Analysis II

Skript zur Vorlesung

H. Führ A. Krieg S. Walcher

Lehrstuhl A für Mathematik
RWTH AACHEN UNIVERSITY

Analysis II

H. Führ A. Krieg S. Walcher

Prof. Dr. Hartmut Führ
Prof. Dr. Aloys Krieg
Prof. Dr. Sebastian Walcher
Lehrstuhl A für Mathematik
RWTH Aachen University
D-52056 Aachen
E-Mail: fuehr@mathA.rwth-aachen.de
krieg@rwth-aachen.de
walcher@mathA.rwth-aachen.de
Internet: http://www.mathA.rwth-aachen.de/

Inhaltsverzeichnis

Analysis II

Vorwort

Gegenstand der Analysis II sind die Differential- und Integralrechnung von reellen Funktionen einer Variablen, die Differentialrechnung von reellen Funktionen mehrerer Variablen sowie eine Einführung in die normierten Vektorräume.

Die Hörerinnen und Hörer sollten neben dieser Vorlesungsmitschrift auch Standardliteratur zur Analysis zu Rate ziehen. Zu empfehlen sind insbesondere:

BARNER, M., und FLOHR, F.: Analysis I, II; 5., 3. Aufl., de Gruyter, Berlin-New York 2000, 1996.

FORSTER, O.: Analysis 1, 2; 11., 10. Aufl., Vieweg, Wiesbaden 2013.

HEUSER, H.: Lehrbuch der Analysis 1, 2; 17., 14. Aufl., Teubner, Wiesbaden 2009, 2008.

KÖNIGSBERGER, K.: Analysis 1, 2; 6., 5. Aufl., Springer-Verlag, Berlin-Heidelberg-New York 2004.

WALTER, W.: Analysis 1, 2; 7., 5. Aufl., Springer-Verlag, Berlin-Heidelberg-New York 2004, 2002.

Hinsichtlich der Anwendungen empfiehlt es sich, auch einen Blick in

MEYBERG, K., und VACHENAUER, P.: Höhere Mathematik 1; 6. Aufl., Springer, Berlin-Heidelberg-New York 2001.

zu werfen.

Aachen, im März 2018

Hartmut FÜHR
Aloys KRIEG
Sebastian WALCHER

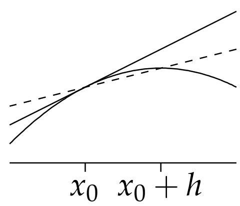
Kapitel VII.

Differentialrechnung

In diesem Kapitel soll die Differentialrechnung von reellwertigen Funktionen einer reellen Variablen beschrieben werden. Dazu sei stets $D \subset \mathbb{R}$, $D \neq \emptyset$.

§1. Die Ableitung einer Funktion

Geometrisch gesehen, möchte man eine Funktion in der Nähe eines Punktes durch eine Gerade approximieren. Die Tangente liefert die beste Approximation. Man versucht nun, die Steigung der Tangente als Grenzwert der Steigung von Sekanten auszurechnen. Für $h \neq 0$ ist die Steigung der Sekante, also der Geraden durch $(x_0, f(x_0))$ und $(x_0 + h, f(x_0 + h))$, gegeben durch



$$\frac{f(x_0 + h) - f(x_0)}{h}.$$

(1.1) Definition. Sei $f : D \to \mathbb{R}$ eine Funktion und x_0 ein innerer Punkt von $D \subset \mathbb{R}$. Man nennt f in x_0 *differenzierbar*, wenn der Limes

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

[im Sinne von V(2.3)] existiert. Dieser Grenzwert heißt *Ableitung* von f im Punkt x_0 oder an der Stelle x_0 und wird mit $f'(x_0) = \frac{df}{dx}(x_0)$ bezeichnet. Man nennt dann die Gerade

$$\{(x, y) \in \mathbb{R} \times \mathbb{R};\ y = f(x_0) + f'(x_0) \cdot (x - x_0)\}$$

die *Tangente* an den Graphen von f im Punkt $(x_0, f(x_0))$.

Die Tangente ist eine lineare Approximation an den Graphen von f. Der leicht zu berechnende Wert $y = f(x_0) + f'(x_0) \cdot (x - x_0)$ sollte in der Nähe des i. A. schwer zu berechnenden Funktionswertes $f(x)$ liegen, solange x nahe bei x_0 ist. Diesen anschaulichen Sachverhalt werden wir in (1.3) mathematisch präzisieren.

Elementar sind die

(1.2) Beispiele. a) Sei $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$, und $x_0 \in \mathbb{R}$. Dann gilt

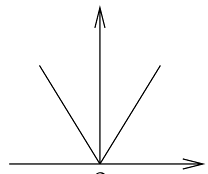
$$\frac{f(x_0+h)-f(x_0)}{h} = \frac{(x_0+h)^2 - x_0^2}{h} = 2x_0 + h \xrightarrow[h\to 0]{} 2x_0.$$

Also ist f in x_0 differenzierbar mit $f'(x_0) = 2x_0$.
b) Wir betrachten $\exp : \mathbb{R} \to \mathbb{R}$ und $x_0 \in \mathbb{R}$. Dann folgt

$$\frac{\exp(x_0+h)-\exp(x_0)}{h} = \exp(x_0)\cdot\frac{\exp(h)-1}{h} = \exp(x_0)\cdot\sum_{k=1}^{\infty}\frac{1}{k!}h^{k-1} \xrightarrow[h\to 0]{} \exp(x_0),$$

da die Potenzreihe in h stetig ist. Also ist exp in x_0 differenzierbar mit $\exp'(x_0) = \exp(x_0)$.
c) Wir betrachten $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$, und $x_0 = 0$. Dann gilt

$$\frac{f(0+h)-f(0)}{h} = \frac{|h|}{h} = \begin{cases} 1 & \text{für } h > 0, \\ -1 & \text{für } h < 0. \end{cases}$$



Also existiert kein Limes für $h \to 0$. Demnach ist f im Punkt 0 nicht differenzierbar. f ist aber stetig, insbesondere auch im Punkt 0.
d) Wir betrachten $f(x) = \sin x$. Dann folgt aus den Additionstheoremen IV(3.6) sowie V(5.2)

$$\begin{aligned} \frac{\sin(x+h)-\sin x}{h} &= \frac{\sin x\cdot\cos h + \cos x\cdot\sin h - \sin x}{h} \\ &= \sin x\cdot\frac{\cos h - 1}{h} + \cos x\cdot\frac{\sin h}{h} \xrightarrow[h\to 0]{} \cos x. \end{aligned}$$

Also ist der Sinus differenzierbar in jedem $x \in \mathbb{R}$ mit $\sin' x = \cos x$. Analog zeigt man, dass der Cosinus differenzierbar in jedem $x \in \mathbb{R}$ ist mit $\cos' x = -\sin x$.

Anschaulich gesprochen, hat der Graph einer differenzierbaren Funktion keine “Ecken”.

Wir fügen eine physikalische Interpretation hinzu. Die Funktion $s : \mathbb{R} \to \mathbb{R}$, $t \mapsto s(t)$, gebe den zurückgelegten Weg zum Zeitpunkt t an. Der Differenzenquotient $\frac{s(t+h)-s(t)}{h}$ beschreibt die Durchschnittsgeschwindigkeit in der Zeit zwischen t und $t+h$. Mit dem Limes $h \to 0$ erhält man natürlich die Momentangeschwindigkeit $s'(t)$ zum Zeitpunkt t.

Wir geben nun eine Charakterisierung der Differenzierbarkeit.

(1.3) Satz. *Sei $f : D \to \mathbb{R}$ eine Funktion und x_0 ein innerer Punkt von $D \subset \mathbb{R}$. Dann sind äquivalent:*

(i) f ist differenzierbar in x_0.

(ii) $\lim\limits_{x\to x_0} \dfrac{f(x)-f(x_0)}{x-x_0}$ existiert.

(iii) Es gibt ein $a \in \mathbb{R}$, so dass zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit der Eigenschaft

$$\left|\frac{f(x_0+h)-f(x_0)}{h} - a\right| < \varepsilon \quad \textit{für alle } 0 < |h| < \delta.$$

(iv) Es gibt ein $b \in \mathbb{R}$ und eine in x_0 stetige Funktion $\varphi : D \to \mathbb{R}$ mit $\varphi(x_0) = 0$ und

$$f(x) = f(x_0) + (x - x_0)b + (x - x_0)\varphi(x) \quad \text{für alle } x \in D.$$

In diesem Fall gilt

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

Beweis. "(i) $\Leftrightarrow$ (ii)" Wir setzen $h = x - x_0$, also $x = x_0 + h$. Dann ist $h \to 0$ äquivalent zu $x \to x_0$.
"(i) $\Leftrightarrow$ (iii)" Man verwende die Definition V(2.3) des Funktionenlimes.
"(ii) $\Rightarrow$ (iv)" Sei $b = \lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$. Wir definieren

$$\varphi : D \to \mathbb{R},\ x \mapsto \begin{cases} \dfrac{f(x) - f(x_0)}{x - x_0} - b, & \text{falls } x \neq x_0, \\ 0, & \text{falls } x = x_0. \end{cases}$$

Dann gilt $f(x) = f(x_0) + (x - x_0)b + (x - x_0)\varphi(x)$ für alle $x \in D$. Aus

$$\lim_{x \to x_0} \varphi(x) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} - b = 0 = \varphi(x_0)$$

folgt die Stetigkeit von φ im Punkt x_0.
"(iv) $\Rightarrow$ (ii)" Aus der Darstellung in (iv) ergibt sich

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} (b + \varphi(x)) = b + \varphi(x_0) = b$$

aufgrund der Stetigkeit von φ in x_0. □

Wir ziehen sofort zwei einfache Folgerungen.

(1.4) Korollar. *Sei $f : D \to \mathbb{R}$ eine Funktion, die in einem inneren Punkt x_0 von $D \subset \mathbb{R}$ differenzierbar ist. Dann gilt:*
a) f ist stetig in x_0.
b) Es gibt ein $\delta > 0$ mit $U_\delta(x_0) \subset D$ und ein $M > 0$, so dass

$$|f(x) - f(x_0)| \leqslant M \cdot |x - x_0| \quad \text{für alle } x \in U_\delta(x_0).$$

Beweis. Wir verwenden für f die Darstellung aus (1.3) (iv).
a) Die rechte Seite ist nach V(1.9) stetig in x_0. Also ist auch f stetig in x_0.
b) Nach V(1.9) gibt es ein $\delta > 0$ mit $U_\delta(x_0) \subset D$ und ein $L > 0$ mit $|\varphi(x)| < L$ für alle $x \in U_\delta(x_0)$. Es folgt mit $M = |b| + L$

$$|f(x) - f(x_0)| = |(x - x_0)b + (x - x_0)\varphi(x)| \leqslant |x - x_0| \cdot (|b| + |\varphi(x)|) \leqslant M \cdot |x - x_0|. \quad \square$$

Aus der Differenzierbarkeit folgt die Stetigkeit, sogar die (lokale) LIPSCHITZ-Stetigkeit in einer geeigneten Umgebung des Punktes x_0. Die Umkehrung ist nicht richtig, wie (1.2) c) zeigt. Es gibt sogar stetige Funktionen $f : \mathbb{R} \to \mathbb{R}$, die in keinem Punkt differenzierbar sind. Man vergleiche BARNER/FLOHR, §8.1.

Wir kommen nun analog zum Limesbegriff für Funktionen zur einseitigen Differenzierbarkeit.

(1.5) Definition. Sei $f : D \to \mathbb{R}$ eine Funktion, $x_0 \in D \subset \mathbb{R}$ und $\delta > 0$ mit der Eigenschaft $[x_0, x_0 + \delta) \subset D$ (bzw. $(x_0 - \delta, x_0] \subset D$). Man nennt f *rechtsseitig* (bzw. *linksseitig*) *differenzierbar* in x_0, wenn der Limes

$$\lim_{h \downarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad \Big[\text{bzw. } \lim_{h \uparrow 0} \frac{f(x_0 + h) - f(x_0)}{h}\Big]$$

existiert. Wir verwenden dafür die Bezeichnung $f'_+(x_0)$ [bzw. $f'_-(x_0)$].

Aus V(2.12) folgt dann unmittelbar das

(1.6) Lemma. *Sei $f : D \to \mathbb{R}$ eine Funktion und x_0 ein innerer Punkt von $D \subset \mathbb{R}$. Dann sind äquivalent:*
(i) f ist differenzierbar in x_0.
(ii) $f'_+(x_0)$ und $f'_-(x_0)$ existieren und haben den gleichen Wert.

Wenn man sich $f(x) = |x|$ im Nullpunkt anschaut, so hat man natürlich $f'_+(0) = 1$ und $f'_-(0) = -1$ (vgl. (1.2)).

Nun führen wir globale Differenzierbarkeit ein.

(1.7) Definition. Sei $f : D \to \mathbb{R}$ eine Funktion und $D \subset \mathbb{R}$ enthalte keine isolierten Punkte.
a) f heißt *differenzierbar auf D*, wenn f in jedem inneren Punkt von D differenzierbar ist und in jedem Randpunkt, der zu D gehört, einseitig differenzierbar ist.
b) Ist f differenzierbar auf D, so heißt die Funktion

$$f' : D \to \mathbb{R},\ x \mapsto f'(x) \text{ (bzw. } f'_+(x) \text{ oder } f'_-(x),\ \text{falls } x \text{ Randpunkt)},$$

die *Ableitung von f*.
Man schreibt auch $f' = \frac{df}{dx} = \frac{d}{dx} f$.

Wir formulieren die Rechenregeln in dem

(1.8) Satz. *Gegeben seien Funktionen $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$, die in einem inneren Punkt x_0 von $D \subset \mathbb{R}$ differenzierbar sind. Dann gilt*
a) Für jedes $\alpha \in \mathbb{R}$ ist αf differenzierbar in x_0 mit $(\alpha f)'(x_0) = \alpha f'(x_0)$.
b) $f + g$ ist differenzierbar in x_0 mit $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.
c) (Produktregel) $f \cdot g$ ist differenzierbar in x_0 mit

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0).$$

d) (Quotientenregel) Gilt $g(x_0) \neq 0$, so ist auch f/g differenzierbar in x_0 mit

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Beweis. a) Es gilt

$$\frac{(\alpha f)(x_0 + h) - (\alpha f)(x_0)}{h} = \alpha \cdot \frac{f(x_0 + h) - f(x_0)}{h} \xrightarrow[h \to 0]{} \alpha f'(x_0).$$

b) Man hat

$$\frac{(f + g)(x_0 + h) - (f + g)(x_0)}{h} = \frac{f(x_0 + h) - f(x_0)}{h} + \frac{g(x_0 + h) - g(x_0)}{h} \xrightarrow[h \to 0]{} f'(x_0) + g'(x_0).$$

c) Weil g nach (1.4) auch stetig in x_0 ist, folgt

$$\begin{aligned}\frac{(f \cdot g)(x_0 + h) - (f \cdot g)(x_0)}{h} &= \frac{f(x_0 + h) - f(x_0)}{h} \cdot g(x_0 + h) + f(x_0) \cdot \frac{g(x_0 + h) - g(x_0)}{h} \\ &\xrightarrow[h \to 0]{} f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0).\end{aligned}$$

d) Nach V(3.9) existiert wegen der Stetigkeit von g in x_0 ein $\delta > 0$, so dass $U_\delta(x_0) \subset D$ und $g(x) \neq 0$ für alle $x \in U_\delta(x_0)$. Für $0 < |h| < \delta$ folgert man dann

$$\begin{aligned}&\frac{\left(\frac{f}{g}\right)(x_0 + h) - \left(\frac{f}{g}\right)(x_0)}{h} \\ &= \frac{1}{g(x_0 + h)g(x_0)} \cdot \left[g(x_0) \cdot \frac{f(x_0 + h) - f(x_0)}{h} - f(x_0) \cdot \frac{g(x_0 + h) - g(x_0)}{h}\right] \\ &\xrightarrow[h \to 0]{} \frac{1}{g(x_0)^2}[g(x_0)f'(x_0) - f(x_0)g'(x_0)].\end{aligned}$$

□

Der Satz besagt, dass die Menge der auf D differenzierbaren Funktionen eine $\mathbb{R}$-Algebra ist. Die Ableitung ist eine lineare Abbildung, wegen c) jedoch kein Algebrenhomomorphismus.

(1.9) Korollar. *a) Jede konstante Funktion ist differenzierbar und die Ableitung ist identisch* 0.
b) Sei $p(x) = a_0 + a_1 x + \ldots + a_n x^n$, $n \in \mathbb{N}_0$, *ein reelles Polynom. Dann ist* $p(x)$ *differenzierbar in jedem Punkt* $x_0 \in \mathbb{R}$ *mit*

$$p'(x_0) = a_1 + 2a_2 x_0 + \ldots + n a_n x_0^{n-1},$$

d. h.

$$\frac{d}{dx} \sum_{k=0}^{n} a_k x^k = \sum_{k=1}^{n} k a_k x^{k-1}.$$

Beweis. a) Aus $f(x) = c$ folgt

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{c - c}{h} = 0 = f'(x_0).$$

b) Wir verwenden eine Induktion nach n, wobei die Behauptung für $n = 0$ nach a) richtig ist. Für $n = 1$ haben wir

$$\frac{p(x_0 + h) - p(x_0)}{h} = \frac{a_0 + a_1(x_0 + h) - (a_0 + a_1 x_0)}{h} = a_1 = p'(x_0).$$

Durch Induktion folgt mit (1.8)

$$\frac{d}{dx} x^{n+1} = \frac{d}{dx}(x \cdot x^n) = \left(\frac{d}{dx} x\right) \cdot x^n + x \cdot \left(\frac{d}{dx} x^n\right) = 1 \cdot x^n + x \cdot n x^{n-1} = (n+1) x^n.$$

Die Behauptung für *grad* $p = n + 1$ erhält man nun aus (1.8) mit der Zerlegung $p(x) = q(x) + a_{n+1} x^{n+1}$, *grad* $q \leqslant n$ oder $q \equiv 0$. □

Als "Grenzwert" von Polynomen erhalten wir Potenzreihen.

(1.10) Satz. *Sei* $a \in \mathbb{R}$ *und* $f(x) = \sum_{n=0}^{\infty} a_n (x - a)^n$ *eine reelle Potenzreihe mit Konvergenzradius* $R > 0$. *Dann ist* f *auf* $U_R(a)$ *differenzierbar mit*

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - a)^{n-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} (x - a)^k$$

und diese Potenzreihe hat ebenfalls den Konvergenzradius R.

Beweis. Sei R' der Konvergenzradius von $g(x) = \sum_{n=1}^{\infty} n a_n (x - a)^{n-1}$. Dann folgt $R' = R$ aus dem ABELschen Lemma IV(4.3). Sei ohne Einschränkung $a = 0$ und $x_0 \in U_R(0)$ sowie $\delta > 0$, so dass $r := |x_0| + \delta < R$. Für $|h| < \delta$, $n \geqslant 2$ gilt dann

$$\left| \sum_{k=2}^{n} \binom{n}{k} h^{k-2} x_0^{n-k} \right| = \left| \sum_{k=2}^{n} \frac{n(n-1)}{k(k-1)} \binom{n-2}{k-2} h^{k-2} x_0^{n-k} \right|$$

$$\leqslant n(n-1) \sum_{k=2}^{n} \binom{n-2}{k-2} |h|^{k-2} |x_0|^{n-k} = n(n-1)(|h| + |x_0|)^{n-2} \leqslant n(n-1) r^{n-2}.$$

Damit folgt für $0 < |h| < \delta$

$$\begin{aligned}
\left|\frac{f(x_0+h)-f(x_0)}{h} - g(x_0)\right| &= \left|\sum_{n=1}^{\infty} a_n \left(\frac{(x_0+h)^n - x_0^n}{h} - n x_0^{n-1}\right)\right| \\
= \left|\sum_{n=2}^{\infty} a_n \sum_{k=2}^{n} \binom{n}{k} h^{k-1} x_0^{n-k}\right| &\leqslant |h| \cdot \sum_{n=2}^{\infty} |a_n| \left|\sum_{k=2}^{n} \binom{n}{k} h^{k-2} x_0^{n-k}\right| \\
\leqslant |h| \cdot \sum_{n=2}^{\infty} n(n-1)|a_n| r^{n-2} &\xrightarrow[h\to 0]{} 0,
\end{aligned}$$

denn $\sum_{n=2}^{\infty} n(n-1)|a_n|r^{n-2}$ konvergiert nach dem ABELschen Lemma IV(4.3). Also ist f differenzierbar in x_0 mit $f'(x_0) = g(x_0)$. □

Nun wenden wir den Satz auf bekannte Potenzreihen an

(1.11) Korollar. *a) Die Funktion* exp, sin *und* cos *sind auf* $\mathbb{R}$ *differenzierbar mit*

$$\exp'(x) = \exp(x), \ \sin' x = \cos x, \ \cos' x = -\sin x, \ x \in \mathbb{R}.$$

b) Der Tangens tan *ist auf* $\mathbb{R} \setminus \left(\frac{\pi}{2} + \mathbb{Z}\pi\right)$ *differenzierbar mit*

$$\tan' x = \frac{1}{\cos^2 x} = 1 + \tan^2 x.$$

Der Cotangens cot *ist auf* $\mathbb{R} \setminus \mathbb{Z}\pi$ *differenzierbar mit*

$$\cot' x = \frac{-1}{\sin^2 x} = -1 - \cot^2 x.$$

Beweis. a) Die Funktionen werden durch Potenzreihen auf $\mathbb{R}$ gegeben und können nach (1.10) differenziert werden. Es folgt

$$\begin{aligned}
\exp'(x) &= \frac{d}{dx}\left(\sum_{n=0}^{\infty} \frac{1}{n!} x^n\right) = \sum_{n=1}^{\infty} \frac{1}{n!} n x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = \exp(x), \\
\sin'(x) &= \frac{d}{dx}\left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = \cos x, \\
\cos'(x) &= \frac{d}{dx}\left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}\right) = \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n-1)!} x^{2n-1} = -\sin x.
\end{aligned}$$

b) Man verwende a) und die Quotientenregel (1.8)

$$\begin{aligned}
\tan' x &= \frac{d}{dx}\left(\frac{\sin x}{\cos x}\right) = \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x} \\
&= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = 1 + \tan^2 x, \\
\cot' x &= \frac{d}{dx}\left(\frac{\cos x}{\sin x}\right) = \frac{\cos' x \sin x - \cos x \sin' x}{\sin^2 x} \\
&= \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} = \frac{-1}{\sin^2 x} = -1 - \cot^2 x.
\end{aligned}$$

□

Nun beschäftigen wir uns mit der Differentiation bei Komposition.

(1.12) Satz. *(Kettenregel)*
Sei $f : D \to \mathbb{R}$ differenzierbar in einem inneren Punkt x_0 von $D \subset \mathbb{R}$. Ist $g : D' \to \mathbb{R}$ eine Funktion mit $f(D) \subset D'$, die in $y_0 = f(x_0)$ differenzierbar ist, so ist $g \circ f : D \to \mathbb{R}$ differenzierbar in x_0 mit

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

Beweis. Weil g differenzierbar in y_0 ist, ist

$$h : D' \to \mathbb{R},\ y \mapsto \begin{cases} \frac{g(y)-g(y_0)}{y-y_0}, & \text{falls } y \neq y_0, \\ g'(y_0), & \text{falls } y = y_0, \end{cases}$$

stetig in y_0 mit der Eigenschaft

$$g(y) - g(y_0) = (y - y_0) \cdot h(y) \quad \text{für alle} \quad y \in D'.$$

Setzt man nun $y = f(x)$, so folgt für alle $x \in D, x \neq x_0$:

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} \cdot h(f(x)) \xrightarrow[x \to x_0]{} f'(x_0) \cdot g'(y_0),$$

denn es gilt $\lim\limits_{x \to x_0} f(x) = f(x_0) = y_0$ wegen der Stetigkeit von f in x_0. □

Es folgt eine einfache

(1.13) Bemerkung. Könnte man voraussetzen, dass es ein $\delta > 0$ mit der Eigenschaft gibt, dass $f(x) - f(x_0) \neq 0$ für alle $x \in D$ mit $0 < |x - x_0| < \delta$, so liefe ein (intuitiv einsichtiger) Beweis wie folgt:

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \frac{f(x) \cdot f(x_0)}{x - x_0}.$$

Da $f(x) \to f(x_0)$ für $x \to x_0$ (Stetigkeit von f), liefert der Grenzübergang die Behauptung. Der tatsächliche (allgemein gültige) Beweis von (1.12) umschifft diese Problematik von $f(x) - f(x_0)$ elegant. □

In der Praxis gilt es, g und f zu bestimmen, wenn man einen konkreten Ausdruck zu differenzieren hat.

(1.14) Beispiele. a) Sei $h : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sin(x^2 + x)$. Man hat

$$h = g \circ f \quad \text{mit} \quad g(y) = \sin y \quad \text{und} \quad f(x) = x^2 + x.$$

Dann gilt $g'(y) = \cos y$ und $f'(x) = 2x + 1$, also mit (1.12)

$$h'(x) = f'(x) \cdot g'(f(x)) = (2x + 1) \cdot \cos(x^2 + x).$$

b) Sei $a > 0$ und $h : \mathbb{R} \to \mathbb{R}$, $x \mapsto a^x = \exp(x \ln a)$. Man hat

$$h = g \circ f \quad \text{mit} \quad g(y) = \exp\, y \quad \text{und} \quad f(x) = x \ln a.$$

Dann gilt $g'(y) = \exp\, y$ und $f'(x) = \ln a$, also mit (1.12)

$$\frac{d}{dx}(a^x) = h'(x) = g'(f(x)) \cdot f'(x) = \exp(x \ln a) \cdot \ln a = a^x \cdot \ln a.$$

Nun beschäftigen wir uns mit der Ableitung der Umkehrfunktion.

(1.15) Satz. *Sei $I \subset \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{R}$ stetig, injektiv sowie differenzierbar in x_0 mit $f'(x_0) \neq 0$. Dann ist die Umkehrfunktion $f^{-1} : W \to \mathbb{R}$, $W = f(I)$, differenzierbar in $y_0 = f(x_0)$ mit*

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

Beweis. f und f^{-1} sind nach II(3.12), V(1.15) und V(3.10) stetig und streng monoton. Sei $(y_n)_{n\geqslant 1}$ eine Folge in $W \backslash \{y_0\}$ mit $\lim\limits_{n\to\infty} y_n = y_0$. Weil f^{-1} injektiv ist, ist $(x_n)_{n\geqslant 1} := (f^{-1}(y_n))_{n\geqslant 1}$ eine Folge in $I \backslash \{x_0\}$. Da f^{-1} stetig ist, gilt $\lim\limits_{n\to\infty} x_n = f^{-1}(y_0) = x_0$. Weil f in x_0 differenzierbar ist, erhält man

$$\begin{aligned}\frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} &= \frac{f^{-1}(f(x_n)) - f^{-1}(f(x_0))}{f(x_n) - f(x_0)} \\ &= \frac{1}{\frac{f(x_n)-f(x_0)}{x_n - x_0}} \xrightarrow[n\to\infty]{} \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.\end{aligned}$$ □

Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$, ist differenzierbar und bijektiv. Die Umkehrfunktion $f^{-1} : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sqrt[3]{x} = sgn\, x \cdot \sqrt[3]{|x|}$, ist in $x_0 = 0$ aber nicht differenzierbar.

Die Ergebnisse von (1.15) verwenden wir in dem

(1.16) Korollar. *a)* $\frac{d}{dx}\ln x = \frac{1}{x}$ *für* $x > 0$.

b) $\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1}$ *für* $x > 0, \alpha \in \mathbb{R}$.

c) $\frac{d}{dx}(x^x) = (1+\ln x)\cdot x^x$ *für* $x > 0$.

d) $\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}}$ *für* $|x| < 1$.

e) $\frac{d}{dx}\arccos x = \frac{-1}{\sqrt{1-x^2}}$ *für* $|x| < 1$.

f) $\frac{d}{dx}\arctan x = \frac{1}{1+x^2}$ *für* $x \in \mathbb{R}$.

g) $\frac{d}{dx}$ *arccot* $x = \frac{-1}{1+x^2}$ *für* $x \in \mathbb{R}$.

Beweis. a) Es gilt $\ln x = f^{-1}(x)$ für $f(x) = \exp\, x = f'(x)$. Also gilt

$$\frac{d}{dx}\ln x = \frac{1}{f'(f^{-1}(x))} = \frac{1}{\exp(\ln x)} = \frac{1}{x} \quad \text{für } x > 0.$$

b) Es gilt nach der Kettenregel und a) für $x > 0$:

$$\frac{d}{dx}(x^\alpha) = \frac{d}{dx}\left(e^{\alpha\ln x}\right) = e^{\alpha\ln x}\cdot\alpha\,\frac{1}{x} = \alpha x^{\alpha-1}.$$

c) Man hat für $x > 0$:

$$\frac{d}{dx}(x^x) = \frac{d}{dx}(e^{x\ln x}) = e^{x\ln x}\left(1\cdot\ln x + x\cdot\frac{1}{x}\right) = (1+\ln x)\cdot x^x.$$

d) Für $f(x) = \sin x$ gilt $f'(x) = \cos x = \sqrt{1-\sin^2 x} > 0,\ |x| < \frac{\pi}{2}$. Also folgt für $|y| < 1$

$$\frac{d}{dy}\arcsin y = \frac{1}{\sqrt{1-\sin^2(\arcsin y)}} = \frac{1}{\sqrt{1-y^2}}.$$

e) Für $f(x) = \cos x$ gilt $f'(x) = -\sin x = -\sqrt{1-\cos^2 x} < 0,\ 0 < x < \pi$. Also hat man für $|y| < 1$

$$\frac{d}{dy}\arccos y = \frac{-1}{\sqrt{1-\cos^2(\arccos y)}} = \frac{-1}{\sqrt{1-y^2}}.$$

f) Es gilt $\tan' x = 1+\tan^2 x > 0$ für $|x| < \frac{\pi}{2}$ nach (1.11). Demnach erhält man für alle $y \in \mathbb{R}$

$$\frac{d}{dy}\arctan y = \frac{1}{1+\tan^2(\arctan y)} = \frac{1}{1+y^2}.$$

g) Man hat $\cot' x = -1-\cot^2 x$ für $0 < x < \pi$ nach (1.11). Daher folgt für alle $y \in \mathbb{R}$

$$\frac{d}{dy}\text{arccot}\, y = \frac{-1}{1+\cot^2(\text{arccot}\, y)} = \frac{-1}{1+y^2}.$$

□

§2. Mittelwertsätze der Differentialrechnung

In diesem Paragrafen sollen die Mittelwertsätze für differenzierbare Funktionen mit ihren Anwendungen hergeleitet werden.

(2.1) Definition. Sei $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$. Man sagt, dass f in x_0 ein *relatives* oder *lokales Maximum* [bzw. *Minimum*] hat, wenn ein $\delta > 0$ existiert, so dass

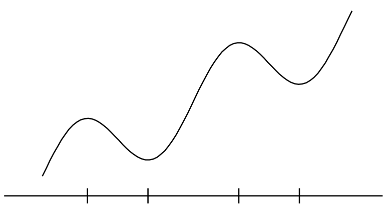
$$f(x) \leqslant f(x_0) \quad [\text{bzw. } f(x) \geqslant f(x_0)] \quad \text{für alle} \quad x \in U_\delta(x_0) \cap D.$$

f hat in x_0 ein *absolutes* oder *globales Maximum* [bzw. *Minimum*], wenn

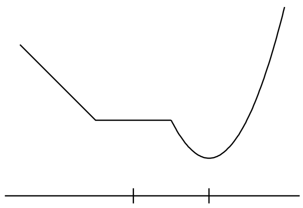
$$f(x) \leqslant f(x_0) \quad [\text{bzw. } f(x) \geqslant f(x_0)] \quad \text{für alle} \quad x \in D.$$

Ein *relatives* (*absolutes*) *Extremum* ist ein relatives (absolutes) Minimum oder Maximum.

Der Veranschaulichung dienen die folgenden Graphen



relative Extrema



relative Minima

(2.2) Satz. *Sei $f : D \to \mathbb{R}$ eine Funktion, x_0 ein innerer Punkt von D und f differenzierbar in x_0. Wenn f in x_0 ein relatives Extremum hat, dann gilt $f'(x_0) = 0$.*

Beweis. f habe in x_0 ohne Einschränkung ein relatives Maximum. Sei $\delta > 0$ mit $U_\delta(x_0) \subset D$ und $f(x) \leqslant f(x_0)$ für alle $x \in U_\delta(x_0)$. Dann gilt für alle $0 < h < \delta$

$$\frac{f(x_0+h) - f(x_0)}{h} \leqslant 0, \quad \text{also } f'_+(x_0) \leqslant 0,$$
$$\frac{f(x_0-h) - f(x_0)}{-h} \geqslant 0, \quad \text{also } f'_-(x_0) \geqslant 0.$$

Weil f in x_0 differenzierbar ist, folgt mit (1.6)

$$f'(x_0) = f'_+(x_0) = f'_-(x_0) = 0. \qquad \square$$

Auf die Voraussetzungen des Satzes gehen wir ein in den

(2.3) Bemerkungen. a) Die Bedingung, dass x_0 ein innerer Punkt von D ist, ist wesentlich, da z. B. $f : [0,1] \to \mathbb{R}$, $x \mapsto x$, in $x_0 = 0$ und $x_1 = 1$ relative Extrema hat. Es gilt aber $f'(x_0) = f'(x_1) = 1$.
b) Die Bedingung $f'(x_0) = 0$ ist notwendig, aber nicht hinreichend für die Existenz eines relativen Extremums in x_0. Man betrachte

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^3,\ f'(x) = 3x^2,\ x_0 = 0,\ f'(0) = 0.$$

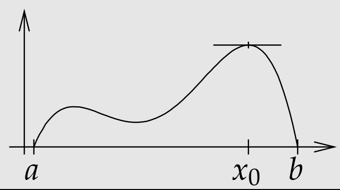
c) Es kann auch ein Extremum vorliegen, ohne dass die Funktion in x_0 differenzierbar ist. Man betrachte

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto |x|,\ x_0 = 0.$$

Wie man auf die Existenz von Nullstellen der Ableitung schließen kann, zeigt der

(2.4) Satz von Michel ROLLE (1652-1719).

Seien $a, b \in \mathbb{R}$, $a < b$, $f : [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar mit $f(a) = f(b) = 0$. Dann existiert (mindestens) ein $x_0 \in (a,b)$ mit $f'(x_0) = 0$.



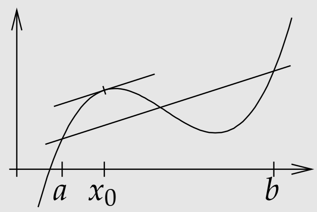
Beweis. Falls $f(x) = 0$ für alle $x \in [a,b]$, wählt man $x_0 = \frac{a+b}{2}$. Andernfalls existiert ohne Einschränkung ein $x_1 \in (a,b)$ mit $f(x_1) > 0$ (sonst betrachte man $-f$). Nach dem Satz vom Minimum und Maximum V(3.7) gibt es ein $x_0 \in [a,b]$ mit $f(x) \leqslant f(x_0)$ für alle $x \in [a,b]$. Wegen $f(x_1) > 0$ gilt $x_0 \in (a,b)$. Dann hat f in x_0 ein Extremum und $f'(x_0) = 0$ folgt aus (2.2). $\square$

Man kann den Satz von ROLLE auch so formulieren, dass zwischen je zwei Nullstellen einer differenzierbaren Funktion auf einem Intervall stets eine Nullstelle der Ableitung liegt.

(2.5) Mittelwertsatz der Differentialrechnung.

Seien $a, b \in \mathbb{R}$, $a < b$, $f : [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann existiert (mindestens) ein $x_0 \in (a,b)$ mit

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$



Die Steigung der Tangente in x_0 ist gleich der Steigung der Sekante durch $(a, f(a))$ und $(b, f(b))$. Nun ist (2.4) wegen $f(a) = f(b) = 0$ ein Spezialfall von (2.5).

Beweis. Wir definieren

$$F : [a, b] \to \mathbb{R},\ x \mapsto f(x) - f(a) - (x - a) \cdot \frac{f(b) - f(a)}{b - a}.$$

Mit f ist F stetig und auf (a, b) differenzierbar. Wegen $F(a) = F(b) = 0$ kann man (2.4) anwenden und erhält ein $x_0 \in (a, b)$ mit

$$0 = F'(x_0) = f'(x_0) - 1 \cdot \frac{f(b) - f(a)}{b - a}.$$ □

(2.6) Beispiel. Für alle $x_1, x_2 \in \mathbb{R}$ gilt

$$|\cos x_1 - \cos x_2| \leqslant |x_1 - x_2|.$$

Zum Nachweis sei o. B. d. A. $x_1 < x_2$. Nach dem Mittelwertsatz gibt es ein $\xi \in (x_1, x_2)$ derart, dass

$$\frac{\cos x_2 - \cos x_1}{x_2 - x_1} = \cos'(\xi) = -\sin \xi.$$

Übergang zum Betrag und die (allgemein gültige) Abschätzung $|\sin \xi| \leqslant 1$ ergibt die Behauptung.

Das spezielle Beispiel ist hier nicht so wichtig, aber die Art, wie der Mittelwertsatz hier für Abschätzungen verwendet wurde, sollte man sich merken.

Eine weitere Verallgemeinerung ist der

(2.7) Satz. *(Verallgemeinerter Mittelwertsatz)*
Seien $a, b \in \mathbb{R}$, $a < b$, $f : [a, b] \to \mathbb{R}$ und $g : [a, b] \to \mathbb{R}$ zwei stetige Funktionen, die auf (a, b) differenzierbar sind. Dann existiert ein $x_0 \in (a, b)$ mit

$$g'(x_0) \cdot (f(b) - f(a)) = f'(x_0) \cdot (g(b) - g(a)).$$

Mit $g(x) = x$ erhält man auch (2.5) als Spezialfall von (2.7).

Beweis. Wir definieren

$$H : [a, b] \to \mathbb{R},\ x \mapsto (f(b) - f(a)) \cdot (g(x) - g(a)) - (f(x) - f(a)) \cdot (g(b) - g(a)).$$

Mit f und g ist auch H stetig und auf (a, b) differenzierbar. Wegen $H(a) = H(b) = 0$ kann man (2.4) anwenden und erhält ein $x_0 \in (a, b)$ mit

$$0 = H'(x_0) = (f(b) - f(a)) \cdot g'(x_0) - f'(x_0) \cdot (g(b) - g(a)).$$ □

Für ein Intervall I bezeichne $\overset{\circ}{I}$ das zugehörige offene Intervall mit den gleichen Grenzen. Als wichtige Folgerung aus dem Mittelwertsatz erhalten wir den

(2.8) Satz. *Seien I ein Intervall und $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$ zwei stetige Funktionen, die auf $\overset{\circ}{I}$ differenzierbar sind. Dann gilt:*

a) $f'(x) = 0$ für alle $x \in \overset{\circ}{I}$ ist äquivalent zur Existenz eines $c \in \mathbb{R}$, so dass $f(x) = c$ für alle $x \in I$.

b) $f'(x) = g'(x)$ für alle $x \in \overset{\circ}{I}$ ist äquivalent zur Existenz eines $c \in \mathbb{R}$, so dass $f(x) = g(x) + c$ für alle $x \in I$.

c) f ist genau dann monoton wachsend, wenn $f'(x) \geqslant 0$ für alle $x \in \overset{\circ}{I}$.

d) f ist genau dann monoton fallend, wenn $f'(x) \leqslant 0$ für alle $x \in \overset{\circ}{I}$.

e) f ist streng monoton wachsend, wenn $f'(x) > 0$ für alle $x \in \overset{\circ}{I}$.

f) f ist streng monoton fallend, wenn $f'(x) < 0$ für alle $x \in \overset{\circ}{I}$.

Beweis. Sei ohne Einschränkung $I = [a, b]$ mit $a < b$.
a) Sei $d \in (a, b]$. Nach dem Mittelwertsatz (2.5) existiert ein $x_0 \in (a, d) \subset (a, b)$ mit

$$0 = f'(x_0) = \frac{f(d) - f(a)}{d - a}, \quad \text{also} \quad f(d) = f(a).$$

Demnach ist f konstant. Umgekehrt hat ein konstantes f die Ableitung 0.
b) Man wende a) auf $f(x) - g(x)$ an.
c), e) Wenn f monoton wächst, so gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \geqslant 0 \quad \text{für alle} \quad x \in I,\ x \neq x_0,$$

also auch $f'(x_0) \geqslant 0$.

Nun gelte andererseits $f'(x) \geqslant 0$ [bzw. $f'(x) > 0$] für alle $x \in \overset{\circ}{I}$. Sind $\alpha, \beta \in [a, b]$ mit $\alpha < \beta$, so existiert nach (2.5) ein $x_0 \in (\alpha, \beta)$ mit

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} = f'(x_0) \geqslant 0 \quad [\text{bzw. } > 0].$$

Es folgt $f(\beta) \geqslant f(\alpha)$ [bzw. $f(\beta) > f(\alpha)$], so dass f [streng] monoton wachsend ist.
d), f) Man wende c), e) auf $-f$ an. □

Als erste Folgerungen notieren wir das

(2.9) Korollar. *Es gilt*

a)
$$\arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \quad \textit{für } |x| < 1,$$

b)
$$\ln(1 + x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} \quad \textit{für } |x| < 1.$$

Beweis. Beide Reihen haben die geometrische Reihe als Majorante und konvergieren daher für $|x| < 1$ absolut.
a) Bezeichnet man die Reihe mit $f(x)$, so folgt aus (1.10) und (1.16) sowie IV(1.5)

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

für $|x| < 1$. Nach (2.8) existiert ein $c \in \mathbb{R}$ mit

$$f(x) = (\arctan x) + c \quad \text{für } |x| < 1.$$

Aus $f(0) = \arctan 0 = 0$ folgt $c = 0$.
b) Bezeichnet man die Reihe mit $g(x)$, so folgt aus (1.10) und (1.16) sowie IV(1.5)

$$g'(x) = \sum_{n=1}^{\infty} (-1)^{n-1} x^{n-1} = \frac{1}{1+x} = \ln'(1+x)$$

für $|x| < 1$. Nach (2.8) existiert ein $c \in \mathbb{R}$ mit

$$g(x) = \ln(1+x) + c.$$

Aus $g(0) = \ln 1 = 0$ folgt $c = 0$. $\square$

Einen neuen Begriff erhalten wir aus der

(2.10) Definition. Gegeben seien Funktionen $F : D \to \mathbb{R}$ und $f : D \to \mathbb{R}$. Man nennt F eine *Stammfunktion von* f, wenn F auf D differenzierbar ist mit $F' = f$.

Aus (2.8) b) folgt direkt das

(2.11) Korollar. *Sei I ein Intervall. Gegeben seien Funktionen $f, F, G : I \to \mathbb{R}$, so dass F eine Stammfunktion von f ist. Dann sind äquivalent:*

(i) G ist ebenfalls eine Stammfunktion von f.

(ii) Es existiert ein $c \in \mathbb{R}$ mit der Eigenschaft

$$F(x) = G(x) + c \quad \text{für alle } \; x \in I.$$

Als weitere Anwendung lösen wir eine einfache Differentialgleichung.

(2.12) Satz. *Für eine Funktion $f : I \to \mathbb{R}$, I Intervall und $\alpha \in \mathbb{R}$ sind äquivalent:*
(i) f ist differenzierbar mit $f' = \alpha f$.
(ii) Es gibt ein $c \in \mathbb{R}$ mit der Eigenschaft $f(x) = ce^{\alpha x}$ für alle $x \in I$.

Beweis. "(ii) $\Rightarrow$ (i)" Aus $f(x) = ce^{\alpha x}$ folgt $f'(x) = \alpha c e^{\alpha x} = \alpha f(x)$ für alle $x \in I$.
"(i) $\Rightarrow$ (ii)" Wir betrachten $g : I \to \mathbb{R}$, $x \mapsto e^{-\alpha x} f(x)$. Dann ist g differenzierbar mit

$$g'(x) = -\alpha e^{-\alpha x} f(x) + e^{-\alpha x} f'(x) = 0 \quad \text{für alle} \quad x \in I.$$

Nach (2.8) ist $g(x)$ konstant c. Also hat man

$$g(x) = c, \quad \text{d.h. } f(x) = ce^{\alpha x} \quad \text{für alle } x \in I. \qquad \square$$

Wir hatten bereits $\lim\limits_{x\to 0} \frac{\sin x}{x} = 1$ gezeigt. Jetzt soll eine einfache Methode angegeben werden, wie man solche Grenzwerte berechnen kann.

(2.13) Satz. *(Regel von Guillaume de* L'HOSPITAL (1661-1704)*)*
Seien $a, b \in \mathbb{R}, a < b$ oder $a = -\infty$. Gegeben seien zwei differenzierbare Funktionen $f : (a,b) \to \mathbb{R}$ und $g : (a,b) \to \mathbb{R}$ mit $g'(x) \neq 0$ für alle $x \in (a,b)$. Gilt

$$\lim_{x\downarrow a} f(x) = \lim_{x\downarrow a} g(x) = 0 \quad \textit{oder} \quad \lim_{x\downarrow a} f(x),\ \lim_{x\downarrow a} g(x) \in \{\infty, -\infty\}$$

und existiert

$$\lim_{x\downarrow a} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{\pm\infty\},$$

so gilt

$$\lim_{x\downarrow a} \frac{f(x)}{g(x)} = L.$$

Eine entsprechende Aussage gilt für linksseitige Limiten.

Beweis. Für $a < \alpha < \beta < b$ gilt nach dem Mittelwertsatz

$$\frac{g(\beta) - g(\alpha)}{\beta - \alpha} = g'(x_0) \neq 0 \quad \text{für ein} \ \ x_0 \in (a,b),$$

also insbesondere $g(\alpha) \neq g(\beta)$. Daher ist g injektiv.
1. Fall: $a \in \mathbb{R}$, $A = 0$.
Wir setzen $f(a) = g(a) = 0$. Dann sind $f, g : [a, x] \to \mathbb{R}$ stetig mit $g(x) \neq 0$ für jedes $x \in (a,b)$. Nach (2.7) existiert ein $\overline{x} \in (a,x)$ mit

$$\frac{f(x)}{g(x)} = \frac{f'(\overline{x})}{g'(\overline{x})} \underset{x\downarrow a}{\longrightarrow} L,$$

denn $x \downarrow a$ impliziert $\overline{x} \downarrow a$.
2. Fall: $a = -\infty$, $A = 0$.
Sei ohne Einschränkung $b < 0$. Wir betrachten dann $\overline{f} : \left(0, -\frac{1}{b}\right) \to \mathbb{R}$, $x \mapsto f\left(-\frac{1}{x}\right)$, und $\overline{g} : \left(0, -\frac{1}{b}\right) \to \mathbb{R}$, $x \mapsto g\left(-\frac{1}{x}\right)$. Aus dem 1. Fall folgt

$$\begin{aligned} \lim_{x\downarrow -\infty} \frac{f(x)}{g(x)} &= \lim_{x\downarrow 0} \frac{\overline{f}(x)}{\overline{g}(x)} = \lim_{x\downarrow 0} \frac{\overline{f}'(x)}{\overline{g}'(x)} = \lim_{x\downarrow 0} \frac{f'\left(-\frac{1}{x}\right) \cdot \frac{1}{x^2}}{g'\left(-\frac{1}{x}\right) \cdot \frac{1}{x^2}} \\ &= \lim_{x\downarrow 0} \frac{f'\left(-\frac{1}{x}\right)}{g'\left(-\frac{1}{x}\right)} = \lim_{x\downarrow -\infty} \frac{f'(x)}{g'(x)} = L. \end{aligned}$$

3. Fall: $a \in \mathbb{R},\ A = \pm\infty,\ L \in \mathbb{R}$.
Sei ohne Einschränkung $A = \infty$. Weil g injektiv und stetig ist, ist g nach V(3.10) streng monoton und wegen $\lim\limits_{x \downarrow a} g(x) = \infty$ ist g dann streng monoton fallend. Zu jedem $\varepsilon > 0$ existiert dann nach Definition des Funktionenlimes ein $\delta > 0$ mit

$$g(t) > 0,\ \Big|\frac{f'(t)}{g'(t)} - L\Big| < \varepsilon \quad \text{für alle } a < t < a + \delta.$$

Für alle $a < x < y < a + \delta$ existiert nach dem verallgemeinerten Mittelwertsatz (2.7) ein $\xi \in (x, y)$ mit

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(y) - f(x)}{g(y) - g(x)}, \quad \text{also} \quad \frac{f(y) - f(x)}{g(y) - g(x)} \in (L - \varepsilon, L + \varepsilon).$$

Nun gilt $\frac{g(y)-g(x)}{-g(x)} > 0$, also

$$(L - \varepsilon) \cdot \frac{g(y) - g(x)}{-g(x)} < \frac{f(y) - f(x)}{-g(x)} < (L + \varepsilon) \cdot \frac{g(y) - g(x)}{-g(x)},$$
$$(L - \varepsilon) \cdot \left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} < \frac{f(x)}{g(x)} < (L + \varepsilon) \cdot \left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)}.$$

Sei nun Œ $\varepsilon < 1$, y fest und $\delta > 0$ so, dass

$$\left|\frac{g(y)}{g(x)}\right| < \varepsilon \quad \text{sowie} \quad \left|\frac{f(y)}{f(x)}\right| < \varepsilon$$

für alle $x \in (a, a + \delta)$. Dann ist $1 - \frac{g(y)}{g(x)} > 1 - \varepsilon > 0$ und damit

$$(L - \varepsilon)(1 - \varepsilon) - \varepsilon < \frac{f(x)}{g(x)} < (L + \varepsilon)(1 + \varepsilon) + \varepsilon.$$

Für $\varepsilon \to 0$ konvergieren die linke und die rechte Seite der Abschätzungskette gegen 0. Also $\lim\limits_{x \downarrow a} \frac{f(x)}{g(x)} = L$.

4. Fall: $a \in \mathbb{R}$, $A = \pm\infty$, $L = \pm\infty$.
Die Behauptung wird analog zum 3. Fall gezeigt.
5. Fall: $a = -\infty$, $A = \pm\infty$.
Die Behauptung wird wie im 2. Fall auf den 3. und 4. Fall zurückgeführt.
Für linksseitige Limiten verwende man

$$\lim_{x \uparrow x_0} F(x) = \lim_{x \downarrow -x_0} F(-x).$$

□

Als einfache Anwendungen erhalten wir aus (2.13)

(i) $$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = \cos 0 = 1,$$

(iii) $$\lim_{x \to 0} \frac{e^x - 1}{\sin x} = \lim_{x \to 0} \frac{e^x}{\cos x} = \frac{e^0}{\cos 0} = 1,$$

(ii) $$\lim_{x \to 1} \frac{x^\alpha - 1}{x - 1} = \lim_{x \to 1} \frac{\alpha x^{\alpha - 1}}{1} = \alpha, \quad \alpha \in \mathbb{R},$$

(iv) $$\lim_{x \to 0} \frac{(\cos x) - 1}{x \sin x} = \lim_{x \to 0} \frac{-\sin x}{\sin x + x \cos x} = \lim_{x \to 0} \frac{-\cos x}{2\cos x - x \sin x} = -\frac{1}{2},$$

da man (2.13) für beide einseitigen Limiten verwenden kann. Man sollte aber den Grenzwert (i) nicht mit der Regel von L'HOSPITAL berechnen, da es sich hier um Differenzenquotienten handelt, die natürlich gegen die Ableitung der Funktion an der betreffenden Stelle konvergieren.

(2.14) Korollar. *Es gilt* $\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = e^a$ *für jedes* $a \in \mathbb{R}$, *insbesondere*

$$\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a.$$

Beweis. Es gilt $\left(1+\frac{a}{x}\right)^x = e^{F(x)}$ mit

$$F(x) = x\ln\left(1+\frac{a}{x}\right) = \frac{\ln\left(1+\frac{a}{x}\right)}{\frac{1}{x}}.$$

Wegen $\lim_{x\to\infty}\ln\left(1+\frac{a}{x}\right) = \ln 1 = 0 = \lim_{x\to\infty}\frac{1}{x}$ kann man (2.13) anwenden und erhält

$$\lim_{x\to\infty}F(x) = \lim_{x\to\infty}\frac{\frac{1}{1+a/x}\cdot\frac{-a}{x^2}}{\frac{-1}{x^2}} = \lim_{x\to\infty}\frac{a}{1+\frac{a}{x}} = a.$$

Weil die Exponentialfunktion stetig ist, folgt

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = \exp\left(\lim_{x\to\infty}F(x)\right) = e^a. \qquad \square$$

§3. TAYLOR-Reihen

Die Tangente ist eine lineare Approximation eines Graphen in einem Punkt. Es ist zu erwarten, dass man eine bessere Approximation erhält, wenn man Polynome betrachtet, d. h. höhere Ableitungen einbezieht.

