Analysis I

Wintersemester 2002/03

W. Ebeling

1 Zahlensysteme

Die Analysis befasst sich mit Funktionen von reellen Zahlen. Zunächst wird eine Übersicht über Zahlensysteme gegeben.

Zum Abzählen bedient man sich der natürlichen Zahlen:

$$\mathbb{N} = \{0, 1, 2, \dots\}$$

bezeichnet die *Menge der natürlichen Zahlen*. (Man beachte, dass wir die Null mit hinzunehmen!) Von L. Kronecker (1823-1891) stammt der Ausspruch: Die natürlichen Zahlen sind vom lieben Gott gemacht, alles andere ist Menschenwerk.

Das Dezimalsystem, in dem wir die natürlichen Zahlen aufschreiben

$$\{0, 1, 2, \dots, 10, 11, 12, \dots, 100, 101, 102, \dots\},$$

wurde in Indien in der Zeit 300 v. – 600 n. Chr. eingeführt. Der Hindu-Mathematiker Brahmagupta rechnete 628 bereits mit den ganzen Zahlen:

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

bezeichnet die *Menge der ganzen Zahlen*.

Aus den ganzen Zahlen lassen sich die rationalen Zahlen konstruieren:

$$\mathbb{Q} = \left\{ \frac{p}{q} \,\middle|\, p, q \in \mathbb{Z}, q \neq 0 \right\}$$

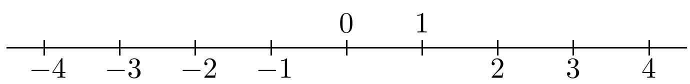
ist die *Menge der rationalen Zahlen*.

Schließlich hat man noch die *Menge der reellen Zahlen* $\mathbb{R}$, deren Elemente in "eineindeutiger" Weise den Punkten einer Geraden entsprechen.

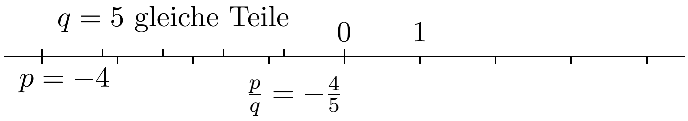
Zwischen den bisher genannten Mengen hat man die folgenden Teilmengenbeziehungen:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Wie erhält man nun die eineindeutige Korrespondenz zwischen den reellen Zahlen $\mathbb{R}$ und einer Geraden? Dazu wählen wir zunächst zwei Punkte 0 und 1 auf der Geraden und tragen dann äquidistant die übrigen ganzen Zahlen ab.

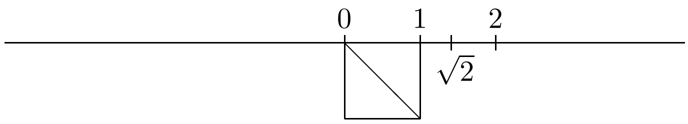


Um nun einen Punkt für die rationale Zahl $\frac{p}{q}$, $p, q \in \mathbb{Z}$, $q > 0$, festzulegen, teilt man die Strecke $\overline{0p}$ auf der Geraden in q gleiche Teile. Der am dichtesten bei 0 gelegene Teilpunkt steht dann für $\frac{p}{q}$.



Man sieht dann ein, dass die rationalen Zahlen "dicht" auf der Geraden liegen. Das heißt, in jeder noch so kleinen Umgebung eines jeden Punktes der Geraden findet man eine rationale Zahl (genauer gesagt, einen Punkt, der eine rationale Zahl repräsentiert). Um einen solchen Punkt zu finden, muss man q genügend groß und p passend wählen.

Allerdings hat man bisher keineswegs alle Punkte der Geraden erwischt. Die Diagonale des Einheitsquadrats hat die Länge $\sqrt{2}$.



Satz 1.1 *Die Zahl $\sqrt{2}$ ist irrational (d.h. nicht rational).*

Beweis. Dieser Beweis ist ein Beispiel für einen *indirekten Beweis.* Wir nehmen an, dass $\sqrt{2}$ rational ist, und leiten daraus einen Widerspruch her. Wir nehmen also an

$$\sqrt{2} = \frac{p}{q}, \quad p, q \in \mathbb{Z}, \quad q \neq 0.$$

Ohne Einschränkung können wir zusätzlich annehmen, dass dieser Bruch bereits gekürzt ist. Insbesondere können wir annehmen, dass p und q nicht beide gerade sind.

Aus $\sqrt{2} = \frac{p}{q}$ folgt aber

$$2q^2 = p^2.$$

Daraus folgt p^2 gerade. Da aber Quadrate ungerader Zahlen wieder ungerade sind, muss also schon p gerade sein, d.h. $p = 2p'$ für ein $p' \in \mathbb{Z}$. Setzen wir dies für p in die obige Gleichung ein, so folgt

$$2q^2 = 4(p')^2.$$

Teilen wir beide Seiten der Gleichung durch 2, so erhalten wir

$$q^2 = 2(p')^2.$$

Daraus folgt nun aber, dass auch q^2 und damit q gerade sein muss. Dies ist aber ein Widerspruch zu unserer Annahme. Also war unsere Annahme, dass $\sqrt{2}$ rational ist, falsch. $\square$

Wie schreibt man nun reelle Zahlen hin? Man kann eine reelle Zahl durch eine periodische oder nicht periodische Dezimalentwicklung darstellen, z.B.

$$\sqrt{2} = 1,4142135\ldots$$

Diese Darstellung ist noch gar nicht so alt. RHAETICUS stellte 1550 eine Sinus-Tabelle auf. Da das Komma erst 1660 erfunden wurde, behalf sich Rhaeticus dadurch, dass er $10^{10} \cdot \sin\varphi$ anstelle von $\sin\varphi$ angab. Rhaeticus berechnete den Sinus auf 10 Stellen genau. Insbesondere hatte er also auch $\sqrt{2} = 2 \cdot \sin 45°$ gut approximiert.

Ein weiteres Beispiel einer reellen Zahl ist die Zahl π. Ein Kreis vom Durchmesser d hat den Umfang $\pi \cdot d$. Die Zahl π ist eine sogenannte transzendente Zahl. Eine Zahl $r \in \mathbb{R}$ heißt *algebraisch* genau dann, wenn r Nullstelle irgendeines Polynoms $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ mit ganzzahligen Koeffizienten $a_n, a_{n-1}, \ldots, a_0$ und $a_n \neq 0$ ist. Eine Zahl $r \in \mathbb{R}$ heißt *transzendent* genau dann, wenn r nicht algebraisch ist. Z. B. sind alle rationalen Zahlen $\frac{p}{q}$, p, q wie oben, algebraisch, da $\frac{p}{q}$ Nullstelle von $qx - p$ ist. Die Zahl $\sqrt{2}$ ist algebraisch, da sie Nullstelle von $x^2 - 2$ ist. Für π zeigte F. V. LINDEMANN ($\star$ 1852 Hannover, † 1939) 1882, dass π transzendent ist. Die Transzendenz von π hängt mit der Quadratur des Kreises insofern zusammen, als sich mit Zirkel und Lineal nur algebraische Zahlen konstruieren lassen. Deswegen ist die Quadratur des Kreises unmöglich. Genaueres lernt man in der Vorlesung Algebra I.

Die reellen Zahlen sind grundlegend für die Analysis. Wie führt man die reellen Zahlen streng mathematisch ein? Unsere bisherigen Betrachtungen waren nämlich nur heuristisch. Die "Zahlengerade" haben wir nicht präzise definiert. Dazu müssten wir genau erklären, was wir unter einer Geraden, einem Punkt und dem Abstand von zwei Punkten verstehen. Eine andere Möglichkeit wäre, die reellen Zahlen direkt durch "Dezimalentwicklungen" einzuführen. Dabei ergeben sich aber folgende Probleme:

- Die Bedeutung von $\ldots$ müsste erklärt werden.
- Summe und Produkt von Dezimalentwicklungen müssten erklärt werden.
- Die Zahl 10 ist aus mathematischer Sicht nicht ausgezeichnet.

Wir gehen deshalb folgenden Weg, der charakteristisch für die Mathematik ist: Wir stellen ein System von Gesetzen zusammen, die von den reellen Zahlen - was immer das sein mag - erfüllt werden und auf die allein bei Beweisen zurückgegriffen wird. Die ausgewählten Gesetze nennt man *Axiome*.

2 Die Axiome der reellen Zahlen (1. Teil)

Die Menge der reellen Zahlen $\mathbb{R}$ hat zwei Verknüpfungen $+$ und $\cdot$, d.h. $a, b \in \mathbb{R}$ ist eindeutig zugeordnet

$$\begin{aligned} a + b &\in \mathbb{R} \quad \text{(Summe)}, \\ a \cdot b &\in \mathbb{R} \quad \text{(Produkt)}. \end{aligned}$$

Diese Verknüpfungen erfüllen die folgenden Axiome:

Axiome der Addition:

(A1) Für alle $a, b \in \mathbb{R}$ gilt

$$a + b = b + a$$

(*Kommutativgesetz*).

(A2) Für alle $a, b, c \in \mathbb{R}$ gilt

$$a + (b + c) = (a + b) + c$$

(*Assoziativgesetz*).

(A3) Es gibt (mindestens) ein Element $0 \in \mathbb{R}$, so dass für alle $a \in \mathbb{R}$ gilt

$$a + 0 = a$$

(*Existenz der Null*).

(A4) Zu jedem $a \in \mathbb{R}$ existiert (mindestens) ein $a' \in \mathbb{R}$, so dass gilt

$$a + a' = 0$$

(*Existenz des Negativen*).

Wir wollen nun zeigen, wie aus diesen Axiomen einige bekannte Aussagen formal hergeleitet werden können.

Satz 2.1 *Es gibt nur eine* 0. *Genauer: Wenn* $a + 0 = a$ *und* $a + \tilde{0} = a$ *für alle* $a \in \mathbb{R}$ *gilt, dann ist* $0 = \tilde{0}$.

Beweis. $0 \overset{\text{Vor.}}{=} 0 + \tilde{0} \overset{(A1)}{=} \tilde{0} + 0 \overset{\text{Vor.}}{=} \tilde{0}$. □

Satz 2.2 *Zu jedem* $a \in \mathbb{R}$ *gibt es nur ein Gegenelement* a'.

Beweis. Es seien $a', a'' \in \mathbb{R}$ mit

$$\begin{aligned} a + a' &= 0, \\ a + a'' &= 0. \end{aligned}$$

Dann gilt

$$\begin{aligned} a'' &\overset{\text{(A3)}}{=} a'' + 0 \overset{\text{Vor.}}{=} a'' + (a + a') \overset{\text{(A2)}}{=} (a'' + a) + a' \\ &\overset{\text{(A1)}}{=} (a + a'') + a' \overset{\text{Vor.}}{=} 0 + a' \overset{\text{(A1)}}{=} a' + 0 \overset{\text{(A3)}}{=} a' \end{aligned}$$

□

Notation Das eindeutig bestimmte a' mit $a + a' = 0$ wird mit $-a$ bezeichnet.

Satz 2.3 *Für jedes $a \in \mathbb{R}$ gilt*

$$-(-a) = a.$$

Beweis. Nach Definition des Negativen von $-a$ gilt

$$(-a) + (-(-a)) = 0.$$

Andererseits gilt

$$(-a) + a \overset{\text{(A1)}}{=} a + (-a) \overset{\text{(A4)}}{=} 0.$$

Aus der Eindeutigkeit des Negativen folgt nun $-(-a) = a$. □

Satz 2.4 *Die Gleichung $a + x = b$ hat eine eindeutig bestimmte Lösung $x = b + (-a)$. (*Eindeutige Lösbarkeit von Differenzaufgaben*)*

Notation $b - a := b + (-a)$.

Beweis. (a) Wir zeigen zunächst, dass $x = b - a$ die Gleichung löst:

$$\begin{aligned} a + (b - a) &= a + (b + (-a)) \quad \text{(Def.)} \\ &= a + ((-a) + b) \quad \text{(A1)} \\ &= (a + (-a)) + b \quad \text{(A2)} \\ &= 0 + b \quad \text{(A4)} \\ &= b \quad \text{(A3)} \end{aligned}$$

(b) Wir zeigen jetzt die Eindeutigkeit der Lösung: Aus

$$\begin{aligned} a + x &= b, \\ a + y &= b \end{aligned}$$

folgt

$$\begin{aligned} x &\overset{\text{(A3,A4)}}{=} ((-a)+a)+x \overset{\text{(A2)}}{=} (-a)+(a+x) \\ &\overset{\text{(Vor.)}}{=} (-a)+(a+y) \overset{\text{(A2)}}{=} ((-a)+a)+y \overset{\text{(A3,A4)}}{=} y. \end{aligned}$$

□

Man kann noch zahllose andere Folgerungen aus den Axiomen ziehen, z.B.

$$a-(b-c)=(a-b)+c.$$

Satz 2.5 *Für alle $a, b \in \mathbb{R}$ gilt*

$$-(a+b)=-a-b.$$

Beweis. Übung! □

Axiome der Multiplikation:

(M1) Für alle $a, b \in \mathbb{R}$ gilt

$$a \cdot b = b \cdot a$$

(*Kommutativgesetz*).

(M2) Für alle $a, b, c \in \mathbb{R}$ gilt

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c$$

(*Assoziativgesetz*).

(M3) Es gibt (mindestens) ein Element $1 \neq 0$ in $\mathbb{R}$, so dass für alle $a \in \mathbb{R}$ gilt

$$a \cdot 1 = a$$

(*Existenz der Eins*).

(M4) Zu jedem $a \in \mathbb{R}$ mit $a \neq 0$ existiert (mindestens) ein $a^* \in \mathbb{R}$, so dass gilt

$$a \cdot a^* = 1$$

(*Existenz des Inversen*).

Wie oben kann man beweisen:

Satz 2.6 *Ist $a \cdot 1 = a$ für alle $a \in \mathbb{R}$ und $a \cdot \tilde{1} = a$ für alle $a \in \mathbb{R}$, so ist $1 = \tilde{1}$. (*Eindeutigkeit der 1*)*

Satz 2.7 *Ist* $a \neq 0$ *ein Element von* $\mathbb{R}$ *und*

$$\begin{aligned} a \cdot a^* &= 1 \quad \textit{und} \\ a \cdot a^{**} &= 1, \end{aligned}$$

so ist $a^* = a^{**}$. *(*Eindeutigkeit des Inversen*)*

Notation Man bezeichnet das zu $a \neq 0$ eindeutig bestimmte multiplikative Gegenelement a^{-1} oder $\frac{1}{a}$, es ist von Null verschieden.

Zur Verbindung von $+$ und $\cdot$ hat man noch das Distributivgesetz als Axiom.

Distributivgesetz

(D) Für alle $a, b, c \in \mathbb{R}$ gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

Satz 2.8 *Für alle* $a, b, c \in \mathbb{R}$ *gilt*

$$(a + b) \cdot c = a \cdot c + b \cdot c.$$

Beweis.

$$(a + b) \cdot c \overset{(M1)}{=} c \cdot (a + b) \overset{(D)}{=} c \cdot a + c \cdot b \overset{(M1)}{=} a \cdot c + b \cdot c.$$

□

Satz 2.9 $a \cdot 0 = 0$ *für alle* $a \in \mathbb{R}$.

Beweis. Es gilt

$$\begin{aligned} a \cdot 0 &= a \cdot (0 + 0) \quad (A3) \\ &= a \cdot 0 + a \cdot 0 \quad (D). \end{aligned}$$

Die Zahl $a \cdot 0$ ist daher Lösung der Differenzaufgabe $a \cdot 0 + x = a \cdot 0$. Aber auch 0 ist eine Lösung dieser Aufgabe. Aus der Eindeutigkeit der Lösung folgt $a \cdot 0 = 0$. □

Satz 2.10 *Für* $a, b \in \mathbb{R}$ *gilt* $a \cdot b = 0$ *genau dann, wenn* $a = 0$ *oder* $b = 0$.

Bemerkung 2.1 Hier müssen wir eine Bemerkung zum Gebrauch des Wortes "oder" in der Mathematik machen. Das Wort "oder" ist in der Mathematik nie ausschließend gemeint im Sinne von "entweder/oder", sondern schließt mit ein, dass beide Aussagen gelten. Sind also A und B Aussagen wie z.B. $a = 0$ und $b = 0$, so bedeutet A *oder* B, dass A oder B gilt oder beide Aussagen A und B gelten. In unserem Fall bedeutet also "$a = 0$ oder $b = 0$", dass $a = 0$ oder $b = 0$ oder $a = b = 0$ gilt.

Beweis. Wir müssen zwei Aussagen beweisen:

(a) Wenn $a \cdot b = 0$ gilt, dann ist $a = 0$ oder $b = 0$.

(b) Wenn $a = 0$ oder $b = 0$ gilt, dann ist $a \cdot b = 0$.

Zu (a): Es sei $a \cdot b = 0$. Angenommen, es ist $a \neq 0$. Dann gilt

$$\begin{aligned} 0 &= a \cdot b \quad \text{(Vor.)} \\ &= a^{-1} \cdot (a \cdot b) \quad \text{(Satz 2.9)} \\ &= (a^{-1} \cdot a) \cdot b \quad \text{(M2)} \\ &= 1 \cdot b \quad \text{(M4)} \\ &= b \quad \text{(M3)}. \end{aligned}$$

Also ist wenigstens eine der beiden Zahlen a oder b gleich 0.
Zu (b): Dies folgt aus Satz 2.9 und (M1). □

Satz 2.11 *Für alle $a, b \in \mathbb{R}$ gilt*

$$(-a) \cdot b = -(a \cdot b).$$

Beweis. Behauptet wird also, dass $(-a) \cdot b$ das Gegenelement zu $a \cdot b$ ist, d.h. $a \cdot b + (-a) \cdot b = 0$. Aber das stimmt natürlich:

$$\begin{aligned} a \cdot b + (-a) \cdot b &= (a + (-a)) \cdot b \quad \text{(Satz 2.8)} \\ &= 0 \cdot b \quad \text{(A4)} \\ &= 0 \quad \text{(M1, Satz 2.9)}. \end{aligned}$$

□

Satz 2.12 *Für alle $a, b, c \in \mathbb{R}$ gilt*

$$a \cdot (b - c) = a \cdot b - a \cdot c.$$

Beweis. Behauptet wird also, dass $a \cdot (b - c)$ Lösung der Differenzaufgabe $a \cdot c + x = a \cdot b$ ist. Aber

$$\begin{aligned} a \cdot c + a \cdot (b - c) &= a \cdot (c + b + (-c)) \quad \text{(D)} \\ &= a \cdot (b + c + (-c)) \quad \text{(A1)} \\ &= a \cdot (b + 0) \quad \text{(A4)} \\ &= a \cdot b \quad \text{(A3)}. \end{aligned}$$

□

Genau wie oben folgen die drei nächsten Sätze:

Satz 2.13 *Für alle $a \in \mathbb{R}$ mit $a \neq 0$ gilt*

$$(a^{-1})^{-1} = a.$$

Satz 2.14 *Für $a \neq 0$ und b aus $\mathbb{R}$ existiert genau ein Element x mit $a \cdot x = b$. (*Eindeutige Lösbarkeit von Divisionsaufgaben*)*

Notation Dieses Element wird man $\frac{b}{a}$ nennen:

$$\frac{b}{a} := b \cdot a^{-1} = a^{-1} \cdot b.$$

Satz 2.15 *Für alle $a, b \in \mathbb{R}$ gilt*

$$(a \cdot b)^{-1} = a^{-1} \cdot b^{-1}.$$

Abschweifung in die abstrakte Algebra

Durch die bisher genannten Axiome (A1)–(A4), (M1)–(M4), (D) sind die reellen Zahlen noch nicht eindeutig festgelegt. Jede beliebige Menge M mit zwei Verknüpfungen, die diese Axiome erfüllt, nennt man einen *Körper*.

Beispiele für Körper sind $\mathbb{R}$ und $\mathbb{Q}$, jedoch ist $\mathbb{Z}$ kein Körper, da das Axiom von der Existenz des Inversen verletzt ist (z.B. besitzt die Zahl $2 \in \mathbb{Z}$ in $\mathbb{Z}$ kein Inverses).

Erstaunlicherweise gibt es Körper mit endlich vielen Elementen. Zum Beipiel einen mit 2 Elementen: Teilt man nämlich die natürlichen Zahlen in die zwei Klassen "ger" der geraden und "ung" der ungeraden Zahlen ein und addiert und multipliziert sie, wie ihre Elemente es tun würden, nämlich

$$\begin{array}{rcl@{\qquad}rcl} \text{ger} + \text{ger} & = & \text{ger} & \text{ger} \cdot \text{ger} & = & \text{ger} \\ \text{ger} + \text{ung} & = & \text{ung} & \text{ger} \cdot \text{ung} & = & \text{ger} \\ \text{ung} + \text{ger} & = & \text{ung} & \text{ung} \cdot \text{ger} & = & \text{ger} \\ \text{ung} + \text{ung} & = & \text{ger} & \text{ung} \cdot \text{ung} & = & \text{ung}, \end{array}$$

so ensteht ein Körper mit 2 Elementen, in dem "ger" der 0 und "ung" der 1 entspricht. Hier ist also insbesondere $1 + 1 = 0$, d.h. $1 = -1$. Diesen Körper bezeichnet man mit $\mathbb{F}_2$.

Zu weiteren Axiomen der reellen Zahlen gehören die Ordnungsaxiome, die die reellen Zahlen allerdings immer noch nicht charakterisieren. Der Zweck dieser Axiome ist es, die positiven Zahlen zu charakterisieren.

Ordnungsaxiome

Es ist eine Menge $P \subset \mathbb{R}$ ausgezeichnet, die *Menge der positiven Zahlen.* Sie erfüllt

(Ord1) Für jedes $a \in \mathbb{R}$ ist genau eine der drei Aussagen $a \in P$, $-a \in P$, $a = 0$ wahr. (*Trichotomiegesetz*)

(Ord2) Aus $a \in P$ und $b \in P$ folgt $a + b \in P$.

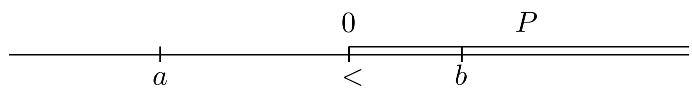
(Ord3) Aus $a \in P$ und $b \in P$ folgt $a \cdot b \in P$.

Wir können nun *definieren*:

Definition

$$\begin{aligned} a > b &\quad \text{genau dann, wenn} \quad a - b \in P, \\ a < b &\quad \text{genau dann, wenn} \quad b - a \in P. \end{aligned}$$

Anschaulich bedeutet dies: Auf der Zahlengeraden sind diejenigen Zahlen positiv, die rechts von 0 liegen. Es gilt $b > a$ genau dann, wenn b rechts von a liegt.



Merke

$$\begin{aligned} a > 0 &\quad \Leftrightarrow \quad a \in P \\ a < 0 &\quad \Leftrightarrow \quad -a \in P \end{aligned}$$

Von den drei Auusagen $a > 0$, $a < 0$, $a = 0$ ist für $a \in \mathbb{R}$ jeweils eine wahr.

Satz 2.16 *Für $a, b \in \mathbb{R}$ ist von den drei Aussagen $a > b$, $a < b$, $a = b$ genau eine erfüllt.*

Beweis. Setze $c := b - a$. Dann gilt

$$\begin{aligned} a < b &\quad \Leftrightarrow \quad b - a \in P \Leftrightarrow c \in P \\ a > b &\quad \Leftrightarrow \quad a - b \in P \Leftrightarrow -c \in P \\ a = b &\quad \Leftrightarrow \quad b - a = 0 \Leftrightarrow c = 0 \end{aligned}$$

Aus (Ord1) folgt die Behauptung. □

Satz 2.17 (Transitivität der Kleiner-Relation) *Für $a, b, c \in \mathbb{R}$ gilt:*

$$a < b \text{ und } b < c \Rightarrow a < c.$$

Beweis.

$$\begin{aligned} & a < b \text{ und } b < c \\ \Rightarrow \quad & b - a \in P \text{ und } c - b \in P \quad \text{(Def.)} \\ \Rightarrow \quad & c - a = (c - b) + (b - a) \in P \quad \text{(Ord2)} \end{aligned}$$

□

Satz 2.18 *Für $a, b, c \in \mathbb{R}$ gilt*

(a) $a < b \Rightarrow a + c < b + c.$

(b) $a < b$ *und* $c > 0 \Rightarrow a \cdot c < b \cdot c.$

(c) $a < b$ *und* $c < 0 \Rightarrow a \cdot c > b \cdot c.$

Beweis. (a) $(b + c) - (a + c) = b - a > 0$ nach Voraussetzung. Daher ist $a + c < b + c$.

(b)

$$\begin{aligned}
& a < b \text{ und } c > 0 \\
\Rightarrow\ & b - a \in P \text{ und } c \in P \quad \text{(Def.)} \\
\Rightarrow\ & (b - a) \cdot c \in P \quad \text{(Ord3)} \\
\Rightarrow\ & b \cdot c - a \cdot c \in P \quad \text{(Satz 2.12, M1)} \\
\Rightarrow\ & a \cdot c < b \cdot c \quad \text{(Def.)}
\end{aligned}$$

(c)

$$\begin{aligned}
& a < b \text{ und } c < 0 \\
\Rightarrow\ & b - a \in P \text{ und } -c \in P \quad \text{(Def.)} \\
\Rightarrow\ & (b - a) \cdot (-c) \in P \quad \text{(Ord3)} \\
\Rightarrow\ & a \cdot c - b \cdot c \in P \quad \text{(Satz 2.12, M1, Satz 2.11)} \\
\Rightarrow\ & a \cdot c > b \cdot c \quad \text{(Def.)}
\end{aligned}$$

□

Satz 2.19 *Für alle $a \in \mathbb{R}$ gilt*

$$a \neq 0 \Rightarrow a^2 := a \cdot a > 0.$$

Beweis. Falls $a > 0$, folgt die Behauptung aus (Ord3). Andernfalls ist $-a > 0$ und

$$(-a) \cdot (-a) = a^2 > 0 \quad \text{(Ord3)}.$$

□

Satz 2.20 *Für alle $a \in \mathbb{R}$ gilt*

$$\begin{aligned}
a > 0 \quad &\Rightarrow \quad a^{-1} = \frac{1}{a} > 0 \\
a < 0 \quad &\Rightarrow \quad a^{-1} = \frac{1}{a} < 0
\end{aligned}$$

Beweis. Es gilt $a^{-1} = a \cdot (a^{-1})^2 > 0$ falls $a > 0$ wegen (Ord3) und dem vorherigen Satz. □

Satz 2.21 *Für alle $a, b \in \mathbb{R}$ gilt*

$$0 < a < b \Rightarrow 0 < b^{-1} < a^{-1}.$$

Beweis.

$$\begin{aligned} a > 0 \text{ und } b > 0 \quad &\Rightarrow \quad a \cdot b > 0 \quad \text{(Ord3)} \\ &\Rightarrow \quad (a \cdot b)^{-1} = a^{-1} \cdot b^{-1} > 0 \quad \text{(Satz 2.15, Satz 2.20)} \end{aligned}$$

Deswegen darf man nach Satz 2.18 die Ungleichung $a < b$ mit $a^{-1} \cdot b^{-1}$ multiplizieren und erhält

$$b^{-1} = a \cdot (a^{-1} \cdot b^{-1}) < b \cdot (a^{-1} \cdot b^{-1}) = a^{-1}.$$

□

Satz 2.22 *Es gilt* $1 > 0$.

Beweis. Aus $1 \neq 0$ folgt nach Satz 2.19 $1^2 > 0$. Wegen $1^2 = 1$ (M3) folgt daraus $1 > 0$. □

Exkurs über Quadratwurzeln

Es sei $a \in \mathbb{R}$.
Frage: Gibt es dann ein $x \in \mathbb{R}$, so dass $x^2 = a$?

Das ist ein Problem, denn die bisherigen Axiome werden auch von der Menge $\mathbb{Q}$ der rationalen Zahlen erfüllt, und wir wissen bereits, dass z.B. die Gleichung $x^2 = 2$ keine Lösung in $\mathbb{Q}$ besitzt.

Satz 2.23 (a) *Falls $a < 0$, dann gibt es keine Lösung von $x^2 = a$.*

(b) *Falls $a = 0$, dann gibt es genau die Lösung $x = 0$ von $x^2 = a$ und keine andere.*

(c) *Falls $a > 0$, dann gibt es keine Lösung von $x^2 = a$ oder genau zwei.*

Beweis. (a) Aus Satz 2.19 folgt: Quadrate sind nicht negativ.

(b) Dass $x = 0$ eine Lösung von $x^2 = 0$ ist, ist klar; dass es keine andere gibt, folgt auch aus Satz 2.19.

(c) Angenommen, es gibt mindestens zwei Lösungen x_1, x_2 von $x^2 = a$. Dann gilt

$$\begin{aligned} & a = x_1^2 = x_2^2 \\ \Rightarrow \quad & 0 = x_1^2 - x_2^2 = (x_1 - x_2) \cdot (x_1 + x_2) \\ \Rightarrow \quad & x_1 = x_2 \text{ oder } x_1 = -x_2 \quad \text{(Satz 2.10)} \end{aligned}$$

Wenn es also eine Lösung x_1 von $x^2 = a$ gibt, dann ist $-x_1$ auch eine Lösung und es gibt keine weiteren. □

Welche Alternative von (c) nun gilt, werden wir später sehen.
Wir führen noch eine neue Bezeichnung ein:

Definition Man setzt $a \leq b$, falls $a < b$ oder $a = b$ gilt.

Für diese Relation gelten (als Folgerungen obiger Regeln) zahllose weitere, wie z.B.

$$a \leq b \text{ und } c \leq 0 \Rightarrow a \cdot c \leq b \cdot c,$$

$$a \geq 0 \text{ und } b \geq 0 \text{ und } a \leq b \Rightarrow a^2 \leq b^2.$$

Wir führen nun den *Absolutbetrag* ein.

Definition Für $a \in \mathbb{R}$ setzt man

$$|a| := \begin{cases} a & \text{falls } a \geq 0, \\ -a & \text{falls } a < 0. \end{cases}$$

Wegen (Ord1) hat man $|a|$ für alle $a \in \mathbb{R}$ definiert. Wir geben nun einige Eigenschaften des Absolutbetrags an.

Satz 2.24 *Es gilt stets $|a| \geq 0$ und $|a| = 0$ genau dann, wenn $a = 0$.*

Satz 2.25 *Für alle $a \in \mathbb{R}$ gilt*

$$|-a| = |a|.$$

Satz 2.26 *Für alle $a, b \in \mathbb{R}$ gilt*

$$|a \cdot b| = |a| \cdot |b|.$$

Beweis. Dies beweist man durch direkte Verifikation in jedem der vier möglichen Fälle

$$\begin{array}{ll} (1)\ a \geq 0, b \geq 0 & (2)\ a \geq 0, b < 0 \\ (3)\ a < 0, b \geq 0 & (4)\ a < 0, b < 0. \end{array}$$

□

Satz 2.27 (Dreiecksungleichung) *Für alle $a, b \in \mathbb{R}$ gilt*

$$|a + b| \leq |a| + |b|.$$

Beweis. Es gilt $a \leq |a|$ und $b \leq |b|$. Daher folgt (mit obigen Sätzen)

$$a + b \leq |a| + |b|.$$

Andererseits gilt $-a \leq |a|$ und $-b \leq |b|$ und daher

$$-(a + b) = -a - b \leq |a| + |b|.$$

Insgesamt ergibt sich also

$$|a+b| \le |a| + |b|.$$

□

Korollar 2.1 *Für alle $a, b \in \mathbb{R}$ gilt*

$$|a+b| \ge |a| - |b|.$$

Beweis. Es sei $c := a + b$ und $d := -b$. Nach der Dreiecksungleichung gilt

$$|c+d| \le |c| + |d|,$$

d.h.

$$|a| \le |a+b| + |-b| = |a+b| + |b|.$$

Addition von $-|b|$ auf beiden Seiten der Ungleichung ergibt die Behauptung. □

Mit Hilfe der Anordnung werden die *Intervalle* definiert.

Definition Es sei $a < b$.

$$\begin{array}{rcll} [a,b] & := & \{x \in \mathbb{R} \,|\, a \le x \le b\} & \textit{abgeschlossenes Intervall}, \\ [a,b) & := & \{x \in \mathbb{R} \,|\, a \le x < b\} & \textit{halboffenes Intervall}\ , \\ (a,b] & := & \{x \in \mathbb{R} \,|\, a < x \le b\} & \textit{halboffenes Intervall}\ , \\ (a,b) & := & \{x \in \mathbb{R} \,|\, a < x < b\} & \textit{offenes Intervall}\ . \end{array}$$

Die Zahl $b - a$ bezeichnet man als die *Länge* des Intervalls. Ebenso definiert man die *uneigentlichen Intervalle*:

$$\begin{array}{rcl} [a,\infty) & := & \{x \in \mathbb{R} \,|\, x \ge a\}, \\ (a,\infty) & := & \{x \in \mathbb{R} \,|\, x > a\}, \\ (-\infty,a] & := & \{x \in \mathbb{R} \,|\, x \le a\}, \\ (-\infty,a) & := & \{x \in \mathbb{R} \,|\, x < a\}. \end{array}$$

3 Vollständige Induktion, elementare Kombinatorik

Die natürlichen Zahlen sind durch den Zählprozeß

$$0, 1, 1+1, 1+1+1, \ldots$$

definiert. In dem von uns betrachteten Körper $\mathbb{R}$ gibt es $0, 1$, also auch $1+1$, $1+1+1$, usw. Diese Zahlen sind im Körper auch alle verschieden. (Diese

Bemerkung müssen wir machen, da im Körper mit 2 Elementen $1+1=0$ gilt, dies also nicht stimmt.) Dies folgt aus den Ordnungsaxiomen: Aus $0<1$ folgt $0+1<1+1$, daraus wiederum $1+1<1+1+1$, usw. In diesem Sinne ist also $\mathbb{N}$ in $\mathbb{R}$ eingebettet, d.h. $\mathbb{N}\subset\mathbb{R}$. Diese Einbettung hat man gerade so gewonnen, dass man alles nimmt, was sich aus 0 und 1 durch Addition gewinnen lässt. Nimmt man sogar alles zu 0 und 1 hinzu, was sich durch Addition und Subtraktion herstellen lässt, so erhält man eine Einbettung von $\mathbb{Z}$ in $\mathbb{R}$. Nimmt man schließlich zu 0 und 1 alles hinzu, was sich durch $+,-,\cdot,:$ gewinnen lässt, so erhält man die Einbettung $\mathbb{Q}\subset\mathbb{R}$.

Mit den natürlichen Zahlen hängt das *Beweisprinzip der vollständigen Induktion* zusammen, das wir nun darstellen wollen.

Es sei $A(n)$ eine Aussage, die von der natürlichen Zahl n abhängt, z.B. die Aussage

$$0+1+2+\cdots+n=\frac{n(n+1)}{2},$$

in Worten: Die Summe der ersten $n+1$ natürlichen Zahlen ist $\frac{n(n+1)}{2}$.

Wir wollen zeigen, dass diese Aussage für alle natürlichen Zahlen n richtig ist. Dazu ist zu zeigen:

(1) *Induktionsanfang*: $A(0)$ ist richtig.

(2) *Induktionsschritt*: Für jede Zahl $n\in\mathbb{N}$ gilt: Ist $A(n)$ richtig, so auch $A(n+1)$.

Gelingt dies zu zeigen, dann ist $A(n)$ für *jede* natürliche Zahl n richtig. Anschaulich ist das klar: Man beweist zuerst $A(0)$. Nach dem Induktionsschritt ist dann auch $A(1)$ wahr, also auch $A(2)$, also $A(3)$, usw.

Variante: Anstelle von 0 kann eine beliebige andere natürliche Zahl n_0 als "Anfangszahl" treten. Die Behauptung folgt mit obiger Beweismethode dann natürlich nur für $A(n)$ mit $n\geq n_0$.

Wir geben nun einige Beispiele zur Einübung der vollständigen Induktion. Hierzu ist es zweckmäßig, das Summen- und Produktzeichen einzuführen.

Es seien $m\leq n$ ganze Zahlen. Für jede ganze Zahl k mit $m\leq k\leq n$ sei $a_k\in\mathbb{R}$. Dann setzt man

$$\begin{aligned}\sum_{k=m}^{n}a_k &:= a_m+a_{m+1}+\cdots+a_n,\\ \prod_{k=m}^{n}a_k &:= a_m\cdot a_{m+1}\cdot\ldots\cdot a_n.\end{aligned}$$

Satz 3.1 *Für alle natürlichen Zahlen n gilt:*

$$\sum_{k=0}^{n}k=\frac{n(n+1)}{2}.$$

Beweis (durch vollständige Induktion nach n).

(1) Induktionsanfang: Für $n = 0$ gilt

$$\sum_{k=0}^{0} k = 0 = \frac{0(0+1)}{2}.$$

(2) Induktionsschritt $n \to n+1$:
Induktionsvoraussetzung: Es gelte

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}.$$

Es ist zu zeigen:

$$\sum_{k=0}^{n+1} k = \frac{(n+1)(n+2)}{2}.$$

Dies sieht man so ein:

$$\begin{aligned}\sum_{k=0}^{n+1} k &= \sum_{k=0}^{n} k + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) \quad \text{(Induktionsvoraussetzung)} \\ &= \frac{(n+1)(n+2)}{2}.\end{aligned}$$

□

Dieser Satz erinnert an die bekannte Geschichte über C. F. Gauß. Als Sechsjähriger sollte er die Zahlen von 1 bis 100 aufaddieren – damit sein Lehrer Ruhe hatte. Gauß war schnell fertig – ohne vollständige Induktion:

$$\sum_{k=0}^{100} k = \frac{100 \cdot 101}{2} = 5050.$$

Wie ist Gauß auf diese Zahl gekommen? Er schrieb die Zahlen nebeneinander und addierte von außen nach innen:

$$\begin{aligned}\sum_{k=0}^{100} k &= (1+100) + (2+99) + \cdots + (50+51) \\ &= 50 \cdot 101 = 5050.\end{aligned}$$

Satz 3.2 *Für alle natürliche Zahlen $n \geq 1$ gilt:*

$$\sum_{k=1}^{n} (2k-1) = n^2.$$

Beweis (durch vollständige Induktion nach n).

(1) Induktionsanfang: $n = 1$: $1 = 1$

(2) Induktionsschritt $n \to n+1$:

Induktionsvoraussetzung:

$$\sum_{k=1}^{n}(2k-1) = n^2.$$

Es ist zu zeigen:

$$\sum_{k=1}^{n+1}(2k-1) = (n+1)^2.$$

Dies sieht man so ein:

$$\begin{aligned}\sum_{k=1}^{n+1}(2k-1) &= \sum_{k=1}^{n}(2k-1) + (2(n+1)-1)\\ &= n^2 + 2n + 1 \quad \text{(Induktionsvoraussetzung)}\\ &= (n+1)^2.\end{aligned}$$

□

Definition Für eine natürliche Zahl $n \geq 1$ setzen wir

$$\begin{aligned}n! &:= 1 \cdot 2 \cdot \ldots \cdot n = \prod_{k=1}^{n} k \quad n \text{ } \textit{Fakultät},\\ 0! &:= 1.\end{aligned}$$

Bemerkung 3.1 Eine alternative Definition ist eine *rekursive* Definition von $n!$:

$$\begin{aligned}0! &:= 1,\\ n! &:= n \cdot (n-1)! \quad \text{für } n \geq 1.\end{aligned}$$

Satz 3.3 *Die Anzahl der möglichen Anordnungen einer n-elementigen Menge $\{A_1, \ldots, A_n\}$ ($n \geq 1$) ist gleich $n!$.*

Beweis (durch vollständige Induktion nach n).

(1) Induktionsanfang $n = 1$: Die einelementige Menge besitzt nur eine Anordnung ihrer Elemente. Andererseits ist $1! = 1$.

(2) Induktionsschritt: Die Behauptung gelte für n-elementige Mengen.

Mögliche Anordnungen der $(n+1)$-elementigen Menge $\{A_1, \ldots, A_{n+1}\}$:

$$\begin{array}{rl} \text{Anordnungen mit } A_1 \text{ an erster Stelle:} & n! \\ \vdots & \vdots \\ \text{Anordnungen mit } A_{n+1} \text{ an erster Stelle:} & n! \\ \hline \text{insgesamt } (n+1)n! = (n+1)! & \end{array}$$

□

Definition Es seien $n, k \in \mathbb{N}$, $0 \leq k \leq n$.

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \quad \textit{Binomialkoeffizient}$$

Lemma 3.1 *Für* $1 \leq k \leq n$ *gilt*

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Beweis. Durch Nachrechnen:

$$\begin{aligned}
\binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} \\
&= \frac{n!k + n!(n+1-k)}{k!(n+1-k)!} \\
&= \frac{n!(k+n+1-k)}{k!(n+1-k)!} \\
&= \frac{(n+1)!}{k!(n+1-k)!} \\
&= \binom{n+1}{k}.
\end{aligned}$$

□

Beispiel 3.1 Es gilt

$$\binom{0}{0} = 1$$

$$\binom{1}{0} = 1 \quad \binom{1}{1} = 1$$

$$\binom{2}{0} = 1 \quad \binom{2}{1} = 2 \quad \binom{2}{2} = 1$$

$$\binom{3}{0} = 1 \quad \binom{3}{1} = 3 \quad \binom{3}{2} = 3 \quad \binom{3}{3} = 1$$

Auf diese Weise entsteht das Pascalsche Dreieck. Das obige Lemma ist gerade die Regel, nach der das Pascalsche Dreieck sukzessive aufgebaut wird.

Satz 3.4 *Die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge $\{A_1, \ldots, A_n\}$ ist gleich $\binom{n}{k}$.*

Bemerkung 3.2 Dieser Satz zeigt, dass die Zahlen $\binom{n}{k}$ ganz sind, was aus ihrer Definition nicht unmittelbar ersichtlich ist.

Beweis (durch vollständige Induktion nach n).

(1) Induktionsanfang: $n = 0$. Dann ist auch $k = 0$ und $\emptyset$ ist die einzige Teilmenge von $\emptyset$. Andererseits gilt $\binom{0}{0} = 1$.

(2) Induktionsschritt $n \to n + 1$: Angenommen, die Behauptung ist richtig für n.

k-elementige Teilmengen der $(n + 1)$-elementigen Menge $\{A_1, \dots , A_{n+1}\}$:

$k = 0$: $1 = \binom{n+1}{0}$.

$k = n + 1$: $1 = \binom{n+1}{n+1}$.

sonst:

Teilmengen, die A_{n+1} nicht enthalten: $\binom{n}{k}$ (nach Induktionsannahme)

Teilmengen, die A_{n+1} enthalten: $\binom{n}{k-1}$ (entsprechen $(k-1)$-elementigen Teilmengen von $\{A_1, \dots , A_n\}$)

Also insgesamt nach Lemma 3.1:

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}.$$

□

Übungsaufgabe 3.1 Wieviele Kombinationen sind beim Lotto "6 aus 49" möglich?
Antwort: $\binom{49}{6} = \frac{49\cdot48\cdot47\cdot46\cdot45\cdot44}{1\cdot2\cdot3\cdot4\cdot5\cdot6} = 13\,983\,816.$

Notation Für eine reelle Zahl a setzen wir

$$\begin{aligned} a^0 &:= 1, \\ a^n &:= a \cdot a \cdot \dots \cdot a \ (n \text{ Faktoren}), \quad n \geq 1. \end{aligned}$$

Satz 3.5 (Binomische Formel) *Es seien a, b reelle Zahlen. Dann gilt für alle natürlichen Zahlen n*

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

Beweis (durch vollständige Induktion nach n).

(1) Induktionsanfang: $n = 0$

$$(a + b)^0 = 1 = \sum_{k=0}^{0} \binom{0}{k} a^{0-k} b^k.$$

(2) Induktionsschritt $n \to n+1$:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n(a+b) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k\right)(a+b) \quad \text{(Induktionsvoraussetzung)} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^{n-1} \binom{n}{k} a^{n-k} b^{k+1} + b^{n+1}
\end{aligned}$$

Wir setzen nun in der zweiten Summe für den Summationsindex $k = j - 1$ ein.

$$= a^{n+1} + \sum_{k=1}^{n} \binom{n}{k} a^{n+1-k} b^k + \sum_{j=1}^{n} \binom{n}{j-1} a^{n+1-j} b^j + b^{n+1}$$

Da sich die beiden Summen bis auf die Binomialkoeffizienten nur in der Indizierung unterscheiden, können wir sie zusammenfassen:

$$\begin{aligned}
&= a^{n+1} + \sum_{k=1}^{n} \left(\binom{n}{k} + \binom{n}{k-1}\right) a^{n+1-k} b^k + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k \quad \text{(nach Lemma 3.1).}
\end{aligned}$$

□

Satz 3.6 (Geometrische Summenformel) *Es sei $q \neq 1$. Dann gilt für alle natürlichen Zahlen $n \geq 0$:*

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}.$$

Beweis (durch vollständige Induktion nach n).

