Analysis II

Lutz Habermann

Oktober 2003

Inhaltsverzeichnis

Kapitel 1

Differentialrechnung in einer Veränderlichen

1.1 Differenzierbare Abbildungen

Im Folgenden bezeichne $\mathbb{K}$ den Körper $\mathbb{R}$ der reellen oder den Körper $\mathbb{C}$ der komplexen Zahlen und E den Vektorraum $\mathbb{K}^n$. Desweiteren sei $I \subseteq \mathbb{R}$ irgendein Intervall.

Definition 1.1.1 (i) *Eine Abbildung $f : I \to E$ heißt* **differenzierbar in** $a \in I$ $:\Longleftrightarrow$ *Der Grenzwert*

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \tag{1.1.1}$$

existiert. Dieser Grenzwert wird dann die **Ableitung** *oder der* **Differentialquotient** *von f in a genannt und mit $f'(a)$ oder $\frac{\mathrm{d}f}{\mathrm{d}x}(a)$ oder auch $\left.\frac{\mathrm{d}}{\mathrm{d}x} f(x)\right|_{x=a}$ bezeichnet.*

(ii) *Eine Abbildung $f : I \to E$ heißt* **differenzierbar** $:\Longleftrightarrow$ *f ist in jedem $a \in I$ differenzierbar. Die Abbildung*

$$f' : I \to E\,, \quad x \mapsto f'(x)\,,$$

heißt dann die **(erste) Ableitung** *von f.*

Den Grenzwert (1.1.1) schreibt man auch in der Form

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}\,.$$

Bemerkung 1.1.2 (i) Geometrisch kann man die Ableitung folgendermaßen interpretieren. Sei $S_{a,x}$ die Sekante an den Graphen

$$\mathrm{Graph}(f) \subseteq I \times E$$

von $f : I \to E$ durch die Punkte $(a, f(a))$ und $(x, f(x))$ für $a, x \in I$ mit $x \neq a$. Dann ist $S_{a,x} = \mathrm{Graph}(\sigma_{a,x})$, wobei $\sigma_{a,x} : \mathbb{R} \to E$ durch

$$\sigma_{a,x}(t) := \frac{f(x) - f(a)}{x - a}\,(t - a) + f(a)$$

definiert ist. Die Steigung von $S_{a,x}$ ist somit der so genannte **Differenzenquotient**

$$\frac{f(x) - f(a)}{x - a}\,.$$

Ist f in a differenzierbar, so konvergiert dieser Quotient für $x \to a$ gegen die Ableitung $f'(a)$. Die Gerade $T_a := \mathrm{Graph}(\tau_a)$, wobei $\tau_a : \mathbb{R} \to E$ durch

$$\tau_a(t) := f'(a)(t-a) + f(a)$$

gegeben ist, ist dann die **Tangente in** $(a, f(a))$ **an** $\mathrm{Graph}(f)$.

(ii) Die Ableitung wird vielfach zur mathematischen Modellierung zeitabhängiger Größen benutzt. Beschreibe z.B. $s : I \to \mathbb{R}^3$ die Bewegung eines Massepunktes P, d.h. $s(t)$ sei der Ort von P zur Zeit t. Dann ist der Differenzenquotient

$$\frac{s(t) - s(t_0)}{t - t_0}$$

die *mittlere Geschwindigkeit* zwischen den Zeiten t und t_0. Somit gibt der Differentialquotient

$$s'(t_0) = \lim_{t \to t_0} \frac{s(t) - s(t_0)}{t - t_0} \ ,$$

falls er existiert, die *momentane Geschwindigkeit* zur Zeit t_0 an. □

Beispiel 1.1.3 (i) Seien $\mathfrak{a}, \mathfrak{b} \in E$ und sei $f : \mathbb{R} \to E$ durch

$$f(x) := \mathfrak{a}x + \mathfrak{b}$$

definiert. Dann ist f differenzierbar und

$$f'(a) = \mathfrak{a} \quad \text{für alle} \quad a \in \mathbb{R} \ .$$

Tatsächlich ist

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{\mathfrak{a}(x-a)}{x-a} = \mathfrak{a} \ .$$

(ii) Sei $k \in \mathbb{N}$ und sei $f : \mathbb{R} \to \mathbb{R}$ durch

$$f(x) := x^k$$

gegeben. Mit Hilfe des binomischen Lehrsatzes sehen wir, dass

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0} \sum_{j=1}^{k} \binom{k}{j} a^{k-j} h^{j-1} = k a^{k-1} \ .$$

Folglich ist f differenzierbar und

$$f'(a) = k a^{k-1} \quad \text{für alle} \quad a \in \mathbb{R} \ .$$

(iii) Sei $\lambda \in \mathbb{C}$. Bekanntlich gilt

$$\lim_{h \to 0} \frac{\mathrm{e}^{\lambda h} - 1}{h} = \lambda \ .$$

Hieraus folgt für die Abbildung

$$f : \mathbb{R} \to \mathbb{C} \ , \quad f(x) := \mathrm{e}^{\lambda x} \ ,$$

dass

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0} \frac{\mathrm{e}^{\lambda h} - 1}{h} \, \mathrm{e}^{\lambda a} = \lambda \mathrm{e}^{\lambda a} \ .$$

Also ist auch diese Abbildung differenzierbar, wobei

$$f'(a) = \lambda \mathrm{e}^{\lambda a} \quad \text{für alle} \quad a \in \mathbb{R} \ .$$

Insbesondere erhalten wir, dass

$$\exp' = \exp \ .$$

□

Die nächsten beiden Sätze liefern äquivalente Definitionen der Differenzierbarkeit.

Satz 1.1.4 *Eine Abbildung $f : I \to E$ ist genau dann in $a \in I$ differenzierbar, wenn es eine in a stetige Abbildung $\hat{f} : I \to E$ mit*

$$f(x) - f(a) = \hat{f}(x)\,(x-a) \quad \textit{für alle} \quad x \in I \tag{1.1.2}$$

gibt. In diesem Fall ist $f'(a) = \hat{f}(a)$.

Beweis. ($\Longrightarrow$) Ist f in a differenzierbar, so leistet die Abbildung

$$\hat{f} : I \to E\ , \quad \hat{f}(x) := \begin{cases} \dfrac{f(x)-f(a)}{x-a} & \text{für} \quad x \neq a \\ f'(a) & \text{für} \quad x = a \end{cases}\ ,$$

das Gewünschte.

($\Longleftarrow$) Sei $\hat{f} : I \to E$ eine in a stetige Abbildung mit (1.1.2), d.h. mit

$$\hat{f}(x) = \frac{f(x)-f(a)}{x-a} \quad \text{für alle} \quad x \in I \setminus \{a\}\ .$$

Dann ist

$$\lim_{x \to a} \frac{f(x)-f(a)}{x-a} = \lim_{x \to a} \hat{f}(x) = \hat{f}(a)\ .$$

Folglich ist f in a differenzierbar und $\hat{f}(a) = f'(a)$. □

Satz 1.1.5 *Eine Abbildung $f : I \to E$ ist genau dann in $a \in I$ differenzierbar, wenn eine lineare Abbildung $L : \mathbb{R} \to E$ mit*

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - L(h)}{h} = 0 \tag{1.1.3}$$

existiert. In diesem Fall ist

$$L(h) = f'(a)h \quad \textit{für} \quad h \in \mathbb{R}\ . \tag{1.1.4}$$

Beweis. ($\Longrightarrow$) Sei f in a differenzierbar. Dann ist

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} - f'(a) = 0\ .$$

Also erfüllt die durch (1.1.4) definierte lineare Abbildung $L : \mathbb{R} \to E$ die Bedingung (1.1.3).

($\Longleftarrow$) Sei $L : \mathbb{R} \to E$ eine lineare Abbildung mit (1.1.3) und sei $\mathfrak{a} := L(1)$. Dann haben wir

$$L(h) = \mathfrak{a}h \quad \text{für} \quad h \in \mathbb{R}$$

und

$$\begin{aligned} \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} &= \lim_{h \to 0} \frac{f(a+h) - f(a) - \mathfrak{a}h}{h} + \mathfrak{a} \\ &= \lim_{h \to 0} \frac{f(a+h) - f(a) - L(h)}{h} + \mathfrak{a} \\ &= \mathfrak{a}\ . \end{aligned}$$

Folglich ist f in a differenzierbar und $f'(a) = \mathfrak{a}$, womit auch (1.1.4) gilt. □

Definition 1.1.6 *Die durch Satz 1.1.5 bestimmte lineare Abbildung L wird das* **Differential von f in** *a genannt. Die affine Abbildung*

$$x \in \mathbb{R} \mapsto L(x-a) + f(a) \in E$$

heißt **lineare Approximation von f in** *a.*

Das Differential von f in a bezeichnen wir mit $\mathrm{D}f(a)$. Wegen (1.1.4) gilt dann

$$\mathrm{D}f(a)(h) = f'(a)h \quad \text{für} \quad h \in \mathbb{R}\,.$$

Satz 1.1.7 *Ist die Abbildung $f : I \to E$ in $a \in I$ differenzierbar, so ist sie in a stetig.*

Beweis. Sei $f : I \to E$ in a differenzierbar. Dann existiert nach Satz 1.1.4 eine in a stetige Abbildung $\hat{f} : I \to E$ mit

$$f(x) = f(a) + \hat{f}(x)\,(x-a) \quad \text{für alle} \quad x \in I\,.$$

Hieraus folgt, dass auch f in a stetig ist. □

Folgerung 1.1.8 *Ist die Abbildung $f : I \to E$ differenzierbar, so ist sie stetig.* □

Wir bemerken, dass es Funktionen $f : \mathbb{R} \to \mathbb{R}$ gibt, die stetig, aber in keinem $a \in \mathbb{R}$ differenzierbar sind. (Vgl. Königsberger, Analysis 1, Abschnitt 9.11 oder Walter, Analysis 1, Abschnitt 12.26.)

Im Folgenden sind wichtige Rechenregeln für differenzierbare Abbildungen zusammengestellt.

Satz 1.1.9 *Seien $f_1, f_2, f : I \to E$ und $g : I \to \mathbb{K}$ in $a \in I$ differenzierbar. Dann gilt:*

(i) **Summenregel:** *Die Abbildung $f_1 + f_2 : I \to E$ ist in a differenzierbar und*

$$(f_1 + f_2)'\,(a) = f_1'(a) + f_2'(a)\,.$$

(ii) **Produktregel:** *Die Abbildung $fg : I \to E$ ist in a differenzierbar und*

$$(fg)'\,(a) = f'(a)\,g(a) + f(a)\,g'(a)\,.$$

(iii) **Quotientenregel:** *Ist $g(x) \neq 0$ für alle $x \in I$, so ist $f/g : I \to E$ in a differenzierbar und*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)\,g(a) - f(a)\,g'(a)}{(g(a))^2}\,.$$

Beweis. Für die Differenzenquotienten der Abbildungen $f_1 + f_2$, fg und f/g gilt

$$\frac{(f_1+f_2)\,(x) - (f_1+f_2)\,(a)}{x-a} = \frac{f_1(x) - f_1(a)}{x-a} + \frac{f_2(x) - f_2(a)}{x-a}\,,$$

$$\frac{(fg)\,(x) - (fg)\,(a)}{x-a} = \frac{f(x) - f(a)}{x-a}\,g(x) + f(a)\,\frac{g(x) - g(a)}{x-a} \quad \text{und}$$

$$\frac{(f/g)(x) - (f/g)(a)}{x-a} = \left[\frac{f(x) - f(a)}{x-a}\,g(a) - f(a)\,\frac{g(x) - g(a)}{x-a}\right]\frac{1}{g(x)\,g(a)}\,.$$

Hieraus folgen mit Hilfe der Rechenregeln für Grenzwerte von Abbildungen und Satz 1.1.7 die Behauptungen. □

Satz 1.1.10 (Kettenregel) *Seien $I_1, I_2 \subseteq \mathbb{R}$ Intervalle, sei $g : I_1 \to I_2$ in $a \in I_1$ differenzierbar und sei $f : I_2 \to E$ in $g(a)$ differenzierbar. Dann ist $f \circ g : I_1 \to E$ in a differenzierbar und*

$$(f \circ g)'(a) = f'(g(a))\, g'(a) \ .$$

Beweis. Nach Voraussetzung und Satz 1.1.4 existieren eine in a stetige Funktion $\hat{g} : I_1 \to \mathbb{R}$ mit

$$g(x) - g(a) = \hat{g}(x)\,(x-a) \quad \text{für alle} \quad x \in I_1$$

und eine in $g(a)$ stetige Funktion $\hat{f} : I_2 \to E$ mit

$$f(y) - f(g(a)) = \hat{f}(y)\,(y - g(a)) \quad \text{für alle} \quad y \in I_2 \ ,$$

wobei $g'(a) = \hat{g}(a)$ und $f'(g(a)) = \hat{f}(g(a))$. Damit haben wir

$$\begin{aligned}(f \circ g)(x) - (f \circ g)(a) &= f(g(x)) - f(g(a)) \\ &= \hat{f}(g(x))\,(g(x) - g(a)) \\ &= \hat{f}(g(x))\,\hat{g}(x)\,(x-a)\end{aligned}$$

für alle $x \in I_1$. Da die Abbildung $x \in I_1 \mapsto \hat{f}(g(x))\,\hat{g}(x) \in E$ in a stetig ist, folgt wiederum mit Satz 1.1.4, dass $f \circ g$ in a differenzierbar ist. Außerdem erhalten wir, dass

$$(f \circ g)'(a) = \hat{f}(g(a))\,\hat{g}(a) = f'(g(a))\,g'(a) \ .$$

□

Für die Ableitung der inversen Abbildung gilt folgende Regel.

Satz 1.1.11 *Sei $f : I \to \mathbb{R}$ streng monoton und stetig. Ist f in $a \in I$ differenzierbar und ist $f'(a) \neq 0$, so ist $f^{-1} : f(I) \to I$ in $f(a)$ differenzierbar und*

$$\left(f^{-1}\right)'(f(a)) = \frac{1}{f'(a)} \ .$$

Beweis. Wir benutzen wieder Satz 1.1.4. Sei f in a differenzierbar und sei $f'(a) \neq 0$. Dann existiert eine stetige Funktion $\hat{f} : I \to \mathbb{R}$ mit

$$f(x) - f(a) = \hat{f}(x)\,(x-a) \quad \text{für alle} \quad x \in I \ ,$$

wobei $f'(a) = \hat{f}(a)$. Da f streng monoton ist und da $f'(a) \neq 0$, gilt $\hat{f}(x) \neq 0$ für alle $x \in I$. Setzen wir also $x = f^{-1}(y)$, so folgt

$$f^{-1}(y) - f^{-1}(f(a)) = \frac{1}{\hat{f}\left(f^{-1}(y)\right)}\,(y - f(a)) \quad \text{für alle} \quad y \in f(I) \ . \tag{1.1.5}$$

Desweiteren implizieren die Voraussetzungen, dass f^{-1} stetig ist. Folglich ist $\dfrac{1}{\hat{f} \circ f^{-1}}$ in $f(a)$ stetig. Dies und (1.1.5) liefern die Behauptung. □

Bemerkung 1.1.12 Satz 1.1.11 ist ohne die Voraussetzung $f'(a) \neq 0$ falsch. Man betrachte z.B.

$$f : \mathbb{R} \to \mathbb{R} \ , \quad f(x) := x^3 \ .$$

Diese Abbildung ist streng monoton wachsend und differenzierbar, wobei $f'(0) = 0$. Die inverse Abbildung

$$f^{-1} : \mathbb{R} \to \mathbb{R} \ , \quad f^{-1}(x) = \operatorname{sgn}(x)\,|x|^{1/3} \ ,$$

ist in $f(0) = 0$ nicht differenzierbar, denn wegen

$$\frac{f^{-1}(h) - f^{-1}(0)}{h} = \frac{\operatorname{sgn}(h)\,|h|^{1/3}}{\operatorname{sgn}(h)\,|h|} = |h|^{-2/3}$$

existiert der Differentialquotient von f^{-1} in 0 nicht. □

In den nachfolgenden Beispielen berechnen wir die Ableitungen weiterer elementarer Funktionen.

Beispiel 1.1.13 Nach Beispiel 1.1.3 und Satz 1.1.9 ist jedes Polynom

$$p : \mathbb{R} \to E \,, \quad p(x) := \mathfrak{a}_m x^m + \mathfrak{a}_{m-1} x^{m-1} + \cdots + \mathfrak{a}_1 x + \mathfrak{a}_0 \,,$$

mit Koeffizienten $\mathfrak{a}_0, \mathfrak{a}_1, \ldots, \mathfrak{a}_m \in E$ differenzierbar, wobei

$$p'(x) = m\mathfrak{a}_m x^{m-1} + (m-1)\mathfrak{a}_{m-1} x^{m-2} + \cdots + 2\mathfrak{a}_2 x + \mathfrak{a}_1 \quad \text{für alle} \quad x \in \mathbb{R} \,.$$

□

Beispiel 1.1.14 Es gilt

$$\cos'(x) = -\sin(x) \quad \text{und} \quad \sin'(x) = \cos(x) \quad \text{für alle} \quad x \in \mathbb{R} \,.$$

Da

$$\cos(x) = \frac{\exp(\mathrm{i}x) + \exp(-\mathrm{i}x)}{2} \quad \text{und} \quad \sin(x) = \frac{\exp(\mathrm{i}x) - \exp(-\mathrm{i}x)}{2\mathrm{i}} \,,$$

können wir nämlich mit Beispiel 1.1.3(iii) und Satz 1.1.9 schließen, dass

$$\cos'(x) = \frac{\mathrm{i}\exp(\mathrm{i}x) - \mathrm{i}\exp(-\mathrm{i}x)}{2} = -\sin(x)$$

und

$$\sin'(x) = \frac{\mathrm{i}\exp(\mathrm{i}x) + \mathrm{i}\exp(-\mathrm{i}x)}{2\mathrm{i}} = \cos(x) \,.$$

Analog sieht man, dass

$$\cosh'(x) = \sinh(x) \quad \text{und} \quad \sinh'(x) = \cosh(x) \quad \text{für alle} \quad x \in \mathbb{R} \,.$$

□

Beispiel 1.1.15 Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ ist streng monoton wachsend und stetig. Nach Beispiel 1.1.3(iii) haben wir außerdem $\exp'(x) = \exp(x) \neq 0$ für alle $x \in \mathbb{R}$. Wir wenden Satz 1.1.11 auf die inverse Funktion $\ln : \mathbb{R}_+ \to \mathbb{R}$ an und erhalten

$$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{x} \quad \text{für alle} \quad x \in \mathbb{R}_+ \,.$$

□

Beispiel 1.1.16 Sei $a \in \mathbb{R}_+$ und sei $f : \mathbb{R} \to \mathbb{R}$ durch

$$f(x) := a^x$$

definiert. Dann ist

$$f'(x) = \ln(a)\, a^x \quad \text{für alle} \quad x \in \mathbb{R} \,.$$

Wegen $a^x = \exp(\ln(a)\, x)$ gilt nämlich aufgrund der Kettenregel

$$f'(x) = \ln(a) \exp'(\ln(a)\, x) = \ln(a) \exp(\ln(a)\, x) = \ln(a)\, a^x \,.$$

□

Wir erklären jetzt induktiv, was man unter höheren Ableitungen von Funktionen einer reellen Veränderlichen versteht.

Definition 1.1.17 *Sei $k = 2, 3, \ldots$.*

(i) *Eine Abbildung $f : I \to \mathbb{R}$ heißt* **einmal differenzierbar** *$:\iff$ f ist differenzierbar. Die Ableitung von f wird dann auch mit $f^{(1)}$ bezeichnet. Eine Abbildung $f : I \to \mathbb{R}$ heißt* **k-mal differenzierbar** *$:\iff$ f ist $(k-1)$-mal differenzierbar und $f^{(k-1)}$ ist differenzierbar. Die Ableitung von $f^{(k-1)}$ wird dann die* **k-te Ableitung** *von f genannt und mit $f^{(k)}$ bezeichnet.*

(ii) *Eine $(k-1)$-mal differenzierbare Abbildung $f : I \to \mathbb{R}$ heißt* **k-mal differenzierbar in** *$a \in I$ $:\iff$ $f^{(k-1)}$ ist in a differenzierbar. Der Differentialquotient $\dfrac{\mathrm{d}f^{(k-1)}}{\mathrm{d}x}(a)$ heißt dann die* **k-te Ableitung von f in** *a. Er wird mit $f^{(k)}(a)$ oder $\dfrac{\mathrm{d}^k f}{\mathrm{d}x^k}(a)$ bezeichnet.*

Statt $f^{(2)}$ und $f^{(3)}$ schreibt man auch f'' bzw. f'''. Außerdem setzt man $f^{(0)} := f$.

Beispiel 1.1.18 Die Abbildung

$$f : \mathbb{R} \to \mathbb{R}\,, \quad f(x) := \operatorname{sgn}(x)\, x^2\,,$$

ist differenzierbar. Sie ist aber nicht zweimal differenzierbar. Tatsächlich ist

$$f'(x) = 2\,|x| \quad \text{für alle} \quad x \in \mathbb{R}$$

und f' in 0 nicht differenzierbar. □

Satz 1.1.19 *Sei $k \in \mathbb{N}$ und seien $f_1, f_2, f : I \to E$ und $g : I \to \mathbb{K}$ k-mal differenzierbar. Dann gilt:*

(i) *Die Abbildung $f_1 + f_2 : I \to E$ ist k-mal differenzierbar und*

$$(f_1 + f_2)^{(k)} = f_1^{(k)} + f_2^{(k)}\,.$$

(ii) **Leibniz-Regel:** *Die Abbildung $fg : I \to E$ ist k-mal differenzierbar und*

$$(fg)^{(k)} = \sum_{j=0}^{k} \binom{k}{j} f^{(k-j)} g^{(j)}\,.$$

Beweis. Übung. □

Definition 1.1.20 (i) *Sei $k \in \mathbb{N}$. Eine Abbildung $f : I \to E$ heißt* **k-mal stetig differenzierbar** *$:\iff$ f ist k-mal differenzierbar und $f^{(k)}$ ist stetig.*

(ii) *Eine Abbildung $f : I \to E$ heißt* **glatt** *oder* **beliebig oft differenzierbar** *$:\iff$ Für jedes $k \in \mathbb{N}$ ist f k-mal differenzierbar.*

Eine einmal stetig differenzierbare Abbildung $f : I \to E$ nennt man auch einfach **stetig differenzierbar**. Wir bezeichnen die Menge aller k-mal stetig differenzierbaren Abbildungen $f : I \to E$ mit $C^k(I, E)$ und die Menge aller glatten Abbildungen $f : I \to E$ mit $C^\infty(I, E)$. Außerdem sei $C^0(I, E)$ die Menge aller stetigen Abbildungen $f : I \to E$. Offensichtlich tragen alle diese Mengen die Struktur eines $\mathbb{K}$-Vektorraumes. Desweiteren gilt aufgrund von Folgerung 1.1.8

$$C^0(I, E) \supseteq C^1(I, E) \supseteq C^2(I, E) \supseteq \cdots$$

und

$$C^\infty(I, E) = \bigcap_{k \in \mathbb{N}_0} C^k(I, E)\,.$$

Das nächste Beispiel zeigt, dass nicht jede differenzierbare Abbildung auch stetig differenzierbar ist.

Beispiel 1.1.21 Wir betrachten

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

Diese Funktion ist differenzierbar. Da

$$f(x) := \begin{cases} 2x \sin(1/x) - \cos(1/x) & \text{für} \quad x \neq 0 \\ 0 & \text{für} \quad x = 0 \end{cases} \ ,$$

ist aber f' in 0 unstetig und somit $f' \notin C^1(\mathbb{R}, \mathbb{R})$. □

1.2 Differenzierbarkeit und reelle Funktionen

Definition 1.2.1 *Sei (X, d) irgendein metrischer Raum und sei $f : X \to \mathbb{R}$. Man sagt:*

(i) f **hat in** $\xi \in X$ **ein globales Maximum** *(bzw.* **globales Minimum***)* $:\Longleftrightarrow$

$$f(x) \leq f(\xi) \ \ (\textit{bzw. } f(x) \geq f(\xi)) \quad \textit{für alle} \quad x \in X\ .$$

(ii) f **hat in** $\xi \in X$ **ein lokales Maximum** *(bzw.* **lokales Minimum***)* $:\Longleftrightarrow$ *Es existiert eine Umgebung U von ξ mit*

$$f(x) \leq f(\xi) \ \ (\textit{bzw. } f(x) \geq f(\xi)) \quad \textit{für alle} \quad x \in U\ .$$

(iii) f **hat in** $\xi \in X$ **ein isoliertes lokales Maximum** *(bzw.* **isoliertes lokales Minimum***)* $:\Longleftrightarrow$ *Es existiert eine Umgebung U von ξ mit*

$$f(x) < f(\xi) \ \ (\textit{bzw. } f(x) > f(\xi)) \quad \textit{für alle} \quad x \in U \setminus \{0\}\ .$$

Hat $f : X \to \mathbb{R}$ in $\xi \in X$ ein globales Maximum (bzw. globales Minimum), so besitzt f dort offensichtlich auch ein lokales Maximum (bzw. lokales Minimum). Die Umkehrung ist nicht richtig. Es gilt jedoch, dass f genau dann in ξ ein lokales Maximum (bzw. lokales Minimum) hat, wenn es eine solche Umgebung U von ξ gibt, dass $f\big|_U$ in ξ ein globales Maximum (bzw. globales Minimum) hat. Für "globales Maximum" (bzw. "globales Minimum") sagt man auch einfach "Maximum" (bzw. "Minimum").

Satz 1.2.2 *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, sei $\xi \in I$ und sei $f : I \to \mathbb{R}$. Hat f in ξ ein lokales Maximum oder ein lokales Minimum und ist f in ξ differenzierbar, so gilt $f'(\xi) = 0$.*

Beweis. Habe f in ξ ein lokales Maximum. Dann gibt es eine Umgebung $U \subseteq I$ von ξ mit

$$f(x) \leq f(\xi) \quad \text{für alle} \quad x \in U\ .$$

Damit gilt

$$\frac{f(x) - f(\xi)}{x - \xi} \leq 0 \quad \text{für alle} \quad x \in U \cap \,]\xi, +\infty[$$

und

$$\frac{f(x) - f(\xi)}{x - \xi} \geq 0 \quad \text{für alle} \quad x \in U \cap \,]-\infty, \xi[\ .$$

Ist f in ξ differenzierbar, so folgt

$$f'(\xi) = \lim_{x \to \xi+} \frac{f(x) - f(\xi)}{x - \xi} \leq 0 \quad \text{und} \quad f'(\xi) = \lim_{x \to \xi-} \frac{f(x) - f(\xi)}{x - \xi} \geq 0\ ,$$

d.h. $f'(\xi) = 0$.

Hat f in ξ ein lokales Minimum, so verfährt man analog. □

Bemerkung 1.2.3 (i) Aus $f'(\xi) = 0$ folgt nicht zwingend, dass f in ξ ein lokales Extremum hat. Man betrachte z.B. $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^3$. Für diese Funktion ist $f'(0) = 0$, aber f hat in 0 kein lokales Extremum.

(ii) Hat $f : [a, b] \to \mathbb{R}$ in a oder b ein lokales Extremum und ist f dort differenzierbar, so gilt nicht notwendig $f'(a) = 0$ bzw. $f'(b) = 0$. Zum Beispiel hat die Abbildung $f : [a, b] \to \mathbb{R}$, $f(x) := x$, in a ein globales Minimum und in b ein globales Maximum, aber $f'(a) = f'(b) = 1$. □

Satz 1.2.4 (Satz von Rolle) *Sei $f : [a, b] \to \mathbb{R}$ stetig, sei $f\big|_{]a,b[}$ differenzierbar und gelte $f(a) = f(b)$. Dann existiert ein $\xi \in \,]a, b[$ mit $f'(\xi) = 0$.*

Beweis. Da f stetig und $[a, b]$ kompakt ist, existieren nach dem Satz von Weierstrass $\xi_1, \xi_2 \in [a, b]$ derart, dass f in ξ_1 ein globales Minimum und in ξ_2 ein globales Maximum hat.

1. Fall: $\xi_1, \xi_2 \in \{a, b\}$. Wegen $f(a) = f(b)$ haben wir in diesem Fall

$$f(a) = f(\xi_1) \leq f(x) \leq f(\xi_2) = f(a) \quad \text{für alle} \quad x \in [a, b] \ ,$$

d.h.

$$f(x) = f(a) \quad \text{für alle} \quad x \in [a, b] \ .$$

Dies impliziert, dass

$$f'(x) = 0 \quad \text{für alle} \quad x \in \,]a, b[\ .$$

2. Fall: $\xi_1 \in \,]a, b[$ oder $\xi_2 \in \,]a, b[$. Nach Satz 1.2.2 gilt dann

$$f'(\xi_1) = 0 \quad \text{oder} \quad f'(\xi_2) = 0 \ .$$

□

Der Satz von Rolle ist ein Spezialfall von

Satz 1.2.5 (Mittelwertsatz der Differentialrechnung) *Sei $f : [a, b] \to \mathbb{R}$ stetig und sei $f\big|_{]a,b[}$ differenzierbar. Dann existiert ein $\xi \in \,]a, b[$ mit*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \ . \tag{1.2.1}$$

Beweis. Für die Abbildung

$$\varphi : [a, b] \to \mathbb{R} \ , \quad \varphi(x) := f(x) - \frac{f(b) - f(a)}{b - a} (x - a) \ ,$$

gelten die Voraussetzungen von Satz 1.2.4. Somit existiert ein $\xi \in \,]a, b[$ mit $\varphi'(\xi) = 0$, d.h. mit (1.2.1). □

Der Mittelwertsatz bedeutet geometrisch, dass es ein $\xi \in \,]a, b[$ derart gibt, dass die Steigung der Tangente in $(\xi, f(\xi))$ an Graph(f) mit der Steigung der Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ übereinstimmt.

Aus dem Mittelwertsatz leiten wir folgende Monotoniekriterien ab.

Satz 1.2.6 *Sei $I \subseteq \mathbb{R}$ ein Intervall und sei $f : I \to \mathbb{R}$ differenzierbar. Dann gilt:*

(i) *Gilt $f'(x) \geq 0$ für alle $x \in I$, so ist f monoton wachsend.*

(ii) *Gilt $f'(x) \leq 0$ für alle $x \in I$, so ist f monoton fallend.*

(iii) *Gilt $f'(x) > 0$ für alle $x \in I$, so ist f streng monoton wachsend.*

(iv) *Gilt $f'(x) < 0$ für alle $x \in I$, so ist f streng monoton fallend.*

(v) *Gilt $f'(x) = 0$ für alle $x \in I$, so ist f konstant.*

Beweis. Seien $a, b \in I$ mit $a < b$. Indem wir Satz 1.2.5 auf $f\big|_{[a,b]}$ anwenden, erhalten wir, dass

$$f(b) - f(a) = f'(\xi)(b - a) \quad \text{für ein} \quad \xi \in \,]a, b[\,. \tag{1.2.2}$$

Gilt nun $f'(x) \geq 0$ für alle $x \in I$, so folgt wegen $b - a > 0$, dass $f(b) - f(a) \geq 0$, d.h. $f(a) \leq f(b)$. Also ist f in diesem Fall tatsächlich monoton wachsend.

Genauso erhält man aus (1.2.2) die restlichen Behauptungen. □

Bemerkung 1.2.7 Wie man sofort sieht, gelten von den Behauptungen (i), (ii) und (v) des Satzes 1.2.6 auch die Umkehrungen. Für die Behauptungen (iii) und (iv) trifft das nicht zu. Zum Beispiel ist $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^3$, streng monoton wachsend, aber $f'(0) = 0$. □

Die nächsten beiden Sätze liefern hinreichende Kriterien für Extrema.

Satz 1.2.8 *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und sei $\xi \in I$. Desweiteren sei $f : I \to \mathbb{R}$ differenzierbar und $f'(\xi) = 0$. Dann gilt:*

(i) *Ist $f'(x) \leq 0$ für alle $x \in I \cap \,]-\infty, \xi[$ und $f'(x) \geq 0$ für alle $x \in I \cap \,]\xi, +\infty[$, so hat f in ξ ein Minimum.*

(ii) *Ist $f'(x) \geq 0$ für alle $x \in I \cap \,]-\infty, \xi[$ und $f'(x) \leq 0$ für alle $x \in I \cap \,]\xi, +\infty[$, so hat f in ξ ein Maximum.*

Beweis. (i) Laut Voraussetzungen und Satz 1.2.6 ist f auf $I \cap \,]-\infty, \xi]$ monoton fallend und auf $I \cap [\xi, +\infty[$ monoton wachsend. Dies ergibt sofort die Behauptung.

(ii) Analog. □

Satz 1.2.9 *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und sei $\xi \in I$. Sei $f : I \to \mathbb{R}$ differenzierbar und in ξ zweimal differenzierbar. Außerdem sei $f'(\xi) = 0$. Dann gilt:*

(i) *Ist $f''(\xi) < 0$, so hat f in ξ ein isoliertes lokales Maximum.*

(ii) *Ist $f''(\xi) > 0$, so hat f in ξ ein isoliertes lokales Minimum.*

Beweis. (i) Sei $f''(\xi) < 0$. Dann existiert ein $\varepsilon > 0$ derart, dass $]\xi - \varepsilon, \xi + \varepsilon[\subseteq I$ und

$$\frac{f'(x) - f'(\xi)}{x - \xi} < 0 \quad \text{für alle} \quad x \in \,]\xi - \varepsilon, \xi + \varepsilon[\setminus \{0\}\,.$$

Damit gilt $f'(x) > f'(\xi)$ für alle $x \in \,]\xi - \varepsilon, \xi[$ und $f'(x) < f'(\xi)$ für alle $x \in \,]\xi, \xi + \varepsilon[$. Da $f'(\xi) = 0$, folgt mit Satz 1.2.6, dass f auf $]\xi - \varepsilon, \xi[$ streng monoton wachsend und auf $]\xi, \xi + \varepsilon[$ streng monoton fallend ist. Hieraus und aus der Tatsache, dass f in ξ stetig ist, erhalten wir

$$f(x) < f(\xi) \quad \text{für alle} \quad x \in \,]\xi - \varepsilon, \xi + \varepsilon[\setminus \{0\}\,.$$

(ii) Analog. □

Beispiel 1.2.10 Zu gegebenen $a_1, \ldots, a_m \in \mathbb{R}$ soll ein $a \in \mathbb{R}$ so bestimmt werden, dass der Ausdruck

$$(a - a_1)^2 + \cdots + (a - a_m)^2$$

minimal ist. Zu untersuchen ist also, in welchen Punkten die Funktion

$$f : \mathbb{R} \to \mathbb{R} , \quad f(x) := (x - a_1)^2 + \cdots + (x - a_m)^2 ,$$

ein Minimum hat. Solche Punkte müssen nach Satz 1.2.2 Nullstellen von f' sein. Da

$$f'(x) = 2(x - a_1) + \cdots + 2(x - a_m) = 2m \left(x - \frac{a_1 + \cdots + a_m}{m} \right) ,$$

gilt $f'(x) = 0$ genau dann, wenn

$$x = a := \frac{a_1 + \cdots + a_m}{m} ,$$

d.h. wenn x das arithmetische Mittel von $a_1, \ldots, a_m$ ist. Da außerdem $f'(x) < 0$ für $x < a$ und $f'(x) > 0$ für $x > a$, folgt mit Satz 1.2.8, dass f in a tatsächlich minimal ist. $\square$

Bei der Untersuchung des Verlaufs von reellen Funktionen sind neben den Extrema und der Monotonie auch die folgenden Eigenschaften von Bedeutung.

Definition 1.2.11 *Seien $I \subseteq \mathbb{R}$ und $J \subseteq I$ Intervalle und sei $f : I \to \mathbb{R}$.*

(i) *f heißt* **konvex** $:\Longleftrightarrow$

$$\forall\, x_1, x_2 \in I \,\forall\, s \in \,]0, 1[\, : f(sx_1 + (1 - s)x_2) \leq sf(x_1) + (1 - s)f(x_2) .$$

(ii) *f heißt* **konkav** $:\Longleftrightarrow$

$$\forall\, x_1, x_2 \in I \,\forall\, s \in \,]0, 1[\, : f(sx_1 + (1 - s)x_2) \geq sf(x_1) + (1 - s)f(x_2) .$$

(ii) *f heißt* **konvex** *(bzw.* **konkav***)* **auf** $J :\Longleftrightarrow f\big|_J$ *ist konvex (bzw. konkav).*

Geometrisch bedeuten diese Begriffe folgendes. Eine Abbildung $f : I \to \mathbb{R}$ ist genau dann konvex (bzw. konkav), wenn für alle $x_1, x_2 \in I$ die Verbindungsstrecke der Punkte $(x_1, f(x_1))$ und $(x_2, f(x_2))$ oberhalb (bzw. unterhalb) von Graph(f) liegt.

Beispiel 1.2.12 (i) Die Abbildung $f : \mathbb{R} \to \mathbb{R}$, $f(x) := x^2$, ist konvex. Für $x_1, x_2, s \in \mathbb{R}$ gilt nämlich

$$\begin{aligned}
&(sx_1 + (1 - s)x_2)^2 \leq sx_1^2 + (1 - s)x_2^2 \\
&\quad\Longleftrightarrow\quad s^2x_1^2 + 2s(1 - s)x_1x_2 + (1 - s)^2x_2^2 \leq sx_1^2 + (1 - s)x_2^2 \\
&\quad\Longleftrightarrow\quad 0 \leq \left(s - s^2\right) x_1^2 - 2s(1 - s)x_1x_2 + \left((1 - s) - (1 - s)^2\right) x_2^2 \\
&\quad\Longleftrightarrow\quad 0 \leq s(1 - s)(x_1 - x_2)^2 .
\end{aligned}$$

Da $s(1 - s) > 0$ für $s \in \,]0, 1[$, ergibt sich hieraus die gewünschte Abschätzung.

(ii) Die Abbildung $g : [0, +\infty[\, \to \mathbb{R}$, $g(x) := \sqrt{x}$, ist konkav. Dies folgt aus (i) und der Tatsache, dass g monoton wachsend ist.

Satz 1.2.13 *Sei $I \subseteq \mathbb{R}$ ein Intervall und sei $f : I \to \mathbb{R}$ differenzierbar. Dann ist f genau dann konvex (bzw. konkav), wenn f' monoton wachsend (bzw. monoton fallend) ist.*

Beweis. Wir beweisen die Aussage für Konvexität. Für Konkavität schließt man analog.

($\Longrightarrow$) Sei f konvex. Wir fixieren $x_1, x_2 \in I$ mit $x_1 < x_2$ und setzen $x(s) := sx_1 + (1-s)x_2$ für $s \in]0,1[$. Dann gilt

$$f(x(s)) \leq sf(x_1) + (1-s)f(x_2) \; . \tag{1.2.3}$$

Dies ist äquivalent zu

$$s(f(x(s)) - f(x_1)) \leq (1-s)(f(x_2) - f(x(s))) \; ,$$

also auch zu

$$\frac{f(x(s)) - f(x_1)}{(1-s)(x_2 - x_1)} \leq \frac{f(x_2) - f(x(s))}{s(x_2 - x_1)} \; ,$$

d.h. zu

$$\frac{f(x(s)) - f(x_1)}{x(s) - x_1} \leq \frac{f(x_2) - f(x(s))}{x_2 - x(s)} \; . \tag{1.2.4}$$

Die letzte Ungleichung impliziert

$$f'(x_1) = \lim_{s \to 1-} \frac{f(x(s)) - f(x_1)}{x(s) - x_1} \leq \lim_{s \to 1-} \frac{f(x_2) - f(x(s))}{x_2 - x(s)} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

und

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = \lim_{s \to 0+} \frac{f(x(s)) - f(x_1)}{x(s) - x_1} \leq \lim_{s \to 0+} \frac{f(x_2) - f(x(s))}{x_2 - x(s)} = f'(x_2) \; .$$

Folglich ist $f'(x_1) \leq f'(x_2)$.

($\Longleftarrow$) Sei f' monoton wachsend und seien x_1, x_2 und $x(s)$ wie oben. Nach dem Mittelwertsatz existieren $\xi_1 \in]x_1, x(s)[$ und $\xi_2 \in]x(s), x_2[$ mit

$$f'(\xi_1) = \frac{f(x(s)) - f(x_1)}{x(s) - x_1} \quad \text{und} \quad f'(\xi_2) = \frac{f(x_2) - f(x(s))}{x_2 - x(s)} \; .$$

Da $\xi_1 < \xi_2$ und f' monoton wachsend ist, folgt (1.2.4), was, wie bereits gezeigt, zu (1.2.3) äquivalent ist. $\square$

Folgerung 1.2.14 *Sei $I \subseteq \mathbb{R}$ ein Intervall und sei $f : I \to \mathbb{R}$ zweimal differenzierbar. Dann ist f genau dann konvex (bzw. konkav), wenn $f''(x) \geq 0$ (bzw. $f''(x) \leq 0$) für alle $x \in I$.*

Beweis. Nach Satz 1.2.13 ist f genau dann konvex (bzw. konkav), wenn f' monoton wachsend (bzw. monoton fallend) ist, was nach Satz 1.2.6 und Bemerkung 1.2.7 bedeutet, dass $f''(x) \geq 0$ (bzw. $f''(x) \leq 0$) für alle $x \in I$. $\square$

Wie der Beweis des folgenden Satzes illustriert, kann man aus der Konvexität bzw. Konkavität von Funktionen eine Reihe wichtiger Ungleichungen ableiten.