(3.1) Definition. Sei $D \subset \mathbb{R}$ eine Menge ohne isolierte Punkte und $f : D \to \mathbb{R}$ eine Funktion. Ist f auf D differenzierbar (für Randpunkte ist einseitige Differenzierbarkeit gemeint), so ist die Ableitung $f' : D \to \mathbb{R}$ definiert. Ist f' ebenfalls differenzierbar, so nennt man die Ableitung von f' die *zweite Ableitung* von f und schreibt

$$(f')'(x_0) = f''(x_0) = f^{(2)}(x_0) = \frac{d^2 f}{dx^2}(x_0),\ x_0 \in D.$$

Rekursiv definiert man die *n-te Ableitung von f* für $n \in \mathbb{N}_0$ durch

$$f^{(0)}(x_0) := f(x_0),\ f^{(1)}(x_0) := f'(x_0),\ f^{(n+1)}(x_0) = (f^{(n)})'(x_0),$$

sofern die Ableitungen existieren. Man schreibt auch

$$f^{(n)}(x_0) = \frac{d^n f}{dx^n}(x_0).$$

Wohl bekannt sind die

(3.2) Beispiele. a) Sei $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^n$, für festes $n \in \mathbb{N}$. Dann gilt $f'(x) = nx^{n-1}$, $f''(x) = n(n-1)x^{n-2}$, $f^{(k)}(x) = n(n-1) \cdot \ldots \cdot (n-k+1)x^{n-k}$ für $1 \leqslant k < n$ und $f^{(n)}(x) = n!$ sowie $f^{(k)}(x) = 0$ für alle $k > n$.
b) Sei $f : \mathbb{R}_+^* \to \mathbb{R}$, $x \mapsto \ln x$. Dann gilt $f'(x) = \frac{1}{x} = x^{-1}$, $f''(x) = -x^{-2}$, $f'''(x) = 2x^{-3}$, $f^{(n+1)}(x) = (-1)^n n! x^{-(n+1)}$ für alle $x > 0$.
c) $\sin' x = \cos x$, $\sin'' x = -\sin x$, $\sin''' x = -\cos x$, $\sin^{(4)} x = \sin x$.
d) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \operatorname{sgn} x \cdot x^2$, ist differenzierbar mit $f'(x) = 2 \operatorname{sgn} x \cdot x = 2|x|$. Also ist f' zwar stetig, aber im Nullpunkt nicht differenzierbar.
e) Die Funktion

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

ist differenzierbar mit

$$f' : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} 2x \sin(1/x) - \cos(1/x), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0. \end{cases}$$

Die Ableitung f' ist in 0 nicht einmal stetig, da

$$\lim_{x \to 0} 2x \sin(1/x) = 0 \quad \text{und} \quad \lim_{x \to 0} \cos(1/x) \quad \text{nicht existiert.}$$

Zur einfacheren Formulierung von Aussagen benötigen wir die

(3.3) Definition. Für $D \subset \mathbb{R}$ ohne isolierte Punkte und $n \in \mathbb{N}_0$ bestehe $C^{(n)}(D)$ aus allen Funktionen auf D, die *n-mal stetig differenzierbar* sind, d. h.

$$C^{(n)}(D) := \{ f : D \to \mathbb{R} \mid f \ n\text{-mal differenzierbar}, \ f^{(n)} \text{ stetig} \}.$$

Schließlich sei

$$C^{\infty}(D) := \bigcap_{n=0}^{\infty} C^{(n)}(D)$$

die Menge der unendlich oft (d. h. beliebig oft) stetig differenzierbaren Funktionen auf D.

Algebraisch gesehen folgt aus den Differenzierbarkeitsregeln (1.8), dass für jedes $n \in \mathbb{N}_0 \cup \{\infty\}$ die Menge $C^{(n)}(D)$ eine $\mathbb{R}$-Algebra ist. Diesen Sachverhalt, den man leicht per Induktion beweisen kann, formulieren wir in dem

(3.4) Lemma. *Seien $D \subset \mathbb{R}$ ohne isolierte Punkte, $n \in \mathbb{N}_0$, $f, g \in C^{(n)}(D)$ und $\alpha \in \mathbb{R}$. Dann gilt*

(i) $\alpha f \in C^{(n)}(D)$ *mit* $(\alpha f)^{(n)} = \alpha f^{(n)}$.

(ii) $f + g \in C^{(n)}(D)$ *mit* $(f+g)^{(n)} = f^{(n)} + g^{(n)}$.

(iii) $f \cdot g \in C^{(n)}(D)$ *mit* $(f \cdot g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} \cdot g^{(n-k)}$.

(iv) Gilt $g(x) \neq 0$ *für alle* $x \in D$, *so folgt* $\frac{f}{g} \in C^{(n)}(D)$.

Unter geeigneten Differenzierbarkeitsvoraussetzungen geben wir nun eine Approximation von f durch ein Polynom an.

(3.5) Satz. *(TAYLORsche Formel mit dem Restglied von LAGRANGE)*
Seien I ein Intervall, $n \in \mathbb{N}$, $f \in C^{(n)}(I)$ und $a \in I$. Dann gibt es eine Funktion $R_{n,a} : I \to \mathbb{R}$, so dass

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_{n,a}(x), \quad x \in I.$$

Zu jedem $x \in I, x \neq a$, existiert ein $\xi = \xi_{a,x}$ echt zwischen a und x (d. h. $\xi \in (a,x)$ bzw. $\xi \in (x,a)$), so dass

$$R_{n,a}(x) = \frac{f^{(n)}(\xi)}{n!}(x-a)^n.$$

Man nennt $\sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k$ das $(n-1)$*te* TAYLOR-*Polynom* von f zum Entwicklungspunkt a. Für ξ gibt es eine Darstellung $\xi = a + \vartheta \cdot (x-a)$ mit $0 < \vartheta < 1$.

Beweis. Sei

$$R_{n,a}(x) := f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Dann gilt trivialerweise die obige Gleichung. Sei nun $x \in I, x \neq a$, und ohne Einschränkung $x > a$. Dann existiert ein $M \in \mathbb{R}$ mit

$$(*) \qquad R_{n,a}(x) = \frac{M}{n!}(x-a)^n.$$

Wir wollen den Satz von ROLLE (2.4) anwenden und betrachten

$$F : [a,x] \to \mathbb{R}, \; t \mapsto f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(t)}{k!}(x-t)^k - \frac{M}{n!}(x-t)^n.$$

Dann gilt $F(a) = F(x) = 0$ nach $(*)$. Wegen $f \in C^{(n)}(I)$ ist F stetig und auf (a,x) differenzierbar. Nach (2.4) existiert ein $\xi \in (a,x)$ mit

$$\begin{aligned} 0 = F'(\xi) &= -\sum_{k=0}^{n-1} \frac{f^{(k+1)}(\xi)}{k!}(x-\xi)^k + \sum_{k=1}^{n-1} \frac{f^{(k)}(\xi)}{(k-1)!}(x-\xi)^{k-1} + \frac{M}{(n-1)!}(x-\xi)^{n-1} \\ &= -\frac{f^{(n)}(\xi)}{(n-1)!}(x-\xi)^{n-1} + \frac{M}{(n-1)!}(x-\xi)^{n-1}. \end{aligned}$$

Aus $x \neq \xi$ folgt $M = f^{(n)}(\xi)$, also die Behauptung. □

Dieser Satz stammt von Brook TAYLOR (1685-1731). Im Fall $n = 1$ hat man

$$f(x) = f(a) + f'(\xi)(x-a),$$

also den Mittelwertsatz der Differentialrechnung.

(3.6) Beispiele. a) Sei $I = \mathbb{R}$, $f(x) = e^x, a = 0$. Dann gilt $f^{(n)}(x) = e^x, f^{(n)}(0) = 1$

$$e^x = \sum_{k=0}^{n-1} \frac{1}{k!} x^k + e^{\xi} \frac{x^n}{n!}$$

mit einem ξ zwischen 0 und x. Man beachte $e^{\xi} \frac{x^n}{n!} \xrightarrow[n\to\infty]{} 0$. Der Limes $n \to \infty$ ergibt daher

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \ x \in \mathbb{R}.$$

Man kann die TAYLORsche Formel aber für Approximationen der konkreten Funktionswerte nutzen. Wegen $e^x \leqslant 1$ für $x \leqslant 0$ hat man z. B.

$$e^{-0,1} = \sum_{k=0}^{4} \frac{1}{k!} (-0,1)^k + R = 0,904.837.5 + R$$

mit

$$R = \frac{e^{\xi}}{5!} (-0,1)^5, \quad R \leqslant 0, \quad |R| \leqslant \frac{1}{120} \cdot 10^{-5} \leqslant 10^{-7}.$$

Damit haben wir die ersten 6 Nachkommastellen des exakten Wertes

$$e^{-0,1} = 0,904.837.418.03\ldots$$

bestimmt. Wegen $e \leqslant 3$ hat man $e^{1/2} \leqslant \sqrt{3} \leqslant 2$, also

$$\sqrt{e} = e^{1/2} = \sum_{k=0}^{5} \frac{1}{k!} \left(\frac{1}{2}\right)^k + R = 1,648.697\ldots + R$$

mit

$$0 \leqslant R = \frac{e^{\xi}}{6!} \left(\frac{1}{2}\right)^6 \leqslant \frac{1}{6! \cdot 2^5} \leqslant 10^{-4},$$

wobei

$$\sqrt{e} = 1,648.772\ldots .$$

b) Sei $I = \mathbb{R}_+^*$, $f(x) = \ln x$, $a = 1$. Dann gilt $f^{(k)}(x) = (-1)^{k-1}(k-1)! \cdot x^{-k}$, $k \in \mathbb{N}$, also

$$\ln x = \sum_{k=1}^{n-1} (-1)^{k-1} \frac{(x-1)^k}{k} + \frac{(-1)^{n-1}(x-1)^n}{n} \xi^{-n}$$

für ein ξ zwischen 1 und x. Wegen $\left|\frac{x-1}{\xi}\right| \leqslant 1$ für $\frac{1}{2} \leqslant x \leqslant 2$ erhält man

$$\ln x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(x-1)^k}{k}, \quad \frac{1}{2} \leqslant x \leqslant 2, \qquad \ln 2 = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}.$$

Die in (2.9) b) gefundene Identität gilt damit auch für $x = 1$.

Als Anwendung von (3.5) betrachten wir Extrema.

(3.7) Satz. *(Kurvendiskussion)*
Sei $n \in \mathbb{N}$, $n \geqslant 2$, $I \subset \mathbb{R}$ ein Intervall, x_0 ein innerer Punkt von I und $f \in C^{(n)}(I)$. Es gelte

$$f^{(k)}(x_0) = 0,\ 1 \leqslant k \leqslant n-1,\ f^{(n)}(x_0) \neq 0.$$

a) Ist n gerade und $f^{(n)}(x_0) > 0$, so hat f ein relatives Minimum an der Stelle x_0.
b) Ist n gerade und $f^{(n)}(x_0) < 0$, so hat f ein relatives Maximum an der Stelle x_0.
c) Ist n ungerade, so hat f kein relatives Extremum in x_0.

Beweis. Man wendet (3.5) an auf $a = x_0$ und erhält

$$f(x) = f(x_0) + \frac{(x-x_0)^n}{n!} f^{(n)}(\xi)$$

für ein ξ echt zwischen x und x_0. Weil $f^{(n)}$ stetig ist, existiert ein $\delta > 0$, so dass $f^{(n)}(x) > 0$ bzw. $f^{(n)}(x) < 0$ für alle $|x - x_0| < \delta$.
a) Es gilt $f(x) - f(x_0) > 0$ für alle $0 < |x - x_0| < \delta$.
b) Es gilt $f(x) - f(x_0) < 0$ für alle $0 < |x - x_0| < \delta$.
c) Gilt $f^{(n)}(x_0) > 0$, so hat man

$$\begin{array}{lll} f(x) - f(x_0) > 0 & \text{für alle} & x_0 < x < x_0 + \delta \quad \text{und} \\ f(x) - f(x_0) < 0 & \text{für alle} & x_0 - \delta < x < x_0. \end{array}$$

Gilt $f^{(n)}(x_0) < 0$, so schließt man durch Betrachtung von $-f$ analog. □

Wir diskutieren ein konkretes

(3.8) Beispiel. Wir betrachten $f_n : \mathbb{R}_+^* \to \mathbb{R}$, $x \mapsto x^n \ln x$, $n \in \mathbb{N}$. Für $x > 0$ hat man $f'(x) = nx^{n-1}\ln x + x^{n-1} = 0$ genau dann, wenn $\ln x = -\frac{1}{n}$, also $x = e^{-1/n} = \frac{1}{\sqrt[n]{e}}$. Es gilt

$$f_n''(x) = x^{n-2}(n(n-1)\ln x + 2n - 1), \qquad f_n''(e^{-1/n}) = n\, e^{-(n-2)/n} > 0.$$

Also hat f_n an der Stelle $\frac{1}{\sqrt[n]{e}}$ ein lokales Minimum.

(3.9) Bemerkung. Das Vorgehen nach Satz (3.7) ist nicht immer optimal. Oft kommt man mit der Beobachtung schneller zum Ziel, dass das Vorzeichenverhalten von f' die Art der kritischen Punkte bestimmt:

– Wechselt das Vorzeichen von f' bei x_0 von " $+$ " nach " $-$ " (gibt es also ein $\rho > 0$, so dass $f'(x) > 0$ für alle $x \in (x_0 - \rho, x_0)$ und $f'(x) < 0$ für alle $x \in (x_0, x_0 + \rho)$, so liegt in x_0 ein lokales Maximum vor.

– Wechselt das Vorzeichen von f' bei x_0 von " $-$ " nach " $+$ ", so liegt in x_0 ein lokales Minimum vor.

– Wechselt das Vorzeichen von f' bei x_0 nicht, so liegt kein lokales Extremum vor.

(3.10) Beispiel. $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto \sin(x^8)$. Dann wechselt

$$f'(x) = \cos(x^8) \cdot 8x^7$$

bei 0 sein Vorzeichen von “ $-$ ” nach “ $+$ ” (beachte $\cos 0 > 0$), also liegt ein lokales Minimum vor. (Dies geht schneller als acht Ableitungen auszurechnen.)

Als direkte Anwendung von (3.5) erhält man in dem Fall, dass das Restglied 0 ist, das

(3.11) Korollar. *Seien $n \in \mathbb{N}$, $I \subset \mathbb{R}$ ein Intervall und $f \in C^{(n)}(I)$. Dann sind äquivalent:*
(i) $f^{(n)}(x) = 0$ für alle $x \in I$.
(ii) f ist eine Polynomfunktion von einem Grad $< n$ oder $f \equiv 0$.

Nun untersuchen wir TAYLOR-Reihen.

(3.12) Definition. Sei $I \subset \mathbb{R}$ ein Intervall, $a \in I$ und $f \in C^{\infty}(I)$. Dann heißt

$$T_{f,a}(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

die TAYLOR-*Reihe von f mit Entwicklungspunkt a*.

Bevor wir zu Beispielen kommen, notieren wir die

(3.13) Bemerkungen. a) Der Konvergenzradius der TAYLOR-Reihe ist nicht notwendig > 0.
b) Falls die TAYLOR-Reihe konvergiert, konvergiert sie - außer für den Entwicklungspunkt - nicht notwendig gegen $f(x)$ (vgl. (3.15) d)).
c) Die TAYLOR-Reihe konvergiert genau für diejenigen $x \in I$ gegen $f(x)$, für die das Restglied $R_{n,a}(x) \xrightarrow[n\to\infty]{} 0$ in (3.5).

Eine ganze Klasse von Beispielen behandeln wir in dem

(3.14) Satz. *Sei $f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n$ eine reelle Potenzreihe mit Entwicklungspunkt a und Konvergenzradius $R > 0$. Dann stimmt die* TAYLOR-*Reihe mit Entwicklungspunkt a mit f überein, d. h.*

$$T_{f,a}(x) = f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n \quad \textit{für} \quad |x-a| < R.$$

Beweis. Aus (1.10) folgt $f \in C^\infty(I)$, $I = (a-R, a+R)$. Es gilt $a_0 = f(a)$. Durch Induktion nach k zeigt man leicht

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdot\ldots\cdot(n-k+1)a_n(x-a)^{n-k}, \qquad f^{(k)}(a) = k!\cdot a_k,$$

wenn man (1.10) verwendet. Es folgt

$$\frac{f^{(k)}(a)}{k!} = a_k \quad \text{für alle} \quad k \in \mathbb{N}_0.$$

□

Nun diskutieren wir einige

(3.15) Beispiele. a) Wir kennen schon einige Potenzreihen

$$\begin{aligned}
\exp(x) &= T_{\exp,0}(x) = \sum_{n=0}^{\infty} \frac{1}{n!}x^n,\ x \in \mathbb{R},\\
\sin x &= T_{\sin,0}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1},\ x \in \mathbb{R},\\
\cos x &= T_{\cos,0}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}x^{2n},\ x \in \mathbb{R}.
\end{aligned}$$

b) Wir hatten für einige Umkehrfunktionen in (2.9) bereits Potenzreihenentwicklungen hergeleitet:

$$\begin{aligned}
\ln x &= T_{\ln,1}(x) = \sum_{n=1}^{\infty} (-1)^{n-1}\frac{(x-1)^n}{n},\ 0 < x < 2,\\
\arctan x &= T_{\arctan,0}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1},\ |x| < 1.
\end{aligned}$$

Obwohl arctan zu $C^\infty(\mathbb{R})$ gehört, konvergiert die TAYLOR-Reihe nur auf $[-1,1]$.
c) Wir betrachten die geometrische Reihe

$$f(x) = \frac{1}{1-x} = T_{f,0}(x) = \sum_{n=0}^{\infty} x^n, \quad |x| < 1.$$

Durch Differentiation folgt daraus

$$\frac{1}{(1-x)^2} = T_{f',0}(x) = \sum_{n=1}^{\infty} nx^{n-1}, \quad |x| < 1.$$

Eine Induktion nach k liefert

$$\begin{aligned}
\frac{k!}{(1-x)^{k+1}} &= T_{f^{(k)},0}(x) = \sum_{n=k}^{\infty} n(n-1)\cdot\ldots\cdot(n-k+1)x^{n-k},\\
\frac{1}{(1-x)^{k+1}} &= \sum_{n=k}^{\infty} \binom{n}{k} x^{n-k} = \sum_{j=0}^{\infty} \binom{j+k}{k} x^j,\ |x| < 1.
\end{aligned}$$

d) Wir betrachten

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} e^{-1/x^2}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

und zeigen $f \in C^\infty(\mathbb{R})$ mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}_0$. Die TAYLOR-Reihe mit Entwicklungspunkt 0 ist also identisch 0 und konvergiert sicher nicht gegen f. Eine Induktion nach n zeigt, dass es ein Polynom $p_n(x)$ vom Grad $3n$ gibt mit

$$f^{(n)}(x) = \begin{cases} p_n\left(\frac{1}{x}\right) e^{-1/x^2}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0. \end{cases}$$

Der Fall $n = 0$ ist klar mit $p_0(x) = 1$. Für $x \neq 0$ hat man

$$\begin{aligned} f^{(n+1)}(x) &= \frac{d}{dx}\left(p_n\left(\frac{1}{x}\right)e^{-1/x^2}\right) = \left(-p_n'\left(\frac{1}{x}\right)\frac{1}{x^2} + 2p_n\left(\frac{1}{x}\right)\frac{1}{x^3}\right) e^{-1/x^2} \\ &= p_{n+1}\left(\frac{1}{x}\right) e^{-1/x^2} \end{aligned}$$

mit

$$p_{n+1}(t) = -p_n'(t)t^2 + 2p_n(t)t^3, \quad \text{grad}\, p_{n+1} = \text{grad}\, p_n + 3.$$

Für $x_0 = 0$ hat man

$$\lim_{h\to 0} \frac{f^{(n)}(h) - f^{(n)}(0)}{h} = \lim_{h\to 0} p_n\left(\frac{1}{h}\right)\frac{1}{h}\, e^{-1/h^2} = \lim_{t\to\pm\infty} \frac{p_n(t)t}{e^{t^2}} = 0$$

nach V(4.1). Es folgt $f^{(n+1)}(0) = 0$.

Nun werden wir Potenzreihen "integrieren".

(3.16) Satz. *Sei $f(x) = \sum_{n=0}^\infty a_n(x-a)^n$ eine reelle Potenzreihe mit Entwicklungspunkt a und Konvergenzradius $R > 0$. Dann hat $f(x)$ auf $(a-R, a+R)$ eine Stammfunktion, nämlich*

$$F(x) = c + \sum_{n=0}^\infty \frac{a_n}{n+1}(x-a)^{n+1}, \quad x \in (a-R, a+R),\ c \in \mathbb{R}.$$

Beweis. Nach dem ABELschen Lemma IV(4.3) hat $F(x)$ ebenfalls den Konvergenzradius R. Es folgt $F'(x) = f(x)$ aus (1.10). □

Als weitere Anwendung erhalten wir den

(3.17) Identitätssatz. *Seien $f(x) = \sum_{n=0}^\infty a_n(x-x_0)^n$ und $g(x) = \sum_{n=0}^\infty b_n(x-x_0)^n$ Potenzreihen, die für $|x-x_0| < R$, $R > 0$, konvergieren. Gibt es eine Folge $(x_k)_{k\geqslant 1}$ in $U_R(x_0)\setminus\{x_0\}$ mit $\lim_{k\to\infty} x_k = x_0$ und $f(x_k) = g(x_k)$ für alle $k \in \mathbb{N}$, so folgt bereits*

$$a_n = b_n \quad \text{für alle} \quad n \in \mathbb{N} \quad \text{und} \quad f(x) = g(x) \quad \text{für alle} \quad x \in U_R(x_0).$$

Beweis. Wir zeigen $a_n = b_n$ durch Induktion nach n. Da $f(x)$ und $g(x)$ auf $U_R(x_0)$ stetig sind, folgt

$$a_0 = f(x_0) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} g(x_k) = g(x_0) = b_0.$$

Wir nehmen an, dass $a_0 = b_0, \dots, a_n = b_n$, $n \in \mathbb{N}_0$, bereits gezeigt wurde. Die beiden Potenzreihen

$$F(x) = \sum_{k=n+1}^{\infty} a_k (x-x_0)^{k-n-1}, \quad G(x) = \sum_{k=n+1}^{\infty} b_k (x-x_0)^{k-n-1}$$

konvergieren ebenfalls für $|x - x_0| < R$. Für $x \neq x_0$ gilt

$$\begin{aligned} F(x) &= \frac{1}{(x-x_0)^{n+1}} \left(f(x) - \sum_{k=0}^{n} a_k (x-x_0)^k \right), \\ G(x) &= \frac{1}{(x-x_0)^{n+1}} \left(g(x) - \sum_{k=0}^{n} b_k (x-x_0)^k \right). \end{aligned}$$

Aus $a_0 = b_0, \dots, a_n = b_n$ erhält man $F(x_k) = G(x_k)$ für alle $k \in \mathbb{N}$. Die Stetigkeit ergibt nun

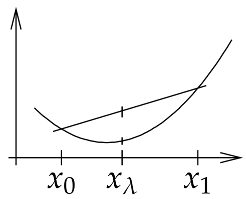
$$a_{n+1} = F(x_0) = \lim_{k\to\infty} F(x_k) = \lim_{k\to\infty} G(x_k) = G(x_0) = b_{n+1}.$$

Damit hat man $a_n = b_n$ für alle $n \in \mathbb{N}$ und somit $f = g$. □

Die Voraussetzungen des Identitätssatzes sind natürlich insbesondere erfüllt, wenn es ein $0 < \varepsilon < R$ gibt, so dass $f(x) = g(x)$ für alle $x \in U_\varepsilon(x_0)$ gilt.

§4. Konvexität

Mit der Konvexität führen wir einen weiteren recht anschaulichen Begriff ein. Er bedeutet, dass die Verbindungsstrecke zwischen je zwei Punkten des Graphen stets nicht unterhalb des Graphen verläuft. Das formalisieren wir in der



(4.1) Definition. Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt *konvex*, wenn für alle $x_0, x_1 \in I$ und $0 \leqslant \lambda \leqslant 1$ gilt

$$f(x_\lambda) \leqslant (1-\lambda) f(x_0) + \lambda f(x_1) \quad \text{für } x_\lambda := (1-\lambda)x_0 + \lambda x_1.$$

Man nennt f *konkav*, wenn $-f$ konvex ist.

Die Bedingung ist offenbar für $\lambda = 0$ und $\lambda = 1$ erfüllt. Wegen $x_\lambda = (1-\lambda')x_1 + \lambda' x_0$, $\lambda' = 1 - \lambda$, genügt es, $x_0 < x_1$ und $0 < \lambda < 1$ zu betrachten.

Zur Illustration betrachten wir $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. Für $0 < \lambda < 1$ gilt

$$\begin{aligned}&(1-\lambda)f(x_0) + \lambda f(x_1) - f(x_\lambda) = (1-\lambda)x_0^2 + \lambda x_1^2 - [(1-\lambda)x_0 + \lambda x_1]^2 \\ &\quad = \lambda(1-\lambda)[x_0^2 + x_1^2 - 2x_0x_1] = \lambda(1-\lambda)(x_0 - x_1)^2 \geqslant 0.\end{aligned}$$

Wir beweisen ein handliches Kriterium.

(4.2) Satz. *Sei I ein Intervall, $f : I \to \mathbb{R}$ stetig und auf dem offenen Intervall $\mathring{I}$ differenzierbar. f ist genau dann konvex, wenn f' auf $\mathring{I}$ monoton wachsend ist.*

Beweis. „$\Rightarrow$" Sei f konvex. Für $x_0, x_1 \in \mathring{I}$ mit $x_0 < x_1, x_\lambda := (1-\lambda)x_0 + \lambda x_1$ und $0 < \lambda < 1$ hat man $x_\lambda - x_0 = \lambda(x_1 - x_0)$, $x_1 - x_\lambda = (1-\lambda)(x_1 - x_0)$, also

$$(x_1 - x_0)[(1-\lambda)f(x_0) + \lambda f(x_1) - f(x_\lambda)] \geqslant 0.$$

Daraus erhält man die Ungleichungen

$$\frac{f(x_\lambda) - f(x_0)}{x_\lambda - x_0} \leqslant \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \frac{f(x_1) - f(x_\lambda)}{x_1 - x_\lambda}.$$

Führt man den Limes $\lambda \downarrow 0$ bzw. $\lambda \uparrow 1$ aus, so folgt aus $x_\lambda \downarrow x_0$ bzw. $x_\lambda \uparrow x_1$ und der Differenzierbarkeit von f

$$f'(x_0) \leqslant \frac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant f'(x_1).$$

Demnach ist f' monoton wachsend auf $\mathring{I}$.

„$\Leftarrow$" Sei f' monoton wachsend auf $\mathring{I}$. Seien $x_0, x_1 \in I, x_0 < x_1$. Für $0 < \lambda < 1$ erhält man aus dem Mittelwertsatz (2.5) $z_0 \in (x_0, x_\lambda)$ und $z_1 \in (x_\lambda, x_1)$ mit

$$\frac{f(x_\lambda) - f(x_0)}{x_\lambda - x_0} = f'(z_0), \qquad \frac{f(x_1) - f(x_\lambda)}{x_1 - x_\lambda} = f'(z_1).$$

Da f' monoton wachsend ist, folgt aus $z_0 < z_1$

$$\frac{f(x_\lambda) - f(x_0)}{x_\lambda - x_0} \leqslant \frac{f(x_1) - f(x_\lambda)}{x_1 - x_\lambda}$$

und daher

$$\begin{aligned} 0 &\leqslant (x_\lambda - x_0)(f(x_1) - f(x_\lambda)) - (x_1 - x_\lambda)(f(x_\lambda) - f(x_0)) \\ &= (x_1 - x_0)[(1-\lambda)f(x_0) + \lambda f(x_1) - f(x_\lambda)]. \end{aligned}$$

Das bedeutet $(1-\lambda)f(x_0) + \lambda f(x_1) \geqslant f(x_\lambda)$. □

Mit (2.7) erhält man sofort das

(4.3) Korollar. *Sei I ein Intervall, $f : I \to \mathbb{R}$ stetig und auf dem offenen Intervall $\mathring{I}$ zweimal differenzierbar. Dann sind äquivalent:*
(i) f ist konvex.
(ii) $f''(x) \geqslant 0$ für alle $x \in \mathring{I}$.

Damit sieht man, dass alle Polynomfunktionen x^n auf $\mathbb{R}_+$ und für gerades $n \in \mathbb{N}$ auf $\mathbb{R}$ konvex sind.

Eine weitere Anwendung ist das

(4.4) Lemma. *Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für alle $x, y \in \mathbb{R}_+$*

$$x^{1/p} \cdot y^{1/q} \leqslant \frac{x}{p} + \frac{y}{q}.$$

Im Fall $p = q = 2$ ist das die Ungleichung zwischen dem geometrischen und arithmetischen Mittel II(1.8).

Beweis. Es genügt, $x, y \in \mathbb{R}_+^*$ zu betrachten, da sonst die linke Seite 0 ist. Wir betrachten

$$f : \mathbb{R}_+^* \to \mathbb{R},\ x \mapsto -\ln x, \quad f'(x) = \frac{-1}{x},\ f''(x) = \frac{1}{x^2} > 0 \text{ für alle } x > 0.$$

Also ist f konvex nach (4.3). Es folgt

$$-\ln\left(\frac{1}{p}x + \frac{1}{q}y\right) \leqslant -\frac{1}{p}\ln x - \frac{1}{q}\ln y.$$

Daraus erhält man

$$\frac{x}{p} + \frac{y}{q} = \exp\left(\ln\left(\frac{1}{p}x + \frac{1}{q}y\right)\right) \geqslant \exp\left(\frac{1}{p}\ln x + \frac{1}{q}\ln y\right) = x^{1/p} \cdot y^{1/q}.$$ □

Dies nutzen wir für die

(4.5) Definition. Sei $p \in \mathbb{R},\ p \geqslant 1$ und $z = (z_1, \dots, z_n)^{tr} \in \mathbb{C}^n$. Dann definiert man die *p-Norm von z* durch

$$\|z\|_p := \left(\sum_{j=1}^{n} |z_j|^p\right)^{1/p}.$$

Im Fall $p = 2$ erhält man die gewöhnliche EUKLIDische Norm. Offenbar gilt $\|0\|_p = 0$ und $\|z\|_p > 0$ für $z \neq 0$ sowie $\|\lambda z\|_p = |\lambda| \cdot \|z\|_p$ für alle $\lambda \in \mathbb{C}$.

(4.6) Satz. (HÖLDERsche Ungleichung)
Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für alle Vektoren $z = (z_1, \ldots, z_n)^{tr}, w = (w_1, \ldots, w_n)^{tr} \in \mathbb{C}^n$

$$\sum_{j=1}^{n} |z_j w_j| \leqslant \|z\|_p \cdot \|w\|_q.$$

Im Fall $p = q = 2$ ist das genau die CAUCHY-SCHWARZsche Ungleichung.

Beweis. Wir dürfen $z \neq 0$ und $w \neq 0$ annehmen, da sonst beide Seiten 0 sind. Mit

$$\alpha_j := \frac{|z_j|^p}{\|z\|_p^p} \in \mathbb{R}_+, \quad \beta_j := \frac{|w_j|^q}{\|w\|_q^q} \in \mathbb{R}_+$$

erhält man $\sum_{j=1}^n \alpha_j = \sum_{j=1}^n \beta_j = 1$ aus (4.5). Aus (4.4) ergibt sich dann

$$\frac{|z_j w_j|}{\|z\|_p \cdot \|w\|_q} = \alpha_j^{1/p} \cdot \beta_j^{1/q} \leqslant \frac{\alpha_j}{p} + \frac{\beta_j}{q},$$

also

$$\frac{1}{\|z\|_p \cdot \|w\|_q} \sum_{j=1}^{n} |z_j w_j| \leqslant \frac{1}{p} \sum_{j=1}^{n} \alpha_j + \frac{1}{q} \sum_{j=1}^{n} \beta_j = \frac{1}{p} + \frac{1}{q} = 1.$$

Das ist die Behauptung. □

Als Anwendung davon erhalten wir den

(4.7) Satz. (MINKOWSKIsche Ungleichung)
Ist $p \geqslant 1$, so gilt für alle $z, w \in \mathbb{C}^n$ die Dreiecksungleichung

$$\|z + w\|_p \leqslant \|z\|_p + \|w\|_p.$$

Beweis. Im Fall $z + w = 0$ ist die Behauptung klar. Sei also $z + w \neq 0$. Für $p = 1$ folgt die Behauptung aus der Dreiecksungleichung für den Betrag in $\mathbb{C}$, also

$$\|z + w\|_1 = \sum_{j=1}^{n} |z_j + w_j| \leqslant \sum_{j=1}^{n} (|z_j| + |w_j|) = \|z\|_1 + \|w\|_1.$$

Ist $p > 1$, so wählt man $q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Sei $x \in \mathbb{R}^n$ mit $x_j := |z_j + w_j|^{p-1}, j = 1, \ldots, n$. Dann gilt $\frac{p}{q} = p - 1, q(p-1) = p$ und

$$x_j^q = |z_j + w_j|^{q(p-1)} = |z_j + w_j|^p, \quad \|x\|_q = \|z + w\|_p^{p/q}.$$

Aus der HÖLDERschen Ungleichung (4.6) ergibt sich

$$\sum_{j=1}^{n} |z_j + w_j|\, |x_j| \leqslant \sum_{j=1}^{n} |z_j x_j| + \sum_{j=1}^{n} |w_j x_j| \leqslant \|z\|_p \|x\|_q + \|w\|_p \|x\|_q.$$

Dann liefert die Definition von x_j

$$\|z+w\|_p^p = \sum_{j=1}^{n} |z_j + w_j|^p \leqslant (\|z\|_p + \|w\|_p) \cdot \|x\|_q = (\|z\|_p + \|w\|_p) \cdot \|z+w\|_p^{p/q}.$$

Wegen $p - \frac{p}{q} = 1$ hat man

$$\|z+w\|_p \leqslant \|z\|_p + \|w\|_p.$$

□

Wir erweitern die Aussagen auf Reihen

(4.8) Korollar. *Seien $a_k, b_k \in \mathbb{C}$, $k \in \mathbb{N}_0$.*

a) Wenn die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergiert, dann auch die Reihe $\sum_{k=0}^{\infty} |a_k|^p$ für jedes $p \geqslant 1$.

b) Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Wenn die Reihen $\sum_{k=0}^{\infty} |a_k|^p$ und $\sum_{k=0}^{\infty} |b_k|^q$ konvergieren, dann auch $\sum_{k=0}^{\infty} |a_k b_k|$ und es gilt

$$\sum_{k=0}^{n} |a_k b_k| \leqslant \left(\sum_{k=0}^{\infty} |a_k|^p \right)^{1/p} \cdot \left(\sum_{k=0}^{\infty} |b_k|^q \right)^{1/q}.$$

c) Ist $p > 1$ und konvergieren $\sum_{k=0}^{\infty} |a_k|^p$ und $\sum_{k=0}^{\infty} |b_k|^p$, dann auch $\sum_{k=0}^{\infty} |a_k + b_k|^p$ und es gilt

$$\left(\sum_{k=0}^{\infty} |a_k + b_k|^p \right)^{1/p} \leqslant \left(\sum_{k=0}^{\infty} |a_k|^p \right)^{1/p} + \left(\sum_{k=0}^{\infty} |a_k|^q \right)^{1/q}.$$

Beweis. a) Wegen der Konvergenz der Reihe ist $(a_k)_{k \geqslant 0}$ eine Nullfolge. Also existiert ein k_0 mit $|a_k| \leqslant 1$, also

$$|a_k^p| \leqslant |a_k| \leqslant 1 \quad \text{für alle} \ \ k \geqslant k_0.$$

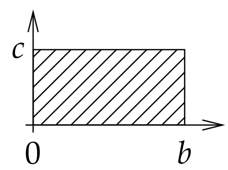
Dann konvergiert die Reihe nach dem Majorantenkriterium.
b),c) Man verwende (4.6) und (4.7) und führe den Limes $n \to \infty$ durch. Die linke Seite konvergiert dann jeweils nach dem Majorantenkriterium. □

Die Ungleichungen gehen zurück auf Otto HÖLDER (1859-1937) und Hermann MINKOWSKI (1864-1909).

Kapitel VIII.

Das RIEMANN-Integral

Die Fläche eines Rechtecks mit den Kantenlängen c und b ist $c \cdot b$. Betrachtet man die konstante Funktion

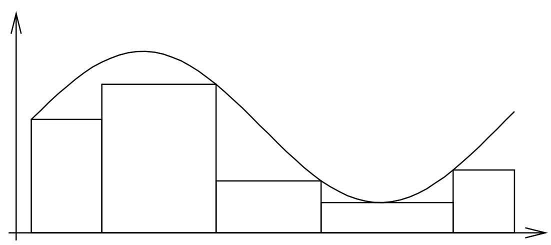


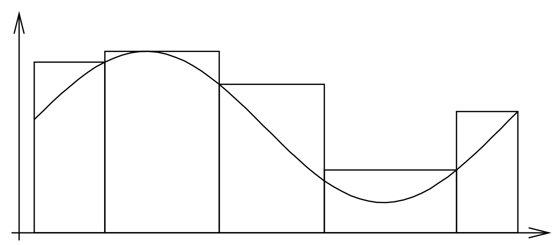
$$f : [0, b] \to \mathbb{R}, \quad x \mapsto c,$$

so ist die Fläche zwischen der x-Achse und dem Graphen von f gerade $c \cdot b$. Man schreibt dafür

$$\int_0^b f(x)\, dx = c \cdot b.$$

Das Ziel der Integralrechnung ist es, diese Fläche für eine möglichst große Klasse von Funktionen zu bestimmen. Dabei soll der Flächeninhalt durch Rechtecke von oben und von unten approximiert werden.



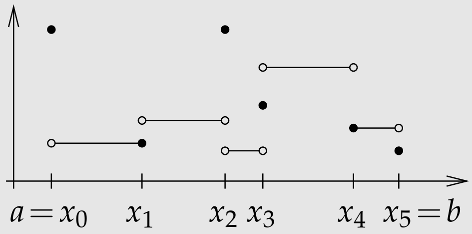


§1. Der Begriff des RIEMANN-Integrals

Um mit Rechtecken arbeiten zu können, interessieren wir uns für stückweise konstante Funktionen.

(1.1) Definition. a) Eine *Zerlegung* des Intervalls $[a,b]$ ist eine endliche reelle Folge $Z = (x_i)_{0 \leqslant i \leqslant n}$, so dass

$$a = x_0 < x_1 < \ldots < x_n = b.$$



b) Eine Funktion $\varphi : [a,b] \to \mathbb{R}$ nennt man eine RIEMANN*sche Treppenfunktion*, wenn es eine Zerlegung $Z = (x_i)_{0 \leqslant i \leqslant n}$ von $[a,b]$ gibt, so dass

$$\varphi\big|_{(x_{i-1},x_i)} \quad \text{konstant ist für } 1 \leqslant i \leqslant n.$$

Mit $T[a,b]$ bezeichnen wir die Menge der Treppenfunktionen auf $[a,b]$.

Den Vergleich von Zerlegungen behandeln wir in der

(1.2) Bemerkung. Sind $Z = (x_i)_{0 \leqslant i \leqslant n}$ und $Z' = (y_j)_{0 \leqslant j \leqslant m}$ Zerlegungen von $[a,b]$, so nennen wir Z' eine *Verfeinerung* von Z, wenn $n \leqslant m$ und es zu jedem $i \in \{1,\ldots,n\}$ ein $j(i) \in \{1,\ldots,m\}$ gibt, so dass $x_i = y_{j(i)}$.
Ist φ eine Treppenfunktion bezüglich der Zerlegung Z und Z' Verfeinerung von Z, so ist φ offenbar auch eine Treppenfunktion bezüglich Z'.
Weiterhin ist klar, dass zwei Zerlegungen von $[a,b]$ eine gemeinsame Verfeinerung besitzen: Man ordne dazu die Elemente beider Folgen der Größe nach an.

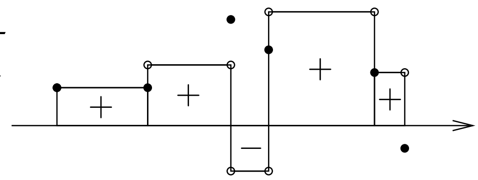
Diese Bemerkung wird nun beim Nachweis einiger grundlegender Eigenschaften verwendet.

Eine erste strukturelle Eigenschaft beschreibt das

(1.3) Lemma. *Für $a, b \in \mathbb{R}$ mit $a < b$ ist $T[a,b]$ ein Unterraum des Vektorraums aller Funktionen $f : [a,b] \to \mathbb{R}$.*

Beweis. Die konstanten Funktionen gehören zu $T[a,b]$. Für $\lambda \in \mathbb{R}$ und $\varphi \in T[a,b]$ gilt $\lambda\varphi \in T[a,b]$. Sind $\varphi, \psi \in T[a,b]$ bezüglich Zerlegungen Z bzw. Z', so betrachte eine gemeinsame Verfeinerung $Z'' = (t_k)_{0 \leqslant k \leqslant p}$ von Z und Z'. Dann sind φ und ψ, also auch $\varphi + \psi$, konstant auf jedem Teilintervall (t_k, t_{k+1}), $0 \leqslant k < p$. Also $\varphi + \psi \in T[a,b]$. $\square$

Wir wollen nun das RIEMANN-Integral für Treppenfunktionen erklären. Dazu addieren wir die Flächen der Rechtecke, wobei Rechtecke unterhalb der x-Achse negativ gewichtet werden.



Wohldefiniertheit ist hier zunächst ein Problem, deshalb zeigen wir:

(1.4) Lemma. *Seien $a, b \in \mathbb{R}$, $a < b$ und $\varphi \in T[a,b]$. Gegeben seien zwei Zerlegungen $(x_i)_{0 \leqslant i \leqslant n}$ und $(y_j)_{0 \leqslant j \leqslant m}$ von $[a,b]$, so dass*

$$\varphi\big|_{(x_{i-1},x_i)} = c_i,\ 1 \leqslant i \leqslant n, \quad \varphi\big|_{(y_{j-1},y_j)} = d_j, \quad 1 \leqslant j \leqslant m.$$

Dann gilt

$$\sum_{i=1}^{n} c_i(x_i - x_{i-1}) = \sum_{j=1}^{m} d_j(y_j - y_{j-1}).$$

Beweis. Für die Zerlegung $Z = (x_i)_{0 \leqslant i \leqslant n}$ setzen wir

$$\int\limits_Z \varphi := \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

1. Fall: $Z' = (y_j)_{0 \leqslant j \leqslant m}$ ist Verfeinerung von Z.
Dann gilt $x_i = y_{k_i}, i = 0, \ldots, n$ und

$$x_{i-1} = y_{k_{i-1}} < y_{k_{i-1}+1} < \ldots < y_{k_i} = x_i, \quad d_j = c_i, \quad k_{i-1} < j \leqslant k_i, \quad 1 \leqslant i \leqslant n.$$

Daraus folgt

$$\int\limits_{Z'} \varphi = \sum_{j=1}^{m} d_j(y_j - y_{j-1}) = \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} c_i(y_j - y_{j-1}) = \sum_{i=1}^{n} c_i(x_i - x_{i-1}) = \int\limits_Z \varphi.$$

2. Fall: Sind Z und Z' beliebig, so gibt es eine gemeinsame Verfeinerung Z^*, und es folgt mit dem 1. Fall

$$\int\limits_Z \varphi = \int\limits_{Z^*} \varphi = \int\limits_{Z'} \varphi.$$

□

Das rechtfertigt die

(1.5) Definition. Seien $a, b \in \mathbb{R}$, $a < b$, $\varphi \in T[a,b]$ mit einer Zerlegung $(x_i)_{0 \leqslant i \leqslant n}$, so dass $\varphi\big|_{(x_{i-1},x_i)} = c_i$, $1 \leqslant i \leqslant n$. Dann ist das RIEMANN-*Integral* von φ über $[a,b]$ erklärt durch

$$\int\limits_a^b \varphi(x)\,dx := \int\limits_a^b \varphi := \sum_{i=1}^{n} c_i(x_i - x_{i-1}).$$

Wir führen eine neue Sprechweise ein.

(1.6) Definition. Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a,b] \to \mathbb{R}$, $g : [a,b] \to \mathbb{R}$ beliebige Funktionen. Man definiert $f \leqslant g$ (bzw. $g \geqslant f$) genau dann, wenn $f(x) \leqslant g(x)$ für alle $x \in [a,b]$.

Man beachte, dass im Allgemeinen weder $f \leqslant g$ noch $g \leqslant f$ gilt. Die folgende Aussage besagt, dass das RIEMANN-Integral ein monotones lineares Funktional auf $T[a,b]$ ist.

(1.7) Lemma. *Seien $a, b \in \mathbb{R}$, $a < b$, φ, $\psi \in T[a,b]$, $\lambda \in \mathbb{R}$. Dann gilt*

a) $\int_a^b \varphi(x) + \psi(x)\, dx = \int_a^b \varphi(x)\, dx + \int_a^b \psi(x)\, dx.$

b) $\int_a^b \lambda\varphi(x)\, dx = \lambda \int_a^b \varphi(x)\, dx.$

c) Aus $\varphi \leqslant \psi$ *folgt* $\int_a^b \varphi(x)\, dx \leqslant \int_a^b \psi(x)\, dx.$

Beweis. Nach (1.2) kann man annehmen, dass φ und ψ bezüglich derselben Zerlegung von $[a,b]$ gegeben sind

$$a = x_0 < \ldots < x_n = b, \quad \varphi\big|_{(x_{i-1},x_i)} = c_i, \quad \psi\big|_{(x_{i-1},x_i)} = d_i, \quad 1 \leqslant i \leqslant n.$$

a) Aus der Definition folgt

$$\begin{aligned} \int\limits_a^b \varphi(x) + \psi(x)\, dx &= \sum_{i=1}^n (c_i + d_i)(x_i - x_{i-1}) = \sum_{i=1}^n c_i(x_i - x_{i-1}) + \sum_{i=1}^n d_i(x_i - x_{i-1}) \\ &= \int\limits_a^b \varphi(x)\, dx + \int\limits_a^b \psi(x)\, dx. \end{aligned}$$

b) Man hat

$$\int\limits_a^b \lambda\varphi(x)\, dx = \sum_{i=1}^n \lambda c_i(x_i - x_{i-1}) = \lambda \sum_{i=1}^n c_i(x_i - x_{i-1}) = \lambda \int\limits_a^b \varphi(x)\, dx.$$

c) Aus $\varphi \leqslant \psi$ folgt $c_i \leqslant d_i$ für $1 \leqslant i \leqslant n$. Wegen $x_i - x_{i-1} > 0$ ergibt sich

$$\int\limits_a^b \varphi(x)\, dx = \sum_{i=1}^n c_i(x_i - x_{i-1}) \leqslant \sum_{i=1}^n d_i(x_i - x_{i-1}) = \int\limits_a^b \psi(x)\, dx.$$

□

Es kann $\varphi \leqslant \psi$ und $\varphi \neq \psi$ gelten, aber trotzdem $\int_a^b \varphi(x)\, \mathrm{d}x = \int_a^b \psi(x)\, \mathrm{d}x$. Man betrachte hierzu das Beispiel

$$\varphi : [0,1] \to \mathbb{R},\ \varphi \equiv 0; \quad \psi : [0,1] \to \mathbb{R},\ \psi(x) = \begin{cases} 1, & x = 0, \\ 0, & x > 0, \end{cases}$$

mit $\int_0^1 \varphi(x)\, dx = \int_0^1 \psi(x)\, dx = 0 \cdot 1 = 0.$

Treppenfunktionen haben wir bereits früher kennen gelernt, z. B. die Signum-Funktion und die Größte-Ganze-Funktion (vgl. II(3.2)) eingeschränkt auf ein Intervall $[a,b]$.

(1.8) Definition. Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \to \mathbb{R}$ beschränkt. Dann sind das *Oberintegral* und das *Unterintegral von f über* $[a, b]$ definiert durch

$$\int_a^{b*} f(x)\,dx \quad := \quad \inf\left\{ \int_a^b \psi(x)\,dx;\ \psi \in T[a,b],\ f \leqslant \psi \right\},$$
$$\int_{a*}^{b} f(x)\,dx \quad := \quad \sup\left\{ \int_a^b \varphi(x)\,dx;\ \varphi \in T[a,b],\ \varphi \leqslant f \right\}.$$

Weil f nach Voraussetzung beschränkt ist, existieren (z. B. konstante) Treppenfunktionen $\widetilde{\varphi}$, $\widetilde{\psi} \in T[a, b]$ mit $\widetilde{\varphi} \leqslant f \leqslant \widetilde{\psi}$, so dass die Mengen, über die das Infimum bzw. Supremum gebildet wird, nicht leer sind. Weil zudem jede Treppenfunktion ψ mit $f \leqslant \psi$ auch $\widetilde{\varphi} \leqslant \psi$ erfüllt und mit (1.7) dann

$$\int_a^b \psi(x)\,\mathrm{d}x \geqslant \int_a^b \widetilde{\varphi}\,\mathrm{d}x$$

gilt, ist das Infinum im ersten Teil der Definition endlich. Analog folgt die Endlichkeit des Supremums im zweiten Teil. Als Warnung sei vermerkt, dass es im Allgemeinen keine Treppenfunktion gibt, die obiges Infimum bzw. Supremum realisiert.