(1) Induktionsanfang $n = 0$:

$$\sum_{k=0}^{0} q^k = 1 = \frac{1-q}{1-q}.$$

(2) Induktionsschritt $n \to n+1$:

$$\begin{aligned}
\sum_{k=0}^{n+1} q^k &= \sum_{k=0}^{n} q^k + q^{n+1} \\
&= \frac{1-q^{n+1}+(1-q)q^{n+1}}{1-q} \quad \text{(Induktionsvoraussetzung)} \\
&= \frac{1-q^{n+1}+q^{n+1}-q^{n+2}}{1-q} \\
&= \frac{1-q^{n+2}}{1-q}.
\end{aligned}$$

□

4 Die Axiome der reellen Zahlen (Teil 2)

Die Menge $\mathbb{Q}$ der rationalen Zahlen erfüllt die Körperaxiome und die Anordnungsaxiome. Der Körper der reellen Zahlen, den wir axiomatisch kennzeichnen wollen, ist aber umfangreicher als $\mathbb{Q}$, er soll ja z.B. auch $\sqrt{2}$ enthalten.

Wir werden nun ein Axiom formulieren, das die Vollständigkeit der reellen Zahlen beschreibt.

Definition Eine Menge M von reellen Zahlen heißt *nach oben beschränkt*, wenn es eine reelle Zahl s gibt, so dass gilt

$$x \leq s \text{ für alle } x \in M.$$

Eine solche Zahl s heißt dann eine *obere Schranke* von M.

Beispiel 4.1 Beispiele für nach oben beschränkte Mengen sind:

$$[a,b], [a,b), (a,b), (-\infty,a), (-\infty,a],$$

für *nicht* nach oben beschränkte Mengen: $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$.

Bemerkung 4.1 Falls s_1 eine obere Schranke von M ist und $s_2 \geq s_1$, dann ist auch s_2 eine obere Schranke von M.

Definition Eine Zahl s_0 heißt *kleinste obere Schranke* oder *Supremum* von M (in Zeichen $s_0 = \sup M$) genau dann, wenn gilt:

(i) s_0 ist eine obere Schranke von M, und

(ii) wenn s eine obere Schranke von M ist, so gilt

$$s_0 \leq s.$$

Satz 4.1 (Eindeutigkeit des Supremums) *Sind s_0 und s_0' kleinste obere Schranken von M, so gilt $s_0 = s_0'$.*

Beweis. Es gilt

$$s_0 \leq s_0',$$

da s_0' eine obere Schranke und s_0 eine kleinste obere Schranke von M ist, und

$$s_0' \leq s_0,$$

da s_0 eine obere Schranke und s_0' eine kleinste obere Schranke von M ist. Deshalb folgt $s_0 = s_0'$. □

Wegen Satz 4.1 können wir von *dem* Supremum von M sprechen.

Beispiel 4.2 $M = \{x \in \mathbb{R} \,|\, x < 1\} = (-\infty, 1)$ ist nach oben beschränkt. Wir zeigen $s_0 = \sup M = 1$. Denn:
(i) $s_0 = 1$ ist obere Schranke von M.
(ii) Angenommen, $s_0 < 1$. Dann gilt

$$\begin{aligned} & 2s_0 < s_0 + 1 < 2 \\ \Rightarrow \quad & s_0 < \frac{s_0+1}{2} < 1 \\ \Rightarrow \quad & \frac{s_0+1}{2} \in M \end{aligned}$$

Daraus folgt, dass s_0 keine obere Schranke von M ist, ein Widerspruch. (N.B. M besitzt kein größtes Element, denn:

$$x \in M \Rightarrow x < 1 \Rightarrow x < \frac{x+1}{2} < 1 \Rightarrow \frac{x+1}{2} \in M.)$$

Die Menge $M = \{x \in \mathbb{R} \,|\, x^2 < 2\}$ ist nach oben beschränkt: z.B. ist $s = \frac{3}{2}$ eine obere Schranke von M.

$$(x \geq \frac{3}{2} \Rightarrow x^2 \geq \frac{9}{4} \Rightarrow x^2 > 2)$$

Existiert die kleinste obere Schranke s_0? Falls ja, dann muss $s_0^2 = 2$ gelten (das werden wir gleich zeigen), aber $\sqrt{2} \notin \mathbb{Q}$. Also kann die Existenz einer kleinsten oberen Schranke in diesem Fall nicht aus den Körper- und Ordungsaxiomen gefolgert werden.

Daher fordern wir die Existenz des Supremums für *jede* nach oben beschränkte Menge:

Vollständigkeitsaxiom

(V) Jede nicht leere, nach oben beschränkte Menge von reellen Zahlen besitzt ein Supremum.

Es gibt auch äquivalente Formulierungen mit Intervallschachtelungen oder Dedekindschen Schnitten.

Bemerkung 4.2 $s_0 = \sup M$ braucht nicht zu M gehören. Wenn $s_0 \in M$ gilt, dann nennen wir s_0 das *Maximum* von M oder das *größte* Element von M, in Zeichen $s_0 = \max M$.

Beispiel 4.3 Es sei

$$M := \{x \in \mathbb{R} \,|\, x^2 < 2\}.$$

Dann ist M nach oben beschränkt. Das Vollständigkeitsaxiom sichert uns für M die Existenz von $s_0 = \sup M$.

Behauptung $s_0^2 = 2$.

Beweis. (1) Es gilt nicht $s_0^2 < 2$.
Angenommen, $s_0^2 < 2$. Diese Annahme führen wir zum Widerspruch, indem wir ein h mit $0 < h < 1$ so bestimmen, dass $s_0 + h \in M$ gilt. Es gilt

$$s_0 + h \in M \Leftrightarrow (s_0 + h)^2 < 2 \Leftrightarrow 2hs_0 + h^2 < 2 - s_0^2.$$

Wegen $h < 1$ gilt

$$2hs_0 + h^2 < 2hs_0 + h.$$

Setze

$$2hs_0 + h = 2 - s_0^2, \text{ d.h. } h := \frac{2 - s_0^2}{2s_0 + 1}.$$

Wegen $s_0 > 1$ gilt $0 < h < 1$ und nach Konstruktion von h gilt $s_0 + h \in M$. Da $s_0 + h \in M$ und $s_0 + h > s_0$, kann s_0 nicht obere Schranke von M sein, Widerspruch.

(2) Es gilt nicht $s_0^2 > 2$.
Angenommen, $s_0^2 > 2$. Dann setzen wir

$$s := s_0 - \frac{s_0^2 - 2}{2s_0}.$$

Dann gilt $s < s_0$ und

$$s^2 = s_0^2 - 2s_0\frac{s_0^2 - 2}{2s_0} + \left(\frac{s_0^2 - 2}{2s_0}\right)^2 > s_0^2 - (s_0^2 - 2) = 2.$$

Also folgt für jedes $x \in M$

$$x^2 < 2 < s^2$$

und daraus $x < s$, d.h. s ist obere Schranke von M, aber $s < s_0$, Widerspruch. □

Aufgrund des Vollständigkeitsaxioms ist also gezeigt (aber nicht konstruktiv!), dass die Gleichung $x^2 = 2$ eine nichtnegative reelle Lösung besitzt:

$$\sqrt{2} := \sup\{x \in \mathbb{R} \,|\, x^2 < 2\}.$$

Analog werden definiert:

- *untere Schranke* von M
- *größte untere Schranke* oder *Infimum* von M: $\inf M$
- *Minimum* oder *kleinstes Element* von M: $\min M$

Es gilt der Satz

Satz 4.2 *Jede nicht leere, nach unten beschränkte Menge M reeller Zahlen hat eine größte untere Schranke.*

Beweis. Die Menge $M' = \{-x \,|\, x \in M\}$ ist nach oben beschränkt, besitzt also nach dem Vollständigkeitsaxiom eine kleinste obere Schranke s_0.

Behauptung $\inf M = -s_0$.

Beweis (a) $-s_0$ ist untere Schranke von M, denn es gilt

$$x \in M \Rightarrow -x \in M' \Rightarrow -x \leq s_0 \Rightarrow -s_0 \leq x.$$

(b) Ist s eine untere Schranke von M, dann gilt $s \leq y$ für alle $y \in M$, also $x \leq -s$ für alle $x \in M'$. Daher gilt $s_0 \leq -s$, also $s \leq -s_0$. Deshalb ist $-s_0$ größte untere Schranke von M. □

Man kann nun beweisen, dass die reellen Zahlen vollständig durch die genannten Axiome charakterisiert sind, d.h. es bis auf Äquivalenz einen und auch nur einen Körper gibt, der alle genannten Axiome erfüllt. Diesen wohlbestimmten Körper nennt man $\mathbb{R}$.

Auf den Beweis von Existenz und Eindeutigkeit verzichten wir und wenden uns nun weiteren Folgerungen aus dem Vollständugkeitsaxiom zu.

Satz 4.3 *Die Menge $\mathbb{N}$ der natürlichen Zahlen ist nicht nach oben beschränkt.*

Beweis. Angenommen, $\mathbb{N}$ ist nach oben beschränkt. Da $\mathbb{N} \neq \emptyset$, gibt es dann ein Supremum s_0 für $\mathbb{N}$. Es gilt also

$$n \leq s_0 \quad \text{für alle } n \in \mathbb{N}.$$

Andererseits gibt es ein $n_0 \in \mathbb{N}$ mit

$$n_0 > s_0 - 1,$$

denn sonst wäre $s_0 - 1$ obere Schranke von $\mathbb{N}$ im Widerspruch dazu, dass s_0 die *kleinste* obere Schranke von $\mathbb{N}$ ist. Für n_0 gilt also

$$n_0 + 1 > s_0,$$

ein Widerspruch. □

Satz 4.4 (Satz von Archimedes) *Zu jedem $a \in \mathbb{R}$ mit $a > 0$ und jedem $b \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$, so dass $na > b$ gilt. (Für diesen Sachverhalt sagt man kurz: Der Körper $\mathbb{R}$ ist* archimedisch angeordnet.*)*

Beweis. Andernfalls würde für alle $n \in \mathbb{N}$

$$na \leq b \quad \text{d.h.} \quad n \leq \frac{b}{a}$$

gelten, $\frac{b}{a}$ wäre also eine obere Schranke für $\mathbb{N}$. □

Satz 4.5 *Zu jedem $\varepsilon > 0$ gibt es eine natürliche Zahl $n \neq 0$ mit $\frac{1}{n} < \varepsilon$.*

Beweis. Da

$$\frac{1}{n} < \varepsilon \Leftrightarrow n\varepsilon > 1,$$

folgt die Behauptung aus dem Satz von Archimedes mit $a = \varepsilon$ und $b = 1$. □

5 Folgen, Grenzwerte

Wir kommen jetzt zu einem der zentralen Begriffe der Analysis, dem des Grenzwertes einer Folge. Seine Bedeutung beruht darauf, dass viele Größen nicht durch einen in endlich vielen Schritten exakt berechenbaren Ausdruck gegeben werden, sondern nur mit beliebiger Genauigkeit approximiert werden können. Eine Zahl mit beliebiger Genauigkeit approximieren heißt sie als Grenzwert einer Folge darstellen.

Definition Unter einer *Folge* reeller Zahlen versteht man eine Funktion

$$\begin{array}{ccc} \mathbb{N} & \longrightarrow & \mathbb{R} \\ n & \longmapsto & a_n \end{array}.$$

Bezeichnung $(a_n)_{n\in\mathbb{N}}$ oder $a_0, a_1, a_2, \ldots$.

Bemerkung 5.1 Es spielt im Prinzip keine Rolle, mit welchem Index man beginnt. Es sei $n_0 \in \mathbb{Z}$. Dann bezeichnet man $(a_n)_{n\geq n_0}$ oder

$$a_{n_0}, a_{n_0+1}, a_{n_0+2}, \ldots$$

auch als Folge.

Beispiel 5.1 (1) Es sei $a_n = 1$ für alle $n \in \mathbb{N}$. Man erhält die konstante Folge $1, 1, 1, \ldots$.

(2) $a_n = \frac{1}{n}$, $(n \geq 1)$: $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$.

(3)

$$a_n = \begin{cases} \frac{1}{n+2} & \text{für } n \text{ ungerade,} \\ \frac{1}{n} & \text{für } n \text{ gerade,} \end{cases} \quad (n \geq 1): \frac{1}{3}, \frac{1}{2}, \frac{1}{5}, \frac{1}{4}, \frac{1}{7}, \frac{1}{6}, \ldots.$$

(4) $a_n = (-1)^n : 1, -1, 1, -1, 1, -1, \ldots$.

(5)

$$a_n = \begin{cases} n & \text{für } n \text{ ungerade,} \\ \frac{1}{n} & \text{für } n \text{ gerade,} \end{cases} \quad (n \geq 1): 1, \frac{1}{2}, 3, \frac{1}{4}, 5, \frac{1}{6}, \ldots.$$

(6) Es sei $q \in \mathbb{R}$, $a_n = q^n$: $1, q, q^2, q^3, \ldots$.

(7) Fest vorgegeben seien die Zahlen $b > 0$ und $c > 0$ sowie das Rekursionsschema

$$\begin{aligned} a_0 &= c, \\ a_{n+1} &= \frac{1}{2}\left(a_n + \frac{b}{a_n}\right). \end{aligned}$$

Definition Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt *konvergent* gegen $a \in \mathbb{R}$ (in Zeichen $\lim_{n\to\infty} a_n = a$ oder $a_n \to a$) genau dann, wenn gilt:

> Zu jedem $\varepsilon > 0$ gibt es eine natürliche Zahl n_0, so dass für alle $n \geq n_0$ gilt: $|a_n - a| < \varepsilon$.

Hier ist es wichtig, sich die logische Struktur klarzumachen. Deswegen schreiben wir diese Definition noch einmal in Kurzform (mit Quantoren) hin:

$$\lim_{n\to\infty} a_n = a \Longleftrightarrow \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0\ |a_n - a| < \varepsilon.$$

Eine andere Formulierung: Es gilt $\lim_{n\to\infty} a_n = a$ genau dann, wenn gilt:

> In jeder ε-Umgebung $(\varepsilon > 0)$ von a liegen fast alle Glieder der Folge.

Hierbei nennt man das Intervall

$$(a - \varepsilon, a + \varepsilon) := \{x \in \mathbb{R} \mid |x - a| < \varepsilon\}$$

eine *ε-Umgebung* von a. *Fast alle* bedeutet alle bis auf endlich viele.

Es gilt also $\lim_{n\to\infty} a_n = a$ genau dann, wenn in jedem Intervall $(a - \varepsilon, a + \varepsilon)$ (für beliebig kleines $\varepsilon > 0$) unendlich viele, außerhalb aber höchstens endlich viele Glieder der Folge liegen.

Definition Eine Folge heißt *divergent*, wenn sie gegen keine reelle Zahl konvergiert.

Formal:

$$(a_n)_{n\in\mathbb{N}} \text{ divergent } \iff \forall a \in \mathbb{R}\ \exists \varepsilon > 0\ \forall n_0 \in \mathbb{N}\ \exists n \geq n_0\ |a_n - a| \geq \varepsilon.$$

Wir untersuchen nun die Beispiele auf Konvergenz bzw. Divergenz.
Beispiel 5.1 (1): Es gilt $\lim_{n\to\infty} a_n = 1$.
Beispiel 5.1 (2): Wir zeigen $\lim_{n\to\infty} \frac{1}{n} = 0$. Es sei $\varepsilon > 0$ beliebig vorgegeben. Wir müssen ein $n_0 \in \mathbb{N}$ finden, so dass für alle $n \geq n_0$ gilt

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon.$$

Es sei $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \varepsilon$. Eine solche Zahl existiert nach Satz 4.5. Dann gilt für alle $n \geq n_0$

$$\frac{1}{n} \leq \frac{1}{n_0} < \varepsilon.$$

Folgen, die konvergent mit dem Grenzwert 0 sind, heißen auch *Nullfolgen*.
Beispiel 5.1 (3): Es gilt ebenfalls $\lim_{n\to\infty} a_n = 0$.
Beispiel 5.1 (4): Wir zeigen, dass die Folge $((-1)^n)_{n\in\mathbb{N}}$ divergent ist. Dazu ist zu zeigen, dass sie gegen kein $a \in \mathbb{R}$ konvergiert. Es sei $a \in \mathbb{R}$. Setze $\varepsilon := 1$. Ist $a \geq 0$ und $n_0 \in \mathbb{N}$ gegeben, so gilt für ungerade n mit $n \geq n_0$

$$|a_n - a| = |-1 - a| = 1 + a \geq 1.$$

Ist $a < 0$ und $n_0 \in \mathbb{N}$ gegeben, so gilt für gerade n mit $n \geq n_0$

$$|a_n - a| = |1 - a| = 1 + (-a) > 1.$$

Bevor wir die anderen Beispiele untersuchen, notieren wir einige einfache Sätze.

Satz 5.1 (Eindeutigkeit des Grenzwertes) *Gilt* $\lim_{n\to\infty} a_n = a$ *und* $\lim_{n\to\infty} a_n = b$, *so gilt* $a = b$.

Beweis. Angenommen, es wäre $a \neq b$. Es sei $\varepsilon = \frac{|a-b|}{2}$.
Da $\lim_{n\to\infty} a_n = a$, gibt es ein $n_1 \in \mathbb{N}$, so dass für alle $n \geq n_1$ gilt: $|a_n - a| < \varepsilon$.
Da $\lim_{n\to\infty} a_n = b$, gibt es ein $n_2 \in \mathbb{N}$, so dass für alle $n \geq n_2$ gilt: $|a_n - b| < \varepsilon$.
Für $n_0 := \max(n_1, n_2)$ und $n \geq n_0$ gilt dann

$$\begin{aligned} |a - b| &= |(a - a_n) + (a_n - b)| \\ &\leq |a_n - a| + |a_n - b| \\ &< 2\varepsilon = |a - b|, \end{aligned}$$

d.h. $|a - b| < |a - b|$, ein Widerspruch. $\square$

Satz 5.2 *Eine Folge* $(a_n)_{n\in\mathbb{N}}$ *konvergiert genau dann gegen* $a \in \mathbb{R}$*, wenn* $(a_n - a)_{n\in\mathbb{N}}$ *eine Nullfolge ist.*

Beweis. Dies folgt unmittelbar aus der Definition. □

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

- *monoton wachsend*, falls $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt.
- *monoton fallend*, falls $a_n \geq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt.
- *nach oben beschränkt*, falls es eine Zahl $M \in \mathbb{R}$ gibt, so dass $a_n \leq M$ für alle $n \in \mathbb{N}$ gilt.
- *nach unten beschränkt*, falls es eine Zahl $m \in \mathbb{R}$ gibt, so dass $m \leq a_n$ für alle $n \in \mathbb{N}$ gilt.
- *beschränkt*, falls sie nach oben und unten beschränkt ist.

Satz 5.3 *Jede konvergente Folge* $(a_n)_{n\in\mathbb{N}}$ *ist beschränkt.*

Beweis. Es sei $\lim_{n\to\infty} a_n = a$. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt

$$|a_n - a| < 1$$

($\varepsilon = 1$ gewählt). Daraus folgt für alle $n \geq N$

$$|a_n| \leq |a| + |a_n - a| \leq |a| + 1.$$

Es sei

$$M := \max\{|a_0|, |a_1|, \ldots, |a_{N-1}|, |a| + 1\}.$$

Dann gilt

$$|a_n| \leq M \text{ für alle } n \in \mathbb{N}.$$

□

Bemerkung 5.2 Die Umkehrung dieses Satzes gilt nicht, siehe Beispiel 5.1 (4).

Beispiel 5.1 (5): Wegen $a_n = n$ für n ungerade ist die Folge nicht beschränkt, also nicht konvergent.

Satz 5.4 (Bernoullische Ungleichung) *Es sei* $x \geq -1$*. Dann gilt für alle* $n \in \mathbb{N}$

$$(1 + x)^n \geq 1 + nx.$$

Beweis (durch vollständige Induktion nach n).

Induktionsanfang: Die Behauptung ist offensichtlich richtig für $n = 0$.

Induktionsschritt $n \to n+1$: Es gilt

$$\begin{aligned} (1+x)^{n+1} &= (1+x)^n(1+x) \\ &\geq (1+nx)(1+x) \quad \text{(Induktionsvoraussetzung)} \\ &= 1+nx+x+nx^2 \\ &\geq 1+(n+1)x. \end{aligned}$$

□

Beispiel 5.1 (6): Das Konvergenzverhalten der Folge $(q^n)_{n\in\mathbb{N}}$ hängt vom Wert von q ab.

- $|q| > 1$: Dann $|q| = 1+h$, $h > 0$. Also folgt aus der Bernoullischen Ungleichung
$$|q^n| = |q|^n = (1+h)^n \geq 1+nh.$$
Nach dem Satz von Archimedes ist also (q^n) nicht beschränkt, also divergent.

- $q = 1$: $\lim_{n\to\infty} q^n = 1$.

- $q = -1$: (q^n) divergent.

- $|q| < 1$: Dann $\left|\frac{1}{q}\right| > 1$, also $\left|\frac{1}{q}\right| = 1+h$, $h > 0$. Für gegebenes $\varepsilon > 0$ findet man ein n_0, so dass für $n \geq n_0$
$$\left|\frac{1}{q^n}\right| \geq 1+nh > \frac{1}{\varepsilon} \quad \text{(wie oben)},$$
also $|q^n| < \varepsilon$. Also gilt $\lim_{n\to\infty} q^n = 0$.

Eine weitere notwendige Bedingung für die Konvergenz einer Folge verwendet den Begriff der Teilfolge.

Man erhält eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$, indem man aus der Folge (a_n) gewisse Glieder wegstreicht, jedoch noch unendlich viele Glieder übriglässt:

Definition Eine *Teilfolge* einer Folge (a_n) ist eine Folge der Form

$$a_{n_0}, a_{n_1}, a_{n_2}, \ldots,$$

wobei die n_j natürliche Zahlen mit

$$n_0 < n_1 < n_2 < \cdots$$

sind.

Satz 5.5 *Jede Teilfolge einer konvergenten Folge konvergiert gegen denselben Grenzwert.*

Beweis. Zu gegebenem $\varepsilon > 0$ ist bis auf endlich viele Ausnahmen $|a_k - a| < \varepsilon$, also auch $|a_{n_k} - a| < \varepsilon$, da $n_k \geq k$. □

Wir besprechen nun einige Rechenregeln für konvergente Folgen.

Satz 5.6 *Es sei $\lambda \in \mathbb{R}$ und es gelte $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann sind die Folgen $(a_n + b_n)$ und (λa_n) konvergent und es gilt*

$$\begin{aligned} \lim_{n\to\infty}(a_n + b_n) &= a + b, \\ \lim_{n\to\infty} \lambda a_n &= \lambda a. \end{aligned}$$

Beweis. (i) Es sei $\varepsilon > 0$ vorgegeben. Dann gibt es zu $\frac{\varepsilon}{2} > 0$ natürliche Zahlen n_0 und m_0, so dass gilt

$$\begin{aligned} |a_n - a| &< \frac{\varepsilon}{2} \quad \text{für } n \geq n_0, \\ |b_n - b| &< \frac{\varepsilon}{2} \quad \text{für } n \geq m_0. \end{aligned}$$

Also gilt für $n \geq \max(n_0, m_0)$:

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(ii) Für $\lambda = 0$ ist die Behauptung klar. Es sei $\lambda \neq 0$. Dann gibt es zu $\frac{\varepsilon}{|\lambda|} > 0$ ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt

$$|a_n - a| < \frac{\varepsilon}{|\lambda|},$$

also

$$|\lambda||a_n - a| = |\lambda a_n - \lambda a| < \varepsilon.$$

□

Satz 5.7 *Es sei $(a_n)_{n\in\mathbb{N}}$ Nullfolge und $(b_n)_{n\in\mathbb{N}}$ beschränkt. Dann ist auch $(a_n b_n)$ Nullfolge.*

Beweis. Es sei $\varepsilon > 0$ vorgegeben und $|b_n| \leq c$ für ein $c > 0$ für alle $n \in \mathbb{N}$. Dann gibt es zu $\frac{\varepsilon}{c}$ ein $n_0 \in \mathbb{N}$, so dass für $n \geq n_0$

$$|a_n| < \frac{\varepsilon}{c}, \text{ also } |a_n b_n| < |a_n| c < \varepsilon.$$

□

Satz 5.8 *Es sei* $\lim_{n\to\infty} a_n = a$ *und* $\lim_{n\to\infty} b_n = b$. *Dann ist* $(a_n b_n)$ *konvergent und es gilt*

$$\lim_{n\to\infty} a_n b_n = ab.$$

Beweis. Nach Satz 5.2 ist $(a_n - a)$ eine Nullfolge und (b_n) ist als konvergente Folge beschränkt (Satz 5.3). Aus dem vorherigen Satz folgt daher, dass $(a_n b_n - a b_n)$ eine Nullfolge ist. Aus Satz 5.6 folgt: $\lim_{n\to\infty} a b_n = ab$. Schreibe: $a_n b_n = (a_n b_n - a b_n) + a b_n$. Nochmalige Anwendung der Rechenregeln von Satz 5.6 ergibt:

$$\lim_{n\to\infty} a_n b_n = \lim_{n\to\infty} (a_n b_n - a b_n) + \lim_{n\to\infty} a b_n = 0 + ab = ab.$$

□

Satz 5.9 *Es sei* $\lim_{n\to\infty} a_n = a$ *und* $\lim_{n\to\infty} b_n = b \neq 0$. *Dann gibt es ein* $n_0 \in \mathbb{N}$, *so dass* $b_n \neq 0$ *für* $n \geq n_0$, *und die Folge* $\left(\frac{a_n}{b_n}\right)_{n \geq n_0}$ *ist konvergent und es gilt*

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

Beweis. Wir betrachten zunächst den Spezialfall $a_n = 1$ für alle $n \in \mathbb{N}$. Wir wollen Satz 5.7 anwenden auf

$$\frac{1}{b_n} - \frac{1}{b} = (b - b_n) \cdot \frac{1}{b b_n} \quad (n \geq n_0).$$

Dazu ist zu zeigen:

Behauptung $\left(\frac{1}{b b_n}\right)$ ist beschränkt.

Beweis. Für fast alle n gilt

$$|b_n - b| < \frac{|b|}{2} (=: \varepsilon).$$

Hieraus folgt

$$|b_n| \geq \varepsilon = \frac{|b|}{2}.$$

Also gilt für fast alle $n \in \mathbb{N}$

$$\left|\frac{1}{b b_n}\right| \leq \frac{2}{|b|^2}.$$

Hieraus folgt wie im Beweis von Satz 5.3 die Behauptung. □

Für den allgemeinen Fall betrachten wir

$$\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$$

und wenden Satz 5.8 an. □

Satz 5.10 *Es sei* $\lim_{n\to\infty} a_n = a$*,* $\lim_{n\to\infty} b_n = b$ *und* $a_n \le b_n$ *für fast alle* n*. Dann gilt* $a \le b$*.*

Warnung Wenn $a_n < b_n$ für fast alle n gilt, dann ist nicht notwendig $a < b$! Beispiel:

$$a_n = 0, \; b_n = \frac{1}{n} \; (n \ge 1), \; a = b = 0.$$

Beweis. Angenommen, $b < a$. Dann ist $\varepsilon := \frac{a-b}{2} > 0$ und es gibt $n_0, m_0 \in \mathbb{N}$ mit

$$\begin{aligned} |a_n - a| &< \varepsilon \text{ für } n \ge n_0, \\ |b_n - b| &< \varepsilon \text{ für } n \ge m_0. \end{aligned}$$

Für $n \ge \max(n_0, m_0)$ gilt dann

$$a_n > a - \varepsilon \text{ und } b_n < b + \varepsilon.$$

Wegen $a - \varepsilon = b + \varepsilon$ folgt daraus aber $b_n < a_n$, ein Widerspruch. □

Satz 5.11 (Einzwängungssatz) *Für fast alle* n *gelte*

$$a_n \le b_n \le c_n.$$

Gilt $\lim_{n\to\infty} a_n = a$ *und* $\lim_{n\to\infty} c_n = a$*, so gilt auch* $\lim_{n\to\infty} b_n = a$*.*

Beweis. Es sei $\varepsilon > 0$ gegeben. In der ε-Umgebung von a liegen fast alle Glieder der Folgen (a_n) und (c_n). Wegen $a_n \le b_n \le c_n$ gilt dasselbe für die Folge b_n. □

Aus dem Vollständigkeitsaxiom kann man ein wichtiges Konvergenzkriterium ableiten.

Satz 5.12 (Ein Konvergenzkriterium) *Jede monoton wachsende und nach oben beschränkte Folge ist konvergent. Jede monoton fallende und nach unten beschränkte Folge ist konvergent.*

Beweis. Wir betrachten den Fall, dass $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und nach oben beschränkt ist. Es sei

$$M := \{a_n \,|\, n \in \mathbb{N}\}.$$

Dann ist M nicht leer und nach oben beschränkt, besitzt also nach dem Vollständigkeitsaxiom ein Supremum $\sup M =: a$.

Behauptung $\lim_{n\to\infty} a_n = a$.

Beweis. Es sei $\varepsilon > 0$ gegeben. Dann gibt es ein $n_0 \in \mathbb{N}$ mit

$$a - \varepsilon < a_{n_0} \leq a,$$

denn sonst wäre a nicht die kleinste obere Schranke von M. Da (a_n) monoton wachsend ist, gilt also für alle $n \geq n_0$

$$a - \varepsilon < a_{n_0} \leq a_n \leq a,$$

also $|a_n - a| < \varepsilon$. □

□

Wir wollen nun Anwendungen dieser Sätze darstellen.

Beispiel 5.2 Jeder Dezimalbruch, z.B.

$$0,\bar{9} = 0,99999999999999999999\ldots,$$

stellt eine reelle Zahl dar: Die Folge

$$0; 0,9; 0,99; 0,999; 0,9999; 0,99999; \ldots$$

ist monoton wachsend und nach oben beschränkt, konvergiert also nach Satz 5.12. Der Dezimalbruch bezeichnet gerade den Grenzwert dieser Folge. In unserem Beispiel ist das die Zahl 1. Also stellt der unendliche periodische Dezimalbruch $0,\bar{9}$ die Zahl 1 dar.

Beispiel 5.1 (7): Es seien $b, c > 0$ vorgegeben. Die Folge $(a_n)_{n\in\mathbb{N}}$ war rekursiv definiert durch

$$\begin{aligned} a_0 &= c, \\ a_{n+1} &= \frac{1}{2}\left(a_n + \frac{b}{a_n}\right) \quad (n \geq 0). \end{aligned}$$

Es gilt

$$a_{n+1} = \frac{1}{2}\left(a_n + \frac{b}{a_n}\right) \geq \sqrt{a_n \cdot \frac{b}{a_n}} = \sqrt{b}.$$

Somit ist $\sqrt{b}$ eine untere Schranke der Folge $(a_n)_{n\geq 1}$ ($a_0 = c$ außer Acht gelassen). Also gilt

$$\begin{aligned} & a_n \geq \sqrt{b} \\ \Rightarrow\ & a_n^2 \geq b \\ \Rightarrow\ & \frac{b}{a_n} \leq a_n \\ \Rightarrow\ & a_{n+1} = \frac{1}{2}\left(a_n + \frac{b}{a_n}\right) \leq a_n \end{aligned}$$

Also ist die Folge $(a_n)_{n\in\mathbb{N}}$ monoton fallend und nach unten beschränkt, also konvergent nach Satz 5.12. Es sei $a = \lim_{n\to\infty} a_n$. Wir berechnen a.

Es gilt $a_n \geq \sqrt{b} > 0$ für alle $n \geq 1$, also folgt aus Satz 5.10 $a > 0$. Nach Satz 5.6 und 5.9 gilt somit

$$a = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{1}{2}\left(a_n + \frac{b}{a_n}\right) = \frac{1}{2}\left(a + \frac{b}{a}\right).$$

Daraus folgt aber $a = \sqrt{b}$. Es gilt also $\lim_{n\to\infty} a_n = \sqrt{b}$.

Damit haben wir eine Methode gefunden, $\sqrt{b}$ approximativ zu berechnen. Dieses Verfahren geht schon auf die Babylonier zurück, es ist zur numerischen Berechnung von Quadratwurzeln bestens geeignet. Man probiere es für $\sqrt{2}$ aus!

Das Problem der stetigen Verzinsung

Eine alte Aufgabe von Jakob Bernoulli (1654–1705): Ein Kapital $a > 0$ wird nach einem Jahr mit 100% verzinst. Nach einem Jahr ist das Kapital also $2a$. Schlägt man die Zinsen schon nach einem halben Jahr hinzu, so ist das Kapital nach einem Jahr:

$$a\left(1 + \frac{1}{2}\right)\left(1 + \frac{1}{2}\right).$$

Schlägt man die Zinsen schon nach $\frac{1}{3}$ Jahr hinzu, so ist das Kapital nach einem Jahr

$$a\left(1 + \frac{1}{3}\right)\left(1 + \frac{1}{3}\right)\left(1 + \frac{1}{3}\right).$$

Teilt man das Jahr in n Teile und schlägt die Zinsen jeweils nach einem $\frac{1}{n}$ Jahr hinzu, so ist das Endkapital nach einem Jahr

$$a\left(1 + \frac{1}{n}\right)^n.$$

Was passiert mit wachsendem n?

Die Folge $\left(1+\frac{1}{n}\right)^n$ ist monoton wachsend:

$$\begin{aligned}
\left(1+\frac{1}{n}\right)^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} = \sum_{k=0}^{n}\frac{n(n-1)\cdots(n-k+1)}{k!}\cdot\frac{1}{n^k} \\
&= \sum_{k=0}^{n}\frac{1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{k!} \quad (\Rightarrow \text{ monoton wachsend}) \\
&\leq \sum_{n=0}^{n}\frac{1}{k!} = \frac{1}{1}+\frac{1}{1}+\frac{1}{2}+\frac{1}{2\cdot 3}+\frac{1}{2\cdot 3\cdot 4}+\cdots+\frac{1}{n!} \\
&\leq 1+1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\frac{1}{2^{n-1}} \\
&= 1+\frac{1-\left(\frac{1}{2}\right)^n}{1-\frac{1}{2}} \quad \text{(geometrische Summenformel)} \\
&< 3.
\end{aligned}$$

Also ist die Folge auch beschränkt, also konvergent.

Definition Die Zahl

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

heißt die *Eulersche Zahl.*

Für $n = 10^6$ ist

$$(1,000001)^{1\ 000\ 000} = 2,7182817.$$

Das ist e auf 5 Stellen genau.

Wir brauchen noch ein weiteres Konvergenzkriterium für Folgen reeller Zahlen. Für *monotone* Folgen ist die Beschränktheit eine notwendige und hinreichende Bedingung für die Konvergenz. Für *beliebige* Zahlenfolgen ist die Beschränktheit jedoch nur eine notwendige Bedingung für die Konvergenz. Allgemein gilt jedoch der folgende Satz, der ganz wesentlich auf dem Vollständigkeitsaxiom beruht:

Satz 5.13 (Satz von Bolzano-Weierstraß) *Jede beschränkte Folge besitzt eine konvergente Teilfolge.*

Dieser Satz folgt aus dem folgenden Lemma:

Lemma 5.1 *Jede Folge (a_n) besitzt eine Teilfolge, die entweder monoton wachsend oder monoton fallend ist.*

Beweis. Wir nennen eine natürliche Zahl n einen "Gipfel" der Folge (a_n), wenn $a_m < a_n$ für alle $m > n$ gilt.

Fall 1. Die Folge hat unendlich viele Gipfel. Wenn $n_0 < n_1 < n_2 < \cdots$ die Gipfel sind, so gilt $a_{n_0} > a_{n_1} > a_{n_2} > \cdots$ und $a_{n_0}, a_{n_1}, a_{n_2}, \ldots$ ist eine monoton fallende Teilfolge von (a_n).

Fall 2. Die Folge hat nur endlich viele Gipfel. In diesem Fall sei n_0 größer als alle Gipfel. Da n_0 kein Gipfel ist, gibt es eine natürliche Zahl n_1 mit $n_1 > n_0$ und $a_{n_1} \geq a_{n_0}$. Da n_1 kein Gipfel ist (n_1 ist größer als n_0 und damit größer als alle Gipfel), gibt es ein $n_2 > n_1$ mit $a_{n_2} \geq a_{n_1}$. Fahren wir auf diese Weise fort, so erhalten wir eine monoton wachsende Teilfolge $a_{n_0}, a_{n_1}, a_{n_2}, \ldots$ von (a_n). □

Wir führen nun den Begriff der Cauchyfolge ein.

Definition Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt eine *Cauchyfolge*, falls gilt: Zu jedem $\varepsilon > 0$ gibt es eine natürliche Zahl n_0, so dass für alle $m, n \in \mathbb{N}$ mit $m, n \geq n_0$ gilt: $|a_n - a_m| < \varepsilon$.

Satz 5.14 (Cauchysches Konvergenzkriterium) *Eine Folge $(a_n)_{n\in\mathbb{N}}$ ist genau dann konvergent, wenn sie eine Cauchyfolge ist.*

Beweis. (a) Wir zeigen zunächst: Jede konvergente Folge ist eine Cauchyfolge.

Wenn $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert, so kann man zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ finden, so dass für jedes $n \geq n_0$ gilt: $|a_n - a| < \frac{\varepsilon}{2}$. Wenn nun $n \geq n_0$ und $m \geq n_0$ ist, so gilt

$$|a_n - a_m| = |a_n - a + a - a_m| \leq |a_n - a| + |a - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(b) Wir zeigen nun den schwierigeren Teil: Jede Cauchyfolge $(a_n)_{n\in\mathbb{N}}$ ist konvergent.

(1) Wir zeigen zunächst: Die Folge $(a_n)_{n\in\mathbb{N}}$ ist beschränkt. Dazu wählen wir $\varepsilon = 1$ in der Definition einer Cauchyfolge. Dann gibt es ein $n_0 \in \mathbb{N}$ so dass für alle $m, n \geq n_0$ gilt

$$|a_m - a_n| < 1.$$

Dies gilt z.B. auch für $n = n_0$ und $m \geq n_0$:

$$|a_m - a_{n_0}| < 1.$$

Das besagt gerade, dass sich für fast alle m (für alle $m \geq n_0$) a_m um weniger als 1 von a_{n_0} unterscheidet. Die endlich vielen Ausnahmen stören die Beschränktheit nicht.

(2) Nach dem Satz von Bolzano-Weierstraß besitzt die Folge $(a_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit dem Grenzwert $a \in \mathbb{R}$.

(3) Wir zeigen: $\lim_{n\to\infty} a_n = a$. Es sei $\varepsilon > 0$ gegeben. Da $\lim_{n\to\infty} a_{n_k} = a$, gibt es ein $k_0 \in \mathbb{N}$, so dass für alle $k \geq k_0$ gilt:

$$|a_{n_k} - a| < \frac{\varepsilon}{2}.$$

Da $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist, gibt es ein n_0, so dass für alle $m, n \geq n_0$ gilt

$$|a_n - a_m| < \frac{\varepsilon}{2}.$$

Für alle $k \geq \max(k_0, n_0)$ gilt also wegen $n_k \geq k$

$$|a_k - a| = |a_k - a_{n_k} + a_{n_k} - a| \leq |a_k - a_{n_k}| + |a_{n_k} - a| < \varepsilon.$$

Also konvergiert $(a_n)_{n\in\mathbb{N}}$ gegen a. $\square$

Verfolgt man die Beweiskette zurück, so sieht man, das wir das Cauchysche Konvergenzkriterium aus dem Vollständigkeitsaxiom abgeleitet haben. Tatsächlich kann man anstelle des Vollständigkeitsaxioms auch das folgende Axiom in das Axiomensystem der reellen Zahlen aufnehmen und das Vollständigkeitsaxiom daraus ableiten.

Axiom *Der Körper $\mathbb{R}$ ist archimedisch angeordnet, und jede Cauchyfolge reeller Zahlen konvergiert in $\mathbb{R}$.*

Dieses Axiom gilt wieder nicht über $\mathbb{Q}$: Die Folge aus Beispiel 5.1 (7) ist eine Folge rationaler Zahlen, wenn b und c aus $\mathbb{Q}$ gewählt werden, die in $\mathbb{R}$ gegen $\sqrt{b}$ konvergiert, also eine Cauchyfolge. Sie konvergiert aber nicht in $\mathbb{Q}$, wenn z.B. $b = 2$ ist. Ebenso gilt der Satz von Bolzano-Weierstraß nicht über $\mathbb{Q}$.

6 Reihen

Wenn man eine Folge gegeben hat, so kann man auch versuchen, eine "Summe"

$$a_0 + a_1 + a_2 + \cdots$$

zu bilden. Wir wollen nun erklären, was wir darunter verstehen wollen. Zunächst kann man die "Partialsummen"

$$s_n = a_0 + a_1 + \cdots + a_n$$

betrachten. Konvergiert die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$, so kann man den Grenzwert als die "unendliche Summe"

$$a_0 + a_1 + a_2 + \cdots$$

betrachten.

Definition Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen. Durch

$$s_n := a_0 + a_1 + \cdots + a_n \quad (\textit{Partialsumme})$$

wird eine Folge $(s_n)_{n\in\mathbb{N}}$ definiert, die *Reihe* genannt wird und mit $\sum_{k=0}^{\infty} a_k$ bezeichnet wird.

Konvergiert die Folge $(s_n)_{n\in\mathbb{N}}$, so wird ihr Grenzwert ebenfalls mit $\sum_{k=0}^{\infty} a_k$ bezeichnet. Man sagt auch $\sum_{k=0}^{\infty} a_k$ *konvergiert.* Den Grenzwert nennt man auch die *Summe* oder den *Wert* der Reihe.

Das Symbol $\sum_{k=0}^{\infty} a_k$ bedeutet also

(a) die Folge $(\sum_{k=0}^{n} a_k)_{n\in\mathbb{N}}$ der Partialsummen

(b) im Falle der Konvergenz den Grenzwert $\lim_{n\to\infty} \sum_{k=0}^{n} a_k$.

Beispiel 6.1 Das wichtigste Beispiel einer Reihe ist die *geometrische Reihe*

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + q^3 + \cdots$$

für $q \in \mathbb{R}$. Die Partialsumme ist

$$s_n = 1 + q + \cdots + q^n.$$

Für $q = 1$ erhält man

$$s_n = \underbrace{1 + 1 + \cdots + 1}_{n+1} = n + 1.$$

Für $q = 1$ konvergiert die Reihe $\sum_{k=0}^{\infty} q^k$ daher nicht.

Es sei $q \neq 1$. Zieht man die beiden Gleichungen

$$\begin{aligned} s_n &= 1 + q + q^2 + \cdots + q^n \\ qs_n &= \phantom{1 + {}} q + q^2 + \cdots + q^n + q^{n+1} \end{aligned}$$

voneinander ab, so erhält man

$$(1 - q)s_n = 1 - q^{n+1}.$$

Für $q \neq 1$ gilt also

$$s_n = \frac{1 - q^{n+1}}{1 - q}.$$

Wir sehen also, dass die Folge $(s_n)_{n\in\mathbb{N}}$ genau dann konvergiert, wenn die Folge $(q^n)_{n\in\mathbb{N}}$ konvergiert. Nach Beispiel 5.1 (6) konvergiert die Folge $(q^n)_{n\in\mathbb{N}}$ aber genau dann, wenn $|q| < 1$ ist, und es gilt für $|q| < 1$

$$\lim_{n\to\infty} q^n = 0.$$

Also folgt für $|q| < 1$

$$\sum_{k=0}^{\infty} q^k = \lim_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q}.$$

Wir erhalten also die sehr wichtige *Summenformel für die geometrische Reihe*

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \quad \text{für } |q| < 1.$$

Insbesondere gilt also

$$\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2.$$

Beispiel 6.2 Es sei $a_k = \frac{1}{k(k+1)}$. Wir untersuchen die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)}.$$

Schreibe

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}.$$

Damit gilt

$$\begin{aligned} s_n &= \left(1-\frac{1}{2}\right) + \left(\frac{1}{2}-\frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1}-\frac{1}{n}\right) + \left(\frac{1}{n}-\frac{1}{n+1}\right) \\ &= 1 - \frac{1}{n+1}. \end{aligned}$$

Daher gilt

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \left(1 - \frac{1}{n+1}\right) = 1.$$

Die Theorien von Folgen und Reihen sind im folgenden Sinn dasselbe. Es sei $(s_n)_{n\in\mathbb{N}}$ eine beliebig vorgegebene Folge. Diese Folge ist die Partialsummenfolge der Reihe $\sum_{k=0}^{\infty} a_k$ mit

$$\begin{aligned} a_0 &= s_0, \\ a_k &= s_k - s_{k-1}, \end{aligned}$$

denn

$$\sum_{k=0}^{n} a_k = s_0 + (s_1 - s_0) + (s_2 - s_1) + \cdots + (s_n - s_{n-1}) = s_n.$$

Umgekehrt hat man ja zur Reihe $\sum_{k=0}^{\infty} a_k$ die Partialsummenfolge $(s_n)_{n\in\mathbb{N}}$. Auf diese Weise entsprechen sich auch die Konvergenzeigenschaften.

