Mathematik I für Elektrotechniker, TUD WS 08/9
Teil b: Analysis

Danksagung. Freundlicherweise hat Prof. Steffen Roch den Latex-file seines Skripts Mathematik I zur Verfügung gestellt. Davon wird im Folgenden reichlich Gebrauch gemacht, insbesondere von der großen Zahl von Beispielen (die Nummerierung wird meist übernommen)und Graphiken. Das vollständige Skript von Prof. Roch finden Sie auf der Lehrmaterialseite des FB Mathematik WS 2007.

Inhaltsverzeichnis

12 Folgen und Konvergenz: Grundlegendes

12.1 Folgen

Wir betrachten zunächst nur Folgen reeller Zahlen. Wir haben den Begriff Folge in Kap.3.10 definiert. Wir denken uns aber einfach eine unendliche Liste von Zahlen, die man z.B. von vorn anfangend mit $0, 1, 2, \ldots$ durchnummerieren kann, also $a_0, a_1, a_2, \ldots$. Die Zahl a_n zusammmen mit ihrer Positionsnummer n heisst dann das n-te *Glied* der Folge.

Wir dürfen die Nummerierung aber auch bei 27 anfangen lassen wie in $b_{27}, b_{28}, \ldots$ oder bei -27 wie in $c_{-27}, c_{-26}, \ldots)$ oder nur die geraden Zahlen benutzen wie in $d_0, d_2, d_4, \ldots$ oder auf eine andere Art mit einer aufsteigenden Zahlenreihe durchnummerieren.

Eine Folge $(a_n)_{n\in\mathbb{N}}$ ist

- *nach oben beschränkt* wenn es ein $c \in \mathbb{R}$ gibt so, dass $a_n \le c$ für alle $n \in \mathbb{N}$
- *nach unten beschränkt* wenn es ein $c \in \mathbb{R}$ gibt so, dass $a_n \ge c$ für alle $n \in \mathbb{N}$
- *beschränkt* wenn es ein $c \in \mathbb{R}$ gibt so, dass $|a_n| \le c$ für alle $n \in \mathbb{N}$
- *monoton wachsend* bzw. *nicht fallend* wenn $a_n \le a_{n+1}$ für alle n
- *streng monoton wachsend* wenn $a_n < a_{n+1}$ für alle n
- *monoton fallend* bzw. *nicht wachsend* wenn $a_n \ge a_{n+1}$ für alle n
- *streng monoton fallend* wenn $a_n > a_{n+1}$ für alle n

12.2 Monotone Nullfolgen

In der Definition von "Intervallschachtelung" haben wir verlangt, dass der Abstand $d_n = b_n - a_n \ge 0$ der Intervallgrenzen unter jede Schranke $\frac{1}{k}$ fällt. Da die Folge der d_n zudem monoton fällt, d.h. $d_n \ge d_{n+1} \ge 0$, konnten wir dies so formulieren

für alle $k \ge 1$ in $\mathbb{N}$ gibt es ein $N \in \mathbb{N}$ mit $d_N \le \frac{1}{k}$

und sprechen auch von einer *monotonen Nullfolge*. Klar ist dann, dass auch

$$d_n \le \frac{1}{k} \quad \text{für alle } n \ge N$$

Will man in einem Beispiel zeigen, dass man in der Tat eine monotone Nullfolge hat, so ist es am besten, zu jedem $\varepsilon = \frac{1}{k}$ ein $N(\varepsilon)$ so anzugeben, dass $d_{N(\varepsilon)} \le \varepsilon$.
Ein wichtiges Beispiel einer monotonen Nullfolge ist $d_n = \frac{1}{2^n}$ mit $N(\varepsilon) \ge \log_2 \varepsilon$. Man darf es sich aber auch mit $N(\frac{1}{k}) = k$ bequem machen.

Beispiel 1 Wir betrachten die Folge

$$\left(\frac{1}{n(n+1))}\right)_{n\ge 1} = \frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \frac{1}{20}, \ldots.$$

Das Verhalten der berechneten Folgenglieder legt die Vermutung nahe, dass es sich um eine monotone Nullfolge handelt. Die Monotonie

$$\frac{1}{n(n+1)} \ge \frac{1}{(n+1)(n+2)}$$

folgt aus $\frac{1}{n} \ge \frac{1}{n+2}$. Wir wollen nun bestätigen, dass es auch eine Nullfolge ist, und wählen dazu ein beliebiges $\varepsilon > 0$. Unsere Aufgabe ist es, eine Zahl $N(\varepsilon)$ zu finden, so dass

$$\frac{1}{n(n+1)} < \varepsilon \qquad \text{für alle } n = N(\varepsilon). \tag{12.1}$$

Nun kann man daran denken, die Ungleichung $\frac{1}{n(n+1)} < \varepsilon$ nach n umzuformen und so $N(\varepsilon)$ zu bestimmen. Das ist natürlich unbequem. Zum Glück müssen wir aber gar nicht das kleinstmögliche $N(\varepsilon)$ bestimmen, sondern nur irgend eins. Wir schätzen daher erst nach oben ab:

$$\frac{1}{n(n+1)} < \frac{1}{n^2}$$

und suchen nun $N(\varepsilon)$ so, dass

$$\frac{1}{n(n+1)} < \frac{1}{n^2} < \varepsilon \quad \text{für} \quad n = N(\varepsilon). \tag{12.2}$$

Nun ist $\frac{1}{n^2} < \varepsilon$ genau dann, wenn $n^2 > \frac{1}{\varepsilon}$ bzw. $n > \frac{1}{\sqrt{\varepsilon}}$. Wir können also für $N(\varepsilon)$ irgendeine natürliche Zahl wählen, die größer ist als $\frac{1}{\sqrt{\varepsilon}}$. Dann gilt (12.2) und damit erst recht (12.1). Also ist $\lim_{n\to\infty} \frac{1}{n(n+1)} = 0$, und nebenbei haben wir auch $\lim_{n\to\infty} \frac{1}{n^2} = 0$ erhalten. ∎

Beispiel 2 $\lim_{n\to\infty} \frac{1}{n^k} = 0 \quad$ für $k > 0$.

Satz 12.1 *Sei $d_n \ge d_{n+1} \ge 0$ für alle n. Dann sind die folgenden Aussagen äquivalent*

(a) Die d_n bilden eine monotone Nullfolge

(b) Es gibt eine monotone Nullfolge a_k so, dass es zu jedem k ein $M(k)$ gibt mit $d_{M(k)} \le a_k$

(c) Zu jedem $\varepsilon > 0$ in $\mathbb{R}$ gibt es ein $N(\varepsilon) \in \mathbb{N}$ mit $d_{N(\varepsilon)} \le \varepsilon$.

Beweis. Um $(a) \Rightarrow (b)$ zu zeigen, setze $a_k = d_k$ und $M(k) = N(\frac{1}{k})$. $(b) \Rightarrow (a)$: Nach Voraussetzung gibt es zu jedem $\varepsilon = \frac{1}{k}$ ein $L(\varepsilon)$ mit $a_{L(\varepsilon)} \le \varepsilon$. Setze nun $N(\varepsilon) = M(L(\varepsilon))$. Aus (c) folgt (a) trivialerweise - jedes $\varepsilon = \frac{1}{k}$ ist in (c) mitgemeint. Gelte nun (a), d.h. wir haben für jedes $\frac{1}{k}$ ein $N(\frac{1}{k})$ mit $d_{N(\frac{1}{k})} \le \frac{1}{k}$. Wir haben $N(\varepsilon)$ für jedes $\varepsilon > 0$ anuzgeben. Nach Archimedes gibt es ein $k \in \mathbb{N}$ mit $\varepsilon \ge \frac{1}{k}$ und wir setzen $N(\varepsilon) = N(\frac{1}{k})$. Dann

$$d_{N(\varepsilon)} = d_{N(\frac{1}{k})} \le \frac{1}{k} \le \varepsilon$$

12.3 Nullfolgen

Sehen wir eine Nullgolge d_n als einen (zeitdiskreten) Prozess, der sich auf den Wert Null einpendelt, so kommt es nur darauf an, dass die Abweichungen $|d_n|$ letztlich unter jede vorgegebene Schranke ε fallen, von endlich vielen Ausreissern abgesehen. Dies führt zu der Definition: Die Folge d_n ist eine *Nullfolge*, falls es zu jedem $\varepsilon > 0$ ein $N(\varepsilon)$ gibt so, dass

$$|d_n| \leq \varepsilon \quad \text{für alle } n \geq N(\varepsilon)$$

Da die Annahme der Monotonie weggefallen ist, bleibt uns die Formulierung "für alle $n \geq N(\varepsilon)$" nicht erspart. Beispiel und Gegenbeispiel

$$d_n = \begin{cases} \frac{1}{2^n} & n \text{ gerade} \\ -\frac{1}{n} & n \text{ ungerade} \end{cases} \quad \text{mit } N(\varepsilon) \geq \frac{1}{\varepsilon}, \qquad d_n = \begin{cases} n & n \text{ Primzahl} \\ \frac{1}{2^n} & n \text{ sonst} \end{cases}$$

Satz 12.2 *a_n ist genau dann eine Nullfolge, wenn $|a_n|$ eine Nullfolge ist. Das Produkt einer Nullfolge mit einer beschränkten Folge ist eine Nullfolge.*

Beweis der zweiten Aussage. Sei (a_n) eine Nullfolge und (b_n) eine beschränkte Folge. Wir wählen K so, dass $|b_n| \leq K$ für alle $n \in \mathbb{N}$. Für jedes n ist dann

$$|a_n b_n - 0| = |a_n b_n| = |a_n|\,|b_n| \leq K|a_n|. \tag{12.3}$$

Sei $\varepsilon > 0$ beliebig. Wegen $\lim a_n = 0$ finden wir ein $N(\varepsilon)$ so, dass

$$|a_n| < \varepsilon/K \quad \text{für alle } n \geq N(\varepsilon).$$

Wegen (12.3) ist dann

$$|a_n b_n - 0| < \varepsilon \quad \text{für alle } n \geq N(\varepsilon) \quad \blacksquare$$

.

12.4 Konvergenz von Folgen

Nun muss ja nicht immer alles gegen Null gehen, also definieren wir:

Die Folge c_n *konvergiert* gegen den *Limes* oder *Grenzwert* c, falls gilt

- Zu jedem $\varepsilon > 0$ in $\mathbb{R}$
  - gibt es ein $N = N(\varepsilon)$ in $\mathbb{N}$ so, dass
    - für alle $n \geq N(\varepsilon)$ in $\mathbb{N}$ gilt

$$|c - c_n| \leq \varepsilon \qquad (\text{ bzw. } |c - c_n| < \varepsilon)$$

Wir schreiben dann

$$c = \lim_{n\to\infty} c_n, \qquad c_n \to c \text{ für } n \to \infty$$

In der Literatur wird meist die äquivalente zweite Version mit $< \varepsilon$ benutzt. Beispiel: wird c durch die Intervallschachtelung $a_n \leq b_n$ approximiert, so gilt $c = \lim_{n\to\infty} c_n$ für jede Folge von "Näherungswerten" c_n so, dass $a_n \leq c_n \leq b_n$ fï alle n.

Umgekehrt haben wir für einen Limes $c = \lim_{n\to\infty} c_n$ die Approximation durch die Intervallschachtelung $a_n = c - |c_n| \leq b_n = c + |c_n|$. Also ist der Limes, wenn er denn existiert, durch die Folge eindeutig bestimmt.

Korollar 12.3 *Die Folge c_n konvergiert genau dann nicht gegen c, wenn es ein $\varepsilon > 0$ gibt so, dass es zu jedem $N \in \mathbb{N}$ ein $n \geq N$ gibt mit $|c_n - c| \geq \varepsilon$.*

Eine Folge c_n heisst *divergent*, wenn sie nicht konvergent ist, d.h. wenn sie gegen kein c konvergiert.

Ob wir bei einer Folge endlich viele Glieder weglassen oder hinzufügen, hat keinen Einfluss auf die Konvergenz. bzw. den Limes. Sind aber unendlich viele Glieder betroffen, so können die Folgen dramatisch sein (siehe $(-1)^n$).

Eine konvergente Folge ist immer beschränkt: Wähle $\varepsilon = 1$ und C als Maximum von 1 und den endlich vielen "Ausreissern" c_n mit $n < N(1)$. Dann $|c - c_n| \leq C$ für alle n, also $|c_n| \leq |c| + C$. Aber nicht jede beschränkte Folge ist konvergent: z.B. $c_n = (-1)^n$.

Hier sind die Rechenregeln für Grenzwerte.

Satz 12.4 *Seien $\left(a_n\right)_{n\geq 1}$ und $\left(b_n\right)_{n\geq 1}$ konvergente Folgen mit $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann sind auch die Folgen $\left(a_n + b_n\right)_{n\geq 1}$, $\left(a_n b_n\right)_{n\geq 1}$ und $\left(c a_n\right)_{n\geq 1}$ mit $c \in \mathbb{R}$ konvergent, und es gilt:*

$$\lim_{n\to\infty} (a_n + b_n) = a + b, \quad \lim_{n\to\infty} a_n b_n = ab, \quad \lim_{n\to\infty} c a_n = ca.$$

Ist $b_n \neq 0$ für alle $n \in \mathbb{N}$ und $b \neq 0$, so konvergiert auch die Folge $\left(\frac{a_n}{b_n}\right)_{n\geq 1}$, und es ist:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}.$$

Ist $a_n \leq b_n$ für alle $n \in \mathbb{N}$, so ist auch $a \leq b$.

Man beweist das, indem man die Bedingungen aus der Definition direkt überprüft. Ein Beweis im Kontext der Stetigkeit wird im nächsten Kapitel gegeben.

Beispiel 5. Für $a_n = \frac{n+1}{2n}$ ist

$$\lim_{n\to\infty} \frac{n+1}{2n} = \lim_{n\to\infty} \frac{1 + \frac{1}{n}}{2} = \frac{1 + \lim_{n\to\infty} \frac{1}{n}}{2} = \frac{1}{2}.$$

Beispiel 6. Es ist $\frac{1}{n} < \frac{2}{n}$, aber $\lim_{n\to\infty} \frac{1}{n} = 0 = \lim_{n\to\infty} \frac{2}{n}$.

Satz 12.5 (Einschließungskriterium) *Seien (a_n), (b_n), (c_n) Folgen reeller Zahlen und c_n zwischen a_n und b_n für alle $n \in \mathbb{N}$. Sind die Folgen (a_n) und (b_n) konvergent und haben sie den gleichen Grenzwert c, so konvergiert auch die Folge (c_n), und ihr Grenzwert ist ebenfalls c.*

Beweis: Zu $\varepsilon > 0$ gibt es n_a und n_b so, dass a_n und b_n zwischen $c - \varepsilon$ und $c + \varepsilon$ für alle $n \geq n_a$ bzw. $n \geq n_b$. Dann c_n zwischen $c - \varepsilon$ und $c + \varepsilon$ für alle $n \geq \max\{n_a, n_b\}$. $\square$

Beispiel 7. Wir benutzen dieses Kriterium, um zu zeigen, dass die Folge $\left(\sqrt[n]{n}\right)_{n\geq 1}$ konvergiert und dass

$$\lim_{n\to\infty} \sqrt[n]{n} = 1. \tag{12.4}$$

Wir zeigen dazu, dass die Folge (b_n) mit $b_n = \sqrt[n]{n} - 1$ eine Nullfolge ist. Dazu benötigen wir den *binomischen Satz*

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

mit den *Binomialkoeffizienten* $\binom{n}{k} := \frac{n!}{k!(n-k)!}$.

Mit diesem Satz erhalten wir zunächst

$$n = \left(\sqrt[n]{n}\right)^n = (1 + b_n)^n = 1 + \binom{n}{1} b_n + \binom{n}{2} b_n^2 + \ldots + \binom{n}{n} b_n^n.$$

Vernachlässigen wir alle Summanden bis auf den ersten und dritten, so folgt für $n \geq 2$

$$n > 1 + \binom{n}{2} b_n^2 = 1 + \frac{n(n-1)}{2} b_n^2,$$

so dass $b_n^2 < \frac{2}{n}$ ist. Da andererseits $b_n \geq 0$ ist, erhalten wir die Abschätzung

$$0 \leq b_n < \frac{\sqrt{2}}{\sqrt{n}} \quad \text{für } n \geq 2\,.$$

Wegen $\lim_{n\to\infty} \frac{\sqrt{2}}{\sqrt{n}} = 0$ liefert das Einschließungskriterium, dass auch $\lim_{n\to\infty} b_n = 0$ ist, d.h. (12.4) gilt. ■

12.5 Bestimmte Divergenz

Eine Folge a_n heisst

- *bestimmt divergent* gegen $+\infty$, wenn zu es jedem $C \in R$ ein N gibt mit $a_n \geq C$ für alle $n \geq N$
- *bestimmt divergent* gegen $-\infty$, wenn zu es jedem $C \in R$ ein N gibt mit $a_n \leq C$ für alle $n \geq N$

Die bestimmt divergenten Folgen sind also insbesondere divergent und nach unter bzw. oben beschränkt. Aus dem Archimedisches Prinzip folgt sofort

- Die Folge $a_n = na$ divergiert bestimmt gegen $+\infty$, falls $a > 0$
- Die Folge $a_n = na$ divergiert bestimmt gegen $-\infty$, falls $a < 0$

Mit ∞ wollen wir nicht rechnen. Aber wir haben folgende nützliche Beziehung

- Sind die $a_n > 0$, so ist a_n genau dann eine Nullfolge, wenn die Folge $\frac{1}{a_n}$ bestimmt (gegen $+\infty$) divergiert.

Der Beweis ergibt sich sofort daraus, dass $a_n \leq \varepsilon \Leftrightarrow \frac{1}{a_n} \geq \frac{1}{\varepsilon}$.

12 Folgen und Konvergenz: Konvergenzkriterien

12.6 Potenzen

Für eine reelle Zahl a ist rekursiv definiert

$$a^0 = 1,\ a^{n+1} = a^n \cdot a$$

Durch Induktion beweist man leicht

$$a^{n+m} = a^n a^m,\ a^{nm} = (a^n)^m,\ (ab)^n = a^n \cdot b^n$$

$$0 < a < b \Rightarrow a^n < b^n, \qquad n < m \Rightarrow \begin{cases} a^n < a^m & \text{falls } 1 < a \\ a^n > a^m & \text{falls } 0 < a < 1 \end{cases}$$

und mit etwas mehr Mühe findet man den “Binomischen Satz” für $(a+b)^n$. Uns langt einstweilen die folgende

$$\textit{Benoulli}\text{-Ungleichung} \quad (1+b)^n \geq 1 + nb \quad \text{für } b \geq -1$$

Diese folgt durch Induktion

$$(1+b)^{n+1} = (1+b)^n(1+b) \geq (1+nb)(1+b) = 1 + nb + b + nb^2 \geq 1 + (n+1)b$$

Satz 12.1 *(1) Falls $0 \leq a < 1$, so ist a^n monotone Nullfolge*

(2) Falls $-1 < a < 0$, so ist a^n alternierende Nullfolge

(3) Falls $a = 1$, so $a^n = 1$ konstant (mit Limes 1)

(4) Falls $a > 1$, so divergiert a^n bestimmt gegen $+\infty$,

(5) Falls $a = -1$, so ist a^n divergent, aber beschränkt

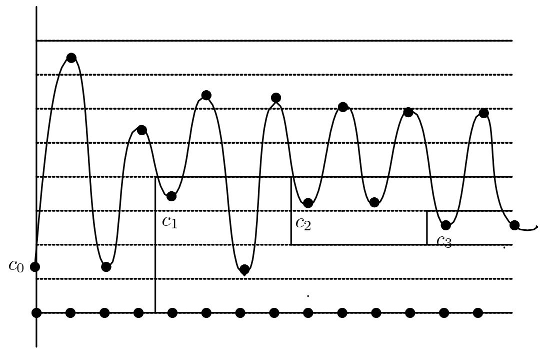
(6) Falls $a < -1$, so ist a^n divergent und weder nach oben noch nach unten beschränkt

Beweis. Wir zeigen zuerst (4): Wir haben $a = 1 + b$ mit $b > 0$, also $a^n \geq 1 + nb$ nach Bernoulli und somit $a_n \to +\infty$. (1) und (2) folgen sofort, indem wir die Kehrwerte betrachten. Der Rest ist einfach. □

12.7 Teilfolgen

Satz 12.2 Bolzano-Weierstrass. *Jede beschränkte Folge reeller Zahlen hat mindestens eine konvergente Teilfolge.*

Dabei heisst *Teilfolge*, dass einfach Folgenglieder ausgelassen werden, aber immer noch unendlich viele übrigbleiben.
Beweis. Sei $a \leq a_n \leq b$ für alle n. Wir bauen uns nach und nach eine Teilfolge c_k. Diese fängt mit $c_0 = a_0$ an. Dann halbieren wir $[a,b]$ fortlaufend. Bei jeder Halbierung muss mindenstens eine der Hälften noch unendlich viele Glieder der Folge enthalten. Wir wählen im k-ten Schritt eine solche Hälfte und fügen zu den in den vorangegangenen Schritten ausgewählten $c_0, \ldots, c_{k-1}$ als c_k das nächste Glied der Folge a_n aus dieser Hälfte hinzu (das können wir natürlich nicht wirklich tun, aber das Axion der bedingten Auswahl tut es für uns). Damit bekommen wir eine Teilfolge der Folge a_n, die von einer Intervallschachtelung eingeschlossen wird, also konvergiert. □



12.8 Monotone Konvergenz

Korollar 12.3 *Jede nach oben durch b beschränkte monoton wachsende Folge $(a_n)_{n\in\mathbb{N}}$ ist konvergent mit Limes a und $b \geq a \geq a_n$ für alle n. Jede nach unten durch b beschränkte monoton fallende Folge $(a_n)_{n\in\mathbb{N}}$ ist konvergent mit Limes a und $b \leq a \leq a_n$ für alle n.*

Aber über den Limes weiß man dann nicht viel mehr. Beweis. Wir betrachten nur den ersten Fall. Die Folge $(a_n)_{n\in\mathbb{N}}$ ist beschränkt, also hat sie nach Bolzano-Weierstrass eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit Limes a. Dann ist aber $(a_n - a)_{n\in\mathbb{N}}$ eine Nullfolge: zu $\varepsilon > 0$ gibt es $K(\varepsilon)$ mit

$$|a_{n_k} - a| < \varepsilon \quad \text{für alle } k \geq K(\varepsilon)$$

Wähle nun $N(\varepsilon) = n_{K(\varepsilon)}$. Dann

$$a - \varepsilon \leq a_{K(\varepsilon)} \leq a_n \leq a + \varepsilon \quad \text{für alle } n \geq N(\varepsilon).$$

Wäre $a < a_{n_0}$ für ein n_0, so $|a_n - a| = a_n - a > \varepsilon = \frac{1}{2}(a_{n_0} - a)$ für alle $n \geq n_0$ im Widerspruch zur Konvergenz. □.

Beispiel 8. Die Folge $(\frac{1}{n^2})_{n\in\mathbb{N}}$ ist monoton fallend und nach unten durch 0 beschränkt. Also konvergiert sie. ■

Beispiel 9. Wir wollen zeigen, dass die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = (1 + \frac{1}{n})^n$ konvergiert. Diese Folge *wächst monoton*: Für $n \geq 2$ ist nämlich

$$\frac{a_n}{a_{n-1}} = \left(\frac{n+1}{n}\right)^n \left(\frac{n-1}{n}\right)^{n-1} = \left(\frac{n^2-1}{n^2}\right)^n \frac{n}{n-1} = \left(1 - \frac{1}{n^2}\right)^n \frac{n}{n-1}.$$

Aus der Bernoullischen Ungleichung folgt

$$\left(1 - \frac{1}{n^2}\right)^n \geq 1 - \frac{n}{n^2} = 1 - \frac{1}{n}.$$

Damit erhalten wir

$$\frac{a_n}{a_{n-1}} \geq \left(1 - \frac{1}{n}\right) \frac{n}{n-1} = 1,$$

also $a_n \geq a_{n-1}$ für alle $n \geq 2$.

Die Folge (a_n) ist aber auch *nach oben beschränkt*: Aus dem binomischen Satz folgt zunächst

$$a_n = \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = 1 + \sum_{k=1}^{n} \binom{n}{k} \frac{1}{n^k}.$$

Für $k \geq 1$ schätzen wir ab:

$$\binom{n}{k} \frac{1}{n^k} = \frac{n!}{k!(n-k)!} \frac{1}{n^k} = \frac{1}{k!} \frac{n(n-1)\dots(n-k+1)}{n \cdot n \dots n} \leq \frac{1}{k!} \leq \frac{1}{2^{k-1}}.$$

Hieraus und mit der Summenformel für die geometrische Reihe folgt:

$$\begin{aligned} a_n &= 1 + \sum_{k=1}^{n} \binom{n}{k} \frac{1}{n^k} \leq 1 + \sum_{k=1}^{n} \frac{1}{2^{k-1}} = 1 + \sum_{k=0}^{n-1} \frac{1}{2^k} \\ &= 1 + \frac{1 - (1/2)^n}{1 - (1/2)} < 1 + \frac{1}{1 - (1/2)} = 3. \end{aligned}$$

Nach dem Monotoniekriterium existiert der Grenzwert $\lim_{n\to\infty}(1+\frac{1}{n})^n$. Dieser Grenzwert heißt *Eulersche Zahl* e. Die Zahl e ist irrational mit

$$e = 2.71\,82\,81\,82\,84\,5\ldots$$

Beispiel 10. Sei $a_n = \sum_{k=0}^n \frac{1}{k!}$. Die Folge $(a_n)_{n\geq 0}$ ist offenbar monoton wachsend, und sie ist auch nach oben beschränkt:

$$a_n = \sum_{k=0}^n \frac{1}{k!} = 1 + \sum_{k=1}^n \frac{1}{k!} \leq 1 + \sum_{k=1}^n \frac{1}{2^{k-1}} = 3$$

(vgl. Beispiel 9). Also konvergiert (a_n), und man kann zeigen, dass

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \sum_{k=0}^n \frac{1}{k!} = e.$$

12.9 Cauchyfolgen.

Wir wollen noch ein weiteres Konvergenzkriterium angeben, welches sich nicht nur auf monotone Folgen anwenden läßt, und welches außerordentlich wichtig für die Begründung der Analysis ist. Dazu definieren wir:

Definition 12.4 *Eine Folge $(a_n)_{n\geq 1}$ reeller Zahlen heißt* Cauchyfolge *oder* Fundamentalfolge*, wenn man zu jedem $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ so finden kann, dass*

$$|a_n - a_m| < \varepsilon \quad \textit{für alle } m,\, n \geq N(\varepsilon).$$

Die Folgenglieder müssen sich also in einem bestimmten Sinn immer näher kommen.

Satz 12.5 (Cauchy-Kriterium) *Eine Folge reeller Zahlen konvergiert genau dann, wenn sie eine Cauchyfolge ist.*

Die Tatsache, dass dieses Kriterium gilt, bezeichnet man als *Vollständigkeit* von $\mathbb{R}$.

Beweis. Es ist leicht zu sehen, dass jede konvergente Folge eine Cauchyfolge ist. Die umgekehrte Behauptung zeigt man in drei Schritten:

1. Schritt Cauchyfolgen sind beschränkt. Das ist ebenfalls leicht zu sehen: Wähle $\varepsilon = 42$. Dann gibt es ein $N(42)$ so, dass $|a_n - a_m| \leq 42$ für alle $n, m \geq N(42)$, also $|a_n - a_{42}| \leq 42$ für alle $n \geq N(42)$ und somit

$$|a_n| \leq c \text{ für alle } n \text{ wobei } c = \max\{|a_{N(42)} + 42, |a_1|, \ldots, |a_{N(42)-1}|\}$$

2. Schritt Nach Bolzano-Weierstraß besitzt eine beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen eine konvergente Teilfolge $a_{n_k} \to a$
3. Schritt $a_n \to a$: Sei $\varepsilon > 0$ gegeben. Wegen der Cauchy-Eigenschaft gibt es N mit

$$|a_n - a_m| < \frac{\varepsilon}{42} \quad \text{für alle } n, m \geq N$$

Aus der Teilfolgeneigenschaft und der Konvergenz der Teilfolge erhält man ein k mit

$$n_k \geq N \text{ und } |a_{n_k} - a| < \frac{\varepsilon}{42}$$

Es folgt mit der Dreiecksungleichung

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\varepsilon}{42} + \frac{\varepsilon}{42} \leq \varepsilon \quad \text{für alle } n \geq N \quad \blacksquare$$

Beispiel 11. Für $n \geq 1$ sei $a_n = \sum_{k=1}^n (-1)^{k-1}\frac{1}{k}$. Wir zeigen, dass $(a_n)_{n\geq 1}$ eine Cauchyfolge ist. Für $m \geq n+2$ ist

$$\begin{aligned} |a_m - a_n| &= \left|\sum_{k=n+1}^m (-1)^{k-1}\frac{1}{k}\right| = \left|(-1)^n \frac{1}{n+1} + \sum_{k=n+2}^m (-1)^{k-1}\frac{1}{k}\right| \\ &= \left|\frac{1}{n+1} - \sum_{k=n+2}^m (-1)^{k-n}\frac{1}{k}\right|. \end{aligned} \tag{12.1}$$

Weiter ist

$$\sum_{k=n+2}^m (-1)^{k-n}\frac{1}{k} = \left(\frac{1}{n+2} - \frac{1}{n+3}\right) + \left(\frac{1}{n+4} - \frac{1}{n+5}\right) + \cdots > 0$$

sowie

$$\sum_{k=n+2}^m (-1)^{k-n}\frac{1}{k} = \frac{1}{n+2} - \left(\frac{1}{n+3} - \frac{1}{n+4}\right) - \left(\frac{1}{n+5} - \frac{1}{n+6}\right) - \cdots < \frac{1}{n+2}.$$

Hieraus und aus (12.1) folgt $|a_m - a_n| \leq \frac{1}{n+1}$, und dies gilt sogar für alle $m \geq n$. Ist nun $\varepsilon > 0$ beliebig vorgegeben, so wählen wir N so, dass $\frac{1}{N+1} < \varepsilon$. Für alle $m, n \geq N$ ist dann

$$|a_n - a_m| \leq \frac{1}{n+1} \leq \frac{1}{N+1} < \varepsilon.$$

Nach dem Cauchy-Kriterium ist die Folge $(a_n)_{n\geq 1}$ konvergent. Man kann zeigen, dass der Grenzwert dieser Folge gleich $\ln 2$ ist. $\blacksquare$

12 Folgen und Konvergenz: Vektorfolgen

12.10 Vektorfolgen und komplexe Zahlenfolgen

Statt Zahlenfolgen können wir genauso Folgen $\vec{v}_n$ $(n \in \mathbb{N})$ von Vektoren $\vec{v}_n$ in einem **endlichdimensionalen** euklidischen Vektorraum V bzw. den Spezialfall $V = \mathbb{C}$ betrachten. Die Folge heißt *konvergent* gegen $\vec{w}$ und $\vec{w}$ *Limes* der Folge, falls gilt

- Zu jedem $\varepsilon > 0$ in $\mathbb{R}$
  - gibt es ein $N = N(\varepsilon) \in \mathbb{N}$ so,dass

∗ für alle $n \geq N(\varepsilon)$ gilt
$$\|\vec{v}_n - \vec{w}\| < \varepsilon.$$

Der Vektor $\vec{w}$ (falls er existiert) ist eindeutig bestimmt und man schreibt

$$\vec{w} = \lim_{n\to\infty} \vec{n}_n, \quad \vec{v}_n \to \vec{w} \text{ für } n \to \infty$$

Die Eindeutigkeit sieht man so: Gälte $\vec{v}_n \to \vec{u}$ mit $\vec{u} \neq \vec{w}$, so $\varepsilon = \frac{1}{3}\|\vec{w}-\vec{u}\| > 0$. Dann gibt es N_w und N_u so, dass $\|\vec{v}_n - \vec{w}\| < \varepsilon$ für alle $n \geq N_w$ und $\|\vec{v}_n - \vec{u}\| < \varepsilon$ für alle $n \geq N_u$. Setze $n = \max\{N_w, N_u\}$. Dann folgt mit der Dreiecksungleichung der Widerspruch

$$\|\vec{w} - \vec{u}\| = \|\vec{w} - \vec{v}_n + \vec{v}_n - \vec{u}\} \leq \|\vec{w} - \vec{w}_n\| + |\vec{v}_n - \vec{u}\| \leq \frac{2}{3}\varepsilon < \|\vec{w} - \vec{u}\|$$

Satz 12.1 *Seien Vektoren $\vec{v}_1, \ldots, \vec{v}_m$ in V, die Zahlenfolgen $(x_{1n})_{n\in\mathbb{N}}, \ldots, (x_{mn})_{n\in\mathbb{N}}$ und die Vektorfolge $(\vec{x}_n)_{n\in\mathbb{N}}$ mit $\vec{x}_n = x_{1n}\vec{v}_1 + \ldots + x_{mn}\vec{v}_m$ gegeben. Hat man $\lim_{n\to\infty} x_{im} = y_i$ für $i = 1, \ldots, m$, so gilt $\lim_{n\to\infty} \vec{x}_n = \vec{y}$. Bilden die $\vec{v}_1, \ldots, \vec{v}_m$ ein Orthonormalsystem, so gilt auch die Umkehrung. Insbesondere folgt aus der Konvergenz von $(\vec{x}_n)_{n\in\mathbb{N}}$ die Konvergenz der $(x_{in})_{n\in\mathbb{N}}$.*

Korollar 12.2 *Jede beschränkte Folge hat eine konvergente Teilfolge und jede Cauchyfolge konvergiert.*

Beweis des Satzes. Zum ersten Teil: Sei $\varepsilon > 0$ gegeben. Setze

$$\varepsilon_i = \frac{\varepsilon}{m\|\vec{v}_i\|}$$

und wähle N_i so, dass
$$|x_{in} - y_n| \leq \varepsilon_i \quad \text{für alle } n \geq N_i$$
Dann mit der Dreiecksungleichung für $n \geq N = \max\{N_1, \ldots, N_m\}$

$$\|\vec{x}_n - \vec{y}\| = \|\sum_{i=1}^{m}(x_{in} - y_i)\vec{v}_i\| \leq \sum_{i=1}^{m}\|(x_{in} - y_i)\vec{v}_i\| = \sum_{i=1}^{m}|x_{in} - y_i| \cdot \|\vec{v}_i\| \leq \varepsilon$$

In der umgekehrten Richtung sei angenommen, dass es für ein i_0 ein $\varepsilon > 0$ gibt so, dass es zu jedem N ein $n \geq N$ gibt mit $|x_{i_0 n} - y_{i_0}| > \varepsilon$ und somit auch

$$\|\vec{x}_n - \vec{y}\| = \sqrt{\sum_{i=1}^{m}(x_{in} - y_i)^2} \geq |x_{i_0 n} - y_{i_0}| > \varepsilon$$

ein Widerspruch zur Konvergenz.

13 Funktionen und Stetigkeit: Grundlegendes

Wir betrachten nur Funktionen mit reellem Definitionsbereich. Auch die Funktionswerte sind zunächst nur reelle Zahlen. Später lassen wir auch Vektoren bzw. komplexe Zahlen als Funktionswerte zu.

13.1 Funktionen

Eine reelle *Funktion* mit *Definitionsbereich* $D = D(f)$ ist durch ihren *Graphen* gegeben, eine Teilmenge G von $\mathbb{R}^2$ so, dass gilt

- *Wohldefiniertheit:* $\forall x \in D\, \forall y, y' \in \mathbb{R}.\ \begin{pmatrix} x \\ y \end{pmatrix} \in G$ und $\begin{pmatrix} x \\ y' \end{pmatrix} \in G \Rightarrow y = y'$
- *Überalldefiniertheit:* $\forall x \in D\, \exists y \in \mathbb{R}.\ \begin{pmatrix} x \\ y \end{pmatrix} \in G$

Man schreibt dann
$$f : D \to \mathbb{R}, \quad f(x) = y \;\Leftrightarrow\; \begin{pmatrix} x \\ y \end{pmatrix} \in G$$

Der *Bildbereich*, kurz das *Bild* von f ist $B(f) = \{f(x) \mid x \in D\}$. $f : D \to \mathbb{R}$ ist

- *nach oben beschränkt* falls es ein C gibt mit $f(x) \leq C$ für alle $x \in D$
- *nach unten beschränkt* falls es ein C gibt mit $f(x) \geq C$ für alle $x \in D$
- *beschränkt* falls es ein C gibt mit $|f(x)| \leq C$ für alle $x \in D$

Ist $D' \subseteq D$, so kann man f auf D' *einschränken* und erhält

$$g : D' \to \mathbb{R} \quad \text{mit} \quad g(x) = f(x) \quad \text{für alle } x \in D'$$

- aber man wird oft für f und g dieselbe Bezeichnung benutzen und den Definitionsbereich im Kontext angeben. Hat man auf $D'' \supseteq D$ eine Funktion $h : D'' \to \mathbb{R}$, die f als Einschränkung auf D hat, so ist h eine *Fortsetzung* von f auf D''. Ist $D \neq D''$, so ist h nicht eindeutig bestimmt; andererseits wird man in der Regel kein h finden, dass zu der gleichen Funktionenklasse wie f gehört.

13.2 Stetige Grössen

In der Physik und in Folge auch in den anderen Naturwissenschaften und in der Technik betrachtet man "skalare Grössen" wie Masse, Länge, Druck, Temperatur, Intelligenz, Zeit, Lautstärke, Erdbebenstärke von denen man meist annimmt, dass sie in bestimmten Grenzen "stetig variieren". Legt man eine Skalierung fest, so werden die Werte dieser Grössen durch reelle Zahlen ausgedrückt. "Stetige Variation" bedeutet, das man sich vorstellt, die Schwankung eine Grösse um einen jeweils festen Wert klein und immer noch kleiner halten zu können und auch den Übergang zwischen zwei deutlich unterschiedlichen Werten durch eine Aufeinanderfolge endlich vieler winziger Schritte bewerkstelligen zu können. Eine radikale Sicht ist die, dass eine Grösse nicht anderes ist als die phantasievoll benannte Funktion $f(x) = x$ auf einem Intervall z.B. $[a, b]$ von reellen Zahlen - d.h. x variiert zwischen den Grenzen a und b: $a \leq x \leq b$.

Was auch immer eine Grösse ist, wichtig ist der Begriff der Annäherung der Grösse x an den festen Wert c, mit Verlaub schreiben wir $x \to c$. Was auch immer das bedeutet, wir haben dann
$$\Delta x \to 0 \quad \text{genau dann, wenn} \quad x = c + \Delta x \;\to c$$

d.h. wir erfinden eine neue Grösse Δx für Differenzen der Werte der Grösse x. Richtiger wäre es, auch noch die Stelle c zu notieren, d.h. z.B. $\Delta(x,c)$ zu schreiben. Präzisere Möglichkeiten zu sagen, dass wir uns für x in der "Nähe" von c interessieren sind

- $|x-c| \leq \delta$ für "kleines" δ
- betrachte Folgen x_n mit $\lim_{n\to\infty} x_n = c$

Der Begriff der Stetigkeit einer Operation oder Funktion soll das Prinzip "kleine Ursache, kleine Wirkung" fassen. Ist die Grösse y eine Funktion $y = f(x)$ der Grösse x, so findet man das so geschrieben

$$\Delta y \to 0 \quad \text{für} \quad \Delta x \to 0$$

oder etwas genauer

$$y = f(x) \to f(c) \quad \text{für} \quad x \to c$$

13.3 Fehlerabschätzung

Bei naturwissenschaftlichen Messungen sind die Messwerte der Grössen x_i meist mit gewissen Fehlern Δx_i behaftet, die man nach Kenntnis des Messverfahren als durch δ_i beschränkt annimmt: $|\Delta x_i| \leq \delta_i$. Bei der Auswertung setzt man diese Werte in Formeln ein, um abgeleitete Grössen, z.B. y, zu berechnen, d.h. man wendet gewisse Operationen und Funktionen auf die Messwerte c_i an und erhält

$$d = f(c_1, \dots c_n)$$

Aufgrund der Fehler in den Messwerten, muss man mit einem Fehler Δy rechnen und versuchen, für diesen eine obere Schranke S zu garantieren. Dies gelingt, falls f stetig ist - und man die Stetigkeit effektiv nachgewiesen hat. Dazu ein Beispiel

$$y = (\cos x_1) \cdot \sin(x_2 + x_3)$$

Wir setzen

$$z_1 = \cos x_1, \;\; z_2 = \sin(x_2 + x_3), \;\; y = z_1 \cdot z_2, \;\; d_1 = \cos c_1, \;\; d_2 = \sin(c_2 + c_3)$$

und haben, wie wir später zeigen werden, die Abschätzungen

$$|\Delta z_1| \leq |\Delta x_1| \leq \delta_1, \;\; |\Delta z_2| \leq |\Delta(x_2 + x_3)| \leq |\Delta x_2| + |\Delta x_3| \leq \delta_2 + \delta_3$$

$$\begin{aligned} |\Delta y| &\leq |\Delta z_1||d_2| + |\Delta z_2||d_1| + |\Delta z_1 \Delta z_2| \\ &\leq \delta_1|d_2| + (\delta_2 + \delta_3)|d_1| + \delta_1(\delta_2 + \delta_3) =_{def} S(c_1, c_2, c_3) \end{aligned}$$

Hier können wir auch noch die $|d_i|$ abschätzen durch $|d_i| \leq 1$ und erhalten eine von den c_i unabhängige Schranke

$$S = \delta_1 + \delta_2 + \delta_3 + \delta_1\delta_2 + \delta_1\delta_3 \geq |\Delta y|$$

Umgekehrt sagt uns das, wie klein wir die δ_i machem müssen (d.h. wieweit wir unsere Messgenauigkeit verbessern müssen) um eine vorgegebene Schranke ε für den Fehler Δy einzuhalten.