(1.9) Beispiele. a) Für jedes $\varphi \in T[a, b]$ gilt

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

b) Sei D die DIRICHLETsche Sprungfunktion auf $[0, 1]$

$$D : [0,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} 1, & \text{falls } x \in \mathbb{Q}, \\ 0, & \text{falls } x \notin \mathbb{Q}. \end{cases}$$

Seien φ, $\psi \in T[0, 1]$ mit $\varphi \leqslant D \leqslant \psi$. Sei $(x_i)_{0 \leqslant i \leqslant n}$ eine Zerlegung von $[0, 1]$ und gilt

$$\varphi\big|_{(x_{i-1},x_i)} = c_i \leqslant D\big|_{(x_{i-1},x_i)} \leqslant d_i = \psi\big|_{(x_{i-1},x_i)},$$

so folgt $c_i \leqslant 0$ und $d_i \geqslant 1$, da jedes nicht-leere offene Intervall sowohl rationale (vgl. II(2.24)) als auch irrationale Punkte (vgl. III(3.18)) enthält. Also hat man

$$\int_0^1 \varphi(x)\,dx \leqslant 0, \quad \int_0^1 \psi(x)\,dx \geqslant 1.$$

Mit den konstanten Funktionen $\varphi \equiv 0$ und $\psi \equiv 1$ folgt

$$\int_{0*}^{1} D(x)\,dx = 0, \quad \int_0^{1*} D(x)\,dx = 1.$$

Eine erste Aussage für Ober- und Unterintegral beinhaltet das folgende

(1.10) Lemma. *Seien $a, b \in \mathbb{R},\ a < b$.*
a) Ist $f : [a,b] \to \mathbb{R}$ eine Funktion und sind $m, M \in \mathbb{R}$ mit

$$m \leqslant f(x) \leqslant M \quad \text{für alle } x \in [a,b],$$

so gilt

$$m(b-a) \leqslant \int_{a*}^{b} f(x)\,dx \leqslant \int_{a}^{b*} f(x)\,dx \leqslant M(b-a).$$

b) Sind $f, g : [a,b] \to \mathbb{R}$ beschränkt, so folgt

$$\begin{aligned} \int_{a}^{b*} f(x)+g(x)\,dx \quad &\leqslant \quad \int_{a}^{b*} f(x)\,dx + \int_{a}^{b*} g(x)\,dx, \\ \int_{a*}^{b} f(x)+g(x)\,dx \quad &\geqslant \quad \int_{a*}^{b} f(x)\,dx + \int_{a*}^{b} g(x)\,dx. \end{aligned}$$

c) Ist $f : [a,b] \to \mathbb{R}$ beschränkt, so gilt für alle $\alpha,\ \beta \in \mathbb{R},\ \alpha \geqslant 0,\ \beta \leqslant 0$:

$$\begin{aligned} \int_{a}^{b*} \alpha f(x)\,dx \quad &= \quad \alpha \int_{a}^{b*} f(x)\,dx, \quad \int_{a*}^{b} \alpha f(x)\,dx = \alpha \int_{a*}^{b} f(x)\,dx, \\ \int_{a}^{b*} \beta f(x)\,dx \quad &= \quad \beta \int_{a*}^{b} f(x)\,dx, \quad \int_{a*}^{b} \beta f(x)dx = \beta \int_{a}^{b*} f(x)\,dx. \end{aligned}$$

Beweis. a) Betrachtet man konstante Treppenfunktionen, so folgt aus $m \leqslant f \leqslant M$ mit (1.7) c) bereits

$$m(b-a) = \int_{a}^{b} m\,dx \leqslant \int_{a*}^{b} f(x)\,dx, \quad \int_{a}^{b*} f(x)\,dx \leqslant \int_{a}^{b} M dx = M(b-a).$$

Für alle $\varphi, \psi \in T[a,b]$ mit $\varphi \leqslant f \leqslant \psi$ gilt nach (1.7) und (1.8)

$$\begin{aligned} \int_{a}^{b} \varphi(x)\,dx \quad &\leqslant \quad \int_{a}^{b} \psi(x)\,dx, \quad \text{also} \\ \int_{a*}^{b} f(x)\,dx \quad &= \quad \sup\left\{ \int_{a}^{b} \varphi(x)\,dx;\ \varphi \in T[a,b],\ \varphi \leqslant f \right\} \leqslant \int_{a}^{b} \psi(x)\,dx, \\ \int_{a*}^{b} f(x)\,dx \quad &\leqslant \quad \inf\left\{ \int_{a}^{b} \psi(x)\,dx;\ \psi \in T[a,b],\ f \leqslant \psi \right\} = \int_{a}^{b*} f(x)\,dx. \end{aligned}$$

b) Sei $\varepsilon > 0$. Dann existieren $\varphi, \psi \in T[a,b]$ mit

$$f \leqslant \varphi, \quad \int_a^b \varphi(x)\,dx \quad \leqslant \quad \int_a^{b*} f(x)dx + \frac{\varepsilon}{2},$$

$$g \leqslant \psi, \quad \int_a^b \psi(x)\,dx \quad \leqslant \quad \int_a^{b*} g(x)\,dx + \frac{\varepsilon}{2}.$$

nach Definition des Infimums. Es folgt $f + g \leqslant \varphi + \psi$ und mit (1.7)

$$\begin{aligned} \int_a^{b*} f(x) + g(x)dx \quad &\leqslant \quad \int_a^b \varphi(x) + \psi(x)\,dx \\ = \int_a^b \varphi(x)\,dx + \int_a^b \psi(x)\,dx \quad &\leqslant \quad \int_a^{b*} f(x)\,dx + \int_a^{b*} g(x)\,dx + \varepsilon. \end{aligned}$$

Da diese Ungleichungen für alle $\varepsilon > 0$ gelten, schließt man

$$\int_a^{b*} f(x) + g(x)\,dx \leqslant \int_a^{b*} f(x)\,dx + \int_a^{b*} g(x)\,dx.$$

Die zweite Behauptung folgt analog oder aus

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

c) Für $\lambda = 0$ ist die Behauptung klar. Ist $\lambda > 0$, so gilt $\varphi \leqslant f$ genau dann, wenn $\lambda\varphi \leqslant \lambda f$. Jetzt folgt die Behauptung aus

$$\sup\{\kappa m;\ m \in M\} = \begin{cases} \kappa \sup M, & \text{falls } \kappa \geqslant 0, \\ \kappa \inf M, & \text{falls } \kappa \leqslant 0, \end{cases}$$

für beschränktes $M \subset \mathbb{R},\ M \neq \emptyset$, und $\kappa \in \mathbb{R}$. □

In Teil b) gilt i. Allg. nicht Gleichheit: Betrachte z. B. $f = D$, $g = 1 - D$ mit der Dirichlet-Funktion.

Wir kommen zu der wichtigen

(1.11) Definition. Seien $a, b \in \mathbb{R}$, $a < b$. Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt RIEMANN- *integrierbar* oder einfach nur *integrierbar*, wenn sie beschränkt ist und $\int_{a*}^b f(x)\,dx = \int_a^{b*} f(x)\,dx$ erfüllt. Dann heißt

$$\int_a^b f(x)\,dx := \int_a^b f(t)\,dt := \int_a^b f(\xi)\,d\xi := \int_a^b f := \int_{a*}^b f(x)\,dx$$

das RIEMANN-*Integral von f über $[a,b]$*. Man nennt a die *Untergrenze* und b die *Obergrenze* des Integrals.

Wir erläutern den Begriff zunächst an Beispielen.

(1.12) Beispiele. a) Nach (1.9) ist jede Treppenfunktion integrierbar. Die in (1.5) und (1.11) definierten Integrale stimmen überein.
b) Die DIRICHLETsche Sprungfunktion über $[0,1]$ ist nach (1.9) nicht integrierbar.
c) Wir betrachten

$$f : [0,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} 0, & \text{falls } x \in \mathbb{R}\backslash\mathbb{Q}, \\ \dfrac{1}{q}, & \text{falls } x \in \mathbb{Q},\ x = \frac{p}{q}, p \in \mathbb{N}_0,\ q \in \mathbb{N},\ \text{gekürzte Bruchdarstellung.} \end{cases}$$

Offenbar gilt $0 \leqslant f$ und damit

$$0 \leqslant \int_{0*}^{1} f(x)\,dx \leqslant \int_{0}^{1*} f(x)\,dx.$$

Sei $\varepsilon > 0$. Dann ist

$$M := \left\{ \frac{p}{q};\ 1 \leqslant q \leqslant \frac{1}{\varepsilon},\ 0 \leqslant p \leqslant q \right\}$$

endlich und wir erhalten eine Treppenfunktion

$$\psi \in T[0,1], \quad \psi(x) = \begin{cases} f(x), & \text{falls } x \in M, \\ \varepsilon, & \text{falls } x \notin M. \end{cases}$$

Aus $f \leqslant \psi$ folgt

$$\int_{0}^{1*} f(x)\,dx \leqslant \int_{0}^{1} \psi(x)\,dx = \varepsilon$$

für jedes $\varepsilon > 0$. Demnach ist f integrierbar mit

$$\int_{0}^{1} f(x)\,dx = 0.$$

Es ist ziemlich aufwändig, Integrale mit der Definition (1.11) zu berechnen. Wir werden in §2 Methoden kennen lernen, die die konkrete Berechnung von Integralen für Standardfunktionen ermöglichen.

Ein für theoretische Zwecke nützliches Kriterium beinhaltet der

(1.13) Satz. *Seien $a, b \in \mathbb{R}$, $a < b$. Eine Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann* RIEMANN-*integrierbar, wenn es zu jedem $\varepsilon > 0$ Treppenfunktionen $\varphi, \psi \in T[a,b]$ gibt, so dass*

$$\varphi \leqslant f \leqslant \psi \quad \text{und} \quad \int_{a}^{b} \psi(x) - \varphi(x)\,dx \leqslant \varepsilon.$$

Beweis. „$\Rightarrow$" Nach Definition des Infimums und des Supremums existieren zu jedem $\varepsilon > 0$ nach (1.8) Treppenfunktionen $\varphi, \psi \in T[a,b]$ mit $\varphi \leqslant f \leqslant \psi$ und

$$\int_a^b \varphi(x)\,dx \geqslant \int_{a*}^b f(x)\,dx - \frac{\varepsilon}{2}, \quad \int_a^b \psi(x)\,dx \leqslant \int_a^{b*} f(x)\,dx + \frac{\varepsilon}{2}.$$

Da f integrierbar ist, folgt

$$\int_a^b \psi(x) - \varphi(x)\,dx = \int_a^b \psi(x)\,dx - \int_a^b \varphi(x)\,dx \leqslant \int_a^{b*} f(x)\,dx + \frac{\varepsilon}{2} - \int_{a*}^b f(x)\,dx + \frac{\varepsilon}{2} = \varepsilon.$$

„$\Leftarrow$" Wegen $\varphi \leqslant f \leqslant \psi$ ist f beschränkt. Mit (1.10) ergibt sich

$$\int_a^b \varphi(x)\,dx \leqslant \int_{a*}^b f(x)\,dx \leqslant \int_a^{b*} f(x)\,dx \leqslant \int_a^b \psi(x)\,dx,$$

also nach Voraussetzung

$$0 \leqslant \int_a^{b*} f(x)\,dx - \int_{a*}^b f(x)\,dx \leqslant \int_a^b \psi(x)\,dx - \int_a^b \varphi(x)dx = \int_a^b \psi(x) - \varphi(x)\,dx \leqslant \varepsilon.$$

Demnach stimmen Ober- und Unterintegral überein. □

Direkt aus dem Beweis folgert man noch das

(1.14) Korollar. *Seien $a, b \in \mathbb{R}$, $a < b$, eine Funktion $f : [a,b] \to \mathbb{R}$ und $I \in \mathbb{R}$ gegeben. Dann sind äquivalent:*

(i) f ist RIEMANN-*integrierbar mit $\int_a^b f(x)\,dx = I$.*

(ii) Zu jedem $\varepsilon > 0$ gibt es Treppenfunktionen $\varphi, \psi \in T[a,b]$ mit

$$\varphi \leqslant f \leqslant \psi, \quad I - \int_a^b \varphi(x)\,dx \leqslant \varepsilon, \quad \int_a^b \psi(x)\,dx - I \leqslant \varepsilon.$$

Als Anwendung erhalten wir die Aussage, dass die Abänderung einer Funktion an endlich vielen Stellen weder die RIEMANN-Integrierbarkeit noch das Integral verändert.

(1.15) Korollar. *Seien $a, b \in \mathbb{R}$, $a < b$, $f : [a,b] \to \mathbb{R}$ eine* RIEMANN-*integrierbare Funktion und $g : [a,b] \to \mathbb{R}$ eine beliebige Funktion, so dass*

$$\{x \in [a,b];\ f(x) \neq g(x)\}$$

endlich ist. Dann ist auch g RIEMANN-*integrierbar mit*

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

Beweis. Sei $\varepsilon > 0$ und $I = \int_a^b f(x)\,dx$. Nach (1.14) gibt es $\varphi, \psi \in T[a,b]$ mit

$$\varphi \leqslant f \leqslant \psi, \quad I - \int_a^b \varphi(x)\,dx \leqslant \varepsilon, \quad \int_a^b \psi(x)\,dx - I \leqslant \varepsilon.$$

Sei $M := \{x \in [a,b];\ f(x) \neq g(x)\}$. Da M endlich ist, sind

$$\varphi^*(x) := \begin{cases} g(x), & x \in M, \\ \varphi(x), & x \notin M, \end{cases} \quad \psi^*(x) := \begin{cases} g(x), & x \in M, \\ \psi(x), & x \notin M, \end{cases}$$

Treppenfunktionen auf $[a,b]$ mit den Eigenschaften

$$\varphi^* \leqslant g \leqslant \psi^*, \quad \int_a^b \varphi(x)\,dx = \int_a^b \varphi^*(x)\,dx, \quad \int_a^b \psi(x)\,dx = \int_a^b \psi^*(x)\,dx,$$

also

$$I - \int_a^b \varphi^*(x)\,dx \leqslant \varepsilon, \quad \int_a^b \psi^*(x)\,dx - I \leqslant \varepsilon.$$

Dann ist g nach (1.14) RIEMANN-integrierbar mit $\int_a^b g(x)\,dx = I$. □

Wir kommen nun zu den wichtigsten Beispielklassen integrierbarer Funktionen.

(1.16) Satz. *Sind $a, b \in \mathbb{R}$, $a < b$, so ist jede monotone Funktion $f : [a,b] \to \mathbb{R}$ integrierbar.*

Beweis. Der Beweis wird für monoton wachsendes f geführt. Für monoton fallendes f läuft die Argumentation analog. Zu $n \in \mathbb{N}$ betrachte man die äquidistante Zerlegung

$$(x_i)_{0 \leqslant i \leqslant n}, \quad x_i := a + i \cdot \frac{b-a}{n}, \ 0 \leqslant i \leqslant n.$$

Wir definieren $\varphi, \psi \in T[a,b]$ durch

$$\varphi(x) := f(x_{i-1}),\ \psi(x) := f(x_i) \quad \text{für} \quad x_{i-1} \leqslant x < x_i,\ 1 \leqslant i \leqslant n,\ \varphi(b) = \psi(b) = f(b).$$

Da f monoton wachsend ist, gilt $\varphi \leqslant f \leqslant \psi$ und

$$\int_a^b \psi(x) - \varphi(x)\,dx = \sum_{i=1}^n (f(x_i) - f(x_{i-1}))(x_i - x_{i-1})$$

$$= \frac{b-a}{n} \left(\sum_{i=1}^n f(x_i) - \sum_{i=1}^n f(x_{i-1}) \right) = \frac{b-a}{n}(f(x_n) - f(x_0)) = \frac{b-a}{n}(f(b) - f(a)).$$

Zu $\varepsilon > 0$ wählt man nun ein $n \in \mathbb{N}$ mit $\frac{b-a}{n}(f(b) - f(a)) < \varepsilon$. Also ist f nach (1.13) integrierbar. □

Das nächste zentrale Beispiel wird behandelt in dem

(1.17) Satz. *Seien $a, b \in \mathbb{R}$, $a < b$. Dann ist jede stetige Funktion $f : [a, b] \to \mathbb{R}$ integrierbar.*

Beweis. Da $[a, b]$ ein kompaktes Intervall ist, ist f nach V(3.12) gleichmäßig stetig auf $[a, b]$. Zu $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $x, y \in [a, b]$ mit $|x - y| < \delta$ gilt

$$(*) \qquad |f(x) - f(y)| < \varepsilon', \quad \varepsilon' := \frac{\varepsilon}{2(b-a)}.$$

Nun wählen wir ein $n \in \mathbb{N}$ mit $\dfrac{b-a}{n} < \delta$ und betrachten wieder die äquidistante Zerlegung

$$(x_i)_{0 \leqslant i \leqslant n}, \; x_i := a + i \cdot \frac{b-a}{n}, \; 0 \leqslant i \leqslant n.$$

Wir definieren $\varphi, \psi \in T[a, b]$ durch

$$\varphi(x) := f(x_i) - \varepsilon', \; \psi(x) := f(x_i) + \varepsilon' \quad \text{für} \quad x_{i-1} \leqslant x < x_i, \; 1 \leqslant i \leqslant n, \; \varphi(b) = \psi(b) = f(b).$$

Wegen $(*)$ folgt $\varphi \leqslant f \leqslant \psi$ aus der Wahl von n. Dann erhält man

$$\int_a^b \psi(x) - \varphi(x)\, dx = 2\varepsilon'(b-a) = \varepsilon.$$

Die Behauptung ergibt sich wieder aus (1.13) □

Wir stellen nun einige elementare Eigenschaften des RIEMANN-Integrals zusammen.

(1.18) Satz. *Seien $a, b \in \mathbb{R}$, $a < b$ und $f, g : [a, b] \to \mathbb{R}$ RIEMANN-integrierbar. Dann gilt:*
a) Für $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ integrierbar mit

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

b) Ist $[c, d] \subset [a, b]$, so ist $f\big|_{[c,d]}$ integrierbar.
c) Sei $a < c < b$ und $h : [a, b] \to \mathbb{R}$ eine Funktion. h ist genau dann integrierbar, wenn $h\big|_{[a,c]}$ und $h\big|_{[c,b]}$ integrierbar sind. In diesem Fall gilt

$$\int_a^b h(x)\, dx = \int_a^c h(x)\, dx + \int_c^b h(x)\, dx.$$

d) Gilt $f \leqslant g$, so folgt

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

Beweis. a) Man verwende (1.10) in

$$\int_a^b f + \int_a^b g = \int_{a*}^b f + \int_{a*}^b g \leqslant \int_{a*}^b (f+g) \leqslant \int_a^{b*} (f+g) \leqslant \int_a^{b*} f + \int_a^{b*} g = \int_a^b f + \int_a^b g.$$

Also gilt überall „=". Insbesondere ist daher $f+g$ integrierbar mit $\int_a^b (f+g) = \int_a^b f + \int_a^b g$. Aus (1.10) c) erhält man $\int_a^b \lambda f = \lambda \int_a^b f$.
b) Man verwende (1.13). Aus $0 \leqslant \psi - \varphi$ und der Definition des RIEMANN-Integrals für Treppenfunktionen folgt sofort

$$0 \leqslant \int_c^d \psi(x) - \varphi(x)\, dx \leqslant \int_a^b \psi(x) - \varphi(x)\, dx \leqslant \varepsilon.$$

c) Man verwende b) und (1.14).
d) Es gilt $0 \leqslant g - f$. Da 0 eine Treppenfunktion ist, ergibt sich

$$0 = \int_a^b 0\, dx \leqslant \int_{a*}^b g(x) - f(x)\, dx = \int_a^b g(x)\, dx - \int_a^b f(x)\, dx.$$

□

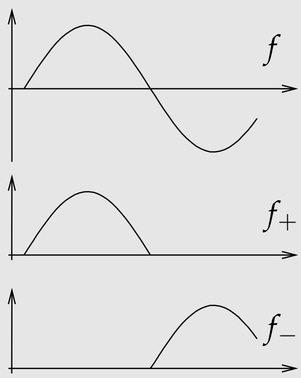
Aus a) und d) folgt, dass das RIEMANN-Integral ein monotones lineares Funktional auf dem Vektorraum $\mathcal{R}[a,b]$ der auf $[a,b]$ RIEMANN-integrierbaren Funktionen ist.

Nun zerlegen wir eine Funktion in ihren Positiv- und Negativteil.

(1.19) Definition. Zu einer Funktion $f : D \to \mathbb{R}$ definieren wir die beiden Funktionen $f_\pm : D \to \mathbb{R}_+$ durch

$$f_+(x) := \max\{f(x),\, 0\} = \begin{cases} f(x), & \text{falls } f(x) \geqslant 0, \\ 0, & \text{falls } f(x) \leqslant 0, \end{cases}$$

$$f_-(x) := -\min\{f(x),\, 0\} = \begin{cases} -f(x), & \text{falls } f(x) \leqslant 0, \\ 0, & \text{falls } f(x) \geqslant 0. \end{cases}$$



Offenbar gilt dann

$$f = f_+ - f_- \quad \text{und} \quad |f| = f_+ + f_-.$$

Den Zusammenhang mit der Integrierbarkeit beinhaltet der

(1.20) Satz. *Gegeben seien $a, b \in \mathbb{R}$, $a < b$ und Funktionen $f, g : [a, b] \to \mathbb{R}$.*
a) f ist genau dann integrierbar, wenn f_+ und f_- integrierbar sind.
b) Mit f ist $|f|$ integrierbar und es gilt die Dreiecksungleichung

$$\left| \int_a^b f(x)\,dx \right| \leqslant \int_a^b |f(x)|\,dx.$$

c) Mit f und g ist auch $f \cdot g$ und jedes Polynom in f integrierbar.

Nach c) ist $\mathcal{R}[a, b]$ sogar eine $\mathbb{R}$-Algebra.

Beweis. a) "$\Leftarrow$" Man verwende $f = f_+ - f_-$ und (1.18) a).
"$\Rightarrow$" Zu $\varepsilon > 0$ wählen wir $\varphi, \psi \in T[a, b]$ nach (1.13) mit $\varphi \leqslant f \leqslant \psi$ und $\int_a^b \psi(x) - \varphi(x)\,dx \leqslant \varepsilon$. Dann gilt $\varphi_\pm, \psi_\pm \in T[a, b]$ mit $\varphi_+ \leqslant f_+ \leqslant \psi_+$ und $\psi_- \leqslant f_- \leqslant \varphi_-$.

Wir führen für die zweite Ungleichungskette den (elementaren) Nachweis ausführlich. Falls $f(x) \geqslant 0$, so $\psi(x) \geqslant 0$, also

$$f_-(x) = \psi_-(x) = 0 \quad \text{und} \quad f_-(x) \geqslant \psi_-(x).$$

Trivialerweise ist auch $f_-(x) = 0 \leqslant \psi_-(x)$. Falls $f(x) < 0$, ist jedenfalls $\varphi(x) < 0$ und $-f_-(x) = f(x) \geqslant \varphi(x) = -\varphi_-(x)$, also $f_-(x) \leqslant \varphi_-(x)$.

Falls zudem $\psi(x) < 0$, so $f_-(x) \geqslant \psi_-(x)$ mit derselben Argumentation. Ist schließlich $\psi(x) \geqslant 0$, so ist $f_-(x) \geqslant 0 = \psi_-(x)$ trivial.

Wegen $\psi_+ - \varphi_+ \leqslant \psi - \varphi$ und $\varphi_- - \psi_- \leqslant \psi - \varphi$ folgt nun

$$\int_a^b \psi_+(x) - \varphi_+(x)\,dx \leqslant \int_a^b \psi(x) - \varphi(x)\,dx \leqslant \varepsilon,$$
$$\int_a^b \varphi_-(x) - \psi_-(x)\,dx \leqslant \int_a^b \psi(x) - \varphi(x)\,dx \leqslant \varepsilon.$$

Also sind f_+ und f_- nach (1.13) integrierbar.
b) $|f| = f_+ + f_-$ ist nach a) und (1.18) integrierbar und es gilt

$$\left| \int_a^b f(x)\,dx \right| = \left| \int_a^b f_+(x)\,dx - \int_a^b f_-(x)\,dx \right| \leqslant \max \left\{ \int_a^b f_+(x)\,dx, \int_a^b f_-(x)\,dx \right\}$$
$$\leqslant \int_a^b f_+(x)\,dx + \int_a^b f_-(x)\,dx = \int_a^b |f(x)|\,dx.$$

c) Wegen $f \cdot g = \frac{1}{4}(f+g)^2 - \frac{1}{4}(f-g)^2$ und (1.18) genügt es zu zeigen, dass f^2 integrierbar ist, denn der Rest folgt per Induktion. Wegen $f^2 = |f|^2$ und b) darf man $f \geqslant 0$ annehmen. Sei

$M > 0$ mit $0 \leqslant f \leqslant M$. Zu $\varepsilon > 0$ existieren nach (1.13) $\varphi, \psi \in T[a,b]$ mit $0 \leqslant \varphi \leqslant f \leqslant \psi \leqslant M$ und $\int_a^b \psi(x) - \varphi(x)\, dx \leqslant \varepsilon/(2M)$. Nun gilt $\varphi^2, \psi^2 \in T[a,b]$ mit $\varphi^2 \leqslant f^2 \leqslant \psi^2$. Wegen

$$\psi^2 - \varphi^2 = (\psi - \varphi) \cdot (\psi + \varphi) \leqslant 2M \cdot (\psi - \varphi)$$

folgt mit (1.18) d)

$$\int_a^b \psi^2(x) - \varphi^2(x)\, dx \leqslant 2M \int_a^b \psi(x) - \varphi(x)\, dx \leqslant \varepsilon.$$

Also ist f^2 nach (1.13) ebenfalls integrierbar. □

Als Verallgemeinerung notieren wir den

(1.21) Satz. *(*HÖLDER*sche Ungleichung)*
Seien $a, b \in \mathbb{R}$, $a < b$ sowie $f, g : [a,b] \to \mathbb{R}$ integrierbar und $p > 1, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann sind auch $|f|^p$ und $|g|^q$ integrierbar und es gilt

$$\int_a^b |f(x)g(x)|\, dx \leqslant \left(\int_a^b |f(x)|^p\, dx \right)^{1/p} \cdot \left(\int_a^b |g(x)|^q\, dx \right)^{1/q}.$$

Beweis. $|f|, |g|$ und $|f \cdot g|$ sind nach (1.20) integrierbar. Sei $M > 0$ mit $|f| \leqslant M$ und $|g| \leqslant M$. Wir zeigen nun die Integrierbarkeit von $|f|^p$. Zu $\varepsilon > 0$ wählen wir nach (1.13)

$$\varphi, \psi \in T[a,b], \quad 0 \leqslant \varphi \leqslant |f| \leqslant \psi \leqslant M, \quad \int_a^b \psi(x) - \varphi(x)\, dx \leqslant \varepsilon', \quad \varepsilon' = \frac{\varepsilon}{pM^{p-1}}.$$

Dann gilt $\varphi^p, \psi^p \in T[a,b]$ mit $0 \leqslant \varphi^p \leqslant |f|^p \leqslant \psi^p \leqslant M^p$. Wendet man für $x \in [a,b]$ den Mittelwertsatz VII(2.5) auf die Funktion

$$[\varphi(x), \psi(x)] \to \mathbb{R}, \quad y \mapsto y^p, \quad \text{mit} \quad \frac{d}{dy}(y^p) = p \cdot y^{p-1}$$

an, so folgt für ein geeignetes $\xi \in [\varphi(x), \psi(x)]$

$$0 \leqslant \psi^p(x) - \varphi^p(x) = p \cdot \xi^{p-1} \cdot (\psi(x) - \varphi(x)) \leqslant p \cdot M^{p-1} \cdot (\psi(x) - \varphi(x)).$$

Daraus ergibt sich

$$\int_a^b \psi^p(x) - \varphi^p(x)\, dx \leqslant p \cdot M^{p-1} \cdot \int_a^b \psi(x) - \varphi(x)\, dx \leqslant \varepsilon.$$

Also ist $|f|^p$ und analog auch $|g|^q$ nach (1.13) integrierbar.
Nun wählt man $\varphi_1, \psi_1 \in T[a,b]$ mit

$$0 \leqslant \varphi_1 \leqslant |g| \leqslant \psi_1 \leqslant M, \quad \int_a^b \psi_1(x) - \varphi_1(x)\, dx \leqslant \varepsilon'.$$

Dann gilt

$$\begin{aligned}
0 &\leqslant \int_a^b \psi(x)\psi_1(x) - \varphi(x)\varphi_1(x)\,dx = \int_a^b [\psi(x)-\varphi(x)]\psi_1(x) + \varphi(x)[\psi_1(x)-\varphi_1(x)]\,dx \\
&\leqslant M\cdot\int_a^b \psi(x)-\varphi(x)\,dx + M\cdot\int_a^b \psi_1(x)-\varphi_1(x)\,dx \leqslant 2M\varepsilon'.
\end{aligned}$$

Wir betrachten eine Zerlegung $(x_i)_{0\leqslant i\leqslant n}$ von $[a,b]$ mit $\varphi\big|_{(x_{i-1},x_i)} = \alpha_i$ und $\varphi_1\big|_{(x_{i-1},x_i)} = \beta_i$, $i = 1,\ldots,n$. Wegen $0 \leqslant \varphi\varphi_1 \leqslant |fg| \leqslant \psi\psi_1$ folgt

$$\begin{aligned}
&\int_a^b |f(x)g(x)|dx \leqslant \int_a^b \psi(x)\psi_1(x)\,dx \leqslant 2M\varepsilon' + \int_a^b \varphi(x)\varphi_1(x)\,dx \\
&= 2M\varepsilon' + \sum_{i=1}^n \alpha_i\beta_i(x_i - x_{i-1}) = 2M\varepsilon' + \sum_{i=1}^n \alpha_i(x_i-x_{i-1})^{1/p}\cdot\beta_i(x_i-x_{i-1})^{1/q} \\
&\leqslant 2M\varepsilon' + \left(\sum_{i=1}^n \alpha_i^p(x_i-x_{i-1})\right)^{1/p}\cdot\left(\sum_{i=1}^n \beta_i^q(x_i-x_{i-1})\right)^{1/q} \\
&= 2M\varepsilon' + \left(\int_a^b \varphi^p(x)\,dx\right)^{1/p}\cdot\left(\int_a^b \varphi_1^q(x)\,dx\right)^{1/q} \\
&\leqslant 2M\varepsilon' + \left(\int_a^b |f(x)|^p\,dx\right)^{1/p}\cdot\left(\int_a^b |g(x)|^q\,dx\right)^{1/q},
\end{aligned}$$

wenn man $\dfrac{1}{p} + \dfrac{1}{q} = 1$ sowie die HÖLDERsche Ungleichung VII(4.6) verwendet. Da die Ungleichung für jedes $\varepsilon > 0$ gilt, folgt die Behauptung. □

Als direkte Anwendung erhalten wir das

(1.22) Korollar. *(CAUCHY-SCHWARZsche Ungleichung/Dreiecksungleichung)*
Sind $a,b \in \mathbb{R}$, $a < b$ und $f,g : [a,b] \to \mathbb{R}$ integrierbar, so sind $|fg|$, $|f|^2$, $|g|^2|$ und $|f+g|^2$ integrierbar und es gilt

a) $\int_a^b |f(x)g(x)|\,dx \leqslant \sqrt{\int_a^b |f(x)|^2\,dx}\cdot\sqrt{\int_a^b |g(x)|^2 dx}.$

b) $\sqrt{\int_a^b |f(x)+g(x)|^2\,dx} \leqslant \sqrt{\int_a^b |f(x)|^2 dx} + \sqrt{\int_a^b |g(x)|^2\,dx}.$

Beweis. a) Das ist die Aussage (1.21) mit $p = q = 2$.
b) Man quadriere und verwende a). □

Wir notieren eine nützliche Erweiterung der Definition.

(1.23) Definition. Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \to \mathbb{R}$ integrierbar. Dann setzt man

$$\int_b^a f(x)\,dx := -\int_a^b f(x)\,dx, \qquad \int_a^a f(x)\,dx = 0.$$

Mit dieser Erweiterung gilt (1.18) c) für beliebige Punkte c, sofern die Funktion über dem größten Intervall integrierbar ist.

Zum Verständnis der Lehrbuchliteratur wichtig ist die

(1.24) Bemerkung. An Stelle von Treppenfunktionen kann man zur Definition des Integrals auch RIEMANN*sche Summen* verwenden:

Sei $f : [a, b] \to \mathbb{R}$ beschränkt und $I \in \mathbb{R}$. Die Funktion f ist genau dann integrierbar mit $\int_a^b f(x)\,dx = I$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für jede Zerlegung $(x_i)_{0 \leqslant i \leqslant n}$ von $[a, b]$ mit $\max\{x_i - x_{i-1};\ 1 \leqslant i \leqslant n\} \leqslant \delta$ und jede Wahl von Stützstellen $\xi_i \in [x_{i-1}, x_i]$, $1 \leqslant i \leqslant n$, gilt

$$\left| I - \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \right| \leqslant \varepsilon.$$

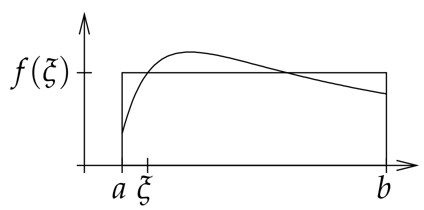
Zum Beweis vergleiche man O. FORSTER, Analysis I, §18.

Bei dem Zugang über RIEMANNsche Summen wird deutlich, dass Integration als ein "Grenzfall von Summation" aufgefasst werden kann. In der Tat ist das – von LEIBNIZ eingeführte – Symbol $\int$ ein auseinandergezogenes "S", welches für "Summe" steht. Die Beweise werden bei diesem intuitiven Zugang nicht einfacher, aber man versteht vor diesem Hintergrund besser, wie sich z. B. die HÖLDERsche Ungleichung von Summen auf Integrale fortsetzt.

§2. Integration und Differentiation

Im gesamten Paragrafen seien $a, b \in \mathbb{R}$, $a < b$.

Als wichtiges Hilfsmittel benötigen wir den



(2.1) Satz. *a) (Mittelwertsatz der Integralrechnung)*
Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann gibt es (mindestens) ein $\xi \in [a,b]$ mit

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

b) (Verallgemeinerter Mittelwertsatz der Integralrechnung) Seien $f, g : [a,b] \to \mathbb{R}$ Funktionen, f stetig, g integrierbar und $g(x) \geqslant 0$ für alle $x \in [a,b]$. Dann existiert (mindestens) ein $\xi \in [a,b]$ mit

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

Beweis. b) Nach (1.17) ist f und nach (1.20) auch $f \cdot g$ integrierbar. Weil f stetig ist, existieren nach V(3.7)

$$m := \min\{f(x);\ x \in [a,b]\},\ M := \max\{f(x);\ x \in [a,b]\}.$$

Aus $g \geqslant 0$ folgt

$$mg(x) \leqslant f(x)g(x) \leqslant Mg(x) \quad \text{für alle} \quad x \in [a,b].$$

Mit (1.18) erhält man daraus

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

Gilt $\int_a^b g(x)\,dx = 0$, so kann man $\xi = (a+b)/2$ wählen, da beide Seiten 0 sind. Gilt $\int_a^b g(x)\,dx \neq 0$, also $\int_a^b g(x)\,dx > 0$, so existiert ein $\mu \in [m,M]$ mit

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

Nach dem Zwischenwertsatz V(3.8) gibt es ein $\xi \in [a,b]$ mit $f(\xi) = \mu$.
a) Man setze $g(x) = 1$ in b). Dann folgt die Behauptung aus $\int_a^b 1\,dx = b-a$. □

Im weiteren Verlauf betrachten wir das Integral als Funktion der oberen Grenze.

(2.2) Satz. *Sei $f : [a,b] \to \mathbb{R}$ integrierbar, $c \in [a,b]$ und*

$$F : [a,b] \to \mathbb{R}, \quad F(x) := \int_c^x f(t)\,dt.$$

Dann gilt:
a) F ist gleichmäßig stetig auf $[a,b]$.
b) Ist f stetig in $x_0 \in [a,b]$, so ist F differenzierbar in x_0 mit $F'(x_0) = f(x_0)$, d. h.

$$\frac{d}{dx}\int_c^x f(t)\,dt\Big|_{x=x_0} = f(x_0).$$

c) Ist f stetig auf $[a,b]$, so ist F differenzierbar auf $[a,b]$ mit $F' = f$, d. h. F ist eine Stammfunktion von f.

Beweis. F existiert nach (1.18) und (1.23).
a) Als integrierbare Funktion ist f beschränkt. Sei also $M > 0$ mit $|f(x)| \leqslant M$ für alle $x \in [a,b]$. Seien $x, x_0 \in [a,b]$. Dann gilt nach (1.23), (1.18) und (1.20)

$$\begin{aligned} |F(x) - F(x_0)| &= \left| \int_c^x f(t)\,dt - \int_c^{x_0} f(t)\,dt \right| = \left| \int_{x_0}^x f(t)\,dt \right| \\ &\leqslant \left| \int_{x_0}^x |f(t)| dt \right| \leqslant \left| \int_{x_0}^x M\,dt \right| = M \cdot |x - x_0|. \end{aligned}$$

Also ist F nach V(3.13) c) LIPSCHITZ-stetig und damit gleichmäßig stetig auf $[a,b]$.
b) Sei $h \neq 0$ mit $x_0 + h \in [a,b]$. Dann gilt

$$\begin{aligned} \left| \frac{F(x_0+h) - F(x_0)}{h} - f(x_0) \right| &= \left| \frac{1}{h} \int_c^{x_0+h} f(t)\,dt - \frac{1}{h} \int_c^{x_0} f(t)\,dt - f(x_0) \right| \\ = \left| \frac{1}{h} \int_{x_0}^{x_0+h} f(t)\,dt - \frac{1}{h} \int_{x_0}^{x_0+h} f(x_0)\,dt \right| &= \left| \frac{1}{h} \int_{x_0}^{x_0+h} f(t) - f(x_0)\,dt \right| \\ \leqslant \frac{1}{|h|} \left| \int_{x_0}^{x_0+h} |\, f(t) - f(x_0) \,|\,dt \right|. \end{aligned}$$

Weil f in x_0 stetig ist, existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass $|f(t) - f(x_0)| < \varepsilon$ für alle $|t - x_0| < \delta$. Also gilt für $0 < |h| < \delta$

$$\left| \frac{1}{h} \int_{x_0}^{x_0+h} |\, f(t) - f(x_0) \,|\,dt \right| \leqslant \frac{1}{|h|} \left| \int_{x_0}^{x_0+h} \varepsilon\,dt \right| = \varepsilon.$$

Es folgt

$$\frac{F(x_0+h) - F(x_0)}{h} \xrightarrow[h \to 0]{} f(x_0) = F'(x_0).$$

c) Man verwende b). □

Wichtig sind die

(2.3) Bemerkungen. a) Man nennt die Aussagen von (2.2) den *Ersten Fundamentalsatz der Differential- und Integralrechnung*.
b) In b) und c) ist in den Randpunkten natürlich einseitige Differenzierbarkeit gemeint.

c) Man kann auf die Stetigkeit in b) nicht verzichten. Wir betrachten $f = \operatorname{sgn}\big|_{[-1,1]}$. Dann gilt

$$F(x) := \int_0^x \operatorname{sgn}(t)\,dt = |x| \quad \text{für} \quad x \in [-1,1].$$

F ist zwar stetig, aber in 0 nicht differenzierbar.
d) Nach (2.2) c) hat jede stetige Funktion auf einem Intervall eine Stammfunktion. Für integrierbare Funktionen ist das nicht mehr richtig. Sei $f : [-1,1] \to \mathbb{R}$, $f(x) = 0$ für $x \neq 0, f(0) = 1$. Wäre F eine Stammfunktion von f, so würde $F'\big|_{[-1,0)} = 0$ und $F'\big|_{(0,1]} = 0$ gelten. Also ist F auf $[-1,0)$ und auf $(0,1]$ jeweils konstant. Weil F stetig ist, ist F auf $[-1,1]$ konstant im Widerspruch zu $F' = f \neq 0$.

Man beachte, dass sich zwei Stammfunktionen von f auf $[a,b]$ nur um eine additive Konstante unterscheiden.

(2.4) Satz. *(Zweiter Fundamentalsatz der Differential- und Integralrechnung)*
Sei $f : [a,b] \to \mathbb{R}$ stetig und F eine Stammfunktion von f. Dann gilt

$$\int_a^b f(t)\,dt = F(t)\big|_a^b := F(b) - F(a).$$

Beweis. $G(x) := \int_a^x f(t)\,dt$ ist nach (2.2) eine Stammfunktion von f. Nach VII(2.11) existiert ein $c \in \mathbb{R}$ mit $G(x) = F(x) + c$. Wegen $G(a) = \int_a^a f(t)\,dt = 0$ erhält man

$$\int_a^b f(t)\,dt = G(b) = G(b) - G(a) = F(b) - F(a).$$

□

Der Satz bleibt mit einem aufwändigeren Beweis richtig, wenn man f nur als integrierbar, statt stetig voraussetzt. Man vergleiche H. Heuser, Lehrbuch der Analysis, Teil 1, 11. Auflage, 79.1.

(2.5) Bezeichnung. Wir führen das *unbestimmte Integral* ein: Die Schreibweise

$$\int f(x)\,dx = F(x)$$

bedeutet (*keine Gleichheit*, sondern)

$$\int_a^b f(x)\,dx = F(x)\big|_a^b = F(b) - F(a)$$

für jedes Intervall $[a,b]$ im Definitionsbereich von f.

Manche Autoren bezeichnen mit $\int f(x)\,dx$ auch die Menge aller Stammfunktionen von $f(x)$.

Aus Kapitel VI kennen wir schon eine ganze Reihe von Stammfunktionen.

(2.6) Beispiele. a) $\int x^n\,dx = \frac{1}{n+1}x^{n+1},\ n \in \mathbb{N}_0$.

b) $\int x^s\,dx = \frac{1}{s+1}x^{s+1}$ für $s \in \mathbb{R},\ s \neq -1$, Definitionsbereich $\mathbb{R}_+^*$.

c) $\int \frac{1}{x}\,dx = \ln|x|$, Definitionsbereich $\mathbb{R}\setminus\{0\}$, d. h. $0 \notin [a,b]$.

d) $\int \sin x\,dx = -\cos x,\ \int \cos x\,dx = \sin x$.

e) $\int \exp(x)\,dx = \exp(x)$.

f) $\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin x$ für $|x| < 1$.

g) $\int \frac{1}{1+x^2}\,dx = \arctan x$.

h) $\int \frac{1}{\cos^2 x}\,dx = \tan x$, falls $\cos x \neq 0$ für alle $x \in [a,b]$.

i) $\int \frac{f'(x)}{f(x)}\,dx = \ln|f(x)|$, falls $f(x) \neq 0$ für alle $x \in [a,b]$, f stetig differenzierbar.

Es gibt einige Verfahren, die beim konkreten Berechnen von Integralen helfen können. Die Produktregel der Differentiation führt zu

(2.7) Satz. *(Partielle Integration)*
Sind $f, g : [a,b] \to \mathbb{R}$ stetig differenzierbar, so gilt

$$\int_a^b f(t)g'(t)\,dt = f(t)g(t)\Big|_a^b - \int_a^b f'(t)g(t)\,dt.$$

Beweis. Als stetige Funktionen sind alle auftretenden Integranden nach (1.17) integrierbar. Die Behauptung folgt aus (2.4) und (1.18):

$$\begin{aligned} f(t)g(t)\Big|_a^b &= \int_a^b \frac{d}{dt}(f(t)g(t))\,dt = \int_a^b f'(t)g(t) + f(t)g'(t)\,dt \\ &= \int_a^b f'(t)g(t)\,dt + \int_a^b f(t)g'(t)\,dt. \end{aligned}$$ □

Wenn man partielle Integration verwenden will, liegt in der Praxis das Problem darin, zu entscheiden, wie man f und g zu wählen hat, damit das auf der rechten Seite entstehende Integral einfacher wird.

Die Kettenregel der Differentiation führt zu

(2.8) Satz. *(Substitution)*
Sei $f : D \to \mathbb{R}$ stetig und $\varphi : [a, b] \to \mathbb{R}$ stetig differenzierbar mit $\varphi([a, b]) \subset D$. Dann gilt

$$\int\limits_a^b f(\varphi(t))\varphi'(t)\,dt = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$$

Ist φ darüber hinaus streng monoton mit $\varphi(a) = \alpha, \varphi(b) = \beta$ und ist ψ die Umkehrfunktion von φ, so gilt

$$\int\limits_\alpha^\beta f(x)\,dx = \int\limits_{\psi(\alpha)}^{\psi(\beta)} f(\varphi(t))\varphi'(t)\,dt.$$

Beweis. Sei F eine Stammfunktion von f. Dann ist $F \circ \varphi$ differenzierbar mit

$$(F \circ \varphi)'(t) = F'(\varphi(t)) \cdot \varphi'(t) = f(\varphi(t)) \cdot \varphi'(t).$$

Aus (2.4) ergibt sich damit

$$\int\limits_a^b f(\varphi(t))\varphi'(t)\,dt = (F \circ \varphi)(t)\Big|_a^b = F(\varphi(b)) - F(\varphi(a)) = F(x)\Big|_{\varphi(a)}^{\varphi(b)} = \int\limits_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$$

Der zweite Teil folgt mit $\psi(\alpha) = a$ und $\psi(\beta) = b$. □

Nützlich sind die

(2.9) Bemerkungen. a) Zur Abkürzung (und als Merkregel) schreibt man oft: $x = \varphi(t)$, $\mathrm{d}x = \varphi'(t)\,\mathrm{d}t$; $x = \alpha \Leftrightarrow t = a$.
b) Gilt die (etwas schärfere) Bedingung $\varphi'(x) \neq 0$ für alle $x \in [a, b]$, so ist eine äquivalente (aber manchmal nützliche) Version die Folgende:

$$\int\limits_\alpha^\beta f(x)\,\mathrm{d}x = \int\limits_{\psi(\alpha)}^{\psi(\beta)} h(\varphi(u))\,\mathrm{d}u,$$

wobei $h(x) := \frac{f(x)}{\psi'(x)}$. Zum Nachweis beachte die Regel für die Ableitung der Umkehrfunktion.
Kurzschreibweise und Merkregel: $u = \psi(x),\ \mathrm{d}u = \psi'(x)\,\mathrm{d}x$.

Der Illustration dienen die

(2.10) Beispiele. a) Wir berechnen $\int_a^b \ln t\, dt$ für $0 < a < b$. Dazu verwenden wir (2.7) mit $f(t) = \ln t, g(t) = t$, $g'(t) = 1$, $f'(t) = \frac{1}{t}$.

$$\int_a^b \ln t \cdot 1\, dt = t \cdot \ln t\Big|_a^b - \int_a^b \frac{1}{t} \cdot t\, dt = t \cdot \ln t\Big|_a^b - \int_a^b dt = (t \ln t - t)\Big|_a^b,$$

$$\int \ln x\, dx = x(\ln x - 1).$$

b) Sei f stetig. Dann gilt für $c \in \mathbb{R}$

$$\int_a^b f(t+c)\, dt = \int_{a+c}^{b+c} f(x)\, dx \quad \text{mit } x = \varphi(t) = t + c,\ dx = dt.$$

Für $n \in \mathbb{Z}$ hat man konkret $\sin(x + n\pi) = (-1)^n \sin x$, also

$$\int_{n\pi}^{(n+1)\pi} \sin x\, \mathrm{d}x = \int_0^\pi \sin(x + n\pi)\, \mathrm{d}x = (-1)^n \int_0^\pi \sin x\, \mathrm{d}x$$
$$= -(-1)^n \cos x\Big|_0^\pi = 2(-1)^n.$$

c) Für stetiges f und $c \in \mathbb{R},\ c \neq 0$ gilt

$$\int_a^b f(ct)\, dt = \frac{1}{c}\int_a^b f(ct)c\, dt = \frac{1}{c}\int_{ac}^{bc} f(x)\, dx \quad \text{mit } x = ct,\ dx = c\, dt.$$

Für $0 \neq n \in \mathbb{Z}$ hat man

$$\int_0^\pi \sin(nx)\, \mathrm{d}x = \frac{1}{n}\int_0^{n\pi} \sin x\, \mathrm{d}x = -\frac{1}{n}\cos x\Big|_0^{n\pi} = \frac{1}{n}\big(1 - (-1)^n\big).$$

d) Für stetiges f hat man mit $x = t^2$, $dx = 2t\, dt$

$$\int_a^b t f(t^2) dt = \frac{1}{2}\int_{a^2}^{b^2} f(x)\, dx.$$

Es gilt

$$\int_0^1 t e^{t^2}\, \mathrm{d}t = \frac{1}{2}\int_0^1 e^x\, \mathrm{d}x = \frac{1}{2}e^x\Big|_0^1 = \frac{1}{2}(e-1).$$

e) Sei $\varphi : [a,b] \to \mathbb{R}$ stetig differenzierbar mit $\varphi(t) \neq 0$ für alle $t \in [a,b]$. Dann gilt

$$\int_a^b \frac{\varphi'(t)}{\varphi(t)}\, dt = \int_{\varphi(a)}^{\varphi(b)} \frac{1}{x}\, dx = \ln|x|\,\Big|_{\varphi(a)}^{\varphi(b)} = \ln|\varphi(t)|\Big|_a^b.$$

Mit $\varphi(t) = \cos t$ erhält man insbesondere

$$\int_a^b \tan t \, dt = -\int_a^b \frac{-\sin t}{\cos t} dt = -\ln(\cos t)\Big|_a^b \quad \text{für } [a,b] \subset \left(\frac{-\pi}{2}, \frac{\pi}{2}\right).$$

f) Für $n \in \mathbb{N}_0$ gilt

$$\int_0^\pi \cos^n \varphi \, d\varphi = \begin{cases} 0, & \text{falls } n \text{ ungerade,} \\ (-1)^{n/2} \frac{1\cdot 3\cdot\ldots\cdot(n-1)}{2\cdot 4\cdot\ldots\cdot n}\pi, & \text{falls } n \text{ gerade.} \end{cases}$$

Diese Formel beweisen wir durch Induktion nach n, wobei die Fälle $n = 0$ und $n = 1$ klar sind. Für $n \geqslant 2$ erhält man mit partieller Integration

$$\int_0^\pi \cos^n \varphi \, d\varphi = \int_0^\pi \cos^{n-1} \varphi \cdot \frac{d}{d\varphi} \sin \varphi \, d\varphi = \cos^{n-1} \varphi \cdot \sin \varphi \Big|_0^\pi + (n-1) \int_0^\pi \cos^{n-2} \varphi \sin^2 \varphi \, d\varphi.$$

Setzt man $\sin^2 \varphi = 1 - \cos^2 \varphi$ ein, so ergibt sich

$$n \int_0^\pi \cos^n \varphi \, d\varphi = -(n-1) \int_0^\pi \cos^{n-2} \varphi \, d\varphi.$$

Die Behauptung folgt nun, wenn man den Wert für $\int_0^\pi \cos^{n-2} \varphi \, d\varphi$ nach Induktionsvoraussetzung einsetzt.

g) Sei $[\alpha, \beta] \subset \mathbb{R}$. Dann

$$\int_\alpha^\beta \frac{\mathrm{d}x}{1+e^x} = \int_{e^\alpha}^{e^\beta} \frac{\mathrm{d}u}{u(1+u)},$$

mit $u = e^x$, $\mathrm{d}u = e^x \, \mathrm{d}x$ (ausführlich $\psi(x) = e^x$, $\psi'(x) = e^x \, \mathrm{d}x$, $h(x) = \frac{1}{e^x(1+e^x)}$). In diesem Beispiel, das den Nutzen von (2.10) illustriert, wurde das Problem auf die Integration einer rationalen Funktion reduziert.

Wesentlich ist das folgende

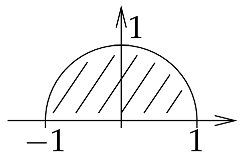
(2.11) Beispiel. Für $-1 \leqslant a < b \leqslant 1$ ist arcsin auf $[a, b]$ streng monoton wachsend. Mit $x = \sin t$, $dx = \cos t \, dt = \sqrt{1 - \sin^2 t} \, dt$, $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ folgt

$$\begin{aligned}
\int_a^b \sqrt{1-x^2} \, dx &= \int_\alpha^\beta \sqrt{1 - \sin^2 t} \cos t \, dt = \int_\alpha^\beta \cos^2 t \, dt \\
&= \int_\alpha^\beta \frac{1}{2} (\cos(2t) + 1) \, dt = \left(\frac{1}{4} \sin(2t) + \frac{1}{2} t\right) \Big|_\alpha^\beta \\
&= \frac{1}{2} (\sin t \cdot \cos t + t) \Big|_\alpha^\beta = \frac{1}{2} (\sin t \cdot \sqrt{1 - \sin^2 t} + t) \Big|_\alpha^\beta \\
&= \frac{1}{2} (x\sqrt{1-x^2} + \arcsin x) \Big|_a^b
\end{aligned}$$

mit $\alpha = \arcsin a$, $\beta = \arcsin b$. Insbesondere folgt

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x)\Big|_{-1}^{1} = \frac{1}{2}(\arcsin 1 - \arcsin(-1)) = \frac{\pi}{2}.$$

Damit haben wir gezeigt, dass auch bei unserer Definition von π die Fläche des Einheitskreises ganau π ist. Mit einer Substitution folgt für $r > 0$



$$\int_{-r}^{r} \sqrt{r^2-x^2}\,dx = \int_{-r}^{r} \sqrt{1-(x/r)^2}\,r\,dx = r^2 \int_{-1}^{1} \sqrt{1-t^2}\,dt = \frac{1}{2}r^2\pi.$$

Also ist die Fläche eines Kreises mit Radius r genau $r^2\pi$. Der Umfang des Kreises wird später berechnet (s. Skript zu Analysis III).

Mit partieller Integration kann man eine weitere Formel für das Restglied im Satz von TAYLOR VII(3.5) beweisen.

(2.12) Satz. *Sei I ein Intervall und $f : I \to \mathbb{R}$ sei $(n+1)$-mal stetig differenzierbar, $n \in \mathbb{N}_0$. Dann gilt für alle $a, x \in I$:*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_{n+1,a}(x)$$

mit

$$R_{n+1,a}(x) = \frac{1}{n!}\int_{a}^{x} (x-t)^n f^{(n+1)}(t)\,dt.$$

Beweis. Das Integral existiert nach (1.17), da der Integrand stetig ist. Wir verwenden eine Induktion nach n. Für $n = 0$ gilt nach (2.4)

$$f(a) + \int_{a}^{x} f'(t)\,dt = f(a) + \left[f(t)\Big|_{a}^{x}\right] = f(x).$$

Beim Induktionsschritt $n-1 \mapsto n$ verwenden wir

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_{n,a}(x)$$

und erhalten mit (2.7)

$$\begin{aligned} R_{n,a}(x) &= \frac{1}{(n-1)!}\int\limits_a^x (x-t)^{n-1} f^{(n)}(t)dt \\ &= -\int\limits_a^x f^{(n)}(t)\frac{d}{dt}\left[\frac{(x-t)^n}{n!}\right] dt \\ &= -f^{(n)}(t)\cdot\frac{(x-t)^n}{n!}\Big|_{t=a}^{t=x} + \frac{1}{n!}\int\limits_a^x (x-t)^n f^{(n+1)}(t)\, dt \\ &= \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1,a}(x) \end{aligned}$$

mit obiger Definition des Restgliedes. Daraus folgt die Behauptung. □

Wir deuten kurz einige physikalische Interpretationen an.

