Skript zur Vorlesung

Analysis 2 für Lehramtsstudierende (GS/HS/RS)

(Sommersemester 2006)

Dieses Geheft enthält in kompakter Form die wesentlichen Inhalte, wie sie in der Vorlesung ,,Analysis für Lehramtsstudierende (GS/HS/RS)" vorgestellt werden.

Es ist zum Gebrauch neben der Vorlesung gedacht und erhebt nicht den Anspruch, ,,in sich selbst verständlich" oder vollständig zu sein.

S. Hilger

Inhaltsverzeichnis

1 Vorspann: Die Tangens–Funktion

Definition Die Funktion

$$\tan \begin{cases} \mathbb{R} \setminus \left\{ (2k+1)\frac{\pi}{2} \,\middle|\, k \in \mathbb{Z} \right\} & \to \mathbb{R} \\ \alpha & \mapsto \tan(\alpha) = \frac{\sin \alpha}{\cos \alpha} \end{cases}$$

heißt *Tangens–Funktion.* Die gleiche Funktion

$$\tan \begin{cases}]-\frac{\pi}{2}, +\frac{\pi}{2}[& \to \mathbb{R} \\ \alpha & \mapsto \tan(\alpha) = \frac{\sin \alpha}{\cos \alpha} \end{cases}$$

mit der eingeschränkten Definitionsmenge $]-\frac{\pi}{2}, +\frac{\pi}{2}[$ nennen wir *eingeschränkte Tangens–Funktion.*

Satz 1 *Die eingeschränkte Tangens–Funktion ist stetig, streng monoton steigend und surjektiv. Sie besitzt also eine stetige, streng monoton steigende Umkehrfunktion*

$$\arctan \begin{cases} \mathbb{R} & \to]-\frac{\pi}{2}, +\frac{\pi}{2}[\\ x & \mapsto \arctan(x) \end{cases} .$$

Beweis Auf dem Intervall $[0, \frac{\pi}{2}[$ sind sowohl die Sinusfunktion als auch die Funktion $\frac{1}{\cos \alpha}$ positiv und streng monoton steigend, also ist auch ihr Produkt, die Tangens–Funktion, auf $[0, \frac{\pi}{2}[$ positiv und streng monoton steigend.
Wegen der Schiefsymmetrie (Graph punktsymmetrisch)

$$\tan(-\alpha) = \frac{\sin(-\alpha)}{\cos(-\alpha)} = \frac{-\sin \alpha}{\cos \alpha} = -\tan \alpha$$

ist die Tangens–Funktion auf dem Intervall $]-\frac{\pi}{2}, 0]$ negativ und ebenfalls streng monoton steigend. Daraus folgt, dass sie auf dem ganzen Intervall $]\frac{\pi}{2}, +\frac{\pi}{2}[$ streng monoton steigend ist.
Weiter existiert für jedes $M > 0$ ein $\alpha_0 \in [0, \frac{\pi}{2}[$, so dass für $\alpha > \alpha_0$:

$$\sin \alpha > \frac{1}{2} \quad \text{und} \quad \frac{1}{\cos \alpha} > 2M \quad \Longrightarrow \quad \tan \alpha > M.$$

Da die Tangens–Funktion streng monoton steigend und schiefsymmetrisch ist, folgt

$$\lim_{\alpha \nearrow +\frac{\pi}{2}} \tan \alpha = +\infty \qquad \text{und} \qquad \lim_{\alpha \searrow -\frac{\pi}{2}} \tan \alpha = -\infty.$$

Mit dem Zwischenwertsatz folgt, dass jede reelle Zahl als Bild angenommen wird. Das ist die Surjektivität. ♦
Überlege die Formeln:

$$\tan^2 \alpha = \frac{\sin^2 \alpha}{\cos^2 \alpha} = \frac{\sin^2 \alpha}{1 - \sin^2 \alpha} = \frac{1 - \cos^2 \alpha}{\cos^2 \alpha} = \frac{1}{\cos^2 \alpha} - 1.$$

2 Differenzierbare Funktionen

2.1 Definition

In diesem Kapitel sei D eine Teilmenge von $\mathbb{R}$ und die Zahl $a \in D$ habe die Eigenschaft, dass es mindestens eine Folge $(x_n) \subseteq D \setminus \{a\}$ gibt mit $\lim_{n\to\infty} x_n = a$.

Man denke dabei immer an solche a, für die es ein $\varepsilon > 0$ gibt, so dass $]a-\varepsilon, a+\varepsilon[\subseteq D$ ist. Wir benützen ab jetzt die abkürzende Schreibweise

$$\lim_{x \overset{\bullet}{\to} a} f(x) \;:=\; \lim_{x\to a, x\neq a} f(x)$$

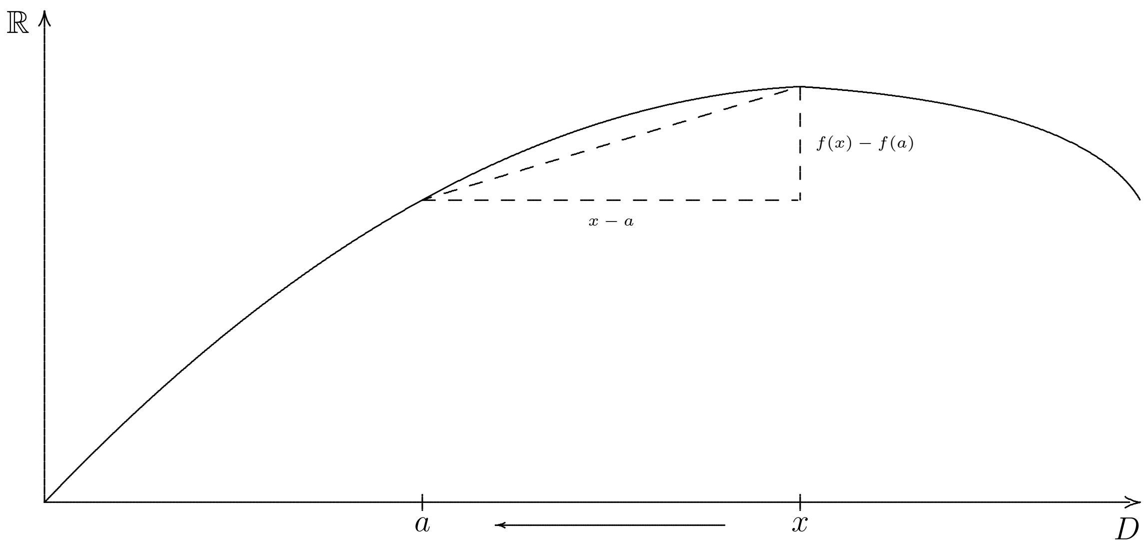
für Funktionsgrenzwerte, bei denen die zulässigen gegen a konvergierenden Folgen $(x_n) \subseteq D$ nicht die Grenzstelle a selbst enthalten dürfen.

Definition Die Funktion $f : D \to \mathbb{R}$ heißt *differenzierbar an der Stelle* a, (kürzer: in a), wenn der Grenzwert

$$\lim_{x \overset{\bullet}{\to} a} \frac{f(x)-f(a)}{x-a} \;=\; \lim_{h \overset{\bullet}{\to} 0} \frac{f(a+h)-f(a)}{h}$$

existiert. Der Grenzwert heißt in diesem Fall auch *Differentialquotient (Neue Rechtschreibung: Differenzialquotient)* oder *Ableitung der Funktion f an der Stelle a.*

Geometrisch kann die Ableitung als die Steigung der Tangente an den Graphen von f im Punkt $(a, f(a))$ interpretiert werden.



Der Quotient

$$\frac{\Delta f}{\Delta x} = \frac{f(x)-f(a)}{x-a}$$

gibt dabei die Steigung der Sekante an. Er heißt auch *Differenzenquotient.* (Beachte, dass die Symbole Δ keine eigenständige Bedeutung haben.)

Existieren die Grenzwerte

$$\lim_{x \nearrow a} \frac{f(x)-f(a)}{x-a} \qquad \text{bzw.} \qquad \lim_{x \searrow a} \frac{f(x)-f(a)}{x-a},$$

so spricht man von der *linksseitigen* bzw. *rechtsseitigen Ableitung* in a.

Überlege, dass die Ableitung an einer Stelle a genau dann existiert, wenn die beiden einseitigen Ableitungen existieren und gleich sind.

Definition Die Funktion heißt *differenzierbar in* D, wenn sie an jeder Stelle $a \in D$ differenzierbar ist. Ist dies der Fall, so kann die Funktion

$$f' = \frac{df}{dx} : \begin{cases} D & \to & \mathbb{R} \\ a & \mapsto & \lim\limits_{x \overset{\bullet}{\to} a} \dfrac{f(x) - f(a)}{x - a} \end{cases}$$

gebildet werden. Sie heißt *Ableitungsfunktion* oder einfach *Ableitung* der Funktion f. (Beachte, dass das zweite Symbol nur als ganzes Sinn macht.)

2.2 Beispiele

1. Ist $f : D \to \mathbb{R}$ eine konstante Funktion, $f(x) = c$ für alle $x \in D$, so gilt

   $$f'(a) = \lim_{x \overset{\bullet}{\to} a} \frac{f(x) - f(a)}{x - a} = \lim_{x \overset{\bullet}{\to} a} \frac{c - c}{x - a} = 0.$$

   Die Ableitung ist die Konstant–Null-Funktion.

2. Die Identität $f = \mathrm{id}_{\mathbb{R}}$ hat als Ableitung die Konstant–Eins–Funktion:

   $$f'(a) = \lim_{x \overset{\bullet}{\to} a} \frac{f(x) - f(a)}{x - a} = \lim_{x \overset{\bullet}{\to} a} \frac{x - a}{x - a} = 1.$$

3. Es sei $n \in \mathbb{N}_0$. Die Potenz–Funktion $f : x \mapsto x^n$ hat als Ableitung die Funktion

   $$f'(x) = n \cdot x^{n-1}.$$

   Dazu überlegen wir zunächst mit Hilfe des binomischen Lehrsatzes

   $$\begin{aligned} x^n - a^n &= [(x-a) + a]^n - a^n = \sum_{k=0}^{n} \binom{n}{k} (x-a)^k \cdot a^{n-k} - a^n \\ &= \sum_{k=1}^{n} \binom{n}{k} (x-a)^k \cdot a^{n-k} \end{aligned}$$

   und dann

   $$\frac{x^n - a^n}{x - a} = \sum_{k=1}^{n} \binom{n}{k} (x-a)^{k-1} \cdot a^{n-k}.$$

   Jetzt kann man den Differentialquotienten ausrechnen:

   $$\begin{aligned} \lim_{x \overset{\bullet}{\to} a} \frac{x^n - a^n}{x - a} &= \lim_{x \overset{\bullet}{\to} a} \sum_{k=1}^{n} \binom{n}{k} (x-a)^{k-1} \cdot a^{n-k} \\ &= \lim_{x \overset{\bullet}{\to} a} \binom{n}{1} (x-a)^{1-1} \cdot a^{n-1} = n \cdot a^{n-1}. \end{aligned}$$

4. Die Hyperbel–Funktion und ihre Ableitung sind gegeben durch

$$f : \begin{cases} \mathbb{R}^* & \to & \mathbb{R} \\ x & \mapsto & \frac{1}{x} \end{cases} \qquad f' : \begin{cases} \mathbb{R}^* & \to & \mathbb{R} \\ x & \mapsto & -\frac{1}{x^2} \end{cases} .$$

Zur Begründung:

$$\lim_{h \overset{\bullet}{\to} 0} \frac{f(a+h)-f(a)}{h} = \lim_{h \overset{\bullet}{\to} 0} \frac{\frac{1}{a+h} - \frac{1}{a}}{h} =$$
$$\lim_{h \overset{\bullet}{\to} 0} \frac{a-(a+h)}{h(a+h)a} = \lim_{h \overset{\bullet}{\to} 0} \frac{-1}{(a+h)a} = -\frac{1}{a^2}.$$

5. Wir betrachten die Betragsfunktion $f(x) = |x|$. Zu jedem $x \neq 0$ gibt es eine ε–Umgebung, so dass die Betragsfunktion mit der Funktion $x \mapsto x$ oder $x \mapsto -x$ übereinstimmt. Deshalb ist sie an diesen Stellen differenzierbar mit

$$f'(x) = \operatorname{sgn}(x).$$

An der Stelle $a = 0$ gilt

$$\lim_{x \searrow a} \frac{|x|-|a|}{x-a} = \lim_{x \searrow a} \frac{x-a}{x-a} = +1$$
$$\lim_{x \nearrow a} \frac{|x|-|a|}{x-a} = \lim_{x \nearrow a} \frac{(-x)-(-a)}{x-a} = -1.$$

Es existieren die einseitigen Ableitungen. Sie sind aber verschieden. Die Betragsfunktion ist nicht in $a = 0$ differenzierbar.

6. Die Funktion exp hat sich selbst als Ableitungsfunktion. Tatsächlich ist aufgrund der Restgliedabschätzung Lemma 46 (Analysis 1) zunächst

$$\Big|\exp(x) - (1+x)\Big| \le |x|^2 \qquad \text{für } |x| \le \frac{3}{2}$$

woraus folgt

$$\Big|\frac{\exp(x)-1}{x} - 1\Big| \le |x| \qquad \text{für } 0 < |x| \le \frac{3}{2}$$

und deshalb

$$\lim_{x \overset{\bullet}{\to} 0} \frac{\exp(x)-1}{x} = 1.$$

Dann geht es so weiter:

$$\lim_{x \overset{\bullet}{\to} a} \frac{\exp(x)-\exp(a)}{x-a} = \lim_{x \overset{\bullet}{\to} a} \frac{\exp(x-a)-1}{x-a} \cdot \exp(a)$$
$$= \lim_{h \overset{\bullet}{\to} 0} \frac{\exp(h)-1}{h} \cdot \exp(a) = \exp(a)$$

7. Um die Ableitung des Sinus zu berechnen, überlegen wir zunächst mit Hilfe der Additionstheoreme (Satz 51, ANA01):

$$\frac{\sin(x+h)-\sin(x)}{h} = \frac{\sin(x+\frac{h}{2}+\frac{h}{2})-\sin(x+\frac{h}{2}-\frac{h}{2})}{h} =$$
$$\frac{2\cos(x+\frac{h}{2})\sin(\frac{h}{2})}{h} = \frac{\cos(x+\frac{h}{2})\sin(\frac{h}{2})}{\frac{h}{2}} \xrightarrow{h \overset{\bullet}{\to} 0} \cos(x),$$

Der Grenzübergang $h \overset{\bullet}{\to} 0$ erfolgt mit Hilfe von $\lim_{h \overset{\bullet}{\to} 0} \frac{\sin(\frac{h}{2})}{\frac{h}{2}} = 1$ (Folgerung 53, ANA01) und der Stetigkeit des Cosinus. Es gilt also

$$\sin'(x) = \cos(x).$$

8. Mit Hilfe des Zusammenhangs $\cos(x) = \sin(x+\frac{\pi}{2})$ oder analog zu gerade eben kann man ausrechnen, dass

$$\cos'(x) = -\sin(x).$$

Satz 2 (Alternativdefinition der Differenzierbarkeit) *Es sei $f : D \to \mathbb{R}$ ein reelle Funktion und $a \in D$.*

(i) f ist an einer Stelle a genau dann differenzierbar mit Ableitung $f'(a)$, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|f(x) - f(a) - f'(a)\cdot(x-a)| \le \varepsilon \cdot |x-a|, \qquad \textit{falls } |x-a| \le \delta.$$

(ii) f ist an einer Stelle a genau dann differenzierbar, wenn eine Konstante $\ell \in \mathbb{R}$ existiert, so dass für alle $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|f(x) - f(a) - \ell\cdot(x-a)| \le \varepsilon \cdot |x-a|, \qquad \textit{falls } |x-a| \le \delta.$$

Bemerkung Die beiden Alternativdefinitionen unterscheiden sich dadurch, dass bei (i) die Ableitung bereits vorgegeben ist, bei (ii) nur die Existenz (irgend)einer Konstante verlangt wird. Natürlich ist — im Nachhinein — diese Konstante eindeutig die Ableitung. Die Alternativdefinitionen muten zunächst sehr abstrakt und umständlich an. Ihr Vorteil besteht aber darin, dass das ,,Dividieren" vermieden wird. Das ist für die ,,Abschätzerei der Differentialrechnung" von Vorteil. Außerdem sind sie der Ansatzpunkt für eine Verallgemeinerung der Differentialrechnung auf mehrdimensionale Funktionen.

Beweis Die Aussage (ii) ergibt sich durch die folgenden Äquivalenzen:

- f ist in a differenzierbar.

$\Longleftrightarrow$ Es existiert der Grenzwert

$$\ell := \lim_{x \overset{\bullet}{\to} a} \frac{f(x)-f(a)}{x-a}.$$

$\Longleftrightarrow$ Es existiert ein $\ell \in \mathbb{R}$, so dass es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit

$$\left|\frac{f(x) - f(a)}{x - a} - \ell\right| \leq \varepsilon, \text{ falls } 0 < |x - a| \leq \delta.$$

$\Longleftrightarrow$ Es existiert ein $\ell \in \mathbb{R}$, so dass es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt mit

$$|f(x) - f(a) - \ell \cdot (x - a)| \leq \varepsilon \cdot |x - a|, \text{ falls } |x - a| \leq \delta.$$

Die Aussage (i) kann man aus diesen Überlegungen ebenfalls ablesen.

♦

Satz 3 (Differenzierbarkeit $\Longrightarrow$ Stetigkeit)

(i) Ist f an der Stelle a differenzierbar, so gibt es ein $\delta > 0$ und eine Konstante $L > 0$, so dass

$$|f(x) - f(a)| \leq L \cdot |x - a|, \qquad \textit{falls } |x - a| \leq \delta.$$

(ii) Ist die Funktion f an der Stelle a differenzierbar, so ist sie an dieser Stelle auch stetig.

Beweis (i) Zu der Zahl 1 gibt es ein $\delta > 0$, so dass

$$|f(x) - f(a) - f'(a) \cdot (x - a)| \leq 1 \cdot |x - a|, \qquad \text{falls } |x - a| \leq \delta.$$

Für diese x gilt dann weiter

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f(a) - f'(a) \cdot (x - a)| + |f'(a) \cdot (x - a)| \\ &\leq 1 \cdot |x - a| + |f'(a)| \cdot |x - a| = \underbrace{(1 + |f'(a)|)}_{=:L} \cdot |x - a|. \end{aligned}$$

(ii) Ist $\varepsilon > 0$ vorgegeben, so wählt man $\delta_1 = \min\{\delta, \frac{\varepsilon}{L+1}\}$ mit δ und L gemäß Teil (i) des Satzes. Dann gilt für alle x mit $|x - a| < \delta_1$

$$|f(x) - f(a)| \leq L \cdot |x - a| \leq \frac{L \cdot \varepsilon}{L + 1} < \varepsilon.$$

♦

2.3 Differentiation und „Grundrechenarten"

Satz 4 *Es seien $f, g : D \to \mathbb{R}$ zwei in a differenzierbare Funktionen.*

(i) Für $\alpha, \beta \in \mathbb{R}$ ist auch die Linearkombination $\alpha f + \beta g$ differenzierbar in a und es gilt:

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

(ii) Das Produkt der beiden Funktionen ist differenzierbar in a und es gilt:

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

(iii) Ist $g(x) \neq 0$ für $x \in D$, so ist die Funktion $\frac{f}{g}$ definiert und differenzierbar in a. Dabei gilt:

$$\Big(\frac{f}{g}\Big)'(a) = \frac{f'(a) \cdot g(a) - f(a) \cdot g'(a)}{g(a)^2}.$$

Beweis Die erste Behauptung ergibt sich sofort aus den entsprechenden Regeln für Grenzwerte (Vgl. Satz 20 Analysis I).

(ii) Für den Beweis der zweiten Behauptung formen wir den zugehörigen Differenzenquotienten um:

$$\frac{f(x)g(x) - f(a)g(a)}{x - a} = f(x) \cdot \frac{g(x) - g(a)}{x - a} + \frac{f(x) - f(a)}{x - a} \cdot g(a)$$

Da f und g in a differenzierbar sind und f stetig in a ist, existiert der Grenzwert $\lim_{x \overset{\bullet}{\to} a}$ der rechten Seite. Deswegen existiert auch der Grenzwert der linken Seite, er hat den im Satz angegebenen Wert.

(iii) Es ist

$$\begin{aligned} \frac{\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}}{x - a} &= \frac{1}{g(x)g(a)} \cdot \frac{f(x)g(a) - f(a)g(x)}{x - a} \\ &= \frac{1}{g(x)g(a)} \cdot \Big[\frac{f(x) - f(a)}{x - a} g(a) - f(a) \frac{g(x) - g(a)}{x - a}\Big]. \end{aligned}$$

Der Grenzwert $\lim_{x \overset{\bullet}{\to} a}$ der rechten Seite existiert und ist wie im Satz angegeben. Deshalb gilt dies auch für die linke Seite. ♦

Beispiele Übung: Beweise mit Hilfe der im Satz angegebenen Regeln, dass die Funktion $x \mapsto x^n$ (geeignete Definitionsmenge) die folgende Ableitung hat:

$$(x^n)' = nx^{n-1}.$$

Die Tangens–Funktion hat die Ableitung

$$\tan' \alpha = \frac{\sin' \alpha \cos \alpha - \sin \alpha \cos' \alpha}{\cos^2 \alpha} = \frac{\cos^2 \alpha + \sin^2 \alpha}{\cos^2 \alpha} = \frac{1}{\cos^2 \alpha}.$$

2.4 Differentiation und Umkehrfunktion

Satz 5 (Differentiation und Umkehrfunktion) *Es seien D und E Intervalle und $f : D \to E$ streng monoton steigend (fallend) und stetig. Gemäß Satz 42 (ANA01) existiert die Umkehrfunktion $g : E \to D$, die ebenfalls streng monoton steigend (fallend) und stetig ist. Ist f in $a \in D$ differenzierbar mit $f'(a) \neq 0$, so ist g im Bildpunkt $b = f(a)$ differenzierbar und es gilt:*

$$f'(a) \cdot g'(b) = 1.$$

Beweis (i) Es sei y_n eine beliebige Folge in $E \setminus \{b\}$ mit $\lim_{n\to\infty} y_n = b$. Mit $x_n := g(y_n)$ gilt aufgrund der Stetigkeit und Bijektivität von g, dass

$$\lim_{n\to\infty} x_n = a, \qquad x_n \neq a.$$

Dann ist

$$\frac{g(y_n) - g(b)}{y_n - b} = \frac{x_n - a}{f(x_n) - f(a)} = \frac{1}{\frac{f(x_n)-f(a)}{x_n-a}}.$$

Der Grenzwert $\lim_{n\to\infty}$ auf der rechten Seite existiert aufgrund der Voraussetzungen des Satzes. Also existiert auch der Grenzwert $\lim_{n\to\infty}$ auf der linken Seite. Da die Folgen y_n beliebig waren, gilt

$$\begin{aligned} g'(b) &= \lim_{y \overset{\bullet}{\to} b} \frac{g(y) - g(b)}{y - b} = \lim_{n\to\infty} \frac{g(y_n) - g(b)}{y_n - b} = \\ &= \lim_{n\to\infty} \frac{1}{\frac{f(x_n)-f(a)}{x_n-a}} = \lim_{x \overset{\bullet}{\to} a} \frac{1}{\frac{f(x)-f(a)}{x-a}} = \frac{1}{f'(a)}. \end{aligned}$$

♦

Beispiele

1. Da die Exponentialfunktion auf $\mathbb{R}$ differenzierbar ist, ist auch ihre Umkehrfunktion, der natürliche Logarithmus, auf seiner Definitionsmenge $\mathbb{R}^+$ differenzierbar. Für ein $b \in \mathbb{R}^+$ und $a = \ln(b)$ gilt

$$\exp'(a) \cdot \ln'(b) = 1$$

und daher

$$\ln'(b) = \frac{1}{\exp'(a)} = \frac{1}{\exp(a)} = \frac{1}{b}.$$

Weiter kann man daraus ableiten, dass

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

Es ist nämlich

$$(1 + \frac{1}{n})^n = \exp\left[\ln[(1+\frac{1}{n})^n]\right] = \exp\left[n \cdot \ln(1+\frac{1}{n})\right] = \exp\left[\frac{\ln(1+\frac{1}{n})}{\frac{1}{n}}\right]$$

Wendet man darauf $\lim_{n\to\infty}$ an, so folgt die Behauptung wegen der Stetigkeit von exp $(*)$ und $\ln'(1) = 1$ $(**)$ so:

$$\lim_{n\to\infty} (1 + \frac{1}{n})^n = \lim_{n\to\infty} \exp\left[\frac{\ln(1+\frac{1}{n})}{\frac{1}{n}}\right] \overset{(*)}{=} \exp\left[\lim_{n\to\infty} \frac{\ln(1+\frac{1}{n})}{\frac{1}{n}}\right] \overset{(**)}{=} \exp(1) = e.$$

2. Die eingeschränkte Sinus–Funktion

$$\sin \begin{cases} [-\frac{\pi}{2}, +\frac{\pi}{2}] & \to & [-1,+1] \\ \alpha & \mapsto & \sin(\alpha) \end{cases}$$

ist differenzierbar und streng monoton steigend, da für $\alpha, \beta \in [-\frac{\pi}{2}, +\frac{\pi}{2}]$ mit $\alpha > \beta$

$$\begin{aligned} &\sin\alpha - \sin\beta = \sin(\frac{\alpha+\beta}{2} + \frac{\alpha-\beta}{2}) - \sin(\frac{\alpha+\beta}{2} - \frac{\alpha-\beta}{2}) = \\ &\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} + \cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} \\ &\qquad - \sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} + \cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} = \\ &\underbrace{2\cos\overbrace{\frac{\alpha+\beta}{2}}^{\in]-\frac{\pi}{2},+\frac{\pi}{2}[}}_{>0}\ \underbrace{\sin\overbrace{\frac{\alpha-\beta}{2}}^{\in]0,+\pi[}}_{>0}. \end{aligned}$$

Es existiert also eine differenzierbare Umkehrfunktion

$$\arcsin : \begin{cases} [-1,+1] & \to & [-\frac{\pi}{2},+\frac{\pi}{2}] \\ x & \mapsto & \arcsin(x) \end{cases}.$$

Ist $b \in [-1,+1]$, so gibt es ein $\alpha \in [-\frac{\pi}{2},+\frac{\pi}{2}]$ mit $\sin\alpha = b$.

$$\arcsin'(b) \cdot \sin'(\alpha) = \arcsin'(b) \cdot \cos(\alpha) = 1.$$

Da $\cos\alpha \geq 0$ für $\alpha \in [-\frac{\pi}{2},+\frac{\pi}{2}]$, erhalten wir $\cos\alpha = \sqrt{1-\sin^2\alpha} = \sqrt{1-b^2}$, insgesamt also

$$\arcsin'(b) = \frac{1}{\sqrt{1-b^2}}.$$

3. Für die Ableitung der Arcus–Tangens–Funktion

$$\arctan : \begin{cases} \mathbb{R} & \to &]-\frac{\pi}{2},+\frac{\pi}{2}[\\ x & \mapsto & \arctan(x) \end{cases}$$

überlegen wir: Ist $b \in \mathbb{R}$, so gibt es ein $\alpha \in]-\frac{\pi}{2},+\frac{\pi}{2}[$ mit $\tan\alpha = b$, es folgt

$$\begin{aligned} 1 &= \arctan'(b) \cdot \tan'(\alpha) = \arctan'(b) \cdot \frac{1}{\cos^2\alpha} = \\ &= \arctan'(b) \cdot \left(1+\tan^2\alpha\right) = \arctan'(b) \cdot \left(1+b^2\right), \end{aligned}$$

also

$$\arctan'(b) = \frac{1}{1+b^2}.$$

2.5 Differentiation und Verknüpfung von Funktionen

Satz 6 (Kettenregel) *Es seien $f : D \to \mathbb{R}$ und $g : E \to \mathbb{R}$ reelle Funktionen mit $f(D) \subseteq E$. Ist f in $a \in D$ differenzierbar und g in $b = f(a)$ differenzierbar, so ist auch die Funktion $g \circ f : D \to \mathbb{R}$ in a differenzierbar und es gilt:*

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

Man spricht auch vom Nachdifferenzieren der Funktion g bzgl. des Arguments $f(a)$.

Beweis Es sei x_n eine beliebige Folge in $D \setminus \{a\}$ mit $\lim_{n\to\infty} x_n = a$. Die Folge $y_n := f(x_n)$ konvergiert aufgrund der Stetigkeit von f in a gegen $b = f(a)$.

Wäre $y_n \neq b$ für alle n, (beispielsweise wenn f bijektiv) so könnten wir umschreiben

$$\frac{g(f(x_n)) - g(f(a))}{x_n - a} = \frac{g(y_n) - g(b)}{y_n - b} \cdot \frac{f(x_n) - f(a)}{x_n - a}.$$

und der Grenzübergang $\lim_{n\to\infty}$ würde sofort die Behauptung liefern.

Da wir aber das nicht voraussetzen können, benutzen wir die Alternativdefinition der Differenzierbarkeit.

Zunächst zerlegen wir den Ausdruck in der Alternativdefinition geeignet:

$$\begin{aligned}&\Big|g(f(x)) - g(f(a)) - g'(b)f'(a)(x-a)\Big| = \\ &\Big|g(y) - g(b) - g'(b)(y-b) + g'(b)(f(x)-f(a)) - g'(b)f'(a)(x-a)\Big| \le \\ &\Big|g(y) - g(b) - g'(b)(y-b)\Big| + |g'(b)|\Big|f(x) - f(a) - f'(a)(x-a)\Big| \overset{\text{Ziel}}{<} \varepsilon|x-a|.\end{aligned}$$

Unser Ziel ist es jetzt, zu einem vorgegebenem $\varepsilon > 0$ ein $\delta > 0$ zu finden, so dass die letzte Ungleichung für $|x-a| < \delta$ erfüllt ist.

Dazu gehen wir der Reihe nach so vor:

1. Es existiert (vgl. Satz 3) ein δ_1 und ein $L \ge 0$, so dass
$$|y-b| = |f(x) - f(a)| \le L \cdot |x-a|, \qquad \text{falls } |x-a| < \delta_1.$$
2. Es existiert (vgl. Satz 2) ein $\delta_2 > 0$, so dass
$$\Big|f(x) - f(a) - f'(a)(x-a)\Big| < \frac{\varepsilon}{2\cdot(|g'(b)|+1)}|x-a|, \qquad \text{falls } |x-a| < \delta_2.$$
3. Es existiert (vgl. Satz 2) ein $\delta_3 > 0$, so dass
$$\Big|g(y) - g(b) - g'(b)(y-b)\Big| < \frac{\varepsilon}{2\cdot(L+1)}|y-b|, \qquad \text{falls } |y-b| < \delta_3.$$

Setzen wir jetzt $\delta := \min\{\delta_1, \delta_2, \frac{\delta_3}{L+1}\}$, so gilt für $x \in D$ mit $|x-a| < \delta$ zunächst

$$|y-b| = |f(x) - f(a)| \le L \cdot |x-a| \le L \cdot \frac{\delta_3}{L+1} < \delta_3$$

und dann

$$\underbrace{\underbrace{\Big|g(y) - g(b) - g'(b)(y-b)\Big|}_{\le \frac{\varepsilon}{2\cdot(L+1)}|y-b|}}_{< \frac{\varepsilon}{2}|x-a|} + \underbrace{|g'(b)|\,\underbrace{\Big|f(x) - f(a) - f'(a)(x-a)\Big|}_{\le \frac{\varepsilon}{2\cdot(|g'(b)|+1)}|x-a|}}_{< \frac{\varepsilon}{2}|x-a|} < \varepsilon|x-a|.$$

♦

Beispiele

1. Ableitung und „Shift" einer Funktion sind vertauschbar: Der h–Shift f_h einer Funktion $f : \mathbb{R} \to \mathbb{R}$ ist definiert durch $f_h(x) := f(x+h)$. Beim Übergang von f nach f_h wird der Graph horizontal um h nach links verschoben. Ist f differenzierbar, so ist auch f_h differenzierbar, es gilt:
$$f_h'(x) = f'(x+h).$$

2. Wird das Argument einer Funktion mit dem Faktor $a \in \mathbb{R}$ gestreckt: $g(x) = f(ax)$, so gilt für die Ableitung:

$$g'(x) = a \cdot f'(ax).$$

3. Die Ableitung der Funktion x^b, $b \in \mathbb{R}$ fest, mit $D = \mathbb{R}^+$, ergibt sich wegen der Definition

$$x^b := \exp(b \cdot \ln x)$$

zu

$$\begin{aligned}(x^b)' &= [\exp(b \cdot \ln x)]' = [\exp'(b \cdot \ln x)] \cdot [b \cdot \ln(x)]' = \\ &= \exp(b \cdot \ln x) \cdot \frac{b}{x} = x^b \cdot \frac{b}{x} = b \cdot x^{b-1}.\end{aligned}$$

2.6 Differentiation und Abschätzung, der Mittelwertsatz

Definition Es sei $f : D \to \mathbb{R}$ eine reelle Funktion und $a \in D$.
(1) Der Punkt $(a, f(a))$ des Graphen heißt wie in der ersten Spalte angegeben, wenn es ein $\varepsilon > 0$ gibt, so dass die Bedingung der zweiten Spalte gilt:

lokales Maximum	$f(x) \leq f(a)$ für alle $x \in D$ mit $\|x-a\| < \varepsilon$
lokales Minimum	$f(x) \geq f(a)$ für alle $x \in D$ mit $\|x-a\| < \varepsilon$
strenges lokales Maximum	$f(x) < f(a)$ für alle $x \in D$ mit $0 < \|x-a\| < \varepsilon$
strenges lokales Minimum	$f(x) > f(a)$ für alle $x \in D$ mit $0 < \|x-a\| < \varepsilon$

(2) Der Punkt $(a, f(a))$ des Graphen heißt wie in der ersten Spalte angegeben, wenn die Bedingung der zweiten Spalte erfüllt ist:

globales Maximum	$f(x) \leq f(a)$ für alle $x \in D$
globales Minimum	$f(x) \geq f(a)$ für alle $x \in D$
strenges globales Maximum	$f(x) < f(a)$ für alle $x \in D$
strenges globales Minimum	$f(x) > f(a)$ für alle $x \in D$

(3) Man spricht jeweils von einem Extremum, wenn es sich um ein Maximum oder Minimum handelt.
(4) In diesem Zusammenhang heißt a auch Stelle des Maximums, Minimums oder Extremums und $f(a)$ der Wert des Maximums, Minimums oder Extremums.

Satz 7 *(i) Ist die Funktion $f : D \to \mathbb{R}$ an der Stelle a differenzierbar und hat sie dort ein lokales Extremum, so gilt $f'(a) = 0$.*

(ii) Die Umkehrung gilt nicht, wie das Beispiel $f(x) = x^3$ und $a = 0$ zeigt.

Beweis O.B.d.A. habe f in a ein lokales Minimum. Es gibt dann ein $\varepsilon > 0$, so dass für alle $x \in D$ mit $|x - a| < \varepsilon$ gilt: $f(x) - f(a) \geq 0$. Es folgt, dass

$$\frac{f(x)-f(a)}{x-a} \leq 0 \text{ für alle } x \in D \cap]a-\varepsilon, a[,$$
$$\frac{f(x)-f(a)}{x-a} \geq 0 \text{ für alle } x \in D \cap]a, a+\varepsilon[.$$

Da f in a differenzierbar ist, kann man in diesen Ungleichungen zum Grenzwert übergehen (vgl. Satz 20(v) ANA01):

$$f'(a) = \lim_{x \overset{\bullet}{\to} a} \frac{f(x)-f(a)}{x-a} = \lim_{x \nearrow a} \frac{f(x)-f(a)}{x-a} \leq 0,$$
$$f'(a) = \lim_{x \overset{\bullet}{\to} a} \frac{f(x)-f(a)}{x-a} = \lim_{x \searrow a} \frac{f(x)-f(a)}{x-a} \geq 0.$$

Es muss also $f'(a) = 0$ sein. ♦

Satz 8 (Satz von Rolle) *Es sei $f : [c, d] \to \mathbb{R}$ eine stetige, in $]c, d[$ differenzierbare Funktion. Gilt $f(c) = f(d)$, so gibt es ein $\xi \in]c, d[$ mit $f'(\xi) = 0$.*

Bemerkung: Beachte, dass nicht die Stetigkeit der Ableitung f' auf $]c, d[$ gefordert ist.

Beweis O.B.d.A. ist die Funktion nicht konstant. Nach Satz 41(ii)/ANA01 nimmt die Funktion ihr globales Maximum oder Minimum in einem $\xi \in]c, d[$ an. Nach Satz 7 gilt $f'(\xi) = 0$. ♦

Satz 9 („Umkehrung" des Satzes von Rolle) *Es sei $f : [c, d] \to \mathbb{R}$ eine stetige, in $]c, d[$ differenzierbare Funktion. Ist $f'(x) \neq 0$ für alle $x \in]c, d[$, so ist f streng monoton steigend oder streng monoton fallend.*

Beweis Wenn f nicht streng monoton wäre, so gäbe es in $[c, d]$ drei Stellen $x_1 < x_2 < x_3$ mit

$$\begin{aligned} &f(x_2) > \max\{f(x_1), f(x_3)\} \qquad \text{oder} \\ &f(x_2) < \min\{f(x_1), f(x_3)\}. \end{aligned}$$

In jedem Fall nimmt die Funktion f in $]c, d[$ ein Maximum oder Minimum an, was eine Nullstelle von f' an dieser Stelle nach sich zieht. Widerspruch. ♦

Satz 10 (Mittelwertsatz) *Die Funktion $f : [c, d] \to \mathbb{R}$ sei stetig in $[c, d]$ und differenzierbar in $]c, d[$. Dann gibt es eine Stelle $\xi \in]c, d[$, so dass*

$$f'(\xi) = \frac{f(d) - f(c)}{d - c}.$$

Zeichnung.

Beweis Es sei g die lineare Funktion, deren Graph (Gerade) durch die beiden Punkte $(c, f(c))$ und $(d, f(d))$ geht:

$$g(x) = f(c) + \frac{f(d) - f(c)}{d - c} \cdot (x - c).$$

Die Differenzfunktion

$$h(x) = f(x) - g(x)$$

hat dann in c und in d Nullstellen, sie ist stetig in $[c, d]$ und differenzierbar in $]c, d[$. Nach dem Satz von Rolle gibt es $\xi \in]c, d[$ mit $h'(\xi) = 0$. Das aber bedeutet gerade

$$f'(\xi) - \frac{f(d) - f(c)}{d - c} = f'(\xi) - g'(\xi) = h'(\xi) = 0.$$

♦

Folgerung 11 (Differentiation und Abschätzung) *Es sei $f : [c, d] \to \mathbb{R}$ eine stetige und in $]c, d[$ differenzierbare Funktion.*

(i) (Geometrie–Sichtweise) Die Ableitung f' (Tangentensteigung) sei in $]c, d[$ durch zwei Konstanten M_1, M_2 beschränkt.

$$M_1 \leq f'(x) \leq M_2 \text{ für alle } x \in]c, d[.$$

Dann ist auch die Steigung jeder Sekante des Graphen von f in gleicher Weise beschränkt:

$$M_1 \leq \frac{f(x) - f(y)}{x - y} \leq M_2 \text{ für } x, y \in [c, d] \text{ mit } x \neq y.$$

Andere Sichtweise: Der Graph der Funktion verläuft zwischen den linearen Funktionen mit Steigung M_1 bzw. M_2:

$$f(y) + M_1 \cdot (x - y) \leq f(x) \leq f(y) + M_2 \cdot (x - y) \text{ für } x, y \in [c, d] \text{ mit } x > y.$$

Der Satz ist auch bei Vorliegen nur einer der beiden Abschätzungen — entsprechend — richtig.