13.4 Stetigkeit an einer Stelle

Wir sagen: f ist an der Stelle c *stetig* und definieren das auf die folgenden beiden äquivalenten Weisen

(1) Für alle Folgen x_n mit $\lim_{n\to\infty} x_n = c$ gilt

$$\lim_{n\to\infty} f(x_n) = f(c)$$

(2) Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass

- für alle x mit $|x-c| \leq \delta$ gilt: $|f(x) - f(c)| \leq \varepsilon$

In der Tat, gelte (1). Wir nehmen an dass (2) nicht gilt. Dann gibt es ein $\varepsilon > 0$ und zu jedem $\delta = \frac{1}{n}$ ein x_n mit $|x_n - c| \leq \delta$ aber $|f(x_n) - f(c)| > \varepsilon$. Es folgt, dass $\lim_{n\to\infty} x_n = c$ aber $f(x_n)$ sicher nicht gegen $f(c)$ konvergiert. Widerspruch. Also gilt (2). Umgekehrt sei (2) angenommen und $x_n \to c$. Ist $\varepsilon > 0$ gegeben, so gibt es nach (2) ein δ so, dass $|f(x) - f(c)| < \varepsilon$ für alle x mit $|x - c| < \delta$. Da $x_n \to c$, gibt es ein n_0 so, dass $|x_n - c| < \delta$ für alle $n \geq n_0$, also auch $|f(x_n) - f(c)| < \varepsilon$. □ Die Bedingung (2) kann man mit mehr Pathos auch so lesen

(2') Für jedes noch so kleine $\varepsilon > 0$ kann man ein (hinreichend kleines) $\delta > 0$ so finden, dass $|\Delta x| \leq \delta$ (an der Stelle c) schon garantiert dass, $|\Delta y| \leq \varepsilon$ an der Stelle $f(c)$

Korollar 13.1 *Ist die Folge $a_n \in D$ mithilfe der auf D definierten Funktion $y = f(x)$ rekursiv definiert, d.h. $a_{n+1} = f(a_n)$, existiert $a = \lim_{n\to\infty} a_n \in D$ und ist f an der Stelle a stetig, so ist a ein* Fixpunkt *von f, d.h. $f(a) = a$.*

13.5 Stetigkeit und stetige Fortsetzung

Ist f auf D definiert und an jeder Stelle $c \in D$ stetig, so nennt man f eine (auf D) *stetige* Funktion. Ist $D' \supseteq D$ und $g : D' \to \mathbb{R}$ eine Fortsetzung von f und stetig auf D', so nennt man g eine *stetige Fortsetzung* von f. Insbesondere ist dann auch f stetig.

$c \in \mathbb{R}$ heißt ein *Häufungspunkt* von D, wenn es beliebig nahe bei c noch von c verschiedene Punkte von D gibt, d.h. alternativ

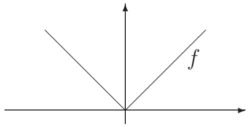
- Zu jeden $\varepsilon > 0$ gibt es $c \neq d \in D$ mit $|c - d| < \varepsilon$
- Es gibt $d_n \in D$ mit $d_n \neq c$ und $d_n \to c$.

Ist $D'' \supseteq D$ und besteht $D'' \setminus D$ nur aus Häufungspunkten von D, so ist D *dicht* in D'' - z.B. ist $\mathbb{Q}$ dicht in $\mathbb{R}$.

Korollar 13.2 *Ist D dicht in D'', so hat $f : D \to \mathbb{R}$ höchstens eine stetige Fortsetzung $h : D'' \to \mathbb{R}$.*

Beweis. Zu jedem $c \in D''$ gibt es $d_n \to c$ mit $d_n \in D$. Mit der Stetigkeit von h folgt $h(c) = \lim_{n\to\infty} h(d_n) = \lim_{n\to\infty} f(d_n)$ - der eindeutig bestimmte Limes. □

Beispiel 1: *Die Betragsfunktion*
Das ist die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ mit dem folgenden Graphen:



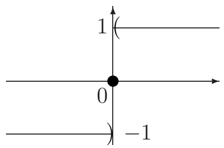
Sie ist auf ganz $\mathbb{R}$ stetig, wie man leicht mit der Dreiecksungleichung erhält. Aus $x_n \to x_0$ folgt nämlich wegen

$$\big||x_n| - |x_0|\big| \le |x_n - x_0| \to 0\,,$$

dass auch $|x_n| \to |x_0|$. ∎

Beispiel 2: *Die Signumfunktion*
Das die Funktion

$$\operatorname{sgn}\,:\mathbb{R}\to\mathbb{R},\quad x\mapsto\begin{cases}1 & \text{falls } x>0\\ 0 & \text{falls } x=0\\ -1 & \text{falls } x<0\end{cases}\;.$$



Schränkt man auf $\mathbb{R}\backslash\{0\}$ ein, so erhält man eine stetige Funktion. Diese erlaubt aber keine stetige Fortsetzung $h:\mathbb{R}\to\mathbb{R}$ - für die müsste gelten: $1=\lim_{n\to\infty}\operatorname{sgn}\frac{1}{n}=\lim_{n\to\infty}h(\frac{1}{n})=h(0)=\lim_{n\to\infty}h(-\frac{1}{n})=\lim_{n\to\infty}\operatorname{sgn}-\frac{1}{n}=-1$.

14 Funktionen und Stetigkeit: Beispiele

14.8 Stetigkeit der arithmetischen Operationen

Satz 14.1 *Addition und Multiplikation reeller Zahlen sind stetig (wobei man beide Argumente simultan und unabhängig voneinander variieren darf. Die Inversion ist stetig, soweit sie ausführbar ist, d.h. für $x\neq 0$.*

Die Abschätzungen für $+$ und $\cdot$ sind klar. Für die Inversion folgt sie aus $\frac{c^2}{2}\le|c(c+\Delta x)|$. Es folgt Satz 12.4.

$$z = x+y$$

$$|\Delta z|\le|\Delta x|+|\Delta y|$$

$$z = xy \qquad\qquad y=\frac{1}{x}$$

$$|\Delta z|\le|c\Delta x|+|d\Delta y|+|\Delta x\Delta y| \qquad\qquad |\Delta y|\le\frac{2}{|c|^2}|\Delta x|\quad\text{für } |\Delta x|\le\frac{|c|}{2}$$

14.9 Substitution

Hat man zwei Funktionen $y=f(x)$ und $z=g(y)$, so kann man z.B. f in g *substituieren* und erhält die *Verkettung* $g\circ f$ (lies g nach f) mit

$$(g\circ f)x = g(f(x))$$

Satz 14.2 *Ist f an der Stelle c und g an der Stelle $d=f(c)$ stetig, so ist $g\circ f$ an der Stelle c stetig.*

Beweis. Sei $\varepsilon>0$ gegeben. Da g an der Stelle d stetig ist, gibt es $\gamma>0$ so, dass

$$|g(y)-g(d)|\le\varepsilon\quad\text{für alle } y \text{ mit } |y-d|\le\gamma$$

Da f an der Stelle c stetig ist, gibt es $\delta>0$ mit

$$|f(x)-f(c)|\le\gamma\quad\text{für alle } x \text{ mit } |x-c|\le\delta$$

Nun $d=f(c)$, also gilt für alle x mit $|x-c|\le\delta$

$$|g(f(x))-g(f(c))|=|g(y)-g(d)|\le\varepsilon\;\text{ mit } y=f(x) \text{ also } |y-d|\le\gamma\quad\square$$

14.10 Rationale Komposition

Sind $f,g:X\to\mathbb{R}$ Funktionen, so erhält man ihre *Summe* $f+g$, ihr *Produkt* fg und – falls $g(x)\neq 0$ für alle $x\in X$ – ihren *Quotienten* f/g durch

$$(f+g)(x)=f(x)+g(x),\quad (fg)(x)=f(x)g(x),\quad (f/g)(x)=f(x)/g(x)\,.$$

Korollar 14.3 *Sind $f,g:X\to\mathbb{R}$ in x_0 stetig, so sind auch $f+g$, fg und – falls $g(x)\neq 0$ für $x\in X$ – auch f/g in x_0 stetig.*

Beispiel 3: *Polynome*
Sei $n\in\mathbb{N}_0$, $a_0,\ldots,a_n\in\mathbb{R}$ und $a_n\neq 0$. Die Funktion $f:\mathbb{R}\to\mathbb{R}$,

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0=\sum_{i=0}^{n}a_ix^i$$

heißt *Polynom n. Grades*, und die a_i heißen seine *Koeffizienten*. Polynome sind auf ganz $\mathbb{R}$ stetig.

Lemma 14.4 *Ist $f(x)=a_nx^n+\ldots+a_1x+a_0$ ein Polynom vom Grad n und $x_0\in\mathbb{R}$, so gibt es ein Polynom $g(x)=b_{n-1}x^{n-1}+\ldots+b_1x+b_0$ vom Grad $n-1$ und ein $c\in\mathbb{R}$ mit*

$$f(x)=(x-x_0)g(x)+c\quad\textit{für alle } x\in\mathbb{R}. \tag{14.1}$$

Aus (14.1) folgt $c=f(x_0)$. Um (14.1) zu zeigen, vergleichen wir die Koeffizienten der Polynome auf der linken bzw. rechten Seite von (14.1):

$$\begin{array}{lrcl}
\text{bei } x^n: & a_n & = & b_{n-1}\\
\text{bei } x^{n-1}: & a_{n-1} & = & b_{n-2}-x_0b_{n-1}\\
 & & \vdots & \\
\text{bei } x^1: & a_1 & = & b_0-x_0b_1\\
\text{bei } x^0: & a_0 & = & c-x_0b_0\,.
\end{array}$$

Hieraus berechnet man $b_{n-1}, b_{n-2}, \ldots, b_1, b_0$ und c:

$$\begin{aligned} b_{n-1} &= a_n \\ b_{n-2} &= a_{n-1} + x_0 b_{n-1} \\ &\vdots \\ b_0 &= a_1 + x_0 b_1 \\ c &= a_0 + x_0 b_0 \,. \end{aligned}$$

■

Ein bequemes Berechnungsverfahren bietet das *Hornerschema*:

$$\begin{array}{llllll} a_n & a_{n-1} & a_{n-2} & \ldots & a_1 & a_0 \\ & x_0 b_{n-1} & x_0 b_{n-2} & \ldots & x_0 b_1 & x_0 b_0 \\ \hline b_{n-1} & b_{n-2} & b_{n-3} & \ldots & b_0 & c\,. \end{array}$$

Beispielsweise ist für $f(x) = x^5 + 2x^4 - 12x + 5$ und $x_0 = 2$:

$$\begin{array}{rrrrrr} 1 & 2 & 0 & 0 & -12 & 5 \\ & 2 & 8 & 16 & 32 & 40 \\ \hline 1 & 4 & 8 & 16 & 20 & 45 \end{array}$$

Also ist $f(2) = 45$ und $x^5 + 2x^4 - 12x + 5 = (x-2)(x^4 + 4x^3 + 8x^2 + 16x + 20) + 45$. ■

Satz 14.5 (Identitätssatz, Koeffizientenvergleich) *Gilt* $a_n x_k^n + \ldots a_1 x_k + a_0 = a'_n x_k^n + \ldots + a'_1 x_k + a'_0$ *für midnestens* $n+1$ *verschiedene* $x_k \in \mathbb{R}$, *so ist* $a_i = a'_i$ *für* $i = 0, \ldots, n$.

Beweis. Wir beweisen durch Induktion über n: Hat $f(x) = \sum_{i=0}^n a_i x^i$ $n+1$ verschiedene Nullstellen, so gilt $a_0 = \ldots = a_n = 0$. Der Fall $n = 0$ ist trivial. Im Induktionsschritt seien die Nullstellen $x_0, \ldots, x_n$ gegeben und die $b_0, \ldots, b_{n-1}, c$ und $g(x)$ wie im Beweis des Lemmas bestimmt. Dann gilt $f(x) = (x-x_0)g(x)+c$ für alle x, insbesondere $0 = f(x_0) = c$ und $0 = f(x_k) = (x_k - x_0)g(x_k)$ für $k > 0$ also $g(x_k) = 0$. Nach Induktionsannahme folgt $b_0 = \ldots = b_{n-1} = 0$ und daraus $a_0 = \ldots = a_n = 0$. Der allgemeine Fall ergibt sich, indem man $(a_n - a'_n)x_k^n + \ldots + (a_0 - a'_0) = 0$ betrachtet. □

Korollar 14.6 *Die Zerlegung* $f(x) = (x - x_0)g(x) + c$ *im Lemma ist eindeutig.*

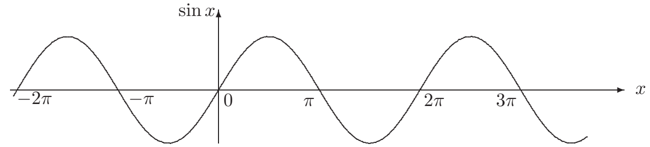
Ist x_0 Nullstelle von f, so kann f faktorisiert werden: $f(x) = (x - x_0)g(x)$. Mehrfache Anwendung dieser Überlegung zeigt: ein Polynom vom Grad n hat höchstens n Nullstellen.

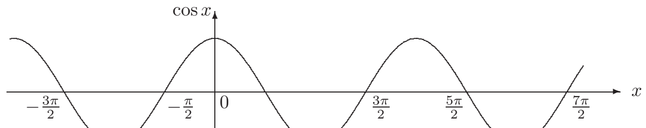
Ist also x_0 eine Nullstelle von f, so gibt es ein Polynom g_1 so, dass $f(x) = (x - x_0)g_1(x)$. Ist x_0 auch Nullstelle von g_1, so gibt es ein Polynom g_2 so, dass $f(x) = (x-x_0)^2 g_2(x)$. Wir fahren so fort und finden eine Zahl $\ell \in \mathbb{N}$ mit $f(x) = (x-x_0)^\ell g_\ell(x)$, aber $g_\ell(x_0) \neq 0$. Dann heißt ℓ die *Ordnung der Nullstelle* x_0 von f. Analog heißt x_0 eine *Polstelle* ℓ. Ordnung der rationalen Funktion $f = g/h$, wenn $g(x_0) \neq 0$ und wenn x_0 Nullstelle ℓ. Ordnung von h ist.

Beispiel 4: *Rationale Funktionen*
Eine rationale Funktion ist der Quotient $f(x)/g(x)$ zweier Polynome f, g. Sie ist auf $\{x \in \mathbb{R} : g(x) \neq 0\}$ erklärt und nach Satz 14.3 und Beispiel 3 auf dieser Menge stetig. ■

14.11 Sinus und Co

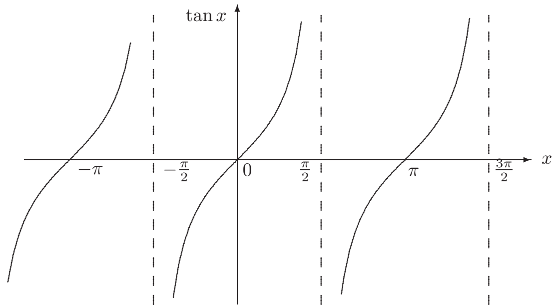
Wir können die Stetigkeit der Funktionen Sinus und Cosinus zeigen, wenn wir akzeptieren, dass die Länge des Bogens mindestens die Länge der Sekante ist. In beiden Fällen haben wir

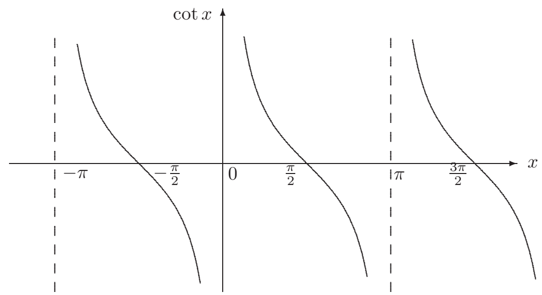




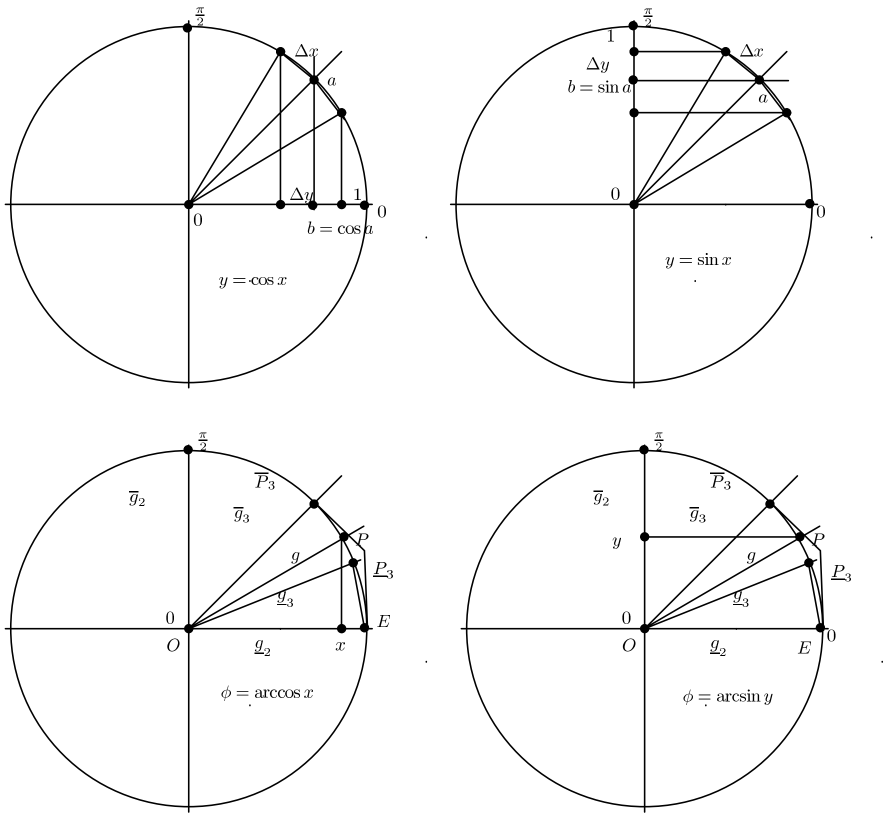
Wir erklären die *Tangensfunktion* tan und die *Cotangensfunktion* cot durch

$$\begin{aligned} \tan: \ & \mathbb{R}\backslash(\tfrac{\pi}{2} + \pi\mathbb{Z}) \to \mathbb{R}, & x \mapsto \tfrac{\sin x}{\cos x} \\ \cot: \ & \mathbb{R}\backslash\pi\mathbb{Z} \to \mathbb{R}\,, & x \mapsto \tfrac{\cos x}{\sin x}\,. \end{aligned}$$





$|\Delta y| \leq |\Delta x|$



Beide Funktionen sind periodisch mit der Periode π:

$$\tan(x + \pi) = \tan x\,, \quad \cot(x + \pi) = \cot x\,,$$

und beide Funktionen sind ungerade:

$$\tan(-x) = -\tan x\,, \quad \cot(-x) = -\cot x\,.$$

Für alle zulässigen x, y gelten die Additionstheoreme

$$\begin{aligned} \tan(x + y) &= \frac{\tan x + \tan y}{1 - \tan x\, \tan y} \\ \cot(x + y) &= \frac{\cot x\, \cot y - 1}{\cot x + \cot y}\,. \end{aligned}$$

Schließlich sind nach Satz 14.3 beide Funktionen tan und cot stetig.

13 Funktionen und Stetigkeit: Eigenschaften

13.6 Zwischenwerte

Nach eine alten Dogma macht die Natur keine Sprünge und haben alle Funktionen, die Naturvorgänge beschreiben sollen, zumindestens stetig zu sein. Wenn Sie z.B. eine Reaktion beobachten, so nimmt die Zahl der Moleküle, für die Reaktion erfolgt ist, mit der Zeit stetig zu: 0 Moleküle zur Zeit 0, 100 zur Zeit 1, ein Tausendstel Molekül zur Zeit $\frac{1}{10}$ usw. wer's glaubt wird seelig. Aber: Stetigkeit ist eine nützliche und oft auch sinnvolle Fiktion. Man kann sie auch so deuten: kein möglicher Wert ist unmöglich. In der Tat

Satz 13.1 Zwischenwertsatz. *Ist f stetig auf dem Intervall $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ und liegt d zwischen $f(a)$ und $f(b)$, so gibt es ein $c \in [a, b]$ mit $f(c) = d$.*

Beweis. Es gibt eine Intervallschachtelung a_n, b_n so, dass d für alle n zwischen $f(a_n)$ und $f(b_n)$ liegt. Dazu $a_0 = a, b_0 = b$ und (mit dem Prinzip der bedingten Auswahl) $a_{n+1}, b_{n+1} \in \{a_n,\ (a_n + b_n)/2,\ b_n\}$ so, dass d zwischen $f(a_{n+1})$ und $f(b_{n+1})$ liegt. Für die durch die Intervallschachtelung definierte reeelle Zahl c gilt $c = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$, also mit der Stetigkeit von f auch

$$f(c) = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n)$$

Mit dem Einschließungskriterium (für $d_n = d$ zwischen $f(a_n)$ und $f(b_n)$) folgt $f(c) = d$. □

13.7 Extremwerte

Satz 13.2 *Eine auf einem Intervall $[a, b]$ stetige Funktion ist beschränkt und nimmt ihr Maximum bzw. Minimum je mindestens einmal an.*

Beweis. Angenommen, f ist nicht nach oben beschränkt. Dann gibt es zu jedem n ein x_n in $[a,b]$ mit $f(x_n) \geq n$. Nach Bolzano-Weierstrass können wir zu einer Teilfolge p_k mit $f(p_k) \geq k$ übergehen, die gegen ein $p \in [a,b]$ konvergiert. Dann $f(p) = \lim_{k\to\infty} f(p_k) \geq m$ für alle m, Widerspruch. Ebenso sieht man, dass f nach unten beschränkt ist.
Wir haben also c_0, d_0 mit $c_0 \leq f(x) \leq d_0$ für alle $x \in [a,b]$. Wir halbieren nun das Intervall $[c_0, d_0]$ fortlaufend so, dass es zu $[c_n, d_n]$ ein $x_n \in [a,b]$ gibt mit $c_n \leq f(x_n)$ und $f(x) \leq d_n$ für alle $x \in [a,b]$ - genau eine der beiden Hälften tut's jeweils. Die Intervallschachtelung liefert uns die obere Schranke $M = \lim_{n\to\infty} d_n$ von f und $M = f(x)$, wobei x der Limes einer konvergenten Teilfolge der x_n ist. $\square$.

Beispiel: Sei

$$Q(x_1, x_2) = ax_1^2 + bx_2^2 + cx_1x_2, \quad f(x) = Q(\cos x, \sin x) \; x \in [0, 2\pi]$$

Dann nimmt f mit Garantie Maximum und Minimum an, sogar je genau zweimal, wenn's nicht konstant ist. Man betrachtet hier eine "quadratische Form" auf dem Einheitskreis; die Vektoren, an denen die Extremwerte angenommen werden, heissen "Eigenvektoren". Diese werden Sie im zweiten Semester kennenlernen, oder in der Physik z.B. als Hauptträgheitsachsen. Das Schöne ist: unser simples Prinzip erlaubt die Existenz mühelos nachzuweisen und dann kann man sich Methoden zur Berechnung ausdenken. Aber mit Rechnen allein ist nichts zu machen.

13.8 Monotonie

Eine Funktion $y = f(x)$ heisst *monoton* (auf dem betrachteten Definitiosbereich D) falls

$$\begin{array}{ll} f(x) \leq f(u) \text{ für alle } x \leq u & \textit{monoton wachsend} \\ f(x) \geq f(u) \text{ für alle } x \leq u & \textit{monoton fallend} \end{array}$$

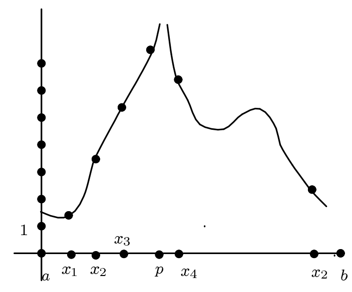
und *streng monoton* falls

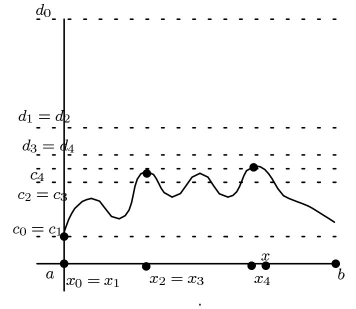
$$\begin{array}{ll} f(x) < f(u) \text{ für alle } x < u & \textit{streng monoton wachsend} \\ f(x) > f(u) \text{ für alle } x < u & \textit{streng monoton fallend} \end{array}$$

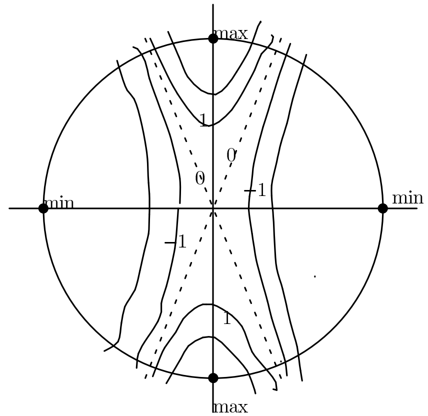
Beispiel: für $n = 1, 3, 5, \ldots$ ist $f(x) = x^n$ ist streng monoton wachsend auf $\mathbb{R}$, für $n = 2, 4, \ldots$ streng monoton fallend für $x \leq 0$ und wachsend für $x \geq 0$.

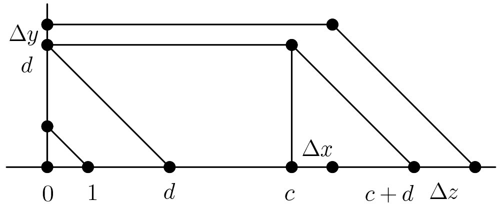
Satz 13.3 *Ist f auf $[a,b]$ monoton und nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an, so ist f auf $[a,b]$ stetig.*

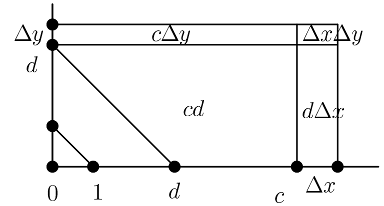
Beweis. Sei f z.B. monoton wachsend. Sei $f(a) < d = f(c) < f(b)$ angenommen und $\varepsilon > 0$ gegeben. Nach Voraussetzung gibt es ein $c_1 \in [a,b]$ mit $f(c_1) = \max\{a, d-\varepsilon\} < d$ und wegen der Monotonie gilt $c_1 < c$. Ebenso gibt es $c_2 > c$ mit $f(c_2) = \min\{b, d+\varepsilon\}$. Wähle nun $\delta = \min\{c - c_1, c_2 - c\}$. Ist $f(c) = f(a)$ oder $f(c) = f(b)$, so hat man nur eine Seite zu betrachten, auf der anderen Seite wirds trivial (die Stetigkeitsbedingung $|f(x) - f(c)| \leq \varepsilon$ bezieht sich nur auf x, für die $f(x)$ auch definiert ist. $\square$

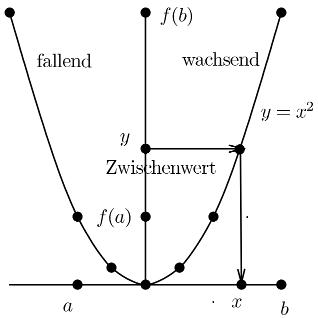


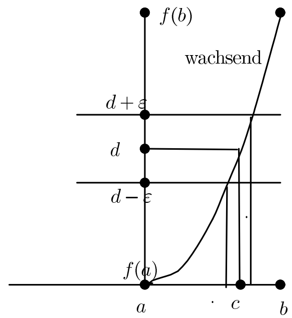


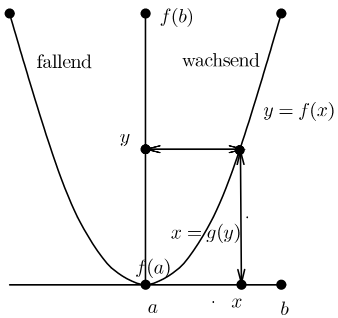


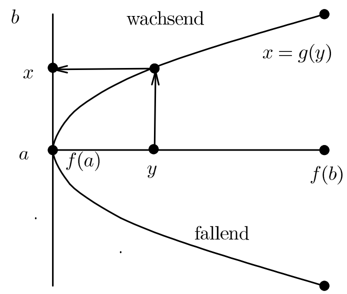












13.9 Umkehr

Die Funktionen $y = f(x)$ und $x = g(y)$ sind gegenseitig *Umkehrfunktionen*, wenn gilt

$$y = f(x) \quad \text{genau dann, wenn} \quad x = g(y)$$

für alle x im Definitionsbereich von f bzw. y im Definitionsbereich von g. Man kann auch so sagen

- Ist $f(x)$ definiert, so ist $g(f(x)) = x$
- Ist $g(y)$ definiert, so ist $f(g(y)) = y$

Also

$$f : D(f) \to B(f) \quad \text{und} \quad g : D(g) = B(f) \to B(g) = D(f)$$

Hat f eine Umkehrfunktion g, so ist diese eindeutig betsimmt und man schreibt auf $g = f^{-1}$.

Satz 13.4 *Eine auf einem Intervall definierte stetige Funktion hat genau dann eine Umkehrfunktion, wenn sie streng monoton ist. Die Umkehrfunktion ist dann ebenfalls auf einem Intervall definiert, stetig und im gleichen Sinne streng monoton.*

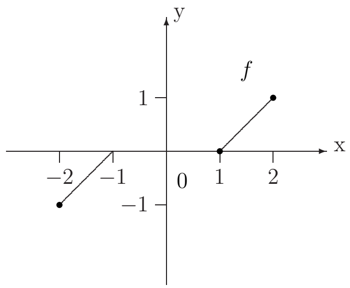
Beweis. Sei f stetig auf einem Intervall und $a < b$ in diesem. Sei g Umkehrfunktion von f. Wäre $f(a) = f(b)$, so $a = g(f(a)) = g(f(b)) = b$. Also sei z.B. $f(a) < f(b)$. Wir behaupten, dass dann $f(a) < f(x) < f(b)$ für alle x mit $a < x < b$. Andernfalls wäre z.B. $f(x) \leq f(a)$, aber dann auch, wie gerade gesehen, $f(x) < f(a)$. Nach dem Zwischenwertsatz gibt es aber dann ein y mit $x \leq y \leq b$ mit $f(y) = f(a)$, Widerspruch wie eben.

Sei umgekehrt f auf einem Intervall $[a, b]$ stetig und streng monoton, z.B. wachsend. Dann $f(a) \leq f(x) \leq f(b)$ für alle $x \in [a, b]$ und nach dem Zwischenwertsatz tritt jedes $y \in [f(a), f(b)]$ als Wert $y = f(x)$ auf, wegen der strengen Monotonie aber genau einmal. Also haben wir mit $g(y) = x \Leftrightarrow y = f(x)$ eine wohldefinierte Funktion auf $[f(a), f(b)]$, die f umkehrt. Diese ist ebenfalls streng monoton, hier wachsend: ist $y = g(x) < y = g(x_1)$, so kann nicht $x_1 \leq x$ sein. □ Beispiel

$$\begin{array}{lll} n = 1, 3, 5 \ldots : & y = x^n, \;\; x = \sqrt[n]{y} & \text{für } x, y \in \mathbb{R} \\ n = 2, 4, 6 \ldots : & y = x^n, \;\; x = +\sqrt[n]{y} & \text{für } x, y \geq 0 \\ n = 2, 4, 6 \ldots : & y = x^n, \;\; x = -\sqrt[n]{y} & \text{für } x \leq 0, \; y \geq 0 \end{array}$$

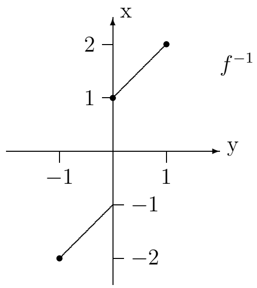
Hier ist es wichtig, dass f auf einem *Intervall* streng monoton wächst. Das zeigt folgendes Beispiel. Sei $X = [-2, -1) \cup [1, 2]$ und

$$f(x) = \begin{cases} x+1 & \text{falls } x \in [-2,-1) \\ x-1 & \text{falls } x \in [1,2] \end{cases}$$



Die Funktion f ist streng monoton wachsend, stetig, und sie bildet X auf $[-1,1]$ ab. Die Umkehrfunktion f^{-1} ist

$$f(y)^{-1} = \begin{cases} y-1 & \text{falls } y \in [-1,0), \\ y+1 & \text{falls } y \in [0,1]. \end{cases}$$



Diese ist an der Stelle 0 unstetig. ∎

13.10 Arcus

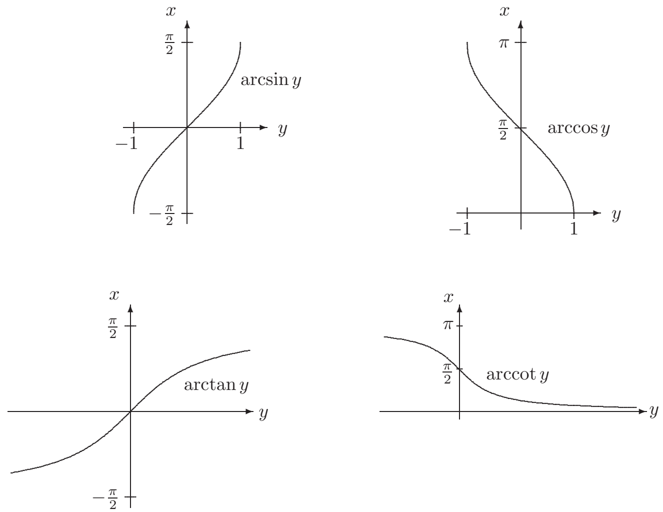
Wir schauen uns noch das Problem der Umkehrfunktion der trigonometrischen Funktionen an. Wegen der Periodizität ist keine der Funktionen $\sin, \cos, \tan, \cot$ auf ihrem gesamten Definitionsbereich injektiv, und es besitzt keine dieser Funktionen auf dem gesamten Definitionsbereich einer Umkehrfunktion. Man sucht sich daher für diese Funktionen (möglichst große) Intervalle, auf denen die Funktionen streng monoton (wachsend oder fallend) sind. Betrachtet man die trigonometrischen Funktionen auf diesen Intervallen, so werden sie injektiv und können umgekehrt werden. Diese Intervalle sind natürlich nicht eindeutig bestimmt (z.B. ist der Sinus auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend und auf $[\frac{\pi}{2}, \frac{3\pi}{2}]$ streng monoton fallend), und jede Wahl eines solchen Intervalls liefert eine andere Umkehrfunktion. Üblich und praktisch ist es, die folgenden Intervalle zu wählen:

$$\begin{array}{rrcl@{\qquad}rrcl}
 & \multicolumn{3}{c}{\textit{Funktion}} & & \multicolumn{3}{c}{\textit{Umkehrfunktion}} \\
\sin: & [-\frac{\pi}{2}, \frac{\pi}{2}] & \to & [-1,1] & \arcsin: & [-1,1] & \to & [-\frac{\pi}{2}, \frac{\pi}{2}], \\
\cos: & [0,\pi] & \to & [-1,1] & \arccos: & [-1,1] & \to & [0,\pi], \\
\tan: & (-\frac{\pi}{2}, \frac{\pi}{2}) & \to & \mathbb{R} & \arctan: & \mathbb{R} & \to & (-\frac{\pi}{2}, \frac{\pi}{2}), \\
\cot: & (0,\pi) & \to & \mathbb{R} & \operatorname{arccot}: & \mathbb{R} & \to & (0,\pi)\,.
\end{array}$$

Die so definierten Umkehrfunktionen heißen *Arkusfunktionen* (lies z.B. *Arkussinus* für arcsin). Da die Funktionen

$$\begin{array}{ll@{\qquad}ll}
[-\frac{\pi}{2}, \frac{\pi}{2}] \to \mathbb{R}, & x \mapsto \sin x & [0,\pi] \to \mathbb{R}, & x \mapsto \cos x \\
(-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}, & x \mapsto \tan x & (0,\pi) \to \mathbb{R}, & x \mapsto \cot x
\end{array}$$

auf den angegebenen Intervallen streng monoton sind, sind ihre Umkehrfunktionen $x \mapsto \arcsin y, \arccos y, \arctan y, \operatorname{arccot} y$ stetig.



Elementargeometrisch sieht das Ganze für sin und cos so aus: Nach Archimedes können wir jeden Strahl g aus dem Zentrum O des Einheitskreises durch eine fortlaufenden Halbierung $\underline{g}_n, \overline{g}_n$ einschachteln (vgl.Skizze) - wir betrachten nur den ersten Quadranten und fangen mit $n = 2$ an. Der Winkel zwischen $\underline{g}_2$ und $\underline{g}_n$ bzw. $\overline{g}_n$ sei

$$\frac{k_n}{2^n} \quad \text{bzw.} \quad \frac{k_n+1}{2^n}$$

mal der Vollwinkel. Für die Schnittpunkte mit dem Kreis gilt dann die Reihenfolge $\underline{P}_n, \underline{P}_{n+1}, P, \overline{P}_{n+1}, \overline{P}_n$ und für die Länge ϕ des Bogens von E nach P

$$\frac{k_n}{2^n}\underline{p}_n \le \phi \le \frac{k_n+1}{2^n}\overline{p}_n$$

Richtig ist: wir definieren ϕ durch diese Intervallschachtelung. Den Punkt P können wir durch seine y-Koordinate y angeben, und haben somit eine Funktion definiert

$$\phi = \arcsin y, \quad 0 \le y \le 1$$

Aus der Geometrie ist sofort klar, dass diese Funktion streng monoton wachsend ist und keinen Zwischenwert auslässt; also ist sie stetig und hat eine streng monotone stetige Umkehrfunktion

$$y = \sin\phi, \quad 0 \le \phi \le \frac{\pi}{2}$$

Nehmen wir für Punkte in vierten Quadranten die Bogenlänge mit negativem Vorzeichen, so kǒnnen wir das erweitern zu dem streng monoton wachsenden und stetigen Umkehr-Funktionen-Paar

$$\phi = \arcsin y, \quad y = \sin\phi, \quad -1 \le y \le 1, \ -\frac{\pi}{2} \le \phi \le \frac{\pi}{2}$$

Wenn wir P durch seine Koordinate x auf der Achse durch OE notieren, so erhalten wir entsprechend das streng monoton fallende und stetige Umkehr-Funktionen-Paar

$$\phi = \arccos x, \quad x = \cos\phi, \quad -1 \le x \le 1, \ 0 \le \phi \le \pi$$

13 Funktionen und Stetigkeit: Limites

13.15 Funktionenlimes

Manchmal möchte man das Verhalten von f an einer Stelle c studieren, wo f garnicht oder falsch definiert ist. Man macht's meist aber nur dann, wenn f in der Nähe von c, c ausgenommen, definiert ist, d.h. wenn es ein $\eta > 0$ gibt so, dass $f(x)$ für alle $x \ne c$ mit $|x-c| \le \eta$ definiert ist. Mindestens soll aber c ein Häufungspunkt des Definitionsbreichs von f sein.