(2.13) Beispiele. a) Bezeichnet $v(t)$ die Geschwindigkeit eines Objekts zum Zeitpunkt t und $s(t)$ die zurückgelegte Strecke, so hat man die Beziehung

$$v(t) = \dot{s}(t).$$

Ist $v(t)$ stetig, so folgt mit (2.4)

$$\int\limits_{t_0}^{t_1} v(t)\, dt = s(t_1) - s(t_0).$$

Also erscheint die zurückgelegte Strecke in diesem Fall als Fläche unter dem Graphen der Funktion $v(t)$.

b) Wirkt eine Kraft $K(x)$ entlang einer Strecke $[a, b]$, so ist

$$\int\limits_a^b K(x)\, dx$$

die von der Kraft geleistete *Arbeit*. Nach dem NEWTONschen Gravitationsgesetz ist damit die Arbeit, die geleistet werden muss, um eine Rakete gegen die Anziehungskraft der Erde auf eine Höhe h über dem Erdboden zu bringen, gerade

$$A(h) = \int\limits_R^{R+h} G\,\frac{mM}{x^2}\, dx = GmM\frac{-1}{x}\Big|_R^{R+h} = GmM\left(\frac{1}{R} - \frac{1}{R+h}\right),$$

wobei R der Erdradius, M die Erdmasse, m die Raketenmasse und G die Gravitationskonstante ist.

Dabei besagt das NEWTON*sche Gravitationsgesetz*, dass die Anziehungskraft K zwischen zwei Körpern gegeben wird durch

$$K = G\cdot\frac{m_1 m_2}{r^2},$$

wobei m_1, m_2 die beiden Massen und r der Abstand der Körper ist.

§3. Uneigentliche Integrale

Es gibt zwei Typen von uneigentlichen Integralen, die den RIEMANNschen Integrationsbegriff aus §1 verallgemeinern. Einmal ist das Integrationsintervall unbeschränkt, beim anderen Typ ist der Integrand unbeschränkt.

(3.1) Definition. a) Eine Funktion $f : [a,\infty) \to \mathbb{R}$ heißt *uneigentlich* RIEMANN-*integrierbar,* wenn $f\big|_{[a,b]}$ für jedes $b > a$ integrierbar ist und $\lim\limits_{b\to\infty} \int_a^b f(x)\,dx$ (in $\mathbb{R}$) existiert. Dann setzt man

$$\int_a^\infty f(x)\,dx := \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

b) Eine Funktion $f : (-\infty,b] \to \mathbb{R}$ heißt *uneigentlich* RIEMANN-integrierbar, wenn $f\big|_{[a,b]}$ für jedes $a < b$ integrierbar ist und $\lim\limits_{a\to-\infty} \int_a^b f(x)\,dx$ (in $\mathbb{R}$) existiert. Dann setzt man

$$\int_{-\infty}^b f(x)\,dx := \lim_{a\to-\infty} \int_a^b f(x)\,dx.$$

c) Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt *uneigentlich* RIEMANN-*integrierbar,* wenn es ein $c \in \mathbb{R}$ gibt, so dass $f\big|_{(-\infty,c]}$ und $f\big|_{[c,\infty)}$ uneigentlich RIEMANN-integrierbar sind. Dann setzt man

$$\int_{-\infty}^\infty f(x)\,dx := \int_{-\infty}^c f(x)\,dx + \int_c^\infty f(x)\,dx.$$

In den Fällen a) - c) heißen $\int_a^\infty f(x)\,dx$, $\int_{-\infty}^b f(x)\,dx$ bzw. $\int_{-\infty}^\infty f(x)\,dx$ *uneigentliches* RIEMANN-*Integral.*
d) Eine Funktion f heißt *uneigentlich absolut integrierbar* in den Fällen a) - c), wenn $|f|$ uneigentlich RIEMANN-integrierbar ist.

Aufgrund von (1.18) ist die Bedingung in c) unabhängig von der Wahl von $c \in \mathbb{R}$.

(3.2) Beispiele. a) Ist $\alpha \in \mathbb{R}$, so existiert $\int_1^\infty \frac{1}{x^\alpha}\,dx$ genau dann, wenn $\alpha > 1$. Ist nämlich $\alpha \neq 1$, so gilt für $b > 1$

$$\int_1^b x^{-\alpha}\,dx = \frac{x^{1-\alpha}}{1-\alpha}\bigg|_1^b = \frac{b^{1-\alpha}}{1-\alpha} - \frac{1}{1-\alpha} \xrightarrow[b\to\infty]{} \begin{cases} \frac{1}{\alpha-1}, & \text{falls } \alpha > 1, \\ \infty, & \text{falls } \alpha < 1. \end{cases}$$

Also gilt

$$\int_1^\infty x^{-\alpha}\,dx = \frac{1}{\alpha-1} \quad \text{für } \alpha > 1.$$

Darüber hinaus hat man

$$\int_1^b \frac{1}{x}\,dx = \ln x\Big|_1^b = \ln b \xrightarrow[b\to\infty]{} \infty.$$

b) $\int_0^\infty e^{-x}\,dx = \lim\limits_{b\to\infty} \int_0^b e^{-x}\,dx = \lim\limits_{b\to\infty} (-e^{-x})\big|_0^b = \lim\limits_{b\to\infty} (-e^{-b}+1) = 1.$

c) $\int_a^\infty \sin x\,dx = \lim\limits_{b\to\infty} \int_a^b \sin x\,dx = \lim\limits_{b\to\infty} (-\cos x)\big|_a^b = \lim\limits_{b\to\infty} (\cos a - \cos b)$ existiert nicht. Also ist der Sinus nicht uneigentlich RIEMANN-integrierbar über $[a,\infty)$.

d) Es gilt $\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx = \pi$, denn

$$\int_0^\infty \frac{1}{1+x^2}\,dx = \lim_{b\to\infty} \arctan x\Big|_0^b = \lim_{b\to\infty} \arctan b = \frac{\pi}{2} = \int_{-\infty}^0 \frac{1}{1+x^2}\,dx.$$

e) Die Schlusskette

$$\int_{-\infty}^\infty x\,\mathrm{d}x = \lim_{R\to\infty} \int_{-R}^R x\,\mathrm{d}x = \lim_{R\to\infty} \tfrac{1}{2}x^2\Big|_{-R}^R = \lim_{R\to\infty} 0 = 0$$

ist offensichtlich illegal: Das erste Gleichheitszeichen folgt nicht aus (3.1). Da $\int_0^\infty x\,\mathrm{d}x$ nicht (in $\mathbb{R}$) existiert, existiert auch $\int_{-\infty}^\infty x\,\mathrm{d}x$ nicht.

Wir leiten ein Kriterium für die Existenz uneigentlicher RIEMANN-Integrale her.

(3.3) Satz. *(CAUCHY-Kriterium)*
Sei $f : [a,\infty) \to \mathbb{R}$ eine Funktion, so dass $f\big|_{[a,b]}$ für jedes $b > a$ integrierbar ist.
a) $\int_a^\infty f(x)\,dx$ existiert genau dann, wenn zu jedem $\varepsilon > 0$ ein $B > 0$ existiert, so dass

$$\left|\int_b^c f(x)\,dx\right| < \varepsilon \quad \text{für alle } c > b > B.$$

b) Gilt $f(x) \geqslant 0$ für alle $x \geqslant a$, so existiert $\int_a^\infty f(x)\,dx$ genau dann, wenn ein $M > 0$ existiert, so dass

$$\int_a^b f(x)\,dx \leqslant M \quad \text{für alle } b > a.$$

Beweis. Wir betrachten

$$F : [a,\infty) \to \mathbb{R}, \quad x \mapsto \int_a^x f(t)\,dt.$$

a) Das uneigentliche Integral $\int_a^\infty f(t)\,dt$ existiert nach (3.1) genau dann, wenn $\lim\limits_{x\to\infty} F(x)$ in $\mathbb{R}$ existiert. Nach dem CAUCHY-Kriterium bedeutet dies, dass es zu jedem $\varepsilon > 0$ ein $x_0 > a$ gibt mit $|F(x_2) - F(x_1)| < \varepsilon$ für alle $x_2 > x_1 > x_0$. Aus (1.18) erhält man aber

$$F(x_2) - F(x_1) = \int\limits_{x_1}^{x_2} f(t)\,dt$$

und somit die Behauptung.
b) Wegen $f \geqslant 0$ ist F nach (1.18) monoton wachsend.
"$\Rightarrow$" Es gilt für alle $x \geqslant a$

$$F(x) \leqslant \lim_{y\to\infty} F(y) = \int\limits_a^\infty f(t)\,dt =: M.$$

"$\Leftarrow$" Da F nach Voraussetzung auch beschränkt ist, liefert III(2.2) für jede Folge $(x_n)_{n\geqslant 1}$ mit $x_n \to \infty$ die Existenz von

$$\sup\Big\{\int\limits_a^R f(t)dt; R > a\Big\} = \lim_{n\to\infty} F(x_n) = \int\limits_a^\infty f(t)\,dt.$$

□

Die Bedingung $f \geqslant 0$ ist natürlich ganz wesentlich, wie Beispiel (3.2) c) zeigt.

Man erhält analoge Aussagen natürlich für $\int_{-\infty}^b f(t)\,dt$ mit der Substitution $t = -s$ vermöge

$$\int\limits_{-\infty}^b f(t)\,dt = \int\limits_{-b}^\infty f(-s)\,ds.$$

Wie bei der Konvergenz von Reihen gelten Vergleichssätze.

(3.4) Satz. *(Vergleichskriterium)*
Gegeben seien Funktionen $f, g : [a,\infty) \to \mathbb{R}$, so dass $f|_{[a,b]}$ und $g|_{[a,b]}$ für jedes $b > a$ RIEMANN-integrierbar sind.
a) Gilt $|f(x)| \leqslant g(x)$ für alle $x \geqslant a$ und existiert $\int_a^\infty g(x)\,dx$, so sind f und $|f|$ uneigentlich integrierbar und es gilt

$$\left|\int\limits_a^\infty f(t)\,dt\right| \leqslant \int\limits_a^\infty |f(t)|\,dt \leqslant \int\limits_a^\infty g(t)\,dt.$$

b) Gilt $|f(x)| \geqslant g(x) \geqslant 0$ und ist g nicht uneigentlich RIEMANN-integrierbar, so ist auch $|f|$ nicht uneigentlich RIEMANN-integrierbar, d. h.

$$\int\limits_a^\infty |f(t)|\,dt = \int\limits_a^\infty g(t)\,dt = \infty.$$

c) Sei $f(x) \geqslant 0$ und $g(x) > 0$ für alle $x \geqslant a$ und es existiere

$$L := \lim_{x\to\infty} \frac{f(x)}{g(x)} \in \mathbb{R}_+.$$

(i) Gilt $L > 0$*, so existiert* $\int_a^\infty f(x)\,dx$ *genau dann, wenn* $\int_a^\infty g(x)\,dx$ *existiert.*

(ii) Gilt $L = 0$ *und existiert* $\int_a^\infty g(x)\,dx$*, so existiert auch* $\int_a^\infty f(x)\,dx$*.*

Beweis. a) Mit f ist auch $|f|$ nach (1.20) auf $[a,b]$ integrierbar. Zu $\varepsilon > 0$ wählen wir nach (3.3) ein $B > 0$ mit

$$\left|\int_b^c g(t)\,dt\right| \leqslant \varepsilon \quad \text{für alle } c > b > B.$$

Wegen $0 \leqslant |f(t)| \leqslant g(t)$ folgt daraus mit (1.20) und (1.18)

$$\left|\int_b^c f(t)\,dt\right| \leqslant \int_b^c |f(t)|\,dt \leqslant \int_b^c g(t)\,dt \leqslant \varepsilon \quad \text{für alle } c > b > B.$$

Also sind f und $|f|$ nach (3.3) uneigentlich RIEMANN-integrierbar mit der Eigenschaft

$$\begin{aligned}\left|\int_a^\infty f(t)\,dt\right| &= \lim_{x\to\infty}\left|\int_a^x f(t)\,dt\right| \leqslant \lim_{x\to\infty}\int_a^x |f(t)|\,dt = \int_a^\infty |f(t)|\,dt \\ &\leqslant \lim_{x\to\infty}\int_a^x g(t)\,dt = \int_a^\infty g(t)\,dt.\end{aligned}$$

b) Aus (3.3) folgt, dass ein $\varepsilon^* > 0$ existiert, so dass es zu jedem $B > a$ auch $c > b > B$ gibt mit $|\int_b^c g(t)\,dt| = \int_b^c g(t)\,dt \geqslant \varepsilon^*$. Aus (1.17) erhält man

$$\int_b^c |f(t)|\,dt \geqslant \int_b^c g(t)\,dt \geqslant \varepsilon^*.$$

Also ist $|f|$ nach (3.3) nicht uneigentlich RIEMANN-integrierbar.
c) (i) Weil der Limes existiert, gibt es ein $C > a$, so dass

$$\frac{1}{2}L \leqslant \frac{f(x)}{g(x)} \leqslant 2L, \quad \text{also } \frac{1}{2}L\,g(x) \leqslant f(x) \leqslant 2L\,g(x) \quad \text{für alle } x \geqslant C.$$

Jetzt folgt die Behauptung aus b).
(ii) Es gibt ein $C > a$, so dass

$$0 \leqslant \frac{f(x)}{g(x)} \leqslant 1, \quad \text{also } 0 \leqslant f(x) \leqslant g(x) \quad \text{für alle } x \geqslant C.$$

Also erhält man die Behauptung wieder aus b). □

Wendet man a) auf $g = |f|$ an, so folgt das

(3.5) Korollar. *Sei $f : [a, \infty) \to \mathbb{R}$ über jedem Intervall $[a, b], b > a$,* RIEMANN*-integrierbar. Ist f uneigentlich absolut integrierbar, so ist f auch uneigentlich integrierbar.*

Wir diskutieren einige

(3.6) Beispiele. a) Für $\alpha \in \mathbb{R}$, $\alpha > 1$ sind

$$\int_1^\infty \frac{\sin x}{x^\alpha}\,dx \quad \text{und} \quad \int_1^\infty \frac{\cos x}{x^\alpha}\,dx$$

uneigentlich absolut integrierbar. Es gilt nämlich

$$\left|\frac{\sin x}{x^\alpha}\right| \leqslant x^{-\alpha} \quad \text{und} \quad \left|\frac{\cos x}{x^\alpha}\right| \leqslant x^{-\alpha} \quad \text{für alle } x \geqslant 1 \quad \text{und} \quad \int_1^\infty x^{-\alpha}\,dx = \frac{1}{\alpha - 1}$$

nach (3.2). Nun folgt die Behauptung aus (3.4).
b) Das uneigentliche RIEMANN-Integral $\int_1^\infty \frac{\sin x}{x}\,dx$ existiert. Mit partieller Integration für $f(x) = \frac{1}{x}$, $f'(x) = \frac{-1}{x^2}$, $g'(x) = \sin x$, $g(x) = -\cos x$ erhält man

$$\int_1^b \frac{\sin x}{x}\,dx = -\frac{\cos x}{x}\Big|_1^b - \int_1^b \frac{\cos x}{x^2}\,dx \xrightarrow[b\to\infty]{} \cos 1 - \int_1^\infty \frac{\cos x}{x^2}\,dx,$$

wenn man a) verwendet.
Andererseits gilt aber $\int_1^\infty \left|\frac{\sin x}{x}\right|\,dx = \infty$, denn für $n \in \mathbb{N}$ und $n\pi \leqslant x \leqslant (n+1)\pi$ ist

$$\frac{1}{x} \geqslant \frac{1}{(n+1)\pi}, \quad \text{also} \quad \left|\frac{\sin x}{x}\right| \geqslant \frac{|\sin x|}{(n+1)\pi}.$$

Es folgt für $n \in \mathbb{N}$:

$$\int_{n\pi}^{(n+1)\pi} \left|\frac{\sin x}{x}\right| dx \geqslant \frac{1}{(n+1)\pi} \int_{n\pi}^{(n+1)\pi} |\sin x|\,dx = \frac{1}{(n+1)\pi} \int_0^\pi \sin x\,dx = \frac{2}{(n+1)\pi},$$

$$\int_1^{N\pi} \left|\frac{\sin x}{x}\right| dx \geqslant \sum_{n=1}^{N-1} \int_{n\pi}^{(n+1)\pi} \left|\frac{\sin x}{x}\right| dx \geqslant \sum_{n=1}^{N-1} \frac{2}{(n+1)\pi} \xrightarrow[N\to\infty]{} \infty.$$

c) $\int_{-\infty}^\infty e^{-x^2}$ existiert, denn es gilt $e^{-x^2} \leqslant e^{-|x|}$ für $|x| \geqslant 1$. Nun verwende man (3.2) b).

Wir beweisen eine weitere wichtige Anwendung.

(3.7) Satz. *(Integral-Vergleichskriterium für Reihen)*
Sei $m \in \mathbb{Z}$ und $f : [m, \infty) \to \mathbb{R}$ eine monoton fallende Funktion. In diesem Fall existiert $\int_m^\infty f(x)\,dx$ genau dann, wenn die Reihe $\sum_{k=m}^\infty f(k)$ konvergiert.

Beweis. Als monotone Funktion ist f nach (1.16) über jedem Intervall $[a,b]$ mit $m \leqslant a < b$ integrierbar. Sei $f \geqslant 0$, da andernfalls weder $\int_m^\infty f(x)\,dx$ noch $\sum_{k=m}^\infty f(k)$ existieren, weil f monoton fällt. Aus der Monotonie erhält man für $b > m$

$$f(k+1) \leqslant \int_k^{k+1} f(x)\,dx \leqslant f(k), \quad k \geqslant m,$$

also

$$\sum_{k=m+1}^{[b]} f(k) \leqslant \int_m^{[b]} f(x)\,dx \leqslant \int_m^{b} f(x)\,dx \leqslant \int_m^{[b+1]} f(x)\,dx \leqslant \sum_{k=m}^{[b]} f(k).$$

Aus dem Monotonie-Kriterium für unendliche Reihen in IV(1.10) und für Integrale in (3.4) folgt dann die Behauptung. □

Mit (3.2) erhält man als direkte Konsequenz das

(3.8) Korollar. *Ist $\alpha \in \mathbb{R}$, so konvergiert die Reihe $\sum_{n=1}^\infty n^{-\alpha}$ genau dann, wenn $\alpha > 1$.*

Der Grenzwert wird mit $\zeta(\alpha)$ bezeichnet. $\zeta : (1,\infty) \to \mathbb{R}$ heißt RIEMANN*sche Zetafunktion*.

Wir wissen, dass die harmonische Reihe $\sum_{n=1}^\infty \frac{1}{n}$ divergiert. Als Verschärfung dieser Aussage diskutieren wir das

(3.9) Beispiel. a) Die Reihe $\sum_{n=2}^\infty \frac{1}{n \ln n}$ divergiert, denn $f : [2,\infty) \to \mathbb{R}$, $x \mapsto \frac{1}{x \ln x}$, ist monoton fallend und es gilt

$$\int_2^b \frac{1}{x \ln x}\,dx = \ln(\ln x)\Big|_2^b = \ln(\ln b) - \ln(\ln 2) \xrightarrow[b\to\infty]{} \infty.$$

b) Für $\alpha > 1$ ist die Reihe $\sum_{n=2}^\infty \frac{1}{n\cdot(\ln n)^\alpha}$ konvergent, denn $f : [2,\infty] \to \mathbb{R}$, $x \mapsto \frac{1}{x\cdot(\ln x)^\alpha}$, ist monoton fallend und es gilt

$$\begin{aligned}\int_2^b \frac{1}{x\cdot(\ln x)^\alpha}\,dx &= \frac{1}{1-\alpha}(\ln x)^{1-\alpha}\Big|_2^b \\ &= \frac{1}{1-\alpha}\left[(\ln b)^{1-\alpha} - (\ln 2)^{1-\alpha}\right] \xrightarrow[b\to\infty]{} \frac{1}{(\alpha-1)\cdot(\ln 2)^{\alpha-1}}.\end{aligned}$$

Wir beschäftigen uns nun mit dem zweiten Typ, den unbeschränkten Integranden.

(3.10) Definition. a) Sei $f : (a,b]$ eine Funktion, die für jedes $a < \alpha < b$ über $[\alpha,b]$ RIEMANN-integrierbar ist. Falls $\lim\limits_{\alpha\downarrow a} \int_\alpha^b f(x)\,dx$ (in $\mathbb{R}$) existiert, nennt man f über $(a,b]$ *uneigentlich* RIEMANN-*integrierbar* und setzt

$$\int_a^b f(x)\,dx := \lim_{\alpha\downarrow a} \int_\alpha^b f(x)\,dx.$$

b) Sei $f : [a,b) \to \mathbb{R}$ eine Funktion, die für jedes $a < \beta < b$ über $[a,\beta]$ RIEMANN-integrierbar ist. Falls $\lim\limits_{\beta\uparrow b} \int_a^\beta f(x)\,dx$ (in $\mathbb{R}$) existiert, nennt man f über $[a,b)$ *uneigentlich* RIEMANN-*integrierbar* und setzt

$$\int\limits_a^b f(x)\,dx := \lim_{\beta\uparrow b} \int\limits_a^\beta f(x)\,dx.$$

c) Ist f auf $(a,c]$ und $[c,b)$ mit $a < c < b$ uneigentlich RIEMANN-integrierbar, so definiert man

$$\int\limits_a^b f(x)\,dx := \int\limits_a^c f(x)\,dx + \int\limits_c^b f(x)\,dx$$

und nennt f über (a,b) *uneigentlich* RIEMANN-*integrierbar*. Die Fälle $a = -\infty$, $b = \infty$ sind entsprechend definiert.
Man nennt $\int_a^b f(x)\,dx$ in den Fällen a) - c) das *uneigentliche* RIEMANN-*Integral*.
d) In den Fällen a) - c) heißt f *uneigentlich absolut integrierbar*, wenn $|f|$ uneigentlich integrierbar ist.

Den Zusammenhang mit früheren Bezeichnungen erläutert die

(3.11) Bemerkung. Ist $f : [a,b] \to \mathbb{R}$ (eigentlich) RIEMANN-integrierbar, so ist das Integral als Funktion der oberen oder unteren Grenze nach (2.2) stetig. Also stimmen dann das eigentliche und uneigentliche RIEMANN-Integral überein. Das rechtfertigt die gleichen Bezeichnungen.

Zu diesem neuen Begriff rechnen wir zunächst einige

(3.12) Beispiele. a) Für welche $\alpha \in \mathbb{R}$ existiert $\int_0^1 x^{-\alpha}\,dx$? Für $\varepsilon > 0$ gilt

$$\int\limits_\varepsilon^1 \frac{1}{x^\alpha}\,dx = \begin{cases} \frac{1}{1-\alpha} x^{1-\alpha}\big|_\varepsilon^1, \\ \ln x\big|_\varepsilon^1, \end{cases} = \begin{cases} \frac{1}{1-\alpha}(1-\varepsilon^{1-\alpha}), & \text{falls } \alpha \neq 1, \\ -\ln\varepsilon, & \text{falls } \alpha = 1, \end{cases} \underset{\varepsilon\downarrow 0}{\longrightarrow} \begin{cases} \frac{1}{1-\alpha}, & \text{falls } \alpha < 1, \\ \infty, & \text{falls } \alpha \geqslant 1. \end{cases}$$

Also existiert das uneigentliche Integral für $\alpha < 1$ mit

$$\int\limits_0^1 x^{-\alpha}\,dx = \frac{1}{1-\alpha}.$$

b) Es gilt

$$\int\limits_0^1 \frac{1}{\sqrt{1-x^2}}\,dx = \lim_{\beta\uparrow 1} \arcsin x\big|_0^\beta = \lim_{\beta\uparrow 1} \arcsin\beta = \arcsin 1 = \frac{\pi}{2},$$

$$\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}}\,dx = \int_{-1}^{0} \frac{dx}{\sqrt{1-x^2}} + \int_{0}^{1} \frac{dx}{\sqrt{1-x^2}} = 2\int_{0}^{1} \frac{dx}{\sqrt{1-x^2}} = \pi.$$

c) $\int_0^1 \sin\left(\frac{1}{x}\right)\,dx$ existiert, denn mit $t = \frac{1}{x}$, $dt = \frac{-1}{x^2}\,dx$ folgt

$$\int_{\varepsilon}^{1} \sin\left(\frac{1}{x}\right)\,dx = \int_{1/\varepsilon}^{1} (\sin t)\left(\frac{-1}{t^2}\right)\,dt = \int_{1}^{1/\varepsilon} \frac{\sin t}{t^2}\,dt \underset{\varepsilon\downarrow 0}{\longrightarrow} \int_{1}^{\infty} \frac{\sin t}{t^2}\,dt.$$

Das letzte Integral existiert nach (3.6).
d) $\int_{-1}^1 \frac{1}{x}\,dx := \int_{-1}^0 \frac{1}{x}\,dx + \int_0^1 \frac{1}{x}\,dx$ existiert nicht, da $\int_{-1}^0 \frac{1}{x}\,dx$ und $\int_0^1 \frac{1}{x}\,dx$ beide nicht existieren. Es gilt aber mit der Substitution $t = -x$

$$\int_{-1}^{-\varepsilon} \frac{1}{x}\,dx + \int_{\varepsilon}^{1} \frac{1}{x}\,dx = -\int_{1}^{\varepsilon} \frac{1}{-t}\,dt + \int_{\varepsilon}^{1} \frac{1}{x}\,dx = 0$$

für alle $\varepsilon > 0$.
e) Das Integral $\int_0^1 \frac{1}{x\ln x}\,\mathrm{d}x$ ist an der Unter- und Obergrenze uneigentlich, da dort der Nenner Nullstellen hat. Für $0 < \alpha < \frac{1}{2} < \beta < 1$ gilt

$$\int_{\alpha}^{1/2} \frac{1}{x\ln x}\,\mathrm{d}x = \ln(-\ln x)\Big|_{\alpha}^{1/2} = \ln(\ln 2) - \ln(-\ln\alpha) \underset{\alpha\downarrow 0}{\longrightarrow} -\infty,$$

$$\int_{1/2}^{\beta} \frac{1}{x\ln x}\,\mathrm{d}x = \ln(-\ln x)\Big|_{1/2}^{\beta} = \ln(-\ln\beta) - \ln(\ln 2) \underset{\beta\uparrow 1}{\longrightarrow} -\infty.$$

Also existiert das uneigentliche Integral nicht.

Mit einer Substitution wie in c) oder wörtlich wie zuvor beweist man den

(3.13) Satz. (CAUCHY-*Kriterium/Vergleichskriterium)*
Seien $f, g : (a,b] \to \mathbb{R}$ Funktionen, die für jedes $a < \alpha < b$ auf $[\alpha, b]$ RIEMANN*-integrierbar sind. Dann gilt*
a) $\int_a^b f(x)\,dx$ existiert genau dann, wenn zu jedem $\varepsilon > 0$ ein $0 < \delta < b - a$ existiert, so dass

$$\left|\int_{\alpha}^{\beta} f(x)\,dx\right| < \varepsilon \quad \text{für alle } a < \alpha < \beta \leqslant a + \delta.$$

b) Gilt $f(x) \geqslant 0$ für alle $x \in (a,b]$, so existiert $\int_a^b f(x)\,dx$ genau dann, wenn es ein $M > 0$ gibt mit der Eigenschaft

$$\int_{\alpha}^{b} f(x)\,dx \leqslant M \quad \text{für alle } a < \alpha < b.$$

c) Gilt $|f(x)| \leqslant g(x)$ für alle $x \in (a,b]$ und existiert $\int_a^b g(x)\,dx$, so sind f und $|f|$ uneigentlich integrierbar und es gilt

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx \leqslant \int_a^b g(x)\,dx.$$

Ist f uneigentlich absolut integrierbar, so ist f auch uneigentlich integrierbar.
d) Gilt $|f(x)| \geqslant g(x) \geqslant 0$ für alle $x \in (a,b]$ und ist g nicht uneigentlich integrierbar, so ist f nicht uneigentlich absolut integrierbar.
e) Sei $f(x) \geqslant 0$ und $g(x) \geqslant 0$ für alle $x \in (a,b]$ und es existiere $L := \lim\limits_{x \downarrow a} \frac{f(x)}{g(x)} \in \mathbb{R}_+$.
(i) Gilt $L > 0$, so existiert $\int_a^b f(x)\,dx$ genau dann, wenn $\int_a^b g(x)\,dx$ existiert.
(ii) Gilt $L = 0$ und existiert $\int_a^b g(x)\,dx$, so existiert auch $\int_a^b f(x)\,dx$.

Der Einordnung dient die

(3.14) Bemerkung. Die Definitionen (3.1) und (3.10) umfassen nicht alle möglichen (und nicht alle relevanten) Typen uneigentlicher Integrale. Zum Beispiel hat man auch (offensichtliche) vernünftige Definitionen für

$$\int_{-1}^{1} \frac{\mathrm{d}x}{\sqrt{|x|}} \quad \text{und} \quad \int_0^\infty \ln x \cdot e^{-x}\,\mathrm{d}x.$$

Aber eine komplette Auflistung würde hier zu weit führen.

§4. Die Gamma-Funktion

Wir führen einen wichtigen Beispieltyp von uneigentlichen Integralen ein.

(4.1) Satz. *Das uneigentliche Integral* $\int_0^\infty t^{x-1}e^{-t}\,dt$ *konvergiert für alle* $x > 0$.

Beweis. Sei $x > 0$. Da die Exponentialfunktion stärker wächst als jede Potenz (vgl. V(4.1)), existiert ein $c > 0$, so dass

$$t^{x-1}e^{-t} \leqslant \frac{c}{t^2} \quad \text{für alle } t \geqslant 1.$$

Da $\int_1^\infty \frac{1}{t^2}\,dt = 1$ nach (3.2) gilt, konvergiert $\int_1^\infty t^{x-1}e^{-t}\,dt$ nach (3.4). Es gilt $t^{x-1}e^{-t} \leqslant t^{x-1}$ für alle $t > 0$. Da $\int_0^1 t^{x-1}\,dt = \frac{1}{x}$ nach (3.12) gilt, konvergiert $\int_0^1 t^{x-1}e^{-t}\,dt$ nach (3.13). Also existiert $\int_0^\infty t^{x-1}e^{-t}\,dt$ für alle $x > 0$. □

Damit kommen wir zu der

(4.2) Definition. Die [nach (4.1)] durch

$$\Gamma : \mathbb{R}_+^* \to \mathbb{R}_+^*,\ x \mapsto \Gamma(x) := \int_0^\infty t^{x-1}e^{-t}\,dt,$$

definierte Funktion heißt *Gamma-Funktion*.

Zunächst schauen wir uns spezielle Werte und eine Funktionalgleichung an.

(4.3) Satz. *Für alle* $x > 0$ *gilt* $\Gamma(x+1) = x \cdot \Gamma(x)$ *sowie*

$$\Gamma(n+1) = n! \quad \text{für alle } n \in \mathbb{N}_0.$$

Beweis. Für $0 < a < b$ gilt mit partieller Integration nach (2.7)

$$\int_a^b t^x e^{-t}\,dt = -t^x e^{-t}\Big|_{t=a}^{t=b} + \int_a^b x t^{x-1}e^{-t}\,dt.$$

Die Grenzübergänge $a \downarrow 0$ und $\lim_{a\downarrow 0} a^x e^{-a} = 0$ sowie $b \to \infty$ und $\lim_{b\to\infty} b^x e^{-b} = 0$ liefern

$$\Gamma(x+1) = \int_0^\infty t^x\, e^{-t}\,dt = x\int_0^\infty t^{x-1}\, e^{-t}\,dt = x \cdot \Gamma(x) \quad \text{für } x > 0.$$

Aus (3.2) folgt $\Gamma(1) = \int_0^\infty e^{-t}\,dt = 1 = 0!$. Eine Induktion liefert

$$\Gamma(n+2) = (n+1)\cdot\Gamma(n+1) = (n+1)\cdot n! = (n+1)!.$$ □

Man könnte fragen, ob Γ durch seine Funktionalgleichung und die Bedingung $\Gamma(1) = 1$ schon eindeutig bestimmt ist. Das ist aber sicher nicht der Fall, da auch

$$G(x) := \Gamma(x) \cdot (1 + \sin(2\pi x))$$

diese Bedingungen erfüllt. Jedoch reicht schon eine relativ schwache Zusatzbedingung für Eindeutigkeit.

Dazu ist es sinnvoll, einen neuen Begriff einzuführen.

(4.4) Definition. Ist I ein Intervall, so heißt eine Funktion $f : I \to \mathbb{R}_+^*$ *logarithmisch konvex*, wenn $\ln \circ f : I \to \mathbb{R}$ konvex ist. Wegen $f = \exp \circ (\ln \circ f)$ ist das genau dann der Fall, wenn

$$f(x_\lambda) \leqslant f(x_0)^{1-\lambda} \cdot f(x_1)^\lambda \quad \text{für } x_\lambda := (1-\lambda)x_0 + \lambda x_1,\ 0 \leqslant \lambda \leqslant 1,\ x_0,\ x_1 \in I.$$

Als erstes Beispiel notieren wir den

(4.5) Satz. *Die Gamma-Funktion $\Gamma : \mathbb{R}_+^* \to \mathbb{R}_+^*$ ist logarithmisch konvex.*

Beweis. Seien $x, y \in \mathbb{R}_+^*$, $0 < \lambda < 1$, $p := \frac{1}{\lambda}$, $q := \frac{1}{1-\lambda} > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. Wir wenden nun auf $f(t) := t^{(x-1)/p} e^{-t/p}$ und $g(t) := t^{(y-1)/q} \cdot e^{-t/q}$ mit $x > 0$, $y > 0$ die HÖLDERsche Ungleichung (1.20) an. Dann folgt für $0 < a < b$

$$\begin{aligned} \int_a^b t^{(x/p)+(y/q)-1} e^{-t}\, dt &= \int_a^b f(t) g(t)\, dt \leqslant \left(\int_a^b f(t)^p\, dt \right)^{1/p} \cdot \left(\int_a^b g(t)^q\, dt \right)^{1/q} \\ &= \left(\int_a^b t^{x-1} e^{-t}\, dt \right)^{1/p} \cdot \left(\int_a^b t^{y-1} e^{-t}\, dt \right)^{1/q}. \end{aligned}$$

Der Grenzübergang $a \downarrow 0$ und $b \to \infty$ liefert

$$\Gamma\left(\frac{x}{p} + \frac{y}{q} \right) \leqslant \Gamma(x)^{1/p} \cdot \Gamma(y)^{1/q}.$$

Also ist Γ logarithmisch konvex. □

Es gilt auch eine Umkehrung dieser Ergebnisse, die die Gamma-Funktion charakterisieren. Die Ergebnisse stammen von Harald BOHR (1887 - 1951).

(4.6) Satz von BOHR. *Sei $F : \mathbb{R}_+^* \to \mathbb{R}_+^*$ eine Funktion mit den Eigenschaften:*

(i) *F ist logarithmisch konvex.*

(ii) *$F(x+1) = x \cdot F(x)$ für alle $x > 0$.*

(iii) *$F(1) = 1$.*

Dann gilt $F(x) = \Gamma(x)$ für alle $x > 0$.

Beweis. Nach (4.3) und (4.5) erfüllt Γ die Bedingungen (i)-(iii). Es genügt also zu zeigen, dass eine Funktion F durch (i)-(iii) eindeutig bestimmt ist. Aus (ii) folgt

$$(*)\qquad F(x+n) = (x+n-1)\cdot\ldots\cdot x\cdot F(x) = F(x)\cdot\prod_{k=0}^{n-1}(x+k) \quad \text{für alle } x>0,\ n\in\mathbb{N}$$

sowie mit (iii) auch $F(n+1) = n!$ durch eine Induktion nach n. Also genügt es zu zeigen, dass F auf $(0,1)$ eindeutig bestimmt ist. Wegen $n+x = (1-x)n + x(n+1)$ folgt aus (i) für $0<x<1$

$$F(n+x) \leqslant F(n)^{1-x}\cdot F(n+1)^x = [(n-1)!]^{1-x}\cdot[n!]^x = (n-1)!\cdot n^x.$$

Mit $n+1 = x(n+x) + (1-x)(n+1+x)$ folgt aus (i), (ii) für $0<x<1$

$$n! = F(n+1) \leqslant F(n+x)^x\cdot F(n+1+x)^{1-x} = F(n+x)\cdot(n+x)^{1-x}.$$

Aus diesen beiden Ungleichungen ergibt sich

$$n!\cdot(n+x)^{x-1} \leqslant F(n+x) \leqslant (n-1)!\cdot n^x, \quad \text{also}$$

$$a_n(x) := \frac{n!\cdot(n+x)^{x-1}}{x\cdot(x+1)\cdot\ldots\cdot(x+n-1)} \leqslant F(x) \leqslant \frac{(n-1)!\cdot n^x}{x\cdot(x+1)\cdot\ldots\cdot(x+n-1)} =: b_n(x),$$

wenn man $(*)$ verwendet. Aufgrund von

$$\frac{a_n(x)}{b_n(x)} = \frac{n(n+x)^{x-1}}{n^x} = \left(1+\frac{x}{n}\right)^{x-1} \xrightarrow[n\to\infty]{} 1$$

erhält man $F(x) = \lim\limits_{n\to\infty} b_n(x)$ (sowie die Existenz von $\lim\limits_{n\to\infty} b_n(x)$) aus dem Sandwich-Lemma III(1.11), denn es ist

$$\frac{a_n(x)}{b_n(x)} \geqslant \frac{F(x)}{b_n(x)} \geqslant 1$$

und $\lim\limits_{n\to\infty}\frac{a_n(x)}{b_n(x)} = 1$. Also ist F eindeutig bestimmt. □

Wir gewinnen aus dem Beweis eine Produktdarstellung der Gamma-Funktion.

(4.7) Satz. *Für alle $x>0$ gilt*

$$\Gamma(x) = \lim_{n\to\infty}\frac{n!\cdot n^x}{x\cdot(x+1)\cdot\ldots\cdot(x+n)}.$$

Beweis. Wegen $\lim_{n\to\infty}\frac{n}{x+n} = 1$ folgt die Behauptung für $0<x<1$ wegen

$$\lim_{n\to\infty} b_n(x) = \Gamma(x)$$

aus dem letzten Beweis. Für $x=1$ sind beide Seiten gleich 1. Es genügt zu zeigen, dass die Formel für $y := x+1$ gilt, wenn sie für $x>0$ gilt. Man hat

$$\begin{aligned}\Gamma(y) &= \Gamma(x+1) = x\cdot\Gamma(x) = \lim_{n\to\infty}\frac{n!\cdot n^x}{(x+1)\cdot\ldots\cdot(x+n)} = \lim_{n\to\infty}\frac{n!\cdot n^{y-1}}{y\cdot\ldots\cdot(y+n-1)}\\ &= \lim_{n\to\infty}\frac{n!\cdot n^{y-1}}{y\cdot\ldots\cdot(y+n-1)}\cdot\lim_{n\to\infty}\frac{n}{y+n} = \lim_{n\to\infty}\frac{n!\cdot n^y}{y\cdot\ldots\cdot(y+n)}.\end{aligned}$$ □

Als speziellen Wert notieren wir noch das

(4.8) Lemma. $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \Gamma\left(\frac{1}{2}\right) \ (= \sqrt{\pi})$.

Beweis. Für $0 < a < b$ liefert die Substitution $x = \sqrt{t}$, $dx = \frac{dt}{2\sqrt{t}}$ in (2.8)

$$\int_a^b e^{-x^2}\,dx = \int_{a^2}^{b^2} e^{-t}\frac{1}{2\sqrt{t}}\,dt = \frac{1}{2}\int_{a^2}^{b^2} t^{-1/2}e^{-t}\,dt.$$

Die Grenzübergänge $a \downarrow 0$ und $b \to \infty$ liefern

$$\int_0^\infty e^{-x^2}\,dx = \frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right).$$

Die Behauptung folgt aus

$$\int_{-\infty}^\infty e^{-x^2}\,dx = 2\int_0^\infty e^{-x^2}\,dx.$$

□

Der Wert des Integrals wird später berechnet (s. Skript zu Analysis III).

Kapitel IX.

Folgen und Reihen von Funktionen

In diesem Kapitel betrachten wir Funktionenfolgen. Insbesondere werden Stetigkeit, Differenzierbarkeit und Integrierbarkeit auf Vertauschbarkeit mit der Limesbildung untersucht. Die zentrale Frage lautet, wann die Grenzfunktion die entsprechende Eigenschaft hat.

§1. Punktweise und gleichmäßige Konvergenz

In diesem Paragrafen werden zwei unterschiedliche Konvergenzbegriffe eingeführt, die auf obige Fragen auch unterschiedliche Antworten liefern.

(1.1) Definition. Es sei D eine nicht-leere Menge. Eine (reelle) *Funktionenfolge auf* D ist eine Folge $(f_n)_{n\geqslant 1}$ von Funktionen $f_n : D \to \mathbb{R}$, $n \geqslant 1$. Man nennt die Funktionenfolge $(f_n)_{n\geqslant 1}$ auf D *punktweise konvergent* gegen die Funktion $f : D \to \mathbb{R}$, wenn zu jedem $x \in D$ und zu jedem $\varepsilon > 0$ ein $N = N(x, \varepsilon) \in \mathbb{N}$ existiert, so dass

$$|f_n(x) - f(x)| < \varepsilon \quad \text{für alle } n \geqslant N.$$

Dafür schreibt man $f_n \to f$.
Man nennt die Funktionenfolge $(f_n)_{n\geqslant 1}$ auf D *gleichmäßig konvergent* gegen die Funktion $f : D \to \mathbb{R}$, wenn es zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ gibt, so dass

$$|f_n(x) - f(x)| < \varepsilon \quad \text{für alle } n \geqslant N \text{ und alle } x \in D.$$

Dafür schreibt man $f_n \underset{\text{glm}}{\longrightarrow} f$.
In beiden Fällen heißt f *Grenzfunktion* von $(f_n)_{n\geqslant 1}$.

Wir interessieren uns i. F. vor allem für $D \subset \mathbb{R}$.

Nützlich sind die folgenden

(1.2) Bemerkungen. a) Offenbar folgt aus der gleichmäßigen Konvergenz die punktweise Konvergenz.
b) Ist $(f_n)_{n\geqslant 1}$ eine Funktionenfolge auf D, so ist $(f_n(x))_{n\geqslant 1}$ für jedes $x \in D$ eine reelle Folge. Damit ist $f_n \to f$ äquivalent zu

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{für alle } x \in D.$$

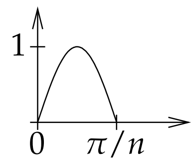
c) Der gewöhnliche Limes reeller Zahlen ist eindeutig bestimmt. Also ist nach a) und b) auch die Grenzfunktion f in (1.1) eindeutig bestimmt, sofern sie existiert.

Der Illustration dienen die

(1.3) Beispiele. a) Sei $D = \mathbb{R}$, $f_n(x) = \cos(nx)$, $n \in \mathbb{N}$. Dann gilt zwar $f_n(0) = 1$, aber $f_n(\pi) = (-1)^n$ für alle $n \in \mathbb{N}$. Also konvergiert $(f_n)_{n\geqslant 1}$ nicht.
b) Sei $D = \mathbb{R}$ und $f_n(x) = \frac{\sin(nx)}{\sqrt{n}}$, $f(x) = 0$, $x \in \mathbb{R}$. Dann gilt $f_n \xrightarrow[\text{glm}]{} f$. Zu $\varepsilon > 0$ wählt man $N \in \mathbb{N}$, $N > \frac{1}{\varepsilon^2}$ und erhält

$$\left|\frac{\sin(nx)}{\sqrt{n}} - 0\right| \leqslant \frac{1}{\sqrt{n}} \leqslant \frac{1}{\sqrt{N}} < \varepsilon \quad \text{für alle } n \geqslant N \text{ und alle } x \in \mathbb{R}.$$

c) Sei $D = \mathbb{R}$, $f_n(x) = \begin{cases} \sin(nx), & \text{falls } 0 \leqslant x \leqslant \frac{\pi}{n}, \\ 0, & \text{sonst.} \end{cases}$, $f(x) = 0$, $\quad x \in \mathbb{R}$.

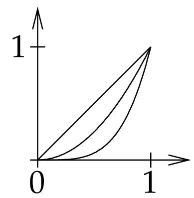


Dann gilt $f_n(x) = 0 = f(x)$ für alle $x \leqslant 0$ und alle $n \in \mathbb{N}$. Ist $x > 0$, so hat man $f_n(x) = 0$ für alle $n \geqslant n_0 := 1 + [\pi/x]$. Es folgt

$$\lim_{n\to\infty} f_n(x) = 0 = f(x) \quad \text{für alle } x \in \mathbb{R}, \quad \text{also } f_n \to f.$$

Sei $\varepsilon = \frac{1}{2}$ und $N \in \mathbb{N}$. Zu $n > N$ wählt man $x := \frac{\pi}{2n}$, also $f_n(x) = \sin\frac{\pi}{2} = 1$. Es folgt $|f_n(x) - f(x)| = 1 > \varepsilon$. Also konvergiert $(f_n)_{n\geqslant 1}$ nicht gleichmäßig gegen f.

d) Sei $D = [0,1]$, $f_n(x) = x^n$, $f(x) = \begin{cases} 0, & \text{falls } x < 1, \\ 1, & \text{falls } x = 1. \end{cases}$

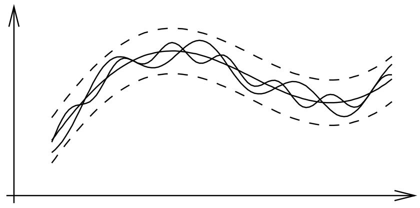


Es gilt $f_n(1) = 1 = f(1)$ für alle $n \geqslant 1$ und $\lim\limits_{n\to\infty} f_n(x) = \lim\limits_{n\to\infty} x^n = 0 = f(x)$ für alle $0 \leqslant x < 1$, also $f_n \to f$.
Sei $\varepsilon = \frac{1}{2}$, $N \in \mathbb{N}$. Zu $n > N$ wählt man $x := \sqrt[n]{\frac{3}{4}} \in (0,1)$. Dann gilt

$$|f_n(x) - f(x)| = \left(\sqrt[n]{\frac{3}{4}}\right)^n = \frac{3}{4} > \frac{1}{2} = \varepsilon.$$

Also konvergiert $(f_n)_{n\geqslant 1}$ nicht gleichmäßig gegen f.

Geometrisch bedeutet gleichmäßige Konvergenz, dass in jedem ε-Schlauch um den Graphen von f fast alle Funktionsgraphen der f_n verlaufen.



(1.4) Definition. Ist $g : D \to \mathbb{R}$ eine Funktion, so definiert man die *Supremumsnorm* von g durch

$$\|g\| := \|g\|_D := \sup\{|g(x)|;\ x \in D\} \in \mathbb{R}_+ \cup \{\infty\}.$$

Ein nützliches Kriterium formulieren wir nun als

(1.5) Lemma. *Die Funktionenfolge $(f_n)_{n\geqslant 1}$ auf D konvergiert genau dann gleichmäßig gegen $f : D \to \mathbb{R}$, wenn $\lim\limits_{n\to\infty} \|f_n - f\| = 0$. (Das bedeutet insbesondere die Existenz eines $N \in \mathbb{N}$ mit $\|f_n - f\| < \infty$ für alle $n \geqslant N$.)*

Beweis. „$\Rightarrow$" Zu $\varepsilon > 0$ existiert nach Voraussetzung ein $N \in \mathbb{N}$, so dass $|f_n(x) - f(x)| < \varepsilon$ für alle $n \geqslant N$ und alle $x \in D$. Also gilt

$$\|f_n - f\| = \sup\{|f_n(x) - f(x)|; x \in D\} \leqslant \varepsilon \quad \text{für alle } \ n \geqslant N.$$

Das bedeutet $\|f_n - f\| \xrightarrow[n\to\infty]{} 0$.

„$\Leftarrow$" Zu $\varepsilon > 0$ existiert nach Voraussetzung ein $N \in \mathbb{N}$ mit $\|f_n - f\| < \varepsilon$ für alle $n \geqslant N$. Also hat man für alle $n \geqslant N$ und alle $x \in D$

$$|f_n(x) - f(x)| \leqslant \|f_n - f\| < \varepsilon.$$

Das bedeutet $f_n \xrightarrow[\text{glm}]{} f$. □

Diese Aussage verwenden wir in dem folgenden

(1.6) Beispiel. Sei $D = [0,1], f_n(x) = n^\alpha x(1-x)^n$, $f(x) = 0$, $x \in [0,1]$, $\alpha \in \mathbb{R}$. Es gilt $f_n(0) = f_n(1) = 0$ und $f_n'(x) = n^\alpha(1-x)^n - n^{\alpha+1}x(1-x)^{n-1} = n^\alpha(1-x)^{n-1}(1-(n+1)x)$. Also sind $x = 1$ und $x = \frac{1}{n+1}$ die Nullstellen von f_n'. Aus $f_n(0) = f_n(1) = 0$ und $f_n\left(\frac{1}{n+1}\right) = \frac{n^\alpha}{n+1}\left(1 - \frac{1}{n+1}\right)^n$ folgt

$$\|f_n - f\| = \frac{n^\alpha}{n+1}\left(1 - \frac{1}{n+1}\right)^n = n^{\alpha-1}\cdot\left(1 - \frac{1}{n+1}\right)^{n+1} \to 0 \Leftrightarrow \alpha < 1.$$

Demnach konvergiert $(f_n)_{n\geqslant 1}$ genau dann gleichmäßig gegen f, wenn $\alpha < 1$.

Wie bei gewöhnlichen Folgen erhalten wir den

(1.7) Satz. *(*CAUCHY*-Kriterium)*
Für eine Funktionenfolge $(f_n)_{n\geqslant 1}$ auf D sind äquivalent:

(i) Es gibt ein $f: D \to \mathbb{R}$ mit $f_n \underset{glm}{\longrightarrow} f$.

(ii) Zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit der Eigenschaft

$$\|f_n - f_m\| < \varepsilon \quad \textit{für alle } n, m \geqslant N.$$

(iii) Zu jedem $\varepsilon > 0$ existiert ein $N = N(\varepsilon) \in \mathbb{N}$, so dass

$$|f_n(x) - f_m(x)| < \varepsilon \quad \textit{für alle } n, m \geqslant N \textit{ und alle } x \in D.$$

Beweis. "(i) $\Rightarrow$ (ii)" Sei $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$, so dass

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \text{für alle } n \geqslant N \text{ und alle } x \in D.$$

Es folgt mit der Dreiecksungleichung

$$|f_n(x) - f_m(x)| \leqslant |f_n(x) - f(x)| + |f(x) - f_m(x)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3}$$

für alle $n, m \geqslant N$ und alle $x \in D$, also auch

$$\|f_n - f_m\| = \sup\{|f_n(x) - f_m(x)|;\ x \in D\} \leqslant \frac{2\varepsilon}{3} < \varepsilon.$$

"(ii) $\Rightarrow$ (iii)" Zu $\varepsilon > 0$ existiert nach Voraussetzung ein $N \in \mathbb{N}$, so dass

$$\|f_n - f_m\| < \varepsilon \quad \text{für alle } n, m \geqslant N,$$

also $|f_n(x) - f_m(x)| < \varepsilon$ für alle $x \in D$ und $n, m \geqslant N$.

"(iii) $\Rightarrow$ (i)" Für jedes $x \in D$ gilt $|f_n(x) - f_m(x)| < \varepsilon$ für alle $n, m > N(\varepsilon)$, so dass $(f_n(x))_{n\geqslant 1}$ eine CAUCHY-Folge ist. Also existiert eine Funktion $f: D \to \mathbb{R}$ mit $f(x) := \lim\limits_{n\to\infty} f_n(x)$. Das bedeutet $f_n \to f$.
Zu $\varepsilon > 0$ wählen wir N wie in der Voraussetzung. Ist $x \in D$ und $n \geqslant N$, so folgt

$$|f_n(x) - f_m(x)| < \varepsilon \quad \text{für alle } m \geqslant N,$$

also

$$|f_n(x) - f(x)| = |f_n(x) - \lim_{m\to\infty} f_m(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leqslant \varepsilon.$$

Man erhält $|f_n(x) - f(x)| < 2\varepsilon$ für alle $n \geqslant N$ und alle $x \in D$, also $f_n \underset{\text{glm}}{\longrightarrow} f$. □

Analog zu den reellen Folgen kann man das Konzept auf Funktionenreihen übertragen.

(1.8) Definition. Sei D eine nicht-leere Menge. Ist $(f_k)_{k\geqslant 1}$ eine (reelle) Funktionenfolge auf D, so heißt die Funktionenfolge $(s_n)_{n\geqslant 1}$ auf D mit $s_n(x) := \sum_{k=1}^{n} f_k(x)$, $x \in D$, eine *Funktionenreihe* auf D, die mit $\sum_{k=1}^{\infty} f_k(x)$ bezeichnet wird. s_n heißt n-te *Partialsummenfunktion*. Man nennt die Funktionenreihe *punktweise* bzw. *gleichmäßig konvergent* , wenn die Partialsummenfolge $(s_n)_{n\geqslant 1}$ punktweise bzw. gleichmäßig konvergiert. Die Funktionenreihe $\sum_{k=1}^{\infty} f_k(x)$ heißt *absolut* bzw. *absolut gleichmäßig konvergent*, wenn die Funktionenreihe $\sum_{k=1}^{\infty} |f_k(x)|$ konvergiert bzw. gleichmäßig konvergiert.