(ii) (Abschätzungs–Sichtweise) Die Ableitung f' sei in $]c, d[$ betragsmäßig durch die Konstante M beschränkt.

$$|f'(x)| \leq M \text{ für alle } x \in]c, d[.$$

Dann gilt:

$$|f(x) - f(y)| \leq M \cdot |x - y| \text{ für } x, y \in [c, d].$$

Beweis (i) Wende für fest ausgesuchte x, y den Mittelwertsatz auf die Funktion $f : [x, y] \to \mathbb{R}$ an. Der Mittelwert

$$f'(\xi) = \frac{f(y) - f(x)}{y - x}$$

ist durch M_1, M_2 beschränkt.
(ii) Aus der Voraussetzung folgt

$$-M \leq f'(x) \leq +M \text{ für alle } x \in]c, d[$$

und daraus mit (i) für $x, y \in [c, d]$, $x \neq y$

$$-M \leq \frac{f(x) - f(y)}{x - y} \leq +M$$

Das ist aber gleichbedeutend mit

$$\left| \frac{f(x) - f(y)}{x - y} \right| \leq M$$

und weiter mit

$$|f(x) - f(y)| \leq M\, |x - y|$$

Auch für $y = x$ stimmt die Aussage des Satzes. ♦

Folgerung 12 *Es seien* $f, g : [c, d] \to \mathbb{R}$ *stetig und in* $]c, d[$ *differenzierbar. Gilt*

$$f(c) = g(c) \text{ und } f'(x) \leq g'(x) \text{ für alle } x \in]c, d[,$$

so ist

$$f(x) \leq g(x) \text{ für alle } x \in [c, d].$$

Beweis Aus

$$0 \leq g'(x) - f'(x) = (g - f)'(x) \text{ für alle } x \in]c, d[$$

folgt mit der Folgerung 11 ($M_1 = 0$) für $x \in]c, d]$

$$0 \leq \frac{[g(x) - f(x)] - [g(c) - f(c)]}{x - c} = \frac{g(x) - f(x)}{x - c}$$

und daraus wegen $x - c > 0$ dann

$$g(x) - f(x) \geq 0.$$

♦

Folgerung 13 *Es sei* $f : [c, d] \to \mathbb{R}$ *stetig und in* $]c, d[$ *differenzierbar. Gilt*

$$f'(x) = 0 \text{ für alle } x \in]c, d[,$$

so ist f *auf* $[c, d]$ *konstant.*

Zum Beweis wende man die Folgerung 11 (ii) mit $M = 0$ an.

2.7 Differentiation und Monotonie

Satz 14 *Es sei wieder $f : [c, d] \to \mathbb{R}$ stetig und in $]c, d[$ differenzierbar. Es gelten die folgenden Äquivalenzen bzw. Implikationen:*

$$f'(x) \geq 0 \text{ für alle } x \in]c,d[\quad \Longleftrightarrow \quad f \text{ ist monoton wachsend in } [c,d]$$

$$f'(x) \leq 0 \text{ für alle } x \in]c,d[\quad \Longleftrightarrow \quad f \text{ ist monoton fallend in } [c,d]$$

$$f'(x) > 0 \text{ für alle } x \in]c,d[\quad \Longrightarrow \quad f \text{ ist streng monoton wachsend in } [c,d]$$

$$f'(x) < 0 \text{ für alle } x \in]c,d[\quad \Longrightarrow \quad f \text{ ist streng monoton fallend in } [c,d]$$

Beweis Wir zeigen zunächst die Implikation $\Longrightarrow$ der ersten Zeile und nehmen dazu an, dass f nicht monoton wachsend ist. Das heißt, es existieren $x, y \in [c, d]$ mit $x < y$ und

$$f(x) > f(y).$$

Daraus folgt aber für die nach dem Mittelwertsatz existierende Stelle ξ

$$f'(\xi) = \frac{f(y) - f(x)}{y - x} < 0,$$

was ein Widerspruch zur Voraussetzung ist. Die $\Longrightarrow$ Implikationen der anderen Zeilen zeigt man völlig analog.
Es bleibt also noch die Umkehrung $\Longleftarrow$ der ersten Zeile (zweite analog) zu zeigen: Ist f monoton wachsend, so gilt für einen beliebigen Differenzenquotienten

$$\frac{f(x) - f(a)}{x - a} \geq 0, \quad x, a \in [c, d], x \neq a.$$

Für festes $a \in]c, d[$ kann der Grenzübergang $\lim_{x \overset{\bullet}{\to} a}$ durchgeführt werden. Es folgt mit Satz 20(v) (ANA01):

$$f'(a) = \lim_{x \overset{\bullet}{\to} a} \frac{f(x) - f(a)}{x - a} \geq 0.$$

♦

Beachte, dass der Satz 20(v) aus ANA01 nicht bei einer strengen Ungleichung angewandt werden kann. Tatsächlich zeigt dass Beispiel der Funktion $f(x) = x^3$, dass der Links–Implikationspfeil in der dritten (analog: vierten) Zeile nicht gesetzt werden kann.

2.8 Die Regeln von l'Hospital

Satz 15 (Regel von l'Hospital für Grenzwerte im Endlichen)
Es seien $f, g :]c, a[\to \mathbb{R}$ zwei in $]c, a[$ differenzierbare Funktionen. g habe keine Nullstelle. Es sei die Voraussetzung

$$\lim_{x \nearrow a} f(x) = 0 \qquad \lim_{x \nearrow a} g(x) = 0.$$

erfüllt.
Wenn der Grenzwert auf der rechten Seite existiert, dann gilt

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

Bemerkung Analoges gilt für die folgenden Situationen:

- Funktionen $f, g :]a, d[\to \mathbb{R}$ und rechtsseitige Grenzwerte $\lim\limits_{x \searrow a}$.
- Funktionen $f, g :]c, d[\setminus \{a\} \to \mathbb{R}$, wobei $c < a < d$ und Grenzwerte $\lim\limits_{x \dot{\to} a}$.

Beweis Wir wollen **zunächst** die ZUSÄTZLICHE VORAUSSETZUNG

$$\lim_{x \nearrow a} f'(x) \text{ existiert} \qquad \lim_{x \nearrow a} g'(x) \text{ existiert.}$$

verwenden, wodurch der Beweis durchsichtig und schnell wird. Ein aufwändigerer Beweis ohne diese Voraussetzung ist im nächsten Unterkapitel aufgeschrieben.

Nach der Festsetzung $f(a) = g(a) = 0$ sind die Funktionen $f, g :]c, a] \to \mathbb{R}$ stetig und in $]c, a[$ differenzierbar.
Gemäß Mittelwertsatz (Satz 10) existieren für jedes $x \in]c, a[$ zwei Zahlen $\xi(x), \eta(x) \in]x, a[$, so dass

$$\frac{f(x) - f(a)}{x - a} = f'(\xi(x)) \qquad \text{und} \qquad \frac{g(x) - g(a)}{x - a} = g'(\eta(x)).$$

Dann gilt

$$\frac{f(x)}{g(x)} = \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} = \frac{f'(\xi(x))}{g'(\eta(x))}. \qquad (*)$$

Wegen $\xi(x), \eta(x) \in]x, a[$ gilt

$$\lim_{x \nearrow a} \xi(x) = a \quad \text{und} \quad \lim_{x \nearrow a} \eta(x) = a,$$

aufgrund unserer ZUSÄTZLICHEN VORAUSSETZUNG existiert der Grenzwert des rechten Ausdrucks in $(*)$

$$\lim_{x \nearrow a} \frac{f'(\xi(x))}{g'(\eta(x))} = \lim_{x \nearrow a} \frac{f'(x)}{g'(x)}.$$

Deshalb existiert auch der Grenzwert des linken Ausdrucks in $(*)$, die beiden Grenzwerte sind gleich. ♦

Beispiele Existiert jeweils der Grenzwert?

(1) Existiert der Grenzwert $\lim\limits_{x \overset{\bullet}{\to} 0} \dfrac{\sin(x)}{x}$?

Wir berechnen die Ableitungen von Zähler und Nenner und prüfen die Konvergenz des Quotienten:

$$f'(x) = \cos(x), \quad g'(x) = 1, \qquad \lim_{x \overset{\bullet}{\to} 0} \frac{\cos(x)}{1} = 1.$$

Also gilt:

$$\lim_{x \overset{\bullet}{\to} 0} \frac{\sin(x)}{x} = 1.$$

(2) Existiert der Grenzwert $\lim\limits_{x \overset{\bullet}{\to} 0} \dfrac{\exp(x) - 1}{x}$?

Wir berechnen die Ableitungen von Zähler und Nenner und prüfen die Konvergenz des Quotienten:

$$f'(x) = \exp(x), \quad g'(x) = 1, \qquad \lim_{x \overset{\bullet}{\to} 0} \frac{\exp(x)}{1} = 1.$$

Also gilt:

$$\lim_{x \overset{\bullet}{\to} 0} \frac{\exp(x) - 1}{x} = 1.$$

(3) Existiert der Grenzwert $\lim\limits_{x \overset{\bullet}{\to} 1} \dfrac{3x^3 - 9x + 6}{x^3 - 5x^2 + 7x - 3}$?

Wir berechnen die Ableitungen von Zähler und Nenner und prüfen die Konvergenz des Quotienten:

$$f'(x) = 9x^2 - 9, \quad g'(x) = 3x^2 - 10x + 7.$$

Es liegt wieder die Situation $\frac{0}{0}$ vor. Wir können versuchen, auf diese Situation die Regel von l'Hospital erneut anzuwenden.

(3b) Existiert der Grenzwert $\lim\limits_{x \overset{\bullet}{\to} 1} \dfrac{9x^2 - 9}{3x^2 - 10x + 7}$?

Wir berechnen erneut die (zweiten) Ableitungen von Zähler und Nenner und prüfen die Konvergenz des Quotienten:

$$f'(x) = 18x, \quad g'(x) = 6x - 10, \quad \lim_{x \overset{\bullet}{\to} 1} \frac{18x}{6x - 10} = \frac{18}{-4} = -\frac{9}{2}.$$

Also gilt:

$$\lim_{x \overset{\bullet}{\to} 1} \frac{9x^2 - 9}{3x^2 - 10x + 7} = -\frac{9}{2}$$

und deshalb

$$\lim_{x \overset{\bullet}{\to} 1} \frac{3x^3 - 9x + 6}{x^3 - 5x^2 + 7x - 3} = -\frac{9}{2}.$$

Man hätte dieses Ergebnis auch durch Kürzen des quadratischen Polynoms $(x-1)^2$ mittels Polynomdivision erhalten können.

Der Satz und die bisherigen Beispiele betreffen die Situation

$$\lim_{x \nearrow a} f(x) = 0, \qquad \lim_{x \nearrow a} g(x) = 0,$$

die wir mit dem Symbol $\boxed{\frac{0}{0}}$ (kein mathematisches Objekt) umschreiben können. Andere Grenzwertsituationen können darauf zurückgeführt werden:

- Existiert der Grenzwert

$$\lim_{x \to a} \frac{f(x)}{g(x)}, \qquad \text{wobei} \qquad \lim_{x \to a} f(x) = \infty, \quad \lim_{x \to a} g(x) = \infty \quad ?$$

  Diese Situation $\boxed{\frac{\infty}{\infty}}$ kann durch eine kleine Umformung auf die Situation $\boxed{\frac{0}{0}}$ zurückgeführt werden:

$$\frac{f(x)}{g(x)} = \frac{\frac{1}{g(x)}}{\frac{1}{f(x)}}.$$

- Existiert der Grenzwert

$$\lim_{x \to a} f(x) \cdot g(x), \qquad \text{wobei} \qquad \lim_{x \to a} f(x) = 0, \quad \lim_{x \to a} g(x) = \infty \quad ?$$

  Diese Situation $\boxed{0 \cdot \infty}$ kann ebenfalls durch eine kleine Umformung auf die Situation $\boxed{\frac{0}{0}}$ zurückgeführt werden:

$$f(x) \cdot g(x) = \frac{f(x)}{\frac{1}{g(x)}}.$$

- Existiert der Grenzwert

$$\lim_{x \to a} \Big[f(x) - g(x) \Big], \qquad \text{wobei} \qquad \lim_{x \to a} f(x) = \infty, \quad \lim_{x \to a} g(x) = \infty \quad ?$$

  Diese Situation $\boxed{\infty - \infty}$ kann durch eine kleine Umformung auf die Situation $\boxed{\frac{\infty}{\infty}}$ und dann — gegebenenfalls — auf $\boxed{0 \cdot \infty}$ zurückgeführt werden:

$$f(x) - g(x) = [\frac{f(x)}{g(x)} - 1] \cdot g(x).$$

Satz 16 (Regel von L'Hospital für Grenzwerte im Unendlichen) *Es seien* $f, g :]c, \infty[\to \mathbb{R}$ *mit* $c \in \mathbb{R} \cup \{-\infty\}$ *zwei differenzierbare Funktionen.* g *habe keine Nullstelle. Es sei die Voraussetzung*

$$\lim_{x\to\infty} f(x) = 0 \qquad \lim_{x\to\infty} g(x) = 0.$$

erfüllt.
Wenn der Grenzwert auf der rechten Seite existiert, dann gilt

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

Beweis Wir führen die Aussage dieses zweiten Satzes durch eine Substitution (bijektive Transformation der Definitionsmenge) auf die des ersten Satzes zurück: Mit der bijektiven Abbildung

$$\begin{cases}]\gamma, +\frac{\pi}{2}[& \to &]c, \infty[\\ y & \mapsto & x = \tan y \end{cases}$$

definieren wir die transformierten Funktionen

$$F \begin{cases}]\gamma, +\frac{\pi}{2}[& \to & \mathbb{R} \\ y & \mapsto & f(\tan y) \end{cases}, \qquad G \begin{cases}]\gamma, +\frac{\pi}{2}[& \to & \mathbb{R} \\ y & \mapsto & g(\tan y) \end{cases}.$$

Dann können wir die Voraussetzungen des aktuellen Satzes in die Voraussetzungen des vorherigen Satzes 15 übersetzen:

$$\begin{aligned}
\lim_{y\nearrow\frac{\pi}{2}} F(y) &= \lim_{y\nearrow\frac{\pi}{2}} f(\tan y) = \lim_{x\to\infty} f(x) = 0, \\
\lim_{y\nearrow\frac{\pi}{2}} G(y) &= \lim_{y\nearrow\frac{\pi}{2}} g(\tan y) = \lim_{x\to\infty} g(x) = 0, \\
\lim_{y\nearrow\frac{\pi}{2}} \frac{F'(y)}{G'(y)} &= \lim_{y\nearrow\frac{\pi}{2}} \frac{[f(\tan y)]'}{[g(\tan y)]'} = \lim_{y\nearrow\frac{\pi}{2}} \frac{f'(\tan y)\cdot\frac{1}{\cos^2 y}}{g'(\tan y)\cdot\frac{1}{\cos^2 y}} \\
&= \lim_{y\nearrow\frac{\pi}{2}} \frac{f'(\tan y)}{g'(\tan y)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = b.
\end{aligned}$$

Damit haben wir die Voraussetzungen von Satz 15 nachgewiesen.

Die Rücktransformation liefert jetzt die Behauptung:

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{F(\arctan x)}{G(\arctan x)} = \lim_{y\nearrow\frac{\pi}{2}} \frac{F(y)}{G(y)} \overset{\text{Satz 15}}{=} b.$$

♦

2.8.1 Beweis des Satzes von l'Hospital

Präposition 17 *Es sei* $f :]c, a] \to \mathbb{R}$ *eine stetige, in* $]c, a[$ *differenzierbare Funktion. Wenn der Grenzwert*

$$\lim_{x\nearrow a} f'(x) = b$$

existiert, dann existiert auch der Grenzwert

$$\lim_{x\nearrow a} \frac{f(x) - f(a)}{x - a} = b.$$

Beweis Es sei ein $\varepsilon > 0$ vorgegeben. Wegen der Voraussetzung existiert ein $x_0 \in \,]c,a[$, so dass

$$b - \varepsilon \leq f'(x) \leq b + \varepsilon \qquad \text{für alle } x_0 < x < a.$$

Mit Folgerung 11 (i), $y = a$, folgt:

$$b - \varepsilon \leq \frac{f(x) - f(a)}{x - a} \leq b + \varepsilon \qquad \text{für alle } x_0 < x < a,$$

was zu

$$\left| \frac{f(x) - f(a)}{x - a} - b \right| \leq \varepsilon$$

umgeformt werden kann. Das ist die Grenzwertaussage. ♦

Präposition 18 *Es seien $f, g : \,]c,a] \to \mathbb{R}$ stetige, in $]c,a[$ differenzierbare Funktionen. g' habe in $]c,a[$ keine Nullstelle. Wenn der Grenzwert*

$$\lim_{x \nearrow a} \frac{f'(x)}{g'(x)} = b$$

existiert, dann existiert auch der Grenzwert

$$\lim_{x \nearrow a} \frac{f(x) - f(a)}{g(x) - g(a)} = b.$$

Bemerkung: Mit $g(x) = x$ enthält diese Präposition die vorherige.

Beweis (1) Gemäß Satz 9 ist die Abbildung $g : \,]c,a[\to \mathbb{R}$ auf jedem kompakten Teilintervall von $]c,a[$, also auch auf ganz $]c,a[$ streng monoton. Wir betrachten nur den Fall, dass g streng monoton wachsend ist. Insgesamt wissen wir, dass die Abbildung

$$g : \,]c,a[\to \,]g(c), g(a)[$$

streng monoton wachsend und differenzierbar mit streng monoton wachsende differenzierbarer Umkehrabbildung

$$h : \,]g(c), g(a)[\to \,]c,a[$$

ist.

(2) Wir definieren die Funktion

$$F = f \circ h : \,]g(c), g(a)[\to \mathbb{R}$$

und überlegen mit der Kettenregel und Satz 5:

$$F'(y) = (f \circ h)'(y) \stackrel{\text{KR}}{=} f'(h(y)) \cdot h'(y) \stackrel{\text{Satz 5}}{=} \frac{f'(h(y))}{g'(h(y))}$$

Es gilt jetzt

$$\lim_{y \nearrow g(a)} F'(y) = \lim_{y \nearrow g(a)} \frac{f'(h(y))}{g'(h(y))} = \lim_{x \nearrow a} \frac{f'(x)}{g'(x)} = b.$$

Daraus folgt mit der ersten Präposition

$$\lim_{y \nearrow g(a)} \frac{F(y)}{y} = b$$

und dann mit Präposition 17

$$\lim_{x \nearrow a} \frac{f(x) - f(a)}{g(x) - g(a)} = \lim_{y \nearrow g(a)} \frac{f(h(y)) - f(h(g(a)))}{y - g(a)} = \lim_{y \nearrow g(a)} \frac{F(y) - F(g(a))}{y - g(a)} = b.$$

♦

3 Integrierbare Funktionen

3.1 Definition

Definition (Zerlegungen) Es sei ein (echtes) Intervall $[c,d]$ gegeben.

(1) Eine **endliche** (streng monoton steigende) Folge

$$c = x_0 < x_1 < \ldots < x_n = d$$

von $n+1$ Zahlen in $[c,d]$ heißt eine *Unterteilung* oder *Zerlegung* $\mathcal{Z}$ (von $[c,d]$ in n Teilintervalle). Die Zahlen x_i heißen *Zerlegungsstellen*.

(2) Eine Zerlegung $\mathcal{Z}_2$ heißt Verfeinerung der Zerlegung $\mathcal{Z}_1$, wenn jede Zerlegungsstelle von $\mathcal{Z}_1$ auch eine von $\mathcal{Z}_2$ ist.

(3) Sind zwei Zerlegungen $\mathcal{Z}_1$ und $\mathcal{Z}_2$ eines Intervalls $[c,d]$ gegeben, so können wir die *gemeinsame Verfeinerung* $\mathcal{Z}_1 \sqcup \mathcal{Z}_2$ bilden, deren Menge der Zerlegungsstellen durch Vereinigung der beiden gegebenen Mengen von Zerlegungsstellen gebildet wird.

Definition (Riemann–Integral) Es sei $f : [c,d] \to \mathbb{R}$ eine reelle beschränkte Funktion, das heißt es gibt eine Konstante $M > 0$, so dass

$$|f(x)| \leq M \qquad \text{für alle} \qquad x \in [c,d].$$

(1a) Die $\mathcal{Z}$*–Obersumme* von f ist die reelle Zahl

$$\overline{S}(f,\mathcal{Z}) \; := \; \sum_{i=1}^{n} \sup_{x \in]x_{i-1},x_i[} \{\, f(x) \,\} \cdot (x_i - x_{i-1})$$

(1b) Die $\mathcal{Z}$*–Untersumme* von f ist die reelle Zahl

$$\underline{S}(f,\mathcal{Z}) \; := \; \sum_{i=1}^{n} \inf_{x \in]x_{i-1},x_i[} \{\, f(x) \,\} \cdot (x_i - x_{i-1})$$

(2a) Das *Oberintegral* von f ist die reelle Zahl

$$\overline{\int_c^d} f(x)\,dx \; := \; \inf_{\mathcal{Z}} \left\{ \overline{S}(f,\mathcal{Z}) \right\}.$$

(2b) Das *Unterintegral* von f ist die reelle Zahl

$$\underline{\int_c^d} f(x)\,dx \; := \; \sup_{\mathcal{Z}} \left\{ \underline{S}(f,\mathcal{Z}) \right\}.$$

(3) Es ist unmittelbar klar, dass immer

$$\underline{\int_c^d} f(x)\,dx \leq \overline{\int_c^d} f(x)\,dx.$$

Die Funktion f heißt *(Riemann–)integrierbar* (auf $[c, d]$), wenn Oberintegral und Unterintegral übereinstimmen. Diese Zahl

$$\int_c^d f(x)\, dx \;:=\; \overline{\int_c^d} f(x)\, dx = \underline{\int_c^d} f(x)\, dx$$

heißt dann das *(Riemann–)Integral* (Bernhard Riemann).

*** Illustration mit Klappfolien ***

Bemerkungen

(1) Die $\mathcal{Z}$–Obersumme kann man sich vorstellen als die (vorzeichenbehaftete) Fläche eines Rechtecksystems, das durch die horizontale Achse, die Abszissen der Intervallgrenzen und $\mathcal{Z}$–Strecken, die von oberhalb auf den Funktionsgraphen ,,aufgesetzt" sind, gebildet wird.

(2) Entsprechend ist die $\mathcal{Z}$–Untersumme als Fläche eines Rechtecksystems aufzufassen, das durch die horizontale Achse, die Abszissen der Intervallgrenzen und $\mathcal{Z}$–Strecken, die von unterhalb an den Funktionsgraphen ,,angepasst" sind, gebildet wird.

(3) Beachte, dass die Werte der Funktion f an den Zerlegungsstellen für die Definition von Ober-, Untersumme, damit Ober-, Unterintegral und Integral ohne Belang sind. Das bedeutet, dass man eine Funktion f an endlich vielen Stellen abändern kann, ohne dass dies einen Einfluß auf die Integrierbarkeit oder den Integralwert hätte. Nützlich ist im folgenden die Abkürzung

$$\operatorname{mod} \mathcal{Z}.$$

Sie bedeutet ,,bis auf Zerlegungsstellen" oder ,,ohne Zerlegungsstellen".

(4) Wenn man das Oberintegral berechnen will, muss man im Prinzip erst die Obersummen bzgl. aller möglichen Zerlegungen berechnen und dann das Infimum (größte untere Grenze) all dieser Zahlen bilden. Dies scheint eine unmögliche Aufgabe zu sein. Wir werden sehen, wie man das — by the help of MATH POWER — bewältigen kann.

(5) In der Mathematik ungleich wichtiger und ,,natürlicher" ist ein anderer Integrationsbegriff, der auf Henri Lebesgue zurückgeht. Der wesentliche Unterschied zum Riemann–Begriff besteht darin, dass man dabei die Definitionsmenge der zu integrierenden Funktion in abzählbar unendlich viele Teilmengen zerlegen darf. Leider ist dies viel aufwändiger, wir können das hier nicht bewerkstelligen.

Wir ergänzen die obige Definition. Ist $f : [c, d] \to \mathbb{R}$ integrierbar und $a, b \in [c, d]$, so setzen wir

$$\int_a^b f(x)\,dx \;:=\; \begin{cases} \int_a^b f(x)\,dx, & \text{falls } a < b \quad \text{(wie oben)}, \\ 0, & \text{falls } a = b, \\ -\int_b^a f(x)\,dx, & \text{falls } a > b. \end{cases}$$

Ein erster Satz, dessen Inhalte „unmittelbar klar" sind, ist:

Satz 19 *Es seien f, g auf $[c, d]$ integrierbar.*
(1) Es gilt dann für beliebige $a, b, e \in [c, d]$

$$\int_a^e f(x)\,dx + \int_e^b f(x)\,dx = \int_a^b f(x)\,dx.$$

(2) Falls $f(x) \leq g(x)$ für alle $x \in [c, d]$ ist, gilt

$$\int_c^d f(x)\,dx \leq \int_c^d g(x)\,dx.$$

3.2 Treppenfunktionen sind integrierbar

Definition (Treppenfunktionen)

(1) Eine Funktion $\varphi : [c, d] \to \mathbb{R}$ heißt (naheliegend) $\mathcal{Z}$–*Treppenfunktion*, wenn sie auf jedem der n offenen Teilintervalle $]x_{i-1}, x_i[$, $i = 1, \ldots, n$, konstant ist.

(2) Eine Funktion $\varphi : [c, d] \to \mathbb{R}$ heißt einfach *Treppenfunktion*, wenn es eine Zerlegung $\mathcal{Z}$ gibt, so dass φ $\mathcal{Z}$–Treppenfunktion ist.

Die Werte einer Treppenfunktion an den Zerlegungsstellen sind für die im folgenden zu entwickelnden Begriffe und Sätze im wesentlichen ohne Bedeutung. Um technische Penibilitäten zu vermeiden, denken wir uns Treppenfunktionen immer $\mod \mathcal{Z}$ definiert. Dabei kann man sich selbst heraussuchen, ob man die Treppenfunktionen an den Zerlegungsstellen geeignet ,,nachdefinieren" will oder ob man ständig mitbedenkt, dass die Zerlegungsstellen ohnehin keine Rolle spielen.

Satz 20 *Eine Treppenfunktion* $\varphi : [c, d] \to \mathbb{R}$ *ist integrierbar, wobei*

$$\int_c^d \varphi(x)\, dx = \sum_{k=1}^n \varphi(\frac{x_i + x_{i-1}}{2}) \cdot (x_i - x_{i-1}).$$

Beweis Ist φ eine $\mathcal{Z}$–Treppenfunktion, so gilt für jedes $i = 1, \ldots, n$

$$\inf_{x \in]x_{i-1}, x_i[} \{\varphi(x)\} = \varphi(\frac{x_i + x_{i-1}}{2}) = \sup_{x \in]x_{i-1}, x_i[} \{\varphi(x)\},$$

und deshalb

$$\underline{S}(f, \mathcal{Z}) = \sum_{k=1}^n \varphi(\frac{x_i + x_{i-1}}{2}) \cdot (x_i - x_{i-1}) = \overline{S}(f, \mathcal{Z}).$$

Das heißt, dass durch die Zerlegung $\mathcal{Z}$ bereits das Infimum der Obersummen und das Supremum der Untersummen realisiert ist. Es folgt

$$\underline{\int_c^d} f(x)\, dx = \sum_{k=1}^n \varphi(\frac{x_i + x_{i-1}}{2}) \cdot (x_i - x_{i-1}) = \overline{\int_c^d} f(x)\, dx.$$

♦

Für die Integrationstheorie wichtig sind Treppenfunktionen wegen des folgenden Satzes.

Satz 21 (Integrierbarkeits–Kriterium mittels Treppenfunktionen)
Betrachte die drei Aussagen über eine Funktion $f : [c,d] \to \mathbb{R}$.

(A) Die Funktion f ist integrierbar.

(B) Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $\mathcal{Z}$ und zwei $\mathcal{Z}$–Treppenfunktionen $\varphi, \psi : [c,d] \to \mathbb{R}$, so dass

$$\varphi(x) \le f(x) \le \psi(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}$$

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \varepsilon.$$

(C) Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $\mathcal{Z}$ und zwei $\mathcal{Z}$–Treppenfunktionen $\varphi, \psi : [c,d] \to \mathbb{R}$, so dass

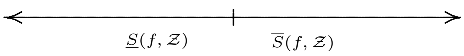
$$\left.\begin{array}{l} \varphi(x) \le f(x) \le \psi(x) \\ \psi(x) - \varphi(x) \le \varepsilon \end{array}\right. \quad \text{für alle } x \in [c,d] \mod \mathcal{Z}.$$

Dann gelten die folgenden Implikationen:

$$(A) \quad \Longleftrightarrow \quad (B) \quad \Longleftarrow \quad (C).$$

Beweis (A) $\Longrightarrow (B)$ Ist f integrierbar, so sind Ober- und Unterintegral gleich:

$$\sup_{\mathcal{Z}} \left\{ \underline{S}(f,\mathcal{Z}) \right\} = \underline{\int_c^d} f(x)\,dx = \overline{\int_c^d} f(x)\,dx = \inf_{\mathcal{Z}} \left\{ \overline{S}(f,\mathcal{Z}) \right\}.$$



Das bedeutet aber, dass es zu jedem $\varepsilon > 0$ Zerlegungen $\mathcal{Z}'$ und $\mathcal{Z}''$ gibt, so dass

$$\overline{S}(f,\mathcal{Z}') - \underline{S}(f,\mathcal{Z}'') \le \varepsilon.$$

Geht man zu der gemeinsamen Verfeinerung $\mathcal{Z} = \mathcal{Z}' \sqcup \mathcal{Z}''$ über und definiert man dann Treppenfunktionen durch

$$\left.\begin{array}{rcl} \varphi(x) & = & \inf\limits_{x \in]x_{i-1},x_i[} \{f(x)\} \\ \psi(x) & = & \sup\limits_{x \in]x_{i-1},x_i[} \{f(x)\} \end{array}\right. \quad \text{für } x \in \,]x_{i-1}, x_i[\text{ bzgl. } \mathcal{Z},$$

so gilt $\varphi(x) \le f(x) \le \psi(x)$, $x \in [c,d] \mod \mathcal{Z}$, und

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx = \overline{S}(f,\mathcal{Z}) - \underline{S}(f,\mathcal{Z}) \le \overline{S}(f,\mathcal{Z}') - \underline{S}(f,\mathcal{Z}'') \le \varepsilon.$$

(B) $\Longrightarrow (A)$ Es sei $n \in \mathbb{N}$. Zu $\varepsilon = \frac{1}{n}$ existieren $\mathcal{Z}_n$–Treppenfunktionen φ_n, ψ_n mit

$$\varphi_n(x) \le f(x) \le \psi_n(x) \quad \text{für alle } x \in [c,d] \mod \mathcal{Z}_n,$$

$$\int_c^d \psi_n(x)\,dx - \int_c^d \varphi_n(x)\,dx \le \frac{1}{n}.$$

Für die zugehörigen $\mathcal{Z}_n$–Ober- und Untersummen gilt:

$$\overline{S}(f, \mathcal{Z}_n) \leq \int_c^d \psi_n(x)\, dx, \qquad \underline{S}(f, \mathcal{Z}_n) \geq \int_c^d \varphi_n(x)$$

deshalb weiter

$$\overline{\int_c^d} f(x)\, dx - \underline{\int_c^d} f(x)\, dx \leq \overline{S}(f, \mathcal{Z}_n) - \underline{S}(f, \mathcal{Z}_n) \leq \int_c^d \psi_n(x)\, dx - \int_c^d \varphi_n(x)\, dx \leq \frac{1}{n}.$$

Da die letzte Ungleichung für alle $n \in \mathbb{N}$ gilt, folgt:

$$\overline{\int_c^d} f(x)\, dx - \underline{\int_c^d} f(x) = 0.$$

Das war zu beweisen.

(C) $\Longrightarrow$ (B). Es sei $\varepsilon > 0$ (wie in (B)) vorgegeben. Zu $\frac{\varepsilon}{d-c}$ existieren gemäß (C) $\mathcal{Z}$–Treppenfunktionen φ, ψ, so dass

$$\left.\begin{aligned} &\varphi(x) \leq f(x) \leq \psi(x) \\ &\psi(x) - \varphi(x) \leq \tfrac{\varepsilon}{d-c} \end{aligned}\right. \qquad \text{für alle } x \in [c, d] \mod \mathcal{Z}.$$

Die Funktion $\psi - \varphi$ ist eine $\mathcal{Z}$–Treppenfunktion, deren Betrag $\leq \frac{\varepsilon}{d-c}$ ist. Wir wollen nicht noch einmal technische Einzelheiten nachvollziehen und schließen etwas schneller als oben:

$$\int_c^d \psi(x)\, dx - \int_c^d \varphi(x)\, dx = \int_c^d [\psi - \varphi](x)\, dx \leq \frac{\varepsilon}{d-c} \cdot (d-c) = \varepsilon.$$

♦

Gibt es überhaupt beschränkte Funktionen, die nicht integrierbar sind?

Sätzchen 22 *Die Dirichlet–Funktion*

$$\chi_{\mathbb{Q}} \begin{cases} [0,1] & \to \quad \mathbb{R} \\ x & \mapsto \quad \begin{cases} 1, & \textit{falls } x \in [0,1] \cap \mathbb{Q}, \\ 0, & \textit{falls } x \in [0,1] \cap \mathbb{R} \setminus \mathbb{Q}. \end{cases} \end{cases}$$

ist nicht Riemann–integrierbar. (Sie ist Lebesgue–integrierbar.)

Beweis Für jedes Paar von Treppenfunktionen φ, ψ mit

$$\varphi(x) \le \chi_{\mathbb{Q}}(x) \le \psi(x) \quad x \in [0,1] \bmod \mathcal{Z}$$

gilt nämlich zwangsläufig:

$$\varphi(x) \le 0 \qquad \text{und} \qquad \psi(x) \ge 1 \quad x \in [0,1] \bmod \mathcal{Z}$$

und damit

$$\int_0^1 \psi(x)\,dx - \int_0^1 \varphi(x) \ge 1.$$

Die letzte Differenz kann also nicht kleiner als $\varepsilon = \frac{1}{2}$ (beispielsweise) gemacht werden. ♦

3.3 Stetige Funktionen sind integrierbar

Satz 23 (Stetigkeit $\Longrightarrow$ Integrierbarkeit) *Ist die Funktion $f : [c,d] \to \mathbb{R}$ stetig, so ist sie integrierbar.*

Der Beweis erfordert ein wenig Vorbereitung.

Präposition 24 *Es sei $f : [c,d] \to \mathbb{R}$ stetig und $a \in [c,d]$. Zu jedem $\varepsilon > 0$ gibt es Konstanten $\delta > 0$, $M^+, M^- \in \mathbb{R}$, so dass*

$$\begin{aligned} &M^- \le f(x) \le M^+ \qquad \textit{für alle } x \in [a-\delta, a+\delta] \cap]c,d[\\ &M^+ - M^- \le \varepsilon. \end{aligned}$$

Beweis Wegen der Stetigkeit von f in a gibt es zu $\frac{\varepsilon}{2}$ ein $\delta > 0$, so dass

$$|f(x) - f(a)| \le \frac{\varepsilon}{2} \qquad \text{für alle } x \in [c,d] \text{ mit } |x-a| \le \delta.$$

Setzt man

$$\begin{aligned} M^- &:= f(a) - \frac{\varepsilon}{2} \\ M^+ &:= f(a) + \frac{\varepsilon}{2}, \end{aligned}$$

so sind die Aussagen des Satzes erfüllt: $M^+ - M^- = \varepsilon$ und

$$M^- = f(a) - \frac{\varepsilon}{2} \le f(x) \le f(a) + \frac{\varepsilon}{2} = M^+ \qquad \text{für } x \in [a-\delta, a+\delta] \cap]c,d[.$$

♦

Beweis (von Satz 23) Wir wollen die Implikation $(C) \Longrightarrow (A)$ aus Satz 21 anwenden. Dazu sei ein $\varepsilon > 0$ vorgegeben.