Wir sagen: $y = f(x)$*geht gegen d für x gegen c* und schreiben

$$y = f(x) \to d \ \text{ für } x \to c \qquad \text{auch } d = \lim_{x\to c} f(x)$$

falls c ein Häufungspunkt von $D(f)$ ist und die Funktion

$$\overline{f}(x) = \begin{cases} f(x) & \text{für } x \ne c \\ d & \text{für } x = c \end{cases}$$

an der Stelle c stetig ist.

Dann bedeuten die folgenden Aussagen und Schreibweisen dasselbe

- $d = \lim_{x\to c} f(x)$
- Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ so, dass für alle $x \ne c$ mit $x \in D(f)$ und $|x-c| \le \delta$ gilt: $|f(x) - d| \le \varepsilon$
- Die Funktion g mit $g(c) = d$ und $g(x) = f(x)$ für $x \ne c$ in $D(f)$ ist an der Stelle c stetig

- Für jede Folge $x_n \ne c$ mit $x_n \in D(f)$ und $x_n \to c$ gilt $f(x_n) \to d$
- $f(x) \to d$ für $x \to c$

Lemma 13.1 *Sind g und h stetig an der Stelle c mit $g(c) = h(c) = d$ und gilt $g(x) \le f(x) \le h(x)$ für alle x mit $0 < |x-c| \le \gamma$, so $f(x) \to d$ für $x \to c$.*

Beweis. Zu $\varepsilon > 0$ wähle $0 < \delta \le \gamma$ so, dass $|g(x) - d| \le \varepsilon$ und $|h(x) - d| \le \varepsilon$ für alle x mit $|x - c| \le \delta$. Dann

$$|f(x) - d| \le \max\{|g(x) - d|, |h(x) - d|\} \le \varepsilon \ \text{ für alle } x \text{ mit } |x-c| \le \delta$$

Beispiel

$$y = \frac{\sin x}{x} \to 1 \ \text{ für } x \to 0$$

Aus der Geometrie folgt nämlich

$$x \le \frac{\sin x}{\cos x} \quad \text{und} \quad \frac{\sin x}{x} \le 1$$

also, da $\cos x > 0$ für $x \to 0$ und $x \sin x \ge 0$

$$\cos x \le \frac{\sin x}{x} \le 1$$

Lemma 13.2 *Ist $g(y_1, y_2)$ stetig und $y_1 = f_1(x) \to d_1$ und $y_2 = f_2(x) \to d_2$ für $x \to c$, so folgt $z = g(f_1(x), f_2(x)) \to g(d_1, d_2)$ für $x \to c$.*

Beweis: Stetigkeit der Substitution. □

Korollar 13.3 *Die Funktionen $f, g : X \to \mathbb{R}$ sollen an der Stelle x_0 einen Grenzwert besitzen. Dann besitzen auch die Funktionen $f+g$, fg und cf mit $c \in \mathbb{R}$ sowie – falls $g(x) \ne 0$ für $x \in X$ und falls $\lim_{x\to x_0} g(x) \ne 0$ – die Funktion f/g einen Grenzwert in x_0, und es gilt*

$$\begin{aligned}
\lim_{x\to x_0} (f(x) + g(x)) &= \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x), \\
\lim_{x\to x_0} (f(x)g(x)) &= \lim_{x\to x_0} f(x) \cdot \lim_{x\to x_0} g(x), \\
\lim_{x\to x_0} (c\, f(x)) &= c \lim_{x\to x_0} f(x), \\
\lim_{x\to x_0} (f(x)/g(x)) &= \lim_{x\to x_0} f(x) / \lim_{x\to x_0} g(x).
\end{aligned}$$

Beispiel

$$\frac{1-\cos x}{x} = \frac{1-\cos^2 x}{x(1+\cos x)} = \frac{\sin x}{x} \cdot \frac{\sin x}{1+\cos x} \to 1 \cdot 0 = 0 \text{ für } x \to 0$$

Beispiel 6 Sei $f : \mathbb{R}\backslash\{2\} \to \mathbb{R}$ gegeben durch $f(x) = \frac{x^2+x-6}{x-2}$. Dann ist $\lim_{x\to 4} \frac{x^2+x-6}{x-2} = \frac{\lim_{x\to 4}(x^2+x-6)}{\lim_{x\to 4}(x-2)} = \frac{14}{2} = 7$, und $\lim_{x\to 2} \frac{x^2+x-6}{x-2} = \lim_{x\to 2} \frac{(x-2)(x+3)}{x-2} = \lim_{x\to 2}(x+3) = 5$. ■

Sei wieder X Teilmenge von $\mathbb{R}$. Ein Punkt $x \in X$ heißt *isoliert*, wenn er kein Häufungspunkt von X ist. Z.B. ist jeder Punkt von $\mathbb{N}$ (also jede natürliche Zahl) isoliert, wenn wir $\mathbb{N}$ als Teilmenge von $\mathbb{R}$ betrachten.
Mit diesen Begriffen können wir eine weitere äquivalente Charakterisierung der Stetigkeit geben:

Korollar 13.4 *Eine Funktion $f : X \to \mathbb{R}$ ist genau dann stetig in $x_0 \in X$, wenn eine der folgenden Bedingungen erfüllt ist:*

(a) *x_0 ist ein isolierter Punkt von X*
(b) *x_0 ist Häufungspunkt von X, der Grenzwert $\lim_{x\to x_0} f(x)$ existiert, und dieser Grenzwert ist gleich $f(x_0)$.*

Das folgt sofort aus den Definitionen. Man beachte, dass für einen isolierten Punkt $x_0 \in X$ eine Folge (x_n) aus X genau dann gegen x_0 konvergiert, wenn $x_n = x_0$ für alle hinreichend großen n.

13.16 Einseitige Grenzwerte

Beispiel 5 Die Signumfunktion aus Beispiel 2 besitzt keinen Grenzwert im Punkt $x_0 = 0$, denn für die Folgen $(\frac{1}{n})$ und $(\frac{-1}{n})$ mit Grenzwert 0 gilt:

$$\lim_{n\to\infty} \mathrm{sgn}\frac{1}{n} = 1, \text{ aber } \lim_{n\to\infty} \mathrm{sgn}(\frac{-1}{n}) = -1.$$

■

In diesem Beispiel hat man aber folgenden Effekt: Nähert sich eine Folge (x_n) von oben der Null (ist also $x_n > 0$ für alle n), so existiert $\lim_{n\to\infty} f(x_n)$ und ist gleich 1. Man sagt auch, dass 1 der *rechtsseitige Grenzwert* von f ist. Analog dazu ist -1 der *linksseitige Grenzwert* von $f = \mathrm{sgn}$. Allgemein definiert man

Definition 13.5 *Sei x_0 ein Häufungspunkt von $X \cap (x_0, \infty)$. Man sagt, dass der* rechtsseitige Grenzwert *von $f : X\backslash\{x_0\} \to \mathbb{R}$ an der Stelle x_0 existiert und gleich $y_0 \in \mathbb{R}$ ist, wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$ aus $X \cap (x_0, \infty)$ für alle n der Grenzwert $\lim_{n\to\infty} f(x_n)$ existiert und gleich y_0 ist. Man schreibt dann auch*

$$y_0 = \lim_{\substack{x\to x_0\\ x>x_0}} f(x) \quad \text{oder} \quad y_0 = \lim_{x\searrow x_0} f(x) \quad \text{oder} \quad y_0 = f(x_0 + 0).$$

Ersetzt man jedes $>$ durch $<$, erhält man den Begriff des linksseitigen Grenzwertes. Es ist also z.B. $\lim_{x\nearrow 0} \mathrm{sgn} x = -1$.

Korollar 13.6 $y_0 = \lim_{x\to x_0} f(x) \Leftrightarrow y_0 = \lim_{x\searrow x_0} f(x)$ *und* $y_0 = \lim_{x\nearrow 0} f(x)$.

13.18 Uneigentliche Grenzwerte

Sei $f : D(F) \to \mathbb{R}$ eine Funktion und p ein Häufungspunkt von $D(f)$.

- Gibt es zu jedem $c \in \mathbb{R}$ ein $\delta > 0$ so, dass $f(x) \geq c$ für alle $x \in D(f)$ mit $|p - x| < \delta$, so hat f an der Stelle p den *uneigentlichen Grenzwert* $+\infty$ und man schreibt $\lim_{x\to p} f(x) = +\infty$ oder $y = f(x) \to +\infty$ für $x \to p$.
- Gibt es zu jedem $c \in \mathbb{R}$ ein $\delta > 0$ so, dass $f(x) \leq c$ für alle $x \in D(f)$ mit $|p - x| < \delta$, so hat f an der Stelle p den *uneigentlichen Grenzwert* $-\infty$ und man schreibt $\lim_{x\to p} f(x) = -\infty$ oder $y = f(x) \to -\infty$ für $x \to p$.

- analog für einseitige Grenzwerte.
- Sei $D(f)$ nicht nach oben beschränkt. Ist $a \in \mathbb{R}$ und gibt es zu jedem $\varepsilon > 0$ ein c so, dass $|a - f(x)| < \varepsilon$ für alle $x \in D(f)$, $x \geq c$, so ist a der *Grenzwert* von f für $x \to \infty$ und man schreibt $\lim_{x\to+\infty} f(x) = a$ oder $y = f(x) \to a$ für $x \to +\infty$.
- Sei $D(f)$ nicht nach unten beschränkt. Ist $a \in \mathbb{R}$ und gibt es zu jedem $\varepsilon > 0$ ein c so, dass $|a - f(x)| < \varepsilon$ für alle $x \in D(f)$, $x \leq c$, so ist a der *Grenzwert* von f für $x \to \infty$ und man schreibt $\lim_{x\to-\infty} f(x) = a$ oder $y = f(x) \to a$ für $x \to -\infty$.
- Sei $D(f)$ nicht nach oben beschränkt. Gibt es zu jedem $a \in \mathbb{R}$ ein c so, dass $f(x) \geq a$ für alle $x \in D(f)$, $x \geq c$, so ist $+\infty$ der *uneigentliche Grenzwert* von f für $x \to +\infty$ und man schreibt $\lim_{x\to+\infty} f(x) = +\infty$ oder $y = f(x) \to +\infty$ für $x \to +\infty$.
- Sei $D(f)$ nicht nach oben beschränkt. Gibt es zu jedem $a \in \mathbb{R}$ ein c so, dass $f(x) \leq a$ für alle $x \in D(f)$, $x \geq c$, so ist $-\infty$ der *uneigentliche Grenzwert* von f für $x \to +\infty$ und man schreibt $\lim_{x\to+\infty} f(x) = -\infty$ oder $y = f(x) \to -\infty$ für $x \to +\infty$.
- Sei $D(f)$ nicht nach unten beschränkt. Gibt es zu jedem $a \in \mathbb{R}$ ein c so, dass $f(x) \geq a$ für alle $x \in D(f)$, $x \leq c$, so ist $+\infty$ der *uneigentliche Grenzwert* von f für $x \to -\infty$ und man schreibt $\lim_{x\to-\infty} f(x) = +\infty$ oder $y = f(x) \to +\infty$ für $x \to +\infty$.
- Sei $D(f)$ nicht nach unten beschränkt. Gibt es zu jedem $a \in \mathbb{R}$ ein c so, dass $f(x) \leq a$ für alle $x \in D(f)$, $x \leq c$, so ist $-\infty$ der *uneigentliche Grenzwert* von f für $x \to -\infty$ und man schreibt $\lim_{x\to-\infty} f(x) = -\infty$ oder $y = f(x) \to -\infty$ für $x \to -\infty$.

13 Funktionen und Stetigkeit: Vektorfunktionen

13.17 Vektorfunktionen

Sei V ein endlichdimensionaler euklidischer Vektorraum, z.B. $V = \mathbb{C}$ und $D \subseteq \mathbb{R}$. Funktionen $f : D \to V$ heissen *Vektorfuntkionen* und werden of so geschrieben

$$f(t) = \vec{x}(t) \in V, \quad (t\ inD)$$

d.h. der Vektor $\vec{x}(t)$ Funktion der Zeit t. Bei der Definiton der Stetigkeit zur Zeit t_0 hat man dann den Abstand in V statt des Absolutbetrags, ansonsten wörtlich dassselbe, d.h. die äquivalenten Bedingungen

- $\Delta\vec{x} \to \vec{0}$ für $\Delta t \to 0$
- $\forall\varepsilon > 0\ \exists\delta > 0\ \forall t \in D.\ |t_0 - t| < \delta \Rightarrow \|\vec{x}(t_0) - \vec{x}(t)\| < \varepsilon$
- Für alle Folgen $t_n \to t_0$ gilt $\vec{x}(t_n) \to \vec{x}(t_0)$

Korollar 13.1 *Ist $\vec{v}_1, \ldots, \vec{v}_m$ eine Basis von V, so gibt es eindeutig bestimmte Funktionen $t \mapsto x_i(t) \in \mathbb{R}$, $(i = 1, \ldots, m)$ mit*

$$\vec{x}(t) = x_1(t)\vec{v}_1 + \ldots + x_m(t)\vec{v}_m$$

und die Vektorfunktion $t \mapsto \vec{x}(t)$ ist genau dann stetig an t_0, wenn es die Funktionen $x_i(t)$ sind.

Das folgt sofort mit der Charakterisierung über die Konvergenz von Folgen und Satz 12.11. Beispiel mit $V = \mathbb{C}$

$$\vec{x}(t) = r(t)(\cos t)\vec{e}_1 + r(t)(\sin t)\vec{e}_2 = r(t)e^{jt}$$

Ist $h : V \to V$ eine affine Abbildung und $f : D \to V$ an t_0 stetig, so ist $h \circ f : D \to V$ stetig an t_0, z.B.

$$t \mapsto \vec{y}(t) = A\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} v_1 \\ v_1 \end{pmatrix}$$

mit stetigen $x_1(t)$ und $x_2(t)$. Nämlich $\vec{y}(t) = y_1(t)\vec{e}_1 + y_2(t)\vec{e}_2$ mit $y_i(t) = a_{i1}x_1(t) + a_{i2}x_2(t) + v_i$

Seien $t \mapsto \vec{x}(t)$ und $t \mapsto \vec{y}(t)$ an t_0 stetige Abbildungen von D in V und $r \in \mathbb{R}$. Dann sind auch die folgenden Abbildungen stetig an t_0

$$t \mapsto (\vec{x} + \vec{y})(t) = \vec{x}(t) + \vec{y}(t), \quad t \mapsto (r\vec{x})(t) = r\vec{x}(t), \quad t \mapsto \langle \vec{x}(t) \mid \vec{y}(t) \rangle$$

wie man sofort mithilfe der $x_i(t)$ und $y_i(t)$ sieht.

Sind $t \mapsto z(t)$ und $w \mapsto w(t)$ an t_0 stetige Abbildungen von D in $\mathbb{C}$ und $h : \mathbb{C} \to \mathbb{C}$ gebrochen linear, so sind auch die folgenden Abbildungen stetig an t_0 (soweit definiert)

$$t \mapsto z(t) \cdot w(t), \; t \mapsto \frac{1}{z(t)}, \quad t \mapsto h(z(t))$$

Zum Beweis benutzen wir die Polardarstellung und bemerken, dass nach der Dreiecksungleichung und dem Umstand, dass die Sehne kürzer als der Bogen ist, gilt

$$|re^{j\phi} - se^{j\psi}| \leq |r - s| + \max\{r, s\} \cdot |\phi - \psi|$$

Für

$$z(t) = r(t)(\cos\phi(t) + j\sin\phi(t)), \quad w(t) = s(t)(\cos\psi(t) + j\sin\psi(t))$$

haben wir also

$$|\Delta(zw)| \leq |\Delta r| + |\Delta s| + \max\{|\Delta s| \cdot |r(t_0)|, |\Delta r| \cdot |s(t_0)|\} \cdot (|\Delta\phi| + |\Delta\psi|)$$

$$|\Delta z^{-1}| \leq \frac{2}{r(t_0)^2}|\Delta r| + (\frac{1}{r(t_0)} + |\Delta r|)|\Delta\phi|$$

Ist $z \mapsto h(z)$ gebrochen linear, so ist es Komposition von affinen Abbildungen und Inversion, also $t \mapsto h(z(t)$ ebenfalls stetig.

14.11 Limes von Vektorfunktionen

Sei V n-dimensionaler euklidischer Vektorraum, z.B. $V = \mathbb{C}$ und

$$f : D(f) \to V, \quad \vec{y} = f(t)$$

und p ein Häufungspunkt von $D(f)$. Dann ist $\vec{a}$ Limes von f für $t \to p$

$$\vec{a} = \lim_{t \to p} f(t), \quad \vec{y} = f(t) \to \vec{a} \text{ für } t \to p$$

falls es zu jedem $\varepsilon > 0$ ein $\delta >$ gibt so, dass für alle t mit $|t - p| < \delta$ gilt: $\|f(t) - f(p)\| < \varepsilon$.

Satz 14.2 *Ist eine Orthonormalbasis $\vec{e}_1, \ldots \vec{e}_n$ gegeben und $f(t) = f_1(t)\vec{e}_1 + \ldots + f_n(t)\vec{e}_n$, so*

$$\vec{y} = f(t) \to \vec{a} \;\Leftrightarrow\; y_i = f_i(t) \to a_i \; (i = 1, \ldots, n)$$

Beweis. Das kann man mit dem entsprechenden Satz über Folgen beweisen. Oder direkt so: Gelte $y_i \to a_i$ für $i = 1, \ldots, n$. Ist ein $\varepsilon > 0$ gegeben, so gibt es δ_i so, dass $|f_i(t) - f_i(p)| < \frac{\varepsilon}{n}$ falls $|t - p| < \delta_i$. Wähle δ als Minimum der δ_i. Dann

$$\|f(t) - f(p)\| = \sqrt{\sum_{i=1}^{n}(f_i(t) - f_i(p))^2} < \varepsilon \quad \text{falls } |t - p| < \delta$$

In der umgekehrten Richtung kann man benutzen, dass man $f(t)$ an der Stelle p durch $f(p) = \vec{a}$ stetig ergänzen kann und dann $f_i(t) = \langle \vec{e}_i \mid f(t) \rangle$ durch $f_i(p) = a_i$ stetig ergaänzt wird, weil das Skalarprodukt stetig in jedem Argument ist. Zu Fuß: Zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle t mit $|t - p| < \delta$ gilt: $\|f(t) - f(p)\| < \varepsilon$. Dann aber auch $|f_i(t) - f_i(p)| < \varepsilon$, da $|f_i(t) - f_i(p)| \leq \|f(t) - f(p)\|$. $\square$

14 Differentiation: Grundlagen und Beispiele

14.1 Differentialquotient

Wir betrachten im Folgenden eine Funktion $y = f(x)$. $f : D(f) \to \mathbb{R}$, an einer Stelle p und setzen voraus, dass $f(x)$ für ein Intervall $|x - p| < \gamma$ definiert ist. Wir setzen

$$\Delta x = x - p, \quad \Delta y = f(x) - f(p)$$

Dabei können wir Δx als eine unabhängige Variable für die Änderung von x auffassen, und haben dann die Änderung

$$\Delta y = (\Delta y)(\Delta x, p) = f(p + \Delta x) - f(p)$$

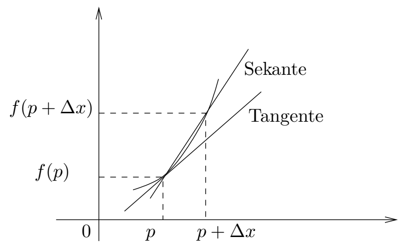
von y als Funktion von Δx bei fester Stelle p. Die Stetigkeit von $y = f(x)$ an der Stelle p bedeutet dann

$$\Delta y \to 0 \text{ für } \Delta x \to 0.$$

Der Quotient

$$\frac{\Delta y}{\Delta x} = \frac{f(p + \Delta x) - f(p)}{\Delta x} \quad \Delta x \neq 0$$

ist ein *Differenzenquotient* und gibt die Steigung der Sekante durch die Punkte $(p, f(p))$ und $(p + \Delta x,\; f(p + \Delta x)$ an.



Wir betrachten die Funktion

$$\frac{\Delta y}{\Delta x} \quad \text{für } \Delta x \to 0,\ \Delta x \neq 0$$

Existiert der Limes, so heisst f an der Stelle p *differenzierbar* und wir haben die *Ableitung* a an der Stelle p mit

$$(1)\quad \frac{\Delta y}{\Delta x} \to a = f'(p) = \frac{\partial y}{\partial x} \quad \text{für } \Delta x \to 0,\ \Delta x \neq 0$$

Beispiele

$$y = x^2,\quad \frac{\Delta y}{\Delta x} = \frac{2p\Delta x + (\Delta x)^2}{\Delta x} = 2p + \Delta x \to 2p \text{ für } \Delta x \to 0$$

$$y = \frac{1}{x},\quad \frac{\partial \frac{1}{x}}{\partial x} = -\frac{1}{x^2}$$

$$\Delta\frac{1}{x} = \frac{1}{p+\Delta x} - \frac{1}{p} = \frac{-\Delta x}{p(p+\Delta x)},\quad \frac{\Delta\frac{1}{x}}{\Delta x} = \frac{-1}{p(p+\Delta x)} \to \frac{-1}{p^2}$$

$$\frac{\partial \sin x}{\partial x} = \cos x,\quad \Delta y = \sin(p+\Delta x) - \sin p = \sin p \cos \Delta x + \cos p \sin \Delta x - \sin p$$

$$\frac{\Delta y}{\Delta x} = -\sin p \frac{1-\cos \Delta x}{\Delta x} + \cos p \frac{\sin \Delta x}{\Delta x} \to -\sin p \cdot 0 + \cos p \cdot 1 = \cos p$$

Satz 14.1 *Ist $f(x)$ an der Stelle p differenzierbar, so auch stetig:*

$$\Delta y = \frac{\Delta y}{\Delta x} \cdot \Delta x \;\to\; \frac{\partial y}{\partial x} \cdot 0 = 0 \;\; \textit{für } \Delta x \to 0,\ \Delta x \neq 0$$

$$(*)\quad \frac{\partial y}{\partial x} \neq 0 \;\Rightarrow\; \Delta y \neq 0 \;\textit{für } \Delta x \to 0,\ \Delta x \neq 0$$

Haben wir nämlich $\frac{\Delta y}{\Delta x} \to a \neq 0$ und z.B. $a > 0$, so $\frac{\Delta y}{\Delta x} \geq \frac{a}{2} > 0$ für hinreichend kleine Δx, also $\Delta y \neq 0$. □

Zur Differenzierbarkeit äquivalente Bedingungen sind

$$(2)\quad \Delta y = \frac{\partial y}{\partial x}\Delta x + \phi(\Delta x)\Delta x \quad \text{mit } \phi(\Delta x) \to 0 \text{ für } \Delta x \to 0,\ \Delta x \neq 0$$

In der Tat, wähle

$$\phi(\Delta x) = \frac{\Delta y}{\Delta x} - \frac{\partial y}{\partial x}$$

$$(3)\quad \Delta y = \frac{\partial y}{\partial x}\Delta x + R(\Delta x) \quad \text{mit } \frac{R(\Delta x)}{\Delta x} \to 0 \text{ für } \Delta x \to 0,\ \Delta x \neq 0$$

In der Tat, wähle nach (2)

$$R(\Delta x) = \phi(\Delta x) \cdot \Delta x$$

(3) besagt, dass man $f(x)$ in der Nähe von p durch eine affin lineare Funktion approximieren kann, mit einem Fehler, der "klein von zweiter Ordnung" ist

$$(4)\quad f(p+\Delta x) = f(p) + \Delta y = f(p) + \frac{\partial y}{\partial x}\Delta x + R(\Delta x)$$

Und aus (3) folgt die Differenzierbarkeit so

$$\frac{\Delta y}{\Delta x} = \frac{\partial y}{\partial x} + \frac{R(\Delta x)}{\Delta x} \to \frac{\partial y}{\partial x} \quad \text{für } \Delta x \to 0,\ \Delta x \neq 0$$

Aus (3) folgt sofort, dass eine konstante Funktion überall Ableitung 0 hat.

14.2 Arithmetische Differentiationsregeln

$$\frac{\partial Cy}{\partial x} = C\frac{\partial y}{\partial x}$$

Beweis: $\Delta Cy = C\Delta y$

$$\text{Summenregel}\quad \frac{\partial(y+z)}{\partial x} = \frac{\partial y}{\partial x} + \frac{\partial z}{\partial x}$$

Beweis: $\Delta(y+z) = \Delta y + \Delta z$

$$\text{Produktregel}\quad \frac{\partial(yz)}{\partial x} = y\frac{\partial z}{\partial x} + z\frac{\partial y}{\partial x}$$

Beweis: $\Delta(yz) = y\Delta z + z\Delta y + \Delta y \Delta z$

$$\frac{\Delta(yz)}{\Delta x} = y\frac{\Delta z}{\Delta x} + z\frac{\Delta y}{\Delta x} + \Delta y\frac{\Delta z}{\Delta x} \to y\frac{\partial z}{\partial x} + z\frac{\partial y}{\partial x} + 0\frac{\partial z}{\partial x}$$

Anwendung

$$\frac{\partial \sum_{k=0}^{n} a_n x^n}{\partial x} = \sum_{k=1}^{n} k a_k x^{k-1}$$

Beweis durch Summenregel und Induktion mithilfe der Produktregel

$$\frac{\partial x^k}{\partial x} = \frac{\partial x^{k-1}x}{\partial x} = x\frac{\partial x^{k-1}}{\partial x} + x^{k-1}\frac{\partial x}{\partial x} = x(k-1)x^{k-2} + x^{k-1} \cdot 1 = kx^{k-1}$$

$$\text{Quotientenregel}\quad \frac{\partial \frac{1}{x}}{\partial x} - \frac{1}{x^2}$$

Beweis.

$$\frac{\Delta y}{\Delta x} = \left(\frac{1}{x+\Delta x} - \frac{1}{x}\right)(\Delta x)^{-1} = \frac{x-(x+\Delta x)}{x^2+x\Delta x}(\Delta x)^{-1} = \frac{-1}{x^2+x\Delta x} \to \frac{-1}{x^2} \text{ für } \Delta x \to 0$$

14.3 Ketten- und Umkehrregel

Satz 14.2 Kettenregel *Sind $y = f(x)$ an der Stelle p und $z = g(y)$ an der Stelle $q = f(p)$ differenzierbar, so ist $z = g(f(x))$ an der Stelle p differenzierbar und es gilt*

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x} \qquad \textit{genauer } (g \circ f)'(p) = g'(f(p)) \cdot f'(p)$$

Beweis: Ist $\frac{\partial y}{\partial x} \neq 0$, so gilt für $\Delta x \to 0,\ \Delta x \neq 0$ wegen Stetigkeit und (*) $\Delta y \to 0,\ \Delta y \neq 0$ und somit

$$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y} \cdot \frac{\Delta y}{\Delta x} \to \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x}$$

Im allgemeinen hat man mit (2)

$$\Delta z = \left(\frac{\partial z}{\partial y} + \psi(\Delta y)\right)\Delta y = \left(\frac{\partial z}{\partial y} + \psi(\Delta y)\right)\left(\frac{\partial y}{\partial x} + \phi(\Delta x)\right)\Delta x$$

$$= \frac{\partial z}{\partial y}\frac{\partial y}{\partial x}\Delta x + \chi(\Delta x)\Delta x$$

$$\chi(\Delta x) = \frac{\partial z}{\partial y}\phi(\Delta x) + \frac{\partial y}{\partial x}\psi(\Delta y) + \psi(\Delta y)\phi(\Delta x) \to 0 \text{ für } \Delta x \to 0$$

also mit (2) die Behauptung. □

Anwendung

$$\frac{\partial \cos x}{\partial x} = \frac{\partial \sin(x+\frac{\pi}{2})}{\partial x} = \cos(x+\frac{\pi}{2}) \cdot \frac{\partial(x+\frac{\pi}{2})}{\partial x} = -\sin x \cdot 1$$

Korollar 14.3 Quotientenregel

$$\frac{\partial \frac{1}{y}}{\partial x} = -\frac{1}{y^2} \cdot \frac{\partial y}{\partial x}$$

$$\frac{\partial x^{-n}}{\partial x} = -\frac{1}{x^{2n}} n x^{n-1} = -n x^{-n-1}$$

Mit der Produktregel folgt auch

$$\frac{\partial \frac{z}{y}}{\partial x} = \frac{1}{y^2} \cdot (y\frac{\partial z}{\partial x} - z\frac{\partial y}{\partial x})$$

Satz 14.4 Umkehrregel *Ist $y = f(x)$ in der Nähe von p stetig und umkehrbar und an der Stelle p differenzierbar mit Ableitung $f'(p) = \frac{\partial y}{\partial x} \neq 0$, so ist die Umkehrfuktion $x = g(y)$ an der Stelle $q = f(p)$ differenzierbar und es gilt*

$$g'(q) = \frac{\partial x}{\partial y} = \frac{1}{\frac{\partial y}{\partial x}} = \frac{1}{f'(p)} = \frac{1}{f'(g(q))}$$

Beweis: $x = g(y)$ ist in der Nähe von q stetig und streng monoton. Also haben wir für $\Delta y \to 0,\ \Delta y \neq 0$ auch $\Delta x \to 0,\ \Delta x \neq 0$ und somit

$$\frac{\Delta x}{\Delta y} = \frac{1}{\frac{\Delta y}{\Delta x}} \to \frac{1}{\frac{\partial y}{\partial x}}$$

Anwendung

$$y = x^2,\ x = \sqrt{y},\quad \frac{\partial \sqrt{y}}{\partial y} = \frac{1}{\frac{\partial x^2}{\partial x}} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}$$

14.4 Trigonometrische Funktionen

Wir wissen schon

$$\frac{\partial \sin x}{\partial x} = \cos x,\quad \frac{\partial \cos x}{\partial x} - \sin x$$

Insbesondere erfüllen beide Funktionen $y = f(x)$ die Differentialgleichung

$$f'' = \frac{\partial \frac{\partial y}{\partial x}}{\partial x} = \frac{\partial^2 y}{(\partial x)^2} = -y$$

und das gilt ebenso für alle Linearkombinationen $f(x) = a\sin x + b\sin y$. Später werden Sie sehen, dass dies genau die Lösungen dieser Differentialgleichungen sind.

Die Quotientenregel liefert

$$\frac{\partial \tan x}{\partial x} = \frac{1}{(\cos x)^2},\quad \frac{\partial \cot x}{\partial x} = \frac{-1}{(\sin x)^2}$$

Für die Umkehrfunktionen erhalten wir

$$y = \sin x\ (-\frac{\pi}{2} \leq x \leq \frac{\pi}{2})\quad x = \arcsin y\ \ y \in (-1,1)$$

$$\frac{\partial y}{\partial x} = \cos x = \sqrt{1-\sin^2 x} = \sqrt{1-y^2},\quad \frac{\partial \arcsin y}{\partial y} = \frac{1}{\sqrt{1-y^2}}$$

$$y = \cos x\ (0 \leq x \leq \pi)\quad x = \arccos y\ \ y \in (-1,1)$$

$$\frac{\partial y}{\partial x} = -\sin x = -\sqrt{1-\sin^2 x} = -\sqrt{1-y^2},\quad \frac{\partial \arccos y}{\partial y} = \frac{-1}{\sqrt{1-y^2}}$$

$$y = \tan x\ (-\frac{\pi}{2} < x < \frac{\pi}{2}),\quad x = \arctan y$$

$$\frac{\partial y}{\partial x} = \frac{1}{(\cos x)^2} = \frac{(\sin x)^2 + (\cos x)^2}{(\cos x)^2} = 1 + (\tan x)^2 = 1 + y^2,\quad \frac{\partial \arctan y}{\partial y} = \frac{1}{1+y^2}$$

$$\frac{\partial\, \text{arccoty}}{\partial y} = \frac{1}{1+y^2}$$

14 Differentiation: Anwendungen

14.5 Relative Extrema

Die Funktion $y = f(x)$ sei in der Nähe von p definiert: d.h. für $|x-p| < \gamma$ mit einem $\gamma > 0$. Wir sagen: $f(x)$ hat ein *relatives Maximum* (bzw. *Minimum*) an der Stelle p, wenn es $0 < \delta \leq \gamma$ gibt so, dass $f(p) \geq f(x)$ (bzw. $f(p) \leq f(x)$ für alle x mit $|x-p| < \delta$. Zusammengefasst: *relatives* oder em lokales Extremum.

Satz 14.1 Notwendiges Kriterium. *Hat $y = f(x)$ ein relatives Extremum an der Stelle p und ist dort differenzierbar, so gilt $f'(p) = \frac{\partial y}{\partial x} = 0$.*

Beweis: Im Falle eines relativen Maximums haben wir $\Delta y \geq 0$, also

$$\frac{\Delta y}{\Delta x}\begin{cases} \leq 0 & \text{für } \Delta x < 0 \\ \geq 0 & \text{für } \Delta x > 0 \end{cases}$$

und somit $0 \leq \frac{dy}{dx} \leq 0$. □. Relative Extremwerte liegen aber oft auch am Rand und werden hier nicht erfasst.

14.6 Mittelwertsatz

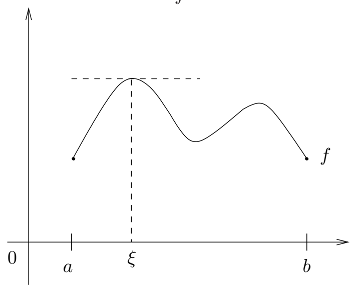
Satz 14.2 Mittelwertsatz. *Sei* $y = f(x)$ *auf dem Intervall* $[a,b]$ *stetig und auf dem offenen Intervall* $]a,b[= \{x \mid a < x < b\}$ *differenzierbar. Dann gibt es* p *mit*

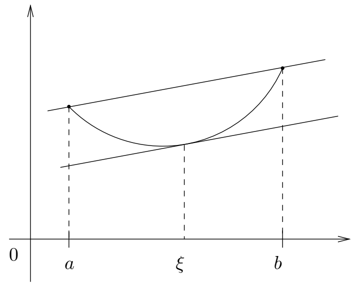
$$f'(p) = \frac{f(b)-f(a)}{b-a} \quad \textit{und} \quad a < p < b$$

Beweis. Wir zeigen zunächst den Spezialfall (Rolle): $f(a) = f(b) = 0$. Wegen der Stetigkeit werden absolutes Maximum und Minimum auf $[a,b]$ angenommen. Ist $f(x)$ nicht konstant 0, ist einer dieser Werte $\neq 0$ und muss an einem Punkt $p \in]a,b[$ angenommen werden. Nach dem Kriterium $f'(p) = 0$. Für den allgemeien Fall wende den Spezialfall an auf

$$g(x) = f(x) - [f(a) + \frac{f(b)-f(a)}{b-a}(x-a)]$$

d.h. subtrahiere von f die Sekante. □





Satz 14.3 Eindeutigkeitssatz. *Sind* $y = f(x)$ *und* $z = g(x)$ *auf dem Intervall* I *differenzierbar und* $f'(x) = g'(x)$ *auf* I*, so gibt es eine Konstante* C *mit*

$$g(x) = f(x) + C \quad \textit{für alle } x \in I$$

Gilt zudem $f(p) = g(p)$ *für ein* $p \in I$*, so* $f(x) = g(x)$ *für alle* $x \in I$.

Beweis. Setze $h(x) = f(x) - g(x)$. Nach Voraussetzung ist $h'(x) = f'(x) - g'(x) = 0$ auf I. Es ist zu zeigen, dass $h(x) = C$ konstant auf I. Wäre das nicht so, hätte man $a < b$ mit $h(a) \neq h(b)$, also nach dem Mittelwertsatz ein p mit $h'(p) \neq 0$.

14.7 Monotonie

Sei $y = f(x)$ auf $[a,b]$ stetig und auf $]a,b[= \{x \mid a < x < b\}$ differenzierbar. Dann

$f'(x) > 0$ auf $]a,b[$ $\Rightarrow f(x)$ streng monoton wachsend auf $[a,b]$
$f'(x) \geq 0$ auf $]a,b[$ $\Rightarrow f(x)$ monoton wachsend auf $[a,b]$
$f'(x) < 0$ auf $]a,b[$ $\Rightarrow f(x)$ streng monoton fallend auf $[a,b]$
$f'(x) \leq 0$ auf $]a,b[$ $\Rightarrow f(x)$ monoton wachsend auf $[a,b]$

Das folgt aus dem Mittelwertsatz. Z.B. für $f'(x) > 0$ so: $f(b) - f(a) = f'(p)(b-a) > 0$. □.

14.8 Natürlicher Logarithmus

Satz 14.4 *Der* natürliche Logarithmus *ist die eindeutig bestimmte Funktion* $x = \ln y$ *mit*

$$\frac{\partial \ln y}{\partial y} = \frac{1}{y}, \quad \ln 1 = 0, \quad y > 0$$

$\ln y$ *ist streng monoton wachsend und es gilt*

$$\ln(yv) = \ln y + \ln v, \quad \ln\frac{1}{y} = -\ln y$$

$$\ln y \to +\infty \textit{ für } y \to +\infty, \quad \ln y \to -\infty \textit{ für } y \to 0$$

Beweis. Die Existenz werden wir später (durch Integration) beweisen. Die Eindeutigkeit folgt sofort aus dem Eindeutigkeitssatz. Dass $\ln y$ streng monoton wächst, ist klar mit Abschn.14.7. Für festes v haben wir

$$\frac{\partial \ln(yv)}{\partial y} = \frac{1}{yv}v = \frac{1}{y} = \frac{\partial(\ln y + \ln v)}{\mathrm{d}y}$$

also folgt mit dem Eindeutigkeitssatz $\ln(yv) = \ln y + \ln v$ für alle v, da für $v = 1$ Gleichheit gilt. Insbesondere folgt $\ln\frac{1}{y} + \ln y = \ln 1 = 0$, also $\ln\frac{1}{y} = -\ln y$ .Weiterhin $\ln 2^n = n\ln 2 \to \infty$ und $\ln 2^{-n} = -n\ln 2 \to -\infty$ für $n \to \infty$, da $\ln 2 > 1$ wegen Monotonie. Also $B(\ln) = \mathbb{R}$. □

14.9 Exponentialfunktion

Satz 14.5 *Es gibt genau eine auf* $\mathbb{R}$ *definierte Funktion* $y = \exp x = e^x$ *mit*

$$(**) \quad \frac{\partial e^x}{\partial x} = e^x, \quad e^0 = 1,$$

die Exponentialfunktion. e^x *ist streng monoton wachsend und es gilt*

$$e^{-x} = \frac{1}{e^x}, \quad e^x > 0, \quad e^{x+u} = e^x \cdot e^u$$

$$x \leq e^x \to +\infty \textit{ für } 0 \leq x \to +\infty, \qquad e^x \to 0 \textit{ für } x \to -\infty$$

e^x *ist die Umkehrfunktion zu* $\ln y$*, d.h.* $y = e^x \Leftrightarrow x = \ln y, \quad y > 0.$

Beweis. Sei e^x die Umkehrfunktion zu $\ln y$. Nach der Umkehrregel gilt

$$\frac{\partial e^x}{\partial x} = \frac{1}{\frac{\partial \ln y}{\partial y}} = \frac{1}{\frac{1}{y}} = y = e^x$$

und $e^0 = 1$ weil $\ln 1 = 0$. Also erfüllt e^x die Bedingung $(**)$.

