n=0}^{\infty} a_n x^n$ mit der Methode des Koeffizientenvergleichs aus $\tan x := \sin x / \cos x$ und den bekannten Potenzreihen für die Sinus- und Kosinusfunktion. Koeffizientenvergleich in

$$\begin{aligned} x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \\ = (a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + \dots) \cdot (1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots) \end{aligned}$$

liefert

$$\begin{aligned} \text{bei } x^0 : \quad 0 &= a_0 \\ \text{bei } x^1 : \quad 1 &= a_1 \\ \text{bei } x^2 : \quad 0 &= a_2 - \frac{1}{2}a_0 \\ \text{bei } x^3 : \quad -\frac{1}{6} &= a_3 - \frac{1}{2}a_1 \\ \text{bei } x^4 : \quad 0 &= a_4 - \frac{1}{2}a_2 + \frac{1}{24}a_0 \end{aligned}$$

und hieraus der Reihe nach $a_0 = 0$, $a_1 = 1$, $a_2 = 0$, $a_3 = \frac{1}{3}$ und $a_4 = 0$. Weitere Rechnungen liefern

$$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{17}{315}x^7 + \dots$$

Es gibt keine „einfache“ Formel für die Koeffizienten dieser Reihe. ■

Beispiel 2. Wir suchen eine explizite Formel für die Glieder der *Fibonacci-Folge*

$$a_0 = 1, \quad a_1 = 1, \quad a_n = a_{n-1} + a_{n-2} \quad \text{für } n \geq 2. \quad (9.7)$$

Wir ordnen dieser Folge die Potenzreihe $f(x) = \sum_{n=0}^{\infty} a_n x^n$ zu und finden mit (9.7)

$$\begin{aligned} f(x) &= 1 + x + \sum_{n=2}^{\infty} a_n x^n = 1 + x + \sum_{n=2}^{\infty} (a_{n-2} + a_{n-1}) x^n \\ &= 1 + x + x^2 \sum_{n=0}^{\infty} a_n x^n + x \sum_{n=1}^{\infty} a_n x^n \\ &= 1 + x + x^2 f(x) + x(f(x) - 1). \end{aligned}$$

Umstellen nach $f(x)$ liefert $f(x) = (1 - x - x^2)^{-1}$. Da die Potenzreihe $1 - x - x^2$ (die ein Polynom ist) überall konvergiert und an der Stelle 0 ungleich 0 ist, folgt mit Satz 9.17, dass auch die Reihe

$$f(x) = \frac{1}{1 - x - x^2} = \sum_{n=0}^{\infty} a_n x^n$$

einen positiven Konvergenzradius besitzt. Wir bestimmen die a_n , indem wir $f(x) = (1 - x - x^2)^{-1}$ geschickt in eine Potenzreihe entwickeln. Dazu schreiben wir

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{5}} \frac{1}{x - \frac{-1-\sqrt{5}}{2}} - \frac{1}{\sqrt{5}} \frac{1}{x - \frac{-1+\sqrt{5}}{2}} && \text{(Partialbruchentwicklung)} \\ &= \frac{1}{\sqrt{5}} \frac{1}{\frac{-1+\sqrt{5}}{2} - x} - \frac{1}{\sqrt{5}} \frac{1}{\frac{-1-\sqrt{5}}{2} - x} \\ &= \frac{2}{\sqrt{5}(-1 + \sqrt{5})} \frac{1}{1 - \frac{2x}{-1+\sqrt{5}}} - \frac{2}{\sqrt{5}(-1 - \sqrt{5})} \frac{1}{1 - \frac{2x}{-1-\sqrt{5}}}. \end{aligned}$$

Ist $|x|$ hinreichend klein, so ist dies gleich (geometrische Reihe!)

$$\begin{aligned} f(x) &= \frac{2}{\sqrt{5}(-1 + \sqrt{5})} \sum_{n=0}^{\infty} \left(\frac{2x}{-1 + \sqrt{5}} \right)^n - \frac{2}{\sqrt{5}(-1 - \sqrt{5})} \sum_{n=0}^{\infty} \left(\frac{2x}{-1 - \sqrt{5}} \right)^n \\ &= \frac{1}{\sqrt{5}} \sum_{n=0}^{\infty} \left(\left(\frac{2}{-1 + \sqrt{5}} \right)^{n+1} - \left(\frac{2}{-1 - \sqrt{5}} \right)^{n+1} \right) x^n, \end{aligned}$$

woraus nach Koeffizientenvergleich folgt

$$\begin{aligned} a_n &= \frac{1}{\sqrt{5}} \left(\left(\frac{2}{-1 + \sqrt{5}} \right)^{n+1} - \left(\frac{2}{-1 - \sqrt{5}} \right)^{n+1} \right) \\ &= \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{n+1} \right). \quad \blacksquare \end{aligned}$$

Wir werden Satz 9.17 im 3. Semester beweisen. Für Interessenten folgt hier ein Beweis, der nur reelle Methoden benutzt. Dafür benötigen wir einige Vorbereitungen. Eine Abbildung $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{C}$ heißt auch eine *Doppelfolge*. Wie bei gewöhnlichen Folgen identifiziert man f häufig mit ihren Werten $f(m, n) =: a_{mn}$ und schreibt $(a_{mn})_{m,n=0}^{\infty}$ für die Doppelfolge. Konvergenz der Doppelfolge $(a_{mn})_{m,n=0}^{\infty}$ gegen $a \in \mathbb{C}$ bedeutet nach Definition

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall m, n \geq n_0 : |a_{mn} - a| < \varepsilon. \quad (9.8)$$

Beispiel 3. Für jede Folge (a_n) wird durch $a_{mn} := a_m - a_n$ eine Doppelfolge festgelegt. Die Folge (a_n) ist genau dann eine Cauchyfolge, wenn die zugeordnete Doppelfolge (a_{mn}) gegen 0 konvergiert. ■

Beispiel 4. Die Doppelfolge (a_{mn}) mit $a_{mn} = (-1)^{m+n}(\frac{1}{m} + \frac{1}{n})$ konvergiert gegen 0. Für $n_0 \geq 2/\varepsilon$ und $m, n \geq n_0$ ist nämlich

$$|a_{mn} - 0| = \left| \frac{1}{m} + \frac{1}{n} \right| \leq \frac{2}{n_0} \leq \varepsilon. \quad \blacksquare$$

Für den Grenzwert a einer konvergenten Doppelfolge (a_{mn}) bezeichnen wir mit $\lim_{m,n \rightarrow \infty} a_{mn} = a$. Wenn für jedes m die Grenzwerte $\lim_{n \rightarrow \infty} a_{mn}$ bzw. für jedes n die Grenzwerte $\lim_{m \rightarrow \infty} a_{mn}$ existieren, kann man auch die *iterierten* Grenzwerte

$$\lim_{m \rightarrow \infty} (\lim_{n \rightarrow \infty} a_{mn}) \quad \text{und} \quad \lim_{n \rightarrow \infty} (\lim_{m \rightarrow \infty} a_{mn}) \quad (9.9)$$

betrachten. Man beachte, dass im Beispiel 4 zwar der Grenzwert $\lim_{m,n \rightarrow \infty} a_{mn}$ existiert, nicht aber die Grenzwerte $\lim_{m \rightarrow \infty} a_{mn}$ und $\lim_{n \rightarrow \infty} a_{mn}$.

Satz 9.18 *Die Doppelfolge (a_{mn}) sei konvergent, und für jedes m bzw. n sollen die Grenzwerte $\lim_{n \rightarrow \infty} a_{mn}$ und $\lim_{m \rightarrow \infty} a_{mn}$ existieren. Dann existieren auch die iterierten Grenzwerte (9.9), und es gilt:*

$$\lim_{m \rightarrow \infty} (\lim_{n \rightarrow \infty} a_{mn}) = \lim_{n \rightarrow \infty} (\lim_{m \rightarrow \infty} a_{mn}) = \lim_{m,n \rightarrow \infty} a_{mn}.$$

Beweis. Sei $a := \lim_{m,n \rightarrow \infty} a_{mn}$, und für jedes m existiere der Grenzwert $\alpha_m := \lim_{n \rightarrow \infty} a_{mn}$. Konvergenz der Doppelfolge (a_{mn}) bedeutet gerade (9.8). Lassen wir in (9.8) $n \rightarrow \infty$ streben, folgt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall m \geq n_0 : |\alpha_m - a| \leq \varepsilon.$$

Dies heißt aber nichts anderes als dass $\lim_{m \rightarrow \infty} \alpha_m = a$. Die zweite Aussage folgt analog. ■

Jede Doppelfolge (a_{mn}) erzeugt eine Doppelfolge (s_{mn}) durch

$$s_{mn} = \sum_{j=0}^m \sum_{k=0}^n a_{jk}.$$

Wenn die Doppelfolge (s_{mn}) gegen ein s konvergiert, so nennen wir die *Doppelreihe* $\sum_{j,k=0}^{\infty} a_{jk}$ konvergent und s ihre Summe, und wir schreiben $s = \sum_{j,k=0}^{\infty} a_{jk}$. Durch Übertragung von Satz 9.18 erhält man sofort das folgende Resultat.

Satz 9.19 *Die Doppelreihe $\sum_{j,k=0}^{\infty} a_{jk}$ sei konvergent, und für jedes j bzw. k sollen die Reihen $\sum_{k=0}^{\infty} a_{jk}$ bzw. $\sum_{j=0}^{\infty} a_{jk}$ konvergieren. Dann konvergieren auch die iterierten Reihen $\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{jk}$ bzw. $\sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_{jk}$, und es gilt*

$$\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{jk} = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_{jk} = \sum_{j,k=0}^{\infty} a_{jk}. \quad (9.10)$$

Hieraus erhält man leicht den folgenden wichtigen *Doppelreihensatz von Cauchy*.

Satz 9.20 *Ist eine der iterierten Reihen $\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{jk}$ bzw. $\sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_{jk}$ absolut konvergent (d.h. konvergiert sie auch noch, wenn a_{jk} durch $|a_{jk}|$ ersetzt wird), dann sind auch die andere iterierte Reihe sowie die Doppelreihe $\sum_{j,k=0}^{\infty} a_{jk}$ absolut konvergent, und es gilt (9.10).*

Beweis. Sei z.B. $\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |a_{jk}|$ konvergent mit der Summe a . Dann ist jede Reihe $\sum_{k=0}^{\infty} a_{jk}$ absolut konvergent. Wegen $\sum_{j=0}^m \sum_{k=0}^n |a_{jk}| \leq a$ konvergiert auch die Doppelreihe absolut. Gleiches gilt wegen $\sum_{j=0}^m |a_{jk}| \leq a$ für jede Reihe $\sum_{j=0}^{\infty} |a_{jk}|$. Aus Satz 9.19 folgt nun die Behauptung. ■

Beweis von Satz 9.17. Ohne Einschränkung der Allgemeinheit sei $a_0 = 1$. Nach Satz 6.19 gibt es ein $\delta \in (0, R)$, so dass $|a_1 z + a_2 z^2 + \dots| < 1$ für alle $|z| < \delta$. Für diese z ist (geometrische Reihe!)

$$\frac{1}{f(z)} = \frac{1}{1 - (-a_1 z - a_2 z^2 - \dots)} = \sum_{j=0}^{\infty} (-a_1 z - a_2 z^2 - \dots)^j.$$

Mit dem Cauchyprodukt schreiben wir für jedes $j \in \mathbb{N}$

$$(-a_1 z - a_2 z^2 - \dots)^j = \sum_{k=0}^{\infty} a_{jk} z^k$$

mit gewissen Koeffizienten a_{jk} . Es ist also

$$\frac{1}{f(z)} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} a_{jk} z^k.$$

Dürften wir hier die Summationsreihenfolge vertauschen, wäre dies die Behauptung. Wenn die Reihe $\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |a_{jk}| |z|^k$ konvergiert, ist nach Satz 9.20 das Vertauschen möglich. Wir zeigen, dass dies für hinreichend kleine $|z|$ tatsächlich gilt.

Da jede Potenzreihe im Innern ihres Konvergenzkreises absolut konvergiert, gibt es ein $\rho \in (0, \delta)$ so, dass

$$|a_1| |z| + |a_2| |z|^2 + \dots < 1 \quad \text{für alle } z \text{ mit } |z| < \rho.$$

Für diese z konvergiert die geometrische Reihe $\sum_{j=0}^{\infty} (|a_1| |z| + |a_2| |z|^2 + \dots)^j$, und diese lässt sich nach Cauchymultiplikation in der Form $\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \alpha_{jk} |z|^k$ schreiben. Offenbar gilt dabei $|a_{jk}| \leq \alpha_{jk}$. Folglich konvergiert die iterierte Reihe $\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |a_{jk}| |z|^k$ für alle z mit $|z| < \rho$. Hieraus und aus dem Doppelreihensatz folgt die Behauptung. ■

9.6 Fourierreihen

Nach den Potenzreihen betrachten wir eine weitere Klasse von Funktionenreihen, die von großer Bedeutung in der Analysis sowie für ihre Anwendungen in Physik und Technik ist: die Fourierreihen. Wir haben im letzten Abschnitt gesehen, dass Potenzreihen im Inneren ihres Konvergenzintervalls beliebig oft differenzierbar sind. Es können also nur „wenige“ Funktionen durch ihre Potenzreihe dargestellt werden: die reell-analytischen. Demgegenüber lassen sich durch Fourierreihen z.B. auch periodische Funktionen darstellen, die nur stückweise differenzierbar sind und deren Ableitungen Sprünge haben. Aus Zeitgründen können wir uns nur einen elementaren Überblick über Fourierreihen verschaffen, obwohl dies der Bedeutung dieser Reihen nicht angemessen ist. Auf weitere Aspekte wird z.B. in der Funktionalanalysis-Vorlesung eingegangen.

9.6.1 Periodische Funktionen

Eine auf \mathbb{R} definierte Funktion f heißt *periodisch mit der Periode L* , wenn

$$f(x + L) = f(x) \quad \text{für alle } x \in \mathbb{R}.$$

Beispielsweise sind die Funktionen $x \mapsto \sin(kx)$ und $x \mapsto \cos(kx)$ 2π -periodisch, und für die Dirichletfunktion ist *jede* rationale Zahl eine Periode.

Durch eine Variablensubstitution kann man Funktionen mit der Periode L auf solche mit der Periode 2π zurückführen. Hat f die Periode L , so hat $F(x) := f(\frac{L}{2\pi}x)$ die Periode 2π :

$$F(x + 2\pi) = f\left(\frac{L}{2\pi}(x + 2\pi)\right) = f\left(\frac{L}{2\pi}x + L\right) = f\left(\frac{L}{2\pi}x\right) = F(x).$$

Wir betrachten daher von nun an nur Funktionen der Periode 2π .

9.6.2 Trigonometrische Reihen

Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt *trigonometrische Reihe*, wenn es Konstanten $a_n \in \mathbb{R}$ für $n \geq 0$ und $b_n \in \mathbb{R}$ für $n \geq 1$ so gibt, dass

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \quad \text{für alle } x \in \mathbb{R}. \quad (9.11)$$

Trigonometrische Reihen sind offenbar 2π -periodisch.

Satz 9.21 *Wenn die Reihen $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ absolut konvergieren, so konvergiert die Reihe (9.11) auf \mathbb{R} absolut bezüglich der Supremumsnorm und gleichmäßig.*

Beweis. Die absolute Konvergenz der Reihe (9.11) folgt aus

$$\|a_n \cos(nx) + b_n \sin(nx)\|_\infty = \sup_{x \in \mathbb{R}} |a_n \cos(nx) + b_n \sin(nx)| \leq |a_n| + |b_n|.$$

Nach Satz 9.9 folgt aus der absoluten die gleichmäßige Konvergenz. ■

Im Weiteren benötigen wir die für alle $m, n \in \mathbb{Z}$ gültigen Identitäten

$$\begin{aligned} \int_0^{2\pi} \cos(mx) \cos(nx) dx &= \begin{cases} 0 & \text{für } m \neq n \\ \pi & \text{für } m = n \neq 0 \\ 2\pi & \text{für } m = n = 0, \end{cases} \\ \int_0^{2\pi} \sin(mx) \sin(nx) dx &= \begin{cases} 0 & \text{für } m \neq n \\ \pi & \text{für } m = n > 0, \end{cases} \\ \int_0^{2\pi} \sin(mx) \cos(nx) dx &= 0. \end{aligned} \quad (9.12)$$

Diese kann man leicht mit Hilfe von Additionstheoremen wie

$$\cos \alpha \cos \beta = \frac{1}{2} (\cos(\alpha - \beta) + \cos(\alpha + \beta))$$

zeigen (Übung). Mit den Identitäten (9.12) erhält man einen Zusammenhang zwischen den Werten einer trigonometrischen Reihe f und ihren Koeffizienten a_n und b_n .

Satz 9.22 (Euler/Fourier) *Die Reihe (9.11) sei auf \mathbb{R} gleichmäßig konvergent. Dann gilt*

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \quad (n \in \mathbb{N}), \\ b_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx \quad (n \in \mathbb{N} \setminus \{0\}). \end{aligned} \quad (9.13)$$

Beweis. Wir multiplizieren beide Seiten von (9.11) mit $\cos(nx)$

$$f(x) \cos(nx) = \frac{a_0}{2} \cos(nx) + \sum_{m=1}^{\infty} (a_m \cos(mx) \cos(nx) + b_m \sin(mx) \cos(nx))$$

und integrieren über $[0, 2\pi]$. Da die Reihe (9.11) gleichmäßig konvergiert, konvergiert auch die Reihe $f(x) \cos(nx)$ gleichmäßig, und Integration und Summation dürfen vertauscht werden. Mit (9.12) erhält man sofort die erste Behauptung in (9.13). Die zweite bekommt man analog durch Multiplikation von f mit $\sin(nx)$. ■

9.6.3 Fourierreihen

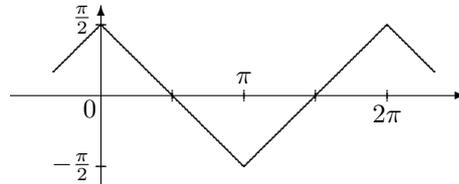
Die Formeln (9.13) erlauben die Bestimmung von Zahlen a_n, b_n auch dann, wenn f nicht als gleichmäßig konvergente trigonometrische Reihe vorausgesetzt wird, sondern nur als Riemann-integrierbar auf $[0, 2\pi]$ (es genügt sogar, dass $\int_0^{2\pi} f(x) dx$ als uneigentliches Integral absolut konvergiert.) Ist f eine solche Funktion, so bestimmen wir gemäß (9.13) Zahlen a_n, b_n , die dann die *Fourierkoeffizienten* von f heißen, und ordnen f die trigonometrische Reihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \quad (9.14)$$

zu, die sogenannte *Fourierreihe* von f .

Beispiel 1. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ die „Sägezahnfunktion“, die auf $[0, 2\pi]$ durch

$$f(x) = \begin{cases} -x + \frac{\pi}{2} & \text{für } x \in [0, \pi) \\ x - \frac{3}{2}\pi & \text{für } x \in [\pi, 2\pi) \end{cases}$$



definiert und 2π -periodisch ist.

Da f eine gerade Funktion ist (d.h. es ist $f(-x) = f(x)$ für alle $x \in \mathbb{R}$), ist $b_n = 0$ für alle $n \geq 1$. Für die a_n erhält man

$$a_n = \frac{1}{\pi} \int_0^{\pi} \left(-x + \frac{\pi}{2}\right) \cos(nx) dx + \frac{1}{\pi} \int_{\pi}^{2\pi} \left(x - \frac{3}{2}\pi\right) \cos(nx) dx.$$

Für $n = 0$ erhält man sofort $a_0 = 0$. Für $n \geq 1$ folgt mit partieller Integration

$$a_n = \frac{2}{\pi} \frac{1}{n^2} (1 - (-1)^n) = \begin{cases} \frac{4}{\pi} \cdot \frac{1}{n^2} & \text{für } n \text{ ungerade,} \\ 0 & \text{für } n \text{ gerade.} \end{cases}$$

Die durch f definierte Fourierreihe ist also

$$\frac{4}{\pi} \left(\cos x + \frac{\cos(3x)}{3^2} + \frac{\cos(5x)}{5^2} + \dots \right).$$

Da die Reihe $\sum a_n$ absolut konvergiert, konvergiert diese Fourierreihe gleichmäßig (Satz 9.21). Es ist aber im Moment nicht klar, ob die Summe dieser trigonometrischen Reihe etwas mit den Werten von f zu tun hat. Der folgende Satz 9.23 wird zeigen, dass tatsächlich

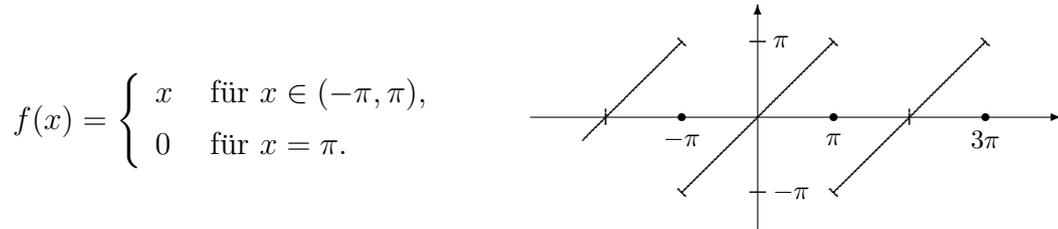
$$f(x) = \frac{4}{\pi} \left(\cos x + \frac{\cos(3x)}{3^2} + \frac{\cos(5x)}{5^2} + \dots \right) \quad (9.15)$$

für alle $x \in \mathbb{R}$ ist. Hieraus folgt z.B. die bemerkenswerte Identität

$$\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots = \frac{\pi^2}{8},$$

indem man in (9.15) $x = 0$ setzt. ■

Beispiel 2. Die 2π -periodische Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sei erklärt durch



Sie ist unstetig in allen Punkten $x = (2k + 1)\pi$ mit $k \in \mathbb{Z}$. Da f ungerade ist, sind alle $a_n = 0$. Für die b_n erhalten wir

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin(nx) dx = 2 \frac{(-1)^{n+1}}{n}$$

(wegen der 2π -Periodizität von f ist es egal, über welches Integral der Länge 2π man integriert). Die Fourierreihe von f lautet also

$$2 \left(\sin x - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} - \dots \right).$$

Es ist weder unmittelbar klar, ob diese Reihe für $x \neq 0$ überhaupt konvergiert, noch ob ihre Grenzfunktion mit f übereinstimmt. ■

9.6.4 Punktweise und gleichmäßige Konvergenz von Fourierreihen

Die zentrale Frage ist also: konvergiert die Fourierreihe einer 2π -periodischen Funktion f in irgendeinem Sinn (punktweise, gleichmäßig, ...), und wenn ja, stimmt ihre Grenzfunktion dann mit f überein?

Einfache Überlegungen zeigen, dass in der Regel nicht einmal punktweise Konvergenz vorliegen kann: zwei Funktionen f und g , die sich nur in endlich vielen Punkten voneinander unterscheiden, besitzen die gleiche Fourierreihe. Selbst für stetige Funktionen kann man nicht garantieren, dass ihre Fourierreihe gegen die Ausgangsfunktion punktweise konvergiert. Jedoch gilt:

Satz 9.23 Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sei 2π -periodisch und Riemann-integrierbar auf $[0, 2\pi]$, und die folgenden einseitigen Grenzwerte sollen an der Stelle $x \in \mathbb{R}$ existieren:

$$f(x+0) := \lim_{h \searrow 0} f(x+h), \quad f(x-0) := \lim_{h \nearrow 0} f(x+h),$$

$$\lim_{h \searrow 0} \frac{f(x+h) - f(x+0)}{h}, \quad \lim_{h \nearrow 0} \frac{f(x+h) - f(x-0)}{h}.$$

Dann konvergiert die Fourierreihe von f an der Stelle x gegen $\frac{f(x+0)+f(x-0)}{2}$.

Die Funktion f soll also in x einseitige Grenzwerte und „einseitige Ableitungen“ besitzen. Dann konvergiert die Fourierreihe gegen das arithmetische Mittel der einseitigen Grenzwerte. In den Beispielen 1 und 2 sind diese Bedingungen in jedem Punkt erfüllt. Man hat also jeweils punktweise Konvergenz der Fourierreihe gegen f auf ganz \mathbb{R} . Einen Beweis dieses Satzes finden Sie in Heuser, Analysis II, Satz 136.4. ■

Ohne die Differenzierbarkeitsannahme gilt dieser Satz nicht. Man kann jedoch ohne solche Annahmen auskommen, wenn man den Begriff der Konvergenz selbst abschwächt.

Wenn die Reihe $\sum_{k=1}^{\infty} a_k$ die Summe s besitzt, wenn also ihre Partialsummen $s_n = a_1 + \dots + a_n$ gegen s konvergieren, so konvergieren auch die arithmetischen Mittel $\sigma_n := \frac{1}{n}(s_1 + \dots + s_n)$ gegen s (Cauchyscher Grenzwertsatz, vgl. Heuser, Analysis I, Satz 27.1). Möglicherweise konvergiert die Folge (σ_n) aber auch dann noch gegen eine Zahl s , wenn die Folge (s_n) nicht konvergiert (Beispiel: $a_n = (-1)^n$). Man sagt dann, dass die Reihe $\sum_{n=1}^{\infty} a_n$ im Sinne von Cesaro konvergiert und dass ihre Summe gleich s ist.

Von Fejer wurde dieses Konzept auf Fourierreihen angewandt. Seien $s_0(x) := a_0/2$ und, für $n \geq 1$,

$$s_n(x) := \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx))$$

die Partialsummen der Fourierreihe und

$$\sigma_n(x) := \frac{1}{n+1}(s_0(x) + s_1(x) + \dots + s_n(x)).$$

Satz 9.24 (Fejer) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ 2π -periodisch und Riemann-integrierbar auf $[0, 2\pi]$. Für jedes $x \in \mathbb{R}$ sollen die einseitigen Grenzwerte $f(x+0)$ und $f(x-0)$ existieren, und es sei $f(x) = \frac{1}{2}(f(x+0)+f(x-0))$. Dann ist $f(x) = \lim_{n \rightarrow \infty} \sigma_n(x)$, d.h. die Fourierreihe von f konvergiert im Sinne von Cesaro punktweise gegen f . Ist f stetig, so konvergieren die σ_n sogar gleichmäßig gegen f .

Ein Beweis steht in Heuser, Analysis II, Sätze 139.3 und 139.5. ■

Eine Anwendung findet die zweite Aussage dieses Satzes beim Beweis des wichtigen Weierstraßschen Approximationssatzes.

Satz 9.25 (Weierstraß) Sei f stetig auf $[a, b]$ und $\varepsilon > 0$. Dann gibt es ein Polynom p mit

$$\|p - f\|_{\infty} = \sup_{x \in [a, b]} |p(x) - f(x)| < \varepsilon.$$

Stetige Funktionen auf Intervallen können also bezüglich der Supremumsnorm beliebig genau durch Polynome approximiert werden. Einen Beweis finden Sie in Heuser, Analysis II, Abschnitt 139, Aufgabe 3. ■

9.6.5 Konvergenz im quadratischen Mittel

Wir haben gesehen, dass die Untersuchung der punktweisen oder gar gleichmäßigen Konvergenz einer Fourierreihe erhebliche Schwierigkeiten bereitet. Ein zu Fourierreihen „passender“ Konvergenzbegriff ist die Konvergenz im quadratischen Mittel.

Seien f, g Riemann-integrierbare Funktionen auf $[0, 2\pi]$. Für andere Intervalle definiert man die folgenden Begriffe analog. Das *Skalarprodukt* von f und g ist die Zahl

$$\langle f, g \rangle := \int_0^{2\pi} f(t)g(t) dt,$$

und die L^2 -Norm von f wird erklärt durch

$$\|f\|_2 := \langle f, f \rangle^{1/2} = \left(\int_0^{2\pi} |f(t)|^2 dt \right)^{1/2}.$$

Dann gilt die Cauchy-Schwarz-Ungleichung

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2. \quad (9.16)$$

Man beachte die Analogie zum Skalarprodukt bzw. zur Euklidischen Norm von Vektoren im \mathbb{R}^n . Die Cauchy-Schwarz-Ungleichung (9.16) wird genauso bewiesen wie in diesem Fall.

Mit (9.16) erhält man leicht die Dreiecksungleichung

$$\|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

Es ist nämlich

$$\begin{aligned} \|f + g\|_2^2 &= \langle f + g, f + g \rangle = \langle f, f \rangle + 2\langle f, g \rangle + \langle g, g \rangle \\ &\leq \|f\|_2^2 + 2\|f\|_2 \|g\|_2 + \|g\|_2^2 = (\|f\|_2 + \|g\|_2)^2. \end{aligned}$$

Dennoch ist $\|\cdot\|_2$ keine Norm auf $R([0, 2\pi])$! (Warum nicht?) Mögliche Auswege sind:

- (A) Wir betrachten nur stetige Funktionen. Auf $C([0, 2\pi])$ ist $\|\cdot\|_2$ eine Norm.
- (B) Man identifiziert zwei Funktionen, wenn $\|f - g\|_2 = 0$.

Wir werden hier beides nicht tun: (A) engt uns zu sehr ein, und (B) schauen wir uns im Rahmen der Lebesgueschen Integrationstheorie im vierten Semester an.

Definition 9.26 Seien $f, f_n : \mathbb{R} \rightarrow \mathbb{R}$ 2π -periodisch und Riemann-integrierbar auf $[0, 2\pi]$. Die Folge (f_n) konvergiert im quadratischen Mittel gegen f , wenn

$$\lim_{n \rightarrow \infty} \|f - f_n\|_2 = 0 \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} \int_0^{2\pi} (f(x) - f_n(x))^2 dx = 0.$$

Noch einmal: die Tatsache, dass $\|\cdot\|_2$ keine Norm auf $R([0, 2\pi])$ ist, bringt einige Komplikationen mit sich (z.B. kann (f_n) gegen zwei verschiedene Funktionen f und g im quadratischen Mittel konvergieren), die wir erst im vierten Semester beheben.

Definition 9.27 Zwei Funktionen $f, g \in R([0, 2\pi])$ heißen orthogonal, wenn $\langle f, g \rangle = 0$. Eine Folge $(u_n)_{n \geq 0}$ von Funktionen heißt ein Orthogonalsystem, wenn $\langle u_m, u_n \rangle = 0$ für alle $m \neq n$, und ein Orthonormalsystem, wenn zusätzlich $\langle u_n, u_n \rangle = 1$ für alle $n \in \mathbb{N}$ ist.

Aus den Identitäten (9.12) wissen wir, dass die Funktionen

$$u_0(x) := \frac{1}{\sqrt{2\pi}}, \quad u_{2n}(x) := \frac{\cos(nx)}{\sqrt{\pi}}, \quad u_{2n-1}(x) := \frac{\sin(nx)}{\sqrt{\pi}} \quad (n \geq 1) \quad (9.17)$$

ein Orthonormalsystem über dem Intervall $[0, 2\pi]$ bilden.

Ist f Riemann-integrierbar und $(u_n)_{n \geq 0}$ ein Orthonormalsystem auf $[0, 2\pi]$, so heißen die Zahlen $c_n := \langle f, u_n \rangle$ die *Fourierkoeffizienten* von f und $\sum_{n=0}^{\infty} c_n u_n$ die *Fourierreihe* von f . Man überlegt sich leicht, dass für das spezielle Orthonormalsystem (9.17) diese Begriffe mit den früher definierten übereinstimmen.

Wann konvergiert nun die Fourierreihe einer Funktion im quadratischen Mittel gegen diese Funktion? Zunächst eine Vorüberlegung.

Satz 9.28 (Besselsche Ungleichung) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ 2π -periodisch und Riemann-integrierbar auf $[0, 2\pi]$, und sei $(u_n)_{n=0}^{\infty}$ ein Orthonormalsystem auf $[0, 2\pi]$. Dann gilt für die Fourierkoeffizienten $c_n = \langle f, u_n \rangle$ die Ungleichung

$$\sum_{n=0}^{\infty} |c_n|^2 \leq \int_0^{2\pi} |f(x)|^2 dx = \|f\|_2^2.$$

Beweis. Wir überlegen uns zunächst für jedes $k \in \mathbb{N}$ die Beziehung

$$\|f - \sum_{n=0}^k c_n u_n\|_2^2 = \|f\|_2^2 - \sum_{n=0}^k |c_n|^2. \quad (9.18)$$

Sei $g = \sum_{n=0}^k c_n u_n$. Dann ist

$$\|f - \sum_{n=0}^k c_n u_n\|_2^2 = \|f - g\|_2^2 = \langle f - g, f - g \rangle = \|f\|_2^2 - 2\langle f, g \rangle + \langle g, g \rangle.$$

Für die Skalarprodukte finden wir

$$\begin{aligned} \langle f, g \rangle &= \left\langle f, \sum_{n=0}^k c_n u_n \right\rangle = \sum_{n=0}^k c_n \langle f, u_n \rangle = \sum_{n=0}^k |c_n|^2, \\ \langle g, g \rangle &= \left\langle \sum_{n=0}^k c_n u_n, \sum_{m=0}^k c_m u_m \right\rangle = \sum_{n=0}^k \sum_{m=0}^k c_n c_m \langle u_n, u_m \rangle = \sum_{n=0}^k |c_n|^2, \end{aligned}$$

womit (9.18) sofort folgt. Da $\|f - g\|_2^2 \geq 0$, folgt aus (9.18) die Behauptung. ■

Folgerung 9.29 Die Fourierreihe $\sum_{n=0}^{\infty} c_n u_n$ von f konvergiert genau dann im quadratischen Mittel gegen f , wenn

$$\sum_{n=0}^{\infty} |c_n|^2 = \|f\|_2^2. \quad (9.19)$$

Dies folgt sofort aus (9.18). Die Beziehung (9.19) heißt die *Parsevalsche Gleichung*.

Satz 9.30 Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ 2π -periodisch und Riemann-integrierbar auf $[0, 2\pi]$, und sei (u_n) das spezielle Orthonormalsystem (9.17). Dann konvergiert die Fourierreihe von f im quadratischen Mittel gegen f .

Die Konvergenz der Fourierreihe gegen f im quadratischen Mittel gilt also ohne einschränkende Voraussetzungen an f und ist deshalb eine „sehr natürliche“ Art der Konvergenz für Fourierreihen. Dafür ist sie schwächer als die gleichmäßige Konvergenz.

Beweis. 1. Schritt. Sei $0 \leq a < 2\pi$. Wir zeigen die Aussage für die Funktion

$$f(x) = \begin{cases} 1 & \text{wenn } x \in [0, a]. \\ 0 & \text{wenn } x \in (a, 2\pi]. \end{cases}$$

Offenbar ist $\|f\|_2^2 = \int_0^a dx = a$, und für die Fourierkoeffizienten gilt

$$\begin{aligned} c_0 &= \langle f, u_0 \rangle = \langle f, \frac{1}{\sqrt{2\pi}} \rangle = \frac{1}{\sqrt{2\pi}} \int_0^a dx = \frac{a}{\sqrt{2\pi}}, \\ c_{2n} &= \langle f, u_{2n} \rangle = \langle f, \frac{\cos(nx)}{\sqrt{\pi}} \rangle = \frac{1}{\sqrt{\pi}} \int_0^a \cos(nx) dx = \frac{\sin(nx)}{n\sqrt{\pi}} \Big|_0^a = \frac{\sin(na)}{n\sqrt{\pi}}, \\ c_{2n-1} &= \langle f, u_{2n-1} \rangle = \langle f, \frac{\sin(nx)}{\sqrt{\pi}} \rangle = \frac{1}{\sqrt{\pi}} \int_0^a \sin(nx) dx \\ &= \frac{-\cos(nx)}{n\sqrt{\pi}} \Big|_0^a = \frac{1 - \cos(na)}{n\sqrt{\pi}}. \end{aligned}$$

Somit ist (die absolute Konvergenz der betrachteten Reihen folgt aus der Konvergenz der Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$; vgl. Kapitel 5)

$$\begin{aligned} \sum_{n=0}^{\infty} |c_n|^2 &= \frac{a^2}{2\pi} + \sum_{n=1}^{\infty} \frac{\sin^2(na)}{n^2\pi} + \sum_{n=1}^{\infty} \frac{(1 - \cos(na))^2}{n^2\pi} \\ &= \frac{a^2}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{\infty} \left(\frac{\sin^2(na)}{n^2} + \frac{1}{n^2} - \frac{2\cos(na)}{n^2} + \frac{\cos^2(na)}{n^2} \right) \\ &= \frac{a^2}{2\pi} + \frac{2}{\pi} \sum_{n=1}^{\infty} \left(\frac{1}{n^2} - \frac{\cos(na)}{n^2} \right). \end{aligned}$$

Mit der Identität

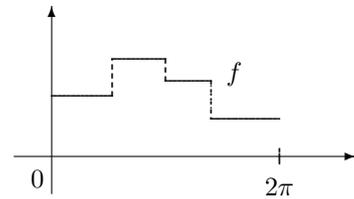
$$\sum_{n=1}^{\infty} \frac{\cos(nx)}{n^2} = \frac{(\pi - x)^2}{4} - \frac{\pi^2}{12}, \quad x \in [0, 2\pi]$$

(Nachrechnen!) erhalten wir weiter

$$\begin{aligned} \sum_{n=0}^{\infty} |c_n|^2 &= \frac{a^2}{2\pi} + \frac{2}{\pi} \frac{\pi^2}{6} - \frac{2}{\pi} \left(\frac{(\pi - a)^2}{4} - \frac{\pi^2}{12} \right) \\ &= \frac{a^2}{2\pi} + \frac{\pi}{3} - \frac{2}{\pi} \left(\frac{\pi^2}{4} - \frac{\pi a}{2} + \frac{a^2}{4} - \frac{\pi^2}{12} \right) = a. \end{aligned}$$

Für diese spezielle Funktion gilt also die Parsevalsche Gleichung (9.19). Aus Folgerung 9.29 folgt die Behauptung.

2. *Schritt.* Wir zeigen die Behauptung für stückweise konstante Funktionen f („Treppenfunktionen“). Für jede derartige Funktion gibt es Funktionen f_1, \dots, f_r von der im 1. Schritt beschriebenen Gestalt sowie Konstanten $\alpha_1, \dots, \alpha_r$ so, dass $f(x) = \sum_{j=1}^r \alpha_j f_j(x)$ für alle $x \in (0, 2\pi]$ mit Ausnahme endlich vieler.



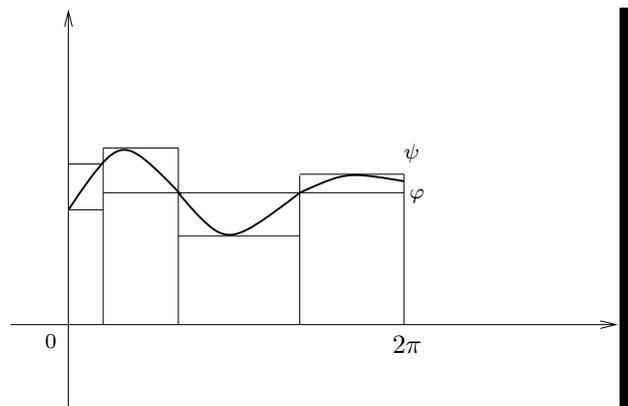
Seien s_n bzw. s_{nj} die n . Partialsummen der Fourierreihen der Funktionen f bzw. f_j . Dann ist offenbar $s_n = \sum_{j=1}^r \alpha_j s_{nj}$ und folglich

$$\|f - s_n\|_2 = \left\| \sum_{j=1}^r \alpha_j (f_j - s_{nj}) \right\|_2 \leq \sum_{j=1}^r |\alpha_j| \|f_j - s_{nj}\|_2.$$

Mit Schritt 1 folgt $\|f - s_n\|_2 \rightarrow 0$.

3. *Schritt.* Wir zeigen die Behauptung für eine beliebige Riemann-integrierbare Funktion f mit $\|f\|_{\infty} \leq 1$ (offenbar genügt es, solche Funktionen zu betrachten). Nach dem Riemannschemen Integrabilitätskriterium gibt es für jedes $\varepsilon > 0$ 2π -periodische Funktionen $\varphi, \psi : \mathbb{R} \rightarrow \mathbb{R}$ mit folgenden Eigenschaften:

- (a) φ, ψ sind Treppenfunktionen.
- (b) $-1 \leq \varphi \leq f \leq \psi \leq 1$.
- (c) $\int_0^{2\pi} (\psi(x) - \varphi(x)) dx \leq \frac{\varepsilon^2}{8}$.



Sei $g := f - \varphi$ und seien $s_{n,f}, s_{n,g}$ bzw. $s_{n,\varphi}$ die n -ten Partialsummen der Fourierreihen von f, g bzw. φ . Dann ist $s_{n,f} = s_{n,g} + s_{n,\varphi}$ und folglich

$$\|f - s_{n,f}\|_2 \leq \|\varphi - s_{n,\varphi}\|_2 + \|g - s_{n,g}\|_2. \quad (9.20)$$

Nach Schritt 2 gibt es ein N so, dass $\|\varphi - s_{n,\varphi}\|_2 < \varepsilon/2$ für alle $n \geq N$. Weiter ist

$$\begin{aligned} \|g - s_{n,g}\|_2^2 &\stackrel{(9.18)}{\leq} \|g\|_2^2 = \int_0^{2\pi} |g(x)|^2 dx = \int_0^{2\pi} |f(x) - \varphi(x)|^2 dx \\ &\leq \int_0^{2\pi} |\psi(x) - \varphi(x)|^2 dx \leq 2 \int_0^{2\pi} (\psi(x) - \varphi(x)) dx \stackrel{(c)}{\leq} \frac{\varepsilon^2}{4}. \end{aligned}$$

(Beachte: wegen $|\psi - \varphi| \leq 2$ ist $|\psi - \varphi|^2 \leq 2(\psi - \varphi)$.) Mit (9.20) folgt schließlich

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N : \|f - s_{n,f}\|_2 < \varepsilon,$$

d.h. die Partialsummen $s_{n,f}$ konvergieren in der L^2 -Norm gegen f . ■

10 Differentialrechnung für Funktionen mehrerer Veränderlicher

Reale Vorgänge hängen in der Regel von mehreren Einflussgrößen ab. Wir befassen uns daher in diesem Abschnitt mit der Differentialrechnung für Funktionen, die auf einer Teilmenge des \mathbb{R}^n definiert sind und in einen Raum \mathbb{R}^m abbilden.

10.1 Lineare Abbildungen und Stetigkeit

Wir bezeichnen wieder mit \mathbb{R}^n den linearen Raum aller Vektoren (x_1, \dots, x_n) mit den Operationen

$$\begin{aligned}(x_1, \dots, x_n) + (y_1, \dots, y_n) &:= (x_1 + y_1, \dots, x_n + y_n), \\ \alpha(x_1, \dots, x_n) &:= (\alpha x_1, \dots, \alpha x_n),\end{aligned}$$

wobei $\alpha \in \mathbb{R}$. Wir kennen bereits mehrere Normen auf \mathbb{R}^n :

$$\begin{aligned}\|(x_1, \dots, x_n)\|_\infty &:= \max_{1 \leq j \leq n} |x_j|, \\ \|(x_1, \dots, x_n)\|_2 &:= \left(\sum_{j=1}^n |x_j|^2 \right)^{1/2}, \\ \|(x_1, \dots, x_n)\|_1 &:= \sum_{j=1}^n |x_j|.\end{aligned}$$

Allgemeiner wird für jedes $p \geq 1$ durch

$$\|(x_1, \dots, x_n)\|_p := \left(\sum_{j=1}^n |x_j|^p \right)^{1/p}$$

eine Norm auf \mathbb{R}^n definiert, die \mathbb{R}^n zu einem normierten linearen Raum macht (↗ Tutorium).

Zwei Normen $\|\cdot\|_A$ und $\|\cdot\|_B$ auf einem linearen Raum X heißen *äquivalent*, wenn es Konstanten $C_1, C_2 > 0$ so gibt, dass

$$C_1 \|x\|_A \leq \|x\|_B \leq C_2 \|x\|_A \quad \text{für alle } x \in X.$$

Die Äquivalenz von Normen ist eine Äquivalenzrelation.

Satz 10.1 *Alle Normen auf \mathbb{R}^n sind untereinander äquivalent.*

Beweis. Sei $\|\cdot\|$ eine Norm auf \mathbb{R}^n . Da die Äquivalenz von Normen eine Äquivalenzrelation ist, genügt es zu zeigen, dass $\|\cdot\|$ zur Maximumnorm $\|\cdot\|_\infty$ äquivalent

ist. Dazu sei $e_j := (0, \dots, 0, 1, 0, \dots, 0)$ (mit der 1 an der j . Stelle) der j . Einheitsvektor im \mathbb{R}^n . Für $x = (x_1, \dots, x_n) = \sum_{j=1}^n x_j e_j$ ist dann

$$\|x\| = \left\| \sum_{j=1}^n x_j e_j \right\| \leq \sum_{j=1}^n |x_j| \|e_j\| \leq \|x\|_\infty \sum_{j=1}^n \|e_j\|,$$

woraus mit $C_2 := \sum_{j=1}^n \|e_j\|$ die Abschätzung

$$\|x\| \leq C_2 \|x\|_\infty \quad \text{für alle } x \in \mathbb{R}^n \quad (10.1)$$

folgt. Aus (10.1) erhalten wir für $x_1, x_2 \in \mathbb{R}^n$

$$\left| \|x_1\| - \|x_2\| \right| \leq \|x_1 - x_2\| \leq C_2 \|x_1 - x_2\|_\infty,$$

d.h. die Abbildung

$$(\mathbb{R}^n, \|\cdot\|_\infty) \rightarrow \mathbb{R}, \quad x \mapsto \|x\| \quad (10.2)$$

ist stetig (sogar Lipschitzstetig). Da die Menge $\{x \in \mathbb{R}^n : \|x\|_\infty = 1\}$ kompakt ist, nimmt die Funktion (10.2) ihr Minimum C_1 auf dieser Menge an, und C_1 ist positiv, da die Null nicht in dieser Menge liegt. Es ist also $\|x\| \geq C_1$ für alle $x \in \mathbb{R}^n$ mit $\|x\|_\infty = 1$ bzw. $\|x\| \geq C_1 \|x\|_\infty$ für alle $x \in \mathbb{R}^n$. ■

Folgerung 10.2 (a) \mathbb{R}^n ist bezüglich jeder Norm vollständig.

(b) Konvergiert eine Folge im \mathbb{R}^n bzgl. einer Norm, so konvergiert sie bzgl. jeder Norm.

(c) Alle Normen auf \mathbb{R}^n liefern die gleichen offenen Mengen.

(d) Die Stetigkeit von Abbildungen $f : X \rightarrow \mathbb{R}^n$ oder $g : \mathbb{R}^n \rightarrow Y$, wobei X und Y metrische Räume sind, hängt nicht von der Wahl der Norm auf \mathbb{R}^n ab.

Seien X, Y lineare Räume über \mathbb{R} . Eine Abbildung $A : X \rightarrow Y$ heißt *linear*, wenn

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay \quad \text{für alle } x, y \in X \text{ und } \alpha, \beta \in \mathbb{R}.$$

Sind X, Y normierte lineare Räume, so heißt eine lineare Abbildung $A : X \rightarrow Y$ *beschränkt*, wenn sie die Einheitskugel von X in eine beschränkte Menge in Y überführt. Die Zahl

$$\|A\| := \sup\{\|Ax\|_Y : x \in X, \|x\|_X \leq 1\}$$

heißt die durch die Normen $\|\cdot\|_X$ und $\|\cdot\|_Y$ induzierte *Operatornorm* von A . Die Norm $\|A\|$ ist also gleich dem Radius der kleinsten Kugel um $0 \in Y$, die das Bild der Einheitskugel von X unter der Abbildung A enthält. Sind beispielsweise \mathbb{R}^n und \mathbb{R}^m mit der Maximumnorm versehen und ist (a_{ij}) die Matrixdarstellung von $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ bezüglich der jeweiligen Standardbasen, so ist

$$\|A\| = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|.$$

Die Menge der linearen beschränkten Abbildungen von X nach Y bezeichnen wir mit $L(X, Y)$. Versehen mit der Operatornorm wird $L(X, Y)$ zu einem normierten linearen Raum (↗ Übung).

Satz 10.3 Für eine lineare Abbildung $A : X \rightarrow Y$ zwischen normierten Räumen sind folgende Aussagen äquivalent:

- (a) A ist beschränkt.
- (b) Es gibt ein $C \geq 0$ so, dass $\|Ax\| \leq C\|x\|$ für alle $x \in X$.
- (c) A ist stetig.
- (d) A ist stetig in $0 \in X$.

Beweis. (a) \Rightarrow (b): Für $x \neq 0$ ist $\|x/\|x\|\| = 1$ und daher $\|A\frac{x}{\|x\|}\| \leq \|A\|$. Also ist $\|Ax\| \leq \|A\| \|x\|$ für alle $x \in X$.

