ANALYSIS II

Ausarbeitung einer Vorlesung

vom Sommersemester 1991

überarbeitet im Sommersemester 2002

Joachim Weidmann

Fachbereich Mathematik

der Universität Frankfurt

Stand: 8. Juli 2002

Inhaltsverzeichnis

1 Fourierreihen

1.1 Vorbemerkungen

Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ heißt *p–periodisch*, wenn gilt

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

Offenbar gilt dann auch $f(x+kp) = f(x)$ für alle $k \in \mathbb{Z}$. Die kleinste Zahl p, für die dies gilt, wird als *die Periode* von f bezeichnet.

Ist f periodisch mit Periode p und $L > 0$, so ist

$$\tilde{f}(x) := f\left(\frac{p}{L}x\right)$$

periodisch mit Periode L. Es bedeutet deshalb keine wesentliche Einschränkung, wenn im folgenden nur 2π–periodische Funktionen betrachtet werden.

Offensichtlich ist jedes *trigonometrische Polynom*

$$\begin{aligned} f(x) \quad &:= \quad \frac{a_0}{2} + \sum_{k=1}^{n} \Big\{ a_k \cos(kx) + b_k \sin(kx) \Big\} \\ &= \quad \frac{1}{2}\Big\{ a_0 + \sum_{k=1}^{n} \Big[(a_k - ib_k)e^{ikx} + (a_k + ib_k)e^{-ikx} \Big] \Big\} \\ &= \quad \sum_{k=-n}^{n} c_k e^{ikx} \end{aligned}$$

2π–periodisch. Dabei gilt offenbar $c_0 = \frac{a_0}{2}$ und

$$c_k = \frac{1}{2}(a_k - ib_k) \text{ für } k > 0, \ = \frac{1}{2}(a_{-k} + ib_{-k}) \text{ für } k < 0,$$

$$a_k = c_k + c_{-k}, \ \ b_k = c_k - c_{-k} \ \text{ für } k > 0.$$

Die Periode ist genau dann 2π, wenn die Menge der k mit $a_k \neq 0$ oder $b_k \neq 0$ (bzw. $c_k \neq 0$ oder $c_{-k} \neq 0$) teilerfremd ist.

Ist f ein solches trigonometrisches Polynom, so sind die Koeffizienten a_k, b_k bzw. c_k eindeutig bestimmt durch

$$\begin{aligned}
a_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(kx)\,\mathrm{d}x \qquad \text{für} \quad k = 0,1,2,\ldots,n,\\
b_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(kx)\,\mathrm{d}x \qquad \text{für} \quad k = 1,2,\ldots,n,\\
c_k &= \frac{1}{2\pi}\int^{2\pi} f(x)e^{-ikx}\,\mathrm{d}x \qquad \text{für} \quad k = -n,-n+1,\cdots,n-1,n.
\end{aligned}$$

Diese Formeln ergeben sich sofort aus den folgenden Beziehungen:

1. $\displaystyle\int\limits_0^{2\pi} \cos(kx)\sin(k'x)\,\mathrm{d}x = 0$ für alle $k \in \mathbb{N}_0, k' \in \mathbb{N}$,
2. $\displaystyle\int\limits_0^{2\pi} \cos(kx)\cos(k'x)\,\mathrm{d}x = 0$ für $k, k' \in \mathbb{N}_0, k \neq k'$,
3. $\displaystyle\int\limits_0^{2\pi} \sin(kx)\sin(k'x)\,\mathrm{d}x = 0$ für $k, k' \in \mathbb{N}, k \neq k'$,
4. $\displaystyle\int\limits_0^{2\pi} [\sin(kx)]^2\,\mathrm{d}x = \pi$ für alle $k \in \mathbb{N}$,
5. $\displaystyle\int\limits_0^{2\pi} [\cos(kx)]^2\,\mathrm{d}x = \pi$ für alle $k \in \mathbb{N}$,
6. $\displaystyle\int\limits_0^{2\pi} [\cos(0x)]^2\,\mathrm{d}x = 2\pi$,

bzw.

7. $\displaystyle\int_0^{2\pi} e^{-ikx}e^{ik'x}\,\mathrm{d}x = \delta_{k.k'}2\pi$, wobei $\delta_{k,k'}$ das Kroneckersymbol ist ($= 1$ für $k = k'$, $\quad = 0$ für $k \neq k'$).

Beweis. 1. Das Integral ändert sich nicht, wenn über (π,π) statt $(0,2\pi)$ integriert wird. Da der Integrand ungerade ist, folgt die Behauptung.

2. Auf Grund des Additionstheorems für den Cosinus ist der Integrand gleich

$$\frac{1}{2}\Big\{\cos(k+k')x + \cos(k-k')x\Big\}.$$

Wegen $k + k' \neq 0$ und $k - k' \neq 0$ verschwindet also das Integral.

Ebenso erhält man in den Fällen 3., 4. und 5. für den Integranden

$$\frac{1}{2}\{\cos(k-k')x - \cos(k+k')x\},$$

$$\frac{1}{2}\{\cos(0x) - \cos(2kx)\}, \textit{ bzw.}$$

$$\frac{1}{2}\{\cos(2kx) + \cos(0x)\}$$

und somit für die Integrale $0, \pi$ und π. Die Aussage 6 ist offensichtlich. Die Aussage 7 ergibt sich durch einfache Integration. ■

Da die Darstellung in der Form $\sum c_k e^{ikx}$ einfacher ist, arbeiten wir ab jetzt mit dieser weiter (natürlich lassen sich alle Aussagen sofort auf die andere Darstellung übertragen).

Die obige Aussage gilt offenbar auch, wenn f die Form

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \Big\{a_k \cos(kx) + b_k \sin(kx)\Big\} = \sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

hat, wobei die Reihe gleichmäßig konvergiert (Beweis!).

Wir stellen uns die *Frage*: Hat jede 2π–periodische Funktion die Form $\sum_{k=-n}^{n} c_k e^{ikx}$, oder kann sie durch solche Funktionen beliebig gut approximiert werden? Was heißt hier „jede" Funktion, und in welchem Sinn soll approximiert werden?

1.2 Konvergenz im Quadratmittel

Im folgenden sei R der Raum der 2π–periodischen Funktionen $f : \mathbb{R} \to \mathbb{R}$, die über $[0, 2\pi]$ Riemann–integrierbar sind (diese sind dann natürlich über jedes beschränkte Intervall Riemann–integrierbar). Für jedes $f \in R$ sind dann die *Fourierkoeffizienten* von f

$$c_k = c_k(f) := \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x)\,\mathrm{d}x, \qquad k \in \mathbb{Z}$$

wohldefiniert. Wir können also die *(formale) Fourierreihe*

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

angeben. Im folgenden sei

$$s_n(x) = s_{f,n}(x) := \sum_{k=-n}^{n} c_k e^{ikx}, \qquad n \in \mathbb{N}_0$$

die n–te Partialsumme der Fourierreihe von f. Es stellt sich also die

Frage: Können Bedingungen an $f \in \mathbb{R}$ angegeben werden, unter denen gilt

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \qquad \text{d. h.} \quad s_n(x) \longrightarrow f(x)?$$

Dabei kann man an gleichmäßige Konvergenz denken, an punktweise Konvergenz, oder an eine u. U. noch schwächere Konvergenz, z. B. die im folgenden erklärte Konvergenz im Quadratmittel.

Wir sagen: Eine Folge (f_n) aus R *konvergiert im Quadratmittel* (oder auch *konvergiert im L^2-Sinn*) gegen eine Funktion $f \in R$, wenn gilt

$$\int_0^{2\pi} |f_n(x) - f(x)|^2 \, \mathrm{d}x \to 0.$$

(Natürlich genügt es hierfür, wenn f_n und f auf $[0, 2\pi]$ erklärt und dort Riemann–integrierbar sind; man kann diese dann 2π–periodisch auf ganz $\mathbb{R}$ fortsetzen.)

Durch

$$\langle f, g\rangle := \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) \, \mathrm{d}x \qquad \text{und} \qquad \|f\| := \langle f, f\rangle^{1/2}$$

sind ein *Semikalarprodukt* und eine *Halbnorm* auf R erklärt (Skalaprodukt bzw. Norm, wenn man Funktionen f und g mit $\int |f - g| \, \mathrm{d}x = 0$ identifiziert; Beweis!) und es gilt offensichtlich: (f_n) konvergiert genau dann im Quadratmittel gegen f, wenn (f_n) im Sinne dieser (Halb–) Norm gegen f konvergiert, d. h. $\|f_n - f\| \to 0$.

Bemerkung Aus der gleichmäßigen Konvergenz folgt natürlich die Konvergenz im Quadratmittel. Die Umkehrung gilt nicht. Zum Beispiel konvergiert die Folge (f_n) mit

$$f_n(x) = \begin{cases} 0 & \text{für} \quad 0 \leq x \leq 2\pi - \frac{1}{n}, \\ n(x + \frac{1}{n} - 2\pi) & \text{für} \quad 2\pi - \frac{1}{n} \leq x \leq 2\pi, \end{cases} \qquad (2\pi\text{–periodisch})$$

im Quadratmittel gegen 0, nicht aber gleichmäßig (vgl. Aufgabe 1.3). Aus der Konvergenz im Quadratmittel folgt aber auch nicht die punktweise Konvergenz (vgl. Aufgabe 1.4).

Im folgenden seien die Funktionen e_k definiert durch

$$e_k : \mathbb{R} \to \mathbb{C}, \quad e_k(x) = e^{ikx} \qquad \text{für} \quad k \in \mathbb{Z},\ x \in \mathbb{R}.$$

Dann gilt offenbar

$$c_k = c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x)\,\mathrm{d}x = \langle e_k, f \rangle$$

und (vgl. obige Eigenschaft 7)

$$\langle e_j, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-j)x}\,\mathrm{d}x = \delta_{jk},$$

d. h. $\{e_k : k \in \mathbb{Z}\}$ ist ein *Orthonormalsystem* bezüglich des Skalarprodukts $\langle \cdot, \cdot \rangle$.

Satz 1.1 (Entwicklungssatz) *Sei $f \in R$; c_k und s_n seien wie oben erklärt. Dann gilt:*

a) $\|f - s_n\|^2 = \|f\|^2 - \sum_{k=-n}^{n} |c_k|^2$,

b) $\sum_{k=-\infty}^{\infty} |c_k|^2 \le \|f\|^2$ (Besselsche Ungleichung),

insbesondere gilt das Riemann–Lebsgue–Lemma*: $c_k \to 0$ für $k \to \infty$.*

c) (s_n) *konvergiert genau dann im Quadratmittel gegen f, wenn gilt*

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \|f\|^2 \qquad \text{(Parsevalsche Gleichung).}$$

(Die Gültigkeit der Parsevalsche Gleichung*, für Treppenfunktionen $f \in R$ wird in Hilfssatz 1.6 bewiesen wird.)*

Beweis. a) Man rechnet leicht nach:

$$\begin{aligned}
\|f - s_n\|^2 &= \frac{1}{2\pi}\int_0^{2\pi} \left| f(x) - \sum_{k=-n}^{n} c_k e^{ikx} \right|^2 \mathrm{d}x \\
&= \frac{1}{2\pi}\left\{ \int_0^{2\pi} |f(x)|^2 \,\mathrm{d}x - \sum_{k=-n}^{n} c_k \int_0^{2\pi} \overline{f(x)} e^{ikx} \,\mathrm{d}x \right. \\
&\qquad \left. - \sum_{k=-n}^{n} \overline{c_k} \int_0^{2\pi} f(x) e^{-ikx} \,\mathrm{d}x + 2\pi \sum_{k=-n}^{n} |c_k|^2 \right\} \\
&= \|f\|^2 - \sum_{k=-n}^{n} \left\{ c_k\overline{c_k} + \overline{c_k} c_k - |c_k|^2 \right\} = \|f\|^2 - \sum_{k=-n}^{n} |c_k|^2.
\end{aligned}$$

b) Dies folgt aus Teil a, da stets $\|f - s_n\|^2 \geq 0$ gilt.

c) Dies folgt direkt aus Teil a. ∎

Korollar 1.2 *Für die Koeffizienten a_k und b_k ergibt sich hieraus*

$$\|f\|^2 \geq \left|\frac{a_0}{2}\right|^2 + \frac{1}{2}\sum_{k=1}^{\infty} \left(|a_k|^2 + |b_k|^2\right),$$

bzw. im Fall der Gültigkeit der Parsevalschen Gleichung auch hier die Gleichheit.

Der *Beweis* ergibt sich aus $c_0 = a_0/2,\ c_{\pm k} = (a_k \pm ib_k)/2$.

Es ist unser Ziel zu zeigen, daß für jedes $f \in R$ die Folge (s_n) im Quadratmittel gegen f konvergiert. Teil c) von Satz 1.1 sagt aus, daß hierfür die Parsevalsche Gleichung zu beweisen ist, was wir in Hilfssatz 1.6 zunächst für Treppenfunktionen tun werden. Dazu werden noch drei Hilfssätze benötigt.

Hilfssatz 1.3 *Für alle $t \in \mathbb{R}\backslash\{2k\pi : k \in \mathbb{Z}\}$ gilt*

$$\sum_{k=1}^{n} \cos(kt) = \frac{\sin(n+\frac{1}{2})t}{2\sin(\frac{1}{2}t)} - \frac{1}{2}.$$

Beweis. Aus $cos(kt) = \frac{1}{2}(e^{ikt} + e^{-ikt})$ folgt mit Hilfe der Summenformel für die geometrische Reihe

$$\begin{aligned}\frac{1}{2} + \sum_{k=1}^{n} \cos(kt) &= \frac{1}{2}\sum_{k=-n}^{n} e^{ikt} = \frac{1}{2}e^{-int}\sum_{k=0}^{2n} e^{ikt} \\ &= \frac{1}{2}e^{-int}\frac{1-e^{(2n+1)it}}{1-e^{it}} = \frac{1}{2}\frac{e^{-i(n+\frac{1}{2})t} - e^{i(n+\frac{1}{2})t}}{e^{-i\frac{t}{2}} - e^{i\frac{t}{2}}};\end{aligned}$$

das ist die Behauptung. ■

Hilfssatz 1.4 *Es gilt*

$$\frac{\pi - x}{2} = \sum_{k=1}^{\infty} \frac{\sin(kx)}{k} \qquad \text{für} \quad 0 < x < 2\pi;$$

für jedes $\delta > 0$ ist dabei die Konvergenz gleichmäßig in $[\delta, 2\pi - \delta]$.

Beweis. Aus Hilfssatz 1.3 folgt für $x \in (0, 2\pi)$ und $n \in \mathbb{N}$

$$\begin{aligned}\sum_{k=1}^{n} \frac{\sin(kx)}{k} &= \sum_{k=1}^{n} \int_{\pi}^{x} \cos(kt)\,\mathrm{d}t = \int_{\pi}^{x} \sum_{k=1}^{n} \cos(kt)\,\mathrm{d}t \\ &= \int_{\pi}^{x} \frac{\sin(n+\frac{1}{2})t}{2\sin(\frac{1}{2}t)}\,\mathrm{d}t - \frac{1}{2}(x-\pi) =: F_n(x) + \frac{\pi - x}{2}\end{aligned}$$

mit

$$\begin{aligned}F_n(x) &= \int_{\pi}^{x} \frac{1}{2\sin(\frac{1}{2}t)} \sin(n+\frac{1}{2})t\,\mathrm{d}t \\ &= -\frac{1}{2\sin(\frac{1}{2}t)} \frac{\cos(n+\frac{1}{2})t}{n+\frac{1}{2}}\Bigg|_{\pi}^{x} - \frac{1}{n+\frac{1}{2}} \int_{\pi}^{x} \frac{\cos(\frac{1}{2}t)}{4\sin^2(\frac{1}{2}t)} \cos(n+\frac{1}{2})t\,\mathrm{d}t.\end{aligned}$$

Für jedes $\delta > 0$ konvergiert $F_n(\cdot)$ in $[\delta, 2\pi - \delta]$ gleichmäßig gegen 0. ■

Hilfssatz 1.5 *Es gilt*

$$\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} = \frac{(x-\pi)^2}{4} - \frac{\pi^2}{12} \qquad \text{für} \quad 0 \le x \le 2\pi.$$

Speziell erhält man für $x = 0$: $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$.

Beweis. Aus Hilfssatz 1.4 folgt für $x, y \in (0, 2\pi)$ durch Integration

$$\begin{aligned}\frac{(x-\pi)^2}{4} - \frac{(y-\pi)^2}{4} &= \int_y^x \frac{t-\pi}{2}\,\mathrm{d}t = -\int_y^x \sum_{k=1}^{\infty} \frac{\sin(kt)}{k}\,\mathrm{d}t \\ &= \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} - \sum_{k=1}^{\infty} \frac{\cos(ky)}{k^2},\end{aligned}$$

also

$$\frac{(x-\pi)^2}{4} = \sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} + c.$$

Aus

$$\frac{\pi^3}{6} = \int_0^{2\pi} \frac{(x-\pi)^2}{4}\,\mathrm{d}x = \int_0^{2\pi} \left(\sum_{k=1}^{\infty} \frac{\cos(kx)}{k^2} + c \right) \mathrm{d}x = 2\pi c$$

folgt $c = \pi^2/12$ und damit die obige Behauptung für alle $x \in (0, 2\pi)$. Für $x = 0$ und $x = 2\pi$ folgt die Behauptung durch Grenzübergang, da beide Seiten stetig sind. ■

Hilfssatz 1.6 *Ist $f \in R$ so, daß die Einschränkung von f auf $[0, 2\pi]$ eine Treppenfunktion ist, so gilt $s_n \to f$ im Quadratmittel.*

Beweis. a) Wir betrachten zunächst den Spezialfall

$$f_a(x) := \begin{cases} 1 & \text{für } 0 \le x \le a, \\ 0 & \text{für } a < x < 2\pi. \end{cases}$$

In diesem Fall berechnet man für $c_k = c_k(f_a)$:

$$\begin{aligned} c_0 &= \frac{a}{2\pi}, \\ c_k &= \frac{1}{2\pi} \int_0^a e^{-ikx}\,\mathrm{d}x = \frac{i}{2\pi k}(e^{-ika} - 1) \qquad \text{für} \quad k \neq 0, \\ |c_k|^2 &= \frac{1}{4\pi^2 k^2}(1 - e^{ika})(1 - e^{-ika}) = \frac{1 - \cos(ka)}{2\pi^2 k^2} \qquad \text{für} \quad k \neq 0 \end{aligned}$$

und somit (unter Verwendung von Hilfssatz 1.5)

$$\begin{aligned}\sum_{k=-\infty}^{\infty} |c_k|^2 &= \frac{a^2}{4\pi^2} + \sum_{k=1}^{\infty} \frac{1-\cos(ka)}{\pi^2 k^2} = \frac{a^2}{4\pi^2} + \frac{1}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2} - \frac{1}{\pi^2}\sum_{k=1}^{\infty}\frac{\cos(ka)}{k^2} \\ &= \frac{a^2}{4\pi^2} + \frac{1}{6} - \frac{1}{\pi^2}\left\{\frac{(\pi-a)^2}{4} - \frac{\pi^2}{12}\right\} = \frac{a}{2\pi} \\ &= \frac{1}{2\pi}\int_0^{2\pi} |f_a(x)|^2\,\mathrm{d}x \;=\; \|f_a\|^2.\end{aligned}$$

Mit Teil c von Satz 1.1 folgt die Behauptung in diesem Spezialfall.

b) Sei jetzt $f \in R$ und $f|_{[0,2\pi]}$ eine Treppenfunktion. Dann läßt sich f schreiben in der Form $f = \sum_{j=1}^{\ell} d_j f_j$, wobei die f_j von der in Teil a des Beweises betrachteten Gestalt sind. Sind $s_{j,n}$ die Partialsummen der zu f_j gehörigen Fourierreihen, so gilt nach Teil a

$$\|f_j - s_{j,n}\| \to 0 \text{für } n \to \infty,\ j = 1, \ldots, \ell$$

und somit aus der Dreiecksungleichung für $\|\cdot\|$

$$\Big\|f - s_n\Big\| = \Big\|\sum_{j=1}^{\ell}(d_j f_j - d_j s_{j,n})\Big\| \le \sum_{j=1}^{\ell}|d_j|\Big\|f_j - s_{j,n}\Big\| \to 0,$$

d. h. es gilt $s_n \to f$ im Quadratmittel. ■

Satz 1.7 *Für jedes $f \in R$ konvergiert die Fourierreihe im Quadratmittel gegen f. Insbesondere gilt die* Parsevalsche Gleichung

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\,\mathrm{d}x.$$

Beweis. Da $f \in R$ ist, ist es beschränkt, d. h. es existiert ein $c \ge 0$ mit

$$|f(x)| \le c \qquad \text{für alle} \quad x \in R,$$

und zu jedem $\varepsilon > 0$ existiert ein $f_\varepsilon \in R$ so, daß $f_\varepsilon|_{[0,2\pi]}$ eine Treppenfunktion ist und

$$\int_0^{2\pi} \Big| f_\varepsilon(x) - f(x) \Big| \, \mathrm{d}x \leq \frac{1}{2c}\varepsilon^2;$$

dabei kann o. E. angenommen werden, daß auch $|f_\varepsilon(x)| \leq c$ gilt. Dann folgt

$$\int_0^{2\pi} \Big| f_\varepsilon(x) - f(x) \Big|^2 \, \mathrm{d}x \leq 2c \int_0^{2\pi} \Big| f_\varepsilon(x) - f(x) \Big| \, \mathrm{d}x \leq \varepsilon^2.$$

Sind s_n und $s_{\varepsilon,n}$ die Partialsummen der Fourierreihen zu f bzw. f_ε und $s_{g,n}$ die zu $g := f - f_\varepsilon$, so gilt

$$f = f_\varepsilon + g \quad \text{und} \quad s_n = s_{\varepsilon_n} + s_{g_n},$$

und somit

$$\begin{aligned} \|f - s_n\| &= \|f_\varepsilon - s_{\varepsilon,n} + g - s_{g,n}\| \leq \|f_\varepsilon - s_{\varepsilon,n}\| + \|g - s_{g,n}\| \\ & \text{(nach Satz 1.1)} \\ &\leq \|f_\varepsilon - s_{\varepsilon,n}\| + \|g\| \\ &\leq \|f_\varepsilon - s_{\varepsilon,n}\| + \varepsilon \ \leq \ 2\varepsilon \quad \text{für große} \ \ n. \end{aligned}$$

Das ist die Konvergenz $s_n \to f$ im Quadratmittel, woraus die anderen Aussagen mit Satz 1.1 folgen. ■

1.3 Gleichmäßige und punktweise Konvergenz

Der Satz 1.7 ist zwar sehr allgemein gültig, liefert allerdings nur eine „recht schwache“Kon ver genz aussage. Gleichmäßige Konvergenz wäre schöner. Natürlich kann man das nur für stetige Funktionen aus R erwarten; es ist aber bekannt, daß dies auch für stetige Funktionen i. allgem. nicht gilt. (Es gibt sogar stetige Funktionen, deren Fourierreihen nicht überall konvergieren; Konstruktionen gehen auf P. du Bois Raymond, H. Lebesgue, H. A. Schwarz, A. Haar und L. Fejer zurück [vgl. z. B. A. Zygmund: Trigonometric series **I**, VIII.1]. Andererseits weiß man nach L. Carleson, Acta Mathematica **116** (1966), daß die Fourierreihe jeder stetigen (sogar jeder L_2-) Funktion „fast überall“konvergiert.) Für etwas „glattere“Funktionen können wir allerdings gleichmäßige Konvergenz sehr leicht beweisen.

Satz 1.8 *Sei $f : \mathbb{R} \to \mathbb{C}$ 2π–periodisch, stetig und stückweise stetig differenzierbar (d. h. es gibt eine Unterteilung $0 = t_0 < t_1 < \ldots < t_r = 2\pi$ von $[0, 2\pi]$ so, daß $f|_{[t_{j-1},t_j]}$ stetig differenzierbar ist). Dann konvergiert die Fourierreihe von f gleichmäßig gegen f.*

Beweis. Sei

$$\varphi : \mathbb{R} \to \mathbb{C}, \quad \varphi(x) := \begin{cases} f'(x) & \text{für} \quad x \neq t_j + 2k\pi, \\ 0 & \text{für} \quad x = t_j + 2k\pi. \end{cases}$$

Dann ist offensichtlich $\varphi \in R$. Seien γ_k die Fourierkoeffizienten von φ. Nach Satz 1.7 gilt

$$\sum_{k=-\infty}^{\infty} |\gamma_k|^2 = \|\varphi\|^2 < \infty.$$

Für die Fourierkoeffizienten c_k von f gilt

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} \,\mathrm{d}x = \frac{1}{2\pi} \sum_{j=1}^{r} \int_{t_{j-1}}^{t_j} f(x) e^{-ikx} \,\mathrm{d}x \\ &= \frac{i}{2k\pi} \sum_{j=1}^{r} f(x) e^{-ikx} \Big|_{t_{j-1}}^{t_j} - \frac{i}{2k\pi} \sum_{j=1}^{r} \int_{t_{j-1}}^{t_j} \varphi(x) e^{-ikx} \,\mathrm{d}x \\ &= 0 - \frac{i}{2k\pi} \int_0^{2\pi} \varphi(x) e^{-ikx} \,\mathrm{d}x = -\frac{i}{k} \gamma_k \end{aligned}$$

und somit (wegen $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$)

$$|c_k| \leq \frac{1}{2}(k^{-2} + |\gamma_k|^2),$$

d. h. $\sum |c_k|$ ist konvergent und somit

$$\sum c_k e^{ikx}$$

gleichmäßig konvergent gegen eine stetige Funktion g.

Wegen

$$\|f - g\| \leq \|f - s_{f,n}\| + \|s_{f,n} - g\| \longrightarrow 0$$

ist $\|f - g\| = 0$. Da f und g stetig sind, folgt hieraus $f(x) = g(x)$ für alle x. ∎

Beispiel 1.9 Sei

$$f(x) = |x| \qquad \text{für} \quad |x| \le \pi, \qquad 2\pi - \text{periodisch fortgesetzt.}$$

Nach Satz 1.8 konvergiert in diesem Fall die Fourierreihe gleichmäßig (dies werden wir auch gleich an den Fourierkoeffizienten sehen).

Da f *gerade* ist, haben wir natürlich eine reine *Cosinusreihe* (Beweis!)

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx)$$

mit

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_{-\pi}^{\pi} |x|\,\mathrm{d}x = \pi, \\ a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} |x|\cos(kx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\cos(kx)\,\mathrm{d}x \\ &= \frac{2}{\pi}\left\{\frac{1}{k}x\sin(kx)\Big|_0^{\pi} - \frac{1}{k}\int_0^{\pi}\sin(kx)\,\mathrm{d}x\right\} = \frac{2}{\pi}k^{-2}\cos(kx)\Big|_0^{\pi} \\ &= \begin{cases} 0 & \text{falls } k \text{ gerade ist,} \\ -4(k^2\pi)^{-1} & \text{falls } k \text{ ungerade ist,} \end{cases} \end{aligned}$$

also

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos(2n-1)x}{(2n-1)^2}.$$

Für $|x| \le \pi$ gilt somit

$$\frac{\pi^2}{8} - \frac{\pi}{4}|x| = \cos x + \frac{\cos(3x)}{3^2} + \frac{\cos(5x)}{5^2} + \dots$$

und somit speziell für $x = 0$

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

eine überraschende Formel für π^2. □

Wenn f nicht stetig ist, kann man zwar keine gleichmäßige Konvergenz der Fourierreihe mehr erwarten, eventuell aber wenigstens punktweise Konvergenz in gewissen Regularitätspunkten.

Satz 1.10 *Sei $f \in R$ und es existieren*

$$\begin{aligned} f_\pm(x_0) &= \lim_{h\to 0\pm} f(x_0+h), \\ f'_\pm(x_0) &= \lim_{h\to 0\pm} \frac{1}{h}\Big(f(x_0+h) - f_\pm(x_0)\Big). \end{aligned}$$

Dann gilt

$$s_n(x_0) \longrightarrow \frac{1}{2}\left(f_+(x_0) + f_-(x_0)\right).$$

Beweis. Man berechnet

$$\begin{aligned} s_n(x_0) = &\sum_{k=-n}^{n} e^{ikx_0} \frac{1}{2\pi} \int_0^{2\pi} e^{-iky} f(y)\,\mathrm{d}y \\ = &\quad \frac{1}{2\pi} \int_{x_0-\pi}^{x_0+\pi} \sum_{k=-n}^{n} e^{ik(x_0-y)} f(y)\,\mathrm{d}y \\ = &\quad \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k=-n}^{n} e^{-iky} f(x_0+y)\,\mathrm{d}y \\ = &\quad \frac{1}{2\pi} \int_0^{\pi} \sum_{k=-n}^{n} e^{-iky} (f(x_0+y) + f(x_0-y))\,\mathrm{d}y \\ &\quad \left(\text{mit } \sum_{k=-n}^{n} e^{-iky} = 2\sum_{k=-n}^{n} \cos ky = \frac{\sin(n+\frac{1}{2})y}{\sin\frac{1}{2}y}, \text{ vgl. Hilfssatz 1.3}\right) \\ = &\quad \frac{1}{2\pi} \left\{ \int_0^{\pi} 2\sum_{k=0}^{n} \cos ky \Big(f_+(x_0) + f_-(x_0)\Big)\,\mathrm{d}y \right. \\ &\quad \left. + \int_0^{\pi} \sin\left[\left(n+\frac{1}{2}\right)y\right] \underbrace{\frac{f(x_0+y) - f_+(x_0) + f(x_0-y) - f_-(x_0)}{\sin(\frac{1}{2}y)}}_{=g(y)}\,\mathrm{d}y \right\} \\ = &\quad \frac{1}{2\pi} \left\{ \int_0^{\pi} \{f_+(x_0) + f_-(x_0)\}\,\mathrm{d}y \right. \end{aligned}$$

$$
\begin{aligned}
&+ \int_0^\pi \frac{1}{2i} \left\{ e^{iny} e^{\frac{i}{2}y} - e^{-iny} e^{-\frac{i}{2}y} \right\} g(y)\,\mathrm{d}y \Bigg\} \\
\longrightarrow\ & \frac{1}{2}\{f_+(x_0) + f_-(x_0)\},
\end{aligned}
$$

wobei das Riem-Lebesgue-Lemma benutzt wurde, und daß auf Grund der Voraussetzung die 2π–periodische Funktion

$$
h(y) := \begin{cases} \exp(\pm\frac{i}{2}y)g(y) & \text{für } y \in [0, 2\pi], \\ 0 & \text{für } y \in (-\pi, 0) \end{cases}
$$

in R liegt. ∎

Beispiel 1.11 Die *Sägefunktion* ist definiert durch

$$
f(x) = x \quad in \quad (-\pi, \pi], \qquad 2\pi - \text{periodisch fortgesetzt auf} \mathbb{R}.
$$

Es ist dann

$$
\begin{aligned}
c_k &= \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x)\,\mathrm{d}x \quad = \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-ikx}\,\mathrm{d}x \\
&= \begin{cases} 0 & \text{für} \quad k = 0 \\ \frac{i}{2k\pi} x e^{-ikx}\big|_{-\pi}^{\pi} - \frac{i}{2k\pi} \int_{-\pi}^{\pi} e^{-ikx}\,\mathrm{d}x \quad = \quad (-1)^k \frac{i}{k} & \text{für} \quad k \neq 0. \end{cases}
\end{aligned}
$$

Für die Fourierreihe der Sägefunktion erhalten wir also

$$
\sum_{k=1}^{\infty} (-1)^k \frac{i}{k} \left(e^{ikx} - e^{-ikx} \right) = 2 \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} \sin kx,
$$

eine reine Sinusreihe, da f ungerade ist. □

1.4 Die Sätze von Weierstraß

Für beliebige stetige 2π–periodische Funktionen kann man zwar nicht die gleichmäßige Konvergenz der Fourierreihe beweisen, trotzdem läßt sich aber jede derartige Funktion gleichmäßig durch trigonometrische Polynome approximieren.

Satz 1.12 (Weierstraßscher Appoximationssatz, periodisch)
Sei $f \in R$ stetig. Dann gibt es zu jedem $\varepsilon > 0$ ein trigonometrisches Polynom p mit

$$|p(x) - f(x)| < \varepsilon \quad \text{für alle} \quad x \in \mathbb{R}.$$

In anderen Worten: Es gibt eine Folge (p_n) von trigonometrischen Polynomen mit $p_n \longrightarrow f$ gleichmäßig.

Beweis. Da f stetig ist, ist es gleichmäßig stetig und es gibt zu jedem $\varepsilon > 0$ eine stetige stückweise lineare Funktion f_ε mit $|f(x) - f_\varepsilon(x)| < \frac{\varepsilon}{2}$ für alle $x \in \mathbb{R}$ (Beweis!). Nach Satz 1.8 konvergiert die Fourierreihe von f_ε gleichmäßig gegen f_ε. Es gibt also ein trigonometrisches Polynom p mit $|p(x) - f_\varepsilon(x)| < \frac{\varepsilon}{2}$ für alle $x \in \mathbb{R}$ und somit $|p(x) - f(x)| < \varepsilon$ für alle $x \in \mathbb{R}$. ■

Satz 1.13 (Weierstraßscher Approximationssatz) *Sei $f : [a, b] \longrightarrow \mathbb{C}$ stetig. Dann gibt es zu jedem $\varepsilon > 0$ ein Polynom p mit $|f(x) - p(x)| < \varepsilon$ für alle $x \in [a, b]$. In anderen Worten: Es gibt eine Folge (p_n) von Polynomen mit $p_n \longrightarrow f$ gleichmäßig in $[a, b]$*

Beweis. Ohne Einschränkung nehmen wir an, daß $[a, b] = [0, 1]$ gilt (andernfalls betrachtet man $\hat{f}(x) := f(a + x(b - a))$). Indem man definiert

$$f(x) = f(1) + \frac{x - 1}{2\pi - 1}(f(0) - f(1)) \quad \text{für} \quad x \in (1, 2\pi),$$

kann f zu einer stetigen 2π–periodischen Funktion g fortgesetzt werden. Nach Satz 1.12 gibt es ein trigonometrisches Polynom $q(x) = \sum_{k=-n}^{n} c_k e^{ikx}$ mit

$$|g(x) - q(x)| \leq \frac{\varepsilon}{2} \quad \text{für alle} \quad x \in \mathbb{R}.$$

Jeder Term $c_k e^{ikx}$ läßt sich auf $[0, 1]$ bis auf $\varepsilon/4n$ durch ein Polynom p_k approximieren (Taylorreihe); exakt für $k = 0$. Mit $p(x) = \sum_{k=-n}^{n} p_k(x)$ gilt also

$$|f(x) - p(x)| \leq |g(x) - q(x)| + |q(x) - p_k(x)| \leq \frac{\varepsilon}{2} + \sum_{k=-n}^{n} |c_k e^{ikx} - p(x)| \leq \varepsilon,$$

d. h. f wird durch Polynome beliebig gut approximiert. ■

1.5 Übungsaufgaben

1.1 Man berechne die Fourierreihe der Funktion $f(x) := |\sin x|$.
Anleitung: Man beachte, daß $f(x)$ und $\cos kx$ gerade und $\sin kx$ ungerade Funktionen sind.

1.2 Ein trigonometrisches Polynom $\sum_{k=-n}^{n} c_k e^{ikx}$ hat genau dann die (minimale) Periode 2π, wenn die Menge $\{k \in \mathbb{N} : |c_k| + |c_{-k}| \neq 0\}$ teilerfremd ist.

1.3 Man zeige, daß die Funktionenfolge

$$f_n(x) = \begin{cases} 0 & \text{für } 0 \le x \le 2\pi - 1/n, \\ n(x + 1/n - 2\pi) & \text{für } 2\pi - 1/n \le x \le 2\pi \end{cases}$$

im Quadratmittel konvergiert, aber nicht gleichmäßig.

1.4 Man konstruiere eine Folge (f_n) von Funktionen auf $[o, 2\pi)]$, die aus immer schmaler werdenden „wandernden Buckeln" bestehen so, daß (f_n) im Quadratmittel (gegen 0) konvergiert, aber nicht punktweise konvergent ist.

1.5 Sei

$$f(x) = \begin{cases} 0 & \text{für } 0 \le x < \pi, \\ 1 & \text{für } \pi \le x < 2\pi, \end{cases} \qquad 2\pi\text{–periodisch}$$

a) Man bestimme die Fourierreihe von f (im wesentlichen eine reine Sinusreihe).

b) Mit Hilfe der Parsevalschen Gleichung zeige man

$$\frac{\pi^2}{8} = \sum_{k=1}^{\infty} (2k-1)^{-2}.$$

c) Die Summe *aller* k^{-2} läßt sich daraus berechnen.

1.6 Man beweise $\sum_{k=1}^{\infty} \frac{1}{k^6} = \dfrac{\pi^6}{945}$.
Anleitung: Man integriere $\sum_{k=1}^{\infty} k^{-2} \cos kt = \dfrac{(t-\pi)^2}{4} - \dfrac{\pi^2}{12}$ von π bis x und wende die Vollständigkeitsrelation an.

1.7 Sei $f : [1, \infty) \longrightarrow \mathbb{C}$ stetig und $\lim_{x\to\infty} f(x)$ existiere. Dann gibt es für jedes $\varepsilon > 0$ ein $n \in \mathbb{N}$ und Konstanten $d_0, d_1, \ldots, d_n$ mit

$$|f(x) - \sum_{j=0}^{n} d_j x^{-j}| < \varepsilon \qquad \text{für alle} \quad x \in [1, \infty).$$

1.8 Zu jedem $f \in C[0, 1]$ gibt es eine Folge gerader Polynome, die gleichmäßig auf $[0, 1]$ gegen f konvergiert.

1.9 Direkter Beweis des Riemannschen Lemmas: Für f mit $f(x) = 1$ in $[a, b]$ und $f(x) = 0$ sonst berechne man $c_k(f)$ und zeige $c_k(f) \longrightarrow 0$ für $k \longrightarrow \infty$. Damit beweise man $c_k(f) \longrightarrow 0$ zunächst für Funktionen aus R, für die $f|_{[0,2\pi)}$ Treppenfunktionen sind und dann für alle $f \in R$.

2 Stetige Funktionen $f : \mathbb{R}^m \longrightarrow \mathbb{R}^n$

2.1 Vorbemerkungen

Es soll zunächst mit einigen Beispielen belegt werden, daß es notwendig ist, nicht nur Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ zu betrachten, sondern auch solche, die auf einen höherdimensionalen Raum $\mathbb{R}^m$ definiert sind und/oder ihre Werte in einem höherdimensionalen Raum $\mathbb{R}^n$ annehmen; dabei wird nur in Spezialfällen $n = m$ sein.

Beispiel 2.1 a) Die *Bewegung eines Körpers* (Massenpunktes) im Raum wird durch eine Funktion

$$x : \mathbb{R} \longrightarrow \mathbb{R}^3, \quad t \mapsto x(t) = \Big(\xi_1(t), \xi_2(t), \xi_3(t)\Big)$$

beschrieben; dabei steht $\mathbb{R} = \mathbb{R}^1$ für die Zeitachse.

b) Der *Temperaturverlauf im Raum* wird durch eine Funktion

$$T : \mathbb{R}^3 \to \mathbb{R}, \quad x \mapsto T(x) = T(\xi_1, \xi_2, \xi_3)$$

beschrieben.

c) Will man beim *Temperaturverlauf* in b auch die *zeitliche Abhängigkeit* berücksichtigen, so hat man eine Funktion

$$T : \mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}, \quad (x, t) \mapsto T(x, t) = T(\xi_1, \xi_2, \xi_3, t).$$

d) Die Kraft, die in einem Punkt des Raumes wirkt, wird durch ein *Kraftfeld* beschrieben,

$$k : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto k(x) = \Big(k_1(\xi_1, \xi_2, \xi_3),\ k_2(\xi_1, \xi_2, \xi_3),\ k_3(\xi_1, \xi_2, \xi_3)\Big).$$

Allgemein wird eine Funktion $f : \mathbb{R}^m \supset A \to \mathbb{R}^m$ als ein *Vektorfeld* (in A) bezeichnet.

e) Will man in d) auch noch eine eventuelle zeitliche Abhängigkeit des Kraftfeldes beschreiben, so hat man

$$k : \mathbb{R}^4 = \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3, \quad (x, t) \mapsto k(x, t).$$

Es tritt also auch hier $\mathbb{R}^4$ in natürlicher Weise als Definitionsbereich auf.

□

Es wird im folgenden immer wieder deutlich werden, daß (gegenüber Funktionen $f : \mathbb{R} \to \mathbb{R}$) nur bei Funktionen $f : \mathbb{R}^m \to \mathbb{R}$ wesentlich neue Verhältnisse auftreten; Funktionen $f : \mathbb{R}^m \to \mathbb{R}^n$ können dagegen im wesentlichen als n–tupel von Funktionen $f_j : \mathbb{R}^m \to \mathbb{R}$ aufgefaßt werden. (Nur die Erhöhung der Zahl der Variablen ist wirklich wesentlich; die Erhöhung der Zahl der Komponenten im Bild schafft keine ernsthaften Probleme.)

Um Stetigkeit von Funktionen erklären zu können, brauchen wir einen *Abstand* in $\mathbb{R}^m$ und einen *Konvergenzbegriff* (dabei handelt es sich um die einfachsten topologischen Begriffe; wir werden später bei Bedarf weitere Begriffe kennenlernen, z.B. offen, abgeschlossen, kompakt).

Für $x = (\xi_1, \ldots, \xi_m)$ und $y = (\eta_1, \ldots, \eta_m)$ aus $\mathbb{R}^m$ ist das (übliche) *Skalarprodukt* erklärt durch

$$\langle x, y\rangle := \sum_{j=1}^{m} \xi_j \eta_j .$$

Es gilt offenbar für $x, y, z \in \mathbb{R}^m, \lambda \in \mathbb{R}$

$$\begin{array}{lll}
\langle x, x\rangle \geq 0; \quad \text{speziell} \ \ \langle x, x\rangle = 0 \Leftrightarrow x = 0, & & \text{(Positivität)}\\
\langle x, y\rangle = \langle y, x\rangle, & & \text{(Symmetrie)}\\
\langle \lambda x, y\rangle = \lambda\langle y, x\rangle \qquad (\text{also} = \langle x, \lambda y\rangle), & & \text{(Homogenität)}\\
\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle . & & \text{(Additivität)}
\end{array}$$

(In $\mathbb{C}^m$ würde man definieren $\langle x, y\rangle := \sum_{j=1}^{m} \overline{\xi_j} \eta_j$; man hat dann $\langle x, y\rangle = \overline{\langle y, x\rangle}$, $\langle \lambda x, y\rangle = \overline{\lambda}\langle x, y\rangle$ und $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$.)

Mit diesem Skalarprodukt wird die (*euklidische*) *Norm* in $\mathbb{R}^m$ (bzw. in $\mathbb{C}^m$) definiert durch

$$|x| = \langle x, x\rangle^{1/2} \quad (\text{ häufig auch} \quad \|\cdot\| \quad \text{statt } |\cdot|),$$

und entsprechend die *euklidische Metrik* in $\mathbb{R}^m$ (bzw. $\mathbb{C}^m$) $\mathrm{d}(x, y) = |x - y|$.

Satz 2.2 (Cauchy–Schwarzsche Ungleichung) *Für alle $x, y \in \mathbb{R}^m$ gilt*

$$|\langle x, y\rangle| \leq |x||y| \qquad \text{Cauchy–Schwarzsche Ungleichung.}$$

In der Schwarzschen Ungleichung gilt Gleichheit genau dann, wenn x und y linear abhängig sind; es gilt $\langle x, y\rangle = |x||y|$ genau dann, wenn $\alpha \geq 0$ existiert mit $x = \alpha y$ oder $y = \alpha x$.

Beweis. Wir beweisen hier nur die Schwarzsche Ungleichung (den Rest der Aussage benutzen wir im folgenden nicht). Für alle $\lambda \in \mathbb{R}$ gilt

$$0 \leq |x + \lambda y|^2 = \lambda^2 |y|^2 + 2\lambda \langle x, y \rangle + |x|^2.$$

Da $\lambda^2 a + 2\lambda b + c \geq 0$ genau dann für alle $\lambda \in \mathbb{R}$ erfüllt ist, wenn $a > 0$ und $b^2 \leq ac$ gilt (man beachte $a = |y^2 \geq 0$ und $c = |x|^2 \geq 0$), folgt $\langle x, y \rangle^2 = b^2 \leq ac = |x|^2 |y|^2$ d. h.$\langle x, y \rangle| \leq |x||y|$.
■

Es gilt für alle $x, y \in \mathbb{R}^m$, $\lambda \in \mathbb{R}$

$$\begin{array}{lll} |x| \geq 0; \quad \text{insbesondere} \quad |x| = 0 \Leftrightarrow x = 0 & & \text{(Positivität)}, \\ |\lambda x| = |\lambda||x| & & \text{(positive Homogenität)}, \\ |x + y| \leq |x| + |y| & & \text{(Dreiecksungleichung)}. \end{array}$$

Dies sind die definierenden Eigenschaften einer Norm. Lediglich die Dreiecksungleichung ist nicht ganz trivial. Sie folgt mit Hilfe der Schwarzschen Ungleichung:

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

Daraus ergeben sich auch die Eigenschaften einer Metrik für $d(x, y) = |x - y|$:

$$\begin{array}{ll} d(x, y) \geq 0; \quad d(x, y) = 0 \Leftrightarrow x = y, & \text{(Positivität)} \\ d(x, y) = d(y, x), & \text{(Symmetrie)} \\ d(x, y) \leq d(x, z) + d(z, y). & \text{(Dreiecksungleichung)} \end{array}$$

Man kann noch zahlreiche andere Normen (und entsprechende Metriken) auf $\mathbb{R}^m$ (bzw. $\mathbb{C}^m$) erklären: gängig ist z. B.

$$|x|_\infty := \max \Big\{ |\xi_j| : j = 1, \ldots, m \Big\}, \text{ und } |x|_1 := \sum_{j=1}^m |\xi_j|.$$

Es gilt offenbar

$$|x|_\infty \leq |x| \leq \sqrt{m} |x|_\infty \text{ und } |x|_\infty \leq |x|_1 \leq m |x|_\infty;$$

man sagt deshalb, die drei Normen sind *äquivalent*. (Entsprechendes gilt auch für $|x|_p := \left\{ \sum_{j=1}^m |\xi_j|^p \right\}^{1/p}$, $1 < p < \infty$.)

Wir können nun *Konvergenz* von Folgen und *Vollständigkeit* von $\mathbb{R}^m$ beschreiben. Eine Folge (x_k) aus $\mathbb{R}^m$, $x_k = (\xi_{k,1}, \xi_{k,2}, \ldots, \xi_{k,m})$, *konvergiert* gegen ein $x \in \mathbb{R}^m$, $x_k \to x$ für $k \to \infty$ (oder $x_k \to x$ oder $\lim\limits_{k\to\infty} x_k = x$ oder $\lim x_k = x$), wenn gilt $|x_k - x| \to 0$ (bzw. gleichwertig: $d(x_k, x) \to 0$) im Sinne der bekannten Konvergenz in $\mathbb{R}$. Offensichtlich ist dies genau dann der Fall, wenn gilt $\xi_{k,j} \to \xi_j$ für $k \to \infty,\ j = 1, \ldots, m$.

Eine Folge (x_k) heißt eine *Cauchyfolge*, wenn gilt: Für jedes $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ so, daß für alle $n, n' \geq n_0$ gilt $|x_n - x_{n'}| < \varepsilon$. Dies ist offenbar genau dann der Fall, wenn $(\xi_{k,j})_{k\in\mathbb{N}}$ für jedes j eine Cauchyfolge in $\mathbb{R}$ ist.

Für äquivalente Normen sind offensichtlich die gleichen Folgen konvergent bzw. Cauchyfolgen; wir werden in der Regel mit der euklidischen Norm arbeiten.

$\mathbb{R}^m$ ist *vollständig*, d. h. jede Cauchyfolge (x_k) in $\mathbb{R}^m$ ist konvergent: Ist (x_k) eine Cauchyfolge, so ist $(\xi_{k,j})_{k\in\mathbb{N}}$ für jedes j eine Cauchyfolge. Da $\mathbb{R}$ vollständig ist, gibt es $\xi_j \in \mathbb{R}$ mit $\xi_{k,j} \to \xi_j$ für $k \to \infty$. Damit folgt $x_k \to x := (\xi_1, \ldots, \xi_m)$, d. h. (x_k) ist konvergent.

2.2 Stetige Funktionen

Sei $A \subset \mathbb{R}^m$. Eine Funktion $f : A \to \mathbb{R}^n$ heißt *stetig im Punkt* $x_0 \in A$, wenn gilt:

> zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ so, daß für alle $x \in A$ mit $|x - x_0| < \delta$ gilt $|f(x) - f(x_0)| < \varepsilon$

(dabei ist mit $|\cdot|$ in $\mathbb{R}^m$ bzw. $\mathbb{R}^n$ die jeweilige euklidische Norm gemeint; allerdings ändert sich der Stetigkeitsbegriff auf Grund der obigen Anmerkung nicht, wenn man eine oder beide Normen durch äquivalente Normen ersetzt). Völlig analog zum 1–dimensionalen Fall zeigt man:

Satz 2.3 *$f : A \to \mathbb{R}^n$ ist genau dann stetig (ε–δ–stetig) in $x_0 \in A$, wenn für jede Folge (x_k) aus A mit $x_k \to x_0$ gilt $f(x_k) \to f(x_0)$ (folgenstetig).*

Beweis. $\Rightarrow$: Sei $x_n \to 0$. Da f stetig ist, existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für $|x - x_0| < \delta$. Zu diesem δ gibt es ein $n_0 \in \mathbb{N}$ mit $|x_n - x_0| < \delta$ für $n > n_0$. Also ist $|f(x_n) - f(x_o)| < \varepsilon$ für

$n > n_0$, d. h. es gilt $f(x_n) \to f(x_0)$.
$\Leftarrow$: Nehmen wir an, daß f nicht stetig ist. Dann gibt es ein $\varepsilon > 0$ so, daß für jedes $n \in \mathbb{N}$ ($\delta = 1/n > 0$) ein x_n mit $|x_n - x_0| < 1/n$ existiert mit $|f(x_n) - f(x_0)| \geq \varepsilon$. Es gilt also $x_n \to x_0$ und $f(x_n) \not\to f(x_0)$, d. h. f ist nicht folgenstetig. ■

Die Funktion f heißt *stetig*, wenn es in jedem Punkt von A stetig ist.