Wir betrachten die Zahlen $y \in [c,d]$ mit der folgenden Eigenschaft:

$(\mathcal{A}_y)$ Es gibt eine Zerlegung $\mathcal{Z}$ von $[c,y]$ und zwei $\mathcal{Z}$–Treppenfunktionen $\varphi, \psi : [c,y] \to \mathbb{R}$ mit

$$\left.\begin{aligned} &\varphi(x) \le f(x) \le \psi(x) \\ &\psi(x) - \varphi(x) \le \varepsilon \end{aligned}\right\} \quad \text{für alle } x \in [c,y] \mod \mathcal{Z}.$$

Es sei $Y \subseteq [c,d]$ die Menge der Zahlen, die diese Eigenschaft $(\mathcal{A}_y)$ erfüllen. Wir setzen

$$a := \sup Y = \sup \left\{ y \in [c,d] \;\middle|\; (\mathcal{A}_y) \text{ ist erfüllt} \right\}.$$

Es gilt offenbar $a \ge c$. Zur Beweisführung nehmen wir an, dass $a < d$. Gemäß der Präposition gibt es $\delta > 0$ und Konstanten M^+, M^- mit

$$\begin{aligned} &M^- \le f(x) \le M^+ \qquad \text{für alle } x \in [a-\delta, a+\delta] \cap]c,d[\\ &M^+ - M^- \le \varepsilon. \end{aligned}$$

Gemäß Definition von a ist die Eigenschaft $\mathcal{A}_{a-\delta}$ erfüllt. Das heißt, auf dem Intervall $[c, a-\delta]$ existieren bereits die richtigen Treppenfunktionen. Ergänzt man diese Treppenfunktionen mit Hilfe der Konstanten M^-, M^+ zu Treppenfunktionen auf dem Intervall $[c, a+\delta]$, so ist die Eigenschaft $\mathcal{A}_y$ für alle $c \le y \le a+\delta$ erfüllt. Das steht aber im Widerspruch zur Definition von a. ♦

*** Zeichnung ***

3.4 Monotone Funktionen sind integrierbar

Satz 25 *Ist eine Funktion $f : [c, d] \to \mathbb{R}$ monoton steigend oder monoton fallend, so ist sie integrierbar.*

Beweis Wir wenden die Implikation $(B) \implies (A)$ aus Satz 21 an und beschränken uns dabei auf den Fall ,,monoton wachsend". Dazu sei $\varepsilon > 0$ vorgegeben. Es sei n eine natürliche Zahl mit

$$n \geq \frac{[f(d) - f(c)] \cdot (d - c)}{\varepsilon}$$

und die (äquidistante) Zerlegung $\mathcal{Z}$ gegeben durch

$$x_i \ := \ c + i \cdot \frac{d - c}{n}, \qquad i = 0, \ldots, n.$$

Wir definieren $\mathcal{Z}$–Treppenfunktionen durch die intervallweise Festlegung

$$\varphi(x) \ := \ f(x_{i-1}) \leq f(x_i) =: \psi(x), \qquad \text{falls } x \in]x_{i-1}, x_i[.$$

Aufgrund des Monoton–Wachstums von f gilt $\varphi(x) \leq f(x) \leq \psi(x)$ für $x \in [c, d] \mod \mathcal{Z}$. Außerdem ist

$$\int_c^d \psi(x)\, dx - \int_c^d \varphi(x)\, dx = \int_c^d [\psi - \varphi](x)\, dx =$$
$$\sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \cdot (x - x_{i-1}) = \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \cdot \frac{d - c}{n} =$$
$$[f(d) - f(c)] \cdot \frac{d - c}{n} \leq \varepsilon.$$

♦

*** Zeichnung bzw. Folie ****

Beispiel:
Es sei die Funktion $f : [c,d] \to \mathbb{R}$, $x \mapsto x^2$ gegeben. Wir berechnen für $c \geq 0$ die Obersumme bzgl. der äquidistanten Zerlegung $\mathcal{Z}_n$

$$x_i = c + \frac{i}{n}(d-c).$$

$$\begin{aligned}
\overline{\mathcal{S}}(f, \mathcal{Z}_n) &= \sum_{i=1}^{n} \sup_{x \in]x_{i-1}, x_i[} \{f(x)\} \cdot (x_i - x_{i-1}) \\
&= \sum_{i=1}^{n} x_i^2 \cdot \frac{d-c}{n} \\
&= \frac{d-c}{n} \cdot \sum_{i=1}^{n} [c + \frac{i}{n}(d-c)]^2 \\
&= \frac{d-c}{n} \cdot \sum_{i=1}^{n} [c^2 + 2\frac{i}{n}c(d-c) + i^2 \left(\frac{d-c}{n}\right)^2] \\
&= \frac{d-c}{n} \cdot \left[\sum_{i=1}^{n} c^2 + \frac{2}{n}c(d-c)\sum_{i=1}^{n} i + \left(\frac{d-c}{n}\right)^2 \sum_{i=1}^{n} i^2\right] \\
&= \frac{d-c}{n} \cdot \left[n \cdot c^2 + \frac{2}{n}c(d-c)\frac{n(n+1)}{2} + \left(\frac{d-c}{n}\right)^2 \frac{n(n+1)(2n+1)}{6}\right] \\
&= (d-c)c^2 + c(d-c)^2\frac{n+1}{n} + (d-c)^3\frac{(n+1)(2n+1)}{6n^2}
\end{aligned}$$

Diese Werte sind monoton fallend für größer werdendes n. Das Infimum ist zugleich der Grenzwert. Es gilt ($*$ müsste noch etwas feiner begründet werden)

$$\begin{aligned}
\inf_{\mathcal{Z}} \overline{\mathcal{S}}(f, \mathcal{Z}) &\overset{*}{=} \inf_{n \in \mathbb{N}} \overline{\mathcal{S}}(f, \mathcal{Z}_n) = \lim_{n \to \infty} \overline{\mathcal{S}}(f, \mathcal{Z}_n) \\
&= (d-c)c^2 + c(d-c)^2 + \frac{(d-c)^3}{3} = (d-c) \cdot \left[c^2 + cd - c^2 + \frac{d^2 - 2cd + c^2}{3}\right] \\
&= \frac{d-c}{3} \cdot \left[cd + d^2 + c^2\right] \\
&= \frac{d^3 - c^3}{3}.
\end{aligned}$$

3.5 Linearkombinationen sind integrierbar

Sätzchen 26 *Sind φ, ψ Treppenfunktionen und $\alpha, \beta \in \mathbb{R}$, so ist auch die Funktion*

$$\alpha\,\varphi + \beta\,\psi$$

eine Treppenfunktion.

Beweis Ist φ eine $\mathcal{Z}'$–Treppenfunktion und ψ eine $\mathcal{Z}''$–Treppenfunktion, so ist $\alpha\,\varphi + \beta\,\psi$ eine $\mathcal{Z}$–Treppenfunktion, wobei $\mathcal{Z} = \mathcal{Z}' \sqcup \mathcal{Z}''$ die gemeinsame Verfeinerung ist. Der genauere Beweis ist einfach, ein bißchen technisch. ♦

Satz 27 (Linearität) *Sind die beiden Funktionen $f, g : [c,d] \to \mathbb{R}$ integrierbar und $\alpha, \beta \in \mathbb{R}$, so ist auch die Funktion $\alpha f + \beta g$ integrierbar und es gilt:*

$$\int_c^d [\alpha f + \beta g](x)\,dx = \alpha \int_c^d f(x)\,dx + \beta \int_c^d g(x)\,dx.$$

Das bedeutet, die Menge der integrierbaren Funktionen ist ein $\mathbb{R}$–Vektorraum (vgl. LAL01). Die Menge der Treppenfunktionen ist ein Untervektorraum.

Beweis Wir zeigen zunächst, dass $f+g$ integrierbar ist und wenden dafür das Kriterium aus Satz 21 an. Dazu sei ein $\varepsilon > 0$ vorgegeben. Zu $\frac{\varepsilon}{2}$ existieren $\mathcal{Z}_1$–Treppenfunktionenen φ_1, ψ_1, so dass

$$\varphi_1(x) \le f(x) \le \psi_1(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}_1$$

$$\int_c^d \psi_1(x)\,dx - \int_c^d \varphi_1(x)\,dx \le \frac{\varepsilon}{2}.$$

und $\mathcal{Z}_2$–Treppenfunktionenen φ_2, ψ_2, so dass

$$\varphi_2(x) \le g(x) \le \psi_2(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}_2$$

$$\int_c^d \psi_2(x)\,dx - \int_c^d \varphi_2(x)\,dx \le \frac{\varepsilon}{2}.$$

Dann sind $\varphi_1 + \varphi_2$ und $\psi_1 + \psi_2$ $\mathcal{Z}_1 \sqcup \mathcal{Z}_2$–Treppenfunktionen, für die gilt:

$$\varphi_1(x) + \varphi_2(x) \le f(x) + g(x) \le \psi_1(x) + \psi_2(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}_1 \sqcup \mathcal{Z}_2$$

$$\int_c^d [\psi_1 + \psi_2](x)\,dx - \int_c^d [\varphi_1 + \varphi_2](x)\,dx \le \varepsilon.$$

Als nächstes überlegen wir, dass mit einer Funktion f auch die Funktion $-f$ integrierbar ist. Dies folgt aber mit Hilfe des Kriteriums, wenn man alle beteiligten Funktionen mit -1 multipliziert und die Rollen von $-\varphi$ und $-\psi$ vertauscht.

Ist f integrierbar und $\alpha > 0$, $\varepsilon > 0$ vorgegeben, so existieren $\mathcal{Z}$–Treppenfunktionen mit der Eigenschaft, dass

$$\varphi(x) \le f(x) \le \psi(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}$$

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \frac{\varepsilon}{\alpha}.$$

Multipliziert man diese beiden Gleichungen mit $\alpha > 0$, so entsteht das Kriterium für die Integrierbarkeit von αf

Da die Funktion $\alpha f + \beta g$ schrittweise aus f und g durch die obigen Operationen aufgebaut werden kann, ist der Satz gezeigt. ♦

3.6 Produkte sind integrierbar

Für eine Funktion $f : D \to \mathbb{R}$ definieren wir den *Positiv-* und *Negativteil*

$$f_+(x) := \begin{cases} f(x), & \text{falls } f(x) > 0, \\ 0, & \text{falls } f(x) \le 0. \end{cases} \qquad f_-(x) := \begin{cases} -f(x), & \text{falls } f(x) < 0, \\ 0, & \text{falls } f(x) \ge 0. \end{cases}$$

Es gilt dann

$$f = f_+ - f_- \qquad |f| = f_+ + f_-.$$

Satz 28 (Betrags–Integrierbarkeit) *Es sei $f : [c,d] \to \mathbb{R}$ integrierbar. Dann sind auch $f_+, f_-, |f|$ integrierbar. Es gilt:*

$$\Big| \int_c^d f(x)\,dx \Big| \le \int_c^d |f(x)|\,dx.$$

(Beachte die Ähnlichkeit zur Dreiecksungleichung)

Beweis Nach Voraussetzung gibt es zu $\varepsilon > 0$ Treppenfunktionen φ, ψ mit

$$\varphi(x) \le f(x) \le \psi(x) \ \text{ für alle } \ x \in [c,d] \ \bmod \mathcal{Z}_1$$
$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \varepsilon.$$

Für die zugehörigen Positivteile gilt:

$$\varphi_+(x) \le f_+(x) \le \psi_+(x) \ \text{ für alle } \ x \in [c,d] \ \bmod \mathcal{Z}_1$$
$$\int_c^d \psi_+(x)\,dx - \int_c^d \varphi_+(x)\,dx \le \int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \varepsilon.$$

Also ist auch f_+ integrierbar. Wegen $f_- = (-f)_+$ ist auch f_- integrierbar. Wegen $|f| = f_+ + f_-$ ist auch $|f|$ integrierbar. Dabei kann man schließen

$$\begin{aligned} f \le |f| \quad &\Longrightarrow \quad \int_c^d f(x)\,dx \le \int_c^d |f(x)|\,dx \\ -f \le |f| \quad &\Longrightarrow \quad -\int_c^d f(x)\,dx = \int_c^d [-f(x)]\,dx \le \int_c^d |f(x)|\,dx \end{aligned}$$

In jedem Fall gilt also

$$\Big| \int_c^d f(x)\,dx \Big| \le \int_c^d |f(x)|\,dx.$$

♦

Satz 29 (Integrierbarkeit von Produkten) *Es seien $f, g : [c,d] \to \mathbb{R}$ integrierbar.*
(i) Für jedes $p \in [1, \infty[$ ist die Funktion $|f|^p$ integrierbar.
(ii) Dann ist auch die Produktfunktion $f \cdot g$ integrierbar.

Beweis (i) Es sei $M := \sup\{|f(x)| \big| x \in [c,d]\}$. Ist ein $\varepsilon > 0$ vorgegeben, so existieren (vgl. Satz 28) $\mathcal{Z}$–Treppenfunktionen φ, ψ, so dass

$$0 \le \varphi(x) \le |f(x)| \le \psi(x) \le M \text{ für alle } x \in [c,d] \mod \mathcal{Z} \qquad (*)$$

$$\int_c^d \psi(x)\,dx - \int_c^d \varphi(x)\,dx \le \frac{\varepsilon}{p \cdot M^{p-1}}.$$

Sind φ_i, ψ_i die Werte der beiden Treppenfunktionen im Intervall $]x_{i-1}, x_i[$ der Zerlegung $\mathcal{Z}$, so gilt

$$|\psi_i^p - \varphi_i^p| \le p \cdot \xi^{p-1} |\psi_i - \varphi_i|,$$

wobei $\xi \in [\varphi_i, \psi_i] \subseteq [0, M]$ ein Mittelwert ist. (Mittelwertsatz 10,11 angewandt auf die Funktion x^p und die beiden Stellen φ_i, ψ_i). Wir formulieren um und schätzen weiter ab:

$$|\psi^p(x) - \varphi^p(x)| \le p \cdot M^{p-1} |\psi_i(x) - \varphi_i(x)| \quad \text{für alle } x \in [c,d] \mod \mathcal{Z}$$

Daraus folgt aber

$$\int_c^d |\psi^p(x) - \varphi^p(x)|\,dx \le p \cdot M^{p-1} \int_c^d |\psi(x) - \varphi(x)|\,dx \le p \cdot M^{p-1} \cdot \frac{\varepsilon}{p \cdot M^{p-1}} = \varepsilon.$$

Die Ungleichung

$$\varphi^p(x) \le |f|^p(x) \le \psi^p(x) \text{ für alle } x \in [c,d] \mod \mathcal{Z}$$

folgt direkt aus $(*)$.

(ii) Dies ist ganz einfach, da

$$fg = \frac{(f+g)^2 - (f-g)^2}{4}.$$

♦

3.7 Integration und Abschätzung, der Mittelwertsatz

Satz 30 (Abschätzungssatz, Mittelwertsatz) *Es sei die Funktionen $\mu : [c,d] \to \mathbb{R}$ eine nicht–negative integrierbare Funktion („eine Gewichtung auf dem Intervall $[c,d]$"). Dann gilt für eine integrierbare Funktion $f : [c,d] \to \mathbb{R}$:*
(1)

$$\inf_{x\in[c,d]}\{f(x)\} \cdot \int_c^d \mu(x)\,dx \le \int_c^d f(x)\mu(x)\,dx \le \sup_{x\in[c,d]}\{f(x)\} \cdot \int_c^d \mu(x)\,dx.$$

(2)

$$\inf_{x\in[c,d]}\{f(x)\} \cdot (d-c) \le \int_c^d f(x)\,dx \le \sup_{x\in[c,d]}\{f(x)\} \cdot (d-c).$$

(3) Ist f stetig, so gibt es ein $\xi \in [c,d]$, so dass

$$\int_c^d f(x)\mu(x)\,dx = f(\xi) \cdot \int_c^d \mu(x)\,dx.$$

(4) Ist f stetig, so gibt es ein $\xi \in [c,d]$, so dass

$$\int_c^d f(x)\,dx = f(\xi) \cdot (d-c).$$

Beweis Es gilt einfach:

$$\inf_{x\in[c,d]}\{f(x)\} \cdot \mu(x) \le f(x)\mu(x) \le \sup_{x\in[c,d]}\{f(x)\} \cdot \mu(x)$$

und deshalb mit Satz 19 (2) die Behauptung (1). Die Behauptung (2) ist ein Spezialfall von (1) für $\mu \equiv 1$. Die Behauptungen (3) und (4) ergeben sich aus dem Zwischenwertsatz.

Für den Fall $\int_c^d \mu(x)\,dx = 0$ folgt die Aussage (3) sofort aus (1).

Im Fall $\int_c^d \mu(x)\,dx > 0$ gilt mit (1)

$$\min_{x\in[c,d]}\{f(x)\} = \inf_{x\in[c,d]}\{f(x)\} \le \frac{\int_c^d f(x)\mu(x)\,dx}{\int_c^d \mu(x)\,dx} \le \sup_{x\in[c,d]}\{f(x)\} = \max_{x\in[c,d]}\{f(x)\}.$$

Also existiert eine Zwischenstelle $\xi \in [c,d]$ mit

$$f(\xi) = \frac{\int_c^d f(x)\mu(x)\,dx}{\int_c^d \mu(x)\,dx}.$$

♦

3.8 Der Hauptsatz der Differential- und Integralrechnung

Satz 31 (Hauptsatz der Differential- und Integralrechnung (HDI)) *Es sei $J \subseteq \mathbb{R}$ ein Intervall und $f : J \to \mathbb{R}$ eine stetige Funktion, $a \in J$ eine fest gewählte Stelle.*

Für eine andere Funktion $F : J \to \mathbb{R}$ sind die beiden folgenden Aussagen äquivalent:

*(A) F ist die **Integralfunktion (Unbestimmtes Integral)** zu f mit unterer Grenze a, das heißt*

$$F(x) = \int_a^x f(y)\,dy.$$

*(B) F ist die **Stammfunktion** von f mit Nullstelle bei a, das heißt es gilt*

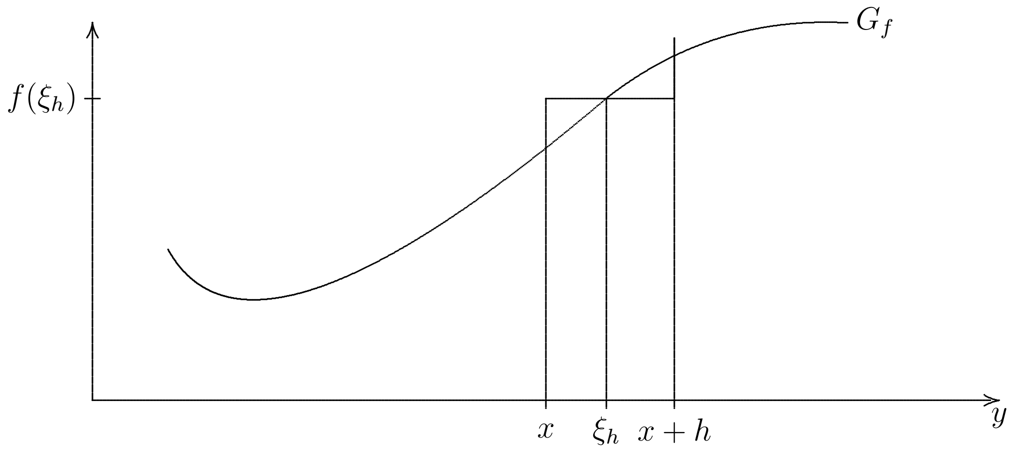
$$F'(x) = f(x) \quad \text{für alle } x \in J \qquad \text{und} \quad F(a) = 0.$$

Bemerkung: Ist x ein Endpunkt des Intervalls J, so ist die Ableitung als einseitige Ableitung definiert.

Beweis Wir beweisen die Folgerung $(A) \Longrightarrow (B)$:
Gemäß Aussage (A) sei

$$F(x) = \int_a^x f(y)\,dy$$

eine Integralfunktion zu f mit unterer Grenze a im Intervall J. Für ein beliebiges $x \in J$ müssen wir $F'(x) = f(x)$ beweisen.



Wir betrachten für $h > 0$, wobei $x, x+h \in J$, den Differenzenquotienten

$$\frac{F(x+h) - F(x)}{h} = \frac{\int_a^{x+h} f(y)\,dy - \int_a^x f(y)\,dy}{h} = \frac{\int_x^{x+h} f(y)\,dy}{h}.$$

Im Intervall $[x; x+h]$ gibt es gemäß Mittelwertsatz eine Stelle ξ, so dass

$$\int_x^{x+h} f(y)\,dy = f(\xi) \cdot h,$$

insgesamt also

$$\frac{F(x+h)-F(x)}{h} = \frac{\int_x^{x+h} f(y)\,dy}{h} = \frac{f(\xi)\cdot h}{h} = f(\xi).$$

Wir lassen jetzt $h \searrow 0$ gehen. Wegen $\xi \in [x; x+h]$ gilt

$$\lim_{h\searrow 0} \xi = x$$

und, da die Funktion f als stetig vorausgesetzt ist,

$$\lim_{h\searrow 0} f(\xi) = f(x).$$

Mit der Gleichung oben folgt insgesamt:

$$\lim_{h\searrow 0} \frac{F(x+h)-F(x)}{h} = \lim_{h\searrow 0} f(\xi) = f(x).$$

Eine analoge Überlegung mit $h < 0$ führt auf

$$\lim_{h\nearrow 0} \frac{F(x+h)-F(x)}{h} = f(x),$$

insgesamt also

$$F'(x) = \lim_{h \overset{\bullet}{\to} 0} \frac{F(x+h)-F(x)}{h} = f(x).$$

Zum Schluss bemerken wir noch, dass

$$F(a) = \int_a^a f(y)\,dy = 0.$$

Damit ist die Aussage (B) gezeigt.

Wir beweisen die Folgerung $(B) \Longrightarrow (A))$
Gemäß Aussage (B) sei F eine Stammfunktion von f im Intervall J mit Nullstelle a:

$$F'(x) = f(x) \quad \text{für} \quad x \in J, \qquad F(a) = 0.$$

Nach der (bereits bewiesenen) Folgerung $(A) \Longrightarrow (B)$ ist

$$G(x) = \int_a^x f(y)\,dy$$

ebenfalls eine Stammfunktion zu f mit $G(a) = 0$. Für die Differenz $H(x) = F(x) - G(x)$ dieser beiden Stammfunktionen von f gilt nun:

$$H'(x) = F'(x) - G'(x) = f(x) - f(x) = 0.$$

Nach Folgerung 13 ist diese Funktion konstant, und zwar wegen

$$H(a) = F(a) - G(a) = 0 - 0 = 0$$

gleich Null. Daraus folgt mit Folgerung 13

$$F(x) = G(x) = \int_a^x f(y)\,dy.$$

Das ist die zu beweisende Aussage (A).

♦

3.9 Anwendungen des HDI: Berechnung von Integralen

Man sagt, eine Funktion ist *elementar*, wenn sie als Verknüpfung (Grundrechenarten, Hintereinanderausführung) von Polynom–, Exponential–, trigonometrischen Funktionen und deren Umkehrfunktionen gebildet werden kann. Bei Zugrundelegung geeigneter Definitionsmengen sind elementare Funktionen differenzierbar, also

$$\text{elementar} \quad \Longrightarrow \quad \text{differenzierbar} \quad \Longrightarrow \quad \text{stetig} \quad \Longrightarrow \quad \text{integrierbar.}$$

Aufgrund der Differentiationsregeln ist einsichtig, dass die Ableitungsfunktion einer beliebigen elementaren Funktion wieder eine elementare Funktion ist.

Beachte aber, dass es elementare Funktionen gibt, deren Stammfunktion nicht elementar ist. Beispiele dafür sind die Funktionen $\sin(x^2)$ oder $x^\alpha \cdot (1-x)^\beta$.

Der HDI erleichtert erheblich die Berechnung von Integralen vieler elementarer Funktionen.

Als abkürzende Schreibweise für die Differenz von zwei Funktionswerten führen wir ein:

$$\Big[f(x)\Big]_{x=a}^{x=b} = \Big[f(x)\Big]_{x=a}^{x=b} = f(x)\Big|_{x=a}^{x=b} = f(x)\Big|_a^b := f(b) - f(a).$$

Beispiele Die Formel

$$\int_c^d x^\alpha \, dx = \frac{x^{\alpha+1}}{\alpha+1}\Big|_c^d$$

ist gültig, wenn die Ausdrücke auf den beiden Seiten wohldefiniert sind, das heißt, eine der folgenden Bedingungen erfüllt ist:

$$\begin{array}{l} \alpha \geq 0 \\ \alpha \in \mathbb{Z}, \alpha \leq -2 \quad \text{und} \quad (c,d \in \mathbb{R}^+ \text{ oder } c,d \in \mathbb{R}^-), \\ \alpha \in \mathbb{R} \setminus \{-1\} \quad \text{und} \quad c,d \in \mathbb{R}^+. \\ \alpha > -1 \qquad (**) \end{array}$$

Der Nachweis, dass die letzte Bedingung $(**)$ ebenfalls hinreichend ist für die obige Formel, erfordert eine Ausweitung des Integralbegriffs auf Funktionen f, die im Integrationsintervall $[c,d]$ Polstellen haben (Uneigentliche Integrale: Sie gehören nicht zum GS/HS/RS–Grundstock).

Für den Sonderfall $\alpha = -1$ und $(c,d \in \mathbb{R}^+$ oder $c,d \in \mathbb{R}^-)$ gilt:

$$\int_c^d x^{-1} \, dx = \ln|x| \Big|_c^d.$$

Für die natürliche Exponentialfunktion und die allgemeinere Exponentialfunktion $(a > 0)$ gelten:

$$\int_c^d \exp(x) \, dx = \exp(x)\Big|_c^d, \qquad \int_c^d a^x \, dx = \frac{a^x}{\ln(a)}\Big|_c^d.$$

Aus den Ableitungsformeln für trigonometrischen Funktionen ergeben sich die folgenden Integrationsformeln:

$$\begin{aligned}
\int_c^d \sin(x)\,dx &= -\cos(x)\Big|_c^d \\
\int_c^d \cos(x)\,dx &= \sin(x)\Big|_c^d \\
\int_c^d \frac{1}{\cos^2(x)}\,dx &= \tan(x)\Big|_c^d, \qquad \text{wobei } (k-\frac{1}{2})\pi < c,d < (k+\frac{1}{2})\pi \text{ für ein } k \in \mathbb{Z}.
\end{aligned}$$

Aus den Ableitungsformeln für die Umkehrfunktionen trigonometrischen Umkehrfunktionen ergeben sich die folgenden Integrationsformeln:

$$\begin{aligned}
\int_c^d \frac{1}{1+x^2}\,dx &= \arctan(x)\Big|_c^d \\
\int_c^d \frac{1}{\sqrt{1-x^2}}\,dx &= \arcsin(x)\Big|_c^d, \qquad \text{wobei } -1 < c,d < +1.
\end{aligned}$$

3.10 Partielle Integration

Mit Hilfe des HDI kann die Leibniz'sche Produktregel in eine Integrationstechnik umgeschrieben werden:

Satz 32 (Partielle Integration) *Es seien $f,g : [c,d] \to \mathbb{R}$ differenzierbare Funktionen, ihre Ableitungsfunktionen $f', g' : [c,d] \to \mathbb{R}$ (in den Endpunkten einseitig definiert) seien stetig. Dann gilt:*

$$\int_c^d f(x)g'(x)\,dx + \int_c^d f'(x)g(x)\,dx = \Big[f(x)g(x)\Big]_{x=c}^{x=d}.$$

Beweis Ist die Funktion F definiert durch $F(x) = f(x)g(x) - f(c)g(c)$, so gilt:

$$F'(x) = f(x)g'(x) + f'(x)g(x) \qquad \text{und} \qquad F(c) = 0.$$

Aufgrund der Implikation $(B) \Longrightarrow (A)$ aus dem HDI gilt:

$$\int_c^x [f(y)g'(y) + f'(y)g(y)]\,dy = F(x) = f(x)g(x) - f(c)g(c).$$

Für $x = d$ (und Ersetzung der Integrationsvariablen y durch x) ist dies die Aussage des Satzes. ♦

Beispiele (1) Für $c, d > 0$ ist

$$\begin{aligned}
\int_c^d \ln(x)\,dx &= \int_c^d \underbrace{1}_{f'(x)} \cdot \underbrace{\ln(x)}_{g(x)}\,dx = [\underbrace{x}_{f(x)} \cdot \underbrace{\ln(x)}_{g(x)}]_c^d - \int_c^d \underbrace{x}_{f(x)} \cdot \underbrace{\frac{1}{x}}_{g'(x)}\,dx \\
&= [x \cdot \ln(x) - x]_c^d.
\end{aligned}$$

(2) Für $c, d \in \mathbb{R}$ ist

$$\begin{aligned}\int_c^d \sin^2(x)\,dx &= \int_c^d \underbrace{\sin(x)}_{f'(x)} \cdot \underbrace{\sin(x)}_{g(x)}\,dx = [\underbrace{-\cos(x)}_{f(x)} \cdot \underbrace{\sin(x)}_{g(x)}]_c^d - \int_c^d \underbrace{-\cos(x)}_{f(x)} \cdot \underbrace{\cos(x)}_{g'(x)}\,dx \\ &= [\underbrace{-\cos(x)}_{f(x)} \cdot \underbrace{\sin(x)}_{g(x)}]_c^d + \int_c^d [1-\sin^2(x)]\,dx.\end{aligned}$$

Daraus folgt:

$$2 \cdot \int_c^d \sin^2(x)\,dx = [x - \cos(x)\cdot\sin(x)]_c^d.$$

Allgemeiner kann man mit Hilfe der Zerlegung

$$\sin^m(x) = \sin(x) \cdot \sin^{m-1}(x)$$

Integrale von Sinus–Potenzen bestimmen.

3.11 Die Substitutionsregel

Satz 33 (Substitutionsregel) *Es sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion und die Funktion $\vartheta : [c,d] \to \mathbb{R}$ differenzierbar mit stetiger Ableitungsfunktion $\vartheta' : [c,d] \to \mathbb{R}$ (in den Endpunkten einseitig). Das Bild von ϑ sei in $[a,b]$ enthalten: $\vartheta([c,d]) \subseteq [a,b]$. Dann gilt:*

$$\int_c^d f(\vartheta(y)) \cdot \vartheta'(y)\,dy = \int_{\vartheta(c)}^{\vartheta(d)} f(x)\,dx.$$

Beweis Ist die Funktion F irgendeine Stammfunktion von f, so gilt gemäß Kettenregel

$$(F \circ \vartheta)'(y) = F'(\vartheta(y)) \cdot \vartheta'(y) = f(\vartheta(y)) \cdot \vartheta'(y)$$

Die Implikation $(B) \Longrightarrow (A)$ aus dem HDI liefert:

$$\int_c^x f(\vartheta(y)) \cdot \vartheta'(y)\,dy = F(\vartheta(x)) - F(\vartheta(c)) = \int_{\vartheta(c)}^{\vartheta(x)} f(y)\,dy$$

Mit $x = d$ (und Ersetzung der Integrationsvariablen y durch x auf der rechten Seite) folgt der Satz. ♦

** Zeichnung **

Beispiele (1) Die lineare Transformation $\vartheta : y \mapsto x = \alpha x + \beta$ für die Definitionsmenge führt auf die Formel

$$\int_c^d \alpha \cdot f(\alpha y + \beta)\,dy = \int_{\alpha c+\beta}^{\alpha d+\beta} f(x)\,dx.$$

Beachte die Spezialfälle $\alpha = -1, \beta = 0$ (Spiegelung) und $\alpha = 1$ (Verschiebung).

(2) Beispiel: Für $\alpha \neq 0$ ist

$$\int_c^d (\alpha y + \beta)^2\,dy = \frac{1}{\alpha} \int_{\alpha c+\beta}^{\alpha d+\beta} x^2\,dx = \frac{(\alpha d + \beta)^3 - (\alpha c + \beta)^3}{3\alpha}.$$

Satz 34 (Logarithmische Integration) *Es sei* $g : [c,d] \to \mathbb{R}^+$ *(oder* $\mathbb{R}^-$*) differenzierbar mit stetiger Ableitung. Dann gilt*

$$\int_c^d \frac{g'(y)}{g(y)}\, dy = (\ln|g(y)|)\Big|_c^d.$$

Beweis Dies ist eine direkte Anwendung der Substitutionsregel mit den Ersetzungen $f(x) \rightsquigarrow \frac{1}{x}$ und $\vartheta(y) \rightsquigarrow g(y)$.

$$\int_c^d \frac{g'(y)}{g(y)}\, dy = \int_c^d f(g(y)) \cdot g'(y)\, dy = \int_{g(c)}^{g(d)} \frac{1}{x}\, dx = \ln|x|\Big|_{x=g(c)}^{x=g(d)} = \ln|g(y)|\Big|_{y=c}^{y=d}.$$

♦

Beispiel Es gilt für $c, d \in \left]-\frac{\pi}{2}, +\frac{\pi}{2}\right[$

$$\int_c^d \tan(y)\, dy = -\int_c^d \frac{-\sin(y)}{\cos(y)}\, dy = -(\ln|\cos(y)|)\Big|_c^d.$$

4 Analytische Funktionen

4.1 Taylor–Approximation

Definition Es sei $J \subseteq \mathbb{R}$ ein Intervall.

(1) Für $n \in \mathbb{N}_0$ heißt eine Funktion $f : J \to \mathbb{R}$ heißt *n–mal differenzierbar*, wenn — der Reihe nach — für $k = 0, \ldots, n$ die Ableitungsfunktion

$$f^{(0)} := f \qquad f^{(k)} := (f^{(k-1)})'$$

existiert. (An Intervallenden benutzen wir die einseitige Ableitung).

(2) Die Funktion f heißt *n–mal stetig differenzierbar*, wenn zusätzlich die n–te Ableitung stetig ist.

(3) Die Funktion f heißt *∞–oft (stetig) differenzierbar*, wenn sie n–mal (stetig) differenzierbar ist für jedes $n \in \mathbb{N}_0$.

Beispiele Vorbemerkung: Die Alternativdefinition 2(i) der Differenzierbarkeit zeigt, dass für eine Funktion f mit $f(0) = 0$ in 0 die Ableitung 0 existiert, falls

$$|f(x)| = |f(x) - 0 - 0 \cdot (x - 0)| \leq x^2 = x \cdot x$$

(Setze $\delta = \varepsilon$).

(1) Die Funktion

$$x \mapsto \operatorname{sgn}(x) \cdot \frac{x^2}{2}$$

ist einmal differenzierbar. Die Ableitung $x \mapsto |x|$ ist stetig, aber nicht differenzierbar.

(2) Die Funktion

$$x \mapsto \begin{cases} x^2 \cdot \sin(\frac{1}{x}), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0. \end{cases}$$

ist differenzierbar. Die Ableitung

$$x \mapsto \begin{cases} 2x \cdot \sin(\frac{1}{x}) - \cos(\frac{1}{x}), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0. \end{cases}$$

ist aber nicht stetig in 0.

Definition Es sei $a \in J \subseteq \mathbb{R}$ eine Stelle im Intervall J.

(1) Existiert für eine Funktion $f : J \to \mathbb{R}$ die n–te Ableitung an der Stelle a, so kann man ihr das *Taylorpolynom n–ten Grades mit Entwicklungsstelle a*

$$\sum_{k=0}^{n} f^{(k)}(a) \, \frac{(x-a)^k}{k!} =$$
$$f(a) + f'(a) \, \frac{(x-a)}{1} + f'(a) \, \frac{(x-a)}{1} + f''(a) \, \frac{(x-a)^2}{2} + \ldots + f^{(n)}(a) \, \frac{(x-a)^n}{n!}$$

zuordnen.

(2) Die Differenz $r_{n+1}(x)$ zwischen Funktion und zugeordnetem Taylorpolynom n–ten Grades wird als *(Taylor–)Restglied $(n+1)$–ter Ordnung* bezeichnet:

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

Es gilt offenbar:

$$r_n(x) - r_{n+1}(x) = f^{(n)}(a)\,\frac{(x-a)^n}{n!}.$$

Beispiel Prüfe selbst nach: Ist f ein Polynom m–ten Grades

$$f(x) = p_0 + p_1 x + p_2 x^2 + \ldots + p_m x^m,$$

so enthält das zugeordnete Taylorpolynom n–ten Grades, $n \le m$, mit Entwicklungsstelle $a = 0$ genau die Glieder bis zum Grad n. Demzufolge enthält das Restglied $(n+1)$–ter Ordnung genau die Glieder ab Ordnung $n+1$.

Satz 35 (Taylor–Formel) *Die auf einem Intervall definierte Funktion $f : J \to \mathbb{R}$ sei $(n+1)$–mal stetig differenzierbar. Es sei $a \in J$ eine Entwicklungsstelle.*

(1) Für das Restglied r_{n+1} gilt:

$$r_{n+1}(x) = \int_a^x f^{(n+1)}(y)\,\frac{(x-y)^n}{n!}\,dy.$$

(2) Zu jedem $x \in J$ existiert ein ξ zwischen a und x, so dass

$$r_{n+1}(x) = f^{(n+1)}(\xi)\,\frac{(x-a)^{n+1}}{(n+1)!}.$$

Man nennt diesen Ausdruck das Lagrange–Restglied.

Beweis (1) Für $n = 0$ ist dies einfach der HDI. Um den Induktionsschritt $n \to n+1$ einzusehen, berechnen wir die Differenz der beiden Ausdrücke für die Restglieder $(n+2)$–ter und $(n+1)$–ter Ordnung:

$$\begin{aligned}
r_{n+1}(x) - r_{n+2}(x) &= \int_a^x f^{(n+1)}(y)\frac{(x-y)^n}{n!}\,dy - \int_a^x f^{(n+2)}(y)\frac{(x-y)^{n+1}}{(n+1)!}\,dy \\
&= \int_a^x \underbrace{f^{(n+1)}(y)}_{F(y)}\,\underbrace{\frac{(x-y)^n}{n!}}_{G'(y)}\,dy + \int_a^x \underbrace{f^{(n+2)}(y)}_{F'(y)}\,\underbrace{\frac{-(x-y)^{n+1}}{(n+1)!}}_{G(y)}\,dy \\
&= \Big[\underbrace{f^{(n+1)}(y)}_{F(y)} \cdot \underbrace{\frac{-(x-y)^{n+1}}{(n+1)!}}_{G(y)}\Big]_{y=a}^{y=x} = f^{(n+1)}(a) \cdot \frac{(x-a)^{n+1}}{(n+1)!}.
\end{aligned}$$

Ist also der Ausdruck für das $(n+1)$–te Restglied richtig, so ist er auch für das $(n+2)$–te richtig.

(2) Gemäß Mittelwertsatz der Integralrechnung existiert zwischen a und x ein ξ, so dass

$$\begin{aligned} r_{n+1}(x) &= \int_a^x f^{(n+1)}(y)\,\frac{(x-y)^n}{n!}\,dy = f^{(n+1)}(\xi)\int_a^x \frac{(x-y)^n}{n!}\,dy \\ &= f^{(n+1)}(\xi)\left[\frac{-(x-y)^{n+1}}{(n+1)!}\right]_{y=a}^{y=x} = f^{(n+1)}(\xi)\,\frac{(x-a)^{n+1}}{(n+1)!}. \end{aligned}$$

♦

Satz 36 (Taylor Approximationssatz) *Es sei $f : J \to \mathbb{R}$ n–mal stetig differenzierbar. Dann gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass*

$$\underbrace{\left| f(x) - \sum_{k=0}^{n} f^{(k)}(a)\,\frac{(x-a)^k}{k!} \right|}_{=r_{n+1}(x)} \le \varepsilon\,|x-a|^n, \qquad \textit{falls } |x-a| \le \delta.$$

Beweis Wir verwenden das Lagrange–Restglied. Es ist dann für ein geeignetes ξ zwischen a und x

$$|r_{n+1}(x)| = |r_n(x) - f^{(n)}(a)\,\frac{(x-a)^n}{n!}| = |f^{(n)}(\xi) - f^{(n)}(a)| \cdot \left|\frac{(x-a)^n}{n!}\right|.$$

Wegen der Stetigkeit von $f^{(n)}$ kann man den ersten Faktor rechts kleiner als ein vorgegebenes $\varepsilon > 0$ machen, wenn nur $|x-a|$, und damit $|\xi - a|$, kleiner als ein geeignetes $\delta > 0$ gemacht wird. ♦

4.2 Gleichmäßige Konvergenz von Funktionenfolgen

Es sei in diesem Abschnitt $(f_n)_{n\in\mathbb{N}}$ eine Folge von reellen Funktionen $f_n : D \to \mathbb{R}$. Es stellt sich die Frage nach der Konvergenz. Eine weniger wichtige Definition ist zunächst:

Definition Die Funktionenfolge (f_n) heißt *punktweise konvergent* in einem festen $x \in D$, wenn die Folge $(f_n(x))_{n\in\mathbb{N}}$ konvergent in $\mathbb{R}$ ist. Sie heißt *punktweise konvergent* (schlechthin), wenn dies für alle $x \in D$ zutrifft. Es kann dann die *Grenzfunktion*

$$f(x) \ := \ \lim_{n\to\infty} f_n(x)$$

mit Definitionsmenge D gebildet werden.

Beispiele (1) Es sei $D =]-1,+1]$ und $f_n(x) \ := \ x^n$. Es gilt

$$\lim_{n\to\infty} x^n = \begin{cases} 0, & \text{falls } -1 < x < +1, \\ 1, & \text{falls } x = +1. \end{cases}$$

Die Grenzfunktion $f :]-1,+1] \to \mathbb{R}$ ist nicht stetig in $x = 1$, obwohl dies für alle Funktionen der Folge zutrifft.