(b) \Rightarrow (c): Da A linear ist, ist

$$\|Ax - Ay\| = \|A(x - y)\| \leq C\|x - y\|$$

für beliebige $x, y \in X$. Also ist A sogar Lipschitzstetig.

(c) \Rightarrow (d): Dies ist klar.

(d) \Rightarrow (a): Sei $\delta > 0$ so, dass $\|Ax - A0\| = \|Ax\| \leq 1$ falls $\|x\| \leq \delta$. Dann ist $\|A(\delta x)\| \leq 1$ für alle $\|x\| \leq 1$, d.h. $\|A\| \leq 1/\delta$. ■

Satz 10.4 Jede lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^m ist stetig.

Beweis. Sei $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear. Wegen Folgerung 10.2 können wir annehmen, dass \mathbb{R}^n mit der Maximumnorm versehen ist. Für $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$ ist

$$\|Ax\| = \left\| \sum_{j=1}^n Ax_j e_j \right\| \leq \sum_{j=1}^n |x_j| \|Ae_j\| \leq \|x\|_\infty \sum_{j=1}^n \|Ae_j\|.$$

Nach Satz 10.3, Implikation (b) \Rightarrow (c), ist A stetig. ■

Da jeder endlichdimensionale lineare Raum über \mathbb{R} zu einem Raum \mathbb{R}^n isomorph ist, gelten die Sätze 10.1 und 10.4 entsprechend für beliebige endlichdimensionale lineare Räume über \mathbb{R} . Auch für endlichdimensionale Räume über \mathbb{C} bleiben diese Sätze richtig.

Als Ergänzung zum Thema Stetigkeit schauen wir uns noch die Vertauschbarkeit von Grenzübergängen an. Genauer: wir fragen, wann für $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gilt

$$\lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y) = \lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y).$$

Beispiel 1. Auf $(0, 1] \times (0, 1]$ sei $f(x, y) = x^y$. Dann ist

$$\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} x^y = \lim_{y \rightarrow 0} 0 = 0, \quad \text{aber} \quad \lim_{x \rightarrow 0} \lim_{y \rightarrow 0} x^y = \lim_{x \rightarrow 0} 1 = 1.$$

Beispiel 2. Auf $(0, 1] \times (0, 1]$ sei

$$f(x, y) = \frac{x \sin \frac{1}{x} + y}{x + y}.$$

Dann ist $\lim_{x \rightarrow 0} f(x, y) = 1$ und $\lim_{y \rightarrow 0} f(x, y) = \sin \frac{1}{x}$. Folglich ist

$$\lim_{y \rightarrow 0} \lim_{x \rightarrow 0} f(x, y) = 1,$$

aber $\lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x, y)$ existiert nicht. ■

Satz 10.5 Sei $f : (a, b) \times (c, d) \rightarrow \mathbb{R}$ und $(x_0, y_0) \in [a, b] \times [c, d]$. Der Grenzwert

$$A := \lim_{(x, y) \rightarrow (x_0, y_0)} f(x, y) \quad (10.3)$$

soll existieren, und für jedes $y \in (c, d)$ existiere der Grenzwert

$$\varphi(y) := \lim_{x \rightarrow x_0} f(x, y). \quad (10.4)$$

Dann existiert auch der iterierte Grenzwert $\lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y)$ und ist gleich A . Eine analoge Aussage gilt, wenn für jedes $x \in (a, b)$ der Grenzwert

$$\psi(x) := \lim_{y \rightarrow y_0} f(x, y) \quad (10.5)$$

existiert. Existieren also alle Grenzwerte (10.3) – (10.5), so ist

$$\lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y) = \lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y) = A.$$

Beweis. Wir zeigen nur die erste Aussage. Existenz des Grenzwertes (10.3) bedeutet:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall (x, y) \neq (x_0, y_0) \text{ mit } \|(x, y) - (x_0, y_0)\|_\infty < \delta : |f(x, y) - A| < \varepsilon.$$

Mit der Definition der Maximumnorm ist also

$$\forall (x, y) \neq (x_0, y_0) \text{ mit } |x - x_0| < \delta, |y - y_0| < \delta : |f(x, y) - A| < \varepsilon.$$

Wir fixieren ein $y \neq y_0$ mit $|y - y_0| < \delta$ und lassen $x \rightarrow x_0$ streben. Dann folgt

$$\forall \varepsilon > 0 \exists \delta > 0 \forall y \neq y_0 \text{ mit } |y - y_0| < \delta : |\varphi(y) - A| \leq \varepsilon.$$

Also existiert $\lim_{y \rightarrow y_0} \varphi(y)$ und ist gleich A . ■

10.2 Partielle Differenzierbarkeit

In diesem Abschnitt sei $U \subseteq \mathbb{R}^n$ offen und nicht leer und $f : U \rightarrow \mathbb{R}$. Ist $x = (x_1, \dots, x_n) \in U$, so schreiben wir statt $f(x)$ auch $f(x_1, \dots, x_n)$.

Definition 10.6 Die Funktion f heißt in $x \in U$ partiell differenzierbar bzgl. x_i , falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{h} \quad (10.6)$$

existiert. Dieser Grenzwert heißt die partielle Ableitung von f bzgl. x_i an der Stelle x und wird mit $(D_i f)(x)$, $\frac{\partial f}{\partial x_i}(x)$ oder $f_{x_i}(x)$ bezeichnet. Die Funktion f heißt partiell differenzierbar in x (auf U), wenn alle partiellen Ableitungen von f in x (in jedem Punkt von U) existieren. Ist außerdem jede dieser partiellen Ableitungen in x (bzw. auf U) stetig, so heißt f stetig partiell differenzierbar in x (bzw. auf U).

Ist e_i der i . Einheitsvektor von \mathbb{R}^n , so ist demnach

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

Beispiel 1. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ gegeben durch

$$f(x_1, \dots, x_n) = \sqrt{x_1^2 + \dots + x_n^2} \quad (= \|x\|_2). \quad (10.7)$$

Beim partiellen Differenzieren betrachten wir nur die i . der Variablen x_1, \dots, x_n als veränderlich und die übrigen als fixiert. Wir können daher die bekannten Differentiationsregeln für Funktionen einer reellen Veränderlichen anwenden und erhalten: Die Funktion (10.7) ist auf $\mathbb{R}^n \setminus \{0\}$ partiell differenzierbar, und

$$\frac{\partial f}{\partial x_i}(x) = \frac{\partial f}{\partial x_i}(x_1^2 + \dots + x_i^2 + \dots + x_n^2)^{1/2} = \frac{1}{2} (x_1^2 + \dots + x_n^2)^{-1/2} \cdot 2x_i = \frac{x_i}{\|x\|_2}.$$

Offenbar ist diese Funktion sogar *stetig* partiell differenzierbar auf $\mathbb{R}^n \setminus \{0\}$. ■

Beispiel 2. Bei Funktionen auf \mathbb{R}^2 oder \mathbb{R}^3 schreibt man oft $f(x, y)$ bzw. $f(x, y, z)$ statt $f(x_1, x_2)$ bzw. $f(x_1, x_2, x_3)$. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ erklärt durch $f(0, 0) = 0$ und

$$f(x, y) = \frac{xy}{(x^2 + y^2)^2} \quad \text{falls } (x, y) \neq (0, 0).$$

Für $(x, y) \neq (0, 0)$ finden wir die partiellen Ableitungen sofort:

$$f_x(x, y) = \frac{y}{(x^2 + y^2)^2} - 4 \frac{x^2 y}{(x^2 + y^2)^3}, \quad f_y(x, y) = \frac{x}{(x^2 + y^2)^2} - 4 \frac{xy^2}{(x^2 + y^2)^3}.$$

An der Stelle $(x, y) = (0, 0)$ arbeiten wir mit Definition 10.6 und finden

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0.$$

Analog ist $f_y(0, 0) = 0$. Also ist f auf ganz \mathbb{R}^2 partiell differenzierbar. Man beachte, dass f in $(0, 0)$ nicht stetig ist! Es ist nämlich

$$f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{(1/n^2)}{(2/n^2)^2} = \frac{n^2}{4} \rightarrow \infty \quad \text{für } n \rightarrow \infty.$$

Dieses Beispiel zeigt, dass aus der partiellen Differenzierbarkeit *nicht* die Stetigkeit folgt. Später werden wir sehen, dass dagegen aus der *stetigen* partiellen Differenzierbarkeit die Stetigkeit folgt. ■

Ist $f : U \rightarrow \mathbb{R}$ partiell differenzierbar und sind alle partiellen Ableitungen $\frac{\partial f}{\partial x_i} = D_i f : U \rightarrow \mathbb{R}$ wieder partiell differenzierbar, so heißt f *zweimal partiell differenzierbar*, und wir schreiben $\frac{\partial^2 f}{\partial x_j \partial x_i} = D_j D_i f = f_{x_i x_j}$ für die partielle Ableitung von $\frac{\partial f}{\partial x_i}$ nach x_j . Allgemein heißt f *k-mal partiell differenzierbar* ($k \geq 2$) wenn f $(k-1)$ -mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $k-1$ partiell differenzierbar sind. Schließlich heißt f *k-mal stetig partiell differenzierbar*, wenn f k -mal partiell differenzierbar ist und alle partiellen Ableitungen bis zur k . Ordnung stetig sind.

Beispiel 3. Wir erklären $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ durch $f(0, 0) = 0$ und

$$f(x, y) = xy \frac{x^2 - y^2}{x^2 + y^2} \quad \text{für } (x, y) \neq (0, 0).$$

Für alle Punkte $(x, y) \neq (0, 0)$ erhalten wir

$$f_x(x, y) = y \frac{x^4 - y^4 + 4x^2 y^2}{(x^2 + y^2)^2}, \quad f_y(x, y) = x \frac{x^4 - y^4 - 4x^2 y^2}{(x^2 + y^2)^2}.$$

Für $(x, y) = (0, 0)$ folgt mit der Definition der Ableitung

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = 0 \quad \text{und} \quad f_y(0, 0) = 0.$$

Für die gemischten zweiten Ableitungen in $(0, 0)$ finden wir schließlich

$$\begin{aligned} f_{xy}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_x(0, h) - f_x(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{-h - 0}{h} = -1, \\ f_{yx}(0, 0) &= \lim_{h \rightarrow 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h - 0}{h} = 1. \end{aligned} \tag{10.8}$$

Die Reihenfolge der partiellen Ableitungen darf also i. Allg. nicht vertauscht werden. Der folgende Satz gibt Bedingungen an, unter denen dieses Vertauschen erlaubt ist. ■

Satz 10.7 (H. A. Schwarz) Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ und $x^* \in U$. Alle partiellen Ableitungen erster Ordnung von f sollen auf U existieren. Weiter existiere die zweite Ableitung $\frac{\partial^2 f}{\partial x_i \partial x_j}$ auf U , und diese sei in x^* stetig. Dann existiert auch $\frac{\partial^2 f}{\partial x_j \partial x_i}$ in x^* , und es gilt

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x^*) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x^*).$$

Beweis. Wir beschränken uns auf den Fall $n = 2$, schreiben (x, y) statt (x_1, x_2) und setzen voraus, dass $\frac{\partial^2 f}{\partial y \partial x}$ auf U existiert und in x^* stetig ist. Weiter nehmen wir an, dass $(0, 0) \in U$ und $x^* = (0, 0)$ (andernfalls verschieben wir U geeignet). Schließlich wählen wir $\delta > 0$ so, dass $(-\delta, \delta) \times (-\delta, \delta) \subseteq U$ und arbeiten im weiteren ausschließlich auf diesem Quadrat.

Wir beginnen mit einer Vorüberlegung. Seien $h, k \in (-\delta, \delta) \setminus \{0\}$ fixiert. Der Mittelwertsatz, angewandt auf die Funktion

$$F : (-\delta, \delta) \rightarrow \mathbb{R}, \quad s \mapsto f(s, k) - f(s, 0)$$

liefert die Existenz eines $\xi = \xi(h, k)$ zwischen 0 und h so, dass

$$F(h) - F(0) = hF'(\xi) = h \left(\frac{\partial f}{\partial x}(\xi, k) - \frac{\partial f}{\partial x}(\xi, 0) \right).$$

Erneute Anwendung des Mittelwertsatzes, nun auf die Funktion

$$G : (-\delta, \delta) \rightarrow \mathbb{R}, \quad t \mapsto \frac{\partial f}{\partial x}(\xi, t),$$

liefert die Existenz eines $\eta = \eta(h, k)$ zwischen 0 und k so, dass

$$\frac{\partial f}{\partial x}(\xi, k) - \frac{\partial f}{\partial x}(\xi, 0) = G(k) - G(0) = k G'(\eta) = k \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta).$$

Es ist also

$$F(h) - F(0) = f(h, k) - f(h, 0) - f(0, k) + f(0, 0) = hk \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta)$$

mit gewissen $\xi = \xi(h, k)$ und $\eta = \eta(h, k)$. Damit wird

$$\begin{aligned} \frac{f_y(h, 0) - f_y(0, 0)}{h} &= \lim_{k \rightarrow 0} \frac{f(h, k) - f(h, 0) - f(0, k) + f(0, 0)}{hk} \\ &= \lim_{k \rightarrow 0} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \end{aligned} \quad (10.9)$$

mit gewissen Zahlen ξ und η , die von h und k abhängen und für die $|\xi| \leq |h|$ und $|\eta| \leq |k|$ ist. Für jedes fixierte h existiert also der Grenzwert (10.9).

Außerdem existiert wegen der Stetigkeit von $\frac{\partial^2 f}{\partial y \partial x}$ in $(0, 0)$ und wegen $(\xi, \eta) \rightarrow (0, 0)$ für $(h, k) \rightarrow (0, 0)$ der Grenzwert

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \quad \text{und ist gleich} \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

Nach Satz 10.5 existiert dann auch der iterierte Grenzwert

$$\lim_{h \rightarrow 0} \left(\lim_{k \rightarrow 0} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) \right) \quad \text{und ist gleich} \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

Wegen (10.9) bedeutet dies, dass der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f_y(h, 0) - f_y(0, 0)}{h} = \frac{\partial^2 f}{\partial x \partial y}(0, 0) \quad \text{existiert und gleich} \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) \text{ ist.}$$

Das ist die Behauptung. ■

Entsprechend gilt natürlich $f_{xxy} = f_{yxx} = f_{xyx}$, falls nur eine dieser Ableitungen existiert und stetig ist.

Eine Funktion $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}^m$, gegeben durch

$$f(x) = f(x_1, \dots, x_n) = \left(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n) \right),$$

heißt *partiell differenzierbar* bzw. *stetig partiell differenzierbar*, wenn jede ihrer Komponenten diese Eigenschaft besitzt.

Beispiel 4. Sei $U \subseteq \mathbb{R}^n$ offen. Ist $f : U \rightarrow \mathbb{R}$ partiell differenzierbar, so heißt

$$(\text{grad } f)(x) := \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \in \mathbb{R}^n$$

der *Gradient* von f in x . Ist $f : U \rightarrow \mathbb{R}^n$ partiell differenzierbar, so heißt

$$(\text{div } f)(x) := \frac{\partial f_1}{\partial x_1}(x) + \dots + \frac{\partial f_n}{\partial x_n}(x) \in \mathbb{R}$$

die *Divergenz* von f in x . Ist schließlich $n = 3$ und $f : U \rightarrow \mathbb{R}^3$ partiell differenzierbar, so heißt

$$(\text{rot } f)(x) := \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \right) \in \mathbb{R}^3$$

die *Rotation* von f in x . Ist $n = 3$ und $f : U \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar, so gilt

$$\text{rot grad } f = 0. \tag{10.10}$$

Die erste Komponente des Vektors $\operatorname{rot} \operatorname{grad} f$ ist nämlich

$$\frac{\partial}{\partial x_2} \frac{\partial f}{\partial x_3} - \frac{\partial}{\partial x_3} \frac{\partial f}{\partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_3} - \frac{\partial^2 f}{\partial x_3 \partial x_2} = 0$$

nach dem Satz von Schwarz. Ebenso sind die übrigen Komponenten gleich 0. Ähnlich zeigt man, dass für jede zweimal stetig partiell differenzierbare Funktion $f : U \rightarrow \mathbb{R}^3$ gilt

$$\operatorname{div} \operatorname{rot} f = 0. \quad (10.11)$$

10.3 Differenzierbarkeit

In diesem Abschnitt sehen wir uns an, wie sich das Konzept der Differentiation auf Funktionen mehrerer Veränderlicher übertragen lässt. Dazu erinnern wir an die Zerlegungsformel für differenzierbare Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$:

$$f(x+h) = f(x) + \alpha h + r(h) \quad \text{mit } \alpha \in \mathbb{R} \text{ und } \lim_{h \rightarrow 0} \frac{r(h)}{h} = 0.$$

Wir können diese Formel und damit die Grundidee der Differentiation wie folgt interpretieren: Die Funktion f wird im Punkt x lokal durch die lineare Funktion $h \mapsto \alpha h$ approximiert. Im Weiteren sei $\|\cdot\|$ die Euklidische Norm.

Definition 10.8 Sei $U \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}^m$ heißt differenzierbar (oft auch Frechét-differenzierbar oder total differenzierbar) in $x \in U$, wenn es eine lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sowie eine in einer Umgebung W von $0 \in \mathbb{R}^n$ definierte Funktion r mit Werten in \mathbb{R}^m und mit $\lim_{h \rightarrow 0} r(h)/\|h\| = 0$ so gibt, dass

$$f(x+h) = f(x) + Ah + r(h) \quad \text{für alle } h \in W. \quad (10.12)$$

Ist f in jedem Punkt $x \in U$ differenzierbar, so heißt f differenzierbar auf U .

Man beachte, dass A und r von x abhängen. Eine ganz analoge Definition, bei der man zusätzlich die Stetigkeit von A fordern muss, trifft man für beliebige Banachräume.

Wir überlegen uns die folgenden Beziehungen zwischen den eingeführten Differenzierbarkeitsbegriffen:

$$\begin{array}{ccccc} f \text{ stetig partiell} & \text{Satz 10.11} & f \text{ differen-} & \text{Satz 10.9} & f \text{ partiell} \\ \text{differenzierbar} & \implies & \text{zierbar} & \implies & \text{differenzierbar} \\ & & \Downarrow \text{Satz 10.10} & & \\ & & f \text{ stetig.} & & \end{array}$$

Satz 10.9 Sei $U \subseteq \mathbb{R}^n$ offen und $f = (f_1, \dots, f_m)^T : U \rightarrow \mathbb{R}^m$ in $x \in U$ differenzierbar. Dann ist jede Funktion $f_i : U \rightarrow \mathbb{R}$ in x partiell differenzierbar, und die Matrixdarstellung von $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ aus (10.12) bezüglich der Standardbasen von \mathbb{R}^n bzw. \mathbb{R}^m ist

$$A = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}. \quad (10.13)$$

Die Abbildung A in (10.12) ist also eindeutig bestimmt. Sie heißt die *Ableitung* von f in x und wird mit $(Df)(x)$ oder $f'(x)$ bezeichnet. Die Matrix (10.13) heißt die *Jacobi-Matrix* von f in x und wird oft mit $J_f(x)$ bezeichnet.

Beweis. Es gelte (10.12) mit $A = (a_{ij})_{i,j=1}^{mn}$, $r = (r_1, \dots, r_m)^T$, $h = (h_1, \dots, h_n)^T$. Mit diesen Bezeichnungen folgt für jedes $i = 1, \dots, m$

$$f_i(x+h) = f_i(x) + \sum_{j=1}^n a_{ij} h_j + r_i(h) \quad \text{für alle } h \in W, \quad (10.14)$$

wobei $\lim_{h \rightarrow 0} \frac{r_i(h)}{\|h\|} = 0$, da $|r_i| \leq \|r\|_\infty$. Wir fixieren nun ein j zwischen 1 und n und wählen $h = (0, \dots, 0, h_j, 0, \dots, 0) = h_j e_j$. Für hinreichend kleine h_j liegen diese Vektoren in W , und (10.14) reduziert sich auf

$$f_i(x + h_j e_j) = f_i(x) + a_{ij} h_j + r_i(h).$$

Hieraus folgt

$$\frac{f_i(x + h_j e_j) - f_i(x)}{h_j} = a_{ij} + \frac{r_i(h)}{h_j} \quad \text{für } h_j \neq 0.$$

Wegen $|h_j| = \|h\|$ können wir $h_j \rightarrow 0$ streben lassen und erhalten

$$\frac{\partial f_i}{\partial x_j}(x) = \lim_{h_j \rightarrow 0} \frac{f_i(x + h_j e_j) - f_i(x)}{h_j} = a_{ij}. \quad \blacksquare$$

Satz 10.10 Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ in $x \in U$ differenzierbar. Dann ist f in x stetig.

Beweis. Da lineare Abbildungen von \mathbb{R}^n nach \mathbb{R}^m stetig sind (Satz 10.4) und $0 \in \mathbb{R}^n$ in $0 \in \mathbb{R}^m$ überführen, geht die rechte Seite von (10.12) für $h \rightarrow 0$ gegen $f(x)$. Also existiert $\lim_{h \rightarrow 0} f(x+h)$ und ist gleich $f(x)$. \blacksquare

Satz 10.11 Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^m$ in $x \in U$ stetig partiell differenzierbar. Dann ist f in x differenzierbar.

Beweis. Es genügt, diese Aussage für jede Komponente von f zu zeigen, d.h. wir nehmen $m = 1$ an. Für $h = (h_1, \dots, h_n)$ nahe bei 0 definieren wir

$$x^{(0)} := x \quad \text{und} \quad x^{(i)} := x + \sum_{j=1}^i h_j e_j \quad \text{für } i = 1, \dots, n.$$

Insbesondere ist $x^{(n)} = x + h$. Da sich $x^{(i)}$ und $x^{(i-1)}$ nur in der i . Komponente unterscheiden, gibt es nach dem Mittelwertsatz für Funktionen *einer* Veränderlichen für jedes $i = 1, \dots, n$ ein $t_i \in [0, 1]$ so, dass

$$f(x^{(i)}) - f(x^{(i-1)}) = \frac{\partial f}{\partial x_i}(x^{(i-1)} + t_i h_i e_i) \cdot h_i.$$

Aufsummieren liefert

$$f(x + h) - f(x) = f(x^{(n)}) - f(x^{(0)}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) h_i + r(h)$$

(Teleskopsumme) mit

$$r(h) = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x^{(i-1)} + t_i h_i e_i) - \frac{\partial f}{\partial x_i}(x) \right) h_i.$$

Da alle partiellen Ableitungen in x nach Voraussetzung stetig sind, ist

$$\lim_{h \rightarrow 0} \left(\frac{\partial f}{\partial x_i}(x^{(i-1)} + t_i h_i e_i) - \frac{\partial f}{\partial x_i}(x) \right) = 0$$

für jedes $i = 1, \dots, n$. Folglich ist $\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$. ■

Wir nennen stetig partiell differenzierbare Funktionen daher auch kurz *stetig differenzierbar*.

Keine der im Schema vor Satz 10.9 angegebenen Implikationen lässt sich umkehren: In Beispiel 2 aus Abschnitt 10.2 ist eine partiell differenzierbare, aber nicht differenzierbare Funktion angegeben. Ein Beispiel für eine differenzierbare, aber nicht stetig differenzierbare Funktion steht in Heuser, Ana II, Pkt. 164, Aufg. 7.

Die aus Kapitel 7 bekannten Differentiationsregeln übertragen sich ohne Änderung auf den allgemeinen Fall.

Satz 10.12 *Seien $U \subseteq \mathbb{R}^n$ offen und $f, g : U \rightarrow \mathbb{R}^m$ in $x \in U$ differenzierbar. Dann ist für alle $\alpha, \beta \in \mathbb{R}$ auch die Funktion $\alpha f + \beta g$ in x differenzierbar, und es gilt*

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x).$$

Die Produkt- und Quotientenregel vermerken wir nur für skalarwertige Funktionen.

Satz 10.13 Seien $U \subseteq \mathbb{R}^n$ offen und $f, g : U \rightarrow \mathbb{R}$ in $x \in U$ differenzierbar. Dann sind auch die Funktionen $fg : U \rightarrow \mathbb{R}$ und (falls $g(x) \neq 0$) $f/g : U \rightarrow \mathbb{R}$ in x differenzierbar, und es ist

$$\begin{aligned}(fg)'(x) &= g(x)f'(x) + f(x)g'(x), \\ (f/g)'(x) &= \frac{g(x)f'(x) - f(x)g'(x)}{g^2(x)}.\end{aligned}$$

Die Beweise dieser beiden Sätze werden wie im Fall $n = 1$ geführt und sind HA.

Satz 10.14 (Kettenregel) Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offene Mengen, und seien $g : U \rightarrow \mathbb{R}^m$ und $f : V \rightarrow \mathbb{R}^k$ Funktionen mit $g(U) \subseteq V$. Ist g in $x_0 \in U$ und f in $g(x_0) \in V$ differenzierbar, so ist die zusammengesetzte Funktion $f \circ g : U \rightarrow \mathbb{R}^k$ in x_0 differenzierbar, und es gilt

$$(f \circ g)'(x_0) = f'(g(x_0)) \circ g'(x_0).$$

Das \circ auf der rechten Seite steht für die Verkettung der linearen Abbildungen $g'(x_0)$ und $f'(g(x_0))$. In Matrixschreibweise bedeutet dies gerade das Bilden des Matrixprodukts der $k \times m$ -Matrix $f'(g(x_0))$ mit der $m \times n$ -Matrix $g'(x_0)$.

Beweis. Differenzierbarkeit von g in x_0 bzw. f in $g(x_0)$ bedeutet

$$g(x) - g(x_0) = g'(x_0)(x - x_0) + r(x - x_0), \quad (10.15)$$

$$f(y) - f(g(x_0)) = f'(g(x_0))(y - g(x_0)) + s(y - g(x_0)), \quad (10.16)$$

wobei

$$\lim_{x \rightarrow x_0} \frac{r(x - x_0)}{\|x - x_0\|} = 0, \quad \lim_{y \rightarrow g(x_0)} \frac{s(y - g(x_0))}{\|y - g(x_0)\|} = 0.$$

Wir setzen in (10.16) $y = g(x)$ und dann (10.16) in (10.15) ein und erhalten

$$f(g(x)) - f(g(x_0)) = f'(g(x_0))g'(x_0)(x - x_0) + t(x_0, x)$$

mit

$$t(x_0, x) = f'(g(x_0))r(x - x_0) + s(g(x) - g(x_0)).$$

Wir müssen zeigen, dass

$$\frac{t(x_0, x)}{\|x - x_0\|} = f'(g(x_0)) \frac{r(x - x_0)}{\|x - x_0\|} + \frac{s(g(x) - g(x_0))}{\|x - x_0\|}$$

für $x \rightarrow x_0$ gegen 0 strebt. Für den ersten Summanden ist dies wegen der Stetigkeit der linearen Abbildung $f'(g(x_0))$ und wegen $\lim_{h \rightarrow 0} r(h)/\|h\| = 0$ klar. Für

den zweiten Summanden beachten wir, dass $s(0) = 0$. Für $g(x) = g(x_0)$ ist also $s(g(x) - g(x_0)) = 0$, und für $g(x) \neq g(x_0)$ haben wir

$$\frac{s(g(x) - g(x_0))}{\|x - x_0\|} = \frac{s(g(x) - g(x_0))}{\|g(x) - g(x_0)\|} \frac{\|g(x) - g(x_0)\|}{\|x - x_0\|}.$$

Wegen der Stetigkeit von g in x_0 und wegen $\lim_{h \rightarrow 0} s(h)/\|h\| = 0$ ist

$$\lim_{x \rightarrow x_0} \frac{s(g(x) - g(x_0))}{\|g(x) - g(x_0)\|} = 0.$$

Schließlich folgt aus der Zerlegungsformel für g

$$\frac{\|g(x) - g(x_0)\|}{\|x - x_0\|} = \frac{\|g'(x_0)(x - x_0) + r(x - x_0)\|}{\|x - x_0\|} \leq \|g'(x_0)\| + \frac{\|r(x - x_0)\|}{\|x - x_0\|},$$

dass der Quotient $\|g(x) - g(x_0)\|/\|x - x_0\|$ in einer Umgebung von x_0 beschränkt bleibt. ■

Wir sehen uns die Kettenregel für einige Spezialfälle an.

Beispiel 1. Die reellwertigen Funktionen f bzw. x_1, \dots, x_n seien auf der offenen Menge $U \subseteq \mathbb{R}^n$ bzw. auf dem offenen Intervall $I \subseteq \mathbb{R}$ definiert, und die verkettete Funktion $F(t) := f(x_1(t), \dots, x_n(t))$ soll auf I erklärt sein. Sind alle Funktionen f und x_i auf I differenzierbar, so ist auch F auf I differenzierbar, und

$$\frac{dF}{dt} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt}.$$

Genauer: für $t_0 \in I$ ist

$$\frac{dF}{dt}(t_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_1(t_0), \dots, x_n(t_0)) \frac{dx_i}{dt}(t_0).$$

Der Beweis folgt sofort aus der Kettenregel, angewandt auf die äußere Funktion f und die innere Funktion $g(t) := (x_1(t), \dots, x_n(t))^T : I \rightarrow \mathbb{R}^n$. ■

Beispiel 2. Die reellwertigen Funktionen f bzw. u_1, \dots, u_n seien auf der offenen Menge $U \subseteq \mathbb{R}^n$ bzw. der offenen Menge $V \subseteq \mathbb{R}^m$ definiert. Wir betrachten die verkettete Funktion

$$F : (x_1, \dots, x_m) \mapsto f(u_1(x_1, \dots, x_m), \dots, u_n(x_1, \dots, x_m))$$

auf V . Ist f auf U differenzierbar und jede Funktion u_i auf V partiell differenzierbar, so ist F auf V partiell differenzierbar, und es gilt

$$\frac{\partial F}{\partial x_i} = \frac{\partial f}{\partial u_1} \frac{\partial u_1}{\partial x_i} + \dots + \frac{\partial f}{\partial u_n} \frac{\partial u_n}{\partial x_i} \quad \text{für } 1 \leq i \leq n.$$

Dies folgt sofort aus Beispiel 1. Ist sogar jede der Funktionen u_i differenzierbar, so ist F auf V differenzierbar, und es ist

$$F' = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_m} \right) \text{ auf } V.$$

Beispiel 3. Sei $x : (0, 1) \rightarrow \mathbb{R}^n \setminus \{0\}$ eine differenzierbare Funktion. Dann ist die verkettete Funktion

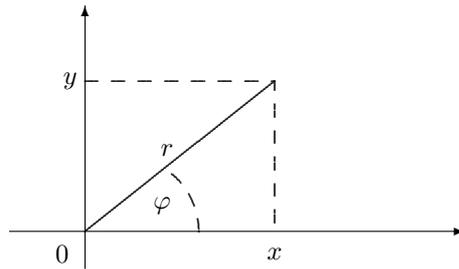
$$f : (0, 1) \rightarrow \mathbb{R}, \quad t \mapsto \|x(t)\|_2$$

nach Beispiel 1 aus Abschnitt 10.2 und nach Satz 10.14 differenzierbar, und die Kettenregel liefert wie in Beispiel 1

$$f'(t) = \sum_{i=1}^n \frac{x_i(t)}{\|x(t)\|_2} x'_i(t) = \frac{\langle x(t), x'(t) \rangle}{\|x(t)\|_2},$$

wobei $\langle \cdot, \cdot \rangle$ für das übliche Skalarprodukt im \mathbb{R}^n steht. ■

Beispiel 4. Wir sehen uns die Anwendung der Kettenregel bei der Transformation von Differentialausdrücken an. Durch Einführung von Polarkoordinaten $x = r \cos \varphi$, $y = r \sin \varphi$ auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ wird aus einer Funktion $u = u(x, y)$ eine Funktion $v = v(r, \varphi) = u(r \cos \varphi, r \sin \varphi)$.



Wir zeigen, dass dabei beispielsweise der Ausdruck $x \frac{\partial u}{\partial y} - y \frac{\partial u}{\partial x}$ in $\frac{\partial v}{\partial r}$ übergeht. Aus $x = r \cos \varphi$ und $y = r \sin \varphi$ folgt $r = \sqrt{x^2 + y^2}$ und $\varphi = \arctan \frac{y}{x}$. Differenzieren liefert

$$\frac{\partial r}{\partial x} = \frac{2x}{2\sqrt{x^2 + y^2}} = \frac{r \cos \varphi}{r} = \cos \varphi$$

und analog $\frac{\partial r}{\partial y} = \sin \varphi$ sowie

$$\frac{\partial \varphi}{\partial x} = \frac{1}{1 + (\frac{y}{x})^2} \frac{-y}{x^2} = \frac{-y}{x^2 + y^2} = \frac{-r \sin \varphi}{r^2} = -\frac{\sin \varphi}{r}$$

und analog $\frac{\partial \varphi}{\partial y} = \frac{\cos \varphi}{r}$. Demzufolge ist nach Beispiel 2

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial v}{\partial \varphi} \frac{\partial \varphi}{\partial x} = \frac{\partial v}{\partial r} \cos \varphi - \frac{\partial v}{\partial \varphi} \frac{\sin \varphi}{r}, \\ \frac{\partial u}{\partial y} &= \frac{\partial v}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial v}{\partial \varphi} \frac{\partial \varphi}{\partial y} = \frac{\partial v}{\partial r} \sin \varphi + \frac{\partial v}{\partial \varphi} \frac{\cos \varphi}{r}. \end{aligned}$$

Hieraus folgt schließlich

$$\begin{aligned} x \frac{\partial u}{\partial y} - y \frac{\partial u}{\partial x} &= r \cos \varphi \left(\frac{\partial v}{\partial r} \sin \varphi + \frac{\partial v}{\partial \varphi} \frac{\cos \varphi}{r} \right) - r \sin \varphi \left(\frac{\partial v}{\partial r} \cos \varphi - \frac{\partial v}{\partial \varphi} \frac{\sin \varphi}{r} \right) \\ &= \frac{\partial v}{\partial r} (r \sin \varphi \cos \varphi - r \sin \varphi \cos \varphi) + \frac{\partial v}{\partial \varphi} (\cos^2 \varphi + \sin^2 \varphi) = \frac{\partial v}{\partial \varphi}. \end{aligned}$$

Es ist also beispielsweise genau dann $x \frac{\partial u}{\partial y} - y \frac{\partial u}{\partial x} = 0$, wenn $\frac{\partial v}{\partial \varphi} = 0$, d.h. genau dann, wenn v nur von r abhängt, also rotationssymmetrisch ist. ■

10.4 Richtungsableitungen

Definition 10.15 Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$, $x \in U$ und $v \in \mathbb{R}^n$ ein Vektor der Länge $\|v\|_2 = 1$. Man sagt, dass f eine Ableitung in Richtung des Vektors v besitzt, wenn der Grenzwert

$$\lim_{t \rightarrow 0} \frac{f(x + tv) - f(x)}{t}$$

existiert. Wir bezeichnen diesen Grenzwert mit $\frac{\partial f}{\partial v}$.

Die partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ sind spezielle Richtungsableitungen in Richtung der Einheitsvektoren $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ mit der 1 an i . Stelle.

Satz 10.16 Seien U, x, f wie in Definition 10.15, und f sei in x differenzierbar. Dann existiert für jeden Einheitsvektor $v = (v_1, \dots, v_n)^T \in \mathbb{R}^n$ die Ableitung von f im Punkt x in Richtung v , und es gilt

$$\frac{\partial f}{\partial v}(x) = f'(x)v = \frac{\partial f}{\partial x_1}(x)v_1 + \dots + \frac{\partial f}{\partial x_n}(x)v_n. \quad (10.17)$$

Beweis. Für t hinreichend nahe bei 0 ist nach Definition von $f'(x)$

$$\frac{f(x + tv) - f(x)}{t} = \frac{f'(x) \cdot tv + r(tv)}{t} = f'(x)v + \frac{r(tv)}{t}.$$

Aus $\lim_{t \rightarrow 0} \frac{\|r(tv)\|}{|t|} = \lim_{t \rightarrow 0} \frac{\|r(tv)\|}{\|tv\|} = 0$ folgt die Behauptung. ■

Mit Hilfe des Gradienten $\text{grad} f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$ können wir (10.17) auch als

$$\frac{\partial f}{\partial v}(x) = \langle (\text{grad} f)(x), v \rangle \quad (10.18)$$

schreiben, wobei $\langle \cdot, \cdot \rangle$ das übliche Skalarprodukt auf \mathbb{R}^n ist. Wie im \mathbb{R}^2 führt man den Winkel $\varphi \in [0, \pi]$ zwischen zwei Vektoren $x, y \in \mathbb{R}^n \setminus \{0\}$ durch

$$\langle x, y \rangle = \|x\| \|y\| \cos \varphi$$

ein. Ist also $(\text{grad } f)(x) \neq 0$, so folgt aus (10.18) für den Winkel φ zwischen v und $(\text{grad } f)(x)$:

$$\frac{\partial f}{\partial v}(x) = \|(\text{grad } f)(x)\| \cos \varphi.$$

Folgerung 10.17 *Seien U, f, x, v wie in Definition 10.15, und sei $(\text{grad } f)(x) \neq 0$. Dann ist die Richtungsableitung $\frac{\partial f}{\partial v}(x)$ genau dann maximal, wenn $\cos \varphi = 1$ bzw. $\varphi = 0$, d.h. wenn $(\text{grad } f)(x)$ und v die gleiche Richtung haben. Der Gradient zeigt also in die Richtung des stärksten Anstieges von f in x . Diese Tatsache wird bei der numerischen Lösung von Extremalaufgaben benutzt.*

10.5 Der Mittelwertsatz

Am Ende von Abschnitt 7.6 haben wir gesehen, dass der Mittelwertsatz in seiner gewohnten Form für vektorwertige Funktionen nicht mehr gilt. Man hat jedoch für reellwertige Funktionen auf \mathbb{R}^n die folgende Version.

Satz 10.18 (Mittelwertsatz) *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ differenzierbar. Weiter sei $x + th \in U$ für alle $t \in [0, 1]$. Dann existiert ein $\tau \in (0, 1)$ mit*

$$f(x + h) - f(x) = f'(x + \tau h)h.$$

Beweis. Wir betrachten die differenzierbare Funktion

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

Nach dem Mittelwertsatz 7.28 für Funktionen einer Veränderlichen gibt es ein $\tau \in (0, 1)$ mit

$$f(x + h) - f(x) = g(1) - g(0) = \frac{dg}{dt}(\tau).$$

Nach der Kettenregel (Beispiel 1 aus 10.3) ist weiter

$$\frac{dg}{dt}(\tau) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x + \tau h)h_i = f'(x + \tau h)h. \quad \blacksquare$$

Für Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $m \geq 2$ ist die folgende Version des Mittelwertsatzes die nächstbeste.

Satz 10.19 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar und $x + th \in U$ für alle $t \in [0, 1]$. Dann ist*

$$f(x + h) - f(x) = \int_0^1 f'(x + \tau h)h \, d\tau.$$

Ist insbesondere $\|f'(x + th)\| \leq M$ für alle $t \in [0, 1]$, so folgt

$$\|f(x + h) - f(x)\| \leq M\|h\|.$$

Das Integral über die vektorwertige Funktion $f'(x + \tau h)h$ ist komponentenweise erklärt.

Beweis. Für $g : [0, 1] \rightarrow \mathbb{R}^m$, $g(t) := f(x + th)$ ist $g'(t) = f'(x + th)h$ und daher

$$f(x + h) - f(x) = g(1) - g(0) = \int_0^1 g'(\tau) d\tau = \int_0^1 f'(x + \tau h)h d\tau,$$

womit die erste Aussage gezeigt ist. Die zweite folgt aus

$$\begin{aligned} \left\| \int_0^1 f'(x + \tau h)h d\tau \right\| &\leq \int_0^1 \|f'(x + \tau h)h\| d\tau \\ &\leq \int_0^1 \|f'(x + \tau h)\| \|h\| d\tau \leq \int_0^1 M \|h\| d\tau = M \|h\|. \end{aligned} \quad \blacksquare$$

10.6 Der Satz von Taylor

Wir lernen nun den Satz von Taylor für Funktionen mehrerer Veränderlicher kennen. Dabei beschränken wir uns auf reellwertige Funktionen. Wir führen zunächst einige Bezeichnungen ein, die helfen, die Übersicht über die zahlreichen Summanden in der Taylorentwicklung zu behalten.

Ein *Multiindex* ist ein n -Tupel $(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$. Für jeden Multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$ sei $|\alpha| := \alpha_1 + \dots + \alpha_n$ seine *Ordnung* und $\alpha! := \alpha_1! \dots \alpha_n!$ seine *Fakultät*. Für jede $|\alpha|$ -mal partiell differenzierbare Funktion $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ setzen wir

$$D^\alpha f := D_1^{\alpha_1} \dots D_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}},$$

und für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ sei

$$x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}.$$

Als Vorbereitung für den Satz von Taylor zeigen wir:

Satz 10.20 Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ k -mal stetig partiell differenzierbar, und sei $x + th \in U$ für alle $t \in [0, 1]$. Dann ist die Funktion

$$g : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto f(x + th)$$

k -mal stetig differenzierbar, und es gilt

$$\frac{d^k g}{dt^k}(\tau) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D^\alpha f)(x + \tau h)h^\alpha \quad (10.19)$$

(die Summation erfolgt über alle Multiindizes der Ordnung k).

Beweis. Wir zeigen zuerst mit vollständiger Induktion, dass

$$\frac{d^k g}{dt^k}(\tau) = \sum_{i_1, \dots, i_k=1}^n (D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \cdot \dots \cdot h_{i_k}. \quad (10.20)$$

Für $k = 1$ haben wir dies bereits im Beweis von Satz 10.18 getan. Aus der Kettenregel folgt nämlich

$$\begin{aligned} \frac{dg}{dt}(\tau) &= \frac{df(x_1 + th_1, \dots, x_n + th_n)}{dt}(\tau) = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt} \\ &= \frac{\partial f}{\partial x_1} h_1 + \dots + \frac{\partial f}{\partial x_n} h_n = \sum_{i_1=1}^n \frac{\partial f}{\partial x_{i_1}}(x + \tau h) h_{i_1} \\ &= \sum_{i_1=1}^n (D_{i_1} f)(x + \tau h) h_{i_1}. \end{aligned}$$

Nehmen wir an, dass (10.20) für ein $k - 1 \geq 1$ richtig ist, so folgt analog

$$\begin{aligned} \frac{d^k g}{dt^k}(\tau) &= \frac{d}{dt} \left(\sum_{i_1, \dots, i_{k-1}=1}^n (D_{i_{k-1}} \dots D_{i_1} f)(x + th) h_{i_1} \dots h_{i_{k-1}} \right) (\tau) \\ &= \sum_{i_k=1}^n D_{i_k} \left(\sum_{i_1, \dots, i_{k-1}=1}^n (D_{i_{k-1}} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_{k-1}} \right) h_{i_k} \\ &= \sum_{i_1, \dots, i_k=1}^n (D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_k}. \end{aligned}$$

Damit ist (10.20) gezeigt. Wir überlegen uns nun, dass die rechte Seite von (10.20) gleich der rechten Seite von (10.19) ist. Grundidee ist, dass es nach dem Satz von Schwarz auf die Reihenfolge der partiellen Ableitungen nicht ankommt und wir daher die partiellen Ableitungen umsortieren können. Kommt unter den Indizes i_1, \dots, i_k die Zahl $m \in \{1, 2, \dots, n\}$ genau α_m -mal vor, so ergibt ein Umsortieren und Zusammenfassen gleicher Ableitungen

$$(D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_k} = (D_1^{\alpha_1} \dots D_n^{\alpha_n} f)(x + \tau h) h_1^{\alpha_1} \dots h_n^{\alpha_n}.$$

Da es genau $\frac{k!}{\alpha_1! \dots \alpha_n!} = \frac{k!}{\alpha!}$ k -Tupel (i_1, \dots, i_k) gibt, in denen die Zahl m genau α_m -mal vorkommt (beachten Sie, dass $\alpha_1 + \dots + \alpha_n = k$), gilt

$$\begin{aligned} \frac{d^k g}{dt^k}(\tau) &= \sum_{i_1, \dots, i_k}^n (D_{i_k} \dots D_{i_1} f)(x + \tau h) h_{i_1} \dots h_{i_k} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D_1^{\alpha_1} \dots D_n^{\alpha_n} f)(x + \tau h) h_1^{\alpha_1} \dots h_n^{\alpha_n} \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} (D^\alpha f)(x + \tau h) h^\alpha. \quad \blacksquare \end{aligned}$$

Satz 10.21 (Taylor) Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ $(k+1)$ -mal stetig partiell differenzierbar, und $x + th \in U$ für alle $t \in [0, 1]$. Dann gibt es ein $\tau \in (0, 1)$ so, dass

$$f(x+h) = \underbrace{\sum_{|\alpha| \leq k} \frac{1}{\alpha!} (D^\alpha f)(x) h^\alpha}_{\text{Taylorpolynom der Ordnung } k} + \underbrace{\sum_{|\alpha|=k+1} \frac{(D^\alpha f)(x + \tau h)}{\alpha!} h^\alpha}_{\text{Restglied}}.$$

Beispiel. Um das Taylorpolynom der Ordnung 2

$$\begin{aligned} f(0) &+ (D_1 f)(0)h_1 + (D_2 f)(0)h_2 + \frac{1}{2} (D_1^2 f)(0)h_1^2 \\ &+ \frac{1}{2} (D_2^2 f)(0)h_2^2 + (D_1 D_2 f)(0)h_1 h_2, \end{aligned}$$

der Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x_1, x_2) \mapsto e^{x_1^2 + \cos x_2}$ im Punkt $x = (0, 0)$ aufzustellen, bestimmen wir die partiellen Ableitungen von f bis zur 2. Ordnung:

$$\begin{aligned} (D_1 f)(x_1, x_2) &= 2x_1 e^{x_1^2 + \cos x_2} && \Rightarrow (D_1 f)(0) = 0, \\ (D_2 f)(x_1, x_2) &= -\sin x_2 e^{x_1^2 + \cos x_2} && \Rightarrow (D_2 f)(0) = 0, \\ (D_1^2 f)(x_1, x_2) &= (2 + 4x_1^2) e^{x_1^2 + \cos x_2} && \Rightarrow (D_1^2 f)(0) = 2e, \\ (D_2^2 f)(x_1, x_2) &= (-\cos x_2 + \sin^2 x_2) e^{x_1^2 + \cos x_2} && \Rightarrow (D_2^2 f)(0) = -e, \\ (D_1 D_2 f)(x_1, x_2) &= -2x_1 \sin x_2 e^{x_1^2 + \cos x_2} && \Rightarrow (D_1 D_2 f)(0) = 0. \end{aligned}$$

Das gesuchte Taylorpolynom ist somit $(h_1, h_2) \mapsto e + eh_1^2 - \frac{e}{2} h_2^2$. ■

Beweis von Satz 10.21. Die Funktion $g : [0, 1] \rightarrow \mathbb{R}$, $t \mapsto f(x + th)$ ist nach Satz 10.20 $(k+1)$ -mal stetig differenzierbar. Der Satz von Taylor für Funktionen einer Veränderlichen behauptet die Existenz eines $\tau \in (0, 1)$ so, dass

$$f(x+h) = g(1) = \sum_{m=0}^k \frac{g^{(m)}(0)}{m!} + \frac{g^{(k+1)}(\tau)}{(k+1)!}.$$

Wieder nach Satz 10.20 ist

$$\frac{g^{(m)}(0)}{m!} = \sum_{|\alpha|=m} \frac{(D^\alpha f)(x)}{\alpha!} h^\alpha \quad \text{sowie} \quad \frac{g^{(k+1)}(\tau)}{(k+1)!} = \sum_{|\alpha|=k+1} \frac{(D^\alpha f)(x + \tau h)}{\alpha!} h^\alpha,$$

woraus die Behauptung folgt. ■

Anmerkung 1. Seien α, β Multiindizes gleicher Länge. Man rechnet leicht nach, dass für $f(x) := x^\beta$ gilt

$$(D^\alpha f)(x) = \begin{cases} \frac{\beta!}{(\beta - \alpha)!} x^{\beta - \alpha} & \text{falls } \beta \geq \alpha, \\ 0 & \text{sonst} \end{cases}$$

($\beta - \alpha$ und $\beta \geq \alpha$ ist jeweils komponentenweise zu verstehen). Insbesondere ist

$$(D^\alpha f)(0) = \begin{cases} \beta! & \text{falls } \alpha = \beta \\ 0 & \text{sonst.} \end{cases}$$

Hieraus folgt, dass wie bei Funktionen einer Veränderlichen alle partiellen Ableitungen in x bis zur k . Ordnung der Funktion f mit denen ihres Taylorpolynoms der Ordnung k übereinstimmen. ■

Anmerkung 2. Man kann den Satz von Taylor auch für vektorwertige Funktionen $f = (f_1, \dots, f_m)^T : U \rightarrow \mathbb{R}^m$ formulieren und beweisen. Definieren wir für solche Funktionen $D_i f := (D_i f_1, \dots, D_i f_m)^T : U \rightarrow \mathbb{R}^m$, so sieht das entsprechende Taylorpolynom (welches nun ein Vektor ist) formal genauso aus wie in Satz 10.21. Für $m > 1$ muss jedoch das Restglied modifiziert werden (vgl. Abschnitt 10.5 für den Mittelwertsatz). ■

Wir sehen uns die Polynome $P_m(h) := \sum_{|\alpha|=m} \frac{(D^\alpha f)(x)}{\alpha!} h^\alpha$ für $m = 0, 1, 2$ genauer an.