Satz 1.2.15 (Youngsche Ungleichung) *Seien $p, q \in \mathbb{R}_+$ mit $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Dann gilt*

$$x^{1/p} y^{1/q} \leq \frac{x}{p} + \frac{y}{q} \quad \textit{für alle} \quad x, y \in \mathbb{R}_+ \; . \tag{1.2.5}$$

Beweis. Da der natürliche Logarithmus wegen

$$\ln''(x) = -\frac{1}{x^2} < 0 \quad \text{für alle} \quad x \in \mathbb{R}_+$$

nach Folgerung 1.2.14 konkav ist, gilt

$$\frac{\ln(x)}{p} + \frac{\ln(y)}{q} \leq \ln\left(\frac{x}{p} + \frac{y}{q}\right) \quad \text{für alle} \quad x, y \in \mathbb{R}_+ \; .$$

Durch Anwendung der Exponentialfunktion folgt (1.2.5). $\square$

1.3 Verallgemeinerungen des Mittelwertsatzes und die Regel von de l'Hospital

Satz 1.3.1 (Verallgemeinerter Mittelwertsatz der Differentialrechnung) *Seien die Funktionen $f, g : [a, b] \to \mathbb{R}$ stetig, seien $f\big|_{]a,b[}$ und $g\big|_{]a,b[}$ differenzierbar und gelte $g'(x) \neq 0$ für alle $x \in]a, b[$. Dann ist $g(a) \neq g(b)$ und es existiert ein $\xi \in]a, b[$ mit*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)} .$$

Beweis. Wäre $g(a) = g(b)$, so würde nach dem Satz von Rolle ein $\xi \in]a, b[$ mit $g'(\xi) = 0$ existieren. Da das den Voraussetzungen widerspricht, gilt $g(a) \neq g(b)$. Der Rest der Behauptung folgt analog zum Beweis von Satz 1.2.5 durch Betrachtung der Funktion

$$\varphi : [a, b] \to \mathbb{R} , \quad \varphi(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} (g(x) - g(a)) .$$

□

Bemerkung 1.3.2 Offensichtlich erhält man Satz 1.2.5 aus Satz 1.3.1, indem man $g(x) = x$ setzt.

□

Die Bestimmung von $\lim\limits_{x\to a} f(x)/g(x)$ für zwei reelle Funktionen f und g kann sich recht schwierig gestalten, falls $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$. Sind f und g jedoch zusätzlich in a differenzierbar und gilt $g'(a) \neq 0$, so hat man

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{\dfrac{f(x) - f(a)}{x - a}}{\dfrac{g(x) - g(a)}{x - a}} \to \frac{f'(a)}{g'(a)} \quad \text{für} \quad x \to a .$$

Eine Verallgemeinerung dieser Überlegung ist

Satz 1.3.3 (Regel von de l'Hospital) *Seien $f, g :]a, b[\to \mathbb{R}$ differenzierbar und sei $g'(x) \neq 0$ für alle $x \in]a, b[$. Außerdem gelte einer der beiden folgenden Fälle:*

(a) *$f(x) \to 0$ und $g(x) \to 0$ für $x \to a+$,*

(b) *$f(x) \to +\infty$ und $g(x) \to +\infty$ für $x \to a+$.*

Existiert dann $\lim\limits_{x\to a+} f'(x)/g'(x)$, so existiert auch $\lim\limits_{x\to a+} f(x)/g(x)$ und

$$\lim_{x\to a+} \frac{f(x)}{g(x)} = \lim_{x\to a+} \frac{f'(x)}{g'(x)} .$$

Beweis. Es gelte der Fall (a). Wir setzen f und g durch $f(a) := 0$ und $g(a) := 0$ zu stetigen Funktionen auf $[a, b[$ fort. Nach Satz 1.3.1 gibt es dann zu jedem $x \in]a, b[$ ein $\xi(x) \in]a, x[$ mit

$$\frac{f(x)}{g(x)} = \frac{f'(\xi(x))}{g'(\xi(x))} .$$

Da mit $x \to a+$ auch $\xi(x) \to a+$ gilt, folgt hieraus die Behauptung.

Es gelte jetzt der Fall (b). Sei

$$\alpha := \lim_{x \to a+} \frac{f'(x)}{g'(x)}$$

und sei $\varepsilon > 0$ beliebig. Dann existiert ein $c \in]a, b[$ derart, dass

$$\left| \frac{f'(x)}{g'(x)} - \alpha \right| < \varepsilon \quad \text{für alle} \quad x \in]a, c[\ .$$

Nach Satz 1.3.1 gilt damit

$$\left| \frac{f(x) - f(c)}{g(x) - g(c)} - \alpha \right| < \varepsilon \quad \text{für alle} \quad x \in]a, c[\ . \tag{1.3.1}$$

Wegen $f(x) \to +\infty$ und $g(x) \to +\infty$ für $x \to a+$, können wir o.B.d.A. annehmen, dass

$$f(x) \neq 0 \quad \text{und} \quad g(x) \neq 0 \quad \text{für} \quad x \in]a, c]\ .$$

Somit haben wir

$$\frac{f(x)}{g(x)} - \frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f(x) - f(c)}{g(x) - g(c)} \left(\frac{1 - g(c)/g(x)}{1 - f(c)/f(x)} - 1 \right) \quad \text{für} \quad x \in]a, c]\ .$$

Da nach (1.3.1)

$$\left| \frac{f(x) - f(c)}{g(x) - g(c)} \right| < \alpha + \varepsilon \quad \text{für alle} \quad x \in]a, c[$$

und da

$$\lim_{x \to a+} \frac{1 - g(c)/g(x)}{1 - f(c)/f(x)} = 1\ ,$$

existiert folglich ein $d \in]a, c]$ mit

$$\left| \frac{f(x)}{g(x)} - \frac{f(x) - f(c)}{g(x) - g(c)} \right| < \varepsilon \quad \text{für alle} \quad x \in]a, d[\ . \tag{1.3.2}$$

Aus (1.3.1) und (1.3.2) erhalten wir mittels Dreiecksungleichung

$$\left| \frac{f(x)}{g(x)} - \alpha \right| < 2\varepsilon \quad \text{für alle} \quad x \in]a, d[\ .$$

Damit ist die Behauptung auch für den Fall (b) bewiesen. □

Satz 1.3.4 *Seien $f, g :]a, +\infty[\to \mathbb{R}$ differenzierbar, sei $g'(x) \neq 0$ für alle $x \in]a, +\infty[$ und gelte*

$$\lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} g(x) = 0 \quad \textit{oder} \quad \lim_{x \to +\infty} f(x) = \lim_{x \to +\infty} g(x) = +\infty\ .$$

Existiert $\lim_{x \to +\infty} f'(x)/g'(x)$, so existiert auch $\lim_{x \to +\infty} f(x)/g(x)$ und

$$\lim_{x \to +\infty} \frac{f(x)}{g(x)} = \lim_{x \to +\infty} \frac{f'(x)}{g'(x)}\ .$$

Beweis. Wir setzen voraus, dass $\lim_{x \to +\infty} f'(x)/g'(x)$ existiert, nehmen o.B.d.A. $a > 0$ an und definieren $f_1, g_1 :]0, 1/a[\to \mathbb{R}$ durch

$$f_1(y) := f\left(\frac{1}{y}\right) \quad \text{und} \quad g_1(y) := g\left(\frac{1}{y}\right)\ .$$

Da dann

$$f_1'(y) = -\frac{1}{y^2} f'\left(\frac{1}{y}\right) \quad \text{und} \quad g_1'(y) = -\frac{1}{y^2} g'\left(\frac{1}{y}\right)$$

und somit

$$\lim_{y \to 0+} \frac{f_1'(y)}{g_1'(y)} = \lim_{y \to 0+} \frac{f'(1/y)}{g'(1/y)} = \lim_{x \to +\infty} \frac{f'(x)}{g'(x)}\ ,$$

erhält man die Behauptung, indem man Satz 1.3.3 auf f_1 und g_1 für $y \to 0+$ anwendet. □

Bemerkung 1.3.5 Offensichtlich gelten die Sätze 1.3.3 und 1.3.4 für $x \to b-$ bzw. $x \to -\infty$ in analoger Form. Damit kann die Regel von de l'Hospital auch zur Berechnung von Grenzwerten $\lim\limits_{x\to\xi} f(x)/g(x)$ benutzt werden. Außerdem kann man leicht einsehen, dass die oben genannten Sätze auch gültig bleiben, wenn $f'(x)/g'(x)$ gegen $\pm\infty$ konvergiert. Dazu ist lediglich der Beweis von Satz 1.3.3 für den Fall (b) zu modifizieren (vgl. Walter, Analysis 1, Abschnitt 10.11). □

Beispiel 1.3.6 (i) Es ist

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} = \lim_{x\to 0} \frac{(1+x)^{-1}}{1} = 1$$

und

$$\lim_{x\to+\infty} \frac{\ln(x)}{x} = \lim_{x\to+\infty} \frac{x^{-1}}{1} = 0 \ .$$

(ii) Zur Berechnung von $\lim\limits_{x\to 0+} x\ln(x)$ kann man

$$\lim_{x\to 0+} x\ln(x) = \lim_{x\to 0+} \frac{\ln(x)}{x^{-1}} = \lim_{x\to 0+} \frac{x^{-1}}{-x^{-2}} = -\lim_{x\to 0+} x = 0$$

schließen.

(iii) Es ist

$$\lim_{x\to 0} \cos(x)^{1/x} = 1 \ ,$$

denn

$$\lim_{x\to 0} \ln\left(\cos(x)^{1/x}\right) = \lim_{x\to 0} \frac{\ln(\cos(x))}{x} = -\lim_{x\to 0} \frac{\sin(x)}{\cos(x)} = 0 \ ,$$

woraus sich aufgrund der Stetigkeit der Exponentialfunktion

$$\lim_{x\to 0} \cos(x)^{1/x} = \lim_{x\to 0} \exp\left(\ln\left(\cos(x)^{1/x}\right)\right) = \exp\left(\lim_{x\to 0} \ln\left(\cos(x)^{1/x}\right)\right) = \exp(0) = 1$$

ergibt. □

Ein Anwendungsbereich der Regel von de l'Hospital ist die Untersuchung des asymptotischen Vehaltens von Funktionen.

Definition 1.3.7 *Sei $M \subseteq \mathbb{R}$ mit $\sup M = +\infty$. Zwei Funktionen $f, g : M \to \mathbb{K}$ heißen* **asymptotisch gleich für** $x \to +\infty$ *(in Zeichen: $f(x) \sim g(x)$ für $x \to +\infty$)* $:\Longleftrightarrow$

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = 1 \ .$$

Analog für $x \to -\infty$.

Beispiel 1.3.8 Wir zeigen, dass

$$\mathrm{e} - \left(1 + \frac{1}{x}\right)^x \sim \frac{\mathrm{e}}{2\,x} \quad \text{für} \quad x \to +\infty \ .$$

Dazu substituieren wir $x = 1/y$. Es ist dann zu verifizieren, dass

$$\lim_{y\to 0+} \frac{\mathrm{e} - (1+y)^{1/y}}{y} = \frac{\mathrm{e}}{2} \ .$$

Diesen Grenzwert erhält man durch zweimalige Anwendung von Satz 1.3.3. Es ist nämlich

$$\begin{aligned}\lim_{y\to 0+}\frac{\mathrm{e}-(1+y)^{1/y}}{y} &= -\lim_{y\to 0+}(1+y)^{1/y}\frac{(1+y)^{-1}y-\ln(1+y)}{y^2}\\ &= \mathrm{e}\lim_{y\to 0+}\frac{\ln(1+y)-(1+y)^{-1}y}{y^2}\end{aligned}$$

und

$$\lim_{y\to 0+}\frac{\ln(1+y)-(1+y)^{-1}y}{y^2} = \lim_{y\to 0+}\frac{(1+y)^{-2}}{2} = \frac{1}{2}\,.$$

□

Beispiel 1.3.8 zeigt u.a., dass die Konvergenz von $(1+1/x)^x$ gegen e für $x \to +\infty$ “langsam” ist.

Wir wollen jetzt der Frage nachgehen, inwieweit man den Mittelwertsatz für komplex- und vektorwertige Funktionen, also für Funktionen $f : I \to \mathbb{C}$ bzw. $f : I \to E$ verallgemeinern kann. Dabei sei hier und im Folgenden wieder $E = \mathbb{K}^n$.

Beispiel 1.3.9 Die Abbildung $f : [0, 2\pi] \to \mathbb{C}$, $f(x) := \mathrm{e}^{\mathrm{i}x}$, ist differenzierbar und es gilt $f(0) = f(2\pi) = 1$. Es ist aber $f'(x) \neq 0$ für alle $x \in [0, 2\pi]$, denn $f'(x) = \mathrm{i}\mathrm{e}^{\mathrm{i}x}$. □

Also gilt Satz 1.2.5 nicht für komplex- und vektorwertige Funktionen. Man kann jedoch folgende Abschätzung beweisen, die für Funktionen $f : I \to \mathbb{R}$ sofort aus Satz 1.2.5 folgt.

Satz 1.3.10 *Sei $f : [a, b] \to E$ stetig, sei $f\big|_{]a,b[}$ differenzierbar und gelte*

$$\|f'(x)\| \le d \quad \textit{für alle} \quad x \in\,]a, b[\tag{1.3.3}$$

mit einem $d \in \mathbb{R}$. Dann ist

$$\|f(b) - f(a)\| \le d(b - a)\,. \tag{1.3.4}$$

Beweis. Seien $\varepsilon > 0$ und $y \in\,]a, b[$ beliebig und sei M die Menge aller $\xi \in [y, b]$ mit

$$\|f(x) - f(y)\| \le (d + \varepsilon)(x - y) \quad \text{für} \quad y \le x \le \xi\,.$$

Dann ist M beschränkt und nichtleer, denn $M \subseteq [y, b]$ und $y \in M$. Also existiert $c := \sup M$. Da f stetig ist, ist $c \in M$. Somit gilt

$$\|f(x) - f(y)\| \le (d + \varepsilon)(x - y) \quad \text{für} \quad y \le x \le c\,. \tag{1.3.5}$$

Wir wollen zeigen, dass $c = b$. Dazu nehmen wir an, dass $c < b$. Aufgrund der Differenzierbarkeit von $f\big|_{]a,b[}$ können wir dann ein $\delta > 0$ so wählen, dass $c + \delta \le b$ und

$$\left\|\frac{f(x) - f(c)}{x - c} - f'(c)\right\| \le \varepsilon \quad \text{für alle} \quad x \in\,]c, c + \delta]\,,$$

d.h.

$$\|f(x) - f(c) - f'(c)(x - c)\| \le \varepsilon(x - c) \quad \text{für alle} \quad x \in\,]c, c + \delta]\,.$$

Da wegen (1.3.3)

$$\|f(x) - f(c)\| - d(x - c) \le \|f(x) - f(c)\| - \|f'(c)\|\,(x - c) \le \|f(x) - f(c) - f'(c)(x - c)\|\,,$$

für $x \in [c, b]$, folgt

$$\|f(x) - f(c)\| \le (d + \varepsilon)(x - c) \quad \text{für alle} \quad x \in\,]c, c + \delta]\,.$$

Wir schließen weiter, dass

$$\begin{aligned}\|f(x)-f(y)\| &\le \|f(x)-f(c)\| + \|f(c)-f(y)\| \\ &\le (d+\varepsilon)(x-c) + (d+\varepsilon)(c-y) \\ &= (d+\varepsilon)(x-y)\end{aligned}$$

für alle $x \in]c, c+\delta]$. Dies und (1.3.5) implizieren, dass $c+\delta \in M$, was im Widerspruch zu $c = \sup M$ steht. Also ist $c = b$.

Damit haben wir gezeigt, dass

$$\|f(b)-f(y)\| \le (d+\varepsilon)(b-y) \quad \text{für alle} \quad \varepsilon > 0 \text{ und } y \in]a, b[\ .$$

Wegen der Stetigkeit von f folgt

$$\|f(b)-f(a)\| \le (d+\varepsilon)(b-a) \quad \text{für alle} \quad \varepsilon > 0\ .$$

Die letzte Aussage ist äquivalent zu (1.3.4) und der Satz ist bewiesen. □

Satz 1.3.10 wird gelegentlich auch als Mittelwertsatz zitiert. Als Konsequenzen dieses Satzes haben wir:

Satz 1.3.11 *Sei $f : [a, b] \to E$ stetig differenzierbar. Dann ist f Lipschitz-stetig.*

Beweis. Da $f' : [a, b] \to E$ stetig und $[a, b]$ kompakt ist, ist auch $f'([a, b])$ kompakt und damit insbesondere beschränkt. Also existiert ein $\alpha \in \mathbb{R}$ mit

$$\|f'(x)\| \le \alpha \quad \text{für alle} \quad x \in [a, b]\ .$$

Durch Anwendung von Satz 1.3.10 erhält man

$$\|f(x_1)-f(x_2)\| \le \alpha|x_1 - x_2| \quad \text{für alle} \quad x_1, x_2 \in [a, b]\ .$$

□

Satz 1.3.12 *Sei $I \subseteq \mathbb{R}$ ein Intervall, sei $f : I \to E$ differenzierbar und gelte $f'(x) = 0$ für alle $x \in I$. Dann ist f konstant.*

Beweis. Nach Voraussetzung gilt

$$\|f'(x)\| = 0 \quad \text{für alle} \quad x \in I\ .$$

Mit Satz 1.3.10 folgt

$$\|f(x_1)-f(x_2)\| \le 0 \quad \text{für alle} \quad x_1, x_2 \in I\ ,$$

d.h.

$$f(x_1) = f(x_2) \quad \text{für alle} \quad x_1, x_2 \in I\ .$$

□

Satz 1.3.12 verallgemeinert Satz 1.2.6(v). Als Anwendungen haben wir:

Folgerung 1.3.13 *Sei $I \subseteq \mathbb{R}$ ein Intervall, sei $f : I \to E$ differenzierbar und gelte*

$$f'(x) = \mathfrak{a} \quad \textit{für alle} \quad x \in I$$

mit einem $\mathfrak{a} \in E$. Dann ist

$$f(x) = \mathfrak{a}x + \mathfrak{b} \quad \textit{für alle} \quad x \in I \tag{1.3.6}$$

mit einem $\mathfrak{b} \in E$.

Beweis. Die Abbildung $\varphi : I \to E$, $\varphi(x) := f(x) - \mathfrak{a}x$, erfüllt die Voraussetzungen von Satz 1.3.12. Folglich gibt es ein $\mathfrak{b} \in E$ mit $\varphi(x) = \mathfrak{b}$ für alle $x \in I$, d.h. mit (1.3.6). $\square$

Folgerung 1.3.14 *Sei $\lambda \in \mathbb{C}$, sei $f : \mathbb{R} \to \mathbb{C}$ differenzierbar und gelte*

$$f'(x) = \lambda f(x) \quad \textit{für alle} \quad x \in \mathbb{R}.$$

Dann ist

$$f(x) = f(0)\mathrm{e}^{\lambda x} \quad \textit{für alle} \quad x \in \mathbb{R}.$$

Beweis. Man betrachte $\varphi : \mathbb{R} \to \mathbb{C}$, $\varphi(x) := f(x)\mathrm{e}^{-\lambda x}$, und nutze wieder Satz 1.3.12. $\square$

1.4 Differentiation von Funktionenfolgen und -reihen

Wir beginnen mit zwei Beispielen.

Beispiel 1.4.1 Seien $f_k : \mathbb{R} \to \mathbb{R}$, $k \in \mathbb{N}$, durch

$$f_k(x) := \sqrt{x^2 + \frac{1}{k}}$$

definiert. Dann sind alle f_k differenzierbar, die Folge (f_k) ist gleichmäßig konvergent, aber deren Grenzfunktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) := |x|$, ist nicht differenzierbar. $\square$

Beispiel 1.4.2 Seien $f_k : \mathbb{R} \to \mathbb{R}$, $k \in \mathbb{N}$, durch

$$f_k(x) := \frac{1}{k}\sin(kx)$$

definiert. Dann sind alle f_k differenzierbar, die Folge (f_k) konvergiert gleichmäßig gegen 0, aber (f_k') ist nicht mal punktweise konvergent, denn $f_k'(x) = \cos(kx)$ und folglich ist z.B. $(f_k'(\pi))$ divergent. $\square$

Sei wieder $E = \mathbb{K}^n$.

Satz 1.4.3 *Seien $f_k : [a, b] \to E$, $k \in \mathbb{N}$, differenzierbar und gelte:*

(1) *Es existiert ein $x_0 \in [a, b]$ derart, dass $(f_k(x_0))$ konvergent ist.*

(2) *Die Folge (f_k') ist gleichmäßig konvergent.*

Dann konvergiert die Folge (f_k) gleichmäßig gegen eine differenzierbare Abbildung $f : [a, b] \to E$ und

$$f'(x) = \lim_{k\to\infty} f_k'(x) \quad \textit{für alle} \quad x \in [a, b]\ . \tag{1.4.1}$$

Beweis. Sei $\varepsilon > 0$. Nach Eigenschaft (1) können wir ein $x_0 \in [a, b]$ so wählen, dass $(f_k(x_0))$ konvergent und somit eine Cauchy-Folge ist. Dann existiert ein $k_1 \in \mathbb{N}$ mit

$$\|f_k(x_0) - f_l(x_0)\| < \frac{\varepsilon}{2} \quad \text{für alle} \quad k, l \geq k_1\ . \tag{1.4.2}$$

Wegen der Eigenschaft (2) existiert außerdem ein $k_2 \in \mathbb{N}$ mit

$$\|f_k'(x) - f_l'(x)\| < \frac{\varepsilon}{2(b-a)} \quad \text{für alle} \quad x \in [a, b] \text{ und } k, l \geq k_2\ . \tag{1.4.3}$$

Indem wir Satz 1.3.10 auf $f_k - f_l$ anwenden, erhalten wir aus (1.4.3), dass

$$\|f_k(x) - f_l(x) - f_k(\xi) - f_l(\xi)\| \le \frac{\varepsilon}{2(b-a)} |x - \xi| \quad \text{für alle} \quad x, \xi \in [a,b] \text{ und } k, l \ge k_2 , \tag{1.4.4}$$

woraus sich weiter

$$\|f_k(x) - f_l(x) - f_k(\xi) - f_l(\xi)\| \le \frac{\varepsilon}{2} \quad \text{für alle} \quad x, \xi \in [a,b] \text{ und } k, l \ge k_2 \tag{1.4.5}$$

ergibt. Sei $k_0 := \max\{k_1, k_2\}$. Da

$$\|f_k(x) - f_l(x)\| \le \|f_k(x) - f_l(x) - f_k(x_0) - f_l(x_0)\| + \|f_k(x_0) - f_l(x_0)\| ,$$

liefern (1.4.2) und (1.4.5), dass

$$\|f_k(x) - f_l(x)\| < \varepsilon \quad \text{für alle} \quad x \in [a,b] \text{ und } k, l \ge k_0 . \tag{1.4.6}$$

Folglich ist $(f_k(x))$ für jedes $x \in [a,b]$ eine Cauchy-Folge. Da E vollständig ist, ist dann $(f_k(x))$ für jedes $x \in [a,b]$ konvergent. Wir definieren $f : [a,b] \to E$ durch

$$f(x) := \lim_{k\to\infty} f_k(x) .$$

Aus (1.4.6) folgt durch Bildung des Grenzwertes für $l \to \infty$, dass

$$\|f_k(x) - f(x)\| \le \varepsilon \quad \text{für alle} \quad x \in [a,b] \text{ und } k \ge k_0 .$$

Also konvergiert (f_k) gleichmäßig gegen f.

Es bleibt zu zeigen, dass f differenzierbar ist und (1.4.1) gilt. Wir fixieren ein $\xi \in [a,b]$ und definieren $\varphi_k : [a,b] \to E$, $k \in \mathbb{N}$, durch

$$\varphi_k(x) := \begin{cases} \dfrac{f_k(x) - f_k(\xi)}{x - \xi} & \text{für} \quad x \neq \xi \\ f_k'(\xi) & \text{für} \quad x = \xi \end{cases} .$$

Dann sind alle φ_k stetig. Außerdem gilt wegen (1.4.4) für $x \neq \xi$ und $k, l \ge k_2$

$$\|\varphi_k(x) - \varphi_l(x)\| \le \frac{1}{|x-\xi|} \|f_k(x) - f_l(x) - f_k(\xi) + f_l(\xi)\| \le \frac{\varepsilon}{2(b-a)} .$$

Aufgrund der Setigkeit von $\varphi_k - \varphi_l$ gilt diese Abschätzung auch für $x = \xi$. Also haben wir

$$\|\varphi_k(x) - \varphi_l(x)\| \le \frac{\varepsilon}{2(b-a)} \quad \text{für alle} \quad x \in [a,b] \text{ und } k, l \ge k_2 .$$

Damit ist $(\varphi_k(x))$ für jedes $x \in [a,b]$ eine Cauchy-Folge und wie oben sieht man, dass (φ_k) gleichmäßig gegen die Abbildung $\varphi : [a,b] \to E$, $\varphi(x) := \lim\limits_{k\to\infty} \varphi_k(x)$, konvergiert. Hieraus und aus der Stetigkeit der φ_k folgt, dass auch φ stetig ist. Somit gilt

$$\varphi(\xi) = \lim_{x\to\xi} \varphi(x) = \lim_{x\to\xi} \lim_{k\to\infty} \varphi_k(x) = \lim_{x\to\xi} \lim_{k\to\infty} \frac{f_k(x) - f_k(\xi)}{x - \xi} = \lim_{x\to\xi} \frac{f(x) - f(\xi)}{x - \xi} .$$

Also ist f in ξ differenzierbar und

$$f'(\xi) = \varphi(\xi) = \lim_{k\to\infty} \varphi_k(\xi) = \lim_{k\to\infty} f_k'(\xi) .$$

$\square$

Eine im wesentlichen nur andere Formulierung von Satz 1.4.3 ist

Satz 1.4.4 *Seien $f_k : [a,b] \to E$, $k \in \mathbb{N}$, differenzierbar und gelte:*

(1) *Es existiert ein $x_0 \in [a,b]$ derart, dass $\sum_{k=1}^{\infty} f_k(x_0)$ konvergent ist.*

(2) *Die Reihe $\sum_{k=1}^{\infty} f_k'$ ist gleichmäßig konvergent.*

Dann konvergiert die Reihe $\sum_{k=1}^{\infty} f_k$ gleichmäßig gegen eine differenzierbare Abbildung $f : [a,b] \to E$ und

$$f'(x) = \sum_{k=1}^{\infty} f_k'(x) \quad \textit{für alle} \quad x \in [a,b] \,. \tag{1.4.7}$$

□

Die Beziehungen (1.4.1) und (1.4.7) kann man auch in der Form

$$\frac{\mathrm{d}}{\mathrm{d}x} \lim_{k\to\infty} f_k = \lim_{k\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_k \qquad \text{bzw.} \qquad \frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=1}^{\infty} f_k = \sum_{k=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_k$$

schreiben, also als Vertauschungsregel zwischen Grenzwert von Folgen bzw. Reihen und Ableitung. Als Anwendung von Satz 1.4.4 haben wir

Satz 1.4.5 *Sei $f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k$ eine Potenzreihe mit Zentrum $x_0 \in \mathbb{R}$, Koeffizienten $a_k \in \mathbb{K}$ und Konvergenzradius $\rho > 0$. Dann ist $f : \,]x_0 - \rho, x_0 + \rho[\, \to \mathbb{K}$ differenzierbar und*

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1} \quad \textit{für alle} \quad x \in \,]x_0 - \rho, x_0 + \rho[\,.$$

Beweis. Sei $\delta \in \,]0, \rho[$ und seien $f_k : [x_0 - \delta, x_0 + \delta] \to \mathbb{K}$, $k \in \mathbb{N}$, durch

$$f_k(x) := a_k(x - x_o)^k$$

definiert. Die Abbildungen f_k sind offensichtlich differenzierbar, wobei

$$f_0'(x) = 0 \qquad \text{und} \qquad f_k'(x) = k a_k (x - x_0)^{k-1} \quad \text{für} \quad k \in \mathbb{N} \,.$$

Außerdem ist $\sum_{k=0}^{\infty} f_k(x_0)$ konvergent. Sei ρ' der Konvergenzradius von $\sum_{k=1}^{\infty} k a_k (x-x_0)^{k-1}$. Dann ist $\rho' = \rho$, denn die Konvergenzradien von $\sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}$ und $\sum_{k=0}^{\infty} k a_k (x - x_0)^k$ stimmen überein und folglich gilt

$$\rho' = \left(\limsup_{k\to\infty} \sqrt[k]{|k a_k|} \right)^{-1} = \left(\limsup_{k\to\infty} \sqrt[k]{k} \sqrt[k]{|a_k|} \right)^{-1} = \left(\limsup_{k\to\infty} \sqrt[k]{|a_k|} \right)^{-1} = \rho \,.$$

Damit konvergiert $\sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}$ auf $[x_0 - \delta, x_0 + \delta]$ gleichmäßig, d.h. $\sum_{k=1}^{\infty} f_k'$ ist gleichmäßig konvergent. Mit Satz 1.4.4 folgt, dass $f\big|_{[x_0 - \delta, x_0 + \delta]}$ differenzierbar ist und

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1} \quad \text{für alle} \quad x \in \,]x_0 - \delta, x_0 + \delta[\,.$$

Da dies für alle $\delta \in \,]0, \rho[$ gilt, ist damit der Satz bewiesen. □

Beispiel 1.4.6 Die so genannte **Zeta-Funktion** $\zeta :]1,+\infty[\to \mathbb{R}$ ist durch

$$\zeta(s) := \sum_{k=1}^{\infty} \frac{1}{k^s}$$

definiert. Wir wollen zeigen, dass diese Funktion differenzierbar ist und

$$\zeta'(s) = -\sum_{k=2}^{\infty} \frac{\ln(k)}{k^s} \quad \text{für alle} \quad s \in]1,+\infty[\, .$$

Dazu fixieren wir ein $\delta > 0$ und ein $\varepsilon \in]0,\delta[$. Wir setzen $a := \delta - \varepsilon$ und definieren $g : [1,+\infty[\to \mathbb{R}$ durch

$$g(x) := \frac{\ln(x)}{x^a} \, .$$

Da $a > 0$ und

$$g'(x) = \frac{1 - a\ln(x)}{x^{a+1}} \quad \text{für alle} \quad x \in [1,+\infty[\, ,$$

hat g ein Maximum und ist somit nach oben beschränkt. Also gibt es ein $c \in \mathbb{R}_+$ mit

$$\frac{\ln(x)}{x^a} \le c \quad \text{für alle} \quad x \in [1,+\infty[\, .$$

Damit schließen wir, dass für alle $s \in [1+\delta,+\infty[$ und alle $k \in \mathbb{N}$

$$\left|\frac{\ln(k)}{k^s}\right| = \frac{\ln(k)}{k^s} \le \frac{\ln(k)}{k^{1+\delta}} = \frac{\ln(k)}{k^a} \cdot \frac{1}{k^{1+\varepsilon}} \le \frac{c}{k^{1+\varepsilon}} \, .$$

Diese Abschätzung impliziert, dass die Reihe $\sum\limits_{k=2}^{\infty} \dfrac{\ln(k)}{k^s}$ auf dem Intervall $[1+\delta,+\infty[$ gleichmäßig konvergiert. Mit Satz 1.4.4 erhalten wir, dass $\zeta\big|_{[1+\delta,+\infty[}$ differenzierbar ist und

$$\zeta'(s) = -\sum_{k=2}^{\infty} \frac{\ln(k)}{k^s} \quad \text{für alle} \quad s \in]1+\delta,+\infty[\, .$$

Da dies für alle $\delta > 0$ gilt, folgt die gewünschte Aussage. $\square$

Beispiel 1.4.7 Wir entwickeln $\arctan : \mathbb{R} \to]-\pi/2,\pi/2[$ um 0 in eine Potenzreihe. Da

$$\arctan'(x) = \frac{1}{1+x^2} \quad \text{für alle} \quad x \in \mathbb{R} \, ,$$

haben wir

$$\arctan'(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} \quad \text{für alle} \quad x \in]-1,1[\, .$$

Für die Potenzreihe $g(x) = \sum\limits_{k=0}^{\infty} \dfrac{(-1)^k}{2k+1}\, x^{2k+1}$ gilt nach Satz 1.4.5 ebenfalls

$$g'(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} \quad \text{für alle} \quad x \in]-1,1[\, .$$

Da außerdem $\arctan(0) = g(0)$, folgt

$$\arctan(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}\, x^{2k+1} \quad \text{für alle} \quad x \in]-1,1[\, .$$

Diese Reihenentwicklung liefert zusammen mit der so genannten **Machinschen Formel**

$$\frac{\pi}{4} = 4\arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right)$$

gute Näherungen für π (vgl. Königsberger, Analysis 1, Abschnitt 8.11). $\square$

1.5 Die Taylor-Entwicklung

Bezeichne wieder I ein Intervall in $\mathbb{R}$.

Definition 1.5.1 *Sei $x_0 \in I$ und sei $f : I \to \mathbb{K}$ k-mal differenzierbar. Dann heißt*

$$T_k(f,x_0) : \mathbb{R} \to \mathbb{K}\ , \quad T_k(f,x_0)(x) := \sum_{j=0}^{k} \frac{f^{(j)}(x_0)}{j!}\,(x-x_0)^j\ ,$$

k-tes Taylor-Polynom von f in x_0 *und*

$$R_k(f,x_0) : I \to \mathbb{K}\ , \quad R_k(f,x_0)(x) := f(x) - T_k(f,x_0)(x)\ ,$$

k-tes Restglied von f in x_0.

Bemerkung 1.5.2 Wie man unmittelbar sieht, ist $T_k(f,x_0)$ das eindeutig bestimmte Polynom $p : \mathbb{R} \to \mathbb{K}$ vom Grad höchstens k mit

$$p^{(j)}(x_0) = f^{(j)}(x_0) \quad \text{für} \quad j = 0,1,\ldots,n\ .$$

□

Satz 1.5.3 (Satz von Taylor) *Sei $x_0 \in I$ und sei $f : I \to \mathbb{R}$ $(k+1)$-mal differenzierbar, $k \in \mathbb{N}_0$. Dann gilt:*

(i) **Cauchy-Form des Restgliedes:** *Zu jedem $x \in I$ existiert ein $s \in\ ]0,1[$ mit*

$$R_k(f,x_0)(x) = \frac{f^{(k+1)}(x_0 + s(x-x_0))}{k!}\,(1-s)^k(x-x_0)^{k+1}\ . \tag{1.5.1}$$

(ii) **Lagrange-Form des Restgliedes:** *Zu jedem $x \in I$ existiert ein $t \in\ ]0,1[$ mit*

$$R_k(f,x_0)(x) = \frac{f^{(k+1)}(x_0 + t(x-x_0))}{(k+1)!}\,(x-x_0)^{k+1}\ . \tag{1.5.2}$$

Beweis. Sei $x \in I$ fixiert, gelte o.B.d.A. $x \neq x_0$ und sei $\varphi : I \to \mathbb{R}$ durch

$$\varphi(y) := f(x) - T_k(f,y)(x)$$

definiert. Dann ist φ differenzierbar und es gilt $\varphi(x) = 0$ und $\varphi(x_0) = R_k(f,x_0)(x)$. Nach dem Mittelwertsatz der Differentialrechnung existiert demnach ein $s \in\ ]0,1[$ mit

$$R_k(f,x_0)(x) = -\varphi'(x_0 + s(x-x_0))(x-x_0)\ .$$

Da

$$\varphi'(y) = -\sum_{j=0}^{k} \frac{f^{(j+1)}(y)}{j!}\,(x-y)^j + \sum_{j=1}^{k} \frac{f^{(j)}(y)}{(j-1)!}\,(x-y)^{j-1} = -\frac{f^{(k+1)}(y)}{k!}\,(x-y)^k\ ,$$

folgt (1.5.1) und (i) ist bewiesen.

Zum Beweis von (ii) betrachten wir zusätzlich die Abbildung

$$\psi : I \to \mathbb{R}\ , \quad \psi(y) := (x-y)^{k+1}\ .$$

Auch diese Abbildung ist differenzierbar, wobei

$$\psi'(y) = -(k+1)(x-y)^k\ .$$

Außerdem gilt $\psi(x) = 0$ und $\psi(x_0) = (x - x_0)^{k+1}$. Damit existiert nach dem verallgemeinerten Mittelwertsatz der Differentialrechnung ein $t \in]0,1[$ mit

$$\frac{R_k(f,x_0)(x)}{(x-x_0)^{k+1}} = \frac{\varphi'(x_0 + t(x-x_0))}{\psi'(x_0 + t(x-x_0))} = \frac{f^{(k+1)}(x_0 + t(x-x_0))}{(k+1)!} ,$$

d.h. mit (1.5.2). □

Bemerkung 1.5.4 Offensichtlich ergibt der Satz von Taylor für $k = 0$ gerade wieder den Mittelwertsatz der Differentialrechnung. □

Als erste Anwendung von Satz 1.5.3 betrachten wir

Beispiel 1.5.5 Wir berechnen $\sqrt[3]{9}$ mit einer Fehlerdifferenz kleiner als 10^{-3}. Dafür benutzen wir folgende Überlegung. Sei $a \in]0,1[$ und sei $f :]-1,+\infty[\to \mathbb{R}$ durch

$$f(x) := (1+x)^a$$

gegeben. Dann ist f beliebig oft differenzierbar und es gilt $f(0) = 1$ und

$$f^{(j)}(x) = a(a-1)\cdots(a-j+1)(1+x)^{a-j} \quad \text{für} \quad j \in \mathbb{N} .$$

Somit ist

$$f(x) = \sum_{j=0}^{k} \binom{a}{j} x^j + R_k(f,0)(x) \quad \text{für} \quad k \in \mathbb{N}_0 .$$

Sei $k \in \mathbb{N}$ und sei $x \in \mathbb{R}$. Nach der Lagrange-Form des Restgliedes ist

$$R_k(f,0)(x) = \frac{f^{(k+1)}(tx)}{(k+1)!}\, x^{k+1} = \binom{a}{k+1}(1+tx)^{a-k-1}x^{k+1}$$

mit einem $t \in]0,1[$. Da

$$(1+tx)^{a-k-1} < 1 \quad \text{für} \quad x > 0 \text{ und } t \in]0,1[$$

und

$$\left|\binom{a}{k+1}\right| = \frac{a \cdot |a-1| \cdot |a-2| \cdots |a-k|}{(k+1)!} \leq \frac{a \cdot 1 \cdot 2 \cdots k}{(k+1)!} = \frac{a}{k+1} ,$$

folgt

$$|R_k(f,0)(x)| \leq \frac{a}{k+1}\, x^{k+1} \quad \text{für} \quad k \in \mathbb{N} \text{ und } x > 0 . \tag{1.5.3}$$

Sei jetzt $a = 1/3$. Wegen (1.5.3) und

$$f\left(\frac{1}{8}\right) = \left(\frac{9}{8}\right)^{1/3} = \frac{\sqrt[3]{9}}{2}$$

haben wir dann

$$\left|\sqrt[3]{9} - 2\sum_{j=0}^{2} \binom{1/3}{j} \frac{1}{8^j}\right| \leq \frac{2}{9 \cdot 8^3} = \frac{1}{2304} < 10^{-3} .$$

Also ist

$$2\sum_{j=0}^{2} \binom{1/3}{j} \frac{1}{8^j} = 2\left(1 + \frac{1}{3 \cdot 8} - \frac{1}{9 \cdot 64}\right) = \frac{599}{288}$$

eine Näherung von $\sqrt[3]{9}$ mit einem Fehler kleiner als 10^{-3}. □

Eine weitere Anwendung von Satz 1.5.3 ist die folgende Verallgemeinerung von Satz 1.2.9.

Satz 1.5.6 *Sei I offen, sei $x_0 \in I$ und sei $f : I \to \mathbb{R}$ k-mal differenzierbar, $k \geq 2$. Außerdem gelte*

$$f^{(j)}(x_0) = 0 \quad \textit{für} \quad j = 1, \dots, k-1 \; . \tag{1.5.4}$$

Ist dann k gerade und $f^{(k)}(x_0) < 0$ (bzw. $f^{(k)}(x_0) > 0$), so hat f in x_0 ein isoliertes lokales Maximum (bzw. isoliertes lokales Minimum).