(2) Für die auf $[0,1]$ definierten Treppenfunktionen

$$\varphi_n(x) \ := \ \begin{cases} n, & \text{falls }]0, \frac{1}{n}[, \\ 0, & \text{sonst,} \end{cases}$$

gilt $\int_0^1 \varphi_n(x)\,dx = 1$. Für jedes feste $x \in [0,1]$ gilt aber

$$\varphi(x) \ := \ \lim_{n\to\infty} \varphi_n(x) = 0$$

und daher $\int_0^1 \varphi(x)\,dx = 0$.

(3) Die Funktionen

$$f_n(x) = |x|^{1+\frac{1}{n}}, \qquad n \in \mathbb{N},$$

sind alle differenzierbar. Für $n = 1$ ergibt sich als Graph die Normalparabel, für $n \to \infty$ konvergiert die Funktionenfolge punktweise gegen die nicht differenzierbare Betragsfunktion.

Wie kann man gewährleisten, dass sich bestimmte Eigenschaften der Funktionen f_n in der Folge auf die Grenzfunktion f übertragen?

Definition
(1) Ist $f : D \to \mathbb{R}$ eine beschränkte Funktion, so heißt

$$\|f\| \ := \ \sup_{x\in D} |f(x)|$$

die Supremumsnorm von f. (Vgl. LAL01: Das ist eine Norm auf dem Vektorraum der beschränkten Funktionen.)

(2) Die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ heißt *gleichmäßig konvergent (gegen die Grenzfunktion f)*, wenn die Zahlenfolge $(\|f_n - f\|)_{n\in\mathbb{N}}$ eine Nullfolge ist:

$$\lim_{n\to\infty} \|f_n - f\| = 0.$$

Bildlich: Zu jedem ε–Streifen um den Graphen von f muss es ein $N \in \mathbb{N}$ geben, so dass für alle $n \geq N$ die Graphen von f_n in diesem Streifen drin liegen müssen.

Man kann dies auch ohne Bezugnahme auf die Supremumsnorm formulieren:
Die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ heißt gleichmäßig konvergent gegen die Grenzfunktion f, es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n \geq N$ gilt:

$$|f_n(x) - f(x)| \leq \varepsilon \qquad \text{für alle } x \in D.$$

Wesentlich ist, dass das N unabhängig von $x \in D$ gefunden werden muss.

Satz 37 *Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge stetiger Funktionen, die gleichmäßig gegen die Grenzfunktion f konvergiert.*
(i) Die Grenzfunktion f ist stetig.

(ii) Ist $[c,d] \subseteq D$, so gilt für die zugehörigen Integrale:

$$\int_c^d f(x)\,dx = \lim_{n\to\infty} \int_c^d f_n(x)\,dx.$$

(Integration und Limesbildung sind vertauschbar.)

(iii) Für jedes $n \in \mathbb{N}$ sei F_n eine Stammfunktionen von f_n und F eine Stammfunktion von f. Konvergiert die Folge der Stammfunktionen an (mindestens) einer Stelle $a \in [c,d]$ gegen die Stammfunktion der Grenzfunktion,

$$\lim_{n\to\infty} F_n(a) = F(a),$$

so konvergiert die Folge der Stammfunktionen F_n gleichmäßig gegen F.

(iv) Aus der Differenzierbarkeit der Funktionen f_n kann — auch bei gleichmäßiger Konvergenz von (f_n) — ***nicht*** *auf die Differenzierbarkeit der Grenzfunktion f geschlossen werden.*

Beweis (i) Es sei $x \in D$ und $\varepsilon > 0$ vorgegeben. Wir wollen zeigen, dass es ein $\delta > 0$ gibt, so dass

$$|f(y) - f(x)| \leq \varepsilon, \qquad \text{falls } |y - x| \leq \delta.$$

Da (f_n) gleichmäßig konvergiert, gibt es zu $\frac{\varepsilon}{3}$ ein $N \in \mathbb{N}$, so dass

$$|f_n(z) - f(z)| \leq \frac{\varepsilon}{3}, \qquad \text{für alle } z \in D, \quad \text{falls } n \geq N.$$

Insbesondere ist

$$|f_N(z) - f(z)| \leq \frac{\varepsilon}{3} \qquad \text{für alle } z \in D. \qquad (*)$$

Da f_N stetig ist, gibt es zu $\frac{\varepsilon}{3}$ ein $\delta > 0$, so dass

$$|f_N(y) - f_N(x)| \leq \frac{\varepsilon}{3}, \qquad \text{falls } |y - x| \leq \delta. \qquad (**)$$

Insgesamt gilt für alle $y \in D$ mit $|y - x| \leq \delta$

$$|f(y) - f(x)| \leq \underbrace{|f(y) - f_N(y)|}_{(*)\,\leq\frac{\varepsilon}{3}} + \underbrace{|f_N(y) - f_N(x)|}_{(**)\,\leq\frac{\varepsilon}{3}} + \underbrace{|f_N(x) - f(x)|}_{(*)\,\leq\frac{\varepsilon}{3}} \leq \varepsilon.$$

(ii) Es sei $\varepsilon > 0$ vorgegeben. Für ein festes $n \in \mathbb{N}$ gilt gemäß Satz 28 und Satz 30 (2)

$$\begin{aligned}\Big| \int_c^d f(x)\,dx - \int_c^d f_n(x)\,dx \Big| &= \Big| \int_c^d [f(x) - f_n(x)]\,dx \Big| \le \int_c^d |f(x) - f_n(x)|\,dx \\ &\le (d-c) \cdot \sup_{x \in [c,d]} \{|f(x) - f_n(x)|\} \le \varepsilon,\end{aligned}$$

da aufgrund der gleichmäßigen Konvergenz das Supremum kleiner als $\frac{\varepsilon}{(d-c)}$ gemacht werden kann.

(iii) Für die Stammfunktionen F_n und F gilt aufgrund des HDI:

$$F_n(x) = F_n(a) + \int_a^x f_n(y)\,dy, \qquad F(x) = F(a) + \int_a^x f(y)\,dy.$$

Es ist dann

$$\begin{aligned}\Big| F(x) - F_n(x) \Big| &= \Big| F(a) + \int_a^x f(y)\,dy - F_n(a) - \int_a^x f_n(y)\,dy \Big| = \\ &\le \Big| F(a) - F_n(a) \Big| + \Big| \int_a^x [f(y) - f_n(y)]\,dy \Big|.\end{aligned}$$

Zu einem vorgegebenen $\varepsilon > 0$ kann — gemäß Voraussetzung des Satzes und Aussage (2) — ein $N \in \mathbb{N}_0$ gefunden werden kann, so dass für alle $n \ge N$ gilt

$$\Big| \int_a^x [f(y) - f_n(y)]\,dy \Big| \le \frac{\varepsilon}{2}, \qquad \Big| F(a) - F_n(a) \Big| \le \frac{\varepsilon}{2}$$

gilt und damit für alle $x \in [c, d]$

$$\Big| F(x) - F_n(x) \Big| \le \varepsilon.$$

(iv) Wir zeigen, dass die Funktionenfolge $f_n(x) = |x|^{1+\frac{1}{n}}$ aus Beispiel (4) sogar gleichmäßig gegen die Betragsfunktion konvergiert. Zunächst ist

$$\lim_{x \to \infty} \frac{1}{x} = 0 \quad \Longrightarrow \quad \lim_{x \to \infty} \frac{1}{x} \cdot \ln\Big(\frac{1}{x}\Big) = -\lim_{x \to \infty} \frac{\ln x}{x} \overset{\text{l'H}}{=} 0.$$

Daraus folgt für die Funktion

$$\lim_{x \searrow 0} (x \cdot \ln x) = \lim_{x \to \infty} \frac{1}{x} \cdot \ln\Big(\frac{1}{x}\Big) = 0,$$

so dass die Funktion $f(x) = x \cdot \ln x$ auf dem Intervall $]0, 1]$ ein Betragsmaximum hat:

$$M := \max_{x \in]0,1]} \{|x \cdot \ln x|\}.$$

Für ein festes $x \in]0, 1]$ und $\alpha \in [1, 2]$ gibt es aufgrund des Mittelwertsatzes ein $\xi \in [1, \alpha] \subseteq [1, 2]$, so dass

$$\Big| x - x^\alpha \Big| \le \frac{dx^\alpha}{d\alpha}\Big|_{\alpha = \xi} \cdot |1 - \alpha| = |\ln x \cdot x^\xi| \cdot |1 - \alpha| \le |\ln x \cdot x| \cdot |1 - \alpha| \le M \cdot |1 - \alpha|.$$

Daraus folgt, dass für alle $x \in [-1, +1]$ gilt:

$$\Big| |x| - |x|^\alpha \Big| \le M \cdot |1 - \alpha|.$$

Das aber bedeutet, dass für $\alpha = 1 + \frac{1}{n} \to 1$ die Funktionenfolge gleichmäßig konvergiert.

♦

Satz 38 (Gleichmäßige Konvergenz von Funktionenreihen, Weierstrass)
Es seien $g_k : D \to \mathbb{R}$ die Glieder einer Folge $(g_k)_{k\in\mathbb{N}_0}$ von reellen Funktionen und

$$f_n := \sum_{k=0}^{n} g_k$$

die zugehörige Partialsummenfolge. Wenn die Zahlenreihe

$$\sum_{k=0}^{\infty} \|g_k\|$$

konvergiert, dann konvergiert die Partialsummenfolge gleichmäßig (und absolut). Man sagt, die Funktionenreihe konvergiert gleichmäßig absolut.

Beweis Für jedes feste $x \in D$ gilt

$$|g_k(x)| \leq \sup_{y\in D}\{|g_k(y)|\} = \|g_k\|.$$

Daher stellt die Zahlenreihe $\sum_{k=0}^{\infty} \|g_k\|$ eine konvergent Reihe für die Reihe

$$\sum_{k=0}^{\infty} g_k(x)$$

dar, die also nach dem Majorantenkriterium (ANA01, Satz 32) absolut konvergiert. Wir definieren die Funktion f durch diese Grenzwerte

$$f(x) := \lim_{n\to\infty} f_n(x) = \sum_{k=0}^{\infty} g_k(x)$$

und zeigen, dass die Funktionenfolge f_n gleichmäßig gegen f konvergiert. Zu einem vorgegebenen $\varepsilon > 0$ gibt es nach Voraussetzung des Satzes ein $N \in \mathbb{N}$, so dass für alle $n \geq N$

$$\sum_{k=n}^{\infty} \|g_k\| \leq \sum_{k=N}^{\infty} \|g_k\| < \varepsilon.$$

Für alle $n \geq N$ und alle $x \in D$ folgern wir weiter (mit ANA01, Satz 31)

$$\Big|f(x) - f_n(x)\Big| = \Big|\sum_{k=n+1}^{\infty} g_k(x)\Big| \leq \sum_{k=n+1}^{\infty} |g_k(x)| \leq \sum_{k=n+1}^{\infty} \|g_k\| < \varepsilon.$$

♦

4.3 Potenzreihen

Definition Ist $(p_k)_{k\in\mathbb{N}_0}$ eine Folge reeller Zahlen und $a \in \mathbb{R}$, so heißt der Ausdruck

$$\sum_{k=0}^{\infty} p_k(x-a)^k$$

(formale) Potenzreihe mit der Koeffizientenfolge $(p_k)_{k\in\mathbb{N}_0}$ *und Entwicklungsstelle* a.

Die beiden Begriffe „Ausdruck" und „formal" sollen zum Ausdruck bringen, dass die Frage der Konvergenz noch nicht geklärt ist und deshalb noch nicht von einer Funktion (mit Definitionsvariabler x) gesprochen werden kann. Erst, wenn für x eine konkrete reelle Zahl eingesetzt wird, entscheidet sich die Konvergenz.

Potenzreihen können addiert und mit einer reellen Zahl (Skalar) vervielfacht werden, indem man diese Operation mit den Koeffizientenfolgen durchführt.

$$\sum_{k=0}^{\infty} p_k\, x^k + \sum_{k=0}^{\infty} q_k\, x^k = \sum_{k=0}^{\infty} (p_k + q_k)\, x^k, \qquad \alpha \cdot \sum_{k=0}^{\infty} p_k\, x^k = \sum_{k=0}^{\infty} (\alpha \cdot p_k)\, x^k.$$

Beispiele

(1) Jedes Polynom

$$p_0 + p_1 x + p_2 x^2 + \ldots + p_n x^n$$

stellt eine Potenzreihe (mit $p_k = 0$ für $k \geq n+1$) dar. Offenbar „konvergiert" sie für jedes $x \in \mathbb{R}$.

(2) Die Exponentialreihe

$$\sum_{k=0}^{\infty} \frac{1}{k!}\, x^k$$

ist eine Potenzreihe mit $p_k = \frac{1}{k!}$. Wir wissen bereits, dass sie für jedes $x \in \mathbb{R}$ konvergiert.

(3) Kosinus und Sinus waren durch die Potenzreihen

$$\sum_{k=0}^{\infty} p_k x^k \qquad \text{mit} \qquad p_k = \begin{cases} \frac{(-1)^\ell}{(2\ell)!}, & \text{falls } k = 2\ell \text{ gerade,} \\ 0, & \text{sonst,} \end{cases}$$

bzw.

$$\sum_{k=0}^{\infty} p_k x^k \qquad \text{mit} \qquad p_k = \begin{cases} \frac{(-1)^\ell}{(2\ell+1)!}, & \text{falls } k = 2\ell+1 \text{ ungerade,} \\ 0, & \text{sonst,} \end{cases}$$

definiert. Sie konvergieren ebenfalls für jedes $x \in \mathbb{R}$.

(4) Setzt man $p_k = 1$ für alle $k \in \mathbb{N}_0$, so entsteht die geometrische Reihe

$$\sum_{k=0}^{\infty} x^k.$$

Wir wissen bereits, dass sie für jedes $x \in \mathbb{R}$ mit $|x| < 1$ konvergiert und für $|x| \geq 1$ divergiert.

(5) Die Potenzreihe

$$\sum_{k=0}^{\infty} k^k x^k$$

konvergiert nur für $x = 0$. Für jedes $x \neq 0$ gibt es nämlich ein $N \in \mathbb{N}$, so dass $N \cdot |x| \geq 2$. Dann ist für $k \geq N$

$$|k^k x^k| = |k \cdot x|^k \geq |N \cdot x|^k \geq 2^k.$$

Also ist die Reihe $\sum_{k=0}^{\infty} 2^k$ eine „divergente Minorante".

(Bei all diesen Beispielen war die Entwicklungsstelle $a = 0$. Das ist auch der wesentliche Fall.)

Eine erste wesentliche Beobachtung über Potenzreihen ist in dem folgenden Lemma enthalten:

Lemma 39

(i) Jede Potenzreihe konvergiert an ihrer Entwicklungsstelle.

(ii) Konvergiert die Potenzreihe $\sum_{k=0}^{\infty} p_k x^k$ für ein festes $y \in \mathbb{R}$, so konvergiert sie auch (absolut) für alle $x \in \mathbb{R}$ mit $|x| < |y|$.

(iii) Konvergiert die Potenzreihe $\sum_{k=0}^{\infty} p_k (x-a)^k$ für ein festes $y \in \mathbb{R}$, so konvergiert sie auch (absolut) für alle $x \in \mathbb{R}$ mit $|x - a| < |y - a|$.

Beweis (i) muss nicht bewiesen werden.

(ii) Da die Reihe $\sum_{k=0}^{\infty} p_k\, y^k$ konvergiert, ist $(p_k\, y^k)_{k \in \mathbb{N}_0}$ eine Nullfolge und daher beschränkt:

$$|p_k\, y^k| \leq M \quad \text{für alle } k \in \mathbb{N}_0.$$

Für alle x mit $|x| < |y|$ gilt dann

$$|p_k\, x^k| = |p_k\, y^k| \cdot \left|\frac{x}{y}\right|^k \leq M \cdot \vartheta^k, \qquad \text{wobei } \vartheta := \left|\frac{x}{y}\right| < 1.$$

Die Reihe $\sum_{k=0}^{\infty} M \cdot \vartheta^k$ ist eine geometrische (absolut konvergente) Reihe und stellt daher eine Majorante für die Reihe $\sum_{k=0}^{\infty} p_k\, x^k$ dar.

Für den Beweis von (iii) muss man in (ii) nur x durch $x - a$ und y durch $y - a$ ersetzen. ♦

Aus dem Lemma folgt unmittelbar, dass die Menge der $x \in \mathbb{R}$, für die eine Potenzreihe konvergiert, ein Intervall ist, dessen obere und untere Grenze symmetrisch bzgl. a liegen. Beachte, dass das Lemma keine Aussage darüber macht, ob die Potenzreihe an den beiden Grenzen selbst konvergiert.

Definition (1) Für eine gegebene Potenzreihe $\sum_{k=0}^{\infty} p_k (x-a)^k$ heißt die Menge

$$\mathcal{K} := \left\{ x \in \mathbb{R} \;\middle|\; \sum_{k=0}^{\infty} p_k\, (x-a)^k \text{ konvergiert} \right\}$$

das *Konvergenzintervall (der Potenzreihe).*

(2) Die Zahl ϱ mit $0 \leq \varrho \leq +\infty$, definiert durch

$$\inf \mathcal{K} = a - \varrho, \qquad \sup \mathcal{K} = a + \varrho,$$

heißt *Konvergenzradius (der Potenzreihe).*

Bemerkung Die Theorie der Potenzreihen lässt sich auch für $z \in \mathbb{C}$ anstelle $x \in \mathbb{R}$ aufbauen. Die nahezu gleichen Überlegungen zeigen dabei, dass anstelle eines Konvergenzintervalls eine Konvergenzkreisscheibe mit Mittelpunkt $a \in \mathbb{C}$ in Erscheinung tritt. Deshalb der Name Konvergenz**radius**.

Beispiele

(1) Polynome, die Exponentialfunktion und die beiden Funktionen sin und cos haben den Konvergenzradius $\varrho = +\infty$.

(2) Ist r eine beliebige positive Zahl, so konvergiert die Potenzreihe

$$\sum_{k=0}^{\infty} \left(\frac{x}{r}\right)^k$$

genau dann, wenn $\left|\frac{x}{r}\right| < 1 \iff |x| < r$. Der Konvergenzradius ist also $\varrho = r$.

(3) Wir betrachten die Potenzreihe

$$\sum_{k=1}^{\infty} \frac{1}{k} x^k$$

Sie wird für $x < 1$ von der geometrischen Reihe majorisiert. Wir wissen, dass sie für $x = 1$ divergiert (harmonische Reihe) und für $x = -1$ konvergiert (Leibniz–Kriterium). Es ist also $\varrho = 1$.

Präposition 40

(1) Für $0 < b < 1$ und $n \in \mathbb{N}_0$ gilt:

$$\lim_{x \to \infty} x^n \cdot b^x = 0.$$

(2) Die beiden Potenzreihen

$$\sum_{k=0}^{\infty} p_k (x-a)^k \qquad \textit{und} \qquad \sum_{k=0}^{\infty} k p_k (x-a)^k$$

haben den gleichen Konvergenzradius ϱ.

Beweis (1) Per Induktion: Es ist $(b^x)' = \ln b \cdot b^x < 0$, deshalb ist die Funktion $x \mapsto b^x$ monoton fallend. Für die geometrische Folge gilt $\lim_{n\to\infty} b^n = 0$, also ist der Induktionsanfang gesetzt: Es gilt auch $\lim_{x\to\infty} b^x = 0$.

Wir nehmen jetzt an, die Aussage sei für $n-1$ gezeigt. Es sei $f(x) = x^n$ und $g(x) = b^{-x}$. Es gilt dann

$$\frac{f'(x)}{g'(x)} = \frac{n \cdot x^{n-1}}{-\ln b \cdot b^{-x}} = -\frac{n}{\ln b} \cdot x^{n-1} \cdot b^x \overset{x\to\infty}{\longrightarrow} 0$$

und dann mit l'Hospital

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} x^n \cdot b^x = 0.$$

(2) Es genügt, den Fall $a = 0$ zu betrachten. Für ein beliebiges x ist

$$|p_k x^k| \le |k p_k\, x^k|,$$

so dass aus der absoluten Konvergenz der rechten Reihe mit dem Majorantenkriterium auf die Konvergenz der linken Reihe geschlossen werden kann. Deshalb ist der Konvergenzradius ϱ_r der rechten Potenzreihe in (2) kleiner oder gleich dem Konvergenzradius ϱ_ℓ der linken Potenzreihe.

Es bleibt $\varrho_r \ge \varrho_\ell$ zu zeigen. Dazu sei y mit $0 < |y| < \varrho_\ell$ vorgegeben. Es existiert ein r mit $|y| < r < \varrho_\ell$, so dass die Reihe $\sum_{k=0}^\infty p_k\, r^k$ absolut konvergiert. Nach (1) (mit $n = 1$) ist die Folge

$$k \cdot \left|\frac{y}{r}\right|^k$$

eine Nullfolge, also durch M beschränkt. Dann gilt weiter

$$\Big|k p_k y^k\Big| = \Big|k \cdot (\frac{y}{r})^k \cdot p_k\, r^k\Big| \le M \Big|p_k\, r^k\Big|.$$

Damit haben wir eine absolut konvergente Majorante für die Reihe

$$\sum_{k=1}^\infty k p_k\, y^k$$

gefunden, sie konvergiert also absolut. Insgesamt haben wir bewiesen, dass

$$\varrho_r \ge y \qquad \text{für alle } y \text{ mit } y < \varrho_\ell$$

gilt. Das heißt aber gerade $\varrho_r \ge \varrho_\ell$.

♦

Satz 41 (Gleichmäßige Konvergenz einer Potenzreihe)
Ist $r < \varrho$, so konvergiert die Potenzreihe $\sum_{k=0}^\infty p_k\,(x-a)^k$ gleichmäßig absolut auf $[a-r, a+r]$.

Beweis (Nur für $a = 0$) Für alle x mit $|x| \le r$ gilt

$$|p_k\, x^k| \le |p_k r^k|$$

und daher

$$\|p_k\, x^k\| := \sup_{x\in[-r,+r]} \{|p_k\, x^k|\} \le |p_k r^k|.$$

Damit stellt die Reihe $\sum_{k=0}^\infty p_k r^k$ eine konvergente Majorante für die Reihe

$$\sum_{k=0}^\infty \|p_k\, x^k\|$$

dar. Nach dem Satz 38 von Weierstrass konvergiert die Potenzreihe $\sum_{k=0}^\infty p_k x^k$ gleichmäßig auf $[-r, +r]$. ♦

Satz 42 (Eigenschaften der Grenzfunktion einer Potenzreihe)
Es sei $\sum_{k=0}^{\infty} p_k(x-a)^k$ eine Potenzreihe mit Konvergenzradius ϱ und

$$f : \begin{cases}]a-\varrho, a+\varrho[& \to \mathbb{R} \\ x & \mapsto \sum_{k=0}^{\infty} p_k(x-a)^k \end{cases}$$

die Grenzfunktion.

(i) f ist auf $]a-\varrho, a+\varrho[$ stetig.

(ii) f ist auf $[c,d] \subseteq\,]a-\varrho, a+\varrho[$ integrierbar. Das Integral darf gliedweise berechnet werden.

$$\int_c^d \sum_{k=0}^{\infty} p_k(x-a)^k\,dx = \sum_{k=0}^{\infty} \int_c^d p_k(x-a)^k\,dx = \sum_{k=0}^{\infty} \left[\frac{p_k}{k+1}\,(x-a)^{k+1}\right]_{x=c}^{x=d}.$$

(iii) f besitzt eine Stammfunktion F auf $]-\varrho, +\varrho[$. Die durch $F(a) = 0$ festgelegte Stammfunktion ist gegeben durch die Potenzreihe

$$F(x) = \sum_{k=0}^{\infty} \frac{p_k}{k+1}\,(x-a)^{k+1},$$

(„Gliedweise Stammfunktion"). Ihr Konvergenzradius ist ϱ.

(iv) f ist auf $]a-\varrho, a+\varrho[$ unendlich oft differenzierbar. Dabei gilt

$$\begin{aligned} f'(x) &= \sum_{k=1}^{\infty} kp_k\,(x-a)^{k-1} = \sum_{k=0}^{\infty}(k+1)a_{k+1}\,(x-a)^k \\ f^{(n)}(x) &= \sum_{k=n}^{\infty} \frac{k!}{(k-n)!} p_k\,(x-a)^{k-n}, \qquad \textit{insbesondere } f^{(n)}(a) = n!\,a_n. \end{aligned}$$

(„Gliedweise Ableitung"). Die Potenzreihen haben den Konvergenzradius ϱ.

Beweis (Generell nur für $a = 0$). Für $r < \varrho$ ist die Funktionen–Partialsummenfolge

$$f_n(x) := \sum_{k=0}^{n} p_k\,x^k$$

gleichmäßig konvergent auf $[-r, +r]$. Damit sind die Aussagen des Satzes 37 (i) – (iii) anwendbar. Die Grenzfunktion hat also die Eigenschaften (i) – (iii) auf $[-r, +r]$. Da dies für alle $r < \varrho$ so ist, sind die Eigenschaften auch auf

$$]-\varrho, +\varrho[\;=\; \bigcup_{n \geq N} [-\varrho + \frac{1}{n}, +\varrho - \frac{1}{n}], \qquad (N \text{ groß genug})$$

erfüllt.

Die Präposition zeigt, dass die in der ersten Zeile von (iv) angegebene Potenzreihe den gleichen Konvergenzradius ϱ hat wie die ursprüngliche Potenzreihe und daher gleichmäßig auf $[-r, +r]$ für jedes $r < \varrho$ konvergiert. Das aber bedeutet gemäß Satz 37 (iii), dass die Grenzfunktion der ursprünglichen Potenzreihe auf $]-\varrho, +\varrho[$ eine Stammfunktion ist.

Die anderen Aussagen für $n \geq 2$ werden einfach mit Induktion bewiesen. ♦

4.4 Zwei Aussagen über den Konvergenzradius

Ist eine Folge $(a_k)_{k\in\mathbb{N}}$ gegeben, so können wir weitere Folgen bilden:

- $\sigma_\ell := \sup\limits_{k\geq\ell}\{a_k\}$ (Folge der „Rest–Suprema")
- $\iota_\ell := \inf\limits_{k\geq\ell}\{a_k\}$ (Folge der „Rest–Infima")

Es ist klar, dass $(\sigma_\ell)_{\ell\in\mathbb{N}}$ eine monoton fallende, $(\iota_\ell)_{\ell\in\mathbb{N}}$ eine monoton wachsende Folge mit Werten in $\overline{\mathbb{R}} = [-\infty, +\infty]$ ist.

Daraus folgt, dass jede der beiden Folgen einen (eigentlichen oder uneigentlichen) Grenzwert hat. Für sie gibt es eine besondere Bezeichnung:

Definition Die beiden Grenzwerte heißen *Limes superior* bzw. *Limes inferior.*

$$\begin{aligned} \limsup_{k\to\infty} a_k &:= \lim_{\ell\to\infty} \sigma_\ell = \lim_{\ell\to\infty} \sup_{k\geq\ell}\{a_k\} \in [-\infty, +\infty], \\ \liminf_{k\to\infty} a_k &:= \lim_{\ell\to\infty} \iota_\ell = \lim_{\ell\to\infty} \inf_{k\geq\ell}\{a_k\} \in [-\infty, +\infty]. \end{aligned}$$

Wir halten einige Eigenschaften dieser Grenzwerte in einem Satz fest:

Satz 43
(1) Der Limes superior S einer Folge ist durch die folgende Eigenschaft eindeutig festgelegt: Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass

$$\begin{aligned} \alpha_k &\leq S + \varepsilon \quad \textit{für alle } k \geq N, \\ \alpha_k &\geq S - \varepsilon \quad \textit{für unendlich viele } k \geq N. \end{aligned}$$

(2) Der Limes inferior I einer Folge ist durch die folgende Eigenschaft eindeutig festgelegt: Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass

$$\begin{aligned} \alpha_k &\geq I - \varepsilon \quad \textit{für alle } k \geq N, \\ \alpha_k &\leq I + \varepsilon \quad \textit{für unendlich viele } k \geq N. \end{aligned}$$

(3) Hat die Folge $(a_k)_{k\in\mathbb{N}}$ selbst einen eigentlichen oder uneigentlichen Grenzwert, so gilt

$$\liminf_{k\to\infty} a_k = \lim_{k\to\infty} a_k = \limsup_{k\to\infty} a_k.$$

Beispiel Es sei $(a_\ell)_{\ell\in\mathbb{N}}$ eine konvergente Folge mit Grenzwert a. Wir schieben Nullen dazwischen:

$$\widetilde{a}_k := \begin{cases} a_\ell, & \text{falls } k = 2\,\ell \text{ gerade }, \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt

$$\limsup_{k\to\infty} \widetilde{a}_k = \begin{cases} a, & \text{falls } a \geq 0, \\ 0, & \text{falls } a < 0, \end{cases} \qquad \liminf_{k\to\infty} \widetilde{a}_k = \begin{cases} a, & \text{falls } a \leq 0, \\ 0, & \text{falls } a > 0. \end{cases}$$

Satz 44 (Formeln für den Konvergenzradius)
Es sei $\sum_{k=0}^{\infty} p_k(x-a)^k$ *eine Potenzreihe mit Konvergenzradius* ϱ. *Es gilt*

$$\varrho \quad = \quad \frac{1}{\lim\limits_{\ell\to\infty} \sup\limits_{k\geq\ell} \sqrt[k]{|p_k|}} \qquad \text{(Formel von Cauchy–Hadamard)}$$

$$\lim_{\ell\to\infty} \inf_{k\geq\ell} \left|\frac{p_k}{p_{k+1}}\right| \quad \leq \quad \varrho \quad \leq \quad \lim_{\ell\to\infty} \sup_{k\geq\ell} \left|\frac{p_k}{p_{k+1}}\right|$$

Dabei sind die Rechenregeln $\frac{1}{0} = \infty$, $\frac{1}{\infty} = 0$ *anwendbar.*
Existiert der (eigentliche oder uneigentliche) Grenzwert auf der rechten Seite, so gilt

$$\varrho = \lim_{k\to\infty} \left|\frac{p_k}{p_{k+1}}\right|$$

Beweis Die beiden Aussagen sind Übertragungen von Wurzel– bzw. Quotientenkriterium auf Potenzreihen. Wir beweisen nur die erste, können dabei wieder $a=0$ annehmen.

(1) Wir bezeichnen den Cauchy–Hadamard–Ausdruck mit

$$R := \frac{1}{\lim\limits_{\ell\to\infty} \sup\limits_{k\geq\ell} \sqrt[k]{|p_k|}} = \lim_{\ell\to\infty} \inf_{k\geq\ell} \underbrace{\frac{1}{\sqrt[k]{|p_k|}}}_{=:\alpha_k}.$$

Gemäß der Charakterisierung (2) des Limes inferior in Satz 43 gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass

$$\frac{1}{\sqrt[k]{|p_k|}} \quad \geq \quad R - \varepsilon \qquad \text{für alle } k \geq N, \qquad (*)$$
$$\frac{1}{\sqrt[k]{|p_k|}} \quad \leq \quad R + \varepsilon \qquad \text{für unendlich viele } k \geq N. \qquad (**)$$

(2) Es sei jetzt (O.B.d.A.) $R > 0$ und $y \in \mathbb{R}$ beliebig, so dass $0 < |y| < R$. Dann gibt es ein $\varepsilon > 0$, so dass $|y| < R - \varepsilon$. Aufgrund von (*) gibt es zu diesem ε ein $N \in \mathbb{N}$, so dass

$$\sqrt[k]{|p_k \cdot y^k|} = \sqrt[k]{|p_k|} \cdot |y| \leq \frac{|y|}{R-\varepsilon} = q < 1 \qquad \text{für alle } k \geq N.$$

Gemäß Wurzel–Kriterium (ANA01, Satz 34) konvergiert die Reihe $\sum_{k=0}^{\infty} p_k y^k$ absolut. Da y eine beliebige Zahl mit $|y| < R$ war, können wir $R \leq \varrho$ folgern.

(3) Wir betrachten jetzt ein $y \in \mathbb{R}$ mit $|y| > R$ und setzen $\varepsilon = |y| - R$. Aufgrund von $(**)$ gibt es $N \in \mathbb{N}$, so dass

$$\frac{1}{\sqrt[k]{|p_k|}} \leq R + \varepsilon = |y| \qquad \text{für unendlich viele } k \geq N,$$

woraus folgt, dass

$$1 \leq |p_k y^k| \qquad \text{für unendlich viele } k \geq N.$$

Das aber bedeutet, dass die Reihe $\sum_{k=0}^{\infty} p_k y^k$ nicht konvergieren kann, es folgt:

$$R \geq \varrho.$$

(4) Die zweite Formel wird hier nicht bewiesen. ♦

4.5 Analytische Funktionen

Definition (und Bezeichnungen) Es sei $a \in \mathbb{R}$.

$\mathcal{C}_a^\infty$ sei die Menge der reellen Funktionen f, die auf einem ,,Intervall" $[a-r, a+r]$ mit $r > 0$ definiert und dort ∞–oft differenzierbar sind.

$\mathcal{P}_a^{\text{form}}$ sei die Menge der (formalen) Potenzreihen mit Entwicklungsstelle a.

$\mathcal{P}_a^{\text{konv}}$ sei die Menge der Potenzreihen mit Entwicklungsstelle a, die einen positiven Konvergenzradius haben, also in einem echten Intervall um a konvergieren.

Zu jeder Funktion in $f \in \mathcal{C}_a^\infty$ können wir die *Taylorreihe*

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

bilden. Dies ist eine formale Potenzreihe mit Entwicklungsstelle a. Über ihre Konvergenz wissen wir a priori NICHTS.

Wir kennen inzwischen zwei ,,Operatoren", die eine Beziehung zwischen ∞–oft differenzierbaren Funktionen und Potenzreihen aufbauen.

$$\begin{array}{ccc} \mathcal{C}_a^\infty & \xrightarrow{\text{Taylorreihe}} & \mathcal{P}_a^{\text{form}} \\ \mathcal{C}_a^\infty & \xleftarrow[\text{Grenzfunktion}]{} & \mathcal{P}_a^{\text{konv}} \end{array}$$

Dazu einige Erläuterungen:

- Der Operator ,,Taylorreihe" der oberen Zeile des Diagramms ist nicht injektiv. Dies zeigt das Beispiel der Funktion

  $$\eta(x) := \begin{cases} \exp(-\frac{1}{x}), & \text{falls } x > 0, \\ 0, & \text{falls } x \le 0. \end{cases}$$

  Sie ist ∞–oft differenzierbar (Beweis?), es gilt

  $$\eta^{(n)}(0) = 0$$

  für alle $n \in \mathbb{N}_0$. Insbesondere ist ihre Taylorreihe mit Entwicklungsstelle 0 gleich Null. Die Funktion η besitzt also die gleiche Taylorreihe wie die Nullfunktion, obwohl sie in jeder (beliebig kleinen) Umgebung von 0 nicht mit der Nullfunktion übereinstimmt.

- Ein Satz von Borel besagt, dass es zu jeder (formalen) Potenzreihe eine ∞–oft differenzierbare Funktion gibt, die diese Potenzreihe als Taylorreihe besitzt. Mit anderen Worten, der Taylorreihe–Operator ist surjektiv. Das heißt insbesondere, dass die Taylorreihe einer $\mathcal{C}^\infty$–Funktion den Konvergenzradius $\varrho = 0$ haben kann, also nur in der Entwicklungsstelle konvergiert. Beim Beweis des Satzes von Borel muss zu jeder gegebenen Potenzreihe eine passende $\mathcal{C}^\infty$ Funktion konstruiert werden. Dabei wird gerade die Funktion η von weiter oben benutzt.

Definition Eine $\mathcal{C}^\infty$–Funktion $f : D \to \mathbb{R}$ heißt *analytisch in a*, wenn es ein $\varepsilon > 0$ gibt, so dass

(a) $]a-\varepsilon, a+\varepsilon[\subseteq D$,

(b) die Taylorreihe von f mit Entwicklungsstelle a

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

einen Konvergenzradius $\varrho \geq \varepsilon$ besitzt und

(c) die Grenzfunktion der Taylorreihe auf $]a-\varepsilon, a+\varepsilon[$ mit f übereinstimmt.

Eine $\mathcal{C}^\infty$–Funktion $f : D \to \mathbb{R}$ heißt *analytisch (schlechthin)*, wenn sie analytisch in jedem $a \in J$ ist.

Beachte nochmals, dass die Taylorreihe für jede $\mathcal{C}^\infty$–Funktion existiert.
Die Bedingung (b) ist durch die Existenz der Taylorreihe nicht gewährleistet.
Die Bedingung (c) ist durch (b) nicht gewährleistet.

Satz 45 *Ist*

$$\sum_{k=0}^{\infty} p_k (x-a)^k$$

eine Potenzreihe mit positivem Konvergenzradius ϱ, so ist die Grenzfunktion f analytisch auf $]a-\varepsilon, a+\varepsilon[$.

Beweis Dass f in a analytisch ist, folgt aus Satz 42 (iv). Den Beweis, dass f auch in jedem anderen $x \in\,]a-\varepsilon, a+\varepsilon[$ analytisch ist, lassen wir weg. ♦

5 Topologie des $\mathbb{R}^d$

5.1 Grundlage

Im folgenden steht im Mittelpunkt die Menge

$$\mathbb{R}^d := \underbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}_{d\text{–mal}} = \{(x_1, x_2, \ldots, x_d) | x_i \in \mathbb{R} \text{ für alle } i = 1, \ldots, d\}.$$

Oft (Staatsexamen) geht es nur um die Fälle $d = 2$ oder $d = 3$. In der Linearen Algebra steht im Vordergrund, dass man die Elemente des $\mathbb{R}^d$ addieren und skalar multiplizieren kann: Der $\mathbb{R}^d$ ist ein (reeller) Vektorraum. Der Nullvektor $\vec{0} = (0, \ldots, 0)$ heißt auch *Ursprung.*

Oft — je nach Kontext — werden für die Elemente des $\mathbb{R}^d$ auch die Schreibweisen

$$x = (x_1, x_2, \ldots, x_d) \qquad \text{oder} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix}$$

benutzt. Beachte, dass in der Matrixtheorie Zeilenvektoren und Spaltenvektoren unterschieden werden.

Eine *reelle Abbildung ,,mehrerer Variabler"* ist einfach eine auf einer Teilmenge $X \subseteq \mathbb{R}^d$ definierte Abbildung

$$f : \begin{cases} X & \to \mathbb{R} \\ x = (x_1, \ldots, x_d) & \mapsto f(x) = f(x_1, \ldots, x_d). \end{cases}$$

Umgekehrt werden wir uns auch für *Kurven*

$$\gamma : \begin{cases} J & \to \mathbb{R}^d \\ t & \mapsto (\gamma_1(t), \ldots, \gamma_d(t)), \end{cases}$$

wobei J ein Intervall in $\mathbb{R}$ ist, interessieren.

