Analysis I

WS 2016/2017

Basierend auf einem Skript von E. Kuwert

Guofang Wang

Mathematisches Institut
Universität Freiburg

Inhaltsverzeichnis

Kapitel 1

Grundlagen

0 Mengenlehre und Abbildungen

0.1 Etwas Mengenlehre

Georg Cantor: Eine Menge ist die ungeordnete Zusammenfassung verschiedener Elemente zu einem Ganzen.

Bezeichnung. Ist M eine solche Menge, so führen wir die folgenden Bezeichnungen ein:

$$a \in M$$

heißt: "a ist Element von M" (oder: "a gehört zu M", "a liegt in M", "a ist aus M").

$$a \notin M$$

heißt: "a ist nicht Element von M".

$$M = \{a, b, \cdots\}$$

heißt: " M ist die Menge, die aus den Elemneten a, b usw. besteht".

$$M = \{a : a \text{ hat die Eigenschft } E\}$$

heißt: "M ist die Menge aller Elementen a, die die Eigenschaft E haben, und nur diese Elemente ligen in M."

Beispiel 0.1 *i) Für $a \neq b$ gilt*

$$\{a, b\} = \{b, a\} = \{a, b, a\}.$$

ii) $\mathbb{N} = \{1, 2, 3, 4, \cdots\}$.

iii) $\emptyset = \{\}$*: leere Menge.*

iv) $\{n \in \mathbb{N} \,|\, n \textit{ teilt } 15\} = \{1, 3, 5, 15\}$.

Quantoren:
"$\forall$": "für alle" z.B. $\forall\, n \in \mathbb{N} : (n+1)^2 > n+1$.
"$\exists$": "es gbit" z.B. $\forall\, n \in \mathbb{N} \exists\, k \in \mathbb{N} : k > n$.
":=": "gilt per Definition" z.B. $M := \{1, 2, 3\}$.
"$(A) \Rightarrow (B)$": "Aus der Aussage (A) folgt die Aussage (B)" z.B $n > 5 \Rightarrow n^2 > 20$.
Sei M, N zwei Mengen. M heißt in N *enthalten*, in Zeichen:

$$M \subset N,$$

wenn gilt: ist $x \in M$, so auch $x \in N$. Man nennt M dann auch *Teilmenge* von N. Zwei Menge M, N sind genau dann *gleich*, wann

$$M \subset N \text{ und } N \subset M.$$

Dass heißt:

$$\text{“}M = N \Leftrightarrow \text{“}M \subset N \text{ und } N \subset M$$

Definition 0.1 (Verknüpfungen) *Seien M, N beliegige Mengen:*

- *Vereinigung: $M \cup N := \{x \,|\, x \in M \text{ oder } x \in N\}$.*
- *Durchschnitt: $M \cap N := \{x \,|\, x \in M \text{ und } x \in N\}$.*
- *Differenz: $M \backslash N := \{x \,|\, x \in M, x \notin N\}$*
- *Produktmenge: $M \times N := \{(x, y) \,|\, x \in M \text{ und } y \in N\}$*
- *Potenmenge: $\mathcal{P}(M) :=$ Die Menge aller Teilmengen von M.*

Sei $M \subset X$. Das *Komplement* von M in X ist $M^c := X \backslash M$.

Für diese Verknüpfungen gelten die folgenden Regeln:

$$\begin{aligned}
(A \cup B) \cup C &= A \cup (B \cup C), \\
(A \cap B) \cap C &= A \cap (B \cap C), \\
A \cup (B \cap C) &= (A \cup B) \cap (A \cup C), \\
A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\
A \cup \emptyset &= A, \\
A \cap \emptyset &= \emptyset.
\end{aligned}$$

Beweis: Wir zeigen nur

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Es folgt aus

$$\begin{aligned}
x \in A \cap (B \cup C) &\Leftrightarrow (x \in A) \text{ und } (x \in B \cup C) \\
&\Leftrightarrow (x \in A) \text{ und } (x \in B \text{ oder } x \in C) \\
&\Leftrightarrow (x \in A \text{ und } x \in B) \text{ oder } (x \in A \text{ und } x \in C) \\
&\Leftrightarrow (x \in A \cap B) \text{ oder } (x \in A \cap C) \\
&\Leftrightarrow x \in (A \cap B) \cup (A \cap C).
\end{aligned}$$

$\square$

Für das Komplement haben wir die folgenden Eigenschaften:
Seien $A, B \subset X$ mit Komplement $A^c = X \backslash A$ und $B^c = X \backslash B$. Dann gilt:

i) $(A \cup B)^c = A^c \cap B^c$.

ii) $(A \cap B)^c = A^c \cup B^c$.

iii) ist $A \subset B$, so ist $B^c \subset A^c$.

Beweis: Wir zeigen nur i). Sei $x \in X$. i) folgt aus

$$\begin{aligned} x \in (A \cup B)^c &\Leftrightarrow x \notin (A \cup B) \\ &\Leftrightarrow (x \notin A) \text{ und } (x \notin B) \\ &\Leftrightarrow (x \in A^c) \text{ und } (x \in B^c) \\ &\Leftrightarrow x \in A^c \cap B^c. \end{aligned}$$

□

0.2 Abbildungen

Definition 0.2 *Seien M, N Mengen. Eine* Abbildung *f von M nach N*

$$f : M \to N$$

ordnet jedem Element $x \in M$ ein Element $f(x) \in N$ zu.

Man bezeichnet die Abbildung $f : M \to N$ auch mit $x \mapsto f(x)$. Nämlich

$$f : N \to M, x \mapsto f(x).$$

Für $N = \mathbb{R}$ wird f auch *Funktion* genannt. Man nennt M den *Definitionsbereich* von f, N den *Zielraum* von f, und $f(M) := \{f(x) : x \in M\}$ heißt *Wertebereich* von f oder *Bild* von M unter f. Für $A \subset M$ nennt man die Menge $f(A) := \{f(x) : x \in A\}$ *Bild* von A und für $B \subset N$ die Menge

$$f^{-1}(B) := \{x \in N : f(x) \in B\} \quad \text{Urbild } \textit{von } B.$$

Eine Abbildung (Funktion) wird erklärt durch die Angabe:

- des Definitionsbereiches (hier M)
- des Bild- oder Wertebereiches (hier N)
- der Abbildungsvorschrift (hier $x \mapsto f(x)$).

Definition 0.3 *Die Abbildung $f : M \to N$ heißt*

- *injektiv, wenn aus $f(x) = f(x')$ folgt $x = x'$,*
- *surjektiv, wenn zu jedem $y \in N$ gibt es ein $x \in M$ mit $f(x) = y$,*
- *bijektiv, wenn f ist injektiv und surjektiv.*

Definition 0.4 *Seien M, N, L Mengen und $f : M \to N$ eine Abbildung. Sei $x \in M$*

i) Zu der Abbildung $g : N \to L$ ist die Komposition $g \circ f : M \to L$ die Abbildung definiert durch $(g \circ f)(x) = g(f(x))$.

ii) Die Abbildung $g : N \to M$ heißt inverse Abbildung zu f, in Zeichen $g = f^{-1}$, wenn gilt: $g \circ f = id_M$ und $f \circ g = id_N$. Hierbei bezeichnet $id_M : M \to M$ die Identitätsabbildung, $id_M(x) = x$ für alle $x \in M$.

1 Körperaxiome und Anordnungsaxiome

Wir setzen in dieser Vorlesung die reellen Zaheln als gegeben aus. Mit $\mathbb{R}$ bezeichnen wir die Menge aller reellen Zahlen, auf der folgende Strukturen gegeben sind:

> *eine Verknüpfung $+$, die je zwei $a, b \in \mathbb{R}$ ein $a + b \in \mathbb{R}$ zuordnet (*Addition*),*
> *eine Verknüpfung $\cdot$, die je zwei $a, b \in \mathbb{R}$ ein $ab \in \mathbb{R}$ zuordnet (*Multiplikation*),*
> *eine Relation $a > b$, die für $a, b \in \mathbb{R}$ zutrifft oder nicht (*Anordnung*).*

Es sollen dabei folgende drei Gruppen von Axiomen gelten:

> *Die Körperaxiome* (K)
> *Die Anordnungsaxiome* (A1, A2 und A3)
> *Das Vollständigkeitsaxiom* (V).

1.1 Körperaxiome

Wir wollen in der Folge die einzelnen Axiome vorstellen und beginnen mit den Körperaxiomen, die die arithmetischen Eigenschaften von $\mathbb{R}$ regeln, das heißt die Addition und die Multiplikation. Sie lauten wie folgt:

Körperaxiome

	$+$	$\bullet$
Assoziativgesetz:	$(a + b) + c = a + (b + c)$	$(a \cdot b) \cdot c = a \cdot (b \cdot c)$
Kommutativgesetz:	$a + b = b + a$	$a \cdot b = b \cdot a$
Neutrales Element:	*Es gibt Zahlen $0 \in \mathbb{R}$ und $1 \in \mathbb{R}$ mit $1 \neq 0$, so dass für alle $a \in \mathbb{R}$ gilt:*	
	$a + 0 = a$	$a \cdot 1 = a$
Inverses Element:	*Zu jedem $a \in \mathbb{R}$ gibt es Lösungen $x, y \in \mathbb{R}$ der Gleichungen*	
	$a + x = 0$	$a \cdot y = 1$ *falls* $a \neq 0$
Distributivgesetz:	$a \cdot (b + c) = a \cdot b + a \cdot c.$	

Definition 1.1 *Eine Menge $\mathbb{K}$ mit Verknüpfungen $+$ und $\cdot$, so dass die obigen Axiome erfüllt sind, heißt* Körper *(englisch:* field*).*

In der Algebra treten viele verschiedene Körper auf, ein extremes Beispiel ist

Beispiel 1.1 $\mathbb{K} = \{0, 1\}$ *mit $1 + 1 = 0$ und den sonst üblichen Rechenregeln.*

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \textit{und} \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

Aus der Körperaxiomen haben wir direkte

Folgerung 1.1 *(1). Die neutralen Elemente sind durch die Axiome eindeutig bestimmt. (2). Die inversen Elemente sind ebenfalls eindeutig bestimmt.*

BEWEIS: (1). Zum Beispiel wären $0_1 \in \mathbb{R}$ und $0_2 \in \mathbb{R}$ neutrale Elemente für die Addition, so folgt mit dem Kommutativgesetz

$$0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2.$$

Das Argument für die Eindeutigkeit von $1 \in \mathbb{R}$ ist natürlich ganz analog.

(2). Für zwei Lösungen $x_{1,2}$ der Gleichung $a + x = 0$ folgt mit dem Assoziativgesetz und dem Kommutativgesetz

$$x_1 = x_1 + 0 = x_1 + (a + x_2) = (x_1 + a) + x_2 = (a + x_1) + x_2 = 0 + x_2 = x_2.$$

Wieder ist das Argument für die Multiplikation analog. □

Bezeichnung. Wir bezeichnen die Lösung x der Gleichung $a + x = 0$ mit $-a$ sowie die Lösung y der Gleichung $ay = 1$ mit $1/a$ oder a^{-1}, und vereinbaren die Notation

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

Aus den Körperaxiomen leiten sich unter anderem folgende Rechengesetze ab.

Satz 1.1 (Rechnen in $\mathbb{R}$) *Für reelle Zahlen a, b gelten folgende Aussagen:*

$$\begin{cases} -(-a) = a & -(a+b) = (-a) + (-b), \\ (a^{-1})^{-1} = a & (ab)^{-1} = a^{-1}b^{-1} \text{ für } b \neq 0, \\ a \cdot 0 = 0 & a(-b) = -(ab), \\ (-a)(-b) = ab & a(b-c) = ab - ac. \end{cases} \tag{1.1}$$

$$ab = 0 \quad \Rightarrow \quad a = 0 \text{ oder } b = 0 \quad (\textit{Nullteilerfreiheit}). \tag{1.2}$$

BEWEIS: Die erste Zeile von (1.1) folgt mit der Eindeutigkeit der Inversen aus

$$(-a) + a = a + (-a) = 0 \qquad (a + b) + ((-a) + (-b)) = (a + (-a)) + (b + (-b)) = 0 + 0 = 0.$$

Der Beweis der zweiten Zeile ist analog. Nun gilt $a \cdot 0 + a \cdot 0 = a \cdot (0 + 0) = a \cdot 0$. Nach Addition von $-a \cdot 0$ folgt $a \cdot 0 = 0$. Daraus ergibt sich weiter

$$ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0,$$

also $a(-b) = -ab$, und dann

$$(-a)(-b) = -((-a)b) = -(b(-a)) = -(-ba) = ba = ab.$$

Die letzte Aussage von (1.1) folgt mit

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

Ist in (1.2) $a \neq 0$, so folgt schließlich

$$0 = ab \cdot \frac{1}{a} = (a \cdot \frac{1}{a}) \cdot b = 1 \cdot b = b.$$

Damit sind alle Aussagen des Satzes gezeigt. $\square$

Auch die folgenden Regeln der Bruchrechnung lassen sich aus den Körperaxiomen herleiten, dies sei jedoch den Lesern überlassen.

Folgerung 1.2 (Bruchrechnung) *Für $a, b, c, d \in \mathbb{R}$ mit $c, d \neq 0$ gilt:*

(1) $\frac{a}{c} + \frac{b}{d} = \frac{ad + bc}{cd}$,

(2) $\frac{a}{c} \cdot \frac{b}{d} = \frac{ab}{cd}$,

(3) $\frac{a/c}{b/d} = \frac{ad}{bc}$, *falls zusätzlich $b \neq 0$.*

Für reelle Zahlen $a_m, \ldots, a_n$ setzen wir

$$\sum_{k=m}^{n} a_k = a_m + \ldots + a_n \quad (m, n \in \mathbb{Z}).$$

Mit des Kommutativgesetzes und des Assoziativgestzes kann man leicht zeigen, dass es wohldefiniert ist. Der Index k durchläuft dabei die ganzen Zahlen von der unteren Grenze $k = m$ bis zur oberen Grenze $k = n$. Der Laufindex k kann substituiert werden, wobei die Grenzen umzurechnen sind. Zum Beispiel liefert $k = j + 1$

$$\sum_{k=m}^{n} a_k = \sum_{j=m-1}^{n-1} a_{j+1} = a_m + \ldots + a_n.$$

Es ist praktisch den Fall zuzulassen, dass die untere Grenze größer als die obere Grenze ist, und in diesem Fall die Summe gleich Null zu setzen. Dass heißt

$$\sum_{k=m}^{n} a_k = 0 \qquad \text{falls } n < m.$$

Das Produktzeichen ist ganz analog erklärt:

$$\prod_{k=m}^{n} a_k = a_m \cdot \ldots \cdot a_n.$$

Das leere Produkt wird gleich Eins gesetzt, d.h.,

$$\prod_{k=m}^{n} a_k = 1 \qquad \text{falls } n < m.$$

1.2 Anordnungsaxiome

Als nächstes formulieren wir die Anordnungsaxiome.

(**A1**) Für jedes $a \in \mathbb{R}$ gilt genau eine der drei Aussagen $a > 0$, $a = 0$ oder $-a > 0$.

(**A2**) Aus $a, b > 0$ folgt $a + b > 0$ und $ab > 0$.

(**A3**) Zu jedem $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, existiert ein $n \in \mathbb{N}$ mit $1/n < \varepsilon$ (*Archimedisches Axiom*).

Wir diskutieren erste die erste zwei Anordnungsaxiome. Das drittes, das Archimedische Axiom, werden wir erst später benutzen, wenn es für den Grenzwertbegriff gebraucht wird.

Bezeichnung. Statt $-a > 0$ schreiben wir auch $a < 0$, und statt $a - b > 0$ auch $a > b$.

Hier einige Folgerungen aus den Anordnungsaxiomen (A1) und (A2).

Satz 1.2 (Rechnen mit Ungleichungen)

(1) *Für $a, b \in \mathbb{R}$ gilt genau eine der Relationen $a > b$, $a = b$ oder $a < b$.* (Trichotomie)

(2) *Aus $a > b$, $b > c$ folgt $a > c$* (Transitivität).

(3) *Aus $a > b$ folgt*

$$\begin{cases} a + c > b + c & \\ ac > bc, & \text{wenn } c > 0 \\ ac < bc, & \text{wenn } c < 0 \end{cases}$$

(4) *Aus $a > b$ und $c > d$ folgt*

$$\begin{cases} a + c > b + d, & \\ ac > bd, & \text{falls } b, d > 0 \end{cases}$$

(5) *Für $a \neq 0$ ist $a^2 > 0$. Insbesondere, gilt $1 > 0$.*

(6) *Aus $a > 0$ folgt $1/a > 0$.*

(7) *Aus $a > b$, $b > 0$ folgt $1/a < 1/b$.*

Beweis: (1) folgt direkt aus (A1) und der Definition von $a > b$. Für (2) rechnen wir

$$a - c = (a - b) + (b - c) > 0 \quad \text{nach (A2)}.$$

Ähnlich ergeben sich (3) und (4):

$$\begin{aligned} (a + c) - (b + c) &= a - b > 0, && \\ ac - bc &= (a - b)c > 0 && \text{im Fall } c > 0 \text{ nach (A2)}, \\ bc - ac &= (a - b)(-c) > 0 && \text{im Fall } c < 0 \text{ nach (A2)}, \\ (a + c) - (b + d) &= (a - b) + (c - d) > 0 && \text{nach (A2)}, \\ ac - bd &= ac - bc + bc - bd && \\ &= (a - b)c + b(c - d) > 0 && \text{nach (2) und (A2)}. \end{aligned}$$

Die Positivität von Quadraten folgt aus der Fallunterscheidung

$$a^2 = \begin{cases} a \cdot a > 0 & \text{im Fall } a > 0, \\ (-a) \cdot (-a) > 0 & \text{im Fall } -a > 0. \end{cases}$$

Dabei haben wir die Regel $(-a)(-b) = ab$ aus (1.1) benutzt. Nach (A1) ist $a^2 > 0$ für $a \neq 0$ bewiesen. Aussage (6) ergibt sich nun mit

$$\frac{1}{a} = \left(\frac{1}{a}\right)^2 \cdot a > 0 \quad \text{nach (5) und (A2).}$$

Zu guter Letzt haben wir für (7)

$$\frac{1}{b} - \frac{1}{a} = \frac{1}{ab}(a-b) > 0 \quad \text{mit (6) und (A2).}$$

□

Definition 1.2 (Betrag einer reellen Zahl) *Der Betrag von $a \in \mathbb{R}$ ist*

$$|a| = \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0. \end{cases} \tag{1.3}$$

Ein anschauliches Modell der reellen Zahlen ist die Zahlengerade. Ist $a < b$, so liegt b rechts von a im Abstand $|a - b|$. Insbesondere ist $|a|$ der Abstand zum Nullpunkt.

Satz 1.3 (Rechnen mit Beträgen) *Für $a, b \in \mathbb{R}$ gelten folgende Aussagen:*

(1) $|-a| = |a|$ *und* $a \leq |a|$.

(2) $|a| \geq 0$*; aus Gleichheit folgt* $a = 0$.

(3) $|ab| = |a| \cdot |b|$.

(4) $|a + b| \leq |a| + |b|$. *(Dreiecke-Ungleichung)*

(5) $|a - b| \geq ||a| - |b||$.

Beweis: Aus Definition 1.2 folgt

$$|-a| = \begin{cases} -a & \text{falls } -a \geq 0 \\ -(-a) & \text{falls } -a \leq 0 \end{cases} = \begin{cases} -a & \text{falls } a \leq 0 \\ a & \text{falls } a \geq 0 \end{cases} = |a|.$$

Weiter folgt (2) aus

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

In (3) bleiben die linke und rechte Seite gleich, wenn wir a durch $-a$ ersetzen, dasselbe gilt bezüglich b. Also können wir $a, b \geq 0$ annehmen, und erhalten $|ab| = ab = |a| \cdot |b|$ wie verlangt. Für (4) schätzen wir mit (1) wie folgt ab:

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

Schließlich gilt $|a| = |a - b + b| \leq |a - b| + |b|$ nach (4), also $|a - b| \geq |a| - |b|$. Durch Vertauschen von a und b folgt (5). □

Wie wir später sehen werden, spielen Ungleichungen in der Analysis eine große Rolle.

2 Vollständige Induktion

Wir unterbrechen jetzt die Diskussion der Axiome der reellen Zahlen, um das Beweisverfahren der vollständigen Induktion kennenzulernen. Wir setzen voraus, dass die natürlichen Zahlen $\mathbb{N}$ als Folge 1, $1+1=2$, $2+1=3$, $\ldots$ gegeben sind. Ausgehend von $1 \in \mathbb{N}$ wird also jede natürliche Zahl erreicht, indem die 1 endlich oft addiert wird. Darauf beruht das

Induktionsprinzip: Sei $M \subset \mathbb{N}$ eine Menge mit den beiden Eigenschaften

(1) $1 \in M$,

(2) $n \in M \quad \Rightarrow \quad n+1 \in M$.

Dann gilt schon $M = \mathbb{N}$.

Die Beschreibung der natürlichen Zahlen als Folge $1, 2, 3, \ldots$ ist keine strenge Definition, da die Pünktchen nicht präzisiert werden. Demzufolge können wir auch das Induktionsprinzip nicht rigoros begründen, sondern nehmen es schlicht als gegeben hin. Auf der Basis der Axiome (K), (A1) und (A2) kann aber eine strenge Definition der natürlichen Zahlen als Teilmenge der reellen Zahlen gegeben werden, wobei das Induktionsprinzip dann als Satz gefolgert wird. Dies wird zum Beispiel in den Büchern von Barner & Flohr sowie Hildebrandt ausgeführt. Das Beweisverfahren der vollständigen Induktion ergibt sich direkt aus dem Induktionsprinzip.

Satz 2.1 (Beweisverfahren der vollständigen Induktion) *Gegeben sei eine Folge von Aussagen $A(n)$ für $n \in \mathbb{N}$. Es möge gelten:*

(1) $A(1)$ *ist wahr. (Induktionsanfang)*

(2) $A(n)$ *ist wahr* $\quad \Rightarrow \quad A(n+1)$ *ist wahr. (Induktionsschluß)*

Dann sind alle Aussagen $A(n)$ wahr.

Beweis: Wir betrachten die Menge $M = \{n \in \mathbb{N} : A(n) \text{ ist wahr}\}$. Nach Voraussetzung gilt $1 \in M$, und mit $n \in M$ ist auch $n+1 \in M$. Das Induktionsprinzip ergibt $M = \mathbb{N}$, das heißt alle Aussagen $A(n)$ sind wahr. $\square$

Ein Induktionsbeweis funktioniert immer in zwei Schritten:

- Induktionsanfang ($n = 1$): Beweis der Behauptung für $n = 1$.
- Induktionsschluß ($n \Rightarrow n+1$): Beweis, dass aus der Wahrheit der Behauptung für $A(n)$ (*Induktionsannahme*) die Wahrheit der Behauptung für $A(n+1)$ folgt.

Bemerkung. Statt bei $n = 1$ kann die Induktion auch bei einer anderen Zahl starten. Zum Beispiel ist zu jeder ganzen Zahl $n \geq n_0$ eine Aussage $A(n)$ gegeben. Vollstandige Induktion kann sinngemäß auch in dieser Situation angewendet werden. Als Induktionsanfang hat man $A(n_0)$ zu beweisen und der Induktionsschluß $A(n) \rightarrow A(n+1)$ ist für die $n \geq n_0$ zu erbringen.

Beispiel 2.1 (arithmetische Summe) Wir zeigen die Summenformel

$$A(n): \quad 1+2+\cdots+n=\sum_{k=1}^{n} k=\frac{n(n+1)}{2}.$$

Beweis: *Induktionsanfang.* Für $n=1$ ist sowohl die linke als auch die rechte Seite gleich Eins, also gilt der Induktionsanfang.
Induktionsschluß $(n \Rightarrow n+1)$. Jetzt berechnen wir unter Verwendung von $A(n)$

$$\sum_{k=1}^{n+1} k=\left(\sum_{k=1}^{n} k\right)+(n+1) \stackrel{A(n)}{=} \frac{n(n+1)}{2}+(n+1)=\frac{(n+1)((n+1)+1)}{2}.$$

Damit ist $A(n)$ für alle $n \in \mathbb{N}$ bewiesen. □

Nach dem neunjährigen Gauß kommen wir natürlich auch ohne Induktion zum Ziel:

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

Der Vorteil dieses Arguments ist, dass wir die Formel nicht vorher raten müssen. □

Beispiel 2.2 (geometrische Summe) Sei $x \in \mathbb{R}$, $x \neq 1$. Dann gilt für alle $n \in \mathbb{N}_0$

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

Beweis: *Induktionsanfang* $(n=0)$. Wir zeigen das wieder durch vollständige Induktion, wobei wir bei $n=0$ beginnen:

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

Induktionsschluß $(n \Rightarrow n+1)$. Jetzt gelte die Formel für ein $n \in \mathbb{N}$. Dann folgt

$$\sum_{k=0}^{n+1} x^{k}=\left(\sum_{k=0}^{n} x^{k}\right)+x^{n+1} \stackrel{A(n)}{=} \frac{1-x^{n+1}}{1-x}+x^{n+1}=\frac{1-x^{(n+1)+1}}{1-x},$$

womit die Behauptung $A(n+1)$ gezeigt ist. □

Auch hier haben wir ein alternatives Argument, nämlich den sogenannten Teleskopsummentrick:

$$1-x^{n+1}=1-x+x-\ldots-x^{n}+x^{n}-x^{n+1}=\sum_{k=0}^{n}\left(x^{k}-x^{k+1}\right)=(1-x) \sum_{k=0}^{n} x^{k}.$$

□

Ebenfalls mit dem Beweisverfahren der vollständigen Induktion zeigen wir folgende nützliche Ungleichung.

Satz 2.2 (Bernoullische Ungleichung) *Für $x \in \mathbb{R}$, $x \geq -1$, und $n \in \mathbb{N}_0$ gilt*

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

Beweis: Wir führen Induktion über $n \in \mathbb{N}_0$. (Die Induktion fäng von $n = 0$ an.) Für $n = 0$ gilt nach Definition $(1+x)^0 = 1 = 1 + 0 \cdot x$. Wegen $1 + x \geq 0$ folgt weiter

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

□

Wir wollen als nächstes die Elemente gewisser Mengen zählen.

Satz 2.3 (Schubfachprinzip) *Ist $f : \{1, \ldots, m\} \to \{1, \ldots, n\}$ injektiv, so folgt $m \leq n$.*

Beweis: Wir fassen die Behauptung als Aussage $A(n)$ auf, die jeweils für alle $m \in \mathbb{N}$ zu zeigen ist, und führen einen Induktionsbeweis.
(*IA*) Für $n = 1$ haben wir eine injektive Abbildung $f : \{1, \ldots, m\} \to \{1\}$ und es folgt sofort $m = 1$, also der Induktionsanfang.
(*IS*) Sei nun eine injektive Abbildung $f : \{1, \ldots, m\} \to \{1, \ldots, n+1\}$ gegeben. Zu zeigen ist $m \leq n+1$, was für $m = 1$ offensichtlich ist. Für $m \geq 2$ konstruieren wir eine injektive Abbildung

$$\tilde{f} : \{1, \ldots, m-1\} \to \{1, \ldots, n\},\ k \mapsto \tilde{f}(k).$$

Mit der Induktionsannahme folgt dann $m - 1 \leq n$ beziehungsweise $m \leq n + 1$ wie gewünscht.
Die Funkition $\tilde{f}$ ist wie folgt konstruiert: Im Fall $f(k) \in \{1, \ldots, n\}$ für alle $k = 1, \ldots, m-1$ können wir einfach $\tilde{f}(k) = f(k)$ setzen. Andernfalls gibt es genau ein $i \in \{1, \ldots, m-1\}$ mit $f(i) = n + 1$. Da f nach Voraussetzung injektiv ist, folgt $f(m) \neq n + 1$, das heißt $f(m) \in \{1, \ldots, n\}$, und wir können setzen

$$\tilde{f}(k) = \begin{cases} f(k) & \text{für } k = 1, \ldots, m-1,\ k \neq i, \\ f(m) & \text{für } k = i. \end{cases}$$

Es ist leicht zu sehen (*Übung*), dass in jedem der Fälle $\tilde{f}$ injektiv ist. □

Definition 2.1 (Zahl der Elemente) *Eine nichtleere Menge M heißt endlich, wenn für ein $n \in \mathbb{N}$ eine Bijektion $f : \{1, \ldots, n\} \to M$ existiert; andernfalls heißt sie unendlich. Im endlichen Fall heißt n Anzahl der Elemente von M (Symbol: $\#M = n$). Die leere Menge wird ebenfalls als endlich bezeichnet mit $\#\emptyset = 0$.*

Lemma 2.1 *Die Anzahl einer endlichen Menge ist* wohldefiniert

Beweis: Wir mussen zeigen dass die Zahl n mit einer Bijektion $f : \{1, \ldots, n\} \to M$ eindeutig bestimmt ist. Denn ist $\tilde{f} : \{1, \ldots, m\} \to M$ ebenfalls bijektiv, so haben wir die bijektiven, insbesondere injektiven, Abbildungen

$$\tilde{f}^{-1} \circ f : \{1, \ldots, n\} \to \{1, \ldots, m\} \quad \text{sowie} \quad f^{-1} \circ \tilde{f} : \{1, \ldots, m\} \to \{1, \ldots, n\}.$$

Aus dem Schubfachprinzip, Satz 2.3, folgt dann $n \leq m$ und $m \leq n$, also $m = n$. $\square$

Der Satz garantiert also, dass die scheinbare Uneindeutigkeit in der Definition der Anzahl nicht vorhanden ist. Die Mathematiker haben dafür die schöne Formulierung, die Anzahl sei *wohldefiniert*.

Das Produkt $n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot \ldots \cdot n$ wird als *n-Fakultät* bezeichnet; dabei ist per Definition $0! = 1$ (in Konsistenz mit unserer Vereinbarung zum leeren Produkt). Der folgende Satz beantwortet die Frage nach der Anzahl der möglichen Anordnungen (oder Umordnungen oder Permutationen) von n Dingen.

Satz 2.4 (Zahl der Permutationen) *Für $n \in \mathbb{N}$ sei S_n die Menge der bijektiven Abbildungen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. Dann gilt $\#S_n = n!$.*

Beweis: Wir zeigen die Behauptung durch Induktion, wobei der Induktionsanfang $n = 1$ offensichtlich ist. Es ist praktisch, jedes $\sigma \in S_n$ mit dem n-Tupel $(\sigma_1, \ldots, \sigma_n)$ zu identifizieren, wobei $\sigma_i = \sigma(i)$. Die Menge S_{n+1} ist disjunkte Vereinigung der Teilmengen

$$S_{n+1,k} = \{\tau \in S_{n+1} : \tau_k = n+1\} \quad \text{für } k = 1, \ldots, n+1.$$

Beispielsweise ist in aufzählender Form

$$S_{4,2} = \{(1,4,2,3), (2,4,3,1), (3,4,1,2), (1,4,2,3), (2,4,1,3), (3,4,1,2)\}.$$

Jedem $\sigma = (\sigma_1, \ldots, \sigma_n) \in S_n$ können wir die Permutation $(\sigma_1, \ldots, \sigma_{k-1}, n+1, \sigma_k, \ldots, \sigma_n)$ in $S_{n+1,k}$ zuordnen, und diese Abbildung ist bijektiv (*nachprüfen!*). Also folgt aus der Induktionsannahme $\#S_{n+1,k} = \#S_n = n!$ und

$$\#S_{n+1} = \sum_{k=1}^{n+1} \#S_{n+1,k} = (n+1) \cdot n! = (n+1)!.$$

$\square$

Definition 2.2 (Binomialkoeffizienten) *Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$ setzen wir*

$$\binom{\alpha}{k} = \frac{\alpha \cdot (\alpha - 1) \cdot \ldots \cdot (\alpha - k + 1)}{1 \cdot 2 \cdot \ldots \cdot k}, \qquad \textit{sowie} \quad \binom{\alpha}{0} = 1.$$

Lemma 2.2 (Additionstheorem für Binomialkoeffizienten) *Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$ erfüllen die Binomialkoeffizienten die Formel*

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

Beweis: Für $k = 1$ ist leicht zu sehen, dass die Formel richtig ist. Für $k \geq 2$ berechnen wir

$$\begin{aligned}
\binom{\alpha}{k} + \binom{\alpha}{k-1} &= \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} + \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+2)}{1 \cdot 2 \cdot \ldots \cdot (k-1)} \\
&= \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+2) \cdot (\alpha-k+1+k)}{1 \cdot 2 \cdot \ldots \cdot k} \\
&= \frac{(\alpha+1) \cdot \alpha \cdot \ldots \cdot ((\alpha+1)-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} \\
&= \binom{\alpha+1}{k}.
\end{aligned}$$

□

Im Fall $\alpha = n \in \mathbb{N}_0$ erlaubt Lemma 2.2 die rekursive Berechnung der Binomialkoeffizienten $\binom{n}{k}$ nach dem Dreiecksschema von Blaise Pascal (1623-1662). (Das Pascalsche Dreieck nennt man auch Dreiecke von Yang Hui, vor 1303)

n=0							1							
n=1						1		1						
n=2					1		2		1					
n=3				1		3		3		1				
n=4			1		4		6		4		1			
n=5		1		5		10		10		5		1		
n=6	1		6		15		20		15		6		1	

Ebenfalls für $\alpha = n \in \mathbb{N}_0$ folgt durch Erweitern der Binomialkoeffizienten mit $(n-k)!$ die alternative Darstellung

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \text{ für } n \in \mathbb{N}_0,\ k \in \{0, 1, \ldots, n\}, \tag{2.4}$$

und daraus weiter die am Diagramm ersichtliche Symmetrieeigenschaft

$$\binom{n}{k} = \binom{n}{n-k} \text{ für } n \in \mathbb{N}_0,\ k \in \{0, 1, \ldots, n\}. \tag{2.5}$$

Satz 2.5 (Zahl der Kombinationen) *Sei $n \in \mathbb{N}_0$ und $k \in \{0, 1, \ldots, n\}$. Dann ist die Anzahl der k-elementigen Teilmengen von $\{1, \ldots, n\}$ gleich $\binom{n}{k}$.*

Beweis: Die Behauptung gilt für $k = 0$ und beliebiges n, denn die leere Menge ist die einzige null-elementige Teilmenge von $\{1, \ldots, n\}$, und nach Definition ist $\binom{n}{0} = 1$. Insbesondere gilt die Behauptung für $n = 0$. Wir führen nun Induktion über n, wobei wir die Behauptung jeweils für alle $k \in \{0, 1, \ldots, n\}$ zeigen. Im Induktionsschluss müssen wir die Anzahl der k-elementigen Teilmengen von $\{1, \ldots, n+1\}$ bestimmen, wobei wir $k \geq 1$ annehmen können. Diese Teilmengen zerfallen in zwei disjunkte Klassen:

Klasse 1: Die Menge enthält die Nummer $n+1$ nicht.

Klasse 2: Die Menge enthält die Nummer $n+1$.

Klasse 1 besteht genau aus den k-elementigen Teilmengen von $\{1,\ldots,n\}$, Klasse 2 ergibt sich durch Hinzufügen des Elements $n+1$ zu jeder der $(k-1)$-elementigen Teilmengen von $\{1,\ldots,n\}$. Insgesamt ist die Zahl der k-elementigen Teilmengen von $\{1,\ldots,n+1\}$ nach Induktionsannahme also gleich

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

nach Lemma 2.2, womit der Satz bewiesen ist. □

Satz 2.6 (Binomische Formel) *Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}. \tag{2.6}$$

Beweis: Ausmultiplizieren des n-fachen Produkts $(a+b)^n = (a+b)\cdot(a+b)\cdot\ldots\cdot(a+b)$ mit dem Distributivgesetz und Ordnen nach dem Kommutativgesetz liefert Terme der Form $a^k b^{n-k}$ für $k \in \{0,1,\ldots,n\}$. Die Häufigkeit eines solchen Terms ist gleich der Anzahl der Möglichkeiten, aus den n Klammern k Klammern auszusuchen, in denen a als Faktor genommen wird; in den restlichen Klammern muss dann der Faktor b gewählt werden. Nach Satz 2.5 kommt $a^k b^{n-k}$ also genau $\binom{n}{k}$ mal vor. □

Alternativ folgt der Satz auch durch vollständige Induktion über n, und zwar gilt der Induktionsanfang $n=1$ wegen

$$n=1: \quad (a+b)^1 = \binom{1}{0} a^0\, b^1 + \binom{1}{1} a^1 b^0.$$

Der Induktionsschluss ergibt sich wie folgt:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b) \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \qquad \text{(nach Induktionsannahme)} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} a^k\, b^{n-(k-1)} + \sum_{k=1}^{n} \binom{n}{k} a^k b^{n+1-k} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] a^k b^{n+1-k} + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k} \qquad \text{nach Lemma 2.2.}
\end{aligned}$$

Die folgende Umformulierung des Induktionsprinzips erscheint absolut selbstverständlich; wir erinnern aber daran, dass wir keine strenge Definition der natürlichen Zahlen gegeben haben.

Satz 2.7 (Prinzip der kleinsten natürlichen Zahl) *Jede nichtleere Menge $M \subset \mathbb{N}$ besitzt ein kleinstes Element.*

BEWEIS: Wir beweissen dem Satz durch eine Widerspruch-Argumente. Angenommen, es gibt kein kleinstes Element in $M \subset \mathbb{N}$. Wie zeigen dann durch Induktion $\{1, \ldots, n\} \cap M = \emptyset$. Für $n = 1$ ist das richtig, denn sonst wäre $1 \in M$ das kleinste Element. Ist die Behauptung für $n \in \mathbb{N}$ gezeigt, so gilt sie auch für $n + 1$, denn sonst wäre $n + 1$ kleinstes Element in M. Also folgt $\{1, \ldots, n\} \cap M = \emptyset$ für alle $n \in \mathbb{N}$. Aber dann ist M die leere Menge im Widerspruch zur Voraussetzung. $\square$

Als Anwendung zeigen wir nun, dass die rationalen Zahlen auf der Zahlengeraden dicht verteilt sind. Dafür brauchen wir nun das Archimedisches Axiom (A3). Wir holen das hier wieder.

(A3) Zu jedem $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, existiert ein $n \in \mathbb{N}$ mit $1/n < \varepsilon$ (*Archimedisches Axiom*).

Es gibt dann auch zu jedem $K \in \mathbb{R}$, $K > 0$, ein $n \in \mathbb{N}$ mit $n > K$, wie sich mit der Wahl $\varepsilon = 1/K > 0$ in (A3) sofort ergibt.

Satz 2.8 ($\mathbb{Q}$ ist dicht in $\mathbb{R}$) *Zu je zwei reellen Zahlen $a, b \in \mathbb{R}$ mit $a < b$ gibt es eine rationale Zahl $q \in \mathbb{Q}$ mit $a < q < b$.*

BEWEIS: Wir können $b > 0$ annehmen, denn sonst gehen wir zu $a' = -b$, $b' = -a$ über. Ausserdem ist o.B.d.A. $a \geq 0$, da wir andernfalls einfach $q = 0$ setzen. Nach A3 existiert ein $n \in \mathbb{N}$ mit $1/n < b - a$. Betrachte nun die Menge

$$M = \{k \in \mathbb{N} : k/n > a\} \subset \mathbb{N}.$$

M ist nichtleer: falls $a = 0$ so ist zum Beispiel $1 \in M$, andernfalls gibt es nach A3 ein $k \in \mathbb{N}$ mit $1/k < 1/na$ bzw. $k/n > a$. Sei $m \in M$ das kleinste Element nach Satz 2.7. Dann ist einerseits $m/n > a$ und andererseits $(m - 1)/n \leq a$, also

$$a < \frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} < a + (b - a) = b.$$

Die Zahl $q = m/n$ leistet somit das Verlangte. $\square$

Ebenso wichtig wie der Beweis durch vollstädige Induktion ist die Konstruktion durch vollständige Induktion, *rekursive Definition* genannt. Es soll jeder natürlichen Zahl n ein Element $f(n)$ einer Menge X zugeordnet werden durch

(I) die Angabe von $f(1)$ und

(II) eine Vorschrift F, die für jedes $n \in \mathbb{N}$ ein Element $f(n + 1)$ aus den Elementen $f(1), f(2), \cdots, f(n)$ zu berechnen gestattet:

$$f(n + 1) = F(f(1), \cdots, f(n)).$$

(Rekursionsvorschrift)

Beispiel 2.3 *Man erklärt die Potenzen einer Zahl durch*

(I) $x^1 := x$ *und*

(II) die Rekursionsformel $x^{n+1} := x^n \cdot x$ *für jedes* $n \in \mathbb{N}$.

Kapitel 2

Konvergenz

1 Grenzwerte von Folgen

Definition 1.1 *Eine Folge reeller Zahlen ist eine Abbildung*

$$\mathbb{N} \to \mathbb{R}, \quad n \mapsto a_n.$$

Die Zahl a_n heißt das *n-te Glied* der Folge, die Folge insgesamt wird mit $(a_n)_{n\in\mathbb{N}}$ bzw. kurz mit (a_n) bezeichnet. Oft wird die Folge durch das Bildungsgesetz angegeben, durch Aufzählen der ersten Folgenglieder oder durch die rekursive Definition definiert.
Zum Beispiel ist die Folge der Quadratzahlen gegeben durch $a_n = n^2$, bzw. alternativ aufzählend $a_n = 1, 4, 9, 16, \cdots$.

Beispiel 1.1 (Folge der Fibonaccizahlen) *Ist $a_1 = 0$ und $a_2 = 1$, und für $n \geq 3$ ist a_n durch die Rekursionsvorschrift $a_n := a_{n-1} + a_{n-2}$ gegeben.*

Definition 1.2 (Konvergenz von Folgen) *Die Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert mit $n \to \infty$ gegen $a \in \mathbb{R}$, falls gilt:*

Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{R}$, so dass für alle $n > N$ gilt: $|a_n - a| < \varepsilon$.

Die Zahl a heißt dann Grenzwert *der Folge und wir schreiben kurz*

$$\lim_{n\to\infty} a_n = a \quad \text{oder} \quad a_n \to a \text{ für } n \to \infty.$$

(gelesen: a_n strebt gegen a für n gegen unendlich) Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt konvergent, wenn es ein $a \in \mathbb{R}$ gibt, das Grenzwert der Folge ist; andernfalls heißt die Folge divergent.

Mit den Quantoren $\forall$ (für alle), $\exists$ (existiert) und $\Rightarrow$ (daraus folgt) lässt sich die Definition der Konvergenz auch wie folgt fassen:

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{R} : \quad \Big(n > N \quad \Rightarrow \quad |a_n - a| < \varepsilon \Big).$$

Beispiel 1.2 (Harmonische Folge) Die Folge $a_n = 1/n$ konvergiert gegen $a = 0$. Denn zu gegebenem $\varepsilon > 0$ wählen wir $N = 1/\varepsilon$, und es folgt für alle $n > N$

$$|a_n - a| = |1/n - 0| = 1/n < 1/N = \varepsilon.$$

□

Je kleiner die geforderte Abweichung $\varepsilon > 0$ vom Grenzwert und damit die Genauigkeit der Approximation sein soll, desto größer muss im allgemeinen die Zahl N in der Definition des Grenzwerts gewählt werden, das heißt wie in obigem Beispiel hängt N von ε ab, $N = N(\varepsilon)$. Eine Ausnahme bildet hier nur die konstante Folge.