$m = 0$. Notwendigerweise ist $\alpha = (0, \dots, 0)$ und daher $P_0(h) = f(x)$.

$m = 1$. Die einzigen n -Tupel $\alpha \in \mathbb{N}^n$ mit $|\alpha| = 1$ sind die "Einheitsvektoren" $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ mit der 1 an der j . Stelle. Wegen $D^{e_j} f = D_j f$, $e_j! = 1$ und $h^{e_j} = h_j$ ist

$$P_1(h) = \sum_{j=1}^n (D_j f)(x) h_j = \langle (\text{grad } f)(x), h \rangle.$$

$m = 2$. Wir haben im Beweis von Satz 10.20 gesehen, dass

$$P_2(h) = \sum_{|\alpha|=2} \frac{(D^\alpha f)(x)}{\alpha!} h^\alpha = \frac{1}{2} \sum_{i,j=1}^n (D_i D_j f)(x) h_i h_j.$$

Um dies kompakter zu schreiben, bezeichnen wir für jede zweimal stetig partiell differenzierbare Funktion $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$ die Matrix

$$\left((D_i D_j f)(x) \right)_{i,j=1}^n = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{i,j=1}^n$$

mit $(\text{Hess } f)(x)$ und nennen sie die *Hesse-Matrix* oder den *Hessian* von f in x . Nach dem Satz von Schwarz ist die Matrix $(\text{Hess } f)(x)$ symmetrisch. Mit der Hesse-Matrix können wir $P_2(h)$ schreiben als

$$P_2(h) = \frac{1}{2} \langle (\text{Hess } f)(x) h, h \rangle.$$

Folgerung 10.22 Ist $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ dreimal stetig partiell differenzierbar und $x + th \in U$ für alle $t \in [0, 1]$, so ist

$$f(x + h) = c + \langle a, h \rangle + \frac{1}{2} \langle Ah, h \rangle + R_2(x, h)$$

mit $c = f(x)$, $a = (\text{grad } f)(x)$, $A = (\text{Hess } f)(x)$ und einem Restglied R_2 wie im Satz 10.21.

10.7 Lokale Extrema

Wir benutzen nun Folgerung 10.22 zur Untersuchung des lokalen Verhaltens von Funktionen $f : \mathbb{R}^n \supseteq U \rightarrow \mathbb{R}$. Für offenes $U \subseteq \mathbb{R}^n$ sei $C^k(U)$ die Menge aller k -mal stetig partiell differenzierbaren Funktionen $f : U \rightarrow \mathbb{R}$.

Eine Funktion $f : U \rightarrow \mathbb{R}$ besitzt in $x_0 \in U$ ein *lokales Minimum* (bzw. ein *lokales Maximum*), wenn für alle x aus einer Umgebung $V \subseteq U$ von x_0 gilt

$$f(x_0) \leq f(x) \quad (\text{bzw.} \quad f(x_0) \geq f(x)).$$

Tritt die Gleichheit nur für $x = x_0$ ein, nennen wir x_0 ein *isoliertes* lokales Minimum (bzw. Maximum).

Satz 10.23 (Notwendige Bedingung) Sei U offen und $f : U \rightarrow \mathbb{R}$ partiell differenzierbar. Besitzt f in $x_0 \in U$ ein lokales Extremum (Minimum oder Maximum), so ist $(\text{grad } f)(x_0) = 0$.

Beweis. Für $i = 1, \dots, n$ betrachten wir die Funktionen $g_i : t \mapsto f(x_0 + te_i)$. Diese sind auf einem Intervall $(-\varepsilon, \varepsilon)$ mit $\varepsilon > 0$ definiert und differenzierbar, und sie besitzen in $t = 0$ ein lokales Extremum. Nach Lemma 7.9 ist $g_i'(0) = 0$. Nun ist aber $g_i'(0) = \frac{\partial f}{\partial x_i}(x_0)$. Also ist $(\text{grad } f)(x_0) = 0$. ■

Das Verschwinden aller partiellen Ableitungen in x_0 ist also eine *notwendige* Bedingung für das Vorliegen eines lokalen Extremums. Wie bei Funktionen einer Veränderlichen erhält man hinreichende Bedingungen durch Betrachten der zweiten Ableitungen. Wir treffen dazu einige Vorbereitungen.

Definition 10.24 Sei A eine symmetrische ($A = A^T$) reelle $n \times n$ Matrix. A heißt

- (a) positiv definit, wenn $\langle Ax, x \rangle > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$.
- (b) positiv semidefinit, wenn $\langle Ax, x \rangle \geq 0$ für alle $x \in \mathbb{R}^n$.
- (c) negativ definit (semidefinit), wenn $-A$ positiv definit (semidefinit) ist.
- (d) indefinit, wenn es $x, y \in \mathbb{R}^n$ gibt mit $\langle Ax, x \rangle > 0$ und $\langle Ay, y \rangle < 0$.

Aus der linearen Algebra kennen wir Kriterien für die Definitheit. So gilt:

Satz 10.25 Sei A eine symmetrische reelle $n \times n$ - Matrix. Dann ist A

- (a) positiv definit \Leftrightarrow alle Eigenwerte von A sind positiv.
- (b) positiv semidefinit \Leftrightarrow alle Eigenwerte von A sind nichtnegativ.
- (c) indefinit \Leftrightarrow A hat sowohl positive als auch negative Eigenwerte.

Satz 10.26 (Hurwitz-Kriterium) Eine symmetrische reelle $n \times n$ - Matrix $A = (a_{ij})_{i,j=1}^n$ ist genau dann positiv definit, wenn

$$\det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0 \quad \text{für alle } k = 1, \dots, n.$$

Wir formulieren nun hinreichende Kriterien für das Vorliegen von Extremwerten.

Satz 10.27 Sei $U \subseteq \mathbb{R}^n$ offen, $f \in C^2(U)$, $x_0 \in U$ und $(\text{grad } f)(x_0) = 0$. Dann gilt: Ist $(\text{Hess } f)(x_0)$

- (a) positiv definit, so hat f in x_0 ein isoliertes Minimum,
- (b) negativ definit, so hat f in x_0 ein isoliertes Maximum,
- (c) indefinit, so besitzt f in x_0 kein lokales Extremum.

Ist $(\text{Hess } f)(x_0)$ nur semidefinit, so ist keine Entscheidung möglich.

Beweis. Wir zeigen nur Aussage (a) und nehmen der Einfachheit halber $f \in C^3(U)$ an. Nach Folgerung 10.22 gilt für alle $x_0 + x$ aus einer Umgebung von x_0

$$f(x_0 + x) = f(x_0) + \frac{1}{2} \langle Ax, x \rangle + R(x) \quad (10.21)$$

mit $A := (\text{Hess } f)(x_0)$. Da $S := \{y \in \mathbb{R}^n : \|y\| = 1\}$ kompakt ist, nimmt die stetige Funktion

$$S \rightarrow \mathbb{R}, \quad y \mapsto \langle Ay, y \rangle$$

auf S ihr Minimum α an. Da $\langle Ay, y \rangle > 0$ für alle $y \in S$, ist insbesondere $\alpha > 0$. Ist nun $x \in \mathbb{R}^n \setminus \{0\}$, so ist $x/\|x\| \in S$ und folglich

$$\left\langle A \frac{x}{\|x\|}, \frac{x}{\|x\|} \right\rangle \geq \alpha > 0 \quad \text{bzw.} \quad \langle Ax, x \rangle \geq \alpha \|x\|^2. \quad (10.22)$$

Die letztere Abschätzung gilt offenbar auch für $x = 0$ und damit für alle $x \in \mathbb{R}^n$. (Umgekehrt folgt aus dieser Abschätzung natürlich die positive Definitheit von A ; beide Aussagen sind also äquivalent.)

Weiter: für das Restglied $R(x)$ in (10.21) haben wir wegen $|x_i| \leq \|x\|$ offenbar die Abschätzung

$$|R(x)| = \left| \sum_{|\alpha|=3} \frac{(D^\alpha f)(x_0 + \tau x)}{\alpha!} x^\alpha \right| \leq C \|x\|^3$$

für alle x aus einer hinreichend kleinen Umgebung von 0. Wählen wir diese Umgebung so klein, dass auch $C\|x\| \leq \alpha/4$, so wird

$$|R(x)| \leq C\|x\|^3 = C\|x\| \|x\|^2 \leq \frac{\alpha}{4}\|x\|^2.$$

Hieraus und aus (10.21) und (10.22) folgt schließlich

$$f(x_0 + x) \geq f(x_0) + \frac{\alpha}{2}\|x\|^2 - \frac{\alpha}{4}\|x\|^2 = f(x_0) + \frac{\alpha}{4}\|x\|^2$$

für alle x aus einer (hinreichend kleinen) Umgebung von 0. Also besitzt f in x_0 ein isoliertes lokales Minimum. ■

Beispiel. Für $f(x, y) = x^2 + \frac{1}{2}y^2 + xy + 10$ ist $(\text{grad } f)(0, 0) = 0$, und die Matrix

$$(\text{Hess } f)(0, 0) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

ist positiv definit (Hurwitz-Kriterium). Also besitzt f in $(0, 0)$ ein lokales isoliertes Minimum. ■

10.8 Parameterabhängige Integrale

In diesem Abschnitt betrachten wir folgendes Problem: Wird eine Funktion f zweier Veränderlicher bzgl. einer der Veränderlichen integriert, so hängt das Ergebnis von der zweiten Veränderlichen ab. So ist für $f(x, y) = x^y$ mit $x, y > 0$

$$\int_1^2 f(x, y) dx = \int_1^2 x^y dx = \frac{x^{y+1}}{y+1} \Big|_1^2 = \frac{2^{y+1} - 1}{y+1}.$$

Die Frage ist, unter welchen Voraussetzungen an f diese Abhängigkeit stetig oder sogar differenzierbar ist.

Satz 10.28 Sei $U \subseteq \mathbb{R}^n$ offen und $D = [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall. Die Funktion $f : D \times U \rightarrow \mathbb{R}$ sei stetig. Dann ist auch die Funktion

$$F : U \rightarrow \mathbb{R}, \quad x \mapsto \int_a^b f(t, x) dt$$

stetig. Hat f außerdem stetige partielle Ableitungen $\frac{\partial f}{\partial x_i} : D \times U \rightarrow \mathbb{R}$, $i = 1, \dots, n$, so ist auch F stetig partiell differenzierbar, und es gilt

$$\frac{\partial}{\partial x_i} \int_a^b f(t, x) dt = \int_a^b \frac{\partial f}{\partial x_i}(t, x) dt.$$

Unter den getroffenen Voraussetzungen dürfen Integration und Differentiation also vertauscht werden.

Beweis. Wir überlegen uns zuerst die Stetigkeit von F . Sei $x \in U$. Da U offen ist, gibt es ein $r > 0$ so, dass $U_{2r}(x) \subseteq U$. Dann liegt aber erst recht die abgeschlossene Kugel $\overline{U_r(x)} = \{y \in \mathbb{R}^n : \|x - y\| \leq r\}$ in U . Weiter: die Menge

$$D \times \overline{U_r(x)} = [a, b] \times \{y \in \mathbb{R}^n : \|x - y\| \leq r\}$$

ist abgeschlossen und beschränkt in $\mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$, also kompakt. Daher ist f auf dieser Menge sogar gleichmäßig stetig (Satz 6.41). Insbesondere gibt es zu jedem $\varepsilon > 0$ ein $\delta \in (0, r)$ so, dass

$$|f(t, x + h) - f(t, x)| < \frac{\varepsilon}{b - a}$$

für alle $t \in [a, b]$ und alle $h \in \mathbb{R}^n$ mit $\|h\| < \delta$. Integration liefert

$$|F(x + h) - F(x)| \leq \int_a^b |f(t, x + h) - f(t, x)| dt \leq \frac{\varepsilon}{b - a} (b - a) = \varepsilon$$

für alle h mit $\|h\| < \delta$. Also ist F in x stetig.

Für den Beweis der zweiten Aussage sei wieder $x \in U$ beliebig. Wir wählen $r \in \mathbb{R}$ so, dass $x + he_i \in U$ für alle $h \in (-r, r)$. Wir zeigen, dass die Funktion

$$g(t, x, h) := \begin{cases} \frac{f(t, x + he_i) - f(t, x)}{h} & \text{falls } h \neq 0, \\ \frac{\partial f}{\partial x_i}(t, x) & \text{falls } h = 0 \end{cases}$$

auf der Menge

$$\tilde{U} := \{(t, x, h) \in [a, b] \times U \times (-r, r)\}$$

stetig ist. In allen Punkten $(t, x, h) \in \tilde{U}$ mit $h \neq 0$ ist dies klar. Wir zeigen noch die Stetigkeit in allen Punkten $(t, x, 0) \in \tilde{U}$. Seien (t_n, x_n, h_n) Punkte aus \tilde{U} mit $\lim_{n \rightarrow \infty} (t_n, x_n, h_n) = (t, x, 0)$ und $h_n \neq 0$ (für $h_n = 0$ ist die Aussage wieder offensichtlich). Nach dem Mittelwertsatz für Funktionen einer Veränderlichen, angewandt auf die Funktion

$$a : s \mapsto f(t_n, x_n + sh_n e_i),$$

gibt es ein $\xi_n \in [0, 1]$ mit

$$\frac{a(1) - a(0)}{h_n} = \frac{f(t_n, x_n + h_n e_i) - f(t_n, x_n)}{h_n} = \frac{\partial f}{\partial x_i}(t_n, x_n + \xi_n h_n e_i).$$

Für $n \rightarrow \infty$ ist $h_n \rightarrow 0$ und folglich auch $\xi_n h_n \rightarrow 0$. Wegen der Stetigkeit von $\frac{\partial f}{\partial x_i}$ ist daher

$$\begin{aligned} g(t_n, x_n, h_n) &= \frac{f(t_n, x_n + h_n e_i) - f(t_n, x_n)}{h_n} \\ &= \frac{\partial f}{\partial x_i}(t_n, x_n + \xi_n h_n e_i) \rightarrow \frac{\partial f}{\partial x_i}(t, x) = g(t, x, 0). \end{aligned}$$

Somit ist g stetig. Wenden wir die Aussage des ersten Teils des Satzes an, erhalten wir

$$\begin{aligned} \frac{\partial F}{\partial x_i}(x) &= \frac{\partial}{\partial x_i} \int_a^b f(t, x) dt = \lim_{h \rightarrow 0} \int_a^b g(t, x, h) dt \\ &= \int_a^b g(t, x, 0) dt = \int_a^b \frac{\partial f}{\partial x_i}(t, x) dt. \end{aligned}$$

Schließlich hängt diese Funktion – wieder nach dem ersten Teil des Satzes – stetig von x ab. ■

Beispiel. Für $|t| < 1$ berechnen wir das Integral

$$F(t) := \int_0^\pi \ln(1 - 2t \cos x + t^2) dx.$$

Um Satz 10.28 benutzen zu können, wählen wir ein $a \in (|t|, 1)$ und betrachten F auf dem kompakten Intervall $[-a, a]$. (Man beachte, dass $1 - 2t \cos x + t^2 > 0$ für $|t| < 1$.) Die Funktion

$$f(x, t) := \ln(1 - 2t \cos x + t^2)$$

ist nach t stetig partiell differenzierbar und hat die Ableitung

$$\frac{\partial f}{\partial t} = \frac{2t - 2 \cos x}{1 - 2t \cos x + t^2} \quad \text{auf } [0, \pi] \times [-a, a].$$

Aus der zweiten Aussage von Satz 10.28 folgt

$$F'(t) = \int_0^\pi \frac{2t - 2 \cos x}{1 - 2t \cos x + t^2} dx.$$

Offenbar ist $F'(0) = 0$. Für $t \neq 0$ substituieren wir

$$s := \tan \frac{x}{2} \quad \text{bzw.} \quad x = 2 \arctan s.$$

Dann ist $\frac{dx}{ds} = \frac{2}{1+s^2}$ bzw. $dx = \frac{2ds}{1+s^2}$ sowie

$$\cos x = \frac{\cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}}{\cos^2 \frac{x}{2} + \sin^2 \frac{x}{2}} = \frac{1 - \tan^2 \frac{x}{2}}{1 + \tan^2 \frac{x}{2}} = \frac{1 - s^2}{1 + s^2}.$$

Das gesuchte Integral geht damit über in

$$\int_0^\infty \frac{2t - 2\frac{1-s^2}{1+s^2}}{1 - 2t\frac{1-s^2}{1+s^2} + t^2} \cdot \frac{2}{1+s^2} ds = 4 \int_0^\infty \frac{s^2(1+t) - (1-t)}{(1+t)^2 s^4 + 2(1+t^2)s^2 + (1-t)^2} ds.$$

Partialbruchzerlegung liefert

$$4 \frac{s^2(1+t) - (1-t)}{(1+t)^2 s^4 + 2(1+t^2)s^2 + (1-t)^2} = \frac{2}{t} \left(\frac{1}{s^2+1} + \frac{t^2-1}{(1+t)^2 s^2 + (1-t)^2} \right).$$

Mit dem bekannten Integral

$$\int \frac{1}{ax^2+c} dx = \frac{1}{\sqrt{ac}} \arctan \sqrt{\frac{a}{c}} x \quad \text{für } ac > 0$$

erhalten wir

$$F'(t) = \frac{2}{t} \left(\arctan s - \arctan\left(\frac{1+t}{1-t}s\right) \right) \Big|_0^\infty = 0$$

(beachten Sie: $\arctan 0 = 0$ und $\lim_{s \rightarrow \infty} \arctan s = \pi/2$). Also ist $F'(t) = 0$ für alle $t \in (-1, 1)$, d.h. F eine konstante Funktion. Wir bestimmen ihren (einzigsten) Wert, indem wir im Ausgangsintegral $t = 0$ setzen:

$$F(t) = F(0) = \int_0^\pi \ln 1 dx = 0. \quad \blacksquare$$

Als weitere Anwendung von Satz 10.28 betrachten wir iterierte Integrale und ihre Berechnung. Sei $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ eine stetige Funktion. Nach Satz 10.28 ist die Funktion

$$F(y) := \int_a^b f(x, y) dx$$

auf $[c, d]$ stetig und kann folglich integriert werden. Man bezeichnet

$$\int_c^d F(y) dy = \int_c^d \int_a^b f(x, y) dx dy$$

als *iteriertes Integral*. Der folgende Satz sagt, dass es auf die Reihenfolge der Integrationen nicht ankommt.

Satz 10.29 Sei $f : [a, b] \times [c, d] \rightarrow \mathbb{R}$ stetig. Dann ist

$$\int_c^d \int_a^b f(x, y) dx dy = \int_a^b \int_c^d f(x, y) dy dx.$$

Beweis. Wir erklären $\varphi : [c, d] \rightarrow \mathbb{R}$ durch

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

Dann ist $\varphi(c) = 0$, φ ist nach Satz 10.28 differenzierbar, und es gilt

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

Hieraus folgt

$$\int_c^d \int_a^b f(x, y) dx dy = \int_c^d \varphi'(y) dy = \varphi(d) - \varphi(c) = \varphi(d) = \int_a^b \int_c^d f(x, y) dy dx,$$

was wir zeigen wollten. ■

Ein analoger Satz gilt für n -fache Integrale einer stetigen Funktion auf einem Quader im \mathbb{R}^n .

Abschließend betrachten wir noch uneigentliche Parameterintegrale der Gestalt

$$\int_a^\infty f(x, y) dx, \quad y \in [c, d].$$

Da das Integrationsintervall nicht mehr kompakt ist, lässt sich der Beweis von Satz 10.28 nicht unmittelbar übertragen, und wir benötigen stärkere Voraussetzungen.

Definition 10.30 *Das Integral $\int_a^\infty f(x, y) dx$ heißt auf $[c, d]$ gleichmäßig konvergent, wenn es für jedes $y \in [c, d]$ konvergiert und wenn für jedes $\varepsilon > 0$ ein $b_0 \geq a$ existiert, so dass*

$$\left| \int_b^\infty f(x, y) dx \right| < \varepsilon \quad \text{für alle } b \geq b_0 \text{ und alle } y \in [c, d].$$

Man beachte, dass b_0 unabhängig von y ist.

Satz 10.31 (a) *Sei $f : [a, \infty) \times [c, d] \rightarrow \mathbb{R}$ stetig und beschränkt, und das Integral $\int_a^\infty f(x, y) dx$ konvergiere gleichmäßig auf $[c, d]$. Dann ist die Funktion*

$$\varphi : [c, d] \rightarrow \mathbb{R}, \quad y \mapsto \int_a^\infty f(x, y) dx \tag{10.23}$$

stetig. Außerdem konvergiert das Integral $\int_a^\infty \int_c^d f(x, y) dy dx$, und es gilt

$$\int_c^d \int_a^\infty f(x, y) dx dy = \int_a^\infty \int_c^d f(x, y) dy dx.$$

(b) *Sei $f : [a, \infty) \times [c, d] \rightarrow \mathbb{R}$ stetig, beschränkt und nach y partiell differenzierbar, die Ableitung $\frac{\partial f}{\partial y}$ sei auf $[a, \infty) \times [c, d]$ stetig und beschränkt, das Integral $\int_a^\infty \frac{\partial f}{\partial y}(x, y) dx$ konvergiere gleichmäßig, und das Integral $\int_a^\infty f(x, c) dx$ sei konvergent. Dann konvergiert das Integral $\int_a^\infty f(x, y) dx$ für jedes $y \in [c, d]$, die Funktion (10.23) ist differenzierbar, und*

$$\varphi'(y) = \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx.$$

Beweis. (a) Für alle $y, y + h \in [c, d]$ und $b \geq a$ ist

$$\begin{aligned}\varphi(y + h) - \varphi(y) &= \int_a^\infty (f(x, y + h) - f(x, y)) dx & (10.24) \\ &= \int_a^b (f(x, y + h) - f(x, y)) dx + \int_b^\infty f(x, y + h) dx - \int_b^\infty f(x, y) dx.\end{aligned}$$

Sei $\varepsilon > 0$. Wir wählen b so groß, dass der Betrag der letzten beiden Integrale in (10.24) jeweils kleiner als $\varepsilon/3$ wird (gleichmäßige Konvergenz!). Weiter wissen wir aus Satz 10.28, dass $y \mapsto \int_a^b f(x, y) dx$ eine stetige Funktion ist. Daher wird auch das erste Integral in (10.24) kleiner als $\varepsilon/3$, wenn nur h hinreichend klein ist, etwa für $h \leq h_0$. Für alle $h \leq h_0$ ist also

$$|\varphi(y + h) - \varphi(y)| < \varepsilon.$$

Für die zweite Aussage von (a) sei wieder $\varepsilon > 0$. Wir wählen $b_0 \geq a$ so, dass

$$\left| \int_a^\infty f(x, y) dx - \int_a^{b_0} f(x, y) dx \right| < \varepsilon \quad \text{für alle } y \in [c, d] \text{ und } b \geq b_0$$

(gleichmäßige Konvergenz!). Integrieren liefert

$$\left| \int_c^d \int_a^\infty f(x, y) dx dy - \int_c^d \int_a^{b_0} f(x, y) dx dy \right| < \varepsilon(d - c).$$

Mit Satz 10.29 vertauschen wir die Integrationsreihenfolge im 2. Integral und erhalten

$$\left| \int_c^d \int_a^\infty f(x, y) dx dy - \int_a^{b_0} \int_c^d f(x, y) dy dx \right| < \varepsilon(d - c).$$

Da dies für jedes $b \geq b_0$ gilt, folgt die Behauptung.

(b) Nach Teil (a) konvergiert für jedes $t \in [c, d]$ das uneigentliche Integral

$$\int_a^\infty \int_c^t \frac{\partial f}{\partial y}(x, y) dy dx,$$

und es gilt

$$\int_c^t \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx dy = \int_a^\infty \int_c^t \frac{\partial f}{\partial y}(x, y) dy dx. \quad (10.25)$$

Da $\frac{\partial f}{\partial y}$ stetig ist, folgt mit dem Hauptsatz der Differential- und Integralrechnung

$$\int_c^t \frac{\partial f}{\partial y}(x, y) dy = f(x, t) - f(x, c).$$

Integration bzgl. x über $[a, \infty)$ liefert wegen (10.25) (und wegen der Konvergenz von $\int_a^\infty f(x, c) dx$ nach Voraussetzung) für $t \in [c, d]$

$$\varphi(t) = \int_a^\infty f(x, t) dx = \int_a^\infty f(x, c) dx + \int_c^t \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx dy.$$

Nun ist $y \mapsto \int_a^\infty \frac{\partial f}{\partial y}(x, y) dx$ stetig nach Teil (a). Wieder nach dem Hauptsatz der Differential- und Integralrechnung ist φ differenzierbar, und

$$\varphi'(t) = \int_a^\infty \frac{\partial f}{\partial y}(x, t) dx. \quad \blacksquare$$

Aufgabe. Zeigen Sie mit Hilfe dieses Satzes, dass die Eulersche Gammafunktion

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt, \quad x \geq 1$$

stetig und sogar unendlich oft differenzierbar ist. (Für $x \in (0, 1)$ ist der Integrand unbeschränkt und Satz 10.31 nicht mehr unmittelbar anwendbar.)

11 Kurvenintegrale

Wir haben bisher ausschließlich Integrale über Intervallen betrachtet. Ein Ziel dieses Kapitels ist es, Integrale über Kurven zu erklären. Besonders interessiert uns die Frage, wann ein solches Integral nur vom Anfangs- und Endpunkt der Kurve abhängt.

11.1 Wege und Kurven

Unter einem *Weg* im \mathbb{R}^n verstehen wir eine stetige Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^n$. Die Punkte $\gamma(a)$ und $\gamma(b)$ heißen *Anfangs-* bzw. *Endpunkt* des Weges. Ist $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ein Weg, so heißt sein Wertebereich

$$\Gamma := \{\gamma(t) \in \mathbb{R}^n : t \in [a, b]\}$$

die zugehörige *Kurve*. Man beachte: ein Weg γ ist eine Abbildung, die zugehörige Kurve eine Punktmenge. Man sagt auch, dass durch γ eine Parametrisierung der Kurve Γ gegeben ist.

Ein Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$, $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))^T$ heißt (*stetig*) *differenzierbar*, wenn jede seiner Komponenten $\gamma_i : [a, b] \rightarrow \mathbb{R}$ (stetig) differenzierbar ist. In diesem Fall heißt

$$\dot{\gamma}(t) = \gamma'(t) = (\gamma'_1(t), \dots, \gamma'_n(t))^T$$

die *Ableitung* (oder der Geschwindigkeitsvektor) von γ in t , und die Zahl

$$\|\dot{\gamma}(t)\|_2 = \left(\sum_{j=1}^n |\gamma'_j(t)|^2 \right)^{1/2}$$

die *Geschwindigkeit* von γ in t . Falls $\gamma'(t_0) \neq 0$, so beschreibt

$$g : \mathbb{R} \rightarrow \mathbb{R}^n, \quad t \mapsto \gamma(t_0) + \gamma'(t_0)t$$

die *Tangente* an γ im Punkt t_0 .

Beispiele. (a) Für jedes $n \in \mathbb{Z} \setminus \{0\}$ ist

$$\gamma_n : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad t \mapsto (\cos(nt), \sin(nt))^T$$

ein Weg. Alle Wege γ_n beschreiben die gleiche Kurve im \mathbb{R}^2 , nämlich die Einheitskreislinie.

(b) Eine Ellipse um den Ursprung mit den Hauptachsen a, b wird durch den Weg $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, $t \mapsto (a \cos t, b \sin t)^T$ parametrisiert.

(c) Für $a, b \in \mathbb{R}^n$ wird durch $\gamma : [0, 1] \rightarrow \mathbb{R}^n$, $t \mapsto a + t(b - a)$ ein Weg definiert. Die zugehörige Kurve ist die Strecke $[a, b]$.

(d) Die *Neilsche Parabel* $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto (t^2, t^3)$ ist überall differenzierbar, obwohl die zugehörige Kurve eine Spitze in 0 hat.

(e) Eine Schraubenlinie im \mathbb{R}^3 lässt sich durch den Weg $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3, t \mapsto (\cos t, \sin t, t)^T$ beschreiben.

(f) Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ definiert einen Weg $\gamma : [a, b] \rightarrow \mathbb{R}^2, t \mapsto (t, f(t))$. Die zugehörige Kurve ist der Graph der Funktion. ■

Wie das erste dieser Beispiele zeigt, kann ein Weg Teile einer Kurve mehrfach durchlaufen. Will man dies ausschließen, muss man verlangen, dass je zwei Punkten $t_1, t_2 \in [a, b]$ mit $t_1 \neq t_2$ unterschiedliche Punkte $\gamma(t_1), \gamma(t_2)$ entsprechen. Da wir auch geschlossene Wege betrachten wollen (d.h. solche mit $\gamma(a) = \gamma(b)$), nehmen wir den Anfangs- und Endpunkt von dieser Forderung aus.

Definition 11.1 Ein Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt ein *Jordanweg*, wenn für beliebige Punkte $s, t \in [a, b]$ mit $s < t$ und $\gamma(s) = \gamma(t)$ folgt: $s = a$ und $t = b$. Eine Kurve heißt *Jordankurve*, wenn sie durch einen Jordanweg beschrieben werden kann.

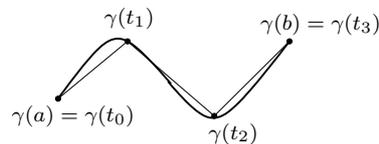
Mit möglicher Ausnahme ihrer Endpunkte sind Jordankurven also *doppelpunktfrei*. Die oben betrachteten Beispiele haben diese Eigenschaft.

Kurven (oder Wege) sind wesentlich kompliziertere Objekte als es unsere Anschauung erwarten lässt. So gibt es Kurven, die ein Quadrat im \mathbb{R}^2 komplett ausfüllen (Peano-Kurve), und Kurven, die in keinem Punkt eine Tangente besitzen (Koch'sche Schneeflocke). Umso bemerkenswerter ist der folgende Satz, der unserer Anschauung perfekt entspricht, dessen Beweis jedoch außerordentlich schwierig ist.

Satz 11.2 (Jordanscher Kurvensatz) Jede geschlossene Jordankurve $\Gamma \subseteq \mathbb{R}^2$ zerlegt den \mathbb{R}^2 in zwei Gebiete G_1, G_2 , die von ihr berandet werden (d.h. es ist $\mathbb{R}^2 = G_1 \cup \Gamma \cup G_2$ und $\partial G_1 = \partial G_2 = \Gamma$). Genau eines dieser Gebiete – es heißt das Innengebiet von Γ – ist beschränkt.

11.2 Rektifizierbare Wege und Bogenlänge

Wir wollen nun die Länge eines Weges $\gamma : [a, b] \rightarrow \mathbb{R}^n$ definieren und berechnen. Sei $Z := \{t_0, \dots, t_m\}$ mit $a = t_0 < t_1 < \dots < t_m = b$ eine Zerlegung von $[a, b]$. Wir verbinden für jedes i die Punkte $\gamma(t_i)$ und $\gamma(t_{i+1})$ durch eine Strecke und erhalten einen Polygonzug der Länge



$$L(Z, \gamma) := \sum_{i=0}^{m-1} \|\gamma(t_{i+1}) - \gamma(t_i)\|_2.$$

Wir erwarten, dass sich bei Verfeinerung von Z die Länge des Polygonzuges der „Länge von γ annähert“. Da sich bei Verfeinerung von Z die Zahl $L(Z, \gamma)$ niemals verkleinert, definieren wir:

Definition 11.3 Ein Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt rektifizierbar, wenn $\sup L(Z, \gamma)$ endlich ist, wobei das Supremum über alle Zerlegungen Z von $[a, b]$ genommen wird. In diesem Fall heißt $L(\gamma) := \sup L(Z, \gamma)$ die Weglänge von γ .

Beispiel. Der durch die stetige Funktion

$$f : [0, 1] \rightarrow \mathbb{R}, \quad t \mapsto \begin{cases} 0 & \text{für } t = 0, \\ t \cos \frac{\pi}{t} & \text{für } t \in (0, 1] \end{cases}$$

definierte Weg $\gamma : [0, 1] \rightarrow \mathbb{R}^2, t \mapsto (t, f(t))$ ist nicht rektifizierbar. Für die Punkte $t_n := 1/n$ gilt nämlich

$$\begin{aligned} \|\gamma(t_{n+1}) - \gamma(t_n)\|_2 &\geq |f(t_{n+1}) - f(t_n)| = \left| \frac{\cos((n+1)\pi)}{n+1} - \frac{\cos(n\pi)}{n} \right| \\ &= \left| \frac{(-1)^{n+1}}{n+1} - \frac{(-1)^n}{n} \right| = \frac{1}{n+1} + \frac{1}{n} > \frac{1}{n}, \end{aligned}$$

und die harmonische Reihe $\sum_{n=1}^{\infty} 1/n$ divergiert. ■

Für jeden Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ und jedes $c \in (a, b)$ sind auch $\gamma_1 := \gamma|_{[a, c]}$ und $\gamma_2 := \gamma|_{[c, b]}$ Wege. Man sieht leicht, dass γ genau dann rektifizierbar ist, wenn γ_1 und γ_2 rektifizierbar sind und dass in diesem Fall $L(\gamma) = L(\gamma_1) + L(\gamma_2)$ gilt.

Schließlich definieren wir die *Weglängenfunktion*

$$s : [a, b] \rightarrow \mathbb{R}, \quad t \mapsto \begin{cases} 0 & \text{für } t = a, \\ L(\gamma|_{[a, t]}) & \text{für } t \in (a, b]. \end{cases}$$

Satz 11.4 Für jeden rektifizierbaren Weg ist seine Weglängenfunktion stetig.

Einen Beweis finden Sie in Heuser, Analysis II, Satz 177.3. Unter stärkeren Voraussetzungen an γ wollen wir nun Weglängen berechnen.

Satz 11.5 Der Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ sei stetig differenzierbar. Dann ist γ rektifizierbar, die Weglängenfunktion s von γ ist stetig differenzierbar, und für alle $t \in [a, b]$ gilt $s'(t) = \|\gamma'(t)\|_2$. Die Länge $L(\gamma)$ von γ ist gleich

$$L(\gamma) = \int_a^b s'(t) dt = \int_a^b \sqrt{\gamma'_1(t)^2 + \dots + \gamma'_n(t)^2} dt, \quad (11.1)$$

wobei $\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))^T$.

Beweis. Wir benutzen im Beweis Integrale von vektorwertigen Funktionen, die wir komponentenweise erklären.

Sei $Z = \{t_0, \dots, t_m\}$ eine Zerlegung von $[a, b]$, d.h. $a = t_0 < t_1 < \dots < t_m = b$. Mit der Dreiecksungleichung für Integrale ist

$$\|\gamma(t_{i+1}) - \gamma(t_i)\|_2 = \left\| \int_{t_i}^{t_{i+1}} \gamma'(t) dt \right\|_2 \leq \int_{t_i}^{t_{i+1}} \|\gamma'(t)\|_2 dt,$$

und Aufsummieren liefert

$$L(Z, \gamma) = \sum_{i=0}^{m-1} \|\gamma(t_{i+1}) - \gamma(t_i)\|_2 \leq \int_a^b \|\gamma'(t)\|_2 dt.$$

Alle Polygonzuglängen sind also durch eine von Z unabhängige Konstante nach oben beschränkt. Folglich ist γ rektifizierbar, und

$$L(\gamma) = \sup_Z L(Z, \gamma) \leq \int_a^b \|\gamma'(t)\|_2 dt. \quad (11.2)$$

Sei nun $t \in [a, b]$ und $h > 0$ so, dass $t + h \leq b$. Dann ist die Länge der Strecke von $\gamma(t)$ bis $\gamma(t+h)$ nicht größer als die Länge des Weges von $\gamma(t)$ bis $\gamma(t+h)$:

$$\|\gamma(t+h) - \gamma(t)\|_2 \leq s(t+h) - s(t).$$

Wenden wir (11.2) speziell auf den Weg $\gamma|_{[t, t+h]}$ an, folgt

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\|_2 \leq \frac{s(t+h) - s(t)}{h} \leq \frac{1}{h} \int_t^{t+h} \|\gamma'(r)\|_2 dr.$$

Für $h \searrow 0$ strebt die linke Seite dieser Abschätzung gegen $\|\gamma'(t)\|_2$. Die rechte Seite konvergiert nach dem Mittelwertsatz für Integrale gegen den gleichen Wert:

$$\frac{1}{h} \int_t^{t+h} \|\gamma'(r)\|_2 dr = \|\gamma'(\xi_h)\|_2 \quad \text{mit } t \leq \xi_h \leq t+h.$$

Also existiert die rechtsseitige Ableitung von s in t und ist gleich $\|\gamma'(t)\|$. Analog zeigt man die Differenzierbarkeit von links. Aus $s'(t) = \|\gamma'(t)\|$ folgt die Behauptung (11.1). ■

Beispiel 1 (Kreise und Ellipsen). Für den Weg

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad t \mapsto (a \cos t, a \sin t) \quad \text{mit } a > 0$$

finden wir

$$L(\gamma) = \int_0^{2\pi} \sqrt{\gamma_1'(t)^2 + \gamma_2'(t)^2} dt = \int_0^{2\pi} \sqrt{a^2(\sin^2 t + \cos^2 t)} dt = \int_0^{2\pi} a dt = 2\pi a.$$

Die durch γ beschriebene Kurve ist eine Kreislinie vom Radius a . Analog führt die Berechnung der Länge des Weges

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, t \mapsto (a \cos t, b \sin t) \quad \text{mit } a > b > 0$$

(dessen zugehörige Kurve eine Ellipse ist) auf das Integral

$$L(\gamma) = \int_0^{2\pi} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt = a \int_0^{2\pi} \sqrt{1 - \varepsilon^2 \cos^2 t} dt \quad (11.3)$$

mit $\varepsilon := \frac{1}{a} \sqrt{a^2 - b^2}$ (die sog. *numerische Exzentrizität* der Ellipse). Das Integral in (11.3) ist (für $\varepsilon > 0$) ein sog. *elliptisches Integral*. Es lässt sich nicht mit Hilfe elementarer Funktionen geschlossen darstellen. ■

Beispiel 2 (Funktionsgraphen). Ist $f : [a, b] \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und $\gamma : [a, b] \rightarrow \mathbb{R}^2, t \mapsto (t, f(t))$ der durch f induzierte Weg, so reduziert sich (11.1) auf

$$L(\gamma) = \int_a^b \sqrt{1 + f'(t)^2} dt. \quad \blacksquare$$

Haben wir in Beispiel 1 mit $L(\gamma) = 2\pi a$ tatsächlich den Kreisumfang (d.h. die Länge einer *Kurve*) berechnet? Was wir berechnet haben, ist die Länge eines *Weges*. Um hieraus zu einem vernünftigen Begriff einer *Kurvenlänge* zu gelangen, müssen wir zunächst dafür sorgen, dass der Weg jeden Teil der Kurve nur einmal durchläuft, d.h. wir betrachten ausschließlich Jordanwege bzw. Jordankurven. Selbst für Jordankurven ist damit die Kurvenlänge noch nicht eindeutig festgelegt. Es könnte ja sein, dass ein- und dieselbe Jordankurve durch verschiedene Jordanwege mit verschiedenen Weglängen parametrisiert werden kann. Der folgende Satz klärt dieses Problem.

Satz 11.6 *Sei γ eine Jordankurve, die eine Darstellung durch einen rektifizierbaren Jordanweg besitzt. Dann sind alle Jordandarstellungen rektifizierbar und haben ein- und dieselbe Weglänge.*

Die gemeinsame Weglänge nennen wir die *Länge einer Kurve* Γ . Erst mit diesem Satz können wir sagen, dass ein Kreis mit Radius a einen Umfang $2\pi a$ besitzt (und dass auch Ellipsen einen Umfang besitzen, auch wenn wir ihn nicht elementar angeben können).

Beweisidee. Wir zeigen die Aussage nur für stetig differenzierbare Wege. Einen Beweis für den allgemeinen Fall finden Sie in Heuser, Analysis II, Satz 178.3.

Zunächst eine Vorbemerkung. Ist $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ein Weg mit zugehöriger Kurve Γ und ist $\varphi : [\alpha, \beta] \rightarrow [a, b]$ eine stetige Bijektion, so ist $\gamma \circ \varphi : [\alpha, \beta] \rightarrow \mathbb{R}^n$ ebenfalls ein Weg, der auf die Kurve Γ führt. Da φ stetig und bijektiv ist, tritt einer der folgenden Fälle ein (Satz 6.42) :

(a) φ wächst streng monoton. Dann heißt φ *orientierungserhaltend*.

(b) φ fällt streng monoton. Dann heißt φ *orientierungsumkehrend*.

Sind insbesondere φ und $\varphi^{(-1)}$ stetig differenzierbar, so folgt aus $\varphi^{(-1)}(\varphi(t)) = t$ mit der Kettenregel $\varphi^{(-1)'(\varphi(t))} \cdot \varphi'(t) = 1$, d.h. es ist $\varphi'(t) \neq 0$ für alle $t \in [\alpha, \beta]$. Es ist klar, dass φ genau dann orientierungserhaltend (bzw. -umkehrend) ist, wenn $\varphi'(t) > 0$ (bzw. < 0) für alle $t \in [\alpha, \beta]$ ist.

Sei nun $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg und $\varphi : [\alpha, \beta] \rightarrow [a, b]$ eine stetig differenzierbare und orientierungserhaltende Bijektion. Dann ist

$$\begin{aligned} L(\gamma \circ \varphi) &= \int_{\alpha}^{\beta} \|(\gamma \circ \varphi)'(t)\|_2 dt = \int_{\alpha}^{\beta} \|\gamma'(\varphi(t)) \varphi'(t)\|_2 dt \\ &= \int_{\alpha}^{\beta} \|\gamma'(\varphi(t))\|_2 \varphi'(t) dt = \int_a^b \|\gamma'(x)\|_2 dx = L(\gamma). \end{aligned}$$

Ein ähnlicher Beweis erfolgt für orientierungsumkehrendes φ . ■

Sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbar. Dann kann – wie wir gesehen haben – die Länge des Weges γ durch $\int_a^b \|\gamma'(t)\|_2 dt$ berechnet werden. Ist Γ die durch γ definierte Kurve und $f : \Gamma \rightarrow \mathbb{R}^m$ eine Funktion, für die $f \circ \gamma$ Riemann-integrierbar ist, so definiert man allgemeiner das *Integral von f entlang γ* durch

$$\int_{\gamma} f := \int_a^b f(\gamma(t)) \|\gamma'(t)\|_2 dt \quad (11.4)$$

(für $m > 1$ berechnen wir das Integral auf der rechten Seite komponentenweise). Insbesondere ist also $\int_{\gamma} 1 = L(\gamma)$, und ist $\gamma : [a, b] \rightarrow [a, b]$ die identische Abbildung, so ist

$$\int_{\gamma} f = \int_a^b f(t) dt$$

das übliche Riemann-Integral. Man kann – ähnlich wie im Beweis von Satz 11.6 – zeigen, dass $\int_{\gamma} f$ von der Parametrisierung von Γ unabhängig ist.

11.3 Wegintegrale

In die Definition des Kurvenintegrals (11.4) geht nur die Länge des Geschwindigkeitsvektors $\gamma'(t)$ ein, nicht seine Richtung. Dies ist für Anwendungen oft nicht ausreichend (man denke etwa an einen Körper, der sich entlang eines Weges γ in einem Kraftfeld bewegt und bei dem die verrichtete Arbeit berechnet werden soll). In diesem Abschnitt werden wir uns einen angemessenen Begriff eines Kurvenintegrals erarbeiten. Das folgende Beispiel soll die einzuführenden Begriffe motivieren.

Beispiel. Auf einer offenen Menge $U \subseteq \mathbb{R}^3$ sei ein Vektorfeld $F = (F_1, F_2, F_3) :$

$U \rightarrow \mathbb{R}^3$ gegeben, das wir uns als zeitlich konstantes Kraftfeld denken. Ist $\gamma = (\gamma_1, \gamma_2, \gamma_3) : [a, b] \rightarrow U$ ein stetig differenzierbarer Weg, so interpretieren wir das Integral

$$\int_a^b \langle F(\gamma(t)), \dot{\gamma}(t) \rangle dt = \int_a^b \sum_{i=1}^3 F_i(\gamma(t)) \gamma'_i(t) dt \quad (11.5)$$

als Arbeit, die man aufwenden muss, um sich vom Punkt $\gamma(a)$ zum Punkt $\gamma(b)$ entlang des Weges γ zu bewegen. Dies ist gerechtfertigt, da man für ein kleines Wegstück näherungsweise annehmen kann, dass F konstant ist und

$$\gamma(t) = \gamma(t_i) + \frac{t - t_i}{t_{i+1} - t_i} (\gamma(t_{i+1}) - \gamma(t_i))$$

gilt. Dann ist

$$\langle F(\gamma(t_i)), \dot{\gamma}(t_i) \rangle (t_{i+1} - t_i) = \langle F(\gamma(t_i)), \gamma(t_{i+1}) - \gamma(t_i) \rangle,$$

und dieser Ausdruck ist proportional zur Weglänge, zur Größe des Kraftfeldes und zum Kosinus des Winkels α zwischen Weg- und Kraftvektor:

$$\langle F(\gamma(t_i)), \gamma(t_{i+1}) - \gamma(t_i) \rangle = \cos \alpha \|F(\gamma(t_i))\| \|\gamma(t_{i+1}) - \gamma(t_i)\|.$$

Steht insbesondere das Kraftfeld senkrecht zur Wegrichtung, so wird keine Arbeit verrichtet. ■

Wir werden uns also mit Integralen der Gestalt (11.5) befassen müssen. Dazu treffen wir folgende Definition.

Definition 11.7 Seien $\gamma = (\gamma_1, \dots, \gamma_n) : [a, b] \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg mit zugehöriger Kurve Γ und $f = (f_1, \dots, f_n) : \Gamma \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Dann definieren wir das Wegintegral von f entlang γ durch

$$\int_{\gamma} f \cdot dx := \int_a^b \langle f(\gamma(t)), \dot{\gamma}(t) \rangle dt = \int_a^b \sum_{i=1}^n f_i(\gamma(t)) \gamma'_i(t) dt. \quad (11.6)$$

Anstelle von (11.6) findet man auch die Schreibweise

$$\int_{\gamma} f_1 dx_1 + \dots + f_n dx_n,$$

die man im Rahmen der Theorie der Pfaffschen Formen verstehen kann (vgl. Barner/Flohr, Analysis II, Abschnitt 17.1). Unter Verwendung von Riemann-Stieltjes-Integralen kann man $\int_{\gamma} f \cdot dx$ für beliebige rektifizierbare Wege γ und stetige Vektorfelder f auf Γ definieren und berechnen (vgl. Heuser, Analysis 2, Abschnitt 180).

Der Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt *stückweise stetig differenzierbar*, wenn es eine Zerlegung $a = x_0 < x_1 < \dots < x_m = b$ des Intervalles $[a, b]$ so gibt, dass die Wege

$$\gamma^{(i)} := \gamma|_{[x_i, x_{i+1}]} : [x_i, x_{i+1}] \rightarrow \mathbb{R}^n, \quad i = 0, \dots, m-1$$

stetig differenzierbar sind. Für stückweise stetig differenzierbare Wege γ und stetige Vektorfelder f definieren wir das Wegintegral durch

$$\int_{\gamma} f \cdot dx := \sum_{i=0}^{m-1} \int_{\gamma^{(i)}} f \cdot dx.$$

Die folgenden Eigenschaften von Wegintegralen sind leicht zu sehen. Dabei ist γ ein stückweise stetig differenzierbarer Weg, und f und g sind stetige Vektorfelder.

$$(a) \quad \int_{\gamma} (f + g) \cdot dx = \int_{\gamma} f \cdot dx + \int_{\gamma} g \cdot dx, \quad \int_{\gamma} (cf) \cdot dx = c \int_{\gamma} f \cdot dx.$$

(b) Für jeden Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ bezeichne $\gamma^- : [a, b] \rightarrow \mathbb{R}^n$, $t \mapsto \gamma(a + b - t)$ den *entgegengesetzten Weg*. Offenbar beschreiben γ und γ^- die gleiche Kurve, die in unterschiedlicher Richtung durchlaufen wird. Dann gilt

$$\int_{\gamma} f \cdot dx = - \int_{\gamma^-} f \cdot dx.$$

(c) Ist $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ein stückweise stetig differenzierbarer Weg und $c \in (a, b)$, so sind auch $\gamma_1 := \gamma|_{[a, c]}$ und $\gamma_2 := \gamma|_{[c, b]}$ stückweise stetig differenzierbare Wege, und es gilt

$$\int_{\gamma} f \cdot dx = \int_{\gamma_1} f \cdot dx + \int_{\gamma_2} f \cdot dx.$$

$$(d) \quad \left| \int_{\gamma} f \cdot dx \right| \leq \max_{t \in \Gamma} \|f(t)\|_2 \cdot L(\gamma).$$

Wir untersuchen nun, inwieweit $\int_{\gamma} f \cdot dx$ tatsächlich vom Weg γ oder nur von der durch γ definierten Kurve abhängt.