Satz 6.1 (Rechenregeln für Summen von Reihen) (1) *Ist* $\sum_{k=0}^{\infty} a_k$ *und* $\sum_{k=0}^{\infty} b_k$ *konvergent, so ist auch* $\sum_{k=0}^{\infty}(a_k + b_k)$ *konvergent und es gilt*

$$\sum_{k=0}^{\infty}(a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k.$$

(2) *Ist* $\sum_{k=0}^{\infty} a_k$ *konvergent und* $c \in \mathbb{R}$, *so ist auch* $\sum_{k=0}^{\infty} c \cdot a_k$ *konvergent und es gilt*

$$\sum_{k=0}^{\infty} c \cdot a_k = c \cdot \sum_{k=0}^{\infty} a_k.$$

(3) *Ist* $\sum_{k=0}^{\infty} a_k$ *und* $\sum_{k=0}^{\infty} b_k$ *konvergent und* $a_k \leq b_k$, *so gilt*

$$\sum_{k=0}^{\infty} a_k \leq \sum_{k=0}^{\infty} b_k.$$

Beweis. Es sei

$$s_n := \sum_{k=0}^{n} a_k, \qquad t_n := \sum_{k=0}^{n} b_k.$$

Aus dem Assoziativ- und Kommutativgesetz der Addition folgt

$$s_n + t_n = \sum_{k=0}^{n}(a_k + b_k).$$

Nach Satz 5.6 gilt

$$\begin{aligned} \sum_{k=0}^{\infty}(a_k + b_k) &= \lim_{n\to\infty}(s_n + t_n) = \lim_{n\to\infty} s_n + \lim_{n\to\infty} t_n \\ &= \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k. \end{aligned}$$

Die anderen Aussagen werden entsprechend auf Satz 5.6 und 5.10 zurückgeführt. □

Wir behandeln nun Konvergenzkriterien für Reihen.

Satz 6.2 (Cauchysches Konvergenzkriterium) *Die Reihe* $\sum_{k=0}^{\infty} a_k$ *konvergiert genau dann, wenn gilt: Zu jedem* $\varepsilon > 0$ *gibt es ein* $n_0 \in \mathbb{N}$, *so dass für alle* $n \geq m \geq n_0$ *gilt:*

$$|a_m + \cdots + a_n| < \varepsilon.$$

Beweis. Es gilt

$$s_n - s_{m-1} = a_m + \cdots a_n.$$

Die angegebene Bedingung bedeutet also gerade, dass die Partialsummenfolge $(s_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist. Nach Satz 5.14 ist aber $(s_n)_{n\in\mathbb{N}}$ genau dann konvergent, wenn $(s_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist. □

Beispiel 6.3 Wir betrachten die Reihe

$$\sum_{k=0}^{\infty}(-1)^k\frac{1}{2k+1}.$$

Wir zeigen mit dem Cauchyschen Konvergenzkriterium, dass diese Reihe konvergiert. Für $n > m$ gilt

$$\begin{aligned}\left|\sum_{k=m}^{n} a_k\right| &= \left|\frac{1}{2m+1} - \frac{1}{2m+3} + \frac{1}{2m+5} - \cdots\right| \\ &= \left|\frac{1}{2m+1} - \left(\frac{1}{2m+3} - \frac{1}{2m+5}\right) - \cdots\right| \\ &< \frac{1}{2m+1}.\end{aligned}$$

Für $n = m$ gilt

$$\left|\sum_{k=m}^{n} a_k\right| = \frac{1}{2m+1}.$$

Wählt man nun zu vorgegebenem $\varepsilon > 0$ die Zahl n_0 so, dass

$$\frac{1}{2n_0+1} < \varepsilon,$$

dann gilt für alle $n \geq m \geq n_0$

$$\left|\sum_{k=m}^{n} a_k\right| < \varepsilon.$$

Also ist das Cauchykriterium erfüllt und die Reihe konvergiert.

Der Grenzwert ist übrigens $\frac{\pi}{4}$, die Reihe wurde schon von G. W. Leibniz (1646-1716) betrachtet.

Korollar 6.1 *Ist $\sum_{k=0}^{\infty} a_k$ konvergent, so gilt*

$$\lim_{n\to\infty} a_n = 0.$$

Beweis. Man nehme im Cauchyschen Konvergenzkriterium $m = n$. □

Satz 6.3 (Beschränkheitskriterium) *Eine Reihe* $\sum_{k=0}^{\infty} a_k$ *mit* $a_k \geq 0$ *konvergiert genau dann, wenn die Partialsummenfolge* $(s_n)_{n\in\mathbb{N}}$ *beschränkt ist.*

Beweis. Die Partialsummenfolge $(s_n)_{n\in\mathbb{N}}$ einer solchen Reihe ist eine monoton wachsende Folge, die genau dann konvergiert, wenn sie beschränkt ist (Satz 5.12). □

Beispiel 6.4 Die *harmonische Reihe*

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

konvergiert nicht, denn

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{>\frac{1}{2}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{>\frac{1}{2}} + \underbrace{\frac{1}{9} + \cdots + \frac{1}{16}}_{>\frac{1}{2}} + \cdots .$$

Man findet also eine Teilfolge der Partialsummenfolge, die nicht beschränkt ist und daher nicht konvergiert.

Definition Eine Reihe $\sum_{k=0}^{\infty} a_k$ heißt *alternierende Reihe*, falls $a_{2j} \geq 0$ und $a_{2j+1} \leq 0$ für alle $j \in \mathbb{N}$ (oder umgekehrt).

Satz 6.4 (Leibniz-Kriterium) *Es sei*

$$a_0 \geq a_1 \geq a_2 \geq \cdots \geq 0$$

und es gelte

$$\lim_{n\to\infty} a_n = 0.$$

Dann konvergiert die Reihe

$$\sum_{n=0}^{\infty} (-1)^n a_n = a_0 - a_1 + a_2 - a_3 + a_4 - \cdots .$$

Beweis. Von der Folge (s_n) betrachten wir die beiden Teilfolgen (s_{2m}) und (s_{2m+1}). Die Teilfolge (s_{2m}) ist monoton fallend:

$$s_{2m+2} = s_{2m} - a_{2m+1} + a_{2m+2} \leq s_{2m}.$$

Die Teilfolge (s_{2m+1}) ist monoton wachsend:

$$s_{2m+3} = s_{2m+1} + a_{2m+2} - a_{2m+3} \geq s_{2m+1}.$$

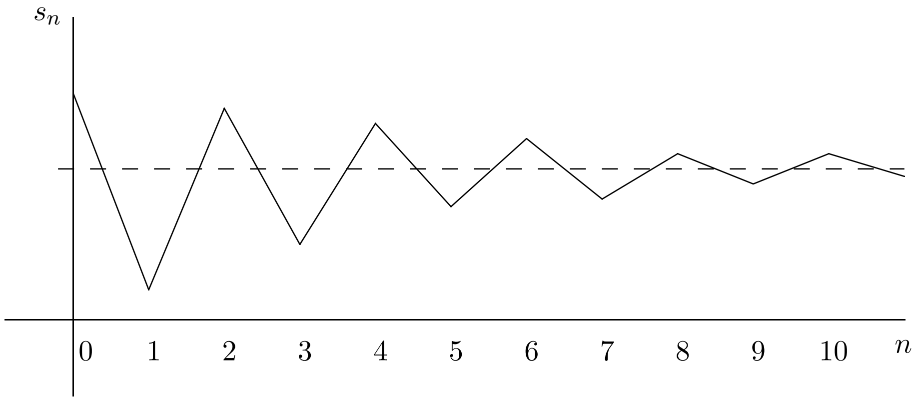


Abbildung 1: Partialsummen einer alternierenden Reihe

Außerdem gilt wegen $s_{2m+1} - s_{2m} = -a_{2m+1} \leq 0$

$$s_{2m+1} \leq s_{2m} \quad \text{für alle } m \in \mathbb{N}.$$

Die Folge (s_{2m}) ist durch s_1 nach unten, (s_{2m+1}) durch s_0 nach oben beschränkt. Beide Teilfolgen sind also konvergent. Es sei

$$\begin{aligned} s &:= \lim_{m\to\infty} s_{2m}, \\ s' &:= \lim_{m\to\infty} s_{2m+1}. \end{aligned}$$

Wir zeigen nun, dass $s = s'$ und dass die gesamte Folge $(s_n)_{n\in\mathbb{N}}$ gegen s konvergiert.

Zunächst ist

$$s' - s = \lim_{m\to\infty}(s_{2m+1} - s_{2m}) = \lim_{m\to\infty}(-a_{2m+1}) = 0.$$

Es sei nun $\varepsilon > 0$ vorgegeben. Dann gibt es $n_1, n_2 \in \mathbb{N}$, so dass

$$\begin{aligned} |s_{2m} - s| &< \varepsilon \quad \text{für } 2m \geq n_1, \\ |s_{2m+1} - s| &< \varepsilon \quad \text{für } 2m+1 \geq n_2. \end{aligned}$$

Für $n \geq n_0 := \max(n_1, n_2)$ gilt dann

$$|s_n - s| < \varepsilon.$$

□

Aus dem Leibniz-Kriterium folgt sofort, dass die Reihe von Beispiel 6.3 konvergiert.

Umordnung von Reihen, absolute Konvergenz

Eine Summe von endlich vielen Zahlen ist unabhängig von der Reihenfolge. Bei einer Reihe kann jedoch das Umordnen der Glieder zu seltsamen Resultaten führen.

Beispiel 6.5 Wir betrachten die alternierende Reihe

$$\sum_{k=0}^{\infty}(-1)^k\frac{1}{k+1}.$$

Nach dem Leibniz-Kriterium konvergiert diese Reihe gegen eine Zahl a. Nach dem Beweis des Leibnizkriteriums gilt

$$s_1=\frac{1}{2}\leq a\leq s_0=1.$$

Wir ordnen nun die Reihe um:

$$\begin{aligned}
a &= 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots\\
&= 1-\frac{1}{2}-\frac{1}{4}+\frac{1}{3}-\frac{1}{6}-\frac{1}{8}+\frac{1}{5}-\frac{1}{10}-\frac{1}{12}+\cdots\\
&\ \text{(auf ein positives Glied folgen stets zwei negative Glieder)}\\
&= \left(1-\frac{1}{2}\right)-\frac{1}{4}+\left(\frac{1}{3}-\frac{1}{6}\right)-\frac{1}{8}+\left(\frac{1}{5}-\frac{1}{10}\right)-\frac{1}{12}+\cdots\\
&= \frac{1}{2}-\frac{1}{4}+\frac{1}{6}-\frac{1}{8}+\frac{1}{10}-\frac{1}{12}+\cdots\\
&= \frac{1}{2}\left(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots\right)\\
&= \frac{1}{2}a.
\end{aligned}$$

Es gilt also $a=\frac{1}{2}a$, und es folgt $a=0$, ein Widerspruch.

Definition Es sei $\sigma:\mathbb{N}\to\mathbb{N}$ bijektiv und $b_k=a_{\sigma(k)}$. Dann heißt die Reihe $\sum_{k=0}^{\infty}b_k$ eine *Umordnung* der Reihe $\sum_{k=0}^{\infty}a_k$.

Bemerkung 6.1 Die Reihe $\sum_{k=0}^{\infty}a_k$ ist dann auch eine Umordnung der Reihe $\sum_{k=0}^{\infty}b_k$ und zwar mit der Bijektion $\sigma^{-1}:\mathbb{N}\to\mathbb{N}$ und $a_k=b_{\sigma^{-1}(k)}$.

Satz 6.5 *Es sei $a_k\geq 0$ für alle $k\in\mathbb{N}$ und die Reihe $\sum_{k=0}^{\infty}a_k$ konvergent mit der Summe a. Dann ist auch jede Umordnung $\sum_{k=0}^{\infty}a_{\sigma(k)}$ konvergent mit der Summe a.*

Beweis. Setze

$$\begin{aligned} s_n &= \sum_{k=0}^{n} a_k, \\ s'_n &= \sum_{k=0}^{m} a_{\sigma(k)}. \end{aligned}$$

Die Folgen (s_n) und (s'_m) sind monoton wachsend. Wir zeigen, dass mit (s_n) auch die Folge (s'_m) durch a nach oben beschränkt ist.

Es sei m gegeben. Setze

$$n_m := \max\{\sigma(0), \sigma(1), \ldots, \sigma(m)\}.$$

Dann gilt

$$s'_m \leq s_{n_m} \leq a,$$

da alle Reihenglieder nicht negativ sind und alle Summanden von s'_m auch in s_{n_m} vorkommen.

Die umgeordnete Reihe ist also konvergent mit einer Summe $a' \leq a$ (Satz 5.10).

Da die ursprüngliche Reihe eine Umordnung der zweiten Reihe ist, folgt entsprechend $a \leq a'$. Also $a = a'$. □

Auch bei Reihen, in denen es Glieder mit verschiedenen Vorzeichen gibt, kann jede Umordnung wieder zu einer konvergenten Reihe mit derselben Summe führen. Es sind dies genau die absolut konvergenten Reihen.

Definition Eine Reihe $\sum_{k=0}^{\infty} a_k$ heißt *absolut konvergent*, wenn die Reihe $\sum_{k=0}^{\infty} |a_k|$ konvergent ist.

Satz 6.6 *Wenn eine Reihe absolut konvergent ist, dann ist sie konvergent.*

Bemerkung 6.2 Die Umkehrung ist falsch, wie das Beispiel $\sum_{k=0}^{\infty}(-1)^k \frac{1}{k+1}$ zeigt.

Beweis. Nach dem Cauchyschen Konvergenzkriterium gibt es zu jedem $\varepsilon > 0$ ein n_0, so dass für alle $m, n \geq n_0$ gilt

$$\sum_{k=m}^{n} |a_k| < \varepsilon.$$

Wegen der Dreiecksungleichung ist dann aber erst recht auch

$$\left| \sum_{k=m}^{n} a_k \right| < \varepsilon.$$

□

Satz 6.7 (Umordnungssatz) *Es sei* $\sum_{k=0}^{\infty} a_k$ *absolut konvergent. Dann ist auch jede Umordnung* $\sum_{k=0}^{\infty} a_{\sigma(k)}$ *absolut konvergent und es gilt*

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_{\sigma(k)}.$$

Beweis. Nach Satz 6.5 konvergiert $\sum_{k=0}^{\infty} |a_{\sigma(k)}|$ als Umordnung von $\sum_{k=0}^{\infty} |a_k|$.

Noch zu zeigen:

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_{\sigma(k)}.$$

(Satz 6.5 liefert nur $\sum_{k=0}^{\infty} |a_k| = \sum_{k=0}^{\infty} |a_{\sigma(k)}|$.)

Es sei n gegeben. Es sei k_n die kleinste Zahl $k \in \mathbb{N}$ mit

$$\{0, 1, 2, \ldots, n\} \subseteq \{\sigma(0), \sigma(1), \ldots, \sigma(k)\}.$$

Dann gilt $k_n \geq n$ und die Differenz

$$s'_{k_n} - s_n$$

(Notation wie beim Beweis von Satz 6.5) enthält nur solche Glieder, deren Index größer als n ist. Da $\sum_{k=0}^{\infty} a_k$ absolut konvergent ist, gibt es zu vorgegebenem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass für alle $m, n \geq n_0$

$$\sum_{k=m}^{n} |a_k| < \varepsilon.$$

Also gilt für alle $n \geq n_0$

$$|s'_{k_n} - s_n| < \varepsilon,$$

d.h. $(s'_{k_n} - s_n)_{n\in\mathbb{N}}$ ist eine Nullfolge. Wir wissen: (s'_n) und (s_n) konvergieren beide, da die Folgen absolut konvergent sind. Daraus folgt $\lim_{n\to\infty} s'_n = \lim_{n\to\infty} s_n$, also

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_{\sigma(k)}.$$

□

Bemerkung 6.3 Ist $\sum_{k=0}^{\infty} a_k$ konvergent, aber nicht absolut konvergent, dann kann man durch Umordnung jede beliebige Zahl s als Summe erhalten:

Setze

$$a_k^+ = \frac{a_k + |a_k|}{2} = \begin{cases} a_k & \text{falls } a_k > 0, \\ 0 & \text{falls } a_k \leq 0, \end{cases}$$

$$a_k^- = \frac{a_k - |a_k|}{2} = \begin{cases} 0 & \text{falls } a_k \geq 0, \\ a_k & \text{falls } a_k < 0. \end{cases}$$

Es gilt
$$a_k = a_k^+ + a_k^- .$$
Wegen der Konvergenz der Reihe $\sum_{k=0}^{\infty} a_k$ sind die Reihen $\sum_{k=0}^{\infty} a_k^+$ und $\sum_{k=0}^{\infty} a_k^-$ beide konvergent oder beide divergent. (Denn angenommen, eine der beiden Reihen konvergiert. Dann konvergiert auch die andere als Differenz zweier konvergenter Reihen.) Konvergieren beide, so wäre wegen $|a_k| = a_k^+ - a_k^-$ die Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergent im Gegensatz zur Voraussetzung.

Weglassen von Nullen ergibt die Folgen

$$\begin{array}{lcl} (a_k^+) & \rightsquigarrow & p_0, p_1, p_2, \ldots \\ (a_k^-) & \rightsquigarrow & q_0, q_1, q_2, \ldots \end{array}$$

Wähle nun: k_0 kleinste natürliche Zahl mit
$$\sum_{n=0}^{k_0} p_n > s,$$
k_1 kleinste natürliche Zahl mit
$$\sum_{n=0}^{k_0} p_n + \sum_{n=0}^{k_1} q_n < s,$$
k_2 kleinste natürliche Zahl mit
$$\sum_{n=0}^{k_0} p_n + \sum_{n=0}^{k_1} q_n + \sum_{n=k_0+1}^{k_2} p_n > s,$$
usw. Es ist klar, welcher Umordnung der ursprünglichen Reihe dies entspricht.

Wegen der Wahl von $k_0, k_1, k_2, \ldots$ unterscheiden sich die zugehörigen Teilsummen höchstens um $p_{k_0}, q_{k_1}, p_{k_2}, \ldots$ von der Zahl s. Da die Folgen (p_n) und (q_n) Nullfolgen sind, konvergieren die Teilsummen der umgeordneten Reihe gegen s.

Man kann sogar so umordnen, dass die Reihe $\sum_{k=0}^{\infty} a_{\sigma(k)}$ divergent ist! (Beweis als Aufgabe).

Satz 6.8 *Eine Reihe $\sum_{k=0}^{\infty} a_k$ ist genau dann absolut konvergent, wenn jede ihrer Umordnungen konvergent ist.*

Beweis. "$\Rightarrow$": Satz 6.7.

"$\Leftarrow$": Wäre $\sum_{k=0}^{\infty} a_k$ nicht absolut konvergent, so könnte man eine Umordnung finden, die divergent ist. $\square$

Nun kommen wir zur *Multiplikation von Reihen.* Was ist

$$\left(\sum_{j=0}^{\infty} a_j\right) \cdot \left(\sum_{k=0}^{\infty} b_k\right)?$$

Wir ordnen die vorkommenden Produkte in einem quadratischen Schema an:

$$\begin{tikzcd}
a_0b_0 \arrow[r] & a_0b_1 \arrow[dl] & a_0b_2 \arrow[r] & a_0b_3 \arrow[dl] & \cdots \\
a_1b_0 \arrow[d] & a_1b_1 \arrow[ur] & a_1b_2 \arrow[dl] & a_1b_3 \arrow[ur] & \cdots \arrow[dl] \\
a_2b_0 \arrow[ur] & a_2b_1 \arrow[dl] & a_2b_2 \arrow[ur] & a_2b_3 \arrow[dl] & \cdots \\
a_3b_0 \arrow[d] & a_3b_1 \arrow[ur] & a_3b_2 \arrow[dl] & a_3b_3 \arrow[ur] & \cdots \\
\vdots \arrow[ur] & \vdots & \vdots \arrow[ur] & \vdots &
\end{tikzcd}$$

Es ist nicht von vornherein klar, welche Reihenfolge zu wählen ist. Nahe liegt die Anordnung nach "Schrägzeilen". Gilt nun

$$\left(\sum_{j=0}^{\infty} a_j\right) \cdot \left(\sum_{k=0}^{\infty} b_k\right) = \sum_{n=0}^{\infty} c_n \text{ mit } c_n = \sum_{j=0}^{n} a_j b_{n-j}?$$

Die Antwort gibt der folgende Satz, den wir nicht beweisen wollen.

Satz 6.9 (Produkt von Reihen) *Die Reihen* $\sum_{j=0}^{\infty} a_j$ *und* $\sum_{k=0}^{\infty} b_k$ *seien absolut konvergent und es sei*

$$c_n = \sum_{j=0}^{n} a_j b_{n-j}.$$

Dann ist auch die Reihe $\sum_{n=0}^{\infty} c_n$ *absolut konvergent und es gilt*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{j=0}^{\infty} a_j\right) \cdot \left(\sum_{k=0}^{\infty} b_k\right).$$

Bemerkung 6.4 Wenn beide Reihen nicht absolut konvergent sind, braucht die Aussage nicht richtig zu sein.

Wir geben nun einige Kriterien für die absolute Konvergenz von Reihen an.

Satz 6.10 (Majorantenkriterium) *Es sei $\sum_{n=0}^{\infty} b_n$ eine konvergente Reihe mit lauter nichtnegativen Gliedern und $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $|a_n| \leq b_n$ für alle $n \in \mathbb{N}$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$ absolut.*

Notation Man nennt dann $\sum_{n=0}^{\infty} b_n$ *Majorante* von $\sum_{n=0}^{\infty} a_n$.

Beweis. Die monoton wachsende Folge $(\sum_{k=0}^{n} |a_k|)$ ist durch $\sum_{k=0}^{\infty} b_k$ nach oben beschränkt, also konvergent. □

Beispiel 6.6 Wir beweisen die Konvergenz der Reihen

$$\sum_{n=1}^{\infty} \frac{1}{n^k}, \quad k \geq 2,$$

mit Hilfe des Majorantenkriteriums.

Für $k \geq 2$ und alle $n \geq 1$ gilt

$$\frac{1}{n^k} \leq \frac{1}{n^2} \leq \frac{2}{n(n+1)} \left(\Leftrightarrow n(n+1) \leq 2n^2\right).$$

Nach Beispiel 6.2 konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$, also auch (Satz 6.1) die Reihe $\sum_{n=1}^{\infty} \frac{2}{n(n+1)}$. Diese Reihe ist daher Majorante von $\sum_{n=1}^{\infty} \frac{1}{n^k}$.

Bemerkung 6.5 Satz 6.10 impliziert folgendes Divergenzkriterium: Es sei $\sum_{n=0}^{\infty} b_n$ eine divergente Reihe mit lauter nicht negativen Gliedern und $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $a_n \geq b_n$ für alle $n \in \mathbb{N}$. Dann ist auch die Reihe $\sum_{n=0}^{\infty} a_n$ divergent. (Denn andernfalls wäre $\sum_{n=0}^{\infty} a_n$ eine konvergente Majorante von $\sum_{n=0}^{\infty} b_n$, also müsste auch $\sum_{n=0}^{\infty} b_n$ konvergieren.)

Satz 6.11 *Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe. Gibt es ein q, $0 \leq q < 1$, so dass für alle $n \in \mathbb{N}$ gilt $|a_n| \leq q^n$, dann ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.*

Warnung Sind alle $a_n \leq q = 1$, so folgt die Konvergenz nicht!

Beweis. Die Voraussetzung bedeutet, dass die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ eine konvergente Majorante von $\sum_{n=0}^{\infty} |a_n|$ ist. □

Man formuliert Satz 6.11 normalerweise, indem man auf beiden Seiten der Ungleichung $|a_n| \leq q^n$ die n-te Wurzel zieht. Dazu müssen wir zeigen, dass es zu $a \in \mathbb{R}$, $a \geq 0$, genau ein $x \geq 0$ gibt mit $x^n = a$. Dieses x bezeichnet man als die n-te Wurzel aus a, Schreibweise $\sqrt[n]{a}$. Den Beweis müssen wir an dieser Stelle noch vertagen.

Satz 6.12 (Wurzelkriterium) *Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe. Gibt es ein q, $0 \leq q < 1$, so dass für fast alle $n \in \mathbb{N}$*

$$\sqrt[n]{|a_n|} \leq q$$

ist, dann ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

Leichter zu handhaben als das Wurzelkriterium, dafür aber nicht so weitreichend, ist das Quotientenkriterium.

Satz 6.13 (Quotientenkriterium) *Es sei* $\sum_{n=0}^{\infty} a_n$ *eine Reihe. Gibt es ein* q*,* $0 \leq q < 1$*, so dass für fast alle* $n \in \mathbb{N}$

$$\left|\frac{a_{n+1}}{a_n}\right| \leq q,$$

dann ist $\sum_{n=0}^{\infty} a_n$ *absolut konvergent.*

Bemerkung 6.6 Es ist natürlich, so etwas zu erwarten. Bei der geometrischen Reihe ist der Quotient zweier aufeinanderfolgender Reihenglieder konstant gleich q. Bei obiger Reihe $\sum_{n=0}^{\infty} a_n$ können also für fast alle n die Quotienten der Reihenglieder durch $q < 1$ abgeschätzt werden.

Beweis. Es sei also für $n \geq n_0$

$$\left|\frac{a_{n+1}}{a_n}\right| \leq q.$$

Daraus ergibt sich mittels vollständiger Induktion

$$|a_{n_0+r}| \leq |a_{n_0}| \cdot q^r.$$

Durch Weglassen endlich vieler Glieder der Reihe $\sum_{n=0}^{\infty} |a_n|$ erhält man die Reihe $\sum_{r=0}^{\infty} |a_{n_0+r}|$, die die Reihe $\sum_{r=0}^{\infty} |a_{n_0}| \cdot q^r$ als konvergente Majorante besitzt. □

Beispiel 6.7 Wir zeigen mit Hilfe des Quotientenkriteriums, dass die Reihe

$$\sum_{n=0}^{\infty} \frac{n^2}{2^n}$$

konvergiert: Mit $a_n := \frac{n^2}{2^n}$ gilt für alle $n \geq 3$:

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^2 2^n}{2^{n+1} n^2} = \frac{1}{2}\left(1 + \frac{1}{n}\right)^2 \leq \frac{1}{2}\left(1 + \frac{1}{3}\right)^2 = \frac{8}{9} = q < 1.$$

Beispiel 6.8 Wir untersuchen Quotienten- und Wurzelkriterium für die *divergente* harmonische Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n} \quad \text{(siehe Beispiel 6.4).}$$

Hier gilt zwar

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n}{n+1} < 1 \quad \text{für alle } n \geq 1,$$

wegen $\lim_{n\to\infty} \frac{n}{n+1} = 1$ gibt es jedoch *kein* $q < 1$ mit

$$\left|\frac{a_{n+1}}{a_n}\right| \leq q \quad \text{für fast alle } n \in \mathbb{N}.$$

Ähnliches gilt für das Wurzelkriterium: Man kann zeigen

$$\sqrt[n]{|a_n|} = \sqrt[n]{\frac{1}{n}} = \frac{1}{\sqrt[n]{n}} < 1 \quad \text{für alle } n \geq 2,$$

und $\left(\frac{1}{\sqrt[n]{n}}\right)_{n\geq 3}$ ist monoton wachsend und konvergiert gegen 1. Das Wurzelkriterium versagt also auch.

Auch bei der *konvergenten* Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

versagen das Quotienten- und das Wurzelkriterium:

$$\begin{aligned}
\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{n^2}{(n+1)^2} = 1, \\
\lim_{n\to\infty} \sqrt[n]{|a_n|} &= \lim_{n\to\infty} \frac{1}{\sqrt[n]{n^2}} = \lim_{n\to\infty} \frac{1}{\sqrt[n]{n}} \cdot \frac{1}{\sqrt[n]{n}} = 1.
\end{aligned}$$

7 Die Exponentialreihe

Es sei $x \in \mathbb{R}$. Zu diesem x betrachten wir die *Exponentialreihe* (Potenzreihe)

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Satz 7.1 *Für jedes $x \in \mathbb{R}$ ist die Exponentialreihe $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ absolut konvergent.*

Beweis. Die Behauptung folgt aus dem Quotientenkriterium, denn für alle $n \geq 2|x|$ gilt

$$\left|\frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}}\right| = \frac{|x|}{n+1} \leq \frac{1}{2}.$$

□

Satz 7.2 (Abschätzung des Restglieds) *Es gilt*

$$\exp(x) = \sum_{n=0}^{N} \frac{x^n}{n!} + r_{N+1}(x),$$

wobei

$$|r_{N+1}(x)| \leq 2\frac{|x|^{N+1}}{(N+1)!} \quad \textit{für } |x| \leq 1 + \frac{N}{2}.$$

Beweis. Wir schätzen den Rest

$$r_{N+1}(x) = \sum_{n=N+1}^{\infty} \frac{x^n}{n!}$$

mittels der geometrischen Reihe ab. Es ist

$$\begin{aligned}
|r_{N+1}(x)| &\leq \sum_{n=N+1}^{\infty} \left|\frac{x^n}{n!}\right| \\
&= \frac{|x|^{N+1}}{(N+1)!}\left(1 + \frac{|x|}{N+2} + \frac{|x|^2}{(N+2)(N+3)} + \right. \\
&\qquad \left. \cdots + \frac{|x|^k}{(N+2)\cdots(N+k+1)} + \cdots\right) \\
&\leq \frac{|x|^{N+1}}{(N+1)!}\left(1 + \frac{|x|}{N+2} + \left(\frac{|x|}{N+2}\right)^2 + \cdots + \left(\frac{|x|}{N+2}\right)^k + \cdots\right) \\
&\leq \frac{|x|^{N+1}}{(N+1)!}\left(\frac{1}{1 - \frac{|x|}{N+2}}\right).
\end{aligned}$$

Für $|x| \leq \frac{N+2}{2} = 1 + \frac{N}{2}$ gilt deshalb

$$|r_{N+1}(x)| \leq 2\frac{|x|^{N+1}}{(N+1)!}.$$

□

Wir untersuchen nun die Folge $\left(\left(1 + \frac{x}{n}\right)^n\right)_{n\geq 1}$ für festes $x \geq 0$. Für $x = 1$ hatten wir schon gesehen, dass diese Folge gegen e konvergiert.

Satz 7.3 *Es sei $x \geq 0$. Dann gilt:*

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

insbesondere

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

Beweis.

$$\begin{aligned}\left(1+\frac{x}{n}\right)^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{x^k}{n^k}\\ &= \sum_{k=0}^{n}\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{x^k}{k!}\\ &= \sum_{k=0}^{n}1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\cdot\frac{x^k}{k!}\\ &\le \exp(x).\end{aligned}$$

Also ist die Folge $\left(\left(1+\frac{x}{n}\right)^n\right)_{n\geq 1}$ $(x\geq 0)$ nach oben beschränkt. Da sie auch monoton wachsend ist, ist sie konvergent und es folgt

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \le \exp(x).$$

Es gilt (man beachte, dass die Summe bis m läuft)

$$\left(1+\frac{x}{n}\right)^n \ge \sum_{k=0}^{m}1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\cdot\frac{x^k}{k!}\quad\text{für } m\le n.$$

Es sei nun m fest und n laufe. Dann konvergiert die linke Seite und ebenso die rechte Seite. Die rechte Seite konvergiert gegen $\sum_{k=0}^{m}\frac{x^k}{k!}$. Also gilt

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \ge \sum_{k=0}^{m}\frac{x^k}{k!}\quad\text{für } \textit{jedes } m\in\mathbb{N}.$$

Also ist auch

$$\exp(x)=\sum_{k=0}^{\infty}\frac{x^k}{k!}\le\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n.$$

□

Bemerkung 7.1 Man darf nicht einfach in der Formel

$$\sum_{k=0}^{n}\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{x^k}{k!}$$

zum Limes übergehen, da man dann für verschiedene n die Partialsummen *völlig verschiedener* Reihen bildet.

Wie kann man $a\cdot\exp(x)$ deuten? Lässt man ein Kapital a bei stetiger Verzinsung zu 100% genau x Jahre liegen, so wächst es auf $a\cdot\exp(x)$. Dies folgt wieder wie oben: $a\left(1+\frac{x}{n}\right)^n$ ist das Kapital nach x Jahren, wenn die Zinsen alle $\frac{x}{n}$ Jahre zugeschlagen werden. Der Grenzübergang zeigt die Behauptung.

Eine andere Deutung: Lässt man ein Kapital a bei stetiger Verzinsung zu $x \cdot 100\%$ ein Jahr liegen, so ist das Endkapital $a \exp(x)$.

Hieraus lässt sich folgende Eigenschaft der Exponentialreihe erraten:

$$\exp(x+y) = \exp(x) \cdot \exp(y).$$

Die Zahl $\exp(x+y)$ ist das Endkapital bei stetiger Verzinsung zu 100%, wenn man ein Anfangskapital 1 genau $(x+y)$ Jahre liegen lässt. Die Zahl $\exp(x)\exp(y)$ ist das Endkapital, wenn man dasselbe Anfangskapital erst x Jahre stehen lässt und dann das bis dahin gewonnene Kapital weitere y Jahre verzinsen lässt.

Satz 7.4 (Funktionalgleichung der Exponentialfunktion) *Für alle $x, y \in \mathbb{R}$ gilt*

$$\exp(x+y) = \exp(x)\exp(y).$$

Beweis. Wir wenden Satz 6.9 auf die absolut konvergenten Reihen

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \text{ und } \exp(y) = \sum_{n=0}^{\infty} \frac{y^n}{n!}$$

an. Nach dem binomischen Lehrsatz ist

$$\begin{aligned} c_n &= \sum_{k=0}^{n} \frac{x^{n-k}}{(n-k)!} \cdot \frac{y^k}{k!} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\ &= \frac{1}{n!}(x+y)^n. \end{aligned}$$

Deshalb folgt

$$\exp(x+y) = \sum_{n=0}^{\infty} c_n = \exp(x)\exp(y).$$

□

Korollar 7.1 (i) *Für alle $x \in \mathbb{R}$ gilt* $\exp(x) > 0$.

(ii) *Für alle $x \in \mathbb{R}$ gilt* $\exp(-x) = (\exp(x))^{-1}$.

(iii) *Für jede ganze Zahl n ist* $\exp(n) = e^n$.

Beweis.

Zu (ii):

$$\exp(x)\exp(-x) = \exp(x-x) = \exp(0) = 1.$$

Daraus folgt insbesondere $\exp(x) \neq 0$ und

$$\exp(-x) = \frac{1}{\exp(x)}.$$

Zu (i): Für alle $x \geq 0$ ist

$$\exp(x) = 1 + x + \frac{x^2}{2} + \cdots \geq 1 > 0.$$

Falls $x < 0$, ist $-x > 0$, also $\exp(-x) > 0$ und

$$\exp(x) = \frac{1}{\exp(-x)} > 0.$$

Zu (iii): Wir zeigen zunächst mittels vollständiger Induktion, dass für alle $n \in \mathbb{N}$ gilt: $\exp(n) = e^n$.

Induktionsanfang $n = 0$: $\exp(0) = 1 = e^0$.

Induktionsschritt $n \to n+1$:

$$\exp(n+1) \stackrel{\text{F.G.}}{=} \exp(n)\exp(1) \stackrel{\text{I.V.}}{=} e^n e = e^{n+1}.$$

Damit ist $\exp(n) = e^n$ für alle $n \in \mathbb{N}$ bewiesen. Aus (ii) folgt

$$\exp(-n) = \frac{1}{\exp(n)} = \frac{1}{e^n} = e^{-n} \quad \text{für alle } n \in \mathbb{N}.$$

Daher gilt $\exp(n) = e^n$ für alle ganzen Zahlen n. $\square$

Zusammenfassung: Die Exponentialreihe definiert eine Abbildung

$$\begin{array}{rccl} \exp: & \mathbb{R} & \longrightarrow & \mathbb{R}_+^* := \{r \in \mathbb{R} \mid r > 0\} \\ & x & \longmapsto & \exp(x) \end{array}.$$

Hier sind $\mathbb{R}$ und $\mathbb{R}_+^*$ Gruppen, $\mathbb{R}$ mit der Addition, $\mathbb{R}_+^*$ mit der Multiplikation. Die Funktionalgleichung zeigt, dass exp ein *Homomorphismus* dieser Gruppen ist.

Korollar 7.2 *Für alle $x_1, x_2 \in \mathbb{R}$ gilt: Falls $x_1 < x_2$, so ist $\exp(x_1) < \exp(x_2)$.*

Beweis. Falls $x_1 < x_2$, so gilt

$$\exp(x_2) = \exp(x_1 + (x_2 - x_1)) = \underbrace{\exp(x_1)}_{>0} \cdot \underbrace{\exp(x_2 - x_1)}_{>1} > \exp(x_1).$$

$\square$

Satz 7.5 (i) *Für alle $x \in \mathbb{R}$ gilt: $1 + x \leq \exp(x)$.*

(ii) *Für alle $x \in \mathbb{R}$, $x < 1$, gilt: $\exp(x) \leq \frac{1}{1-x}$.*

(iii) *Für alle $x, y \in \mathbb{R}$ mit $y - x < 1$ gilt:*

$$(y - x) \cdot \exp(x) \leq \exp(y) - \exp(x) \leq \frac{y - x}{1 + x - y} \cdot \exp(x).$$

(iv) *Zu jedem $a \in \mathbb{R}$ existiert ein $L > 0$ mit*

$$x, y \leq a \Rightarrow |\exp(x) - \exp(y)| \leq L|x - y|.$$

Beweis. Zu (i): Für $x \geq 0$ ist

$$\exp(x) = 1 + x + \frac{x^2}{2} + \cdots \geq 1 + x,$$

für $x \leq -1$ gilt

$$1 + x \leq 0 < \exp(x)$$

nach Korollar 7.1.

Es sei also $-1 < x < 0$. Nach Satz 7.2 gilt

$$\exp(x) = 1 + x + \frac{x^2}{2} + r_3(x),$$

$$|r_3(x)| \leq 2 \cdot \frac{|x|^3}{3!} = \frac{|x|^3}{3}.$$

Also gilt

$$\frac{x^2}{2} + r_3(x) \geq \frac{x^2}{2} + \frac{x^2}{3} \geq 0,$$

und daher $1 + x \leq \exp(x)$.

Zu (ii): Aus

$$1 - x \leq \exp(-x) = \frac{1}{\exp(x)} \quad \text{(nach (i))}$$

folgt für $1 - x > 0$

$$\exp x \leq \frac{1}{1 - x}.$$

Der Beweis von (iii) und (iv) soll in den Übungen erbracht werden. □

8 Funktionen

Was eine Abbildung

$$\begin{array}{rccc} f: & M & \longrightarrow & N \\ & x & \longmapsto & f(x) \end{array}$$

ist, wissen Sie schon aus der Vorlesung Lineare Algebra I, nämlich eine Zuordnung, die jedem Element x der Menge M genau ein Element $f(x)$ der Menge N als Bild zuordnet.

Insbesondere betrachten wir Abbildungen $f : D \to \mathbb{R}$, wobei $D \subseteq \mathbb{R}$ ist. Solche Abbildungen nennen wir auch (*reelle*) *Funktionen*. Die Menge D heißt *Definitionsbereich* von f.

Etwas ungenau, aber insbesondere in der Physik üblich, ist die Schreibweise $f(x)$ für die Funktion f ($f(x)$ ist ja nur der Funktionswert $f(x) \in \mathbb{R}$). Wir interpretieren stets $f(x)$ richtig, wenn im Zusammenhang von Funktionen die Rede ist. Dann fasst man x in $f(x)$ als Leerstelle auf, in die etwas eingesetzt werden kann.

Wir betrachten nun Beispiele für Funktionen.

Polynome

Ein *Polynom* mit reellen Koeffizienten ist ein Ausdruck

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

mit $a_i \in \mathbb{R}$. Der Buchstabe x ist hier nur ein Symbol, eine Unbestimmte. Die a_i bezeichnet man als die *Koeffizienten* des Polynoms. Die Menge aller Polynome mit reellen Koeffizienten bezeichnet mam mit $\mathbb{R}[x]$. Ist $a_n \neq 0$, so heißt n der *Grad* des Polynoms $a_0 + a_1 x + \cdot + a_n x^n$.

Polynome kann man addieren und multiplizieren:

$$\begin{aligned}
&(a_0 + a_1 x + \cdots + a_n x^n) + (b_0 + b_1 x + \cdots + b_m x^m) \\
&\quad = \quad (a_0 + b_0) + (a_1 + b_1)x + \cdots, \\
&(a_0 + a_1 x + \cdots + a_n x^n) \cdot (b_0 + b_1 x + \cdots + b_m x^m) \\
&\quad = \quad a_0 b_0 + (a_1 b_0 + a_0 b_1)x + (a_2 b_0 + a_1 b_1 + a_0 b_2)x^2 + \cdots \\
&\quad = \quad c_0 + c_1 x + \cdots + c_{n+m} x^{n+m},
\end{aligned}$$

wobei $a_i = 0$ für $i > n$ und $b_j = 0$ für $j > m$ gesetzt wird und

$$c_k := \sum_{i=0}^{k} a_i b_{k-i} \quad \text{für } k = 0, 1, \ldots, n+m.$$

Jedes Polynom

$$a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{R}[x]$$

liefert nun eine Funktion $f : \mathbb{R} \to \mathbb{R}$ durch

$$\alpha \in \mathbb{R} \mapsto f(\alpha) = a_0 + a_1 \alpha + \cdots + a_n \alpha^n \in \mathbb{R}.$$

Diese Funktion nennt man auch eine *ganzrationale Funktion*. Abkürzend schreibt man für das Polynom $a_0 + a_1 x + \cdots + a_n x^n$ auch $f(x)$. Zunächst ist diese Schreibweise gefährlich, weil es um verschiedene Objekte geht. Es könnte ja sein, dass verschiedene Polynome gleiche Funktionen definieren. Wir werden aber gleich zeigen, dass dies nicht der Fall ist.

Definition Eine Zahl $\alpha \in \mathbb{R}$ heißt *Nullstelle* des Polynoms $f(x)$ genau dann, wenn $f(\alpha) = 0$.

Satz 8.1 *Ist $\alpha \in \mathbb{R}$ Nullstelle des Polynoms $f(x)$, dann gibt es ein $g(x) \in \mathbb{R}[x]$, so dass*

$$f(x) = (x - \alpha) \cdot g(x).$$

Beweis. Durch Ausmultiplizieren sieht man

$$x^p - \alpha^p = (x - \alpha)(x^{p-1} + \alpha x^{p-2} + \alpha^2 x^{p-3} + \cdots + \alpha^{p-2} x + \alpha^{p-1}).$$

Es sei $f(\alpha) = 0$. Dann gilt

$$\begin{aligned} f(x) &= a_0 + a_1 x + \cdots + a_n x^n - (a_0 + a_1\alpha + \cdots + a_n\alpha^n) \\ &= a_1(x - \alpha) + a_2(x^2 - \alpha^2) + \cdots + a_n(x^n - \alpha^n). \end{aligned}$$

□

Satz 8.2 *Ein Polynom vom Grad n hat höchstens n verschiedene Nullstellen.*

Beweis. Es sei $f(x)$ ein Polynom vom Grad $n > 0$. Hat $f(x)$ keine Nullstelle, so ist man fertig. Ist für $\alpha \in \mathbb{R}$ $f(\alpha) = 0$, so kann man nach Satz 8.1 $f(x) = (x - \alpha)g(x)$ für ein Polynom $g(x)$ schreiben. Der Grad von $g(x)$ ist aber $n - 1$. Also kann man den Satz durch Induktion beweisen; ein Polynom vom Grad 0 hat keine Nullstellen. □

Satz 8.3 *Es seien $f_1(x), f_2(x)$ zwei Polynome, die dieselbe Funktion von $\mathbb{R}$ nach $\mathbb{R}$ liefern. Dann sind $f_1(x)$ und $f_2(x)$ als Polynome gleich, d.h. alle Koeffizienten sind gleich.*

Beweis. Wir betrachten das Polynom $f_1(x) - f_2(x)$. Wenn die Polynome $f_1(x)$ und $f_2(x)$ als Polynome verschieden wären, dann wäre dieses Polynom nicht das Nullpolynom, hätte also einen Grad $n \geq 0$. Nach Satz 8.2 hat $f_1(x) - f_2(x)$ höchstens n Nullstellen. Nach Voraussetzung hat $f_1(x) - f_2(x)$ aber jedes $\alpha \in \mathbb{R}$ als Nullstelle. Dies sind unendlich viele. Also gilt $f_1(x) = f_2(x)$. □

Man veranschaulicht sich eine Funktion durch einen Graphen. Es sei $f : D \to \mathbb{R}$ eine Funktion. Dann heißt die Menge

$$\{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \in D, y = f(x)\}$$

der *Graph* der Funktion f. Der Graph der Funktion $f(x) = x^3 - x$ ist in Abb. 2 angegeben.

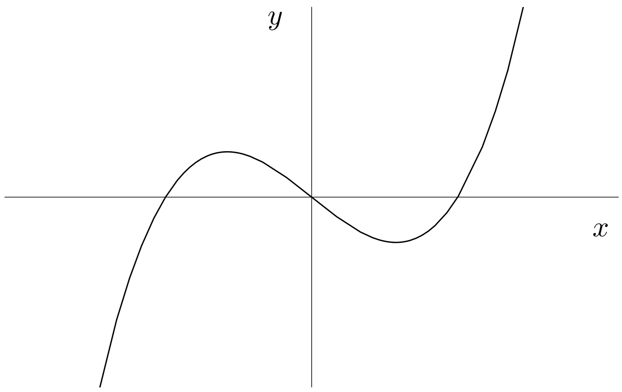


Abbildung 2: Graph von $f(x) = x^3 - x$

Rationale Funktionen

Sind g und h ganzrationale Funktionen und ist $M = \{\alpha \in \mathbb{R} \mid h(\alpha) = 0\}$ die Menge der Nullstellen von h, so wird durch

$$f(x) = \frac{g(x)}{h(x)}$$

eine Funktion $f : \mathbb{R} \setminus M \to \mathbb{R}$ erklärt. Funktionen dieses Typs heißen *rationale Funktionen*.