Beispiel 1.3 (Konstante Folge) Ist $a_n = a$ für alle $n \in \mathbb{N}$, so folgt $\lim_{n\to\infty} a_n = a$. Denn für $\varepsilon > 0$ gilt $|a_n - a| = 0 < \varepsilon$ für alle $n > 0$, also können wir immer $N = 0$ wählen. □

Übrigens ist es egal, ob in der Definition des Grenzwerts statt $N \in \mathbb{R}$ die Bedingung $N \in \mathbb{N}$ verlangt wird, denn wir können statt $N \in \mathbb{R}$ ja immer die nächstgrößere natürliche Zahl nehmen. Überhaupt kann N immer vergrößert werden, und in der Regel besteht kein Interesse daran, dass kleinstmögliche $n_\varepsilon \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ für $n \geq n_\varepsilon$ zu finden. Dies ist natürlich anders, wenn ein Grenzwert numerisch berechnet werden soll, aber für den Nachweis der Konvergenz reicht es völlig, irgendeine Schranke zu finden, von der ab die Ungleichung gilt.

Beispiel 1.4 (Geometrische Folge) Sei $q \in \mathbb{R}$ mit $|q| < 1$. Dann gilt $\lim_{n\to\infty} q^n = 0$. Um das zu zeigen, können wir $q \neq 0$ voraussetzen und haben dann $1/|q| > 1$, also gilt $1/|q| = 1 + x$ für ein $x > 0$. Es folgt mit der Bernoulli-Ungleichung, Satz 2.2,

$$|q^n - 0| = |q|^n = \frac{1}{(1+x)^n} \leq \frac{1}{1+nx} \leq \frac{1}{nx} < \varepsilon$$

für alle $n > 1/(\varepsilon x)$. Wir können also $N = 1/(\varepsilon x)$ wählen. □

Beispiel 1.5 Die Folge $a_n = (-1)^n$, also $a_n = -1, 1, -1, \ldots$ ist nicht konvergent. Denn angenommen es wäre $\lim_{n\to\infty} a_n = a$ für ein $a \in \mathbb{R}$. Zu $\varepsilon = 1$ gibt es dann ein $N \in \mathbb{R}$ mit $|a_n - a| < 1$ für $n > N$, also gilt für $n > N$

$$2 = |a_n - a_{n+1}| = |a_n - a + a - a_{n+1}| \leq |a_n - a| + |a - a_{n+1}| < 1 + 1 = 2,$$

ein Widerspruch. □

Der Begriff des Grenzwerts wird anschaulicher, indem wir folgende Teilmengen von $\mathbb{R}$ einführen.

Definition 1.3 (ε-Umgebung) *Die ε-Umgebung von $a \in \mathbb{R}$ ist die Menge*

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

Eine Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert genau dann gegen $a \in \mathbb{R}$, wenn die Folgenglieder ab einer gewissen Nummer in der ε-Umgebung von a liegen, egal wie klein $\varepsilon > 0$ gewählt ist.

Satz 1.1 (Eindeutigkeit des Grenzwerts) *Falls die Folge $(a_n)_{n\in\mathbb{N}}$ konvergent ist, so ist ihr Grenzwert eindeutig bestimmt.*

Beweis: Wir beginnen mit einer Vorüberlegung, und zwar behaupten wir

$$0 < \varepsilon \leq \frac{1}{2}|a - a'| \quad \Rightarrow \quad U_\varepsilon(a) \cap U_\varepsilon(a') = \emptyset. \tag{1.1}$$

Denn ist $x \in U_\varepsilon(a) \cap U_\varepsilon(a')$, so folgt mit der Dreiecksungleichung

$$|a - a'| = |a - x + x - a'| \le |x - a| + |x - a'| < 2\varepsilon.$$

Seien nun $a, a' \in \mathbb{R}$ Grenzwerte der Folge $(a_n)_{n\in\mathbb{N}}$. Zu jedem $\varepsilon > 0$ gibt es dann $N, N' \in \mathbb{R}$ mit $a_n \in U_\varepsilon(a)$ für $n > N$, sowie $a_n \in U_\varepsilon(a')$ für $n > N'$. Wäre $a \neq a'$, so wählen wir $\varepsilon = \frac{1}{2}|a - a'| > 0$ und erhalten für $n > \max(N, N')$

$$a_n \in U_\varepsilon(a) \cap U_\varepsilon(a') = \emptyset,$$

ein Widerspruch. □

Unser nächstes Ziel ist es, einige Rechenregeln für Grenzwerte zu erarbeiten. Wir beginnen mit der

Definition 1.4 (Beschränktheit von Folgen) *Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt*

a) nach oben (bzw. nach unten) beschränkt, wenn es ein $K \in \mathbb{R}$ gibt mit $a_n \le K$ (bzw. $a_n \ge K$) für alle $n \in \mathbb{N}$.

b) beschränkt, wenn sie nach oben und nach unten beschränkt ist.

Beispiel 1.6 Die Folge $a_n = n$ ist nach unten beschränkt, denn es ist zum Beispiel $a_n \ge 0$ für alle n. Sie ist aber nicht nach oben beschränkt: angenommen, es gibt ein $K \in \mathbb{R}$ mit $a_n \le K$ für alle $n \in \mathbb{N}$. Dann ist $K \ge a_1 = 1 > 0$, also auch $1/K > 0$, und nach Archimedes gibt es ein $n \in \mathbb{N}$ mit $1/n < 1/K$, also $a_n = n > K$, ein Widerspruch.

Satz 1.2 (konvergent ⇒ beschränkt) *Jede konvergente Folge ist beschänkt.*

Beweis: Sei $\lim_{n\to\infty} a_n = a$. Wähle zu $\varepsilon = 1$ ein $N \in \mathbb{R}$ mit $|a_n - a| < 1$ für $n > N$. Wir können $N \in \mathbb{N}$ annehmen, andernfalls ersetzen wir N durch die nächstgrößere natürliche Zahl. Es gilt dann

$$\begin{aligned} n > N \quad &\Rightarrow \quad |a_n| = |a_n - a + a| \le |a_n - a| + |a| < 1 + |a| \\ n \le N \quad &\Rightarrow \quad |a_n| \le \max(|a_1|, \ldots, |a_N|). \end{aligned}$$

Wir haben also $|a_n| \le K$ für alle n, wobei $K = \max(|a_1|, \ldots, |a_N|, 1 + |a|)$. □

Satz 1.3 (Rechenregeln für Grenzwerte) *Es gelte $a_n \to a, b_n \to b$ mit $n \to \infty$.*

a) *Für alle $\lambda, \mu \in \mathbb{R}$ ist $(\lambda a_n + \mu b_n)_{n\in\mathbb{N}}$ konvergent mit Grenzwert $\lim_{n\to\infty}(\lambda a_n + \mu b_n) = \lambda a + \mu b$.*

b) *Die Folge $(a_n \cdot b_n)_{n\in\mathbb{N}}$ ist konvergent mit Grenzwert $\lim_{n\to\infty}(a_n \cdot b_n) = a \cdot b$.*

c) *Falls $b \neq 0$, so gibt es ein $N_0 \in \mathbb{R}$ mit $b_n \neq 0$ für $n > N_0$ und die Folge $(a_n/b_n)_{n>N_0}$ ist konvergent mit Grenzwert $\lim_{n\to\infty} a_n/b_n = a/b$.*

Beweis: Wir beginnen mit dem Beweis von b). Nach Satz 1.2 gibt es ein $K > 0$ mit $|a_n| \le K$ für alle $n \in \mathbb{N}$, und außerdem mit $|b| \le K$. Dann gilt für alle $n \in \mathbb{N}$

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| \\ &\le |a_n| \cdot |b_n - b| + |a_n - a| \cdot |b| \\ &\le K(|a_n - a| + |b_n - b|). \end{aligned}$$

Zu $\varepsilon > 0$ gibt es nun ein $N \in \mathbb{R}$ mit $|a_n - a| < \varepsilon/(2K)$ sowie $|b_n - a| < \varepsilon/(2K)$ für $n > N$. Also folgt für $n > N$

$$|a_n b_n - ab| < K\left(\frac{\varepsilon}{2K} + \frac{\varepsilon}{2K}\right) = \varepsilon.$$

Für a) reicht es wegen b), den Fall $\lambda = \mu = 1$ zu betrachten. Zu $\varepsilon > 0$ gibt es ein $N \in \mathbb{R}$ mit $|a_n - a| < \varepsilon/2$ und $|b_n - b| < \varepsilon/2$ für $n > N$. Es folgt für $n > N$

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

Für c) können wir uns auf den Fall $a_n = b = 1$ beschränken, denn sonst schreiben wir

$$\frac{a_n}{b_n} = \frac{a_n}{b} \cdot \frac{1}{b'_n} \text{ mit } b'_n = \frac{b_n}{b}$$

und wenden b) an. Es gibt nun ein $N_0 \in \mathbb{R}$ mit $|b_n - 1| \leq \frac{1}{2}$ für $n > N_0$, also

$$|b_n| = |1 - (1 - b_n)| \geq 1 - |1 - b_n| \geq \frac{1}{2} > 0.$$

Damit ist die erste Aussage gezeigt. Zu $\varepsilon > 0$ wähle weiter $N \geq N_0$ mit $|1 - b_n| < \varepsilon/2$ für $n > N$, und somit

$$\left|\frac{1}{b_n} - 1\right| = \frac{|1 - b_n|}{b_n} < 2 \cdot \frac{\varepsilon}{2} = \varepsilon.$$

□

Bemerkung: Für jede Menge X ist die Menge der Funktionen $f : X \to \mathbb{R}$ ein $\mathbb{R}$-Vektorraum, mit der Addition $(f + g)(x) = f(x) + g(x)$ und der Skalarmultiplikation $(\lambda f)(x) = \lambda f(x)$. Dies gilt auch für $X = \mathbb{N}$, das heißt die Menge der reellen Zahlenfolgen $(a_n)_{n\in\mathbb{N}}$ ist ein $\mathbb{R}$-Vektorraum, wobei die Addition von $a = (a_n)_{n\in\mathbb{N}}$ und $b = (b_n)_{n\in\mathbb{N}}$ sowie die Skalarmultiplikation mit $\lambda \in \mathbb{R}$ wie folgt gegeben sind:

$$(a + b)_n = a_n + b_n \quad \text{und} \quad (\lambda a)_n = \lambda a_n \quad \text{für alle } n \in \mathbb{N}.$$

Überlegen Sie, dass die nachstehenden Mengen Untervektorräume bilden, die der Reihe nach ineinander enthalten sind:

Nullfolgen	Die Menge aller $(a_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} a_n = 0$;
Konvergente Folgen	Die Menge aller $(a_n)_{n\in\mathbb{N}}$, so dass $\lim_{n\to\infty} a_n$ existiert;
Beschränkte Folgen	Die Menge aller $(a_n)_{n\in\mathbb{N}}$, für die ein $K \geq 0$ existiert mit $|a_n| \leq K$ für alle $n \in \mathbb{N}$.

Hier zwei Anwendungen der Rechenregeln für Grenzwerte.

Beispiel 1.7 (Grenzwerte rationaler Funktionen) Seien $p, q : \mathbb{R} \to \mathbb{R}$ Polynome vom Grad $m, n \in \mathbb{N}_0$, das heißt für alle $x \in \mathbb{R}$ gilt

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \ldots + a_0 \quad \text{und} \quad q(x) = b_n x^n + b_{n-1} x^{n-1} + \ldots + b_0,$$

wobei $a_i, b_j \in \mathbb{R}$ mit $a_m, b_n \neq 0$. Wir bestimmen im Fall $m \leq n$ das Verhalten von $p(k)/q(k)$ für $k \to \infty$, und zwar liefert mehrfache Anwendung der Konvergenzregeln in Satz 1.3

$$\frac{p(k)}{q(k)} = k^{m-n} \frac{a_m + a_{m-1}k^{-1} + \ldots + a_0 k^{-m}}{b_n + b_{n-1}k^{-1} + \ldots + b_0 k^{-n}} \to \begin{cases} a_m/b_n & \text{falls } m = n, \\ 0 & \text{falls } m < n. \end{cases}$$

Beispiel 1.8 (geometrische Reihe) Für $-1 < q < 1$ betrachten wir die Folge

$$a_n = 1 + q + \ldots + q^n = \sum_{k=0}^{n} q^k.$$

Dann ergibt sich aus Beispiel 2.2, Beispiel 1.4 und Satz 1.3

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

Wir schreiben hierfür auch $\sum_{k=0}^{\infty} q^k = 1/(1-q)$.

Definition 1.5 *Folgen, deren Folgenglieder Summen sind, heißen Reihen.*

Die Reihen spielen eine große Rolle in der Analysis und werden in Kürze ausführlicher untersucht.

Satz 1.4 (Grenzwerte und Ungleichungen) *Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergent, mit Grenzwerten $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann gelten folgende Aussagen:*

a) *Ist $a_n \le b_n$ für alle n, so folgt $a \le b$.*

b) *Gilt $c \le a_n \le d$ für alle n mit $c, d \in \mathbb{R}$, so folgt $c \le a \le d$.*

c) *Ist $a_n \le c_n \le b_n$ und gilt $a = b$, so konvergiert auch die Folge $(c_n)_{n\in\mathbb{N}}$ gegen $a = b$.*

Beweis: Nach Voraussetzung gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{R}$ mit $a_n > a - \varepsilon$ und $b_n < b + \varepsilon$ für alle $n > N$. Die Voraussetzung in a) liefert dann $a - \varepsilon < b + \varepsilon$ beziehungsweis $(a-b)/2 < \varepsilon$ für jedes $\varepsilon > 0$, also $a \le b$. Aussage b) folgt unmittelbar aus a), indem wir c, d als konstante Folgen auffassen. Unter den Voraussetzungen in c) folgt für $n > N$ die Ungleichungskette

$$a - \varepsilon < a_n \le c_n \le b_n < b + \varepsilon = a + \varepsilon,$$

also $\lim_{n\to\infty} c_n = a$ nach Definition des Grenzwerts. □

Achtung: aus $a_n < b_n$ folgt *nicht* $a < b$, sondern nur $a \le b$. Die Striktheit von Ungleichungen geht beim Übergang zu Grenzwerten im allgemeinen verloren. Zum Beispiel gilt $1/n > 0$ für alle $n \in \mathbb{N}$, aber $\lim_{n\to\infty} 1/n = 0$.

Beispiel 1.9 (n-te Wurzel) Hier betrachten wir für $a > 0$ und $n \in \mathbb{N}$ die Gleichung $x^n = a$. Es gibt höchstens eine Lösung $x > 0$, denn für $x, y > 0$ mit $x > y$ folgt $x^n > y^n$, oder

$$x, y > 0 \text{ und } x^n \le y^n \quad \Rightarrow \quad x \le y.$$

Die Existenz der Lösung wird im nächsten Kapitel aus dem Vollständigkeitsaxiom hergeleitet. Damit gibt es genau eine Lösung $x > 0$, die mit $a^{1/n}$ oder $\sqrt[n]{a}$ bezeichnet wird, und es gilt

$$a, b > 0 \text{ und } a \le b \quad \Rightarrow \quad a^{1/n} \le b^{1/n}. \tag{1.2}$$

Wir behaupten nun

$$\lim_{n\to\infty} a^{1/n} = 1.$$

Für $a \geq 1$ ist $a^{1/n} \geq 1$ nach (1.2). Wir setzen $\xi_n = a^{1/n} - 1 \geq 0$ und schließen aus der Bernoulli-Ungleichung, Satz 2.2,

$$a = (1 + \xi_n)^n \geq 1 + n\xi_n \quad \Rightarrow \quad 0 \leq \xi_n \leq (a-1)/n.$$

Also folgt $\lim_{n\to\infty} \xi_n = 0$ bzw. $\lim_{n\to\infty} a^{1/n} = 1$ mit Satz 1.4 c). Für $a < 1$ gilt

$$\left(a^{1/n} \cdot (a^{-1})^{1/n}\right)^n = a \cdot a^{-1} = 1 \quad \Rightarrow \quad a^{1/n} = \frac{1}{(a^{-1})^{1/n}},$$

und die Behauptung folgt aus dem vorigen Fall mit Satz 1.3 c). □

Definition 1.6 (Uneigentliche Konvergenz) *Die Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert uneigentlich (oder divergiert bestimmt) gegen $+\infty$, falls gilt:*

Zu jedem $K > 0$ gibt es ein $N \in \mathbb{R}$, so dass $a_n > K$ für alle $n > N$.

Wir schreiben $\lim_{n\to\infty} a_n = +\infty$ oder $a_n \to +\infty$ mit $n \to \infty$. Uneigentliche Konvergenz gegen $-\infty$ ist analog definiert.

Beispiel 1.10 Für $q > 1$ gilt $\lim_{n\to\infty} q^n = +\infty$. Denn zu gegebenem $K > 0$ gibt es nach Beispiel 1.4 ein $N \in \mathbb{R}$ mit $(1/q)^n < 1/K$ für $n > N$, also $q^n > K$ für $n > N$. Insgesamt haben wir für das Verhalten der Folge q^n mit $n \to \infty$ folgende Tabelle:

$$\begin{array}{cl} q > 1 & \Rightarrow \ \lim_{n\to\infty} q^n = +\infty, \\ q = 1 & \Rightarrow \ \lim_{n\to\infty} q^n = 1, \\ -1 < q < 1 & \Rightarrow \ \lim_{n\to\infty} q^n = 0, \\ q \leq -1 & \Rightarrow \ (q^n) \text{ nicht konvergent.} \end{array}$$

Der Fall $-1 < q < 1$ wurde in Beispiel 1.4 behandelt, und der Fall $q \leq -1$ folgt mit etwas Überlegung aus Beispiel 1.5 (Übungsaufgabe).

Für eine Folge mit $a_n > 0$ für alle n ist $\lim_{n\to\infty} a_n = +\infty$ äquivalent zu $\lim_{n\to\infty} 1/a_n = 0$ (Übungsaufgabe). Zum Schluss dieses Kapitels führen wir noch folgende Bezeichnungen für Teilmengen von $\mathbb{R}$ ein:

$(a,b) = \{x \in \mathbb{R} : a < x < b\}$	offenes Intervall
$[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}$	abgeschlossenes Intervall
$[a,b) = \{x \in \mathbb{R} : a \leq x < b\}$	rechtsseitig offen, linksseitig abgeschlossen
$(a,b] = \{x \in \mathbb{R} : a < x \leq b\}$	linksseitig offen, rechtsseitig abgeschlossen
$\lvert I \rvert = b - a$ für ein Intervall I	Intervalllänge

Hierbei sind $+\infty$ und $-\infty$ als offene Intervallgrenzen zugelassen, zum Beispiel ist $(-\infty, 1] = \{x \in \mathbb{R} : -\infty < x \leq 1\}$. Den ε-Umgebungen bei der Definition der Konvergenz entsprechen bei uneigentlicher Konvergenz gegen $+\infty$ die Intervalle $(K, +\infty)$: ab einem gewissen Index müssen alle Folgenglieder in $(K, +\infty)$ liegen, egal wie groß K gewählt ist.

Satz 1.5 (Konvergenz von Kehrwerten) *Für eine Folge $(a_n)_{n\in\mathbb{N}}$ gilt:*

(1) *Aus $a_n \to +\infty$ (bzw. $a_n \to -\infty$) folgt $1/a_n \to 0$.*

(2) *Aus $a_n \to 0$ und $a_n > 0$ (bzw. $a_n < 0$) folgt $1/a_n \to +\infty$ (bzw. $1/a_n \to -\infty$).*

Beweis: Der Beweis wird in den Anwesenheitsübungen besprochen. □

2 Vollständigkeit der reellen Zahlen

Die bisher eingeführten Axiome (K) sowie (A1) bis (A3) gelten selbstverständlich auch für die rationalen Zahlen. Dennoch sind die rationalen Zahlen für die Analysis ungeeignet. Wir beginnen mit folgender Beobachtung der Pythagoräer.

Satz 2.1 (Irrationalität von $\sqrt{2}$) *Die Gleichung $x^2 = 2$ ist in $\mathbb{Q}$ nicht lösbar.*

Beweis: (durch Widerspruch) Angenommen, die Gleichung $x^2 = 2$ hat eine rationale Lösung, also $x = p/q$ mit $p, q \in \mathbb{N}$. Durch fortgesetztes Kürzen können wir annehmen, dass höchstens eine der Zahlen p und q gerade ist. Nun gilt

$$p^2 = 2q^2 \quad \Rightarrow \quad p^2 \text{ gerade} \quad \Rightarrow \quad p \text{ gerade} \quad \Rightarrow \quad p = 2p_1 \text{ mit } p_1 \in \mathbb{N}.$$

Daraus folgt weiter

$$2q^2 = 4p_1^2 \quad \Rightarrow \quad q^2 \text{ gerade} \quad \Rightarrow \quad q \text{ gerade}\,.$$

Also sind doch p, q beide gerade, ein Widerspuch. □

In $\mathbb{R}$ ist die Gleichung $x^2 = 2$ und allgemeiner die Gleichung $x^n = a$ für beliebige $n \in \mathbb{N}$, $a > 0$ lösbar. Dies könnte man als Grund für die Erweiterung $\mathbb{Q} \subset \mathbb{R}$ anführen, ähnlich wie die Erweiterungen $\mathbb{N} \subset \mathbb{Z}$ beziehungsweise $\mathbb{Z} \subset \mathbb{Q}$ durch die Lösbarkeit von Gleichungen motiviert waren. Dies geht aber am Kern der Sache vorbei: einerseits bleibt die Gleichung $x^2 = -1$ unlösbar, andererseits bilden die reellen Nullstellen von beliebigen Polynomen mit rationalen Koeffizienten nur eine relativ kleine Teilmenge von $\mathbb{R}$, wie wir noch zeigen werden.

Das Ziel der Analysis ist es, neue Objekte – Zahlen, Funktionen, Operationen – durch Grenzprozesse zu konstruieren. Unsere Definition des Grenzwerts setzt voraus, dass wir den Grenzwert der Folge bereits kennen. Das Vollständigkeitsaxiom muss die Existenz von Grenzwerten in Situationen garantieren, in denen der Grenzwert a priori nicht bekannt ist. Wir betrachten hierzu zwei Beispiele.

Beispiel 2.1 (Zinseszinsrechnung) Wird ein Euro für ein Jahr mit einem Zinssatz x angelegt, so beträgt die Ausszahlung $E_1(x) = 1 + x$. Die Idee des Zinseszinses ist es, den Zeitraum in kürzere Abschnitte zu unterteilen und den Zins anteilig pro Abschnitt anzurechenen mit dem Effekt, dass der schon angerechnete Teil des Zinses seinerseits Zinsen produziert. Zum Beispiel ergibt das bei monatlicher Verzinsung nach einem Monat $1 + \frac{x}{12}$, nach zwei Monaten $(1 + \frac{x}{12})(1 + \frac{x}{12}) = (1 + \frac{x}{12})^2$, und nach zwölf Monaten $E_{12}(x) = (1 + \frac{x}{12})^{12}$. Allgemein ergibt sich nach einem Jahr bei Unterteilung in n Zeiteinheiten

$$E_n(x) = \Big(1 + \frac{x}{n}\Big)^n \qquad \text{für } x \in \mathbb{R},\, n \in \mathbb{N}. \tag{2.1}$$

Es stellt sich ganz natürlich die Frage nach einer kontinuierlichen Verzinsung, also nach dem Grenzwert $\lim_{n\to\infty} E_n(x)$.

Beispiel 2.2 (Dezimalbruchdarstellung) Für $a \in \mathbb{R}$ gibt es $k_0 \in \mathbb{Z}$ und $k_j \in \{0, 1, \ldots, 9\}$ für $j \in \mathbb{N}$, so dass folgende Darstellung als unendlicher Dezimalbruch gilt:

$$a = \lim_{n\to\infty} a_n \quad \text{mit} \quad a_n = \sum_{j=0}^{n} k_j \cdot 10^{-j} = k_0, k_1 k_2 \ldots k_n.$$

Um das zu zeigen, definieren wir induktiv $a_n = a_{n-1} + k_n \cdot 10^{-n}$ mit $a_n \leq a < a_n + 10^{-n}$. Für $n = 0$ setzen wir $k_0 = \max\{k \in \mathbb{Z} : k \leq a\}$ und haben wie gewünscht

$$a_0 = k_0 \leq a < k_0 + 1 = a_0 + 10^{-0}.$$

Sind $k_0, k_1, \ldots, k_{n-1}$ bereits gefunden für ein $n \in \mathbb{N}$, so definieren wir $k_n = \max\{k \in \mathbb{Z} : a_{n-1} + k \cdot 10^{-n} \leq a\}$. Nach Induktionsannahme gilt $a_{n-1} \leq a < a_{n-1} + 10 \cdot 10^{-n}$ und folglich $k_n \in \{0, 1, \ldots, 9\}$. Weiter liefert die Wahl von k_n

$$a_n = a_{n-1} + k_n \cdot 10^{-n} \leq a < a_{n-1} + (k_n + 1) \cdot 10^{-n} = a_n + 10^{-n}.$$

Die Darstellung als unendlicher Dezimalbruch gilt wegen

$$|a - a_n| \leq 10^{-n} \to 0 \text{ mit } n \to \infty.$$

Umgekehrt stellt sich aber nun die Frage, ob jede Dezimalbruchfolge $a_n = k_0, k_1 k_2 \ldots k_n$ gegen eine gewisse, reelle Zahl konvergiert.

In beiden Beispielen brauchen wir wie gesagt eine Charakterisierung konvergenter Folgen, die ohne die Kenntnis des Grenzwerts auskommt. Die Idee von Augustin Louis Cauchy (1789–1857) besteht darin, die Glieder der Folge nicht mit dem unbekannten Grenzwert, sondern *untereinander* zu vergleichen.

Definition 2.1 (Cauchyfolge) *Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt Cauchyfolge, wenn gilt:*

Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{R}$, so dass $|a_n - a_m| < \varepsilon$ für alle $n, m > N$.

Beim Nachweis dieser Eigenschaft reicht es aus, die Zahlen $n, m > 0$ mit $n < m$ zu betrachten, denn die Definition ist symmetrisch in n und m und für $n = m$ ist nichts zu tun.

(V) Vollständigkeitsaxiom: Jede Cauchyfolge ist konvergent.

Damit sind die Axiome (KAV) der reellen Zahlen komplett. Je nach Autor werden auch andere Aussagen als Vollständigkeitsaxiom zugrunde gelegt, die aber natürlich äquivalent sind und sich bei uns als Folgerungen ergeben werden.

Satz 2.2 *Eine Folge ist genau dann konvergent, wenn sie eine Cauchyfolge ist.*

Beweis: Eine Cauchyfolge ist konvergent nach dem Vollständigkeitsaxiom. Sei umgekehrt $\lim_{n\to\infty} a_n = a$. Zu $\varepsilon > 0$ gibt es dann ein $N \in \mathbb{N}$ mit $|a_n - a| < \varepsilon/2$ für $n > N$, und für $n, m > N$ folgt

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

□

Als erste Anwendung der Vollständigkeit betrachten wir die Dezimaldarstellung und zeigen

Satz 2.3 (Konvergenz von Dezimalbrüchen) *Seien $k_0 \in \mathbb{Z}$ und $k_j \in \{0, 1, \ldots, 9\}$ für $j \in \mathbb{N}$. Dann konvergiert der Dezimalbruch $a_n = \sum_{j=0}^{n} k_j \cdot 10^{-j}$ gegen eine reelle Zahl.*

BEWEIS: Für $n < m$ schätzen wir wie folgt ab, wobei wir die Formel für die geometrische Reihe, Beispiel 1.8, und Beispiel 1.4 verwenden:

$$|a_m - a_n| = \sum_{j=n+1}^{m} k_j \cdot 10^{-j} \le 10^{-(n+1)} \sum_{j=0}^{\infty} 9 \cdot 10^{-j} \le 10^{-n} < \varepsilon \text{ für } n > N.$$

□

Als zweite Anwendung definieren wir die Eulersche Zahl e, und betrachten dazu die Folge

$$a_n = \sum_{k=0}^{n} \frac{1}{k!}.$$

Für $n \in \mathbb{N}$ und $m > n$ berechnen wir

$$\begin{aligned} a_m - a_n &= \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \ldots + \frac{1}{(n+2) \cdot \ldots \cdot (n+m-n)} \right) \\ &\le \frac{1}{(n+1)!} \left(1 + 2^{-1} + 2^{-2} + \ldots + 2^{m-n-1} \right). \end{aligned}$$

Mit der Formel für die geometrische Reihe, siehe Beispiel 1.8, folgt die Abschätzung

$$\frac{1}{(n+1)!} \le a_m - a_n \le \frac{2}{(n+1)!} \qquad \text{für } m > n. \tag{2.2}$$

Da $1/(n+1)! \le 1/(n+1) \to 0$ mit $n \to \infty$, gibt es zu $\varepsilon > 0$ ein $N \in \mathbb{R}$ mit

$$|a_m - a_n| \le \frac{2}{(n+1)!} < \varepsilon \qquad \text{für } n, m > N.$$

Aus dem Vollständigkeitsaxiom folgt die Existenz des Grenzwerts $\lim_{n\to\infty} a_n$, so dass folgende Definition sinnvoll ist.

Definition 2.2 (Eulersche Zahl)

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

Satz 2.4 (Irrationalität der Eulerschen Zahl) $e = \sum_{k=0}^{\infty} 1/k!$ *ist nicht rational.*

BEWEIS: Aus Abschätzung (2.2) folgt mit $m \to \infty$

$$\frac{1}{(n+1)!} \le e - a_n \le \frac{2}{(n+1)!} \qquad \text{für alle } n \in \mathbb{N}.$$

Nach Multiplikation mit $n!$ ergibt sich hieraus für $n \ge 2$

$$0 < \frac{1}{n+1} \le n!\, e - \sum_{k=0}^{n} \frac{n!}{k!} \le \frac{2}{(n+1)} < 1.$$

Wäre e rational, so wäre der mittlere Term eine ganze Zahl für n hinreichend groß, ein Widerspruch. □

Das nun folgende Konvergenzkriterium ist überaus nützlich. Es ist ein *hinreichendes* Kriterium für Konvergenz, ist aber nicht notwendig für die Konvergenz einer Folge $(a_n)_{n\in\mathbb{N}}$.

Definition 2.3 (Monotone Folge) *Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt monoton wachsend, wenn*

$$a_{n+1} \geq a_n \quad \text{für alle } n \in \mathbb{N}.$$

Manche Autoren bezeichnen diese Eigenschaft auch als nichtfallend, und reservieren den Begriff wachsend für eine Folge mit $a_{n+1} > a_n$. Das bezeichnen wir als *streng monoton wachsend.*

Satz 2.5 (Konvergenzkriterium der Monotonie und Beschränktheit) *Jede monoton wachsende, nach oben beschränkte Folge ist eine Cauchyfolge und damit konvergent.*

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und $a_n \leq K < \infty$ für alle $n \in \mathbb{N}$. Für $\varepsilon > 0$ betrachte

$$M = \{j \in \mathbb{N}_0 : \text{es gibt ein } n \in \mathbb{N} \text{ mit } a_n \geq a_1 + j\varepsilon\}.$$

Offenbar ist $0 \in M$, und für $j \in M$ gilt $j \leq (K - a_1)/\varepsilon$. Sei $k \in M$ maximal. Dann gibt es ein $N \in \mathbb{N}$ mit $a_N \geq a_1 + k\varepsilon$, und für alle $n \geq N$ folgt

$$a_1 + k\varepsilon \leq a_N \leq a_n < a_1 + (k+1)\varepsilon.$$

Damit gilt $|a_n - a_m| \leq \varepsilon$ für $n, m \geq N$, das heißt (a_n) ist eine Cauchyfolge. □

Beispiel 2.3 (Die Zahl e und Zinsrechnung) Zu diesem Zeitpunkt ist nicht einsichtig, warum zur Definition der Eulerschen Zahl die Formel $e = \sum_{k=0}^{\infty} 1/k!$ gewählt wurde. Darum zeigen wir nun die alternative und vielleicht aus der Schule bekannte Darstellung

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n. \tag{2.3}$$

Mit anderen Worten: wenn ein Euro für ein Jahr mit Zinssatz $x = 1$ bzw. 100 Prozent kontinuierlich verzinst wird, ist das Endkapital $e \approx 2,71828\ldots$ Euro, statt 2 Euro bei jährlicher Verzinsung, vergleiche Beispiel 2.1. Zum Beweis von (2.3) bemerken wir zunächst, dass die Folge $b_n = \left(1 + \frac{1}{n}\right)^{1/n}$ monoton wachsend und nach oben beschränkt ist; dies wurde in Aufgabe 4, Serie 3, gezeigt. Nach Satz 2.5 existiert der Grenzwert $b = \lim_{n\to\infty} b_n$. Mit dem Binomischen Lehrsatz erhalten wir

$$b_n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{1}{k!}\, \frac{n}{n} \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n} \leq \sum_{k=0}^{n} \frac{1}{k!} = a_n.$$

Mit $n \to \infty$ folgt $b \leq e$ nach Satz 1.4. Die umgekehrte Abschätzung ist etwas subtiler: für beliebiges $m \in \mathbb{N}$ und $n \geq m$ gilt, da die Summanden größer gleich Null sind,

$$b_n \geq \sum_{k=0}^{m} \frac{1}{k!}\, \frac{n}{n} \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n}.$$

Indem wir hier $n \to \infty$ gehen lassen, folgt mit Satz 1.4

$$b \geq \sum_{k=0}^{m} \frac{1}{k!},$$

und $m \to \infty$ liefert $b \geq e$ und damit $b = e$ wie gewünscht. Tatsächlich liefert das Argument $\lim_{n\to\infty} b_n = e$, ohne dass die Konvergenz der Folge b_n vorausgesetzt wird.

Als nächste Anwendung des Vollständigkeitsaxioms diskutieren wir nun das Intervallschachtelungsprinzip.

Definition 2.4 (Intervallschachtelung) *Eine Intervallschachtelung ist eine Folge von abgeschlossenen Intervallen $I_n = [a_n, b_n] \subset \mathbb{R}$ mit $I_{n+1} \subset I_n$ für alle n und $|I_n| = b_n - a_n \to 0$ mit $n \to \infty$.*

Satz 2.6 (Intervallschachtelungsprinzip) *Zu jeder Intervallschachtelung $(I_n)_{n\in\mathbb{N}}$ gibt es genau ein $x \in \mathbb{R}$ mit $x \in \bigcap_{n\in\mathbb{N}} I_n$. Es gilt $x = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.*

Beweis: Nach Voraussetzung haben wir

$$a_1 \leq a_2 \leq \ldots \leq a_n \leq \ldots \quad \ldots \leq b_n \leq \ldots \leq b_2 \leq b_1.$$

Aus Satz 2.5 folgt die Existenz der Grenzwerte $a = \lim_{n\to\infty} a_n$ bzw. $b = \lim_{n\to\infty} b_n$. Dann gilt nach den Satz 1.3 und Satz 1.4

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

Setze $x := a = b$. Dann ist $a_n \leq a = x = b \leq b_n$, also $x \in I_n$ für alle n. Sei $y \in \mathbb{R}$ mit $y \in I_n$ für alle n, das heißt $a_n \leq y \leq b_n$. Durch Grenzübergang ergibt sich nach Satz 1.4 $a \leq y \leq b$, also $y = x$. □

Satz 2.7 (Existenz der n-ten Wurzel) *Zu jedem $a > 0$ und $n \in \mathbb{N}$ gibt es genau ein $x > 0$ mit $x^n = a$. Bezeichnung: $x = \sqrt[n]{a} = a^{1/n}$.*

Beweis: Die Eindeutigkeit wurde schon in Beispiel 1.9 aus den Anordnungsaxiomen gefolgert. Wir konstruieren die Lösung mit dem Verfahren der fortgesetzten Intervallhalbierung: Bestimme $I_k = [a_k, b_k]$ für $k = 1, 2, \ldots$, so dass mit $m_k = \frac{a_k + b_k}{2}$ gilt:

$$I_1 = [a_1, b_1] \text{ mit } a_1^n \leq a \leq b_1^n;$$

$$I_{k+1} = \begin{cases} [a_k, m_k] & \text{falls } m_k^n \geq a \\ [m_k, b_k] & \text{falls } m_k^n < a. \end{cases}$$

Es folgt $I_{k+1} \subset I_k$ für alle k und $|I_k| = 2^{1-k}|I_1| \to 0$ mit $k \to \infty$. Sei $x \in \mathbb{R}$ wie in Satz 2.6 gegeben. Per Induktion ergibt sich aus der Definition von I_{k+1} die Ungleichung $a_k^n \leq a \leq b_k^n$ für alle $k \in \mathbb{N}$, und hieraus mit den Sätzen 1.3 und 1.4

$$x^n = \lim_{n\to\infty} a_k^n \leq a \leq \lim_{n\to\infty} b_k^n = x^n.$$

□

Für rationale Exponenten $r = p/q$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ wird die Potenz erklärt durch $a^r = (a^p)^{1/q}$. Dies ist wohldefiniert, denn aus $p_1/q_1 = p_2/q_2$ folgt

$$\left(\left(a^{p_1}\right)^{1/q_1}\right)^{q_1 q_2} = \left(\left(\left(a^{p_1}\right)^{1/q_1}\right)^{q_1}\right)^{q_2} = \left(a^{p_1}\right)^{q_2} = a^{p_1 q_2} = a^{p_2 q_1} = \left(\left(a^{p_2}\right)^{1/q_2}\right)^{q_1 q_2},$$

also $(a^{p_1})^{1/q_1} = (a^{p_2})^{1/q_2}$. Weiter zeigt man leicht die Potenzgesetze (für ganzzahlige Exponenten sind diese Regeln klar und wurden schon benutzt)

$$(i)\ a^s a^r = a^{r+s} \quad (ii)\ (a^r)^s = a^{rs} \quad (iii)\ a^r b^r = (ab)^r.$$

Zum Beispiel gilt mit $r = k/m$ und $s = p/q$

$$(a^r a^s)^{mq} = (a^{k/m})^{mq} (a^{p/q})^{mq} = \Big(\big((a^k)^{1/m}\big)^m\Big)^q \cdot \Big(\big((a^p)^{1/q}\big)^q\Big)^m = (a^k)^q \cdot (a^p)^m = a^{kq+pm}.$$

Dies bedeutet

$$a^r a^s = \big(a^{kq+pm}\big)^{1/mq} = a^{(kq+pm)/mq} = a^{k/m+p/q} = a^{r+s}.$$

Die anderen beiden Potenzgesetze werden ähnlich verifiziert.

Definition 2.5 (Teilfolge) *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und $(n_k)_{k\in\mathbb{N}}$ eine Folge natürlicher Zahlen mit $n_1 < n_2 < n_3 \ldots$ Dann heißt die Folge $(a_{n_k})_{k\in\mathbb{N}}$ Teilfolge von $(a_n)_{n\in\mathbb{N}}$.*

Durch Induktion ergibt sich sofort $n_k \geq k$: es ist $n_1 \geq 1$ und $n_{k+1} \geq n_k + 1 \geq k + 1$. Die Teilfolge entsteht aus der ursprünglichen Folge durch *Auswahl der Nummern n_k*. Da Folgen Abbildungen von $\mathbb{N}$ nach $\mathbb{R}$ sind, ist eine Teilfolge formal als Verkettung von zwei Abbildungen definiert: der Ausgangsfolge $a : \mathbb{N} \to \mathbb{R}$, $n \longmapsto a_n$ und der Folge $\mathbb{N} \to \mathbb{N}$, $k \longmapsto n_k$, die die Indizes auswählt. Am Beispiel $a_n = (-1)^n/n^3$ und $n_k = 2k - 1$ sieht das wie folgt aus:

$$\begin{array}{ccccc}
\mathbb{N} & \xrightarrow{n_k=2k-1} & \mathbb{N} & \xrightarrow{a_n=(-1)^n/n^3} & \mathbb{R} \\[2ex]
k & \longmapsto & n_k = 2k-1 & \longmapsto & a_{n_k} = -1/(2k-1)^3
\end{array}$$

Definition 2.6 (Häufungspunkt von Folgen) *$a \in \mathbb{R}$ heißt Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$, wenn eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ gibt, die mit $k \to \infty$ gegen a konvergiert.*

Beispielsweise hat die Folge $a_n = (-1)^n + 1/n^2$ den Häufungspunkt $+1$, denn mit $n_k = 2k$ gilt $a_{n_k} = a_{2k} = 1 + 1/(2k)^2 \to 1$ mit $k \to \infty$. Auch -1 ist ein Häufungspunkt der Folge, denn für $n_k = 2k - 1$ ist $a_{n_k} = a_{2k-1} = -1 + 1/(2k-1)^2 \to -1$ mit $k \to \infty$.

Lemma 2.1 *Die Zahl $a \in \mathbb{R}$ ist genau dann Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$, wenn für jedes $\varepsilon > 0$ die Menge $\{n \in \mathbb{N} : a_n \in U_\varepsilon(a)\}$ unendlich viele Elemente hat.*

Beweis: Wenn $a \in \mathbb{R}$ Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ ist, so gilt nach Definition $a_{n_k} \to a$ mit $k \to \infty$ für eine Teilfolge n_k. Zu jedem $\varepsilon > 0$ gibt es dann ein $K \in \mathbb{R}$ mit $a_{n_k} \in U_\varepsilon(a)$ für alle $k > K$. Die Abbildung $k \mapsto n_k$ ist injektiv wegen $n_1 < n_2 < \ldots$, also ist die Menge $\{n_k : k > K\}$ nicht endlich nach dem Schubfachprinzip. Dies beweist die eine Richtung der Äquivalenz. Umgekehrt wählen wir induktiv n_k mit $n_1 < n_2 < \ldots$, so dass $a_{n_k} \in U_{1/k}(a)$. Die Induktion bricht nicht ab, da $a_n \in U_{1/k}(a)$ für unendlich viele n gilt. Es folgt dann $|a_{n_k} - a| < 1/k \to 0$ mit $k \to \infty$. □

Satz 2.8 (Bolzano-Weierstraß) *Jede beschränkte Folge reeller Zahlen hat eine konvergente Teilfolge, also mindestens einen Häufungspunkt.*

Beweis: Konstruktion durch fortgesetzte Intervallhalbierung: wähle eine obere Schranke b_1 und eine untere Schranke a_1 für $(x_n)_{n\in\mathbb{N}}$, also $x_n \in [a_1, b_1]$ für alle $n \in \mathbb{N}$. Wir nehmen nun induktiv an, dass $I_k = [a_k, b_k]$ schon gefunden ist mit der Eigenschaft

$$(*) \quad x_n \in I_k \quad \text{für unendlich viele } n.$$

Setze $m_k = \frac{1}{2}(a_k + b_k)$ und definiere

$$I_{k+1} = [a_{k+1}, b_{k+1}] = \begin{cases} [m_k, b_k], & \text{falls } x_n \in [m_k, b_k] \text{ für unendlich viele } n, \\ [a_k, m_k] & \text{sonst.} \end{cases}$$

Es ist offensichtlich, dass $(*)$ auch für das Intervall I_{k+1} gilt. Nach dem Intervallschachtelungsprinzip aus Satz 2.6 gibt es ein $x \in \mathbb{R}$ mit $x \in I_k$ für alle $k \in \mathbb{N}$. Zu $\varepsilon > 0$ gibt es nun ein $K \in \mathbb{R}$ mit $|I_k| < \varepsilon$ für $k > K$, also $I_k \subset U_\varepsilon(x)$ für $k > K$. Damit gilt auch $x_n \in U_\varepsilon(x)$ für unendlich viele n. Aus Lemma 2.1 schließen wir, dass x ein Häufungspunkt der Folge $(x_n)_{n\in\mathbb{N}}$ ist. □

Definition 2.7 (Limes superior/inferior) *Für eine reelle Folge $(x_n)_{n\in\mathbb{N}}$ und $x^* \in \mathbb{R} \cup \{\pm\infty\}$ gilt $\limsup_{n\to\infty} x_n = x^*$, falls folgende zwei Bedingungen erfüllt sind:*

(i) *es gibt eine Teilfolge x_{n_k} mit $x_{n_k} \to x^*$ für $k \to \infty$,*

(ii) *für alle $x > x^*$ ist die Menge $\{n \in \mathbb{N} : x_n > x\}$ endlich.*

Entsprechend bedeutet $\liminf_{n\to\infty} x_n = x_*$ mit $x_* \in \mathbb{R} \cup \{\pm\infty\}$:

(i) *es gibt eine Teilfolge x_{n_k} mit $x_{n_k} \to x_*$ für $k \to \infty$,*

(ii) *für alle $x < x_*$ ist die Menge $\{n \in \mathbb{N} : x_n < x\}$ endlich.*

Der Limes superior ist nicht notwendig obere Schranke der Folge, zum Beispiel gilt $\limsup_{n\to\infty} 1/n = 0$. Er kann auch $-\infty$ sein, zum Beispiel für $x_n = -n$. Während wir den Grenzwert $\lim_{n\to\infty} x_n$ nur dann bilden können, wenn die Folge konvergiert, ist der größte Häufungspunkt $x^* = \limsup_{n\to\infty} x_n$ und der kleinste Häufungspunkt $x_* = \liminf_{n\to\infty} x_n = x_*$ immer definiert, wenn wir jeweils die Möglichkeit $x^* = \pm\infty$ bzw. $x_* = \pm\infty$ zulassen. Dies soll nun bewiesen werden, wobei wir uns o.B.d.A. auf den Limes superior beschränken.