Für $g(x) = e^x e^{-x}$ gilt $g'(x) = e^x e^{-x} - e^x e^{-x} = 0$ und $g(0) = 1$, also $g(x) = 1$ konstant und somit $e^x \neq 0$ und $e^{-x} = \frac{1}{e^x}$ für alle x. Wäre $e^x < 0$ für ein x, so gäbe es wegen $e^0 = 1$ nach dem Zwischenwertsatz ein p mit $e^p = 0$ - Widerspruch. Für $x \geq 0$ wächst $e^x - x$ monoton, da Ableitung ≥ 0, also $e^x - x \geq e^0 - 0 = 1$.

Diese Aussagen folgen für jede Funktion $y = f(x)$, die $(**)$ erfüllt. Insbesondere ist eine solche Funktion umkehrbar und $B(f) =]0, \infty[$. Für die Umkehrfunktion $g(y) = x$ folgt nach der Umkehrregel

$$\frac{\partial x}{\partial y} = \frac{1}{\frac{\partial y}{\partial x}} = \frac{1}{y}$$

und $g(1) = 0$ weil $f(0) = 1$. Nach dem Eindeutigkeitssatz ist $g(y) = \ln y$ also $f(x) = e^x$. Damit ist die Eindeutigkeit gezeigt.

Für festes u betrachte nun $h(x) = \frac{1}{e^u}e^{x+u}$. Nun $h'(x) = h(x)$ und $h(0) = 1$, also $h(x) = e^x$. $\square$

14.10 Potenzen und Logarithmen

Definiere

$$c = b^a = e^{a \ln b} \quad \Leftrightarrow \quad a = \log_b c \quad \text{für } b, c > 0$$

Die bekannten Regeln für das Rechnen mit Potenzen und Logarithmen leitet man leicht her:

$$(xu)^a = x^a u^a$$

$$b^0 = 1,\ \ b^{x+u} = b^x b^u,\ \ b^{-x} = \frac{1}{b^x},\ b^{xu} = (b^x)^u$$

$$\log_b 1 = 0,\ \ \log_b(yv) = \log_b y + \log_b v,\ \ \log_b \frac{1}{y} = -\log_b y,\ \ \log_b(y^v) = v \log_b y$$

und man erhält die Paare von Umkehrfunktionen mit ihren Ableitungen (nach der Kettenregel)

$$y = x^a = e^a \ln x,\ \ x = y^{\frac{1}{a}} \ \ (= \sqrt[n]{y} \text{ falls } a = n \in \mathbb{N}) \quad x, y > 0$$

$$\frac{\partial x^a}{\partial y} = a x^{a-1},\ \ \frac{\partial y^{\frac{1}{a}}}{\partial y} = \frac{1}{a} y^{-\frac{a-1}{a}} \ \ (= \frac{1}{n} \frac{1}{\sqrt[n]{x^{n-1}}})$$

sowie

$$y = b^x = e^x \ln b,\ \ x = \log_b y = \frac{1}{\ln b} \ln y,\ \ y > 0 \ \text{ für } b > 0$$

$$\frac{\partial b^x}{\partial x} = \ln b b^x,\ \ \frac{\partial \log_b y}{\partial y} = \frac{1}{\ln b} \cdot \frac{1}{y}$$

14 Differentiation: Vektorfunktionen

14.12 Ableitung von Vektorfunktionen

Sei $f(t)$ für $|p - t| < \gamma$ definiert - z.B. Ort eines Massenpunktes zur Zeit t. Die Differenz von f an der Stelle p ist definiert als

$$\Delta\vec{y} = \Delta(\Delta t,\, p) = f(p + \Delta t) - f(t)$$

und gibt z.B. die Ortsänderung an. Die Stetigkeit von $\vec{y} = f(t)$ an der Stelle p bedeutet dann

$$\Delta\vec{y} \to \vec{0} \text{ für } \Delta t \to 0$$

also die Gleichmäßigkeit der Bewegung. Der Quotient

$$\frac{\Delta\vec{y}}{\Delta t} = \frac{f(p + \Delta t) - f(p)}{\Delta t} \qquad \Delta t \neq 0$$

ist ein *Differenzenquotient*, und kann z.B. eine Geschwindigkeit nach Grösse und Richtung beschreiben. Wir betrachten die Funktion

$$\frac{\Delta\vec{y}}{\Delta t} \quad \text{für } \Delta t \to 0,\ \Delta t \neq 0$$

Existiert der Limes, so heisst f an der Stelle p *differenzierbar* und wir haben die *Ableitung* $\vec{a}$ and der Stelle p mit

$$(1)\ \ \frac{\Delta\vec{y}}{\Delta t} \to \vec{a} = f'(p) = \frac{\partial\vec{y}}{\partial t} = \dot{\vec{y}} \quad \text{für } \Delta t \to 0,\ \Delta t \neq 0$$

Satz 14.1 *Ist $f(x)$ an der Stelle p differenzierbar, so auch stetig:*

$$\Delta\vec{y} = \frac{\Delta\vec{y}}{\Delta t} \cdot \Delta t \ \to \ \frac{\partial\vec{y}}{\partial t} \cdot 0 = 0 \ \ \textit{für } \Delta t \to 0,\ \Delta t \neq 0$$

$$(*) \quad \frac{\partial y}{\partial x} \neq 0 \ \Rightarrow\ \Delta y \neq 0 \ \textit{für } \Delta x \to 0,\ \Delta x \neq 0$$

Haben wir nämlich $\frac{\Delta\vec{y}}{\Delta t} \to \vec{a} \neq \vec{0}$, so $\|\frac{\Delta\vec{y}}{\Delta t}\| \geq \frac{\|a\}}{2} > 0$ für hinreichend kleine Δt, also $\Delta\vec{y} \neq 0$. $\square$

Zur Differenzierbarkeit äquivalente Bedingungen sind wie bei reellwertigen Funktionen

$$(2)\ \ \Delta\vec{y} = \frac{\partial\vec{y}}{\partial t}\Delta t + \phi(\Delta t)\Delta t \quad \text{mit } \phi(\Delta t) \to \vec{0} \text{ für } \Delta t \to 0,\ \Delta t \neq 0$$

$$(3)\ \ \Delta\vec{y} = \frac{\partial\vec{y}}{\partial t}\Delta t + R(\Delta t) \quad \text{mit } \frac{R(\Delta t)}{\Delta t} \to \vec{0} \text{ für } \Delta t \to 0,\ \Delta t \neq 0$$

Satz 14.2 *Ist eine Basis $\vec{e}_1, \ldots \vec{e}_n$ gegeben und $f(t) = f_1(t)\vec{e}_1 + \ldots + f_n(t)\vec{e}_n$, so ist f an der Stelle p differenzierbar genau dann, wenn es die $y_i = f_i(t)$ $i = 1, \ldots, n$ sind, und es gilt*

$$\frac{\partial\vec{y}}{\partial t} = \vec{a} \ \Leftrightarrow\ \frac{\partial y_i}{\partial t} = a_i \ (i = 1, \ldots, n)$$

Beweis. Ist $\phi(\Delta t) = \sum_i \phi_i(\delta t)\vec{e}_i$, so

$$\lim_{\Delta t \to 0} \phi(\Delta t) = 0 \ \Leftrightarrow\ \lim_{\Delta t \to 0} \phi_i(\Delta t) = 0 \text{ für } i = 1, \ldots, n$$

14.13 Differentiationsregeln für Vektorfunktionen

$$\frac{\partial C\vec{y}}{\partial t} = C\frac{\partial\vec{y}}{\partial t}$$

Beweis: $\Delta C\vec{y} = C\Delta\vec{y}$

$$\text{Summenregel} \quad \frac{\partial(\vec{y} + \vec{z})}{\partial t} = \frac{\partial\vec{y}}{\partial t} + \frac{\partial\vec{z}}{\partial t}$$

Beweis: $\Delta(\vec{y}+\vec{z}) = \Delta\vec{y}+\Delta\vec{z}$

$$\text{Produktregel}\quad \frac{\partial(y\vec{z})}{\partial t} = y\frac{\partial \vec{z}}{\partial t} + \frac{\partial y}{\partial t}\vec{z} \quad (y = f(t) \in \mathbb{R})$$

Beweis: $\Delta(y\vec{z}) = y\Delta\vec{z} + (\Delta y)\vec{z} + \Delta y \Delta\vec{z}$

$$\frac{\Delta(y\vec{z})}{\Delta t} = y\frac{\Delta\vec{z}}{\Delta t} + \frac{\Delta y}{\Delta t}\vec{z} + \Delta y\frac{\Delta\vec{z}}{\Delta t} \to y\frac{\partial\vec{z}}{\partial t} + \frac{\partial y}{\partial t}\vec{z} + 0\frac{\partial\vec{z}z}{\partial t}$$

$$\text{Skalarproduktregel}\quad \frac{\partial\langle\vec{y}\,|\,\vec{z}\rangle}{\partial t} = \langle\vec{y}\,|\,\frac{\partial\vec{z}}{\partial t}\rangle + \langle\vec{z}\,|\,\frac{\partial y}{\partial t}\rangle$$

Beweis:

$$\Delta\langle\vec{y}\,|\,\vec{z}\rangle = \langle\vec{y}+\Delta\vec{y}\,|\,\vec{z}+\Delta\vec{z}\rangle - \langle\vec{y}\,|\,\vec{z}\rangle = \langle\vec{y}\,|\,\Delta\vec{z}\rangle + \langle\vec{z}\,|\,\Delta\vec{y}\rangle + \langle\Delta\vec{y}\,|\,\Delta\vec{z}\rangle$$

$$\frac{\Delta\langle\vec{y}\,|\,\vec{z}\rangle}{\Delta t} = \langle\vec{y}\,|\,\frac{\Delta\vec{z}}{\Delta t}\rangle + \langle\vec{z}\,|\,\frac{\Delta\vec{y}}{\Delta t}\rangle + \langle\Delta\vec{y}\,|\,\frac{\Delta\vec{z}}{\Delta t}\rangle \to \langle\vec{y}\,|\,\frac{\partial\vec{z}}{\partial t}\rangle + \langle\vec{z}\,|\,\frac{\partial y}{\partial t}\rangle + \langle\vec{0}\,|\,\frac{\partial z}{\partial t}\rangle$$

Satz 14.3 Kettenregel *Sei $y = f(x)$ reellwertig. Sind $y = f(x)$ an der Stelle p und $\vec{z} = g(y)$ an der Stelle $q = f(p)$ differenzierbar, so ist $\vec{z} = g(f(x))$ an der Stelle p differenzierbar und es gilt*

$$\frac{\partial\vec{z}}{\partial x} = \frac{\partial\vec{z}}{\partial y}\cdot\frac{\partial y}{\partial x} \qquad \textit{genauer } (g\circ f)'(p) = g'(f(p))\cdot f'(p)$$

Beweis wie für reellwertiges g.

14.14 Komplexwertige Funktionen

Wir definieren

$$e^{(a+bj)t} = e^{at}e^{bjt} = e^{at}(\cos bt + j\sin bt)$$

Satz 14.4

$$\frac{\partial e^{(a+bj)t}}{\partial t} = (a+bj)e^{(a+bj)t}$$

Beweis. $e^{bjt} = \cos bt + j\sin bt$ und mit der Kettenregel

$$\frac{\partial\cos bt}{\partial t} = -b\sin bt, \quad \frac{\partial\sin bt}{\partial t} = b\cos bt$$

also

$$\frac{\partial e^{bjt}}{\partial t} = -b\sin t + j\cos bt = jbe^{jbt}$$

Nun mit der Produktregel

$$\frac{\partial e^{(a+bj)t}}{\partial t} = e^{at}\frac{\partial e^{bjt}}{\partial t} + \frac{\partial e^{at}}{\partial t}e^{bjt} = e^{at}bje^{bjt} + ae^{at}e^{bjt} = (a+bj)e^{at}e^{bjt} = (a+bj)e^{(a+bj)t}$$

Wie für reelle Funktionen erhält man

$$\text{Produktregel}\quad \frac{\partial yz}{\partial t} = y\frac{\partial z}{\partial t} + z\frac{\partial y}{\partial t}$$

$$\text{Quotientenregel}\quad \frac{\partial\frac{z}{y}}{\partial x} = \frac{1}{y^2}\cdot(y\frac{\partial z}{\partial x} - z\frac{\partial y}{\partial x})$$

Beweis.

$$\frac{\Delta\frac{1}{y}}{\Delta t} = (\frac{1}{y+\Delta y} - \frac{1}{y})\frac{1}{\Delta t} = \frac{y-(y+\Delta y)}{y^2+y\Delta y}\frac{1}{\Delta t} = \frac{-1}{y^2+y\Delta y}\frac{\Delta y}{\Delta t} \to \frac{-1}{y^2}\frac{\partial y}{\partial t} \text{ für } \Delta t\to 0$$

14 Differentiation: Differentiale

Um Missverständnisse beim Gebrauch von Differentialen zu vermeiden, werden wir die bei Funktionen mehrerer Variablen übliche Schreibweise für die Ableitung auch im Falle von Funktionen $y = f(x)$ einer Variablen benutzen

$$f'(p) = \frac{\partial f}{\partial x} = \frac{\partial y}{\partial x} = \frac{\partial}{\partial x}y = \frac{\partial}{\partial x}f$$

Beachte, dass es bei $\frac{\partial y}{\partial x}$ um einen Ausdruck geht, dessen Bestandteile ∂y und ∂x keine Bedeutung haben. Dagegen kann $\frac{\partial}{\partial x}$ als *Differentialoperator* verstanden werden.

14.15 Differentiale

Sei $y = f(x)$ reellwertig. Den homogen linearen Anteil der Funktion in (4) nennt man auch das *Differential* $\mathrm{d}y$ von $y = f(x)$ an der Stelle p

$$\mathrm{d}y = \mathrm{d}y(p) : \mathbb{R}\to\mathbb{R}, \quad \mathrm{d}y(p)(\Delta x) = f'(p)\Delta x$$

oder, wenn man die Stelle p im Hinterkopf hat

$$\mathrm{d}y(\Delta x) = \frac{\partial y}{\partial x}\Delta x.$$

Versteht man nun $\mathrm{d}x$ als eine Bezeichnung für die identische Funktion $\mathrm{d}x(\Delta x) = \Delta x$ so kann man den Differentialquotienten als den Quotienten der beiden Funktionen auffassen (wieder an der Stelle p) mit dem Limes $f'(p)$ für $0\neq\Delta x\to 0$

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y(\Delta x)}{\mathrm{d}x(\Delta x)} = \frac{f'(p)\Delta x}{\Delta x} \to f'(p) = \frac{\partial y}{\partial x} \text{ für } 0\neq\Delta x\to 0$$

Für vektorwertige Funktionen besagt (3), dass man $f(t)$ in der Nähe von p durch eine affin lineare Funktion approximieren kann, mit einem Fehler, der "klein von zweiter Ordnung" ist

$$(4)\quad f(p+\Delta t) = f(p) + \Delta\vec{y} = f(p) + \frac{\partial\vec{y}}{\partial t}\Delta t + R(\Delta t)$$

Den homogen linearen Anteil der Funktion in (4) nennt man auch das *Differential* $\mathrm{d}\vec{y}$ von $\vec{y} = f(t)$ an der Stelle p und hat die Analogie zum reellwertigen Fall.

14.16 Infinitesimale Differentiale

Hat man gelernt, mit infinitesimalen Zahlen $\mathrm{d}x$ (d.h. $|\mathrm{d}x| < \frac{1}{n}$ für alle n, aber $\mathrm{d}x \neq 0$ - insbesondere ist $\mathrm{d}x$ keine reelle Zahl!) zu rechnen und setzt eine solche für Δx in (4) ein, so ergibt sich

$$f(p+\mathrm{d}x) = f(p) + f'(p)\mathrm{d}x + R(\mathrm{d}x) = f(p) + \mathrm{d}y + R(\mathrm{d}x)$$

wenn man das Argument $\mathrm{d}x$ der Funktion $\mathrm{d}y$ unterschlägt. Also

$$\mathrm{d}y = f'(p)\mathrm{d}x$$

und $\mathrm{d}y$ ist bis auf einen Fehler, der infinitesimal von zweiter Ordnung (relativ zu $\mathrm{d}x$) ist, die Änderung von y bei der gegebenen Änderung $\mathrm{d}x$ von x. Dividiert man durch $\mathrm{d}x$ (das darf man, weil $\mathrm{d}x \neq 0$), so folgt

$$\frac{f(p+\mathrm{d}x) - f(p)}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}x} + \frac{R(\mathrm{d}x)}{\mathrm{d}x}$$

mit infinitesimalem $\frac{R(\mathrm{d}x)}{\mathrm{d}x}$. Damit ist im Falle der Differenzierbarkeit an der Stelle p, $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\partial y}{\partial x}$ die eindeutig bestimmte reelle Zahl mit infinitesimalem Abstand zu $\frac{f(p+\mathrm{d}x)-f(p)}{\mathrm{d}x}$. Aber

$\mathrm{d}y$ ist **nicht** die Änderung von y bei Änderung $\mathrm{d}x$ von x, sondern nur in "erster Näherung"

Entsprechend geht es für vektorwertige Funktionen und infinitesimale Vektor $\vec{v} \neq \vec{0}$ mit $\|\vec{v}\| < \frac{1}{n}$ für alle n. Dann ist im Falle der Differenzierbarkeit an der Stelle p, $\frac{\mathrm{d}\vec{y}}{\mathrm{d}t} = \frac{\partial \vec{y}}{\partial t}$ der eindeutig bestimmte reelle Vektor mit infinitesimalem Abstand zu $\frac{f(p+\mathrm{d}t)-f(p)}{\mathrm{d}t}$.

Differentiale sind hier also im Prinzip genau dasselbe, wie oben, nur dass man dass es genügt, zu jeder unabhängigen Variablen x ein einziges infinitesimales Argument $\mathrm{d}x$ in die Rechnung einzuführen. Das ist of bequemer als das Rechnen mit Limites oder ε und δ. Aber man muss beachten, dass $\mathrm{d}x \neq 0$ und dass der Übergang von einer Nonstandard-Zahl zu der nächsten reellen Zahl nicht umkehrbar ist.

14.17 Grössen und Differentiale

Obiger Sichtweise von Differentialen liegt der funktionale Zusammenhang $y = f(x)$ zugrunde: dieser Bezug muss immer gegeben sein, wenn man von dem Differential $\mathrm{d}y$ spricht (auch wenn man das als infinitesimal ansieht).

In Naturwissenschaft und Technik spricht man aber traditionell (und vorteilhaft) von *Größen* und ihren Differentialen und hat aus dem Zusammenhang auch eine hinreichend verläßliche Art mit dieser Vorstellungsweise umzugehen.

In üblichen mathematischen Begriffen kann man ein *System von Grössen* $\xi_1, \ldots, \xi_n$ als eine Teilmenge S von $\mathbb{R}^n$ fassen, d.h. die Grössen ξ_i nehmen reelle Werte x_i an und ein n-Tupel $(x_1, \ldots, x_n) \in S$ entspricht einem *Zustand* des Systems S. Der *Wertebereich* W_k von ξ_k besteht dann aus allen $x_k \in \mathbb{R}$, zu denen es $(x_1, \ldots, x_n) \in S$ gibt.

Die Grösse ξ_i *hängt funktional von* der Größe ξ_k *ab*, falls für alle Zustände $(x_1, \ldots, x_n)$ und $(z_1, \ldots, z_n)$ des Systems S gilt: Ist $x_k = z_k$ so auch $x_i = z_i$, d.h. die Zuordnung $x_k \mapsto x_i$

ist durch eine Funktion $f_{ik} : W_k \to \mathbb{R}$ gegeben. Sind f_{ik} und f_{kl} gegeben, so folgt $x_i = f_{il}(x_l)$ aus $x_i = f_{ik}(x_k)$ und $x_k = f_{kl}(x_l)$ und man hat, im Falle der Differenzierbarkeit, $f'_{il}(p) = f'_{ik}(f_{kl}(p)) \cdot f'_{kl}(p)$ nach der Kettenregel.

Sei nun J eine Menge von Indexpaaren (i,k) so, dass für $(i,k) \in J$ die Größe ξ_i von ξ_k funktional abhängt und die Funktion f_{ik} auf W_k die Ableitung f'_{ik} hat. Ein J-*kompatibles System von Differentialen* für S ist dann ein System von Abbildungen $\mathrm{d}x_k : W_k \to \mathbb{R} \setminus \{0\}$ $(k = 1, \ldots, n)$ so, dass gilt

- $\mathrm{d}x_i(f_{ik}(p)) = f'_{ik}(p) \cdot \mathrm{d}x_k(p)$ für alle $p \in W_k$ und $(i,k) \in J$.
- äquivalent: $f'_{ik}(p) = \frac{\mathrm{d}x_i(f_{ik}(p))}{\mathrm{d}x_k(p)}$
- kryptisch formuliert: $\mathrm{d}x_i = \frac{\mathrm{d}x_i}{\mathrm{d}x_k}\mathrm{d}x_k$, besser $\mathrm{d}x_i = \frac{\partial x_i}{\partial x_k}\mathrm{d}x_k$

Hat man das, so darf man unbesorgt mit den Differentialen $\mathrm{d}x_k$ rechnen und $\mathrm{d}x_k$ als nur von ξ_k abhängige Größe ansehen. Die Preisfrage ist aber, ob es ein solches kompatibles System von Differentialen gibt. Eine hinreichende Bedingung (und gewagte Annahme) ist, dass es eine Größe ξ_{i_0} gibt, von der alle ξ_k funktional abhängen mit differenzierbarer und umkehrbarer Abbildung $f_{ki_0} : W_{i_0} \to W_k$. Dann wähle man z.B. ein (klitzekleines) $\varepsilon \neq 0$ und setze $\mathrm{d}x_k(p) = f'_{ki_0}(q)\varepsilon$ mit $f_{ki_0}(q) = p$ (also $\mathrm{d}x_{i_0} = \varepsilon$).

Es ist nicht die Frage, ob man Differentiale als unendlich klein ansehen darf, sondern wie man sie als Änderungsgrössen, (in erster Näherung) von Grössen ansehen darf.

Darauf hat schon Euler hingewiesen.

14.18 Fehlerabschätzung

Für $y = x^2$ haben wir den Fehler $\Delta y = \Delta y(p, \Delta x)$ beim Nachweis der Stetigkeit so bestimmt

$$\Delta y(p, \delta x) = 2p\Delta x + (\Delta x)^2$$

Vernachlässigen wir den quadratischen Term $(\Delta x)^2$ so ergibt sich das Differential

$$Dey(p, \Delta x) \approx 2p\Delta x = \mathrm{d}y(p, \Delta x)$$

und wir können $|\Delta y|$ näherungsweise so abschätzen

$$|\Delta y(p)| \lessapprox |2p|\delta \quad \text{für } |\Delta x| \leq \delta$$

Allgemein für in der Nähe von p stetig differenzierbares $y = f(x)$

$$|\Delta y(p)| \lessapprox |f'(p)|\delta \quad \text{für } |\Delta x| \leq \delta$$

und damit ein *relativer Fehler*

$$|\frac{\Delta y}{y}|(p) = \frac{|\Delta y(p)|}{|fp|} \lessapprox \frac{|f'(p)|}{|f(p)|}\delta$$

14 Differentiation: Höhere Ableitungen

14.19 Konvexität und höhere Ableitungen

Im vergangenen Abschnitt haben wir mit Hilfe der Ableitungen das Monotonieverhalten einer Funktion charakterisiert. Nun schauen wir uns das Krümmungsverhalten an.

Definition 14.1 *Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt* konvex*, wenn für alle $x_1, x_2 \in [a,b]$ und für alle $t \in [0,1]$ gilt:*

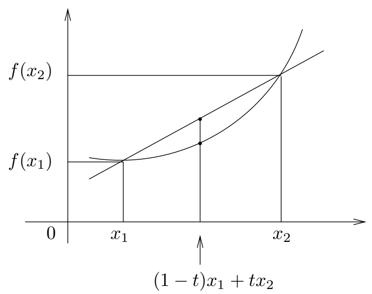
$$f((1-t)x_1 + t\,x_2) \leq (1-t)f(x_1) + t\,f(x_2)\,. \tag{14.1}$$

Steht in (14.1) *das Relationszeichen $\geq$, so heißt f* konkav *auf $[a,b]$.*

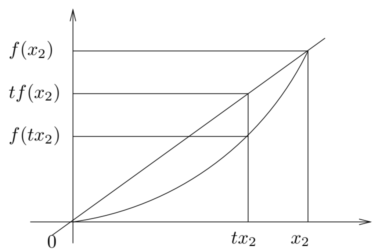
Man beachte, dass für beliebige Zahlen $x_1, x_2 \in \mathbb{R}$ mit $x_1 < x_2$ gilt

$$\{(1-t)x_1 + tx_2 : t \in [0,1]\} = [x_1, x_2]\,.$$

Geometrische Deutung: Der Graph einer konvexen (konkaven) Funktion liegt unterhalb (oberhalb) der Verbindungsstrecke zweier beliebiger seiner Punkte.



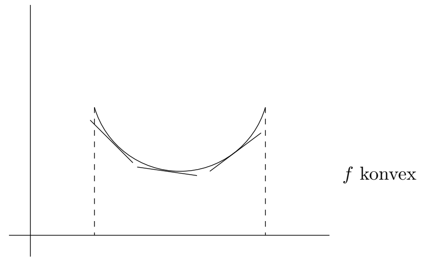
Diese Deutung sieht man am einfachsten ein, wenn man den Graphen von f so verschiebt, dass $x_1 = f(x_1) = 0$:



Spezialfall $x_1 = f(x_2) = 0$

Die Konvexität differenzierbarer Funktionen läßt sich wie folgt charakterisieren.

Satz 14.2 *Sei f stetig auf $[a,b]$ und differenzierbar auf (a,b). Dann ist f genau dann konvex (konkav) auf $[a,b]$, wenn die Ableitung f' auf (a,b) monoton wachsend (fallend) ist.*



Nun wissen wir aus 14.7, dass man die Monotonie einer differenzierbaren Funktion f mit Hilfe ihrer Ableitung f' beschreiben kann. In Satz 14.2 benötigen wir die Monotonie der Ableitung. Diese könnten wir mit Mitteln der Differentialrechnung untersuchen, wenn wir wüssten, dass f' differenzierbar ist.
Differenzierbare Funktionen auf (a,b), deren Ableitung f' auf (a,b) differenzierbar ist, heißen *zweimal differenzierbar*, und $(f')'$ heißt *zweite Ableitung* von f. Statt $(f')'$ schreibt man f''. Ganz analog erklärt man k-mal differenzierbare Funktionen und k-te Ableitungen. Für die k-te Ableitung von f im Punkt $x_0 \in (a,b)$ schreibt man

$$f^{(k)}(x_0) \quad \text{oder} \quad \frac{d^k f}{dx^k}(x_0)\,.$$

Mitunter ist es zweckmäßig, die Funktion f selbst als 0-te Ableitung von f zu betrachten. Man schreibt dann $f = f^{(0)}$.

Aus Satz 14.2 und Abschn.14.7 erhalten wir:

Satz 14.3 *Sei f auf $[a,b]$ stetig und auf (a,b) zweimal differenzierbar. Dann ist f genau dann konvex (konkav) auf $[a,b]$, wenn $f''(x) \geq 0$ ($f''(x) \leq 0$) für alle $x \in (a,b)$.*

Beispiele Für die Exponentialfunktion ist $(e^x)'' = (e^x)' = e^x$. Also ist exp auf ganz $\mathbb{R}$ konvex. Für die Sinusfunktion ist $(\sin x)'' = (\cos x)' = -\sin x$. Also ist sin z.B. auf $[0,\pi]$ konkav und auf $[\pi, 2\pi]$ konvex. ■

14.20 Der Satz von Taylor

Wir schreiben den Mittelwertsatz

$$\exists \xi \in (x_0, x): \quad \frac{f(x) - f(x_0)}{x - x_0} = f'(\xi)$$

in der Form

$$\exists \xi \in (x_0, x): \quad f(x) = f(x_0) + f'(\xi)(x - x_0) \tag{14.2}$$

und interpretieren ihn neu: *Der Term $f'(\xi)(x-x_0)$ beschreibt den Fehler, den man begeht, wenn man die Funktion f durch die konstante Funktion $f(x_0)$ ersetzt.* Man kann nun daran denken, die Funktion f genauer zu approximieren, indem man nicht nur (wie in (14.2)) konstante Funktionen, sondern Polynome zur Approximation zuläßt. Es ist nahe liegend, diese Polynome P so zu wählen, dass nicht nur die Funktionswerte $f(x_0)$ und $P(x_0)$ übereinstimmen (wie in (14.2)), sondern auch die Werte der Ableitungen $f'(x_0) = P'(x_0)$, $f''(x_0) = P''(x_0), \ldots, f^{(n)}(x_0) = P^{(n)}(x_0)$, wobei n der Grad des Polynoms P ist. Wir überlegen uns zunächst, wie ein solches Polynom aussieht und machen dazu den Ansatz

$$P(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \ldots + a_n(x-x_0)^n .$$

Offenbar ist $P(x_0) = a_0$. Weiter ist

$$P'(x) = a_1 + 2a_2(x-x_0) + \ldots + na_n(x-x_0)^{n-1}$$

und daher $P'(x_0) = a_1$. Aus

$$P''(x) = 2a_2 + \ldots + n(n-1)a_n(x-x_0)^{n-2}$$

folgt $a_2 = \frac{1}{2}P''(x_0)$. Allgemein findet man

$$a_k = \frac{1}{k!}\, P^{(k)}(x_0)\,,$$

was mit den Vereinbarungen $0! = 1$ und $P^{(0)} = P$ auch für $k = 0$ richtig ist. Wir erhalten damit

$$P(x) = \sum_{k=0}^{n} a_k(x-x_0)^k = \sum_{k=0}^{n} \frac{1}{k!}\, P^{(k)}(x_0)(x-x_0)^k .$$

Ist also f eine in x_0 n-mal differenzierbare Funktion, so wird durch

$$T_n(x,x_0) := \sum_{k=0}^{n} \frac{1}{k!}\, f^{(k)}(x_0)(x-x_0)^k \tag{14.3}$$

ein Polynom vom Grad $\leq n$ definiert, welches im Punkt x_0 in allen Ableitungen bis zur n-ten mit der Funktion f übereinstimmt. Dieses Polynom heißt das *Taylorpolynom vom Grad n von f im Punkt x_0*. Der Fehler, der beim Ersetzen einer Funktion durch ihr Taylorpolynom gemacht wird, wird in folgendem Satz beschrieben. Dazu vereinbaren wir, eine Funktion $f : [a,b] \to \mathbb{R}$ *n-mal stetig differenzierbar* zu nennen, wenn sie auf $[a,b]$ n-mal differenzierbar und ihre n-te Ableitung stetig ist.

Satz 14.4 (Taylor) *Sei $f : [a,b] \to \mathbb{R}$ n-mal stetig differenzierbar, und auf (a,b) existiere die $n{+}1$-te Ableitung. Dann gibt es ein $\xi \in (a,b)$ so, dass*

$$f(b) = \underbrace{\sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a)(b-a)^k}_{= T_n(b,a)} + \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}\,(b-a)^{n+1}}_{=: R_n(b,a)} . \tag{14.4}$$

Dabei ist $T_n(b,a)$ das Taylorpolynom vom Grad n von f in a, und $R_n(b,a)$ heißt das Restglied nach Lagrange.

Beweisidee Man definiert eine Funktion $h : [a,b] \to \mathbb{R}$ durch

$$h(x) := f(b) - f(x) - \frac{f'(x)}{1!}(b-x) - \ldots - \frac{f^{(n)}(x)}{n!}(b-x)^n - m\,\frac{(b-x)^{n+1}}{(n+1)!}\,,$$

wobei m so gewählt wird, dass $h(a) = 0$. Die Funktion h ist stetig auf $[a,b]$, differenzierbar auf (a,b), und es ist $h(a) = h(b) = 0$. Ihre Ableitung im Punkt $x \in (a,b)$ ist

$$h'(x) = -\frac{f^{(n+1)}(x)}{n!}\,(b-x)^n + m\,\frac{(b-x)^n}{n!} . \tag{14.5}$$

(Nachrechnen!) Nach Rolle gibt es ein $\xi \in (a,b)$ mit $h'(\xi) = 0$. Aus (14.5) erhält man damit für m

$$-\frac{f^{(n+1)}(\xi)}{n!}(b-\xi)^n + m\,\frac{(b-\xi)^n}{n!} = 0 \quad \text{bzw.} \quad m = f^{(n+1)}(\xi) .$$

Man setzt diesen Wert für m in die Definition von h ein, wählt $x = a$ und erhält (14.4). ■

Beispiel 6 Die Sinusfunktion ist auf ganz $\mathbb{R}$ beliebig oft differenzierbar. Wir können also für jeden Punkt $x_0 \in \mathbb{R}$ Taylorpolynome beliebig hoher Ordnung bilden. Mit $f(x) = \sin x, f'(x) = \cos x, f''(x) = -\sin x, f'''(x) = -\cos x, f^{(4)}(x) = \sin x$ erhalten wir für $a = x_0 = 0$

$$\frac{f^{(k)}(0)}{k!} = \begin{cases} 0 & \text{für } k \text{ gerade} \\ 1/k! & \text{für } k = 1,5,9,13,\ldots \\ -1/k! & \text{für } k = 3,7,11,15,\ldots . \end{cases}$$

Die ersten Taylorpolynome von f im Punkt 0 sind also

$$\begin{aligned} T_0(x,0) &= 0\,, \\ T_1(x,0) &= x\,, \\ T_2(x,0) &= T_1(x,0)\,, \\ T_3(x,0) &= x - \frac{x^3}{3!}\,, \\ T_4(x,0) &= T_3(x,0)\,, \end{aligned}$$

und der Satz von Taylor ergibt für $n = 2m-1, m \in \mathbb{N}$,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^{m-1}\frac{x^{2m-1}}{(2m-1)!} + R_{2m-1}(x,0).$$

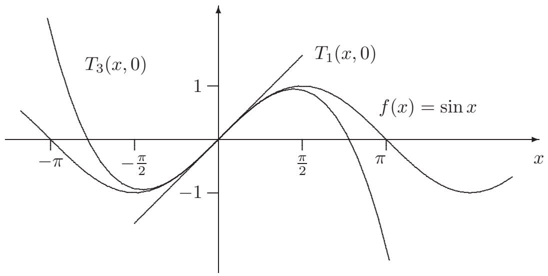
Das Restglied ist von der Gestalt

$$\frac{f^{(2m)}(\xi)}{(2m)!}x^{2m} \quad \text{mit einem } \xi \in (0,x),$$

und wegen $|f^{(2m)}(\xi)| \leq 1$ können wir das Restglied abschätzen:

$$|R_{2m-1}(x,0)| \leq \frac{|x|^{2m}}{(2m)!}.$$

Damit haben wir die Möglichkeit, $\sin x$ mit einer vorgegebenen Genauigkeit näherungsweise zu berechnen.



■

Beispiel 7 Sei $f(x) = \ln(x+1), x_0 = a = 0$ sowie $x = b \in (-1, \infty)$. Dann ist $f'(x) = \frac{1}{1+x}$ und $f^{(k)}(x) = (-1)^{k-1}(k-1)!(1+x)^{-k}$ für $k \geq 2$ und daher

$$f(0) = 0, \quad \frac{f^{(k)}(0)}{k!} = \frac{(-1)^{k-1}}{k} \quad \text{für } k \geq 1.$$

Einsetzen in die Taylorsche Formel liefert

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n-1}\frac{x^n}{n} + (-1)^n \frac{1}{(1+\xi)^{n+1}} \frac{x^{n+1}}{n+1}$$

mit einem $\xi \in (0, x)$. Für $x = 1$ ist $\xi \in (0,1)$ und folglich

$$|R_n(1,0)| = \left|(-1)^n \frac{1}{(1+\xi)^{n+1}} \frac{x^{n+1}}{n+1}\right| < \frac{1}{n+1}.$$

Wegen $R_n(1,0) \to 0$ ist klar: Die Reihe $\sum_{k=1}^{\infty} (-1)^{k-1}\frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$ konvergiert gegen $\ln 2$. ■

Beispiel 8 Für $f(x) = e^x$ ist $f^{(k)}(0) = 1$ für jedes $k \in \mathbb{N}_0$ und somit

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{e^{\xi}}{(n+1)!}x^{n+1}$$

mit einem $\xi \in (0, x)$. Das n-te Taylorpolynom ist also gerade die n-te Partialsumme der Reihe, mit der exp oft definiert wird. Wir kommen auf diesen Zusammenhang später zurück. ■

14.23 Bezeichnungen

Für Funktionen $y = f(t) \in \mathbb{R}$ mit $t \in \mathbb{R}$ und $p \in D(f)$ benutzen wir

- $\Delta y = \Delta f = (\Delta f)(p, \Delta t) = f(p + \Delta t) - f(p)$ *Differenz* oder *Änderung* von f an der Stelle p bei Änderung Δt
  $(\Delta f)(p) : \mathbb{R} \to \mathbb{R}$ ist eine Abbildung: $(\Delta f)(p)(\Delta t) = \Delta f(p, \Delta t)$
- $\frac{\partial y}{\partial t} = \frac{\partial f}{\partial t}(p) = f'(p)$ *Ableitung, Differentialquotient* von f an der Stelle p
  $f'(p) \in \mathbb{R}$ ist eine Zahl

- $\mathrm{d}y = \mathrm{d}f = (\mathrm{d}f)(p, \mathrm{d}t) = f'(p) \cdot \mathrm{d}t$ *Differential* von f an Stelle p bei Änderung $\mathrm{d}t$
  $(\mathrm{d}f)(p) : \mathbb{R} \to \mathbb{R}$ ist eine homogen-lineare Abbildung: $(\mathrm{d}f)(p)(\mathrm{d}t) = (\mathrm{d}f)(p, t)$

Δt wird meist als sehr klein, $\mathrm{d}t$ auch als infinitesimal gedacht, beide $\neq 0$. Dann ist, Differenzierbarkeit vorausgesetzt

$$(\Delta y)(p, \Delta t) \approx (\mathrm{d}y)(p, \Delta t)$$

$$\text{genauer: } \frac{(\Delta y)(p, \Delta t) - (\mathrm{d}y)(p, \Delta t)}{\Delta t} \to 0 \text{ für } \Delta t \to 0$$

$$\frac{(\Delta y)(p, \mathrm{d}t) - (\mathrm{d}y)(p, \mathrm{d}t)}{\mathrm{d}t} \text{ infinitesimal für infinitesimale } \mathrm{d}t$$

14.24 Satz von Lagrange-McLaurin-Taylor

Satz 14.5 *Sei $\delta = b - a > 0$ und $f : [a, b] \to \mathbb{R}$. Dann sind äquivalent*

(i) $f(a+t) = \sum_{k=0}^{n} a_k t^k + R(t)$ für alle $t \in [0, \delta]$ mit $R^{(k)}(0) = 0$ für $0 \leq k \leq n$ und $R(t)$ auf $[0, \delta]$ n-mal stetig differenzierbar und auf $]0, \delta[$ $n+1$-mal differenzierbar

(ii) $a_k = \frac{1}{k!}f^{(k)}(a)$ für $0 \leq k \leq n$ und $y = f(x)$ auf $[a, b]$ n-mal stetig differenzierbar und auf $]a, b[$ $n+1$-mal differenzierbar.

Die Funktion $R(t)$ in (i) ist eindeutig bestimmt durch die Bedingung

$$R^{(k)}(0) = 0 \text{ für } 0 \leq k \leq n \text{ und } f^{(n+1)}(a+t) = R^{(n+1)}(t) \text{ für alle } t \in]0, \delta[$$

Weiterhin gilt

$$\text{zu jedem } t \in [0, \delta] \text{ gibt es } \xi(t) \in [0, t] \text{ mit } \quad R(t) = \frac{f^{(n+1)}(a + \xi(t))}{(n+1)!}t^{n+1}$$

$$\text{Gilt } |f^{(n+1)}(a+\xi)| \leq M \text{ für alle } \xi \in [0, t] \text{ so folgt } \quad |R(t)| \leq \frac{M}{(n+1)!}|t|^{n+1}$$

Entsprechendes gilt für b und $t \in [-\delta, 0]$.

