Skript zur Vorlesung

Analysis I

Wintersemester 2015/16

Robert Haller-Dintelmann

16. Februar 2016

Inhaltsverzeichnis

Teil I.

Zahlen und Mengen

1. Grundlegende Begriffe

Wir wollen uns zu Beginn kurz mit den Grundbausteinen der mathematischen Sprache, wie sie sich in den letzten gut 100 Jahren entwickelt hat, vertraut machen. Dazu führen wir die Begriffe Menge und Abbildung ein und besprechen einige logische Begriffe wie Aussage, Implikation oder Äquivalenz. Dieser Abschnitt ist aber definitiv keine grundlegende Einführung in die Mengenlehre oder die Logik, denn damit könnten wir uns problemlos ein ganzes Semester beschäftigen, sondern er dient vielmehr dazu, Sprechweisen zu vermitteln, die nicht Selbstzweck, sondern bequemes Hilfsmittel zur Darstellung mathematischer Inhalte sind.

1.1. Mengen und Aussagen

Beispiele von Mengen sind: Die Menge aller Studierenden in einem Hörsaal, ein Dreieck (als Punktmenge der Ebene), die Menge aller Dreiecke in der Ebene oder die Mengen $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, also die Mengen der natürlichen, ganzen, rationalen, reellen, bzw. komplexen Zahlen. Den Begriff der Menge definieren wir hier nicht, sondern legen ihn naiv zu Grunde; wir stellen uns damit auf den Standpunkt der naiven (und nicht der axiomatischen) Mengenlehre.

Wenn wir Mengen bilden, ist unser Ausgangspunkt immer eine gegebene, unter Umständen sehr großen Grundmenge G, aus der Elemente ausgesondert und zu neuen Mengen zusammengefasst werden. Auf diese Weise vermeidet man Bildungen wie die „Menge aller Mengen", die zu Widersprüchen führen.

Mengen kann man, solange sie klein genug sind, einfach durch das Aufzählen ihrer Elemente angeben, z. B.

$$M_1 = \{1, 2, 3, 4, 5\}.$$

Es ist aber häufig angenehmer, sie durch die Angabe einer definierenden Eigenschaft, die genau für die Elemente der Menge, und nur für diese, wahr ist, zu beschreiben. Für unsere Menge M_1 könnte das so aussehen:[1]

$$M_1 = \{x \in \mathbb{N} : x < 6\} \quad \text{oder} \quad M_1 = \{x \in \mathbb{N} : x - 5 \text{ ist keine natürliche Zahl}\}.$$

Allgemein schreibt man

$$M = \{x \in G : E(x)\},$$

[1] Man beachte, dass wir hier, wie auch im Rest der Vorlesung, Null nicht als natürliche Zahl ansehen

wobei G die Grundmenge ist, aus der die Elemente der Menge M ausgesondert werden sollen und $E(x)$ eine Aussageform, die bei Einsetzen eines Elements aus G zu einer *Aussage* wird, d. h. zu einem Satz der entweder wahr oder falsch ist. M enthält dann genau die Elemente, für die $E(x)$ eine wahre Aussage ist. Betrachtet man das weitere Beispiel

$$M_2 = \{n \in \mathbb{N} : n \text{ gerade}\},$$

so sieht man schnell den Vorteil dieser Methode gegenüber der reinen Aufzählung. Für die weiteren Betrachtungen in diesem Abschnitt ist stets eine Grundmenge G als gegeben anzunehmen.

Definition 1.1. *Es seien M und N Mengen. Dann verwenden wir die folgenden Notationen:*

(a) $a \in M$: a ist in M enthalten; $a \notin M$: a gehört nicht zu M.

(b) $N \subseteq M$: N ist eine Teilmenge *von M, d. h. jedes Element von N ist auch in M enthalten. Eine solche Teilmengenbeziehung nennt man auch eine* Inklusion*.*

(c) $N \supseteq M$: N ist Obermenge *von M, d. h. M ist eine Teilmenge von N.*

(d) $N = M$: Beide Mengen enthalten genau die gleichen Elemente.

(e) $\emptyset$: Dieses Symbol bezeichnet die leere Menge*, d. h. eine Menge, die kein Element enthält.*

Bemerkung 1.2. Man beachte, dass damit zwei Mengen M und N gleich sind, wenn sowohl $M \subseteq N$, als auch $N \subseteq M$ gilt. Die Unterteilung in diese zwei Teilschritte ist häufig eine gute Strategie um die Gleichheit von zwei Mengen nachzuweisen.

Definition 1.3. *Es seien M und N zwei Mengen. Dann heißt*

(a) $M \cup N = \{x \in G : x \in M \text{ oder } x \in N\}$[2] die Vereinigung *von M und N.*

(b) $M \cap N = \{x \in G : x \in M \text{ und } x \in N\}$ der Durchschnitt *oder auch nur* Schnitt *der Mengen M und N.*

(c) $M^c = \{x \in G : x \notin M\}$ das Komplement *von M (in G).*

(d) $M \setminus N = \{x \in M : x \notin N\}$ die Mengendifferenz *von M und N.*

(e) $M \times N = \{(m, n) : m \in M \text{und } n \in N\}$ das kartesische Produkt *von M und N.*

[2] „oder“ wird in der Mathematik immer im Sinne von „das eine, das andere oder beides“ verstanden, also nicht im Sinne von „entweder ... oder“.

Mit diesen Begriffen kann man nun schon etwas Mathematik betreiben. Wir sammeln die wichtigsten Regeln für obige Mengenoperationen im folgenden Satz.

Satz 1.4. *Es seien A, B und C Teilmengen einer Grundmenge G. Dann gelten*

(a) $A \cup B = B \cup A$ *und* $A \cap B = B \cap A$ (Kommutativgesetze),

(b) $(A \cup B) \cup C = A \cup (B \cup C)$ *und* $(A \cap B) \cap C = A \cap (B \cap C)$ (Assoziativgesetze),

(c) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *und* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributivgesetze),

(d) $(A \cup B)^c = A^c \cap B^c$ *und* $(A \cap B)^c = A^c \cup B^c$ (Regeln von De Morgan).

Beweis. Wir behandeln hier das erste Distributivgesetz und die erste Regel von De Morgan, die weiteren verbleiben als Übungsaufgabe. Für das Distributivgesetz zeigen wir zuerst (vgl. Bemerkung 1.2)

$$A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C),$$

und zwar folgendermaßen: Sei $x \in A \cup (B \cap C)$. Dann ist also $x \in A$ oder $x \in B \cap C$. Betrachten wir zunächst den Fall $x \in A$. Dann gilt natürlich auch $x \in A \cup B$ und $x \in A \cup C$, denn diese Mengen sind ja größer als A. Also ist $x \in (A \cup B) \cap (A \cup C)$ und wir sind fertig. Betrachten wir also den Fall $x \in B \cap C$. Dann ist $x \in B$ und $x \in C$, also gilt wieder $x \in A \cup B$ und $x \in A \cup C$, dieses Mal, weil x sowohl in B als auch in C liegt. Daraus folgt wieder $x \in (A \cup B) \cap (A \cup C)$ und wir haben $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$ gezeigt.
Um die im ersten Distributivgesetz behauptete Gleichheit zu zeigen, müssen wir nun noch die umgekehrte Inklusion

$$(A \cup B) \cap (A \cup C) \subseteq A \cup (B \cap C)$$

zeigen. Dazu sei $x \in (A \cup B) \cap (A \cup C)$. Dann ist x sowohl in $A \cup B$, als auch in $A \cup C$. Wir betrachten die beiden Fälle $x \in A$ und $x \notin A$. (Man beachte, dass wir dann alle mögliche Fälle berücksichtigt haben!) Ist $x \in A$, so haben wir sofort auch $x \in A \cup (B \cap C)$, was unser Ziel war. Es bleibt also der Fall $x \notin A$. Da dann x in $A \cup B$ ist, ohne in A zu sein, muss x zwangsläufig in B sein, denn wie sollte es sonst da hineinkommen? Genauso folgt $x \in C$ aus $x \in A \cup C$. Also ist x in $B \cap C$ und damit auch $x \in A \cup (B \cap C)$ und wir haben auch die zweite Inklusion und damit die Gleichheit

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

gezeigt.

Für die erste De Morgan'sche Regel zeigen wir wieder zuerst

$$(A \cup B)^c \subseteq A^c \cap B^c.$$

Sei dazu $x \in (A \cup B)^c$. Dann ist $x \notin (A \cup B)$, d. h. x ist nicht in der Vereinigung von A und B. Damit kann x weder in A noch in B sein, denn sonst würde es ja in dieser Vereinigung liegen. Es ist also $x \notin A$ und $x \notin B$, d. h. $x \in A^c$ und $x \in B^c$, was schließlich $x \in A^c \cap B^c$ nach sich zieht.
Die zweite Inklusion

$$(A \cup B)^c \supseteq A^c \cap B^c$$

geht folgendermaßen: Es sei $x \in A^c \cap B^c$. Dann ist $x \in A^c$ und $x \in B^c$. Also ist x nicht in A und nicht in B, es ist also auch nicht in der Vereinigung von A und B, was gerade $x \in (A \cup B)^c$ bedeutet. $\square$

Bemerkung 1.5. Zuweilen bildet man Schnitte oder Vereinigungen von vielen, zum Teil sogar unendlich vielen Mengen. Dazu ist folgende Notation nützlich. Ist I irgendeine Menge (Man nennt I in diesem Zusammenhang *Indexmenge*) und ist für jedes $i \in I$ eine Menge M_i gegeben, so ist

$$\begin{aligned} \bigcup_{i \in I} M_i &= \{x \in G : \text{ es gibt ein } j \in I \text{ mit } x \in M_j\} \quad \text{und} \\ \bigcap_{i \in I} M_i &= \{x \in G : x \in M_j \text{ für alle } j \in I\}. \end{aligned}$$

Ist $I = \mathbb{N}$, so schreibt man auch oft $\bigcup_{n=1}^{\infty} M_n$, bzw. $\bigcap_{n=1}^{\infty} M_n$ statt $\bigcup_{n \in \mathbb{N}} M_n$ und $\bigcap_{n \in \mathbb{N}} M_n$.

1.2. Zeichenerklärung für die logischen Symbole

Im Alltagsgebrauch wird das Gleichheitszeichen für zwei grundverschiedene Dinge verwendet; wer schon einmal programmiert hat, wird das Problem kennen. Zum einen bedeutet $a = b$ die *Aussage*, dass a und b gleich sind. Zum anderen schreibt man oft $a = 2$, wenn man a auf den Wert 2 setzen will, also auf diese Weise *a definiert*. Auch in mathematischen Texten wird oft nicht exakt auf die Unterscheidung geachtet. Jetzt zum Anfang sollten wir uns aber bemühen, hier zu unterscheiden und führen dazu die folgende sehr gebräuchliche Notation ein:

- Das Zeichen „:=“ bedeutet „per Definition gleich“.
- Das Zeichen „=“ steht in der Gleichheits-Aussage.

All- und Existenzquantor

Oft werden in mathematischen Texten die folgenden Zeichen verwendet:

- Der *Allquantor* $\forall$ bedeutet „für alle“. Beispiel: „$\forall n \in \mathbb{N} : 2n$ ist gerade“ ist eine wahre Aussage.

- Der *Existenzquantor* $\exists$ steht für „es existiert". Beispiel: „$\exists x \in \mathbb{R} : x^2 = -1$" ist eine falsche Aussage.

Mit diesen beiden unscheinbaren Zeichen kann man schon sehr effizient einfache Dinge kompliziert aufschreiben. Für ein etwas elaborierteres Beispiel, das die beiden Quantoren kombiniert, definieren wir S als die Menge aller Städte und W als die Menge aller Wege auf der Erde und betrachten die Aussage

$$\exists s \in S \; \forall w \in W : w \text{ führt nach } s. \tag{1.1}$$

Übersetzt: Es gibt eine Stadt s, zu der jeder Weg hinführt. Meistens wird s dann Rom genannt.
An diesem Beispiel sieht man zum einen, warum der exzessive Gebrauch von Quantoren in geschriebener Mathematik eher verpönt ist, zum anderen kann man aber nun auch gut demonstrieren, wozu Quantoren unter anderem sehr praktisch sind. Dazu sei die Frage gestellt: Was ist die Verneinung der obigen Aussage, d. h. formulieren Sie eine Aussage, die genau dann wahr ist, wenn (1.1) falsch ist und genau dann falsch ist, wenn (1.1) wahr ist.
Es gilt nun die folgende *Regel zum Verneinen von Aussagen*:

Jedes $\exists$ wird ein $\forall$, jedes $\forall$ ein $\exists$ und die Bedingung am Ende wird verneint.

Im obigen Beispiel also

$$\forall s \in S \; \exists w \in W : w \text{ führt nicht nach } s,$$

d. h. für jede Stadt gibt es einen Weg, der nicht zu ihr führt.
Darauf mag man auch noch ohne Quantoren kommen, wenn die Aussagen komplizierter werden, ist das „Quantoren-Umklappen" manchmal ein vernünftiger Weg, vgl. z. B. Seite 135.

Implikation und Äquivalenz

Sind A und B zwei Aussagen, so bezeichnet man mit

- „$A \Longrightarrow B$" die Aussage „Aus A folgt B" oder „A impliziert B" (*Implikation*).
- „$A \iff B$" die Aussage „A gilt genau dann, wenn B gilt" oder „A ist äquivalent zu B" (*Äquivalenz*).

Die Wahrheitswerte dieser Aussagen ergeben sich aus der folgenden Wahrheitstafel:

A	B	$A \Longrightarrow B$	$A \iff B$
w	w	w	w
w	f	f	f
f	w	w	f
f	f	w	w

Man beachte, dass für eine falsche Aussage A die Implikation $A \Longrightarrow B$ immer wahr ist. Das finden manche auf den ersten Blick unintuitiv, es ist aber sehr sinnvoll. Man sieht das z. B. durch Übertreibung ins Absurde: „Wenn der Mond aus fünfeckigem Käse ist, dann ist die Erde eine Scheibe.“
Oft hat man in Beweisen die Äquivalenz zweier Aussagen A und B nachzuweisen. Dazu ist es meist von Vorteil, diese Aufgabe in die beiden Teilprobleme $A \Longrightarrow B$ und $B \Longrightarrow A$ aufzuteilen und diese beiden Implikationen getrennt zu beweisen.

Übungsaufgabe: Machen Sie sich anhand einer Wahrheitstafel klar, dass die beiden Aussagen „$A \iff B$“ und „$A \Longrightarrow B$ und $B \Longrightarrow A$“ tatsächlich die gleichen Wahrheitswerte haben.

Hat man sogar „$A \iff B \iff C \iff D \iff \dots \iff P$“ zu beweisen, so hilft das Prinzip des *Ringschlusses*: Man zeigt „$A \Longrightarrow B \Longrightarrow C \Longrightarrow \dots \Longrightarrow P \Longrightarrow A$“. Machen Sie sich auch hier klar, das damit wirklich die obige Aussage gezeigt ist!

1.3. Abbildungen

Wir betrachten nun den für alle Teilbereiche der Mathematik wichtigen Begriff der Abbildung (oder auch Funktion).

Definition 1.6. *Es seien A und B nicht-leere Mengen und es sei jedem Element a aus A genau ein Element $f(a)$ in B zugeordnet. Diese Zuordnung nennt man dann eine* Abbildung *oder auch* Funktion *f und schreibt*

$$f : A \to B, \quad a \mapsto f(a) \qquad \text{oder} \qquad f : \begin{cases} A \to B \\ a \mapsto f(a). \end{cases}$$

Dabei heißt A die Definitionsmenge *und B die* Zielmenge *von f. Weiter ist $f(A) := \{f(a) : a \in A\} \subseteq B$ das* Bild *von f und ist $b \in B$ gegeben, so heißt jedes Element $a \in A$ mit $f(a) = b$ ein* Urbild *von b. Ist schließlich $C \subseteq B$, so nennt man die Menge*

$$f^{-1}(C) := \{a \in A : f(a) \in C\}$$

das Urbild *der Menge C.*

Beispiel 1.7. (a) Die Zuordnung, die aus der Menge aller Studierenden jedem Studienfach die dort Eingeschriebenen zuordnet, ist keine Abbildung, wohl aber die, die jedem Studienfach die Anzahl der dort Eingeschriebenen zuordnet. Überlegen Sie sich warum.

(b) Im weiteren Verlauf werden wir es vor allem mit Funktionen zu tun haben, die zwischen Mengen von Zahlen definiert sind. Beispiele wären hier das Potenzieren mit zwei in $\mathbb{R}$, der Menge der reellen Zahlen,

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x) := x^2$$

oder die Wurzelfunktion

$$\sqrt{} : [0, \infty) \to \mathbb{R}, \quad x \mapsto \sqrt{x}.$$

Weiter definieren wir die Verkettung, d. h. die Nacheinanderausführung von Abbildungen.

Definition 1.8. *Es seien A, B und C Mengen und $f : A \to B$ sowie $g : B \to C$ Funktionen. Dann heißt die Funktion $g \circ f$ (lies „g* nach *f“), gegeben durch*

$$g \circ f : A \to C, \quad a \mapsto (g \circ f)(a) := g(f(a)),$$

die Verkettung *von f und g.*

Wir wollen einigen besonders schönen Eigenschaften von Funktionen einen Namen geben.

Definition 1.9. *Es seien A und B zwei Mengen sowie $f : A \to B$ eine Funktion. Dann heißt f*

(a) surjektiv *genau dann, wenn $f(A) = B$ gilt.*

(b) injektiv *genau dann, wenn für $x, y \in A$ mit $f(x) = f(y)$ stets $x = y$ gilt.*

(c) bijektiv *genau dann, wenn f injektiv und surjektiv ist.*

Bijektive Abbildungen sind deshalb besonders wichtig, weil sie sich umkehren, d. h. rückgängig machen lassen.

Satz 1.10. *Es seien A, B Mengen und $f : A \to B$ eine Funktion. Dann ist f genau dann bijektiv, wenn es für jedes $b \in B$ genau ein $a \in A$ gibt, so dass $f(a) = b$ ist. In diesem Fall existiert eine Abbildung $f^{-1} : B \to A$, so dass*

$$f^{-1}(f(a)) = a \quad \textit{für alle } a \in A \quad \textit{und} \quad f(f^{-1}(b)) = b \quad \textit{für alle } b \in B$$

gilt.

Beweis. **1. Schritt:** Wir zeigen: f bijektiv $\Longrightarrow$ für alle $b \in B$ existiert genau ein $a \in A$ mit $f(a) = b$.

Da f surjektiv ist, gibt es zu jedem $b \in B$ mindestens ein $a \in A$ mit $f(a) = b$. Nehmen wir an, es gäbe mehr als eins, d. h. es gäbe $a_1, a_2 \in A$ mit $f(a_1) =$

$f(a_2) = b$, so folgt aus der Injektivität von f sofort $a_1 = a_2$, es kann also nur genau ein solches $a \in A$ geben.

2. Schritt: Wir zeigen: Für alle $b \in B$ existiert genau ein $a \in A$ mit $f(a) = b \Longrightarrow f$ bijektiv.

Nach Voraussetzung sind alle $b \in B$ in $f(A)$ enthalten, also ist f surjektiv. Seien nun $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$ gegeben. Da jedes $b \in B$ nur genau ein Urbild hat, muss dann $a_1 = a_2$ sein, d. h. f ist auch injektiv.

3. Schritt: Wir zeigen: f bijektiv $\Longrightarrow$ es existiert $f^{-1} : B \to A$ mit $f^{-1}(f(a)) = a$ für alle $a \in A$ und $f(f^{-1}(b)) = b$ für alle $b \in B$.

Für jedes $b \in B$ definieren wir $f^{-1}(b) := a$, wobei $a \in A$ das nach dem ersten Schritt eindeutig bestimmte Element mit $f(a) = b$ ist. Dann ist $f^{-1}(f(a))$ das Element von A, das in f eingesetzt $f(a)$ ergibt, also $f^{-1}(f(a)) = a$ für alle $a \in A$. Sei nun $b \in B$. Dann ist $f^{-1}(b)$ das Element von A mit $f(f^{-1}(b)) = b$ und wir sind fertig. $\square$

Definition 1.11. *Es seien A, B zwei Mengen und $f : A \to B$ bijektiv. Dann heißt die Abbildung f^{-1} aus Satz 1.10* Umkehrfunktion *von f.*

Beispiel 1.12. Betrachten wir noch einmal die beiden Abbildungen aus Beispiel 1.7 (b), so ist die erste Funktion

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x) := x^2$$

weder injektiv (denn $f(1) = f(-1)$, aber $1 \neq -1$) noch surjektiv (denn $-1 \notin f(\mathbb{R})$). Betrachten wir dagegen

$$\tilde{f} : \mathbb{R} \to \{x \in \mathbb{R} : x \geq 0\}, \quad x \mapsto \tilde{f}(x) := x^2,$$

so ist diese nun surjektiv, denn jede positive reelle Zahl ist das Quadrat einer reellen Zahl, aber weiterhin nicht injektiv, denn das Problem mit $f(1) = f(-1)$ bleibt bestehen. Das können wir lösen, indem wir nun noch den Definitionsbereich einschränken, d. h. wir betrachten

$$\hat{f} : \{x \in \mathbb{R} : x \geq 0\} \to \{x \in \mathbb{R} : x \geq 0\}, \quad x \mapsto \hat{f}(x) := x^2.$$

Nun ist $\hat{f}$ tatsächlich bijektiv (Nachweisen!) und die oben erwähnte Wurzelfunktion ist die Umkehrabbildung.

Wie in obigem Beispiel will man oft eine gegebene Funktion nur auf einem Teil ihres Definitionsbereiches untersuchen. Dazu vereinbaren wir die folgende Notation.

Definition 1.13. *Seien A, B Mengen, $f : A \to B$ eine Funktion und $M \subseteq A$. Dann ist $f|_M$ die* Einschränkung von f auf M*, d. h. $f|_M : M \to B$ ist gegeben durch $f|_M(x) = f(x)$ für $x \in M$.*

2. Die reellen Zahlen

Die Grundmenge der Analysis ist die Menge der reellen Zahlen, geschrieben $\mathbb{R}$. Diese führen wir *axiomatisch* ein, d. h. wir postulieren eine gewisse Anzahl von Grundannahmen, genannt Axiome, deren Gültigkeit wir zu Grunde legen, ohne sie beweisen zu können. Schließlich kann auch die Mathematik nichts aus dem luftleeren Raum heraus beweisen, mit irgendwelchen Voraussetzungen muss man anfangen. Dieser Herangehensweise werden Sie im weiteren Studium noch vielfach begegnen.

Körperaxiome

In $\mathbb{R}$ sind zwei Abbildungen („Verknüpfungen") $+ : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ und $\cdot : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ gegeben, genannt *Addition* und *Multiplikation*, die jedem Paar von Elementen $a, b \in \mathbb{R}$ ein $a + b \in \mathbb{R}$, bzw. ein $a \cdot b \in \mathbb{R}$ zuordnen. Dabei sollen die folgenden Axiome gelten.

(A1) $a + (b + c) = (a + b) + c$ für alle $a, b, c \in \mathbb{R}$ (*Assoziativgesetz der Addition*).

(A2) Es gibt ein Element $0 \in \mathbb{R}$, so dass $a + 0 = a$ für alle $a \in \mathbb{R}$ (*Nullelement*).

(A3) Für jedes $a \in \mathbb{R}$ gibt es ein $-a \in \mathbb{R}$, so dass $a + (-a) = 0$ gilt (*additives inverses Element*).

(A4) $a + b = b + a$ für alle $a, b \in \mathbb{R}$ (*Kommutativgesetz der Addition*).

(A5) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in \mathbb{R}$ (*Assoziativgesetz der Multiplikation*).

(A6) Es gibt ein Element $1 \in \mathbb{R}$ mit $1 \neq 0$, so dass $a \cdot 1 = a$ für alle $a \in \mathbb{R}$ (*Einselement*).

(A7) Für jedes $a \in \mathbb{R} \setminus \{0\}$ gibt es ein $a^{-1} \in \mathbb{R}$, so dass $a \cdot a^{-1} = 1$ gilt (*multiplikatives inverses Element*).

(A8) $a \cdot b = b \cdot a$ für alle $a, b \in \mathbb{R}$ (*Kommutativgesetz der Multiplikation*).

(A9) $a \cdot (b + c) = a \cdot b + a \cdot c$ für alle $a, b, c \in \mathbb{R}$ (*Distributivgesetz*).

Bemerkung 2.1. Alle bekannten Rechenregeln für „+" und „·" lassen sich aus (A1) – (A9) ableiten.

Wir betrachten die folgenden Aussagen als Beispiele:

Satz 2.2. *(a) Es gibt genau ein Nullelement in* $\mathbb{R}$*.*

(b) $a \cdot 0 = 0$ *für alle* $a \in \mathbb{R}$*.*

(c) Gilt $a \cdot b = 0$ *für zwei reelle Zahlen* a, b*, so ist* $a = 0$ *oder* $b = 0$*.*

Beweis. (a) Sei $\tilde{0} \in \mathbb{R}$ ein weiteres Nullelement, d. h. für alle $a \in \mathbb{R}$ gilt $a + \tilde{0} = a$. Insbesondere gilt also für $a = 0$ damit $0 + \tilde{0} = 0$. Mit (A2) für $a = \tilde{0}$ haben wir außerdem $\tilde{0} + 0 = \tilde{0}$. Also können wir mit Hilfe von (A4) folgern:

$$\tilde{0} = \tilde{0} + 0 = 0 + \tilde{0} = 0.$$

(b) Nach (A2) gilt $0 + 0 = 0$, also ist auch für jedes $a \in \mathbb{R}$

$$a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$$

unter Zuhilfenahme von (A9). Daraus folgt nun

$$\begin{aligned} 0 &\overset{\text{(A3)}}{=} a \cdot 0 + \big(-(a \cdot 0)\big) = (a \cdot 0 + a \cdot 0) + \big(-(a \cdot 0)\big) \\ &\overset{\text{(A1)}}{=} a \cdot 0 + \big(a \cdot 0 + (-(a \cdot 0))\big) \overset{\text{(A3)}}{=} a \cdot 0 + 0 \overset{\text{(A2)}}{=} a \cdot 0, \end{aligned}$$

also $a \cdot 0 = 0$.

(c) Es seien $a, b \in \mathbb{R}$ mit $a \cdot b = 0$. Im Falle $a = 0$ sind wir fertig, wir betrachten also den Fall $a \neq 0$. Dann gibt es nach (A7) ein Element $a^{-1} \in \mathbb{R}$ mit $a \cdot a^{-1} = 1$. Also ist in diesem Fall

$$b \overset{\text{(A6)}}{=} b \cdot 1 = b \cdot (a \cdot a^{-1}) \overset{\text{(A5)}}{=} (b \cdot a) \cdot a^{-1} \overset{\text{(A8)}}{=} (a \cdot b) \cdot a^{-1} = 0 \cdot a^{-1} \overset{\text{(b)}}{=} 0,$$

d. h. $b = 0$. Damit folgt die Behauptung. □

Definition 2.3 (Schreibweisen)**.**

(a) Den „·“ für die Multiplikation lassen wir meist weg und schreiben einfach „ab“ statt „$a \cdot b$“.

(b) Wir setzen $a - b := a + (-b)$ *für* $a, b \in \mathbb{R}$ (Subtraktion).

(c) Sind $a, b \in \mathbb{R}$ *und gilt* $b \neq 0$*, so schreiben wir* $\frac{a}{b} := a \cdot b^{-1}$ (Division).

Anordnungsaxiome

In $\mathbb{R}$ ist eine Relation $\leq$ mit den folgenden Eigenschaften gegeben.

(A10) Für jede Wahl von $a, b \in \mathbb{R}$ gilt stets $a \leq b$ oder $b \leq a$ (*Totalordnung*).

(A11) Gelten für zwei Zahlen $a, b \in \mathbb{R}$ die beiden Aussagen $a \leq b$ und $b \leq a$, so ist $a = b$ (*Antisymmetrie*).

(A12) Wenn für drei Zahlen $a, b, c \in \mathbb{R}$ sowohl $a \leq b$ als auch $b \leq c$ gilt, so ist auch $a \leq c$ (*Transitivität*).

(A13) Sind $a, b, c \in \mathbb{R}$ und gilt $a \leq b$, so ist auch $a + c \leq b + c$.

(A14) Sind $a, b, c \in \mathbb{R}$ und gilt $a \leq b$ und $0 \leq c$, so ist auch $ac \leq bc$.

Definition 2.4 (Schreibweisen)**.** *Seien $a, b \in \mathbb{R}$. Wir setzen*

(a) $b \geq a$ genau dann, wenn $a \leq b$ ist,

(b) $a < b$ genau dann, wenn $a \leq b$ und $a \neq b$ ist,

(c) $a > b$ genau dann, wenn $b < a$ ist.

Bemerkung 2.5. Alle Regeln für Ungleichungen lassen sich aus den Axiomen (A1) – (A14) ableiten.

Wir betrachten wieder ein Beispiel.

Satz 2.6. *Seien $a, b, c \in \mathbb{R}$. Gilt $a < b$ und $0 < c$, so ist $ac < bc$.*

Beweis. Insbesondere ist $a \leq b$ und $0 \leq c$, also ist wegen (A14) $ac \leq bc$. Noch zu zeigen ist $ac \neq bc$. Wir nehmen an, es gelte $ac = bc$. Dann ist $(a - b)c = 0$ und wegen Satz 2.2 (c) können wir daraus $a - b = 0$ oder $c = 0$ folgern. Nach Voraussetzung ist $c = 0$ falsch, also gilt $a - b = 0$, d. h. $a = b$, was auch ein Widerspruch ist. Damit ist $ac \neq bc$. □

Übungsaufgabe 2.7. Zeigen Sie: Sind $a, b, c \in \mathbb{R}$ mit $a \leq b$ und $c \leq 0$, so gilt $ac \geq bc$.

Wir definieren nun die Betragsfunktion, ein fundamentales Hilsmittel in der gesamten Analysis. Anschaulich gesprochen misst diese den Abstand einer reellen Zahl zur Null.

Definition 2.8. *Sei $a \in \mathbb{R}$. Dann ist der* Betrag *von a, symbolisiert durch $|a|$, gegeben durch*

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

Es gelten die folgenden Rechenregeln.

Satz 2.9. *Für alle $a, b \in \mathbb{R}$ gilt*

(a) $|a| \geq 0$

(b) $|a| = |-a|$,

(c) $\pm a \leq |a|$,

(d) $|ab| = |a| \cdot |b|$,

(e) $|a| = 0$ *genau dann, wenn* $a = 0$,

(f) $|a + b| \leq |a| + |b|$ (Dreiecksungleichung),

(g) $\big||a| - |b|\big| \leq |a - b|$ (umgekehrte Dreiecksungleichung).

Beweis. (a)-(d) Übungsaufgabe.

(e) Zu zeigen ist: $|a| = 0 \iff a = 0$. Die Implikation „$\Longleftarrow$" folgt direkt aus der Definition des Betrages. Für die umgekehrte Implikation beobachten wir, dass für alle $a > 0$ auch $|a| = a > 0$ ist und dass für alle $a < 0$ genauso $|a| = -a > 0$ ist. Also gilt $|a| = 0$ nur für $a = 0$.

(f) Wir betrachten zunächst den Fall $a + b \geq 0$. Dann gilt nach Definition des Betrags $|a + b| = a + b$ und mit Hilfe von (c) ist $|a + b| = a + b \leq |a| + |b|$.

Ist dagegen $a + b < 0$, so gilt $|a + b| = -(a + b) = -a + (-b)$, woraus wieder mit (c) $|a + b| \leq |a| + |b|$ folgt.

(g) Mit Hilfe von (f) gilt $|a| = |a - b + b| \leq |a - b| + |b|$ und damit haben wir $|a| - |b| \leq |a - b|$. Analog erhält man durch Vertauschen der Rollen von a und b die Ungleichung $|b| - |a| \leq |b - a| = |a - b|$. Da $|b| - |a| = -(|a| - |b|)$ ist, sehen wir damit $|a| - |b| \leq |a - b|$ und $-(|a| - |b|) \leq |a - b|$, woraus nach der Definition des Betrages die Behauptung folgt. □

Als Vorbereitung für das 15. Axiom führen wir einige Schreibweisen und Begriffe ein, die bei der Untersuchung von Teilmengen von $\mathbb{R}$ entscheidend sind. Besonders wichtige solcher Teilmengen sind die Intervalle. Da hier verschiedene Bezeichnungsweisen üblich sind, einigen wir uns für diese Vorlesung auf die folgenden Notationen.

Definition 2.10. *Es seien zwei Zahlen $a, b \in \mathbb{R}$ mit $a < b$ gegeben. Dann heißen*

- $(a, b) := \{x \in \mathbb{R} : a < x < b\}$ offenes Intervall*,*
- $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$ abgeschlossenes Intervall*,*
- $(a, b] := \{x \in \mathbb{R} : a < x \leq b\}$ *und*
- $[a, b) := \{x \in \mathbb{R} : a \leq x < b\}$ halboffene Intervalle*.*

Um auch die Fälle von Halbstrahlen abzudecken, definieren wir weiter:

- $[a, \infty) := \{x \in \mathbb{R} : a \leq x\}$,
- $(a, \infty) := \{x \in \mathbb{R} : a < x\}$,

- $(-\infty, a] := \{x \in \mathbb{R} : x \leq a\}$,
- $(-\infty, a) := \{x \in \mathbb{R} : x < a\}$,
- $(-\infty, \infty) := \mathbb{R}$.

Definition 2.11. *Es sei M eine nicht-leere Teilmenge von $\mathbb{R}$.*

(a) M heißt nach oben beschränkt *genau dann, wenn ein $C \in \mathbb{R}$ existiert, so dass $x \leq C$ für alle $x \in M$ gilt.*

In diesem Fall heißt C eine obere Schranke *von M.*

(b) M heißt nach unten beschränkt *genau dann, wenn ein $C \in \mathbb{R}$ existiert, so dass $x \geq C$ für alle $x \in M$ gilt.*

In diesem Fall heißt C eine untere Schranke *von M.*

(c) Ist C eine obere Schranke von M und für jede weitere obere Schranke $\tilde{C}$ von M gilt $C \leq \tilde{C}$, so heißt C Supremum *von M. Wir bezeichnen es mit* $\sup M$.

(d) Ist C eine untere Schranke von M und für jede weitere untere Schranke $\tilde{C}$ von M gilt $C \geq \tilde{C}$, so heißt C Infimum *von M. Wir bezeichnen es mit* $\inf M$.