Wegen

$$s_m(x) - s_n(x) = \sum_{k=n+1}^{m} f_k(x) \quad \text{für alle } m > n$$

hat das CAUCHY-Kriterium (1.7) in diesem Fall die Form

(1.9) Korollar. *(CAUCHY-Kriterium)*
Sei $(f_k)_{k\geqslant 1}$ eine Funktionenfolge auf D. Die Funktionenreihe $\sum_{k=1}^{\infty} f_k(x)$ konvergiert genau dann gleichmäßig auf D, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass

$$\left| \sum_{k=n+1}^{m} f_k(x) \right| < \varepsilon \quad \textit{für alle } m > n \geqslant N \textit{ und alle } x \in D.$$

Ein wichtiges Kriterium formulieren wir in dem

(1.10) Satz. *(WEIERSTRASSsches Majorantenkriterium)*
Gegeben sei eine Funktionenreihe $\sum_{k=1}^{\infty} f_k(x)$ auf D. Wenn es eine Folge $(M_k)_{k\geqslant 1}$ reeller Zahlen gibt, so dass

$$|f_k(x)| \leqslant M_k \quad \textit{für alle } x \in D,\ k \in \mathbb{N} \textit{ und } \quad \sum_{k=1}^{\infty} M_k < \infty,$$

dann konvergiert die Funktionenreihe $\sum_{k=1}^{\infty} f_k(x)$ gleichmäßig und absolut gleichmäßig auf D. Insbesondere ist die Funktionenreihe $\sum_{k=1}^{\infty} f_k(x)$ gleichmäßig konvergent, wenn $\sum_{k=1}^{\infty} \|f_k\| < \infty$.

Beweis. Sei $\varepsilon > 0$. Aus dem CAUCHY-Kriterium für die Konvergenz von $\sum_{k=1}^{\infty} M_k$ folgt die Existenz eines $N \in \mathbb{N}$ mit $\sum_{k=n+1}^{m} M_k < \varepsilon$ für alle $m > n \geqslant N$. Also gilt für alle $m > n \geqslant N$ und alle $x \in D$ mit der Dreiecksungleichung

$$\left| \sum_{k=n+1}^{m} f_k(x) \right| \leqslant \sum_{k=n+1}^{m} |f_k(x)| \leqslant \sum_{k=n+1}^{m} M_k < \varepsilon.$$

Dann folgt die Behauptung aus (1.9). □

Wir betrachten nun als wichtigste Beispielklasse Potenzreihen $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ als Funktionenreihen mit den Partialsummenfunktionen $s_n(x) = \sum_{k=0}^{n} a_k(x-x_0)^k$, die Polynomfunktionen sind.

(1.11) Satz. *Gegeben sei eine reelle Potenzreihe*

$$\sum_{k=0}^{\infty} a_k(x-x_0)^k$$

mit dem Konvergenzradius $R > 0$. *Für jedes* $0 \leqslant r < R$ *konvergiert die Potenzreihe absolut und gleichmäßig auf* $[x_0 - r, x_0 + r]$.

Beweis. Zur Erinnerung: Es ist $R = \sup\{\rho \geqslant 0;\ \sum |a_k|\rho^k \text{ konvergiert}\}$ und $\sum |a_k|r^k$ konvergiert für $0 \leqslant r < \rho$. Ist also $0 \leqslant r < R$, so gilt $|a_k(x-x_0)^k| \leqslant |a_k|r^k$ für alle $x \in [x_0 - r, x_0 + r]$. Weil $\sum_{k=0}^{\infty} |a_k|r^k$ wegen $r < R$ konvergiert, folgt die Behauptung aus (1.10). □

Die Potenzreihe konvergiert i. Allg. **nicht gleichmäßig** auf $(x_0 - R, x_0 + R)$. Zum Beispiel konvergiert die geometrische Reihe $\sum_{k=0}^{\infty} x^k$ nicht gleichmäßig auf (-1,1) (vgl. (2.2)), wohl aber $\sum_{k=1}^{\infty} \frac{1}{k^2}x^k$ auf (-1,1). Die Exponentialreihe konvergiert nicht gleichmäßig auf $\mathbb{R}$.

(1.12) Beispiel. Die RIEMANNsche Zetafunktion

$$\zeta(x) := \sum_{k=1}^{\infty} \frac{1}{k^x},\ x \in (1, \infty),$$

konvergiert nach VIII(3.8) punktweise auf $(1, \infty)$. Ist $\varepsilon > 0$, so konvergiert sie aufgrund von $\left|\frac{1}{k^x}\right| \leqslant \frac{1}{k^{1+\varepsilon}}$ für $x \geqslant 1 + \varepsilon$ nach (1.10) gleichmäßig auf $[1 + \varepsilon, \infty)$. Allerdings konvergiert sie nicht gleichmäßig auf $(1, \infty)$, denn für $m > n$ gilt

$$\|s_m(x) - s_n(x)\| = \left\| \sum_{k=n+1}^{m} \frac{1}{k^x} \right\| = \sum_{k=n+1}^{m} \frac{1}{k} \xrightarrow[m\to\infty]{} \infty.$$

§2. Vertauschungssätze

In diesem Paragrafen untersuchen wir die Eigenschaften der Grenzfunktion von Funktionenfolgen. Dazu benötigen wir die

(2.1) Definition. Eine Funktionenfolge $(f_n)_{n\geqslant 1}$ auf D heißt *gleichgradig beschränkt*, wenn es ein $M \in \mathbb{R}$ gibt mit der Eigenschaft

$$|f_n(x)| \leqslant M \quad \text{für alle } n \in \mathbb{N} \quad \text{und alle } x \in D.$$

Als erste Aussage notieren wir den

(2.2) Satz. *Ist die Funktionenfolge $(f_n)_{n\geqslant 1}$ auf D gleichmäßig konvergent gegen f und ist jedes f_n einzeln beschränkt, d. h. es gibt für alle $n \in \mathbb{N}$ ein $M_n \in \mathbb{R}$ mit $|f_n(x)| \leqslant M_n$ für alle $x \in D$, so ist $(f_n)_{n\geqslant 1}$ gleichgradig beschränkt und die Grenzfunktion f ist ebenfalls beschränkt. Darüber hinaus gilt*

$$\lim_{n\to\infty} \|f_n\| = \|f\|.$$

Aus $f_n \xrightarrow[glm]{} f$ und $\|f_n\| < \infty$ für alle $n \in \mathbb{N}$ folgt also $\|f\| < \infty$ und $\|f_n\| \xrightarrow[n\to\infty]{} \|f\|$.

Beweis. Zu $\varepsilon = 1$ existiert nach (1.8) ein $N \in \mathbb{N}$, so dass $|f_n(x) - f_m(x)| < 1$ für alle $n, m \geqslant N$ und $x \in D$. Man erhält

$$|f_n(x)| \leqslant |f_n(x) - f_N(x)| + |f_N(x)| \leqslant 1 + M_N \quad \text{für alle } n \geqslant N,\ x \in D.$$

Mit $M := 1 + \max\{M_1, \dots, M_N\}$ folgt $|f_n(x)| \leqslant M$ für alle $n \in \mathbb{N}$, $x \in D$ und damit auch $|f(x)| = \lim\limits_{n\to\infty} |f_n(x)| \leqslant M$ für alle $x \in D$.

Sei $\varepsilon > 0$. Dann wählen wir $x_n, y \in D$ mit der Eigenschaft

$$\|f_n\| \leqslant |f_n(x_n)| + \varepsilon/2, \quad \|f\| \leqslant |f(y)| + \varepsilon/2$$

sowie ein $N \in \mathbb{N}$ mit $\|f_n - f\| < \varepsilon/2$ für alle $n \geqslant N$. Dann gilt für alle $n \geqslant N$ mit der Dreiecksungleichung

$$\begin{aligned}
\|f_n\| - \|f\| &\leqslant |f_n(x_n)| + \varepsilon/2 - \|f\| \\
&\leqslant |f_n(x_n) - f(x_n)| + |f(x_n)| - \|f\| + \varepsilon/2 \\
&\leqslant \|f_n - f\| + \varepsilon/2 < \varepsilon
\end{aligned}$$

sowie

$$\begin{aligned}
\|f\| - \|f_n\| &\leqslant |f(y)| + \varepsilon/2 - \|f_n\| \\
&\leqslant |f(y) - f_n(y)| + |f_n(y)| - \|f_n\| + \varepsilon/2 \\
&\leqslant \|f - f_n\| + \varepsilon/2 < \varepsilon.
\end{aligned}$$

Also gilt $\big|\|f_n\| - \|f\|\big| < \varepsilon$ für alle $n \geqslant N$ und damit $\lim\limits_{n\to\infty} \|f_n\| = \|f\|$. □

Als Anwendung diskutieren wir ein

(2.3) Beispiel. Die geometrische Reihe $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ konvergiert nicht gleichmäßig auf $(-1, 1)$. Andernfalls wäre wegen

$$|s_n(x)| = \left|\sum_{k=0}^{n} x^k\right| \leqslant \sum_{k=0}^{n} |x|^k \leqslant n + 1 \quad \text{für alle } x \in (-1, 1)$$

auch die Grenzfunktion $\frac{1}{1-x}$ beschränkt auf $(-1, 1)$. Es gilt aber $\lim\limits_{x\uparrow 1} \frac{1}{1-x} = \infty$.

Nun untersuchen wir die Stetigkeit.

(2.4) Satz. *Sei $D \subset \mathbb{R}$, und die Funktionenfolge $(f_n)_{n\geqslant 1}$ sei auf D gleichmäßig konvergent gegen $f : D \to \mathbb{R}$.*
a) Ist jedes f_n stetig in $x_0 \in D$, so ist f stetig in x_0, d. h.

$$\lim_{n\to\infty}\big(\lim_{x\to x_0} f_n(x)\big) = \lim_{x\to x_0}\big(\lim_{n\to\infty} f_n(x)\big),$$

falls x_0 kein isolierter Punkt von D ist.
b) Ist jedes f_n stetig, so ist f stetig.

Beweis. b) Man verwende a) für alle $x_0 \in D$.
a) Falls x_0 kein Häufungspunkt von D ist, sind alle f_n und f stetig in x_0. Sei nun x_0 ein Häufungspunkt von D und $\varepsilon > 0$. Wir benutzen die Identität

$$f(x) - f(x_0) = \big(f(x) - f_n(x)\big) + \big(f_n(x) - f_n(x_0)\big) + \big(f_n(x_0) - f(x_0)\big)$$

und schätzen die Terme rechts einzeln ab.
Weil $(f_n)_{n\geqslant 1}$ gleichmäßig gegen f konvergiert, existiert ein $N \in \mathbb{N}$, so dass

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \text{für alle } n \geqslant N \text{ und alle } x \in D.$$

Da f_N stetig in x_0 ist, existiert ein $\delta > 0$, so dass

$$|f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D \text{ mit } |x - x_0| < \delta.$$

Dann folgt für alle $x \in D$ mit $|x - x_0| < \delta$

$$|f(x) - f(x_0)| \leqslant |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Also ist f stetig in x_0 und es gilt

$$\lim_{n\to\infty}\big(\lim_{x\to x_0} f_n(x)\big) = \lim_{n\to\infty} f_n(x_0) = f(x_0) = \lim_{x\to x_0} f(x) = \lim_{x\to x_0}\big(\lim_{n\to\infty} f_n(x)\big). \qquad \square$$

Für Reihen ergibt sich unmittelbar das

(2.5) Korollar. *Wenn die Funktionenreihe $\sum_{k=1}^{\infty} f_k(x)$ auf $D \subset \mathbb{R}$ gleichmäßig konvergiert und jedes f_k, $k \in \mathbb{N}$, auf D stetig ist, so ist auch die Grenzfunktion $\sum_{k=1}^{\infty} f_k(x)$ stetig auf D.*

Wendet man (2.5) auf Potenzreihen an, so erhält man einen neuen Beweis der Stetigkeit von Potenzreihen auf $(x_0 - R, x_0 + R)$, da $s_n(x)$ als Polynomfunktion stetig ist und die Potenzreihe für jedes $0 \leqslant r < R$ auf $[x_0 - r, x_0 + r]$ gleichmäßig konvergiert. Die gleichmäßige Konvergenz ist hinreichend, aber nicht notwendig für die Stetigkeit der Grenzfunktion (vgl. (1.3) c)). Die punktweise Konvergenz ist nicht hinreichend, wie das Beispiel (1.3) d) zeigt.

Nun wenden wir uns der Integrierbarkeit zu.

(2.6) Satz. *Gegeben sei eine Funktionenfolge* $(f_n)_{n\geqslant 1}$ *auf dem Intervall* $[a,b]$*, die gleichmäßig gegen die Funktion* $f:[a,b]\to\mathbb{R}$ *konvergiert.*
a) Ist jedes f_n*,* $n\geqslant 1$*, integrierbar, so ist auch* f *integrierbar und es gilt*

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

b) Ist $c\in[a,b]$ *fest, so gilt für alle* $x\in[a,b]$*:*

$$\int_c^x f(t)\,\mathrm{d}t = \lim_{n\to\infty}\int_c^x f_n(t)\,\mathrm{d}t.$$

Beweis. Sei $\varepsilon>0$. Dann existiert ein $N_0\in\mathbb{N}$, so dass $|f_n(x)-f_m(x)|<\varepsilon$ für alle $x\in[a,b]$ und alle $n,m\geqslant N_0$. Daraus folgt

$$\left|\int_a^b f_n(x)\,dx - \int_a^b f_m(x)\,dx\right| = \left|\int_a^b (f_n(x)-f_m(x))\,dx\right|$$
$$\leqslant \int_a^b |f_n(x)-f_m(x)|\,dx \leqslant \int_a^b \varepsilon\,dx = \varepsilon(b-a) \quad \text{für } n,m\geqslant N_0.$$

Demnach ist $\left(\int_a^b f_n(x)\,dx\right)_{n\geqslant 1}$ eine CAUCHY-Folge in $\mathbb{R}$ und $I := \lim\limits_{n\to\infty}\int_a^b f_n(x)\,dx$ existiert. Sei $N_1\in\mathbb{N}$ mit der Eigenschaft

$$\left|I - \int_a^b f_n(x)\,dx\right| < \varepsilon \quad \text{für alle } n\geqslant N_1.$$

Dann wählen wir $N_2\in\mathbb{N}$, so dass

$$|f_n(x)-f(x)|<\varepsilon \quad \text{für alle } x\in[a,b],\ n\geqslant N_2.$$

Sei $N=\max\{N_1,N_2\}$. Nun wählen wir $\varphi,\psi\in T[a,b]$ mit

$$\varphi\leqslant f_N\leqslant\psi,\quad \int_a^b (f_N(x)-\varphi(x))\,dx\leqslant\varepsilon,\quad \int_a^b (\psi(x)-f_N(x))\,dx\leqslant\varepsilon.$$

Dann sind $\varphi-\varepsilon$, $\psi+\varepsilon$ Treppenfunktionen mit $\varphi-\varepsilon\leqslant f\leqslant\psi+\varepsilon$ und

$$0\leqslant\int_a^b(\psi(x)+\varepsilon)\,dx - I = \varepsilon(b-a)+\int_a^b\psi(x)\,dx-\int_a^b f_N(x)\,dx+\int_a^b f_N(x)\,dx - I$$
$$\leqslant \varepsilon(b-a)+\varepsilon+\varepsilon=\varepsilon(2+b-a),\quad \text{sowie}$$
$$0\leqslant I-\int_a^b(\varphi(x)-\varepsilon)\,dx\leqslant\varepsilon(2+b-a).$$

Nach VIII(1.14) ist f integrierbar mit $\int_a^b f(x)\,dx = I$. Damit ist Teil a) bewiesen und b) folgt mit denselben Argumenten. □

Die gleichmäßige Konvergenz ist hinreichend, aber nicht notwendig für die Integrierbarkeit der Grenzfunktion, wie (1.3) zeigt:
$f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$, konvergieren punktweise gegen $f : [0,1] \to \mathbb{R}, f(x) = 0$ für $0 \leqslant x < 1$, $f(1) = 1$, jedoch konvergiert $(f_n)_{n\geqslant 1}$ nicht gleichmäßig gegen f. Aber f_n und f sind integrierbar mit

$$\int_0^1 x^n\,dx = \frac{1}{n+1} \xrightarrow[n\to\infty]{} 0 = \int_0^1 f(x)\,dx.$$

Andererseits ist die punktweise Konvergenz nicht hinreichend. Wir betrachten die Funktionen $f_n : [0,\pi] \to \mathbb{R}$, $f_n(x) = n\sin(nx)$ für $0 \leqslant x \leqslant \frac{\pi}{n}$ und $f_n(x) = 0$ für $x > \frac{\pi}{n}$. Dann konvergiert $(f_n)_{n\geqslant 1}$ punktweise gegen $f \equiv 0$. Aus $\|f_n\| = n$ folgt, dass $(f_n)_{n\geqslant 1}$ nicht gleichmäßig gegen f konvergiert. Es gilt

$$\int_0^\pi f_n(x)\,dx = \int_0^{\pi/n} n\sin(nx)\,dx = -\cos(nx)\Big|_0^{\pi/n} = 2 \neq 0 = \int_0^\pi f(x)\,dx.$$

Es gibt aber noch schwächere Voraussetzungen, unter denen die Vertauschbarkeit richtig bleibt, z. B. der Satz von ARZELA, OSGOOD, LEBESGUE. Man vergleiche BARNER/FLOHR, Analysis I, §10.4.

(2.7) Korollar. *Sei $(f_k)_{k\geqslant 1}$ eine Folge von integrierbaren Funktionen $f_k : [a,b] \to \mathbb{R}$ und $c \in [a,b]$. Wenn die Funktionenreihe $\sum_{k=1}^\infty f_k(x)$ gleichmäßig auf $[a,b]$ konvergiert, ist die Grenzfunktion ebenfalls integrierbar und es gilt*

$$\int_c^x \sum_{k=1}^\infty f_k(t)\,dt = \sum_{k=1}^\infty \int_c^x f_k(t)\,dt \quad \text{für alle } x \in [a,b].$$

Beweis. Man verwende (2.6) und $\int_c^x s_n(t)\,dt = \sum_{k=1}^n \int_c^x f_k(t)\,dt$ aufgrund der Linearität des Integrals. □

Wir wenden uns nun der Differenzierbarkeit zu, wo die Dinge nicht so einfach sind.

(2.8) Beispiele. a) Wir betrachten

$$f_n : [-1,1] \to \mathbb{R},\ x \mapsto \sqrt{x^2 + \frac{1}{n}}, \quad f : [-1,1] \to \mathbb{R},\ x \mapsto |x|.$$

Dann gilt für alle $x \in [-1,1]$

$$|f_n(x) - f(x)| = \sqrt{x^2 + \frac{1}{n}} - |x| = \frac{1/n}{\sqrt{x^2 + 1/n} + |x|} \leqslant \frac{1/n}{\sqrt{1/n}} = \frac{1}{\sqrt{n}}.$$

Man erhält $\|f_n - f\| = \frac{1}{\sqrt{n}} \xrightarrow[n\to\infty]{} 0$, also $f_n \xrightarrow[\text{glm}]{} f$. Nun sind alle f_n differenzierbar, nicht aber die Grenzfunktion f.

b) $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = \frac{\sin(n^2x)}{n}$, konvergiert gleichmäßig gegen $f \equiv 0$, da $\|f_n\| = \frac{1}{n} \xrightarrow[n\to\infty]{} 0$. Alle f_n und f sind differenzierbar. Aber $f_n'(x) = n\cos(n^2x)$ konvergiert nicht.

Das Ergebnis formulieren wir in diesem Fall als

(2.9) Satz. *Sei $(f_n)_{n\geqslant 1}$ eine Folge stetig differenzierbarer Funktionen auf einem beschränkten Intervall I. Die Funktionenfolge $(f_n')_{n\geqslant 1}$ konvergiere auf I gleichmäßig gegen eine Funktion $g : I \to \mathbb{R}$. Wenn ein $c \in I$ existiert, so dass $(f_n(c))_{n\geqslant 1}$ konvergiert, so konvergiert $(f_n)_{n\geqslant 1}$ gleichmäßig gegen eine differenzierbare Funktion $f : I \to \mathbb{R}$ mit $f' = g$, d. h.*

$$f_n' \xrightarrow[glm]{} g,\ (f_n(c))_{n\geqslant 1} \quad \textit{konvergiert} \quad \Longrightarrow \quad \exists f : f_n \xrightarrow[glm]{} f, f' = g.$$

Beweis. Nach VIII(2.4) gilt

$$f_n(x) = f_n(c) + \int_c^x f_n'(t)\,dt \quad \text{für alle } \ x \in I,\ n \in \mathbb{N}.$$

Zu $\varepsilon > 0$ wählen wir ein $N \in \mathbb{N}$ mit der Eigenschaft

$$|f_n(c) - f_m(c)| < \frac{\varepsilon}{2}, \quad \|f_n' - f_m'\| < \frac{\varepsilon}{2L} \quad \text{für alle } \ n, m \geqslant N,$$

wobei $L > 0$ die Länge von I bezeichne. Dann gilt für alle $x \in I, n, m \geqslant N$

$$|f_n(x) - f_m(x)| = \left| f_n(c) + \int_c^x f_n'(t)\,dt - f_m(c) - \int_c^x f_m'(t)\,dt \right|$$
$$\leqslant |f_n(c) - f_m(c)| + \left| \int_c^x |f_n'(t) - f_m'(t)|\,dt \right| \leqslant \frac{\varepsilon}{2} + \|f_n' - f_m'\| \cdot |x - c| < \varepsilon.$$

Also ist $(f_n)_{n\geqslant 1}$ nach (1.7) gleichmäßig konvergent gegen eine Funktion $f : I \to \mathbb{R}$. Nach (2.4) ist g stetig und mit (2.6) folgt

$$f(c) + \int_c^x g(t)\,dt = f(c) + \int_c^x \lim_{n\to\infty} f_n'(t)\,dt = \lim_{n\to\infty} \left(f(c) + \int_c^x f_n'(t)\,dt \right)$$
$$= \lim_{n\to\infty} f_n(x) = f(x).$$

Also ist f nach VIII(2.2) differenzierbar mit $f' = g$. □

Das entsprechende Ergebnis für Reihen formulieren wir als

(2.10) Korollar. *Sei $\sum_{k=1}^{\infty} f_k(x)$ eine Funktionenreihe auf einem beschränkten Intervall I, die in mindestens einem Punkt $c \in I$ konvergiert. Sind alle $f_k, k \geqslant 1$, stetig differenzierbar und konvergiert $\sum_{k=1}^{\infty} f_k'(x)$ gleichmäßig auf I, so konvergiert auch $\sum_{k=1}^{\infty} f_k(x)$ gleichmäßig gegen eine differenzierbare Funktion und es gilt*

$$\frac{d}{dx}\sum_{k=1}^{\infty} f_k(x) = \sum_{k=1}^{\infty} \frac{d}{dx} f_k(x).$$

Man kann mit (2.10) einen neuen Beweis für die gliedweise Differenzierbarkeit von Potenzreihen führen.

Darüber hinaus kann man in (2.9) und (2.10) auf die Stetigkeit der Ableitungen verzichten. Dann ist der Beweis aber sehr viel aufwändiger.

(2.11) Definition. Eine Funktionenfolge $(f_n)_{n\geqslant 1}$ auf $D \subset \mathbb{R}$ heißt *gleichgradig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$|f_n(x) - f_n(y)| < \varepsilon \quad \text{für alle} \quad x, y \in D \quad \text{mit} \quad |x - y| < \delta \quad \text{und alle} \quad n \in \mathbb{N}.$$

Insbesondere ist dann jedes f_n auch gleichmäßig stetig auf D.

(2.12) Satz. *Sei $(f_n)_{n\geqslant 1}$ eine gleichmäßig konvergente Folge von Funktionen $f_n : K \to \mathbb{R}$, wobei $K \subset \mathbb{R}$ kompakt ist. Wenn alle f_n, $n \geqslant 1$, stetig sind, dann ist die Funktionenfolge $(f_n)_{n\geqslant 1}$ gleichgradig stetig und gleichgradig beschränkt.*

Beweis. Weil K kompakt und f_n stetig ist, ist f_n nach V(3.7) beschränkt. Dann ist $(f_n)_{n\geqslant 1}$ nach (2.2) gleichgradig beschränkt. Sei $\varepsilon > 0$. Weil $(f_n)_{n\geqslant 1}$ gleichmäßig gegen eine Funktion f konvergiert, existiert nach (1.7) ein $N \in \mathbb{N}$, so dass

$$|f_n(x) - f_m(x)| < \frac{\varepsilon}{3} \quad \text{für alle } n, m \geqslant N,\ x \in K.$$

Jedes f_n ist stetig auf dem Kompaktum K, also gleichmäßig stetig nach V(3.12). Daher existiert zu jedem $n \in \mathbb{N}$ ein $\delta_n > 0$ mit

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3} \quad \text{für alle } x, y \in K \text{ mit } |x - y| < \delta_n.$$

Sei $\delta := \min\{\delta_1, \ldots, \delta_N\} > 0$. Dann gilt für $1 \leqslant n \leqslant N$

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3} < \varepsilon \quad \text{für alle } x, y \in K \text{ mit } |x - y| < \delta.$$

Man hat für alle $n > N$ und alle $x, y \in K$ mit $|x - y| < \delta$

$$|f_n(x) - f_n(y)| \leqslant |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Also ist $(f_n)_{n\geqslant 1}$ gleichgradig stetig auf K. □

Es gibt auch eine Art Umkehrung, die wir ohne Beweis angeben (vgl. H. HEUSER, Lehrbuch der Analysis, Teil 1, S. 563).

(2.13) Auswahlsatz von ARZELA-ASCOLI. *Sei $(f_n)_{n\geqslant 1}$ eine Funktionenfolge auf einem Kompaktum K, die gleichgradig stetig und gleichgradig beschränkt ist. Dann existiert eine Teilfolge $(f_{n_k})_{k\geqslant 1}$, die auf K gleichmäßig konvergiert.*

Einige Vertauschungssätze lassen sich unter schwächeren Voraussetzungen zeigen.

Wir führen hierzu einen neuen Begriff ein.

(2.14) Definition. Eine Funktionenfolge $(f_n)_{n\geqslant 1}$ auf $D \subset \mathbb{R}$ heißt *lokal gleichmäßig konvergent*, wenn es zu jedem $x_0 \in D$ ein $\varepsilon = \varepsilon(x_0) > 0$ gibt, so dass $(f_n|_{(x_0-\varepsilon,x_0+\varepsilon)\cap D})_{n\geqslant 1}$ gleichmäßig konvergiert.

In dem Fall, dass der Definitionsbereich kompakt ist, haben wir das

(2.15) Lemma. *Sei $K \subset \mathbb{R}$ kompakt und $(f_n)_{n\geqslant 1}$ eine lokal gleichmäßig konvergente Funktionenfolge auf K. Dann ist $(f_n)_{n\geqslant 1}$ gleichmäßig konvergent auf K.*

Beweis. Sei f die Grenzfunktion von $(f_n)_{n\geqslant 1}$. Angenommen, $(f_n)_{n\geqslant 1}$ konvergiert nicht gleichmäßig. Dann existiert ein $\varepsilon^* > 0$ und zu jedem $N \in \mathbb{N}$ ein $n_N \geqslant N$ sowie ein $x_N \in K$ mit der Eigenschaft

$$(*) \qquad |f_{n_N}(x_N) - f(x_N)| \geqslant \varepsilon^*.$$

Also ist $(x_N)_{N\geqslant 1}$ eine Folge im Kompaktum K, die nach dem Satz von BOLZANO-WEIERSTRASS III(2.4) eine konvergente Teilfolge $x_{N_k} \to x_0 \in K$ besitzt. Nach (2.14) existiert ein $\delta > 0$, so dass die Funktionenfolge auf $(x_0 - \delta, x_0 + \delta) \cap K$ gleichmäßig konvergiert. Also existiert ein $n_0 \in \mathbb{N}$, so dass

$$(**) \qquad |f_n(x) - f(x)| < \varepsilon^* \quad \text{für alle } \; x \in K \cap (x_0 - \delta, x_0 + \delta), \; n \geqslant n_0.$$

Andererseits existiert ein $k_0 \geqslant n_0$ mit $x_{N_{k_0}} \in K \cap (x_0 - \delta, x_0 + \delta)$. Für $m = n_{N_{k_0}} \geqslant n_0$ gilt dann $|f_m(x_{N_{k_0}}) - f(x_{N_{k_0}})| \geqslant \varepsilon^*$ nach $(*)$ im Widerspruch zu $(**)$. Also ist $(f_n)_{n\geqslant 1}$ gleichmäßig konvergent auf K. □

Die wichtigste Klasse diskutieren wir in dem

(2.16) Beispiel. Ist $\sum_{k=0}^{\infty} a_k(x - x_0)^k$ eine reelle Potenzreihe mit Konvergenzradius $R > 0$, so konvergiert sie gleichmäßig auf $[x_0 - r, x_0 + r]$ für jedes $0 \leqslant r < R$, also lokal gleichmäßig auf $(x_0 - R, x_0 + R)$. Wie schon nach (1.11) bemerkt, ist die Konvergenz auf $(x_0 - \mathbb{R}, x_0 + \mathbb{R})$ i. Allg. nicht gleichmäßig.

Lokale Eigenschaften wie Stetigkeit und Differenzierbarkeit kann man schon mit lokal gleichmäßiger Konvergenz übertragen. Aus (2.9) und (2.10) folgt das

(2.17) Korollar. *a) Ist* $(f_n)_{n\geqslant 1}$ *eine Folge von stetigen Funktionen, die lokal gleichmäßig gegen eine Funktion* f *konvergiert, dann ist auch* f *stetig.*
b) Sei $(f_n)_{n\geqslant 1}$ *eine Folge von stetig differenzierbaren Funktionen, so dass* $(f_n)_{n\geqslant 1}$ *punktweise gegen eine Funktion* f *und* $(f_n')_{n\geqslant 1}$ *lokal gleichmäßig gegen eine Funktion* g *konvergiere. Dann konvergiert* $(f_n)_{n\geqslant 1}$ *lokal gleichmäßig gegen* f*, und* f *ist differenzierbar mit* $f' = g$*.*

Kapitel X.

Normierte Vektorräume

In diesem Kapitel geht es darum, topologische Eigenschaften wie Offenheit, Kompaktheit, Konvergenz und Stetigkeit im allgemeinen Kontext von normierten Vektorräumen zu entwickeln.

§1. Topologie normierter Vektorräume

Wohlbekannt aus der Linearen Algebra ist der Begriff des $\mathbb{R}$-*Vektorraums* V.

Man nennt bekanntlich $(V, \langle\cdot,\cdot\rangle)$ einen *euklidischen Vektorraum*, wenn die Abbildung

$$\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}, \quad (x,y) \mapsto \langle x,y\rangle,$$

ein *Skalarprodukt* ist, d. h. die Eigenschaften

$$\begin{aligned} \langle \alpha x + \beta y, z\rangle &= \alpha\langle x,z\rangle + \beta\langle y,z\rangle \quad \text{für alle } x,y,z \in V,\ \alpha,\beta \in \mathbb{R}, \\ \langle x,y\rangle &= \langle y,x\rangle \quad \text{für alle } x,y \in V, \\ \langle x,x\rangle &> 0 \quad \text{für alle } 0 \neq x \in V \end{aligned}$$

hat. In diesem Fall gilt die CAUCHY-SCHWARZsche Ungleichung

$$\langle x,y\rangle^2 \leqslant \langle x,x\rangle \cdot \langle y,y\rangle \quad \text{für alle } x,y \in V.$$

Als Standardbeispiel haben wir den Spaltenraum $\mathbb{R}^n$ mit dem kanonischen Skalarprodukt

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n, \quad \langle x,y\rangle := x^t y = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

(1.1) Definition. Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $\|\cdot\| : V \to \mathbb{R}$, $x \mapsto \|x\|$, heißt *Norm* auf V, wenn gilt

(N.1) $\|0\| = 0,\ \|x\| > 0$ für alle $0 \neq x \in V$,

(N.2) $\|\alpha x\| = |\alpha| \cdot \|x\|$ für alle $x \in V$, $\alpha \in \mathbb{R}$,

(N.3) (*Dreiecksungleichung*) $\|x+y\| \leqslant \|x\| + \|y\|$ für alle $x,y \in V$.

Man nennt dann $(V, \|\cdot\|)$ einen *normierten Vektorraum* oder einfach einen *normierten Raum*.

Eine zu (N.3) analoge Abschätzung formulieren wir in dem

(1.2) Lemma. *Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so gilt die zweite Dreiecksungleichung*

$$\Big|\|x\| - \|y\|\Big| \leqslant \|x - y\| \quad \textit{für alle } x, y \in V.$$

Beweis. Aus (N.3) und (N.2) mit $\alpha = -1$ folgt

$$\|x\| = \|(x-y) + y\| \leqslant \|x - y\| + \|y\|,$$
$$\|y\| = \|(y-x) + x\| \leqslant \|y - x\| + \|x\| = \|x - y\| + \|x\|.$$

Daraus erhält man

$$\Big|\|x\| - \|y\|\Big| = \max\{\|x\| - \|y\|,\ \|y\| - \|x\|\} \leqslant \|x - y\|. \qquad \square$$

Dem anschaulichen Verständnis dienen die

(1.3) Beispiele. a) Ist $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum, so sei

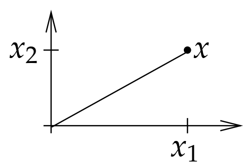
$$\|x\| := \sqrt{\langle x, x\rangle}.$$

Dann sind (N.1) und (N.2) offenbar erfüllt und die CAUCHY-SCHWARZsche Ungleichung liefert für alle $x, y \in V$

$$\|x+y\|^2 = \langle x+y,\ x+y\rangle = \langle x,x\rangle + \langle y,y\rangle + 2\langle x,y\rangle$$
$$\leqslant \langle x,x\rangle + \langle y,y\rangle + 2\sqrt{\langle x,x\rangle \cdot \langle y,y\rangle} = (\|x\| + \|y\|)^2,$$

also (N.3). Demnach ist $\|\cdot\|$ eine Norm auf V.

b) Ist $V = \mathbb{R}^n$ und $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ das kanonische Skalarprodukt, so spricht man von der *euklidischen Norm* auf dem $\mathbb{R}^n$:



$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}.$$

$\|x\|_2$ beschreibt im $\mathbb{R}^2$ und $\mathbb{R}^3$ den anschaulichen Abstand des Punktes x vom Nullpunkt. Für $n = 1$ erhalten wir den reellen Betrag.

c) Sei $p \in \mathbb{R}$, $p \geqslant 1$ und $(\mathbb{R}^n, \|\cdot\|_p)$ mit

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}.$$

Dann sind (N.1) und (N.2) wieder klar und man erhält (N.3) aus der MINKOWSKIschen Ungleichung VII(4.7). Also ist $\|\cdot\|_p$ eine Norm auf $\mathbb{R}^n$, die so genannte *p-Norm*, die im Fall $p = 2$

mit der euklidischen Norm übereinstimmt.
d) Ebenso ist auf $\mathbb{R}^{m\times n} = \left\{ (x_{ij})_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}};\ x_{ij}\in\mathbb{R} \right\}$ für $p\in\mathbb{R}$, $p\geqslant 1$ durch

$$\left\| (x_{ij})_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}} \right\|_p := \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |x_{ij}|^p \right)^{1/p}$$

eine Norm gegeben. (Dies ist schon wegen der Isomorphie von $\mathbb{R}^{m\times n}$ und $\mathbb{R}^{mn}$ klar.)
e) Für $x\in\mathbb{R}^n$ sei $\|x\|_\infty := \max\{|x_i|;\ 1\leqslant i\leqslant n\}$. Dann sind (N.1) und (N.2) wieder klar. Gilt $\|x\|_\infty = |x_i|$ und $\|y\|_\infty = |y_j|$, so folgt

$$\begin{aligned} \|x+y\|_\infty &= \max\{|x_1+y_1|,\dots,|x_n+y_n|\} \leqslant \max\{|x_1|+|y_1|,\dots,|x_n|+|y_n|\} \\ &\leqslant |x_i|+|y_j| = \|x\|_\infty + \|y\|_\infty. \end{aligned}$$

Also ist $(\mathbb{R}^n, \|\cdot\|_\infty)$ ein normierter Vektorraum. Man nennt $\|\cdot\|_\infty$ die *Maximumsnorm*.
f) Sei $\mathcal{C}([a,b])$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen $f:[a,b]\to\mathbb{R}$ und

$$\|f\|_\infty := \sup\{|f(x)|;\ x\in[a,b]\}.$$

Als stetige Funktion nimmt f auf dem kompakten Intervall $[a,b]$ nach V(3.7) sein Minimum und Maximum an. Wieder sind (N.1) und (N.2) klar und (N.3) folgt wie in e). Also ist $(\mathcal{C}([a,b]), \|\cdot\|_\infty)$ ein normierter Vektorraum.
g) Für $f\in\mathcal{C}([a,b])$ definiert man

$$\|f\|_2 := \sqrt{\int_a^b f(t)^2\,dt}.$$

Dann gelten wieder (N.1) und (N.2). Man erhält die Dreiecksungleichung aus VIII(1.22), so dass $(\mathcal{C}([a,b]), \|\cdot\|_2)$ ein normierter Vektorraum ist.

Einen neuen Begriff beschreiben wir in der

(1.4) Definition. Ist $(V,\|\cdot\|)$ ein normierter Vektorraum, $x\in V$, $r\in\mathbb{R}$, $r>0$, so heißt

$$K_r(x) := K_r^{(V,\|\cdot\|)}(x) := \{y\in V;\ \|y-x\|<r\}$$

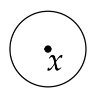
die *offene Kugel in* $(V,\|\cdot\|)$ mit Mittelpunkt x und Radius r.

Nun diskutieren wir erneut einige Beispiele aus (1.3).

(1.5) Beispiele. a) Sei $V=\mathbb{R}$ und $\|x\| = |x|$. Dann ist $K_r(x) = (x-r, x+r)$ das offene Intervall der Länge $2r$ mit Mittelpunkt x.

b) Sei $V=\mathbb{R}^2$ und $\|x\| = \|x\|_2$. Dann ist

$$K_r(x) = \{y\in\mathbb{R}^2;\ (y_1-x_1)^2+(y_2-x_2)^2<r^2\}$$



die offene Kreisscheibe mit Mittelpunkt x und Radius r.

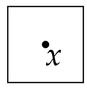
c) Ist $V = \mathbb{R}^3$ und

$$\|x\| = \|x\|_2, \quad \text{so ist} \quad K_r(x) = \{y \in \mathbb{R}^3;\ (y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2 < r^2\}$$

die offene Vollkugel mit Mittelpunkt x und Radius r.

d) Sei $V = \mathbb{R}^2$ und $\|x\| = \|x\|_\infty$. Dann ist

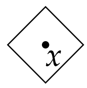
$$K_r(x) = \{y \in \mathbb{R}^2;\ |y_1 - x_1| < r,\ |y_2 - x_2| < r\}$$



das offene achsenparallele Quadrat mit Mittelpunkt x und Kantenlänge $2r$.

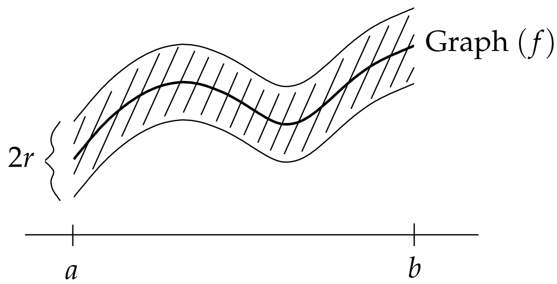
e) Sei $V = \mathbb{R}^2$ und $\|x\| = \|x\|_1$. Dann ist

$$K_r(x) = \{y \in \mathbb{R}^2;\ |y_1 - x_1| + |y_2 - x_2| < r\}$$



ein offenes Quadrat mit Mittelpunkt x und Diagonallänge $2r$.

f) In $\big(\mathcal{C}([a,b]), \|\cdot\|_\infty\big)$ ist eine analoge Visualisierung von $K_r(f)$ nicht möglich. Immerhin kann man sich aber noch vorstellen, dass die Graphen der Funktionen in $K_r(f)$ in einem Streifen der Breite $2r$ um den Graphen von f verlaufen.



(In $\big(\mathcal{C}([a,b]), \|\cdot\|_2\big)$ sind auch solche Veranschaulichungen nicht mehr möglich.)

Wie schon zuvor in III(3.3) für $\mathbb{R}$ treffen wir nun die

(1.6) Definition. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum, $x \in V$, $U \subset V$. Man nennt U eine *Umgebung von x* in $(V, \|\cdot\|)$, wenn es ein $r > 0$ gibt mit $K_r(x) \subset U$.

Einfache Eigenschaften dieses Begriffs formulieren wir als

(1.7) Lemma. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann gilt:*
a) Ist U eine Umgebung von x und $U \subset U' \subset V$, so gilt $x \in U$ und U' ist ebenfalls eine Umgebung von x.
b) Der Durchschnitt endlich vieler Umgebungen von $x \in V$ ist wieder eine Umgebung von x.
c) Es gilt das HAUSDORFF*sche Trennungsaxiom: Sind $x, x' \in V$ mit $x \neq x'$, so existieren Umgebungen U von x und U' von x' mit $U \cap U' = \emptyset$.*

Beweis. a) Nach (1.4) gibt es ein $r > 0$ mit $x \in K_r(x) \subset U \subset U'$.
b) Seien $r_i > 0$ mit $K_{r_i}(x) \subset U_i$ für $1 \leqslant i \leqslant m$ und $r := \min\{r_1, \ldots, r_m\}$. Dann gilt $K_r(x) \subset K_{r_i}(x)$ für $i = 1, \ldots, m$ und somit

$$K_r(x) \subset (U_1 \cap \ldots \cap U_m).$$

c) Sei $\|x - x'\| = 2r > 0$. Dann gilt $K_r(x) \cap K_r(x') = \emptyset$, denn $y \in K_r(x) \cap K_r(x')$ impliziert

$$\|x - x'\| = \|(x - y) + (y - x')\| \leqslant \|x - y\| + \|y - x'\| < 2r$$

als Widerspruch. □

Damit kommen wir zu einer ganzen Reihe topologischer Begriffe.

(1.8) Definition. Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum, $x \in V$, $M \subset V$.

a) Man nennt x einen *inneren Punkt* von M, wenn ein $r > 0$ mit $K_r(x) \subset M$ existiert (also M eine Umgebung von x ist). Die Menge aller inneren Punkte von M wird mit $\overset{\circ}{M}$ bezeichnet und das *Innere* oder der *offene Kern* von M genannt. M heißt *offen*, wenn jeder Punkt von M auch innerer Punkt ist, also $M = \overset{\circ}{M}$.

b) Man nennt x einen *Häufungspunkt* von M, wenn in jeder Umgebung von x ein Punkt aus $M \backslash \{x\}$ liegt. Ist $x \in M$ kein Häufungspunkt, gibt es also ein $r > 0$ mit $M \cap K_r(x) = \{x\}$, so nennt man x einen *isolierten Punkt* von M.

c) Der *Abschluss* oder die *abgeschlossene Hülle* von M ist definiert als

$$\overline{M} := \{x \in V;\ x \in M \text{ oder } x \text{ Häufungspunkt von } M\}.$$

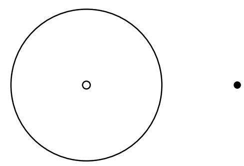
M heißt *abgeschlossen*, wenn $M = \overline{M}$.

d) Es heißt x *Randpunkt* von M, wenn jede Umgebung von x ein Element von M und ein Element von $V \backslash M$ enthält. Der *Rand* von M ist die Menge aller Randpunkte von M; Bezeichnung ∂M.

e) Man nennt M *beschränkt*, wenn ein $r > 0$ existiert, so dass $\|x\| < r$ für alle $x \in M$ (also $M \subset K_r(0)$).

Im Fall $(\mathbb{R}, |\cdot|)$ stimmen diese Definitionen mit den früheren (vgl. III, §3) überein. Man sollte sich aber bewusst machen, dass die Begriffe und Eigenschaften von der gegebenen Norm abhängen. Der Anschauung dienen die

(1.9) Beispiele.



a) Sei $V = \mathbb{R}^2$, $\|x\| = \|x\|_2 = \sqrt{x_1^2 + x_2^2}$ und

$$M = \{x \in \mathbb{R}^2;\ 0 < x_1^2 + x_2^2 \leqslant 1 \text{ oder } x = \begin{pmatrix} 2 \\ 0 \end{pmatrix}\}.$$

$\binom{2}{0}$ ist ein isolierter Punkt von M, also auch kein innerer Punkt von M, da

$$K_{1/2}\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) \cap M = \left\{\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right\}.$$

$\binom{0}{0}$ ist ein Häufungspunkt von M, aber kein Element von M, denn für $0 < r < 1$ gilt

$$\begin{pmatrix} r/2 \\ 0 \end{pmatrix} \in K_r\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) \cap M.$$

Jedes $x \in \mathbb{R}^2$ mit $0 < \|x\| < 1$ ist innerer Punkt und Häufungspunkt von M, denn für $r = \min\{\|x\|, 1 - \|x\|\} > 0$ gilt

$$y \in K_r(x) \quad \Rightarrow \quad \|y\| = \|(y-x)+x\| \leqslant \|y-x\| + \|x\| < r + \|x\| \leqslant 1 \text{ und } y \neq 0,$$

also $K_r(x) \subset M$.
Jedes $x \in \mathbb{R}^2$ mit $\|x\| = 1$ ist Häufungspunkt und Randpunkt von M, denn für $0 < r < 1$ gilt

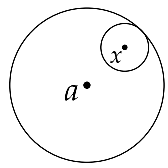
$$\Big(1 - \frac{r}{2}\Big)\, x \in K_r(x) \cap M \quad \text{und} \quad \Big(1 + \frac{r}{2}\Big)\, x \in K_r(x) \cap (\mathbb{R}^2 \backslash M).$$

Jedes $x \in \mathbb{R}^2$ mit $\|x\| > 1$ und $x \neq \binom{2}{0}$ ist weder Häufungspunkt noch Randpunkt von M, denn es gilt $K_r(x) \cap M = \emptyset$ für $r := \min\{\|x\| - 1, \|x - \binom{2}{0}\|\} > 0$. Insgesamt folgt

$$\begin{aligned} \overset{\circ}{M} &= \{x \in \mathbb{R}^2;\ 0 < \|x\| < 1\}, \\ \overline{M} &= \{x \in \mathbb{R}^2;\ \|x\| \leqslant 1\} \cup \left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}, \\ \partial M &= \{x \in \mathbb{R}^2;\ \|x\| = 1\} \cup \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}. \end{aligned}$$

b) Sei $(V, \|\cdot\|)$ ein normierter Vektorraum, $a \in V, r > 0$. Dann ist $K_r(a)$ offen, denn für $x \in K_r(a)$ gilt nach der Dreiecksungleichung

$$K_\delta(x) \subset K_r(a) \quad \text{für } \delta := r - \|x - a\| > 0.$$



Ist $\|x - a\| > r$, so gilt $K_\delta(x) \cap K_r(a) = \emptyset$ für $\delta := \|x-a\| - r$. Gilt $\|x - a\| = r$, so folgt für $y_\pm := a + (1 \pm \rho/2)(x - a)$ bereits $y_\pm \in K_\rho(x)$ und $y_+ \notin K_r(a)$ sowie $y_- \in K_r(a)$ für alle $0 < \rho < 2$. Man erhält

$$\overline{K_r(a)} = \{x \in V;\ \|x - a\| \leqslant r\}, \quad \partial K_r(a) = \{x \in V;\ \|x - a\| = r\}.$$

c) Für gegebenes $\alpha \in \mathbb{R}$ definiere $f_\alpha : [0,1] \to \mathbb{R}$, $x \mapsto \alpha \cdot \exp(-\alpha^2 x)$. Setze

$$M := \{f_\alpha;\ \alpha \in \mathbb{R}\}.$$

Dann gilt:

- In $\big(\mathcal{C}([0,1]), \|\cdot\|_\infty\big)$ ist M unbeschränkt, M ist nicht offen und $M \subset \partial M$. Denn offenbar ist $\|f_\alpha\|_\infty = |\alpha|$, woraus die Unbeschränktheit folgt. Zu $f_\alpha \in M$ und $r > 0$ betrachte $g \in \mathcal{C}([0,1])$ mit $g(x) = f_\alpha(x) + \frac{r}{2} \cdot x$. Dann $g \in K_r(f_\alpha)$ wegen $\|g - f_\alpha\|_\infty = \frac{r}{2} < r$. Wäre $g \in M$, also $g = f_\beta$ für ein β, so $\alpha\, e^{-\alpha^2 x} + \frac{r}{2} x = \beta\, e^{-\beta^2 x}$ für alle $x \in [0,1]$, was nach zweimaligem Differenzieren zu $\alpha = \beta$ und Widerspruch führt.

- In $\big(\mathcal{C}([0,1]), \|\cdot\|_2\big)$ ist M beschränkt, da

$$\int_0^1 f_\alpha^2(x)\,\mathrm{d}x = \frac{1}{2}(1 - e^{-\alpha^2}) \leqslant \frac{1}{2}.$$

Die Feststellungen im letzten Beispiel sollten vor allem als Warnung dienen, dass jedenfalls in Funktionenräumen seltsame Phänomene nicht auszuschließen sind.

Elementare Eigenschaften der zuvor definierten Begriffe beschreiben wir in dem

(1.10) Satz. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $M \subset V$.*
a) Es gilt

$$\overset{\circ}{M} \subset M \subset \overline{M}, \quad \overline{M} = M \cup \partial M = \overset{\circ}{M} \cup \partial M, \quad \overset{\circ}{M} \cap \partial M = \emptyset.$$

b) $\emptyset$ und V sind offen. Beliebige Vereinigungen und endliche Durchschnitte offener Mengen sind wieder offen.
c) M ist genau dann abgeschlossen, wenn $V \backslash M$ offen ist.
d) M ist genau dann offen, wenn $V \backslash M$ abgeschlossen ist.
e) $\emptyset$ und V sind abgeschlossen. Endliche Vereinigungen und beliebige Durchschnitte abgeschlossener Mengen sind wieder abgeschlossen.

Beweis. a) Direkt aus den Definitionen folgt $\overset{\circ}{M} \subset M \subset \overline{M}$. Ist weiter $x \in \overset{\circ}{M}$, so gibt es ein $r > 0$ mit $K_r(x) \subset M$, somit nach Definition $x \notin \partial M$; also $\overset{\circ}{M} \cap \partial M = \emptyset$. Weiter ist $\overset{\circ}{M} \cup \partial M \subset M \cup \partial M$ offensichtlich. Ist $x \in \partial M$ und $x \notin M$, so gibt es zu jeder Umgebung von x einen Punkt aus M, also aus $M \backslash \{x\}$. Damit ist x Häufungspunkt von M und nach Definition $M \cup \partial M \subset \overline{M}$. Es bleibt $\overline{M} \subset \overset{\circ}{M} \cup \partial M$ zu zeigen: Ist $x \in \overline{M}$ nicht innerer Punkt von M, so gibt es in jeder Umgebung von x einen Punkt aus $M \backslash \{x\}$ (da x Häufungspunkt) und einen Punkt aus $V \backslash M$; insgesamt $x \in \partial M$.
b) Aus (1.8) folgt sofort, dass $\emptyset$ und V offen sind, da $K_1(x) \subset V$ für alle $x \in V$.
Seien $I \neq \emptyset$ eine Indexmenge, U_i, $i \in I$, offen und $x \in U := \cup_{i \in I} U_i$. Dann existiert ein $i \in I$ mit $x \in U_i$ und ein $r > 0$ mit $K_r(x) \subset U_i$, da U_i offen ist. Es folgt $K_r(x) \subset U$, so dass U offen ist.
Sei nun $I = \{1, \ldots, n\}$ und $x \in U' := \bigcap_{i=1}^n U_i$. Da alle U_i offen sind, gibt es $r_i > 0$ mit der Eigenschaft $K_{r_i}(x) \subset U_i$, $i = 1, \ldots, n$. Sei $r = \min\{r_1, \ldots, r_n\} > 0$. Wegen

$$K_r(x) \subset K_{r_i}(x) \subset U_i, \; i = 1, \ldots, n,$$

hat man

$$K_r(x) \subset \bigcap_{i=1}^n U_i = U',$$

so dass auch U' offen ist.
c) "$\Rightarrow$" Sei M abgeschlossen und $x \in V \backslash M$. Dann ist x kein Häufungspunkt von M. Somit existiert ein $r > 0$ mit $K_r(x) \cap M = \emptyset$, also $K_r(x) \subset V \backslash M$. Demnach ist $V \backslash M$ offen.
"$\Leftarrow$" Sei $V \backslash M$ offen und $x \in V \backslash M$. Dann existiert ein $r > 0$ mit $K_r(x) \subset V \backslash M$, d. h. $K_r(x) \cap M = \emptyset$. Also ist x kein Häufungspunkt von M. Demnach gehören alle Häufungspunkte von M bereits zu M und M ist abgeschlossen.
d) Die Behauptung folgt aus c) mit $V \backslash (V \backslash M) = M$.
e) Sei $M^c := V \backslash M$. Dann sind $\emptyset = V^c$ und $V = \emptyset^c$ nach b) und c) abgeschlossen. Seien A_i, $i \in I$, abgeschlossen, also A_i^c offen. Dann ist

$$\bigcap_{i \in I} A_i = \left(\bigcup_{i \in I} A_i^c \right)^c$$

nach b) und c) abgeschlossen. Gilt $I = \{1, \dots, n\}$, so ist

$$\bigcup_{i=1}^{n} A_i = \left(\bigcap_{i=1}^{n} A_i^c \right)^c$$

ebenfalls nach b) und c) abgeschlossen. □

f

§2. Konvergenz und Vollständigkeit

In diesem Paragrafen formulieren wir die Konvergenz von Folgen im allgemeineren Kontext von normierten Vektorräumen.