Beweis. Beweis von $(i) \Leftrightarrow (ii)$ durch Induktion über n. Beachte dass

$$\frac{\partial f(a+t)}{\partial t}(t_0) = \frac{\partial f(x)}{\partial x}(a + t_0)$$

und dass Polynome beliebig oft differenzierbar sind. Damit folgen aus den Differenzierbarkeitsannahmen für $f(x)$ die für $R(t)$ und umgekehrt.

Der Fall $n = 1$ ist gerade die Charakterisierung der Ableitung. Sei $n > 1$ und (i) angenommen. Ableiten vom (i) nach t ergibt

$$f'(a+t) = \sum_{k=1}^{n} k a_k t^{k-1} + R'(t), \quad R'^{(k)}(0) = R^{(k+1)}(0) = 0 \text{ für } 0 \leq k \leq n-1$$

Also nach Induktionsannahme

$$ka_k = \frac{1}{(k-1)!} f'^{(k-1)}(a) = \frac{1}{(k-1)!} f^{(k)}(a), \quad a_k = \frac{1}{(k)!} f^{(k)}(a)$$

Gilt (ii) so setze $R(t) = f(a+t) - \sum_{k=0}^{n} a_k t^k$. Dann gilt für die m-te Ableitung, $m \leq n+1$:

$$f^{(m)}(a+t) = m! a_m + \sum_{k=m+1}^{n} ((k-m+1) \cdots k) a_k t^{k-m} + R^{(k)}(t)$$

$$R^{(m)}(0) = f^{(m)}(a+0) - m! a_m = 0 \quad \text{für } m \leq n$$

Mit $m = k+1$ folgt

$$f^{(n+1)}(a+t) = R^{(n+1)}(t) \text{ für alle } t \in]0, \delta[$$

Die Eindeutigkeit von $R(t)$ ergibt sich nun durch $n+1$-malige Anwendung des Eindeutigkeitssatzes. Man kann, wie Lagrange, $R(t)$ als Integral darstellen und dann ξ nach dem Mittelwertsatz der Integralrechnung wählen, um das Restglied wie oben zu erhalten. Das werden wir bei Gelegenheit tun. □

14.21 Anwendungen auf die Untersuchung von Funktionsgraphen

Wir sehen uns nun genauer an, wie sich für genügend oft differenzierbare Funktionen ihr Monotonie- und Krümmungsverhalten sowie lokale Extrema mit Hilfe der Differentialrechnung effektiv untersuchen lassen. Wir wissen bereits:

- f hat in x_0 ein lokales Extremum $\Rightarrow$ $f'(x_0) = 0$ (Satz 14.5)
- f monoton wachsend (fallend) $\Leftrightarrow$ $f' \geq 0$ ($f' \leq 0$) (Abschn.14.7)
- f konvex (konkav) $\Leftrightarrow$ $f'' \geq 0$ ($f'' \leq 0$). (Satz 4.17)

Für lokale Extrema haben wir bisher nur eine notwendige Bedingung angegeben. Wir ergänzen diese durch hinreichende Bedingungen. Anschaulich klar ist die folgende Bedingung.

Satz 14.6 *Sei f differenzierbar auf (a, b), $x_0 \in (a, b)$, $f'(x_0) = 0$, und f' wechsle in x_0 das Vorzeichen. Dann besitzt f ein lokales Maximum in x_0, wenn das Vorzeichen von $+$ nach $-$ wechselt für größer werdendes x, und ein lokales Minimum bei Wechsel von $-$ nach $+$.*

Satz 14.7 *Sei $n \geq 2$, f auf (a, b) n-mal stetig differenzierbar, und $x_0 \in (a, b)$. Weiter sei $f^{(k)}(x_0) = 0$ für $k = 1, 2, \ldots, n-1$ und $f^{(n)}(x_0) \neq 0$. Ist n gerade, so besitzt f in x_0 einen lokalen Extremwert, und zwar ein lokales Minimum, falls $f^{(n)}(x_0) > 0$ und ein lokales Maximum für $f^{(n)}(x_0) < 0$. Ist n ungerade, so liegt in x_0 kein Extremwert vor.*

Wir wollen uns dies für $n = 2$ klarmachen, d.h. wir zeigen unter den Voraussetzungen des Satzes

$$\begin{aligned} f'(x_0) = 0,\ f''(x_0) > 0 \quad &\Rightarrow \quad f \text{ hat lokales Minimum in } x_0, \\ f'(x_0) = 0,\ f''(x_0) < 0 \quad &\Rightarrow \quad f \text{ hat lokales Maximum in } x_0. \end{aligned}$$

Sei $x \in (a, b)$. Der Satz von Taylor liefert uns die Existenz eines $\xi \in (x_0, x)$ mit

$$f(x) - f(x_0) = \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(\xi)}{2!}(x - x_0)^2 = \frac{f''(\xi)}{2!}(x - x_0)^2 .$$

Sei beispielsweise $f''(x_0) > 0$. Wegen der Stetigkeit von f'' gibt es dann eine Umgebung $U \subseteq (a, b)$ von x_0, auf der f'' positiv ist. Für $x \in U$ ist auch $\xi \in U$, und wir erhalten

$$f(x) - f(x_0) = \frac{f''(\xi)}{2!}(x - x_0)^2 > 0 \quad \text{für alle } x \in U \backslash \{x_0\} .$$

Also besitzt f in x_0 ein (sogar echtes) lokales Minimum. Den Fall $f''(x_0) < 0$ behandelt man analog.

Ein Punkt $x_0 \in (a, b)$ heißt *Wendepunkt* von $f : (a, b) \to \mathbb{R}$, wenn es eine Umgebung $(x_0 - \varepsilon, x_0 + \varepsilon) \subseteq (a, b)$ von x_0 gibt, so dass f auf $(x_0 - \varepsilon, x_0)$ konvex (bzw. konkav) und auf $(x_0, x_0 + \varepsilon)$ konkav (bzw. konvex) ist.

Satz 14.8 (a) *Sei f in (a, b) zweimal stetig differenzierbar und $x_0 \in (a, b)$ ein Wendepunkt von f. Dann ist $f''(x_0) = 0$.*

(b) *Sei $n \geq 3$, f auf (a, b) n-mal stetig differenzierbar und $x_0 \in (a, b)$. Weiter sei $f^{(k)}(x_0) = 0$ für $k = 2, \ldots, n-1$ und $f^{(n)}(x_0) \neq 0$. Dann ist x_0 ein Wendepunkt für f, falls n ungerade ist, und kein Wendepunkt, falls n gerade ist.*

Die Begründung folgt wieder mit dem Satz von Taylor.

Beispiel 9 Wir betrachten die Funktion $f : [-1, 1] \to \mathbb{R}$, $x \mapsto x^2$. Für diese ist $f'(x) = 2x$ und $f''(x) = 2$ für alle $x \in (-1, 1)$. Nach Satz 14.7 hat f in $x_0 = 0$ ein lokales Minimum, und nach Satz 14.8 hat f keine Wendepunkte. Man beachte, dass wir in $(-1, 1)$ keine lokalen Maxima gefunden haben. Da die Funktion f aber stetig ist, muss sie auf $[-1, 1]$ globale Maxima und Minima besitzen. Wie wir gesehen haben, kann das globale Maximum nur in den Endpunkten ± 1 des Intervalles angenommen werden. Wegen $f(1) = f(-1) = 1$ ist klar: Die Funktion f nimmt in ± 1 ihr globales Maximum an und in 0 ihr globales Minimum. ■

14.22 Anwendung auf die Bestimmung von Grenzwerten

Die folgende *Regel von de l'Hospital* hilft bei der Bestimmung von Grenzwerten von Brüchen der Gestalt "0/0".

Satz 14.9 *Die Funktionen f und g seien auf (a, b) n-mal stetig differenzierbar, in $x_0 \in (a, b)$ gelte $f(x_0) = f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ sowie $g(x_0) = g'(x_0) = \ldots = g^{(n-1)}(x_0) = 0$ und $g^{(n)}(x_0) \neq 0$. Dann existiert der Grenzwert $\lim_{x \to x_0} \frac{f(x)}{g(x)}$, und dieser Grenzwert ist gleich $\frac{f^{(n)}(x_0)}{g^{(n)}(x_0)}$.*

Beweis Der Satz von Taylor liefert

$$\frac{f(x)}{g(x)} = \frac{0+0+\ldots+0+(x-x_0)^n f^{(n)}(\xi_x)/n!}{0+0+\ldots+0+(x-x_0)^n g^{(n)}(\eta_x)/n!} = \frac{f^{(n)}(\xi_x)}{g^{(n)}(\eta_x)}$$

mit gewissen Zahlen $\xi_x, \eta_x \in (x_0, x)$. (Man beachte, dass wegen $g^{(n)}(x_0) \neq 0$ und wegen der Stetigkeit von $g^{(n)}$ auch $g^{(n)}(\eta_x) \neq 0$ für alle η_x aus einer geeigneten Umgebung von x_0 und dass dann auch $g(x) \neq 0$ für alle $x \neq x_0$, die nahe genug an x_0 liegen.) Für $x \to x_0$ konvergieren ξ_x und η_x gegen x_0. Aus der Stetigkeit von $f^{(n)}$ und $g^{(n)}$ sowie aus $g^{(n)}(x_0) \neq 0$ folgt

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f^{(n)}(\xi_x)}{g^{(n)}(\eta_x)} = \frac{f^{(n)}(x_0)}{g^{(n)}(x_0)}. \qquad \blacksquare$$

Beispiel 10 Für $a > 0$ und $\alpha, \beta > 0$ ist

$$\lim_{x\to a} \frac{x^\alpha - a^\alpha}{x^\beta - a^\beta} = \lim_{x\to a} \frac{\alpha x^{\alpha-1}}{\beta x^{\beta-1}} = \frac{\alpha a^{\alpha-1}}{\beta a^{\beta-1}} = \frac{\alpha}{\beta} a^{\alpha-\beta}. \qquad \blacksquare$$

Beispiel 11 Die Regel von l'Hospital gilt auch, wenn unbestimmte Ausdrücke der Gestalt "∞/∞" vorliegen (vgl. Barner/Flohr, S. 276 – 277). So ist z.B.

$$\lim_{x\searrow 0} x \ln x = \lim_{x\searrow 0} \frac{\ln x}{\frac{1}{x}} = \lim_{x\searrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x\searrow 0} -x = 0.$$

In diesem Beispiel haben wir einen unbestimmten Ausdruck "$0 \cdot \infty$" in die Form "∞/∞" gebracht und darauf l'Hospital angewandt. Ähnlich lassen sich zahlreiche weitere Grenzwerte, die auf unbestimmte Ausdrücke wie "$\infty - \infty$" oder "1^∞" führen, berechnen. $\blacksquare$

14.25 Bernoulli-l'Hospital

Satz 14.10 *Seien $a > b$ in $\mathbb{R} \cup \{\infty, -\infty\}$ und f, g auf $]a, b[$ differenzierbar und $g'(x) \neq 0$ für alle $x \in]a, b[$. Gelte*

$$\lim_{a<x\to a} f(x) = 0 = \lim_{a<x\to a} g(x) \quad oder \quad \lim_{a<x\to a} g(x) = +\infty \quad oder \quad \lim_{a<x\to a} g(x) = -\infty$$

Dann folgt

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

falls der zweite Limes existiert. Das gilt auch für uneigentliche Limites. Entsprechendes gilt für $b > x \to b$ bzw. $x \to c$ mit f, g differenzierbar auf $]a, c[\cup]c, b[$.

Beweis im Falle eigentlicher Limites und differenzierbarer $f, g : [a, b] \to \mathbb{R}$ mit $g'(a) \neq 0$. Wegen der Stetigkeit von f, g gilt $f(a) = g(a) = 0$. Also

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

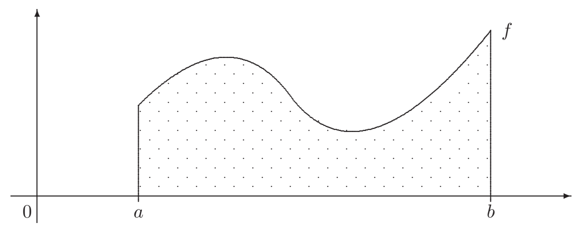
Allgemein: Siehe Heuser, Analysis I, §50. Beispiel

$$\lim_{x\to+\infty} \frac{e^{ax}}{x} = \lim_{x\to+\infty} \frac{ae^{ax}}{1} = +\infty$$

15 Integration

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

Flächenberechnungen Gegeben ist eine Funktion $f : [a, b] \to [0, \infty)$. Gesucht ist der Inhalt der von den Geraden $x = a, x = b, y = 0$ und vom Graphen der Funktion f berandeten Fläche. Genau genommen müssen wir dazu vorab die Frage klären, was wir unter dem Inhalt einer kompliziert geformten Fläche verstehen wollen. Es geht uns also nicht nur um eine Berechnungsvorschrift für Flächeninhalte, sondern auch um deren Definition.



Umkehrung des Differenzierens Kann man aus der Ableitung einer Funktion die Funktion selbst rekonstruieren? Gibt es für jede Funktion f eine Funktion F mit $F' = f$? Wenn ja, wie viele solcher Funktionen gibt es?

Zur Beantwortung dieser und anderer Fragen wurden verschiedene Integralbegriffe entwickelt, von denen wir einen der einfachsten - das Riemann-Integral - nun kennen lernen wollen. Eine nahe liegende Idee zur Berechnung des Inhalts eines komplizierten Gebietes (etwa des oben skizzierten) ist es, das Gebiet durch andere Gebiete „anzunähern", deren Flächenberechnung einfacher ist, etwa durch Gebiete, die aus endlich vielen Rechtecken zusammengesetzt sind. Wir hoffen, dass wir dem wahren Flächeninhalt näherkommen, wenn wir die Approximation „verfeinern", indem wir z.B. die Rechtecke schmaler machen. Diese vagen Vorstellungen wollen wir nun präzisieren. Dabei wollen wir auch die historische Entwicklung des Integralbegriffs und die aktuelle Sichtweise in den Ingenieurwissenschaften berücksichtigen.

15.1 Grundlegendes zur Integration

Wir betrachten nur beschränkte Funktionen $f : [a, b] \to \mathbb{R}$, also $m \leq f(x) \leq M$ für alle $x \in [a, b]$.

15.1.1 Zerlegungen und Treppenfunktionen

Ene *Zerlegung* Z des Intervalls $[a, b]$ wird gegeben durch $a = z_0 < z_1 < \ldots < z_{n-1} < z_n = b$ (mit beliebigem n). Dazu gehören die *Linienelemente* $(\Delta x)_k = z_{k+1} - z_k$. Die *Maschenweite* von Z ist $\min\{(\Delta x)_k \mid k < n\}$. (Man darf auch $z_k = z_{k+1}$ zulassen, aber das bringt nichts Neues und so passt's besser zur üblichen Vorstellung.)

Eine *Treppenfunktion* $f : [a,b] \to \mathbb{R}$ zur Zerlegung Z ist konstant auf jedem der offenen Intervalle $]z_k, z_{k+1}[$. Das Integral einer Treppenfunktion f ist definiert als

$$\int_a^b f(x)\mathrm{d}x = \sum_{k=0}^{n-1} f(\xi_k)(\Delta x)_k \quad \text{mit } \xi_k \in]z_k, z_{k+1}[$$

Lemma 15.1 *Das Integral einer Treppenfunktion hängt nicht von der gewählten Zerlegung ab.*

Beweis: Ist $z_h \leq w \leq z_{h+1}$ für ein h, so ergibt Einfügen von w eine Zerlegung für f und es ergibt sich dasselbe Integral, weil für $\xi \in]z_h, w[$ und $\xi' \in]w, z_{h+1}[$ gilt

$$f(\xi_h) = f(\xi) = f(\xi') \text{ also } f(\xi_h)(z_{h+1} - z_h) = f(\xi)(w - z_h) + f(\xi')(z_{h+1} - w).$$

Ist allgemeiner Z' mit $a = z'_0 < z'_1 \ldots < z'_m = b$ eine weitere Zerlegung für die Treppenfunktion f, so betrachte man die *gemeinsame Verfeinerung* Z'', die man dadurch erhält, dass man die Menge $\{z_0, \ldots, z_n, z'_0, \ldots, z'_m\}$ nach der Größe der Elemente anordnet. Durch Einzelschritte wie eben kommt man von Z wie von Z' zu Z'', also folgt die Gleichheit der das Integral definierenden Summen. $\Box$

15.1.2 Riemannsches Integral

Satz 15.2 *Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Für $c \in \mathbb{R}$ sind äquivalent*

(i) Es gibt Treppenfunktionen $\underline{f}_n, \overline{f}_n : [a,b] \to \mathbb{R}$ mit $\underline{f}_n \leq f \leq \overline{f}_n$ für alle n (o.B.d.A. $\underline{f}_{n-1} \leq \underline{f}_n \leq \overline{f}_n \leq \overline{f}_{n-1}$) und $\lim_{n\to\infty} \int_a^b \underline{f}_n(x)\mathrm{d}x = c = \lim_{n\to\infty} \int_a^b \overline{f}_n(x)\mathrm{d}x$

(ii) Für alle $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle Zerlegungen Z von $[a,b]$ mit Maschenweite $\leq \delta$ und alle $\xi_k \in]z_k, z_{k+1}[$ gilt: $|c - \sum_{k=0}^b f(\xi_k)(\Delta x)_k| \leq \varepsilon$

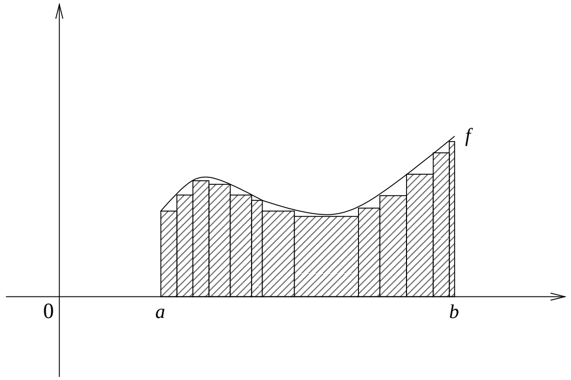
(iii) Für alle $\varepsilon > 0$ gibt es eine Zerlegung Z von $[a,b]$ so, dass für alle $\xi_k \in]z_k, z_{k+1}[$ gilt: $|c - \sum_{k=0}^b f(\xi_k)(\Delta x)_k| \leq \varepsilon$

Es gibt höchstens ein c, das (iii) erfüllt und es gilt $m(b-a) \leq c \leq M(b-a)$ falls $m \leq f(x) \leq M$ für alle $x \in [a,b]$.

Gilt (i), so heißt $f : [a,b] \to \mathbb{R}$ (*Riemann*) *integrierbar* und $c = \int_a^b f(x)\mathrm{d}x$ das *Integral* von f auf dem Intervall $[a,b]$. (ii) und (iii) geben Anlass, für eine Zerlegung Z und *Zwischenvektor* $\xi = (\xi_k)$ mit $\xi_k \in]z_k, z_{k+1}[$ die *Riemann-Summe* zu definieren als

$$R(Z, \xi, f) = \sum_{k=0}^{n} f(\xi_k)(\Delta x)_k$$

Beschränktes $f : [a,b] \to \mathbb{R}$ ist integrierbar genau dann, wenn es Treppenfunktionen $\underline{f}_n \leq f \leq \overline{f}_n$ gibt mit $\int_a^b \overline{f}_n(x)\mathrm{d}x - \int_a^b \underline{f}_n(x)\mathrm{d}x \to 0$ für $n \to \infty$.



Beweis. (iii) folgt trival aus (ii). Um $(iii) \Rightarrow (i)$ zu zeigen, sei $\varepsilon = \frac{1}{n}$ und Z die zugehörige Zerlegung. sei $\eta = \varepsilon/(b-a)$. Da f auf $]z_k, z_{k+1}[$ nach oben und unten beschränkt ist, gibt es (dank Archimedes) $\overline{\xi}_k, \underline{\xi}_k \in]z_k, z_{k+1}[$ so, dass

$$f(\underline{\xi}_k) - \eta \leq f(x) \leq f(\overline{\xi}_k) + \eta \quad \text{für alle } x \in]x_k, x_{k+1}[.$$

Definiere Treppenfunktionen

$$\begin{aligned} \overline{f}_n(x) &= f(\overline{\xi}_k) + \eta \quad &\text{für } z_k < x < z_{k+1} \\ \underline{f}_n(x) &= f(\underline{\xi}_k) - \eta \quad &\text{für } z_k < x < z_{k+1} \end{aligned}$$

und $\overline{f}_n(z_k) = \underline{f}_n(z_k) = f(z_k)$. Dann folgt $\underline{f}_n \leq f \leq \overline{f}_n$ und

$$\sum_{k=0}^{n-1} f(\underline{\xi}_k)(\Delta x)_k - \int_a^b \underline{f}_n(x)\mathrm{d}x \leq \varepsilon$$

also mit Voraussetzung (iii)

$$|c - \int_a^b \underline{f}_n(x)\mathrm{d}x| \leq 2\varepsilon.$$

Es folgt

$$c = \lim_{n\to\infty} \int_a^b \underline{f}_n(x)\mathrm{d}x$$

und entsprechend $c = \lim_{n\to\infty} \int_a^b \overline{f}_n(x)\mathrm{d}x$. Insbesondere ist c durch f eindeutig bestimmt. Indem man $\underline{f}_n$ durch $\max\{\underline{f}_n, \underline{f}_{n-1}\}$ und $\overline{f}_n$ durch $\min\{\overline{f}_n, \overline{f}_{n-1}\}$ ersetzt, kann man $\underline{f}_{n-1} \leq \underline{f}_n \leq \overline{f}_n \leq \overline{f}_{n-1}$ erreichen.

Sei nun (i) angenommen. Wegen der Beschränktheit von f gibt es m, M mit $m \leq f(x) \leq M$ für alle $x \in [a,b]$. Wir dürfen also auch $m \leq \underline{f}_n(x) \leq \overline{f}_n(x) \leq M$ annehmen (indem wir zu $\max\{m, \underline{f}_n(x)\}$ und $\min\{M, \overline{f}_n(x)\}$ übergehen). Sei $\varepsilon > 0$ gegeben. Nach Voraussetzung gibt es ein n_0 so, dass

$$\int_a^b \overline{f}_{n_0}(x)\mathrm{d}x - \int_a^b \underline{f}_{n_0}(x)\mathrm{d}x \quad \leq \frac{\varepsilon}{3}.$$

Sei N größer als die Anzahl der Teilpunkte in den gewählten Zerlegungen zu $\underline{f}_{n_0}$ bzw. $\overline{f}_{n_0}$. Wähle $\delta > 0$ mit

$$\delta(M-m)N \leq \frac{\varepsilon}{3}$$

Seien nun eine Zerlegung Z von Maschenweite $\leq \delta$ und $\xi_k \in]z_k, z_{k+1}[$ gegeben. Definiere die Treppenfunktion g durch

$$g(x) = \begin{cases} f(\xi_k) & \text{für } z_k < x < z_{k+1} \\ f(z_k) & \text{für } x = z_k \end{cases}$$

Dann gilt

$$\int_a^b \underline{f}_{n_0}(x)\mathrm{d}x - \frac{\varepsilon}{3} \leq \int_a^b g(x)\mathrm{d}x \leq \int_a^b \overline{f}_{n_0}(x)\mathrm{d}x + \frac{\varepsilon}{3}.$$

Ist nämlich ein Linienelement $]z_k, z_{k+1}[$ von Z ganz in einem von $\underline{f}_{n_0}$ enthalten, so gilt auf diesem $\underline{f}_{n_0}(x) \leq f(x)$. Andererseits gibt es höchstens N Linienelemente von Z, die nicht ganz in einem Linienlement von $\underline{f}_{n_0}$ enthalten sind und für jedes solche gilt

$$|g(\xi_k)(\Delta x)_k - \underline{f}_{n_0}(\xi_k)(\Delta x)_k| \leq \delta(M-m).$$

Entsprechendes gilt für $\overline{f}_{n_0}$. Es folgt, wie in (ii) behauptet,

$$|c - \int_a^b g(x)\mathrm{d}x| \leq \varepsilon.$$

Hat man Treppenfunktionen wie im Kasten, so hat man Treppenfunktionen

$$\underline{g}_n(x) = \max\{\underline{f}_k(x) \mid k \leq n\} \leq f(x) \leq \overline{g}_n(x) = \min\{\overline{f}_n(x) \mid k \leq n\}$$

und somit Intervallschachtelung

$$[\int_a^b \underline{g}_n(x)\mathrm{d}x,\ \int_a^n \overline{g}_n(x)\mathrm{d}x]$$

und diese bestimmt $c = \int_a^b f(x)\mathrm{d}x$. □

Beispiel 1 Sei $0 \leq a < b$. Auf dem Intervall $[a,b]$ betrachten wir die Funktion $f : x \mapsto x^2$. Wir wählen eine gleichmäßige Zerlegung $Z:\ z_0 < z_1 <, \ldots, < z_n$ mit

$$z_k = a + k\Delta x \quad \text{für } k = 0,1,\ldots,n,\ \Delta x := \frac{b-a}{n}$$

also $z_k - z_{k-1} = \Delta x$. Da f auf $[a,b]$ monoton wächst, ist

$$m_k = \min\{f(x) \mid x_k \leq x \leq x_{k+1}\} = f(z_k) = x_k^2$$

$$M_k = \max\{f(x) \mid x_k \leq x \leq x_{k+1}\} = f(z_{k+1}) = z_{k+1}^2,$$

und wir definieren

$$\underline{f}_n(x) = m_k \text{ für } z_k \leq x < z_{k+1},\ \ \overline{f}_n(x) = M_k \text{ für } z_k < x \leq z_{k+1}$$

Also

$$\int_a^b \underline{f}_n(x)\mathrm{d}x = (\Delta x)\sum_{k=0}^{n-1} z_k^2 \quad \text{und} \quad \int_a^b \overline{f}_n(x)\mathrm{d}x = (\Delta x)\sum_{k=1}^{n} z_k^2.$$

Also ist

$$\int_a^b \overline{f}_n(x)\mathrm{d}x - \int_a^b \underline{f}_n(x)\mathrm{d}x = (\Delta x)(z_n^2 - z_0^2) = (\Delta x)(b^2 - a^2) = \frac{(b-a)(b^2-a^2)}{n}.$$

und somit f auf $[a,b]$ integrierbar. ■

Beispiel 2 Für die durch

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

definierte *Dirichlet-Funktion* auf $[a,b]$ ist jede Obersumme gleich $b-a$ und jede Untersumme gleich 0. Also ist f nicht Riemann-integrierbar. ■

Wie in Beispiel 1 beweist man

Satz 15.3 *Jede monotone Funktion $f : [a,b] \to \mathbb{R}$ ist integrierbar.*

15.1.3 Eigenschaften des Riemannintegrals

Satz 15.4 *Seien $f, g : [a,b] \to \mathbb{R}$ Riemann-integrierbar.*

(a) *Für jedes $c \in \mathbb{R}$ ist cf Riemann-integrierbar, und es gilt*

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

(b) *Die Summe $f+g$ ist Riemann-integrierbar, und es gilt*

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

(c) *Das Produkt fg ist Riemann-integrierbar.*

Die Aussagen (a) und (b) kann man am einfachsten so einsehen: Ist Z eine Zerlegung und ξ ein entsprechender Zwischenvektor, so gilt

$$R(Z,\xi,cf+g) = cR(Z,\xi,f) + R(Z,\xi,g)\,.$$

Der Beweis von (c) ist etwas schwieriger. ■

Satz 15.5 *Seien $f, g : [a,b] \to \mathbb{R}$ Riemann-integrierbar.*

(a) *Ist $f(x) \leq g(x)$ für alle $x \in [a,b]$, so ist*

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

(b) *Die Funktion* $|f| : [a,b] \to \mathbb{R}$, $x \mapsto |f(x)|$, *ist Riemann-integrierbar, und*

$$\Big|\int_a^b f(x)dx\Big| \le \int_a^b |f(x)|dx\,.$$

Die Aussage in (*b*) heißt auch die *Dreiecksungleichung* für Integrale. Beachten Sie die Analogie zu den bekannten Ungleichungen

$$|a_1 + a_2| \le |a_1| + |a_2| \quad \text{und} \quad \Big|\sum_{i=1}^n a_i\Big| \le \sum_{i=1}^n |a_i|\,.$$

Zum Beweis Aussage (a) ist klar, denn für jede Riemannsumme gilt

$$R(Z,\xi,f) \le R(Z,\xi,g)\,.$$

Schwieriger ist der Beweis, dass $|f|$ Riemann-integrierbar ist, falls f diese Eigenschaft besitzt. Wenn man aber davon ausgeht, dass $|f|$ Riemann-integrierbar ist, wird der Beweis von (*b*) einfach: Aus $-|f| \le f \le |f|$ folgt nämlich mit Teil (a)

$$-\int_a^b |f(x)|dx \le \int_a^b f(x)dx \le \int_a^b |f(x)|dx\,,$$

und das ist die Behauptung. ■

Zur Integration über Teilintervalle gibt folgender Satz Auskunft.

Satz 15.6 (a) *Ist* $f : [a,b] \to \mathbb{R}$ *Riemann-integrierbar und* $a \le c < d \le b$, *so ist* f *auch auf* $[c,d]$ *Riemann-integrierbar.*

(b) *Sei* $f : [a,b] \to \mathbb{R}$ *und* $a < c < b$. *Ist* f *auf* $[a,c]$ *und auf* $[c,b]$ *Riemann-integrierbar, so ist* f *auch auf* $[a,b]$ *Riemann-integrierbar, und es gilt*

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

Beweis. (b) ist klar mit (i) aus Satz 15.2. Bei (a) können wir o.B.d.A. annehmen, dass $f \ge 0$ (nach Addition von $m \le f$) und $d = b$. Seien $\underline{f}_n$ und $\overline{f}_n$ für f gemäß (i) in Satz 15.2 gegeben. Seien g und h die Einschränkungen von f auf $[a,c]$ bzw. $[c,b]$ und entsprechend $\underline{g}_n$ usw. Dann

$$\lim_{n\to\infty}(\int_a^c \underline{g}_n + \int_c^b \underline{h}_n) = \int_a^b f = \lim_{n\to\infty}(\int_a^c \overline{g}_n + \int_c^b \overline{h}_n)$$

also

$$\lim_{n\to\infty}[(\int_a^c \overline{g}_n - \int_a^c \underline{g}_n) + (\int_c^b \overline{h}_n - \int_c^b \underline{h}_n)] = 0$$

also wegen der Nichtnegativität der Summanden

$$\lim_{n\to\infty}(\int_a^c \overline{g}_n - \int_a^c \underline{g}_n) = 0 = \lim_{n\to\infty}(\int_c^b \overline{h}_n - \int_c^b \underline{h}_n) \quad \square$$

Bisher haben wir das Integral $\int_a^b f(x)dx$ für $a < b$ definiert. Mitunter sind die folgenden *Vereinbarungen* nützlich:

(a)
$$\int_a^a f(x)dx := 0\,.$$

(b) Ist $f : [a,b] \to \mathbb{R}$ integrierbar auf $[a,b]$ (mit $a < b$), so sei

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

Mit diesen Vereinbarungen gilt:

Ist $f : [a,b] \to \mathbb{R}$ *Riemann-integrierbar, so ist für beliebige Punkte* $\alpha, \beta, \gamma \in [a,b]$

$$\int_\alpha^\beta f(x)dx + \int_\beta^\gamma f(x)dx = \int_\alpha^\gamma f(x)dx\,.$$

und die Funktion $F(u) = \int_c^u f(x)dx$ ist auf $[a,b]$ stetig - sogar *Lipshitz-stetig* nämlich

$$|F(u + \Delta u) - F(u)| \le |\Delta u|(M - m) \text{ falls } m \le f(x) \le M$$

15.2 Integration stetiger Funktionen

Bei der Definition der Stetigkeit von $f : D \to \mathbb{R}$] wird zu jedem $p \in D$ und $\varepsilon > 0$ ein $\delta(p,\varepsilon) > 0$ verlangt so, dass $|f(x) - f(p)| \le \varepsilon$ falls $|x - p| \le \delta(p,\varepsilon)$. Diese Definition ist 'lokal'. Das Beispiel der Funktion $f(x) = \frac{1}{x}$ auf $]0,1]$, zeigt, dass i.A. δ nicht unabhängig von p gewählt werden kann. Es soll gezeigt werde, dass dies aber der Fall ist, wenn $D = [a,b]$ ein abgeschlossenes Intervall ist. Das ist eine wesentliche Voraussetzung, um die Rolle von Integralen in Natur- und Technikwissenschaften diskutieren zu können.

Lemma 15.7 *Ist* $c \in [a,b]$ *und sind für* $f : [a,b] \to V$ *die Einschränkungen* $f|[a,c]$ *und* $f|[c,b]$ *stetig, so ist auch* $f : [a,b] \to V$ *stetig.*

Beweis. Für $p \ne c$ gehört jede Folge $x_n \to p$ ab einer Stelle zu $[a,c]$ oder $[c,b]$, also $\lim_{n\to\infty} f(x_n) = f(p)$. An der Stelle c hat man nach Voraussetzung rechts- wie linksseitigen Limes $f(c)$. Die Behauptung folgt nun mit Kor.13.18. □

15.2.1 Kompaktheit

Gegeben eine Abbildung $\delta : \mathbb{R} \to \mathbb{R}_{\ge 0}$, sei U_x das offene Intervall $]x - \delta(x),\, x + \delta(x)[$ (d.h. leer, falls $\delta(x) = 0$). Eine Teilmenge D von $\mathbb{R}$ heisse *kompakt* für δ, wenn es entweder ein $c \in D$ gibt, so dass $c \notin U_x$ für alle $x \in \mathbb{R}$ oder es endlich viele $x_1, \ldots, x_n \in \mathbb{R}$ gibt so, dass

$$D \subseteq U_{x_1} \cup \ldots \cup U_{x_n}$$

Dabei o.B.d.A. $\delta(x_i) > 0$.

Satz 15.8 *Ein abgeschlossenes Intervall* $[a,b]$ *ist kompakt bzgl. jeder Funktion* δ.

Beweis: Sind $[a,c]$ und $[c,b]$ kompakt bzgl. δ, so offensichtlich auch $[a,b]$. Sei nun angenommen, dass $[a,b]$ nicht kompakt bzgl. δ. Durch fortlaufende Halbierung bekommen wir eine Intervallschachtelung von Intervallen $[a_n,b_n] \subseteq [a,b]$, die bzgl. δ nicht kompakt sind. Sei c die dadurch bestimmte reelle Zahl und $x \in [a,b]$ mit $c \in U_x$ (gäbe es kein solches x, wäre $[a,b]$ trivialerweise kompakt für δ). Sei $\varepsilon = \min\{c - (x - \delta(x),\, x + \delta(x) - c\}$. Dann gibt es ein n mit

$$x - \delta(x) \le c - \varepsilon < a_n \le c \le b_n < c + \varepsilon \le x + \delta(x)$$

also $[a_n,b_n]$ kompakt bzgl. δ, ein Widerspruch zur Annahme. □

15.2.2 Gleichmäßige Stetigkeit

Eine Funktion $f : D \to \mathbb{R}$ heisst *gleichmäßig stetig* auf D, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so gibt, dass

$$\text{für alle } x, x' \in D: \; |x - x'| \le \delta \;\Rightarrow\; |f(x) - f(x')| \le \varepsilon$$

Satz 15.9 *Ist $f : [a,b] \to \mathbb{R}$ stetig, so auch gleichmäßig stetig.*

Beweis. Sei $\varepsilon > 0$ gegeben. Nach Voraussetzung gibt es zu jedem $p \in [a,b]$ ein $\delta(p) > 0$ so, dass $|f(p) - f(x)| \le \frac{\varepsilon}{2}$ für alle $x \in [a,b]$ mit $|x - p| \le \delta(p)$. Setze $\delta(x) = 0$ für $x \notin [a,b]$. Wegen der Kompaktheit kann man $[a,b]$ durch endlich viele (nichtleere) U_x überdecken. Wenn keine Lücke bleiben soll, müssen sich diese hinreichend überlappen, d.h. es gibt $x_1 < \ldots < x_n \in [a,b]$ mit $x_i + \delta(x_i) > x_{i+1} - \delta(x_{i+1})$. Setze

$$\delta = \frac{1}{2}\min\{x_i + \delta(x_i) - (x_{i+1} - \delta(x_{i+1})) \mid i = 1, \ldots n-1\}.$$

Ist nun $|x - x'| \le \delta$, so $x, x' \in\,]x_i - \delta(x_i),\, x_i + \delta(x_i)[$ für ein $i = 1, \ldots, n$, also

$$|f(x) - f(x')| \le |f(x) - f(x_i)| + |f(x_i) - f(x')| \le 2\frac{\varepsilon}{2} = \varepsilon \quad \square$$

15.2.3 Integrierbarkeit stetiger Funktionen

Satz 15.10 *Jede stetige Funktion $f : [a,b] \to \mathbb{R}$ ist integrierbar und es gilt*

(i) Zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ so, dass für alle $a \le c \le d \le b$ mit $d - c \le \delta$ und alle $\xi \in [c,d]$ gilt

$$\left|\int_c^d f(x)\mathrm{d}x - f(\xi)(d-c)\right| \le \varepsilon(d-c)$$

(ii) Mittelwertsatz*: Es gibt $\xi \in [a,b]$ mit $\int_a^b f(x)\mathrm{d}x = f(\xi)(b-a)$.*

Beweis. Wegen der gleichmäßigen Stetigkeit von f kann man zu jeden $\varepsilon > 0$ ein $\delta > 0$ wählen so, dass $|f(x) - f(p)| \le \frac{\varepsilon}{2}$ falls $|x - p| \le \delta$. Sei $\varepsilon = \frac{1}{n}$. Wähle eine Zerlegung $a = z_0 < z_1 < \ldots < z_m = b$ so, dass $z_{k+1} - z_k \le \delta$ für alle k und beliebige $\xi_k \in [z_k, z_{k+1}]$. Definiere Treppenfunktionen durch

$$\begin{array}{ll} \overline{f}_n(x) = f(\xi_k) + \frac{\varepsilon}{2} & \text{für } z_k < x < z_{k+1} \\ \underline{f}_n(x) = f(\xi_k) - \frac{\varepsilon}{2} & \text{für } z_k < x < z_{k+1} \end{array}$$

und $\overline{f}_n(z_k) = \underline{f}_n(z_k) = f(z_k)$. Dann folgt $\underline{f}_n \le f \le \overline{f}_n$ und

$$\int_a^b \overline{f}_n \mathrm{d}x - \int_a^b \underline{f}_n \mathrm{d}x \le \varepsilon(b-a) = \frac{b-a}{n}.$$

Also $\lim_{n\to\infty} \int_a^b \overline{f}_n \mathrm{d}x = \lim_{n\to\infty} \int_a^b \underline{f}_n \mathrm{d}x$, d.h. f ist integrierbar. Sei schließlich $0 < d - c \le \delta$ und $\xi \in [c,d]$. Die Einschränkung g von f auf $[c,d]$ ist integrierbar. Definiere auf $[c,d]$ die Treppenfunktionen

$$\begin{array}{ll} \overline{g}(x) = f(\xi) + \frac{\varepsilon}{2} & \text{für } c < x < d \\ \underline{g}(x) = f(\xi) - \frac{\varepsilon}{2} & \text{für } c < x < d \end{array}$$

und $\overline{g}(x) = \underline{g}(x) = g(x)$ für $x = c, d$. Dann $\underline{g} \le g \le \overline{g}$ und

$$f(\xi)(d-c) - \frac{\varepsilon(d-c)}{2} = \int_c^d \overline{g}\mathrm{d}x \le \int_c^d g(x)\mathrm{d}x \le \int_c^d \underline{g}\mathrm{d}x = f(\xi)(d-c) + \frac{\varepsilon(d-c)}{2}$$

und es folgt die Behauptung (i). (ii) ergibt sich sofort mit dem Zwischenwertsatz aus der Stetigkeit von $f(x)(b-a)$ und $(b-a)\min f \le \int_a^b f(x)\mathrm{d}x \le (b-a)\max f$. □.

Wir vermerken noch eine nützliche Verallgemeinerung von Satz 15.10 (ii).