Als Merkregel kann man behalten: Das Supremum ist (falls es existiert) die kleinste obere Schranke einer Menge und das Infimum die größte untere Schranke. Man beachte, dass Supremum und Infimum nicht unbedingt existieren müssen, so ist zum Beispiel die Menge $\mathbb{R}$ weder nach oben noch nach unten beschränkt, hat also weder ein Supremum noch ein Infimum.
Falls eine Menge M aber ein Supremum (bzw. ein Infimum) hat, so ist dieses eindeutig bestimmt. Wir überlegen uns das für das Supremum: Angenommen es gäbe zwei Suprema C_1 und C_2. Dann sind sowohl C_1 als auch C_2 obere Schranken von M. Da also C_1 ein Supremum und C_2 eine obere Schranke von M ist, gilt nach der Definition des Supremums $C_1 \leq C_2$. Umgekehrt ist aber auch C_2 ein Supremum und C_1 eine obere Schranke von M. Also gilt $C_2 \leq C_1$. Nach Axiom (A11) gilt dann $C_2 = C_1$. Das motiviert die nachstehende Definition.

Definition 2.12. *Es sei $M \subseteq \mathbb{R}$ und $M \neq \emptyset$.*

(a) Existiert $\sup M$ und gilt $\sup M \in M$, so heißt $\sup M$ das Maximum *von M. Wir bezeichnen es mit* $\max M$.

(b) Existiert $\inf M$ und gilt $\inf M \in M$, so heißt $\inf M$ das Minimum *von M. Wir bezeichnen es mit* $\min M$.

Beispiel 2.13. (a) $M = (0,1]$. Dann ist M nach oben und nach unten beschränkt. Eine obere Schranke ist 17, eine untere ist 0. Weiter gilt $\sup M = \max M = 1$ und $\inf M = 0$. M hat aber kein Minimum, denn $0 \notin M$!

(b) $M = (-\infty, -1)$. Dann ist M nach oben aber nicht nach unten beschränkt. Weiter gilt $\sup M = -1$, aber M hat kein Maximum. Da M kein Infimum hat, erübrigt sich die Suche nach einem Minimum.

Wir kommen nun zum letzten Axiom der reellen Zahlen.

Vollständigkeitsaxiom

(A15) Jede Teilmenge $M \subseteq \mathbb{R}$ mit $M \neq \emptyset$, die nach oben beschränkt ist, besitzt ein Supremum.

Dieses Axiom wird auch *Archimedisches Axiom* genannt.

Diskussionsanregung: Warum reichen die Axiome (A1) – (A14) noch nicht aus, um die Menge der reellen Zahlen eindeutig zu beschreiben? Denken Sie dabei an die rationalen Zahlen. Warum gilt das Vollständigkeitsaxiom für die Menge der rationalen Zahlen nicht?

Mit diesem Axiom können wir die entsprechende Aussage über das Infimum beweisen.

Satz 2.14. *Ist $M \subseteq \mathbb{R}$, $M \neq \emptyset$ und M nach unten beschränkt, so existiert* $\inf M$.

Beweis. Wir setzen $\tilde{M} := \{-x : x \in M\}$. Nach Voraussetzung existiert eine untere Schranke C_* von M. Für diese gilt also $C_* \leq x$ für alle $x \in M$. Damit ist $-x \leq -C_*$ für alle $x \in M$, also ist $-C_*$ eine obere Schranke von $\tilde{M}$. Nach Axiom (A15) existiert also $s := \sup \tilde{M}$. Weiter gilt $-x \leq s$ für alle $x \in M$, also ist $-s \leq x$ für alle diese x. Das bedeutet, dass $-s$ eine untere Schranke von M ist. Wir müssen noch zeigen, dass $-s$ die größte untere Schranke von M ist. Sei also σ eine weitere untere Schranke von M. Dann ist wie oben $-\sigma$ eine obere Schranke von $\tilde{M}$. Da s das Supremum von $\tilde{M}$ ist, muss also $s \leq -\sigma$, und damit $\sigma \leq -s$ gelten. Also ist $-s = \inf M$. □

Um die bisher eingeführten Begriffe abzurunden, fehlt noch der folgende.

Definition 2.15. *Eine Menge $M \subseteq \mathbb{R}$ mit $M \neq \emptyset$ heißt* beschränkt *genau dann, wenn M nach oben und nach unten beschränkt ist.*

Zum Abschluss dieses Kapitels beweisen wir noch einen Satz, der ein „im realen Leben“ nachprüfbares Kriterium angibt, ob eine obere (bzw. untere) Schranke tatsächlich das Supremum (bzw. Infimum) ist.

Satz 2.16. *(a) Ist $A \subseteq \mathbb{R}$ nicht-leer und beschränkt, so gilt* $\inf A \leq \sup A$.

(b) Ist $A \subseteq \mathbb{R}$ nach oben (bzw. unten) beschränkt und $B \subseteq A$ nicht-leer, so ist auch B nach oben (bzw. unten) beschränkt und es gilt $\sup B \leq \sup A$ *(bzw.* $\inf B \geq \inf A$*).*

(c) Ist $A \subseteq \mathbb{R}$ nicht-leer und C eine obere (bzw. untere) Schranke von A, so ist $C = \sup A$ (bzw. $C = \inf A$) genau dann, wenn für jedes $\varepsilon > 0$ ein $a \in A$ existiert, für das $a > C - \varepsilon$ (bzw. $a < C + \varepsilon$) gilt.

Beweis. (a) Sei $x \in A$ beliebig gewählt. Dann gilt $x \geq \inf A$ und $x \leq \sup A$. Also ist $\inf A \leq x \leq \sup A$.

(b) Für jedes $b \in B$ gilt $b \in A$ und damit $b \leq \sup A$ (bzw. $b \geq \inf A$). Also ist B nach oben (bzw. unten) beschränkt und $\sup A$ (bzw. $\inf A$) ist eine obere (bzw. untere) Schranke von B. Also ist $\sup B \leq \sup A$ (bzw. $\inf B \geq \inf A$).

(c) Wir beweisen für diese Aussage nur noch den „sup"-Fall.

„$\Rightarrow$" Sei $C = \sup A$ und $\varepsilon > 0$. Dann ist $C - \varepsilon$ kleiner als das Supremum von A, also keine obere Schranke von A. Damit existiert ein $a \in A$, so dass $a > C - \varepsilon$ ist.

„$\Leftarrow$" Wir müssen zeigen, dass für jede andere obere Schranke $\tilde{C}$ von A gilt $C \leq \tilde{C}$. Sei also $\tilde{C}$ eine weitere solche Schranke und wir nehmen an, es gelte $C > \tilde{C}$. Dann ist $\varepsilon := C - \tilde{C} > 0$. Nach Voraussetzung existiert zu diesem ε nun ein $a \in A$, so dass $a > C - \varepsilon$ ist. Wir haben aber $C - \varepsilon = C - (C - \tilde{C}) = \tilde{C}$, also $a > \tilde{C}$. Dann kann aber $\tilde{C}$ keine obere Schranke von A sein. Widerspruch. $\square$

Übungsaufgabe 2.17. Beweisen Sie die folgende Aussage: $M \neq \emptyset$ ist genau dann beschränkt, wenn es ein $C > 0$ gibt, so dass $|x| \leq C$ für alle $x \in M$ gilt.

3. Die natürlichen Zahlen

Sie alle kennen natürlich die natürlichen Zahlen. Aber wie definieren wir diese nur aus unseren 15 Axiomen heraus? Mehr haben wir ja im Moment noch nicht zur Verfügung. Wir führen hier eine Möglichkeit der Definition vor, die darauf basiert, dass wir zumindest die Existenz einer natürlichen Zahl, nämlich der Eins, durch das Axiom (A6) gesichert haben.
Ein Vorteil dieses Zugangs ist, dass wir das wichtige Beweisverfahren der vollständigen Induktion damit leicht herleiten können.

Definition 3.1. *Eine Menge $A \subseteq \mathbb{R}$ heißt eine* Induktionsmenge*, falls gilt*

(a) $1 \in A$ *und*

(b) ist $x \in A$, so ist auch stets $x + 1 \in A$.

Beispiel 3.2. Beispiele von Induktionsmengen sind $\mathbb{R}$ oder $\{1\} \cup [2, \infty)$.

Definition 3.3. *Den Durchschnitt aller Induktionsmengen bezeichnen wir mit $\mathbb{N}$. Das ist die Menge der* natürlichen Zahlen*.*

Wir sammeln ein paar grundlegende Eigenschaften von $\mathbb{N}$.

Satz 3.4. *(a) $\mathbb{N}$ ist eine Induktionsmenge.*

(b) $\mathbb{N}$ ist nicht nach oben beschränkt (Satz von Archimedes)*.*

(c) Für alle $x \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$ mit $x < n$.

(d) Ist A eine Induktionsmenge und $A \subseteq \mathbb{N}$, so ist $A = \mathbb{N}$ (Prinzip der vollständigen Induktion)*.*

Beweis. (a) In jeder Induktionsmenge ist 1 enthalten, also gilt auch $1 \in \mathbb{N}$, da $\mathbb{N}$ der Durchschnitt aller Induktionsmengen ist.

Sei nun $x \in \mathbb{N}$. Dann gilt $x \in A$ für jede Induktionsmenge A. Nach der Definition einer Induktionsmenge ist damit auch $x+1$ in jeder Induktionsmenge enthalten. Also ist $x + 1 \in \mathbb{N}$.

(b) Wir nehmen an, $\mathbb{N}$ wäre nach oben beschränkt. Nach dem Vollständigkeitsaxiom gibt es dann $s := \sup \mathbb{N} \in \mathbb{R}$ und $s - 1$ ist keine obere Schranke von $\mathbb{N}$. Also muss es ein größeres Element $n \in \mathbb{N}$ geben, d. h. $n > s-1$, bzw. $n+1 > s$. Da $\mathbb{N}$ nach (a) eine Induktionsmenge ist, gilt aber $n+1 \in \mathbb{N}$, also gilt $n + 1 \leq s$, da s ja das Supremum von $\mathbb{N}$ ist. Das ist ein Widerspruch.

(c) Wir nehmen an, es gäbe ein $x \in \mathbb{R}$ mit $x \geq n$ für alle $n \in \mathbb{N}$. Dann wäre x eine obere Schranke von $\mathbb{N}$, was im Widerspruch zu (b) steht.

(d) Sei $A \subseteq \mathbb{N}$ eine Induktionsmenge. Da $\mathbb{N}$ der Schnitt aller Induktionsmengen ist, gilt $\mathbb{N} \subseteq A$. Zusammen mit der Voraussetzung $A \subseteq \mathbb{N}$ gilt also $A = \mathbb{N}$. □

Dieses Wissen können wir nun nutzen, um zu zeigen, dass die so definierte Menge $\mathbb{N}$ mit unserer Vorstellung der natürlichen Zahlen übereinstimmt.

Satz 3.5. *(a) Für alle $n \in \mathbb{N}$ gilt $n \geq 1$.*

(b) Für jedes $n \in \mathbb{N}$ ist die Menge $A_n := (\mathbb{N} \cap [1, n]) \cup [n + 1, \infty)$ eine Induktionsmenge.

(c) Es seien $n \in \mathbb{N}$ und $x \in \mathbb{R}$ mit $n < x < n + 1$ gegeben. Dann gilt $x \notin \mathbb{N}$.

Beweis. (a) Wir setzen $A := \{n \in \mathbb{N} : n \geq 1\}$. Dann gilt offensichtlich $1 \in A$ und wenn $n \in A$ ist, so gilt wegen $n \geq 1$ auch $n + 1 \geq 1 + 1 \geq 1$, also ist auch $n + 1 \in A$ und damit A eine Induktionsmenge. Wegen $A \subseteq \mathbb{N}$ gilt damit wegen Satz 3.4 (d) $A = \mathbb{N}$.

(b) Wir setzen $A := \{n \in \mathbb{N} : A_n \text{ ist eine Induktionsmenge}\}$. Der Beweis verläuft nun in drei Schritten:

Induktionsanfang (1 ist in A): Nach Beispiel 3.2 ist $A_1 = \{1\} \cup [2, \infty)$ eine Induktionsmenge. Also ist $1 \in A$.

Induktionsvoraussetzung: Es sei $n \in A$, d. h. wir wählen ein $n \in \mathbb{N}$, so dass A_n eine Induktionsmenge ist.

Induktionsschluss (zeige, dass auch $n + 1 \in A$ gilt): Es ist $A_{n+1} = (\mathbb{N} \cap [1, n + 1]) \cup [n + 2, \infty)$. Also ist $1 \in A_{n+1}$. Sei nun $x \in A_{n+1}$ Dann gilt entweder $x \geq n + 2$ oder $1 \leq x \leq n + 1$, wobei im zweiten Fall $x \in \mathbb{N}$ ist. Im ersten Fall ist $x+1 \geq n+3 \geq n+2$, also haben wir sofort $x+1 \in A_{n+1}$. Im zweiten Fall machen wir uns zunutze, dass A_n eine Induktionsmenge ist (Induktionsvoraussetzung!) und deshalb $\mathbb{N} \subseteq A_n$ gilt. Das liefert uns, dass $x \in A_n$ ist. Damit ist entweder $1 \leq x \leq n$ oder $x \geq n + 1$, d. h. entweder wir haben $2 \leq x + 1 \leq n + 1$ oder $x + 1 \geq n + 2$. Also ist $x + 1 \in A_{n+1}$.

(c) Wir nehmen an, es wäre doch $x \in \mathbb{N}$. Da nach (b) $\mathbb{N} \subseteq A_n$ gilt, hätten wir dann $x \in A_n$. Das impliziert aber $x \leq n$ oder $x \geq n + 1$. Widerspruch. □

Die Menge A wird üblicherweise bei einem Induktionsbeweis nicht mehr erwähnt, da die Methode immer die gleiche ist. Wir wollen uns das an einem weiteren, sehr typischen Beispiel für einen Induktionsbeweis anschauen.

Satz 3.6. *Für alle $n \in \mathbb{N}$ gilt*

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

Beweis.
Induktionsanfang: Es gilt $1 = 1 \cdot (1+1)/2$, also ist die Aussage für $n = 1$ richtig.
Induktionsvoraussetzung: Die Aussage des Satzes sei für ein bestimmtes $n \in \mathbb{N}$ richtig, d. h. es gilt

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

für dieses $n \in \mathbb{N}$.
Induktionsschluss: Mit Hilfe der Induktionsvoraussetzung gilt

$$\begin{aligned} 1 + 2 + \cdots + n + (n+1) &= \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2} = \frac{(n+1)((n+1)+1)}{2}. \end{aligned}$$

Also stimmt die Aussage auch für $n + 1$. □

Bemerkung 3.7. Die Pünktchen-Schreibweise im obigen Beweis für die Summation von n Zahlen ist reichlich schwerfällig und führt oft zu unpräzisen Formulierungen. Deshalb hat sich dafür eine sehr praktische Schreibweise eingebürgert. Sind $n, N \in \mathbb{Z}$ mit $n < N$ und reelle Zahlen $a_n, a_{n+1}, \ldots, a_N$ gegeben, so schreiben wir

$$\sum_{k=n}^{N} a_k := a_n + a_{n+1} + \cdots + a_{N-1} + a_N$$

mit dem sogenannten *Summenzeichen*. Die Aussage von Satz 3.6 lässt sich damit z. B. so hinschreiben:

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

Im nächsten Satz zeigen wir, dass $\mathbb{N}$ eine sogenannte *wohlgeordnete* Menge ist, d. h. jede nicht-leere Teilmenge besitzt ein Minimum.

Satz 3.8 (Wohlordnungsprinzip)**.** *Ist $M \neq \emptyset$ eine Teilmenge von $\mathbb{N}$, so existiert* $\min M$.

Beweis. Nach Satz 3.5 (a) ist 1 eine untere Schranke von $\mathbb{N}$ und damit auch von M. Wir setzen $A := \{\gamma \in \mathbb{N} : \gamma \text{ ist untere Schranke von } M\}$. Wir zeigen nun, dass $A \neq \mathbb{N}$ ist. Sei dazu ein $x \in M$ gegeben. Dann gilt nach der Definition von M, dass $x \geq \gamma$ für alle $\gamma \in A$ ist. Das ist für $A = \mathbb{N}$ nach Satz 3.4 (b) nicht möglich.

Also ist A keine Induktionsmenge. Da $1 \in A$ gilt, bedeutet dies, dass es ein $\gamma_0 \in A$ geben muss, für das $\gamma_0 + 1 \notin A$ gilt. Da γ_0 in A und damit in $\mathbb{N}$ liegt, ist auch $\gamma_0 + 1 \in \mathbb{N}$, aber $\gamma_0 + 1$ ist keine untere Schranke von M (sonst wäre es in A). Das heißt, es gibt ein $n_0 \in M$ mit $n_0 < \gamma_0 + 1$. Zusammengefasst haben wir nun $\gamma_0 \leq n_0 < \gamma_0 + 1$. Da aber alle drei beteiligten Zahlen in dieser Ungleichungskette in $\mathbb{N}$ liegen, gilt nach Satz 3.5 (c) $n_0 = \gamma_0$. Damit ist n_0 eine untere Schranke von M (da γ_0 eine ist) und es ist $n_0 \in M$. Somit folgt $n_0 = \min M$. $\square$

Definition 3.9. *Wir definieren*

$$\begin{aligned}
\mathbb{N}_0 &:= \mathbb{N} \cup \{0\}, \\
\mathbb{Z} &:= \mathbb{N}_0 \cup \{-n : n \in \mathbb{N}\} \qquad \text{(ganze Zahlen)}, \\
\mathbb{Q} &:= \left\{ \frac{m}{n} : m \in \mathbb{Z},\ n \in \mathbb{N} \right\} \text{ (rationale Zahlen)}.
\end{aligned}$$

4. Folgen und Abzählbarkeit

Ein grundlegendes Hilfsmittel der Analysis sind die Folgen.

Definition 4.1. *Sei X eine beliebige nicht-leere Menge. Eine Funktion $a : \mathbb{N} \to X$ heißt eine* Folge in X.
Bei Folgen schreibt man traditionell a_n statt $a(n)$, $n \in \mathbb{N}$, für das n-te Folgenglied.
Die gesamte Folge wird üblicherweise mit (a_n) oder $(a_n)_{n=1}^{\infty}$ oder $(a_n)_{n\in\mathbb{N}}$ oder $(a_n)_{n\geq 1}$ oder $(a_1, a_2, a_3, \dots)$ bezeichnet.

Beispiel 4.2. (a) Ist $X = \mathbb{R}$ (der für uns im Folgenden interessanteste Fall), so spricht man von einer *reellen Folge* oder *Zahlenfolge*. Ein Beispiel einer solchen Folge ist

$$a_n = \frac{1}{n}, \quad n \in \mathbb{N}, \text{ bzw. } (a_n) = \Big(\frac{1}{n}\Big) = \Big(1, \frac{1}{2}, \frac{1}{3}, \dots\Big).$$

(b) Für $X = \{0, 1\}$ ist $(a_n) = (1, 0, 1, 0, 1, 0, \dots)$ eine Folge in $\{0, 1\}$.

Mit Hilfe von Folgen können wir nun Begriffe einführen um die „Größe“ unendlich großer Mengen zu klassifizieren. Ein erster Schritt zur Beschreibung der Unendlichkeit.

Definition 4.3. *Sei X eine beliebige nicht-leere Menge.*

(a) X heißt endlich, *wenn ein $n \in \mathbb{N}$ und eine Funktion $f : \{1, 2, \dots, n\} \to X$ existieren, so dass f surjektiv ist.*

(In diesem Fall ist $X = \{f(1), f(2), \dots, f(n)\}$.)

(b) X heißt unendlich, *wenn X nicht endlich ist.*

(c) X heißt abzählbar, *wenn eine Folge (a_n) in X existiert, die surjektiv ist, d. h. es gilt $X = \{a_1, a_2, a_3, \dots\}$.*

(d) X heißt abzählbar unendlich, *wenn X unendlich und abzählbar ist.*

(e) X heißt überabzählbar, *wenn X nicht abzählbar ist.*

Bemerkung 4.4. Anschaulich bedeutet Abzählbarkeit, dass man die Elemente der Menge X mit den natürlichen Zahlen durchnummerieren kann.
Beachten Sie, dass nach dieser Definition auch endliche Mengen abzählbar sind.

Beispiel 4.5. (a) $\mathbb{N}$ ist abzählbar (unendlich), denn $\mathbb{N} = \{a_1, a_2, a_3, \dots\}$ mit $(a_n) = (n)$.

(b) $\mathbb{Z}$ ist abzählbar, denn $\mathbb{Z} = \{a_1, a_2, a_3, \dots\}$ mit $a_1 = 0,\ a_2 = 1,\ a_3 = -1,\ a_4 = 2, a_5 = -2, \dots$, also $a_{2n} = n$ und $a_{2n-1} = -n + 1$.

(c) $\mathbb{Q}$ ist abzählbar. Das mag zuerst verblüffend sein, doch man kann die positiven Brüche tatsächlich nach folgendem Schema ordnen:

$$
\begin{array}{cccccccccc}
1 & \rightarrow & 2 & & 3 & \rightarrow & 4 & & 5 & \dots \\
 & \swarrow & & \nearrow & & \swarrow & & & & \\
\frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & \frac{4}{2} & & \frac{5}{2} & \dots \\
\downarrow & \nearrow & & \swarrow & & & & & & \\
\frac{1}{3} & & \frac{2}{3} & & \frac{3}{3} & & \frac{4}{3} & & \frac{5}{3} & \dots \\
 & \swarrow & & & & & & & & \\
\frac{1}{4} & & \frac{2}{4} & & \frac{3}{4} & & \frac{4}{4} & & \frac{5}{4} & \dots \\
 & & & & & & & & & \\
\frac{1}{5} & \dots & & & & & & & & \\
\vdots & & & & & & & & &
\end{array}
$$

Durchnummerieren in Pfeilrichtung liefert also die Abzählung $\{x \in \mathbb{Q} : x > 0\} = \{a_1, a_2, a_3, \dots\}$. Setzt man nun $b_1 = 0,\ b_{2n} = a_n,\ b_{2n+1} = -a_n,\ n \in \mathbb{N}$, so gilt $\mathbb{Q} = \{b_1, b_2, b_3, \dots\}$.

(d) $\mathbb{R}$ ist überabzählbar (Beweis später).

(e) Die Menge X aller Folgen in $\{0, 1\}$ ist überabzählbar. Um das zu beweisen, nehmen wir an, X wäre abzählbar unendlich, d. h. $X = \{f_1, f_2, f_3, \dots\}$, wobei $f_j = (a_{j1}, a_{j2}, a_{j3}, \dots)$ und $a_{jk} \in \{0, 1\}$ für alle $j, k \in \mathbb{N}$.

Wir definieren nun eine Folge in $\{0, 1\}$ wie folgt:

$$
a_j := \begin{cases} 1, & \text{falls } a_{jj} = 0, \\ 0, & \text{falls } a_{jj} = 1 \end{cases}
$$

für jedes $j \in \mathbb{N}$. Dann ist $(a_j)_{j=1}^{\infty}$ in X, also gibt es nach Annahme ein $m_0 \in \mathbb{N}$, so dass $(a_j) = f_{m_0}$ gilt. Das heißt aber, dass $a_{m_0} = a_{m_0 m_0}$ ist, ein Widerspruch, denn wir haben die Folge (a_j) gerade so konstruiert, dass dies nicht gilt.

Das Beweisverfahren dieses Überabzählbarkeitsbeweises heißt *Cantorsches Diagonalverfahren.*

Satz 4.6. *Es sei A eine abzählbare Menge und $B \subseteq A$ sei nicht-leer. Dann ist auch B abzählbar.*

Beweis. Da A abzählbar ist, gibt es eine Folge in A, so dass $A = \{a_1, a_2, a_3, \dots\}$ gilt. Wähle nun ein beliebiges $b \in B$ fest aus (B ist nicht-leer!). Damit definieren wir eine Folge

$$b_n = \begin{cases} b, & \text{falls } a_n \notin B, \\ a_n, & \text{falls } a_n \in B, \end{cases} \qquad n \in \mathbb{N}.$$

Sei nun $x \in B$. Dann gilt nach Voraussetzung $x \in A$, also gibt es ein $m \in \mathbb{N}$, für das $x = a_m$ ist. Da damit $a_m \in B$ ist, haben wir $b_m = a_m = x$ (so war (b_n) definiert), also ist $x \in \{b_1, b_2, b_3, \dots\}$. Damit haben wir $B \subseteq \{b_1, b_2, b_3, \dots\}$ gezeigt. Da aber auch offensichtlich $B \supseteq \{b_1, b_2, b_3, \dots\}$ gilt, ist damit $B = \{b_1, b_2, b_3, \dots\}$. $\square$

Vereinigt man endlich viele abzählbare Mengen, so ist es nicht verwunderlich, dass man wieder eine abzählbare Menge erhält. Nicht mehr so offensichtlich ist, dass das auch für Vereinigungen von abzählbar unendlich vielen abzählbaren Mengen gilt. Das ist der Inhalt des folgenden abschließenden Satzes in diesem Kapitel.

Satz 4.7. *Es seien $X_1, X_2, X_3, \dots$ abzählbar viele abzählbare Mengen. Dann ist auch die Menge $\bigcup_{j=1}^{\infty} X_j$ abzählbar.*

Beweis. Übung.

5. Binomialkoeffizienten, Fakultäten und Wurzeln

Definition 5.1. *(a) Für $a \in \mathbb{R}$ und $n \in \mathbb{N}$ schreiben wir für die* Potenzen

$$a^n := \underbrace{a \cdot a \cdot a \cdot \ldots \cdot a}_{n \text{ Faktoren}}.$$

Weiter gilt $a^0 := 1$.

Ist außerdem $a \neq 0$, so schreiben wir

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

(b) Für $n \in \mathbb{N}$ definieren wir die Fakultät *von n als*

$$n! := 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$$

und wir vereinbaren $0! := 1$.

(c) Schließlich ist für $k, n \in \mathbb{N}_0$ mit $0 \leq k \leq n$ der Binomialkoeffizient *gegeben durch*

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

Satz 5.2. *(a) Ist $x \geq -1$ und $n \in \mathbb{N}$ so gilt*

$$(1+x)^n \geq 1 + nx \qquad \text{(Bernoullische Ungleichung).}$$

(b) Für die Binomialkoeffizienten gelten für alle $n \in \mathbb{N}$ und alle $k \in \mathbb{N}$ mit $1 \leq k \leq n$ die folgenden Identitäten:

$$\binom{n}{0} = 1 = \binom{n}{n}, \qquad \binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

(c) Für alle $a, b \in \mathbb{R}$ und alle $n \in \mathbb{N}_0$ gilt

$$\begin{aligned} a^{n+1} - b^{n+1} &= (a-b)(a^n + a^{n-1}b + a^{n-2}b^2 + \cdots + ab^{n-1} + b^n) \\ &= (a-b)\sum_{k=0}^{n} a^{n-k}b^k. \end{aligned}$$

(d) *Für alle $a, b \in \mathbb{R}$ und alle $n \in \mathbb{N}$ gilt die* Binomialformel

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

Beweis. (a) Übung: Führen Sie einen Induktionsbeweis.

(b) Es gilt $\binom{n}{0} = \dfrac{n!}{0!(n-0)!} = \dfrac{n!}{n!} = 1$ und $\binom{n}{n} = \dfrac{n!}{n!(n-n)!} = \dfrac{n!}{n!} = 1$ sowie

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

(c) Es gilt

$$\begin{aligned}(a-b)\sum_{k=0}^{n} a^{n-k}b^k &= \sum_{k=0}^{n} a^{n-k+1}b^k - \sum_{k=0}^{n} a^{n-k}b^{k+1} \\ &= a^{n+1} + \sum_{k=1}^{n} a^{n-k+1}b^k - \sum_{k=0}^{n-1} a^{n-k}b^{k+1} - b^{n+1} \\ &= a^{n+1} + \sum_{k=0}^{n-1} a^{n-k}b^{k+1} - \sum_{k=0}^{n-1} a^{n-k}b^{k+1} - b^{n+1} \\ &= a^{n+1} - b^{n+1}.\end{aligned}$$

(d) Übung. □

Zur Definition von Wurzeln benötigen wir zunächst den folgenden Hilfssatz.

Lemma 5.3. *Sind $x, y \geq 0$ und $n \in \mathbb{N}$, so gilt $x \leq y$ genau dann, wenn $x^n \leq y^n$ ist.*

Beweis. Übungsaufgabe.

Satz 5.4. *Es sei $a \geq 0$ und $n \in \mathbb{N}$. Dann gibt es genau ein $x \in \mathbb{R}$, so dass $x \geq 0$ und $x^n = a$ gilt.*

Beweis. Wir zeigen zunächst die Eindeutigkeit. Seien also $x, y \geq 0$ gegeben, so dass $x^n = a = y^n$ gilt. Dann gilt insbesondere $x^n \leq y^n$ und $y^n \leq x^n$ und mit Lemma 5.3 folgt dann $x \leq y$ und $y \leq x$, also $x = y$.

Für die Existenz stellen wir zunächst fest, dass die Sache für $a = 0$ durch $x = 0$ gelöst wird. Sei also $a > 0$. Wir betrachten die Menge $M := \{x \in \mathbb{R} : x \geq 0 \text{ und } x^n \leq a\}$. Dann ist in jedem Fall $0 \in M$, also $M \neq \emptyset$. Da $n \in \mathbb{N}$ ist, gilt $a \leq na \leq 1+na$ und daher mit der Bernoullischen Ungleichung, vgl. Satz 5.2 (a), auch $a \leq (1+a)^n$. Damit folgern wir nun für alle $x \in M$ die Abschätzung $x^n \leq a \leq (1+a)^n$, aus der mit Lemma 5.3 sofort $x \leq 1+a$ folgt. Also ist M nach oben beschränkt. Nach dem Vollständigkeitsaxiom existiert damit $s := \sup M$, unser Kandidat für die Wurzel. Wir nehmen also an, es gelte $s^n \neq a$ und unterscheiden zwei Fälle.
1. Fall $s^n < a$: Für jedes $m \in \mathbb{N}$ gilt nach Satz 5.2 (d)

$$\left(s + \frac{1}{m}\right)^n = \sum_{k=0}^{n} \binom{n}{k} s^{n-k} \frac{1}{m^k} = s^n + \sum_{k=1}^{n} \binom{n}{k} s^{n-k} \underbrace{\frac{1}{m^k}}_{\leq 1/m}$$
$$\leq s^n + \frac{1}{m} \sum_{k=1}^{n} \binom{n}{k} s^{n-k} =: s^n + \frac{1}{m}\alpha.$$

Man beachte, dass das so definierte α in jedem Fall größer oder gleich Null ist. Nach dem Satz von Archimedes (Satz 3.4 (b)) und da $a \neq s^n$ ist, existiert ein $m \in \mathbb{N}$, so dass $m \geq \alpha/(a - s^n)$ gilt. Da nach Annahme $a - s^n > 0$ gilt, ist dann aber $\alpha/m < a - s^n$ und weiter $s^n + \alpha/m < a$. Also gilt für dieses m nun $(s + 1/m)^n < a$, was uns $s + 1/m \in M$ liefert. Da s das Supremum dieser Menge war, gilt damit $s + 1/m \leq s$, d. h. $1/m \leq 0$. Widerspruch!
2. Fall $s^n > a$: Wir rechnen für jedes $m \in \mathbb{N}$ (man beachte, dass $s \neq 0$ sein muss, da sonst $s^n = 0 < a$ wäre)

$$\left(s - \frac{1}{m}\right)^n = \left[s\left(1 - \frac{1}{ms}\right)\right]^n = s^n \left(1 - \frac{1}{ms}\right)^n.$$

Wählen wir unser $m \geq 1/s$ (die Möglichkeit hierzu sichert wieder der Satz von Archimedes), dann gilt $-1/(ms) \geq -1$, wir können folglich die Bernoullische Ungleichung aus Satz 5.2 (a) anwenden und erhalten

$$\left(s - \frac{1}{m}\right)^n \geq s^n \left(1 - \frac{n}{ms}\right).$$

Wir machen nun bei Bedarf unser m noch einmal größer, damit $m \geq 1/s$ und $m > (ns^n)/(s(s^n - a))$ gilt. Formen wir die zweite Ungleichung ein wenig um und beachten dabei wieder, dass nach Annahme $s^n - a > 0$ gilt:

$$m > \frac{ns^n}{s(s^n - a)} \iff (s^n - a)m > \frac{n}{s}s^n \iff -a > \frac{n}{ms}s^n - s^n$$
$$\iff a < s^n \left(1 - \frac{n}{ms}\right),$$

so sehen wir, dass für ein solches $m \in \mathbb{N}$ die Ungleichung $(s - 1/m)^n > a$ gilt. Da s das Supremum von M ist, ist $s - 1/m$ keine obere Schranke von M, also gibt es ein $x_0 \in M$ mit $x_0 > s - 1/m$. Nach Lemma 5.3 gilt dann auch

$$x_0^n \geq \left(s - \frac{1}{m}\right)^n > a,$$

was im Widerspruch zu $x_0 \in M$ steht. □

Definition 5.5. *Zu gegebenen $a \geq 0$ und $n \in \mathbb{N}$ bezeichnen wir die nach obigem Satz eindeutig existierende Zahl x, für die $x^n = a$ gilt, als n-*te Wurzel von *a und schreiben*

$$x = \sqrt[n]{a} \quad \textit{oder} \quad x = a^{1/n}.$$

Ist $n = 2$, so sagt man einfach Wurzel von *a und schreibt kurz $x = \sqrt{a}$. Außerdem setzen wir wieder*

$$a^{-1/n} = \frac{1}{a^{1/n}}.$$

An dieser Stelle sind zwei Warnungen angebracht. Erstens ziehen wir hier nur Wurzeln aus nicht-negativen Zahlen und es ist stets $\sqrt[n]{a} \geq 0$, auch wenn z. B. die Gleichung $x^2 = 3$ *zwei* reelle Lösungen besitzt, nämlich $\sqrt{3}$ und $-\sqrt{3}$. Zweitens gilt deshalb $\sqrt{b^2} = |b|$. Die Betragsstriche werden sehr gerne vergessen.
Wir wissen nun für $a > 0$, was $a^{\pm n}$ und $a^{\pm 1/n}$ ist. Im nächsten Schritt wollen wir uns mit a^q für beliebige rationale Zahlen q befassen. Ist $q = m/n$ mit $m, n \in \mathbb{N}$, so ist es naheliegend $a^q := (\sqrt[n]{a})^m$ zu definieren. Das ist aber etwas voreilig, denn wir sollten uns zunächst überlegen, dass das Ergebnis nicht von der speziellen Wahl von n und m abhängt.
Seien also $m, n, p, r \in \mathbb{N}$, so dass $q = m/n = p/r$ ist. Dann gilt $mr = np$ und es folgt für jedes $a \geq 0$

$$\begin{aligned} \left((\sqrt[n]{a})^m\right)^r &= (\sqrt[n]{a})^{mr} = (\sqrt[n]{a})^{np} = \left((\sqrt[n]{a})^n\right)^p = a^p \quad \text{und} \\ \left((\sqrt[r]{a})^p\right)^r &= (\sqrt[r]{a})^{pr} = \left((\sqrt[r]{a})^r\right)^p = a^p. \end{aligned}$$

Also ist dank der Eindeutigkeit der Wurzel $(\sqrt[n]{a})^m = (\sqrt[r]{a})^p$. Damit ist die folgende Definition gerechtfertigt.