Satz 11.8 Seien $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg, $f : \Gamma \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld und $\varphi : [\alpha, \beta] \rightarrow [a, b]$ eine stetig differenzierbare Bijektion mit $\varphi(\alpha) = a$ und $\varphi(\beta) = b$. Dann ist $\gamma \circ \varphi : [\alpha, \beta] \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg mit der gleichen Bildkurve und den gleichen Anfangs- und Endpunkten wie γ , und es gilt

$$\int_{\gamma} f \cdot dx = \int_{\gamma \circ \varphi} f \cdot dx.$$

Beweis. Mit Ketten- und Substitutionsregel finden wir

$$\begin{aligned}
 \int_{\gamma \circ \varphi} f \cdot dx &= \int_{\alpha}^{\beta} \langle f((\gamma \circ \varphi)(t)), (\gamma \circ \varphi)'(t) \rangle dt \\
 &= \int_{\alpha}^{\beta} \langle f((\gamma \circ \varphi)(t)), \gamma'(\varphi(t))\varphi'(t) \rangle dt \\
 &= \int_{\alpha}^{\beta} \langle f((\gamma \circ \varphi)(t)), \gamma'(\varphi(t)) \rangle \varphi'(t) dt \\
 &= \int_a^b \langle f(\gamma(s)), \gamma'(s) \rangle ds = \int_{\gamma} f \cdot dx. \quad \blacksquare
 \end{aligned}$$

Analog zeigt man, dass für jede stetig differenzierbare Bijektion $\varphi : [\alpha, \beta] \rightarrow [a, b]$ mit $\varphi(\alpha) = b$ und $\varphi(\beta) = a$ gilt:

$$\int_{\gamma \circ \varphi} f \cdot dx = - \int_{\gamma} f \cdot dx.$$

Bei Umkehrung der Orientierung ändert das Kurvenintegral also sein Vorzeichen.

Beispiel 1. Ein Punkt P bewege sich auf dem Weg

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^3, \quad t \mapsto (a \cos t, a \sin t, \frac{h}{2\pi} t)$$

von $\gamma(0) = (a, 0, 0)$ nach $\gamma(2\pi) = (a, 0, h)$. Dabei wirke eine Kraft $F(x, y, z) = (F_1, F_2, F_3) = -\alpha(x, y, z)$ mit $\alpha > 0$. Für die geleistete Arbeit finden wir

$$\begin{aligned}
 \int_{\gamma} F \cdot dx &= \int_0^{2\pi} (F_1(\gamma(t))\gamma_1'(t) + F_2(\gamma(t))\gamma_2'(t) + F_3(\gamma(t))\gamma_3'(t)) dt \\
 &= \int_0^{2\pi} \left((-\alpha a \cos t)(-a \sin t) + (-\alpha a \sin t)(a \cos t) + \left(-\alpha \frac{h}{2\pi} t\right)\left(\frac{h}{2\pi}\right) \right) dt \\
 &= -\alpha \frac{h^2}{(2\pi)^2} \int_0^{2\pi} t dt = -\frac{\alpha h^2}{(2\pi)^2} \frac{(2\pi)^2}{2} = -\frac{\alpha h^2}{2}.
 \end{aligned}$$

Bewegt sich P unter Einfluss der gleichen Kraft entlang des Weges

$$\gamma : [0, h] \rightarrow \mathbb{R}^3, \quad t \mapsto (a, 0, t)$$

von $(a, 0, 0)$ nach $(a, 0, h)$, so ergibt sich wegen

$$\int_{\gamma} F \cdot dx = \int_0^h F_3(\gamma(t))\gamma_3'(t) dt = \int_0^h (-\alpha t) dt = -\frac{\alpha h^2}{2}$$

die gleiche geleistete Arbeit. ■

Dieses Resultat ist kein Zufall. Verantwortlich für die beobachtete Wegunabhängigkeit des Integrals ist eine spezielle Eigenschaft der Funktion F .

Definition 11.9 Sei $U \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : U \rightarrow \mathbb{R}^n$ heißt Gradientenfeld (oder vollständiges Differential), wenn es eine differenzierbare Funktion $\varphi : U \rightarrow \mathbb{R}$ gibt, so dass

$$f(x) = (\text{grad } \varphi)(x) \quad \text{für alle } x \in U. \quad (11.7)$$

Die Funktion φ heißt eine Stammfunktion von f (und in der Physik heißt $-\varphi$ ein Potential von f).

Beispiel 2. Die Funktion $f(x, y, z) = -\alpha(x, y, z)$ aus Beispiel 1 ist ein Gradientenfeld, da z.B. für $\varphi(x, y, z) = -\frac{\alpha}{2}(x^2 + y^2 + z^2)$ die Beziehung (11.7) gilt. Physikalisch interessanter ist folgendes Beispiel. Ist eine Masse m im Nullpunkt eines Koordinatensystems konzentriert, so übt sie auf einen Punkt mit der Masse 1, der sich in $(x, y, z) \in \mathbb{R}^3$ befindet, eine Kraft f der Stärke

$$\|f\| = G \frac{m}{\|(x, y, z)\|_2^2} = \frac{Gm}{x^2 + y^2 + z^2}$$

aus (Newtonsches Gravitationsgesetz). Diese Kraft weist zum Nullpunkt, hat also die Richtung $\frac{-(x, y, z)}{\|(x, y, z)\|_2}$. Daher ist

$$f(x, y, z) = -\frac{Gm}{\|(x, y, z)\|_2^3} (x, y, z) = -\frac{Gm}{(x^2 + y^2 + z^2)^{3/2}} (x, y, z).$$

Auch diese Funktion ist ein Gradientenfeld; für die Funktion

$$\varphi : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}, \quad (x, y, z) \mapsto \frac{Gm}{\sqrt{x^2 + y^2 + z^2}}$$

ist nämlich $\text{grad } \varphi = f$. ■

Satz 11.10 Sei $U \subseteq \mathbb{R}^n$ offen und $F : U \rightarrow \mathbb{R}$ stetig partiell differenzierbar. Sind $a, b \in U$ und ist γ ein stückweise stetig differenzierbarer Weg mit Anfangspunkt a und Endpunkt b , der komplett in U verläuft, so ist

$$\int_{\gamma} \text{grad } F \cdot dx = F(b) - F(a). \quad (11.8)$$

Das Wegintegral über ein stetiges Gradientenfeld längs eines stückweise glatten Weges hängt also nur von Anfangs- und Endpunkt, nicht vom konkreten Verlauf des Weges ab. Man kann Satz 11.10 als Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung betrachten.

Beweis. Sei zunächst γ stetig differenzierbar. Dann ist

$$\int_{\gamma} (\text{grad } F)(x) \cdot dx = \int_{\alpha}^{\beta} \langle (\text{grad } F)(\gamma(t)), \gamma'(t) \rangle dt.$$

Nach der Kettenregel ist $\langle (\text{grad } F)(\gamma(t)), \gamma'(t) \rangle$ gerade die Ableitung der Funktion $\phi(t) := F(\gamma(t)) = (F \circ \gamma)(t)$. Wir erhalten daher mit dem Hauptsatz der Differential- und Integralrechnung

$$\int_{\gamma} (\text{grad } F)(x) \cdot dx = \int_{\alpha}^{\beta} \phi'(t) dt = \phi(\beta) - \phi(\alpha).$$

Schließlich ist $\phi(\beta) - \phi(\alpha) = F(\gamma(\beta)) - F(\gamma(\alpha)) = F(b) - F(a)$. Der allgemeine Fall eines stückweise stetig differenzierbaren Weges ergibt sich durch Zusammensetzen der einzelnen Integrale. ■

Ist γ ein geschlossener Weg (d.h. ist $\gamma(\alpha) = \gamma(\beta)$) so erhalten wir insbesondere

$$\int_{\gamma} \text{grad } F \cdot dx = 0.$$

Beispiel 3. Auf $U := \mathbb{R}^2 \setminus \{0\}$ betrachten wir das Vektorfeld

$$f(x, y) := \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

und den geschlossenen Weg $\gamma : [0, 2\pi] \rightarrow U$, $t \mapsto (\cos t, \sin t)$. Wegen $\gamma'(t) = (-\sin t, \cos t)$ ist

$$\int_{\gamma} f \cdot dx = \int_0^{2\pi} \left(\frac{-\sin t}{\sin^2 t + \cos^2 t} (-\sin t) + \frac{\cos t}{\sin^2 t + \cos^2 t} (\cos t) \right) dt = \int_0^{2\pi} dt = 2\pi.$$

Die stetige Funktion f ist also kein Gradientenfeld, da das Integral über die geschlossene Kurve γ nicht verschwindet. Dies ist ein wichtiger Unterschied zu Funktionen einer Veränderlichen. Im \mathbb{R}^1 besitzt *jede* auf einem Intervall stetige Funktion eine Stammfunktion. ■

11.4 Ergänzungen zu zusammenhängenden Mengen

Dem in Abschnitt 6.8 entwickelten Zusammenhangsbegriff stellen wir einen zweiten gegenüber.

Definition 11.11 *Wir nennen einen metrischer Raum (X, d) wegzusammenhängend, wenn es für je zwei Punkte $x, y \in X$ einen Weg $\gamma : [a, b] \rightarrow X$ mit $\gamma(a) = x$ und $\gamma(b) = y$ gibt.*

Eine Teilmenge $U \subseteq \mathbb{R}^n$ heißt *konvex*, wenn sie mit je zwei Punkten x, y auch deren Verbindungsstrecke $[x, y] := \{\lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}$ enthält, und U heißt *sternförmig*, wenn es einen Punkt $z \in U$ so gibt, dass $[z, x] \subseteq U$ für jedes $x \in U$. Der Punkt z heißt dann ein *Zentrum* von U . Offenbar sind konvexe Mengen sternförmig, und jeder ihrer Punkte ist ein Zentrum. Sternförmige und insbesondere konvexe Mengen sind wegzusammenhängend.

Satz 11.12 *Wegzusammenhängende metrische Räume sind zusammenhängend.*

Beweis. Sei (X, d) ein wegzusammenhängender metrischer Raum, und seien U_1, U_2 nichtleere offene Teilmengen von X mit $U_1 \cap U_2 = \emptyset$ und $U_1 \cup U_2 = X$. Dann gibt es Punkte $u_1 \in U_1$ und $u_2 \in U_2$ und einen Weg $\gamma : [a, b] \rightarrow X$ mit $\gamma(a) = u_1$ und $\gamma(b) = u_2$. Sei Γ die durch γ definierte Kurve. Nach Satz 6.48 ist Γ zusammenhängend. Andererseits gilt

$$U_1 \cap \Gamma \neq \emptyset, \quad U_2 \cap \Gamma \neq \emptyset, \quad U_1 \cap U_2 = \emptyset \quad \text{und} \quad \Gamma \subseteq U_1 \cup U_2,$$

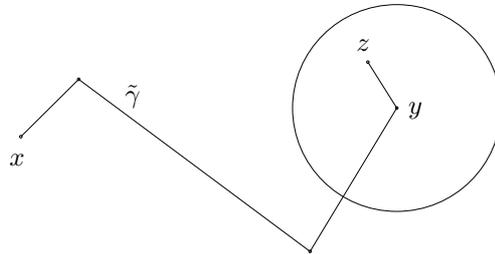
ein Widerspruch. ■

Die Umkehrung von Satz 11.12 gilt im Allgemeinen nicht. Die Menge

$$\{(0, y) \in \mathbb{R}^2 : y \in [-1, 1]\} \cup \{(x, y) \in \mathbb{R}^2 : x \in (0, 1], y = \sin \frac{1}{x}\}$$

ist zwar zusammenhängend, jedoch nicht wegzusammenhängend (HA). Die Umkehrung von Satz 11.12 gilt aber für offene Mengen. Eine zusammenhängende offene Menge heißt ein *Gebiet*.

Satz 11.13 *Gebiete im \mathbb{R}^n sind wegzusammenhängend. Genauer: ist $U \subseteq \mathbb{R}^n$ ein Gebiet und sind $x, y \in U$, dann gibt es einen Polygonzug $\gamma : [0, 1] \rightarrow U$ mit $\gamma(0) = x$ und $\gamma(1) = y$.*



Beweis. Sei $x \in U$. Wir bezeichnen mit U_x die Menge aller Punkte $y \in U$, für die es einen Polygonzug $\gamma : [0, 1] \rightarrow U$ mit $\gamma(0) = x$ und $\gamma(1) = y$ gibt. Wir zeigen zuerst, dass U_x offen ist.

Sei $y \in U_x$. Da U offen ist, gibt es eine Umgebung $U_\epsilon(y)$, die ganz in U liegt. Sei $z \in U_\epsilon(y)$, und sei $\tilde{\gamma} : [0, 1] \rightarrow U$ ein Polygonzug mit $\tilde{\gamma}(0) = x$ und $\tilde{\gamma}(1) = y$. Wir verlängern $\tilde{\gamma}$ um die Strecke $[y, z]$:

$$\gamma : [0, 1] \rightarrow U, \quad t \mapsto \begin{cases} \tilde{\gamma}(2t) & \text{für } t \in [0, 1/2], \\ y + (2t - 1)(z - y) & \text{für } t \in [1/2, 1]. \end{cases}$$

Dann ist γ ein Polygonzug in U , der x mit z verbindet. Da $z \in U_\epsilon(y)$ beliebig war, folgt $U_\epsilon(y) \subseteq U_x$, d.h. U_x ist offen.

Wir zeigen nun, dass auch $U \setminus U_x$ offen ist. Ist $y \in U \setminus U_x$, so gibt es wie oben eine Umgebung $U_\epsilon(y)$, die in U liegt. Läge ein Punkt $z \in U_\epsilon(y)$ in U_x , so könnten

wir wie oben den Polygonzug von x nach z zu einem Polygonzug von x nach y verlängern, d.h. es wäre $y \in U_x$. Dieser Widerspruch zeigt, dass $U_\varepsilon(y) \subseteq U \setminus U_x$, d.h. $U \setminus U_x$ ist offen.

Für die offenen Mengen U_x und $U \setminus U_x$ gilt nun

$$U_x \cap (U \setminus U_x) = \emptyset \quad \text{und} \quad U_x \cup (U \setminus U_x) = U.$$

Da U zusammenhängend ist, muss eine der Mengen U_x und $U \setminus U_x$ leer sein. Wegen $x \in U_x$ ist $U \setminus U_x$ leer. Somit ist $U_x = U$. ■

11.5 Stammfunktionen und Wegunabhängigkeit von Kurvenintegralen

Lemma 11.14 Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld, und $c \in \mathbb{R}$.

- (a) Ist F Stammfunktion von f , so ist auch $F + c$ Stammfunktion von f .
 (b) Ist U ein Gebiet, und sind F_1 und F_2 Stammfunktionen von f , so ist $F_1 - F_2$ eine konstante Funktion.

Beweis. Aussage (a) ist klar, da F und $F + c$ den gleichen Gradienten besitzen. Zu (b): Da F_1 und F_2 Stammfunktionen von f sind, gilt $\text{grad}(F_1 - F_2) = 0$. Wir zeigen: Ist $F : U \rightarrow \mathbb{R}^n$ stetig differenzierbar auf dem Gebiet U und ist $\text{grad} F = 0$, so ist F eine Konstante. Seien $x, y \in U$. Nach Satz 11.13 gibt es einen stückweise stetig differenzierbaren Weg $\gamma : [0, 1] \rightarrow U$ mit $\gamma(0) = x$ und $\gamma(1) = y$. Aus Satz 11.10 wissen wir, dass

$$F(y) - F(x) = \int_{\gamma} \text{grad} F \cdot dx = \int_{\gamma} 0 \cdot dx = 0.$$

Also ist F konstant. ■

Am Ende von Abschnitt 11.3 haben wir gesehen, dass nicht jedes stetige Vektorfeld eine Stammfunktion besitzt. Der folgende Satz stellt einen Zusammenhang her zwischen der Existenz einer Stammfunktion und der Wegunabhängigkeit von Kurvenintegralen.

Satz 11.15 Sei $U \subseteq \mathbb{R}^n$ ein Gebiet und f ein stetiges Vektorfeld auf U . Dann besitzt f genau dann eine Stammfunktion auf U , wenn für jeden geschlossenen stückweise stetig differenzierbaren Weg γ in U das Integral $\int_{\gamma} f \cdot dx$ verschwindet.

Beweis. Die Implikation \Rightarrow haben wir uns bereits im Anschluss an Satz 11.10 überlegt. Nehmen wir nun also an, dass $\int_{\gamma} f \cdot dx = 0$ für jeden geschlossenen Weg γ in U .

Wir fixieren einen Punkt $x_0 \in U$. Für jeden Punkt $y \in U$ gibt es nach Satz 11.13 einen stückweise stetig differenzierbaren Weg $\gamma : [0, 1] \rightarrow U$ mit $\gamma(0) = x_0$

und $\gamma(1) = y$. Wir möchten definieren

$$F(y) := \int_{\gamma} f \cdot dx. \quad (11.9)$$

Um auf diese Weise eine Funktion F festlegen zu können, müssen wir uns vergewissern, dass $\int_{\gamma} f \cdot dx$ nicht von der Wahl des Weges von x_0 nach y abhängt. Sei also $\eta : [0, 1] \rightarrow U$ ein weiterer stückweise stetig differenzierbarer Weg mit $\eta(0) = x_0$ und $\eta(1) = y$. Wir betrachten den Weg

$$\alpha : [0, 2] \rightarrow U, \quad t \mapsto \begin{cases} \gamma(t) & \text{für } t \in [0, 1], \\ \eta(2-t) & \text{für } t \in [1, 2]. \end{cases}$$

Offenbar ist α wieder stückweise stetig differenzierbar, und es gilt $\alpha(0) = \alpha(2) = x_0$, d.h. der Weg α ist geschlossen. Nach Voraussetzung ist daher

$$0 = \int_{\alpha} f \cdot dx = \int_{\alpha|_{[0,1]}} f \cdot dx + \int_{\alpha|_{[1,2]}} f \cdot dx = \int_{\gamma} f \cdot dx - \int_{\eta} f \cdot dx,$$

da im zweiten Teilstück von α der Weg η rückwärts durchlaufen wird. Somit ist tatsächlich $\int_{\gamma} f \cdot dx$ nur vom Anfangs- und Endpunkt von γ abhängig. Wir schreiben daher auch $\int_{x_0}^y f \cdot dx$ statt $\int_{\gamma} f \cdot dx$.

Wir zeigen nun, dass F eine Stammfunktion von f ist. Ist $f = (f_1, \dots, f_n)$, so haben wir zu zeigen, dass F stetig differenzierbar in jedem Punkt $y \in U$ ist und dass $\frac{\partial F}{\partial x_i}(y) = f_i(y)$ für alle i ist.

Sei $y \in U$. Ist $\varepsilon > 0$ hinreichend klein, so liegt mit y auch die komplette Umgebung $U_{\varepsilon}(y)$ in U . Für alle $z \in U_{\varepsilon}(y)$ ist dann

$$F(z) = \int_{x_0}^z f \cdot dx = \int_{x_0}^y f \cdot dx + \int_y^z f \cdot dx = F(y) + \int_y^z f \cdot dx.$$

Wir betrachten den Weg

$$\gamma_z : [0, 1] \rightarrow U_{\varepsilon}(y), \quad t \mapsto y + t(z - y),$$

der y mit z verbindet. Mit $z - y := h$ erhalten wir

$$\begin{aligned} F(y+h) - F(y) &= \int_y^{y+h} f \cdot dx = \int_{\gamma_z} f \cdot dx \\ &= \int_0^1 \sum_{i=1}^n f_i(y+th) h_i dt = \sum_{i=1}^n \int_0^1 f_i(y+th) dt \cdot h_i. \end{aligned}$$

Wählen wir speziell h so, dass alle Komponenten bis auf die i -te verschwinden, so folgt

$$F(y + h_i e_i) - F(y) = \int_0^1 f_i(y + th_i e_i) dt \cdot h_i,$$

woraus wir mit dem ersten Teil von Satz 10.28 erhalten

$$\begin{aligned}\frac{\partial F}{\partial x_i}(y) &= \lim_{h_i \rightarrow 0} \frac{F(y + h_i e_i) - F(y)}{h_i} = \lim_{h_i \rightarrow 0} \int_0^1 f_i(y + th_i e_i) dt \\ &= \int_0^1 \lim_{h_i \rightarrow 0} f_i(y + th_i e_i) dt = \int_0^1 f_i(y) dt = f_i(y).\end{aligned}\quad \blacksquare$$

Die Bedingung, dass die Integrale über alle geschlossenen Wege verschwinden, ist oft schwierig zu überprüfen. Unser nächstes Ziel ist ein Kriterium, bei dem nicht Integrations-, sondern Differentiationseigenschaften eine Rolle spielen und welches meist leichter zu überprüfen ist.

Definition 11.16 Sei $U \subseteq \mathbb{R}^n$ offen und $f = (f_1, \dots, f_n)$ ein stetig differenzierbares Vektorfeld. Dann heißt f geschlossen, wenn

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} \quad \text{für alle } i, j \in \{1, \dots, n\}.$$

Besitzt das stetig differenzierbare Vektorfeld f eine Stammfunktion F , so ist f geschlossen. Aus $f_i = \frac{\partial F}{\partial x_i}$ folgt nämlich mit dem Satz von Schwarz (Satz 10.7)

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i} = \frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial f_j}{\partial x_i}.$$

Die Geschlossenheit von f ist also *notwendig* für die Existenz einer Stammfunktion; sie ist im Allgemeinen jedoch nicht *hinreichend*, wie das Beispiel

$$f(x, y) = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right) \quad \text{auf } U = \mathbb{R}^2 \setminus \{0\} \quad (11.10)$$

zeigt. Für $f_1(x, y) := -\frac{y}{x^2 + y^2}$ und $f_2(x, y) := \frac{x}{x^2 + y^2}$ ist nämlich

$$\frac{\partial f_1}{\partial y} = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial f_2}{\partial x}.$$

Das Vektorfeld f ist also geschlossen, besitzt aber – wie wir bereits wissen – keine Stammfunktion. ■

Ob die Geschlossenheit eines Vektorfeldes hinreichend für die Existenz einer Stammfunktion ist, hängt von den Eigenschaften des Gebietes U ab. Wir sehen uns eine einfache Version eines solchen Resultates an.

Satz 11.17 Ist das Gebiet $U \subseteq \mathbb{R}^n$ sternförmig, so besitzt jedes geschlossene Vektorfeld f auf U eine Stammfunktion.

Beweis. Wir können o.E.d.A. annehmen, dass 0 zu U gehört und ein Zentrum von U ist (andernfalls verschieben wir U geeignet). Für $x \in U$ sei

$$\gamma_{\alpha, x} : [0, 1] \rightarrow U, \quad t \mapsto tx$$

und

$$F(x) := \int_{\gamma_x} f \cdot dx = \int_0^1 \sum_{i=1}^n f_i(tx) x_i dt.$$

(Man beachte, dass wegen der Sternförmigkeit von U der Weg γ_x komplett in U verläuft). Mit dem zweiten Teil von Satz 10.28 erhalten wir, dass F differenzierbar ist und dass

$$\begin{aligned} \frac{\partial F}{\partial x_j}(x) &= \sum_{i=1}^n \int_0^1 \frac{\partial}{\partial x_j} (f_i(tx) x_i) dt \\ &= \sum_{i=1}^n \int_0^1 \frac{\partial f_i(tx)}{\partial x_j} t x_i dt + \sum_{i=1}^n \int_0^1 f_i(tx) \frac{\partial x_i}{\partial x_j} dt \\ &= \int_0^1 \sum_{i=1}^n \frac{\partial f_j(tx)}{\partial x_i} t x_i dt + \int_0^1 f_j(tx) dt. \end{aligned}$$

Für $L(t) := f_j(tx)$ finden wir mit der Kettenregel

$$L'(t) = \sum_{i=1}^n \frac{\partial f_j(tx)}{\partial x_i} x_i,$$

und wir erhalten

$$\begin{aligned} \frac{\partial F}{\partial x_j}(x) &= \int_0^1 L'(t) t dt + \int_0^1 L(t) dt = \int_0^1 (tL(t))' dt \\ &= tL(t)|_0^1 = L(1) = f_j(x). \end{aligned} \quad \blacksquare$$

Wir kommen noch einmal auf Beispiel (11.10) zurück. Natürlich ist $\mathbb{R}^2 \setminus \{0\}$ nicht sternförmig. Nehmen wir allerdings aus \mathbb{R}^2 die negative Halbachse $(-\infty, 0]$ heraus, so ist $\mathbb{R} \setminus (-\infty, 0]$ sternförmig, und f besitzt auf diesem kleineren Gebiet nach Satz 11.17 eine Stammfunktion.

Das folgende Vorgehen zum Auffinden einer Stammfunktion eines geschlossenen Vektorfeldes $f = (f_1, f_2)$ auf einem Gebiet $U \subseteq \mathbb{R}^2$ ist praktikabel. Wir machen den Ansatz

$$\frac{\partial F}{\partial x} = f_1, \quad \frac{\partial F}{\partial y} = f_2$$

mit einer Funktion $F : U \rightarrow \mathbb{R}$ (von der wir ohne weitere Voraussetzungen nicht wissen, ob sie existiert). Aus $\frac{\partial F}{\partial x} = f_1$ folgt

$$F(x, y) = \int f_1(x, y) dx + g(y)$$

mit einer von y abhängenden Integrationskonstanten g . Wir leiten dies formal nach y ab und erhalten mit $\frac{\partial F}{\partial y} = f_2$:

$$\frac{\partial F}{\partial y} = \frac{\partial}{\partial y} \int f_1(x, y) dx + g'(y) = f_2(x, y),$$

also

$$g'(y) = f_2(x, y) - \frac{\partial}{\partial y} \int f_1(x, y) dx.$$

Durch Integration bezüglich y gewinnt man g und damit F .

Beispiel. Für $x \neq 0$, $y \neq \frac{\pi}{2} + k\pi$ und $k \in \mathbb{Z}$ sei

$$f_1(x, y) = -\frac{\tan y}{x^2} + 2xy + x^2, \quad f_2(x, y) = \frac{1}{x \cos^2 y} + x^2 + y^2.$$

Dann ist $f = (f_1, f_2)$ geschlossen, und der Ansatz $\frac{\partial F}{\partial x} = f_1$, $\frac{\partial F}{\partial y} = f_2$ liefert

$$F(x, y) = \int \left(-\frac{\tan y}{x^2} + 2xy + x^2 \right) dx = \frac{\tan y}{x} + x^2 y + \frac{x^3}{3} + g(y).$$

Wir differenzieren nach y und setzen die Ableitung gleich f_2 :

$$\frac{1}{x \cos^2 y} + x^2 + g'(y) \stackrel{!}{=} f_2(x, y) = \frac{1}{x \cos^2 y} + x^2 + y^2.$$

Dann folgt $g(y) = \frac{y^3}{3} + C$ und $F(x, y) = \frac{\tan y}{x} + x^2 y + \frac{x^3}{3} + \frac{y^3}{3} + C$. ■

12 Gleichungen und Mannigfaltigkeiten

Wir wenden uns nun einer der typischen Aufgaben der Mathematik zu, dem Lösen von Gleichungen. Sind X, Y geeignete Mengen und ist $F : X \rightarrow Y$ eine entsprechende Abbildung, so sind in Verbindung mit der Gleichung $F(x) = y$ folgende Fragen von Interesse:

- Für welche $y \in Y$ hat die Gleichung $F(x) = y$ eine Lösung $x \in X$?
- Wenn die Gleichung $F(x) = y$ für ein $y \in Y$ lösbar ist, wieviele Lösungen hat sie dann?
- Falls die Gleichung $F(x) = y$ für alle $y \in Y$ eindeutig lösbar ist, wie hängen dann die Lösungen x von den rechten Seiten y ab? Ist diese Abhängigkeit stetig oder gar differenzierbar?
- Falls die Gleichung $F(x) = y$ mehrere Lösungen besitzt, wie kann man dann die Menge aller Lösungen geeignet darstellen?
- Wie findet man Lösungen?

Die erste Frage ist eng mit der Surjektivität von F verknüpft und die zweite mit der Injektivität. Ist F surjektiv und injektiv (also bijektiv), so besitzt F eine Umkehrabbildung F^{-1} , und der dritte Punkt fragt nach der Stetigkeit bzw. Differenzierbarkeit von F^{-1} . Schließlich wird uns die vierte Frage auf den Begriff einer Mannigfaltigkeit führen.

Wir wollen insbesondere sehen, inwieweit die uns zugänglichen Mittel der Analysis bei der Beantwortung dieser Fragen helfen. Wir betrachten deshalb meist differenzierbare Funktionen $F : U \rightarrow V$, wobei $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offene Mengen sind. Beginnen werden wir aber mit einem Resultat, das nicht auf diesen Rahmen beschränkt ist.

12.1 Der Banachsche Fixpunktsatz

Sei (X, d) ein metrischer Raum. Eine Abbildung $f : X \rightarrow X$ heißt *Kontraktion*, wenn es eine Zahl $L < 1$ (die *Kontraktionskonstante*) so gibt, dass

$$d(f(x), f(y)) \leq Ld(x, y) \quad \text{für alle } x, y \in X.$$

Kontraktionen sind Lipschitzstetig und insbesondere stetig.

Ein Punkt $x \in X$ heißt *Fixpunkt* von $f : X \rightarrow X$, wenn $f(x) = x$. Für die identische Abbildung von X ist jeder Punkt ein Fixpunkt.

Satz 12.1 (Banachscher Fixpunktsatz) *Sei (X, d) ein nichtleerer vollständiger metrischer Raum und $f : X \rightarrow X$ eine Kontraktion mit einer Kontraktionskonstanten L . Dann gilt*

- (a) die Abbildung f besitzt genau einen Fixpunkt x^* .
 (b) für jeden Startvektor $x_0 \in X$ konvergiert die für $n \geq 1$ durch $x_n := f(x_{n-1})$ definierte Folge (x_n) gegen x^* .
 (c) $d(x_n, x^*) \leq \frac{1}{1-L} d(x_n, x_{n+1}) \leq \frac{L^n}{1-L} d(x_0, x_1)$.

Beweis. Sei $x_0 \in X$ beliebig gewählt und $x_n := f(x_{n-1})$ für $n \geq 1$. Wir zeigen, dass (x_n) eine Cauchyfolge ist. Für $m \geq 1$ haben wir

$$\begin{aligned} d(x_n, x_{n+m}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+m-1}, x_{n+m}) \\ &\leq (1 + L + \dots + L^{m-1}) d(x_n, x_{n+1}) \\ &= \frac{1 - L^m}{1 - L} d(x_n, x_{n+1}) \leq \frac{1}{1 - L} d(x_n, x_{n+1}) \\ &\leq \frac{L^n}{1 - L} d(x_0, x_1). \end{aligned}$$

Wegen $0 \leq L < 1$ wird die rechte Seite kleiner als jedes vorgegebene $\varepsilon > 0$, wenn nur n hinreichend groß ist. Also ist (x_n) eine Cauchyfolge.

Da X vollständig ist, konvergiert (x_n) gegen ein $x^* \in X$. Aus der Stetigkeit von f folgt schließlich

$$f(x^*) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x^*;$$

x^* ist also Fixpunkt von f . Die Abbildung f kann keine weiteren Fixpunkte besitzen. Aus $x^* = f(x^*)$ und $y^* = f(y^*)$ folgt nämlich

$$d(x^*, y^*) = d(f(x^*), f(y^*)) \leq L d(x^*, y^*),$$

also $d(x^*, y^*) = 0$. Die in (c) angegebenen Abschätzungen folgen sofort aus

$$d(x_n, x_{n+m}) \leq \frac{1}{1-L} d(x_n, x_{n+1}) \leq \frac{L^n}{1-L} d(x_0, x_1),$$

wenn man $m \rightarrow \infty$ streben lässt. ■

Der Banachsche Fixpunktsatz ist ein wichtiges Werkzeug der Analysis, das Ihnen auch in anderen Situationen wiederbegegnet wird (z.B. in der Vorlesung über Differentialgleichungen im 3. Semester). In der numerischen Mathematik ist der Banachsche Fixpunktsatz ein Instrument, um die Konvergenz von Näherungsverfahren zu beweisen und Lösungen von Fixpunktgleichungen näherungsweise zu berechnen.

Beispiel 1. Sei X das Stadtgebiet von Darmstadt (das wir als abgeschlossene Teilmenge des \mathbb{R}^2 auffassen und das deshalb vollständig ist). Irgendwo in Darmstadt breiten wir einen Stadtplan von Darmstadt aus und erklären eine Abbildung $f : X \rightarrow X$ wie folgt. Jedem Punkt $x \in X$ wird derjenige Punkt $f(x) \in X$ zugeordnet, über dem das Bild von x auf dem Stadtplan liegt. Ist $1 : n$ der Maßstab des

Stadtplans, so ist f eine Kontraktion mit der Kontraktionskonstanten $L = 1/n$. Nach dem Banachschen Fixpunktsatz gibt es genau einen Punkt x in Darmstadt mit $f(x) = x$, d.h. x liegt genau unter demjenigen Punkt des Stadtplanes, der x abbildet.

Beispiel 2. Sei $A \in L(\mathbb{R}^n)$ eine lineare Abbildung mit $\|A\| < 1$. Wir wollen die lineare Gleichung

$$(I - A)x = y, \quad y \in \mathbb{R}^n, \quad (12.1)$$

lösen. Dazu betrachten wir die Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $z \mapsto Az + y$, mit der wir (12.1) als Fixpunktgleichung $f(x) = x$ schreiben können. Wegen

$$\|f(z_1) - f(z_2)\| = \|(Az_1 + y) - (Az_2 + y)\| = \|A(z_1 - z_2)\| \leq \|A\| \|z_1 - z_2\|$$

ist f eine Kontraktion mit der Kontraktionskonstanten $\|A\| < 1$. Nach Satz 12.1 hat f genau einen Fixpunkt, d.h. die Gleichung (12.1) hat genau eine Lösung in \mathbb{R}^n . Diese können wir näherungsweise berechnen. Wir wählen z.B. $x_0 = y$ als Startvektor. Dann ist $x_1 = f(x_0) = Ay + y$,

$$x_2 = f(x_1) = Ax_1 + y = A(Ay + y) + y = A^2y + Ay + y$$

und allgemein, mit $A^0 := I$,

$$x_n = \sum_{k=0}^n A^k y \quad \text{für } n \geq 1.$$

Folglich ist

$$(I - A)^{-1}y = x = \lim_{n \rightarrow \infty} x_n = \sum_{k=0}^{\infty} A^k y.$$

Man kann auch leicht direkt beweisen, dass die Reihe $\sum_{k=0}^{\infty} A^k$ in $L(\mathbb{R}^n)$ konvergiert (die geometrische Reihe $\sum_{k=0}^{\infty} \|A\|^k$ ist eine konvergente Majorante) und dass

$$\sum_{k=0}^{\infty} A^k = (I - A)^{-1} \quad \text{für } \|A\| < 1.$$

Die Reihe $\sum_{k=0}^{\infty} A^k$ heißt die *Neumann-Reihe* von A . ■

12.2 Der Satz über die Umkehrfunktion

Wir sehen uns nun an, wie man den Satz über die Umkehrfunktion auf Funktionen mehrerer Veränderlicher verallgemeinert.

Definition 12.2 Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen. Eine Bijektion $f : U \rightarrow V$ heißt Diffeomorphismus, wenn sowohl f als auch die Umkehrabbildung $f^{-1} : V \rightarrow U$ stetig differenzierbar sind.

Lemma 12.3 Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen und $f : U \rightarrow V$ eine stetig differenzierbare Funktion mit differenzierbarer Umkehrfunktion. Dann gilt:

(a) für jedes $x \in U$ ist $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^m)$ invertierbar, und

$$(f'(x))^{-1} = (f^{-1})'(f(x)), \quad (12.2)$$

(b) $m = n$,

(c) $f : U \rightarrow V$ ist ein Diffeomorphismus.

Beweis. Sei $x \in U$ und $y = f(x)$. Mit der Kettenregel folgt aus $f^{-1} \circ f = id_U$ und $f \circ f^{-1} = id_V$, dass

$$\begin{aligned} id_{\mathbb{R}^n}(x) &= (f^{-1} \circ f)'(x) = (f^{-1})'(f(x)) \cdot f'(x), \\ id_{\mathbb{R}^m}(y) &= (f \circ f^{-1})'(y) = f'(f^{-1}(y)) \cdot (f^{-1})'(y) = f'(x) \cdot (f^{-1})'(f(x)). \end{aligned}$$

Also ist $f'(x)$ invertierbar, und es gilt (12.2). Aus der Invertierbarkeit von $f'(x) \in L(\mathbb{R}^n, \mathbb{R}^m)$ folgt weiter $m = n$ (lineare Algebra). Schließlich ist

$$(f^{-1})' : V \rightarrow L(\mathbb{R}^n), \quad y \mapsto (f'(f^{-1}(y)))^{-1}. \quad (12.3)$$

Für (c) müssen wir noch die Stetigkeit dieser Funktion zeigen. Nach Voraussetzung sind f^{-1} und f' stetig. Damit ist auch ihre Verkettung $f' \circ f^{-1}$ stetig, und wir müssen noch die Stetigkeit der Abbildung

$$GL(\mathbb{R}^n) \rightarrow GL(\mathbb{R}^n), \quad A \mapsto A^{-1} \quad (12.4)$$

zeigen, wobei $GL(\mathbb{R}^n)$ für die Gruppe der invertierbaren Abbildungen aus $L(\mathbb{R}^n)$ steht. Dies folgt leicht aus der expliziten Formel zur Berechnung der inversen Matrix von A , wonach $A^{-1} = \frac{1}{\det A} \tilde{A}$ mit

$$\tilde{A} = (\tilde{a}_{ij})_{i,j=1}^n \quad \text{und} \quad \tilde{a}_{ij} = (-1)^{i+j} \det A_{ji}.$$

Hier ist A_{ji} diejenige Untermatrix von A , die durch Streichen der j -ten Zeile und i -ten Spalte von A entsteht. Es ist klar, dass jeder Eintrag von \tilde{A} (und damit \tilde{A} selbst) sowie $\det A$ stetig von (den Einträgen von) A abhängen. Also ist die Abbildung (12.4) stetig. Damit ist die Stetigkeit von $(f^{-1})'$ in (12.3) gezeigt. Die Funktion f^{-1} ist also *stetig* differenzierbar, d.h. f ist ein Diffeomorphismus. ■

Anmerkung. Sei $C^k(U, \mathbb{R}^m)$ die Menge der k -mal stetig differenzierbaren Funktionen $f : U \rightarrow \mathbb{R}^m$. Verlangt man in Lemma 12.3 zusätzlich $f \in C^k(U, \mathbb{R}^m)$, so kann man $f^{-1} \in C^k(V, \mathbb{R}^n)$ zeigen. Dies geschieht wieder mit Hilfe der Darstellung (12.3). Man kann sich überlegen, dass die Abbildung (12.4) sogar beliebig oft differenzierbar ist. ■

Beispiel. Dieses Beispiel soll noch einmal auf Unterschiede zwischen Funktionen

einer bzw. mehrerer Veränderlicher aufmerksam machen. Sei $f : U \rightarrow V$ eine stetig differenzierbare und surjektive Funktion, deren Ableitung $f'(x)$ in jedem Punkt $x \in U$ invertierbar ist. Sind U, V offene Intervalle in \mathbb{R} , so folgt hieraus, dass f bijektiv ist. Die Invertierbarkeit von $f'(x)$ bedeutet im Eindimensionalen nämlich gerade, dass $f'(x) \neq 0$ ist. Ist dies für alle $x \in U$ der Fall, so ist f streng monoton, also injektiv. Im Mehrdimensionalen trifft dies jedoch nicht mehr zu, wie folgendes Beispiel zeigt. Sei

$$f : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}, \quad (r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi).$$

Die Jacobimatrix

$$f'(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

hat die Determinante $r \neq 0$ und ist folglich immer invertierbar. Die Funktion f ist offenbar auch surjektiv, wegen $f(r, \varphi) = f(r, \varphi + 2\pi)$ jedoch nicht injektiv. ■

Die Funktion f aus diesem Beispiel besitzt jedoch eine „lokale Umkehrfunktion“. Ist etwa $U := (0, \infty) \times (-\pi, \pi)$, so ist $f|_U : U \rightarrow \mathbb{R}^2$ injektiv, und $V := f(U) = \mathbb{R}^2 \setminus (-\infty, 0]$ ist offen in \mathbb{R}^2 . Man kann also eine Umkehrfunktion $f^{-1} : V \rightarrow U$ definieren. Wir wollen uns nun die Differenzierbarkeitseigenschaften solcher „lokaler Umkehrfunktionen“ ansehen.

Definition 12.4 Sei $U \subseteq \mathbb{R}^n$ offen. Eine stetig differenzierbare Abbildung $f : U \rightarrow \mathbb{R}^n$ heißt lokal um $x \in U$ invertierbar, wenn es offene Umgebungen $U_1 \subseteq U$ von x und V_1 von $f(x)$ so gibt, dass $f|_{U_1} : U_1 \rightarrow V_1$ ein Diffeomorphismus ist. Die Abbildung $(f|_{U_1})^{-1} : V_1 \rightarrow U_1$ heißt dann eine lokale Umkehrfunktion von f . Schließlich heißt f ein lokaler Diffeomorphismus, wenn f um jeden Punkt $x \in U$ lokal invertierbar ist.

Satz 12.5 (Satz über die Umkehrfunktion) Sei $U \subseteq \mathbb{R}^n$ offen, $x_0 \in U$, und $f : U \rightarrow \mathbb{R}^n$ stetig differenzierbar. Die Funktion f ist genau dann um x_0 lokal invertierbar, wenn $f'(x_0)$ invertierbar ist.

Verlangt man $f \in C^k(U, \mathbb{R}^n)$, so folgt aus der Anmerkung nach Lemma 12.3, dass die lokale Umkehrfunktion ebenfalls zu C^k gehört.

Beweis. Ist f um x_0 lokal invertierbar, so folgt die Invertierbarkeit von $f'(x_0)$ aus Lemma 12.3. Sei umgekehrt $f'(x_0)$ invertierbar. Wir zeigen, dass dann f um x_0 lokal invertierbar ist. Ohne Einschränkung der Allgemeinheit können wir dabei annehmen, dass

$$x_0 = 0, \quad f(0) = 0 \quad \text{und} \quad f'(0) = I \tag{12.5}$$

ist. Andernfalls ersetzen wir U durch $\tilde{U} = U - x_0$ und f durch

$$\tilde{f}(x) := (f'(x_0))^{-1} \cdot (f(x_0 + x) - f(x_0)).$$

Hat man dann eine lokale Umkehrfunktion von \tilde{f} um $0 \in \tilde{U}$ gefunden, so ist

$$f^{-1}(y) = \tilde{f}^{-1}(f'(x_0)^{-1}(y - f(x_0))) + x_0 \quad (12.6)$$

die lokale Umkehrfunktion von f um x_0 . Aus der Definition von \tilde{f} folgt nämlich

$$x + x_0 = \tilde{f}^{-1}((f'(x_0))^{-1}(f(x_0 + x) - f(x_0))) + x_0,$$

und die Substitution $f(x_0 + x) = y$ liefert (12.6).

Wir zeigen, dass man unter den Voraussetzungen (12.5) die Gleichung $f(x) = y$ für kleine y nach x auflösen kann. Wie möchten dabei den Banachschen Fixpunktsatz benutzen und schreiben $f(x) = y$ als Fixpunktgleichung. Dazu definieren wir

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

Offenbar ist $f(x) = y$ genau dann, wenn $g_y(x) = x$.

Wir suchen einen geeigneten vollständigen metrischen Raum X , auf dem g_y kontraktiv wirkt. Sei zunächst $y = 0$ und $g := g_0$. Mit f ist auch g stetig differenzierbar. Da $g'(0) = I - I = 0$ ist, gibt es ein $r > 0$ so, dass

$$U_{2r}(0) \subseteq U \quad \text{und} \quad \|g'(x)\| \leq 1/2 \quad \text{für alle } x \text{ mit } \|x\| \leq r.$$

Sei $X := \{x \in \mathbb{R}^n : \|x\| \leq r\}$. Da X im vollständigen metrischen Raum \mathbb{R}^n abgeschlossen ist, ist X selbst vollständig. Nach Satz 10.19 gilt weiter

$$\|g(x) - g(x')\| \leq \frac{1}{2} \|x - x'\| \quad \text{für alle } x, x' \in X. \quad (12.7)$$

Mit $x' = 0$ folgt hieraus insbesondere $\|g(x)\| = \frac{1}{2} \|x\| \leq r/2$ für $x \in X$.

Sei nun $\|y\| \leq r/2$. Dann ist g_y wegen $\|g_y(x)\| \leq \|g(x)\| + \|y\| \leq r$ für alle $x \in X$ eine Abbildung von X in X , die wegen $g_y(x) - g_y(x') = g(x) - g(x')$ und (12.7) eine Kontraktion mit der Kontraktionskonstanten $1/2$ ist. Nach dem Banachschen Fixpunktsatz gibt es also für jedes y mit $\|y\| \leq r/2$ genau ein x mit $\|x\| \leq r$ so, dass $g_y(x) = x$ bzw. $f(x) = y$.

Sei $U_1 := \{x \in \mathbb{R}^n : \|x\| < r, \|f(x)\| < r/2\}$ und $V_1 := f(U_1)$. Wie wir soeben gesehen haben, ist $f|_{U_1} : U_1 \rightarrow V_1$ bijektiv. Es existiert also die Umkehrabbildung $\varphi := (f|_{U_1})^{-1} : V_1 \rightarrow U_1$. Wir zeigen, dass U_1 und V_1 offen sind und φ stetig ist. Für U_1 ist dies klar (Urbilder offener Mengen bzgl. stetiger Abbildungen sind offen). Für V_1 zeigen wir $V_1 = \{y \in \mathbb{R}^n : \|y\| < r/2\}$, woraus die Offenheit von V_1 folgt. Die Inklusion \subseteq ist klar. Für die umgekehrte Inklusion sei $\|y\| < r/2$. Dann existiert genau ein x mit $\|x\| \leq r$ und $f(x) = y$. Zu zeigen ist, dass sogar $\|x\| < r$. Für beliebige Punkte $x_1, x_2 \in \overline{U_r(0)}$ ist

$$\begin{aligned} \|x_1 - x_2\| &= \|g(x_1) + f(x_1) - g(x_2) - f(x_2)\| \\ &\leq \|g(x_1) - g(x_2)\| + \|f(x_1) - f(x_2)\| \\ &\leq \frac{1}{2} \|x_1 - x_2\| + \|f(x_1) - f(x_2)\|, \end{aligned}$$

also

$$\|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\|. \quad (12.8)$$

Wir setzen hierin $x_2 = 0$ und $x_1 = x$ mit $f(x) = y$ und $\|y\| < r/2$. Dann folgt $\|x\| \leq 2\|y\| < r$, also $x \in U_1$ und $y \in V_1$. Schreiben wir noch in (12.8) $x_1 = \varphi(y_1)$ und $x_2 = \varphi(y_2)$, so erhalten wir

$$\|\varphi(y_1) - \varphi(y_2)\| \leq 2\|y_1 - y_2\| \quad \text{für alle } y_1, y_2 \in V_1, \quad (12.9)$$

also die Stetigkeit von φ .

Weiter ist $f'(x)$ für alle $x \in U_1$ invertierbar. Für $x \in U_1$ ist nämlich

$$\|f'(x) - I\| = \|g'(x)\| \leq 1/2,$$

und die Invertierbarkeit von $f'(x)$ folgt wie in Abschnitt 12.1 (Neumann-Reihe).