Lemma 2.2 *Sei $\limsup_{n\to\infty} x_n = x^* < \infty$. Dann gilt*

$$x > x^* \quad \Rightarrow \quad x \text{ ist kein Häufungspunkt von } (x_n).$$

Beweis: Setze $\varepsilon = \frac{1}{2}(x - x^*) > 0$. Für $y \in U_\varepsilon(x)$ folgt $y - x^* = x - x^* - (x - y) > 2\varepsilon - \varepsilon = \varepsilon$, mit anderen Worten $U_\varepsilon(x) \subset \{y \in \mathbb{R} : y > x^* + \varepsilon\}$. Nach Definition 2.7(ii) ist $x_n \in U_\varepsilon(x)$ nur für endlich viele $n \in \mathbb{N}$, das heißt x ist kein Häufungspunkt nach Lemma 2.1. □

Satz 2.9 (Existenz des Limes superior) *Für jede Folge $(x_n)_{n\in\mathbb{R}}$ gibt es genau ein $x^* \in \mathbb{R} \cup \{\pm\infty\}$ mit $\limsup_{n\to\infty} x_n = x^*$.*

Beweis:

Fall 1: (x_n) ist nicht nach oben beschränkt.
Dann ist $\{n : x_n \geq b\}$ unendlich für alle $b \in \mathbb{R}$. Bestimme induktiv $n_1 < n_2 < \ldots$ mit $x_{n_k} \geq k$. Es folgt $\limsup_{n\to\infty} x_n = +\infty$.

Fall 2: Es gibt ein $b_1 \in \mathbb{R}$ mit $x_n \leq b_1$ für alle $n \in \mathbb{N}$.

Fall 2.1: $\{n : x_n \geq a\}$ ist endlich für alle $a \in \mathbb{R}$.
Dann gilt $x_n \to -\infty$ und es folgt $\limsup_{n\to\infty} x_n = -\infty$.

Fall 2.2: Es gibt ein $a_1 \in \mathbb{R}$, so dass $\{n : x_n \in [a_1, b_1]\}$ unendlich ist.
In diesem Fall wenden wir das Intervallhalbierungsverfahren aus dem Beweis von Satz 2.8 an und behaupten

$$\{n : x_n > b_k\} \text{ ist endlich für alle } k.$$

Für $k = 1$ ist das richtig, da b_1 obere Schranke. Sei die Behauptung schon für $k \in \mathbb{N}$ gezeigt.

$b_{k+1} = b_k \quad \Rightarrow \quad$ Die Behauptung gilt nach Induktionsannahme.

$b_{k+1} = m_k \quad \Rightarrow \quad$ Die Menge $\{n : x_n > b_{k+1}\} = \{n : x_n \in (m_k, b_k]\} \cup \{n : x_n > b_k\}$ ist endlich nach Fallunterscheidung sowie Induktionsannahme.

Sei nun $x^* := \lim_{k\to\infty} b_k$. Für $x > x^*$ ist $(x, +\infty) \subset (b_k, +\infty)$ für k hinreichend groß, also ist $\{n : x_n > x\}$ endlich und es folgt $\limsup_{n\to\infty} = x^*$. Damit ist die Existenz bewiesen.

Angenommen es gibt $x_1^* < x_2^*$ mit den Eigenschaften (i) und (ii). Wähle $x \in (x_1^*, x_2^*)$. Wegen (ii) für x_1^* ist dann $\{n : x_n > x\}$ endlich. Dann kann aber (i) für x_2^* nicht gelten, ein Widerspruch. □

Die Begriffe Häufungspunkt, Limes superior und Limes inferior sind gewöhnungsbedürftig, und wir werden bei Gelegenheit mehr Beispiele betrachten. Die logische Abfolge der *zentralen theoretischen Aussagen* in diesem Abschnitt war folgende:

Vollständigkeitsaxiom:
Cauchyfolgen sind konvergent

$\Downarrow$

Konvergenzkriterium der
Monotonie und Beschränktheit

$\Downarrow$

Intervallschachtelungsprinzip

$\Downarrow$

Bolzano-Weierstraß: Auswahlsatz

$\Downarrow$

Cauchyfolgen sind konvergent

Die Implikationen sind so zu verstehen, dass jeweils nur die jeweils vorangehende Eigenschaft von $\mathbb{R}$ im Beweis des darauf folgenden Resultats benutzt wurde. Die letzte Implikation werden wir dabei gleich noch zeigen. Es folgt, dass jede der vier Eigenschaften als Axiom für $\mathbb{R}$ benutzt werden könnte - die anderen Eigenschaften würden als Sätze folgen. Im nächsten Abschnitt werden wir eine weitere, äquivalente Eigenschaft kennenlernen, nämlich den Satz vom Supremum (Satz 2.10). In dieser Vorlesung wird die Konvergenz der Cauchyfolgen als grundlegendes Axiom gewählt.

Beweis: *Auswahlsatz von Bolzano-Weierstraß $\Rightarrow$ Cauchyfolgen sind konvergent*
Wir zeigen zunächst, dass eine Cauchyfolge $(a_n)_{n\in\mathbb{N}}$ beschränkt ist. Sei $n_0 \in \mathbb{N}$ mit $|a_n -$

$a_m| \leq 1$ für $n, m \geq n_0$. Dann folgt

$$n \geq n_0 \Rightarrow |a_n| \leq \underbrace{|a_n - a_{n_0}|}_{\leq 1} + |a_{n_0}|.$$

Also gilt $|a_n| \leq \max(|a_1|, |a_2|, \ldots, |a_{n_0-1}|, |a_{n_0}|+1) = K$. Sei nun $N \in \mathbb{R}$ mit $|a_n - a_m| < \frac{\varepsilon}{2}$ für $n, m > N$. Wir wenden den Auswahlsatz von Bolzano-Weierstraß an: es gibt eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit $a_{n_k} \to a$ mit $k \to \infty$. Wähle $k_0 \in \mathbb{N}$ mit $|a_{n_{k_0}} - a| < \frac{\varepsilon}{2}$ und $n_{k_0} > N$. Für $n > N$ folgt

$$|a_n - a| \leq \underbrace{|a_n - a_{n_{k_0}}|}_{<\frac{\varepsilon}{2}} + \underbrace{|a_{n_{k_0}} - a|}_{<\frac{\varepsilon}{2}} < \varepsilon.$$

Somit konvergiert die ganze Folge gegen a. □

Definition 2.8 *Die Menge $M \subset \mathbb{R}$ heißt*

$$\begin{aligned} \textit{nach oben beschränkt} \quad &\Leftrightarrow \quad \exists\, K \in \mathbb{R} \textit{ mit } x \leq K \textit{ für alle } x \in M, \\ \textit{nach unten beschränkt} \quad &\Leftrightarrow \quad \exists\, K \in \mathbb{R} \textit{ mit } x \geq K \textit{ für alle } x \in M. \end{aligned}$$

Die Zahl K heißt dann obere bzw. untere Schranke. Weiter heißt M

$$\textit{beschränkt} \quad \Leftrightarrow \exists\, K \geq 0 \textit{ mit } |x| \leq K \textit{ für alle } x \in M.$$

Eine Menge ist genau dann beschränkt, wenn sie nach oben und unten beschränkt ist. Denn aus $|x| \leq K$ folgt $-K \leq x \leq K$, und aus $K_1 \leq x \leq K_2$ folgt umgekehrt $|x| \leq \max(|K_1|, |K_2|)$.

Beispiel 2.4 Die Menge $[0, 1)$ ist nach oben beschränkt, eine obere Schranke ist zum Beispiel $K = 2011$. Es gibt in $[0, 1)$ kein größtes Element, denn es gilt der Schluss

$$x \in [0, 1) \quad \Rightarrow \quad x < \frac{x+1}{2} \in [0, 1).$$

Unter den oberen Schranken von $[0, 1)$ gibt es aber eine kleinste, nämlich die Zahl 1.

Definition 2.9 (Supremum/Infimum) *Die Zahl $a \in \mathbb{R} \cup \{\pm\infty\}$ heißt Supremum (bzw. Infimum) der Menge $M \subset \mathbb{R}$, wenn folgendes gilt:*

(1) *$x \leq a$ für alle $x \in M$ (bzw. $x \geq a$ für alle $x \in M$),*

(2) *Für alle $a' < a$ (bzw. $a' > a$) gibt es ein $x \in M$ mit $x > a'$ (bzw. $x < a'$).*

Notation: $a = \sup M$ (bzw. $a = \inf M$). Ist M nach oben (bzw. unten) beschränkt, so bezeichnen wir $\sup M$ auch als kleinste obere Schranke (bzw. $\inf M$ als größte untere Schranke).

Satz 2.10 *Jede Menge $M \subset \mathbb{R}$ hat genau ein Supremum $S \in \mathbb{R} \cup \{\pm\infty\}$.*

Beweis: Wir zeigen als erstes die Eindeutigkeit. Angenommen es gibt $S_{1,2} \in \mathbb{R} \cup \{\pm\infty\}$, die beide die Definition 2.9 erfüllen. Wir können $S_1 < S_2$ annehmen. Nach Eigenschaft (2) bzgl. S_2 gibt es ein $x \in M$ mit $x > S_1$, im Widerspruch zur Eigenschaft (1) bezüglich S_1.

Man sieht leicht $\sup \emptyset = -\infty$, und $\sup M = +\infty$ falls M nicht nach oben beschränkt ist. Im verbleibenden Fall wählen wir ein Element a_1 von M und eine obere

Schranke b_1 von M, und konstruieren wie folgt eine Intervallschachtelung $I_n = [a_n, b_n]$ für $n \geq 1$, wobei $m_n = (a_n + b_n)/2$:

$$I_{n+1} = [a_{n+1}, b_{n+1}] = \begin{cases} [m_n, b_n], & \text{falls } [m_n, b_n] \cap M \neq \emptyset \\ [a_n, m_n], & \text{sonst.} \end{cases}$$

Nach dem Intervallschachtelungsprinzip, Satz 2.6, gibt es genau ein $S \in \mathbb{R}$ mit $S \in I_n$ für alle n, und genauer gilt $a_n \to S$ sowie $b_n \to S$. Für $x \in M$ sieht man durch Induktion $x \leq b_n$ für alle n, also $x \leq \lim_{n\to\infty} b_n = S$. Andererseits gilt, ebenfalls induktiv, $M \cap I_n \neq \emptyset$ für alle n, das heißt es gibt $x_n \in M$ mit $a_n \leq x_n \leq b_n$. Ist $S' < S$, so gilt also $x_n > S'$ für hinreichend große n. Damit ist $\sup M = S$ gezeigt. □

Folgerung 2.1 *Sei $M \subset \mathbb{R}$ nichtleer. Dann gibt es eine Folge $x_n \in M$ (bzw. $x'_n \in M$) mit $x_n \to \sup M$ (bzw. $x'_n \to \inf M$).*

Beweis: Wir zeigen die Aussage für das Supremum. Ist M nach oben beschränkt, so wurde eine solche Folge im Beweis des vorangehenden Satzes konstruiert. Ist M nicht nach oben beschränkt, so gibt es zu $n \in \mathbb{N}$ ein $x_n \in M$ mit $x_n \geq n$, also $x_n \to +\infty = \sup M$. □

3 Mächtigkeit der Mengen und komplexe Zahlen

Jetzt wollen wir uns einer neuen Frage zuwenden: wie kann man die unendlichen Mengen $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ der Größe nach vergleichen? Gibt es wirklich mehr rationale Zahlen als natürliche? Mehr reelle als rationale? Die Antwort lautet witzigerweise, dass es gleich viele natürliche, ganze und rationale Zahlen gibt, aber mehr reelle Zahlen. Alle genannten Mengen enthalten unendlich viele Elemente. Die Präzisierung der Begriffe *gleichviele* und *mehr* geht auf Georg Cantor (1872) zurück.

Definition 3.1 *Eine Menge A heißt gleichmächtig zur Menge B (Notation: $A \sim B$), wenn es eine bijektive Abbildung (Bijektion) $\varphi : A \to B$ gibt.*

Lemma 3.1 *Die Relation $A \sim B$ (A ist gleichmächtig zu B) ist eine Äquivalenzrelation, das heißt für alle Mengen A, B, C gilt:*

(i) $A \sim A$

(ii) $A \sim B \quad \Rightarrow B \sim A$

(iii) $A \sim B$ und $B \sim C \quad \Rightarrow A \sim C.$

Beweis:

(i) Wähle als Bijektion die Identität $id_A : A \to A$, $a \longmapsto a$.

(ii) Sei $\varphi : A \to B$ Bijektion. Wähle dann $\varphi^{-1} : B \to A$.

(iii) Seien $\varphi : A \to B$ und $\psi : B \to C$ bijektiv. Wähle dann $\psi \circ \varphi : A \to C$.

□

Definition 3.2 *Eine Menge A heißt*

- *endlich, wenn sie gleichmächtig zu einer Menge $\{1, 2, \ldots, k\}$ mit $k \in \mathbb{N}$ ist;*
- *abzählbar unendlich, wenn sie gleichmächtig zu $\mathbb{N}$ ist;*
- *abzählbar, wenn sie endlich oder abzählbar unendlich ist;*
- *überabzählbar, wenn sie nicht abzählbar ist.*

Die Menge der natürlichen Zahlen ist nicht endlich, denn andernfalls gibt es ein $n \in \mathbb{N}$ und eine Bijektion $f : \mathbb{N} \to \{1, \ldots, n\}$ im Widerspruch zum Schubfachprinzip (Satz 2.3).

Lemma 3.2 *Falls eine surjektive Abbildung $\varphi : \mathbb{N} \to A$ existiert, so ist A abzählbar.*

Beweis: Sei $a_n = \varphi(n)$. Die naheliegende Idee ist, bei der Abzählung induktiv diejenigen a_n auszulassen, die bereits vorher auftraten, und die restlichen entsprechend neu zu nummerieren. Dazu setzen wir $n_1 = 1$ und konstruieren induktiv eine Teilfolge a_{n_k} durch die Vorschrift

$$n_{k+1} = \min\{n > n_k : a_n \neq a_{n_j} \text{ für } j = 1, \ldots, k\}.$$

Falls die Rekursion nach einem n_k abbricht, ist die Abbildung $f : \{1, \ldots, k\} \to A$, $k \mapsto a_{n_k}$ bijektiv und damit ist A endlich. Andernfalls ist die Abbildung $f : \mathbb{N} \to A$, $k \mapsto a_{n_k}$ bijektiv und A ist abzählbar undendlich. □

Satz 3.1 *Die Mengen $\mathbb{Z}$ und $\mathbb{Q}$ sind abzählbar.*

Beweis: Für $\mathbb{Z}$ wähle die surjektive (sogar bijektive!) Abbildung

$$\varphi : \mathbb{N} \to \mathbb{Z},\ \varphi(n) = \begin{cases} (n-1)/2 & n \text{ ungerade} \\ -n/2 & n \text{ gerade.} \end{cases}$$

Das Argument zeigt, dass wir uns beim Beweis der Abzählbarkeit von $\mathbb{Q}$ auf die Menge $\mathbb{Q}^+ = \{p/q : p, q \in \mathbb{N}\}$ beschränken können. Die Idee ist dann, diagonal nach folgendem Schema abzuzählen:

$$\begin{array}{llcccccccccccc}
 & p = & 1 & & 2 & & 3 & & 4 & & 5 & & 6 \\
q = & & & & & & & & & & & & \\
1 & & 1/1 & & 2/1 & & 3/1 & & 4/1 & & 5/1 & & 6/1 \\
 & & & \swarrow & & \swarrow & & \swarrow & & \swarrow & & \swarrow & \\
2 & & 1/2 & & 2/2 & & 3/2 & & 4/2 & & \cdot & & \\
 & & & \swarrow & & \swarrow & & \swarrow & & \cdot & & & \\
3 & & 1/3 & & 2/3 & & 3/3 & & \cdot & & & & \\
 & & & \swarrow & & \swarrow & & \cdot & & & & & \\
4 & & 1/4 & & 2/4 & & \cdot & & & & & & \\
 & & & \swarrow & & \cdot & & & & & & & \\
5 & & 1/5 & & & & & & & & & &
\end{array}$$

Die k-te Diagonale enthält k Elemente, also enthalten die Diagonalen mit kleinerer Nummer insgesamt $1 + \ldots + (k-1) = (k-1)k/2$ Einträge. Wir behaupten, dass jedes $n \in \mathbb{N}$ in genau einem k-Abschnitt $\{(k-1)k/2 + i : i = 1, \ldots, k\}$ liegt. Zur Existenz sei $k \in \mathbb{N}$ maximal mit $(k-1)k/2 < n$. Dann gilt

$$(k-1)k/2 < n \leq k(k+1)/2 = (k-1)k/2 + k.$$

Für $l > k$ ist $(l-1)l/2+1 > k(k+1)/2 = (k-1)k/2+k$, das heißt die beiden Abschnitte sind disjunkt, womit die Eindeutigkeit gezeigt ist. Somit ist folgende Abbildung wohldefiniert:

$$\varphi : \mathbb{N} \to \mathbb{Q},\ \varphi(n) = \frac{k-i+1}{i}$$

$$\text{für } n = \frac{(k-1)k}{2} + i \quad (k \in \mathbb{N},\ 1 \leq i \leq k).$$

Die Abbildung ist surjektiv, und zwar gilt $\varphi(n) = p/q$ für $i = q$ und $k = p + q - 1$. Die Abzählbarkeit von $\mathbb{Q}$ folgt nun mit Lemma 3.2. □

Durch Anordnung in einem quadratischen Schema und analoge Abzählung läßt sich ganz analog zeigen, daß eine abzählbare Vereinigung von jeweils abzählbaren Mengen abermals abzählbar ist.

Satz 3.2 ($\mathbb{R}$ ist nicht abzählbar) *Es gibt keine surjektive Abbildung $\varphi : \mathbb{N} \to \mathbb{R}$.*

Beweis: Sei $\varphi : \mathbb{N} \to \mathbb{R}$ eine beliebige Abbildung und $\varphi(n) = x_n$. Wir konstruieren eine Intervallschachtelung $(I_n)_{n \in \mathbb{N}}$ mit $x_n \notin I_n$ für jedes n. Ist I_n schon bestimmt, so zerlege I_n in drei abgeschlossene Teilintervalle gleicher Länge und wähle für I_{n+1} ein Teilintervall, das x_{n+1} nicht enthält (im Zweifelsfall das rechte). Um I_1 zu definieren, wenden wir dieses

Argument an auf $I_0 = [0,1]$. Nach Satz 2.6 gibt es ein $x \in \bigcap_{n\in\mathbb{N}} I_n \in \mathbb{R}$, also $x \neq x_n$ für alle $n \in \mathbb{N}$. $\square$

In Zukunft brauchen wir den Begriff der Konvergenz auch für Punkte im $\mathbb{R}^n$. Für $n = 1$ haben wir als Modell die Zahlengerade benutzt, entsprechend betrachten wir für $n = 2$ die Ebene mit kartesischen Koordinaten $z = (x, y)$ und für $n = 3$ den dreidimensionalen Raum mit Koordinaten $p = (x, y, z)$; dabei finde ich den zweidimensionalen Fall besonders anschaulich. Der $\mathbb{R}^n$ ist eine mathematische Verallgemeinerung:

$$\mathbb{R}^n = \{x = (x_1, \ldots, x_n) : x_i \in \mathbb{R}\} = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n \text{ mal}}.$$

Genauer Definitionen finden Sie in der Vorlesung LA1.
Für uns hat der Fall $n = 2$ der Euklidischen Ebene eine besondere Bedeutung, denn aus dem Vektorraum $(\mathbb{R}^2, +)$ wird mit einer geeigneten Multiplikation der Körper $\mathbb{C} = (\mathbb{R}^2, +, \cdot)$ der komplexen Zahlen. Die Verknüpfungen sind dabei wie folgt definiert:

Die Addition ist die Vektorraumaddition von $\mathbb{R}^2$. Die Standardbasis wird mit $(1, 0) = 1$ und $(0, 1) = i$ bezeichnet, das heißt jedes $z = (x, y) \in \mathbb{R}^2$ besitzt die Basisdarstellung $z = x + iy$. Damit lautet die Addition von $x_k + iy_k$, $k = 1, 2$,

$$\begin{aligned}(x_1 + iy_1) + (x_2 + iy_2) &= (x_1, y_1) + (x_2, y_2)\\ &= (x_1 + x_2, y_1 + y_2)\\ &= (x_1 + x_2) + i(y_1 + y_2).\end{aligned}$$

Mit diesen Vereinbarungen gilt $x = x + i0 = (x, 0)$ für $x \in \mathbb{R}$, das heißt $\mathbb{R}$ wird mit der x-Achse in $\mathbb{R}^2$ identifiziert. Für $z = x + iy$ heißt $x =: \operatorname{Re} z \in \mathbb{R}$ der Realteil und $y = \operatorname{Im} z \in \mathbb{R}$ der Imaginärteil von z.

Die Multiplikation ergibt sich durch die Forderung $i^2 = -1$ und Ausmultiplizieren nach den Körpergesetzen:

$$\begin{aligned}(x_1 + iy_1)(x_2 + iy_2) &= x_1x_2 + i^2y_1y_2 + i(x_1y_2 + x_2y_1)\\ &= x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1).\end{aligned}$$

Für $\lambda \in \mathbb{R}$ ist $\lambda(x + iy) = (\lambda + i0)(x + iy) = \lambda x + i\lambda y = (\lambda x, \lambda y) = \lambda(x, y)$, das heißt Multiplikation mit $\lambda \in \mathbb{R} \subset \mathbb{C}$ ist einfach Skalarmultiplikation im $\mathbb{R}$-Vektorraum $\mathbb{R}^2$. Dagegen liefert die Multiplikation mit i für $(x, y) = x + iy \in \mathbb{C}$

$$i(x, y) = i(x + iy) = -y + ix = (-y, x).$$

Der Vektor $(-y, x)$ entsteht anschaulich aus (x, y) durch Drehung um $90°$ im mathematisch positiven Sinn, das heißt gegen den Uhrzeigersinn.

Die Körpergesetze in $\mathbb{C}$ folgen leicht aus den Definitionen, nur die Bestätigung des Assoziativgesetzes der Multiplikation erfordert etwas Rechenarbeit. Um das inverse Element der Multiplikation anzugeben, ist ein weiterer Begriff nützlich: für $z = x + iy \in \mathbb{C}$ heißt $\overline{z} = x - iy \in \mathbb{C}$ die zu z konjugiert komplexe Zahl. Anschaulich ergibt sich $\overline{z}$ aus z durch Spiegelung an der x-Achse, insbesondere gilt $\overline{\overline{z}} = z$.

Lemma 3.3 *Für die komplexe Konjugation gelten folgende Regeln:*

(1) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, $\overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$,

(2) $\overline{z} = z \quad \Leftrightarrow \quad z \in \mathbb{R}$,

(3 *Für* $z = x + iy$ *ist* $\operatorname{Re} z = \frac{1}{2}(z + \overline{z})$ *und* $\operatorname{Im} z = \frac{1}{2i}(z - \overline{z})$.

Beweis: Die Beweise erfolgen alle durch Nachrechnen, zum Beispiel gilt

$$\overline{z_1}\,\overline{z_2} = (x_1 - iy_1)(x_2 - iy_2) = x_1x_2 - y_1y_2 - i(x_1y_2 + y_1x_2) = \overline{z_1 z_2}.$$

□

Die komplexen Zahlen entstehen aus $\mathbb{R}^2$ durch Hinzunahme der Multiplikation als zusätzliche Struktur. Man nennt die Euklidische Norm

$$|z| = \sqrt[+]{x^2 + y^2} \quad \text{für } z = x + iy$$

auch den Betrag der komplexen Zahl z.

Lemma 3.4 *Für den Betrag einer komplexen Zahl gelten folgende Regeln:*

(1) $|z|^2 = z\,\overline{z}$.

(2) $|z_1 z_2| = |z_1||z_2|$.

(3) $|\operatorname{Re} z| \le |z|$, $|\operatorname{Im}(z)| \le |z|$.

Beweis: Für Aussage (1) berechnen wir mit $z = x + iy$

$$z\overline{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2.$$

Die Formel (2) folgt aus (1), Lemma 3.3(1) und den Körpergesetzen:

$$|z_1 z_2|^2 = (z_1 z_2)(\overline{z_1 z_2}) = z_1\,\overline{z_1}\,z_2\overline{z_2} = |z_1|^2\,|z_2|^2.$$

Die Ungleichungen (3) sind offensichtlich. □

Damit können wir nun das inverse Element der Multiplikation leicht hinschreiben, und zwar folgt aus Lemma 3.4(1) für $z = x + iy \neq 0$

$$\frac{1}{z} = \frac{\overline{z}}{z\,\overline{z}} = \frac{x - iy}{x^2 + y^2}.$$

4 Reihen

Viele Funktionen in der Analysis können als unendliche Reihen definiert beziehungsweise dargestellt werden. Darum ist es wichtig, die Konvergenzfrage für Reihen zu untersuchen. Als erste Anwendung der Konvergenzaussagen definieren wir die Exponentialfunktion. Um später den Zusammenhang zu den trigonometrischen Funktionen zu sehen, arbeiten wir direkt in $\mathbb{C}$ statt nur in $\mathbb{R}$, zumal sich die Argumente nur wenig unterscheiden.

Definition 4.1 *Eine Folge $(S_n)_{n\in\mathbb{N}_0}$ heißt Reihe mit Gliedern $a_n \in \mathbb{C}$, falls gilt:*

$$S_n = \sum_{k=0}^{n} a_k \quad \textit{für alle } n \in \mathbb{N}_0.$$

Die Reihe heißt konvergent mit Wert $S \in \mathbb{C}$, wenn die Folge $(S_n)_{n\in\mathbb{N}_0}$ gegen S konvergiert:

$$\sum_{k=0}^{\infty} a_k := \lim_{n\to\infty} \sum_{k=0}^{n} a_n = \lim_{n\to\infty} S_n = S.$$

Die Zahl S_n wird auch als n-te Partialsumme der Reihe bezeichnet. Oft wird die Reihe in der Form $a_0 + a_1 + a_2 + \ldots$ durch ihre ersten Glieder angegeben. Leider ist es auch üblich, die Reihe selbst - unabhängig von der Frage der Konvergenz - ebenfalls mit dem Symbol $\sum_{k=0}^{\infty} a_k$ zu bezeichnen.

Beispiel 4.1 Die Reihe $\dfrac{1}{1\cdot 2} + \dfrac{1}{2\cdot 3} + \dfrac{1}{3\cdot 4} + \ldots$ wird auch mit $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ bezeichnet. Gemeint ist jeweils die Folge $(S_n)_{n\in\mathbb{N}}$ mit den Gliedern

$$S_1 = \frac{1}{1\cdot 2} = \frac{1}{2}, \quad S_2 = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} = \frac{2}{3}, \quad S_3 = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} = \frac{3}{4}, \ldots$$

Für die betrachtete Reihe gilt, für alle $n \in \mathbb{N}$,

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

Also ist die Reihe konvergent mit Wert

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

Die Regeln für die Addition von konvergenten Reihen sowie die Multiplikation von konvergenten Reihen mit reellen oder komplexen Zahlen ergeben sich direkt aus den entsprechenden Regeln für konvergente Folgen, siehe Satz 1.3 a) in Kapitel 1. So ist für konvergente Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ und $\lambda, \mu \in \mathbb{R}$ auch die Reihe $\sum_{k=0}^{\infty} (\lambda a_k + \mu b_k)$ konvergent und es gilt

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

Durch Abänderung, Hinzufügen oder Weglassen von endlich vielen Gliedern in einer Reihe $\sum_{k=0}^{\infty} a_k$ wird die Konvergenz oder Divergenz nicht beeinflusst, sondern nur der Wert der Reihe. Zum Beispiel haben wir für $n \in \mathbb{N}_0$ und $m \geq n$

$$\sum_{k=n}^{m} a_k = \sum_{k=0}^{m} a_k - \sum_{k=0}^{n-1} a_k,$$

und mit $m \to \infty$ folgt, falls die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert,

$$\sum_{k=n}^{\infty} a_k = \sum_{k=0}^{\infty} a_k - \sum_{k=0}^{n-1} a_k. \tag{4.1}$$

Wie bei Folgen kann der Anfangsindex einer Reihe statt $k = 0$ auch eine andere Zahl sein, zum Beispiel $k = 1$ wie in Beispiel 4.1 Allgemein ist eine Reihe nichts anderes als eine Folge $(S_n)_{n\in\mathbb{N}_0}$, die gegeben ist durch die Differenzen $a_n = S_{n+1} - S_n$ und das erste Folgenglied $S_0 = a_0$. Es stellen sich nun zwei Fragen:

- Wie kann ich den Gliedern a_n der Reihe ansehen, ob die Reihe konvergiert bzw. divergiert?
- Im Fall der Konvergenz: welchen Wert hat die Reihe?

Bei der zweiten Frage ist zum Beispiel gemeint, ob eine bereits definierte Zahl wie $\sqrt{2}$, e, π, ... als Grenzwert einer Reihe dargestellt werden kann. Im Folgenden steht aber die erste Frage im Zentrum des Interesses. Dabei ist das nächste Beispiel fundamental.

Beispiel 4.2 (Geometrische Reihe) Die geometrische Reihe $\sum_{k=0}^{\infty} z^k$ mit $z \in \mathbb{C}$ konvergiert genau für $|z| < 1$. Der Konvergenzbeweis ist derselbe wie in Beispiel 1.8, und zwar folgt aus Beispiel 2.2

$$\sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z} \to \frac{1}{1-z} \qquad \text{mit } n \to \infty.$$

Dagegen gilt für $|z| \geq 1$ mit $S_n = \sum_{k=0}^{n} z^k$

$$|S_{n+1} - S_n| = |z^{n+1}| = |z|^{n+1} \geq 1,$$

so dass $(S_n)_{n\in\mathbb{N}}$ keine Cauchyfolge sein kann.

Beispiel 4.3 (Unendliche Dezimalbrüche) Ist $(k_j)_{j\in\mathbb{N}_0}$ eine Folge von Ziffern $k_j \in \{0, 1, \ldots, 9\}$, so ist die Reihe $\sum_{j=0}^{\infty} k_j\, 10^{-j}$ konvergent, vgl. Satz 2.3.

Beispiel 4.4 (Harmonische Reihe) Die harmonische Reihe $\sum_{k=1}^{\infty} 1/k$ ist bestimmt divergent gegen $+\infty$. Dies zeigen wir, indem wir in den Partialsummen wie folgt Klammern setzen:

$$\underbrace{\left(\frac{1}{1}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{2} + \frac{1}{3}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7}\right)}_{\geq 1/2} + \underbrace{\left(\frac{1}{8} + \frac{1}{9} + \ldots + \frac{1}{15}\right)}_{\geq 1/2} + \ldots$$

Die Summe der $1/k$ mit $2^m \leq k < 2^{m+1}$ ist nach unten abgeschätzt durch $2^m \cdot 2^{-(m+1)} = 1/2$.

Nach diesen ersten Beispielen kommen nun zur allgemeinen Konvergenzfrage für Reihen, und beginnen mit einem notwendigen Kriterium.

Satz 4.1 (Nullfolgentest) *Ist die Reihe $\sum_{k=0}^{\infty} a_k$ konvergent, so folgt $\lim_{k\to\infty} a_k = 0$.*

Beweis: Nach Voraussetzung ist die Folge der Partialsummen $S_n = \sum_{k=0}^n a_k$ konvergent mit Grenzwert $S \in \mathbb{C}$, also folgt $a_n = S_n - S_{n-1} \to S - S = 0$. $\square$

Als nächstes formulieren wir unsere Konvergenzkriterien für Folgen neu in der Situation von Reihen. Die Cauchyfolgeneigenschaft sieht wie folgt aus.

Satz 4.2 (Konvergenzkriterium von Cauchy) *Eine Reihe $\sum_{k=0}^\infty a_k$ konvergiert dann und nur dann, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{R}$ gibt, so dass gilt:*

$$\left| \sum_{k=n}^{m} a_k \right| < \varepsilon \quad \text{für alle } n, m \in \mathbb{N}_0 mit m \geq n > N.$$

Beweis: Satz 2.2. $\square$

Satz 4.3 (Reihen mit Gliedern $a_k \geq 0$) *Eine reelle Reihe $\sum_{k=0}^\infty a_k$ mit $a_k \geq 0$ für alle k konvergiert genau dann, wenn die Folge der Partialsummen $S_n = \sum_{k=0}^n a_k$ nach oben beschränkt ist.*

Beweis: Da $a_k \geq 0$, ist die Folge (S_n) der Partialsummen monoton wachsend. Die Behauptung folgt aus Kapitel 1, Satz 2.5 und Satz 1.2. $\square$

Beispiel 4.5 Für $s > 1$ ist die Reihe $\sum_{k=1}^\infty k^{-s}$ konvergent. Wie summieren wie in Beispiel 4.4 über $2^m \leq k < 2^{m+1}$ und erhalten pro Abschnitt

$$\sum_{2^m \leq k < 2^{m+1}} k^{-s} \leq 2^m \cdot (2^m)^{-s} = (2^{1-s})^m.$$

Es folgt für $n < 2^{M+1}$ mit der geometrischen Reihe, Beispiel 1.8,

$$\sum_{k=1}^{n} k^{-s} \leq \sum_{m=0}^{M} \left(\sum_{2^m \leq k < 2^{m+1}} k^{-s} \right) \leq \sum_{m=0}^{M} (2^{1-s})^m \leq \frac{1}{1 - 2^{1-s}}.$$

Die Behauptung folgt nun aus Satz 4.3. Wir haben damit eine wohldefinierte Funktion

$$\zeta : (1, \infty) \to \mathbb{R},\, \zeta(s) = \sum_{k=1}^{\infty} k^{-s},$$

die sogenannte Riemannsche Zetafunktion. Sie spielt bei der Untersuchung von Primzahlen eine fundamentale Rolle.

Die Reihe $\sum_{k=1}^\infty (-1)^{k-1}/k = 1 - 1/2 + 1/3 - + \ldots$ ist ein Beispiel für eine sogenannte alternierende Reihe. Während die entsprechende Reihe mit nur positiven Vorzeichen nicht konvergiert – es ist die harmonische Reihe aus Beispiel 4.4 – ist die Reihe mit dem Vorzeichenwechsel konvergent. Dies ergibt sich aus dem nächsten Satz, bei dessen Beweis wieder Monotonieargumente eine wesentliche Rolle spielen.

Satz 4.4 (Leibniz-Kriterium für alternierende Reihen) *Sei a_k, $k \in \mathbb{N}_0$, eine reelle, monoton fallende Nullfolge (also insbesondere $a_k \geq 0$). Dann ist die Reihe $\sum_{k=0}^\infty (-1)^k a_k$ konvergent. Außerdem gilt für alle $n \in \mathbb{N}_0$ die Abschätzung*

$$0 \leq (-1)^n \sum_{k=n}^{\infty} (-1)^k a_k \leq a_n.$$

BEWEIS: Wir betrachten die Partialsummen $S_n = \sum_{k=0}^n (-1)^k a_k$. Für alle $n \in \mathbb{N}_0$ gelten die Ungleichungen

$$\begin{aligned}
S_{2n+2} - S_{2n} &= (-1)^{2n+2} a_{2n+2} + (-1)^{2n+1} a_{2n+1} = a_{2n+2} - a_{2n+1} \leq 0, \\
S_{2n+3} - S_{2n+1} &= (-1)^{2n+3} a_{2n+3} + (-1)^{2n+2} a_{2n+2} = a_{2n+2} - a_{2n+3} \geq 0, \\
S_{2n+1} - S_{2n} &= (-1)^{2n+1} a_{2n+1} = -a_{2n+1} \leq 0.
\end{aligned}$$

Also ist die Folge $(S_{2n})_{n\in\mathbb{N}_0}$ monoton fallend, die Folge $(S_{2n+1})_{n\in\mathbb{N}_0}$ monoton wachsend. Wegen $S_{2n} \geq S_{2n+1}$ ist die Folge S_{2n} nach unten beschränkt durch S_1, die Folge S_{2n+1} nach oben beschränkt durch S_0. Nach Satz 2.5 sind die Folgen S_{2n} sowie S_{2n+1} konvergent. Aber $S_{2n+1} - S_{2n} \to 0$, das heißt beide Folgen, und damit die gesamte Folge, konvergieren gegen denselben Grenzwert $S \in \mathbb{R}$. Nun ist $S \in [S_1, S_0]$ mit $S_0 = a_0$ und $S_1 = a_1 - a_0 \geq 0$, also gilt die Abschätzung im Fall $n = 0$. Anwendung auf die Folge $b_k = a_{n+k}$ liefert nun die Abschätzung für beliebige $n \in \mathbb{N}_0$. □

Bei der alternierenden Reihe $1 - 1/2 + 1/3 - + \ldots$ stößt man auf Merkwürdigkeiten, wenn man die Summationsreihenfolge ändert. Während die Ausgangsreihe konvergent ist, ist die durch Umordnung entstehende Reihe

$$\left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} + \frac{1}{5} - \frac{1}{4}\right) + \left(\frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} - \frac{1}{6}\right) + \ldots.$$

bestimmt divergent gegen $+\infty$, denn die Summe der positiven Zahlen in der m-ten Klammer ist mindestens $2^{m-1} \cdot 2^{-(m+1)} = 1/4$. Dies ist ein interessantes Phänomen, jedoch sind wir in erster Linie an Reihen interessiert, deren Konvergenz stabiler ist. Dabei ist der folgende Begriff zentral.

Definition 4.2 *Die Reihe $\sum_{k=0}^\infty a_k$ mit $a_k \in \mathbb{C}$ heißt absolut konvergent, wenn die Reihe $\sum_{k=0}^\infty |a_k|$ konvergiert, das heißt $\sum_{k=0}^\infty |a_k| < \infty$.*

Satz 4.5 (absolut konvergent ⇒ konvergent) *Wenn die Reihe $\sum_{k=0}^\infty a_k$ absolut konvergiert, so ist sie konvergent und es gilt*

$$\left|\sum_{k=0}^\infty a_k\right| \leq \sum_{k=0}^\infty |a_k|.$$

BEWEIS: Die Dreiecksungleichung besagt

$$\left|\sum_{k=n}^m a_k\right| \leq \sum_{k=n}^m |a_k| \quad \text{für } m \geq n \geq 0.$$

Aus dem Cauchykriterium für $\sum_{k=0}^\infty |a_k|$ folgt deshalb das Cauchykriterium für $\sum_{k=0}^\infty a_k$, das heißt die Reihe konvergiert nach Satz 4.2. Die Abschätzung folgt, indem wir in der Dreiecksungleichung $n = 0$ setzen und $m \to \infty$ gehen lassen. □

Der folgende Satz fasst die wesentlichen Kriterien für die absolute Konvergenz von Reihen zusammen.

Satz 4.6 (Tests für absolute Konvergenz) *Sei $\sum_{k=0}^\infty a_k$ eine Reihe mit $a_k \in \mathbb{C}$. Ist eine der drei folgenden Bedingungen erfüllt, so ist $\sum_{k=0}^\infty a_k$ absolut konvergent:*

(a) Majorantenkriterium (M-Test): *Es gilt* $|a_k| \leq c_k \in [0,\infty)$ *mit* $\sum_{k=0}^\infty c_k < \infty$.

(b) Quotientenkriterium: *Es gibt ein* $\theta \in [0,1)$ *und ein* $n \in \mathbb{N}$ *mit*

$$\frac{|a_{k+1}|}{|a_k|} \leq \theta \textit{ für alle } k \geq n \quad (\textit{wobei } a_k \neq 0 \textit{ für } k \geq n).$$

(c) Wurzelkriterium: *Es gibt ein* $\theta \in [0,1)$ *und ein* $n \in \mathbb{N}$ *mit* $\sqrt[k]{|a_k|} \leq \theta$ *für alle* $k \geq n$.