5.2 Die euklidische Norm im $\mathbb{R}^d$

Definition (1) Die *euklidische Norm* im $\mathbb{R}^d$ ist definiert als die Abbildung

$$\|\cdot\| \begin{cases} \mathbb{R}^d & \to & \mathbb{R}_0^+ \\ x & \mapsto & \|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_d^2}. \end{cases}$$

(2) Wird bei mathematischen Überlegungen über den $\mathbb{R}^d$ die euklidische Norm benutzt, so spricht man auch vom *euklidischen Vektorraum* $\mathbb{R}^d$.

(3) Die Abbildung

$$d : \begin{cases} \mathbb{R}^d \times \mathbb{R}^d & \to & \mathbb{R}_0^+ \\ (x,y) & \mapsto & d(x,y) = \|x - y\| \end{cases}$$

heißt (euklidische) *Abstandsfunktion* oder *Metrik* (auf $\mathbb{R}^d$).

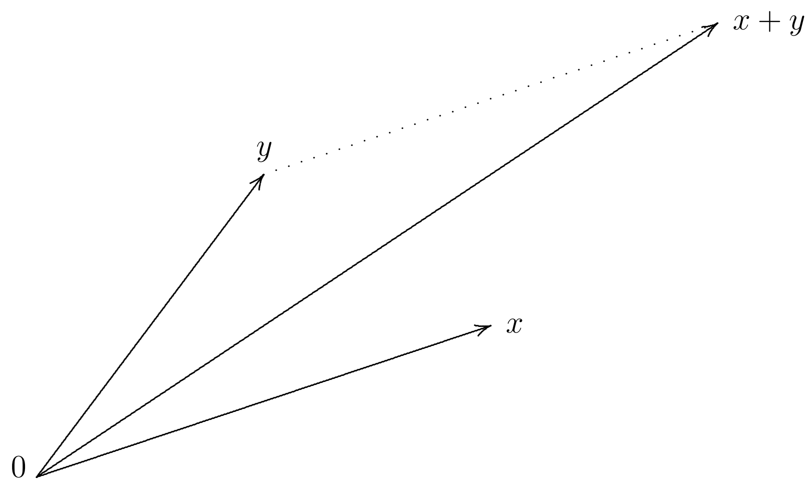
Satz 46 (Eigenschaften der euklidischen Norm) *Für $x, y \in \mathbb{R}^d$ und $\alpha \in \mathbb{R}$ sind die folgenden Eigenschaften erfüllt.*

(i) $\|x\| = 0 \quad \Longrightarrow \quad x = \vec{0}$,

(ii) $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$,

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *(Dreiecksungleichung)*

(iv) $\Big| \|x\| - \|y\| \Big| \leq \|x - y\|$



Beweis Die Aussagen (i) und (ii) sind einfach nachzurechnen. Die Aussage (iv) beruht auf (iii). Die Dreiecksungleichung (iii) wird in der Linearen Algebra mit Hilfe des Begriffs des Skalarprodukts und der „Cauchy–Schwarz–Ungleichung" bewiesen(vgl. LAL02, Satz 18 (EU 2)). ♦

5.3 Besondere Teilmengen im $\mathbb{R}^d$

Definition [Kugeln, Sphären] Es sei $a \in \mathbb{R}^d$ und $r > 0$ (Das impliziert $r \in \mathbb{R}$).

(1) Wir definieren die Mengen

$$\begin{aligned} U_r(a) &:= \Big\{x \in \mathbb{R}^d \Big| \|x-a\| < r\Big\} && (\textit{Offene Umgebung oder Kugel}) \\ B_r(a) &:= \Big\{x \in \mathbb{R}^d \Big| \|x-a\| \le r\Big\} && (\textit{Abgeschlossener Ball}) \\ S_r(a) &:= \Big\{x \in \mathbb{R}^d \Big| \|x-a\| = r\Big\} && (\textit{Sphäre}) \end{aligned}$$

jeweils mit *Radius* r und *Mittelpunkt* a.

Es sei X eine Teilmenge des euklidischen $\mathbb{R}^d$.

(2) X heißt *beschränkt* (in $\mathbb{R}^d$), wenn es ein $M > 0$ gibt, so dass

$$\|x\| \le M \qquad \text{für alle } x \in X.$$

(3) X heißt *offen* (in $\mathbb{R}^d$), wenn es zu jedem $x \in X$ ein $\varepsilon > 0$ gibt, so dass

$$U_\varepsilon(x) \subseteq X.$$

Bildlich gesprochen: Jedes Element von X hat ein bißchen Abstand zum Rand.

(4) X heißt *abgeschlossen* (in $\mathbb{R}^d$), wenn die Komplementärmenge $\mathbb{R}^d \setminus X$ offen ist.

Beispiele (0) Überlege, was diese Begriffe in $\mathbb{R} = \mathbb{R}^1$ bedeuten.

(1) Umgebungen, Bälle und Sphären sind beschränkt.

(2) Eine offene Umgebung $U_r(a)$ (im Sinne von (1)) ist offen (im Sinne von (3)). Ist nämlich $x \in U_r(a)$ und dabei $r_1 := \|x-a\| < r$, so gilt für alle $y \in U_{r_2}(x)$ mit $r_2 := \frac{r-r_1}{2}$:

$$\|y-a\| \le \|y-x\| + \|x-a\| \le r_2 + r_1 = \frac{r-r_1}{2} + r_1 = \frac{r}{2} + \frac{r_1}{2} < r,$$

also $y \in U_r(a)$. Insgesamt bedeutet dies $U_{r_2}(x) \subseteq U_r(a)$.

(3) Bälle $B_r(a)$ und Sphären $S_r(a)$ sind abgeschlossen: Beweis als Übung.

5.4 Folgen im $\mathbb{R}^d$

Wir wiederholen und verallgemeinern die Definitionen im Umfeld von Folgen.

Definition Es sei $(a_{,n})_{n\in\mathbb{N}}$ eine *Folge im* $\mathbb{R}^d$, d.h. eine Abbildung

$$a_{,\bullet} \begin{cases} \mathbb{N} & \to & \mathbb{R}^d \\ n & \mapsto & a_{,n} = (a_{1,n}, \ldots, a_{d,n}) \end{cases}$$

Für jedes $j = 1, \ldots, d$ heißt die Folge $(a_{j,n})_{n\in\mathbb{N}}$ eine *Koordinatenfolge*. Das Komma vor einem Index bedeutet, dass es sich um einen Folgenindex handelt. Wir benutzen ab jetzt diese Schreibweise, um Folgen–Indices und Koordinaten–Indices unterscheiden zu können.

- Die Folge heißt *Cauchy–Folge*, wenn es zu jedem $\varepsilon > 0$ eine Zahl $N \in \mathbb{N}$ gibt, so dass für alle $n, m \in \mathbb{N}$ mit $n \geq N$, $m \geq N$ gilt:

  $$\|a_{,n} - a_{,m}\| < \varepsilon.$$

  Ganz grob: ,,Die Folgenglieder rücken immer näher aneinander".

- Es sei $a \in \mathbb{R}^d$. Die Folge heißt *konvergent gegen (den Grenzwert)* a, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt:

  $$\|a_{,n} - a\| < \varepsilon.$$

  In diesem Fall schreibt man $\lim_{n\to\infty} a_{,n} = a$.

- Die Folge heißt *divergent*, wenn sie nicht konvergent ist.

- Die Folge heißt *Nullfolge*, wenn sie gegen $\vec{0}$ konvergiert.

Satz 47
(i) Die folgenden Aussagen über eine Folge $(a_{,n})_{n\in\mathbb{N}}$ sind äquivalent:

1. *Die Folge $(a_{,n})_{n\in\mathbb{N}}$ ist eine Cauchy–Folge.*

2. *Für jedes $j = 1, \ldots, d$ ist die Koordinatenfolge $(a_{j,n})_{n\in\mathbb{N}}$ eine Cauchy–Folge.*

(ii) Die folgenden Aussagen über eine Folge $(a_{,n})_{n\in\mathbb{N}}$ sind äquivalent:

1. *Die Folge $(a_{,n})_{n\in\mathbb{N}}$ ist konvergent gegen $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$.*

2. *Es ist* $\lim\limits_{n\to\infty} \|a_{,n} - a\| = 0$ *(in $\mathbb{R}$).*

3. *Es ist* $\lim\limits_{n\to\infty} [(a_{1,n} - a_1)^2 + (a_{2,n} - a_2)^2 + \ldots + (a_{d,n} - a_d)^2] = 0.$

4. *Für jedes $j = 1, \ldots, d$ konvergiert die Koordinatenfolgen $(a_{j,n})_{n\in\mathbb{N}}$ gegen a_j:*

$$\begin{aligned} \lim_{n\to\infty} a_{1,n} &= a_1 \\ \lim_{n\to\infty} a_{2,n} &= a_2 \\ &\vdots \\ \lim_{n\to\infty} a_{d,n} &= a_d. \end{aligned}$$

Beweis Wir zeigen nur (ii) mittels einer Kette von Äquivalenzen:

$\lim a_{,n} = a \quad (\text{im } \mathbb{R}^d)$

$\iff$ Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass $\|a_{,n} - b\| < \varepsilon$ für alle $n \geq N$.

$\iff$ Es ist $\lim_{n\to\infty} \|a_{,n} - a\| = 0$.

$\iff$ Es ist $\lim\limits_{n\to\infty} \sqrt{(a_{1n} - a_1)^2 + (a_{2n} - a_2)^2 + \ldots + (a_{dn} - a_d)^2} = 0.$

$\iff$ Es ist $\lim\limits_{n\to\infty} [(a_{1n} - a_1)^2 + (a_{2n} - a_2)^2 + \ldots + (a_{dn} - a_d)^2] = 0.$

$\iff$ Es ist $\lim\limits_{n\to\infty} (a_{j,n} - a_j)^2 = 0 \quad$ für alle $j = 1, \ldots, n$.

$\iff$ Es ist $\lim\limits_{n\to\infty} (a_{j,n} - a_j) = 0 \quad$ für alle $j = 1, \ldots, n$.

$\iff$ Es ist $\lim\limits_{n\to\infty} a_{j,n} = a_j \quad$ für alle $j = 1, \ldots, n$.

♦

Satz 48
(i) Eine konvergente Folge ist eine Cauchy–Folge.

(ii) Der $\mathbb{R}^d$ ist vollständig im folgenden Sinne:

Jede Cauchy–Folge ist konvergent.

Beweis Der Beweis besteht darin, dass man diese Aussagen mittels Satz 47 auf die entsprechenden Aussagen in $\mathbb{R}$ (vgl. Sätze 18 und 22, ANA01) zurückführt:

Die Folge $(a_{,n})_{n\in\mathbb{N}}$ im $\mathbb{R}^d$ ist Cauchy–Folge.

$\iff$ Jede Koordinatenfolge $(a_{j,n})_{n\in\mathbb{N}}$, $j = 1, \ldots, d$, ist eine Cauchy–Folge.

$\iff$ Jede Koordinatenfolge $(a_{j,n})_{n\in\mathbb{N}}$, $j = 1, \ldots, d$, ist konvergent.

$\iff$ Die Folge $(a_{,n})_{n\in\mathbb{N}}$ im $\mathbb{R}^d$ ist konvergent.

♦

5.5 Kompakte Mengen im $\mathbb{R}^d$

Definition Eine Teilmenge des euklidischen $\mathbb{R}^d$ heißt *kompakt (in $\mathbb{R}^d$)*, wenn sie beschränkt und abgeschlossen ist.

Satz 49 *Es sei X eine Teilmenge von $\mathbb{R}^d$.*

(1) X ist genau dann beschränkt, wenn jede Folge in X eine konvergente Teilfolge enthält.

(2) X ist genau dann abgeschlossen, wenn jede konvergente Folge in X ihren Grenzwert in X hat.

(3) X ist genau dann kompakt, wenn jede Folge in X eine konvergente Teilfolge mit Grenzwert in X hat.

Beweis (1) Wenn X unbeschränkt ist, so gibt es zu jedem $n \in \mathbb{N}$ ein $a_{,n} \in X$ mit $\|a_{,n}\| \geq n$. Diese Folge besitzt keine konvergente Teilfolge.

Ist umgekehrt X beschränkt, so ist jede Folge $a_{,n} \subseteq X$ beschränkt, das heißt, es existiert ein $M \geq 0$, so dass

$$\|a_{,n}\| \leq M \qquad \text{für alle } n \in \mathbb{N}.$$

Dann ist aber auch jede der Koordinatenfolgen $(a_{i,n})_{n\in\mathbb{N}}$ beschränkt.

Wir zeigen jetzt per Induktion über die Koordinatennummern $i = 0, \ldots, d$ die folgende Aussage:

Es existiert eine (streng monoton steigende) Auswahlfunktion natürlicher Zahlen

$$\begin{cases} \mathbb{N} & \to & \mathbb{N} \\ k_i & \mapsto & n_{k_i} \end{cases}$$

(der Index i als die Nummer des aktuellen Induktionsschritts kennzeichnet den Namen des Folgenindex k_i), so dass die ersten i gemäß dieser Auswahl gebildeten Koordinaten–Teilfolgen

$$k_i \mapsto a_{1,n_{k_i}} \qquad \ldots\ldots \qquad k_i \mapsto a_{i,n_{k_i}} \qquad (*)_i$$

jeweils konvergieren.

Der Induktionsanfang ist trivial: Man wählt einfach die Auswahl $k_0 = \mathrm{id}_{\mathbb{N}}$, was bedeutet, dass gar keine echte Teilfolge ausgewählt wurde. Konvergenzeigenschaften müssen hier nicht bewiesen werden.

Induktionsschritt $i \mapsto i + 1$. Wir betrachten die $(i + 1)$–te Koordinatenfolge und deren entsprechend dem vorhergehenden Schritt ausgewählte Teilfolge:

$$k_i \mapsto a_{i+1,n_{k_i}}$$

Diese Folge ist beschränkt, besitzt also gemäß dem Satz von Bolzano–Weierstraß (ANA01, Satz 24) eine konvergente Teilfolge

$$k_{i+1} \mapsto a_{i+1,n_{k_{i+1}}} \qquad (**)_{i+1}$$

Die Tatsache, dass es sich um eine Teilfolge handelt, wird dabei dadurch ausgedrückt, dass zu jedem $k_{i+1} \in \mathbb{N}$ ein $k_i \in \mathbb{N}$ existiert, so dass $n_{k_{i+1}} = n_{k_i}$.

Diese Auswahl bringt es aber mit sich, dass alle Koordinaten–Teilfolgen

$$k_{i+1} \mapsto a_{1,n_{k_{i+1}}}, \qquad \ldots\ldots \qquad k_{i+1} \mapsto a_{i+1,n_{k_{i+1}}}$$

jeweils konvergieren, die ersten i deshalb, weil sie Teilfolgen der nach Induktionsvoraussetzung konvergenten Teilfolgen $(*)_i$ sind, die letzte deshalb, weil wir sie in $(**)_{i+1}$ so konstruiert haben.

Damit ist die Induktion durchgeführt.

Setzen wir jetzt in der soeben bewiesenen Aussage $i = d$, so heißt das:

Es existiert eine Auswahlfunktion

$$\begin{cases} \mathbb{N} & \to & \mathbb{N} \\ k_d & \mapsto & n_{k_d} \end{cases}$$

so dass alle d Koordinaten–Teilfolgen

$$k_d \mapsto a_{1,n_{k_d}}, \qquad \ldots\ldots \qquad k_d \mapsto a_{d,n_{k_d}}$$

jeweils konvergieren. Das aber bedeutet, dass die Teilfolge

$$k_d \mapsto a_{,n_{k_d}}$$

der ursprünglich gegebenen Folge $a_{,n}$ im $\mathbb{R}^d$ konvergiert.

(2) Wäre a in der offenen Menge $\mathbb{R}^d \setminus X$ enthalten, so gäbe es ein $\varepsilon > 0$, so dass die offene ε–Umgebung $U_\varepsilon(a)$ ganz in $\mathbb{R}^d \setminus X$ liegt. Dann können aber die in X liegenden Folgenglieder $a_{,n}$ nicht auf einen Abstand $< \varepsilon$ an a ,,herankommen".

Umgekehrt sei X eine Menge, die ,,gegenüber Folgengrenzwerten abgeschlossen" ist.

Wir nehmen an, dass ihr Komplement $\mathbb{R}^d \setminus X$ nicht offen ist. Das bedeutet, dass es ein $a \in \mathbb{R}^d \setminus X$ gibt, dessen offene Umgebungen $U_{1/n}(a)$ für alle $n \in \mathbb{N}$ nicht ganz zu $\mathbb{R}^d \setminus X$ gehören. Es gibt also für jedes $n \in \mathbb{N}$ ein $a_{,n} \in U_{1/n}(a) \cap X$. Die so gebildete Folge $(a_{1,n})_{n\in\mathbb{N}}$ liegt in X und konvergiert gegen a. Das würde aber gemäß der ,,Folgenabgeschlossenheit" $a \in X$ nach sich ziehen. Widerspruch.

(3) Wie aus der Definition der Kompaktheit zu entnehmen ist, ist (3) eine Kombination von (1) und (2). ♦

Korollar 50 *Eine kompakte Teilmenge K von $\mathbb{R}$ enthält ihr Minimum und Maximum.*

Beweis Wir beschränken uns auf das Maximum. Da K beschränkt ist, ist $\sup K$ endlich. Gemäß Definition von sup existiert eine Folge $(a_{,n}) \subseteq \mathbb{R}$ mit $\lim_{n\to\infty} a_{,n} = \sup K$. Da K abgeschlossen ist, gilt

$$\sup K = \lim_{n\to\infty} a_{,n} \in K, \qquad \text{also } \sup K = \max K.$$

♦

5.6 Stetige Funktionen im $\mathbb{R}^d$

Satz 51 (und Definition) *Es sei $a \in X \subseteq \mathbb{R}^d$ und $f : X \to \mathbb{R}^\ell$ eine Funktion. Die folgenden Aussagen sind äquivalent:*

(1) (Definition) Die Funktion f heißt stetig in a.

(2) (Folgen–Definition) Für Folgen $(x_{,n})_{n\in\mathbb{N}} \subseteq X$ besteht die Implikation

$$\lim_{n\to\infty} x_{,n} = a \quad \Longrightarrow \quad \lim_{n\to\infty} f(x_{,n}) = f(a).$$

(3) Für Folgen $(x_{,n})_{n\in\mathbb{N}} \subseteq X$ besteht die Implikation

$$\lim_{n\to\infty} \|x_{,n} - a\| = 0 \quad \Longrightarrow \quad \lim_{n\to\infty} \|f(x_{,n}) - f(a)\| = 0.$$

(4) (ε–δ–Definition) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass

$$\|f(x) - f(a)\| < \varepsilon \quad \text{für alle } x \text{ mit } \quad \|x - a\| < \delta.$$

Beweis Die Äquivalenz (2) $\Longleftrightarrow$ (3) ist klar mit Satz 47. Die beiden Aussagen (3) und (4) haben Folgen von reellen Zahlen zum Inhalt. Deshalb ist die Äquivalenz (3) $\Longleftrightarrow$ (4) bereits bekannt (Satz 37, ANA01).

♦

Stetigkeit auf X heißt wieder Stetigkeit in jedem $a \in X$.

Definition Es sei $X \subseteq \mathbb{R}^d$ und $f : X \to \mathbb{R}^\ell$ eine Funktion. Für festes $a \in X$ und ein festes $j \in \{1, \ldots, d\}$ heißt die Funktion

$$f(a_1, \ldots, a_{j-1}, \underset{\substack{\uparrow\\ j}}{\cdot}, a_{j+1}, \ldots, a_d) : \left\{ \begin{array}{rcl} X_j & \to & \mathbb{R}^\ell \\ x & \mapsto & f(a_1, \ldots, a_{j-1}, x, a_{j+1}, \ldots, a_d) \end{array} \right.$$

die *partielle Funktion für die j–te Stelle*. Dabei enthält X_j genau diejenigen Zahlen $x \in \mathbb{R}$, für die $(a_1, \ldots, a_{j-1}, x, a_{j+1}, \ldots, a_d)$ in X enthalten ist.

Satz 52 *(1) Eine Funktion $f : X \to \mathbb{R}^\ell$ ist genau dann stetig, wenn für jedes $j = 1, \ldots, \ell$ die Koordinatenfunktion*

$$f_j \left\{ \begin{array}{rcl} X & \to & \mathbb{R} \\ x & \mapsto & f(x)_j \end{array} \right.$$

stetig ist.

(2) Aus der Stetigkeit der partiellen Funktionen kann nicht auf die Stetigkeit der Funktionen geschlossen werden.

Beweis (1) Dies liegt daran, dass Folgen im $\mathbb{R}^\ell$ genau dann konvergieren, wenn alle Koordinatenfolgen konvergieren.

(2) Für die Funktion

$$f : \left\{ \begin{array}{rcl} \mathbb{R}^2 & \to & \mathbb{R} \\ (x_1, x_2) & \mapsto & \left\{ \begin{array}{ll} 1, & \text{falls } x_1 \cdot x_2 \neq 0, \\ 0, & \text{sonst} \end{array} \right. \end{array} \right.$$

sind die partiellen Funktionen $f(\cdot\,,0)$ und $f(0,\cdot\,)$ konstant Null, also stetig. f ist aber nicht stetig in $(0,0)$, da in jeder Umgebung von $(0,0)$ der Funktionswert 1 angenommen wird.

(3) Noch interessanteres Gegenbeispiel: Betrachte (Zeichnung) die Funktion

$$f : \begin{cases} \mathbb{R}^2 & \to \quad \mathbb{R} \\ (x_1, x_2) & \mapsto \quad \begin{cases} 1, & \text{falls } 0 < x_2 < x_1^2, \\ 0, & \text{sonst.} \end{cases} \end{cases}$$

Jede Folge, die ganz auf einer Geraden durch den Ursprung gegen den Ursprung konvergiert, ist schließlich eine Nullfolge. Trotzdem ist f nicht stetig in $(0,0)$. ♦

Satz 53 *(Geeignete) Linearkombinationen, Produkte, Quotienten und Hintereinanderausführungen stetiger Funktionen sind wieder stetig.*

Beweis Ohne Beweis! ♦

Satz 54 *Die folgenden Aussagen über eine Funktion $f : \mathbb{R}^d \to \mathbb{R}^\ell$ sind äquivalent:*

(1) f ist stetig.

(2) Das Urbild $f^{-1}(Y)$ jeder offenen Teilmenge $Y \subseteq \mathbb{R}^\ell$ ist offen.

(3) Das Urbild $f^{-1}(Y)$ jeder abgeschlossenen Teilmenge $Y \subseteq \mathbb{R}^\ell$ ist abgeschlossen.

Beweis (1) $\Longrightarrow$ (3): Es sei $Y \subseteq \mathbb{R}^\ell$ abgeschlossen und $(a_{,n})_{n\in\mathbb{N}}$ eine konvergente Folge in $f^{-1}(Y)$. Dann ist auch die durch $b_{,n} := f(a_{,n})$ definierte Folge $(b_{,n})_{n\in\mathbb{N}} \subseteq Y$ konvergent. Ihr Grenzwert b ist wegen der Abgeschlossenehit von Y in Y enthalten. Es gilt insgesamt:

$$f(a) = f(\lim_{n\to\infty} a_{,n}) = \lim_{n\to\infty} f(a_{,n}) = \lim_{n\to\infty} b_{,n} = b \in Y,$$

also $a \in f^{-1}(Y)$. Also ist $f^{-1}(Y)$ abgeschlossen.

(3) $\Longrightarrow$ (2) beweisen wir durch die folgende Implikationskette:

$$\begin{aligned} & Y \subseteq \mathbb{R}^\ell \text{ offen} \\ \Longrightarrow\ & \mathbb{R}^\ell \setminus Y \text{ abgeschlossen} \\ \Longrightarrow\ & f^{-1}(\mathbb{R}^\ell \setminus Y) \text{ abgeschlossen} \\ \Longrightarrow\ & f^{-1}(\mathbb{R}^\ell) \setminus f^{-1}(Y) \text{ abgeschlossen} \\ \Longrightarrow\ & \mathbb{R}^d \setminus f^{-1}(Y) \text{ abgeschlossen} \\ \Longrightarrow\ & f^{-1}(Y) \text{ offen.} \end{aligned}$$

Dabei haben wir für die dritte Schlussfolgerung benutzt, dass das Urbild einer der Differenz zweier Mengen A und B gleich der Mengendifferenz der Urbilder ist:

$$f^{-1}(A \setminus B) = f^{-1}(A) \setminus f^{-1}(B).$$

Das ist eine Eigenschaft beliebiger Funktionen, die wir in ANA01 am Anfang beweisen hätten sollen (können). Das ist eine Übung für Sie.

(2) $\Longrightarrow$ (1): Wir betrachten ein beliebieges $a \in X$, es sei $\varepsilon > 0$ vorgegeben. Das Urbild der offenen Menge $U_\varepsilon(f(a))$ ist nach Voraussetzung (2) offen, es existiert also ein $\delta > 0$, so dass die offene δ–Kugel $U_\delta(a)$ um a in $f^{-1}(U_\varepsilon(f(a)))$ enthalten ist. Daraus folgt aber:

$$\begin{aligned} \|x-a\| < \delta \quad &\Longrightarrow \quad x \in U_\delta(a) \subseteq f^{-1}(U_\varepsilon(f(a))) \quad \Longrightarrow \quad f(x) \in U_\varepsilon(f(a)) \\ &\Longrightarrow \quad \|f(x)-f(a)\| < \varepsilon. \end{aligned}$$

Also ist f stetig in a. Da a beliebig war, ist f überhaupt stetig. ♦

Satz 55 *Ist $K \subseteq \mathbb{R}^d$ eine kompakte Teilmenge und $f : K \to \mathbb{R}^\ell$ stetig, so ist auch die Bildmenge $f(K)$ kompakt.*

Beweis Der Beweis beruht im wesentlichen auf der Charakterisierung von kompakten Mengen in Satz 49.

Es sei $(b_{,n}) \subseteq f(K)$ eine beliebige Folge im Bild von K. Zu jedem $n \in \mathbb{N}$ gibt es ein $a_{,n} \in K$, so dass $f(a_{,n}) = b_{,n}$. Die Folge $n \mapsto a_{,n}$ besitzt gemäß Satz 49 eine Teilfolge $k \mapsto a_{,n_k}$, deren Grenzwert

$$a = \lim_{k\to\infty} a_{,n_k} \in K$$

existiert. Da f stetig ist, existiert auch

$$b := \lim_{k\to\infty} f(a_{,n_k}) = f(\lim_{k\to\infty} a_{,n_k}) = f(a) \in f(K).$$

Also besitzt $(b_{,n})_{n\in\mathbb{N}}$ eine in $f(K)$ konvergente Teilfolge. Wieder mit Satz 49 folgt die Behauptung. ♦

Korollar 56 *Eine stetige Funktion $f : K \to \mathbb{R}$ auf einer kompakten Menge $K \subseteq \mathbb{R}^d$ nimmt ihr Minimum und Maximum an, d.h. es gibt $a, b \in K$, so dass*

$$f(a) = \min f(K) \qquad \textit{und} \qquad f(b) = \max f(K).$$

Der Beweis ist einfach eine Kombination von Satz 55 und Korollar 50.

Definition Es sei $X \subseteq \mathbb{R}^d$ und $f : X \to \mathbb{R}^\ell$ eine Funktion. f heißt *gleichmäßig stetig*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass

$$\|f(x) - f(y)\| < \varepsilon, \qquad \text{falls } \|x - y\| < \delta.$$

Beachte, dass das δ **unabhängig** von einer festen Stelle $a \in [c, d]$ gewählt sein muss.

Satz 57 *(1) Ist $f : [c, d] \to \mathbb{R}$ stetig, so ist f sogar gleichmäßig stetig.*

(2) Allgemeiner: Ist X kompakt, so ist $f : X \to \mathbb{R}^\ell$ gleichmäßig stetig.

Beweis
Wir beweisen nur die erste Aussage. Dafür sei ein $\varepsilon > 0$ vorgegeben. Wir betrachten die folgende Familie von Aussagen $(\mathcal{A}_z)_{z \in [c,d]}$:

Es existiert ein δ_z, so dass

$$\|f(x) - f(y)\| < \varepsilon, \qquad \text{falls } x, y \in [c, z] \text{ und } |x - y| < \delta_z.$$

Wenn $\mathcal{A}_d$ gezeigt ist, ist die Aussage (1) des Satzes bewiesen. Wir nehmen an, es gibt ein $z \in [c, d]$, so dass $\mathcal{A}_z$ nicht gilt, und setzen dann

$$a := \sup\{z | \mathcal{A}_z \text{ gilt nicht } \}.$$

Da f in a stetig ist, gibt es ein $\gamma > 0$, so dass

$$\|f(x) - f(a)\| < \frac{\varepsilon}{2}, \qquad \text{falls } x \in [a - \gamma, a + \gamma]. \qquad (*)$$

Wir setzen jetzt

$$\delta_{a+\gamma} := \min\{\delta_{a - \frac{\gamma}{2}}, \frac{\gamma}{2}\} > 0$$

und zeigen, dass $\mathcal{A}_{a+\gamma}$ gilt. Es seien also $x, y \in [c, a + \gamma]$ mit $|x - y| < \delta_{a+\gamma}$.

1. Fall: $x, y \in [c, a - \frac{\gamma}{2}]$. Da $\mathcal{A}_{a-\frac{\gamma}{2}}$ wahr und $\delta_{a+\gamma} < \delta_{a-\frac{\gamma}{2}}$ ist, gilt $\|f(x) - f(y)\| < \varepsilon$.

2. Fall: $x, y \in [a - \gamma, a + \gamma]$. Dann gilt aufgrund von $(*)$

$$\|f(x) - f(y)\| \le \|f(x) - f(a)\| + \|f(a) - f(y)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

3. Fall: $x \in [c, a - \gamma], y \in [a - \gamma, a + \gamma]$. Dann folgt aber mit $|x - y| < \delta_{a+\gamma} \le \frac{\gamma}{2}$, dass

$$y < x + \frac{\gamma}{2} \le a - \gamma + \frac{\gamma}{2} = a - \frac{\gamma}{2},$$

also $x, y \in [c, a - \frac{\gamma}{2}]$. Das ist aber wieder Fall 1. ♦

6 Kurven im $\mathbb{R}^w$

6.1 Definitionen und Beispiele

Definition (1) Es sei J ein Intervall. Eine stetige Abbildung

$$\gamma : \begin{cases} J & \to & \mathbb{R}^w \\ t & \mapsto & \gamma(t) = (\gamma_1(t), \ldots, \gamma_w(t)) \end{cases}$$

heißt *Kurve im* $\mathbb{R}^w$. Eine Kurve wird also durch w stetige Funktionen, die Koordinatenfunktionen, $\gamma_i : J \to \mathbb{R}$ beschrieben.

In diesem Zusammenhang heißt die Definitionsmenge auch *Parameterbereich*, die Bildmenge

$$\gamma(J) = \Big\{\gamma(t) \,\Big|\, t \in J\Big\} \subseteq \mathbb{R}^w$$

heißt *Spur* der Kurve. Oft kommt es vor (beachte aber genau den Unterschied), dass zwischen Kurve (Abbildung) und Spur (Menge) nicht klar unterschieden wird.

Physikalische Vorstellung: Die Variable t wird als Zeit aufgefasst. Die Funktion γ gibt den Ort (im $\mathbb{R}^w$) eines (punktförmig gedachten) Teilchens zum Zeitpunkt t an.

(2) Die Kurve γ heißt *(stetig) differenzierbar (in t)*, wenn alle Koordinatenfunktionen (stetig) differenzierbar (in t) sind.

(3) Ist eine Kurve in $\tau \in J$ differenzierbar, so heißt der Vektor

$$\gamma'(\tau) \ := \ (\gamma_1'(\tau), \ldots, \gamma_w'(\tau))$$

der *Tangentenvektor an die Kurve γ zum Parameterwert τ*. Die Gerade

$$\{\gamma(\tau) + s\gamma'(\tau) | s \in \mathbb{R}\}$$

heißt die *Tangente an die Kurve γ zum Parameterwert τ* — oder (nicht ganz eindeutig:) im Punkt $\gamma(\tau)$.

(4) Eine differenzierbare Kurve γ heißt *regulär*, wenn $\gamma'(t) \neq 0$ für alle $t \in J$. Sie heißt *singulär* für $\tau \in J$, falls $\gamma'(\tau) = 0$.

Ist $\gamma : [c, d] \to \mathbb{R}^w$ eine Kurve und $\vartheta : [c_2, d_2] \to [c, d]$ eine streng monoton steigende (oder fallende) surjektive Abbildung, so ist

$$\gamma \circ \vartheta : \begin{cases} [c_2, d_2] & \to & \mathbb{R}^w \\ s & \mapsto & \gamma(\vartheta(s)) \end{cases}$$

eine Kurve mit der gleichen Spur wie γ. Mit Hilfe von ϑ wurde der Parameter t in den Parameter s *transformiert*. Die Abbildung ϑ heißt in diesem Zusammenhang eine (*orientierungserhaltende* bzw. *orientierungsumkehrende*) *Parametertransformation.*

Beispiele (1) Sind a, b zwei positive Zahlen, so wird durch die Abbildung

$$\gamma \begin{cases} \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (a \cdot \cos t, b \cdot \sin t) \end{cases}$$

eine *Ellipse* mit dem Ursprung als Mittelpunkt und den beiden Halbachsen a und b beschrieben.

Für $a = b$ ergibt sich eine *Kreislinie.*

(2) Die Abbildung

$$\gamma : \begin{cases} \mathbb{R} & \to & \mathbb{R}^3 \\ t & \mapsto & (\cos t, \sin t, t) \end{cases}$$

beschreibt eine *Schraubenlinie* im $\mathbb{R}^3$.

(3) Ist $f : J \to \mathbb{R}$ eine stetige Funktion, so kann der Graph als Kurve im $\mathbb{R}^2$

$$\gamma_f : \begin{cases} \mathbb{R} & \to & \mathbb{R}^2 \\ t & \mapsto & (t, f(t)) \end{cases}$$

aufgefasst werden.

6.2 Bogenlänge

Wir betrachten jetzt Kurven γ, die ein kompaktes Intervall $J = [c, d]$ als Parameterbereich haben.

Für eine vorgegebene Zerlegung $\mathcal{Z}$ (vgl. Abschnitt 3.1)

$$c = t_0 < t_1 < \ldots < t_n = d$$

des Intervalls können wir die Gesamtlänge des die Punkte

$$\gamma(t_0), \quad \gamma(t_1), \quad \ldots \quad \gamma(t_n)$$

verbindenden Polygonzugs (Vereinigung der Verbindungsstrecken) im $\mathbb{R}^w$ berechnen:

$$\mathcal{L}(\gamma, \mathcal{Z}) \ := \ \|\gamma(t_1) - \gamma(t_0)\| + \|\gamma(t_2) - \gamma(t_1)\| + \ldots + \|\gamma(t_n) - \gamma(t_{n-1})\|.$$

Gehen wir zu einer Verfeinerung $\widetilde{\mathcal{Z}}$ von $\mathcal{Z}$ über so gilt wegen der Dreiecksungleichung:

$$\mathcal{L}(\gamma, \mathcal{Z}) \leq \mathcal{L}(\gamma, \widetilde{\mathcal{Z}}).$$

Definition Als *Länge* der Kurve definieren wir die Obergrenze all dieser Längen:

$$\mathcal{L}(\gamma) \ := \ \sup_{\mathcal{Z}} \{\mathcal{L}(\gamma, \mathcal{Z})\}.$$

Ist $\mathcal{L}(\gamma) < \infty$, so heißt die Kurve γ *rektifizierbar.*

In diese Definition gehen

- nur Punkte $\gamma(t_i)$ auf der Spur von γ in einer bestimmten Reihenfolge,
- nicht aber die Kurve als Abbildung $\gamma : [c, d] \to \mathbb{R}$

ein. Das bedeutet, dass die Definition unabhängig von einer Parametertransformation ist. Eigentlich müsste man das präziser beschreiben und beweisen.

Satz 58 *(1) Es gibt stetige Kurven, die nicht rektifizierbar sind.*

(2) Ist die Kurve $\gamma : [c, d] \to \mathbb{R}^w$ stetig differenzierbar, so ist γ rektifizierbar und es gilt:

$$\mathcal{L}(\gamma) = \int_c^d \|\gamma'(t)\| \, dt = \int_c^d \sqrt{[\gamma_1'(t)]^2 + [\gamma_2'(t)]^2 + \ldots + [\gamma_w'(t)]^2} \, dt.$$

Beweis (1) Als Beispiel nehmen wir die auf $[0, 1]$ definierte Kurve im $\mathbb{R}^2$:

$$\gamma(t) = \begin{cases} \begin{pmatrix} t \\ t \sin(\frac{\pi}{t} - \frac{\pi}{2}) \end{pmatrix}, & \text{falls } t > 0, \\ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & \text{sonst.} \end{cases}$$

Wir wählen eine feste gerade natürliche Zahl n und dann eine Zerlegung $\mathcal{Z}_n$ gemäß der zweiten Zeile der folgenden Tabelle

i	0	1	2	$\ldots$	$n-3$	$n-2$	$n-1$	n
t_i	0	$\frac{1}{n}$	$\frac{1}{n-1}$	$\ldots$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{1}$
$\sin(\frac{\pi}{t} - \frac{\pi}{2})$	/	-1	$+1$	$\ldots$	-1	$+1$	-1	$+1$
$\gamma(t_i)$	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{n} \\ \frac{-1}{n} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{n-1} \\ \frac{+1}{n-1} \end{pmatrix}$	$\ldots$	$\begin{pmatrix} \frac{1}{4} \\ \frac{-1}{4} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{3} \\ \frac{+1}{3} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ \frac{-1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{1} \\ \frac{+1}{1} \end{pmatrix}$

mit $t_i := \frac{1}{n-i+1}$ für $i \geq 1$. Für $i \geq 2$ ist

$$\|\gamma(t_i) - \gamma(t_{i-1})\| = \sqrt{(t_i - t_{i-1})^2 + (t_i + t_{i-1})^2} \geq t_i + t_{i-1} \geq 2t_{i-1}$$

Es ergibt sich als Länge des zugehörigen Polygonzugs:

$$\begin{aligned} \mathcal{L}(\gamma, \mathcal{Z}_n) &= \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\| \geq \sum_{i=1}^{n} 2t_{i-1} = \\ &= 2 \cdot [\frac{1}{n} + \frac{1}{n-1} + \ldots + \frac{1}{2}]. \end{aligned}$$

Sie wächst für $n \to \infty$ gegen ∞.

(2) Es sei ein (beliebig kleines) $\varepsilon > 0$ vorgegeben.