Eine rationale Funktion kann auf ihrer Definitionsmenge mit einer ganzrationalen Funktion übereinstimmen: Z.B. gilt für $x \in \mathbb{R} \setminus \{-1, 1\}$

$$\frac{1 - x^4}{1 - x^2} = 1 + x^2.$$

Aber: Die Funktionen

$$\begin{array}{rcl} x & \mapsto & 1 + x^2, \quad x \in \mathbb{R}, \\ x & \mapsto & \dfrac{1 - x^4}{1 - x^2}, \quad x \in \mathbb{R} \setminus \{-1, 1\}, \end{array}$$

sind verschieden, weil ihre Definitionsbereiche verschieden sind! Aber die Funktionen

$$\begin{array}{rcl} x & \mapsto & 1 - x^2, \quad x \in \mathbb{R}, \\ x & \mapsto & \dfrac{1 - x^4}{1 + x^2}, \quad x \in \mathbb{R}, \end{array}$$

sind gleich!

Exponentialfunktion

Die Exponentialreihe definiert eine Funktion

$$\begin{array}{rccl} \exp: & \mathbb{R} & \longrightarrow & \mathbb{R}_+^* := \{r \in \mathbb{R} \mid r > 0\} \\ & x & \longmapsto & \exp(x) \end{array}.$$

Diese Funktion heißt die *Exponentialfunktion.* Die Exponentialfunktion ist streng monoton wachsend.

Definition Eine Funktion $f : D \to \mathbb{R}$ heißt *monoton wachsend*, wenn für alle $x_1, x_2 \in D$ gilt:

$$x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2).$$

Die Funktion f heißt *streng monoton wachsend*, wenn für alle $x_1, x_2 \in D$ gilt:

$$x_1 < x_2 \Rightarrow f(x_1) < f(x_2).$$

Entsprechend werden die Eigenschaften *monoton fallend* und *streng monoton fallend* definiert. Ist f entweder streng monoton wachsend oder streng monoton fallend, so heißt f *streng monoton.*

Man beachte: konstante Funktionen sind sowohl monoton wachsend als auch monoton fallend.

Nach Korollar 7.2 ist exp streng monoton wachsend.

Satz 8.4 *Ist $f : D \to \mathbb{R}$ streng monoton, so ist f injektiv, es existiert also eine Umkehrfunktion $f^{-1} : f(D) \to D$ und diese ist ebenfalls streng monoton.*

Warnung Die Umkehrfunktion f^{-1} ist nicht zu verwechseln mit der Funktion $\frac{1}{f}$!

Beweis. Es sei f etwa streng monoton wachsend. Dann gilt

$$\begin{array}{cl} & x_1 \neq x_2 \\ \Rightarrow & \text{entweder } x_1 < x_2 \text{ oder } x_2 < x_1 \\ \Rightarrow & \text{entweder } f(x_1) < f(x_2) \text{ oder } f(x_2) < f(x_1) \\ \Rightarrow & f(x_1) \neq f(x_2). \end{array}$$

Also ist f injektiv und es existiert die Umkehrfunktion $f^{-1} : f(D) \to D$.

Die Funktion f^{-1} ist streng monoton wachsend: Andernfalls gäbe es y_1, y_2 mit $y_1 < y_2$ und $f^{-1}(y_1) \geq f^{-1}(y_2)$. Setze

$$x_1 := f^{-1}(y_1), \quad x_2 := f^{-1}(y_2).$$

Dann gilt

$$x_1 \geq x_2 \text{ und } y_1 = f(x_1) < f(x_2) = y_2.$$

Das widerspricht der Voraussetzung, dass f streng monoton wachsend ist.

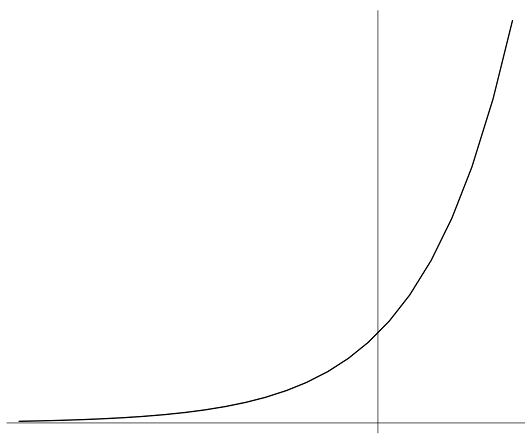
Für streng monoton fallende Funktionen verläuft der Beweis analog. $\square$

Bemerkung 8.1 Die strenge Monotonie ist eine hinreichende, aber keine notwendige Bedingung für die Umkehrbarkeit einer reellen Funktion. Z.B. bildet $f : [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x & \text{für } x \in \mathbb{Q}, \\ 1-x & \text{für } x \notin \mathbb{Q}, \end{cases}$$

das Intervall $[0,1]$ bijektiv auf $[0,1]$ ab, ist aber nicht (streng) monoton.

Nach Satz 8.4 ist die Exponentialfunktion also injektiv. Außerdem wissen wir nach Korollar 7.1, dass $\exp(n) = e^n$ für alle ganzen Zahlen n gilt. Die Folge $(e^n)_{n\in\mathbb{N}}$ nimmt beliebig große Werte an, die Folge $(e^{-n})_{n\in\mathbb{N}}$ wiederum nimmt beliebig kleine Werte an. Der Graph der Exponentialfunktion sieht also folgendermaßen aus:



Aus der Graphik ist anschaulich klar, dass $\exp : \mathbb{R} \to \mathbb{R}_+^*$ auch surjektiv ist. Das ist aber kein Beweis. Den Beweis werden wir im Zusammenhang mit der Untersuchung stetiger Funktionen erbringen.

9 Logarithmus und allgemeine Potenz

Definition Die Umkehrfunktion von $\exp : \mathbb{R} \to \mathbb{R}_+^*$

$$\ln : \mathbb{R}_+^* \to \mathbb{R}$$

(manchmal auch log genannt) heißt der *natürliche Logarithmus*.

Satz 9.1 (Funktionalgleichung des Logarithmus) *Für alle* $x, y \in \mathbb{R}_+^*$ *gilt*

$$\begin{aligned} \ln(xy) &= \ln x + \ln y, \\ \ln(\frac{x}{y}) &= \ln x - \ln y. \end{aligned}$$

Beweis. Diese Gleichungen folgen daraus, dass exp ein Gruppenhomomorphismus von $(\mathbb{R}_+^*, \cdot)$ nach $(\mathbb{R}, +)$ ist. Denn dann ist die Umkehrabbildung ebenfalls ein Gruppenhomomorphismus. Die Gleichungen lassen sich aber auch leicht direkt ableiten: Wir setzen $\xi := \ln(x)$ und $\eta := \ln(y)$. Dann gilt nach Definition $\exp(\xi) = x$ und $\exp(\eta) = y$. Aus der Funktionalgleichung von exp folgt

$$\exp(\xi + \eta) = \exp(\xi)\exp(\eta) = xy.$$

Daraus folgt

$$\ln(xy) = \xi + \eta = \ln x + \ln y.$$

Die andere Gleichung folgt hieraus. □

Definition Für $a \in \mathbb{R}$, $a > 0$, und $x \in \mathbb{R}$ definieren wir die *allgemeine Potenz zur Basis a* durch

$$a^x := \exp(x \ln a).$$

Satz 9.2 (i) *Für alle* $x, y \in \mathbb{R}$ *gilt*

$$a^{x+y} = a^x a^y.$$

(ii) *Für* $n \in \mathbb{N}$ *stimmt die Definition von* a^n *mit der alten Definition* $a^n = a \cdots a$ *(n Faktoren) überein.*

(iii) *Für alle* $p \in \mathbb{Z}$ *und* $q \in \mathbb{N}$ *mit* $q \geq 2$ *gilt*

$$a^{\frac{p}{q}} = \sqrt[q]{a^p}.$$

Beweis. (i) folgt unmittelbar aus der Funktionalgleichung der Exponentialfunktion.

Zu (ii): Durch vollständige Induktion zeigt man

$$\exp(nx) = (\exp(x))^n$$

für alle $n \in \mathbb{N}$ und $x \in \mathbb{R}$. Mit $x = \ln a$ folgt (ii).

Zu (iii): Die Behauptung (iii) folgt aus

$$a^p = \exp(p \ln a) = \exp(q \cdot \frac{p}{q} \ln a) = \left(\exp(\frac{p}{q} \ln a)\right)^q = \left(a^{\frac{p}{q}}\right)^q.$$

□

Satz 9.3 *Für alle* $a, b \in \mathbb{R}_+^*$ *und* $x, y \in \mathbb{R}$ *gilt:*

(i) $(a^x)^y = a^{xy}$.

(ii) $a^x b^x = (ab)^x$.

(iii) $\left(\frac{1}{a}\right)^x = a^{-x}$.

Beweis. Zu (i): Wegen $a^x = \exp(x \ln a)$ ist $\ln(a^x) = x \ln a$, also

$$(a^x)^y = \exp(y \ln(a^x)) = \exp(yx \ln a) = a^{xy}.$$

Die Behauptungen (ii) und (iii) sind ebenso einfach zu beweisen. □

Die Funktion $x \mapsto a^x$ $(x \in \mathbb{R})$ ist für $a > 1$ streng monoton wachsend und für $0 < a < 1$ streng monoton fallend. In beiden Fällen wird $\mathbb{R}$ bijektiv auf $\mathbb{R}_+^*$ abgebildet. Falls $a > 0$, $a \neq 1$, existiert also die Umkehrfunktion von $x \mapsto a^x$, sie wird *Logarithmus zur Basis a* genannt und mit

$$\log_a : \mathbb{R}_+^* \to \mathbb{R}$$

bezeichnet. Mit dem natürlichen Logarithmus besteht der Zusammenhang

$$\log_a x = \frac{\ln x}{\ln a},$$

denn aus $x = a^y = \exp(y \cdot \ln a)$ folgt

$$\ln x = y \ln a = \log_a x \cdot \ln a.$$

10 Komplexe Zahlen, Trigonometrische Funktionen

Mit Hilfe der Exponentialfunktion kann man die folgenden Funktionen definieren:

Definition Die *Hyperbelfunktionen* sind definiert als:

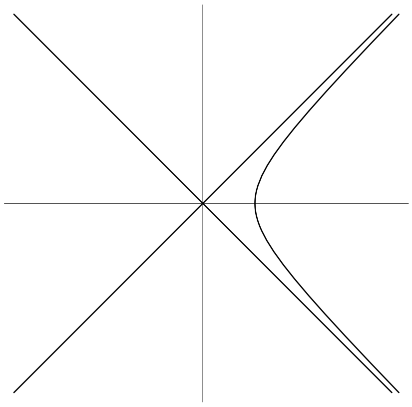
$$\begin{aligned} \cosh x &:= \frac{1}{2}(e^x + e^{-x}) \quad \textit{cosinus hyperbolicus} \\ \sinh x &:= \frac{1}{2}(e^x - e^{-x}) \quad \textit{sinus hyperbolicus} \end{aligned}$$

Warum heißen diese Funktionen Hyperbelfunktionen? Wir ordnen einem $x \in \mathbb{R}$ das Paar $(\xi, \eta) = (\cosh x, \sinh x) \in \mathbb{R}^2$ zu. Wenn nun x die reelle Zahlengerade durchläuft, so durchlaufen die Bildpunkte (ξ, η) eine Kurve.

Wie sieht diese Kurve aus? Dazu bemerken wir, dass für $\xi = \cosh x$ und $\eta = \sinh x$ gilt:

$$\begin{aligned} \xi + \eta &= e^x \quad > 0, \\ \xi - \eta &= e^{-x} \quad > 0, \\ \xi^2 - \eta^2 &= 1. \end{aligned}$$

Also haben wir es mit einer Hyperbel zu tun. Die Ebene $\mathbb{R}^2$ wird durch die Geraden $\xi + \eta = 0$ und $\xi - \eta = 0$ in 4 Quadranten eingeteilt, wenn auch nicht in die üblichen. Wegen $\xi + \eta > 0$ und $\xi - \eta > 0$ für Punkte auf der Kurve liegt die Kurve ganz im rechten Quadranten.



Aus den Reihendarstellungen

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!}, \\ e^{-x} &= \sum_{n=0}^{\infty} \frac{(-x)^n}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^n}{n!}, \end{aligned}$$

erhält man

$$\cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots .$$

Entsprechend erhält man

$$\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots .$$

Diese Reihen konvergieren absolut für jedes $x \in \mathbb{R}$.

Nachdem wir die Hyperbel $\xi^2 - \eta^2 = 1$ parametrisiert haben, soll nun auch der Kreis parametrisiert werden. Der Kreis ist durch die Gleichung $\xi^2 + \eta^2 = 1$ gegeben. Nun wenden wir einen Kunstgriff an: Von der Hyperbel kommt man zum Kreis, wenn man η durch $\frac{1}{i}\eta$ ersetzt, wobei $i^2 = -1$ ist. Aus $(\cosh x)^2 - (\sinh x)^2 = 1$ folgt dann, dass für jedes x auch

$$(\cosh x)^2 + (\frac{1}{i}\sinh x)^2 = 1$$

gilt. Um die Sache nun noch schlimmer zu machen, ersetzen wir auch noch x durch ix:

$$(\cosh(ix))^2 + (\frac{1}{i}\sinh(ix))^2 = 1.$$

Nun gilt erstaunlicherweise

$$\begin{aligned} \cosh(ix) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \quad \in \mathbb{R}, \\ \frac{1}{i}\sinh(ix) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \quad \in \mathbb{R}. \end{aligned}$$

Wir werden definieren

Definition Für jedes $x \in \mathbb{R}$ sei

$$\begin{aligned} \cos(x) &:= \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots, \\ \sin(x) &:= \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots. \end{aligned}$$

Unsere Argumentation wollen wir nun auf sichere Füße stellen. Dazu führen wir zunächst die komplexen Zahlen ein.

Die Menge $\mathbb{R}\times\mathbb{R}$ aller (geordneten) Paare reeller Zahlen bildet zusammen mit der Addition und Multiplikation

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2, y_1 + y_2) \\ (x_1, y_1) \cdot (x_2, y_2) &:= (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2) \end{aligned}$$

einen Körper. Dieser Körper heißt der Körper $\mathbb{C}$ der *komplexen Zahlen.* Das Nullelement ist $(0, 0)$, das Einselement ist $(1, 0)$.

Da

$$\begin{aligned} (x_1, 0) + (x_2, 0) &= (x_1 + x_2, 0) \text{ und} \\ (x_1, 0) \cdot (x_2, 0) &= (x_1 x_2, 0) \end{aligned}$$

gilt, kann man $x \in \mathbb{R}$ mit $(x, 0) \in \mathbb{C}$ identifizieren:

$$\begin{aligned} \mathbb{R} &\hookrightarrow \mathbb{C} \\ x &\mapsto (x, 0) \end{aligned}$$

Die Menge $\mathbb{R}$ wird so eine Teilmenge von $\mathbb{C}$. Setzt man noch $i := (0,1)$, so erhält man die gebräuchliche Schreibweise für komplexe Zahlen:

$$(x,y) = (x,0) + (0,1)\cdot(y,0) = x + iy \quad (x,y \in \mathbb{R}).$$

Es gilt

$$i^2 = (0,1)\cdot(0,1) = (-1,0) = -1.$$

Für eine komplexe Zahl $z = x + iy$ $(x,y \in \mathbb{R})$ werden *Realteil* und *Imaginärteil* wie folgt definiert:

$$\mathrm{Re}(z) := x, \qquad \mathrm{Im}(z) := y.$$

Zwei komplexe Zahlen z und z' sind also genau dann gleich, wenn $\mathrm{Re}(z) = \mathrm{Re}(z')$ und $\mathrm{Im}(z) = \mathrm{Im}(z')$ gilt.

Es sei $z = x + iy \in \mathbb{C}$ $(x,y \in \mathbb{R})$. Dann heißt $\bar{z} := x - iy$ die zu z *konjugiert komplexe Zahl*. In der Gaußschen Zahlenebene entsteht $\bar{z}$ aus z durch Spiegelung an der reellen Achse. Es gilt

$$z = \bar{z} \Leftrightarrow z \in \mathbb{R}.$$

Aus der Definition folgt

$$\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z}), \quad \mathrm{Im}(z) = \frac{1}{2i}(z - \bar{z}).$$

Einfach nachzurechnen sind folgende Rechenregeln für die Konjugation:

Satz 10.1 *Für alle $z, z_1, z_2 \in \mathbb{C}$ gilt*

(a) $\bar{\bar{z}} = z$.

(b) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$.

(c) $\overline{z_1 z_2} = \overline{z_1}\cdot\overline{z_2}$.

Es sei $z = x + iy \in \mathbb{C}$. Dann ist

$$z\bar{z} = (x + iy)(x - iy) = x^2 + y^2$$

eine nichtnegative reelle Zahl. Man setzt

$$|z| := \sqrt{z\bar{z}}.$$

Die Zahl $|z|$ heißt der *Betrag* von z. Da $|z| = \sqrt{x^2 + y^2}$, ist der Betrag von z gleich dem Abstand des Punktes z vom Nullpunkt der Gaußschen Zahlenebene.

Für $z \in \mathbb{R}$ stimmt der Betrag mit dem Absolutbetrag für reelle Zahlen überein. Für alle $z \in \mathbb{C}$ gilt $|z| = |\bar{z}|$.

Satz 10.2 *Für alle $z, z_1, z_2 \in \mathbb{C}$ gilt*

(a) $|z| \geq 0$; $|z| = 0 \Leftrightarrow z = 0$.

(b) $|z_1 + z_2| \leq |z_1| + |z_2|$ *(Dreiecksungleichung).*

(c) $|z_1 z_2| = |z_1||z_2|$.

Beweis. (a) ist klar.
Zu (c): Nach Definition des Betrags ist

$$|z_1 z_2|^2 = (z_1 z_2)(\overline{z_1 z_2}) = z_1 z_2 \overline{z_1}\,\overline{z_2} = (z_1 \overline{z_1})(z_2 \overline{z_2}) = |z_1|^2 |z_2|^2.$$

Wurzelziehen liefert die Behauptung.
Zu (b): Für jede komplexe Zahl z gilt

$$\mathrm{Re}(z) \leq |z|.$$

Daraus folgt

$$\mathrm{Re}(z_1 \overline{z_2}) \leq |z_1 \overline{z_2}| = |z_1||\overline{z_2}| = |z_1||z_2|.$$

Also gilt

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1} + \overline{z_2}) \\ &= z_1 \overline{z_1} + z_1 \overline{z_2} + z_2 \overline{z_1} + z_2 \overline{z_2} \\ &= |z_1|^2 + 2\mathrm{Re}(z_1 \overline{z_2}) + |z_2|^2 \\ &\leq |z_1|^2 + 2|z_1||z_2| + |z_2|^2 \\ &= (|z_1| + |z_2|)^2. \end{aligned}$$

Wurzelziehen liefert wieder die Behauptung. □

Bemerkung 10.1 Ein Körper, für dessen Elemente z ein Betrag $|z| \in \mathbb{R}$ definiert ist, so dass die drei Eigenschaften aus dem vorherigen Satz gelten, heißt ein *bewerteter Körper.* Die Körper $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ sind bewertete Körper.

Wir können nun auch Folgen und Reihen komplexer Zahlen betrachten. Die Konvergenzbegriffe für Folgen und Reihen übertragen sich wörtlich. Ohne Beweis geben wir den folgenden Satz an. Der Beweis ist Routine.

Satz 10.3 *Es sei $(c_n)_{n\in\mathbb{N}}$ eine Folge komplexer Zahlen. Die Folge konvergiert genau dann, wenn die beiden Folgen $(\mathrm{Re}(c_n))_{n\in\mathbb{N}}$ und $(\mathrm{Im}(c_n))_{n\in\mathbb{N}}$ konvergieren. Im Falle der Konvergenz gilt*

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} \mathrm{Re}(c_n) + i \lim_{n\to\infty} \mathrm{Im}(c_n).$$

Korollar 10.1 *Es sei $(c_n)_{n\in\mathbb{N}}$ eine Folge komplexer Zahlen mit* $\lim_{n\to\infty} c_n = c$*. Dann konvergiert auch die Folge $(\overline{c_n})_{n\in\mathbb{N}}$ und es gilt*

$$\lim_{n\to\infty} \overline{c_n} = \bar{c}.$$

Beweis. Dies folgt daraus, dass $\mathrm{Re}(\overline{c_n}) = \mathrm{Re}(c_n)$ und $\mathrm{Im}(\overline{c_n}) = -\mathrm{Im}(c_n)$. □

Wir können nun für jedes $z \in \mathbb{C}$ die Exponentialreihe

$$\exp z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

betrachten. Genau wie für reelle z zeigt man, dass diese Reihe auch für komplexe Zahlen z absolut konvergent ist. Die Abschätzung des Restglieds (Satz 7.2) und die Funktionalgleichung der Exponentialfunktion (Satz 7.4) gelten auch für die komplexe Exponentialfunktion. Die Beweise übertragen sich wörtlich. Ferner gilt $\exp(z) \neq 0$ für alle $z \in \mathbb{C}$. Auch dies wird genau wie für $z \in \mathbb{R}$ bewiesen.

Satz 10.4 *Für jedes $z \in \mathbb{C}$ gilt*

$$\exp(\bar{z}) = \overline{\exp(z)}.$$

Beweis. Es sei

$$s_n(z) := \sum_{k=0}^{n} \frac{z^k}{k!}, \quad \tilde{s}_n(z) := \sum_{k=0}^{n} \frac{\bar{z}^k}{k!}.$$

Nach den Rechenregeln für die Konjugation gilt für alle $n \in \mathbb{N}$

$$\tilde{s}_n(z) = \sum_{k=0}^{n} \frac{\bar{z}^k}{k!} = \sum_{k=0}^{n} \left(\overline{\frac{z^k}{k!}}\right) = \left(\overline{\sum_{k=0}^{n} \frac{z^k}{k!}}\right) = \overline{s_n(z)}.$$

Aus dem vorhergehenden Korollar folgt daher

$$\begin{aligned} \exp(\bar{z}) &= \lim_{n\to\infty} \tilde{s}_n(z) = \lim_{n\to\infty} \overline{s_n(z)} \\ &= \overline{\lim_{n\to\infty} s_n(z)} = \overline{\exp(z)}. \end{aligned}$$

□

Wir betrachten nun die Funktion $x \mapsto \exp(ix)$ $(x \in \mathbb{R})$. Was ist die Wertemenge dieser Funktion? Nach dem vorhergehenden Satz gilt

$$\begin{aligned} |\exp(ix)|^2 &= \exp(ix)\overline{\exp(ix)} = \exp(ix)\exp(\overline{ix}) \\ &= \exp(ix)\exp(-ix) = \exp(0) = 1. \end{aligned}$$

Die Menge

$$S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$$

heißt auch *Einheitskreis*. Mit $z_1, z_2 \in S^1$ ist auch ihr Produkt $z_1 \cdot z_2 \in S^1$. Außerdem gilt $1 = (1, 0) \in S^1$ und mit $z \in S^1$ ist auch $\frac{1}{z} \in S^1$. Deshalb bildet $(S^1, \cdot)$ eine Gruppe.

Nach der Funktionalgleichung der Exponentialfunktion ist die Abbildung $x \mapsto \exp(ix)$ $(x \in \mathbb{R})$ ein Homomorphismus der Gruppe $(\mathbb{R}, +)$ in die Gruppe $(S^1, \cdot)$.

Für $\exp(ix)$ schreiben wir auch e^{ix}.

Definition Für $x \in \mathbb{R}$ sei

$$\begin{aligned} \cos x &:= \operatorname{Re}(e^{ix}) \\ \sin x &:= \operatorname{Im}(e^{ix}) \end{aligned}$$

Wegen

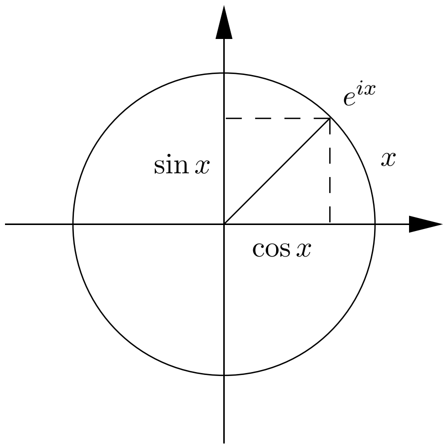
$$\begin{aligned} \operatorname{Re}(e^{ix}) &= \frac{1}{2}(e^{ix} + e^{-ix}) = \cosh(ix), \\ \operatorname{Im}(e^{ix}) &= \frac{1}{2i}(e^{ix} - e^{-ix}) = \frac{1}{i}\sinh(ix), \end{aligned}$$

stimmt diese Definition mit der alten überein.

Es ist also

$$e^{ix} = \cos x + i \sin x \quad (\textit{Eulersche Formel}).$$

Dies können wir geometrisch so deuten: e^{ix} ist ein Punkt des Einheitskreises und $\cos x$ bzw. $\sin x$ sind die Projektionen dieses Punktes auf die reelle bzw. imaginäre Achse.



Das Argument x kann man als die orientierte Länge des Bogens e^{it}, $0 \le t \le x$ (bzw. $0 \ge t \ge x$, falls x negativ ist), deuten.

Satz 10.5 *Für alle $x \in \mathbb{R}$ gilt:*

(a) $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$, $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$.

(b) $\cos(-x) = \cos x$, $\sin(-x) = -\sin x$.

(c) $\cos^2 x + \sin^2 x = 1$.

Beweis. Dies folgt unmittelbar aus der Definition. □

Satz 10.6 (Additionstheoreme) *Für alle $x, y \in \mathbb{R}$ gilt*

(a) $\cos(x + y) = \cos x \cos y - \sin x \sin y$.

(b) $\sin(x + y) = \sin x \cos y + \cos x \sin y$.

Beweis. Aus der Eulerschen Formel und der Funktionalgleichung der Exponentialfunktion ergibt sich

$$\begin{aligned} &\cos(x+y) + i\sin(x+y) \\ &= e^{i(x+y)} = e^{ix+iy} = e^{ix}e^{iy} \\ &= (\cos x + i\sin x)(\cos y + i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) + i(\sin x \cos y + \cos x \sin y). \end{aligned}$$

Vergleicht man Real- und Imaginärteil, so erhält man die Behauptung. □

Korollar 10.2 *Für alle $x, y \in \mathbb{R}$ gilt*

(a) $\sin x - \sin y = 2\cos\frac{x+y}{2}\sin\frac{x-y}{2}$.

(b) $\cos x - \cos y = -2\sin\frac{x+y}{2}\sin\frac{x-y}{2}$.

Beweis. Wir setzen

$$u := \frac{x+y}{2}, \quad v := \frac{x-y}{2}.$$

Dann gilt $x = u + v$ und $y = u - v$. Aus dem vorhergehenden Satz folgt

$$\begin{aligned} \sin x - \sin y &= \sin(u+v) - \sin(u-v) \\ &= (\sin u \cos v + \cos u \sin v) - (\sin u \cos(-v) + \cos u \sin(-v)) \\ &= 2\cos u \sin v \\ &= 2\cos\frac{x+y}{2}\sin\frac{x-y}{2}. \end{aligned}$$

Die zweite Gleichung wird analog bewiesen. □

Satz 10.7 (a) *Für alle $x \in \mathbb{R}$ gilt $|\sin x| \le |x|$.*

(b) *Für alle $x, y \in \mathbb{R}$ gilt*

$$\begin{aligned} |\cos y - \cos x| &\le |y - x|, \\ |\sin y - \sin x| &\le |y - x|. \end{aligned}$$

Beweis. Zu (a): Aus der Reihendarstellung von $\sin x$ folgt

$$|\sin x| \leq |x|$$

für $0 \leq x \leq 1$. Wegen $\sin(-x) = -\sin x$ folgt diese Ungleichung für $|x| \leq 1$. Für $|x| > 1$ ist die Ungleichung aber wegen $|\sin x| \leq |e^{ix}| = 1$ trivialerweise erfüllt.

Zu (b): Wegen $|\sin x| \leq 1$ für alle $x \in \mathbb{R}$ und (a) folgt aus Korollar 10.2

$$|\cos y - \cos x| = 2 \left|\sin \frac{x+y}{2}\right| \left|\sin \frac{y-x}{2}\right| \leq 2 \left|\frac{y-x}{2}\right| = |y-x|$$

für alle $x, y \in \mathbb{R}$. Analog wird die zweite Ungleichung bewiesen. □

Satz 10.8 *Die Funktion* cos *hat im Intervall* $[0, 2]$ *genau eine Nullstelle.*

Definition Diese Nullstelle bezeichnet man mit $\frac{\pi}{2}$.

Wir werden diesen Satz erst später vollständig beweisen. Zum Beweis brauchen wir aber drei Hilfssätze, die wir jetzt schon beweisen.

Lemma 10.1 $\cos 2 \leq -\frac{1}{3}$.

Beweis. Aus der Reihendarstellung von cos folgt:

$$\cos 2 = 1 - \frac{4}{2!} + \frac{16}{4!} - r = -\frac{1}{3} - r \quad \text{mit } r > 0.$$

□

Lemma 10.2 $\sin x > 0$ *für alle* $x \in (0, 2]$.

Beweis. Die Reihendarstellung von sin lautet:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \cdots .$$

Für $0 < x \leq 2$ handelt es sich bei der rechten Seite um eine alternierende Reihe mit monoton gegen 0 fallenden Reihengliedern. Die Summe der ersten beiden Glieder ist positiv, die jeweiligen Fehler sind nicht negativ (vgl. den Beweis zum Leibniz-Kriterium). Also ist $\sin x > 0$ für $0 < x \leq 2$. □

Lemma 10.3 *Die Funktion* cos *ist im Intervall* $[0, 2]$ *streng monoton fallend.*

Beweis. Es sei $0 \leq x_1 < x_2 \leq 2$. Dann folgt aus Lemma 10.2 und Korollar 10.2

$$\cos x_2 - \cos x_1 = -2 \sin \frac{x_1 + x_2}{2} \sin \frac{x_2 - x_1}{2} < 0.$$

□

Korollar 10.3 (Spezielle Werte der Exponentialfunktion)

$$e^{i\frac{\pi}{2}} = i, \quad e^{i\pi} = -1, \quad e^{i\frac{3\pi}{2}} = -i, \quad e^{2\pi i} = 1.$$

Beweis. Da $\cos\frac{\pi}{2} = 0$, ist

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

Nach Lemma 10.2 ist $\sin\frac{\pi}{2} > 0$, also $\sin\frac{\pi}{2} = 1$, also

$$e^{i\frac{\pi}{2}} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i.$$

Die restlichen Behauptungen folgen wegen

$$e^{i\frac{n\pi}{2}} = i^n.$$

□

Aus diesem Korollar ergibt sich folgende Wertetabelle für sin und cos:

x	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\sin x$	0	1	0	-1	0
$\cos x$	1	0	-1	0	1

Korollar 10.4 *Für alle $x \in \mathbb{R}$ gilt*

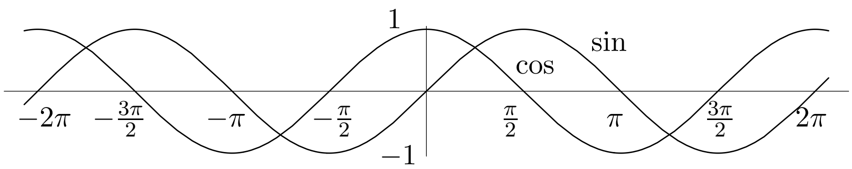
(a) $\cos(x + 2\pi) = \cos x$, $\sin(x + 2\pi) = \sin x$.

(b) $\cos(x + \pi) = -\cos x$, $\sin(x + \pi) = -\sin x$.

(c) $\cos x = \sin(\frac{\pi}{2} - x)$, $\sin x = \cos(\frac{\pi}{2} - x)$.

Beweis. Additionstheoreme und obige Wertetabelle. □

Bemerkung 10.2 Aus dem Korollar folgt, dass man nur die Funktion cos im Intervall $[0, \frac{\pi}{2}]$ zu kennen braucht, um den Gesamtverlauf der Funktionen cos und sin zu kennen.



Korollar 10.5 (Nullstellen von sin und cos) (a) $\{x \in \mathbb{R} \mid \sin x = 0\} = \{k\pi \mid k \in \mathbb{Z}\}$.

(b) $\{x \in \mathbb{R} \mid \cos x = 0\} = \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$.

Beweis. Zu (a): Es gilt $\sin k\pi = 0$ für alle $k \in \mathbb{Z}$ nach Korollar 10.4.

Wir zeigen, dass sin keine weiteren Nullstellen hat: Nach Definition von $\frac{\pi}{2}$ und wegen $\cos(-x) = \cos x$ gilt $\cos x > 0$ für $-\frac{\pi}{2} < x < \frac{\pi}{2}$. Wegen $\sin x = \cos(\frac{\pi}{2} - x)$ folgt daraus

$$\sin x > 0 \quad \text{für } 0 < x < \pi.$$

Wegen $\sin(x + \pi) = -\sin x$ gilt

$$\sin x < 0 \quad \text{für } \pi < x < 2\pi.$$

Daraus folgt, dass 0 und π die einzigen Nullstellen von sin im Intervall $[0, 2\pi)$ sind.

Es sei nun x eine Nullstelle von sin und m die größte ganze Zahl, die kleiner oder gleich $\frac{x}{2\pi}$ ist. Dann gilt

$$x = 2m\pi + \xi, \quad 0 \le \xi < 2\pi,$$

und

$$\sin \xi = \sin(x - 2m\pi) = \sin x = 0.$$

Also ist $\xi = 0$ oder $\xi = \pi$, d.h. $x = 2m\pi$ oder $x = (2m + 1)\pi$.

Zu (b): Dies folgt aus (a) wegen $\cos x = -\sin(x - \frac{\pi}{2})$. □

Korollar 10.6 *Für $x \in \mathbb{R}$ gilt:*

$$e^{ix} = 1 \Leftrightarrow x = k \cdot 2\pi, \quad k \in \mathbb{Z}.$$

Beweis. Wegen

$$\sin \frac{x}{2} = \frac{1}{2i}\left(e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}\right) = \frac{e^{-i\frac{x}{2}}}{2i}(e^{ix} - 1)$$

gilt

$$e^{ix} = 1 \Leftrightarrow \sin \frac{x}{2} = 0.$$

Die Behauptung folgt daher aus Korollar 10.5(a). □

Bemerkung 10.3 Korollar 10.6 bedeutet, dass der Kern des Gruppenhomomorphismus

$$\begin{array}{ccc} (\mathbb{R}, +) & \to & (S^1, \cdot) \\ x & \mapsto & e^{ix} \end{array}$$

die Menge $2\pi\mathbb{Z} := \{2\pi k \mid k \in \mathbb{Z}\}$ ist.

Definition Für $x \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$ setzt man

$$\tan x = \frac{\sin x}{\cos x}.$$

Definition Für $x \in \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\}$ setzt man

$$\cot x = \frac{\cos x}{\sin x}.$$

Satz 10.9 (und Definition) (a) $\cos : [0, \pi] \to [-1, 1]$ *ist bijektiv und streng monoton fallend.*
Umkehrfunktion: $\arccos : [-1, 1] \to [0, \pi]$.

(b) $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$ *ist bijektiv und streng monoton wachsend.*
Umkehrfunktion: $\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$.

(c) $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ *ist bijektiv und streng monoton wachsend.*
Umkehrfunktion: $\arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$.

(d) $\cot : (0, \pi) \to \mathbb{R}$ *ist bijektiv und streng monoton fallend.*
Umkehrfunktion: $\operatorname{arccot} : \mathbb{R} \to (0, \pi)$.

Beweis. (a) Nach Lemma 10.3 ist cos in $[0, 2]$, insbesondere in $[0, \frac{\pi}{2}]$, streng monoton fallend. Da $\cos x = -\cos(\pi - x)$, ist cos auch in $[\frac{\pi}{2}, \pi]$ fallend. Nach Satz 8.4 ist cos injektiv. Den Beweis, dass $\cos : [0, \pi] \to [-1, 1]$ auch surjektiv ist, werden wir nachtragen.

(b) Wegen $\sin x = \cos(\frac{\pi}{2} - x)$ folgt aus (a), dass sin im Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton wachsend ist.

(c) Wir zeigen zunächst, dass tan im Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$ streng monoton wachsend ist. Dazu sei $0 \leq x_1 < x_2 < \frac{\pi}{2}$. Dann gilt $\sin x_1 < \sin x_2$ und $\cos x_1 > \cos x_2 > 0$. Daraus folgt

$$\tan x_1 = \frac{\sin x_1}{\cos x_1} < \frac{\sin x_2}{\cos x_2} = \tan x_2.$$

Die Funktion tan ist also in $[0, \frac{\pi}{2})$ streng monoton wachsend. Wegen

$$\tan(-x) = -\tan x$$

ist tan auch in $(-\frac{\pi}{2}, 0]$ streng monoton wachsend. Den Beweis der Surjektivität der Funktion $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ müssen wir wieder verschieben.

(d) Wegen $\cot x = \frac{1}{\tan x}$ folgt aus (c), dass cot im Intervall $(0, \pi)$ streng monoton fallend ist. □

Satz 10.10 (Polarkoordinaten) *Jede komplexe Zahl z lässt sich schreiben als*

$$z = re^{i\varphi},$$

wobei $\varphi \in \mathbb{R}$ und $r = |z| \in \mathbb{R}_+$. Für $z \neq 0$ ist φ bis auf ein ganzzahliges Vielfaches von 2π eindeutig bestimmt.

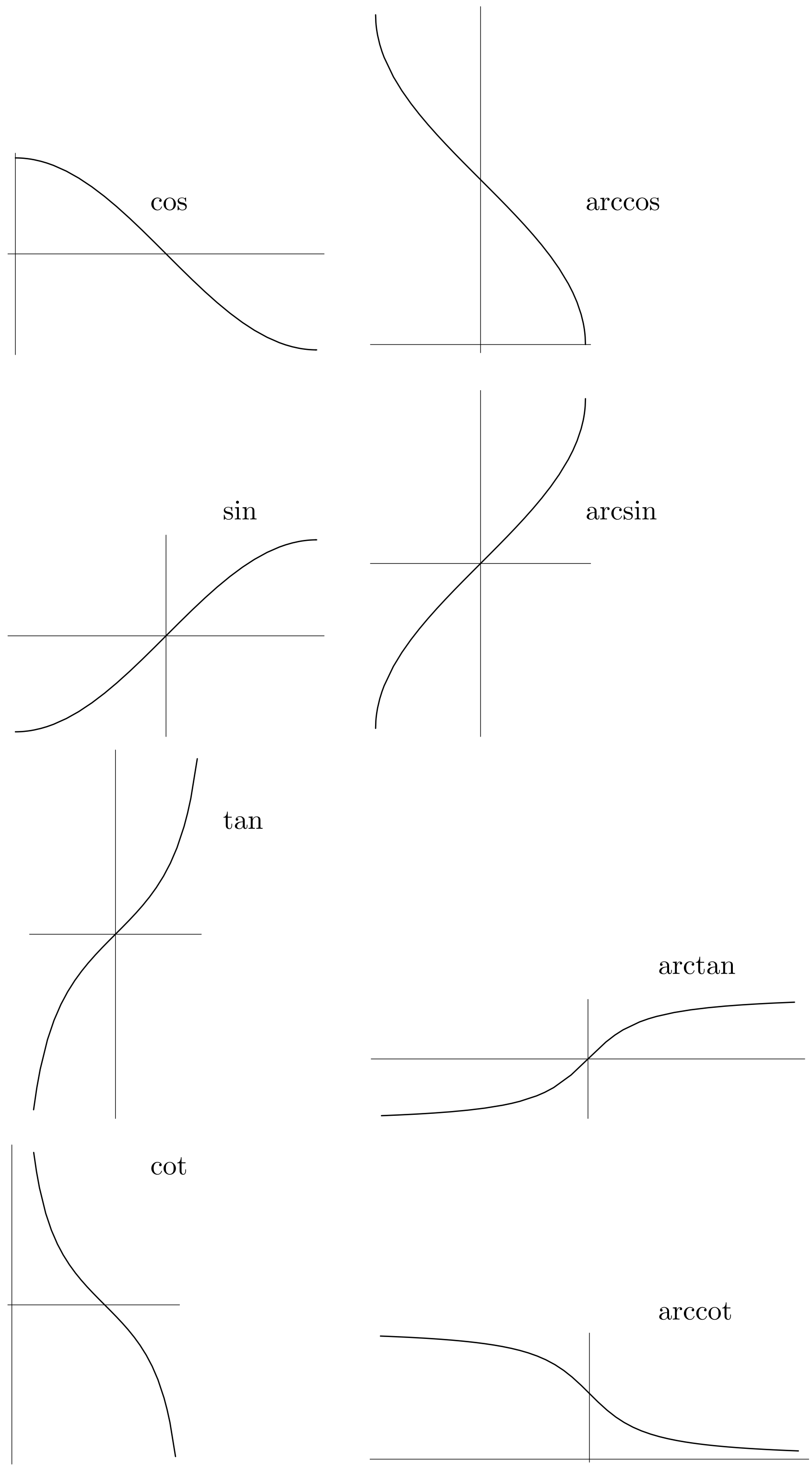
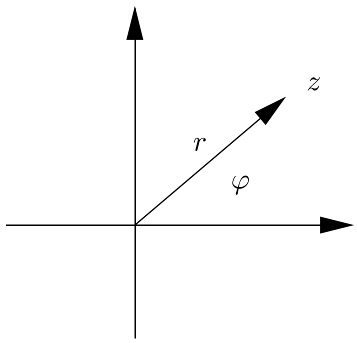


Abbildung 3: Die trigonometrischen Funktionen und ihre Umkehrfunktionen

Definition Die Zahlen r und φ nennt man die *Polarkoordinaten* der komplexen Zahl $z \neq 0$. Die Zahl φ gibt den Winkel (im Bogenmaß) zwischen der positiven reellen Achse und dem Ortsvektor von z an. Man nennt φ auch das *Argument* von z.



Beweis. Für $z = 0$ ist $z = 0 \cdot e^{i\varphi}$ mit beliebigem φ.

Es sei jetzt $z \neq 0$, $r := |z|$ und $\zeta := \frac{z}{r}$. Dann ist $|\zeta| = \frac{|z|}{r} = 1$. Es sei $\zeta = \xi + i\eta$, $\xi, \eta \in \mathbb{R}$. Dann gilt

$$\xi^2 + \eta^2 = 1 \text{ und } |\xi| \leq 1.$$

Deshalb ist

$$\alpha := \arccos \xi$$

definiert. Da $\cos \alpha = \xi$, folgt

$$\sin \alpha = \pm\sqrt{1 - \xi^2} = \pm\eta.$$

Wir setzen

$$\varphi := \begin{cases} \alpha & \text{falls } \sin \alpha = \eta, \\ -\alpha & \text{falls } \sin \alpha = -\eta. \end{cases}$$

Dann gilt

$$e^{i\varphi} = \cos \varphi + i \sin \varphi = \xi + i\eta = \zeta.$$

Damit gilt

$$z = re^{i\varphi}.$$

Die Eindeutigkeit von φ bis auf ein Vielfaches von 2π folgt aus Korollar 10.6. Denn

$$e^{i\varphi} = e^{i\psi} \Rightarrow e^{i(\varphi - \psi)} = 1 \Rightarrow \varphi - \psi = 2k\pi, \; k \in \mathbb{Z}.$$

$\square$

Bemerkung 10.4 Aus Satz 10.10 folgt, dass der Gruppenhomomorphismus

$$\begin{array}{ccc} (\mathbb{R}, +) & \to & (S^1, \cdot) \\ x & \mapsto & e^{ix} \end{array}$$

surjektiv ist.

Bemerkung 10.5 Satz 10.10 erlaubt eine einfache Interpretation der Multiplikation zweier komplexer Zahlen: Es sei $z_1 = r_1 e^{i\varphi_1}$, $z_2 = r_2 e^{i\varphi_2}$. Dann ist

$$z_1 \cdot z_2 = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}.$$

Man erhält also das Produkt zweier komplexer Zahlen, indem man ihre Beträge multipliziert und ihre Argumente addiert.

Korollar 10.7 (n-te Einheitswurzeln) *Es sei $n \in \mathbb{N}$, $n \geq 2$. Die Gleichung $z^n = 1$ hat genau n komplexe Lösungen, nämlich $z = \zeta_k$, wobei*

$$\zeta_k = e^{i\frac{2k\pi}{n}}, \quad k = 0, 1, \ldots, n-1.$$

Beweis. Es sei $z \in \mathbb{C}$ mit $z^n = 1$, $z = re^{i\varphi}$, $r \geq 0$, $0 \leq \varphi < 2\pi$. Wegen

$$1 = |z^n| = |z|^n = r^n$$

ist $r = 1$, also

$$z^n = (e^{i\varphi})^n = e^{in\varphi} = 1.$$

Nach Korollar 10.6 existiert ein $k \in \mathbb{Z}$ mit $n\varphi = 2k\pi$, d.h.

$$\varphi = \frac{2k\pi}{n}.$$

Wegen $0 \leq \varphi < 2\pi$ ist $0 \leq k < n$ und $z = \zeta_k$.

Umgekehrt gilt für jedes k

$$\zeta_k^n = \left(e^{i\frac{2k\pi}{n}}\right)^n = e^{i2k\pi} = 1.$$

□

11 Stetigkeit

Wir haben die Surjektivität der Exponentialfunktion damit begründet, dass wir den Graphen "durchgezogen" zeichnen können. Diese Eigenschaft wollen wir nun mathematisch exakt beschreiben. Dies führt zu dem Begriff der Stetigkeit.