Wieder völlig analog zum 1–dimensionalen Fall zeigt man

> *Summen*, *Produkte* und *Hintereinanderausführungen* (Zusammensetzungen, Kompositionen) stetiger Funktionen sind wieder stetig. Das entsprechende gilt für Quotienten, wenn (wie üblich) die Nullstellen des Nenners als nicht zum Definitionsbereich gehörig betrachtet werden.

Beispiel 2.4 Die Funktion

$$f_{j,k} : \mathbb{R}^m \to \mathbb{R}, \quad x = (\xi_1, \dots, \xi_m) \mapsto \xi_j^k$$

ist stetig. Definieren wir für jeden *Multiindex* $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{N}_0^m$

$$f_\alpha : \mathbb{R}^m \to \mathbb{R}, \quad x = (\xi_1, \dots, \xi_m) \mapsto x^\alpha := \prod_{j=1}^m \xi_j^{\alpha_j},$$

so ist auch f_α stetig. Ebenso ist auch jedes Polynom

$$P : \mathbb{R}^m \to \mathbb{R}, \quad x = (\xi_1, \dots, \xi_m) \mapsto \sum_{|\alpha| \leq k} c_\alpha x^\alpha$$

stetig; hier haben wir für Multiindizes α den *Betrag* bzw. die *Länge* $|\alpha| := \sum_{j=1}^m \alpha_j$ benutzt. Schließlich ist jede *gebrochen rationale Funktion*

$$Q = \frac{P_1}{P_2} : \mathbb{R}^m \backslash \Big\{x \in \mathbb{R}^m : P_2(x) = 0\Big\} \to \mathbb{R}, \quad Q(x) = \frac{P_1(x)}{P_2(x)}$$

(wobei P_1, P_2 Polynome sind) stetig. Die Beweise ergeben sich unmittelbar aus den obigen Aussagen. □

Der folgende Satz macht deutlich, daß sich die Untersuchung stetiger Funktionen $f : \ldots \to \mathbb{R}^n$ auf die Untersuchung von Funktionen $f : \ldots \to \mathbb{R}$ zurückführen läßt.

Satz 2.5 *Sei $A \subset \mathbb{R}^m$, $f : A \to \mathbb{R}^n$, $f(x) = (f_1(x), \ldots, f_n(x))$. Dann ist f genau dann stetig (in $x_0 \in A$), wenn alle f_j stetig (in x_0) sind.*

Der *Beweis* ergibt sich unmittelbar aus der obigen Folgencharakterisierung der Stetigkeit.

Halten wir in $f(x) = f(\xi_1, \ldots, \xi_m)$ alle bis auf eine Variable fest, so erhalten wir die sogenannten *partiellen Funktionen*, z. B.

$$f_{1,x} : \Big\{t \in \mathbb{R} : (\xi_1 + t, \xi_2, \ldots, \xi_m) \in A\Big\} \to \mathbb{R}^m, \quad f_{1,x}(t) = f(\xi_1 + t, \xi_2, \ldots, \xi_m)$$

(und entsprechend $f_{j,x}$ für $j = 2, \ldots, m$). Eine Funktion $f : A \to \mathbb{R}^n$ mit $A \subset \mathbb{R}^m$ heißt *partiell stetig* (d. h. stetig in jeder einzelnen Variablen), wenn für jedes $x \in A$ und jedes $j \in \{1, \ldots, m\}$ die partielle Funktion $f_{j,x}$ stetig ist (Der Leser definiere „partiell stetig“ im Punkt x_0).

Für $x \in A$ und $e \in \mathbb{R}^m$ (o. E. mit $|e| = 1$) definieren wir

$$f_{x,e} : \Big\{t \in \mathbb{R} : x + te \in A\Big\} \to \mathbb{R}^n, \quad f_{x,e}(t) = f(x + te).$$

f heißt *richtungsstetig*, wenn $f_{x,e}$ für alle $x \in A$ und $e \in \mathbb{R}^m$ stetig ist (man definiere „richtungsstetig in x“, „richtungsstetig in Richtung e“). Offensichtlich gilt

Satz 2.6 *f stetig (in x_0) $\Rightarrow$ f richtungsstetig (in x_0) $\Rightarrow$ f partiell stetig (in x_0).*

Gelten auch die umgekehrten Richtungen?

Beispiel 2.7 Wir definieren hier eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ die im Nullpunkt zwar richtungsstetig ist, aber nicht stetig (in allen anderen Punkten des $\mathbb{R}^2$ ist sie stetig).

Dabei setzen wir $f(x) = 1$

- in der abgeschlossenen unteren Halbebene $\Big\{x \in \mathbb{R}^2 : \xi_2 \leq 0\Big\}$, und
- auf und oberhalb der Parabel $\xi_2 = \xi_1^2$, $\Big\{x \in \mathbb{R}^2 : \xi_2 \geq \xi_1^2\Big\}$.

Wir setzen weiter $f(x) = 0$

– auf der Parabel $\xi_2 = \dfrac{1}{2}\xi_1^2$, außer im Nullpunkt (wo sie gleich 1 ist).

Dazwischen wird auf den Parallelen zur ξ_2–Achse linear interpoliert. Somit ist insgesamt

$$f(x) = f(\xi_1, \xi_2) = \begin{cases} 1 & \text{für } \xi_2 \le 0 \text{ und für } \xi_2 \ge \xi_1^2, \\ 1 - 2\dfrac{\xi_2}{\xi_1^2} & \text{für } 0 < \xi_2 \le \frac{1}{2}\xi_1^2, \\ 1 - 2\dfrac{\xi_1^2 - \xi_2}{\xi_1^2} & \text{für } 0 < \frac{1}{2}\xi_1^2 \le \xi_2 < \xi_1^2. \end{cases}$$

Offensichtlich ist f in jedem Punkt außerhalb $x = 0$ stetig. In jeder Richtung durch 0 ist f ebenfalls stetig, da *jede* Gerade durch 0 ein Stück um Null enthält, auf der $f(x) = 1$ gilt (entweder in der unteren Halbebene einschließlich der ξ_1–Achse, oder auf bzw. oberhalb der Parabel $\xi_2 = \xi_1^2$). Da für $x_n = \left(\frac{1}{n}, \frac{1}{2n^2}\right)$ gilt

$$x_n \to 0, \qquad \text{aber} \quad 0 = f(x_n) \not\to f(0) = 1,$$

ist f im Nullpunkt nicht stetig. □

Es sei dem Leser überlassen, ein ähnliches (allerdings einfacheres) Beispiel dafür zu konstruieren, daß aus der partiellen Stetigkeit nicht die Richtungsstetigkeit folgt (z. B. kann man im obigen Beispiel die Parabel $\xi_2 = \xi_1^2$ durch $\xi_2 = |\xi_1|$ ersetzen).

Sei $A \subset \mathbb{R}^m$, $f_k : A \to \mathbb{R}^n$ $(k \in \mathbb{N})$ eine Folge von Funktionen auf A. Die Folge (f_k) heißt (wie im 1–dimensionalen) *gleichmäßig konvergent*, gegen eine Funktion $f : A \to \mathbb{R}^n$, wenn gilt

> zu jedem $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$ so, daß für alle $n \ge n_0$ und alle $x \in A$ gilt $|f_n(x) - f(x)| < \varepsilon$.

Wie im 1–dimensionalen beweist man leicht:

Satz 2.8 *Sei $A \subset \mathbb{R}^m$, $f_k : A \to \mathbb{R}^n$ stetig für alle $k \in \mathbb{N}$, und (f_k) konvergiere gleichmäßig gegen eine Funktion $f : A \to \mathbb{R}^m$. Dann ist auch f stetig.*

2.3 Grenzwerte von Funktionen

Völlig analog zum 1–dimensionalen Fall definieren wir in diesem Abschnitt Grenzwerte von Funktionen für $x \to x_0$.

Sei $A \subset \mathbb{R}^m$; ein $x_0 \in \mathbb{R}^m$ heißt ein *Berührungspunkt* von A, wenn es für jedes $\varepsilon > 0$ (mindestens) ein $x \in A$ gibt mit $|x - x_0| < \varepsilon$. Insbesondere ist also jeder Punkt von A auch Berührungspunkt von A. – Ein $x_0 \in \mathbb{R}^m$ heißt ein *Häufungspunkt* von A, wenn es zu jedem $\varepsilon > 0$ (mindestens) ein $x \in A \setminus \{x_0\}$ gibt mit $|x - x_0| < \varepsilon$. Jeder Häufungspunkt ist auch Berührungspunkt, aber nicht umgekehrt.

Offenbar ist x_0 genau dann Berührungspunkt (bzw. Häufungspunkt), wenn es eine Folge (x_k) aus A (bzw. aus $A \setminus \{x_0\}$) gibt mit $x_k \to x_0$.

Sei $A \subset \mathbb{R}^m$, x_0 Häufungspunkt von A, $f : A \backslash \{x_0\} \to \mathbb{R}^n$. Wir sagen $y \in \mathbb{R}^n$ ist der *Grenzwert* von $f(x)$ für $x \to x_0$, $\lim\limits_{x \to x_0} f(x) = y$, wenn gilt:

– für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ so, daß für alle $x \in A \backslash \{x_0\}$ mit $|x - x_0| < \delta$ gilt $|f(x) - y| < \varepsilon$.

Dies ist offenbar äquivalent zu

– für jede Folge (x_k) aus $A \backslash \{x_0\}$ mit $x_k \to x_0$ gilt $f(x_k) \to y$, bzw.

– die Funktion $\tilde{f} : A \cup \{x_0\} \to \mathbb{R}^n$, $\tilde{f}(x) = \begin{cases} f(x) & \text{für } x \neq x_0, \\ y & \text{für } x = x_0 \end{cases}$ ist in x_0 stetig.

Ist $x_0 \in A$, so existiert also $\lim_{x \to x_0} f(x)$ genau dann, wenn f in x_0 stetig ist.

Beispiel 2.9 a) Die Funktion

$$f : \mathbb{R}^m \backslash \{0\} \to \mathbb{R}^m, \quad f(x) = |x|^{-1} x$$

hat für $x \to 0$ keinen Grenzwert (das ist übrigens schon für $m = 1$ so).

b) Die Funktion

$$f : \Big\{ x \in \mathbb{R}^m : \xi_j \neq 0 \text{ für alle } j \Big\}, \to \mathbb{R}^m, \quad f(x) = \Big(\frac{\sin \xi_1}{\xi_1}, \ldots, \frac{\sin \xi_m}{\xi_m} \Big)$$

hat Grenzwerte in allen Punkten x, die nicht zum Definitionsbereich gehören (die Werte der einzelnen Komponenten sind $\xi_j^{-1} \sin \xi_j$ für die j mit $\xi_j \neq 0$ und 1 für die j mit $\xi_j = 0$).

□

Satz 2.10 *Sei $A \subset \mathbb{R}^m$, $f, g : A \to \mathbb{R}^n$, $h : f(A) \to \mathbb{R}^p$.*

a) *Es gilt $\lim\limits_{x \to x_0} f(x) = y = (\eta_1, \ldots, \eta_n)$ genau dann, wenn für $j = 1, \ldots, n$ gilt $\lim\limits_{x \to x_0} f_j(x) = \eta_j$.*

b) *Existieren die Grenzwerte $\lim\limits_{x \to x_0} f(x)$ und $\lim\limits_{x \to x_0} g(x)$, so existieren auch die folgenden Grenzwerte und es gilt für $\lambda, \mu \in \mathbb{R}$*

   i) $\lim\limits_{x \to x_0} \Big(\lambda f(x) + \mu g(x)\Big) = \lambda \lim\limits_{x \to x_0} f(x) + \mu \lim\limits_{x \to x_0} g(x)$,

   ii) $\lim\limits_{x \to x_0} \Big\langle f(x),\, g(x) \Big\rangle = \Big\langle \lim\limits_{x \to x_0} f(x),\, \lim\limits_{x \to x_0} g(x) \Big\rangle$,

   iii) $\lim\limits_{x \to x_0} |f(x)| = \Big| \lim\limits_{x \to x_0} f(x) \Big|$.

c) *Ist außerdem $h : f(A) \to \mathbb{R}^p$, $y_0 := \lim\limits_{x \to x_0} f(x)$ und existiert $z := \lim\limits_{y \to y_0} h(y)$, so existiert auch der folgende Grenzwert und es gilt*

$$\lim_{x \to x_0} (h \circ f)(x) = z.$$

Die *Beweise* sind offensichtlich, z. B. mit Hilfe der Folgencharakterisierung des Grenzwertes.

2.4 Einige weitere „topologische" Begriffe

Die („offene") *Kugel* um $x \in \mathbb{R}^m$ mit Radius r ist definiert durch

$$K_r(x) = \Big\{ y \in \mathbb{R}^m : |x - y| < r \Big\}$$

(für $r = 0$ ist $K_r(x)$ leer). Eine Teilmenge $A \subset \mathbb{R}^m$ heißt *offen*, wenn zu jedem $x \in A$ ein $r = r(x) > 0$ existiert mit $K_r(x) \subset A$. — Zum Beispiel ist jedes „offene" Intervall in $\mathbb{R}^1$

$$(a, b) = \Big\{ x \in \mathbb{R} : a < x < b \Big\} \quad \text{mit} \quad a \leq b$$

offen in diesem Sinn (man kann z. B. $r(x) = \min\{x - a, b - x\} > 0$ wählen). Jede „offene“ Kugel $K_r(x)$ ist ebenfalls offen im Sinne dieser Definition (z. B. mit $r(x) = r - |z - x| > 0$). Die leere Menge und der ganze Raum sind offensichtlich offen (dies gilt insbesondere für (a, a) und $K_0(z)$).

Eine Teilmenge $A \subset \mathbb{R}^m$ heißt *abgeschlossen*, wenn ihr Komplement $\mathbb{R}^m \backslash A$ offen ist. Z. B. ist jedes „abgeschlossene“ Intervall

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\} \quad \text{mit} \quad a \leq b$$

abgeschlossen und jede „abgeschlossene“ Kugel in $\mathbb{R}^m$

$$\overline{K}_r(z) = \{x \in \mathbb{R}^m : |z - x| \leq r\} \qquad \text{für} \quad r \geq 0$$

ist ebenfalls abgeschlossen. Die leere Menge und der ganze Raum sind offensichtlich auch abgeschlossen.

Satz 2.11 *Eine Teilmenge $A \subset \mathbb{R}^m$ ist genau dann abgeschlossen, wenn der Grenzwert jeder konvergenten Folge aus A wieder in A liegt.*

Beweis. $\Rightarrow$: Wir nehmen an, daß die Behauptung nicht gilt, d. h. es gibt eine Folge (x_k) aus A mit $x_k \to x$ und $x \in B := \mathbb{R}^m \backslash A$. Da A abgeschlossen ist, ist B offen, d. h. es gibt ein $r > 0$ mit $K_r(x) \subset B$. Wegen $x_k \to x$ würden dann aber alle x_k für hinreichend große k in B liegen, im Widerspruch dazu, daß (x_k) eine Folge aus A war.

$\Leftarrow$: Nehmen wir an, daß A nicht abgeschlossen ist (Achtung: das bedeutet *nicht*, daß A offen ist), d. h. $B := \mathbb{R}^m \backslash A$ ist nicht offen. Also gibt es ein $x \in B$ so, daß $K_r(x) \cap A \neq \emptyset$ für alle $r > 0$ gilt. Insbesondere gibt es zu jedem k ein $x_k \in K_{1/k}(x) \cap A \subset A$, d. h. es gilt $x_k \in A$, $x_k \to x$ und $x \notin A$ im Widerspruch zur Voraussetzung. ∎

Eine Teilmenge $A \subset \mathbb{R}^m$ heißt *beschränkt*, wenn ein $C \geq 0$ existiert mit

$$|x| \leq C \text{ für alle } x \in A \text{ (oder } A \subset \overline{K}_C(0) \text{ bzw. } A \subset K_D(0) \text{ für } D > C\text{).}$$

Eine Teilmenge $A \subset \mathbb{R}^m$ heißt *kompakt*, wenn sie beschränkt und abgeschlossen ist. Der folgende Satz ist das m–dimensionale Analogon zum Satz von Bolzano–Weierstraß.

Satz 2.12 *Eine Teilmenge $A \subset \mathbb{R}^m$ ist genau dann kompakt, wenn jede Folge (x_k) aus A eine konvergente Teilfolge mit Grenzwert in A besitzt.*

Beweis. $\Rightarrow$: Sei A kompakt (beschränkt und abgeschlossen), (x_k) eine Folge aus A, $x_k = (\xi_{k,1}, \ldots \xi_{k,m})$. Dann gilt $|\xi_{k,j}| \leq C$ für alle $k \in \mathbb{N}$, $j = 1, \ldots, m$. Also enthält $(k)_{k \in \mathbb{N}}$ eine Teilfolge $(n_1(k))_{k \in \mathbb{N}}$ so, daß $(\xi_{n_1(k),1})$ in $\mathbb{R}$ konvergiert (Satz von Bolzano–Weierstraß). Weiter enthält $(n_1(k))_{k \in \mathbb{N}}$ eine Teilfolge $(n_2(k))_{k \in \mathbb{N}}$ so, daß $(\xi_{n_2(k),2})$ in $\mathbb{R}$ konvergiert; natürlich ist auch $(\xi_{n_2(k),1})$ konvergent. Nach m Schritten erhalten wir eine Teilfolge $(n(k))_{k \in \mathbb{N}} = (n_m(k))_{k \in \mathbb{N}}$ so, daß die Folgen $(\xi_{n(k),j})_{k \in \mathbb{N}}$ für alle $j = 1, \ldots, m$ konvergieren. Dann ist auch $(x_{n(k)})$ konvergent gegen ein $x \in \mathbb{R}^m$. Da A abgeschlossen ist, liegt x in A.

$\Leftarrow$: *A ist beschränkt*: Sonst gäbe es eine Folge (x_k) aus A mit $|x_k| \to \infty$; diese könnte offensichtlich keine konvergente Teilfolge enthalten (denn für jedes $x \in \mathbb{R}^m$ würde gelten $|x_k - x| \geq |x_k| - |x| \to \infty$).

A ist abgeschlossen: Sei (x_k) eine Folge aus A mit $x_k \to x$. Auf Grund der Voraussetzung an A gibt es eine Teilfolge, die gegen einen Punkt aus A konvergiert. Andererseits konvergiert jede Teilfolge einer konvergenten Folge gegen den gleichen Grenzwert. Also liegt x in A. ∎

Satz 2.13 (Satz vom Maximum) *Ist $K \subset \mathbb{R}^m$ kompakt und $f : K \to \mathbb{R}$ stetig, so ist f beschränkt und nimmt in K sein Maximum (und sein Minimum) an.*

Beweis. Der Beweis ist völlig gleich wie im eindimensionalen Fall; es genügt, die Beschränktheit nach oben und die Esxistenz des Maximums zu beweisen.

f ist beschränkt: Nehmen wir an, daß f unbeschränkt ist, d. h. es gibt eine Folge (x_n) in K mit $f(x_n) \to \infty$. Da K kompakt ist, gibt es eine konvergente Teilfolge (x_{n_k}). Mit $x := \lim_{k\to\infty} x_{n_k}$ gilt dann $f(x_{n_k}) \to f(x)$, ein Widerspruch zu $f(x_{n_k}) \to \infty$.

f nimmt sein Maximum an: Da f beschränkt ist, existiert $\sup f$, und es gibt eine Folge (x_n) mit $f(x_n) \to \sup f$. Dann gibt es eine Teilfolge (x_{n_k}) mit $x_{n_k} \to x$ und $f(x) = \lim_{k\to\infty} f(x_{n_k}) = \sup f$. ∎

Eine Funktion $f : A \to \mathbb{R}^n$ $(A \subset \mathbb{R}^m)$ heißt (wie im 1–dimensionalen) *gleichmäßig stetig*, wenn gilt: Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ so, daß für alle $x, y \in A$ mit $|x - y| < \delta$ gilt $|f(x) - f(y)| < \varepsilon$.

Satz 2.14 *Sei $A \subset \mathbb{R}^m$ kompakt, $f : A \to \mathbb{R}^n$ stetig. Dann ist f gleichmäßig stetig.*

Beweis. Der Beweis läuft wie in $\mathbb{R}^1$. Dort haben wir den Satz von Bolzano–Weierstraß benutzt, der hier durch die Kompaktheit ersetzt wird (Übung!). ■

Die folgende Charakterisierung der kompakten Teilmengen von $\mathbb{R}^m$ ist die auch in wesentlich allgemeinerem Kontext in der Topologie übliche Definition. Eine Familie $\{U_\alpha : \alpha \in A\}$ mit einer beliebigen Indexmenge A heißt eine *offene Überdeckung* einer Menge $K \subset \mathbb{R}^m$, wenn alle U_α offene Teilmenge von $\mathbb{R}^m$ sind, und $K \subset \bigcup\limits_{\alpha \in A} U_\alpha$ gilt.

Satz 2.15 (Heine–Borelscher Überdeckungssatz) *Eine Teilmenge $K \subset \mathbb{R}^m$ ist genau dann kompakt, wenn es zu jeder offenen Überdeckung $\{U_\alpha : \alpha \in A\}$ von K ein $n \in \mathbb{N}$ und $\alpha_1, \ldots, \alpha_n \in A$ gibt mit $K \subset \bigcup\limits_{j=1}^{n} U_{\alpha_j}$. (Man sagt kurz: Kist genau dann kompakt, wenn jede offene Überdeckung von K eine* endliche Teilüberdeckung *enthält — noch kürzer, wenn K die* Überdeckungseigenschaft *hat.)*

Beispiel 2.16 Die Menge $M = (0, 1]$ ist *nicht* kompakt. Offenbar ist $\{U_r : r \in (0, 1]\}$ mit $U_r := (\frac{r}{2}, 2r)$ eine offene Überdeckung von M, die keine endliche Teilüberdeckung enthält. □

Beweis von Satz 2.15. $\Rightarrow$: Sei K kompakt, $\{U_\alpha : \alpha \in A\}$ eine offene Überdeckung von K. Es ist zu zeigen, daß $\{U_\alpha\}$ eine endliche Teilüberdeckung enthält.

a) Wir zeigen zunächst: Es gibt ein $\delta > 0$ so, daß für jedes $x \in K$ ein $\alpha(x) \in A$ existiert mit $K(x, \delta) \subset U_{\alpha(x)}$.

Wir nehmen das Gegenteil an, d. h. insbesondere: zu jedem $n \in \mathbb{N}$ ($\delta = 1/n$) gibt es ein $x_n \in K$ mit $K(x_n, \frac{1}{n}) \not\subset U_\alpha$ für alle $\alpha \in A$. Da K kompakt ist, existiert eine Teilfolge (x_{n_k}) mit $x_{n_k} \to x \in K$ für $k \to \infty$. Natürlich gibt es ein $\alpha \in A$ mit $x \in U_\alpha$. Da U_α offen ist, gibt es ein $\varepsilon > 0$ mit $K(x, \varepsilon) \subset U_\alpha$. Wegen $x_{n_k} \to x$ ist $x_{n_k} \in K(x, \frac{\varepsilon}{2})$ für hinreichend große k und somit

$$K\Big(x_{n_k}, \frac{1}{n_k}\Big) \subset K(x, \varepsilon) \subset U_\alpha \text{ für hinreichend große } k \text{ (mit } \frac{1}{n_k} < \frac{\varepsilon}{2}).$$

Das ist im Widerspruch zur Konstruktion der Folge (x_n).

b) Wir zeigen: Zu jedem $\delta > 0$ gibt es ein $n \in \mathbb{N}$ und $x_1, \ldots, x_n \in K$ mit $K \subset \bigcup_{j=1}^{n} K(x_j, \delta)$ (man sagt dafür auch „K ist total beschränkt“). — Zusammen mit Teil a folgt dann die Behauptung: Zu δ aus Teil a wählt man die x_j nach Teil b und α_j so, daß $K(x_j, \delta) \subset U_{\alpha_j}$ gilt.

Wir nehmen an, daß dies nicht gilt, d. h. es gibt ein $\delta > 0$ so, daß K nicht durch endlich viele Kugeln $K(x, \delta)$ überdeckt werden kann. Man kann also für dieses δ induktiv eine Folge (x_n) aus K definieren mit

$$|x_{n+1} - x_j| \geq \delta \text{ für } j = 1, \ldots, n, \text{ alle } n \in \mathbb{N}.$$

Diese Folge kann offensichtlich keine konvergente Teilfolge enthalten. Das ist im Widerspruch zur Kompaktheit.

$\Leftarrow$: K habe die Überdeckungseigenschaft, (x_n) sei eine Folge in K. Es ist zu zeigen, daß (x_n) eine in K konvergente Teilfolge enthält. Wir nehmen an, daß dies nicht der Fall ist, d. h. kein $x \in K$ ist Häufungspunkt von (x_n) (denn die Häufungspunkte sind gerade die Grenzwerte von Teilfolgen). Das bedeutet: Zu jedem $x \in K$ existiert ein $\varepsilon(x) > 0$ so, daß gilt

$$K(x, \varepsilon(x)) \text{ enthält nur endlich viele der } x_n.$$

Natürlich ist $\Big\{K(x, \varepsilon(x)) : x \in K\Big\}$ eine offene Überdeckung von K, die nach Voraussetzung eine endliche Teilüberdeckung $\Big\{K(x_j, \varepsilon(x_j)) : j = 1, \ldots, n\Big\}$ enthält. Da jedes $K(x_j, \varepsilon(x_j))$ nur endlich viele der x_n enthält, ergibt sich ein Widerspruch. ∎

2.5 Übungsaufgaben

2.1 Für welche $a > 0$ ist die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$

$$f(x) = f(\xi_1, \xi_2) = \begin{cases} |x|^{-a}\xi_1\xi_2 & \text{für } x \neq 0, \\ 0 & \text{für } x = 0 \end{cases}$$

partiell stetig, richtungsstetig, stetig ?

2.2 Zwei Normen $|\cdot|_a, |\cdot|_b$ heißen äquivalente Normen, wenn es Konstanten $c_1, c_2 > 0$ gibt mit $|\cdot|_a \leq c_1|\cdot|_b, \quad |\cdot|_b \leq c_2|\cdot|_a$.

a) Sind $|\cdot|_a$ und $|\cdot|_b$ äquivalente Normen, so ist eine Folge (x_n) genau dann bezüglich $|\cdot|_a$ eine Cauchyfolge (konvergent), wenn dies bezüglich $|\cdot|_b$ gilt.

b) Auf $\mathbb{R}^m$ sind durch $|x|_1 = \sum_{k=1}^{m} |\xi_k|, \ |x|_\infty = \max\{|\xi_k| : k = 1, \ldots, m\}$ Normen definiert (die Norm aus der Vorlesung wird mit $|\cdot|_2$ bezeichnet).

c) Die Normen $|\cdot|_1, \ |\cdot|_2$ und $|\cdot|_\infty$ auf $\mathbb{R}^m$ sind äquivalent.

d) $\mathbb{R}^m$ ist auch bezüglich $|\cdot|_1$ und $|\cdot|_\infty$ vollständig.

2.3 Eine Funktion $f : A \to \mathrm{Re}^n$ $(A \subset \mathbb{R}^m)$ heißt *hölderstetig*, wenn ein $\alpha > 0$ und ein $c \geq 0$ existiert mit $|f(x) - f(y)| \leq c|x-y|^\alpha$ für alle $x, y \in A$. Man sagt dann auch, f sei *hölderstetig mit Exponent* α; ist $\alpha = 1$, so nennt man f *lipschitzstetig.*

a) Jede hölderstetige Funktion ist gleichmäßig stetig (Es gibt allerdings auch den Begriff „lokal hölderstetig“ oder „lokal hölderstetig mit Exponent α“)

b) Für A=$[0, 1] \subset \mathbb{R}^1$ und $n = 1$ gebe man eine (gleichmäßig) stetige Funktion an, die *nicht* hölderstetig ist.

2.4 Für eine Funktion $f : \mathbb{R}^m \to \mathbb{R}^n$ zeige man: f ist genau dann stetig, wenn für jede offene Menge $B \subset \mathbb{R}^n$ auch das Urbild $f^{-1}(B) = \{x \in \mathbb{R}^m : f(x) \in B\}$ offen ist.

2.5 In $\mathbb{R}^m$ sind $\emptyset$ und $\mathbb{R}^m$ die einzigen Teilmengen, die offen und abgeschlossen sind.

2.6 Das *Innere* $\overset{\circ}{A}$ einer Teilmenge A von $\mathbb{R}^m$ ist die Menge der *inneren Punkte* $x \in A$, d. h. der Punkte $x \in A$, für die ein $\varepsilon > 0$ existiert mit $K_\varepsilon(x) \subset A$. Man zeige:

a) $\overset{\circ}{A} \cup \overset{\circ}{B} \subset (A \cup B)^\circ$; die Gleichheit gilt i. allg. nicht.

b) $\overset{\circ}{A} \cap \overset{\circ}{B} = (A \cap B)^\circ$

c) $\overset{\circ}{A}$ ist die größte offene Teilmenge von A, d. h. es gilt: $\overset{\circ}{A}$ ist offen und aus $B \subset A$ und B offen folgt $B \subset \overset{\circ}{A}$.

2.7 Der *Abschluß* $\overline{A}$ einer Teilmenge A non $\mathbb{R}^m$ ist die Menge der Berührungspunkte von A, also die Menge aller Grenzwerte von Folgen aus A. Man zeige (beachte die „Dualität“ zur vorhergehenden Aufgabe):

a) $\overline{A}$ ist die kleinste abgeschlossene Teilmenge von $\mathbb{R}^m$, die A enthält (d. h. A ist abgeschlossen und aus B abgeschlossen und $A \subset B$ folgt $\overline{A} \subset B$).

b) $\overline{A} \cup \overline{B} = \overline{A \cup B}$.

c) $\overline{A} \cap \overline{B} \supset \overline{A \cap B}$; die Gleichheit gilt i. allg. nicht.

2.8 a) Die Vereinigung beliebig vieler offener Mengen ist offen.

b) Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.

c) Der Durchschnitt endlich vieler offener Mengen ist offen.

d) Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.

e) Die Aussagen c und d gelten i. all. nicht für abzählbar viele Mengen.

2.9 Man zeige in $\mathbb{R}$ (etwas schwieriger ist es in $\mathbb{R}^m$): Die leere Menge und der ganze Raum sind die einzigen Mengen, die offen und abgeschlossen sind.

3 Differentiation von Funktionen $f : \mathbb{R}^m \to \mathbb{R}^n$

3.1 Differenzierbarkeit und partielle Differenzierbarkeit

Im 1–dimensionalen hatten wir die Differenzierbarkeit in einem Punkt x_0 zunächst mit der Existenz des Grenzwertes der Differenzenquotienten erklärt:

$$f'(x_0) = \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0}.$$

Dies läßt sich offensichtlich so nicht auf Funktionen, die in $\mathbb{R}^m$ definiert sind, übertragen, da wir nicht durch den Vektor $x - x_0$ dividieren können, und Division durch $|x - x_0|$ liefert schon im 1–dimensionalen nicht das „richtige"Resultat.

Wir haben dann im 1–dimensionalen gezeigt, daß f genau dann in x_0 differenzierbar ist, wenn es ein $c \in \mathbb{R}$ gibt mit

$$f(x) = f(x_0) + c(x - x_0) + \varphi(x_0, x),$$

wobei für $\varphi(x_0, x)$ gilt

$$\frac{\varphi(x_0, x)}{x - x_0} \to 0 \text{ für } x \to x_0.$$

Es ist dann c die Ableitung von f im Punkt x_0, $c = f'(x_0)$. In der Nähe von x_0 wird also f „sehr gut" durch die affine Funktion

$$g(x) = f(x_0) + c(x - x_0)$$

approximiert, die die Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$ beschreibt (*lineare Approximierbarkeit*). Dieses Resultat benutzen wir in $\mathbb{R}^m$ als Definition:

Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}^n$, $x_0 \in A$. f heißt im Punkt x_0 (*total*) *differenzierbar*, wenn eine lineare Abbildung $L : \mathbb{R}^m \to \mathbb{R}^n$ existiert mit

$$f(x) = f(x_0) + L(x - x_0) + \varphi(x_0, x),$$

wobei für $\varphi(x_0, x)$ gilt

$$\frac{\varphi(x_0, x)}{|x - x_0|} \to 0 \text{ für } x \to x_0;$$

in anderen Worten, wenn gilt:

$$\frac{f(x) - f(x_0) - L(x - x_0)}{|x - x_0|} \to 0 \text{ für } x \to x_0.$$

Der folgende Satz erlaubt es, diese lineare Abbildung L als *Ableitung* von f im Punkt x_0 zu bezeichnen: $\mathrm{D}f(x_0)$ (oder auch gelegentlich $\mathrm{d}f(x_0)$, $f'(x_0)$).

Satz 3.1 *Sei f in x_0 differenzierbar. Dann ist die lineare Abbildung L eindeutig bestimmt.*

Beweis. Wir nehmen an, daß die obige Bedingung für lineare Abbildungen L_1 und L_2 gilt mit entsprechenden Funktionen $\varphi_1(x_0, x), \varphi_2(x_0, x)$. Dann folgt für die lineare Abbildung $M := L_1 - L_2$

$$\begin{aligned}
\lim_{x \to 0} |x|^{-1} Mx &= \lim_{x \to x_0} |x - x_0|^{-1} M(x - x_0) \\
&= \lim_{x \to x_0} |x - x_0|^{-1} \Big\{ L_1(x - x_0) - L_2(x - x_0) \Big\} \\
&= \lim_{x \to x_0} |x - x_0|^{-1} \Big\{ \varphi_2(x_0, x) - \varphi_1(x_0, x) \Big\} = 0.
\end{aligned}$$

Also gilt für alle $x \in \mathbb{R}^m$

$$Mx = nM\Big(\frac{x}{n}\Big) = |x| \Big\{ \Big|\frac{x}{n}\Big|^{-1} M\Big(\frac{x}{n}\Big) \Big\} \to 0 \text{ für } n \to \infty,$$

und somit

$$Mx = 0 \text{ für alle } x \in \mathbb{R}^m,$$

d. h. M ist die Nullabbildung, $L_1 = L_2$. ∎

$f : A \to \mathbb{R}^n$ heißt *differenzierbar*, wenn es in jedem Punkt von A differenzierbar ist. Dann ist $\mathrm{D}f(\cdot)$ eine neue Funktion auf A, deren Werte lineare Abbildungen von $\mathbb{R}^m$ nach $\mathbb{R}^n$ sind, die *Ableitung* von f.

Im folgenden benutzen wir, daß sich jede lineare Abbildung L von $\mathbb{R}^m$ nach $\mathbb{R}^n$ mit Hilfe einer Matrix

$$(\ell_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,m}} = \begin{pmatrix} \ell_{11} & \dots & \ell_{1m} \\ \vdots & \ddots & \vdots \\ \ell_{n1} & \dots & \ell_{nm} \end{pmatrix}$$

mit n Zeilen und m Spalten darstellen läßt, $(Lx)_k = \sum_{\nu=1}^m l_k \nu \xi_\nu$ für $k = 1, \cdots, m$.

Satz 3.2 *Sei $A \subset \mathbb{R}^m$ offen und $f : A \to \mathbb{R}^n$ differenzierbar (in x_0). Dann ist f stetig (in x_0).*

Beweis. Es genügt, die Behauptung für x_0 zu beweisen. Da f in x_0 differenzierbar ist, gilt

$$f(x) = f(x_0) + L(x - x_0) + \varphi(x_0, x)$$

mit $|x - x_0|^{-1}\varphi(x_0, x) \to 0$ für $x \to x_0$, also ganz sicher $\varphi(x_0, x) \to 0$ für $x \to x_0$. Aus der obigen Darstellung von L folgt (mit $x_0 = (\xi_{0,1}, \dots, \xi_{0,m})^t$,$x = (\xi_1, \dots, \xi_m)^t$)

$$\begin{aligned} L(x - x_0) &= \Big(\sum_{j=1}^m \ell_{1j}(\xi_j - \xi_{0,j}), \dots, \sum_{j=1}^m \ell_{nj}(\xi_j - \xi_{o,j}) \Big)^t \\ &\to 0 \text{ für } x \to x_0 \qquad (\xi_j \to \xi_{0,j},\ j = 1, \dots, m), \end{aligned}$$

also $f(x) \to f(x_0)$ für $x \to x_0$. ■

Für konkrete Berechnungen sind die obigen recht abstrakten Überlegungen kaum hilfreich. Man wird statt dessen mit einzelnen Koordinaten rechnen müssen. Sei f in x_0 differenzierbar, $\mathrm{D}f(x_0) = L = (\ell_{ij})$, also

$$f(x) - f(x_0) - \Big(\dots, \sum_{j=1}^m \ell_{ij}(\xi_j - \xi_{0,j}), \dots \Big)^t = \varphi(x_0, x).$$

Betrachten wir nur die i–te Komponente dieser Gleichung, so erhalten wir nach Division durch $|x - x_0|$

$$|x - x_0|^{-1}\Big\{ f_i(x) - f_i(x_0) - \sum_{j=1}^m \ell_{ij}(\xi_j - \xi_{0,j}) \Big\} \to 0 \text{ für } x \to x_0.$$

Wählen wir hier speziell $x = x_0 + he_k$ ($e_k = k$–ter Einheitsvektor, $k \in \mathbb{R} \setminus \{0\}$), so folgt

$$\begin{aligned} &\frac{h}{|h|}\left\{\frac{f_i(x_0 + he_k) - f_i(x_0)}{h} - l_{ik}\right\} \\ = \quad & \frac{1}{|h|}\left\{f_i(x_0 + he_k) - f_i(x_0) - h\ell_{ik}\right\} \to 0 \ \text{ für } h \to 0, \end{aligned}$$

d. h. es gilt

$$\ell_{ik} = \lim_{h \to 0} \frac{1}{h}\left\{f_i(x_0 + he_k) - f_i(x_0)\right\} =: \frac{\partial f_i}{\partial \xi_k}(x_0).$$

Dieser Grenzwert wird als *partielle Ableitung* von f_i nach der k–ten Variablen im Punkt x_0 bezeichnet. Für die partiellen Ableitungen $\partial f_i / \partial \xi_j$ schreiben wir auch $\mathrm{D}_j f_i$, $\partial_j f_i$, $f_{i,j}$.

Mit der eben durchgeführten Überlegung haben wir bewiesen:

Satz 3.3 *Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}^m$ differenzierbar in x_0. Dann existieren alle partiellen Ableitungen von f in x_0 und es gilt*

$$\mathrm{D}f(x_0) = \left(\frac{\partial f_i}{\partial \xi_j}(x_0)\right)_{\substack{i=1,\dots,n \\ j=1,\dots,m}}.$$

Die im Satz angegebene Matrix wird als *Jacobi–Matrix* (oder *Differentialmatrix* oder *Funktionalmatrix*) von f im Punkt x_0 bezeichnet. Man schreibt dafür auch

$$\left(\frac{\partial(f_1, \dots, f_n)}{\partial(\xi_1, \dots, \xi_m)}\right)(x_0).$$

Eine Funktion $f : A \to \mathbb{R}^n$ ($A \subset \mathbb{R}^m$ offen) heißt *in* $\mathbf{x_0}$ *partiell differenzierbar* wenn in x_0 alle partiellen Ableitungen $\mathrm{D}_j f_i(x_0)$ existieren. Sie heißt *partiell differenzierbar*, wenn sie in jedem Punkt von A partiell differenzierbar ist. Sie heißt *stetig partiell* differenzierbar, wenn alle partiellen Ableitungen stetige Funktionen auf A sind. Der obige Satz besagt, daß jede differenzierbare Funktion auch partiell differenzierbar ist. Das folgende Beispiel zeigt, daß die Umkehrung nicht gilt. Danach zeigen wir dann, daß jede stetig partiell differenzierbare Funktion auch (stetig) differenzierbar ist.

Beispiel 3.4 Die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x) = f(\xi_1, \xi_2) = \begin{cases} \dfrac{\xi_1 \xi_2}{|x|} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0 \end{cases}$$

ist außerhalb des Punktes $x = (\xi_1, \xi_2) = 0$ offensichtlich (sogar stetig) partiell differenzierbar (nachrechnen!). Auf beiden Achsen verschwindet sie identisch. Deshalb existieren in 0 beide partiellen Ableitungen und sind gleich 0. Wäre f in 0 (total) differenzierbar, so müßte $\mathrm{D}f(0) = (0, 0) = 0$ gelten, also

$$\frac{1}{|x|} f(x) = \frac{1}{|x|} \{ f(x) - f(0) - 0x \} \to 0 \text{ für } x \to 0.$$

Wegen $f(t, t) = t/\sqrt{2}$ für alle $t \in \mathbb{R}$ kann dies aber nicht gelten. Also ist f im Nullpunkt nicht differenzierbar. □

Der folgende Satz liefert uns, daß in dem eben betrachteten Beispiel die partiellen Ableitungen von f im Nullpunkt nicht stetig sein können. (Übrigens könnte man in obigem Beispiel den Faktor $(\xi_1^2 + \xi_2^2)^{-1/2}$ durch $(\xi_1^2 + \xi_2^2)^{-1}$ ersetzen. Es würden die gleichen Resultate gelten; an $f(t, t) = \frac{1}{2}$ für $t \neq 0$ erkennt man, daß dieses f im Nullpunkt sogar unstetig ist, ein weiteres Beispiel für eine Funktion, die partiell stetig, aber nicht stetig ist.)

Satz 3.5 *Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}^n$ stetig partiell differenzierbar. Dann ist f (total) differenzierbar. (Da dann alle Glieder der Matrix $\mathrm{D}f(x)$ stetige Funktionen sind, sagt man f ist* stetig differenzierbar.*)*

Beweis. Sei $x_0 \in A$, $L := (\mathrm{D}_j f_i(x_0))_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$. Wir haben zu zeigen: zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ so, daß für alle $x \in A$ mit $|x - x_0| < \delta$ gilt

$$|x - x_0|^{-1} \Big| f(x) - f(x_0) - L(x - x_0) \Big| < \varepsilon.$$

Sei also $\varepsilon > 0$. Da alle $\mathrm{D}_j f_i(\cdot)$ stetig sind, existiert ein $\delta > 0$ mit

$$\Big| \mathrm{D}_j f_i(x) - \mathrm{D}_j f_i(x_0) \Big| < \frac{\varepsilon}{\sqrt{nm}} \text{ für } x \in A \text{ mit } |x - x_0| < \delta.$$

Sei $x \in K_\delta(x_0)$, $t = (\tau_1, \ldots, \tau_m)^t := x - x_0$,

$$t_j := (\tau_1, \ldots, \tau_j, 0, \ldots, 0)^t = \sum_{\ell=1}^{j} \tau_\ell e_\ell \text{ für } j = 0, \ldots, m.$$

Dann gilt für $i = 1, \ldots, n$

$$f_i(x) - f_i(x_0) = \sum_{j=1}^{m} \Big(f_i(x_0 + t_j) - f_i(x_0 + t_{j-1}) \Big).$$

Nach dem (1–dimensionalen) Mittelwertsatz der Differentialrechnung existieren $\sigma_{j,i}$ zwischen 0 und τ_j mit

$$\begin{aligned}
f_i(x_0 + t_j) - f_i(x_0 + t_{j-1}) &= f_i(x_0 + t_{j-1} + \tau_k e_k) - f_i(x_0 + t_{j-1}) \\
&= \tau_j \, \mathrm{D}_j f_i(x_0 + t_{j-1} + \sigma_{j,i} e_j), \\
\Big| f_i(x_0 + t_j) - f_i(x_0 + t_{j-1}) - \tau_j \, \mathrm{D}_j f_i(x_0) \Big| & \\
\leq |\tau_j| \Big| \mathrm{D}_j f_i(x_0 + t_{j-1} + \sigma_{j,i} e_j) - \mathrm{D}_j f_i(x_0) \Big| &\leq |x - x_0| \frac{\varepsilon}{\sqrt{nm}}, \\
\Big| f_i(x) - f_i(x_0) - \sum_{j=1}^{m} \tau_j \, \mathrm{D}_j f_i(x_0) \Big| & \\
\leq \sum_{j=1}^{m} \Big| f_i(x_0 + t_j) - f_i(x_0 + t_{j-1}) - \tau_j \, \mathrm{D}_j f_i(x_0) \Big| &\leq |x - x_0| \frac{\varepsilon}{\sqrt{n}}, \\
\Big| f(x) - f(x_0) - L(x - x_0) \Big|^2 &= \sum_{i=1}^{n} \Big| f_i(x) - f_i(x_0) - \sum_{j=1}^{m} \mathrm{D}_j f_i(x_0) \tau_j \Big|^2 \\
= |x - x_0|^2 \frac{\varepsilon^2}{n} n &= |x - x_0|^2 \, \varepsilon^2,
\end{aligned}$$

und schließlich

$$|x - x_0|^{-1} \Big| f(x) - f(x_0) - L(x - x_0) \Big| \leq \varepsilon \text{ für } |x - x_0| < \delta,$$

was zu beweisen war. ■

Wir betrachten zwei *Spezialfälle*:

m = 1: Eine Funktion $f : \mathbb{R}^1 = \mathbb{R} \to \mathbb{R}^n$ ist gleichbedeutend mit n Funktionen einer Variablen. In diesem Fall ist

$$\mathrm{D}f(x) = f'(x) = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_n'(x) \end{pmatrix} = \Big(f_1'(x), \ldots, f_n'(x)\Big)^t,$$

und für $y \in \mathbb{R}$ ist $Df(x)y = \Big(cf_1'(x), \ldots, cf_n'(x)\Big)^t$.

n = 1: Eine Funktion von m Variablen. Hier ist

$$\mathrm{D}f(x) = \Big(\frac{\partial}{\partial \xi_1} f(x), \ldots, \frac{\partial}{\partial \xi_m} f(x)\Big) = \Big(\,\mathrm{D}_1 f(x), \ldots, \mathrm{D}_m f(x)\Big).$$

Für $y = (\eta_1, \ldots, \eta_m) \in \mathbb{R}^m$ ist

$$\mathrm{D}f(x_0)y = \sum_{j=1}^{m} \frac{\partial f}{\partial \xi_j}(x_0)\eta_j = \Big\langle\, \mathrm{D}f(x_0)^t, y\Big\rangle$$

(wobei allerdings häufig in dem letzten Skalarprodukt das t weggelassen wird). Der Vektor (die $m \times 1$–Matrix) $\mathrm{D}f(x_0)^t$ heißt der *Gradient* von f im Punkt x_0, kurz: $\operatorname{grad} f(x_0)$.