Es verbleibt zu zeigen, dass φ stetig differenzierbar ist. Wir zeigen zunächst die Differenzierbarkeit von φ . Sei $y \in V_1$ und ℓ so, dass $y + \ell \in V_1$. Wir setzen $x = \varphi(y)$ und $x + h = \varphi(y + \ell)$, wobei $h = \varphi(y + \ell) - \varphi(y)$ von ℓ abhängt. Da f in x differenzierbar ist, haben wir

$$f(x + h) - f(x) = f'(x)h + r(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0.$$

Hieraus folgt

$$\begin{aligned} \varphi(y + \ell) - \varphi(y) &= (x + h) - x = h \\ &= (f'(x))^{-1}(f(x + h) - f(x) - r(h)) \\ &= (f'(\varphi(y)))^{-1}\ell - (f'(x))^{-1}r(h), \end{aligned}$$

und die Differenzierbarkeit von φ in y folgt, wenn wir gezeigt haben, dass $\lim_{\ell \rightarrow 0} r(h(\ell))/\|\ell\| = 0$. Nun ist

$$\frac{r(h(\ell))}{\|\ell\|} = \frac{r(h(\ell))}{\|h(\ell)\|} \frac{\|h(\ell)\|}{\|\ell\|} \quad (12.10)$$

(wegen der Bijektivität von φ ist $h \neq 0$ für $\ell \neq 0$). Wegen $h = \varphi(y + \ell) - \varphi(y)$ und der Stetigkeit von φ folgt aus $\ell \rightarrow 0$ auch $h \rightarrow 0$. Also ist

$$\lim_{\ell \rightarrow 0} \frac{r(h(\ell))}{\|h(\ell)\|} = \lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0,$$

und der zweite Faktor in (12.10) bleibt wegen (12.9) beschränkt:

$$\|h(\ell)\| = \|\varphi(y + \ell) - \varphi(y)\| \leq 2\|(y + \ell) - y\| = 2\|\ell\|.$$

Also ist φ in $y \in V_1$ differenzierbar. Die Stetigkeit von φ' folgt aus Lemma 12.3 (c). ■

Der Satz über die Umkehrfunktion ist ein wichtiges Werkzeug der Analysis. In vielen Situationen dient er beispielsweise dazu, geeignete „Koordinaten“ einzuführen, wobei wir uns die lokal invertierbaren Funktionen als „Koordinatenwechsel“ denken (in Analogie zu den Basistransformationen der linearen Algebra).

Der Satz über die Umkehrfunktion ist eine lokale Aussage. Eine globale Variante dieses Satzes erhält man, wenn man von vornherein die Existenz einer (globalen) Umkehrfunktion fordert.

Folgerung 12.6 *Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^n$ sei injektiv und stetig differenzierbar, und $f'(x)$ sei invertierbar für alle $x \in U$. Dann ist $f(U) \subseteq \mathbb{R}^n$ offen, und $f^{-1} : f(U) \rightarrow U$ ist stetig differenzierbar.*

Ist sogar $f \in C^k(U, \mathbb{R}^n)$, so folgt auch $f^{-1} \in C^k(f(U), \mathbb{R}^n)$.

Beweis. Sei $y \in f(U)$ und $f(x) = y$. Nach Satz 12.5 gibt es offene Umgebungen $U_1 \subseteq U$ und $V_1 \subseteq f(U)$ von x bzw. y so, dass $f|_{U_1} : U_1 \rightarrow V_1$ ein Diffeomorphismus ist. Also ist $f(U)$ offen und $f^{-1}|_{V_1} = (f|_{U_1})^{-1}$ ist stetig differenzierbar. ■

12.3 Der Satz über implizite Funktionen

Dieser Satz ist eine wichtige Folgerung aus dem Satz über die Umkehrfunktion. Er eröffnet die Möglichkeit, die Lösungsmenge einer Gleichung wie $f(x, y) = 0$ geeignet zu parametrisieren. Um den Satz über implizite Funktionen besser zu verstehen, betrachten wir zunächst das folgende lineare Problem.

Sei $A : \mathbb{R}^{n+k} \rightarrow \mathbb{R}^k$ eine lineare Abbildung. Wir schreiben die Elemente von \mathbb{R}^{n+k} als Paare $(x, y) \in \mathbb{R}^n \times \mathbb{R}^k$. Da A linear ist, gibt es lineare Abbildungen $A_1 : \mathbb{R}^n \rightarrow \mathbb{R}^k$ und $A_2 : \mathbb{R}^k \rightarrow \mathbb{R}^k$ so, dass

$$A(x, y) = A_1x + A_2y \quad \text{für alle } (x, y) \in \mathbb{R}^n \times \mathbb{R}^k.$$

In dieser Darstellung ist klar, dass sich die Gleichung $A(x, y) = 0$ genau dann nach y auflösen lässt, wenn die lineare Abbildung A_2 invertierbar ist, und dass in diesem Fall

$$A(x, y) = 0 \quad \iff \quad y = -A_2^{-1}A_1x.$$

Die Lösungsmenge ist also der Graph einer linearen Abbildung von \mathbb{R}^n nach \mathbb{R}^k . Der Satz über implizite Funktionen ist eine Verallgemeinerung dieser Beobachtung auf nichtlineare Abbildungen. Wegen der Nichtlinearität erhält man nur eine lokale Aussage.

Satz 12.7 (Satz über implizite Funktionen) Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^k$ offen und $f : U \times V \rightarrow \mathbb{R}^k$ stetig differenzierbar. Für $(x, y) \in U \times V$ spalten wir die Ableitung $f'(x, y)$ von f auf in

$$f'(x, y) = (d_1f(x, y), d_2f(x, y)),$$

wobei

$$\begin{aligned} d_1f(x, y) &= f'(x, y)|_{\mathbb{R}^n \times \{0\}} \in L(\mathbb{R}^n, \mathbb{R}^k), \\ d_2f(x, y) &= f'(x, y)|_{\{0\} \times \mathbb{R}^k} \in L(\mathbb{R}^k, \mathbb{R}^k). \end{aligned}$$

Ist $(x_0, y_0) \in U \times V$ ein Punkt mit $f(x_0, y_0) = 0$ und ist $d_2f(x_0, y_0)$ invertierbar, so existieren offene Umgebungen $U_1 \subseteq U$ von x_0 und $V_1 \subseteq V$ von y_0 sowie eine stetig differenzierbare Abbildung $\eta : U_1 \rightarrow V_1$ mit $\eta(x_0) = y_0$ und

$$\{(x, y) \in U_1 \times V_1 : f(x, y) = 0\} = \{(x, \eta(x)) : x \in U_1\}. \quad (12.11)$$

Insbesondere ist $f(x, \eta(x)) = 0$ für alle $x \in U_1$. Ist auch noch $d_2f(x, \eta(x))$ invertierbar, so folgt

$$\eta'(x) = -(d_2f(x, \eta(x)))^{-1} d_1f(x, \eta(x)). \quad (12.12)$$

Speziell ist für $n = k = 1$

$$\eta'(x) = -\frac{\frac{\partial f}{\partial x}(x, \eta(x))}{\frac{\partial f}{\partial y}(x, \eta(x))}.$$

Die Funktion η wird also *implizit* durch die Gleichung $f(x, \eta(x)) = 0$ definiert. Wir können diese Funktion auffassen als lokale Parametrisierung der Lösungsmenge der Gleichung $f(x, y) = 0$, da diese durch (12.11) lokal als Graph der Funktion η dargestellt wird. Ist $f \in C^m$, so kann man wieder $\eta \in C^m$ zeigen.

Beweis. Wir identifizieren lineare Abbildungen mit ihren Matrixdarstellungen bezüglich der Standardbasis von \mathbb{R}^m . In diesem Sinn hat die Abbildung

$$\varphi : U \times V \rightarrow \mathbb{R}^n \times \mathbb{R}^k, \quad (x, y) \mapsto (x, f(x, y))$$

die Ableitung (= Jacobimatrix)

$$\varphi'(x, y) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ & \ddots & & \ddots \\ 0 & 1 & 0 & 0 \\ \left(\frac{\partial f_i}{\partial x_j}(x, y)\right)_{i,j=1}^{k,n} & & \left(\frac{\partial f_i}{\partial y_\ell}(x, y)\right)_{i,\ell=1}^{k,k} & \end{pmatrix} = \begin{pmatrix} I & 0 \\ d_1f(x, y) & d_2f(x, y) \end{pmatrix}.$$

Aus der Invertierbarkeit von $d_2f(x_0, y_0)$ folgt daher die von $\varphi'(x_0, y_0)$. Da φ stetig differenzierbar ist, existiert nach dem Satz über die Umkehrfunktion eine Umgebung $W \subseteq U \times V$ von (x_0, y_0) so, dass $\varphi|_W : W \rightarrow \varphi(W) \subseteq \mathbb{R}^n \times \mathbb{R}^k$ ein

Diffeomorphismus ist. Die Umkehrfunktion $\psi := (\varphi|_W)^{-1} : \varphi(W) \rightarrow W$ hat dann die Gestalt

$$\psi(x, y) = (x, g(x, y)) \quad \text{mit } g : \varphi(W) \rightarrow \mathbb{R}^k.$$

Wir definieren

$$\eta : \{x \in \mathbb{R}^n : (x, 0) \in \varphi(W)\} \rightarrow \mathbb{R}^k, \quad x \mapsto g(x, 0).$$

Dann ist $\psi(x, 0) = (x, g(x, 0)) = (x, \eta(x))$ und daher

$$(x, 0) = \varphi(\psi(x, 0)) = \varphi(x, \eta(x)) = (x, f(x, \eta(x))),$$

also $f(x, \eta(x)) = 0$. Ist umgekehrt $f(x, y) = 0$ für $(x, y) \in W$, so ist $\varphi(x, y) = (x, 0)$ und daher $(x, y) = \psi(x, 0) = (x, g(x, 0)) = (x, \eta(x))$, also $y = \eta(x)$. Dann ist klar, dass

$$\{(x, y) \in W : f(x, y) = 0\} = \{(x, \eta(x)) \in W : (x, 0) \in \varphi(W)\}. \quad (12.13)$$

Wir wählen nun offene Umgebungen U'_1 von x_0 und V_1 von y_0 so klein, dass $U'_1 \times V_1 \subseteq W$ und dann eine Umgebung U_1 von x_0 mit $\eta(U_1) \subseteq V_1$ (diese lässt sich finden, da η stetig ist). Dann folgt (12.11) aus (12.13), und η ist stetig differenzierbar nach dem Satz über die Umkehrfunktion.

Wir zeigen noch (12.12). Aus $f(x, \eta(x)) = 0$ für alle $x \in U_1$ und

$$f'(x, y) = (d_1 f(x, y), d_2 f(x, y))$$

folgt mit der Kettenregel (wir leiten $f(x, \eta(x)) = 0$ nach x ab):

$$0 = f'(x, \eta(x)) \circ (id, \eta'(x)) = d_1 f(x, \eta(x)) + d_2 (f(x, \eta(x))) \circ \eta'(x).$$

Ist $d_2(f(x, \eta(x)))$ invertierbar, folgt hieraus (12.12). ■

Beispiel 1. Wir betrachten den Einheitskreis, d.h. die Nullstellenmenge von

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^2 + y^2 - 1.$$

Die lokale Auflösbarkeitsbedingung nach y lautet

$$0 \neq \frac{\partial f}{\partial y}(x_0, y_0) = 2y_0.$$

Für $y_0 \neq 0$ existieren also Funktionen η so, dass der Einheitskreis lokal als Graph von η darstellbar ist. Für $y_0 > 0$ ist $\eta(x) = \sqrt{1 - x^2}$, und für $y_0 < 0$ haben wir $\eta(x) = -\sqrt{1 - x^2}$. Für $y_0 = 0$ ist der Einheitskreis nicht lokal als Graph einer Funktion von x darstellbar. Dafür kann man um die Punkte $(1, 0)$ und $(-1, 0)$ den Einheitskreis lokal als Graph einer Funktion von y darstellen. Wegen

$$\frac{\partial f}{\partial x}(x_0, y_0) = 2x_0 \neq 0$$

ist in diesen Punkten nämlich die lokale Auflösbarkeitsbedingung nach x erfüllt. (Beachten Sie: es spielt keine Rolle, wie man den \mathbb{R}^{n+k} in zwei Unterräume E_1 und E_2 aufteilt. Wichtig ist nur, dass $d_2f(x_0, y_0) : E_2 \rightarrow \mathbb{R}^k$ invertierbar ist.) Entsprechend erhalten wir Funktionen $\eta(y) = \sqrt{1 - y^2}$ für $x_0 > 0$ und $\eta(y) = -\sqrt{1 - y^2}$ für $x < x_0$. ■

Beispiel 2. Die Funktion

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad (x, y, z) \mapsto \begin{pmatrix} x^2 - y^2 \\ x^2 - z^2 \end{pmatrix}$$

hat die Jacobimatrix

$$\begin{pmatrix} 2x & -2y & 0 \\ 2x & 0 & -2z \end{pmatrix}.$$

Die Auflösbarkeitsbedingung nach (y, z) ist erfüllt, wenn

$$\det \left(\frac{\partial f}{\partial (y, z)}(x_0, y_0, z_0) \right) = \det \begin{pmatrix} -2y_0 & 0 \\ 0 & -2z_0 \end{pmatrix} = 4y_0z_0 \neq 0.$$

Ist dagegen

$$\det \left(\frac{\partial f}{\partial (x, z)}(x_0, y_0, z_0) \right) = \det \begin{pmatrix} 2x_0 & 0 \\ 2x_0 & -2z_0 \end{pmatrix} = -4x_0z_0 \neq 0,$$

so erhalten wir die lokale Auflösbarkeit nach dem Variablenpaar (x, z) . Hier steht $\frac{\partial f}{\partial (y, z)}$ für die Jacobimatrix von f , betrachtet als Funktion von y und z . ■

Beispiel 3. Sei $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ die Funktion

$$f(x, t) = t^n + \sum_{k=0}^{n-1} x_k t^k.$$

Offenbar ist $f(x, t) = 0$ genau dann, wenn t eine Nullstelle des Polynoms $P_x(t) := f(x, t)$ ist. Sei t_0 eine einfache Nullstelle von P_{x_0} , d.h. sei $P_{x_0}(t_0) = 0$ und $P'_{x_0}(t_0) \neq 0$. Dann ist

$$\frac{\partial f}{\partial t}(x_0, t_0) = P'_{x_0}(t_0) \neq 0.$$

Also ist $f(x, t)$ lokal nach t auflösbar. ■

Wir haben damit die folgende bemerkenswerte Aussage bewiesen:

Die einfachen Nullstellen eines Polynoms hängen lokal beliebig oft differenzierbar von den Koeffizienten des Polynoms ab.

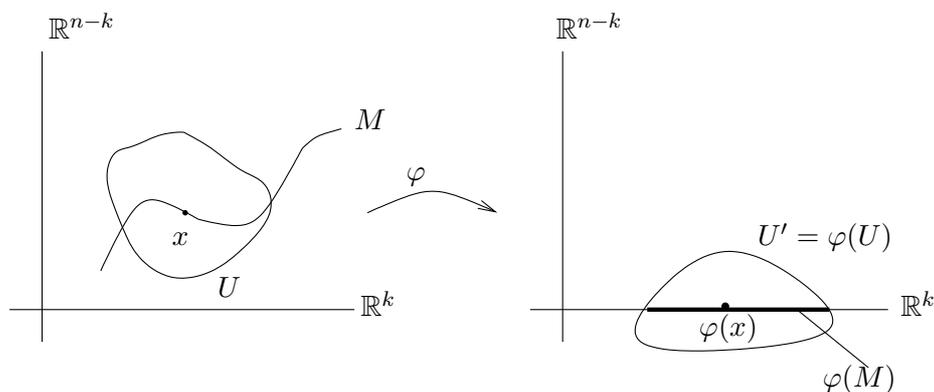
12.4 Untermannigfaltigkeiten des \mathbb{R}^n

Wir haben im vorigen Abschnitt Lösungsmengen von Gleichungen lokal als Funktionsgraphen dargestellt. Der folgende Begriff beschreibt allgemein Teilmengen des \mathbb{R}^n , die sich so darstellen lassen (unabhängig davon, ob sie Lösungsmenge einer Gleichung sind).

Definition 12.8 Eine Teilmenge $M \subseteq \mathbb{R}^n$ heißt k -dimensionale C^m -Untermannigfaltigkeit, wenn gilt: Für jedes $x \in M$ gibt es offene Mengen $U \subseteq \mathbb{R}^n$ und $U' \subseteq \mathbb{R}^n$ mit $x \in U$ sowie einen C^m -Diffeomorphismus

$$\varphi : U \rightarrow U' \quad \text{mit } \varphi(U \cap M) = U' \cap (\mathbb{R}^k \times \{0\}).$$

Eine solche Abbildung heißt Umgebungskarte. Eine Familie $(\varphi_j)_{j \in J}$ von Umgebungskarten $\varphi_j : U_j \rightarrow U'_j$ von M heißt Umgebungsatlas von M , wenn $M \subseteq \bigcup_{j \in J} U_j$.



Eine Mannigfaltigkeit ist also eine Menge, die in geeigneten krummlinigen Koordinaten (beschrieben durch φ) lokal wie der \mathbb{R}^k im \mathbb{R}^n aussieht.

Beispiel 1: Funktionsgraphen. Sei $V \subseteq \mathbb{R}^k$ offen und $f : V \rightarrow \mathbb{R}^n$ stetig differenzierbar. Wir zeigen, dass der Graph von f , d.h. die Menge

$$M := \{(x, f(x)) \in \mathbb{R}^{k+n} : x \in V\}$$

eine k -dimensionale Untermannigfaltigkeit von \mathbb{R}^{k+n} ist. Dazu sei $U := V \times \mathbb{R}^n$. Diese Menge ist offen und enthält M (ist also eine offene Umgebung jedes Punktes von M). Die Abbildung

$$\varphi : U \rightarrow U, \quad (x, y) \mapsto (x, y - f(x))$$

ist ein Diffeomorphismus von U mit der Umkehrabbildung

$$\varphi^{-1} : U \rightarrow U, \quad (x, y) \mapsto (x, y + f(x)).$$

Weiter ist

$$\varphi(U \cap M) = \varphi(M) = V \times \{0\} = U \cap (\mathbb{R}^k \times \{0\}).$$

Die Abbildung φ liefert also einen einelementigen Umgebungsatlas von M . ■

Beispiel 2. Wir zeigen, dass die n -Sphäre

$$\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : \|x\|_2 = 1\} = \{x \in \mathbb{R}^{n+1} : \sum_{i=1}^{n+1} |x_i|^2 = 1\}$$

eine n -dimensionale Untermannigfaltigkeit des \mathbb{R}^{n+1} ist. Dazu betrachten wir für jedes $j = 1, \dots, n+1$ die Mengen

$$U_j^\pm := \{x \in \mathbb{R}^{n+1} : \sum_{i \neq j} x_i^2 < 1, \pm x_j > 0\}.$$

Jede dieser Mengen ist offen, und diese Mengen überdecken \mathbb{S}^n . Weiter: die Abbildungen $\varphi_j^\pm : U_j^\pm \rightarrow \mathbb{R}^{n+1}$,

$$(x_1, \dots, x_{n+1}) \mapsto (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_{n+1}, x_j \mp (1 - \sum_{i \neq j} x_i^2)^{1/2})$$

sind Diffeomorphismen von U_j^\pm auf $\varphi_j^\pm(U_j^\pm)$ (warum?), und es gilt

$$\varphi_j^\pm(U_j^\pm \cap \mathbb{S}^n) = \varphi_j^\pm(U_j^\pm) \cap (\mathbb{R}^n \times \{0\}).$$

Die $2(n+1)$ Umgebungskarten φ_j^\pm , $j = 1, \dots, n+1$, bilden also einen Umgebungsatlas von \mathbb{S}^n . ■

Beispiel 3. Sei $A : \mathbb{R}^n \rightarrow \mathbb{R}^k$ eine lineare Abbildung. Dann ist $\ker A$ eine lineare Mannigfaltigkeit der Dimension $\dim \ker A = n - \text{rang } A$ im Sinne der linearen Algebra. Der Kern von A ist auch eine Untermannigfaltigkeit von \mathbb{R}^n der Dimension $n - \text{rang } A$ im Sinne von Definition 12.8. Dazu schreiben wir \mathbb{R}^n als direkte Summe $\ker A \dot{+} N$, wählen lineare Isomorphismen

$$J_1 : \ker A \rightarrow \mathbb{R}^{n-\text{rang } A}, \quad J_2 : N \rightarrow \mathbb{R}^{\text{rang } A}$$

und definieren

$$\varphi : \mathbb{R}^n = \ker A \dot{+} N \rightarrow \mathbb{R}^n = \mathbb{R}^{n-\text{rang } A} \times \mathbb{R}^{\text{rang } A}, \quad x \dot{+} y \mapsto (J_1 x, J_2 y). \quad \blacksquare$$

Den folgenden Satz kann man als eine globale Version des Satzes über implizite Funktionen auffassen. Er gibt nicht nur Auskunft über die *lokale* Struktur der Lösungsmenge einer Gleichung, sondern über deren *globale* Struktur. Vorab noch eine Definition.

Definition 12.9 Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^k$ eine differenzierbare Funktion. Wir nennen $x \in U$ einen kritischen Punkt und $f(x) \in \mathbb{R}^k$ einen kritischen Wert, wenn

$$\text{rang } f'(x) < k.$$

Ein Punkt $y \in \mathbb{R}^k$ heißt regulärer Wert, wenn sein Urbild $f^{-1}(y)$ keine kritischen Punkte enthält.

Satz 12.10 (Rangatz) Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}^k$ eine C^m -Abbildung. Ist $w \in f(U)$ ein regulärer Wert, so ist das Urbild $f^{-1}(w)$ eine $n-k$ -dimensionale C^m -Untermannigfaltigkeit von \mathbb{R}^n .

Im Spezialfall $n = 2, k = 1$ heißen die Mengen $f^{-1}(w)$ Höhenlinien von f . Allgemeiner spricht man für $n \geq 2$ von Niveauflächen.

Beweis. Wir können davon ausgehen, dass $w = 0$ ist (andernfalls wenden wir die folgenden Überlegungen auf die Funktion $\tilde{f} : U \rightarrow \mathbb{R}^k, x \mapsto f(x) - w$ an). Sei also $u \in U$ ein Punkt mit $f(u) = 0$. Nach Voraussetzung ist

$$\text{rang } f'(u) = \text{rang} \left(\frac{\partial f_i}{\partial x_j} \right)_{i,j=1}^{k,n} = k.$$

O.B.d.A. können wir weiter annehmen, dass die ersten k Spalten der Matrix $\left(\frac{\partial f_i}{\partial x_j} \right)$ linear unabhängig sind (andernfalls nummerieren wir die Koordinaten x_j um). Sei nun

$$\varphi : U \rightarrow \mathbb{R}^n, \quad x \mapsto (f_1(x), \dots, f_k(x), x_{k+1}, \dots, x_n).$$

Dann ist

$$\varphi'(u) = \begin{pmatrix} \left(\frac{\partial f_i}{\partial x_j} \right)_{i=1,j=1}^{k,k} & \left(\frac{\partial f_i}{\partial x_j} \right)_{i=1,j=k+1}^{k,n} \\ 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

eine invertierbare Matrix. Nach dem Satz über die Umkehrfunktion gibt es eine offene Umgebung $W \subseteq U$ von u so, dass $\varphi|_W : W \rightarrow \varphi(W)$ ein Diffeomorphismus ist. Dann ist

$$\varphi(f^{-1}(0) \cap W) = (\{0\} \times \mathbb{R}^{n-k}) \cap \varphi(W).$$

Also ist φ eine Umgebungskarte von $f^{-1}(0)$ um u . Da $u \in f^{-1}(0)$ beliebig gewählt war, folgt die Behauptung. ■

Beispiel 4. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^4 - y^4$. Dann ist $f'(x, y) = (4x^3, -4y^3)$. Folglich ist $(0, 0) \in \mathbb{R}^2$ der einzige kritische Punkt und $w = 0 \in \mathbb{R}$ der einzige kritische Wert von f . Für $w \neq 0$ sind also alle Höhenlinien von f glatte eindimensionale Untermannigfaltigkeiten von \mathbb{R}^2 . Für $w = 0$ ist die zugehörige Höhenlinie

keine Untermannigfaltigkeit (sie besteht aus zwei Geraden, die sich in der Null schneiden). ■

Beispiel 5. Sei $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ symmetrisch und invertierbar und

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad x \mapsto \langle Ax, x \rangle.$$

Aus

$$\langle A(x+h), x+h \rangle - \langle Ax, x \rangle = \langle Ax, h \rangle + \langle Ah, x \rangle + \langle Ah, h \rangle$$

folgt sofort

$$f'(x)h = 2\langle Ax, h \rangle = 2x^T Ah,$$

also $f'(x) = 2x^T A \in L(\mathbb{R}^n, \mathbb{R})$. Da A invertierbar ist, ist $x \in \mathbb{R}^n$ genau dann kritisch, wenn $x = 0 \in \mathbb{R}^n$, und $w = 0 \in \mathbb{R}$ ist der einzige kritische Wert von f . Ist beispielsweise $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ mit $\lambda_i > 0$ und $t > 0$, so erhalten wir, dass die Ellipsoide

$$f^{-1}(t) = \{x \in \mathbb{R}^n : \sum_{i=1}^n \lambda_i x_i^2 = t\}$$

glatte $n-1$ -dimensionale Untermannigfaltigkeiten des \mathbb{R}^n sind. Insbesondere ist \mathbb{S}^{n-1} eine $n-1$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n (vgl. Beispiel 2). ■

12.5 Extrema unter Nebenbedingungen

In Kapitel 10 haben wir Extrema von Funktionen studiert, die auf einer offenen Teilmenge des \mathbb{R}^n definiert sind. Wir betrachten nun eine Situation, die in praktischen Problemen viel häufiger ist: wir suchen Extrema von Funktionen unter Nebenbedingungen, d.h. Extrema von Funktionen auf Untermannigfaltigkeiten des \mathbb{R}^n . Wir wollen dabei *vermeiden*, durch eine geeignete Parametrisierung der Untermannigfaltigkeit das Problem auf die in Kapitel 10 betrachtete Situation zurückführen, da dies in der Regel recht schwierig ist.

Wir präzisieren zunächst die Problemstellung. Sei $U \subseteq \mathbb{R}^{m+n}$ eine offene Menge, und $f : U \rightarrow \mathbb{R}$ sei stetig differenzierbar. Weiter sei $g : U \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Funktion, für die $0 \in \mathbb{R}^n$ ein regulärer Wert ist. Dann ist

$$M := \{x \in U : g(x) = 0\}$$

eine m -dimensionale C^1 -Untermannigfaltigkeit des \mathbb{R}^{m+n} . Wir sagen, dass $u \in U$ ein *lokales Minimum* (bzw. *Maximum*) von f unter der Nebenbedingung M oder $g = 0$ ist, wenn $u \in M$ und wenn es eine Umgebung V von u so gibt, dass $f(x) \geq f(u)$ (bzw. $f(x) \leq f(u)$) für alle $x \in V \cap M$.

Satz 12.11 *Unter den soeben getroffenen Voraussetzungen gilt: Ist $u \in U$ ein lokales Extremum von f unter der Nebenbedingung $g = 0$, so gibt es reelle Zahlen $\lambda_1, \dots, \lambda_n$ so, dass*

$$f'(u) = \sum_{i=1}^n \lambda_i g'_i(u). \quad (12.14)$$

Diese Zahlen λ_i heißen auch *Lagrange-Multiplikatoren*.

Beweis. Nach Satz 12.10 ist M eine m -dimensionale Untermannigfaltigkeit von \mathbb{R}^{m+n} . Wir finden also eine Umgebung $W \subseteq U$ von u und einen Diffeomorphismus $\varphi : W \rightarrow \varphi(W) \subseteq \mathbb{R}^{m+n}$ mit

$$\varphi(W \cap M) = \varphi(W) \cap (\mathbb{R}^m \times \{0\}).$$

Wir schreiben $\varphi(W) \cap (\mathbb{R}^m \times \{0\})$ als

$$\{(y, 0) \in \mathbb{R}^m \times \mathbb{R}^n : y \in V\}$$

mit einer offenen Menge $V \subseteq \mathbb{R}^m$ und definieren

$$\psi : V \rightarrow U, \quad y \mapsto \varphi^{-1}(y, 0). \quad (12.15)$$

Offenbar ist ψ eine Bijektion von V auf $W \cap M$. Die Abbildung ψ ist auch stetig differenzierbar. Mit der linearen Abbildung

$$E : \mathbb{R}^m \rightarrow \mathbb{R}^m \times \mathbb{R}^n, \quad y \mapsto (y, 0)$$

können wir nämlich ψ auffassen als Einschränkung von $\varphi^{-1} \circ E$ auf V . Die Einschränkung $E|_V$ ist aber differenzierbar, und ihre Ableitung ist in jedem Punkt gleich E . Aus $\varphi \circ \psi = E|_V$ folgt außerdem mit der Kettenregel $\varphi'(\psi(y)) \cdot \psi'(y) = E$. Für $v := \varphi(u)$ ist also insbesondere $\text{rang } \psi'(v) = \text{rang } E = m$.

Nach diesen Vorüberlegungen nun zum eigentlichen Beweis. Die Funktion $f \circ \psi : V \rightarrow \mathbb{R}$ hat in v ein lokales Extremum (ohne Nebenbedingungen!). Es ist also

$$(f \circ \psi)'(v) = f'(\psi(v)) \circ \psi'(v) = f'(u) \circ \psi'(v) = 0. \quad (12.16)$$

Außerdem gilt für jedes $i = 1, \dots, n$, dass $g_i(\psi(y)) = 0$ für $y \in V$. Differenzieren liefert $g'_i(\psi(y)) \cdot \psi'(y) = 0$ und insbesondere

$$g'_i(u) \circ \psi'(v) = 0 \quad \text{für } i = 1, \dots, n. \quad (12.17)$$

Damit ist klar, dass $\text{im } \psi'(v) \subseteq \ker g'(u)$. Nun ist aber

$$\dim \ker g'(u) + \text{rang } g'(u) = \dim \ker g'(u) + n = m + n,$$

also $\dim \ker g'(u) = m$. Hieraus folgt wegen $\text{rang } \psi'(v) = m$ sogar

$$\text{im } \psi'(v) = \ker g'(u).$$

Ist nun $w \in \text{im } \psi'(v)$, so folgt aus (12.16) und (12.17), dass

$$f'(u)w - \lambda_1 g'_1(u)w - \dots - \lambda_n g'_n(u)w = 0 \quad (12.18)$$

für beliebige $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Um die Behauptung zu zeigen, müssen wir noch die Vektoren aus dem orthogonalen Komplement W von $\psi'(v)$ betrachten. Sei w_1, \dots, w_n eine Basis von W . Wir bestimmen $\lambda_1, \dots, \lambda_n$ aus dem linearen Gleichungssystem

$$\begin{aligned} f'(u)w_1 &- \lambda_1 g'_1(u)w_1 - \dots - \lambda_n g'_n(u)w_1 &= 0 \\ &\vdots & \\ f'(u)w_n &- \lambda_1 g'_1(u)w_n - \dots - \lambda_n g'_n(u)w_n &= 0. \end{aligned}$$

Dazu müssen wir uns lediglich klarmachen, dass die Matrix

$$\begin{pmatrix} g'_1(u)w_1 & \dots & g'_n(u)w_1 \\ \vdots & & \vdots \\ g'_1(u)w_n & \dots & g'_n(u)w_n \end{pmatrix}$$

invertierbar ist. Wäre sie das nicht, so wären ihre Zeilen linear abhängig. Es gäbe also Zahlen μ_1, \dots, μ_n , die nicht alle gleich Null sind, so dass

$$\mu_1(g'_1(u)w_1, \dots, g'_n(u)w_1) + \dots + \mu_n(g'_1(u)w_n, \dots, g'_n(u)w_n) = 0$$

bzw.

$$g'_1(u)w = \dots = g'_n(u)w = 0$$

mit $w = \mu_1 w_1 + \dots + \mu_n w_n$. Der Vektor w liegt also im Kern von $g'(u)$. Da $w \in W$, und da W senkrecht zu diesem Kern steht, folgt $w = 0$. Da aber die w_i linear unabhängig sind, folgt $\mu_1 = \dots = \mu_n = 0$, ein Widerspruch.

Wir können also die λ_i aus dem Gleichungssystem eindeutig bestimmen. Da die w_i den Raum W aufspannen, folgt

$$f'(u)w - \lambda_1 g'_1(u)w - \dots - \lambda_n g'_n(u)w = 0$$

für alle $w \in W$ und folglich für alle $w \in \mathbb{R}^{m+n}$. ■

Zur praktischen Bestimmung lokaler Extrema unter Nebenbedingungen muss man also das System

$$\begin{aligned} \frac{\partial f}{\partial x_k}(x) &= \lambda_1 \frac{\partial g_1}{\partial x_k}(x) + \dots + \lambda_n \frac{\partial g_n}{\partial x_k}(x), & k &= 1, \dots, m+n, \\ g_\ell(x) &= 0, & \ell &= 1, \dots, n, \end{aligned} \tag{12.19}$$

bestehend aus $m+2n$ Gleichungen für die $m+2n$ Unbekannten $x = (x_1, \dots, x_{m+n})$ und $\lambda_1, \dots, \lambda_n$ lösen und die erhaltenen extremwertverdächtigen Punkte auf ihre Extremaleigenschaften untersuchen. Es ist hilfreich, sich die Funktion

$$G(x_1, \dots, x_{m+n}, \lambda_1, \dots, \lambda_n) := f(x) - \lambda_1 g_1(x) - \dots - \lambda_n g_n(x)$$

zu merken. Sucht man nämlich die extremwertverdächtigen Punkte von G durch partielles Ableiten nach jeder der Variablen x_i, λ_j und anschließendes Nullsetzen, so gelangt man gerade zum System (12.19).

Beispiel 1. Gesucht sind Maximum und Minimum der Funktion $f(x, y) = xy$ unter der Nebenbedingung $g(x, y) := x^2 + y^2 - 1 = 0$, d.h. auf der Einheitskreislinie. Die Matrix $g'(x, y) = (2x, 2y)$ hat offenbar in jedem Punkt der Einheitskreislinie den Rang 1, so dass Satz 12.11 anwendbar ist. Wir betrachten die Funktion

$$G(x, y, \lambda) = f(x, y) - \lambda g(x, y) = xy - \lambda(x^2 + y^2 - 1)$$

und setzen ihre partiellen Ableitungen Null:

$$\begin{aligned} \frac{\partial G}{\partial x} = 0 & \Leftrightarrow y - 2\lambda x = 0, \\ \frac{\partial G}{\partial y} = 0 & \Leftrightarrow x - 2\lambda y = 0, \\ \frac{\partial G}{\partial \lambda} = 0 & \Leftrightarrow x^2 + y^2 = 1. \end{aligned}$$

Wir multiplizieren die erste Gleichung mit y und die zweite mit x und finden $x^2 = y^2$, weshalb zusammen mit der dritten Gleichung für (x, y) die Paare

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

in Frage kommen. Diese lösen tatsächlich das System (die ersten beiden Paare mit $\lambda = 1/2$, die anderen mit $\lambda = -1/2$). Als Funktionswert ergibt sich für die ersten beiden Paare $1/2$, für die anderen $-1/2$. Da die Funktion f Maximum und Minimum auf der Einheitskreislinie besitzen *muss*, ist $1/2$ das Maximum und $-1/2$ das Minimum von f . ■

Beispiel 2. Sei A eine symmetrische $k \times k$ Matrix. Wir suchen die extremwertverdächtigen Punkte von $f(x) = \langle Ax, x \rangle$ auf der Sphäre $\mathbb{S}^{k-1} \subseteq \mathbb{R}^k$, d.h. unter der Nebenbedingung $g(x) = \|x\|^2 - 1 = 0$. Hier ist $n = 1$ und $m = k - 1$, also $m + 2n = k + 1$.

Die Ableitung von $f(x) = \langle Ax, x \rangle$ ist $f'(x) = 2x^T A$ (vgl. Beispiel 5 in Abschnitt 12.4) und die von $g(x) = \|x\|^2 - 1 = \langle x, x \rangle - 1$ ist $g'(x) = 2x^T$. Damit ist klar, dass die Rangbedingung erfüllt ist. Das zu lösende Gleichungssystem ist also

$$\begin{aligned} f'(x) - \lambda g'(x) = 2x^T A - 2\lambda x^T &= 0, \\ g(x) = \|x\|^2 - 1 &= 0, \end{aligned}$$

bzw. nach Transponieren

$$Ax = \lambda x, \quad \|x\|^2 = 1.$$

Ein Punkt $x \in \mathbb{S}^{k-1}$ ist also genau dann extremwertverdächtig, wenn er Eigenvektor zum Eigenwert λ ist. Wegen $f(x) = \langle Ax, x \rangle$ ergibt sich für die zugehörigen Funktionswerte

$$f(x) = \langle Ax, x \rangle = \lambda \langle x, x \rangle = \lambda.$$

Folglich wird f in x genau dann minimal (bzw. maximal), wenn x ein Eigenvektor zum kleinsten (bzw. größten) Eigenwert von A ist. Man beachte: die stetige Funktion f nimmt auf der kompakten Menge \mathbb{S}^{k-1} ihr Maximum an, woraus folgt, dass A mindestens einen reellen Eigenwert besitzen muss. Es ist also

$$\begin{aligned} \max_{x \neq 0} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} &= \text{größter Eigenwert von } A, \\ \min_{x \neq 0} \frac{\langle Ax, x \rangle}{\langle x, x \rangle} &= \text{kleinster Eigenwert von } A. \quad \blacksquare \end{aligned}$$

13 Das Riemann-Integral für Funktionen mehrerer Veränderlicher

In diesem Kapitel kommen wir zur Definition und wesentlichen Eigenschaften des Riemann-Integrals für Funktionen mehrerer Veränderlicher. So wie das Problem der Flächenberechnung eine Motivation für die Einführung des Integrals für Funktionen einer Veränderlichen war, kann nun das Problem der Volumendefinition- und berechnung (oder allgemeiner etwa das Problem der Massebestimmung eines Körpers mit ortsabhängiger Dichte) als Motivation dienen.

In diesem und im folgenden Kapitel kann es lediglich darum gehen, einen ersten Eindruck von der Integration im \mathbb{R}^n zu gewinnen und einige Rechentechniken zu vermitteln. Im vierten Semester wenden wir uns diesem Thema ausführlicher zu.

13.1 Das Riemann-Integral über Intervallen im \mathbb{R}^n

Wir beginnen mit der Integration über den „einfachsten“ Mengen im \mathbb{R}^n , nämlich über achsenparallele Intervalle, Rechtecke, Quader u.s.w., die wir kurz unter dem Namen *Intervall im \mathbb{R}^n* zusammenfassen. Im Unterschied zum \mathbb{R}^1 gibt es natürlich im \mathbb{R}^n weitaus mehr Mengen, über die man integrieren möchte, etwa Kugeln oder Kegel. Wir werden daher später die in diesem Abschnitt angestellten Überlegungen auf allgemeinere Mengen übertragen.

Unter einem *abgeschlossenen Intervall* im \mathbb{R}^n verstehen wir ein Produkt

$$[a_1, b_1] \times \dots \times [a_n, b_n] = \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_i \leq x_i \leq b_i \text{ für alle } i\}.$$

Ein *offenes Intervall* im \mathbb{R}^n ist ein Produkt

$$(a_1, b_1) \times \dots \times (a_n, b_n) = \{(x_1, \dots, x_n) \in \mathbb{R}^n : a_i < x_i < b_i \text{ für alle } i\}.$$

Ist I ein offenes oder abgeschlossenes Intervall wie oben, so erklären wir seinen *Inhalt* durch

$$|I| := (b_1 - a_1) \cdot \dots \cdot (b_n - a_n).$$

Wir erhalten also für

- $n = 1$ gewöhnliche Intervalle, und $|I|$ ist die Intervalllänge,
- $n = 2$ Rechtecke, und $|I|$ ist der Flächeninhalt,
- $n = 3$ Quader, und $|I|$ ist das Volumen.

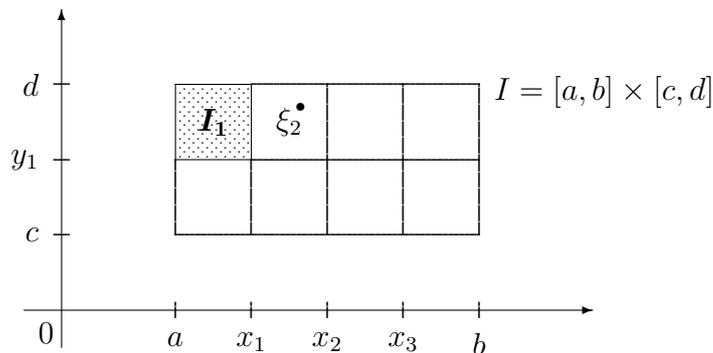
Eine *Zerlegung* Z eines Intervalls $I \subseteq \mathbb{R}^n$ ist ein Produkt $Z_1 \times \dots \times Z_n$, wobei Z_i eine Zerlegung von $[a_i, b_i]$ ist. Die *Teilintervalle* von Z erhält man, indem man im Produkt $T_1 \times \dots \times T_n$ die T_i alle Teilintervalle der Zerlegung Z_i von $[a_i, b_i]$

durchlaufen lässt. Eine Zerlegung Z' heißt *Verfeinerung* von Z , wenn $Z \subseteq Z'$. Ist insbesondere $Z = Z_1 \times \dots \times Z_n$ und $Z' = Z'_1 \times \dots \times Z'_n$, so ist Z' genau dann eine Verfeinerung von Z , wenn jedes Z'_i eine Verfeinerung von Z_i ist.

Unter dem *Feinheitsmaß* (Maschenweite) von $Z = Z_1 \times \dots \times Z_n$ versteht man die Zahl

$$|Z| := \max_{1 \leq i \leq n} |Z_i|,$$

wobei $|Z_i|$ das Feinheitsmaß der Zerlegung Z_i von $[a_i, b_i]$ ist. Man beachte, dass in diese Definition nicht der Inhalt der Teilintervalle von Z eingeht, sondern deren Kantenlänge.



Hat Z die Teilintervalle I_1, \dots, I_r , so heißt jeder Vektor $\xi = (\xi_1, \dots, \xi_r)$ mit $\xi_i \in I_i$ ein *Zwischenvektor* zu Z . Ist schließlich f eine reellwertige Funktion auf I , Z eine Zerlegung von I mit den Teilintervallen I_1, \dots, I_r und $\xi = (\xi_1, \dots, \xi_r)$ ein Zwischenvektor zu Z , so heißt

$$S(Z, \xi, f) := \sum_{i=1}^r f(\xi_i) |I_i|$$

eine *Riemannsumme* für f .

Definition 13.1 Die Funktion $f : I \rightarrow \mathbb{R}$ heißt *Riemann-integrierbar*, wenn für jede Folge $(Z^{(m)})$ von Zerlegungen von I mit $|Z^{(m)}| \rightarrow 0$ und für jede zugehörige Folge von Zwischenvektoren $(\xi^{(m)})$ die Folge $(S(Z^{(m)}, \xi^{(m)}, f))$ der entsprechenden Riemannsummen konvergiert.

Wenn f Riemann-integrierbar ist, dann sieht man wie im Beweis von Satz 8.11, dass alle Folgen $(S(Z^{(m)}, \xi^{(m)}, f))$ gegen den gleichen Wert konvergieren. Dieser heißt das *Riemann-Integral* von f über I , und wir schreiben dafür

$$\int_I f dx, \quad \int_I f(x) dx, \quad \int_I f(x_1, \dots, x_n) d(x_1, \dots, x_n) \quad \text{oder} \quad \int_I f dV.$$

Die folgenden Aussagen beweist man wie für $n = 1$ (Sätze 8.4, 8.23, 8.24, 8.25 und 8.20).

Satz 13.2 Jede auf einem Intervall $I \subseteq \mathbb{R}^n$ Riemann-integrierbare Funktion ist beschränkt.

Satz 13.3 Sind f, g Riemann-integrierbar auf I und $\alpha, \beta \in \mathbb{R}$, so ist auch $\alpha f + \beta g$ Riemann-integrierbar auf I , und es gilt

$$\int_I (\alpha f + \beta g) dx = \alpha \int_I f dx + \beta \int_I g dx.$$

Satz 13.4 Sind f, g Riemann-integrierbar auf I und ist $f(x) \geq g(x)$ für alle $x \in I$, dann ist auch

$$\int_I f dx \geq \int_I g dx.$$

Folgerung 13.5 Für Riemann-integrierbares $f : I \rightarrow \mathbb{R}$ ist

$$\left| \int_I f dx \right| \leq \sup_{x \in I} |f(x)| \cdot |I|.$$

Satz 13.6 Sind f, g Riemann-integrierbar auf I und stimmen f und g auf einer in I dichten Menge überein, so ist bereits

$$\int_I f dx = \int_I g dx.$$

13.2 Integrabilitätskriterien

13.2.1 Charakterisierung über Darboux'sche Integrale

Sei $I \subseteq \mathbb{R}^n$ ein abgeschlossenes Intervall und $f : I \rightarrow \mathbb{R}$ eine beschränkte Funktion. Weiter sei Z eine Zerlegung von I in Teilintervalle I_1, \dots, I_r . Mit den Zahlen

$$m_k := \inf_{x \in I_k} f(x) \quad \text{und} \quad M_k := \sup_{x \in I_k} f(x)$$

definieren wir die Unter- bzw. Obersummen von f bzgl. Z durch

$$U(Z, f) := \sum_{k=1}^r m_k |I_k| \quad \text{bzw.} \quad O(Z, f) := \sum_{k=1}^r M_k |I_k|$$

und nennen

$$\int_{I^*} f dx := \sup_Z U(Z, f) \quad \text{bzw.} \quad \int_I^* f dx := \inf_Z O(Z, f)$$

das untere bzw. obere Darboux'sche Integral von f . Für einen kurzen Moment wollen wir eine Funktion f *Darboux-integrierbar* nennen, wenn

$$\int_{*I} f dx = \int_I^* f dx.$$

Wie Satz 8.8 beweist man:

Satz 13.7 Die beschränkte Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann Darboux-integrierbar, wenn für jedes $\varepsilon > 0$ eine Zerlegung Z von I mit $O(Z, f) - U(Z, f) < \varepsilon$ existiert.

Durch Übertragung des Beweises von Satz 8.9 (Details siehe Heuser, Analysis II, Satz 199.1) erhält man weiter:

Satz 13.8 Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie beschränkt und Darboux-integrierbar ist. In diesem Fall gilt

$$\int_I f \, dx = \int_{I^*} f \, dx = \int_I^* f \, dx.$$

Hieaus erhält man wir für $n = 1$ das *Riemannsches Integrabilitätskriterium*:

Folgerung 13.9 Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie beschränkt ist und wenn für jedes $\varepsilon > 0$ eine Zerlegung Z von I mit $O(Z, f) - U(Z, f) < \varepsilon$ existiert.

13.2.2 Charakterisierung über Nullmengen

Wie im \mathbb{R}^1 nennen wir eine Menge $M \subseteq \mathbb{R}^n$ eine *Nullmenge*, wenn es für jedes $\varepsilon > 0$ höchstens abzählbar viele (offene oder abgeschlossene) Intervalle I_1, I_2, \dots gibt, welche M überdecken (d.h. $M \subseteq \bigcup_{k \geq 1} I_k$) und für die $\sum_{k \geq 1} |I_k| < \varepsilon$ ist.

Beispiel. Für jedes $c \in \mathbb{R}$ und jedes $j = 1, \dots, n$ ist die Hyperebene

$$H := \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_j = c\}$$

eine Nullmenge. Um dies einzusehen, bilden wir für jedes $k \in \mathbb{N}$ das Intervall $I_k := (a_1, b_1) \times \dots \times (a_n, b_n)$ mit

$$a_i := \begin{cases} -k & \text{für } i \neq j \\ c - \frac{\varepsilon}{2^{k+1}(2k)^{n-1}} & \text{für } i = j, \end{cases} \quad b_i := \begin{cases} k & \text{für } i \neq j \\ c + \frac{\varepsilon}{2^{k+1}(2k)^{n-1}} & \text{für } i = j. \end{cases}$$

Dann ist $H \subseteq \bigcup_k I_k$ (die Intervalle I_k werden immer „breiter“ und „flacher“) und

$$\sum_{k=1}^{\infty} |I_k| = \sum_{k=1}^{\infty} (2k)^{n-1} \cdot \frac{2\varepsilon}{2^{k+1}(2k)^{n-1}} = \varepsilon \sum_{k=1}^{\infty} \frac{1}{2^k} = \varepsilon. \quad \blacksquare$$

Es lassen sich die Beweise von Lemma 8.14 (Eigenschaften von Nullmengen) und Satz 8.15 auf den \mathbb{R}^n mit $n > 1$ übertragen. Auch die in diesen Beweisen benutzte *Schwankung* einer Funktion f auf einem Intervall I definiert man wie im \mathbb{R}^1 :

$$\Omega_f(I) := \sup_{t \in I} f(t) - \inf_{t \in I} f(t) = \sup\{|f(s) - f(t)| : s, t \in I\}.$$

Satz 13.10 (Lebesguesches Integrabilitätskriterium) Eine Funktion $f : I \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar auf I , wenn sie beschränkt ist und wenn die Menge ihrer Unstetigkeitsstellen eine Nullmenge bildet, d.h. wenn f fast überall stetig ist.