Umgekehrt ist die Reihe divergent, wenn für ein $n \in \mathbb{N}$ *gilt:*

$$\frac{|a_{k+1}|}{|a_k|} \geq 1 \quad \textit{oder} \quad \sqrt[k]{|a_k|} \geq 1 \textit{ für alle } k \geq n.$$

Für die Reihe $\sum_{k=1}^\infty k^{-s}$ mit $s \geq 1$ haben wir

$$\frac{|a_{k+1}|}{|a_k|} = \left(\frac{k}{k+1}\right)^s < 1 \text{ für alle } k \quad \text{und} \quad \lim_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} = 1.$$

Für $s = 1$ gilt Divergenz nach Beispiel 4.4; also reicht es in (b) *nicht*, nur die Abschätzung $|a_{k+1}|/|a_k| < 1$ vorauszusetzen. Andererseits konvergiert die Reihe für $s > 1$ nach Beispiel 4.5, also kann aus $\lim_{k\to\infty} |a_{k+1}|/|a_k| = 1$ auch nicht auf Divergenz geschlossen werden. Eine ganz analoge Diskussion gilt für das Wurzelkriterium. Die Bedingungen (b) beziehungsweise (c) sind äquivalent zu den folgenden Voraussetzungen:

$$\limsup_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} < 1 \quad \text{bzw.} \quad \limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1.$$

Beweis des Satzes: (a) folgt aus Satz 4.3, denn für alle $n \in \mathbb{N}$ hat man die obere Schranke

$$\sum_{k=0}^n |a_k| \leq \sum_{k=0}^n c_k \leq \sum_{k=0}^\infty c_k < \infty.$$

Voraussetzung (b) liefert per Induktion

$$|a_k| \leq \theta^{k-n}|a_n| \quad \text{für } k \geq n.$$

Da die geometrische Reihe wegen $0 \leq \theta < 1$ nach Beispiel 1.8 konvergiert, folgt die Behauptung aus (a) und wir erhalten außerdem die Abschätzung

$$\sum_{k=n}^\infty |a_k| \leq \frac{|a_n|}{1-\theta}. \tag{4.2}$$

Unter der Voraussetzung (c) gilt

$$|a_k| \leq \theta^k \quad \text{für } k \geq n.$$

Wieder folgt die Behauptung durch durch M-Test mit der geometrischen Reihe. Die Abschätzung lautet hier

$$\sum_{k=n}^\infty |a_k| \leq \frac{\theta^n}{1-\theta}. \tag{4.3}$$

Die Divergenzaussagen folgen unmittelbar aus dem Nullfolgentest, Satz 4.1. $\square$

Als Anwendung kommt nun endlich die

Satz 4.7 (Definition der Exponentialfunktion) *Die Reihe* $\sum_{k=0}^{\infty} z^k/k!$ *ist für alle* $z \in \mathbb{C}$ *absolut konvergent; damit ist die Exponentialfunktion*

$$\exp : \mathbb{C} \to \mathbb{C},\ \exp(z) = \sum_{n=0}^{\infty} \frac{z^k}{k!}$$

wohldefiniert. Für jedes $n \in \mathbb{N}$ *und* $|z| \leq (n+1)/2$ *gilt die Abschätzung*

$$\left| \exp(z) - \sum_{k=0}^{n-1} \frac{z^k}{k!} \right| \leq \frac{2|z|^n}{n!}.$$

Beweis: Für $z = 0$ ist nichts zu zeigen. Seien $z \in \mathbb{C}\backslash\{0\}$ und $n \in \mathbb{N}$ mit $|z| \leq (n+1)/2$. Dann gilt für $k \geq n$ mit $a_k = z^k/k!$

$$\frac{|a_{k+1}|}{|a_k|} = \frac{|z|}{k+1} \leq \frac{1}{2}.$$

Aus dem Quotientenkriterium, Satz 4.6(b), folgt die absolute Konvergenz. Außerdem liefert die Ungleichung (4.2), hier mit $\theta = 1/2$ und $a_n = z^n/n!$,

$$\sum_{k=n}^{\infty} \left| \frac{z^k}{k!} \right| \leq \frac{|z^n/n!|}{1 - 1/2} = \frac{2|z|^n}{n!},$$

das heißt

$$\left| \exp(z) - \sum_{k=0}^{n-1} \frac{z^k}{k!} \right| = \left| \sum_{k=n}^{\infty} \frac{z^k}{k!} \right| \leq \sum_{k=n}^{\infty} \left| \frac{z^k}{k!} \right| \leq \frac{2|z|^n}{n!}.$$

□

Wegen $\mathbb{R} \subset \mathbb{C}$ ist durch Satz 4.7 auch die reelle Exponentialfunktion

$$\exp : \mathbb{R} \to \mathbb{R},\ \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

erklärt. Nach Definition 2.2 gilt $\exp(1) = e \approx 2,7\ldots$. Außerdem sieht man wie in Beispiel 2.3

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n \quad \text{für alle } x \in \mathbb{R}.$$

Eine charakteristische Eigenschaft der Exponentialfunktion ist die Funktionalgleichung $\exp(z+w) = \exp(z)\exp(w)$. Um diese zu beweisen, brauchen wir einen Satz zur Multilplikation von Reihen.

Bei der Multiplikation von zwei Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ treten die doppelt indizierten Produkte $a_k b_l$ mit $k, l \in \mathbb{N}_0$ auf. Es ist zunächst nicht klar, in welcher Reihenfolge diese summiert werden sollen. Schreiben wir die Produkte in einem Schema auf, so dass $a_k b_l$ in der k-ten Zeile und l-ten Spalte steht, so ist folgendes Summationsverfahren naheliegend: erst werden die $n+1$ Produkte $a_k b_l$ in jeder Diagonale $k + l = n$ addiert, dann wird die Konvergenz der resultierenden Reihe studiert. Im folgenden sind stets $k, l \in \mathbb{N}_0$ und wir schreiben zum Beispiel kurz $\sum_{k+l=n}$ für die Summe über alle Paare $k, l \in \mathbb{N}_0$ mit $k + l = n$.

Satz 4.8 (Cauchyprodukt) *Die Reihen* $\sum_{k=0}^{\infty} a_k$ *und* $\sum_{k=0}^{\infty} b_k$ *seien absolut konvergent. Dann ist auch die Reihe*

$$\sum_{k=0}^{\infty} c_n \quad \mathit{mit} \quad c_n = \sum_{k+l=n} a_k b_l = \sum_{k=0}^{n} a_k b_{n-k}$$

absolut konvergent, und es gilt

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \left(\sum_{k+l=n} a_k\, b_l \right) = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right).$$

Beweis: Wir setzen für $N \in \mathbb{N}_0$

$$A_N := \sum_{k=0}^{N} |a_k| \to \sum_{k=0}^{\infty} |a_k| =: A \quad \text{und} \quad B_N := \sum_{k=0}^{N} |b_k| \to \sum_{k=0}^{\infty} |b_k| =: B,$$

wobei nach Voraussetzung $A, B < \infty$. Es folgt

$$\sum_{n=0}^{N} |c_n| \leq \sum_{k+l \leq N} |a_k|\,|b_l| \leq \sum_{k,l \leq N} |a_k|\,|b_l| = A_N B_N \leq AB < \infty.$$

Nach Satz 4.3 ist $\sum_{n=0}^{\infty} c_n$ absolut konvergent und insbesondere konvergent (Satz 4.5). Um den Grenzwert zu identifizieren, reicht es also die geraden Partialsummen $\sum_{n=0}^{2N} c_n$ zu betrachten. Wir berechnen

$$\left| \left(\sum_{k=0}^{2N} a_k \right) \cdot \left(\sum_{l=0}^{2N} b_l \right) - \sum_{n=0}^{2N} c_n \right| = \left| \sum_{k,l \leq 2N} a_k b_l - \sum_{k+l \leq 2N} a_k b_l \right| \leq \sum_{k,l \leq 2N, \max(k,l) > N} |a_k|\,|b_l|,$$

Die rechte Seite ist gleich $A_{2N}\, B_{2N} - A_N\, B_N$, konvergiert also gegen Null für $N \to \infty$. Für die letzte Abschätzung ist es hilfreich, die Indexbereiche zu skizzieren. Jedenfalls ist damit die gewünschte Formel für das Cauchyprodukt bewiesen. □

Wir wenden nun den Produktsatz auf die Exponentialfunktion an.

Satz 4.9 (Funktionalgleichung der Exponentialfunktion) *Für alle* $z, w \in \mathbb{C}$ *gilt*

$$\exp(z + w) = \exp(z) \exp(w).$$

Beweis: Da die Reihen absolut konvergieren, erhalten wir mit Satz 4.8 und der Binomischen Formel, Satz 2.6,

$$\exp(z) \exp(w) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z^k}{k!} \frac{w^{n-k}}{(n-k)!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} = \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = \exp(z+w).$$

□

Wir wollen daraus direkt einige Eigenschaften der Exponentialfunktion herleiten. Dazu brauchen wir noch eine einfache Tatsache.

Lemma 4.1 *Für $a_k \in \mathbb{C}$ sei die Reihe $\sum_{k=0}^{\infty} a_k$ konvergent. Dann ist auch die Reihe $\sum_{k=0}^{\infty} \overline{a}_k$ konvergent und mit $S = \sum_{k=0}^{\infty} a_k$ gilt*

$$\overline{S} = \sum_{k=0}^{\infty} \overline{a}_k.$$

Beweis: Mit $S_n = \sum_{k=0}^{n} a_k$ gilt $\overline{S}_n = \sum_{k=0}^{\infty} \overline{a}_k$ und (wie für jede Folge)

$$|\overline{S} - \overline{S}_n| = |\overline{S - S_n}| = |S - S_n| \to 0 \quad \text{mit } n \to \infty.$$

□

Folgerung 4.1 *Die Exponentialfunktion hat folgende Eigenschaften:*

(1) $\exp(z)\exp(-z) = 1$ *für alle* $z \in \mathbb{C}$, *insbesondere* $\exp(z) \neq 0$.

(2) $\exp(x) > 0$ *für alle* $x \in \mathbb{R}$.

(3) $|exp(iy)| = 1$ *für alle* $y \in \mathbb{R}$.

(4) $\exp(pz) = (\exp(z))^p$ *für alle* $z \in \mathbb{C}$, $p \in \mathbb{Z}$.

Beweis: Behauptung (1) folgt sofort aus Satz 4.9, denn

$$\exp(z)\exp(-z) = \exp\big(z + (-z)\big) = \exp(0) = 1.$$

Für $x > 0$ ist trivialerweise $\exp(x) \in (1, \infty)$, und für $x < 0$ folgt $\exp(x) = 1/\exp(-x) \in (0,1)$ aus Gleichung (1). Aus Lemma 4.1 folgt

$$\overline{\exp(z)} = \overline{\left(\sum_{k=0}^{\infty} \frac{z^k}{k!}\right)} = \sum_{k=0}^{\infty} \overline{\left(\frac{z^k}{k!}\right)} = \sum_{k=0}^{\infty} \frac{\overline{z}^k}{k!} = \exp(\overline{z}).$$

Gleichung (3) ergibt sich nun mit Lemma 3.4:

$$|\exp(iy)|^2 = \exp(iy)\overline{exp(iy)} = \exp(iy)\exp(\overline{iy}) = \exp(iy)\exp(-iy) = 1.$$

Für $p \in \mathbb{N}_0$ folgt Behauptung (4) mit der Funktionalgleichung durch Induktion, und ebenfalls mit der Funktionalgleichung gilt für $p \in \mathbb{N}$

$$1 = \exp(pz)\exp(-pz) = (\exp(z))^p \exp(-pz)), \text{ also } \exp(-pz) = (\exp(z)^p)^{-1} = \exp(z)^{-p}.$$

Damit sind alle Aussagen bewiesen. □

In der Sprache der Algebra besagt die Funktionalgleichung, in Verbindung mit Folgerung 4.1(1), dass die Exponentialfunktion ein Gruppenhomomorphismus von der Gruppe $(\mathbb{C}, +)$ in die Gruppe $(\mathbb{C}\backslash\{0\}, \cdot)$ ist. Die reelle Exponentialabbildung ist analog ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ nach $(\mathbb{R}^+, \cdot)$, wobei $\mathbb{R}^+$ die positiven reellen Zahlen bezeichnet.

Bislang haben wir außer neuerdings der Exponentialfunktion nur Polynome beziehungsweise stückweise Polynome als Funktionen zur Verfügung. Es ist naheliegend, weitere neue Funktionen ebenfalls als Grenzwerte von Folgen bzw. Reihen von Polynomen zu suchen. Dies führt auf den Begriff der Potenzreihe.

Definition 4.3 (Potenzreihen) *Eine komplexe Potenzreihe ist eine Reihe der Form*

$$P(z) = \sum_{k=0}^{\infty} a_k z^k \qquad \text{mit } a_k \in \mathbb{C}.$$

Allgemeiner betrachtet man auch Potenzreihen mit Entwicklungspunkt $z_0 \in \mathbb{C}$, das heißt Reihen der Form $P(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$. Der Einfachheit halber beschränken wir uns jedoch hier auf den Fall $z_0 = 0$.

Definition 4.3 lässt offen, ob beziehungsweise für welche $z \in \mathbb{C}$ die Reihe tatsächlich konvergiert. Natürlich ist $P(0) = a_0$, es kann aber sein, dass die Reihe für alle $z \neq 0$ divergiert, dann ist sie natürlich nicht von Interesse. Ein Beispiel ist $\sum_{n=0}^{\infty} n^n z^n$, wie der Nullfolgentest sofort ergibt. Bei interessanten Reihen (was immer das ist) erwarten wir aber Konvergenz zumindest für manche $z \in \mathbb{C}$. Auf dem Konvergenzgebiet erhalten wir dann eine Funktion, und genau daran sind wir interessiert. Zum Beispiel konvergiert die Exponentialreihe sogar für alle $z \in \mathbb{C}$, und definiert die Exponentialfunktion. Erstaunlicherweise kann über das Konvergenzgebiet einer Potenzreihe eine allgemeine Aussage getroffen werden. Dazu benötigen wir das folgende technische Lemma, dessen Abschätzung (4.5) noch mehrfach benutzt wird.

Lemma 4.2 *Sei $P(z) = \sum_{k=0}^{\infty} a_k z^k$ eine Potenzreihe mit $a_k \in \mathbb{C}$, die für ein $z_0 \neq 0$ konvergiert. Dann gibt es ein $M \in [0, \infty)$ mit*

$$|a_k|\,|z_0|^k \leq M \quad \text{für alle } k \in \mathbb{N}_0. \tag{4.4}$$

Aus (4.4) folgt weiter, dass $P(z)$ für $|z| < |z_0|$ absolut konvergiert und dass für $n \in \mathbb{N}_0$ gilt:

$$|P(z) - P_n(z)| \leq \sum_{k=n+1}^{\infty} |a_k|\,|z|^k \leq \frac{M}{1 - \frac{|z|}{|z_0|}} \left(\frac{|z|}{|z_0|}\right)^{n+1}. \tag{4.5}$$

Hier ist $P_n(z) = \sum_{k=0}^{n} a_k z^k$ die n-te Partialsumme.

Beweis: Für $|z| < |z_0|$ gilt $|a_k z^k| = |a_k|\,|z_0|^k (|z|/|z_0|)^k \leq M(|z|/|z_0|)^k$. Wegen $|z|/|z_0| < 1$ folgt die absolute Konvergenz aus dem Majorantenkriterium durch Vergleich mit der geometrischen Reihe. Genauer gilt

$$|P(z) - P_n(z)| \leq \sum_{k=n+1}^{\infty} |a_k||z|^k \leq M \sum_{k=n+1}^{\infty} \left(\frac{|z|}{|z_0|}\right)^k \leq \frac{M}{1 - \frac{|z|}{|z_0|}} \left(\frac{|z|}{|z_0|}\right)^{n+1}.$$

□

Satz 4.10 (vom Konvergenzradius) *Zu jeder Potenzreihe $P(z) = \sum_{k=0}^{\infty} a_k z^k$ gibt es genau ein $R \in [0, \infty]$, den Konvergenzradius, mit folgender Eigenschaft:*

$$P(z) \text{ ist } \begin{cases} \text{absolut konvergent} & \text{für } |z| < R, \\ \text{divergent} & \text{für } |z| > R. \end{cases}$$

Beweis: Die Eindeutigkeit von R ist klar. Zur Existenz definieren wir

$$R = \sup\{|z| : P(z) \text{ konvergiert}\} \in [0, \infty].$$

Ist $|z| < R$, so gibt es nach Definition ein $z_0 \in \mathbb{C}$ mit $|z| < |z_0| \leq R$, so dass $P(z_0)$ konvergiert. Also konvergiert $P(z)$ absolut wegen Lemma 4.2. Andererseits ist die Reihe divergent für $|z| > R$ nach Definition von R. □

Beispiel 4.6 Als weiteres Beispiel einer Potenzreihe betrachten wir für beliebiges $\alpha \in \mathbb{C}$ die Binomialreihe

$$B_\alpha(z) = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k = 1 + \alpha z + \frac{\alpha(\alpha-1)}{2} z^2 + \ldots$$

Für $\alpha \in \mathbb{N}_0$ bricht die Reihe nach $k = \alpha$ ab, und die Binomische Formel aus Satz 2.6 liefert $B_\alpha(z) = (1+z)^\alpha$. Im folgenden sei nun $\alpha \notin \mathbb{N}_0$. Für $z \neq 0$ ist dann $a_k = \binom{\alpha}{k} z^k \neq 0$ für alle $k \in \mathbb{N}_0$, und es gilt

$$\left|\frac{a_{k+1}}{a_k}\right| = \frac{|\alpha - k|}{k+1}|z| \to |z| \quad \text{mit } k \to \infty.$$

Nach dem Quotientenkriterium konvergiert die Reihe für $|z| < 1$ und divergiert für $|z| > 1$, das heißt der Konvergenzradius ist $R = 1$.

Für eine gegebene reelle Funktion $f : (-R, R) \to \mathbb{R}$ stellt sich die Frage, ob die Funktion durch eine Potenzreihe dargestellt werden kann: gibt es $a_k \in \mathbb{R}$ mit $f(x) = \sum_{k=0}^{\infty} a_k x^k$ für $x \in (-R, R)$, oder zumindest auf einem kleineren Intervall $x \in (-r, r)$? Wenn ja, sind die a_k eindeutig bestimmt? Natürlich kann diese Frage auch über $\mathbb{C}$ gestellt werden, anstelle der Intervalle treten dann Kreisscheiben $B_R(0) = \{z \in \mathbb{C} : |z| < R\}$. Das folgende Ergebnis zeigt, dass bei weitem nicht jede Funktion als Potenzreihe darstellbar ist. Zu der Folgerung vergleiche auch die entsprechende Aussage für Polynome, Satz **??**.

Satz 4.11 (Identitätssatz für Potenzreihen) *Sei $P(z) = \sum_{k=0}^{\infty} a_k z^k$ eine Potenzreihe, die für ein $z_0 \neq 0$ konvergiert. Ist $0 \in \mathbb{C}$ Häufungspunkt der Menge $\{z \in \mathbb{C} : P(z) = 0\}$, so folgt $a_k = 0$ für alle $k \in \mathbb{N}_0$.*

Beweis: Wir nehmen induktiv an, dass schon $a_0 = a_1 = \ldots = a_{n-1} = 0$ gezeigt ist, wobei der Fall $n = 0$ den Induktionsanfang liefert. Für $|z| \leq |z_0|/2$ folgt aus (4.5) die Abschätzung

$$|P(z) - a_n z^n| = |P(z) - P_n(z)| \leq C|z|^{n+1} \quad \text{wobei } C = 2M|z_0|^{-(n+1)}.$$

Nach Voraussetzung gilt $P(z_i) = 0$ für eine Folge $z_i \neq 0$ mit $z_i \to 0$, und Einsetzen von $z = z_i$ ergibt $|a_n|\,|z_i|^n \leq C|z_i|^{n+1}$, also $|a_n| \leq C|z_i| \to 0$ mit $i \to \infty$, das heißt $a_n = 0$. □

Folgerung 4.2 (Koeffizientenvergleich) *Seien $P(z) = \sum_{k=0}^{\infty} a_k z^k$ und $Q(z) = \sum_{k=0}^{\infty} b_k z^k$ Potenzreihen mit positivem Konvergenzradius. Ist der Nullpunkt Häufungspunkt der Menge $\{z \in \mathbb{C} : P(z) = Q(z)\}$, so folgt $a_k = b_k$ für alle $k \in \mathbb{N}_0$.*

Beweis: Die Potenzreihe $F(z) = \sum_{k=0}^{\infty} c_k z^k$ mit $c_k = a_k - b_k$ hat positiven Konvergenzradius, und der Nullpunkt ist Häufungspunkt der Menge $\{z \in \mathbb{C} : F(z) = 0\}$. Die Behauptung folgt damit aus Satz 4.11. $\square$

Zum Schluss dieses Kapitels kommen wir zu der Frage der Umordnung von Reihen zurück und zeigen, dass absolut konvergente Reihen beliebig umgeordnet werden können.

Satz 4.12 (Umordnungssatz) *Sei $\tau : \mathbb{N} \to \mathbb{N}$ bijektiv. Falls $\sum_{k=1}^{\infty} a_k$ absolut konvergent ist, so konvergiert auch die Reihe $\sum_{j=1}^{\infty} a_{\tau(j)}$ absolut und hat denselben Grenzwert.*

Beweis: Setze $S = \sum_{k=1}^{\infty} a_k$. Nach Voraussetzung gilt

$$\sum_{k=n+1}^{\infty} |a_k| = \sum_{k=1}^{\infty} |a_k| - \sum_{k=1}^{n} |a_k| \to 0 \quad \text{mit } n \to \infty,$$

also gibt es zu $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit $\sum_{k=n+1}^{\infty} |a_k| < \varepsilon/2$. Für m hinreichend groß gilt $\tau^{-1}\{1, \ldots, n\} \subset \{1, \ldots, m\}$ beziehungsweise $\{1, \ldots, n\} \subset \tau(\{1, \ldots, m\})$, und es folgt

$$\Big|S - \sum_{j=1}^{m} a_{\tau(j)}\Big| \leq \Big|S - \sum_{j=1, \tau(j) \leq n}^{m} a_{\tau(j)}\Big| + \sum_{j=1, \tau(j) > n}^{m} |a_{\tau(j)}| \leq \underbrace{\Big|S - \sum_{k=1}^{n} a_k\Big|}_{<\varepsilon/2} + \underbrace{\sum_{k=n+1}^{\infty} |a_k|}_{<\varepsilon/2}.$$

$\square$

5 Teilmengen von $\mathbb{R}$ und von $\mathbb{R}^n$

Der $\mathbb{R}^n$ ist eine mathematische Verallgemeinerung:

$$\mathbb{R}^n = \{x = (x_1, \ldots, x_n) : x_i \in \mathbb{R}\} = \underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n \text{ mal}}.$$

Für $x \in \mathbb{R}$ ist $|x|$ der Abstand zum Nullpunkt. Die entsprechende Verallgemeinerung im $\mathbb{R}^n$ ist die Euklidische Länge oder Norm.

Definition 5.1 *Die Euklidische Norm eines Vektors $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ ist*

$$|x| = \sqrt[+]{x_1^2 + \ldots + x_n^2}.$$

Der Euklidische Abstand von zwei Punkten $x, y \in \mathbb{R}^n$ ist $|x - y|$.

Das Argument der Wurzel ist als Summe von Quadraten nichtnegativ, also ist $|x|$ wohldefiniert. Der folgende Satz fasst wesentliche Eigenschaften der Euklidischen Norm zusammen.

Satz 5.1 *Für die Euklidische Norm $|x|$ gilt:*

(1) Positivität: $|x| \geq 0$ *mit Gleichheit genau wenn $x = 0$.*

(2) Halblinearität: $|\lambda x| = |\lambda| \cdot |x|$ *für alle $\lambda \in \mathbb{R}$, $x \in \mathbb{R}^n$.*

(3) Dreiecksungleichung: $|x + y| \leq |x| + |y|$ *für alle $x, y \in \mathbb{R}^n$.*

Gleichheit in der Dreiecksungleichung gilt genau wenn x und y gleichsinnig parallel sind.

Beweis: Nach Definition der Wurzel ist $|x| \geq 0$. Im Gleichheitsfall ist $x_1^2 + \ldots + x_n^2 = 0$, also $x_i = 0$ für alle $i = 1, \ldots, n$ und damit $x = 0$. Die Skalarmultiplikation im $\mathbb{R}^n$ ist komponentenweise definiert durch $\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n)$, also folgt

$$|\lambda x| = \left(\sum_{i=1}^n (\lambda x_i)^2\right)^{1/2} = \left(\lambda^2 \sum_{i=1}^n x_i^2\right)^{1/2} = |\lambda| \left(\sum_{i=1}^n x_i^2\right)^{1/2} = |\lambda| \cdot |x|.$$

Beachten Sie, dass hier das Symbol $|\cdot|$ in zwei Bedeutungen vorkommt: einmal als Betrag der reellen Zahl λ, und einmal als Euklidische Norm des Vektors $x \in \mathbb{R}^n$. Um die Dreiecksungleichung zu beweisen, holen wir etwas weiter aus und führen das Skalarprodukt ein. Vorher erklären wir noch die Bezeichnung: Für drei Punkte x, y, z ergibt sich aus der Dreiecksungleichung

$$|x - z| = |x - y + y - z| \leq |x - y| + |y - z|,$$

das heißt im Dreieck mit den Eckpunkten x, y, z ist jede Seite kürzer als die Summe der beiden anderen Seiten. Im Gleichheitsfall sind $x - y$ und $y - z$ parallel, das heißt die Punkte x, y, z liegen auf einer Geraden.

Definition 5.2 *Das Standardskalarprodukt von $x, y \in \mathbb{R}^n$ ist*

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n.$$

Offenbar gilt $\langle x, x\rangle = |x|^2$ für $x \in \mathbb{R}^n$.

Lemma 5.1 *Für das Standardskalarprodukt $\langle x, y\rangle$ gilt:*

(1) Positivität: $\langle x, x\rangle \geq 0$ *mit Gleichheit genau wenn* $x = 0$.

(2) Symmetrie: $\langle x, y\rangle = \langle y, x\rangle$ *für alle* $x, y \in \mathbb{R}^n$.

(3) Bilinearität: *Für alle* $x, y, z \in \mathbb{R}^n$ *und* $\lambda, \mu \in \mathbb{R}$ *gilt*

$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle \qquad \textit{und} \qquad \langle x, \lambda y + \mu z\rangle = \lambda\langle x, y\rangle + \mu\langle x, z\rangle.$$

Beweis: Die Positivität folgt aus der Positivität der Norm, und die Symmetrie ist offensichtlich. Auch die Bilinearität ergibt sich leicht aus der Definition:

$$\langle \lambda x + \mu y, z\rangle = \sum_{i=1}^{n}(\lambda x_i + \mu y_i)z_i = \lambda\sum_{i=1}^{n} x_i z_i + \mu\sum_{i=1}^{n} y_i z_i = \lambda\langle x, z\rangle + \mu\langle y, z\rangle.$$

$\square$

Es kommt nun eine fundamentale Ungleichung:

Satz 5.2 (Ungleichung von Cauchy-Schwarz) *Für $x, y \in \mathbb{R}^n$ gilt*

$$|\langle x, y\rangle| \leq |x|\,|y|. \tag{5.6}$$

Gleichheit gilt genau dann, wenn x und y linear abhängig sind.

Beweis: Für $a, b \in \mathbb{R}^n$ mit $|a| = |b| = 1$ gilt $|\langle a, b\rangle| \leq 1$, denn die Bilinearität ergibt

$$1 \pm \langle a, b\rangle = \frac{1}{2}\big(|a|^2 + |b|^2 \pm 2\langle a, b\rangle\big) = \frac{1}{2}|a \pm b|^2 \geq 0.$$

Ist nun $x = 0$ oder $y = 0$, so gilt $\langle x, y\rangle = |x|\,|y| = 0$. Andernfalls betrachten wir die Vektoren $a = x/|x|$ sowie $b = y/|y|$. Diese haben Länge Eins, denn nach Halblinearität der Norm gilt

$$\left|\frac{x}{|x|}\right| = \left|\frac{1}{|x|}x\right| = \frac{1}{|x|}\,|x| = 1.$$

Anwendung der gezeigten Ungleichung für a, b liefert

$$\frac{1}{|x|}\frac{1}{|y|}|\langle x, y\rangle| = \left|\frac{1}{|x|}\frac{1}{|y|}\langle x, y\rangle\right| = \left|\left\langle \frac{x}{|x|}, \frac{y}{|y|}\right\rangle\right| = |\langle a, b\rangle| \leq 1,$$

was zu zeigen war. Für $x, y \neq 0$ kann Gleichheit kann nur dann eintreten, wenn $a \pm b = 0$, also wenn x und y parallel sind. $\square$

Damit können wir nun leicht den Beweis der Dreiecksungleichung führen:

$$|x + y|^2 = |x|^2 + 2\langle x, y\rangle + |y|^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2.$$

Mit der Monotonie der Wurzel folgt $|x + y| \leq |x| + |y|$ wie behauptet. Im Gleichheitsfall müssen zunächst x und y parallel sein, außerdem muss aber $\langle x, y\rangle \geq 0$ sein, also sind x, y gleichsinnig parallel.

Im Gegensatz zu $\mathbb{R}$ haben wir im $\mathbb{R}^n$ keine Anordnung zur Verfügung. Deshalb verwenden wir die Ungleichungszeichen $<$, $>$, $\leq$, $\geq$, Formulierungen wie "nach oben bzw. unten beschränkt" und den Begriff der Monotonie *ausschließlich* für reelle Zahlen und *niemals* für Punkte im $\mathbb{R}^n$. Eine Ungleichung $a < b$ ist nur für reelle Zahlen sinnvoll (später auch mal für gewisse Matrizen, aber das tut jetzt nichts zur Sache). Ein Check zeigt jedoch, dass die Begriffe Konvergenz, ε-Umgebung, Beschränktheit, Cauchyfolge und Vollständigkeit mit der *Betragsfunktion* formuliert werden können. Sie lassen sich dann auf den $\mathbb{R}^n$ verallgemeinern, indem wir statt der Betragsfunktion die Euklidische Norm einsetzen.

Definition 5.3 (Konvergenz im $\mathbb{R}^n$) *Die Folge $x_k \in \mathbb{R}^n$ konvergiert gegen $a \in \mathbb{R}^n$, falls gilt:*

Für alle $\varepsilon > 0$ gibt es ein $N \in \mathbb{R}$, so dass für alle $k > N$ gilt: $|x_k - a| < \varepsilon$.

Definition 5.4 (ε-Umgebung in $\mathbb{R}^n$) *Die ε-Umgebung von $a \in \mathbb{R}^n$ ist die Menge*

$$B_\varepsilon(a) = \{x \in \mathbb{R}^n : |x - a| < \varepsilon\}.$$

Sie wird auch als offene Kugel (für $n \geq 3$) bzw. offene Kreisscheibe (für $n = 2$) um a mit Radius $\varepsilon > 0$ bezeichnet. Die abgeschlossene Kugel ist

$$K_\varepsilon(a) = \{x \in \mathbb{R}^n : |x - a| \leq \varepsilon\}.$$

Definition 5.5 (Beschränkte Teilmengen von $\mathbb{R}^n$) *Eine Menge $M \subset \mathbb{R}^n$ heißt beschränkt, wenn gilt:*

Es gibt ein $K \geq 0$ mit $|x| \leq K$ für alle $x \in M$.

Zum Glück müssen wir nun nicht die ganze Theorie von vorne beginnen, sondern wir können alles zurückspielen auf die Resultate für $\mathbb{R}$, indem wir die einzelnen Koordinaten der Punkte $x \in \mathbb{R}^n$ betrachten. Wesentlich dafür sind die folgenden Ungleichungen:

Lemma 5.2 *Für $x \in \mathbb{R}^n$ gilt*

$$\max_{i=1,\ldots,n} |x_i| \leq |x| \leq \sqrt{n} \max_{i=1,\ldots,n} |x_i|.$$

Beweis: Die Ungleichungen folgen mit der Monotonie der Wurzel aus

$$\big(\max_{i=1,\ldots,n} |x_i|\big)^2 \leq x_1^2 + \ldots + x_n^2 = |x|^2 \leq n \big(\max_{i=1,\ldots,n} |x_i|\big)^2.$$

Überlegen Sie, was die Ungleichungen geometrisch bedeuten. $\square$

Satz 5.3 (Eigenschaften bzgl. Norm $\simeq$ Eigenschaften bzgl. der Koordinaten) *Eine Folge $x_k = \big((x_k)_1, \ldots, (x_k)_n\big) \in \mathbb{R}^n$ ist genau dann bezüglich der Euklidischen Norm konvergent (beschränkt, Cauchyfolge), wenn für alle $i = 1, \ldots, n$ die Koordinatenfolgen $\big((x_k)_i\big)_{k \in \mathbb{N}}$ konvergent (beschränkt, Cauchyfolgen) sind.*

Beweis: Für die Beschränktheit folgt die Aussage sofort aus Lemma 5.2. Das Argument für die Konvergenz geht wie folgt:

$$\begin{aligned} x_k \to a \quad &\Leftrightarrow \quad |x_k - a| \to 0 \\ &\Leftrightarrow \quad \max_{i=1,\dots,n} \big|(x_k)_i - a_i\big| \to 0 \quad \text{(wegen Lemma (5.2))} \\ &\Leftrightarrow \quad |(x_k)_i - a_i| \to 0 \quad \text{für alle } i = 1, \dots, n \\ &\Leftrightarrow \quad (x_k)_i \to a_i \quad \text{für alle } i = 1, \dots, n. \end{aligned}$$

Sind schließlich die einzelnen $\big((x_k)_i\big)_{k\in\mathbb{N}}$ Cauchyfolgen, so gibt es zu $\varepsilon > 0$ Zahlen $N_i \in \mathbb{R}$ mit $|(x_k)_i - (x_l)_i| < \varepsilon/\sqrt{n}$ für alle $k, l > N_i$. Es folgt für $k, l > \max_{i=1,\dots,n} N_i$ mit Lemma 5.2

$$|x_k - x_l| \le \sqrt{n} \max_{i=1,\dots,n} |(x_k)_i - (x_l)_i| < \sqrt{n}\,\varepsilon/\sqrt{n} = \varepsilon.$$

Die umgekehrte Richtung ist den Lesern überlassen. □

Folgerung 5.1 *Mit der Euklidischen Norm ist $\mathbb{R}^n$ vollständig: zu jeder Cauchyfolge $(x_k)_{k\in\mathbb{N}}$ gibt es ein $a \in \mathbb{R}^n$, so dass $x_k \to a$ mit $k \to \infty$.*

Beweis: Nach Satz 5.3 sind die Folgen $\big((x_k)_i\big)_{k\in\mathbb{N}}$ Cauchyfolgen in $\mathbb{R}$. Nach dem Vollständigkeitsaxiom gibt es $a_i \in \mathbb{R}$ mit $\lim_{k\to\infty} (x_k)_i = a_i$ für alle $i \in \{1, \dots, n\}$. Wieder nach Satz 5.3 folgt $\lim_{k\to\infty} x_k = a$. □

Beispiel 5.1 Um die Problematik der nächsten Folgerung zu verdeutlichen, betrachten wir hier die Folge $x_k = \big((x_k)_1, (x_k)_2\big)$ von Punkten in $\mathbb{R}^2$ mit

$$(x_k)_1 = \begin{cases} 1 \text{ für } k \text{ gerade,} \\ -1 \text{ für } k \text{ ungerade,} \end{cases} \qquad (x_k)_2 = \begin{cases} 0 \text{ für } k = 0 \bmod 3, \\ 1 \text{ für } k = 1 \bmod 3, \\ 2 \text{ für } k = 2 \bmod 3. \end{cases}$$

Wir wollen eine konvergente Teilfolge finden. Die Koordinatenfolge $\big((x_k)_1\big)_{k\in\mathbb{N}}$ hat die konstante Teilfolge $(x_{2j})_1 = 1$ für $j \in \mathbb{N}$, und die Koordinatenfolge $\big((x_k)_2\big)_{k\in\mathbb{N}}$ hat die konstante Teilfolge $(x_{3j})_2 = 0$ für $j \in \mathbb{N}$. Damit haben wir aber noch *keine* konvergente Teilfolge für die Punkte $x_k \in \mathbb{R}^2$ gefunden, denn weder $(x_{2j})_2$ noch $(x_{3j})_1$ sind konvergent. Stattdessen wählen wir aus der Folge $(x_{2j})_{j\in\mathbb{N}}$ die *weitere* Teilfolge $(x_{2\cdot 3i})_{i\in\mathbb{N}}$, also $j = 3i$, aus. Diese konvergiert in $\mathbb{R}^2$, denn es ist sogar $x_{6i} = (1, 0)$ für alle $i \in \mathbb{N}$.

Folgerung 5.2 (Bolzano-Weierstraß im $\mathbb{R}^n$) *Jede beschränkte Folge $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.*

Beweis: Wir konstruieren sukzessive Teilfolgen, so dass schliesslich alle n Koordinatenfolgen konvergieren. Aus Satz 5.3 folgt dann die Behauptung. Angenommen, für ein $j \in \{1, \dots, n\}$ sei schon eine Teilfolge $k_1 < k_2 < \dots$ gefunden, so dass

$$\lim_{l\to\infty} (x_{k_l})_i = a_i \quad \text{für alle } i = 1, \dots, j-1.$$

Der Fall $j = 1$ ist dabei der Induktionsanfang. Da die Folge $(x_{k_l})_j$, $l \in \mathbb{N}$, beschränkt ist nach Satz 5.3, gibt es nach dem Satz von Bolzano-Weierstraß in $\mathbb{R}$, Satz 2.8, eine Teilfolge $l_1 < l_2 < \dots$ und ein $a_j \in \mathbb{R}$, so dass zusätzlich

$$\lim_{m\to\infty} (x_{k_{l_m}})_j = a_j.$$

Nach dem n-ten Schritt ist eine Teilfolge bestimmt, für die alle Koordinatenfolgen konvergieren. $\square$

Definition 5.6 (Häufungspunkt von Mengen) *Ein Punkt $a \in \mathbb{R}^n$ heißt Häufungspunkt der Menge $M \subset \mathbb{R}^n$, falls für jedes $\varepsilon > 0$ die Menge $B_\varepsilon(a) \cap M$ unendlich viele Elemente hat.*

Ein Häufungspunkt von M ist nicht notwendig Element von M, zum Beispiel ist $0 \in \mathbb{R}$ Häufungspunkt von $M = \{1/n : n \in \mathbb{N}\}$.

Lemma 5.3 *Ein Punkt $a \in \mathbb{R}^n$ ist genau dann Häufungspunkt der Menge M, wenn es eine Folge von Punkten $x_k \in M\backslash\{a\}$ gibt mit $\lim_{k\to\infty} x_k = a$.*

Beweis: Sei a Häufungspunkt gemäß Definition 5.6. Zu $\varepsilon_k = 1/k$ gibt es dann ein $x_k \in B_{1/k}(a) \cap M\backslash\{a\}$. Die Folge $(x_k)_{k\in\mathbb{N}}$ ist wie verlangt.

Sei umgekehrt eine Folge $x_k \in M\backslash\{a\}$ gegeben mit $\lim_{k\to\infty} x_k = a$. Angenommen, für ein $\varepsilon > 0$ ist die Menge $B_\varepsilon(a) \cap M$ endlich, also $B_\varepsilon(a) \cap M\backslash\{a\} = \{a_1, \ldots, a_m\}$ für ein $m \in \mathbb{N}_0$. Dann ist $\varrho := \min_{1\le i\le m} |a_i - a| > 0$. Wegen $\lim_{k\to\infty} |x_k - a| = 0$ folgt für k hinreichend groß

$$x_k \in B_\varrho(a) \cap M\backslash\{a\} = \emptyset,$$

ein Widerspruch. $\square$

Beispiel 5.2 Die Menge $M = \{1/n : n \in \mathbb{N}\}$ hat $0 \in \mathbb{R}$ als einzigen Häufungspunkt. Dagegen ist die Menge der Häufungspunkte von $\mathbb{Q}$ gleich $\mathbb{R}$, denn in jeder Umgebung einer reellen Zahl gibt es eine rationale Zahl, siehe Satz 2.8. Es gibt eine kleine Differenz zwischen dem Begriff des Häufungspunkts für Folgen und für Mengen: die konstante Folge $x_k = a$, $k \in \mathbb{N}$, hat den Häufungspunkt a, dagegen hat die einelementige Menge $\{a\} \subset \mathbb{R}^n$ keinen Häufungspunkt. Überlegen Sie, dass für eine beliebige Folge $(x_k)_{k\in\mathbb{N}}$ im $\mathbb{R}^n$ die Häufungspunkte des Wertebereichs $M = \{x_k : k \in \mathbb{N}\}$ eine Teilmenge der Häufungspunkte der Folge bilden, die aber eine echte Teilmenge sein kann, wie das Beispiel der konstanten Folge zeigt.

Definition 5.7 *Eine Menge*

(i) *$U \subset \mathbb{R}^n$ heißt offen, falls es zu jedem $a \in U$ ein $\varepsilon > 0$ gibt mit $B_\varepsilon(a) \subset U$,*

(ii) *$A \subset \mathbb{R}^n$ heißt abgeschlossen, wenn folgende Implikation stets gilt:*

$$a = \lim_{n\to\infty} x_n \text{ mit } x_n \in A \quad \Rightarrow \quad a \in A.$$

Nach Lemma 5.3 ist eine Menge A genau dann abgeschlossen, wenn jeder Häufungspunkt von A schon Element von A ist.

Beispiel 5.3 Die Kugel $B_\varrho(a) = \{x \in \mathbb{R}^n : |x - a| < \varrho\}$ ist offen. Denn sei $x \in B_\varrho(a)$. Für $\varepsilon := \varrho - |x - a| > 0$ und $y \in B_\varepsilon(x)$ folgt aus der Dreiecksungleichung

$$|y - a| \le |y - x| + |x - a| < \varepsilon + |x - a| = \varrho.$$

Also gilt $B_\varepsilon(x) \subset B_\varrho(a)$, das heißt $B_\varrho(a)$ ist offen. Die Kugel $K_\varrho(a) = \{x \in \mathbb{R}^n : |x-a| \leq \varrho\}$ ist dagegen abgeschlossen: ist $x_k \in K_\varrho(a)$ und $x_k \to x \in \mathbb{R}^n$, so folgt

$$|x-a|^2 = \sum_{i=1}^n (x_i - a_i)^2 = \lim_{k\to\infty} \sum_{i=1}^n \big((x_k)_i - a_i\big)^2 \leq \varrho^2,$$

das heißt $x \in K_\varrho(a)$.

Satz 5.4 *Eine Menge $M \subset \mathbb{R}^n$ ist genau dann offen, wenn $\mathbb{R}^n \backslash M$ abgeschlossen ist.*

Beweis: Sei M offen und $x_k \in \mathbb{R}^n \backslash M$ mit $\lim_{k\to\infty} x_k = a \in \mathbb{R}^n$. Wäre $a \in M$, so folgt $B_\varepsilon(a) \subset M$ für ein $\varepsilon > 0$, und somit $x_k \in M$ für hinreichend große k, ein Widerspruch. Also ist $a \in \mathbb{R}^n \backslash M$, das heißt $\mathbb{R}^n \backslash M$ ist abgeschlossen.

Sei nun $\mathbb{R}^n \backslash M$ abgeschlossen. Wäre M nicht offen, so gibt es ein $a \in M$ mit $B_\varepsilon(a) \backslash M \neq \emptyset$ für alle $\varepsilon > 0$. Finde induktiv $x_k \in B_{1/k}(a)$ mit $x_k \in \mathbb{R}^n \backslash M$. Es folgt $a = \lim_{k\to\infty} x_k \in \mathbb{R}^n \backslash M$, ein Widerspruch. □

Satz 5.5 *Die offenen Teilmengen des $\mathbb{R}^n$ bilden eine Topologie, das heißt es gelten folgende Aussagen:*

(1) *$\emptyset$ und $\mathbb{R}^n$ sind offen.*

(2) *Jede Vereinigung $\bigcup_{\lambda\in\Lambda} U_\lambda$ von offenen Mengen $U_\lambda \subset \mathbb{R}^n$ ist offen.*

(3) *Der Durchschnitt $\bigcap_{i=1}^k U_i$ von endlich vielen offenen Mengen $U_i \subset \mathbb{R}^n$ ist offen.*

Beweis: Aussage (1) ist evident. Ist $x \in \bigcup_{\lambda\in\Lambda} U_\lambda$, so gilt $x \in U_\mu$ für ein $\mu \in \Lambda$. Da U_μ offen, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset U_\mu \subset \bigcup_{\lambda\in\Lambda} U_\lambda$, was zu zeigen war.

Sei nun $x \in \bigcap_{i=1}^k U_i$ wie in (3). Da U_i offen, gibt es ein $\varepsilon_i > 0$ mit $B_{\varepsilon_i}(x) \subset U_i$ für $i = 1, \ldots, k$. Es folgt $\varepsilon := \min_{i=1,\ldots,k} \varepsilon_i > 0$ und $B_\varepsilon(x) \subset B_{\varepsilon_i}(x) \subset U_i$ für alle $i = 1, \ldots, k$, also $B_\varepsilon(x) \subset \bigcap_{i=1}^k U_i$. □

Wir betonen, dass im Gegensatz zur Vereinigung der Durchschnitt von unendlich vielen offenen Mengen in der Regel nicht offen ist, zum Beispiel gilt

$$\{a\} = \bigcap_{n\in\mathbb{N}} B_{1/n}(a).$$

Eine Menge $M \subset \mathbb{R}^n$ muss weder offen noch abgeschlossen sein; dies trifft zum Beispiel auf ein halboffenes Intervall $[a, b)$ zu. Die Mengen $\mathbb{R}$ und $\emptyset$ sind die einzigen Teilmengen von $\mathbb{R}$, die sowohl offen als auch abgeschlossen sind. Dies soll in den Übungen gezeigt werden. Durch Übergang zu den Komplementen sieht man, dass beleibige Durchschnitte und endliche Vereinigungen von abgeschlossenen Mengen abermals abgeschlossen sind.