(2.1) Definition. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum.
a) Eine *Folge* in V ist eine Abbildung $\mathbb{N} \to V, k \mapsto x_k$. Man schreibt dafür $(x_k)_{k\geqslant 1}$.
b) Eine Folge $(x_k)_{k\geqslant 1}$ in V heißt *beschränkt*, wenn es ein $r > 0$ gibt, so dass $\|x_k\| < r$ für alle $k \in \mathbb{N}$.
c) Man nennt eine Folge $(x_k)_{k\geqslant 1}$ in V eine CAUCHY-*Folge*, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass $\|x_k - x_l\| < \varepsilon$ für alle $k, l \geqslant N$.
d) Eine Folge $(x_k)_{k\geqslant 1}$ in V heißt *konvergent*, wenn es ein $x \in V$ gibt, so dass zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert mit der Eigenschaft $\|x_k - x\| < \varepsilon$ für alle $k \geqslant N$. Man nennt x *Limes* oder *Grenzwert* der Folge $(x_k)_{k\geqslant 1}$, schreibt $x_k \xrightarrow[k\to\infty]{} x$ oder $\lim\limits_{k\to\infty} x_k = x$ und sagt, dass $(x_k)_{k\geqslant 1}$ gegen x *konvergiert*.
e) Eine *Nullfolge* in V ist eine Folge $(x_k)_{k\geqslant 1}$ in V mit $\lim\limits_{k\to\infty} x_k = 0 \in V$.
f) Eine Folge in V, die nicht konvergiert, heißt *divergent*.

Die obigen Begriffe verallgemeinern die Konvergenzbegriffe von reellen und komplexen Folgen aus Analysis I und auch von Funktionenfolgen aus Kapitel VII. Erste Eigenschaften formulieren wir in dem

(2.2) Lemma. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum.*
a) Eine Folge $(x_k)_{k\geqslant 1}$ in V konvergiert genau dann gegen $x \in V$, wenn es zu jeder Umgebung U von x ein $N \in \mathbb{N}$ gibt, so dass $x_k \in U$ für alle $k \geqslant N$.
b) Wenn eine Folge konvergiert, ist der Grenzwert eindeutig bestimmt.
c) Jede konvergente Folge ist eine CAUCHY-*Folge.*
d) Jede konvergente Folge und jede CAUCHY-*Folge ist beschränkt.*

Beweis. a) "$\Leftarrow$" Für $U = K_\varepsilon(x)$ hat man $x_k \in K_\varepsilon(x)$ genau dann, wenn $\|x_k - x\| < \varepsilon$.
"$\Rightarrow$" Ist U eine Umgebung von x, so gibt es ein $\varepsilon > 0$ mit $K_\varepsilon(x) \subset U$.
b) Sei x ein Grenzwert von $(x_k)_{k\geqslant 1}$. Gilt $x \neq x'$, so existiert nach (1.7) ein $\varepsilon > 0$ mit $K_\varepsilon(x) \cap K_\varepsilon(x') = \emptyset$. Aus $x_k \xrightarrow[k\to\infty]{} x$ folgt $x_k \in K_\varepsilon(x)$ für alle $k \geqslant N$, also $x_k \notin K_\varepsilon(x')$ für alle $k \geqslant N$.

Demnach konvergiert $(x_k)_{k\geqslant 1}$ nicht gegen x'.
c) Sei $\lim\limits_{k\to\infty} x_k = x$ und $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$ mit $\|x_k - x\| < \varepsilon/2$ für alle $k \geqslant N$. Aus der Dreiecksungleichung folgt nun

$$\|x_k - x_l\| = \|x_k - x + x - x_l\| \leqslant \|x_k - x\| + \|x - x_l\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für alle $k, l \geqslant N$. Demnach ist $(x_k)_{k\geqslant 1}$ eine CAUCHY-Folge.
d) Nach c) genügt es, eine CAUCHY-Folge $(x_k)_{k\geqslant 1}$ zu betrachten. Zu $\varepsilon = 1$ existiert ein $N \in \mathbb{N}$ mit $\|x_k - x_l\| < 1$ für alle $k, l \geqslant N$. Man erhält

$$\|x_k\| - \|x_N\| \leqslant \|x_k - x_N\| < 1, \quad \text{also} \quad \|x_k\| < \|x_N\| + 1 \quad \text{für} \quad k \geqslant N$$

aus (1.2). Es folgt für alle $k \in \mathbb{N}$

$$\|x_k\| < r \quad \text{mit } r := 1 + \max\{\|x_1\|, \ldots, \|x_N\|\}. \qquad \square$$

Wie für reelle Folgen hat man Rechenregeln für Grenzwerte, mit sehr ähnlichen Beweisen.

(2.3) Satz. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum.*

a) Sind $(x_k)_{k\geqslant 1}$ und $(y_k)_{k\geqslant 1}$ konvergente Folgen in V, so auch $(x_k + y_k)_{k\geqslant 1}$ und es gilt

$$\lim_{k\to\infty} (x_k + y_k) = \lim_{k\to\infty} x_k + \lim_{k\to\infty} y_k.$$

b) Ist $(x_k)_{k\geqslant 1}$ eine konvergente Folge in V und $(\gamma_k)_{k\geqslant 1}$ eine konvergente Folge in $\mathbb{R}$, so konvergiert $(\gamma_k \cdot x_k)_{k\geqslant 1}$ in V, und es gilt

$$\lim_{k\to\infty} (\gamma_k \cdot x_k) = \lim_{k\to\infty} \gamma_k \cdot \lim_{k\to\infty} x_k.$$

Insbesondere gilt für $\alpha \in \mathbb{R}$:

$$\lim_{k\to\infty} (\alpha \cdot x_k) = \alpha \cdot \lim_{k\to\infty} x_k.$$

Beweis. Sei $x = \lim x_k$; $y = \lim y_k$ und $\gamma = \lim \gamma_k$.
a) Nach Voraussetzung gibt es zu $\varepsilon > 0$ ein $N_1 \in \mathbb{N}$, so dass $\|x_k - x\| < \frac{\varepsilon}{2}$ für alle $k \geqslant N_1$, sowie ein $N_2 \in \mathbb{N}$ mit $\|y_k - y\| < \frac{\varepsilon}{2}$ für alle $k \geqslant N_2$. Mit der Dreiecksungleichung folgt für alle $k \geqslant N := \max\{N_1, N_2\}$ dann

$$\|x_k + y_k - (x + y)\| \leqslant \|x_k - x\| + \|y_k - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

b) Beide Folgen sind konvergent, also (nach III(1.7) bzw. nach (2.2)) beschränkt. Es gibt also ein reelles $M > 0$ derart, dass $\|x_k\| \leqslant M$ und $|\gamma_k| \leqslant M$ für alle $k \in \mathbb{N}$. Zu $\varepsilon > 0$ gibt es dann ein $N_1 \in \mathbb{N}$ mit $\|x_k - x\| < \frac{\varepsilon}{2M}$ für alle $k \geqslant N_1$, sowie ein $N_2 \in \mathbb{N}$ mit $|\gamma_k - \gamma| < \frac{\varepsilon}{2M}$ für alle $k \geqslant N_2$. Für alle $k \geqslant N := \max\{N_1, N_2\}$ folgt dann

$$\begin{aligned}
&\|\gamma_k \cdot x_k - \gamma \cdot x\| = \|\gamma_k \cdot x_k - \gamma_k \cdot x + \gamma_k \cdot x - \gamma \cdot x\| \\
&\leqslant \|\gamma_k \cdot x_k - \gamma_k \cdot x\| + \|\gamma_k \cdot x - \gamma \cdot x\| \\
&= |\gamma_k| \cdot \|x_k - x\| + |\gamma_k - \gamma| \cdot \|x\| \\
&< M \cdot \frac{\varepsilon}{2M} + \frac{\varepsilon}{2M} \cdot M = \varepsilon.
\end{aligned}$$

Die letzte Aussage folgt mit der konstanten Folge $(\alpha)_{k\geqslant 1}$. $\square$

Wichtig ist, dass man mit Hilfe von Folgen auch die Abgeschlossenheit von Mengen nachweisen kann.

(2.4) Satz. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann sind für $x \in V$ und $M \subset V$ äquivalent:*

(i) $x \in \overline{M}$.

(ii) Es gibt eine Folge $(x_k)_{k\geqslant 1}$ mit $x_k \in M$ für alle $k \in \mathbb{N}$ und $\lim\limits_{k\to\infty} x_k = x$.

Beweis. "(i) $\Rightarrow$ (ii)" Für $x \in M$ setzt man $x_k = x$ für alle $k \in \mathbb{N}$. Ist $x \notin M$, so ist x ein Häufungspunkt von M. Also existiert zu jedem $k \in \mathbb{N}$ ein $x_k \in K_{1/k}(x) \cap M$. Aus $\|x_k - x\| < \frac{1}{k}$ folgt $\lim\limits_{k\to\infty} x_k = x$.
"(ii) $\Rightarrow$ (i)" Sei $(x_k)_{k\geqslant 1}$ eine Folge in M mit $\lim\limits_{k\to\infty} x_k = x$. Zu $r > 0$ existiert dann ein $N \in \mathbb{N}$ mit $x_k \in K_r(x)$ für alle $k \geqslant N$. Aus $x_N \in M$ folgt $K_r(x) \cap M \neq \emptyset$. Also gilt $x \in M$ oder x ist ein Häufungspunkt von M, d.h. $x \in \overline{M}$. $\square$

Wir betrachten nun zahlreiche

(2.5) Beispiele. a) Sei $(V, \|\cdot\|) = (\mathbb{R}^n, \|\cdot\|_\infty)$ aus (1.3) d). Sei $(x_k)_{k\geqslant 1}$ eine Folge von Vektoren im $\mathbb{R}^n$ mit $x_k = (x_{k,1}, \dots, x_{k,n})^t$. Die Folge $(x_k)_{k\geqslant 1}$ konvergiert genau dann gegen $a = (a_1, \dots, a_n)^t \in \mathbb{R}^n$, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt mit

$$\|x_k - a\|_\infty = \max\{|x_{k,j} - a_j| \mid 1 \leqslant j \leqslant n\} < \varepsilon \quad \text{für alle } k \geqslant N.$$

Das ist genau dann der Fall, wenn jede reelle Komponentenfolge $(x_{k,j})_{k\geqslant 1}$ in $\mathbb{R}$ gegen a_j konvergiert für $j = 1, \dots, n$. Es gilt z. B.

$$\begin{pmatrix} (1/2)^k \\ (1+1/k)^k \\ \sqrt[k]{2} \end{pmatrix} \xrightarrow[k\to\infty]{} \begin{pmatrix} 0 \\ e \\ 1 \end{pmatrix}, \quad \text{da } \left(\frac{1}{2}\right)^k \to 0,\ \left(1+\frac{1}{k}\right)^k \to e,\ \sqrt[k]{2} \to 1.$$

b) Sei $(V, \|\cdot\|) = (\mathcal{C}([a,b]), \|\cdot\|_\infty)$ aus (1.3) e). Eine Folge $(f_k)_{k\geqslant 1}$ in $\mathcal{C}([a,b])$ konvergiert genau dann gegen $f \in \mathcal{C}([a,b])$, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $k \geqslant N$ gilt

$$\|f_k - f\|_\infty \leqslant \varepsilon, \quad \text{d.h.} \quad |f_k(x) - f(x)| \leqslant \varepsilon \quad \text{für alle } x \in [a,b].$$

Das bedeutet $f_k \xrightarrow[glm]{} f$.

c) Sei $(V, \|\cdot\|) = (\mathcal{C}([0,2]), \|\cdot\|)$ mit $\|f\| = \sqrt{\int_0^2 f(t)^2\,dt}$ aus (1.3) f). Wir betrachten $(f_k)_{k\geqslant 1}$ in $\mathcal{C}([0,2])$ mit

$$f_k(t) = \begin{cases} t^k, & \text{falls } 0 \leqslant t \leqslant 1, \\ 1, & \text{falls } 1 \leqslant t \leqslant 2. \end{cases}$$

Für $k \geqslant l$ gilt dann

$$\begin{aligned}\|f_k - f_l\|^2 &= \int_0^2 (f_k(t) - f_l(t))^2\,dt = \int_0^1 (t^k - t^l)^2\,dt \\ &= \left(\frac{1}{2k+1}t^{2k+1} - \frac{2}{k+l+1}t^{k+l+1} + \frac{1}{2l+1}t^{2l+1}\right)\Big|_0^1 \\ &= \frac{1}{2k+1} - \frac{2}{k+l+1} + \frac{1}{2l+1} \leqslant \frac{1}{2l+1}.\end{aligned}$$

Also ist $(f_k)_{k\geqslant 1}$ eine CAUCHY-Folge in $(\mathcal{C}([0,2]), \|\cdot\|)$. Angenommen es gibt ein $f \in \mathcal{C}([0,2])$, so dass $(f_k)_{k\geqslant 1}$ gegen f in $(\mathcal{C}([0,2]); \|\cdot\|)$ konvergiert. Dann gilt mit VII(1.22)

$$\begin{aligned}0 &\leqslant \sqrt{\int_0^1 f(t)^2\,dt} \leqslant \sqrt{\int_0^1 (f(t) - f_k(t))^2\,dt} + \sqrt{\int_0^1 f_k(t)^2\,dt} \\ &\leqslant \|f - f_k\| + \frac{1}{\sqrt{2k+1}} \xrightarrow[k\to\infty]{} 0.\end{aligned}$$

Man erhält $\int_0^1 f(t)^2\,dt = 0$ und somit $f(t) = 0$ für alle $0 \leqslant t \leqslant 1$ wegen der Stetigkeit von f. Andererseits gilt

$$0 \leqslant \int_1^2 (f(t) - 1)^2\,dt = \int_1^2 (f(t) - f_k(t))^2\,dt \leqslant \|f - f_k\|^2 \xrightarrow[k\to\infty]{} 0,$$

also $\int_1^2 (f(t) - 1)^2\,dt = 0$ und somit $f(t) = 1$ für alle $1 \leqslant t \leqslant 2$ wegen der Stetigkeit von f. Das ist ein Widerspruch, also konvergiert $(f_k)_{k\geqslant 1}$ nicht in $(\mathcal{C}([0,2]), \|\cdot\|)$.
Man beachte, dass die Konvergenz in $(\mathcal{C}([0,2]), \|\cdot\|)$ nicht die gleichmäßige Konvergenz ist. Für $g_k(t) = \left(\frac{t}{2}\right)^k$ gilt analog

$$\|g_k - 0\| = \frac{1}{\sqrt{4k+2}} \xrightarrow[k\to\infty]{} 0, \quad g_k \xrightarrow[\text{glm}]{\;\;/\;\;} 0.$$

Also konvergiert $(g_k)_{k\geqslant 1}$ in $(\mathcal{C}([0,2]), \|\cdot\|)$ gegen die Nullfunktion, aber $(g_k)_{k\geqslant 1}$ konvergiert weder punktweise noch gleichmäßig gegen die Nullfunktion.

Im Hinblick auf das letzte Beispiel und (2.2) c) formulieren wir die

(2.6) Definition. Ein normierter Vektorraum $(V, \|\cdot\|)$ heißt *vollständig* oder BANACH-*Raum*, wenn in $(V, \|\cdot\|)$ jede CAUCHY-Folge konvergiert.

Wir haben sogleich einige

(2.7) Beispiele. a) $(\mathbb{R}, |\cdot|)$ ist nach III(2.9) vollständig.
b) $(\mathbb{R}^n, \|\cdot\|_\infty)$ ist nach a) und (2.5) a) vollständig.
c) $(\mathcal{C}([a,b]), \|\cdot\|_\infty)$, $a < b$, ist nach (2.5) b) und IX(2.4) vollständig.
d) $(\mathcal{C}([a,b]), \|\cdot\|)$, $a < b$, aus (2.5) c) ist nicht vollständig.

Nun wollen wir in mehreren Schritten zeigen, dass $\mathbb{R}^n$ mit einer beliebigen Norm ebenfalls vollständig ist.

(2.8) Satz von Bolzano-Weierstrass. *Gegeben sei eine beschränkte Folge $(x_k)_{k\geqslant 1}$ in $(\mathbb{R}^n, \|\cdot\|_\infty)$. Dann existiert eine Teilfolge von $(x_k)_{k\geqslant 1}$, die in $(\mathbb{R}^n, \|\cdot\|_\infty)$ konvergiert.*

Beweis. Wir verwenden eine Induktion nach n, wobei $n = 1$ nach III(2.4) klar ist. Sei nun $(x_k)_{k\geqslant 1}$ eine Folge in $\mathbb{R}^{n+1}$, die bezüglich $\|\cdot\|_\infty$ beschränkt ist. Sei $x_k = \binom{y_k}{x_{k,n+1}}, y_k \in \mathbb{R}^n$. Wegen

$$\|y_k\|_\infty \leqslant \max\{\|y_k\|_\infty,\ |x_{k,n+1}|\} = \|x_k\|_\infty$$

sind $(y_k)_{k\geqslant 1}$ und $(x_{k,n+1})_{k\geqslant 1}$ beschränkt. Nach Induktionsvoraussetzung existiert eine Teilfolge $(y_{k_j})_{j\geqslant 1}$, die in $(\mathbb{R}^n, \|\cdot\|_\infty)$ gegen $y \in \mathbb{R}^n$ konvergiert. Weil auch $(x_{k_j,n+1})_{j\geqslant 1}$ beschränkt ist, existiert eine Teilfolge $(x_{k_{j_\nu},n+1})_{\nu\geqslant 1}$, die gegen ein $\xi \in \mathbb{R}$ konvergiert. Weil auch $(y_{k_{j_\nu}})_{\nu\geqslant 1}$ gegen y konvergiert, konvergiert $(x_{k_{j_\nu}})_{\nu\geqslant 1}$ nach (2.5) a) in $(\mathbb{R}^{n+1}, \|\cdot\|_\infty)$ gegen $\binom{y}{\xi} \in \mathbb{R}^{n+1}$. □

Nun beweisen wir die Äquivalenz der Normen auf $\mathbb{R}^n$.

(2.9) Satz. *Sind $\|\cdot\|$ und $\|\cdot\|'$ Normen auf $\mathbb{R}^n$, so gibt es $\alpha > 0$ und $\beta > 0$ mit der Eigenschaft*

$$\alpha\|x\| \leqslant \|x\|' \leqslant \beta\|x\| \quad \textit{für alle} \quad x \in \mathbb{R}^n.$$

Beweis. Es genügt, die Behauptung für $\|\cdot\|' = \|\cdot\|_\infty$ zu beweisen, da aus $\alpha\|x\| \leqslant \|x\|_\infty \leqslant \beta\|x\|$ und $\alpha'\|x\|' \leqslant \|x\|_\infty \leqslant \beta'\|x\|'$ auch $\frac{\alpha}{\beta'}\|x\| \leqslant \|x\|' \leqslant \frac{\beta}{\alpha'}\|x\|$ folgt.
Sei $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$, wobei $e_1, \ldots, e_n$ die kanonische Basis von $\mathbb{R}^n$ sei. Dann gilt

$$\|x\| = \left\|\sum_{j=1}^n x_j e_j\right\| \leqslant \sum_{j=1}^n |x_j| \cdot \|e_j\| \leqslant \max\{|x_1|, \ldots, |x_n|\} \cdot \sum_{j=1}^n \|e_j\| = \|x\|_\infty \cdot \sum_{j=1}^n \|e_j\|.$$

Daraus folgt

$$\alpha\|x\| \leqslant \|x\|_\infty \quad \text{mit } \alpha = \left(\sum_{j=1}^n \|e_j\|\right)^{-1}.$$

Die zweite Ungleichung beweisen wir indirekt und nehmen an, dass es zu jedem $\beta > 0$ ein $x \in \mathbb{R}^n$ mit $\|x\|_\infty > \beta\|x\|$ gilt. Insbesondere gibt es zu $\beta = k \in \mathbb{N}$ einen Vektor $y_k \in \mathbb{R}^n$ mit $\|y_k\|_\infty > k\|y_k\| > 0$. Für $x_k := \frac{1}{\|y_k\|_\infty} y_k \in \mathbb{R}^n$ gilt dann

$$\|x_k\|_\infty = 1, \quad \|x_k\| = \frac{1}{\|y_k\|_\infty}\|y_k\| < \frac{1}{k} \quad \text{für alle } k \in \mathbb{N}.$$

Es folgt $x_k \to 0$ in $(\mathbb{R}^n, \|\cdot\|)$ und $(x_k)_{k\geqslant 1}$ ist in $(\mathbb{R}^n, \|\cdot\|_\infty)$ durch 1 beschränkt. Nach (2.8) gibt es eine Teilfolge $(x_{k_j})_{j\geqslant 1}$, die in $(\mathbb{R}^n, \|\cdot\|_\infty)$ gegen ein x^* konvergiert, d. h.

$$\|x_{k_j} - x^*\|_\infty \xrightarrow[j\to\infty]{} 0.$$

Aus dem bewiesenen Teil ergibt sich

$$0 \leqslant \|x_{k_j} - x^*\| \leqslant \frac{1}{\alpha}\|x_{k_j} - x^*\|_\infty \xrightarrow[j\to\infty]{} 0,$$

also $x_{k_j} \to x^*$ in $(\mathbb{R}^n, \|\cdot\|)$. Wegen $x_k \xrightarrow[k\to\infty]{} 0$ in $(\mathbb{R}^n, \|\cdot\|)$ folgt $x_{k_j} \to 0$ in $(\mathbb{R}^n, \|\cdot\|)$ und dann $x^* = 0$ aus der Eindeutigkeit des Limes in (2.2). Wegen $\|x_{k_j}\|_\infty = \|x_{k_j} - 0\|_\infty = 1$ erhalten wir einen Widerspruch. □

Als direkte Folgerung notieren wir das

(2.10) Korollar. *Seien $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf $\mathbb{R}^n$, $a = (a_1, \dots, a_n)^t \in \mathbb{R}^n$, $(x_k)_{k\geqslant 1}$ eine Folge in $\mathbb{R}^n$ mit $x_k = (x_{k,1}, \dots, x_{k,n})^t$, $k \in \mathbb{N}$. Dann gilt*
a) $M \subset \mathbb{R}^n$ ist genau dann offen bzw. abgeschlossen bzw. beschränkt bzw. eine Umgebung von a in $(\mathbb{R}^n, \|\cdot\|)$, wenn M offen bzw. abgeschlossen bzw. beschränkt bzw. eine Umgebung von a in $(\mathbb{R}^n, \|\cdot\|')$ ist.
b) Es sind äquivalent:

(i) $(x_k)_{k\geqslant 1}$ konvergiert gegen a in $(\mathbb{R}^n, \|\cdot\|)$.

(ii) $(x_k)_{k\geqslant 1}$ konvergiert gegen a in $(\mathbb{R}^n, \|\cdot\|')$.

(iii) $(x_{k,j})_{k\geqslant 1}$ konvergiert in $\mathbb{R}$ gegen a_j für $j = 1, \dots, n$.

Beweis. a) Wir verwenden (2.9) und

$$K_{r/\beta}^{(\mathbb{R}^n,\|\cdot\|)}(a) \subset K_r^{(\mathbb{R}^n,\|\cdot\|')}(a) \subset K_{r/\alpha}^{(\mathbb{R}^n,\|\cdot\|)}(a),$$

denn $\|x - a\| < \frac{r}{\beta}$ impliziert $\|x - a\|' \leqslant \beta\|x - a\| < r$ und $\|x - a'\| < r$ führt analog zu $\|x - a\| \leqslant \frac{1}{\alpha}\|x - a\|' < \frac{r}{\alpha}$. Daraus folgt die Behauptung direkt mit den Definitionen.
b) Wir verwenden wieder (2.9) und (2.5) a) für $(\mathbb{R}^n, \|\cdot\|_\infty)$ in (iii). □

Aus (2.10) und (2.7) b) folgern wir das

(2.11) Korollar. *Ist $\|\cdot\|$ eine beliebige Norm auf $\mathbb{R}^n$, so ist $(\mathbb{R}^n, \|\cdot\|)$ vollständig.*

Wenn wir von Konvergenz, Offenheit oder Abgeschlossenheit in $\mathbb{R}^n$ reden, so meinen wir stets $(\mathbb{R}^n, \|\cdot\|)$ mit einer beliebigen Norm $\|\cdot\|$ auf $\mathbb{R}^n$. Entsprechendes gilt für beliebige endlich dimensionale $\mathbb{R}$-Vektorräume.

§3. Stetige Funktionen

In diesem Paragrafen übertragen wir das Konzept der Stetigkeit von Funktionen auf normierte Vektorräume.

(3.1) Definition. Seien $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume, $M \subset V$, $a \in M$ und $f : M \to V'$ eine Abbildung. Man nennt f *stetig in* a, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in M$ mit $\|x - a\| < \delta$ gilt $\|f(x) - f(a)\|' < \varepsilon$.
f heißt *stetig*, wenn f in jedem Punkt $a \in M$ stetig ist. f heißt *gleichmäßig stetig*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $a, x \in M$ mit $\|x - a\| < \delta$ auch $\|f(x) - f(a)\|' < \varepsilon$ gilt.

Aus der gleichmäßigen Stetigkeit folgt natürlich trivialerweise die Stetigkeit. Eine Charakterisierung der Stetigkeit geben wir in dem folgenden

(3.2) Satz. *Seien $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume, $M \subset V, a \in M$ und $f : M \to V'$ eine Abbildung. Dann sind äquivalent:*

(i) f ist stetig in a.

(ii) Zu jeder Umgebung U' von $f(a)$ existiert eine Umgebung U von a mit $f(M \cap U) \subset U'$.

(iii) Für jede Folge $(x_k)_{k\geqslant 1}$ mit $x_k \in M$ für alle $k \in \mathbb{N}$ und $\lim\limits_{k\to\infty} x_k = a$ gilt $\lim\limits_{k\to\infty} f(x_k) = f(a)$.

Man nennt (iii) das *Folgenkriterium der Stetigkeit*. Falls a ein Häufungspunkt von M ist, schreibt man für (iii) auch
$$\lim_{x\to a} f(x) = f(a).$$

Beweis. "(i) $\Rightarrow$ (ii)" Zu jeder Umgebung U' von $f(a)$ existiert ein $\varepsilon > 0$ mit der Eigenschaft $K_\varepsilon^{(V',\|\cdot\|')}(f(a)) \subset U'$. Nun wählt man zu diesem ε ein $\delta > 0$ gemäß (3.1). Mit $U = K_\delta^{(V,\|\cdot\|)}(a)$ folgt dann aus (3.1)
$$f(M \cap U) \subset K_\varepsilon^{(V',\|\cdot\|')}(f(a)) \subset U'.$$
"(ii) $\Rightarrow$ (iii)" Sei $(x_k)_{k\geqslant 1}$ eine Folge in M mit $\lim\limits_{k\to\infty} x_k = a$ und U' eine Umgebung von $f(a)$. Nach Voraussetzung existiert eine Umgebung U von a mit $f(U \cap M) \subset U'$. Wegen $x_k \to a$ existiert ein $N \in \mathbb{N}$ mit $x_k \in U \cap M$ für alle $k \geqslant N$. Es folgt $f(x_k) \in U'$ für alle $k \geqslant N$, also $f(x_k) \xrightarrow[k\to\infty]{} f(a)$.
"(iii) $\Rightarrow$ (i)" Angenommen, f ist nicht stetig in a. Dann gibt es ein $\varepsilon^* > 0$, so dass zu jedem $\delta = \frac{1}{k}, k \in \mathbb{N}$, ein $x_k \in M$ existiert mit $\|x_k - a\| < \frac{1}{k}$ und $\|f(x_k) - f(a)\|' \geqslant \varepsilon^*$. Es folgt $x_k \xrightarrow[k\to\infty]{} a$, aber $f(x_k) \underset{k\to\infty}{\not\longrightarrow} f(a)$ im Widerspruch zur Voraussetzung. □

Aus dem Folgenkriterium erhält man unter Benutzung von V, §1 und Satz (2.3):

(3.3) Satz. *Seien $(V, \|\cdot\|)$, $(V', \|\cdot\|')$ und $(V'', \|\cdot\|'')$ jeweils normierte Vektorräume $M \subset V$ und $f : M \to V'$, $g : M \to V'$, $h : M' \to V''$ sowie $\varphi : M \to \mathbb{R}$ Abbildungen. Dann gilt:*
a) Sind f, g und φ stetig in $a \in V$, so sind auch $\alpha f + \beta g$, $\alpha, \beta \in \mathbb{R}$, $\varphi \cdot f$ und im Fall $\varphi(a) \neq 0$ auch $\frac{1}{\varphi} \cdot f$ stetig in a.
b) Ist f stetig in a und h stetig in $f(a)$ mit $f(M) \subset M'$, so ist $h \circ f : M \to V''$ stetig in a.

Nun diskutieren wir einige

(3.4) Beispiele. a) Die *Projektionen* $\pi_j : \mathbb{R}^n \to \mathbb{R}$, $x = (x_1, \dots, x_n)^t \mapsto x_j$, sind für $j = 1, \dots, n$ stetig. Zu $\varepsilon > 0$ wählt man $\delta = \varepsilon$ und aus $\|x - a\|_\infty < \varepsilon$ folgt

$$|\pi_j(x) - \pi_j(a)| = |x_j - a_j| \leqslant \|x - a\|_\infty < \varepsilon.$$

Wegen (2.10) hängt die Stetigkeit nach dem Folgenkriterium nicht von der Wahl der Norm ab.
b) Ein *Polynom* $p : \mathbb{R}^n \to \mathbb{R}$ ist eine endliche Linearkombination von Funktionen der Form $\mathbb{R}^n \to \mathbb{R}$, $x = (x_1, \dots, x_n)^t \mapsto x_1^{\nu_1} \cdot \ldots \cdot x_n^{\nu_n} = \pi_1(x)^{\nu_1} \cdot \ldots \cdot \pi_n(x)^{\nu_n}$, $\nu_1, \dots, \nu_n \in \mathbb{N}_0$, also stetig nach a) und (3.3).
c) Eine *rationale Funktion* $f : M \to \mathbb{R}$, $M \subset \mathbb{R}^n$, ist gegeben durch $f(x) = \frac{P(x)}{Q(x)}$, wobei $P(x), Q(x)$ Polynomfunktionen sind und $Q(x) \neq 0$ für alle $x \in M$. Dann ist f nach b) und (3.3) stetig.
d) Eine lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^m$, $x \mapsto Ax$, $A \in \mathbb{R}^{m \times n}$ ist stetig. Dazu sei nämlich $c \geqslant \|A\|_\infty = \max\{|a_{ij}|;\ 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n\}$. Dann gilt

$$\|Ax\|_\infty = \left\| \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} \right\|_\infty \leqslant nc \cdot \|x\|_\infty.$$

Zu $\varepsilon > 0$ wählt man $\delta := \frac{\varepsilon}{nc}$. Aus $\|x - y\|_\infty < \delta$ folgt dann

$$\|Ax - Ay\|_\infty = \|A(x - y)\|_\infty \leqslant nc \cdot \|x - y\|_\infty < \varepsilon.$$

e) $f : \mathbb{R}^2 \to \mathbb{R}$, $x \mapsto x_1^2 + x_2^2$, ist als Polynom stetig. Der Graph $\{\binom{x}{f(x)} \in \mathbb{R}^3;\ x \in \mathbb{R}^2\}$ ist ein *Rotationsparaboloid*: Die Parabel $x_3 = x_1^2$ wird um die x_3-Achse gedreht.
f) $g : \mathbb{R}^2 \to \mathbb{R},\ x \mapsto \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$
ist als rationale Funktion stetig in $\mathbb{R}^2 \backslash \{0\}$. Wegen

$$\begin{pmatrix} 1/k \\ 1/k \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad g\left(\begin{pmatrix} 1/k \\ 1/k \end{pmatrix} \right) = \frac{1}{2} \underset{k \to \infty}{\not\longrightarrow} 0 = g(0)$$

ist g nicht stetig in 0.
g) $h : \mathbb{R}^2 \to \mathbb{R},\ x \mapsto \begin{cases} \frac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$
ist als Verkettung stetiger Funktionen stetig in $\mathbb{R}^2 \backslash \{0\}$. Für $x \neq 0$ gilt $|x_1| \leqslant \|x\|_2$ und $|x_2| \leqslant \|x\|_2$, also $|h(x)| \leqslant \frac{\|x\|_2^2}{\|x\|_2} = \|x\|_2$.
Zu $\varepsilon > 0$ wählt man $\delta = \varepsilon$. Aus $\|x\|_2 < \varepsilon$ folgt $|h(x)| \leqslant \|x\|_2 < \varepsilon$. Also ist h auch stetig in 0.
h) $F : \mathbb{R}^2 \to \mathbb{R},\ x \mapsto \begin{cases} \frac{1}{(x_1 - x_2)^2}, & \text{falls } x_1 \neq x_2, \\ 0, & \text{falls } x_1 = x_2, \end{cases}$

ist als rationale Funktion stetig auf $\mathbb{R} \backslash \{\binom{\alpha}{\alpha};\ \alpha \in \mathbb{R}\}$. Für $\alpha \in \mathbb{R}$ gilt

$$\begin{pmatrix} \alpha + 1/k \\ \alpha \end{pmatrix} \to \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}, \quad F\left(\begin{pmatrix} \alpha + 1/k \\ \alpha \end{pmatrix} \right) = k^2 \underset{k \to \infty}{\not\longrightarrow} 0 = F\left(\begin{pmatrix} \alpha \\ \alpha \end{pmatrix} \right).$$

Also ist F nicht stetig in $\binom{\alpha}{\alpha}, \alpha \in \mathbb{R}$.
i) Wir betrachten $(\mathcal{C}([a,b]), \|\cdot\|_\infty)$ und

$$I : \mathcal{C}([a,b]) \to \mathcal{C}([a,b]), \quad f \mapsto F, \quad F(x) := \int_a^x f(t)\,dt.$$

Es gilt

$$|F(x)| \leqslant \int_a^x |f(t)|\,dt \leqslant (x-a) \cdot \|f\|_\infty, \quad x \in [a,b],$$

also $\|F\|_\infty \leqslant (b-a) \cdot \|f\|_\infty$. Aus $\|f-g\|_\infty < \frac{\varepsilon}{b-a}$ folgt also

$$\|I(f) - I(g)\|_\infty \leqslant (b-a) \cdot \|f-g\|_\infty < \varepsilon,$$

so dass I stetig ist.

Wir leiten ein erstes Kriterium her.

(3.5) Lemma. *Sei $M \subset \mathbb{R}^n$. Eine Abbildung*

$$f : M \to \mathbb{R}^m,\ x \mapsto (f_1(x), \dots, f_m(x))^t,\ f_j = \pi_j \circ f : M \to \mathbb{R},$$

ist genau dann stetig in $a \in M$, wenn alle $f_j, j = 1, \dots, m$, stetig in a sind.

Beweis. "$\Rightarrow$" Mit f und π_j nach (3.4) ist auch $f_j = \pi_j \circ f$ nach (3.3) stetig in a.
"$\Leftarrow$" Zu $\varepsilon > 0$ existiert ein $\delta_j > 0$, so dass aus $\|x-a\|_\infty < \delta_j$ stets $|f_j(x) - f_j(a)| < \varepsilon$ folgt. Sei $\delta = \min\{\delta_1, \dots, \delta_m\}$. Aus $\|x-a\|_\infty < \delta$ erhält man dann $|f_j(x) - f_j(a)| < \varepsilon$ für $j = 1, \dots, m$, also $\|f(x) - f(a)\|_\infty < \varepsilon$. Daher ist f stetig in a. □

Besonders einfach ist auch das folgende Kriterium.

(3.6) Lemma. *Seien $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume, $D \subset V$ und $f : D \to V'$ eine Abbildung, so dass ein $\lambda > 0$ existiert mit der Eigenschaft*

$$\|f(x) - f(y)\|' \leqslant \lambda \cdot \|x-y\| \quad \textit{für alle} \ \ x, y \in D.$$

Dann ist f gleichmäßig stetig auf D. Man nennt f dann LIPSCHITZ*-stetig auf D.*

Beweis. Sei $\varepsilon > 0$ und $\delta := \frac{\varepsilon}{\lambda}$. Für $x, y \in D$ mit $\|x-y\| < \delta$ folgt

$$\|f(x) - f(y)\|' \leqslant \lambda \cdot \|x-y\| < \varepsilon.$$

Also ist f gleichmäßig stetig auf D. □

Nach (3.4) d) sind insbesondere lineare Abbildungen von $\mathbb{R}^n$ in $\mathbb{R}^m$ Lipschitz-stetig. Ein wichtiger Spezialfall liegt vor, wenn die LIPSCHITZ-*Konstante* $\lambda < 1$ erfüllt.

(3.7) BANACHscher Fixpunktsatz. *Sei $(V, \|\cdot\|)$ ein vollständiger, normierter Vektorraum, $D \subset V$ abgeschlossen und $f : D \to D$ eine Abbildung, so dass es ein $0 < \lambda < 1$ gibt mit der Eigenschaft*

$$(*) \qquad \|f(x) - f(y)\| \leqslant \lambda \cdot \|x - y\| \quad \text{für alle} \quad x, y \in D.$$

Dann hat f genau einen Fixpunkt in D, d. h., es gibt genau ein $\overline{x} \in D$ mit $f(\overline{x}) = \overline{x}$. Ist $x_0 \in D$ beliebig, so konvergiert die Iterationsfolge

$$(x_k)_{k\geqslant 1}, x_k := f(x_{k-1}), k \in \mathbb{N},$$

gegen $\overline{x}$.

Ein f mit der Eigenschaft $(*)$ nennt man eine *Kontraktion*.

Beweis. Eindeutigkeit: Gilt $f(\overline{x}) = \overline{x}$ und $f(x') = x'$ für $\overline{x}, x' \in D$, so folgt aus

$$\|\overline{x} - x'\| = \|f(\overline{x}) - f(x')| \leqslant \lambda \cdot \|\overline{x} - x'\|$$

und $\lambda < 1$ sofort $\|\overline{x} - x'\| = 0$, also $\overline{x} = x'$ und damit die Eindeutigkeit.
Existenz: Sei $x_0 \in D$ beliebig und $x_k = f(x_{k-1})$ für alle $k \in \mathbb{N}$. Dann zeigt man per Induktion

$$\|x_{m+k+1} - x_{m+k}\| \leqslant \lambda^k \cdot \|x_{m+1} - x_m\| \quad \text{für alle} \quad m, k \in \mathbb{N}_0.$$

Dabei ist $k = 0$ trivial und

$$\|x_{m+k+2} - x_{m+k+1}\| = \|f(x_{m+k+1}) - f(x_{m+k})\| \leqslant \lambda \cdot \|x_{m+k+1} - x_{m+k}\| \leqslant \lambda^{k+1} \cdot \|x_{m+1} - x_m\|.$$

Für $k, m \in \mathbb{N}$ erhält man damit (und mit dem "Teleskop-Trick")

$$\|x_{m+k} - x_m\| = \left\|\sum_{j=0}^{k-1}(x_{m+j+1} - x_{m+j})\right\| \leqslant \sum_{j=0}^{k-1} \|x_{m+j+1} - x_{m+j}\|$$

$$\leqslant \sum_{j=0}^{k-1} \lambda^j \cdot \|x_{m+1} - x_m\| \leqslant \frac{1}{1-\lambda} \cdot \|x_{m+1} - x_m\| \leqslant \frac{\lambda^m}{1-\lambda}\|x_1 - x_0\| \xrightarrow[m\to\infty]{} 0.$$

Also ist $(x_k)_{k\geqslant 1}$ eine CAUCHY-Folge in $(V, \|\cdot\|)$, die gegen ein $\overline{x} \in V$ konvergiert, weil der normierte Raum $(V, \|\cdot\|)$ vollständig ist. Es gilt $\overline{x} \in \overline{D}$ nach (2.3) und $\overline{D} = D$, weil D abgeschlossen ist. Da f nach (3.6) stetig in D ist, folgt

$$f(\overline{x}) = f(\lim_{k\to\infty} x_k) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} x_{k+1} = \overline{x}.$$

□

(3.8) Beispiel. Es sei $g : [0, \infty) \to \mathbb{R}$, $g(x) = \frac{3}{2} - e^{-x/2}$. Dann gibt es genau ein $x^* \in [0, \infty)$ mit $g(x^*) = x^*$. Dies weisen wir mit Hilfe des BANACHschen Fixpunktsatzes für g auf der abgeschlossenen Menge $[0, \infty)$ nach.
Weil $g'(x) = \frac{1}{2}e^{-x/2}$, also $|g'(x)| \leqslant \frac{1}{2}$ für alle $x \geqslant 0$, folgt mit dem Mittelwertsatz der Differentialrechnung

$$|g(x_2) - g(x_1)| \leqslant \frac{1}{2}|x_2 - x_1| \quad \text{für alle} \quad x_1, x_2 \geqslant 0.$$

Also ist g eine Kontraktion mit $\lambda = \frac{1}{2}$.

Das Beispiel hat eher illustrativen Charakter. Wir werden später aber sehr substantielle Anwendungen von (3.7) sehen.

Eine weitere Klasse stetiger Funktionen beschreiben wir in dem

(3.9) Lemma. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann ist die Norm $\|\cdot\| : V \to \mathbb{R}$ gleichmäßig stetig.*

Beweis. Für $x, y \in V$ gilt $\big|\|x\| - \|y\|\big| \leqslant \|x - y\|$ nach der 2. Dreiecksungleichung (1.2). Dann folgt die gleichmäßige Stetigkeit aus (3.6). □

Als Anwendung formulieren wir das

(3.10) Korollar. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann gilt für $A \subset V, A \neq \emptyset$:*

a) Die Abbildung

$$d(\cdot, A) : V \to \mathbb{R}, \; x \mapsto d(x, A) := \inf\{\|x - y\|; y \in A\},$$

ist gleichmäßig stetig.
b) Es gilt

$$\overline{A} = \{x \in V;\ d(x, A) = 0\}.$$

Beweis. a) Für $x, y \in V$ und $a \in A$ gilt nach der Dreiecksungleichung

$$d(x, A) \leqslant \|x - a\| \leqslant \|x - y\| + \|y - a\|.$$

Daraus folgt

$$d(y, A) = \inf\{\|y - a\|;\ a \in A\} \geqslant d(x, A) - \|x - y\|$$

Man erhält

$$d(x, A) - d(y, A) \leqslant \|x - y\|$$

und aus Symmetriegründen

$$|d(x, A) - d(y, A)| \leqslant \|x - y\|.$$

Dann ergibt sich die gleichmäßige Stetigkeit aus (3.6).
b) Zu $x \in \overline{A}$ gibt es nach (2.4) eine Folge $(a_k)_{k\geqslant 1}$ in A, die gegen x konvergiert. Es folgt

$$d(x, A) = \inf\{\|x - a\|;\ a \in A\} \leqslant \inf\{\|x - a_k\|;\ k \in \mathbb{N}\} = 0,$$

also $d(x, A) = 0$.
Gilt $d(x, A) = 0$, so existiert eine Folge $(a_k)_{k\geqslant 1}$ in A mit $\lim\limits_{k\to\infty} \|x - a_k\| = 0$. Also konvergiert $(a_k)_{k\geqslant 1}$ gegen x und $x \in \overline{A}$ folgt aus (2.4). □

Stetigkeit lässt sich auch noch auf andere Weise charakterisieren.

(3.11) Satz. *Seien $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume und $M \subset V$. Für eine Abbildung $f : M \to V'$ sind äquivalent:*
(i) f ist stetig.
(ii) Zu jeder offenen Teilmenge $U' \subset V'$ existiert eine offene Teilmenge $U \subset V$ mit

$$f^{-1}(U') = U \cap M.$$

(iii) Zu jeder abgeschlossenen Teilmenge $A' \subset V'$ existiert eine abgeschlossene Teilmenge $A \subset V$ mit

$$f^{-1}(A') = A \cap M.$$

Beweis. "(i) $\Rightarrow$ (ii)" Sei $U' \subset V'$ offen und $x \in f^{-1}(U')$, also $f(x) \in U'$. Weil U' offen ist, gibt es ein $\varepsilon_x > 0$ mit

$$K_{\varepsilon_x}(f(x)) \subset U'.$$

Da f in x stetig ist, gibt es ein $\delta_x > 0$ mit

$$f(K_{\delta_x}(x) \cap M) \subset K_{\varepsilon_x}(f(x)) \subset U'.$$

Es folgt

$$f^{-1}(U') = U \cap M \quad \text{mit} \quad U = \bigcup_{x \in f^{-1}(U')} K_{\delta_x}(x).$$

U ist nach (1.10) offen.
"(ii) $\Rightarrow$ (i)" Sei $a \in M$ und $\varepsilon > 0$. Dann ist $U' := K_\varepsilon(f(a))$ offen, also gibt es ein offenes $U \subset V$ mit $f^{-1}(U') = U \cap M$. Weil $a \in U$, gibt es ein $\delta > 0$ so, dass $K_\delta(a) \subset U$. Aus $x \in M$ und $\|x - a\| < \delta$ folgt also $\|f(x) - f(a)\|' < \varepsilon$.
"(ii) $\Rightarrow$ (iii)" Sei $A' \subset V'$ abgeschlossen. Dann ist $U' = V' \backslash A'$ nach (1.10) offen. Nach Voraussetzung existiert ein offenes $U \subset V$ mit $f^{-1}(U') = U \cap M$. Es folgt

$$\begin{aligned} f^{-1}(A') &= f^{-1}(V' \backslash U') = M \backslash f^{-1}(U') \\ &= M \backslash (M \cap U) = (V \backslash U) \cap M. \end{aligned}$$

Weil $V \backslash U$ nach (1.10) abgeschlossen ist, folgt die Behauptung.
"(iii) $\Rightarrow$ (ii)" folgt völlig analog. □

Weil der Durchschnitt zweier offener bzw. abgeschlossener Mengen wieder offen bzw. abgeschlossen ist, folgt das

(3.12) Korollar. *Sei $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume, $M \subset V$ und $f : M \to V'$ eine Abbildung.*
a) Ist M offen, so ist f genau dann stetig, wenn $f^{-1}(U')$ für jedes offene $U' \subset V'$ wieder offen in V ist.
b) Ist M abgeschlossen, so ist f genau dann stetig, wenn $f^{-1}(A')$ für jedes abgeschlossene $A' \subset V'$ wieder abgeschlossen in V ist.

§4. Kompakte und zusammenhängende Mengen

In diesem Paragrafen führen wir den üblichen Kompaktheitsbegriff ein und zeigen, dass der für $\mathbb{R}$ in V(3.1) eingeführte Begriff dazu äquivalent ist. Darüber hinaus geben wir eine axiomatische Charakterisierung der Intervalle an.

(4.1) Definition. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $M \subset V$.
a) $\{F_i;\ i \in I\}$ heißt eine *offene Überdeckung von* M, wenn $I \neq \emptyset, F_i$ offen für jedes $i \in I$ und $M \subset \bigcup_{i\in I} F_i$.
b) Man nennt M *kompakt*, wenn jede offene Überdeckung $\{F_i;\ i \in I\}$ von M eine endliche Teilüberdeckung besitzt, d. h., es gibt ein endliches $J \subset I$ mit $M \subset \cup_{j\in J} F_j$.
c) M heißt *folgenkompakt*, wenn jede Folge $(x_k)_{k\geqslant 1}$ mit $x_k \in M, k \geqslant 1$, eine Teilfolge besitzt, die gegen ein $x \in M$ konvergiert.

Die Folgenkompaktheit ist uns aus V(3.4) in $\mathbb{R}$ vertraut. Der Kompaktheitsbegriff ist neu. Einen Zusammenhang beschreibt der

(4.2) Satz. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Für $M \subset V$ sind äquivalent:*

(i) M ist kompakt.

(ii) M ist folgenkompakt.

Beweis. "(i) $\Rightarrow$ (ii)" Sei M kompakt, $(x_k)_{k\geqslant 1}$ eine Folge in M und

$$A := \{x_k;\ k \in \mathbb{N}\}.$$

Wir nehmen zunächst an, dass A keinen Häufungspunkt in M besitzt. Zu jedem $x \in M$ existiert daher ein $\varepsilon_x > 0$ mit $K_{\varepsilon_x}(x) \cap A \subset \{x\}$. Sei

$$F_k := K_{\varepsilon_{x_k}}(x_k),\ k \geqslant 1,\ F_0 := \bigcup_{x\in M\setminus A} K_{\varepsilon_x}(x).$$

Dann ist $\{F_i;\ i \in \mathbb{N}_0\}$ eine offene Überdeckung von M, die nach Voraussetzung eine endliche Teilüberdeckung besitzt. Wegen $F_0 \cap A = \emptyset$ gibt es ein $N \in \mathbb{N}$ mit $A \subset F_1 \cup \ldots \cup F_N$, also

$$A = \{x_1, \ldots, x_N\}.$$

Dann gibt es eine konstante Teilfolge $(x_{k_j})_{j\geqslant 1}$, die somit auch konvergiert.
Im zweiten Fall können wir annehmen, dass A einen Häufungspunkt x in M besitzt. Sei $k_1 \in \mathbb{N}$ mit $x_{k_1} \in K_1(x)$ und $k_j > k_{j-1}$ mit $x_{k_j} \in K_{1/j}(x)$. Aus $\|x_{k_j} - x\| < \frac{1}{j}$ folgt $x_{k_j} \xrightarrow[j\to\infty]{} x$.
Also ist M folgenkompakt.

"(ii) $\Rightarrow$ (i)" Sei M folgenkompakt. Wir betrachten zunächst speziell eine offene Überdeckung durch Kugeln mit festem Radius.

Beh. 1: Zu jedem $r > 0$ gibt es $x_1, \ldots, x_N \in M$ mit $M \subset \cup_{j=1}^N K_r(x_j)$.

Bew.: Ist die Behauptung falsch, dann existiert ein $r > 0$, so dass $\{K_r(x);\, x \in M\}$ keine endliche Teilüberdeckung besitzt. Wir wählen $x_1 \in M$ und nach Voraussetzung existiert ein $x_{k+1} \in M \setminus \left(\bigcup_{j=1}^{k} K_r(x_j)\right)$, $k \in \mathbb{N}$. Wegen der Folgenkompaktheit gibt es eine Teilfolge $(x_{k_j})_{j\geqslant 1}$, die gegen ein $a \in M$ konvergiert. Zu $\varepsilon = \frac{r}{2}$ gibt es somit ein $j_0 \in \mathbb{N}$, so dass für alle $j \geqslant j_0$.

$$\|x_{k_j} - x_{k_{j_0}}\| \leqslant \|x_{k_j} - a\| + \|a - x_{k_{j_0}}\| < \varepsilon + \varepsilon = r.$$

Man erhält $x_{k_j} \in K_r(x_{k_{j_0}})$ für alle $j \geqslant j_0$. Das widerspricht der Konstruktion der Folge $(x_k)_{k\geqslant 1}$. Also muss eine endliche Teilüberdeckung existieren. □

Nun führen wir den allgemeinen Fall auf den Spezialfall zurück.

Beh. 2: Ist $\mathcal{F} = \{F_i;\, i \in I\}$ eine offene Überdeckung von M, dann gibt es ein $r > 0$, so dass zu jedem $x \in M$ ein $F_i \in \mathcal{F}$ mit $K_r(x) \subset F_i$ existiert.