Definition 5.6. *Es sei $a \geq 0$, $q \in \mathbb{Q}$ mit $q > 0$ und $q = m/n$ ($m, n \in \mathbb{N}$). Dann setzen wir*

$$a^q := (\sqrt[n]{a})^m,$$

und falls $a > 0$ gilt, definieren wir für $q < 0$

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

Es gelten die bekannten Rechenregeln für rationale Exponenten:

- $a^p a^q = a^{p+q}$
- $a^p b^p = (ab)^p$
- $(a^p)^q = a^{pq}$
- $\frac{a^p}{a^q} = a^{p-q}$
- $\frac{a^p}{b^p} = \left(\frac{a}{b}\right)^p$

Teil II.

Folgen und Reihen

6. Konvergente Folgen

Wir wollen uns nun dem zentralen Thema der Analysis zuwenden, der mathematisch exakten Behandlung des unendlich Kleinen und unendlich Großen. Beispielsweise kann es darum gehen, unendlich viele Zahlen aufzuaddieren, wie in der unendlichen Summe

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots,$$

der wir im Folgenden einen exakten Sinn geben werden.
Wie schon in der OWO-Vorlesung deutlich geworden ist, können hier sehr unintuitive Dinge passieren, so dass anschauliche Argumentationen uns schnell in die Irre führen können. Unser Ziel wird also zunächst sein, eine exakte mathematische Definition für solche Grenzwertfragen zu geben. Diese Aufgabe wollen wir in diesem für alles weitere zentralen Kapitel angehen.

Definition 6.1. *Eine Folge (a_n) in $\mathbb{R}$ heißt* konvergent*, wenn es ein $a \in \mathbb{R}$ gibt, so dass für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert mit*

$$|a_n - a| < \varepsilon \textit{ für alle } n \geq n_0.$$

In diesem Fall heißt a der Grenzwert *oder* Limes *der Folge (a_n). Man schreibt dafür*

$$a = \lim_{n\to\infty} a_n \quad \textit{oder} \quad a_n \to a \ (n \to \infty).$$

Ist (a_n) nicht konvergent, so heißt die Folge divergent*.*

Für eine Umformulierung dieser Definition benötigen wir die folgenden Begriffe.

Definition 6.2. *(a) Seien $x_0 \in \mathbb{R}$ und $\varepsilon > 0$ gegeben. Dann heißt die Menge*

$$U_\varepsilon(x_0) := \{x \in \mathbb{R} : |x - x_0| < \varepsilon\} = (x_0 - \varepsilon, x_0 + \varepsilon)$$

ε-Umgebung *von x_0.*

(b) Es sei $A(n)$ eine für jedes $n \in \mathbb{N}$ definierte Aussage. Wir sagen, $A(n)$ gilt für fast alle $n \in \mathbb{N}$ *oder synonym* fast überall*, wenn es ein $m \in \mathbb{N}$ gibt, so dass $A(n)$ für alle $n \geq m$ richtig ist. Die Bezeichnung „fast überall“ wird auch gerne mit „f.ü.“ abgekürzt.*

Damit können wir nun sagen, dass $\lim_{n\to\infty} a_n = a$ ist, genau dann wenn

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} : a_n \in U_\varepsilon(a) \; \forall n \geq n_0$$

bzw. in Worte gefasst:

$$\text{Für alle } \varepsilon > 0 \text{ gilt } a_n \in U_\varepsilon(a) \text{ für fast alle } n \in \mathbb{N}.$$

Aus der letzten Formulierung ist gut ersichtlich, dass es für die Konvergenz einer Folge auf endlich viele Folgenglieder nicht ankommt. Entscheidend ist, was die Folge „janz weit draußen" treibt. Formalisieren kann man diese Beobachtung so:

Bemerkung 6.3. Sind (a_n) und (b_n) reelle Folgen mit $a_n = b_n$ für fast alle $n \in \mathbb{N}$, dann ist (a_n) konvergent genau dann, wenn (b_n) konvergiert, und in diesem Fall gilt $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

Beispiel 6.4. (a) Sei $a_n = 1/n$, $n \in \mathbb{N}$.

Behauptung: (a_n) ist konvergent und es gilt $\lim_{n\to\infty} a_n = 0$.

Beweis. Sei $\varepsilon > 0$. Nach dem Satz von Archimedes (Satz 3.4 (b)) gibt es ein $n_0 \in \mathbb{N}$, so dass $n_0 > 1/\varepsilon$ ist. Damit gilt $1/n_0 < \varepsilon$ und es ist für alle $n \geq n_0$

$$|a_n - a| = |a_n - 0| = |a_n| = \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon. \qquad \square$$

(b) Sei $a_n = (-1)^n/n$, $n \in \mathbb{N}$.

Behauptung: (a_n) ist konvergent und $\lim_{n\to\infty} a_n = 0$.

Beweis. Es gilt $|a_n - 0| = |(-1)^n/n| = 1/n$. Wir können also ab jetzt den Beweis von Beispiel (a) übernehmen. $\square$

(c) Sei $a_n = (-1)^n$, $n \in \mathbb{N}$.

Behauptung: Die Folge (a_n) divergiert.

Beweis. Wir nehmen an, es gäbe ein $a \in \mathbb{R}$ mit $a_n \to a$ $(n \to \infty)$. Dann gibt es zu $\varepsilon = 1$ ein $n_0 \in \mathbb{N}$, so dass für jedes $n \geq n_0$ die Ungleichung $|a_n - a| < 1$ gilt. Für $n \geq n_0$ gilt dann aber mit Hilfe der Dreiecksungleichung

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

Also folgt $2 < 2$, ein Widerspruch. $\square$

(d) Sei $a_n = \dfrac{n^2 + 2n - 1}{n^2 + 2}$, $n \in \mathbb{N}$.

Behauptung: (a_n) konvergiert und $\lim_{n\to\infty} a_n = 1$.

Beweis. Es gilt

$$|a_n - 1| = \left|\frac{n^2 + 2n - 1 - n^2 - 2}{n^2 + 2}\right| = \frac{|2n - 3|}{n^2 + 2} \leq \frac{|2n - 3|}{n^2} \leq \frac{2n + 3}{n^2},$$

wobei wir bei der letzten Abschätzung die Dreiecksungleichung angewendet haben. Nun verwenden wir noch, dass für alle $n \in \mathbb{N}$ gilt $2n+3 \leq 2n+3n = 5n$ und erhalten damit

$$|a_n - 1| \leq \frac{5n}{n^2} = \frac{5}{n}.$$

Sei nun $\varepsilon > 0$. Nach dem Satz von Archimedes existiert analog zu (a) ein $n_0 \in \mathbb{N}$, so dass $n_0 > 5/\varepsilon$ gilt. Dann haben wir nach obiger Abschätzung für alle $n \geq n_0$

$$|a_n - 1| \leq \frac{5}{n} \leq \frac{5}{n_0} < \varepsilon. \qquad \square$$

Definition 6.5. *Eine reelle Folge* (a_n) *heißt* beschränkt*, wenn die Menge der Folgenglieder* $\{a_n : n \in \mathbb{N}\} = \{a_1, a_2, a_3, \dots\}$ *beschränkt ist. Diese Menge besitzt damit ein Infimum und ein Supremum. Wir schreiben dafür*

$$\sup_{n\in\mathbb{N}} a_n := \sup_{n=1}^{\infty} a_n := \sup\{a_n : n \in \mathbb{N}\},$$
$$\inf_{n\in\mathbb{N}} a_n := \inf_{n=1}^{\infty} a_n := \inf\{a_n : n \in \mathbb{N}\}.$$

Beachten Sie, dass eine Folge (a_n) genau dann beschränkt ist, wenn ein $C \geq 0$ existiert, so dass $|a_n| \leq C$ für alle $n \in \mathbb{N}$ gilt (vgl. Übungsaufgabe 2.17).

Satz 6.6. *Sei* (a_n) *eine konvergente Folge in* $\mathbb{R}$. *Dann gilt:*

(a) Der Limes von (a_n) *ist eindeutig bestimmt.*

(b) (a_n) *ist beschränkt.*

Warnung 6.7. Die Umkehrung von Teil (b) ist *falsch*! Es gibt durchaus beschränkte Folgen, die nicht konvergieren, vgl. Beispiel 6.4 (c).

Beweis von Satz 6.6. (a) Es sei $a \in \mathbb{R}$ so, dass $a_n \to a$ $(n \to \infty)$ und es sei $b \in \mathbb{R}$ mit $b \neq a$. Wir wollen zeigen, dass b kein Limes von (a_n) sein kann. Wir setzen dazu $\varepsilon := |a-b|/2$. Dann ist $\varepsilon > 0$ und es gilt $U_\varepsilon(a) \cap U_\varepsilon(b) = \emptyset$. Nach der Definition des Grenzwertes existiert aber ein $n_0 \in \mathbb{N}$, so dass $a_n \in U_\varepsilon(a)$ für alle $n \geq n_0$ gilt. Also gilt $a_n \in U_\varepsilon(b)$ nur für höchstens endlich viele $n \in \mathbb{N}$. Damit kann (a_n) nicht gegen b konvergieren.

(b) Nach der Definition der Konvergenz existiert zu $\varepsilon = 1$ ein $n_0 \in \mathbb{N}$, so dass $|a_n - a| < 1$ für alle $n \geq n_0$ gilt. Wir setzen

$$C := \max\{|a_1|, |a_2|, |a_3|, \dots, |a_{n_0-1}|, 1 + |a|\}.$$

Dann gilt zum einen für alle $n < n_0$ sofort $|a_n| \leq C$ und zum anderen auch für alle $n \geq n_0$, denn

$$|a_n| = |a_n - a + a| \leq |a_n - a| + |a| < 1 + |a| \leq C.$$

Zusammengenommen gilt also $|a_n| \leq C$ für alle $n \in \mathbb{N}$ und somit die Behauptung. $\square$

Der Inhalt des folgenden Satzes enthält äußerst wichtige Rechenregeln für Grenzwerte von Folgen, die sogenannten *Grenzwertsätze*. Mit Hilfe dieser Regeln ist es oft möglich, die Frage nach Konvergenz von komplizierten Folgen auf die Untersuchung einiger einfacher Folgen zurückzuführen.

Satz 6.8. *Es seien (a_n), (b_n) und (c_n) Folgen in $\mathbb{R}$. Dann gilt:*

(a) *(a_n) konvergiert gegen $a \in \mathbb{R}$ genau dann, wenn die Folge $|a_n - a|$, $n \in \mathbb{N}$, gegen Null konvergiert.*

(b) *Gilt $|a_n - a| \leq \alpha_n$ für fast alle $n \in \mathbb{N}$ und ist die Folge (α_n) konvergent mit $\lim_{n\to\infty} \alpha_n = 0$, so ist (a_n) konvergent und es gilt $\lim_{n\to\infty} a_n = a$.*

(c) *Ist $\lim_{n\to\infty} a_n = a$, so gilt $\lim_{n\to\infty} |a_n| = |a|$.*

(d) *Gilt $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$, so folgt*

i) $\lim_{n\to\infty}(a_n + b_n) = a + b$.

ii) $\lim_{n\to\infty}(\alpha a_n) = \alpha a$ *für alle* $\alpha \in \mathbb{R}$.

iii) $\lim_{n\to\infty}(a_n \cdot b_n) = a \cdot b$.

iv) *Ist zusätzlich $a_n \neq 0$ für alle $n \in \mathbb{N}$ und $a \neq 0$, so ist $\lim_{n\to\infty} 1/a_n = 1/a$.*

(e) *Ist $a_n \leq b_n$ für fast alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} a_n = a$ sowie $\lim_{n\to\infty} b_n = b$, so folgt $a \leq b$.*

(f) *Ist $a_n \leq c_n \leq b_n$ für fast alle $n \in \mathbb{N}$ und sind (a_n) und (b_n) konvergent mit $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = a$, so ist auch die Folge (c_n) konvergent und es gilt $\lim_{n\to\infty} c_n = a$* (Sandwich-Theorem).

Beweis. (a) Übungsaufgabe.

(b) Nach Voraussetzung gibt es ein $m \in \mathbb{N}$, so dass $|a_n - a| \leq \alpha_n$ für alle $n \geq m$ gilt. Sei nun $\varepsilon > 0$. Dann existiert nach Voraussetzung ein $n_1 \in \mathbb{N}$, so dass $0 \leq \alpha_n < \varepsilon$ für alle $n \geq n_1$. Wir wählen $n_0 := \max\{m, n_1\}$. Dann gilt für alle $n \geq n_0$ die Abschätzung $|a_n - a| \leq \alpha_n < \varepsilon$.

(c) Nach der umgekehrten Dreiecksungleichung gilt

$$\big||a_n| - |a|\big| \leq |a_n - a| =: \alpha_n.$$

Da (a_n) gegen a konvergiert, konvergiert nach (a) die Folge (α_n) gegen Null. Also können wir mit (b) folgern, dass $(|a_n|)$ gegen $|a|$ konvergiert.

(d) (d)i), (d)ii), (d)iii) Übungsaufgabe.

iv) Für die Behandlung dieser Aussage müssen wir zuerst sicherstellen, dass die Folge (a_n) einen echten „Sicherheitsabstand" zur Null einhält. Dazu bemerken wir, dass nach (c) die Folge $(|a_n|)$ gegen $|a|$ konvergiert. Da nach Voraussetzung $|a| \neq 0$ ist, gibt es zu $\varepsilon := |a|/2 > 0$ ein $m \in \mathbb{N}$, so dass $||a_n| - |a|| < |a|/2$ für alle $n \geq m$ gilt. Somit ist der Abstand von $|a_n|$ zu $|a|$ kleiner als $|a|/2$, insbesondere muss also $|a_n| > |a|/2$ für alle diese n gelten.

Aus dieser Überlegung erhalten wir nun, dass $1/|a_n| \leq 2/|a|$ für alle $n \geq m$ gilt und wir damit für diese n folgendermaßen abschätzen können:

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \left|\frac{a_n - a}{a_n a}\right| = \frac{|a_n - a|}{|a_n||a|} \leq \frac{2|a_n - a|}{|a|^2} = \frac{2}{a^2}|a_n - a| =: \alpha_n.$$

Nach (a) und (d)ii) gilt nun wieder $\alpha_n \to 0$ $(n \to \infty)$ und damit folgt mit Hilfe von (b) die Behauptung.

(e) Wir nehmen an, es wäre $a > b$. Dann ist $\varepsilon := (a-b)/2 > 0$ und dank der Konvergenz von (a_n) und (b_n) gibt es nun ein $n_1 \in \mathbb{N}$, so dass $b_n \in U_\varepsilon(b) = (b - \varepsilon, b + \varepsilon)$ und $a_n \in U_\varepsilon(a) = (a - \varepsilon, a + \varepsilon)$ für alle $n \geq n_1$ gilt. Da

$$b + \varepsilon = b + \frac{a-b}{2} = \frac{a}{2} + \frac{b}{2} = a - \frac{a-b}{2} = a - \varepsilon$$

gilt, haben wir also für diese n auch $b_n < a_n$ im Widerspruch zur Voraussetzung.

(f) Sei $\varepsilon > 0$. Dann gibt es ein $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ sowohl $|a_n - a| < \varepsilon$ als auch $|b_n - a| < \varepsilon$ gilt. Hieraus und aus der Voraussetzung folgern wir für alle diese n

$$a - \varepsilon < a_n \leq c_n \leq b_n < a + \varepsilon.$$

Also ist $a - \varepsilon < c_n < a + \varepsilon$ oder, anders ausgedrückt, $-\varepsilon < c_n - a < \varepsilon$, d. h. $|c_n - a| < \varepsilon$ für alle $n \geq n_0$ und damit konvergiert die Folge (c_n) gegen a. □

Wir wollen nun an zwei Beispielen zeigen, wie mit Hilfe dieses Satzes etwas kompliziertere Grenzwerte angegangen werden können.

Beispiel 6.9. (a) Sei $p \in \mathbb{N}$ fest gewählt und $a_n = 1/n^p$ für $n \in \mathbb{N}$. Dann gilt für alle $n \in \mathbb{N}$ die Ungleichung $n \leq n^p$ und damit

$$0 \leq a_n = \frac{1}{n^p} \leq \frac{1}{n}.$$

Da sowohl die Folge, die konstant Null ist, als auch die Folge $(1/n)$ gegen Null konvergiert, ist damit nach Satz 6.8 (f) auch die Folge (a_n) konvergent und ebenfalls eine *Nullfolge*, d. h. es gilt $\lim_{n\to\infty} a_n = 0$.

(b) Wir untersuchen

$$a_n = \frac{n^2 + 2n + 3}{n^2 + 3}, \quad n \in \mathbb{N}.$$

Dazu kürzen wir den Bruch durch die höchste auftretende Potenz:

$$a_n = \frac{n^2 + 2n + 3}{n^2 + 3} = \frac{1 + \frac{2}{n} + \frac{3}{n^2}}{1 + \frac{3}{n^2}} \longrightarrow \frac{1 + 0 + 0}{1 + 0} = 1 \quad (n \to \infty).$$

Dabei stützen wir uns für die Berechnung des Grenzwertes von $(1/n^2)$ auf obiges Beispiel und zum Zusammenbau des Gesamtausdruckes auf (d)i), (d)ii) und (d)iv) aus Satz 6.8.

Dieses Vorgehen (Kürzen durch die höchste auftretende Potenz) ist bei allen Grenzwerten der Form „Polynom in n geteilt durch Polynom in n“ Erfolg versprechend.

Wir wollen uns als nächstes mit dem Monotonie-Verhalten von Folgen auseinandersetzen und danach mit Hilfe der neuen Begriffe ein weiteres Konvergenzkriterium herleiten.

Definition 6.10. *Eine reelle Folge (a_n) heißt*

(a) monoton wachsend, *wenn $a_{n+1} \geq a_n$ für alle $n \in \mathbb{N}$ gilt.*

(b) monoton fallend, *wenn $a_{n+1} \leq a_n$ für alle $n \in \mathbb{N}$ gilt.*

(c) streng monoton wachsend, *wenn $a_{n+1} > a_n$ für alle $n \in \mathbb{N}$ gilt.*

(d) streng monoton fallend, *wenn $a_{n+1} < a_n$ für alle $n \in \mathbb{N}$ gilt.*

(e) (streng) monoton, *wenn sie (streng) monoton wachsend oder (streng) monoton fallend ist.*

Damit können wir folgendes Konvergenzkriterium beweisen.

Satz 6.11 (Monotonie-Kriterium)**.** *Ist die reelle Folge (a_n) nach oben (bzw. unten) beschränkt und monoton wachsend (bzw. fallend), so ist (a_n) konvergent und es gilt*

$$\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n \quad (\textit{bzw.}\ \lim_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} a_n).$$

Beweis. Es sei (a_n) nach oben beschränkt und monoton wachsend sowie $a := \sup_{n\in\mathbb{N}} a_n$. Wählen wir nun ein $\varepsilon > 0$, so ist sicherlich $a - \varepsilon$ keine obere Schranke von $\{a_n : n \in \mathbb{N}\}$. Damit muss aber ein $n_0 \in \mathbb{N}$ existieren, so dass $a_{n_0} > a - \varepsilon$ ist. Nun haben wir unsere Folge umzingelt, denn es gilt wegen der Monotonie und der Beschränktheit von (a_n) für alle $n \geq n_0$:

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

und hiermit $|a_n - a| < \varepsilon$ für alle $n \geq n_0$. Da $\varepsilon > 0$ beliebig war, sind wir mit der ungeklammerten Aussage fertig. Die Aussage für monoton fallende Folgen beweist man analog. □

Wir betrachten ein Beispiel für die Anwendung dieses Satzes.

Beispiel 6.12. Wir betrachten eine *rekursiv definierte* Folge, die gegeben ist durch

$$a_1 := \sqrt[3]{6} \quad \text{und} \quad a_{n+1} := \sqrt[3]{6 + a_n}, \; n \geq 1.$$

Bei einer in dieser Weise gegebenen Folge ist keine explizite Rechenvorschrift angegeben, wie man das n-te Folgenglied bestimmen kann, sondern nur ein Startwert a_1 und dann eine Vorschrift, wie man aus einem Folgenglied das jeweils nächste berechnen kann. Auch damit ist die Folge eindeutig bestimmt. In unserem Beispiel erhält man für die ersten Folgenglieder

$$a_2 = \sqrt[3]{6 + \sqrt[3]{6}}, \quad a_3 = \sqrt[3]{6 + \sqrt[3]{6 + \sqrt[3]{6}}}, \quad a_4 = \sqrt[3]{6 + \sqrt[3]{6 + \sqrt[3]{6 + \sqrt[3]{6}}}}, \; \ldots.$$

So abstrus dieses Beispiel auch aussieht, in dieser Weise gegebene Folgen treten sehr häufig auf, so liefert z. B. jedes iterative Näherungsverfahren eine solche Folge.

Wie untersuchen wir aber ein solches Monstrum auf Konvergenz? Wir wenden unser Monotoniekriterium an, zeigen also, dass (a_n) nach oben beschränkt und monoton wachsend ist. Genauer gesagt beweisen wir

(a) $a_n < 2$ und $a_{n+1} > a_n$ für alle $n \in \mathbb{N}$,

(b) (a_n) konvergiert und $\lim_{n\to\infty} a_n = 2$.

Beweis. (a) Hier gehen wir induktiv vor:

Induktionsanfang: Es gilt $a_1 = \sqrt[3]{6} < \sqrt[3]{8} = 2$ und $a_2 = \sqrt[3]{6 + a_1} > \sqrt[3]{6} = a_1$, da $a_1 \geq 0$ ist. Also ist die Aussage für $n = 1$ richtig.

Induktionsvoraussetzung: Für ein $n \in \mathbb{N}$ gelte $a_n < 2$ und $a_{n+1} > a_n$.

Induktionsschritt: Es ist mit Hilfe der Induktionsvoraussetzung

$$a_{n+1} = \sqrt[3]{6 + a_n} < \sqrt[3]{6 + 2} = \sqrt[3]{8} = 2$$

und

$$a_{n+2} = \sqrt[3]{6 + a_{n+1}} > \sqrt[3]{6 + a_n} = a_{n+1}.$$

(b) Nach Satz 6.11 wissen wir nun, dass (a_n) konvergiert, und dass $\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n \leq 2$ ist, denn 2 ist eine obere Schranke der Folge. Außerdem wissen wir, dass $a_{n+1}^3 = 6 + a_n$ für alle $n \in \mathbb{N}$ gilt. Nach den Rechenregeln für Grenzwertbildung aus Satz 6.8 konvergieren bei dieser Gleichung die Folgen auf beiden Seiten des Gleichheitszeichens. Gehen wir also in dieser Gleichung zum Limes über, so erhalten wir für $a := \lim_{n\to\infty} a_n$ die Beziehung $a^3 = 6 + a$, bzw. $a^3 - a - 6 = 0$.

Eine Lösung dieser Gleichung ist $a = 2$. Dividieren wir diese ab, so erhalten wir $(a-2)(a^2+2a+3) = 0$ und $a^2 + 2a + 3 = (a+1)^2 + 2 = 0$ hat keine weiteren reellen Lösungen. Also muss $a = 2$ sein. $\square$

Noch ein Kommentar zum Verfahren. Obwohl es nicht immer zum Ziel führt, ist dieses doch ein starkes Hilfsmittel zur Behandlung rekursiver Folgen, das man immer wieder mit Gewinn verwenden kann.

7. Wichtige Beispiele

In diesem Abschnitt wollen wir nun einige weitere wichtige Folgen, die auch im weiteren Verlauf der Vorlesung immer wieder auftreten werden, auf Konvergenz untersuchen. Am Ende dieses Kapitels werden wir die Euler-Zahl e definieren können.

Satz 7.1. *Es sei (a_n) eine konvergente reelle Folge mit $a_n \geq 0$ für alle $n \in \mathbb{N}$. Bezeichnen wir den Limes von (a_n) mit a, so gilt $\sqrt[p]{a_n} \to \sqrt[p]{a}$ $(n \to \infty)$ für alle $p \in \mathbb{N}$.*

Beweis. Sei $p \in \mathbb{N}$ beliebig. Wir betrachten zunächst den Fall $a = 0$. Sei dazu $\varepsilon > 0$ gegeben. Dann gilt auch $\varepsilon^p > 0$, also gibt es ein $n_0 \in \mathbb{N}$, so dass $a_n < \varepsilon^p$ für alle $n \geq n_0$ gilt. Damit ist für diese n auch $\sqrt[p]{a_n} < \varepsilon$, also $|\sqrt[p]{a_n} - 0| < \varepsilon$ und wir sind fertig.
Sei nun $a > 0$. Dann gilt für alle $n \in \mathbb{N}$ nach Satz 5.2 (c) (setze dort $n := p - 1$)

$$\begin{aligned}|a_n - a| = \big|(\sqrt[p]{a_n})^p - (\sqrt[p]{a})^p\big| &= \left|\big(\sqrt[p]{a_n} - \sqrt[p]{a}\big)\sum_{k=0}^{p-1}(\sqrt[p]{a_n})^{p-1-k}(\sqrt[p]{a})^k\right| \\ &= \big|\sqrt[p]{a_n} - \sqrt[p]{a}\big|\sum_{k=0}^{p-1}(\sqrt[p]{a_n})^{p-1-k}(\sqrt[p]{a})^k.\end{aligned}$$

Beachten Sie, dass wir bei der Summe die Beträge weglassen können, da alle Summanden positiv sind, die Summe also in jedem Fall positiv ist. Da auch unser Gesamtausdruck dank des Betrages positiv ist, können wir diesen nun kleiner machen, indem wir in der Summe alle Summanden bis auf den letzten für $k = p-1$ weglassen. Das ist zugegebenermaßen eine grobe Abschätzung, aber, wie wir sehen werden, reicht das aus. Damit erhalten wir

$$|a_n - a| \geq \big|\big(\sqrt[p]{a_n} - \sqrt[p]{a}\big)\big|\,(\sqrt[p]{a})^{p-1}.$$

Setzen wir $c := (\sqrt[p]{a})^{p-1}$, so haben wir damit

$$\big|\sqrt[p]{a_n} - \sqrt[p]{a}\big| \leq \frac{1}{c}|a_n - a|.$$

Man beachte, dass wegen $a > 0$ auch $c > 0$ ist, und damit diese Umformung erlaubt ist.

Es gilt

$$0 \leq \left|\sqrt[p]{a_n} - \sqrt[p]{a}\right| \leq \frac{1}{c}|a_n - a|$$

und der Ausdruck auf der rechten Seite konvergiert für $n \to \infty$ gegen Null. Also gilt nach Satz 6.8 (f) $|\sqrt[p]{a_n} - \sqrt[p]{a}| \to 0$ $(n \to \infty)$ und damit schließlich $\sqrt[p]{a_n} \to \sqrt[p]{a}$ $(n \to \infty)$. □

Satz 7.2. *Sei $q \in \mathbb{R}$ beliebig. Dann ist die Folge $a_n = q^n$, $n \in \mathbb{N}$, genau dann konvergent, wenn $q \in (-1, 1]$ ist und es gilt*

$$\lim_{n\to\infty} a_n = \begin{cases} 1, & \text{falls } q = 1, \\ 0, & \text{falls } q \in (-1, 1). \end{cases}$$

Beweis. Wir betrachten zunächst die einfachen Fälle $q \in \{-1, 0, 1\}$. Ist $q = 0$, so ist $a_n = 0$ für alle $n \in \mathbb{N}$, die Folge konvergiert also gegen Null. Im Falle $q = 1$ findet man genauso wegen $a_n = 1$ für alle $n \in \mathbb{N}$ Konvergenz gegen Eins. Für $q = -1$ erhalten wir die schon aus Beispiel 6.4 (c) bekannte divergente Folge $((-1)^n)$.
Als nächstes wollen wir die Divergenz der Folge im Fall $|q| > 1$ zeigen. Wir setzen dazu $p = |q| - 1 > 0$ und schätzen mit Hilfe der Bernoullischen Ungleichung ab:

$$|q^n| = |q|^n = (1 + p)^n \geq 1 + np \geq np.$$

Damit ist die Folge (a_n) nicht beschränkt, denn gäbe es ein $M \geq 0$ mit $a_n \leq M$ für alle $n \in \mathbb{N}$, so wäre $np \leq |q^n| = |a_n| \leq M$ für alle $n \in \mathbb{N}$, also $\mathbb{N}$ beschränkt. Nach Satz 6.6 (b) kann damit (a_n) nicht konvergieren.
Schließlich zeigen wir Konvergenz für $0 < |q| < 1$. Dann ist $1/|q| > 1$, also haben wir wieder $1/|q| = 1 + p$ für ein $p > 0$. Die Bernoullische Ungleichung zeigt uns wieder

$$\frac{1}{|q^n|} = (1 + p)^n \geq 1 + np \geq np.$$

Also ist $0 < |q^n| \leq 1/np$ für alle $n \in \mathbb{N}$ und wie am Ende des Beweises von Satz 7.1 gilt damit $q^n \to 0$ $(n \to \infty)$. □

In diesem Zusammenhang beweisen wir die folgende Formel.

Satz 7.3 (Geometrische Summenformel)**.** *Es sei $q \neq 1$. Dann gilt für alle $n \in \mathbb{N}$*

$$\sum_{k=0}^{n} q^k = 1 + q + q^2 + q^3 + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q}.$$

Beweis. Es gilt

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

Da $q \neq 1$ ist, liefert Division durch $1 - q$ die Behauptung. $\square$

Satz 7.4. *Es ist* $\lim_{n\to\infty} \sqrt[n]{n} = 1$ *und für jedes* $c > 0$ *gilt* $\lim_{n\to\infty} \sqrt[n]{c} = 1$.

Beweis. Für alle $n \in \mathbb{N}$ gilt $1 = 1^n \leq n$ und damit $1 \leq \sqrt[n]{n}$ nach Lemma 5.3. Also ist $\sqrt[n]{n} = 1 + a_n$ mit einer Folge (a_n), für die $a_n \geq 0$ für alle $n \in \mathbb{N}$ gilt. Wir untersuchen nun die Folge (a_n) auf Konvergenz und erledigen damit sofort auch $(\sqrt[n]{n})$. Es ist für $n \geq 2$ mit Hilfe der Binomialformel

$$n = (\sqrt[n]{n})^n = (1 + a_n)^n = \sum_{k=0}^{n} \binom{n}{k} a_n^k \geq \binom{n}{2} a_n^2 = \frac{n!}{2!(n-2)!} a_n^2 = \frac{n(n-1)}{2} a_n^2.$$

Dabei haben wir alle Summanden, bis auf den mit $k = 2$, weggelassen. Wir formen um und erhalten $a_n^2 \leq 2/(n-1)$ für alle $n \geq 2$. Damit gilt für diese n

$$0 \leq a_n \leq \sqrt{\frac{2}{n-1}}.$$

Da die Folge $(2/(n-1))$, $n \geq 2$, gegen Null strebt, geht nach Satz 7.1 mit $p = 2$ die rechte Seite obiger Ungleichungskette ebenfalls gegen Null. Das Sandwich-Theorem liefert damit $a_n \to 0$ $(n \to \infty)$ und daher $\sqrt[n]{n} \to 1$ $(n \to \infty)$.
Bei der Betrachtung des zweiten Grenzwertes unterscheiden wir zwei Fälle. Ist $c \geq 1$, so wählen wir ein $m \in \mathbb{N}$ mit $m \geq c$ (dieses existiert, da $\mathbb{N}$ sonst beschränkt wäre). Dann gilt nach Lemma 5.3

$$1 \leq \sqrt[n]{c} \leq \sqrt[n]{m} \leq \sqrt[n]{n}$$

für alle $n \geq m$. Da wir den Grenzwert von $\sqrt[n]{n}$ oben schon zu 1 bestimmt haben, liefert das Sandwich-Theorem $\sqrt[n]{c} \to 1$ $(n \to \infty)$.
Ist nun $0 < c < 1$, so ist $1/c > 1$, wofür wir eben

$$\frac{1}{\sqrt[n]{c}} = \sqrt[n]{\frac{1}{c}} \longrightarrow 1 \ (n \to \infty)$$

gezeigt haben. Mit Hilfe von Satz 6.8 (d)iv) ist damit $\lim_{n\to\infty} \sqrt[n]{c} = 1$. $\square$

Wir wenden uns nun zwei besonders wichtigen Folgen zu, die harmlos aussehen, aber viel Zündstoff enthalten:

$$\begin{aligned} a_n &:= \left(1 + \frac{1}{n}\right)^n, & n \in \mathbb{N}, \\ b_n &:= \sum_{k=0}^{n} \frac{1}{k!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}, & n \in \mathbb{N}. \end{aligned}$$

Warnung 7.5. Die Folge (a_n) bietet eine gute Gelegenheit vor einem verbreiteten Fehler bei der Bestimmung von Grenzwerten zu warnen, der Unterteilung in „eiligere“ und „trägere“ n. Falsch ist nämlich folgende Überlegung: Die Folge $(1+1/n)$ geht offensichtlich gegen 1, also geht (a_n) gegen 1^n und das ist immer 1, was zu dem Ergebnis führe (a_n) würde gegen 1 streben. Das ist, wie wir nachher sehen werden, grob falsch. Der Grund ist folgender: Bei obiger Überlegung werden nicht alle n in der Formel gleich behandelt. Das n innerhalb der Klammer wird (quasi als Vorhut) zuerst nach ∞ geschickt, während das n im Exponenten noch warten muss, also zum „trägen“ n ernannt wird. Das geht nicht. Merke: Alle n sind gleich!
Mit der gleichen Berechtigung könnte man auch argumentieren, dass $1 + 1/n$ immer echt größer als 1 ist und da q^n für alle $q > 1$ divergiert, divergiert der Ausdruck in der Klammer, also auch die ganze Folge. Nun ist das andere n zum Warten gezwungen worden, und das Ergebnis ist genauso falsch wie das erste.
Diese Erörterung zeigt aber, was hier passiert. Das $1/n$ in der Klammer bringt den Ausdruck immer näher an 1 während es groß wird und macht es dem n im Exponenten damit immer schwerer, die Werte von a_n zu vergrößern. Die beiden beeinflussen den Wert also in verschiedene Richtungen und die Frage, die wir nun klären müssen, ist, wer dabei erfolgreicher ist: Schafft es das n in der Klammer, die Sache nach 1 zu drücken, oder ist das n im Exponent stärker und die Folge divergiert? Wir werden sehen, dass die beiden sich in magischer Weise im Gleichgewicht halten und die Wahrheit irgendwo dazwischen liegt.