Beweis. Sei k gerade und sei $f^{(k)}(x_0) < 0$. Da $f^{(k-1)}(x_0) = 0$ und somit

$$f^{(k)}(x_0) = \lim_{x \to x_0} \frac{f^{(k-1)}(x)}{x - x_0} \; ,$$

existiert dann ein $\varepsilon > 0$ derart, dass $]x_0 - \varepsilon, x_0 + \varepsilon[\subseteq I$ und

$$\frac{f^{(k-1)}(x)}{x - x_0} < 0 \quad \text{für alle} \quad x \in \,]x_0 - \varepsilon, x_0 + \varepsilon[\setminus \{x_0\} \; .$$

Also ist $f^{(k-1)}(x) > 0$ für alle $x \in \,]x_0 - \varepsilon, x_0[$ und $f^{(k-1)}(x) < 0$ für alle $x \in \,]x_0, x_0 + \varepsilon[$. Da k gerade ist, gilt dann auch

$$f^{(k-1)}(x_0 + t(x - x_0))(x - x_0)^{k-1} < 0 \;\text{ für alle }\; x \in \,]x_0 - \varepsilon, x_0 + \varepsilon[\setminus \{x_0\} \text{ und } t \in \,]0, 1[\; . \tag{1.5.5}$$

Nach Satz 1.5.3 und wegen (1.5.4) existiert zu jedem $x \in I$ ein $t \in \,]0, 1[$ mit

$$f(x) = f(x_0) + \frac{f^{(k-1)}(x_0 + t(x - x_0))}{(k-1)!} \, (x - x_0)^{k-1} \; .$$

Hieraus und aus (1.5.5) folgt

$$f(x) < f(x_0) \quad \text{für alle } x \in \,]x_0 - \varepsilon, x_0 + \varepsilon[\setminus \{x_0\} \; ,$$

was bedeutet, dass f in x_0 ein isoliertes lokales Maximum hat.
Der Beweis für $f^{(k)}(x_0) > 0$ verläuft völlig analog. □

Schaut man sich den letzten Beweis noch einmal an, so sieht man, dass in der Situation von Satz 1.5.6 für ungerades k und $f^{(k)}(x_0) \neq 0$ in x_0 kein lokales Extremum vorliegt. Stattdessen hat f in diesem Fall in x_0 einen so genannten Wendepunkt.

Definition 1.5.7 *Sei I offen und sei $f : I \to \mathbb{R}$. Man sagt, f* **hat in $x_0 \in I$ einen Wendepunkt** $:\Longleftrightarrow$ *Es existiert ein $\varepsilon > 0$ derart, dass $]x_0 - \varepsilon, x_0 + \varepsilon[\subseteq I$ und eine der Abbildungen $f\big|_{]x_0 - \varepsilon, x_0]}$ und $f\big|_{[x_0, x_0 + \varepsilon[}$ konvex und die andere konkav ist.*

Satz 1.5.8 *Sei I offen, sei $x_0 \in I$ und sei $f : I \to \mathbb{R}$ differenzierbar. Dann hat f genau dann in x_0 einen Wendepunkt, wenn f' in x_0 ein lokales Extremum hat.*

Beweis. Die Behauptung ist eine unmittelbare Schlussfolgerung aus Satz 1.2.13. □

Satz 1.5.9 *Sei I offen, sei $x_0 \in I$ und sei $f : I \to \mathbb{R}$ k-mal differenzierbar, $k \geq 3$. Außerdem gelte*

$$f^{(j)}(x_0) = 0 \quad \textit{für} \quad j = 2, \dots, k-1 \; .$$

Ist dann k ungerade und $f^{(k)}(x_0) \neq 0$, so hat f in x_0 einen Wendepunkt.

Beweis. Indem man Satz 1.5.6 auf f' anwendet, erhält man, dass f' in x_0 ein lokales Extremum hat. Nach Satz 1.5.8 hat f damit in x_0 einen Wendepunkt. □

Definition 1.5.10 *Sei $x_0 \in I$ und sei $f \in C^\infty(I, \mathbb{K})$. Dann heißt die Potenzreihe*

$$T(f, x_0)(x) := \sum_{j=0}^{\infty} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j$$

Taylor-Reihe von f in x_0.

Selbst wenn $T(f, x_0)(x)$ für ein $x \neq x_0$ konvergent ist, muss $T(f, x_0)(x)$ nicht notwendig gegen $f(x)$ konvergieren. Wir geben dazu das folgende Beispiel an.

Beispiel 1.5.11 Die Funktion

$$f : \mathbb{R} \to \mathbb{R}\,, \quad f(x) := \begin{cases} \exp(-1/x) & \text{für} \quad x > 0 \\ 0 & \text{für} \quad x \leq 0 \end{cases}\,,$$

ist beliebig oft differenzierbar und

$$f^{(j)}(0) = 0 \quad \text{für alle} \quad j \in \mathbb{N}_0\,.$$

Folglich ist $T(f, 0)(x) = 0$ für alle $x \in \mathbb{R}$, aber $f(x) \neq 0$ für $x > 0$. □

Existiert eine Umgebung $U \subseteq I$ von x_0 derart, dass $T(f, x_0)(x)$ für alle $x \in U$ gegen $f(x)$ konvergiert, so sagt man auch, dass f in x_0 eine Taylor-Entwicklung besitzt.

Satz 1.5.12 *Seien $x_0, x \in I$ und sei $f \in C^\infty(I, \mathbb{K})$. Dann konvergiert $T(f, x_0)(x)$ genau dann gegen $f(x)$, wenn*

$$\lim_{k\to\infty} R_k(f, x_0)(x) = 0\,.$$

Beweis. Dies folgt daraus, dass $T_k(f, x_0)(x)$ die k-te Partialsumme von $T(f, x_0)(x)$ ist und dass $R_k(f, x_0)(x) = f(x) - T_k(f, x_0)(x)$. □

Beispiel 1.5.13 Wir betrachten die Funktion

$$f :]-1, +\infty[\to \mathbb{R}\,, \quad f(x) := \ln(1 + x)\,.$$

Diese Funktion ist glatt und es gilt $f(0) = 0$ sowie

$$f^{(j)}(x) = \frac{(-1)^{j+1}(j-1)!}{(1+x)^j} \quad \text{für alle} \quad j \in \mathbb{N}\,.$$

Demnach ist

$$T(f, 0)(x) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} x^j\,.$$

Sei $x \in]-1, 1[$. Nach der Cauchy-Form des Restgliedes haben wir

$$R_k(f, 0)(x) = \frac{f^{(k+1)}(sx)}{k!} (1-s)^k x^{k+1} = \frac{(-1)^k}{(1+sx)^{k+1}} (1-s)^k x^{k+1}$$

für ein $s \in]0, 1[$. Mittels

$$1 + sx \geq 1 - s|x| \geq 1 - |x| > 0 \quad \text{und} \quad \frac{1-s}{1-s|x|} < 1$$

leiten wir

$$|R_k(f,0)(x)| = \frac{(1-s)^k |x|^{k+1}}{(1+sx)^{k+1}} \le \frac{|x|^{k+1}}{1-|x|}\left(\frac{1-s}{1-s|x|}\right)^k \le \frac{|x|^{k+1}}{1-|x|}$$

ab. Dies impliziert, dass

$$\lim_{k\to\infty} R_k(f,0)(x) = 0 \quad \text{für alle} \quad x \in \,]-1,1[\;.$$

Nach Satz 1.5.12 gilt somit

$$\ln(1+x) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j}\, x^j \quad \text{für alle} \quad x \in \,]-1,1[\;.$$

Die obige Gleichung bleibt auch für $x = 1$ richtig, d.h. es gilt

$$\ln(2) = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots\;. \tag{1.5.6}$$

Um dies einzusehen, benutzen wir die Lagrange-Form des Restgliedes. Danach ist

$$R_k(f,0)(1) = \frac{f^{(k+1)}(t)}{(k+1)!} = \frac{(-1)^k}{(k+1)(1+t)^{k+1}}$$

für ein $t \in \,]0,1[$. Es folgt

$$|R_k(f,0)(1)| \le \frac{1}{k+1}\;.$$

Dies liefert

$$\lim_{k\to\infty} R_k(f,0)(1) = 0$$

und (1.5.6) ist gezeigt. □

In engem Zusammenhang mit der Taylor-Entwicklung steht der Begriff der analytischen Funktion.

Definition 1.5.14 *Sei I offen. Eine Abbildung $f : I \to \mathbb{K}$ heißt (reell) analytisch $:\Longleftrightarrow$ Zu jedem $x_0 \in I$ existieren ein $\varepsilon > 0$ mit $]x_0-\varepsilon, x_0+\varepsilon[\; \subseteq I$ und eine Potenzreihe $\sum\limits_{k=0}^{\infty} a_k(x-x_0)^k$ mit Zentrum x_0 und Koeffizienten $a_k \in \mathbb{K}$ derart, dass*

$$f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k \quad \textit{für alle} \quad x \in \,]x_0-\varepsilon, x_0+\varepsilon[\;.$$

Die Menge aller analytischen Abbildungen $f : I \to \mathbb{K}$ wird üblicherweise mit $C^\omega(I,\mathbb{K})$ bezeichnet. Wie man leicht sieht, ist $C^\omega(I,\mathbb{K})$ ein $\mathbb{K}$-Vektorraum.

Satz 1.5.15 *Sei I offen. Eine Abbildung $f : I \to \mathbb{K}$ ist genau dann analytisch, wenn f glatt ist und zu jedem $x_0 \in I$ ein $\varepsilon > 0$ mit $]x_0-\varepsilon, x_0+\varepsilon[\; \subseteq I$ und*

$$f(x) = T(f,x_0)(x) \quad \textit{für alle} \quad x \in \,]x_0-\varepsilon, x_0+\varepsilon[\tag{1.5.7}$$

existiert.

Beweis. ($\Longrightarrow$) Sei $f : I \to \mathbb{K}$ analytisch und sei $x_0 \in I$. Dann gilt

$$f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k \quad \text{für alle} \quad x \in \,]x_0-\varepsilon, x_0+\varepsilon[\tag{1.5.8}$$

mit gewissen $a_k \in \mathbb{K}$ und einem $\varepsilon > 0$. Hieraus folgt durch wiederholtes Anwenden von Satz 1.4.5, dass f im Intervall $]x_0 - \varepsilon, x_0 + \varepsilon[$ beliebig oft differenzierbar ist und dass

$$f^{(j)}(x) = \sum_{k=j}^{\infty} k(k-1)\cdots(k-j+1)a_k(x-x_0)^{k-j} \quad \text{für alle} \quad x \in]x_0 - \varepsilon, x_0 + \varepsilon[\text{ und } j \in \mathbb{N}_0 \, .$$

Insbesondere ist

$$a_j = \frac{f^{(j)}(x_0)}{j!} \quad \text{für alle} \quad j \in \mathbb{N}_0 \, . \tag{1.5.9}$$

Aus (1.5.8) und (1.5.9) folgt (1.5.7).

($\Longleftarrow$) Das ist offensichtlich. $\square$

Beispiel 1.5.16 (i) Jedes Polynom

$$p : \mathbb{R} \to \mathbb{K} \, , \quad p(x) := \sum_{k=0}^{m} a_k x^k \, ,$$

ist analytisch, denn für alle $x_0, x \in \mathbb{R}$ gilt

$$\begin{aligned} p(x) = \sum_{k=0}^{m} a_k(x - x_0 + x_0)^k &= \sum_{k=0}^{m} a_k \sum_{l=0}^{k} \binom{k}{l} (x-x_0)^l x_0^{k-l} \\ &= \sum_{l=0}^{m} \left(\sum_{k=l}^{m} \binom{k}{l} a_k x_0^{k-l} \right) (x-x_0)^l \, . \end{aligned}$$

(ii) Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ ist analytisch. Wir haben nämlich

$$\exp(x) = \exp(x_0)\exp(x-x_0) = \exp(x_0) \sum_{k=0}^{\infty} \frac{(x-x_0)^k}{k!}$$

für alle $x_0, x \in \mathbb{R}$. $\square$

Den natürlichen Rahmen zur Untersuchung analytischer Funktionen bildet die komplexe Analysis, genauer die so genannte Funktionentheorie. In diesem Kontext kann man relativ einfach beweisen, dass, neben der Summe und dem Produkt, auch die Verkettung analytischer Funktionen und das Reziproke eine nirgends verschwindenden analytischen Funktion wieder analytisch sind. Außerdem kann man zeigen, dass, wie im letzten Beispiel in Spezialfällen nachgewiesen, eine Potenzreihe mit Zentrum $x_0 \in \mathbb{R}$ und Konvergenzradius $\rho > 0$ auf $]x_0 - \rho, x_0 + \rho[$ analytisch ist. Natürlich können diese Aussagen auch auf entsprechende Resultate für Potenzreihen zurückgeführt werden (vgl. z.B. Königsberger, Analysis 1, Abschnitt 14.2).

Kapitel 2

Differentialrechnung in mehreren Veränderlichen

2.1 Differenzierbare Abbildungen

Im Folgenden sei U eine offene Teilmenge von $\mathbb{R}^m$. Desweiteren sei wieder $E = \mathbb{K}^n$, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Definition 2.1.1 (i) *Eine Abbildung $f : U \to E$ heißt* **differenzierbar in** $a \in U$ $:\Longleftrightarrow$ *Es existiert eine lineare Abbildung $L : \mathbb{R}^m \to E$ mit*

$$\lim_{v\to 0} \frac{f(a+v) - f(a) - L(v)}{\|v\|} = 0\ . \tag{2.1.1}$$

(ii) *Eine Abbildung $f : U \to E$ heißt* **differenzierbar** $:\Longleftrightarrow$ *f ist in jedem $a \in U$ differenzierbar.*

Bemerkung 2.1.2 Nach Satz 1.1.5 ist Definition 2.1.1 eine Verallgemeinerung von Definition 1.1.1. □

Lemma 2.1.3 *Sei $f : U \to E$ und sei $a \in U$. Dann existiert höchstens eine lineare Abbildung $L : \mathbb{R}^m \to E$ mit (2.1.1).*

Beweis. Seien $L_1, L_2 : \mathbb{R}^m \to E$ linear und gelte

$$\lim_{v\to 0} \frac{f(a+v) - f(a) - L_1(v)}{\|v\|} = \lim_{v\to 0} \frac{f(a+v) - f(a) - L_2(v)}{\|v\|} = 0\ .$$

Für jedes $v \in \mathbb{R}^n \setminus \{0\}$ haben wir dann

$$\begin{aligned} \frac{L_1(v) - L_2(v)}{\|v\|} &= \lim_{t\to 0+} \frac{L_1(tv) - L_2(tv)}{\|tv\|} \\ &= \lim_{t\to 0+} \frac{f(a+tv) - f(a) - L_2(tv)}{\|tv\|} - \lim_{t\to 0+} \frac{f(a+tv) - f(a) - L_1(tv)}{\|tv\|} = 0\ , \end{aligned}$$

also $L_1(v_1) = L_2(v_2)$. □

Definition 2.1.4 *Ist $f : U \to E$ in $a \in U$ differenzierbar, so wird die durch (2.1.1) bestimmte lineare Abbildung $L : \mathbb{R}^m \to E$ das* **Differential von** f **in** a *genannt und mit $\mathrm{D}f(a)$ bezeichnet.*

Beispiel 2.1.5 Sei $L : \mathbb{R}^m \to E$ linear. Aus

$$L(a+v) - L(a) - L(v) = 0 \quad \text{für alle} \quad a, v \in \mathbb{R}^m$$

folgt, dass L differenzierbar ist und

$$\mathrm{D}L(a) = L \quad \text{für alle} \quad a \in \mathbb{R}^m \; .$$

□

Beispiel 2.1.6 Sei $\Phi : \mathbb{R}^k \times \mathbb{R}^l \to E$ bilinear, d.h. für alle $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^k \times \mathbb{R}^l$ und $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ ist

$$\begin{aligned} &\Phi(\alpha_1 x_1 + \alpha_1 x_2, \beta_1 y_1 + \beta_1 y_2) \\ &\qquad = \alpha_1\beta_1\Phi(x_1, y_1) + \alpha_1\beta_2\Phi(x_1, y_2) + \alpha_2\beta_1\Phi(x_2, y_1) + \alpha_2\beta_2\Phi(x_2, y_2) \; . \end{aligned}$$

Wir zeigen, dass Φ differenzierbar ist und

$$\mathrm{D}\Phi(a,b)(v,w) = \Phi(v,b) + \Phi(a,w) \quad \text{für alle} \quad (a,b), (v,w) \in \mathbb{R}^k \times \mathbb{R}^l \; .$$

Da Φ stetig und $M := \{(x,y) \in \mathbb{R}^k \times \mathbb{R}^l : \|x\| = \|y\| = 1\}$ kompakt ist, existiert ein $c > 0$ mit

$$\|\Phi(x,y)\| \le c \quad \text{für alle} \quad (x,y) \in M \; .$$

Außerdem gilt

$$\|(v,w)\|^2 = \|v\|^2 + \|w\|^2 \ge 2\,\|v\|\,\|w\| \quad \text{für alle} \quad (v,w) \in \mathbb{R}^k \times \mathbb{R}^l \; .$$

Somit ist

$$\begin{aligned} &\lim_{(v,w)\to(0,0)} \frac{\|\Phi(a+v, b+w) - \Phi(a,b) - \Phi(v,b) - \Phi(a,w)\|}{\|(v,w)\|} \\ &\qquad = \lim_{(v,w)\to(0,0)} \frac{\|\Phi(v,w)\|}{\|(v,w)\|} \\ &\qquad \le \lim_{(v,w)\to(0,0)} \sqrt{\frac{\|v\|\,\|w\|}{2}} \left\| \Phi\left(\frac{v}{\|v\|}, \frac{w}{\|w\|} \right) \right\| \\ &\qquad \le \frac{c}{\sqrt{2}} \lim_{(v,w)\to(0,0)} \sqrt{\|v\|\,\|w\|} \\ &\qquad = 0 \; , \end{aligned}$$

was die gewünschte Aussage liefert. □

Satz 2.1.7 *Ist $f : U \to E$ in $a \in U$ differenzierbar, so ist f in a stetig.*

Beweis. Sei $f : U \to E$ in $a \in U$ differenzierbar. Dann ist

$$\lim_{v\to 0} \left(f(a+v) - f(a) - \mathrm{D}f(a)(v) \right) = 0 \; .$$

Da $\mathrm{D}f(a)$ stetig ist und somit

$$\lim_{v\to 0} \mathrm{D}f(a)(v) = 0 \; ,$$

folgt

$$\lim_{v\to 0} \left(f(a+v) - f(a) \right) = \lim_{v\to 0} \left(f(a+v) - f(a) - \mathrm{D}f(a)(v) \right) + \lim_{v\to 0} \mathrm{D}f(a)(v) = 0 \; .$$

□

Offensichtlich gilt

Satz 2.1.8 (Summenregel) *Seien $f_1, f_2 : U \to E$ in $a \in U$ differenzierbar. Dann ist auch $f_1 + f_2 : U \to E$ in $a \in U$ differenzierbar und*

$$\mathrm{D}(f_1 + f_2)(a) = \mathrm{D}f_1(a) + \mathrm{D}f_2(a) \; .$$

□

Als weitere Rechenregeln haben wir:

Satz 2.1.9 (Kettenregel) *Sei U_1 eine offene Teilmenge von $\mathbb{R}^k$ und U_2 eine offene Teilmenge von $\mathbb{R}^l$. Sei $g : U_1 \to U_2$ in $a \in U_1$ differenzierbar und sei $f : U_2 \to E$ in $g(a)$ differenzierbar. Dann ist $f \circ g : U_1 \to E$ in a differenzierbar und*

$$\mathrm{D}(f \circ g)(a) = \mathrm{D}f(g(a)) \circ \mathrm{D}g(a) \; .$$

Beweis. Wir setzen

$$L := \mathrm{D}g(a) \quad \text{und} \quad K := \mathrm{D}f(g(a))$$

sowie

$$\begin{aligned} R(v) &:= g(a+v) - g(a) - L(v) \; , \\ S(w) &:= f(g(a)+w) - f(g(a)) - K(w) \; , \\ T(v) &:= f \circ g(a+v) - f \circ g(a) - K \circ L(v) \end{aligned}$$

für $v \in V_1 \subseteq \mathbb{R}^k$ und $w \in V_2 \subseteq \mathbb{R}^l$, wobei V_1 und V_2 genügend kleine Umgebungen von 0 sind. Nach Voraussetzung ist

$$\lim_{v \to 0} \frac{R(v)}{\|v\|} = 0 \tag{2.1.2}$$

und

$$\lim_{w \to 0} \frac{S(w)}{\|w\|} = 0 \; . \tag{2.1.3}$$

Wir müssen zeigen, dass

$$\lim_{v \to 0} \frac{T(v)}{\|v\|} = 0 \; .$$

Dazu bemerken wir, dass

$$\begin{aligned} T(v) &= f(g(a) + g(a+v) - g(a)) - f(g(a)) - K(L(v)) \\ &= f(g(a) + L(v) + R(v)) - f(g(a)) - K(L(v) + R(v)) + K(R(v)) \\ &= S(L(v) + R(v)) + K(R(v)) \; . \end{aligned}$$

Wegen (2.1.2) ist

$$\lim_{v \to 0} \frac{K(R(v))}{\|v\|} = 0 \; .$$

Es bleibt zu zeigen, dass

$$\lim_{v \to 0} \frac{S(L(v) + R(v))}{\|v\|} = 0 \; . \tag{2.1.4}$$

Nach (2.1.3) ist die Abbildung

$$\tilde{S} : V_2 \to E \; , \quad \tilde{S}(w) := \begin{cases} \dfrac{S(w)}{\|w\|} & \text{für} \quad w \neq 0 \\ 0 & \text{für} \quad w = 0 \end{cases}$$

in 0 stetig, was zusammen mit (2.1.2)

$$\lim_{v \to 0} \tilde{S}(L(v) + R(v)) = 0 \tag{2.1.5}$$

liefert. Außerdem können wir ein $c > 0$ und ein $\varepsilon > 0$ derart wählen, dass

$$\|L(x)\| \leq c \quad \text{für} \quad \|x\| = 1$$

und

$$\frac{\|R(v)\|}{\|v\|} \leq 1 \quad \text{für} \quad 0 < \|v\| < \varepsilon \, ,$$

was

$$\frac{\|L(v) + R(v)\|}{\|v\|} \leq \left\| L\left(\frac{v}{\|v\|}\right)\right\| + \frac{\|R(v)\|}{\|v\|} \leq c + 1 \quad \text{für} \quad 0 < \|v\| < \varepsilon$$

impliziert. Hieraus und aus (2.1.5) folgt

$$\lim_{v \to 0} \frac{S(L(v) + R(v))}{\|v\|} = \lim_{v \to 0} \frac{\|L(v) + R(v)\|}{\|v\|} \tilde{S}(L(v) + R(v)) = 0 \, .$$

Damit ist (2.1.4) gezeigt und der Satz ist bewiesen. □

Satz 2.1.10 (Produktregel) *Seien $f : U \to E$ und $g : U \to \mathbb{K}$ in $a \in U$ differenzierbar. Dann ist auch $fg : U \to E$ in $a \in U$ differenzierbar und*

$$\mathrm{D}(fg)(a)(v) = \mathrm{D}f(a)(v)\, g(a) + f(a)\, \mathrm{D}g(a)(v) \quad \textit{für alle} \quad v \in \mathbb{R}^m \, .$$

Beweis. Seien $\varphi : U \to E \times \mathbb{K}$ und $\Phi : E \times \mathbb{K} \to E$ durch

$$\varphi(x) := (f(x), g(x)) \quad \text{und} \quad \Phi(y, z) := yz$$

definiert. Die Abbildung φ ist nach Voraussetzung in a differenzierbar und

$$\mathrm{D}\varphi(a) = (\mathrm{D}f(a), \mathrm{D}g(a)) \, .$$

Mit Beispiel 2.1.6 und Satz 2.1.9 folgt

$$\begin{aligned}
\mathrm{D}(fg)(a)(v) &= \mathrm{D}(\Phi \circ \varphi)(a)(v) \\
&= \mathrm{D}\Phi(\varphi(a)) \circ \mathrm{D}\varphi(a)(v) \\
&= \mathrm{D}\Phi(f(a), g(a))(\mathrm{D}f(a)(v), \mathrm{D}g(a)(v)) \\
&= \Phi(\mathrm{D}f(a)(v), g(a)) + \Phi(f(a), \mathrm{D}g(a)(v)) \\
&= \mathrm{D}f(a)(v)\, g(a) + f(a)\, \mathrm{D}g(a)(v)
\end{aligned}$$

für $v \in \mathbb{R}^m$. □

Satz 2.1.11 (Mittelwertsatz der Differentialrechnung) *Sei $f : U \to \mathbb{R}$ differenzierbar und seien $a, b \in U$ derart, dass*

$$S(a, b) := \{a + t(b - a) : t \in \,]0, 1[\} \subseteq U \, .$$

Dann existiert ein $\xi \in S(a, b)$ mit

$$f(b) - f(a) = \mathrm{D}f(\xi)(b - a) \, .$$

Beweis. Sei $\varphi : [0, 1] \to U$ durch

$$\varphi(t) := a + t(b - a)$$

definiert und sei $F := f \circ \varphi$. Da F stetig und $F\big|_{]0,1[}$ differenzierbar ist, existiert nach Satz 1.2.5 ein $\vartheta \in \,]0, 1[$ mit

$$F(1) - F(0) = F'(\vartheta) \, .$$

Wir setzen $\xi := \varphi(\vartheta) \in S(a,b)$. Da (vgl. Satz 1.1.5)

$$F'(\vartheta) = \mathrm{D}F(\vartheta)(1) \quad \text{und} \quad \varphi'(\vartheta) = \mathrm{D}\varphi(\vartheta)(1) ,$$

erhalten wir durch Anwendung der Kettenregel

$$\begin{aligned} f(b) - f(a) &= F(1) - F(0) \\ &= \mathrm{D}F(\vartheta)(1) \\ &= \mathrm{D}f(\varphi(\vartheta)) \circ \mathrm{D}\varphi(\vartheta)(1) \\ &= \mathrm{D}f(\xi)(\varphi'(\vartheta)) \\ &= \mathrm{D}f(\xi)(b-a) . \end{aligned}$$

□

Satz 2.1.12 *Sei $U \subseteq \mathbb{R}^m$ offen und zusammenhängend und sei $f : U \to E$ differenzierbar. Dann ist f genau dann konstant, wenn*

$$\mathrm{D}f(x) = 0 \quad \textit{für alle} \quad x \in U . \tag{2.1.6}$$

Beweis. ($\Longrightarrow$) Das ist offensichtlich.

($\Longleftarrow$) Es gelte (2.1.6). Sei $a \in U$ fixiert und sei

$$M := \{x \in U : f(x) = f(a)\} .$$

Dann ist $M \neq \emptyset$, denn $a \in M$. Sei $x \in M$. Da U offen ist, existiert ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subseteq U$. Sei $\xi \in B_\varepsilon(x)$ und sei $\varphi : [0,1] \to B_\varepsilon(x)$ durch

$$\varphi(t) := x + t(\xi - x)$$

definiert und sei $F := f \circ \varphi$. Dann ist F stetig, $F\big|_{]0,1[}$ differenzierbar und

$$F'(t) = \mathrm{D}F(t)(1) = \mathrm{D}f(\varphi(t))(\mathrm{D}\varphi(t)(1)) = 0 \quad \text{für alle} \quad t \in]0,1[,$$

was nach Satz 1.3.12 impliziert, dass F konstant ist. Folglich gilt

$$f(\xi) = f(x) = f(a) \quad \text{für alle} \quad \xi \in B_\varepsilon(x) ,$$

d.h. $B_\varepsilon(x) \subseteq M$. Also ist M eine offene Teilmenge von $\mathbb{R}^m$. Aufgrund der Setigkeit von f ist M außerdem eine abgeschlossene Teilmenge von U und somit $U \setminus M$ eine offene Teilmenge von $\mathbb{R}^m$. Da U nach Voraussetzung zusammenhängend ist, folgt $M = U$, d.h.

$$f(x) = f(a) \quad \text{für alle} \quad x \in U .$$

□

2.2 Richtungsableitungen und partielle Differenzierbarkeit

In den folgenden Überlegungen sei $t \in \mathbb{R}$.

Definition 2.2.1 *Sei $f : U \to E$, sei $a \in U$ und sei $v \in \mathbb{R}^m$. Existiert der Grenzwert*

$$\lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} ,$$

so wird er die **Richtungsableitung von f in a in Richtung** *v genannt und mit $\nabla_v f(a)$ bezeichnet.*

Bemerkung 2.2.2 Die Richtungsableitung von $f : U \to E$ in $a \in U$ in Richtung $v \in \mathbb{R}^m$ existiert offensichtlich genau dann, wenn die durch $\varphi(t) := f(a + tv)$ für t nahe 0 definierte Funktion φ in 0 differenzierbar ist. In diesem Fall ist

$$\nabla_v f(a) = \varphi'(0) ,$$

was man auch in der Form

$$\nabla_v f(a) = \frac{\mathrm{d}}{\mathrm{d}t} f(a + tv)\Big|_{t=0}$$

schreibt. □

Satz 2.2.3 *Sei $f : U \to E$ in $a \in U$ differenzierbar. Dann existieren alle Richtungsableitungen von f in a und*

$$\nabla_v f(a) = \mathrm{D}f(a)(v) \quad \textit{für alle} \quad v \in \mathbb{R}^m .$$

Beweis. Sei $v \in \mathbb{R}^m \setminus \{0\}$. Dann gilt

$$\lim_{t\to 0} \left\| \frac{f(a + tv) - f(a)}{t} - \mathrm{D}f(a)(v) \right\| = \|v\| \lim_{t\to 0} \frac{\|f(a + tv) - f(a) - \mathrm{D}f(a)(tv)\|}{\|tv\|} = 0 ,$$

also

$$\lim_{t\to 0} \frac{f(a + tv) - f(a)}{t} = \mathrm{D}f(a)(v) .$$

Da trivialerweise

$$\nabla_0 f(a) = 0 = \mathrm{D}f(a)(0) ,$$

ist damit der Satz beweisen. □

Das folgende Beispiel zeigt, das man von der Existenz aller Richtungsableitungen in einem Punkt nicht auf Stetigkeit und somit auch nicht auf Differenzierbarkeit in diesem Punkt schließen kann.

Beispiel 2.2.4 Sei $f : \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x_1, x_2) := \begin{cases} \dfrac{x_1 x_2^2}{x_1^2 + x_2^4} & \text{für} \quad (x_1, x_2) \neq 0 \\ 0 & \text{für} \quad (x_1, x_2) = 0 \end{cases}$$

definiert. Dann ist f in 0 unstetig. Es existieren aber alle Richtungsableitungen von f in 0, denn für $t \neq 0$ und $v = (v_1, v_2) \in \mathbb{R}^2 \setminus \{0\}$ gilt

$$\frac{f(tv) - f(0)}{t} = \frac{f(tv)}{t} = \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^4} ,$$

woraus

$$\nabla_v f(0) = \begin{cases} \dfrac{v_2^2}{v_1} & \text{für} \quad v_1 \neq 0 \\ 0 & \text{für} \quad v_1 = 0 \end{cases}$$

folgt. □

Als Rechenregeln für die Richtungsableitung haben wir u.a.:

Satz 2.2.5 *Für $f, f_1, f_2 : U \to E$, $g : U \to \mathbb{K}$, $a \in U$, $v \in \mathbb{R}^m$ und $\alpha \in \mathbb{R}$ gilt:*

(i) *Existiert $\nabla_v f(a)$, so existiert auch $\nabla_{\alpha v} f(a)$ und*

$$\nabla_{\alpha v} f(a) = \alpha \nabla_v f(a) .$$

(ii) *Existieren* $\nabla_v f_1(a)$ *und* $\nabla_v f_2(a)$, *so existiert auch* $\nabla_v(f_1+f_2)(a)$ *und*

$$\nabla_v(f_1+f_2)(a) = \nabla_v f_1(a) + \nabla_v f_2(a)\ .$$

(iii) *Existieren* $\nabla_v f(a)$ *und* $\nabla_v g(a)$, *so existiert auch* $\nabla_v(fg)(a)$ *und*

$$\nabla_v(fg)(a) = \nabla_v f(a)\, g(a) + f(a)\, \nabla_v g(a)\ .$$

Beweis. Die Behauptungen ergeben sich aus Satz 1.1.9 und Bemerkung 2.2.2. $\square$

Sei $\{e_1,\ldots,e_m\}$ die Standardbasis von $\mathbb{R}^m$, d.h.

$$e_1 := (1,0,0,\ldots,0)\ , \quad e_2 := (0,1,0,\ldots)\ , \quad \ldots\ .$$

Definition 2.2.6 (i) *Eine Abbildung* $f : x \in U \mapsto f(x) = f(x_1,\ldots,x_m) \in E$ *heißt* **partiell differenzierbar in** $a \in U :\Longleftrightarrow$ *Die Richtungsableitungen* $\nabla_{e_i} f(a)$, $i = 1,\ldots,m$, *existieren. Die Richtungsableitung* $\nabla_{e_i} f(a)$ *wird dann die* **partielle Ableitung von** f **in** a **nach** x_i *genannt und mit* $\dfrac{\partial f}{\partial x_i}(a)$ *bezeichnet.*

(ii) *Eine Abbildung* $f : U \to E$ *heißt* **partiell differenzierbar** $:\Longleftrightarrow f$ *ist in jedem* $a \in U$ *partiell differenzierbar.*

Die partielle Ableitung einer Abbildung $f : U \to E$ in $a = (a_1,\ldots,a_n) \in U$ nach x_i ist also

$$\begin{aligned}\frac{\partial f}{\partial x_i}(a) &= \frac{\mathrm{d}}{\mathrm{d}t} f(a+te_i)\Big|_{t=0}\\ &= \frac{\mathrm{d}}{\mathrm{d}x_i} f(a_1,\ldots,a_{i-1},x_i,a_{i+1},\ldots,a_m)\Big|_{x_i=a_i}\ .\end{aligned}$$

Beispiel 2.2.7 Die Abbildung

$$f : \mathbb{R}^2 \to \mathbb{R}^2\ , \quad f(x_1,x_2) := \left(\sin\left(x_1x_2^2\right), \cos(x_1)\right)\ ,$$

ist partiell differenzierbar und es gilt

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = \left(x_2^2\cos\left(x_1x_2^2\right), -\sin(x_1)\right) \quad \text{und} \quad \frac{\partial f}{\partial x_2}(x_1,x_2) = \left(2x_1x_2\cos\left(x_1x_2^2\right), 0\right)\ .$$

$\square$

Satz 2.2.8 *Ist* $f : U \to E$ *in* $a \in U$ *differenzierbar, so ist* f *in* a *auch partiell differenzierbar und*

$$\frac{\partial f}{\partial x_i}(a) = \mathrm{D}f(a)(e_i) \quad \textit{für} \quad i = 1,\ldots,m\ .$$

Beweis. Das ist eine unmittelbare Konsequenz von Satz 2.2.3 und Definition 2.2.6. $\square$

Sei $f = (f_1,\ldots,f_n) : U \to \mathbb{R}^n$ in $a \in U$ differenzierbar und sei $v = (v_1,\ldots,v_m) \in \mathbb{R}^m$. Nach Satz 2.2.8 gilt

$$\begin{aligned}\mathrm{D}f(a)(v) = \mathrm{D}f(a)\left(\sum_{i=1}^m e_i v_i\right) &= \sum_{i=1}^m \mathrm{D}f(a)(e_i)\, v_i\\ &= \sum_{i=1}^m \frac{\partial f}{\partial x_i}(a)\, v_i\\ &= \left(\sum_{i=1}^m \frac{\partial f_1}{\partial x_i}(a)\, v_i, \ldots, \sum_{i=1}^m \frac{\partial f_n}{\partial x_i}(a)\, v_i\right)\ ,\end{aligned}$$

d.h.

$$\mathrm{D}f(a)(v) = \left(\frac{\partial f}{\partial x}(a) \cdot v^{\mathrm{T}}\right)^{\mathrm{T}} \;,$$

wobei

$$\frac{\partial f}{\partial x}(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_m}(a) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \dots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix} . \tag{2.2.1}$$

Also ist die Matrix $\frac{\partial f}{\partial x}(a)$ die Matrix-Darstellung der linearen Abbildung $\mathrm{D}f(a) : \mathbb{R}^m \to \mathbb{R}^n$ bezüglich der Standardbasen.

Definition 2.2.9 *Sei $f = (f_1, \dots, f_n) : U \to \mathbb{R}^n$ in $a \in U$ differenzierbar. Die in (2.2.1) definierte $n \times m$-Matrix $\frac{\partial f}{\partial x}(a)$ heißt die* **Jacobi-Matrix von f in** *a.*

Mit Hilfe der partiellen Ableitungen drückt sich die Kettenregel (Satz 2.1.9) folgendermaßen aus.

Satz 2.2.10 (Kettenregel) *Sei U_1 eine offene Teilmenge von $\mathbb{R}^k$ und U_2 eine offene Teilmenge von $\mathbb{R}^l$. Sei*

$$g : x \in U_1 \mapsto (g_1(x), \dots, g_l(x)) = (g_1(x_1, \dots, x_k), \dots, g_l(x_1, \dots, x_k)) \in U_2$$

in $a \in U_1$ differenzierbar und sei

$$f : y \in U_2 \mapsto f(y) = f(y_1, \dots, y_l) \in E$$

in $g(a)$ differenzierbar. Dann gilt

$$\frac{\partial (f \circ g)}{\partial x_i}(a) = \sum_{j=1}^{l} \frac{\partial f}{\partial y_j}(g(a)) \, \frac{\partial g_j}{\partial x_i}(a) \quad \textit{für} \quad i = 1, \dots, k \, .$$

Beweis. Gemäß den Sätzen 2.1.9 und 2.2.8 ist

$$\begin{aligned} \frac{\partial (f \circ g)}{\partial x_i}(a) = \mathrm{D}(f \circ g)(a)\left(e_i^{(k)}\right) &= \mathrm{D}f(g(a))\left(\mathrm{D}g(a)\left(e_i^{(k)}\right)\right) \\ &= \mathrm{D}f(g(a))\left(\sum_{j=1}^{l} e_j^{(l)} \, \frac{\partial g_j}{\partial x_i}(a)\right) \\ &= \sum_{j=1}^{l} \mathrm{D}f(g(a))\left(e_j^{(l)}\right) \frac{\partial g_j}{\partial x_i}(a) \\ &= \sum_{j=1}^{l} \frac{\partial f}{\partial y_j}(g(a)) \, \frac{\partial g_j}{\partial x_i}(a) \end{aligned}$$

für $i = 1, \dots, k$, wobei $\left\{e_1^{(k)}, \dots, e_k^{(k)}\right\}$ und $\left\{e_1^{(l)}, \dots, e_l^{(l)}\right\}$ die Standardbasen von $\mathbb{R}^k$ und bzw. $\mathbb{R}^l$ bezeichnen. □

Beispiel 2.2.4 zeigt, dass man aus partieller Differenzierbarkeit nicht auf Differenzierbarkeit schließen kann. Es gilt jedoch:

Satz 2.2.11 *Sei $f : U \to E$ partiell differenzierbar und seien die partiellen Ableitungen*

$$\frac{\partial f}{\partial x_i} : U \to E \ , \quad i = 1, \dots, m \ ,$$

in $a \in U$ stetig. Dann ist f in a differenzierbar.

Beweis. Seien $L : \mathbb{R}^m \to E$ und $\varphi : U \to E$ durch

$$L(v) := \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(a)\, v_i \quad \text{und} \quad \varphi(x) := f(x) - L(x - a)$$

definiert. Dann ist L offensichtlich linear. Außerdem ist φ partiell differenzierbar und

$$\frac{\partial \varphi}{\partial x_i}(x) = \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(a) \quad \text{für} \quad i = 1, \dots, m \ .$$

Da die partiellen Ableitungen von f in a stetig sind, existiert demnach zu vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ mit $B_\delta(a) \subseteq U$ und

$$\left\| \frac{\partial \varphi}{\partial x_i}(x) \right\| < \frac{\varepsilon}{\sqrt{m}} \quad \text{für alle} \quad x \in B_\delta(a) \text{ und } i = 1, \dots, m \ . \tag{2.2.2}$$

Sei $v \in \mathbb{R}^m$ mit $\|v\| < \delta$ und seien $\psi_i : [0, 1] \to E$, $i = 1, \dots, m$, durch

$$\begin{aligned}
\psi_1(t) &:= \varphi(a_1 + tv_1, a_2, a_3, \dots, a_{m-1}, a_m) \ , \\
\psi_2(t) &:= \varphi(a_1 + v_1, a_2 + tv_2, a_3, \dots, a_{m-1}, a_m) \ , \\
&\ \ \vdots \\
\psi_m(t) &:= \varphi(a_1 + v_1, a_2 + v_2, a_3 + v_3, \dots, a_{m-1} + v_{m-1}, a_m + tv_m)
\end{aligned}$$

definiert. Dann ist

$$\varphi(a + v) - \varphi(a) = \sum_{i=1}^{m} (\psi_i(1) - \psi_i(0)) \ .$$

Außerdem impliziert (2.2.2), dass

$$\|\psi_i'(t)\| < \frac{\varepsilon}{\sqrt{m}} |v_i| \quad \text{für alle} \quad t \in]0, 1[\text{ und } i = 1, \dots, m \ .$$

Mit Hilfe von Satz 1.3.10 und der Cauchy-Schwarzschen Ungleichung folgt

$$\|\varphi(a + v) - \varphi(a)\| \leq \sum_{i=1}^{m} \|\psi_i(1) - \psi_i(0)\| \leq \sum_{i=1}^{m} \frac{\varepsilon}{\sqrt{m}} |v_i| \leq \varepsilon \sqrt{\sum_{i=1}^{m} 1} \sqrt{\sum_{i=1}^{m} \frac{|v_i|^2}{m}} = \varepsilon \|v\| \ .$$

Damit haben wir gezeigt, dass

$$\lim_{v \to 0} \frac{f(a + v) - f(a) - L(v)}{\|v\|} = \lim_{v \to 0} \frac{\varphi(a + v) - \varphi(a)}{\|v\|} = 0 \ ,$$

und der Satz ist bewiesen. □

Wie man bereits anhand von Beispiel 1.1.21 sehen kann, ist die Stetigkeit der partiellen Ableitungen keine notwendige Bedingung für die Differenzierbarkeit.