Kapitel 3

Stetige Funktionen

1 Grenzwerte und Stetigkeit

Eine Funktion auf einer Menge $D \subset \mathbb{R}^n$ mit Werten in $\mathbb{R}^m$ ist bekanntlich eine Abbildung

$$f : D \to \mathbb{R}^m,\, x \mapsto f(x).$$

Im Fall $m = 1$, also $f : D \to \mathbb{R}$, heißt die Funktion reellwertig. D heißt Definitionsbereich und $f(D)$ heißt Bild von f. Der Graph von f ist die Menge

$$\mathrm{Graph}\,(f) = \{(x, f(x)) : x \in D\} \subset D \times \mathbb{R}^m \subset \mathbb{R}^n \times \mathbb{R}^m.$$

Beispiel 1.1 Die relativistische Masse als Funktion der Geschwindigkeit ist gegeben durch

$$m : [0, c) \to \mathbb{R},\, m(v) = \frac{m_0}{\sqrt{1 - (\frac{v}{c})^2}};$$

dabei steht m_0 für die Ruhemasse und c für die Lichtgeschwindigkeit. Der Definitionsbereich der Relativitätstheorie ist also das Geschwindigkeitsintervall $[0, c)$.

Beispiel 1.2 Die Signumfunktion ist definiert durch

$$\mathrm{sign} : \mathbb{R} \to \mathbb{R},\, \mathrm{sign}(x) = \begin{cases} 1 & \text{für } x > 0, \\ 0 & \text{für } x = 0, \\ -1 & \text{für } x < 0. \end{cases}$$

Beispiel 1.3 Die Gaußklammerfunktion oder Größte-Ganze-Funktion ist gegeben durch

$$f : \mathbb{R} \to \mathbb{R},\, f(x) = [x] = \max\,\{n \in \mathbb{N} : n \leq x\}.$$

Beispiel 1.4 Ist $I \subset \mathbb{R}$ ein Intervall, so bezeichnet man Funktionen $c : I \to \mathbb{R}^m$ auch als (parametrisierte) Kurven (*curva*). Oft wird die unabhängige Variable mit $t \in I$ (statt $x \in I$) bezeichnet. Diese Notation folgt Newton, der sich für Bahnkurven von Massenpunkten in Abhängigkeit von der Zeit (*tempus*) interessiert hat und $c(t) = (x(t), y(t), z(t))$ schreibt. Ein explizites Beispiel ist der horizontale Wurf aus der Höhe h mit Geschwindigkeit v:

$$c : \mathbb{R} \to \mathbb{R}^3,\, c(t) = (vt, 0, h - gt^2/2) \quad (g = \text{Erdbeschleunigung}).$$

Beispiel 1.5 Die charakteristische Funktion (Indikatorfunktion) einer Menge $E \subset \mathbb{R}^n$ ist

$$\chi_E : \mathbb{R}^n \to \mathbb{R},\, \chi_E(x) = \begin{cases} 1 & \text{für } x \in E, \\ 0 & \text{für } x \notin E. \end{cases}$$

Beispiel 1.6 Die Euklidische Norm auf $\mathbb{R}^n$ ist die Funktion

$$|\cdot| : \mathbb{R}^n \to \mathbb{R},\, x \mapsto |x| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

In $\mathbb{R}$ sowie in $\mathbb{C}$ sprechen wir auch von der Betragsfunktion.

Beispiel 1.7 Die Abstandsfunktion einer Menge $E \subset \mathbb{R}^n$ ist

$$\operatorname{dist}_E : \mathbb{R}^n \to \mathbb{R},\, \operatorname{dist}_E(x) = \inf\{|y - x| : y \in E\}.$$

Definition 1.1 (Stetigkeit) *Die Funktion $f : D \to \mathbb{R}^m$ heißt stetig im Punkt $x_0 \in D$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass gilt:*

$$|f(x) - f(x_0)| < \varepsilon \quad \textit{für alle } x \in D \textit{ mit } |x - x_0| < \delta. \tag{1.1}$$

f heißt stetig, falls f in allen $x_0 \in D$ stetig ist.

Mithilfe von Umgebungen lässt sich die Stetigkeit in x_0 auch so fassen:

Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $f(D \cap B_\delta(x_0)) \subset B_\varepsilon(f(x_0))$.

Wir wollen einige Beispiele betrachten.

Beispiel 1.8 Konstante Funktionen sind stetig: sei $f(x) = c$ für alle $x \in D$ und ein $c \in \mathbb{R}^m$. Dann folgt

$$|f(x) - f(x_0)| = 0 \quad \text{für alle } x \in D,$$

also können wir zu gegebenem $\varepsilon > 0$ jedes $\delta > 0$ wählen.

Beispiel 1.9 Affinlineare Funktionen $f : \mathbb{R} \to \mathbb{R}$ sind stetig: sei $f(x) = ax+b$ mit $a, b \in \mathbb{R}$, $a \neq 0$. Dann gilt

$$|f(x) - f(x_0)| = |(ax + b) - (ax_0 + b)| = |a|\,|x - x_0|,$$

das heißt wir können $\delta = \varepsilon/|a|$ wählen.

Beispiel 1.10 Die Euklidische Norm $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = |x|$, ist ebenfalls stetig auf ganz $\mathbb{R}^n$, denn aus der Dreiecksungleichung folgt

$$|f(x) - f(x_0)| = ||x| - |x_0|| \leq |x - x_0|.$$

Wir können also $\delta = \varepsilon$ nehmen.

Beispiel 1.11 Die charakteristische Funktion von $\mathbb{Q}$ (oder Dirichlet-Funktion)

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q}, \\ 0 & \text{sonst.} \end{cases}$$

ist nirgends stetig. Denn wäre $\chi_{\mathbb{Q}}$ stetig in $x_0 \in \mathbb{R}$, so gibt es zu $\varepsilon = 1$ ein $\delta > 0$ mit

$$|\chi_{\mathbb{Q}}(x) - \chi_{\mathbb{Q}}(x_0)| < \varepsilon \quad \text{für alle } x \in \mathbb{R} \text{ mit } |x - x_0| < \delta.$$

Ist $x_0 \in \mathbb{R}\backslash\mathbb{Q}$, so können wir $x \in \mathbb{Q}$ wählen, da $\mathbb{Q}$ dicht in $\mathbb{R}$ ist nach Satz 2.8, und erhalten $1 = |\chi_{\mathbb{Q}}(x) - \chi_{\mathbb{Q}}(x_0)| < \varepsilon = 1$, ein Widerspruch. Ist $x_0 \in \mathbb{Q}$, so können wir analog argumentieren, da $\mathbb{R}\backslash\mathbb{Q}$ ebenfalls dicht ist nach Aufgabe 1, Serie 6.

Ein sehr nützliches, hinreichendes Kriterium für die Stetigkeit ist das folgende.

Definition 1.2 *$f : D \to \mathbb{R}^m$ heißt Lipschitzstetig mit Konstante $L \in [0, \infty)$, falls gilt:*

$$|f(x) - f(y)| \leq L|x - y| \quad \textit{für alle } x, y \in D.$$

Die Euklidische Norm ist Lipschitzstetig mit Konstante $L = 1$, vgl. Beispiel 1.10.

Lemma 1.1 *Jede Lipschitzstetige Funktion $f : D \to \mathbb{R}^m$ ist stetig.*

Beweis: Sei f Lipschitzstetig mit Konstante $L > 0$. Zu $x_0 \in D$ und gegebenem $\varepsilon > 0$ wählen wir $\delta = \varepsilon/L > 0$, und erhalten für $|x - x_0| < \delta$

$$|f(x) - f(x_0)| \leq L|x - x_0| < L\delta = \varepsilon.$$

□

Beispiel 1.12 Die Abstandsfunktion $\mathrm{dist}_E : \mathbb{R}^n \to \mathbb{R}$ einer Menge $E \subset \mathbb{R}^n$ ist Lipschitzstetig mit Konstante Eins. Denn es gilt nach der Dreiecksungleichung

$$|y - x_2| \leq |y - x_1| + |x_1 - x_2| \quad \text{für alle } x_{1,2} \in \mathbb{R}^n,\ y \in E.$$

Bilden wir auf beiden Seiten das Infimum über alle $y \in E$, so folgt

$$\mathrm{dist}_E(x_2) \leq \mathrm{dist}_E(x_1) + |x_1 - x_2|.$$

Durch Vertauschen von x_1 mit x_2 ergibt sich

$$|\mathrm{dist}_E(x_1) - \mathrm{dist}_E(x_2)| \leq |x_1 - x_2|.$$

Beispiel 1.13 Lineare Abbildungen $A : \mathbb{R}^n \to \mathbb{R}^m$ sind Lipschitzstetig. Um dies zu zeigen, verwenden wir die Matrixdarstellung bezüglich der Standardbasen. Es gilt für $x \in \mathbb{R}^n$

$$A(x)_i = \sum_{j=1}^{n} a_{ij} x_j \quad \text{wobei } A(x) = \sum_{i=1}^{m} A(x)_i e_i.$$

Die Euklidische Norm der Matrix $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ ist

$$|A| = \Big(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \Big)^{1/2}.$$

Mit der Ungleichung von Cauchy-Schwarz folgt

$$|A(x)|^2 = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} x_j \Big)^2 \leq \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij}^2 \Big) \cdot \Big(\sum_{j=1}^{n} x_j^2 \Big) = |A|^2 |x|^2,$$

das heißt es gilt die Abschätzung

$$|A(x)| \leq |A|\,|x| \quad \text{für alle } x \in \mathbb{R}^n.$$

Insbesondere folgt für $x, y \in \mathbb{R}^n$

$$|A(x) - A(y)| = |A(x-y)| \leq |A|\,|x-y|,$$

d. h. A ist Lipschitzstetig mit Konstante $|A|$.

Satz 1.1 (Folgenkriterium der Stetigkeit) *Für $f : D \to \mathbb{R}^m$ und $x_0 \in D$ sind folgende Aussagen äquivalent:*

(1) *f ist stetig in x_0.*

(2) *Für jede Folge $(x_k)_{k \in \mathbb{N}}$ mit $x_k \in D$ und $\lim_{k \to \infty} x_k = x_0$ gilt: $\lim_{k \to \infty} f(x_k) = f(x_0)$.*

Beweis: Für die Implikation (1) $\Rightarrow$ (2) sei $x_k \in D$ mit $x_k \to x_0$ für $k \to \infty$. Wähle zu $\varepsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Es gibt ein $N \in \mathbb{R}$ mit $|x_k - x_0| < \delta$ für $k > N$, also folgt $|f(x_k) - f(x_0)| < \varepsilon$ für $k > N$.

Wir überlegen nun was es bedeutet, wenn f nicht stetig in x_0 ist. Es gibt dann ein $\varepsilon > 0$, so dass (1.1) für kein $\delta > 0$ erfüllt ist. Wählen wir $\delta_k = 1/k$ für $k = 1, 2, \ldots$, so gibt es $x_k \in D$ mit $|x_k - x_0| < 1/k$, aber $|f(x_k) - f(x_0)| \geq \varepsilon$. Dies zeigt indirekt die Implikation (2) $\Rightarrow$ (1), womit der Satz bewiesen ist. $\square$

Lemma 1.2 *Eine Funktion $f : D \to \mathbb{R}^m$ ist genau dann stetig in $x_0 \in D$, wenn jede Koordinatenfunktion $f_i : D \to \mathbb{R}$ in x_0 stetig ist.*

Beweis: Mit dem vorangegangenen Satz folgt dies aus der Äquivalenz von Konvergenz und Konvergenz der Koordinatenfolgen im $\mathbb{R}^m$, siehe Satz 5.3. $\square$

Lemma 1.3 *Sei $g : D \to \mathbb{R}$ stetig in x_0 mit $g(x_0) \neq 0$. Dann gibt es ein $\delta > 0$ mit*

$$|g(x)| \geq |g(x_0)|/2 > 0 \quad \textit{für alle } x \in D \cap U_\delta(x_0).$$

Beweis: Zu $\varepsilon = |g(x_0)|/2 > 0$ gibt es aufgrund der Stetigkeit ein $\delta > 0$ mit $|g(x) - g(x_0)| < \varepsilon$ für alle $x \in D \cap U_\delta(x_0)$, also folgt mit der Dreiecksungleichung für diese x

$$|g(x)| = |g(x_0) - (g(x_0) - g(x))| \geq |g(x_0)| - |g(x) - g(x_0)| \geq 2\varepsilon - \varepsilon = \varepsilon.$$

$\square$

Satz 1.2 (Stetigkeitsregeln) *Seien $f, g : D \to \mathbb{R}$ stetig in $x_0 \in D$. Dann gilt:*

(1) *Für beliebige $\alpha, \beta \in \mathbb{R}$ ist die Funktion $\alpha f + \beta g$ stetig in x_0.*

(2) *Die Funktion fg ist stetig in x_0.*

(3) *Ist $g(x_0) \neq 0$, so ist die Funktion $f/g : D \cap U_\delta(x_0) \to \mathbb{R}$ für $\delta > 0$ hinreichend klein definiert und stetig in x_0.*

Beweis: Die Funktionen sind grundsätzlich punktweise erklärt, das heißt für alle $x \in D$ gilt

$$(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x),\ (fg)(x) = f(x)g(x) \text{ und } (f/g)(x) = f(x)/g(x).$$

Wir führen die Aussagen auf die entsprechenden Konvergenzregeln für Folgen zurück: sei $(x_n)_{n\in\mathbb{N}}$ eine beliebige Folge mit $x_n \in D$ und $\lim_{n\to\infty} x_n = x_0$. Nach Satz 1.1 gilt $f(x_n) \to f(x_0)$ und $g(x_n) \to g(x_0)$ mit $n \to \infty$. Aus Kapitel I, Satz 1.3 folgt

$$\alpha f(x_n) + \beta g(x_n) \to \alpha f(x_0) + \beta g(x_0) \quad \text{sowie} \quad f(x_n)g(x_n) \to f(x_0)g(x_0).$$

Mit Satz 1.1 ergeben sich die Behauptungen (1) und (2). In der Situation von (3) gibt es nach Lemma 1.3 ein $\delta > 0$ mit $g(x) \neq 0$ auf $D \cap U_\delta(x_0)$, so dass f/g dort definiert ist. Weiter folgt $f(x_n)/g(x_n) \to f(x_0)/g(x_0)$, und mit Satz 1.1 ist auch f/g stetig in x_0. □

Folgerung 1.1 *Die Menge $C^0(D)$ der stetigen Funktionen auf $D \subset \mathbb{R}^n$ ist ein $\mathbb{R}$-Vektorraum. Allgemeiner ist die Menge $C^0(D, \mathbb{R}^k)$ der stetigen Abbildungen $f : D \to \mathbb{R}^k$ ein $\mathbb{R}$-Vektorraum.*

Beweis: Das ist offensichtlich nach Satz 1.2(1) und Lemma 1.2. □

Satz 1.3 (Verkettung stetiger Funktionen) *Seien $f : D \to \mathbb{R}^m$, $g : E \to \mathbb{R}^k$ mit $f(D) \subset E \subset \mathbb{R}^m$. Ist f stetig in x_0 und g stetig in $y_0 = f(x_0)$, so ist $g \circ f : D \to \mathbb{R}^k$ stetig in x_0.*

Beweis: Wir verwenden wieder Satz 1.1. Ist $x_n \in D$ eine beliebige Folge mit $\lim_{n\to\infty} x_n = x_0$, so folgt $f(x_n) \to f(x_0)$ aus der Stetigkeit von f in x_0, und weiter $g(f(x_n)) \to g(f(x_0))$ wegen der Stetigkeit von g in $y_0 = f(x_0)$. Nach Satz 1.1 ist damit die Stetigkeit von $g \circ f$ in x_0 schon gezeigt. □

Beispiel 1.14 Seien $f, g : D \to \mathbb{R}$ stetig. Dann ist auch die Funktion $|f| : D \to \mathbb{R}$ stetig, und ebenso die Funktionen $\max(f, g) = (f+g+|f-g|)/2$ und $\min(f, g) = (f+g-|f-g|)/2$.

Schließlich wollen wir in diesem Abschnitt noch den Grenzwertbegriff für Funktionen diskutieren; allerdings fassen wir uns dabei kurz, weil die Angelegenheit für Folgen ja ausführlich behandelt wurde.

Definition 1.3 (Grenzwert für Funktionen) *Sei $D \subset \mathbb{R}^n$ und x_0 ein Häufungspunkt von D. Die Funktion $f : D \to \mathbb{R}^m$ konvergiert für $x \to x_0$ gegen $a \in \mathbb{R}^m$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass gilt:*

$$|f(x) - a| < \varepsilon \quad \textit{für alle } x \in D \textit{ mit } 0 < |x - x_0| < \delta.$$

Notation: $\lim_{x\to x_0} f(x) = a$ *oder* $f(x) \to a$ *für* $x \to x_0$.

Für die Existenz und den Wert von $\lim_{x\to x_0} f(x)$ ist es egal, ob die Funktion in x_0 definiert ist bzw. welchen Funktionswert sie dort hat. Die Beziehung zwischen Grenzwert und Stetigkeit ist wie folgt:

Lemma 1.4 (Stetigkeit und Grenzwert) *Sei $x_0 \in D$ ein Häufungspunkt von D. Für die Funktion $f : D \to \mathbb{R}^m$ sind folgende Aussagen äquivalent:*

(1) $\lim_{x\to x_0} f(x) = f(x_0)$.

(2) *f ist stetig in x_0.*

Beweis: Das folgt direkt aus den gegebenen Definitionen. □

Für reellwertige Funktionen können wir den Konvergenzbegriff ausdehnen, indem wir (uneigentliche) Konvergenz gegen $\pm\infty$ zulassen.

Definition 1.4 (Uneigentlicher Grenzwert) *Sei $x_0 \in \mathbb{R}$ ein Häufungspunkt von D, und $f : D \to \mathbb{R}$. Dann gilt $\lim_{x\to x_0} f(x) = +\infty$, wenn es zu jedem $K > 0$ ein $\delta > 0$ gibt mit*

$$f(x) > K \text{ für alle } x \in D \text{ mit } 0 < |x - x_0| < \delta.$$

Entsprechend wird $\lim_{x\to x_0} f(x) = -\infty$ erklärt.

Im Fall $D \subset \mathbb{R}$ besteht die Möglichkeit, Grenzwerte von $f(x)$ für $x \to \pm\infty$ zu betrachten.

Definition 1.5 (Grenzwert bei $\pm\infty$) *Sei $f : D \to \mathbb{R}^m$ mit $D \subset \mathbb{R}$ nicht nach oben beschränkt. Dann gilt $\lim_{x\to+\infty} f(x) = a \in \mathbb{R}^m$, wenn es zu $\varepsilon > 0$ ein $K \in \mathbb{R}$ gibt mit*

$$|f(x) - a| < \varepsilon \quad \text{für alle } x \in D \text{ mit } x > K.$$

Analog wird der Grenzwert für $x \to -\infty$ erklärt.

Sowohl bei der Stetigkeit als auch beim Grenzwert spielt der zugrundeliegende Definitionsbereich eine Rolle. Wir schreiben $\lim_{x\to x_0, x\in D} f(x)$, wenn wir den gewählten Definitionsbereich hervorheben möchten. In $\mathbb{R}$ werden oft einseitige Grenzwerte gebraucht:

$$\lim_{x\searrow x_0} f(x) = \lim_{x\to x_0, x>x_0} f(x) \quad \text{und} \quad \lim_{x\nearrow x_0} f(x) = \lim_{x\to x_0, x<x_0} f(x).$$

Zum Beispiel ist $\lim_{x\searrow 0} \operatorname{sign}(x) = +1$ und $\lim_{x\nearrow 0} \operatorname{sign}(x) = -1$, während der Grenzwert $\lim_{x\to 0} \operatorname{sign}(x)$ nicht existiert. Wir formulieren jetzt die wichtigsten Regeln für Grenzwerte von Funktionen. Diese sind für Folgen bzw. aus der Diskussion der Stetigkeit schon bekannt, deshalb verzichten wir auf die Beweise und formulieren die Aussagen der Einfachheit halber nur für $n = m = 1$, also reellwertige Funktionen einer Variablen.

Satz 1.4 (Rechenregeln für Grenzwerte) *Es gelten folgende Aussagen:*

(1) *Seien $a, b \in \mathbb{R}$ und $x_0 \in \mathbb{R} \cup \{\pm\infty\}$. Aus $f(x) \to a$, $g(x) \to b$ für $x \to x_0$ folgt*

$$\begin{aligned} \alpha f(x) + \beta g(x) &\to \alpha a + \beta b \quad (\alpha, \beta \in \mathbb{R}), \\ f(x)g(x) &\to ab, \\ f(x)/g(x) &\to a/b, \quad \text{falls } b \neq 0. \end{aligned}$$

(2) *Seien* $f : D \to \mathbb{R}$, $g : E \to \mathbb{R}$ *mit* $f(D) \subset E$. *Gilt* $f(x) \to y_0$ *mit* $x \to x_0$ *und ist* g *stetig in* y_0, *so folgt* $(g \circ f)(x) \to g(y_0)$ *mit* $x \to x_0$.

(3) *Sei* $f(x) \to a$, $g(x) \to b$ *für* $x \to x_0$. *Ist* $f(x) \leq g(x)$ *für* $0 < |x - x_0| < \delta$, *so folgt* $a \leq b$.

(4) *Ist* $f > 0$ *auf* D, *so ist* $\lim_{x \to x_0} f(x) = 0$ *äquivalent zu* $\lim\limits_{x \to x_0} \dfrac{1}{f(x)} = +\infty$.

Beispiel 1.15 (Rationale Funktionen) Seien p, q reelle Polynome vom Grad m bzw. n, also $p(x) = a_0 + a_1 x + \ldots + a_m x^m$ bzw. $q(x) = b_0 + b_1 x + \ldots + b_n x^n$ mit $a_i, b_j \in \mathbb{R}$ und $a_m, b_n \neq 0$. Setze $N = \{x \in \mathbb{R} : q(x) = 0\}$ und definiere

$$f : \mathbb{R}\backslash N \to \mathbb{R},\ f(x) = \frac{p(x)}{q(x)}.$$

Wann hat f eine stetige Fortsetzung bei $x_0 \in N$? Sei dazu x_0 eine ν-fache Nullstelle von q mit $\nu \geq 1$, und eine μ-fache Nullstelle von p, evtl. $\mu = 0$ falls $p(x_0) \neq 0$. Also gilt

$$p(x) = (x - x_0)^{\mu}\, \tilde{p}(x) \quad \text{bzw.} \quad q(x) = (x - x_0)^{\nu}\, \tilde{q}(x)$$

für Polynome $\tilde{p}, \tilde{q}$ mit $\tilde{p}(x_0), \tilde{q}(x_0) \neq 0$. Es folgt

$$f(x) = (x - x_0)^{\mu - \nu} \frac{\tilde{p}(x)}{\tilde{q}(x)} \quad \text{für alle } x \in \mathbb{R}\backslash N,$$

und wir erhalten aus den Rechenregeln für Grenzwerte $\lim_{x \to x_0} f(x) = 0$ im Fall $\mu > \nu$ und $\lim_{x \to x_0} f(x) = \tilde{p}(x_0)/\tilde{q}(x_0) \neq 0$ im Fall $\mu = \nu$. In diesen Fällen ist die Funktion stetig fortsetzbar nach Lemma 1.4, dagegen hat f im Fall $\mu < \nu$ in x_0 eine Polstelle. Sei zum Beispiel $\tilde{p}(x_0)/\tilde{q}(x_0) > 0$, dann ergibt sich für die einseitigen Grenzwerte bei x_0

	$x \searrow x_0$	$x \nearrow x_0$
$\nu - \mu$ gerade	$+\infty$	$+\infty$
$\nu - \mu$ ungerade	$+\infty$	$-\infty$

Das Verhalten von f für $x \to \pm\infty$ wurde bereits in Kapitel 1, Beispiel 1.7 untersucht.

2 Zwischenwertsatz und klassische Funktionen

In diesem Abschnitt haben wir es mit Funktionen zu tun, die auf einem Intervall definiert sind. Eine Menge $I \subset \mathbb{R}$ ist genau dann ein Intervall, wenn $(a, b) \subset I$ mit $a = \inf I$ und $b = \sup I$. Hier sind also unendliche Intervallgrenzen zugelassen, aber das Intervall soll stets Teilmenge von $\mathbb{R}$ sein, das heißt unendliche Intervallgrenzen sind offen. Der Durchschnitt von zwei Intervallen $I_{1,2}$ mit den Grenzen $a_{1,2}$ und $b_{1,2}$ ist wieder ein Intervall mit Grenzen $a = \max(a_1, a_2)$ und $b = \min(b_1, b_2)$. Nach unserer Definition ist die leere Menge auch ein Intervall.

Satz 2.1 (Zwischenwertsatz) *Sei* $f : [a, b] \to \mathbb{R}$ *stetig. Dann gibt es zu jedem* y_0 *zwischen* $f(a)$ *und* $f(b)$ *ein* $x_0 \in [a, b]$ *mit* $f(x_0) = y_0$.

Bemerkung. Die Gleichung $f(x) = y_0$ kann mehrere Lösungen in $[a, b]$ besitzen, das heißt x_0 ist im allgemeinen nicht eindeutig bestimmt.

Beweis: Sei oBdA $f(a) \leq y_0 \leq f(b)$. Dann ist die Menge $M = \{x \in [a, b] : f(x) \leq y_0\}$ nichtleer, da $a \in M$. Wir behaupten $f(x_0) = y_0$ für $x_0 = \sup M$. Für $n \in \mathbb{N}$ ist $x_0 - 1/n$ keine obere Schranke von M, also gibt es $x_n \in M$ mit $x_0 - 1/n < x_n \leq x_0$. Da f stetig, folgt $f(x_0) = \lim_{n\to\infty} f(x_n) \leq y_0$. Andererseits ist x_0 obere Schranke von M, also gilt $f(x) > y_0$ für $x_0 < x \leq b$. Ist $x_0 < b$, so folgt $f(x_0) = \lim_{x\searrow x_0} f(x) \geq y_0$. Im Fall $x_0 = b$ gilt $f(x_0) \geq y_0$ sowieso nach Voraussetzung. $\square$

Folgerung 2.1 *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ stetig. Dann ist $f(I)$ ein Intervall mit Endpunkten $\alpha = \inf_{x\in I} f(x)$ und $\beta = \sup_{x\in I} f(x)$.*

Beweis: Zu $y \in (\alpha, \beta)$ gibt es $x_1, x_2 \in I$ mit $f(x_1) < y < f(x_2)$. Dann gibt es nach Satz 2.1 ein $x \in [x_1, x_2]$ (bzw. $x \in [x_2, x_1]$) mit $f(x) = y$. Es folgt $(\alpha, \beta) \subset f(I)$. $\square$

Jede streng monotone Funktion $f : I \to \mathbb{R}$ ist injektiv und hat damit eine Umkehrfunktion $g : f(I) \to \mathbb{R}$. Es stellt sich die Frage nach der Stetigkeit von g. Zum nächsten Lemma vgl. das entsprechende Resultat für Folgen, Satz 2.5 in Kapitel 1.

Lemma 2.1 (Einseitige Grenzwerte monotoner Funktionen) *Sei $f : I = (a, b) \to \mathbb{R}$ monoton wachsend. Dann gilt*

$$\lim_{x\searrow a} f(x) = \inf\{f(x) : x \in I\} \quad \text{und} \quad \lim_{x\nearrow b} f(x) = \sup\{f(x) : x \in I\}.$$

Beweis: Mit $\alpha = \inf\{f(x) : x \in I\} \in [-\infty, \infty)$ gilt $f(x) \geq \alpha$ für alle $x \in I$. Andererseits gibt es zu jedem $\alpha' > \alpha$ ein $x' \in I$ mit $f(x') < \alpha'$, und somit $f(x) < \alpha'$ für alle $x < x'$. Die Behauptung für den rechtsseitigen Grenzwert folgt analog. $\square$

Satz 2.2 (Monotonie und Umkehrfunktion) *Sei I ein Intervall mit Endpunkten $a < b$, und $f : I \to \mathbb{R}$ sei streng monoton wachsend und stetig. Dann gilt:*

(1) *Die Umkehrfunktion $g : f(I) \to \mathbb{R}$ ist ebenfalls streng monoton wachsend und stetig.*

(2) *$I^* = f(I)$ ist ein Intervall mit Endpunkten $\alpha = \lim_{x\searrow a} f(x) < \lim_{x\nearrow b} f(x) = \beta$.*

(3) *$\lim_{y\searrow\alpha} g(y) = a$ und $\lim_{y\nearrow\beta} g(y) = b$.*

Beweis: Wäre g nicht streng monoton wachsend, so gibt es $y_1, y_2 \in f(I)$ mit $y_1 < y_2$, aber $g(y_2) \leq g(y_1)$. Aus der Monotonie von f folgt

$$y_2 = f\big(g(y_2)\big) \leq f\big(g(y_1)\big) = y_1,$$

im Widerspruch zur Annahme. Um die Stetigkeit von g im Punkt $y_0 = f(x_0)$ zu zeigen, sei $\varepsilon > 0$ gegeben. Nach Folgerung 2.1 ist dann $f(U_\varepsilon(x_0) \cap I)$ ein Intervall mit den Grenzen

$$\alpha' = \inf\{f(x) : x \in U_\varepsilon(x_0) \cap I\} \text{ und } \beta' = \sup\{f(x) : x \in U_\varepsilon(x_0) \cap I\}.$$

Im Fall $a < x_0 < b$ gilt $\alpha' < f(x_0) = y_0 < \beta'$, da f streng monoton wachsend, also gibt es ein $\delta > 0$ mit $U_\delta(y_0) \subset f(U_\varepsilon(x_0) \cap I)$. Es folgt

$$g(U_\delta(y_0)) \subset (g \circ f)(U_\varepsilon(x_0) \cap I) \subset U_\varepsilon(x_0),$$

was zu zeigen war. Für $x_0 = a$ ist $\alpha = f(x_0) = y_0 < \beta'$, also gilt für hinreichend kleines $\delta > 0$

$$U_\delta(y_0) \cap f(I) \subset [y_0, y_0 + \delta) \subset f(U_\varepsilon(x_0) \cap I),$$

und es folgt $g\big(U_\delta(y_0) \cap f(I)\big) \subset U_\varepsilon(x_0)$. Aussage (2) folgt direkt aus Folgerung 2.1 und Lemma 2.1, und (3) ergibt sich aus (1) und (2), angewandt auf die Funktion g statt f. □

Als Ergänzung zu Satz 2.2 bemerken wir noch

$$a \in I \Leftrightarrow \alpha \in I^* \quad \text{und} \quad b \in I \Leftrightarrow \beta \in I^*. \tag{2.1}$$

Denn für $a \in I$ ist $f(a) = \inf\{f(x) : x \in I\} = \alpha$, also $\alpha \in f(I)$. Sei umgekehrt $\alpha = f(x)$ für ein $x \in I$. Wäre $x > a$, so gibt es $x' \in (a, x)$ mit $f(x') < f(x) = \alpha$, Widerspruch. Also folgt $a = x \in I$.

Wir wollen nun den Satz anwenden, um die Umkehrfunktion der reellen Exponentialfunktion zu konstruieren. Als erstes müssen wir dafür die Stetigkeit zeigen; das tun wir natürlich gleich in $\mathbb{C}$.

Satz 2.3 *Die Exponentialfunktion* $\exp : \mathbb{C} \to \mathbb{C}$ *ist stetig.*

Beweis: Für die Stetigkeit in $z_0 = 0$ verwenden wir die Abschätzung aus Kapitel 1, Satz 4.7, und zwar ergibt der Fall $n = 1$ für alle $|z| \leq 1$

$$|\exp(z) - \exp(0)| = |\exp(z) - 1| \leq 2|z|.$$

Es folgt $|\exp(z) - \exp(0)| < \varepsilon$ für $|z| < \delta = \min(1, \varepsilon/2)$. Für die Stetigkeit in $z_0 \neq 0$ setzen wir die Funktionalgleichung ein: es gilt $\exp(z - z_0) \to 1$ mit $z \to z_0$, also

$$\exp(z) = \exp(z_0)\exp(z - z_0) \to \exp(z_0) \quad \text{mit } z \to z_0.$$

□

Satz 2.4 (Definition des Logarithmus) *Die Funktion* $\exp : (-\infty, \infty) \to (0, \infty)$ *ist streng monoton wachsend, stetig und bijektiv, und es gilt*

$$\lim_{x \to -\infty} \exp(x) = 0 \quad \textit{und} \quad \lim_{x \to \infty} \exp(x) = \infty. \tag{2.2}$$

Die Umkehrfunktion $\log : (0, \infty) \to (-\infty, \infty)$ *heißt (natürlicher) Logarithmus. Die Funktion ist ebenfalls streng monoton wachsend, stetig und bijektiv, und es gilt*

$$\lim_{y \searrow 0} \log(y) = -\infty \quad \textit{und} \quad \lim_{y \to \infty} \log(y) = \infty. \tag{2.3}$$

Weiter ist $\log(1) = 0$ *und* $\log(e) = 1$, *und* $\log$ *erfüllt die Funktionalgleichung*

$$\log(y_1 y_2) = \log(y_1) + \log(y_2) \quad \textit{für alle } y_1, y_2 > 0. \tag{2.4}$$

Beweis: Nach Definition der Exponentialfunktion gilt

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} > 1 + x \quad \text{für } x > 0.$$

Es folgt $\lim_{x\to\infty} \exp(x) = \infty$, sowie für $x > 0$ mit der Funktionalgleichung

$$\exp(-x) = \frac{1}{\exp(x)} \to 0 \quad \text{mit } x \to \infty.$$

Weiter liefert die Funktionalgleichung für $x_1 < x_2$

$$\frac{\exp(x_2)}{\exp(x_1)} = \exp(x_2 - x_1) > 1,$$

das heißt exp ist streng monoton wachsend und stetig nach Satz 2.3. Nach Satz 2.2 ist nun $\exp : (-\infty, \infty) \to (0, \infty)$ bijektiv, und die Umkehrfunktion $\log : (0, \infty) \to (-\infty, \infty)$ erfüllt (2.3). Außerdem gilt $\log(1) = \log(\exp(0)) = 0$, $\log(e) = \log(\exp(1)) = 1$. Die Funktionalgleichung (2.4) ergibt sich aus $\exp(x_1)\exp(x_2) = \exp(x_1 + x_2)$, indem wir $x_k = \log(y_k)$ einsetzen und den Logarithmus nehmen. □

Definition 2.1 (Potenz mit reellen Exponenten) *Für $a > 0$, $x \in \mathbb{R}$ definieren wir*

$$a^x = \exp\left(x \log(a)\right).$$

Für rationale Exponenten $r = p/q$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ ist diese Definition konsistent mit der bereits gegebenen Definition als eindeutig bestimmte Lösung y der Gleichung $y^q = a^p$, denn $\exp(x\log(a)) > 0$ und nach Folgerung 4.1 gilt

$$\exp\left(r\log(a)\right)^q = \exp\left(q \cdot r\log(a)\right) = \exp\left(p\log(a)\right) = \exp\left(\log(a)\right)^p = a^p.$$

Insbesondere können wir also $\exp(x) = e^x$ schreiben. Durch die klassische Definition der Potenz ist die Funktion $x \mapsto a^x$ für alle $x \in \mathbb{Q}$ und damit auf einer dichten Teilmenge von $\mathbb{R}$ schon definiert. Die Exponentialfunktion setzt diese Funktion auf alle $x \in \mathbb{R}$ fort, genauer ist es die eindeutig bestimmte Fortsetzung, die stetig ist (Übungsaufgabe). Die Regeln der Potenzrechnung für $a, b > 0$

$$\begin{aligned} a^x\, a^y &= a^{x+y} & (a^x)^y &= a^x y \\ \left(\frac{1}{a}\right)^x &= a^{-x} & a^x\, b^x &= (ab)^x \end{aligned}$$

ergeben sich leicht aus den Definitionen und den Funktionalgleichungen von exp bzw. log. Weiter sind nun auch die Potenzfunktionen für beliebige reelle Exponenten erklärt:

$$f : (0, \infty) \to \mathbb{R},\ f(x) = x^\alpha = \exp\left(\alpha\log(x)\right).$$

Folgerung 2.2 (Charakterisierung von exp durch die Funktionalgleichung) *Die Funktion $f : \mathbb{R} \to \mathbb{R}$ habe folgenden beiden Eigenschaften:*

(1) $f(x + y) = f(x)\, f(y)$ *für alle $x, y \in \mathbb{R}$,*

(2) $\lim_{x\to 0} f(x) = f(0)$ *(Stetigkeit bei $x_0 = 0$).*

Dann ist entweder $a := f(1) > 0$ und $f(x) = a^x$ für alle x, oder f ist identisch Null.

Beweis: Es ist $f(1) = f\left(\frac{1}{2}\right)^2 \geq 0$. Im Fall $f(1) = 0$ folgt für alle $x \in \mathbb{R}$

$$f(x) = f\big((x-1)+1\big) = f(x-1)\, f(1) = 0.$$

Im Fall $a > 0$ kann wie oben argumentiert werden, das heißt man zeigt mit der Funktionalgleichung $f(x) = a^x$ für alle $x \in \mathbb{Q}$, und folgert $f(x) = a^x$ für alle $x \in \mathbb{R}$ aus der Stetigkeit. □

Nachdem die Diskussion der reellen Exponentialfunktion so ergiebig war, wollen wir nun auch die komplexe Exponentialfunktion anschauen; das wird sogar noch interessanter. Wir betrachten mit $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$ die Abbildung

$$c : \mathbb{R} \to \mathbb{S}^1 \subset \mathbb{C},\; c(t) = \exp(it) = e^{it}, \qquad \text{wobei } i = \sqrt{-1}. \tag{2.5}$$

In Folgerung 4.1 hatten wir schon gezeigt, dass c tatsächlich in den Einheitskreis abbildet, und zwar gilt

$$|c(t)|^2 = c(t)\overline{c(t)} = e^{it}\overline{e^{it}} = e^{it}e^{-it} = 1.$$

Das Ziel ist nun zu verstehen, wie sich der Punkt $c(t)$ in Anhängigkeit von t auf dem Einheitskreis bewegt, genauer wollen wir sehen, dass $c(t)$ periodisch ist und in jeder Periode den Einheitskreis einmal im mathematisch positiven Sinn (also gegen den Uhrzeigersinn) durchläuft. Erstmal geben wir den Komponenten von $c(t) = e^{it}$ einen Namen.

Definition 2.2 *Die Funktionen Cosinus und Sinus* $\cos, \sin : \mathbb{R} \to \mathbb{R}$ *sind definiert durch*

$$\cos t = \operatorname{Re} e^{it} = \frac{e^{it} + e^{-it}}{2} \qquad \textit{und} \qquad \sin t = \operatorname{Im} e^{it} = \frac{e^{it} - e^{-it}}{2i}. \tag{2.6}$$

Es gilt also die Eulersche Formel

$$e^{it} = \cos t + i\, \sin t \qquad \textit{für alle } t \in \mathbb{R}. \tag{2.7}$$

Aus der Definition folgen die Potenzreihendarstellungen der Funktionen Sinus und Cosinus.

Satz 2.5 *Für alle* $t \in \mathbb{R}$ *gelten die absolut konvergenten Potenzreihendarstellungen*

$$\begin{aligned} \cos t &= \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} = 1 - \frac{t^2}{2!} \pm \ldots \qquad \textit{und} \\ \sin t &= \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} = t - \frac{t^3}{3!} \pm \ldots. \end{aligned}$$

Beweis: Die absolute Konvergenz ergibt sich mit Satz 4.3 aus der Abschätzung

$$\sum_{k=0}^{N} \frac{|t|^{2k}}{(2k)!} + \sum_{k=0}^{N} \frac{|t|^{2k+1}}{(2k+1)!} \leq \sum_{n=0}^{\infty} \frac{|t|^n}{n!} = e^{|t|} < \infty.$$

Mit $i^{2k} = (-1)^k$ sowie $i^{2k+1} = i\,(-1)^k$ berechnen wir

$$\sum_{n=0}^{2N+1} \frac{(it)^n}{n!} = \sum_{k=0}^{N} (-1)^k \frac{t^{2k}}{(2k)!} + i \sum_{k=0}^{N} (-1)^k \frac{t^{2k+1}}{(2k+1)!}.$$

Für $N \to \infty$ ist die rechte Seite konvergent, die linke Seite geht gegen $e^{it} = \cos t + i \sin t$. □

Lemma 2.2 *Die Funktionen* $\cos, \sin : \mathbb{R} \to \mathbb{R}$ *sind stetig.*

BEWEIS: $c : \mathbb{R} \to \mathbb{R}^2 = \mathbb{C}$, $c(t) = e^{it}$, ist stetig als Verkettung der stetigen Abbildungen $\varphi : \mathbb{R} \to \mathbb{C}$, $\varphi(t) = it$, und $\exp : \mathbb{C} \to \mathbb{C}$, vergleiche Satz 2.3. Dann sind auch $\cos, \sin$ als Koordinatenfunktionen von $c(t)$ stetig. $\square$

Lemma 2.3 *Die Funktionen Cosinus und Sinus haben folgende Eigenschaften:*

(1) $\cos^2 t + \sin^2 t = 1$ *für alle* $t \in \mathbb{R}$.

(2) *Cosinus ist eine gerade und Sinus eine ungerade Funktion:*

$$\cos(-t) = \cos t \quad \textit{und} \quad \sin(-t) = -\sin t.$$

(3) *Für alle* $\alpha, \beta \in \mathbb{R}$ *gelten die Additionstheoreme*

$$\begin{aligned} \cos(\alpha + \beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta, \\ \sin(\alpha + \beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta. \end{aligned}$$

BEWEIS: Behauptung (1) folgt direkt aus $|e^{it}| = 1$, Behaupung (2) gilt nach Definition und (3) ergibt sich aus der Funktionalgleichung der Exponentialfunktion und Vergleich der Real- und Imaginärteile:

$$\cos(\alpha + \beta) + i\sin(\alpha + \beta) = e^{i(\alpha+\beta)} = e^{i\alpha}\, e^{i\beta} = (\cos\alpha + i\sin\alpha)\,(\cos\beta + i\sin\beta).$$

$\square$

Ersetzen wir in den Additionstheoremen α, β durch $(\alpha + \beta)/2, (\alpha - \beta)/2$, so erhalten wir

$$\begin{aligned} \cos\alpha &= \cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} - \sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} \\ \sin\alpha &= \sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} + \cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}. \end{aligned}$$

Durch Vertauschen von α und β und Subtraktion ergibt sich mit Lemma 2.3(2)

$$\begin{aligned} \cos\alpha - \cos\beta &= -2\sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} \\ \sin\alpha - \sin\beta &= 2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} \end{aligned} \tag{2.8}$$

Wir benötigen den folgenden Grenzwert.

Lemma 2.4 *Es gilt* $\lim_{t\to 0} \dfrac{\sin t}{t} = 1$.