3.2 Gradient und Divergenz

Wie soeben definiert sei für $A \subset \mathbb{R}^m$ offen und eine differenzierbare Funktion $f : A \to \mathbb{R}$ der Gradient gegeben durch

$$\operatorname{grad} f(x) = \mathrm{D}f(x)^t = \begin{pmatrix} \mathrm{D}_1 f(x) \\ \vdots \\ \mathrm{D}_m f(x) \end{pmatrix}.$$

Satz 3.6 (Geometrische Deutung des Gradienten) *Sei $A \subset \mathbb{R}^m$ offen, $x_0 \in A$, und $f : A \to \mathbb{R}$ differenzierbar in x_0.*

a) *Die Richtungsableitung in Richtung e mit $|e| = 1$ ist $\langle \operatorname{grad} f(x_0), e\rangle$.*

b) *Ist $\operatorname{grad} f(x_0) = 0$, so sind alle Richtungsableitungen 0.*

c) *Ist* $\operatorname{grad} f(x_0) \neq 0$*, so zeigt der Vektor* $\operatorname{grad} f(x_0)$ *in die Richtung in der* f *am schnellsten wächst (größte Richtungsableitung), der Betrag* $|\operatorname{grad} f(x_0)|$ *von* $\operatorname{grad} f(x_0)$ *gibt die Richtungsableitung in dieser Richtung an.*

Beweis. Sei $e \in \mathbb{R}^m$, $|e| = 1$. Dann gilt für die Richtungsableitung in Richtung e im Punkt x_0

$$
\begin{aligned}
\frac{\partial}{\partial t} f(x_0 + te)\Big|_{t=0} &:= \lim_{t\to 0} \frac{1}{t}\Big(f(x_0+te) - f(x_0)\Big) \\
&= \lim_{t\to 0} \frac{1}{t}\Big(f(x_0) + \langle \operatorname{grad} f(x_0), te\rangle + \varphi(x_0, x_0+te) - f(x_0)\Big) \\
&\ (\text{wegen } \frac{1}{t}\varphi(x_0, x_0+te) \to 0 \text{ für } t \to 0) \\
&\overset{(*)}{=} \langle \operatorname{grad} f(x_0), e\rangle \\
&\leq |\operatorname{grad} f(x_0)| \\
&\ (\text{ falls } \operatorname{grad} f(x_0) \neq 0 \text{ gilt }) \\
&= \Big\langle \operatorname{grad} f(x_0), |\operatorname{grad} f(x_0)|^{-1} \operatorname{grad} f(x_0)\Big\rangle \\
&\ (\text{mit } (*) \text{ für } e = |\operatorname{grad} f(x_0)|^{-1} \operatorname{grad} f(x_0)) \\
&= \frac{\partial}{\partial t} f\Big(x_0 + t|\operatorname{grad} f(x_0)|^{-1} \operatorname{grad} f(x_0)\Big)\Big|_{t=0}.
\end{aligned}
$$

Bis zur Gleichung $(*)$ ist alles auch im Falle $\operatorname{grad} f(x_0) = 0$ durchführbar. Damit folgen die Teile a und b. Insgesamt haben wir auch gezeigt, daß die Richtungsableitung in Richtung des Gradienten gleich $|\operatorname{grad} f(x_0)|$ ist, und daß die Ableitung in dieser Richtung am größten ist; das ist Teil c. ■

Beispiel 3.7 Für $f : \mathbb{R}^m \to \mathbb{R}$, $f(x) = |x| = \{\xi_1^2 + \ldots + \xi_m^2\}^{1/2}$ ist

$$\operatorname{grad} f(x) = \frac{1}{|x|}(\xi_1, \ldots, \xi_m)^t = \frac{1}{|x|} x \quad \text{für } x \neq 0;$$

im Punkt $x = 0$ ist f nicht (partiell) differenzierbar. □

Da die (1–dimensionale) Produktregel sich in offensichtlicher Weise auf die partiellen Ableitungen von Produkten übertragen läßt, erhält man für differenzierbare Funktionen $f, g : \mathbb{R}^m \to \mathbb{R}$

$$\operatorname{grad}(fg)(x) = f(x) \operatorname{grad} g(x) + g(x) \operatorname{grad} f(x).$$

Ist $A \subset \mathbb{R}^m$, so wird eine Funktion $f : A \to \mathbb{R}^m$ als *Vektorfeld* bezeichnet. Falls f differenzierbar ist, definiert man die *Divergenz* von f durch

$$\operatorname{div} f(x) = \sum_{j=1}^{m} \mathrm{D}_j f_j(x) = \sum_{j=1}^{m} \frac{\partial f_j}{\partial \xi_j}(x).$$

(In Kapitel 10 (Die klassischen Integralsätze) werden wir die Divergenz als „Quellenstärke" erkennen.) Mit dem „formalen Vektor"

$$\nabla = \text{Nabla} := \begin{pmatrix} \partial/\partial\xi_1 \\ \vdots \\ \partial/\partial\xi_m \end{pmatrix} = \begin{pmatrix} \mathrm{D}_1 \\ \vdots \\ \mathrm{D}_m \end{pmatrix}$$

können wir abkürzend (formal) schreiben:

$$\begin{aligned} \operatorname{grad} f(x) &= \nabla f(x) && \text{für } f : \mathbb{R}^m \to \mathbb{R}, \\ \operatorname{div} f(x) &= \langle \nabla, f(x) \rangle && \text{für } f : \mathbb{R}^m \to \mathbb{R}^m. \end{aligned}$$

Ist $A \subset \mathbb{R}^m$ und sind $f : A \to \mathbb{R}^m$ und $g : A \to \mathbb{R}$ differenzierbar, so gilt

$$\begin{aligned} \operatorname{div}(gf)(x) &= \sum_{j=1}^{m} \mathrm{D}_j(gf_j) = \sum_{j=1}^{m} \Big\{ f_j \, \mathrm{D}_j g + g \, \mathrm{D}_j f_j \Big\} \\ &= \langle \operatorname{grad} g(x), f(x) \rangle + g(x) \operatorname{div} f(x). \end{aligned}$$

Beispiel 3.8 a) Für $f : \mathbb{R}^m \to \mathbb{R}^m$, $f(x) = x$ ist

$$\operatorname{div} f(x) = \sum_{j=1}^{m} \mathrm{D}_j \xi_j = \sum_{j=1}^{m} 1 = m.$$

b) Für $f : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}^m$, $f(x) = |x|^{-1} x$ ist für $x \neq 0$

$$\begin{aligned} \operatorname{div} f(x) &= \operatorname{div} \left(\frac{1}{|x|} \cdot x \right) = \Big\langle \operatorname{grad} \frac{1}{|x|}, x \Big\rangle + \frac{1}{|x|} \operatorname{div} x \\ &= -\Big\langle \frac{1}{|x|^3} x, x \Big\rangle + \frac{m}{|x|} = \frac{m-1}{|x|}. \end{aligned}$$

□

3.3 Differentiationsregeln

Für die Arbeit mit differenzierbaren Funktionen und explizite Berechnungen benötigen wir die folgenden Differentiationsregeln, deren Beweise weitgehend den 1–dimensionalen Beweisen entsprechen.

Satz 3.9 a) Linearität*: Ist $A \subset \mathbb{R}^m$ offen und sind $f, g : A \to \mathbb{R}^n$ differenzierbar in $x_0 \in A$, $a, b \in \mathbb{R}$, so ist auch $af + bg$ differenzierbar in x_0 und es gilt*

$$\mathrm{D}(af+bg)(x) = a\,\mathrm{D}f(x_0) + b\,\mathrm{D}g(x_0).$$

(Die entsprechenden Formeln gelten für die partiellen Ableitungen.)

b) Produktregel*: Ist $A \subset \mathbb{R}^m$ offen und sind $f : A \to \mathbb{R}^n, g : A \to \mathbb{R}$ differenzierbar in x_0, so ist auch gf differenzierbar in x_0 mit*

$$\begin{array}{ccccc} \mathrm{D}(gf)(x_0) & = & f(x_0)\cdot \mathrm{D}g(x_0) & + & g(x_0)\cdot \mathrm{D}f(x_0), \\ n\times m & & (n\times 1)\times(1\times m) & & ((1\times 1)\times(n\times m) \end{array}$$

die lineare Abbildung $\mathrm{D}(gf)(x_0) : \mathbb{R}^m \to \mathbb{R}^n$ ist gegeben durch

$$\mathrm{D}(gf)(x_0) : y \mapsto (\,\mathrm{D}g(x_0)y)f(x_0) + g(x_0)\,\mathrm{D}f(x_0)y,$$

speziell gilt für die partiellen Ableitungen

$$\mathrm{D}_j(gf)(x_0) = \mathrm{D}_j g(x_0)f(x_0) + g(x_0)\,\mathrm{D}_j f(x_0).$$

c) Kettenregel*: Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}^n$ differenzierbar in x_0, $B \subset \mathbb{R}^n$ offen, $f(A) \subset B$, $g : B \to \mathbb{R}^p$ differenzierbar in $f(x_0)$. Dann ist $h = g \circ f$ differenzierbar in x_0 und es gilt*

$$h(x_0) = \mathrm{D}g(f(x_0)) \cdot \mathrm{D}f(x_0) : \mathbb{R}^m \to \mathbb{R}^p.$$

Beweis. a) Nach Voraussetzung gilt

$$\begin{aligned} f(x) &= f(x_0) + \mathrm{D}f(x_0)(x-x_0) + \varphi_f(x_0,x), \\ g(x) &= g(x_0) + \mathrm{D}g(x_0)(x-x_0) + \varphi_g(x_0,x) \end{aligned}$$

mit $|x-x_0|^{-1}\varphi_f(x_0,x) \to 0$ und $|x-x_0|^{-1}\varphi_g(x-x_0) \to 0$ für $x \to x_0$. Damit folgt

$$(af+bg)(x) = (af+bg)(x_0) + (a\,\mathrm{D}f(x_0) + b\,\mathrm{D}g(x_0))(x-x_0) + \varphi(x_0,x)$$

mit

$$|x-x_0|^{-1}\varphi(x_0,x) = |x-x_0|^{-1}\Big\{a\varphi_f(x_0,x)+b\varphi_g(x_0,x)\Big\} \to 0 \text{ für } x\to x_0.$$

Damit folgt die Behauptung.

b) Für f und g gelten analoge Formeln wie in Teil a (wobei lediglich zu beachten ist, daß jetzt g seine Werte in $\mathbb{R}$ hat). Damit folgt

$$\begin{aligned} g(x)f(x) &= g(x_0)f(x_0) + (\,\mathrm{D}g(x_0)(x-x_0))f(x_0) \\ &\quad + g(x_0)\,\mathrm{D}f(x_0)(x-x_0) + \varphi(x_0,x), \end{aligned}$$

wobei $\varphi(x_0,x)$ aus Termen besteht, die mindestens einen Faktor $\varphi_f(x_0,x)$ bzw. $\varphi_g(x_0,x)$ enthalten, oder zweimal den Faktor $(x-x_0)$. Also gilt $|x-x_0|^{-1}\varphi(x_0,x)\to 0$ für $x\to x_0$. Damit ist die Behauptung erwiesen.

c) Es gilt

$$\begin{aligned} h(x) &= g(f(x)) \\ &= g(f(x_0)) + \mathrm{D}g(f(x_0))(f(x)-f(x_0)) + \varphi_g(f(x_0),f(x)) \\ &= h(x_0) + \mathrm{D}g(f(x_0))\Big\{\,\mathrm{D}f(x_0)(x-x_0)+\varphi_f(x_0,x)\Big\} + \varphi_g(f(x_0),f(x)) \\ &= h(x_0) + \mathrm{D}g(f(x_0))\,\mathrm{D}f(x_0)(x-x_0) \\ &\quad + \mathrm{D}g(f(x_0))\varphi_f(x_0,x) + \varphi_g(f(x_0),f(x)) \\ &= h(x_0) + \mathrm{D}g(f(x_0))\,\mathrm{D}f(x_0)(x-x_0) + \varphi(x_0,x) \end{aligned}$$

mit

$$\varphi(x_0,x) = \mathrm{D}_g(f(x_0))\varphi_f(x_0,x) + \varphi_g(f(x_0),f(x)).$$

Es bleibt $|x-x_0|^{-1}\varphi(x_0,x)\to 0$ für $x\to x_0$ zu zeigen. Für den ersten Term gilt

$$\frac{1}{|x-x_0|}\mathrm{D}g(f(x_0))\varphi_f(x_0,x) = \mathrm{D}g(f(x_0))\frac{\varphi_f(x_0,x)}{|x-x_0|} \to 0 \text{ für } x\to x_0.$$

Für den zweiten Term gilt

$$\frac{1}{|x-x_0|}\varphi_g(f(x_0),f(x)) = \begin{cases} 0 & \text{falls } f(x)=f(x_0), \\ \dfrac{\varphi_g(f(x_0),f(x))}{|f(x)-f(x_0)|}\dfrac{|f(x)-f(x_0)|}{|x-x_0|} & \text{falls } f(x)\neq f(x_0). \end{cases}$$

Wegen

$$\begin{aligned}
&\frac{\varphi_g(f(x_0), f(x))}{|f(x) - f(x_0)|} \to 0 \text{ für } x \to x_0 \quad (\text{also } f(x) \to f(x_0)),\\
&\frac{|f(x) - f(x_0)|}{|x - x_0|} \leq \underbrace{\frac{|\,\mathrm{D}f(x_0)(x - x_0)|}{|x - x_0|}}_{\text{beschränkt}} + \underbrace{\frac{\varphi_f(x_0, x)}{|x - x_0|}}_{\to 0} \quad \text{beschränkt für } x \to x_0
\end{aligned}$$

konvergiert auch dieser Ausdruck gegen 0 für $x \to x_0$. ∎

Korollar 3.10 *Aus Teil c des obigen Satzes ergibt sich für die einzelnen partiellen Ableitungen von $h = g \circ f$:*

$$\mathrm{D}_j h_i(x) = \sum_{\ell=1} \mathrm{D}_\ell g_i(f(x))\, \mathrm{D}_j f_\ell(x).$$

3.4 Mittelwertsätze

Wir beweisen nun noch verschiedene Verallgemeinerungen des 1–dimensionalen Mittelwertsatzes der Differentialrechnung.

Satz 3.11 (Mittelwertsatz) *Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}$ differenzierbar, $x, y \in A$ und die Verbindungsstrecke zwischen x und y*

$$\Big\{x + t(y - x) : 0\Big\}$$

liege ebenfalls in A. Dann existiert ein $\vartheta \in (0, 1)$ mit

$$\begin{aligned}
f(y) - f(x) &= \mathrm{D}f(x + \vartheta(y - x))(y - x)\\
&= \Big\langle \operatorname{grad} f(x + \vartheta(y - x)), y - x \Big\rangle
\end{aligned}$$

($x + \vartheta(y - x)$st ein Punkt zwischen x und y, Zwischenpunkt*).*

Beweis. Wir betrachten

$$g : [0,1] \to \mathbb{R}^m, \qquad g(t) := x + t(y - x)$$

und

$$h := f \circ g : [0,1] \to \mathbb{R}.$$

Dann ist h stetig auf $[0,1]$ und differenzierbar in $(0,1)$ mit

$$h'(t) = \mathrm{D}h(t) = \mathrm{D}f(g(t))\,\mathrm{D}g(t) = \Big\langle \operatorname{grad} f(g(t)), y - x \Big\rangle.$$

Nach dem 1–dimensionalen Mittelwertsatz der Differentialrechnung gilt also

$$\begin{aligned} f(y) - f(x) &= h(1) - h(0) = h'(\vartheta) \cdot 1 \\ &= \Big\langle \operatorname{grad} f(x + \vartheta(y - x)), y - x \Big\rangle. \end{aligned}$$

mit einem $\delta \in (0,1)$. ■

Korollar 3.12 a) *Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}^n$ differenzierbar und*

$$|\,\mathrm{D}_j f_i(x)| \le M \text{ für alle } x \in A,\ 1 \le j \le m,\ 1 \le i \le n.$$

Dann gilt für alle x und $y \in A$, deren Verbindungsstrecke ebenfalls in A liegt,

$$|f(y) - f(x)| \le M\sqrt{nm}|y - x|.$$

b) *Sei $A \subset \mathbb{R}^m$ offen und* zusammenhängend *(d. h. zu beliebigen $x, y \in A$ existieren Punkte $x_1, x_2, \ldots, x_\ell \in A$ so, daß alle Verbindungsstrecken $\overline{xx_1}, \overline{x_1x_2}, \ldots, \overline{x_{\ell-1}x_\ell}, \overline{x_\ell y}$ in A liegen; vgl. auch Aufgabe 3.8). Ist $f : A \to \mathbb{R}^m$ differenzierbar mit $\mathrm{D}f(x) = 0$ für alle $x \in A$, so ist f konstant in A.*

c) *Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}^n$ stetig differenzierbar, $x, y \in A$ und die Verbindungsstrecke von x nach y liege ebenfalls in A. Dann existiert ein ξ zwischen x und y ($\xi = x + \vartheta(y - x)$ mit $\vartheta \in (0,1)$) so, daß gilt*

$$|f(y) - f(x)| \le |\,\mathrm{D}f(\xi)||x - y|$$

(dabei ist für eine lineare Abbildung $L : \mathbb{R}^m \to \mathbb{R}^n$ mit $|L| = \sup\{|Lx| : x \in \mathbb{R}^m, |x| \le 1\}$ die Norm *der linearen Abbildung L gemeint). Insbesondere gilt $|f(y) - f(x)| \le C|x - y|$ mit $C = \sup\{|\,\mathrm{d}f(\xi)| : \xi$ zwischen X und $Y\}$*

Beweis. a) Nach dem eben bewiesenen Mittelwertsatz gibt es für jedes $i \in \{1, \cdots, n\}$ ein $\vartheta_i \in (0,1)$ mit

$$\begin{aligned}|f_i(y) - f_i(x)| &= \left|\Big\langle \operatorname{grad} f_i(x + \vartheta_i(y-x)), (y-x)\Big\rangle\right| \\ &\leq \sum_{j=1}^{m} M|\eta_j - \xi_j| \leq M\sqrt{m}|y-x|\end{aligned}$$

Damit folgt

$$|f(y) - f(x)| = \Big\{\sum_{i=1}^{n} |f_i(y) - f_i(x)|^2\Big\}^{1/2} \leq M\sqrt{nm}|y-x|.$$

b) Aus Teil a folgt mit $M = 0$

$$f(y) = f(x_\ell) = \ldots = f(x_1) = f(x).$$

c) *Spezialfall*: Es gibt ein $k \in \{1, \ldots, n\}$ mit $f_i(x) = f_i(y)$ für $i \neq k$. Dann folgt aus dem obigen Mittelwertsatz mit einem z zwischen x und y

$$\begin{aligned}|f(y) - f(x)| &= |f_k(y) - f_k(x)| = |\,\mathrm{D}f_k(z)(y-x)| \\ &\leq |\,\mathrm{D}f_k(z)||y-x| \leq |\,\mathrm{D}f(z)||y-x|.\end{aligned}$$

Allgemeiner Fall. Durch geeignete Translation T (wobei einer der Punkte $f(x)$ oder $f(y)$ in den 0–Punkt verschoben wird) und Drehung U in $\mathbb{R}^n$ kann erreicht werden, daß für $\tilde{f} := UTf$ der obige Spezialfall eintritt, d. h. es gilt mit einem z zwischen x und y

$$\begin{aligned}|f(y) - f(x)| &= |\tilde{f}(x) - \tilde{f}(y)| \leq |\,\mathrm{D}\tilde{f}(z)||y-x||\,\mathrm{D}(UTf)(z)||y-x| \\ &= |U\,\mathrm{D}Tf(z)||y-x| = |U\,\mathrm{D}f(z)||y-x| = |\,\mathrm{D}f(z)||y-x|,\end{aligned}$$

da für jede lineare Abbildung L gilt $|UL| = |L|$. ■

3.5 Übungsaufgaben

3.1 Man zeige: Die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) = \begin{cases} (x^2+y^2)\sin\Big\{\dfrac{1}{(x^2+y^2)^{1/2}}\Big\} & \text{für } (x,y) \neq 0, \\ 0 & \text{für } (x,y) = 0 \end{cases}$$

ist differenzierbar, aber nicht stetig differenzierbar.

3.2 Man konstruiere eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, die richtungsdifferenzierbar ist, aber nicht differenzierbar.
Anleitung: Man kann z. B. das Beispiel 2.7 für eine Funktion, die richtungsstetig ist, aber nicht stetig, entsprechend modifizieren.

3.3 a) Die Abbildung

$$f : (0,\infty) \times (0,\pi) \times [0,2\pi) \to \mathbb{R}^3 \backslash \{(0,0,\xi_3) : \xi_3 \in \mathbb{R}\},$$
$$f(r,\vartheta,\varphi) = (r\sin\vartheta\cos\varphi, r\sin\vartheta\sin\varphi, r\cos\vartheta)$$

ist bijektiv (*Polarkoordinaten* oder *Kugelkoordinaten*); man beschreibe die geometrische Bedeutung von r, ϑ, φ in Bezug auf $x = f(r,\vartheta,\varphi)$ (Skizze).

b) Man berechne die Jacobi–Matrix und deren Determinante.

3.4 Sei $g : [0,\infty) \to \mathbb{R}$; die Funktion $f : \mathbb{R}^m \to \mathbb{R}$ sei definiert durch

$$f(x) = g(|x|) \text{ für alle } x \in \mathbb{R}^m$$

(eine solche Funktion f heißt *sphärisch symmetrisch*).
Man zeige: f ist genau dann (überall) differenzierbar, wenn g in $(0,\infty)$ differenzierbar ist und in 0 die rechtsseitige Ableitung $g'(0) = 0$ hat. Man berechne den Gradienten von f.

3.5 Sei $A \subset \mathbb{R}^m$ offen und zusammenhängend.

a) Ist $f : A \to \mathbb{R}$ stetig, so ist $f(A)$ ein Intervall.

b) Ist $f : A \to \mathbb{R}$ stetig differenzierbar und $Df(x) \neq 0$ für alle $x \in A$, so ist $f(A)$ offen.

c) Teil b gilt ohne die Forderung $Df(x) \neq 0$ i. allg. nicht.

3.6 Sei $f : \mathbb{R}^2 \to \mathbb{R}$ stetig, $-\infty < a < b < \infty$, $F : \mathbb{R} \to \mathbb{R}$, $\quad F(x) := \int\limits_a^b f(x,y)\,\mathrm{d}y$.

a) F ist stetig.

b) Ist f stetig differenzeirbar, so ist auch F stetig differenzierbar mit

$$F'(x) = \int\limits_a^b \frac{\partial}{\partial x} f(x,y)\,\mathrm{d}y.$$

3.7 a) Sind $f, g : \mathbb{R}^m \to \mathbb{R}^n$ differenzierbar mit $Df = Dg$ und $f(x_0) = g(x_0)$ für ein $x_0 \in \mathbb{R}^m$, so gilt $f \equiv g$.

b) Ist $f : \mathbb{R}^2 \to \mathbb{R}$ stetig, so ist

$$g : \mathbb{R}^2 \to \mathbb{R}, \quad g(x, y) = \int\limits_{x_0}^{x} f(s, y)\,\mathrm{d}s$$

stetig.

c) Ist $f : \mathbb{R}^2 \to \mathbb{R}$ stetig, so gilt

$$\int\limits_{y_0}^{y} \left\{ \int\limits_{x_0}^{x} f(s, t)\,\mathrm{d}s \right\} \mathrm{d}t = \int\limits_{x_0}^{x} \left\{ \int\limits_{y_0}^{y} f(s, t)\,\mathrm{d}t \right\} \mathrm{d}s.$$

Anleitung: Beide Seiten sind (total) differenzierbar mit gleicher Ableitung.

3.8 Eine offene Teilmenge $A \subset \mathbb{R}^m$ ist genau dann zusammenhängend im Sinne dieser Vorlesung (vgl. Korollar 3.12 b: d. h. zwei beliebige Punkte aus A könnenn durch einen Polygonzug in A verbunden werden), wenn aus B, C offen, $B \cap C = \emptyset$ und $A = B \cup C$ folgt: $B = \emptyset$ oder $C = \emptyset$ (bzw. $A = A$ oder $C = A$).

4 Höhere Ableitungen und der Taylorsche Satz

4.1 Höhere Ableitungen

Sei $A \subset \mathbb{R}^m$ offen. Eine Funktion $f : A \longrightarrow \mathbb{R}^n$ heißt *zweimal stetig differenzierbar*, wenn f stetig differenzierbar ist und alle partiellen Ableitungen $\mathrm{D}_j f : A \longrightarrow \mathbb{R}^n$ ebenfalls stetig differenzierbar sind. Die partiellen Ableitungen $\mathrm{D}_i \mathrm{D}_j f = \mathrm{D}_i(\mathrm{D}_j f)$ von $\mathrm{D}_j f$ heißen zweite partielle Ableitungen von f. Entsprechend definiert man induktiv p, *mal stetig differenzierbar* $(p > 2)$ und die *p–ten partiellen* Ableitungen $\mathrm{D}_{j_1} \mathrm{D}_{j_2} \dots \mathrm{D}_{j_p} f$, wobei es zunächst durchaus auf die Reihenfolge der D_{j_k} ankommt. Der folgende Satz zeigt, daß diese Reihenfolge „in der Regel“ doch vertauschbar ist.

Satz 4.1 *Sei $A \subset \mathbb{R}^m$ offen und $f : A \longrightarrow \mathbb{R}^n$ zweimal stetig differenzierbar. Dann gilt*

$$\mathrm{D}_i \mathrm{D}_j f = \mathrm{D}_j \mathrm{D}_i f \text{ für alle } i, j = 1, 2, \dots, m$$

(kurz: die Differentiationsreihenfolge ist vertauschbar).

Beweis. Wir reduzieren das Problem schrittweise auf einen (zumindest formal) besonders einfachen Spezialfall. Zunächst ist klar, daß es genügt, eine einzelne Komponente von f zu betrachten,

$$\mathrm{D}_i \mathrm{D}_j f_k = \mathrm{D}_j \mathrm{D}_i f_k \text{ für } k = 1, \dots, n,$$

d. h. wir dürfen ohne weiteres annehmen, daß $n = 1$ gilt.

Für $i = j$ ist offenbar nichts zu beweisen. Es genügt also o. E. den Fall $i = 1$ *und* $j = 2$ zu betrachten (andernfalls können wir die Koordinaten umnumerieren). Da die weiteren Koordinaten bei diesen Überlegungen außer Betracht bleiben, können wir o. E. annehmen, daß $m = 2$ gilt.

Weiterhin können wir unsere Untersuchungen o. E. *im Nullpunkt* durchführen (andernfalls führen wir eine geeignete Translation durch). All dies erspart uns unnötige Schreibarbeit.

Ingesamt genügt es also, für $A \subset \mathbb{R}^2$ offen mit $0 = (0, 0) \in A$, $f : A \to \mathbb{R}$ zweimal stetig differenzierbar zu zeigen:

$$\mathrm{D}_1 \mathrm{D}_2 f(0, 0) = \mathrm{D}_2 \mathrm{D}_1 f(0, 0).$$

Da A offen ist, gibt es eine Kugel $K_r(0) \subset A$. Die folgenden Überlegungen sind deshalb für $s^2 + t^2 < r^2$ durchführbar. Für beliebige s, t mit $s^2 + t^2 < r^2$ gilt

$$
\begin{aligned}
f(s,t) - f(s,0) - f(0,t) + f(0,0) &= \Big(f(s,t) - f(s,0)\Big) - \Big(f(0,t) - f(0,0)\Big) \\
&\quad \text{(mit } g_t(x) := f(x,t) - f(x,0)) \\
&= g_t(s) - g_t(0) \\
&\quad \text{(Mittelwertsatz der 1–dim. Diff.– u. Int.–R. mit } s_1 = s_1(s,t) \text{ zwischen 0 und } s) \\
&= g_t'(s_1)(s-0) = s\Big(\mathrm{D}_1 f(s_1,t) - \mathrm{D}_1 f(s_1,0)\Big) \\
&\quad \text{(mit } h_{s_1} := \mathrm{D}_1 f(s_1,x)) \\
&= s(h_{s_1}(t) - h_{s_1}(0)) \\
&\quad \text{(Mittelwertsatz der 1–dim. Diff.– u. Int.–R. mit } t_1 = t_1(s,t) \text{ zwischen 0 und } t) \\
&= s t h_{s_1}'(t_1) = st\,\mathrm{D}_2\,\mathrm{D}_1 f(s_1,t_1).
\end{aligned}
$$

Völlig analog zeigt man

$$
\begin{aligned}
f(s,t) - f(s,0) - f(0,t) + f(0,0) &= \Big(f(s,t) - f(0,t)\Big) - \Big(f(s,0) - f(0,0)\Big) \\
&= t\Big(\mathrm{D}_2 f(s,t_2) - \mathrm{D}_2 f(0,t_2)\Big) = ts\,\mathrm{D}_1\,\mathrm{D}_2 f(s_2,t_2)
\end{aligned}
$$

mit $s_2 = s_2(s,t)$ zwischen 0 und s, $t_2 = t_2(s,t)$ zwischen 0 und t. Damit ergibt sich für alle s, t mit $s \neq 0$,$t \neq 0$ und $s^2 + t^2 < r^2$

$$\mathrm{D}_2\,\mathrm{D}_1 f(s_1,t_1) = \mathrm{D}_1\,\mathrm{D}_2 f(s_2,t_2)$$

für geeignete s_1, s_2) zwischen 0 und s sowie t_1, t_2 zwischen 0 und t. Für $s, t \to 0$ gilt $s_j, t_j \to 0$, und somit folgt für $s, t \to 0$

$$\mathrm{D}_2\,\mathrm{D}_1 f(0,0) = \mathrm{D}_1\,\mathrm{D}_2 f(0,0)$$

wegen der Stetigkeit von $D_1 D_2 f$ und $D_2 D_1 f$. ■

Korollar 4.2 *Ist $A \subset \mathbb{R}^m$ offen und $f : A \to \mathbb{R}^n$ p mal stetig differenzierbar, so können in $\mathrm{D}_{j_1} \ldots \mathrm{D}_{j_p} f$ die partiellen Ableitungen beliebig vertauscht werden.*

Für $A \subset \mathbb{R}^m$ offen und $p \in \mathbb{N}$ sei $C^p(A, \mathbb{R}^n)$ der Vektorraum der p mal stetig differenzierbaren Funktionen $f : A \to \mathbb{R}^n$. Für $f \in C^p(A, \mathbb{R}^n), q \in \mathbb{N}$ mit $q \leq p$ definieren wir

$$\mathrm{D}_j^q f := \underbrace{\mathrm{D}_j \ldots \mathrm{D}_j}_{q \text{ mal}} f \qquad (\text{speziell } \mathrm{D}_j^0 f := f).$$

Für $\{j_1, \ldots, j_q\} \in \{1, \ldots, m\}^q$ bezeichne α_i die Anzahl wie oft die Zahl i unter den j_k vorkommt. Dann gilt für $f \in C^p(A, \mathbb{R}^n)$ mit $p \geq q$

$$\mathrm{D}_{j_1} \ldots \mathrm{D}_{j_q} f = \mathrm{D}_1^{\alpha_1} \ldots \mathrm{D}_m^{\alpha_m} f.$$

Für *Multiindizes* $\alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}_0^m$ sei im folgenden

$$\begin{aligned}
|\alpha| &:= \sum_{j=1}^m \alpha_j \qquad \text{der } \textit{Betrag} \text{ bzw. die } \textit{Länge} \text{ von } \alpha, \\
\alpha! &:= \prod_{j=1}^m \alpha_j! \qquad \text{die } \textit{Fakultät} \text{ von } \alpha, \\
x^\alpha &:= \xi_1^{\alpha_1} \ldots \xi_m^{\alpha_m} = \prod_{j=1}^m \xi_j^{\alpha_j} \qquad \text{für } x = (\xi_1, \ldots, \xi_m) \in \mathbb{R}^m, \\
\mathrm{D}^\alpha f &:= \mathrm{D}_1^{\alpha_1} \ldots \mathrm{D}_m^{\alpha_m} f = \left(\prod_{j=1}^m \mathrm{D}_j^\alpha \right) f \qquad \text{für } f \in C^p(A, \mathbb{R}^n), |\alpha| \leq p.
\end{aligned}$$

4.2 Der Taylorsche Satz

Mit den Resultaten und Bezeichnungen des vorigen Abschnitts sind wir nun in der Lage, den aus Analysis I bekannten Taylorschen Satz auf Funktionen von m Variablen zu übertragen.

Satz 4.3 (Taylorscher Satz) *Sei $A \subset \mathbb{R}^m$ offen, $p \geq 0, f \in C^{p+1}(A, \mathbb{R}), x_0 \in A$. Dann gilt für alle $x \in A$, für die die Verbindungsstrecke von x_0 und x in A liegt*

$$f(x) = \sum_{|\alpha| \leq p} \frac{1}{\alpha!} \mathrm{D}^\alpha f(x_0)(x - x_0)^\alpha + R_p(x)$$

mit

$$R_p(x) = \sum_{|\alpha| = p+1} \frac{1}{\alpha!} \mathrm{D}^\alpha f\Big(x_0 + \vartheta(x - x_0)\Big)(x - x_0)^\alpha \qquad \textit{mit einem} \quad \vartheta \in (0, 1).$$

Zur Verdeutlichung geben wir dieses Resultat für $p = 0, 1, 2$ expliziter an:

$$\mathbf{p = 0:} \qquad f(x) = f(x_0) + \sum_{j=1}^{m} \mathrm{D}_j f\Big(x_0 + \vartheta(x - x_0)\Big)(\xi_j - \xi_{0,j}),$$

$$\begin{aligned}\mathbf{p = 1:} \qquad f(x) &= f(x_0) + \sum_{j=1}^{m} \mathrm{D}_j f(x_0)(\xi_j - \xi_{0,j}) \\ &\quad + \frac{1}{2} \sum_{j,k=1}^{m} \mathrm{D}_j \mathrm{D}_k f\Big(x_0 + \vartheta(x - x_0)\Big)(\xi_j - \xi_{0,j})(\xi_k - \xi_{0,k}),\end{aligned}$$

$$\begin{aligned}\mathbf{p = 2:} \qquad f(x) &= f(x_0) + \sum_{j=1}^{m} \mathrm{D}_j f(x_0)(\xi_j - \xi_{0,j}) \\ &\quad + \frac{1}{2} \sum_{j,k=1}^{m} \mathrm{D}_j \mathrm{D}_k f(x_0)(\xi_j - \xi_{0,j})(\xi_k - \xi_{0,k}) + R_2(x).\end{aligned}$$

Es sei noch angemerkt, daß die Terme zweiter Ordnung als quadratische Form der *Hesse–Matrix* $H(\cdot) = \Big(\mathrm{D}_j \mathrm{D}_k f(\cdot)\Big)_{j,j=1,\ldots,m}$ aufgefaßt werden können, z. B.

$$\sum_{j,k=1}^{m} \mathrm{D}_j \mathrm{D}_k f(x_0)(\xi_j - \xi_{0,j})(\xi_k - \xi_{0,k}) = \Big\langle H(x_0)(x - x_0), x - x_0 \Big\rangle.$$

Beweis. Der Satz wird bewiesen durch Rückführung auf den 1–dimensionalen Fall. Dazu sei

$$\begin{aligned}&y = (\eta_1, \ldots, \eta_m) := (x - x_0), \\ &g : [-\varepsilon, 1 + \varepsilon] \to \mathbb{R}, \\ &g(t) := f(x_0 + ty) = (f \circ h)(t) \quad \text{mit} \quad h(t) := x_0 + ty,\end{aligned}$$

wobei $\varepsilon > 0$ so klein gewählt sei, daß das Geradenstück $\{x_0 + ty : -\varepsilon \leq t \leq 1 + \varepsilon\}$ ganz in A liegt) (ein solches ε existiert, weil A offen ist und die Verbindungsstrecke zwischen x_0 und $x = x_0 + y$ in A liegt). Offensichtlich ist g stetig differenzierbar und mit der Kettenregel folgt

$$\begin{aligned}g'(t) &= \mathrm{D}g(t) = \mathrm{D}f(h(t))\,\mathrm{D}h(t) = \sum_{j=1}^{m} \mathrm{D}_j f(h(t)) \underbrace{\frac{d}{\mathrm{d}t} h_j(t)}_{=\eta_j} \\ &= \sum_{j=1}^{m} \eta_j \mathrm{D}_j f(x_0 + ty).\end{aligned}$$

Ist $p \geq 1$, so ist auch g' stetig differenzierbar und es gilt

$$\begin{aligned} g''(t) &= \mathrm{D}g'(t) = \sum_{j=1}^m \eta_j \frac{d}{\mathrm{d}t} \mathrm{D}_j f(x_0 + ty) \\ &\qquad \text{(mit der gleichen Rechnung wie oben)} \\ &= \sum_{j=1}^m \eta_j \sum_{i=1}^m \eta_i \mathrm{D}_i \mathrm{D}_j f(x_0+ty) = \Big(\sum_{j=1}^m \eta_j \mathrm{D}_j\Big)^2 f(x_0+ty), \end{aligned}$$

wobei der Term $\Big(\sum_{j=1}^m \eta_j \mathrm{D}_j\Big)^2$ formal auszumultiplizieren ist. Allgemeiner erhält man durch induktive Anwendung dieses Verfahrens für $k \leq p+1$

$$g^{(k)}(t) = \Big(\sum_{j=1}^m \eta_j \mathrm{D}_j\Big)^k f(x_0+ty),$$

wobei auch hier die k–te Potenz formal auszurechnen ist.

Nach dem *Polynomischen Satz* gilt für beliebige $z = (\zeta_1, \ldots, \zeta_m) \in \mathbb{R}^m$ und $k \in \mathbb{N}$ [1]

$$\Big(\sum \zeta_j\Big)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} z^\alpha.$$

In dieser Formel kann natürlich $z = (\xi_1, \ldots, \xi_m)$ durch $(\eta_1 \mathrm{D}_1, \ldots, \eta_m \mathrm{D}_m)$ ersetzt werden. Damit erhalten wir

$$\begin{aligned} \frac{1}{k!} g^{(k)}(t) &= \frac{1}{k!}\Big(\sum_{j=1}^m \eta_j \mathrm{D}_j\Big)^k f(x_0+ty) \\ &= \Big(\sum_{|\alpha|=k} \frac{(\eta_1 \mathrm{D}_1, \ldots, \eta_m \mathrm{D}_m)^\alpha}{\alpha!}\Big) f(x_0+ty) = \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \mathrm{D}^\alpha f(x_0+ty). \end{aligned}$$

[1] *Beweis* durch Induktion nach m. Offensichtlich richtig für $m=1$. Sei die Formel richtig für $m-1$. Dann folgt

$$\begin{aligned} \Big(\sum_{j=1}^m \xi_j\Big)^k &= \Big[\Big(\sum_{j=1}^{m-1} \xi_j\Big) + \xi_m\Big]^k \qquad \text{(binomische Formel)} \\ &= \sum_{\alpha_m=0}^k \binom{k}{\alpha_m} \Big(\sum_{j=1}^{m-1} \xi_j\Big)^{k-\alpha_m} \xi_m^{\alpha_m} \qquad (\text{Induktionsannahme}, m-1) \\ &= \sum_{\alpha_m=0}^k \frac{k!}{\alpha_m!(k-\alpha_m)!} \sum_{|\beta|=k-\alpha_m} \frac{(k-\alpha_m)!}{\beta!} (\xi_1, \cdots, \xi_{m-1})^\beta \xi_m^{\alpha_m} \qquad (\beta \in \mathbb{N}_0^{m-1}) \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} z^\alpha \qquad (\alpha = (\beta, \alpha_m)). \end{aligned}$$

Die Taylorsche Formel für g liefert für $t \in [0,1]$ mit einem $\delta \in (0,1)$

$$g(t) = \sum_{k=0}^{p} \frac{1}{k!} g^{(k)}(0) t^k + \tilde{R}_p(t), \quad \tilde{R}_p(t) = \frac{1}{(p+1)!} t^p g^{(p+1)}(\vartheta t),$$

also

$$\begin{aligned} f(x) &= g(1) = \sum_{k=0}^{p} \frac{1}{k!} g^{(k)}(0) + \tilde{R}_p(1) \\ &= \sum_{k=0}^{p} \sum_{|\alpha|=k} \frac{1}{\alpha!} (x-x_0)^\alpha \, \mathrm{D}^\alpha f(x_0) + \tilde{R}_p(1) \\ &= \sum_{|\alpha| \le p} \frac{1}{\alpha!} \, \mathrm{D}^\alpha f(x_0)(x-x_0)^\alpha + R_p(x) \end{aligned}$$

mit

$$\begin{aligned} R_p(x) &= \tilde{R}_p(1) = \frac{1}{(p+1)!} g^{(p+1)}(\vartheta) = \sum_{|\alpha|=p+1} \frac{y^\alpha}{\alpha!} \, \mathrm{D}^\alpha f(x_0 + \vartheta y) \\ &= \sum_{|\alpha|=p+1} \frac{1}{\alpha!} \, \mathrm{D}^\alpha f\Big(x_0 + \vartheta(x-x_0)\Big)(x-x_0)^\alpha. \end{aligned}$$

Das ist das behauptete Taylorpolynom mit dem entsprechenden Restglied. ■

4.3 Lokale Extrema differenzierbarer Funktionen

Wie im 1–dimensionalen ist der Tasylorsche Satz unter anderem geeignet, Bedingungen für das Vorliegen lokaler Extrema (im Inneren des Gebietes A) anzugeben. Die folgenden Definitionen sind völlig analog zu denen im 1–dimensionalen:

Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}, x_0 \in A$. Man sagt, f hat in x_0 ein *lokales Maximum* (*lokales Minimum*), wenn ein $\delta > 0$ existiert mit

$$f(x) \le f(x_0) \quad \Big(f(x) \ge f(x_0)\Big) \qquad \text{für alle} \quad x \in A \quad \text{mit} \quad |x - x_0| < \delta;$$

f hat in x_0 ein *strenges lokales Maximum* (*strenges lokales Minimum*), wenn ein $\delta > 0$ existiert mit

$$f(x) < f(x_0) \quad \Big(f(x) > f(x_0)\Big) \qquad \text{für alle} \quad x \in A \quad \text{mit} \quad 0 < |x - x_0| < \delta.$$

Satz 4.4 (Notwendige Bedingung für Extremum) *Sei $A \subset \mathbb{R}^m$ offen, $x_0 \in A$ und die Funktion $f : A \to \mathbb{R}$ habe in x_0 ein lokales Maximum (bzw. lokales Minimum).*

a) *Ist f differenzierbar im Punkt x_0, so gilt $\mathrm{D}f(x_0) = 0$ (d. h. $\operatorname{grad} f(x_0) = 0$ oder gleichbedeutend $\mathrm{D}_j f(x_0) = 0$ für $j = 1, \cdots, m$).*

b) *Ist f zweimal stetig differenzierbar, so gilt $\mathrm{D}f(x_0) = 0$ und die Hessematrix $(\mathrm{D}_i\, \mathrm{D}_j f(x_0))_{i,j=1,\cdots,m}$ ist negativ semidefinit (bzw. positiv semidefinit).*

Beweis. a) Die Funktionen $f_j : (-\varepsilon, \varepsilon) \to \mathbb{R}$, $g_j(t) = f(x_0 + te_j)$ ist bei $t = 0$ differenzierbar und hat dort ein lokales Extremum. Also gilt $D_j f(x_0) = g'(0) = 0$.

b) Nach Teil a gilt jedenfalls $Df(x_0) = 0$. Nehmen wir an, daß $(\mathrm{D}_i\, \mathrm{D}_j f(x_0))$ *nicht* negativ semidefinit ist. Dann existiert ein $y \in \mathbb{R}^m$ mit $|y| = 1$ und

$$\Big\langle \Big(\mathrm{D}_i\, \mathrm{D}_j f(x_0)\Big) y, y \Big\rangle > 0.$$

Da f zweimal stetig differenzierbar ist, existiert ein $\delta > 0$ so, daß

$$\Big\langle \Big(\mathrm{D}_i\, \mathrm{D}_j f(x)\Big) y, y \Big\rangle > 0 \qquad \text{für alle} \quad x \in A \quad \text{mit} \quad |x - x_0| < \delta.$$

Damit folgt aus dem Taylorschen Satz für alle $t \in (0, \delta]$:

$$\begin{aligned} f(x_0 + ty) &= f(x_0) + 0 + \sum_{|\alpha|=2} \frac{1}{\alpha!} \mathrm{D}^\alpha f(x_0 + \vartheta t y) t^{|\alpha|} y^\alpha \\ &= f(x_0) + \frac{1}{2} \sum_{i,j=1}^{m} t^2\, \mathrm{D}_i\, \mathrm{D}_j f(x_0 + \vartheta t y) \eta_i \eta_j \\ &= f(x_0) + \frac{1}{2} t^2 \Big\langle \Big(\mathrm{D}_i\, \mathrm{D}_j f(x_0 + \vartheta t y)\Big) y, y \Big\rangle > f(x_0), \end{aligned}$$

im Widerspruch zur Voraussetzung, daß bei x_0 ein lokales Maximum vorliegt. ∎

Satz 4.5 (Hinreichende Bedingung für Extremum) *Sei $A \subset \mathbb{R}^m$ offen, $f \in C^2(A, \mathbb{R})$, $x_0 \in A$. Gilt $\mathrm{D}f(x_0) = 0$ und ist die Hesse–Matrix $\Big(\mathrm{D}_i\, \mathrm{D}_j f(x_0)\Big)_{i,j=1,\ldots,m}$ negativ definit (positiv definit), so hat f bei x_0 ein strenges lokales Maximum (strenges lokales Minimum).*

Beweis. Sei $\mathrm{D}f(x_0) = 0$ und $\Big(\mathrm{D}_i\,\mathrm{D}_j f(x_0)\Big)$ negativ definit. Es gilt also

$$\sum_{i,j=1}^{m} \mathrm{D}_i\,\mathrm{D}_j f(x_0)\eta_i\eta_j < 0 \qquad \text{für alle} \quad y \in \mathbb{R}^m \quad \text{mit} \quad |y| = 1.$$

Da die Einheitsshpäre in $\mathbb{R}^m$ kompakt ist, nimmt die stetige Funktion $y \mapsto \sum\limits_{i,j=1}^{m} \mathrm{D}_i\,\mathrm{D}_j f(x_0)\eta_i\eta_j$ dort ihr Maximum an, und der Wert in diesem Maximum ist notwendig negativ, d. h. es existiert ein $\lambda > 0$ mit

$$\sum_{i,j=1}^{m} \mathrm{D}_1\,\mathrm{D}_j f(x_0)\eta_i\eta_j \leq -\lambda \qquad \text{für alle} \quad y \in \mathbb{R}^m \quad \text{mit} \quad |y| = 1.$$

Da alle $\mathrm{D}_i\,\mathrm{D}_j f(\cdot)$ stetig sind, existiert ein $\delta > 0$ mit

$$\Big|\mathrm{D}_i\,\mathrm{D}_j f(x) - \mathrm{D}_i\,\mathrm{D}_j f(x_0)\Big| \leq \frac{\lambda}{2m^2} \qquad \text{für} \quad x \in A \quad \text{mit} \quad |x - x_0| < \delta,$$

also

$$\begin{aligned}
\sum_{i,j=1}^{m} \mathrm{D}_i\,\mathrm{D}_j f(x)\eta_i\eta_j &= \sum_{i,j=1}^{m} \mathrm{D}_i\,\mathrm{D}_j f(x_0)\eta_i\eta_j + \sum_{i,j=1}^{m} \Big(\mathrm{D}_i\,\mathrm{D}_j f(x) - \mathrm{D}_i\,\mathrm{D}_j f(x_0)\Big)\eta_i\eta_j \\
&\leq -\lambda + m^2\frac{\lambda}{2m^2} = -\frac{\lambda}{2} < 0
\end{aligned}$$

für alle $x \in A$ mit $|x - x_0| < \delta$ und $y \in \mathbb{R}^m$ mit $|y| = 1$. Für *alle* $y \in \mathbb{R}^m$ und diese x gilt also

$$\sum_{i,j=1}^{m} \mathrm{D}_i\,\mathrm{D}_j f(x)\eta_i\eta_j < 0.$$

Damit folgt aus dem Taylorschen Satz für $p = 1$ und für $x \in A$ mit $|x - x_0| < \delta$

$$\begin{aligned}
f(x) - f(x_0) &= \sum_{|\alpha|=2} \frac{1}{\alpha!}\mathrm{D}^\alpha f\Big(x_0 + \vartheta(x - x_0)\Big)(x - x_0)^\alpha \\
&= \frac{1}{2}\sum_{i,j=1}^{m} \mathrm{D}_i\,\mathrm{D}_j f\Big(x_0 + \vartheta(x - x_0)\Big)(\xi_i - \xi_{i,0})(\xi_j - \xi_{j,0}) < 0,
\end{aligned}$$

d.h. f hat in x_0 ein striktes lokales Maximum. (Entsprechend zeigt man im anderen Fall, daß ein striktes lokales Minimum vorliegt.) ■

Wir wenden nun diese Überlegungen auf ein einfaches Beispiel an:

Beispiel 4.6 Auf $A = \mathbb{R}^2$ sei

$$f(x,y) := x^4 + y^4 - 4c^2xy \qquad (\text{mit} \quad c > 0).$$

Offensichtlich ist f beliebig oft stetig differenzierbar, also sicher in $C^2(\mathbb{R}^2, \mathbb{R})$. Die im folgenden benötigten Ableitungen sind

$$\begin{aligned}
&\mathrm{D}_1 f(x,y) = 4x^3 - 4c^2y, \quad \mathrm{D}_1^2 f(x,y) = 12x^2,\\
&\mathrm{D}_2 f(x,y) = 4y^3 - 4c^2x, \quad \mathrm{D}_2^2 f(x,y) = 12y^2,\\
&\mathrm{D}_1\,\mathrm{D}_2 f(x,y) = \mathrm{D}_2\,\mathrm{D}_1 f(x,y) = -4c^2.
\end{aligned}$$

Lokale Extrema sind nur möglich in den Punkten $(x,y) \in \mathbb{R}^2$ mit

$$\mathrm{D}_1 f(x,y) = \mathrm{D}_2 f(x,y) = 0,$$

in diesem Fall also

$$x^3 - c^2y = 0, \quad y^3 - c^2x = 0.$$

Die erste Gleichung liefert $y = c^{-2}x^3$; Einsetzen in die zweite Gleichung ergibt

$$c^{-6}x^9 - c^2x = 0, \quad \text{bzw.} \quad c^2x(c^{-8}x^8 - 1) = 0$$

mit den (reellen) Lösungen $x_1 = 0$ und $x_{2,3} = \pm c$. Insgesamt erhalten wir:

$$(x_1,y_1) = (0,0), \quad (x_2,y_2) = (c,c), \quad (x_3,y_3) = (-c,-c).$$

Nur in diesen drei Punkten brauchen wir die Hesse–Matrix $\Big(\mathrm{D}_i\,\mathrm{D}_j f(x,y)\Big)$ zu untersuchen.

Für $(x, y) = (x_1, y_1) = (0, 0)$ gilt

$$\Big(\mathrm{D}_i\,\mathrm{D}_j f(0,0)\Big) = \begin{pmatrix} 0 & -4c^2 \\ -4c^2 & 0 \end{pmatrix}.$$

Diese Matrix ist weder positiv noch negativ definit. Das erkennt man einerseits daran, daß sie die Eigenwerte $\pm 4c^2$ hat. Man sieht es aber auch durch Ausrechnen der quadratischen Form für die Vektoren $(1,1)^t$ und $(1,-1)^t$ (dies sind gerade die Eigenvektoren):

$$\left\langle \begin{pmatrix} 0 & -4c^2 \\ -4c^2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = -8c^2 < 0,$$
$$\left\langle \begin{pmatrix} 0 & -4c^2 \\ -4c^2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = 8c^2 > 0.$$

Folglich kann dort kein Extremum vorliegen.