Wir führen jetzt einen weiteren Begriff ein. Dabei benutzen wir den Fakt, dass zu jeder linearen Abbildung $L : \mathbb{R}^m \to \mathbb{R}$ genau ein Vektor $w \in \mathbb{R}^m$ mit

$$L(v) = \langle w, v \rangle \quad \text{für alle} \quad v \in \mathbb{R}^m$$

existiert.

Definition 2.2.12 *Sei $f : U \to \mathbb{R}$ in $a \in U$ differenzierbar. Der Vektor $w \in \mathbb{R}^m$ mit*

$$\mathrm{D}f(a)(v) = \langle w, v \rangle \quad \textit{für alle} \quad v \in \mathbb{R}^m$$

wird der **Gradient von f in** *a genannt und mit* $\operatorname{grad}(f)(a)$ *bezeichnet.*

Satz 2.2.13 *Sei $f : U \to \mathbb{R}$ in $a \in U$ differenzierbar. Dann ist*

$$\operatorname{grad}(f)(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_m}(a) \right) .$$

Beweis. Für $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$ haben wir nach Satz 2.2.8

$$\mathrm{D}f(a)(v) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(a)\, v_i = \left\langle \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_m}(a) \right), (v_1, \ldots, v_m) \right\rangle .$$

□

Der folgende Satz besagt, dass der Gradient diejenige Richtung angibt, in der die Funktion am stärksten wächst.

Satz 2.2.14 *Sei $f : U \to \mathbb{R}$ in $a \in U$ differenzierbar und sei $v \in \mathbb{R}^m$ mit $\|v\| = 1$. Dann gilt:*

(i) $|\nabla_v f(a)| \leq \|\operatorname{grad}(f)(a)\|$.

(ii) *Im Fall* $\operatorname{grad}(f)(a) \neq 0$ *ist*

$$\nabla_v f(a) = \|\operatorname{grad}(f)(a)\| \quad \Longleftrightarrow \quad v = \frac{\operatorname{grad}(f)(a)}{\|\operatorname{grad}(f)(a)\|} .$$

Beweis. Sei $v \in \mathbb{R}^m$ mit $\|v\| = 1$. Als erstes bemerken wir, dass

$$\nabla_v f(a) = \mathrm{D}f(a)(v) = \langle \operatorname{grad}(f)(a), v \rangle .$$

Nach der Cauchy-Schwarzschen Ungleichung ist

$$|\langle \operatorname{grad}(f)(a), v \rangle| \leq \|\operatorname{grad}(f)(a)\| ,$$

wobei die Gleichheit genau dann eintritt, wenn $\operatorname{grad}(f)(a)$ und v linear abhängig sind. Ist also $\operatorname{grad}(f)(a) \neq 0$ und

$$\langle \operatorname{grad}(f)(a), v \rangle = \|\operatorname{grad}(f)(a)\| , \tag{2.2.3}$$

so ist $v = \alpha \operatorname{grad}(f)(a)$ für ein $\alpha \in \mathbb{R}$. Setzt man das in die Gleichung (2.2.3) ein, so erhält man $\alpha = 1/\|\operatorname{grad}(f)(a)\|$, also

$$v = \frac{\operatorname{grad}(f)(a)}{\|\operatorname{grad}(f)(a)\|} . \tag{2.2.4}$$

Da aus (2.2.4) offensichtlich (2.2.3) folgt, ist damit der Satz bewiesen. □

Als nächstes definieren wir die höheren partiellen Ableitungen.

Definition 2.2.15 *Eine Abbildung $f : U \to E$ heißt* **einmal partiell differenzierbar** $:\Longleftrightarrow$ *f ist partiell differenzierbar. Die partiellen Ableitungen von f werden dann* **partielle Ableitungen erster Ordnung von** *f genannt. Die Abbildung f heißt* **k-mal partiell differenzierbar** *für $k = 2, 3, \ldots :\Longleftrightarrow f$ ist $(k-1)$-mal partiell differenzierbar und alle partiellen Ableitungen $(k-1)$-ter Ordnung von f sind partiell differenzierbar. Deren partielle Ableitungen werden dann* **partielle Ableitungen k-ter Ordnung von** *f genannt.*

Ist $f : x \in U \mapsto f(x) = f(x_1, \ldots, x_m) \in E$ zweimal partiell differenzierbar, so bezeichnen wir die partielle Ableitung von $\dfrac{\partial f}{\partial x_i}$ nach x_j mit $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$, d.h. wir setzen

$$\frac{\partial^2 f}{\partial x_j \partial x_i} := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) .$$

Analog setzen wir

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}} := \frac{\partial}{\partial x_{i_1}}\left(\frac{\partial^{k-1} f}{\partial x_{i_2} \cdots \partial x_{i_k}}\right)$$

für $k = 3, 4, \ldots$ und $i_1, i_2, \ldots, i_k = 1, \ldots, m$, falls f k-mal partiell differenzierbar ist. Wie üblich werden wir die Abkürzung

$$\frac{\partial^k f}{\partial x_i^k} := \frac{\partial^k f}{\partial x_i \cdots \partial x_i}$$

verwenden.

Beispiel 2.2.16 Die partiellen Ableitungen erster Ordnung von

$$f : \mathbb{R}^2 \to \mathbb{R}^2 , \quad f(x_1, x_2) := \left(x_1^3 x_2^2, x_1^2 + x_2^2\right) ,$$

sind

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \left(3\, x_1^2 x_2^2, 2\, x_1\right) \quad \text{und} \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = \left(2\, x_1^3 x_2, 2\, x_2\right) .$$

Als partielle Ableitungen zweiter Ordnung erhalten wir

$$\begin{aligned} \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) &= \left(6\, x_1 x_2^2, 2\right) , \\ \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2) &= \left(2\, x_1^3, 2\right) , \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2) &= \left(6\, x_1^2 x_2, 0\right) . \end{aligned}$$

□

Beispiel 2.2.17 Sei $f : \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x_1, x_2) := \begin{cases} \dfrac{x_1^3 x_2 - x_1 x_2^3}{x_1^2 + x_2^2} & \text{für} \quad (x_1, x_2) \neq 0 \\ 0 & \text{für} \quad (x_1, x_2) = 0 \end{cases}$$

definiert. Dann haben wir

$$\frac{\partial f}{\partial x_1}(x_1, x_2) := \begin{cases} \dfrac{x_1^4 x_2 + 4\, x_1^2 x_2^3 - x_2^5}{\left(x_1^2 + x_2^2\right)^2} & \text{für} \quad (x_1, x_2) \neq 0 \\ 0 & \text{für} \quad (x_1, x_2) = 0 \end{cases}$$

und

$$\frac{\partial f}{\partial x_2}(x_1, x_2) := \begin{cases} \dfrac{x_1^5 - 4\, x_1^3 x_2^2 - x_1 x_2^4}{\left(x_1^2 + x_2^2\right)^2} & \text{für} \quad (x_1, x_2) \neq 0 \\ 0 & \text{für} \quad (x_1, x_2) = 0 \end{cases} .$$

Folglich ist

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(0) = \lim_{t \to 0} \frac{1}{t} \frac{\partial f}{\partial x_1}(0, t) = -1 \quad \text{und} \quad \frac{\partial^2 f}{\partial x_1 \partial x_2}(0) = \lim_{t \to 0} \frac{1}{t} \frac{\partial f}{\partial x_2}(t, 0) = 1 .$$

Es ist somit

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(0) \neq \frac{\partial^2 f}{\partial x_1 \partial x_2}(0) .$$

□

Im allgemeinen darf also die Reihenfolge der partiellen Ableitungen nicht vertauscht werden.

Wir verallgemeinern Definition 1.1.20.

Definition 2.2.18 (i) *Sei $k \in \mathbb{N}$. Eine Abbildung $f : U \to E$ heißt* **k-mal stetig differenzierbar** *oder* **C^k-Abbildung** $:\Longleftrightarrow$

(1) *f ist k-mal partiell differenzierbar.*

(2) *f und alle partiellen Ableitungen bis zur Ordnung k von f sind stetig.*

(ii) *Eine Abbildung $f : U \to E$ heißt* **glatt**, **beliebig oft differenzierbar** *oder* **C^∞-Abbildung** $:\Longleftrightarrow$ *Für jedes $k \in \mathbb{N}$ ist f k-mal stetig differenzierbar.*

Wir bezeichnen die Menge aller k-mal stetig differenzierbaren Abbildungen $f : U \to E$ mit $C^k(U, E)$ und die Menge aller glatten Abbildungen $f : U \to E$ mit $C^\infty(U, E)$. Zusätzlich sei $C^0(U, E)$ die Menge aller stetigen Abbildungen $f : U \to E$. Dann gilt

$$C^0(U, E) \supseteq C^1(U, E) \supseteq C^2(U, E) \supseteq \cdots$$

und

$$C^\infty(U, E) = \bigcap_{k \in \mathbb{N}_0} C^k(U, E) \ .$$

Satz 2.2.19 (Satz von Schwarz) *Ist $f \in C^2(U, E)$, so ist*

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} \quad \textit{für} \quad i, j = 1, \ldots, m \ .$$

Beweis. Offensichtlich brauchen wir nur den Fall $m = 2$ und $E = \mathbb{R}$ zu betrachten. Sei also V eine offene Teilmenge von $\mathbb{R}^2$ und sei $f \in C^2(V, \mathbb{R})$. Wir müssen zeigen, dass

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_2} \ .$$

Dazu fixieren wir ein $a = (a_1, a_2) \in V$. Wir wählen $h_1, h_2 > 0$ derart, dass

$$[a_1, a_1 + h_1] \times [a_2, a_2 + h_2] \subseteq V \ ,$$

und setzen

$$\alpha := f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) - f(a_1, a_2 + h_2) + f(a_1, a_2) \ .$$

Außerdem definieren wir $\varphi_1 : [a_1, a_1 + h_1] \to \mathbb{R}$ und $\varphi_2 : [a_2, a_2 + h_2] \to \mathbb{R}$ durch

$$\varphi_1(t) := f(t, a_2 + h_2) - f(t, a_2) \quad \text{und} \quad \varphi_2(t) := f(a_1 + h_1, t) - f(a_1, t) \ .$$

Dann ist

$$\alpha = \varphi_1(a_1 + h_1) - \varphi_1(a_1) = \varphi_2(a_2 + h_2) - \varphi_2(a_2) \ .$$

Da φ_1 stetig und $\varphi_1\big|_{]a_1, a_1 + h_1[}$ differenzierbar ist, existiert nach Satz 1.2.5 ein $\xi_{11} \in \,]a_1, a_1 + h_1[$ mit

$$\varphi_1(a_1 + h_1) - \varphi_1(a_1) = h_1\, \varphi_1'(\xi_{11}) \ ,$$

d.h. mit

$$\alpha = h_1 \left(\frac{\partial f}{\partial x_1}(\xi_{11}, a_2 + h_2) - \frac{\partial f}{\partial x_1}(\xi_{11}, a_2) \right) \ .$$

Durch Anwendung von Satz 1.2.5 auf die Funktion

$$t \in [a_2, a_2 + h_2] \mapsto \frac{\partial f}{\partial x_1}(\xi_{11}, t) \in \mathbb{R}$$

folgt, dass ein $\xi_{12} \in \,]a_2, a_2 + h_2[$ mit

$$\alpha = h_1 h_2 \, \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_{11}, \xi_{12}) \, . \tag{2.2.5}$$

Analog erhält man durch Betrachtung der Funktion φ_2, dass ein $\xi_{21} \in \,]a_1, a_1 + h_1[$ und ein $\xi_{22} \in \,]a_2, a_2 + h_2[$ mit

$$\alpha = h_1 h_2 \, \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_{21}, \xi_{22}) \tag{2.2.6}$$

existieren. Wegen $h_1, h_2 > 0$ implizieren (2.2.5) und (2.2.6)

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_{11}, \xi_{12}) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_{21}, \xi_{22}) \, .$$

Sei nun $\varepsilon > 0$. Da auch die partiellen Ableitungen zweiter Ordnung von f stetig sind, existiert ein $\delta > 0$ mit $B_\delta(a) \subseteq V$ und

$$\left| \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) \right| < \varepsilon \quad \text{sowie} \quad \left| \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) - \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) \right| < \varepsilon \quad \text{für alle} \quad x \in B_\delta(a) \, .$$

Nach der obigen Überlegung finden wir $\xi_1, \xi_2 \in B_\delta(a)$ mit

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_2) \, .$$

Damit erhalten wir

$$\begin{aligned} \left| \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) - \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \right| &= \left| \left(\frac{\partial^2 f}{\partial x_2 \partial x_1}(a) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1) \right) + \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_2) - \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) \right) \right| \\ &\leq \left| \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1) \right| + \left| \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) - \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_2) \right| \\ &\leq 2\,\varepsilon \, . \end{aligned}$$

Da dabei $\varepsilon > 0$ beliebig ist, folgt

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(a) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(a)$$

□

Eine unmittelbare Konsequenz des letzten Satzes ist

Folgerung 2.2.20 *Die partiellen Ableitungen bis zur Ordnung k einer Abbildung $f \in C^k(U, E)$ sind unabhängig von der Reihenfolge, in der differenziert wird.* □

2.3 Der Satz von Taylor und lokale Extrema

Sei auch in diesem Abschnitt U eine offene Teilmenge von $\mathbb{R}^m$. Als erstes wollen wir den Satz von Taylor (Satz 1.5.3) für den Fall von Funktionen mehrerer Veränderlicher verallgemeinern. Dazu führen wir zunächst einige Bezeichnungen ein.

Sei p ein so genannter **Multiindex**, d.h. $p = (p_1, \ldots, p_m) \in \mathbb{N}_0^m$. Dann sei

$$|p| := p_1 + \cdots + p_m \quad \text{und} \quad p! := (p_1!) \cdots (p_m!) \, .$$

Die Zahl $|p| \in \mathbb{N}_0$ nennt man **Ordnung** von p. Für $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$ sei

$$v^p := v_1^{p_1} \cdots v_m^{p_m} \ .$$

Ist $f \in C^k(U, E)$ und $|p| \leq k$, so setzen wir

$$\partial^p f := \frac{\partial^{|p|} f}{\partial x_1^{p_1} \cdots \partial x_m^{p_m}} \ ,$$

wobei dies so zu verstehen ist, dass im Fall $p_i = 0$ keine partielle Ableitung nach x_i auftritt. Zum Beispiel ist

$$\begin{aligned}
\partial^p f &= f && \text{für} \quad p = (0, 0, 0, \ldots, 0) \ , \\
\partial^p f &= \frac{\partial f}{\partial x_1} && \text{für} \quad p = (1, 0, 0, \ldots, 0) \ , \\
\partial^p f &= \frac{\partial^3 f}{\partial x_1 \partial x_2^2} && \text{für} \quad p = (1, 2, 0, \ldots, 0) \ .
\end{aligned}$$

Definition 2.3.1 *Sei $f \in C^k(U, \mathbb{K})$ und $a \in U$. Dann heißt*

$$T_k(f, a) : \mathbb{R}^m \to \mathbb{K} \ , \quad T_k(f, a)(x) := \sum_{|p| \leq k} \frac{\partial^p f(a)}{p!} \, (x - a)^p \ ,$$

***k-tes Taylor-Polynom von f in** a und*

$$R_k(f, a) : U \to \mathbb{K} \ , \quad R_k(f, a)(x) := f(x) - T_k(f, a)(x) \ ,$$

***k-tes Restglied von f in** a.*

Zum Beispiel ist

$$T_2(f, a)(x) = f(a) + \sum_{i=1}^m \frac{\partial f}{\partial x_i}(a) \, (x_i - a_i) + \frac{1}{2} \sum_{i,j=1}^m \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \, (x_i - a_i)(x_j - a_j) \ .$$

Satz 2.3.2 *Sei $f \in C^{k+1}(U, \mathbb{R})$ und seien $a, x \in U$ derart, dass $S(a, x) \subseteq U$, wobei wieder $S(a, x) = \{a + t(x - a) : t \in \,]0, 1[\}$. Dann existiert ein $\xi \in S(a, x)$ mit*

$$R_k(f, a)(x) = \sum_{|p| = k+1} \frac{\partial^p f(\xi)}{p!} \, (x - a)^p \ .$$

Beweis. Sei $v := x - a$ und sei ein offenes Intervall $I \subseteq \mathbb{R}$ so gewählt, dass $[0, 1] \subseteq I$ und

$$a + tv \in U \quad \text{für alle} \quad t \in I \ .$$

Wir definieren $\varphi : I \to \mathbb{R}$ durch

$$\varphi(t) := f(a + tv) \ .$$

Mit Hilfe von Satz 2.2.10 berechnen wir

$$\begin{aligned}
\varphi'(t) &= \sum_{i=1}^m \frac{\partial f}{\partial x_i}(a + tv) \, v_i = \sum_{|p|=1} \partial^p f(a + tv) \, v^p \ , \\
\varphi''(t) &= \sum_{i,j}^m \frac{\partial f}{\partial x_i \partial x_j}(a + tv) \, v_i v_j = 2 \sum_{|p|=2} \frac{\partial^p f(a + tv)}{p!} \, v^p \ , \\
&\ \, \vdots \\
\varphi^{(k+1)}(t) &= \sum_{i_1, \ldots, i_{k+1}=1}^m \frac{\partial^{k+1} f}{\partial x_{i_1} \cdots \partial x_{i_{k+1}}}(a + tv) \, v_{i_1} \cdots v_{i_{k+1}} = (k+1)! \sum_{|p|=k+1} \frac{\partial^p f(a + tv)}{p!} \, v^p \ .
\end{aligned}$$

Folglich ist $\varphi \in C^{k+1}(I, \mathbb{R})$. Nach Satz 1.5.3 existiert dann ein $\vartheta \in]0,1[$ mit

$$R_k(\varphi, 0)(1) = \frac{\varphi^{(k+1)}(\vartheta)}{(k+1)!} \ .$$

Indem wir $\xi := a + \vartheta v \in S(a, x)$ setzen, erhalten wir

$$R_k(\varphi, 0)(1) = \sum_{|p|=k+1} \frac{\partial^p f(\xi)}{p!} \, v^p \ .$$

Da nach den obigen Berechnungen

$$T_k(\varphi, 0)(1) = \sum_{i=1}^{k} \frac{\varphi^{(i)}(0)}{i!} = \sum_{i=1}^{k} \sum_{|p|=i} \frac{\partial^p f(a)}{p!} \, v^p = T_k(f, a)(x)$$

und somit

$$R_k(\varphi, 0)(1) = \varphi(0) - T_k(\varphi, 0)(1) = f(a) - T_k(f, a)(x) = R_k(f, a)(x) \ ,$$

liefert das die Behauptung. □

Für $k = 0$ ist Satz 2.3.2 der Mittelwertsatz (Satz 2.1.11).

Folgerung 2.3.3 *Sei $f \in C^k(U, \mathbb{R})$ und $a \in U$. Dann ist*

$$\lim_{x \to a} \frac{R_k(f, a)(x)}{\|x - a\|^k} = 0 \ .$$

Beweis. Sei $\varepsilon > 0$. Wir wählen ein $\delta > 0$ derart, dass $B_\delta(a) \subseteq U$ und

$$\sum_{|p|=k} \frac{|\partial^p f(y) - \partial^p f(a)|}{p!} < \varepsilon \quad \text{für alle} \quad y \in B_\delta(a) \ .$$

Da

$$|v^p| \le \max\left\{ |v_i|^{|p|} : i = 1, \ldots, m \right\} \le \|v\|^{|p|}$$

für jeden Multiindex $p \in \mathbb{N}_0^m$ und $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$, folgt

$$\sum_{|p|=k} \left| \frac{\partial^p f(y) - \partial^p f(a)}{p!} \, v^p \right| \le \varepsilon \|v\|^k \quad \text{für alle} \quad y \in B_\delta(a) \text{ und } v \in \mathbb{R}^m \ . \tag{2.3.1}$$

Sei jetzt $x \in B_\delta(a)$. Nach Satz 2.3.2 existiert ein $\xi \in S(a, x)$ mit

$$f(x) = T_{k-1}(f, a)(x) + \sum_{|p|=k} \frac{\partial^p f(\xi)}{p!} \, (x - a)^p \ .$$

Damit gilt

$$\begin{aligned} R_k(f, a)(x) &= f(x) - T_k(f, a)(x) \\ &= f(x) - T_{k-1}(f, a)(x) - \sum_{|p|=k} \frac{\partial^p f(a)}{p!} \, (x - a)^p \\ &= \sum_{|p|=k} \frac{\partial^p f(\xi) - \partial^p f(a)}{p!} \, (x - a)^p \end{aligned}$$

und folglich

$$|R_k(f,a)(x)| \leq \sum_{|p|=k} \left| \frac{\partial^p f(\xi) - \partial^p f(a)}{p!} (x-a)^p \right| .$$

Das impliziert zusammen mit (2.3.1), dass

$$\frac{|R_k(f,a)(x)|}{\|x-a\|^k} \leq \varepsilon \quad \text{für alle} \quad x \in B_\delta(a) \setminus \{a\} ,$$

womit die Behauptung bewiesen ist. □

Im zweiten Teil dieses Abschnitts wollen wir Bedingungen für lokale Extrema von Funktionen mehrerer Veränderlicher herleiten. Wir beginnen mit folgenden Definitionen.

Definition 2.3.4 *Sei $f : U \to E$ differenzierbar. Ein $a \in U$ heißt* **stationärer** *oder* **kritischer Punkt von** f $:\Longleftrightarrow \mathrm{D}f(a) = 0$.

Nach Satz 2.2.8 ist a genau dann ein kritischer Punkt von f, wenn

$$\frac{\partial f}{\partial x_i}(a) = 0 \quad \text{für} \quad i = 1, \ldots, m .$$

Definition 2.3.5 *Sei $f \in C^2(U, \mathbb{R})$ und $a \in U$. Die Bilinearform*

$$\nabla^2 f(a) : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R} , \quad \nabla^2 f(a)(v,w) := \sum_{i,j=1}^m \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, v_i w_j ,$$

heißt **Hessesche Form von** f **in** a *und die $m \times m$-Matrix*

$$\mathrm{H}f(a) := \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \ldots & \dfrac{\partial^2 f}{\partial x_1 \partial x_m}(a) \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_m \partial x_1}(a) & \ldots & \dfrac{\partial^2 f}{\partial x_m \partial x_m}(a) \end{pmatrix}$$

Hesse-Matrix von f **in** a.

Der Satz von Schwarz impliziert, dass $\nabla^2 f(a)$ und $\mathrm{H}f(a)$ symmetrisch sind.

In den folgenden Betrachtungen sei $B : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ eine symmetrische Bilinearform.

Definition 2.3.6 (i) *B heißt* **positiv definit** *(in Zeichen: $B > 0$) $:\Longleftrightarrow$ Für alle $v \in \mathbb{R}^m \setminus \{0\}$ ist $B(v,v) > 0$.*

(ii) *B heißt* **negativ definit** *(in Zeichen: $B < 0$) $:\Longleftrightarrow$ Für alle $v \in \mathbb{R}^m \setminus \{0\}$ ist $B(v,v) < 0$.*

(iii) *B heißt* **positiv semidefinit** *(in Zeichen: $B \geq 0$) $:\Longleftrightarrow$ Für alle $v \in \mathbb{R}^m$ ist $B(v,v) \geq 0$.*

(iv) *B heißt* **negativ semidefinit** *(in Zeichen: $B \leq 0$) $:\Longleftrightarrow$ Für alle $v \in \mathbb{R}^m$ ist $B(v,v) \leq 0$.*

(v) *B heißt* **indefinit** $:\Longleftrightarrow$ *Es ist weder $B \geq 0$ noch $B \leq 0$.*

Seien die reellen Zahlen B_{ij}, $i,j = 1, \ldots, m$, dadurch bestimmt, dass

$$B(v,w) = \sum_{i,j=1}^m B_{ij}\, v_i w_j \quad \text{für alle} \quad v, w \in \mathbb{R}^m ,$$

d.h. es sei

$$B_{ij} := B(e_i, e_j) \quad \text{für} \quad i, j = 1, \ldots, m \; ,$$

und sei Spec(B) die Menge der Eigenwerte der $m \times m$-Matrix $(B_{ij})_{i,j=1,\ldots,m}$. Da diese Matrix reell und symmetrisch ist, ist Spec$(B) \subseteq \mathbb{R}$. Wie man leicht sieht, gilt

$$\begin{aligned}
B > 0 &\iff \operatorname{Spec}(B) \subseteq \,]0, +\infty[\; , \\
B < 0 &\iff \operatorname{Spec}(B) \subseteq \,]-\infty, 0[\; , \\
B \geq 0 &\iff \operatorname{Spec}(B) \subseteq [0, +\infty[\; , \\
B \leq 0 &\iff \operatorname{Spec}(B) \subseteq \,]-\infty, 0] \; .
\end{aligned}$$

Der nächste Satz, den wir hier ohne Beweis anführen, gibt uns ein einfaches Kriterium für die Definitheit einer symmetrischen Bilinearform.

Satz 2.3.7 *Die symmetrische Bilinearform B ist genau dann positiv definit, wenn*

$$\begin{vmatrix} B_{11} & \ldots & B_{1k} \\ \vdots & & \vdots \\ B_{k1} & \ldots & B_{kk} \end{vmatrix} > 0 \qquad \textit{für} \qquad k = 1, \ldots, m \; .$$

Sie ist genau dann negativ definit, wenn

$$(-1)^k \begin{vmatrix} B_{11} & \ldots & B_{1k} \\ \vdots & & \vdots \\ B_{k1} & \ldots & B_{kk} \end{vmatrix} > 0 \qquad \textit{für} \qquad k = 1, \ldots, m \; .$$

□

Nach diesen Vorbereitungen beweisen wir jetzt ein notwendiges Kriterium für lokale Extrema.

Satz 2.3.8 *Sei $f \in C^2(U, \mathbb{R})$ und $a \in U$. Hat f in a ein lokales Maximum (bzw. lokales Minimum), so ist a ein stationärer Punkt von f und $\nabla^2 f(a) \leq 0$ (bzw. $\nabla^2 f(a) \geq 0$).*

Beweis. Habe f in a ein lokales Maximum und sei $v \in \mathbb{R}^m$. Wir müssen zeigen, dass

$$\mathrm{D}f(a)(v) = 0 \quad \text{und} \quad \nabla^2 f(a)(v, v) \leq 0 \; . \tag{2.3.2}$$

Dazu definieren wir $\varphi : \,]-\varepsilon, \varepsilon[\, \to \mathbb{R}$ für ein genügend kleines ε durch

$$\varphi(t) := f(a + tv) \; .$$

Dann hat φ in 0 ein lokales Maximum. Außerdem ist φ zweimal differenzierbar. Somit gilt

$$\varphi'(0) = 0 \quad \text{und} \quad \varphi''(0) \leq 0 \; .$$

Da

$$\varphi'(t) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(a + tv)\, v_i \quad \text{und} \quad \varphi''(t) = \sum_{i,j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + tv)\, v_i v_j \; ,$$

folgt

$$\sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(a)\, v_i = 0 \quad \text{und} \quad \sum_{i,j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, v_i v_j \leq 0 \; ,$$

was nichts anderes als (2.3.2) ist.

Der Beweis für ein lokales Minimum verläuft analog. □

Als hinreichendes Kriterium für lokale Extrema haben wir:

Satz 2.3.9 *Sei $f \in C^2(U,\mathbb{R})$ und $a \in U$. Ist a ein stationärer Punkt von f und ist $\nabla^2 f(a) < 0$ (bzw. $\nabla^2 f(a) > 0$), so hat f in a ein isoliertes lokales Maximum (bzw. isoliertes lokales Minimum).*

Beweis. Sei $\mathrm{D}f(a) = 0$ und $\nabla^2 f(a) < 0$. Nach Folgerung 2.3.3 gilt dann

$$\lim_{v\to 0} \|v\|^{-2} \left(f(a+v) - f(a) - \frac{1}{2}\nabla^2 f(a)(v,v) \right) = 0 \ . \tag{2.3.3}$$

Die Abbildung

$$v \in \mathbb{R}^m \mapsto \nabla^2 f(a)(v,v) \in \mathbb{R}^m$$

ist stetig und nimmt folglich auf der kompakten Menge $S^{m-1} := \{x \in \mathbb{R}^m : \|x\| = 1\}$ ein Maximum c an. Da

$$\nabla^2 f(a)(v,v) = \nabla^2 f(a)\left(\frac{v}{\|v\|}, \frac{v}{\|v\|}\right) \|v\|^2 \quad \text{für} \quad v \in \mathbb{R}^m \setminus \{0\} \ ,$$

haben wir somit

$$\nabla^2 f(a)(v,v) \le c\,\|v\|^2 \quad \text{für} \quad v \in \mathbb{R}^m \ . \tag{2.3.4}$$

Wegen $\nabla^2 f(a) < 0$ gilt außerdem

$$c < 0 \ . \tag{2.3.5}$$

Wir benutzen jetzt (2.3.3). Danach können wir ein $\delta > 0$ so wählen, dass

$$\left| f(a+v) - f(a) - \frac{1}{2}\nabla^2 f(a)(v,v) \right| \le -\frac{c}{4}\,\|v\|^2 \quad \text{für} \quad \|v\| < \delta \ .$$

Das liefert zusammen mit (2.3.4), dass

$$f(a+v) - f(a) \le -\frac{c}{4}\,\|v\|^2 + \frac{1}{2}\nabla^2 f(a)(v,v) \le \frac{c}{4}\,\|v\|^2 \quad \text{für} \quad \|v\| < \delta \ ,$$

also

$$f(a+v) \le f(a) + \frac{c}{4}\,\|v\|^2 \quad \text{für} \quad \|v\| < \delta \ .$$

Mit (2.3.5) folgt, dass f in a ein isoliertes lokales Maximum hat.

Analog leitet man aus $\mathrm{D}f(a) = 0$ und $\nabla^2 f(a) > 0$ ab, dass f in a ein isoliertes lokales Minimum hat. □

Beispiel 2.3.10 Wir wollen untersuchen, in welchen Punkten die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R} \ , \quad f(x) := (x_1+1)x_1^2 + (x_1-1)x_2^2 \ ,$$

lokale Extrema hat. Da

$$\mathrm{grad}(f)(x) = \big((3x_1+2)x_1 + x_2^2, 2(x_1-1)x_2\big) \ ,$$

sind

$$a := (0,0) \quad \text{und} \quad b := (-2/3, 0)$$

die stationären Punkte von f. Die Hesse-Matrizen von f sind

$$\mathrm{H}f(x) = \begin{pmatrix} 6x_1+2 & 2x_2 \\ 2x_2 & 2x_1-2 \end{pmatrix} .$$

Insbesondere ist

$$\mathrm{H}f(a) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{und} \quad \mathrm{H}f(b) = \begin{pmatrix} -2 & 0 \\ 0 & -10/3 \end{pmatrix} .$$

Folglich ist $\nabla^2 f(a)$ indefinit. Nach Satz 2.3.8 folgt, dass f in a kein lokales Extremum hat. Die Hessesche Form $\nabla^2 f(b)$ ist negativ definit. Also hat f nach Satz 2.3.9 in b ein lokales Maximum. □

Aus der Semidefinitheit der Hesseschen Form kann man nicht schließen, ob ein lokales Extremum vorliegt oder nicht.

Beispiel 2.3.11 Seien $f, g : \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x) := x_1^2 + x_2^3 \quad \text{und} \quad g(x) := x_1^2 + x_2^4$$

definiert. Dann ist

$$\operatorname{grad}(f)(x) = \left(2x_1, 3x_2^2\right) \quad \text{und} \quad \operatorname{grad}(g)(x) = \left(2x_1, 4x_2^3\right) \ .$$

Somit ist 0 ein stationärer Punkt von f und von g. Die Hesse-Matrizen von f und g in 0 sind

$$\mathrm{H}f(0) = \mathrm{H}g(0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \ .$$

Folglich sind $\nabla^2 f(0)$ und $\nabla^2 g(0)$ semidefinit. Die Funktion f hat in 0 kein lokales Extremum. Dagegen hat g in 0 ein isoliertes globales Minimum. □

Zum Schluss dieses Abschnitts geben wir noch den folgenden Begriff an.

Definition 2.3.12 *Sei $f \in C^2(U, \mathbb{R})$. Ein $a \in U$ heißt* **Sattelpunkt von** f $:\Longleftrightarrow$ *a ist ein stationärer Punkt von f und $\nabla^2 f(a)$ ist indefinit.*

Beispiel 2.3.13 Die Abbildung

$$f : \mathbb{R}^2 \to \mathbb{R} \ , \quad f(x) := x_1^2 - x_2^2 \ ,$$

hat 0 als Sattelpunkt. Das gleiche gilt für die Funktion aus Beispiel 2.3.10. Der Graph der Abbildung

$$g : \mathbb{R}^2 \to \mathbb{R} \ , \quad g(x) := x_1^3 - 3x_1 x_2^2 \ ,$$

wird auch als **Affensattel** bezeichnet. Aber 0 ist kein Sattelpunkt von g. □

2.4 Lokale Umkehrbarkeit und implizite Funktionen

Ist $I \subseteq \mathbb{R}$ ein offenes Intervall und ist $f \in C^1(I, \mathbb{R})$ mit $f'(x) \neq 0$ für alle $x \in I$, so folgt mit Satz 1.2.6, dass f injektiv, also invertierbar ist. In höheren Dimensionen gilt das nicht.

Beispiel 2.4.1 Offensichlich ist

$$f : \mathbb{R}^2 \to \mathbb{R}^2 \ , \quad f(x) := \left(\mathrm{e}^{x_1} \cos(x_2), \mathrm{e}^{x_1} \sin(x_2)\right) \ ,$$

eine C^1-Abbildung. Außerdem ist

$$\frac{\partial f}{\partial x}(x) = \begin{pmatrix} \mathrm{e}^{x_1} \cos(x_2) & -\mathrm{e}^{x_1} \sin(x_2) \\ \mathrm{e}^{x_1} \sin(x_2) & \mathrm{e}^{x_1} \cos(x_2) \end{pmatrix}$$

und somit

$$\det\left(\frac{\partial f}{\partial x}(x)\right) = \mathrm{e}^{2x_1} \ .$$

Folglich ist $\mathrm{D}f(x)$ für jedes $x \in \mathbb{R}^2$ ein Isomorphismus. Aber f ist nicht injektiv. □

In diesem Abschnitt werden wir u.a. beweisen, dass eine C^1-Abbildung f, deren Differentiale $\mathrm{D}f(x)$ Isomorphismen sind, zumindest lokal umkehrbar (siehe nachstehende Definition) ist.

Definition 2.4.2 *Sei (X, d) ein metrischer Raum und M ein nichtleere Menge. Eine Abbildung $f : X \to M$ heißt* **lokal umkehrbar** *oder* **lokal invertierbar** $:\iff$ *Zu jedem $a \in X$ existiert eine Umgebung U_a von a derart, dass $f\big|_{U_a}$ injektiv ist.*

Wir führen jetzt eine wichtige Klasse von Abbildungen ein.

Definition 2.4.3 *Seien $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen und sei $k \in \mathbb{N} \cup \{\infty\}$. Eine Abbildung $f : U \to V$ heißt* **C^k-Diffeomorphismus** $:\iff$

(1) *f ist bijektiv.*

(2) *f und f^{-1} sind C^k-Abbildungen.*

Ist f ein C^k-Diffeomorphismus und ist $l \in \mathbb{N}$ und $l \leq k$, so ist f offensichtlich auch ein C^l-Diffeomorphismus.

Lemma 2.4.4 *Seien $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen und sei $f : U \to V$ ein C^1-Diffeomorphismus. Für alle $x \in U$ ist dann $\mathrm{D}f(x)$ ein Isomorphismus und*

$$\mathrm{D}\left(f^{-1}\right)(f(x)) = (\mathrm{D}f(x))^{-1} .$$

Beweis. Es ist $f^{-1} \circ f = \mathrm{id}_U$ und $f \circ f^{-1} = \mathrm{id}_V$. Mittels der Kettenregel folgt

$$\mathrm{D}\left(f^{-1}\right)(f(x)) \circ \mathrm{D}f(x) = \mathrm{D}\left(f^{-1} \circ f\right)(x) = \mathrm{D}\,\mathrm{id}_U(x) = \mathrm{id}_{\mathbb{R}^m}$$

und

$$\begin{aligned}\mathrm{D}f(x) \circ \mathrm{D}\left(f^{-1}\right)(f(x)) &= \mathrm{D}f\left(f^{-1}(f(x))\right) \circ \mathrm{D}\left(f^{-1}\right)(f(x)) \\ &= \mathrm{D}\left(f \circ f^{-1}\right)(f(x)) = \mathrm{D}\,\mathrm{id}_V(f(x)) = \mathrm{id}_{\mathbb{R}^n}\end{aligned}$$

für alle $x \in U$. Das ergibt die Behauptung. □

Eine unmittelbare Konsequenz von Lemma 2.4.4 ist

Folgerung 2.4.5 *Seien U, V und f wie in Lemma 2.4.4. Dann ist $m = n$.* □

Als nächstes beweisen wir

Satz 2.4.6 *Seien $U, V \subseteq \mathbb{R}^m$ offen, sei $k \in \mathbb{N}$ und sei $f : U \to V$ eine C^k-Abbildung mit folgenden Eigenschaften.*

(1) *Für jedes $x \in U$ ist $\mathrm{D}f(x)$ ein Isomorphismus.*

(2) *f ist bijektiv und f^{-1} ist stetig.*

Dann ist f ein C^k-Diffeomorphismus.

Beweis. Sei $g := f^{-1}$. Als erstes zeigen wir, dass g differenzierbar ist. Dazu fixieren wir ein $b \in V$. Sei $a := g(b)$ und $L := \mathrm{D}f(a)$. Außerdem setzen wir

$$R(x) := f(x) - f(a) - L(x - a) \quad \text{für} \quad x \in U$$

und

$$S(y) := g(y) - g(b) - L^{-1}(y - b) \quad \text{für} \quad y \in V .$$

Wir müssen verifizieren, dass

$$\lim_{y\to b} \frac{\|S(y)\|}{\|y-b\|} = 0 \ . \tag{2.4.1}$$

Die Differenzierbarkeit von f in a besagt, dass

$$\lim_{x\to a} \frac{R(x)}{\|x-a\|} = 0 \ . \tag{2.4.2}$$

Außerdem ist

$$\begin{aligned} R(g(y)) &= f(g(y)) - f(a) - L(g(y)-a)) \\ &= y - b - L(g(y) - g(b)) \end{aligned}$$

und somit

$$S(y) = -L^{-1}(R(g(y))) \ . \tag{2.4.3}$$

Wegen der Stetigkeit von L^{-1} gibt es ein $c > 0$ derart, dass

$$\left\|L^{-1}(w)\right\| \leq c \quad \text{für alle} \quad w \in \mathbb{R}^m \text{ mit } \|w\| = 1 \ ,$$

was zusammen mit der Linearität von L^{-1}

$$\left\|L^{-1}(w)\right\| \leq c\,\|w\| \quad \text{für alle} \quad w \in \mathbb{R}^m \tag{2.4.4}$$

impliziert. Aufgrund von (2.4.2) existiert ein $\varepsilon > 0$

$$\|R(x)\| \leq \frac{1}{2c}\,\|x-a\| \quad \text{für alle} \quad x \in B_\varepsilon(a) \cap U \ . \tag{2.4.5}$$

Da g stetig ist, existiert weiter ein $\delta > 0$ mit

$$\|g(y) - a\| < \varepsilon \quad \text{für alle} \quad y \in B_\delta(b) \cap V \ . \tag{2.4.6}$$

Aus (2.4.3)-(2.4.6) schließen wir für $y \in B_\delta(b) \cap V$

$$\|S(y)\| = \left\|L^{-1}(R(g(y)))\right\| \leq c\,\|R(g(y))\| \leq \frac{1}{2}\,\|g(y)-a\| \ .$$

Da nach Definition von $S(y)$ und (2.4.4)

$$\|g(y) - a\| \leq c\,\|y-b\| + \|S(y)\| \ ,$$

folgt

$$\|g(y) - a\| \leq 2c\,\|y-b\| \quad \text{für alle} \quad y \in B_\delta(b) \cap V \ . \tag{2.4.7}$$

Aus (2.4.3), (2.4.4) und (2.4.7) erhalten wir für $y \in (B_\delta(b) \setminus \{b\}) \cap V$

$$\frac{\|S(y)\|}{\|y-b\|} = \frac{\left\|L^{-1}(R(g(y)))\right\|}{\|y-b\|} \leq c\,\frac{\|R(g(y))\|}{\|y-b\|} = c\,\frac{\|g(y)-a\|}{\|y-b\|}\,\frac{\|R(g(y))\|}{\|g(y)-a\|} \leq 2c^2\,\frac{\|R(g(y))\|}{\|g(y)-a\|} \ .$$

Da aufgrund von (2.4.2) und der Stetigkeit von g

$$\lim_{y\to b} \frac{R(g(y))}{\|g(y)-a\|} = 0 \ ,$$

ist damit (2.4.1) gezeigt.