Warnung 7.6. Obige Überlegungen sollten auch eine Warnung vor einer weiteren bösen Falle sein: Grenzwerte lassen sich im Allgemeinen nicht vertauschen! Die Ausführungen aus Warnung 7.5 bedeuten genau, dass

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

sowie dass

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

nicht existiert.
Es geht aber noch böser:

$$\begin{aligned}
\lim_{n\to\infty} \lim_{k\to\infty} \left(\frac{k}{k+1}\right)^n &= \lim_{n\to\infty} \left(\lim_{k\to\infty} \frac{k}{k+1}\right)^n = \lim_{n\to\infty} 1^n = 1, \\
\lim_{k\to\infty} \lim_{n\to\infty} \left(\frac{k}{k+1}\right)^n &= \lim_{k\to\infty} 0 = 0.
\end{aligned}$$

Mit diesem Problem werden wir uns immer wieder herumschlagen müssen und das bei weitem nicht nur in diesem Semster.

Nach diesem reichlichen Packen an Vorbemerkungen wenden wir uns jetzt der Behandlung der beiden oben angegebenen Folgen zu.

Satz 7.7. *Die Folgen* (a_n) *und* (b_n) *konvergieren und* $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

Beweis. Wir beginnen damit, die Konvergenz von (b_n) mit Hilfe des Monotoniekriteriums zu beweisen. Da $m! > 0$ für jedes $m \in \mathbb{N}$ gilt, sehen wir

$$b_{n+1} = \sum_{k=0}^{n+1} \frac{1}{k!} = \sum_{k=0}^{n} \frac{1}{k!} + \frac{1}{(n+1)!} = b_n + \frac{1}{(n+1)!} > b_n.$$

Die Folge (b_n) ist also streng monoton wachsend und es bleibt noch Beschränktheit zu zeigen. Dazu schreiben wir b_n aus

$$b_n = 1 + 1 + \frac{1}{2} + \frac{1}{2\cdot 3} + \frac{1}{2\cdot 3\cdot 4} + \cdots + \frac{1}{2\cdot 3\cdot \ldots \cdot n}$$

und beobachten, dass wir den Ausdruck größer machen, wenn wir in den Nennern alle Faktoren, die größer als 2 sind, durch 2 ersetzen:

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

Wir verwenden nun die geometrische Summenformel aus Satz 7.3, denn damit haben wir

$$b_n < 1 + \sum_{k=0}^{n-1} \left(\frac{1}{2}\right)^k = 1 + \frac{1 - (\frac{1}{2})^n}{1 - \frac{1}{2}} = 1 + 2 - 2\left(\frac{1}{2}\right)^n < 3.$$

Mit Hilfe des Monotoniekriteriums (s. Satz 6.11) wissen wir nun also, dass die Folge (b_n) konvergent ist, und dass für $b := \lim_{n\to\infty} b_n$ sogar $b \leq 3$ gilt.
Wir wenden uns der Folge (a_n) zu. Um für diese Folge Monotonie nachzuweisen, verwenden wir einen kleinen Trick: Wir betrachten den Quotienten zweier benachbarter Folgenglieder. Es gilt

$$\begin{aligned}
\frac{a_{n+1}}{a_n} &= \left(1 + \frac{1}{n+1}\right)^{n+1} \left(1 + \frac{1}{n}\right)^{-n} = \left(1 + \frac{1}{n+1}\right)\left(1 + \frac{1}{n+1}\right)^n \left(\frac{n+1}{n}\right)^{-n} \\
&= \left(1 + \frac{1}{n+1}\right)\left(\frac{n+2}{n+1} \cdot \frac{n}{n+1}\right)^n = \left(1 + \frac{1}{n+1}\right)\left(\frac{n^2 + 2n + 1 - 1}{(n+1)^2}\right)^n \\
&= \left(1 + \frac{1}{n+1}\right)\left(1 - \frac{1}{(n+1)^2}\right)^n.
\end{aligned}$$

Mit der Bernoullischen Ungleichung finden wir dafür die Abschätzung

$$\begin{aligned}\frac{a_{n+1}}{a_n} &\geq \left(1+\frac{1}{n+1}\right)\left(1-\frac{n}{(n+1)^2}\right) = 1+\frac{1}{n+1}-\frac{n}{(n+1)^2}-\frac{n}{(n+1)^3}\\ &= 1+\frac{n^2+2n+1-n^2-n-n}{(n+1)^3} = 1+\frac{1}{(n+1)^3} > 1.\end{aligned}$$

Nun haben wir $a_{n+1}/a_n > 1$ für alle $n \in \mathbb{N}$. Daraus folgt, da alle Folgenglieder positiv sind, sofort $a_n < a_{n+1}$ für alle $n \in \mathbb{N}$, also ist auch die Folge (a_n) streng monoton wachsend. Die Beschränktheit dieser Folge spielen wir nun auf die Beschränktheit von (b_n) zurück. Nach der Binomialformel haben wir

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

Es ist $\binom{n}{0} \cdot 1/n^0 = 1$ und $\binom{n}{1} \cdot 1/n^1 = n/n = 1$. Also gilt

$$\begin{aligned}&= 1+1+\sum_{k=2}^{n} \binom{n}{k} \frac{1}{n^k} = 1+1+\sum_{k=2}^{n} \frac{n!}{k!(n-k)!} \cdot \frac{1}{n^k}\\ &= 1+1+\sum_{k=2}^{n} \frac{1}{k!} \cdot \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}.\end{aligned}$$

Im hinteren Bruch können wir nun ein n kürzen. Dann bleiben sowohl im Zähler als auch im Nenner genau $k-1$ Faktoren übrig, die wir folgendermaßen zusammenfassen:

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

Schließlich beobachten wir, dass der Ausdruck in jeder Klammer nun kleiner als 1 ist und erhalten

$$< 1+1+\sum_{k=2}^{n} \frac{1}{k!} = b_n.$$

Also haben wir für jedes $n \in \mathbb{N}$ nun $a_n < b_n < 3$ und damit ist auch die Folge (a_n) beschränkt und zusammen mit der oben gezeigten Monotonie folgt damit die Konvergenz. Wir setzen $a := \lim_{n\to\infty} a_n$.

Es bleibt nun noch $a = b$ zu zeigen. Da wir schon $a_n \leq b_n$ für alle $n \in \mathbb{N}$ gezeigt haben, wissen wir bereits $a \leq b$. Es bleibt also nur die umgekehrte Ungleichung zu zeigen. Sei dazu ein $j \in \mathbb{N}$ mit $j \geq 2$ fest gewählt und $n \geq j$. Dann gilt wegen der obigen Rechnung

$$\begin{aligned} a_n &= 1 + 1 + \sum_{k=2}^{n} \frac{1}{k!}\Big(1 - \frac{1}{n}\Big)\Big(1 - \frac{2}{n}\Big)\cdots\Big(1 - \frac{k-1}{n}\Big) \\ &\geq 1 + 1 + \sum_{k=2}^{j} \frac{1}{k!}\Big(1 - \frac{1}{n}\Big)\Big(1 - \frac{2}{n}\Big)\cdots\Big(1 - \frac{k-1}{n}\Big). \end{aligned}$$

Wir gehen nun in dieser Ungleichung auf beiden Seiten zum Grenzwert $n \to \infty$ über. Man beachte, dass wir es hierbei nicht mit Schwierigkeiten wie unendlicher Summierung oder unendlichen Produkten zu tun bekommen. Wir haben „lediglich" eine endliche Summe von einem endlichen Produkt von Folgen, die alle konvergieren, Satz 6.8 ist also hier anwendbar.
Da auf jeden Fall $k - 1 < j \leq n$ ist, geht jeder einzelne Klammerausdruck dabei gegen 1, also geht jeder einzelne Summand gegen $1/k!$ und damit strebt der gesamte Ausdruck auf der rechten Seite genau gegen $1 + 1 + \sum_{k=2}^{j} 1/k! = b_j$ und wir erhalten somit aus unserer Ungleichung $a = \lim_{n\to\infty} a_n \geq b_j$ für jedes $j \in \mathbb{N}$ mit $j \geq 2$. Nun können wir schließlich den Grenzübergang $j \to \infty$ durchführen und erhalten $a \geq \lim_{j\to\infty} b_j = b$. □

Der Grenzwert dieser beiden Folgen ist so wichtig, dass wir ihm einen eigenen Namen verpassen.

Definition 7.8. *Die Zahl*

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

heißt Eulersche Zahl.

Wir haben im Beweis des obigen Satzes schon gesehen, dass $\mathrm{e} \leq 3$ gilt, im weiteren Verlauf der Vorlesung werden wir noch zeigen, dass $2 < \mathrm{e} < 3$ ist. Tatsächlich ist e eine irrationale Zahl mit

$$\mathrm{e} \approx 2,718281828459.$$

Mit Stand 2010 waren mehr als eine Billiarde Nachkommastellen berechnet.

8. Oberer und unterer Limes

Wie wir gesehen haben, sind beschränkte Folgen nicht automatisch konvergent. Trotzdem kann man einiges über ihr Verhalten im Unendlichen aussagen.
Sei also (a_n) eine beschränkte Folge in $\mathbb{R}$. Für jedes $n \in \mathbb{N}$ betrachten wir das Endstück

$$W_n := \{a_k : k \geq n\} = \{a_n, a_{n+1}, a_{n+2}, \dots\}$$

der Folge und setzen

$$\sigma_n := \sup W_n \qquad \varrho_n := \inf W_n.$$

Man beachte, dass diese beiden Zahlen jeweils existieren, denn $W_n \neq \emptyset$ und dank der Beschränktheit von (a_n) ist W_n beschränkt. Aber nicht nur W_n ist beschränkt, auch die beiden Folgen (σ_n) und (ϱ_n) teilen diese Eigenschaft.
Weiterhin gilt offensichtlich für jedes $n \in \mathbb{N}$ die Inklusion $W_{n+1} \subseteq W_n$. Damit haben wir nach Satz 2.16 (b) die Ungleichungen $\sigma_{n+1} \leq \sigma_n$ und $\varrho_{n+1} \geq \varrho_n$ für alle $n \in \mathbb{N}$. Das heißt (σ_n) ist monoton fallend und (ϱ_n) ist monoton wachsend. Wir können also erneut das Monotoniekriterium zücken und erhalten die Konvergenz beider Folgen.
Das ermöglicht die folgende Definition.

Definition 8.1. *Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$ und (σ_n) und (ϱ_n) seien die soeben konstruierten Folgen. Dann heißt*

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} \sigma_n \text{ oberer Limes } \textit{oder} \text{ Limes superior } \textit{von } (a_n),$$
$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} \varrho_n \text{ unterer Limes } \textit{oder} \text{ Limes inferior } \textit{von } (a_n).$$

Bemerkung 8.2. Nach dem Monotoniekriterium gilt $\lim_{n\to\infty} \sigma_n = \inf \sigma_n$ und $\lim_{n\to\infty} \varrho_n = \sup \varrho_n$. Also kann man den Limes superior bzw. inferior auch folgendermaßen beschreiben:

$$\limsup_{n\to\infty} a_n = \inf \Big\{ \sup\{a_n, a_{n+1}, a_{n+2}, \dots\} : n \in \mathbb{N}\Big\},$$
$$\liminf_{n\to\infty} a_n = \sup \Big\{ \inf\{a_n, a_{n+1}, a_{n+2}, \dots\} : n \in \mathbb{N}\Big\}.$$

Beispiel 8.3. (a) Für die uns schon bekannte beschränkte aber nicht konvergente Folge $a_n = (-1)^n$, $n \in \mathbb{N}$, gilt $W_n = \{-1, 1\}$ für alle $n \in \mathbb{N}$ und damit ist $\sigma_n = 1$ und $\varrho_n = -1$ unabhängig von $n \in \mathbb{N}$. Also ist $\limsup_{n\to\infty} a_n = 1$ und $\liminf_{n\to\infty} a_n = -1$.

(b) Für die Folge $a_n = 1/n$, $n \in \mathbb{N}$, gilt $W_n = \{1/n, 1/(n+1), 1/(n+2), \dots\}$, also ist $\sigma_n = 1/n$ und $\varrho_n = 0$ für alle $n \in \mathbb{N}$. Damit haben wir in diesem Fall $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = 0$.

Wir beweisen nun die den Grenzwertsätzen entsprechenden Aussagen für Limes superior und inferior.

Satz 8.4. *Es seien (a_n) und (b_n) beschränkte Folgen in $\mathbb{R}$. Dann gilt*

(a) $\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n$.

(b) Ist $a_n \le b_n$ für fast alle $n \in \mathbb{N}$, so gilt

$$\limsup_{n\to\infty} a_n \le \limsup_{n\to\infty} b_n \quad \textit{und} \quad \liminf_{n\to\infty} a_n \le \liminf_{n\to\infty} b_n.$$

(c) $\limsup_{n\to\infty}(a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$ *und*

$\liminf_{n\to\infty}(a_n + b_n) \ge \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$.

(d) Ist $\alpha \ge 0$, so gilt

$$\limsup_{n\to\infty}(\alpha a_n) = \alpha \limsup_{n\to\infty} a_n,$$
$$\liminf_{n\to\infty}(\alpha a_n) = \alpha \liminf_{n\to\infty} a_n.$$

(e) $\limsup_{n\to\infty}(-a_n) = -\liminf_{n\to\infty} a_n$ *und*

$\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n$.

Beweis. Wir beweisen jeweils nur die Aussagen für den Limes superior.
Zur Folge (a_n) definieren wir wie oben die Menge W_n und die Folgen (σ_n) und (ϱ_n).

(a) Nach Satz 2.16 (a) gilt immer $\varrho_n \le \sigma_n$. Also ist $\lim_{n\to\infty} \varrho_n \le \lim_{n\to\infty} \sigma_n$ und das ist nach der Definition von Limes superior und inferior genau die Behauptung.

(b) Wir definieren uns für die Folge (b_n) die Menge $V_n := \{b_k : k \ge n\}$ und die Folge $s_n := \sup V_n$ für jedes $n \in \mathbb{N}$. Nach Voraussetzung existiert ein $m \in \mathbb{N}$, so dass $a_n \le b_n$ für alle $n \ge m$ gilt. Sei nun so ein $n \ge m$ fest gewählt und $x \in W_n$. Dann gibt es ein $k \ge n$ mit $x = a_k$ und wir haben $x = a_k \le b_k \le s_n$. Damit haben wir aber $x \le s_n$ für alle $x \in W_n$ gezeigt, was $\sigma_n \le s_n$ impliziert. Da dies für alle $n \ge m$ gilt, folgt mit Satz 6.8 (e)

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sigma_n \le \lim_{n\to\infty} s_n = \limsup_{n\to\infty} b_n.$$

(c) Wir übernehmen die Bezeichnung s_n aus dem letzten Beweispunkt und definieren für jedes $n \in \mathbb{N}$ die Menge

$$U_n := \{a_n + b_n, a_{n+1} + b_{n+1}, a_{n+2} + b_{n+2}, \dots\},$$

sowie $r_n := \sup U_n$. Sei nun $n \in \mathbb{N}$ fest und $x \in U_n$. Dann gibt es ein $k \in \mathbb{N}$ mit $k \geq n$, so dass $x = a_k + b_k$ ist. Also folgt nach Definition von σ_n und s_n sofort $x \leq \sigma_n + s_n$. Da wir diese Überlegung für jedes $n \in \mathbb{N}$ anstellen können, gilt damit nach Satz 6.8 (d)i)

$$\limsup_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} r_n \leq \lim_{n\to\infty} \sigma_n + \lim_{n\to\infty} s_n = \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

(d) Übungsaufgabe.

(e) Wir setzen $\tilde{W}_n := \{-a_k : k \geq n\}$ und $\tilde{\sigma}_n := \sup \tilde{W}_n$. Dann gilt $\tilde{\sigma}_n = -\varrho_n$ (vgl. den Beweis von Satz 2.14). Damit haben wir $\limsup_{n\to\infty}(-a_n) = -\liminf_{n\to\infty} a_n$. □

Wir verdeutlichen uns durch ein Beispiel, dass in (c) im Allgemeinen nicht „=“ gilt.

Beispiel 8.5. Für jedes $n \in \mathbb{N}$ setzen wir $a_n = (-1)^n$ und $b_n = (-1)^{n+1}$. Dann gilt $a_n + b_n = 0$ für jedes $n \in \mathbb{N}$, also ist auch der Limes superior und der Limes inferior der Summenfolge 0. Aber es gilt $\limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n = 2$ und $\liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n = -2$.

Übungsaufgabe 8.6. Es seien (a_n) eine beschränkte und (b_n) eine konvergente Folge in $\mathbb{R}$. Zeigen Sie

$$\limsup_{n\to\infty}(a_n \cdot b_n) = \lim_{n\to\infty} b_n \limsup_{n\to\infty} a_n \quad \text{und} \quad \liminf_{n\to\infty}(a_n \cdot b_n) = \lim_{n\to\infty} b_n \liminf_{n\to\infty} a_n.$$

Gilt für zwei beschränkte Folgen (a_n) und (b_n) in $\mathbb{R}$ auch $\limsup_{n\to\infty}(a_n \cdot b_n) = \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n$?

9. Teilfolgen und Häufungswerte

Definition 9.1. *Sei* (a_n) *eine reelle Folge,* $\varphi : \mathbb{N} \to \mathbb{N}$ *eine streng monoton wachsende Abbildung und* $b_n := a_{\varphi(n)}$*,* $n \in \mathbb{N}$*. Dann heißt die Folge* (b_n) *eine* Teilfolge *von* (a_n)*.*

Dieser Begriff sieht zunächst etwas sperrig aus, ist aber genau das, wonach sich der Name anhört: Ein Teil der Folge. Man wählt also eine gewisse (in jedem Fall unendliche) Anzahl von Folgengliedern aus der Folge aus, das ist die Bildmenge von φ, und lässt die anderen weg, ohne die Reihenfolge der Folge zu ändern. Um diesen zweiten Punkt zu gewährleisten, muss φ streng monoton wachsend sein. Wir betrachten zwei Beispiele.

Beispiel 9.2. (a) Für $\varphi(k) = 2k$, $k \in \mathbb{N}$, haben wir $b_k = a_{2k}$, also $(b_n) = (a_2, a_4, a_6, \dots)$ und wir haben so die Teilfolge der Folgenglieder mit geradem Index ausgewählt.

(b) Ist $\varphi(k) := k^2$ für $k \in \mathbb{N}$, so ist $(b_n) = (a_1, a_4, a_9, \dots)$.

Definition 9.3. *Es sei* (a_n) *eine reelle Folge. Eine Zahl* $\alpha \in \mathbb{R}$ *heißt* Häufungswert *von* (a_n)*, wenn für jedes* $\varepsilon > 0$ *die Menge* $\{n \in \mathbb{N} : a_n \in U_\varepsilon(\alpha)\}$ *unendlich ist.*

Diese Definition sieht unserer Umformulierung der Konvergenz-Definition aus Kapitel 6 sehr ähnlich. Wir stellen die beiden noch einmal gegenüber:

- Es ist $\lim_{n\to\infty} a_n = a$, falls
$$\forall\, \varepsilon > 0 : a_n \in U_\varepsilon(a) \text{ für fast alle } n \in \mathbb{N}.$$
- Die Folge (a_n) hat a als Häufungswert, falls
$$\forall\, \varepsilon > 0 : a_n \in U_\varepsilon(a) \text{ für unendlich viele } n \in \mathbb{N}.$$

Nun sieht man deutlich, dass die Anforderung Häufungswert zu sein schwächer ist als die der Konvergenz. Tatsächlich feststellen können wir das mit dem schon mehrfach bemühten Beispiel $a_n = (-1)^n$, $n \in \mathbb{N}$.

Beispiel 9.4. Es sei also $(a_n) = ((-1)^n)$. Diese Folge hat genau zwei Häufungswerte, nämlich 1 und -1.
Das sieht man so: Für jedes gerade $n \in \mathbb{N}$ gilt $a_n = 1$, also liegen für jedes $\varepsilon > 0$ alle Folgenglieder mit geradem Index in $U_\varepsilon(1)$ und da das unendlich viele sind, ist 1 Häufungswert. Genauso ist -1 Häufungswert. Das sieht man durch Betrachtung der Folgenglieder mit ungeradem Index.
Es bleibt also noch zu zeigen, dass alle anderen reellen Zahlen keine Häufungswerte sind. Dazu erinnern wir uns zunächst an die kleine Logikeinführung und verneinen die Aussage „$\beta \in \mathbb{R}$ ist Häufungspwert von (a_n)“. Das ergibt: Es gibt ein $\varepsilon_0 > 0$, so dass $a_n \in U_{\varepsilon_0}(\beta)$ für höchstens endlich viele $n \in \mathbb{N}$ gilt.
Sei also $\beta \in \mathbb{R}$ mit $\beta \neq 1$ und $\beta \neq -1$. Wähle nun $\varepsilon_0 > 0$ so klein, dass $1, -1 \notin U_\varepsilon(\beta)$ gilt. Das geht, da β einen echt positiven Abstand sowohl von 1 als auch von -1 haben muss. Dann gilt $a_n \in U_{\varepsilon_0}(\beta)$ für gar kein $n \in \mathbb{N}$. Also kann β kein Häufungswert sein.

Beispiel 9.5. (a) Die Folge $(a_n) = (n)$ hat gar keinen Häufungswert, denn für jedes $\beta \in \mathbb{R}$ gilt $a_n > \beta + 1$ für fast alle $n \in \mathbb{N}$.

(b) Wie wir in Beispiel 4.5 (c) gesehen haben, ist $\mathbb{Q}$ abzählbar. Es gibt also eine Folge (a_n), so dass $\mathbb{Q} = \{a_n : n \in \mathbb{N}\}$ gilt. Sei nun $\alpha \in \mathbb{R}$ beliebig und $\varepsilon > 0$. Dann liegen im Intervall $(\alpha - \varepsilon, \alpha + \varepsilon)$ unendlich viele rationale Zahlen, d. h. α ist Häufungswert der Folge (a_n). Wir haben also eine Folge gefunden, für die *jede* reelle Zahl ein Häufungswert ist.

Mit den Zusammenhängen zwischen Teilfolgen, Häufungswerten und Konvergenz befasst sich der folgende Satz.

Satz 9.6. *Es sei (a_n) eine reelle Folge. Dann gilt*

(a) Eine Zahl $\alpha \in \mathbb{R}$ ist Häufungswert von (a_n) genau dann, wenn eine Teilfolge (b_k) von (a_n) existiert, die gegen α konvergiert.

(b) Ist (a_n) konvergent und (b_k) eine Teilfolge von (a_n), so ist auch (b_k) konvergent und es gilt $\lim_{k\to\infty} b_k = \lim_{n\to\infty} a_n$.

(c) Ist (a_n) konvergent, so hat (a_n) genau einen Häufungswert, nämlich den Grenzwert $\lim_{n\to\infty} a_n$.

Beweis. (a) Wir zeigen zunächst die Richtung von links nach rechts. Sei also α ein Häufungswert von (a_n). Dann existiert insbesondere für $\varepsilon = 1$ ein $n_1 \in \mathbb{N}$ mit $|a_{n_1} - \alpha| < 1$. Da es auch für $\varepsilon = 1/2$ unendlich viele Folgenglieder von (a_n) in der $1/2$-Umgebung von α gibt, muss es auch ein $n_2 \in \mathbb{N}$ mit $n_2 > n_1$ geben, so dass $|a_{n_2} - \alpha| < 1/2$ gilt. Genauso finden wir ein $n_3 \in \mathbb{N}$ mit $n_3 > n_2$, so dass $|a_{n_3} - \alpha| < 1/3$ gilt.

Verfahren wir immer weiter so, erhalten wir schließlich eine Folge von Indizes $n_1, n_2, n_3, \dots$ mit $n_1 < n_2 < n_3 < \dots$, so dass

$$|a_{n_k} - \alpha| \leq \frac{1}{k} \quad \text{für alle } k \in \mathbb{N} \tag{9.1}$$

gilt. Setzen wir nun $b_k := a_{n_k}$, $k \in \mathbb{N}$, so ist (b_k) eine Teilfolge von (a_n), von der noch zu zeigen ist, dass sie gegen α konvergiert. Sei dazu $\varepsilon > 0$. Dann gibt es ein $k_0 \in \mathbb{N}$, so dass $1/k_0 < \varepsilon$ ist und mit (9.1) gilt für alle $k \geq k_0$

$$|b_k - \alpha| = |a_{n_k} - \alpha| \leq \frac{1}{k} \leq \frac{1}{k_0} < \varepsilon.$$

Wir wenden uns nun der Richtung von rechts nach links zu. Sei also (b_k) eine Teilfolge von (a_n) mit $b_k \to \alpha$ $(k \to \infty)$. Zu gegebenem $\varepsilon > 0$ existiert dann ein $k_0 \in \mathbb{N}$, so dass $|b_k - \alpha| < \varepsilon$ für alle $k \geq k_0$ gilt. Damit ist aber $a_{n_k} \in U_\varepsilon(\alpha)$ für alle $k \geq k_0$, also für unendlich viele Indizes. Damit ist α Häufungswert von (a_n).

(b) Wir setzen $a := \lim_{n\to\infty} a_n$ und $(a_{n_k}) = (b_k)$. Sei nun $\varepsilon > 0$ gegeben. Dann gibt es ob der Konvergenz von (a_n) ein $n_0 \in \mathbb{N}$, so dass $|a_n - a| < \varepsilon$ für alle $n \geq n_0$ gilt. Nun wählen wir $k_0 \in \mathbb{N}$ so groß, dass $n_{k_0} \geq n_0$ gilt. Dann ist für alle $k \geq k_0$ nämlich $n_k \geq n_{k_0} \geq n_0$, weshalb $|b_k - a| = |a_{n_k} - a| < \varepsilon$ für alle $k \geq k_0$ folgt.

(c) Offensichtlich ist $a := \lim_{n\to\infty} a_n$ ein Häufungswert von (a_n), denn in jeder ε-Umgebung liegen ja fast alle (also insbesondere unendlich viele) Folgenglieder. Wir zeigen, dass es keine weiteren Häufungswerte geben kann. Sei dazu $\alpha \in \mathbb{R}$ ein Häufungswert von (a_n). Dann gibt es wegen (a) eine Teilfolge (b_k) von (a_n), die gegen α konvergiert. Nach (b) gilt dann aber $\alpha = \lim_{k\to\infty} b_k = a$, also ist a der einzige mögliche Häufungswert. □

Beispiel 9.7. Wir bestimmen alle Häufungswerte von

$$a_n = (-1)^n \left(1 + \frac{1}{n}\right)^n, \quad n \in \mathbb{N}.$$

Zunächst ist $a_{2n} = (-1)^{2n}(1 + 1/(2n))^{2n} = (1 + 1/(2n))^{2n}$, $n \in \mathbb{N}$, sowohl eine Teilfolge von (a_n) als auch von $((1 + 1/n)^n)$, von der wir ja schon wissen, dass sie gegen die Eulersche Zahl e strebt. Nach Satz 9.6 (b) konvergiert (a_{2n}) gegen e und mit (a) aus dem selben Satz ist e ein Häufungswert von (a_n). Für die Teilfolge $a_{2n-1} = -(1 + 1/(2n-1))^{2n-1}$ erhält man genauso, dass $-$e ein Häufungswert der Folge ist.
Sei nun $\alpha \in \mathbb{R}$ mit $\alpha \neq \mathrm{e}$ und $\alpha \neq -\mathrm{e}$. Weiterhin setzen wir

$$\varepsilon := \frac{1}{2} \min\{|\mathrm{e} - \alpha|, |-\mathrm{e} - \alpha|\}, \qquad U := U_\varepsilon(\mathrm{e}) \cup U_\varepsilon(-\mathrm{e}).$$

Dann ist $\varepsilon > 0$ und es gilt $U \cap U_\varepsilon(\alpha) = \emptyset$. Aus der Konvergenz der beiden obigen Teilfolgen ersehen wir, dass $a_n \in U_\varepsilon(\mathrm{e})$ für fast alle geraden $n \in \mathbb{N}$ und $a_n \in U_\varepsilon(-\mathrm{e})$ für fast alle ungeraden $n \in \mathbb{N}$ gilt. Also ist $a_n \in U$ für fast alle $n \in \mathbb{N}$. Andersherum betrachtet bedeutet dies, dass höchstens endlich viele Folgenglieder in $U_\varepsilon(\alpha)$ liegen können. Also ist α kein Häufungswert von (a_n).

Lemma 9.8. *Sei (a_n) eine beliebige Folge in $\mathbb{R}$. Dann enthält (a_n) eine monotone Teilfolge.*

Beweis. Um diese Folge elegant konstruieren zu können, nennen wir ein $j \in \mathbb{N}$ „niedrig“, wenn $a_n \geq a_j$ für alle $n \geq j$ gilt. Wir unterscheiden zwei Fälle.
Gibt es nur endlich viele niedrige Indizes, so existiert ein $m \in \mathbb{N}$, so dass alle $n \geq m$ nicht niedrig sind. Wir setzen $n_1 := m$. Da n_1 damit nicht niedrig ist, gibt es ein $n_2 \in \mathbb{N}$ mit $n_2 > n_1$, so dass $a_{n_2} < a_{n_1}$ ist. Nun ist aber auch $n_2 > n_1 > m$, also nicht niedrig. Wir erhalten in gleicher Weise ein $n_3 > n_2$ mit $a_{n_3} < a_{n_2}$. Treiben wir dieses Spielchen immer weiter, so finden wir eine streng monoton fallende Teilfolge (a_{n_k}) von (a_n).
Gibt es dagegen unendlich viele niedrige Indizes, so nummerieren wir diese mit $n_1, n_2, n_3, \ldots$ aufsteigend durch, so dass $n_1 < n_2 < n_3 < \ldots$ gilt. Da jedes dieser n_j niedrig ist und $n_{j+1} > n_j$ für alle $j \in \mathbb{N}$ gilt, haben wir auch immer $a_{n_{j+1}} \geq a_{n_j}$. Also ist in diesem Fall (a_{n_j}) monoton wachsend. □

Wir wenden uns nun dem Satz von Bolzano-Weierstraß zu, dem zentralen Satz dieses Kapitels. Bekommt ein Satz einen Namen verpasst, so ist das ein starkes Indiz dafür, dass es sich dabei um ein wichtiges Resultat handelt.

Satz 9.9 (Satz von Bolzano-Weierstraß)**.** *Jede beschränkte Folge in $\mathbb{R}$ besitzt eine konvergente Teilfolge.*

Mit Hilfe unserer Überlegungen in Satz 9.6 können diesen Satz auch folgendermaßen umformulieren:

Jede beschränkte Folge in $\mathbb{R}$ hat mindestens einen Häufungswert.

Der Beweis dieses Satzes ist mit unserem Lemma von oben schon fast erledigt. Er geht so:

Beweis. Sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Dann gibt es nach Lemma 9.8 eine monotone Teilfolge (b_k) von (a_n). Da (a_n) aber beschränkt ist, ist auch (b_k) beschränkt. Nach dem Monotonie-Kriterium ist (b_k) also konvergent. □

Mit Hilfe des Satzes von Bolzano-Weierstraß können wir nun die eingängige Beschreibung des Limes superior als größtem Häufungswert und des Limes inferior als kleinstem Häufungswert der Folge beweisen. Dazu setzen wir für eine beschränkte Folge (a_n) in $\mathbb{R}$

$$H(a_n) := \{\alpha \in \mathbb{R} : \alpha \text{ ist Häufungswert von } (a_n)\}.$$

Satz 9.10. *Es sei* (a_n) *eine beschränkte Folge in* $\mathbb{R}$. *Dann gilt*

$$\limsup_{n\to\infty} a_n = \max H(a_n), \qquad \liminf_{n\to\infty} a_n = \min H(a_n).$$

Beweis. Zum Beweis dieses Satzes gehen wir in zwei Schritten vor. Erstens zeigen wir, dass $\limsup_{n\to\infty} a_n$ (und analog auch $\liminf_{n\to\infty} a_n$) in jedem Fall ein Häufungswert der Folge (a_n) ist. Danach weisen wir für jedes $\alpha \in H(a_n)$ die Ungleichung $\liminf_{n\to\infty} a_n \leq \alpha \leq \limsup_{n\to\infty} a_n$ nach.
Wir kümmern uns zunächst um den ersten Schritt. Für jedes $k \in \mathbb{N}$ schreiben wir wieder $W_k := \{a_n : n \geq k\}$ und $\sigma_k = \sup W_k$. Nach der Definition des Limes superior gilt $\alpha := \limsup_{k\to\infty} a_k = \lim_{k\to\infty} \sigma_k$. Sei $k \in \mathbb{N}$ fest gewählt. Dann ist $\sigma_k - 1/k$ keine obere Schranke von W_k, denn σ_k ist das Supremum. Folglich existiert ein Index $n_k \geq k$, so dass $\sigma_k - 1/k < a_{n_k}$ ist. Setzen wir $c_k := a_{n_k}$, so gilt für diese Folge

$$\sigma_k - \frac{1}{k} < c_k \leq \sigma_k \quad \text{für alle } k \in \mathbb{N}.$$

Wir beobachten, dass die Ausdrücke auf der linken und rechten Seite beide gegen α konvergieren. Gehen wir nun in dieser Ungleichung zum Grenzwert über, so erhalten wir mit dem Sandwich-Theorem (Satz 6.8 (f)) $\lim_{n\to\infty} c_n = \alpha$.
An dieser Stelle ist eine *Warnung* angebracht, denn nun liegt der Schluss nahe, dass wir mit (c_n) eine Teilfolge von (a_n) gefunden haben, die gegen α konvergiert. Aber Achtung: (c_n) muss nicht unbedingt eine Teilfolge von (a_n) sein! Wir haben bisher nicht garantiert, dass $n_1 < n_2 < n_3 < \dots$ gilt, denn wir haben für jedes k immer nur ein $n_k \geq k$ ausgewählt. Dieses kann sich aber häufiger wiederholt haben oder wir könnten sogar in der Zählung zurück gesprungen sein.
Die Idee mit der Teilfolge ist aber doch recht brauchbar, denn wenn wir eine Teilfolge (b_n) von (a_n) mit $\lim_{n\to\infty} b_n = \alpha$ haben, so ist α Häufungswert von (a_n) und dies peilen wir ja an. Wie konstruieren wir also aus (c_n) unser (b_n)?
Da $n_k \geq k$ für alle $k \in \mathbb{N}$ gilt, ist zumindest die Menge der Indizes $\{n_k : k \in \mathbb{N}\}$ unendlich groß. Wir können also aus der Folge (n_k) dieser Indizes eine Teilfolge (m_k) mit $m_1 < m_2 < m_3 < \dots$ auswählen und setzen dann $b_k := a_{m_k}$. Dann ist (b_k) eine Teilfolge von (c_k) und damit auch von (a_n). Als Teilfolge der konvergenten Folge (c_k) konvergiert (b_k) in jedem Fall gegen α und dieses ist damit Häufungswert von (a_n).
Wir wenden uns unserem zweiten Beweisschritt zu. Sei dazu $\alpha \in H(a_n)$. Nach Satz 9.6 (a) gibt es eine Teilfolge (a_{n_k}) von (a_n) mit $a_{n_k} \to \alpha$ $(k \to \infty)$. Außerdem gilt $\varrho_n \leq a_n \leq \sigma_n$ für alle $n \in \mathbb{N}$, wobei ϱ_n und σ_n wie in Kapitel 8 definiert sind. Insbesondere gilt also

$$\varrho_{n_k} \leq a_{n_k} \leq \sigma_{n_k} \quad \text{für alle } k \in \mathbb{N}.$$

In dieser Ungleichung existieren nun alle beteiligten Grenzwerte, so dass wir $k \to \infty$ streben lassen können. Das liefert $\liminf_{n\to\infty} a_n \leq \alpha \leq \limsup_{n\to\infty} a_n$ und damit die Behauptung. $\square$

Beispiel 9.11. In Beispiel 9.7 haben wir für die Folge $a_n = (-1)^n(1+1/n)^n$, $n \in \mathbb{N}$, schon alle Häufungswerte bestimmt und bekamen $H(a_n) = \{-\mathrm{e}, \mathrm{e}\}$. Hiermit und mit obigem Satz wissen wir nun, dass

$$\limsup_{n\to\infty}(-1)^n\left(1+\frac{1}{n}\right)^n = \mathrm{e} \quad \text{und} \quad \liminf_{n\to\infty}(-1)^n\left(1+\frac{1}{n}\right)^n = -\mathrm{e}$$

gilt.