Für $(x, y) = (x_{2/3}, y_{2/3}) = (\pm c, \pm c)$ gilt

$$\Big(\mathrm{D}_i\,\mathrm{D}_j f(x_{2,3}, y_{2,3})\Big) = \begin{pmatrix} 12c^2 & -4c^2 \\ -4c^2 & 12c^2 \end{pmatrix} = 4c^2 \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

Diese Matrix ist positiv definit. Das erkennt man daran, daß die Eigenwerte $8c^2$ und $16c^2$ sind. Man kann aber auch direkt die quadratische Form nach unten abschätzen:

$$\begin{aligned} \left\langle \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} 3a-b \\ -a+3b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle \\ &= 3a^2 - 2ab + 3b^2 \geq 3(a^2+b^2) - a^2 - b^2 = 2(a^2+b^2) \\ &> 0 \quad \text{für} \quad (a,b) \neq 0. \end{aligned}$$

Also liegen bei (c, c) und $(-c, -c)$ strenge lokale Minima vor. Es gilt

$$f(\pm c, \pm c) = -2c^4.$$

Für alle anderen $(x, y) \in \mathbb{R}^2$ gilt

$$\begin{aligned} f(x,y) &\geq x^4 + y^4 - 2c^2x^2 - 2c^2y^2 = (x^2-c^2)^2 + (y^2-c^2)^2 - 2c^4 \\ &= -2c^4 \qquad \text{falls} \quad x^2 = y^2 = c^2, \\ &> -2c^4 \qquad \text{in allen anderen Fällen.} \end{aligned}$$

Außerdem gilt für die in der vorletzten Formelzeile ausgeschlossenen Punkte mit $x^2 = y^2 = c^2$

$$f(c,-c) = f(-c,c) = 6c^4 > -2c^4.$$

Somit ist klar, daß f in $(x,y) = (\pm c, \pm c)$ seine „globalen“Minima hat; weitere Extrema liegen nicht vor.

Auf der Geraden $x = y = t$ erhält man

$$f(t,t) = 2t^4 - 4c^2t^2;$$

diese Funktion hat ein lokales Maximum in $t = 0$ und außerdem (die bereits gefundenen) Minima in $t = \pm c$.

Auf der Geraden $x = -y = t$ erhält man

$$f(t,-t) = 2t^4 + 4c^2t^2;$$

diese Funktion hat ein globales Minimum in $t = 0$ (und sonst kein Extremum). □

4.4 Übungsaufgaben

4.1 Eine Funktion $f : \mathbb{R}^m \to \mathbb{R}$ heißt *konvex*, wenn für alle $x, y \in \mathbb{R}^m$ und $\alpha \in [0,1]$ gilt $f(x + \alpha(y - x)) \le f(x) + \alpha(f(y) - f(x))$.

a) f ist genau dann konvex, wenn die Menge $\{(x,z) \in \mathbb{R}^m \times \mathbb{R} : z > f(x)\}$ „oberhalb des Graphen von f“ konvex ist.

b) Ist $f : \mathbb{R}^m \to \mathbb{R}$ konvex, so ist die Menge $A := \{x \in \mathbb{R}^m : f(x) \le 0\}$ konvex (d. h. mit zwei Punkten liegt auch deren Verbindungsstrecke in A).

c) Ist $f : \mathbb{R}^m \to \mathbb{R}$ zweimal stetig differenzierbar und ist die Matrix $(D_iD_jf(x))$ positiv definit für alle $x \in \mathbb{R}^m$, so ist f konvex.

4.2 Sei $A = \{x = (\xi_1, \xi_2) \in \mathbb{R}^2 : |x|^2 = \xi_1^2 + \xi_2^2 \le 1\}$, $f : A \to \mathbb{R}$, $\quad f(x) = (1 - |x|^2)^{1/2}$ (setzt man $\xi_3 = f(\xi_1, \xi_2)$ für $x = (\xi_1, \xi_2) \in A$, so beschreibt f die obere Hälfte der Einheitssphäre in $\mathbb{R}^3$). Man bestimme das Taylorpolynom $T_{0,3}(x)$ der Ordnung 3 für den Entwicklungspunkt 0.

4.3 Welche Kantenlängen hat ein Quader mit dem Volumen 1, wenn die Oberfläche ohne die Grundfläche minimal sein soll? (Minimaler Materialverbrauch für eine Kiste ohne Deckel.)

4.4 Sei $A = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : -1 \le \xi_1 \le 1, -1 \le \xi_2 \le 1\}$. Man bestimme den maximalen und den minimalen Wert der Funktion $f : A \to \mathbb{R}$, $f(\xi_1, \xi_2) = \xi_1^2 + \xi_2^2 - \xi_1\xi_2 + \xi_1$ (man beachte den Rand von A).

4.5 Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}$, zweimal stetig differenzierbar $\Big(\Delta f = \sum_{j=1}^{m} \frac{\partial^2}{\partial \xi_j^2} f\Big)$.

a) Gilt $\Delta f > 0$, so hat f in A kein lokales Maximum.

b) Ist $f : \overline{A} \to \mathbb{R}$ stetig und gilt $\Delta f = 0$ in A (d. h. f ist *harmonisch* in A), so nimmt f sein Maximum (Minimum) auf den Rand $\overline{A} \setminus A$ an. *Anleitung*: für $g(x) = f(x) + k^{-1}\xi_1^2$ gilt $\Delta f > 0$.

5 Kurven und Kurvenintegrale

5.1 Kurven

Anschaulich ist eine Kurve (speziell in $\mathbb{R}^2$) ein Gebilde, das in einem Zug (d. h. ohne abzusetzen) gezeichnet werden kann. Für mathematische Zwecke, aber auch für Anwendungen, insbesondere in der Physik, ist natürlich eine präzisere Definition nötig:

Als einen *Weg* γ in $\mathbb{R}^m$ bezeichnen wir eine stetige Abbildung $\gamma : I \to \mathbb{R}^m$, wobei $I = [a, b]$ ein beschränktes abgeschlossenes Intervall ist (in einigen Situationen werden wir auch unbeschränkte Intervalle zulassen),

$$\gamma = (\gamma_1, \ldots, \gamma_m) \qquad \gamma_j : I \to \mathbb{R} \;\; (j = 1, \ldots, m).$$

(Stellen wir uns das Intervall $[a, b]$ als Zeitintervall vor, so beschreibt $\gamma(t)$ eine Bewegung in $\mathbb{R}^m$.)

Zwei Wege $\gamma : [a, b] \to \mathbb{R}^m$ und $\tilde{\gamma} : [\tilde{a}, \tilde{b}] \to \mathbb{R}^m$ heißen *äquivalent*, $\gamma \sim \tilde{\gamma}$, wenn es eine strikt wachsende stetige Funktion $\varphi : [a, b] \to [\tilde{a}, \tilde{b}]$ gibt mit

$$\gamma(t) = \tilde{\gamma}(\varphi(t)) \text{ für alle } t \in [a, b], \text{ bzw. } \tilde{\gamma}(s) = \gamma(\varphi^{-1}(s)) \text{ für alle } s \in [\tilde{a}, \tilde{b}].$$

Bei zwei äquivalenten Wegen wird also die gleiche Punktmenge in gleicher Weise durchlaufen (d. h. u. a., daß ggf. die gleichen Punktmengen mehrfach durchlaufen werden.). Es ist offentsichtlich, daß hierdurch eine *Äquivalenzrelation* definiert wird: Es gilt $\gamma \sim \gamma$, aus $\gamma_1 \sim \gamma_2$ folgt $\gamma_2 \sim \gamma_1$; aus $\gamma_1 \sim \gamma_2$ und $\gamma_2 \sim \gamma_3$ folgt $\gamma_1 \sim \gamma_3$.

Jede Äquivalenzklasse Γ von Wegen bezüglich dieser Äquivalenzrelation (d. h. die Menge aller Wege, die zu einem Weg γ äquivalent sind) heißt eine *Kurve*. Die verschiedenen Wege einer Äquivalenzklasse Γ nennen wir auch *Parameterdarstellungen* der Kurve Γ.

Eine Kurve Γ heißt *geschlossen*, wenn für eine (und damit für jede) Parameterdarstellung γ gilt $\gamma(a) = \gamma(b)$, d. h. wenn *Anfangspunkt* und *Endpunkt* zusammenfallen. Γ heißt *doppelpunktfrei*, wenn eine/jede Parameterdarstellung γ injektiv ist (mit der möglichen Ausnahme $\gamma(a) = \gamma(b)$, falls Γ geschlossen ist). Γ heißt *(stetig) differenzierbar*, wenn es eine (stetig) differenzierbare Parameterdarstellung γ gibt. Γ heißt *glatt*, wenn es eine stetig differenzierbare Parameterdarstellung gibt mit $\gamma'(t) \neq 0$ für alle $t \in I$ (tatsächlich genügt es hierfür, daß es zu jedem x auf der Kurve eine Parameterdarstellung γ gibt mit $\gamma(t) = x$ für ein geeignetes t, die in einer Umgebung U von t stetig differenzierbar ist mit $\gamma'(s) \neq 0$ für $s \in U$; insbesondere am Beispiel der Neil'schen Parabel werden wir sehen, daß die Existenz einer stetig differenzierbaren Parameterdarstellung nicht reicht, um „Glattheit“ der Kurve im anschaulichen Sinn zu gewährleisten).

Beispiel 5.1 (a) Der *Kreis* mit Radius r um den Nullpunkt in $\mathbb{R}^2$:

(i) 1 mal durchlaufen

$$\gamma_1 : [0, 2\pi] \to \mathbb{R}^2, \quad t \mapsto (r\cos t, r\sin t),$$

(ii) k mal durchlaufen

$$\gamma_k : [0, 2k\pi] \to \mathbb{R}^2, \quad t \mapsto (r\cos t, r\sin t),$$

oder (dazu äquivalent)

$$\tilde{\gamma}_k : [0, 2\pi] \to \mathbb{R}^2, \quad t \mapsto (r\cos kt, r\sin kt).$$

(b) Für $a, v \in \mathbb{R}^m$ wird durch

$$\gamma : \mathbb{R} \to \mathbb{R}^m, \quad t \mapsto a + tv$$

eine *Gerade* beschrieben, die mit der „Geschwindigkeit“ $|v|$ durchlaufen wird.

(c) Für $r > 0$ und $c \in \mathbb{R}$ wird durch

$$\gamma : \mathbb{R} \to \mathbb{R}^3, \quad t \mapsto (r\cos t, r\sin t, ct)$$

eine *Schraubenlinie* mit *Ganghöhe* $2c\pi$ beschrieben.

(d) Ist $f : I \to \mathbb{R}$ eine stetige reellwertige Funktion, so ist

$$\gamma_f : I \to \mathbb{R}^2, \quad t \mapsto (t, f(t))$$

ein Weg (der Graph von f); ist f stetig differenzierbar, so ist die durch γ_f beschriebene Kurve in jedem Fall glatt, denn es ist $\gamma'(t) = (1,\ f'(t)) \neq 0$ für alle t (auch dann, wenn $f'(t) = 0$ sein sollte). Die Beispiele aus Teil (a) zeigen, daß natürlich nicht jede Kurve als Graph darstellbar ist (in der Regel ist dies aber stückweise möglich, vgl. §23).

□

Fassen wir, wie schon oben erwähnt, t als die Zeit auf, so ist $\gamma'(t_0)$ die (vektorielle) *Momentangeschwindigkeit* zur Zeit t_0. Die Geschwindigkeit (genauer der Betrag der Geschwindigkeit) ist also

$$|\gamma'(t_0)| = \Big\{ \sum_{j=1}^{m} \gamma_j'(t_0)^2 \Big\}^{1/2}.$$

Ist $\gamma'(t_0) \neq 0$, so wird durch den Vektor

$$|\gamma'(t_0)|^{-1}\gamma'(t_0) = \textit{Tangentialvektor} \text{ im Punkt } \gamma(t_0)$$

die *Richtung der Kurve* im Punkt $\gamma(t_0)$ beschrieben; insbesondere hat also eine glatte Kurve in jedem Punkt $\gamma(t_0)$ eine wohldefinierte Richtung.

Nach Definition der Ableitung ist die Gerade

$$g : \mathbb{R} \to \mathbb{R}^m, \quad t \mapsto \gamma(t_0) + \gamma'(t_0)(t - t_0)$$

die einzige Gerade mit der Eigenschaft

$$\frac{1}{t - t_0}\,(\gamma(t) - g(t)) \to 0 \text{ für } t \to t_0\,.$$

Eine Kurve Γ heißt *stückweise glatt*, wenn es eine Parameterdarstellung $\gamma : I \to \mathbb{R}^m$ gibt und $t_1, t_2, \ldots, t_{k-1}$ mit $a = t_0 < t_1 < t_2 < \ldots < t_{k-1} < t_k = b$ so, daß $\gamma|_{[t_{j-1},t_j]}$ stetig differenzierbar ist mit $\gamma'(s) \neq 0$ für $s \in [t_{j-1}, t_j]$ $(j = 1, \ldots, k)$; dabei sind in den Randpunkten t_{j-1} bzw. t_j die rechtsseitige bzw. linksseitige Ableitung gemeint.

Beispiel 5.2 (a) Die *Neil'sche Parabel* wird beschrieben durch

$$\gamma : \mathbb{R} \to \mathbb{R}^2, \ t \mapsto (t^2, t^3).$$

Sie kann als Graph der Funktion (x als Funktion von y)

$$f_1 : \mathbb{R} \to \mathbb{R}, \ f_1(s) = \left(\sqrt[3]{s}\right)^2$$

aufgefaßt werden, oder als Vereinigung der Graphen der Funktionen (y als Funktion von x)

$$f_j : [o, \infty) \to \mathbb{R}, \;\; f_1(s) := -(\sqrt{s})^3, \;\; f_2(s) := (\sqrt{s})^3.$$

. Bei $\gamma(0) = (0,0)$ liegt eine *Spitze* vor. Es kann deshalb keine stetig differenzierbare Parameterdarstellung geben, die in diesem Punkt eine von 0 verschiedene Ableitung hat. Die Parameterdarstellung

$$\tilde{\gamma} : \mathbb{R} \to \mathbb{R}^2, \quad t \mapsto \begin{cases} (t, t^{3/2}) & \text{für } t > 0, \\ (-t, -(-t)^{3/2}) & \text{für } t < 0 \end{cases}$$

zeigt jedoch, daß die Kurve stückweise glatt ist.

(b) Die Parameterdarstellung

$$\gamma : \mathbb{R} \to \mathbb{R}^2, \;\; t \mapsto (t^3, t^3)$$

erweckt den Eindruck, daß die dadurch dargestellte Kurve nicht glatt sein könnte ($\gamma'(0) = (0,0)$). Tatsächlich ist eine hierzu äquivalente Parameterdarstellung

$$\tilde{\gamma} : \mathbb{R} \to \mathbb{R}^2, \;\; t \mapsto (t, t),$$

der man sofort ansieht, daß es sich um ein Gerade handelt, die natürlich glatt ist ($\tilde{\gamma}'(t) = (1,1)$).

□

Sind $\gamma_1 : I_1 \to \mathbb{R}^m$ und $\gamma_2 : I_2 \to \mathbb{R}^m$ Parameterdarstellungen von zwei Kurven Γ_1 und Γ_2 mit $\gamma_1(t_1) = \gamma_2(t_2)$ für geeignete $t_1 \in I_1, t_2 \in I_2$ mit $\gamma_1'(t_1) \neq 0$ und $\gamma_2'(t_2) \neq 0$, so ist der *Schnittwinkel* ϑ der Kurven Γ_1 und Γ_2 in $\gamma_2(t_1) = \gamma_2(t_2)$ definiert durch

$$\vartheta = \arccos \frac{\langle \gamma_1'(t_1), \gamma_2'(t_2) \rangle}{|\gamma_1'(t_1)||\gamma_2'(t_1)|},$$

wobei $\arccos = (\cos|_{[0,\pi]})^{-1}$ ist. (Für $m = 2, 3$ ist dies der Winkel im Sinne der Elementargeometrie zwischen den beiden Tangenten, $0 \leq \vartheta \leq \pi$.)

5.2 Kurvenlänge

Wir wollen nun versuchen, die Länge einer Kurve Γ zu definieren. Sei $\gamma : I = [a,b] \to \mathbb{R}^m$ eine Parameterdarstellung von Γ,

$$Z : a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$$

eine *Zerlegung* des Intervalls $I = [a,b]$. Die Länge des durch die Punkte $\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_{n-1})$, $\gamma(t_n)$ erzeugten Polygonzuges ist

$$L_\gamma(Z) = L_\gamma(t_0, \ldots, t_n) := \sum_{i=1}^{n} |\gamma(t_i) - \gamma(t_{i-1})|.$$

Die Kurve Γ mit einer Parameterdarstellung γ heißt *rektifizierbar* mit *Länge* L_Γ (man schreibt dafür auch $|\Gamma|$), wenn

$$L_\Gamma := \sup \Big\{ L_\gamma(Z) : Z \ \text{ Zerlegungen von } I \Big\} < \infty$$

gilt; wenn Γ nicht rektifizierbar ist, sagt man auch, Γ habe unendliche Länge.

Die Rektifizierbarkeit bzw. die Länge L_Γ hängen nicht von der Wahl der Parameterdarstellung ab; sind γ und $\tilde{\gamma}$ äquivalente Parameterdarstellungen, $\gamma = \tilde{\gamma} \circ \phi$, so ist $Z : a = t_0 < \ldots < t_n = b$ genau dann eine Zerlegung von I, wenn $\widetilde{Z} : \tilde{a} = \phi(t_0) < \ldots < \phi(t_n) = \tilde{b}$ eine Zerlegung von $\widetilde{I}$ ist.

Tatsächlich kann man die Länge L_Γ (bzw. ∞, falls Γ nicht rektifizierbar ist) als Limes von Folgen $(L_\gamma(Z_n))$ mit geeigneten Zerlegungsfolgen (Z_n) bestimmen. Dazu sei die *Feinheit der Zerlegung* $Z : a = t_0 < \ldots < t_n = b$ definiert durch

$$\eta(Z) = \max \Big\{ |t_j - t_{j-1}| : j = 1, \ldots, n \Big\}.$$

Satz 5.3 *Sei Γ eine Kurve mit Parameterdarstellung $\gamma : I = [a,b] \to \mathbb{R}^m$. Ist (Z_n) eine Zerlegungsfolge von I mit $\eta(Z_n) \to 0$ für $n \to \infty$, so gilt:*
– $L_\gamma(Z_n) \to L_\Gamma$, falls Γ rektifizierbar ist,
– $L_\gamma(Z_n) \to \infty$, falls Γ nicht rektifizierbar ist.

Beweis. Offenbar genügt es zu zeigen: Zu jeder Zerlegung Z_0 und jedem $\varepsilon > 0$ existiert ein $\delta > 0$ so, daß für jede Zerlegung Z_0 und jede Zerlegung Z mit Feinheit $\eta(Z) < \delta$ gilt

$$L_\gamma(Z) \geq L_\gamma(Z_0) - \varepsilon.$$

Sei N die Zahl der Zerlegungspunkte von Z_0 (das entspricht oben der Zahl $n-1$). Da γ gleichmäßig stetig ist, existiert ein $\delta > 0$ mit $|\gamma(s) - \gamma(t)| < \varepsilon/2N$ für $|s-t| < \delta$. Sei nun Z eine Zerlegung mit Feinheit $\eta(Z) < \delta$, Z' die Zerlegung, die man erhält, wenn man zu Z die Zerlegungspunkte von Z_0 (als höchstens N Stück) hinzunimmt. Dann ist jedenfalls $L_\gamma(Z') \geq L_\gamma(Z_0)$. Da höchstens N Zerlegungspunkt mit jewseils zwei Intervallen der Länge $< \delta$ zur Zerlegung Z hinzugenommen wurden, gilt andererseits $L_\gamma(Z) \geq L_\gamma(Z') - 2N\varepsilon/2N = L_\gamma(Z') - \varepsilon > L_\gamma(z_0) - \varepsilon$. Das ist die gewünschte Ungleichung. ■

Beispiel 5.4 Wir betrachten

$$\gamma : [0,1] \to \mathbb{R}^2,\ \gamma(t) = \begin{cases} (0,0) & \text{für } t = 0, \\ (t, t \sin \pi/2t) & \text{für } t \neq 0. \end{cases}$$

Die hierdurch definierte Kurve ist *nicht rektifizierbar*, denn für

$$t_0 = 0,\ t_j = \{2(n-j)+1\}^{-1} \text{ für } j = 1, \ldots, n-1,\ t_n = 1$$

gilt

$$L_\gamma(t_0, \ldots, t_n) > \sum_{j=1}^{n-1} \frac{1}{2(n-j)+1} \to \infty \text{ für } n \to \infty.$$

Man könnte vielleicht glauben, daß dies daran liegt, daß γ in 0 nicht differenzierbar ist. Das ist aber nicht der Fall, denn die Kurve

$$\tilde{\gamma} : [0,1] \to \mathbb{R}^2,\ \tilde{\gamma}(t) = \begin{cases} (0,0) & \text{für } t = 0, \\ (t, t^2 \sin t^{-2}) & \text{für } t \neq 0 \end{cases}$$

ist in 0 differenzierbar (allerdings nicht stetig differenzierbar), aber ebenfalls nicht rektifizierbar (Beweis!). □

Satz 5.5 *Jede stetig differenzierbare Kurve Γ ist rektifizierbar. Für jede stetig differenzierbare Parameterdarstellung $\gamma : [a,b] \to \mathbb{R}^m$ gilt*

$$L_\Gamma = \int\limits_a^b |\gamma'(t)|\,\mathrm{d}t = \int\limits_a^b \Big\{\sum_{j=1}^m \gamma_j'(t)^2\Big\}^{1/2}\,\mathrm{d}t.$$

Beweis. Ist $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$ eine Zerlegung von I, so gilt für jedes $k \in \{1,\ldots,n\}$

$$\begin{aligned} |\gamma(t_k) - \gamma(t_{k-1})| &= \Big\{\sum_{j=1}^m \Big(\gamma_j(t_k) - \gamma_j(t_{k-1})\Big)^2\Big\}^{1/2} \\ &\quad \text{(Mittelwertsatz der Differential– und Integral–Rechnung)} \\ &= \Big\{\sum_{j=1}^m \Big[\gamma_j'(\tau_{j,k})(t_k - t_{k-1})\Big]^2\Big\}^{1/2} \qquad (\text{mit } \tau_{j,k} \in (t_{k-1},t_k)) \\ &= (t_k - t_{k-1})\Big\{\sum_{j=1}^m \gamma_j'(\tau_{j,k})^2\Big\}^{1/2}. \end{aligned}$$

(Wären alle $\tau_{j,k}$ $(j = 1,\ldots,m)$ gleich einem τ_k, so wäre der Beweis hier schon fertig, denn die Aufsummation der Terme würde eine Riemannsumme für das Integral $\int_a^b |\gamma(t)|\,\mathrm{d}t$ zur Zerlegung $a = t_0, t_1, \ldots, t_n = b$ liefern.) Andererseits gilt nach dem Mittelwertsatz der Integralrechnung, da $|\gamma'(t)|$ stetig ist, mit $\tau_k \in (t_{k-1}, t_k)$

$$\int_{t_{k-1}}^{t_k} |\gamma'(t)|\,\mathrm{d}t = (t_k - t_{k-1})|\gamma'(\tau_k)| = (t_k - t_{k-1})\Big\{\sum_{j=1}^m \gamma_j'(\tau_k)^2\Big\}^{1/2}.$$

Die Funktionen γ_j' sind auf $[a,b]$ gleichmäßig stetig, d. h. es gilt

$$\forall \varepsilon > 0\, \exists\, \delta > 0 \text{ so, daß } |\gamma_j'(s) - \gamma_j'(t)| < \frac{\varepsilon}{\sqrt{m}(b-a)} \text{ für } |s-t| < \delta \text{ und } j = 1,\ldots,m.$$

Wählen wir die Zerlegung so fein, daß $t_k - t_{k-1} < \delta$ gilt, so folgt also

$$\begin{aligned} &\Big||\gamma(t_k) - \gamma(t_{k-1})| - \int\limits_{t_{k-1}}^{t_k} |\gamma'(t)|\,\mathrm{d}t\Big| \\ &= (t_k - t_{k-1})\left|\Big\{\sum_{j=1}^m \gamma_j'(\tau_{j,k})^2\Big\}^{1/2} - \Big\{\sum_{j=1}^m \gamma_j'(\tau_k)^2\Big\}^{1/2}\right| \\ &\quad \Big(\text{wegen } \Big||x| - |y|\Big| \le |x-y| \text{ für } x,y \in \mathbb{R}^m\Big) \\ &\le (t_k - t_{k-1})\Big\{\sum_{j=1}^m \Big(\gamma_j'(\tau_{j,k}) - \gamma_j'(\tau_k)\Big)^2\Big\}^{1/2} \\ &\le (t_k - t_{k-1})\Big\{\sum_{j=1}^m \frac{\varepsilon^2}{m(b-a)^2}\Big\}^{1/2} = (t_k - t_{k-1})\frac{\varepsilon}{b-a}. \end{aligned}$$

Summation über $k = 1, \ldots, m$ liefert:

$$\Big| L_\gamma(t_0, t_1, \ldots, t_{n-1}, t_n) - \int_a^b |\gamma'(t)|\, \mathrm{d}t \Big| \leq \sum_{k=1}^{m} (t_k - t_{k-1}) \frac{\varepsilon}{b-a} = \varepsilon,$$

d. h. $L_\gamma(t_0, t_1, \cdots, t_n)$ konvergiert bei Verfeinerung der Zerlegung gegen $\int_a^b |\gamma'(t)|\, \mathrm{d}t$.

(Für einen anderen Beweis vgl. z. B. Endl–Luh, Analysis II: Ist $L_\gamma(a, t)$ die Länge der Kurve von $\gamma(a)$ bis $\gamma(t)$, so ist $L_\gamma(a, \cdot)$ stetig differenzierbar mit Ableitung $|\gamma(t)|$.) ■

Beispiel 5.6 Die *Zykloide* beschreibt die Bahn eines Punktes auf der Peripherie eines Kreises vom Radius 1, der auf der x–Achse abrollt. An einer Skizze erkennt man leicht die folgende Parameterdarstellung:

$$\gamma : \mathbb{R} \to \mathbb{R}^2, \ \mapsto (t - \sin t, 1 - \cos t).$$

Offenbar ist γ stetig differenzierbar mit

$$\begin{aligned} \gamma'(t) &= (1 - \cos t, \sin t), \\ |\gamma'(t)|^2 &= (1 - \cos t)^2 + \sin^2 t \\ &= 2 - 2\cos t = 2\Big\{\sin^2 \frac{t}{2} + \cos^2 \frac{t}{2} - \cos^2 \frac{t}{2} + sin^2 \frac{t}{2}\Big\} = 4 \sin^2 \frac{t}{2} \\ |\gamma'(t)| &= 2|\sin \frac{t}{2}|. \end{aligned}$$

Die Länge des Kurvenstücks mit Parameterdarstellung $\gamma|_{[0,2\pi]}$ ist somit

$$L = \int\limits_0^{2\pi} 2|\sin \frac{t}{2}|\, \mathrm{d}t = 4 \int\limits_0^{\pi} \sin u \, \mathrm{d}u = 8.$$

Es ist überraschend und bemerkenswert, daß diese Kurvenlänge ganzzahlig ist. □

Häufig ist es günstig, Kurven in $\mathbb{R}^2$ in *Polarkoordinaten* darzustellen, Radius (= Abstand vom Nullpunkt) als Funktion vom Winkel φ:

$$\varphi \mapsto r(\varphi).$$

In rechtwinkligen Polarkoordinaten erhält man daraus die Parameterdarstellung

$$\gamma(\varphi) = \Big(r(\varphi)\cos\varphi, r(\varphi)\sin\varphi\Big),$$

und, falls $r(\cdot)$ differenzierbar ist,

$$\begin{aligned}
\gamma'(\varphi) &= \Big(r'(\varphi)\cos\varphi - r(\varphi)\sin\varphi, r'(\varphi)\sin\varphi + r(\varphi)\cos(\varphi)\Big), \\
|\gamma'(\varphi)| &= \Big\{\Big(r'(\varphi)\cos\varphi - r(\varphi)\sin\varphi\Big)^2 + \Big(r'(\varphi)\sin\varphi + r(\varphi)\cos\varphi\Big)^2\Big\}^{1/2} \\
&= \Big\{r'(\varphi)^2 + r(\varphi)^2\Big\}^{1/2}.
\end{aligned}$$

Damit folgt:

Satz 5.7 *Ist $\varphi \mapsto r(\varphi)$ eine stetig differenzierbare Darstellung einer Kurve in $\mathbb{R}^2$ in Polarkoordinaten, so gilt für die Länge der Kurve Γ für φ zwischen a und b*

$$L_\Gamma = \int_a^b \Big\{r'(\varphi)^2 + r(\varphi)^2\Big\}^{1/2} \mathrm{d}\varphi.$$

Diese Formeln für die Kurvenlänge lassen sich natürlich sofort auf stückweise stetig differenzierbare Kurven übertragen.

5.3 Krümmung ebener Kurven

Für die geometrische Gestalt einer glatten Kurve Γ in der Nähe eines Punktes der Kurve ist außer der Richtung in diesem Punkt auch die „Änderungsgeschwindigkeit“ der Richtung gemessen an der Weglänge interessant (also Richtungsänderung pro Längeneinheit des Weges; nicht Richtungsänderung pro Zeiteinheit, was natürlich von der Wahl der Parameterdarstellung abhängen würde).

Ist die Richtungsänderung (gemessen im Bogenmaß) pro Längeneinheit des Weges über ein Wegstück konstant $= \kappa$, so beschreibt dieses Wegstück offenbar ein Stück einer Kreislinie mit Radius $r = 1/\kappa$. Dieser Kreis liegt links von der Kurve (man beachte, daß die Kurve eine Richtung hat), wenn $\kappa > 0$ ist; er liegt rechts von der Kurve, wenn $\kappa < 0$ ist.

Sei also $\gamma : [a, b] \to \mathbb{R}^2$ eine stetig differenzierbare Parameterdarstellung mit $\gamma'(t) \neq 0$, $\alpha(t)$ der Winkel zwischen der Richtung der positiven x–Achse und der Richtung $v(t) := |\gamma'(t)|^{-1}\gamma'(t)$ der Kurve im Punkt $\gamma(t)$; es gilt also (mod π)

$$\alpha(t) = \begin{cases} \arctan \dfrac{\gamma_2'(t)}{\gamma_1'(t)}, & \text{falls } \gamma_1'(t) \neq 0, \\ \arccos \dfrac{\gamma_1'(t)}{\gamma_2'(t)}, & \text{falls } \gamma_2'(t) \neq 0. \end{cases}$$

Daraus ergibt sich, wenn γ zweimal stetig differenzierbar ist, für die Winkeländerung pro Zeiteinheit im ersten Fall

$$\alpha'(t) = \frac{1}{1 + (\gamma_2'(t)/\gamma_1'(t))^2} \, \frac{\gamma_2''\gamma_1'(t) - \gamma_1''\gamma_2'(t)}{\gamma_1'(t)^2} = \frac{\gamma_2''(t)\gamma_1'(t) - \gamma_1''(t)\gamma_2'(t)}{|\gamma'(t)|^2};$$

und im zweiten Fall erhält man das gleiche Resultat. Da wir aber die Winkeländerung pro Längeneinheit des Weges haben wollen, müssen wir zusätzlich durch die Geschwindigkeit $|\gamma'(t)|$ dividieren. Damit folgt für die *Krümmung* einer zweimal stetig differenzierbaren Kurve

$$\kappa(t) = \frac{\gamma_2''(t)\gamma_1'(t) - \gamma_1''(t)\gamma_2'(t)}{|\gamma'(t)|^3} = \frac{\gamma_2''(t)\gamma_1'(t) - \gamma_1''(t)\gamma_2'(t)}{(\gamma_1'(t)^2 + \gamma_2'(t)^2)^{3/2}}.$$

Mit $r(t) = \kappa(t)^{-1}$ läßt sich damit leicht in jedem Punkt der Kurve der *Krümmungsradius* berechnen.

5.4 Kurvenintegrale

Zur Motivation denken wir uns einen Körper, der unter der Einwirkung eines Kraftfeldes (Vektorfeldes) $k : \mathbb{R}^3 \to \mathbb{R}^3$ eine glatte Kurve Γ mit einer stetig differenzierbaren Parameterdarstellung $\gamma : I \to \mathbb{R}^3$ durchläuft. Wir wollen die durch das Kraftfeld geleistete Arbeit beim Durchlaufen der Kurve bestimmen. Dazu betrachten wir eine Zerlegung des Zeitintervalls $I = [a, b] : a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$. Die in dem „kleinen“ Zeitintervall $[t_{k-1}, t_k]$

geleistete Arbeit ist (mit $t_k^* \in [t_{k-1}, t_k]$) „näherungsweise“ gleich

$$\begin{aligned}
&\Big\langle k(\gamma(t_k^*)), \gamma(t_k) - \gamma(t_{k-1})\Big\rangle \\
=\ & \Big\langle k(\gamma(t_k^*)), \frac{1}{t_k - t_{k-1}}\Big(\gamma(t_k) - \gamma(t_{k-1})\Big)\Big\rangle (t_k - t_{k-1}) \\
& (\text{Mittelwertsatz der Differential– und Integralrechnung} \\
& \text{mit } t_{j,k} \in (t_{k-1}, t_k)) \\
=\ & \sum_{j=1}^{3} k_j(\gamma(t_k^*))\gamma_j'(t_{j,k})(t_k - t_{k-1}) \\
=\ & \sum_{j=1}^{3} k_j(\gamma(t_k^*))\gamma_j'(t_k^*)(t_k - t_{k-1}) \\
& + \sum_{j=1}^{3} k_j(\gamma(t_k^*))\Big(\gamma_j'(t_{j,k}) - \gamma_j'(t_k^*)\Big)(t_k - t_{k-1}) \\
=\ & \Big\langle k(\gamma(t_k^*)), \gamma'(t_k^*)\Big\rangle (t_k - t_{k-1}) \\
& + \sum_{j=1}^{3} k_j(\gamma(t_k^*))\Big(\gamma_j'(t_{j,k}) - \gamma_j'(t_k^*)\Big)(t_k - t_{k-1}).
\end{aligned}$$

Die in dem Zeitintervall $[a, b]$ geleistete Arbeit ist also „näherungsweise“ gleich

$$\begin{aligned}
&\sum_{k=1}^{n} \Big\langle k(\gamma(t_k^*)), \gamma'(t_k^*)\Big\rangle (t_k - t_{k_1}) \\
&\quad + \sum_{j=1}^{3} \sum_{k=1}^{n} k_j(\gamma(t_k^*))\Big(\gamma_j'(t_{j,k}) - \gamma_j'(t_k^*)\Big)(t_k - t_{k-1}).
\end{aligned}$$

Da die γ_j' gleichmäßig stetig sind, strebt hier der zweite Summand gegen 0 wenn die Feinheit der Zerlegung gegen 0 geht. Der erste ist eine Riemannsumme für

$$\int_a^b \Big\langle k(\gamma(t)), \gamma'(t)\Big\rangle \,\mathrm{d}t$$

und strebt somit, unabhängig von der Wahl der Zerlegungen und der Stützstellen t_k^*, bei Verfeinerung gegen diesen Wert.

Wir definieren deshalb für eine stetig differenzierbare Kurve Γ mit stetig differenzierbarer Parameterdarstellung $\gamma : I \to \mathbb{R}^m$ das *Kurvenintegral* des stetigen Vektorfeldes $k : \mathbb{R}^m \to \mathbb{R}^m$ längs der Kurve Γ durch

$$\int_\Gamma k(x)\,\mathrm{d}x = \int_\Gamma k_1(x)\,\mathrm{d}\xi_1 + \ldots + k_m(x)\,\mathrm{d}\xi_m := \int_a^b \Big\langle k(\gamma(t)), \gamma'(t)\Big\rangle \,\mathrm{d}t.$$

Aus den Überlegungen, die zur Definition geführt haben, folgt, daß dieses Kurvenintegral von der Wahl der (stetig differenzierbaren) Parameterdarstellung unabhängig ist: Ist $\tilde{\gamma} : \tilde{I} \to \mathbb{R}^m$ eine zu γ äquivalente Parameterdarstellung, $\gamma = \tilde{\gamma} \circ \varphi$, so ist

$$\begin{aligned} &\sum_{k=1}^{n} \Big\langle k(\tilde{\gamma}(\tilde{t}_{k-1})), \tilde{\gamma}(\tilde{t}_k) - \tilde{\gamma}(\tilde{t}_{k-1}) \Big\rangle \\ &\quad = \sum_{k=1}^{n} \Big\langle k\Big(\gamma(\varphi^{-1}(\tilde{t}_{k-1}))\Big), \gamma(\varphi^{-1}(\tilde{t}_k)) - \gamma(\varphi^{-1}(\tilde{t}_{k-1})) \Big\rangle \\ &\quad = \sum_{k=1}^{n} \Big\langle k(t_{k-1}), \gamma(t_k) - \gamma(t_{k-1}) \Big\rangle \end{aligned}$$

mit $t_k = \varphi^{-1}(\tilde{t}_k)$, d. h. eine Näherungssumme für $\tilde{\gamma}$ läßt sich gleichzeitig als Näherungssumme für γ auffassen.

Ist Γ eine glatte Kurve, so kann auch die *Kurvenlänge* als Kurvenintegral über ein geeignetes Vektorfeld definiert werden: Ist γ eine Parameterdarstellung mit $\gamma'(t) \neq 0$, so ist offenbar

$$L_\Gamma = \int_a^b |\gamma'(t)| \,\mathrm{d}t = \int_a^b \Big\langle \frac{\gamma'(t)}{|\gamma'(t)|}, \gamma'(t) \Big\rangle \,\mathrm{d}t = \int_\Gamma \frac{\gamma'(\gamma^{-1}(x))}{|\gamma'(\gamma^{-1}(x))|} \,\mathrm{d}x,$$

das ist das Kurvenintegral über das *Tangentialfeld* (das allerdings nur auf der Kurve Γ selbst definiert ist).

Alle diese Überlegungen lassen sich natürlich ohne weiteres auf stückweise glatte Kurven übertragen.

Sei $V : \mathbb{R}^m \to \mathbb{R}$ stetig differenzierbar, $k : \mathbb{R}^m \to \mathbb{R}^m$ ein Vektorfeld. Gilt $k(x) = \operatorname{grad} V(x)$, so heißt V ein *Potential* von k, k ein *Gradientenfeld* mit Potential V. (In der Physik verlangt man meist $k(x) = -\operatorname{grad} V(x)$.)

Sei $k : \mathbb{R}^m \to \mathbb{R}^m$ ein Vektorfeld, $A \subset \mathbb{R}^m$ offen. Man sagt, das Kurvenintegral von k ist (in A) *wegunabhängig*, wenn das Kurvenintegral des Vektorfeldes k längs aller (stückweise) glatten Kurven (in A) mit gleichem Anfangs– und Endpunkt den gleichen Wert hat; oder gleichbedeutend: wenn das Kurvenintegral über alle geschlossenen Kurven (in A) verschwindet.

Satz 5.8 *Sei $A \subset \mathbb{R}^m$ offen, $k : A \to \mathbb{R}^m$ ein stetiges Vektorfeld. Dann gilt: k ist genau dann ein Gradientenfeld, wenn das Kurvenintegral von k in A wegunabhängig ist.*

Beweis. $\Rightarrow$: Sei $k = \operatorname{grad} V$ mit stetig differenzierbarem $V : \mathbb{R}^m \to \mathbb{R}$. Dann gilt

$$\begin{aligned}\int\limits_\Gamma k(x)\,\mathrm{d}x &= \int\limits_a^b \Big\langle k(\gamma(t)), \gamma'(t)\Big\rangle\,\mathrm{d}t = \int\limits_a^b \sum_{j=1}^m \mathrm{D}_j V(\gamma(t))\gamma_j'(t)\,\mathrm{d}t \\ &= \int\limits_a^b \frac{d}{\mathrm{d}t} V(\gamma(t))\,\mathrm{d}t = V(\gamma(b)) - V(\gamma(a)),\end{aligned}$$

d. h. das Integral hängt nur vom Anfangs– und Endpunkt von Γ ab und ist somit wegunabhängig.

$\Leftarrow$: Sei $x_0 \in A$ fest. Da das Kurvenintegral wegunabhängig ist, können wir definieren

$$V(x) := \int\limits_{\Gamma_x} k(z)\,\mathrm{d}z$$

mit einer beliebigen stetig differenzierbaren Kurve Γ_x von x_0 nach x für jedes $x \in A$, das von x_0 aus auf einer Kurve erreichbar ist (das ist eine *Zusammenhangskomponente* von A; das Verfahren kann so auf jeder Zusammenhangskomponente, unabhängig von den anderen Zusammenhangskomponenten, durchgeführt werden). Der so definierte Wert hängt nach Voraussetzung nicht von der Wahl der Kurve Γ_x ab. Es bleibt zu zeigen, daß $\mathrm{D}_j V(x) = k_j(x)$ gilt für $j = 1, \ldots, m$. Da A offen ist, ist für $x \in A$ und hinreichend kleines h auch $x + te_j$ aus A für alle t zwischen 0 und h und $j = 1, \ldots, m$. Also gilt

$$\begin{aligned}\mathrm{D}_j V(x) &= \lim_{h\to 0} \frac{1}{h}\Big(V(\xi_1, \ldots, \xi_j + h, \ldots, \xi_m) - V(x)\Big) \\ &= \lim_{h\to 0} \frac{1}{h}\Big\{\int_{\Gamma_{x+he_j}} k(z)\,\mathrm{d}z - \int_{\Gamma_x} k(z)\,\mathrm{d}z\Big\} \\ &\qquad (\text{ mit einer beliebigen Kurve } \Gamma \text{ von } x \text{ nach } x + he_j, \\ &\qquad \text{z. B. mit } \gamma(t) = x + te_j \text{ für } t \in [0, h]) \\ &= \lim_{h\to 0} \frac{1}{h}\int_\Gamma k(z)\,\mathrm{d}z = \lim_{h\to 0} \frac{1}{h}\int_0^h k_j(\xi_1, \ldots, \xi_j + t, \ldots, \xi_m)\,\mathrm{d}t \\ &= k_j(x),\end{aligned}$$

d. h. V ist ein Potential von k. ■

Bemerkung 5.9 Bei einer Bewegung auf einer geschlossenen Kurve in einem Kraftfeld mit Potential (Gradientenfeld) wird bei einem vollen Umlauf keine Arbeit geleistet (ein solches Feld wird deshalb auch als *konservativesFeld* bezeichnet).

Beispiel 5.10 Das Vektorfeld

$$k : \mathbb{R}^2 \to \mathbb{R}^2, \qquad k(x) = (\xi_2, -\xi_1)$$

besitzt kein Potential, denn offensichtlich verschwinden die Kurvenintegrale (z. B.) über Kreise um den Nullpunkt nicht (Nachrechnen; das Integral über einem Kreis um 0 mit Radius r in positiver Richtung ist $-2\pi r^2$). — Man kann dies aber auch ohne diese Rechnung sehen: Wäre nämlich

$$k(x) = \operatorname{grad} V(x) = (\,\mathrm{D}_1 V(x),\, \mathrm{D}_2 V(x))$$

mit einer stetig differenzierbaren Funktion V, so wäre (da k stetig differenzierbar ist) V 2 mal stetig differenzierbar und es würde gelten

$$\mathrm{D}_1 k_2(x) - \mathrm{D}_2 k_1(x) = \mathrm{D}_1\, \mathrm{D}_2 V(x) - \mathrm{D}_2\, \mathrm{D}_1 V(x) = 0.$$

Tatsächlich ist aber

$$\mathrm{D}_1 k_2(x) - \mathrm{D}_2 k_1(x) = -1 - 1 = -2 \neq 0.$$

Der Ausdruck $\mathrm{D}_1 k_2(x) - \mathrm{D}_2 k_1(x)$ wird als *Rotation* des Vektorfeldes k bezeichnet. Wir werden später mehr darüber erfahren (vgl. Kapitel 10), insbesondere, daß es sich dabei um so etwas wie eine Wirbeldichte handelt. □

5.5 Übungsaufgaben

5.1 Man berechne die Länge einer Wirkung der Schraubenlinie

$$\gamma : [0, 2\pi] \to \mathbb{R}^3, \;\; \gamma(t) = (r\cos t, r\sin t, tc)$$

a) mit Hilfe der Formel aus dem Satz 5.5,

b) elementargeometrisch.

5.2 Die *Kardioide* (Herzkurve) hat in Polarkoordinaten die Darstellung

$$r(\varphi) = a(1 + \cos\varphi) \text{ für } 0 \leq \varphi \leq 2\pi, \quad a > 0.$$

Man skizziere den Verlauf dieser Kurve und berechne ihre Länge. (Das Integral $\int (1 + \cos\varphi)^{1/2}\,\mathrm{d}\varphi$ kann man z. B. mit Hilfe der Substitution $\varphi = \arccos t$ berechnen.)

5.3 Die Kurve $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = (e^{ct}\cos t, e^{ct}\sin t)$ heißt *logarithmische Spirale*.

a) Man skizziere die Kurve und stelle sie in der Form „ϕ Funktion von r".

b) Man zeige, daß die Einschränkung der Kurve auf den Parameterbereich $(-\infty, 0]$ endliche Länge hat und berechne sie.

c) Man zeige, daß die logarithmische Spirale jeden Kreis um den Nullpunkt in genau einem Punkt schneidet und berechne den Schnittwinkel.

5.4 Man zeige, daß das Vektorfeld $f : \mathbb{R}^2 \to \mathbb{R}^2$

a) $f(x, y) = (y, y - x)$ kein Potential besitzt,

b) $f(x, y) = (y, x - y)$ ein Potential besitzt und gebe eines an.

5.5 Für eine stetige Funktion $g : (0, \infty) \to \mathbb{R}$ sei

$$f : \mathbb{R}^m \backslash \{0\} \to \mathbb{R}^m, \; f(x) = g(|x|)x.$$

a) Man zeige, daß f ein Potential besitzt.

b) Man berechne ein Potential für das Vektorfeld $f(x) = |x|^{-k}x$.

c) Man berechen die Difergenz von f und zeige, daß f genau dann divergenzfrei ist, wenn $g(t) = ct^{-m}$ gilt.

5.6 a) Man skizziere das Vektorfeld in $\mathbb{R}^2 \backslash \{0\}$

$$k(x) = |x|^{-2}(\xi_2, -\xi_1), \; x \in \mathbb{R}^2 \backslash \{0\}$$

und zeige, daß es kein Potential besitzt.

b) Auf $\mathbb{R}^2 \backslash \{(\xi_1, 0) : \xi_1 \leq 0\}$ ist k durch ein Potential darstellbar.
Anleitung: Man glaube zunächst die Behauptung, setze z. B. $V(1, 0) = 0$, bestimme V auf der positiven ξ_1–Achse durch Integration und auf allen Kreisen um Null (ohne den Punkt auf der negativen ξ_1–Achse) durch Integration längs dieses Kreises.

6 Umkehrfunktion und implizit definierte Funktion

6.1 Umkehrfunktion

Es sei zunächst an den 1–dimensionalen Fall erinnert: Ist $f : I \to \mathbb{R}$ stetig und streng monoton, so hat f eine stetige Umkehrfunktion (Inverse) $f^{-1} : f(I) \to I$. Ist zusätzlich f differenzierbar in x_0 und gilt $f'(x_0) \neq 0$, so ist f^{-1} an der Stelle $y_0 := f(x_0)$ differenzierbar mit

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)} \text{ bzw. } (f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Ist f stetig differenzierbar mit $f'(x) \neq 0$ für alle $x \in I$, so ist auch f^{-1} stetig differenzierbar. (Die strenge Monotonie von f folgt z. B. aus der stetigen Differenzierbarkeit von f und $f'(x) \neq 0$ für alle $x \in I$.)

In diesem Abschnitt wollen wir dieses Resultat auf Funktionen mehrerer Variablen verallgemeinern. Der Begriff der Monotonie läßt sich offenbar nicht ohne weiteres verallgemeinern. Dagegen gibt es eine sinnvolle Verallgemeinerung von „$f'(x) \neq 0$“, wie wir am folgenden Spezialfall erkennen: $\det \mathrm{D}f(x) \neq 0$.

Beispiel 6.1 Sei L eine $m \times m$–Matrix, $a \in \mathbb{R}^m$, f die affine Funktion

$$f : \mathbb{R}^m \to \mathbb{R}^m, \; f(x) = a + Lx.$$

Offenbar ist f in jedem $x \in \mathbb{R}^m$ differenzierbar mit konstanter Ableitung $\mathrm{D}f(x) = L$, denn es gilt

$$\begin{aligned} f(z) &= a + Lz = a + Lx + L(z - x) \\ &= f(x) + L(z - x) + \varphi_x(z) \text{ mit } \varphi_x(z) \equiv 0; \end{aligned}$$

man kann aber auch leicht die partiellen Ableitungen von f explizit berechnen. Da $y = a+Lx$ genau dann für alle y eindeutig nach x auflösbar ist, wenn L invertierbar ist, $x = L^{-1}(y-a)$, ist f genau dann invertierbar, wenn L invertierbar ist, und es gilt

$$f^{-1} : \mathbb{R}^m \to \mathbb{R}^m, \; f^{-1}(y) = L^{-1}(y - a) = b + L^{-1}y \text{ mit } b = -L^{-1}a.$$

In diesem elementaren Fall ist also f genau dann invertierbar, wenn $Df(x)$ invertierbare Matrix ist. $\square$

Da eine differenzierbare Funktion $f : A \to \mathbb{R}^m$ mit $A \subset \mathbb{R}^m$ in der Nähe von x_0 durch die affine Funktion

$$F_{x_0} : \mathbb{R}^m \to \mathbb{R}^m,\ F_{x_0}(x) = f(x_0) + \mathrm{D}f(x_0)(x - x_0)$$

sehr gut approximiert wird, liegt es nahe, daß f in der Nähe von x_0 umkehrbar ist, wenn die Matrix $\mathrm{D}f(x_0)$ invertierbar ist: annähernd sollte also für y nahe $f(x_0)$ gelten

$$f^{-1}(y) \sim F_{x_0}^{-1}(y) = (\,\mathrm{D}f(x_0))^{-1}(y - f(x_0)) + x_0;$$

die Ableitung von f^{-1} bei x_0 sollte also $(\,\mathrm{D}f(x_0))^{-1}$ sein. (Da die Approximation nur lokal gut ist, ist allerdings auch zu erwarten, daß diese Aussage nur lokal gilt; genau das ist die Behauptung des folgenden Satzes.)