Als nächstes beweisen wir, dass g eine C^1-Abbildung ist. Für f und $g = (g_1, \ldots, g_m)$ gilt nach der Kettenregel

$$\mathrm{D}f(g(y)) \circ \mathrm{D}g(y) = \mathrm{id}_{\mathbb{R}^m} \quad \text{für alle} \quad y \in V \ .$$

Folglich ist

$$\mathrm{D}g(y) = (\mathrm{D}f(g(y)))^{-1} \quad \text{für alle} \quad y \in V \ ,$$

d.h.

$$\frac{\partial g}{\partial y}(y) = \left(\frac{\partial f}{\partial x}(g(y))\right)^{-1} \quad \text{für alle} \quad y \in V \ .$$

Mit Hilfe der Formel für die inverse Matrix ergibt sich, dass

$$\frac{\partial g_i}{\partial y_j} = \psi_{ij} \circ \frac{\partial f}{\partial x} \circ g \quad \text{für} \quad i, j = 1, \ldots, m \ , \tag{2.4.8}$$

wobei wir $\frac{\partial f}{\partial x}$ als eine Abbildung von U in $\mathbb{R}^{m^2}$ verstehen und die ψ_{ij} gewisse rationale Funktionen von m^2 Veränderlichen sind. Da g und die ψ_{ij} stetig sind, folgt aus (2.4.8), dass alle partiellen Ableitungen erster Ordnung von g stetig sind. Also ist $g \in C^1(V, \mathbb{R}^m)$.

Schließlich erhalten wir durch vollständige Induktion, dass $g \in C^k(V, \mathbb{R}^m)$, falls $f \in C^k(U, \mathbb{R}^m)$. Ist nämlich $g \in C^l(V, \mathbb{R}^m)$ mit $1 \leq l < k$, so ist nach (2.4.8) jede partielle Ableitung $\frac{\partial g_i}{\partial y_j}$ die Komposition einer glatten, einer C^{k-1}- und einer C^l-Abbildung und somit selbst eine C^l-Abbildung. Folglich ist $g \in C^{l+1}(V, \mathbb{R}^m)$. □

Sei wieder $E = \mathbb{K}^n$ und bezeichne $\mathcal{L}(\mathbb{R}^m, E)$ den $\mathbb{K}$-Vektorraum aller linearen Abbildungen $L : \mathbb{R}^m \to E$.

Definition 2.4.7 *Sei $L \in \mathcal{L}(\mathbb{R}^m, E)$. Die reelle Zahl*

$$\|L\| := \max \{\|L(v)\| : v \in \mathbb{R}^m \ \textit{und} \ \|v\| = 1\}$$

heißt **Operatornorm von** L.

Lemma 2.4.8 *Für $L, L_1, L_2 \in \mathcal{L}(\mathbb{R}^m, E)$, $\lambda \in \mathbb{K}$ und $v \in \mathbb{R}^m$ gilt:*

(i) $\|L\| \geq 0$,

(ii) $\|L\| = 0 \iff L = 0$,

(iii) $\|\lambda L\| = \lambda \cdot \|L\|$,

(iv) $\|L_1 + L_2\| \leq \|L_1\| + \|L_2\|$,

(v) $\|L(v)\| \leq \|L\| \cdot \|v\|$.

Beweis. (i)-(iv) Übung.

(v) Sei o.B.d.A. $v \neq 0$. Aufgrund der Linearität von L und wegen (iii) ist dann

$$\|L(v)\| = \left\| L\left(\frac{v}{\|v\|}\right) \|v\| \right\| = \left\| L\left(\frac{v}{\|v\|}\right) \right\| \cdot \|v\| \leq \|L\| \cdot \|v\| \ .$$

□

Bemerkung 2.4.9 Die Eigenschaften (i)-(iv) von Lemma 2.4.8 besagen, dass die Abbildung

$$L \in \mathcal{L}(\mathbb{R}^m, E) \mapsto \|L\| \in \mathbb{R}$$

eine Norm ist. □

Sei wieder $S(x, y) = \{x + t(y - x) : t \in]0, 1[\}$ für $x, y \in \mathbb{R}^m$.

Definition 2.4.10 *Eine Menge $M \subseteq \mathbb{R}^m$ heißt* **konvex** $:\iff$ *Für alle $x, y \in M$ ist $S(x, y) \subseteq M$.*

Der folgende Satz wird zuweilen als **Schrankensatz** zitiert.

Satz 2.4.11 *Sei $U \subseteq \mathbb{R}^m$ offen, sei $M \subseteq U$ konvex und sei $f \in C^1(U, E)$. Gilt*

$$\|\mathrm{D}f(x)\| \leq c \quad \textit{für alle} \quad x \in M \tag{2.4.9}$$

mit einem $c \in \mathbb{R}$, so ist

$$\|f(y) - f(x)\| \leq c\,\|y - x\| \quad \textit{für alle} \quad x, y \in M\ .$$

Beweis. Es gelte (2.4.9). Seien $x, y \in M$ fixiert, sei $\varphi : [0, 1] \to U$ durch

$$\varphi(t) := x + t(y - x)$$

definiert und sei $F := f \circ \varphi$. Dann ist F stetig und $F\big|_{]0,1[}$ differenzierbar. Außerdem ist

$$F'(t) = \mathrm{D}f(\varphi(t))(\varphi'(t)) = \mathrm{D}f(\varphi(t))(y - x) \quad \text{für alle} \quad t \in]0, 1[\ .$$

Mit Lemma 2.4.8(v) ergibt sich

$$\|F'(t)\| \leq \|\mathrm{D}f(\varphi(t))\| \cdot \|y - x\| \leq c\,\|y - x\| \quad \text{für alle} \quad t \in]0, 1[\ .$$

Durch Anwendung von Satz 1.3.10 auf F folgt

$$\|f(y) - f(x)\| \leq c\,\|y - x\|\ .$$

□

Der nachstehende Satz liefert die oben angeführte Aussage zur lokalen Umkehrbarkeit von C^k-Abbildungen.

Satz 2.4.12 *Sei $U \subseteq \mathbb{R}^m$ offen, sei $f \in C^k(U, \mathbb{R}^m)$ für ein $k \in \mathbb{N}$ und sei $a \in U$ derart, dass $\mathrm{D}f(a)$ ein Isomorphismus ist. Dann gibt es eine Umgebung $U_a \subseteq U$ von a mit:*

(1) *$f(U_a)$ ist offen.*

(2) *$f\big|_{U_a} : U_a \to f(U_a)$ ist ein C^k-Diffeomorphimus.*

Beweis. Wir können o.B.d.A. $a = 0$, $f(a) = 0$ und $\mathrm{D}f(a) = \mathrm{id}_{\mathbb{R}^m}$ annehmen. Andernfalls betrachte man statt f die Abbildung

$$x \mapsto (\mathrm{D}f(a))^{-1}(f(x + a) - f(a))\ .$$

Diese bildet 0 auf 0 ab und hat $\mathrm{id}_{\mathbb{R}^m}$ als Differential in 0.

Mit Hilfe des Banachschen Fixpunktsatzes konstruieren wir eine Umgebung U_0 von 0, für die $f(U_0)$ offen und $f\big|_{U_0}$ invertierbar ist. Dazu definieren wir $F_y : U \to \mathbb{R}^m$ für $y \in \mathbb{R}^m$ durch

$$F_y(x) := y + x - f(x)\ .$$

Dann ist

$$\mathrm{D}F_y(x) = \mathrm{id}_{\mathbb{R}^m} - \mathrm{D}f(x) \quad \text{für alle} \quad x \in U\ . \tag{2.4.10}$$

Da f stetig differenzierbar ist und da $\mathrm{D}f(0) = \mathrm{id}_{\mathbb{R}^m}$, gibt es ein $\varepsilon > 0$ derart, dass $\bar{B}_{2\varepsilon}(0) := \{x \in \mathbb{R}^m : \|x\| \leq 2\varepsilon\}$ eine Teilmenge von U ist und

$$\| \mathrm{id}_{\mathbb{R}^m} - \mathrm{D}f(x)\| \leq \frac{1}{2} \quad \text{für alle} \quad x \in \bar{B}_{2\varepsilon}(0)\ . \tag{2.4.11}$$

Aus (2.4.10) und (2.4.11) ergibt sich nach Satz 2.4.11, dass

$$\|F_y(x_1) - F_y(x_2)\| \leq \frac{1}{2}\,\|x_1 - x_2\| \quad \text{für alle} \quad x_1, x_2 \in \bar{B}_{2\varepsilon}(0) \text{ und } y \in \mathbb{R}^m\ . \tag{2.4.12}$$

Weiter folgt

$$\|F_y(x)\| \leq \|F_y(x) - F_y(0)\| + \|y\| < 2\varepsilon \quad \text{für alle} \quad x \in \bar{B}_{2\varepsilon}(0) \text{ und } y \in B_\varepsilon(0)\ . \tag{2.4.13}$$

Sei $y \in B_\varepsilon(0)$. Nach (2.4.13) gilt $F_y\left(\bar{B}_{2\varepsilon}(0)\right) \subseteq \bar{B}_{2\varepsilon}(0)$ und nach (2.4.12) ist

$$F_y\big|_{\bar{B}_{2\varepsilon}(0)} \colon \bar{B}_{2\varepsilon}(0) \to \bar{B}_{2\varepsilon}(0)$$

kontrahierend. Außerdem ist $\bar{B}_{2\varepsilon}(0)$ als abgeschlossene Teilmenge eines vollständigen metrischen Raumes selbst vollständig. Nach dem Banachschen Fixpunktsatz hat somit $F_y\big|_{\bar{B}_{2\varepsilon}(0)}$ genau einen Fixpunkt x_y. Wegen (2.4.13) gilt sogar $x_y \in B_{2\varepsilon}(0)$. Wir setzen jetzt

$$V_0 := B_\varepsilon(0) \quad \text{und} \quad U_0 := f^{-1}(V_0) \cap B_{2\varepsilon}(0)\ .$$

Dann sind U_0 und V_0 offen. Da $f(x) = y$ äquivalent zu $F_y(x) = x$ ist, gilt

$$f(x_y) = y \quad \text{für alle} \quad y \in V_0\ .$$

Folglich ist $f(U_0) = V_0$. Insbesondere ist $f(U_0)$ offen. Desweiteren ist $f\big|_{U_0} \colon U_0 \to V_0$ injektiv, denn für $x \in U_0$ und $y \in V_0$ gilt $f(x) = y$ dann und nur dann, wenn $x = x_y$.

Es bleibt zu zeigen, dass $f\big|_{U_0} \colon U_0 \to V_0$ ein C^k-Diffeomorphismus ist. Dafür bemerken wir zunächst, dass die Umkehrabbildung $g : V_0 \to U_0$ von $f\big|_{U_0}$ stetig ist. Für $y_1, y_2 \in V_0$ haben wir nämlich

$$\begin{aligned} g(y_1) - g(y_2) &= F_0(g(y_1)) - F_0(g(y_2)) + f(g(y_1)) - f(g(y_2)) \\ &= F_0(g(y_1)) - F_0(g(y_2)) + y_1 - y_2\ , \end{aligned}$$

woraus wir unter Benutzung von (2.4.12)

$$\begin{aligned} \|g(y_1) - g(y_2)\| &\leq \|F_0(g(y_1)) - F_0(g(y_2))\| + \|y_1 - y_2\| \\ &\leq \frac{1}{2}\,\|g(y_1) - g(y_2)\| + \|y_1 - y_2\|\ , \end{aligned}$$

also

$$\|g(y_1) - g(y_2)\| \leq 2\,\|y_1 - y_2\|$$

erhalten. Außerdem ist $\mathrm{D}f(x)$ für jedes $x \in U_0$ ein Isomorphismus. Ist nämlich $x \in U_0$ und $v \in \mathbb{R}^m$ und gilt $\mathrm{D}f(x)(v) = 0$, so folgt nach (2.4.11)

$$\|v\| = \|(\mathrm{id}_{\mathbb{R}^m} - \mathrm{D}f(x))(v)\| \leq \frac{1}{2}\,\|v\|\ ,$$

was $v = 0$ impliziert. Damit erfüllt $f\big|_{U_0} \colon U_0 \to V_0$ die Voraussetzungen von Satz 2.4.6. Folglich ist diese Abbildung ein C^k-Diffeomorphismus. □

Wir ziehen eine wichtige Folgerung.

Folgerung 2.4.13 *Sei $U \subseteq \mathbb{R}^m$ offen, sei $f \in C^k(U, \mathbb{R}^m)$ für ein $k \in \mathbb{N} \cup \{\infty\}$ und sei $\mathrm{D}f(x)$ für jedes $x \in U$ ein Isomorphismus. Dann gilt:*

(i) *Ist $V \subseteq U$ offen, so ist auch $f(V)$ ist offen.*

(ii) *Ist f injektiv, so ist $f : U \to f(U)$ ein C^k-Diffeomorphismus.*

Beweis. (i) Sei $V \subseteq U$ offen und o.B.d.A. $V \neq \emptyset$. Wir wenden Satz 2.4.12 auf die Abbildung $f\big|_V$ an. Danach existiert zu jedem $x \in V$ eine Umgebung $V_x \subseteq V$ von x mit der Eigenschaft, dass $f\big|_V(V_x) = f(V_x)$ offen ist. Folglich ist $f(V) = \bigcup_{x \in V} f(V_x)$ die Vereinigung offener Mengen und somit selbst offen.

(ii) Sei f injektiv. Die Umkehrabbildung $g : f(U) \to U$ von f ist nach (i) stetig. Mit Satz 2.4.6 folgt, dass f ein C^k-Diffeomorphismus ist. □

Diffeomorphismen kann man auch als Koordinatentransformationen verstehen. Als wichtige Beispiele haben wir:

Beispiel 2.4.14 Die Abbildung

$$\begin{aligned} f : \mathbb{R}_+ \times]0, 2\pi[\quad &\to \quad \mathbb{R}^2 \setminus ([0, +\infty[\times \{0\}) , \\ (r, \theta) \quad &\mapsto \quad (r \cos(\theta), r \sin(\theta)) , \end{aligned}$$

ist bijektiv und glatt. Außerdem ist

$$\det \left(\frac{\partial f}{\partial(r, \theta)}(r, \theta) \right) = \begin{vmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{vmatrix} = r > 0$$

und somit $\mathrm{D}f(r, \theta)$ für alle $(r, \theta) \in \mathbb{R}_+ \times]0, 2\pi[$ ein Isomorphismus. Also ist f nach Folgerung 2.4.13 ein C^∞-Diffeomorphimus. Man nennt (r, θ) **Polarkoordinaten** von $(x_1, x_2) = f(r, \theta)$. □

Beispiel 2.4.15 Das vorhergehende Beispiel impliziert, dass

$$\begin{aligned} f : \mathbb{R}_+ \times]0, 2\pi[\times \mathbb{R} \quad &\to \quad \mathbb{R}^3 \setminus ([0, +\infty[\times \{0\} \times \mathbb{R}) , \\ (r, \theta, t) \quad &\mapsto \quad (r \cos(\theta), r \sin(\theta), t) , \end{aligned}$$

ein C^∞-Diffeomorphimus. Man nennt (r, θ, t) **Zylinderkoordinaten** von $(x_1, x_2, x_3) = f(r, \theta, t)$. □

Beispiel 2.4.16 Die **Polarkoordinaten** (r, θ_1, θ_2) in $\mathbb{R}^3$ sind durch den C^∞-Diffemorphismus

$$\begin{aligned} f : \mathbb{R}_+ \times]0, 2\pi[\times \mathbb{R} \quad &\to \quad \mathbb{R}^3 \setminus ([0, +\infty[\times \{0\} \times \mathbb{R}) , \\ (r, \theta_1, \theta_2) \quad &\mapsto \quad (r \cos(\theta_1) \cos(\theta_2), r \sin(\theta_1) \cos(\theta_2), r \sin(\theta_2)) , \end{aligned}$$

gegeben. □

Im zweiten Teil dieses Abschnitts betrachten wir folgendes Problem. Gegeben seien eine C^1-Abbildung $f = (f_1, \ldots, f_n) : W \to \mathbb{R}^n$ von einer offenen Menge $W \subseteq \mathbb{R}^m \times \mathbb{R}^n$ in $\mathbb{R}^n$ und eine Nullstelle $(a, b) \in W$ von f. Wir fragen:

1. Kann man die Gleichung
$$f(x, y) = 0 ,$$
also das Gleichungssystem
$$f_i(x_1, \ldots, x_m, y_1, \ldots, y_n) = 0 , \quad i = 1, \ldots, n ,$$

in einer Umgebung von (a,b) nach $y = (y_1, \ldots, y_n)$ auflösen? Das heißt, kann man eine Umgebung $W_{(a,b)} \subseteq W$ von (a,b) und eine Abbildung $g : U \to \mathbb{R}^n$ von einer offenen Menge $U \subseteq \mathbb{R}^n$ in $\mathbb{R}^n$ mit

$$\{(x,y) \in W_{(a,b)} : f(x,y) = 0\} = \{(x, g(x)) : x \in U\} ,$$

also mit

$$f^{-1}(\{0\}) \cap W_{(a,b)} = \operatorname{Graph}(g)$$

finden?

2. Ist eine solche Abbildung g differenzierbar?

Die folgenden beiden Beispiele illustrieren typische Situationen, die dabei auftreten können.

Beispiel 2.4.17 Sei $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ durch

$$f(x,y) := x^2 + y^2 - 1$$

definiert. Dann ist

$$f^{-1}(\{0\}) = \left\{(x,y) : |x| \leq 1 \text{ und } |y| = \sqrt{1-x^2}\right\} .$$

Demnach ist

$$f^{-1}(\{0\}) \cap (\mathbb{R} \times]0, +\infty[) = \operatorname{Graph}(g_+) \quad \text{und} \quad f^{-1}(\{0\}) \cap (\mathbb{R} \times]-\infty, 0[) = \operatorname{Graph}(g_-) ,$$

wobei $g_+, g_- :]-1,1[\to \mathbb{R}$ durch

$$g_+(x) := \sqrt{1-x^2} \quad \text{und} \quad g_-(x) := -\sqrt{1-x^2}$$

gegeben sind. Ist also $(a,b) \in f^{-1}(\{0\})$ und $b \neq 0$, so kann man $f(x,y) = 0$ in einer Umgebung vom (a,b) nach y auflösen. In den Fällen $(a,b) = (1,0)$ und $(a,b) = (-1,0)$ ist das aber nicht möglich. □

Beispiel 2.4.18 Sei $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ durch

$$f(x,y) := x - y^3$$

definiert. Dann ist

$$f^{-1}(\{0\}) = \operatorname{Graph}(g) .$$

wobei $g : \mathbb{R} \to \mathbb{R}$ durch

$$g(x) := \operatorname{sgn}(x) \sqrt[3]{|x|}$$

gegeben ist. Aber g ist in 0 nicht differenzierbar. □

Definition 2.4.19 *Sei $W \subseteq \mathbb{R}^m \times \mathbb{R}^n$ offen und sei*

$$f = (f_1, \ldots, f_l) : (x,y) \in W \mapsto f(x,y) = f(x_1, \ldots, x_m, y_1, \ldots, y_n) \in \mathbb{R}^l ,$$

in $(a,b) \in W$ differenzierbar. Die lineare Abbildung

$$\mathrm{D}_x f(a,b) : \mathbb{R}^m \to \mathbb{R}^l , \quad \mathrm{D}_x f(a,b)(v) := \mathrm{D}f(a,b)(v,0) ,$$

heißt das **partielle Differential von f in (a,b) nach x**. *Analog heißt die lineare Abbildung*

$$\mathrm{D}_y f(a,b) : \mathbb{R}^n \to \mathbb{R}^l , \quad \mathrm{D}_y f(a,b)(w) := \mathrm{D}f(a,b)(0,w) ,$$

das **partielle Differential von f in (a,b) nach y**.

Damit ist

$$\mathrm{D}f(a,b)(v,w) = \mathrm{D}_x f(a,b)(v) + \mathrm{D}_y f(a,b)(w) \quad \text{für alle} \quad (v,w) \in \mathbb{R}^m \times \mathbb{R}^n \ .$$

Die Matrixdarstellungen von $\mathrm{D}_x f(a,b)$ und $\mathrm{D}_y f(a,b)$ bezüglich der Standardbasen sind

$$\frac{\partial f}{\partial x}(a,b) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a,b) & \dots & \frac{\partial f_1}{\partial x_m}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_l}{\partial x_1}(a,b) & \dots & \frac{\partial f_l}{\partial x_m}(a,b) \end{pmatrix}$$

bzw.

$$\frac{\partial f}{\partial y}(a,b) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & \dots & \frac{\partial f_1}{\partial y_n}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_l}{\partial y_1}(a,b) & \dots & \frac{\partial f_l}{\partial y_n}(a,b) \end{pmatrix} .$$

Satz 2.4.20 (Satz über implizite Funktionen) *Sei $W \subseteq \mathbb{R}^m \times \mathbb{R}^n$ offen, sei*

$$f : (x,y) \in W \mapsto f(x,y) \in \mathbb{R}^n$$

eine C^k-Abbildung für ein $k \in \mathbb{N}$ und sei $(a,b) \in W$ derart, dass $f(a,b) = 0$ und $\mathrm{D}_y f(a,b)$ ein Isomorphismus ist. Dann existieren eine Umgebung $U_a \subseteq \mathbb{R}^m$ von a, eine Umgebung $V_b \subseteq \mathbb{R}^n$ von b und eine C^k-Abbildung $g : U_a \to V_b$ mit $U_a \times V_b \subseteq W$ und

$$\forall\, (x,y) \in U_a \times V_b : (f(x,y) = 0 \iff y = g(x)) \ . \tag{2.4.14}$$

Ferner kann U_a so gewählt werden, dass

$$\mathrm{D}g(x) = -(\mathrm{D}_y f(x, g(x)))^{-1} \circ \mathrm{D}_x f(x, g(x)) \quad \textit{für alle} \quad x \in U_a \ . \tag{2.4.15}$$

Beweis. Sei $F \in C^k(W, \mathbb{R}^m \times \mathbb{R}^n)$ durch

$$F(x,y) := (x, f(x,y))$$

definiert. Da $\mathrm{D}_y f(a,b)$ ein Isomorphismus ist und da

$$\mathrm{D}F(x,y)(v,w) = (v, \mathrm{D}f(x,y)(v,w)) = (v, \mathrm{D}_x f(x,y)(v) + \mathrm{D}_y f(x,y)(w)) \tag{2.4.16}$$

für $(x,y) \in W$ und $(v,w) \in \mathbb{R}^m \times \mathbb{R}^n$, ist dann auch $\mathrm{D}F(a,b)$ ein Isomorphismus. Also existiert nach Satz 2.4.12 eine Umgebung $W_{(a,b)} \subseteq W$ von (a,b) mit der Eigenschaft, dass $F(W_{(a,b)})$ offen und

$$F\big|_{W_{(a,b)}} : W_{(a,b)} \to F(W_{(a,b)})$$

ein C^k-Diffeomorphimus ist. Sei $G = (G_1, G_2) : F(W_{(a,b)}) \to W_{(a,b)} \subseteq \mathbb{R}^m \times \mathbb{R}^n$ die Umkehrabbildung von $F\big|_{W_{(a,b)}}$. Für $(\xi, \eta) \in F(W_{(a,b)})$ ist

$$(\xi,\eta) = F(G_1(\xi,\eta), G_2(\xi,\eta)) = (G_1(\xi,\eta), f(G_1(\xi,\eta), G_2(\xi,\eta)))$$

und somit

$$G_1(\xi,\eta) = \xi \ .$$

Da für $(x,y) \in W_{(a,b)}$

$$f(x,y) = 0 \iff F(x,y) = (x,0) \iff (x,y) = G(x,0) \ ,$$

folgt

$$\forall (x,y) \in W_{(a,b)} : (f(x,y) = 0 \iff y = G_2(x,0)) \ . \tag{2.4.17}$$

Wegen $G_2(a,0) = b$ und der Stetigkeit von G_2 können wir Umgebungen U_a von a und V_b von b so wählen, dass $U_a \times V_b \subseteq W_{(a,b)}$ und

$$G_2(x,0) \in V_b \quad \text{für alle} \quad x \in U_a \ .$$

Wir definieren $g : U_a \to V_b$ durch

$$g(x) := G_2(x,0) \ .$$

Da G eine C^k-Abbildung ist, ist auch g eine C^k-Abbildung. Außerdem genügt g aufgrund von (2.4.17) der Bedingung (2.4.14). Schließlich erhält man aus

$$f(x,g(x)) = 0 \quad \text{für alle} \quad x \in U_a \ ,$$

dass

$$0 = \mathrm{D}f(x,g(x))(v, \mathrm{D}g(x)(v)) = \mathrm{D}_x f(x,g(x))(v) + \mathrm{D}_y f(x,g(x)) \circ \mathrm{D}g(x)(v) \ ,$$

für alle $v \in \mathbb{R}^m$. Da nach (2.4.16) mit $\mathrm{D}F(x,g(x))$ auch $\mathrm{D}_y f(x,g(x))$ ein Isomorphismus ist, folgt (2.4.15). □

2.5 Extrema mit Nebenbedingungen

Ist $L \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, so sei $\mathrm{rg}(L)$ der Rang und $\mathrm{Ker}(L)$ der Kern von L, d.h.

$$\mathrm{rg}(L) := \dim L(\mathbb{R}^m) \quad \text{und} \quad \mathrm{Ker}(L) := L^{-1}(\{0\}) \ .$$

Satz 2.5.1 *Sei $U \subseteq \mathbb{R}^m$ offen, sei $\Phi \in C^1(U, \mathbb{R}^n)$ und sei $a \in U$ derart, dass $\Phi(a) = 0$ und $\mathrm{rg}(\mathrm{D}\Phi(a)) = n$. Dann existiert zu jedem $v \in \mathrm{Ker}(\mathrm{D}\Phi(a))$ eine C^1-Abbildung $\gamma :]-\varepsilon, \varepsilon[\to U$ mit*

(1) $\Phi(\gamma(t)) = 0$ *für alle* $t \in]-\varepsilon, \varepsilon[$,

(2) $\gamma(0) = a$ *und* $\gamma'(0) = v$.

Beweis. Wir schreiben $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ als (x', x'') mit $x' = (x_1, \ldots, x_{m-n}) \in \mathbb{R}^{m-n}$ und $x'' = (x_{m-n+1}, \ldots, x_m) \in \mathbb{R}^n$. Entsprechend ist $a = (a', a'')$. Wegen $\mathrm{rg}(\mathrm{D}\Phi(a)) = n$ können wir o.B.d.A. annehmen, dass

$$\det\left(\frac{\partial \Phi}{\partial x''}(a)\right) \neq 0 \ .$$

Andernfalls ordne man die Komponenten $x_1, \ldots, x_m$ von x geeignet um. Somit ist das partielle Differential $\mathrm{D}_{x''}\Phi(a)$ ein Isomorphimus. Nach dem Satz über implizite Funktionen gibt es Umgebungen $V' \subseteq \mathbb{R}^{m-n}$ von a' und $V'' \subseteq \mathbb{R}^n$ von a'' und ein C^1-Abbildung $g : V' \to V''$ mit

$$\Phi^{-1}(\{0\}) \cap (V' \times V'') = \{(x', g(x')) : x' \in V'\} \ .$$

Wir definieren $G : V' \to \mathbb{R}^m$ durch

$$G(x') := (x', g(x')) \ .$$

Dann ist $\Phi \circ G = 0$. Folglich gilt

$$\mathrm{D}\Phi(a)(\mathrm{D}G(a')(v')) = \mathrm{D}(\Phi \circ G)(a')(v') = 0 \quad \text{für alle} \quad v' \in \mathbb{R}^{m-n} \ ,$$

d.h.

$$\mathrm{D}G(a')(\mathbb{R}^{m-n}) \subseteq \mathrm{Ker}(\mathrm{D}\Phi(a)) \ . \tag{2.5.1}$$

Da $\mathrm{D}G(a')(v') = (v', \mathrm{D}g(a')(v'))$, ist

$$\mathrm{rg}(\mathrm{D}g(a')) = m - n \ . \tag{2.5.2}$$

Die Voraussetzung $\mathrm{rg}(\mathrm{D}\Phi(a)) = n$ impliziert

$$\dim \mathrm{Ker}(\mathrm{D}\Phi(a)) = m - n \ . \tag{2.5.3}$$

Aus (2.5.1), (2.5.1) und (2.5.3) folgt

$$\mathrm{D}G(a')(\mathbb{R}^{m-n}) = \mathrm{Ker}(\mathrm{D}\Phi(a)) \ .$$

Also existiert zu jedem $v \in \mathrm{Ker}(\mathrm{D}\Phi(a))$ ein $v' \in \mathbb{R}^{m-n}$ mit $v = \mathrm{D}G(a')(v')$. Wir setzen

$$\gamma(t) := G(a' + tv')$$

und erhalten so eine C^1-Abbildung $\gamma : \,]-\varepsilon, \varepsilon[\, \to U$ mit $\Phi \circ \gamma = 0$, $\gamma(0) = G(a') = a$ und $\gamma'(0) = \mathrm{D}G(a')(v') = v$. □

Definition 2.5.2 *Sei (X, d) ein metrischer Raum und seien $f : X \to \mathbb{R}$ und $\Phi : X \to \mathbb{R}^n$. Man sagt:*

(i) f **hat in** $a \in X$ **ein lokales Maximum unter der Nebenbedingung** $\Phi = 0 :\Longleftrightarrow$ $\Phi(a) = 0$ *und es existiert ein* $\varepsilon > 0$ *mit*

$$f(x) \leq f(a) \quad \textit{für alle} \quad x \in B_\varepsilon(a) \cap \Phi^{-1}(\{0\}) \ .$$

(ii) f **hat in** $a \in X$ **ein lokales Minimum unter der Nebenbedingung** $\Phi = 0 :\Longleftrightarrow$ $\Phi(a) = 0$ *und es existiert ein* $\varepsilon > 0$ *mit*

$$f(x) \geq f(a) \quad \textit{für alle} \quad x \in B_\varepsilon(a) \cap \Phi^{-1}(\{0\}) \ .$$

Äquivalent kann man formulieren, dass $f : X \to \mathbb{R}$ genau dann in $a \in X$ ein lokales Maximum (bzw. Minimum) unter der Nebenbedingung $\Phi = 0$ hat, wenn $f\big|_{\Phi^{-1}(\{0\})}$ in a ein lokales Maximum (bzw. Minimum) hat, wobei $\Phi^{-1}(\{0\})$ mit der von d induzierten Metrik versehen ist.

Satz 2.5.3 (Multiplikatorregel von Lagrange) *Sei $U \subseteq \mathbb{R}^m$ offen, sei $f : U \to \mathbb{R}$ differenzierbar, sei $\Phi = (\Phi_1, \ldots, \Phi_n) \in C^1(U, \mathbb{R}^n)$ und sei $a \in U$ derart, dass $\Phi(a) = 0$ und $\mathrm{rg}(\mathrm{D}\Phi(a)) = n$. Ferner habe f in a ein lokales Extremum unter der Nebenbedingung $\Phi = 0$. Dann ist $\mathrm{grad}(f)(a)$ eine Linearkombination der Vektoren $\mathrm{grad}(\Phi_i)(a)$, $i = 1, \ldots, n$, d.h. es existieren $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit*

$$\mathrm{grad}(f)(a) = \sum_{i=1}^{n} \lambda_i \, \mathrm{grad}(\Phi_i)(a) \ . \tag{2.5.4}$$

Beweis. Seien E_1 und E_2 die linearen Hüllen von $\{\mathrm{grad}(f)(a)\}$ und $\{\mathrm{grad}(\Phi_1)(a), \ldots, \mathrm{grad}(\Phi_n)(a)\}$ und seien $E_1^\perp$ und $E_2^\perp$ ihre orthogonalen Komplemente. Zu zeigen ist $E_1 \subseteq E_2$. Da diese Inklusion äquivalent zu $E_2^\perp \subseteq E_1^\perp$ ist, genügt es zu zeigen, dass für $v \in \mathbb{R}^m$

$$\langle \mathrm{grad}(\Phi_i)(a), v \rangle = 0 \ \text{ für } \ i = 1, \ldots, n \quad \Longrightarrow \quad \langle \mathrm{grad}(f)(a), v \rangle = 0 \ ,$$

d.h.

$$\mathrm{D}\Phi(a)(v) = 0 \quad \Longrightarrow \quad \mathrm{D}f(a)(v) = 0$$

gilt.

Sei also $v \in \mathbb{R}^m$ und gelte $\mathrm{D}\Phi(a)(v) = 0$. Da $\mathrm{rg}(\mathrm{D}\Phi(a)) = n$, existiert nach Satz 2.5.1 eine C^1-Abbildung $\gamma : \,]-\varepsilon, \varepsilon[\, \to U$ mit $\Phi \circ \gamma = 0$, $\gamma(0) = a$ und $\gamma'(0) = v$. Die Funktion $F := f \circ \gamma$ hat nach Voraussetzung in 0 ein lokales Extremum. Somit ist $F'(0) = 0$. Da

$$F'(0) = \mathrm{D}f(\gamma(0))(\gamma'(0)) = \mathrm{D}f(a)(v) \ ,$$

folgt $\mathrm{D}f(a)(v) = 0$. □

Bemerkung 2.5.4 (i) Die Zahlen $\lambda_1, \ldots, \lambda_n$ in (2.5.4) werden **Lagrange-Multiplikatoren** genannt.

(ii) Die Multiplikatorregel von Lagrange liefert nur eine notwendige Bedingung für Extrema mit Nebenbedingungen. Man muss folglich extra untersuchen, ob in den mit Hilfe von Satz 2.5.3 ermittelten Punkten tatsächlich Extrema vorliegen. □

Beispiel 2.5.5 Wir wollen die Extrema der Einschränkung von

$$f : \mathbb{R}^2 \to \mathbb{R}\,, \quad f(x) := x_1 x_2\ ,$$

auf $S^1 := \left\{x \in \mathbb{R}^2 : \|x\| = 1\right\}$ bestimmen. Offensichtlich ist S^1 die Nullstellenmenge der C^∞-Abbildung

$$\Phi : \mathbb{R}^2 \to \mathbb{R}\ , \quad \Phi(x) := x_1^2 + x_2^2 - 1\ .$$

Somit sind die Extrema von f unter der Nebenbedingung $\Phi = 0$ zu berechnen.

Habe f in $a \in S^1$ ein Extremum unter der Nebenbedingung $\Phi = 0$. Da $\operatorname{grad}(\Phi)(x) \neq 0$ und folglich $\operatorname{rg}(\mathrm{D}\Phi(x)) = 1$ für alle $x \in S^1$, können wir Satz 2.5.3 anwenden. Demnach existiert ein $\lambda \in \mathbb{R}$ mit

$$\operatorname{grad}(f)(a) = \lambda \operatorname{grad}(\Phi)(a)\ .$$

Also gilt

$$a_1 = 2\lambda a_2\ , \quad a_2 = 2\lambda a_1 \quad \text{und} \quad a_1^2 + a_2^2 = 1\ .$$

Das ist genau dann der Fall, wenn

$$|\lambda| = \frac{1}{2} \quad \text{und} \quad |a_1| = |a_2| = \frac{1}{\sqrt{2}}\ .$$

Somit ist

$$a \in C := \left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)\right\}\ .$$

Aufgrund der Kompaktheit von S^1 nimmt die Abbildung $f\big|_{S^1}$ ihr Minimum und ihr Maximum an, d.h. es existieren $\xi_1, \xi_2 \in S^1$ mit

$$f(\xi_1) = \max\left\{f(x) : x \in S^1\right\} \quad \text{und} \quad f(\xi_2) = \min\left\{f(x) : x \in S^1\right\}\ .$$

Da $\xi_1, \xi_2 \in C$ und $f(C) = \{1/2, -1/2\}$, erhalten wir

$$\max\left\{f(x) : x \in S^1\right\} = \frac{1}{2} \quad \text{und} \quad \min\left\{f(x) : x \in S^1\right\} = -\frac{1}{2}\ .$$

Außerdem sehen wir, dass das Maximum in den Punkten $\left(1/\sqrt{2}, 1/\sqrt{2}\right)$ und $\left(-1/\sqrt{2}, -1/\sqrt{2}\right)$ und das Minimum in den Punkten $\left(1/\sqrt{2}, -1/\sqrt{2}\right)$ und $\left(-1/\sqrt{2}, 1/\sqrt{2}\right)$ angenommen wird. □

Kapitel 3

Integralrechnung

3.1 Das unbestimmte Integral

Sei im Folgenden $I \subset \mathbb{R}$ irgendein Intervall und $E = \mathbb{K}^n$ mit $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Definition 3.1.1 *Sei $f : I \to E$. Eine Abbildung $F : I \to E$ heißt* **Stammfunktion von** f $:\Longleftrightarrow$ *F ist differenzierbar und $F' = f$. Die Menge aller Stammfunktionen von f wird das* **unbestimmte Integral von** f *genannt und mit $\int f(x)\,\mathrm{d}x$ bezeichnet.*

Satz 3.1.2 *Für jede Abbildung $f : I \to E$ gilt:*

(i) *Ist F eine Stammfunktion von f und ist $c \in E$, so ist auch*

$$F + c : I \to E\ , \quad (F + c)(x) := F(x) + c\ ,$$

eine Stammfunktion von f.

(ii) *Sind F_1 und F_2 Stammfunktionen von f, so ist $F_1 - F_2$ konstant.*

Beweis. (i) Mit F ist auch $F + c$ differenzierbar und

$$(F + c)' = F' = f\ .$$

(ii) Seien F_1 und F_2 Stammfunktionen von f. Dann ist $F_1 - F_2$ differenzierbar und

$$(F_1 - F_2)' = F_1' - F_2' = f - f = 0\ .$$

Nach Satz 1.2.6(v) folgt, dass $F_1 - F_2$ konstant ist. □

Satz 3.1.2 rechtfertigt die folgende Schreibweise. Ist F eine Stammfunktion von $f : I \to E$, so schreibt man

$$\int f(x)\,\mathrm{d}x = F(x) + c\ .$$

Im folgenden Beispiel sind Grundintegrale zusammengestellt. Zum Nachweis leite man jeweils die Stammfunktion ab.

Beispiel 3.1.3 (i) Sei $\alpha \in \mathbb{R}$ und $\alpha \neq -1$. Dann ist

$$\int x^\alpha\,\mathrm{d}x = \frac{x^{\alpha+1}}{\alpha + 1} + c\ ,$$

wobei für I folgende Einschränkungen gelten. Ist $\alpha \in \mathbb{Z}$ und $\alpha \leq -2$, so sei $0 \notin I$. Ist $\alpha \notin \mathbb{Z}$, so wird $I \subseteq \mathbb{R}_+$ vorausgesetzt.

(ii) Falls $0 \notin I$, ist

$$\int \frac{1}{x}\,\mathrm{d}x = \ln(|x|) + c\ .$$

(iii) $\displaystyle\int \exp(x)\,\mathrm{d}x = \exp(x) + c\ .$

(iv) $\displaystyle\int \cos(x)\,\mathrm{d}x = \sin(x) + c \quad \text{und} \quad \int \sin(x)\,\mathrm{d}x = -\cos(x) + c\ .$

(v) $\displaystyle\int \cosh(x)\,\mathrm{d}x = \sinh(x) + c \quad \text{und} \quad \int \sinh(x)\,\mathrm{d}x = \cosh(x) + c\ .$

(vi) $\displaystyle\int \frac{1}{x^2+1}\,\mathrm{d}x = \arctan(x) + c\ .$

(vii) Ist $I \subseteq\ ]-1,1[$, so gilt

$$\int \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \arcsin(x) + c\ .$$

□

Wir geben jetzt die wichtigsten Rechenregeln der unbestimmten Integration und einige Beispiele dazu an.