Satz 15.11 Verallgemeinerter Mittelwertsatz *Seien $f, g : [a,b] \to \mathbb{R}$ Riemann-integrierbar und $g \ge 0$ auf $[a,b]$. Dann gibt es ein $\eta \in [m,M]$ so, dass*

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

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

15.2.4 Summationstheorem

Theorem 15.12 *Sei $W(a,b) \in \mathbb{R}$ für $a \le b$ im offenen Intervall $I \subseteq \mathbb{R}$ definiert und gelte die* Additivität

$$W(a,b) = W(a,c) + W(c,b) \quad \textit{falls } c \textit{ zwischen } a \textit{ und } b$$

Sei $f : I \to \mathbb{R}$ auf I stetig und gelte

(∗) Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass für alle Δx mit $|\Delta x| \le \delta$ und alle $p \in I$ gilt:

- *Es gibt $\xi \in [p,\, p + \Delta x]$ mit $|W(p,\, p + \Delta x) - f(\xi)\Delta x| \le \varepsilon|\Delta x|$*

Dann gilt für alle $a \le b$ in I: $W(a,b) = \int_a^b f(x)\mathrm{d}x$.

Bei der ursprünglichen Auffassung von Integralen ist $\int_a^b f(x)\mathrm{d}x$ die reelle Zahl, die von der Summe $\sum_a^b f(x)\mathrm{d}x$ mit infinitesimalen $\mathrm{d}x$ nur infinitesimalen Abstand hat. Die Voraussetzung (∗) des Theorems besagt dann, dass $\alpha = W(p,\, p + \mathrm{d}x) - f(\xi)\mathrm{d}x$ infinitesimal von zweiter Ordung ist, d.h. dass $\frac{\alpha}{\mathrm{d}x}$ infinitesimal ist. In dieser Form war das Theorem die Grundlage für die Einführung physikalischer Größen als Integrale. In den Natur- und Technikwissenschaften wird das heute so formuliert:

Von der Summe $W = \sum_k f(\xi_k)(\Delta x)_k$ mit den Linienelementen $(\Delta x)_k$ geht man mit $(\Delta x)_k \to 0$ zum Integral $W = \int f(x)\mathrm{d}x$ über (unter Bezugnahme auf den Fall stückweise konstanter Funktionen).

Ist die Stetigkeit von $f(x)$ gegeben, so handelt es sich demnach um eineAnwendung des Theorems. In der gängigen mathematischen Literatur, wird bestenfalls bei der Definition einer Größe durch ein Integral eine Motivation im Sinne des Theorems gegeben. Die infinitesimale Form findet man in Lehrbüchern zur Nonstandard-Analysis. Das Theorem verallgemeinert sich in offensichtlicher Weise auf Bereichsintegrale und mit mehr Anstrengung auf Kurven- und Flächenintegrale - das Problem ist die Zuordnung von geradlinigen Strecken- bzw. Flächenelementen zu gekrümmten Kurven- bzw. Flächenstücken.

Beweis. Sei $a \geq b$ und $\varepsilon > 0$ gegeben. Wähle δ gemäß $(*)$ und nach dem vorangehenden Satz so, dass

$$|\int_p^{p+\Delta x} f(x)\mathrm{d}x - f(\xi)|\Delta x| \leq \varepsilon\Delta x| \quad \text{für alle } p \in [a,b],\ |\Delta| \leq \delta \text{ und passendes } \xi \in [p,\, p+\Delta x].$$

Wähle ein Zerlegung $a = z_0 < z_1 \ldots < z_n = b$ von Maschenweite $\leq \delta$ und $\xi_k \in [z_k, z_{k+1}]$ so, dass $|W(z_k, z_{k+1}) - f(\xi_k)(z_{k+1} - z_k)| \leq \varepsilon(z_{k+1} - z_k)$. Dann folgt

$$|W(z_k, z_{k+1}) - \int_{z_k}^{z_{k+1}} f(x)\mathrm{d}x| \leq$$

$$\leq |W(z_k, z_{k+1}) - f(\xi_k)(z_{k+1} - z_k)| + |f(\xi_k)(z_{k+1} - z_k) - \int_{z_k}^{z_{k+1}} f(x)\mathrm{d}x| \leq 2\varepsilon(z_{k+1} - z_k)$$

also mit der Additiviät und Dreiecksungleichung

$$|W(a,b) - \int_a^b f(x)\mathrm{d}x| \leq \sum_{k=0}^{n-1} |W(z_k, z_{k+1}) - \int_{z_k}^{z_{k+1}} f(x)\mathrm{d}x| \leq 2\varepsilon(b-a)$$

Dies gilt für alle $\varepsilon > 0$, also nach Archimedes $W(a,b) = \int_a^b f(x)\mathrm{d}x$. □

15.2.5 Integration vektorwertiger Funktionen

Die vektorwertige Funktion $\vec{y} = \vec{y}(t) \in V$ sei auf dem Zeitintervall $I \subseteq \mathbb{R}$ definiert. Weiterhin sei für alle $a \leq b$ in I $\vec{w}(a,b)$ definiert so, dass gilt

- $\vec{w}(a,b) = \vec{w}(a,c) + \vec{w}(c,b)$ für alle $a \leq c \leq b$ (Additivität)

$(*)$ Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass für alle Δt mit $|\Delta t| \leq \delta$ und alle $p \in I$ gilt:

  – Es gibt $\xi \in [p,\, p+\Delta t]$ mit $\|\vec{w}(p,\, p+\Delta t) - \vec{y}(\xi)\Delta t\| \leq \varepsilon|\Delta t|$

Dann heisse $\vec{w}$ eine *Integration* für $\vec{y}$.

Satz 15.13 *Ist I beschränkt und abgeschlossen und $\vec{y}(t)$ stetig so gibt es eine eindeutig bestimmte Integration $\vec{w}$ zu $\vec{y}$. Bzgl. einer Orthonormalbasis $\vec{e}_1, \ldots, \vec{e}_m$ von V ist $\vec{w}$ gegeben durch*

$$\vec{w}(a,b) = \int_a^b \vec{y}(t)\mathrm{d}t := (\sum_{k=1}^m \int_a^b y_k(t)\mathrm{d}t)\vec{e}_k, \quad \textit{wobei } \vec{y}(t) = \sum_{k=1}^m y_k(t)\vec{e}_k$$

Beweis. Definiere

$$w_k(a,b) = \int_a^b y_k(y)\mathrm{d}t, \quad \vec{w}(a,b) = \sum_k w_k(a,b)\vec{e}_k$$

Dann folgt die Additiviät aus der Additiviät des Riemann-Intergrals. Da die $y_k(t)$ auf I stetig sind, gibt nach Satz 15.10 (i) es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so, dass für alle Δt mit $|\Delta t| \leq \delta$ und alle $p \in I$ und $\xi \in [p, p+\Delta t]$

$$|w_k(p,\, p+\Delta t) - y_k(\xi)\Delta t| \leq \frac{\varepsilon}{m}|\Delta t|$$

Mit der Dreicksungleichung folgt

$$\|\vec{w}(p,\, p+\Delta t) - \vec{y}(\xi)\Delta t\| \leq \varepsilon|\Delta t|$$

Also $\vec{w}$ ein Integral für $\vec{y}$. Ist umgekehrt $\vec{w}$ ein Integral für $\vec{y}$ und $\vec{w}(a,b) = \sum_k w_k(a,b)\vec{e}_k$, so folgt die Additivität der w_k aus der Eindeutigkeit der Vektorzelegung und man schließt von $(*)$ auf $|w_k(p,\, p+\Delta t) - y_k(\xi)\Delta t| \leq \varepsilon|\Delta t|$. Also $w_k(a,b) = \int_a^b y_k(t)dt$ nach dem Summationstheorem. □

Beispiel: Trägheitskompass. Sei $\vec{x}(t)$ der Ort der Masse M zur Zeit t. Die Geschwindigkeit zur Zeit t ist die erste Ableitung

$$\vec{v}(t) = \frac{\partial \vec{x}}{\partial t}(t)$$

Durch Integration erhält man

$$\vec{w}(a,b) = \vec{x}(b) - \vec{x}(a) = \int_a^b \vec{v}(t)\mathrm{d}t$$

d.h. man kann $\vec{x}(b)$ bestimmen, wenn man $\vec{v}(t)$ und den Anfangswert $\vec{x}(a)$ kennt

$$\vec{x}(b) = \vec{x}(a) + \int_a^b \vec{v}(t)\mathrm{d}t$$

Die Beschleunigung $\vec{a}(t)$ ist die Ableitung der Geschwindigkeit. Aus dieser und der Anfangsgeschwindigkeit $\vec{v}(a)$ erhält man entsprechend die Geschwindigkeit

$$\vec{v}(t) = \vec{v}(a) + \int_a^t \vec{a}(t)\mathrm{d}t$$

Also kann man $\vec{x}(b)$ aus $\vec{x}(a)$, $\vec{v}(a)$ und der Funktion $\vec{a}(t)$ bestimmen. Das war das Prinzip der Trägheitsnavigation.

Beispiel. Komplexwertige Funktionen. Ist $f: D \to \mathbb{C}$ stetig so gilt

$$\int f(t)\mathrm{d}t = \int f_1(t)\mathrm{d}t + j\int f_2(t)\mathrm{d}t\ + C, \quad C \in \mathbb{C} \quad \text{mit } f(t) = f_1(t) + jf_2(t),\ f_k(t) \in \mathbb{R}$$

15.2.6 Hauptsatz

Satz 15.14 *Ist $f : [a,b] \to \mathbb{R}$ stetig und $c \in [a,b]$, so ist die Funktion $F(\tau)$ auf $[a,b]$ differenzierbar, wobei*

$$F(\tau) = \int_c^\tau f(t)\mathrm{d}t \quad \textit{und} \quad \frac{\partial F}{\partial t}(\tau) = f(\tau) \quad \tau \in [a,b]$$

Beweis.

$$\Delta F(\tau, \Delta\tau) = F(\tau + \Delta\tau) - F(\tau) = \int_\tau^{\tau+\Delta\tau} f(t)\mathrm{d}t$$

Nach Satz 15.10 (i) gibt es also zu jedem $\varepsilon > 0$ ein $\delta > 0$ so, dass für alle $|\Delta\tau| \le \delta$ gilt

$$|F(\tau + \delta\tau) - F(\tau) - f(\tau)\Delta\tau| \le \varepsilon|\Delta\tau|$$

und es folgt

$$|\frac{F(\tau + \Delta\tau) - F(\tau)}{\Delta\tau} - f(\tau)| \le \varepsilon \quad \Box$$

Korollar 15.15 *Ist $f : [a,b] \to \mathbb{R}$ stetig, $c \in [a,b]$ und $G : [a,b] \to \mathbb{R}$ eine differenzierbare Funktion mit $\frac{\partial G}{\partial t}(\tau) = f(\tau)$ für alle $\tau \in [a,b]$, so gibt es eine Konstante C mit*

$$G(\tau) = \int_c^\tau f(t)\mathrm{d}t + C \quad \textit{für alle } \tau \in [a,b]$$

Insbesondere

$$C = G(c), \quad \int_c^\tau f(t)\mathrm{d}t = G(\tau) - G(c)$$

$$G(c) = 0 \ \Leftrightarrow\ G(\tau) = \int_c^\tau f(t)\mathrm{d}t \quad \textit{für alle } \tau \in [a,b]$$

Beweis: Eindeutigkeitssatz. $\Box$.

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

15.3 Einige Integrationstechniken

Die Hauptsätze der Differential- und Integralrechnung reduzieren die Berechnung eines Riemann-Integrals über eine Funktion f auf die Bestimmung einer Stammfunktion von f. Wir lernen nun einige Aussagen kennen, die diese Aufgabe erleichtern. Allerdings bleibt die Bestimmung einer Stammfunktion (im Gegensatz zur „umgekehrten“ Aufgabe, der Bestimmung einer Ableitung) ein schwieriges Problem. Im Gegensatz zu den Regeln der Differentiation, die uns die Differentiation beliebig komplizierter Funktionen ermöglichen, führen die Integrationsregeln, die wir uns nun ansehen werden, nicht immer zum Ziel. Mehr noch: bereits für so einfache Funktionen wie $x \mapsto 1/\ln x$ und $x \mapsto e^{-x^2}$, die nach Satz 15.13 eine Stammfunktion auf $(1,\infty)$ bzw. auf $\mathbb{R}$ besitzen, ist es *nicht* möglich, diese Stammfunktion mit Hilfe endlicher Ausdrücke von „elementaren“ Funktionen (wie Potenz-, Exponential- und Winkelfunktionen sowie deren Umkehrfunktionen) zu beschreiben.

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

15.3.1 Stammfunktionen

Definition 15.16 *Sind $F, f : [a,b] \to \mathbb{R}$ Funktionen, und ist F differenzierbar auf $[a,b]$ mit $F'(x) = f(x)$ für alle $x \in [a,b]$, so heißt F eine* Stammfunktion *von f.*

Satz 15.17 (a) *Ist $F : [a,b] \to \mathbb{R}$ Stammfunktion von $f : [a,b] \to \mathbb{R}$ und $C \in \mathbb{R}$, so ist auch $F + C$ eine Stammfunktion von f.*

(b) *Je zwei Stammfunktionen von $f : [a,b] \to \mathbb{R}$ unterscheiden sich nur um eine Konstante.*

Beweis Aussage (a) ist klar. Für den Beweis von Aussage (b) seien F_1, F_2 Stammfunktionen einer Funktion f auf $[a,b]$, d.h. es ist $F_1' = F_2' = f$. Dann ist $(F_1 - F_2)' = F_1' - F_2' = 0$, und dem Eindeutigkeitssatz ist die Funktion $F_1 - F_2$ konstant. ■

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

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

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

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

$$\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C \quad \text{auf} \quad \begin{cases} \mathbb{R} & \text{falls} \quad \alpha = 0,1,2,\ldots \\ \mathbb{R}\backslash\{0\} & \text{falls} \quad \alpha = -2,-3,-4,\ldots \\ (0,\infty) & \text{falls} \quad \alpha \in \mathbb{R}\backslash\{-1\}, \end{cases}$$

$$\int x^{-1}dx = \ln|x| + C \quad \text{auf} \quad \mathbb{R}\backslash\{0\},$$

$$\int e^x dx = e^x + C \quad \text{auf} \ \mathbb{R},$$

$$\int \sin x\, dx = -\cos x + C, \quad \int \cos x\, dx = \sin x + C \quad \text{auf} \ \mathbb{R},$$

$$\int \sinh x\, dx = \cosh x + C, \quad \int \cosh x\, dx = \sinh x + C \quad \text{auf} \ \mathbb{R},$$

$$\int \frac{1}{1+x^2}\, dx = \arctan x + C \quad \text{auf} \ \mathbb{R},$$

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin x + C \quad \text{auf } (-1,1).$$

Dabei

$$\sinh x = \frac{1}{2}(e^x - e^{-x}), \quad \cosh x = \frac{1}{2}(e^x + e^{-x})$$

Anmerkung 1 Das Korollar das Hauptsatzes gilt auch in der folgenden allgemeineren Form:

Ist $f : [a,b] \to \mathbb{R}$ Riemann-integrierbar, und besitzt f auf $[a,b]$ eine Stammfunktion F, so gilt $\int_a^b f(t\mathrm{d}t) = F(b) - F(a)$.

Allerdings gibt es Riemann-integrierbare Funktionen, die keine Stammfunktion besitzen, wie die Funktion

$$f : [-1,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} -1 & \text{für } x \in [-1,0) \\ 1 & \text{für } x \in [0,1] \end{cases}$$

zeigt. Diese ist Riemann-integrierbar, da sie auf $[-1,1]$ mit Ausnahme des Punktes $x=0$ stetig ist. Es gibt jedoch keine auf $[-1,1]$ differenzierbare Funktion F mit $F' = f$. Für $x \in [-1,0)$ müsste nämlich $F(x) = -x+c$ mit einem $c \in \mathbb{R}$ sein, und für $x \in (0,1]$ müsste $F(x) = x+d$ mit $d \in \mathbb{R}$ sein. Die Funktion F ist nur für $c = d$ stetig in $x = 0$. Dann stimmt $F(x)$ mit $|x| + c$ überein. Die Betragsfunktion ist aber in $x = 0$ nicht differenzierbar. ■

Anmerkung 2 Eine Funktion, die eine Stammfunktion besitzt, muss nicht Riemann-integrierbar sein. Beispielsweise ist die Funktion

$$F : [-1,1] \to \mathbb{R}, \quad x \mapsto \begin{cases} x^2 \cos\frac{1}{x^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$

differenzierbar auf $[-1,1]$, ihre Ableitung F' ist aber unbeschränkt und damit nicht Riemann-integrierbar. ■

Beispiele Sei $f(x) = x^2$ auf $[-1,2]$. Dann ist $F(x) = \frac{x^3}{3}$ eine Stammfunktion von f (denn es ist $F'(x) = \frac{3}{3}x^2 = x^2 = f(x)$), und demnach ist

$$\int_{-1}^{2} x^2 dx = F(2) - F(-1) = \frac{8}{3} - \frac{(-1)}{3} = 3\,.$$

Ganz ähnlich findet man

$$\int_0^{\pi} \sin x\,dx = -\cos x\big|_0^{\pi} = -(-1) - (-1) = 2\,,$$

wobei wir die Schreibweise $F(x)\big|_a^b := F(b) - F(a)$ benutzt haben. ■

15.3.2 Linearität des Integrals

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

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

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

$$\int \big(\alpha f(x) + \beta g(x)\big)dx = \alpha \int f(x)dx + \beta \int g(x)dx\,. \tag{15.1}$$

Satz 15.18 *Besitzen $f, g : [a,b] \to \mathbb{R}$ Stammfunktionen und sind $\alpha, \beta \in \mathbb{R}$, so besitzt auch $\alpha f + \beta g$ eine Stammfunktion auf $[a,b]$, und es gilt* (15.1).

Beispielsweise darf man Polynome gliedweise integrieren:

$$\int \sum_{j=0}^{n} a_j x^j dx = \sum_{j=0}^{n} a_j \int x^j dx = \sum_{j=0}^{n} a_j \frac{x^{j+1}}{j+1} + C\,.$$

15.3.3 Substitutionsregel

Lemma 15.19 *Sei $x = x(t)$ auf $[a,b]$ differenzierbar und $f(x)$ auf dem Wertebereich $[c,d]$ von x definiert. Ist $F(x)$ eine Stammfunktion von $f(x)$ so ist $F(x(t))$ eine Stammfunktion von $\phi(t) = f(x(t))\frac{\partial x}{\partial t}$, d.h.*

$$\int f(x)\mathrm{d}x = \int f(x(t))\frac{\partial x}{\partial t}\mathrm{d}t + C$$

Ist x stetig differenzierbar und $f(x)$ stetig, so gilt

$$\int_{x(a)}^{x(b)} f(x)\mathrm{d}x = \int_a^b f(x(t))\frac{\partial x}{\partial t}\mathrm{d}t$$

Beweis. Die erste Aussage folgt aus der Kettenregel. Nun mit dem Korollar des Hauptsatzes

$$\int_{x(a)}^{x(b)} f(x)\mathrm{d}x = F(x(b)) - F(x(a)) = \int_a^b f(x(t))\frac{\partial x}{\partial t}\mathrm{d}t \quad \square$$

15.3.4 Differentiale

Um die Regeln korrekt und in Übereinstimmung mit der Praxis in Natur- und Ingenieurwissenschaften formulieren zu können, bedienen wir uns der Differentiale. Für eine auf $[a,b]$ definierte und differenzierbare Funktion $y = f(t)$ ist das *Differential* an der Stelle $p \in [a,b]$ die homogen lineare Funktion

$$\mathrm{d}f(p,\mathrm{d}t) = \frac{\partial f}{\partial t}(p)\cdot \mathrm{d}t = \frac{\partial y}{\partial t}\mathrm{d}t \quad \mathrm{d}t \in p - a \leq \mathrm{d}t \leq b, \;\; \mathrm{d}t \neq 0$$

oder wenn man die Stelle p nicht explizit erwähnt

$$\mathrm{d}y = \frac{\partial y}{\partial t}\mathrm{d}t.$$

Das Differential an der Stelle p ist natürlich schon dann bekannt, wenn man es für ein einziges $\mathrm{d}t \neq 0$ kennt.

Seien nun $x = x(t)$ und $y = y(t)$ auf $[a, b]$ differenzierbar und sei y eine differenzierbare Funktion von x, also nach der Kettenregel

$$y(t) = y = y(x) = y(x(t)), \quad \frac{\partial y}{\partial t}(p) = \frac{\partial y}{\partial x}(x(p)) \cdot \frac{\partial x}{\partial t}(p)$$

$$\mathrm{d}y(p, \mathrm{d}t) = \frac{\partial y}{\partial t}(p) \cdot \mathrm{d}t = \frac{\partial y}{\partial x}(x(p)) \cdot \frac{\partial x}{\partial t}(p) \cdot \mathrm{d}t = \frac{\partial y}{\partial x}(x(p)) \cdot \mathrm{d}x(p, \mathrm{d}t)$$

d.h. wir können $\mathrm{d}y(p)$ auch als Differential bzgl. x verstehen. Dementsprechend haben wir die folgende *Konsistenzvoraussetzung* für den problemlosen Umgang mit Differentialen

♡ Alle betrachteten Größen sind stetig differenzierbare Funktionen einer vorgegebenen unabhängigen Variablen $t \in [a, b]$.

Dann gilt unzweideutig

$$\mathrm{d}y = \frac{\partial y}{\partial x}\mathrm{d}x$$

wie auch immer $y = y(x)$ differenzierbare Funktion von x, und

$$\frac{\partial y}{\partial x}(p) = \frac{\mathrm{d}y}{\mathrm{d}x}(p) \quad \text{falls } \frac{\partial x}{\partial t}(p) \neq 0$$

$$\frac{\partial y}{\partial x}(p) = 0 \quad \text{falls } \frac{\partial x}{\partial t}(p) = 0$$

Natürlich genügt es, wenn man das Differential $\mathrm{d}x(p)$ als eine Funktion versteht, die nur für sehr kleine $|\mathrm{d}t| \neq 0$ bzw. für infinitesimale $\mathrm{d}t$ definiert ist.

15.3.5 Integrationsregel

Der folgende Satz fasst die üblichen Integrationsregeln zusammen und zeigt, dass das sogenannte "formale Rechnen" legitim und sinnvoll ist, wenn die Konsistenzbedingung für Differentiale erfüllt ist. Wenn Sie sich vergewissert haben, dass die Konsistenzbedingung bei Ihrer Rechnung erfüllt ist, notieren Sie das einfach durch ein ♡. Der Vorteil dieser Rechnung ist die intuitive Notation und die Option, die Argumentwerte weitgehend zu unterdrücken (da diese über die Abhängigkeit von t gekoppelt sind). Merke: "Formal" ist das Unwort der Analysis.

Satz 15.20 *Seien $x = x(t)$, $y = y(t)$ und $z = z(t)$ auf $[a, b]$ differenzierbar und $f(x)$, $g(y)$, $h(z)$ stetige Funktionen auf den jeweiligen Wertebereichen. Für die Differentiale gelte*

$$f(x(p))\mathrm{d}x(p) = cg(y(p))\mathrm{d}y(p) + h(z(p))\mathrm{d}z(p) \text{ für alle } p \in [a, b]$$

$$\textit{kurz} \quad f(x)\mathrm{d}x = cg(y)\mathrm{d}y + h(z)\mathrm{d}z$$

Dann gilt: Für alle Stammfunktionen F, G, H von f, g bzw. h gibt es eine Konstante C mit

$$F(x(t)) = c\,G(y(t) + H(z(t)) \quad \text{für alle } t \in [a, b]$$

kurz

$$\int f(x)\mathrm{d}x = c\int g(y)\mathrm{d}y + \int h(z)\mathrm{d}z \ + C$$

Sind x, y, z stetig differenzierbar, so gilt

$$\int_{x(a)}^{x(b)} f(x)\mathrm{d}x = c\int_{y(a)}^{y(b)} g(y)\mathrm{d}y + \int_{z(a)}^{z(b)} h(z)\mathrm{d}z$$

Linearitäts- und Substitutionsregel sind Spezialfälle. Beweis. Nach Voraussetzung haben wir

$$f(x)\frac{\partial x}{\partial t}\mathrm{d}t = cg(y)\frac{\partial y}{\partial t}\mathrm{d}t + h(z)\frac{\partial z}{\partial t}\mathrm{d}t$$

also

$$\phi(t) := f(x)\frac{\partial x}{\partial t} = cg(y)\frac{\partial y}{\partial t} + h(z)\frac{\partial z}{\partial t}$$

Dann

$$\frac{\partial F}{\partial t} = \frac{\partial F}{\partial x}\frac{\partial x}{\partial t} = f(x)\frac{\partial x}{\partial t} = \phi(t)$$

$$\frac{\partial(cG + H)}{\partial t} = c\frac{\partial G}{\partial t} + \frac{\partial H}{\partial t} = c\frac{\partial G}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial H}{\partial z}\frac{\partial z}{\partial t} = cg(y)\frac{\partial y}{\partial t} + h(z)\frac{\partial z}{\partial t} = \phi(t)$$

Das ist die Aussage für unbestimmte Integrale und die für bestimmte folgt mit der Substitutionsregel.

15.3.6 Anwendung der Substitution

Unter der Voraussetzung ♡ gilt $\frac{\partial x}{\partial u} = \frac{\mathrm{d}x}{\mathrm{d}u}$ und es folgt $\frac{\partial x}{\partial u}\mathrm{d}u = \mathrm{d}x$ und somit

$$f(x(u))\frac{\partial x}{\partial u}\mathrm{d}u = f(x)\mathrm{d}x \quad \text{für alle (stetigen) Funktionen } f(x)$$

also

$$\boxed{\begin{aligned} \int f(x(u))\frac{\partial x}{\partial u}\mathrm{d}u &= \int f(x)\mathrm{d}x + C \\ \int_{u(a)}^{u(b)} f(x(u))\frac{\partial x}{\partial u}\mathrm{d}u &= \int_{x(a)}^{x(b)} f(x)\mathrm{d}x + C \end{aligned}}$$

Beispiel 5 $f(x) = x$

$$\int x(u)\frac{\partial x}{\partial u}\mathrm{d}u = \int x\mathrm{d}x = \frac{1}{2}(x(u))^2 + C$$

Beispiel 6 Für $f(x) = 1/x$ und $x(u) \neq 0$ erhält man

$$\int \frac{\frac{\partial x}{\partial u}(u)}{x(u)}\mathrm{d}u = \int \frac{1}{x}\mathrm{d}x = \ln|x(u)| + C\,.$$

Beispiel 8 Auf $\mathbb{R}$ suchen wir $\int \cos u \sin^2 u\,\mathrm{d}u$. Mit $f(x) = x^2$ und $x(u) = \sin u$ ist $\frac{\partial x}{\partial u}(u) = \cos u$, und wir erhalten

$$\int \cos u \sin^2 u\,\mathrm{d}u = \int x(u)^2\frac{\partial x}{\partial u}\mathrm{d}u = \int x^2\mathrm{d}x = \frac{x^3}{3} + C = \frac{\sin^3 u}{3} + C\,.$$

Häufig geht es besser mit folgendem Spezialfall der Integrationsregel

Ist $y = y(x)$ und $f(x)\mathrm{d}x = cg(y)\mathrm{d}y$ also $f(x) = cg(y)\frac{\partial y}{\partial x}$, so

$$\int f(x)\mathrm{d}x = c\int g(y)\mathrm{d}y, \quad \int_c^d f(x)\mathrm{d}c = c\int_{y(c)}^{y(d)} g(y)\mathrm{d}y$$

Beispiel 7 $f(x) = g(\alpha x + \beta)$ mit $\alpha \neq 0$ und $g(y)$ stetig, so ist mit $y = \alpha x + \beta$

$$\frac{\partial y}{\partial x} = \alpha, \quad f(x) = \frac{1}{\alpha} g(y)\frac{\partial y}{\partial x}, \quad f(x)\mathrm{d}x = \frac{1}{\alpha} g(y)\mathrm{d}y$$

$$\int f(x)\mathrm{d}x = \frac{1}{\alpha}\int g(y)\mathrm{d}y + C$$

$$\int_c^d f(x)\,\mathrm{d}x = \frac{1}{\alpha}\int_{\alpha c+\beta}^{\alpha d+\beta} g(x)\mathrm{d}x + C$$

Beispiel $\int e^{2\sin x}\cos x$, $y(x) = 2\sin x$, $g(y) = e^y$

$$e^{2\sin x}\cos x\mathrm{d}x = \frac{1}{2}e^y\mathrm{d}y$$

$$\int e^{2\sin x}\cos x\mathrm{d}x = \frac{1}{2}\int e^y\mathrm{d}y = \frac{1}{2}e^y + C = \frac{1}{2}e^{\sin x} + C$$

$$\int_0^{\frac{\pi}{2}} e^{2\sin x}\cos x\mathrm{d}x = \frac{1}{2}\int_0^2 e^y\mathrm{d}y = \frac{1}{2}(e^2 - e^0)$$

Beispiel 8 Auf $\mathbb{R}$ suchen wir $\int \cos x \sin^2 x\,\mathrm{d}x$. Mit $y = \sin x$ und $g(y) = y^2$

$$\sin^2 x\cos x\,\mathrm{d}x = y^2\mathrm{d}y$$

$$\int \sin^2 x\,\cos x\mathrm{d}x = \int y^2\mathrm{d}y = \frac{y^3}{3} + C = \frac{\sin^3 x}{3} + C$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2 x\,\cos x\mathrm{d}x = \int_{-1}^{1} y^2\mathrm{d}y = \frac{1^3}{3} - \frac{(-1)^3}{3} = \frac{2}{3}$$

Ist $\frac{\partial x}{\partial u}(u) \neq 0$ für alle u und der Wertebereich von $x(u)$ der Definitionsbereich von x, so hat $x = x(u)$ eine Umkehrfunktion $u = u(x)$ und man kann

$$\Phi(u) = \int f(x(u))\frac{\partial x}{\partial u}\mathrm{d}u + C$$

nach x auflösen

$$\int f(x)\mathrm{d}x = F(x) + C = \Phi(u(x)) + C$$

$$\int_c^d f(x)\mathrm{d}x = \Phi(u(d)) - \Phi(u(c))$$

Beweis. Da $\frac{\partial x}{\partial u}$ stetig und nie Null ist, ist $\frac{\partial x}{\partial u}$ nach dem Zwischenwertsatz entweder positiv für alle u oder negativ. Also ist $x(u)$ entweder streng monoton wachsend oder streng monoton fallend. Hieraus folgt die Existenz der Umkehrfunktion $u = u(x)$. Alles weitere folgt aus dem Substitutionslemma. □

Beispiel 9 Wir suchen $\int \sqrt{a^2 - x^2}\mathrm{d}x$ auf $(-a, a)$. Dazu substituieren wir $x := a\sin u$ (man beachte, dass die Ableitung $\frac{\partial x}{\partial u}u = a\cos u$ auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ nicht verschwindet), wir haben

$$\sqrt{a^2 - x^2}\mathrm{d}x = \sqrt{a^2 - a^2\sin^2 u}\,a\cos u\,\mathrm{d}u$$

und wir gelangen zu folgendem Integral

$$\begin{aligned}\int \sqrt{a^2 - x^2}\mathrm{d}x &= \int \sqrt{a^2 - a^2\sin^2 u}\,a\cos u\,\mathrm{d}u = a^2\int \cos^2 u\,\mathrm{d}u \\ &= \frac{a^2}{2}(u + \sin u\,\cos u) + C \qquad \text{(nach Beispiel 4)}\end{aligned}$$

Für die Rücksubstitution $u = \arcsin\frac{x}{a}$ schreiben wir $\cos u = \sqrt{1 - \sin^2 u}$, $\sin u = \frac{x}{a}$ und erhalten

$$\begin{aligned}\int \sqrt{a^2 - x^2}\,\mathrm{d}x &= \frac{a^2}{2}\left(\arcsin\frac{x}{a} + \frac{x}{a}\sqrt{1 - \left(\frac{x}{a}\right)^2}\right) + C \\ &= \frac{1}{2}\left(a^2\arcsin\frac{x}{a} + x\sqrt{a^2 - x^2}\right) + C.\end{aligned}$$

■

Beispiel 10 Wir suchen $\int \frac{1}{\sin x}\,\mathrm{d}x$ auf $(0, \pi)$. Die Substitution $x = 2\arctan u$ führt wegen $\frac{\partial x}{\partial u} = \frac{2}{1+u^2}$ und

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

auf das Integral

$$\int \frac{1}{\sin x}\mathrm{d}x = \int \frac{1 + u^2}{2u}\frac{2}{1 + u^2}\,\mathrm{d}u = \int \frac{1}{u}\mathrm{d}u = \ln|u| + C\,.$$

Rücksubstitution $u = \tan\frac{x}{2}$ liefert

$$\int \frac{1}{\sin x}\,dx = \ln|\tan\frac{x}{2}| + C\,.$$

15.3.7 Partielle Integration

Für $w(t) = u(t)\cdot v(t)$ haben wir nach der Produktregel $\frac{\partial uv}{\partial t} = v\frac{\partial u}{\partial t} + u\frac{\partial v}{\partial t}$ also das Differential

$$\mathrm{d}w = \mathrm{d}uv = v\frac{\partial u}{\partial t}\mathrm{d}t + u\frac{\partial v}{\partial t}\mathrm{d}t = v\mathrm{d}u + u\mathrm{d}v$$

(unter Voraussetzung ♡) und es folgt

$$uv = w = \int 1\,\mathrm{d}w = \int v\frac{\partial u}{\partial t}\mathrm{d}t + \int u\frac{\partial v}{\partial t}\mathrm{d}t + C = \int v\,\mathrm{d}u + \int u\,\mathrm{d}v + C$$

$$\int_{w(a)}^{w(b)} 1\,\mathrm{d}w = \int_a^b v\frac{\partial u}{\partial t}\mathrm{d}t + \int_a^b u\frac{\partial v}{\partial t}\mathrm{d}t = \int_{u(a)}^{u(b)} v\,\mathrm{d}u + \int_{v(a)}^{v(b)} u\,\mathrm{d}v$$

$$u(t)v(t)\Big|_a^b := u(b)v(b) - u(a)v(a) = \int_{w(a)}^{w(b)} 1\,\mathrm{d}w$$

Bei der Berechnung von $\int f(t)g(t)\mathrm{d}t$ wenden wir (zur Vereinheitlichung der Notation) partielle Integration auf $v\mathrm{d}u = f(t)(g(t)\mathrm{d}t)$ oder auf $v\mathrm{d}u = g(t)(f(t)\mathrm{d}t)$ an und erhalten

$$\int f(t)g(t)\mathrm{d}t = uv - \int u\mathrm{d}v, \quad u = \int \mathrm{d}u$$

Beispiel 1 Wir wollen $\int t\sin t\,dt$ bestimmen.

$$t\cdot(\sin t\mathrm{d}t) = v\cdot\mathrm{d}u = \mathrm{d}uv - u\mathrm{d}v, \quad u = \int \mathrm{d}u = \int \sin t\mathrm{d}t = -\cos t$$

$$\int t\sin t\mathrm{d}t = -t\cos t - \int -\cos t\mathrm{d}t = -t\cos t + \sin t + C$$

Hätten wir dagegen $\frac{\partial u}{\partial t} = t$ und $v(t) = \sin t$ gewählt, so hätten wir

$$\int t\sin t\,\mathrm{d}t = \frac{t^2}{2}\sin t - \int \frac{t^2}{2}\cos t\,\mathrm{d}t$$

bekommen. Das Integral $\int t^2\cos t\,dt$ ist aber komplizierter als das Ausgangsintegral. ∎

Beispiel 2 Bei $\int \ln t\,dt$ hilft ein Trick: Wir wählen $\mathrm{d}t = 1\mathrm{d}u$

$$\ln t\mathrm{d}t = \ln t(1\mathrm{d}t) = v\mathrm{d}u = \mathrm{d}uv - u\mathrm{d}v, \quad u = \int 1\mathrm{d}t = t$$

$$\int \ln t\mathrm{d}t = t\ln t - \int t\frac{1}{t}\mathrm{d}t = t\ln t - \int 1\mathrm{d}t = t\ln t - t + C$$

Beispiel 3 Für $\int x^2e^{3x}\mathrm{d}x$ wenden wir partielle Integration zweimal an:

$$\begin{aligned}
\int x^2e^{3x}\mathrm{d}x &= x^2\frac{e^{3x}}{3} - \int \frac{e^{3x}}{3}2x\mathrm{d}x = \frac{1}{3}x^2e^{3x} - \frac{2}{3}\int xe^{3x}\mathrm{d}x \\
&= \frac{1}{3}x^2e^{3x} - \frac{2}{3}(x\frac{e^{3x}}{3} - \int \frac{e^{3x}}{3}1\mathrm{d}x) \\
&= \frac{1}{3}x^2e^{3x} - \frac{2}{9}xe^{3x} + \frac{2}{9}\int e^{3x}dx \\
&= (\frac{1}{3}x^2 - \frac{2}{9}x + \frac{2}{27})e^{3x} + C\,.
\end{aligned}$$

∎

Beispiel 4 Für $\int \cos^2 dx$ hilft wieder ein einfacher Trick:

$$\begin{aligned}
\int \cos^2 x\,dx &= \int \cos x\cdot\cos x\,dx = \sin x\,\cos x - \int \sin x(-\sin x)dx \\
&= \sin x\cos x + \int \sin x\sin x\,dx \\
&= \sin x\cos x + \int (1-\cos^2 x)dx \\
&= \sin x\cos x + x - \int \cos^2 x\,dx\,.
\end{aligned}$$

Hieraus folgt schließlich

$$2\int \cos^2 x\,dx = \sin x\cos x + x$$

bzw.

$$\int \cos^2 x\,dx = \frac{1}{2}(\sin x\cos x + x) + C\,.$$

In diesem Beispiel hätte man einfacher benutzen können, dass $\cos^2 x = \frac{1}{2}\cdot(1+\cos 2x)$ ist. Damit bekommt man sofort

$$\int \cos^2 x\,dx = \frac{1}{2}\int dx + \frac{1}{2}\int \cos 2x\,dx = \frac{1}{2}(x + \frac{\sin 2x}{2}) + C\,.$$

∎

16 Rationale Funktionen

16.1 Polynome

16.1.1 Defintion

Im Folgenden sei K ein Körper, z.B. $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$. Ein *Polynom mit Koeffizienten* mit a_k in K ist ein formaler Ausdruck

$$p(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 = \sum_{i=0}^n a_ix^i$$

Dieses Polynom kann aber auch durch jeden andern Ausdruck dargestellt werden, der sich durch Anwenden der Körperaxiome, ausgenommen die Existenz Inverser a^{-1}, und Rechnen in K ergibt. Insbesondere ist die Reihenfolge der Summanden und Hinzufügung oder Wegnahme von Summanden $0x^k$ unwesentlich. Die Gesamtheit der Polynome mit Koeffizienten in K wird als $K[x]$ notiert.

16.1.2 Auswertung

Ist $\alpha \in K$, so kann man $p(x)$ an α *auwerten*

$$p(\alpha) = a_n\alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1\alpha + a_0$$

und erhält somit eine *Polynomfunktion*

$$t \mapsto p(t) = a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0 \in K \quad (t \in K)$$

Für $K \subseteq \mathbb{C}$ ist diese beliebig oft differenzierbar und integrierbar.

Um ein Polynom auszuwerten schreibt man es zur Einsparung von Multiplikationen in der folgenden Form (*Hornerschema*)

$$(((\ldots(((a_n x + a_n - 1)x + a_{n-2})x + a_{n-3})\ldots)x + a_2)x + a_1)x + a_0$$

16.1.3 Grad

Ist $a_n \neq 0$ und $a_k = 0$ für alle $k > n$ so ist das Polynom vom *Grad* $\deg p(x) = n$ aund a_n der *Leitkoeffizient*. In diesem Falle sind in der obigen Darstellung die Koeffizienten eindeutig bestimmt. $p(x)$ ist hier *normiert*, falls $a_n = 1$.

Das Polynom mit $a_k = 0$ für alle k ist das *Nullpolynom* $\mathbf{0}$ und $\deg \mathbf{0} = -\infty$. Jedes Polynom $\neq \mathbf{0}$ ist von der Form $cq(x)$ mit $c \neq 0$ und normiertem $q(x)$.