Wir benutzen die folgende Ausdrucksweise: Ist $x_0 \in A \subset \mathbb{R}^m$, so heißt U eine offene Umgebung von x_0 in A, wenn U offen ist mit $x_0 \in U \subset A$ (es gibt also eine ganze Kugel um x_0, die noch in U liegt); im Falle $A = \mathbb{R}^m$ nennen wir U einfach eine offene Umgebung von x_0.

Satz 6.2 (Umkehrfunktion) *Sei $A \subset \mathbb{R}^m$ offen, $f : A \to \mathbb{R}^m$ stetig differenzierbar, $x_0 \in A,\ y_0 := f(x_0),\ \mathrm{D}f(x_0)$ invertierbar. Dann gibt es eine offene Umgebung U von x_0 in A und eine offene Umgebung V von y_0 so, daß*

$$f : U \to V$$

bijektiv ist und $\mathrm{D}f(x)$ für alle $x \in U$ invertierbar ist. Die Umkehrabbildung $g := (f|_U)^{-1} : V \to U$ ist stetig differenzierbar mit

$$\mathrm{D}g(f(x)) = (\,\mathrm{D}f(x))^{-1} \text{ für } x \in U$$

bzw.

$$\mathrm{D}g(y) = (\,\mathrm{D}f(g(y)))^{-1} \text{ für } y \in V.$$

Beispiel 6.3 Wie schon oben angemerkt wurde, ist nicht zu erwarten, daß die Aussage des Satzes (insbesondere die Invertierbarkeit von f) global gilt. Dies soll durch folgendes Beispiel belegt werden: Sei

$$A = \mathbb{R}^2 \backslash \{0\} \subset \mathbb{R}^2, \; f(\xi_1, \xi_2) = (\xi_1^2 - \xi_2^2, 2\xi_1\xi_2).$$

Dann gilt

$$\mathrm{D}f(\xi_1, \xi_2) = \begin{pmatrix} 2\xi_1 & -2\xi_2 \\ 2\xi_2 & 2\xi_1 \end{pmatrix},$$

$$\det \mathrm{D}f(\xi_1, \xi_2) = 4(\xi_1^2 + \xi_2^2) \neq 0 \text{ für } x \in A.$$

Trotzdem ist f nicht injektiv, denn es gilt für alle $x \in A$

$$f(-\xi_1, -\xi_2) = f(\xi_1, \xi_2).$$

(Setzt man $z = \xi_1 + i\xi_2$, so ist offenbar $f(\xi_1, \xi_2) = (\mathrm{Re}\, z^2, \mathrm{Im}\, z^2)$, d. h. es handelt sich bei f um die Darstellung der komplexen Quadrierung in $\mathbb{R}^2$.) □

Beweis von Satz 6.2. Da A offen ist, und f stetig differenzierbar, existiert ein $\delta > 0$ mit $\boldsymbol{U := K_\delta(x_0)} \subset A$ und [2]

$$|\mathrm{D}f(x) - \mathrm{D}f(x_0)| \leq \frac{1}{2}\left|\mathrm{D}f(x_0)^{-1}\right|^{-1} := \lambda \text{ für } x \in A \text{ mit } |x - x_0| < \delta_0.$$

Wir werden zeigen: Mit diesem U und $\boldsymbol{V := f(U)}$ gilt die Behaauptung des Satzes.

***$\mathrm{D}f(x)$ ist invertierbar für $x \in U$*:** Für alle $x \in U$ und $y \in \mathbb{R}^m \backslash \{0\}$ gilt (dabei benutzen wir, daß für eine invertierbare Matrix L gilt $|y| \leq |L^{-1}||Ly|$ bzw. $|Ly| \geq |L^{-1}|^{-1}|y|$)

$$\begin{aligned} |\mathrm{D}f(x)y| &\geq |\mathrm{D}f(x_0)y| - \left|\mathrm{D}f(x)y - \mathrm{D}f(x_0)y\right| \\ &\geq \left|\mathrm{D}f(x_0)^{-1}\right|^{-1}|y| - \frac{1}{2}\left|\mathrm{D}f(x_0)^{-1}\right|^{-1}|y| \\ &= \frac{1}{2}\left|\mathrm{D}f(x_0)^{-1}\right|^{-1}|y| > 0, \end{aligned}$$

[2] dabei ist für eine $m \times m$–Matrix L mit $|L|$ die „Operatorennorm“ von L gemeint, $|L| := \sup\{|Lx| : x \in \mathbb{R}^m, |x| \leq 1\}$. Es gilt z. B. $|L| \leq \sum_{i,j=1}^m |\ell_{ij}|$ und $|L| \leq \left\{\sum_{i,j=1}^m |\ell_{ij}|^2\right\}^{1/2}$; insbesondere ist $|L|$ klein, wenn alle a_{ij} klein sind.

d. h. für $x \in U$ ist $\mathrm{D}f(x)$ invertierbar.

***f** ist auf U injektiv:* Für $x \in U$ schreiben wir f in der Form

$$f(x) = Df(x_0)x + F(x) \text{ mit } F(x) := f(x) - \mathrm{D}f(x_0)x.$$

Dann gilt für alle $x, x' \in U$ mit $x \neq x'$ (wieder benutzen wir $|Ly| \geq |L^{-1}|^{-1}|y|$)

$$\begin{aligned}
|f(x) - f(x')| &= |\,\mathrm{D}f(x_0)(x - x') + F(x) - F(x')| \\
&\geq |\,\mathrm{D}f(x_0)(x - x')| - |F(x) - F(x')| \\
&\geq |Df(x_0)^{-1}|^{-1}|x - x'| - |DF(\xi)||x - x'| \\
&= 2\lambda|x - x'| - |Df(\xi) - Df(x_0)||x - x'| \\
&\geq 2\lambda|x - x'| - \lambda|x - x'| = \lambda|x - x'| > 0.
\end{aligned}$$

Also ist $f|_U$ injektiv.

Damit ist jedenfalls $\boldsymbol{f : U \to V}$ *bijektiv*.

$\boldsymbol{V}$ *ist offen*: Dazu zeigen wir: Ist

$$y' = f(x') \in V = f(U),\ \delta' := \delta - |x' - x_0|,\ \varepsilon = \varepsilon(x') := \frac{\lambda}{2}\delta',$$

so ist $K_\varepsilon(y') \in V$, d. h.: zu jedem $z \in K_\varepsilon(y')$ gibt es ein $x \in U = K_\delta(x_0)$ mit $f(x) = z$.

Wegen $|z - y'| < \varepsilon = \dfrac{\lambda}{2}\delta'$ gibt es ein $r < \delta'$ mit $|z - y'| < \dfrac{\lambda}{2}r$.

Wir betrachten die Hilfsfunktion

$$\psi : \overline{K}_r(x') \to \mathbb{R}_+,\ x \mapsto |f(x) - z|^2.$$

Gesucht ist also ein $x \in K_r(x') \subset U$ mit $f(x) = z$, d. h. mit $\psi(x) = 0$; da ψ nicht negativ ist, ist dies ein Minimum von ψ). – Da ψ auf $\overline{K}_r(x')$ stetig ist, nimmt es dort sein Minimum an; es ist zu zeigen, daß dieser minimale Wert 0 ist. Für x auf dem Rand von $\overline{K}_r(x')$, d. h. $|x - x'| = r$, gilt

$$\begin{aligned}
\psi(x) &= |f(x) - z|^2 \geq \Big\{|f(x) - f(x')| - |y' - z|\Big\}^2 \\
&\geq \Big\{\lambda|x - x'| - \frac{\lambda}{2}r\Big\}^2 = \Big\{\lambda r - \frac{\lambda}{2}r\Big\}^2 = \frac{1}{4}\lambda^2 r^2.
\end{aligned}$$

Andererseits gilt im Mittelpunkt x' von $\overline{K}_r(x')$

$$\psi(x') = |f(x') - z|^2 = |y' - z|^2 < \left(\frac{1}{2}\lambda r\right)^2 = \frac{1}{4}\lambda^2 r^2.$$

Also nimmt ψ sein Minimum im Innern der Kugel $K_r(x')$ an, etwa in $\tilde{x} \in K_r(x')$. Dann muß aber gelten

$$\begin{aligned} 0 &= \operatorname{grad}\psi(\tilde{x}) = \Big(\dots, \mathrm{D}_j \sum_{k=1}^m |f_k(\tilde{x}) - z_k|^2, \dots\Big) \\ &= \Big(\dots, 2\sum_{k=1}^m \mathrm{D}_j f_k(\tilde{x})(f_k(\tilde{x}) - z_k), \dots\Big) = 2(\mathrm{D}f(\tilde{x}))^t(f(\tilde{x}) - z). \end{aligned}$$

Da $\mathrm{D}f(\tilde{x})$ invertierbar ist, gilt dies auch für $(\mathrm{D}f(\tilde{x}))^t$, und somit ist $f(\tilde{x}) = z$, d. h. $\tilde{x}$ ist das gesuchte x mit $z = f(x)$. Damit ist bewiesen, daß V offen ist.

Es bleibt schließlich nur noch zu zeigen $\boldsymbol{g := (f|_U)^{-1}}$ *ist in* $\boldsymbol{y_0 = f(x_0)}$ *differenzierbar mit* $\boldsymbol{Dg(y_0) = Df(x_0)^{-1}}$**:** (O. E. können wir uns auf x_0 beschränken, da jedes $x \in A$ als Ausgangspunkt gewählt werden kann.) Für $y \in \mathbb{R}^m$ mit $|y| < \varepsilon$ (ε wie oben) gilt

$$g(y_0 + y) = x_0 + x \in U = K_\delta(x_0).$$

Wegen

$$y = f(x_0 + x) - y_0 = f(x_0 + x) - f(x_0)$$

folgt aus den obigen Überlegungen

$$|y| \geq \lambda|x| \text{ bzw. } |x| \leq \frac{1}{\lambda}|y|.$$

Außerdem gilt (da f in x_0 die Ableitung $\mathrm{D}f(x_0)$ hat)

$$y - \mathrm{D}f(x_0)x = f(x_0 + x) - f(x_0) - \mathrm{D}f(x_0)x = \varphi_{f,x_0}(x_0 + x),$$

$$\mathrm{D}f(x_0)^{-1}y - x = \mathrm{D}f(x_0)^{-1}\varphi_{f,x_0}(x_0 + x)$$

mit

$$\frac{1}{|x|}\varphi_{f,x_0}(x_0+x) \to 0 \text{ für } x \to 0.$$

Damit erhalten wir

$$\begin{aligned}
&|g(y_0+y)-g(y_0)-\mathrm{D}f(x_0)^{-1}y| = |(x_0+x)-x_0-\mathrm{D}f(x_0)^{-1}y| \\
&\quad = \quad |x-\mathrm{D}f(x_0)^{-1}y| = |\mathrm{D}f(x_0)^{-1}\varphi_{f,x_0}(x_0+x)| \\
&\quad \le \quad |\mathrm{D}f(x_0)^{-1}||\varphi_{f,x_0}(x_0+x)|
\end{aligned}$$

mit

$$\frac{1}{|y|}|\mathrm{D}f(x_0)^{-1}||\varphi_{f,x_0}(x_0+x)| \le \frac{1}{\lambda}|\mathrm{D}f(x_0)^{-1}|\frac{|\varphi_{f,x_0}(x_0+x)|}{|x|} \to 0$$

für $y \to 0$ (was mit $x \to 0$ gleichbedeutend ist). Also gilt

$$g(y_0+y) = g(y_0) + \mathrm{D}f(x_0)^{-1}y + \varphi_{g,y_0}(y_0+y)$$

mit

$$\frac{1}{|y|}\varphi_{g,y_0}(y_0+y) \to 0 \text{ für } y \to 0,$$

d. h. g ist in $y_0 = f(x_0)$ differenzierbar mit $\mathrm{D}g(y_0) = \mathrm{D}f(x_0)^{-1}$. ■

6.2 Implizit definierte Funktionen

Ist $g : \mathbb{R}^2 \to \mathbb{R}$ eine „vernünftige“ (z. B. stetig differenzierbare) Funktion, so wird durch die Nullstellenmenge von g

$$M := \left\{(x,y) \in \mathbb{R}^2 : g(x,y) = 0\right\}$$

in vielen Fällen eine Kurve in $\mathbb{R}^2$ beschrieben. So ist z. B. für $g(x) = x^2+y^2-1$ die Menge M gerade die Einheitskreislinie. Da sich Kurven in der Regel (zumindest stückweise) als Graphen darstellen lassen, kommt man zu der

Frage: Gibt es eine Funktion $f : U \to \mathbb{R}$ (mit $U \subset \mathbb{R}$) so, daß der Graph von f die Menge M beschreibt? In anderen Worten, kann die Gleichung $g(x, y) = 0$ (wenigstens lokal) nach y aufgelöst werden, $y = f(x)$ mit $g(x, f(x)) = 0$?

Offensichtlich gilt: M ist sicher dann der Graph einer (allerdings nicht notwendig stetigen und überall definierten) Funktion f, wenn die Funktion

$$g_x : y \mapsto g(x, y)$$

für jedes x streng monoton ist; denn dann gibt es für jedes x höchstens ein y mit $g_x(y) = g(x, y) = 0$. — Dies gilt aber schon für so einfache Funktionen wie die oben genannte

$$g(x, y) = x^2 + y^2 - 1$$

nicht global. Offenbar läßt sich aber die Einheitskreislinie stückweise als Graph von stetig differenzierbaren Funktionen darstellen, nämlich

$$\begin{aligned} y &= \pm\sqrt{1 - x^2} \text{ für } -1 < x < 1, \\ x &= \pm\sqrt{1 - y^2} \text{ für } -1 < y < 1. \end{aligned}$$

Bevor wir den allgemeinen Fall studieren, geben wir den Satz für den bisher betrachteten einfachen Fall in $\mathbb{R}^2$ an, da hier die Situation geometrisch noch leicht erfaßbar ist.

Satz 6.4 *Sei $A \subset \mathbb{R}^2$ offen, $g \in C^1(A) = C^1(A, \mathbb{R})$ (dem Raum der stetig differenzierbaren Funktionen auf A mit Werten in $\mathbb{R}$),*

$$(x_0, y_0) \in M := \Big\{(x, y) \in A : g(x, y) = 0\Big\}, \quad \mathrm{D}_2 g(x_0, y_0) \neq 0.$$

Dann existieren offene Intervalle $U, V \subset \mathbb{R}$ mit

$$x_0 \in U, \;\; y_0 \in V, \;\; U \times V \subset A$$

und eine stetig differenzierbare Funktion $f : U \to V$ so, daß

$$M \cap (U \times V) = \Gamma_f \;\; (= \;\; \textit{Graph von } f := \{(x, f(x)) : x \in U\})$$

gilt, und

$$\mathrm{D}f(x) = f'(x) = -\frac{\mathrm{D}_1 g(x, f(x))}{\mathrm{D}_2 g(x, f(x))} \text{ für } x \in U.$$

Die Funktion f heißt eine *durch g* implizit definierte Funktion *(wie schon bemerkt, gewinnt man sie formal durch Auflösen der Gleichung $g(x,y) = 0$ nach y; aber selbst wenn diese Auflösung nicht explizit möglich ist, erlaubt die obige Formel die Berechnung der Ableitung).*

Beweis. Wegen $\mathrm{D}_2 g(x_0, y_0) \neq 0$ und der Stetigkeit von $\mathrm{D}_2 g$ existiert ein Rechteck $U \times V$ um (x_0, y_0) mit

$$\mathrm{D}_2 g(x,y) \neq 0 \text{ für } (x,y) \in U \times V.$$

Deshalb gibt es (vgl. obige Überlegung) zu jedem $x \in U$ höchstens ein $y \in V$ mit $g(x,y) = 0$, d. h. $M \cap (U \times V)$ ist Graph einer Funktion f. Die Tatsache, daß f auf einem ganzen Intervall erklärt und differenzierbar ist, werden wir im folgenden Satz allgemeiner beweisen. Wenn wir die Differenzierbarkeit von f als bekannt annehmen, ist es leicht, f' zu berechnen: Für $h(x) := g(x, f(x))$ gilt offensichtlich $h(x) \equiv 0$, also (mit $k(x) := (x, f(x))$)

$$\begin{aligned} 0 &= h'(x) = \mathrm{D}h(x) = \mathrm{D}g(k(x))\,\mathrm{D}k(x) \\ &= \Big\langle (\mathrm{D}_1 g(x, f(x)), \mathrm{D}_2 g(x, f(x)), (1, f'(x)) \Big\rangle \\ &= \mathrm{D}_1 g(x, f(x)) + f'(x)\,\mathrm{D}_2 g(x, f(x)). \end{aligned}$$

Hieraus folgt die gewünschte Formel. ∎

Korollar 6.5 *Sei $A \subset \mathbb{R}^2$ offen, $g : A \to \mathbb{R}$ stetig differenzierbar mit*

$$\operatorname{grad} g(x,y) \neq 0 \textit{ in } M := \{(x,y) \in A : g(x,y) = 0\}.$$

Dann ist M endliche Vereinigung von Graphen stetiger Funktionen f_j, zum Teil y als Funktion von x, zum Teil x als Funktion von y.

Beweis. In jedem $(x_0, y_0) \in M$ ist $\mathrm{D}_1 g(x_0, y_0) \neq 0$ oder $\mathrm{D}_2 g(x_0, y_0) \neq 0$. Deshalb gibt es ein ganzes Rechteck $U \times V$ um (x_0, y_0) mit $\mathrm{D}_1 g(x, y) \neq 0$ oder $\mathrm{D}_2 g(x, y) \neq 0$ für alle $(x, y) \in U \times V$ und somit eine Funktion $f : U \to V$ oder $\tilde{f} : V \to U$ mit $M \cap (U \times V) = \Gamma_f$ oder $M \cap (U \times V) = \Gamma_{\tilde{f}}^{-1}$ (wobei Γ_h^{-1} aus Γ_h durch Vertauschen der Komponenten entsteht). In konkreten Fällen ist in der Regel klar, daß sich M durch endlich viele solche Rechtecke überdecken läßt; im allgemeinen Fall folgt dies aus dem Satz von Heine–Borel 2.15 (vgl. auch Satz 9.6). ■

Satz 6.6 (Implizite Funktionen) *Sei $A \subset \mathbb{R}^{m+n}$ offen, $g \in C^1(A, \mathbb{R}^n)$ (dem Raum der stetig differenzierbaren Funktionen auf A mit Werten in $\mathbb{R}^n$); die Punkte aus $\mathbb{R}^{m+n}$ schreiben wir im folgenden in der Form*

$$(x, y) = (\xi_1, \ldots, \xi_m, \eta_1, \ldots, \eta_n).$$

Es sei

$$\begin{aligned} L_1(x, y) &:= (\mathrm{D}_j g_i(x, y))_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} = \begin{pmatrix} \mathrm{D}_1 g_1(x, y) & \ldots & \mathrm{D}_m g_1(x, y) \\ \vdots & \ddots & \vdots \\ \mathrm{D}_1 g_n(x, y) & \ldots & \mathrm{D}_m g_n(x, y) \end{pmatrix}, \\ L_2(x, y) &:= (\mathrm{D}_{m+j} g_i(x, y))_{i,j=1,\ldots,n} = \begin{pmatrix} \mathrm{D}_{m+1} g_1(x, y) & \ldots & \mathrm{D}_{m+n} g_1(x, y) \\ \vdots & \ddots & \vdots \\ \mathrm{D}_{m+1} g_n(x, y) & \ldots & \mathrm{D}_{m+n} g_n(x, y) \end{pmatrix} \end{aligned}$$

und somit

$$\mathrm{D}g(x, y) = \Big(L_1(x, y) \quad L_2(x, y) \Big).$$

Für ein

$$(x_0, y_0) \in M := \Big\{ (x, y) \in A : g(x, y) = 0 \Big\}$$

sei die Matrix $L_2(x_0, y_0)$ invertierbar.

Dann existieren offene Mengen $U \subset \mathbb{R}^m$ und $V \subset \mathbb{R}^n$ mit

$$(x_0, y_0) \in U \times V \subset A$$

und eine stetig differenzierbare Funktion $f : U \to V$ *mit*

$$M \cap (U \times V) = \Gamma_f = \textit{Graph von } f,$$
$$\mathrm{D}f(x) = -L_2(x, f(x))^{-1} L_1(x, f(x)).$$

Vorbemerkung zum Beweis: In der Nähe von (x_0, y_0) gilt mit „großer Genauigkeit“

$$g(x, y) \sim g(x_0, y_0) + L_1(x_0, y_0)(x - x_0) + L_2(x_0, y_0)(y - y_0).$$

Die Auflösung der Gleichung $g(x, y) = 0$ nach y in der Nähe eines Punktes (x_0, y_0) mit $g(x_0, y_0) = 0$ sollte also möglich sein, wenn die Gleichung

$$L_1(x_0, y_0)(x - x_0) + L_2(x_0, y_0)(y - y_0) = 0$$

nach y auflösbar ist. Dies ist genau dann möglich, wenn $L_2(x_0, y_0)$ invertierbar ist; die Lösung ist

$$f(x) = y = y_0 - L_2(x_0, y_0)^{-1} L_1(x_0, y_0)(x - x_0),$$

und als Ableitung (insbesondere im Punkt x_0) erhalten wir

$$\mathrm{D}f(x_0) = L_2(x_0, y_0)^{-1} L_1(x_0, y_0).$$

Beweis. Wir definieren die folgenden Hilfsfunktionen

$$\begin{aligned}
G &: A \to \mathbb{R}^{m+n} \quad (x, y) \mapsto (x, g(x, y)),\\
j_1 &: \mathbb{R}^m \to \mathbb{R}^{m+n} \quad x \mapsto (x, 0) \qquad \text{(Injektion)},\\
p_2 &: \mathbb{R}^{m+n} \to \mathbb{R}^n, \quad (x, y) \mapsto y \qquad \text{(Projektion)}.
\end{aligned}$$

Im folgenden sei E_m die m–dimensionale Einheitsmatrix ist, $0_{m,n}$ die m–zeilige und n–spaltige Nullmatrix. Damit gilt offensichtlich

$$\begin{aligned}
\mathrm{D}G(x_0, y_0) &= \begin{pmatrix} E_m & O_{mn} \\ L_1(x_0, y_0) & L_2(x_0, y_0) \end{pmatrix},\\
\det \mathrm{D}G(x_0, y_0) &= \det L_2(x_0, y_0) \neq 0
\end{aligned}$$

da $L_2(x_0, y_0)$ invertierbar ist. Also gibt es eine Umgebung von (x_0, y_0) in A, ohne Einschränkung in der Form $U_1 \times V$ wählbar mit offenen Umgebungen U_1 und V von x_0 bzw.

y_0, so, daß $G : U_1 \times V \to \mathbb{R}^{m+n}$ eine stetig differenzierbare Inverse $(G|_{U_1\times V})^{-1}$ hat; wegen $G(x_0, y_0) = (x_0, g(x_0, y_0)) = (x_0, 0)$ ist $(G|_{U\times V})^{-1}$ jedenfalls auf $U \times \{0\}$ definiert für eine geeignete Umgebung $U \subset U_1$ von x_0 in $\mathbb{R}^m$.

Wir zeigen, daß

$$f := p_2 \circ G^{-1} \circ j_1 : U \to V$$

die gewünschte Funktion ist:

- j_1 bildet U auf $U \times \{0\}$ ab, $x \mapsto (x, 0)$,
- G^{-1} bildet $U \times \{0\}$ in $U \times V$ ab, $(x, 0) \mapsto (x, y)$ mit $g(x, y) = 0$,
- p_2 bildet $U \times V$ nach V ab, $(x, y) \mapsto y$, wobei für die hier in Frage stehenden (x, y) das y die Eigenschaft hat, daß $g(x, y) = 0$ ist.

Insgesamt bildet also f jedes $x \in U$ auf das $y \in V$ ab, für das $g(x, y) = 0$ gilt; wir haben also die gesuchte Funktion von f auf einer Umgebung U von x_0 „explizit“ dargestellt.

Nun bleibt nur noch die Ableitung zu berechnen:

$$\mathrm{D}f(x) = \mathrm{D}p_2\Big(G^{-1}(j_1(x))\Big)\,\mathrm{D}G^{-1}(j_1(x))\,\mathrm{D}j_1(x).$$

Der erste und der letzte Faktor sind einfach zu berechnen,

$$\mathrm{D}p_2(\cdot) = \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix}, \quad \mathrm{D}j_1 = \begin{pmatrix} E_m \\ 0_{n,m} \end{pmatrix}.$$

Den mittleren Faktor erhalten wir aus dem Satz über die Umkehrfunktion (Satz 6.2). $\mathrm{D}G(x, y)$ haben wir bereits oben angegeben (wobei für (x_0, y_0) jetzt (x, y) zu setzen ist). Man sieht leicht, daß gilt

$$(\,\mathrm{D}G(x, y))^{-1} = \begin{pmatrix} E_m & 0_{m,n} \\ -L_2^{-1}L_1 & L_2^{-1} \end{pmatrix} \quad \text{mit} \quad L_j := L_j(x, y),$$

also, mit der Formel für die Ableitung der Umkehrfunktion,

$$\begin{aligned} \mathrm{D}G^{-1}(j_1(x)) &= \mathrm{D}G^{-1}(x, 0) = \Big(\mathrm{D}G(G^{-1}(x, 0))\Big)^{-1} \\ &= \Big(\mathrm{D}G(x, f(x))\Big)^{-1} = \begin{pmatrix} E_m & 0_{m,n} \\ -M_2^{-1}M_1 & M_2^{-1} \end{pmatrix} \quad \text{mit} \quad M_j := L_j(x, f(x)), \end{aligned}$$

und damit erhalten wir

$$\begin{aligned}
\mathrm{D}f(x) &= \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix} \begin{pmatrix} E_m & 0_{m,n} \\ -M_2^{-1}M_1 & M_2^{-1} \end{pmatrix} \begin{pmatrix} E_m \\ 0_{n,m} \end{pmatrix} \\
&= \begin{pmatrix} 0_{n,m} & E_n \end{pmatrix} \begin{pmatrix} E_m \\ -M_2^{-1}M_1 \end{pmatrix} = -M_2^{-1}M_1 \\
&= -L_2(x, f(x))^{-1} L_1(x, f(x)).
\end{aligned}$$

Das ist die behauptete Formel. ■

Im folgenden Abschnitt verwenden wir dieses Resultat zur Beschreibung von Hyperflächen in $\mathbb{R}^m$. In diesem Fall ist $n = 1$ und an die Stelle von m im obigen Satz tritt $m - 1$.

6.3 Übungsaufgaben

6.1 Man zeige, daß die Abbildung

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \;\; f(x, y) = (x + y^2, y + x^2)$$

in $\{(x, y) \in \mathbb{R}^2 : |(x, y)| \leq \frac{1}{4}\}$ umkehrbar ist und berechne die Ableitung von f^{-1} in $(0, 0)$. (Man benutze "U" aus dem Beweis des Satzes über die Umkehrfunktion.)

6.2 Sei $A \subset \mathbb{R}^m$ offen und $f : A \to \mathbb{R}^n$ $(n \leq m)$ stetig differenzierbar. Hat $Df(x)$ für jedes $x \in A$ den Rang n, so ist $f(A)$ offen. (Der Spezialfall $n = 1$ wurde in Aufgabe 3.5 behandelt, während der Fall $n = m$ unmittelbar aus dem Satz über die Umkehrfunktion folgt.)

6.3 Sei $A \subset \mathbb{R}^2$ offen, $z : A \to \mathbb{R}$ sei so, daß

$$\left.\begin{array}{l} 3z(x, y)^2 - x \neq 0 \\ z(x, y)^3 - x\, z(x, y) - y = 0 \end{array}\right\} \quad \text{für alle} \quad (x, y) \in A$$

gilt. Man zeige, daß z die partielle Differentialgleichung

$$z_{xy}(3z^2 - x)^3 + 3z^2 + x = 0$$

erfüllt (dabei ist $z_{xy} = \dfrac{\partial}{\partial x}\dfrac{\partial}{\partial y} z = \dfrac{\partial}{\partial y}\dfrac{\partial}{\partial x} z$).

6.4 Man zeige, daß das Gleichungssystem

$$
\begin{aligned}
x^2 + y^2 - 2z^2 &= 0, \\
x^2 + 2y^2 + z^2 &= 4
\end{aligned}
$$

für hinreichend kleine x durch positive Funktionen $y(x)$ und $z(x)$ aufgelöst werden kann und berechne ohne explizite Auflösung y' (in Abhängigkeit von x und $y(x)$) und z' (in Abhängigkeit von x und $z(x)$).

7 Hyperflächen in $\mathbb{R}^m$ und bedingte Extrema

7.1 Hyperflächen in $\mathbb{R}^m$

In $\mathbb{R}^m$ wird durch

$$M := \{x \in \mathbb{R}^m : g(x) = 0\}$$

eine *Hyperfläche* (Dimension $= m-1$) definiert, wenn $g : \mathbb{R}^m \to \mathbb{R}$ stetig differenzierbar ist mit

$$\operatorname{grad} g(x) \neq 0 \text{ für alle } x \in M.$$

Ist $x_0 = (\xi_{01}, \dots, \xi_{0m}) \in M$ und z. B. $\mathrm{D}_m g(x_0) \neq 0$, so läßt sich nach Satz 6.6 ein Stück von M um x_0 als Graph einer Funktion

$$\xi_m = h_m(\xi_1, \dots, \xi_{m-1}) \quad (h_m : B \to \mathbb{R} \text{ mit } B \subset \mathbb{R}^{m-1})$$

schreiben und es gilt für $j = 1, \dots, m-1$

$$\mathrm{D}_j h_m(\xi_1, \dots, \xi_{m-1}) = -\frac{\mathrm{D}_j g(\xi_1, \dots, \xi_{m-1}, h_m(\xi_1, \dots, \xi_{m-1}))}{\mathrm{D}_m g(\xi_1, \dots, \xi_{m-1}, h_m(\xi_1, \dots, \xi_{m-1}))}.$$

Entsprechendes gilt, wenn $\mathrm{D}_j g(x_0) \neq 0$ ist für ein $j = 1, \dots, m-1$. Die gesamte Menge M läßt sich aus solchen Teilstücken zusammensetzen.

Beispiel 7.1 Für $g(x) = \xi_1^2 + \xi_2^2 + \xi_3^2 - 1$ ist

$$M = \{x \in \mathbb{R}^3 : g(x) = 0\} \text{ die } \textit{Einheitssphäre} \text{ in } \mathbb{R}^3;$$

offenbar gilt $\operatorname{grad} g(x) = 2x \neq 0$ für alle $x \in M$. Es gilt z. B.

$$\xi_3 = h_3(\xi_1, \xi_2) = \pm\sqrt{1 - \xi_1^2 - \xi_2^2}$$

und somit für $j = 1, 2$

$$\begin{aligned} \mathrm{D}_j h_3(\xi_1, \xi_2) &= -\frac{\mathrm{D}_j g(\xi_1, \xi_2, \pm\sqrt{1-\xi_1^2-\xi_2^2})}{\mathrm{D}_3 g(\xi_1, \xi_2, \pm\sqrt{1-\xi_1^2-\xi_2^2})} \\ &\quad \Big(\text{wegen } \mathrm{D}_k g(\eta_1, \eta_2, \eta_3) = 2\eta_k \text{ für } k = 1, 2, 3\Big) \\ &= \mp\frac{2\xi_j}{2\sqrt{1-\xi_1^2-\xi_2^2}}. \end{aligned}$$

□

Im Fall des obigen Beispiels, wo h_3 explizit bekannt ist, können die Ableitungen natürlich (sogar etwas schneller) direkt aus der Darstellung von h_3 berechnet werden. Der wesentliche Vorteil der allgemeinen Aussage ist, daß unter sehr schwachen Bedingungen ein solches h_3 existiert, und daß dies stetig differenzierbar ist.

Sei M eine (zunächst beliebige) Teilmenge von $\mathbb{R}^m$, $x_0 \in M$. Wir sagen *der Vektor $z \in \mathbb{R}^m$ steht in x_0 senkrecht auf M*, wenn z in x_0 senkrecht auf jeder in M liegenden (im Punkt x_0) differenzierbaren Kurve durch x_0 steht (d. h. senkrecht auf dem Tangentialvektor in x_0 an jede Kurve in M durch x_0; in etwas anderer Sprechweise: senkrecht auf dem *Tangentialraum* an M in x_0).

Satz 7.2 *Sei $A \subset \mathbb{R}^m$ offen, $g : A \to \mathbb{R}$ stetig differenzierbar,*

$$M = \{x \in A : g(x) = 0\}, \quad x_0 \in M, \quad \operatorname{grad} g(x_0) \neq 0.$$

Ein Vektor $z \in \mathbb{R}^m$ steht genau dann in x_0 senkrecht auf M, wenn ein $\lambda \in \mathbb{R}$ existiert mit

$$z = \lambda \operatorname{grad} g(x_0) \ \textit{(i.a. Worten, wenn gilt } z \in L\{\operatorname{grad} g(x_0)\}\textit{)}.$$

Beweis. Wir zeigen im folgenden:

(i) $\operatorname{grad} g(x_0)$ steht in x_0 senkrecht auf jeder in x_0 differenzierbaren Kurve in M durch x_0 (d. h. $\operatorname{grad} g(x_0)$ steht in x_0 senkrecht auf M),

(ii) der Raum der in x_0 senkrecht auf M stehenden Vektoren ist höchstens 1–dimensional.

Insgesamt folgt damit die Behauptung.

(i) Sei $\gamma : [a,b] \to M$ stetig mit $\gamma(c) = x_0$ für ein $c \in (a,b)$, γ differenzierbar in c. Dann gilt $g(\gamma(t)) \equiv 0$ für $t \in [a,b]$ und somit

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} g(\gamma(t))\Big|_{t=c} = \mathrm{D}g(\gamma(c))\,\mathrm{D}\gamma(c) = \langle \operatorname{grad} g(x_0), \gamma'(c) \rangle,$$

d. h. $\operatorname{grad} g(x_0)$ steht in x_0 senkrecht auf γ.

(ii) Ohne Einschränkung ist M in einer Umgebung von x_0 als Graph einer stetig differenzierbaren Funktion darstellbar, etwa in der Form

$$\xi_m = h_m(\xi_1, \ldots, \xi_{m-1}) \quad (h_m : B \to \mathbb{R} \text{ mit } B \subset \mathbb{R}^{m-1}).$$

Weiter sei o.E. $x_0 = 0 = (0, \ldots, 0)$. Wir betrachten für hinreichend kleines $\varepsilon > 0$ die $m-1$ Kurven durch x_0

$$\begin{aligned}
\gamma_1 \ &: \ [-\varepsilon, \varepsilon] \to \mathbb{R}^m, \quad t \mapsto (t, 0, \ldots, 0, h_m(t, 0, \ldots, 0)),\\
&\ \vdots\\
\gamma_{m-1} \ &: \ [-\varepsilon, \varepsilon] \to \mathbb{R}^m, \quad t \mapsto (0, \ldots, 0, t, h_m(0, \ldots, 0, t)).
\end{aligned}$$

Diese sind offenbar differenzierbar mit den Tangentialrichtungen (Ableitungen) in x_0

$$\begin{aligned}
\gamma_1'(0) &= (1, 0, \ldots, 0, \mathrm{D}_1 h(0, \ldots, 0)),\\
&\ \vdots\\
\gamma_{m-1}'(0) &= (0, \ldots, 0, 1, \mathrm{D}_{m-1} h(0, \ldots, 0)).
\end{aligned}$$

Da diese $m-1$ Tangentialvektoren linear unabhängig sind, kann es (bis auf einen Faktor) höchstens einen Vektor geben, der auf allen senkrecht steht. ■

Der folgende Satz enthält den vorhergehenden als Spezialfall

Satz 7.3 *Sei $A \subset \mathbb{R}^m$ offen,*

$$\left.\begin{aligned}
&g_j : A \to \mathbb{R} \text{ stetig differenzierbar,}\\
&M_j := \{x \in A : g_j(x) = 0\}
\end{aligned}\right\} \quad \text{für } j = 1, \ldots, k,$$

$$M := \bigcap_{j=1}^{k} M_j, \ x_0 \in M,$$

$$\Big\{ \operatorname{grad} g_j(x_0) : j = 1, \ldots, k \Big\} \text{ linear unabhängig.}$$

Ein Vektor $z \in \mathbb{R}^m$ steht genau dann in x_0 senkrecht auf M, wenn gilt

$$z \in L\Big\{\operatorname{grad} g_j(x_0) : j = 1, \ldots, k\Big\}.$$

Beweis. In Analogie zum vorhergehenden Satz beweisen wir:

(i) Jedes $z \in L\Big\{\operatorname{grad} g_j(x_0) : j = 1, \ldots, k\Big\}$ steht in x_0 senkrecht auf M,

(ii) der Raum der Vektoren, die in x_0 senkrecht auf M stehen, ist höchstens k–dimensional.

Wegen $\dim L\Big\{\operatorname{grad} g_j(x_0) : j = 1, \ldots, k\Big\} = k$ folgt damit die Behauptung.

(i) Jeder Vektor $\operatorname{grad} g_j(x_0)$ steht in x_0 senkrecht auf M_j und somit auf $M \subset M_j$. Da dies für alle j gilt, steht jedes $z \in L\Big\{\operatorname{grad} g_j(x_0) : j = 1, \ldots, k\Big\}$ in x_0 senkrecht auf M.

(ii) Da $\Big\{\operatorname{grad} g_j(x_0) : j = 1, \ldots, k\Big\}$ linear unabhängig ist, hat die Matrix

$$\begin{pmatrix} D_1 g_1(x_0) & \ldots & D_m g_1(x_0) \\ \vdots & & \vdots \\ D_1 g_k(x_0) & \ldots & D_m g_k(x_0) \end{pmatrix} \quad \text{den Rang } k.$$

Durch Umnumerierung der Koordinaten (von x) kann man also erreichen, daß die Matrix

$$\begin{pmatrix} D_{m-k+1} g_1(x_0) & \ldots & D_m g_1(x_0) \\ \vdots & & \vdots \\ D_{m-k+1} g_k(x_0) & \ldots & D_m g_k(x_0) \end{pmatrix} \quad \text{invertierbar}$$

ist. Damit ist die Voraussetzung des Satzes über implizierte Funktionen (mit $m-k$, $\quad k$ statt n, m) erfüllt für

$$f : \mathbb{R}^m = \mathbb{R}^{(m-k)+k} \to \mathbb{R}^k, \quad x \mapsto (g_1(x), \ldots, g_k(x)),$$

und ein Stück der Menge $M = \Big\{x \in \mathbb{R}^m : f(x) = 0\Big\}$ um x_0 kann als Graph geschrieben werden:

$$(\xi_{m-k+1}, \ldots, \xi_m) = h(\xi_1, \ldots, \xi_{m-k}) \quad (h : B \to \mathbb{R}^k \text{ mit } B \subset \mathbb{R}^{m-k}).$$

Damit können wir (für hinreichend kleines $\varepsilon > 0$) $m-k$ Kurven in M durch x_0 angeben (o.E.sei wieder $x_0 = 0$):

$$\begin{aligned}\gamma_1 &: [-\varepsilon, \varepsilon] \to \mathbb{R}^m, \quad t \mapsto (t, 0, \ldots, 0, h(t, 0, \ldots, 0)),\\ &\vdots\\ \gamma_{m-k} &: [-\varepsilon, \varepsilon] \to \mathbb{R}^m, \quad t \mapsto (0, \ldots, 0, t, h(0, \ldots, 0, t))\end{aligned}$$

mit den $m-k$ linear unabhängigen Tangentialvektoren in $x_0 = 0$

$$\begin{aligned}\gamma_1'(0) &= (1, 0, \ldots, 0, \mathrm{D}_1 h(0, \ldots, 0)),\\ &\vdots\\ \gamma_{m-k}'(0) &= (0, \ldots, 0, 1, \mathrm{D}_{m-k} h(0, \ldots, 0)).\end{aligned}$$

Also kann es höchstens k linear unabhängige Vektoren geben, die in x_0 auf M (d. h. auf $\gamma_1'(0), \ldots, \gamma_{m-k}'(0)$) senkrecht stehen. ■

7.2 Bedingte Extrema

Sei $A \subset \mathbb{R}^m$ offen, $M \subset A$, $f : A \to \mathbb{R}$, $x_0 \in M$. Man sagt, f nimmt in x_0 ein *lokales Minimum auf* M (*bedingtes lokales Minimum*) an, wenn ein $\delta > 0$ existiert mit

$$f(x) \geq f(x_0) \text{ für alle } \boldsymbol{x \in M} \text{ mit } |x - x_0| < \delta.$$

Entsprechend definiert man *lokales Maximum auf* M (*bedingtes lokales Maximum*) oder allgemeiner *bedingte lokale Extrema.*

Im Unterschied zu früher suchen wir also nicht Extrema in bezug auf die offene Menge A, sondern in bezug auf die Menge $M \subset A$. Diese ist i.allg. (insbesondere im folgenden) nicht offen, so daß die früher entwickelten Methoden (vgl. Abschnitt 4.3) nicht mehr anwendbar sind. Wir suchen Extrema unter der zusätzlichen Bedingung $x \in M$, weshalb man von bedingten Extrema spricht.

Im folgenden entwickeln wir eine Methode zum Auffinden von Extrema unter Bedingungen der Form

$$x \in M = \Big\{x \in \mathbb{R}^m : g_j(x) = 0 \text{ für } j = 1, \ldots, k\Big\}.$$

Satz 7.4 (Lagrange–Multiplikation) *Sei $A \subset \mathbb{R}^m$ offen,*

$$\left.\begin{array}{l} g_j : A \to \mathbb{R} \text{ stetig differenzierbar,} \\ M_j := \{x \in A : g_j(x) = 0\} \end{array}\right\} \quad \text{für } j = 1, \dots, k,$$
$$M := \bigcap_{j=1}^{k} M_j, \quad x_0 \in M,$$
$$\Big\{ \operatorname{grad} g_j(x_0) : j = 1, \dots, k \Big\} \text{ linear unabhängig.}$$

Ist $f : A \to \mathbb{R}$ stetig differenzierbar und hat f in x_0 ein bedingtes lokales Extremum auf M, so gibt es reelle Zahlen $\lambda_1, \dots, \lambda_k$ mit

$$\operatorname{grad} f(x_0) = \sum_{j=1}^{k} \lambda_j \operatorname{grad} g_j(x_0) \, (d.\,h. \operatorname{grad} f(x_0) \in L\Big\{ \operatorname{grad} g_j(x_0) : j = 1, \dots, k \Big\})$$

(die Zahlen λ_j heißen Lagrange–Multiplikatoren*).*

Beweis. Ist $\gamma : [a, b] \to M$ eine (in x_0) differenzierbare Kurve in M durch x_0, $\gamma(c) = x_0$ für ein $c \in (a, b)$, so hat die Funktion

$$h : [a, b] \to \mathbb{R}, \quad h(t) = f(\gamma(t))$$

in c ein lokales Extremum. Da h in c differenzierbar ist, gilt also

$$0 = h'(c) = \langle \operatorname{grad} f(\gamma(c)), \gamma'(c) \rangle,$$

d. h. $\operatorname{grad} f(x_0) = \operatorname{grad} f(\gamma(c))$ steht in $x_0 = \gamma(c)$ senkrecht auf γ. Da dies für alle γ gilt, steht $\operatorname{grad} f(x_0)$ in x_0 senkrecht auf M. Damit folgt die Behauptung aus dem vorhergehenden Satz. ■

Um Punkte x_0 zu finden, in denen *möglicherweise* bedingte Extrema vorliegen, hat man also das (nichtlineare) Gleichungssystem

$$\begin{array}{ll} g_j(x_0) = 0 \quad j = 1, \dots, k & (k \text{ Gleichungen}) \\ \operatorname{grad} f(x_0) = \sum\limits_{j=1}^{k} \lambda_j \operatorname{grad} g_j(x_0) & (m \text{ Gleichungen}) \end{array}$$

nach $x_0 = (\xi_{01}, \dots, \xi_{0m}), \lambda_1, \dots, \lambda_k$ aufzulösen (wobei wir uns allerdings in der Regel für die Lagrange–Multiplikatoren nicht interessieren). Es handelt sich also um $k + m$ Gleichungen für $k + m$ Unbekannte; wir haben somit eine gute Chance, daß das System lösbar ist.

Beispiel 7.5 In $A = \mathbb{R}^2$ sei $g(x) = \xi_1^2 + \xi_2^2 - 1$, also

$$M = \left\{x \in \mathbb{R}^2 : \xi_1^2 + \xi_2^2 = 1\right\} \quad \text{die Einheitskreislinie.}$$

Wir wollen die Extrema der Funktion

$$f(x) = \xi_1 + 2\xi_2$$

auf M bestimmen. Das zu lösende Gleichungssystem ist

$$\begin{aligned} \xi_1^2 + \xi_2^2 - 1 &= 0, & (1)\\ \Big(\operatorname{grad} f(x) - \lambda \operatorname{grad} g(x) = \Big)\ (1,2) - \lambda(2\xi_1, 2\xi_2) &= 0. & (2) \end{aligned}$$

Aus (2) folgt $\lambda \neq 0$ und

$$\xi_1 = \frac{1}{2\lambda}, \quad \xi_2 = \frac{1}{\lambda}. \tag{3}$$

Einsetzen in (1) liefert

$$\frac{1}{4\lambda^2} + \frac{1}{\lambda^2} - 1 = 0, \quad \lambda^2 = \frac{5}{4}, \quad \lambda = \pm\frac{1}{2}\sqrt{5}.$$

Einsetzen in (3) ergibt

$$\xi_1 = \pm\frac{1}{\sqrt{5}}, \quad \xi_2 = \pm\frac{2}{\sqrt{5}}.$$

Die Funktionswerte in diesen Punkten sind

$$f\left(\pm\frac{1}{\sqrt{5}}, \pm\frac{2}{\sqrt{5}}\right) = \pm\frac{5}{\sqrt{5}} = \pm\sqrt{5} \sim \pm 2,236\ldots.$$

Da die lineare Funktion f auf der Kreislinie M sicher ein (striktes) Maximum und Minimum hat, liegt also in $\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ das Maximum und in $\left(-\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)$ das Minimum auf M vor. $\square$

Beispiel 7.6 In $A = \mathbb{R}^3$ sei

$$g_1(x) = \xi_1^2 + \xi_3^2 - 1, \quad g_2(x) = \xi_1 + \xi_2,$$

also ist

$$M = \left\{ x \in \mathbb{R}^3 : \xi_1^2 + \xi_3^2 - 1 = 0,\ \xi_1 + \xi_2 = 0 \right\}$$

der Schnitt

- eines Zylindermantels mit Radius 1 längs der ξ_2–Achse und
- einer Ebene senkrecht zur ξ_1–ξ_2–Ebene durch die Winkelhalbierende im 2.und 4.Quadranten

Wir wollen die Extrema der Funktion

$$f(x) = \xi_1 - \xi_2$$

auf M bestimmen. Das zu lösende Gleichungssystem ist

$$\begin{aligned} \xi_1^2 + \xi_3^2 - 1 &= 0, & (1)\\ \xi_1 + \xi_2 &= 0, & (2)\\ (1,-1,0) - \lambda_1(2\xi_1, 0, 2\xi_3) - \lambda_2(1,1,0) &= 0. & (3) \end{aligned}$$

Aus der 2.Komponente von (3) folgt

$$\lambda_2 = -1$$

und somit aus der 1.Komponente von (3)

$$\lambda_1 \neq 0.$$

Damit folgt aus der 3.Komponente von (3)

$$\xi_3 = 0.$$

Aus (1) folgt dann

$$\xi_1^2 = 1, \quad \xi_1 = \pm 1,$$

und somit aus (2)

$$\xi_2 = \mp 1.$$

Die Funktionswerte von f in diesen Punkten sind

$$f(\pm 1, \mp 1, 0) = \pm 2.$$

Es liegt also in $(1, -1, 0)$ das Maximum und in $(-1, 1, 0)$ das Minimum der Funktion f auf M vor, da auch hier offensichtlich ist, daß f auf M ein (striktes) Maximum und Minimum hat. □

7.3 Übungsaufgaben

7.1 Welcher Punkt der Fläche $z = x^2 + y^2$ liegt dem Punkt $(1, 1, \frac{1}{2})$ am nächsten?

7.2 Welche Kantenlängen hat ein Quader mit Volumen 1 und minimaler Oberfläche?

7.3 Man bestimme den in die Einheitskugel in $\mathbb{R}^m$ einbeschriebenen Quader mit dem größten Volumen.

7.4 Welche Kantenlängen hat ein Quader mit Volumen 1 und minimaler Summe der Kantenlängen?

7.5 Man löse Aufgabe 4.3 mit Hilfe der Lagrange–Multiplikatoren nochmals: Welche Kantenlänge hat ein Quader mit dem Volumen 1, wenn die Oberfläche ohne die Grundfläche minimal sein soll?

7.6 Für einen trapezförmigen Kanal ist die Fläche Q des Querschnitts vorgegeben.

a) Wie müssen die Basisbreite a, die Höhe h und der Neigungswinkel α der Böschung gewählt werden, damit die Benetzungsfläche minimal wird?

b) Wie sieht es aus, wenn die Benetzungsfläche möglichst groß sein soll?

c) Unter der Annahme, daß pro Flächeneinheit der Benetzungsfläche 2 mal so viel versickert wir pro Flächeneinheit der Oberfläche verdunstet, minimiere man den Flüssigkeitsverlust.