Schritt 1: Die Funktionen γ_j' sind stetig, also mit Satz 57 sogar gleichmäßig stetig. Daher gibt es ein $\delta > 0$, so dass für alle $j = 1, \ldots, w$

$$\|\gamma_j'(t) - \gamma_j'(s)\| \leq \frac{\varepsilon}{w(d-c)}, \qquad \text{wenn nur } |t - s| < \delta.$$

Schritt 2: Wir zeigen: Hat eine Zerlegung

$$\mathcal{Z} : c = t_0 < t_1 < \ldots < t_n = d$$

von $[c, d]$ eine Körnigkeit $< \delta$, d.h.

$$t_i - t_{i-1} < \delta \qquad \text{für alle } i = 1, \ldots, n,$$

so gilt

$$\left|\mathcal{L}(\gamma, \mathcal{Z}) - \int_c^d \|\gamma'(t)\| \, dt\right| \leq \varepsilon.$$

Es sei nun i eine der Nummern aus der Zerlegung $\mathcal{Z}$. Gemäß Mittelwertsatz der Integralrechnung 30(4) — $t \mapsto \|\gamma'(t)\|$ ist stetig — gibt es ein $\tau \in [t_{i-1}, t_i]$, so dass

$$\int_{t_{i-1}}^{t_i} \|\gamma'(t)\| \, dt = \|\gamma'(\tau)\| \cdot [t_i - t_{i-1}].$$

Gemäß Mittelwertsatz 10 der Differentialrechnung gibt es für jede Koordinatennummer $j \in \{1, \ldots, w\}$ eine Stelle $\tau_j \in [t_{i-1}, t_i]$, so dass

$$|\gamma_j(t_i) - \gamma_j(t_{i-1})| = \gamma_j'(\tau_j) \cdot [t_i - t_{i-1}].$$

Zusammen gilt dann, wenn wir noch zweimal die Dreiecksungleichung anwenden

$$\begin{aligned}
&\left|\|\gamma(t_i) - \gamma(t_{i-1})\| - \int_{t_{i-1}}^{t_i} \|\gamma'(t)\| \, dt\right| = \\
&\left|\|(\gamma_1'(\tau_1), \gamma_2'(\tau_2), \ldots, \gamma_w'(\tau_w))\| - \|(\gamma_1'(\tau), \gamma_2'(\tau), \ldots, \gamma_w'(\tau))\|\right| \cdot [t_i - t_{i-1}] \leq \\
&\|((\gamma_1'(\tau_1) - \gamma_1'(\tau), \gamma_2'(\tau_2) - \gamma_2'(\tau), \ldots, \gamma_w'(\tau_w) - \gamma_w'(\tau))\| \cdot [t_i - t_{i-1}] \leq \\
&\sum_{j=1}^{w} |\gamma_j'(\tau_j) - \gamma_j'(\tau)| \cdot [t_i - t_{i-1}] \leq \\
&\sum_{j=1}^{w} \frac{\varepsilon}{w(d-c)} \cdot [t_i - t_{i-1}] = \frac{\varepsilon}{d-c} \cdot [t_i - t_{i-1}].
\end{aligned}$$

Summieren wir über alle Teilintervalle der Zerlegung $\mathcal{Z}$, so ergibt sich:

$$\begin{aligned}
&\left|\mathcal{L}(\gamma, \mathcal{Z}) - \int_c^d \|\gamma'(t)\| \, dt\right| = \\
&\left|\sum_{i=1}^{n} \left[\|\gamma(t_i) - \gamma(t_{i-1})\| - \int_{t_{i-1}}^{t_i} \|\gamma'(t)\| \, dt\right]\right| \leq \\
&\sum_{i=1}^{n} \frac{\varepsilon}{d-c} \cdot [t_i - t_{i-1}] = \frac{\varepsilon}{d-c} \cdot (d-c) = \varepsilon.
\end{aligned}$$

Schritt 3: Es sei jetzt $\mathcal{Z}'$ eine beliebige Zerlegung von $[c, d]$. Dann gibt es eine Verfeinerung $\mathcal{Z}$ mit einer Körnigkeit $< \delta$ und wir erhalten:

$$\mathcal{L}(\gamma, \mathcal{Z}') \leq \mathcal{L}(\gamma, \mathcal{Z}) \leq \int_c^d \|\gamma'(t)\| \, dt + \varepsilon.$$

Es folgt, da $\mathcal{Z}'$ völlig beliebig war:

$$\mathcal{L}(\gamma) \leq \int_c^d \|\gamma'(t)\| \, dt + \varepsilon.$$

Andererseits ist

$$\mathcal{L}(\gamma) \geq \mathcal{L}(\gamma, \mathcal{Z}) \geq \int_c^d \|\gamma'(t)\| \, dt - \varepsilon.$$

In einem geschrieben ist dies:

$$\int_c^d \|\gamma'(t)\| \, dt - \varepsilon \leq \mathcal{L}(\gamma) \leq \int_c^d \|\gamma'(t)\| \, dt + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung des Satzes. ♦

Bemerkung Ist $\vartheta : [c_2, d_2] \to [c, d]$ eine stetig differenzierbare streng monoton steigende (bzw. fallende) Abbildung, also eine Parametertransformation, so gilt mit der Kettenregel (KR) und der Substitutionsregel (SR: Satz 33)

$$\int_{c_2}^{d_2} \|(\gamma \circ \vartheta)'(s)\| \, ds \overset{KR}{=} \int_{c_2}^{d_2} \|(\gamma'(\vartheta(s)) \cdot \vartheta'(s)\| \, ds =$$
$$\left\{ \begin{array}{ll} \int_{c_2}^{d_2} \|(\gamma'(\vartheta(s))\| \cdot \vartheta'(s) \, ds, & \text{(falls } \vartheta \text{ steigend)} \\ -\int_{c_2}^{d_2} \|(\gamma'(\vartheta(s))\| \cdot \vartheta'(s) \, ds, & \text{(falls } \vartheta \text{ fallend)} \end{array} \right\} \overset{SR}{=} \int_c^d \|\gamma'(t)\| \, dt.$$

Damit wird auf andere Weise bestätigt, dass die Länge einer stetig differenzierbaren Kurve γ unabhängig von der konkreten Parameterdarstellung ist.

Korollar 59 *Die Graph–Kurve γ_f einer stetig differenzierbaren Funktion $f : [c, d] \to \mathbb{R}$ hat die Länge*

$$\mathcal{L}(\gamma_f) = \int_c^d \sqrt{1 + [f'(t)]^2} \, dt.$$

Beispiele (1) Die Länge der Kreislinie

$$\gamma : \begin{cases} [0, 2\pi] & \to & \mathbb{R}^2 \\ t & \mapsto & (r\cos t, r\sin t) \end{cases}$$

(vgl. Beispiel 1 oben) ist

$$\mathcal{L}(\gamma) = \int_0^{2\pi} \sqrt{(-r\sin t)^2 + (r\cos t)^2}\; dt = 2\pi.$$

(2) Wir betrachten die Kurve, die ein Randpunkt auf der Umfangslinie eines auf der x–Achse nach rechts abrollenden Einheitskreises während einer Umdrehung beschreibt.

Zeichnung

Die Kurve ist gegeben durch die Funktion

$$\gamma \begin{cases} [0, 2\pi] & \to & \mathbb{R} \\ t & \mapsto & \begin{pmatrix} t \\ 1 \end{pmatrix} + \begin{pmatrix} \cos(\frac{\pi}{2} - t) \\ \sin(\frac{\pi}{2} - t) \end{pmatrix} = \begin{pmatrix} t + \sin t \\ 1 + \cos t \end{pmatrix} \end{cases},$$

was wir noch erläutern wollen:

- Der erste Vektor $\begin{pmatrix} t \\ 1 \end{pmatrix}$ beschreibt die Position des Mittelpunktes zur Zeit t. Beim Start befindet sich der Mittelpunkt auf der Höhe 1 über dem Nullpunkt der x–Achse.
- Der zweite Vektor $\begin{pmatrix} \cos(\frac{\pi}{2} - t) \\ \sin(\frac{\pi}{2} - t) \end{pmatrix}$ beschreibt die Position des Randpunktes relativ zum Mittelpunkt. Beim Start ($t = 0$) befindet sich der Randpunkt in der 12–Uhr–Position. Die Variable t geht über ein Minuszeichen ein, da das Rad eine Rechts–Drehung (mathematisch negativ) ausführt.
- Das Intervall $[0, 2\pi]$ beschreibt aufgrund der 2π–Periodizität genau eine Volldrehung. Dabei bewegt sich auch der Mittelpunkt um die Strecke 2π weiter.

Für die Länge der Kurve ergibt sich jetzt

$$\begin{aligned}
\mathcal{L}(\gamma) &= \int_0^{2\pi} \|\gamma'(t)\|\, dt = \int_0^{2\pi} \sqrt{[1+\cos t]^2 + [\sin t]^2}\, dt \\
&= \int_0^{2\pi} \sqrt{2 + 2\cos t}\, dt = \int_0^{2\pi} \sqrt{2 + 2(\cos^2 \frac{t}{2} - \sin^2 \frac{t}{2})}\, dt \\
&= \int_0^{2\pi} \sqrt{4\,\cos^2(\frac{t}{2})}\, dt = \int_0^{2\pi} 2\,|\cos(\frac{t}{2})|\, dt = 4 \int_0^{\pi} \cos(\frac{t}{2})\, dt \\
&= 8 \int_0^{\frac{\pi}{2}} \cos s\, ds = 8\Big[\sin s\Big]_{s=0}^{s=\frac{\pi}{2}} = 8.
\end{aligned}$$

Der Randpunkt legt also während einer Umdrehung den 8–fachen Radius zurück.

7 Differenzierbarkeit von Funktionen mehrerer Variabler

Im folgenden sei X eine offene Teilmenge von $\mathbb{R}^d$, $a \in X$ und $f{:}X \to \mathbb{R}^w$ eine Funktion. Die Vektoren aus $\mathbb{R}^d$ (**D**efinitionsmenge) bzw. $\mathbb{R}^w$ (**W**ertemenge) werden generell als Spaltenvektoren aufgefasst (Im Staatsexamen sind nur die Fälle $d + w \leq 3$ relevant).

7.1 Partielle Differenzierbarkeit

Definition Es seien eine Koordinatennummer $j \in \{1, \ldots, w\}$ für die Wertemenge und eine Koordinatennummer $k \in \{1, \ldots, d\}$ für die Definitionsmenge fixiert.

Die j–te Koordinatenfunktion $f_j : X \to \mathbb{R}$ heißt in a *partiell differenzierbar* nach der k–ten Variablen x_k, wenn die partielle Funktion

$$f_j(a_1, \ldots, a_{k-1}, \underset{\substack{\uparrow \\ k}}{\cdot}, a_{k+1}, \ldots, a_d) : \begin{cases} X_k & \to \quad \mathbb{R}^w \\ \quad x & \mapsto \quad f_j(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_d) \end{cases}$$

an der Stelle a_k (wie gewöhnlich) differenzierbar ist. Die Ableitung heißt *partielle Ableitung der j–ten Koordinatenfunktion nach der k–ten Variablen* und wird notiert als:

$$\frac{\partial f_j}{\partial x_k}(a) \;:=\; \frac{d}{dx} f_j(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_d)\Big|_{x=a_k}.$$

Die Funktion f heißt *partiell differenzierbar auf X*, wenn die obige Definition für alle Stellen in X und für alle Variablen $j \in \{1, \ldots, w\}$ und $k = 1, \ldots, d$ zutrifft. In diesem Fall bilden die partiellen Ableitungen $\frac{\partial f_j}{\partial x_k} : X \to \mathbb{R}$ einen Satz von $w \cdot d$ Funktionen auf X.

Beispiel Wir betrachten die Funktion

$$f \begin{cases} \mathbb{R}^+ \times \mathbb{R} & \to \quad \mathbb{R}^2 \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \mapsto \quad \begin{pmatrix} x_1^2\, x_2 + \sin(x_1) \cdot e^{x_2} \\ x_1^{x_2} \end{pmatrix} \end{cases}$$

Unter Beachtung von $x_1^{x_2} = e^{x_2 \cdot \ln x_1}$ kann man berechnen:

$$\frac{\partial f_1}{\partial x_1}(x_1, x_2) \;=\; 2x_1\, x_2 + \cos(x_1) \cdot e^{x_2} \qquad \frac{\partial f_1}{\partial x_2}(x_1, x_2) \;=\; x_1^2 + \sin(x_1) \cdot e^{x_2}$$

$$\frac{\partial f_2}{\partial x_1}(x_1, x_2) \;=\; x_2 \cdot x_1^{x_2 - 1} \qquad \frac{\partial f_2}{\partial x_2}(x_1, x_2) \;=\; \ln(x_1) \cdot x_1^{x_2}$$

7.2 Totale Differenzierbarkeit

Definition f heißt *(total) differenzierbar an der Stelle* a mit der *Jacobi–Matrix* (oder *Funktional–Matrix*)

$$J_a = J = \begin{pmatrix} J_{11} & \cdots & J_{1d} \\ \vdots & & \vdots \\ J_{w\,1} & \cdots & J_{w\,d} \end{pmatrix} \in \mathbb{R}^{w\times d},$$

als *Ableitung*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für $\|x-a\| < \delta$ gilt:

$$\|f(x) - f(a) - J\cdot(x-a)\| < \varepsilon \cdot \|x-a\|.$$

Wenn f an allen Stellen von X differenzierbar ist, dann heißt f *differenzierbar* schlechthin. In diesem Fall ist die Abbildung

$$Df : \begin{cases} X & \to & \mathbb{R}^{w\times d} \\ a & \mapsto & J_a \end{cases}$$

definiert.

In dieser Definition nehmen wir Bezug auf die Lineare Algebra: J ist eine $w\times d$–Matrix mit w Zeilen und d Spalten. Sie wird mit dem Vektor $x - a \in \mathbb{R}^d$ multipliziert, das Ergebnis ist ein Vektor aus $\mathbb{R}^w$:

$$J\cdot(x-a) = \begin{pmatrix} J_{11} & \cdots & J_{1d} \\ \vdots & & \vdots \\ J_{w\,1} & \cdots & J_{w\,d} \end{pmatrix} \cdot \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_d - a_d \end{pmatrix} :=$$

$$\begin{pmatrix} J_{11}\cdot(x_1 - a_1) + \ldots + J_{1d}(x_d - a_d) \\ \vdots \\ J_{w\,1}\cdot(x_1 - a_1) + \ldots + J_{w\,d}(x_d - a_d) \end{pmatrix}.$$

Beispiele (1) Ist $d = w = 1$, so zeigt ein Vergleich der definierenden Ungleichung mit der in Satz 2, dass dann der neue Differenzierbarkeitsbegriff mit dem alten übereinstimmt. Die „alte" Ableitung $f'(a)$ ist als 1×1–Matrix ein Spezialfall der „neuen" Ableitung $Df(a)$.

(2) Ist $F : \mathbb{R}^d \to \mathbb{R}^w$ eine lineare Abbildung, die dann als Multiplikation mit der Matrix F aufgefasst werden kann, so ist

$$\|F(x) - F(a) - F\cdot(x-a)\| = 0 \qquad \text{für alle } x, a \in \mathbb{R}^d.$$

Damit ist die Differenzierbarkeit an allen Stellen a gegeben. Die Ableitungsfunktion ist konstant: $DF(a) = F$ für alle $a \in \mathbb{R}^d$.

Weitere Beispiele folgen nach den nächsten Sätzen.

Der folgende Satz gestattet es, die (totale) Differenzierbarkeit ohne Rückgriff auf Matrixrechnung zu fassen:

Satz 60 *(i) Die beiden folgenden Aussagen sind äquivalent:*

(A) Die Funktion $f : X \to \mathbb{R}^w$ *ist differenzierbar in* $a \in X$ *mit Ableitung* $J \in \mathbb{R}^{w \times d}$.

(B) Jede der w *Koordinatenfunktionen* $f_j : X \to \mathbb{R}$ *ist differenzierbar in* a *mit Ableitung*

$$\begin{pmatrix} J_{j1} & \cdots & J_{jd} \end{pmatrix} \qquad (j\text{–te Zeile von } J)$$

(ii) Die beiden folgenden Aussagen sind äquivalent:

(A) Die Funktion $f : X \to \mathbb{R}$ *(also* $w = 1$*) ist differenzierbar in* $a \in X$ *mit Ableitung*

$$J = \begin{pmatrix} J_1 & \cdots & J_d \end{pmatrix} \in \mathbb{R}^{1 \times d} \qquad (\text{Zeilenvektor})\ .$$

(B) Es gibt d *Zahlen* $J_1, \ldots, J_d$*, so dass für jedes* $\varepsilon > 0$ *ein* $\delta > 0$ *existiert mit*

$$|f(x) - f(a) - \underbrace{[J_1(x_1 - a_1) + \ldots + J_d(x_d - a_d)]}|< \varepsilon \cdot \|x - a\|, \ \text{falls } x \in U_\delta(a)\ .$$

Bemerkungen (1) Für den Fall einer Kurve $\gamma : \mathbb{R}^1 \to \mathbb{R}^w$ zeigt die Äquivalenz von (i), dass die Definition der Differenzierbarkeit in Abschnitt 6.1 mit der aktuellen übereinstimmt.

(2) Der Ausdruck oberhalb der geschweiften Klammer in Aussage (B) von (ii) kann auch als Skalarprodukt aufgefasst werden:

$$\Big[J_1(x_1 - a_1) + \ldots + J_d(x_d - a_d)\Big] = \langle \begin{pmatrix} J_1 \\ \vdots \\ J_d \end{pmatrix}, \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_d - a_d \end{pmatrix} \rangle.$$

Wenn man herausstellen will, dass das Skalarprodukt involviert ist, benutzt man die Schreibweise $\operatorname{grad} f$ oder ∇f anstelle von Df. für die als Spaltenvektor aufzufassende Ableitung. Die Unterscheidung von Matrixmultiplikation und Skalarprodukt scheint überflüssig, sie gewinnt aber an Relevanz, wenn man andere Skalarprodukte als das Standard–Skalarprodukt betrachtet oder Variablentransformationen in der Definitionsmenge $X \subseteq \mathbb{R}^d$ durchführt.

Der folgende Satz stellt eine Verbindung zwischen partieller und totaler Differenzierbarkeit her.

Satz 61 *Betrachte die folgenden Aussagen:*

(A) Alle Koordinatenfunktionen f_j $(j = 1, \ldots, w)$ *sind auf einer offenen Umgebung* $U_\varepsilon(a)$ *von* a *(nach allen Variablen) partiell differenzierbar und alle partiellen Ableitungen*

$$\frac{\partial f_j}{\partial x_k}, \quad j = 1, \ldots, w, \quad k = 1, \ldots, d$$

sind auf dieser Umgebung stetig.

(B) f *ist in* a *total differenzierbar.*

(C) Alle Koordinatenfunktionen f_j $(j = 1, \ldots, w)$ sind in a (nach allen Variablen) partiell differenzierbar.

(D) f ist in a Lipschitz–stetig, d.h. es existiert ein $\delta > 0$ und eine Konstante $L \geq 0$, so dass

$$|f(x) - f(a)| \leq L \cdot \|x - a\| \quad \text{für alle } x \in U_\delta(a).$$

(E) f ist stetig in a.

Dann gelten die Implikationen:

$$\begin{array}{ccccc} (A) & \Longrightarrow & (B) & \Longrightarrow & (C) \\ & & \Downarrow & & \\ & & (D) & \Longrightarrow & (E) \end{array}$$

Ist (B) erfüllt, so gilt

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_d}(a) \\ \vdots & & \vdots \\ \frac{\partial f_w}{\partial x_1}(a) & \cdots & \frac{\partial f_w}{\partial x_d}(a) \end{pmatrix}$$

Beweis Der Beweis von $(B) \Longrightarrow (D) \Longrightarrow (E)$ ist fast gleich wie der des entsprechenden Satzes 3 aus der 1–dimensionalen Differentialrechnung. Man muss nur die Zahlen x, a und $f'(a)$ durch die Vektoren x, a bzw. die Matrix $Df(a)$, die Betragsfunktion $|\cdot|$ durch die Normfunktion $\|\cdot\|$ ersetzen. Schließlich muss die Gleichung

$$|f'(a) \cdot (x - a)| = |f'(a)| \cdot |x - a|$$

durch die Ungleichung

$$\|Df(a) \cdot (x - a)\| \leq \|Df(a)\| \cdot \|x - a\|$$

ersetzt werden. Für $d = 1$ oder $w = 1$ ist ja Df_a ein Spaltenvektor bzw. ein Zeilenvektor, in diesem Fall ist die gerade die aus LAL02 bekannte Cauchy–Schwarz–Ungleichung. Der allgemeinere Fall $d \geq 2$ und $w \geq 2$ ist in LAL02 nicht behandelt, aber nicht prinzipiell schwieriger.

$(B) \Longrightarrow (C)$ Wir zeigen die Aussage nur für $d = 2$ und stellvertretend für $j = k = 1$. Dazu sei $\varepsilon > 0$ vorgegeben. Es gibt nach Voraussetzung eine $w \times 2$–Matrix J und $\delta > 0$, so dass für $x \in U_\delta(a)$

$$\|f(x_1, x_2) - f(a_1, a_2) - J \cdot \begin{pmatrix} x_1 - a_1 \\ x_2 - a_2 \end{pmatrix}\| < \varepsilon \cdot \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\|$$

Da der Betrag einer Koordinate eines Vektors immer kleiner oder gleich der Norm ist, gilt weiter

$$|f_1(x_1, x_2) - f_1(a_1, a_2) - \Big[J_{11}(x_1 - a_1) + J_{12}(x_2 - a_2)\Big]| < \\ \varepsilon \cdot \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\|$$

Dann gilt aber speziell für alle $z \in \mathbb{R}$ mit $\begin{pmatrix} z \\ a_2 \end{pmatrix} \in U_\delta(a)$

$$|f_1(z,a_2) - f_1(a_1,a_2) - J_{11}(z-a_1)| < \varepsilon \cdot \left\| \begin{pmatrix} z \\ a_2 \end{pmatrix} - \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \right\| = \varepsilon |z - a_1|.$$

Das aber bedeutet gerade, dass die partielle Funktion $f_1(\cdot\,, a_2)$ in a_1 differenzierbar ist mit Ableitung J_{11}. Wegen der Eindeutigkeit der Ableitung gilt:

$$\frac{\partial f_1}{\partial x_1}(a_1,a_2) = J_{11}.$$

Nun zu $(A) \Longrightarrow (B)$. Das wesentliche an diesem Beweis tritt schon beim Fall $d = 2$ und $w = 1$ auf. Es sei wieder ein $\varepsilon > 0$ vorgegeben. Wähle ein $\delta > 0$, so dass das folgende gilt:

- Die partielle Ableitung $\frac{\partial f}{\partial x_1}$ ist stetig auf $U_\delta(a)$,
- $\left| \frac{\partial f}{\partial x_1}(x) - \frac{\partial f}{\partial x_1}(a) \right| < \frac{\varepsilon}{2}, \qquad$ falls $x \in U_\delta(a)$,
- $\left| f(a_1,x_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_2}(a_1,a_2) \cdot (x_2 - a_2) \right| < \frac{\varepsilon}{2} \cdot |x_2 - a_2|, \qquad$ falls $|x_2 - a_2| < \delta$.

Es sei jetzt $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in U_\delta(a)$ beliebig. Wir wenden den MWS der Differentialrechnung auf die partielle Funktion $f(\cdot\,, x_2)$ an: Es gibt ein ξ zwischen x_1 und a_1, so dass

$$f(x_1,x_2) - f(a_1,x_2) = \frac{\partial f}{\partial x_1}(\xi, x_2) \cdot (x_1 - a_1).$$

Wegen $\begin{pmatrix} \xi \\ x_2 \end{pmatrix} \in U_\delta(a)$ gilt dann weiter

$$\begin{aligned}
&\left| f(x) - f(a) - [\frac{\partial f}{\partial x_1}(a) \cdot (x_1 - a_1) + \frac{\partial f}{\partial x_2}(a) \cdot (x_2 - a_2)] \right| \le \\
&\left| \overbrace{f(x_1,x_2) - f(a_1,x_2)}^{\frac{\partial f}{\partial x_1}(\xi,x_2)\cdot(x_1-a_1)} - \frac{\partial f}{\partial x_1}(a_1,a_2) \cdot (x_1 - a_1) \right| + \\
&\left| f(a_1,x_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_2}(a_1,a_2) \cdot (x_2 - a_2) \right| < \\
&\frac{\varepsilon}{2} \cdot |x_1 - a_1| + \frac{\varepsilon}{2} \cdot |x_2 - a_2| \le \frac{\varepsilon}{2} \cdot \|x - a\| + \frac{\varepsilon}{2} \cdot \|x - a\| = \varepsilon \cdot \|x - a\|.
\end{aligned}$$

♦

Beispiel Die Implikation $(A) \Longrightarrow (B)$ zeigt, dass die Funktion f aus dem Beispiel in Abschnitt 7.1 auf X total differenzierbar ist mit Ableitung

$$\begin{aligned}
Df(x_1,x_2) &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_1,x_2) & \frac{\partial f_1}{\partial x_2}(x_1,x_2) \\ \frac{\partial f_2}{\partial x_1}(x_1,x_2) & \frac{\partial f_2}{\partial x_2}(x_1,x_2) \end{pmatrix} \\
&= \begin{pmatrix} 2x_1\,x_2 + \cos(x_1) \cdot e^{x_2} & x_1^2 + \sin(x_1) \cdot e^{x_2} \\ x_2 \cdot x_1^{x_2 - 1} & \ln(x_1) \cdot x_1^{x_2} \end{pmatrix}.
\end{aligned}$$

Satz 62 (Kettenregel)
Es seien $f : X \to \mathbb{R}^w$ und $g : E \to \mathbb{R}^v$ reelle Funktionen mit $f(X) \subseteq E \subseteq \mathbb{R}^w$.

Ist f in $a \in X$ differenzierbar und g in $b = f(a)$ differenzierbar, so ist auch $g \circ f : X \to \mathbb{R}^v$ differenzierbar mit Ableitung

$$\underbrace{D(g \circ f)(a)}_{v \times d} = \underbrace{Dg(b)}_{v \times w} \cdot \underbrace{Df(a)}_{w \times d}.$$

Unterhalb sieht man, dass die Matrixdimensionen genau den Dimensionen von Definitions– und Wertemengen entsprechen.

Beweis Der Beweis besteht wieder in einer geeigneten Modifizierung des Beweises von Satz 6 über die 1–dimensionale Kettenregel.

Dabei wird noch die Cauchy–Schwarz–Ungleichung für die Norm von Matrizen

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|,$$

die in LAL02 nicht behandelt wurde, eine Rolle. ♦

7.3 Extrema und Sattelpunkte

Definition Es sei $f : X \to \mathbb{R}$ eine reellwertige Funktion und $a \in X$.
(1) Der Punkt $(a, f(a))$ des Graphen heißt wie in der ersten Spalte angegeben, wenn es ein $\varepsilon > 0$ gibt, so dass die Bedingung der zweiten Spalte gilt:

lokales Maximum	$f(x) \leq f(a)$ für alle $x \in U_\varepsilon(a) \cap X$
lokales Minimum	$f(x) \geq f(a)$ für alle $x \in U_\varepsilon(a) \cap X$
strenges lokales Maximum	$f(x) < f(a)$ für alle $x \in U_\varepsilon(a) \cap X$ mit $x \neq a$
strenges lokales Minimum	$f(x) > f(a)$ für alle $x \in U_\varepsilon(a) \cap X$ mit $x \neq a$

(2) Der Punkt $(a, f(a))$ des Graphen heißt wie in der ersten Spalte angegeben, wenn die Bedingung der zweiten Spalte erfüllt ist:

globales Maximum	$f(x) \leq f(a)$ für alle $x \in X$
globales Minimum	$f(x) \geq f(a)$ für alle $x \in X$
strenges globales Maximum	$f(x) < f(a)$ für alle $x \in X$ mit $x \neq a$
strenges globales Minimum	$f(x) > f(a)$ für alle $x \in X$ mit $x \neq a$

(3) Man spricht jeweils von einem Extremum, wenn es sich um ein Maximum oder Minimum handelt.

(4) In diesem Zusammenhang heißt a auch Stelle des Maximums, Minimums oder Extremums und $f(a)$ der Wert des Maximums, Minimums oder Extremums.

Satz 63 *Hat die differenzierbare Funktion $f : X \to \mathbb{R}$ bei dem inneren Punkt $a \in X$ ein Extremum, so gilt*

$$Df(a) = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Dabei heißt a ein innerer Punkt, wenn ein $\varepsilon > 0$ existiert, so dass $U_\varepsilon(a) \subseteq X$.

Beweis Da a ein innerer Punkt ist, gibt es ein $\varepsilon > 0$, so dass die partielle Funktion

$$f(a_1, \ldots, a_{j-1}, \underset{\substack{\uparrow \\ j}}{\cdot}, a_{j+1}, \ldots, a_d) : \begin{cases}]a_j - \varepsilon, a_j - \varepsilon[, & \to & \mathbb{R} \\ x & \mapsto & f(a_1, \ldots, a_{j-1}, x, a_{j+1}, \ldots, a_d) \end{cases}$$

wohl–definiert ist. Sie hat ein Extremum bei a_j, deshalb ist aufgrund von Satz 7 $\frac{\partial f}{\partial x_j}(a) = 0$.
Da dies für alle $j \in \{1, \ldots, d\}$ gilt, ist aufgrund von Satz 61

$$Df(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \frac{\partial f}{\partial x_2}(a) & \cdots & \frac{\partial f}{\partial x_d}(a) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}.$$

♦

Definition Es sei $X \subseteq \mathbb{R}^2$ und $a \in X$ ein innerer Punkt. Man sagt, die Funktion $f : X \to \mathbb{R}$ hat bei a einen *(strengen) Sattelpunkt*, wenn es zwei reguläre Kurven

$$\alpha, \beta :]-\varepsilon, +\varepsilon[\to X \qquad \text{mit } \alpha(0) = \beta(0) = a$$

gibt, so dass

- die zusammengesetzte Funktion $f \circ \alpha :]-\varepsilon, +\varepsilon[\to \mathbb{R}$ bei 0 ein strenges Maximum hat,
- die zusammengesetzte Funktion $f \circ \beta :]-\varepsilon, +\varepsilon[\to \mathbb{R}$ bei 0 ein strenges Minimum hat,
- die beiden Vektoren $\alpha'(0), \beta'(0)$ linear unabhängig in $\mathbb{R}^2$ sind.

Die dritte Bedingung ist nicht automatisch durch die ersten beiden erfüllt, wie man erwarten könnte. Als Beispiel diene

$$f(x_1, x_2) = -x_1^6 + x_1^2 \cdot x_2$$

Die Kurven

$$\alpha(t) = \begin{pmatrix} t \\ 0 \end{pmatrix} \qquad \beta(t) = \begin{pmatrix} t \\ t^2 \end{pmatrix}$$

haben beide den Tangentialvektor $\alpha'(0) = \beta'(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und es gilt

$$(f \circ \alpha)(t) = -t^6 \qquad \text{(Maximum)}$$

$$(f \circ \beta)(t) = -t^6 + t^4 = t^4(1 - t^2) = \begin{cases} > 0, & \text{falls } 0 < |t| < 1, \\ = 0, & \text{falls } t = 0. \end{cases} \qquad \text{(Minimum)}$$

Zeichnung

Satz 64 *Hat die Funktion $f : X \to \mathbb{R}$ bei a einen (strengen) Sattelpunkt und ist sie da differenzierbar, so ist die Ableitung Null:*

$$Df(a) = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

Beweis Es seien α, β die beiden Kurven, die die Sattelpunkteigenschaft hervorbringen. Aufgrund der Kettenregel Satz 62 ist das folgende 2×2–Gleichungssystem für die beiden Koordinaten in dem Zeilenvektor $f'(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \frac{\partial f}{\partial x_2}(a) \end{pmatrix}$ erfüllt:

$$\frac{\partial f}{\partial x_1}(a) \cdot \alpha_1'(0) + \frac{\partial f}{\partial x_2}(a) \cdot \alpha_2'(0) = f'(\alpha(0)) \cdot \alpha'(0) \overset{KR}{=} (f \circ \alpha)'(0) = 0$$

$$\frac{\partial f}{\partial x_1}(a) \cdot \beta_1'(0) + \frac{\partial f}{\partial x_2}(a) \cdot \beta_2'(0) = f'(\beta(0)) \cdot \beta'(0) \overset{KR}{=} (f \circ \beta)'(0) = 0$$

Dass der jeweils letzte Ausdruck verschwindet, liegt daran, dass die Funktionen $f \circ \alpha$ und $f \circ \beta$ in 0 jeweils ein Extremum haben. Da die beiden Vektoren $\alpha'(0), \beta'(0)$ linear unabhängig in $\mathbb{R}^2$ sind, hat das LGS eine eindeutige Lösung, nämlich

$$Df(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \frac{\partial f}{\partial x_2}(a) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

♦

8 Einfache Gewöhnliche Differentialgleichungen

8.1 Beispiele

Beispiel 1: Betrachte die (Differential–)Gleichung:

$$y' = -x \cdot y.$$

Gesucht ist eine Funktion y mit Variabler x, die diese Gleichung erfüllt. Überpüfen Sie (mit Hilfe der Kettenregel), dass die *Gauß'sche Glockenfunktion*

$$y(x) = e^{-\frac{x^2}{2}}$$

eine Lösung darstellt. Gibt es noch andere Lösungen?

Beispiel 2: Es seien K und R zwei positive Konstanten. Die Differentialgleichung

$$\ddot{x} = -Kx - R\dot{x}$$

beschreibt die Bewegung eines Körpers, der von einer Schraubenfeder angetrieben wird und dabei einer Reibung unterliegt.

- Der (doppelte) Punkt über dem x bedeutet die erste (bzw. zweite) Ableitung nach der Zeitvariablen t. Die Ausdrücke $\dot{x}$ bzw. $\ddot{x}$ stehen also für die Geschwindigkeit bzw. Beschleunigung des Körpers.
- Die Beschleunigung (Bremsung) $\ddot{x}$ (linke Seite) wird hervorgerufen durch die Federkraft $-Kx$ (Gemäß Hooke'schem Gesetz ist sie proportional zur Dehnung x der Feder) und durch die Reibung $-R\dot{x}$ (sie ist proportional zur Geschwindigkeit).

Wir probieren einen Ansatz:

$$\begin{aligned} x(t) &= e^{-\varrho t} \cdot \sin(\omega \cdot t) \\ \dot{x}(t) &= -\varrho e^{-\varrho t} \cdot \sin(\omega \cdot t) + \omega e^{-\varrho t} \cdot \cos(\omega \cdot t) \\ \ddot{x}(t) &= \varrho^2 e^{-\varrho t} \cdot \sin(\omega \cdot t) - 2\varrho\omega e^{-\varrho t} \cdot \cos(\omega \cdot t) - \omega^2 e^{-\varrho t} \cdot \sin(\omega \cdot t) \end{aligned}$$

Setzen wir diese Funktion und ihre Ableitungen in die Differentialgleichung ein, so ergibt sich

$$\varrho^2 - \omega^2 = -K + R\varrho \qquad -2\varrho\omega = -R\omega^2$$

und daraus

$$\varrho = \frac{R}{2} \qquad \omega = \pm\sqrt{K - \frac{R^2}{4}}.$$

Weisen Sie direkt nach, dass die gedämpfte Schwingung

$$x(t) = e^{-\frac{R}{2}} \cdot \sin\left(\sqrt{K - \frac{R^2}{4}} \cdot t\right)$$

eine Lösung der Differentialgleichung darstellt.

Gibt es noch andere Lösungen? Gibt es systematische Verfahren zur Lösung? Weiß man von jeder Differentialgleichung, ob es eine Lösung gibt?

8.2 Definition

Wir vereinbaren zunächst für dieses Kapitel, dass N eine feste natürliche Zahl, D eine offene Teilmenge von $\mathbb{R}^{1+N}$ ist. I sei ein offenes Intervall von $\mathbb{R}$.

Definition Ist $f : D \to \mathbb{R}$ eine stetige Funktion und $y : I \to \mathbb{R}$ eine N–mal differenzierbare Funktion, so heißt y *Lösung der Different(z)ialgleichung (N–ter Ordnung)*

$$y^{(N)} = f(x, y, y', y'', \ldots, y^{(N-1)})$$

wenn für alle $x \in I$ gilt:

$$y^{(N)}(x) = f(x, y(x), y'(x), y''(x), \ldots, y^{(N-1)}(x)).$$

(Erinnerung: $y^{(i)}$ bezeichnet die *i–te Ableitung* von y.) Die Funktion f heißt in diesem Zusammenhang *Rechte Seite*.

Eine Lösung $y : I \to \mathbb{R}$ heißt *maximal* oder mit *maximaler Definitionsmenge*, wenn es eine ,,umfassendere" weitere Lösung $\widetilde{y} : \widetilde{I} \to \mathbb{R}$ mit

$$I \subsetneqq \widetilde{I} \quad \text{und} \quad y(x) = \widetilde{y}(x) \text{ für } x \in I$$

nicht gibt.

Bemerkungen

- Eine Differentialgleichung beschreibt also die Aufgabe, bei gegebener ,,Rechter Seite" f eine unbekannte differenzierbare Funktion $x \mapsto y$ zu finden. Bei Einsetzung dieser Funktion und ihrer Ableitungen in die Rechte Seite soll auf der linken Seite die N–te Ableitung (linke Seite) herauskommen.

- Da auf der linken Seite die N–te Ableitung explizit (aufgelöst) auftritt, spricht man auch von einer *expliziten* Differentialgleichung (N–ter Ordnung). Beispielsweise ist

  $$\sin[(y')^3] + y^2 = \cos(x)$$

  keine explizite Differentialgleichung (erster Ordnung). Für die expliziten Differentialgleichungen existiert eine geschlossene Theorie. Nicht–explizite Differentialgleichungen kann man unter Umständen mit ad–hoc Methoden lösen.

- Es sei

  $$y^{(N)} = f(x, y, y', \ldots, y^{(N-1)})$$

  eine Differentialgleichung N–ter Ordnung und $(x_0, y_0, y_1, \ldots, y_{N-1})$ ein Punkt in der Definitionsmenge D. Stellt man an eine Lösung y der Differentialgleichung zusätzlich die *Anfangsbedingung*, dass

  $$y(x_0) = y_0, \qquad y'(x_0) = y_1, \qquad \ldots y^{(N-1)}(x_0) = y_{N-1},$$

  so spricht man von einem *Anfangswertproblem (AWP)*. Wir werden anhand einiger Beispiele sehen, dass ,,im Normalfall" AWPe genau eine Lösung besitzen.