BEWEIS: Wir verwenden die Abschätzung der Exponentialfunktion aus Satz 4.7 mit $z = it$ und $n = 2$. Wegen $|\mathrm{Im}\, z| \le |z|$ gilt danach für $|t| \le 3/2$

$$|\sin t - t| \le |e^{it} - (1 + it)| \le t^2.$$

Nach Division durch $|t| > 0$ folgt mit $t \to 0$ die Behauptung. $\square$

Satz 2.6 (Definition von π) *Es gibt eine eindeutig bestimmte Zahl $\tau \in (0,\infty)$ mit folgenden Eigenschaften:*

(1) *Auf $(0,\tau)$ gilt $\cos t, \sin t > 0$.*

(2) *Auf $[0,\tau)$ ist $\cos$ streng monoton fallend und $\sin$ streng monoton wachsend.*

(3) *$\cos\tau = 0$ und $\sin\tau = 1$, also $e^{i\tau} = i$.*

Wir definieren $\pi = 2\tau$.

Beweis: Als erstes zeigen wir, dass die Funktionen cos und sin auf einem hinreichend kleinen Intervall $(0,t)$ strikt positiv sind. Wegen $\cos 0 = 1$ folgt das für den Cosinus sofort aus der Stetigkeit. Für den Sinus gibt es wegen Lemma 2.4 ein $t > 0$, so dass für $0 < t' < t$ gilt:

$$\frac{\sin t'}{t'} \geq 1/2 \quad \Rightarrow \quad \sin t' \geq t'/2 > 0.$$

Also ist die Menge $M = \{t > 0 : \cos t', \sin t' > 0 \text{ für alle } t' \in (0,t)\}$ nichtleer und wir setzen $\tau = \sup M$. Behauptung (1) gilt dann nach Definition. Für $0 \leq t_1 < t_2 < \tau$ folgt mit den Formeln (2.8), Lemma 2.3(2) und Aussage (1), da $0 < (t_2-t_1)/2 \leq (t_2+t_1)/2 < \tau$,

$$\begin{aligned}
\cos t_2 - \cos t_1 &= -2\sin\frac{t_2+t_1}{2}\sin\frac{t_2-t_1}{2} < 0,\\
\sin t_2 - \sin t_1 &= 2\cos\frac{t_2+t_1}{2}\sin\frac{t_2-t_1}{2} > 0.
\end{aligned}$$

Dies zeigt Behauptung (2), insbesondere existieren die Grenzwerte $\xi = \lim_{t\nearrow\tau}\cos t \in [0,1)$ und $\eta = \lim_{t\nearrow\tau}\sin t \in (0,1]$, vgl. Lemma 2.1. Um zu zeigen, dass τ endlich ist, betrachten wir für ein $t_0 \in (0,\tau)$ die Folge $(e^{it_0})^k$, $k \in \mathbb{N}$. Wäre $\tau = +\infty$, so folgt aus dem Bewiesenen

$$(e^{it_0})^k = e^{ikt_0} = \cos kt_0 + i\sin kt_0 \to \xi + i\eta \quad \text{mit } k \to \infty.$$

Aber $|(e^{it_0})^{k+1} - (e^{it_0})^k| = |(e^{it_0})^k(e^{it_0}-1)| = |e^{it_0}-1| > 0$, das heißt die Folge ist keine Cauchyfolge, ein Widerspruch. Also ist $\tau < \infty$, und mit der Stetigkeit folgt $\cos\tau = \xi$ und $\sin\tau = \eta$. Nun ist $\eta > 0$, also muss $\xi = 0$ gelten, denn sonst wären cos und sin beide strikt positiv nahe bei τ, also τ keine obere Schranke für M. Wegen $\xi^2 + \eta^2 = 1$ und $\eta > 0$ ist $\eta = +1$. Schließlich ist klar, dass τ durch (1) und (3) eindeutig festgelegt ist. □

Mit Satz 2.6 erhalten wir

$$e^{i(t+\pi/2)} = e^{i\pi/2}e^{it} = ie^{it}, \tag{2.9}$$

bzw. wegen $e^{i(t+\pi/2)} = \cos(t+\pi/2) + i\sin(t+\pi/2)$ und $ie^{it} = -\sin t + i\cos t$

$$\cos(t+\pi/2) = -\sin t \quad \text{und} \quad \sin(t+\pi/2) = \cos t. \tag{2.10}$$

Damit können wir eine Wertetabelle der Funktionen erstellen. Die Ziffern I bis IV bedeuten die gegen den Uhrzeiger nummerierten Quadranten.

t	0		$\pi/2$		π		$3\pi/2$		2π
e^{it}	1	I	i	II	-1	III	$-i$	IV	1
$\cos$	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0	$\nearrow$	1
$\sin$	0	$\nearrow$	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0

Folgerung 2.3 *Die Funktionen* e^{it}, $\cos t$ *und* $\sin t$ *sind* 2π*-periodisch, und* 2π *ist die kleinste Periode für jede der Funktionen.*

Beweis: Es gilt $e^{2\pi i} = 1$ und damit $e^{i(t+2\pi)} = e^{it}e^{2\pi i} = e^{it}$ für alle t, also sind die Funktionen 2π-periodisch. Wäre e^{it} periodisch mit Periode $p \in (0, 2\pi)$, so folgt $e^{ip} = e^{i0} = 1$, was der Wertetabelle widerspricht. Weiter ist eine Periode von Cosinus oder Sinus auch Periode von e^{it} wegen (2.10). Damit ist die Folgerung bewiesen. $\square$

Wir können zum Beispiel nun die Nullstellen der trigonometrischen Funktionen angeben.

$$e^{it} = 1 \quad \Leftrightarrow \quad t = 2k\pi \quad \text{mit } k \in \mathbb{Z}, \tag{2.11}$$
$$\cos t = 0 \quad \Leftrightarrow \quad t = \pi/2 + k\pi \quad \text{mit } k \in \mathbb{Z}, \tag{2.12}$$
$$\sin t = 0 \quad \Leftrightarrow \quad t = k\pi \quad \text{mit } k \in \mathbb{Z}. \tag{2.13}$$

Satz 2.7 (Arcusfunktionen) *Die Funktionen* $\cos : [0, \pi] \to [-1, 1]$ *und* $\sin : [-\pi/2, \pi/2] \to [-1, 1]$ *sind streng monoton fallend bzw. wachsend und bijektiv. Ihre Umkehrfunktionen* $\arccos : [-1, 1] \to [0, \pi]$ *bzw.* $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ *sind gleichfalls stetig, streng monoton fallend bzw. wachsend und es gilt*

$$\lim_{x \searrow -1} \arccos x = \pi \quad \textit{und} \quad \lim_{x \nearrow 1} \arccos x = 0,$$
$$\lim_{y \searrow -1} \arcsin y = -\pi/2 \quad \textit{und} \quad \lim_{y \nearrow 1} \arcsin y = \pi/2.$$

Beweis: Alle Aussagen folgen aus Satz 2.2 und unserer Wertetabelle. $\square$

Satz 2.8 (Polarkoordinaten) *Zu jedem* $z \in \mathbb{C}\backslash\{0\}$ *gibt es eindeutig bestimmte* $r > 0$, $\vartheta \in [0, 2\pi)$ *mit* $z = re^{i\vartheta}$.

Beweis: Wir behaupten, dass die Abbildung $c : [0, 2\pi) \to \mathbb{S}^1$, $c(t) = e^{it}$, bijektiv ist. Denn aus $e^{it_1} = e^{it_2}$ folgt $e^{i(t_1 - t_2)} = 1$, also $t_1 - t_2 = 2k\pi$ mit $k \in \mathbb{Z}$. Für $t_1, t_2 \in (0, 2\pi)$ folgt dann $t_1 = t_2$. Für die Surjektivität definieren wir für $x + iy \in \mathbb{S}^1$

$$\vartheta = \begin{cases} \arccos x & \text{für } y \geq 0 \\ 2\pi - \arccos x & \text{für } y < 0. \end{cases}$$

Wegen $\arccos x \in [0, \pi]$ und $\sin \vartheta = \sqrt{1 - \cos^2 \vartheta}$ für $\vartheta \in [0, \pi]$ folgt für $y \geq 0$

$$e^{i\vartheta} = \cos(\arccos x) + i \sin(\arccos x) = x + i\sqrt{1 - x^2} = x + iy = z.$$

Aus $\cos(2\pi - \arccos x) = \cos x$ und $\sin(2\pi - x) = -\sin x$ folgt für $y < 0$

$$e^{i\vartheta} = \cos(\arccos x) - i \sin(\arccos x) = x - i\sqrt{1 - x^2} = x + iy = z.$$

Für allgemeines $z \in \mathbb{C}\backslash\{0\}$ setzen wir $r = |z| > 0$ und wählen $\vartheta \in [0, 2\pi)$ mit $e^{i\vartheta} = z/|z|$. Es folgt $z = re^{i\vartheta}$ wie gewünscht. $\square$

Die durch den Satz definierte Funktion $\arg : \mathbb{C}\backslash\{0\} \to [0, 2\pi)$, die jedem z den Polarwinkel $\vartheta = \arg(z)$ zuordnet, hat längs der positiven x-Achse einen Sprung:

$$\lim_{y \searrow 0} \arg(x + iy) = 0 \quad \text{und} \quad \lim_{y \nearrow 0} \arg(x + iy) = 2\pi \quad \text{für alle } x > 0.$$

Ansonsten ist die Funktion stetig. In der Schule werden die Funktionen Cosinus, Sinus und Tangens durch die Längenverhältnisse am rechtwinkligen Dreieck eingeführt. Um dies jedoch zu einer exakten Definition zu machen, muss ein Konzept zur Messung von Winkeln zur Verfügung stehen, es muss also die Länge eines Einheitskreisbogens definiert werden. An dieser Stelle wollen dies nur ad hoc betrachten. Um die Länge eines Bogens

$$c : [0, \alpha] \to \mathbb{S}^1,\ c(t) = e^{it},$$

zu bestimmen, wählen wir $t_k = k\alpha/n$ mit $k = 0, 1, \ldots, n$ und erhalten die Näherung

$$L_n = \sum_{k=1}^{n} |c(t_k) - c(t_{k-1})| = \sum_{k=1}^{n} |e^{i(k-1)\alpha/n}(e^{i\alpha/n} - 1)| = n|e^{i\alpha/n} - 1|.$$

Nach Satz 4.7 gilt für $|z| \leq 3/2$ die Abschätzung $|e^z - (1 + z)| \leq |z|^2$, also ergibt sich

$$|L_n - \alpha| = n\big||e^{i\alpha/n} - 1| - |i\alpha/n|\big| \leq n\,|e^{i\alpha/n} - 1 - i\alpha/n| \leq n\,(\alpha/n)^2 \to 0 \quad \text{mit } n \to \infty.$$

Die Länge von c auf $[0, \alpha]$ ist also gleich α, das heißt c bildet längentreu ab und insbesondere ist die Länge des Vollkreises 2π. wie es sein sollte.

Kapitel 4

Differentialrechnung für Funktionen einer Variablen

1 Die Ableitung: Definition und Regeln

Im diesem Abschnitt betrachten wir stets reellwertige oder vektorwertige Funktionen einer Variablen, die auf einem offenen Intervall $I \subset \mathbb{R}$ definiert sind.

Definition 1.1 (Ableitung) *Die Funktion $f : I \to \mathbb{R}^n$ hat in $x_0 \in I$ die Ableitung $a \in \mathbb{R}^n$ (Notation: $f'(x_0) = a$), falls gilt:*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = a. \tag{1.1}$$

Wir nennen f differenzierbar in x_0, falls es ein $a \in \mathbb{R}^n$ mit (1.1) gibt, falls also der in (1.1) betrachtete Grenzwert existiert.

Eine alternative Formulierung ergibt sich durch die Substitution $x = x_0 + h$:

$$f'(x_0) = a \quad \Leftrightarrow \quad \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = a.$$

Leibniz interessierte sich für die Definition der Ableitung im Zusammenhang mit dem Problem, die Tangente an eine ebene Kurve in einem gegebenen Punkt zu definieren. Nehmen wir dazu an, dass die Kurve als Graph einer Funktion $f : I \to \mathbb{R}$ gegeben ist, und dass die Tangente im Punkt $(x_0, f(x_0))$ gesucht ist. Der Differenzenquotient

$$\frac{f(x) - f(x_0)}{x - x_0} \quad (x_0, x \in I,\, x \neq x_0)$$

ist geometrisch die Steigung der Sekante durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$. Die Existenz der Ableitung bedeutet, dass die Sekantensteigungen für $x \to x_0$ gegen den Wert $f'(x_0)$ konvergieren. Die Tangente wird nun definiert als die Gerade, die durch den Punkt $(x_0, f(x_0))$ geht und die Steigung $f'(x_0)$ hat. Daraus ergibt sich ihre Gleichung

$$y = f(x_0) + f'(x_0)(x - x_0) \quad \text{für alle } x \in \mathbb{R}.$$

Im Fall von vektorwertigen Funktionen, also $n \geq 2$, ist der Grenzwert in (1.1) wie üblich bezüglich der Euklidischen Norm aufzufassen. Der in Kapitel 2 bewiesene Satz 5.3 über die Äquivalenz von Normkonvergenz und Konvergenz der Koordinaten besagt hier Folgendes:

Lemma 1.1 *Die Funktion $f : I \to \mathbb{R}^n$ ist genau dann in $x_0 \in I$ differenzierbar, wenn alle Koordinatenfunktionen $f_i : I \to \mathbb{R}$, $i = 1, \ldots, n$, in x_0 differenzierbar sind. Die Ableitung kann dann koordinatenweise berechnet werden, das heißt es gilt:*

$$f'(x_0) = ((f_1)'(x_0), \ldots, (f_n)'(x_0)) \in \mathbb{R}^n.$$

Newton entwickelte den Differentialkalkül (Englisch: Calculus) unter anderem um die Keplerschen Gesetze für die Planetenbewegung zu begründen, genauer konnte er diese Gesetze alle aus dem Gravitationsgesetz ableiten. Dazu wird die Bewegung eines Planeten durch eine Abbildung

$$f : I \to \mathbb{R}^3, \quad f(t) = \big(x(t), y(t), z(t)\big),$$

beschrieben, also durch dessen Koordinaten zur Zeit $t \in I$ bezüglich eines Euklidischen Koordinatensystems. Erstes Ziel ist dann die Definition der Momentangeschwindigkeit als Vektor in $\mathbb{R}^3$. Die vektorielle Durchschnittsgeschwindigkeit auf dem Zeitintervall $[t_0, t]$ ist der Quotient von Weg und Zeit, also gleich

$$\frac{f(t) - f(t_0)}{t - t_0} \in \mathbb{R}^3.$$

Die Momentangeschwindigkeit $v(t_0)$ zum Zeitpunkt $t = t_0$ ist deshalb als vektorielle Ableitung zu definieren, wobei Newton einen Punkt statt eines Strichs benutzt hat:

$$v(t_0) = f'(t_0) = (x'(t_0), y'(t_0), z'(t_0)) \in \mathbb{R}^3.$$

Definition 1.2 (Ableitungsfunktion) *Die Funktion $f : I \to \mathbb{R}^n$ heißt differenzierbar auf I (oder einfach differenzierbar), falls f in jedem Punkt $x_0 \in I$ differenzierbar ist. Die hierdurch gegebene Funktion*

$$f' : I \to \mathbb{R}^n, \quad x_0 \mapsto f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \in \mathbb{R}^n$$

heißt Ableitungsfunktion oder schlicht Ableitung von f.

Beispiel 1.1 Für eine konstante Funktion, also $f(x) = c \in \mathbb{R}^n$ für alle $x \in I$, gilt für $x \neq x_0$:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0 \quad \Rightarrow \quad f'(x_0) = 0 \text{ bzw. } f' = 0.$$

Beispiel 1.2 Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x$, gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1 \quad \text{für alle } x \neq x_0,$$

also folgt $f'(x_0) = 1$.

Beispiel 1.3 Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$, ist nicht differenzierbar in $x_0 = 0$:

$$\lim_{x \searrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \searrow 0} \frac{x}{x} = 1 \quad \text{und} \quad \lim_{x \nearrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \nearrow 0} \frac{-x}{x} = -1.$$

Die rechts- und linksseitige Ableitung existieren in $x_0 = 0$, sie sind aber verschieden.

Für die Ableitung der Exponentialfunktion gehen wir ähnlich wie bei der Stetigkeit vor, indem wir erst den Fall $x = 0$ mit einer expliziten Abschätzung behandeln und dann die Funktionalgleichung einsetzen.

Satz 1.1 *Die Funktion* $\exp : \mathbb{R} \to \mathbb{R}$ *ist differenzierbar mit Ableitung* $\exp' = \exp$.

Beweis: Für $x = 0$ folgt die Aussage aus Satz 4.7 in Kapitel 2, und zwar mit $n = 2$:

$$\left| \frac{e^x - e^0}{x - 0} - e^0 \right| = \left| \frac{e^x - 1}{x} - 1 \right| = \frac{|e^x - (1 + x)|}{|x|} \leq |x| \to 0 \qquad \text{mit } x \to 0.$$

Für $x \neq 0$ schließen wir weiter mit der Funktionalgleichung

$$\frac{e^{x+h} - e^x}{h} = e^x \, \frac{e^h - 1}{h} \to e^x \qquad \text{mit } h \to 0.$$

$\square$

Satz 1.2 *Die Abbildung* $c : \mathbb{R} \to \mathbb{C}$, $c(t) = e^{it}$, *ist differenzierbar mit Ableitung* $c' = ic$. *Insbesondere sind auch* $\cos$ *und* $\sin$ *differenzierbar und es gilt*

$$\cos' = -\sin \qquad \textit{und} \qquad \sin' = \cos\,.$$

Beweis: Für die Ableitung in $t = 0$ verwenden wir wieder Satz 4.7 in Kapitel 2, mit $n = 2$,

$$\left| \frac{c(t) - c(0)}{t - 0} - ic(0) \right| = \left| \frac{e^{it} - 1}{t} - i \right| = \frac{|e^{it} - (1 + it)|}{|t|} \leq |t| \to 0 \qquad \text{mit } t \to 0.$$

Für beliebige t ergibt sich mit der Funktionalgleichung

$$\frac{c(t + h) - c(t)}{h} = \frac{e^{i(t+h)} - e^{it}}{h} = e^{it} \, \frac{e^{ih} - 1}{h} \to ie^{it} = ic(t) \qquad \text{für } h \to 0.$$

Lemma 1.1 besagt nun, dass cos und sin als Koordinatenfunktionen von c ebenfalls differenzierbar sind mit Ableitung $c'(t) = \cos'(t) + i \sin'(t)$. Wegen $ic(t) = -\sin t + i \cos t$ folgt $\cos' = -\sin$ und $\sin' = \cos$ wie behauptet. $\square$

Der Satz besagt anschaulich, dass der Punkt $c(t) = e^{it}$ den Einheitskreis mit konstanter Absolutgeschwindigkeit $|c'(t)| = 1$ durchläuft. Der Umlaufsinn ist dabei positiv, das heißt der Vektor ic' weist ins Innere des Kreises.

Das folgende Lemma gibt eine geringfügige Umformulierung des Begriffs der Differenzierbarkeit. Danach ist f genau dann in x_0 differenzierbar, wenn f mit einer affinlinearen Funktion übereinstimmt bis auf einen Fehler, der schneller als linear verschwindet.

Lemma 1.2 *Genau dann hat* $f : I \to \mathbb{R}^n$ *in* $x_0 \in I$ *die Ableitung* $a \in \mathbb{R}^n$, *wenn für die Differenzfunktion* $r(h) = f(x_0 + h) - (f(x_0) + ah)$ *gilt:*

$$\lim_{h \to 0} \frac{r(h)}{h} = 0. \tag{1.2}$$

Beweis: Für $h \neq 0$ gilt

$$\frac{f(x_0+h)-f(x_0)}{h} - a = \frac{f(x_0+h)-(f(x_0)+a\,h)}{h} = \frac{r(h)}{h}.$$

Hieraus ergibt sich die Behauptung. □

Satz 1.3 *Ist $f : I \to \mathbb{R}^n$ differenzierbar in $x_0 \in I$, so ist f auch stetig in x_0.*

Beweis: Nach Lemma 1.4 in Kapitel 3 reicht es zu zeigen, dass $\lim_{h\to 0} f(x_0+h) = f(x_0)$. Aber für $h \neq 0$ gilt nach Lemma 1.2

$$f(x_0+h) = f(x_0) + ah + h\,\frac{r(h)}{h} \to f(x_0) \text{ mit } h \to 0.$$

□

Satz 1.4 (Differentiationsregeln) *Seien $f, g : I \to \mathbb{R}$ differenzierbar in $x_0 \in I$. Dann sind auch die Funktionen $\alpha f + \beta g$ $(\alpha, \beta \in \mathbb{R})$, fg und f/g (im Fall $g(x_0) \neq 0$) in x_0 differenzierbar mit folgenden Ableitungen:*

(1) *Linearität:*
$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$$

(2) *Produktregel:*
$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$$

(3) *Quotientenregel:*
$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$$

Beweis: Wir müssen jeweils für $x \neq x_0$ die Differenzenquotienten bilden und zeigen, dass diese mit $x \to x_0$ gegen das gewünschte konvergieren. Für (1) haben wir

$$\begin{aligned} \frac{(\alpha f(x) + \beta g(x)) - (\alpha f(x_0) + \beta g(x_0))}{x - x_0} &= \alpha\,\frac{f(x)-f(x_0)}{x-x_0} + \beta\,\frac{g(x)-g(x_0)}{x-x_0} \\ &\to \alpha f'(x_0) + \beta g'(x_0). \end{aligned}$$

Natürlich gilt die Aussage mit demselben Argument auch für vektorwertige Funktionen. Die Produktregel folgt durch „Mischen der Terme“:

$$\begin{aligned} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x)-f(x_0)}{x-x_0}\,g(x) + f(x_0)\,\frac{g(x)-g(x_0)}{x-x_0} \\ &\to f'(x_0)g(x_0) + f(x_0)g'(x_0), \end{aligned}$$

wobei die Stetigkeit von g in x_0 benutzt wurde (Satz 1.3). Wir zeigen die Quotientenregel zunächst für die Funktion $1/g$, also $f \equiv 1$. Es gibt ein $\delta > 0$ mit $g(x) \neq 0$ für $|x - x_0| < \delta$ nach Kapitel 3, Lemma 1.3. Für diese $x \in I$ gilt

$$\begin{aligned} \frac{1}{x-x_0}\left(\frac{1}{g(x)} - \frac{1}{g(x_0)}\right) &= -\,\frac{1}{g(x)g(x_0)}\,\frac{g(x)-g(x_0)}{x-x_0} \\ &\to -\,\frac{1}{g(x_0)^2}\,g'(x_0) \quad \text{mit } x \to x_0. \end{aligned}$$

Für beliebiges f schreiben wir $f/g = f \cdot \dfrac{1}{g}$ und verwenden die Produktregel. □

Beispiel 1.4 Für $f_n(x) = x^n$ folgt aus Beispiel 1.2, also $f_1' = 1$, und der Produktregel

$$f_n'(x) = (f_1 f_{n-1})'(x) = f_1'(x) f_{n-1}(x) + f_1(x) f_{n-1}'(x) = x^{n-1} + x\, f_{n-1}'(x),$$

und damit per Induktion $f_n'(x) = nx^{n-1}$. Allgemeiner ergibt sich mit Satz 1.4(1) für Polynome $p(x) = \sum_{k=0}^n a_k x^k$ die Formel

$$p'(x) = \sum_{k=1}^n k a_k x^{k-1} = \sum_{j=0}^{n-1} (j+1) a_{j+1} x^j.$$

Beispiel 1.5 Für $f(x) = x^{-n}$, $n \in \mathbb{N}$, gilt $f'(x) = -nx^{-n-1}$ nach der Quotientenregel:

$$f'(x) = -\frac{nx^{n-1}}{(x^n)^2} = -nx^{-n-1}.$$

Satz 1.5 (Kettenregel) *Seien $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ mit $f(I) \subset J$. Ist f in $x_0 \in I$ differenzierbar und g in $y_0 = f(x_0) \in J$ differenzierbar, so ist auch $g \circ f : I \to \mathbb{R}$ in x_0 differenzierbar und hat die Ableitung*

$$(g \circ f)'(x_0) = g'\big(f(x_0)\big)\, f'(x_0).$$

Beweis: Wir betrachten wieder für $x \neq x_0$ den Differenzenquotienten:

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \begin{cases} \dfrac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \, \dfrac{f(x) - f(x_0)}{x - x_0} & \text{falls } f(x) \neq f(x_0) \\ 0 & \text{falls } f(x) = f(x_0). \end{cases}$$

Wir definieren die Funktion

$$a : J \to \mathbb{R},\ a(y) = \begin{cases} \dfrac{g(y) - g(f(x_0))}{y - f(x_0)} & \text{für } y \neq f(x_0) \\ g'(f(x_0)) & \text{für } y = f(x_0). \end{cases}$$

Es ist a stetig in $f(x_0)$ nach Kapitel 3, Lemma 1.4. Mit $x \to x_0$ folgt, da $f(x) \to f(x_0)$,

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = a(f(x)) \frac{f(x) - f(x_0)}{x - x_0} \to a(f(x_0))\, f'(x_0) = g'(f(x_0)) f'(x_0).$$

$\square$

Satz 1.6 (Differenzierbarkeit der Umkehrfunktion) *Sei $f : I \to \mathbb{R}$ streng monoton und stetig auf dem offenen Intervall I, und differenzierbar in x_0 mit $f'(x_0) \neq 0$. Dann ist $I^* = f(I)$ ein offenes Intervall, und die Umkehrfunktion $g : I^* \to I$ ist differenzierbar in $y_0 = f(x_0)$ mit Ableitung*

$$g'(y_0) = \frac{1}{f'\big(g(y_0)\big)}.$$

Beweis: Nach Satz 2.2 in Kapitel 3 mit Zusatz (2.1) ist I^* ein offenes Intervall und $g : I^* \to \mathbb{R}$ streng monoton und stetig, insbesondere $g(y) \to g(y_0) = x_0$ mit $y \to y_0$. Für $y \neq y_0$ folgt

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{g(y) - g(y_0)}{f(g(y)) - f(g(y_0))} = \frac{1}{\frac{f(g(y)) - f(g(y_0))}{g(y) - g(y_0)}} \to \frac{1}{f'(x_0)}.$$

□

Seien I, I^* offene Intervalle, $f : I \to I^*$ bijektiv und in x_0 differenzierbar. Ist die Umkehrfunktion $g : I^* \to I$ in $y_0 = f(x_0)$ auch differenzierbar, so folgt schon aus der Kettenregel

$$g(f(x)) = x \quad \Rightarrow \quad g'(f(x_0))f'(x_0) = 1.$$

Insbesondere muss $f'(x_0) \neq 0$ sein, und es gilt die Formel für die Ableitung der Umkehrfunktion. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, ist streng monoton wachsend und stetig; ihre Umkehrfunktion lautet

$$g : \mathbb{R} \to \mathbb{R},\ g(y) = \begin{cases} y^{1/3} & \text{für } y \geq 0 \\ -|y|^{1/3} & \text{für } y \leq 0. \end{cases}$$

Die Umkehrfunktion ist nicht differenzierbar in $y = 0$, denn es gilt

$$\frac{g(y) - g(0)}{y - 0} = |y|^{-2/3} \to +\infty \text{ mit } y \to 0.$$

Die Voraussetzung des Satzes ist hier verletzt, es ist $f'(0) = 0$.

Die beiden vorangegangengen Regeln sind in der von Leibniz eingeführten Notation besonders suggestiv. Leibniz schreibt Funktionen in der Form $y = y(x)$ und bezeichnet die Ableitung mit dem Symbol $\frac{dy}{dx}$, das auch als *Differentialquotient* bezeichnet wird. Formal ergeben sich die Kettenregel und die Regel für die Ableitung der Umkehrfunktion dann aus der Bruchrechnung:

$$\begin{aligned} y = y(x),\ z = z(y) \quad &\Rightarrow \quad \frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}, \\ y = y(x),\ x = x(y) \quad &\Rightarrow \quad \frac{dx}{dy} = \Big(\frac{dy}{dx}\Big)^{-1}. \end{aligned}$$

Bei der Anwendung dieser saloppen Notation ist jedoch darauf zu achten, wo die jeweiligen Funktionen definiert sind.

Satz 1.7 *Die Funktion* $\log : (0, \infty) \to \mathbb{R}$ *ist differennzierbar mit Ableitung*

$$\log'(y) = \frac{1}{y}.$$

Beweis: Dies folgt aus Satz 1.1 und Satz 1.6, genauer ist

$$\log'(y) = \frac{1}{\exp'(\log y)} = \frac{1}{\exp(\log y)} = \frac{1}{y}.$$

$\square$

Als eine Anwendung von Satz 1.7 haben wir einen zweiten Beweis von

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

Es folgt aus

$$\lim_{n\to\infty} n\log\left(1+\frac{1}{n}\right) = \lim_{n\to\infty}\frac{\log(1+\frac{1}{n})}{\frac{1}{n}} = \log'(1) = 1$$

und die Stetigkeit von exp. Denn

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

Beispiel 1.6 Die Funktion $f : (0,\infty) \to \mathbb{R}$, $f(x) = x^\alpha$ mit $\alpha \in \mathbb{R}$ ist Verkettung der Funktionen $\exp : \mathbb{R} \to \mathbb{R}$ und $h : \mathbb{R} \to \mathbb{R}$, $h(x) = \alpha\log(x)$. Mit der Kettenregel berechnen wir

$$f'(x) = \exp'(h(x))h'(x) = \exp(\alpha\log(x))\,\frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

Beispiel 1.7 Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = a^x$ mit $a > 0$, ist Verkettung von exp und $h(x) = \log(a)x$, deshalb folgt

$$f'(x) = \exp'(h(x))h'(x) = \exp(\log(a)x)\ \log(a) = \log(a)a^x.$$

Beispiel 1.8 Die Differenzierbarkeit der Arcusfunktionen auf dem offenen Intervall $(-1,1)$ folgt aus den Sätzen 1.2 und 1.6. Beachtet man $\cos^2 + \sin^2 = 1$ sowie $\arccos x \in (0,\pi)$ bzw. $\arcsin x \in (-\pi/2, \pi/2)$, so erhält man

$$\begin{aligned}
\arccos'(x) &= \frac{1}{\cos'(\arccos x)} = -\frac{1}{\sin(\arccos x)} = -\frac{1}{\sqrt{1-x^2}}\\
\arcsin'(x) &= \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}}.
\end{aligned}$$

2 Mittelwertsatz und Anwendungen

Es ist eine zentrales Problem der Analysis, aus Eigenschaften der Ableitung auf Eigenschaften der Funktion selbst zu schließen. Der Mittelwertsatz ist dafür ein einfaches und effektives Hilfsmittel. Für seinen Beweis müssen wir allerdings erst über Extremwerte – Maxima und Minima – von stetigen Funktionen sprechen. Da sich keine Unterschiede ergeben, betrachten wir dabei gleich reellwertige Funktionen auf einer Teilmenge des $\mathbb{R}^n$.

Wir benötigen einen Satz, der die Existenz von Extremalstellen für eine stetige Funktion $f : D \to \mathbb{R}$ allgemein garantiert. Ohne Voraussetzungen an D kann das nicht gehen, wie etwa das folgende Beispiel zeigt:

$$f : D = (0,\infty) \to \mathbb{R},\ f(x) = \frac{1}{1+x}.$$

Hier wird weder das Supremum $\sup_{x\in D} f(x) = 1$ noch das Infimum $\inf_{x\in D} f(x) = 0$ durch die Funktion angenommen.

Satz 2.1 (Existenz von Extremalstellen) *Sei $D \subset \mathbb{R}^n$ nichtleer, abgeschlossen und beschränkt, und sei $f : D \to \mathbb{R}$ stetig. Dann ist f beschränkt und nimmt sein Infimum und Supremum an, das heißt es gibt $x_0, x_1 \in D$ mit*

$$f(x_0) = \inf_{x \in D} f(x) \quad \textit{und} \quad f(x_1) = \sup_{x \in D} f(x).$$

Beweis: Setze $\alpha = \inf_{x \in D} f(x) \in [-\infty, \infty)$. Wir zeigen die Existenz eines $x_0 \in D$ mit $f(x_0) = \alpha$, daraus folgt insbesondere $\alpha > -\infty$, das heißt f ist nach unten beschränkt. Für das Supremum kann analog argumentiert werden.

Nach Definition des Infimums gibt es eine Folge $x_k \in D$ mit $f(x_k) \to \alpha$. Da D beschränkt, ist die Folge x_k beschränkt. Nach Bolzano-Weierstraß, siehe Satz 2.8 bzw. Folgerung 5.2 in Kapitel 2, gibt es eine Teilfolge x_{k_j} und ein $x_0 \in \mathbb{R}^n$ mit $x_{k_j} \to x_0$ für $j \to \infty$. Aus D abgeschlossen, siehe Definition 5.7 in Kapitel 2, folgt $x_0 \in D$. Aber dann gilt wegen der Stetigkeit von f

$$-\infty < f(x_0) = \lim_{j \to \infty} f(x_{k_j}) = \alpha.$$

□

Im Beweis spielte die folgende Eigenschaft von D eine Rolle:

Definition 2.1 (Folgenkompaktheit) *Eine Menge $D \subset \mathbb{R}^n$ heißt folgenkompakt, wenn es zu jeder Folge $x_k \in D$ eine Teilfolge x_{k_j} gibt, die konvergiert und deren Grenzwert in D liegt:*

$$\lim_{j \to \infty} x_{k_j} = x_0 \in D.$$

Wir halten aus dem Beweis von Satz 2.1 folgende Beobachtung fest:

Satz 2.2 *Für $D \subset \mathbb{R}^n$ sind äquivalent:*

(1) *D ist folgenkompakt.*

(2) *D ist abgeschlossen und beschränkt.*

Beweis: Die Implikation (2) $\Rightarrow$ (1) wurde im Beweis von Satz 2.1 gezeigt. Jetzt setzen wir (1) voraus. Um die Abgeschlossenheit von D zu zeigen, sei $x_k \in D$ eine beliebige Folge mit $x_k \to x_0 \in \mathbb{R}^n$. Da D folgenkompakt, gibt es eine Teilfolge x_{k_j}, die gegen ein $x_0' \in D$ konvergiert. Aus der Eindeutigkeit des Grenzwerts folgt $x_0 = x_0' \in D$, also ist D abgeschlossen. Wäre D nicht beschränkt, so gibt es zu $k \in \mathbb{N}$ ein $x_k \in D$ mit $|x_k| \geq k$. Da D kompakt, existiert eine konvergente Teilfolge x_{k_j}, und diese ist nach Satz 1.2 in Kapitel 2 beschränkt, ein Widerspruch. □

Wir kommen jetzt wieder zurück zur eindimensionalen Situation.

Definition 2.2 (Lokale Extrema) *Die Funktion $f : (a, b) \to \mathbb{R}$ hat in $x_0 \in (a, b)$ ein lokales Minimum, falls es ein $\delta > 0$ gibt, so dass gilt:*

$$f(x_0) \leq f(x) \quad \textit{für alle } x \in U_\delta(x_0) = (x_0 - \delta, x_0 + \delta).$$

Ist sogar $f(x_0) < f(x)$ für $x \in U_\delta(x_0)$, $x \neq x_0$, so heißt das lokale Minimum isoliert. Ein (isoliertes) lokales Maximum ist entsprechend definiert.

Satz 2.3 (notwendige Bedingung für Extrema) *Die Funktion $f : (a,b) \to \mathbb{R}$ habe in $x_0 \in (a,b)$ ein lokales Extremum. Ist f in x_0 differenzierbar, so gilt $f'(x_0) = 0$.*

Beweis: Sei x_0 lokales Minimum von f. Dann gibt es ein $\delta > 0$ mit $f(x) \geq f(x_0)$ für $x \in U_\delta(x_0)$, also gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \quad \begin{cases} \geq 0 & \text{für } x \in (x_0, x_0 + \delta), \\ \leq 0 & \text{für } x \in (x_0 - \delta, x_0). \end{cases}$$

Mit $x \searrow x_0$ folgt $f'(x_0) \geq 0$, mit $x \nearrow x_0$ folgt $f'(x_0) \leq 0$. □

Die Funktion $f(x) = x^3$ erfüllt $f'(0) = 0$, aber in $x = 0$ liegt kein lokales Extremum vor. Die Bedingung $f'(x_0) = 0$ ist notwendig für eine lokale Extremalstelle einer differenzierbaren Funktion, aber sie ist nicht hinreichend.

Satz 2.4 (Mittelwertsatz der Differentialrechnung) *Sei $f : [a,b] \to \mathbb{R}$ stetig und differenzierbar auf (a,b). Dann gibt es ein $\xi \in (a,b)$ mit*

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

Beweis: Wir zeigen die Behauptung zuerst im Fall $f(a) = f(b) = 0$ (Satz von Rolle). Wir brauchen dann ein $\xi \in (a,b)$ mit $f'(\xi) = 0$. Nach Satz 2.1 gibt es $\xi_1, \xi_2 \in [a,b]$ mit mit

$$f(\xi_1) = \inf_{x \in [a,b]} f(x) \quad \text{und} \quad f(\xi_2) = \sup_{x \in [a,b]} f(x).$$

Ist $\xi_1 \in (a,b)$, so folgt $f'(\xi_1) = 0$ nach Satz 2.3 und wir können $\xi = \xi_1$ wählen. Analog, wenn $\xi_2 \in (a,b)$. Im verbleibenden Fall $\xi_1, \xi_2 \in \{a,b\}$ folgt $\inf f = \sup f = 0$ bzw. $f(x) = 0$ für alle $x \in [a,b]$, und damit auch $f'(x) = 0$ für alle x. Seien nun $f(a), f(b)$ beliebig. Definiere $h : [a,b] \to \mathbb{R}$ durch Abziehen der Sekante:

$$h(x) = f(x) - \left(f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \right).$$

Es gilt $h(a) = h(b) = 0$. Also existiert ein $\xi \in (a,b)$ mit

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

□

Folgerung 2.1 (Monotoniekriterium) *Sei f differenzierbar auf (a,b), stetig auf $[a,b]$. Dann gelten folgende Aussagen:*

$$\begin{aligned} f'(x) \geq 0 \text{ für alle } x \in (a,b) \quad &\Rightarrow \quad f \text{ ist wachsend auf } [a,b] \\ f'(x) \leq 0 \text{ für alle } x \in (a,b) \quad &\Rightarrow \quad f \text{ ist fallend auf } [a,b] \\ f'(x) = 0 \text{ für alle } x \in (a,b) \quad &\Rightarrow \quad f \text{ ist konstant.} \end{aligned}$$

Bei strikter Ungleichung folgt strenge Monotonie auf $[a,b]$.

BEWEIS: Sei $a \leq x_1 < x_2 \leq b$. Nach dem Mittelwertsatz gibt es ein $\xi \in (x_1, x_2)$, so dass gilt:

$$f(x_2) - f(x_1) = f'(\xi)\underbrace{(x_2 - x_1)}_{>0} \begin{cases} \geq 0 & \text{wenn } f'(\xi) \geq 0 \\ > 0 & \text{wenn } f'(\xi) > 0 \\ \leq 0 & \text{wenn } f'(\xi) \leq 0 \\ < 0 & \text{wenn } f'(\xi) < 0 \\ = 0 & \text{wenn } f'(\xi) = 0 \end{cases}.$$

□

Folgerung 2.2 (Schrankensatz) *Ist $f : [a, b] \to \mathbb{R}$ stetig und differenzierbar auf (a, b), so gilt für $a \leq x_1 < x_2 \leq b$:*

$$f'(x) \geq m \text{ für alle } x \in (a, b) \quad \Rightarrow \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq m$$

$$f'(x) \leq M \text{ für alle } x \in (a, b) \quad \Rightarrow \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq M.$$

BEWEIS: Wir zeigen die erste Aussage. Mit $g(x) = mx$ gilt $(f - g)' = f' - g' \geq m - m = 0$. Nach Folgerung 2.1 ist $f - g$ wachsend, ds heißt $f(x_2) - m\,x_2 \geq f(x_1) - m\,x_1$. Die Behauptung folgt. □

Der Mittelwertsatz gilt nicht für vektorwertige Funktionen $f : [a, b] \to \mathbb{R}^n$, $n \geq 2$. Technisch liegt das daran, dass die Stelle ξ aus dem Satz für die einzelnen Koordinatenfunktionen im allgemeinen verschieden ist. Ein konkretes Gegenbeispiel ist $c : [0, 2\pi] \to \mathbb{C}$, $c(t) = e^{it}$:

$$\frac{c(2\pi) - c(0)}{2\pi - 0} = 0, \qquad \text{aber} \qquad |c'(t)| = 1 \text{ für alle } t \in [0, 2\pi].$$

Dennoch können wir auch im vektorwertigen Fall eine Version des Schrankensatzes beweisen.

Folgerung 2.3 (f' beschränkt $\Rightarrow$ f Lipschitz) *Sei $f : [a, b] \to \mathbb{R}^n$ stetig und auf (a, b) differenzierbar. Falls $|f'(x)| \leq L$ für alle $x \in (a, b)$, so folgt*

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

BEWEIS: Um die Aussage auf den reellwertigen Fall zu reduzieren, betrachten wir für einen beliebigen Vektor $v \in \mathbb{R}^n$ mit $|v| = 1$ die Funktion

$$\varphi : [a, b] \to \mathbb{R}, \, \varphi(x) = \langle v, f(x) \rangle = \sum_{i=1}^{n} v_i f_i(x).$$

Aus der Ungleichung von Cauchy-Schwarz, siehe Satz 5.2 in Kapitel 2, folgt

$$\varphi'(x) = \langle v, f'(x) \rangle \leq |v|\,|f'(x)| \leq L.$$

Für $x_1, x_2 \in [a, b]$ mit $x_1 > x_2$ erhalten wir aus Folgerung 2.2

$$\langle v, f(x_1) - f(x_2) \rangle = \varphi(x_1) - \varphi(x_2) \leq L(x_1 - x_2) = L|x_1 - x_2|.$$

Wir können $f(x_1) \neq f(x_2)$ annehmen, denn sonst ist nichts zu zeigen. Nun wählen wir $v = \big(f(x_1) - f(x_2)\big)/\big|f(x_1) - f(x_2)\big|$, insbesondere $|v| = 1$, und erhalten

$$|f(x_1) - f(x_2)| = \langle f(x_1) - f(x_2), v\rangle \leq L|x_1 - x_2|.$$

□

Wir kommen jetzt zu höheren Ableitungen.

Definition 2.3 *Die k-te Ableitung von $f : (a,b) \to \mathbb{R}$ in x_0 ist induktiv definiert durch*

$$f^{(k)}(x_0) = \big(f^{(k-1)}\big)'(x_0).$$

Damit $f^{(k)}(x_0)$ definiert ist, müssen also die Ableitungen bis Ordnung $k-1$ in einer Umgebung von x_0 definiert sein, und $f^{(k-1)}$ muss in x_0 differenzierbar sein.

Natürlich schreiben wir f' und f'' statt $f^{(1)}$ bzw. $f^{(2)}$. Mit der zweiten Ableitung gewinnen wir genauere Informationen über lokale Extrema.

Satz 2.5 *Sei $f : (a,b) \to \mathbb{R}$ differenzierbar. In $x_0 \in (a,b)$ gelte $f'(x_0) = 0$, und $f''(x_0)$ sei definiert. Dann gilt:*

(1) *Ist $f''(x_0) > 0$, so hat f in x_0 ein isoliertes, lokales Minimum.*

(2) *Hat f in x_0 ein lokales Minimum, so folgt $f''(x_0) \geq 0$.*

Analoge Aussagen gelten mit umgekehrten Ungleichungen für Maxima.