Satz 3.1.4 *Seien $f_1, f_2 : I \to E$ und seien $\lambda_1, \lambda_2 \in \mathbb{K}$. Ist F_1 eine Stammfunktion von f_1 und F_2 eine Stammfunktion von f_2, so ist $\lambda_1 F_1 + \lambda_2 F_2$ eine Stammfunktion von $\lambda_1 f_1 + \lambda_2 f_2$, d.h.*

$$\int (\lambda_1 f_1(x) + \lambda_2 f_2(x))\,\mathrm{d}x = \lambda_1 \int f_1(x)\,\mathrm{d}x + \lambda_2 \int f_2(x)\,\mathrm{d}x\ .$$

Beweis. Das folgt aus Satz 1.1.9. □

Beispiel 3.1.5 Nach Beispiel 3.1.3 und Satz 3.1.4 ist

$$\int \left(4x^2 + 3 - \frac{2}{x}\right)\mathrm{d}x = \frac{4}{3}x^3 + 3x - 2\ln(|x|) + c$$

für $I =]-\infty, 0[$ und $I =]0, +\infty[$. □

Satz 3.1.6 (Partielle Integration) *Seien $f_1, f_2 : I \to \mathbb{K}$ differenzierbar. Ist F eine Stammfunktion von $f_1 f_2'$, so ist $f_1 f_2 - F$ eine Stammfunktion von $f_1' f_2$, d.h.*

$$\int f_1'(x) f_2(x)\,\mathrm{d}x = f_1(x) f_2(x) - \int f_1(x) f_2'(x)\,\mathrm{d}x\ .$$

Beweis. Sei F eine Stammfunktion von $f_1 f_2'$. Nach der Produktregel gilt dann

$$(f_1 f_2 - F)' = f_1' f_2 + f_1 f_2' - F' = f_1' f_2\ .$$

Das liefert die Behauptung. □

Beispiel 3.1.7 (i) Es ist

$$\int \exp(x)\,x\,\mathrm{d}x = \int \exp'(x)\,x\,\mathrm{d}x = \exp(x)\,x - \int \exp(x)\,\mathrm{d}x = \exp(x)\,x - \exp(x) + c\ .$$

(ii) Aus

$$\begin{aligned}\int \sin^2(x)\,\mathrm{d}x &= -\int \cos'(x)\sin(x)\,\mathrm{d}x\\ &= -\cos(x)\sin(x) + \int \cos^2(x)\,\mathrm{d}x\\ &= -\cos(x)\sin(x) + \int \left(1-\sin^2(x)\right)\mathrm{d}x\\ &= -\cos(x)\sin(x) + x - \int \sin^2(x)\,\mathrm{d}x\end{aligned}$$

erhält man

$$\int \sin^2(x)\,\mathrm{d}x = \frac{1}{2}(x - \cos(x)\sin(x)) + c\ .$$

□

Satz 3.1.8 (Substitutionsregel) *Seien $I_1, I_2 \subseteq \mathbb{R}$ Intervalle, sei $\varphi : I_1 \to I_2$ differenzierbar und sei $f : I_2 \to E$. Ist F eine Stammfunktion von f, so ist $F \circ \varphi$ eine Stammfunktion von $(f \circ \varphi)\varphi'$, d.h.*

$$\int f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int f(y)\,\mathrm{d}y\Big|_{y=\varphi(x)}\ .$$

Beweis. Sei F eine Stammfunktion von f. Nach der Kettenregel gilt dann

$$(F\circ\varphi)' = (F'\circ\varphi)\varphi' = (f\circ\varphi)\varphi'\ .$$

Das ergibt die Behauptung. □

Beispiel 3.1.9 (i) Sei $\varphi : I \to \mathbb{R}$ differenzierbar und gelte $\varphi(x) \neq 0$ für alle $x \in I$. Dann ist

$$\int \frac{\varphi'(x)}{\varphi(x)}\,\mathrm{d}x = \int \frac{1}{y}\,\mathrm{d}y\Big|_{y=\varphi(x)} = \ln(|y|)\big|_{y=\varphi(x)} = \ln(|\varphi(x)|)\ .$$

Zum Beispiel gilt

$$\int \frac{x}{x^2+1}\,\mathrm{d}x = \frac{1}{2}\ln\left(x^2+1\right) + c\ . \tag{3.1.1}$$

(ii) Sei $f : I \to E$ und seien $a, b \in \mathbb{R}$ mit $a \neq 0$. Indem man $\varphi(x) := ax + b$ substituiert, erhält man

$$\int f(ax+b)\,\mathrm{d}x = \frac{1}{a}\int f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \frac{1}{a}\int f(y)\,\mathrm{d}y\Big|_{y=ax+b}\ .$$

Insbesondere ist

$$\int \frac{1}{(x+b)^k}\,\mathrm{d}x = -\frac{1}{k-1}\,\frac{1}{(x+b)^{k-1}} + c \quad \text{für} \quad k = 2, 3, \ldots \tag{3.1.2}$$

und

$$\int \frac{1}{x+b}\,\mathrm{d}x = \ln(|x+b|) + c \tag{3.1.3}$$

auf $I = \left]-\infty, -b\right[$ und $I = \left]-b, +\infty\right[$. □

Zur Formulierung des nächsten Beispiels führen wir die die folgenden Funktionen ein.

Definition 3.1.10 *Der* **Areasinus hyperbolicus** $\operatorname{arsinh} : \mathbb{R} \to \mathbb{R}$ *ist die inverse Funktion von* $\sinh : \mathbb{R} \to \mathbb{R}$. *Der* **Areacosinus hyperbolicus** $\operatorname{arcosh} : [1, +\infty[\to [0, +\infty[$ *ist die inverse Funktion von*

$$x \in [0, +\infty[\mapsto \cosh(x) \in [1, +\infty[\ .$$

Lemma 3.1.11 *Es gilt:*

(i) $\operatorname{arsinh}(x) = \ln\left(x + \sqrt{x^2+1}\right)$ *für* $x \in \mathbb{R}$,

(ii) $\operatorname{arcosh}(x) = \ln\left(x + \sqrt{x^2-1}\right)$ *für* $x \geq 1$.

Beweis. Übung. $\square$

Beispiel 3.1.12 (i) Mittels der Substitution $x = \sinh(t)$ berechnen wir

$$\int \frac{1}{\sqrt{x^2+1}}\,\mathrm{d}x\bigg|_{x=\sinh(t)} = \int \frac{1}{\sqrt{\sinh^2(t)+1}}\,\sinh'(t)\,\mathrm{d}t = \int \frac{\cosh(t)}{\cosh(t)}\,\mathrm{d}t = t + c\,.$$

Folglich ist

$$\int \frac{1}{\sqrt{x^2+1}}\,\mathrm{d}x = \operatorname{arsinh}(x) + c\,.$$

(ii) Sei $I \subseteq \,]1, +\infty[$. Die Substitution $x = \cosh(t)$ mit $t > 0$ liefert dann

$$\int \frac{1}{\sqrt{x^2-1}}\,\mathrm{d}x\bigg|_{x=\cosh(t)} = \int \frac{1}{\sqrt{\cosh^2(t)-1}}\,\cosh'(t)\,\mathrm{d}t = \int \frac{\sinh(t)}{\sinh(t)}\,\mathrm{d}t = t + c\,,$$

also

$$\int \frac{1}{\sqrt{x^2-1}}\,\mathrm{d}x = \operatorname{arcosh}(x) + c\,,$$

$\square$

Zur Integration der rationalen Funktionen verwendet man

Satz 3.1.13 (Partialbruchzerlegung) *Seien* $p(z) = a_k z^k + a_{k-1}z^{k-1} + \cdots + a_1 z + a_0$ *und* $q(z) = z^l + b_{l-1}z^{l-1} + \cdots + b_1 z + b_0$ *zwei Polynome und gelte*

$$q(z) = (z-\lambda_1)^{l_1}(z-\lambda_2)^{l_2}\cdots(z-\lambda_m)^{l_m} \quad \textit{für} \quad z \in \mathbb{C}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_m \in \mathbb{C}$. *Ist* $k < l$, *so existieren Zahlen* $\mu_{ij} \in \mathbb{C}$ *derart, dass*

$$\begin{aligned}\frac{p(z)}{q(z)} = &\frac{\mu_{11}}{z-\lambda_1} + \frac{\mu_{12}}{(z-\lambda_1)^2} + \cdots + \frac{\mu_{1l_1}}{(z-\lambda_1)^{l_1}}\\ &+ \frac{\mu_{21}}{z-\lambda_2} + \frac{\mu_{22}}{(z-\lambda_2)^2} + \cdots + \frac{\mu_{2l_2}}{(z-\lambda_2)^{l_2}}\\ &+ \cdots\\ &+ \frac{\mu_{m1}}{z-\lambda_m} + \frac{\mu_{m2}}{(z-\lambda_m)^2} + \cdots + \frac{\mu_{ml_m}}{(z-\lambda_m)^{l_m}}\end{aligned}$$

für $z \in \mathbb{C}$ *mit* $q(z) \neq 0$.

Beweis. Vgl. Heuser, Lehrbuch der Analysis, Teil 1, Abschnitt 69. $\square$

Beispiel 3.1.14 (i) Wir berechnen

$$\int \frac{3x+1}{(x-1)(x+1)^2}\,\mathrm{d}x\,,$$

wobei $1, -1 \notin I$. Nach Satz 3.1.13 gibt es Zahlen α, β_1, β_2 mit

$$\frac{3x+1}{(x-1)(x+1)^2} = \frac{\alpha}{x-1} + \frac{\beta_1}{x+1} + \frac{\beta_2}{(x+1)^2} \quad \text{für} \quad x \in \mathbb{R} \setminus \{1, -1\} ,$$

d.h. mit

$$3x+1 = \alpha(x+1)^2 + \beta_1(x-1)(x+1) + \beta_2(x-1)$$

bzw.

$$3x+1 = (\alpha+\beta_1)x^2 + (2\alpha+\beta_2)x + \alpha - \beta_1 - \beta_2$$

für $x \in \mathbb{R}$. Ein Vergleich der Koeffizienten liefert $\alpha = \beta_2 = 1$ und $\beta_1 = -1$. Mit (3.1.2) und (3.1.3) folgt

$$\begin{aligned} \int \frac{3x+1}{(x-1)(x+1)^2}\,\mathrm{d}x &= \int \frac{1}{x-1}\,\mathrm{d}x - \int \frac{1}{x+1}\,\mathrm{d}x + \int \frac{1}{(x+1)^2}\,\mathrm{d}x \\ &= \ln(|x-1|) - \ln(|x+1|) - \frac{1}{x+1} + c . \end{aligned}$$

(ii) Zur Berechnung von

$$\int \frac{x^2-x+1}{(x-1)(x^2+1)}\,\mathrm{d}x$$

auf einem Intervall I mit $1 \notin I$ bestimmen wir $\alpha, \beta_1, \beta_2 \in \mathbb{C}$ mit

$$\frac{x^2-x+1}{(x-1)(x^2+1)} = \frac{\alpha}{x-1} + \frac{\beta_1}{x-\mathrm{i}} + \frac{\beta_2}{x+\mathrm{i}} \quad \text{für} \quad x \in \mathbb{R} \setminus \{1\} ,$$

also mit

$$x^2-x+1 = \alpha(x-\mathrm{i})(x+\mathrm{i}) + \beta_1(x-1)(x+\mathrm{i}) + \beta_2(x-1)(x-\mathrm{i}) \quad \text{für} \quad x \in \mathbb{R} .$$

Indem wir in der letzten Gleichung nacheinander $x = 1$, $x = \mathrm{i}$ und $x = -\mathrm{i}$ einsetzen, erhalten wir

$$1 = 2\alpha , \quad -\mathrm{i} = 2\mathrm{i}(\mathrm{i}-1)\beta_1 \quad \text{und} \quad \mathrm{i} = 2\mathrm{i}(\mathrm{i}+1)\beta_2 ,$$

d.h. $\alpha = 1/2$, $\beta_1 = (1+\mathrm{i})/4$ und $\beta_2 = (1-\mathrm{i})/4$. Somit ist

$$\frac{x^2-x+1}{(x-1)(x^2+1)} = \frac{1}{2}\,\frac{1}{x-1} + \frac{1+\mathrm{i}}{4}\,\frac{1}{x-\mathrm{i}} + \frac{1-\mathrm{i}}{4}\,\frac{1}{x+\mathrm{i}} = \frac{1}{2}\left(\frac{1}{x-1} + \frac{x-1}{x^2+1}\right) .$$

Mit (3.1.1) und (3.1.3) folgt

$$\begin{aligned} \int \frac{x^2-x+1}{(x-1)(x^2+1)}\,\mathrm{d}x &= \frac{1}{2}\left(\int \frac{1}{x-1}\,\mathrm{d}x + \int \frac{x}{x^2+1}\,\mathrm{d}x - \int \frac{1}{x^2+1}\,\mathrm{d}x\right) \\ &= \frac{1}{4}\left(2\ln(|x-1|) + \ln\left(x^2+1\right) - 2\arctan(x)\right) + c . \end{aligned}$$

□

Wie im nächsten Beispiel angedeutet, lassen sich in einer Reihe von Fällen Integrale durch geeignete Substitutionen auf Integrale rationaler Funktionen zurückführen.

Beispiel 3.1.15 (i) Die Substitution $y = \exp(x)$ führt auf

$$\begin{aligned} \int \frac{1}{\cosh(x)}\,\mathrm{d}x &= \int \frac{2}{\exp(x)+\exp(-x)}\,\mathrm{d}x = \int \frac{2\exp'(x)}{\exp^2(x)+1}\,\mathrm{d}x \\ &= \int \frac{2}{y^2+1}\,\mathrm{d}y\bigg|_{y=\exp(x)} = 2\arctan(\exp(x)) + c . \end{aligned}$$

(ii) Mit der Substitution $y = \cos(x)/2$ erhält man

$$\int \frac{\sin(x)}{\cos^2(x)+4}\,\mathrm{d}x = -\int \frac{\cos'(x)}{\cos^2(x)+4}\,\mathrm{d}x = -\frac{1}{2}\int \frac{1}{y^2+1}\,\mathrm{d}y\bigg|_{y=\cos(x)/2}$$
$$= -\frac{1}{2}\arctan\left(\frac{\cos(x)}{2}\right) + c\,.$$

□

Bemerkung 3.1.16 Nicht in jedem Fall ist die Stammfunktion einer elementaren Funktion wieder eine elementare Funktion. (Eine Funktion $f : I \to \mathbb{R}$ heißt **elementar**, wenn sie sich in endlich vielen Schritten mittels Addition, Multiplikation, Division, Verkettung und Invertierung aus Konstanten und den Funktionen $\mathrm{id}_{\mathbb{R}}$, $\exp : \mathbb{R} \to \mathbb{R}$ und $\sin : \mathbb{R} \to \mathbb{R}$ bilden lässt.) Zum Beispiel ist

$$\int \frac{1}{\sqrt{(1-x^2)\,(1-\alpha x^2)}}\,\mathrm{d}x\,, \tag{3.1.4}$$

wobei $I \subseteq\,]-1,1[$ und $\alpha \in\,]0,1[$, keine elementare Funktion. Integrale der Form (3.1.4) werden **elliptische Integrale erster Art** genannt. □

3.2 Das Riemannsche Integral

Der Ausgangspunkt für die anschließenden Überlegungen ist die folgende Frage. Sei $f : [a,b] \to \mathbb{R}$ beschränkt und gelte $f(x) \geq 0$ für alle $x \in [a,b]$. Unter welchen Bedingungen und in welcher Weise kann man für die so genannte **Ordinatenmenge**

$$\mathrm{Ord}(f) := \big\{(x,y) \in \mathbb{R}^2 : a \leq x \leq b \text{ und } 0 \leq y \leq f(x)\big\}$$

einen Flächeninhalt definieren? Der Ansatz zur Beantwortung dieser Frage ist die Approximation von $\mathrm{Ord}(f)$ durch Rechtecke.

Definition 3.2.1 *Eine endliche Menge Z heißt Zerlegung von $[a,b]$ $:\Longleftrightarrow$ $Z \subseteq [a,b]$ und $a, b \in Z$.*

Ist $Z = \{t_0, t_1, \ldots, t_k\}$ eine Zerlegung von $[a,b]$, so nehmen wir im Weiteren stets

$$a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b$$

an. Die Menge aller Zerlegungen von $[a,b]$ bezeichnen wir mit $\mathfrak{Z}([a,b])$.

Definition 3.2.2 *Eine Zerlegung $Z' \in \mathfrak{Z}([a,b])$ heißt* **Verfeinerung** *von $Z \in \mathfrak{Z}([a,b])$ $:\Longleftrightarrow$ $Z \subseteq Z'$.*

Lemma 3.2.3 *Zu je zwei Zerlegungen $Z_1, Z_2 \in \mathfrak{Z}([a,b])$ gibt es eine gemeinsame Verfeinerung Z'.*

Beweis. Seien $Z_1, Z_2 \in \mathfrak{Z}([a,b])$. Dann ist $Z' := Z_1 \cup Z_2 \in \mathfrak{Z}([a,b])$ eine Verfeinerung sowohl von Z_1 als auch von Z_2. □

Definition 3.2.4 *Sei $f : [a,b] \to \mathbb{R}$ beschränkt und sei $Z = \{t_0, t_1, \ldots, t_k\} \in \mathfrak{Z}([a,b])$. Dann heißt*

$$\underline{S}(f,Z) := \sum_{i=1}^{k} (\inf f([t_{i-1}, t_i]))\,(t_i - t_{i-1})$$

die **untere Summe** *von f bezüglich Z und*

$$\overline{S}(f,Z) := \sum_{i=1}^{k} (\sup f([t_{i-1}, t_i]))\,(t_i - t_{i-1})$$

die **obere Summe** *von f bezüglich Z.*

Lemma 3.2.5 *Sei $f : [a,b] \to \mathbb{R}$ beschränkt.*

(i) *Ist $Z \in \mathfrak{Z}([a,b])$ und ist Z' eine Verfeinerung von Z, so gilt*

$$\underline{S}(f,Z) \leq \underline{S}(f,Z') \quad \textit{und} \quad \overline{S}(f,Z) \geq \overline{S}(f,Z') \;.$$

(ii) *Für alle $Z_1, Z_2 \in \mathfrak{Z}([a,b])$ gilt*

$$\underline{S}(f,Z_1) \leq \overline{S}(f,Z_2) \;.$$

Beweis. (i) Sei $Z = \{t_0, \ldots, t_k\} \in \mathfrak{Z}([a,b])$, sei $Z' = \{t'_0, \ldots, t'_l\}$ eine Verfeinerung von Z und seien $\sigma(0), \ldots, \sigma(k) \in \{0, \ldots, l\}$ durch

$$t_i = t'_{\sigma(i)}$$

bestimmt. Ist $\sigma(i-1) < j \leq \sigma(i)$, so ist $[t'_{j-1}, t'_j] \subseteq [t_{i-1}, t_i]$ und somit

$$\inf f([t'_{j-1}, t'_j]) \geq \inf f([t_{i-1}, t_i]) \quad \text{und} \quad \sup f([t'_{j-1}, t'_j]) \leq \sup f([t_{i-1}, t_i]) \;.$$

Es folgt

$$\begin{aligned}
\underline{S}(f,Z') &= \sum_{j=1}^{l} (\inf f([t'_{j-1}, t'_j]))\, (t'_j - t'_{j-1}) \\
&= \sum_{i=1}^{k} \sum_{j=\sigma(i-1)+1}^{\sigma(i)} (\inf f([t'_{j-1}, t'_j]))\, (t'_j - t'_{j-1}) \\
&\geq \sum_{i=1}^{k} \sum_{j=\sigma(i-1)+1}^{\sigma(i)} (\inf f([t_{i-1}, t_i]))\, (t'_j - t'_{j-1}) \\
&= \sum_{i=1}^{k} (\inf f([t_{i-1}, t_i])) \sum_{j=\sigma(i-1)+1}^{\sigma(i)} (t'_j - t'_{j-1}) \\
&= \sum_{i=1}^{k} (\inf f([t_{i-1}, t_i]))\, (t_i - t_{i-1}) \\
&= \underline{S}(f,Z)
\end{aligned}$$

und analog

$$\overline{S}(f,Z') \leq \overline{S}(f,Z) \;.$$

(ii) Seien $Z_1, Z_2 \in \mathfrak{Z}([a,b])$ und sei Z' eine gemeinsame Verfeinerung von Z_1 und Z_2. Nach (i) gilt dann

$$\underline{S}(f,Z_1) \leq \underline{S}(f,Z') \leq \overline{S}(f,Z') \leq \overline{S}(f,Z_2) \;.$$

$\square$

Definition 3.2.6 *Sei $f : [a,b] \to \mathbb{R}$ beschränkt. Das* **untere Riemann-Integral** *von f ist*

$$J_*(f) := \sup \left\{ \underline{S}(f,Z) : Z \in \mathfrak{Z}([a,b]) \right\} \;.$$

Das **obere Riemann-Integral** *von f ist*

$$J^*(f) := \inf \left\{ \overline{S}(f,Z) : Z \in \mathfrak{Z}([a,b]) \right\} \;.$$

Lemma 3.2.7 *Für jede beschränkte Funktion $f : [a,b] \to \mathbb{R}$ ist*

$$J_*(f) \leq J^*(f) \;.$$

Beweis. Das ist eine Konsequenz von Lemma 3.2.5(ii). □

Definition 3.2.8 (i) *Eine Funktion* $f : [a, b] \to \mathbb{R}$ *heißt* **Riemann-integrierbar** $:\Longleftrightarrow$ f *ist beschränkt und* $J_*(f) = J^*(f)$.

(ii) *Ist* $f : [a, b] \to \mathbb{R}$ *Riemann-integrierbar, so wird*

$$\int_a^b f(x)\,\mathrm{d}x := J_*(f) = J^*(f)$$

das **Riemann-Integral** *oder das* **bestimmte Integral** *von* f *genannt.*

Die eingangs gestellte Frage beantworten wir jetzt wie folgt.

Definition 3.2.9 *Sei* $f : [a, b] \to \mathbb{R}$ *beschränkt und gelte* $f(x) \geq 0$ *für alle* $x \in [a, b]$*. Wir sagen,* $\mathrm{Ord}(f)$ **hat einen Flächeninhalt** $:\Longleftrightarrow$ f *ist Riemann-integrierbar. Ist dies der Fall, so ist*

$$|\,\mathrm{Ord}(f)| := \int_a^b f(x)\,\mathrm{d}x$$

der **Flächeninhalt** *von* $\mathrm{Ord}(f)$.

Beispiel 3.2.10 Jede konstante Funktion $f : [a, b] \to \mathbb{R}$ ist Riemann-integrierbar. Gilt nämlich $f(x) = \alpha$ für alle $x \in [a, b]$, so ist

$$\underline{S}(f, Z) = \overline{S}(f, Z) = \alpha\,(b - a) \quad \text{für alle} \quad Z \in \mathfrak{Z}([a, b])$$

Somit ist $J_*(f) = J^*(f)$ und

$$\int_a^b \alpha\,\mathrm{d}x = \alpha\,(b - a)\ .$$

□

Beispiel 3.2.11 Die Dirichlet-Funktion auf $[a, b]$, d.h. die Funktion

$$f : [a, b] \to \mathbb{R}\ , \quad f(x) := \begin{cases} 1 & \text{für} \quad x \in [a, b] \cap \mathbb{Q} \\ 0 & \text{für} \quad x \in [a, b] \setminus \mathbb{Q} \end{cases}\ ,$$

ist nicht Riemann-integrierbar, denn

$$\underline{S}(f, Z) = 0 \quad \text{und} \quad \overline{S}(f, Z) = b - a \quad \text{für alle} \quad Z \in \mathfrak{Z}([a, b])\ ,$$

woraus $J_*(f) = 0$ und $J^*(f) = b - a$, also insbesondere $J_*(f) \neq J^*(f)$ folgt. □

Satz 3.2.12 (Riemannsches Integrierbarkeitskriterium) *Sei* $f : [a, b] \to \mathbb{R}$ *beschränkt. Dann ist* f *genau dann Riemann-integrierbar, wenn*

$$\forall\,\varepsilon > 0\,\exists\,Z_1, Z_2 \in \mathfrak{Z}([a, b]) : \overline{S}(f, Z_2) - \underline{S}(f, Z_1) < \varepsilon\ . \tag{3.2.1}$$

Beweis. ($\Longrightarrow$) Sei f Riemann-integrierbar und sei $\varepsilon > 0$. Nach Definition von $J_*(F)$ und $J^*(F)$ existieren $Z_1, Z_2 \in \mathfrak{Z}([a, b])$ mit

$$\underline{S}(f, Z_1) > J_*(F) - \frac{\varepsilon}{2} \quad \text{und} \quad \overline{S}(f, Z_2) < J^*(F) + \frac{\varepsilon}{2}\ .$$

Wegen $J_*(F) = J^*(F)$ folgt

$$\overline{S}(f, Z_2) - \underline{S}(f, Z_1) < J^*(F) + \frac{\varepsilon}{2} - J_*(F) + \frac{\varepsilon}{2} = \varepsilon\ .$$

($\Longleftarrow$) Gelte (3.2.1). Da

$$\underline{S}(f, Z_1) \le J_*(F) \le J^*(F) \le \overline{S}(f, Z_2) \quad \text{für alle} \quad Z_1, Z_2 \in \mathfrak{Z}([a,b]) \, ,$$

folgt

$$0 \le J^*(F) - J_*(F) < \varepsilon \quad \text{für alle} \quad \varepsilon > 0 \, .$$

Also ist $J_*(F) = J^*(F)$, d.h. f ist Riemann-integrierbar. □

Folgerung 3.2.13 *Sei $f : [a,b] \to \mathbb{R}$ beschränkt und existiere eine Folge (Z_k) in $\mathfrak{Z}([a,b])$ mit*

$$\lim_{k\to\infty} \big(\overline{S}(f, Z_k) - \underline{S}(f, Z_k)\big) = 0 \, .$$

Dann ist f Riemann-integrierbar und

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{k\to\infty} \underline{S}(f, Z_k) = \lim_{k\to\infty} \overline{S}(f, Z_k) \, . \tag{3.2.2}$$

Beweis. Die Riemann-Integrierbarkeit von f erhält man mit Satz 3.2.12, die Beziehung (3.2.2) mit

$$|J_*(f) - \underline{S}(f, Z_k)| = J_*(f) - \underline{S}(f, Z_k) \le \overline{S}(f, Z_k) - \underline{S}(f, Z_k)$$

und

$$\big|J^*(f) - \overline{S}(f, Z_k)\big| = \overline{S}(f, Z_k) - J^*(f) \le \overline{S}(f, Z_k) - \underline{S}(f, Z_k) \, .$$

□

Beispiel 3.2.14 Die unteren und oberen Summen von $\mathrm{id}_{[a,b]}$ bezüglich der Zerlegungen

$$Z_k := \left\{ a + \frac{i}{k}\,(b-a) : i = 0, \ldots, n \right\} \, , \quad k \in \mathbb{N} \, ,$$

sind

$$\begin{aligned}
\underline{S}\left(\mathrm{id}_{[a,b]}, Z_k\right) &= \sum_{i=1}^{k} \left(\inf \left[a + \frac{i-1}{k}\,(b-a), a + \frac{i}{k}\,(b-a) \right] \right) \frac{b-a}{k} \\
&= \sum_{i=1}^{k} \left(a + \frac{i-1}{k}\,(b-a) \right) \frac{b-a}{k} = a(b-a) + \frac{k-1}{2k}\,(b-a)^2
\end{aligned}$$

und

$$\begin{aligned}
\overline{S}\left(\mathrm{id}_{[a,b]}, Z_k\right) &= \sum_{i=1}^{k} \left(\sup \left[a + \frac{i-1}{k}\,(b-a), a + \frac{i}{k}\,(b-a) \right] \right) \frac{b-a}{k} \\
&= \sum_{i=1}^{k} \left(a + \frac{i}{k}\,(b-a) \right) \frac{b-a}{k} = a(b-a) + \frac{k+1}{2k}\,(b-a)^2 \, .
\end{aligned}$$

Damit gilt

$$\lim_{k\to\infty} \underline{S}\left(\mathrm{id}_{[a,b]}, Z_k\right) = \lim_{k\to\infty} \overline{S}\left(\mathrm{id}_{[a,b]}, Z_k\right) = a(b-a) + \frac{(b-a)^2}{2} = \frac{b^2-a^2}{2} \, ,$$

was nach Folgerung 3.2.13 die Riemann-Integrierbarkeit von $\mathrm{id}_{[a,b]}$ und

$$\int_a^b x\,\mathrm{d}x = \frac{b^2-a^2}{2}$$

impliziert. □

Beispiel 3.2.15 Sei $a < c < b$, seien $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ und sei $f : [a,b] \to \mathbb{R}$ durch

$$f(x) := \begin{cases} \alpha_1 & \text{für} \quad x \in [a,c[\\ \alpha_2 & \text{für} \quad x = c \\ \alpha_3 & \text{für} \quad x \in]c,b] \end{cases}$$

definiert. Für $k \in \mathbb{N}$ setzen wir

$$Z_k := \left\{a, c - \frac{1}{k}, c + \frac{1}{k}, b\right\} .$$

Ist k genügend groß, so ist $Z_k \in \mathfrak{Z}([a,b])$ und es gilt

$$\underline{S}(f, Z_k) = \alpha_1 \left(c - a - \frac{1}{k}\right) + (\min\{\alpha_1, \alpha_2, \alpha_3\}) \frac{2}{k} + \alpha_3 \left(b - c - \frac{1}{k}\right)$$

und

$$\overline{S}(f, Z_k) = \alpha_1 \left(c - a - \frac{1}{k}\right) + (\max\{\alpha_1, \alpha_2, \alpha_3\}) \frac{2}{k} + \alpha_3 \left(b - c - \frac{1}{k}\right) .$$

Es folgt

$$\lim_{k\to\infty} \underline{S}(f, Z_k) = \lim_{k\to\infty} \overline{S}(f, Z_k) = \alpha_1(c-a) + \alpha_3(b-c) .$$

Somit ist f Riemann-integrierbar und

$$\int_a^b f(x)\,\mathrm{d}x = \alpha_1(c-a) + \alpha_3(b-c) .$$

$\square$

Wir wollen zeigen, dass jede stetige Funktion $f : [a,b] \to \mathbb{R}$ Riemann-integrierbar ist. Dazu beweisen wir zunächst

Satz 3.2.16 *Sei $f : [a,b] \to [c,d]$ Riemann-integrierbar und sei $\varphi : [c,d] \to \mathbb{R}$ stetig. Dann ist auch $\varphi \circ f : [a,b] \to \mathbb{R}$ Riemann-integrierbar.*

Beweis. Da $[c,d]$ kompakt ist, ist φ und somit auch $\varphi \circ f$ beschränkt. Außerdem ist φ sogar gleichmäßig stetig. Also existiert zu vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ mit

$$y_1, y_2 \in [c,d] \ \text{und}\ |y_1 - y_2| < \delta \quad \Longrightarrow \quad |\varphi(y_1) - \varphi(y_2)| < \varepsilon . \tag{3.2.3}$$

Dabei können wir

$$\delta < \varepsilon \tag{3.2.4}$$

annehmen. Da f Riemann-integrierbar ist, existieren nach Satz 3.2.12 zwei Zerlegungen $Z_1, Z_2 \in \mathfrak{Z}([a,b])$ mit

$$\overline{S}(f, Z_2) - \underline{S}(f, Z_1) < \delta^2 .$$

Sei $Z = \{t_0, \ldots, t_k\}$ eine gemeinsame Verfeinerung von Z_1 und Z_2. Nach Lemma 3.2.5(i) gilt dann auch

$$\overline{S}(f, Z) - \underline{S}(f, Z) < \delta^2 . \tag{3.2.5}$$

Wir setzen jetzt

$$\begin{aligned} N_1 &:= \{i \in \{1, \ldots, k\} : \sup f([t_{i-1}, t_i]) - \inf f([t_{i-1}, t_i]) < \delta\} , \\ N_2 &:= \{i \in \{1, \ldots, k\} : \sup f([t_{i-1}, t_i]) - \inf f([t_{i-1}, t_i]) \geq \delta\} . \end{aligned}$$

Ist $i \in N_1$, so gilt

$$|f(x_1) - f(x_2)| < \delta \quad \text{für alle} \quad x_1, x_2 \in [t_{i-1}, t_i] ,$$

was nach (3.2.3)

$$|\varphi \circ f(x_1) - \varphi \circ f(x_2)| < \varepsilon \quad \text{für alle} \quad x_1, x_2 \in [t_{i-1}, t_i]$$

impliziert. Also ist

$$\sup \varphi \circ f([t_{i-1}, t_i]) - \inf \varphi \circ f([t_{i-1}, t_i]) \leq \varepsilon \quad \text{für alle} \quad i \in N_1 \ . \tag{3.2.6}$$

Weiter haben wir

$$\begin{aligned}
\sum_{i \in N_2} (t_i - t_{i-1}) &= \frac{1}{\delta} \sum_{i \in N_2} \delta \, (t_i - t_{i-1}) \\
&\leq \frac{1}{\delta} \sum_{i \in N_2} (\sup f([t_{i-1}, t_i]) - \inf f([t_{i-1}, t_i])) \, (t_i - t_{i-1}) \\
&\leq \frac{1}{\delta} \sum_{i=1}^{k} (\sup f([t_{i-1}, t_i]) - \inf f([t_{i-1}, t_i])) \, (t_i - t_{i-1}) \\
&= \frac{1}{\delta} \left(\overline{S}(f, Z) - \underline{S}(f, Z) \right) \ .
\end{aligned}$$

Hieraus und aus (3.2.5) folgt

$$\sum_{i \in N_2} (t_i - t_{i-1}) < \delta \ . \tag{3.2.7}$$

Wir setzen

$$\alpha := \max \{ |\varphi(y)| : y \in [c, d] \}$$

und schließen mit (3.2.4), (3.2.6) und (3.2.7)

$$\begin{aligned}
\overline{S}(\varphi \circ f, Z) - \underline{S}(\varphi \circ f, Z) &= \sum_{i=1}^{k} (\sup \varphi \circ f([t_{i-1}, t_i]) - \inf \varphi \circ f([t_{i-1}, t_i])) \, (t_i - t_{i-1}) \\
&= \sum_{i \in N_1} (\sup \varphi \circ f([t_{i-1}, t_i]) - \inf \varphi \circ f([t_{i-1}, t_i])) \, (t_i - t_{i-1}) \\
&\quad + \sum_{i \in N_2} (\sup \varphi \circ f([t_{i-1}, t_i]) - \inf \varphi \circ f([t_{i-1}, t_i])) \, (t_i - t_{i-1}) \\
&\leq \sum_{i \in N_1} \varepsilon \, (t_i - t_{i-1}) + \sum_{i \in N_2} 2\alpha \, (t_i - t_{i-1}) \\
&\leq \varepsilon \, (b - a) + 2\alpha\delta \\
&\leq \varepsilon \, (b - a + 2\alpha) \ .
\end{aligned}$$

Damit haben wir gezeigt, dass zu jedem $\varepsilon > 0$ ein $Z \in \mathfrak{Z}([a, b])$ mit

$$\overline{S}(\varphi \circ f, Z) - \underline{S}(\varphi \circ f, Z) \leq \varepsilon \, (b - a + 2\alpha)$$

existiert. Dies liefert nach Satz 3.2.12 die Behauptung. □

Folgerung 3.2.17 *Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ ist Riemann-integrierbar.*

Beweis. Sei $f : [a, b] \to \mathbb{R}$ stetig. Da $\mathrm{id}_{[a,b]}$ nach Beispiel 3.2.14 Riemann-integrierbar ist, folgt mit Satz 3.2.16, dass $f = f \circ \mathrm{id}_{[a,b]}$ Riemann-integrierbar ist. □

Die nächsten beiden Sätze beinhalten grundlegende Rechenregeln für das Riemann-Integral.

Satz 3.2.18 *Seien $f, g : [a, b] \to \mathbb{R}$ Riemann-integrierbar und sei $\alpha \in \mathbb{R}$. Dann gilt:*

(i) *$f+g$ ist Riemann-integrierbar und*

$$\int_a^b (f(x)+g(x))\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x \;.$$

(ii) *αf ist Riemann-integrierbar und*

$$\int_a^b \alpha f(x)\,\mathrm{d}x = \alpha \int_a^b f(x)\,\mathrm{d}x \;.$$

(iii) *fg ist Riemann-integrierbar.*

(iv) *Gilt*

$$g(x) \le f(x) \quad \textit{für alle} \quad x \in [a,b] \;,$$

so ist

$$\int_a^b g(x)\,\mathrm{d}x \le \int_a^b f(x)\,\mathrm{d}x \;.$$

(v) *$|f|$ ist Riemann-integrierbar und*

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \le \int_a^b |f(x)|\,\mathrm{d}x \;. \tag{3.2.8}$$

Beweis. (i) Als erstes stellen wir fest, dass mit f und g auch $f+g$ beschränkt ist. Weiter haben wir für jedes $Z = \{t_0, \ldots, t_k\} \in \mathfrak{Z}([a,b])$

$$\begin{aligned}
\overline{S}(f+g,Z) &= \sum_{i=1}^{k} (\sup\,(f+g)([t_{i-1},t_i]))\,(t_i - t_{i-1}) \\
&\le \sum_{i=1}^{k} (\sup f([t_{i-1},t_i]) + \sup g([t_{i-1},t_i]))\,(t_i - t_{i-1}) \\
&= \overline{S}(f,Z) + \overline{S}(g,Z) \;.
\end{aligned}$$

Folglich ist

$$J^*(f+g) \le \inf\,\left\{\overline{S}(f,Z) + \overline{S}(g,Z) : Z \in \mathfrak{Z}([a,b])\right\} \;. \tag{3.2.9}$$

Außerdem ist

$$\inf\,\left\{\overline{S}(f,Z) + \overline{S}(g,Z) : Z \in \mathfrak{Z}([a,b])\right\} = J^*(f) + J^*(g) \;. \tag{3.2.10}$$

Zunächst gilt nämlich

$$\begin{aligned}
&\inf\,\left\{\overline{S}(f,Z) + \overline{S}(g,Z) : Z \in \mathfrak{Z}([a,b])\right\} \\
&\qquad \ge \inf\,\left\{\overline{S}(f,Z) : Z \in \mathfrak{Z}([a,b])\right\} + \inf\,\left\{\overline{S}(g,Z) : Z \in \mathfrak{Z}([a,b])\right\} \\
&\qquad = J^*(f) + J^*(g) \;.
\end{aligned}$$

Wäre

$$\inf\,\left\{\overline{S}(f,Z) + \overline{S}(g,Z) : Z \in \mathfrak{Z}([a,b])\right\} > J^*(f) + J^*(g) \;,$$

so würden $Z_1, Z_2 \in \mathfrak{Z}([a,b])$ mit

$$\overline{S}(f,Z) + \overline{S}(g,Z) > \overline{S}(f,Z_1) + \overline{S}(g,Z_2) \quad \text{für alle} \quad Z \in \mathfrak{Z}([a,b])$$

existieren. Im Widerspruch dazu würde aber nach Lemma 3.2.5(i) für jede gemeinsame Verfeinerung Z' von Z_1 und Z_2

$$\overline{S}(f,Z') + \overline{S}(g,Z') \le \overline{S}(f,Z_1) + \overline{S}(g,Z_2)$$

gelten. Damit ist (3.2.10) gezeigt. Aus (3.2.9) und (3.2.10) folgt

$$J^*(f+g) \leq J^*(f) + J^*(g) = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x\ .$$

Analog verifiziert man

$$J_*(f+g) \geq \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x\ .$$

Also gilt

$$\int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x \leq J_*(f+g) \leq J^*(f+g) \leq \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x\ ,$$

was unmittelbar auf die Behauptung (i) führt.

(ii) Übung.

(iii) Wir benutzen (i) und (ii). Da f und g Riemann-integrierbar sind, sind auch $f+g$ und $f-g$ Riemann-integrierbar. Mittels Satz 3.2.16 für $\varphi : [c,d] \to \mathbb{R}$, $\varphi(t) := t^2$, wobei das Intervall $[c,d]$ so gewählt ist, dass $(f+g)([a,b]) \subseteq [c,d]$ und $(f-g)([a,b]) \subseteq [c,d]$, erhalten wir, dass $(f+g)^2$ und $(f-g)^2$ Riemann-integrierbar sind. Also ist auch

$$fg = \frac{1}{4}\left((f+g)^2 - (f-g)^2\right)$$

Riemann-integrierbar.

(iv) Übung.

(v) Da f Riemann-integrierbar ist, ist nach Satz 3.2.16 auch $|f|$ Riemann-integrierbar. Da außerdem

$$-|f(x)| \leq f(x) \leq |f(x)| \quad \text{für alle} \quad x \in [a,b]\ ,$$

folgt mit (ii) und (iv)

$$-\int_a^b |f(x)|\,\mathrm{d}x \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b |f(x)|\,\mathrm{d}x\ .$$

Damit ist auch (3.2.8) gezeigt. □

Satz 3.2.19 *Sei $a < c < b$. Eine Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn $f\big|_{[a,\,c]}$ und $f\big|_{[c,\,b]}$ Riemann-integrierbar sind, und in diesem Fall gilt*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x\ . \tag{3.2.11}$$

Beweis. Offensichtlich ist $f : [a,b] \to \mathbb{R}$ genau dann beschränkt, wenn $f\big|_{[a,\,c]}$ und $f\big|_{[c,\,b]}$ beschränkt sind. Wir nehmen zuerst an, dass f Riemann-integrierbar ist. Die Funktion

$$g : [a,b] \to \mathbb{R}\ , \quad g(x) := \begin{cases} 1 & \text{für} \quad x \in [a,c] \\ 0 & \text{für} \quad x \in \,]c,b] \end{cases}\ ,$$

ist nach Beispiel 3.2.15 Riemann-integrierbar. Nach Satz 3.2.18(iii) ist dann auch $f_1 := fg$ Riemann-integrierbar. Außerdem gilt für jedes $Z \in \mathfrak{Z}([a,b])$

$$\begin{aligned}
\underline{S}(f_1, Z) &\leq \underline{S}(f_1, Z \cup \{c\}) = \underline{S}\left(f\big|_{[a,\,c]}, (Z \cup \{c\}) \cap [a,c]\right)\ ,\\
\overline{S}(f_1, Z) &\geq \overline{S}(f_1, Z \cup \{c\}) = \overline{S}\left(f\big|_{[a,\,c]}, (Z \cup \{c\}) \cap [a,c]\right)\ .
\end{aligned}$$

Das liefert
$$J_*(f_1) \le J_*\left(f\big|_{[a,\,c]}\right) \le J^*\left(f\big|_{[a,\,c]}\right) \le J^*(f_1)\ .$$
Somit ist $f\big|_{[a,\,c]}$ Riemann-integrierbar. Analog erhält man die Riemann-Integrierbarkeit von $f\big|_{[c,\,b]}$.