Der folgende Satz bringt unsere Einsichten in Zusammenhang mit dem Begriff der Konvergenz.

Satz 9.12. *Es sei (a_n) eine beschränkte Folge in $\mathbb{R}$. Dann sind die folgenden Aussagen äquivalent:*

(a) (a_n) *ist konvergent.*

(b) (a_n) *hat genau einen Häufungswert.*

(c) $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$.

Ist eine dieser Aussagen erfüllt, so gilt

$$\lim_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n.$$

Beweis. Die Äquivalenz (b) $\iff$ (c) folgt direkt aus Satz 9.10 und (a)$\Rightarrow$(b) ist genau der Inhalt von Satz 9.6 (c). Es bleibt uns also nur noch die Implikation (b)$\Rightarrow$(a) zu zeigen.
Dazu vergegenwärtigen wir uns erneut, dass für die beiden Folgen (σ_n) und (ϱ_n) aus der Definition von Limes superior und Limes inferior $\varrho_n \le a_n \le \sigma_n$ für alle $n \in \mathbb{N}$ gilt. Gehen wir nun in dieser Ungleichung zum Grenzwert über, so konvergiert der Ausdruck auf der linken Seite gegen $\liminf_{n\to\infty} a_n$, der auf der rechten Seite gegen $\limsup_{n\to\infty} a_n$. Da diese beiden aber nach Voraussetzung identisch sind, muss nach dem Sandwich-Theorem die Folge (a_n) konvergieren und es gilt $\lim_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$, womit auch gleichzeitig der Nachsatz des Theorems bewiesen wäre. □

Übungsaufgabe 9.13. (Prinzip der Intervallschachtelung) Seien I_n, $n \in \mathbb{N}$, abgeschlossene und nicht-leere Intervalle in $\mathbb{R}$ mit $I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$, so dass die Länge von I_n für n gegen Unendlich gegen Null geht. Dann ist $\bigcap_{n\in\mathbb{N}} I_n = \{x_*\}$ für ein $x_* \in \mathbb{R}$.

Übungsaufgabe 9.14. Erklären Sie jemandem aus ihrem Semester, warum die Umkehrung der Aussage in Teil (c) von Satz 9.6 im Allgemeinen falsch ist.

10. Cauchy-Folgen

Die in dieser Vorlesung bisher behandelten Konvergenz-Kriterien haben meist den Nachteil, dass man zu ihrer Anwendung schon eine Idee haben muss, was denn der Grenzwert der Folge sein könnte. Ein solches Orakel ist im Allgemeinen nicht zu erwarten. Oft reicht es auch schon aus, zu wissen, dass eine Folge konvergiert, ohne den genauen Wert des Limes zu kennen. Dementsprechend ist es wünschenswert, ein Kriterium zu kennen, das etwas über die Konvergenz der Folge auch ohne ein solches Vorwissen aussagt. Das ist das Ziel dieses Abschnitts. Das dabei eingeführte Konzept der Cauchy-Folgen wird ein weit über diese Motivation herausgehendes Eigenleben entwickeln und ist in verschiedenen Bereichen der Analysis von elementarer Wichtigkeit.

Anschaulich gesprochen, wollen wir zeigen, dass eine reelle Folge konvergiert, wenn die Abstände zwischen verschiedenen Folgengliedern im Unendlichen beliebig klein werden. Das gibt uns ein Kriterium an die Hand, das nur die relative Lage der Folgenglieder zueinander und eben nicht die absolute Lage des Grenzwerts berücksichtigt.

Um ein solches Kriterium herzuleiten, beginnen wir andersherum und überlegen uns, wie sich der Abstand zweier weit draußen liegender Folgenglieder bei einer konvergenten Folge verhält.

Sei also (a_n) eine konvergente Folge in $\mathbb{R}$ mit Grenzwert a. Dann gibt es für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass $|a_n - a| \leq \varepsilon/2$ für alle $n \geq n_0$ gilt. Wählen wir nun einen weiteren beliebigen Index $m \geq n_0$, so ist mit der Dreiecksungleichung

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

Diese wichtige Eigenschaft präzisieren wir nun in einer Definition.

Definition 10.1. *Eine reelle Folge (a_n) heißt* Cauchy-Folge, *wenn für jedes $\varepsilon > 0$ ein Index $n_0 \in \mathbb{N}$ existiert, so dass gilt*

$$|a_n - a_m| < \varepsilon \quad \text{für alle } n, m \geq n_0.$$

Satz 10.2 (Cauchy-Kriterium). *Eine Folge in $\mathbb{R}$ ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.*

Beweis. Den Beweis, dass jede konvergente Folge eine Cauchy-Folge ist, haben wir schon in obiger Vorüberlegung erbracht. Wir wenden unser Augenmerk also der anderen Beweisrichtung zu.

Sei dazu (a_n) eine Cauchy-Folge. Wir überlegen uns zunächst, dass diese Folge dann beschränkt sein muss. Nach Definition der Cauchy-Folge gibt es ein $n_0 \in \mathbb{N}$, so dass $|a_n - a_m| < 1$ für alle $n, m \geq n_0$ ist. Insbesondere haben wir für den Spezialfall $m = n_0$ die Ungleichung $|a_n - a_{n_0}| < 1$ für alle $n \geq n_0$. Damit gilt für alle $n \geq n_0$

$$|a_n| = |a_n - a_{n_0} + a_{n_0}| \leq |a_n - a_{n_0}| + |a_{n_0}| < 1 + |a_{n_0}|.$$

Damit haben wir alle bis auf endlich viele Folgenglieder beschränkt, also ist

$$|a_n| \leq \max\{|a_1|, |a_2|, |a_3|, \ldots, |a_{n_0-1}|, 1 + |a_{n_0}|\} \quad \text{für alle } n \in \mathbb{N}.$$

Für die beschränkte Folge (a_n) existieren nun also der Limes inferior und der Limes superior. Wir nehmen an, unsere Folge (a_n) divergiere. Dann gilt die strikte Ungleichung

$$\alpha := \liminf_{n\to\infty} a_n < \liminf_{n\to\infty} a_n =: \beta,$$

denn wären die beiden gleich, so wäre nach Satz 9.12 die Folge (a_n) konvergent. Die Idee für das weitere Vorgehen ist nun, dass sich sowohl in jeder Umgebung von α als auch in jeder Umgebung von β jeweils unendlich viele Folgenglieder tummeln müssen, die Folge also (für kleine Umgebungen) sehr oft zwischen den Umgebungen von α und β hin- und herspringen muss. Anschaulich kann dann aber, da zwischen α und β ein echter Abstand ist, der Abstand der Folgenglieder zueinander nicht beliebig klein werden. Diese Überlegung wollen wir im Folgenden formalisieren.
Wir wählen zwei weitere reelle Zahlen λ, μ, so dass $\alpha < \lambda < \mu < \beta$ gilt und setzen $\varepsilon := \mu - \lambda$. Dann ist $\varepsilon > 0$, es existiert also ein $n_0 \in \mathbb{N}$, so dass

$$|a_n - a_m| < \varepsilon = \mu - \lambda \quad \text{für alle } n, m \geq n_0 \tag{10.1}$$

gilt. Andererseits ist α Häufungswert von (a_n), also gilt $a_n < \lambda$ für unendlich viele $n \in \mathbb{N}$. Insbesondere gibt es einen Index $j \geq n_0$, so dass $a_j < \lambda$ ist. Ebenso ist, da β Häufungswert ist, $a_m > \mu$ für unendlich viele m, es gibt also einen Index $k \geq n_0$ mit $a_k > \mu$. Für diese beiden Indizes erhalten wir aber wegen $a_j < \lambda < \mu < a_k$ die Beziehung

$$|a_k - a_j| = a_k - a_j > \mu - \lambda = \varepsilon,$$

was wegen $j, k \geq n_0$ im Widerspruch zu (10.1) steht. □

Bemerkung 10.3. (a) Man kann das Cauchy-Kriterium auch folgendermaßen umformulieren: Eine reelle Folge (a_n) ist eine Cauchy-Folge genau dann, wenn

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \forall n \geq n_0 \ \forall k \in \mathbb{N}_0 : |a_n - a_{n+k}| < \varepsilon.$$

(b) Die Richtigkeit der Aussage, dass jede Cauchy-Folge konvergiert, ist eng mit dem Vollständigkeitsaxiom verknüpft. So ist diese Aussage z. B. in $\mathbb{Q}$ falsch! Machen Sie sich das anhand einer Folge in $\mathbb{R}$, die gegen eine irrationale Zahl strebt, also z. B. $(1 + 1/n)^n$, $n \in \mathbb{N}$, klar.

11. Unendliche Reihen

Wir wollen uns nun mit der spannenden Theorie der unendlichen Reihen, also des unendlichen Summationsprozesses, beschäftigen. Die Besonderheiten, die hier im Zusammenhang mit der Unendlichkeit auftreten, wollen wir mit dem in den vorherigen Kapiteln entwickelten Werkzeug der Konvergenz in den Griff bekommen. Gerade für dieses Kapitel sei noch einmal die Warnung ausgesprochen: Vertrauen Sie beim Kontakt mit dem Unendlichen nur sehr eingeschränkt Ihrer Intuition!
Zunächst erweitern wir unseren Folgenbegriff ein wenig, indem wir auch Folgen zulassen, deren Indizes nicht ausgerechnet bei 1 anfangen, wie das bisher der Fall war, sondern z. B. bei 7, 0 oder -20.

Definition 11.1. *Es sei $p \in \mathbb{Z}$ und $X \neq \emptyset$ eine Menge. Eine Abbildung*

$$a : \{p, p+1, p+2, \dots\} \longrightarrow X$$

nennen wir in Erweiterung von Definition 4.1 ebenfalls Folge *in X und bezeichnen diese mit $(a_n)_{n=p}^{\infty}$. Ist $p = 1$, so schreiben wir weiterhin (a_n) für $(a_n)_{n=1}^{\infty}$.*

Für eine solche Folge $(a_n)_{n=p}^{\infty}$ können wir durch $b_n := a_{n+p-1}$, $n \in \mathbb{N}$, eine Folge (b_n) konstruieren, die unserer alten Definition entspricht.
Das Zurückspielen der unendlichen Summation auf den Begriff der Konvergenz einer Folge geschieht in der folgenden Definition.

Definition 11.2. *(a) Es sei (a_n) eine Folge in $\mathbb{R}$. Wir setzen $s_n := a_1 + a_2 + a_3 + \dots + a_n$, $n \in \mathbb{N}$. Dann heißt die Folge (s_n) eine* (unendliche) Reihe *und wir schreiben dafür das Symbol*

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \dots .$$

*(b) Für jedes $n \in \mathbb{N}$ heißt $s_n = \sum_{k=1}^{n} a_k$ die n-*te Teilsumme *oder* Partialsumme *der Reihe.*

(c) Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt konvergent *(bzw.* divergent*), wenn die Folge (s_n) konvergiert (bzw. divergiert).*

(d) Ist die Reihe $\sum_{n=1}^{\infty} a_n$ konvergent, so heißt $\lim_{n\to\infty} s_n$ der Reihenwert *oder die* Reihensumme *und wir schreiben*

$$\sum_{n=1}^{\infty} a_n := \lim_{n\to\infty} s_n.$$

Bemerkung 11.3. (a) Beachten Sie, dass das Symbol $\sum_{n=1}^{\infty} a_n$ nun zwei Bedeutungen hat. Es steht zum einen für die Folge der Partialsummen und ist damit ein abstraktes Zeichen, dass unabhängig davon sinnvoll ist, ob die Partialsummen konvergieren, d. h. die unendliche Summation definiert ist. Gleichzeitig symbolisiert es den Grenzwert der Partialsummen, also den Reihenwert, und bevor man es in dieser Bedeutung verwendet, muss natürlich zunächst geklärt werden, ob die Folge (s_n) überhaupt konvergiert.

Diese Doppelbelegung hat sich fest eingebürgert, daher machen wir uns diesen Sprachgebrauch im Folgenden zu eigen.

(b) Der Name des Summationsindexes spielt für die Bedeutung des Symbols keine Rolle, d. h. es ist

$$\sum_{n=1}^{\infty} a_n = \sum_{k=1}^{\infty} a_k = \sum_{\nu=1}^{\infty} a_\nu.$$

(c) Ist $p \in \mathbb{Z}$ und $(a_n)_{n=p}^{\infty}$ eine Folge, so definiert man entsprechend die Folge der Partialsummen als $s_n := a_p + a_{p+1} + a_{p+2} + \cdots + a_n$ für alle $n \geq p$ und schreibt für die Folge (s_n) wieder $\sum_{n=p}^{\infty} a_n$, genau so wie für den Reihenwert, falls dieser existiert.

Wir werden auch im Folgenden Definitionen und Sätze immer für den Fall $p = 1$ angeben, um nicht zu viele Notationen zu produzieren. Diese gelten dann stets in diesem Sinne auch für allgemeines $p \in \mathbb{Z}$.

Beispiel 11.4. (a) Wir haben im Kapitel 7 bereits eine Reihe kennengelernt, wir haben sie dort nur noch nicht so genannt, nämlich $\sum_{k=0}^{\infty} \frac{1}{k!}$ mit $s_n = 1 + 1 + 1/2! + 1/3! + \cdots + 1/n!$. Wir haben dort in Satz 7.7 gesehen, dass diese Folge konvergiert mit $\lim_{n\to\infty} s_n = \mathrm{e}$. Also ist dieses nach Definition eine konvergente Reihe mit Reihenwert

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

(b) Um die Reihe

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

zu untersuchen, beobachten wir, dass für jedes $n \in \mathbb{N}$ gilt

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

Also ist

$$\begin{aligned}
s_n &= a_1 + a_2 + \cdots + a_n \\
&= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\
&= 1 + \left(\frac{1}{2} - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n}\right) - \frac{1}{n+1} = 1 - \frac{1}{n+1}.
\end{aligned}$$

Für $n \to \infty$ folgt also $s_n \to 1$ und damit haben wir auch hier eine konvergente Reihe mit Reihenwert 1, in Formeln

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

Eine solche Summe, bei der sich jeweils aufeinanderfolgende Summanden so wegheben, dass nur am Anfang und am Ende etwas übrigbleibt, nennt man *Teleskopsumme*.

(c) In diesem und im folgenden Punkt wollen wir zwei weitere Reihen untersuchen, die einen Namen haben, was auch in diesem Fall ein deutlicher Hinweis auf ihre Bedeutung ist. Da ist zum einen die *geometrische Reihe*, die für $x \in \mathbb{R}$ gegeben ist durch

$$\sum_{k=0}^{\infty} x^k \quad \text{(geometrische Reihe).}$$

In Satz 7.3 haben wir schon gesehen, dass für die Partialsummen dieser Reihe

$$s_n = 1 + x + x^2 + x^3 + \cdots + x^n = \begin{cases} \dfrac{1 - x^{n+1}}{1 - x}, & \text{falls } x \neq 1, \\ n + 1, & \text{falls } x = 1 \end{cases}$$

gilt.

Für $x = 1$ konvergiert diese Folge offensichtlich nicht und nach Satz 7.2 konvergiert die Folge (x^{n+1}) genau dann, wenn $x \in (-1, 1]$ ist und dann ist der Grenzwert 0, falls $x \neq 1$ ist. Fazit: Unsere Folge (s_n) und damit *die geometrische Reihe konvergiert genau dann, wenn* $|x| < 1$ *gilt* und in diesem Fall ist

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

(d) Abschließend betrachten wir die *harmonische Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots.$$

Für diese Reihe gilt

$$\begin{aligned} s_{2n} &= \underbrace{1 + \frac{1}{2} + \dots + \frac{1}{n}}_{=s_n} + \underbrace{\frac{1}{n+1}}_{\geq 1/(2n)} + \underbrace{\frac{1}{n+2}}_{\geq 1/(2n)} + \dots + \frac{1}{2n} \\ &\geq s_n + n \cdot \frac{1}{2n} = s_n + \frac{1}{2}. \end{aligned}$$

Nehmen wir an, die Folge (s_n), und damit die harmonische Reihe, wären konvergent, so existiert $s := \lim_{n\to\infty} s_n$. Da dann aber auch die Teilfolge (s_{2n}) gegen s konvergiert, liefert uns obige Ungleichung den Widerspruch $s \geq s + 1/2$. Also ist die *harmonische Reihe divergent*, obwohl die Folge, über die summiert wird, gegen Null strebt.

Wir beweisen nun die ersten Aussagen über Konvergenz von Reihen. Dazu können wir insbesondere unsere Konvergenzkriterien für Folgen auf den Fall von Reihen übertragen.

Satz 11.5. *Es sei (a_n) eine reelle Folge und $s_n := \sum_{k=1}^{n} a_k$, $n \in \mathbb{N}$. Dann gilt*

(a) **(Monotonie-Kriterium)** *Ist $a_n \geq 0$ für alle $n \in \mathbb{N}$ und die Folge (s_n) nach oben beschränkt, so ist $\sum_{k=1}^{\infty} a_n$ konvergent.*

(b) **(Cauchy-Kriterium)** *Die Reihe $\sum_{n=1}^{\infty} a_n$ ist genau dann konvergent, wenn für jedes $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert, so dass*

$$\left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon \quad \textit{für alle } n, m \in \mathbb{N} \textit{ mit } m > n > n_0 \textit{ gilt.}$$

Beweis. (a) Da alle a_n nicht-negativ sind, gilt für jedes $n \in \mathbb{N}$

$$s_{n+1} = \sum_{k=1}^{n} a_k + a_{n+1} = s_n + a_{n+1} \geq s_n,$$

die Folge (s_n) ist also monoton wachsend. Da sie nach Voraussetzung auch beschränkt ist, folgt die Behauptung direkt aus dem Monotonie-Kriterium für Folgen (Satz 6.11).

(b) Seien $m, n \in \mathbb{N}$ mit $n < m$. Dann gilt

$$|s_m - s_n| = |a_{n+1} + a_{n+2} + \cdots + a_m| = \left| \sum_{k=n+1}^{m} a_k \right|.$$

Die Behauptung folgt damit direkt aus dem Cauchy-Kriterium für Folgen (Satz 10.2) □

Außerdem gelten für konvergente Reihen die folgenden Aussagen.

Satz 11.6. *Es sei $\sum_{k=1}^{\infty} a_k$ eine konvergente Reihe in $\mathbb{R}$. Dann gilt:*

(a) Für jedes $\nu \in \mathbb{N}$ ist die Reihe $\sum_{k=\nu}^{\infty} a_k$ ebenfalls konvergent.

(b) Für $r_\nu := \sum_{k=\nu}^{\infty} a_k$, $\nu \in \mathbb{N}$, konvergiert die Folge (r_ν) und es ist $\lim_{\nu\to\infty} r_\nu = 0$.

(c) Die Folge (a_n) konvergiert gegen 0.

Beweis. Wir setzen wieder $s_n := \sum_{k=1}^{n} a_k$, $n \in \mathbb{N}$, und $s := \lim_{n\to\infty} s_n = \sum_{k=1}^{\infty} a_k$.

(a) Sei $\nu \in \mathbb{N}$. Für jedes $m \in \mathbb{N}$ mit $m \geq \nu$ setzen wir nun $\sigma_m = \sum_{k=\nu}^{m} a_k$. Dann gilt

$$\sigma_m = \sum_{k=1}^{m} a_k - \sum_{k=1}^{\nu-1} a_k = s_m - s_{\nu-1}.$$

Die Folge auf der rechten Seite der Gleichung ist für $m \to \infty$ konvergent, also konvergiert auch (σ_m) mit $\sigma_m \to s - s_{\nu-1}$ $(m \to \infty)$.

(b) Nach dem vorherigen Punkt gilt $r_\nu = s - s_{\nu-1}$ für jedes $\nu \in \mathbb{N}$. Da $\lim_{\nu\to\infty} s_{\nu-1} = s$ gilt, ist damit (r_ν) eine Nullfolge.

(c) Hierzu beobachten wir, dass $a_n = s_n - s_{n-1}$ für jedes $n \in \mathbb{N}$ gilt. Daraus folgt für $n \to \infty$ die Behauptung. □

Warnung 11.7. In (c) im obigen Satz gilt die Umkehrung *nicht*, wie das Beispiel der harmonischen Reihe, vgl. Beispiel 11.4 (d), zeigt.

Auch die Linearität des Grenzwerts überträgt sich auf Reihen.

Satz 11.8. *Es seien $\sum_{k=1}^{\infty} a_k$ und $\sum_{k=1}^{\infty} b_k$ zwei konvergente Reihen in $\mathbb{R}$ sowie $\alpha, \beta \in \mathbb{R}$. Dann konvergiert auch die Reihe $\sum_{k=1}^{\infty}(\alpha a_k + \beta b_k)$ mit*

$$\sum_{k=1}^{\infty}(\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^{\infty} a_k + \beta \sum_{k=1}^{\infty} b_k.$$

Beweis. Für $n \in \mathbb{N}$ setzen wir $s_n := \sum_{k=1}^n a_k$ und $t_n := \sum_{k=1}^n b_k$ sowie $r_n := \sum_{k=1}^n (\alpha a_k + \beta b_k)$. Dann gilt für jedes $n \in \mathbb{N}$

$$r_n = \sum_{k=1}^n (\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^n a_k + \beta \sum_{k=1}^n b_k = \alpha s_n + \beta t_n \longrightarrow \alpha \lim_{n\to\infty} s_n + \beta \lim_{n\to\infty} t_n$$

nach den Grenzwertsätzen für Folgen (Satz 6.8). □

Bis jetzt haben wir unsere Erkenntnisse über die Konvergenz von Folgen in die Sprache der Reihen übersetzt. Wir kommen nun zum ersten Begriff, der speziell zum Verständnis der Konvergenz von Reihen beiträgt.

Definition 11.9. *Eine Reihe* $\sum_{k=1}^\infty a_k$ *heißt* absolut konvergent, *wenn die Reihe* $\sum_{k=1}^\infty |a_k|$ *konvergiert.*

Wir zeigen zunächst, dass absolute Konvergenz ein stärkerer Begriff als Konvergenz ist.

Satz 11.10. *Es sei* $\sum_{k=1}^\infty a_k$ *eine absolut konvergente Reihe. Dann ist sie insbesondere konvergent und es gilt die* verallgemeinerte Dreiecksungleichung

$$\left| \sum_{k=1}^\infty a_k \right| \leq \sum_{k=1}^\infty |a_k|.$$

Beweis. Wir verwenden das Cauchy-Kriterium. Sei dazu $\varepsilon > 0$ gegeben. Dann gibt es dank der absoluten Konvergenz ein $n_0 \in \mathbb{N}$, so dass $\sum_{k=n+1}^m |a_k| < \varepsilon$ für alle $n, m \in \mathbb{N}$ mit $m > n > n_0$ gilt. Mit der Dreiecksungleichung gilt dann sofort auch

$$\left| \sum_{k=n+1}^m a_k \right| = |a_{n+1} + a_{n+2} + \cdots + a_m| \leq |a_{n+1}| + |a_{n+2}| + \cdots + |a_m| = \sum_{k=n+1}^m |a_k| < \varepsilon.$$

Also ist die Reihe nach dem Cauchy-Kriterium konvergent.
Setzen wir $s_n := \sum_{k=1}^n a_k$ und $\sigma_n := \sum_{k=1}^n |a_k|$ für jedes $n \in \mathbb{N}$, so erhalten wir wie oben mit der Dreiecksungleichung $|s_n| \leq \sigma_n$ für alle n und damit im Grenzwert $n \to \infty$

$$\left| \sum_{k=1}^\infty a_k \right| = |\lim_{n\to\infty} s_n| = \lim_{n\to\infty} |s_n| \leq \lim_{n\to\infty} \sigma_n = \sum_{k=1}^\infty |a_k|.$$ □

Diskussionsanregung: Warum durften wir bei der letzten Rechnung im obigen Beweis den Limes aus dem Betrag ziehen?

Wir betrachten nun noch ein Beispiel einer konvergenten Reihe, die nicht absolut konvergiert. Damit ist auch die natürlicherweise im Raum stehende Frage geklärt, ob vielleicht Konvergenz und absolute Konvergenz einfach das Selbe sind; das sind sie nicht!

Beispiel 11.11. Wir betrachten die *alternierende harmonische Reihe* oder auch *Leibniz-Reihe*

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \dots.$$

Diese ist *nicht* absolut konvergent, denn wenn wir Betragsstriche unter das Summenzeichen schreiben, erhalten wir genau die harmonische Reihe und diese konvergiert nach Beispiel 11.4 (d) nicht.
Im Kontrast dazu wollen wir nun zeigen, dass die alternierende harmonische Reihe konvergiert (im weiteren Verlauf der Vorlesung werden wir sogar noch sehen, dass der Reihenwert $\ln(2)$ beträgt, vgl. Beispiel 24.9). Sei dazu wieder $a_n := (-1)^{n+1}/n$ und $s_n := \sum_{k=1}^{n} a_k$ für jedes $n \in \mathbb{N}$.
Betrachten wir zunächst nur die geraden n, so finden wir

$$s_{n+2} = s_n + a_{n+1} + a_{n+2} = s_n + \underbrace{\frac{1}{n+1} - \frac{1}{n+2}}_{\geq 0} \geq s_n.$$

Die Teilfolge (s_{2n}) von (s_n) ist also monoton wachsend. Analog erkennen wir, dass die Teilfolge (s_{2n-1}) der ungeraden Folgenglieder monoton fällt. Außerdem gilt für jedes $k \in \mathbb{N}$

$$s_{2k+1} = s_{2k} + a_{2k+1} = s_{2k} + \frac{1}{2k+1} \geq s_{2k},$$

d. h. die ungeraden Folgenglieder liegen immer oberhalb des vorhergehenden geraden Folgenglieds. Setzen wir diese beiden Einsichten zusammen, bekommen wir für alle $k \in \mathbb{N}$.

$$s_2 \leq s_4 \leq \ldots \leq s_{2k} \leq s_{2k+1} \leq \ldots \leq s_3 \leq s_1.$$

Somit sind beide Teilfolgen auch beschränkt; die der geraden Folgenglieder nach oben durch s_1 und die der ungeraden nach unten durch s_2. Nach dem Monotonie-Kriterium sind daher beide Teilfolgen konvergent. Wir setzen $s := \lim_{n\to\infty} s_{2n}$ und $\sigma := \lim_{n\to\infty} s_{2n-1}$.
Nehmen wir uns noch einmal die Identität $s_{2n+1} = s_{2n} + a_{2n+1}$ vor und lassen darin n nach ∞ streben, so gelangen wir wegen $\lim_{n\to\infty} a_n = 0$ zu der Erkenntnis $\sigma = s + 0$, also $\sigma = s$.
Sei nun $\varepsilon > 0$ gegeben. Dann gilt für fast alle $k \in \mathbb{N}$ sowohl $s_{2k} \in U_\varepsilon(s)$ als auch $s_{2k-1} \in U_\varepsilon(s)$. Da aber jede natürliche Zahl entweder gerade oder ungerade ist, liegt damit $s_k \in U_\varepsilon(s)$ für fast alle $k \in \mathbb{N}$ und wir haben gezeigt, dass (s_n) und damit auch die Reihe konvergent ist.

12. Konvergenzkriterien für Reihen

Die Untersuchung einer Reihe auf Konvergenz kann ein verzwicktes Problem darstellen. In diesem Abschnitt wollen wir einige Kriterien beweisen, die dabei helfen können. Ein – eher seltener – Glücksfall liegt dabei vor, wenn wir einen Reihenwert tatsächlich exakt bestimmen können. In der Regel beschränken wir uns darauf, einer Reihe ansehen zu wollen, *ob* sie konvergiert.
Die meisten Konvergenz-Kriterien sind auf spezielle Klassen von Reihen zugeschnitten oder sie sind nicht komplett trennscharf, d. h. sie liefern in manchen Fällen einfach keine Entscheidung, ob eine konkrete Reihe konvergiert oder divergiert.

Satz 12.1 (Leibniz-Kriterium)**.** *Es sei (b_n) eine monoton fallende Folge mit $b_n > 0$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} b_n = 0$. Dann ist die Reihe $\sum_{n=1}^{\infty}(-1)^{n+1}b_n$ konvergent.*

Der Prototyp einer solchen Reihe ist die alternierende harmonische Reihe aus Beispiel 11.11. Schauen wir uns den Beweis der Konvergenz dieser Reihe in diesem Beispiel noch einmal genau an, so stellen wir fest, dass dabei nur verwendet wurde, dass die Folge $(1/n)$ eine monoton fallende Nullfolge mit positiven Folgengliedern ist. Wir können also die Beweisführung vollständig übernehmen, um diesen Satz zu beweisen. Wir verzichten deshalb darauf, das hier noch einmal auszuführen.

Satz 12.2. *Es seien (a_n) und (b_n) zwei reelle Folgen.*

(a) **(Majorantenkriterium)** *Ist $|a_n| \leq b_n$ für fast alle $n \in \mathbb{N}$ und konvergiert die Reihe $\sum_{n=1}^{\infty} b_n$, so ist die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent (man nennt (b_n) eine* konvergente Majorante*).*

(b) **(Minorantenkriterium)** *Ist $a_n \geq b_n \geq 0$ für fast alle $n \in \mathbb{N}$ und divergiert die Reihe $\sum_{n=1}^{\infty} b_n$, so divergiert auch die Reihe $\sum_{n=1}^{\infty} a_n$ (man nennt (b_n) eine* divergente Minorante*).*

Beweis. (a) Nach Voraussetzung gibt es ein $k \in \mathbb{N}$, so dass $|a_n| \leq b_n$ für alle $n \geq k$ gilt. Seien nun $n, m \in \mathbb{N}$ mit $m > n > k$ gegeben. Dann gilt

$$\sum_{j=n+1}^{m} |a_j| \leq \sum_{j=n+1}^{m} b_j = \left| \sum_{j=n+1}^{m} b_j \right|.$$

Zu gegebenem $\varepsilon > 0$ wird nach dem Cauchy-Kriterium der Ausdruck rechts für hinreichend große n, m kleiner als ε. Damit ist für diese n, m auch der

Ausdruck links kleiner als ε und die Behauptung folgt aus dem Cauchy-Kriterium.

(b) Den Beweis des Minorantenkriteriums können wir einfach auf das Majorantenkriterium zurückspielen. Wenn wir annehmen, dass $\sum_{n=1}^{\infty} a_n$ konvergent ist, so wäre (a_n) eine konvergente Majorante von $\sum_{j=1}^{\infty} b_j$ und somit diese Reihe konvergent, was ja gerade nicht sein soll. Also divergiert $\sum_{n=1}^{\infty} a_n$. □

Wir wollen diese sehr nützlichen Kriterien an zwei Beispielen erproben.

Beispiel 12.3. (a) Es sei $\alpha \in \mathbb{Q}$ mit $0 < \alpha \leq 1$. Dann ist die Reihe

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

divergent, denn für alle $n \in \mathbb{N}$ gilt $n^\alpha \leq n$ und damit $1/n^\alpha \geq 1/n$. Da die Reihe $\sum_{n=1}^{\infty} 1/n$ divergiert, ist $(1/n)$ eine divergente Minorante für unsere Reihe und diese damit auch divergent.

Wenn wir im späteren Verlauf der Vorlesung die allgemeine Potenz n^x für alle $x \in \mathbb{R}$ definiert haben werden, verallgemeinert sich auch dieser Beweis sofort und die Aussage bleibt richtig für alle $\alpha \in (0, 1]$.

(b) Nun stellt sich natürlich sofort die Frage, wie das denn für Exponenten α aussieht, die größer als 1 sind. Um eine Intuition zu bekommen, untersuchen wir versuchsweise die Reihe

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

Dazu zeigen wir zunächst mit Hilfe des Majorantenkriteriums, dass die Reihe

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

konvergiert. Warum? Weil dann wegen

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

auch obige Reihe konvergiert.

Wir beobachten dazu, dass für jedes $n \in \mathbb{N}$ gilt

$$\frac{1}{(n+1)^2} = \frac{1}{n^2 + 2n + 1} \leq \frac{1}{n^2 + n} = \frac{1}{n(n+1)}.$$

Wir wissen aber seit Beispiel 11.4 (b), dass die Reihe über $1/(n(n+1))$ konvergiert, also haben wir mit dieser Folge eine konvergente Majorante gefunden und mithin ist $\sum_{n=1}^{\infty} 1/n^2$ konvergent.

Wir haben gesehen, dass es für die Konvergenz der Reihe notwendig ist, dass über eine Nullfolge summiert wird, aber dass dies alleine kein hinreichendes Kriterium darstellt. Die Folgenglieder müssen sich in einem gewissen Sinne „schnell genug“ der Null annähern. Wir brauchen also Techniken, um diese Annäherungsgeschwindigkeit zu messen. Die zwei folgenden Sätze basieren auf dieser Idee.