Folgerung 13.11 Jede auf I stetige Funktion ist Riemann-integrierbar.

Folgerung 13.12 Mit f und g sind auch die Funktionen $|f|$, fg , $\max\{f, g\}$ und $\min\{f, g\}$ Riemann-integrierbar auf I .

Folgerung 13.13 Sei g Riemann-integrierbar auf I , $g(I) \subseteq [a, b]$ und f stetig auf $[a, b]$. Dann ist auch $f \circ g$ Riemann-integrierbar auf I .

Beweis. Ist $f \circ g$ in $x \in I$ unstetig, so muss auch g in x unstetig sein. Also ist $\Delta(f \circ g) \subseteq \Delta(g)$. Damit ist $\Delta(f \circ g)$ eine Nullmenge. ■

Für Riemann-integrierbares f ist also z.B. auch die Funktion $\sqrt{|f(x)|}$ Riemann-integrierbar.

Folgerung 13.14 Sind f, g Riemann-integrierbar auf I , und sind f, g fast überall gleich, so ist

$$\int_I f \, dx = \int_I g \, dx.$$

13.3 Der Satz von Fubini

Der nachfolgende Satz gibt uns in vielen Fällen ein bequemes Verfahren in die Hand, Riemann-Integrale auf mehrdimensionalen Intervallen zu berechnen.

Satz 13.15 (Fubini) Seien $I_k \subseteq \mathbb{R}^k$ und $I_l \subseteq \mathbb{R}^l$ abgeschlossene Intervalle und $I := I_k \times I_l \subseteq \mathbb{R}^{k+l} := \mathbb{R}^k \times \mathbb{R}^l$. Weiter sei $f : I \rightarrow \mathbb{R}$ Riemann-integrierbar auf I , und für jedes $y \in I_l$ existiere das Riemann-Integral

$$g(y) := \int_{I_k} f(x, y) \, dx.$$

Dann ist die Funktion g auf I_l Riemann-integrierbar, und es gilt

$$\int_I f(x, y) \, d(x, y) = \int_{I_l} \left(\int_{I_k} f(x, y) \, dx \right) dy. \quad (13.1)$$

Beweis. Die Existenz des iterierten Integrals sowie die Identität (13.1) folgen aus dem Satz über iterierte Grenzwerte von Doppelfolgen (Satz 9.18). Wir sehen uns einige Details des Beweises an.

Für $m \in \mathbb{N}$ sei $Z_k^{(m)}$ eine Zerlegung des Intervalls I_k und $\xi_k^{(m)} = (\xi_{k,i}^{(m)})_{i=1}^k$ ein zugehöriger Zwischenvektor, und für $n \in \mathbb{N}$ sei $Z_l^{(n)}$ eine Zerlegung von I_l

mit Zwischenvektor $\xi_l^{(n)} = (\xi_{l,j}^{(n)})_{j=1}^l$. Dann definiert $Z^{(m,n)} := Z_k^{(m)} \times Z_l^{(n)}$ eine Zerlegung von I , und $\xi^{(m,n)} := (\xi_k^{(m)}, \xi_l^{(n)})$ ist ein zugehöriger Zwischenvektor. Wir erhalten

$$\begin{aligned} S(Z^{(m,n)}, \xi^{(m,n)}, f) &:= \sum_{(i,j)} f(\xi_{k,i}^{(m)}, \xi_{l,j}^{(n)}) |I_{k,i}^{(m)}| |I_{l,i}^{(n)}| \\ &= \sum_{j=1}^l \left(\sum_{i=1}^k f(\xi_{k,i}^{(m)}, \xi_{l,j}^{(n)}) |I_{k,i}^{(m)}| \right) |I_{l,i}^{(n)}|. \end{aligned} \quad (13.2)$$

Seien nun die Zerlegungsfolgen $(Z_k^{(m)})_{m \geq 0}$, $(Z_l^{(n)})_{n \geq 0}$ so beschaffen, dass $|Z_k^{(m)}| \rightarrow 0$ und $|Z_l^{(n)}| \rightarrow 0$ für $m \rightarrow \infty$ bzw. $n \rightarrow \infty$. Für die Produktzerlegung $Z^{(m,n)}$ gilt dann offenbar $|Z^{(m,n)}| \rightarrow 0$ für $(m,n) \rightarrow \infty$. Da f auf I Riemann-integrierbar ist, konvergiert die linke Seite von (13.2) für $(m,n) \rightarrow \infty$ gegen $\int_I f(x,y) d(x,y)$. Außerdem wissen wir aus der Voraussetzung, dass für $m \rightarrow \infty$ für jedes feste $\xi_{l,j}^{(n)}$ der Klammerterm auf der rechten Seite von (13.2) konvergiert:

$$\lim_{m \rightarrow \infty} \sum_i f(\xi_{k,i}^{(m)}, \xi_{l,j}^{(n)}) |I_{k,i}^{(m)}| = \int_{I_k} f(x, \xi_{l,j}^{(n)}) dx = g(\xi_{l,j}^{(n)}).$$

Nach dem erwähnten Satz 9.18 existiert dann auch der iterierte Grenzwert

$$\lim_{n \rightarrow \infty} \sum_j g(\xi_{l,j}^{(n)}) |I_{l,j}^{(n)}| = \int_{I_l} g(y) dy = \int_{I_l} \left(\int_{I_k} g(x,y) dx \right) dy$$

und stimmt mit

$$\lim_{(m,n) \rightarrow \infty} S(Z^{(m,n)}, \xi^{(m,n)}, f) = \int_I f(x,y) d(x,y)$$

überein. ■

Folgerung 13.16 (Vertauschen der Integrationsreihenfolge) *Mit den Bezeichnungen aus Satz 13.15 gilt: Ist f auf $I = I_k \times I_l$ Riemann-integrierbar, und existieren die Integrale*

$$\int_{I_k} f(x,y) dx \text{ für jedes } y \in I_l \quad \text{und} \quad \int_{I_l} f(x,y) dy \text{ für jedes } x \in I_k,$$

so existieren alle iterierten Integrale, und es gilt

$$\int_{I_k} \left(\int_{I_l} f(x,y) dy \right) dx = \int_{I_l} \left(\int_{I_k} f(x,y) dx \right) dy = \int_I f(x,y) d(x,y).$$

Für stetiges f wissen wir dies bereits aus Satz 10.29.

Durch wiederholtes Anwenden des Satzes von Fubini und von Folgerung 13.16 erhalten wir:

Folgerung 13.17 Ist f stetig auf $I = [a_1, b_1] \times \dots \times [a_n, b_n]$, so ist

$$\int_I f(x_1, \dots, x_n) d(x_1, \dots, x_n) = \int_{a_1}^{b_1} (\dots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \dots) dx_1.$$

Dabei darf die Reihenfolge der Integrationen beliebig vertauscht werden.

Beispiel. Auf $I = [0, 1] \times [0, 1] \times [0, 1]$ sei $f(x, y, z) := xyz$. Dann ist

$$\begin{aligned} \int_I f(x, y, z) d(x, y, z) &= \int_0^1 \int_0^1 \int_0^1 xyz \, dx \, dy \, dz \\ &= \frac{1}{2} \int_0^1 \int_0^1 yz \, dy \, dz = \frac{1}{4} \int_0^1 z \, dz = \frac{1}{8}. \quad \blacksquare \end{aligned}$$

13.4 Integration über Jordan-messbaren Mengen

In diesem Abschnitt geht es um die Integration auf komplizierteren Mengen als Intervallen. Für jede nichtleere Menge $B \subseteq \mathbb{R}^n$ und jede Funktion $f : B \rightarrow \mathbb{R}$ definieren wir

$$f_B : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f_B(x) := \begin{cases} f(x) & \text{falls } x \in B, \\ 0 & \text{falls } x \notin B, \end{cases}$$

d.h. f_B setzt die Funktion f durch 0 auf ganz \mathbb{R}^n fort.

Definition 13.18 Sei $B \subseteq \mathbb{R}^n$ nichtleer und beschränkt und $I \subseteq \mathbb{R}^n$ ein abgeschlossenes Intervall mit $B \subseteq I$. Die Funktion $f : B \rightarrow \mathbb{R}$ heißt Riemann-integrierbar auf B , wenn die Funktion f_B Riemann-integrierbar auf I ist. In diesem Fall heißt

$$\int_B f \, dx := \int_I f_B \, dx$$

das Riemannintegral von f über B .

Man kann zeigen, dass diese Definition unabhängig von der Wahl von I ist. Man beachte auch, dass diese Definition auch im \mathbb{R}^1 etwas Neues bietet, da B kein Intervall sein muss.

Ob eine Funktion f auf einer Menge B integrierbar ist, hängt sowohl von f als auch von B ab. Insbesondere erwartet man von B , dass wenigstens so einfache Funktionen wie $\chi : B \rightarrow \mathbb{R}, x \mapsto 1$ Riemann-integrierbar sind. Die entsprechende Funktion

$$\chi_B : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \chi_B(x) = \begin{cases} 1 & \text{wenn } x \in B \\ 0 & \text{wenn } x \notin B \end{cases}$$

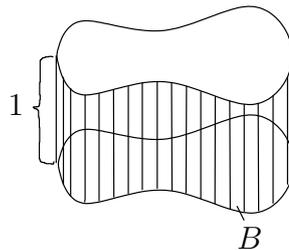
heißt die charakteristische Funktion von B .

Definition 13.19 Eine nichtleere beschränkte Menge $B \subseteq \mathbb{R}^n$ heißt Jordan-messbar, wenn ihre charakteristische Funktion χ_B Riemann-integrierbar ist. In diesem Fall heißt

$$|B| := \int_I \chi_B dx = \int_B 1 dx = \int_B dx$$

der (n -dimensionale) Jordan-Inhalt von B .

Geometrische Deutung:

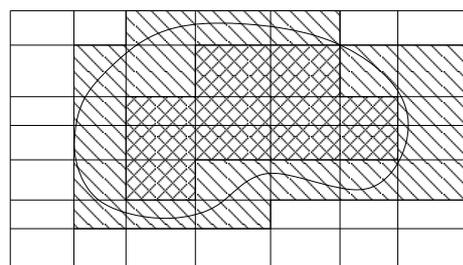


$\int_B 1 dx$ beschreibt das Volumen eines Zylinders über B mit der Höhe 1. Dieses ist gleich Grundfläche \times Höhe, also gleich $|B|$. ■

Deutung über Ober- und Untersummen: Seien $B \subseteq \mathbb{R}^n$ nichtleer und beschränkt, $I \subseteq \mathbb{R}^n$ ein Intervall mit $B \subseteq I$ und Z eine Zerlegung von I in Teilintervalle I_1, \dots, I_r . Dann ist

$$\inf_{x \in I_k} \chi_B(x) = \begin{cases} 1 & \text{falls } I_k \subseteq B, \\ 0 & \text{sonst,} \end{cases}$$

$$\sup_{x \in I_k} \chi_B(x) = \begin{cases} 1 & \text{falls } I_k \cap B \neq \emptyset, \\ 0 & \text{sonst.} \end{cases}$$



$$\begin{array}{cc} \text{[Schraffiert]} & O(Z, \chi_B) \quad \text{[Nicht schraffiert]} & U(Z, \chi_B) \end{array}$$

Für die zugehörigen Unter- bzw. Obersummen gilt

$$U(Z, \chi_B) = \sum_k' |I_k|, \quad O(Z, \chi_B) = \sum_k'' |I_k|,$$

wobei \sum' bzw. \sum'' über alle k mit $I_k \subseteq B$ bzw. mit $I_k \cap B \neq \emptyset$ erstreckt wird. Die Darboux'schen Integrale

$$\int_{*I} \chi_B dx, \quad \int_I^* \chi_B dx$$

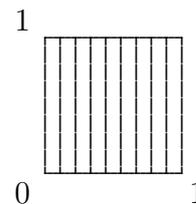
heißen *innerer* bzw. *äußerer Inhalt* von B . Aus dem Riemannsches Integrabilitätskriterium folgt sofort:

Folgerung 13.20 *Eine nichtleere beschränkte Menge $B \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn ihr innerer und ihr äußerer Inhalt übereinstimmen. In diesem Fall ist*

$$|B| = \int_{*I} \chi_B dx = \int_I^* \chi_B dx.$$

Beispiel. Sei $B := \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1, x \text{ rational}\}$. Als umgebendes Intervall wählen wir das Quadrat $[0, 1] \times [0, 1]$.

Dann gilt für jede Zerlegung Z von $[0, 1] \times [0, 1]$ $U(Z, \chi_B) = 0$ und $O(Z, \chi_B) = 1$. Also ist der innere Inhalt von B gleich 0 und der äußere gleich 1. B ist somit *nicht* Jordan-messbar, und wir schreiben B *keinen* Flächeninhalt zu.



Eine Anwendung des Lebesgueschen Integrabilitätskriteriums liefert sofort

Satz 13.21 *Eine nichtleere beschränkte Menge $B \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn ihr Rand ∂B eine Nullmenge ist.*

Beweis. Nach Definition und dem Lebesgueschen Kriterium ist B genau dann Jordan-messbar, wenn die Menge der Unstetigkeitsstellen von χ_B eine Nullmenge ist. Man macht sich leicht klar, dass χ_B genau dann in $x \in \mathbb{R}^n$ unstetig ist, wenn x ein Randpunkt von B ist. ■

Satz 13.22 (Allgemeines Lebesguesches Integrabilitätskriterium) *Sei $B \subseteq \mathbb{R}^n$ nichtleer und beschränkt, und B sei Jordan-messbar. Eine Funktion $f : B \rightarrow \mathbb{R}$ ist genau dann auf B Riemann-integrierbar, wenn sie auf B beschränkt und fast überall stetig ist.*

Beweis. Sei $I \subseteq \mathbb{R}^n$ ein Intervall mit $B \subseteq I$. Sei f auf B Riemann-integrierbar. Dann ist f_B auf I Riemann-integrierbar. Nach Satz 13.10 ist f_B beschränkt und fast überall stetig auf I . Dann ist f auch beschränkt und fast überall stetig auf B . Ist umgekehrt f beschränkt und fast überall stetig auf B , so ist f_B beschränkt auf I , und für die Menge der Unstetigkeitsstellen gilt: $\Delta(f_B) \subseteq \Delta(f) \cup \partial B$. Nach Satz 13.21 ist ∂B eine Nullmenge. Also ist $\Delta(f_B)$ Nullmenge, d.h. f_B ist auf I Riemann-integrierbar, und f ist auf B Riemann-integrierbar. ■

Es ist nun klar, dass auch die Folgerungen 13.11 – 13.14 entsprechend für Integrale über Jordan-messbare Mengen gelten. Beispielsweise hat man:

Folgerung 13.23 *Stetige Funktionen auf kompakten Jordan-messbaren Mengen sind Riemann-integrierbar.*

Wir überlegen uns nun, wie das Integral bei fester Funktion f vom Integrationsbereich abhängt. Dazu vereinbaren wir:

$$\int_{\emptyset} f \, dx := 0.$$

Aus Satz 13.21 bzw. dem Lebesgueschen Integrabilitätskriterium folgt sofort: Sind A und B Jordan-messbar, so sind auch $A \cup B$, $A \cap B$ und $A \setminus B$ Jordan-messbar (die Ränder dieser Mengen liegen in $\partial A \cup \partial B$ und sind folglich Nullmengen). Weiter: Ist f auf einer Jordan-messbaren Menge B integrierbar, so ist f auch auf jeder Jordan-messbaren Teilmenge von B integrierbar.

Satz 13.24 *Seien $A, B \subseteq \mathbb{R}^n$ Jordan-messbar und f auf A und B Riemann-integrierbar. Dann gilt:*

$$\int_{A \cup B} f \, dx + \int_{A \cap B} f \, dx = \int_A f \, dx + \int_B f \, dx.$$

Beweis. Die Existenz aller Integrale folgt aus den Vorbemerkungen. Wir zeigen die Behauptung zuerst im Fall $A \cap B = \emptyset$. Dazu wählen wir ein Intervall $I \subseteq \mathbb{R}^n$ mit $A \cup B \subseteq I$. Dann ist

$$\begin{aligned} \int_A f \, dx + \int_B f \, dx &= \int_I f_A \, dx + \int_I f_B \, dx = \int_I (f_A + f_B) \, dx \\ &= \int_I f_{A \cup B} \, dx = \int_{A \cup B} f \, dx. \end{aligned} \quad (13.3)$$

Im Fall $A \cap B \neq \emptyset$ schreiben wir A , B und $A \cup B$ als Vereinigung paarweise disjunkter Mengen

$$A = (A \cap B) \cup (A \setminus B), \quad B = (A \cap B) \cup (B \setminus A), \quad A \cup B = (A \cap B) \cup (A \setminus B) \cup (B \setminus A)$$

und erhalten durch wiederholte Anwendung von (13.3)

$$\begin{aligned} \int_A f \, dx + \int_B f \, dx &= \int_{A \cap B} f \, dx + \int_{A \setminus B} f \, dx + \int_{B \setminus A} f \, dx + \int_{A \cap B} f \, dx \\ &= \int_{A \cup B} f \, dx + \int_{A \cap B} f \, dx. \quad \blacksquare \end{aligned}$$

Folgerung 13.25 *Für Jordan-messbare Mengen A , B gilt*

$$|A \cup B| + |A \cap B| = |A| + |B|.$$

Folgerung 13.26 Für Jordan-messbare Mengen A, B mit $A \subseteq B$ gilt $|A| \leq |B|$.

Beweis. Auf jedem Intervall I mit $B \subseteq I$ gilt $\chi_A \leq \chi_B$ und daher

$$|A| = \int_A dx = \int_I \chi_A dx \leq \int_I \chi_B dx = \int_B dx = |B|. \quad \blacksquare$$

Wir sagen, dass sich zwei Mengen A, B *nicht überlappen*, wenn sie nur Randpunkte gemeinsam haben, d.h. wenn $A \cap B \subseteq \partial A \cup \partial B$.

Satz 13.27 Seien A, B sich nicht überlappende Jordan-messbare Mengen, und sei f auf A und B Riemann-integrierbar. Dann ist

$$\int_{A \cup B} f dx = \int_A f dx + \int_B f dx.$$

Beweis. Nach Satz 13.24 genügt es zu zeigen, dass $\int_{A \cap B} f dx = 0$. Da außerdem

$$\left| \int_{A \cap B} f dx \right| \leq \|f\|_\infty |A \cap B|$$

ist, genügt es zu zeigen, dass $|A \cap B| = 0$. Da $A \cap B \subseteq \partial A \cup \partial B$, und da ∂A und ∂B kompakte Nullmengen sind, folgt diese Aussage aus der folgenden Behauptung:

Kompakte Nullmengen sind Jordan-messbar und besitzen den Jordan-Inhalt 0.

Wir beweisen diese Behauptung. Sei N eine kompakte Nullmenge. Da N Nullmenge ist, gibt es für jedes $\varepsilon > 0$ eine Überdeckung von N durch abzählbar viele offene Intervalle I_1, I_2, \dots mit $|I_1| + |I_2| + \dots < \varepsilon$. Da N kompakt ist, überdecken bereits endlich viele dieser Intervalle, etwa I_1, \dots, I_n , die Menge N , und es gilt $|I_1| + \dots + |I_n| < \varepsilon$. Da auch ∂N von diesen Intervallen überdeckt wird, ist ∂N Nullmenge und daher N Jordan-messbar. Für den Inhalt von N erhalten wir mit Folgerung 13.25 und Folgerung 13.26:

$$|N| \leq \left| \bigcup_{i=1}^n I_n \right| \leq \sum_{i=1}^n |I_n| < \varepsilon.$$

Da dies für jedes $\varepsilon > 0$ gilt, ist $|N| = 0$. ■

Folgerung 13.28 Für sich nicht überlappende Jordan-messbare Mengen A, B gilt $|A \cup B| = |A| + |B|$.

Man kann auch leicht die Umkehrung zur im Beweis von Satz 13.27 formulierten Behauptung beweisen: Jede Jordan-messbare Menge mit Inhalt 0 ist Nullmenge.

Folgerung 13.29 Eine beschränkte Menge $B \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn ihr Rand Jordan-messbar ist und den Inhalt 0 hat.

Beweis. Der Rand ∂B ist kompakt, und kompakte Mengen sind genau dann Nullmengen, wenn sie Jordan-messbar sind und den Inhalt 0 besitzen. ■

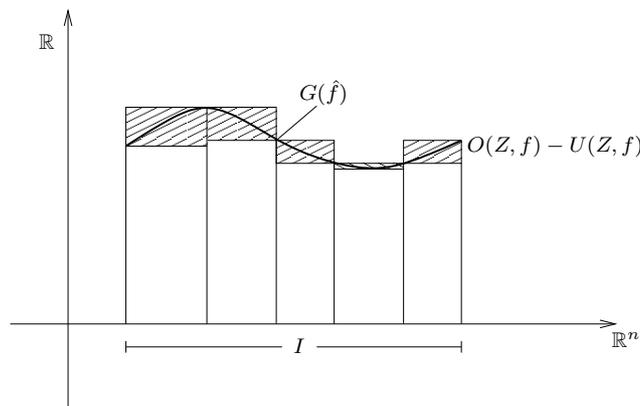
Wir haben nun schon viel über Jordan-messbare Mengen erfahren, wissen aber immer noch nicht, ob so einfache Mengen wie ein Kreis im \mathbb{R}^2 oder eine Kugel im \mathbb{R}^3 Jordan-messbar sind. Auf Grund von Folgerung 13.29 benötigen wir noch Kriterien dafür, dass eine beschränkte Menge Jordan-messbar ist und den Jordan-Inhalt 0 besitzt. Solche Mengen nennen wir auch *Jordansche Nullmengen*. Wir geben zwei solcher Kriterien an.

Satz 13.30 Sei $B \subseteq \mathbb{R}^n$ Jordan-messbar und $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar. Dann ist der Graph von f , d.h. die Menge

$$G(f) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in B, y = f(x)\},$$

eine Jordansche Nullmenge im \mathbb{R}^{n+1} .

Beweis. Sei $I \subseteq \mathbb{R}^n$ ein Intervall, welches B umfaßt, und sei \hat{f} die Einschränkung von f_B auf I . Nach dem Riemanschen Integrabilitätskriterium gibt es für jedes $\varepsilon > 0$ eine Zerlegung Z von I mit $O(Z, \hat{f}) - U(Z, \hat{f}) < \varepsilon$. Damit haben wir sofort eine Überdeckung von $G(\hat{f})$ durch endlich viele abgeschlossene Intervalle mit einer Inhaltssumme $< \varepsilon$. Wegen $G(f) \subseteq G(\hat{f})$ gilt dies erst recht für $G(f)$. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung. ■



Beispiel. Nun können wir auch zeigen, dass Kreise und Kugeln Jordan-messbar sind. Zunächst ist $[-1, 1] \subseteq \mathbb{R}^1$ Jordan-messbar (Intervall). Auf $[-1, 1]$ sind die Funktionen $f^+(x) := \sqrt{1-x^2}$ und $f^-(x) := -\sqrt{1-x^2}$ stetig und folglich Riemann-integrierbar. Nach Satz 13.30 sind die Graphen

$$G(f^\pm) = \{(x, y) \in \mathbb{R}^2 : x \in [-1, 1], y = \pm\sqrt{1-x^2}\}$$

Jordansche Nullmengen. Wegen $G(f^+) \cup G(f^-) = \{(x, y) : x^2 + y^2 = 1\}$ ist die Einheitskreislinie im \mathbb{R}^2 eine Jordansche Nullmenge. Folgerung 13.29 zeigt dann, dass die Einheitskreisscheibe $B_2 := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ Jordan-messbar ist.

Analog betrachten wir auf der (nun als Jordan-messbar erkannten) Menge B_2 die Funktionen

$$g^\pm(x, y) = \pm\sqrt{1 - x^2 - y^2} : B_2 \rightarrow \mathbb{R}$$

und erhalten wie oben, dass die Oberfläche der Einheitskugel im \mathbb{R}^3 eine Jordansche Nullmenge und damit die Einheitskugel selbst Jordan-messbar ist. Durch vollständige Induktion überträgt man dieses Resultat auf Kugeln im \mathbb{R}^n . ■

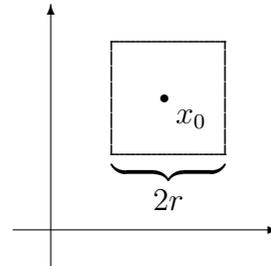
Satz 13.31 Sei $N \subseteq \mathbb{R}^n$ eine Jordansche Nullmenge und $f : N \rightarrow \mathbb{R}^m$ eine Lipschitz-stetige Funktion. Falls $m \geq n$, so ist $f(N)$ eine Jordansche Nullmenge in \mathbb{R}^m .

Im folgenden Beweis dieses Satzes ist es bequem, mit Würfeln statt mit Intervallen zu arbeiten. Ein Intervall $[a_1, b_1] \times \dots \times [a_n, b_n] \subseteq \mathbb{R}^n$ heißt *Würfel*, wenn $b_1 - a_1 = b_2 - a_2 = \dots = b_n - a_n$. Die Zahl $b_1 - a_1$ heißt die *Kantenlänge des Würfels*. Man kann Würfel im \mathbb{R}^n als Kugeln bezüglich der Maximumnorm

$$\|(x_1, \dots, x_n)^T\|_\infty := \max_{1 \leq i \leq n} |x_i|$$

betrachten: Für jeden Vektor $x_0 \in \mathbb{R}^n$ und jedes $r > 0$ ist

$$I := \{x \in \mathbb{R}^n : \|x - x_0\|_\infty \leq r\}$$



der Würfel mit Mittelpunkt x_0 und achsenparallelen Kanten der Länge $2r$. Man überlegt sich leicht, dass sich jede Jordansche Nullmenge im \mathbb{R}^n für jedes $\varepsilon > 0$ durch endlich viele Würfel I_1, \dots, I_n mit *gleicher* Kantenlänge und mit $\sum |I_i| < \varepsilon$ überdecken lässt (genauer: Heuser, Analysis II, S. 462).

Beweis von Satz 13.31. Da sich Würfel bequem mit der Maximumnorm beschreiben lassen, arbeiten wir sowohl in \mathbb{R}^n als auch in \mathbb{R}^m mit Norm $\|\cdot\|_\infty$.

Nach Voraussetzung gibt es ein $L > 0$, so dass

$$\|g(x) - g(y)\|_\infty \leq L\|x - y\|_\infty \quad \text{für alle } x, y \in N.$$

(Die Lipschitz-Stetigkeit einer Funktion ist unabhängig von der gewählten Norm. Die Lipschitz-Konstante L hängt dagegen sehr wohl von der Norm ab.) Für beliebig vorgegebenes $\varepsilon > 0$ überdecken wir N durch die k Würfel

$$I_i = \{x \in \mathbb{R}^n : \|x - \xi_i\|_\infty \leq r\}, \quad i = 1, \dots, k$$

mit den Mittelpunkten ξ_i und der Kantenlänge $2r < 1$ so, dass $\sum_{i=1}^k |I_i| = k(2r)^n < \varepsilon$. Da $N = \bigcup_{i=1}^k (N \cap I_i)$, ist $g(N) = \bigcup_{i=1}^k g(N \cap I_i)$ (↗ Übung).

Aus jeder Menge $N \cap I_i$ wählen wir einen Punkt η_i . Für jedes $x \in N \cap I_i$ ist dann

$$\|x - \eta_i\|_\infty \leq \|x - \xi_i\|_\infty + \|\xi_i - \eta_i\|_\infty \leq 2r,$$

woraus mit der Lipschitz-Stetigkeit von g folgt:

$$\|g(x) - g(\eta_i)\|_\infty \leq L\|x - \eta_i\|_\infty \leq 2rL.$$

Folglich ist $g(N \cap I_i)$ enthalten im Würfel mit Mittelpunkt $g(\eta_i)$ und mit der Kantenlänge $4rL$. Die Menge $g(N)$ kann also durch k Würfel überdeckt werden, für deren Inhaltssumme gilt:

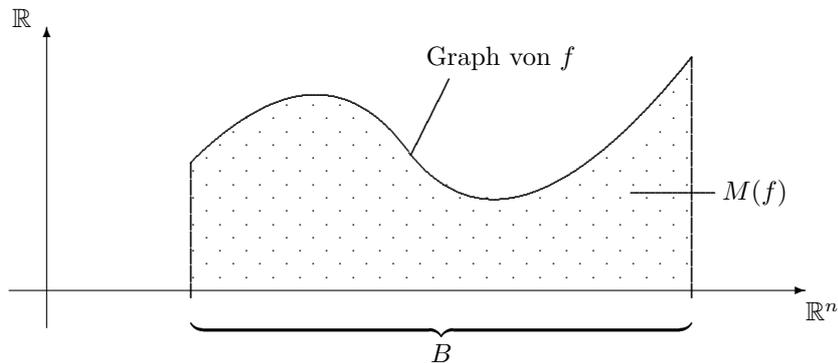
$$k(4rL)^m = k \cdot (2r)^n (2r)^{-n} (4rL)^m = L^m (2r)^{m-n} 2^m \cdot k(2r)^n < (2L)^m \varepsilon$$

(beachte: $2r < 1$, $m - n \geq 0$). Also ist $g(N)$ eine Jordansche Nullmenge. ■

13.5 Inhalt von Ordinatenmengen

Nachdem wir nun wissen, dass Kreise Jordan-messbar sind, möchten wir nun auch ihren Flächeninhalt berechnen. Allgemeiner geht es darum, Jordan-Inhalte so genannter Ordinatenmengen zu bestimmen. Die *Ordinatenmenge* $M(f)$ einer Funktion $f : \mathbb{R}^n \supseteq B \rightarrow \mathbb{R}^+$ ist die Menge

$$M(f) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in B, 0 \leq y \leq f(x)\}.$$



Im Fall $B = [a, b] \subseteq \mathbb{R}^1$ haben wir als Flächeninhalt von $M(f)$ definiert:

$$\text{Inhalt von } M(f) := \int_a^b f(x) dx.$$

Wir zeigen nun, dass der so definierte Flächeninhalt mit dem Jordanschen Inhalt von $M(f)$ übereinstimmt.

Satz 13.32 Sei $B \subseteq \mathbb{R}^n$ Jordan-messbar, $f : B \rightarrow \mathbb{R}$ Riemann-integrierbar und $f \geq 0$. Dann ist die Menge $M(f) \subseteq \mathbb{R}^{n+1}$ Jordan-messbar, und es gilt

$$|M(f)| = \int_B f dx.$$

Beweis. Wir beweisen zuerst die Jordan-Messbarkeit von $M(f)$. Dazu müssen wir zeigen, dass $\partial M(f)$ eine Nullmenge im \mathbb{R}^{n+1} ist.

Sei $S := \sup_{x \in B} f(x)$. Ein Punkt $(x, y) \in \mathbb{R}^n \times \mathbb{R}$ liegt sicher im Inneren von $M(f)$, wenn x im Inneren von B liegt, $0 < y < f(x)$ ist, und f in x stetig ist. Also ist $\partial M(f)$ sicher in der Vereinigung der folgenden Mengen enthalten:

$$\begin{aligned} A_1 &:= \{(x, 0) : x \in B\}, \\ A_2 &:= \{(x, f(x)) : x \in B\}, \\ A_3 &:= \{(x, y) : x \in \partial B, 0 \leq y \leq S\}, \\ A_4 &:= \{(x, y) : x \in B, f \text{ in } x \text{ unstetig}, 0 \leq y \leq S\}. \end{aligned}$$

Wir zeigen, dass jede dieser Mengen eine Nullmenge ist. A_1 ist Teil der Hyperebene $\{(x, 0) : x \in \mathbb{R}^n\}$ und hat daher den Jordan-Inhalt 0 (vgl. das Beispiel aus Abschnitt 13.2.2). A_2 ist der Graph von f und hat nach Satz 13.30 den Jordan-Inhalt 0. Für A_3 geben wir uns ein $\varepsilon > 0$ vor und überdecken ∂B durch n -dimensionale Intervalle I_1, I_2, \dots mit der Inhaltssumme $\sum |I_i| < \varepsilon/S$. Dann überdecken die $(n+1)$ -dimensionalen Intervalle $I_1 \times [0, S], I_2 \times [0, S], \dots$ die Menge A_3 , und für deren Inhaltssumme gilt: $\sum |I_i \times [0, S]| < \varepsilon/S \cdot S = \varepsilon$. Da die Menge der Unstetigkeitsstellen von f eine Nullmenge ist, können wir auf analoge Weise zeigen, dass auch A_4 Nullmenge ist. Damit ist die Jordan-Messbarkeit von $M(f)$ klar.

Für die Inhaltsformel sei I ein Intervall, welches B umfasst. Dann liegt $M(f)$ im Intervall $I \times [0, S]$, und es gilt nach Fubini:

$$\begin{aligned} |M(f)| &= \int_{I \times [0, S]} \chi_{M(f)} d(x, y) = \int_I \left(\int_0^S \chi_{M(f)}(x, y) dy \right) dx \\ &= \int_I \left(\int_0^{f_B(x)} 1 dy \right) dx = \int_I f_B(x) dx = \int_B f dx. \quad \blacksquare \end{aligned}$$

Offenbar gilt auch die folgende Verallgemeinerung von Satz 13.32.

Satz 13.33 Seien $B \subseteq \mathbb{R}^n$ Jordan-messbar und $f_1, f_2 : B \rightarrow \mathbb{R}$ Riemann-integrierbar, und sei $f_1(x) \leq f_2(x)$ für alle $x \in B$. Dann ist die Menge

$$M(f_1, f_2) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : x \in B, f_1(x) \leq y \leq f_2(x)\}$$

Jordan-messbar und

$$|M(f_1, f_2)| = \int_B (f_2 - f_1) dx.$$

Beispiel. Seien $(x_0, y_0) \in \mathbb{R}^2$, $r > 0$ sowie $B = [x_0 - r, x_0 + r]$, und sei

$$f_1(x) := y_0 - \sqrt{r^2 - (x - x_0)^2}, \quad f_2(x) := y_0 + \sqrt{r^2 - (x - x_0)^2}.$$

Dann ist $M(f_1, f_2)$ die Kreisscheibe mit Mittelpunkt (x_0, y_0) und Radius r . Für ihren Inhalt finden wir

$$\begin{aligned}
 |M(f_1, f_2)| &= 2 \int_{x_0-r}^{x_0+r} \sqrt{r^2 - (x - x_0)^2} dx = 2 \int_{-r}^r \sqrt{r^2 - t^2} dt \\
 &= 2 \left[\frac{1}{2} t + \sqrt{r^2 - t^2} + \frac{1}{2} r^2 \arcsin \frac{t}{r} \right] \Big|_{-r}^r = \pi r^2. \quad \blacksquare
 \end{aligned}$$

13.6 Integration über Normalbereiche

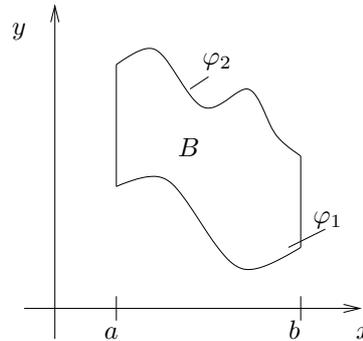
Unter einem *Normalbereich bzgl. der x-Achse* versteht man eine Menge $B \subseteq \mathbb{R}^2$ der Gestalt

$$B = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\}, \quad (13.4)$$

wobei φ_1, φ_2 stetige Funktionen auf $[a, b]$ mit $\varphi_1(x) \leq \varphi_2(x)$ für alle $x \in [a, b]$ sind.

Satz 13.34 Sei B wie in (13.4) und $f : B \rightarrow \mathbb{R}$ stetig. Dann ist

$$\int_B f d(x, y) = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy dx.$$



Beweis. Sei $I := [a, b] \times [m, M]$ ein Rechteck, welches B umfasst. Nach Fubini ist

$$\int_B f(x, y) d(x, y) = \int_a^b \int_m^M f_B(x, y) dy dx.$$

Für jedes feste $x \in [a, b]$ ist offenbar

$$\int_m^M f_B(x, y) dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy. \quad \blacksquare$$

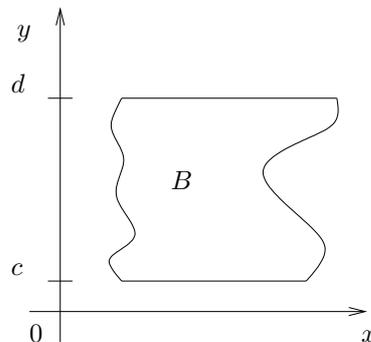
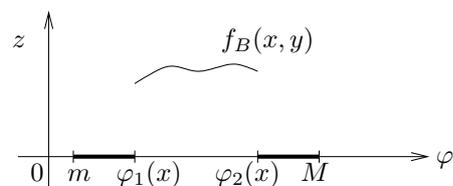
Analog heißt $B :=$

$$\{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}$$

ein *Normalbereich bezüglich der y-Achse*. Es gilt

$$\int_B f(x, y) d(x, y) = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx dy.$$

Ähnlich erklärt man Normalbereiche im \mathbb{R}^3 .



Sei $B_1 := [a, b]$. Auf B_1 sind stetige Funktionen φ_1, ψ_1 gegeben mit $\varphi_1(x) \leq \psi_1(x)$ für alle $x \in B_1$. Dann sei

$$B_2 := \{(x, y) : x \in B_1, \varphi_1(x) \leq y \leq \psi_1(x)\} \subseteq \mathbb{R}^2.$$

Auf B_2 sind stetige Funktionen φ_2, ψ_2 gegeben mit $\varphi_2(x, y) \leq \psi_2(x, y)$ für alle $(x, y) \in B_2$. Dann ist

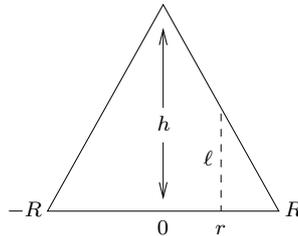
$$B_3 := \{(x, y, z) : (x, y) \in B_2, \varphi_2(x, y) \leq z \leq \psi_2(x, y)\} \subseteq \mathbb{R}^3$$

ein Normalbereich in \mathbb{R}^3 , und es gilt

$$\int_{B_3} f(x, y, z) d(x, y, z) = \int_a^b \int_{\varphi_1(x)}^{\psi_1(x)} \int_{\varphi_2(x, y)}^{\psi_2(x, y)} f(x, y, z) dz dy dx.$$

Beispiel. Ein Kreiskegel mit Radius R und Höhe h ist ein Normalbereich im \mathbb{R}^3 . Wir haben etwa

$$\begin{aligned} B_1 &= [-R, R], \\ B_2 &= \{(x, y) : x \in [-R, R], -\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2}\} \\ &\quad (= \text{Grundfläche des Kegels}), \\ B_3 &= \{(x, y, z) : (x, y) \in B_2, 0 \leq z \leq h - \frac{h}{R}\sqrt{x^2 + y^2}\}. \end{aligned}$$



(Beachte: (x, y) hat von $(0, 0)$ den Abstand $r := \sqrt{x^2 + y^2}$.) Die Länge ℓ der gestrichelten Strecke ist nach dem Strahlensatz

$$h : R = \ell : (R - r) \quad \text{bzw.} \quad \ell = \frac{hR - hr}{R} = h - \frac{h}{R} r.$$

Für das Volumen dieses Kegels finden wir daher

$$\begin{aligned} V &= \int_{B_3} d(x, y, z) = \int_{-R}^R \int_{-\sqrt{R^2 - x^2}}^{\sqrt{R^2 - x^2}} \int_0^{h - \frac{h}{R}\sqrt{x^2 + y^2}} dz dy dx \\ &= \int_{-R}^R \int_{-\sqrt{R^2 - x^2}}^{\sqrt{R^2 - x^2}} \left(h - \frac{h}{R}\sqrt{x^2 + y^2} \right) dy dx \\ &= \int_{-R}^R \left(hy - \frac{h}{R} \left(\frac{y}{2}\sqrt{x^2 + y^2} + \frac{x^2}{2} \ln(y + \sqrt{x^2 + y^2}) \right) \right) \Big|_{-\sqrt{R^2 - x^2}}^{\sqrt{R^2 - x^2}} dx \end{aligned}$$

$$\begin{aligned}
&= \int_{-R}^R \left(2h\sqrt{R^2 - x^2} - h\sqrt{R^2 - x^2} - \frac{x^2 h}{2R} \ln \frac{R + \sqrt{R^2 - x^2}}{R - \sqrt{R^2 - x^2}} \right) dx \\
&= 2h \int_0^R \left(\sqrt{R^2 - x^2} - \frac{x^2}{2R} \ln \frac{R + \sqrt{R^2 - x^2}}{R - \sqrt{R^2 - x^2}} \right) dx.
\end{aligned}$$

Wir substituieren $x = R \sin t$, $dx = R \cos t dt$ und erhalten

$$\begin{aligned}
V &= 2h \int_0^{\pi/2} \left(R \cos t - \frac{R}{2} \sin^2 t \ln \frac{1 + \cos t}{1 - \cos t} \right) R \cos t dt \\
&= 2hR^2 \int_0^{\pi/2} \left(\cos^2 t - \frac{1}{2} \sin^2 t \cos t \ln \frac{1 + \cos t}{1 - \cos t} \right) dt.
\end{aligned}$$

Das Integral $\sin^2 t \cos t \ln \frac{1+\cos t}{1-\cos t}$ kann durch partielle Integration bestimmt werden (der Faktor $\sin^2 t \cos t$ wird integriert und liefert $\frac{1}{3} \sin^3 t$; der Faktor $\ln \frac{1+\cos t}{1-\cos t}$ wird differenziert und ergibt $\frac{-2}{\sin t}$). Eingesetzt findet man schließlich

$$V = 2hR^2 \left(\frac{\pi}{4} - \frac{1}{6} \sin^3 \varphi \ln \frac{1 + \cos t}{1 - \cos t} \Big|_0^{\pi/2} - \frac{2}{6} \int_0^{\pi/2} \sin^2 \varphi d\varphi \right),$$

also

$$V = \frac{1}{3} \pi h R^2. \quad \blacksquare$$

13.7 Die Substitutionsregel

Nach dieser aufwändigen Rechnung für ein elementares Resultat fragt man sich, ob man nicht von vornherein die Rechnung hätte vereinfachen können durch eine andere Beschreibung des Kegels, etwa in Zylinderkoordinaten (die Substitutionsregel mussten wir ja ohnehin verwenden). Beschreiben wir die Grundfläche in Polarkoordinaten, so wird der Kegel offenbar beschrieben durch

$$\{(r, \varphi, z) : r \in [0, R], \varphi \in [0, 2\pi], z \in [0, h - hr/R]\},$$

was eine wesentlich einfachere Integration erwarten lässt. *Frage:* Wie haben wir im Integral $\int_B f(x, y, z) d(x, y, z)$ den Ausdruck $d(x, y, z)$ zu transformieren, wenn wir von (x, y, z) zu neuen Koordinaten, etwa r, φ und z , übergehen?

A Motivation. Zu berechnen ist das Integral

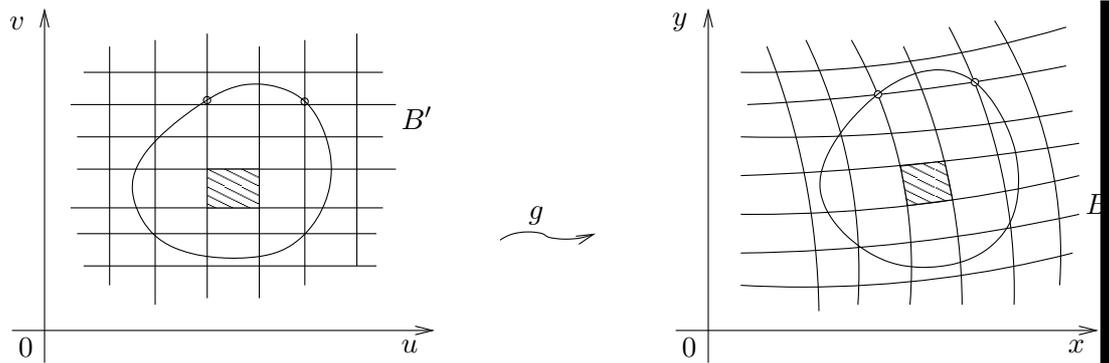
$$\int_B f(x, y) d(x, y)$$

über einem Bereich $B \subseteq \mathbb{R}^2$, versehen mit x, y -Koordinaten. Die Substitution $x := \varphi(u, v)$, $y := \psi(u, v)$ führt neue Veränderliche ein. Durch diese Substitution

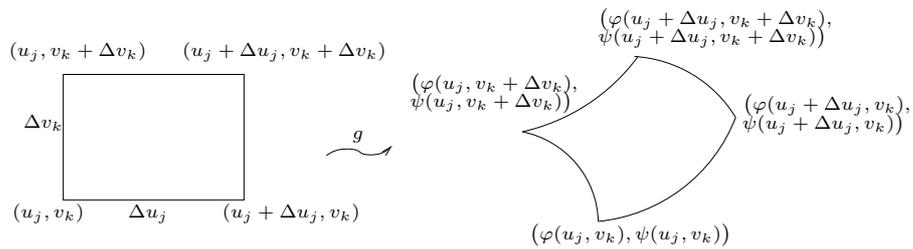
werde ein Bereich $B' \subseteq \mathbb{R}^2$ mit Koordinaten u, v (wir sagen auch: ein Bereich der uv -Ebene) injektiv auf B abgebildet; genauer: die Abbildung

$$g : B' \rightarrow B, \quad (u, v) \mapsto g(u, v) = (\varphi(u, v), \psi(u, v))$$

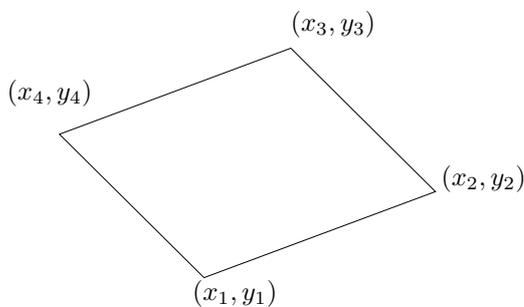
ist eine Bijektion von B' auf B . Diese Abbildung übersetzt ein Rechtecknetz über B' in ein „krummliniges Netz“ über B :



Wir sehen uns genauer an, wie das schraffierte Rechteck in der uv -Ebene auf das „krummlinige“ schraffierte Parallelogramm in der xy -Ebene abgebildet wird:

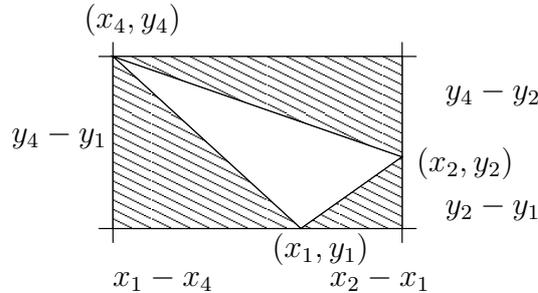


Der Flächeninhalt dieses Rechtecks (vgl. die Definition des Riemann-Integrals über B') beträgt $\Delta u_j \Delta v_k$. Um den Flächeninhalt des „krummlinigen“ Parallelogramms zu berechnen, nehmen wir an, dass Δu_j und Δv_k so klein sind, dass das „krumme“ Parallelogramm fast ein echtes Parallelogramm ist. Für den Flächeninhalt eines Parallelogramms mit den Eckpunkten $(x_1, y_1), \dots, (x_4, y_4)$ gilt



$$\text{Fläche} = |(x_2 - x_1)(y_4 - y_1) - (x_4 - x_1)(y_2 - y_1)|. \quad (13.5)$$

Dies läßt sich leicht ableiten, indem man z.B. von der unten gezeichneten Rechteckfläche die schraffierten Dreiecksflächen subtrahiert:



$$\begin{aligned} \frac{1}{2} \text{ Parallelogrammfläche} &= \text{Fläche des weißen Dreiecks} \\ &= (y_4 - y_1)((x_2 - x_1) + (x_1 - x_4)) - \frac{1}{2}(y_2 - y_1)(x_2 - x_1) \\ &\quad - \frac{1}{2}(x_1 - x_4)(y_4 - y_1) - \frac{1}{2}(y_4 - y_2)(x_2 - x_4). \end{aligned}$$

Mit Hilfe von Determinanten läßt sich (13.5) schreiben als

$$\text{Fläche} = \left| \det \begin{pmatrix} x_2 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_4 - y_1 \end{pmatrix} \right|.$$

Die Fläche des „krummen“ Parallelogramms ist also ungefähr gleich

$$\left| \det \begin{pmatrix} \varphi(u_j + \Delta u_j, v_k) - \varphi(u_j, v_k) & \varphi(u_j, v_k + \Delta v_k) - \varphi(u_j, v_k) \\ \psi(y_j + \Delta u_j, v_k) - \psi(u_j, v_k) & \psi(u_j, v_k + \Delta v_k) - \psi(u_j, v_k) \end{pmatrix} \right|.$$

Wir nehmen nun an, dass φ und ψ differenzierbar sind. Für kleines Δu_j ist dann

$$\varphi(u_j + \Delta u_j, v_k) - \varphi(u_j, v_k) \approx \frac{\partial \varphi}{\partial u}(u_j, v_k) \cdot \Delta u_j.$$

Der Ausdruck (13.4) ist daher ungefähr gleich

$$\left| \det \begin{pmatrix} \frac{\partial \varphi}{\partial u}(u_j, v_k) & \frac{\partial \varphi}{\partial v}(u_j, v_k) \\ \frac{\partial \psi}{\partial u}(u_j, v_k) & \frac{\partial \psi}{\partial v}(u_j, v_k) \end{pmatrix} \right| \Delta u_j \Delta v_k.$$

Diese Matrix ist aber nichts anderes als die Jacobi-Matrix von g an der Stelle (u_j, v_k) . Mit anderen Worten: (13.4) ist etwa gleich

$$|\det g'(u_j, v_k)| \cdot \Delta u_j \Delta v_k.$$

Wir erwarten daher näherungsweise

$$\int_B f(x, y) d(x, y) \approx \sum_{j,k} f(\varphi(u_j, v_k), \psi(u_j, v_k)) |\det g'(u_j, v_k)| \Delta u_j \Delta v_k$$

und somit die Substitutionsregel

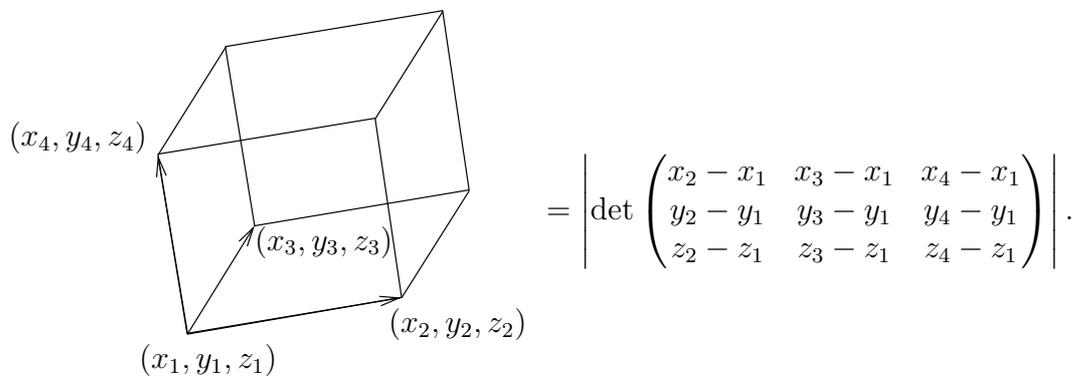
$$\int_B f(x, y) d(x, y) = \int_{B'} f(\varphi(u, v), \psi(u, v)) |\det g'(u, v)| d(u, v).$$

Es zeigt sich, dass die hier „abgeleitete“ Formel unter entsprechenden Voraussetzungen tatsächlich gilt und dass sie auf Funktionen mehrerer Veränderlicher verallgemeinert werden kann.