Beweis: Da $f'(x_0) = 0$ nach Voraussetzung, gilt

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

Unter der Voraussetzung (1) gibt es ein $\delta > 0$ mit $f'(x) < 0$ auf $(x_0 - \delta, x_0)$ und $f'(x) > 0$ auf $(x_0, x_0 + \delta)$. Nach Folgerung 2.1 ist f dann streng monoton fallend auf $(x_0 - \delta, x_0)$ und und streng monoton wachsend auf $(x_0, x_0 + \delta)$, hat also in x_0 ein isoliertes, lokales Minimum. Wäre $f''(x_0) < 0$ in (2), so hätte f in x_0 ein isoliertes lokales Maximum im Widerspruch zur Voraussetzung. □

Die Funktion $f(x) = x^4$ zeigt, dass in einem isolierten Minimum $f''(x_0) = 0$ gelten kann. Häufig ist man nicht wirklich an den lokalen Minima, sondern am globalen Minimum interessiert. Dafür ist der Begriff der Konvexität relevant.

Definition 2.4 *Sei $I \subset \mathbb{R}$ ein Intervall. $f : I \to \mathbb{R}$ heißt konvex, falls gilt:*

$$f\big((1-t)x_0 + tx_1\big) \leq (1-t)f(x_0) + tf(x_1) \quad \forall x_0, x_1 \in I,\ t \in [0,1]. \tag{2.1}$$

Gilt dies mit $\geq$ statt mit $\leq$, so heißt f konkav.

Die Sekante durch $\big(x_0, f(x_0)\big)$ und $\big(x_1, f(x_1)\big)$ hat die Geradengleichung

$$y = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) =: g(x).$$

An der Stelle $x(t) = (1-t)x_0 + tx_1$ gilt

$$g\big(x(t)\big) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}\, t(x_1 - x_0) = (1-t)f(x_0) + tf(x_1).$$

Die Konvexität bedeutet also, dass der Graph von f stets unterhalb der Sekanten liegt,

Satz 2.6 (Konvexitätskriterien) *Sei $f : (a,b) \to \mathbb{R}$ differenzierbar. Dann sind folgende Aussagen äquivalent:*

(1) *f' ist monoton wachsend auf (a,b).*

(2) *f ist konvex.*

(3) $f(x_1) \geq f(x_0) + f'(x_0)(x_1 - x_0)$ *für alle $x_0, x_1 \in (a,b)$.*

Ist f zweimal differenzierbar auf (a,b), so ist außerdem äquivalent:

(4) $f'' \geq 0$.

Beweis: Wir zeigen (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1). Sei also (1) erfüllt. Wäre f nicht konvex, so gibt es $x_0, x_1 \in I$ mit der Eigenschaft, dass die stetige Funktion

$$g : [0,1] \to \mathbb{R},\ g(t) = f\big((1-t)x_0 + tx_1\big) - \big((1-t)f(x_0) + tf(x_1)\big)$$

für gewisse $t \in [0,1]$ strikt positiv ist. Da $g(0) = g(1) = 0$, nimmt g ein strikt positives Maximum in $t_0 \in (a,b)$ an, siehe Satz 2.1, und es gilt $g'(t_0) = 0$ wegen Satz 2.3. Aber $g'(t)$ ist monoton wachsend, denn es gilt

$$g'(t) = f'\big(x_0 + t(x_1 - x_0)\big) - \big(f(x_1) - f(x_0)\big).$$

Also ist $g'(t) \geq 0$ für $t \geq t_0$, und mit Folgerung 2.1 folgt $g(1) \geq g(t_0) > 0$, Widerspruch.

Sei jetzt f konvex, das heißt für $t \in (0,1)$ und $x_0, x_1 \in (a,b)$ mit $x_0 \neq x_1$ gilt

$$(x_1 - x_0)\,\frac{f\big(x_0 + t(x_1 - x_0)\big) - f(x_0)}{t(x_1 - x_0)} \leq f(x_1) - f(x_0).$$

Mit $t \searrow 0$ folgt $(x_1 - x_0)f'(x_0) \leq f(x_1) - f(x_0)$, dies ist die behauptete Ungleichung (3). Schließlich setzen wir (3) voraus. Indem wir dort x_0 und x_1 vertauschen und addieren, folgt

$$\big(f'(x_1) - f'(x_0)\big)(x_1 - x_0) \geq 0,$$

womit (1) gezeigt ist. Ist f zweimal differenzierbar, so impliziert $f'' \geq 0$ Bedingung (1) nach Folgerung 2.1, und umgekehrt folgt $f'' \geq 0$ aus (1) einfach durch Betrachtung des Differenzenquotienten. $\square$

Beispiel 2.1 (Youngsche Ungleichung) Für $x, y \geq 0$ gilt die Ungleichung

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \qquad \text{falls } p, q \in (1,\infty) \text{ mit } \frac{1}{p} + \frac{1}{q} = 1.$$

Dazu betrachten wir für festes $y > 0$ die Funktion

$$f : (0,\infty) \to \mathbb{R},\ f(x) = xy - \frac{x^p}{p}.$$

Es gilt $f'(x) = y - x^{p-1}$ und $f''(x) = -(p-1)x^{p-2} \leq 0$, das heißt f ist konkav nach Satz 2.6. Aber $f'(x_0) = 0$ für $x_0 = y^{1/(p-1)}$, also folgt mit Satz 2.6(3) wegen $p/(p-1) = q$

$$xy - \frac{x^p}{p} = f(x) \leq f(x_0) + f'(x_0)(x - x_0) = y^{1/(p-1)}y - \frac{y^{p/(p-1)}}{p} = \frac{y^q}{q}.$$

Definition 2.5 *Sei $I \subset \mathbb{R}$ ein offenes Intervall und $k \in \mathbb{N}_0$. Wir bezeichnen mit $C^k(I)$ den $\mathbb{R}$-Vektorraum der k mal stetig differenzierbaren Funktionen $f : I \to \mathbb{R}$, das heißt*

$$C^k(I) = \{f : I \to \mathbb{R} : f^{(i)} : I \to \mathbb{R} \text{ sind definiert und stetig für } i = 0, 1, \ldots, k\}.$$

Weiter definieren wir $C^\infty(I)$ als den $\mathbb{R}$-Vektorraum der unendlich oft differenzierbaren Funktionen, also ist

$$C^\infty(I) = \bigcap_{k \geq 0} C^k(I).$$

Der Umgang mit C^∞-Funktionen ist besonders angenehm, weil die Klasse im Gegensatz zu den Räumen $C^k(I)$ unter der Bildung von Ableitungen abgeschlossen ist. Es ist klar, dass Polynome unendlich oft differenzierbar sind, ebenso die Exponentialfunktion sowie Cosinus und Sinus. Diese Funktionen sind alle durch konvergente Potenzreihen dargestellt. Wir wollen nun eine Funktion $f \in C^\infty(\mathbb{R})$ konstruieren, die für $x > 0$ strikt positiv und für $x \leq 0$ gleich Null ist. Eine solche Funktion kann in keiner Umgebung von $x = 0$ durch eine Potenzreihe dargestellt werden, denn ihre Nullstellenmenge hat in $x = 0$ einen Häufungspunkt; nach dem Identitätssatz für Potenzreihen, siehe Satz 4.11 in Kapitel 2, müsste f dann die Nullfunktion sein. Wir brauchen zur Konstruktion folgende Aussagen zum Verhalten von e^x für $x \to \pm\infty$.

Satz 2.7 (Wachstum von exp **und** log**)** *Es gelten folgende Aussagen:*

$$\lim_{x\to\infty} x^{-s} e^x = \infty \quad \text{und} \quad \lim_{x\to\infty} x^s e^{-x} = 0 \quad \text{für jedes } s > 0, \tag{2.2}$$

$$\lim_{y\to\infty} y^{-s} \log y = 0 \quad \text{und} \quad \lim_{y\searrow 0} y^s \log y = 0 \quad \text{für jedes } s > 0. \tag{2.3}$$

Beweis: Wir wählen $n \in \mathbb{N}$ mit $n > s$ und erhalten für $x \geq 0$

$$x^{-s} e^x = x^{-s} \sum_{k=0}^{\infty} \frac{x^k}{k!} \geq \frac{x^{n-s}}{n!} \to \infty \quad \text{mit } x \to \infty.$$

Die zweite Aussage in (2.2) ergibt sich durch Übergang zum Kehrwert. Mit der Substitution $x = -s \log y$ folgt weiter für $y \searrow 0$

$$0 = \lim_{x\to\infty} e^{-x} x = \lim_{y \searrow 0} y^s(-s \log y).$$

Setzen wir $y = 1/z$ mit $z \to \infty$, so folgt auch der linke Grenzwert in (2.3). □

Folgerung 2.4 *Die Funktion*

$$f(x) = \begin{cases} e^{-1/x} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$$

ist unendlich oft differenzierbar.

Beweis: Wir zeigen durch Induktion, dass es Polynome p_n gibt mit

$$f^{(n)}(x) = \begin{cases} p_n(\frac{1}{x})\, e^{-1/x} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0. \end{cases}$$

Für $n = 0$ ist das richtig mit $p_0(s) \equiv 1$. Ist die Aussage für ein $n \in \mathbb{N}_0$ gezeigt, so folgt:

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

Für $x \neq 0$ gilt der Induktionsschluss also mit $p_{n+1}(s) = s^2\,(p_n(s) - p_n'(s))$. Zu zeigen bleibt $f^{(n+1)}(0) = 0$. Für den linksseitigen Differenzenquotienten ist das klar, für $x > 0$ berechnen wir mit (2.2)

$$\frac{f^{(n)}(x) - f^{(n)}(0)}{x} = \frac{1}{x}\, p_n\Big(\frac{1}{x}\Big)\, e^{-1/x} \to 0 \qquad \text{mit } x \searrow 0.$$

□

Am Schluß beweisen wir noch eine Verallgemeinerung des Mittelwertsatzes, die vor allem bei der Ermittlung von Grenzwerten gute Dienste leistet.

Satz 2.8 (Zweiter Mittelwertsatz der Differentialrechnung) *Seien $f, g : [a, b] \to \mathbb{R}$ stetig, f und g differenzierbar auf (a, b). Dann existiert ein $\xi \in (a, b)$ mit*

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

Beweis: Setze

$$h(x) := (g(x) - g(a))(f(b) - f(a)) - (f(x) - f(a))(g(b) - g(a))$$

für $x \in [a, b]$ und wende den Satz von Rolle an. □

Als Anwendung haben wir die folgende, oft nützliche Regel.

Satz 2.9 (Regel von de l'Hospital)
Seien $f, g : (a, b] \to \mathbb{R}$ differenziebar. Es gelte

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

$$g(x) \neq 0, \quad g'(x) \neq 0 \qquad \text{für } x \in (a, b],$$

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

Dann ist

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

Beweis: Wir setzen noch $f(a) = g(a) = 0$; dann sind f und g stetige Funktionen auf $[a, b]$. Sei $\epsilon > 0$. Nach Voraussetzung existiert ein $\delta > 0$ mit

$$\left|\frac{f'(y)}{g'(y)} - c\right| < \epsilon \qquad \text{für } y \in (a, a + \delta)$$

Sei jetzt $x \in (a, a + \delta)$. Nach Satz 2.8 existiert ein $\xi \in (a, x)$ mit

$$\left|\frac{f(x)}{g(x)}\right| = \left|\frac{f(\xi)}{g(\xi)}\right|.$$

Es folgt
$$\left|\frac{f(x)}{g(x)} - c\right| = \left|\frac{f(\xi)}{g(\xi)} - c\right| < \epsilon.$$
Domit ist
$$\lim_{x\to a} \frac{f(x)}{g(x)} = c$$
gezeigt. $\square$

Nützlich ist oft auch die folgende Version der Regel von de l'Hospital, bei der Zähler und Nenner des fraglichen Quotienten nicht gegen 0 konvergieren, sondern bestimmt divergieren.

Satz 2.10 *Seien $f, g : (a, b] \to \mathbb{R}$ differenziebar. Es gelte*
$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty,$$
$$g(x) \neq 0, \quad g'(x) \neq 0 \quad \textit{für } x \in (a, b],$$
$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = c.$$
Dann ist
$$\lim_{x\to a} \frac{f(x)}{g(x)} = c.$$
Den Beweis finden Sie in dem Skript von Herrn R. Schneider.

Beispiel 2.2 $\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1$. *(Lemma 2.4 in Kapitel 3.)*

Beispiel 2.3 $\lim_{x\to 0} x^s \log x = 0$ *für* $s > 0$. *((2.3) in Satz 2.7)*
$$-x^s \log x = \frac{\log \frac{1}{x}}{x^{-s}} =: \frac{f(x)}{g(x)},$$
$$f'(x) = -\frac{1}{x}, \quad g'(x) = -sx^{-s-1},$$
$$\frac{f'(x)}{g'(x)} = \frac{1}{s}x^s, \quad \lim_{x\to 0} \frac{f'(x)}{g'(x)} = 0.$$
Damit folgt aus der Regel von de l'Hospital
$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = 0.$$

Kapitel 5

Integralrechnung

1 Das Riemannsche Integral

Das Integral einer nichtnegativen Funktion $f : I = [a,b] \to \mathbb{R}$ ist anschaulich der Flächeninhalt des Gebiets $\{(x,y) : x \in I, 0 < y < f(x)\}$. Allerdings haben wir den Flächeninhalt von Teilmengen des $\mathbb{R}^2$ noch gar nicht definiert, außer vielleicht von einfachen Gebieten wie Rechtecken, so dass die gegebene Beschreibung nicht zur Definition taugen kann. Dennoch lassen wir uns im Folgenden von dieser geometrischen Vorstellung leiten.

Definition 1.1 (Zerlegung) *Eine Zerlegung Z des Intervalls $I = [a,b]$ ist eine geordnete Menge von Punkten $a = x_0 \leq x_1 \leq \ldots \leq x_N = b$. Wir setzen $I_k = [x_{k-1}, x_k]$ sowie $\Delta x_k = x_k - x_{k-1}$ für $k = 1, \ldots, N$, und definieren die Feinheit von Z durch*

$$\Delta(Z) = \max_{1 \leq k \leq N} \Delta x_k. \tag{1.1}$$

Definition 1.2 (Riemannsche Summe) *Sei $f : I = [a,b] \to \mathbb{R}$. Die Riemannsche Summe von f zur Zerlegung Z und den Stützstellen $\xi_k \in I_k$ ist*

$$S_{Z,\xi}(f) = \sum_{k=1}^{N} f(\xi_k)\, \Delta x_k \in \mathbb{R}.$$

Die Riemannsche Summe sollte einen Näherungswert für das noch zu definierende Integral darstellen. Eine konkrete Wahl der Zerlegung und der Stützstellen, zum Beispiel die äquidistante Zerlegung mit Intervallmittelpunkten als Stützstellen, führt auf ein numerisches Verfahren zur Approximation des Integrals. Im allgemeinen ist aber nicht gefordert, dass die Zerlegung äquidistant ist, auch können die Stützstellen beliebig in den I_k gewählt werden. Um zu einer sinnvollen Definition zu gelangen, sollte bei Verfeinerung der Zerlegung der Approximationsfehler kleiner werden. Dies führt auf folgenden Begriff der Integrierbarkeit.

Definition 1.3 (Riemann-Integral) *Eine beschränkte Funktion $f : I = [a,b] \to \mathbb{R}$ heißt (Riemann-)integrierbar mit Integral $S \in \mathbb{R}$, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für jede Zerlegung Z und jede Wahl ξ der Stützstellen gilt:*

$$\Delta(Z) < \delta \quad \Rightarrow \quad |S_{Z,\xi}(f) - S| < \varepsilon.$$

Wir nennen dann S das (bestimmte) Integral von f auf $[a,b]$ und schreiben

$$S = \int_a^b f \in \mathbb{R}.$$

Beispiel 1.1 Die konstante Funktion $f : I = [a,b] \to \mathbb{R}$, $f(x) = c$, ist integrierbar mit

$$\int_a^b f = c\,(b-a).$$

Denn für jede Zerlegung Z und jede Wahl der $\xi_k \in I_k$ gilt

$$S_{Z,\xi}(f) = \sum_{k=1}^{N} c\,\Delta x_k = c\sum_{k=1}^{N}(x_k - x_{k-1}) = c\,(b-a).$$

Beispiel 1.2 Die Dirichletfunktion

$$\chi_{\mathbb{Q}} : [0,1] \to \mathbb{R},\ \chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q}, \\ 0 & \text{für } x \notin \mathbb{Q} \end{cases}$$

ist nicht (Riemann-) integrierbar. In jedem I_k mit $\Delta x_k > 0$ gibt es rationale und irrationale Punkte. Für rationale Stützstellen ist $S_{Z,\xi}(\chi_{\mathbb{Q}}) = 1$, für irrationale dagegen $S_{Z,\xi}(\chi_{\mathbb{Q}}) = 0$.

Definition 1.4 (Supremumsnorm) *Für $f : I \to \mathbb{R}$ setzen wir*

$$\|f\|_I = \sup\{|f(x)| : x \in I\}.$$

Die Menge der beschränkten Funktionen auf I, also $\|f\|_I < \infty$, bezeichnen wir mit $\mathcal{B}(I)$.

Die Supremumsnorm hat folgende Eigenschaften, analog zur Euklidischen Norm:

Positivität: $\|f\|_I \geq 0$ mit Gleichheit genau wenn $f = 0$,

Halblinearität: $\|\lambda f\|_I = |\lambda|\,\|f\|_I$ für $\lambda \in \mathbb{R}$.

Dreiecksungleichung: $\|f+g\|_I \leq \|f\|_I + \|g\|_I$.

Satz 1.1 (Linearität des Integrals) *Die Menge $\mathcal{R}(I)$ der Riemann-integrierbaren Funktionen $f : I \to \mathbb{R}$ ist ein Untervektorraum von $\mathcal{B}(I)$ und das Integral $\mathcal{R}(I) \to \mathbb{R}$, $f \mapsto \int_a^b f$, ist ein lineares Funktional. Es gilt also für $\lambda, \mu \in \mathbb{R}$*

$$\int_a^b (\lambda\, f + \mu\, g) = \lambda \int_a^b f + \mu \int_a^b g.$$

Beweis: Es gilt $S_{Z,\xi}(\lambda f + \mu g) = \lambda S_{Z,\xi}(f) + \mu S_{Z,\xi}(g)$, also folgt für $\Delta(Z)$ hinreichend klein

$$\Big|S_{Z,\xi}(\lambda f + \mu g) - \Big(\lambda \int_a^b f + \mu \int_a^b g\Big)\Big| \leq |\lambda|\,\Big|S_{Z,\xi}(f) - \int_a^b f\Big| + |\mu|\,\Big|S_{Z,\xi}(g) - \int_a^b g\Big| < \varepsilon.$$

□

Um zu zeigen, dass stetige Funktionen integrierbar sind, gehen wir in drei Schritten vor:

a) Stückweise konstante Funktionen (Treppenfunktionen) sind integrierbar.

b) Läßt sich eine Funktion gut durch integrierbare Funktionen approximieren, so ist sie integrierbar.

c) Stetige Funktionen lassen sich gut durch Treppenfunktionen approximieren, und sind damit integrierbar.

Natürlich ist unter anderem noch zu klären, was die gute Approximation eigentlich sein soll. Wir beginnen unser Programm, indem wir zunächst zwei Eigenschaften des Integrals zeigen.

Lemma 1.1 *Seien $f, \tilde{f} : I \to \mathbb{R}$ mit $\tilde{f}(x) = f(x)$ für alle $x \in I \backslash \{p_1, \ldots, p_r\}$. Mit $f \in \mathcal{R}(I)$ ist dann auch $\tilde{f} \in \mathcal{R}(I)$ und es gilt $\int_a^b \tilde{f} = \int_a^b f$.*

Beweis: Wir zeigen die Aussage im Fall eines Ausnahmepunkts $p \in I$; der allgemeine Fall folgt daraus per Induktion. Es gilt für jede Zerlegung Z mit Stützstellen ξ_k

$$S_{Z,\xi}(\tilde{f}) - S_{Z,\xi}(f) = \sum_{k=1}^{N} (\tilde{f}(\xi_k) - f(\xi_k))\,\Delta x_k = \sum_{\{k:\xi_k = a\}} (\tilde{f}(p) - f(p))\,\Delta x_k.$$

Es ist $\xi_k = p$ höchstens für zwei k mit $\Delta x_k > 0$. Damit schätzen wir ab

$$\begin{aligned} \left|S_{Z,\xi}(\tilde{f}) - \int_a^b f\right| &\leq \left|S_{Z,\xi}(\tilde{f}) - S_{Z,\xi}(f)\right| + \left|S_{Z,\xi}(f) - \int_a^b f\right| \\ &\leq 2\left|\tilde{f}(p) - f(p)\right| \Delta(Z) + \left|S_{Z,\xi}(f) - \int_a^b f\right|, \end{aligned}$$

und die rechte Seite geht gegen Null mit $\Delta(Z) \to 0$. □

Wie Beispiel 1.2 zeigt, ist Lemma 1.1 nicht richtig für eine abzählbare Ausnahmemenge.

Lemma 1.2 *Sei $I = I_1 \cup \ldots \cup I_n$ eine Zerlegung von $I = [a, b]$ in Intervalle $I_k = [a_{k-1}, a_k]$. Ist $f : I \to \mathbb{R}$ auf jedem der Teilintervalle I_k integrierbar, so folgt $f \in \mathcal{R}(I)$ und*

$$\int_a^b f = \sum_{k=1}^{n} \int_{a_{k-1}}^{a_k} f.$$

Beweis: Es reicht den Fall $n = 2$ zu betrachten. Sei $I = I' \cup I''$ mit $I' = [a, p]$ und $I'' = [p, b]$, und sei $f : I \to \mathbb{R}$ integrierbar auf I' und I'', insbesondere $\|f\|_I = \max(\|f\|_{I'}, \|f\|_{I''}) < \infty$. Ist Z eine Zerlegung von I mit Punkten $a = x_0 \leq \ldots \leq x_N = b$ sowie Stützstellen $\xi_1, \ldots, \xi_N$, so gilt $x_{r-1} \leq p < x_r$ für ein $r \in \{1, \ldots, N\}$, und wir definieren Zerlegungen Z',Z'' von I', I'' mit Stützstellen ξ', ξ'' wie folgt:

$$\begin{aligned} Z' &= \{a = x_0 \leq \ldots \leq x_{r-1} \leq p\} \quad \xi' = \{\xi_1, \ldots, \xi_{r-1}, p\}, \\ Z'' &= \{p < x_r \leq \ldots \leq x_N = b\} \quad \xi'' = \{p, \xi_{r+1}, \ldots, \xi_N\}. \end{aligned}$$

Offenbar gilt $\Delta(Z'), \Delta(Z'') \leq \Delta(Z)$. Die Bilanz der Riemannschen Summen lautet

$$|S_{Z,\xi}(f) - \big(S_{Z',\xi'}(f) + S_{Z'',\xi''}(f)\big)| = |(f(\xi_r) - f(p))\,\Delta x_r| \leq 2\,\|f\|_I\,\Delta(Z).$$

Es folgt mit der Dreiecksungleichung

$$\Big|S_{Z,\xi}(f) - \Big(\int_a^p f + \int_p^b f\Big)\Big| \le 2\,\|f\|_I\,\Delta(Z) + \Big|S_{Z',\xi'}(f) - \int_a^p f\Big| + \Big|S_{Z'',\xi''}(f) - \int_p^b f\Big|.$$

Die rechte Seite geht mit $\Delta(Z) \to 0$ gegen Null. □

Folgerung 1.1 *Sei $f : I \to \mathbb{R}$ eine Treppenfunktion, d. h. es gibt eine Unterteilung $a = a_0 < a_1 < \ldots < a_n = b$ und $c_i \in \mathbb{R}$ $(i = 1, \ldots, n)$ mit $f(x) = c_i$ für alle $x \in (a_{i-1}, a_i)$. Dann ist f integrierbar und es gilt*

$$\int_a^b f = \sum_{i=1}^n (a_i - a_{i-1}) c_i.$$

Beweis: Nach Lemma 1.1 ist $f : [a_{i-1}, a_i] \to \mathbb{R}$ integrierbar für alle $i = 1, \ldots, n$. Aus Lemma 1.2 folgt die Behauptung. □

Damit ist der erste Schritt unseres Programms erledigt. Der zweite Schritt besteht darin, für einen geeigneten Begriff von Konvergenz $f_k \to f$ folgende Aussage zu zeigen:

$$f_k \to f \text{ mit } f_k \text{ integrierbar} \quad \Rightarrow \quad f \text{ integrierbar und } \int_a^b f = \lim_{k\to\infty} \int_a^b f_k.$$

Welcher Konvergenzbegriff ist dabei zu wählen? Zweifellos ist es naheliegend, es mit der punktweisen Konvergenz zu probieren:

$$f_k \to f \text{ punktweise} \quad \Leftrightarrow \quad f_k(x) \to f(x) \quad \text{für alle } x \in I.$$

Aber wie die folgenden Beispiele zeigen, ist die punktweise Konvergenz zu schwach.

Beispiel 1.3 Sei $q_1, q_2, \ldots$ eine Abzählung der rationalen Zahlen in $[0, 1]$. Definiere

$$\chi_n : [0,1] \to \mathbb{R},\ \chi_n(x) = \begin{cases} 1 & \text{für } x \in \{q_1, \ldots, q_n\} \\ 0 & \text{sonst .} \end{cases}$$

Die Folge χ_n konvergiert punktweise gegen die Funktion $\chi_{\mathbb{Q}}$, die nach Beispiel 1.2 nicht integrierbar ist.

Beispiel 1.4 Betrachte die Treppenfunktionen

$$f_k : [0,1] \to \mathbb{R},\ f_k(x) = \begin{cases} k & \text{für } 0 < x < \frac{1}{k} \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $\lim_{k\to\infty} f_k(x) = 0$ für alle $x \in [0, 1]$, denn

$$f_k(x) = 0 \quad \begin{cases} \text{für alle } k, & \text{falls } x = 0 \\ \text{für } k \ge \frac{1}{x}, & \text{falls } x > 0. \end{cases}$$

Also konvergiert f_k punktweise gegen $f \equiv 0$. Aber es ist

$$0 = \int_0^1 f \ne \lim_{k\to\infty} \int_0^1 f_k = \lim_{k\to\infty} \frac{1}{k} \cdot k = 1.$$

Satz 1.2 *Für $f \in \mathcal{R}(I)$ gilt die Ungleichung*

$$\Big|\int_a^b f\Big| \leq |b-a| \, \|f\|_I.$$

BEWEIS: Für jede Zerlegung Z mit Stützstellen ξ_k gilt mit der Dreiecksungleichung

$$|S_{Z,\xi}(f)| \leq \sum_{k=1}^{N} |f(\xi_k)| \, \Delta x_k \leq \|f\|_I \sum_{k=1}^{N} \Delta x_k = |b-a| \, \|f\|_I. \tag{1.2}$$

Die Abschätzung für das Integral folgt. □

Für $f, f_k \in \mathcal{R}(I)$ haben wir nun

$$\Big|\int_a^b f_k - \int_a^b f\Big| = \Big|\int_a^b (f_k - f)\Big| \leq |b-a| \, \|f_k - f\|_I,$$

so dass aus $\|f_k - f\|_I \to 0$ auch die Konvergenz der Integrale folgt. Dies motiviert den folgenden Konvergenzbegriff.

Definition 1.5 (Gleichmäßige Konvergenz) *Die Folge von Funktionen $f_k : I \to \mathbb{R}$ konvergiert gleichmäßig $f : I \to \mathbb{R}$, falls gilt:*

$$\|f_k - f\|_I \to 0 \quad \textit{mit } k \to \infty.$$

In Quantorensprache sieht punktweise bzw. gleichmäßige Konvergenz wie folgt aus:

$$\begin{array}{lllll} \forall \varepsilon > 0 & \forall x \in D & \exists N & \forall k > N: & |f_k(x) - f(x)| < \varepsilon \quad \text{(punktweise)}, \\ \forall \varepsilon > 0 & \exists N & \forall x \in D & \forall k > N: & |f_k(x) - f(x)| < \varepsilon \quad \text{(gleichmäßig)}. \end{array}$$

Der Unterschied ist, dass bei punktweiser Konvergenz die Schranke N von $x \in I$ abhängen darf, also $N = N(\varepsilon, x)$, während bei gleichmäßiger Konvergenz die Schranke N für alle x gleich gewählt werden kann. Im Beispiel 1.4 gilt etwa, falls $\varepsilon \leq 1$ und $x > 0$,

$$|f_k(x) - f(x)| < \varepsilon \quad \Rightarrow \quad k \geq \frac{1}{x},$$

so dass $N(\varepsilon, x) \geq 1/x$ sein muss, also nicht unabhängig von x gewählt werden kann. Klar ist, dass gleichmäßige Konvergenz punktweise Konvergenz impliziert. Deshalb kann man in zwei Schritten vorgehen, um eine Funktionenfolge f_k auf gleichmäßige Konvergenz zu prüfen:

(1) Konvergiert die Folge punktweise? Wenn nicht, so erst recht nicht gleichmäßig. Wenn ja, so ist die punktweise Grenzfunktion die einzig mögliche Kandidatin für den gleichmäßigen Grenzwert.

(2) Nun bestimme $\|f_k - f\|_I$ bzw. schätze diese Norm ab. Gilt $\|f_k - f\|_I \to 0$ mit $k \to \infty$, so ist f_k gleichmäßig konvergent gegen f. Wenn nicht, so ist f_k zwar punktweise, aber nicht gleichmäßig konvergent.

Im Beispiel 1.4 gilt $f_k \to f$ mit $f = 0$ punktweise auf $[0,1]$, aber nicht gleichmäßig, denn

$$\sup_{x \in [0,1]} |f_k(x) - f(x)| = k \to \infty \quad \text{mit } k \to \infty.$$

Satz 1.3 (Integral und gleichmäßige Konvergenz) *Konvergiert die Folge $f_k \in \mathcal{R}(I)$ gleichmäßig gegen $f : I \to \mathbb{R}$, also $\|f_k - f\|_I \to 0$, so ist $f \in \mathcal{R}(I)$ und es gilt*

$$\int_a^b f = \lim_{k\to\infty} \int_a^b f_k.$$

Beweis: Die Funktion f ist beschränkt wegen $\|f\|_I \leq \|f - f_k\|_I + \|f_k\|_I < \infty$. Mit $S_k = \int_a^b f_k$ gilt für k, l hinreichend groß nach Satz 1.2

$$|S_k - S_l| \leq |b-a| \, \|f_k - f_l\|_I \leq |b-a| \left(\|f_k - f\|_I + \|f - f_l\|_I \right) < \varepsilon.$$

Wir setzen $S = \lim_{k\to\infty} S_k$ und zeigen, dass f integrierbar ist mit $\int_a^b f = S$. Für jede Zerlegung Z mit Stützstellen ξ_j und jedes $k \in \mathbb{N}$ haben wir mit (1.2)

$$\begin{aligned} |S_{Z,\xi}(f) - S| &\leq |S_{Z,\xi}(f) - S_{Z,\xi}(f_k)| + |S_{Z,\xi}(f_k) - S_k| + |S_k - S| \\ &\leq (b-a) \, \|f - f_k\|_I + |S_{Z,\xi}(f_k) - S_k| + |S_k - S|. \end{aligned}$$

Zu gegbenem $\varepsilon > 0$ wählen wir erst $k \in \mathbb{N}$ mit $(b-a) \, \|f - f_k\|_I < \varepsilon/3$ und $|S_k - S| < \varepsilon/3$. Für $\Delta(Z) < \delta$ ist dann auch $|S_{Z,\xi}(f_k) - S_k| < \varepsilon/3$, da f_k integrierbar ist. Damit ist der Satz bewiesen. □

Damit ist auch der zweite Schritt unseres Programms abgeschlossen. Es bleibt jetzt nachzuweisen, dass stetige Funktionen $f : I = [a,b] \to \mathbb{R}$ gleichmäßig durch Treppenfunktionen approximiert werden können.

Satz 1.4 *Sei $D \subset \mathbb{R}^n$ folgenkompakt und $f : D \to \mathbb{R}$ stetig. Dann ist f gleichmäßig stetig, das heißt zu $\varepsilon > 0$ gibt es ein $\delta > 0$ mit*

$$x, x' \in D, \quad |x - x'| < \delta \quad \Rightarrow \quad |f(x) - f(x')| < \varepsilon. \tag{1.3}$$

Beweis. Andernfalls gibt es für ein $\varepsilon > 0$ Punkte $x_n, x_n' \in D$, so dass gilt:

$$|x_n - x_n'| \to 0, \text{ aber } |f(x_n) - f(x_n')| \geq \varepsilon.$$

Da D kompakt, konvergiert die Folge x_n nach Übergang zu einer Teilfolge gegen ein $x_0 \in D$. Offenbar gilt dann auch $\lim_{n\to\infty} x_n' = x_0$. Da f stetig, ergibt sich der Widerspruch

$$\varepsilon \leq \lim_{n\to\infty} |f(x_n) - f(x_n')| = |f(x_0) - f(x_0)| = 0,$$

□

Beispiel 1.5 Ist $f : D \to \mathbb{R}$ stetig, aber D nicht kompakt, so muss f nicht gleichmäßig stetig sein. Betrachte zum Beispiel $f : (0, \frac{1}{\pi}] \to \mathbb{R}$, $f(x) = \sin \frac{1}{x}$. Mit $x_n = (n\pi)^{-1}$, $x_n' = (n\pi + \pi/2)^{-1}$ gilt $x_n, x_n' \to 0$, aber $|f(x_n) - f(x_n')| = 1$ für alle n.

Satz 1.5 *Sei $I = [a,b]$ ein kompaktes Intervall. Dann ist jede stetige Funktion $f : I \to \mathbb{R}$ Riemann-integrierbar.*

Beweis: Wir konstruieren zu $\varepsilon > 0$ eine Treppenfunktion $\varphi : I \to \mathbb{R}$ mit $\|\varphi - f\|_I \leq \varepsilon$. Die Behauptung ergibt sich dann aus Folgerung 1.1 und Satz 1.3. Nach Satz 1.4 gibt es ein $\delta > 0$, so dass gilt:

$$x, x' \in I, \quad |x - x'| < \delta \quad \Rightarrow \quad |f(x) - f(x')| < \varepsilon.$$

Wähle eine beliebige Zerlegung Z mit Feinheit $\Delta(Z) < \delta$ und den Unterteilungspunkten $a = x_0 \leq \ldots \leq x_N = b$, und setze

$$\varphi(x) = \begin{cases} f(x_k) & \text{für } x \in (x_{k-1}, x_k] \text{ mit } k \in \{1, \ldots, N\}, \\ f(x_0) & \text{für } x = x_0. \end{cases}$$

Ist $x \in (x_{k-1}, x_k]$, so gilt $|x - x_k| < \delta$ und somit $|\varphi(x) - f(x)| = |f(x_k) - f(x)| < \varepsilon$ nach Wahl von δ. Da $\varphi(x_0) = f(x_0)$, folgt insgesamt $\|\varphi - f\|_I \leq \varepsilon$. □

Es ist nützlich, die Integrierbarkeit auch für stückweise stetige Funktionen $f : [a, b] \to \mathbb{R}$ zu haben, das heißt es gibt eine Zerlegung $a = a_0 < \ldots < a_N = b$, so dass f auf jedem Teilintervall $[a_{k-1}, a_k]$ nach eventueller Abänderung in den Endpunkten a_{k-1} und a_k stetig ist. Diese Verallgemeinerung folgt natürlich sofort aus Satz 1.5 und Lemma 1.2.

Während die bisherige Darstellung des Riemannintegrals ohne Änderungen im vektorwertigen Fall zutrifft, spielt bei den folgenden Aussagen die Anordnung von $\mathbb{R}$ eine Rolle.

Satz 1.6 (Monotonie des Integrals) *Sind $f, g \in \mathcal{R}(I)$, so gilt:*

$$f \leq g \quad \Rightarrow \quad \int_a^b f \leq \int_a^b g.$$

Insbesondere gilt $\int_a^b f \leq \int_a^b |f|$, falls f, $|f| \in \mathcal{R}(I)$.

Beweis: Für jede Zerlegung Z mit Stützstellen ξ_k gilt

$$S_{Z,\xi}(f) = \sum_{k=1}^{N} f(\xi_k)\Delta x_k \leq \sum_{k=1}^{N} g(\xi_k)\Delta x_k = S_{Z,\xi}(g).$$

□

Folgerung 1.2 (Mittelwertsatz der Integralrechnung) *Seien $f, \varphi : [a, b] \to \mathbb{R}$ stetig und $\varphi \geq 0$. Dann gibt es ein $\xi \in [a, b]$ mit*

$$\int_a^b f\varphi = f(\xi) \int_a^b \varphi.$$

Im Spezialfall $\varphi = 1$ folgt $\int_a^b f = f(\xi)(b - a)$.

Beweis: Wir können annehmen, dass $\int_a^b \varphi = 1$. Setze $m = \inf_{x \in I} f(x)$ und $M = \sup_{x \in I} f(x)$. Dann gilt $m\varphi \leq f\varphi \leq M\varphi$, also

$$m = \int_a^b m\varphi \leq \int_a^b f\varphi \leq \int_a^b M\varphi = M.$$

Nach dem Zwischenwertsatz gibt es ein $\xi \in [a,b]$ mit $f(\xi) = \int_a^b f\varphi$. $\square$

Der Vollständigkeit halber wollen wir noch eine alternative Definition des Riemann-Integrals erklären. Sei $f : [a,b] \to \mathbb{R}$ beschränkt und Z eine Zerlegung von $[a,b]$ in Teilintervalle I_k der Länge Δx_k. Dann sind Ober- und Untersumme von f bzgl. Z definiert durch

$$\overline{S}_Z(f) = \sum_{k=1}^N (\sup_{I_k} f)\, \Delta x_k \quad \text{und} \quad \underline{S}_Z(f) = \sum_{k=1}^N (\inf_{I_k} f)\, \Delta x_k.$$

Nach Definition gilt $\underline{S}_Z(f) \leq \overline{S}_Z(f)$. Nehmen wir zu Z den Unterteilungspunkt $\xi \in [x_{k-1}, x_k]$ hinzu, so folgt mit $I'_k = [x_{k-1}, \xi]$ und $I''_k = [\xi, x_k]$

$$\begin{aligned} \underline{S}_{Z\cup\{\xi\}}(f) - \underline{S}_Z(f) &= (\inf_{I'_k} f)\,(\xi - x_{k-1}) + (\inf_{I''_k} f)\,(x_k - \xi) - (\inf_{I_k} f)\,(x_k - x_{k-1}) \\ &= (\inf_{I'_k} f - \inf_{I_k} f)\,(\xi - x_{k-1}) + (\inf_{I''_k} f - \inf_{I_k} f)\,(x_k - \xi). \end{aligned}$$

Mit $\inf_{I'_k} f, \inf_{I''_k} f, -\inf_{I_k} f \leq \|f\|_I$ sowie $0 \leq x_k - x_{k-1} \leq \Delta(Z)$ erhalten wir

$$0 \leq \underline{S}_{Z\cup\{\xi\}}(f) - \underline{S}_Z(f) \leq 2\|f\|_I\, \Delta(Z),$$

und analog für die Obersummen

$$0 \geq \overline{S}_{Z\cup\{\xi\}}(f) - \overline{S}_Z(f) \geq -2\|f\|_I\, \Delta(Z).$$

Für beliebige Zerlegungen Z, Z' ergibt sich per Induktion

$$\underline{S}_Z(f) \leq \underline{S}_{Z\cup Z'}(f) \leq \overline{S}_{Z\cup Z'}(f) \leq \overline{S}_{Z'}(f). \tag{1.4}$$

Bezeichnet N die Zahl der Teilintervalle von Z', so folgt ebenfalls induktiv

$$\underline{S}_{Z\cup Z'}(f) - 2N\,\|f\|_I\, \Delta(Z) \leq \underline{S}_Z(f) \leq \overline{S}_Z(f) \leq \overline{S}_{Z\cup Z'}(f) + 2N\|f\|_I\, \Delta(Z). \tag{1.5}$$

Wir definieren nun das Ober- bzw. Unterintegral von f durch

$$\begin{aligned} \overline{S}(f) &= \inf\{\overline{S}_Z(f) : Z \text{ ist Zerlegung von } [a,b]\}, \\ \underline{S}(f) &= \sup\{\underline{S}_Z(f) : Z \text{ ist Zerlegung von } [a,b]\}. \end{aligned}$$

Aus (1.4) folgt $\underline{S}(f) \leq \overline{S}(f)$.

Satz 1.7 *Eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann (Riemann-) integrierbar, wenn ihr Ober- und Unterintegral übereinstimmen, und es gilt dann*

$$\int_a^b f = \underline{S}(f) = \overline{S}(f).$$

Beweis: Ist f Riemannintegrierbar, so folgt für $\Delta(Z) < \delta$

$$\int_a^b f - \varepsilon \leq \inf_\xi S_{Z,\xi}(f) = \underline{S}_Z(f) \leq \underline{S}(f) \leq \overline{S}(f) \leq \overline{S}_Z(f) = \sup_\xi S_{Z,\xi}(f) \leq \int_a^b f + \varepsilon.$$

Umgekehrt sei Z' eine Zerlegung mit $\underline{S}_{Z'}(f) > \underline{S}(f) - \varepsilon/2$, und N sei die Zahl der Teilintervalle von Z'. Für jede Zerlegung Z gilt $\underline{S}_{Z\cup Z'}(f) \geq \underline{S}_{Z'}(f)$, also folgt mit (1.5)

$$S_{Z,\xi}(f) \geq \underline{S}_Z(f) \geq \underline{S}_{Z\cup Z'}(f) - 2N\|f\|_I\, \Delta(Z) > \underline{S}(f) - \varepsilon/2 - 2N\|f\|_I\, \Delta(Z).$$

Für $\Delta(Z) < \delta$ folgt $S_{Z,\xi}(f) > \underline{S}(f) - \varepsilon$. Die Abschätzung nach oben ist analog. $\square$

2 Ableitung und Integral

Wir kommen nun zu dem zentralen, von Newton und Leibniz studierten Zusammenhang zwischen Differentiation und Integration.

Definition 2.1 *Sei $f : (a,b) \to \mathbb{R}$. Eine differenzierbare Funktion $F : (a,b) \to \mathbb{R}$ heißt Stammfunktion von f, wenn gilt:*

$$F' = f \quad \Leftrightarrow \quad F'(x) = f(x) \quad \text{für alle } x \in (a,b).$$

Satz 2.1 *Ist F eine Stammfunktion von f auf (a,b), so ist jede Stammfunktion von f auf (a,b) von der Form $F + c$, für eine Konstante $c \in \mathbb{R}$.*

Beweis: Sei G auch Stammfunktion von f auf (a,b). Es folgt

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

Nach Folgerung 2.1 in Kapitel 4 gibt es eine Konstante $c \in \mathbb{R}$ mit $G - F = c$, also $G = F + c$.
□

Die Gleichung $F' = f$ ist ein elementares Beispiel für eine Differentialgleichung. Folgerung 2.1 sagt aus, dass eine Lösung der Gleichung bis auf eine additive Konstante $c \in \mathbb{R}$ eindeutig bestimmt ist. Es schließt sich hier unmittelbar die Frage nach der Existenz an:

Für welche f ist die Differentialgleichung $F' = f$ auf dem Intervall (a,b) lösbar?