Bew.: Andernfalls existiert zu $r = 2^{-k}$ ein $x_k \in M$ mit $K_{2^{-k}}(x_k) \not\subset F_i$ für jedes $i \in I$. Nach Voraussetzung gibt es eine Teilfolge $x_{k_j} \xrightarrow[j\to\infty]{} a \in M$. Da F eine offene Überdeckung ist, gibt es ein $i \in I$ mit $a \in F_i$. Weil F_i offen ist, gibt es ein $j \in \mathbb{N}$ mit $K_{2^{-j}}(a) \subset F_i$. Nun wählen wir ein $j_0 > j$ mit $x_{k_{j_0}} \in K_{2^{-j-1}}(a)$, also

$$K_{2^{-k_{j_0}}}(x_{k_{j_0}}) \subset K_{2^{-j-1}}(x_{k_{j_0}}) \subset K_{2^{-j}}(a) \subset F_i$$

im Widerspruch zur Konstruktion der Folge. □

Sei nun $\mathcal{F} = \{F_i;\, i \in I\}$ eine beliebige offene Überdeckung von M. Wir wählen nach Beh. 2 ein $r > 0$, so dass zu jedem $x \in M$ ein $F_{i_x} \in \mathcal{F}$ mit $K_r(x) \subset F_{i_x}$ existiert. Nach Beh. 1 gibt es $x_1, \ldots, x_N \in M$ mit $M \subset \bigcup_{j=1}^{N} K_r(x_j) \subset \bigcup_{j=1}^{N} F_{i_{x_j}}$. Also haben wir eine endliche Teilüberdeckung gefunden. □

Im nächsten Schritt untersuchen wir das Verhalten unter stetigen Abbildungen.

> **(4.3) Satz.** *Seien $(V, \|\cdot\|), (V', \|\cdot\|')$ normierte Vektorräume, $K \subset V$ kompakt und die Abbildung $f : K \to V'$ stetig. Dann gilt:*
> *a) $f(K)$ ist kompakt.*
> *b) f ist gleichmäßig stetig.*

Beweis. a) Sei $(y_k)_{k\geqslant 1}$ eine Folge in $f(K)$. Dann existiert zu jedem $k \in \mathbb{N}$ ein $x_k \in K$ mit $y_k = f(x_k)$. Da K kompakt ist, existiert nach (4.2) eine Teilfolge $x_{k_j} \xrightarrow[j\to\infty]{} a \in K$. Weil f stetig ist, folgt $y_{k_j} = f(x_{k_j}) \xrightarrow[j\to\infty]{} f(a) \in f(K)$. Also ist $f(K)$ nach (4.2) ebenfalls kompakt.
b) Sei $\varepsilon > 0$. Wegen der Stetigkeit existiert zu jedem $a \in K$ ein $r_a > 0$, so dass

$$(*) \qquad x \in K,\ \|x - a\| < 2r_a \ \Rightarrow\ \|f(x) - f(a)\|' < \frac{\varepsilon}{2}.$$

Also ist $\{K_{r_a}(a); a \in K\}$ eine offene Überdeckung von K. Nach (4.1) gibt es $a_1, \ldots, a_N \in K$ mit der Eigenschaft

$$K \subset \bigcup_{j=1}^{N} K_{r_{a_j}}(a_j).$$

Sei $\delta := \min\{r_{a_1}, \dots, r_{a_N}\}$. Zu $x, y \in K$ mit $\|x - y\| < \delta$ gibt es zunächst ein $1 \leqslant j \leqslant N$ mit $x \in K_{r_{a_j}}(a_j)$. Dann folgt

$$\|y - a_j\| \leqslant \|y - x\| + \|x - a_j\| < \delta + r_{a_j} \leqslant 2r_{a_j}.$$

Aus $(*)$ erhält man dann

$$\|f(x) - f(y)\|' \leqslant \|f(x) - f(a_j)\|' + \|f(a_j) - f(y)\|' < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist f gleichmäßig stetig auf K. □

Als Anwendung erhalten wir das

(4.4) Korollar. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Ist $A \subset V$ abgeschlossen und $K \subset V$ kompakt mit $K \cap A = \emptyset$, so ist der Abstand der Mengen K und A*

$$d(K, A) = \inf\{\|x - y\|;\ x \in K,\ y \in A\}$$

positiv.

Beweis. Es gilt

$$\inf\{\|x - y\|;\ x \in K,\ y \in A\} = \inf\{d(x, A);\ x \in K\}.$$

Wegen $K \cap A = \emptyset$ gilt $d(x, A) > 0$ für alle $x \in K$ nach (3.10). Darüber hinaus ist

$$d(\cdot, A)(K) = \{d(x, A);\ x \in K\} \subset \mathbb{R}_+^*$$

nach (4.3) kompakt und besitzt nach V(3.6) ein Minimum, das wiederum positiv ist. □

Als nächstes untersuchen wir Abgeschlossenheit und Beschränktheit.

(4.5) Lemma. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und $K \subset V$ kompakt. Dann ist K abgeschlossen und beschränkt.*

Beweis. $\{K_n(0);\ n \in \mathbb{N}\}$ ist eine offene Überdeckung von K, die eine endliche Teilüberdeckung besitzt. Also existiert ein $N \in \mathbb{N}$ mit $K \subset K_N(0)$. Daher ist K beschränkt.
Ist x ein Häufungspunkt von K, so existiert nach (2.4) eine Folge $(x_k)_{k \geqslant 1}$ in K mit $x_k \xrightarrow[k \to \infty]{} x$. Weil auch jede Teilfolge gegen x konvergiert, folgt $x \in K$ aus der Folgenkompaktheit. Also ist K abgeschlossen. □

Die Umkehrung ist nur unter einer Zusatzvoraussetzung richtig, z. B., wenn V endlich dimensional ist.

(4.6) Satz von HEINE-BOREL. *Für $K \subset \mathbb{R}^n$ sind äquivalent:*

(i) K ist kompakt.

(ii) K ist folgenkompakt.

(iii) K ist abgeschlossen und beschränkt.

Beweis. "(i) $\Leftrightarrow$ (ii), (i) $\Rightarrow$ (iii)" Die Aussagen wurden in (4.2) und (4.5) bewiesen.
"(iii) $\Rightarrow$ (ii)" Sei $(x_k)_{k\geqslant 1}$ eine Folge in K. Dann ist $(x_k)_{k\geqslant 1}$ beschränkt. Nach dem Satz von BOLZANO-WEIERSTRASS (2.8) existiert eine konvergente Teilfolge $x_{k_j} \xrightarrow[j\to\infty]{} a \in \mathbb{R}^n$. Weil K abgeschlossen ist, folgt $a \in K$ aus (2.4). □

Nun diskutieren wir einige

(4.7) Beispiele. a) Ein normierter Vektorraum $(V, \|\cdot\|), V \neq \{0\}$, ist nicht kompakt, da er nicht beschränkt ist. Für $r > 0, a \in V$ ist $K_r(a)$ nicht kompakt, da nicht abgeschlossen.
b) Für $V = \mathbb{R}^n$, $a \in \mathbb{R}$ ist $\overline{K_r(a)} = \{x \in \mathbb{R}^n;\ \|x-a\| \leqslant r\}$ kompakt nach (4.6). Insbesondere sind alle abgeschlossenen Kreise, Kugeln und Quader kompakt.
c) Sei $(V, \|\cdot\|) = (\mathcal{C}([0,1]), \|\cdot\|_\infty)$ und $K = \{f \in \mathcal{C}([0,1]);\ \|f\|_\infty \leqslant 1\}$. Dann ist K abgeschlossen und beschränkt. Die Folge $(f_k)_{k\geqslant 1}$, $f_k(x) = x^k$, $\|f_k\|_\infty = 1$, besitzt keine konvergente Teilfolge in $(\mathcal{C}([0,1]), \|\cdot\|_\infty)$. Also ist K nicht kompakt.

Eine weitere Anwendung des Kompaktheitsbegriffes formulieren wir in dem

(4.8) Satz. *Seien $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume,* dim $V < \infty, U \subset V$ *offen und $f_k : U \to V'$, $k \in \mathbb{N}$, Funktionen. Dann sind äquivalent:*
(i) Die Funktionenfolge $(f_k)_{k\geqslant 1}$ ist lokal gleichmäßig konvergent.
(ii) Die Funktionenfolge $(f_k)_{k\geqslant 1}$ konvergiert auf jedem Kompaktum in U gleichmäßig.

Beweis. "(ii) $\Rightarrow$ (i)" Sei $a \in U$. Da U offen ist, gibt es ein $r > 0$ mit $K_{2r}(a) \subset U$. Dann ist $\overline{K_r(a)}$ eine kompakte Umgebung von a, auf der die Funktionenfolge $(f_k)_{k\geqslant 1}$ nach Voraussetzung gleichmäßig konvergiert.
"(i) $\Rightarrow$ (ii)" Nach Voraussetzung existiert eine Funktion $f : U \to V'$ und zu jedem $a \in U$ ein $r_a > 0$ mit $K_{r_a}(a) \subset U$ sowie zu $\varepsilon > 0$ ein $N_a \in \mathbb{N}$ mit der Eigenschaft

$$(*) \qquad \|f_k(x) - f(x)\|' < \varepsilon \quad \text{für alle} \quad x \in K_{r_a}(a), \quad k \geqslant N_a.$$

Sei $K \subset U$ kompakt. Dann ist $\{K_{r_a}(a);\ a \in U\}$ eine offene Überdeckung von K. Nach (4.1) gibt es $a_1, \ldots, a_s \in U$ mit

$$K \subset \bigcup_{j=1}^{s} K_{r_{aj}}(a_j).$$

Sei $N := \max\{N_{a_1}, \ldots, N_{a_s}\}$. Zu $x \in K$ gibt es ein $1 \leqslant j \leqslant s$ mit $x \in K_{r_{a_j}}(a_j)$. Wegen $N \geqslant N_{a_j}$ liefert $(*)$

$$\|f_k(x) - f(x)\|' < \varepsilon \quad \text{für alle } k \geqslant N.$$

Also konvergiert $\left(f_k\big|_K\right)_{k\geqslant 1}$ gleichmäßig gegen $f\big|_K$. □

In Verbindung mit der Stetigkeit sind auch Zusammenhangsbegriffe von Interesse.

(4.9) Definition. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Eine Menge $M \subset V$ heißt *wegzusammenhängend* oder *kurvenzusammenhängend*, wenn es zu allen Punkten $a, b \in M$ eine stetige Abbildung $\varphi : [0,1] \to V$ gibt mit $\varphi([0,1]) \subset M$, $\varphi(0) = a$ und $\varphi(1) = b$. Jedes solche φ nennt man eine *Kurve* in V.

Wir führen zudem noch einen weiteren Begriff ein.

(4.10) Definition. Es sei V ein $\mathbb{R}$-Vektorraum. Eine Teilmenge M von V heißt *konvex*, wenn für alle $x, y \in M$ und alle reellen $\lambda \in [0,1]$ auch $(1-\lambda)x + \lambda y \in M$. M enthält also mit x, y auch die *Verbindungsstrecke*

$$\{(1-\lambda)x + \lambda y;\ 0 \leqslant \lambda \leqslant 1\}.$$

(4.11) Beispiele. Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum.

a) Jede konvexe Teilmenge M von V ist wegzusammenhängend. Denn zu $a, b \in M$ definiere $\varphi : [0,1] \to V$,

$$\varphi(t) = (1-t)a + tb.$$

Dann ist φ als Polynomabbildung stetig, es gilt $\varphi(0) = a$, $\varphi(1) = b$ und für alle $t \in [0,1]$ ist $\varphi(t) \in M$ nach Definition.

b) V und $\emptyset$ sind konvex, also wegzusammenhängend.

c) Für jedes $a \in V$ und $r > 0$ ist $K_r(a)$ konvex, also wegzusammenhängend.

Zum Nachweis der Konvexität seien $x, y \in K_r(a)$ und $\lambda \in [0,1]$. Dann ist

$$\begin{aligned} &\|(1-\lambda)x + \lambda y - a\| = \|(1-\lambda)(x-a) + \lambda(y-a)\| \\ \leqslant\ &(1-\lambda)\cdot\|x-a\| + \lambda\cdot\|y-a\| < (1-\lambda)r + \lambda r = r \end{aligned}$$

mit der Dreiecksungleichung.

Diese Fakten verwenden wir auch bei der Charakterisierung des Begriffes in $\mathbb{R}$.

(4.12) Satz. *Die wegzusammenhängenden Teilmengen von $\mathbb{R}$ sind genau die Intervalle und die leere Menge.*

Beweis. Intervalle sind konvex, also wegzusammenhängend.
Sei nun $M \subset \mathbb{R}$ kein Intervall. Dann gibt es $a, b, c \in \mathbb{R}$ mit $a < b < c$, so dass $a, c \in M$ und $b \notin M$. Zu jeder stetigen Funktion $\varphi : [0,1] \to \mathbb{R}$ mit $\varphi(0) = a$ und $\varphi(1) = c$ existiert nach dem Zwischenwertsatz V(3.8) ein $\xi \in (0,1)$ mit $\varphi(\xi) = b$, also $\varphi([0,1]) \not\subset M$. Demnach ist M nicht wegzusammenhängend. □

Schließlich untersuchen wir noch im allgemeinen Rahmen das Verhalten unter stetigen Abbildungen.

(4.13) Satz. *Seien $(V, \|\cdot\|)$, $(V', \|\cdot\|')$ normierte Vektorräume, $M \subset V$ wegzusammenhängend. Ist $f : M \to V'$ stetig, so ist $f(M)$ ebenfalls wegzusammenhängend.*

Beweis. Seien $x, y \in f(M)$. Dann gibt es $a, b \in M$ mit $f(a) = x, f(b) = y$. Da M wegzusammenhängend ist, gibt es ein stetiges $\varphi : [0,1] \to V$ mit $\varphi([0,1]) \subset M$, $\varphi(0) = a$ und $\varphi(1) = b$. Nach (3.3) ist dann auch $f \circ \varphi : [0,1] \to V'$ stetig mit

$$(f \circ \varphi)([0,1]) \subset f(M), \quad (f \circ \varphi)(0) = x \quad \text{und} \quad (f \circ \varphi)(1) = y.$$

Also ist $f(M)$ wegzusammenhängend. □

Wir führen noch einen etwas allgemeineren Begriff ein.

(4.14) Definition. Sei $(V; \|\cdot\|)$ ein normierter Vektorraum. $M \subset V$ heißt *zusammenhängend*, wenn für jede Wahl von offenen Mengen

$$U, U' \subset V \quad \text{mit} \quad U \cap U' \cap M = \emptyset \quad \text{und} \quad M \subset U \cup U'$$

bereits folgt

$$M \subset U \quad \text{oder} \quad M \subset U'.$$

Wir geben eine äquivalente Charakterisierung mit abgeschlossenen Mengen.

(4.15) Satz. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. $M \subset V$ ist genau dann zusammenhängend, wenn für jede Wahl von abgeschlossenen Mengen*

$$A, A' \subset V \quad \textit{mit} \quad A \cap A' \cap M = \emptyset \quad \textit{und} \quad M \subset A \cup A'$$

bereits folgt

$$M \subset A \quad \textit{oder} \quad M \subset A'.$$

Beweis. "$\Rightarrow$" Seien $A, A' \subset V$ abgeschlossen mit $A \cap A' \cap M = \emptyset$ und $M \subset A \cup A'$. Dann sind $U := V \backslash A$ und $U' := V \backslash A'$ offen mit

$$U \cap U' \cap M = \big(V \backslash (A \cup A')\big) \cap M = \emptyset, \quad U \cup U' = \big(V \backslash (A \cap A')\big) \supset M.$$

Es folgt $M \subset U = V \backslash A$, also $M \subset A'$ wegen $M \subset A \cup A'$, oder $M \subset U' = V \backslash A'$, also $M \subset A$.
"$\Leftarrow$" Seien $U, U' \subset V$ offen mit $U \cap U' \cap M = \emptyset$ und $M \subset U \cup U'$. Man definiert dann $A := V \backslash U$, $A' := V \backslash U'$ und geht analog vor. □

Wir kommen nun zu wichtigen Beispielen.

(4.16) Satz. *Die zusammenhängenden Teilmengen von $\mathbb{R}$ sind genau die Intervalle und die leere Menge.*

Beweis. Sei $I = [a, b]$ ein Intervall. Seien A, A' abgeschlossen mit $I \cap A \cap A' = \emptyset, I \subset A \cup A'$. Sei ohne Einschränkung $a \in A$. Wir betrachten

$$c := \sup\{t \in I;\ [a, t] \subset A\}.$$

Dann gilt $a \leqslant c \leqslant b$ und $c \in A$, da A abgeschlossen ist. Wäre $c < b$, so würde es zu jedem $\varepsilon > 0$ ein $x \in I$ mit

$$c < x < c + \varepsilon, \quad x \notin A, \quad \text{also} \quad x \in A',$$

geben. Dann wäre c ein Häufungspunkt von A'. Weil A' abgeschlossen ist, folgt $c \in A'$. Das ist ein Widerspruch. Also gilt $c = b$ und $I \subset A$.
Die anderen Intervalle behandelt man analog.

Sei nun $M \subset \mathbb{R}$ kein Intervall. Dann existieren $a, b, c \in \mathbb{R}$ mit

$$a < b < c, \quad a, c \in M, \quad b \notin M.$$

In diesem Fall sind $U := (-\infty, b),\ U' := (b, \infty)$ offen mit

$$U \cap U' = \emptyset, \quad M \subset U \cup U', \quad M \not\subset U,\ M \not\subset U'.$$

Also ist M nicht zusammenhängend. $\square$

Nun beschreiben wir das Verhalten unter stetigen Abbildungen.

(4.17) Satz. *Seien $(V, \|\cdot\|), (V', \|\cdot\|')$ normierte Vektorräume, $M \subset V$ zusammenhängend und $f : M \to V'$ stetig. Dann ist auch $f(M)$ zusammenhängend.*

Beweis. Seien $A_1', A_2' \subset V'$ abgeschlossen mit $f(M) \subset A_1' \cup A_2'$ und $A_1' \cap A_2' \cap f(M) = \emptyset$. Nach (3.11) gibt es abgeschlossene Mengen $A_1, A_2 \subset V$ mit

$$f^{-1}(A_1') = A_1 \cap M, \quad f^{-1}(A_2') = A_2 \cap M.$$

Aus $A_1' \cap A_2' \cap f(M) = \emptyset$ erhält man $A_1 \cap A_2 \cap M = \emptyset$. Darüber hinaus ergibt die Inklusion $f(M) \subset A_1' \cup A_2'$ auch $M \subset A_1 \cup A_2$. Da M zusammenhängend ist, folgt $M \subset A_1$ oder $M \subset A_2$, also $f(M) = f(M \cap A_1) \subset A_1'$ oder $f(M) = f(M \cap A_2) \subset A_2'$. Demnach ist auch $f(M)$ zusammenhängend. $\square$

Die erste Verbindung zum Wegzusammenhang erhalten wir in dem

(4.18) Lemma. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Ist $M \subset V$ wegzusammenhängend, so ist M auch zusammenhängend.*

Beweis. Angenommen, es gibt abgeschlossene Teilmengen A, A' von V mit $A \cap A' \cap M = \emptyset$, $M \subset A \cup A'$, so dass $M \not\subset A$ und $M \not\subset A'$. Dann gibt es ein $a \in A \cap M$ und ein $a' \in A' \cap M$. Weil M wegzusammenhängend ist, existiert eine stetige Abbildung

$$\varphi : [0, 1] \to V \quad \text{mit} \quad \varphi([0, 1]) \subset M,\ \varphi(0) = a,\ \varphi(1) = a'.$$

Nach (3.12) sind $\varphi^{-1}(A)$ und $\varphi^{-1}(A')$ disjunkte abgeschlossene Teilmengen von $\mathbb{R}$ mit

$$\varphi^{-1}(A) \cup \varphi^{-1}(A') = [0, 1], \quad 0 \in \varphi^{-1}(A), \quad 1 \in \varphi^{-1}(A').$$

Das ist ein Widerspruch zu (4.16). $\square$

Es gibt zusammenhängende Teilmengen von $\mathbb{R}^n, n \geqslant 2$, die nicht wegzusammenhängend sind, z. B.

$$\left\{ \begin{pmatrix} x \\ \sin(1/x) \end{pmatrix}; x > 0 \right\} \cup \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix};\ |y| \leqslant 1 \right\} \subset \mathbb{R}^2.$$

Solche Beispiele werden in der Topologie behandelt.

Wir führen eine spezielle Klasse von Kurven ein.

(4.19) Definition. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Eine Abbildung $\varphi : [0,1] \to V$ heißt *Polygon*, wenn es eine (endliche) Zerlegung $0 = t_0 < t_1 < \ldots < t_n = 1$ von $[0,1]$ und Punkte $a_i, b_i \in V$ gibt, so dass gilt

$$\varphi(t) = a_i + tb_i \quad \text{für} \quad t_{i-1} \leqslant t \leqslant t_i \quad \text{und} \quad i = 1, \ldots, n.$$

Ist $M \subset V$ und $\varphi([0,1]) \subset M$, so spricht man von einem *Polygon in* M mit *Anfangspunkt* $\varphi(0)$ und *Endpunkt* $\varphi(1)$.

Man beachte, dass $a_i + t_i b_i = a_{i+1} + t_i b_{i+1}$ für $1 \leqslant i < n$ gilt. Also ist φ stetig und somit eine Kurve. Als Umkehrung von (4.18) notieren wir den

(4.20) Satz. *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Für ein offenes $M \subset V$ sind dann äquivalent:*

(i) M ist zusammenhängend.

(ii) M ist wegzusammenhängend.

(iii) Zu allen $a, b \in M$ gibt es ein Polygon in M mit Anfangspunkt a und Endpunkt b.

Beweis. "(iii) $\Rightarrow$ (ii)" Jedes Polygon ist eine Kurve.
"(ii) $\Rightarrow$ (i)" Man verwende (4.18).
"(i) $\Rightarrow$ (iii)" Sei $a \in M$ und U die Menge aller Punkte $b \in M$, zu denen ein Polygon in M mit Anfangspunkt a und Endpunkt b existiert. Sei $b \in U$. Weil M offen ist, gibt es ein $r > 0$ mit $K_r(b) \subset M$. Sei φ ein Polygon in M mit Anfangspunkt a und Endpunkt b. Dann ist für $x \in K_r(b)$

$$(*) \qquad \psi : [0,1] \to V, \quad \psi(t) = \begin{cases} \varphi(2t), & \text{falls} \quad 0 \leqslant t \leqslant \frac{1}{2}, \\ b + (2t-1)(x-b), & \text{falls} \quad \frac{1}{2} \leqslant t \leqslant 1, \end{cases}$$

ein Polygon in M mit Anfangspunkt a und Endpunkt x. Es folgt $x \in U$, also $K_r(b) \subset U$, so dass U offen ist.
Sei nun $x \in \overline{U} \cap M$. Da M offen ist, gibt es ein $r > 0$ mit $K_r(x) \subset M$. Wegen $x \in \overline{U}$ existiert ein $b \in K_r(x) \cap U$. Dann liefert wiederum das Polygon $(*)$ sofort $x \in U$. Es folgt $\overline{U} \cap M \subset U$.
Sei $U' := V \backslash \overline{U}$. Dann ist U' offen mit

$$U \cup U' \supset M, \quad U \cap U' = \emptyset.$$

Wegen $a \in U$ folgt $M \subset U$, da M zusammenhängend ist. Es folgt $U = M$ und damit (iii). $\square$

Diese Aussage nutzen wir für die

(4.21) Definition. Eine nicht-leere (weg-) zusammenhängende offene Teilmenge eines normierten Vektorraums $(V, \|\cdot\|)$ nennt man ein *Gebiet*.

Kapitel XI.

Differentialrechnung in $\mathbb{R}^n$

In diesem Kapitel werden die Ergebnisse der klassischen Differentialrechnung aus Kapitel VII auf Funktionen in mehreren Variablen übertragen. Wir werden dabei zwei Differenzierbarkeitsbegriffe einführen und auf die Zusammenhänge eingehen.

§1. Richtungsableitungen und partielle Differenzierbarkeit

Zunächst betrachten wir einen Begriff, der – nach Festlegen einer "Richtung" – Differenzierbarkeit auf den bekannten Fall einer Variablen reduziert.

(1.1) Definition. Es seien $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume, $D \subset V$ und $a \in D$ sowie $f : D \to V'$.
Gibt es zu $v \in V \setminus \{0\}$ ein $\rho > 0$ derart, dass $\{a + tv;\ |t| < \rho\} \subset D$ und

$$D_v f(a) := \lim_{t \to 0} \frac{1}{t}\big(f(a + tv) - f(a)\big)$$

existiert, so heißt dieser Grenzwert die *Richtungsableitung* von f im Punkt a in Richtung v.

(1.2) Beispiele. a) Betrachte $f : \mathbb{R}^2 \to \mathbb{R}$, $x \mapsto x_1\, e^{x_2} - \sin(x_1^2 x_2^3)$.

- Für $a = \binom{0}{0}$, $v = \binom{1}{2}$ ist $a + tv = \binom{t}{2t}$ und

$$\frac{1}{t}\big(f(a + tv) - f(a)\big) = \frac{1}{t}\big(t\, e^{2t} - \sin(8t^5)\big) \quad \to 1 \ \text{ für } \ t \to 0.$$

  Also $D_v f(\binom{0}{0}) = 1$.

- Für beliebiges $a \in \mathbb{R}^2$ und $v = \binom{1}{0}$ gilt

$$\gamma(t) := f(a + tv) - f(a) = (a_1 + t)e^{a_2} - \sin\big((a_1 + t)^2 a_2^3\big).$$

  Offensichtlich ist γ eine differenzierbare Funktion von t mit

$$\gamma'(t) = e^{a_2} - \cos\big((a_1 + t)^2 a_2^3\big) \cdot 2(a_1 + t),$$

  und insbesondere

$$\lim_{t \to 0} \frac{1}{t}\big(f(a + tv) - f(a)\big) = \gamma'(0) = e^{a_2} - 2a_1 \cos(a_1^2 a_2^3).$$

- Für beliebiges $a \in \mathbb{R}^2$ und $v = \binom{0}{1}$ gilt
$$\vartheta(t) := f(a+tv) - f(a) = a_1\, e^{a_2+t} - \sin\big(a_1^2(a_2+t)^3\big)$$
mit $\vartheta'(t) = a_1\, e^{a_2+t} - \cos\big(a_1^2(a_2+t)^3\big)\cdot 3(a_2+t)^2$; damit ergibt sich
$$\lim_{t\to 0}\frac{1}{t}\big(f(a+tv)-f(a)\big) = \vartheta'(0) = a_1 e^{a_2} - 3a_2^2\cos(a_1^2 a_2^3).$$

b) Wir diskutieren die Richtungsableitungen folgender Funktion in $a = 0$.
$$g : \mathbb{R}^2 \to \mathbb{R}^2, \quad g(x) = \begin{cases} x, & \text{falls } x_2 = 0 \text{ oder } x_1/x_2 \in \mathbb{Q}, \\ 0 & \text{sonst.}\end{cases}$$
Sei $v \in \mathbb{R}^2\backslash\{0\}$, $v = \binom{v_1}{v_2}$. Ist $v_2 = 0$ oder $v_1/v_2 \in \mathbb{Q}$, so gilt
$$g(0+tv) - g(0) = tv - 0, \quad \text{also}$$
$D_v g(0) = v$. Andernfalls ist
$$g(0+tv) - g(0) = 0 - 0 = 0, \quad \text{also}$$
$D_v g(0) = 0$.

c) Es soll auch ein Beispiel in einem unendlich-dimensionalen Vektorraum betrachtet werden. Sei dazu $(V, \|\cdot\|) = \big(\mathcal{C}([0,1]), \|\cdot\|_\infty\big)$ und $F : V \to \mathbb{R}$ gegeben durch $F(\varphi) = \varphi(0)^2$. Dann existiert $D_\psi F(\varphi)$ für jedes $0 \neq \psi \in \mathcal{C}([0,1])$, denn
$$\frac{1}{t}\left(\big((\varphi+t\psi)(0)\big)^2 - \varphi(0)^2\right) = \frac{1}{t}\big(2t\varphi(0)\psi(0) + t^2\psi(0)^2\big)$$
und der Grenzwert für $t \to 0$ ergibt
$$D_\psi F(\varphi) = 2\varphi(0)\psi(0).$$
Im Folgenden werden wir uns aber auf endlich-dimensionale Räume beschränken.

(1.3) Bemerkungen. a) Ist $V = \mathbb{R}$, so nennt man in der Situation von (1.1) die Richtungsableitung $D_{(1)}F(a)$ in Richtung $v = (1)$ auch kurz die *Ableitung* von F in a (wenn sie existiert) und schreibt kurz $F'(a)$.

b) In praktischen Rechnungen betrachtet man – wieder in der Situation von (1.1) – oft die Funktion $\gamma_v : (-\rho, \rho) \to V'$,
$$t \mapsto f(a+tv) - f(a).$$
Dann ist die Existenz von $D_v f(a)$ äquivalent zur Existenz von $\gamma'(0)$ (i. S. von Teil (a)).

c) Falls $\lambda \neq 0$, gilt
$$D_{\lambda v} f(a) = \lambda \cdot D_v f(a),$$
sofern einer dieser Grenzwerte existiert. Denn
$$\frac{1}{t}\Big(F\big(a+t(\lambda v)\big) - F(a)\Big) = \lambda\cdot\left[\frac{1}{(\lambda t)}\Big(F\big(a+(\lambda t)v\big) - F(a)\Big)\right].$$
Der Name “Richtungsableitung in Richtung v” ist also nicht ganz zutreffend.

Im Folgenden betrachten wir vor allem den $\mathbb{R}^n$.

Es sei stets

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \; e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

die kanonische Basis des $\mathbb{R}^n$. Man nennt nun die Richtungsableitung in Richtung e_k auch die *partielle Ableitung* bezüglich der Variablen x_k. Genauer:

(1.4) Definition. Sei $U \subset \mathbb{R}^n$ offen, $a = (a_1, \dots, a_n)^t \in U$ und $F : U \to \mathbb{R}^m$ eine Funktion.
a) F heißt in a *partiell differenzierbar bezüglich der k-ten Koordinate*, wenn $1 \leqslant k \leqslant n$ und

$$\begin{aligned} D_kF(a) := \frac{\partial F}{\partial x_k}(a) &:= D_{e_k}F(a) = \lim_{x \to a_k} \frac{1}{x - a_k} \left[F((a_1, \dots, a_{k-1}, x, a_{k+1}, \dots, a_n)^t) - F(a) \right] \\ &= \lim_{t \to 0} \frac{1}{t}(F(a + te_k) - F(a)) \in \mathbb{R}^m \end{aligned}$$

existiert.
b) $F : U \to \mathbb{R}^m$ heißt *partiell differenzierbar*, wenn $D_kF(a)$ für alle $a \in U$ und alle $k = 1, \dots, n$ existiert. Die $D_kF : U \to \mathbb{R}^m$ nennt man die *partiellen Ableitungen* (oder die *partiellen Ableitungsfunktionen*) von F. Existieren alle partiellen Ableitungen auf U, so setzt man

$$DF(x) := (D_1F(x), \dots, D_nF(x)) \in \mathbb{R}^{m \times n}$$

für $x \in U$. Man nennt $DF(x)$ auch *Funktionalmatrix*, oder *Jacobi-Matrix*, oder *Differential* von F an der Stelle x.
c) $F : U \to \mathbb{R}^m$ heißt *stetig partiell differenzierbar*, wenn F partiell differenzierbar ist und alle partiellen Ableitungen $D_kF : U \to \mathbb{R}^m$ für $k = 1, \dots, n$ stetig sind.

Im Fall $n = m = 1$ stimmt die partielle Differenzierbarkeit also mit dem Differenzierbarkeitsbegriff (vgl. VII(1.3)) überein.

Man erinnere sich, dass $F : U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$, stetig ist, wenn zu jedem $a \in U$ und $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$x \in U, \|x - a\| < \delta \;\Rightarrow\; \|F(x) - F(a)\|' < \varepsilon,$$

wobei $\|\cdot\|$ und $\|\cdot\|'$ beliebige vorgegebene Normen auf dem $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ sind.

(1.5) Lemma. *Sei $U \subset \mathbb{R}^n$ offen, $a = (a_1, \dots, a_n)^t \in U$, $F : U \to \mathbb{R}^m$ eine Funktion mit $F = (F_1, \dots, F_m)^t$ und den Komponentenfunktionen $F_j : U \to \mathbb{R}$, $j = 1, \dots, m$, sowie $g : U \to \mathbb{R}$ eine Funktion und $1 \leqslant k \leqslant n$.*
a) Die Funktion F ist genau dann in a bezüglich der k-ten Koordinate partiell differenzierbar, wenn

jedes $F_j : U \to \mathbb{R}$, $j = 1, \dots, m$, in a bezüglich der k-ten Koordinate partiell differenzierbar ist. Dann gilt

$$D_k F(a) = (D_k F_1(a), \dots, D_k F_m(a))^t.$$

b) Sei $\delta > 0$ und $\{a + te_k; |t| < \delta\} \subset U$. In diesem Fall ist g genau dann in a bezüglich der k-ten Koordinate partiell differenzierbar, wenn die Funktion

$$\varphi_k : (a_k - \delta, a_k + \delta) \to \mathbb{R}, \quad t \mapsto g((a_1, \dots, a_{k-1}, t, a_{k+1}, \dots, a_n)^t),$$

in a_k differenzierbar ist. Dann gilt

$$D_k g(a) = \varphi_k'(a_k).$$

Beweis. a) Nach X(2.5) a) und X(2.10) b) gilt

$$D_k F(a) = \lim_{t \to 0} \frac{1}{t}(F(a + te_k) - F(a))$$

$$= \lim_{t \to 0} \frac{1}{t} \begin{pmatrix} F_1(a + te_k) - F_1(a) \\ \vdots \\ F_m(a + te_k) - F_m(a) \end{pmatrix} = \begin{pmatrix} \lim\limits_{t \to 0} \frac{1}{t}(F_1(a + te_k) - F_1(a)) \\ \vdots \\ \lim\limits_{t \to 0} \frac{1}{t}(F_m(a + te_k) - F_m(a)) \end{pmatrix} = \begin{pmatrix} D_k F_1(a) \\ \vdots \\ D_k F_m(a) \end{pmatrix}.$$

b) folgt direkt mit (1.3) b). □

Nun diskutieren wir einige

(1.6) Beispiele. a) Wir betrachten

$$F : \mathbb{R}^3 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} (\sin x)e^y \\ y^2 z + z^2 x \end{pmatrix}.$$

Dann gilt

$$D_1 F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (\cos x)e^y \\ z^2 \end{pmatrix}, \quad D_2 F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (\sin x)e^y \\ 2yz \end{pmatrix}, \quad D_3 F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ y^2 + 2zx \end{pmatrix}.$$

b) Nun untersuchen wir

$$f : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \|x\|_2 = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Für $1 \leqslant k \leqslant n$ und $x \neq 0$ gilt

$$D_k f(x) = \frac{1}{2}(x_1^2 + \ldots + x_n^2)^{-1/2} \cdot 2x_k = \frac{x_k}{\|x\|_2},$$

also auch

$$Df(x) = \frac{1}{\|x\|_2} x^t \quad \text{für} \;\; x \neq 0.$$

Aber f ist in $x = 0$ nicht partiell differenzierbar, denn es gilt für $t \neq 0$

$$\frac{f(te_k) - f(0)}{t} = \frac{\sqrt{t^2}}{t} = \frac{|t|}{t}.$$

c) Die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{cases}$$

ist nicht stetig in $\binom{0}{0}$ (vgl. X(3.4) f), da

$$\lim_{k\to\infty} f\begin{pmatrix} 1/k \\ 1/k \end{pmatrix} = \frac{1}{2} \neq f\begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

f ist aber in $\binom{0}{0}$ partiell differenzierbar mit

$$D_1 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lim_{x\to 0} \frac{1}{x}\left[f\begin{pmatrix} x \\ 0 \end{pmatrix} - f\begin{pmatrix} 0 \\ 0 \end{pmatrix} \right] = 0 = D_2 f\begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

d) Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ nach der k-ten Koordinate partiell differenzierbar. Ist $g : \mathbb{R} \to \mathbb{R}$ differenzierbar, so ist $g \circ f : U \to \mathbb{R}$ nach der k-ten Koordinate partiell differenzierbar mit

$$D_k(g \circ f)(a) = g'(f(a)) \cdot D_k f(a),$$

denn mit (1.5) folgt

$$D_k(g \circ f)(a) = \frac{d}{dt}(g \circ \varphi)(t)\Big|_{t=a_k} = g'(\varphi(t)) \cdot \varphi'(t)\Big|_{t=a_k} = g'(f(a)) \cdot D_k f(a).$$

Nützlich, vor allem in der Physik, sind die folgenden Begriffe.

(1.7) Definition. Sei $U \subset \mathbb{R}^n$ offen.
a) Ist $f : U \to \mathbb{R}$ partiell differenzierbar, so nennt man

$$\operatorname{grad} f : U \to \mathbb{R}^n, \quad \operatorname{grad} f = \begin{pmatrix} D_1 f \\ \vdots \\ D_n f \end{pmatrix} = (Df)^t$$

den *Gradienten von f*.
b) Ist $F : U \to \mathbb{R}^n$ partiell differenzierbar, so heißt

$$\operatorname{div} F : U \to \mathbb{R}, \quad \operatorname{div} F = \sum_{k=1}^{n} D_k F_k, \quad F = (F_1, \ldots, F_n)^t,$$

die *Divergenz von F*.

Wir betrachten einige

(1.8) Beispiele. a) Die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \sqrt{1+x^2}\sin(xy),$$

erfüllt

$$\operatorname{grad} f : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \frac{x}{\sqrt{1+x^2}}\sin(xy) + y\sqrt{1+x^2}\cos(xy) \\ x\sqrt{1+x^2}\cos(xy) \end{pmatrix}.$$

b) Sei $\varphi : \mathbb{R}_+^* \to \mathbb{R}$ differenzierbar und $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$, $x \mapsto \varphi(\|x\|_2)$. Dann gilt nach (1.6)

$$\operatorname{grad} f(x) = \varphi'(\|x\|_2) \cdot (x_1/\|x\|_2, \dots, x_n/\|x\|_2)^t = \frac{\varphi'(\|x\|_2)}{\|x\|_2} \cdot x \in \mathbb{R}^n \quad \text{für} \quad 0 \neq x \in \mathbb{R}^n.$$

c) Sei $\varphi : \mathbb{R}_+^* \to \mathbb{R}$ differenzierbar und $F : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}^n$, $x \mapsto \varphi(\|x\|_2)x$. Dann gilt für $k = 1, \dots, n$

$$\frac{\partial}{\partial x_k}(\varphi(\|x\|_2)x_k) = \left[\frac{\partial}{\partial x_k}(\varphi(\|x\|_2))\right] \cdot x_k + \varphi(\|x\|_2) = \varphi'(\|x\|_2)\frac{x_k^2}{\|x\|_2} + \varphi(\|x\|_2).$$

Daraus folgt

$$\operatorname{div} F(x) = \sum_{k=1}^{n} \left(\varphi'(\|x\|_2)\frac{x_k^2}{\|x\|_2} + \varphi(\|x\|_2) \right) = n \cdot \varphi(\|x\|_2) + \|x\|_2 \cdot \varphi'(\|x\|_2).$$

Nun leiten wir eine Produktregel her.

(1.9) Lemma. *a) Seien $U \subset \mathbb{R}^n$ offen und $f, g : U \to \mathbb{R}$ partiell differenzierbar. Dann ist $f \cdot g : U \to \mathbb{R}$ partiell differenzierbar mit*

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

b) Seien $U \subset \mathbb{R}^n$ offen und $F : U \to \mathbb{R}^n$ sowie $g : U \to \mathbb{R}$ partiell differenzierbar. Dann ist $g \cdot F : U \to \mathbb{R}^n$ partiell differenzierbar mit

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

Beweis. a) Es gilt für $k = 1, \dots, n$

$$\frac{\partial}{\partial x_k} f(x)g(x) = \left(\frac{\partial f}{\partial x_k}(x)\right) g(x) + f(x) \left(\frac{\partial g}{\partial x_k}(x)\right),$$

also die Behauptung.
b) Man hat für $k = 1, \dots, n$

$$\frac{\partial}{\partial x_k}(g(x)F_k(x)) = \frac{\partial g}{\partial x_k}(x) \cdot F_k(x) + g(x) \cdot \frac{\partial F_k}{\partial x_k}(x).$$

Daraus ergibt sich

$$\operatorname{div}(g \cdot F) = \sum_{k=1}^{n} \left(\frac{\partial g}{\partial x_k} \cdot F_k + g \cdot \frac{\partial F_k}{\partial x_k} \right) = \langle \operatorname{grad} g, F \rangle + g \cdot \operatorname{div} F.$$

□

Nun beschäftigen wir uns mit höheren Ableitungen.

(1.10) Definition. Sei $U \subset \mathbb{R}^n$ offen und $F : U \to \mathbb{R}^m$ eine Funktion.
a) F heißt *zweimal partiell differenzierbar*, wenn F partiell differenzierbar ist und alle partiellen Ableitungsfunktionen

$$D_i D_k F := D_i(D_k F) = \frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_k} F\right) =: \frac{\partial^2 F}{\partial x_i \partial x_k}$$

für $1 \leqslant i,\ k \leqslant n$ existieren.
b) F heißt *r-mal partiell differenzierbar*, falls alle partiellen Ableitungsfunktionen p-ter Ordnung

$$D_{i_p} D_{i_{p-1}} \ldots D_{i_1} F := D_{i_p}(\ldots(D_{i_1} F)) =: \frac{\partial^p F}{\partial x_{i_p} \ldots \partial x_{i_1}}, i_\nu \in \{1, \ldots, n\},$$

für $1 \leqslant p \leqslant r$ existieren.
c) F heißt *r-mal stetig partiell differenzierbar*, wenn alle partiellen Ableitungsfunktionen von der 1. bis zur r-ten Ordnung existieren und stetig sind.

Wir diskutieren zwei

(1.11) Beispiele. a) Die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x^3 \sin y,$$

erfüllt

$$\begin{aligned}
D_1 f &= 3x^2 \sin y,\ D_1 D_1 f = 6x \sin y,\ D_1 D_1 D_1 f = 6 \sin y, \ldots \\
D_2 f &= x^3 \cos y,\ D_2 D_2 f = -x^3 \sin y, \ldots \\
D_1(D_2 f) &= D_1(x^3 \cos y) = 3x^2 \cos y, \\
D_2(D_1 f) &= D_2(3x^2 \sin y) = 3x^2 \cos y.
\end{aligned}$$

b) Die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} xy \dfrac{x^2 - y^2}{x^2 + y^2}, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{cases}$$

ist partiell differenzierbar mit

$$\begin{aligned}
D_1 f \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{cases} y \dfrac{x^2 - y^2}{x^2 + y^2} + \dfrac{4x^2 y^3}{(x^2 + y^2)^2}, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{cases} \\
D_2 f \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{cases} x \dfrac{x^2 - y^2}{x^2 + y^2} - \dfrac{4x^3 y^2}{(x^2 + y^2)^2}, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{cases}
\end{aligned}$$

Daraus folgt

$$\begin{aligned} D_2D_1f\begin{pmatrix}0\\0\end{pmatrix} &= \lim_{y\to 0}\frac{1}{y}\left(D_1f\begin{pmatrix}0\\y\end{pmatrix} - D_1f\begin{pmatrix}0\\0\end{pmatrix}\right) = \lim_{y\to 0}\frac{-y}{y} = -1,\\ D_1D_2f\begin{pmatrix}0\\0\end{pmatrix} &= \lim_{x\to 0}\frac{1}{x}\left(D_2f\begin{pmatrix}x\\0\end{pmatrix} - D_2f\begin{pmatrix}0\\0\end{pmatrix}\right) = \lim_{x\to 0}\frac{x}{x} = 1. \end{aligned}$$

Also kann durchaus $D_1D_2f \neq D_2D_1f$ gelten. Ein allgemeines Kriterium für die Gleichheit liefert der folgende Satz von Hermann Amandus SCHWARZ (1843 - 1921).

(1.12) Satz von SCHWARZ. *Sei $U \subset \mathbb{R}^n$ offen und $F : U \to \mathbb{R}^m$ zweimal stetig partiell differenzierbar. Dann gilt*

$$D_jD_kF = D_kD_jF \quad \text{für alle } 1 \leqslant j,\, k \leqslant n.$$

Beweis. Ohne Einschränkung nehmen wir $m = 1$ an, denn andernfalls betrachten wir jede Komponentenfunktion $F_j : U \to \mathbb{R}$ einzeln, sowie $n = 2$, da alle anderen Koordinaten konstant sind. Der Fall $j = k$ ist trivial, sei also $j = 1$, $k = 2$.
Zu $\binom{a}{b} \in U$ wählen wir $r > 0$, so dass

$$K_r^{\|\cdot\|_\infty}\begin{pmatrix}a\\b\end{pmatrix} := \left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2; |x-a| < r,\ |y-b| < r\right\} \subset U.$$

Für $\binom{x}{y} \in K_r\binom{a}{b}$ liefert der Mittelwertsatz der Differentialrechnung VII(2.5) angewendet auf die Funktionen

$$t \mapsto F\begin{pmatrix}t\\y\end{pmatrix} - F\begin{pmatrix}t\\b\end{pmatrix} \quad \text{bzw.} \quad t \mapsto D_1F\begin{pmatrix}\xi\\t\end{pmatrix}$$

die Gleichung

$$\begin{aligned} &F\begin{pmatrix}x\\y\end{pmatrix} - F\begin{pmatrix}x\\b\end{pmatrix} - F\begin{pmatrix}a\\y\end{pmatrix} + F\begin{pmatrix}a\\b\end{pmatrix}\\ &= (x-a)\cdot\left(D_1F\begin{pmatrix}\xi\\y\end{pmatrix} - D_1F\begin{pmatrix}\xi\\b\end{pmatrix}\right) = (x-a)\cdot(y-b)\cdot D_2D_1F\begin{pmatrix}\xi\\\eta\end{pmatrix} \end{aligned}$$

für ein ξ zwischen a und x sowie für ein η zwischen y und b. Wendet man nun den Mittelwertsatz auf die Funktionen

$$u \mapsto F\begin{pmatrix}x\\u\end{pmatrix} - F\begin{pmatrix}a\\u\end{pmatrix} \quad \text{bzw.} \quad u \mapsto D_2F\begin{pmatrix}u\\\eta'\end{pmatrix}$$

an, so folgt

$$\begin{aligned} &F\begin{pmatrix}x\\y\end{pmatrix} - F\begin{pmatrix}x\\b\end{pmatrix} - F\begin{pmatrix}a\\y\end{pmatrix} + F\begin{pmatrix}a\\b\end{pmatrix}\\ &= (y-b)\cdot\left(D_2F\begin{pmatrix}x\\\eta'\end{pmatrix} - D_2F\begin{pmatrix}a\\\eta'\end{pmatrix}\right) = (y-b)\cdot(x-a)\cdot D_1D_2F\begin{pmatrix}\xi'\\\eta'\end{pmatrix} \end{aligned}$$

für ein η' zwischen y und b sowie ein ξ' zwischen x und a.
Also gilt für $x \neq a$ und $y \neq b$

$$D_2 D_1 F \begin{pmatrix} \xi \\ \eta \end{pmatrix} = D_1 D_2 F \begin{pmatrix} \xi' \\ \eta' \end{pmatrix} \quad \text{für} \quad x \neq a,\, y \neq b.$$

Für $x \to a$, $y \to b$ hat man $\xi \to a$, $\xi' \to a$, $\eta \to b$, $\eta' \to b$ und wegen der Stetigkeit von $D_1 D_2 F$ sowie $D_2 D_1 F$ auch

$$D_2 D_1 F \begin{pmatrix} a \\ b \end{pmatrix} = D_1 D_2 F \begin{pmatrix} a \\ b \end{pmatrix}.$$

□

Weil jede Permutation ein Produkt von Transpositionen ist, liefert eine Induktion sofort das

(1.13) Korollar. *Sei $U \subset \mathbb{R}^n$ offen und $F : U \to \mathbb{R}^m$ r-mal stetig partiell differenzierbar. Dann gilt für jede Permutation $\pi \in S_r$ sowie für alle $i_1, \dots, i_r \in \{1, \dots, n\}$*

$$D_{i_r} D_{i_{r-1}} \dots D_{i_1} F = D_{i_{\pi(r)}} D_{i_{\pi(r-1)}} \dots D_{i_{\pi(1)}} F.$$

Allgemeiner vereinbaren wir eine Bezeichnung in der

(1.14) Bemerkung. Ist $F : U \to \mathbb{R}^m$ r-mal stetig partiell differenzierbar, so definiert man für $a \in U$ und $k \leqslant r$ die *k-te Ableitung* von F in a rekursiv durch

$$D^k F(a)(v_1, \dots, v_k) := D\big(D^{k-1} F(x)(v_1, \dots, v_{k-1})\big)|_{x=a} v_k.$$

Beachte, dass $D^k F(a)$ eine k-multilineare Abbildung von $\underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{k-\text{mal}}$ in $\mathbb{R}^m$ ist. Nach dem Satz von SCHWARZ und dem Korollar ist $D^k F(a)$ eine symmetrische multilineare Abbildung.

Die nächsten Begriffe spielen bei partiellen Differentialgleichungen eine wichtige Rolle.

(1.15) Definition. Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Dann ist

$$\Delta f := \operatorname{div}\,(\operatorname{grad}\, f) = \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k^2}.$$

Man nennt f *harmonisch*, wenn $\Delta f = 0$.

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}$$

heißt LAPLACE-*Operator*.

Zum Abschluss diskutieren wir ein

(1.16) Beispiel. Sei $\varphi : \mathbb{R}_+^* \to \mathbb{R}$ zweimal stetig differenzierbar und

$$f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R},\ x \mapsto \varphi(\|x\|), \quad \|x\| := \|x\|_2.$$

Dann gilt

$$\frac{\partial}{\partial x_k} f = \varphi'(\|x\|)\frac{x_k}{\|x\|}, \quad \frac{\partial^2}{\partial x_k^2} f = \varphi''(\|x\|)\frac{x_k^2}{\|x\|^2} + \varphi'(\|x\|)\left(\frac{1}{\|x\|} - \frac{x_k^2}{\|x\|^3}\right),$$

$$\Delta f = \sum_{k=1}^n \frac{\partial^2 f}{\partial x_k^2} = \varphi''(\|x\|) + \varphi'(\|x\|)\frac{n-1}{\|x\|}.$$

Im Fall $n = 2$, $\varphi(t) = \ln t$ ist $\Delta(\ln \|x\|) = 0$, also $x \mapsto \ln \|x\|$ harmonisch. Im Fall $n \geqslant 3$, $\varphi(t) = t^{2-n}$ ist $\Delta(\|x\|^{2-n}) = 0$, also $x \mapsto \|x\|^{2-n}$ harmonisch.

§2. Totale Differenzierbarkeit

Bei der totalen Differenzierbarkeit handelt es sich um einen weiteren Begriff, der die Abhängigkeit von mehreren Variablen deutlicher zum Vorschein bringt. Nach VII(1.3) ist eine Funktion $f : \mathbb{R} \to \mathbb{R}$ genau dann differenzierbar in $a \in \mathbb{R}$, wenn es ein $c \in \mathbb{R}$ und eine in a stetige Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ gibt, so dass

$$f(x) = f(a) + c \cdot (x-a) + (x-a) \cdot \varphi(x), \quad x \in \mathbb{R}, \quad \varphi(a) = 0,$$

denn es gilt

$$\varphi(x) = \frac{f(x) - f(a)}{x - a} - c, \quad x \neq a, \quad c = f'(a).$$

Dieser Ansatz wird auf n Variablen verallgemeinert.

(2.1) Definition. Seien $U \subset \mathbb{R}^n$ offen, $a \in U$ und $F : U \to \mathbb{R}^m$ eine Funktion. F heißt *total differenzierbar* in a, wenn es eine Matrix $T \in \mathbb{R}^{m \times n}$ und eine in a stetige Funktion $\varphi : U \to \mathbb{R}^m$ gibt, so dass

$$F(x) = F(a) + T \cdot (x-a) + \|x-a\| \cdot \varphi(x), \quad x \in U, \quad \varphi(a) = 0.$$

F heißt *total differenzierbar*, wenn F in allen Punkten $a \in U$ total differenzierbar ist.