B Determinanten und Volumina von Parallelepipeden. Mit Hilfe der Determinante

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} := a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

kann man das Volumen eines Parallelepipeds im \mathbb{R}^3 beschreiben:



$$= \left| \det \begin{pmatrix} x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_3 - y_1 & y_4 - y_1 \\ z_2 - z_1 & z_3 - z_1 & z_4 - z_1 \end{pmatrix} \right|.$$

Zur Erinnerung: Sei $A = (a_{ij})_{i,j=1}^n$ eine $n \times n$ -Matrix, und A_{ij} sei die $(n-1) \times (n-1)$ -Matrix, die aus A durch Streichen der i -ten Zeile und der j -ten Spalte entsteht. Dann definiert man:

$$\det A := a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} + \dots + (-1)^{n-1} a_{1n} \det A_{1n}.$$

Die Determinante einer $n \times n$ -Matrix steht in engem Zusammenhang zum Volumen eines Parallelepipeds im \mathbb{R}^n .

C Der allgemeine Substitutionssatz.

Satz 13.35 Sei $G \subseteq \mathbb{R}^n$ offen, und $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ sei injektiv und stetig differenzierbar. Die Determinante $\det g'(t)$ sei auf G entweder überall positiv oder überall negativ. Weiter sei T eine kompakte und Jordan-messbare Teilmenge von G , und f sei eine auf $g(T)$ stetige reellwertige Funktion. Dann ist $g(T)$ kompakt und Jordan-messbar, f ist auf $g(T)$ Riemann-integrierbar, und es gilt

$$\int_{g(T)} f(x) dx = \int_T f(g(t)) |\det g'(t)| dt. \quad (13.6)$$

Die Formel (13.6) gilt auch dann noch, wenn – entgegen den obigen Voraussetzungen – die Determinante $\det g'(t)$ auf einer Teilmenge N von T verschwindet oder wenn $g|_N$ auf einer Teilmenge N von T nicht injektiv ist, sofern N den Jordan-Inhalt 0 hat.

Der Beweis kann z.B. mit vollständiger Induktion nach n erfolgen, ist aber recht aufwändig (vgl. Heuser, S. 475–485). Wir werden in der Vorlesung zur „Maß- und Integrationstheorie“ einen Beweis kennenlernen.

D Beispiele: Transformation auf Polar-, Zylinder- und Kugelkoordinaten. Der Zusammenhang zwischen den kartesischen Koordinaten (x, y) und den Polarkoordinaten (r, φ) eines Punktes im \mathbb{R}^2 ist gegeben durch

$$x = r \cos \varphi, \quad y = r \sin \varphi.$$

Ein Integral gemäß Satz 13.35 auf Polarkoordinaten zu transformieren heißt, die Substitution

$$\begin{pmatrix} x \\ y \end{pmatrix} = g(r, \varphi) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}$$

vorzunehmen. Es ist

$$g'(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

und

$$\det g'(r, \varphi) = \cos \varphi \cdot r \cos \varphi - (-r \sin \varphi) \sin \varphi = r(\cos^2 \varphi + \sin^2 \varphi) = r.$$

Für $r > 0$ ist also $\det g'(r, \varphi)$ stets positiv. Weiter ist klar, dass g den Bereich

$$\{(r, \varphi) : r > 0, 0 \leq \varphi < 2\pi\}$$

injektiv auf $\mathbb{R}^2 \setminus \{0\}$ abbildet und dass g insbesondere auf dem Gebiet

$$G := \{(r, \varphi) : r > 0, 0 < \varphi < 2\pi\}$$

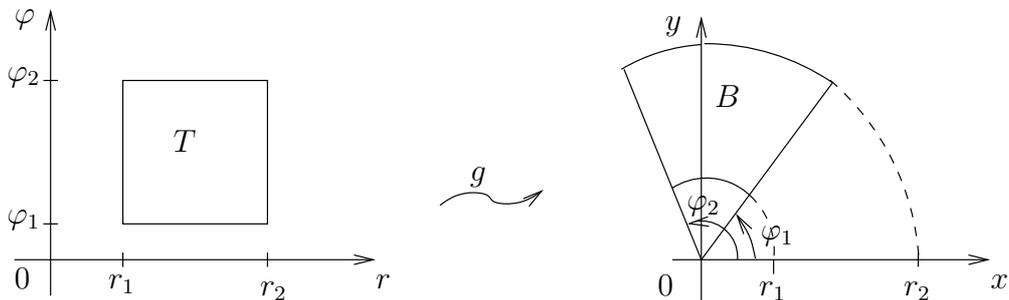
injektiv ist. Auf diesem Gebiet können wir somit Satz 13.35 anwenden und erhalten:

Folgerung 13.36 Ist $B = g(T)$, wobei $T \subseteq G$ kompakt und Jordan-messbar ist, so kann $\int_B f(x, y) d(x, y)$ auf Polarkoordinaten transformiert werden. Dabei gilt

$$\int_B f(x, y) d(x, y) = \int_T f(r \cos \varphi, r \sin \varphi) r d(r, \varphi). \quad (13.7)$$

Ist beispielsweise T ein achsenparalleles Rechteck, so ist $B = g(T)$ ein Kreisring, und wir haben nach Fubini

$$\int_B f(x, y) d(x, y) = \int_{\varphi_1}^{\varphi_2} \int_{r_1}^{r_2} f(r \cos \varphi, r \sin \varphi) r dr d\varphi. \quad (13.8)$$



In der Praxis ist T häufig ein Rechteck, welches im Streifen

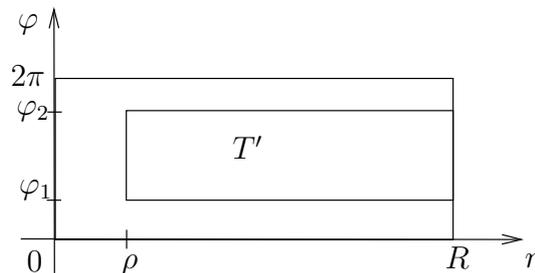
$$\{(r, \varphi) : r \geq 0, 0 \leq \varphi \leq 2\pi\}$$

liegt und Teile des Randes dieses Streifens enthält. Wir überlegen uns, dass die Formel (13.8) auch in diesem Fall noch gilt. Dazu sei

$$T = \{(r, \varphi) : 0 \leq r \leq R, 0 \leq \varphi \leq 2\pi\},$$

und $T' \subseteq T$ sei ein Rechteck der Gestalt

$$T' = \{(r, \varphi) : \varrho \leq r \leq R, \varphi_1 \leq \varphi \leq \varphi_2\} \quad \text{mit } \varrho > 0 \text{ und } 0 < \varphi_1 < \varphi_2 < 2\pi.$$



Auf T' gilt (13.8), d.h. es ist

$$\int_{g(T')} f(x, y) d(x, y) = \int_{T'} f(r \cos \varphi, r \sin \varphi) r d(r, \varphi),$$

und aus

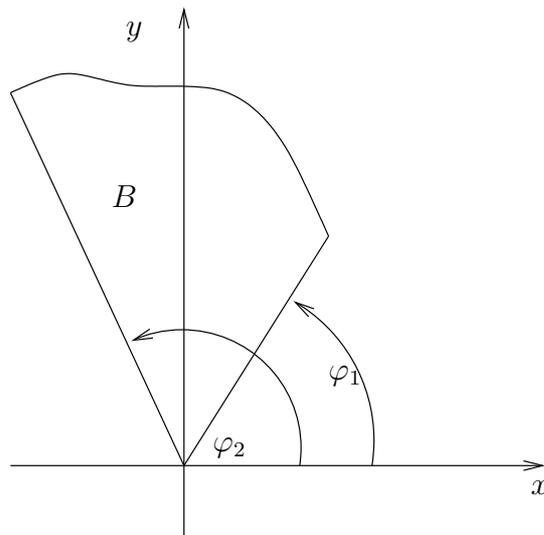
$$\begin{aligned} & \left| \int_T f(r \cos \varphi, r \sin \varphi) r d(r, \varphi) - \int_{T'} f(r \cos \varphi, r \sin \varphi) r d(r, \varphi) \right| \\ &= \left| \int_{T \setminus T'} f(r \cos \varphi, r \sin \varphi) r d(r, \varphi) \right| \\ &\leq \sup_{r, \varphi} |f(r \cos \varphi, r \sin \varphi) r| \cdot |T \setminus T'| \end{aligned}$$

folgt, dass $\int_{T'} f(r \cos \varphi, r \sin \varphi) r d(r, \varphi)$ gegen $\int_T f(r \cos \varphi, r \sin \varphi) r d(r, \varphi)$ strebt, wenn $\varrho \rightarrow 0$, $\varphi_1 \rightarrow 0$ und $\varphi_2 \rightarrow 2\pi$.

Ähnlich erhält man, dass auch $\int_{g(T')} f(x, y) d(x, y)$ gegen $\int_{g(T)} f(x, y) d(x, y)$ strebt. Zusammengefasst:

Folgerung 13.37 *Ist T ein Rechteck $[r_1, r_2] \times [\varphi_1, \varphi_2]$ mit $0 \leq r_1 < r_2$ und $0 \leq \varphi_1 < \varphi_2 \leq 2\pi$, und ist f auf $g(T) = B$ stetig, so gilt (13.8).*

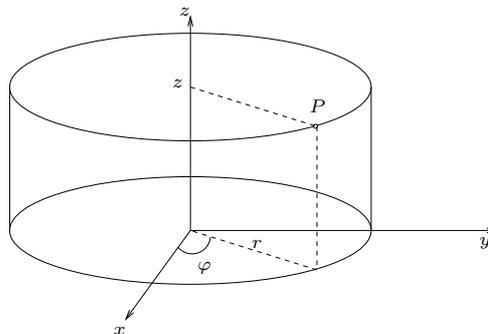
Anwendung: Sei $B \subseteq \mathbb{R}^2$ die Menge aller Punkte (r, φ) mit $0 \leq r \leq \varrho(\varphi)$ und $\varphi_1 \leq \varphi \leq \varphi_2$, wobei ϱ auf $[\varphi_1, \varphi_2]$ stetig und positiv sei. Dann ist der Inhalt von B gleich



$$|B| = \int_B 1 d(x, y) = \int_{\varphi_1}^{\varphi_2} \int_0^{\varrho(\varphi)} r dr d\varphi = \int_{\varphi_1}^{\varphi_2} \frac{r(\varphi)^2}{2} d\varphi. \quad \blacksquare$$

Zylinderkoordinaten. Die Zylinderkoordinaten (r, φ, z) eines Punktes $P \in \mathbb{R}^3$ mit kartesischen Koordinaten (x, y, z) sind

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z.$$



Für die Substitution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(r, \varphi, z) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix}$$

gilt

$$\det g'(r, \varphi, z) = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi & 0 \\ \sin \varphi & r \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = r.$$

Ähnlich wie bei Polarkoordinaten sieht man, dass die Substitutionsformel

$$\int_B f(x, y, z) d(x, y, z) = \int_T f(r \cos \varphi, r \sin \varphi, z) r d(r, \varphi, z) \quad (13.9)$$

sicher immer dann gilt, wenn T eine kompakte und Jordan-messbare Teilmenge des Gebietes

$$G = \{(r, \varphi, z) : r > 0, 0 < \varphi < 2\pi, -\infty < z < \infty\}$$

und $B = g(T)$ ist. Die Formel (13.9) gilt aber auch dann noch, wenn T ein Quader

$$T = [r_1, r_2] \times [\varphi_1, \varphi_2] \times [z_1, z_2] \quad \text{mit } 0 \leq r_1 < r_2, 0 \leq \varphi_1 < \varphi_2 \leq 2\pi$$

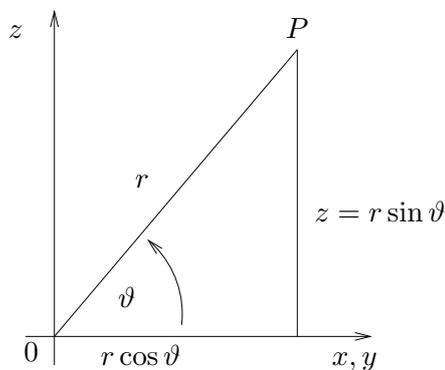
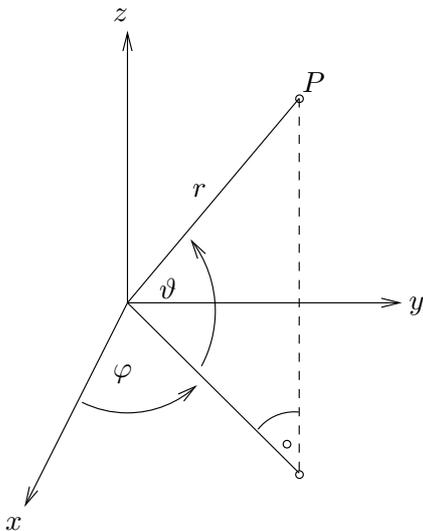
ist. In diesem speziellen Fall geht (13.9) über in

$$\int_B f(x, y, z) d(x, y, z) = \int_{z_1}^{z_2} \int_{\varphi_1}^{\varphi_2} \int_{r_1}^{r_2} f(r \cos \varphi, r \sin \varphi, z) r dr d\varphi dz.$$

Kugelkoordinaten. Der Zusammenhang zwischen Kugelkoordinaten (r, ϑ, φ) und kartesischen Koordinaten (x, y, z) eines Punktes $P \in \mathbb{R}^3$ wird hergestellt durch

$$x = r \cos \vartheta \cos \varphi, \quad y = r \cos \vartheta \sin \varphi, \quad z = r \sin \vartheta;$$

dabei ist $r \geq 0$, $-\frac{\pi}{2} \leq \vartheta \leq \frac{\pi}{2}$ und $0 \leq \varphi < 2\pi$.



Für die Substitution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(r, \vartheta, \varphi) = \begin{pmatrix} r \cos \vartheta \cos \varphi \\ r \cos \vartheta \sin \varphi \\ r \sin \vartheta \end{pmatrix}$$

findet man

$$\det g'(r, \vartheta, \varphi) = \det \begin{pmatrix} \cos \vartheta \cos \varphi & -r \sin \vartheta \cos \varphi & -r \cos \vartheta \sin \varphi \\ \cos \vartheta \sin \varphi & -r \sin \vartheta \sin \varphi & r \cos \vartheta \cos \varphi \\ \sin \vartheta & r \cos \vartheta & 0 \end{pmatrix} = -r^2 \cos \vartheta.$$

Diese Determinante ist negativ, wenn $r > 0$ und $-\pi/2 < \vartheta < \pi/2$. Wie oben gilt für stetiges f die Transformationsformel

$$\int_B f(x, y, z) d(x, y, z) = \int_T f(r \cos \vartheta \cos \varphi, r \cos \vartheta \sin \varphi, r \sin \vartheta) r^2 \cos \vartheta d(r, \vartheta, \varphi),$$

wenn $B = g(T)$ und T eine kompakte und Jordan-messbare Teilmenge von

$$G = \{(r, \vartheta, \varphi) : r > 0, -\frac{\pi}{2} < \vartheta < \frac{\pi}{2}, 0 < \varphi < 2\pi\}$$

ist. Diese Formel gilt auch dann noch, wenn

$$T = [r_1, r_2] \times [\vartheta_1, \vartheta_2] \times [\varphi_1, \varphi_2]$$

mit $0 \leq r_1 < r_2$, $-\frac{\pi}{2} \leq \vartheta_1 < \vartheta_2 \leq \frac{\pi}{2}$ und $0 \leq \varphi_1 < \varphi_2 \leq 2\pi$ ist. Insbesondere ist dann nach Fubini

$$\begin{aligned} \int_B f(x, y, z) d(x, y, z) \\ = \int_{\varphi_1}^{\varphi_2} \int_{\vartheta_1}^{\vartheta_2} \int_{r_1}^{r_2} f(r \cos \vartheta \cos \varphi, r \cos \vartheta \sin \varphi, r \sin \vartheta) r^2 \cos \vartheta dr d\vartheta d\varphi. \end{aligned}$$

Beispiel: Volumen einer Kugel. Die Kugel B mit Mittelpunkt $(0, 0, 0)$ und Radius R ist das Bild des Quaders

$$T = [0, R] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times [0, 2\pi]$$

unter der oben beschriebenen Transformation. Wir finden daher

$$\begin{aligned} V = \int_B d(x, y, z) &= \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \int_0^R r^2 \cos \vartheta dr d\vartheta d\varphi \\ &= \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \frac{R^3}{3} \cos \vartheta d\vartheta d\varphi \\ &= \int_0^{2\pi} \frac{R^3}{3} \sin \vartheta \Big|_{-\pi/2}^{\pi/2} d\varphi = \frac{2}{3} R^3 \int_0^{2\pi} d\varphi = \frac{4}{3} \pi R^3. \quad \blacksquare \end{aligned}$$

Man vergleiche diese Rechnung mit der mühsamen Herleitung der Formel für das Kegelvolumen in Abschnitt 13.6.

14 Oberflächenintegrale und Integralsätze

Nachdem wir im vergangenen Kapitel gesehen haben, wie man das Volumen eines dreidimensionalen Körpers (z.B. das Volumen einer Kugel) mit Hilfe der Integralrechnung bestimmen kann, wenden wir uns nun den Flächeninhalten gekrümmter Flächen zu (wie z.B. der Oberfläche einer Kugel), für die wir einen geeigneten Integralbegriff entwickeln.

14.1 Flächen, Tangenten und Normalen

Wir haben früher Kurven als Bilder von Intervallen bzgl. stetiger Abbildungen beschrieben. Ganz analog definieren wir nun Flächenstücke als Bilder ebener Bereiche bzgl. geeigneter Abbildungen.

Definition 14.1 Sei $D \subseteq \mathbb{R}^2$ eine beschränkte offene Menge, und ihre Abschließung \overline{D} sei Jordan-messbar. Weiter sei $F : \overline{D} \rightarrow \mathbb{R}^3$ eine stetig differenzierbare Funktion mit

$$\text{rang } F'(x_1, x_2) = 2 \quad \text{für alle } (x_1, x_2) \in D. \quad (14.1)$$

Dann heißt das Bild von \overline{D} unter F , d.h. die Menge

$$\mathcal{F} := \{F(x_1, x_2) \in \mathbb{R}^3 : (x_1, x_2) \in \overline{D}\} \quad (14.2)$$

ein Flächenstück im \mathbb{R}^3 und die Abbildung $F : \overline{D} \rightarrow \mathcal{F}$ heißt eine Parameterdarstellung des Flächenstücks \mathcal{F} oder kurz eine Fläche.

Genau wie bei Wegen und Kurven unterscheiden wir sorgfältig zwischen der Abbildung F und ihrem Bild \mathcal{F} . Offenbar kann es für ein- und dasselbe Flächenstück \mathcal{F} verschiedene Parametrisierungen geben.

Wichtige Vereinbarung. Wir haben früher die Differenzierbarkeit einer Abbildung $F : X \rightarrow \mathbb{R}^3$ nur für offene Mengen $X \subseteq \mathbb{R}^2$ erklärt. Die stetige Differenzierbarkeit von $F : \overline{D} \rightarrow \mathbb{R}^3$ haben wir wie folgt zu verstehen: Die Funktion F läßt sich zu einer stetig differenzierbaren Funktion $F : G \rightarrow \mathbb{R}^3$ auf eine offene Menge $G \supset \overline{D}$ fortsetzen. Diese Menge G ist für die Definition des Flächenstücks offenbar unerheblich. Wir möchten jedoch auch in den Randpunkten von D die partiellen Ableitungen F_{x_1} und F_{x_2} bilden können und verlangen daher, dass sich F auf eine offene Umgebung G von \overline{D} stetig differenzierbar fortsetzen läßt. ■

Die Rangbedingung (14.1) wird nur für Punkte aus D gefordert; auf dem Rand $\partial D = \overline{D} \setminus D$ muss sie nicht erfüllt sein. Mit

$$F(x_1, x_2) = (F_1(x_1, x_2), F_2(x_1, x_2), F_3(x_1, x_2))^T$$

lautet die Rangbedingung (14.1)

$$\text{rang} \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} \\ \frac{\partial F_3}{\partial x_1} & \frac{\partial F_3}{\partial x_2} \end{pmatrix} = 2 \quad (14.3)$$

in allen Punkten von D .

Beispiel 1: Kugeloberfläche. Sei $D = (0, 2\pi) \times (-\pi/2, \pi/2)$ und $r > 0$. Dann ist \bar{D} das Rechteck $[0, 2\pi] \times [-\pi/2, \pi/2]$. Sei weiter

$$F(u, v) = (r \cos u \cos v, r \sin u \cos v, r \sin v)$$

für $(u, v) \in \bar{D}$. Dann ist F stetig differenzierbar auf \bar{D} (und sogar auf ganz \mathbb{R}^2), und die Funktionalmatrix

$$F'(u, v) = \begin{pmatrix} -r \sin u \cos v & -r \cos u \sin v \\ r \cos u \cos v & -r \sin u \sin v \\ 0 & r \cos v \end{pmatrix}$$

hat für alle $(u, v) \in D$ den Rang 2. Das zugehörige Flächenstück \mathcal{F} ist die Oberfläche einer Kugel um den Nullpunkt mit dem Radius r . ■

Beispiel 2: Funktionsgraphen. Sei $f : \bar{D} \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und

$$F : \bar{D} \rightarrow \mathbb{R}^3, \quad F(u, v) := (u, v, f(u, v))^T.$$

Dann ist auch F stetig differenzierbar auf \bar{D} , und die Funktionalmatrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix}$$

von F hat offenbar den Rang 2. Das durch F definierte Flächenstück ist gerade der Graph von f . Mit $D = \{(u, v) \in \mathbb{R}^2 : u^2 + v^2 \leq 1\}$ und $f(u, v) = uv$ erhält man beispielsweise eine Sattelfläche (hyperbolisches Paraboloid). ■

Beispiel 3: Ebenen. Seien $a = (a_1, a_2, a_3)^T$ und $b = (b_1, b_2, b_3)^T$ linear unabhängige Vektoren und $x_0 \in \mathbb{R}^3$. Auf $\bar{D} = [-1, 1] \times [-1, 1]$ hat die Funktion $F(u, v) = ua + vb + x_0$ die Funktionalmatrix

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{pmatrix},$$

die wegen der linearen Unabhängigkeit von a und b den Rang 2 hat. Das zugehörige Flächenstück ist ein Teil der durch a und b aufgespannten und durch x_0

verlaufenden Ebene. ■

Sei $F : \overline{D} \rightarrow \mathbb{R}^3$ die Parametrisierung eines Flächenstücks \mathcal{F} . Ist durch $X = (X_1, X_2)^T : [a, b] \rightarrow \overline{D}$ ein stetig differenzierbarer Weg in \overline{D} gegeben, so ist

$$Y : [a, b] \rightarrow \mathbb{R}^3, \quad Y(t) = F(X(t))$$

ein stetig differenzierbarer Weg, der komplett in \mathcal{F} verläuft. Für $t_0 \in [a, b]$ wird die Richtung der Tangente an den Weg Y im Punkt $Y(t_0) = F(X(t_0))$ beschrieben durch den Vektor

$$\dot{Y}(t_0) = F_u(X(t_0))\dot{X}_1(t_0) + F_v(X(t_0))\dot{X}_2(t_0) \quad (14.4)$$

(Kettenregel!) mit

$$F_u(X(t_0)) = \frac{\partial F}{\partial u}(X(t_0)) = \left(\frac{\partial F_1}{\partial u}(X(t_0)), \frac{\partial F_2}{\partial u}(X(t_0)), \frac{\partial F_3}{\partial u}(X(t_0)) \right)$$

und

$$F_v(X(t_0)) = \frac{\partial F}{\partial v}(X(t_0)) = \left(\frac{\partial F_1}{\partial v}(X(t_0)), \frac{\partial F_2}{\partial v}(X(t_0)), \frac{\partial F_3}{\partial v}(X(t_0)) \right).$$

Die Rangbedingung (14.1) ist äquivalent zur linearen Unabhängigkeit der beiden Vektoren $F_u(X(t_0))$ und $F_v(X(t_0))$ für alle $X(t_0) \in D$. Wir schreiben $X(t_0) = (u_0, v_0)$. Betrachten wir alle Wege X durch den Punkt (u_0, v_0) , so können in (14.4) die Ableitungen $\dot{X}_1(t_0)$ und $\dot{X}_2(t_0)$ beliebige Werte annehmen. Die Vektoren $F_u(u_0, v_0)$ und $F_v(u_0, v_0)$ spannen also die komplette *Tangentialebene an die Fläche F im Punkt $F(u_0, v_0)$* auf. Eine Beschreibung der Tangentialebene in Parameterform lautet daher

$$\{F(u_0, v_0) + \lambda F_u(u_0, v_0) + \mu F_v(u_0, v_0) : \lambda, \mu \in \mathbb{R}\}. \quad (14.5)$$

Den *Normalenvektor an die Fläche $F : \overline{D} \rightarrow \mathbb{R}^3$ im Punkt $F(u, v)$* mit $(u, v) \in D$ erklären wir durch das Vektorprodukt

$$N(u, v) := \frac{F_u(u, v) \times F_v(u, v)}{\|F_u(u, v) \times F_v(u, v)\|}. \quad (14.6)$$

(Man beachte, dass $F_u(u, v) \times F_v(u, v) \neq 0$, da beide Vektoren linear unabhängig sind.) Aus den Eigenschaften des Vektorprodukts wissen wir, dass der Vektor $N(u, v)$ senkrecht auf $F_u(u, v)$ und $F_v(u, v)$ und damit auf der gesamten Tangentialebene (14.5) steht. Außerdem hat er die Länge 1, so dass man auch vom *Normaleneinheitsvektor* spricht.

Es stellt sich die Frage, inwieweit die eingeführten Begriffe (Tangential- und Normalenvektoren) von der Parametrisierung F oder nur vom Flächenstück \mathcal{F} abhängen. Dazu zunächst eine Definition.

Definition 14.2 (a) Seien D, E offen in \mathbb{R}^2 . Eine bijektive Abbildung $\varphi : \overline{D} \rightarrow \overline{E}$ heißt ein Diffeomorphismus, wenn φ und die Umkehrabbildung φ^{-1} stetig differenzierbar sind.

(b) Zwei Parameterdarstellungen $F : \overline{D} \rightarrow \mathbb{R}^3$ und $G : \overline{E} \rightarrow \mathbb{R}^3$ heißen äquivalent, wenn es einen Diffeomorphismus $\varphi : \overline{D} \rightarrow \overline{E}$ gibt mit $F = G \circ \varphi$ und $\det \varphi' > 0$ auf \overline{D} .

Sind F und G äquivalent, so sagt man auch, dass sie durch eine Parametertransformation φ auseinander hervorgehen. Es gilt nun:

Bei äquivalenten Parametertransformationen bleiben die wesentlichen Flächengrößen und Flächeneigenschaften (wie Tangentialebenen und Normalenvektoren) unverändert.

Anschaulich ist klar, dass jedes Flächenstück in jedem Punkt genau einen Tangentialraum, aber zwei Normaleneinheitsvektoren (nach „oben“ und nach „unten“) besitzt. Ist $F = G \circ \varphi$, so liefern F und G den gleichen Normalenvektor, wenn $\det \varphi' > 0$, und sie liefern entgegengesetzte Normalenvektoren, wenn $\det \varphi' < 0$. In letzterem Fall sagt man auch, dass die *Orientierung* von F gewechselt wird.

Beispiel 4. Seien D und F wie in Beispiel 1 (Kugeloberfläche). Dann ist

$$\begin{aligned} F_u(u, v) &= (-r \sin u \cos v, r \cos u \cos v, 0)^T, \\ F_v(u, v) &= (-r \cos u \sin v, -r \sin u \sin v, r \cos v)^T, \\ F_u(u, v) \times F_v(u, v) &= (r^2 \cos u \cos^2 v, r^2 \sin u \cos^2 v, r^2 \sin v \cos v), \\ \|F_u(u, v) \times F_v(u, v)\| &= r^2 \cos v \end{aligned}$$

und damit

$$N(u, v) = (\cos u \cos v, \sin u \cos v, \sin v)$$

für alle $(u, v) \in (0, 2\pi) \times (-\pi/2, \pi/2)$. ■

Beispiel 5. Sind D , f und F wie in Beispiel 2 (Funktionsgraphen), so ist

$$\begin{aligned} F_u &= \left(1, 0, \frac{\partial f}{\partial u}\right)^T, & F_v &= \left(0, 1, \frac{\partial f}{\partial v}\right)^T, \\ F_u \times F_v &= \left(-\frac{\partial f}{\partial u}, -\frac{\partial f}{\partial v}, 1\right), & \|F_u \times F_v\| &= \sqrt{\left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2 + 1} \end{aligned}$$

und damit schließlich

$$N(u, v) = \frac{1}{\sqrt{\left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2 + 1}} \left(-\frac{\partial f}{\partial u}, -\frac{\partial f}{\partial v}, 1\right). \quad \blacksquare$$

Beispiel 6. Sei $D = E = \{(u, v) \in \mathbb{R}^2 : u^2 + v^2 < 1\}$. Die Parametrisierungen

$$F : \overline{D} \rightarrow \mathbb{R}^3, \quad F(u, v) = (u, v, uv)^T$$

und

$$G : \bar{E} \rightarrow \mathbb{R}^3, \quad G(u, v) = (u, -v, -uv)^T$$

liefern das gleiche Flächenstück. Wir bestimmen die Normalenvektoren im Punkt $F(0, 0) = G(0, 0) = (0, 0, 0)^T$. Zunächst ist

$$F_u(u, v) = (1, 0, v)^T, \quad F_v(u, v) = (0, 1, u)^T,$$

also

$$F_u(u, v) \times F_v(u, v) = (-v, -u, 1)^T \quad \text{und} \quad F_u(0, 0) \times F_v(0, 0) = (0, 0, 1)^T.$$

Andererseits ist

$$G_u(u, v) = (1, 0, -v)^T, \quad G_v(u, v) = (0, -1, -u)^T$$

und damit

$$G_u(u, v) \times G_v(u, v) = (-v, u, -1)^T \quad \text{und} \quad G_u(0, 0) \times G_v(0, 0) = (0, 0, -1)^T.$$

Bei Verwendung von F erhalten wir also $(0, 0, 1)^T$ als Normaleneinheitsvektor im Punkt $(0, 0, 0)^T$ an die Sattelfläche, und bei Verwendung von G den Vektor $(0, 0, -1)^T$. Man beachte, dass zwar F und G durch den Diffeomorphismus

$$\varphi : \bar{D} \rightarrow \bar{E}, \quad (u, v)^T \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = (u, -v)^T$$

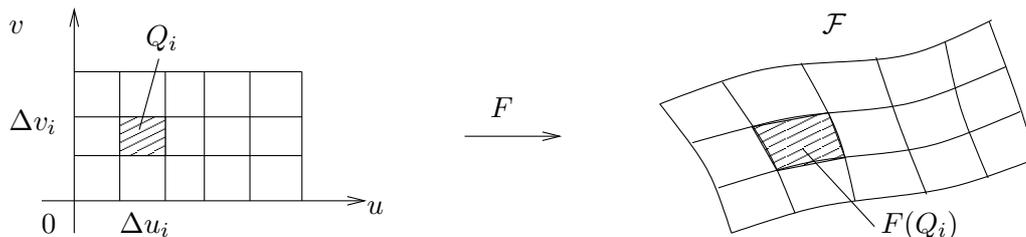
auseinander hervorgehen, dass aber

$$\det \varphi'(u, v) = \det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -1 < 0$$

ist. Also sind F und G nicht zueinander äquivalent. ■

14.2 Flächenintegrale

Wir wollen nun Flächenintegrale definieren. Als Motivation gehen wir ähnlich vor wie bei der „Herleitung“ der Substitutionsregel in Abschnitt 13.7. Sei $F : \bar{D} \rightarrow \mathbb{R}^3$ eine Parametrisierung eines Flächenstückes \mathcal{F} , wobei wir der Einfachheit halber annehmen, dass \bar{D} ein achsenparalleles Rechteck in der uv -Ebene ist. Das Rechteck \bar{D} sei in Teilrechtecke Q_1, \dots, Q_m zerlegt. Ihre Bilder $F(Q_1), \dots, F(Q_m)$ nennen wir *Maschen*. Aus diesen Maschen setzt sich das Flächenstück \mathcal{F} zusammen.



Ist die Rechteckzerlegung von \bar{D} fein genug, so haben die Maschen nahezu die Gestalt eines Parallelogramms. Ist Q_i ein Teilrechteck in \bar{D} mit den Seitenlängen Δu_i und Δv_i und ist (u_i, v_i) der linke untere Eckpunkt von Q_i , so ist die Masche $F(Q_i)$ ungefähr gleich dem Parallelogramm, das von den Vektoren

$$F(u_i + \Delta u_i, v_i) - F(u_i, v_i) \quad \text{und} \quad F(u_i, v_i + \Delta v_i) - F(u_i, v_i)$$

aufgespannt wird. Diese Vektoren sind nach Definition der partiellen Ableitungen in etwa gleich

$$F_u(u_i, v_i)\Delta u_i \quad \text{und} \quad F_v(u_i, v_i)\Delta v_i.$$

Diese beiden Vektoren spannen ein Parallelogramm auf, dessen Flächeninhalt gleich der Länge des Vektorprodukts dieser Vektoren ist, also gleich der Länge von

$$\Delta\sigma_i := (F_u(u_i, v_i) \times F_v(u_i, v_i)) \Delta u_i \Delta v_i.$$

Man kann daher $\|\Delta\sigma_i\|$ als ungefähren Flächeninhalt der Masche $F(Q_i)$ ansehen und

$$\sum_{i=1}^m \|\Delta\sigma_i\| = \sum_{i=1}^m \|F_u(u_i, v_i) \times F_v(u_i, v_i)\| \Delta u_i \Delta v_i \quad (14.7)$$

als Näherung für den Flächeninhalt des Flächenstückes \mathcal{F} . Ist D kein Rechteck, so schöpfen wir \bar{D} von innen durch rechteckzerlegte Bereiche aus. Lassen wir auf der rechten Seite von (14.7) die Zerlegung immer feiner werden, d.h. lassen wir $\max_i\{\Delta u_i, \Delta v_i\}$ gegen 0 streben, so gelangen wir zum Integral

$$\iint_{\bar{D}} \|F_u(u, v) \times F_v(u, v)\| d(u, v).$$

Um sicherzustellen, dass dies tatsächlich dem Flächeninhalt von \mathcal{F} entspricht, müssen wir noch (ähnlich wie bei Kurven) garantieren, dass nicht Teile des Flächenstücks mehrfach durchlaufen werden.

Definition 14.3 Die Parameterdarstellung $F : \bar{D} \rightarrow \mathbb{R}^3$ und das durch sie definierte Flächenstück heißen doppelpunktfrei, wenn F auf D eineindeutig ist.

Definition 14.4 Ist $F : \bar{D} \rightarrow \mathbb{R}^3$ eine doppelpunktfreie Parameterdarstellung eines Flächenstückes \mathcal{F} , so ist der Flächeninhalt von \mathcal{F} die Zahl

$$I(\mathcal{F}) := \iint_{\bar{D}} \|F_u(u, v) \times F_v(u, v)\| d(u, v). \quad (14.8)$$

Das Integral in (14.8) schreibt man auch als $\int_{\mathcal{F}} d\sigma$, wobei $d\sigma$ symbolisch für

$$\|F_u(u, v) \times F_v(u, v)\| d(u, v)$$

steht und *Flächenelement* heißt. Definition 14.4 wird wie folgt verallgemeinert:

Definition 14.5 Durch $F : \bar{D} \rightarrow \mathbb{R}^3$ sei ein Flächenstück \mathcal{F} gegeben, und $H : \mathcal{F} \rightarrow \mathbb{R}$ sei stetig auf \mathcal{F} . Dann heißt das Integral

$$\iint_{\bar{D}} H(F(u, v)) \|F_u(u, v) \times F_v(u, v)\| d(u, v) =: \int_{\mathcal{F}} H d\sigma \quad (14.9)$$

das Flächenintegral von H über \mathcal{F} (oder auch Flächenintegral erster Art).

Als Motivation kann man sich ein elektrostatisch geladenes Flächenstück \mathcal{F} vorstellen, wobei die Ladungsdichte H in jedem Punkt von \mathcal{F} bekannt und die Gesamtladung gesucht ist.

Man kann (und muss) sich überlegen, dass die Integrale in (14.8) und (14.9) bei äquivalenten Parametertransformationen und Orientierungswechsel invariant bleiben. Im Falle der Doppelpunktfreiheit ist die Schreibweise rechts in (14.9) völlig eindeutig, da je zwei zugehörige Parameterdarstellungen entweder äquivalent oder entgegengesetzt orientiert sind.

Beispiel 7. Seien D und F wie in Beispiel 1 (Kugeloberfläche). Mit Beispiel 4 erhalten wir

$$I(\mathcal{F}) = \iint_{\bar{D}} r^2 \cos v d(u, v) = r^2 \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \cos v dv du = 4\pi r^2$$

als Flächeninhalt der Kugeloberfläche. ■

Beispiel 8. Sind D , f und F wie in Beispiel 2 (Funktionsgraphen), so ist

$$I(\mathcal{F}) = \int_{\mathcal{F}} d\sigma = \iint_{\bar{D}} \sqrt{1 + f_u^2 + f_v^2} d(u, v)$$

der Flächeninhalt des Funktionsgraphen. Vergleichen Sie dieses Resultat mit der entsprechenden Formel für die Kurvenlänge des Graphen einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ (vgl. Beispiel 2 in Abschnitt 11.2). ■

Definition 14.6 Durch $F : \bar{D} \rightarrow \mathbb{R}^3$ sei ein Flächenstück \mathcal{F} gegeben, und $H : \mathcal{F} \rightarrow \mathbb{R}^3$ sei ein stetiges Vektorfeld auf \mathcal{F} . Dann heißt

$$\int_{\mathcal{F}} H \cdot d\vec{\sigma} := \iint_{\bar{D}} H(F(u, v)) \cdot (F_u(u, v) \times F_v(u, v)) d(u, v) \quad (14.10)$$

das Flächenintegral von H über \mathcal{F} (auch Flächenintegral zweiter Art). (Der Punkt steht für das Skalarprodukt.)

Schreiben wir

$$\begin{aligned} F_u(u, v) \times F_v(u, v) &= \frac{F_u(u, v) \times F_v(u, v)}{\|F_u(u, v) \times F_v(u, v)\|} \|F_u(u, v) \times F_v(u, v)\| \\ &= N(u, v) \|F_u(u, v) \times F_v(u, v)\|, \end{aligned}$$

so geht das Integral auf der rechten Seite von (14.10) über in

$$\iint_{\overline{D}} H(F(u, v)) \cdot N(u, v) \|F_u(u, v) \times F_v(u, v)\| d(u, v), \quad (14.11)$$

so dass man das Integral $\int_{\mathcal{F}} H \cdot d\vec{\sigma}$ zweiter Art auch als das Integral $\int_{\mathcal{F}} H \cdot N d\sigma$ erster Art auffassen kann. Flächenintegrale zweiter Art sind ebenfalls invariant bezüglich äquivalenter Parametertransformationen. Bei einem Orientierungswechsel ändern sie jedoch ihr Vorzeichen, da die Normalenvektoren ihre Richtung ändern.

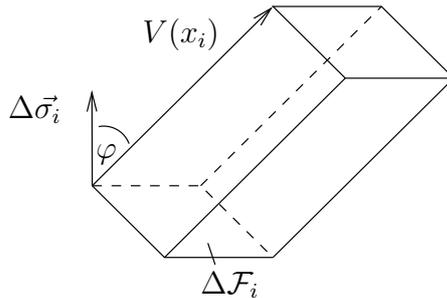
Zur *Motivation* der Flächenintegrale zweiter Art stellen wir uns ein stationäres (zeitunabhängiges) Geschwindigkeitsfeld $V : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ einer strömenden Flüssigkeit vor und fragen nach der Flüssigkeitsmenge, die ein gegebenes Flächenstück \mathcal{F} pro Zeiteinheit durchfließt. Sinnvollerweise soll \mathcal{F} orientiert sein, d.h. wir können uns etwa vorstellen, dass \mathcal{F} eine „Unterseite“ und eine „Oberseite“ hat und dass die Normalenvektoren in Richtung der Oberseite zeigen.

Wie bei der Motivation zum Flächeninhalt denken wir uns \mathcal{F} in Maschen unterteilt, die näherungsweise Parallelogrammform haben. Es sei $\Delta\vec{\sigma}_i$ der *Flächenvektor* eines solchen Parallelogramms $\Delta\mathcal{F}_i$, d.h. $\Delta\vec{\sigma}_i$ steht senkrecht auf $\Delta\mathcal{F}_i$ und zeigt in Normalenrichtung, und $\|\Delta\vec{\sigma}_i\|$ ist gleich dem Flächeninhalt von $\Delta\mathcal{F}_i$.

Dann ist $|V(x_i) \cdot \Delta\vec{\sigma}_i|$ (mit einem $x_i \in \Delta\mathcal{F}_i$) das Flüssigkeitsvolumen, das näherungsweise pro Zeiteinheit durch $\Delta\mathcal{F}_i$ fließt. Pro Zeiteinheit schiebt sich nämlich ein Parallelepiped mit dem Grundflächeninhalt $\|\Delta\vec{\sigma}_i\|$ und der Höhe $\|V(x_i)\| \cos \varphi$ (vgl. die folgende Skizze), d.h. mit dem Volumen

$$\|V(x_i)\| \|\Delta\vec{\sigma}_i\| \cos \varphi = V(x_i) \cdot \Delta\vec{\sigma}_i$$

durch $\Delta\mathcal{F}_i$. (In erster Näherung nehmen wir V als konstant auf $\Delta\mathcal{F}_i$ an.)



Fließt die Flüssigkeit aus der Seite von $\Delta\mathcal{F}_i$ heraus, in die der Flächenvektor $\Delta\vec{\sigma}_i$ zeigt, so ist $V(x_i) \cdot \Delta\vec{\sigma}_i \geq 0$ und andernfalls ≤ 0 . Das Vorzeichen von $V(x_i) \cdot \Delta\vec{\sigma}_i$ gibt also an, in welche Richtung $\Delta\mathcal{F}_i$ durchflossen wird.

Die Summation der Durchflüsse über alle Maschen

$$\sum_i V(x_i) \cdot \Delta\vec{\sigma}_i$$

und der Übergang zu beliebig kleinen Maschenweiten führen zum Flächenintegral 2. Art

$$U = \int_{\mathcal{F}} V(x) \cdot d\vec{\sigma}.$$

Die Größe $|U|$ gibt also das Gesamtvolumen an, welches pro Zeiteinheit das Flächenstück \mathcal{F} durchströmt, wobei die Anteile der beiden Strömungsrichtungen durch \mathcal{F} gegeneinander aufgerechnet sind. Das Vorzeichen von U gibt an, an welcher Seite des Flächenstücks mehr herausfließt: Ist $U > 0$, so strömt mehr Flüssigkeit in Richtung der Normalenvektoren von \mathcal{F} , ist $U < 0$, so in entgegengesetzter Richtung. Man nennt U auch den *Fluss* von V durch \mathcal{F} . ■

Beispiel 9. Seien D und F wieder wie in Beispiel 1 (Kugeloberfläche), und sei $H : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ die identische Abbildung, d.h. $H(x, y, z) = (x, y, z)$. Dann ist nach Beispiel 4

$$N(u, v) = (\cos u \cos v, \sin u \cos v, \sin v),$$

und mit

$$H(F(u, v)) = F(u, v) = (r \cos u \cos v, r \sin u \cos v, r \sin v)$$

erhalten wir mit (14.11) für das Flächenintegral 2. Art

$$\int_{\mathcal{F}} H \cdot d\vec{\sigma} = \int_{\mathcal{F}} H \cdot N \, d\sigma = \int_{\mathcal{F}} r \, d\sigma = r \int_{\mathcal{F}} d\sigma.$$

Das Flächenintegral $\int_{\mathcal{F}} d\sigma$ ist gleich dem Flächeninhalt der Kugeloberfläche. Wir haben es in Beispiel 7 berechnet und erhalten damit

$$\int_{\mathcal{F}} H \cdot d\vec{\sigma} = 4\pi r^3. \quad \blacksquare$$

14.3 Die Divergenz eines Vektorfeldes

Unsere letzten Ziele in diesem Kapitel sind die Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung

$$\int_a^b f'(x) \, dx = f(b) - f(a) = f(x)|_a^b$$

und der Formel der partiellen Integration

$$\int_a^b u'(x)v(x) \, dx = - \int_a^b u(x)v'(x) \, dx + u(x)v(x)|_a^b$$

auf mehrdimensionale Integrale. Es stellt sich die Frage, wodurch die Ableitungen f', u', v' und die Randterme $f|_a^b$ und $uv|_a^b$ zu ersetzen sind.