16.1.4 Summe und Produkt

Ist $q(x) = b_m x^m + b_{m-1} x^{m-1} + \ldots + b_1 x + b_0$ ein weiteres Polynom, so rechnet man

$$p(x) + q(x) = (a_k + b_k)x^k + \ldots + (a_1 + b_1)x + a_0 + b_0 \quad k = \max\{n, m\}$$

$$p(x)q(x) = a_n b_m x^{n+m} + \ldots + (a_i b_0 + \ldots + a_0 b_i)x^i + \ldots + a_0 b_0$$

und es gilt

$$\deg p(x) + \deg q(x) \leq \max\{\deg p(x), \deg q(x)\}, \quad \deg p(x)q(x) = \deg p(x) + \deg q(x)$$

Es gelten die Rechenregeln wie in einem Körper - ausgenommen Division. Aber es gilt die *Kürzungsregel*

$$p(x)q(x) = p(x)r(x) \Rightarrow q(x) = r(x) \quad \text{falls } p(x) \neq \mathbf{0}$$

Diese folgt aus

$$p(x)q(x) = \mathbf{0} \Rightarrow q(x) = \mathbf{0} \quad \text{falls } p(x) \neq \mathbf{0}$$

16.1.5 Polynomdivision

Lemma 16.1 *Sind $p(x)$ und $q(x) = b_0 + \ldots + b_m x^m$ in $K[x]$ Polynome der Grade $n \geq m$ (also $b_m \neq 0$), so gibt es eindeutig bestimte Polynome $s(x)$ und $r(x)$ von Grad $n-m$ bzw. höchstens $m-1$ so, dass $p(x) = s(x)q(x) + r(x)$.*

Beweis und Algorithmus. Zur Berechnung geht man wie bei der Division von Dezimalzahlen vor: setze $c_{n-m} = \frac{a_n}{b_m}$ und $r_1(x) = p(x) - c_{n-m}x^{n-m}q(x)$. Dann hat r_1 höchstens Grad $n-1$ und $p(x) = c_{n-m}x^{n-m}q(x) + r_1(x)$. Man setzt nun das Verfahren mit $r_1(x)$ anstelle von $p(x)$ fort und erhält $r_1(x) = c_{n-m-1}x^{n-m-1}q(x) + r_2(x)$ mit $r_2(x)$ von Grad höchstens $n-2$. Schliesslich erhält man $r_{n-m}(x) = c_0 q(x) + r_{n-m+1}(x)$ mit $r(x) = r_{n-m+1}(x)$ von Grad höchstens $m-1$. Fasst man zusammen, so ergibt das $p(x) = (c_{n-m}x^{n-m} + \ldots + c_0)q(x) + r(x)$. Ist $p(x) = \times s(x)q(x) + \times r(x)$ so folgt $(s(x) - \times s(x))q(x) = \times r(x) - r(x)$ mit $\deg \times r(x) - r(x) < \deg q(x)$, also $\deg s(x) - \times s(x) < 0$ und somit $s(x) - \times s(x) = \mathbf{0} = \times r(x) - r(x)$. □

Korollar 16.2 *Zu $p(x) \in K[x]$ und $a \in K$ gibt es eindeutig bestimmte $r \in K$ und $q(x) \in K[x]$ mit*

$$p(x) = q(x)(x-a)^k + r(x), \ \deg r(x) < k$$

Es folgt $p(a) = r(a)$.

Korollar 16.3 Entwickeln. *Zu $p(x) \in K[x]$ vom Grad n und $\alpha \in K$ gibt es eindeutig bestimmte $d_0, \ldots, d_n$ in K mit*

$$p(x) = d_0 + d_1(x-\alpha) + \ldots d_n(x-\alpha)^n \quad d_0 = p[\alpha], ; d_n \neq 0$$

Beweis durch Induktion über $n \geq \deg p(x)$. Durch Division $p(x) = d_n(x-\alpha)^n + r(x)$ mit $\deg r(x) < n$, also $r(x) = d_0 + d_1(x-\alpha) + \ldots d_{n-1}(x-\alpha)^{n-1}$ nach Induktion. □

16.1.6 Nullstellen

Ist $\alpha \in K$ und $p(\alpha) = 0$, so heisst α *Nullstelle* von $p(x)$.

Korollar 16.4 Abspaltung. *Ist $p(x) \in K[x]$ ein Polynom vom Grad $n > 0$ und $\alpha \in K$ eine Nullstelle, so gibt es ein (eindeutig bestimmtes) Polynom $q(x) \in K[x]$ vom Grad $n-1$ so, dass $p(x) = q(x)(x-\alpha)$. Ist $K \subseteq \mathbb{C}$ so folgt*

$$p'(a) = q(a)$$

Beweis. Division ergibt $p(x) = q(x)(x-\alpha) + \beta$ mit konstantem β. Einsetzen ergibt $0 = p(\alpha) = q(\alpha)(\alpha - \alpha) + \beta = \beta$. Eindeutigkeit nach der Kürzungsregel. Ableiten nach der Produktregel ergibt $p'(x) = q'(x)(x-a) + q(x)$ also $p'(a) = q(a)$. □

$x - \alpha$ heisst der Linearfaktor zur Nullstelle α. Hat man $p(x) = q(x)(x-\alpha)^k$ mit $q(\alpha) \neq 0$ (und das ist eindeutig bestimmt), so ist α eine k-fache Nullstelle von $p(x)$.

Korollar 16.5 *Ein Polynom von Grad $n > 0$ hat höchstens n verschiedene Nullstellen.*

Korollar 16.6 *Ist K unendlich, so ist ein Polynom schon eindeutig bestimmt durch die zugehörige Polynomfunktion*

$$t \mapsto p(t) \in K \text{ für } t \in K.$$

Bemerkung 16.7 *Jede rationale Nullstelle von $p(x) = \sum_{i=0}^n a_i x^i \in \mathbb{Z}[x]$ ist von der Form rs^{-1} mit ganzen $r|a_0$ und $s|a_n$*

Beweis. Sei $p(\alpha) = 0$. O.B.d.A. $\alpha = rs^{-1}$ mit r, s teilerfremd. Es folgt $a_n r^n s^{-n} = -\sum_{i<n} a_i r^i s^{-i}$ also $a_n r^n s^{-1} \in \mathbb{Z}$. Also $s|(a_n r^n)$ und wegen GGT(s, r^n) =GGT$(s, r) = 1$ dann $s|a_n$. Andererseits ist r Nullstelle von $s^n p(x)$, also Teiler von $-s^n a_0 = \sum_{i=1}^n s^n a_i r^i$ und damit von a_0.

Lemma 16.8

$$g(x) = q(x)(x-a)^k \Rightarrow q(a) = \frac{1}{k!} g^{(k)}(a)$$

Beweis. Wir zeigen durch Induktion

$$g^{(l)}(x) = q_l(x)(x-a)^{k-l+1} + k\cdots(k-l+1)\cdot q(x)(x-a)^{k-l}$$

mit passenden $q_l(x)$. Ableiten ergibt

$$g^{(l+1)}(x) = q_l'(x)(x-a)^{k-l+1} + q_l(x)(k-l+1)(x-a)^{k-l} + k\cdots(k-l+1)\cdot q'(x)(x-a)^{k-l} +$$

$$+k\cdots(k-l+1)\cdot(k-l)q(x)(x-a)^{k-l-1}\Box$$

Also

$$g^{(k)}(x) = q_k(x)(x-a) + k!q(x)$$

und die Behauptung folgt durch Einsetzen von a. □

16.1.7 Euklidischer Algorithmus

Wir definieren nun die Teilbarkeit in $K[x]$ analog zu der in $\mathbb{Z}$

$$p(x) \text{ teilt } q(x) \;\Leftrightarrow\; \exists r(x).\quad q(x) = p(x)r(x)$$

$p(x)$ und $q(x)$ sind *teilerfremd*, wenn ihre einzigen gemeinsamen Teiler Kostanten sind,

Satz 16.9 (Bezout) *Zu teilerfremden $p(x), q(x) \in K[x]$ gibt es $r(x), s(x) \in K[x]$ mit*

$$1 = p(x)r(x) + q(x)s(x) \quad \deg r(x) < \deg q(x), \quad \deg s(x) < \deg p(x)$$

Diese bestimmt man mit dem Euklidischen Algorithmus für Polynome.

Beweis exemplarisch in $\mathbb{Z}$. Die beiden Startzeilen sind trivial. Dann zieht man immer ein Vielfaches der letzten aktuelle Zeile von der vorletzten ab (Division mit Rest der Einträge in der ersten Spalte). Dabei entsteht wieder eine gültige Relation.

$$\begin{array}{rcrcr}
98 &=& 1\cdot 98 &+& 0\cdot 27\\
27 &=& 0\cdot 98 &+& 1\cdot 27\\
17 &=& 1\cdot 98 &+& -3\cdot 27\\
10 &=& -1\cdot 98 &+& 4\cdot 27\\
7 &=& 2\cdot 98 &+& -7\cdot 27\\
3 &=& -3\cdot 98 &+& 11\cdot 27\\
1 &=& 8\cdot 98 &+& -29\cdot 27
\end{array}$$

16.1.8 Faktorzerlegung

Ein nicht konstantes Polynom $p(x)$ heisst *unzerlegbar* oder *irreduzibel*, wenn in jeder Zerlegung $p(x) = q(x)r(x)$ einer der Faktoren konstant ist. Diese spielen die Rolle der Primzahlen. Man bemerkt, dass *ein irreduzibles $p(x)$ ein Produkt $q(x)r(x)$ nur dann teilt, wenn es mindestens einen der Faktoren teilt*: sonst sind z.B. $p(x)$ und $q(x)$ teilerfremd, also gibt es $a(x)p(x) + b(x)q(x) = 1$ und $p(x)$ teilt $r(x) = a(x)p(x)r(x) + b(x)q(x)r(x)$. Wie für $\mathbb{Z}$ beweist man

Satz 16.10 *Jedes nichtkonstante Polynom $p(x) \in K[x]$ hat eine bis auf die Reihenfolge eindeutige Darstellung mit normierten irreduziblen $p_i(x) \in K[x]$ und $a \in K$*

$$p(x) = ap_1(x)\cdots p_n(x)$$

Beweis: Die Existenz folgt durch Ordnungsinduktion über den Grad: Ist $p(x)$ nicht schon irreduzibel, so $p(x) = q(x)r(x)$ mit Polynomen kleineren Grades, also $q(x) = bq_1(x)\cdots q_k(x)$ und $r(x) = cr_1(x)\cdots r_l(x)$ und somit $p(x) = bcqq_1(x)\cdots q_k(x)\cdot r_1(x)\cdots r_l(x)$.

Die Eindeutigkeit folgt nun durch Induktion über die Anzahl der Faktoren: Ist

$$p(x) = ap_1(x)\cdots p_n(x) = bq_1(x)\cdots q_m(x)$$

so $a = b$ da die $p_i(x)$ und $q_i(x)$ alle normiert sind. Auch teilt $p_1(x)$ eines der $q_i(x)$, nach Umnummerierung also $q_1(x)$ und wegen Normiertheit folgt $p_1(x) = q_1(x)$. Nun kürzt man $p_1(x)$ und beruft sich auf die Induktionsannahme. □.

16.1.9 Fundamentalsatz der Algebra

Satz 16.11 *Die einzigen normierten irreduziblen Polynome in $\mathbb{C}[x]$ sind die linearen $x - \alpha$.*

Beweis. Es sei

$$\mu := \inf\{|w| \mid w = p(\alpha),\ \alpha \in \mathbb{C}\}.$$

Dieses Infimum wird auch angenommen, d.h. es gibt ein $\alpha_0 \in \mathbb{C}$ mit $p(\alpha_0) = \mu$ (Begründung in Stichworten: Man sieht leicht, dass $|p(\alpha)| \to \infty$ für $|\alpha| \to \infty$. Daher ist μ ein Grenzwert von der Form $\mu = \lim_{\alpha\to\alpha_0} |p(\alpha)|$ und die Behauptung folgt aus der Stetigkeit von p). Wir haben die Annahme $\mu > 0$ auf einen Widerspruch zu führen. Durch Entwickeln

$$p(x) = p(\alpha_0) + d_1(x-\alpha_0) + d_2(x-\alpha_0)^2 + \cdots + d_n(x-\alpha_0)^n.$$

Sei $k \le n$ der kleinste Index mit $d_k \neq 0$. Ein solches k existiert, denn sonst wäre $p(x)$ konstant. Setzt man

$$d := d_k \text{ und } R(y) := d_{k+1}y^{k+1} + \cdots + d_ny^n$$

so hat man

$$p(\alpha) = p(\alpha_0) + d\epsilon^k + R(\epsilon) \qquad \text{für alle } \alpha = \alpha_0 + \epsilon \in \mathbb{C}$$

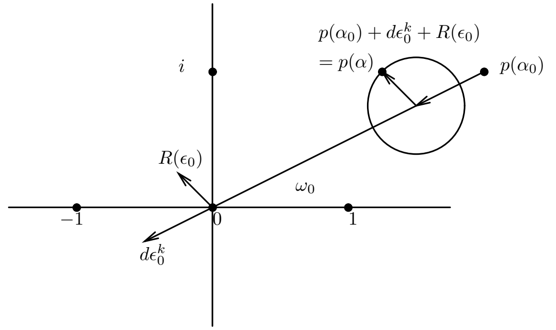
Wir möchten $\epsilon \in \mathbb{C}$ so wählen, dass $d\epsilon^k$ die zu $p(\alpha_0)$ entgegengesetzte Richtung hat. Genauer gesagt schreiben wir $p(\alpha_0) = r(\cos\omega_0 + i\sin\omega_0)$ in Polarkoordinaten. Dann soll $d\epsilon^k$ ein positives reelles Vielfaches von $\cos(\pi+\omega_0) + i\sin(\pi+\omega_0)$ sein (siehe Zeichnung). Die komplexe Zahl $d \neq 0$ ist fest und lautet in Polarkoordinaten $d = s(\cos\omega_1 + i\sin\omega_1)$. Schreibt man auch ϵ als $\epsilon = t(\cos\omega_2 + i\sin\omega_2)$, so

$$d\epsilon^k = st^k(\cos(\omega_1 + k\omega_2) + i\sin(\omega_1 + k\omega_2)),$$

und man sieht, dass man mit $\omega_2 := (\pi + \omega_0 - \omega_1)k^{-1}$ passende ϵ erhält - für jedes $t > 0$. Weil entweder $R(y)$ das Nullpolynom ist, oder ein Polynom in y mit allen Exponenten $> k$, gilt

$$\frac{|R(t(\cos\omega_2 + i\sin\omega_2))|}{|d(t(\cos\omega_2 + i\sin\omega_2))^k|} \to 0 \quad \text{für} \quad t \to 0.$$

Es gibt also eine reelle Zahl $t_0 > 0$, derart dass für $\epsilon_0 := t_0(\cos\omega_2 + i\sin\omega_2)$ gilt $|R(\epsilon_0)| \le |d\epsilon_0^k|/2$. Aus der Skizze ist klar, dass für $\alpha := \alpha_0 + \epsilon_0$ gilt $|p(\alpha)| < |p(\alpha_0)| = \mu$. Dies ist der gewünschte Widerspruch.



□

16.1.10 Reelle Faktorisierung

Satz 16.12 *Ist $p(x) = a_n x^n + \ldots + a_0$ ein Polynom mit reellen Koeffizienten $a_0, \ldots, a_n$, so ist eine komplexe Zahl α Nullstelle von $p(x)$ genau dann, wenn auch ihre Konjugierte $\overline{\alpha}$ Nullstelle von $p(x)$ ist; man hat eine Zerlegung*

$$p(x) = a_n(x-\alpha_1)^{n_1}\ldots(x-\alpha_l)^{n_l} q_{l+1}(x)^{n_{l+1}}\ldots q_m(x)^{n_m}$$

in lineare und quadratische reelle Polynome, die den reellen Nullstellen bzw. den Paaren konjugiert komplexer Nullstellen entsprechen. Dabei ist das Polynom zu $\alpha = a + bj$ und $\overline{\alpha} = a - bj$ das folgende

$$q(x) = (x-\alpha)(x-\overline{\alpha}) = (x-(a+bj))(x-(a-bj)) = x^2 - 2ax + a^2 + b^2.$$

Beweis. Die erste Behauptung ergibt sich sofort daraus, dass die Konjugation mit Addition und Multiplikation verträglich ist und reelle Zahlen festlässt. Dann fasst man im Fundamentalsatz die Paare konjugierter zusammen. □

16.2 Rationale Funktionen

16.2.1 Körper der rationalen Funktionen

Wie von $\mathbb{Z}$ zu $\mathbb{Q}$ kann man vom Polynomring $K[x]$ zum *Körper $K(x)$ der rationalen Funktionen* übergehen. Man betrachte Quotienten, d.h. Paare

$$\frac{p(x)}{q(x)} \quad \text{mit } q(x) \neq \mathbf{0}$$

und setzt zwei solche Quotienten gleich

$$\frac{p(x)}{q(x)} = \frac{r(x)}{s(x)} \quad \Leftrightarrow \; p(x)s(x) = q(x)r(x)$$

Adddition und Multiplaktion sind wie gewohnt definiert

$$\frac{p(x)}{q(x)} + \frac{r(x)}{s(x)} = \frac{p(x)s(x) + q(x)r(x)}{q(x)s(x)}, \quad \frac{p(x)}{q(x)} \cdot \frac{r(x)}{s(x)} = \frac{p(x)r(x)}{q(x)s(x)}$$

Für $K \subseteq \mathbb{C}$ können wir natürlich auch die Funktion

$$t \mapsto \frac{p(t)}{q(t)} \quad t \in K, \; q(t) \neq 0$$

betrachten und mit dem Quotienten identifizieren. Diese Funktionen sind auf ihrem Definitionsbereich beliebig oft differenzierbar und integrierbar.

16.2.2 Partialbruchzerlegung

Lemma 16.13 *Seien $f(x), g(x) \in K[x]$ gegeben und a k-fache Nullstelle von $g(x)$*

$$g(x) = q(x)(x-a)^k, \quad q(a) \neq 0$$

Dann gibt es eindeutig bestimmte $A \in K$ und $h(x) \in K[x]$ so dass

$$\frac{f(x)}{g(x)} = \frac{A}{(x-a)^k} + \frac{h(x)}{q(x)(x-a)^{k-1}}$$

nämlich

$$A = \frac{f(a)}{q(a)} = \frac{k!f(a)}{g^{(k)}(a)} \quad \text{und } f(x) - Aq(x) = h(x)(x-a)$$

wobei $h(x)$ durch Polynomdivision bestimmt werden kann.

Beweis: Ist A wie angegeben, so $f(a) - Aq(a) = 0$, also $f(x) - Aq(x)$ durch $x - a$ teilbar. Es folgt

$$\frac{f(x)}{g(x)} = \frac{Aq(x)}{q(x)(x-a)^k} + \frac{h(x)}{q(x)(x-a)^{k-1}}$$

Umgekehrt folgt durch Multiplikation mit $g(x)$

$$f(x) = Aq(x) + h(x)(x-a)$$

und durch Einsetzen von a für x dass $f(a) = Aq(a)$. Damit ist $A = \frac{f(a)}{q(a)}$ eindeutig bestimmt - beachte $q(a) \neq 0$. Also $f(x) - Aq(x) = h(x)(x-a)$ und die die Eindeutigkeit von $h(x)$ folgt aus der Kürzungsregel. □

Satz 16.14 *Zu Polynomen $f(x), g(x) \in K[x]$, $K = \mathbb{R}, \mathbb{C}$, mit $g(x) = \prod_{i=1}^n p_i(x)^{k_i}$ mit irreduziblen normierten $p_i(x)$ gibt es eine Darstellung*

$$\frac{f(x)}{g(x)} = r(x) + \sum_{i=1}^{n}\sum_{k=1}^{k_i} \frac{r_{ik}(x)}{p_i(x)^k} \quad \text{mit } \deg r_{ik}(x) < \deg p_i(x)$$

Dabei ist $r(x) = \mathbf{0}$ genau dann, wenn $\deg g(x) < \deg f(x)$. Die Darstellung ist eindeutig. Über $\mathbb{C}$ gilt $\deg p_i(x) = 1$. Über $\mathbb{R}$ hat man $\deg p_i(x) = 1$ bzw. Paare konjugierter komplexer Partialbrüche

$$\frac{A}{(x-\alpha)^k} + \frac{A^*}{(x-\alpha^*)^k} = \frac{ax+b}{(x^2+cx+d)^k}$$

$$c = -(\alpha + \alpha^*), \; d = \alpha\alpha^*, a = A + A^*, b = -(A\alpha^* + A^*\alpha) \in \mathbb{R}$$

Beweis. Über $\mathbb{C}$ mit dem Lemma und Induktion über $\deg g(x)$ - und Berufung auf den Fundamentalsatz. Möchte man die Partialbruchzerlegung reeller rationaler Funktionen über $\mathbb{R}$ bestimmen, so ist mit jeder komplexen Nullstelle α von $g(x)$ auch die konjugierte α^* eine Nullstelle - weil $(a+b)^* = a^* + b^*$ und $(ab)^* = a^*b^*$. Also kann man die Behauptung auf den komplexen Fall zurückführen. □

Korollar 16.15 *Ist $g(x) = c(x-a_1)\cdots(x-a_n)$ mit lauter verschiedenen a_k und $\deg f(x) < \deg g(x)$ so gilt*

$$\frac{f(x)}{g(x)} = \frac{f(a_1)}{g'(a_1)} \cdot \frac{1}{x-a_1} + \ldots + \frac{f(a_n)}{g'(a_n)} \cdot \frac{1}{x-a_n}$$

16.2.3 Mehrfache Nullstellen

Lemma 16.16 *Sei $K \subseteq \mathbb{C}$ und $f(x), g(x) \in K[x]$ gegeben und $g(x) = q(x)(x-a)^k$, $q(a) \neq 0$. Dann gibt es eindeutig bestimmte $A, B \in K$ und $r(x) \in K[x]$ so dass*

$$\frac{f(x)}{g(x)} = \frac{A}{(x-a)^k} + \frac{B}{(x-a)^{k-1}} + \frac{r(x)}{q(x)(x-a)^{k-2}}$$

Nämlich

$$A = \frac{f(a)}{q(a)}, \quad B = \frac{f'(a) - Aq'(a)}{q(a)}$$

Beweis. Existenz und Eindeutigkeit vin A, B und $r(x)$. nach dem Satz über Partialbruchzerlegung. Durch Multiplikation mit $g(x)$ folgt

$$f(x) = Aq(x) + Bq(x)(x-a) + r(x)(x-a)^2$$

durch Differenzieren

$$f'(x) = Aq'(x) + Bq(x) + Bq'(x)(x-a) + (r'(x) + 2(x-a))(x-a)$$

und durch Einsetzen von a

$$f'(a) = Aq'(a) + Bq(a) \quad \square$$

Damit kann man die Partialbrüche zu doppelten Nullstellen bestimmen.

16.2.4 Quadratische Faktoren

Lemma 16.17 *Seien $f(x), g(x) \in \mathbb{R}[x]$ und $x^2 + cx + d$ ein irreduzibler Faktor von $g(x)$ der Vielfachheit k*

$$g(x) = q(x)(x^2+cx+d)^k, \quad x^2+cx+d \text{ kein Teiler von } q(x)$$

Dann gibt es eindeutig bestimmte $A, B \in \mathbb{R}$ und $h(x) \in \mathbb{R}[x]$ mit

$$(*) \quad \frac{f(x)}{g(x)} = \frac{Ax+B}{(x^2+cx+d)^k} + \frac{h(x)}{q(x)(x^2+cx+d)^{k-1}}$$

Dabei sind A, B die eindeutig bestimmten Lösungen des linearen Gleichungssystems

$$rA = \mu, \quad sA + rB - a = \mu c, \quad sB - b = \mu d$$

wobei $ax+b$ bzw. $rx+s$ die Reste von $f(x)$ bzw. $q(x)$ bei Division durch x^2+cx+d sind.

Beweis. Die Existenz und Eindeutigkeit wurde schon im Satz über die Partialbruchzerlegung bewiesen. Seien nun α, α^* die komplexen Nullstellen von x^2+cx+d also

$$x^2 + cx + d = (x-\alpha)(x-\alpha^*)$$

Ist $f(x) = m(x)(x^2+cx+d) + ax + b$ so folgt $f(\alpha) = a\alpha + b$ und $f(\alpha^*) = a\alpha^* + b$. Ebenso $q(\alpha) = r\alpha + s$ und $q(\alpha^*) = r\alpha^* + s$. Multipliziert man (*) mit $(x^2+cx+d)^k$ so folgt

$$\frac{f(x)}{q(x)} = Ax + B + \frac{h(x)}{q(x)}(x^2+cx+d)$$

und durch Einsetzen von α und α^*

$$\frac{a\alpha+b}{r\alpha+s} = A\alpha + B, \quad \frac{a\alpha^*+b}{r\alpha^*+s} = A\alpha + B$$

Also sind $\alpha \neq \alpha^*$ die Nullstellen des quadratischen Polynoms

$$(Ax+B)(rx+s) - (ax+b)$$

und dieses somit von der Form $\mu(x^2+cx+d)$ mit eindeutig bestimmtem $\mu \in \mathbb{R}$. Also

$$rAx^2 + (sA + rB - a)x + sB - b = \mu x^2 + \mu cx + \mu d$$

und die Behauptung folgt durch Koeffizientenvergleich. □.

16.2.5 Beispiele

Beispiel.

$$f(x) = x^2, \; g(x) = (x-1)^2(x^2+1) = (x-1)^2(x-j)(x+j)$$

$$\frac{f(x)}{g(x)} = \frac{x^2}{(x-1)^2(x^2+1)} = \frac{A_1}{x-1} + \frac{A_2}{(x-1)^2} + \frac{C}{x-j} + \frac{D}{x+j}$$

$$= \frac{A_1}{x-1} + \frac{A_2}{(x-1)^2} + \frac{Ax+B}{x^2+1}$$

Für $a = 1$: $q(x) = x^2+1$, $q'(x) = 2x$, $f'(x) = 2x$

$$A_2 = \frac{f(a)}{q(a)} = \frac{1}{2}, \quad A_1 = \frac{f'(a) - A_2 q'(a)}{q(a)} = \frac{2 - \frac{1}{2}2}{2} = \frac{1}{2}$$

Die Reste von x^2 bzw. $(x-1)^2$ bei Division durch x^2+1 sind -1 bzw. $-2x$ also

$$(Ax+b)(-2x) + 1 = \mu x^2 + \mu$$

und es folgt

$$\mu - 1, \; B = 0, \; A = \frac{-1}{2}$$

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

Komplexe Rechnung: Für $a = j$: $q(x) = (x-1)^2(x+j)$

$$C = \frac{f(a)}{q(a)} = \frac{-1}{(j-1)^2 2j} = \frac{-1}{-2j2j} = \frac{-1}{4}$$

Da $f(x), g(x) \in \mathbb{R}[x]$ folgt

$$D = C^* = -\frac{1}{4}$$

Nun

$$-\frac{1}{4}\left(\frac{1}{x-j} + \frac{1}{x+j}\right) = -\frac{1}{4}\frac{x+j+x-j}{(x-j)(x+j)} = \frac{-\frac{1}{2}x}{x^2+1}$$

Also

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

Beispiel.

$$\frac{x^3+x}{(x-1)^2(x^2+x+1)} = \frac{A_1}{x-1} + \frac{A_2}{(x-1)^2} + \frac{Ax+B}{x^2+x+1}$$

$a = 1$, $q(x) = x^2+x+1$, $f'(x) = 3x+1, q'(x) = 2x+1$

$$A_2 = \frac{f(a)}{q(a)}\frac{2}{3}, \quad A_1 = \frac{f'(a) - A_2 q'(a)}{q(a)} = \frac{1}{3}(4 - \frac{2}{3}3) = \frac{2}{3}$$

$x^2 + cx + d = x^2 + x + 1$. Reste von $x^3 - x$ bzw. $q(x) = (x-1)^2$ sind

$$x+1, \quad -3x$$

Es folgt

$$(Ax+B)(-3x) - (x+1) = -3Ax^2 + (-3B-1)x - 1 = \mu x^2 + \mu x + \mu$$

$$-3A = \mu, \ (-3B-1) = \mu, \mu = 1$$

$$A = \frac{1}{3}, \ B = 0$$

Mühsamer ist die komplexe Rechnung. $x^2 + cx + d$ hat Nullstellen

$$\zeta = -\frac{1}{2} + \frac{\sqrt{3}}{2}j, \quad \zeta^* = \zeta^2 = -\frac{1}{2} - \frac{\sqrt{3}}{2}$$

Es gilt $\zeta^3 = 1$. Die Partialbrüche

$$\frac{C}{x-\zeta}, \quad \frac{C^*}{x-\zeta^*}$$

bestimmen wir mit

$$C = \frac{f}{\zeta}(\zeta-1)^2(\zeta-\zeta^2) = \frac{\zeta}{(\zeta-1)^3} = -\frac{1}{3\sqrt{3}}\zeta j = \frac{1}{6} + \frac{1}{6\sqrt{3}}j$$

da

$$(\zeta-1)^3 = \zeta^3 - 3\zeta^2 + 3\zeta - 1 = 3(\zeta - \zeta^*) = 3\sqrt{3}j$$

Es folgt

$$A = C + C^* = \operatorname{Re} C = \frac{1}{3}$$

$$B = C\zeta^* + C^*\zeta = \frac{-j}{3\sqrt{3}}\zeta\zeta^2 + \frac{j}{3\sqrt{3}}\zeta\zeta^2 = 0$$

16.2.6 Partialbruchzerlegung über beliebigen Körpern

Wir wollen rationale Funktionen in einfacherer Weise darstellen. Dazu die folgenden drei Reduktionsschritte

- Ausdividieren $f(x) = q(x)g(x) + r(x), \ \deg r(x) < \deg g(x)$

$$\frac{f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)} \qquad \deg r(x) < \deg g(x)$$

- Bezout $a(x)h(x) + b(x)g(x) = 1$

$$\frac{f(x)}{g(x)h(x)} = \frac{f(x)a(x)}{g(x)} + \frac{f(x)b(x)}{h(x)} \quad \text{für teilerfremde } g(x),\ h(x)$$

- Zerlegen $f(x) = q(x)g(x) + r(x), \ \deg r(x) < \deg g(x)$

$$\frac{f(x)}{g(x)^n} = \frac{r(x)}{g(x)^n} + \frac{q(x)}{g(x)^{n-1}}, \qquad \deg r(x) < \deg g(x)$$

Dann gilt der Satz über Partialburchzerlegung wie oben formuliert.

Beweis. Existenz. Wir führen exemplarisch eine Partialbruchzerlegung von $\frac{73}{60}$ in $\mathbb{Q}$ aus, indem wir folgende Schritte anwenden

$$\frac{a}{b} = q + \frac{r}{b} \quad \text{mit } a = qr + b \ \ 0 \leq r < b$$

$$\frac{c}{ab} = \frac{sc}{a} + \frac{rc}{b} \quad \text{mit } ra + sb = 1 \text{ für teilerfremde } a, b$$

$$\frac{a}{p^k} = \frac{q}{p^{k-1}} + \frac{r}{p^k} \quad \text{mit } a = qp + r, \ 0 \leq r < p$$

Lösung

$$\frac{71}{60} = 1 + \frac{11}{60}$$

$$60 = 5 \cdot 12, \quad 1 = 5 \cdot 5 - 2 \cdot 12, \quad \frac{11}{60} = \frac{55}{12} - \frac{22}{5} = \frac{7}{12} - \frac{2}{5}$$

$$12 = 3 \cdot 4, \quad 1 = 4 - 3, \quad \frac{7}{12} = \frac{28}{12} - \frac{21}{12} = 1 + \frac{1}{3} - \frac{3}{4}$$

$$3 = 1 \cdot 2 + 1, \quad \frac{3}{4} = \frac{2}{4} + \frac{1}{4} = \frac{1}{2} + \frac{1}{4}$$

$$\frac{71}{60} = 1 + \frac{11}{60} = 1 + \frac{7}{12} - \frac{2}{5} = 2 - \frac{2}{5} + \frac{1}{3} - \frac{1}{2} - \frac{1}{4}$$

Hat man eine Partialbruchzerlegung wie im Satz, bringt man die Partialbrüche auf den Hauptnenner $g(x)$ und addiert sie auf, so erhält man im Zähler einen Grad $< \deg g(x)$. Damit ist $r(x)$ das Ergebnis bei der Division mit Rest und eindeutig bestimmt.

Zum weiteren Beweis der Eindeutigkeit dürfen wir also $r(x) = 0$ annehmen. Seien also zwei Zerlegungen gegeben

$$\sum_{i=1}^{n}\sum_{k=1}^{k_i}\frac{r_{ik}(x)}{p_i(x)^k} = \sum_{i=1}^{n}\sum_{k=1}^{k_i}\frac{s_{ik}(x)}{p_i(x)^k} \quad \text{mit } \deg r_{ik}(x), \deg s_{ik}(x) < \deg p_i(x)$$

Sei k maximal so, dass für ein i $r_{ik}(x) \neq s_{ik}(x)$. Indem man in der Gleichung alle gleichen Terme streicht und dann mit dem Hauptnenner $N(x)$ multipliziert, erhält man in jedem Summanden einen Faktor $p_i(x)$ in Zähler, ausser in

$$\frac{r_{ik}(x)N(x)}{p_i(x)^k} \text{ und } \frac{s_{ik}(x)N(x)}{p_i(x)^k}$$

Also ist $p_i(x)$ Teiler von $h(x)(r_{ik}(x) - s_{ik}(x))$ wobei

$$h(x) = \frac{N(x)}{p_i(x)^k}$$

nach Wahl von k zu $p_i(x)$ teilerfremd ist. Also ist $p(x)$ Teiler von $r_{ik}(x) - s_{ik}(x)$, Da $\deg(r_{ik}(x) - s_{ik}(x) < \deg p_i(x)$ folgt $r_{ik}(x) - s_{ik}(x) = \mathbf{0}$. □

Für teilerfremde Polynome erhält man

$$\frac{1}{(x-a)(x-b)} = \frac{1}{a-b}\left(\frac{1}{x-a} + \frac{1}{x-b}\right)$$

$$\frac{1}{(x-a)(x^2+cx+d)} = \frac{1}{a^2+ab+c}\left(\frac{1}{x-a} - \frac{x+a+c}{x^2+cx+d}\right)$$

$$\frac{1}{f(x)g(x)} = \frac{q(x)+1}{r\cdot f(x)} - \frac{q(x)}{r\cdot g(x)}, \quad \text{mit } f(x) = x^2+ax+b,\ g(x) = x^2+cx+d$$

$$q(x) = \frac{x}{a-c} + \frac{1}{a-c}\left(c - \frac{b-d}{a-c}\right),\ r = d - \frac{b-d}{a-c}\left(c - \frac{b-d}{a-c}\right)$$

Beispiel

$$\frac{1}{(x-1)^2(x^2+x+1)} = -\frac{x-2}{3(x-1)^2} + \frac{x+1}{3(x^2+x+1)} \quad \text{mit } q(x) = \frac{-1}{3}(x+1),\ r = 1$$

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

16.2.7 Höhere Vielfachheiten

Statt des euklidischen Algorithmus kann man auch einen Ansatz benutzen, um für $g(x) = q(x)p(x)^k$, $p(x)$ teilerfremd zu $q(x)$ ein $r(x)$ mit $\deg r(x) < k \deg p(x)$ so zu bestimmen, dass

$$\frac{f(x}{g(x)} = \frac{r(x)}{p(x)^k} + \frac{h(x)}{q(x)}$$

Nämlich

$$f(x) = r(x)q(x) + h(x)p(x)^k$$

also $p(x)^k$ Teiler von $r(x)q(x) - f(x)$. Geht man zu den Resten $\tilde{f}(x)$ bzw. $\tilde{(q)}(x)$ von $f(x)$ bzw. $q(x)$ bei Division durch $p(x)^k$ über, so gilt

$$r(x)\tilde{q}(x) - \tilde{f}(x) = s(x)p(x)^k$$

mit einem $s(x)$ von $\deg s(x) < k \deg p(x)$. Koeffizientenvergleich liefert ein eindeutig lösbares lineares Gleichungssystem für die Koeffizienten von $r(x)$ und $s(x)$.

Die Partialbrüche zu den Potenzen von $p(x)$ erhält man dann durch

- Fortlaufende Division mit Rest durch $p(x)$ beginnend mit $r(x)$
- (Taylor)Entwicklung an a falls $q(x) = x - a$

Für $K \subseteq \mathbb{C}$ und $p(x) = x - a$ kann man die Partialbrüche auch mittels Differentiation bestimmen. Aus

$$f(x) = A_k + A_{k-1}p(x) + A_2p(x)^{k-2} + A_1p(x)^{k-1} + h(x)p(x)^k$$

folgt

$$(A_{k-m}q(x)p(x)^{k-m})^{(m)} = f^{(m)}(x) - \sum_{l=k-m+1}^{k}(A_lq(x)p(x)^{k-l})^{(m)} + t_m(x)p(x)$$

mit

$$t_m(x)p(x) = \frac{\partial^m}{\partial x^m}(\sum_{l=1}^{k-m-1}A_lq(x)p(x)^{k-l} + h(x)p(x)^k)$$

Durch Berechnung der Ableitungen und Einsetzen von a kann man so $A_k, \dots, A_1$ bestimmen.

16.2.8 Partialbruchzerlegung durch Ansatz

Seien Q, R Polynome mit $Q \not\equiv 0$. Zur Bestimmung einer Stammfunktion der rationalen Funktion $\frac{R}{Q}$ (die auf $\mathbb{R}$ mit Ausnahme der Nullstellen von Q definiert ist) geht man wie folgt vor:

1. Schritt Ist der Grad von R größer oder gleich dem von Q, so liefert eine Polynomdivision von R durch Q Polynome P und S mit

$$\frac{R}{Q} = S + \frac{P}{Q},$$

wobei nun der Grad von P kleiner als der von Q ist. Für das Polynom S kann man stets eine Stammfunktion angeben. Wir betrachten von nun an nur noch P/Q.

2. Schritt Man zerlegt das Nennerpolynom Q multiplikativ in Polynome ersten und zweiten Grades mit reellen Koeffizienten. Dass dies möglich ist, folgt aus nachstehendem Satz

Satz 16.18 (Fundamentalsatz der Algebra) *Für jedes Polynom $Q(x) = \sum_{i=0}^n q_i x^i$ mit $q_i \in \mathbb{R}$ und $q_n \neq 0$ gibt es reelle Zahlen b_i, c_j, d_j sowie natürliche Zahlen k_i, m_j, r und s so, dass*

$$Q(x) = q_n \prod_{i=1}^{r} (x - b_i)^{k_i} \prod_{j=1}^{s} (x^2 + 2c_j x + d_j)^{m_j} \quad \text{für } x \in \mathbb{R} \tag{16.1}$$

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

Zur Bestimmung der b_i, c_j, d_j kann man beispielsweise alle komplexen Nullstellen $\lambda_1, \ldots \lambda_n$ von Q ermitteln. Dann ist $Q(x) = q_n(x - \lambda_1) \ldots (x - \lambda_n)$. Die Terme $(x - \lambda)(x - \overline{\lambda})$ mit $\lambda \notin \mathbb{R}$ werden zu

$$(x - \lambda)(x - \overline{\lambda}) = x^2 - (\lambda + \overline{\lambda})x + |\lambda|^2$$

zusammengefasst. Die exakte Bestimmung der Nullstellen von Q ist oft unmöglich.