7.7 Man bestimme das Maximum und das Minimum der Funktion $f(x, y) = (x-1)^2 + y^2$ im Bereich $\{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ mit $g(x, y) = x^2 + y^2 - 4$. Dabei bestimme man die Extrema auf dem Rand auf drei verschiedene Arten:

- mit Hilfe geometrischer Überlegungen,
- mit Hilfe der Lagrange–Multiplikatoren,
- durch Elimination einer Variablen.

8 Das Riemannintegral in $\mathbb{R}^m$

8.1 Definition des Riemannintegrals

Wir betrachten (zunächst) *beschränkte* Funktionen $f : I \to \mathbb{R}$ auf einen *Intervall* (*Quader*) in $\mathbb{R}^m$

$$I = \Big\{x \in \mathbb{R}^m : a_i \le \xi_i \le b_i \text{ für } i = 1, \ldots, m\Big\}.$$

Mit Z bezeichnen wir *Zerlegungen* von I in Teilintervalle

$$Q_j = \Big\{x \in \mathbb{R}^m : a_{ji} \le \xi_i \le b_{ji} \text{ für } i = 1, \ldots, m\Big\} \quad (j = 1, \ldots, n = n(Z)),$$

die sich höchstens in Randpunkten, Randlinien, Randflächen, ... überschneiden dürfen; $I = \cup Q_j$. Zu jeder solchen Zerlegung definieren wir die *Untersumme* bzw. *Obersumme* zu f durch

$$\begin{aligned} \underline{S}_Z(f) &:= \sum_j |Q_j| \inf\{f(x) : x \in Q_j\}, \\ \overline{S}_Z(f) &:= \sum_j |Q_j| \sup\{f(x) : x \in Q_j\}, \end{aligned}$$

wobei $|Q_j| := \prod_{i=1}^m (b_{ji} - a_{ji})$ das *Volumen* von Q_j ist; in diesem Sinne haben die Schnittmengen der Q_j stets das Volumen 0.

Damit können wir das *Unterintegral* und das *Oberintegral* von f definieren durch

$$\int_* f(x)\,\mathrm{d}x := \sup_Z \underline{S}_Z(f) \text{ bzw. } \int^* f(x)\,\mathrm{d}x := \inf_Z \overline{S}_Z(f),$$

wobei inf und sup über alle Zerlegungen Z von I zu erstrecken sind.

Die beschränkte Funktion $f : I \to \mathbb{R}$ heißt *Riemann–integrierbar*, wenn gilt

$$\int_* f(x)\,\mathrm{d}x = \int^* f(x)\,\mathrm{d}x.$$

Dies ist offenbar genau dann der Fall, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung Z gibt, für die $\overline{S}_Z(f) - \underline{S}_Z(f) < \varepsilon$ gilt (*Riemannsches Integrabilitätskirterium*). Man definiert dann das Riemannintegral durch

$$\int_I f(x)\,\mathrm{d}x = \int f(x)\,\mathrm{d}x := \int_* f(x)\,\mathrm{d}x = \int^* f(x)\,\mathrm{d}x.$$

Wenn Verwechslungen ausgeschlossen sind, lassen wir den Integrationsbereich weg.

Ist $f : I \to \mathbb{R}$ stetig, so ist es offensichtlich Riemann–integrierbar: Da f gleichmäßig stetig ist, kommen sich bei hinreichen feiner Zerlegung die Ober– und die Oper– und die Untersumme beliebig nahe. Mit Satz 8.2 wird diese Aussage verallgemeinert.

Wie in $\mathbb{R}^1$ kann das Integral einer Riemannintegrierbaren Funktion auch mittels Riemannsummen berechnet werden:

Ist $D := \{Q_1, \ldots, Q_n; x_1, \ldots, x_n\}$ eine *Zerlegung* von I in Teilintervalle $Q_1, \ldots, Q_n$ mit *Stützstellen* $x_j \in Q_j$, so heißt

$$R_D(f) := \sum_{j=1}^{m} |Q_j| f(x_j)$$

die zu D gehörige *Riemannsumme* (deren Wert also nicht nur von den Q_j abhängt, sondern auch von der Wahl der Stützstellen x_j). Als *Feinheit* der Zerlegung D bezeichnen wir

$$\eta_D := \max\{\text{ Durchmesser von } Q_j : j = 1, \ldots, n\}.$$

Satz 8.1 *(Riemannsummen) Ist $f : I \to \mathbb{R}$ Riemann–integrierbar, so gilt für jede Folge (D_k) von Zerlegungen mit Stützstellen, für die $\eta_{D_k} \to 0$ gilt, ${}_{D_k}(f) \to \int f(x)\,\mathrm{d}x$.*

Beweis. Offenbar genügt es zu zeigen: Zu jeder Zerlegung Z und jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ so, daß für jede Zerlegung D mit Feinheit $\eta_D < \delta$ gilt

$$\underline{S}_Z(f) - \varepsilon \leq R_D \leq \overline{S}_Z(f) + \varepsilon.$$

Sei $F = F(Z)$ die Gesamtfläche ($= (m-1)$–dimensionales Volumen der Begrenzungsflächen der Intervalle von Z, D eine Zerlegung mit Stützstellen der Feinheit $\eta_D \leq \dfrac{\varepsilon}{2FC} =: \delta$ (mit

$C := \sup\{|f(x)| : x \in I\}$). Wählen wir für D' die Zerlegung, die sich durch Übereinanderlegen von Z und D ergibt; in den Intervallen von D', die auch in D vorkommen, wählen wir die alten Stützstellen, in den neu entstehenden Intervallen können beliebige Stützstellen gewählt werden. Das Gesamvolumen der neu entstehenden Intervalle ist auf Grund der Feinheit der Zerlegung D offenbar $\leq F\delta$. Damit folgt (da D' eine Verfeinerung von Z ist)

$$\begin{aligned} \underline{S}_Z(f) &\leq R_{D'}(f) \leq R_D(f) + 2C \cdot F\delta \leq R_D(f) + \varepsilon, \\ \overline{S}_Z(f) &\geq R_{D'}(f) \geq R_D(f) - 2C \cdot F\delta \geq R_D(f) - \varepsilon. \end{aligned}$$

Insgesamt folgt die gewünschte Ungleichung. ■

Damit ergeben sich sofort die zu erwartenden Eigenschaften des Riemannintegrals:

— *Positivität*: Ist $f \geq c$, so gilt $\int f(x)\,\mathrm{d}x \geq c|I|$. Insbesondere gilt $\int f(x)\,\mathrm{d}x \geq 0$ falls $f \geq 0$ gilt, $\int f(x)\,\mathrm{d}x > 0$ falls $f \geq c > 0$ gilt auf einem Teilintervall Q mit $|Q| > 0$.

— *Linearität*: $\int (af(x) + bg(x))\,\mathrm{d}x = a \int f(x)\,\mathrm{d}x + b \int g(x)\,\mathrm{d}x$.

— Sind $f, g : I \to \mathbb{R}$ Riemann–integrierbar, so ist auch fg Riemann–integrierbar.

— Sind $f_n (n \in \mathbb{N})$ Riemannintegrierbar und gilt $f_n \to f$ gleichmäßig, so ist auch f Riemann-integrierbar und es gilt $\int f_n(x)\,\mathrm{d}x \to \int f(x)\,\mathrm{d}x$.

Eine Menge $N \subset \mathbb{R}^m$ heißt eine *Jordan–Nullmenge* (Menge mit dem Jordan–Maß 0), wenn zu jedem $\varepsilon > 0$ endlich viele Intervalle existieren, deren Vereinigung N enthält, und deren Gesamtvolumen $< \varepsilon$ ist.[3] Man sieht leicht:

— Jede endliche Menge ist Jordan–Nullmenge,

— jede beschränkte Menge mit (nur) endlich vielen Häufungspunkten ist Jordan–Nullmenge,

— ist A ein Intervall in $\mathbb{R}^{m-1}$, $h : A \to \mathbb{R}$ stetig differenzierbar, so ist

$$M = \Big\{x \in \mathbb{R}^m : (\xi_1, \ldots, \xi_{m-1}) \in A,\ \xi_m = h(\xi_1, \ldots, \xi_{m-1})\Big\}$$

eine Jordan–Nullmenge (man kann sie in eine Vereinigung von, bezüglich ξ_m, beliebig flachen Quadern einschließen). Da auch endliche Vereinigungen von Jordan–Nullmengen wieder Jordan–Nullmengen sind, sind die im letzten Abschnitt betrachteten Hyperflächen in der Regel Jordan–Nullmengen.

[3] Im Gegensatz dazu heißt eine Menge $N \subset \mathbb{R}^m$ eine *Lebesgue–Nullmenge*, wenn zu jedem $\varepsilon > 0$ abzählbar viele Intervalle existieren, deren Vereinigung N enthält, und deren Gesamtvolumen $< \varepsilon$ ist (d. h. deren Volumina summierbar sind mit Summe $< \varepsilon$.

Satz 8.2 *Ist $f : I \to \mathbb{R}$ beschränkt und außerhalb einer Jordan–Nullmenge stetig, so ist f Riemann–integrierbar.*[4]

Beweis. Zu $\varepsilon > 0$ sei O_ε die Vereinigung endlich vieler offener Intervalle, die die Unstetigkeiten von f enthalten und ein Gesamtvolumen $< \varepsilon$ haben. In der abgeschlossenen Restmenge $A_\varepsilon := I \setminus O_\varepsilon$ ist f stetig, also gleichmäßig stetig. Die über A_ε erstreckten Ober– bzw. Untersummen kommen sich also bei geeigneter Wahl der Zerlegung beliebig nahe. Ist C die Schranke für f, so beträgt der Unterschied zwischen Ober– und Untersumme auf O_ε höchstens $2C\varepsilon$. Folglich kommen sich Ober– und Untersumme (auf ganz I) beliebig nahe, d. h. f ist integrierbar. ∎

Sei A eine Teilmenge von $\mathbb{R}^m$. Ein Punkt $x \in \mathbb{R}^m$ heißt ein *Randpunkt* von A, wenn in jeder Umgebung von x (oder Kugel um x) mindestens je ein Punkt aus A und aus $\mathbb{R}^m \setminus A$ liegt. Offensichtlich ist x genau dann Randpunkt von A, wenn es Randpunkt von $\mathbb{R}^m \setminus A$ ist. Die Menge ∂A der Randpunkte von A wird als *Rand* von A bezeichnet. Man mache sich an Beispielen den Begriff klar!

Die *charakteristische Funktion* χ_A einer Teilmenge A von $\mathbb{R}^m$ ist definiert durch

$$\chi_A(x) := \begin{cases} 1 & \text{für } x \in A\,, \\ 0 & \text{für } x \in \mathbb{R}^m \setminus A\,. \end{cases}$$

Satz 8.3 *Sei $A \subset \mathbb{R}^m$ beschränkt. Die charakteristische Funktion χ_A von A ist genau dann Riemann–integrierbar, wenn ∂A eine Jordan–Nullmenge ist. Man sagt dann A ist Jordan–meßbar und definiert das* Volumen *(bzw. das* Jordan–Maß*) von A durch*

$$|A| := \int \chi_A(x)\,\mathrm{d}x\,.$$

Der *Beweis* von „$\Leftarrow$“ folgt direkt aus obigem Satz, da die Randpunkte von A genau die Unstetigkeitspunkte von χ_A sind. Den Teil „$\Rightarrow$“ wollen wir hier nicht beweisen; er spielt im folgenden keine Rolle.

[4]Tatsächlich gilt: Eine Funktion ist *genau dann* Riemann–integrierbar, wenn sie außerhalb einer Lebesgue–Nullmenge stetig ist.

Ist $f : I \to \mathbb{R}$ Riemann–integrierbar und $A \subset I$ Jordan–meßbar, so definiert man

$$\int\limits_A f(x)\,\mathrm{d}x := \int\limits_I \chi_A(x) f(x)\,\mathrm{d}x\,.$$

Wenn eine Funktion $f : \mathbb{R}^m \to \mathbb{R}$ außerhalb eines (geeigneten) Intervalls I verschwindet und über I integrierbar ist, schreiben wir auch einfach $\int f(x)\,\mathrm{d}x$ statt $\int_I f(x)\,\mathrm{d}x$.

8.2 Berechnung m–dimensionaler Integrale

Wir wenden uns jetzt der Frage zu, wie m–dimensionale Integrale explizit berechnet werden können.

Satz 8.4 (Satz von Fubini) *Sei I wie oben, $f : I \to \mathbb{R}$ sei Riemann–integrierbar.*

a) Es gilt

$$\int\limits_I f(x)\,\mathrm{d}x = \int\limits_{a_m}^{b_m} \cdots \int\limits_{a_2}^{b_2} \left\{ \int\limits_{a_1}^{b_1} f(\xi_1, \xi_2, \ldots, \xi_m)\,\mathrm{d}\xi_1 \right\} \mathrm{d}\xi_2 \ldots \xi_m\,,$$

falls die auf der rechten Seite vorkommenden 1–dimensionalen Integrale existieren (als Riemannintegrale). Die Integrationsreihenfolge ist beliebig vertauschbar, genauer: Die obige Gleichung gilt für jede Integrationsreihenfolge, für die alle Integrale auf der rechten Seite existieren.

b) Auf die Einschränkung „falls die auf der rechten Seite vorkommenden Integrale existieren“ in Teil a kann verzichtet werden, wenn man die iterierten Integrale entweder alle als Unterintegrale oder als Oberintegrale versteht.

Bemerkung 8.5 Daß die „iterierten Integrale“ i. allg. nicht existieren müssen, erkennt man leicht an folgendem Beispiel. Sei

$$f : I := [0,1] \times [0,1] \to \mathbb{R}\,,$$
$$f(\xi_1, \xi_2) = \begin{cases} 1 & \text{falls } \xi_2 \neq \frac{1}{2}\,,\ \xi_1 \text{ beliebig}\,, \\ 1 & \text{falls } \xi_2 = \frac{1}{2}\,,\ \xi_1 \text{ rational}\,, \\ 0 & \text{falls } \xi_2 = \frac{1}{2}\,,\ \xi_1 \text{ irrational}\,. \end{cases}$$

Dann ist f über I integrierbar mit Integral 1 . Für $\xi_2 = \frac{1}{2}$ existiert das Integral $\int_0^1 f(\xi_1, \frac{1}{2})\, \mathrm{d}\xi_1$ nicht; dagegen gilt

$$\int_I f(x)\,\mathrm{d}x = \int_0^1 \int_0^1 f(\xi_1,\xi_2)\,\mathrm{d}\xi_2\,\mathrm{d}\xi_1 = \int_0^1 1\,\mathrm{d}\xi_1 .$$

Natrülich kann f so modifiziert werden, daß es endlich viele ξ_1 und ξ_2 gibt, für die das Integral über ξ_2 bzw. ξ_1 nicht existiert.

Beweis von Satz 8.4. a) Für charakteristische Funktionen von Intervallen ist die Behauptung offensichtlich richtig (nachrechnen); die iterierten Integrale existieren. Das gilt auch für Treppenfunktionen. Für jede Zerlegung $Z = \{Q_1, \ldots, Q_n\}$ von I definieren wir die Treppenfunktionen

$$\begin{aligned} \underline{f}_Z(x) &:= \inf\{f(y) : y \in Q_j\} \\ \overline{f}_Z(z) &:= \sup\{f(y) : y \in Q_j\} \end{aligned} \qquad \text{für} \quad x \in Q_j,\ j = 1, \ldots, n .$$

Dann gilt, falls die iterierten Integrale über f existieren (zweite Formelzeile),

$$\begin{aligned} \int_I \underline{f}_Z(x)\,\mathrm{d}x &= \int_{a_m}^{b_m} \cdots \int_{a_1}^{b_1} \underline{f}_Z(\xi_1,\ldots,\xi_m)\,\mathrm{d}\xi_1 \ldots \mathrm{d}\xi_m \\ &\le \int_{a_m}^{b_m} \cdots \int_{a_1}^{b_1} f(\xi_1,\ldots,\xi_m)\,\mathrm{d}\xi_1 \ldots \mathrm{d}\xi_m \\ &\le \int_{a_m}^{b_m} \cdots \int_{a_1}^{b_1} \overline{f}_Z(\xi_1,\ldots,\xi_m)\,\mathrm{d}\xi_1 \ldots \mathrm{d}\xi_m = \int_I \overline{f}_Z(x)\,\mathrm{d}x . \end{aligned}$$

Für eine geeignete Folge (Z_k) von Zerlegungen gilt

$$\int_I \underline{f}_{Z_k}(x)\,\mathrm{d}x \to \int f(x)\,\mathrm{d}x, \quad \int_I \overline{f}_{Z_k}(x)\,\mathrm{d}x \to \int f(x)\,\mathrm{d}x \qquad \text{für} \quad k \to \infty .$$

Also gilt

$$\int f(x)\,\mathrm{d}x = \int_{a_m}^{b_m} \cdots \int_{a_1}^{b_1} f(\xi_1,\ldots,\xi_m)\,\mathrm{d}\xi_1 \ldots \mathrm{d}\xi_m .$$

Da die Integrationsreihenfolge bei Treppenfunktionen vertauschbar ist, folgt auch die Vertauschbarkeit der Integrationsreihenfolge, soweit die iterierten Integrale existieren.

b) Dies ergibt sich sofort aus

$$\int_{a_m}^{b_m} \cdots \int_{a_1}^{b_1} \underline{f}_Z \leq \int_{*a_m}^{b_m} \cdots \int_{*a_1}^{b_1} f \leq \int_{a_m}^{*b_m} \cdots \int_{a_1}^{*b_1} f \leq \int_{a_m}^{b_m} \cdots \int_{a_1}^{b_1} \overline{f}_Z.$$

Die iterierten Unter– bzw. Oberintegrale sind also gleich dem m–dimensionalen Integral. ■

In $\mathbb{R}^{p+q} = \mathbb{R}^p \times \mathbb{R}^q$ mit $p, q \geq 1$ werden (Jordan–meßbare) Integrationsgebiete häufig wie folgt beschrieben, wobei wir $x \in \mathbb{R}^{p+q}$ in der Form (y, z) mit $y \in \mathbb{R}^p$ und $z \in \mathbb{R}^q$ schreiben:

$$A = \{x = (y, z) \in \mathbb{R}^{p+q} : y \in A',\ z \in B_y\},$$

wobei A' und B_y Mengen in $\mathbb{R}^p$ bzw. $\mathbb{R}^q$ sind.

Satz 8.6 *Sei A wie oben Jordan–meßbar und $f : A \to \mathbb{R}$ über A Riemann–integrierbar. Dann gilt*

$$\int_A f(x)\,\mathrm{d}x = \int_{A'} \int_{B_y} f(y, z)\,\mathrm{d}z\,\mathrm{d}x,$$

falls die iterierten Integrale auf der rechten Seite existieren (bzw.: wobei die Integrale auf der rechten Seite beide als Ober– oder Unterintegrale zu verstehen sind).

Beweis. Sei $A \subset I = Q_1 \times Q_2$, wobei q_1 und Q_2 Intervalle in $\mathbb{R}^p$ bzw. $\mathbb{R}^q$ sind. Dann gilt

$$\begin{aligned}
\int_A f(x)\,\mathrm{d}x &= \int_I \chi_A(x) f(x)\,\mathrm{d}x = \int_{Q_1} \int_{Q_2} \chi_A(y, z) f(y, z)\,\mathrm{d}y\,\mathrm{d}z \\
&= \int_{Q_1} \int_{Q_2} \chi_{A'}(y) \chi_{B_y}(z) f(y, z)\,\mathrm{d}y\,\mathrm{d}z \\
&= \int_{Q_1} \chi_{A'}(y) \int_{B_y} f(y, z)\,\mathrm{d}y\,\mathrm{d}z = \int_{A'} \int_{B_y} f(y, z)\,\mathrm{d}y\,\mathrm{d}z,
\end{aligned}$$

wobei für die iterierten Integrale die üblichen Anmerkungen gelten. ■

Speziell in $\mathbb{R}^2$ stellt man A häufig in der Form

$$A = \{x = (\xi_1, \xi_2) \in \mathbb{R}^2 : a_1 \le \xi_1 \le b_1,\ \phi(\xi_1) \le \xi_2 \le \psi(\xi_1)\},$$

oder als Vereinigung solcher Mengen dar (wobei eventuell die Rollen von ξ_1 und ξ_2 vertauscht sind). Dann gilt im gleichn Sinn wie im Satz formuliert

$$\int_A f(x)\,\mathrm{d}x = \int_{a_1}^{b_1} \int_{\phi(\xi_1)}^{\psi(\xi_1)} \mathrm{d}\xi_2\,\mathrm{d}\xi_1.$$

Beispiel 8.7 Sei A die rechte Hälfte des Einheitskreises, $f(x) = \xi_1$. Wir wollen $\int_A f(x)\,\mathrm{d}x$ berechnen. Offensichtlich ist

$$A = \left\{x \in \mathbb{R}^2 : 0 \le \xi_1 \le 1, -\sqrt{1-\xi_1^2} \le \xi_2 \le \sqrt{1-\xi_1^2}\right\}$$

also

$$\begin{aligned}\int\limits_A f(x)\,\mathrm{d}x &= \int\limits_0^1 \int\limits_{-\sqrt{1-\xi_1^2}}^{\sqrt{1-\xi_1^2}} \xi_1\,\mathrm{d}\xi_2\,\mathrm{d}\xi_1 = \int\limits_0^1 \xi_1 \int\limits_{-\sqrt{1-\xi_1^2}}^{\sqrt{1-\xi_1^2}} \mathrm{d}\xi_2\,\mathrm{d}\xi_1 \\ &= \int\limits_0^1 2\xi_1\sqrt{1-\xi_1^2}\,\mathrm{d}\xi_1 = -\frac{2}{3}\left(1-\xi_1^2\right)^{3/2}\Big|_0^1 = \frac{2}{3}.\end{aligned}$$

Damit ist der Schwerpunkt diese Halbkreises bei $\xi_2 = 0$ und $\xi_1 = 2/3$ dvidiert durch die Fläche (=Gesamtmasse) $\pi/2$, also $4/3\pi \approx 0.42$. □

Entsprechende Überlegungen sind in $\mathbb{R}^m$ möglich, wenn A z. B. die Form

$$\begin{aligned}A = \Big\{x \in \mathbb{R}^m \quad &: \quad (\xi_1 \ldots, \xi_{m-1}) \in A_{m-1} \subset \mathbb{R}^{m-1}, \\ &\varphi_m(\xi_1, \ldots, \xi_{m-1}) \le \xi_m \le \psi_m(\xi_1, \ldots, \xi_{m-1})\Big\}\end{aligned}$$

hat, wobei u. U. A_{m-1} entsprechend dargestellt werden kann. *Beispiel*: Wir betrachten

$$K_{m,r} \quad = \quad \left\{x \in \mathbb{R}^m : |x| \le r\right\} = \left\{x \in \mathbb{R}^m : \xi_1^2 + \ldots + \xi_{m-1}^2 + \xi_m^2 \le r^2\right\}$$

$$
\begin{aligned}
&= \Big\{x \in \mathbb{R}^m : \xi_1^2 + \ldots + \xi_{m-1}^2 \le r^2, \\
&\qquad -(r^2 - \xi_1^2 - \ldots - \xi_{m-1}^2)^{1/2} \le \xi_m \le (r^2 - \xi_1^2 - \ldots - \xi_{m-1}^2)^{1/2}\Big\} \\
&\;\vdots \\
&= \Big\{x \in \mathbb{R}^m : -r^2 \le \xi_1 \le r^2, -(r^2 - \xi_1^2)^{1/2} \le \xi_2 \le (r^2 - \xi_1^2)^{1/2}, \ldots, \\
&\qquad -(r^2 - \xi_1^2 - \ldots - \xi_{m-1}^2)^{1/2} \le \xi_m \le (r^2 - \xi_1^2 - \ldots - \xi_{m-1}^2)^{1/2}\Big\}.
\end{aligned}
$$

Damit kann also z. B. das Integral über die Einheitskugel $K_{3,r}$ in $\mathbb{R}^3$ geschrieben werden in der Form

$$
\int_{K_3} f(x)\,\mathrm{d}x = \int_{-r}^{r} \int_{-\sqrt{r^2-\xi_1^2}}^{\sqrt{r^2-\xi_1^2}} \int_{-\sqrt{r^2-\xi_1^2-\xi_2^2}}^{\sqrt{r^2-\xi_1^2-\xi_2^2}} f(\xi_1,\xi_2,\xi_3)\,\mathrm{d}\xi_3\,\mathrm{d}\xi_2\,\mathrm{d}\xi_1.
$$

8.3 Transformationsformel (Substitutionsregel)

Für die explizite Berechnung von Integralen ist häufig die folgende *Transformationsformel* (*Substitutionsregel*) nützlich:

Satz 8.8 (Transformationsformel) *Seien $A, B \subset \mathbb{R}^m$ Jordan–meßbar, $\phi : A \to B$ stetig differenzierbar und surjektiv (evtl. bis auf eine Jordan–Nullmenge), es gebe eine Jordan–Nullmenge $N \subset A$ so, daß ϕ in $A' = A \setminus N$ eine stetig differenzierbare Inverse hat. Eine Funktion $f : B \to \mathbb{R}$ ist genau dann Riemann–integrierbar, wenn $f \circ \phi : A \to \mathbb{R}$ Riemann–integrierbar, und es gilt*

$$
\int_B f(x)\,\mathrm{d}x = \int_A f(\phi(y))|\det \mathrm{D}\phi(y)|\,\mathrm{d}y.
$$

Vorbemerkung zum Beweis. Da ϕ stetig differenzierbar ist, ist $\phi(N)$ wieder eine Jordan–Nullmenge (Beweis!). Wir können also in A auf die Menge N und in B auf die Menge $\phi(N)$ verzichten, d. h. es genügt, die obige Gleichung mit A' statt A und $B' = B \setminus \phi(N)$ statt B zu beweisen; es ist dann ϕ stetig differenzierbar mit stetig differenzierbarer Inversen. — Zum Beweis benötigen wir die folgenden Hilfssätze.

Hilfssatz 8.9 *Ist L eine invertierbare reelle $m \times m$–Matrix, so gibt es orthogonale $m \times m$–Matrizen U_1 und U_2 und eine Diagonalmatrix $D = \operatorname{diag}(d_1, \ldots, d_m)$ mit $d_j > 0$ für $j = 1, \ldots, m$ so, daß gilt $L = U_1 D U_2$.*

Beweis. L^*L ist symmetrisch und invertierbar mit positiven Eigenwerten $d_1^2, \ldots, d_m^2$. Also gibt es eine orthogonale Matrix U so, daß gilt:

$$L^*L = U^{-1}D^2U, \quad UL^*LU^{-1} = D \quad \text{mit } D = \operatorname{diag}(d_1, \ldots, d_m).$$

Mit $U_1 := LU^{-1}D^{-1}$ und $U_2 := U$ gilt dann

$$U_1^*U_1 = (D^{-1}UL^*)(LU^{-1}D^{-1}) = D^{-1}\tilde{D}D^{-1} = D^{-1}D^2D^{-1} = I,$$

d. h. U_1 und U_2 sind orthogonal, und mit

$$U_1DU_2 = (LU^{-1}D^{-1})DU = L(U^{-1}D^{-1}DU) = L$$

folgt die Behauptung. [5] ■

Hilfssatz 8.10 *Für jede Jordan–meßbare Menge $M \subset \mathbb{R}^m$ gilt:*

a) *Für jedes $a \in \mathbb{R}^m$ ist auch $M+a = \{x+a : x \in M\}$ Jordan–meßbar mit $|M+a| = |M|$; das Jordan–Maß ist* translationsinvariant.

b) *Ist $D = \operatorname{diag}(d_1, \ldots, d_m)$, so ist auch $DM = \{Dx : x \in M\}$ Jordan–meßba mit $|DM| = |M| \prod_{j=1}^{m} |d_j|$.*

c) *Ist U eine orthogonale Matrix, so ist auch $UM = \{Ux : x \in M\}$ Jordan–meßbar mit $|UM| = |M|$.*

d) *Ist ϕ eine invertierbare affine Abbildung, $\phi(x) = a + Lx$, mit einer (invertierbaren) Matrix L, so ist auch $\phi(M) = \{\phi(x) : x \in M\}$ Jordan–meßbar mit $|\phi(M)| = |M||\det L| = |M||\det D\phi|$.*

[5] Alternativ zu diesem Beweis kann man $|L| := U^{-1}DU$ definieren. Es gilt dann $|L|^2 = L^*L$ und $\Big||L|x\Big| = \langle |L|x, |L|x\rangle = \langle |L|^2x, x\rangle = \langle L^*Lx, x\rangle = |Lx|^2$. Also wird durch $V : |L|x \mapsto Lx$ eine unitäre Abbildung V definiert, mit der gilt $L = V|L| = VU^{-1}DU$.

Beweis. a) Offensichtlich ist $|Q+a| = |Q|$ für jeden Quader Q. Damit folgt die Translationsinvarianz der Ober– und Unterintegrale, also auch der Integrierbarkeit und des Integrals; mit χ_M ist also auch χ_{M+a} integrierbar mit gleichem Integral, also $|M+a| = |M|$.

b) Da χ_M Riemann–integrierbar ist, gibt es zu jedem $\varepsilon > 0$ Würfel $Q_1^{(i)}, \ldots, Q_{n_i}^{(i)}$ in M und $Q_1^{(a)}, \ldots, Q_{n_a}^{(a)}$, die M überdecken mit

$$\sum_{j=1}^{n_a} \left|Q_j^{(a)}\right| - \sum_{j=1}^{n_i} \left|Q_j^{(i)}\right| < \varepsilon.$$

Da $DQ_j^{(a)}$ bzw. $DQ_j^{(i)}$ Quader sind, deren i-te Kantenlänge jeweils gegenüber der von $Q_j^{(a)}$ bzw. $Q_j^{(i)}$ mit d_j multipliziert ist, folgt die Behauptung.

c) Das orthogonale Bild jedes Quaders und jeder Kugel ist offenbar (da die Randmenge eine Jordan–Nullmenge ist) Jordan–meßbar (es handelt sich um einen verdrehten Quader bzw. wieder um eine Kugel). Sei Q_0 irgendein fester Quader (z. B. der Einheitswürfel): $|UQ_0| = \alpha|Q_0|$. Wegen Teil a und Teil b gilt dann für jeden Quader in $\mathbb{R}^m$ und somit (vgl. Beweis von Teil b) für jede Jordan–meßbare Menge M: $|UM| = \alpha|M|$. Es bleibt zu zeigen, daß $\alpha = 1$ gilt. dies erkennt man aber sofort bei Anwendung auf eine Kugel um 0, die unter U auf sich abgebildet wird.

d) Dies ergibt sich durch Zusammensetzen der Teile a, b und c. ∎

Hilfssatz 8.11 a) *Sei $A \subset \mathbb{R}^m$ eine offene Menge, $Q \subset A$ ein Quader in $\mathbb{R}^m$, und $\phi : A \to \mathbb{R}^m$ stetig differenzierbar mit stetig differenzierbarer Inverser. Dann ist $\phi(Q)$ Jordan–meßbar mit*

$$|\phi(Q)| = \int_Q |\det \mathrm{D}\phi(x)| \,\mathrm{d}x.$$

b) *Dies gilt für jede Jordanmeßbare Teilmeng C von A, d. h.:*

$$|\phi(C)| = \int_C |\det \mathrm{D}\phi(x)| \,\mathrm{d}x, \quad |C| = \int_{\phi^{-1}(C)} |\det \mathrm{D}\phi(x)| \,\mathrm{d}x.$$

Beweis. Es genügt offenbar, Teil a zu beweisen. $\phi(Q)$ ist Jordan–meßbar, denn der Rand von $\phi(Q)$ besteht aus Hyperflächenstücken.

Nehmen wir an, daß $|\phi(Q)| > \int_Q |\det \mathrm{D}\phi(x)|\,\mathrm{d}x$ gilt. Dann gibt es ein $\varepsilon > 0$ mit

$$|\phi(Q)| \geq \int_Q |\det \mathrm{D}\phi(x)|\,\mathrm{d}x + \varepsilon|Q|.$$

Durch Halbieren der Kantenlängen zerlegen wir Q in 2^m Quader. Unter diesen muß einer sein, wir bezeichnen ihn mit Q^1, für den gilt

$$|\phi(Q^1)| \geq \int_{Q_1} |\det \mathrm{D}\phi(x)|\,\mathrm{d}x + \varepsilon|Q^1|.$$

Iterieren wir dieses Verfahren, so erhalten wir eine Folge (Q^k) von ineinander geschachtelten Quadern, deren Kantenlänge das 2^{-k}–fache der Kantenlänge von Q sind, mit

$$\frac{|\phi(Q^k)|}{|Q^k|} \geq \frac{1}{|Q^k|} \int_{Q^k} |\det \mathrm{D}\phi(x)|\,\mathrm{d}x + \varepsilon$$

für $k \in \mathbb{N}$. Da die Q^k einen gemeinsamen Punkt x_0 enthalten, o. E. können wir annehmen, daß dieser gleich 0 ist, gilt mit $L := \mathrm{D}\phi(0)$

$$\liminf_{k\to\infty} \frac{|\phi(Q^k)|}{|Q^k|} \geq \lim_{k\to\infty} \frac{1}{|Q^k|} \int_{Q^k} |\det \mathrm{D}\phi(x)|\,\mathrm{d}x + \varepsilon = |\det L| + \varepsilon. \qquad (*)$$

Andererseits folgt aus der Differenzierbarkeit von ϕ

$$|\phi(x) - \phi(0) - Lx| \leq |x|\delta(|x|) \;\text{ mit } \delta(r) \to 0 \;\text{ für } r \to 0 \qquad (**)$$

(o. E. sei $\delta(\cdot)$ wachsend).

Sei Q_0 der Quader mit Zentrum 0, der durch entsprechende Verschiebung von Q entsteht, $\widetilde{Q}_0 := LQ_0$, also nach Hilfssatz 8.10 $|Q| = |Q_0| = |\widetilde{Q}_0| / \det L$. Dann gibt es eine Folge (x_k) mit $x_k \in 2^{-k}Q_0$ so, daß

$$Q^k = x_k + 2^{-k}Q_0$$

gilt. Wegen $0 \in Q^k$ gilt $|x| \le 2^{-k}d$, also $|x|\delta(x) \le 2^{-k}d\delta(2^{-k}d)$, für alle $x \in Q^k$. Mit $A_\varepsilon := \{x \in \mathbb{R}^m : d(x,A) \le \varepsilon\}$ folgt also aus (**)

$$\begin{aligned} \phi(Q^k) &= \phi(x_k + 2^{-k}Q_0) \subset \phi(0) + Lx_k + (2^{-k}\widetilde{Q}_0)_{\varepsilon_k} \\ &= \phi(0) + Lx_k + 2^{-k}(\widetilde{Q}_0)_{d\delta(2^{-k}d)}, \end{aligned}$$

also $|\phi(Q^k)| \le 2^{-km}|(\widetilde{Q}_0)_{d\delta(2^{-k}d)}|$ und, wegen $\delta(r) \to 0$ für $r \to 0$,

$$\frac{|\phi(Q^k)|}{|Q^k|} \le \frac{2^{-km}|(\widetilde{Q}_0)_{d\delta(2^{-k}d)}|}{2^{-km}|Q_0|} = \det L \frac{|(\widetilde{Q}_0)_{d\delta(82^{-k}d)}|}{|\widetilde{Q}_0|} \to \det L \quad \text{für } k \to \infty.$$

Das ist ein Widerspruch zu (*).

Entsprechend zeigt man, daß $|\phi(Q)| < \int\limits_Q |\det \mathrm{D}\,\phi(x)| dx$ nicht gelten kann. Hier wird benutzt

$$\phi(Q^k) \supset \phi(0) + Lx_k + (2^{-k}\widetilde{Q}_0)_{-\varepsilon_k}, \qquad |\phi(Q^k)| \ge 2^{-km}|(\widetilde{Q}_0)_{-d\delta(2^{-k}d)}|$$

mit $M_{-\varepsilon} := \{x \in \mathbb{R}^m : d(x, \mathbb{R}^m \setminus M) \ge \varepsilon\}$. ■

Beweis von Satz 8.8. Aus Hilfssatz 8.11 b folgt die Behauptung für Treppenfunktionen:

$$\begin{aligned} &\int_B f(s)\,\mathrm{d}x = \sum_k f_k|Q_k| = \sum_k f_k \int_{\phi^{-1}(Q_k)} |\det \mathrm{D}\phi(x)|\,\mathrm{d}x \\ &= \sum_k \int_{\phi^{-1}(Q_k)} f(\phi(x))|\det \mathrm{D}\phi(x)|\,\mathrm{d}x = \int_{\phi^{-1}(B)} f(\phi(x)|\det \mathrm{D}\phi(x)|\,\mathrm{d}x. \end{aligned}$$

Daraus ergibt sich die Behauptung für beliebige Riemann–integrierbare Funktionen, da sie sich zwischen Treppenfunktionen mit beliebig kleiner Integraldifferenz einschließen lassen. Insbesondere folgt die Integriebarkeit von $f(\phi(\cdot))|\det D\phi(\cdot)|$.

Aus der Integrierbarkeit von $f \circ \phi|\det \mathrm{D}\,\phi|$ folgt umgekehrt die Integrierbarkeit von f wegen $f(x) = f(\phi(\phi^{-1}(x)))|\det \mathrm{D}\phi(\phi^{-1}(x))||\det \mathrm{D}\phi^{-1}(x)|$. ■

8.4 Polarkoordinaten

Wir wenden dieses Resultat nun auf Polarkoordinaten in $\mathbb{R}^2$ und$\mathbb{R}^3$ an:

Beispiel 8.12 *Polarkoordinaten* in $\mathbb{R}^2$: Die Abbildung

$$\begin{aligned} &\phi : [0,\infty) \times [0,2\pi) \to \mathbb{R}^2\,, \\ &\phi(r,\varphi) \;=\; (r\cos\varphi,\, r\sin\varphi) \end{aligned}$$

ist stetig differenzierbar mit

$$\begin{aligned} \mathrm{D}\phi(r,\varphi) \;&=\; \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}\,, \\ \det \mathrm{D}\phi(r,\varphi) \;&= r\cos^2\varphi + r\sin^2\varphi = r\,. \end{aligned}$$

ϕ ist offenbar surjektiv und mit der Jordan–Nullmenge

$$N = \Big\{\,(0,\varphi) : 0 \le \varphi < 2\pi\Big\}$$

ist ϕ auf $(0,\infty) \times [0,2\pi) \;=\; [0,\infty) \times [0,2\pi) \setminus N$ injektiv.

Zum Beispiel bildet ϕ die Menge

$$A \;=\; \Big\{(r,\varphi) :\; 0 \le r \le R\,,\; 0 \le \varphi < 2\pi\Big\}$$

auf die Kreisscheibe mit Radius R

$$B = K_{2,R} = \Big\{x \in \mathbb{R}^2 : |x| \le R\Big\}$$

ab. Also ist

$$\begin{aligned} \int\limits_{K_{2,R}} f(x)\,\mathrm{d}x \;&=\; \int\limits_{A} f(r\cos\varphi,\, r\sin\varphi)\, r\, \mathrm{d}(r,\varphi) \\ &=\; \int\limits_0^R r \int\limits_0^{2\pi} f(r\cos\varphi,\, r\sin\varphi)\,\mathrm{d}\varphi\;\mathrm{d}r\,. \end{aligned}$$

Ist f nur von r abhängig, $f(x) = \tilde{f}(|x|)$ (sphärisch symmetrisch), so folgt

$$= \int_0^R r 2\pi \tilde{f}(r)\,\mathrm{d}r = 2\pi \int_0^R r\tilde{f}(r)\,\mathrm{d}r\,.$$

$$\int_{K_{2,R}} 1\,\mathrm{d}x = 2\pi \int_0^R r\,\mathrm{d}r \;=\; R^2\pi\,.$$

Damit ergibt sich für die Fläche der Kreisscheibe mit Radius r □

Beispiel 8.13 *Polarkoordinaten* in $\mathbb{R}^3$: Wir betrachten die Abbildung

$$\begin{aligned}
&\phi :\; [0,\infty)\times[0,2\pi)\times[0,\pi] \to \mathbb{R}^3\,,\\
&\phi(r,\varphi,\vartheta) \;=\; (r\cos\varphi\sin\vartheta,\; r\sin\varphi\sin\vartheta, r\cos\vartheta)\,.
\end{aligned}$$

Offenbar ist ϕ stetig differenzierbar mit

$$\mathrm{D}\phi(r,\varphi,\vartheta) \;=\; \begin{pmatrix} \cos\varphi\sin\vartheta & -r\sin\varphi\sin\vartheta & r\cos\varphi\cos\vartheta \\ \sin\varphi\sin\vartheta & r\cos\varphi\sin\vartheta & r\sin\varphi\cos\vartheta \\ \cos\vartheta & 0 & -r\sin\vartheta \end{pmatrix}\,,$$

und durch Entwicklung nach der letzten Zeile

$$\begin{aligned}
\det \mathrm{D}\phi(r,\varphi,\vartheta) \;&=\; \cos\vartheta(-r^2\sin^2\varphi\sin\vartheta\cos\vartheta - r^2\cos^2\varphi\sin\vartheta\cos\vartheta)\\
&\qquad - r\sin\vartheta(r\cos^2\varphi\sin^2\vartheta + r\sin^2\varphi\sin^2\vartheta)\\
&=\; -r^2\sin\vartheta\cos^2\vartheta - r^2\sin\vartheta\sin^2\vartheta \;=\; -r^2\sin\vartheta\,,\\
|\det \mathrm{d}\phi(r,\varphi,\vartheta)| \;&=\; r^2\sin\vartheta\,.
\end{aligned}$$

ϕ ist offenbar surjektiv, und mit der Jordan–Nullmenge

$$\begin{aligned}
N \;=\;& \Big\{(0,\varphi,\vartheta) :\; 0\le\varphi<2\pi,\; 0\le\vartheta\le\pi\Big\}\\
&\cup\Big\{(r,\varphi,0) :\; 0<r<\infty,\; 0\le\varphi<2\pi\Big\}\\
&\cup\Big\{(r,\varphi,\pi) :\; 0<r<\infty,\; 0\le\varphi<2\pi\Big\}
\end{aligned}$$

ist ϕ auf $[0,\infty)\times[0,2\pi)\times[0,\pi]\setminus N$ injektiv. □

Zum Beispiel bildet ϕ die Menge

$$A = \Big\{(r, \varphi, \vartheta) :\ 0 \le r \le R,\ 0 \le \varphi < 2\pi\,, 0 \le \vartheta \le \pi\Big\}$$

auf die Kugel

$$B \;=\; K_{3,R} \;=\; \Big\{\, x \in \mathbb{R}^3 : |x| \le R \,\Big\}$$

ab. Es ist also

$$\begin{aligned}
&\int\limits_{K_{3,R}} f(x)\,\mathrm{d}x \\
&\quad= \int\limits_{A} f(r\cos\varphi\sin\vartheta, r\sin\varphi\sin\vartheta, r\cos\vartheta) r^2 \sin\vartheta\,\mathrm{d}(r,\varphi,\vartheta) \\
&\quad= \int\limits_0^R r^2 \int\limits_0^{2\pi}\int\limits_0^{\pi} f(r\cos\varphi\sin\vartheta, r\sin\varphi\sin\vartheta, r\cos\vartheta)\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi\,\mathrm{d}r\,.
\end{aligned}$$

Ist f nur von r abhängig, $f(x) = \tilde{f}(|x|)$ (sphärisch symmetrisch), so folgt

$$= \int\limits_0^R r^2 2\pi \int\limits_0^{\pi} \tilde{f}(r)\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}r = 4\pi\int\limits_0^R r^2\tilde{f}(r)\,\mathrm{d}r$$

$$\int\limits_{K_{3,R}} 1\,\mathrm{d}x \;=\; 4\pi\int\limits_0^R r^2\,\mathrm{d}r \;=\; \frac{4}{3}\pi R^3$$

und somit für das Volumen der Kugel mit Radius R. #

Beispiel 8.14 Wir wollen das Integral über einen Kegel $K_{r,h}$ der Höhe h über einem Kreis mit Radius r berechnen:

$$\begin{aligned}
K_{r,h} \;&=\; \Big\{x \in \mathbb{R}^3\,:\, 0 \le \xi_3 \le h,\, \sqrt{\xi_1^2+\xi_2^2} \le r - \xi_3\frac{r}{h}\,\Big\} \\
&=\; \Big\{x \in \mathbb{R}^3\,:\, 0 \le \xi_3 \le h,\, 0 \le |\xi_2| \le r\Big(1 - \frac{\xi_3}{h}\Big)\,, \\
&\qquad\qquad 0 \le |\xi_1| \le \sqrt{(1-\xi_3/h)^2 r^2 - \xi_2^2} =: \varrho(\xi_2,\xi_3)\,\Big\}\,.
\end{aligned}$$

Es ist also

$$\begin{aligned}\int_{K_{r,h}} f(x)\,\mathrm{d}x &= \int_0^h \int_{-r(1-\xi_3/h)}^{r(1-\xi_3/h)} \int_{-\varrho(\xi_2,\xi_3)}^{\varrho(\xi_2,\xi_3)} f(\xi_1,\xi_2,\xi_3)\,\mathrm{d}\xi_1\,\mathrm{d}\xi_2\,\mathrm{d}\xi_3 \\ &= \int_0^h \int_{\{|(\xi_1,\xi_2)|\le r(1-\xi_3/h)\}} f(\xi_1,\xi_2,\xi_3)\,\mathrm{d}(\xi_1,\xi_2)\,\mathrm{d}\xi_3\,.\end{aligned}$$

Ist f nur von $\varrho = |(\xi_1,\xi_2)|$ abhängig, $f(x) = \tilde{f}(|(\xi_1,\xi_2)|)$, so folgt (vgl. Beispiel 8.12)

$$= \int_0^h \int_{K_{2,r(1-\xi_3/h)}} \tilde{f}(|(\xi_1,\xi_2)|)\,\mathrm{d}(\xi_1,\xi_2)\,\mathrm{d}\xi_3 = \int_0^h 2\pi \int_0^{r(1-\xi_3/h)} \tilde{f}(\varrho)\,\mathrm{d}\varrho\,\mathrm{d}\xi_3\,.$$

Ist f nur von ξ_3 abhängig, $f(x) = \hat{f}(\xi_3)$, so folgt

$$= \int_0^h \hat{f}(\xi_3) \int_{K_{2,r(1-\xi_3/h)}} \mathrm{d}(\xi_1,\xi_2)\,\mathrm{d}\xi_3 = \int_0^h \hat{f}(\xi_3)\pi r^2\Big(1-\frac{\xi_3}{h}\Big)^2\,\mathrm{d}\xi_3\,.$$

Speziell erhalten wir für das Volumen des Kegels $K_{r,h}$

$$\int_{k_{r,h}} 1\,\mathrm{d}x = \int_0^h \pi r^2\Big(1-\frac{\xi_3}{h}\Big)^2\,\mathrm{d}\xi_3 \qquad = -\frac{1}{3}\pi r^2 h\Big(1-\frac{\xi_3}{h}\Big)^3\Big|_0^h = \frac{1}{3}\pi r^2 h\,.$$

Wie bei den vorhergehenden Beispielen können wir eine Parameterdarstellung des Kegels verwenden:

$$\begin{aligned}&\phi : [0,2\varpi)\times[0,r]\times[0,h]\to\mathbb{R}^3,\\ &\phi\Big(\varphi,\varrho,\xi\Big) = (\varrho(1-\frac{\xi}{h})\cos\varphi, \varrho(1-\frac{\xi}{h})\sin\varphi, \xi\Big).\end{aligned}$$

Es gilt

$$D\phi(\varphi,\varrho\xi) = \begin{pmatrix} -\varrho(1-\frac{\xi}{h})\sin\varphi & (1-\frac{\xi}{h})\cos\varphi & 0 \\ \varrho(1-\frac{\xi}{h})\cos\varphi & (1-\frac{\xi}{h}\sin\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\begin{aligned}\det D\phi(\varphi,\varrho\xi) &= -\varrho(1-\frac{\xi}{h})^2\sin^2\varphi - \varrho(1-\frac{\xi}{h})^2\cos^2\varphi = \varrho(1-\frac{\xi}{h})^2,\\ \int_{K_{2,r}} f(x)\,\mathrm{d}x &= \int_0^{2\pi}\int_0^r\int_0^h f\Big(\varrho(1-\frac{\xi}{h})\cos\varphi, \varrho(1-\frac{\xi}{h})\sin\varphi, \xi\Big)\varrho(1-\frac{\xi}{h})^2\,\mathrm{d}(\varphi,\varrho,\xi).\end{aligned}$$

Ist speziell f nur von ξ abhängig, $f(x) = \hat{f}(\xi)$, so gilt

$$\begin{aligned} &= \int_0^h \hat{f}(\xi) \int_0^{2\pi} \int_0^r \varrho(1-\frac{\xi}{h})\,\mathrm{d}\varrho\,\mathrm{d}\varphi\,\mathrm{d}\xi \\ &= \pi r^2 \int_0^h \hat{f}(\xi)(1-\frac{\xi}{h})^2\,\mathrm{d}\xi \\ \text{Volumen} &= -\frac{1}{3}\pi r^2 h(1-\frac{\xi}{h})^3\Big|_0^h = \frac{1}{3}\pi r^2 h, \end{aligned}$$

wie wir dies bereits oben erhalten hatten. Die Berechnung des Integrals über eine Funktion, die nur von $\varrho = |(\xi_1, \xi_2)|$ abhängt, ist in dieser Darstellung nicht so einfach. □

8.5 Uneigentliche Integrale

Analog zu unseren Untersuchungen in $\mathbb{R}^1$ definieren wir uneigentliche Integrale in $\mathbb{R}^m$: Das Integral $\int_A f(x)\,\mathrm{d}x$ heißt *uneigentlich*, wenn der Integrationsbereich und/oder die Funktion f unbeschränkt sind/ist. Das uneigentliche Integral wird als Limes von Integralen über beschränkte Bereiche A_n, auf denen f Riemann–integrierbar (also insbesondere beschränkt) ist, erklärt; dabei muß natürlich der gesamte Integrationsbereich ausgeschöpft werden:

$$A = \bigcup_n A_n \qquad \text{(evtl. bis auf Jordan–Nullmengen).}$$

Die Folge (A_n) nennen wir eine Ausschöpfung der Menge A.