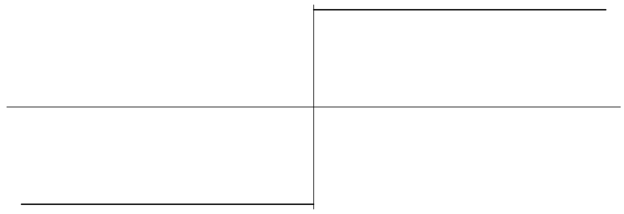
Vage umschrieben hat eine stetige Funktion die folgende Eigenschaft: Ändert man das Argument nur wenig, so ändert sich auch der Funktionswert nur wenig. Dies wollen wir nun präzisieren:

Definition Eine Funktion $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, heißt *stetig in $a \in D$* genau dann, wenn gilt: Für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} x_n = a$ gilt $\lim_{n\to\infty} f(x_n) = f(a)$.

Beispiel 11.1 (1) Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x$, also die Identität auf $\mathbb{R}$, ist stetig in jedem $a \in D = \mathbb{R}$.

(2) Wir betrachten $D = \mathbb{R}$ und $f : D \to \mathbb{R}$ mit

$$f(x) = \begin{cases} -1, & x < 0, \\ 1, & x \geq 0. \end{cases}$$



Diese Funktion ist stetig in jedem Punkt $a \neq 0$, aber nicht stetig an der Stelle $a = 0$.

Beweis. Die Folge $\left(\frac{1}{n}\right)_{n \geq 1}$ konvergiert gegen 0 und es gilt

$$\lim_{n\to\infty} f(\frac{1}{n}) = 1 = f(0).$$

Aber die Folge $\left(-\frac{1}{n}\right)_{n \geq 1}$ konvergiert ebenfalls gegen 0, es gilt jedoch

$$\lim_{n\to\infty} f(\frac{1}{n}) = -1 \neq f(0).$$

Die Folge $1, -1, \frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \ldots$ konvergiert auch gegen 0 und die Folge der Bildwerte ist $1, -1, 1, -1, 1, -1, \ldots$, eine divergente Folge. □

(3) Die konstante Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto c$, ist in jedem Punkt $a \in \mathbb{R}$ stetig.

Satz 11.1 *Es seien $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$ in $a \in D$ stetig. Dann sind auch $f + g$, $c \cdot f$ $(c \in \mathbb{R})$, $f \cdot g$ stetig in $a \in D$. Ebenso $\frac{f}{g}$, falls $g(x) \neq 0$ für alle $x \in D$.*

Beweis. Für jede Folge $(x_n)_{n\in\mathbb{N}}$ in D mit $\lim\limits_{n\to\infty} x_n = a$ gilt nach Voraussetzung

$$\lim_{n\to\infty} f(x_n) = f(a) \text{ und } \lim_{n\to\infty} g(x_n) = g(a).$$

Nach den Rechenregeln für konvergente Folgen sind die Folgen $((f+g)(x_n))$, $((cf)(x_n))$, $((f \cdot g)(x_n))$, $((\frac{f}{g})(x_n))$ somit auch konvergent und es gilt:

$$\begin{aligned}
\lim_{n\to\infty}(f+g)(x_n) &= (f+g)(a),\\
\lim_{n\to\infty}(cf)(x_n) &= (cf)(a),\\
\lim_{n\to\infty}(f\cdot g)(x_n) &= (f\cdot g)(a),\\
\lim_{n\to\infty}\left(\frac{f}{g}\right)(x_n) &= \frac{f(a)}{g(a)}.
\end{aligned}$$

□

Definition Eine Funktion $f\colon D \to \mathbb{R}$ heißt *stetig* genau dann, wenn f stetig in jedem $a \in D$ ist.

Bemerkung 11.1 Aus Satz 11.1 folgt, dass die stetigen Funktionen $f\colon D \to \mathbb{R}$ einen Vektorraum über $\mathbb{R}$ bilden.

Korollar 11.1 *Jede ganzrationale Funktion ist stetig. Jede rationale Funktion ist stetig in jedem Punkt ihres Definitionsbereichs.*

Beweis. Die konstanten Funktionen und die identische Funktion $x \mapsto x$ (jeweils mit der Definitionsmenge $\mathbb{R}$) sind stetig. Da die ganzrationale Funktion

$$x \mapsto a_0 + a_1 x + \cdots + a_n x^n$$

durch Multiplikation und Addition aus diesen Funktionen aufgebaut wird, folgt die Behauptung für die ganzrationalen Funktionen aus Satz 11.1. Auch aus Satz 11.1 ergibt sich dann, dass die rationalen Funktionen stetig in jedem Punkt ihres Definitionsbereichs sind. □

Unsere Definition der Stetigkeit hat den Nachteil, dass man *alle* gegen a konvergenten Folgen $(x_n)_{n\in\mathbb{N}}$ in Betracht zu ziehen hat. Der folgende Satz gibt eine äquivalente Bedingung an, die praktisch besser zu handhaben ist. Diese Bedingung wird oft auch als Definition der Stetigkeit genommen.

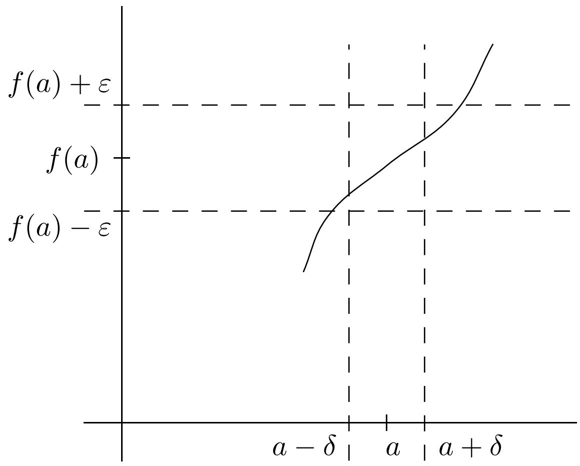
Satz 11.2 (ε-δ-Definition) *Eine Funktion $f : D \to \mathbb{R}$ ist genau dann in $a \in D$ stetig, wenn gilt: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in D$ mit $|x-a| < \delta$ gilt: $|f(x) - f(a)| < \varepsilon$.*

In Quantorenkurzform: f ist stetig in $a \in D$

$$\iff \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D\ (|x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon).$$

In Worte gefasst: f ist genau dann in $a \in D$ stetig, wenn man zu jedem (noch so kleinen) $\varepsilon > 0$ ein $\delta > 0$ finden kann, so dass sich die Funktionswerte

$f(x)$ aller $x \in D$, die sich von a um weniger als δ unterscheiden, von $f(a)$ um weniger als ε unterscheiden.



Beweis. „$\Leftarrow$": Es sei also $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in D$ für jedes $n \in \mathbb{N}$ und $\lim_{n\to\infty} x_n = a$ gegeben. Zu zeigen haben wir:

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

d.h. für alle $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass $|f(x_n) - f(a)| < \varepsilon$ für alle $n \geq n_0$ ist.

Es sei also $\varepsilon > 0$ gegeben. Nach Voraussetzung gibt es ein $\delta > 0$, so dass für alle $x \in D$ mit $|x - a| < \delta$ gilt:

$$|f(x) - f(a)| < \varepsilon.$$

Da (x_n) konvergent mit dem Grenzwert a ist, gibt es ein $n_0 \in \mathbb{N}$, so dass

$$|x_n - a| < \delta \text{ für alle } n \geq n_0.$$

Also gilt auch

$$|f(x_n) - f(a)| < \varepsilon \text{ für alle } n \geq n_0.$$

„$\Rightarrow$": (Indirekter Beweis) Angenommen, die Bedingung im Satz ist nicht erfüllt. Dann gibt es ein $\varepsilon > 0$, so dass für jedes $\delta > 0$ ein $x \in D$ mit $|x - a| < \delta$ existiert, für das gilt

$$|f(x) - f(a)| \geq \varepsilon.$$

Wir wählen nun $\delta = \frac{1}{n}$. Dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in D$ mit $|x_n - a| < \frac{1}{n}$ und $|f(x_n) - f(a)| \geq \varepsilon > 0$.

Wir haben also eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = a$ gefunden, deren „Bildfolge" $(f(x_n))_{n\in\mathbb{N}}$ nicht gegen $f(a)$ konvergieren kann. $\square$

Der Nachweis der Stetigkeit einer Funktion $f\colon D \to \mathbb{R}$ im Punkte $a \in D$ bedeutet also folgendes: Zunächst wird eine beliebige ε-Umgebung von $f(a)$ vorgegeben. Dazu ist eine δ-Umgebung von a so zu bestimmen, dass alle Punkte dieser Umgebung – soweit sie zu D gehören – ihre Bilder in der vorgegebenen ε-Umgebung von $f(a)$ haben.

Wie findet man zu beliebig vorgegebenem $\varepsilon > 0$ ein passendes $\delta > 0$ mit der formulierten Eigenschaft? Man versucht aus der Ungleichung $|f(x) - f(a)| < \varepsilon$ durch Umformung eine Bedingung für $|x - y|$ zu gewinnen und danach ein $\delta > 0$ passend zu wählen.

Beispiel 11.2 Die Quadratwurzelfunktion

$$\begin{aligned} \mathbb{R}_+ &\to \mathbb{R}_+ \\ x &\mapsto \sqrt{x} \end{aligned}$$

in an jeder Stelle $a \in \mathbb{R}_+$ stetig.

Beweis. Ist $a = 0$ und $\varepsilon > 0$ vorgegeben, so kann man $\delta = \varepsilon^2$ wählen. Dann gilt für alle x mit $0 \leq x < \varepsilon^2 = \delta$

$$\sqrt{x} < \varepsilon.$$

Sei nun $a > 0$ und $\varepsilon > 0$ vorgegeben. Wir formen um

$$\sqrt{x} - \sqrt{a} = \frac{x - a}{\sqrt{x} + \sqrt{a}}.$$

Für $\delta = (\sqrt{a})\varepsilon$ gilt:

$$|x - a| < \delta \Rightarrow |\sqrt{x} - \sqrt{a}| < \frac{\sqrt{a} \cdot \varepsilon}{\sqrt{x} + \sqrt{a}} \leq \frac{\sqrt{a}}{\sqrt{a}} \cdot \varepsilon = \varepsilon.$$

□

Satz 11.3 *Es sei $f\colon D \to \mathbb{R}$ stetig in $a \in D$ und $f(a) > 0$. Dann gibt es ein $\delta > 0$, so dass für jedes $x \in D$ mit $|x - a| < \delta$ gilt: $f(x) > 0$. Entsprechend gilt: Falls $f(a) < 0$, dann gibt es ein $\delta > 0$, so dass für jedes $x \in D$ mit $|x - a| < \delta$ gilt: $f(x) < 0$.*

Beweis. Betrachte den Fall $f(a) > 0$. Wähle $\varepsilon = f(a) > 0$. Hierzu gibt es ein $\delta > 0$, so dass für alle $x \in D$ mit $|x - a| < \delta$ gilt:

$$|f(x) - f(a)| < \varepsilon = f(a),$$

$$\text{d. h. } f(x) \in (\underbrace{f(a) - f(a)}_{=0}, f(a) + f(a)),$$

also $f(x) > 0$.

Entsprechend im Fall $f(a) < 0$: Wähle $\varepsilon = -f(a)$, oder wende den ersten Fall auf die Funktion $-f$ an. $\square$

Wir beweisen nun einige wichtige allgemeine Sätze über stetige Funktionen. Im Folgenden sei $[a,b] \subset \mathbb{R}$ ein abgeschlossenes Intervall mit $a,b \in \mathbb{R}$, $a < b$.

Satz 11.4 *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann ist f beschränkt in $[a,b]$, d. h. es gibt ein $K \in \mathbb{R}$, so dass $|f(x)| < K$ für alle $x \in [a,b]$.*

Bemerkung 11.2 Für nicht abgeschlossene Intervalle ist dieser Satz i.A. falsch! Beispiel:

$$f\colon \begin{array}{ll} (0,1] & \to \mathbb{R} \\ x & \mapsto \frac{1}{x} \end{array}$$

Beweis. Angenommen f ist nicht beschränkt. Dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in [a,b]$, so dass $|f(x_n)| \geq n$ (Sonst wäre n eine Schranke). Da die Folge $(x_n)_{n\in\mathbb{N}}$ beschränkt ist, hat sie nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit

$$\lim_{k\to\infty} x_{n_k} =: c.$$

Nach Satz 5.10 gilt $c \in [a,b]$. Aber die Bildfolge $(f(x_{n_k}))_{k\in\mathbb{N}}$ ist nicht einmal beschränkt, geschweige denn konvergent, da nach Konstruktion

$$|f(x_{n_k})| \geq n_k \geq k$$

für jedes $k \in \mathbb{N}$ gilt. Das ist aber ein Widerspruch zu der Voraussetzung, dass $f\colon [a,b] \to \mathbb{R}$ stetig ist. $\square$

Satz 11.5 *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann nimmt f Maximum und Minimum auf $[a,b]$ an, d. h. es gibt $x,y \in [a,b]$, so dass $f(x) = \sup f([a,b])$ und $f(y) = \inf f([a,b])$.*

Bemerkung 11.3 Dieser Satz ist wiederum i.A. falsch für nicht abgeschlossene Intervalle. Beispiel $f\colon (1,2) \to \mathbb{R}$, $x \mapsto x$.

Beweis. Wir geben nur den Beweis für das Maximum. Der Übergang von f zu $-f$ liefert dann die Behauptung für das Minimmum.

Nach dem eben bewiesenen Satz ist die Menge

$$f([a,b]) := \{f(x) \mid x \in [a,b]\}$$

beschränkt. Nach dem Vollständigkeitsaxiom besitzt $f([a,b])$ daher ein Supremum $s := \sup f([a,b])$. Die Zahl s ist die kleinste obere Schranke, also ist $s - \frac{1}{n}$ keine obere Schranke, also gibt es zu $n \in \mathbb{N}$ ein $x_n \in [a,b]$, so dass

$$(*) \qquad s - \frac{1}{n} < f(x_n) \leq s.$$

Die Folge $(x_n)_{n\in\mathbb{N}}$ besitzt eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit

$$\lim_{k\to\infty} x_{n_k} =: x \in [a,b].$$

Aus der Stetigkeit von f folgt

$$f(x) = \lim_{k\to\infty} f(x_{n_k}) \underset{(*)}{=} s.$$

$\square$

Satz 11.6 (Zwischenwertsatz) *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig und $f(a) < 0 < f(b)$. Dann gibt es ein $z \in [a,b]$ mit $f(z) = 0$.*

Bemerkung 11.4 Die Aussage ist anschaulich klar: Der Graph von f ist eine Kurve, die die Punkte $(a, f(a))$ und $(b, f(b))$ verbindet. Die Stetigkeit von f bedeutet gerade, das diese Kurve ohne Absetzen gezeichnet werden kann. Da $(a, f(a))$ unterhalb und $(b, f(b))$ oberhalb der x-Achse liegt, muss diese Kurve daher die x-Achse schneiden.

Die Aussage wird falsch, wenn man nur innerhalb der rationalen Zahlen arbeitet: Beispiel: $D = \{x \in \mathbb{Q} \mid 1 \le x \le 2\}$,

$$f\colon \begin{array}{l} D \to \mathbb{R} \\ x \mapsto x^2 - 2 \end{array}$$

Beweis. Wir betrachten die Menge

$$M := \{x \in [a,b] \mid f(x) \le 0\}.$$

Wegen $a \in M$ folgt $M \ne \emptyset$, M ist durch b nach oben beschränkt. Nach dem Vollständigkeitsaxiom existiert

$$z := \sup M.$$

Um zu zeigen, dass $f(z) = 0$ gilt, widerlegen wir die Annahmen

$$f(z) < 0 \tag{1}$$

$$f(z) > 0 \tag{2}$$

Zu (1): Nach Satz 11.3 gibt es ein $\delta > 0$, so dass für alle $x \in [a,b]$ mit $|x - z| < \delta$ gilt: $f(x) < 0$. Da $z \ne b$ (wäre $z = b$, so hätte man nach (1) $f(b) < 0$ im Widerspruch zur Voraussetzung), gilt $z < b$. Somit gibt es ein $x > z$ mit $f(x) < 0$. Daher kann z nicht obere Schranke von M sein.

Zu (2): Nach Satz 11.3 gibt es ein $\delta > 0$, so dass für alle $x \in [a,b]$ mit $|x - z| < \delta$ gilt: $f(x) > 0$. Da $z > a$ ist (für $z = a$ hätte man $f(a) > 0$ im Widerspruch zur Voraussetzung), kann dann aber z nicht die *kleinste* obere Schranke von M sein. $\square$

Satz 11.7 (Variante des Zwischenwertsatzes) *Es sei* $f\colon [a,b] \to \mathbb{R}$ *stetig und* c *eine reelle Zahl zwischen* $f(a)$ *und* $f(b)$. *Dann existiert ein* $z \in [a,b]$ *mit* $f(z) = c$.

Beweis. Sei etwa $f(a) < c < f(b)$. Wende den vorherigen Satz auf die Funktion
$$\begin{aligned} g\colon\ & [a,b] \to \mathbb{R} \\ & x \ \mapsto f(x) - c \end{aligned}$$
an. □

Wir geben nun eine Anwendung des Zwischenwertsatzes.

Satz 11.8 *Ein Polynom*

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$

ungeraden Grades n *besitzt mindestens eine Nullstelle in* $\mathbb{R}$.

Bemerkung 11.5 Ein Polynom geraden Grades braucht keine reelle Nullstelle zu besitzen, wie das Beispiel $x \mapsto x^2 + 2$ zeigt.

Beweis. O. B. d. A. sei $a_n > 0$. Für $x \neq 0$ ist

$$f(x) = x^n \cdot \left(\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \right).$$

Es gibt ein $K > 0$, so dass für alle $x \in \mathbb{R}$ mit $|x| > K$ gilt

$$\left| \frac{a_0}{x^n} + \cdots + \frac{a_{n-1}}{x} \right| < \frac{a_n}{2}.$$

Für $|x| > K$ gilt dann sogar

$$\frac{f(x)}{x^n} = \frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \geq \frac{a_n}{2} > 0.$$

Für $x > K$ ist deswegen $f(x) > 0$, für $x < -K$ ist $f(x) < 0$. Also hat nach dem Zwischenwertsatz f eine Nullstelle im Intervall $[-K, K]$. □

Nun tragen wir die noch ausstehenden Beweise nach.

Satz 11.9 *Die Exponentialfunktion* $\exp\colon \mathbb{R} \to \mathbb{R}_+^*$ *ist stetig und surjektiv.*

Beweis. (i) Wir zeigen zunächst die Stetigkeit. In Satz 7.5 haben wir bewiesen: Zu jedem $a \in \mathbb{R}$ existiert ein $L > 0$ mit

$$x, y \leq a \Rightarrow |\exp(x) - \exp(y)| \leq L|x - y|.$$

Es sei $b \in \mathbb{R}$ gegeben. Setze $a := b + 1$. Es sei $L > 0$ die Konstante, die zu a nach der zitierten Aussage existiert. Es sei $\varepsilon > 0$ gegeben. Setze

$$\delta := \min(1, \frac{\varepsilon}{L}).$$

Dann gilt für alle $x \in \mathbb{R}$ mit $|x - b| < \delta$

$$|\exp(x) - \exp(b)| \leq L|x - b| < L\delta = \varepsilon.$$

Also ist exp stetig in b. Da $b \in \mathbb{R}$ beliebig war, ist exp stetig.

(ii) Zum Beweis der Surjektivität sei $y > 0$ gegeben. Dann gibt es ein $n \in \mathbb{N}$, so dass $y \in [e^{-n}, e^n]$, denn die Folge $(e^{-n})_{n\in\mathbb{N}}$ ist eine Nullfolge und die Folge $(e^n)_{n\in\mathbb{N}}$ wächst unbeschränkt. Aus dem Zwischenwertsatz folgt nun, dass es ein $x \in [-n, n]$ gibt, so dass $\exp(x) = y$ ist. □

Zusammen mit Satz 11.1 folgt:

Korollar 11.2 *Die Funktionen* $\sinh : \mathbb{R} \to \mathbb{R}$ *und* $\cosh : \mathbb{R} \to \mathbb{R}$ *sind stetig.*

Satz 11.10 *Die Funktionen* $\sin : \mathbb{R} \to \mathbb{R}$ *und* $\cos : \mathbb{R} \to \mathbb{R}$ *sind stetig.*

Beweis. Es sei $a \in \mathbb{R}$ und $\varepsilon > 0$ gegeben. Wähle $\delta = \varepsilon$. Dann gilt für alle $x \in \mathbb{R}$ mit $|x - a| > \delta$ wegen Satz 10.7

$$|\sin x - \sin a| \leq |x - a| < \varepsilon = \delta.$$

Entsprechend für cos. □

Zusammen mit Satz 11.1 folgt:

Korollar 11.3 *Die Funktionen* tan *und* cot *sind stetig auf ihren Definitionsbereichen.*

Nun tragen wir den Beweis von Satz 10.8 nach.

Beweis von Satz 10.8. Da $\cos 0 = 1$ und $\cos 2 \leq -\frac{1}{3}$ (Lemma 10.1), besitzt die Funktion cos nach dem Zwischenwertsatz im Intervall $[0, 2]$ mindestens eine Nullstelle. Nach Lemma 10.3 gibt es nicht mehr als eine Nullstelle. □

Für den Beweis von Satz 10.9 benötigen wir noch den folgenden Satz:

Satz 11.11 *Es sei* $f : [a, b] \to \mathbb{R}$ *eine stetige, streng monoton wachsende Funktion und* $c := f(a)$, $d := f(b)$. *Dann gilt:*

(i) f *bildet* $[a, b]$ *bijektiv auf* $[c, d]$ *ab, und*

(ii) *die Umkehrabbildung* $f^{-1} : [c, d] \to [a, b]$ *ist auch stetig.*

Beweis. Zu (i): f ist injektiv nach Satz 8.4. Aus $a < x < b$ folgt $c < f(x) < d$, also bildet f das Intervall $[a,b]$ in $[c,d]$ ab. Die Surjektivität von f folgt aus dem Zwischenwertsatz.

Zu (ii): Es sei $q \in [c,d]$. Dann gibt es $p \in [a,b]$ mit $f(p) = q$. Zu vorgegebenem $\varepsilon > 0$ müssen wir nun ein $\delta > 0$ finden, so dass für alle $x \in [c,d]$ gilt

$$f(p) - \delta < x < f(p) + \delta \Rightarrow p - \varepsilon < f^{-1}(x) < p + \varepsilon.$$

Wir setzen zunächst voraus, dass $c < q < d$, also auch $a < p < b$. O.B.d.A. sei $\varepsilon > 0$ so gewählt, dass $p \pm \varepsilon \in [a,b]$ (eventuell muss man das gegebene ε verkleinern). Wegen

$$p - \varepsilon < p < p + \varepsilon$$

gilt

$$f(p - \varepsilon) < f(p) < f(p + \varepsilon).$$

Wähle

$$\delta := \min\{f(p+\varepsilon) - f(p), f(p) - f(p-\varepsilon)\}.$$

Dieses δ tut es: Es sei $x \in [c,d]$ mit

$$f(p) - \delta < x < f(p) + \delta.$$

Dann gilt

$$f(p - \varepsilon) < x < f(p + \varepsilon),$$

also

$$p - \varepsilon < f^{-1}(x) < p + \varepsilon.$$

Gilt $q = c$ oder $q = d$, etwa $q = c$, so können wir ε nur so verkleinern, dass $p + \varepsilon \in [a,b]$. Wähle dann $\delta = f(p+\varepsilon) - f(p)$. Analog im Fall $q = d$.
□

Beweis von Satz 10.9. (a), (b) Aus dem schon bewiesenen Teil und Satz 11.11 folgt, dass cos und sin bijektiv sind.

(c) Wir müssen noch zeigen, dass $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ surjektiv ist. Dazu zeigen wir zunächst, dass für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n < \frac{\pi}{2}$ und $\lim_{n\to\infty} x_n = \frac{\pi}{2}$ die Folge $(\tan x_n)_{n\in\mathbb{N}}$ nach oben unbeschränkt ist.

Es sei (x_n) eine solche Folge, o.B.d.A. sei $x_n > 0$ für alle $n \in \mathbb{N}$. Dann gilt

$$\lim_{n\to\infty} \cot(x_n) = \frac{\lim_{n\to\infty} \cos x_n}{\lim_{n\to\infty} \sin x_n} = \frac{\cos \frac{\pi}{2}}{\sin \frac{\pi}{2}} = \frac{0}{1} = 0.$$

Wegen $\tan x_n = \frac{1}{\cot x_n}$ wächst daher die Folge $(\tan x_n)_{n\in\mathbb{N}}$ unbeschränkt.

Wegen $\tan(-x) = -\tan x$ ist für jede Folge (x_n) mit $-\frac{\pi}{2} < x_n$ und $\lim_{n\to\infty} x_n = -\frac{\pi}{2}$ die Folge $(\tan x_n)$ nach unten unbeschränkt.

Der Rest der Behauptung folgt analog zum Beweis von Satz 11.9 aus dem Zwischenwertsatz.

(d) Der Beweis der Surjektivität von $\cot : (0, \pi) \to \mathbb{R}$ verläuft analog. □

Vor weiteren Anwendungen schieben wir noch einige Bemerkungen über die Restriktion von Funktionen ein. Eine Funktion $f : D \to \mathbb{R}$ ist nur vollständig definiert, wenn man ihren Definitionsbereich D mit angibt. Dann erst kann man von Stetigkeit dieser Funktion reden. Ist nun $D_1 \subseteq D$, so heißt

$$\begin{array}{rccc} f|_{D_1}: & D_1 & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

die *Restriktion* oder *Einschränkung von* f *auf* D_1. Sie unterscheidet sich von f nur durch den neuen, kleineren Definitionsbereich. Wenn es aus dem Zusammenhang klar ist, schreibt man manchmal auch etwas ungenauer f anstelle von $f|_{D_1}$.

Satz 11.12 *Es sei $f : D \to \mathbb{R}$ eine Funktion, $D_1 \subseteq D$, $a \in D_1$. Ist f stetig in a, so ist auch $f|_{D_1}$ stetig in a.*

Beweis. Jede Folge $(x_n)_{n\in\mathbb{N}}$ von Elementen aus D_1 ist auch eine Folge von Elementen aus D. Deshalb folgt aus $\lim_{n\to\infty} x_n = a$

$$\lim_{n\to\infty} f|_{D_1}(x_n) = \lim_{n\to\infty} f(x_n) = f(a).$$

□

Dass der Satz nicht ganz trivial ist, sieht man daran, dass die Umkehrung falsch ist: Jede Funktion auf $D_1 = \{a\}$ ist stetig, da die einzige Folge in D_1 die Folge $a, a, a, \ldots$ ist, deren Bildfolge $f(a), f(a), f(a), \ldots$ offenbar gegen $f(a)$ konvergiert. Es gilt aber der folgende Satz.

Satz 11.13 *Es sei $f : D \to \mathbb{R}$ eine Funktion und $a \in D_1 \subseteq D$. Es gebe ein $\eta > 0$, so dass $(a - \eta, a + \eta) \subseteq D_1$. Dann ist $f : D \to \mathbb{R}$ stetig in a genau dann, wenn $f|_{D_1}$ stetig in a ist.*

Beweis. "$\Rightarrow$" ist klar.

"$\Leftarrow$": Es sei $(x_n)_{n\in\mathbb{N}}$ eine Folge von Elementen aus D mit $\lim_{n\to\infty} x_n = a$. Dann gilt für fast alle $n \in \mathbb{N}$ die Ungleichung $|x_n - a| < \eta$, also sind nach Voraussetzung fast alle $x_n \in D_1$. Die endlich vielen Ausnahmen stören nicht, also gilt

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f|_{D_1}(x_n) = f|_{D_1}(a) = f(a).$$

□

Korollar 11.4 *Die Funktion $\ln : \mathbb{R}_+^* \to \mathbb{R}$ ist überall stetig.*

Beweis. Es sei $s \in \mathbb{R}_+^*$ und $t := \ln s$. Dann ist $e^t = s$. Wir legen nun um t ein abgeschlossenes Intervall $[a, b]$ herum, etwa $[t-1, t+1]$. Da exp auf $\mathbb{R}$ stetig ist, ist auch die Einschränkung von exp auf das Intervall $[t-1, t+1]$ stetig und das Intervall $[t-1, t+1]$ wird durch $\exp|_{[t-1,t+1]}$ auf $[e^{t-1}, e^{t+1}] =: [c, d]$ abgebildet. Nach Satz 11.11 ist $\ln|_{[c,d]}$ stetig. Es gilt $c < d$ und $s \in (c, d)$. Wir können also ein $\eta > 0$ finden, so dass $(s - \eta, s + \eta) \subseteq [c, d]$. Nach Satz 11.13 ist also auch ln in s stetig (und nicht nur $\ln|_{[c,d]}$!). □

Korollar 11.5 *Die Funktionen* $\arccos : [-1, 1] \to [0, \pi]$, $\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ *und* $\operatorname{arccot} : \mathbb{R} \to (0, \pi)$ *sind stetig auf ihren jeweiligen Definitionsbereichen.*

Beweis. Für $\arccos : [-1, 1] \to [0, \pi]$ und $\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ folgt dies direkt aus Satz 11.11, für die anderen beiden Funktionen argumentieren wir wie im Beweis des vorhergehenden Korollars. □

Satz 11.14 *Es sei $f: D_1 \to \mathbb{R}$ stetig in $a \in D_1$, $g: D_2 \to \mathbb{R}$, $f(D_1) \subseteq D_2$, g stetig in $f(a)$. Dann ist $g \circ f$ stetig in a. ($g \circ f$ ist die durch $(g \circ f)(x) = f(g(x))$ für $x \in D_1$ definierte Abbildung $g \circ f : D_1 \to \mathbb{R}$ und heißt die* Hintereinanderschaltung *der Funktionen $f : D_1 \to D_2$ und $g : D_2 \to \mathbb{R}$.)*

Beweis. Es sei $(x_n)_{n \in \mathbb{N}}$ eine beliebige Folge mit $\lim\limits_{n \to \infty} x_n = a$. Nach Voraussetzung gilt

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

und damit weiter

$$\lim_{n \to \infty} g(f(x_n)) = g(f(a)),$$

d.h.

$$\lim_{n \to \infty} (g \circ f)(x_n) = (g \circ f)(a).$$

□

Satz 11.15 *Es sei $a > 0$. Dann ist die Funktion $x \mapsto a^x = e^{x \cdot \ln a}$ von $\mathbb{R}$ nach $\mathbb{R}_+^*$ stetig.*

Beweis. Diese Funktion ist die Hintereinanderschaltung der Funktionen $x \mapsto x \cdot \ln a$ (ein Polynom!) und $y \mapsto e^y$, die stetig sind. □

Entsprechend zeigt man:

Satz 11.16 *Es sei $r \in \mathbb{R}$. Dann ist die Funktion $x \mapsto x^r = e^{r \ln x}$ von $\mathbb{R}_+^*$ nach $\mathbb{R}$ stetig.*

Bemerkung 11.6 Für $r \neq 0$ ist die Funktion $x^r : \mathbb{R}_+^* \to \mathbb{R}_+^*$ bijektiv als Hintereinanderschaltung der bijektiven Abbildungen

$$f: \begin{array}{ccc} \mathbb{R}_+^* & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \ln x \end{array}, \quad g: \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & r \cdot y \end{array}, \quad h: \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}_+^* \\ z & \longmapsto & e^z \end{array}.$$

Insbesondere lässt sich jedes $y \in \mathbb{R}_+^*$ auf genau eine Weise als $y = x^r$ schreiben, nämlich mit $x = y^{1/r}$. Insbesondere ist also für $k \in \mathbb{N}$, $k \geq 2$, auch die *k-te Wurzel* $\mathbb{R}_+^* \to \mathbb{R}_+^*$, $x \mapsto x^{1/k} =: \sqrt[k]{x}$, wohldefiniert und stetig.

Wir führen nun einen *Grenzwertbegriff für Funktionen* ein.

Definition Es sei $f\colon D \to \mathbb{R}$ eine Funktion und $a \in \mathbb{R}$ ein Punkt derart, dass es mindestens eine Folge $(a_n)_{n\in\mathbb{N}}$, $a_n \in D$, $a_n \neq a$ gibt mit $\lim\limits_{n\to\infty} a_n = a$ (z.B. $a \in D$). Man schreibt

$$\lim_{x\to a} f(x) = c,$$

falls für jede Folge $(x_n)_{n\in\mathbb{N}}$, $x_n \in D$, $x_n \neq a$, mit $\lim\limits_{n\to\infty} x_n = a$ gilt:

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

Bemerkung 11.7 Die Definition der Stetigkeit kann mit dieser Bezeichnung so formuliert werden: Eine Funktion $f\colon D \to \mathbb{R}$ ist in $a \in D$ stetig genau dann, wenn $\lim\limits_{x\to a} f(x) = f(a)$.

Andererseits folgt aus dem Beweis von Satz 11.2, dass gilt: $\lim\limits_{x\to a} f(x) = c \Leftrightarrow$ Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $x \in D$ mit $0 < |x-a| < \delta$ gilt: $|f(x) - c| < \varepsilon$.

12 Differenzierbarkeit

Es sei $f\colon D \to \mathbb{R}$ eine Funktion und $a \in D$. Einen Ausdruck der Form

$$\frac{f(a+h) - f(a)}{h},$$

wobei $a + h \in D$ und $h \neq 0$, nennen wir *Differenzenquotient* von f an der Stelle a. Anschaulich gibt der Differenzenquotient die Steigung der Sekante des Graphen von f zwischen den Punkten $(a, f(a))$ und $(a+h, f(a+h))$ an.

Definition Die Funktion f heißt *differenzierbar* in a, falls der Grenzwert

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

existiert. Hierbei sind bei der Limesbildung nur solche Folgen (h_n) mit $\lim\limits_{n\to\infty} h_n = 0$ zugelassen, für die $h_n \neq 0$ und $x + h_n \in D$ für alle $n \in \mathbb{N}$ gilt, und es wird vorausgesetzt, dass es mindestens eine solche Folge gibt.

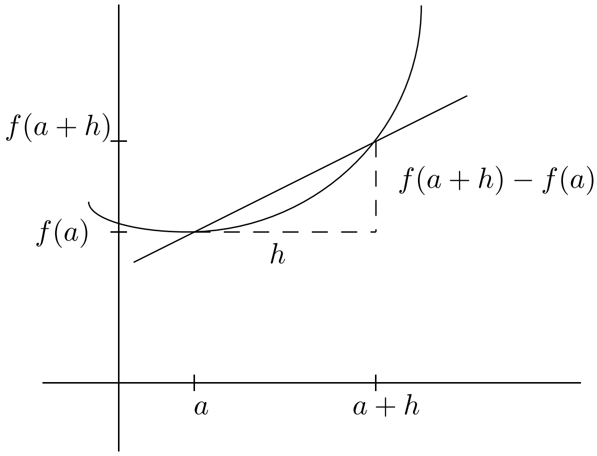


Abbildung 4: Steigung der Sekante von $(a, f(a))$ nach $(a+h, f(a+h))$

Bemerkung 12.1 Dass der Grenzwert

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

existiert, bedeutet: Für jede Nullfolge (h_n) mit $h_n \neq 0$ und $x + h_n \in D$ konvergiert die Folge der Differenzenquotienten

$$\left(\frac{f(a+h_n)-f(a)}{h_n}\right)_{n\in\mathbb{N}}.$$

Der Limes ist unabhängig von der gewählten Nullfolge: Es sei (k_n) eine andere solche Nullfolge. Betrachte die Folge $(h_1, k_1, h_2, k_2, h_3, k_3, \dots)$. Dann konvergiert die zugehörige Folge von Differenzenquotienten. Alle deren Teilfolgen konvergieren aber gegen denselben Grenzwert.

Definition Ist f differenzierbar in a, so heißt

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

die *Ableitung* von f an der Stelle a.

Anschaulich gibt $f'(a)$ die Steigung der Tangente an den Graphen von f in Punkte $(a, f(a))$ an.

Definition Eine Funktion $f\colon D \to \mathbb{R}$ heißt *differenzierbar*, falls f in jedem Punkt $a \in D$ differenzierbar ist.

Man kann die Ableitung auch darstellen als

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

Das bedeutet, dass die Funktion

$$\begin{aligned} R\colon\ & D \backslash \{a\} \to \mathbb{R} \\ & x \quad \mapsto R(x) := \tfrac{f(x)-f(a)}{x-a} \end{aligned}$$

in a stetig fortsetzbar ist. Denn wenn wir als Funktionswert in a den Wert $R(a) := f'(a)$ definieren, so ist R in a stetig. Also gilt:

Satz 12.1 *Die Funktion f ist differenzierbar in a mit Ableitung $f'(a)$ genau dann, wenn die Funktion*

$$R(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{für } x \neq a \\ f'(a) & \text{für } x = a \end{cases}$$

stetig in a ist.

Bemerkung 12.2 Es sei

$$\varphi(x) = R(x) - f'(a).$$

Dann gilt

$$f(x) = f(a) + f'(a)(x - a) + \varphi(x)(x - a).$$

Nach Satz 12.1 ist f genau dann in a differenzierbar, wenn $\lim_{x \to a} \varphi(x) = 0$ ist. Daher kann man $f'(a)(x - a)$ als *lineare Approximation* der Funktion f auffassen.

Beispiel 12.1 Betrachte $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^n$. Dann ist

$$R(x) = \frac{x^n - a^n}{x - a} = x^{n-1} + ax^{n-2} + \cdots + a^{n-1}$$

für $x \neq a$. Offensichtlich läßt sich $R(x)$ zu einer in a stetigen Funktion erweitern durch

$$R(a) = f'(a) = na^{n-1}.$$

Satz 12.2 *Ist die Funktion $f\colon D \to \mathbb{R}$ im Punkt $a \in D$ differenzierbar, so ist sie in a auch stetig.*

Beweis. Für alle $x \in D$ gilt

$$f(x) = f(a) + (x - a)R(x)$$

Nach Satz 12.1 ist $R(x)$ stetig in a, also auch $f(x)$. □

Nun einige Differenzierbarkeitsregeln:

Satz 12.3 *Es seien $f, g\colon D \to \mathbb{R}$ in $a \in D$ differenzierbare Funktionen und $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen $f+g, \lambda f, fg\colon D \to \mathbb{R}$ in a differenzierbar und es gilt*

$$\begin{aligned}
(f+g)'(a) &= f'(a) + g'(a) \\
(\lambda f)'(a) &= \lambda f'(a) \\
(fg)'(a) &= f'(a)g(a) + f(a)g'(a) \qquad \textit{(Produktregel)}
\end{aligned}$$

Ist $g(x) \neq 0$ für alle $x \in D$, so ist auch die Funktion $\frac{f}{g}\colon D \to \mathbb{R}$ in a differenzierbar und es gilt

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2} \qquad \textit{(Quotientenregel)}$$

Bemerkung 12.3 Differenzierbare Funktionen bilden einen Vektorraum über $\mathbb{R}$.

Beweis. (a) Die ersten beiden Beziehungen folgen unmittelbar aus den Rechenregeln für Grenzwerte von Folgen.

(b) *Produktregel*:

$$\begin{aligned}
(fg)'(a) &= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h} \\
&= \lim_{h\to 0} \frac{1}{h}[f(a+h)(g(a+h) - g(a)) + (f(a+h) - f(a))g(a)] \\
&= \lim_{h\to 0} f(a+h)\frac{g(a+h) - g(a)}{h} + \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} g(a)
\end{aligned}$$

wegen Stetigkeit von f in a:

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

(c) *Quotientenregel*: Wir behandeln zunächst den Spezialfall $f = 1$.

$$\begin{aligned}
\left(\frac{1}{g}\right)'(a) &= \lim_{h\to 0} \frac{1}{h}\left(\frac{1}{g(a+h)} - \frac{1}{g(a)}\right) \\
&= \lim_{h\to 0} \frac{1}{g(a+h)g(a)}\left(\frac{g(a) - g(a+h)}{h}\right) \\
&= \frac{-g'(a)}{g(a)^2}.
\end{aligned}$$

Der allgemeine Fall folgt hieraus mit Hilfe der Produktregel:

$$\begin{aligned}
\left(\frac{f}{g}\right)'(a) &= \left(f \cdot \frac{1}{g}\right)'(a) \\
&= f'(a)\frac{1}{g(a)} + f(a)\frac{-g'(a)}{g(a)^2} \\
&= \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.
\end{aligned}$$

□

Wir berechnen nun die Ableitungen einiger bekannter Funktionen.

Beispiel 12.2 (1) Für eine konstante Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = c$, gilt $(a \in \mathbb{R})$:

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

(2) Für die identische Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x$, gilt $(a \in \mathbb{R})$

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h \to 0} \frac{a + h - a}{h} = 1.$$

Um weitere Beispiele zu studieren benötigen wir einen Satz.

Satz 12.4 *Die Reihe*

$$\sum_{k=0}^{\infty} a_k x^k \qquad (a_k \in \mathbb{R}, x \in \mathbb{R})$$

sei absolut konvergent für $x \in (-b, b)$, $b > 0$. Sei $f\colon (-b, b) \to \mathbb{R}$ die Funktion

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

Dann ist f stetig in 0.

Beweis. Es sei $0 < \tilde{b} < b$ und $x \in \mathbb{R}$ mit $|x| \le \tilde{b}$. Dann gilt

$$\begin{aligned} |f(x) - f(0)| &= \left| \sum_{k=1}^{\infty} a_k x^k \right| \\ &= |x| \left| \sum_{k=1}^{\infty} a_k x^{k-1} \right| \le |x| \cdot \sum_{k=1}^{\infty} |a_k| \tilde{b}^{k-1}. \end{aligned}$$

Es sei $\varepsilon > 0$ gegeben. Wir setzen

$$M := \sum_{k=1}^{\infty} |a_k| \tilde{b}^{k-1}.$$

Dann gilt für $\delta := \min\{\tilde{b}, \frac{\varepsilon}{M}\}$ und für $x \in \mathbb{R}$ mit $|x| < \delta$

$$|f(x) - f(0)| < \varepsilon.$$

□

Beispiel 12.3 (3) $\exp\colon \mathbb{R} \to \mathbb{R}$, $a \in \mathbb{R}$:

$$\begin{aligned}\exp'(a) &= \lim_{h\to 0} \frac{\exp(a+h) - \exp(a)}{h}\\ &= \lim_{h\to 0} \exp(a)\frac{\exp(h)-1}{h}\\ &= \exp(a) \lim_{h\to 0} \frac{\exp(h)-1}{h}\end{aligned}$$

Nun gilt

$$\frac{\exp(h)-1}{h} = 1 + \frac{h}{2!} + \frac{h^2}{3!} + \frac{h^3}{4!} + \cdots$$

Diese Potenzreihe ist absolut konvergent für jedes h und damit nach Satz 12.4 stetig in 0. Es gilt also

$$\lim_{h\to 0} \frac{\exp(h)-1}{h} = 1.$$

Also folgt

$$\exp'(a) = \exp(a)$$

(4) $\sin\colon \mathbb{R} \to \mathbb{R}$, $\cos\colon \mathbb{R} \to \mathbb{R}$.

Wir berechnen zunächst die Ableitung von sin und cos in 0:

$$\sin'(0) = \lim_{h\to 0} \frac{\sin h}{h} = 1,$$

da

$$\frac{\sin(h)}{h} = 1 - \frac{h^2}{3!} + \frac{h^4}{5!} - \frac{h^6}{7!} \pm \cdots$$

und die rechte Seite eine absolut konvergente Potenzreihe ist, die nach Satz 12.4 stetig in 0 ist.

Ebenso zeigt

$$\frac{\cos(h)-1}{h} = -\frac{h}{2!} + \frac{h^3}{4!} - \frac{h^5}{6!} \pm \cdots,$$

dass

$$\cos'(0) = \lim_{h\to 0} \frac{\cos(h)-1}{h} = 0.$$

Nun berechnen wir die Ableitung an einer beliebigen Stelle a: Dazu:

$$\sin(a+h) = \sin a \cdot \cos(h) + \cos a \cdot \sin(h),$$

also

$$\begin{aligned}\sin'(a) &= \lim_{h\to 0} \frac{\sin(a+h) - \sin a}{h}\\ &= \lim_{h\to 0} \left[\sin a \frac{\cos(h)-1}{h} + \cos a \frac{\sin(h)}{h}\right]\\ &= \cos a\end{aligned}$$

Entsprechend

$$\begin{aligned}\cos'(a) &= \lim_{h\to 0} \frac{\cos(a+h)-\cos a}{h}\\ &= \lim_{h\to 0}\left[\cos a \frac{\cos h - 1}{h} - \sin a \frac{\sin h}{h}\right]\\ &= -\sin a.\end{aligned}$$

(5) $\tan\colon \mathbb{R}\backslash\{\frac{\pi}{2}+k \mid k\in\mathbb{Z}\}\to\mathbb{R}$ nach Quotientenregel:

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

(6) Die Funktion $f:\mathbb{R}\to\mathbb{R}$, $x\mapsto |x|$, ist nicht in 0 differenzierbar: Es sei $h_n = (-1)^n\frac{1}{n}$ $(n\geq 1)$. Es gilt $\lim_{n\to\infty} h_n = 0$, aber die Folge (q_n) mit

$$q_n := \frac{|0+h_n|-|0|}{h_n} = \frac{\frac{1}{n}-0}{(-1)^n\frac{1}{n}} = (-1)^n$$

konvergiert nicht.