- Eine (explizite) Differentialgleichung *erster Ordnung* hat die Form

  $$y' = f(x, y).$$

  Es ist eine differenzierbare Funktion $y : I \to \mathbb{R}$ mit $y'(x) = f(x, y(x))$ für alle $x \in I$ gesucht. Ein zugeordnetes AWP hat dann die Form

  $$\begin{cases} y' = f(x, y) \\ y_(x_0) = y_0 \end{cases}$$

- Oft werden als Variablen auch t statt x und x statt y verwendet. Der Buchstabe erinnert dann an ,,time (Zeit)". Der Ableitungsbeistrich wird bei dieser Notation durch einen Punkt ersetzt. Beispielsweise lauten dann explizite Differentialgleichungen erster bzw. zweiter Ordnung dann

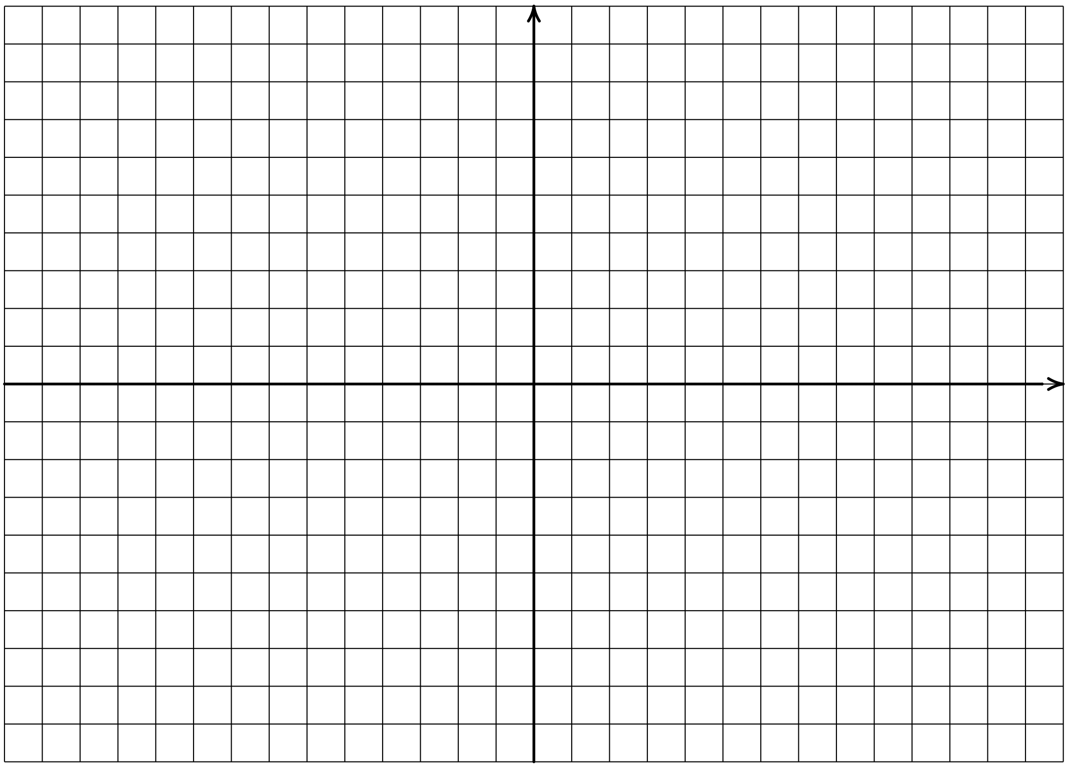
  $$\dot{x} = f(t, x) \qquad \text{bzw.} \qquad \ddot{x} = f(t, x, \dot{x}).$$

- Eine Differentialgleichung und ihre Lösung können auf die folgende Weise geometrisch interpretiert werden. Die Ableitung y' der gesuchten Funktion an einer Stelle x kann als Steigung der Tangente an den Graphen der Funktion aufgefasst werden. Diese Steigung ist also gerade als Funktion von x und y gegeben.

  In einem x–y–Koordinatensystem kann man die Zuordnung $(x, y) \mapsto y'(x)$ dadurch graphisch veranschaulichen, dass man an einigen (günstig ausgewählten) Punkten $(x, y) \in I \times \mathbb{R}$ kleine ,,Tangentchen", sogenannte *Linienelemente*, einzeichnet. Es entsteht ein *Richtungsfeld*. Hilfreich sind dabei *Isoklinen*, dass sind Kurven in der (x, y)–Ebene, entlang derer die $f(x, y)$ konstant ist. Das graphische Lösen der Differentialgleichungs–Aufgabe besteht dann darin, Funktionsgraphen zu finden, die sich an die Linienelemente anschmiegen.

Beispiel: Wir zeichnen das Richtungsfeld für die Differentialgleichung aus dem Anfangsbeispiel:

$$y' = -x \cdot y$$



8.3 Differentialgleichungen mit getrennten Variablen

Definition Eine (explizite) Differentialgleichung erster Ordnung heißt *mit getrennten Variablen*, wenn es zwei Funktionen $h : D_1 \to \mathbb{R}$, $g : D_2 \to \mathbb{R}$ (D_1, D_2 offene Intervalle) gibt, so dass die rechte Seite die Form

$$f(x,y) = h(x) \cdot g(y), \quad (x,y) \in D = D_1 \times D_2$$

annimmt.

Beispiel: Wir betrachten das AWP erster Ordnung

$$\begin{cases} y' = -x \cdot y \\ y(0) = 1. \end{cases} \qquad \Big\| \qquad \begin{cases} y' = h(x) \cdot g(y) \\ y(x_0) = y_0 \end{cases}$$

und lösen es mit einem formalen (mathematisch nicht abgesicherten) Verfahren. Wir schreiben die Differentialgleichung als

$$\frac{dy}{dx} = -x \cdot y \qquad \Big\| \qquad \frac{dy}{dx} = h(x) \cdot g(y)$$

und dann als Gleichung zwischen ,,Differentialen"

$$\frac{1}{y} \cdot dy = -x \cdot dx \qquad \Big\| \qquad \frac{1}{g(y)} \cdot dy = h(x) \cdot dx$$

die integriert werden können. Die Anfangsbedingung legt dabei Integrationsgrenzen fest:

$$\int_1^y \frac{1}{\hat{y}} \cdot d\hat{y} = \int_0^x (-\hat{x}) \cdot d\hat{x} \qquad \Big\| \qquad \int_{y_0}^y \frac{1}{g(\hat{y})} \cdot d\hat{y} = \int_{x_0}^x h(\hat{x}) \cdot d\hat{x}$$

(Der ,,Dach-Akzent" ist für die Unterscheidung von Integrationsvariablen und Integrationsgrenzen nötig).

Als Resultat erhalten wir:

$$\ln(y) - \ln(1) = -\frac{x^2}{2} + \frac{0^2}{2} \qquad \Big\| \qquad G(y) = H(x),$$

wobei G die Stammfunktion von $\frac{1}{g}$ mit $G(y_0) = 0$ und H die Stammfunktion h mit $H(x_0) = 0$ ist. Wir wenden – wenn möglich – die Umkehrfunktion der linken Seite an und erhalten:

$$y = e^{-\frac{x^2}{2}} \qquad \Big\| \qquad y = G^{-1}(H(x)).$$

Man rechnet leicht nach, dass die Funktion auf der linken Seite das zugehörige AWP löst. Das allgemeine Verfahren auf der rechten Seite beinhaltet aber problematische Vorgehensweisen:

- Falls $g(y) = 0$ an einer Stelle y ist, darf man nicht einfach durch $g(y)$ dividieren.
- Existiert überhaupt die Umkehrfunktion von G?

Der folgende Satz und sein Beweis stellen das Verfahren auf eine sichere Grundlage.

Satz 65 *Es seien D_1, D_2 offene Intervalle, $x_0 \in D_2$ und $h : D_1 \to \mathbb{R}$, $g : D_2 \to \mathbb{R}$ stetige Funktionen. Wir betrachten das AWP*

$$\begin{cases} y' = h(x) \cdot g(y) \\ y(x_0) = y_0. \end{cases}$$

(i) Ist $g(y_0) \neq 0$, so gibt es ein offenes Intervall I mit $x_0 \in I \subseteq D_1$ und eine Lösung $y : I \to \mathbb{R}$ des AWPs. Sie kann mit Hilfe des obigen Verfahrens konstruiert werden.

Die Lösung ist eindeutig in dem folgenden Sinne: Ist $\widetilde{y} : \widetilde{I} \to \mathbb{R}$ eine weitere Lösung, so stimmt sie auf dem Schnittintervall $I \cap \widetilde{I}$ mit y überein.

(ii) Ist $g(y_0) = 0$, so ist die konstante Funktion

$$y(x) \equiv y_0$$

eine Lösung des AWPs. Es kann auch andere Lösungen geben.

Beweis (1) Da $g(y_0) \neq 0$ ist, gibt es ein offenes Intervall $J_1 \subseteq D_2$ um y_0 mit $g(y) \neq 0$ für alle $y \in J_1$ (Vgl. Lemma 38, ANA01). Daher ist die Funktion $\frac{1}{g}$ auf J_1 definiert und stetig.

(2) $\frac{1}{g}$ ist stetig und besitzt daher gemäß HDI (Satz 31) eine Stammfunktion G auf J_1.

(3) Da $G'(y_0) = \frac{1}{g(y_0)} \neq 0$, gibt es ein offenes Intervall $J_2 \subseteq J_1$, $y_0 \in J_2$, auf dem G streng monoton, deshalb umkehrbar, ist (Sätze 9(ANA02), 42(ANA01)). Ist $J_3 = G(J_2)$ das Bildintervall, so haben wir die bijektive Zuordnung

$$J_2 \quad \underset{G^{-1}}{\overset{G}{\rightleftarrows}} \quad J_3.$$

Wegen $G(y_0) = 0$ gilt $0 \in J_3$ und $G^{-1}(0) = y_0$.

(4) Die Funktion h ist stetig auf D_1, sie besitzt daher eine Stammfunktion H mit $H(x_0) = 0$. Es sei I ein offenes Intervall in D_1 mit $x_0 \in I$ und $H(I) \subseteq J_3$.

(5) Dann ist die Funktion

$$y = G^{-1} \circ H : I \to \mathbb{R}$$

wohldefiniert. Wir zeigen, dass es sich um eine Lösung des AWPs handelt.

(6) Die Anfangsbedingung ist erfüllt

$$y(x_0) = G^{-1}(H(x_0)) = G^{-1}(0) = y_0.$$

(7) Wegen

$$y'(x) = (G^{-1})'(H(x)) \cdot H'(x) = \frac{1}{G'(G^{-1}(H(x))} \cdot H'(x) = g(y(x)) \cdot h(x), \qquad x \in I$$

(Kettenregel Satz 6, Satz 5) ist y auch Lösung der Differentialgleichung.

(8) Ist $\widetilde{y} : \widetilde{I} \to \mathbb{R}$ eine weitere Lösung des AWPs, so gilt für $x \in I \cap \widetilde{I}$:

$$H(x) = \int_{x_0}^{x} h(\hat{x}) d\hat{x} \stackrel{(*)}{=} \int_{x_0}^{x} \frac{\widetilde{y}'(\hat{x})}{g(\widetilde{y}(\hat{x}))} d\hat{x} \stackrel{(**)}{=} \int_{y_0}^{\widetilde{y}(x)} \frac{1}{g(z)}\, dz = G(\widetilde{y}(x)).$$

Die Umformung $(*)$ beruht darauf, dass $\widetilde{y}$ Lösung ist, in $(**)$ haben wir die Substitution $z = \widetilde{y}(x)$ durchgeführt. Die gesamte Gleichung besagt gerade, dass für $x \in I$

$$\widetilde{y}(x) = G^{-1}(H(x)) = y(x)$$

ist.

(ii) Die erste Aussage kann man direkt verifizieren. Die zweite Aussage werden wir weiter unten anhand des Beispiels 2 begründen. ♦

Die Differentialgleichung mit getrennte Variablen beinhaltet zwei interessante Sonderfälle ab:

- Bei der Differentialgleichung

  $$y' = h(x)$$

  hängt die rechte Seite gar nicht von der Funktion y ab. Als Lösung eines zugehörigen AWPs

  $$\begin{cases} y' = h(x) \\ y(x_0) = y_0 \end{cases}$$

  ergibt sich die Stammfunktion von h:

  $$y(x) = \int_{x_0}^{x} h(\hat{x})\, d\hat{x} + y_0$$

  Dies würde sich auch bei Anwendung des obigen Satzes auf diese Situation ergeben. Die Linienelemente haben in diesem Fall entlang vertikaler Isoklinen konstante Steigung.

- Eine Differentialgleichung

  $$y' = g(y)$$

  mit einer von x unabhängigen rechten Seite heißt *autonom*. Es gilt der folgende Satz:

  Satz 66 *Ist $y : I \to \mathbb{R}$ eine Lösung der obigen autonomen Differentialgleichung, so ist für jedes $\xi \in \mathbb{R}$ auch die verschobene („geshiftete") Funktion*

  $$\begin{cases} I + \xi & \to & \mathbb{R} \\ x & \mapsto & y(x - \xi) \end{cases}$$

  eine Lösung.

Dies ergibt sich aus der Kettenregel

$$\frac{d}{dx}y(x-\xi) = y'(x-\xi) = g(y(x-\xi)).$$

Dies wird auch bei Betrachtung des Richtungsfelds anschaulich klar. Es ist unabhängig von x. Deshalb kann der Graph einer Lösung in x–Richtung um ξ verschoben werden, er bleibt dabei Graph einer Lösung.

Die Lösung eines AWPs

$$\begin{cases} y' = g(y) \\ y(x_0) = y_0 \end{cases}$$

mit einer von x unabhängigen rechten Seite ergibt sich gemäß obigen Satz als

$$y(x) = G^{-1}(x - x_0),$$

wobei G^{-1} die Umkehrfunktion der Funktion $y \mapsto \int_{y_0}^{y} \frac{d\hat{y}}{g(\hat{y})}$ mit geeignetem Definitionsintervall ist.

Beispiele

1. Betrachte das AWP
$$\begin{cases} y' = y^2 \\ y(x_0) = y_0 \end{cases}$$
Für $y_0 = 0$ ist die Lösung trivial, im anderen Fall erhält man gemäß obigem Verfahren zunächst die Gleichung
$$\int_{y_0}^{y} \frac{d\hat{y}}{\hat{y}^2} = \int_{x_0}^{x} 1\, d\hat{x},$$
die mit Hilfe der Stammfunktionen in
$$-\frac{1}{y} + \frac{1}{y_0} = x - x_0$$
überführt wird. Als Lösung ergibt sich:
$$y(x) = \frac{1}{-x + x_0 + \frac{1}{y_0}}.$$
Diese Lösung hat für $x \to x_0 + \frac{1}{y_0}$ eine Polstelle.
2. Wir betrachten das AWP
$$\begin{cases} y' = 2\sqrt{|y|} \\ y(0) = 0 \end{cases}$$
und überzeugen uns davon, dass für je zwei reelle Zahlen $k \leq 0 \leq \ell$ die abschnittsweise durch Parabeläste definierte Funktion $y_{k,\ell} : \mathbb{R} \to \mathbb{R}$
$$y_{k,\ell}(x) = \begin{cases} (x-\ell)^2, & \text{falls } x > \ell, \\ 0, & \text{falls } k \leq x \leq \ell, \\ -(x-k)^2, & \text{falls } x < k, \end{cases}$$

eine Lösung des AWPs darstellt. Die Anfangsbedingung ist erfüllt. Dass die drei Zeilen die Differentialgleichung erfüllen, weisen wir beispielhaft für den „schwierigsten" Fall der dritten Zeile nach: Die Ableitung der unteren Zeile ist

$$\left[-(x-k)^2\right]' = -2(x-k) = 2|x-k| = 2\sqrt{|-(x-k)^2|},$$

entsprechendes gilt für die mittlere und obere Zeile. Wir haben also ein Beispiel eines AWPs gefunden, das unendlich viele Lösungen hat. Eine umfassendere Theorie der Differentialgleichungen klärt auf, das dies an der Nicht–Differenzierbarkeit der rechten Seite $g(y)$ an der Stelle y_0 liegt.

3. Zur Übung:

$$\begin{cases} y' = \frac{y^2}{x} \\ y(0) = 0 \end{cases}$$

Mit Hilfe der Methode der Trennung der Variablen kann man für eine weitere Klasse von Differentialgleichungen Lösungen auffinden. Sie haben die Form

$$y' = f(ax + by + c),$$

wobei $a, b, c \in \mathbb{R}$ Konstanten sind. O.B.d.A. kann $b \neq 0$ angenommen werden.

Ist $y : I \to \mathbb{R}$ eine reelle Funktion, so kann man ihr durch eine Transformation

$$u(x) := ax + by(x) \qquad y(x) = \frac{1}{b} \cdot (u(x) - ax)$$

eine andere Funktion $u : I \to \mathbb{R}$ bijektiv zuordnen. Ist y eine Lösung der obigen Differentialgleichung, so gilt

$$u'(x) = a + by'(x) = a + bf(ax + by(x) + c) = a + bf(u(x)),$$

u ist also Lösung der der transformierten Differentialgleichung,

$$u' = a + bf(u)$$

mit getrennten Variablen.
Als Beispiel betrachten wir das AWP

$$\begin{cases} y' = (x + y)^2 \\ y(0) = 0 \end{cases}$$

Mit $a = b = 1, c = 0$ kann sie der obigen Beispielklasse zugeordnet werden. Mit der Transformation $u(x) = x + y(x)$ erhalten wir das transformierte AWP

$$\begin{cases} u' = 1 + u^2 \\ u(0) = 0 \end{cases}$$

mit der Lösung $u(x) = \tan(x)$. Also ist $y(x) = \tan(x) - x$ eine Lösung des ursprünglichen AWPs.

8.4 Lineare Differentialgleichungen

Definition Eine (explizite) Differentialgleichung N–ter Ordnung heißt *linear*, wenn die Definitionsmenge der rechten Seite f die Form $D = I \times \mathbb{R}^N$, hat und dann die rechte Seite sich als

$$y^{(N)} = a_0(x)y + a_1(x)y' + \ldots + a_{N-1}(x)y^{(N-1)} + b(x) \qquad \text{(iLDG)}$$

darstellen lässt mit stetigen Funktionen $a_i : I \to \mathbb{R}$, $i = 0, \ldots, N-1$ und $b : I \to \mathbb{R}$. Wegen des Terms $b(x)$ nennt man die Differentialgleichung zusätzlich *inhomogen*. Wird die Funktion $b \equiv 0$ gesetzt, so spricht man von einer (zugehörigen) *homogenen* Differentialgleichung

$$y^{(N)} = a_0(x)y + a_1(x)y' + \ldots + a_{N-1}(x)y^{(N-1)}. \qquad \text{(hLDG)}$$

Allgemein lässt sich der folgende Satz über lineare Differentialgleichungen formulieren:

Satz 67 *(i) Superpositionsprinzip für homogene Differentialgleichungen: Sind die Funktionen y und $\widetilde{y}$ zwei Lösungen von (hLDG), so ist auch die Linearkombination*

$$\alpha y + \beta \widetilde{y}$$

eine Lösung. Die Lösungen bilden also einen $\mathbb{R}$–Vektorraum.

(ii) Ist $\widehat{y}$ eine fest vorgegeben Lösung der inhomogenen Differentialgleichung (iLDG), so lässt sich jede Lösung in der Form

$$y = \widehat{y} + \widetilde{y}$$

darstellen, wobei $\widetilde{y}$ eine Lösung der zugehörigen homogenen Differentialgleichung (hLDG) ist.

Beweis Das rechnet man jeweils einfach nach. Für (i):

$$\begin{aligned}
(\alpha y + \beta \widetilde{y})^{(N)}(x) &= \alpha y^{(N)}(x) + \beta \widetilde{y}^{(N)}(x) \\
&= \alpha \left[a_0(x)y(x) + a_1(x)y'(x) + \ldots + a_{N-1}(x)y^{(N-1)}(x)\right] + \\
&\quad\ \beta \left[a_0(x)\widetilde{y}(x) + a_1(x)\widetilde{y}'(x) + \ldots + a_{N-1}(x)\widetilde{y}^{(N-1)}(x)\right] \\
&= a_0(x)(y + \widetilde{y})(x) + a_1(x)(y + \widetilde{y})'(x) + \ldots + a_{N-1}(x)(y + \widetilde{y})^{(N-1)}(x).
\end{aligned}$$

Für (ii): Tatsächlich ist die Differenz zweier Lösungen y und $\widehat{y}$ der inhomogenen Differentialgleichung (LDG) eine Lösung der zugehörigen homogenen Differentialgleichung.

$$\begin{aligned}
(y - \widehat{y})^{(N)}(x) &= y^{(N)}(x) - \widehat{y}^{(N)}(x) \\
&= \left[a_0(x)y(x) + a_1(x)y'(x) + \ldots + a_{N-1}(x)y^{(N-1)}(x) + b(t)\right] - \\
&\quad\ \left[a_0(x)\widetilde{y}(x) + a_1(x)\widetilde{y}'(x) + \ldots + a_{N-1}(x)\widetilde{y}^{(N-1)}(x) + b(t)\right] = \\
&= a_0(x)(y - \widetilde{y})(x) + a_1(x)(y - \widetilde{y})'(x) + \ldots + a_{N-1}(x)(y - \widetilde{y})^{(N-1)}(x)
\end{aligned}$$

♦

Wir betrachten zwei Sonderfälle.

8.5 Lineare Differentialgleichungen erster Ordnung

Es geht also um Differentialgleichungen der Form

$$y' = a(x)y + b(x)$$

mit stetigen Funktionen $a, b : I \to \mathbb{R}$.

Wir definieren für $x_1, x_2 \in I$ die *Übergangsfunktion*

$$\Phi_a(x_1, x_2) \ := \ \exp\left(\int_{x_1}^{x_2} a(\hat{x})d\hat{x}\right)$$

und notieren einige Eigenschaften (Nachrechnen!) für $x_1, x_2, x_3, x, x_0 \in I$

$$\begin{aligned}
&\Phi_a(x_1, x_1) = 1, \\
&\Phi_a(x_1, x_2) \cdot \Phi_a(x_2, x_3) = \Phi_a(x_1, x_3), \\
&\Phi_a(x_1, x_2) = \frac{1}{\Phi_a(x_2, x_1)}. \\
&\Phi_a'(x, x_1) = a(x) \cdot \Phi_a(x, x_1) \qquad \text{(Ableitung nach } x\text{)}
\end{aligned}$$

Satz 68 (Lösungen der homogenen Differentialgleichung)

(i) Das AWP

$$\begin{cases} y' = a(x)y \\ y(x_0) = y_0 \end{cases}$$

hat die eindeutige Lösung auf I

$$y(x) = \Phi_a(x, x_0) \cdot y_0.$$

(ii) Die Menge aller Lösungen der Differentialgleichung

$$y' = a(x)y$$

ist gegeben durch

$$y(x) = \Phi_a(x, x_0) \cdot C, \qquad C \in \mathbb{R}.$$

Beweis Der Satz 65 („getrennte Variablen") würde es erlauben, die obige Lösung aufzufinden. Da wir aber die gesuchte Lösung schon zur Verfügung haben, rechnen wir die Lösungseigenschaft einfach nach:

$$y'(x) \ = \ \Phi'(x, x_0) \cdot y_0 = a(x) \cdot \Phi(x, x_0) \cdot y_0 = a(x) \cdot y(x).$$

Die Anfangsbedingung ist ebenfalls erfüllt.

Es bleibt noch die Eindeutigkeit zu zeigen: Dazu sei $y_0 \neq 0$, was bedeutet, dass y nirgends in I den Wert 0 annimmt. Ist $\widetilde{y} : \widetilde{I} \to \mathbb{R}$ mit $\widetilde{I} \subseteq I$ eine weitere Lösung des AWPs, so gilt für die Funktion $\frac{\widetilde{y}}{y}$ gemäß Quotientenregel

$$\left(\frac{\widetilde{y}}{y}\right)'(x) = \frac{\widetilde{y}'(x)y(x) - y(x)\widetilde{y}(x)}{y^2(x)} = \frac{a(x)\widetilde{y}(x) \cdot a(x)y(x) - a(x)y(x) \cdot a(x)\widetilde{y}(x)}{y^2(x)} = 0.$$

Wegen $\left(\frac{\widetilde{y}}{y}\right)(x_0) = \frac{y_0}{y_0} = 1$ muss nach Folgerung 13 die Funktion $\frac{\widetilde{y}}{y} \equiv 1$ sein. $\widetilde{y}$ stimmt also mit y überein.

Für $y_0 = 0$ ist $y \equiv 0$ die einzige Lösung. Eine andere Lösung $\widehat{y}$ wäre in einem offenen Intervall $\widehat{I} \subseteq I$ ungleich Null. Für sie könnte man dann die vorige Überlegung anwenden.

♦

Um eine Lösung für die inhomogene Differentialgleichung

$$y' = a(x)\,y + b(x)$$

zu finden, wendet man das Prinzip der

Variation der Konstanten
der Lösung der zugehörigen homogenen Differentialgleichung

an. Dies bedeutet dass man als Ansatz für die Lösung der inhomogenen Differentialgleichung die Konstante C in der Lösungsformel 68 (ii) durch eine Funktion $C(x)$ ersetzt:

$$y(x) = \Phi_a(x, x_0) \cdot C(x).$$

Man setzt diesen Ansatz in (iLDG) ein und erhält mit der Produktregel

$$\begin{aligned} a(x)\,y(x) + b(x) &= y'(x) \\ &= \Phi(x, x_0) \cdot C'(x) + a(x) \cdot \Phi_a(x, x_0) \cdot C(x) \\ &= a(x) \cdot y(x) + \Phi(x, x_0) \cdot C'(x). \end{aligned}$$

was auf die Gleichung

$$C'(x) = \Phi_a(x_0, x) \cdot b(x)$$

führt, die dann nur noch integriert werden muss. Die Integrations–Konstante ist dabei durch die Anfangsbedingung festgelegt:

$$C(x) = \int_{x_0}^{x} \Phi_a(x_0, \hat{x}) \cdot b(\hat{x})\, d\hat{x} + y_0$$

Die gesamt Lösung kann also in einem Satz notiert werden.

Satz 69 (Lösungen der inhomogenen Differentialgleichung)

(i) Für die eindeutige Lösung des inhomogenen AWPs

$$\begin{cases} y' = a(x)y + b(x) \\ y(x_0) = y_0 \end{cases}$$

gilt

$$\begin{aligned} y(x) &= \Phi_a(x, x_0) \cdot \left[\int_{x_0}^{x} \Phi_a(x_0, \hat{x}) \cdot b(\hat{x})\, d\hat{x} + y_0\right] \\ &= \int_{x_0}^{x} \Phi_a(x, \hat{x}) \cdot b(\hat{x})\, d\hat{x} + \Phi_a(x, x_0) \cdot y_0. \end{aligned}$$

(Beachte die Besonderheit, dass die Variable x sowohl als Integrationsgrenze als auch im Integranden auftritt.)

(ii) Die Gesamtheit aller Lösungen des AWPs ist gegeben durch

$$y(x) = \int_{x_0}^{x} \Phi_a(x, \hat{x}) \cdot b(\hat{x})\, d\hat{x} + \Phi_a(x, x_0) \cdot C, \qquad C \in \mathbb{R}.$$

Der Ausdruck (ii) gibt die Beobachtung aus Satz 67 wieder. Die allgemeine Lösung der inhomogenen Differentialgleichung setzt sich aus einer speziellen Lösung der inhomogenen Differentialgleichung und der allgemeinen Lösung der homogenen Differentialgleichung zusammen.

8.6 Exkurs: Nullstellen von Polynomen

Es sei

$$p(z) = a_N z^N + a_{N-1} z^{N-1} + \ldots + a_1 z + a_0$$

ein Polynom mit Koeffizienten $a_N, \ldots, a_1, a_0 \in \mathbb{C}$. Im Fall $a_N \neq 0$ hat es den Grad N, es kann dann

$$p(z) = a_N \cdot (z^N + \frac{a_{N-1}}{a_N} z^{N-1} + \ldots + \frac{a_1}{a_N} z + \frac{a_0}{a_N})$$

geschrieben werden. Der Koeffizient von z^N in dem Polynom in Klammern ist 1. Ein solches Polynom heißt *normiert*. Da dieser Übergang zum normierten Polynom die Nullstellen nicht verändert, beschränken wir uns im folgenden nur auf solche normierten Polynome.

Satz 70 (und Definition) *Es sei*

$$p(z) = z^N + \frac{a_{N-1}}{a_N} z^{N-1} + \ldots + \frac{a_1}{a_N} z + \frac{a_0}{a_N}$$

ein normiertes Polynom N–ten Grades und $1 \leq \ell \leq N$. Die folgenden Aussagen über eine Zahl $\lambda \in \mathbb{C}$ sind äquivalent:

(i) (Def) λ ist eine ℓ–fache Nullstelle des Polynoms.

(ii) λ ist Nullstelle der Polynome $p, p', p'', \ldots, p^{(\ell-1)}$.

(iii) Es gibt ein normiertes Polynom q_ℓ vom Grad $N - \ell$, so dass

$$p(z) = (z - \lambda)^\ell \cdot q_\ell(z).$$

Anders formuliert: Aus p können ℓ Linearfaktoren $(z - \lambda)$ abgespalten *werden.*

Beweis (iii) $\Longrightarrow$ (ii) Hat p die Darstellung von (iii), so zeigt die Leibniz–Produktregel, dass in jeder Ableitung $p^{(j)}$ für $j = 0, \ldots, \ell - 1$ ein Faktor $(z - \lambda)$ auftritt. Deshalb ist $p^{(j)}(\lambda) = 0$.

(ii) $\Longrightarrow$ (iii) Wir zeigen mit Induktion über $j = 1, \ldots, \ell$: Aus p können j Linearfaktoren $(z - \lambda)$ *abgespalten* werden. Induktionsanfang $j = 1$: Es gilt

$$\begin{aligned} p(z) &= p(z) - p(\lambda) = \\ &= a_1(z - \lambda) + \ldots + a_{N-1}(z^{N-1} - \lambda^{N-1}) + (z^N - \lambda^N) \end{aligned}$$

Für $k = 1, \ldots, N$ ist aber weiter

$$z^k - \lambda^k = (z - \lambda) \cdot (\underbrace{z^{k-1} + z^{k-2}\lambda^1 + z^{k-2}\lambda^2 + \ldots + z\lambda^{k-2} + \lambda^{k-1}}_{=: e_k(z,\lambda)})$$

und deshalb

$$p(z) = (z - \lambda) \cdot (\underbrace{a_1 + a_2 e_2(z,\lambda) + \ldots + a_{N-1} e_{N-1}(z,\lambda) + e_N(z,\lambda)}_{=:q_1(z)}).$$

Wir nehmen jetzt an, dass die Aussage für $j < \ell$ bewiesen ist. Das Polynom hat also die Darstellung

$$p(z) = (z - \lambda)^j \cdot q_j(z).$$

Es ist $j \leq \ell - 1$ und daher nach Voraussetzung in (ii): $p^{(j)}(\lambda) = 0$. Eine genauere Überlegung mit Hilfe einer j–fachen Anwendung der Leibniz–Produktregel (oder mit den Satz von Taylor) zeigt, dass dann

$$q_j(\lambda) = 0$$

sein muss. Das aber bedeutet (vgl. Induktionsanfang), dass aus dem Polynom q_j ein Faktor $z - \lambda$ abgespalten werden kann. Insgesamt kann also

$$p(z) = (z - \lambda)^j \cdot q_j(z) = (z - \lambda)^j \cdot (z - \lambda)\, q_{j+1}(z) = (z - \lambda)^{j+1}\, q_{j+1}(z)$$

mit einem Polynom q_{j+1} vom Grad $N - (j+1)$ geschrieben werden. ♦

Nicht beweisen können wir hier den

Satz 71 (Fundamentalsatz der Algebra)

Es sei p ein normiertes Polynom vom Grad $N \geq 1$ mit komplexen Koeffizienten.

(i) Das Polynom p hat eine Nullstelle $\lambda \in \mathbb{C}$.

(ii) Es gibt einen Datensatz

$$(\ell_1, \lambda_1), (\ell_2, \lambda_2), \ldots, (\ell_q, \lambda_q) \in \mathbb{N} \times \mathbb{C}$$

mit paarweise verschiedenen λ_j und

$$\ell_1 + \ldots + \ell_q = N,$$

so dass

$$p(z) = (z - \lambda_1)^{\ell_1} \cdot (z - \lambda_2)^{\ell_2} \cdot \ldots \cdot (z - \lambda_q)^{\ell_q}.$$

Das Polynom zerfällt also komplett in Linearfaktoren.

Dabei folgt (ii) aus (i) aufgrund des vorhergehenden Satzes.

Die letzte Aussage kann auch auf Polynome mit reellen Koeffizienten angewandt werden. Über die komplexen Nullstellen λ_j, die in den Linearfaktoren auftreten, kann man hier genaueres sagen:

Satz 72

Es sei p ein normiertes Polynom vom Grad $N \geq 1$ mit ***reellen*** *Koeffizienten.*

(i) Ist $\lambda \in \mathbb{C}$ eine ℓ–fache Nullstelle, so ist auch die konjugiert komplexe Zahl $\overline{\lambda}$ eine ℓ–fache Nullstelle. Das ist natürlich nur für echt komplexe Zahlen (mit Imaginärteil ungleich Null) interessant.

(ii) Es gibt einen Datensatz, bestehend aus

$$\begin{array}{rl} (I) & (\ell_1,\lambda_1),(\ell_2,\lambda_2),\ldots,(\ell_r,\lambda_r) \in \mathbb{N}\times\mathbb{R} \\ (II) & (m_1,\mu_1),(m_2,\mu_2),\ldots,(m_s,\mu_s) \in \mathbb{N}\times(\mathbb{C}\setminus\mathbb{R}) \\ (\overline{II}) & (m_1,\overline{\mu}_1),(m_2,\overline{\mu}_2),\ldots,(m_s,\overline{\mu}_s) \in \mathbb{N}\times(\mathbb{C}\setminus\mathbb{R}) \end{array}$$

mit paarweise verschiedenen zweiten Komponenten $\lambda_*, \mu_*, \overline{\mu}_*$ und

$$\ell_1 + \ldots + \ell_r + 2(m_1 + \ldots + m_s) = N,$$

so dass

$$\begin{aligned} p(z) \; = \; & (z-\lambda_1)^{\ell_1}\cdot(z-\lambda_2)^{\ell_2}\cdot\ldots\cdot(z-\lambda_r)^{\ell_r}\cdot \\ & \quad (z^2 + -2\,\mathrm{Re}(\mu_1) + |\mu_1|^2)^{m_1}\cdot\ldots\cdot(z^2 - 2\,\mathrm{Re}(\mu_s) + |\mu_s|^2)^{m_s}. \end{aligned}$$

Das Polynom zerfällt also komplett in reelle Linearfaktoren und reelle quadratische Faktoren.

Beweis (i) Ist $p(\lambda) = 0$, so ist auch $p(\overline{\lambda}) = \overline{p(\lambda)} = 0$. Die gleiche Überlegung gilt für die Ableitungen.

(ii) Je zwei Potenzen von Linearfaktoren $(z-\mu_j)^{m_j}$ und $(z-\overline{\mu_j})^{m_j}$ für $j = 1,\ldots,s$ mit echt–komplexen μ_j können zur Potenz der quadratischen Faktoren

$$(z-\mu_j)^{m_j}(z-\overline{\mu_j})^{m_j} = (z^2 - 2\,\mathrm{Re}(\mu_j) + |\mu_j|^2)^{m_j}$$

zusammengefasst werden. ♦

8.7 Lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten

Es handelt sich um Gleichungen der Form

$$y^{(N)} + a_{N-1}y^{(N-1)} + \ldots + a_1 y' + a_0 y = 0.$$

mit Konstanten $a_i \in \mathbb{R}$, $i = 0,1,\ldots,N-1$. Sie können natürlich auch explizit geschrieben werden. Man beachte, dass die Rechte Seite

$$-\big(a_{N-1}y^{(N-1)} + \ldots + a_1 y' + a_0 y\big)$$

auf ganz $I = \mathbb{R}$ definiert ist.

Der im Anfangsbeispiel 2 erfolgreiche Ansatz führt insgesamt auf die Lösung in dem folgenden Satz.

Satz 73 *Es sei*

$$y^{(N)} + a_{(N-1)}y^{(N-1)} + \ldots + a_1 y' + a_0 y = 0$$

eine lineare Differentialgleichung N–ter Ordnung mit konstanten reellen Koeffizienten. Betrachte für das zugehörige (charakteristische) Polynom

$$z^N + a_{(N-1)}z^{N-1} + \ldots + a_1 z + a_0 = 0$$

den in Satz 72 beschriebenen Datensatz.

(i) Gehört (ℓ, λ) zum Datensatz (I), so sind die ℓ Funktionen

$$y^I_{(j,\lambda)}(x) = x^{j-1} \cdot \exp(\lambda x), \qquad j = 1, \ldots, \ell$$

Lösungen der Differentialgleichung.

(ii) Gehört (m, μ) zum Datensatz (II), so sind die $2m$ Funktionen

$$\begin{aligned} y^{II}_{(j,\mu)}(x) &= x^{j-1} \cdot e^{(\mathrm{Re}\,\mu)x} \cdot \cos[(\mathrm{Im}\,\mu)x] \\ y^{III}_{(j,\mu)}(x) &= x^{j-1} \cdot e^{(\mathrm{Re}\,\mu)x} \cdot \sin[(\mathrm{Im}\,\mu)x], \qquad j = 1, \ldots, m \end{aligned}$$

Lösungen der Differentialgleichung.

(iii) In (i) und (ii) sind insgesamt

$$\ell_1 + \ldots + \ell_r + 2(m_1 + \ldots + m_s) = N,$$

Lösungen aufgelistet, die wir mit

$$\widetilde{y}_1, \ldots, \widetilde{y}_N$$

durchnummerieren. Dann ist die Menge aller Lösungen der Differentialgleichung gegeben durch

$$y = \widetilde{y}_1 \cdot C_1 + \widetilde{y}_2 \cdot C_2 + \ldots + \widetilde{y}_N \cdot C_N, \qquad C_1, \ldots, C_N \in \mathbb{R}.$$

(iv) Die (eindeutige) Lösung eines AWPs N-ter Ordnung

$$\begin{cases} y^{(N)} + a_{(N-1)}y^{(N-1)} + \ldots + a_1 y' + a_0 y = 0 \\ y(x_0) = y_0 \\ \quad \vdots \\ y^{(N-1)}(x_0) = y_{N-1} \end{cases}$$

mit N Anfangsbedingungen kann aus der Menge aller Lösungen (iii) durch Lösung des inhomogenen Linearen Gleichungssystems

$$\begin{aligned} \widetilde{y}_1(x_0) \cdot C_1 + \widetilde{y}_2(x_0) \cdot C_2 + \ldots + \widetilde{y}_N(x_0) \cdot C_N &= y_0 \\ &\vdots \\ \widetilde{y}_1^{(N-1)}(x_0) \cdot C_1 + \widetilde{y}_2^{(N-1)}(x_0) \cdot C_2 + \ldots + \widetilde{y}_N^{(N-1)}(x_0) \cdot C_N &= y_{N-1} \end{aligned}$$

gewonnen werden.

Diese allgemeinen Sätze sind vor allem im Hinblick auf Lineare Differentialgleichungen zweiter Ordnung interessant: Vgl. Aulbach S. 269 – 274.

8.8 Beispiel eines Differentialgleichungssystems

Wir betrachten nur das eine im StEx aufgetretene Beispiel (F05,T3 A4) eines Differentialgleichungssystems, das zweidimensional, erster Ordnung mit konstanten Koeffizienten ist. Ein weiteres Beispiel H02 T3 A5.

Bestimmen Sie die Menge der auf $\mathbb{R}$ definierten Lösungsfunktionen des Differenzialgleichungssystems

$$\begin{aligned} y_1'(t) &= \ \ y_1(t) + 4y_2(t) \\ y_2'(t) &= 2y_1(t) + 3y_2(t) \end{aligned}$$

Es empfiehlt sich, eine andere Notation (mit Variabler t) zu wählen:

$$\begin{aligned} x' &= \ \ x + 4y \\ y' &= 2x + 3y \end{aligned}$$

HINWEISE zur Lösung: Wähle eine der beiden Gleichungen aus (hier: die erste) und ...