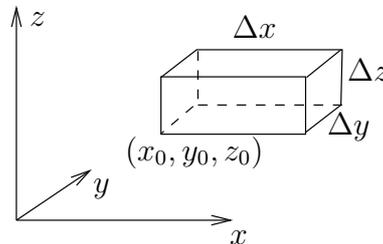
Unser erstes Ziel ist der *Gaußsche Integralsatz* im \mathbb{R}^3 . Seine anschauliche Bedeutung ist völlig einleuchtend:

Die Flüssigkeitsmenge, die durch die Oberfläche eines räumlichen Gebietes herausströmt, ist gleich der Flüssigkeitsmenge, die die Quellen in diesem Gebiet hervorbringen.

Wie kann man die Flüssigkeitsmenge, die eine Quelle im Punkt $(x_0, y_0, z_0) \in \mathbb{R}^3$ hervorbringt, mathematisch beschreiben? Wir betrachten eine stationäre (zeitunabhängige) Strömung einer inkompressiblen Flüssigkeit (d.h. mit konstanter Dichte), die im Punkt $(x, y, z) \in \mathbb{R}^3$ die Geschwindigkeit

$$V(x, y, z) = (V_1(x, y, z), V_2(x, y, z), V_3(x, y, z))$$

hat. Im Punkt (x_0, y_0, z_0) heften wir einen kleinen achsenparallelen Quader Q mit den Seitenlängen Δx , Δy und Δz an.



Dann ist das Flüssigkeitsvolumen, das pro Zeiteinheit in Richtung der positiven x -Achse durch die linke bzw. rechte Seitenwand des Quaders fließt, näherungsweise gleich

$$V_1(x_0, y_0, z_0)\Delta y\Delta z \quad \text{bzw.} \quad V_1(x_0 + \Delta x, y_0, z_0)\Delta y\Delta z.$$

Das Volumen, das pro Zeiteinheit aus dem Quader Q in der positiven x -Richtung austritt, ist also etwa gleich

$$\begin{aligned} & (V_1(x_0 + \Delta x, y_0, z_0) - V_1(x_0, y_0, z_0))\Delta y\Delta z \\ &= \frac{V_1(x_0 + \Delta x, y_0, z_0) - V_1(x_0, y_0, z_0)}{\Delta x} \Delta x\Delta y\Delta z \\ &\approx \frac{\partial V_1}{\partial x}(x_0, y_0, z_0)\Delta x\Delta y\Delta z. \end{aligned}$$

Wir stellen in ähnlicher Weise die Massenbilanz für den Fluß in positiver y - und z -Richtung auf und erhalten als Endresultat, dass das Flüssigkeitsvolumen, das pro Zeiteinheit aus Q austritt, ungefähr gleich

$$\left(\frac{\partial V_1}{\partial x}(x_0, y_0, z_0) + \frac{\partial V_2}{\partial y}(x_0, y_0, z_0) + \frac{\partial V_3}{\partial z}(x_0, y_0, z_0) \right) \Delta x\Delta y\Delta z \quad (14.12)$$

ist. Mit der in Abschnitt 10.2 eingeführten *Divergenz*

$$(\operatorname{div} F)(x) = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i}(x), \quad x \in D$$

eines stetig differenzierbaren Vektorfeldes $F = (F_1, \dots, F_n) : D \rightarrow \mathbb{R}^n$ können wir (14.12) schreiben als

$$(\operatorname{div} V)(x_0, y_0, z_0) \Delta x \Delta y \Delta z.$$

Dividieren wir diesen Wert durch das Volumen $\Delta x \Delta y \Delta z$ von Q und ziehen wir Q auf den Punkt (x_0, y_0, z_0) zusammen, so können wir $(\operatorname{div} V)(x_0, y_0, z_0)$ als *Quelldichte* der Strömung im Punkt (x_0, y_0, z_0) interpretieren. Ist $(\operatorname{div} V)(x_0, y_0, z_0) > 0$, so ist (x_0, y_0, z_0) eine *Quelle* im eigentlichen Sinn (ihr entströmt Flüssigkeit). Im Fall $(\operatorname{div} V)(x_0, y_0, z_0) < 0$ heißt (x_0, y_0, z_0) eine *Senke* (da in diesem Punkt Flüssigkeit verschwindet).

Einige Rechenregeln für die Divergenz. Sei $D \subseteq \mathbb{R}^n$ offen, und seien $F, G : D \rightarrow \mathbb{R}^n$ und $\varphi : D \rightarrow \mathbb{R}$ stetig differenzierbar und $\mu \in \mathbb{R}$. Dann gilt

$$\begin{aligned} \operatorname{div}(F + G) &= \operatorname{div} F + \operatorname{div} G, \\ \operatorname{div}(\mu F) &= \mu \operatorname{div} F, \\ \operatorname{div}(\varphi F) &= \varphi \operatorname{div} F + \operatorname{grad} \varphi \cdot F. \end{aligned}$$

14.4 Der Gaußsche Integralsatz im Raum

Wir können nun die anschauliche Aussage des Gaußschen Integralsatzes in Formeln fassen. Es sei G ein geeigneter räumlicher Bereich und ∂G sein Rand (seine Oberfläche). Die genauen Voraussetzungen geben wir später an. In jedem Randpunkt haben wir zwei Normaleneinheitsvektoren: einen, der in den Körper hineinzeigt (die sogenannte *innere Normale*) und einen, der von G weg zeigt (die *äußere Normale*). Es sei $N : \partial G \rightarrow \mathbb{R}^3$ das Vektorfeld der äußeren Einheitsnormalen. Weiter sei V das Geschwindigkeitsfeld einer inkompressiblen Flüssigkeit. Nach Abschnitt 14.2 (Interpretation des Flächeninhalts 2. Art) ist der Durchfluß durch ∂G in Richtung der äußeren Normalen (also das, was aus G *herausfließt*), gleich

$$\iint_{\partial G} V \cdot N \, d\sigma. \quad (14.13)$$

Die durch Quellen und Senken in G hervorgebrachte Flüssigkeitsmenge erhalten wir dagegen durch Aufintegrieren der Quelldichte über G :

$$\iiint_G (\operatorname{div} V)(x) \, dx. \quad (14.14)$$

Nach dem Gaußschen Integralsatz sind die Integrale (14.13) und (14.14) gleich. Nun zur exakten Formulierung.

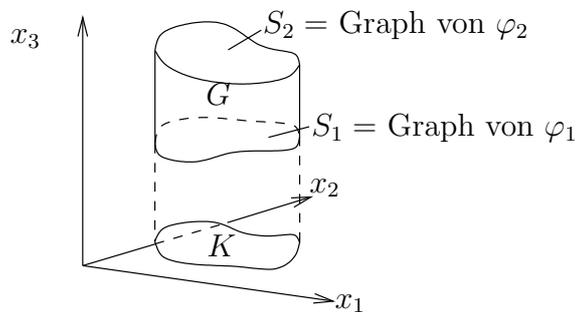
Definition 14.7 Eine Menge $G \subseteq \mathbb{R}^3$ heißt C^1 -Normalbereich bzgl. der x_1x_2 -Ebene, wenn es eine kompakte Menge $K \subseteq \mathbb{R}^2$ und stetig differenzierbare Funktionen $\varphi_1, \varphi_2 : K \rightarrow \mathbb{R}$ so gibt, dass

$$G = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : (x_1, x_2) \in K, \varphi_1(x_1, x_2) \leq x_3 \leq \varphi_2(x_1, x_2)\}$$

gilt und dass der Rand ∂K durch einen stückweise stetig differenzierbaren Weg darstellbar ist. Analog erklärt man C^1 -Normalbereiche bezüglich der x_2x_3 - und x_1x_3 -Ebene. Die Menge G heißt ein C^1 -Normalbereich, wenn sie ein C^1 -Normalbereich bezüglich der x_1x_2 -, der x_2x_3 - und der x_1x_3 -Ebene ist.

(Es sei noch einmal an unsere Vereinbarung erinnert: Eine Funktion $f : K \rightarrow \mathbb{R}$ auf einer kompakten Menge K heißt stetig differenzierbar, wenn sie zu einer stetig differenzierbaren Funktion auf einer offenen Menge $G \supseteq K$ fortgesetzt werden kann.)

Einen C^1 -Normalbereich bzgl. der x_1x_2 -Ebene kann man sich so vorstellen:



Der obere Deckel S_2 ist ein Flächenstück im \mathbb{R}^3 mit der Parameterdarstellung

$$F_2 : K \rightarrow \mathbb{R}^3, \quad (x_1, x_2) \mapsto (x_1, x_2, \varphi_2(x_1, x_2)).$$

Der durch F_2 bestimmte Normaleneinheitsvektor (vgl. Beispiel 5 aus 14.1)

$$N_2(u, v) = \frac{1}{\sqrt{\left(\frac{\partial \varphi_2}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_2}{\partial x_2}\right)^2 + 1}} \left(-\frac{\partial \varphi_2}{\partial x_1}, -\frac{\partial \varphi_2}{\partial x_2}, 1 \right)$$

ist der *äußere* Normaleneinheitsvektor für G (da die z -Komponente positiv ist). Der untere Deckel S_1 wird beschrieben durch

$$F_1 : K \rightarrow \mathbb{R}^3, \quad (x_1, x_2) \mapsto (x_1, x_2, \varphi_1(x_1, x_2)),$$

und der zugehörigen Normaleneinheitsvektor ist

$$N_1(u, v) = \frac{1}{\sqrt{\left(\frac{\partial \varphi_1}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_1}{\partial x_2}\right)^2 + 1}} \left(-\frac{\partial \varphi_1}{\partial x_1}, -\frac{\partial \varphi_1}{\partial x_2}, 1 \right).$$

Dieser zeigt ebenfalls in Richtung der positiven z -Achse (also in G hinein), so dass der *äußere* Normalenvektor an G auf S_1 gleich $-N_1(u, v)$ ist.

Satz 14.8 (Gaußscher Integralsatz im \mathbb{R}^3) Sei $G \subseteq \mathbb{R}^3$ ein C^1 -Normalbereich und $H : G \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Weiter bezeichne $N : \partial G \rightarrow \mathbb{R}^3$ das äußere Normalenfeld von G . Dann ist

$$\iiint_G (\operatorname{div} H)(x) \, dx = \iint_{\partial G} H \cdot N \, d\sigma. \quad (14.15)$$

Beweis. Es genügt, die Aussage zu beweisen, wenn H die Gestalt $H = (0, 0, H_3)^T$ hat (man kann ja H als Summe dreier derartiger Ausdrücke schreiben, und die Divergenz sowie die Integrale in (14.15) sind linear in H). Mit Satz 13.34 erhalten wir dann wegen $\operatorname{div} H = \frac{\partial H_3}{\partial x_3}$

$$\begin{aligned} \iiint_G (\operatorname{div} H)(x) \, dx &= \iint_K \left(\int_{\varphi_1(x_1, x_2)}^{\varphi_2(x_1, x_2)} \frac{\partial H_3}{\partial x_3}(x_1, x_2, x_3) \, dx_3 \right) d(x_1, x_2) \\ &= \iint_K \left(H_3(x_1, x_2, \varphi_2(x_1, x_2)) - H_3(x_1, x_2, \varphi_1(x_1, x_2)) \right) d(x_1, x_2). \end{aligned}$$

Auf dem oberen Deckel S_2 ist

$$H \cdot N = \frac{1}{\sqrt{\left(\frac{\partial \varphi_2}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_2}{\partial x_2}\right)^2 + 1}} H_3$$

und damit

$$\begin{aligned} &\iint_K H_3(x_1, x_2, \varphi_2(x_1, x_2)) \, d(x_1, x_2) \\ &= \iint_K (H \cdot N) \sqrt{\left(\frac{\partial \varphi_2}{\partial x_1}\right)^2 + \left(\frac{\partial \varphi_2}{\partial x_2}\right)^2 + 1} \, d(x_1, x_2) = \iint_{S_2} H \cdot N \, d\sigma. \end{aligned}$$

Analog zeigt man, dass

$$- \iint_K H_3(x_1, x_2, \varphi_1(x_1, x_2)) \, d(x_1, x_2) = \iint_{S_1} H \cdot N \, d\sigma.$$

Auf dem noch fehlenden Randstück $\partial G \setminus (S_1 \cup S_2)$ von G , also auf

$$S_3 := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : (x_1, x_2) \in \partial K, \varphi_1(x_1, x_2) \leq x_3 \leq \varphi_2(x_1, x_2)\},$$

existiert der äußere Normalenvektor N ebenfalls (bis auf endlich viele Geradenstücke $\{y\} \times [\varphi_1(y), \varphi_2(y)]$ mit $y \in \partial K$, in denen die Parametrisierung von ∂K nicht differenzierbar ist), und die Komponente von N in x_3 -Richtung ist 0. Daher ist $H \cdot N = 0$ auf S_3 . Zusammengefasst erhalten wir

$$\iiint_G (\operatorname{div} H)(x) \, dx = \sum_{j=1}^3 \iint_{S_j} H \cdot N \, d\sigma = \iint_{\partial G} H \cdot N \, d\sigma. \quad \blacksquare$$

Beispiel 1. Wir kommen noch einmal zurück auf Beispiel 9 aus Abschnitt 14.2. Dort waren D und F wie in Beispiel 1 aus Abschnitt 14.1 (Kugeloberfläche) und $H(x, y, z) = (x, y, z)$, und wir haben erhalten, dass

$$\iint_{\partial G} H \cdot d\vec{\sigma} = \iint_{\partial G} H \cdot N \, d\sigma = 4\pi r^3.$$

Andererseits ist $\operatorname{div} H = 3$, so dass mit dem Volumen V der Kugel G gilt

$$\iiint_G (\operatorname{div} H)(x) \, dx = 3 \iiint_G dx = 3V.$$

Der Gaußsche Satz liefert nach Gleichsetzen

$$V = \frac{4}{3} \pi r^3. \quad \blacksquare$$

Folgerung 14.9 (Partielle Integration in \mathbb{R}^3) Sei $G \subseteq \mathbb{R}^3$ ein C^1 -Normalbereich und $f, g : G \rightarrow \mathbb{R}$ seien stetig differenzierbare Funktionen. Weiter sei $N = (N_1, N_2, N_3) : \partial G \rightarrow \mathbb{R}$ das äußere Normalenfeld an G . Dann gilt für $i = 1, 2, 3$

$$\iiint_G f \frac{\partial g}{\partial x_i} \, dx = - \iiint_G \frac{\partial f}{\partial x_i} g \, dx + \iint_{\partial G} f g N_i \, d\sigma. \quad (14.16)$$

Beweis. Sei z.B. $i = 1$. Für die stetig differenzierbare Funktion

$$F : G \rightarrow \mathbb{R}^3, \quad F(x) = (f(x)g(x), 0, 0) \quad (14.17)$$

ist

$$\operatorname{div} F = \frac{\partial(fg)}{\partial x_1} = \frac{\partial f}{\partial x_1} g + f \frac{\partial g}{\partial x_1}$$

und $F \cdot N = fg N_1$. Die Behauptung folgt sofort, wenn man die Funktion (14.17) in den Gaußschen Integralsatz (14.15) einsetzt. \blacksquare

Anmerkung 1. Wir haben in Satz 14.8 und Folgerung 14.9 statt \int_G und $\int_{\partial G}$ die Schreibweisen \iiint_G und $\iint_{\partial G}$ benutzt, um deutlich zu machen, dass (dreidimensionale) Volumenintegrale und (zweidimensionale) Oberflächenintegrale auftreten.

Anmerkung 2. Sind F_1, \dots, F_n Flächenstücke, die sich nicht überlappen, so definiert man das Flächenintegral über $F = F_1 \cup \dots \cup F_n$ durch

$$\int_F = \int_{F_1} + \dots + \int_{F_n}.$$

Dies haben wir im Beweis des Gaußschen Satzes benutzt ($\partial G = S_1 \cup S_2 \cup S_3$).

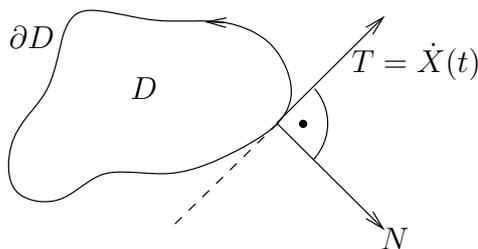
Anmerkung 3. Satz 14.8 und Folgerung 14.9 gelten auch unter schwächeren Voraussetzungen. Z.B. genügt es, dass sich G als endliche Vereinigung sich nicht überlappender C^1 -Normalbereiche schreiben läßt. Man beachte auch, dass der Normalenvektor nicht in jedem Randpunkt definiert ist (z.B. nicht entlang der Kanten eines Quaders).

14.5 Der Gaußsche Integralsatz in der Ebene

Der Gaußsche Integralsatz und seine Folgerung gelten (bei geeigneter Definition des Flächenintegrals) für jede Raumdimension. Den Gaußschen Integralsatz im \mathbb{R}^2 kann man durch geeignete Reduzierung um eine Koordinate wie folgt aus dem Gaußschen Integralsatz im \mathbb{R}^3 gewinnen.

Sei $D \subseteq \mathbb{R}^2$ ein Normalbereich bezüglich der x_1 - und der x_2 -Achse (vgl. Definition 14.8). In diesem Fall nennen wir D einfach einen *Normalbereich*. Der Rand ∂D sei eine geschlossene Kurve, die durch einen stückweise stetig differenzierbaren Weg $X : [a, b] \rightarrow \mathbb{R}^2$ parametrisiert werde. Durchläuft t das Intervall $[a, b]$ von a nach b , so wandert $X(t)$ entlang ∂D in einer bestimmten Richtung. Wir wollen annehmen, dass dabei G stets links des Weges liegt. Man sagt auch, dass D vom Weg X *positiv umlaufen* wird oder dass der Rand ∂D *positiv orientiert* ist (anschaulich: im Gegenuhrzeigersinn).

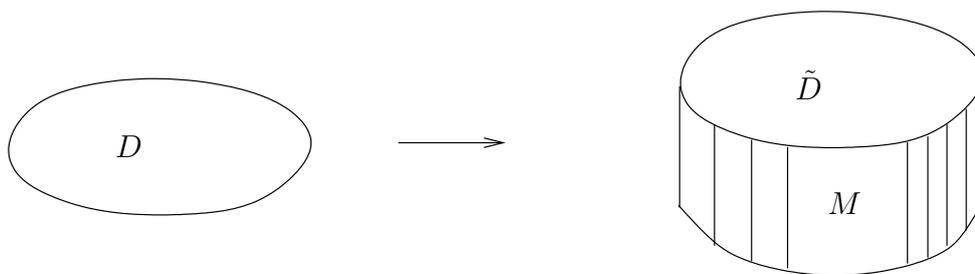
Der Tangentialvektor an ∂D im Punkt $X(t) = (X_1(t), X_2(t))$ ist gegeben durch $T := (\dot{X}_1(t), \dot{X}_2(t))$, und $N := (\dot{X}_2(t), -\dot{X}_1(t))$ ist ein auf T senkrecht stehender Vektor, der nach außen zeigt. Man beachte, dass diese Vektoren nur in Punkten definiert sind, in denen X stetig differenzierbar ist.



Auf D sei ein stetig differenzierbares Vektorfeld $V = (V_1, V_2)^T : D \rightarrow \mathbb{R}^2$ gegeben. Wir bilden aus D den räumlichen Bereich

$$\tilde{D} := D \times [0, 1] = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : (x_1, x_2) \in D, x_3 \in [0, 1]\},$$

d.h. \tilde{D} ist eine Scheibe (= ein Zylinder) der Dicke 1, bei dem Boden und Deckel die Form von D haben. Offenbar ist \tilde{D} ein Normalbereich im \mathbb{R}^3 .



Außerdem erweitern wir V um eine Komponente:

$$\tilde{V} : \tilde{D} \rightarrow \mathbb{R}^3, \quad (x_1, x_2, x_3) \mapsto (V_1(x_1, x_2), V_2(x_1, x_2), 0)^T.$$

Der Gaußsche Integralsatz (14.15), angewandt auf \tilde{D} und \tilde{V} , liefert

$$\iint_{\partial \tilde{D}} \tilde{V} \cdot N \, d\sigma = \iiint_{\tilde{D}} \operatorname{div} \tilde{V} \, dx. \quad (14.18)$$

Beim Flächenintegral auf der linken Seite heben sich die Anteile des Bodens und des Deckels weg, da die zugehörigen Normalenvektoren entgegengesetzt sind, sonst jedoch alles gleich ist. Es ist insbesondere $\tilde{V}(x_1, x_2, 0) = \tilde{V}(x_1, x_2, 1)$. Also verbleibt nur das Integral über die Mantelfläche $M := \partial D \times [0, 1]$. Aus (14.18) folgt somit

$$\iint_M \tilde{V} \cdot N \, d\sigma = \iiint_{\tilde{D}} \operatorname{div} \tilde{V} \, dx. \quad (14.19)$$

Die Mantelfläche hat eine Parameterdarstellung

$$F : M = [a, b] \times [0, 1] \rightarrow \mathbb{R}^3, \quad (t, z) \mapsto (X_1(t), X_2(t), z)^T,$$

woraus man den äußeren (noch unnormierten!) Normalenvektor im Punkt $F(t, z)$

$$n(F(t, z)) = (\dot{X}_2(t), -\dot{X}_1(t), 0)^T$$

erhält. Mit (14.11) bekommen wir für die linke Seite von (14.19)

$$\begin{aligned} \iint_M \tilde{V} \cdot N \, d\sigma &= \iint_M \tilde{V} \cdot \frac{n}{\|n\|} \, d\sigma \\ &\stackrel{(14.11)}{=} \iint_{[a,b] \times [0,1]} \tilde{V}(F(t, z)) \cdot n(F(t, z)) \, d(t, z) \\ &= \int_0^1 \int_a^b (V_1(X(t)), V_2(X(t)), 0) \cdot (\dot{X}_2(t), -\dot{X}_1(t), 0) \, dt \, dz \\ &= \int_a^b (V_1(X(t))\dot{X}_2(t) - V_2(X(t))\dot{X}_1(t)) \, dt, \end{aligned}$$

während auf der rechten Seite von (14.19)

$$\operatorname{div} \tilde{V} = \frac{\partial V_1}{\partial x_1} + \frac{\partial V_2}{\partial x_2} = \operatorname{div} V$$

ist. Daher ist

$$\iiint_{\tilde{D}} \operatorname{div} \tilde{V} \, d(x_1, x_2, x_3) = \int_0^1 \iint_D \operatorname{div} V \, d(x_1, x_2) \, dx_3 = \iint_D \operatorname{div} V \, d(x_1, x_2).$$

Zusammengefaßt erhalten wir

Satz 14.10 (Gaußscher Integralsatz im \mathbb{R}^2) Sei $D \subseteq \mathbb{R}^2$ ein Normalbereich (bzgl. der x_1 - und der x_2 -Achse), dessen Rand ∂D durch einen stückweise stetig differenzierbaren Weg $X = (X_1, X_2)^T : [a, b] \rightarrow \mathbb{R}^2$ parametrisiert wird, der D positiv umläuft. Weiter sei $V = (V_1, V_2)^T : D \rightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_a^b (V_1(X(t))\dot{X}_2(t) - V_2(X(t))\dot{X}_1(t)) dt = \iint_D (\operatorname{div} V)(x) dx. \quad (14.20)$$

Man beachte, dass auf der linken Seite von (14.20) ein Wegintegral entlang des Weges X steht. Nehmen wir die Umbenennung

$$W_1 := V_2, \quad W_2 := -V_1, \quad W := (W_1, W_2)^T$$

vor, so geht nach Multiplikation mit -1 die linke Seite von (14.20) über in

$$\int_a^b (W_1(X(t))\dot{X}_1(t) + W_2(X(t))\dot{X}_2(t)) dt,$$

d.h. in das Wegintegral $\int_{\partial D} W \cdot dX$ über den durch X parametrisierten Rand von D , während auf der rechten Seite von (14.20) $-\operatorname{div} V = -\frac{\partial V_1}{\partial x_1} - \frac{\partial V_2}{\partial x_2}$ zu ersetzen ist durch $\frac{\partial W_2}{\partial x_1} - \frac{\partial W_1}{\partial x_2}$. Man schreibt oft

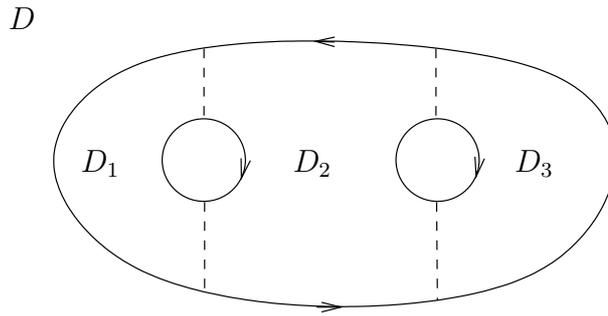
$$\operatorname{rot} W := \frac{\partial W_2}{\partial x_1} - \frac{\partial W_1}{\partial x_2} \quad (14.21)$$

und nennt $\operatorname{rot} W$ die (skalarwertige) *Rotation* des Vektorfeldes $W = (W_1, W_2)^T$. Mit diesen Bezeichnungen erhalten wir die folgende Version des Gaußschen Satzes im \mathbb{R}^2 .

Satz 14.11 (Greenscher Integralsatz) Seien D und X wie in Satz 14.10, und $W = (W_1, W_2)^T : \bar{D} \rightarrow \mathbb{R}^2$ sei ein stetig differenzierbares Vektorfeld. Dann ist

$$\int_{\partial D} W \cdot dX = \int_D \left(\frac{\partial W_2}{\partial x_1} - \frac{\partial W_1}{\partial x_2} \right) d(x_1, x_2) = \int_D (\operatorname{rot} W)(x) dx.$$

Die Sätze 14.10 und 14.11 gelten auch unter allgemeineren Bedingungen an die Menge D . So darf D die Abschließung eines beschränkten einfach zusammenhängenden Gebietes sein, dessen Rand sich durch einen stückweise stetig differenzierbaren Weg parametrisieren läßt. Es genügt sogar, dass D sich in endlich viele derartige Mengen zerlegen läßt.



Man beachte die Orientierung des (aus mehreren Stücken bestehenden) Randes.

14.6 Der Stokessche Integralsatz

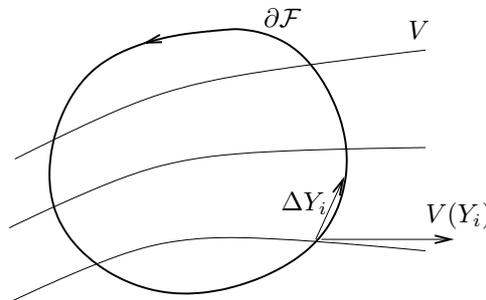
Wir beginnen wieder mit einer kurzen Motivation. Sei $M \subseteq \mathbb{R}^3$ offen und $V : M \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld, das wir als Geschwindigkeitsfeld einer strömenden Flüssigkeit deuten. In M sei ein Flächenstück \mathcal{F} gegeben, dessen Rand durch einen stückweise stetig differenzierbaren und doppeltpunktfreien Weg $Y : [a, b] \rightarrow \mathbb{R}^3$ parametrisiert wird. Unter der *Zirkulation von V entlang $\partial\mathcal{F}$* versteht man das Kurvenintegral

$$\int_{\partial\mathcal{F}} V \cdot dY = \int_a^b \sum_{i=1}^3 V_i(Y(t)) \dot{Y}_i(t) dt. \quad (14.22)$$

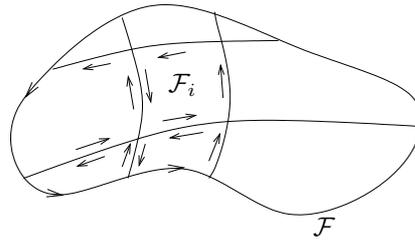
Denkt man sich dieses Integral durch Riemann-Summen

$$\sum_j V(Y_j) \cdot \Delta Y_j$$

approximiert, so entspricht jeder Summand $V(Y_i) \cdot \Delta Y_i$ einer Geschwindigkeitskomponenten in der Durchlaufrichtung der Kurve. Die Summation dieser Komponenten ist ein Maß dafür, wie stark die Kurve $\partial\mathcal{F}$ umströmt wird, d.h. wie stark die Flüssigkeit längs der Kurve zirkuliert.



Wir zerlegen das Flächenstück \mathcal{F} in endlich viele kleine Maschen \mathcal{F}_i .



Bei entsprechender Orientierung der Ränder der \mathcal{F}_i erhalten wir für die Zirkulation

$$\int_{\partial\mathcal{F}} V \cdot dY = \sum_i \int_{\partial\mathcal{F}_i} V \cdot dY = \sum_i \frac{1}{I(\mathcal{F}_i)} \int_{\partial\mathcal{F}_i} V \cdot dY I(\mathcal{F}_i), \quad (14.23)$$

wobei $I(\mathcal{F}_i)$ für den Flächeninhalt von \mathcal{F}_i steht. Nun verfeinern wir die Maschen. Man kann zeigen (vgl. etwa Burg/Haf/Wille IV, S. 156-158): Wenn man eine Masche \mathcal{F}_i auf einen Punkt $x_i \in \mathcal{F}$ zusammenzieht, dann strebt der Quotient

$$\frac{1}{I(\mathcal{F}_i)} \int_{\partial\mathcal{F}_i} V \cdot dY \quad (14.24)$$

gegen einen festen Wert, nämlich gegen $(\text{rot } V)(x_i) \cdot N(x_i)$. Dieser Ausdruck heißt auch die *Wirbelstärke* von V in x_i . Hierbei ist $N(x_i)$ die gemeinsame Flächennormale der zu x_i zusammengezogenen Maschen, und die Rotation $\text{rot } V$ ist wie in Abschnitt 10.2 durch

$$(\text{rot } F)(x) = \left(\frac{\partial F_3}{\partial x_2}(x) - \frac{\partial F_2}{\partial x_3}(x), \frac{\partial F_1}{\partial x_3}(x) - \frac{\partial F_3}{\partial x_1}(x), \frac{\partial F_2}{\partial x_1}(x) - \frac{\partial F_1}{\partial x_2}(x) \right)^T$$

erklärt. Man beachte, dass man für ein Vektorfeld der Gestalt

$$F(x_1, x_2, x_3) = (F_1(x_1, x_2), F_2(x_1, x_2), 0)$$

als Rotation gerade das Vektorfeld $(0, 0, \text{rot}(F_1, F_2))$ mit der in (14.21) eingeführten Rotation eines zweidimensionalen Vektorfeldes erhält.

Für kleine $\Delta\mathcal{F}_i$ können wir also (14.24) näherungsweise durch

$$(\text{rot } V)(x_i) \cdot N(x_i)$$

ersetzen, und aus (14.23) wird

$$\int_{\partial\mathcal{F}} V \cdot dY \approx \sum_i (\text{rot } V)(x_i) \cdot N(x_i) I(\mathcal{F}_i). \quad (14.25)$$

Die rechte Seite ist eine Riemannsumme für das Flächenintegral

$$\int_{\mathcal{F}} \text{rot } V \cdot N \, d\sigma.$$

Der folgende Satz von Stokes sagt, dass unter geeigneten Voraussetzungen an \mathcal{F} die Zirkulation $\int_{\partial\mathcal{F}} V \cdot dY$ tatsächlich gleich diesem Flächenintegral über die Wirbelstärken ist.

Satz 14.12 (Stokes'scher Integralsatz) Sei $D \subseteq \mathbb{R}^2$ ein Normalbereich, dessen Rand ∂D durch einen stückweise stetig differenzierbaren Weg $X : [a, b] \rightarrow \mathbb{R}^2$ parametrisiert wird, der D einmal positiv umläuft. Weiter sei $F : D \rightarrow \mathbb{R}^3$ eine zweimal stetig differenzierbare Abbildung, die ein Flächenstück $\mathcal{F} = F(D)$ parametrisiert. Der Weg $Y : [a, b] \rightarrow \mathbb{R}^3$, $Y(t) = F(X(t))$ definiert eine orientierte Kurve in \mathcal{F} , die wir den Rand $\partial \mathcal{F}$ von \mathcal{F} nennen. Schließlich sei $N : \mathcal{F} \rightarrow \mathbb{R}^3$ das durch F wie in (14.6) festgelegte Normalenfeld. Dann gilt für jedes stetig differenzierbare Vektorfeld $H : \mathcal{F} \rightarrow \mathbb{R}^3$

$$\int_{\mathcal{F}} \operatorname{rot} H \cdot N \, d\sigma = \int_{\partial \mathcal{F}} H \cdot dY. \quad (14.26)$$

Beweis. Es genügt, Vektorfelder H der Form $(P, 0, 0)^T$ zu betrachten. Der Beweis für $H = (0, Q, 0)^T$ und $H = (0, 0, R)^T$ verläuft analog.

Zuerst schreiben wir das Kurvenintegral $\int_{\partial \mathcal{F}} H \cdot dY$ über der Kurve $\partial \mathcal{F}$ in ein Kurvenintegral über ∂D um:

$$\begin{aligned} \int_{\partial \mathcal{F}} (P, 0, 0)^T \cdot dY &= \int_a^b P(Y(t)) \dot{Y}_1(t) \, dt \\ &= \int_a^b P(F(X(t))) (F_1 \circ X)'(t) \, dt \\ &= \int_a^b P(F(X(t))) \left(\frac{\partial F_1}{\partial x_1} \dot{X}_1(t) + \frac{\partial F_1}{\partial x_2} \dot{X}_2(t) \right) \, dt \\ &= \int_{\partial D} \left((P \circ F) \frac{\partial F_1}{\partial x_1}, (P \circ F) \frac{\partial F_1}{\partial x_2} \right) \cdot dX, \end{aligned}$$

wobei wir in der dritten Zeile die Kettenregel benutzt haben. Nach Satz 14.11 (Greenscher Integralsatz) ist dieses Integral gleich

$$\int_D \operatorname{rot} \left((P \circ F) \frac{\partial F_1}{\partial x_1}, (P \circ F) \frac{\partial F_1}{\partial x_2} \right) (x) \, dx. \quad (14.27)$$

Wir berechnen die skalare (zweidimensionale) Rotation nach (14.21):

$$\begin{aligned} &\operatorname{rot} \left((P \circ F) \frac{\partial F_1}{\partial x_1}, (P \circ F) \frac{\partial F_1}{\partial x_2} \right) \\ &= \frac{\partial}{\partial x_1} \left((P \circ F) \frac{\partial F_1}{\partial x_2} \right) - \frac{\partial}{\partial x_2} \left((P \circ F) \frac{\partial F_1}{\partial x_1} \right) \\ &= \frac{\partial(P \circ F)}{\partial x_1} \frac{\partial F_1}{\partial x_2} - \frac{\partial(P \circ F)}{\partial x_2} \frac{\partial F_1}{\partial x_1} + (P \circ F) \left(\frac{\partial^2 F_1}{\partial x_2 \partial x_1} - \frac{\partial^2 F_1}{\partial x_1 \partial x_2} \right), \end{aligned}$$

wobei wir die Produktregel benutzt haben. Der letzte Summand ist nach dem Satz von Schwarz gleich 0. Mit der Kettenregel erhalten wir für die ersten beiden

Summanden

$$\begin{aligned} & \left(\frac{\partial P}{\partial x_1} \frac{\partial F_1}{\partial x_1} + \frac{\partial P}{\partial x_2} \frac{\partial F_2}{\partial x_1} + \frac{\partial P}{\partial x_3} \frac{\partial F_3}{\partial x_1} \right) \frac{\partial F_1}{\partial x_2} - \\ & - \left(\frac{\partial P}{\partial x_1} \frac{\partial F_1}{\partial x_2} + \frac{\partial P}{\partial x_2} \frac{\partial F_2}{\partial x_2} + \frac{\partial P}{\partial x_3} \frac{\partial F_3}{\partial x_2} \right) \frac{\partial F_1}{\partial x_1} \\ & = \frac{\partial P}{\partial x_3} \left(\frac{\partial F_3}{\partial x_1} \frac{\partial F_1}{\partial x_2} - \frac{\partial F_3}{\partial x_2} \frac{\partial F_1}{\partial x_1} \right) - \frac{\partial P}{\partial x_2} \left(\frac{\partial F_1}{\partial x_1} \frac{\partial F_2}{\partial x_2} - \frac{\partial F_2}{\partial x_1} \frac{\partial F_1}{\partial x_2} \right). \end{aligned}$$

Sei $n = (n_1, n_2, n_3) = \frac{\partial F}{\partial x_1} \times \frac{\partial F}{\partial x_2} = \left(\frac{\partial F_1}{\partial x_1}, \frac{\partial F_2}{\partial x_1}, \frac{\partial F_3}{\partial x_1} \right) \times \left(\frac{\partial F_1}{\partial x_2}, \frac{\partial F_2}{\partial x_2}, \frac{\partial F_3}{\partial x_2} \right)$, d.h. $N = n/\|n\|$ ist der Normaleneinheitsvektor zu \mathcal{F} (vgl. (14.6)). Nach Definition des Vektorprodukts ist

$$n_2 = \frac{\partial F_3}{\partial x_1} \frac{\partial F_1}{\partial x_2} - \frac{\partial F_3}{\partial x_2} \frac{\partial F_1}{\partial x_1} \quad \text{und} \quad n_3 = \frac{\partial F_1}{\partial x_1} \frac{\partial F_2}{\partial x_2} - \frac{\partial F_2}{\partial x_1} \frac{\partial F_1}{\partial x_2}$$

und damit zusammengefasst

$$\text{rot} \left((P \circ F) \frac{\partial F_1}{\partial x_1}, (P \circ F) \frac{\partial F_1}{\partial x_2} \right) = \frac{\partial P}{\partial x_3} n_2 - \frac{\partial P}{\partial x_2} n_3. \quad (14.28)$$

Wegen $\text{rot } H = \text{rot} (P, 0, 0)^T = \left(0, \frac{\partial P}{\partial x_3}, -\frac{\partial P}{\partial x_2} \right)^T$ können wir (14.28) schreiben als

$$\text{rot} \left((P \circ F) \frac{\partial F_1}{\partial x_1}, (P \circ F) \frac{\partial F_1}{\partial x_2} \right) = ((\text{rot } H) \circ F) \cdot n.$$

Wir setzen dies in (14.27) ein und erhalten schließlich

$$\begin{aligned} \int_{\partial F} H \cdot dY &= \int_D \text{rot} \left((P \circ F) \frac{\partial F_1}{\partial x_1}, (P \circ F) \frac{\partial F_1}{\partial x_2} \right) dx \\ &= \int_D ((\text{rot } H) \circ F)(x) \cdot n \, dx \\ &= \int_{\mathcal{F}} \text{rot } H \cdot d\vec{\sigma} = \int_{\mathcal{F}} \text{rot } H \cdot N \, d\sigma \end{aligned}$$

nach der Definition 14.6 des Flächenintegrals. ■

Beispiel 2. Sei $D = \{(u, v) \in \mathbb{R}^2 : u \in [0, 2\pi], v \in [0, \pi/2]\}$, $r > 0$ und

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad F(u, v) = (r \cos u \cos v, r \sin u \cos v, r \sin v).$$

Das Flächenstück $F(D)$ ist die obere Halbkugelfläche um den Ursprung mit dem Radius r . Ein Vektorfeld H sei durch

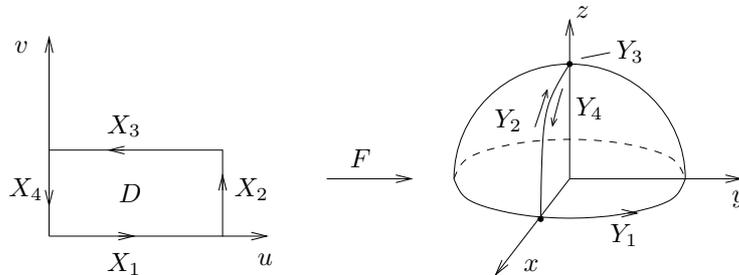
$$H : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad H(x, y, z) = (-y, x, 1)$$

gegeben. Der Rand ∂D des Rechtecks D läßt sich durch vier 4 Wege beschreiben:

$$\begin{aligned} X_1(t) &= (t, 0) & \text{mit } t \in [0, 2\pi], \\ X_2(t) &= (2\pi, t) & \text{mit } t \in [0, \pi/2], \\ X_3(t) &= (-t, \pi/2) & \text{mit } t \in [-2\pi, 0], \\ X_4(t) &= (0, -t) & \text{mit } t \in [-\pi/2, 0]. \end{aligned}$$

Die zugehörigen Parameterdarstellungen $Y_k(t) = F(X_k(t))$ für die Kurve $\partial F(D)$ sind

$$\begin{aligned} Y_1(t) &= (r \cos t, r \sin t, 0) & \text{mit } t \in [0, 2\pi], \\ Y_2(t) &= (r \cos t, 0, r \sin t) & \text{mit } t \in [0, \pi/2], \\ Y_3(t) &= (0, 0, r) & \text{mit } t \in [-2\pi, 0], \\ Y_4(t) &= (r \cos t, 0, -r \sin t) & \text{mit } t \in [-\pi/2, 0]. \end{aligned}$$



(Man beachte, dass $F(\partial D)$ *nicht* mit der Kreislinie in der xy -Ebene zusammenfällt!) Nun ist

$$\begin{aligned} \int_{\partial F(D)} H \cdot dY &= \int_0^{2\pi} H(Y_1(t)) \cdot \dot{Y}_1(t) dt + \int_0^{\pi/2} H(Y_2(t)) \cdot \dot{Y}_2(t) dt \\ &\quad + \int_{-2\pi}^0 H(Y_3(t)) \cdot \dot{Y}_3(t) dt + \int_{-\pi/2}^0 H(Y_4(t)) \cdot \dot{Y}_4(t) dt. \end{aligned}$$

Wir berechnen das erste Integral:

$$\begin{aligned} \int_0^{2\pi} H(Y_1(t)) \cdot \dot{Y}_1(t) dt &= \int_0^{2\pi} (-r \sin t, r \cos t, 1) \cdot (-r \sin t, r \cos t, 0) dt \\ &= \int_0^{2\pi} (r^2 \sin^2 t + r^2 \cos^2 t) dt = 2\pi r^2. \end{aligned}$$

Das dritte Integral ist wegen $\dot{Y}_3 = 0$ gleich 0, und das zweite und vierte Integral heben sich gegenseitig auf (dies ist sofort klar, wenn man sich die Wege Y_2, Y_3 und Y_4 ansieht). Also ist

$$\int_{\partial F(D)} H \cdot dY = 2\pi r^2.$$

Nach dem Satz von Stokes ist dann auch $\iint_{F(D)} \operatorname{rot} H \cdot N \, d\sigma = 2\pi r^2$, was wir zur Übung nachrechnen. Zunächst ist $\operatorname{rot} H = (0, 0, 2)$ und

$$N = (\cos u \cos v, \sin u \cos v, \sin v) \quad \text{sowie} \quad \|F_u \times F_v\| = r^2 \cos v$$

nach Beispiel 4 aus Abschnitt 14.1. Damit wird

$$\begin{aligned} \iint_{F(D)} \operatorname{rot} H \cdot N \, d\sigma &= \iint_D (\operatorname{rot} H)(F(u, v)) \cdot N(u, v) \|F_u \times F_v\| \, d(u, v) \\ &= \iint_D 2 \sin v \cdot r^2 \cos v \, d(u, v) \\ &= r^2 \int_0^{2\pi} \int_0^{\pi/2} \sin(2v) \, dv \, du = 2\pi r^2. \quad \blacksquare \end{aligned}$$

Auch der Satz von Stokes läßt sich unter schwächeren Voraussetzungen zeigen. Wir vermerken noch eine interessante Konsequenz des Stokes'schen Satzes.

Folgerung 14.13 *Sei B ein stückweise glatt berandeter Bereich im \mathbb{R}^3 und $H : B \rightarrow \mathbb{R}^3$ stetig differenzierbar. Dann ist*

$$\iint_{\partial B} \operatorname{rot} H \cdot N \, d\sigma = 0.$$

Kurz gesagt: *Der Wirbelfluß durch eine geschlossene Fläche ist Null.*

Zum Beweis schneidet man aus ∂B ein geeignetes kleines Flächenstück \mathcal{F} heraus. Auf der verbleibenden Fläche ist nach Stokes

$$\iint_{\partial B \setminus \mathcal{F}} \operatorname{rot} H \cdot N \, d\sigma = \int_{\partial \mathcal{F}} H \cdot dX.$$

Zieht man \mathcal{F} auf einen Punkt zusammen, so geht das Integral auf der rechten Seite gegen Null, da die Weglänge von $\partial \mathcal{F}$ gegen Null strebt. \blacksquare

14.7 Einige weitere Differential- und Integralformeln

14.7.1 Der Nabla-Operator

Der *symbolische* Vektor $\nabla := \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ heißt *Nabla-Operator*. Formal rechnet man mit ihm wie mit einem Vektor aus \mathbb{R}^3 . In diesem Sinne ist also für stetig differenzierbare Vektorfelder V und Skalarfelder φ

$$\begin{aligned} \nabla \varphi &= \operatorname{grad} \varphi, \\ \nabla \cdot F &= \operatorname{div} F, \\ \nabla \times F &= \operatorname{rot} F. \end{aligned}$$

Ist das Skalarfeld φ zweimal stetig differenzierbar, so erhält man

$$(\nabla \cdot \nabla)\varphi = \nabla \cdot \operatorname{grad} \varphi = \varphi_{xx} + \varphi_{yy} + \varphi_{zz}.$$

Der Operator

$$\Delta := \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

heißt *Laplace-Operator*.

14.7.2 Mehrfache Anwendungen der Differentialoperatoren

Das Vektorfeld F und das Skalarfeld φ seien zweimal stetig differenzierbar. Dann gilt

$$\operatorname{div} \operatorname{rot} F = 0, \quad (14.29)$$

$$\operatorname{rot} \operatorname{grad} \varphi = 0, \quad (14.30)$$

$$\operatorname{div} \operatorname{grad} \varphi = \Delta \varphi. \quad (14.31)$$

Man rechnet dies mit dem Satz von Schwarz leicht nach. Die ersten beiden Formeln besagen: *Wirbelfelder sind divergenzfrei* und *Gradientenfelder sind wirbelfrei*.

14.7.3 Produktregeln

Die Vektorfelder F, G und die Skalarfelder φ, ψ seien stetig differenzierbar. Dann gelten z.B. die folgenden Produktregeln, die man leicht nachrechnet:

$$\operatorname{grad}(\varphi\psi) = \varphi \operatorname{grad} \psi + \psi \operatorname{grad} \varphi, \quad (14.32)$$

$$\operatorname{div}(\varphi F) = \varphi \operatorname{div} F + F \cdot \operatorname{grad} \varphi, \quad (14.33)$$

$$\operatorname{rot}(\varphi F) = \varphi \operatorname{rot} F + \operatorname{grad} \varphi \times F, \quad (14.34)$$

$$\operatorname{div}(F \times G) = G \cdot \operatorname{rot} F - F \cdot \operatorname{rot} G. \quad (14.35)$$

Weitere Beziehungen finden Sie in der Literatur.

14.7.4 Die Greenschen Formeln

Es sei G wie im Gaußschen Integralsatz im Raum (Satz 14.8), und $f, g : G \rightarrow \mathbb{R}$ seien so oft stetig differenzierbar, wie es die folgenden Formeln verlangen. Aus Formel (14.33) (mit $F = \operatorname{grad} g$) erhalten wir

$$\operatorname{div}(f \operatorname{grad} g) = f \Delta g + \operatorname{grad} f \cdot \operatorname{grad} g,$$

wobei wir noch (14.31) benutzt haben. Integration über G und Anwendung des Gaußschen Integralsatzes auf der linken Seite liefern

$$\iint_{\partial G} f \operatorname{grad} g \cdot N \, d\sigma = \iiint_G (f \Delta g + \operatorname{grad} f \cdot \operatorname{grad} g) \, dx$$

mit dem äußeren Normalenvektor N an ∂G . Mit der Richtungsableitung

$$\frac{\partial g}{\partial N} = \text{grad } g \cdot N$$

(vgl. Abschnitt 10.4) erhalten wir die *erste Greensche Integralformel*

$$\iint_{\partial G} f \frac{\partial g}{\partial N} d\sigma = \iiint_G (f \Delta g + \text{grad } f \cdot \text{grad } g) dx. \quad (14.36)$$

Vertauscht man hierin f mit g und subtrahiert die erhaltene Formel von (14.36), so erhält man die *zweite Greensche Integralformel*

$$\iint_{\partial G} \left(f \frac{\partial g}{\partial N} - g \frac{\partial f}{\partial N} \right) d\sigma = \iiint_G (f \Delta g - g \Delta f) dx. \quad (14.37)$$

Im Spezialfall $g = 1$ folgt hieraus

$$\iint_{\partial G} \frac{\partial f}{\partial N} d\sigma = \iiint_G \Delta f dx. \quad (14.38)$$

Man kann somit Volumenintegrale über Laplacesche Differentialausdrücke Δf in Flächenintegrale umschreiben. Die Greenschen Formeln sind außerordentlich nützlich beim Studium von partiellen Differentialgleichungen.