Satz 12.4 (Wurzelkriterium)**.** *Es sei (a_n) eine reelle Zahlenfolge.*

(a) Ist die Folge $(\sqrt[n]{|a_n|})$ beschränkt und $\alpha := \limsup_{n\to\infty} \sqrt[n]{|a_n|}$, so ist die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent, wenn $\alpha < 1$ ist und divergent, falls $\alpha > 1$ gilt.

(b) Ist die Folge $(\sqrt[n]{|a_n|})$ unbeschränkt, so divergiert die Reihe $\sum_{n=1}^{\infty} a_n$.

Beachten Sie, dass dieser Satz im Falle $\alpha = 1$ keine Aussage trifft. Das ist kein Versehen, sondern nicht zu ändern, denn es gibt bei Folgen mit $\alpha = 1$ sowohl solche, für welche die Reihe konvergiert, als auch solche, für die sie divergiert. Spuckt die Überprüfung dieses Kriteriums also $\alpha = 1$ aus, so muss man sich etwas Neues einfallen lassen.

Beweis von Satz 12.4. Wir betrachten zunächst den Fall, dass die Folge $(\sqrt[n]{|a_n|})$ beschränkt ist und $\alpha < 1$. Dazu wählen wir ein $q \in (\alpha, 1)$. Dann gilt $\sqrt[n]{|a_n|} \le q$ für fast alle $n \in \mathbb{N}$. Um das einzusehen, nehmen wir an, es gäbe unendlich viele $n \in \mathbb{N}$ für die $\sqrt[n]{|a_n|} > q$ gilt. Dann könnten wir die Teilfolge dieser Elemente auswählen. Diese wäre dann als Teilfolge der beschränkten Folge $(\sqrt[n]{|a_n|})$ auch beschränkt, sie besäße also nach dem Satz von Bolzano-Weierstraß einen Häufungswert. Dieser wäre dann auch ein Häufungswert der gesamten Folge, und außerdem wäre er größer oder gleich q und damit insbesondere echt größer als α. Da α der größte Häufungswert von $(\sqrt[n]{|a_n|})$ ist, ergibt sich ein Widerspruch.

Für fast alle $n \in \mathbb{N}$ gilt damit $\sqrt[n]{|a_n|} \le q$, d. h. $|a_n| \le q^n$. Nun ist aber wegen $0 < q < 1$ die Reihe $\sum_{n=1}^{\infty} q^n$ konvergent (geometrische Reihe), und damit haben wir eine konvergente Majorante für unsere Reihe gefunden, die damit auch selbst konvergiert.

Ist $\alpha > 1$ oder die Folge $(\sqrt[n]{|a_n|})$ sogar unbeschränkt, dann gibt es unendlich viele $n \in \mathbb{N}$, für die $\sqrt[n]{|a_n|} \ge 1$ gilt. Das heißt aber, dass für diese n auch $|a_n| \ge 1^n = 1$ gilt. Damit konvergiert die Folge (a_n) sicher nicht gegen Null, was aber nach Satz 11.6 (c) eine notwendige Voraussetzung für die Konvergenz der Reihe ist. Also divergiert die Reihe $\sum_{n=1}^{\infty} a_n$ in diesen Fällen. $\square$

Wir betrachten wieder einige Beispiele für die Anwendung dieses Satzes, bei denen wir auch sehen werden, dass im Fall $\alpha = 1$ tatsächlich sowohl Konvergenz als auch Divergenz der Reihe auftreten können.

Beispiel 12.5. (a) Für die Reihe $\sum_{n=1}^{\infty} 1/n$ ist $a_n = 1/n$ und damit $\sqrt[n]{|a_n|} = 1/\sqrt[n]{n}$. Diese Folge konvergiert gegen 1. Also ist für diese Reihe $\alpha = 1$ und sie divergiert, vgl. Beispiel 11.4 (d).

(b) Für die Reihe $\sum_{n=1}^{\infty} 1/n^2$ gilt wegen $\lim_{n\to\infty} \sqrt[n]{n^2} = \lim_{n\to\infty}(\sqrt[n]{n})^2 = 1^2 = 1$ genauso $\alpha = 1$, aber wie wir bereits gesehen haben, ist die Reihe in diesem Fall konvergent, vgl. Beispiel 12.3 (b).

(c) Wir betrachten die Reihe $\displaystyle\sum_{n=1}^{\infty} \left(1 + \frac{1}{n}\right)^{n^2}$.

Für diese gilt mit $a_n = (1 + 1/n)^{n^2}$

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = \mathrm{e}.$$

Also ist $\alpha = \mathrm{e} > 1$ und damit die Reihe divergent.

(d) Wir untersuchen schließlich die Reihe $\sum_{n=1}^{\infty} a_n$ mit

$$a_n := \begin{cases} \left(\frac{1}{2}\right)^n, \text{ falls } n \text{ gerade}, \\ \left(\frac{1}{3}\right)^n, \text{ falls } n \text{ ungerade}. \end{cases}$$

Dann gilt $\sqrt[n]{|a_n|} = 1/2$ für alle geraden n und $\sqrt[n]{|a_n|} = 1/3$ für alle ungeraden. Der größte Häufungswert dieser Folge ist $1/2$, also ist $\alpha = 1/2 < 1$ und unsere Reihe damit absolut konvergent.

Satz 12.6 (Quotientenkriterium). *Es sei (a_n) eine reelle Folge mit $a_n \neq 0$ für alle $n \in \mathbb{N}$. Dann gilt:*

(a) Ist die Folge $q_n := |a_{n+1}/a_n|$, $n \in \mathbb{N}$, beschränkt und gilt

$$\alpha := \limsup_{n\to\infty} q_n < 1,$$

so konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ absolut.

(b) Ist $q_n \geq 1$ für fast alle $n \in \mathbb{N}$, so divergiert die Reihe $\sum_{n=1}^{\infty} a_n$.

Beweis. (a) Wie im Beweis des Wurzelkriteriums wählen wir ein $q \in (\alpha, 1)$ und folgern, dass $q_n \leq q$ für fast alle $n \in \mathbb{N}$ gilt. Dann können wir ohne Beschränkung der Allgemeinheit annehmen, das $q_n \leq q$ für alle $n \in \mathbb{N}$ gilt, denn die endlich vielen anderen Summanden sind für die Konvergenz unerheblich. Damit gilt $|a_{n+1}| \leq q|a_n|$ für alle $n \in \mathbb{N}$, also ist

$$|a_2| \leq q|a_1|, \quad |a_3| \leq q|a_2| \leq q^2|a_1|, \quad |a_4| \leq q|a_3| \leq q^3|a_1|, \ \dots$$

und somit allgemein $|a_n| \leq q^{n-1}|a_1|$ für alle $n \in \mathbb{N}$.

Nun ist wieder mit Hilfe der geometrischen Reihe und $0 < q < 1$ die Reihe

$$\sum_{n=1}^{\infty} |a_1| q^{n-1} = |a_1| \sum_{n=0}^{\infty} q^n$$

konvergent und dient damit als konvergente Majorante.

(b) Sei $k \in \mathbb{N}$ so groß, dass $q_n = |a_{n+1}|/|a_n| \geq 1$ für alle $n \geq k$ gilt. Dann ist

$$|a_{k+1}| \geq |a_k|, \quad |a_{k+2}| \geq |a_{k+1}| \geq |a_k|, \quad |a_{k+3}| \geq |a_{k+2}| \geq |a_k|, \; \ldots$$

und allgemein $|a_{k+j}| \geq |a_k| \neq 0$ für alle $j \in \mathbb{N}$. Damit kann (a_n) wieder keine Nullfolge sein, d. h. die Reihe über (a_n) ist divergent. □

Wahrscheinlich ist Ihnen aufgefallen, dass keines dieser Kriterien eine Aussage über den Reihenwert trifft. Tatsächlich ist die exakte Bestimmung des Reihenwertes im Allgemeinen ein äußerst schwieriges Problem.

Beispiel 12.7. Als Beispielanwendung für das Quotientenkriterium untersuchen wir für jedes $x \in \mathbb{R}$ die Reihe

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

Die Konvergenz dieser Reihe ist offensichtlich für $x = 0$, da sie dann nur aus einem Summanden besteht, und für $x = 1$ schon bekannt (vgl. Satz 7.7). Sei nun $x \neq 0$, dann ist zunächst jeder Summand von Null verschieden, so dass wir das Quotientenkriterium anwenden können. In diesem Fall ist $a_n = x^n/n!$, so dass wir für q_n erhalten:

$$q_n = \left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{x^{n+1}}{(n+1)!} \cdot \frac{n!}{x^n} \right| = \frac{|x|}{n+1} \longrightarrow 0 \; (n \to \infty).$$

Damit ist $\limsup_{n\to\infty} |a_{n+1}/a_n| = 0$, d. h. die Reihe ist absolut konvergent und zwar für jedes beliebige $x \in \mathbb{R}$.

Das Ergebnis des vorhergehenden Beispiels rechtfertigt die folgende Definition.

Definition 12.8. *Wir definieren die* Exponentialfunktion $E : \mathbb{R} \to \mathbb{R}$ *durch*

$$E(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}.$$

Im weiteren Verlauf der Vorlesung werden wir zeigen, dass der Name dieser Funktion insofern gerechtfertigt ist, als $\mathrm{e}^r = E(r)$ für alle $r \in \mathbb{Q}$ gilt. Dieses ermöglicht uns dann schließlich durch die Setzung $\mathrm{e}^x := E(x)$ für beliebiges $x \in \mathbb{R}$, die allgemeine Potenz mit reellen Exponenten zu definieren.
Wir wollen nun noch die Frage aus Beispiel 12.3 (b) endgültig beantworten, für welche $\alpha > 0$ die Reihe über $1/n^\alpha$ konvergiert.

Satz 12.9. *Ist $\alpha \in \mathbb{Q}$ mit $\alpha > 1$, so konvergiert die Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}.$$

Beweis. Wie üblich setzen wir $s_n = \sum_{k=1}^{n} 1/k^\alpha$, $n \in \mathbb{N}$. Zu gegebenem $n \in \mathbb{N}$ sei nun $k \in \mathbb{N}$ mit $2^k - 1 \geq n$ gewählt. Damit ist

$$\begin{aligned} s_n \leq s_{2^k-1} = 1 + \underbrace{\frac{1}{2^\alpha} + \frac{1}{3^\alpha}}_{\leq 2 \cdot \frac{1}{2^\alpha}} + \underbrace{\frac{1}{4^\alpha} + \cdots + \frac{1}{7^\alpha}}_{\leq 4 \cdot \frac{1}{4^\alpha}} + \underbrace{\frac{1}{8^\alpha} + \cdots + \frac{1}{15^\alpha}}_{\leq 8 \cdot \frac{1}{8^\alpha}} + \dots \\ + \underbrace{\frac{1}{(2^{k-1})^\alpha} + \cdots + \frac{1}{(2^k - 1)^\alpha}}_{\leq 2^{k-1} \cdot \frac{1}{(2^{k-1})^\alpha}} \\ \leq \sum_{j=0}^{k-1} \frac{2^j}{(2^j)^\alpha} = \sum_{j=0}^{k-1} \left(\frac{1}{2^{\alpha-1}} \right)^j. \end{aligned}$$

Setzt man nun $q := 1/2^{\alpha-1}$, so ist wegen $\alpha > 1$ der Exponent positiv und deshalb $0 < q < 1$. Damit können wir mit Hilfe der geometrischen Reihe weiter abschätzen:

$$s_n \leq \sum_{j=0}^{k-1} q^j \leq \sum_{j=0}^{\infty} q^j = \frac{1}{1-q}.$$

Also ist die Folge (s_n) nach oben beschränkt und da alle Summanden unserer Reihe positiv sind, folgt die Behauptung aus dem Monotonie-Kriterium für Reihen, Satz 11.5 (a). $\square$

13. Umordnungen von Reihen

Summiert man endlich viele Zahlen auf, so ist die Reihenfolge der Zahlen egal, es gilt das Kommutativgesetz. Wie Sie schon in der OWO-Vorlesung gesehen haben, geht diese alltägliche Gewissheit verloren, wenn wir unendlich viele Zahlen aufsummieren, also Reihen betrachten.
Wir formulieren zunächst das „Umsortieren“ der Summation mathematisch präzise.

Definition 13.1. *Es sei (a_n) eine reelle Folge und $\varphi : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Wir setzen nun $b_n := a_{\varphi(n)}$ für jedes $n \in \mathbb{N}$. Dann heißt die Folge (b_n) eine* Umordnung *von (a_n) und die Reihe $\sum_{n=1}^{\infty} b_n$ eine* Umordnung *von $\sum_{n=1}^{\infty} a_n$.*

Bemerkung 13.2. Beachten Sie, dass wegen der Bijektivität von φ gewährleistet ist, dass wir die Folge, bzw. Reihe wirklich nur umsortieren und dabei nicht ein Folgenglied doppelt verwenden oder eines weglassen. Damit lässt sich jede Umordnung auch rückgängig machen. Das wird auch daraus ersichtlich, dass wegen der Bijektivität die Umkehrabbildung φ^{-1} existiert mit $\varphi(\varphi^{-1}(n)) = n$ für jedes $n \in \mathbb{N}$, für die dann $b_{\varphi^{-1}(n)} = a_{\varphi(\varphi^{-1}(n))} = a_n$ folgt. Also ist, wann immer (b_n) eine Umordnung von (a_n) ist, auch (a_n) eine Umordnung von (b_n).

Wir beginnen unsere Untersuchungen mit der folgenden kurzen Überlegung.

Lemma 13.3. *Es sei $\varphi : \mathbb{N} \to \mathbb{N}$ bijektiv und $n_0 \in \mathbb{N}$ gegeben. Dann ist $\varphi(n) \geq n_0$ für fast alle $n \in \mathbb{N}$.*

Beweis. Wir nehmen an, es wäre $\varphi(n) < n_0$ für unendlich viele n. Für diese unendlich vielen n wäre dann aber $\varphi(n) \in \{1, 2, \dots, n_0 - 1\}$. Da diese Menge nur endlich ist, ist das ein Widerspruch zur Bijektivität von φ. □

Satz 13.4. *Es sei (a_n) eine reelle Folge und (b_n) eine Umordnung von (a_n). Dann gilt:*

(a) Ist (a_n) eine konvergente Folge, so ist auch (b_n) konvergent und es gilt $\lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n$.

(b) Ist die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent, so ist auch ihre Umordnung $\sum_{n=1}^{\infty} b_n$ absolut konvergent und die Reihenwerte stimmen überein.

Beweis. Es sei $\varphi : \mathbb{N} \to \mathbb{N}$ die zur Umordnung gehörende bijektive Abbildung, es gelte also $b_n = a_{\varphi(n)}$ für alle $n \in \mathbb{N}$. Zum Beweis der Aussage für Folgen setzen wir $a := \lim_{n\to\infty} a_n$ und wählen $\varepsilon > 0$. Dann gibt es ein $m_0 \in \mathbb{N}$, so dass $|a_n - a| < \varepsilon$ für alle $n \geq m_0$ gilt. Nach Lemma 13.3 gibt es nun ein $n_0 \in \mathbb{N}$, so dass auch $\varphi(n) \geq m_0$ für alle $n \geq n_0$ gilt. Für alle diese n ist dann

$$|b_n - a| = |a_{\varphi(n)} - a| < \varepsilon.$$

Wir wenden uns der zweiten Aussage des Satzes zu. Hierzu betrachten wir zunächst den Spezialfall, dass $a_n \geq 0$ für alle $n \in \mathbb{N}$ gilt. Dann setzen wir $\sigma_n := \sum_{k=1}^{n} b_k$, $n \in \mathbb{N}$. Sei nun $j \in \mathbb{N}$ und dazu ein $N \in \mathbb{N}$ so gewählt, dass $\{b_1, \dots, b_j\} \subseteq \{a_1, \dots, a_N\}$ gilt. Dann wissen wir, dass $\sigma_j \leq a_1 + a_2 + \cdots + a_N$ ist. Da alle Folgenglieder von (a_n) positiv sind, machen wir diesen Ausdruck höchstens größer, wenn wir alle restlichen auch noch dazu addieren und erhalten

$$\sigma_j \leq \sum_{k=1}^{N} a_k \leq \sum_{k=1}^{\infty} a_k =: s.$$

Diese Setzung von s ist sinnvoll, da nach Voraussetzung diese Reihe konvergiert. Insbesondere ist damit die Partialsummen-Folge (σ_n) beschränkt und wegen der Positivität der Summanden konvergiert die Reihe $\sum_{k=1}^{\infty} b_k$ nach dem Monotonie-Kriterium, vgl. Satz 11.5 (a). Somit ist diese auch absolut konvergent, denn im Falle, dass alle Folgenglieder positiv sind, fallen diese beiden Begriffe zusammen. Außerdem gilt

$$\sum_{n=1}^{\infty} b_n = \lim_{n\to\infty} \sigma_n \leq \sum_{n=1}^{\infty} a_n = s.$$

Nun ist aber (vgl. Bemerkung 13.2) auch $\sum_{n=1}^{\infty} a_n$ eine Umordnung der Reihe über b_n. Man kann also diesen Beweis mit vertauschten Rollen von a_n und b_n noch einmal führen und erhält dann die umgekehrte Ungleichung

$$s = \sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n.$$

Also sind die beiden Reihenwerte gleich.

Es bleibt noch der Fall einer Folge (a_n) zu betrachten, deren Folgenglieder beliebige Vorzeichen haben. Mit dem ersten Teil dieses Beweises ist dann aber in jedem Fall die Reihe $\sum_{n=1}^{\infty} |b_n|$ konvergent und es gilt $\sum_{n=1}^{\infty} |b_n| = \sum_{n=1}^{\infty} |a_n|$, denn bei diesen Reihen sind alle Summanden nicht-negativ. Also ist die Reihe über b_n absolut konvergent.

Es bleibt noch die Gleichheit der Reihenwerte zu zeigen. Dazu setzen wir

$$\alpha_n := a_n + |a_n|, \qquad \beta_n := b_n + |b_n|, \qquad n \in \mathbb{N}.$$

Dann gilt $\alpha_n \geq 0$ und $\beta_n \geq 0$ für alle $n \in \mathbb{N}$ und (β_n) ist eine Umordnung von (α_n) (mit der gleichen Funktion φ). Außerdem ist wegen der absoluten Konvergenz von $\sum_{n=1}^{\infty} a_n$ die Reihe $\sum_{n=1}^{\infty} \alpha_n$ konvergent, ja sogar absolut konvergent, denn alle Folgenglieder von (α_n) sind nicht-negativ. Nun können wir für die Reihe $\sum_{n=1}^{\infty} \beta_n$ wieder den schon bewiesenen Fall von oben verwenden und wissen hiermit, dass auch diese Reihe absolut konvergiert und $\sum_{n=1}^{\infty} \alpha_n = \sum_{n=1}^{\infty} \beta_n$ gilt. Damit ergibt sich schließlich

$$\begin{aligned}\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} (\beta_n - |b_n|) &= \sum_{n=1}^{\infty} \beta_n - \sum_{n=1}^{\infty} |b_n| = \sum_{n=1}^{\infty} \alpha_n - \sum_{n=1}^{\infty} |a_n| \\ &= \sum_{n=1}^{\infty} (\alpha_n - |a_n|) = \sum_{n=1}^{\infty} a_n.\end{aligned}$$

Beachten Sie, dass das Auseinanderziehen und Zusammenfassen der Summation jeweils nach Satz 11.8 erlaubt ist, da wir nun von allen beteiligten Reihen wissen, dass sie absolut konvergent sind. □

In obigem Satz ist die Voraussetzung, dass die Konvergenz der Reihe absolut ist, wichtig, ohne diese wird die Aussage grob falsch. Der folgende Satz ist hoffentlich Warnung genug, diese Voraussetzung nicht zu vergessen.

Satz 13.5 (Riemannscher Umordnungssatz)**.** *Es sei $\sum_{n=1}^{\infty} a_n$ eine konvergente Reihe, die nicht absolut konvergiert. Dann gibt es für jedes $S \in \mathbb{R}$ eine Umordnung (b_n) von (a_n) mit $\sum_{n=1}^{\infty} b_n = S$.*
Weiter gibt es Umordnungen (c_n) und (d_n) von (a_n) mit $\sum_{n=1}^{\infty} c_n = \infty$ und $\sum_{n=1}^{\infty} d_n = -\infty$.

Beweis. Die Beweis-Idee kennen Sie aus der Owo-Vorlesung. Ich hoffe Sie sehen es als Herausforderung, damit nun den allgemeinen Satz selbst zu beweisen.

Wichtig sind unsere Überlegungen zu Umordnungen auch für die Frage, wie das Produkt zweier Reihen vernünftig definiert werden kann, denn dabei muss man unendlich viele Terme „ausmultiplizieren“ und sich dabei für eine Reihenfolge entscheiden. Dazu ist es gut zu wissen, dass es für das Ergebnis auf diese Reihenfolge nicht ankommt.
Gegeben seien also zwei Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$. Dann erhält man beim formalen Ausmultiplizieren alle Produkte der Form $a_j b_k$ mit $j, k \in \mathbb{N}_0$ je genau einmal. Wir ordnen nun all diese Produkte in einer Folge $(p_n)_{n \geq 0}$ an, wobei jedes einzelne $a_j b_k$ genau einmal vorkommen soll. Genauer gesagt, wählen wir eine bijektive Abbildung $\Phi : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ und setzen $p_{\Phi(j,k)} = a_j b_k$ für alle Paare $j, k \in \mathbb{N}_0$. Wir nennen dann $\sum_{n=0}^{\infty} p_n$ eine *Produktreihe* von $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$.

Beispiel 13.6. Wir geben hier zwei Beispiele solcher Anordnungen der Produkte, damit man sieht, dass solche existieren, die Wahl aber keineswegs offfensichtlich vorgegeben ist. Dazu schreiben wir uns die zu nummerierenden Folgenglieder in ein Schema und nummerieren in diesem auf verschiedenen Wegen:

(a) Anordnung nach Diagonalen:

$$\begin{array}{ccccc}
a_0b_0 & & a_0b_1 & & a_0b_2\dots \\
& \swarrow & & \swarrow & & \swarrow \\
a_1b_0 & & a_1b_1 & & a_1b_2\dots \\
& \swarrow & & \swarrow & \\
a_2b_0 & & a_2b_1 & & a_2b_2\dots \\
\vdots & \swarrow & \vdots & & \vdots
\end{array}$$

Wir wählen also

$$p_0 = a_0b_0,\ p_1 = a_0b_1,\ p_2 = a_1b_0,\ p_3 = a_0b_2,\ p_4 = a_1b_1,\ p_5 = a_2b_0,\ \dots$$

(b) Anordnung nach Quadraten:

$$\begin{array}{ccccc}
a_0b_0 & & a_0b_1 & & a_0b_2\dots \\
& & \downarrow & & \downarrow \\
a_1b_0 & \leftarrow & a_1b_1 & & a_1b_2\dots \\
& & & & \downarrow \\
a_2b_0 & \leftarrow & a_2b_1 & \leftarrow & a_2b_2\dots \\
\vdots & & \vdots & & \vdots
\end{array}$$

Wir wählen also

$$p_0 = a_0b_0,\ p_1 = a_0b_1,\ p_2 = a_1b_1,\ p_3 = a_1b_0,\ p_4 = a_0b_2,\ p_5 = a_1b_2,\ \dots$$

Dank Satz 13.4 können wir nun hoffen, dass es zumindest für absolut konvergente Reihen egal ist, welche der Anordnungen wir wählen. Tatsächlich gilt das folgende Resultat.

Satz 13.7. *Es seien $\sum_{n=0}^\infty a_n$ und $\sum_{n=0}^\infty b_n$ zwei absolut konvergente Reihen und $\sum_{n=0}^\infty p_n$ irgendeine ihrer Produktreihen. Dann konvergiert auch $\sum_{n=0}^\infty p_n$ absolut und es gilt*

$$\sum_{n=0}^\infty p_n = \left(\sum_{n=0}^\infty a_n\right)\left(\sum_{n=0}^\infty b_n\right).$$

Beweis. In einem ersten Schritt zeigen wir die Behauptung für den speziellen Fall der Anordnung nach Quadraten von oben. Wir setzen $s_n := \sum_{k=0}^n p_k$ und $\sigma_n := \sum_{k=0}^n |p_k|$, wobei (p_k) die Anordnung nach Quadraten sei. Wir wählen ein $n \in \mathbb{N}_0$ fix und betrachten alle Summanden von σ_n. Jeder einzelne ist von der

Form $a_j b_k$ mit $j, k \in \mathbb{N}_0$. Sei nun N der höchste Wert für j oder k, der dabei vorkommt. Dann gilt durch Hinzunahme von weiteren positiven(!) Summanden:

$$\begin{aligned}\sigma_n = |p_0| + |p_1| + \cdots + |p_n| &\le (|a_0| + \cdots + |a_N|)(|b_0| + \cdots + |b_N|) \\ &\le \left(\sum_{n=0}^{\infty} |a_n|\right)\left(\sum_{n=0}^{\infty} |b_n|\right).\end{aligned}$$

Also ist die Folge (σ_n) beschränkt. Nach dem Monotoniekriterium für Reihen ist damit die Reihe $\sum_{n=0}^{\infty} p_n$ absolut konvergent.
Damit konvergiert aber auch die Folge (s_n). Es sei $s := \lim_{n\to\infty} s_n = \sum_{n=0}^{\infty} p_n$. Für die Anordnung nach Quadraten im Speziellen gilt die Beziehung

$$s_{n^2+2n} = (a_0 + \cdots + a_n)(b_0 + \cdots + b_n),$$

die man z. B. durch Induktion beweisen kann. Lassen wir in dieser Gleichung nun $n \to \infty$ streben, erhalten wir die gewünschte Aussage

$$s = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

Sei nun $\sum_{n=0}^{\infty} \tilde{p}_n$ eine weitere Produktreihe von $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$. Dann ist diese eine Umordnung von $\sum_{n=0}^{\infty} p_n$ und damit gilt nach Satz 13.4 (b) die Behauptung. □

Am häufigsten arbeitet man aber mit der oben angegebenen Anordnung nach Diagonalen. Diese formalisieren wir in der folgenden Definition.

Definition 13.8. *Es seien zwei Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ gegeben. Wir setzen*

$$c_0 := a_0 b_0, \quad c_1 := a_0 b_1 + a_1 b_0, \quad c_2 := a_0 b_2 + a_1 b_1 + a_2 b_0, \ \ldots,$$

bzw. allgemein

$$c_n := a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0 = \sum_{k=0}^{n} a_k b_{n-k}.$$

Die Reihe

$$\sum_{n=0}^{\infty} c_n$$

heißt dann das Cauchy-Produkt *von $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$.*

Satz 13.9. *Sind $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei absolut konvergente Reihen, so ist das Cauchy-Produkt $\sum_{n=0}^{\infty} c_n$ der beiden auch absolut konvergent und es gilt*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

Der Beweis dieses Satzes ergibt sich sofort aus Satz 13.7 und dem folgenden Lemma über blockweises Aufsummieren.

Lemma 13.10. *Es sei $\sum_{n=0}^{\infty} a_n$ konvergent und $(n_1, n_2, n_3, \dots)$ eine streng monoton wachsende Folge in $\mathbb{N}$. Setzen wir*

$$b_0 := a_0 + a_1 + \dots + a_{n_1}$$

und für jedes $j \in \mathbb{N}$

$$b_j = a_{n_j+1} + a_{n_j+2} + \dots + a_{n_{j+1}},$$

so ist auch die Reihe $\sum_{n=0}^{\infty} b_n$ konvergent und es gilt

$$\sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} a_n.$$

Beweis. Für jedes $n \in \mathbb{N}$ setzen wir $s_n := \sum_{j=0}^{n} a_j$, $\sigma_n := \sum_{j=0}^{n} b_j$ und $s := \lim_{n\to\infty} s_n$. Dann gilt für jedes $k \in \mathbb{N}$

$$\sigma_k = a_1 + a_2 + \dots + a_{n_{k+1}} = s_{n_{k+1}}.$$

Also ist (σ_k) eine Teilfolge von (s_n) und da diese konvergiert gilt insbesondere auch $\sigma_k \to s$ $(k \to \infty)$, was die Behauptung beweist. □

Wir wollen das Cauchy-Produkt nun einmal in Aktion sehen und betrachten das folgende Beispiel.

Beispiel 13.11. Es seien $x, y \in \mathbb{R}$. Dann sind die Werte $E(x)$ und $E(y)$ der Exponentialfunktion durch absolut konvergente Reihen gegeben. Wir erhalten also mit dem Cauchyprodukt:

$$E(x)E(y) = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{y^n}{n!}\right) = \sum_{n=0}^{\infty} c_n,$$

wobei

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

ist. Dabei haben wir im letzten Schritt die Binomialformel verwendet. Setzen wir das wieder oben ein, erhalten wir den schönen Zusammenhang

$$E(x)E(y) = \sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \frac{1}{n!}(x+y)^n = E(x+y),$$

die sogenannte *Funktionalgleichung der Exponentialfunktion.*

14. Potenzreihen

Definition 14.1. *Es sei $(a_n)_{n\geq 0}$ eine Folge in $\mathbb{R}$. Eine Reihe der Form*

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots$$

heißt Potenzreihe.

Zwei Beispiele von Potenzreihen haben wir mit der Exponentialfunktion und der geometrischen Reihe schon gesehen. Wir wollen hier nun eine allgemeine Theorie solcher Reihen entwickeln, da diese ein wichtiges Hilfsmittel in der Analysis sind. Wie schon bei der Exponentialreihe ist es auch für jede Potenzreihe klar, dass sie für $x = 0$ konvergiert. Zur weiteren Untersuchung der Konvergenz einer Potenzreihe dient der folgende Satz, der sich aus dem Wurzelkriterium ableitet.

Satz 14.2 (Satz von Hadamard). *Es sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe. Dann gilt:*

(a) Ist die Folge $(\sqrt[n]{|a_n|})$ unbeschränkt, so konvergiert die Potenzreihe nur für $x = 0$.

(b) Ist die Folge $(\sqrt[n]{|a_n|})$ beschränkt und $\varrho := \limsup_{n\to\infty} \sqrt[n]{|a_n|} = 0$, so konvergiert die Potenzreihe für alle $x \in \mathbb{R}$ absolut.

(c) Ist die Folge $(\sqrt[n]{|a_n|})$ beschränkt und $\varrho > 0$, so ist die Potenzreihe für alle $x \in \mathbb{R}$ mit $|x| < 1/\varrho$ absolut konvergent und für alle $x \in \mathbb{R}$ mit $|x| > 1/\varrho$ divergent.

Wir erben vom Wurzelkriterium auch die Unsicherheit, die dort für den Grenzfall auftrat, dass der betrachtete Limes superior den Wert 1 annimmt. Diese macht sich hier insofern bemerkbar, als wir im dritten Punkt dieses Satzes für $x = 1/\varrho$ und $x = -1/\varrho$, also auf dem Rand des Konvergenzgebietes, keine Aussage treffen können.
Beachten Sie, dass der Konvergenzbereich einer Potenzreihe damit immer $\{0\}$ oder $\mathbb{R}$ oder von der Form $(-r, r)$, $[-r, r)$, $(-r, r]$ oder $[-r, r]$ für $r = 1/\varrho$ ist.
Somit ist die in diesem Zusammenhang definierte Zahl r eine wichtige charakteristische Größe der Potenzreihe, die einen Namen verdient hat.

Definition 14.3. *Es sei* $\sum_{n=0}^{\infty} a_n x^n$ *eine Potenzreihe und, im Falle, dass* $(\sqrt[n]{|a_n|})$ *beschränkt ist,* ϱ *wie in Satz 14.2. Dann heißt die Zahl*

$$r := \begin{cases} 0, & \text{falls in obigem Satz (a) gilt,} \\ \infty, & \text{falls in obigem Satz (b) gilt,} \\ \dfrac{1}{\varrho}, & \text{falls in obigem Satz (c) gilt,} \end{cases}$$

der Konvergenzradius *der Potenzreihe.*

Wir wollen nun noch den Satz von Hadamard beweisen.

Beweis von Satz 14.2. Wir zeigen zunächst (a). Dazu nehmen wir uns ein $x \neq 0$. Dann ist $\sqrt[n]{|a_n x^n|} = |x| \sqrt[n]{|a_n|}$ für jedes $n \in \mathbb{N}$ und diese Folge ist nach Voraussetzung unbeschränkt. Damit folgt die Divergenz der Reihe aus dem Wurzelkriterium (Satz 12.4).
Auch (b) und (c) ergeben sich aus dem Wurzelkriterium, denn hier gilt

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = \limsup_{n\to\infty} |x| \sqrt[n]{a_n} = |x| \cdot \varrho.$$

Dieser Limes superior ist, je nachdem, ob $|x| < 1/\varrho$ oder $|x| > 1/\varrho$ gilt, größer oder kleiner als 1, was nach dem Wurzelkriterium eben genau absolute Konvergenz von Divergenz der Reihe scheidet. □

Wir betrachten einige Beispiele, die verdeutlichen, dass am Rand des Konvergenzintervalls alles passieren kann.

Beispiel 14.4. (a) Wir beginnen mit $a_n := 1$, $n \in \mathbb{N}_0$, d. h. der Potenzreihe

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

Ja genau, das ist unsere altbekannte geometrische Reihe und insofern ist es auch nicht verwunderlich, dass hier $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} 1 = 1$ und damit der Konvergenzradius 1 ist, d. h. diese Potenzreihe konvergiert absolut für $|x| < 1$ und divergiert für $|x| > 1$. Über das Konvergenzverhalten am Rand des Konvergenzintervalls gibt uns der Konvergenzradius keine Auskunft. Für diese Reihe haben wir offensichtlich sowohl für $x = -1$ als auch für $x = 1$ eine divergente Reihe.

(b) Nun sei $a_n = 1/n$, $n \in \mathbb{N}$, wir betrachten also

$$\sum_{n=1}^{\infty} \frac{x^n}{n}.$$

Auch hier ist der Konvergenzradius 1, denn

$$\limsup_{n\to\infty} \sqrt[n]{\frac{1}{n}} = \limsup_{n\to\infty} \frac{1}{\sqrt[n]{n}} = \frac{1}{\lim_{n\to\infty} \sqrt[n]{n}} = 1.$$

Damit konvergiert diese Reihe für $x \in (-1,1)$ absolut und divergiert für $|x| > 1$. An den Rändern des Intervalls $(-1,1)$ erhalten wir für $x = 1$ genau die harmonische Reihe (divergent) und für $x = -1$ die alternierende harmonische Reihe (konvergent), so dass wir also Konvergenz für alle $x \in [-1,1)$ und Divergenz für alle anderen x haben.