Um die Frage zu beantworten, müssen wir die Definition des Integrals noch etwas erweitern, indem wir für eine Riemann-integrierbare Funktion $f : [a,b] \to \mathbb{R}$ setzen:

$$\int_b^a f := -\int_a^b f \quad \text{und} \quad \int_a^a f = 0.$$

Es folgt dann für beliebige $a, b, c \in \mathbb{R}$, sofern f auf allen Intervallen Riemann-integrierbar ist,

$$\int_a^b f + \int_b^c f = \int_a^c f. \tag{2.1}$$

Für $a \leq b \leq c$ gilt dies nach Lemma 1.2, und allgemein folgt (2.1) dann durch Vertauschung von a, b und c. Weiter verwenden wir im Folgenden die Notation

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

Dies ist unter anderem dann nützlich, wenn die Funktion f außer von x noch von weiteren Variablen $y, z, \ldots$ abhängt, es wird nämlich spezifiziert, bezüglich welcher Variablen integriert werden soll. Außerdem erionnert die Notation an die Riemannschen Summen, mit denen das Integral definiert wurde.

Satz 2.2 (Hauptsatz der Differential- und Integralrechnung) *Sei $f : I = [a,b] \to \mathbb{R}$ stetig. Dann ist für jedes $x_0 \in I$ die Funktion*

$$F : I \to \mathbb{R}, \quad F(x) = \int_{x_0}^x f(\xi)\,d\xi$$

eine Stammfunktion von f, das heißt es gilt $F'(x) = f(x)$ für alle $x \in I$.

Bemerkung. In den Endpunkten des Intervalls ist dies im Sinne der einseitigen Ableitungen $F'_+(a) = f(a)$ bzw. $F'_-(b) = f(b)$ zu verstehen.

BEWEIS: Die Funktion F ist wohldefiniert nach Satz 1.5. Wir berechnen

$$\begin{aligned}\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| &= \frac{1}{|h|}\left|\int_{x_0}^{x+h} f(\xi)\,d\xi - \int_{x_0}^{x} f(\xi)\,d\xi - hf(x)\right| \\ &= \frac{1}{|h|}\left|\int_{x}^{x+h} \big(f(\xi)-f(x)\big)\,d\xi\right| \\ &\le \frac{1}{|h|}\,|h| \sup_{|\xi-x|\le|h|} |f(\xi)-f(x)| \quad \text{(Satz 1.2).}\end{aligned}$$

Da f im Punkt x stetig ist, geht die rechte Seite mit $h \to 0$ gegen Null. □

Folgerung 2.1 *Sei $f \in C^0(I)$ mit $I = [a,b]$, und $F : [a,b] \to \mathbb{R}$ sei eine Stammfunktion von f auf I, das heißt $F'(x) = f(x)$ für alle $x \in [a,b]$*[1]*. Dann gilt für jedes $x_0 \in I$*

$$F(x) = F(x_0) + \int_{x_0}^{x} f(\xi)\,d\xi. \tag{2.2}$$

BEWEIS: Nach Satz 2.1 und Satz 2.2 gibt es ein $c \in \mathbb{R}$ mit

$$F(x) = c + \int_{x_0}^{x} f(\xi)\,d\xi \quad \text{für alle } x \in (a,b).$$

Weiter ist jede Stammfunktion F von f auf $[a,b]$ stetig in a und b nach Kapitel 4, Satz 1.3. Durch Grenzübergang $x \searrow a$ bzw. $x \nearrow b$ folgt die Gleichung für alle $x \in [a,b]$. Setze nun $x = x_0$ und erhalte $F(x_0) = c$. □

Folgerung 2.2 (Berechnung von bestimmten Integralen mit einer Stammfunktion) *Die Funktion $F : I = [a,b] \to \mathbb{R}$ sei Stammfunktion von $f \in C^0(I)$ auf I. Dann gilt*

$$\int_a^b f(x)\,dx = F(b) - F(a) =: \big[F(x)\big]_{x=a}^{x=b}$$

BEWEIS: Folgt aus (2.2) mit $x_0 = a$, $x = b$. □

Über den Hauptsatz bzw. Folgerung 2.2 lassen sich Differentiationsregeln aus Kapitel 4.1 in Integrationsregeln übersetzen. Dies wird im Folgenden durchgeführt.

Satz 2.3 (Partielle Integration) *Seien $f, g \in C^1(I)$ mit $I = [a,b]$. Dann gilt*

$$\int_a^b fg' = \big[f(x)\,g(x)\big]_{x=a}^{x=b} - \int_a^b f'g.$$

BEWEIS: Es gilt nach der Produktregel

$$(fg)' = f'g + fg' \quad \text{auf } I = [a,b].$$

Folgerung 2.2 liefert die Behauptung. □

[1] In den Intervallgrenzen bedeutet das $F'_+(a) = f(a)$ bzw. $F'_-(b) = f(b)$.

Satz 2.4 (Substitutions- oder Transformationsregel) *Sei $I = [a,b]$, $I^* = [\alpha,\beta]$ und $\varphi \in C^1(I)$ mit $\varphi(I) \subset I^*$. Dann gilt für $f \in C^0(I^*)$*

$$\int_{\varphi(a)}^{\varphi(b)} f(y)\,dy = \int_a^b f\big(\varphi(x)\big)\varphi'(x)\,dx.$$

Beweis: Wähle nach Satz 2.2 eine Stammfunktion $F \in C^1(I^*)$ von f. Dann gilt

$$\begin{aligned}
\int_{\varphi(a)}^{\varphi(b)} f(y)\,dy &= \big[F(y)\big]_{x=\varphi(a)}^{x=\varphi(b)} \quad \text{(Folgerung 2.2)} \\
&= \big[F\big(\varphi(x)\big)\big]_{x=a}^{x=b} \\
&= \int_a^b (F \circ \varphi)'(x)\,dx \quad \text{(Folgerung 2.2)} \\
&= \int_a^b F'\big(\varphi(x)\big)\varphi'(x)\,dx \qquad \text{(Kettenregel)} \\
&= \int_a^b f\big(\varphi(x)\big)\,\varphi'(x)\,dx.
\end{aligned}$$

□

Für die Anwendung der Substitutionsregel ist folgendes *Kochrezept* nützlich: bei einem gegebenen Integral $\int_a^b f(y)\,dy$ möchten wir $y = y(x)$ substituieren. Dazu berechnen wir

$$y = y(x) \quad \Rightarrow \quad dy = y'(x)\,dx.$$

Für die Intervallgrenzen bestimmen wir durch Auflösen nach x die Umkehrfunktion

$$x = x(y) \quad \Rightarrow \quad a = x(\alpha),\, b = x(\beta).$$

Damit gilt

$$\int_\alpha^\beta f(y)\,dy = \int_a^b f\big(y(x)\big)\,y'(x)\,dx.$$

Im folgenden zeigen wir an einigen Beispielen, wie die Integrationsregeln angewandt werden. Zunächst erhalten wir direkte Integrationsformeln immer dann, wenn die Stammfunktion bekannt ist. Hier sind einige Beispiele.

Beispiel 2.1

$$\begin{aligned}
\int_a^b x^\alpha\,dx &= \left[\frac{x^{\alpha+1}}{\alpha+1}\right]_{x=a}^{x=b} \quad (\alpha \in \mathbb{R}\backslash\{-1\},\, a,b > 0), \\
\int_a^b \frac{dx}{x} &= \big[\log x\big]_{x=a}^{x=b} \quad (a,b > 0), \\
\int_a^b \frac{dx}{\sqrt{1-x^2}} &= \big[\arcsin x\big]_{x=a}^{x=b}, \quad (-1 < a,b < 1).
\end{aligned}$$

Ein Beispiel für die Anwendung der partiellen Integration ist

Beispiel 2.2

$$\int_1^x \log u\,du = \int_1^x 1\cdot \log u\,du = \big[u\log u\big]_{u=1}^{u=x} - \int_1^x u\frac{1}{u}\,du = x\log x - (x-1).$$

Eine schöne Anwendung von Satz 2.3 ist das

Beispiel 2.3 (Wallis-Produkt) Wir berechnen hier $A_n = \int_0^{\pi/2} \sin^n x\,dx$. Offenbar gilt

$$A_0 = \pi/2 \quad \text{und } A_1 = \big[-\cos x\big]_{x=0}^{x=\pi/2} = 1.$$

Für $n \geq 1$ leiten wir durch partielle Integration eine Rekursionsformel her:

$$\begin{aligned}
A_{n+1} &= \int_0^{\pi/2} \sin x\, \sin^n x\,dx \\
&= \big[\underbrace{-\cos x\, \sin^n x}_{=0}\big]_{x=0}^{x=\pi/2} + n\int_0^{\pi/2} \cos^2 x\, \sin^{n-1} x\,dx \\
&= n\int_0^{\pi/2} \sin^{n-1} x\,dx - n\int_0^{\pi/2} \sin^{n+1} x\,dx,
\end{aligned}$$

wobei wir $\cos^2 x = 1 - \sin^2 x$ benutzt haben. Es folgt

$$A_{n+1} = \frac{n}{n+1}A_{n-1} \quad (n \geq 1).$$

Durch Induktion erhalten wir

$$\begin{aligned}
A_{2n} &= \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdot\ldots\cdot\frac{1}{2}\cdot A_0 = \left(\prod_{k=1}^{n}\frac{2k-1}{2k}\right)\frac{\pi}{2}, \\
A_{2n+1} &= \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdot\ldots\cdot\frac{2}{3}\cdot A_1 = \left(\prod_{k=1}^{n}\frac{2k}{2k+1}\right) 1.
\end{aligned}$$

Es folgt weiter

$$\frac{A_{2n}}{A_{2n+1}} = \frac{\pi}{2}\cdot\prod_{k=1}^{n}\frac{4k^2-1}{4k^2}.$$

Nun gilt $A_{n+1} = \int_0^{\pi/2}\sin^{n+1} x\,dx \leq \int_0^{\pi/2}\sin^n x\,dx = A_n$, und es folgt

$$1 \leq \frac{A_n}{A_{n+1}} \leq \frac{A_{n-1}}{A_{n+1}} = \frac{n+1}{n} \longrightarrow 1.$$

Daraus ergibt sich die Produktdarstellung von Wallis

$$\frac{\pi}{2} = \prod_{k=1}^{\infty}\frac{4k^2}{4k^2-1} = \frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdot\ldots$$

Als nächstes behandeln wir Beispiele zur Substitutionsregel.

Beispiel 2.4 (Lineare Parameterwechsel) Mit der Substitution $y = (x - x_0)/m$ haben wir $dy = 1/m\, dx$ und $x = x_0 + my$, also als neue Grenzen $a = x_0 + m\alpha$ und $b = x_0 + m\beta$. Es folgt

$$\int_\alpha^\beta f(y)\, dy = \frac{1}{m} \int_{x_0+m\alpha}^{x_0+m\beta} f\Big(\frac{x - x_0}{m}\Big)\, dx.$$

Beispiel 2.5 (Integration von Ableitungen)

$$\begin{aligned}
\int_a^b \frac{f'(x)}{f(x)}\, dx &= \int_a^b (\log f)'(x)\, dx = \big[\log f(x)\big]_{x=a}^{x=b} \quad (f > 0),\\
\int_a^b u\sqrt{1+u^2}\, du &= \frac{1}{3}\int_a^b \big((1+u^2)^{\frac{3}{2}}\big)'\, du = \Big[\frac{1}{3}(1+u^2)^{\frac{3}{2}}\Big]_{u=a}^{u=b}.\\
\int_a^b F'\big(f(x)\big)f'(x)\, dx &= \big[F\big(f(x)\big)\big]_{x=a}^{x=b}.
\end{aligned}$$

Beispiel 2.6 (Flächeninhalt unter Hyperbel) Zu berechnen ist das Integral

$$A(x) = \int_1^x \sqrt{u^2 - 1}\, du \qquad \text{für } x \geq 1.$$

Wir substituieren $u = \cosh t$ und erhalten $du = \sinh t\, dt$, $t = \text{Arcosh } u$, also

$$\begin{aligned}
A(x) &= \int_0^{\text{Arcosh } x} \sinh^2 t\, dt\\
&= \frac{1}{4}\int_0^{\text{Arcosh } x} (e^{2t} + e^{-2t} - 2)\, dt\\
&= \Big[\frac{1}{8}(e^{2t} - e^{-2t})\Big]_{t=0}^{t = \text{Arcosh } x} - \frac{1}{2}\text{Arcosh } x\\
&= \frac{1}{2}\Bigg(\Big[\frac{e^t + e^{-t}}{2}\,\frac{e^t - e^{-t}}{2}\Big]_{t=0}^{t = \text{Arcosh } x} - \text{Arcosh } x\Bigg)\\
&= \frac{1}{2}\big(x\sqrt{x^2-1} - \text{Arcosh } x\big).
\end{aligned}$$

Für rationale Funktionen, also Quotienten von Polynomen, hat man ein spezielles Integrationsverfahren, die Partialbruchzerlegung, die wir hier nur an einem Beispiel vorführen:

Beispiel 2.7 (Partialbruchzerlegung) Um das Integral $\int_{-1/2}^{1/2} \frac{dx}{1-x^2}$ zu berechnen, machen wir den Ansatz

$$\frac{1}{1-x^2} = \frac{1}{(1+x)(1-x)} \stackrel{!}{=} \frac{A}{1-x} + \frac{B}{1+x} = \frac{(A-B)x + (A+B)}{1-x^2}.$$

Der Koeffizientenvergleich ergibt $A = B = 1/2$, also folgt

$$\int_{-1/2}^{1/2} \frac{dx}{1-x^2} = \int_{-1/2}^{1/2} \Big(\frac{1/2}{1-x} + \frac{1/2}{1+x}\Big)\, dx = \frac{1}{2}\Big[\log\frac{1+x}{1-x}\Big]_{x=-\frac{1}{2}}^{x=\frac{1}{2}} = \frac{1}{2}(\log 3 - \log 1/3) = \log 3.$$

Bei der Definition des Riemannschen Integrals ist das Definitionsintervall $I = [a, b]$ nach Voraussetzung kompakt. Wir wollen ganz kurz erläutern, wie auch unendliche Integrationsintervalle $-\infty \leq a < b \leq \infty$ und in den Intervallgrenzen unbeschränkte Funktionen im Rahmen des Riemann-Integrals behandelt werden können.

Definition 2.2 (uneigentliches Riemann-Integral) *Sei $I = [a, b)$ mit $a < b \leq \infty$. Die Funktion $f : I \to \mathbb{R}$ sei Riemann-integrierbar auf $[a, b']$ für alle $b' < b$. Falls $\lim_{x \nearrow b} \int_a^x f(\xi)\, d\xi$ existiert, so heißt das Integral $\int_a^b f(\xi)\, d\xi$ konvergent (oder existent) und wir setzen*

$$\int_a^b f(\xi)\, d\xi = \lim_{x \nearrow b} \int_a^x f(\xi)\, d\xi.$$

Die Konvergenzaussagen für das uneigentliche Integral sind analog zu den Konvergenzkriterien für Reihen. Das folgende Lemma entspricht dabei dem Cauchykriterium.

Lemma 2.1 *In der Situation von Definition 2.2 ist $\int_a^b f$ genau dann konvergent, wenn es zu jedem $\varepsilon > 0$ ein $b' < b$ gibt mit*

$$\Big| \int_{x_1}^{x_2} f \Big| < \varepsilon \quad \text{für alle } x_1, x_2 > b'.$$

Beweis: Setze $F(x) = \int_a^x f$ für $x \in [a, b)$. Existiert das uneigentliche Integral, das heißt $F(x) \to S$ mit $x \nearrow b$, so gibt es ein $b' < b$ mit $|F(x) - S| < \varepsilon/2$ für $x > b'$ und es folgt

$$|F(x_1) - F(x_2)| \leq |F(x_1) - S| + |F(x_2) - S| < \varepsilon \quad \text{für } x_{1,2} > b'.$$

Umgekehrt wählen wir eine Folge $x_k \nearrow b$. Nach Voraussetzung ist dann $F(x_k)$ Cauchyfolge, das heißt $S = \lim_{k \to \infty} F(x_k)$ existiert. Aber dann folgt sogar $\lim_{x \nearrow b} F(x) = S$. $\square$

Hier sind einige Beispiele von uneigentlichen Riemann-Integralen.

Beispiel 2.8

$$\begin{aligned}
\int_1^\infty x^\alpha\, dx &= \begin{cases} \lim_{R \nearrow \infty} \left[\dfrac{x^{\alpha+1}}{\alpha+1}\right]_{x=1}^{x=R} = -\dfrac{1}{\alpha+1} & \text{für } \alpha < -1, \\ \text{divergent} & \text{für } \alpha \geq -1. \end{cases} \\
\int_0^1 x^\alpha\, dx &= \begin{cases} \text{divergent} & \text{für } \alpha \leq -1, \\ \lim_{\varepsilon \searrow 0} \left[\dfrac{x^{\alpha+1}}{\alpha+1}\right]_{x=\varepsilon}^{x=1} = \dfrac{1}{\alpha+1} & \text{für } \alpha > -1. \end{cases} \\
\int_0^1 \frac{dx}{\sqrt{1-x^2}} &= \lim_{x \nearrow 1} \int_0^x \frac{d\xi}{\sqrt{1-\xi^2}} = \lim_{x \nearrow 1} \arcsin x = \pi/2.
\end{aligned}$$

In den vorangegangenen Beispielen sind die Integranden positiv. Das uneigentliche Integral $\int_a^b f$ heißt absolut konvergent, wenn $\int_a^b |f|$ konvergiert. Aus Lemma 2.1 folgt sofort, dass ein absolut konvergentes Integral konvergiert. Wie bei Reihen impliziert die Konvergenz aber nicht umgekehrt die absolute Konvergenz. Ein simples Beispiel ist das uneigentliche Integral $\int_1^\infty f(x)\, dx$, wobei

$$f(x) = \frac{(-1)^{k-1}}{k} \text{ für } k = [x].$$

Für die Existenz des uneigentlichen Integrals auf einem beidseitig offenen Intervall (a,b) wählt man einen Zwischenpunkt $c \in (a,b)$ und verlangt die Existenz der Integrale auf $(a,c]$ und $[c,b)$. Das Integral über (a,b) ergibt sich dann als Summe. Es ist leicht zu sehen, dass diese Definition nicht von der Wahl des Zwischenpunkts c abhängt. Ein Beispiel ist

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \pi.$$

3 Vertauschungssätze für konvergente Folgen von Funktionen

In diesem Abschnitt betrachten wir Folgen $f_n : I \to \mathbb{R}$ von Funktionen, die punktweise gegen eine Grenzfunktion $f : I \to \mathbb{R}$ konvergieren, das heißt

$$f(x) = \lim_{n\to\infty} f_n(x) \quad \text{für alle } x \in I.$$

Wir interessieren uns dafür, unter welchen Voraussetzungen sich die Stetigkeit bzw. Differenzierbarkeit der Funktionen f_n auf die Grenzfunktion f überträgt. Ein wichtiger Fall ist die Frage, ob eine reelle Potenzreihe $P(x) = \sum_{k=0}^{\infty} a_k x^k$ innerhalb des Konvergenzintervalls $(-R,R)$ eine differenzierbare Funktion darstellt. Hier ist klar, dass die approximierenden Polynome $P_n(x) = \sum_{k=0}^{n} a_k x^k$ differenzierbar sind mit Ableitung

$$P_n'(x) = \sum_{k=1}^{n} k a_k x^{k-1} = \sum_{k=0}^{n-1} (k+1) a_{k+1} x^k.$$

Wir werden zeigen, dass $P(x)$ auf $(-R,R)$ differenzierbar ist und dass gilt

$$P'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} x^k.$$

Dies bedeutet, dass die Potenzreihe gliedweise differenziert werden kann, das heißt die Ableitung kann mit dem Grenzwert $n \to \infty$ der Reihe vertauscht werden. Bei der Exponentialfunktion konnten wir die Frage der Stetigkeit und Differenzierbarkeit in $x \in \mathbb{R}$ mithilfe der Funktionalgleichung auf den Fall $x = 0$ reduzieren, wo wir dann die Abschätzung aus Satz 4.7 zur Verfügung hatten. Für eine allgemeine Potenzreihe haben wir kein Äquivalent der Funktionalgleichung und brauchen deshalb allgemeinere Resultate.

Die punktweise Konvergenz ist im allgemeinen nicht einmal ausreichend für die Stetigkeit der Grenzfunktion – hier zwei typische Beispiele.

Beispiel 3.1 Die Folge $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$, konvergiert mit $n \to \infty$ punktweise gegen die Funktion

$$f(x) = \begin{cases} 0 & \text{für } 0 \le x < 1, \\ 1 & \text{für } x = 1. \end{cases}$$

Die Folge $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = \arctan(nx)$, konvergiert punktweise gegen

$$f(x) = \begin{cases} \pi/2 & \text{für } x > 0 \\ -\pi/2 & \text{für } x < 0 \\ 0 & \text{für } x = 0. \end{cases}$$

Satz 3.1 (Gleichmäßige Konvergenz und Stetigkeit) *Seien $f_n : D \to \mathbb{R}^p$, $n \in \mathbb{N}$, stetige Funktionen auf $D \subset \mathbb{R}^m$, die gleichmäßig gegen $f : D \to \mathbb{R}^p$ konvergieren, das heißt*

$$\|f_n - f\|_D = \sup_{x \in D} |f_n(x) - f(x)| \to 0 \quad \textit{mit } n \to \infty.$$

Dann ist f ebenfalls stetig auf D.

Beweis: Sei $x_0 \in D$ gegeben. Für beliebige $x \in D$, $n \in \mathbb{N}$ gilt die Abschätzung

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &\leq 2\,\|f - f_n\|_D + |f_n(x) - f_n(x_0)|. \end{aligned}$$

Zu $\varepsilon > 0$ wählen wir erst $n \in \mathbb{N}$ mit $\|f - f_n\|_D < \varepsilon/3$, dann $\delta > 0$ mit $|f_n(x) - f_n(x_0)| < \varepsilon/3$ für $|x - x_0| < \delta$. Es folgt $|f(x) - f(x_0)| < 2\varepsilon/3 + \varepsilon/3 = \varepsilon$ für $|x - x_0| < \delta$. □

Folgerung 3.1 *Sei $P(z) = \sum_{k=0}^{\infty} a_k z^k$ eine Potenzreihe mit Konvergenzradius $R > 0$ und n-ten Partialsummen $P_n(z) = \sum_{k=0}^{n} a_k z^k$. Dann konvergiert P_n gleichmäßig gegen P auf jeder Kreisscheibe $B_\varrho(0) = \{z \in \mathbb{C} : |z| < \varrho\}$ mit $\varrho < R$, und P ist stetig auf $B_R(0)$.*

Beweis: Wähle $r \in (\varrho, R)$. Nach der Definition des Konvergenzradius in Satz 4.10 konvergiert $P(z)$ für $z = r$, also ist $a_k r^k$ eine Nullfolge und es gibt ein $M \in [0, \infty)$ mit $|a_k| r^k \leq M$. In dieser Situation liefert Lemma 4.2 die Abschätzung

$$\|P - P_n\|_{B_\varrho(0)} = \sup_{|z| < \varrho} |P(z) - P_n(z)| \leq \frac{M}{1 - \frac{\varrho}{r}} \left(\frac{\varrho}{r}\right)^{n+1} \to 0 \quad \text{mit } n \to \infty.$$

Dies beweist die gleichmäßige Konvergenz auf $B_\varrho(0)$. Nach Satz 3.1 ist P somit stetig auf $B_\varrho(0)$ für alle $\varrho < R$, also auf ganz $B_R(0)$. □

Jetzt kommen wir zur Frage der Differenzierbarkeit der Grenzfunktion.

Satz 3.2 (Vertauschung von Konvergenz und Ableitung) *Sei $f_n \in C^1(I)$ eine Folge von Funktionen auf $I = (a, b)$, die punktweise gegen $f : I \to \mathbb{R}$ konvergiert. Falls die Folge f_n' gleichmäßig gegen $g : I \to \mathbb{R}$ konvergiert, so ist $f \in C^1(I)$ und $f' = g$.*

Beweis: Für $x_0 \in I$ gilt nach Satz 2.2, dem Hauptsatz der Differential- und Integralrechnung,

$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(\xi)\, d\xi \quad \text{für alle } x \in I.$$

Da f_n' gleichmäßig gegen g konvergiert, folgt nach Satz 1.3 mit $n \to \infty$

$$f(x) = f(x_0) + \int_{x_0}^{x} g(\xi)\, d\xi \quad \text{für alle } x \in I.$$

Die Funktion g ist stetig nach Satz 3.1, also folgt aus dem Hauptsatz $f \in C^1(I)$ und $f' = g$. □

Für Funktionen $f : I \to \mathbb{R}^p$ oder $f : I \to \mathbb{C}$ gilt der Satz entsprechend. Dies ergibt sich sofort durch Anwendung auf die einzelnen Koordinatenfunktionen, vgl. Lemma 1.1 in Kapitel 4. Um den Satz auf Potenzreihen anzuwenden, brauchen wir folgende Hilfsaussage.

Lemma 3.1 *Sei $P(z) = \sum_{k=0}^{\infty} a_k z^k$ eine komplexe Potenzreihe mit Konvergenzradius $R \in [0,\infty]$. Dann hat die formal differenzierte Reihe*

$$Q(z) = \sum_{k=1}^{\infty} k a_k z^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} z^k$$

denselben Konvergenzradius R.

Beweis: Es gilt $|a_k z^k| \leq |k a_k z^{k-1}| \cdot |z|$ für $k \geq 1$, also hat $Q(z)$ höchstens den Konvergenzradius R. Zu $r \in (0,R)$ gibt es ein $M \in [0,\infty)$ mit $|a_k| r^k \leq M$, da $P(r)$ konvergiert. Es folgt für alle $z \in B_r(0)$

$$|k a_k z^{k-1}| = k |a_k| r^k \, \frac{|z|^{k-1}}{r^k} \leq \frac{kM|z|^{k-1}}{r^k}.$$

Die rechte Reihe konvergiert aber nach dem Quotientenkriterium, denn

$$\frac{(k+1)M|z|^k}{r^{k+1}} \left(\frac{kM|z|^{k-1}}{r^k} \right)^{-1} = \frac{(k+1)|z|}{kr} \to \frac{|z|}{r} < 1 \quad \text{mit } k \to \infty.$$

Also ist $Q(z)$ für $z \in B_R(0)$ absolut konvergent. □

Satz 3.3 (Differenzierbarkeit von Potenzreihen) *Die Potenzreihe $P(z) = \sum_{k=0}^{\infty} a_k z^k$ mit $a_k \in \mathbb{C}$ habe den Konvergenzradius $R > 0$. Dann ist die Funktion*

$$P : (-R,R) \to \mathbb{C},\ P(x) = \sum_{k=0}^{\infty} a_k x^k$$

differenzierbar, und ihre Ableitung ergibt sich durch gliedweise Differentiation:

$$P'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} x^k \quad \textit{für alle } x \in (-R,R). \tag{3.1}$$

Beweis: Die Funktionen $P_n(x) = \sum_{k=0}^{n} a_k x^k$ konvergieren auf $(-R,R)$ punktweise gegen $P(x)$, und nach Lemma 3.1 und Folgerung 3.1 konvergieren die P_n' lokal gleichmäßig, das heißt gleichmäßig auf jedem Teilintervall $(-\varrho,\varrho) \subset (-R,R)$ mit $0 < \varrho < R$, gegen die stetige Funktion $Q : (-R,R) \to \mathbb{R}$, $Q(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$. Die Behauptung folgt nun aus Satz 3.2. □

Jetzt stehen alle Hilfsmittel zur Verfügung, um einige interessante Potenzreihenentwicklungen herzuleiten.

Beispiel 3.2 Für $x \in (-1,1) \subset \mathbb{R}$ gilt die Reihendarstellung

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k = x - \frac{x^2}{2} + \frac{x^3}{3} - + \ldots.$$

Denn die Reihe $P(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k$ konvergiert für $x \in (-1,1)$ nach dem Quotientenkriterium, also folgt aus Satz 3.3 und der Formel für die geometrische Reihe

$$P'(x) = \sum_{k=0}^{\infty} (-1)^k x^k = \frac{1}{1+x} \qquad \text{für } x \in (-1,1).$$

Wegen $P(0) = 0 = \log 1$ ergibt sich $P(x) = \log(1+x)$.

Beispiel 3.3 Ähnlich zeigen wir für $x \in (-1, 1) \subset \mathbb{R}$ die Reihenentwicklung

$$\arctan x = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - + \ldots .$$

Denn $P(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}$ konvergiert für $x \in (-1, 1)$ nach dem Quotientenkriterium, und Satz 3.3 liefert

$$P'(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} = \frac{1}{1+x^2} \qquad \text{für } x \in (-1, 1).$$

Wegen $P(0) = 0 = \arctan 0$ ergibt sich $P(x) = \arctan x$.

Beispiel 3.4 Die Binomialreihe mit Parameter $\alpha \in \mathbb{R}$ ist definiert durch

$$B_\alpha(z) = \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k.$$

Die Reihe hat Konvergenzradius $R = 1$, siehe Beispiel 4.6 in Kapitel 2. Für $x \in (-1, 1)$ berechnen wir mit Satz 3.3

$$B'_\alpha(x) = \sum_{k=0}^{\infty} (k+1) \binom{\alpha}{k+1} x^k = \sum_{k=0}^{\infty} (k+1) \frac{\alpha - k}{k+1} \binom{\alpha}{k} x^k = \alpha B_\alpha(x) - x B'_\alpha(x),$$

das heißt $B'_\alpha(x) = \dfrac{\alpha}{1+x} B_\alpha(x)$. Es folgt mit $f(x) = (1+x)^{-\alpha}$

$$(f B_\alpha)'(x) = f'(x) B_\alpha(x) + f(x) B'_\alpha(x) = f(x) B_\alpha(x) \Big(-\frac{\alpha}{1+x} + \frac{\alpha}{1+x} \Big) = 0.$$

Wegen $B_\alpha(0) = 1 = f(0)$ ergibt sich die folgende Darstellung (Newton 1665)

$$(1+x)^\alpha = B_\alpha(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \qquad \text{für alle } x \in (-1, 1).$$

In der Physik wird oft die Näherung $(1+x)^\alpha = 1 + \alpha x$ benutzt, für $|x| << 1$.

Satz 3.4 (Abelscher Grenzwertsatz) *Ist die Potenzreihe $P(x) = \sum_{n=0}^{\infty} a_n x^n$ für $x = 1$ konvergent, so gilt*

$$\lim_{x \nearrow 1} P(x) = P(1).$$

Beweis: Nach Voraussetzung hat P Konvergenzradius $R \geq 1$. Wir berechnen mit $P_n(x) = \sum_{k=0}^{n} a_k x^k$ für $0 \leq x < 1$

$$\begin{aligned}
P_n(1) - P_n(x) &= \sum_{k=0}^{n} a_k (1 - x^k) \\
&= (1-x) \sum_{k=0}^{n} a_k \sum_{j=0}^{k-1} x^j \\
&= (1-x) \sum_{j=0}^{n-1} x^j \sum_{k=j+1}^{n} a_k \\
&= (1-x) \sum_{j=0}^{n-1} x^j \big(P_n(1) - P_j(1) \big).
\end{aligned}$$

Zu $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit $|P_n(1) - P_j(1)| < \varepsilon/2$ für $j, n > N$, also gilt

$$(1-x) \sum_{j=N+1}^{n-1} x^j \, |P_n(1) - P_j(1)| \leq (1-x) \, \frac{\varepsilon}{2} \sum_{j=N+1}^{\infty} x^j \leq \frac{\varepsilon}{2}.$$

Es folgt

$$|P_n(1) - P_n(x)| \leq (1-x) \sum_{j=0}^{N} |P_n(1) - P_j(1)| + \frac{\varepsilon}{2},$$

und schließlich mit $n \to \infty$

$$|P(1) - P(x)| \leq (1-x) \sum_{j=0}^{N} |P(1) - P_j(1)| + \frac{\varepsilon}{2} < \varepsilon$$

für x hinreichend nahe bei Eins. $\square$

Beispiel 3.5 Nach dem Leibnizkriterium konvergiert die Reihe $x - \frac{1}{2}x^2 + \frac{1}{3}x^3 \pm \ldots$ auch für $x = 1$, also folgt aus Satz 3.4 (Mercator 1668)

$$\log 2 = \lim_{x \nearrow 1} \log(1+x) = 1 - \frac{1}{2} + \frac{1}{3} - + \ldots.$$

Ebenso konvergiert die Reihe des arctan auch für $x = 1$, und es ergibt sich die Darstellung (Gregory 1671, Leibniz 1674)

$$\frac{\pi}{4} = \arctan 1 = \lim_{x \nearrow 1} \arctan x = 1 - \frac{1}{3} + \frac{1}{5} - + \ldots.$$

4 Die Taylorentwicklung

Als Anwendung führen wir die Taylorentwicklung ein. Die Idee der Taylorentwicklung ist es, eine gegebene Funktion f mit einem Polynom zu vergleichen, das an einer festen Stelle x_0 mit f „von höherer Ordnung" übereinstimmt, das heißt einschließlich einer Reihe von Ableitungen.

Lemma 4.1 *Sei $I = (a,b) \subset \mathbb{R}$, $x_0 \in I$ und $n \in \mathbb{N}_0$. Zu $f \in C^n(I)$ gibt es genau ein Polynom $P : \mathbb{R} \to \mathbb{R}$ vom Grad höchstens n mit $P^{(k)}(x_0) = f^{(k)}(x_0)$ für $k = 0, 1, \ldots, n$, und zwar*

$$P(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j. \tag{4.1}$$

Das Polynom P heißt Taylorpolynom n-ten Grades von f mit Entwicklungspunkt x_0 und wird mit $P^n_{x_0} f$ bezeichnet.

Beweis: Es gilt $\left(\frac{d}{dx}\right)^k (x - x_0)^j \,|_{x=x_0} = k! \, \delta_{jk}$. Daraus ergibt sich allgemein für $k \leq n$

$$P(x) = \sum_{j=0}^{n} a_j (x - x_0)^j \Rightarrow P^{(k)}(x_0) = k! \, a_k. \tag{4.2}$$

Für P wie in (4.1) gilt $a_k = \frac{f^{(k)}(x_0)}{k!}$, also folgt $P^{(k)}(x_0) = f^{(k)}(x_0)$ und P leistet das Verlangte.

Für die Eindeutigkeit reicht es aus zu zeigen, dass ein Polynom P vom Grad höchstens n mit $P^{(k)}(x_0) = 0$ für $k = 0, 1, \ldots, n$ das Nullpolynom ist. Nun gilt

$$x^k = (x_0 + x - x_0)^k = \sum_{i=0}^{k} \binom{k}{i} x_0^{k-i}(x - x_0)^i.$$

Also gilt $P(x) = \sum_{j=0}^{n} a_j(x - x_0)^j$ mit geeigneten $a_j \in \mathbb{R}$, und aus (4.2) folgt $a_j = \frac{1}{j!} P^{(j)}(x_0) = 0$. Somit ist P das Nullpolynom. □

Folgerung 4.1 *Ist f ein Polynom vom Grad höchstens n, so gilt $P^n_{x_0} f = f$, das heißt das n-te Taylorpolynom von f ist f selbst.*

Definition 4.1 *Für $f \in C^n(I)$ und $x_0 \in I$ heißt die Funktion*

$$R^n_{x_0} f : (a, b) \to \mathbb{R}, \quad R^n_{x_0} f(x) = f(x) - P^n_{x_0} f(x)$$

Restglied der Taylorentwicklung n-ter Ordnung mit Entwicklungspunkt x_0.

Knackpunkt bei der Taylorentwicklung ist die Abschätzung des Restglieds; diese besagt, wie gut die Funktion f durch das Taylorpolynom $P^n_{x_0} f$ approximiert wird. Entscheidend für die Abschätzung sind die in den folgenden beiden Sätzen gelieferten Darstellungen.

Satz 4.1 (Integraldarstellung des Restglieds) *Sei $f \in C^{n+1}(I)$ für ein $n \in \mathbb{N}_0$ und $x_0 \in I$. Dann hat das Restglied $R^n_{x_0} f(x) = f(x) - P^n_{x_0} f(x)$ die Darstellung*

$$R^n_{x_0} f(x) = \frac{1}{n!} \int_{x_0}^{x} (x - y)^n f^{(n+1)}(y)\, dy. \tag{4.3}$$

Beweis: Durch Induktion über $n \in \mathbb{N}_0$. Der Fall $n = 0$ folgt aus dem Hauptsatz der Differential- und Integralrechnung:

$$R^0_{x_0} f(x) = f(x) - f(x_0) = \int_{x_0}^{x} f'(y)\, dy.$$

Der Induktionsschritt $n - 1 \Rightarrow n \geq 1$ beruht auf partieller Integration:

$$\begin{aligned}
R^n_{x_0} f(x) &= R^{n-1}_{x_0} f(x) - \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \quad (\text{Definition } R^n_{x_0}) \\
(\text{Induktion}) &= \frac{1}{(n-1)!} \int_{x_0}^{x} (x - y)^{n-1} f^{(n)}(y)\, dy - \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \\
(\text{part. Int.}) &= \frac{1}{(n-1)!} \left[-\frac{(x - y)^n}{n} f^{(n)}(y) \right]_{y=x_0}^{y=x} + \frac{1}{n!} \int_{x_0}^{x} (x - y)^n f^{(n+1)}(y)\, dy \\
&\quad - \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \\
&= \frac{1}{n!} \int_{x_0}^{x} (x - y)^n f^{(n+1)}(y)\, dy.
\end{aligned}$$

□

Satz 4.2 (Lagrangedarstellung des Restglieds) *Unter den Voraussetzungen von Satz 4.1 existiert ein $\xi \in [x_0, x]$ (bzw. $\xi \in [x, x_0]$ für $x \leq x_0$), so dass gilt:*

$$R^n_{x_0} f(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}. \tag{4.4}$$

BEWEIS: Sei $x \geq x_0$ (sonst analog). Dann ist $(x-y)^n \geq 0$ auf $[x_0, x]$. Nach dem MWS der Integralrechnung (Folgerung 1.2) existiert ein $\xi \in [x_0, x]$ mit

$$R^n_{x_0} f(x) = f^{(n+1)}(\xi) \int_{x_0}^{x} \frac{(x-y)^n}{n!}\, dy = f^{(n+1)}(\xi)\, \frac{(x-x_0)^{n+1}}{(n+1)!}$$

□

Beispiel 4.1 *Betrachte für $x \in (-1, 1)$ die Funktion*

$$\begin{array}{rclrcl} f(x) & = & (1-x)^{-1/2}, & f(0) & = & 1 \\ f'(x) & = & \frac{1}{2}(1-x)^{-3/2}, & f'(0) & = & 1/2 \\ f''(x) & = & \frac{3}{4}(1-x)^{-5/2}, & f''(0) & = & 3/4. \end{array}$$

Das Taylorpolynom erster Ordnung bei $x_0 = 0$ lautet

$$P^1_0 f(x) = \sum_{k=0}^{1} \frac{f^{(k)}(0)}{k!} x^k = 1 + \frac{1}{2}x,$$

mit der Lagrange-Restglieddarstellung

$$R^1_0 f(x) = \frac{f''(\xi)}{2} x^2 = \frac{3}{8}(1-\xi)^{-5/2} x^2 \quad \text{mit } \xi \in [0, x].$$

Definition 4.2 *Für $f \in C^\infty(I)$ und $x_0 \in I$ heißt die Reihe*

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

Taylorreihe von f mit Entwicklungspunkt x_0.

Die Taylorreihe ist eine Potenzreihe, mit Entwicklungspunkt $x_0 \in \mathbb{R}$. Insbesondere hat sie einen Konvergenzradius $R \in [0, +\infty]$:

$$\begin{array}{rcccl} |x - x_0| & < & R & \Rightarrow & \text{absolute Konvergenz} \\ |x - x_0| & > & R & \Rightarrow & \text{Divergenz} \end{array}$$

Selbst wenn die Taylorreihe einen positiven Konvergenzradius hat, muss sie nicht notwendig gegen die gegebene Funktion konvergieren.

Beispiel 4.2 *Betrachte $f : \mathbb{R} \to \mathbb{R}$ mit*

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$$

Dann gilt $f \in C^\infty(\mathbb{R})$ und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$, siehe Folgerung 2.6 in Kapitel 3. Damit sind alle Koeffizienten der Taylorreihe Null und diese konvergiert gegen die Nullfunktion, nicht gegen f.

Definition 4.3 *Eine Funktion $f : (a,b) \to \mathbb{R}$ heißt analytisch, wenn es zu jedem $x_0 \in (a,b)$ eine Umgebung $U_\delta(x_0) \subset (a,b)$ gibt, auf der die Funktion f als Potenzreihe mit Entwicklungspunkt x_0 darstellbar ist.*

Satz 4.3 *Seien $a_k \in \mathbb{R}$ für $k \in \mathbb{N}_0$. Für $f : I = (a,b) \to \mathbb{R}$ und $x_0 \in I$ sind folgende Aussagen äquivalent:*

(1) *f besitzt auf I die Potenzreihendarstellung*

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k \quad \text{für alle } x \in I.$$

(2) *$f \in C^\infty(I)$, $a_k = \frac{f^{(k)}(x_0)}{k!}$ und für alle $x \in I$ gilt $R^n_{x_0} f(x) \to 0$ mit $n \to \infty$.*

Beweis: (1) $\Rightarrow$ (2) : Nach Voraussetzung gilt für den Konvergenzradius der Potenzreihe

$$R \geq \max(|x_0 - a|, |x_0 - b|).$$

Nach Satz 3.3 ist $f \in C^\infty(I)$ und die Ableitungen können durch gliedweise Differentiation berechnet werden. Also folgt

$$f^{(k)}(x_0) = \sum_{j=0}^{\infty} a_j \, k! \, \delta_{jk} = k! a_k.$$

Folglich ist die Potenzreihe die Taylorreihe von f mit Entwicklungspunkt x_0, und es folgt

$$R^n_{x_0} f(x) = f(x) - \sum_{k=0}^{n} a_k \, x^k \to 0 \quad \text{nach Voraussetzung.}$$

(2) $\Rightarrow$ (1) : Nach Voraussetzung gilt für $n \in \mathbb{N}_0$

$$\begin{aligned} f(x) &= \sum_{k=0}^{n} a_k (x - x_0)^k + \underbrace{R^n_{x_0} f(x)}_{\to 0 \text{ mit } n \to \infty} \\ \Rightarrow \quad f(x) &= \sum_{k=0}^{\infty} a_k (x - x_0)^k. \end{aligned}$$

□

Beispiel 4.3 *Mit Hilfe der Restgliedabschätzung können wir erneut die Potenzreihenda-stellung der Funktionen* exp, cos *und* sin *beweisen. Wir betrachten zum Beispiel $f(x) = \cos x$. Es gilt dann*

$$f^{(n)}(x) = \begin{cases} (-1)^k \cos x & n = 2k \\ (-1)^k \sin x & n = 2k - 1 \end{cases}.$$

Insbesondere

$$f^{(n)}(0) = \begin{cases} (-1)^k & n = 2k \\ 0 & n = 2k - 1 \end{cases}.$$

Die Taylorreihe mit Entwicklungspunkt Null lautet also

$$P_0 f(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\, x^{2k}.$$

Für das Restglied folgt mit Lagrange

$$R_0^{2k} f(x) = \frac{(-1)^{k+1} \sin \xi}{(2k+1)!}\, x^{2k+1} \to 0 \quad \textit{mit } k \to \infty.$$