Jetzt nehmen wir an, dass $f\big|_{[a,\,c]}$ und $f\big|_{[c,\,b]}$ Riemann-integrierbar sind. Wegen
$$\begin{aligned}\underline{S}(f, Z_1 \cup Z_2) &= \underline{S}\left(f\big|_{[a,\,c]}, Z_1\right) + \underline{S}\left(f\big|_{[c,\,b]}, Z_2\right)\ ,\\ \overline{S}(f, Z_1 \cup Z_2) &= \overline{S}\left(f\big|_{[a,\,c]}, Z_1\right) + \overline{S}\left(f\big|_{[c,\,b]}, Z_2\right)\end{aligned}$$
für alle $Z_1 \in \mathfrak{Z}([a,c])$ und $Z_2 \in \mathfrak{Z}([c,b])$ gilt
$$J_*\left(f\big|_{[a,\,c]}\right) + J_*\left(f\big|_{[c,\,b]}\right) \le J_*(f) \le J^*(f) \le J^*\left(f\big|_{[a,\,c]}\right) + J^*\left(f\big|_{[c,\,b]}\right)\ .$$
Folglich ist f Riemann-integrierbar und es gilt die Gleichung (3.2.11). □

Wir treffen folgende Festlegungen. Für jedes $a \in \mathbb{R}$ sei
$$\int_a^a f(x)\,\mathrm{d}x := 0\ .$$
Ist $f : [a,b] \to \mathbb{R}$ Riemann-integrierbar, so sei
$$\int_b^a f(x)\,\mathrm{d}x := -\int_a^b f(x)\,\mathrm{d}x\ .$$
Die Riemann-Integrierbarkeit und das Riemann-Integral vektorwertiger Funktionen definieren wir folgendermaßen.

Definition 3.2.20 (i) *Eine Funktion* $f = (f_1, \ldots, f_n) : [a,b] \to \mathbb{R}^n$ *heißt* **Riemann-integrierbar** $:\Longleftrightarrow$ $f_1, \ldots, f_n$ *sind Riemann-integrierbar. In diesem Fall sei*
$$\int_a^b f(x)\,\mathrm{d}x := \left(\int_a^b f_1(x)\,\mathrm{d}x, \ldots, \int_a^b f_n(x)\,\mathrm{d}x\right)\ .$$

(i) *Eine Funktion* $f : [a,b] \to \mathbb{C}^n$ *heißt* **Riemann-integrierbar** $:\Longleftrightarrow$ *Der Realteil* $\mathrm{Re}(f)$ *und der Imaginärteil* $\mathrm{Im}(f)$ *von* f *sind Riemann-integrierbar. Es sei dann*
$$\int_a^b f(x)\,\mathrm{d}x := \int_a^b \mathrm{Re}(f(x))\,\mathrm{d}x + \mathrm{i}\int_a^b \mathrm{Im}(f(x))\,\mathrm{d}x$$

Sei wieder $E = \mathbb{K}^n$ mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Wie man einfach sieht, gelten die oben gezeigten Eigenschaften des Riemann-Integrals für Funktionen $f : [a,b] \to E$ in entsprechender Weise. Insbesondere impliziert die Stetigkeit die Riemann-Integrierbarkeit und für jede Riemann-integrierbare Funktion f haben wir die Abschätzung
$$\left\|\int_a^b f(x)\,\mathrm{d}x\right\| \le \int_a^b \|f(x)\|\,\mathrm{d}x\ . \tag{3.2.12}$$
Wir beschließen diesen Abschnitt mit

Satz 3.2.21 (Hauptsatz der Differential- und Integralrechnung) *Sei* $f : [a,b] \to E$ *stetig. Dann gilt:*

(i) *Die Abbildung*

$$F : [a,b] \to E \,, \quad F(x) := \int_a^x f(t)\,\mathrm{d}t \,,$$

ist eine Stammfunktion von f.

(ii) *Für jede Stammfunktion G von f gilt*

$$\int_a^b f(x)\,\mathrm{d}x = G(b) - G(a) \,. \tag{3.2.13}$$

Beweis. (i) Sei $\varepsilon > 0$ und sei $\xi \in [a,b]$. Wegen der Stetigkeit von f gibt es ein $\delta > 0$ mit

$$\|f(t) - f(\xi)\| < \varepsilon \quad \text{für alle} \quad t \in [a,b] \cap \,]\xi - \delta, \xi + \delta[\,.$$

Mittels (3.2.12) folgern wir für alle $x \in [a,b]$ mit $0 < |x - \xi| < \delta$

$$\begin{aligned}
\left\| \frac{F(x) - F(\xi)}{x - \xi} - f(\xi) \right\| &= \left\| \frac{1}{x - \xi} \int_\xi^x f(t)\,\mathrm{d}t - f(\xi) \right\| \\
&= \left\| \frac{1}{x - \xi} \int_\xi^x (f(t) - f(\xi))\,\mathrm{d}t \right\| \\
&\le \frac{1}{x - \xi} \int_\xi^x \|f(t) - f(\xi)\|\,\mathrm{d}t \\
&\le \frac{\varepsilon}{x - \xi} \int_\xi^x \mathrm{d}t \\
&= \varepsilon \,.
\end{aligned}$$

Damit haben wir gezeigt, dass

$$\lim_{x \to \xi} \frac{F(x) - F(\xi)}{x - \xi} = f(\xi) \,.$$

Also ist F in ξ differenzierbar und $F'(\xi) = f(\xi)$.

(ii) Offensichtlich gilt (3.2.13) für $G = F$. Da nach Satz 3.1.2 jede Stammfunktion von f von der Gestalt $F + c$ mit einem $c \in E$ ist, ergibt das die Behauptung. □

Bemerkung 3.2.22 (i) Die Beziehung (3.2.13) schreibt man häufig in der Form

$$\int_a^b f(x)\,\mathrm{d}x = G(x)\Big|_a^b \,.$$

(ii) Nicht jede Riemann-integrierbare Funktion besitzt eine Stammfunktion. Zum Beispiel ist die Funktion $\mathrm{sgn}\big|_{[-1,1]}$ Riemann-integrierbar (vgl. Beispiel 3.2.15). Sie hat aber keine Stammfunktion. □

3.3 Weitere Eigenschaften des Riemannschen Integrals

Satz 3.3.1 (Mittelwertsatz der Integralrechnung) *Sei $f : [a,b] \to \mathbb{R}$ stetig, sei $g : [a,b] \to \mathbb{R}$ Riemann-integrierbar und gelte $g(x) \ge 0$ für alle $x \in [a,b]$. Dann existiert ein $\xi \in [a,b]$ mit*

$$\int_a^b f(x) g(x)\,\mathrm{d}x = f(\xi) \int_a^b g(x)\,\mathrm{d}x \,.$$

Beweis. Wir setzen

$$c_1 := \min f([a,b]) \quad \text{und} \quad c_2 := \max f([a,b]) \, .$$

Dann gilt

$$c_1 \leq f(x) \leq c_2 \quad \text{für alle} \quad x \in [a,b] \, ,$$

also wegen $g(x) \geq 0$ auch

$$c_1 g(x) \leq f(x)g(x) \leq c_2 g(x) \quad \text{für alle} \quad x \in [a,b] \, ,$$

was nach Satz 3.2.18

$$c_1 \int_a^b g(x)\,\mathrm{d}x \leq \int_a^b f(x)g(x)\,\mathrm{d}x \leq c_2 \int_a^b g(x)\,\mathrm{d}x \tag{3.3.1}$$

impliziert.

Wir nehmen zuerst an, dass

$$\int_a^b g(x)\,\mathrm{d}x = 0 \, .$$

Dann ist (3.3.1) zu

$$0 \leq \int_a^b f(x)g(x)\,\mathrm{d}x \leq 0 \, ,$$

d.h. zu

$$\int_a^b f(x)g(x)\,\mathrm{d}x = 0$$

äquivalent. Damit ist die Behauptung in diesem Fall gezeigt.

Gelte jetzt

$$\int_a^b g(x)\,\mathrm{d}x \neq 0$$

und somit wegen $g(x) \geq 0$

$$\int_a^b g(x)\,\mathrm{d}x > 0$$

und seien $x_1, x_2 \in [a,b]$ so gewählt, dass

$$f(x_1) = c_1 \quad \text{und} \quad f(x_2) = c_2 \, .$$

Dann bedeutet (3.3.1), dass

$$f(x_1) \leq \frac{\int_a^b f(x)g(x)\,\mathrm{d}x}{\int_a^b g(x)\,\mathrm{d}x} \leq f(x_2) \, .$$

Mittels Zwischenwertsatz folgt, dass ein $\xi \in [a,b]$ mit

$$f(\xi) = \frac{\int_a^b f(x)g(x)\,\mathrm{d}x}{\int_a^b g(x)\,\mathrm{d}x}$$

existiert. Damit ist die Behauptung auch in diesem Fall bewiesen. □

Folgerung 3.3.2 *Ist $f : [a,b] \to \mathbb{R}$ stetig, so existiert ein $\xi \in [a,b]$ mit*

$$\int_a^b f(x)\,\mathrm{d}x = f(\xi)\,(b-a) \, .$$

Beweis. Das ist die Aussage von Satz 3.3.1 für den Fall, dass g konstant 1 ist. □

Der nächste Satz besagt, dass man für gleichmäßig konvergente Funktionenfolgen Integral und Grenzwert vertauschen kann.

Satz 3.3.3 *Seien $f_k : [a,b] \to \mathbb{R}$, $k \in \mathbb{N}$, Riemann-integrierbar und konvergiere $(f_k)_{k\in\mathbb{N}}$ gleichmäßig gegen $f : [a,b] \to \mathbb{R}$. Dann ist f Riemann-integrierbar und*

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{k\to\infty} \int_a^b f_k(x)\,\mathrm{d}x \,. \tag{3.3.2}$$

Beweis. Sei $\varepsilon > 0$. Da (f_k) gleichmäßig gegen f konvergiert, existiert ein $k_0 \in \mathbb{N}$ derart, dass für alle $k \geq k_0$ und alle $x \in [a,b]$

$$|f(x) - f_k(x)| < \varepsilon \,,$$

d.h.

$$f_k(x) - \varepsilon < f(x) < f_k(x) + \varepsilon \,.$$

Demnach ist f beschränkt. Außerdem folgt für $Z = \{t_0, \ldots, t_l\} \in \mathfrak{Z}([a,b])$ und $k \geq k_0$

$$\begin{aligned}\overline{S}(f,Z) &= \sum_{i=1}^{l} (\sup f([t_{i-1},t_i]))\,(t_i - t_{i-1}) \\ &\leq \sum_{i=1}^{l} (\sup f_k([t_{i-1},t_i]) + \varepsilon)\,(t_i - t_{i-1}) = \overline{S}(f_k,Z) + \varepsilon(b-a) \,.\end{aligned}$$

Also gilt

$$\overline{S}(f,Z) \leq \overline{S}(f_k,Z) + \varepsilon(b-a) \quad \text{für alle} \quad Z \in \mathfrak{Z}([a,b]) \text{ und } k \geq k_0 \,. \tag{3.3.3}$$

Analog erhält man

$$\underline{S}(f,Z) \geq \underline{S}(f_k,Z) - \varepsilon(b-a) \quad \text{für alle} \quad Z \in \mathfrak{Z}([a,b]) \text{ und } k \geq k_0 \,. \tag{3.3.4}$$

Da f_{k_0} Riemann-integrierbar ist, existieren nach Satz 3.2.12 Zerlegungen $Z_1, Z_2 \in \mathfrak{Z}([a,b])$ mit

$$\overline{S}(f_{k_0}, Z_2) - \underline{S}(f_{k_0}, Z_1) < \varepsilon \,.$$

Wegen (3.3.3) und (3.3.4) ist dann

$$\overline{S}(f, Z_2) - \underline{S}(f, Z_1) < \varepsilon(1 + 2(b-a)) \,.$$

Wiederum nach Satz 3.2.12 folgt, dass f Riemann-integrierbar ist.

Aus (3.3.3) und (3.3.4) ergibt sich nun, dass für alle $k \geq k_0$

$$\int_a^b f_k(x)\,\mathrm{d}x - \varepsilon(b-a) \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b f_k(x)\,\mathrm{d}x + \varepsilon(b-a) \,,$$

d.h.

$$\left| \int_a^b f(x)\,\mathrm{d}x - \int_a^b f_k(x)\,\mathrm{d}x \right| \leq \varepsilon(b-a) \,.$$

Damit ist auch (3.3.2) gezeigt. □

Folgerung 3.3.4 *Sind $f_k : [a,b] \to \mathbb{R}$, $k \in \mathbb{N}$, Riemann-integrierbar und konvergiert die Reihe $\sum\limits_{k=1}^{\infty} f_k$ gleichmäßig gegen $f : [a,b] \to \mathbb{R}$, so ist auch f Riemann-integrierbar und es gilt*

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{\infty} \int_a^b f_k(x)\,\mathrm{d}x \,.$$

Beweis. Man wende Satz 3.3.3 auf die Partialsummen von $\sum_{k=1}^{\infty} f_k$ an. $\square$

Eine Antwort auf die Frage , wann Integral und Ableitung vertauschbar sind, gibt

Satz 3.3.5 *Sei $U \subseteq \mathbb{R}$ offen, sei $f : (t,x) \in U \times [a,b] \mapsto f(t,x) \in \mathbb{R}$ stetig und sei $\Phi : U \to \mathbb{R}$ durch*

$$\Phi(t) := \int_a^b f(t,x)\,\mathrm{d}x$$

definiert. Dann gilt:

(i) Φ *ist stetig.*

(ii) *Existiert die partielle Ableitung* $\frac{\partial f}{\partial t} : U \times [a,b] \to \mathbb{R}$ *und ist diese stetig, so ist* Φ *stetig differenzierbar und*

$$\Phi'(t) = \int_a^b \frac{\partial f}{\partial t}(t,x)\,\mathrm{d}x \quad \textit{für alle} \quad t \in U\ .$$

Beweis. (i) Sei $\theta \in U$ fixiert und sei $\alpha > 0$ so gewählt, dass $[\theta-\alpha, \theta+\alpha] \subseteq U$. Da die Einschränkung von f auf die kompakte Menge $[\theta-\alpha, \theta+\alpha] \times [a,b]$ gleichmäßig stetig ist, gibt es zu vorgegebenem $\varepsilon > 0$ ein $\delta \in]0,\alpha]$ mit

$$|f(t,x) - f(\theta,x)| < \frac{\varepsilon}{b-a} \quad \text{für alle} \quad x \in [a,b] \text{ und } t \in]\theta-\delta, \theta+\delta[\ .$$

Hieraus folgt

$$|\Phi(t) - \Phi(\theta)| = \left| \int_a^b f(t,x) - f(\theta,x)\,\mathrm{d}x \right| \le \int_a^b |f(t,x) - f(\theta,x)|\,\mathrm{d}x \le \int_a^b \frac{\varepsilon}{b-a}\,\mathrm{d}x = \varepsilon$$

für $t \in]\theta-\delta, \theta+\delta[$. Somit ist Φ stetig in θ und (i) ist bewiesen.

(ii) Wir setzen voraus, dass $\frac{\partial f}{\partial t} : U \times [a,b] \to \mathbb{R}$ existiert und stetig ist. Seien θ und α wie in (i) und sei $g :]-\alpha, \alpha[\times [a,b] \to \mathbb{R}$ durch

$$g(h,x) := \begin{cases} \dfrac{f(\theta+h,x) - f(\theta,x)}{h} & \text{für} \quad h \neq 0 \\ \dfrac{\partial f}{\partial t}(\theta,x) & \text{für} \quad h = 0 \end{cases}$$

definiert. Offensichtlich ist g in jedem (h,x) mit $h \neq 0$ stetig. Außerdem existiert nach dem Mittelwertsatz der Differentialrechnung zu jedem $(h,x) \in]-\alpha,\alpha[\times [a,b]$ mit $h \neq 0$ ein $s \in]0,1[$ derart, dass

$$g(h,x) = \frac{\partial f}{\partial t}(\theta + sh, x)\ .$$

Wegen der Stetigkeit von $\frac{\partial f}{\partial t}$ folgt, dass g auch in allen $(0,x)$ stetig ist. Indem wir (i) auf g anwenden, erhalten wir

$$\Phi'(\theta) = \lim_{h\to 0} \frac{\Phi(\theta+h) - \Phi(\theta)}{h} = \lim_{h\to 0} \int_a^b g(h,x)\,\mathrm{d}x = \int_a^b g(0,x)\,\mathrm{d}x = \int_a^b \frac{\partial f}{\partial t}(\theta,x)\,\mathrm{d}x\ .$$

Die Stetigkeit von Φ' folgt wieder nach (i) aus der Stetigkeit von $\frac{\partial f}{\partial t}$. $\square$

Wir wollen jetzt noch das so genannte Lebesguesche Integrierbarkeitskriterium angeben. Dazu benötigen wir

Definition 3.3.6 *Eine Menge $A \subseteq \mathbb{R}$ heißt* **Lebesguesche Nullmenge** $:\Longleftrightarrow$ *Zu jedem $\varepsilon > 0$ existiert eine Folge $([a_k, b_k])_{k\in\mathbb{N}}$ von abgeschlossenen Intervallen mit*

$$A \subseteq \bigcup_{k=1}^{\infty} [a_k, b_k] \quad \textit{und} \quad \sum_{k=1}^{\infty} (b_k - a_k) < \varepsilon \, .$$

Lemma 3.3.7 (i) *Jede endliche Teilmenge von $\mathbb{R}$ ist eine Lebesguesche Nullmenge.*

(ii) *Die Vereinigung höchstens abzählbar vieler Lebesguescher Nullmengen ist wieder eine Lebesguesche Nullmenge. Insbesondere ist jede abzählbare Teilmenge von $\mathbb{R}$ eine Lebesguesche Nullmenge.*

Beweis. (i) Übung.

(ii) Vgl. Heuser, Lehrbuch der Analysis, Teil 1, Abschnitt 84. □

Satz 3.3.8 (Lebesguesches Integrierbarkeitskriterium) *Sei $f : [a, b] \to \mathbb{R}$ beschränkt. Dann ist f genau dann Riemann-integrierbar, wenn die Menge der Unstetigkeitsstellen von f eine Lebesguesche Nullmenge ist.*

Beweis. Vgl. Heuser, Lehrbuch der Analysis, Teil 1, Abschnitt 84. □

Bemerkung 3.3.9 (i) Nach Lemma 3.3.7(ii) und Satz 3.3.8 ist jede beschränkte Funktion $f : [a, b] \to \mathbb{R}$ mit höchstens abzählbar vielen Unstetigkeitsstellen Riemann-integrierbar. Insbesondere ist jede monotone Funktion $f : [a, b] \to \mathbb{R}$ Riemann-integrierbar.

(i) Wie man leicht sieht, behalten die Ergebnisse dieses Abschnitts mit Ausnahme von Satz 3.3.1 und Folgerung 3.3.2 auch für vektorwertige Funktionen ihre Gültigkeit.

3.4 Uneigentliche Riemannsche Integrale

In diesem Abschnitt wollen wir das Riemannsche Integral zum einen für unbeschränkte Intervalle und zum anderen für unbeschränkte Funktionen verallgemeinern. Solche Integrale nennt man **uneigentliche** (Riemannsche) Integrale.

Definition 3.4.1 (i) *Sei $f : [a, +\infty[\to E$ derart, dass $f\big|_{[a,\,t]}$ für jedes $t > a$ Riemann-integrierbar ist. Man sagt,* **das Integral** $\int_a^{+\infty} f(x)\,\mathrm{d}x$ **ist konvergent** $:\Longleftrightarrow$ *Der Grenzwert $\lim\limits_{t\to+\infty} \int_a^t f(x)\,\mathrm{d}x$ existiert. In diesem Fall sei*

$$\int_a^{+\infty} f(x)\,\mathrm{d}x := \lim_{t\to+\infty} \int_a^t f(x)\,\mathrm{d}x \, .$$

Analog definiert man $\int_{-\infty}^b f(x)\,\mathrm{d}x$ für eine Funktion $f :]-\infty, b] \to E$.

(ii) *Sei $f : [a, b[\to E$ derart, dass $f\big|_{[a,\,t]}$ für jedes $t \in]a, b[$ Riemann-integrierbar ist. Man sagt,* **das Integral** $\int_a^b f(x)\,\mathrm{d}x$ **ist konvergent** $:\Longleftrightarrow$ *Der Grenzwert $\lim\limits_{t\to b-} \int_a^t f(x)\,\mathrm{d}x$ existiert. Es sei dann*

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{t\to b-} \int_a^t f(x)\,\mathrm{d}x \, .$$

Analog definiert man $\int_a^b f(x)\,\mathrm{d}x$ für eine Funktion $f :]a, b] \to E$.

(iii) *Ein uneigentliches Integral heißt* **divergent** $:\iff$ *Es ist nicht konvergent.*

Beispiel 3.4.2 Es ist

$$\int_0^{+\infty} \mathrm{e}^{-x}\,\mathrm{d}x = 1\ ,$$

denn

$$\lim_{t\to+\infty} \int_0^t \mathrm{e}^{-x}\,\mathrm{d}x = \lim_{t\to+\infty} \left(1-\mathrm{e}^{-t}\right) = 1\ .$$

□

Beispiel 3.4.3 Das Integral $\int_0^{+\infty} \cos(x)\,\mathrm{d}x$ ist divergent, denn

$$\int_0^t \cos(x)\,\mathrm{d}x = \sin(t)$$

und der Grenzwert $\lim\limits_{t\to+\infty} \sin(t)$ existiert nicht. □

Beispiel 3.4.4 Das uneigentliche Integral $\int_1^{+\infty} x^s\,\mathrm{d}x$ ist genau dann konvergent, wenn $s < -1$, und in diesem Fall ist

$$\int_1^{+\infty} x^s\,\mathrm{d}x = -\frac{1}{s+1}\ .$$

Für jedes $t > 1$ gilt nämlich

$$\int_1^t x^s\,\mathrm{d}x = \begin{cases} \dfrac{1}{s+1}\left(t^{s+1}-1\right) & \text{für} \quad s \neq -1 \\ \ln(t) & \text{für} \quad s = -1 \end{cases}\ .$$

□

Beispiel 3.4.5 Es gilt

$$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \frac{\pi}{2}\ ,$$

denn

$$\lim_{t\to 1-} \int_0^t \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{t\to 1-} \arcsin(t) = \frac{\pi}{2}\ .$$

□

Beispiel 3.4.6 Das Integral $\int_0^1 x^s\,\mathrm{d}x$ ist genau dann konvergent, wenn $s > -1$, und in diesem Fall gilt

$$\int_0^1 x^s\,\mathrm{d}x = \frac{1}{s+1}\ .$$

Das folgt aus

$$\int_t^1 x^s\,\mathrm{d}x = \begin{cases} \dfrac{1}{s+1}\left(1-t^{s+1}\right) & \text{für} \quad s \neq -1 \\ -\ln(t) & \text{für} \quad s = -1 \end{cases}$$

für jedes $t \in\,]0,1[$. □

Satz 3.4.7 *Sei $-\infty < a < \beta \leq +\infty$ und seien $f, g : [a,\beta[\, \to \mathbb{R}$.*

(i) **Majorantenkriterium:** *Gilt $|f(x)| \leq g(x)$ für alle $x \in [a,\beta[$ und ist $\int_a^\beta g(x)\,\mathrm{d}x$ konvergent, so ist auch $\int_a^\beta f(x)\,\mathrm{d}x$ konvergent.*

(ii) **Minorantenkriterium:** *Gilt* $0 \le g(x) \le f(x)$ *für alle* $x \in [a, \beta[$ *und ist* $\int_a^\beta g(x)\,\mathrm{d}x$ *divergent, so ist auch* $\int_a^\beta f(x)\,\mathrm{d}x$ *divergent.*

Beweis. Übung. □

Bemerkung 3.4.8 Satz 3.4.7 gilt für $-\infty \le \alpha < b < +\infty$ und Funktionen $f, g :]\alpha, b] \to \mathbb{R}$ in analoger Weise. □

Beispiel 3.4.9 Sei $s \in \mathbb{R}_+$. Dann sind die Integrale $\int_0^1 x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$ und $\int_1^{+\infty} x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$ konvergent. Zur Verifikation benutzen wir das Majorantenkriterium. Die Konvergenz des ersten Integrals erhalten wir aus

$$x^{s-1}\mathrm{e}^{-x} \le x^{s-1} \quad \text{für alle} \quad x > 0$$

und Beispiel 3.4.6. Ferner gibt es wegen $\lim\limits_{x\to+\infty} x^{s+1}\mathrm{e}^{-x} = 0$ ein $c \ge 1$ mit

$$x^{s-1}\mathrm{e}^{-x} \le x^{-2} \quad \text{für alle} \quad x \ge c\,.$$

Diese Ungleichung und Beispiel 3.4.4 implizieren die Konvergenz von $\int_c^{+\infty} x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$ und somit auch von $\int_1^{+\infty} x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$. □

Zwischen der Konvergenz von Reihen und der Konvergenz uneigentlicher Integrale besteht folgender Zusammenhang.

Satz 3.4.10 *Sei* $f : [1, +\infty[\to \mathbb{R}$ *monoton fallend und gelte* $f(x) \ge 0$ *für alle* $x \in [1, +\infty[$. *Dann konvergiert die Reihe* $\sum\limits_{k=1}^{\infty} f(k)$ *genau dann, wenn das Integral* $\int_1^{+\infty} f(x)\,\mathrm{d}x$ *konvergent ist.*

Beweis. Zunächst stellen wir fest, dass $f\big|_{[1,\,t]}$ nach Bemerkung 3.3.9(i) für jedes $t > 1$ Riemann-integrierbar ist. Da

$$f(k+1) \le f(x) \le f(k) \quad \text{für alle} \quad x \in [k, k+1] \text{ und } k \in \mathbb{N}$$

und folglich

$$f(k+1) \le \int_k^{k+1} f(x)\,\mathrm{d}x \le f(k) \quad \text{für alle} \quad k \in \mathbb{N}\,,$$

gilt

$$\sum_{k=2}^{j+1} f(k) \le \int_1^{j+1} f(x)\,\mathrm{d}x \le \sum_{k=1}^{j} f(k) \quad \text{für alle} \quad j \in \mathbb{N}\,. \tag{3.4.1}$$

Wegen $f(x) \ge 0$ haben wir außerdem

$$1 < s \le t \quad \Longrightarrow \quad \int_1^s f(x)\,\mathrm{d}x \le \int_1^t f(x)\,\mathrm{d}x\,. \tag{3.4.2}$$

Aus (3.4.1) und (3.4.2) folgt die Behauptung. □

Beispiel 3.4.11 Die Reihe $\sum\limits_{k=1}^{\infty} k\exp\left(-k^2\right)$ konvergiert. Das ergibt sich mit Satz 3.4.10 aus folgenden Eigenschaften der Funktion $f : [1, +\infty[\to \mathbb{R}$, $f(x) := x\exp\left(-x^2\right)$. Erstens ist $f(x) \ge 0$ für alle $x \in [1, +\infty[$. Zweitens ist f wegen

$$f'(x) = \left(1 - 2x^2\right)\exp\left(-x^2\right) < 0 \quad \text{für alle} \quad x \in [1, +\infty[$$

monoton fallend. Und drittens ist $\int_1^{+\infty} f(x)\,\mathrm{d}x$ konvergent, denn

$$\lim_{t\to+\infty} \int_1^t f(x)\,\mathrm{d}x = \frac{1}{2}\lim_{t\to+\infty}\left(\frac{1}{\mathrm{e}} - \exp\left(-t^2\right)\right) = \frac{1}{2\mathrm{e}}\,.$$

□

Bisher haben wir uneigentliche Integrale über halboffenen Intervallen betrachtet. Integrale über offenen Intervallen erklären wir folgendermaßen.

Definition 3.4.12 *Sei* $-\infty \leq \alpha < c < \beta \leq +\infty$ *und sei* $f :]\alpha, \beta[\to E$ *derart, dass* $f\big|_I$ *für jedes kompakte Intervall* $I \subseteq]\alpha, \beta[$ *Riemann-integrierbar ist. Man sagt,* **das Integral** $\int_\alpha^\beta f(x)\,\mathrm{d}x$ **ist konvergent** $:\iff$ *Sowohl* $\int_\alpha^c f(x)\,\mathrm{d}x$ *als auch* $\int_c^\beta f(x)\,\mathrm{d}x$ *sind konvergent. In diesem Fall sei*

$$\int_\alpha^\beta f(x)\,\mathrm{d}x := \int_\alpha^c f(x)\,\mathrm{d}x + \int_c^\beta f(x)\,\mathrm{d}x \,.$$

Wie man einfach überprüft, ist diese Definition unabhängig von der Wahl von $c \in]\alpha, \beta[$.

Beispiel 3.4.13 Nach Beispiel 3.4.9 ist das uneigentliche Integral $\int_0^s x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$ für jedes $s \in \mathbb{R}_+$ konvergent. □

Definition 3.4.14 *Die Funktion*

$$\Gamma : \mathbb{R}_+ \to \mathbb{R}\,, \quad \Gamma(s) := \int_0^{+\infty} x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x\,,$$

wird **Gamma-Funktion** *genannt.*

Satz 3.4.15 *Es gilt:*

(i) $\Gamma(s+1) = s \cdot \Gamma(s)$ *für alle* $s \in \mathbb{R}_+$,

(ii) $\Gamma(1) = 1$,

(iiii) $\Gamma(k) = (k-1)!$ *für alle* $k \in \mathbb{N}$.

Beweis. (i) Wegen

$$\int_a^1 x^s\mathrm{e}^{-x}\,\mathrm{d}x = -x^s\mathrm{e}^{-x}\Big|_a^1 + s\int_a^1 x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$$

für $a \in]0,1[$ und analog

$$\int_1^b x^s\mathrm{e}^{-x}\,\mathrm{d}x = -x^s\mathrm{e}^{-x}\Big|_1^b + s\int_1^b x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x$$

für $b \in]1, +\infty[$, gilt

$$\int_0^1 x^s\mathrm{e}^{-x}\,\mathrm{d}x = -\frac{1}{\mathrm{e}} + s\int_0^1 x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x \quad \text{und} \quad \int_1^{+\infty} x^s\mathrm{e}^{-x}\,\mathrm{d}x = \frac{1}{\mathrm{e}} + s\int_1^{+\infty} x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x\,.$$

Folglich ist

$$\Gamma(s+1) = \int_0^1 x^s\mathrm{e}^{-x}\,\mathrm{d}x + \int_1^{+\infty} x^s\mathrm{e}^{-x}\,\mathrm{d}x = s\left(\int_0^1 x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x + \int_1^{+\infty} x^{s-1}\mathrm{e}^{-x}\,\mathrm{d}x\right) = s \cdot \Gamma(s)\,.$$

(ii) Da e^{-x} auf $[0,t]$ für alle $t > 0$ Riemann-integrierbar ist, ist

$$\Gamma(1) = \int_0^{+\infty} \mathrm{e}^{-x}\,\mathrm{d}x = \lim_{t\to+\infty}\int_0^t \mathrm{e}^{-x}\,\mathrm{d}x = \lim_{t\to+\infty}\left(1 - \mathrm{e}^{-t}\right) = 1\,.$$

(iii) Das folgt mit vollständiger Induktion aus (i) und (ii). □

3.5 Kurven und Integration

Definition 3.5.1 (i) *Eine* **(parametrisierte)** C^1**-Kurve** *in* $\mathbb{R}^n$ *ist eine stetig differenzierbare Abbildung* $\gamma : I \to \mathbb{R}^n$ *von einem Intervall* $I \subseteq \mathbb{R}$ *in* $\mathbb{R}^n$.

(ii) *Eine* C^1*-Kurve* $\gamma : I \to \mathbb{R}^n$ *heißt* **regulär** $:\Longleftrightarrow$ *Für alle* $t \in I$ *ist* $\gamma'(t) \neq 0$.

(iii) *Eine* C^1*-Kurve* $\gamma : [a,b] \to \mathbb{R}^n$ *heißt* **geschlossen** $:\Longleftrightarrow$ $\gamma(a) = \gamma(b)$ *und* $\gamma'(a) = \gamma'(b)$. *Sie heißt* **einfach geschlossen** $:\Longleftrightarrow$ *Sie ist geschlossen und* $\gamma\big|_{[a,b[}$ *ist injektiv.*

Beispiel 3.5.2 Beispiele für C^1-Kurven sind

(i) die Geraden
$$\gamma : \mathbb{R} \to \mathbb{R}^n\ , \quad \gamma(t) := x + tv\ ,$$
wobei $x, v \in \mathbb{R}^n$,

(ii) die Kreislinien
$$\gamma : [0, 2\pi] \to \mathbb{R}^2\ , \quad \gamma(t) := (r\cos(t), r\sin(t))\ ,$$
wobei $r > 0$,

(iii) die Ellipsen
$$\gamma : [0, 2\pi] \to \mathbb{R}^2\ , \quad \gamma(t) := (a\cos(t), b\sin(t))\ ,$$
wobei $a, b > 0$,

(iv) die Schraubenlinien
$$\gamma : \mathbb{R} \to \mathbb{R}^3\ , \quad \gamma(t) := (r\cos(t), r\sin(t), at)\ ,$$
wobei $r > 0$ und $a \neq 0$,

(v) die Zykloide
$$\gamma : \mathbb{R} \to \mathbb{R}^2\ , \quad \gamma(t) := (t - \sin(t), 1 - \cos(t))\ .$$

Wie man unmittelbar sieht, sind die Kurven aus (i)-(iv) regulär. Dagegen ist die Zykloide nicht regulär. Die Kreislinien und Ellipsen sind Beispiele für einfach geschlossene C^1-Kurven. □

Im Weiteren betrachten wir nur C^1-Kurven mit kompaktem Definitionsbereich.

Definition 3.5.3 *Sei* $\gamma : [a,b] \to \mathbb{R}^n$ *eine* C^1*-Kurve. Dann wird*
$$L(\gamma) := \int_a^b \|\gamma'(t)\|\ \mathrm{d}t$$
die **Bogenlänge** *oder einfach* **Länge** *von* γ *genannt.*

Beispiel 3.5.4 (i) Seien $x, y \in \mathbb{R}^n$. Die Länge der C^1-Kurve
$$\gamma : [0,1] \to \mathbb{R}^n\ , \quad \gamma(t) := x + t(y - x)\ ,$$
ist
$$L(\gamma) = \int_0^1 \|y - x\|\, \mathrm{d}t = \|y - x\|\ .$$

(ii) Die Kreislinie

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

hat die Länge

$$L(\gamma) = \int_0^{2\pi} r \,\mathrm{d}t = 2\pi r .$$

Wie die folgenden Überlegungen zeigen, ist die Bogenlänge invariant unter Umparametrisierungen.

Definition 3.5.5 (i) *Eine Abbildung $\varphi : [c, d] \to [a, b]$ heißt* **Parametertransformation** $:\iff$ *φ ist bijektiv und stetig differenzierbar und es gilt*

$$\varphi'(s) \neq 0 \quad \textit{für alle} \quad s \in [c, d] .$$

(ii) *Wir nennen zwei C^1-Kurven $\gamma : [a, b] \to \mathbb{R}^n$ und $\rho : [c, d] \to \mathbb{R}^n$* **schwach äquivalent** $:\iff$ *Es gibt eine Parametertransformation $\varphi : [c, d] \to [a, b]$ mit $\rho = \gamma \circ \varphi$.*

Bemerkung 3.5.6 Nach Satz 1.1.11 ist mit φ auch φ^{-1} eine Parametertransformation. Folglich ist die schwache Äquivalenz von C^1-Kurven eine Äquivalenzrelation. □

Satz 3.5.7 *Seien $\gamma : [a, b] \to \mathbb{R}^n$ und $\rho : [c, d] \to \mathbb{R}^n$ schwach äquivalente C^1-Kurven. Dann haben γ und ρ die gleiche Länge.*

Beweis. Sei $\varphi : [c, d] \to [a, b]$ eine Parametertransformation mit $\rho = \gamma \circ \varphi$. Da φ entweder streng monoton wachsend oder streng monoton fallend ist, gilt dann

$$\begin{aligned} L(\gamma) = \int_a^b \|\gamma'(t)\| \,\mathrm{d}t &= \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} \|\gamma'(\varphi(s))\|\,\varphi'(s)\,\mathrm{d}s \\ &= \int_c^d \|\gamma'(\varphi(s))\varphi'(s)\| \,\mathrm{d}s = \int_c^d \|\rho'(s)\| \,\mathrm{d}s = L(\rho) . \end{aligned}$$

□

Bemerkung 3.5.8 (i) Unter einem C^1**-Bogen** versteht man eine Äquivalenzklasse von schwach äquivalenten C^1-Kurven. Nach Satz 3.5.7 kann man die Länge $L(C)$ eines C^1-Bogens C durch

$$L(C) := L(\gamma)$$

für eine C repräsentierende C^1-Kurve γ definieren.

(ii) Ist C ein C^1-Bogen und ist $\gamma : [a, b] \to \mathbb{R}^n$ ein Repräsentant von C, so nennt man die Menge $\gamma([a, b]) \subseteq \mathbb{R}^n$ die **Spur** oder die **Trajektorie** von C. Offensichtlich hängt die Spur nicht von der Wahl des Repräsentanten γ ab. Ist ein (und damit jeder) Repräsentant von C regulär und außerdem injektiv oder einfach geschlossen, so ist C durch seine Spur eindeutig bestimmt.

(iii) Es sei darauf hingewiesen, dass der Begriff einer Kurve nicht einheitlich verwendet wird. Verschiedentlich ist mit Kurve ein Bogen bzw. dessen Spur gemeint. □

Wir erklären jetzt, wie man längs einer Kurve integrieren kann.

Definition 3.5.9 *Sei $\gamma = (\gamma_1, \ldots, \gamma_n) : [a, b] \to \mathbb{R}^n$ eine C^1-Kurve, sei $\gamma([a, b]) \subseteq M \subseteq \mathbb{R}^n$ und sei $F = (F_1, \ldots, F_n) : M \to \mathbb{R}^n$ stetig. Dann heißt*

$$\begin{aligned} \int_\gamma (F_1(x)\,\mathrm{d}x_1 + \cdots + F_n(x)\,\mathrm{d}x_n) &:= \int_a^b (F_1(\gamma(t))\gamma_1'(t) + \cdots + F_n(\gamma(t))\gamma_n'(t)) \,\mathrm{d}t \\ &= \int_a^b \langle F(\gamma(t)), \gamma'(t)\rangle \,\mathrm{d}t \end{aligned}$$

das **Kurvenintegral von F längs γ**.

Das Kurvenintegral von F längs γ werden wir abkürzend auch als

$$\int_\gamma \langle F(x), \mathrm{d}x\rangle$$

schreiben.

Beispiel 3.5.10 (i) Für die Kreislinie $\gamma : [0, 2\pi] \to \mathbb{R}^2$, $\gamma(t) := (r\cos(t), r\sin(t))$, ist

$$\int_\gamma (x_2\,\mathrm{d}x_1 - x_1\,\mathrm{d}x_2) = -r^2 \int_0^{2\pi} \left(\sin^2(t) + \cos^2(t)\right)\,\mathrm{d}t = -2\pi r^2 \;.$$

(ii) Sei $\gamma : [0, 1] \to \mathbb{R}^3$ durch $\gamma(t) := \left(t, t^2, t^3\right)$ definiert. Dann ist

$$\int_\gamma (x_2x_3\,\mathrm{d}x_1 + x_1x_3\,\mathrm{d}x_2 + x_1x_2\,\mathrm{d}x_3) = \int_0^1 \left(t^5 + 2t^5 + 3t^5\right)\,\mathrm{d}t = 1 \;.$$

□

Definition 3.5.11 *Wir nennen zwei C^1-Kurven $\gamma : [a, b] \to \mathbb{R}^n$ und $\rho : [c, d] \to \mathbb{R}^n$* **stark äquivalent** $:\iff$ *Es gibt eine monoton wachsende Parametertransformation $\varphi : [c, d] \to [a, b]$ mit $\rho = \gamma \circ \varphi$.*

Bei der starken Äquivalenz sind also im Gegensatz zur schwachen Äquivalenz nur solche Umparametrisierungen zugelassen, die die Orientierung, d.h. den Durchlaufsinn der Kurve nicht ändern.