(c) Schließlich nehmen wir $a_n = 1/n^2$, d. h. die Potenzreihe

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

Für diese ist ebenso der Konvergenzradius 1, aber diese Reihe konvergiert an beiden Randpunkten des Konvergenzintervalls, denn für $x = 1$ erhalten wir die nach Beispiel 12.3 (b) konvergente Reihe über $1/n^2$ und für $x = -1$ die nach dem Leibniz-Kriterium konvergente Reihe $\sum_{n=0}^{\infty}(-1)^n/n^2$. Also ist in diesem Fall das Konvergenzintervall der Potenzreihe $[-1,1]$ und wir haben Divergenz für alle $x \in \mathbb{R}$ mit $|x| > 1$.

Der Satz von Hadamard kann uns auch noch zu einer anderen, eher unerwarteten Erkenntnis verhelfen. Wir haben in Beispiel 13.11 mit Hilfe des Quotientenkriteriums gesehen, dass die Exponentialreihe $\sum_{n=0}^{\infty} x^n/n!$ für alle $x \in \mathbb{R}$ konvergiert. Nach dem Satz von Hadamard bedeutet dies, dass $\limsup_{n\to\infty} 1/\sqrt[n]{n!} = 0$ sein muss. Da diese Folge außerdem positiv ist, kann sie keinen weiteren Häufungswert haben, denn der müßte wegen der Positivität größer oder gleich Null sein. Beachten wir nun noch, dass diese Folge beschränkt ist, so konvergiert sie nach Satz 9.12 und es gilt

$$\lim_{n\to\infty} \frac{1}{\sqrt[n]{n!}} = \limsup_{n\to\infty} \frac{1}{\sqrt[n]{n!}} = 0.$$

So haben wir mit Hilfe eines Satzes über Potenzreihen einen ansonsten schwierig zu berechnenden Grenzwert bestimmt.
Natürlich kann man auch mit dem Quotientenkriterium Konvergenzuntersuchungen bei Potenzreihen anstellen. Wir führen das exemplarisch an einem prominenten Beispiel vor.

Beispiel 14.5. Wir betrachten die Potenzreihe

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots,$$

d. h. die Potenzreihe mit der Folge

$$a_n = \begin{cases} \dfrac{(-1)^n}{n!}, \text{ falls } n \text{ gerade,} \\ \qquad 0, \text{ falls } n \text{ ungerade.} \end{cases}$$

Wir setzen $t := x^2$, damit unsere Potenzreihe die Form $\sum_{n=0}^{\infty}(-1)^n/(2n)!\ t^n$ bekommt und betrachten die Folge $b_n = (-1)^n/(2n)!\ t^n$, $n \in \mathbb{N}$. Für den im Quotientenkriterium betrachteten Quotienten gilt

$$\left|\frac{b_{n+1}}{b_n}\right| = \left|\frac{t^{n+1}}{(2n+2)!} \cdot \frac{(2n)!}{t^n}\right| = \frac{|t|}{(2n+2)(2n+1)},$$

und dieser Ausdruck geht für jedes $t \in \mathbb{R}$ gegen Null für $n \to \infty$. Also konvergiert nach dem Quotientenkriterium für Reihen die Potenzreihe $\sum_{n=0}^{\infty}(-1)^n/(2n)!\ t^n$ für jedes $t \in \mathbb{R}$ absolut und damit ebenso unsere ursprünglich betrachtete Reihe für jedes $x \in \mathbb{R}$. Der Konvergenzradius ist also unendlich.

Was war daran nun prominent? Wie bei der Exponentialfunktion ist durch diese Potenzreihe nun eine Funktion auf ganz $\mathbb{R}$ gegeben, der wir einen Namen geben wollen.

Definition 14.6. *Für jedes $x \in \mathbb{R}$ nennen wir die Zahl*

$$\cos(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

den Cosinus *von x.*

Ganz ähnlich wie in obigem Beispiel kann man für die Reihe in der folgenden Definition ebenso absolute Konvergenz für alle $x \in \mathbb{R}$ zeigen.

Definition 14.7. *Für jedes $x \in \mathbb{R}$ nennen wir die Zahl*

$$\sin(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

den Sinus *von x.*

Wir wollen uns nun dem Cauchy-Produkt zweier Potenzreihen zuwenden.

Satz 14.8. *Es seien $\sum_{n=0}^{\infty} a_n x^n$ und $\sum_{n=0}^{\infty} b_n x^n$ Potenzreihen mit Konvergenzradien $r_1 > 0$ bzw. $r_2 > 0$ (dabei sind $r_1 = \infty$ oder $r_2 = \infty$ zugelassen). Wir setzen*

$$c_n := \sum_{k=0}^{n} a_k b_{n-k},\ n \in \mathbb{N}_0, \quad \textit{und} \quad R := \min\{r_1, r_2\}.$$

Dann ist das Cauchy-Produkt der beiden Potenzreihen die Potenzreihe $\sum_{n=0}^{\infty} c_n x^n$*, deren Konvergenzradius mindestens* R *beträgt und es gilt*

$$\sum_{n=0}^{\infty} c_n x^n = \left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) \quad \textit{für alle } |x| < R.$$

Beweis. Für $|x| < R$ sind beide Potenzreihen absolut konvergent, also konvergiert nach Satz 13.9 auch das Cauchy-Produkt der beiden absolut und es gilt

$$\begin{aligned}\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k x^k b_{n-k} x^{n-k} = \sum_{n=0}^{\infty}\sum_{k=0}^{n} x^n a_n b_{n-k} \\ &= \sum_{n=0}^{\infty} x^n \sum_{k=0}^{n} a_k b_{n-k} = \sum_{n=0}^{\infty} c_n x^n.\end{aligned}$$ □

Wir können abschnließend noch die folgenden Eigenschaften der Exponentialfunktion zeigen.

Satz 14.9. *(a)* $E(x)E(y) = E(x+y)$ *für alle* $x, y \in \mathbb{R}$.

(b) $E(0) = 1$ *und* $E(1) = \mathrm{e}$.

(c) $E(x) > 0$ *und* $E(x)^{-1} = E(-x)$ *für alle* $x \in \mathbb{R}$.

(d) $E(r) = \mathrm{e}^r$ *für alle* $r \in \mathbb{Q}$.

(e) Die Funktion $E : \mathbb{R} \to \mathbb{R}$ *ist streng monoton wachsend, d. h. für alle* $x, y \in \mathbb{R}$ *mit* $x < y$ *gilt* $E(x) < E(y)$.

Beweis. (a) siehe Beispiel 13.11.

(b) $E(0) = 1$ ist klar und $E(1) = \mathrm{e}$ ergibt sich aus der Definition der Eulerzahl.

(c) Mit Hilfe der ersten beiden Punkte gilt für alle $x \in \mathbb{R}$

$$1 = E(0) = E(x - x) = E(x)E(-x).$$

Also gilt $E(x) \neq 0$ und $E(-x) = E(x)^{-1}$. Um die Positivität von $E(x)$ nachzuweisen, betrachten wir zunächst den Fall $x \geq 0$. Dann gilt

$$E(x) = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} \geq 1 > 0.$$

Ist $x < 0$, so ist $E(x) = E(-x)^{-1} > 0$, da $-x > 0$ ist.

(d) Es sei zunächst $n \in \mathbb{N}$. Dann gilt

$$E(n) = E(\underbrace{1 + \cdots + 1}_{n \text{ Stück}}) = (E(1))^n = \mathrm{e}^n$$

und genauso

$$\mathrm{e} = E(1) = E\left(n \cdot \frac{1}{n}\right) = \left(E\left(\frac{1}{n}\right)\right)^n.$$

Also ist $E(1/n) = \mathrm{e}^{1/n}$.

Somit gilt für positive rationale Zahlen $r := m/n$ mit $n, m \in \mathbb{N}$ die Gleichung

$$E(r) = E\left(\underbrace{\frac{1}{n} + \cdots + \frac{1}{n}}_{m \text{ Stück}}\right) = \left(E\left(\frac{1}{n}\right)\right)^m = \left(\mathrm{e}^{1/n}\right)^m = \mathrm{e}^{m/n} = \mathrm{e}^r.$$

Ist $r \in \mathbb{Q}$ kleiner als Null, so ist $-r > 0$ und es gilt $E(r) = E(-r)^{-1} = (\mathrm{e}^{-r})^{-1} = \mathrm{e}^r$.

(e) Es seien $x, y \in \mathbb{R}$ mit $x < y$ gegeben. Dann ist $y - x > 0$ und damit gilt $E(y - x) > 1$ (vgl. den Beweis von (c)). Nun ist aber damit

$$1 < E(y - x) = E(y)E(-x) = \frac{E(y)}{E(x)},$$

also $E(y) > E(x)$. □

Bemerkung 14.10. Sei $x_0 \in \mathbb{R}$ und (a_n) eine reelle Folge. In Erweiterung von Definition 14.1 nennt man auch eine Reihe der Form

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

eine *Potenzreihe*. Die Zahl x_0 heißt dann *Entwicklungspunkt* der Potenzreihe. Alle in diesem Abschnitt für Potenzreihen bewiesenen Resultate gelten sinngemäß auch für Potenzreihen dieses allgemeineren Typs, wobei der Konvergenzbereich einer Potenzreihe mit Konvergenzradius $r > 0$ dann das Intervall $(x_0 - r, x_0 + r)$ ist.

15. g-adische Entwicklungen

Wir führen nun die Dezimalbruchdarstellung der reellen Zahlen ein. Am Ende des Abschnitts werden wir dann den schon vor langer Zeit versprochenen Beweis der Überabzählbarkeit von $\mathbb{R}$ führen.
In diesem gesamten Abschnitt ist $g \geq 2$ eine natürliche Zahl, die Basis des untersuchten Stellenwertsystems.

Definition 15.1. *Es sei $a \in \mathbb{R}$. Dann heißt die größte ganze Zahl, die kleiner oder gleich a ist, also die Zahl*

$$[a] := \max\{k \in \mathbb{Z} : k \leq a\},$$

die Gauß-Klammer *von a.*

Man beachte, dass immer die Ungleichungskette

$$[a] \leq a < [a] + 1$$

gilt.
Sei nun $a \in \mathbb{R}$ mit $a > 0$ gewählt. Wir erhalten die g-adische Entwicklung von a folgendermaßen: Wir setzen zunächst

$$z_0 := [a].$$

Dann gilt $z_0 \leq a < z_0 + 1$ und $z_0 \in \mathbb{N}_0$. Nehmen wir weiter

$$z_1 := \big[(a - z_0) \cdot g\big],$$

so ist $z_1 \leq (a - z_0)g \leq z_1 + 1$ und $z_1 \in \{0, 1, \ldots, g-1\}$, denn $a - z_0 < 1$. Also haben wir nun

$$z_0 + \frac{z_1}{g} \leq a < z_0 + \frac{z_1}{g} + \frac{1}{g}.$$

Im nächsten Schritt setzen wir dann

$$z_2 := \left[\left(a - z_0 - \frac{z_1}{g}\right)g^2\right]$$

und haben damit

$$z_0 + \frac{z_1}{g} + \frac{z_2}{g^2} \leq a < z_0 + \frac{z_1}{g} + \frac{z_2}{g^2} + \frac{1}{g^2}.$$

Allgemein definieren wir rekursiv

$$z_{n+1} := \left[\left(a - \sum_{j=0}^{n} \frac{z_j}{g^j}\right) \cdot g^{n+1}\right], \quad n \in \mathbb{N}.$$

Dann hat die Folge (z_n) die folgenden Eigenschaften, die sich direkt aus der Konstruktion ergeben.

Satz 15.2. *(a) $z_n \in \mathbb{N}_0$ für alle $n \in \mathbb{N}_0$.*

(b) $z_n \leq g - 1$ für alle $n \in \mathbb{N}$ (ohne Null!).

(c) $\displaystyle\sum_{j=0}^{n} \frac{z_j}{g^j} \leq a < \sum_{j=0}^{n} \frac{z_j}{g^j} + \frac{1}{g^n}$ für alle $n \in \mathbb{N}_0$.

(d) Ist $(\tilde{z}_n)$ eine weitere Folge, die die Punkte (a) – (c) erfüllt, dann gilt $(\tilde{z}_n) = (z_n)$.

Nun betrachten wir die Reihe

$$\sum_{n=0}^{\infty} \frac{z_n}{g^n}.$$

Da $z_n \leq g - 1$ für alle $n \geq 1$ gilt, ist $0 \leq z_n/g^n \leq (g-1)/g^n$ für alle diese n. Die Reihe $(g-1)\sum_{n=0}^{\infty} 1/g^n$ ist aber konvergent (geometrische Reihe), also haben wir eine konvergente Majorante gefunden. Damit konvergiert unsere Reihe absolut nach dem Majorantenkriterium. Da wegen (c) aus obigem Satz

$$\sum_{j=0}^{\infty} \frac{z_j}{g^j} = \lim_{n\to\infty} \sum_{j=0}^{n} \frac{z_j}{g^j} \leq a \leq \lim_{n\to\infty} \left(\sum_{j=0}^{n} \frac{z_j}{g^j} + \frac{1}{g^n}\right) = \sum_{j=0}^{\infty} \frac{z_j}{g^j}$$

gilt, erhalten wir

$$a = \sum_{n=0}^{\infty} \frac{z_n}{g^n}. \tag{15.1}$$

Definition 15.3. *Sei $a \in (0, \infty)$ und $g \in \mathbb{N}$ mit $g \geq 2$. Dann heißt die Darstellung aus* (15.1) *mit der oben konstruierten Folge (z_n) die g*-adische Entwicklung *von a. Man schreibt*

$$a = z_0\,,\, z_1\, z_2\, z_3\, z_4\, \ldots$$

Dass wir tatsächlich die alltägliche Dezimalbruchdarstellung konstruiert haben, verdeutlichen wir uns anhand zweier Beispiele.

Beispiel 15.4. Wir wählen $g = 10$ und $a = 1$. Dann ist $z_0 = 1$ und $z_1 = [(1 - z_0)g] = 0$. Genauso sieht man, dass $z_n = 0$ sogar für alle $n \geq 1$ gilt, also ist

$$1 = 1{,}00000\ldots$$

Wählen wir $a = 1/2$, so ist $z_0 = 0$ und $z_1 = [(1/2 - 0) \cdot 10] = 5$ sowie $z_2 = [(1/2 - 1/2) \cdot 100] = 0$, was auch wieder für alle $n \geq 2$ zutrifft. Also haben wir

$$\frac{1}{2} = 0,5000000\ldots$$

Bemerkung 15.5. Man beachte, dass (d) aus Satz 15.2 uns garantiert, dass die g-adische Entwicklung eindeutig ist. Tatsächlich hätte es noch eine andere, gleichwertige Möglichkeit gegeben, wenn wir auf gleiche Weise eine Folge $(\tilde{z}_n)$ konstruiert hätten mit

$$\sum_{j=0}^{n} \frac{\tilde{z}_j}{g^j} < a \leq \sum_{j=0}^{n} \frac{\tilde{z}_j}{g^j} + \frac{1}{g^n}$$

für jedes $n \in \mathbb{N}_0$. Dann wäre für $a = 1$ in obigem Beispiel $\tilde{z}_0 < 1 \leq \tilde{z}_0 + 1$, also $z_0 = 0$ und $\tilde{z}_n = 9$ für alle $n \geq 1$, so dass

$$1 = 0,9999999\ldots.$$

Für $a = 1/2$ bekäme man entsprechend

$$\frac{1}{2} = 0,499999999\ldots.$$

Nun darf jedeR eine eigene Wahl treffen, dieses Skript bleibt bei der oben in der Definition angegebenen Variante.

Wir zeigen, dass so lange Schlangen von Neunen, bzw. $(g-1)$-en in unserem System nicht auftreten können.

Satz 15.6. *Es sei $a \geq 0$ und $z_0, z_1 z_2 z_3 \ldots$ die g-adische Entwicklung von a gemäß Definition 15.3. Dann gilt $z_k \neq g-1$ für unendlich viele $k \in \mathbb{N}$.*

Beweis. Wir nehmen an, es wäre $z_k = g-1$ für fast alle $k \in \mathbb{N}$. Sei also $m \in \mathbb{N}$ so groß, dass $z_k = g-1$ für alle $k \geq m$ gilt. Dann gilt

$$\begin{aligned} a &= \sum_{n=0}^{\infty} \frac{z_n}{g^n} = \sum_{n=0}^{m-1} \frac{z_n}{g^n} + \sum_{n=m}^{\infty} \frac{z_n}{g^n} = \sum_{n=0}^{m-1} \frac{z_n}{g^n} + (g-1) \sum_{n=m}^{\infty} \frac{1}{g^n} \\ &= \sum_{n=0}^{m-1} \frac{z_n}{g^n} + \frac{g-1}{g^m} \sum_{n=m}^{\infty} \frac{1}{g^{n-m}} = \sum_{n=0}^{m-1} \frac{z_n}{g^n} + \frac{g-1}{g^m} \sum_{n=0}^{\infty} \frac{1}{g^n} \\ &= \sum_{n=0}^{m-1} \frac{z_n}{g^n} + \frac{g-1}{g^m} \cdot \frac{1}{1-\frac{1}{g}} = \sum_{n=0}^{m-1} \frac{z_n}{g^n} + \frac{1}{g^{m-1}} \end{aligned}$$

und das ist ein Widerspruch zu Satz 15.2 (c). □

Nun kommen wir zum versprochenen Nachweis der Überabzählbarkeit von $\mathbb{R}$. Wir verwenden wieder das Cantorsche Diagonalverfahren.

Satz 15.7. $\mathbb{R}$ *ist überabzählbar.*

Beweis. Wir zeigen, dass sogar das Intervall $[0,1)$ schon so reichhaltig ist. Dazu nehmen wir an, es gelte $[0,1) = \{a_n : n \in \mathbb{N}\}$ für eine reelle Folge (a_n). Dann können wir jedes a_n durch seine 3-adische Entwicklung darstellen. Demnach ist für jedes $n \in \mathbb{N}$

$$a_n = 0, z_1^{(n)}\, z_2^{(n)}\, z_3^{(n)} \ \ldots \quad \text{mit} \quad z_k^{(n)} \in \{0,1,2\} \quad \text{für alle } k \in \mathbb{N}.$$

Wir definieren nun eine Folge (z_k) durch

$$z_k := \begin{cases} 1, \text{ falls } z_k^{(k)} \in \{0,2\}, \\ 0, \text{ falls } z_k^{(k)} = 1 \end{cases}$$

und setzen $a := \sum_{k=0}^{\infty} z_k/3^k$. Dann ist $a \in [0,1)$, denn es gilt sogar

$$0 \le a = \sum_{k=0}^{\infty} \frac{z_k}{3^k} \le \sum_{k=1}^{\infty} \frac{1}{3^k} = \left(\frac{1}{1-\frac{1}{3}} - 1\right) = \frac{3}{2} - 1 = \frac{1}{2}.$$

Die Folge (z_k) erfüllt mit diesem a die Punkte (a)–(c) aus Satz 15.2, also gibt diese Folge nach Satz 15.2 (d) die 3-adische Entwicklung von a an.
Nach unserer Annahme muss nun also ein Folgenglied a_k mit $a = a_k$ und damit ein Index $k \in \mathbb{N}$ existieren, so dass $z_j^{(k)} = z_j$ für alle $j \in \mathbb{N}$ gilt. Dann ist aber insbesondere $z_k^{(k)} = z_k$ und das ist ein Widerspruch zu unserer Wahl von z_k. $\square$

16. Komplexe Zahlen

Wir wollen nun noch ein letztes Mal unseren Zahlraum erweitern und die komplexen Zahlen einführen. Betrachten wir die bisherigen Zahlraumerweiterungen, so erlauben diese jeweils die Lösung weiterer Gleichungen: in $\mathbb{N}$ kann man die Gleichung $x+3=5$, aber nicht $x+5=3$ lösen. Diese zweite ist aber in $\mathbb{Z}$ lösbar, wo wiederum $3x=5$ nicht lösbar ist. Die Lösung dieser Gleichung findet sich in $\mathbb{Q}$, wo aber $x^3=5$ nicht lösbar ist. So kommt man schließlich nach $\mathbb{R}$.
Die Gleichung, die uns nun in $\mathbb{R}$ Kopfzerbrechen macht, ist $x^2=-1$. Um diese Gleichung zu lösen, müssen wir offensichtlich nicht-reelle Zahlen bemühen, eben die komplexen. Sie werden in den Algebra-Vorlesungen sehen, dass man damit insofern am Ende der Zahlraumerweiterei ist, als jede polynomiale Gleichung in den komplexen Zahlen lösbar ist.

Definition 16.1. *Wir betrachten die Menge*

$$\mathbb{C} := \mathbb{R}^2 = \{(x,y) : x,y \in \mathbb{R}\}$$

der komplexen Zahlen *aller geordneten Paare von reellen Zahlen mit der* Addition

$$(x_1,y_1) \oplus (x_2,y_2) := (x_1+x_2, y_1+y_2) \quad \textit{für alle } (x_1,y_1),\ (x_2,y_2) \in \mathbb{C}$$

und der Multiplikation

$$(x_1,y_1) \odot (x_2,y_2) := (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1) \quad \textit{für alle } (x_1,y_1),\ (x_2,y_2) \in \mathbb{C}.$$

Ist $z=(x,y)\in\mathbb{C}$, so heißt x der Realteil *(Notation $x=\mathrm{Re}(z)$) und y der* Imaginärteil *(Notation $y=\mathrm{Im}(z)$) von z.*

Beispiel 16.2. Wir berechnen einige wichtige Produkte:

$$\begin{aligned}
(0,1)^2 &= (0,1)\odot(0,1) = (0\cdot 0 - 1\cdot 1, 0\cdot 1 + 1\cdot 0) = (-1,0) \\
(x,0)\odot(1,0) &= (x\cdot 1 - 0\cdot 0, x\cdot 0 + 0\cdot 1) = (x,0) \quad \text{für alle } x\in\mathbb{R} \\
(y,0)\odot(0,1) &= (y\cdot 0 - 0\cdot 1, y\cdot 1 + 0\cdot 0) = (0,y) \quad \text{für alle } y\in\mathbb{R}.
\end{aligned}$$

Üblicherweise wird eine andere Schreibweise für die komplexen Zahlen verwendet, die wir auch im Folgenden verwenden wollen. Dazu definiert man die imaginäre Einheit

$$\mathrm{i} := (0,1)$$

und interpretiert die komplexen Zahlen als Erweiterung der reellen, indem man die reelle Zahl x mit der komplexen Zahl $(x, 0)$ identifiziert (deshalb auch die Bezeichung als Realteil). Das ergibt mit den Rechnungen aus obigem Beispiel für alle $x, y \in \mathbb{R}$

$$(x, y) = (x, 0) \oplus (0, y) = (x, 0) \odot (1, 0) \oplus (y, 0) \odot (0, 1) = x \cdot 1 + y \cdot \mathrm{i} = x + \mathrm{i}y.$$

Bemerkung 16.3. (a) Nun können wir auch mühelos die Gleichung $z^2 = -1$ lösen, denn es gilt nach der ersten Berechnung in Beispiel 16.2

$$\mathrm{i}^2 = (0, 1) \odot (0, 1) = (-1, 0) = -1,$$

genauso wie übrigens auch $(-\mathrm{i})^2 = (0, -1) \odot (0, -1) = -1$ gilt.

(b) Der Vorteil davon, wenn man die komplexe Zahl $z = (x, y) \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ als $x + \mathrm{i}y$ schreibt, ist, dass man nun rechnen kann wie gewohnt, solange man immer $\mathrm{i}^2 = -1$ beachtet, denn für alle $x_1, x_2, y_1, y_2 \in \mathbb{R}$ gilt nach den gewohnten Rechenregeln

$$(x_1 + \mathrm{i}y_1) + (x_2 + \mathrm{i}y_2) = x_1 + x_2 + \mathrm{i}(y_1 + y_2),$$

was mit der Definition von i und der Identifikation der reellen Zahlen mit dem Realteil genau der Defintion der Addition oben entspricht.

Für die Multiplikation sieht man

$$\begin{aligned}(x_1 + \mathrm{i}y_1) \cdot (x_2 + \mathrm{i}y_2) &= x_1x_2 + \mathrm{i}x_1y_2 + \mathrm{i}y_1x_2 + \mathrm{i}^2y_1y_2 \\ &= x_1x_2 + \mathrm{i}(x_1y_2 + x_2y_1) - y_1y_2 \\ &= x_1x_2 - y_1y_2 + \mathrm{i}(x_1y_2 + x_2y_1),\end{aligned}$$

was wieder genau zur Definition passt.

(c) Im Folgenden verwenden wir nicht weiter die schwerfälligen Bezeichnungen $\oplus$ und $\odot$ für die Verknüpfungen in $\mathbb{C}$, sondern schreiben auch zwischen komplexen Zahlen wieder $+$, bzw. $\cdot$, da wir ja nach obigen Überlegungen mit den komplexen Zahlen weiter so rechnen können wie gewohnt.

Satz 16.4. *Die Menge $\mathbb{C}$ mit den Verknüpfungen $\oplus$ und $\odot$ ist im algebraischen Sinne ein Körper, d. h. sie erfüllt die Axiome (A1) bis (A9) aus Kapitel 2.*

Der Beweis besteht aus einfachem Nachrechnen und ist eine gute Übung, um sich an das Rechnen mit komplexen Zahlen (in beiden Schreibweisen) zu gewöhnen. Veranschaulichen kann man sich die komplexen Zahlen am besten in der *komplexen Zahlenebene* (auch *Gaußsche Zahlenebene* genannt), s. Abbildung 16.1.

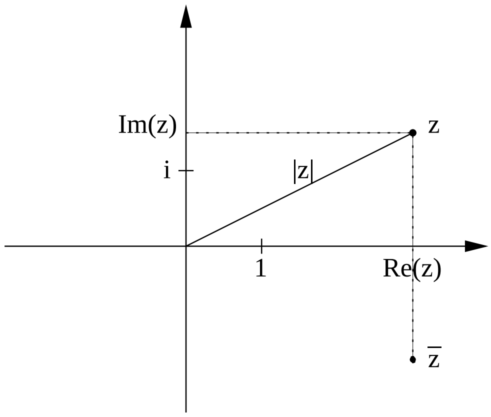

Abbildung 16.1.: Die Gaußsche Zahlenebene

Definition 16.5. *Ist $z = x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ eine komplexe Zahl, so heißt*

$$\overline{z} := x - \mathrm{i}y \qquad \textit{die zu } z \text{ konjugiert komplexe } \textit{Zahl,}$$
$$|z| := \sqrt{x^2 + y^2} \qquad \textit{der } \text{Betrag } \textit{von } z.$$

Offensichtlich gilt $\overline{\overline{z}} = z$ für alle $z \in \mathbb{C}$. Einige weitere bemerkenswerte Eigenschaften der komplexen Konjugation sammeln wir im folgenden Satz.

Satz 16.6. *Sind $z, z_1, z_2 \in \mathbb{C}$, so gilt*

(a) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ *und* $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$.

(b) $z + \overline{z} = 2 \cdot \mathrm{Re}(z)$ *und* $z - \overline{z} = 2\,\mathrm{i} \cdot \mathrm{Im}(z)$.

Beweis. (a) Es ist für $z_1 = x_1 + \mathrm{i}y_1$ und $z_2 = x_2 + \mathrm{i}y_2$ durch einfaches Nachrechnen

$$\begin{aligned}\overline{z_1 + z_2} &= \overline{x_1 + x_2 + \mathrm{i}(y_1 + y_2)} = x_1 + x_2 - \mathrm{i}(y_1 + y_2)\\ &= x_1 - \mathrm{i}y_1 + x_2 - \mathrm{i}y_2 = \overline{z_1} + \overline{z_2}\end{aligned}$$

und

$$\begin{aligned}\overline{z_1 z_2} &= \overline{(x_1x_2 - y_1y_2 + \mathrm{i}(x_1y_2 + x_2y_1)} = x_1x_2 - y_1y_2 - \mathrm{i}(x_1y_2 + x_2y_1)\\ &= x_1x_2 - \mathrm{i}x_1y_2 - \mathrm{i}x_2y_1 - y_1y_2 = (x_1 - \mathrm{i}y_1)(x_2 - \mathrm{i}y_2) = \overline{z_1} \cdot \overline{z_2}.\end{aligned}$$

(b) $z + \overline{z} = \mathrm{Re}(z) + \mathrm{i}\mathrm{Im}(z) + \mathrm{Re}(z) - \mathrm{i}\mathrm{Im}(z) = 2 \cdot \mathrm{Re}(z)$ und
$z - \overline{z} = \mathrm{Re}(z) + \mathrm{i}\mathrm{Im}(z) - \mathrm{Re}(z) + \mathrm{i}\mathrm{Im}(z) = 2\mathrm{i} \cdot \mathrm{Im}(z)$. □

Für den Betrag sehen wir sofort, dass $|z| = |\overline{z}|$ für alle $z \in \mathbb{C}$ ist. Außerdem gelten die folgenden Aussagen.

Satz 16.7. *Für alle komplexen Zahlen* $z = x + \mathrm{i}y$ *mit* $x, y \in \mathbb{R}$ *und* $z_1, z_2 \in \mathbb{C}$ *gilt*

(a) $z \cdot \overline{z} = |z|^2$,

(b) $\mathrm{Re}(z) \leq |z|$ *und* $\mathrm{Im}(z) \leq |z|$,

(c) $|z| \geq 0$ *und* $|z| = 0$ *genau dann, wenn* $z = 0$,

(d) $|z_1 z_2| = |z_1||z_2|$,

(e) $|z_1 + z_2| \leq |z_1| + |z_2|$ (Dreiecksungleichung)

Beweis. (a) $z\overline{z} = (x + \mathrm{i}y)(x - \mathrm{i}y) = x^2 + \mathrm{i}yx - \mathrm{i}yx - \mathrm{i}^2 y^2 = x^2 + y^2 = (\sqrt{x^2 + y^2})^2 = |z|^2$.

(b) Folgt für den Realteil sofort aus $|x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} = |z|$ und genauso für den Imaginärteil.

(c) Ist $|z| = 0$, so ist $\sqrt{x^2 + y^2} = 0$ und damit auch $x^2 + y^2 = 0$. Diese Gleichung (in $\mathbb{R}$!) hat nun nur die Lösung $x = y = 0$, also ist in diesem Fall $z = 0$.

(d) Nachrechnen.

(e) Nach (a) und Satz 16.6 ist

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} = (z_1 + z_2)(\overline{z_1} + \overline{z_2}) \\ &= z_1\overline{z_1} + z_1\overline{z_2} + \overline{z_1}z_2 + z_2\overline{z_2} = |z_1|^2 + |z_2|^2 + z_1\overline{z_2} + \overline{z_1\overline{z_2}} \\ &= |z_1|^2 + |z_2|^2 + 2 \cdot \mathrm{Re}(z_1\overline{z_2}) \end{aligned}$$

Nach (b) und (d) gilt nun $\mathrm{Re}(z_1\overline{z_2}) \leq |z_1\overline{z_2}| = |z_1||z_2|$. Also ist

$$|z_1 + z_2|^2 \leq |z_1|^2 + |z_2|^2 + 2|z_1||z_2| = (|z_1| + |z_2|)^2. \qquad \square$$

Eine natürliche Frage ist nun: Welche Erkenntnisse, die wir bisher in $\mathbb{R}$ gewonnen haben, gelten auch in $\mathbb{C}$ weiter? Da nach Satz 16.4 auch in $\mathbb{C}$ die Körperaxiome $(A1)$ bis $(A9)$ gelten, können wir hoffen, dass das eine ganze Menge ist.
Doch zunächst ein Satz dazu, was *nicht* geht. Es gibt auf $\mathbb{C}$ keine mit der algebraischen (also der durch die Verknüpfungen gegebenen) Struktur verträgliche Ordnung. Man kann also 2 komplexe Zahlen nicht der Größe nach vergleichen. Dementsprechend machen auch alle direkt mit Hilfe der Ordnung definierten Begriffe wie Supremum, Minimum, Monotonie usw. in $\mathbb{C}$ keinen Sinn.
Wohl kann man aber *Beträge* von komplexen Zahlen vergleichen, da man ja dazu nur die Ordnung in $\mathbb{R}$ braucht, schließlich ist der Betrag jeder komplexen Zahl reell. Das werden wir uns oft zu Nutze machen.
Wir wollen im Rest dieses Abschnitts einen Schnelldurchlauf durch die komplexen Folgen und Reihen machen und die Exponentialfunktion in $\mathbb{C}$ definieren. Hier

und im weiteren Verlauf der Vorlesung werden wir feststellen, dass viele Dinge in $\mathbb{C}$ genauso wie in $\mathbb{R}$ funktionieren. Um nicht immer mühsam „in $\mathbb{R}$ oder $\mathbb{C}$“ schreiben zu müssen, wird im weiteren Skript der Buchstabe $\mathbb{K}$ verwendet, wann immer man sowohl $\mathbb{R}$ als auch $\mathbb{C}$ einsetzen kann.
Wir übertragen also zunächst den Begriff der Konvergenz. Diesen können wir auf Konvergenz in $\mathbb{R}$ zurückspielen (vgl. Satz 6.8 (a)).

Definition 16.8. *Es sei (z_n) eine Folge in $\mathbb{C}$ und $z_0 \in \mathbb{C}$. Dann gilt*

(a) (z_n) konvergiert gegen z_0, *wenn die reelle Folge $(|z_n - z_0|)$ in $\mathbb{R}$ gegen Null konvergiert.*

(b) Das Symbol $\sum_{n=1}^{\infty} z_n$ bezeichnet wieder die Folge (s_n) mit $s_n := \sum_{k=1}^{n} z_k$ und die Reihe $\sum_{n=1}^{\infty} z_n$ *heißt* konvergent, *falls (s_n) konvergiert. In diesem Fall gilt $\sum_{n=1}^{\infty} z_n := \lim_{n\to\infty} s_n$.*

Satz 16.9. *Es sei (z_n) eine Folge in $\mathbb{C}$ und $z_0 \in \mathbb{C}$. Dann gilt*

(a) (z_n) konvergiert genau dann gegen z_0, wenn die beiden Folgen $(\mathrm{Re}(z_n))$ und $(\mathrm{Im}(z_n))$ in $\mathbb{R}$ konvergieren und

$$\lim_{n\to\infty} \mathrm{Re}(z_n) = \mathrm{Re}(z_0) \quad \textit{sowie} \quad \lim_{n\to\infty} \mathrm{Im}(z_n) = \mathrm{Im}(z_0)$$

gilt.

Insbesondere ist der Limes einer in $\mathbb{C}$ konvergenten Folge eindeutig bestimmt.