Anders formuliert: F ist genau dann in a total differenzierbar, wenn es ein $T \in \mathbb{R}^{m \times n}$ gibt derart, dass

$$\lim_{x \to a} \frac{1}{\|x-a\|}\big(F(x) - F(a) - T(x-a)\big) = 0.$$

Im Fall $m = 1$ bedeutet die totale Differenzierbarkeit die Existenz eines Vektors $g = T^t \in \mathbb{R}^n$ und einer in a stetigen Funktion $\varphi : U \to \mathbb{R}$, so dass

$$F(x) = F(a) + \langle g, x-a \rangle + \|x-a\| \cdot \varphi(x), \ \ x \in U, \ \ \varphi(a) = 0.$$

Die ersten einfachen Eigenschaften formulieren wir in dem

(2.2) Lemma. *Seien $U \subset \mathbb{R}^n$ offen, $a \in U$, $F : U \to \mathbb{R}^m$ eine Funktion mit $F = (F_1, \dots, F_m)^t$, $F_j : U \to \mathbb{R}$, $j = 1, \dots, m$.*
a) Die Definition (2.1) ist unabhängig von der Wahl der Norm $\| \cdot \|$.
b) Im Fall $n = m = 1$ ist F genau dann total differenzierbar in a, wenn F differenzierbar in a (im Sinne der Analysis I) ist.
c) Ist F total differenzierbar in a, so ist F stetig in a.
d) F ist genau dann total differenzierbar in a, wenn jede Komponentenfunktion F_j total differenzierbar in a ist.

Beweis. a) Seien $\| \cdot \|$ und $\| \cdot \|'$ Normen auf dem $\mathbb{R}^n$ und es gelte (2.1). Nach X(2.9) gibt es $\alpha > 0, \beta > 0$ mit der Eigenschaft

$$\alpha \|x\| \leqslant \|x\|' \leqslant \beta \|x\| \quad \text{für alle} \quad x \in \mathbb{R}^n.$$

Also ist (unter Benutzung der äquivalenten Charakterisierung):

$$\frac{1}{\|x-a\|'} \cdot \|F(x) - F(a) - T(x-a)\|' \leqslant \frac{\beta}{\alpha} \cdot \frac{1}{\|x-a\|} \cdot \|F(x) - F(a) - T(x-a)\|$$

und der Grenzwert links existiert und ist gleich 0, wenn derjenige rechts gleich 0 ist. Die umgekehrte Beweisrichtung geht analog.
b) ist offensichtlich.
c) Es gilt

$$\lim_{x \to a} F(x) = \lim_{x \to a} [F(a) + T \cdot (x-a) + \|x-a\| \cdot \varphi(x)] = F(a).$$

d) Die Behauptung ergibt sich aus

$$\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_m \end{pmatrix}, \; c = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}, \; \lim_{x \to a} \varphi(x) = \begin{pmatrix} \lim\limits_{x \to a} \varphi_1(x) \\ \vdots \\ \lim\limits_{x \to a} \varphi_m(x) \end{pmatrix}.$$

□

Wir diskutieren ein paar einfache

(2.3) Beispiele. a) Sei $A = (\alpha_{ij}) \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ und $F : \mathbb{R}^n \to \mathbb{R}^m, x \mapsto Ax + b$. Wegen

$$F(x) = Ax + b = Aa + b + A(x-a) = F(a) + A(x-a) + \|x-a\| \cdot 0$$

ist F in jedem Punkt $a \in \mathbb{R}^n$ total differenzierbar mit $T = A$.
b) Sei $A = (\alpha_{ij}) \in \mathbb{R}^{n \times n}$. Wir betrachten die quadratische Form

$$f : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \langle x, Ax \rangle = x^t A x = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j.$$

Dann gilt für $x \neq a$:

$$\begin{aligned} f(x) &= \langle x, Ax \rangle = \langle a + (x-a),\ A(a + (x-a)) \rangle \\ &= \langle a, Aa \rangle + \langle a, A(x-a) \rangle + \langle x-a, Aa \rangle + \langle x-a, A(x-a) \rangle \\ &= f(a) + [(A + A^t)a]^t \cdot (x-a) + \|x-a\|^2 \cdot f\left(\frac{1}{\|x-a\|}(x-a) \right). \end{aligned}$$

f ist als Polynom stetig und damit auf dem Kompaktum $\{y \in \mathbb{R}^n;\ \|y\| = 1\}$ beschränkt. Es folgt

$$\varphi(x) = \|x-a\| \cdot f\left(\frac{1}{\|x-a\|}(x-a)\right) \xrightarrow[x \to a]{} 0.$$

Also ist φ stetig in a und f total differenzierbar in a mit $T = [(A + A^t)a]^t$.
c) Die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{cases}$$

ist nach (1.3) c) partiell differenzierbar, aber nicht stetig in $\binom{0}{0}$. Nach (2.2) b) ist f somit nicht total differenzierbar in $\binom{0}{0}$.
d) Wir betrachten

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{cases} (x^2+y^2)\sin\left(\dfrac{1}{x^2+y^2}\right), & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{cases}$$

Dann gilt

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\| \cdot \varphi\begin{pmatrix} x \\ y \end{pmatrix},$$

wobei

$$\varphi\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \sqrt{x^2+y^2}\sin\left(\dfrac{1}{x^2+y^2}\right), & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ 0, & \text{falls } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{cases}$$

stetig in $\binom{0}{0}$ ist. Also ist f total differenzierbar in $\binom{0}{0}$ mit $T = (0,0)$. Es gilt

$$\begin{aligned} D_1 f\begin{pmatrix} x \\ y \end{pmatrix} &= 2x\sin\left(\frac{1}{x^2+y^2}\right) - \frac{2x}{x^2+y^2}\cos\left(\frac{1}{x^2+y^2}\right) \quad \text{für} \quad \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ D_1 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \lim_{x\to 0} \frac{1}{x}\left(f\begin{pmatrix} x \\ 0 \end{pmatrix} - f\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \lim_{x \to 0} x \sin\left(\frac{1}{x^2}\right) = 0. \end{aligned}$$

Wegen

$$D_1 f\begin{pmatrix} 1/\sqrt{k\pi} \\ 0 \end{pmatrix} = (-1)^{k+1} 2\sqrt{k\pi} \not\xrightarrow[k \to \infty]{} 0 = D_1 f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ist $D_1 f$ nicht stetig in $\binom{0}{0}$.

Nun beschreiben wir einen ersten Zusammenhang zwischen partieller und totaler Differenzierbarkeit, sowie totaler Differenzierbarkeit und Richtungsableitungen.

(2.4) Satz. *Seien $U \subset \mathbb{R}^n$ offen, $a \in U$ und $F : U \to \mathbb{R}^m$ total differenzierbar in a, also*

$$F(x) = F(a) + T \cdot (x-a) + \|x-a\| \cdot \varphi(x), \quad \varphi(a) = 0,$$

mit einer Matrix $T \in \mathbb{R}^{m\times n}$ und einer in a stetigen Funktion $\varphi : U \to \mathbb{R}^m$. Dann ist F in a partiell differenzierbar. T ist eindeutig bestimmt und es gilt

$$T = (DF)(a).$$

Weiterhin gilt für alle $v \in \mathbb{R}^m \backslash \{0\}$ die Gleichung

$$D_v F(a) = DF(a) \cdot v.$$

Also spielt T genau die Rolle der Ableitung im Fall $m = n = 1$.

Beweis. Sei F total differenzierbar. Für $v \in V\backslash\{0\}$ ist dann

$$\frac{1}{t}\big(F(a+tv) - F(a)\big) = Tv + \frac{1}{t}|t| \cdot \|v\| \cdot \varphi(tv),$$

und

$$\lim_{t\to 0} \left\| \frac{|t|}{t} \|v\| \varphi(tv) \right\| = \lim_{t\to 0} \|v\| \cdot \|\varphi(tv)\| = 0.$$

Also ist die letzte Aussage gezeigt. Für $v = e_i$ ergibt sich speziell

$$D_i F(a) = T \cdot e_i, \quad 1 \leqslant i \leqslant n;$$

also $T = DF(a)$. Dies zeigt auch die Eindeutigkeit von T. □

Wir haben aber schon in Beispiel (2.3) c) gesehen, dass aus der partiellen Differenzierbarkeit nicht die totale Differenzierbarkeit folgt. Man benötigt für diesen Schluss eine Zusatzvoraussetzung.

(2.5) Satz. *Seien $U \subset \mathbb{R}^n$ offen, $a \in U$ und $F : U \to \mathbb{R}^m$ partiell differenzierbar sowie in a stetig partiell differenzierbar. Dann ist f in a total differenzierbar und insbesondere auch stetig.*

Wir haben also die folgende Implikationskette, in der man keine Implikation umkehren kann.

$$\begin{array}{ccc} & f \text{ stetig partiell differenzierbar} & \\ & \Downarrow & \\ & f \text{ total differenzierbar} & \\ \swarrow & & \searrow \\ f \text{ partiell differenzierbar} & & f \text{ stetig} \end{array}$$

Ein Beispiel einer total differenzierbaren Funktion, die nicht stetig partiell differenzierbar ist, wurde in (2.3) d) angegeben.

Beweis. Nach (2.2) kann man ohne Einschränkung $m = 1$ annehmen. Sei $\delta > 0$, so dass $\{x \in \mathbb{R}^n; \|x-a\|_\infty < \delta\} \subset U$. Für $\|x-a\|_\infty < \delta$ gilt

$$F(x) - F(a) = \sum_{k=1}^{n} \left(F\begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ x_k \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix} - F\begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ a_k \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix} \right) = \sum_{k=1}^{n} (D_kF)\begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ \xi_k \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix} \cdot (x_k - a_k)$$

für ein ξ_k zwischen a_k und x_k nach dem Mittelwertsatz VII(2.5). Es folgt

$$F(x) - F(a) = \sum_{k=1}^{n} (D_kF)(a) \cdot (x_k - a_k) + \|x-a\|_2 \cdot \varphi(x),$$

wobei

$$\varphi(x) = \sum_{k=1}^{n} \Bigg(D_kF\Bigg(\begin{pmatrix} x_1 \\ \vdots \\ x_{k-1} \\ \xi_k \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix} \Bigg) - D_kF\Bigg(\begin{pmatrix} a_1 \\ \vdots \\ a_{k-1} \\ a_k \\ a_{k+1} \\ \vdots \\ a_n \end{pmatrix} \Bigg) \Bigg) \cdot \frac{x_k - a_k}{\|x-a\|_2} \quad \text{für } x \neq a.$$

Weil D_kF für $k = 1, \dots, n$ in a stetig ist und $\frac{|x_k - a_k|}{\|x-a\|_2} \leqslant 1$ gilt, erhält man

$$\lim_{x \to a} \varphi(x) = 0 =: \varphi(a).$$

Also ist F total differenzierbar in a. □

Diesen Satz nehmen wir zum Anlass für die

(2.6) Definition. Sei $U \subset \mathbb{R}^n$ offen. Eine Funktion $F : U \to \mathbb{R}^m$ heißt *stetig differenzierbar*, wenn F total differenzierbar und die Abbildung $DF : U \to \mathbb{R}^{m \times n}$ stetig ist.

Aus (2.5) und (2.4) ergibt sich damit sofort das

(2.7) Korollar. *Sei $U \subset \mathbb{R}^n$ offen. Für eine Funktion $F : U \to \mathbb{R}^m$ sind äquivalent:*

(i) F ist stetig partiell differenzierbar.

(ii) F ist stetig differenzierbar.

Im nächsten Schritt beweisen wir die *Kettenregel*.

(2.8) Satz. *Seien $U \subset \mathbb{R}^n$ offen, $V \subset \mathbb{R}^m$ offen, $F : U \to \mathbb{R}^m$ total differenzierbar in $a \in U$ mit $F(U) \subset V$ sowie $G : V \to \mathbb{R}^p$ total differenzierbar in $b = F(a)$. Dann ist $H := G \circ F : U \to \mathbb{R}^p$ total differenzierbar in a mit*

$$(D(G \circ F))(a) = (DG)(F(a)) \cdot (DF)(a),$$

d. h.

$$\frac{\partial H_i}{\partial x_j}(a) = \sum_{k=1}^{m} \frac{\partial G_i}{\partial y_k}(F(a)) \cdot \frac{\partial F_k}{\partial x_j}(a), \quad 1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant n.$$

Beweis. Sei

$$F(x) = F(a) + (DF)(a) \cdot (x-a) + \|x-a\|_2 \cdot \varphi(x), \quad \varphi(a) = 0,$$

wobei φ stetig in a ist. Darüber hinaus erhalten wir eine in $b = F(a)$ stetige Funktion ψ mit der Eigenschaft

$$G(y) = G(b) + (DG)(b) \cdot (y-b) + \|y-b\|_2 \cdot \psi(y), \quad \psi(b) = 0.$$

Dann gilt für alle $x \in U$

$$\begin{aligned} H(x) &= G(F(x)) = G(b) + (DG)(b) \cdot (F(x)-b) + \|F(x)-b\|_2 \cdot \psi(F(x)) \\ &= H(a) + (DG)(b) \cdot (DF)(a) \cdot (x-a) + \|x-a\|_2 \cdot (DG)(b) \cdot \varphi(x) \\ &\quad + \|F(x)-F(a)\|_2 \cdot \psi(F(x)) \\ &= H(a) + [(DG)(b) \cdot (DF)(a)] \cdot (x-a) + \|x-a\|_2 \cdot \vartheta(x), \end{aligned}$$

wobei $\vartheta(a) = 0$ und für $x \neq a$

$$\vartheta(x) = (DG)(b) \cdot \varphi(x) + \left\| \frac{1}{\|x-a\|_2} \cdot (DF)(a) \cdot (x-a) + \varphi(x) \right\|_2 \cdot \psi(F(x)) \xrightarrow[x \to a]{} 0.$$

Denn zunächst ist $z \mapsto (DF)(a)z$ LIPSCHITZ-stetig, somit gibt es ein $K > 0$ derart, dass

$$\|(DF)(a) \cdot (x-a)\|_2 \leqslant K \cdot \|x-a\|_2.$$

Weiter ergibt sich also

$$\left\| \frac{1}{\|x-a\|_2}(DF)(a) \cdot (x-a) + \varphi(x) \right\|_2 \leqslant K + \|\varphi(x)\|_2,$$

und dies ist beschränkt in einer Umgebung von a, da $\lim_{x\to a} \varphi(x) = 0$.
Schließlich ist $\lim_{x\to a} F(x) = F(a)$, also aus Stetigkeitsgründen

$$\lim_{x\to a} \psi(F(x)) = \psi(F(a)) = 0.$$

Also ist ϑ stetig in a und H somit total differenzierbar mit

$$(DH)(a) = (DG)(b) \cdot (DF)(a). \qquad \square$$

Warnung: Die Formel bleibt nicht richtig, wenn man G nur als partiell differenzierbar vorausgesetzt.

Zur Veranschaulichung diskutieren wir ein

(2.9) Beispiel. Wir betrachten

$$F : \mathbb{R}^3 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} xy \\ ye^z \end{pmatrix}, \quad G : \mathbb{R}^2 \to \mathbb{R}^4, \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} \sin u \\ \cos v \\ u^2 \\ uv^2 \end{pmatrix}, \quad H = G \circ F.$$

Dann gilt

$$\begin{aligned} DF \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} y & x & 0 \\ 0 & e^z & ye^z \end{pmatrix}, \quad DG \begin{pmatrix} u \\ v \end{pmatrix}, = \begin{pmatrix} \cos u & 0 \\ 0 & -\sin v \\ 2u & 0 \\ v^2 & 2uv \end{pmatrix}, \\ DH \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= DG \left(F \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) \cdot DF \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos(xy) & 0 \\ 0 & -\sin(ye^z) \\ 2xy & 0 \\ y^2 e^{2z} & 2xy^2 e^z \end{pmatrix} \begin{pmatrix} y & x & 0 \\ 0 & e^z & ye^z \end{pmatrix} \\ &= \begin{pmatrix} y\cos(xy) & x\cos(xy) & 0 \\ 0 & -e^z \sin(ye^z) & -ye^z \sin(ye^z) \\ 2xy^2 & 2x^2 y & 0 \\ y^3 e^{2z} & 3xy^2 e^{2z} & 2xy^3 e^{2z} \end{pmatrix}. \end{aligned}$$

Der eigentliche Wert der Kettenregel liegt aber in der Theorie, wie wir noch sehen werden.

Kapitel XII.

Ergänzungen

In diesem Kapitel sind einige Sätze und Fakten zusammengestellt, deren Kenntnis und Verständnis für die Veranstaltung Analysis II nicht unbedingt nötig ist. Allerdings können sie im weiteren Verlauf von Analysis-nahen Lehrveranstaltungen relevant werden.

§1. Monotone und konvexe Funktionen

In diesem Abschnitt sollen nochmal monotone und konvexe Funktionen betrachtet werden. Insbesondere geht es um (Un-)Stetigkeitsstellen dieser Klassen von Funktionen. Wie sich zeigt, ergeben die Eigenschaften Monotonie bzw. Konvexität hier starke Restriktionen.

Wir beginnen mit der Charakterisierung einer speziellen Klasse von Unstetigkeitsstellen.

(1.1) Definition. Sei $f\colon D \to \mathbb{R}$ eine Funktion und $x_0 \in \mathbb{R}$ innerer Punkt von $D \cup \{x_0\}$. Man nennt x_0 *Sprungstelle* von f, wenn $\lim\limits_{x\uparrow x_0} f(x)$ und $\lim\limits_{x\downarrow x_0} f(x)$ beide existieren, aber verschieden sind.

Damit lassen sich Unstetigkeitsstellen monotoner Funktionen analysieren; insbesondere "gibt es nicht so viele von diesen".

(1.2) Satz. *Seien $a < b$ und $f : [a, b] \to \mathbb{R}$ eine monoton wachsende Funktion.*
a) Für jedes $x_0 \in (a, b]$ und jedes $x_1 \in [a, b)$ existieren

$$\lim_{x\uparrow x_0} f(x) = \sup\{f(x) \mid a \leqslant x < x_0\} \quad \textit{und} \quad \lim_{x\downarrow x_1} f(x) = \inf\{f(x) \mid x_1 < x \leqslant b\}.$$

b) Ist f nicht stetig in $x_0 \in (a, b)$, so ist x_0 eine Sprungstelle von f.
c) Die Anzahl der Unstetigkeitsstellen von f ist abzählbar.

Beweis. a) Sei $L := \sup\{f(x) \mid a \leqslant x < x_0\}$. Da f monoton wachsend ist, gilt $L \leqslant f(x_0)$. Zu $\varepsilon > 0$ existiert ein $x' \in [a, x_0)$ mit $L - \varepsilon < f(x') \leqslant L$. Weil f monoton wächst, folgt $L - \varepsilon < f(x) \leqslant L$ für alle $x \in [x', x_0)$. Mit $\delta := x_0 - x'$ erhält man $L - \varepsilon < f(x) \leqslant L$, also $|f(x) - L| < \varepsilon$, für alle x mit $0 < x_0 - x < \delta$. Es folgt $\lim\limits_{x \uparrow x_0} f(x) = L$.
Der zweite Teil wird analog bewiesen.
b) Man verwende a) und (1.1).
c) Die Menge der Unstetigkeitsstellen in (a, b) ist nach a)

$$M = \{x_0 \in (a,b);\ \lim_{x \uparrow x_0} f(x) < \lim_{x \downarrow x_0} f(x)\}.$$

Zu $x_0 \in M$ wählt man $r(x_0) \in \mathbb{Q}$ mit

$$\lim_{x \uparrow x_0} f(x) < r(x_0) < \lim_{x \downarrow x_0} f(x)$$

nach II(2.26). Dann gilt $f(x) < r(x_0) < f(x')$ für alle $a \leqslant x < x_0 < x' \leqslant b$. Daraus folgt sofort $r(x_0) < r(x_1)$ für alle $x_0, x_1 \in M$ mit $x_0 < x_1$. Demnach ist die Abbildung $r : M \to \mathbb{Q}$ streng monoton wachsend und somit injektiv. Weil $\mathbb{Q}$ nach II(4.6) abzählbar ist, ist dann auch M abzählbar. $\square$

Der Einordnung dienen die

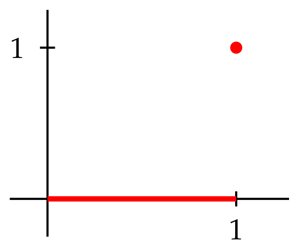
(1.3) Bemerkungen. a) Natürlich kann man monoton fallende Funktionen f analog behandeln (oder $-f$ betrachten). Auf die Monotonie von f kann man in c) nicht verzichten, da die DIRICHLETsche Sprungfunktion in jedem Punkt aus $\mathbb{R}$ unstetig ist.
b) Satz (1.2) gilt auch im Fall, dass x^* innerer Punkt von $[a, b]$ und $f\colon [a, b] \setminus \{x^*\}$ monoton ist. Im Beweis von a) muss nur für $x_0 = x^*$ die Begründung der Existenz von L modifiziert werden. Dazu nimm ein $x_1 > x^*$. Wegen $f(x) \leqslant f(x_1)$ für alle $x \in [a, x^*)$ ist $\{f(x);\ a \leqslant x < x_0\}$ nach oben beschränkt.

Nun wollen wir konvexe Funktionen betrachten. Nach VII(4.1) heißt eine reellwertige Funktion f auf einem Intervall I konvex, wenn für alle $x_0, x_1 \in I$ und $\lambda \in [0, 1]$ gilt

$$f\big((1-\lambda)x_0 + \lambda x_1\big) \leqslant (1-\lambda) f(x_0) + \lambda f(x_1).$$

Es ist nicht richtig, dass konvexe Funktionen schon stetig sind.

(1.4) Beispiel. Sei $f\colon [0,1] \to \mathbb{R}$, $x \mapsto \begin{cases} 0 & \text{für } x < 1, \\ 1 & \text{für } x = 1. \end{cases}$
Dann ist f konvex: Für $x_0 < 1$ und $x_1 < 1$ ergibt die Bedingung auf beiden Seiten 0, ist also trivialerweise erfüllt. Für (etwa) $x_0 < 1$ und $x_1 = 1$ ergibt die linke Seite Null und die rechte Seite λ, falls $\lambda < 1$, und beide Seiten ergeben 1, wenn $\lambda = 1$. Aber f ist in 1 nicht stetig.



Wir zeigen aber nun, dass Unstetigkeit bei konvexen Funktionen nur in Randpunkten auftreten kann.

(1.5) Satz. *Sei $I \subset \mathbb{R}$ ein nicht ausgeartetes Intervall, x^* ein innerer Punkt von I und $f\colon I \to \mathbb{R}$ konvex. Dann ist f stetig in x^*.*

Beweis. (i) Betrachte

$$g\colon I\backslash\{x^*\} \to \mathbb{R}, \quad g(x)\colon = \frac{f(x)-f(x^*)}{x-x^*}.$$

Ist $x_0 < x^* < x_1$ mit $x_0, x_1 \in I$ und $\lambda \in (0,1)$ derart, dass $x^* = (1-\lambda)x_0 + \lambda x_1$, so ist

$$g(x_0) = \frac{f(x_0)-f(x^*)}{x_0-x^*} \leqslant \frac{f(x_1)-f(x^*)}{x_1-x^*} = g(x_1)$$

nach dem Beweis von VII(4.2), erster Teil. Ist $x^* < x_0 < x_1$ und λ so, dass $x_0 = (1-\lambda)x^* + \lambda x_1$, so ergibt derselbe Beweis

$$\frac{f(x_0)-f(x^*)}{x_0-x^*} \leqslant \frac{f(x_1)-f(x^*)}{x_1-x^*}$$

(mit passend geänderten Bezeichnungen). Also ist $g|_{(x^*,\infty)\cap I}$ monoton wachsend und analog folgt, dass $g|_{(-\infty,x^*)\cap I}$ monoton wächst. Also ist g auf $I\backslash\{x^*\}$ monoton wachsend.

(ii) Nach (1.2) existieren nun $\lim\limits_{x\uparrow x^*} g(x)$ und $\lim\limits_{x\downarrow x^*} g(x)$. Wegen $f(x)-f(x^*) = (x-x^*)\cdot g(x)$ und $\lim\limits_{x\to x^*}(x-x^*) = 0$ folgt nun

$$\lim_{x\uparrow x^*}\big(f(x)-f(x^*)\big) = \lim_{x\downarrow x^*}\big(f(x)-f(x^*)\big) = 0.$$

Also ist f in x^* stetig. □

§2. Mehr zu Potenzreihen und TAYLORreihen.

Wir betrachten wieder reelle Potenzreihen und insbesondere ihr Konvergenzverhalten in Randpunkten.

(2.1) ABELscher Grenzwertsatz. *Sei $f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k$ eine reelle Potenzreihe mit Konvergenzradius $0 < R < \infty$, die im Punkt $x = x_0 + R$ konvergiere. Dann konvergiert die Potenzreihe gleichmäßig auf $[x_0, x_0+R]$ und es gilt*

$$\lim_{x\uparrow x_0+R} f(x) = \lim_{x\uparrow x_0+R} \sum_{k=0}^{\infty} a_k(x-x_0)^k = \sum_{k=0}^{\infty} a_k R^k.$$

Analoges gilt natürlich für $x_0 - R$ statt $x_0 + R$.

Beweis. Sei ohne Einschränkung $x_0 = 0$ und $R = 1$, da wir andernfalls die Reihe $\sum_{k=0}^{\infty} b_k y^k$ mit $b_k = a_k R^k$ und $y = (x - x_0)/R$ betrachten können. Sei

$$s_n(x) = \sum_{k=0}^{n} a_k x^k, \quad s_n = \sum_{k=0}^{n} a_k = s_n(1).$$

Für $0 \leqslant x \leqslant 1$ und $n, p \in \mathbb{N}$ gilt (vgl. VII(4.1))

$$\begin{aligned}
s_{n+p}(x) - s_n(x) &= a_{n+1}x^{n+1} + \ldots + a_{n+p}x^{n+p} \\
&= (s_{n+1} - s_n)x^{n+1} + \ldots + (s_{n+p} - s_{n+p-1})x^{n+p} \\
&= (s_{n+1} - s_n)x^{n+1} + ((s_{n+2} - s_n) - (s_{n+1} - s_n))\, x^{n+2} + \ldots \\
&\quad + \left((s_{n+p} - s_n) - (s_{n+p-1} - s_n)\right) x^{n+p} \\
&= (s_{n+1} - s_n) \left(x^{n+1} - x^{n+2}\right) + (s_{n+2} - s_n) \left(x^{n+2} - x^{n+3}\right) + \ldots \\
&\quad + (s_{n+p-1} - s_n) \left(x^{n+p-1} - x^{n+p}\right) + (s_{n+p} - s_n)x^{n+p}.
\end{aligned}$$

Weil $(s_n)_{n\geqslant 1}$ konvergiert, gibt es zu $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass $|s_{n+k} - s_n| < \varepsilon$ für alle $n, k \in \mathbb{N}$, $n \geqslant N$. Aus $0 \leqslant x \leqslant 1$ folgt $x^k - x^{k+1} \geqslant 0$. Daher hat man für alle $n \geqslant N$, $p \in \mathbb{N}$

$$\begin{aligned}
|s_{n+p}(x) - s_n(x)| &\leqslant \varepsilon \left(\left(x^{n+1} - x^{n+2}\right) + \left(x^{n+2} - x^{n+3}\right) + \ldots + \left(x^{n+p-1} - x^{n+p}\right) + x^{n+p} \right) \\
&= \varepsilon x^{n+1} \leqslant \varepsilon.
\end{aligned}$$

Also konvergiert $(s_n(x))_{n\geqslant 1}$ nach IX(1.7) gleichmäßig auf $[0,1]$. Da alle $s_n(x)$ als Polynomfunktionen stetig sind, folgt die fehlende Behauptung aus IX(2.4). □

Bereits früher (vgl. IV(1.11)) angekündigt wurden die

(2.2) Beispiele. a) Es gilt $\ln(1 + x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k}$ für alle $-1 < x < 1$ mit $R = 1$. Die Reihe $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ konvergiert für $x = 1$ nach dem LEIBNIZ-Kriterium IV(1.11). Also folgt aus (2.1)

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \lim_{x\uparrow 1} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} = \lim_{x\uparrow 1} \ln(1 + x) = \ln 2.$$

b) Es gilt $\arctan x = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1}$ mit $R = 1$. Analog folgt

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \lim_{x\uparrow 1} \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} = \lim_{x\uparrow 1} \arctan x = \arctan 1 = \frac{\pi}{4}.$$

Ist f beliebig oft differenzierbar, so hatten wir f die TAYLOR-Reihe um x_0 (vgl. VII(3.12))

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

zugeordnet, die aber nicht immer gegen f konvergiert.

(2.3) Satz. *Sei I ein Intervall und $f : I \to \mathbb{R}$ beliebig oft differenzierbar.*
a) Ist die Folge $\left(f^{(n)}\right)_{n\geqslant 0}$ der Ableitungen von f gleichgradig beschränkt auf I, so gilt für jedes $x_0 \in I$

$$f(x) = T_{f,x_0}(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \quad \text{für alle} \quad x \in I.$$

b) Gibt es ein $M \in \mathbb{R}$, $x_0 \in I$ und $\rho > 0$ mit $U_\rho(x_0) \subset I$ sowie $|f^{(n)}(x)| \leqslant \frac{n!\cdot M}{\rho^n}$ für alle $n \in \mathbb{N}$ und alle $x \in U_\rho(x_0)$, so folgt

$$f(x) = T_{f,x_0}(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \quad \text{für alle} \quad x \in U_\rho(x_0).$$

Beweis. Aus der TAYLOR-Formel VII(3.5) folgt für alle $x \in I$, $x \neq x_0$

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{n!}(x-x_0)^n + R_n(x), \quad R_n(x) = \frac{(x-x_0)^n}{n!} \cdot f^{(n)}(\xi)$$

für ein $\xi = \xi_n \in (x_0, x)$ bzw. $\xi \in (x, x_0)$ und alle $n \in \mathbb{N}$.
a) Gilt $|f^{(n)}(x)| \leqslant M$ für alle $x \in I$ und alle $n \in \mathbb{N}_0$ und bezeichnet $L > 0$ die Länge eines beschränkten Teilintervalls I', so folgt für alle $x, x_0 \in I'$

$$|R_n(x)| = \frac{|x-x_0|^n}{n!}|f^{(n)}(\xi)| \leqslant \frac{L^n}{n!}M \xrightarrow[n\to\infty]{} 0.$$

Weil I' beliebig ist, folgt die Behauptung.
b) Für alle $x \in U_\rho(x_0)$ gilt $|x-x_0| < \rho$ und

$$|R_n(x)| \leqslant \left(\frac{|x-x_0|}{\rho}\right)^n M \xrightarrow[n\to\infty]{} 0.$$

Also konvergiert die TAYLOR-Reihe in diesen Fällen gegen die Funktion f. □

§3. Funktionenfolgen in normierten Vektorräumen

In Kapitel IX haben wir Folgen von reellwertigen Funktionen (einer reellen Variablen) betrachtet. Die Begriffe, und insbesondere der Begriff gleichmäßige Konvergenz, lassen sich auf normierte Vektorräume verallgemeinern.

(3.1) Definition. Sei $M \neq \emptyset$ eine Menge, $(V', \|\cdot\|')$ ein normierter Vektorraum und $(f_k)_{k\geqslant 1}$ eine Folge von Funktionen $f_k : M \to V'$. Die *Funktionenfolge* $(f_k)_{k\geqslant 1}$ heißt *punktweise konvergent* gegen die Funktion $f : M \to V'$, wenn $\lim\limits_{k\to\infty} f_k(x) = f(x)$ für alle $x \in M$ gilt, d. h. wenn zu jedem $x \in M$ und jedem $\varepsilon > 0$ ein $N = N(x, \varepsilon) \in \mathbb{N}$ existiert mit $\|f_k(x) - f(x)\|' < \varepsilon$

für alle $k \geqslant N$. Die *Funktionenfolge* $(f_k)_{k\geqslant 1}$ heißt *gleichmäßig konvergent* gegen die Funktion $f : M \to V'$, wenn es zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$ gibt, so dass

$$\|f_k(x) - f(x)\|' < \varepsilon \quad \text{für alle } \ k \geqslant N \text{ und alle } \ x \in M.$$

Ist zudem M Teilmenge eines normierten Vektorraums $(V, \|\cdot\|)$, so heißt die Funktionenfolge $(f_k)_{k\geqslant 1}$ *lokal gleichmäßig konvergent*, wenn es zu jedem Punkt $a \in M$ eine Umgebung $U \subset M$ von a gibt, so dass $(f_k|_U)_{k\geqslant 1}$ gleichmäßig konvergiert.

Aus der (lokal) gleichmäßigen Konvergenz folgt wieder die punktweise Konvergenz. Wörtlich wie in IX(2.4) zeigt man den

(3.2) Satz. *Seien $(V, \|\cdot\|)$ und $(V', \|\cdot\|')$ normierte Vektorräume. Ist $(f_k)_{k\geqslant 1}$ eine Folge von stetigen Funktionen $f_k : D \to V'$, $D \subset V$, die lokal gleichmäßig gegen $f : D \to V'$ konvergiert, dann ist auch f stetig.*

Wir wollen auch noch Differenzierbarkeit für Funktionenfolgen auf $\mathbb{R}^n$ diskutieren. Dabei wird nicht die "bestmögliche" allgemeine Version vorgestellt, sondern eine, die für den praktischen Gebrauch meistens reicht.

(3.3) Satz. *Sei $U \subset \mathbb{R}^n$ offen und nicht-leer, und für jedes $k \in \mathbb{N}$ sei $f_k : U \to \mathbb{R}^m$ stetig differenzierbar. Weiter konvergiere die Folge der f_k lokal gleichmäßig gegen $f : U \to \mathbb{R}^m$, und für jedes $i \in \{1, \dots, n\}$ konvergiere die Folge $(D_i f_k)_{k\geqslant 1}$ der partiellen Ableitungen lokal gleichmäßig gegen eine Funktion $g_i : U \to \mathbb{R}^m$.*
Dann ist f stetig partiell differenzierbar auf U, und $D_i f = g_i$ für $1 \leqslant i \leqslant n$.

Beweis. (Skizze) Wegen XI(1.5) und XI(2.7) kann $m = 1$ angenommen werden. Setze

$$H_k := \begin{pmatrix} D_1 f_k \\ \vdots \\ D_n f_k \end{pmatrix}, \quad k \in \mathbb{N} \quad \text{sowie} \quad G := \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}.$$

Nach (3.2) sind G und $f = \lim\limits_{k\to\infty} f_k$ stetig auf U.
Da Differenzierbarkeit eine lokale Eigenschaft ist, können wir i. F. annehmen, dass $U = K_\rho(a)$ für ein $\rho > 0$ (insbesondere U konvex), und $H_k \underset{glm}{\to} G$, $f_k \underset{glm}{\to} f$.
Setze für festes $x \in U$ nun $\gamma_x : [0,1] \to U$, $t \mapsto a + t(x-a)$. Dann gilt für alle k:

$$f_k(x) = f_k(a) + \int\limits_{\gamma_x} H_k \, \mathrm{d}x,$$

also

$$f_k(x) = f_k(a) + \int\limits_0^1 \langle H_k(a + t(x-a)), x - a\rangle \, \mathrm{d}t.$$

Weil (H_k) gleichmäßig gegen G konvergiert, ist auch die Funktionenfolge der

$$\psi_k : [0,1] \to \mathbb{R}, \quad t \mapsto \langle H_k(a + t(x-a)), x - a \rangle$$

gleichmäßig konvergent gegen

$$\vartheta : [0,1] \to \mathbb{R}, \quad t \mapsto \langle G(a + t(x-a)), x - a \rangle.$$

Mit dem Vertauschungssatz IX(2.6) ist also

$$f(x) = \lim_{k\to\infty} f_k(x) = f(a) + \int_0^1 \langle G(a + t(x-a)), x - a \rangle \, \mathrm{d}t,$$

und dies gilt für alle $x \in K_\rho(a)$. Nun folgt $D_i f = g_i$. $\square$

Anhang: Fakten aus der Linearen Algebra

Wir werden hier einige Begriffe und Ergebnisse aus der Linearen Algebra zusammenstellen, die (evtl. etwas später) in der zugehörigen Veranstaltung ausführlicher behandelt werden. Für den Umgang in der Analysis reicht zunächst diese Sammlung.

A. 1. Vektorräume in der Analysis

(A 1.1) Vektorräume. Gegeben sei ein Körper K. Eine Klasse von (Standard-) Beispielen für K-*Vektorräume* erhält man bekanntlich wie folgt:
Sei M eine nicht-leere Menge und

$$V :\, = \{f;\ f : M \to K\}$$

die Menge aller Abbildungen von M nach K. Man definiert Operationen $+ : V \times V \to V$ und $\cdot : K \times V \to V$ "punktweise", also für $f, g \in V$, $\lambda \in K$:

$$\begin{aligned}(f+g)(x) :&= f(x) + g(x), \quad x \in M,\\ (\lambda f)(x) :&= \lambda \cdot f(x), \quad x \in M,\end{aligned}$$

wobei rechts Addition und Multiplikation in K zu nehmen sind.
Mit diesen Verknüpfungen wird $(V, K, +, \cdot)$ zu einem K-Vektorraum.
Jede Teilmenge W von V, die bezüglich der Operationen abgeschlossen ist, ist mit den eingeschränkten Operationen selbst ein Vektorraum. Man spricht von einem *Untervektorraum*.

In der Analysis geht es in der Regel um $K = \mathbb{R}$ oder $K = \mathbb{C}$ und M ist Teilmenge eines normierten Vektorraums.
Eine leichte Verallgemeinerung obiger Konstruktion geht von Abbildungen $M \to X$ aus, wobei X ein (normierter) Vektorraum ist.

(A 1.2) Folgenräume. a) Die Menge aller reellen Folgen $(a_n)_{n\in\mathbb{N}}$ – das ist die Menge aller Abbildungen $\mathbb{N} \to \mathbb{R}$ – ist mit den punktweisen Verknüpfungen ein $\mathbb{R}$-Vektorraum.

b) Ein prominenter Untervektorraum ist die Menge aller konvergenten Folgen, wie aus bekannten Grenzwertsätzen folgt. Ist z. B. $(a_n)_{n\in\mathbb{N}}$ konvergent mit Grenzwert a und $(b_n)_{n\in\mathbb{N}}$ konvergent mit Grenzwert b, so ist auch $(a_n + b_n)_{n\in\mathbb{N}}$ konvergent mit Grenzwert $a + b$.
Ein Untervektorraum hiervon ist z. B. die Menge aller Nullfolgen.

c) Andere Untervektorräume (unvollständige Aufzählung):

- Menge aller beschränkten Folgen
- Menge aller CAUCHY-Folgen
- Menge aller Folgen $(a_n)_{n\in\mathbb{N}}$, für die $\sum_{n=1}^{\infty} |a_n|$ konvergiert.

Eine weitere Verallgemeinerung sind die

(A 1.3) Funktionenräume. Reellwertige Funktionen einer reellen Variablen:
Es sei $D \subset \mathbb{R}$ und $V = \{f;\ f : D \to \mathbb{R}\}$. Folgende Untervektorräume von V sind z. B. relevant für die Analysis:

- $\mathcal{C}(D)$: Menge aller stetigen Funktionen.
- Falls D offen, $\mathcal{C}^1(D)$: Menge aller stetig differenzierbaren Funktionen.
- Falls $D = [a, b]$, $\mathcal{R}\big([a, b]\big)$: Menge aller RIEMANN-integrierbaren Funktionen.

Dass hier wirklich Untervektorräume vorliegen, ist nicht trivial, sondern folgt aus Sätzen, die wir bewiesen haben.
Beispiel: Sind f und g stetig in $x_0 \in D$, so auch $f + g$.

Natürlich kann man die Begriffe (und teils die Ergebnisse) auf Funktionen ausweiten, die von einer Teilmenge eines normierten Vektorraums augehen oder Werte in einem normierten Vektorraum haben.

A. 2. Matrixdarstellungen linearer Abbildungen

Wir benötigen im Zusammenhang mit der Differentialrechnung folgendes Faktum für den Fall $K = \mathbb{R}$:

(A 2.1) Proposition. *Es sei $T : K^n \to K^m$ eine lineare Abbildung. Dann gibt es eine Matrix*

$$A \in \mathrm{Mat}(m \times n, K),$$

so dass

$$T(v) = A \cdot v \quad \textit{für alle} \ \ v \in K^n.$$

Diese Aussage wird (in größerem Rahmen) in der Linearen Algebra I bewiesen. Wir liefern hier einen Beweis, der für die spezielle Situation ausreicht.

Beweis. (i) Es sei $(e_1, \ldots, e_n)$ die Standardbasis des K^n und $a_j := T(e_j)$, $1 \leqslant j \leqslant n$. Für jedes

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = v_1 e_1 + \cdots + v_n e_n \in K^n$$

gilt dann wegen Linearität:

$$T(v) = T\left(\sum_{j=1}^{n} v_j e_j\right) = \sum_{j=1}^{n} v_j T(e_j) = \sum_{j=1}^{n} v_j a_j.$$

(ii) Andererseits gilt – nach Definition von $A \cdot v$ – für die Matrix $A = (a_1, \ldots, a_n)$ in Spaltendarstellung auch

$$A \cdot v = (a_1, \ldots, a_n) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{j=1}^{n} v_j a_j.$$

□

A. 3. Definite und indefinite reelle symmetrische Matrizen

Wir fassen hier einige Fakten und Sätze zusammen, die bei der Diskussion relativer Extrema von Funktionen mehrerer Veränderlicher relevant sind.

(A 3.1) Matrizen und symmetrische Bilinearformen. Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch, d. h. $A^{\mathrm{tr}} = A$. Durch

$$\alpha\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad (x, y) \mapsto x^{\mathrm{tr}} A y,$$

ist dann eine symmetrische Bilinearform auf $\mathbb{R}^n$ definiert. Insbesondere gilt also

$$\alpha(x, y) = \alpha(y, x) \quad \text{für alle} \quad x, y \in \mathbb{R}^n.$$

Ist umgekehrt $\beta\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform und $(e_1, \ldots, e_n)$ die Standardbasis des $\mathbb{R}^n$, so definiert man die symmetrische Matrix

$$B = \big(\beta(e_i, e_j)\big)_{1 \leqslant i, j \leqslant n} \in \mathrm{Mat}(n \times n, \mathbb{R}).$$

Wegen der Bilinearität von β gilt dann

$$\beta(x, y) = x^{\mathrm{tr}} B y \quad \text{für alle} \quad x, y \in \mathbb{R}^n.$$

Speziell betrachten wir

(A 3.2) Definite und indefinite Bilinearformen. Eine symmetrische Bilinearform α auf $\mathbb{R}^n$ heißt

- *positiv semi-definit*, wenn $\alpha(x,x) \geqslant 0$ für alle $x \in \mathbb{R}^n$
- *positiv definit*, wenn $\alpha(x,x) > 0$ für alle $x \in \mathbb{R}^n \backslash \{0\}$
- *negativ semi-definit*, wenn $\alpha(x,x) \leqslant 0$ für alle $x \in \mathbb{R}^n$
- *negativ definit*, wenn $\alpha(x,x) < 0$ für alle $x \in \mathbb{R}^n \backslash \{0\}$
- *indefinit*, wenn es $z, w \in \mathbb{R}^n$ gibt mit $\alpha(z,z) < 0$ und $\alpha(w,w) > 0$.

Die Namensgebung wird auf die zugeordneten symmetrischen Matrizen übertragen.

(A 3.3) Beispiel. Eine Diagonalmatrix $D = \operatorname{diag}(\mu_1, \ldots, \mu_n)$ ist

- positiv semi-definit, falls alle $\mu_j \geqslant 0$;
- positiv definit, falls alle $\mu_j > 0$;
- negativ semi-definit, falls alle $\mu_j \leqslant 0$;
- negativ definit, falls alle $\mu_j < 0$;
- indefinit, falls es Indizes k, l gibt mit $\mu_k < 0$, $\mu_l > 0$.
  (Bestätigung via $x^{\mathrm{tr}} D x = \sum \mu_j x_j^2$.)

(A 3.4) Beobachtung. Es sei A symmetrisch, $Q \in \operatorname{Mat}(n \times n, \mathbb{R})$ invertierbar. Dann gilt

$$A \text{ positiv semi-definit} \quad \Longleftrightarrow \quad QAQ^{\mathrm{tr}} \text{ positiv semi-definit}$$

usw. für die restlichen Eigenschaften.

Beweis. Zu jedem $x \in \mathbb{R}^n$ gibt es genau ein $x^* \in \mathbb{R}^n$ mit $x = Q^{\mathrm{tr}} x^*$ und somit

$$x^{\mathrm{tr}} A x = (x^*)^{\mathrm{tr}} (QAQ^{\mathrm{tr}}) x^*. \qquad \square$$

(A 3.5) Untersuchung von Definitheits-Eigenschaften.
Man kann Zeilen- und Spaltenumformungen für diesen Zweck benutzen.
Zur Erinnerung: Eine elementare Zeilenumformung entspricht der Multiplikation mit einer Elementarmatrix C von links. Die korrespondierende Spaltenumformung entspricht der Multiplikation mit C^{tr} von rechts.

Ist eine symmetrische Matrix A gegeben, so führe an A Zeilenumformungen und in gleicher Reihenfolge entsprechende Spaltenumformungen durch, bis Diagonalgestalt erreicht ist. Letzterer ist Definitheit (Indefinitheit ...) einfach anzusehen.

Unter (A 3.8) finden Sie Beispiele.

(A 3.6) Satz. Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch und $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A mit Vielfachheiten gezählt. Bekanntlich sind alle Eigenwerte von A reell und A ist diagonalisierbar. Genau dann ist A

- positiv semi-definit, wenn alle $\lambda_j \geqslant 0, 1 \leqslant j \leqslant n$;
- positiv definit, wenn alle $\lambda_j > 0, 1 \leqslant j \leqslant n$;
- negativ semi-definit, wenn alle $\lambda_j \leqslant 0, 1 \leqslant j \leqslant n$;
- negativ definit, wenn alle $\lambda_j < 0, 1 \leqslant j \leqslant n$;
- indefinit, wenn es k, l gibt mit $\lambda_k < 0, \lambda_l > 0$.

(A 3.7) Satz. (HURWITZ-*Kriterium)*
Es sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ *symmetrisch und* A_k *entstehe aus* A *durch Streichen der letzten* $n-k$ *Zeilen und Spalten* $(1 \leqslant k \leqslant n-1)$, *weiter* $A_n := A$. *Dann gilt:*

$$A \ \text{positiv definit} \iff \det A_k > 0 \ \text{für} \ 1 \leqslant k \leqslant n.$$

Man beachte zusätzlich: A negativ definit $\iff -A$ positiv definit.

(A 3.8) Beispiele. In diesen Beispielen wird der Algorithmus aus (A 3.5) illustriert.

a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$

Simultane Zeilen- und Spaltenumformungen:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \left(\underset{\text{Zeile2}-2\cdot\text{Zeile1}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \right) \underset{\text{Spalte2}-2\cdot\text{Spalte1}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Nach Kriterium (A 3.5) ist A indefinit.

b) $A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & a \end{pmatrix}$; dabei $a \in \mathbb{R}$.

$$\begin{pmatrix} 1 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & a \end{pmatrix} \left(\underset{\substack{\text{Zeile2}+\text{Zeile1} \\ \text{Zeile3}-\text{Zeile1}}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & a-1 \end{pmatrix} \right) \underset{\substack{\text{Spalte2}+\text{Spalte1} \\ \text{Spalte3}-\text{Spalte1}}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & a-1 \end{pmatrix}$$

$$\left(\underset{\text{Zeile3}-2\cdot\text{Zeile2}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & a-5 \end{pmatrix} \right) \underset{\text{Spalte3}-2\cdot\text{Spalte2}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & a-5 \end{pmatrix}$$

Also A positiv definit für alle $a > 5$;
A positiv semi-definit für alle $a \geqslant 5$;
A indefinit für alle $a < 5$.

c) $A = \begin{pmatrix} 10 & 1 & 0 & 2 \\ 1 & -1 & 4 & 5 \\ 0 & 4 & 11 & 0 \\ 2 & 5 & 0 & 12 \end{pmatrix}$ ist indefinit, da in der Diagonale ein positives und ein negatives Element steht. (Explizit $e_1^{\mathrm{tr}} A e_1 = 10$, $e_2^{\mathrm{tr}} A e_2 = -1$.)

Durch Betrachten der Diagonalelemente lässt sich also oft schon Definitheit ausschließen.

d) $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix}$
ist sicher nicht negativ semi-definit (alle Diagonalelemente $\geqslant 0$) und sicher nicht positiv definit (da 0 in Diagonale).
Um zwischen positiver Semidefinitheit und Indefinitheit zu entscheiden, muss jedoch der GAUSS-Algorithmus angewandt werden. Der erste Schritt ist eine Vertauschung, um die 0 links oben loszuwerden.

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{pmatrix} \left(\underset{\text{Zeile2} \leftrightarrow \text{Zeile1}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix} \right) \underset{\text{Spalte2} \leftrightarrow \text{Spalte1}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

$$\left(\underset{\substack{\text{Zeile2}-\frac{1}{2}\cdot\text{Zeile1} \\ \text{Zeile3}-\frac{1}{2}\cdot\text{Zeile1}}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 2 & 1 & 1 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix} \right) \underset{\substack{\text{Spalte2}-\frac{1}{2}\cdot\text{Spalte1} \\ \text{Spalte3}-\frac{1}{2}\cdot\text{Spalte1}}}{\underset{\uparrow}{\longrightarrow}} \begin{pmatrix} 2 & 1 & 1 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix}$$

Hier sieht man aus den Diagonalelementen: A ist indefinit.

A. 4. Determinanten

(A 4.1) Definition. Die *Determinante* einer $n \times n$-Matrix $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ über einem Körper K kann wie folgt rekursiv definiert werden.

- Für $n = 1$ gilt $A = \det(a_{11}) = a_{11}$.
- Für $n > 1$ hat man die Rekursion

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \cdot \det A^{(1,j)},$$

wobei $A^{(1,j)}$ diejenige $(n-1) \times (n-1)$-Matrix ist, welche aus A durch Streichen von Zeile 1 und Spalte j entsteht.

(A 4.2) Beispiele.

$\triangleright$ $n = 2$: $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Big(= (-1)^{1+1} a_{11} \cdot \det(a_{22}) + (-1)^{1+2} a_{12} \cdot \det(a_{21}) \Big) = a_{11}a_{22} - a_{12}a_{21}$

$\triangleright$ $n = 3$: $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \cdot \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \cdot \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \cdot \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Wir geben keine komplette Liste von Eigenschaften an, sondern beschränken uns auf das für uns direkt Notwendige.

(A 4.3) Einige Eigenschaften. a) Die Definition unter (A 4.1) benutzt bei der Rekursion die Entwicklung nach der ersten Zeile. Man erhält das selbe Resultat, wenn man nach Zeile i entwickelt:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det A^{(i,j)}.$$

Und man erhält auch das selbe Resultat, wenn man nach Spalte j entwickelt:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \cdot \det A^{(i,j)}.$$

b) Multiplikationssatz: Für alle $A, B \in \mathrm{Mat}(n \times n; K)$ gilt

$$\det(A \cdot B) = \det A \cdot \det B.$$

c) Die Abbildung

$$\det : \mathrm{Mat}(n \times n; \mathbb{R}) \to \mathbb{R}, \quad X \mapsto \det X,$$

ist eine Polynomabbildung, also insbesondere stetig und stetig differenzierbar.

Anwendungen auf die Invertierbarkeit von Matrizen beschreiben wir in

(A 4.4) Determinanten testen Invertierbarkeit.

a) Für $A \in \mathrm{Mat}(n \times n; K)$ gilt: A ist invertierbar genau dann, wenn $\det A \neq 0$.

b) "Formel für die Inverse": Falls $\det A \neq 0$, so ist

$$A^{-1} = \frac{1}{\det A} \cdot \operatorname{adj} A,$$

wobei der Eintrag von $\operatorname{adj} A$ ("Adjunkte von A") an der Stelle (i, j) gleich $(-1)^{i+j} \cdot \det A^{(j,i)}$ ist.

c) Bekanntes und für Rechnungen nützliches Beispiel:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \cdot \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11}. \end{pmatrix}$$

d) Anwendung auf reelle Matrizen:
Im (irgendwie) normierten Vektorraum $\mathrm{Mat}(n \times n; \mathbb{R})$ ist die Menge $\mathrm{GL}\,(n; \mathbb{R})$ der invertierbaren Matrizen nicht-leer und auch offen, da det stetig und

$$\mathrm{GL}\,(n; \mathbb{R}) = \{X \in \mathrm{Mat}(n \times n; \mathbb{R});\ \det X \neq 0\}.$$

Weiter ist die Abbildung

$$\iota : \mathrm{GL}\,(n; \mathbb{R}) \to \mathrm{GL}\,(n; \mathbb{R}), \quad X \mapsto X^{-1},$$

eine rationale Abbildung, also insbesondere stetig und stetig differenzierbar.

Symbolverzeichnis

Index