Wir sagen, das *uneigentliche Integral* $\int_A f(x)\,\mathrm{d}x$ existiert, wenn $\lim_{n\to\infty} \int_{A_n} f(x)\,\mathrm{d}x$ existiert (falls f das Vorzeichen wechselt, kann die Existenz von der Wahl der Ausschöpfung abhängen; bei positiven (bzw. negativen) Funktionen ist dies nicht der Fall (Beweis!), und nur diesen Fall werden wir im folgenden betrachten).

Beispiel 8.15 In $\mathbb{R}^m$ $(m > 1)$ untersuchen wir das (für $\alpha > 0$ bei $x = 0$ uneigentliche) Integral

$$\int\limits_{K_R} \frac{1}{|x|^\alpha}\,\mathrm{d}x \qquad (\alpha \in \mathbb{R})\,.$$

Für A_n betrachten wir $K_R \setminus K_{1/n}$ oder allgemein $A_\varepsilon = K_R \setminus K_\varepsilon$ für $\varepsilon \to 0$, d. h. wir haben zu untersuchen

$$\lim_{\varepsilon\to 0} \int_{K_R\setminus K_\varepsilon} \frac{1}{|x|^\alpha}\,\mathrm{d}x = \lim_{\varepsilon\to 0} F_m \int_\varepsilon^R r^{m-1-\alpha}\,\mathrm{d}r$$

$$= \lim_{\varepsilon\to 0} F_m \begin{cases} \frac{r^{m-\alpha}}{m-\alpha}\Big|_\varepsilon^R & \text{für } m-\alpha \neq 0\,, \\ \ln r\Big|_\varepsilon^R & \text{für } m-\alpha = 0\,, \end{cases}$$

wobei F_m die Oberfläche der Sphäre mit Radius 1 in $\mathbb{R}^m$ ist. Der Limes existiert genau dann, wenn $m - \alpha > 0$ ist; dann gilt

$$\int_{K_R} \frac{1}{|x|^\alpha}\,\mathrm{d}x = \frac{F_m}{m-\alpha} R^{m-\alpha} \qquad \text{für} \quad \alpha < m\,.$$

Das uneigentliche Integral existiert genau dann, wenn $\alpha < m$ ist. □

Beispiel 8.16 In $\mathbb{R}^m$ $(m > 1)$ untersuchen wir (bei ∞) das uneigentliche Integral

$$\int\limits_{\mathbb{R}^m\setminus K_R} \frac{1}{|x|^\alpha}\,\mathrm{d}x \qquad (\alpha \in \mathbb{R})\,.$$

Für A_n wählen wir $K_n \setminus K_R$ oder allgemeiner $A_s = K_s \setminus K_R$ für $s \to \infty$, d. h. wir haben zu untersuchen

$$\lim_{s\to\infty} \int_{K_s\setminus K_R} \frac{1}{|x|^\alpha}\,\mathrm{d}x = \lim_{s\to\infty} F_m \int_R^s r^{m-1-\alpha}\,\mathrm{d}s$$

$$= \lim_{s\to\infty} F_m \begin{cases} \frac{r^{m-\alpha}}{m-\alpha}\Big|_R^s & \text{für } m-\alpha \neq 0\,, \\ \ln r\Big|_R^s & \text{für } m-\alpha = 0\,. \end{cases}$$

Dieser Limes existiert genau dann, wenn $m - \alpha < 0$ ist; dann gilt

$$\int_{\mathbb{R}^m\setminus K_R} \frac{1}{|x|^\alpha}\,\mathrm{d}x = \frac{F_m}{\alpha-m} R^{m-\alpha} \qquad \text{für} \quad \alpha > m\,.$$

Das uneigentliche Integral existiert genau dann, wen $\alpha > 0$ ist. □

8.6 Parameterabhängige Integrale und Zerlegung der Eins

Wir diskutieren die Frage der Differenzierbarkeit bzw. der Differentiation parameterabhängiger Integrale nach dem Parameter. Der Einfachheit halber werden wir nur Integrale bezüglich einer Variablen betrachten.

Satz 8.17 (Differentiation parameterabhängiger Integrale) *Sei $f : \mathbb{R}^2 \to \mathbb{R}$ stetig und besitze eine stetige partielle Ableitung nach der zweiten Variablen, $g, h : \mathbb{R} \to \mathbb{R}$ seien stetig differenzierbar. Dann gilt:*

(a) $$\frac{\mathrm{d}}{\mathrm{d}y} \int\limits_a^b f(x,y)\,\mathrm{d}x = \int\limits_a^b \mathrm{D}_2 f(x,y)\,\mathrm{d}x\,,$$

(b) $$\frac{\mathrm{d}}{\mathrm{d}y} \int\limits_a^{h(y)} f(x,y)\,\mathrm{d}x = h'(y) f(h(y),y) + \int\limits_a^{h(y)} \mathrm{D}_2 f(x,y)\,\mathrm{d}x\,,$$

(c) $$\frac{\mathrm{d}}{\mathrm{d}y} \int\limits_{g(y)}^{b} f(x,y)\,\mathrm{d}x = -g'(y) f(g(y),y) + \int\limits_{g(y)}^{b} \mathrm{D}_2 f(x,y)\,\mathrm{d}x\,,$$

(d) $$\frac{\mathrm{d}}{\mathrm{d}y} \int\limits_{g(y)}^{h(y)} f(x,y)\,\mathrm{d}x = h'(y) f(h(y),y) - g'(y) f(g(y),y) + \int\limits_{g(y)}^{h(y)} \mathrm{D}_2 f(x,y)\,\mathrm{d}x\,.$$

Beweis. Zunächst gilt offensichtlich nach dem Hauptsatz der Differential–und Integralrechnung

$$\frac{\mathrm{d}}{\mathrm{d}z} \int\limits_a^z f(x,y)\,\mathrm{d}x = f(z,y) \tag{i}$$

und, durch explizite Berechnung der Differentialquotienten

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}y} \int\limits_a^z f(x,y)\,\mathrm{d}x &= \lim_{h\to 0} \frac{1}{h} \int\limits_a^z (f(x,y+h) - f(x,y))\,\mathrm{d}x \\
&= \lim_{h\to 0} \frac{1}{h} \int\limits_a^z \int\limits_y^{y+h} \mathrm{D}_2 f(x,t)\,\mathrm{d}t\,\mathrm{d}x
\end{aligned}$$

$$= \lim_{h\to 0} \frac{1}{h} \int_y^{y+h} \int_a^z \mathrm{D}_2 f(x,t)\,\mathrm{d}x\,\mathrm{d}t = \int_a^z \mathrm{D}_2 f(x,y)\,\mathrm{d}x\,.$$

(a) Dies folgt aus (ii) mit $z = b$.

(b) Mit $k(y) := (h(y), y)$, $\ell(z,w) := \int_a^z f(x,w)\,\mathrm{d}x$ folgt

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}y} \int_a^{h(y)} f(x,y)\,\mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}y}\ell(h(y),y) = \frac{\mathrm{d}}{\mathrm{d}y}\ell(k(y)) \\
&= \mathrm{D}\ell(k(y))\,\mathrm{D}k(y) \\
&= \Big\langle \Big(\mathrm{D}_1\ell(h(y),y), \mathrm{D}_2\ell(h(y),y)\Big), \Big(k_1'(y), k_2'(y)\Big)\Big\rangle \\
&\quad \text{(mit (i) für } \mathrm{D}_1\ell \text{ und (ii) für } \mathrm{D}_2\ell\text{)} \\
&= \Big\langle \Big(f(h(y),y), \int_a^{h(y)} \mathrm{D}_2 f(x,y)\,\mathrm{d}x\Big), \Big(h'(y), 1\Big)\Big\rangle \\
&= h'(y) f(h(y),y) + \int_a^{h(y)} \mathrm{D}_2 f(x,y)\,\mathrm{d}x\,.
\end{aligned}$$

(c) Wird analog zu (b) bewiesen.

(d) Mit der Zerlegung

$$\int_{g(y)}^{h(y)} f(x,y)\,\mathrm{d}x = \int_{g(y)}^{c} f(x,y)\,\mathrm{d}x + \int_c^{h(y)} f(x,y)\,\mathrm{d}x$$

ergibt sich die Aussage (d) aus (b) und (c). ■

Satz 8.18 (Zerlegung der Eins) *Sei $K \subset \mathbb{R}^m$ kompakt, $\{O_1, \ldots, O_n\}$ eine Überdeckung von K aus (endlich vielen) Jordan–meßbaren offenen Mengen. Dann gibt es beliebig oft stetig differenzierbare Funktionen $\varphi_1, \ldots, \varphi_n (\in C_0^\infty(\mathbb{R}^m))$ mit*

$$\sum_{j=1}^n \varphi_j(x) = 1 \qquad \textit{für} \quad x \in K\,, \qquad \varphi_j(x) = 0 \quad \textit{für} \quad x \notin O_j\,.$$

(Das entsprechende gilt für eine unendliche, jedoch lokal endliche, Überdeckung.)

Beweis. Für jedes $\delta > 0$ und $j = 1, \ldots, n$ sei

$$O_j^\delta := \Big\{ x \in O_j : |x-y| > \delta \quad \text{für alle} \quad y \in \mathbb{R}^m \backslash O_j \Big\}$$

(das ist die Menge der Punkte von O_j, die vom Rand von O_j einen Abstand $> \delta$ haben). Offenbar ist O_j^δ offen.

Es gibt ein $\varepsilon > 0$ so, daß $K \subset O^\varepsilon := \bigcup\limits_{j=1}^{n} O_j^{2\varepsilon}$ gilt. Dies folgt leicht daraus, daß $\{O^\varepsilon : \varepsilon > 0\}$ eine offene Überdeckung von K ist; es gibt also eine endliche Teilüberdeckung $\{O^{\varepsilon_1}, \ldots, O^{\varepsilon_k}\}$; mit $\varepsilon := \min\{\varepsilon_j : j = 1, \ldots, k\}$ folgt die Behauptung.

Sei nun für jedes $\varepsilon > 0$

$$\varphi_\varepsilon(x) := \begin{cases} c_\varepsilon \exp\left(-\dfrac{1}{\varepsilon^2 - x^2}\right) & \text{für } |x| < \varepsilon\,, \\ 0 & \text{für } |x| \geq \varepsilon \end{cases}$$

und c_ε so, daß $\int \varphi_\varepsilon(x)\,\mathrm{d}x = 1$ gilt. Für

$$\begin{aligned} \chi_j &= \text{charakteristische Funktion von} \quad O_j^\varepsilon \ (j = 1, \ldots, n)\,, \\ \chi_0 &= \text{charakteristische Funktion von} \quad \mathbb{R}^m \backslash \bigcup_{j=1}^{n} O_j^\varepsilon \end{aligned}$$

definieren wir

$$\tilde{\varphi}_j(x) := \int \varphi_\varepsilon(x-y)\, \chi_j(y)\,\mathrm{d}y \qquad (j = 0, 1, \ldots, n)\,.$$

Dann gilt (vgl. Satz 8.17 a)

$$\tilde{\varphi}_j \in C_0^\infty(\mathbb{R}^m) \quad \text{mit} \quad \tilde{\varphi}_j(x) = \begin{cases} 1 & \text{in } O_j^{2\varepsilon}\,, \\ 0 & \text{außerhalb } O_j \end{cases} \qquad \text{für} \quad j = 1, \ldots, n\,,$$

$$\tilde{\varphi}_0 \in C^\infty(\mathbb{R}^m) \quad \text{mit} \quad \tilde{\varphi}_0(x) = \begin{cases} 1 & \text{in } \mathbb{R}^m \backslash \bigcup\limits_{j=1}^{n} O_j\,, \\ 0 & \text{in } O^\varepsilon = \bigcup_{j=1}^{n} O_j^{2\varepsilon}\,. \end{cases}$$

Also ist $\sum_{k=0}^{n} \widetilde{\varphi}(x) \neq 0$ für alle $x \in \mathbb{R}^m$. Definieren wir

$$\varphi_j(x) := \left\{ \sum_{k=0}^{n} \tilde{\varphi}_k(x) \right\}^{-1} \tilde{\varphi}_j(x) \qquad \text{für} \quad j = 1, \ldots, n\,,$$

so folgt die Behauptung des Satzes. ■

8.7 Übungsaufgaben

8.1 Sei $A \subset \mathbb{R}^m$ beschränkt, A' die Menge der Häufungspunkte von A, A'' die Menge der Häufungspunkte von A'. Ist A'' endlich, so ist A eine Jordan–Nullmenge.

8.2 In Fortsetzung von Aufgabe 8.1 sei $A^{(j+1)}$ die Menge der Häufungspunkte von $A^{(j)}$ für $j \geq 2$. Ist $A^{(k)} = \emptyset$ für ein $k \in \mathbb{N}$, so ist A eine Jordan–Nullmenge.

8.3 Sei P ein reelles Polynom vom Grad $r > 0$ in m Variablen. Für jede kompakte Teilmenge $K \subset \mathbb{R}^m$ und jedes $\lambda \in \mathbb{R}$ ist

$$M := \{x \in K : P(x) = \lambda\}$$

eine Jordan–Nullmenge.

Anleitung: O. E. $\lambda = 0$, Induktion nach r, Induktionsverankerung $r = 1$. Beim Induktionsschritt ist (z. B.) zu zeigen:

a) $N_0 = \{x \in K : P(x) = 0, \operatorname{grad} P(x) = 0\}$ ist eine Jordan–Nullmenge.

b) Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$ so, daß

$$N_\varepsilon^- := \{x \in K : P(x) = 0,\ |\operatorname{grad} P(x)| \leq \delta\}$$

in der Vereinigung K_ε endlich vieler Intervalle mit Gesamtvolumen $\leq \varepsilon$ enthalten ist.

c) $N_\varepsilon^+ := M \cap (K \setminus K_\varepsilon)$ ist eine Vereinigung endlich vieler Hyperflächen.

8.4 Man gebe eine Riemann–integrierbare Funktion an, deren Unstetigkeitspunkte keine Jordan–Nullmenge bilden.

8.5 Der Schwerpunkt eines homogenen Körpers B in $\mathbb{R}^3$ ist gegeben durch

$$\frac{1}{V_B}\Big(\int\limits_B x\,d(x,y,z),\ \int\limits_B y\,d(x,y,z),\ \int\limits_B z\,d(x,y,z)\Big),$$

wobei V_B das Volumen des Körpers B ist. Man berechne den Schwerpunkt der oberen Hälfte der Einheitskugel.

8.6 Man bestimme den Schwerpunkt des Kegels über dem Halbkreis $\{(\xi_1,\xi_2) : \xi_2 \geq 0, \xi_1^2 + \xi_2^2 \leq r\}$ mit der Spitze in $(0,0,h)$.

8.7 Man berechne das Volumen und den Schwerpunkt der oberen Hälfte des Ellipsoids

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \qquad (a,b,c > 0)$$

mit Hilfe der Transformation

$$x = ar\sin\vartheta\cos\varphi, \quad y = br\sin\vartheta\sin\varphi, \quad z = cr\cos\vartheta\,.$$

8.8 Man berechne

(a) $\displaystyle A = \int\limits_{\mathbb{R}^2} \exp(-|x|^2)\,\mathrm{d}x\,,$ (b) $\displaystyle B = \int\limits_{-\infty}^{\infty} \exp(-t^2)\,\mathrm{d}t\,.$

Anleitung: Man zeige $A = B^2$.

8.9 Man berechne

$$\int\limits_0^\pi \left\{\int\limits_x^\pi \frac{\sin y}{y}\,\mathrm{d}y\right\}\mathrm{d}x\,.$$

8.10 1. Man skizziere die Archimedische Spirale

$$r(\varphi) = \varphi \qquad \text{für} \quad 0 \leq \varphi \leq 2\pi$$

(und mache insbesondere das Verhalten der Kurve im Nullpunkt deutlich).

2. Man berechne die Fläche, die durch die Archimedische Spirale und das Intervall $[0, 2\pi]$ auf der x–Achse begrenzt wird. *Anleitung*: Man kann das Flächenstück als Bild des Dreiecks $((0,0), (2\pi, 0), (2\pi, 2\pi)$ der (φ, r)–Ebene darstellen.

8.11 Für welche $\alpha, \beta \in \mathbb{R}$, $\beta \le 1$, existiert das Integral

$$\int_A |x|^{-\alpha} \, \mathrm{d}x$$

mit $A = \left\{ x \in \mathbb{R}^m : \xi_1 \ge 1, \xi_2^2 + \ldots + \xi_m^2 \le \xi_1^{2\beta} \right\}$?

8.12 Ist $f : \mathbb{R}^m \to \mathbb{R}$ über jede kompakte Teilmenge von $\mathbb{R}^m \setminus \{0\}$ Riemann–integrierbar (z. B. stetig in $\mathbb{R}^m \setminus \{0\}$), und gilt

a) $|f(x)| \le |x|^{-\alpha}$ mit $\alpha < m$, so existiert das uneigentliche Integral $\int_{K_R} f(x) \, \mathrm{d}x$ und ist unabhängig von der Ausschöpfung.

b) $|f| \le |x|^{-\alpha}$ mit $\alpha > m$, so existiert das uneigentliche Integral $\int_{\mathbb{R}^m \setminus K_R} f(x) \, \mathrm{d}x$ und ist unabhängig von der Ausschöpfung.

8.13 Man berechne den Schwerpunkt des Kegels über dem Halbkreis $\{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_2 > 0, \; \xi_1^2 + \xi_2^2 \le r\}$ mit Spitze in $(0, 0, h)$.

8.14 Für $m \ge 3$ sei

$$\phi_m : [0, \infty) \times [0, 2\pi) \times \underbrace{[0, \pi] \times \ldots \times [0, \pi]}_{m-2 \text{ mal}} \to \mathbb{R}$$

$$\begin{aligned} \phi_m(r, \varphi, \vartheta_1, \ldots, \vartheta_{m-2}) = (&r \cos\varphi \sin\vartheta_1 \ldots \sin\vartheta_{m-2}, r \sin\varphi \sin\vartheta_1 \ldots \sin\vartheta_{m-2} \\ &r \cos\vartheta_1 \sin\vartheta_2 \ldots \sin\vartheta_{m-2}, \ldots, r\cos\vartheta_{m-3} \sin\vartheta_{m-2}, r\cos\vartheta_{m-2}) \end{aligned}$$

(m–dimensionale *Polarkoordinaten*).

a) ϕ_m ist surjektiv; die Einschränkung auf $(0, \infty) \times [0, 2\pi) \times (0, \pi) \times \ldots \times (0, \pi)$ ist injektiv.

b) Für $m = 4$ berechne man die Funktionaldeterminante $\det D\phi_4(r, \varphi, \vartheta_1, \vartheta_2)$.

c) Mit dieser Parameterdarstellung berechne man das Volumen der Kugel mit Radius R in $\mathbb{R}^4$.

8.15 Das *Trägheitsmoment* eines Körpers K mit Dichte $\varrho : K \to \mathbb{R}$ in bezug auf eine Drehachse g ist $T_g := \int_K \varrho(x)\, \mathrm{d}(x,g)\, \mathrm{d}x$, wobei $\mathrm{d}(x,g)$ der Abstand des Punktes x von der Achse g ist.

Man berechne das Trägheitsmoment einer homogenen Kugel mit Dichte ϱ und Radius r um eine Achse durch den Mittelpunkt.

8.16 Zwei gerade Kreiszylinder mit gleichem Radius R liegen so, daß ihre Achsen sich senkrecht schneiden. Man bestimme das Volumen der innerhalb beider Zylinder liegenden Menge.

8.17 Sei A eine Jordan–meßbare Menge in $\mathbb{R}^{p+q}$ (besonders interessant ist der Fall $p = 1$) mit $A \subset \{(y,z) \in \mathbb{R}^p \times \mathbb{R}^q : y \in Q\}$ für ein geeignetes Intervall $Q \subset \mathbb{R}^p$. Die Schnitte $A_y := \{z \in \mathbb{R}^q : (y,z) \in A\}$ seien Jordan–meßbar in $\mathbb{R}^q$. Dann ist die Funktion $Q \to \mathbb{R}$, $y = |A_y|$ integrierbar, und es gilt *Cavalierisches Prinzip*

$$|A| = \int_Q |A_y|\, \mathrm{d}y.$$

Anleitung: Vgl. Satz 8.6

9 Oberflächenintegrale; der allgemeine Satz von Stokes

Im folgenden definieren wir Integrale über Hyperflächen (insbesondere Kurven in $\mathbb{R}^2$ und Flächen in $\mathbb{R}^3$) und beweisen den grundlegenden Satz über den Zusammenhang zwischen Volumen– und Oberflächenintegral (in $\mathbb{R}^1$ ist dies der Hauptsatz der Differential– und Integralrechnung, der eine Beziehung zwischen dem Integral der Funktion f über ein Intervall und den Randwerten einer Stammfunktion F von f herstellt, bzw. zwischen den Randwerten einer Funktion f und dem Integral der Ableitung f' über das Intervall).

9.1 Oberflächenintegrale

Wir nehmen im folgenden zunächst an, daß eine *Hyperfläche* M beschrieben wird durch:

$$M = \Big\{x = (y, z) : y \in A, z = h(y)\Big\} \quad (A \subset \mathbb{R}^{m-1})$$

mit einer stetig differenzierbaren Funktion

$$h : \mathbb{R}^{m-1} \supset A \to \mathbb{R},$$

wobei es natürlich keine Rolle spielt, daß hier $z(\cdot)$ für die letzte Komponente steht.

Das Integral einer (zunächst stetigen) Funktion $f : M \to \mathbb{R}$ über die Fläche M denken wir uns, im Sinne von Riemannsummen, definiert als Grenzwert von Summen über Terme der Form

(Fläche eines kleinen Flächenstücks in M)$\times$
(Funktionswert von f in einem Punkt dieses Flächenstücks).

Es dürfte bequemer sein, dieses „Integral“als ein Integral über die „ebene“Fläche A, die Projektion von M auf $\mathbb{R}^{m-1}$, zu definieren; dabei sind natürlich die entsprechenden Flächenstücke mit einem Faktor, der die Neigung des Flächenstücks auf M gegenüber A berücksichtigt, zu multiplizieren: dieser Faktor ist nach Pythagoras $(1 + |\operatorname{grad} h(y)|^2)^{1/2}$. Damit kommen wir zu der Definition

$$\int_M f(x)\, \mathrm{d}o(x) := \int_A f(y, h(y))\Big(1 + |\operatorname{grad} h(y)|^2\Big)^{1/2} \mathrm{d}y;$$

dabei soll „ $\mathrm{d}o(x)$“ andeuten, daß es sich um ein Oberflächenintegral handelt.

Läßt sich M nicht als ganzes als Graph darstellen, so gehen wir zumindest davon aus, daß man M in kleinere Teile zerlegen kann, die sich als Graphen darstellen lassen, und definieren das Integral stückweise. (Nach dem Satz über implizite Funktionen (Satz 6.6), in Verbindung mit dem in 26.2 bewiesenen Satz von Heine–Borel, ist dies z. B. dann der Fall, wenn M mit Hilfe einer stetig differenzierbaren Funktion $g : \mathbb{R}^m \to \mathbb{R}$ beschrieben wird durch

$$M := \Big\{x \in \mathbb{R}^m : g(x) = 0\Big\} \quad \text{mit} \quad \operatorname{grad} g(x) \neq 0 \text{ für } x \in M.$$

Wie kann man nun das Oberflächenintegral berechnen, wenn M in einer beliebigen Parameterdarstellung gegeben ist?

Dazu betrachten wir zunächst die durch die Darstellung als Graph erzeugte spezielle Parameterdarstellung von M

$$\phi : \mathbb{R}^{m-1} \supset A \to \mathbb{R}^m, \quad \phi(y) := \Big(y, h(y)\Big) = \Big(\eta_1, \dots, \eta_{m-1}, h(y)\Big).$$

Mit diesem ϕ gilt (einfache Rechnung!)

$$\mathrm{D}\phi(y) = \begin{pmatrix} \mathrm{D}_1\phi_1(y) & \dots & \mathrm{D}_{m-1}\phi_1(y) \\ \vdots & & \vdots \\ \mathrm{D}_1\phi_m(y) & \dots & \mathrm{D}_{m-1}\phi_m(y) \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \\ \mathrm{D}_1 h(y) & \dots & \mathrm{D}_{m-1} h(y) \end{pmatrix},$$

$$1 + |\operatorname{grad} h(y)|^2 = \sum_{j=1}^m |\det \mathrm{D}\phi^{(j)}(y)|^2 =: G_\phi(y)$$

mit

$$\phi^{(j)} := (\phi_1, \dots, \hat{\phi}_j, \dots, \phi_m) \quad (\text{das ist} \quad (\phi_1, \dots, \phi_m) \textit{ ohne } j\text{–te Komponente}),$$

also

$$\int_M f(x)\,\mathrm{d}o(x) = \int_A f(\phi(y)) G_\phi(y)^{1/2}\,\mathrm{d}y, ;$$

$G_\phi(\cdot)$ heißt die *Gramsche Determinante* von $\mathrm{D}\phi$. Diese Bezeichnung geht darauf zurück, daß nach einem Satz der Linearen Algebra gilt (vgl. G. Fischer: Lineare Algebra, vieweg studium, Grundkurs Mathematik; Determinanten–Multiplikationstheorem)

$$G_\phi(y) = \mathrm{Det}\,\Big((\langle\, \mathrm{D}_i\phi(y),\, \mathrm{D}_j\phi(y)\rangle)_{i,j=1,\dots,m}\Big);$$

das ist die Gramsche Determinante der m Vektoren $\{\,\mathrm{D}_1\phi(y),\ldots,\mathrm{D}_m\phi\}$ vgl.z. B. B.Walter: Analysis II, §8.8. Die folgenden Überlegungen könnten entsprechend mit dieser Darstellung von $G_\phi(\cdot)$ durchgeführt werden. (Die Parameterdarstellung nennt man *regulär* in y, wenn $G_\phi(y) \neq 0$ ist; dies wird zunächst keine Rolle spielen, vgl. auch Abschnitt 10.4)

Wir zeigen nun die Unabhängigkeit des zuletzt angegebenen Integrals von der Wahl der (stetig differenzierbaren) Parameterdarstellung. Da jede Darstellung als Graph als eine spezielle Parameterdarstellung betrachtet werden kann, ergibt sich damit gleichzeitig die Unabhängigkeit der Integrale von der Wahl der Darstellung als Graph.

Satz 9.1 *Sei M wie oben gegeben. Ist $\psi : B \to M(B \subset \mathbb{R}^{m-1})$ eine beliebige stetig differenzierbare Parameterdarstellung von M, so gilt*

$$\int_M f(x)\,\mathrm{d}o(x) = \int_B f(\psi(z))G_\psi(z)^{1/2}\,\mathrm{d}z.$$

Die angegebenen Formeln für das Oberflächenintegral liefern also für alle Darstellungen als Graph bzw. für alle Parameterdarstellungen den gleichen Wert.

Beweis. Wie eben gezeigt wurde, gilt

$$\int_M f(x)\,\mathrm{d}o(x) = \int_A f(\phi(y))G_\phi(y)^{1/2}\,\mathrm{d}y$$

(Substitution $k = \phi^{-1}\circ\psi : B \to A$; auf Grund der speziellen Struktur von ϕ gilt offenbar $k(y) = \Big(\psi_1(y),\ldots,\psi_{m-1}(y)\Big)$, d. h. k ist eine stetig differenzierbare Variablentransformation)

$$= \int_B f\Big(\phi(k(z))\Big)G_\phi(k(z))^{1/2}|\det \mathrm{D}k(z)|\,\mathrm{d}z$$

$$= \int_B f(\psi(z))\Big\{\sum_{j=1}^m\Big(\det \mathrm{D}\phi^{(j)}(k(z))\Big)^2\Big\}^{1/2}|\det \mathrm{D}k(z)|\,\mathrm{d}z$$

$$= \int_B f(\psi(z))\Big\{\sum_{j=1}^m\Big(\det \mathrm{D}\phi^{(j)}(k(z))\det \mathrm{D}k(z)\Big)^2\Big\}^{1/2}\,\mathrm{d}z$$

$$\Big(\det \mathrm{D}\phi^{(j)}(k(z))\det \mathrm{D}k(z) = \det\Big\{\mathrm{D}\phi^{(j)}(k(z))\,\mathrm{D}k(z)\Big\}$$

$$= \det \mathrm{D}(\phi^{(j)}\circ k)(z) = \det \mathrm{D}(\phi^{(j)}\circ\phi^{-1}\circ\psi)(z)$$

$$= \det \mathrm{D}\Big(\phi \circ \phi^{-1} \circ \psi(z) \text{ ohne } j\text{–te Komponente}\Big)$$
$$= \det \mathrm{D}\Big(\psi(z) \text{ ohne } j\text{–te Komponente}\Big) = \det \mathrm{D}\psi^{(j)}(z)\Big)$$
$$= \int_B f(\psi(z))\Big\{\sum_{j=1}^{m}(\det \mathrm{D}\psi^{(j)}(z))^2\Big\}^{1/2} \mathrm{d}z$$
$$= \int_B f(\psi(z))G_\psi(z)^{1/2}\,\mathrm{d}z.$$

womit die Behauptung bewiesen ist. ■

Die Formel zur Berechnung von Oberflächenintegralen, die eine Darstellung von M als Graph zu Grunde legt, ist in der Regel besonders einfach auszuwerten. Wir werden deshalb im folgenden meist diese Darstellung benutzen.

Wir betrachten einige Beispiele. Zunächst in $\mathbb{R}^2$; Hyperflächen in $\mathbb{R}^2$ sind natürlich Kurven. Ist eine Kurve Γ in $\mathbb{R}^2$ als *Graph* dargestellt, $z = h(y)$ für $a \leq y \leq b$, so erhalten wir mit der ersten Definition

$$\int_\Gamma f(x)\,\mathrm{d}o(x) = \int_\Gamma f(x)\,\mathrm{d}s(x) = \int_a^b f(y, h(y))\sqrt{1 + h'(y)^2}\,\mathrm{d}y,$$

während wir mit einer beliebigen stetig differenzierbaren *Parameterdarstellung* $\phi : [a, b] \to \mathbb{R}^2$ aus der zweiten Darstellung erhalten

$$\int_\Gamma f(x)\,\mathrm{d}o(x) = \int_b^a f(\phi(t))\sqrt{\phi_1'(t)^2 + \phi_2'(t)^2}\,\mathrm{d}t.$$

Beispiel 9.2 Für *Kreislinie* $S_{1,r}$ mit Radius r um 0

$$h : [-r, r] \to \mathbb{R}, \quad h(y) = \pm\sqrt{r^2 - y^2},$$
$$\begin{aligned}\int_{S_{1,r}} f(x)\,\mathrm{d}o(x) &= \int_{-r}^{r} f\Big(y, \sqrt{r^2 - y^2}\Big)\sqrt{1 + \frac{y^2}{r^2 - y^2}}\,\mathrm{d}y \\ &\quad + \int_{-r}^{r} f\Big(y, -\sqrt{r^2 - y^2}\Big)\sqrt{1 + \frac{y^2}{r^2 - y^2}}\,\mathrm{d}y \\ &= \int_{-r}^{r} \Big\{f\Big(y, \sqrt{r^2 - y^2}\Big) + f\Big(y, -\sqrt{r^2 - y^2}\Big)\Big\}\frac{r}{\sqrt{r^2 - y^2}}\,\mathrm{d}y.\end{aligned}$$

Integration über $f(\cdot) \equiv 1$ liefert dies den Kreisumfang $2\int_{-r}^{r} r(r^2-y^2)^{-1/2}\,\mathrm{d}y = ???$; es sei dem Leser überlassen, die Integration auszuführen, auf dem folgenden Weg geht es allerdings einfacher.

Andererseits erhalten wir aus der Parameterdarstellung

$$\phi : [0,2\pi] \to \mathbb{R}^2, \quad \phi(t) = (r\cos t, r\sin t), \quad G_\phi(t) = r^2,$$

von $S_{1,r}$ für das Integral die Formel

$$\int_{S_{1,r}} f(x)\,\mathrm{d}o(x) = \int_0^{2\pi} f(r\cos t, r\sin t) r\,\mathrm{d}t.$$

für den Kreisumfang $\int_0^{2\pi} r\,\mathrm{d}t = 2\pi r$. □

Ist eine Fläche M in $\mathbb{R}^3$ als *Graph* dargestellt, $z = h(y)$ für $y \in A \subset \mathbb{R}^2$, so erhalten wir mit der ersten Definition

$$\int_M f(x)\,\mathrm{d}o(x) := \int_A f(y,h(y))\sqrt{1 + \mathrm{D}_1 h(y)^2 + \mathrm{D}_2 h(y)^2}\,\mathrm{d}y.$$

Ist die Fläche M in einer beliebigen stetig differenzierbaren *Parameterdarstellung* $\phi : \mathbb{R}^2 \supset A \to \mathbb{R}^3$ gegeben, so erhalten wir aus der zweiten Definition

$$\begin{aligned}\int_M f(x)\,\mathrm{d}o(x) &= \int_A f(\phi(x))G_\phi(y)^{1/2}\,\mathrm{d}y \\ &= \int_A f(\phi(y))\Big\{\Big(\mathrm{D}_1\phi_1(y)\,\mathrm{D}_2\phi_2(y) - \mathrm{D}_2\phi_1(y)\,\mathrm{D}_1\phi_2(y)\Big)^2 \\ &\quad + \Big(\mathrm{D}_1\phi_2(y)\,\mathrm{D}_2\phi_3(y) - \mathrm{D}_2\phi_2(y)\,\mathrm{D}_1\phi_3(y)\Big)^2 \\ &\quad + \Big(\mathrm{D}_1\phi_3(y)\,\mathrm{D}_2\phi_1(y) - \mathrm{D}_2\phi_3(y)\,\mathrm{D}_1\phi_1(y)\Big)^2\Big\}^{1/2}\,\mathrm{d}y.\end{aligned}$$

Beispiel 9.3 Für das Integral über die *Kugeloberfläche* $S_{2,r}$ mit Radius r um 0

$$h : K_r \to \mathbb{R}^3, \quad h(y) = \pm\sqrt{r^2 - |y|^2},$$

$$\begin{aligned}\int_{S_{2,r}} f(x)\,\mathrm{d}o(x) &= \int_{K_r} f\Big(y, \sqrt{r^2-|y|^2}\Big)\sqrt{1 + \frac{|y|^2}{r^2-|y|^2}}\,\mathrm{d}y \\ &\quad + \int_{K_r} f\Big(y, -\sqrt{r^2-|y|^2}\Big)\sqrt{1 + \frac{|y|^2}{r^2-|y|^2}}\,\mathrm{d}y \\ &= \int_{K_r} \Big\{f\Big(y,\sqrt{r^2-|y|^2}\Big) + f\Big(y,-\sqrt{r^2-|y|^2}\Big)\Big\}\frac{r}{\sqrt{r^2-|y|^2}}\,\mathrm{d}y.\end{aligned}$$

Integration über $f(\cdot) \equiv 1$ liefert Kugeloberfläche $2\int_{K_r} r(r^2 - |y|^2)^{-1/2}\,\mathrm{d}y$.

Andererseits hat $S_{2,r}$ die Parameterdarstellung

$$\phi : [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3, \quad \phi(\varphi, \vartheta) = \Big(r\cos\varphi\sin\vartheta, r\sin\varphi\sin\vartheta, r\cos\vartheta\Big)$$

mit der Gramschen Determinante

$$\begin{aligned}
G_\phi(\varphi, \vartheta) &= \left\{\det\begin{pmatrix} -r\sin\varphi\sin\vartheta & r\cos\varphi\cos\vartheta \\ r\cos\varphi\sin\vartheta & r\sin\varphi\cos\vartheta \end{pmatrix}\right\}^2 \\
&\quad + \left\{\det\begin{pmatrix} r\cos\varphi\sin\vartheta & r\sin\varphi\cos\vartheta \\ 0 & -r\sin\vartheta \end{pmatrix}\right\}^2 \\
&\quad + \left\{\det\begin{pmatrix} 0 & -r\sin\vartheta \\ -r\sin\varphi\sin\vartheta & r\cos\varphi\cos\vartheta \end{pmatrix}\right\}^2 \\
&= r^4\Big\{(-\sin^2\varphi\sin\vartheta\cos\vartheta - \cos^2\varphi\sin\vartheta\cos\vartheta)^2 \\
&\quad + (-\cos\varphi\sin^2\vartheta)^2 + (-\sin\varphi\sin^2\vartheta)^2\Big\} \\
&= r^4\Big\{\sin^2\vartheta\cos^2\vartheta + \cos^2\varphi\sin^4\vartheta + \sin^2\varphi\sin^4\vartheta\Big\} \\
&= r^4\Big\{\sin^2\vartheta\cos^2\vartheta + \sin^2\vartheta\sin^2\vartheta\Big\} \;=\; r^4\sin^2\vartheta
\end{aligned}$$

(das ist vielleicht nicht ganz überraschend, vgl. Beispiel 8.13). Daraus ergibt sich für das Integral über $S_{2,r}$

$$\int_{S_{2,r}} f(x)\,\mathrm{d}o(x) = \int_0^\pi\int_0^{2\pi} f(r\cos\varphi\sin\vartheta, r\sin\varphi\sin\vartheta, r\cos\vartheta) r^2\sin\vartheta\,\mathrm{d}\varphi\,\mathrm{d}\vartheta.$$

Für die Kugeloberfläche folgt also $\int_0^\pi\int_0^{2\pi} r^2\sin\delta\,\mathrm{d}\phi\,\mathrm{d}\delta = 4\pi r^2$. □

Beispiel 9.4 Als weiteres Beispiel betrachten wir den *Kegelmantel* $M_{r,h}$ des Kegels der Höhe h über dem Kreis mit Radius r: Als Graph dargestellt haben wir (wobei, um eine Verwechslung mit der Höhe $h(\cdot)$ zu vermeiden jetzt $g(\cdot)$ statt h benutzt wird)

$$\begin{aligned}
&g(y) = h - \frac{h}{r}|y| \text{ für } y \in \mathbb{R}^2 \text{ mit } |y| \le r, \\
&\int_{M_{r,h}} f(x)\,\mathrm{d}o(x) = \int_{|y|\le r} f\Big(y, h - \frac{h}{r}|y|\Big)\left(1 + \frac{h^2}{r^2}\right)^{1/2}\mathrm{d}y \\
&\qquad = \frac{\sqrt{r^2+h^2}}{r}\int_{|y|\le r} f\Big(y, h - \frac{h}{r}|y|\Big)\,\mathrm{d}y.
\end{aligned}$$

Speziell erhalten wir für die Oberfläche dieses Kegelmantels

$$\begin{aligned} F(M_{r,h}) &= \int_{M_{r,h}} 1 \, \mathrm{d}o(x) = \frac{\sqrt{r^2+h^2}}{r} \int_{|y|\le r} 1 \, \mathrm{d}y \\ &= \frac{\sqrt{r^2+h^2}}{r} \pi r^2 = \sqrt{r^2+h^2}\pi r = \pi r L; \end{aligned}$$

man beachte, daß hier $L := \sqrt{r^2+h^2}$ die Länge der Mantellinie ist während $2\pi r$ der Kreisumfang ist, d. h. es gilt Mantelfläche $= \frac{1}{2}$ (Länge der Mantellinie) $\times$ (Kreisumfang), wie man dies aus der Elementargeometrie kennt.

Betrachtet man statt dessen die Parameterdarstellung

$$\phi : [0, r] \times [0.2\pi] \to \mathbb{R}^3, \quad \phi(\varrho, \varphi) = (\varrho \cos\varphi, \varrho \sin\varphi, h(1 - \varrho/r))$$

mit

$$\begin{aligned} G_\phi(\varrho,\varphi)^2 &= \det\begin{pmatrix} \sin\varphi & \varrho\cos\varphi \\ -h/r & 0 \end{pmatrix}^2 + \det\begin{pmatrix} \cos\varphi & -\varrho\sin\varphi \\ -h/r & 0 \end{pmatrix}^2 + \det\begin{pmatrix} \cos\varphi & -\varrho\sin\varphi \\ \sin\varphi & \varrho\cos\varphi \end{pmatrix}^2 \\ &= \left(\frac{\varrho h}{r}\cos\varphi\right)^2 + \left(\varrho\frac{h}{r}\sin\varphi\right)^2 + \varrho^2(\cos^2\varphi + \sin^2\varphi) \\ &= \varrho^2 h^2/r^2 + \varrho^2 = L^2\varrho^2/r^2. \end{aligned}$$

Also gilt

$$\begin{aligned} \int_{M_{r,h}} f(x) \, \mathrm{d}o(x) &= \int_0^r \int_0^{2\pi} f(\phi(\varrho,\varphi)) G_\phi(\varrho,\varphi) \, \mathrm{d}\varphi \, \mathrm{d}\varrho \\ &= \int_0^r \int_0^{2\pi} f(\varrho\cos\varphi, \varrho\sin\varphi, h(1-\varrho/r)) \frac{L\varrho}{r} \, \mathrm{d}\phi \, \mathrm{d}\varrho. \end{aligned}$$

Speziell für erhält man $f = 1$

$$F(M_{r,h}) = \int_0^r \int_0^{2\pi} \frac{L\varrho}{r} \, \mathrm{d}\varphi \, \mathrm{d}\varrho = \pi r L.$$

wieder die Oberfläche des Kegelmantels. □

9.2 Der allgemeine Satz von Stokes

Als Vorstufe für den allgemeinen Satz von Stokes beweisen wir den

Satz 9.5 *Sei Q ein Intervall in $\mathbb{R}^m$,*

$$Q = \Big\{x \in \mathbb{R}^m : a_j \le \xi_j \le b_j \textit{ für } j = 1, \ldots, m\Big\}.$$

Wir schreiben Q in der Form $Q = A \times [a_m, b_m]$ mit

$$A = \Big\{y \in \mathbb{R}^{m-1} : a_j \le \eta_j \le b_j \textit{ für } j = 1, \ldots, m-1\Big\}.$$

Mit einer stetig differenzierbaren Funktion $h : A \to \mathbb{R}$ sei

$$M := \Big\{(y, h(y)) : y \in A\Big\} \textit{ und } \Omega := \Big\{(y, \xi_m) \in Q : \xi_m \gtrless h(y)\Big\}.$$

a) *Für jedes $y \in A$, d. h. $(y, h(y)) \in M$, ist*

$$n(y, h(y)) = \pm\Big\{1 + |\operatorname{grad} h(y)|^2\Big\}^{-1/2}\Big(D_1 h(y), \ldots, D_{m-1} h(y), -1\Big)$$

die bezüglich Ω äußere Normale auf M im Punkt $(y, h(y))$.

b) *Ist $f : Q \to \mathbb{R}$ stetig differenzierbar mit $f(x) \equiv, 0$ in der Nähe des Randes von Q, so gilt, wobei $n_j(\cdot)$ die j–te Komponente von $n(\cdot)$ ist,*

$$\int_\Omega D_j f(x)\,\mathrm{d}x = \int_M f(x) n_j(x)\,\mathrm{do}(x) \textit{ für } j = 1, \ldots, m.$$

(Das entsprechende Resultat gilt, wenn in der Darstellung von M die k–te Komponente $(1 \le k < m)$ als Funktion der übrigen $m-1$ Komponenten aufgefaßt wird.)

Beweis. O.E. betrachten wir nur den Fall, wo in der Definition von Ω das $>$–Zeichen und in $n(y, h(y))$ das $+$–Zeichen stehen; der Leser führe selbst die Modifikationen für den anderen Fall durch.

a) Wir können M auch schreiben in der Form

$$M = \Big\{(y,\xi_m) \in \mathbb{R}^m : g(y,\xi_m) = 0\Big\} \text{ mit } g(y,\xi_m) := h(y) - \xi_m.$$

Dann wissen wir, daß

$$\operatorname{grad} g(y, h(y) = \Big(D_1 h(y), \ldots, D_{m-1} h(y), -1\Big)$$

die (nichtnormierte) Normale auf M im Punkt $(y, h(y))$ ist in die Richtung zeigt, wo $g > 0$ ist, d. h. $\xi_m < h(y)$; d. h. bezüglich Ω nach außen. Damit folgt die Aussage über $n(y, h(y))$ durch Normierung.

b) Es bleibt die Integralformel zu beweisen. Wir unterscheiden zwei Fälle:

$j = m$**:**

$$\begin{aligned}
\int_\Omega \mathrm{D}_m f(x)\,\mathrm{d}x &= \int_A \int_{h(y)}^{b_m} \mathrm{D}_m f(y,t)\,\mathrm{d}t\,\mathrm{d}y = \int_A \Big\{f(y,b_m) - f(y,h(y))\Big\}\,\mathrm{d}y \\
&\qquad \text{(da } f(y,t) \equiv 0 \text{ für } t \text{ nahe } b_m) \\
&= -\int_A f(y,h(y))\,\mathrm{d}y \\
&= -\int_A f(y,h(y)) \underbrace{\Big\{1 + |\operatorname{grad} h(y)|^2\Big\}^{-1/2}}_{=-n_m(y,h(y))} \Big\{1 + |\operatorname{grad} h(y)|^2\Big\}^{1/2}\,\mathrm{d}y \\
&\qquad \text{(nach Definition des Oberflächenintegrals)} \\
&= \int_M f(x) n_m(x)\,\mathrm{d}o(x).
\end{aligned}$$

$j < m$**:**

$$\begin{aligned}
\int_\Omega \mathrm{D}_j f(x) t D x &= \int_A \int_{h(y)}^{b_m} \mathrm{D}_j f(y,t)\,\mathrm{d}t\,\mathrm{d}y \\
&\qquad \text{(Differentiation parameterabhängiger Integrale, Satz 8.17)} \\
&= \int_A \Big\{ \mathrm{D}_j \int_{h(y)}^{b_m} f(y,t)\,\mathrm{d}t + f(y,h(y))\,\mathrm{D}_j h(y)\Big\}\,\mathrm{d}y \\
&\qquad \text{(wegen } A = A' \times [a_j, b_j],\ y = (y', \eta_j)) \\
&= \int_{A'} \int_{a_j}^{b_j} \mathrm{D}_j \int_{h(y',\eta_j)}^{b_m} f(y',\eta_j,t)\,\mathrm{d}t\,\mathrm{d}\eta_j\,\mathrm{d}y' + \int_A f(y,h(y))\,\mathrm{D}_j h(y)\,\mathrm{d}y
\end{aligned}$$

$$= \int_{A'} \Big\{ \int_{h(y',b_j)}^{b_m} f(y',b_j,t)\,\mathrm{d}t - \int_{h(y',a_j)}^{b_m} f(y',a_j,t)\,\mathrm{d}t \Big\}\,\mathrm{d}y'$$

$$+ \int_A f(y,h(y))\,\mathrm{D}_j h(y)\,\mathrm{d}y$$

(ins erste Integral gehen nur Werte von f auf dem Rand von Ω ein)

$$= 0 + \int_A f(y,h(y))\,\underbrace{\mathrm{D}_j h(y)\Big\{1+|\operatorname{grad} h(y)|^2\Big\}^{-1/2}}_{=n_j(y,h(y))}\Big\{1+|\operatorname{grad} h(y)|^2\Big\}^{1/2}\,\mathrm{d}y$$

$$= \int_M f(x) n_j(x)\,\mathrm{d}o(x).$$

Damit ist die Formel für alle j bewiesen. ■

Mit Hilfe des eben bewiesenen Satzes und des Satzes von Heine–Borel (Satz 2.15) sind wir nun in der Lage, den allgemeinen Satz von Stokes zu beweisen.

Satz 9.6 (Allgemeiner Stokesscher Satz) *Sei Ω eine beschränkte offene Menge in $\mathbb{R}^m$ mit stetig differenzierbarem Rand $M = \partial\Omega$. Ist $n(x)$ für jedes $x \in M$ die bezüglich Ω äußere Normale auf M im Punkt x und $f : \Omega \cup M \to \mathbb{R}$ stetig differenzierbar, so gilt für $j = 1, \ldots, m$*

$$\int_\Omega \mathrm{D}_j f(x)\,\mathrm{d}x = \int_M f(x) n_j(x)\,\mathrm{d}o(x).$$

Beweis. Für jedes $x \in \Omega \cup M$ sei Q_x ein offenes Intervall so, daß

entweder $Q_x \cap M$ als Graph darstellbar ist,
oder $Q_x \cap M = \emptyset$.

Dann ist $\{Q_x : x \in \Omega \cup M\}$ eine offene Überdeckung der kompakten Menge $\Omega \cup M$. Also genügen endlich viele $Q_1 := Q_{x_1}, \ldots, Q_n := Q_{x_n}$ um Ω zu überdecken.