Satz 12.5 (Kettenregel) *Es seien $f\colon D_1\to\mathbb{R}$ und $g\colon D_2\to\mathbb{R}$ Funktionen mit $f(D_1)\subseteq D_2$. Die Funktion f sei differenzierbar in $a\in D_1$ und g sei differenzierbar in $b:=f(a)\in D_2$. Dann ist $g\circ f\colon D_1\to\mathbb{R}$ differenzierbar in a und es gilt*

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

Beweis. Es gilt

$$\begin{aligned} f(x) &= f(a)+R(x)(x-a) \qquad (x\in D_1),\\ g(y) &= g(b)+S(y)(y-b) \qquad (y\in D_2),\end{aligned}$$

wobei $R(x)$ bzw. $S(y)$ die Differenzenquotientenfunktionen von f bzw. g bzgl. a bzw. $b=f(a)$ sind. Also ist

$$g(f(x)) = g(f(a)) + S(f(x))\cdot R(x)(x-a),$$

und $T(x) := S(f(x))\cdot R(x)$ ist die Differenzenquotientenfunktion von $g\circ f$ bzgl. a. Nach Voraussetzung ist R stetig in a und S stetig in $b=f(a)$. Nach Satz 11.14 und 11.1 ist deshalb auch T stetig in a und es gilt

$$T(a) = S(f(a))\cdot R(a) = g'(f(a))\cdot f'(a).$$

□

Satz 12.6 (Ableitung der Umkehrfunktion) *Es sei $a < b$ und $f\colon [a,b] \to \mathbb{R}$ streng monoton steigend und stetig, $f\colon [a,b] \to [c,d]$ bijektiv. Außerdem sei f in $x_0 \in (a,b)$ differenzierbar und $f'(x_0) \neq 0$.*

Dann ist $f^{-1}\colon [c,d] \to [a,b]$ in $f(x_0)$ differenzierbar und es gilt

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

Bemerkung 12.4 Wüssten wir schon, dass f^{-1} in $f(x_0)$ differenzierbar ist, dann ist die Formel klar: Man wendet die Kettenregel auf $f^{-1} \circ f = \mathrm{id}$ an:

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

Insbesondere gilt also: Aus f differenzierbar in x_0 und f^{-1} differenzierbar in $f(x_0)$ folgt $f'(x_0) \neq 0$. Also ist die Voraussetzung $f'(x_0) \neq 0$ notwendig.

Beweis des Satzes. Es sei $y_0 = f(x_0)$, $y \in [c,d] \backslash \{y_0\}$. Dann gilt

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

Lassen wir y eine beliebige Folge $(y_n)_{n \in \mathbb{N}}$ mit $\lim\limits_{n \to \infty} y_n = y_0$ durchlaufen, so ist nach der Stetigkeit von f^{-1}:

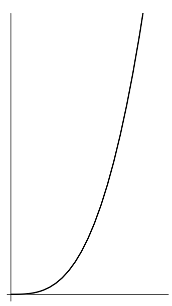
$$\lim_{n \to \infty} \underbrace{f^{-1}(y_n)}_{=: x_n} = f^{-1}(y_0) = x_0.$$

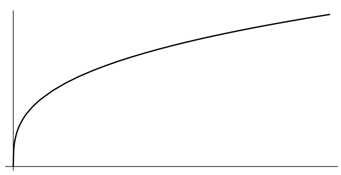
Also gilt:

$$\lim_{n \to \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0}{y_n - y_0} = \frac{1}{\lim\limits_{n \to \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}} = \frac{1}{f'(x_0)}.$$

□

Beispiel 12.4 $x \mapsto x^3$:





Man beachte den Nullpunkt.

Bemerkung 12.5 (Notation) Die Leibnizsche Notation für $f'(x)$ ist

$$\frac{df(x)}{dx}$$

Diese Schreibweise wird gerne benutzt und hat den folgenden Ursprung: In klassischer Notation schreibt man auch $y = f(x)$, wenn man von einer Funktion f spricht. Einen Differenzenquotienten schreibt man klassisch wie folgt:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} =: \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} =: \frac{dy}{dx}.$$

Diesen letzten Quotienten nennt man *Differentialquotient*, und dy und dx nennt man *Differentiale*. Die Symbole dy und dx werden als infinitesimal kleine Differenzen aufgefaßt. Rechnen darf man damit wie mit reellen Zahlen.

Diese klassische Notation bringt zwar gewisse begriffliche Schwierigkeiten mit sich, ist aber ungemein praktisch, weil formales Herumrechnen zum richtigen Resultat führt.

Beispiel 12.5 Der obige Satz liest sich dann so: $y = f(x)$, also $x = f^{-1}(y)$. Dann ist

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

Das ist die richtige Formel und obiger Beweis ist ihr nachgebildet.

Beispiel 12.6 Ein Rechenbeispiel: $y = f(x) = x^2$ (für $x > 0$). Dann ist $x = \sqrt{y} = y^{\frac{1}{2}}$, also

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{2x} = \frac{1}{2\sqrt{y}} = \frac{1}{2} y^{-\frac{1}{2}}.$$

Um den Punkt anzugeben, in dem der Differentialquotient gebildet wird, sind folgende Notationen üblich:

$$f'(a) = \frac{df(a)}{dx} = \frac{df}{dx}(a) = \left.\frac{df(x)}{x}\right|_{x=a}.$$

Beispiel 12.7 (7) $\ln \colon \mathbb{R}_+^* \to \mathbb{R}$,

$$\ln'(y_0) = \frac{1}{\exp'(\ln y_0)} = \frac{1}{\exp(\ln y_0)} = \frac{1}{y_0}.$$

(8) $x \mapsto x^\alpha$, $\alpha \in \mathbb{R}$,

$$\frac{dx^\alpha}{dx} = \exp'(\alpha \ln x) \frac{d}{dx}(\alpha \ln x) = \exp(\alpha \ln x) \frac{\alpha}{x} = x^\alpha \cdot \frac{\alpha}{x} = \alpha x^{\alpha - 1}.$$

(9) $\arcsin\colon [-1,1] \to \mathbb{R}$.
Setze $y = \arcsin x$:

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{\cos(y)} = \frac{1}{\sqrt{1-x^2}} \text{ für } x \in (-1,1),$$

da $\sin y = x$ und $\cos y = \sqrt{1-x^2}$. Also

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} \text{ für } -1 < x < 1.$$

(10) $\arctan\colon \mathbb{R} \to \mathbb{R}$.
Setze $y = \arctan x$. Es gilt

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \cos^2(y).$$

Nun gilt aber

$$x^2 = \tan^2 y = \frac{\sin^2 y}{\cos^2 y} = \frac{1-\cos^2 y}{\cos^2 y} = \frac{1}{\cos^2 y} - 1,$$

also

$$\cos^2(y) = \frac{1}{1+x^2}.$$

Deshalb gilt

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

Wir kommen jetzt zu den ersten Anwendungen der Differentiation.

Definition Es sei $f\colon (a,b) \to \mathbb{R}$ eine Funktion. Man sagt, f habe in $x_0 \in (a,b)$ ein *lokales Maximum* (bzw. *Minimum*), wenn $\delta > 0$ existiert, so dass $(x_0-\delta, x_0+\delta) \subseteq (a,b)$ und $f(x) \leq f(x_0)$ (bzw. $f(x) \geq f(x_0)$) für alle $x \in (x_0-\delta, x_0+\delta)$.

Man sagt f habe in $x_0 \in (a,b)$ ein *isoliertes lokales Maximum* (bzw. *Minimum*), wenn $\delta > 0$ existiert, so dass $(x_0-\delta, x_0+\delta) \subseteq (a,b)$ und $f(x) < f(x_0)$ (bzw. $f(x) > f(x_0)$) für alle $x \neq x_0$, $x \in (x_0-\delta, x_0+\delta)$.

Extremum = Maximum oder Minimum.

Satz 12.7 *Die Funktion $f\colon (a,b) \to \mathbb{R}$ besitze im Punkt $x_0 \in (a,b)$ ein lokales Extremum und sei in x_0 differenzierbar. Dann ist $f'(x_0) = 0$.*

Beweis. Wir betrachten wieder die Differenzenquotientenfunktion von f an der Stelle x_0:

$$R(x) = \begin{cases} \frac{f(x)-f(x_0)}{x-x_0} & \text{falls } x \neq x_0 \\ f'(x_0) & \text{sonst.} \end{cases}$$

Nach Satz 12.1 ist R stetig in x_0. Die Funktion f habe etwa ein lokales Maximum in x_0. Dann ist in einer δ-Umgebung von x_0

$$\begin{aligned} R(x) &\geq 0 \text{ für alle } x < x_0, \\ R(x) &\leq 0 \text{ für alle } x > x_0, \end{aligned}$$

also wegen der Stetigkeit von R in x_0

$$R(x_0) = f'(x_0) = 0.$$

□

Bemerkung 12.6 (1) Die Bedingung $f'(x) = 0$ ist nur eine notwendige, aber nicht hinreichende Bedingung für ein lokales Extremum: Für $f(x) = x^3$ gilt z.B. $f'(0) = 0$, diese Funktion besitzt aber in 0 kein lokales Extremum.

(2) Wichtig ist in Satz 12.7, dass dort das *offene* Intervall (a, b) betrachtet wird. Nach Satz 11.5 nimmt jede in einem abgeschlossenen Intervall stetige Funktion $f\colon [a, b] \to \mathbb{R}$ ihr absolutes Maximum und ihr absolutes Minimum an. Liegt ein Extremum jedoch am Rand, so ist dort nicht notwendig $f'(x_0) = 0$, wie man schon an $f\colon [0, 1] \to \mathbb{R}, x \mapsto x$, sieht.

Definition Es sei $f\colon (a, b) \to \mathbb{R}$ eine differenzierbare Funktion. Ein Punkt $x_0 \in (a, b)$ mit $f'(x_0) = 0$ heißt *kritischer Punkt* von f. Der Wert $f(x_0)$ von f an der Stelle x_0 heißt *kritischer Wert*.

Wir werden später eine Bedingung dafür angeben, wann ein kritischer Punkt ein lokales Extremum ist.

Satz 12.8 (Satz von Rolle) *Es sei $a < b$ und $f\colon [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) = f(b)$. Die Funktion f sei in (a, b) differenzierbar. Dann existiert ein $x_0 \in (a, b)$ mit $f'(x_0) = 0$.*

Bemerkung 12.7 Der Satz von Rolle besagt insbesondere, dass zwischen zwei Nullstellen einer differenzierbaren Funktion eine Nullstelle der Ableitung liegt.

Beweis. Weil f stetig auf $[a, b]$ ist, nimmt f Maximum M und Minimum m auf $[a, b]$ an.

1. Fall: Maximum und Minimum werden am Rande angenommen, also $M = m = f(a) = f(b)$, also f konstant, also $f'(x) = 0$ für jedes $x \in (a, b)$.

2. Fall: M oder m werden in (a, b) angenommen. Dann hat f ein Extremum $x_0 \in (a, b)$. Nach Satz 12.7 gilt $f'(x_0) = 0$. □

Satz 12.9 (Mittelwertsatz) *Es sei $a < b$, $f\colon [a, b] \to \mathbb{R}$ stetig und $f\colon (a, b) \to \mathbb{R}$ differenzierbar. Dann gibt es ein $x_0 \in (a, b)$, so dass*

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

Bemerkung 12.8 Geometrisch bedeutet der Mittelwertsatz, dass die Steigung der Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ gleich der Steigung der Tangente an den Graphen von f an einer gewissen Zwischenstelle $(x_0, f(x_0))$ ist.

Beweis des Satzes. Wir definieren eine Hilfsfunktion $g\colon [a,b] \to \mathbb{R}$ durch

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

Die Funktion g erfüllt die Voraussetzungen des Satzes von Rolle, also gibt es ein $x_0 \in (a, b)$, so dass

$$0 = g'(x_0) = f'(x_0) - \frac{f(b) - f(a)}{b - a}.$$

□

Korollar 12.1 *Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige, in (a,b) differenzierbare Funktion. Für die Ableitung gelte*

$$m \le f'(x) \le M \text{ für alle } x \in (a, b)$$

mit gewissen Konstanten $m, M \in \mathbb{R}$. Dann gilt für alle $x, y \in \mathbb{R}$ mit $a \le x \le y \le b$ die Abschätzung

$$m(y - x) \le f(y) - f(x) \le M(y - x).$$

Beweis. Umformulierung des Mittelwertsatzes. □

Korollar 12.2 *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig und in (a,b) differenzierbar mit $f'(x) = 0$ für alle $x \in (a,b)$. Dann ist f konstant.*

Beweis. Dies ist der Fall $m = M = 0$ von Korollar 12.1. □

Als Anwendung geben wir nun eine Charakterisierung der Exponentialfunktion durch ihre Differentialgleichung.

Satz 12.10 *Es sei $c \in \mathbb{R}$ eine Konstante und $f\colon \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion mit*

$$f'(x) = cf(x) \text{ für alle } x \in \mathbb{R}.$$

Es sei $A := f(0)$. Dann gilt

$$f(x) = Ae^{cx} \text{ für alle } x \in \mathbb{R}.$$

Beweis. Wir betrachten die Funktion

$$g(x) := f(x)e^{-cx}.$$

Nach der Produktregel für die Ableitung ist

$$g'(x) = f'(x)e^{-cx} - cf(x)e^{-cx} = (f'(x) - cf(x))e^{-cx} = 0$$

für alle $x \in \mathbb{R}$, also ist g nach dem obigen Korollar konstant. Da $g(0) = f(0) = A$, ist $g(x) = A$ für alle $x \in \mathbb{R}$, also

$$f(x) = Ae^{cx} \text{ für alle } x \in \mathbb{R}.$$

□

Bemerkung 12.9 Speziell erhält man aus Satz 12.10: Die Funktion $\exp\colon \mathbb{R} \to \mathbb{R}$ ist die eindeutig bestimmte differenzierbare Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit $f' = f$ und $f(0) = 1$. Satz 12.10 kann man auch so ausdrücken: Die Menge der differenzierbaren Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ mit $f' = cf$ bildet einen eindimensionalen Unterraum des Vektorraums aller differenzierbaren Funktionen von $\mathbb{R}$ nach $\mathbb{R}$.

Satz 12.11 *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig und in (a,b) differenzierbar. Für alle $x \in (a,b)$ gelte*

$$f'(x) > 0 \qquad (\textit{bzw. } f'(x) < 0).$$

Dann ist f streng monoton wachsend (bzw. fallend).

Beweis. Wir behandeln nur den Fall, dass $f'(x) > 0$ für alle $x \in (a,b)$ ist. Es seien also $x_1, x_2 \in [a,b]$ mit $x_1 < x_2$. Nun ist $(x_1, x_2) \subseteq (a,b)$ und daher erfüllt $f\colon [x_1, x_2] \to \mathbb{R}$ die Voraussetzungen des Mittelwertsatzes. Es gibt also ein $x_0 \in (x_1, x_2)$, so dass

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_0) > 0.$$

Dies ist aber nur möglich, wenn $f(x_2) - f(x_1) > 0$, also $f(x_2) > f(x_1)$.

Der andere Fall ist analog zu behandeln. □

Wir geben nun ein hinreichendes Kriterium für ein Extremum.

Satz 12.12 *Es sei $f\colon (a,b) \to \mathbb{R}$ eine differenzierbare Funktion und $x_0 \in (a,b)$ ein kritischer Punkt von f, d. h. $f'(x_0) = 0$. Außerdem gebe es ein $\delta > 0$, so dass $(x_0 - \delta, x_0 + \delta) \subseteq (a,b)$ und*

$$\begin{aligned} f'(x) > 0 \quad &\textit{für} \quad x_0 - \delta < x < x_0, \\ f'(x) < 0 \quad &\textit{für} \quad x_0 < x < x_0 + \delta. \end{aligned}$$

Dann hat f in x_0 ein isoliertes lokales Maximum.

Bemerkung 12.10 Entsprechend ist der umgekehrte „Vorzeichenwechsel“ von f' hinreichend für ein isoliertes lokales Minimum.

Beweis. Falls $x_0 - \delta < x < x_0$, so ist

$$\frac{f(x_0) - f(x)}{x_0 - x} = f'(\xi) \text{ für ein } \xi \in (x, x_0),$$

also $f(x) < f(x_0)$, da $f'(\xi) > 0$.

Falls $x_0 < x < x_0 + \delta$, so ist

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\eta) \text{ für ein } \eta \in (x_0, x),$$

also auch $f(x) < f(x_0)$, da $f'(\eta) < 0$. □

Satz 12.13 (Zweiter Mittelwertsatz) *Es seien $f, g\colon [a, b] \to \mathbb{R}$ stetig und in (a, b) differenzierbar. Dann gibt es ein $x_0 \in (a, b)$, so dass*

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

Bemerkung 12.11 Für $g(x) = x$ ergibt sich der erste Mittelwertsatz. Der zweite Mittelwertsatz ist vor allen Dingen für die Ermittlung von Grenzwerten nützlich. Zur Quotientenform gelangt man unter zusätzlichen Voraussetzungen: Aus $g(b) - g(a) \neq 0$ folgt

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

Falls zusätzlich f', g' keine gemeinsame Nullstelle in (a, b) haben, folgt

$$g'(x_0) \neq 0 \text{ und } \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}, \; a < x_0 < b.$$

Insbesondere sind beide Voraussetzungen erfüllt, wenn $g'(x) \neq 0$ für alle $x \in (a, b)$ (Beweis als Aufgabe).

Beweis. Wir definieren eine Hilfsfunktion $h\colon [a, b] \to \mathbb{R}$ durch

$$h(x) = (f(b) - f(a))(g(x) - g(a)) - (g(b) - g(a))(f(x) - f(a)).$$

Die Funktion h erfüllt die Voraussetzungen des Satzes von Rolle, da $h(a) = h(b) = 0$. Also gibt es ein $x_0 \in (a, b)$ mit

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

□

Eine Anwendung des zweiten Mittelwertsatzes ist die Regel von de l'Hospital. Wenn $\lim\limits_{x\to a} f(x)$ und $\lim\limits_{x\to a} g(x)$ existieren und $\lim\limits_{x\to a} g(x) \neq 0$ ist, dann gilt

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)}.$$

Wenn aber $\lim\limits_{x\to a} g(x) = 0$ ist, dann wissen wir im Falle $\lim\limits_{x\to a} f(x) \neq 0$, dass $\frac{f(x)}{g(x)}$ an der Stelle a keinen Grenzwert hat. Es bleibt also die Aufgabe, unter der Voraussetzung

$$\lim_{x\to a} f(x) = 0 \text{ und } \lim_{x\to a} g(x) = 0$$

die Frage nach der Existenz des Grenzwerts

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

zu beantworten. Es genügt den Fall des rechtsseitigen Grenzwerts zu beantworten. Für differenzierbares f und g gilt

Satz 12.14 (Regel von de L'Hospital) *Die Funktionen f und g seien für $x > a$ definiert und differenzierbar und es sei $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ sowie $g(x) \neq 0$ und $g'(x) \neq 0$.*

Wenn $\lim\limits_{x\to a} \frac{f'(x)}{g'(x)}$ existiert, dann existiert auch $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ und es gilt:

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

Beweis. Die Funktionen f und g können stetig nach a fortgesetzt werden, so dass wir $f(a) = g(a) = 0$ annehmen können. Auf $[a, x]$ werden wir den zweiten Mittelwertsatz an: danach gibt es ein y mit $a < y < x$, so dass gilt

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

Wir zeigen, dass für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \neq a$ und $\lim_{n\to\infty} x_n = a$ die Folge $\frac{f(x_n)}{g(x_n)}$ konvergiert.

Zu jedem x_n wählen wir nach dem zweiten Mittelwertsatz ein y_n mit $a < y_n < x_n$; es gilt $\lim_{n\to\infty} y_n = a$. Nach Voraussetzung existiert

$$\lim_{n\to\infty} \frac{f'(y_n)}{g'(y_n)}.$$

Aus

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(y_n)}{g'(y_n)}$$

folgt somit die Behauptung. $\square$

13 Taylorentwicklung

Wir definieren zunächst höhere Ableitungen. Es sei $f\colon D \to \mathbb{R}$ eine differenzierbare Funktion. Dann bezeichnet man die Ableitung mit $f'\colon D \to \mathbb{R}$. Ist f' wiederum differenzierbar – das braucht nicht zu sein –, so nennt man

$$f'' := (f')'\colon D \to \mathbb{R}$$

die *zweite Ableitung* von f. Allgemein definiert man die höheren Ableitungen von f rekursiv: Ist $f^{(k)}\colon D \to \mathbb{R}$ erklärt (d.h. $f, f', \ldots, f^{(k-1)}$ sind differenzierbar) und nochmals differenzierbar, so setzt man

$$f^{(k+1)} := (f^{(k)})'.$$

Man beachte, dass bei der Definition von $f^{(k)}(a)$ die Existenz der Ableitungsfunktion $f', f'', \ldots, f^{(k-1)}$ vorausgesetzt wird; wenn nur $f^{(k-1)}(a)$ vorhanden ist, kann die k-te Ableitung an der Stelle $a \in D$ bei unserem Vorgehen nicht definiert werden.

Definition Die Funktion f heißt *k-mal differenzierbar* genau dann, wenn $f', f'', \ldots, f^{(k)}$ existieren.

Nach Satz 12.2 sind dann alle Funktionen $f, f', \ldots, f^{(k-1)}$ stetig, für $f^{(k)}$ braucht das jedoch nicht zuzutreffen. Ist auch $f^{(k)}$ noch eine stetige Funktion, so nennt man f *k-mal stetig differenzierbar*.

Die Menge aller k-mal stetig differenzierbaren Funktion $f\colon D \to \mathbb{R}$ bezeichnen wir mit $C^k(D)$; sinngemäß setzen wir $C^0(D) = C(D)$. Alle $C^k(D)$ sind reelle Vektorräume; es gilt

$$C^0(D) \supseteq C^1(D) \supseteq \cdots \supseteq C^k(D) \supseteq \cdots$$

Existiert die Ableitungsfunktion $f^{(k)}$ für jedes $k \in \mathbb{N}$, so heißt f *unendlich oft differenzierbar*. Die Menge aller dieser Funktionen wird mit $C^\infty(D)$ bezeichnet; auch sie ist ein Vektorraum über $\mathbb{R}$.

Beispiel 13.1 $f(x) = x^\nu$

$$f^{(k)}(x) = \nu \cdot (\nu - 1) \cdot \ldots \cdot (\nu - k + 1)x^{\nu-k}.$$

Es soll nun die Frage untersucht werden, ob und wie eine vorgegebene Funktion f als Potenzreihe geschrieben werden kann:

$$f(x) = \sum_{\nu=0}^{\infty} c_\nu (x-a)^\nu.$$

Die rechte Seite bezeichnet man als die *Potenzreihenentwicklung von f um den Punkt a.*

Wir betrachten zunächst ein Polynom

$$f(x) = \sum_{\nu=0}^{n} c_\nu x^\nu.$$

Dieses kann man schon als seine Potenzreihenentwicklung um den Nullpunkt $a = 0$ auffassen: Setzt man nämlich $c_\nu := 0$ für $\nu > n$, so ist

$$f(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu.$$

Es sei nun a irgendeine reelle Zahl. Dann suchen wir Koeffizienten b_ν, so dass

$$f(x) = \sum_{\nu=0}^{\infty} b_\nu (x-a)^\nu.$$

Die erste Methode besteht darin, in $f(x) = \sum_{\nu=0}^{\infty} c_\nu x^\nu$ für die Variable $x = (x-a)+a$ einzusetzen, mit Hilfe des Binomischen Satzes nach Potenzen von $x - a$ zu sortieren und die Koeffizienten abzulesen.

Die zweite Methode verwendet die Differentialrechnung. Da f ein Polynom ist, ist sicher $b_\nu = 0$ für $\nu > n$. Polynome sind beliebig oft differenzierbar. Setzen wir kurz $g_\nu(x) = (x-a)^\nu$, so ist nach der Kettenregel

$$g_\nu^{(k)}(x) = \nu(\nu-1)(\nu-2)\cdots(\nu-k+1)(x-a)^{\nu-k}.$$

Diese Formel gilt für alle $k \in \mathbb{N}$. Insbesondere folgt

$$g_\nu^{(k)}(x) = 0 \quad \text{falls } k > \nu,$$

da die 0 als Faktor auftritt. Also ist

$$f^{(k)}(x) = \sum_{\nu=0}^{\infty} b_\nu \cdot \nu(\nu-1)(\nu-2)\cdots(\nu-k+1)(x-a)^{\nu-k}.$$

Insbesondere ist

$$f^{(k)}(a) = b_k \cdot k!.$$

Wir erhalten damit:

Satz 13.1 *Die Koeffizienten b_ν der Potenzreihenentwicklung*

$$f(x) = \sum_{\nu=0}^{\infty} b_\nu (x-a)^\nu$$

eines Polynoms $f(x) = \sum_{\nu=0}^{n} c_\nu x^\nu$ lauten

$$b_\nu = \frac{f^{(\nu)}(a)}{\nu!}.$$

Für jedes Polynom höchstens n-ten Grades gilt also die Gleichung

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Es ist zu vermuten, dass der rechts stehende Ausdruck auch für nicht ganzrationale, aber genügend oft differenzierbare Funktionen eine „gute Approximation" in der Umgebung von a darstellt. Entsprechend wie beim Mittelwertsatz stellt sich die Frage nach dem Fehler, der gemacht wird, wenn man $f(x)$ durch diese Approximation ersetzt.

Die Anwendung des Satzes von Rolle auf eine geeignet gewählte Hilfsfunktion zeigt, dass man den Fehler mit Hilfe der $(n+1)$-ten Ableitung an einer geeigneten „Zwischenstelle" darstellen kann.

Satz 13.2 (Taylorsche Formel) *Die Funktion $f\colon [a,b] \to \mathbb{R}$ sei n-mal stetig differenzierbar und in (a,b) existiere auch die $(n+1)$-te Ableitung. Dann gibt es ein $c \in (a,b)$ derart, dass gilt:*

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

Beweis. Wir wenden den Satz von Rolle an auf

$$g(x) = f(b) - f(x) - f'(x)(b-x) - \cdots - \frac{f^{(n)}(x)}{n!}(b-x)^n - m\frac{(b-x)^{n+1}}{(n+1)!}.$$

Es gilt $g(b) = 0$. Die Zahl $m \in \mathbb{R}$ sei so gewählt, dass $g(a) = 0$.
Die Funktion g ist stetig in $[a,b]$ und differenzierbar in (a,b).

$$\begin{aligned} g'(x) &= -f'(x) + f'(x) - f''(x)(b-x) + \cdots \\ &\quad + \frac{f^{(n)}(x)}{(n-1)!}(b-x)^{n-1} - \frac{f^{(n+1)}(x)}{n!}(b-x)^n + m\frac{(b-x)^n}{n!} \\ &= -\frac{f^{(n+1)}(x)}{n!}(b-x)^n + m\frac{(b-x)^n}{n!}. \end{aligned}$$

Nach dem Satz von Rolle gibt es ein $c \in (a,b)$ mit $g'(c) = 0$. Aus

$$0 = -\frac{f^{(n+1)}(c)}{n!}(b-c)^n + m\frac{(b-c)^n}{n!}$$

folgt daher

$$m = f^{(n+1)}(c).$$

Einsetzen von $x = a$ und $m = f^{(n+1)}(c)$ in die Gleichung für $g(x)$ liefert dann die Taylorsche Formel. □

Andere Formulierungen des Taylorschen Satzes:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(z)}{(n+1)!}(x-a)^{n+1} \quad (a < z < x \text{ oder } x < z < a)$$

$$f(x+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}h^k + \frac{f^{(n+1)}(x+\vartheta h)}{(n+1)!}h^{n+1} \qquad (0 < \vartheta < 1)$$

Das Polynom

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

nennt man auch das *Taylorpolynom* n-ter Ordnung der Funktion f an der Stelle a. Den Ausdruck

$$R_n := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

nennt man auch das *Restglied* der Taylorentwicklung von f an der Stelle a von der Ordnung n. R_n hängt von x und a ab und gibt den Fehler an, den man macht, wenn man f durch sein Taylorpolynom n-ter Ordnung ersetzt.

Satz 13.2 liefert für die Zahl R_n:

$$R_n = \frac{f^{(n+1)}(z)}{(n+1)!}(x-a)^{n+1} \quad (a < z < x \text{ oder } x < z < a)$$

Da von der Stelle z nur bekannt ist, dass sie zwischen a und b liegt, kann man R_n i.A. nur abschätzen.

Beispiel 13.2 Logarithmusfunktion $\ln\colon \mathbb{R}_+^* \to \mathbb{R}$. Für $f(x) = \ln x$ berechnet man

$$\begin{aligned} f^{(k)}(x) &= (-1)^{k-1} \cdot (k-1)! x^{-k} \qquad (k \geq 1) \\ f^{(k)}(1) &= (-1)^{k-1} \cdot (k-1)! \\ f(1) &= 0. \end{aligned}$$

Nach der Taylorschen Formel folgt für alle $h > -1$ mit $0 < \vartheta < 1$:

$$\ln(1+h) = \sum_{k=1}^{n} (-1)^{k-1} \frac{h^k}{k} + (-1)^n \frac{1}{(1+\vartheta h)^{n+1}} \frac{h^{n+1}}{n+1}.$$

Wir untersuchen das Restglied

$$R_n = (-1)^n \frac{1}{(1+\vartheta h)^{n+1}} \frac{h^{n+1}}{n+1}, \quad 0 < \vartheta < 1.$$

Für $h > 0$ folgt

$$|R_n| < \frac{h^{n+1}}{n+1}.$$

(Eine bessere Abschätzung ist zunächst nicht möglich, da ϑ beliebig nahe bei 0 liegen kann.)

Für $-1 < h < 0$ folgt

$$|R_n| < \frac{1}{n+1} \frac{|h|^{n+1}}{(1-|h|)^{n+1}},$$

da $1 + \vartheta h > 1 - |h|$ (auch hier ist eine bessere Abschätzung zunächst nicht möglich, da ϑ beliebig nahe bei 1 liegen kann).

Diese Abschätzung für $|R_n|$ kann sehr schlecht sein, z. B. ergibt sich für $h = -\frac{2}{3}$

$$|R_n| \le \frac{2^{n+1}}{n+1},$$

also etwa $|R_9| \le \frac{1024}{10}$. Tatsächlich gilt aber sogar $|R_9| < 0{,}005$, wie ein direkter Vergleich zeigt.

Wir wollen nun unsere Überlegungen auf die Berechnung von $\ln 2$ anwenden. Setzen wir $h = 1$, so finden wir

$$\ln 2 = \sum_{k=1}^{n} (-1)^{k-1} \cdot \frac{1}{k} + R_n$$

mit $|R_n| < \frac{1}{n+1}$. Wegen $\lim_{n\to\infty} R_n = 0$ folgt hieraus das theoretisch interessante, numerisch aber nicht brauchbare Ergebnis

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \cdots$$

Für die numerische Berechnung von Werten der Logarithmusfunktion besser geeignet ist die folgende Taylor-Formel:

Für $-1 < x < 1$ gilt

$$\begin{aligned} \ln \frac{1+x}{1-x} &= 2 \cdot \left(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{2n-1}}{2n-1} \right) \\ &\quad + \frac{x^{2n+1}}{2n+1} \left(\frac{1}{(1+\vartheta x)^{2n+1}} + \frac{1}{(1-\vartheta x)^{2n+1}} \right) \end{aligned}$$

mit $0 < \vartheta < 1$.

Beweis. Für $f(x) = \ln \frac{1+x}{1-x}$ ergibt sich

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

(da $f'(x) = \frac{1-x}{1+x}\left(\frac{1-x+1+x}{(1-x)^2}\right) = \frac{1}{1+x} + \frac{1}{1-x}$, und Induktion) also

$$f^{(k)}(0) = \begin{cases} 0 & \text{falls } k \text{ gerade} \\ 2\cdot(k-1)! & \text{falls } k \text{ ungerade.} \end{cases}$$

Weiter folgt

$$f^{(2n+1)}(\vartheta x) = (2n)!\left(\frac{1}{(1+\vartheta x)^{2n+1}} + \frac{1}{(1-\vartheta x)^{2n+1}}\right)$$

und damit nach Einsetzen die Behauptung. □

Zur Berechnung von $\ln 2$ setzen wir $x = \frac{1}{3}$ und erhalten dann für das Restglied

$$R_{2n} = \frac{1}{(2n+1)3^{2n+1}}\left(\frac{1}{\left(1+\frac{1}{3}\vartheta\right)^{2n+1}} + \frac{1}{\left(1-\frac{1}{3}\vartheta\right)^{2n+1}}\right)$$

mit $0 < \vartheta < 1$.

Wegen $\left(1+\frac{1}{3}\vartheta\right) > 1$ und $\left(1-\frac{1}{3}\vartheta\right) > \frac{2}{3}$ folgt die Abschätzung

$$R_{2n} < \frac{1}{2n+1}\left(\frac{1}{2^{2n+1}} + \frac{1}{3^{2n+1}}\right).$$

Außerdem gilt $R_{2n} > 0$.

Für $n = 5$ ergibt sich

$$0 < R_{10} < 0{,}00005.$$

Die rationale Zahl

$$2\cdot\left(\frac{1}{3} + \frac{1}{3\cdot 3^3} + \frac{1}{5\cdot 3^5} + \frac{1}{7\cdot 3^7} + \frac{1}{9\cdot 3^9}\right) = 0{,}\underbrace{69314}6047\cdots$$

ist deshalb ein Näherungswert, der von $\ln 2$ höchstens um $5\cdot 10^{-5}$ abweicht. Wir erhalten also

$$0{,}\underbrace{6931}46 < \ln 2 < 0{,}\underbrace{6931}97$$

Taschenrechner: $\ln 2 = 0{,}6931472$.

Wir geben noch eine weitere Anwendung des Taylorschen Satzes. Satz 13.2 liefert in einfacher Weise hinreichende Bedingungen für innere lokale Extrema genügend oft stetig differenzierbarer Funktionen.

Satz 13.3 (Minimaxkriterium) *Die Funktion $f\colon (a,b) \to \mathbb{R}$ sei n-mal stetig differenzierbar, und es gelte für ein $x_0 \in (a,b)$:*

$$f'(x_0) = f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0 \text{ und } f^{(n)}(x_0) \neq 0.$$

Ist n ungerade, so hat f in x_0 kein lokales Extremum. Ist n gerade, so hat f in x_0 ein lokales Minimum, falls $f^{(n)}(x_0) > 0$, und ein lokales Maximum, falls $f^{(n)}(x_0) < 0$.

Beweis. Wegen $f'(x_0) = \cdots = f^{(n-1)}(x_0) = 0$ lautet die Taylorsche Formel

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

und $x < z < x_0$ oder $x_0 < z < x$. Nach Voraussetzung ist $f^{(n)}$ stetig und $f^{(n)}(x_0) \neq 0$; deshalb gilt $f^{(n)}(x) \neq 0$ für alle x aus einer geeigneten Umgebung von x_0. Die Zahlen $f^{(n)}(x)$ und $f^{(n)}(x_0)$ haben dort dasselbe Vorzeichen.

Ist nun n ungerade, so hat $f(x) - f(x_0)$ für $x > x_0$ ein anderes Vorzeichen als für $x < x_0$, d. h. x_0 kein Extremum.

Ist aber n gerade, so folgt für alle x aus der Umgebung von x_0

$$f(x) - f(x_0) \geq 0 \text{ falls } f^{(n)}(x_0) > 0$$

bzw.

$$f(x) - f(x_0) \leq 0 \text{ falls } f^{(n)}(x_0) < 0.$$

□

Bemerkung 13.1 Selbst wenn f unendlich oft differenzierbar ist, liefert Satz 13.3 nicht immer eine Entscheidung, da $f^{(k)}(x_0) = 0$ für alle $k \geq 1$ gelten kann.

Übungsaufgabe 13.1 Die Funktionen f und g seien n-mal stetig differenzierbar in (a, b), und es gelte für $x_0 \in (a, b)$:

$$f^{(k)}(x_0) = g^{(k)}(x_0) = 0 \text{ für } k = 0, 1, \ldots, n - 1,$$

jedoch

$$g^{(n)}(x_0) \neq 0.$$

Dann existiert $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ und es gilt

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

Es sei nun $a < b$ und $f \in C^\infty((a, b))$. Für beliebiges $x_0 \in (a, b)$ existieren dann alle Ableitungen und man kann zu f die folgende Potenzreihe bilden:

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

Man nennt sie die *Taylorreihe* von f bezüglich x_0.

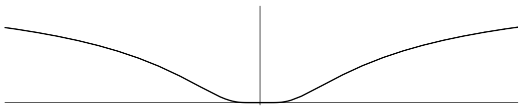
Zwei Fragen müssen geklärt werden:

1. Für welche x konvergiert die Taylorreihe?

2. Wenn die Taylorreihe konvergiert, konvergiert sie dann gegen $f(x)$?

Es ist möglich, dass die Taylorreihe nur für $x = x_0$ konvergiert. Die Antwort auf die zweite Frage ist i.A. nein. Denn z.B. gehört zu

$$f(x) = \begin{cases} \exp\left(-\frac{1}{|x|}\right) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0 \end{cases}$$



und $x_0 = 0$ die Potenzreihe, deren sämtliche Koeffizienten gleich 0 sind, als Taylorreihe. Für $x \neq 0$ konvergiert sie daher nicht gegen $f(x)$.

Eine notwendige und hinreichende Bedingung dafür, dass f durch die zugehörige Taylorreihe dargestellt wird, ist die folgende Bedingung: Wir betrachten die Folge $(R_n)_{n\in\mathbb{N}}$ der Restglieder:

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Dann gilt:

$$f \text{ wird durch Taylorreihe dargestellt } \Leftrightarrow \lim_{n\to\infty} R_n(x) = 0.$$

Mit Hilfe der Darstellung des Restgliedes

$$R_n(x) = \frac{f^{(n+1)}(z)}{(n+1)!}(x - x_0)^{n+1} (x < z < x_0 \text{ bzw. } x_0 < z < x)$$

(„Restglieddarstellung von Lagrange“)

nach Satz 13.2 gelingt in manchen Fällen der Nachweis, dass $\lim_{n\to\infty} R_n(x) = 0$, aber nicht immer. So zeigen die Restgliedabschätzungen für $\ln(1+h)$

$$|R_n(h)| < \frac{h^{n+1}}{n+1} \text{ für } h > 0$$

$$|R_n(h)| < \frac{1}{n+1} \frac{|h|^{n+1}}{(1-|h|)^{n+1}} \text{ für } -1 < h < 0,$$

dass die Gleichung

$$\ln x = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}(x-1)^k$$

in den Intervallen $[1, 2]$ und $[\frac{1}{2}, 1]$ besteht. Tatsächlich ist jedoch diese Potenzreihenentwicklung für ln im Intervall $0 < x \leq 2$ gültig.

Wir wollen deshalb noch andere Darstellungen für das Restglied in der Taylorschen Formel herleiten.

Satz 13.4 *Die Funktion $f\colon [a,b] \to \mathbb{R}$ sei n-mal stetig differenzierbar und in (a,b) existiere auch die $(n+1)$-te Ableitung. Es sei $x_0 \in [a,b]$ und*

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Dann gilt:

(i) Zu jedem $p > 0$ und $x \neq x_0$, $x \in [a,b]$, gibt es ein z mit $x < z < x_0$ bzw. $x_0 < z < x$ und

$$R_n(x) = \frac{f^{(n+1)}(z)}{p \cdot n!} \left(\frac{x-z}{x-x_0}\right)^{n-p+1} \cdot (x-x_0)^{n+1}$$

(„Restglieddarstellung von Schlömilch“)

(ii) Zu jedem $x \in [a,b]$, $x \neq x_0$, gibt es ein z mit $x_0 < z < x$ bzw. $x < z < x_0$ und

$$R_n(x) = \frac{f^{(n+1)}(z)}{n!} \cdot \left(\frac{x-z}{x-x_0}\right)^{n} \cdot (x-x_0)^{n+1}.$$

(„Restglieddarstellung von Cauchy“)

Beweis. Anstelle des Satzes von Rolle beim Beweis von Satz 13.2 verwenden wir den zweiten Mittelwertsatz und können damit eine weitgehend willkürlich wählbare Funktion h ins Spiel bringen.

Es sei $x_0 = a$ und g die im Beweis von Satz 13.2 eingeführte Hilfsfunktion mit $m = 0$. Somit gelten

$$\begin{aligned} g(x) &= f(b) - \sum_{k=0}^{n} \frac{f^{(n)}(x)}{k!}(b-x)^k, \\ g(a) &= R_n(b), \quad g(b) = 0, \\ g'(x) &= -\frac{f^{(n+1)}(x)}{n!}(b-x)^n. \end{aligned}$$

Weiter sei $h\colon [a,b] \to \mathbb{R}$ irgendeine streng monotone differenzierbare Funktion mit $h'(x) \neq 0$ für $x \in (a,b)$. Nach dem zweiten Mittelwertsatz gibt es dann ein $c \in (a,b)$ so dass

$$\frac{g(b)-g(a)}{h(b)-h(a)} = \frac{g'(c)}{h'(c)}.$$

Daraus folgt

$$R_n(b) = \frac{f^{(n+1)}(c)}{n!} \cdot \frac{h(b)-h(a)}{h'(c)} \cdot (b-c)^n.$$

Wir wählen nun speziell $h(x) = (b-x)^p$ mit $p > 0$. Dann folgt

$$R_n(b) = \frac{f^{(n+1)}(c)}{n!} \frac{(b-a)^p}{p} (b-c)^{n-p+1}.$$

Durch Umbenennung folgt daraus

$$R_n(x) = \frac{f^{(n+1)}(z)}{p \cdot n!} \cdot \left(\frac{x-z}{x-x_0}\right)^{n-p+1} \cdot (x-x_0)^{n+1}$$

mit $x < z < x_0$ bzw. $x_0 < z < x$. Dies ist die Restglieddarstellung (i).

Die Restglieddarstellung (ii) erhält man daraus für $p = 1$. □

Bemerkung 13.2 Setzt man $p = n+1$ in der Restglieddarstellung von Schlömilch, so erhält man die obige Restglieddarstellung von Lagrange.

Beispiel 13.3 Mit dem Cauchyschen Restglied läßt sich zeigen, dass die angegebene Potenzreihe den natürlichen Logarithmus im Intervall $0 < x \leq 1$ darstellt:

Für $x_0 = 1$ gilt wegen $f^{(n+1)}(z) = (-1)^n \frac{n!}{z^{n+1}}$:

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

Aus $0 < x < z < 1$ folgt

$$\frac{x-z}{z(x-1)} < 1$$

und damit

$$|R_n(x)| < \frac{1}{z}|x-1|^{n+1} < \frac{1}{x}|x-1|^{n+1}.$$

Also gilt $\lim_{n\to\infty} R_n(x) = 0$ für $0 < x \leq 1$.

14 Funktionenfolgen, gleichmäßige Konvergenz

Für die Einführung des Integrals benötigen wir den Begriff der gleichmäßigen Konvergenz von Funktionenfolgen.

Für jedes $n \in \mathbb{N}$ sei eine Funktion

$$f_n : D \to \mathbb{R}$$

gegeben; dann nennen wir die Folge $(f_n)_{n\in\mathbb{N}}$ eine in D erklärte *Funktionenfolge*.

Für festes $x \in D$ entsteht aus der Funktionenfolge (f_n) durch Einsetzen von x die Zahlenfolge $(f_n(x))$.

Definition (a) Die Funktionenfolge (f_n) heißt *punktweise konvergent* gegen die Grenzfunktion f, wenn für jedes $x \in D$ die Zahlenfolge $(f_n(x))$ gegen den Grenzwert $f(x)$ konvergiert, d.h. wenn gilt:
Zu jedem $x \in D$ und $\varepsilon > 0$ existiert ein $n_0 = n_0(x, \varepsilon) \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt

$$|f_n(x) - f(x)| < \varepsilon.$$

(b) Die Funktionenfolge (f_n) heißt *gleichmäßig konvergent* gegen die Grenzfunktion f, wenn gilt:
Zu jedem $\varepsilon > 0$ existiert ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass für alle $x \in D$ und alle $n \geq n_0$ gilt

$$|f_n(x) - f(x)| < \varepsilon.$$

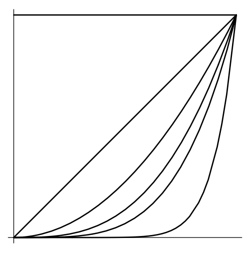
In Quantorenschreibweise:

$$\begin{aligned} &(f_n) \text{ punktweise konvergent gegen } f \\ &\quad :\Leftrightarrow \quad \forall x \in D\ \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0\ |f_n(x) - f(x)| < \varepsilon. \\ &(f_n) \text{ gleichmäßig konvergent gegen } f \\ &\quad :\Leftrightarrow \quad \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall x \in D\ \forall n \geq n_0\ |f_n(x) - f(x)| < \varepsilon. \end{aligned}$$

Der Unterschied ist also der, dass im Fall gleichmäßiger Konvergenz n_0 nur von ε, nicht aber von x, abhängt. Konvergiert eine Funktionenfolge gleichmäßig, so auch punktweise und die Grenzfunktionen stimmen überein. Die Umkehrung gilt jedoch nicht, wie folgende Beispiele zeigen.

Beispiel 14.1 Es sei $D = [0, 1]$ und $f_n(x) = x^n$. Die Funktionenfolge (f_n) ist punktweise, aber nicht gleichmäßig konvergent gegen die Funktion f wobei

$$f(x) = \begin{cases} 0 & \text{für } 0 \leq x < 1, \\ 1 & \text{für } x = 1. \end{cases}$$



Beweis. Für $0 \leq x < 1$ gilt $\lim_{n\to\infty} x^n = 0$. Für $x = 1$ erhält man die konstante Folge $1, 1, \ldots$ mit Grenzwert 1.