Bemerkung 3.5.12 Da für eine Parametertransformation φ mit φ auch φ^{-1} monoton wachsend ist, ist die starke Äquivalenz von C^1-Kurven wieder eine Äquivalenzrelation. □

Satz 3.5.13 *Seien $\gamma : [a, b] \to \mathbb{R}^n$ und $\rho : [c, d] \to \mathbb{R}^n$ stark äquivalente C^1-Kurven, sei $\gamma([a, b]) \subseteq M \subseteq \mathbb{R}^n$ und sei $F : M \to \mathbb{R}^n$ stetig. Dann gilt*

$$\int_\gamma \langle F(x), \mathrm{d}x\rangle = \int_\rho \langle F(x), \mathrm{d}x\rangle \;.$$

Beweis. Sei $\varphi : [c, d] \to [a, b]$ eine monoton wachsende Parametertransformation mit $\rho = \gamma \circ \varphi$. Dann haben wir

$$\begin{aligned}\int_a^b F_i(\gamma(t))\gamma_i'(t)\,\mathrm{d}t &= \int_c^d F_i(\gamma(\varphi(s)))\gamma_i'(\varphi(s))\varphi'(s)\,\mathrm{d}s \\ &= \int_c^d F_i(\rho(s))\rho_i'(s)\,\mathrm{d}s\end{aligned}$$

für $i = 1, \ldots, n$. Das ergibt die Behauptung. □

Bemerkung 3.5.14 Ist C ein mit einer Orientierung versehener C^1-Bogen, d.h. eine Äquivalenzklasse von stark äquivalenten C^1-Kurven, so wird das Kurvenintegral längs C durch

$$\int_C (F_1(x)\,\mathrm{d}x_1 + \cdots + F_n(x)\,\mathrm{d}x_n) := \int_\gamma (F_1(x)\,\mathrm{d}x_1 + \cdots + F_n(x)\,\mathrm{d}x_n)$$

definiert, wobei $\gamma : [a, b] \to \mathbb{R}^n$ den orientierten Bogen C repräsentiert und F wie in Definition 3.5.9 ist. □

Beispiel 3.5.15 Sei U eine offene Teilmenge von $\mathbb{R}^3$ und beschreibe die stetige Abbildung

$$F = (F_1, F_2, F_3) : U \to \mathbb{R}^3$$

ein Kraftfeld auf U. Ferner sei $\gamma : [a, b] \to U$ eine C^1-Kurve. Dann ist

$$A(F;\gamma) := \int_\gamma (F_1(x)\,\mathrm{d}x_1 + F_2(x)\,\mathrm{d}x_2 + F_3(x)\,\mathrm{d}x_3)$$

die Arbeit, die bei der Verschiebung eines Massepunktes längs γ geleistet werden muss. Ist das Feld konstant, d.h.

$$F(x) = c = (c_1, c_2, c_3) \quad \text{für alle} \quad x \in U \ ,$$

so ist

$$A(F;\gamma) = \int_a^b (c_1\gamma_1'(t) + c_2\gamma_2'(t) + c_3\gamma_3'(t))\ \mathrm{d}t = \sum_{i=1}^{3} c_i(\gamma_i(b) - \gamma_i(a)) = \langle c, \gamma(b) - \gamma(a)\rangle \ .$$

Insbesondere hängt in diesem Fall die geleistete Arbeit $A(F;\gamma)$ nur von dem Anfangs- und dem Endpunkt von γ ab. □

Das letzte Beispiel führt uns zu der Frage, wann das Kurvenintegral nur von Anfangs- und Endpunkt der Kurve und nicht vom konkreten Kurvenverlauf abhängt.

Satz 3.5.16 *Sei U eine offene Teilmenge von $\mathbb{R}^n$ und sei $f : U \to \mathbb{R}$ stetig differenzierbar. Für jede C^1-Kurve $\gamma : [a, b] \to U$ gilt dann*

$$\int_\gamma \langle \mathrm{grad}(f)(x), \mathrm{d}x\rangle = f(\gamma(b)) - f(\gamma(a)) \ .$$

Beweis. Sei $\gamma : [a, b] \to U$ eine C^1-Kurve. Mit Hilfe der Kettenregel und des Hauptsatzes der Differential- und Integralrechnung folgern wir

$$\begin{aligned} f(\gamma(a)) - f(\gamma(b)) &= \int_a^b (f \circ \gamma)'(t)\,\mathrm{d}t \\ &= \int_a^b \left(\frac{\partial f}{\partial x_1}(\gamma(t))\gamma_1'(t) + \cdots + \frac{\partial f}{\partial x_n}(\gamma(t))\gamma_n'(t)\right) \mathrm{d}t = \int_\gamma \langle \mathrm{grad}(f)(x), \mathrm{d}x\rangle \ . \end{aligned}$$

□

Die Aussage von Satz 3.5.16 kann umgekehrt werden. Das ist der Inhalt von

Satz 3.5.17 *Sei U eine offene Teilmenge von $\mathbb{R}^n$, sei $F : U \to \mathbb{R}^n$ stetig und sei $f : U \to \mathbb{R}$ derart, dass für alle C^1-Kurven $\gamma : [a, b] \to U$*

$$\int_\gamma \langle F(x), \mathrm{d}x\rangle = f(\gamma(b)) - f(\gamma(a))$$

gilt. Dann ist f stetig differenzierbar und $F = \mathrm{grad}(f)$, d.h.

$$F_i = \frac{\partial f}{\partial x_i} \quad \textit{für} \quad i = 1, \ldots, n \ . \tag{3.5.1}$$

Beweis. Da F stetig ist, genügt es (3.5.1) zu zeigen. Sei $\{e_1, \ldots, e_n\}$ die Standardbasis von $\mathbb{R}^n$ und seien $x \in U$ und $i \in \{1, \ldots, n\}$ fixiert. Wir wählen $\delta > 0$ derart, dass $B_\delta(x) \subseteq U$ und definieren C^1-Kurven $\gamma_{+,s} : [0, s] \to \mathbb{R}^n$ und $\gamma_{-,s} : [0, s] \to \mathbb{R}^n$ für $s \in]0, \delta[$ durch

$$\gamma_{+,s}(t) = x + te_i \quad \text{und} \quad \gamma_{-,s}(t) = x - te_i \ .$$

Nach Voraussetzung gilt dann

$$\lim_{s\to 0+} \frac{f(x+se_i)-f(x)}{s} = \lim_{s\to 0+} \frac{f(\gamma_{+,s}(s))-f(\gamma_{+,s}(0))}{s} = \lim_{s\to 0+} \frac{1}{s}\int_{\gamma_{+,s}} \langle F(x), \mathrm{d}x\rangle$$
$$= \lim_{s\to 0+} \frac{1}{s}\int_0^s F_i(x+te_i)\,\mathrm{d}t$$

und

$$\lim_{s\to 0-} \frac{f(x+se_i)-f(x)}{s} = -\lim_{s\to 0+} \frac{f(x-se_i)-f(x)}{s} = -\lim_{s\to 0+} \frac{f(\gamma_{-,s}(s))-f(\gamma_{-,s}(0))}{s}$$
$$= -\lim_{s\to 0+} \int_{\gamma_{-,s}} \langle F(x), \mathrm{d}x\rangle$$
$$= \lim_{s\to 0+} \int_0^s F_i(x-te_i)\,\mathrm{d}t\ .$$

Da nach dem Hauptsatz der Differential- und Integralrechnung

$$\lim_{s\to 0+} \frac{1}{s}\int_0^s F_i(x+te_i)\,\mathrm{d}t = \lim_{s\to 0+} \int_0^s F_i(x-te_i)\,\mathrm{d}t = F_i(x)\ ,$$

ist damit (3.5.1) gezeigt. □

Im Zusammenhang mit den letzten beiden Sätzen stellt sich die Frage, wann zu einer Abbildung $F : U \to \mathbb{R}^n$ eine Funktion $f : U \to \mathbb{R}$ mit $F = \mathrm{grad}(f)$ existiert. Ist F stetig differenzierbar, so ist nach dem Satz von Schwarz

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i} \quad \text{für} \quad i,j = 1,\dots,n \tag{3.5.2}$$

eine notwendige Bedingung dafür. Die unten angegebene Version des Lemmas von Poincaré basagt, dass diese Bedingung für sternförmige Mengen U auch hinreichend ist.

Definition 3.5.18 *Eine Menge $U \subseteq \mathbb{R}^n$ heißt* **sternförmig** $:\Longleftrightarrow$ *Es gibt ein $y \in U$ mit*

$$y + t(x-y) \in U \quad \textit{für alle} \quad x \in U \ \textit{und}\ t \in [0,1]\ .$$

Beispiel 3.5.19 Alle konvexen Teilmengen von $\mathbb{R}^n$ sind sternförmig. Insbesondere sind $\mathbb{R}^n$ und jede offene Kugel $B_R(x) \subseteq \mathbb{R}^n$ sternförmig. □

Satz 3.5.20 (Lemma von Poincaré) *Sei $U \subseteq \mathbb{R}^n$ offen und sternförmig und sei $F \in C^1(U,\mathbb{R}^n)$. Gilt (3.5.2), so existiert eine Funktion $f \in C^2(U,\mathbb{R})$ mit $F = \mathrm{grad}(f)$.*

Beweis. Gelte (3.5.2) und sei $f : U \to \mathbb{R}$ durch

$$f(x) := \sum_{j=1}^n x_j \int_0^1 F_j(tx)\,\mathrm{d}t$$

definiert. Nach Satz 3.3.5 ist f stetig differenzierbar und

$$\frac{\partial f}{\partial x_i}(x) = \int_0^1 F_i(tx)\,\mathrm{d}t + \sum_{j=1}^n x_j \frac{\partial}{\partial x_i}\int_0^1 F_j(tx)\,\mathrm{d}t = \int_0^1 F_i(tx)\,\mathrm{d}t + \sum_{j=1}^n x_j \int_0^1 t\frac{\partial F_j}{\partial x_i}(tx)\,\mathrm{d}t$$
$$= \int_0^1 \left(F_i(tx) + t\sum_{j=1}^n x_j \frac{\partial F_i}{\partial x_j}(tx)\right)\mathrm{d}t = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(tF_i(tx))\,\mathrm{d}t = F_i(x)$$

für $i = 1,\dots,n$. □

Das folgende Beispiel zeigt, dass Satz 3.5.20 falsch wird, wenn man die Voraussetzung der Sternförmigkeit von U weglässt.

Beispiel 3.5.21 Sei $U = \mathbb{R}^2 \setminus \{0\}$ und sei $F = (F_1, F_2) : U \to \mathbb{R}^2$ durch

$$F_1(x_1, x_2) := -\frac{x_2}{x_1^2 + x_2^2} \quad \text{und} \quad F_2(x_1, x_2) := \frac{x_1}{x_1^2 + x_2^2}$$

definiert. Dann gilt

$$\frac{\partial F_1}{\partial x_2} = \frac{\partial F_2}{\partial x_1} \, .$$

Für das Kurvenintegral von F längs der Kreislinie $\gamma : [0, 2\pi] \to \mathbb{R}^2$, $\gamma(t) := (\cos(t), \sin(t))$, berechnet man aber

$$\int_\gamma \langle F(x), \mathrm{d}x \rangle = \int_0^{2\pi} \left(-\frac{\gamma_2(t)\gamma_1'(t)}{\gamma_1(t)^2 + \gamma_2(t)^2} + \frac{\gamma_1(t)\gamma_2'(t)}{\gamma_1(t)^2 + \gamma_2(t)^2} \right) \mathrm{d}t = \int_0^{2\pi} \left(\sin^2(t) + \cos^2(t) \right) \, \mathrm{d}t = 2\pi \, .$$

Folglich existiert keine Funktion $f : U \to \mathbb{R}$ mit $F = \operatorname{grad}(f)$, denn sonst müsste das angegebene Kurvenintegral nach Satz 3.5.16 den Wert 0 haben. □

Für eine Reihe von Betrachtungen ist es vorteilhaft, den in Definition 3.5.1 eingeführten Kurvenbegriff etwas abzuschwächen. Insbesondere möchte man auch längs Kurven mit "Ecken", also z.B. längs eines Polygons in einfacher Weise integrieren können. Dazu wird definiert:

Definition 3.5.22 (i) *Eine* **stückweise stetig differenzierbare Kurve** *in* $\mathbb{R}^n$ *ist eine stetige Abbildung* $\gamma : [a, b] \to \mathbb{R}^n$, *für die es eine Zerlegung* $a = t_0 < t_1 < \cdots < t_k = b$ *von* $[a, b]$ *derart gibt, dass die Einschränkungen* $\gamma\big|_{[t_{i-1}, t_i]}$, $i = 1, \ldots, k$, *stetig differenzierbar sind.*

(ii) *Eine stückweise stetig differenzierbare Kurve* $\gamma : [a, b] \to \mathbb{R}^n$ *heißt* **geschlossen** $:\Longleftrightarrow \gamma(a) = \gamma(b)$. *Sie heißt* **einfach geschlossen** $:\Longleftrightarrow$ *Sie ist geschlossen und* $\gamma\big|_{[a, b[}$ *ist injektiv.*

Ist $\gamma : [a, b] \to \mathbb{R}^n$ eine stückweise stetig differenzierbare Kurve und sind $a = t_0 < t_1 < \cdots < t_k = b$ wie in Definition 3.5.22, so sagen wir, dass γ aus den C^1-Kurven $\gamma_{(i)} = \gamma\big|_{[t_{i-1}, t_i]}$, $i = 1, \ldots, k$, zusammengesetzt ist.

Die Bogenlänge und das Kurvenintegral lassen sich in naheliegender Weise für stückweise stetig differenzierbare Kurven verallgemeinern:

Definition 3.5.23 *Sei* $\gamma : [a, b] \to \mathbb{R}^n$ *eine aus den* C^1*-Kurven* $\gamma_{(i)}$, $i = 1, \ldots, k$, *zusammengesetzte stückweise stetig differenzierbare Kurve, sei* $\gamma([a, b]) \subseteq M \subseteq \mathbb{R}^n$ *und sei* $F = (F_1, \ldots, F_n) : M \to \mathbb{R}^n$ *stetig. Die* **Bogenlänge** *von* γ *ist*

$$L(\gamma) := \sum_{i=1}^{k} L\left(\gamma_{(i)}\right) \, . \tag{3.5.3}$$

Das **Kurvenintegral von** F **längs** γ *ist*

$$\int_\gamma (F_1(x)\,\mathrm{d}x_1 + \cdots + F_n(x)\,\mathrm{d}x_n) := \sum_{i=1}^{k} \int_{\gamma_{(i)}} (F_1(x)\,\mathrm{d}x_1 + \cdots + F_n(x)\,\mathrm{d}x_n) \, . \tag{3.5.4}$$

Bemerkung 3.5.24 Sind $\gamma_{(i)} : [a_i, b_i] \to \mathbb{R}^n$, $i = 1, \ldots, k$, irgendwelche C^1-Kurven mit

$$\gamma_{(i)}(b_i) = \gamma_{(i+1)}(a_{i+1}) \quad \text{für} \quad i = 1, \ldots, k-1 \, ,$$

so kann man diese nach affiner Umparametrisierung zu einer stückweise stetig differenzierbaren Kurve γ zusammensetzen. Aufgrund der Sätze 3.5.7 und 3.5.13 gelten auch dann die Gleichungen (3.5.3) und (3.5.4). □

3.6 Mehrfache Integrale, Volumen und Integralsätze

Indem man anstelle von abgeschlossenen Intervallen in $\mathbb{R}$ so genannte **n-dimensionale Quader** $[a_1, b_1] \times \cdots \times [a_n, b_n]$ in $\mathbb{R}^n$ betrachtet und dann sämtliche Begriffe aus den Definitionen 3.2.1, 3.2.4, 3.2.6 und 3.2.8 überträgt, gelangt man zum Riemannschen Integral auf beschränkten Teilmengen von $\mathbb{R}^n$. In gewissen Fällen kann dieses Integral als ein mehrfaches Integral ausgedrückt und damit auf die Integration von Funktionen einer Veränderlichen zurückgeführt werden. Für Details verweisen wir auf Heuser, Lehrbuch der Analysis, Teil 2, Kapitel XXIII oder Walter, Analysis 2, §7.

Definition 3.6.1 *Eine Menge $A \subseteq \mathbb{R}^n$ heißt* **Normalbereich** $:\Longleftrightarrow$ *A ist durch Ungleichungen der Gestalt*

$$\begin{aligned} a &\le x_1 \le b\,, \\ \varphi_2(x_1) &\le x_2 \le \psi_2(x_1)\,, \\ \varphi_3(x_1, x_2) &\le x_3 \le \psi_3(x_1, x_2)\,, \\ &\vdots \\ \varphi_n(x_1, x_2, \ldots, x_{n-1}) &\le x_n \le \psi_n(x_1, x_2, \ldots, x_{n-1}) \end{aligned} \tag{3.6.1}$$

gegeben, wobei a, b reelle Zahlen und $\varphi_2, \psi_2, \ldots, \varphi_n, \psi_n$ stetige Funktionen mit $a \le b$ und

$$\varphi_i(x_1, \ldots, x_{i-1}) \le \psi_i(x_1, \ldots, x_{i-1}) \quad \text{für} \quad i = 2, \ldots, n$$

sind.

Beispiel 3.6.2 Die Ordinatenmenge $\mathrm{Ord}(\varphi) = \{(x_1, x_2) \in \mathbb{R}^2 : a \le x_1 \le b \text{ und } 0 \le x_2 \le \varphi(x_1)\}$ einer stetigen Funktion $\varphi : [a, b] \to [0, +\infty[$ ist ein Normalbereich in $\mathbb{R}^2$. □

Ist $A \subseteq \mathbb{R}^n$ beschränkt und ist $f : A \to \mathbb{R}$ Riemann-integrierbar, so bezeichne $\int_A f(x)\,\mathrm{d}x$ das Riemann-Integral von f. In den nächsten beiden Sätzen sind die für die Berechnung wichtigsten Eigenschaften des Riemannschen Integrals von Funktionen mehrerer Veränderlicher zusammengestellt.

Satz 3.6.3 *Seien $A, B \subseteq \mathbb{R}^n$ beschränkt.*

(i) *Sind $f_1, f_2 : A \to \mathbb{R}$ Riemann-integrierbar, so ist auch $\alpha_1 f_1 + \alpha_2 f_2 : A \to \mathbb{R}$ für alle $\alpha_1, \alpha_2 \in \mathbb{R}$ Riemann-integrierbar und*

$$\int_A (\alpha_1 f_1(x) + \alpha_2 f_2(x))\,\mathrm{d}x = \alpha_1 \int_A f_1(x)\,\mathrm{d}x + \alpha_2 \int_A f_2(x)\,\mathrm{d}x\,.$$

(ii) *Ist $f : A \cup B \to \mathbb{R}$ derart, dass $f\big|_A$ und $f\big|_B$ Riemann-integrierbar sind, so sind auch f, $f\big|_{A \setminus B}$, $f\big|_{B \setminus A}$ und $f\big|_{A \cap B}$ Riemann-integrierbar und*

$$\begin{aligned} \int_{A \cup B} f(x)\,\mathrm{d}x &= \int_{A \setminus B} f(x)\,\mathrm{d}x + \int_B f(x)\,\mathrm{d}x \\ &= \int_{B \setminus A} f(x)\,\mathrm{d}x + \int_A f(x)\,\mathrm{d}x \\ &= \int_{A \setminus B} f(x)\,\mathrm{d}x + \int_{B \setminus A} f(x)\,\mathrm{d}x + \int_{A \cap B} f(x)\,\mathrm{d}x\,. \end{aligned}$$

(iii) *Ist A ein durch die Ungleichungen (3.6.1) gegebener Normalbereich und ist $f : A \to \mathbb{R}$ stetig, so ist f Riemann-integrierbar und*

$$\int_A f(x)\,\mathrm{d}x = \int_a^b \left(\int_{\varphi_2(x_1)}^{\psi_2(x_1)} \cdots \left(\int_{\varphi_n(x_1,\ldots,x_{n-1})}^{\psi_n(x_1,\ldots,x_{n-1})} f(x_1, x_2, \ldots, x_n)\,\mathrm{d}x_n \right) \cdots \mathrm{d}x_2 \right) \mathrm{d}x_1 \ .$$

□

Beispiel 3.6.4 Sei $D := \{x \in \mathbb{R}^2 : |x| \leq 1\}$. Dann ist

$$\int_D x_2^2\,\mathrm{d}x = \frac{\pi}{4} \ .$$

Da nämlich D durch die Ungleichungen

$$-1 \leq x_1 \leq 1 \quad \text{und} \quad -\sqrt{1 - x_1^2} \leq x_2 \leq \sqrt{1 - x_1^2}$$

beschrieben ist, haben wir

$$\begin{aligned} \int_D x_2^2\,\mathrm{d}x = \int_{-1}^1 \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} x_2^2\,\mathrm{d}x_2\,\mathrm{d}x_1 &= \frac{2}{3} \int_{-1}^1 \left(1 - x_1^2\right)^{3/2} \mathrm{d}x_1 \\ &= \frac{2}{3} \int_{-\pi/2}^{\pi/2} \cos^4(t)\,\mathrm{d}t = \frac{\pi}{4} \ . \end{aligned}$$

Dabei wurde die Substitution $x_1 = \sin(t)$ benutzt. □

In Analogie zur Substitutionsregel für Integrale von Funktionen einer Veränderlichen gilt:

Satz 3.6.5 *Sei $U \subseteq \mathbb{R}^n$ offen, sei $\Phi : U \to \Phi(U)$ ein Lipschitz-stetiger C^1-Diffeomorphismus und sei $A \subseteq U$ beschränkt. Dann ist eine Funktion $f : \Phi(A) \to \mathbb{R}$ genau dann Riemann-integrierbar, wenn $(f \circ \Phi) \left|\det\left(\frac{\partial \Phi}{\partial x}\right)\right| : A \to \mathbb{R}$ Riemann-integrierbar ist, und in diesem Fall gilt*

$$\int_{\Phi(A)} f(y)\,\mathrm{d}y = \int_A f(\Phi(x)) \left|\det\left(\frac{\partial \Phi}{\partial x}(x)\right)\right| \mathrm{d}x \ .$$

Beispiel 3.6.6 Wir berechnen $\int_A x_1\,\mathrm{d}x_1$, wobei

$$A := \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 \geq 0 \text{ und } 1 \leq x_1^2 + x_2^2 \leq 2\right\} \ .$$

Sei $\Phi = (\Phi_1, \Phi_2) :]0, 3[\times]-\pi, \pi[\to \mathbb{R}^2$ durch

$$\Phi(r, \theta) := (r\cos(\theta), r\sin(\theta))$$

definiert. Nach Satz 2.4.11 und Beispiel 2.4.14 erfüllt Φ die Voraussetzungen von Satz 3.6.5. Wegen $A = \Phi([1, 2] \times [0, \pi/2])$ haben wir somit

$$\begin{aligned} \mathrm{vol}(A) = \int_A x_1\,\mathrm{d}x &= \int_{[1,2]\times[0,\pi/2]} \Phi_1(r, \theta) \left|\det\left(\frac{\partial \Phi}{\partial(r, \theta)}(r, \theta)\right)\right| \mathrm{d}(r, \theta) \\ &= \int_1^2 \int_0^{\pi/2} r^2\cos(\theta)\,\mathrm{d}\theta\,\mathrm{d}r = \int_1^2 r^2\,\mathrm{d}r = \frac{7}{3} \ . \end{aligned}$$

□

Beispiel 3.6.7 Sei $\varphi : [a,b] \to [0,+\infty[$ stetig. Dann ist

$$\int_{\mathrm{Ord}(\varphi)} \mathrm{d}x = \int_a^b \int_0^{\varphi(x_1)} \mathrm{d}x_2 \, \mathrm{d}x_1 = \int_a^b \varphi(x_1) \, \mathrm{d}x_1 \; .$$

Das heißt, das Riemann-Integral der konstanten Funktion $x \in \mathrm{Ord}(\varphi) \mapsto 1 \in \mathbb{R}$ stimmt mit dem in Definition 3.2.9 erklärten Flächeninhalt von $\mathrm{Ord}(\varphi)$ überein. □

Das nutzen wir, um auch für allgemeinere Mengen einen Flächeninhalt bzw. ein Volumen zu definieren.

Definition 3.6.8 *Eine Menge $A \subseteq \mathbb{R}^n$ heißt* **Jordan-messbar** $:\iff$ *A ist beschränkt und die konstante Funktion $x \in A \mapsto 1 \in \mathbb{R}$ ist Riemann-integrierbar. In diesem Fall wird*

$$\mathrm{vol}(A) := \int_A \mathrm{d}x$$

das **Volumen** *von A genannt. Ist $n = 2$, so nennt man $\mathrm{vol}(A)$ auch den* **Flächeninhalt** *von A.*

Für das Volumen von so genannten Rotationskörpern berechnen wir:

Satz 3.6.9 *Sei $\varphi : [a,b] \to [0,+\infty[$ stetig und sei $A(\varphi)$ der durch die Rotation von $\mathrm{Ord}(\varphi)$ um die x_1-Achse enstehende Körper, d.h.*

$$A(\varphi) := \left\{ x \in \mathbb{R}^3 : a \le x_1 \le b \text{ und } x_2^2 + x_3^2 \le \varphi(x_1)^2 \right\} \; .$$

Dann ist

$$\mathrm{vol}(A(\varphi)) = \pi \int_a^b \varphi(t)^2 \, \mathrm{d}t \; .$$

Beweis. Der Rotationskörper $A(\varphi)$ kann durch die Ungleichungen

$$a \le x_1 \le b \; , \quad -\varphi(x_1) \le x_2 \le \varphi(x_1) \quad \text{und} \quad -\sqrt{\varphi(x_1)^2 - x_2^2} \le x_3 \le \sqrt{\varphi(x_1)^2 - x_2^2}$$

beschrieben werden. Unter Benutzung der Substitution $x_2 = \varphi(x_1)\sin(s)$ erhalten wir

$$\begin{aligned} \mathrm{vol}(A(\varphi)) &= 2 \int_a^b \int_{-\varphi(x_1)}^{\varphi(x_1)} \sqrt{\varphi(x_1)^2 - x_2^2} \, \mathrm{d}x_2 \, \mathrm{d}x_1 \\ &= 2 \int_a^b \varphi(x_1)^2 \, \mathrm{d}x_1 \int_{-\pi/2}^{\pi/2} \cos^2(s) \, \mathrm{d}s = \pi \int_a^b \varphi(x_1)^2 \, \mathrm{d}x_1 \; . \end{aligned}$$

□

Wir wollen im Folgenden einen fundamentalen Zusammenhang zwischen Integralen über ebenen Gebieten und Kurvenintegralen ableiten. Zur Formulierung benötigen wir eine Vorbereitung.

Definition 3.6.10 *Eine* **geschlossene Jordan-Kurve** *in $\mathbb{R}^n$ ist eine stetige Abbildung $\gamma : [a,b] \to \mathbb{R}^n$ mit der Eigenschaft, dass $\gamma(a) = \gamma(b)$ und $\gamma\big|_{[a,b[}$ injektiv ist.*

Insbesondere ist jede einfach geschlossene stückweise stetig differenzierbare Kurve eine geschlossene Jordan-Kurve.

Die Aussage des nächsten Satzes ist anschaulich plausibel, aber schwer zu beweisen.

Satz 3.6.11 (Jordanscher Kurvensatz) *Eine geschlossene Jordankurve $\gamma : [a,b] \to \mathbb{R}^2$ unterteilt $\mathbb{R}^2$ in zwei offene und zusammenhängende Mengen. Das heißt, es gibt zwei offene und zusammenhängende Mengen $U_1, U_2 \subseteq \mathbb{R}^2$ mit*

$$U_1 \cup U_2 = \mathbb{R}^2 \setminus \gamma([a,b]) \quad \textit{und} \quad U_1 \cap U_2 = \emptyset \, .$$

Dabei ist genau eine der Mengen U_1, U_2 beschränkt. □

Definition 3.6.12 *Sei $\gamma : [a,b] \to \mathbb{R}^2$ eine geschlossene Jordan-Kurve in $\mathbb{R}^2$ und seien U_1, U_2 wie in Satz 3.6.11. Die Vereinigung von $\gamma([a,b])$ mit derjenigen der Mengen U_1, U_2, welche beschränkt ist, wird die* **von γ eingeschlossene Fläche** *genannt. Wir bezeichnen diese Menge mit $\Omega(\gamma)$.*

Jetzt können wir das oben angekündigte Resultat formulieren.

Satz 3.6.13 (Satz von Stokes in der Ebene) *Sei $\gamma : [a,b] \to \mathbb{R}^2$ eine positiv orientierte aus regulären C^1-Kurven zusammengesetzte einfach geschlossene stückweise stetig differenzierbare Kurve, sei U eine offene Teilmenge von $\mathbb{R}^2$ mit $\Omega(\gamma) \subseteq U$ und sei $F = (F_1, F_2) : U \to \mathbb{R}^2$ stetig differenzierbar. Dann gilt*

$$\int_\gamma \langle F(x), \mathrm{d}x \rangle = \int_{\Omega(\gamma)} \left(\frac{\partial F_2}{\partial x_1}(x) - \frac{\partial F_1}{\partial x_2}(x) \right) \mathrm{d}x \, .$$

Beweis. Indem man $\Omega(\gamma)$ in geeigneter Weise zerlegt, kann man o.B.d.A. annehmen, dass

$$\Omega(\gamma) = \mathrm{Ord}(\varphi)$$

für eine C^1-Funktion $\varphi : [0,b] \to [0,+\infty[$ und dass γ aus den C^1-Kurven

$$\begin{aligned}
&\gamma_{(1)} : [0,b] \to \mathbb{R}^2 \, , && \gamma_{(1)}(t) := (t,0) \, , \\
&\gamma_{(2)} : [0,\varphi(b)] \to \mathbb{R}^2 \, , && \gamma_{(2)}(t) := (b,t) \, , \\
&\gamma_{(3)} : [-b,0] \to \mathbb{R}^2 \, , && \gamma_{(3)}(t) := (-t,\varphi(-t)) \, , \\
&\gamma_{(4)} : [-\varphi(0),0] \to \mathbb{R}^2 \, , && \gamma_{(4)}(t) := (0,-t) \, ,
\end{aligned}$$

zusammengesetzt ist. Damit gilt

$$\begin{aligned}
\int_\gamma \langle F(x), \mathrm{d}x \rangle &= \sum_{i=1}^4 \int_{\gamma_{(i)}} \langle F(x), \mathrm{d}x \rangle \\
&= \int_0^b F_1(t,0)\,\mathrm{d}t + \int_0^{\varphi(b)} F_2(b,t)\,\mathrm{d}t \\
&\quad - \int_0^b (F_1(t,\varphi(t)) + F_2(t,\varphi(t))\varphi'(t))\,\mathrm{d}t - \int_0^{\varphi(0)} F_2(0,t)\,\mathrm{d}t \, .
\end{aligned}$$

Wegen

$$\frac{\mathrm{d}}{\mathrm{d}x_1} \int_0^{\varphi(x_1)} F_2(x_1,x_2)\,\mathrm{d}x_2 = F_2(x_1,\varphi(x_1))\varphi'(x_1) + \int_0^{\varphi(x_1)} \frac{\partial F_2}{\partial x_1}(x_1,x_2)\,\mathrm{d}x_2$$

erhält man andererseits

$$\begin{aligned}
&\int_{\Omega(\gamma)} \left(\frac{\partial F_2}{\partial x_1}(x) - \frac{\partial F_1}{\partial x_2}(x) \right) \mathrm{d}x \\
&= \int_0^b \int_0^{\varphi(x_1)} \frac{\partial F_2}{\partial x_1}(x_1,x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1 - \int_0^b \int_0^{\varphi(x_1)} \frac{\partial F_1}{\partial x_2}(x_1,x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1 \\
&= \int_0^b \frac{\mathrm{d}}{\mathrm{d}x_1} \int_0^{\varphi(x_1)} F_2(x_1,x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1 - \int_0^b F_2(x_1,\varphi(x_1))\varphi'(x_1)\,\mathrm{d}x_1 \\
&\quad - \int_0^b \int_0^{\varphi(x_1)} \frac{\partial F_1}{\partial x_2}(x_1,x_2)\,\mathrm{d}x_2\,\mathrm{d}x_1 \\
&= \int_0^{\varphi(b)} F_2(b,x_2)\,\mathrm{d}x_2 - \int_0^{\varphi(0)} F_2(0,x_2)\,\mathrm{d}x_2 - \int_0^b F_2(x_1,\varphi(x_1))\varphi'(x_1)\,\mathrm{d}x_1 \\
&\quad - \int_0^b (F_1(x_1,\varphi(x_1)) - F_1(x_1,0))\,\mathrm{d}x_1 \ .
\end{aligned}$$

Das ergibt die Behauptung. □

Folgerung 3.6.14 *Sei γ wie in Satz 3.6.13. Dann ist*

$$\mathrm{vol}(\Omega(\gamma)) = \int_\gamma x_1\,\mathrm{d}x_2 = -\int_\gamma x_2\,\mathrm{d}x_1 = \frac{1}{2}\int_\gamma (x_1\,\mathrm{d}x_2 - x_2\,\mathrm{d}x_1) \ .$$

Beweis. Seien $F, G : \mathbb{R}^2 \to \mathbb{R}^2$ durch

$$F(x_1,x_2) := (0,x_1) \quad \text{und} \quad G(x_1,x_2) := (-x_2,0)$$

definiert. Nach Satz 3.6.13 haben wir dann

$$\int_\gamma x_1\,\mathrm{d}x_2 = \int_\gamma \langle F(x), \mathrm{d}x\rangle = \int_{\Omega(\gamma)} \left(\frac{\partial F_2}{\partial x_1}(x) - \frac{\partial F_1}{\partial x_2}(x) \right) \mathrm{d}x = \int_{\Omega(\gamma)} \mathrm{d}x = \mathrm{vol}(\Omega(\gamma))$$

und genauso

$$-\int_\gamma x_2\,\mathrm{d}x_1 = \int_\gamma \langle G(x), \mathrm{d}x\rangle = \int_{\Omega(\gamma)} \left(\frac{\partial G_2}{\partial x_1}(x) - \frac{\partial G_1}{\partial x_2}(x) \right) \mathrm{d}x = \int_{\Omega(\gamma)} \mathrm{d}x = \mathrm{vol}(\Omega(\gamma)) \ .$$

Die dritte Gleichung folgt aus der ersten und zweiten. □

Beispiel 3.6.15 Die von der Ellipse $\gamma : [0,2\pi] \to \mathbb{R}^2$, $\gamma(t) := (a\cos(t), b\sin(t))$, eingeschlossene Fläche hat den Flächeninhalt

$$\mathrm{vol}(\Omega(\gamma)) = \pi ab \ .$$

Nach Folgerung 3.6.14 ist nämlich

$$\begin{aligned}
\mathrm{vol}(\Omega(\gamma)) &= \frac{1}{2}\int_\gamma (x_1\,\mathrm{d}x_2 - x_2\,\mathrm{d}x_1) \\
&= \frac{1}{2}\int_0^{2\pi} \left(ab\cos^2(t) + ab\sin^2(t)\right)\,\mathrm{d}t = \pi ab \ .
\end{aligned}$$

□

Wir wollen jetzt noch weitere Konsequenzen aus dem Satz von Stokes ziehen.

Definition 3.6.16 *Sei $U \subseteq \mathbb{R}^n$ offen.*

(i) *Ein* **Vektorfeld** *auf* U *ist eine Abbildung* $X : U \to \mathbb{R}^n$.

(ii) *Ist* $X = (X_1, \ldots, X_n)$ *ein stetig differenzierbares Vektorfeld auf* U, *so nennt man die Abbildung*

$$\operatorname{div}(X) : U \to \mathbb{R}\ , \quad \operatorname{div}(X)(x) := \sum_{i=1}^{n} \frac{\partial X_i}{\partial x_i}(x)\ ,$$

die **Divergenz** *von* X.

Definition 3.6.17 *Sei* $\gamma = (\gamma_1, \gamma_2) : [a, b] \to \mathbb{R}^2$ *eine positiv orientierte einfach geschlossene reguläre* C^1*-Kurve. Dann heißt die Abbildung*

$$\nu : [a, b] \to \mathbb{R}^2\ , \quad \nu(t) := (\gamma_2'(t), -\gamma_1'(t))\ ,$$

das **äußere Normalenfeld an** $\Omega(\gamma)$.

Den Vektor $\nu(t)$ erhält man, indem man $\gamma'(t)$ in negativem Sinn um dem Winkel $\pi/2$ dreht. Insbesondere sind die Vektoren $\gamma'(t)$ und $\nu(t)$ orthogonal zueinander. Außerdem weist $\nu(t)$ aus $\Omega(\gamma)$ heraus.

Satz 3.6.18 (Gaussscher Integralsatz in der Ebene) *Sei* $\gamma : [a, b] \to \mathbb{R}^2$ *eine positiv orientierte einfach geschlossene reguläre* C^1*-Kurve, sei* U *eine offene Teilmenge von* $\mathbb{R}^2$ *mit* $\Omega(\gamma) \subseteq U$ *und sei* X *ein stetig differenzierbares Vektorfeld auf* U. *Dann gilt*

$$\int_{\Omega(\gamma)} \operatorname{div}(X)(x)\, \mathrm{d}x = \int_a^b \langle X(\gamma(t)), \nu(t)\rangle\, \mathrm{d}t\ ,$$

wobei ν *das äußere Normalenfeld an* $\Omega(\gamma)$ *ist.*

Beweis. Wir wenden Satz 3.6.13 auf

$$F : U \to \mathbb{R}^2\ , \quad F(x) := (-X_2(x), X_1(x))\ ,$$

an und erhalten

$$\begin{aligned}\int_{\Omega(\gamma)} \operatorname{div}(X)(x)\, \mathrm{d}x &= \int_{\Omega(\gamma)} \left(\frac{\partial F_2}{\partial x_1}(x) - \frac{\partial F_1}{\partial x_2}(x)\right) \mathrm{d}x = \int_\gamma \langle F(x), \mathrm{d}x\rangle \\ &= \int_a^b (-X_2(\gamma(t))\gamma_1'(t) + X_1(\gamma(t))\gamma_2'(t))\, \mathrm{d}t = \int_a^b \langle X(\gamma(t)), \nu(t)\rangle\, \mathrm{d}t\ .\end{aligned}$$

□

Folgerung 3.6.19 *Seien* γ, ν, U *und* X *wie in Satz 3.6.18 und sei* $f \in C^1(U, \mathbb{R})$. *Dann ist*

$$\int_{\Omega(\gamma)} (f(x) \operatorname{div}(X)(x) + \langle \operatorname{grad}(f)(x), X(x)\rangle)\, \mathrm{d}x = \int_a^b f(\gamma(t))\, \langle X(\gamma(t)), \nu(t)\rangle\, \mathrm{d}t\ .$$

Beweis. Wie man einfach nachrechnet, ist

$$\operatorname{div}(fX) = f \operatorname{div}(x) + \langle \operatorname{grad}(f), X\rangle\ .$$

Folglich gelangt man zu der Behauptung, indem man Satz 3.6.18 auf das Vektorfeld fX anwendet.

□

Definition 3.6.20 *Sei $U \subseteq \mathbb{R}^n$ offen. Die Abbildung*

$$\Delta : C^2(U,\mathbb{R}) \to C^0(U,\mathbb{R}) \ , \quad \Delta(f) := \mathrm{div}(\mathrm{grad}(f)) \ ,$$

heißt der **Laplace-Operator** *auf U.*

Lemma 3.6.21 *Sei $U \subseteq \mathbb{R}^n$ offen.*

(i) *Für alle $f \in C^2(U,\mathbb{R})$ gilt*

$$\Delta(f) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} \ .$$

(ii) *Der Laplace-Operator ist linear, d.h. für alle $f_1, f_2 \in C^2(U,\mathbb{R})$ und alle $\alpha_1, \alpha_2 \in \mathbb{R}$ ist*

$$\Delta(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 \Delta(f_1) + \alpha_2 \Delta(f_2) \ .$$

(iii) *Für alle $f_1, f_2 \in C^2(U,\mathbb{R})$ ist*

$$\Delta(f_1 f_2) = f_1 \Delta(f_2) + f_2 \Delta(f_1) + 2 \langle \mathrm{grad}(f_1), \mathrm{grad}(f_2) \rangle \ .$$

Beweis. Die Behauptungen (i) und (ii) sind offensichtlich, die Behauptung (iii) ist leicht nachzurechnen. □

Satz 3.6.22 (Greensche Formel) *Seien γ, ν und U wie in Satz 3.6.18. Für alle $f_1 \in C^1(U,\mathbb{R})$ und alle $f_2 \in C^2(U,\mathbb{R})$ gilt dann*

$$\int_{\Omega(\gamma)} (f_1(x)\,\Delta(f_2)(x) + \langle \mathrm{grad}(f_1)(x), \mathrm{grad}(f_2)(x) \rangle \, \mathrm{d}x = \int_a^b f_1(\gamma(t))\,\langle \mathrm{grad}(f_2)(\gamma(t)), \nu(t) \rangle \, \mathrm{d}t \ .$$

Beweis. Das ist die Aussage von Folgerung 3.6.19 für $X = \mathrm{grad}(f_2)$. □