- löse sie nach der ,,anderen" Funktion auf:

$$y = \frac{1}{4}\,(x' - x)$$

- leite sie ein weiteres Mal ab:

$$x'' = x' + 4y'$$

Setzt man nun die beiden Gleichungen für die ,,andere" Funktion geschickt in die letzte Gleichung ein, so erhält man

$$\begin{aligned} x'' &= x' + 4y' \\ &= x' + 4[2x + 3y] \\ &= x' + 4[2x + \frac{3}{4}(x' - x)] \\ &= 4x' + 5x \end{aligned}$$

und damit eine einzige Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten:

$$x'' - 4x' - 5x = 0 \qquad (*)$$

Wir lösen sie mit Hilfe des letzten Satzes. Das charakteristische Polynom

$$z^2 - 4z - 5 = (z + 1) \cdot (z - 5)$$

hat die beiden Nullstellen:

$$\lambda_1 = 5, \qquad \lambda = -1,$$

so dass sich als Lösungen der Differentialgleichung $(*)$ die beiden Funktionen

$$x_1(t) = e^{5t}, \qquad x_2(t) = e^{-t}$$

ergeben. Die Menge aller Lösungen von $(*)$ ist somit gegeben durch

$$x(t) = e^{5t} \cdot C_1 \ + \ e^{-t} \cdot C_2$$

Daraus kann die Funktion $y = \frac{x'-x}{4}$ berechnet werden, es ergibt sich dann als Menge der Lösungen des gegebenen Differentialgleichungssystems

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2 \middle| \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} e^{5t} \cdot C_1 + e^{-t} \cdot C_2 \\ e^{5t} \cdot C_1 - \frac{1}{2} e^{-t} \cdot C_2 \end{pmatrix}, C_1, C_2 \in \mathbb{R} \right\}.$$

9 Elementare Inhaltslehre

9.1 Das mehrdimensionale Riemann–Integral

Wir betrachten den Vektorraum $\mathbb{R}^N$. In den Anwendungen werden wir es wieder mit den Fällen $N = 2, 3$ zu tun haben. Für $N = 1$ ergibt sich im folgenden das uns schon bekannte eindimensionale Riemann–Integral.

Die $2N$ Zahlen $c_i, d_i \in \mathbb{R}$ mit $c_i < d_i$, $i = 1, \ldots, N$ definieren den (achsenparallelen) abgeschlossenen *Quader*

$$Q \ := \ [c_1, d_1] \times [c_2, d_2] \times [c_3, d_3] \times \ldots \times [c_N, d_N].$$

Ist für jedes feste i eine Zerlegung $\mathcal{Z}_i = (x_{0,i}, \ldots, x_{n_i,i})$ des Intervalls $[c_i, d_i]$ mit $n_i + 1$ Zahlen gegeben, so erhält man daraus eine Zerlegung $\mathcal{Z}$ des Quaders in

$$n = n_1 \cdot \ldots \cdot n_N$$

Teilquader der Form

$$\begin{aligned} Q_{\vec{j}} = Q_{j_1,j_2,\ldots,j_N} \ :=& \ [x_{j_1-1}, x_{j_1}] \times \ldots \times [x_{j_N-1}, x_{j_N}], \\ & j_1 \in \{1, \ldots, n_1\}, \ldots, j_N \in \{1, \ldots, n_N\}, \end{aligned}$$

die sich an den Rändern überschneiden. Dabei ist $\vec{j} = (j_1, j_2, \ldots, j_N)$ eine Abkürzung für die „Quadernummer". Mit

$$Q^\circ_{\vec{j}} = Q^\circ_{j_1,j_2,\ldots,j_N} \ := \ ]x_{j_1-1}, x_{j_1}[\times \ldots \times]x_{j_N-1}, x_{j_N}[$$

wird der zugehörige offene Quader (ohne Ränder) bezeichnet. Es gilt

$$Q \ = \ \bigcup_{\vec{j}} Q_{\vec{j}} \ = \ \bigcup_{\vec{j}} Q_{j_1,j_2,\ldots,j_N},$$

wobei das Symbol $\vec{j}$ unterhalb bedeutet, dass alle möglichen Kombinationen der j's, das heißt alle Quader der Zerlegung erfasst werden.

Sind zwei Zerlegungen $\mathcal{Z}$ und $\mathcal{Z}'$ gegeben, so kann man wieder definieren, was es heißt, dass die Zerlegung $\mathcal{Z}'$ feiner als die andere $\mathcal{Z}$ ist (das bedeutet $\mathcal{Z}'_i$ enthält alle Zerlegungsstellen von $\mathcal{Z}_i$ für alle i) und was eine gemeinsame Verfeinerung $\mathcal{Z} \sqcup \mathcal{Z}'$ (Vereinigung der Zerlegungsstellen bzgl. jeder Koordinate) ist.

Mit $\mathrm{vol}(Q_{\vec{j}})$ bezeichnen wir das *Volumen* eines (Teil–)Quaders:

$$\mathrm{vol}(Q_{\vec{j}}) = (x_{j_1} - x_{j_1-1}) \cdot \ldots \cdot (x_{j_N} - x_{j_N-1}).$$

Definition (Riemann–Integral) Es sei $f : Q \to \mathbb{R}$ eine reelle beschränkte Funktion.

(1a) Die $\mathcal{Z}$–*Obersumme* von f ist die reelle Zahl

$$\overline{S}(f, \mathcal{Z}) := \sum_{\vec{j}} \sup_{x \in Q_{\vec{j}}^\circ} \{ f(x) \} \cdot \mathrm{vol}(Q_{\vec{j}}).$$

Das bedeutet, dass auf jedem Teilquader der Zerlegung das Supremum der Funktion f gebildet wird, dann diese Suprema — multipliziert jeweils mit dem Volumen des zugehörigen Teilquaders — aufaddiert werden.

(1b) Die $\mathcal{Z}$–*Untersumme* von f ist die reelle Zahl

$$\underline{S}(f, \mathcal{Z}) := \sum_{\vec{j}} \inf_{x \in Q_{\vec{j}}^\circ} \{ f(x) \} \cdot \mathrm{vol}(Q_{\vec{j}}).$$

(2a) Das *Oberintegral* von f über dem Quader Q ist die reelle Zahl

$$\overline{\int}_Q f(x)\,dx := \inf_{\mathcal{Z}} \left\{ \overline{S}(f, \mathcal{Z}) \right\}.$$

(2b) Das *Unterintegral* von f ist die reelle Zahl

$$\underline{\int}_Q f(x)\,dx := \sup_{\mathcal{Z}} \left\{ \underline{S}(f, \mathcal{Z}) \right\}.$$

(3) Es ist unmittelbar klar, dass immer gilt:

$$\underline{\int}_Q f(x)\,dx \le \overline{\int}_Q f(x)\,dx.$$

Die Funktion f heißt *(Riemann–)integrierbar* (auf dem Quader Q), wenn Oberintegral und Unterintegral übereinstimmen. Diese Zahl

$$\int_Q f(x)\,dx := \overline{\int}_Q f(x)\,dx = \underline{\int}_Q f(x)\,dx$$

heißt dann das *(Riemann–)Integral* der Funktion f auf dem Quader Q.

Satz 74 *Eine reelle beschränkte Funktion $f : Q \to \mathbb{R}$ ist genau dann integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung $\mathcal{Z}$ von Q gibt, so dass*

$$\overline{S}(f, \mathcal{Z}) - \underline{S}(f, \mathcal{Z}) < \varepsilon.$$

Beweis Die Aussage dieses Satzes ist für die eindimensionale Situation im wesentlichen gleich der Äquivalenz $(A) \iff (B)$ aus Satz 21. Dort wurde lediglich mit Hilfe des Begriffs der Treppenfunktion gearbeitet. Der Beweis dort kann — geeignet (leicht) modifiziert — übernommen werden. ♦

Satz 75 *Treppenfunktionen $\varphi : Q \to \mathbb{R}$ — sie sind auf den offenen Teilquadern einer Zerlegung konstant — sind integrierbar.*

9.2 Rückführung auf wiederholte Integration

Etwas Notation ($\widehat{\ }$ bedeutet „Weglassen”):

Es sei $x \in Q = [c_1, d_1] \times \ldots \times [c_N, d_N]$ ein Punkt in einem N–dimensionalen Quader und $k \in \{1, \ldots, N\}$ eine Koordinatennummer.

- Durch „Weglassen” der k–ten Koordinate entsteht der $N-1$–dimensionale projizierte Quader

$$\begin{aligned} Q_{\widehat{k}} &:= [c_1, d_1] \times \ldots \times \widehat{[c_k, d_k]} \times \ldots \times [c_N, d_N] \\ &:= [c_1, d_1] \times \ldots \times [c_{k-1}, d_{k-1}] \times [c_{k+1}, d_{k+1}] \times \ldots \times [c_N, d_N] \subseteq \mathbb{R}^{N-1}. \end{aligned}$$

- Mit

$$\begin{aligned} x_{\widehat{k}} &:= (x_1, \ldots, \widehat{x_k}, \ldots, x_N) \\ &:= (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_N) \qquad \in Q_{\widehat{k}} \subseteq \mathbb{R}^{N-1} \end{aligned}$$

  bezeichnen wir den durch Weglassen der k-ten Koordinaten entstehenden projizierten Punkt.

- Ist $g : Q_{\widehat{k}} \to \mathbb{R}$ eine integrierbare Funktion, so bezeichnen wir das zugehörige Integral mit

$$\int_{Q_{\widehat{k}}} g(\widehat{x}_k)\, dx_{\widehat{k}}\,.$$

- Falls für jedes $x_{\widehat{k}} \in Q_{\widehat{k}}$ das unten angegebene eindimensionale Integral von f über die k–te Koordinate existiert, so entsteht durch Belassen der anderen Koordinaten eine neue Funktion

$$f_{\widehat{k}} : \begin{cases} Q_{\widehat{k}} & \to \quad \mathbb{R} \\ x_{\widehat{k}} & \mapsto \quad \int_{c_k}^{d_k} f(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n)\, dy \end{cases}$$

Satz 76 *Es seien die integrierbare Funktion $f : Q \to \mathbb{R}$ und ein $k \in \{1, \ldots, N\}$ gegeben. Existiert die Funktion $f_{\widehat{k}} : Q_{\widehat{k}} \to \mathbb{R}$ und ist sie über $Q_{\widehat{k}}$ integrierbar, so gilt:*

$$\int_Q f(x)\, dx = \int_{Q_{\widehat{k}}} \underbrace{\int_{c_k}^{d_k} f(x_1, x_2, \ldots, x_k, \ldots, x_{N-1}, x_N)\, dx_k}_{f_{\widehat{k}}(x_{\widehat{k}})}\, dx_{\widehat{k}}.$$

Die N–dimensionale Integration wurde in eine 1–dimensionale und eine $(N-1)$–dimensionale zerlegt.

Beweis Es sei $\mathcal{Z}$ eine Zerlegung von Q und $Q_{\vec{j}}$ ein Teilquader der Zerlegung. Für ein beliebiges $x = (x_1, \ldots, x_N) \in Q_{\vec{j}}$ ist

$$\inf_{z \in Q_{\vec{j}}} \{f(z)\} \leq f(x) \leq \sup_{z \in Q_{\vec{j}}} \{f(z)\}.$$

Mit dem Mittelwertsatzes der eindimensionalen Integralrechnung (Satz 30) folgt:

$$\begin{aligned}\inf_{z\in Q_{\vec{j}}}\{f(z)\}\cdot(x_{j_k}-x_{j_k-1}) &\leq \int_{x_{j_k-1}}^{x_{j_k}} f(x_1,\ldots,x_{k-1},y,x_{k+1},\ldots,x_N)\,dy \\ &\leq \sup_{z\in Q_{\vec{j}}}\{f(z)\}\cdot(x_{j_k}-x_{j_k-1}).\end{aligned}$$

Wir summieren in x_k–Richtung auf:

$$\begin{aligned}\sum_{j_k=1}^{n_k}\inf_{z\in Q_{\vec{j}}}\{f(z)\}\cdot(x_{j_k}-x_{j_k-1}) &\leq \int_{c_k}^{d_k} f(x_1,\ldots,x_{k-1},y,x_{k+1},\ldots,x_N)\,dy \\ &\leq \sum_{j_k=1}^{n_k}\sup_{z\in Q_{\vec{j}}}\{f(z)\}\cdot(x_{j_k}-x_{j_k-1}).\end{aligned}$$

Die Funktionen in dieser Ungleichungskette werden über $(Q_{\vec{j}})_{\widehat{k}}$ integriert:

$$\begin{aligned}&\sum_{j_k=1}^{n_k}\inf_{z\in Q_{\vec{j}}}\{f(z)\}\cdot\underbrace{(x_{j_k}-x_{j_k-1})\cdot\mathrm{vol}((Q_{\vec{j}})_{\widehat{k}})}_{\mathrm{vol}(Q_{\vec{j}})}\leq \\ &\int_{(Q_{\vec{j}})_{\widehat{k}}}\int_{c_k}^{d_k} f(x_1,\ldots,x_{k-1},y,x_{k+1},\ldots,x_N)\,dy\,dx_{\widehat{k}}\leq \\ &\sum_{j_k=1}^{n_k}\sup_{z\in Q_{\vec{j}}}\{f(z)\}\cdot(x_{j_k}-x_{j_k-1})\cdot\underbrace{(x_{j_k}-x_{j_k-1})\cdot\mathrm{vol}((Q_{\vec{j}})_{\widehat{k}})}_{\mathrm{vol}(Q_{\vec{j}})}\end{aligned}$$

Eine weitere Summation über alle Indices

$$(\vec{j})_{\widehat{k}}=(j_1,\ldots,j_{k-1},j_{k+1},\ldots,j_N),\qquad j_i=1,\ldots,n_i \text{ für } i=1,\ldots,N,$$

die die Teilquader der Zerlegung des projizierten Quaders $Q_{\widehat{k}}$ nummerieren, ergibt:

$$\begin{aligned}&\underline{\mathcal{S}}(f,\mathcal{Z})=\sum_{\vec{j}}\inf_{z\in Q_{\vec{j}}}\{f(z)\}\cdot\mathrm{vol}(Q_{\vec{j}})\leq \\ &\int_{Q_{\widehat{k}}}\int_{c_k}^{d_k} f(x_1,\ldots,x_{k-1},y,x_{k+1},\ldots,x_N)\,dy\,dx_{\widehat{k}}\leq \\ &\sum_{\vec{j}}\sup_{z\in Q_{\vec{j}}}\{f(z)\}\cdot(x_{j_k}-x_{j_k-1})\cdot\mathrm{vol}(Q_{\vec{j}})=\overline{\mathcal{S}}(f,\mathcal{Z})\end{aligned}$$

Da dies für jede beliebige Zerlegung $\mathcal{Z}$ gilt, folgt weiter

$$\underline{\int_Q} f(x)\,dx\leq\int_{Q_{\widehat{k}}}\int_{c_k}^{d_k} f(x_1,\ldots,x_{k-1},y,x_{k+1},\ldots,x_N)\,dy\,dx_{\widehat{k}}\leq\overline{\int_Q} f(x)\,dx.$$

Da f integrierbar ist, folgt daraus die Behauptung. Die Integrationsvariable y kann dabei noch in x_k umbenannt werden. ♦

Satz 77 *Es sei die integrierbare Funktion $f : Q \to \mathbb{R}$ gegeben. Die Gleichung*

$$\int_Q f(x)\,dx =$$
$$\int_{c_N}^{d_N} \int_{c_{N-1}}^{d_{N-1}} \ldots \int_{c_2}^{d_2} \int_{c_1}^{d_1} f(x_1, x_2, \ldots, x_{N-1}, x_N)\,dx_1\,dx_2\,\ldots\,dx_{N-1}\,dx_N$$

gilt, wenn die sukzessiven eindimensionalen Integrale existieren. Dabei kann auch jede andere Reihenfolge der Koordinatennummern verwendet werden.

Beweis Man wendet den letzten Satz nacheinander auf die einzelnen Koordinatennummern — insgesamt $(N-1)$–mal — an. ♦

9.3 Stetige Funktionen sind integrierbar

Satz 78 *Ist die Funktion $f : Q \to \mathbb{R}$ stetig, so ist sie integrierbar.*

Beweis Es sei ein $\varepsilon > 0$ vorgegeben.

Wie am Anfang des obigen Beweises kann man begründen, dass zu $\frac{\varepsilon}{\mathrm{vol}(Q)}$ ein $\delta > 0$ existiert, so dass

$$|f(x) - f(x')| \le \frac{\varepsilon}{\mathrm{vol}(Q)}, \qquad \text{falls} \quad ||x - x'||_N < \delta.$$

Es sei jetzt $\mathcal{Z}$ eine Zerlegung, bei der jede Längsseite eines beliebigen Teilquaders kleiner als $\frac{\delta}{N}$ ist. Dann gilt für zwei beliebige Punkte x, x' innerhalb des gleichen Teilquaders

$$||x - x'|| < \delta \qquad \text{und deshalb} \quad |f(x) - f(x')| \le \frac{\varepsilon}{\mathrm{vol}(Q)},$$

dann weiter für jeden einzelnen Quader $Q_{\vec{j}}$ der Zerlegung

$$\sup_{x \in Q_{\vec{j}}} \{f(x)\} - \inf_{x \in Q_{\vec{j}}} \{f(x)\} \le \frac{\varepsilon}{\mathrm{vol}(Q)}.$$

Es gilt weiter

$$\begin{aligned} \overline{\mathcal{S}}(f, \mathcal{Z}) - \underline{\mathcal{S}}(f, \mathcal{Z}) &= \sum_{\vec{j}} \left[\sup_{x \in Q_{\vec{j}}} \{f(x)\} - \inf_{x \in Q_{\vec{j}}} \{f(x)\} \right] \cdot \mathrm{vol}(Q_{\vec{j}}) \\ &\le \sum_{\vec{j}} \frac{\varepsilon}{\mathrm{vol}(Q)} \cdot \mathrm{vol}(Q_{\vec{j}}) = \frac{\varepsilon}{\mathrm{vol}(Q)} \cdot \sum_{\vec{j}} \mathrm{vol}(Q_{\vec{j}}) = \varepsilon. \end{aligned}$$

Nach Satz 74 ist f integrierbar. ♦

Satz 79 *Ist die Funktion $f : Q \to \mathbb{R}$ stetig, so ist für jedes $k \in \{1, \ldots, N\}$ auch die Funktion $f_{\widehat{k}} : Q_{\widehat{k}} \to \mathbb{R}$ stetig.*

Beweis Es sei ein $\varepsilon > 0$ vorgegeben. Der Quader Q ist kompakt, also ist f sogar gleichmäßig stetig (vgl. Satz 57). Das bedeutet, dass zu $\frac{\varepsilon}{d_k - c_k}$ ein $\delta > 0$ existiert, so dass

$$|f(x) - f(x')| \leq \frac{\varepsilon}{d_k - c_k}, \qquad \text{falls} \quad ||x - x'||_N < \delta, \quad x, x' \in \mathbb{R}^N.$$

(Der Index N kennzeichnet, dass es sich um die (euklidische) Norm im $\mathbb{R}^N$ handelt.) Es seien jetzt $x_{\widehat{k}}, x_{\widehat{k}}{}' \in Q_{\widehat{k}} \subseteq \mathbb{R}^{N-1}$, so dass

$$\|x_{\widehat{k}} - x_{\widehat{k}}{}'\|_{N-1} < \delta.$$

Dann ist auch für beliebiges $y \in [c_k, d_k]$

$$\|(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n) - (x'_1, \ldots, x'_{k-1}, y, x'_{k+1}, \ldots, x'_n)\|_N < \delta$$

und deshalb

$$\big|f(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n) - f(x'_1, \ldots, x'_{k-1}, y, x'_{k+1}, \ldots, x'_n)\big| < \frac{\varepsilon}{d_k - c_k}.$$

Dann gilt:

$$\begin{aligned}
&\big|f_{\widehat{k}}(x_{\widehat{k}}) - f_{\widehat{k}}(x_{\widehat{k}}{}')\big| = \\
&\left|\int_{c_k}^{d_k} f(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n) - f(x'_1, \ldots, x'_{k-1}, y, x'_{k+1}, \ldots, x'_n)\, dy\right| \leq \\
&\int_{c_k}^{d_k} \big|f(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_n) - f(x'_1, \ldots, x'_{k-1}, y, x'_{k+1}, \ldots, x'_n)\big|\, dy \leq \\
&\int_{c_k}^{d_k} \frac{\varepsilon}{d_k - c_k}\, dy = \frac{\varepsilon}{d_k - c_k} \cdot (d_k - c_k) = \varepsilon.
\end{aligned}$$

Also ist $f_{\widehat{k}}$ (gleichmäßig) stetig. ♦

9.4 Integration über kompakte Teilmengen des $\mathbb{R}^N$

Ab jetzt sei K eine zusammenhängende kompakte Teilmenge des $\mathbb{R}^N$, deren ,,Rand" aus endlich vielen genügend glatten $(N-1)$–dimensionalen (Hyper–) Flächenstücken besteht.

Die vier Begriffe ,,zusammenhängend", ,,Rand", ,,glatt", ,,(Hyper–)Flächenstücke" bleiben hier undefiniert. Wir müssen uns hier auf die Anschauung berufen, da eine mathematisch–befriedigende Einführung zu aufwändig wäre.

Als Beispiele halten Sie sich bitte 2– oder 3–dimensionale Polyeder (durch Strecken bzw. ebene Flächenstücke begrenzt), Kugeln B (vgl. Abschnitt 5.3) und andere Rotationskörper vor Augen.

Da K als kompakte Menge beschränkt ist, gibt es einen (achsenparallelen) Quader Q, in dem K ganz enthalten ist.
Ist $f : K \to \mathbb{R}$ eine Funktion, so definieren wir eine Erweiterung $f_Q : Q \to \mathbb{R}$ durch

$$f_Q(x) = \begin{cases} f(x), & \text{falls } x \in K, \\ 0, & \text{falls } x \in Q \setminus K. \end{cases}$$

und definieren: $f : K \to \mathbb{R}$ heißt *integrierbar (über K)*, wenn f_Q über Q integrierbar ist. Als *Integral (über K)* wird festgesetzt:

$$\int_K f(x)\,dx \;:=\; \int_Q f_Q(x)\,dx.$$

Man kann leicht beweisen, dass diese Definitionen von dem ausgewählten Quader Q unabhängig sind.

Ohne Beweis stellen wir fest:

Satz 80 *Ist die Funktion $f : K \to \mathbb{R}$ stetig, so ist sie integrierbar. In diesem Fall ist die in Satz 77 angegebene Bedingung erfüllt, es kann sukzessive koordinatenweise integriert werden.*

9.5 Volumina von kompakten Teilmengen

Insbesondere ist die Konstant–Eins–Funktion stetig, also über K integrierbar. Wir definieren das *Volumen* von K (wie zu Beginn des letzten Abschnitts beschrieben) als

$$\mathrm{vol}(K) = \mathrm{vol}_N(K) \;:=\; \int_K 1\,dx = \int_Q \chi_K(x)\,dx.$$

Der Index N kann hinzugefügt werden, wenn man unterstreichen will, dass es sich um das Volumen im N–dimensionalen euklidischen Raum handelt. Die Funktion $\chi_K : \mathbb{R}^N \to \mathbb{R}$ ist die charakteristische Funktion von K:

$$\chi_K(x) = \begin{cases} 1, & \text{falls } x \in K, \\ 0, & \text{falls } x \notin K. \end{cases}$$

Beispiel Ist $K = Q = [c_1, d_1] \times \ldots \times [c_N, d_N]$ ein Quader, so ergibt sich

$$\mathrm{vol}(Q) = \int_Q 1\,dx = (d_1 - c_1) \cdot \ldots \cdot (d_N - c_N) = \mathrm{vol}(Q),$$

die neue Definition des Volumens stimmt also in diesem Fall mit der alten überein.

Für eine kompakte Menge $K \subseteq \mathbb{R}^N$, eine Koordinatennummer $k \in \{1, \ldots, N\}$ und eine fest gegebene Höhe $y \in \mathbb{R}$ definieren wir den $(N-1)$–dimensionalen Schnitt als die Menge im $\mathbb{R}^{N-1}$

$$K_{\widehat{k}}(y) \quad := \quad \Big\{ (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_N) \in \mathbb{R}^{N-1} \;\Big|\; (x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_N) \in K \Big\}.$$

**** Zeichnung für $N = 2$ *****

Ist beispielsweise $K = Q$ ein Quader, so ergibt sich

$$K_{\widehat{k}}(y) = \begin{cases} Q_{\widehat{k}}, & \text{falls } y \in [c_k, d_k], \\ \varnothing, & \text{sonst.} \end{cases}$$

Satz 81 (Prinzip von Cavalieri) *Ist K eine kompakte Teilmenge wie oben beschrieben und $k \in \{1, \ldots, N\}$ eine Koordinatennummer, so kann man das Volumen von K aus den Volumina der Schnitte durch Aufintegration berechnen:*

$$\mathrm{vol}_N(K) = \int_c^d \mathrm{vol}_{N-1}(K_{\widehat{k}}(y))\, dy.$$

Dabei ist die Zahl c so klein, d so groß, gewählt, dass $K_{\widehat{k}}(y) = \emptyset$ für $y \notin [c, d]$.

Beweis Wir wählen einen Quader

$$Q = [c_1, d_1] \times \ldots \times [c_{k-1}, d_{k-1}] \times [c, d] \times [c_{k+1}, d_{k+1}] \times \ldots \times [c_N, d_N]$$

(also $c = c_k, d = d_k$) so aus, dass $K \subseteq Q$. Es ist dann für alle $y \in [c, d]$

$$K_{\widehat{k}}(y) \subseteq Q_{\widehat{k}} = [c_1, d_1] \times \ldots \times [c_{k-1}, d_{k-1}] \times [c_{k+1}, d_{k+1}] \times \ldots \times [c_N, d_N]$$

und deshalb gemäß Definition des Volumens

$$\mathrm{vol}_{N-1}(K_{\widehat{k}}(y)) = \int_{Q_{\widehat{k}}} \chi_{K_{\widehat{k}}(y)}(x_{\widehat{k}})\, dx_{\widehat{k}}.$$

Eine leichte, aber sorgfältige Überprüfung mit Hilfe der Definitionen ergibt, dass für festes $y \in [c, d]$ und $x_{\widehat{k}} = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_N) \in Q_{\widehat{k}}$ gilt

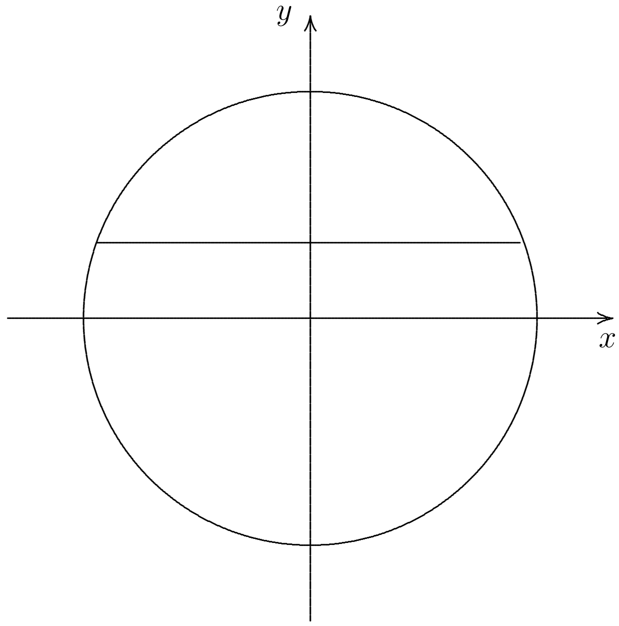
$$\chi_K(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_N) = \chi_{K_{\widehat{k}}(y)}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_N).$$

Wir wenden dann den SatzRiemann Fubini über die sukzessive Integration an, wobei wir über die k–te Koordinate als letztes integrieren, und erhalten:

$$\begin{aligned} \mathrm{vol}_N(K) &= \int_Q \chi_K(x)\, dx = \int_c^d \int_{Q_{\widehat{k}}} \chi_K(x_1, \ldots, x_k, y, x_{k+1}, \ldots, x_N)\, dx_{\widehat{k}}\, dy \\ &= \int_c^d \underbrace{\int_{Q_{\widehat{k}}} \chi_{K_{\widehat{k}}(y)}(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_N)\, dx_{\widehat{k}}}_{\mathrm{vol}_{N-1}(K_{\widehat{k}}(y))}\, dy \end{aligned}$$

♦

Beispiel



Wir berechnen das zweidimensionale Volumen der Kreisfläche (Radius r)

$$K = \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; x^2 + y^2 \le r\Big\}.$$

Als Schnitt in der Höhe y, $-r \le y \le +r$ ergibt sich

$$K_{\widehat{2}}(y) = \Big\{x \in \mathbb{R} \;\Big|\; x \le \sqrt{r^2 - y^2} \le r\Big\}$$

mit der Länge

$$\mathrm{vol}_1(K_{\widehat{2}})(y) = 2 \cdot \sqrt{r^2 - y^2}.$$

Wir wenden jetzt das Prinzip von Cavalieri an:

$$\begin{aligned}
\mathrm{vol}_2(K) &= \int_{-r}^{+r} \mathrm{vol}_1(K_{\widehat{2}})\,dy = \int_{-r}^{+r} 2 \cdot \sqrt{r^2 - y^2}\,dy \\
&= 2 \cdot \frac{1}{2} \cdot \Big[y \cdot \sqrt{r^2 - y^2} + r^2 \arcsin \frac{y}{r}\Big]_{y=-r}^{y=+r} \\
&= r^2 \cdot \Big[\arcsin \frac{y}{r}\Big]_{y=-r}^{y=+r} = r^2 \cdot [\arcsin(+1) - \arcsin(-1)] = r^2 \cdot \pi.
\end{aligned}$$

Die Fläche des oberen Halbkreises kann auch als Fläche unter dem Graphen der Funktion $y = \sqrt{r^2 - x^2}$ aufgefasst werden. Aus der Schule wissen Sie, dass der Inhalt dieser Fläche dann gegeben ist durch

$$\int_{-r}^{+r} \sqrt{r^2 - x^2}\,dx.$$

Das dieses Vorgehen ganz allgemein gerechtfertigt ist, zeigt der folgende Satz:

Satz 82 *Es sei eine Menge $K \subseteq \mathbb{R}^N$ gegeben als Menge der Punkte unterhalb des Graphen einer integrierbaren nicht–negativen Funktion*

$$f \left\{ \begin{array}{rcl} K_{\widehat{N}} & \to & \mathbb{R}_0 \\ (x_1, \ldots, x_{N-1}) & \mapsto & f(x_1, \ldots, x_{N-1}) \end{array} \right. ,$$

also als

$$K = \Big\{x \in \mathbb{R}^N \;\Big|\; 0 \le x_N \le f(x_1, \ldots, x_{N-1})\Big\}.$$

Dann gilt

$$\mathrm{vol}_N(K) = \int_{K_{\widehat{N}}} f(x_{\widehat{N}})\,dx_{\widehat{N}}.$$

Beweis Es sei m eine obere Schranke von f, also $f(x_{\widehat{N}}) \le m$, und dann

$$Q = \underbrace{[c_1, d_1] \times \ldots \times [c_{N-1}, d_{N-1}]}_{Q_{\widehat{N}}} \times [0, m]$$

ein Quader, in dem die Menge K enthalten ist. Dann gilt

$$\begin{aligned}
\mathrm{vol}_N(K) &= \int_Q \chi_K(x)\,dx = \int_{Q_{\widehat{N}}} \int_0^m \chi_K(x_1,\ldots,x_{N-1},y)\,dy\,dx\widehat{N} \\
&= \int_{Q_{\widehat{N}}} \int_0^m \chi_K(x_1,\ldots,x_{N-1},y)\,dy\,dx_{\widehat{N}} \\
&= \int_{Q_{\widehat{N}}} \int_0^m \chi_{K_{\widehat{N}}}(x_1,\ldots,x_{N-1}) \cdot \chi_{[0,f(x_1,\ldots,x_{N-1})]}(y)\,dy\,dx_{\widehat{N}} \\
&= \int_{Q_{\widehat{N}}} \chi_{K_{\widehat{N}}}(x_1,\ldots,x_{N-1}) \int_0^m \chi_{[0,f(x_1,\ldots,x_{N-1})]}(y)\,dy\,dx_{\widehat{N}} \\
&= \int_{Q_{\widehat{N}}} \chi_{K_{\widehat{N}}}(x_1,\ldots,x_{N-1}) f(x_1,\ldots,x_{N-1})]\,dx_{\widehat{N}} \\
&= \int_{K_{\widehat{N}}} f(x_{\widehat{N}})\,dx_{\widehat{N}}.
\end{aligned}$$

♦

9.6 Volumina dreidimensionaler Rotationskörper

Zu einem festen Punkt $(x,y,z) \in \mathbb{R}^3$ sei

$$L_{(x,y,z)} = \{(\widetilde{x},\widetilde{y},z) \in \mathbb{R}^3 | \sqrt{\widetilde{x}^2+\widetilde{y}^2} = \sqrt{x^2+y^2}\}$$

die Kreislinie durch (x,y,z), die auf gleicher Höhe z um die z–Achse läuft.

Definition Gilt für eine kompakte Teilmenge $K \subseteq \mathbb{R}^3$

$$(x,y,z) \in K \quad \Longrightarrow \quad L_{(x,y,z)} \subseteq K,$$

so heißt K ein *Rotationskörper (bzgl. der z–Achse)*.

Beispiel Der Zylinder mit Radius r und Höhe h

$$Z = \{(x,y,z) | 0 \le \sqrt{x^2+y^2} \le r,\ 0 \le z \le h\}$$

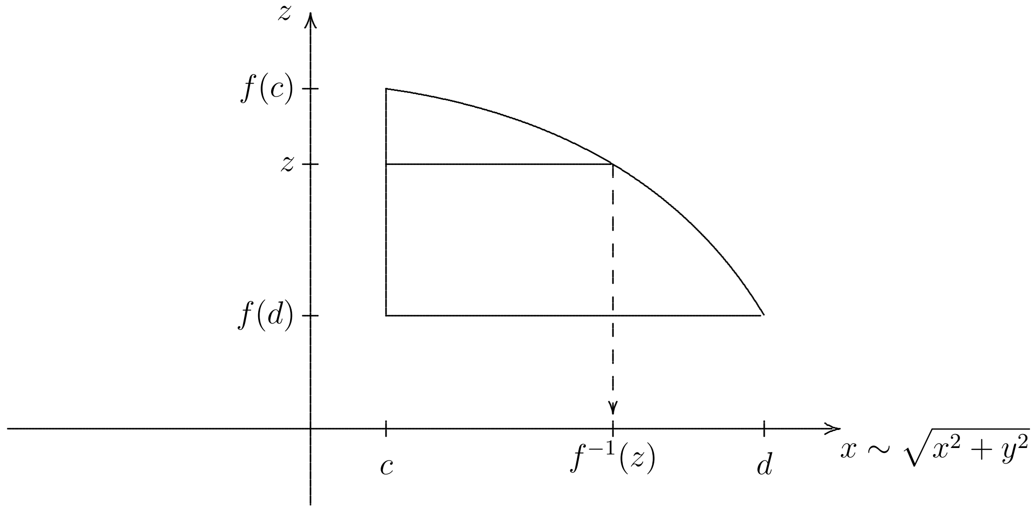
hat gemäß Cavalieri–Prinzip das Volumen

$$\mathrm{vol}(Z) = \int_0^h r^2\pi\,dz = h \cdot r^2\pi.$$

Es seien $0 \le c \le d$ reelle Zahlen und $f : [c,d] \to \mathbb{R}_0^+$ eine streng monoton fallende Funktion. Lässt man den Graphen G_f „um die z–Achse rotieren", so wird dabei ein Rotationskörper

$$K_f = \{(x,y,z) \in \mathbb{R}^3 | c \le \sqrt{x^2+y^2} \le d, \quad f(d) \le z \le f(\sqrt{x^2+y^2})\}$$

überstrichen — bzw. definiert.



Die Schnittmenge in einer festen Höhe z ist ein Kreisring mit Innenradius c und Außenradius $f^{-1}(z)$

$$K_{\hat{3}}(z) = \{(x,y) \in \mathbb{R}^2 | c \leq \sqrt{x^2+y^2} \leq f^{-1}(z)\}.$$

Die Fläche ist gegeben als Differenz von zwei Kreisflächen

$$\mathrm{vol}_2(K_{\hat{3}}) = \pi \cdot [f^{-1}(z)]^2 - \pi \cdot c^2.$$

Mit dem Cavalieri–Prinzip ergibt sich dann der erste Teil des folgenden Satzes, der zweite ist völlig analog.

Satz 83 *Es sei $f : [c,d] \to \mathbb{R}_0^+$ streng monoton fallend bzw. streng monoton steigend. Dann gilt für den zugehörigen Rotationskörper*

$$\mathrm{vol}_3(K_f(z)) = \begin{cases} \pi \cdot \int_{f(d)}^{f(c)} ([f^{-1}(z)]^2 - c^2)\, dz, & \textit{bzw.} \\ \pi \cdot \int_{f(c)}^{f(d)} (d^2 - [f^{-1}(z)]^2)\, dz. & \end{cases}$$

Beispiele

1. Die Funktion $f(x) = h \cdot (1 - \frac{x}{r})$ auf $[0,r]$ beschreibt einen geraden Kegel K. Die Umkehrfunktion ist $f^{-1}(z) = r \cdot (1 - \frac{z}{h})$, somit ergibt sich als Volumen des Kegels

$$\begin{aligned} \mathrm{vol}(K) &= \pi \cdot \int_0^h [r \cdot (1 - \frac{z}{h})]^2\, dz = \\ &= \pi \cdot r^2 \cdot \int_0^h (1 - \frac{2z}{h} + \frac{z^2}{h^2})\, dz \\ &= \pi \cdot r^2 \cdot (h - h + \frac{h}{3}) = \frac{\pi r^2 h}{3} \end{aligned}$$

2. Die Funktion $f(x) = \sqrt{r^2 - x^2}$ auf $[0, r]$ beschreibt eine Kugelhälfte mit Radius r. Die Umkehrfunktion ist $f^{-1}(z) = \sqrt{r^2 - z^2}$, somit ergibt sich als Volumen der Kugelhälfte

$$\begin{aligned}\mathrm{vol}(K) &= \pi \cdot \int_0^r [\sqrt{r^2 - z^2}]^2 \, dz = \pi \cdot \int_0^r (r^2 - z^2) \, dz = \\ &= \pi \cdot \left[r^2 z - \frac{z^3}{3} \right]_{z=0}^{z=r} = \frac{2}{3} \cdot \pi r^3.\end{aligned}$$

Mit dem letzten Satz ist die Grundlage dafür gegeben, auch Volumina von komplexeren Rotationskörpern zu berechnen.

- Die Funktion ist nicht–negativ und definiert damit die Unterseite des Rotationskörpers.
- Der Körper ist vertikal aus mehreren Teilen des obigen Typs zusammengesetzt.
- Der Körper ist radial ($\sim$ horizontal) aus mehreren Teilen des obigen Typs zusammengesetzt.