Wir bestimmen bei vorgegebenem $\varepsilon > 0$ zu jedem $x \in D$ das kleinst mögliche n_0. Für $x = 0$ ist $n_0 = 1$ und für $x = 1$ ist $n_0 = 0$. Für $0 < x < 1$ ist n_0 die kleinste natürliche Zahl mit

$$n_0 > \frac{\ln \varepsilon}{\ln x}$$

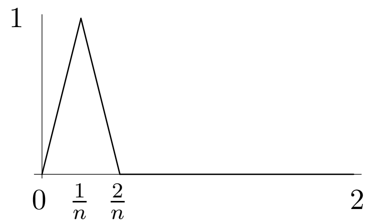
(denn für $n \geq n_0$ gilt:

$$|x^n| < \left|x^{\frac{\ln \varepsilon}{\ln x}}\right| = \left|\exp\left(\frac{\ln \varepsilon}{\ln x} \ln x\right)\right| = \varepsilon).$$

Man sieht daran, dass man bei gegebenem $\varepsilon > 0$ kein von x unabhängiges n_0 finden kann, dass also (f_n) nicht gleichmäßig konvergiert. □

Beispiel 14.2 Es sei $D = [0, 2]$ und (für $n \geq 1$)

$$f_n(x) = \begin{cases} nx & \text{für } 0 \leq x \leq \frac{1}{n}, \\ 2 - nx & \text{für } \frac{1}{n} < x < \frac{2}{n}, \\ 0 & \text{für } \frac{2}{n} \leq x \leq 2. \end{cases}$$



Die Folge (f_n) konvergiert punktweise, aber nicht gleichmäßig gegen die Nullfunktion auf $[0, 2]$.

Beweis. (a) Für $x = 0$ ist $f_n(x) = 0$ für alle $n \in \mathbb{N}$.

(b) Zu jedem $x \in (0, 2]$ existiert ein $n_0 \geq 1$, so dass

$$\frac{2}{n} \leq x \quad \text{für alle } n \geq n_0.$$

Es ist daher für $n \geq n_0$ die dritte Zeile in der Definition von f_n zutreffend, d.h. es gilt $f_n(x) = 0$ für $n \geq n_0$.

(c) Die Funktionenfolge (f_n) konvergiert nicht gleichmäßig gegen 0, denn für kein $n \geq 1$ gilt

$$|f_n(x) - 0| < 1 \text{ für alle } x \in [0, 2].$$

□

Beispiel 14.3 Es sei $D = \mathbb{R}$ und

$$f_n(x) = \frac{n}{n^2 + x^2} \quad (n \geq 1).$$

Die Funktionenfolge (f_n) konvergiert gleichmäßig gegen die Nullfunktion in $\mathbb{R}$. Denn für $n \geq 1$ folgt nämlich:

$$|f_n(x)| = \frac{1}{n}\left(1 + \frac{x^2}{n^2}\right)^{-1} \leq \frac{1}{n}.$$

Anstatt für jedes $x \in D$ das kleinst mögliche n_0 aufzusuchen, kann man auch bei festem n das Supremum der Wertemenge der Funktion $|f_n - f|$ bestimmen. Man erhält dann eine Folge (ε_n) nicht negativer Zahlen mit der Eigenschaft

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

Wenn $(\varepsilon_n)_{n\in\mathbb{N}}$ eine Nullfolge ist, konvergiert (f_n) gleichmäßig, und umgekehrt.

Diese Überlegung führt dazu, den Begriff der *Norm* einer beschränkten Funktion zu erklären:

Definition Es sei $f : D \to \mathbb{R}$ beschränkt. Die nicht negative reelle Zahl

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

heißt *Norm* von f.

Bemerkung 14.1 Wird das Supremum von f in D angenommen, so schreibt man auch

$$\|f\| = \max_{x\in D} |f(x)|.$$

Die Norm besitzt folgende Eigenschaften:

Satz 14.1 *Es seien $f, g : D \to \mathbb{R}$ beschränkt, $\lambda \in \mathbb{R}$. Dann gilt:*

(1) $\|f\| \geq 0$, $\|f\| = 0 \Leftrightarrow f(x) = 0$ *für alle* $x \in D$.

(2) $\|\lambda f\| = |\lambda| \cdot \|f\|$.

(3) $\|f + g\| \leq \|f\| + \|g\|$.

Beweis. Zu (1): $\|f\| \geq 0$ ist klar,

$$\|f\| = 0 \Leftrightarrow \sup_{x\in D} |f(x)| = 0 \Leftrightarrow f(x) = 0 \text{ für alle } x \in D.$$

Zu (2):

$$\begin{aligned} \|\lambda f\| &= \sup_{x\in D} |\lambda f(x)| = \sup_{x\in D} |\lambda||f(x)| \\ &= |\lambda| \sup_{x\in D} |f(x)| = |\lambda|\|f\|. \end{aligned}$$

Zu (3): Für alle $x \in D$ gilt nach Definition:

$$|f(x) \leq \|f\| \quad \text{und} \quad |g(x)| \leq \|g\|.$$

Daraus folgt

$$|(f + g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|,$$

also auch
$$\|f+g\| = \sup_{x\in D} |(f+g)(x)| \le \|f\| + \|g\|.$$

□

Bemerkung 14.2 Ist V ein reeller Vektorraum und $\| \ \| : V \to \mathbb{R}$ eine Abbildung mit den Eigenschaften (1), (2), (3) von Satz 14.1, so nennt man $\| \ \|$ eine *Norm* auf V. Die von uns hier betrachtete Norm ist eine Norm auf dem Vektorraum aller beschränkten Funktionen $f : D \to \mathbb{R}$.

Satz 14.2 *Die Funktionenfolge (f_n) konvergiert genau dann gleichmäßig gegen f, wenn gilt*
$$\lim_{n\to\infty} \|f_n - f\| = 0.$$

Beweis. "$\Rightarrow$": Wenn (f_n) gleichmäßig gegen f konvergiert, so gibt es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass für alle $n \ge n_0$ und $x \in D$ gilt
$$|f_n(x) - f(x)| < \varepsilon,$$
d.h. aber für $n \ge n_0$
$$\|f_n - f\| = \sup_{x\in D} |f_n(x) - f(x)| \le \varepsilon.$$

"$\Leftarrow$": Wenn $\lim_{n\to\infty} \|f_n - f\| = 0$, dann existiert zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass für $n \ge n_0$ gilt:
$$\|f_n - f\| < \varepsilon,$$
d.h. für alle $n \ge n_0$ und für alle $x \in D$ gilt
$$|f_n(x) - f(x)| < \varepsilon.$$

□

Satz 14.3 *Die Funktionenfolge (f_n) konvergiere gleichmäßig gegen f und alle f_n seien stetig in $a \in D$. Dann ist auch f stetig in a.*

Beweis. Wir schreiben $f(x) - f(a)$ in der Form
$$f(x) - f(a) = f(x) - f_n(x) + f_n(x) - f_n(a) + f_n(a) - f(a).$$
Zu vorgegebenem $\varepsilon > 0$ wählen wir ein $n \in \mathbb{N}$, so dass
$$\|f_n - f\| < \frac{\varepsilon}{3}.$$
Die Stetigkeit von f_n an der Stelle a liefert ein $\delta > 0$, so dass für alle $x \in D$ mit $|x - a| < \delta$ gilt:
$$|f_n(x) - f_n(a)| < \frac{\varepsilon}{3}.$$

Für $x \in D$ mit $|x - a| < \delta$ hat man also nach der Dreiecksungleichung

$$|f(x) - f(a)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| < \varepsilon.$$

□

Schließlich führen wir noch den Begriff der *gleichmäßigen Stetigkeit* ein. Wir erinnern zunächst an den Begriff der Stetigkeit:

$f : D \to \mathbb{R}$ ist stetig

$$\Leftrightarrow \quad \forall x_0 \in D \; \forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D \; (|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon).$$

Nun heißt:

$f : D \to \mathbb{R}$ *gleichmäßig stetig*

$$\Leftrightarrow \quad \forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in D \; \forall x_0 \in D \; (|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon).$$

Der Unterschied ist der, dass bei der Stetigkeit δ von ε und x_0 abhängen darf, bei der gleichmäßigen Stetigkeit aber nur von ε. Zur Deutlichkeit fassen wir die Definition noch einmal in Worte:

Definition Eine Funktion $f\colon D \to \mathbb{R}$ heißt *gleichmäßig stetig* genau dann, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in D$ und alle $x_0 \in D$ gilt: Aus $|x - x_0| < \delta$ folgt $|f(x) - f(x_0)| < \varepsilon$.

Hält man x_0 fest, so ergibt sich die Stetigkeit der Funktion f an der Stelle $x_0 \in D$. Aus der gleichmäßigen Stetigkeit folgt also die gewöhnliche Stetigkeit.

Die Umkehrung gilt i.A. nicht, aber:

Satz 14.4 *Jede auf einem abgeschlossenen beschränkten Intervall stetige Funktion $f\colon [a, b] \to \mathbb{R}$ ist dort gleichmäßig stetig.*

Beweis. Angenommen, f sei nicht gleichmäßig stetig. Dann gibt es ein $\varepsilon > 0$, so dass zu jedem $n \in \mathbb{N}$, $n \geq 1$, Punkte $x_n, y_n \in [a, b]$ existieren mit

$$|x_n - y_n| < \frac{1}{n} \text{ und } |f(x_n) - f(y_n)| \geq \varepsilon.$$

Nach dem Satz von Bolzano-Weierstraß besitzt die beschränkte Folge $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$. Für ihren Grenzwert gilt (nach Satz 5.10)

$$\lim_{k\to\infty} x_{n_k} =: c \in [a, b].$$

Wegen $|x_{n_k} - y_{n_k}| < \frac{1}{n_k}$ ist auch

$$\lim_{k\to\infty} y_{n_k} = c.$$

Da f stetig ist, folgt daraus

$$\lim_{k\to\infty} (f(x_{n_k}) - f(y_{n_k})) = f(c) - f(c) = 0.$$

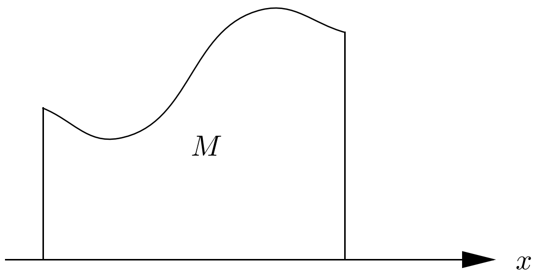
Dies ist ein Widerspruch zu $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon$ für alle $k \in \mathbb{N}$. □

15 Integration

Wir wollen nun den Integralbegriff einführen. Zweck dieser Einführung ist es, Flächen auszumessen:

Wir gehen aus von einer Funktion $f\colon [a,b] \to \mathbb{R}$ mit $f(x) \geq 0$ für alle $x \in [a,b]$. Es interessiert uns die Fläche

$$M = \{(x,y) \mid x \in [a,b], 0 \leq y \leq f(x)\}.$$



Welchen Inhalt hat diese Fläche? Man kann den Flächeninhalt zu bestimmen versuchen, indem man M durch Rechtecke approximiert.

Dazu nun einige Vorbereitungen. Es sei im Folgenden $[a,b]$ immer ein Intervall mit $a < b$.

Definition Eine endliche Menge von Intervallen

$$\mathfrak{Z} = \{J_1, \dots, J_q\}$$

mit den Eigenschaften

$$\bigcup_{k=1}^{q} J_k = [a,b] \text{ und } J_l \cap J_m = \emptyset \text{ für } l \neq m$$

heißt *Zerlegung* von $[a,b]$.

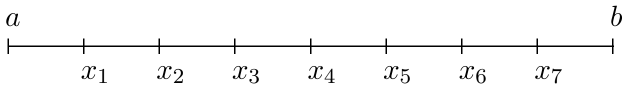
Die Intervalle von $\mathfrak{Z}$ können abgeschlossen, halboffen oder auch offen sein. Es ist zugelassen, dass Intervalle auftreten, die nur aus einem Punkt bestehen. Auch die leere Menge ist als Intervall zugelassen. Also haben wir folgende Intervalle (für $\alpha \leq \beta$): $[\alpha,\beta], (\alpha,\beta], [\alpha,\beta), (\alpha,\beta), \emptyset$.

Beispiel 15.1 (1) Äquidistante Zerlegung mit q Teilintervallen:

$$\text{Teilpunkte: } x_k := a + k\frac{b-a}{q}, \quad k = 0, 1, \dots, q,$$

$$J_k := [x_{k-1}, x_k), \quad k = 1, \dots, q-1,$$

$$J_q := [x_{q-1}, x_q].$$



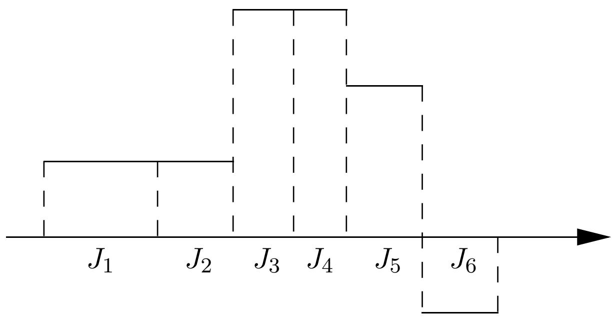
(2) Offene Intervalle und einelementige Intervalle:

$$a = x_0 < x_1 < \cdots < x_n = b,$$

$$J_{2k} = (x_{k-1}, x_k), \quad k = 1, 2, \ldots, n,$$
$$J_{2k+1} = \{x_k\}, \quad k = 0, 1, \ldots, n.$$

Definition Eine Funktion $t\colon [a, b] \to \mathbb{R}$ heißt *Treppenfunktion* über $[a, b]$, wenn es eine Zerlegung $\mathfrak{Z} = \{J_1, J_2, \ldots, J_q\}$ von $[a, b]$ und reelle Zahlen $c_1, \ldots, c_q$ gibt, so dass gilt

$$t(x) = c_k \text{ für jedes } x \in J_k.$$



Eine Treppenfunktion ist also „intervallweise konstant". Die zur Treppenfunktion t gehörende Zerlegung $\mathfrak{Z}$ ist keineswegs eindeutig bestimmt.

Wir wollen nun stetige Funktionen durch Treppenfunktionen approximieren.

Satz 15.1 *Es sei $f\colon [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann gibt es zu jedem $\varepsilon > 0$ eine Treppenfunktion $t\colon [a, b] \to \mathbb{R}$ für die gilt:*

$$||f - t|| < \varepsilon.$$

Beweis. Da f im abgeschlossenen und beschränkten Intervall $[a, b]$ stetig ist, ist f nach Satz 14.4 dort gleichmäßig stetig.

Daher gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$, so dass für alle $x, y \in [a, b]$ mit $|x - y| < \delta$ gilt: $|f(x) - f(y)| < \varepsilon$.

Es sei also $\varepsilon > 0$ gegeben. Dann wählen wir eine Zerlegung $\mathfrak{Z}$ wie in Beispiel 15.1(2), so dass $|x_{k-1} - x_k| < \delta$ für $k = 1, 2, \ldots, n$. Nun definieren wir eine Treppenfunktion $t\colon [a, b] \to \mathbb{R}$ wie folgt:

$$t(x) := f(x_{k-1}) \text{ für } x \in (x_{k-1}, x_k),$$

$$t(x_k) := f(x_k).$$

Wir untersuchen jetzt $|f(x) - t(x)|$ für $x \in [a, b]$.

Falls $x = x_k$ für ein k:

$$|f(x) - t(x)| = |f(x_k) - f(x_k)| = 0.$$

Falls $x \in (x_{k-1}, x_k)$ für ein k:

$$|f(x) - t(x)| = |f(x) - f(x_{k-1})| < \varepsilon,$$

da $|x - x_{k-1}| < x_k - x_{k-1} < \delta$.

Wir haben also für jedes $x \in [a, b]$

$$|f(x) - t(x)| < \varepsilon$$

erhalten. Daraus folgt

$$||f - t|| = \sup_{x \in [a,b]} |f(x) - t(x)| \leq \varepsilon.$$

□

Übungsaufgabe 15.1 Warum darf im Satz das <-Zeichen stehen?

Satz 15.2 *Es sei $f \colon [a, b] \to \mathbb{R}$ stetig. Dann gibt es eine Folge $(t_n)_{n \in \mathbb{N}}$ von Treppenfunktionen, die gleichmäßig gegen f konvergiert.*

Beweis. Setze $\varepsilon := \frac{1}{n}$ und wähle nach Satz 15.1 eine Treppenfunktion t_n mit $||f - t_n|| < \varepsilon = \frac{1}{n}$. Diese so gewählte Folge tut es. □

Um das Integral einer stetigen Funktion zu erklären, gehen wir nun wie folgt vor. Zunächst definieren wir das Integral einer Treppenfunktion. Das Integral einer stetigen Funktion f definieren wir, indem wir f durch Treppenfunktionen gleichmäßig approximieren, deren Integrale das Integral von f approximieren.

Dazu definieren wir als Länge $\lambda(J)$ des Intervalles J mit den Endpunkten α und β, $\alpha \leq \beta$:

$$\lambda(J) := \beta - \alpha.$$

Die Fläche eines Rechtecks ist das Produkt der Seiten.

Die Treppenfunktion $t \colon [a, b] \to \mathbb{R}$ sei definiert mit Hilfe der Zerlegung $\mathfrak{Z} = \{J_1, J_2, \ldots, J_q\}$ von $[a, b]$ und der reellen Zahlen $c_1, c_2, \ldots, c_q$ ($t(x) = c_k$ auf J_k). Dann sei das Integral von t bezüglich der Zerlegung $\mathfrak{Z}$

$$I_{\mathfrak{Z}}(t) := \sum_{k=1}^{q} c_k \cdot \lambda(J_k).$$

Da die Zerlegung $\mathfrak{Z}$ zur Treppenfunktion t nicht eindeutig bestimmt ist, könnte das Integral $I_{\mathfrak{Z}}(t)$ auch von $\mathfrak{Z}$ abhängen.

Satz 15.3 *Das Integral $I_{\mathfrak{Z}}(t)$ hängt nicht von $\mathfrak{Z}$ ab.*

Beweis. Eine Zerlegung $\mathfrak{Z}$ des Intervalls $[a, b]$ ist ja eine Zerlegung in disjunkte Teilintervalle. Diese Zerlegung kann „verfeinert“ werden, indem man die Intervalle J_k der Zerlegung $\mathfrak{Z}$ weiter zerlegt:

$$J_k = \bigcup_{l=1}^{q_k} \tilde{J}_l^{(k)}.$$

Dabei sind $\tilde{J}_l^{(k)}$ ebenfalls disjunkte Intervalle.

Es gilt aber offenbar

$$c_k \cdot \lambda(J_k) = \sum_{l=1}^{q_k} c_k \lambda(\tilde{J}_l^k).$$

Ist also die Zerlegung $\mathfrak{Z}'$ durch Verfeinerung aus der Zerlegung $\mathfrak{Z}$ entstanden, so ist

$$I_{\mathfrak{Z}}(t) = I_{\mathfrak{Z}'}(t);$$

$\mathfrak{Z}'$ ist natürlich auch eine Zerlegung zur Treppenfunktion t.

Zu zwei vorgelegten Zerlegungen $\tilde{\mathfrak{Z}}$ und $\tilde{\tilde{\mathfrak{Z}}}$ zur Treppenfunktion t kann man nun eine gemeinsame Verfeinerung $\mathfrak{Z}$ finden: Ist $\tilde{\mathfrak{Z}} = \{\tilde{J}_1, \dots, \tilde{J}_k\}, \tilde{\tilde{\mathfrak{Z}}} = \{\tilde{\tilde{J}}_1, \dots, \tilde{\tilde{J}}_l\}$, so

$$\mathfrak{Z} = \{\tilde{J}_i \cap \tilde{\tilde{J}}_l \mid i = 1, \dots, k, \; j = 1, \dots, l\}.$$

Also ist $I_{\tilde{\mathfrak{Z}}}(t) = I_{\mathfrak{Z}}(t) = I_{\tilde{\tilde{\mathfrak{Z}}}}(t)$. □

Definition Es sei t eine Treppenfunktion und $\mathfrak{Z}$ irgendeine Zerlegung zu t. Dann ist

$$I(t) := I_{\mathfrak{Z}}(t)$$

das *Integral zur Treppenfunktion t*.

Offenbar ist die Menge aller Treppenfunktionen $t \colon [a, b] \to \mathbb{R}$ ein Vektorraum über $\mathbb{R}$ und es gilt der

Satz 15.4 *Für alle Treppenfunktionen t_1, t_2, t und jede reelle Zahl α gilt:*

(i) $I(t_1 + t_2) = I(t_1) + I(t_2)$,

(ii) $I(\alpha t) = \alpha I(t)$,

d.h. I ist eine lineare Abbildung vom obengenannten Vektorraum nach $\mathbb{R}$.

Beweis. Zu (ii): Ist t eine Treppenfunktion bezüglich $\mathfrak{Z} = \{J_1, \dots , J_q\}$, so auch $\alpha \cdot t$. Nimmt t die Werte c_k an, so αt die Werte $\alpha \cdot c_k$. Also ist

$$I(\alpha t) = \sum_k \alpha c_k \lambda(J_k) = \alpha \sum_k c_k \lambda(J_k) = \alpha \cdot I(t).$$

Zu (i): Wir wählen – wie beim Beweis von Satz 15.3 – eine gemeinsame Zerlegung $\mathfrak{Z}$ für t_1 und t_2. Dann gilt

$$\begin{aligned} I(t_1) + I(t_2) &= \sum_k c_k \lambda(J_k) + \sum_k d_k \lambda(J_k) \\ &= \sum_k (c_k + d_k) \cdot \lambda(J_k) \\ &= I(t_1 + t_2), \end{aligned}$$

wobei t_1 die Werte c_k und t_2 die Werte d_k annimmt. □

Eine Abschätzung:

Satz 15.5 *Für jede Treppenfunktion t gilt:*

$$|I(t)| \le (b - a) \cdot ||t||.$$

Beweis.

$$\begin{aligned} |I(t)| &= \left| \sum_{k=1}^{q} c_k \cdot \lambda(J_k) \right| \\ &\le \sum_{k=1}^{q} |c_k| \cdot \lambda(J_k) \\ &\le \max_k |c_k| \cdot \sum_{k=1}^{q} \lambda(J_k) \\ &= ||t|| \cdot (b - a), \end{aligned}$$

denn es ist

$$||t|| := \sup_{x \in [a,b]} |t(x)| = \max_k |c_k|,$$

da die Funktion t nur die endlich vielen Werte c_k annimmt. □

Satz 15.6 *Wenn die Folge $(t_n)_{n\in\mathbb{N}}$ von Treppenfunktionen gleichmäßig gegen f konvergiert, dann konvergiert die zugehörige Folge $(I(t_n))_{n\in\mathbb{N}}$.*

Beweis. Es wird gezeigt, dass $(I(t_n))_{n\in\mathbb{N}}$ eine Cauchyfolge ist.

Nach Voraussetzung gibt es zu $\varepsilon > 0$ ein n_0, so dass für alle $n \ge n_0$

$$||f - t_n|| < \varepsilon.$$

Also gilt für $n, m \geq n_0$

$$||t_n - t_m|| \leq ||t_n - f|| + ||f - t_m|| < 2\varepsilon.$$

Daher ist

$$\begin{aligned} |I(t_n) - I(t_m)| &= |I(t_n - t_m)| && \text{(Satz 15.4)} \\ &\leq ||t_n - t_m||(b - a) && \text{(Satz 15.5)} \\ &< 2(b - a)\varepsilon. \end{aligned}$$

Da der konstante Faktor $2(b-a)$ nicht stört, ist nachgewiesen, dass $(I(t_n))_{n\in\mathbb{N}}$ eine Cauchyfolge und deshalb konvergent ist. □

Definition Es sei $f\colon [a, b] \to \mathbb{R}$ eine stetige Funktion und $(t_n)_{n\in\mathbb{N}}$ eine Folge von Treppenfunktionen, die gleichmäßig gegen f konvergiert. Wir definieren das Integral $\int_a^b f(x)\,dx$ von f über $[a, b]$ durch

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

Wir müssen noch zeigen, dass diese Definition unabhängig von der gewählten Folge von Treppenfunktionen ist:

Satz 15.7 *Konvergieren* $(t_n)_{n\in\mathbb{N}}$ *und* $(\tilde{t}_n)_{n\in\mathbb{N}}$ *gleichmäßig gegen* f, *so ist*

$$\lim_{n\to\infty} I(t_n) = \lim_{n\to\infty} I(\tilde{t}_n).$$

Beweis. Auch die „gemischte“ Folge von Treppenfunktionen

$$t_1, \tilde{t}_1, t_2, \tilde{t}_2, \ldots$$

konvergiert gleichmäßig gegen f, also konvergiert auch die Folge

$$I(t_1), I(\tilde{t}_1), I(t_2), I(\tilde{t}_2), \ldots ,$$

deren Teilfolgen denselben Grenzwert haben. □

Bemerkung 15.1 Auf die dargestellte Art kann man das Integral für jede Funktion $f\colon [a, b] \to \mathbb{R}$ einführen, die gleichmäßiger Limes einer Folge von Treppenfunktionen sind. Zum Beispiel kann f selbst eine Treppenfunktion sein.

Wir wollen nun zeigen, wie man aufgrund der Definition des Integrals die Integrale einfacher Funktionen berechnen kann.

Definition Jede endliche Folge $x_0, x_1, \dots, x_q$ mit der Eigenschaft

$$a = x_0 < x_1 < \dots < x_q = b$$

heißt eine *Einteilung* E von $[a, b]$.

Die positive Zahl

$$\varphi(E) := \max_{k=1,\dots,q} (x_k - x_{k-1})$$

nennen wir den *Feinheitsgrad* der Einteilung E.

Zu einer Einteilung E von $[a, b]$ haben wir bereits in Beispiel 15.1(2) eine Zerlegung $\mathfrak{Z}$ angegeben.

Zu dieser Zerlegung und einer vorgelegten Funktion f kann man durch beliebige Wahl von Punkten

$$\xi_k \in [x_{k-1}, x_k]$$

eine Treppenfunktion $t\colon [a, b] \to \mathbb{R}$ definieren durch

$$t(x) = \begin{cases} f(\xi_k) & \text{für } x \in (x_{k-1}, x_k),\ k = 1, 2, \dots, q, \\ f(x_k) & \text{für } x = x_k,\ k = 0, 1, \dots, q. \end{cases}$$

Definition Die Zahl

$$S = \sum_{k=1}^{q} f(\xi_k)(x_k - x_{k-1}) = I(t)$$

nennt man eine zur Einteilung E gehörige *Riemannsche Summe* von f.

Die Riemannsche Summe hängt also von E und der Wahl der ξ_k ab.

Spezialfälle:

(a) *Riemannsche Untersumme* von f zur Einteilung E:

$$\sum_{k=1}^{q} m_k (x_k - x_{k-1}),$$

wobei $m_k :=$ Minimum von f auf $[x_{k-1}, x_k]$ (wird wegen der Stetigkeit von f an einer Stelle $\xi_k \in [x_{k-1}, x_k]$ angenommen)

(b) *Riemannsche Obersomme* von f zur Einteilung E:

$$\sum_{k=1}^{q} M_k (x_k - x_{k-1}),$$

mit $M_k :=$ Maximum von f auf $[x_{k-1}, x_k]$.

Satz 15.8 *Es sei* $f\colon [a,b] \to \mathbb{R}$ *eine stetige Funktion und* $(E_n)_{n\in\mathbb{N}}$ *eine Folge von Einteilungen von* $[a,b]$ *mit* $\lim_{n\to\infty} \varphi(E_n) = 0$.

Dann gilt

$$\lim_{n\to\infty} \sum_{k=1}^{q_n} f(\xi_k^{(n)}) \cdot (x_k^{(n)} - x_{k-1}^{(n)}) = \int_a^b f(x)\,dx.$$

Bemerkung 15.2 Jede Folge von Riemannschen Summen zu den Einteilungen E_n konvergiert also gegen das Integral von f, sofern nur die Feinheitsgrade der Einteilungen eine Nullfolge bilden.

Beweis. Wir führen Treppenfunktionen t_n ein durch

$$t_n(x) = \begin{cases} f(\xi_k^{(n)}) & \text{falls } x \in (x_{k-1}^{(n)}, x_k^{(n)}), \\ f(x_k^{(n)}) & \text{falls } x = x_k^{(n)}. \end{cases}$$

Es reicht zu zeigen, dass die Folge (t_n) gleichmäßig gegen f konvergiert, denn die Integrale $I(t_n)$ sind ja gerade die angegebenen Riemannschen Summen.

Da f stetig ist, gibt es zu vorgegebenem $\varepsilon > 0$ ein $\delta > 0$, so dass für alle $x, y \in [a,b]$ mit $|x-y| < \delta$ folgt $|f(x) - f(y)| < \varepsilon$. Sei $n_0 \in \mathbb{N}$, so dass für $n \geq n_0$ gilt: $\varphi(E_n) < \delta$. Dann ist

$$|f(x) - t_n(x)| < \varepsilon$$

für alle $x \in [a,b]$: Denn ist $x = x_k^{(n)}$ für ein k und ein n, so steht links 0. Ist aber $x \in (x_{k-1}^{(n)}, x_k^{(n)})$, so ist ja $t_n(x) = f(\xi_k^{(n)})$ mit $\xi_k^{(n)} \in [x_{k-1}^{(n)}, x_k^{(n)}]$. Nun ist

$$|x - \xi_k^{(n)}| \leq |x_k - x_{k-1}| \leq \varphi(E_n) < \delta,$$

also für $n \geq n_0$

$$|f(x) - t_n(x)| = |f(x) - f(\xi_k^{(n)})| < \varepsilon.$$

□

Wir führen nun anhand von Beispielen die Integration einer stetigen Funktion mit Hilfe von Riemannschen Summen vor.

Beispiel 15.2 (1) $f\colon [0,a] \to \mathbb{R}$, $x \mapsto x^2$. Wir wählen eine äquidistante Einteilung E_n mit dem Feinheitsgrad $\varphi(E_n) = \frac{a}{n}$:

$$x_0^{(n)} = 0, x_1^{(n)} = \frac{a}{n}, \ldots, x_k^{(n)} = \frac{ka}{n}, \ldots, x_n^{(n)} = a.$$

Dann gilt $\lim_{n\to\infty}\varphi(E_n)=0$. Die Obersumme zu E_n ist

$$\begin{aligned} S_n &= \sum_{k=1}^{n}\left(\frac{ka}{n}\right)^2\cdot\frac{a}{n} \\ &= \left(\frac{a}{n}\right)^3\cdot\sum_{k=1}^{n}k^2 = \left(\frac{a}{n}\right)^3\cdot\frac{n(n+1)(2n+1)}{6} \\ &= \frac{a^3}{6}\cdot 1\cdot\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right). \end{aligned}$$

Es gilt

$$\lim_{n\to\infty} S_n = \frac{a^3}{3} = \int_0^a x^2\,dx$$

(2) $f\colon [a,b]\to\mathbb{R}, x\mapsto\frac{1}{x}$, $0<a<b$. Als n-te Einteilung wählen wir

$$a, ac_n, ac_n^2, \dots, ac_n^n, \text{ wobei } c_n := \sqrt[n]{\frac{b}{a}},$$

also $x_k^{(n)} = ac_n^k$. Die Teilpunkte sind also Glieder einer geometrischen Folge, d.h. man hat konstante Quotienten aufeinanderfolgender Glieder. Die Riemannsche Obersumme ist

$$\begin{aligned} S_n &= \sum_{k=1}^{n}\frac{1}{ac_n^{k-1}}(ac_n^k - ac_n^{k-1}) \\ &= n\left(\sqrt[n]{\frac{b}{a}}-1\right). \end{aligned}$$

Was können wir über den Feinheitsgrad $\varphi(E_n)$ sagen? Die Intervallängen sind $ac_n^k - ac_n^{k-1} = ac_n^k\left(1-\frac{1}{c_n}\right)$. Wegen $c_n>1$ ist

$$\varphi(E_n) = ac_n^n - ac_n^{n-1} = b - \frac{b}{c_n} = b\left(1-\frac{1}{c_n}\right),$$

also ist $\lim_{n\to\infty}\varphi(E_n)=0$.

Nach Satz 15.8 ist also

$$\begin{aligned} \int_a^b \frac{1}{x}\,dx &= \lim_{n\to\infty}\frac{\left(\frac{b}{a}\right)^{\frac{1}{n}}-1}{\frac{1}{n}} \\ &= \lim_{x\to 0}\frac{\left(\frac{b}{a}\right)^{x}-1}{x} \\ &= \lim_{x\to 0}\frac{\ln\left(\frac{b}{a}\right)\cdot\left(\frac{b}{a}\right)^x}{1}\text{(de l'Hospital)} \\ &= \ln b - \ln a. \end{aligned}$$

Die Berechnung des Integrals einer stetigen Funktion als Grenzwert von Riemannschen Summen ist im allgemeinen mit großem Rechenaufwand verbunden. Das zeigen die vorherigen Beispiele. Deshalb wird man auch nach anderen Methoden zur Berechnung von Integralen suchen. Dabei ist es vorteilhaft, bestimmte Eigenschaften des Integrals, die im folgenden hergeleitet werden, auszunutzen.

Satz 15.9 (Linearität des Integrals) *Für alle in $[a, b]$ stetigen Funktionen f_1, f_2, f und jede reelle Zahl α gilt:*

(a) $\int\limits_a^b (f_1 + f_2)(x)\,dx = \int\limits_a^b f_1(x)\,dx + \int\limits_a^b f_2(x)\,dx.$

(b) $\int\limits_a^b \alpha \cdot f(x)\,dx = \alpha \int\limits_a^b f(x)\,dx.$

Bemerkung 15.3 Dieser Satz besagt also, dass das Integral eine lineare Abbildung vom Vektorraum $C^0([a, b])$ der auf $[a, b]$ stetigen Funktionen in die reellen Zahlen ist.

Beweis des Satzes. Zu (a): Es sei $(\hat{t}_n)$ bzw. $(\tilde{t}_n)$ eine gleichmäßig gegen f_1 bzw. f_2 konvergierende Folge von Treppenfunktionen. Nach Satz 15.4 gilt für alle $n \in \mathbb{N}$

$$I(\hat{t}_n + \tilde{t}_n) = I(\hat{t}_n) + I(\tilde{t}_n).$$

Wegen der gleichmäßigen Konvergenz der Folgen $(\hat{t}_n)$ und $(\tilde{t}_n)$ folgt daraus

$$\lim_{n\to\infty} I(\hat{t}_n + \tilde{t}_n) = \lim_{n\to\infty} I(\hat{t}_n) + \lim_{n\to\infty} I(\tilde{t}_n) = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx.$$

Läßt sich beweisen, dass die Folge $(\hat{t}_n + \tilde{t}_n)_{n\in\mathbb{N}}$ gleichmäßig gegen $f_1 + f_2$ konvergiert, dann folgt

$$\int_a^b (f_1 + f_2)(x)\,dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx.$$

Also ist noch zu zeigen:

$$\lim_{n\to\infty} ||(f_1 + f_2) - (\hat{t}_n + \tilde{t}_n)|| = 0.$$

Nun sind $(||f_1 - \hat{t}_n||)_{n\in\mathbb{N}}$ und $(||f_2 - \tilde{t}_n||)_{n\in\mathbb{N}}$ nach Voraussetzung Nullfolgen. Aus der Dreiecksungleichung

$$||f_1 + f_2) - (\hat{t}_n + \tilde{t}_n)|| = ||(f_1 - \hat{t_n}) + (f_2 - \tilde{t_n})|| \leq ||f_1 - \hat{t}_n|| + ||f_2 - \tilde{t}_n||$$

für alle $n \in \mathbb{N}$ folgt daher, dass auch $(||(f_1+f_2)-(\hat{t}_n+\tilde{t}_n)||)_{n\in\mathbb{N}}$ eine Nullfolge ist.

Der Beweis zu (b) verläuft entsprechend, wobei man verwendet, dass $||\alpha f - \alpha t_n|| = |\alpha|\,||f - t_n||$ ist. □

Satz 15.10 *Für jede in* $[a, b]$ *stetige Funktion* f *gilt:*

$$\left|\int_a^b f(x)\,dx\right| \leq (b-a)||f||.$$

Beweis. Nach Satz 15.8 gibt es eine gleichmäßig gegen f konvergierende Folge (t_n) von Treppenfunktionen, wobei alle Funktionswerte von t_n auch Funktionswerte von f sind. Es gilt daher

$$||t_n|| \leq ||f|| \text{ für alle } n \in \mathbb{N},$$

folglich auch für jedes n

$$|I(t_n)| \overset{\text{Satz 15.5}}{\leq} ||t_n||(b-a) \leq ||f||(b-a).$$

Also ist auch

$$\left|\lim_{n\to\infty} I(t_n)\right| = \left|\int_a^b f(x)\,dx\right| = \lim_{n\to\infty} |I(t_n)| \leq ||f||(b-a).$$

□

Satz 15.11 *Wenn* f *stetig und* $f(x) \geq 0$ *für alle* $x \in [a, b]$ *ist, dann gilt*

$$\int_a^b f(x)\,dx \geq 0.$$

Beweis. Man bestimme $\int_a^b f(x)\,dx$ als Grenzwert einer Folge von Riemannschen Summen. Diese sind alle ≥ 0. □

Korollar 15.1 (Monotonie des Integrals) *Sind* f_1 *und* f_2 *stetig auf* $[a, b]$ *und ist* $f_1(x) \leq f_2(x)$ *für alle* $x \in [a, b]$, *dann folgt*

$$\int_a^b f_1(x)\,dx \leq \int_a^b f_2(x)\,dx.$$

Beweis.

$$\begin{aligned}
&f_2(x) - f_1(x) \geq 0 \text{ für alle } x \in [a, b] \\
\Rightarrow &\int_a^b (f_2(x) - f_1(x))\,dx \geq 0 \text{ (Satz 15.11)} \\
\Rightarrow &\int_a^b f_2(x)\,dx - \int_a^b f_1(x)\,dx \geq 0 \text{ (Linearität)} \\
\Rightarrow &\int_a^b f_1(x)\,dx \leq \int_a^b f_2(x)\,dx.
\end{aligned}$$

□

Satz 15.12 *Für jede in $[a,b]$ stetige Funktion f gilt:*

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

Beweis. Aus der Stetigkeit von f folgt die Stetigkeit von $(-f)$ und $|f|$. Es gilt

$$|f(x)| \geq f(x) \text{ und } |f(x)| \geq -f(x) \text{ für } x \in [a,b].$$

Wegen der Monotonie und Linearität des Integrals gilt auch

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

also

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

Ein zweiter Beweis mit Riemannschen Summen und Dreiecksungleichung ist möglich. □

Satz 15.13 *Wenn m das Minimum und M das Maximum der stetigen Funktion f in $[a,b]$ ist, dann gilt:*

$$(b-a)\cdot m \leq \int_a^b f(x)\,dx \leq M\cdot(b-a)$$

Beweis. Aus

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

folgt wegen der Monotonie

$$m(b-a) = \int_a^b m\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b M\,dx = M(b-a).$$

□

Korollar 15.2 (Mittelwertsatz der Integralrechnung) *Für jede in $[a,b]$ stetige Funktion f existiert ein $\xi \in [a,b]$ mit*

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

Beweis. Aufgrund von Satz 15.13 gilt

$$m \leq \frac{1}{b-a}\int_a^b f(x)\,dx \leq M.$$

Weil m das Minimum und M das Maximum von f in $[a, b]$ ist, existiert nach dem Zwischenwertsatz ein $\xi \in [a, b]$ mit

$$f(\xi) = \frac{1}{b-a} \int_a^b f(x)\, dx.$$

$\square$

Im Folgenden soll der Zusammenhang zum Riemannschen Integral kurz angedeutet werden. Zu jeder Einteilung $a = x_0 < x_1 < \cdots < x_q = b$ haben wir die Obersumme und Untersumme von f erklärt:

$$\sum_{k=1}^{q} M_k(x_k - x_{k-1}) \text{ und } \sum_{k=1}^{q} m_k(x_k - x_{k-1}).$$

Nimmt man nun eine Folge von Einteilungen E_n mit $\lim_{n\to\infty} \varphi(E_n) = 0$, so dass E_n eine Verfeinerung von E_{n-1} ist für jedes n, so bilden die Untersummen eine monoton steigende, die Obersummen eine monoton fallende Folge: Denn geht man von der gegebenen Einteilung zu einer neuen über, indem man einen Punkt u etwa zwischen x_{k-1} und x_k einfügt, so gilt

$$\min_{x\in[x_{k-1},x_k]} f(x) \le \min_{x\in[x_{k-1},u]} f(x),\ \min_{x\in[u,x_k]} f(x),$$

$$\max_{x\in[x_{k-1},u]} f(x),\ \max_{x\in[u,x_k]} f(x) \le \max_{x\in[x_{k-1},x_k]} f(x),$$

also

$$m_k(x_k - x_{k-1}) \le m_k^{(1)}(u - x_{k-1}) + m_k^{(2)}(x_k - u),$$

$$M_k^{(1)}(u - x_{k-1}) + M_k^{(2)}(x_k - u) \le M_k(x_k - x_{k-1}).$$

Beide Folgen konvergieren gegen das Integral von f (Satz 15.8).

Es folgt

$$\text{Untersumme } \le \int_a^b f(x)\, dx \le \text{ Obersumme.}$$

Insbesondere gilt

$$\begin{aligned} \int_a^b f(x)\, dx &= \text{ Supremum aller Untersummen bzgl. aller Einteilungen} \\ &= \text{ Infimum aller Obersummen bzgl. aller Einteilungen.} \end{aligned}$$

Die rechte Seite ist genau die Definition des *Riemannschen Integrals von f*.

Man beachte insbesondere, dass $m(b-a)$ und $M(b-a)$ die primitivsten denkbaren Unter- und Obersummen sind.

Wir wollen die Integraldefinition noch etwas erweitern.

Definition Es sei $f\colon [a,b] \to \mathbb{R}$ stetig und $c, d \in [a,b]$. Man setzt

$$\int_c^d f(x)\,dx := \begin{cases} \int_c^d f|_{[c,d]}(x)\,dx & \text{falls } c < d, \\ 0 & \text{falls } c = d, \\ -\int_d^c f(x)\,dx & \text{falls } d < c. \end{cases}$$

Im Fall $c < d$ bedeutet dies lediglich, dass man f auf $[c,d]$ einschränkt und die alten Definitionen verwendete. Die Definition im Fall $d < c$ ist plausibel: Eine Seite mit umgedrehter Orientierung wird negativ gezählt.

Satz 15.14 (Additivität bzgl. des Intervalls) *Wenn $f\colon [a,b] \to \mathbb{R}$ stetig ist, dann gilt für alle $c, d, e \in [a,b]$*

$$\int_c^d f(x)\,dx + \int_d^e f(x)\,dx = \int_c^e f(x)\,dx.$$

Beweis. 1. Fall: $c < d < e$:

Es sei $(t_n^{(1)})$ eine auf $[c,d]$ gleichmäßig gegen f konvergierende Folge von Treppenfunktionen, $(t_n^{(2)})$ entsprechend auf $[d,e]$.

Die Funktion $t_n\colon [c,e] \to \mathbb{R}$ mit

$$t_n(x) = \begin{cases} t_n^{(1)}(x) & \text{für } x \in [c,d], \\ t_n^{(2)}(x) & \text{für } x \in (d,e], \end{cases}$$

ist Treppenfunktion auf $[c,e]$, und $(t_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen f. Nun gilt

$$I(t_n) = I(t_n^{(1)}) + I(t_n^{(2)}) \text{ für jedes } n \in \mathbb{N},$$

für die Grenzwerte also

$$\int_c^e f(x)\,dx = \int_c^d f(x)\,dx + \int_d^e f(x)\,dx.$$

2. Fall: $d < c < e$:

Dann ist

$$\int_d^c f + \int_c^e f = \int_d^e f \text{ nach Fall 1,}$$

also

$$\int_c^e f = \int_d^e f - \int_d^c f = \int_d^e f + \int_c^d f$$

nach Definition.

Analog zeigt man die anderen Fälle, bei denen alle drei Zahlen c, d, e voneinander verschieden sind. Sind nicht alle drei Zahlen verschieden, so ist nach Definition eins der Integrale gleich 0 und die Behauptung folgt. $\square$

Wir wollen nun den Zusammenhang zwischen der Differential- und Integralrechnung darstellen und gleichzeitig eine weittragende Methode zur Berechnung von Integralen angeben. Zunächst führen wir den Begriff der Stammfunktion ein.

Ist das Integral $\int_a^b f(x)\,dx$ der stetigen Funktion f zu berechnen, so kann man wie folgt vorgehen, wenn F ein Funktion ist, deren Ableitung f ist ($F' = f$):

Man wählt eine beliebige Einteilung $x_0 = a, x_1, \dots, x_q = b$. Dann ist nach dem Mittelwertsatz der Differentialrechnung

$$F(x_k) - F(x_{k-1}) = f(\xi_k)(x_k - x_{k-1})$$

für eine Zwischenstelle $\xi_k \in [x_{k-1}, x_k]$. Die zugehörige Riemannsche Summe ist also bei dieser Wahl der ξ_k

$$\sum_{k=1}^{q} f(\xi_k)(x_k - x_{k-1}) = \sum_{k=1}^{q} F(x_k) - F(x_{k-1}) = F(b) - F(a).$$

Nimmt man nun irgendeine Folge von Einteilungen E_n, für die die Feinheit $\varphi(E_n)$ gegen 0 konvergiert, und wählt man Riemannsche Summen auf die oben geschilderte Weise mit dem Mittelwertsatz für F, so ist nach Satz 15.8

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} [F(b) - F(a)] = F(b) - F(a).$$

Man kann also das Integral $\int_a^b f(x)\,dx$ einer stetigen Funktion berechnen, indem man eine Funktion F sucht, deren Ableitung f ist. Eine solche Funktion heißt Stammfunktion von f.