$\{\varphi_1, \ldots, \varphi_n\}$ sei eine zur Überdeckung $\{Q_1, \ldots, Q_n\}$ gehörige Zerlegung der Eins (vgl.Satz 8.18). Dann gilt (wegen $\sum_{k=1}^n \varphi_k(x) = 1$ für alle $x \in \Omega \cup M$)

$$\int_\Omega \mathrm{D}_j f(x)\,\mathrm{d}x = \int_\Omega \mathrm{D}_j\Big(\sum_{k=1}^n \varphi_k(x) f(x)\Big)\,\mathrm{d}x = \sum_{k=1}^n \int_\Omega \mathrm{D}_j(\varphi_k(x) f(x))\,\mathrm{d}x$$

(da $\varphi_k f(x) = 0$ außerhalb von Q_k)

$$= \sum_{k=1}^n \int_{\Omega\cap Q_k} \mathrm{D}_j(\varphi_k(x) f(x))\,\mathrm{d}x.$$

Wir untersuchen die einzelnen Integrale und unterscheiden dazu zwei Fälle:

$\mathbf{Q_k} \subset \mathbf{\Omega}$ (d. h. $Q_k \cap M = \emptyset$):

$$\begin{aligned}
\int_{\Omega \cap Q_k} \mathrm{D}_j(\varphi_k(x) f(x))\,\mathrm{d}x &= \int_{Q_k} \mathrm{D}_j(\varphi_k(x) f(x))\,\mathrm{d}x \\
&= \int_{Q_k'} \int_{a_j}^{b_j} \mathrm{D}_j(\varphi_k(x) f(x))\,\mathrm{d}\xi_j\,\mathrm{d}x' \\
&= \int_{Q_k'} \Big\{\varphi_k(x', b_j) f(x', b_j) - \varphi_k(x', a_j) f(x', a_j)\Big\}\,\mathrm{d}x' \\
&= 0 \quad (\text{da } \varphi_k f(x) = 0 \text{ am Rand von } Q_k) \\
&= \int_M \varphi_k(x) f(x) n_j(x)\,\mathrm{d}o(x) \quad (\text{da } \varphi_k(x) = 0 \text{ auf } M).
\end{aligned}$$

$\mathbf{Q_k} \cap \mathbf{M}$ *ist als Graph darstellbar:* ein ξ_l als Funktion h der übrigen $m-1$ Koordinaten; $\Omega \cap Q_k$ ist dann die Menge der Punkte aus Q_k mit $\xi_l >$ oder $\xi_l < h(\xi_1, \ldots, \xi_{l-1}, \xi_{l+1}, \ldots \xi_m))$: Mit Satz 9.5 folgt also

$$\begin{aligned}
\int_{\Omega \cap Q_k} \mathrm{D}_j(\varphi_k(x) f(x))\,\mathrm{d}x &= \int_{Q_k \cap M} \varphi_k(x) f(x) n_j(x)\,\mathrm{d}o(x) \\
&\quad (\text{da } \varphi_k(x) f(x) = 0 \text{ am Rand von } Q_k) \\
&= \int_M \varphi_k(x) f(x) n_j(x)\,\mathrm{d}o(x).
\end{aligned}$$

Insgesamt erhalten wir also

$$\begin{aligned}
\int_\Omega \mathrm{D}_j f(x)\,\mathrm{d}x &= \sum_{k=1}^n \int_{\Omega \cap Q_k} \mathrm{D}_j(\varphi_k(x) f(x))\,\mathrm{d}x = \sum_{k=1}^n \int_M \varphi_k(x) f(x) n_j(x)\,\mathrm{d}o(x) \\
&= \int_M f(x) n_j(x)\,\mathrm{d}o(x),
\end{aligned}$$

womit der Satz bewiesen ist. ■

Eine genaue Durchsicht der obigen Beweise zeigt, daß die Aussage des allgmeinen Stokesschen Satzes auch dann gilt, wenn der Rand von Ω gewissen Ecken und/oder Kanten hat. Bei Bedard überlegt man das am besten für den Einzelfall, da es schwierig und auch nicht lohnend ist, eine möglichst allgemeine Form des Satzes anzugeben.

Als Anwendung betrachten wir den *Auftrieb* eines in eine inkompressible Flüssigkeit eingetauchten Körpers. Die physikalische Aussage, die wir dabei benutzen ist: in einer Tiefe t unter der Oberfläche der Flüssigkeit herrscht ein Druck pro Flächeneinheit von $t\varrho$, wobei ϱ die Dichte der Flüssigkeit ist (und zwar ist dieser Druck unabhängig von der Lage des Flächenstücks). Dazu kommt natürlich noch der evtl. an der Oberfläche herrschende Druck, der aber, wie man leicht sieht, keinen Beitrag liefert; ebenso spielt es keine Rolle, wie tief der Körper unter der Oberfläche ist, sofern er nur vollständig eingetaucht ist.

Satz 9.7 *Der Auftrieb eines (glatten) Körpers K in einem flüssigen oder gasförmigen Medium ist (unabhängig von der Form des Körpers und der Eintauchtiefe) gleich dem Gewicht der durch den Körper verdrängten Menge des Mediums; die Kraft wirkt senkrecht nach oben (hat also keine horizontale Komponente).*

Beweis. Wir gehen von der physikalischen Annahme aus, daß die spezifische Dichte des Mediums nur von ξ_3 (der senkrechten Komponente) abhängt, $\varrho(x) = \tilde{\varrho}(\xi_3)$ (bei einer inkompressiblen Flüssigkeit hat man konstantes ϱ, was den Beweis kaum vereinfacht).

Der Druck in einem Punkt mit 3–ter Komponente ξ_3 ist also

$$p(x) = \tilde{p}(\xi_3) = \tilde{p}(0) + \int_{\xi_3}^{0} \tilde{\varrho}(t)\,\mathrm{d}t.$$

In einem Punkt x der Oberfläche M des Körpers mit äußerer Normale $n(x)$ wirkt auf einer Flächeneinheit der Oberfläche also die die Kraft $p(x)n(x)$. Damit folgt für die j–te Komponente der insgesamt auf den Körper wirkenden Kraft k:

$$\begin{aligned} k_j &= \int_M \{-p(x)n_j(x)\}\,\mathrm{d}o(x) = -\int_M \tilde{p}(\xi_3)n_j(x)\,\mathrm{d}o(x) \\ &\qquad \text{(allg. Stokesscher Satz mit } f(x) = p(x)) \\ &= \begin{cases} -\int_K D_j\tilde{p}(\xi_3)\,\mathrm{d}x = 0 & \text{für } j = 1,2, \\ -\int_K D_3\tilde{p}(\xi_3)\,\mathrm{d}x = \int_K \tilde{\varrho}(\xi_3)\,\mathrm{d}x = G(K) & \text{für } j = 3 \end{cases} \end{aligned}$$

Damit sind alle Aussagen des Satzes bewiesen. ■

9.3 Integrale auf allgemeineren Mannigfaltigkeiten

Sei $A \subset \mathbb{R}^k$ mit $k < m$ eine offene Jordan–meßbare Teilmenge, $\phi : A \to \mathbb{R}^m$ stetig differenzierbar,

$$\phi^{(j_1,\ldots,j_k)}(y) = \Big(\phi_{j_1}(y), \ldots, \phi_{j_k}(y)\Big) \qquad 1 \le j_1 < j_2 < \ldots < j_k \le m.$$

Dann ist die *Gramsche Determinante* von ϕ definiert durch

$$G_\phi(y) = \sum_{1\le j_1<j_2<\ldots<j_k\le m} \Big(\det \mathrm{D}\phi^{(j_1,\ldots,j_k)}(y)\Big)^2;$$

es folgt wieder aus der Linearen Algebra, daß gilt (vgl. G. Fischer: Lineare Algebra; Determinanten–Multiplikationstheorem)

$$G_\phi(y) = \det\Big((\langle\, \mathrm{D}_i\phi(y),\, \mathrm{D}_j\phi(y)\rangle)_{i,j=1,\ldots,k}\Big).$$

Ist $G_\phi(y) \neq 0$ für alle $y \in A$, d. h. ϕ ist *regulär*, so heißt das Bild $M = \phi(A)$ eine differenzierbare Mannigfaltigkeit (allgemeiner ist eine differenzierbare Mannigfaltigkeit ein Gebilde, das stückweise in dieser Form dargestellt werden kann; für die Parameterdarstellungen ϕ, ψ von zwei Teilstücken muß dann $\phi^{-1} \circ \psi$, dort wo es definiert ist [d. h. auf dem Schnitt der Wertebereiche von ϕ und ψ] stetig differenzierbar sein mit $\det \mathrm{D}(\phi^{-1} \circ \psi)(y) \neq 0$, *Verträglichkeit der „lokalen Karten“.*) Das Integral einer (z. B. stetigen) Funktion auf M wird definiert durch

$$\int_M f(x)\, \mathrm{d}o(x) := \int_A f(\phi(y)) G_\phi(y)^{1/2}\, \mathrm{d}y.$$

Der Faktor $G_\phi(\cdot)^{1/2}$ steht hier, weil er gleich dem Volumen des durch die Vektoren $D_1\phi(\cdot), \ldots, D_k\phi(\cdot)$ aufgespannten Parallelflachs ist, dem differentiellen Bild des Einheitswürfels im Punkt $y \in A \subset \mathbb{R}^k$. [6] Wie in Abschnitt 26.1 kann die Unabhängigkeit dieses Integrals von der

[6] Sind $x_1, \ldots, x_k$ linear unabhängige Vektoren in $\mathbb{R}^m$, $\{e_1, \ldots, e_k\}$ eine ONB in $L\{x_1, \ldots, x_k\}$, $x_j = \sum_{l=1}^k c_{jl} e_l$ für $j = 1, \ldots, k$, so gilt für das Volumen V des durch $x_1, \ldots, x_k$ aufgespannten Parallesflachs

$$\begin{aligned} V^2 &= \left(\det \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ \vdots & \vdots & \\ c_{k1} & \cdots & c_{kk} \end{pmatrix}\right)^2 = \left(\det \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ \vdots & \vdots & \\ c_{k1} & \cdots & c_{kk} \end{pmatrix}\right)^2 \left(\det \begin{pmatrix} c_{11} & \cdots & c_{k1} \\ \vdots & \vdots & \\ c_{1k} & \cdots & c_{kk} \end{pmatrix}^2\right) \\ &= \det \begin{pmatrix} \sum c_{1l}c_{1l} & \cdots & \sum c_{1l}c_{kl} \\ \vdots & & \\ \sum c_{kl}c_{1l} & \cdots & \sum c_{kl}c_{kl} \end{pmatrix} = \det \begin{pmatrix} \langle x_1, x_1\rangle & \cdots & \langle x_1, x_k\rangle \\ \langle x_k, x_1\rangle & \cdots & \langle x_k, x_k\rangle \end{pmatrix}. \end{aligned}$$

Parameterdarstellung bewiesen werden.

Beispiel 9.8 *Kurve in* $\mathbb{R}^{\mathbf{m}}$: In diesem Fall ist A ein Intervall $[a, b]$ und

$$G_\phi(y) = \sum_{1 \le j \le m} \phi_j'(y)^2,$$

$$\int_M f(x)\,\mathrm{d}o(x) = \int_a^b f(\phi(y)) \Big\{ \sum_{j=1}^m \phi_j'(y)^2 \Big\}^{1/2} \mathrm{d}y.$$

Dies ist das bereits bekannte Integral über eine Kurve, vgl. S. 128. Für beliebige m hatten wir das bisher nicht betrachtet, außer im Fall $f \equiv 1$, wo wir die bereits bekannte Formel für die Kurvenlänge erhalten (vgl. §22). □

9.4 Übungsaufgaben

9.1 Man stelle die Oberfläche der Sphäre mit Radius r in $\mathbb{R}^3$ in Polarkoordinaten dar und berechne damit die Oberfläche (als Integral über die Funktion $f \equiv 1$).

9.2 Man berechne die Gramsche Determinante für die Polarkoordinatendarstellung der Fläche

$$\Big\{ (x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \Big\}.$$

9.3 Man berechne die Oberfläche der Einheitsspähre in $\mathbb{R}^4$ nach dem Vorbild von Beispiel 9.2 und 9.3.

10 Die klassischen Integralsätze von Gauß, Green und Stokes

10.1 Die Sätze von Gauß und Green

Die beiden Sätze dieses Abschnitts sind einfache Folgerungen aus dem allgemeinen Satz von Stokes. Es sei hier Ω eine offene Teilmenge von $\mathbb{R}^m$ mit einem stetig differenzierbaren Rand $M = \partial\Omega$ wie dies z. B. in Satz 9.6 vorausgesetzt wurde.

Für ein stetig differenzierbares Vektorfeld $f : \Omega \to \mathbb{R}^m$ ist die *Divergenz* definiert durch

$$\operatorname{div} f(x) := \sum_{j=1}^{m} \mathrm{D}_j f_j(x).$$

Der folgende Satz liefert eine anschauliche physikalische Deutung der Divergenz.

Satz 10.1 (Gaußscher Divergenzsatz) *Sei $\Omega \subset \mathbb{R}^m$ offen mit stetig differenzierbarem Rand M; für jedes $x \in M$ sei $n(x)$ die normierte bezüglich Ω äußere Normale auf M im Punkt x. Dann gilt*

$$\int_{\Omega} \operatorname{div} f(x)\,\mathrm{d}x = \int_{M} \langle f(x), n(x)\rangle\,\mathrm{d}o(x).$$

Bemerkung 10.2 a) Dieser Satz stellt offenbar die direkte Verallgemeinerung des Hauptsatzes der Differential– und Integralrechnung in $\mathbb{R}^1$ auf den $\mathbb{R}^m$ dar (für ein Intervall $[a, b]$ zeigt die äußere Normale $n(x)$ in a nach links, in b nach rechts, ist also in a negativ, in b positiv).

b) Faßt man das Vektorfeld f als *Geschwindigkeitsfeld* einer Strömung in $\mathbb{R}^m$ auf, so gibt offenbar $\langle f(x), n(x)\rangle$ den Fluß pro Flächeneinheit durch ein „kleines" Flächenstück von M um den Punkt x von Ω nach $\mathbb{R}^m\backslash\Omega$ an. Das Integral auf der rechten Seite gibt also den gesamten Fluß von Ω durch M nach $\mathbb{R}^m\backslash\Omega$ an. Da dies entsprechend für Teilgebiete von Ω gilt, kann (nach Grenzübergang für immer kleinere Volumina) $\operatorname{div} f(x)$ als *Quellendichte* des Vektorfeldes f aufgefaßt werden; das Integral auf der linken Seite stellt dann die gesamte *Quellenstärke* des Vektorfeldes f in Ω dar.

Beweis von Satz10.1. Durch Anwendung von Satz 9.6 auf f_j statt f erhalten wir

$$\int_\Omega \operatorname{div} f(x)\,\mathrm{d}x = \sum_{j=1}^m \int_\Omega \mathrm{D}_j f_j(x)\,\mathrm{d}x = \sum_{j=1}^m \int_M f_j(x) n_j(x)\,\mathrm{d}o(x)$$

$$= \int_M \sum_{j=1}^m f_j(x) n_j(x)\,\mathrm{d}o(x) = \int_M \langle f(x), n(x)\rangle\,\mathrm{d}o(x),$$

was zu beweisen war. ■

Beispiel 10.3 In $\mathbb{R}^m$ betrachten wir das Vektorfeld

$$f : \mathbb{R}^m \to \mathbb{R}^m, \quad f(x) = x.$$

Dann gilt

$$\operatorname{div} f(x) = \sum_{j=1}^m \mathrm{D}_j x_j = \sum_{j=1}^m 1 = m.$$

Wählen wir z.B.

$\Omega = K_{m,r}$ Kugel mit Radius r um 0,
$M = S_{m-1,r}$ Sphäre mit Radius r um 0,

so folgt für $V(K_{m,r})$ = Volumen von $K_{m,r}$ und $O(S_{m-1,r})$ = Oberfläche von $S_{m-1,r}$ aus obigem Satz mit $f(x) = x$

$$\begin{aligned} mV(K_{m,r}) &= \int_{K_{m,r}} m\,\mathrm{d}x = \int_{K_{m,r}} \operatorname{div} f(x)\,\mathrm{d}x = \int_{S_{m-1,r}} \langle f(x), n(x)\rangle\,\mathrm{d}o(x) \\ &= \int_{S_{m-1,r}} r\,\mathrm{d}o(x) = rO(S_{m-1,r}). \end{aligned}$$

Dies bestätigt insbesondere die folgenden bekannten Formeln:

$\mathbf{m} = \mathbf{2}$: $V(K_{2,r}) = \pi r^2, \quad O(S_{1,r}) = 2\pi r.$

$\mathbf{m} = \mathbf{3}$: $V(K_{3,r}) = \dfrac{4}{3}\pi r^3, \quad O(S_{2,r})4\pi r^2.$

Als weiteres Beispiel betrachten wir

m = 4: Nach unserer Regel über iterative Integration (Satz 8.4)gilt

$$\begin{aligned} V(K_{4,r}) &= \int\limits_{K_{4,r}} 1\,\mathrm{d}x = \int\limits_{-r}^{r} \int\limits_{K_{3,(r^2-t^2)^{1/2}}} 1\,\mathrm{d}x'\,\mathrm{d}t = \int\limits_{-r}^{r} V\Big(K_{3,(r^2-t^2)^{1/2}}\Big)\,\mathrm{d}t \\ &= \frac{4}{3}\pi \int\limits_{-r}^{r} (r^2-t^2)^{3/2}\,\mathrm{d}t. \end{aligned}$$

Mit der Substitution

$$t = \varphi(\tau) \text{ mit } \varphi : \Big[-\frac{\pi}{2}, \frac{\pi}{2}\Big] \to \mathbb{R},\ \varphi(\tau) = r\sin\tau,\ \varphi'(\tau) = r\cos\tau$$

erhalten wir also [7]

$$\begin{aligned} V(K_{4,r}) &= \frac{4}{3}\pi r^4 \int\limits_{-\pi/2}^{\pi/2} \cos^4\tau\,\mathrm{d}\tau \\ &= \frac{4}{3}\pi r^4 \Big\{\Big(\frac{1}{4}\cos^3\tau + \frac{3}{8}\cos\tau\Big)\sin\tau + \frac{3}{8}\tau\Big\}\Big|_{-\pi/2}^{\pi/2} \\ &= \frac{4}{3}\pi r^4 \Big\{\frac{3}{8}\frac{\pi}{2} + \frac{3}{8}\frac{\pi}{2}\Big\} = \frac{1}{2}\pi^2 r^4. \end{aligned}$$

Mit obiger Formel ergibt sich daraus (vgl. auch Aufgabe 9.3)

$$O(S_{3,r}) = \frac{4}{r}V(K_{4,r}) = 2\pi^2 r^3.$$

---

[7] Dabei benutzen wir

$$\begin{aligned} \int \cos^4\tau\,\mathrm{d}\tau &= \int \cos^3\cos\tau\,\mathrm{d}\tau = \cos^3\tau\sin\tau + 3\int \cos^2\tau\sin^2\tau\,\mathrm{d}\tau \\ &= \cos^3\tau\sin\tau + 3\int\cos^2\tau\,\mathrm{d}\tau - 3\int\cos^4\tau\,\mathrm{d}\tau \\ &= \frac{1}{4}\cos^3\tau\sin\tau + \frac{3}{4}\int\cos^2\tau\,\mathrm{d}\tau \\ &= \frac{1}{4}\cos^3\tau\sin\tau + \frac{3}{4}\cos\tau\sin\tau + \frac{3}{4}\sin^2\tau\,\mathrm{d}\tau \\ &= \frac{1}{4}\cos^3\tau\sin\tau + \frac{3}{4}\cos\tau\sin\tau + \frac{3}{4}\int 1\,\mathrm{d}\tau - \frac{3}{4}\int\cos^2\tau\,\mathrm{d}\tau \\ &= \frac{1}{4}\cos^3\tau\sin\tau + \frac{3}{8}\cos\tau\sin\tau + \frac{3}{8}\tau. \end{aligned}$$

Formeln für Volumen der Kugeln und Oberflächen der Sphären beliebiger Dimension findet man z. B. in O.Forster: Analysis 3, Beispiel 5.7. □

Der folgende Satz ist besonders nützlich beim Umgang mit dem *Laplaceschen Differentialausdruck* $\Delta = \sum_{j=1}^m D_j^2$.

Satz 10.4 (Greensche Formel) *Seien $\Omega \subset \mathbb{R}^m$ offen mit stetig differenzierbarem Rand M, $f, g : \Omega \cup M \to \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt*

$$\begin{aligned}&\int_\Omega \Big\{ f(x)\Delta g(x) - (\Delta f(x))g(x) \Big\} \,\mathrm{d}x \\ &= \int M \Big\{ f(x) \Big\langle \operatorname{grad} g(x), n(x) \Big\rangle - g(x) \Big\langle \operatorname{grad} f(x), n(x) \Big\rangle \Big\} \,\mathrm{d}(x).\end{aligned}$$

(Dabei ist $\langle \operatorname{grad} h(x), n(x)\rangle =:$ „$\partial h(x)/\partial n(x)$" für $x \in M$ die Richtungsableitung von h in Richtung der äußeren Normalen auf M, kurz die Normalableitung *von h auf M.)*

Bemerkung 10.5 Dieses Resultat entspricht der 1–dimensionalen Formel

$$\begin{aligned}&\int_a^b \Big\{ f(x)g''(x) - f''(x)g(x) \Big\} \,\mathrm{d}x \\ &= \Big\{ f(b)g'(b) - f'(b)g(b) \Big\} - \Big\{ f(a)g'(a) - f'(a)g(a) \Big\}.\end{aligned}$$

Beweis von Satz 10.4. Mit Hilfe des Divergenzsatzes (Satz 10.1) folgt

$$\begin{aligned}&\int_\Omega \Big\{ f(x)\Delta g(x) - (\Delta f(x))g(x) \Big\} \,\mathrm{d}x \\ &= \int_\Omega \operatorname{div} \Big\{ f(x) \operatorname{grad} g(x) - g(x) \operatorname{grad} f(x) \Big\} \,\mathrm{d}x \\ &= \int_M \Big\{ \Big\langle f(x) \operatorname{grad} g(x), n(x) \Big\rangle - \Big\langle g(x) \operatorname{grad} f(x), n(x) \Big\rangle \Big\} \,\mathrm{d}o(x).\end{aligned}$$

Das ist die gewünschte Gleichung. ■

Besonders wichtig ist der folgende Spezialfall der Greenschen Formel:

Korollar 10.6 *Sind f und g wie im Satz 10.4 und gilt außerdem $f(x) = g(x) = 0$ auf M (oder $f(x) = 0$ und $\operatorname{grad} f(x) = 0$, oder $g(x) = 0$ und $\operatorname{grad} g(x) = 0$), so folgt*

$$\int_\Omega \Big\{ f(x)\Delta g(x) - g(x)\Delta f(x) \Big\}\, \mathrm{d}x = 0.$$

10.2 Der klassische Satz von Stokes in der Ebene

Wir beweisen hier zunächst den klassischen Stokesschen Satz im Spezialfall der Ebene $\mathbb{R}^2$. Den entsprechenden Satz im Raum $\mathbb{R}^3$ werden wir im folgenden Abschnitt beweisen.

Für ein stetig differenzierbares Vektorfeld $f : \Omega \to \mathbb{R}^2$ ($\Omega \subset \mathbb{R}^2$) definieren wir die *Rotation*

$$\operatorname{rot} f(x) := \mathrm{D}_1 f_2(x) - \mathrm{D}_2 f_1(x).$$

(Würde man f als Vektorfeld in $\mathbb{R}^3$ betrachten, das nicht von ξ_3 abhängt mit $f_3(x) = 0$, so wäre dies gerade die dritte Komponente der Rotation von f, vgl. Abschnitt 10.3; allerdings wären in diesem Fall die anderen beiden Komponenten gleich Null.)

Satz 10.7 (Stokesscher Satz in der Ebene) *Sei Ω ein Gebiet in $\mathbb{R}^2$ stetig differenzierbarer Randkurve $M = \Gamma$, die so orientiert sei, daß Ω links von Γ liegt (Umlauf um Ω entgegen dem Uhrzeigersinn; mathematisch positiv) $\gamma : [a,b] \to \mathbb{R}^2$, $\gamma(a) = \gamma(b)$.*
Ist $f : \Omega \cup \Gamma \to \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld, so gilt

$$\int_\Omega \operatorname{rot} f(x)\, \mathrm{d}x = \int_\Gamma f_1(x)\, \mathrm{d}\xi_1 + f_2(x)\, \mathrm{d}\xi_2 = \int_a^b \Big\langle f(\gamma(t)), \gamma'(t) \Big\rangle\, \mathrm{d}t.$$

Beweis. Mit dem Gaußschen Satz (angewandt auf f_2 mit $j = 1$ und auf f_1 mit $j = 2$) erhalten wir

$$\begin{aligned} \int_\Omega \operatorname{rot} f(x)\, \mathrm{d}x &= \int_\Omega \Big(\mathrm{D}_1 f_2(x) - \mathrm{D}_2 f_1(x) \Big)\, \mathrm{d}x \\ &= \int_M \Big\{ f_2(x) n_1(x) - f_1(x) n_2(x) \Big\}\, \mathrm{d}o(x) \end{aligned}$$

$$= \int_M \Big\langle \Big(f_1(x), f_2(x)\Big), \Big(-n_2(x), n_1(x)\Big)\Big\rangle \,\mathrm{d}o(x)$$

$$= \int_a^b \Big\langle \Big(f_1(\gamma(t)), f_2(\gamma(t))\Big), \Big(-n_2(\gamma(t)), n_1(\gamma(t))\Big)\Big\rangle |\gamma'(t)| \,\mathrm{d}t.$$

Der Vektor $(-n_2(x), n_1(x))$ entsteht aus dem äußeren Normalenvektor $n(x) = (n_1(x), n_2(x))$ durch Drehung um $90^\circ = \pi/2$ entgegen dem Uhrzeigersinn, nämlich durch Anwenden der Matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}.$$

Also ist $(-n_2(x), n_1(x))$ der Tangentialvektor im Punkt x an die Kurve Γ, und zwar in Richtung der Orientierung von Γ. Folglich können wir wie folgt weiterrechnen:

$$= \int_a^b \Big\langle f(\gamma(t)), \frac{\gamma'(t)}{|\gamma'(t)|}\Big\rangle |\gamma'(t)| \,\mathrm{d}t = \int_a^b \Big\langle f(\gamma(t)), \gamma'(t)\Big\rangle \,\mathrm{d}t.$$

Das ist die behauptete Gleichung. ■

Betrachtet man speziell ein Vektorfeld f mit $\operatorname{rot} f(x) \equiv 1$, z. B. $f(x) = (-\xi_2, 0)$ oder $(0, \xi_1)$ oder $\frac{1}{2}(-\xi_2, \xi_1)$, so liefert der obige Satz eine Formel für die Fläche, ausgedrückt durch das Kurvenintegral von f über die Randkurve (vgl. Aufgabe 10.3).

10.3 Der klassische Satz von Stokes im $\mathbb{R}^3$

Wir denken uns jetzt eine („krumme") Fläche F in $\mathbb{R}^3$ dargestellt als Graph über einem ebenen Flächenstück $\tilde{F}$ im $\mathbb{R}^2$. Den allgemeinsten interessierenden Fall kann man sich aus solchen Flächenstücken zusammengesetzt denken (dopelt auftretende Kurvenstücke heben sich dabei im Integral weg, da sie in unterschiedlichen Richtungen durchlaufen werden). Sei also

$$F = \Big\{\Big(\xi_1, \xi_2, h(\xi_1, \xi_2)\Big) \in \mathbb{R}^3 : (\xi_1, \xi_2) \in \tilde{F}\Big\}$$

mit einer stetig differenzierbaren Funktion $h : \widetilde{F} \to \mathbb{R}$. $\tilde{\Gamma}$ sei die Randkurve von $\tilde{F}$, $\tilde{\gamma} : [a,b] \to \mathbb{R}^2$, wobei $\tilde{F}$ wieder links von $\widetilde{\Gamma}$ liege. Dann ist die Randkurve Γ von F gegeben durch

$$\gamma : [a,b] \to \mathbb{R}^3, \quad \gamma(t) = \Big(\tilde{\gamma}(t), h(\tilde{\gamma}(t))\Big).$$

Weiter sei $n(x)$ die bezüglich ξ_3 nach oben zeigende Normale auf F im Punkt $x \in F$. (Damit kann die „Orientierung“der Fläche F und der Randkurve Γ auch ohne Bezug auf $\widetilde{F}$ und $\widetilde{\Gamma}$ wie folgt beschrieben werden. Schreitet man aufrecht im Sinne der Normalen $n(x)$ in Richtung der Kurve Γ voran, so liegt F links. So verstanden darf $n(x)$ auch nach unten zeigen; dann wäre auch die Orientierung der Randkurve zu ändern.)

Für ein stetig differenzierbares Vektorfeld $f : A \to \mathbb{R}^3$ sei die *Rotation* von f definiert durch

$$\operatorname{rot} f(x) := \Big(\mathrm{D}_2 f_3(x) - \mathrm{D}_3 f_2(x), \mathrm{D}_3 f_1(x) - \mathrm{D}_1 f_3(x), \mathrm{D}_1 f_2(x) - \mathrm{D}_2 f_1(x)\Big).$$

Wie schon oben angemerkt, entspricht die dritte Komponente gerade der Rotation in $\mathbb{R}^2$. Mit diesen Vereinbarungen gilt

Satz 10.8 (Stokesscher Integralsatz in $\mathbb{R}^3$) *Seien F und Γ wie oben, $f : F \cup \Gamma \to \mathbb{R}^3$ stetig differenzierbar (d. h. f ist auf einer offenen $\mathbb{R}^3$–Umgebung von $F \cup \Gamma$ stetig differenzierbar). Dann gilt*

$$\begin{aligned}\int_F \Big\langle \operatorname{rot} f(x), n(x) \Big\rangle \, \mathrm{d}o(x) &= \int_\Gamma f_1(x)\, \mathrm{d}\xi_1 + f_2(x)\, \mathrm{d}\xi_2 + f_3(x)\, \mathrm{d}\xi_3 \\ &= \int_a^b \Big\langle f(\gamma(t)), \gamma'(t) \Big\rangle \, \mathrm{d}t.\end{aligned}$$

(Im Beweis tauchen an einer Stelle zweite Ableitungen von h auf, d. h. es wird eigentlich benutzt, daß h zweimal stetig differenzierbar ist. Durch geeignete Approximation folgt die Aussage aber auch für stetig differenzierbares h.)

Beweis. Mit den obigen Bezeichnungen gilt

$$\int_\Gamma \Big\langle f(\gamma(t)), \gamma'(t) \Big\rangle \, \mathrm{d}t$$

$$
\begin{aligned}
&= \int_a^b \Big\langle f\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big), \Big(\tilde\gamma_1'(t), \tilde\gamma_2'(t),\, \mathrm{D}_1 h(\tilde\gamma(t))\tilde\gamma_1'(t) + \mathrm{D}_2 h(\tilde\gamma(t))\tilde\gamma_2'(t)\Big)\Big\rangle \,\mathrm{d}t \\
&= \int_a^b \Big\{ f_1\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big)\tilde\gamma_1'(t) + f_2\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big)\tilde\gamma_2'(t) \\
&\qquad + f_3\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big)\Big(\mathrm{D}_1 h(\tilde\gamma(t))\tilde\gamma_1'(t) + \mathrm{D}_2 h(\tilde\gamma(t))\tilde\gamma_2'(t)\Big)\Big\} \,\mathrm{d}t \\
&= \int_a^b \Big\{ \underbrace{\Big(f_1\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big) + f_3\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big)\,\mathrm{D}_1 h(\tilde\gamma(t))\Big)}_{=:k_1(\tilde\gamma(t))} \tilde\gamma_1'(t) \\
&\qquad + \underbrace{\Big(f_2\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big) + f_3\Big(\tilde\gamma(t), h(\tilde\gamma(t))\Big)\,\mathrm{D}_2 h(\tilde\gamma(t))\Big)}_{=:k_2(\tilde\gamma(t))} \tilde\gamma_2'(t)\Big\} \,\mathrm{d}t \\
&= \int_a^b \langle k(\tilde\gamma(t)), \tilde\gamma'(t)\rangle \,\mathrm{d}t
\end{aligned}
$$

(Satz von Stokes in der Ebene)

$$
\begin{aligned}
&= \int_{\tilde F} \operatorname{rot} k(\xi_1,\xi_2)\,\mathrm{d}(\xi_1,\xi_2) \\
&= \int_{\tilde F} \Big\{ \frac{\partial}{\partial \xi_1}\Big(f_2(\xi_1,\xi_2,h(\xi_1,\xi_2)) + f_3(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_2 h(\xi_1,\xi_2)\Big) \\
&\qquad - \frac{\partial}{\partial \xi_2}\Big(f_1(\xi_1,\xi_2,h(\xi_1,\xi_2)) + f_3(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_1 h(\xi_1,\xi_2)\Big)\Big\}\,\mathrm{d}(\xi_1,\xi_2) \\
&= \int_{\tilde F} \Big\{ \mathrm{D}_1 f_2(\xi_1,\xi_2,h(\xi_1,\xi_2)) + \mathrm{D}_3 f_2(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_1 h(\xi_1,\xi_2) \\
&\qquad + \mathrm{D}_1 f_3(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_2 h(\xi_1,\xi_2) \\
&\qquad + \mathrm{D}_3 f_3(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_1 h(\xi_1,\xi_2)\,\mathrm{D}_2 h(\xi_1,\xi_2) \qquad (*) \\
&\qquad + f_3(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_1\,\mathrm{D}_2 h(\xi_1,\xi_2) \qquad (\dagger) \\
&\qquad - \mathrm{D}_2 f_1(\xi_1,\xi_2,h(\xi_1,\xi_2)) - \mathrm{D}_3 f_1(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_2 h(\xi_1,\xi_2) \\
&\qquad - \mathrm{D}_2 f_3(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_1 h(\xi_1,\xi_2) \\
&\qquad - \mathrm{D}_3 f_3(\xi_1,\xi_2,h(\xi_1,\xi_2))\,\mathrm{D}_2 h(\xi_1,\xi_2)\,\mathrm{D}_1 h(\xi_1,\xi_2) \qquad (*) \\
&\qquad - f_3(\xi_1,\xi_2,h(\xi_1\xi_2))\,\mathrm{D}_2\,\mathrm{D}_1 h(\xi_1,\xi_2)\Big\}\,\mathrm{d}(\xi_1,\xi_2) \qquad (\dagger)
\end{aligned}
$$

(die Zeilen mit $(*)$ bzw. $(\dagger)$ heben sich weg)

$$= \int_{\tilde{F}} \Big\{ \Big(\mathrm{D}_1 f_2(\xi_1,\xi_2,h(\xi_1,\xi_2)) - \mathrm{D}_2 f_1(\xi_1,\xi_2,h(\xi_1,\xi_2)) \Big) \cdot 1$$

$$-\Big(\mathrm{D}_2 f_3(\xi_1,\xi_2,h(\xi_1,\xi_2)) - \mathrm{D}_3 f_2(\xi_1,\xi_2,h(\xi_1,\xi_2)) \Big) \mathrm{D}_1 h(\xi_1,\xi_2)$$

$$-\Big(\mathrm{D}_3 f_1(\xi_1,\xi_2,h(\xi_1,\xi_2)) - \mathrm{D}_1 f_3(\xi_1,\xi_2,h(\xi_1,\xi_2)) \Big) \mathrm{D}_2 h(\xi_1,\xi_2) \Big\} \,\mathrm{d}(\xi_1,\xi_2)$$

$$= \int_{\tilde{F}} \Big\langle \operatorname{rot} f(\xi_1,\xi_2,h(\xi_1,\xi_2)), \Big(-\mathrm{D}_1 h(\xi_1,\xi_2), -\mathrm{D}_2 h(\xi_1,\xi_2), 1 \Big) \Big\rangle \,\mathrm{d}(\xi_1,\xi_2)$$

$\Big($wegen $M = \{x \in \mathbb{R}^3 : g(x) := \xi_3 - h(\xi_1,\xi_2) = 0\}$ ist

$$(-\mathrm{D}_1 h(\xi_1,\xi_2), -\mathrm{D}_2 h(\xi_1,\xi_2), 1) = \sqrt{1 + |\operatorname{grad} h(\xi_1,\xi_2)|^2}\, n(\xi_1,\xi_2,h(\xi_1,\xi_2))\Big)$$

$$= \int_{\tilde{F}} \Big\langle \operatorname{rot} f(\xi_1,\xi_2,h(\xi_1,\xi_2)), n(\xi_1,\xi_2,h(\xi_1,\xi_2)) \Big\rangle \sqrt{1 + |\operatorname{grad} h(\xi_1,\xi_2)|^2} \,\mathrm{d}(\xi_1,\xi_2)$$

$$= \int_{F} \Big\langle \operatorname{rot} f(x), n(x) \Big\rangle \,\mathrm{d}o(x),$$

womit der Satz bewiesen ist. ■

Bemerkung 10.9 Das Kurvenintegral über die (geschlossene) Randkurve Γ von F liefert die Arbeit die geleistet wird, wenn sich ein Körper unter Einwirkung des Kraftfeldes f längs Γ bewegt; man nennt dieses Integral die *Zirkulation* des Feldes f längs der geschlossenen Kurve Γ. Der Stokessche Satz besagt, daß die Zirkulation von f längs Γ gleich dem Fluß des Feldes $\operatorname{rot} f$ durch die Fläche F ist. Da der Stokessche Satz auch für alle Teilstücke von F entsprechend gilt, kann man $\operatorname{rot} f(\cdot)$ als die *Wirbeldichte* (Zirkulationsdichte) von f bezeichnen.

10.4 Wegunabhängigkeit von Kurvenintegralen

Über das Resultat von Satz 5.8 hinaus (Wegunabhängigkeit des Kurvenintegrals genau dann, wenn das Vektorfeld ein Gradientenfeld ist) beweisen wir hier zwei weitere Kriterien für die Wegunabhängigkeit des Kurvenintegrals eines Vektorfeldes, die i.allg. wesentlich einfacher zu verifizieren sind. Wir betrachten zunächst Vektorfelder im gesamten Raum $\mathbb{R}^m$ $(m = 2, 3)$ und weiter unten auf gewissen Teilmengen; abschließend werden wir dann den Fall beliebiger Dimension $m \geq 2$ betrachten.

Satz 10.10 *Sei $f : \mathbb{R}^m \to \mathbb{R}^m$ ($m = 2$ oder 3) stetig differenzierbar. Das Kurvenintegral von f ist genau dann wegunabhängig, wenn gilt rot $f(x) = 0$.*

Beweis. $\Rightarrow$: Ist das Kurvenintegral von f wegunabhängig, so wissen wir aus Satz 5.8, daß gilt $f(x) = \operatorname{grad} V(x)$ mit einer stetig differenzierbaren Funktion $V : \mathbb{R}^m \to \mathbb{R}$ (*Potential* von f). Da f selbst wieder stetig differenzierbar ist, ist V tatsächlich zweimal stetig differenzierbar und man rechnet leicht nach (im Fall $m = 2$ steht hier nur die dritte Komponente)

$$\begin{aligned}\operatorname{rot} f(x) &= \operatorname{rot}\operatorname{grad} V(x)\\ &= \Big((\mathrm{D}_2\,\mathrm{D}_3 - \mathrm{D}_3\,\mathrm{D}_2)V(x), (\mathrm{D}_3\,\mathrm{D}_1 - \mathrm{D}_1\,\mathrm{D}_3)V(x), (\mathrm{D}_1\,\mathrm{D}_2 - \mathrm{D}_2\,\mathrm{D}_1)V(x)\Big)\\ &= 0.\end{aligned}$$

$\Leftarrow$: Wir nehmen an, daß rot $f(x) = 0$ gilt. Seien γ_i zweimal stetig differenzierbare Parameterdarstellungen von Γ_i wegen $(i = 1, 2)$ mit gleichen Anfangs– und Endpunkten,

$$\phi(s,t) := (1-s)\gamma_1(t) + s\gamma_2(t).$$

Offenbar ist ϕ stetig differenzierbar mit $\phi(0,t) = \gamma_1(t)$ und $\phi(1,t) = \gamma_2(t)$. Wir können ϕ als Parameterdarstellung einer Fläche F in $\mathbb{R}^3$ auffassen, deren Rand Γ gleich $\Gamma_1 - \Gamma_2$ ist (damit ist die Orientierung von F, und somit die Normale $n(\cdot)$, festgelegt). Also gilt

$$\int_{\Gamma_1} - \int_{\Gamma_2} \sum_{j=1}^{m} f_j(x)\,\mathrm{d}\xi_j = \int_{\Gamma} \ldots = \int_F \Big\langle \operatorname{rot} f(x), n(x) \Big\rangle\,\mathrm{d}o(x) = 0,$$

da rot $f(x)$ verschwindet. ■

Im folgenden sei $\Omega \subset \mathbb{R}^m$ offen. Wir nennen zwei stetig differenzerbare Kurven Γ_1 und Γ_2 in Ω *homotop* bezüglich Ω, wenn sie gleiche Anfangs– und Endpunkte haben und wenn es eine stetig differenzierbare Funktion

$$\phi : [0,1] \times [a,b] \to \Omega$$

gibt mit $\phi(s,a) = \gamma_i(a)$, $\phi(s,b) = \gamma_i(b)$ für alle s und

$$\begin{aligned}\gamma_1(\cdot) &= \phi(0,\cdot) \text{ ist eine Parameterdarstellung von } \Gamma_1,\\ \gamma_2(\cdot) &= \phi(1,\cdot) \text{ ist eine Parameterdarstellung von } \Gamma_2.\end{aligned}$$

In diesem Sinne haben wir im Beweis des obigen Satzes gezeigt, daß bezüglich $\Omega = \mathbb{R}^m$ alle Wege mit gleichen Anfangs– und Endpunkten homotop sind.

Eine offene Teilmenge $\Omega \subset \mathbb{R}^m$ heißt *einfach zusammenhängend*, wenn zwei beliebige stetig differenzierbare Kurven in Ω mit gleichen Anfangs– und Endpunkten bezüglich Ω homotop sind (d. h., wenn man die eine Kurve durch „glattes Verbiegen"in die andere überführen kann, ohne dabei Ω zu verlassen und Anfangs– und Endpunkte zu verschieben).

Satz 10.11 *Sei $m = 2$ oder 3 und $\Omega \subset \mathbb{R}^m$ offen und einfach zusammenhängend, $f : \Omega \to \mathbb{R}^m$ stetig differenzierbar. Dann gilt: Das Kurvenintegral von f ist in Ω genau dann wegunabhängig, wenn gilt* $\operatorname{rot} f(x) = 0$ *für alle $x \in \Omega$.*

Beweis. $\Rightarrow$: wie oben (für diese Richtung wird „einfach zusammenhängend" nicht benutzt).

$\Leftarrow$: Läuft ebenfalls wie oben, wenn man für ϕ die durch die Homotopie der beiden Wege gegebene Funktion verwendet. ■

Ein völlig entsprechendes Resultat kann elementar mit nur wenig mehr Aufwand für beliebige Dimension, sogar ohne Verwendung des Satzes von Stokes, direkt bewiesen werden; dabei werden die Resultate dieses und des vorhergehenden Kapitels nicht benötigt. Außer Kenntnisse über Kurvenintegrale wird nur die Differentiation parameterabhängiger Integrale benötigt.

Satz 10.12 *Sei $\Omega \subset \mathbb{R}^m$ offen und einfach zusammenhängend, $f : \Omega \to \mathbb{R}^m$ stetig differenzierbar. Dann gilt: f ist genau dann ein Gradientenfeld (und somit das Kurvenintegral wegunabhängig), wenn die folgende* Integrabilitätsbedingung *erfüllt ist:*

$$f_{j,k} := \frac{\partial}{\partial \xi_k} f_j(x) = f_{k,j} := \frac{\partial}{\partial \xi_j} f_k(x) \text{ für } j, k = 1, \ldots, m,\ x \in A.$$

Beweis. a) Sei $x_0 \in \Omega$ und $K(x_0, r_0)$ eine offene Kugel um x_0, die ganz in Ω liegt. Wenn f tatsächlich der Gradient eines Potentials V ist, dann muß sich V bei vorgegebenem Wert $V(x_0)$ durch das Kurvenintegral von f längs einer beliebigen Kurve von x_0 nach x in Ω berechnen lassen. Wir definieren deshalb $V : K(x_0, r_0) \to \mathbb{R}$ durch

$$V(x) := \int_0^1 \Big\langle f(x_0 + t(x - x_0)), x - x_0 \Big\rangle \,\mathrm{d}t = \int_0^1 \sum_{j=1}^m f_j(x_0 + t(x - x_0))(\xi_j - \xi_{0,j}) \,\mathrm{d}t,$$

d. h. $V(x)$ ist das Kurvenintegral von f (dessen Wegunabhängigkeit wir noch nicht kennen) über die geradlinige Verbindung von x_0 nach x. V ist offensichtlich stetig auf $K(x_0, r)$ und nach Satz 8.17a (Differentiation parameterabhängiger Integrale) gilt

$$\begin{aligned}
\frac{\partial}{\partial \xi_k} V(x) &= \int_0^1 \left\{ \sum_{j=1}^m t f_{j,k}(x_0 + t(x - x_0))(\xi_j - \xi_{0,j}) + f_k(x_0 + t(x - x_0)) \right\} \mathrm{d}t \\
&\quad \text{(Integrabilitätsbedingung: } f_{j,k} = f_{k,j}) \\
&= \int_0^1 \left\{ t \sum_{j=1}^m f_{k,j}(x_0 + t(x - x_0))(\xi_j - \xi_{0,j}) + f_k(x_0 + t(x - x_0)) \right\} \mathrm{d}t \\
&= \int_0^1 \left\{ t \left\langle \operatorname{grad} f_k(x_0 + t(x - x_0)), x - x_0 \right\rangle + f_k(x_0 + t(x - x_0)) \right\} \mathrm{d}t \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \left\{ t f_k(x_0 + t(x - x_0)) \right\} \mathrm{d}t \\
&= \left\{ t f_k(x_0 + t(x - x_0)) \right\} \Big|_{t=0}^{t=1} \;=\; f_k(x).
\end{aligned}$$

Also ist V auf $K(x_0, r_0)$ ein Potential von f; nach Satz 5.8

b) Wir zeigen nun, daß das Kurvenintegral von f über geschlossene Kurven in Ω verschwindet (d. h. es ist in Ω wegunabhängig): Sei Γ eine beliebige geschlossene Kurve in Ω. Da Ω einfach zusammenhängend ist, gibt es eine glatte Fläche M in Ω, deren Randkurve Γ ist (man kann sich M als das Gebilde vorstellen, auf dem Γ stetig differenzierbar zu einem Punkt zusammengezogen werden kann, ohne dabei Ω zu verlassen). Nun denken wir uns die Fläche M in so kleine „Dreiecke“ zerlegt, daß jedes dieser „Dreiecke“ in einer Kugel enthalten ist, die ganz in Ω liegt. Nach Teil a verschwinden die Kurvenintegrale über jedes dieser Dreiecke. Also verschwindet auch das Integral über Γ (da sich Γ als Summe über alle „Dreieckswege“ auffassen läßt, wobei alle innerhalb Γ liegenden Strecken in beiden Richtungen durchlaufen werden, sich also wegheben). ■

10.5 Übungsaufgaben

10.1 Sei S die Oberfläche der Kugel mit Radius R um den Nullpunkt in $\mathbb{R}^3$, $n(x)$ die äußere Normale von S in x. Mit Hilfe des Divergenzsatzes berechne man

$$\int\limits_S \left\langle \left(\xi_1^3, \xi_2^3, \xi_3^3\right), n(x) \right\rangle \mathrm{d}o(x),$$

Ergebnis: $\frac{12}{5}\pi R^5$.

10.2 Mit Hilfe des Stokes'schen Satzes berechne man das Kurvenintegral von

$$k(x) = \left(2\xi_2, -2\xi_1, 2\xi_1\xi_3^2\right)$$

über die Kurve $\gamma : [0, 2\pi] \to \mathbb{R}^3, t \mapsto (\cos t, \sin t, 5)$.
Ergebnis: -8π.

10.3

1. Sei Ω ein Gebiet in der Ebene, γ die Randkurve von Ω (so, daß Ω links von γ liegt), $\gamma : [a, b] \to \mathbb{R}^2$. Dann ist die Fläche $F(\Omega)$ des Gebietes Ω gleich der Hälfte des Kurvenintegrals des Vektorfeldes $k(x, y) = (-y, x)$ über γ, d. h.

$$F(\Omega) = \frac{1}{2} \int\limits_\gamma (-\xi_2, \xi_1)\, \mathrm{d}x = \frac{1}{2} \int\limits_a^b \left\langle (-\gamma_2(t), \gamma_1(t)), \gamma'(t) \right\rangle \mathrm{d}t.$$

2. Man berechne die Fläche unter einer Periode der Zykloide (vgl. § 5 der Vorlesung).
   Ergebnis: 3π.

Literatur

[1] Blatter, C.: Analysis 2. Springer–Lehrbuch, Springer–Verlag, Berlin/Heidelberg/New York/London/Paris/Tokyo/HongKong/Barcelona/Budapest, 1992

[2] Diendonné, J.: Grundzüge der modernen Analysis. Vieweg, Braunschweig/Wiesbaden, 1971

[3] Endl, K. und Luk, W.: Analysis, eine integrierte Darstellung. Aula–Verlag, Wiesbaden 1989

[4] Forster, O.: Analysis 1. Rororo Vieweg, Hamburg 1976

[5] —: Analysis 2. Rororo Vieweg, Habmurg 1977

[6] —: Analysis 3. Friedrich Vieweg u. Sohn, Braunschweig/Wiesbaden 1981

[7] Heuser, H.: Lehrbuch der Analysis, Teil 2. Mathematische Leitfäden, B. G. Teubner, Stuttgart 1981

[8] Königsberger, K.: Analysis 2. Springer–Lehrbuch, Springer–Verlag, Berlin/Heidelberg/New York/London/Paris/Tokyo/HongKong/Barcelona/Budapest, 1993

[9] Walter, W.: Analysis 2, Springer–Lehrbuch, Springer–Verlag, Berlin/Heidelberg/New York/London/Paris/Tokyo/HongKong/Barcelona/Budapest, 1991