3. Schritt Ist die Zerlegung (16.1) gefunden, macht man den Ansatz

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^{r} \sum_{k=1}^{k_i} \frac{A_{ik}}{(x - b_i)^k} + \sum_{j=1}^{s} \sum_{m=1}^{m_j} \frac{B_{jm}x + C_{jm}}{(x^2 + 2c_j x + d_j)^m} \tag{16.2}$$

mit zu bestimmenden Zahlen A_{ik}, B_{jm} und C_{jm}. Hierdurch wird die rationale Funktion P/Q in einfachere rationale Funktionen zerlegt. Falls alle Nullstellen $b_1, \ldots, b_n$ von Q reell und einfach sind, reduziert sich der Ansatz (16.2) auf

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^{n} \frac{A_i}{x - b_i} .$$

Satz 16.19 (Partialbruchzerlegung) *Sei Q wie in* (16.1)*, und sei P ein Polynom, dessen Grad kleiner als der von Q ist. Dann gibt es Zahlen A_{ik}, B_{jm} und C_{jm} so, dass* (16.2) *gilt, und diese Zahlen sind eindeutig bestimmt.*

Die Zahlen A_{ik}, B_{jm}, C_{jm} können beispielsweise ermittelt werden, indem man (16.2) mit Q multipliziert und durch Koeffizientenvergleich ein lineares Gleichungssystem für die gesuchten Größen aufstellt. Auch das Einsetzen der Nullstellen von Q in die entstehenden Polynome kann hilfreich sein.

Achtung: Macht man den Ansatz, ohne darauf zu achten, dass der Grad von P kleiner als der Grad von Q ist, so geben die aus dem Gleichungssystem bestimmten Werte i.A. keine Lösung der Aufgabe, P/Q in Partialbrüche zu zerlegen.

16.2.9 Beispiel

Man bestimme Partialbruchzerlegung von

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

1. Schritt Polynomdivision

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

2. Schritt Faktorisierung des Nennerpolynoms

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

3. Schritt Partialbruchzerlegung. Der Ansatz

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

liefert nach Multiplikation mit $Q(x) = x^4 - x^3 - x + 1 = (x - 1)^2(x^2 + x + 1)$

$$x^3 + x = A_1(x - 1)(x^2 + x + 1) + A_2(x^2 + x + 1) + (Bx + C)(x - 1)^2 \tag{16.3}$$

bzw. nach Ausmultiplizieren und Zusammenfassen

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

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

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

Die Lösung dieses Gleichungssystems ist

$$A_1 = \frac{2}{3}, \quad A_2 = \frac{2}{3}, \quad B = \frac{1}{3}, \quad C = 0.$$

Die zu integrierende Funktion ist also

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

Alternativ hätte man z.B. in (16.3) $x = 1$ einsetzen können und so A_2 sofort gefunden. Dies entspricht einer Anwendung von Lemma 16.13.

Korollar 16.20 *Benutzt man den Ansatz, so bestimmt man den Partialbruch zur höchsten Potenzen eines Linearfaktors durch Einsetzen der Nullstelle im Koeffizientenvergleich.*

16.3 Integration rationaler Funktionen

16.3.1 Integrale rationaler Grundfunktionen

Zu allen in (16.2) vorkommenden Brüchen lassen sich durch partielle Integration und Substitution die Stammfunktionen effektiv bestimmen. Einige der folgenden Regeln müssen dazu wiederholt angewandt werden:

$$
\begin{aligned}
\int \frac{dx}{(x-b)^k} &= \begin{cases} \frac{1}{1-k}\,(x-b)^{1-k} & \text{falls } k>1 \\ \ln|x-b| & \text{falls } k=1\,, \end{cases} \\
\int \frac{dx}{x^2+2cx+d} &= \frac{1}{\sqrt{d-c^2}} \arctan \frac{x+c}{\sqrt{d-c^2}}\,, \\
\int \frac{dx}{(x^2+2cx+d)^m} &= \frac{x+c}{2(m-1)(d-c^2)(x^2+2cx+d)^{m-1}} \\
&\quad + \frac{(2m-3)}{2(m-1)(d-c^2)} \int \frac{dx}{(x^2+2cx+d)^{m-1}} \quad \text{für } m \geq 2\,, \\
\int \frac{\alpha x+\beta}{x^2+2cx+d}\,dx &= \frac{\alpha}{2}\ln(x^2+2cx+d) + (\beta-\alpha c)\int \frac{dx}{x^2+2cx+d}\,, \\
\int \frac{\alpha x+\beta}{(x^2+2cx+d)^m}\,dx &= \frac{-\alpha}{2(m-1)(x^2+2cx+d)^{m-1}} \\
&\quad + (\beta-\alpha c)\int \frac{dx}{(x^2+2cx+d)^{m-1}} \quad \text{für } m\geq 2\,.
\end{aligned}
$$

16.3.2 Beispiel

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

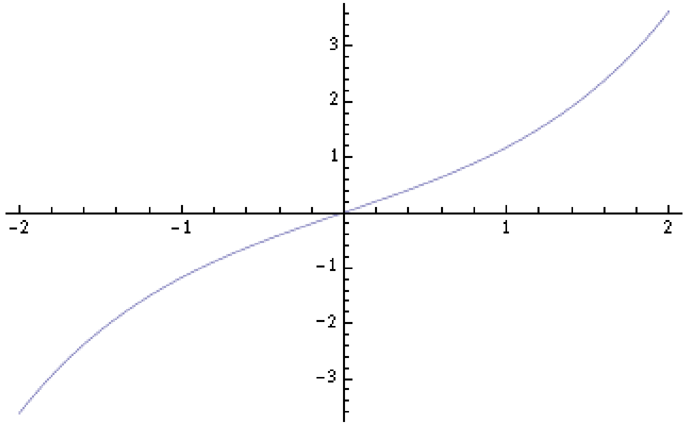
4. Schritt Unbestimmte Integration

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

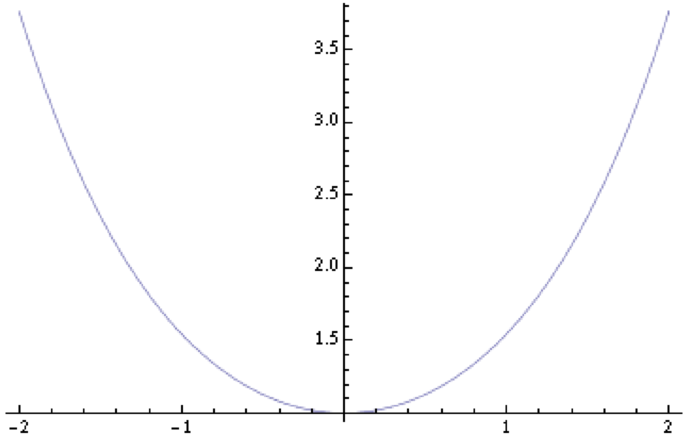
17 Hyperbelfunktionen, Flächen, uneigentliche Integrale und Reihen

17.1 Hyperbelfunktionen

$$\sinh x = \frac{1}{2}(e^x - e^{-x})$$



$$\cosh x \frac{1}{2}(e^x + e^{-x})$$



Es folgt sofort

$$\frac{\partial \sinh x}{\partial x} = \cosh x = \int \sinh x, \quad \frac{\partial \cosh x}{\partial x} = \sinh x = \int \cosh x$$

Die Funktionalgleichungen

$$\sinh(x+y) = \sinh x \cos hx + \cosh x \sinh y$$

$$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y$$

$$\cosh^2 x - \sinh^2 x = 1$$

folgen mit der Funktionalgleichung der Exponentialfunktion oder dem Eindeutigkeitssatz, z.B. $\frac{\partial \cosh^x - sinh^x}{\partial x} = 0$.

$\cosh x > 0$ für alle x also $\sinh x$ auf ganz $\mathbb{R}$ streng monoton wachsend. $\sinh x = 0 \Leftrightarrow e^x = e^{-x} \Leftrightarrow x = 0$ also hat $\sinh x \to \infty$ für $x \to \infty$ $\sinh x \to -\infty$ für $x \to -\infty$, also hat $\sinh x$ auf ganz $\mathbb{R}$ definierte Umkehrfunktion (aus $(e^x)^2 - 2ye^x - 1 = 0$ folgt $e^x = y + \sqrt{y^2+1}$)

$$\operatorname{Arsinh} y = \ln(y + \sqrt{y^2+1}) \text{ auf } \mathbb{R}$$

also

$$\frac{\partial \operatorname{Arsinh} y}{\partial y} = \frac{1}{\cosh x} = \frac{1}{\sqrt{1+y^2}}$$

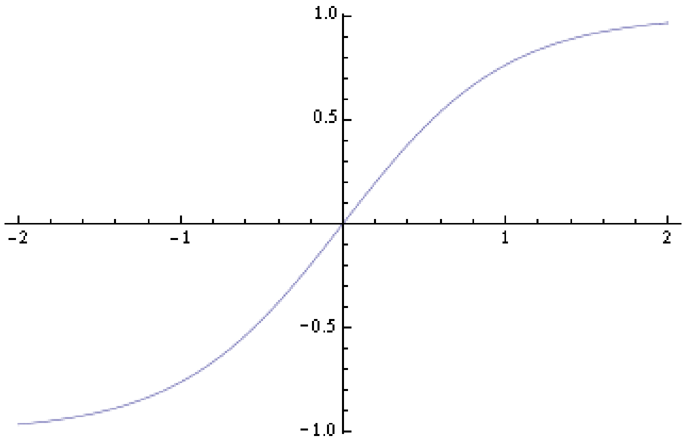
Es folgt, dass $\cosh x$ konvex ist mit Minimum 1 bei $x = 0$ hat und $\cosh x \to \infty$ für $x \to \pm\infty$. Umkehrfunktion

$$\operatorname{arcosh} y = \ln(y + \sqrt{y^2-1} \text{ auf } [1, \infty[$$

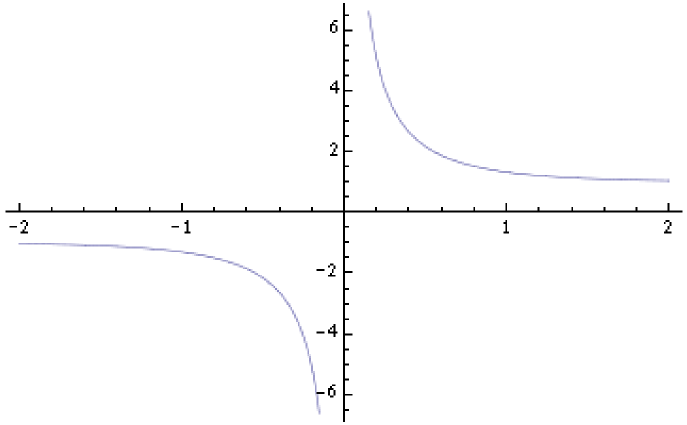
also

$$\frac{\partial \operatorname{Arcosh} y}{\partial y} = \frac{1}{\sinh x} = \frac{1}{\sqrt{y^2-1}}$$

$$\tanh x = \frac{\sinh x}{\cosh x}$$



$$\coth x = \frac{\cosh x}{\sinh x}$$



17.2 Flächeninhalte

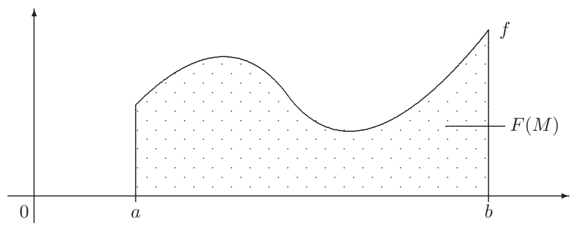
Eines der Motive zur Einführung des Riemann-Integrals war der Wunsch, Flächeninhalte zu definieren und zu berechnen.

Ist $f : [a, b] \to [0, \infty)$ Riemann-integrierbar, so definieren wir als Flächeninhalt der Menge

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

die Zahl

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

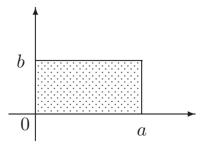


Mit dieser Definition lassen sich auch die Inhalte komplizierter Mengen definieren und berechnen, wenn man akzeptiert, daß der Flächeninhalt die folgenden (aus unserer Erfahrung heraus plausiblen) Eigenschaften aufweist:

(a) Geht M' aus M durch Verschiebung, Drehung oder Spiegelung an einer Geraden hervor, so ist $F(M') = F(M)$.

(b) Kann man M in zwei disjunkte Teilmengen A, B zerlegen, von denen jede einen Flächeninhalt besitzt, so ist $F(M) = F(A) + F(B)$.

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



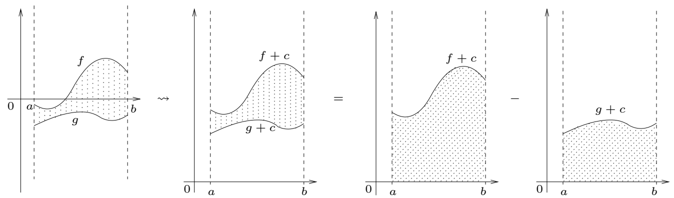
Beispiel 2 Die Dirichletfunktion

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

ist auf keinem Intervall $[a,b]$ mit $a < b$ Riemann-integrierbar. Unsere Definition erlaubt es daher nicht, der Menge $M = \{(x,y) \in \mathbb{R}^2, 0 \le x \le 1,\ 0 \le y \le f(x)\}$ einen Flächeninhalt zuzuschreiben. ■

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

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



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

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

Da wir dem Funktionsgraphen von g den Flächeninhalt 0 zuordnen können, können wir $\int_a^b \big(f(x) - g(x)\big)dx$ auch als Flächeninhalt der Menge

$$\{(x,y) \in \mathbb{R}^2 :\ a \le x \le b,\ g(x) \le y \le f(x)\}$$

betrachten. Eine genauere Untersuchung des Begriffes Flächeninhalt erfolgt im Rahmen der Maßtheorie. ■

Beispiel 4 Oft ist der Graph einer Funktion f in Parameterdarstellung gegeben, etwa

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

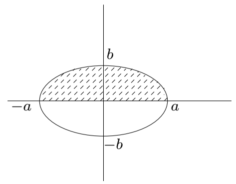
Unter entsprechenden Voraussetzungen an x und y Substitutionsregel) gilt dann

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

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

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

eine Ellipse beschrieben. Für ihren Flächeninhalt findet man (weil $f(x)$ stetig und $x(t)$ stetig differenzierbar ist)

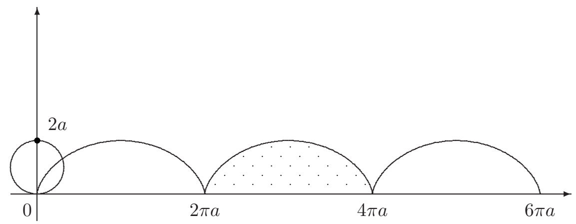


$$\begin{aligned} F(M) &= 2\int_{-a}^{a} f(x)dx = 2\int_\pi^0 y(t)\,\dot{x}(t)dt = 2ab\int_\pi^0 \sin t(-\sin t)dt \\ &= 2ab\int_0^\pi \sin^2 t\,dt = 2ab\Big(-\frac{\sin t\cos t}{2} + \frac{1}{2}t\Big)\Big|_0^\pi = \pi ab\,. \end{aligned}$$

Beispiel 5 Durch

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

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



Für die Fläche unter einem Zykloidenbogen findet man (wieder durch formale Rechnung, die man wie in Beispiel 4 präzisieren kann)

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

17.3 Uneigentliche Integrale

17.3.1 Konvergenz

Bisher haben wir das Integral $\int_a^b f(x)dx$ definiert unter der Voraussetzung, dass f eine beschränkte Funktion auf dem beschränkten abgeschlossenen Intervall $[a,b]$ ist. Sind diese Voraussetzungen nicht alle erfüllt, so lassen sich in einigen Fällen durch nahe liegende Grenzwertbildungen so genannte *uneigentliche* Riemann-Integrale definieren.

Wir beginnen mit dem Fall eines unendlichen Integrationsintervalles.

Definition 17.1 *Die Funktion $f : [a,\infty) \to \mathbb{R}$ sei auf jedem Intervall $[a,t]$ mit $t > a$ Riemann-integrierbar. Existiert der Grenzwert*

$$\lim_{t\to\infty} \int_a^t f(x)dx\,, \tag{17.1}$$

so bezeichnen wir ihn mit $\int_a^\infty f(x)dx$ und nennen ihn uneigentliches Riemann-Integral *von f auf $[a,\infty)$. Man sagt auch, dass f* uneigentlich Riemann-integrierbar *ist oder dass $\int_a^\infty f(x)dx$* konvergiert. *Existiert der Grenzwert* (17.1) *nicht, so heißt $\int_a^\infty f(x)dx$* divergent. *Schließlich heißt $\int_a^\infty f(x)dx$* absolut konvergent, *wenn das uneigentliche Integral $\int_a^\infty |f(x)|dx$ konvergiert.*

Analoge Definitionen trifft man für

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

sowie für

$$\int_{-\infty}^\infty f(x)dx = \int_{-\infty}^a f(x)dx + \int_a^\infty f(x)dx = \lim_{s\to-\infty} \int_s^a f(x)dx + \lim_{t\to+\infty} \int_a^t f(x)dx.$$

Wie bei Reihen gelten die folgenden Aussagen.

Satz 17.2 *Konvergiert das uneigentliche Riemann-Integral $\int_a^\infty f(x)dx$ absolut, so konvergiert es.*

17.3.2 Vergleichkriterium

Satz 17.3 (Vergleichskriterium) *Die Funktionen $f, g : [a,\infty) \to \mathbb{R}$ seien Riemann-integrierbar auf jedem Intervall $[a,t]$ mit $t > a$. Ist $|f(x)| \le g(x)$ für alle $x \ge a$ und existiert $\int_a^\infty g(x)dx$, so ist das uneigentliche Integral $\int_a^\infty f(x)dx$ absolut konvergent, und es gilt*

$$\left|\int_a^\infty f(x)dx\right| \le \int_a^\infty |f(x)|dx \le \int_a^\infty g(x)dx.$$

Ist dagegen $0 \le g(x) \le f(x)$ für alle $x \ge a$ und divergiert $\int_a^\infty g(x)dx$, so divergiert auch $\int_a^\infty f(x)dx$.

Beispiel 1 Es ist

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

■

Beispiel 2 Es ist

$$\lim_{s\to-\infty}\int_s^0 \frac{1}{1+x^2}\,dx = \lim_{s\to-\infty} \arctan x\Big|_s^0 = \lim_{s\to-\infty} -\arctan s = \frac{\pi}{2}$$

und daher

$$\int_{-\infty}^0 \frac{1}{1+x^2}\,dx = \frac{\pi}{2}.$$

Beispiel 3 Wir zeigen, dass $\int_0^\infty x^n e^{-x}dx = n!$ für $n \ge 0$. Es ist nämlich

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

eine Stammfunktion des Integranden, wie man durch Differenzieren leicht bestätigt. Außerdem ist $\lim_{x\to\infty} F(x) = 0$, denn es ist

$$\lim_{x\to\infty}\frac{x^k}{e^x} = 0 \quad \text{für jedes } k \ge 0,$$

wie man mit der l'Hospitalschen Regel sofort sieht. Also ist

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

Beispiel 4 Wir zeigen, dass $\int_0^\infty \frac{\sin x}{x}\,dx$ konvergiert. An der Stelle 0 ist der Integrand nicht definiert. Wegen $\lim_{x\to0}\frac{\sin x}{x} = 1$ lässt sich die Funktion $x \mapsto \frac{\sin x}{x}$ aber zu einer auf $[0,\infty)$ stetigen Funktion fortsetzen, wenn man ihren Wert an der Stelle 0 durch 1 festlegt. Insbesondere existiert $\int_0^1 \frac{\sin x}{x}\,dx$ als (eigentliches) Riemann-Integral und wir müssen noch die Konvergenz des Integrals $\int_1^\infty \frac{\sin x}{x}\,dx$ zeigen. Partielle Integration liefert für jedes $t > 1$

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

Offenbar existiert der Grenzwert

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

und es verbleibt, die Existenz des Grenzwertes $\lim_{t\to\infty}\int_1^t \frac{\cos x}{x^2}\,dx$ bzw. die Konvergenz des uneigentlichen Integrals $\int_1^\infty \frac{\cos x}{x^2}\,dx$ zu zeigen. Wegen

$$\left|\frac{\cos x}{x^2}\right| \le \frac{1}{x^2} \quad \text{für } x \ge 1$$

und Beispiel 1 existiert dieses uneigentliche Integral nach dem Vergleichskriterium. ■

17.3.3 Integrale unbeschränkter Funktionen

Wir sehen uns eine weitere Verallgemeinerung des Integralbegriffes auf Funktionen an, die auf ganz $[a, b]$ definiert und gegebenenfalls unbeschränkt sind.

Definition 17.4 *Die Funktion* $f : [a, b) \to \mathbb{R}$ *sei für jedes* $c \in (a, b)$ *auf* $[a, c]$ *Riemann-integrierbar. Existiert der Grenzwert*

$$\lim_{c \nearrow b} \int_a^c f(x)\, dx,$$

so bezeichnet man ihn mit $\int_a^b f(x)\, dx$ *und nennt* f *uneigentlich integrierbar auf* $[a, b]$.

Ganz analog definiert man diesen Begriff für Funktionen auf links halboffenen Intervallen.

Beispiel 6 Es ist

$$\begin{aligned}
\int_0^1 \frac{1}{x^\alpha}\, dx &= \lim_{s \searrow 0} \int_s^1 \frac{1}{x^\alpha}\, dx = \lim_{s \searrow 0} \begin{cases} \frac{1}{1-\alpha} - \frac{s^{1-\alpha}}{1-\alpha} & \text{falls } \alpha \neq 1 \\ -\ln s & \text{falls } \alpha = 1 \end{cases} \\
&= \begin{cases} \infty & \text{falls } \alpha \geq 1 \quad \text{(Divergenz)} \\ \frac{1}{1-\alpha} & \text{falls } \alpha < 1 \quad \text{(Konvergenz).} \end{cases}
\end{aligned}$$

■

Beispiel 7 Es ist

$$\begin{aligned}
\int_0^1 \ln x\, dx &= \lim_{s \searrow 0} \int_s^1 \ln x\, dx = \lim_{s \searrow 0} (x \ln x - x)\Big|_s^1 \\
&= -1 - \lim_{s \searrow 0} (s \ln s - s) = -1.
\end{aligned}$$

■

Beispiel 8 Es ist

$$\begin{aligned}
\int_0^1 \frac{1}{\sqrt{1-x^2}}\, dx &= \lim_{t \nearrow 1} \int_0^t \frac{1}{\sqrt{1-x^2}}\, dx = \lim_{t \nearrow 1} \arcsin x \Big|_0^t \\
&= \lim_{t \nearrow 1} \arcsin t - \arcsin 0 = \arcsin 1 = \frac{\pi}{2}.
\end{aligned}$$

■

17.4 Reihen

17.4.1 Definition und Beispiele

Mit Hilfe der Körperaxiome lassen sich Summen *endlich* vieler Zahlen erklären. Wir untersuchen in diesem Abschnitt Summen *unendlich* vieler Zahlen - so genannte Reihen. Dazu betrachten wir Reihen als spezielle Folgen.

Definition 17.5 *Sei* $(a_n)_{n\geq 0}$ *eine Folge reeller Zahlen, und sei* $s_n := \sum_{k=0}^n a_k$. *Die Folge* $(s_n)_{n\geq 0}$ *heißt die zu* (a_n) *gehörende Reihe. Die Zahlen* a_n *heißen* Glieder *der Reihe, und die* s_n *ihre* Partialsummen. *Ist die Folge* $(s_n)_{n\geq 0}$ *konvergent und* s *ihr Grenzwert, so heißt die Reihe* konvergent, *die Zahl* s *heißt ihre* Summe, *und man schreibt* $s = \sum_{k=0}^\infty a_k$. *Nichtkonvergente Reihen heißen divergent.*

Eine Reihe ist also die Folge ihrer Partialsummen. Häufig wählt man $\sum_{k=0}^\infty a_k$ auch als *Bezeichnung* für die Reihe (s_n). Aus dem Kontext wird in der Regel klar, ob $\sum_{k=0}^\infty a_k$ für die Reihe selbst oder für ihre Summe steht.

Beispiel 12 Die geometrische Reihe. Sei $q \in \mathbb{R}\backslash\{1\}$ und $a_n = q^n$ für $n \geq 0$. Aus Beispiel 2, Abschnitt 1.1, wissen wir, dass

$$s_n = \sum_{k=0}^n a_k = \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}.$$

Damit ist klar: die *geometrische Reihe* $\sum_{k=0}^\infty q^k$ konvergiert genau dann, wenn $|q| < 1$. In diesem Fall ist $\frac{1}{1-q}$ ihre Summe. ■

Beispiel 13 Die harmonische Reihe. Das ist die Reihe $\sum_{k=1}^\infty \frac{1}{k}$. Sie divergiert, denn für jedes $n \geq 1$ ist

$$|s_{2n} - s_n| = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} \geq n\, \frac{1}{2n} = \frac{1}{2},$$

d.h. (s_n) ist keine Cauchyfolge und deshalb erst recht nicht konvergent. ■

Beispiel 14 Aus den Beispielen 10 und 11 wissen wir: die so genannte *Leibniz-Reihe* $\sum_{k=1}^\infty (-1)^{k-1} \frac{1}{k}$ konvergiert (gegen $\ln 2$), und auch die Reihe $\sum_{k=0}^\infty \frac{1}{k!}$ konvergiert (gegen e). ■

Beispiel 15 Für die Reihe $\sum_{k=2}^\infty \frac{1}{k(k-1)}$ haben wir

$$s_n = \sum_{k=2}^n \frac{1}{k(k-1)} = \sum_{k=2}^n \left(\frac{1}{k-1} - \frac{1}{k} \right) = 1 - \frac{1}{n}.$$

Also konvergiert diese Reihe, und ihre Summe ist 1. ■

17.4.2 Konvergenzkriterien

Da die Konvergenz von Reihen über die Konvergenz der Folge ihrer Partialsummen erklärt ist, kann man Konvergenzkriterien für Folgen auf Reihen übertragen.

Satz 17.6 (Cauchy-Kriterium für Reihen) *Die Reihe* $\sum_{k=0}^\infty a_k$ *konvergiert genau dann, wenn es für jedes* $\varepsilon > 0$ *ein* $N(\varepsilon) \in \mathbb{N}$ *so gibt, dass*

$$\left| \sum_{k=n+1}^m a_k \right| = |a_{n+1} + a_{n+2} + \cdots + a_m| < \varepsilon \quad \textit{für alle } m > n \geq N(\varepsilon).$$

Man beachte, dass

$$s_m - s_n = \sum_{k=0}^{m} a_k - \sum_{k=0}^{n} a_k = \sum_{k=n+1}^{m} a_k.$$

Aus dem Cauchy-Kriterium bekommt man das folgende *notwendige Konvergenzkriterium*, indem man $m = n+1$ setzt.

Wenn $\sum_{k=0}^{\infty} a_k$ konvergiert, dann ist $\lim_{n\to\infty} a_n = 0$.

Die Umkehrung gilt natürlich *nicht*, wie die harmonische Reihe zeigt.

Ein spezielles Konvergenzkriterium hat man für *alternierende* Reihen:

Satz 17.7 (Leibniz-Kriterium für alternierende Reihen) *Für die Reihe $\sum_{k=0}^{\infty}(-1)^k b_k$ gelte $b_k \geq 0$ und $b_k \geq b_{k+1}$ für alle $k \geq 0$, und es sei $\lim_{k\to\infty} b_k = 0$. Dann ist diese Reihe konvergent.*

Dieses Kriterium kann man mit der gleichen Idee beweisen, die wir in Beispiel 11 für die alternierende Reihe $\sum_{k=1}^{\infty}(-1)^{k-1}\frac{1}{k} = \sum_{k=0}^{\infty}(-1)^k\frac{1}{k+1}$ genutzt haben.

Das Leibniz-Kriterium liefert z.B. die Konvergenz der Reihen

$$\sum_{k=1}^{\infty}(-1)^{k-1}\frac{1}{k^2} \quad \text{und} \quad \sum_{k=0}^{\infty}(-1)^k\frac{1}{k!}.$$

Schauen wir uns noch Summen und Differenzen von Reihen an. Aus Satz 12.4 erhalten wir sofort:

Satz 17.8 *Sind $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ konvergente Reihen, so ist für beliebige Zahlen α und β auch die Reihe $\sum_{k=0}^{\infty}(\alpha a_k + \beta b_k)$ konvergent, und ihre Summe ist*

$$\sum_{k=0}^{\infty}(\alpha a_k + \beta b_k) = \alpha\sum_{k=0}^{\infty} a_k + \beta\sum_{k=0}^{\infty} b_k.$$

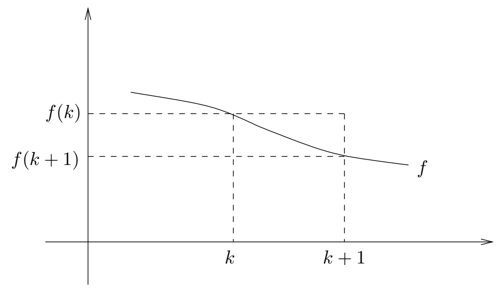
Man beachte aber, dass man Reihen nicht beliebig umordnen darf (während man bei endlichen Summen das Kommutativgesetz hat). Auch Produkte konvergenter Reihen bereiten gewisse Schwierigkeiten. Diese Schwierigkeiten lassen sich vermeiden, wenn man einen stärkeren Konvergenzbegriff zu Grunde legt.

17.4.3 Integralkriterium

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

Satz 17.9 *Sei $f : [1,\infty) \to [0,\infty)$ monoton fallend. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} f(n)$ genau dann, wenn das Integral $\int_1^{\infty} f(x)dx$ konvergiert.*

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



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

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

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

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

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

Beispiel 5 Aus Beispiel 1 (uneigentl. Integrale) wissen wir, dass $\int_1^{\infty} x^{-\alpha}dx$ für alle $\alpha > 1$ konvergiert. Also konvergiert $\sum_{n=1}^{\infty}\frac{1}{n^{\alpha}}$ für alle $\alpha > 1$. ■

17.5 Absolut konvergente Reihen

17.5.1 Definition

Definition 17.10 *Die Reihe $\sum_{k=0}^{\infty} a_k$ heißt absolut konvergent, wenn die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergiert.*

Beispielsweise ist die Reihe $\sum_{k=1}^{\infty}(-1)^k\frac{1}{k}$ konvergent (Leibnitz-Kriterium), jedoch nicht absolut konvergent (harmonische Reihe). Dagegen ist $\sum_{k=0}^{\infty}(-1)^k\frac{1}{k!}$ eine konvergente (Leibniz-Kriterium) und auch absolut konvergente Reihe (Reihe für e). Reihen mit ausschließlich nichtnegativen Gliedern sind genau dann absolut konvergent, wenn sie konvergieren.

Satz 17.11 *Absolut konvergente Reihen sind konvergent.*

Beweis Die Reihe $\sum_{k=0}^{\infty} a_k$ sei absolut konvergent, und wir setzen $s_n := \sum_{k=0}^{n} a_k$ sowie $S_n := \sum_{k=0}^{n} |a_k|$. Für $m > n$ ist dann

$$|s_m - s_n| = |a_{n+1} + a_{n+2} + \ldots + a_m| \leq |a_{n+1}| + \ldots + |a_m| = S_m - S_n.$$

Da die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergiert, gibt es für jedes $\varepsilon > 0$ ein N so, dass $|S_m - S_n| < \varepsilon$ für alle $m, n \geq N$. Dann ist aber auch $|s_m - s_n| < \varepsilon$ für alle $m, n \geq N$. Das Cauchy-Kriterium liefert die Konvergenz der Reihe $\sum_{k=0}^{\infty} a_k$. ■

17.5.2 Kriterien für absolute Konvergenz

Wir sehen uns nun einige Kriterien für die absolute Konvergenz von Reihen an.

Satz 17.12 (Vergleichskriterium) (a) *Ist die Reihe $\sum_{k=0}^{\infty} b_k$ absolut konvergent und gilt $|a_k| \le |b_k|$ für alle k, so ist auch die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergent.*

(b) *Ist die Reihe $\sum_{k=0}^{\infty} b_k$ divergent und ist $0 \le b_k \le a_k$ für alle k, so divergiert auch die Reihe $\sum_{k=0}^{\infty} a_k$.*

Im Fall (a) heißt $\sum_{k=0}^{\infty} |b_k|$ eine *konvergente Majorante* für $\sum_{k=0}^{\infty} a_k$, und im Fall (b) eine *divergente Minorante*.

Aussage (a) kann man sich z.B. so klarmachen: Für jede Partialsumme s_n von $\sum_{k=0}^{\infty} |a_k|$ gilt:

$$s_n = |a_0| + |a_1| + \ldots + |a_n| \le |b_0| + |b_1| + \ldots + |b_n| \le \sum_{k=0}^{\infty} |b_k| < \infty.$$

Die Folge (s_n) ist also nach oben beschränkt, und sie ist offenbar monoton wachsend. Nach dem Monotoniekriterium konvergiert diese Folge. ■

Für die Anwendung dieses Kriteriums ist es offenbar wichtig, einen großen Vorrat an Vergleichsreihen zu besitzen.

Beispiel 16 Die Glieder der Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ lassen sich nach oben abschätzen durch

$$\frac{1}{n^2} \le \begin{cases} 1, & \text{wenn } n = 1 \\ \frac{1}{n(n-1)}, & \text{wenn } n > 1. \end{cases}$$

Die Reihe $1 + \sum_{k=2}^{\infty} \frac{1}{k(k-1)}$ ist aber konvergent, wie wir aus Beispiel 15 wissen. Also konvergiert auch die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$. Ihre Summe (die wir auf diese Weise nicht erhalten können) ist übrigens $\pi^2/6$. Diese Reihe kann nun ihrerseits wieder als konvergente Majorante für Reihen wie $\sum_{k=1}^{\infty} \frac{1}{k^r}$ mit $r \ge 2$ dienen. ■

Beispiel 17 Die Reihe $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ divergiert, da die harmonische Reihe eine divergente Minorante für diese Reihe ist. ■

Satz 17.13 (Quotientenkriterium) *Sei (a_n) eine Folge reeller Zahlen mit $a_n \ne 0$ für alle hinreichend großen n.*

(a) *Wenn es ein $q \in (0,1)$ und ein $N \in \mathbb{N}$ so gibt, dass*

$$\left|\frac{a_{k+1}}{a_k}\right| \le q < 1 \quad \textit{für alle } k \ge N,$$

so ist die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergent.

(b) *Gibt es ein $N \in \mathbb{N}$ so, dass*

$$\left|\frac{a_{k+1}}{a_k}\right| \ge 1 \quad \textit{für alle } k \ge N,$$

so divergiert die Reihe $\sum_{k=0}^{\infty} a_k$.

Falls der Grenzwert $q^* := \lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|$ existiert, so läßt sich das Quotientenkriterium auch so fassen:

Die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert absolut für $q^* < 1$, und sie divergiert für $q^* > 1$.

Im Fall $q^* = 1$ ist keine Aussage möglich, wie die Reihen

$$\sum_{n=1}^{\infty} \frac{1}{n} \quad \text{und} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} \tag{17.2}$$

zeigen. In beiden Fällen ist $q^* = 1$. Die erste Reihe divergiert aber, während die zweite konvergiert.

Satz 17.14 (Wurzelkriterium) *Sei (a_n) eine Folge reeller Zahlen.*

(a) *Wenn es ein $q \in (0,1)$ und ein $N \in \mathbb{N}$ so gibt, dass*

$$\sqrt[k]{|a_k|} \le q < 1 \quad \textit{für alle } k \ge N,$$

so ist die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergent.

(b) *Gibt es ein $N \in \mathbb{N}$ so, dass*

$$\sqrt[k]{|a_k|} \ge 1 \quad \textit{für alle } k \ge N,$$

so divergiert die Reihe $\sum_{k=0}^{\infty} a_k$.

Existiert der Grenzwert $q^* := \lim_{k\to\infty} \sqrt[k]{|a_k|}$, so kann man das Wurzelkriterium wie folgt formulieren:

Die Reihe $\sum_{k=0}^{\infty} a_k$ ist absolut konvergent für $q^* < 1$ und divergent für $q^* > 1$.

Für $q^* = 1$ ist wieder keine Entscheidung möglich, wie die Reihen (17.2) zeigen. Die Sätze 17.13 und 17.14 werden mit dem Vergleichskriterium bewiesen, wobei die geometrische Reihe als Vergleichsreihe dient. Ist etwa

$$\sqrt[k]{|a_k|} \le q < 1 \quad \text{für } k \ge N,$$

so ist $|a_k| \le q^k$, und $\sum_{k=N}^{\infty} q^k$ ist eine konvergente Majorante für $\sum_{k=N}^{\infty} a_k$. ■

Beispiel 18 Für die Reihe $\sum_{k=1}^{\infty} a_k$ mit $a_k = \frac{(k!)^2}{(2k)!}$ ist

$$\frac{a_{k+1}}{a_k} = \frac{(k+1)!\,(k+1)!\,(2k)!}{(2k+2)!\,k!\,k!} = \frac{(k+1)^2}{(2k+1)(2k+2)} = \frac{k^2+2k+1}{4k^2+6k+2}.$$

Wegen $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \frac{1}{4} < 1$ konvergiert diese Reihe. ■

Beispiel 19 Für alle $x \in \mathbb{R}$ konvergiert die Reihe $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ absolut, denn

$$\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty} \frac{|x|}{k+1} = 0.$$

17.5.3 Produkt von Reihen

Wir sehen uns nun noch Produkte von Reihen an. Das Produkt der beiden endlichen Summen $a_0 + \ldots + a_n$ und $b_0 + \ldots + b_m$ ist gleich der Summe über alle Produkte $a_i b_j$ mit $0 \le i \le n$ und $0 \le j \le m$. Analog enstehen beim formalen Multiplizieren der Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ die Produkte

$$\begin{array}{cccc} a_0b_0 & a_0b_1 & a_0b_2 & \ldots \\ a_1b_0 & a_1b_1 & a_1b_2 & \ldots \\ a_2b_0 & a_2b_1 & a_2b_2 & \ldots \\ \vdots & \vdots & \vdots & \end{array} \tag{17.3}$$

Es ist keineswegs klar, wie diese Produkte in einer „neuen“ Reihe $\sum_{n=0}^{\infty} c_n$, dem „Produkt“ der Reihen $\sum a_k$ und $\sum b_k$, angeordnet werden sollen. Man kann aber zeigen, dass es auf die Art der Anordnung überhaupt nicht ankommt, falls beide Reihen absolut konvergieren. Insbesondere für Potenzreihen, die wir später kennen lernen, ist folgende Variante des Durchlaufens der Produkte (17.3) interessant

$$\begin{array}{ccccc} a_0b_0 & & a_0b_1 & & a_0b_2 \\ & \diagup & & \diagup & \\ a_1b_0 & & a_1b_1 & & a_1b_2 \\ & \diagup & & & \\ a_2b_0 & & a_2b_1 & & a_2b_2 \end{array}$$

Die Summe der Elemente der n-ten Diagonale ist gerade $c_n := \sum_{k=0}^{n} a_{n-k} b_k$.

Definition 17.15 *Sind* $\sum_{n=0}^{\infty} a_n, \sum_{n=0}^{\infty} b_n$ *zwei Reihen, so heißt die Reihe* $\sum_{n=0}^{\infty} c_n$ *mit* $c_n := \sum_{k=0}^{n} a_{n-k} b_k$ *das* Cauchy-Produkt *der Ausgangsreihen.*

Satz 17.16 *Die Reihen* $\sum_{n=0}^{\infty} a_n$ *und* $\sum_{n=0}^{\infty} b_n$ *seien absolut konvergent. Dann ist auch ihr Cauchy-Produkt* $\sum_{n=0}^{\infty} c_n$ *absolut konvergent, und es gilt*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n \right) \left(\sum_{n=0}^{\infty} b_n \right).$$

Beispiel 20 Für $x, y \in \mathbb{R}$ sind die Reihen $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ und $\sum_{n=0}^{\infty} \frac{y^n}{n!}$ absolut konvergent (vgl. Beispiel 19). Wir berechnen das Cauchy-Produkt dieser Reihen:

$$\begin{aligned} c_n &= \sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k y^{n-k} \\ &= \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \frac{(x+y)^n}{n!}, \end{aligned}$$

wobei wir die binomische Formel benutzt haben. Nach Satz 17.16 ist

$$\left(\sum_{k=0}^{\infty} \frac{x^k}{k!} \right) \left(\sum_{k=0}^{\infty} \frac{y^k}{k!} \right) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}.$$

Inhaltsverzeichnis