Dabei soll $F' = f$ auf dem ganzen Intervall $[a, b]$ gelten. Wir müssen also auch erklären, wann eine Funktion $g\colon [a, b] \to \mathbb{R}$ differenzierbar ist. Deswegen müssen wir unsere Definition der Differenzierbarkeit noch etwas erweitern: Es sei $g\colon [a, b] \to \mathbb{R}$ eine Funktion. Die Funktion g heißt *rechtsseitig differenzierbar* an der Stelle a dann und nur dann, wenn

$$\lim_{\substack{h\to 0\\ h>0}} \frac{g(a+h) - g(a)}{h}$$

existiert.

Die Funktion g heißt *linksseitig differenzierbar* an der Stelle b dann und nur dann, wenn

$$\lim_{\substack{h\to 0\\ h<0}} \frac{g(b+h) - g(b)}{h}$$

existiert. (Man bildet nur einseitige Grenzwerte, damit die Quotienten überhaupt definiert sind.)

Eine Funktion $g\colon [a,b] \to \mathbb{R}$ heißt *differenzierbar* genau dann, wenn g differenzierbar in (a,b), rechtsseitig differenzierbar in a und linksseitig differenzierbar in b ist. Ist obendrein die Ableitung stetig, so heißt g *stetig differenzierbar* in $[a,b]$.

Definition Eine in einem Intervall J differenzierbare Funktion F heißt *Stammfunktion von f auf J*, wenn $F'(x) = f(x)$ für alle $x \in J$ gilt.

Dabei kann J wie gesagt ein beliebiges Intervall sein.

Bemerkung 15.4 Ist F Stammfunktion von f auf J, und ist k eine konstante Funktion auf J, dann ist auch $F + k$ eine Stammfunktion von f auf J.

Satz 15.15 *Sind F und G Stammfunktionen von f auf J, so ist $F - G$ eine auf J konstante Funktion.*

Beweis. Auf J gilt für die Ableitung der Funktion $F - G$

$$(F - G)' = f - f = 0 \text{ (Nullfunktion).}$$

Nach Korollar 12.2 folgt, dass $F - G$ konstant auf J ist. □

Aus Satz 15.15 folgt, dass man aus einer Stammfunktion von f auf J jede andere durch Addition einer konstanten Funktion erhält.

Wir wollen nun zeigen, dass jede stetige Funktion eine Stammfunktion besitzt. Das ist der berühmte Hauptsatz der Differential- und Integralrechnung.

Satz 15.16 (Hauptsatz der Differential- und Integralrechnung) *Für jede stetige Funktion $f\colon [a,b] \to \mathbb{R}$ ist die auf $[a,b]$ durch*

$$F(x) = \int_c^x f(t)\,dt \ \textit{mit } c \in [a,b]$$

definierte Funktion F eine Stammfunktion von f auf $[a,b]$.

Beweis. Für $h \neq 0$ gilt

$$\begin{aligned}\frac{F(x+h) - F(x)}{h} &= \frac{1}{h}\left(\int_c^{x+h} f(t)\,dt - \int_c^x f(t)\,dt\right)\\ &= \frac{1}{h}\int_x^{x+h} f(t)\,dt \text{ (nach Satz 15.14).}\end{aligned}$$

Ist nun $h > 0$, dann gilt nach dem Mittelwertsatz

$$\frac{1}{h}\int_x^{x+h} f(t)\,dt = f(\xi_h),$$

wobei ξ_h eine Zwischenstelle aus $[x, x+h]$ ist. Im Fall $h < 0$ ist

$$\frac{1}{h}\int_x^{x+h} f(t)\,dt = -\frac{1}{h}\int_{x+h}^{x} f(t)\,dt \text{ (nach Definition)} = f(\xi_h),$$

wobei $\xi_h \in [x+h, x]$ ist (ebenfalls nach dem Mittelwertsatz der Integralrechnung).

Durchläuft also h ein Nullfolge, so durchlaufen die zugehörigen ξ_h eine gegen x konvergente Folge; wegen der Stetigkeit von f durchlaufen die $f(\xi_h)$ dann eine gegen $f(x)$ konvergente Folge. Also ist

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,dt = f(x).$$

□

Aus dem Hauptsatz ergibt sich unmittelbar der folgende Satz, der auch oft Hauptsatz genannt sind.

Satz 15.17 *Die Funktion f sei auf $[a, b]$ stetig. Dann gilt für* jede *Stammfunktion F von f auf $[a, b]$:*

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

Beweis. Nach dem Hauptsatz ist

$$x \mapsto \int_a^x f(t)\,dt$$

eine Stammfunktion von f auf dem Intervall $[a, b]$. Wenn F irgendeine Stammfunktion von f ist, muss es also eine Konstante k geben mit

$$F(x) = k + \int_a^x f(t)\,dt.$$

Durch Einsetzen von $x = a$ erhält man $F(a) = k$, so dass für alle $x \in [a, b]$ die Gleichung

$$\int_a^x f(t)\,dt = F(x) - F(a)$$

bestehen muss, insbesondere auch für $x = b$. □

Bemerkung 15.5 Statt $F(b) - F(a)$ schreibt man auch

$$[F(x)]_a^b \text{ oder } F(x)|_a^b.$$

Eine Stammfunktion F von f nennt man auch „*unbestimmtes Integral*“ und schreibt dafür

$$\int f(x)\,dx.$$

Die Zahl $\int_a^b f(x)\,dx$ nennt man dann im Gegensatz dazu „*bestimmtes Integral*“.

Mit dem Hauptsatz haben wir nun ein Mittel in der Hand, um die Integrale vieler elementarer Funktionen bequem zu berechnen. Aus jedem in der Differentialrechnung gewonnenen Resultat (sofern die Ableitung stetig ist) ergibt sich eine entsprechende Aussage über Integrale. Man kann Tabellen von Funktionen und ihren Stammfunktionen aufstellen, z.B. findet man solche Tabellen in

Bronstein-Semendjajew: Taschenbuch der Mathematik.

Aufgrund des Hauptsatzes der Differential- und Integralrechnung lassen sich Aussagen der Differentialrechnung in Aussagen der Integralrechnung übertragen. So erhält man auch weitere Methoden zur Berechnung von Integralen.

Für in $[a, b]$ differenzierbare Funktionen u und v gilt die Produktregel

$$(uv)' = u'v + uv'.$$

Es ist also uv eine Stammfunktion von $(u'v + uv')$ auf $[a, b]$. Falls u' und v' stetig sind, gilt nach Satz 15.17:

$$[u(x) \cdot v(x)]_a^b = \int_a^b u'(x)v(x)\,dx + \int_a^b u(x)v'(x)\,dx.$$

Hieraus gewinnt man die folgende, auch Produktintegration oder partielle Integration genannte Regel:

Satz 15.18 (Partielle Integration) *Sind u und v stetig differenzierbare Funktionen über $[a, b]$, so gilt*

$$\int_a^b u(x)v'(x)\,dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)\,dx.$$

Beispiel 15.3 a) $\int_0^\pi x \cdot \sin x\,dx$:

Wähle $u(x) = x, v'(x) = \sin x, u'(x) = 1, v(x) = -\cos x$. Nach Satz 15.18

$$\int_0^\pi x \cdot \sin x\,dx = [x(-\cos x)]_0^\pi - \int_0^\pi 1 \cdot (-\cos x)\,dx = \pi + [\sin x]_0^\pi = \pi.$$

b) $\int_0^{\frac{\pi}{2}} \cos^2 x\,dx$:

Wähle $u(x) = \cos x$, $v'(x) = \cos x$, $u'(x) = -\sin x$, $v(x) = \sin x$.

Nach Satz 15.18

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \cos^2 x\,dx &= [\cos x \sin x]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} (-\sin x)\sin x\,dx \\ &= \int_0^{\frac{\pi}{2}} \sin^2 x\,dx \\ &= \int_0^{\frac{\pi}{2}} (1 - \cos^2 x)\,dx \\ &= \int_0^{\frac{\pi}{2}} dx - \int_0^{\frac{\pi}{2}} \cos^2\,dx.\end{aligned}$$

Also

$$2 \cdot \int_0^{\frac{\pi}{2}} \cos^2 x\, dx = \int_0^{\frac{\pi}{2}} dx = \frac{\pi}{2}$$

und damit

$$\int_0^{\frac{\pi}{2}} \cos^2 x\, dx = \int_0^{\frac{\pi}{2}} \sin^2 x\, dx = \frac{\pi}{4}.$$

c)

$$\begin{aligned}\int_a^b \ln x\, dx &= \int_a^b \ln x \cdot 1\, dx \text{ mit } \ln x = u, 1 = v', v := x\\ &= [x \cdot \ln x]_a^b - \int_a^b \frac{x}{x}\, dx\\ &= [x \cdot \ln x - x]_a^b, \quad a, b > 0.\end{aligned}$$

Die Kettenregel der Differentialrechnung überträgt sich in die Integralrechnung wie folgt:

Satz 15.19 (Substitutionsregel) *Es sei* $f\colon [a,b] \to \mathbb{R}$ *stetig, und* $\varphi\colon [\alpha,\beta] \to [a,b] \subset \mathbb{R}$ *stetig differenzierbar. Dann gilt*

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\, dx = \int_\alpha^\beta f(\varphi(u)) \cdot \varphi'(u)\, du.$$

Bemerkung 15.6 Diese Regel kann man sich wie folgt merken: $\int_?^? f(x)\, dx$ ist zu bestimmen. In welchen Grenzen? Wir setzen $x = \varphi(u)$, dann „ist“ $dx = \varphi'(u)\, du$ (Kettenregel). Bei dieser Übersetzung geht der linke Integrand in den rechten Integrand über. Hat man dann die rechten Grenzen, läuft also u von α bis β, so läuft $x = \varphi(u)$ von $\varphi(\alpha)$ bis $\varphi(\beta)$.

Beweis von Satz 15.19. Da f stetig ist, gibt es zu f eine Stammfunktion F. Dies ist eine differenzierbare Funktion auf $[a,b]$, so dass die Hintereinanderschaltung $F \circ \varphi$ erklärt und differenzierbar auf $[\alpha,\beta]$ ist. Nach der Kettenregel gilt nun für alle $u \in [\alpha,\beta]$

$$(F \circ \varphi)'(u) = F'(\varphi(u)) \cdot \varphi'(u).$$

Wegen $F' = f$ ergibt sich

$$\begin{aligned}\int_\alpha^\beta f(\varphi(u)) \cdot \varphi'(u)\, du &= \int_\alpha^\beta F(\varphi(u))'\, du \text{ (Kettenregel)}\\ &= F(\varphi(\beta)) - F(\varphi(\alpha)) \text{ (Satz 15.17)}\\ &= \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\, dx \text{ (erneut Satz 15.17).}\end{aligned}$$

□

Bemerkung 15.7 Die Umkehrbarkeit von φ ist für die Gültigkeit der Substitutionsregel nicht erforderlich. Wenn zusätzlich vorausgesetzt wird, dass φ^{-1} existiert, kann man die Substitutionsregel auch in der Form

$$\int_a^b f(x)\,dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(u))\cdot\varphi'(u)\,du$$

schreiben.

Beispiel 15.4 (a) $\int_a^b (pu+q)^n\,du$ mit $p \neq 0$ und $n \neq -1$.
Ansatz: $\varphi(u) = pu+q, f(x) = x^n, \varphi'(u) = p$.
Also gilt nach der Substitutionsregel

$$\begin{aligned}\int_a^b (pu+q)^n\,du &= \frac{1}{p}\int_a^b \underbrace{(pu+q)^n}_{f(\varphi(u))}\underbrace{p}_{\varphi'(u)}\,du \\ &= \frac{1}{p}\int_{pa+q}^{pb+q} \underbrace{x^n}_{f(x)}\,dx \\ &= \frac{1}{p}\left[\frac{x^{n+1}}{n+1}\right]_{pa+q}^{pb+q} \\ &= \frac{1}{p}\left[\frac{(pu+q)^{n+1}}{n+1}\right]_a^b.\end{aligned}$$

(b) $\int_0^r \sqrt{r^2-x^2}\,dx$ (Fläche eines Viertelkreises mit dem Radius r).
Wir setzen $x = \varphi(u) = r\cdot\sin u$, also $dx = r\cos u\,du$. Deshalb

$$\begin{aligned}\int_0^r \sqrt{r^2-x^2}\,dx &= \int_0^{\frac{\pi}{2}} \sqrt{r^2-r^2\sin^2 u}\cdot r\cos u\,du \\ &= \int_0^{\frac{\pi}{2}} r^2\cos^2 u\,du \\ &= r^2\int_0^{\frac{\pi}{2}} \cos^2 u\,du \\ &= r^2\frac{\pi}{4} \text{ (nach Beispiel (b) zur partiellen Integration).}\end{aligned}$$

Nun versuchen wir $\int_0^x \sqrt{r^2-\xi^2}\,d\xi$ zu bestimmen, also eine Stammfunktion zu $\sqrt{r^2-x^2}$ für $|x| \leq r$.
Wir setzen $\xi = r\sin u$, also $d\xi = r\cos u\,du$. Deshalb

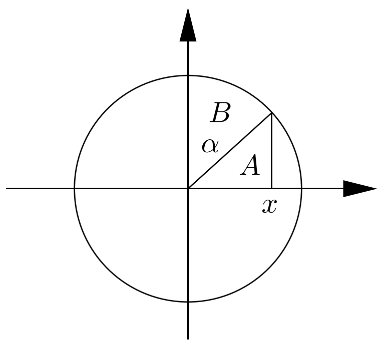
$$\int_0^x \sqrt{r^2-\xi^2}d\xi = \int_0^v \sqrt{r^2-r^2\sin^2 u}\cdot r\cos u\,du = r^2\int_0^v \cos^2 u\,du,$$

$x = r\sin v$, also $v = \arcsin\frac{x}{r}$. Die Definitionsbereiche sehen dabei so aus:

$x \in [-r, r]$, also $v \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Dann ist aber

$$\begin{aligned}
\int_0^x \sqrt{r^2 - \xi^2} d\xi &= r^2 \int_0^v \cos u \cos u \, du \\
&= r^2 [\sin u \cos u]_0^v + r^2 \int_0^v \sin^2 u \, du \\
&= r^2 \sin v \cos v + r^2 \int_0^v (1 - \cos^2 u) \, du \\
&= r^2 \sin v \cos v + r^2 v - r^2 \int_0^v \cos^2 u \, du \\
&= \frac{1}{2} r^2 (\sin v \cos v + v) \\
&= \frac{1}{2} r^2 \left(\frac{x}{r} \sqrt{1 - \frac{x^2}{r^2}} + \arcsin \frac{x}{r} \right) \\
&= \frac{1}{2} x \sqrt{r^2 - x^2} + \frac{1}{2} r^2 \arcsin \frac{x}{r}.
\end{aligned}$$

Dieses Integral können wir auch als Flächeninhalt interpretieren:



Der Term $\frac{1}{2} x \sqrt{r^2 - x^2}$ ist der Flächeninhalt des Dreiecks (A) und der Term $\frac{1}{2} r^2 \cos \arcsin \frac{x}{r}$ der Flächeninhalt des Kreissektors (B). Letzterer ist nämlich gleich Radius mal halber Bogenlänge, die Bogenlänge aber ist gleich Radius mal Winkel α; da $\sin \alpha = \frac{x}{r}$, ist $\alpha = \arcsin \frac{x}{r}$.

Durch den Hauptsatz der Differential- und Integralrechnung wird die Berechnung von Integralen für stetige Integranden zurückgeführt auf die Bestimmung von Stammfunktionen. Für die elementaren Funktionen ist zwar die Existenz von Stammfunktionen (aufgrund des Hauptsatzes) gesichert, doch gehören diese Stammfunktionen selber in vielen Flächen nicht mehr zu den elementaren Funktionen, wie z. B. das „elliptische Integral“ $\int_0^x \sqrt{1 - k^2 \sin^2 t} \, dt$ (mit $0 < k < 1$).

Wir geben nun noch eine Anwendung der Regel der partiellen Integration.

Dazu sei J ein offenes Intervall und $f\colon J \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion. Ferner sei $a, b \in J, n \geq 1$. Dann ist offenbar

$$f(b) = f(a) + \int_a^b f'(x)\,dx.$$

Das sieht schon wie die Taylorsche Formel aus, auf die wir hinaus wollen. Der Punkt a spielt dabei die Rolle des Entwicklungspunktes, wenn $f(b)$ zu bestimmen ist. Das Integral wird das Restglied sein, das wir nun mittels partieller Integration weiter umformen:

$$\begin{aligned}
&\int_a^b \underbrace{f'(x)}_{u(x)}\underbrace{(-1)}_{v'(x)}\,dx \qquad (v(x) := b - x)\\
&= \quad [f'(x)(b-x)]_a^b - \int_a^b f''(x)(b-x)\,dx\\
&= \quad -f'(a)(b-a) - \int_a^b f''(x)(b-x)\,dx.
\end{aligned}$$

Also ist

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

Allgemein gilt nun

Satz 15.20 (Taylor-Formel mit Restglied in Integraldarstellung)
Es sei $f\colon J \to \mathbb{R}$ $(n+1)$-mal stetig differenzierbar, und $a, b \in J$. Dann ist

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

Beweis. Wir führen den Beweis durch Induktion nach n, den Induktionsanfang haben wir bereits gemacht. Induktionsschritt von $n-1$ nach n:

$$\begin{aligned}
&\int_a^b \underbrace{f^{(n)}(x)}_{u(x)}\underbrace{\frac{(b-x)^{n-1}}{(n-1)!}}_{v'(x)}\,dx \qquad (v(x) := -(b-x)^n/n!)\\
&= \quad -\left[f^{(n)}(x)\frac{(b-x)^n}{n!}\right]_a^b + \int_a^b f^{(n+1)}(x)\frac{(b-x)^n}{n!}\,dx\\
&= \quad f^{(n)}(a)\frac{(b-a)^n}{n!} + \int_a^b f^{(n+1)}(x)\frac{(b-x)^n}{n!}\,dx.
\end{aligned}$$

Nach Induktionsannahme ist die Behauptung richtig für $n-1$, also ist sie auch richtig für n. $\square$

Damit haben wir eine vierte Restglieddarstellung hergeleitet.

16 Funktionenfolgen und Potenzreihen

Wir kehren nun zu der Untersuchung von Funktionenfolgen zurück und studieren, wann Integration und Differentation von Funktionen mit der Limesbildung vertauschbar sind. Wir behandeln dann systematisch Potenzreihen.

Satz 16.1 *Es sei* $f_n \colon [a,b] \to \mathbb{R}$, $n \in \mathbb{N}$, *eine Folge stetiger Funktionen, die gleichmäßig gegen* $f \colon [a,b] \to \mathbb{R}$ *konvergiert. Dann gilt*

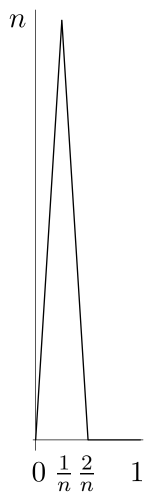
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx.$$

Beweis. Nach Satz 14.3 ist f stetig. Deshalb ist das Integral definiert. Es gilt nun

$$\begin{aligned} \left| \int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx \right| &= \left| \int_a^b (f(x) - f_n(x))\,dx \right| \\ &\le \int_a^b |f(x) - f_n(x)|\,dx \le (b-a) \cdot ||f - f_n||. \end{aligned}$$

Da $(||f - f_n||)_{n\in\mathbb{N}}$ eine Nullfolge bildet, folgt, dass auch die Integraldifferenzen eine Nullfolge bilden. □

Bemerkung 16.1 Satz 16.1 gilt nicht für punktweise Konvergenz: Für $n \ge 2$ sei $f_n \colon [0,1] \to \mathbb{R}$ definiert durch



$$f_n(x) := \max\left\{ n - n^2 \cdot \left| x - \tfrac{1}{n} \right|, 0 \right\}$$

Wie in Beispiel 14.2 kann man zeigen, dass (f_n) punktweise gegen die Nullfunktion auf $[0,1]$ konvergiert. Es gilt

$$\int_0^1 f_n(x)\,dx = 1 \text{ für alle } n \ge 2,$$

aber

$$\int_0^1 0\,dx = 0.$$

Satz 16.2 *Es seien $f_n\colon [a,b] \to \mathbb{R}$ stetig differenzierbare Funktionen ($n \in \mathbb{N}$), die punktweise gegen die Funktion $f\colon [a,b] \to \mathbb{R}$ konvergieren. Die Folge der Ableitungen $f_n'\colon [a,b] \to \mathbb{R}$ konvergiere gleichmäßig gegen eine Funktion $g\colon [a,b] \to \mathbb{R}$. Dann ist f differenzierbar und es gilt*

$$f'(x) = \lim_{n\to\infty} f_n'(x) \text{ für alle } x \in [a,b].$$

Beweis. Nach Satz 14.3 ist g stetig. Nach Satz 16.1 gilt für alle $x \in [a,b]$

$$\begin{aligned}\int_a^x g(\xi)\,d\xi &= \lim_{n\to\infty} \int_a^x f_n'(\xi)\,d\xi \\ &= \lim_{n\to\infty} [f_n(x) - f_n(a)] \text{ (Hauptsatz)} \\ &= f(x) - f(a)\end{aligned}$$

Differentiation ergibt:

$$f'(x) = g(x)$$

für alle $x \in [a,b]$ (nach dem Hauptsatz der Differential- und Integralrechnung). □

Bemerkung 16.2 Selbst wenn (f_n) gleichmäßig gegen eine differenzierbare Funktion f konvergiert, gilt i.A. nicht, dass die Folge (f_n') gegen f' konvergiert, wie folgendes Beispiel zeigt:

$$f_n\colon \mathbb{R} \to \mathbb{R}, \quad f_n(x) := \frac{1}{n}\sin nx \quad (n \geq 1).$$

Da $||f_n|| = \frac{1}{n}$ konvergiert die Folge (f_n) gleichmäßig gegen die Nullfunktion. Die Folge der Ableitungen $f_n'(x) = \cos nx$ konvergiert jedoch nicht gegen die Nullfunktion.

Man kann nun auch Funktionenreihen $\sum_{n=0}^{\infty} f_n$ betrachten.

Definition Eine Funktionenreihe $\sum_{n=0}^{\infty} f_n$ heißt *gleichmäßig konvergent*, wenn die Folge $(s_n)_{n\in\mathbb{N}}$ ihrer Partialsummen gleichmäßig konvergent ist.

Satz 16.3 (Konvergenzkriterium von Weierstraß) *Gegeben seien Funktionen $f_n\colon [a,b] \to \mathbb{R}$, $n \in \mathbb{N}$. Wenn $||f_n|| \leq c_n$ für alle $n \in \mathbb{N}$ gilt und die Reihe $\sum_{n=0}^{\infty} c_n$ konvergiert, dann konvergiert die Funktionenreihe $\sum_{n=0}^{\infty} f_n$ gleichmäßig.*

Beweis. a) Wir zeigen zunächst, dass $\sum_{n=0}^{\infty} f_n$ punktweise gegen eine gewisse Funktion $F\colon [a,b] \to \mathbb{R}$ konvergiert.

Es sei $x \in [a,b]$. Da

$$|f_n(x)| \leq ||f_n|| \leq c_n \text{ für alle } n \in \mathbb{N},$$

konvergiert nach dem Majorantenkriterium die Reihe $\sum_{n=0}^{\infty} f_n(x)$ absolut. Wir setzen

$$F(x) := \sum_{n=0}^{\infty} f_n(x).$$

Damit ist eine Funktion $F\colon [a,b] \to \mathbb{R}$ definiert.

b) Es sei $F_n = \sum_{k=0}^{n} f_k$. Wir beweisen jetzt, dass die Folge (F_n) gleichmäßig gegen F konvergiert.

Sei $\varepsilon > 0$ vorgegeben. Aus der Konvergenz von $\sum_{n=0}^{\infty} ||f_n||$ folgt, dass es ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq n_0$

$$\sum_{k=n+1}^{\infty} ||f_k|| < \varepsilon.$$

Dann gilt für alle $n \geq n_0$ und $x \in [a,b]$

$$|F_n(x) - F(x)| = \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} ||f_k|| < \varepsilon.$$

□

Beispiel 16.1

$$\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$$

Diese Reihe konvergiert gleichmäßig auf $\mathbb{R}$, denn für $f_n(x) := \frac{\cos nx}{n^2}$ gilt

$$||f_n|| = \frac{1}{n^2} \text{ und } \sum_{n=1}^{\infty} \frac{1}{n^2}$$

konvergiert.

Wir wollen nun Potenzreihen betrachten.

Es sei eine Folge (a_n) und eine reelle Zahl x_0 gegeben, $a_n \in \mathbb{R}$ für alle $n \in \mathbb{N}$, x sei eine beliebige reelle Zahl. Dann heißt die Reihe

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

Potenzreihe mit der Koeffizientenfolge (a_n) und der Entwicklungsstelle x_0.

Die wichtigste Frage ist zunächst, für welche x die Reihe konvergiert. Für dieses Problem genügt es, denn Fall $x_0 = 0$ zu betrachten.

Satz 16.4 *Wenn die Potenzreihe* $\sum_{n=0}^{\infty} a_n x^n$ *für* $x = x_1 \neq 0$ *konvergiert, so konvergiert sie für jedes* x *mit* $|x| < |x_1|$ *sogar absolut.*

Beweis. Da die Reihe in x_1 konvergiert, bilden die Glieder $a_n x_1^n$ eine Nullfolge, sind also insbesondere beschränkt. Für alle $n \in \mathbb{N}$ ist also

$$|a_n x_1^n| \leq M.$$

Es folgt für $n \in \mathbb{N}$

$$|a_n x^n| \leq M \cdot \left|\frac{x}{x_1}\right|^n$$

Die geometrische Reihe $\sum_{n=0}^{\infty} M \cdot q^n$ mit $q = \left|\frac{x}{x_1}\right|$ ist also eine Majorante für $\sum_{n=0}^{\infty} |a_n x^n|$. □

Definition Es sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe. Dann heißt

$$r := \sup\left\{\varrho \;\middle|\; \varrho \geq 0 \text{ und } \sum |a_n|\varrho^n \text{ konvergent}\right\}$$

der *Konvergenzradius* der Potenzreihe, falls das Supremum existiert. Andernfalls setzen wir $r := \infty$. Das Intervall $(-r, r)$ heißt das *Konvergenzintervall* der Potenzreihe.

Bemerkung 16.3 Es gilt $r \in \mathbb{R}_+ \cup \{\infty\}$.

Satz 16.5 *Ist* $|x| < r$*, d.h.* $x \in (-r, r)$*, so ist* $\sum_{n=0}^{\infty} a_n x^n$ *absolut konvergent.*

Ist $|x| > r$*, so ist* $\sum_{n=0}^{\infty} a_n x^n$ *divergent.*

Bemerkung 16.4 Über $x = \pm r$ werden keine Aussagen gemacht: Für $x = \pm r$ ist sowohl Divergenz wie auch Konvergenz möglich: Z.B. ist der Konvergenzradius von $\sum_{n=0}^{\infty} x^n$ gleich 1. Man hat Divergenz an beiden Rändern des Konvergenzintervalls.

Auch der Konvergenzradius von $\sum_{n=1}^{\infty} (-1)^n \frac{x^n}{n}$ ist gleich 1 (Taylorreihe von $\ln(1 + x)$, siehe § 13). Die Reihe konvergiert für $x = 1$ nach dem Leibnitzkriterium, aber nicht für $x = -1$ (harmonische Reihe).

Beweis von Satz 16.5. Ist $|x| < r$, so gibt es – weil $|x|$ nicht das Supremum der obengenannten Menge ist – ein ϱ, so dass $|x| < \varrho < r$ und $\sum_{n=0}^{\infty} a_n \varrho^n$ konvergiert. Nach Satz 16.4 konvergiert dann $\sum_{n=0}^{\infty} a_n x^n$ absolut.

Ist hingegen $\sum_{n=0}^{\infty} a_n x^n$ konvergent für ein x mit $|x| > r$, so findet sich ein ϱ, $|x| > \varrho > r$, für das $\sum_{n=0}^{\infty} |a_n|\varrho^n$ konvergiert. Widerspruch, da r Supremum. □

Satz 16.6 *Jede Potenzreihe* $\sum_{n=0}^{\infty} a_n x^n$ *ist in jedem abgeschlossenen und beschränkten Teilintervall – etwa in* $[-c, c]$ *– des Konvergenzintervalls der Potenzreihe gleichmäßig konvergent.*

Beweis. Aus
$$|a_n x^n| \leq |a_n| c^n \text{ für } |x| \leq c < r$$
folgt wegen der absoluten Konvergenz von $\sum_{n=0}^{\infty} a_n c^n$ die gleichmäßige Konvergenz von $\sum_{n=0}^{\infty} a_n x^n$ im abgeschlossenen Intervall $[-c, c]$. □

Bemerkung 16.5 Eine Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ braucht im offenen Intervall $(-r, r)$ – wobei r der Konvergenzradius ist – nicht gleichmäßig zu konvergieren: Man vergleiche die geometrische Reihe $\sum_{n=0}^{\infty} x^n$.

Satz 16.7 *Die durch formales, gliedweises Differenzieren aus einer Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ entstehende Reihe*
$$\sum_{n=1}^{\infty} n \cdot a_n x^{n-1}$$
hat denselben Konvergenzradius wie die ursprüngliche Reihe.

Beweis. Für $x_1 \neq 0$ konvergiere $\sum_{n=0}^{\infty} a_n x_1^n$. Sei $|a_n x_1^n| < M$, dann ist
$$|n \cdot a_n \cdot x^n| \leq nM \cdot \left|\frac{x}{x_1}\right|^n.$$
Also ist $\sum_{n=0}^{\infty} n \cdot q^n$ mit $q = \left|\frac{x}{x_1}\right|$ konvergente (nach dem Quotientenkriterium) Majorante. Also konvergiert die Reihe $\sum_{n=1}^{\infty} n a_n x^{n-1}$ für $|x| < |x_1|$ absolut.

Umgekehrt: $\sum_{n=0}^{\infty} n a_n \cdot x_1^n$ konvergiere. Wie oben ist für $n > 0$
$$|a_n x^n| \leq |n \cdot a_n x^n| = |n \cdot a_n \cdot x_1^n| \cdot \left|\frac{x}{x_1}\right|^n,$$
also ist $\sum_{n=0}^{\infty} M \cdot \left|\frac{x}{x_1}\right|^n$ konvergente Majorante zu $\sum_{n=0}^{\infty} |a_n x^n|$.

Folglich haben $\sum_{n=0}^{\infty} a_n x^n$ und $\sum_{n=1}^{\infty} n a_n x^{n-1}$ denselben Konvergenzradius (falscher konstanter Faktor $|x|$ stört nicht). □

Satz 16.8 *Gegeben sei eine Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ mit Konvergenzradius $r > 0$. In $(-r, r)$ sei $f(x) := \sum_{n=0}^{\infty} a_n x^n$. Dann ist f in $(-r, r)$ differenzierbar und für alle $x \in (-r, r)$ gilt:*
$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

Bemerkung 16.6 Man drückt dies auch so aus: Eine Potenzreihe darf gliedweise differenziert werden.

Beweis von Satz 16.8. Die Teilsummen

$$s_k(x) = \sum_{n=0}^{k} a_n x^n$$

sind in $\mathbb{R}$ und damit auch im Konvergenzintervall der Potenzreihe differenzierbar und es gilt

$$s_k'(x) = \sum_{n=1}^{k} n a_n x^{n-1}.$$

Die Potenzreihe $\sum_{n=1}^{\infty} n a_n x^{n-1}$ hat nach Satz 16.7 denselben Konvergenzradius wie die Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$, nach Satz 16.6 konvergiert sie gleichmäßig in jedem abgeschlossenem Teilintervall von $(-r, r)$. Nach Satz 16.2 konvergiert sie dort gegen die Ableitung f'.

Sei $x \in (-r, r)$. Dann kann man ein $q \in \mathbb{R}$ finden, so dass $0 \leq |x| \leq q < r$ ist. Dann ist $[-q, q]$ ein abgeschlossenes Teilintervall von $(-r, r)$, das x enthält. Also gilt

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

□

Korollar 16.1 *Es sei $f(x) := \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius r. Dann ist $f\colon (-r, r) \to \mathbb{R}$ beliebig oft differenzierbar und es gilt*

$$a_n = \frac{f^{(n)}(0)}{n!}.$$

Beweis. Wiederholte Anwendung von Satz 16.8 ergibt

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) a_n x^{n-k}.$$

Insbesondere folgt daraus

$$f^{(k)}(0) = k! a_k, \text{ d.h. } a_k = \frac{f^{(k)}(0)}{k!}.$$

□

Korollar 16.2 *Es sei $f(x) := \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius r. Dann stimmt die Taylorreihe von f um 0 mit $\sum_{n=0}^{\infty} a_n x^n$ überein und konvergiert gegen f.*

Beweis. Dies folgt unmittelbar aus der Definition der Taylorreihe und Korollar 16.1. □

Mit dem Satz über die gliedweise Differenzierbarkeit steht uns ein neues Hilfsmittel zur Verfügung, mit dem wir für manche unendlich oft differenzierbaren Funktionen die Darstellbarkeit durch ihre Taylorreihen beweisen können, ohne direkt eine Restgliedabschätzung durchführen zu müssen.

Beispiel 16.2 $\ln\colon \mathbb{R}_+^* \to \mathbb{R}$.

Wir haben bereits gesehen, dass die Taylorreihe von ln bzgl. $x_0 = 1$ lautet:

$$f(x) := \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot (x-1)^n.$$

Wir wollen noch einmal auf andere Weise zeigen, dass die durch diese Potenzreihe dargestellte Funktion f in ihrem Konvergenzintervall $0 < x < 2$ mit ln übereinstimmt. Dazu bilden wir die Ableitung:

$$\begin{aligned} \ln' x - f'(x) &= \frac{1}{x} - \sum_{n=1}^{\infty} (-1)^{n-1}(x-1)^{n-1} \\ &= \frac{1}{x} - \frac{1}{1+(x-1)} = \frac{1}{x} - \frac{1}{x} = 0. \end{aligned}$$

Es muß also für $0 < x < 2$

$$f(x) = \ln x + c$$

gelten. Einsetzen von $x = 1$ zeigt, dass $c = 0$ ist. Damit haben wir

$$\ln x = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (x-1)^n \qquad (0 < x < 2).$$

Dass diese Gleichung auch für $x = 2$ noch richtig ist, haben wir schon früher bewiesen.

Beispiel 16.3 (Arcus-Tangens-Reihe)

Wir berechnen die Taylorreihe von $\arctan x$ für $|x| < 1$ nach Satz 16.1:

$$\begin{aligned} \arctan x &= \int_0^x \frac{1}{1+t^2}\, dt \\ &= \int_0^x \left(\sum_{n=0}^{\infty} (-1)^n t^{2n} \right) dt \\ &= \sum_{n=0}^{\infty} (-1)^n \int_0^x t^{2n}\, dt = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}. \end{aligned}$$

Als weiteres Beispiel betrachten wir die binomische Reihe.

Für $a > 0$ und $b \in \mathbb{R}$ war $a^b = \exp(b \ln a)$. Wir können also für $x > -1$ und $\mu \in \mathbb{R}$ die Funktion $(1+x)^\mu$ betrachten. Ist $\mu \in \mathbb{N}$, so ist nach dem binomischen Lehrsatz

$$(1+x)^\mu = \sum_{n=0}^{\mu} \binom{\mu}{n} x^n.$$

Wir suchen nun ein Analogon für nichtnatürliche μ. Dazu definieren wir zunächst

$$\binom{\mu}{n} := \frac{\mu(\mu-1)\cdots(\mu-n+1)}{n!} \text{ für } \mu \in \mathbb{R}, n \in \mathbb{N}.$$

(Ein Produkt von 0 Faktoren ist per definitionem gleich 1. Also $\binom{\mu}{0} = 1$.)

Es gilt nun

Satz 16.9 (Binomische Reihe) *Es sei $\mu \in \mathbb{R}$. Dann gilt für $|x| < 1$*

$$(1+x)^\mu = \sum_{n=0}^{\infty} \binom{\mu}{n} x^n.$$

Beweis. a) Berechnung der Taylorreihe von $f(x) = (1+x)^\mu$ mit Entwicklungspunkt 0:

$$\begin{aligned} f^{(n)}(x) &= \mu(\mu-1)\cdots(\mu-n+1)(1+x)^{\mu-n} \\ &= n!\binom{\mu}{n}(1+x)^{\mu-n}. \end{aligned}$$

Da also $\frac{f^{(n)}(0)}{n!} = \binom{\mu}{n}$, lautet die Taylorreihe von f

$$\sum_{n=0}^{\infty} \binom{\mu}{n} x^n.$$

b) Wir zeigen, dass die Taylorreihe für $|x| < 1$ konvergiert mit dem Quotientenkriterium. Wir dürfen annehmen, dass $\mu \notin \mathbb{N}$ und $x \neq 0$. Es sei

$$a_n := \binom{\mu}{n} x^n.$$

Dann gilt

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{\binom{\mu}{n+1} x^{n+1}}{\binom{\mu}{n} x^n} \right| = |x| \cdot \left| \frac{\mu - n}{n+1} \right|.$$

Da

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = |x| \cdot \lim_{n\to\infty} \left| \frac{\mu - n}{n+1} \right| = |x| < 1,$$

existiert zu q mit $|x| < q < 1$ ein n_0, so dass

$$\left|\frac{a_{n+1}}{a_n}\right| \leq q \text{ für alle } n \geq n_0.$$

Also konvergiert die Taylorreihe für $|x| < 1$.

c) Wir beweisen jetzt, dass die Taylorreihe gegen f konvergiert.

Wir setzen

$$g(x) := \sum_{n=0}^{\infty} \binom{\mu}{n} x^n.$$

Die Ableitung von g ist

$$g'(x) = \sum_{n=0}^{\infty} \binom{\mu}{n} n x^{n-1} = \mu \sum_{k=0}^{\infty} \binom{\mu-1}{k} x^k.$$

Multiplikation mit $(1+x)$ ergibt

$$\begin{aligned}(1+x)g'(x) &= \mu \sum_{k=0}^{\infty} \left(\binom{\mu-1}{k} + \binom{\mu-1}{k-1}\right) x^k \\ &= \mu \sum_{k=0}^{\infty} \binom{\mu}{k} x^k = \mu g(x).\end{aligned}$$

Wir betrachten nun die Funktion h, die definiert ist durch

$$h(x) = \frac{g(x)}{(1+x)^{\mu}} \quad \text{für } -1 < x < 1.$$

Die Funktion h ist differenzierbar und es gilt

$$h'(x) = \frac{(1+x)^{\mu} g'(x) - \mu(1+x)^{\mu-1} g(x)}{(1+x)2\mu} = 0.$$

Die Funktion h muss also konstant sein. Einsetzen von $x = 0$ zeigt, dass $h(x) = 1$ für alle $x \in (-1,1)$ gelten muss. □

Beispiel 16.4 (1) $\mu = 1$:

Da

$$\binom{-1}{\mu} = \frac{(-1)\cdot(-1-1)\cdots(-1-n+1)}{n!} = (-1)^n,$$

ergibt sich für $\mu = -1$ aus der binomischen Reihe

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n \text{ für } |x| < 1.$$

Die geometrische Reihe ist also ein Spezialfall der binomischen Reihe.

(2) $\mu = -\frac{1}{2}$:

$$\begin{aligned}\binom{-\frac{1}{2}}{n} &= \frac{(-\frac{1}{2})(-\frac{1}{2}-1)\cdots(-\frac{1}{2}-n+1)}{n!} \\ &= (-1)^n \cdot \frac{1}{2^n} \cdot \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{n!} \\ &= (-1)^n \cdot \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots 2n}.\end{aligned}$$

Nach Satz 16.9 gilt

$$(1+x)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n} x^n \text{ für } |x| < 1.$$

Mit $|x| < 1$ ist $|-x^2| < 1$, also erhalten wir durch Einsetzen

$$(1-x^2)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots 2n} x^{2n}.$$

In jedem kompakten Teilintervall von $(-1, 1)$ ist die rechts stehende Potenzreihe sogar gleichmäßig konvergent, so dass für $|x| < 1$ über das Intervall $[0, x]$ bzw. $[x, 0]$ gliedweise integriert werden kann. Also

$$\begin{aligned}\arcsin x &= \int_0^x \frac{1}{\sqrt{1-t^2}}\, dt \\ &= \sum_{n=0}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdots 2n} \cdot \int_0^x t^{2n}\, dt \\ &= \sum_{n=0}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdots 2n \cdot (2n+1)} x^{2n+1} \text{ für } |x| < 1.\end{aligned}$$

Ohne Beweis notieren wir, dass die Gleichheit dieser Reihe und arcsin auch an der Stelle $x = 1$, also auch in $x = -1$ gilt. Einsetzen von $x = 1$ liefert

$$\frac{\pi}{2} = \sum_{0}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)} \cdot \frac{1}{2n+1},$$

eine bekannte Entwicklung für $\frac{\pi}{2}$.

17 Uneigentliche Integrale

Bei der Definition des Integrals für stetige Funktionen war die Voraussetzung wesentlich, dass das Integrationsintervall beschränkt und abgeschlossen ist. Wir wollen nun die Integraldefinition auf den Fall ausdehnen, dass das Integrationsintervall unbeschränkt ist. Man spricht dann von einem *uneigentlichen Integral*.

Definition Es sei $f : [a, \infty) \to \mathbb{R}$ eine stetige Funktion. Falls der Grenzwert

$$\lim_{n\to\infty} \int_a^n f(x)\, dx$$

existiert, heißt das Integral $\int_a^\infty f(x)\, dx$ *konvergent* und man setzt

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

Beispiel 17.1 Das Integral $\int_1^\infty \frac{1}{x^s}\, dx$ konvergiert für $s > 1$. Es gilt nämlich

$$\int_1^n \frac{1}{x^s}\, dx = \left[-\frac{1}{s-1} \cdot \frac{1}{x^{s-1}} \right]_1^n = \frac{1}{s-1} \left(1 - \frac{1}{n^{s-1}} \right).$$

Da $\lim_{n\to\infty} \frac{1}{n^{s-1}} = 0$, folgt

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

Für $s = 1$ gilt

$$\int_1^n \frac{1}{x^s}\, dx = \ln n - \ln 1,$$

und für $s < 1$ gilt

$$\int_1^n \frac{1}{x^s}\, dx = \frac{1}{1-s}(n^{1-s} - 1).$$

Daher gilt

$$\int_1^\infty \frac{1}{x^s}\, dx \text{ konvergiert nicht für } s \leq 1.$$

Mit Hilfe eines uneigentlichen Integrals kann man manchmal einfach entscheiden, ob eine Reihe konvergiert oder divergiert.

Satz 17.1 (Integralkriterium für Reihen) *Es sei $f : [1, \infty) \to (0, \infty)$ eine monoton fallende stetige Funktion. Die Reihe*

$$\sum_{n=1}^\infty f(n)$$

konvergiert genau dann, wenn das uneigentliche Integral

$$\int_1^\infty f(x)\, dx$$

konvergiert.

Beweis. Für jedes $k \in \mathbb{N}$ folgt aus $k \leq x \leq k+1$

$$f(k+1) \leq f(x) \leq f(k)$$

und daraus

$$f(k+1) \leq \int_k^{k+1} f(x)\,dx \leq f(k).$$

Hieraus ergibt sich durch Addition:

$$f(2) + \cdots + f(n) \leq \int_1^n f(x)\,dx \leq f(1) + \cdots + f(n-1).$$

Daraus folgt unmittelbar die Behauptung. □

Beispiel 17.2 Aus dem Integralkriterium folgt:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \text{ ist } \begin{cases} \text{konvergent für } s > 1, \\ \text{divergent für } 0 < s \leq 1. \end{cases}$$

Bemerkung 17.1 Betrachtet man die Summe der Reihe als Funktion von s, so erhält man die *Riemannsche Zetafunktion*.

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} \quad (s > 1).$$

Euler wusste schon, dass

$$\zeta(2) = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90}, \quad \zeta(6) = \frac{\pi^6}{945}, \text{ etc.}$$

In der Vorlesung Funktionentheorie wird diese Funktion ins Komplexe fortgesetzt. Die *Riemannsche Vermutung* ist eine berühmte, bisher ungelöste Vermutung, wie die komplexen Nullstellen dieser Funktion aussehen.

Literatur

[1] M. Barner, F. Flohr: Analysis 1. 3. Auflage. Walter de Gruyter 1987.

[2] O. Forster: Analysis 1. 5., überarb. Auflage. Vieweg, Wiesbaden, 1999. ISBN 3-528-47224-3

[3] K. Königsberger: Analysis 1. Springer-Verlag 1990.

[4] S. L. Salas, E. Hille: Calculus. Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford, 1994. ISBN 3-86025-130-9

[5] M. Spivak: Calculus. Third Edition. Publish or Perish, Houston, 1994. ISBN 0-914098-89-6 (amerikanisch)