(b) Die Reihe $\sum_{n=1}^{\infty} z_n$ ist genau dann konvergent, wenn die beiden Reihen $\sum_{n=1}^{\infty} \mathrm{Re}(z_n)$ und $\sum_{n=1}^{\infty} \mathrm{Im}(z_n)$ in $\mathbb{R}$ konvergieren. In diesem Fall gilt

$$\sum_{n=1}^{\infty} z_n = \sum_{n=1}^{\infty} \mathrm{Re}(z_n) + \mathrm{i} \sum_{n=1}^{\infty} \mathrm{Im}(z_n).$$

Beweis. (a) Wir setzen $x_n := \mathrm{Re}(z_n)$ und $y_n := \mathrm{Im}(z_n)$ für alle $n \in \mathbb{N}_0$. Nun gilt für jedes $n \in \mathbb{N}$

$$|x_n - x_0| = \sqrt{(x_n - x_0)^2} \leq \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2} = |z_n - z_0|$$

und genauso $|y_n - y_0| \leq |z_n - z_0|$. Wenn also (z_n) gegen z_0 konvergiert, so haben wir auch $\lim_{n\to\infty} x_n = x_0$ und $\lim_{n\to\infty} y_n = y_0$. Haben wir umgekehrt diese beiden Grenzwertbeziehungen, so gilt

$$\begin{aligned} |z_n - z_0| &= \sqrt{\mathrm{Re}(z_n - z_0)^2 + \mathrm{Im}(z_n - z_0)^2} \\ &= \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2} \longrightarrow 0 \quad (n \to \infty), \end{aligned}$$

wegen Satz 7.1.

(b) Folgt aus (a), da Reihen auch nur Folgen sind. □

Auch der Begriff der absoluten Konvergenz überträgt sich wortwörtlich.

Definition 16.10. *Eine Reihe* $\sum_{n=1}^{\infty} z_n$ *mit* $(z_n) \subseteq \mathbb{C}$ *heißt* absolut konvergent, *falls die Reihe* $\sum_{n=1}^{\infty} |z_n|$ *konvergiert.*

Satz 16.11. *Es sei* (z_n) *eine komplexe Folge. Dann gilt*

(a) Ist $\sum_{n=1}^{\infty} z_n$ *absolut konvergent, so ist die Reihe* $\sum_{n=1}^{\infty} z_n$ *konvergent und es gilt*

$$\Big|\sum_{n=1}^{\infty} z_n\Big| \leq \sum_{n=1}^{\infty} |z_n| \quad \text{(Dreiecksungleichung)}$$

(b) Die Reihe $\sum_{n=1}^{\infty} z_n$ *konvergiert genau dann, wenn für jedes* $\varepsilon > 0$ *ein* $n_0 \in \mathbb{N}$ *existiert mit*

$$\Big|\sum_{k=n+1}^{m} z_k\Big| = |z_{n+1} + z_{n+2} + \cdots + z_m| < \varepsilon \quad \textit{für alle } m > n \geq n_0.$$

(Cauchy-Kriterium)

Der Beweis dieses Satzes überträgt sich direkt aus dem Reellen. Gleiches gilt für den folgenden Satz über das wie in $\mathbb{R}$ definierte Cauchy-Produkt.

Definition 16.12. *Es seien* $\sum_{n=0}^{\infty} z_n$ *und* $\sum_{n=0}^{\infty} w_n$ *zwei komplexe Reihen. Dann heißt die Reihe*

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} z_k w_{n-k}$$

das Cauchy-Produkt *dieser beiden Reihen.*

Satz 16.13. *Es seien* $\sum_{n=0}^{\infty} z_n$ *und* $\sum_{n=1}^{\infty} w_n$ *absolut konvergente Reihen in* $\mathbb{C}$. *Dann ist auch das Cauchy-Produkt der beiden Reihen absolut konvergent und es gilt*

$$\Big(\sum_{n=0}^{\infty} z_n\Big)\Big(\sum_{n=0}^{\infty} w_n\Big) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} z_k w_{n-k}.$$

Da sich also herausstellt, dass unsere Ergebnisse über Reihen in großen Teilen auch im Komplexen gelten, ist es nun natürlich, auch über Potenzreihen mit komplexen Variablen und Koeffizienten nachzudenken. Das wird sich als eine sehr fruchtbare Idee (auch für zukünftige Semester) erweisen. So können wir z. B. nun auch problemlos die Exponentialfunktion für komplexe Zahlen definieren. Tatsächlich kann man leicht feststellen, dass der Satz von Hadamard 14.2 und auch der Satz 14.8 über das Cauchy-Produkt von Potenzreihen genauso in $\mathbb{C}$ funktionieren.

Beispiel 16.14. (a) Wir betrachten zunächst die komplexe geometrische Reihe, also die Potenzreihe

$$\sum_{n=0}^{\infty} z^n, \quad z \in \mathbb{C}.$$

Für $z = 1$ ist diese Reihe bekanntermaßen nicht konvergent und für alle $z \in \mathbb{C} \setminus \{1\}$ gilt genau wie im reellen

$$\sum_{n=0}^{N} z^n = \frac{1 - z^{N+1}}{1 - z}$$

und dieser Ausdruck konvergiert für $N \to \infty$, falls $|z| < 1$. In diesem Fall ist

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}, \qquad |z| < 1.$$

Die komplexen Zahlen mit $|z| < 1$ sind genau die, die innerhalb des Kreises mit Radius 1 um den Ursprung in der komplexen Zahlenebene liegen. Nun wird auch der bisher u.U. eher verwirrende Begriff Konvergenz*radius* klar.

(b) Der Kandidat für die Definition einer komplexen Exponentialfunktion ist natürlich

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}, \qquad z \in \mathbb{C}.$$

Tatsächlich wissen wir schon, dass die Reihe $\sum_{n=0}^{\infty} |z|^n/n!$ für alle $z \in \mathbb{C}$ konvergiert, denn $|z|$ ist ja reell. Also ist die Reihe $\sum_{n=0}^{\infty} z^n/n!$ für alle $z \in \mathbb{C}$ sogar absolut konvergent. Das rechtfertigt die nun folgende Definition.

Definition 16.15. *Die Funktion*

$$\mathrm{e}^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}, \qquad z \in \mathbb{C},$$

heißt (komplexe) Exponentialfunktion.

Wir sammeln einige Eigenschaften dieser Funktion.

Satz 16.16. *(a) Für alle $z_1, z_2 \in \mathbb{C}$ gilt $\mathrm{e}^{z_1+z_2} = \mathrm{e}^{z_1} \cdot \mathrm{e}^{z_2}$.*

(b) Für alle $z \in \mathbb{C}$ gilt $\mathrm{e}^z \neq 0$ und $1/\mathrm{e}^z = \mathrm{e}^{-z}$.

(c) Für alle $t \in \mathbb{R}$ gilt $\mathrm{e}^{\mathrm{i}t} = \cos(t) + \mathrm{i}\sin(t)$.

(d) Ist $z = x + \mathrm{i}y \in \mathbb{C}$ mit $x, y \in \mathbb{R}$, so gilt $\mathrm{e}^z = \mathrm{e}^x \cos(y) + \mathrm{i}\mathrm{e}^x \sin(y)$.

(e) *Für alle $t \in \mathbb{R}$ und alle $n \in \mathbb{N}$ gilt die* Formel von De Moivre*:*

$$\cos(nt) + \mathrm{i}\sin(nt) = \big(\cos(t) + \mathrm{i}\sin(t)\big)^n.$$

Beweis. (a) Genau wie in $\mathbb{R}$ berechnet man dieses über das Cauchy-Produkt.

(b) Aus (a) folgt sofort für jedes $z \in \mathbb{C}$

$$1 = \mathrm{e}^0 = \mathrm{e}^{z-z} = \mathrm{e}^z\mathrm{e}^{-z},$$

und damit $\mathrm{e}^z \neq 0$ sowie $\mathrm{e}^{-z} = \frac{1}{\mathrm{e}^z}$.

(c) Für alle $t \in \mathbb{R}$ können wir wegen der absoluten Konvergenz der Exponentialreihe die Summation umordnen und zunächst über alle geraden n und dann über alle ungeraden n summieren:

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}t} &= \sum_{n=0}^{\infty} \frac{\mathrm{i}^n t^n}{n!} = \sum_{k=0}^{\infty} \frac{\mathrm{i}^{2k} t^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{\mathrm{i}^{2k+1} t^{2k+1}}{(2k+1)!} \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} + \mathrm{i}\sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!} = \cos(t) + \mathrm{i}\sin(t).
\end{aligned}$$

(d) Es ist mit (a) und (c)

$$\mathrm{e}^z = \mathrm{e}^{x+\mathrm{i}y} = \mathrm{e}^x\mathrm{e}^{\mathrm{i}y} = \mathrm{e}^x\cos y + \mathrm{i}\mathrm{e}^x\sin(y).$$

(e) Auch dieses folgt sofort aus (c) und (a), denn

$$\cos(nt) + \mathrm{i}\sin(nt) = \mathrm{e}^{\mathrm{i}nt} = \big(\mathrm{e}^{\mathrm{i}t}\big)^n = \big(\cos(t) + \mathrm{i}\sin(t)\big)^n. \qquad \square$$

Teil III.

Funktionen

17. Grenzwerte bei Funktionen

In diesem Abschnitt übertragen wir Begriffe, die wir im Kontext der Folgen kennen gelernt haben, auf Funktionen von $D \subseteq \mathbb{R}$ nach $\mathbb{R}$, bzw. $\mathbb{C}$. In diesem Zusammenhang sei an die Konvention zum Buchstaben $\mathbb{K}$ von Seite 97 erinnert. Wir beginnen mit dem Monotoniebegriff.

Definition 17.1. *Es sei $I \subseteq \mathbb{R}$ ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt*

(a) monoton wachsend, *falls für alle $x, y \in I$ mit $x \leq y$ gilt, dass $f(x) \leq f(y)$ ist.*

(b) monoton fallend, *falls für alle $x, y \in I$ mit $x \leq y$ gilt, dass $f(x) \geq f(y)$ ist.*

(c) streng monoton wachsend, *falls für alle $x, y \in I$ mit $x < y$ gilt, dass $f(x) < f(y)$ ist.*

(d) streng monoton fallend, *falls für alle $x, y \in I$ mit $x < y$ gilt, dass $f(x) > f(y)$ ist.*

(e) (streng) monoton, *falls sie (streng) monoton wachsend oder (streng) monoton fallend ist.*

Definition 17.2. *Es sei $D \subseteq \mathbb{R}$ eine Menge. Eine Zahl $x_0 \in \mathbb{R}$ heißt* Häufungspunkt *von D genau dann, wenn für jedes $\varepsilon > 0$ ein $x \in D$ existiert mit*

$$x \in U_\varepsilon(x_0) \setminus \{x_0\}.$$

Um ein bisschen mehr Vorstellung von diesem Begriff zu kommen, betrachten wir zwei äquivalente Formulierungen dieser Definition.

Satz 17.3. *Es sei $D \subseteq \mathbb{R}$ und $x_0 \in \mathbb{R}$. Dann gilt*

(a) x_0 ist ein Häufungspunkt von D genau dann, wenn für alle $\varepsilon > 0$ die Menge $D \cap U_\varepsilon(x_0)$ unendlich ist.

(b) x_0 ist ein Häufungspunkt von D genau dann, wenn eine Folge (x_n) in D existiert mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$, die gegen x_0 konvergiert.

Beweis. (a) Übungsaufgabe

(b) Ist x_0 ein Häufungspunkt von D, so existiert nach der Definition des Häufungspunktes für jedes $n \in \mathbb{N}$ ein $x_n \in D \cap U_{1/n}(x_0) \setminus \{x_0\}$, d. h. für jedes $n \in \mathbb{N}$ sind die Zahlen x_n in D, es gilt $x_n \neq x_0$ und wir haben $|x_n - x_0| \leq 1/n$. Also tut die Folge (x_n) das gewünschte.

Es sei nun für die andere Beweisrichtung eine Folge (x_n) in D gegeben mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$, die gegen x_0 konvergiert. Dann gibt es für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass $|x_n - x_0| < \varepsilon$ für alle $n \geq n_0$ gilt. Also ist für jedes $\varepsilon > 0$ die Zahl $x_{n_0} \in D \cap U_\varepsilon(x_0) \setminus \{x_0\}$ und x_0 ist ein Häufungspunkt von D. □

Mit Hilfe dieser Umformulierungen kann man sich leicht die folgenden Beispiele veranschaulichen.

Beispiel 17.4. (a) Ist D endlich, so hat D keinen Häufungspunkt.

(b) Ist $D = (0, 1]$, so ist die Menge aller Häufungspunkte von D genau das Intervall $[0, 1]$.

(c) Ist $D = \{1/n : n \in \mathbb{N}\}$, so ist genau 0 ein Häufungspunkt von D.

Die Definition des Grenzwertes einer Funktion an einer Stelle spielen wir auf den Konvergenzbegriff bei Folgen zurück.

Definition 17.5. *Es sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{K}$ eine Funktion sowie $a \in \mathbb{K}$.*

(a) Ist x_0 ein Häufungspunkt von D, so sagen wir, dass f für x gegen x_0 den Grenzwert *a hat, wenn für jede Folge (x_n) in D, die gegen x_0 konvergiert und für die $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, die Folge $(f(x_n))$ gegen a konvergiert. Wir schreiben dafür*

$$\lim_{x \to x_0} f(x) = a.$$

(b) Ist x_0 ein Häufungspunkt von $D_+ := \{x \in D : x > x_0\}$, so hat f für x gegen x_0 den rechtsseitigen Grenzwert *a, wenn für jede Folge (x_n) in D_+, die gegen x_0 konvergiert, die Folge $(f(x_n))$ gegen a konvergiert. Wir schreiben*

$$\lim_{x \to x_0+} f(x) = a.$$

(c) Ist x_0 ein Häufungspunkt von $D_- := \{x \in D : x < x_0\}$, so hat f für x gegen x_0 den linksseitigen Grenzwert *a, wenn für jede Folge (x_n) in D_-, die gegen x_0 konvergiert, die Folge $(f(x_n))$ gegen a konvergiert. Wir schreiben*

$$\lim_{x \to x_0-} f(x) = a.$$

Bemerkung 17.6. (a) Eine besondere Betonung beim Lesen dieser Definition sollte jeweils auf den Worten „*jede* Folge“ liegen.

(b) Wie bei den Grenzwerten für Folgen gibt es auch hier die alternativen Schreibweisen

$$f(x) \to a \ (x \to x_0), \quad f(x) \to a \ (x \to x_0+) \text{ bzw. } f(x) \to a \ (x \to x_0-).$$

(c) In den folgenden Sätzen und Definitionen werden wir alle Aussagen jeweils nur für Grenzwerte von Funktionen machen. Diese gelten dann sinngemäß auch für den rechtsseitigen und linksseitigen Grenzwert, wenn dieser sinnvoll ist.

Die Subtilitäten dieser Definition verdeutlichen wir uns an einem Beispiel.

Beispiel 17.7. (a) Wir setzen $D = (0, 1]$ und

$$f(x) = \begin{cases} x^2, & 0 < x < \frac{1}{2}, \\ 1, \quad \text{falls} & \frac{1}{2} \leq x < 1, \\ 7, & x = 1. \end{cases}$$

Wie wir schon im vorherigen Beispiel gesehen haben, ist jedes $x \in [0, 1]$ ein Häufungspunkt von D. Wir können also Grenzwertbetrachtungen in allen diesen Punkten anstellen. Wir untersuchen das Verhalten in den interessanten Stellen 0, 1/2 und 1. Es gilt

$$\lim_{x \to 0} f(x) = 0 \quad \text{und} \quad \lim_{x \to 1} f(x) = 1.$$

Wir begründen die erste Aussage. Es sei (x_n) eine beliebige Nullfolge in D ($x_n \neq 0$ gilt sowieso, da $0 \notin D$). Dann gibt es ein $n_0 \in \mathbb{N}$, so dass $x_n < 1/2$ für alle $n \geq n_0$ gilt. Für diese n ist dann $f(x_n) = x_n^2$. Nach den Grenzwertsätzen für Folgen konvergiert die Folge (x_n^2) gegen Null, womit $\lim_{x \to 0} f(x) = 0$ gezeigt ist. Den zweiten Limes bestimmt man analog.

Was ist aber nun mit $\lim_{x \to 1/2} f(x)$? Dieser Grenzwert existiert nicht, denn es gilt

$$\lim_{n \to \infty} f\Big(\frac{1}{2} - \frac{1}{n}\Big) = \frac{1}{4} \quad \text{und} \quad \lim_{n \to \infty} f\Big(\frac{1}{2} + \frac{1}{n}\Big) = 1.$$

Nach unserer Definition müsste der Grenzwert aber für jede Folge, die gegen 1/2 konvergiert, ohne 1/2 zu treffen, der selbe sein.

Das Problem tritt hier auf, weil wir uns einmal von links und einmal von rechts an die kritische Stelle 1/2 angenähert haben. Genau dafür haben wir aber die Begriffe des links- bzw. rechtsseitigen Grenzwertes. Tatsächlich zeigt man wie oben, dass diese existieren und dass gilt

$$\lim_{x\to 1/2+} f(x) = 1 \quad \text{und} \quad \lim_{x\to 1/2-} f(x) = \frac{1}{4}.$$

(b) Ohne es so zu nennen, haben wir schon einmal solche Grenzwerte berechnet. Satz 7.1 können wir nun folgendermaßen formulieren:

$$\lim_{x\to x_0} \sqrt[p]{x} = \sqrt[p]{x_0} \quad \text{für alle } p \in \mathbb{N} \text{ und alle } x_0 \in [0,\infty).$$

Wir beweisen ein alternatives Kriterium, wann $\lim_{x\to x_0} f(x) = a$ gilt.

Satz 17.8. *Es sei $D \subseteq \mathbb{R}$ und x_0 ein Häufungspunkt von D sowie $f : D \to \mathbb{K}$ eine Funktion und $a \in \mathbb{K}$. Dann gilt $\lim_{x\to x_0} f(x) = a$ genau dann, wenn für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert, so dass*

$$|f(x) - a| < \varepsilon \quad \text{für alle } x \in D \text{ mit } 0 < |x - x_0| < \delta \tag{17.1}$$

gilt.

Beweis. Wir beweisen zunächst „$\Longrightarrow$“. Dazu nehmen wir an, die ε-Bedingung sei falsch, d. h. es gibt ein $\varepsilon_0 > 0$, so dass (17.1) für kein $\delta > 0$ gilt. Für dieses ε_0 gibt es also zu jedem $\delta > 0$ ein $x = x(\delta) \in D$, für das $0 < |x - x_0| < \delta$, aber $|f(x) - a| \geq \varepsilon_0$ gilt. Wählen wir für jedes $n \in \mathbb{N}$ speziell $\delta = 1/n$, so erhalten wir eine Folge (x_n) in D mit

$$0 < |x_n - x_0| < 1/n \quad \text{aber} \quad |f(x_n) - a| \geq \varepsilon_0 \quad \text{für alle } n \in \mathbb{N}.$$

Das bedeutet, dass die Folge (x_n) gegen x_0 konvergiert und $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, aber $f(x_n)$ nicht gegen a konvergiert. Das ist ein Widerspruch zu $\lim_{x\to x_0} f(x) = a$.
Zum Beweis von „$\Longleftarrow$“ sei (x_n) eine Folge in D, die gegen x_0 konvergiert, mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$. Wir müssen nun zeigen, dass die Folge $(f(x_n))$ konvergiert und den Limes a hat. Sei also $\varepsilon > 0$ gegeben. Nach Voraussetzung existiert dann ein $\delta > 0$, so dass (17.1) gilt. Weiter ergibt sich aus der Konvergenz der Folge (x_n) die Existenz eines $n_0 \in \mathbb{N}$, so dass $0 < |x_n - x_0| < \delta$ für alle $n \geq n_0$ gilt. Damit haben wir aber nach (17.1) für alle $n \geq n_0$

$$|f(x_n) - a| < \varepsilon$$

und sind damit fertig. □

Satz 17.9. *Es sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{K}$ eine Funktion und x_0 ein Häufungspunkt von D. Weiter sei für jede Folge (x_n) in D, die gegen x_0 konvergiert und für die $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, die Folge $(f(x_n))$ konvergent. Dann existiert $\lim_{x\to x_0} f(x)$.*

Beweis. Es seien (x_n) und (y_n) Folgen in D, welche die Voraussetzungen des Satzes erfüllen. Wir müssen nun zeigen, dass

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

gilt. Dazu definieren wir eine Folge (z_n) durch $z_{2n-1} = x_n$ und $z_{2n} = y_n$ für alle $n \in \mathbb{N}$, d. h. es ist

$$(z_1, z_2, z_3, \dots) = (x_1, y_1, x_2, y_2, x_3, y_3, \dots).$$

Dann gilt auch bei dieser Folge $z_n \in D$ und $z_n \neq x_0$ für alle $n \in \mathbb{N}$ und z_n konvergiert gegen x_0. Also existiert nach Voraussetzung der Limes $a := \lim_{n\to\infty} f(z_n)$. Die Folgen $(f(x_n))$ und $(f(y_n))$ sind aber Teilfolgen von $(f(z_n))$, also konvergieren diese nach Satz 9.6 (b) gegen den selben Grenzwert, d. h.

$$\lim_{n\to\infty} f(x_n) = a = \lim_{n\to\infty} f(y_n). \qquad \square$$

Auch für Grenzwerte von Funktionen können wir Grenzwertsätze formulieren.

Satz 17.10. *Es sei $D \subseteq \mathbb{R}$ und x_0 ein Häufungspunkt von D. Desweiteren seien drei Funktionen $f, g, h : D \to \mathbb{K}$ gegeben, so dass die Grenzwerte*

$$a = \lim_{x\to x_0} f(x) \quad \textit{und} \quad b = \lim_{x\to x_0} g(x)$$

existieren. Dann gilt:

(a) Die Grenzwerte für $x \to x_0$ von $f + g$, fg und $|f|$ existieren und es gilt

$$\lim_{x\to x_0}\big(f(x) + g(x)\big) = a + b, \quad \lim_{x\to x_0}\big(f(x)g(x)\big) = ab, \quad \lim_{x\to x_0} |f(x)| = |a|.$$

(b) Ist $\mathbb{K} = \mathbb{R}$ und gibt es ein $\delta > 0$, so dass für alle $x \in (D \cap U_\delta(x_0)) \setminus \{x_0\}$ die Ungleichung $f(x) \leq g(x)$ gilt, so folgt $a \leq b$.

(c) Gilt $\mathbb{K} = \mathbb{R}$ sowie $a = b$ und gibt es ein $\delta > 0$ mit

$$f(x) \leq h(x) \leq g(x) \quad \textit{für alle } x \in (D \cap U_\delta(x_0)) \setminus \{x_0\},$$

so existiert auch der Limes von h für $x \to x_0$ und es gilt $\lim_{x\to x_0} h(x) = a$.

(d) Ist $b \neq 0$, so existiert ein $\delta > 0$, so dass $|g(x)| \geq |b|/2$ für alle $x \in (D \cap U_\delta(x_0)) \setminus \{x_0\}$ ist. Wir können also die Funktion

$$\frac{f}{g} : \big(D \cap U_\delta(x_0)\big) \setminus \{x_0\} \to \mathbb{K} \quad \textit{mit} \quad \frac{f}{g}(x) := \frac{f(x)}{g(x)}$$

definieren. Für diese existiert dann der Limes für x gegen x_0 mit

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)} = \frac{a}{b}.$$

Beweis. Der Beweis dieses Satzes ergibt sich direkt aus den Grenzwertsätzen für Folgen in Satz 6.8. Nur die erste Aussage in (d) wollen wir kurz beweisen.
Es sei $\varepsilon := |b|/2 > 0$. Da $\lim_{x\to x_0} g(x) = b$ gilt, existiert dann nach Satz 17.8 ein $\delta > 0$ mit $|g(x) - b| < \varepsilon = |b|/2$ für alle $x \in (D \cap U_\delta(x_0)) \setminus \{x_0\}$. Also gilt für alle diese x mit der Dreiecksungleichung

$$|b| = |b - g(x) + g(x)| \leq |b - g(x)| + |g(x)| \leq \frac{|b|}{2} + |g(x)|,$$

d. h. $|g(x)| \geq |b|/2$. □

Wir wollen im Folgenden auch untersuchen, wie sich Funktionen „im Unendlichen", also für sehr große (bzw. kleine) x verhalten. Dazu würden wir gerne „$\lim_{x\to\infty}$" definieren. Das ist aber mit dem derzeitigen Instrumentarium nicht möglich, da es keine Folgen gibt, die „gegen unendlich konvergieren" (eine solche Formulierung an sich sorgt schon für kalte Schauer auf dem Rücken…). Wir führen deshalb die folgenden Begriffe ein.

Definition 17.11. *Es sei (x_n) eine reelle Folge. Wir sagen*

(a) (x_n) divergiert bestimmt gegen ∞ *(in Zeichen $x_n \to \infty$ $(n \to \infty)$), falls für jedes $C > 0$ ein $n_0 = n_0(C) \in \mathbb{N}$ existiert, so dass $x_n \geq C$ für alle $n \geq n_0$ gilt.*

(b) (x_n) divergiert bestimmt gegen $-\infty$ *(in Zeichen $x_n \to -\infty$ $(n \to \infty)$), falls für jedes $C > 0$ ein $n_0 = n_0(C) \in \mathbb{N}$ existiert, so dass $x_n \leq -C$ für alle $n \geq n_0$ gilt.*

Übungsaufgabe 17.12. (a) Ist (x_n) eine bestimmt divergente Folge gegen plus oder minus unendlich, so gilt $x_n \neq 0$ für fast alle $n \in \mathbb{N}$. Weiter ist die Folge $(1/x_n)_{n\geq k}$ für ein geeigenetes $k \in \mathbb{N}$ eine Nullfolge.

(b) Achtung: ist (y_n) eine Nullfolge mit $y_n \neq 0$ für alle $n \in \mathbb{N}$, so ist $(1/y_n)$ im Allgemeinen nicht bestimmt divergent. Geben Sie hierzu ein Beispiel an.

(c) Beweisen Sie, dass für jede Nullfolge (y_n) mit $y_n \neq 0$ für alle $n \in \mathbb{N}$ die Folge $(1/|y_n|)$ bestimmt gegen ∞ divergiert.

Wir können nun auch unsere Grenzwertdefinition für Funktionen entsprechend ausweiten.

Definition 17.13. *Es sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und x_0 ein Häufungspunkt von D. Wir schreiben*

$$\lim_{x\to x_0} f(x) = \infty \quad \big(\textit{bzw. } \lim_{x\to x_0} f(x) = -\infty\big),$$

wenn für jede Folge (x_n) in D, die gegen x_0 konvergiert und für die $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, die Folge $(f(x_n))$ bestimmt gegen ∞ (bzw. $-\infty$) divergiert.

Definition 17.14. *Es sei $D \subseteq \mathbb{R}$ nicht nach oben (bzw. unten) beschränkt, $f : D \to \mathbb{K}$ eine Funktion und $a \in \mathbb{R} \cup \{\infty, -\infty\}$, bzw. $a \in \mathbb{C}$. Wir sagen*

$$\lim_{x\to\infty} f(x) = a \quad \big(\textit{bzw.} \lim_{x\to-\infty} f(x) = a\big),$$

wenn für jede Folge (x_n) in D, die bestimmt gegen ∞ (bzw. $-\infty$) divergiert, $f(x_n) \to a$ $(n \to \infty)$ gilt.

Beispiel 17.15. (a) Es gilt

$$\lim_{x\to\infty} \frac{1}{x} = 0, \qquad \lim_{x\to 0+} \frac{1}{x} = \infty, \qquad \lim_{x\to 0-} \frac{1}{x} = -\infty.$$

(b) Wir untersuchen abermals die Exponentialfunktion und interessieren uns für ihr Verhalten für sehr große und sehr kleine $x \in \mathbb{R}$. Sei zunächst $x > 0$ und $p \in \mathbb{N}$ beliebig. Es ist

$$E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{p+1}}{(p+1)!} + \ldots.$$

Da $x > 0$ ist, können wir

$$E(x) \geq \frac{x^{p+1}}{(p+1)!}$$

abschätzen und erhalten

$$\frac{E(x)}{x^p} \geq \frac{x}{(p+1)!}, \quad \text{d. h.} \quad \lim_{x\to\infty} \frac{E(x)}{x^p} = \infty.$$

Das bedeutet, dass die Exponentialfunktion schneller wächst als jede Potenzfunktion. Insbesondere gilt also

$$\lim_{x\to\infty} E(x) = \infty.$$

Das Verhalten für sehr kleine x können wir nun mit Hilfe von $E(x) = E(-x)^{-1}$ bestimmen. Ist (x_n) eine bestimmt gegen $-\infty$ divergente Folge, so divergiert die Folge $(-x_n)$ bestimmt gegen ∞. Da für jedes $n \in \mathbb{N}$

$$E(x_n) = \frac{1}{E(-x_n)}$$

gilt, und die Folge $(E(-x_n))$ nach unseren obigen Überlegungen bestimmt nach ∞ divergiert, folgt mit Hilfe der Übungsaufgabe 17.12 sofort

$$\lim_{x\to-\infty} E(x) = 0.$$

18. Stetigkeit

Wir kommen nun zu einem fundamentalen Begriff der Analysis, der Stetigkeit. Flapsig gesprochen, bedeutet Stetigkeit, dass ein kleines Wackeln an der Variablen den Funktionswert einer Funktion auch nur wenig verändert, d. h. kleine Störungen haben auch nur kleine Wirkungen. Damit ist Stetigkeit, zumeist unbemerkt, eine häufige Grundannahme unseres Lebens.
Streng mathematisch formuliert liest sich das so:

Definition 18.1. *Es sei $D \subseteq \mathbb{K}$ und $x_0 \in D$. Eine Funkion $f : D \to \mathbb{K}$ heißt* stetig in x_0*, wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, x_0) > 0$ gibt, so dass*

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

gilt.
Weiter heißt f stetig auf D*, wenn f in jedem Punkt $x_0 \in D$ stetig ist. Wir setzen*

$$C(D) := \{f : D \to \mathbb{K} \mid f \textit{ stetig auf } D\}.$$

Wir können stetige Funktionen auch äquivalent mit Hilfe von Folgen charakterisieren:

Satz 18.2. *Es sei $D \subseteq \mathbb{K}$ und $f : D \to \mathbb{K}$ eine Funktion. Dann ist f genau dann in $x_0 \in D$ stetig, wenn für jede Folge (x_n) in D, die gegen x_0 konvergiert, auch die Folge $(f(x_n))$ konvergiert und $\lim_{n\to\infty} f(x_n) = f(x_0)$ gilt* (Folgenstetigkeit)*.*

Beweis. Sei zunächst f in x_0 stetig, (x_n) eine Folge in D, die gegen x_0 konvergiert und $\varepsilon > 0$. Dann existiert nach der Definition der Stetigkeit ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Da (x_n) gegen x_0 konvergiert, gibt es nun weiter ein $n_0 \in \mathbb{N}$ mit $|x_n - x_0| < \delta$ für alle $n \geq n_0$. Also gilt für all diese n auch $|f(x_n) - f(x_0)| < \varepsilon$ und wir haben Konvergenz der Folge $(f(x_n))$ gegen $f(x_0)$ gezeigt.
Es bleibt die umgekehrte Implikation zu zeigen. Dazu nehmen wir an, f wäre nicht stetig in x_0. Das bedeutet (vgl. Abschnitt 1.2): es gibt ein $\varepsilon_0 > 0$, so dass es für jedes $\delta > 0$ ein $x = x(\delta) \in D$ gibt mit $|x - x_0| < \delta$, aber $|f(x) - f(x_0)| \geq \varepsilon_0$. Insbesondere gilt das für alle δ der Form $1/n$ mit $n \in \mathbb{N}$. Also gibt es für jedes $n \in \mathbb{N}$ ein $x_n \in D$ mit $|x_n - x_0| < 1/n$ und $|f(x_n) - f(x_0)| \geq \varepsilon_0$. Wir betrachten nun die Folge (x_n). Wegen $|x_n - x_0| < 1/n$ für alle $n \in \mathbb{N}$ konvergiert diese Folge gegen x_0. Andererseits bleibt die Folge $(f(x_n))$ aber immer mindestens ε_0 von $f(x_0)$ entfernt, im Widerspruch zur Voraussetzung, nach der $(f(x_n))$ gegen $f(x_0)$ konvergieren muss. $\square$

Diskussionsanregung: Auf der Menge $D = [0,1] \cup \{2\}$ sei die Funktion

$$f(x) = \begin{cases} x^2, \text{ falls } x \in [0,1], \\ 43, \text{ falls } x = 2, \end{cases}$$

gegeben. Skizzieren Sie den Graphen und diskutieren Sie, ob diese Funktion auf D stetig ist.
Der folgende Satz erlaubt wieder, wie schon Satz 6.8 für Folgen, Stetigkeitsuntersuchungen komplizierter Funktionen auf die Untersuchung einfacherer Bausteine zu reduzieren. Sein Beweis ergibt sich aus der Kombination von Satz 18.2 und den entsprechenden Aussagen aus Satz 6.8.

Satz 18.3. *Es sei $D \subseteq \mathbb{K}$ und $f, g : D \to \mathbb{K}$ seien stetig in $x_0 \in D$. Dann sind die Funktionen $f + g$, fg und $|f|$ stetig in x_0.*
Ist $x_0 \in \tilde{D} := \{x \in D \mid g(x) \neq 0\}$, so ist die Funktion $f/g : \tilde{D} \to \mathbb{K}$ stetig in x_0.

Übungsaufgabe 18.4. Es sei $D \subseteq \mathbb{K}$ und $f : D \to \mathbb{K}$ eine Funktion sowie $x_0 \in D$.

(a) Ist zusätzlich x_0 ein Häufungspunkt von D, so ist f in x_0 genau dann stetig, wenn $\lim_{x \to x_0} f(x) = f(x_0)$ gilt.

(b) Beweisen Sie Satz 18.3.

(c) Zeigen Sie: Ist $\mathbb{K} = \mathbb{C}$ und f stetig in x_0, so sind auch $\overline{f}$, $\mathrm{Re} f$ und $\mathrm{Im} f$ stetig in x_0.

Für Funktionen gibt es zusätzlich zu Addition und Multiplikation auch noch die Hintereinanderausführung als Verknüpfung. Tatsächlich erhält auch diese die Stetigkeit.

Satz 18.5. *Es sei $D \subseteq \mathbb{K}$ und die Funktion $f : D \to \mathbb{R}$ sei stetig in $x_0 \in D$. Weiter seien $E \subseteq \mathbb{K}$ mit $f(D) \subseteq E$ und eine weitere Funktion $g : E \to \mathbb{K}$ gegeben, die in $f(x_0)$ stetig ist. Dann ist die Funktion $g \circ f : D \to \mathbb{K}$ stetig in x_0.*

Beweis. Sei (x_n) eine Folge in D mit $x_n \to x_0$ $(n \to \infty)$. Dann folgt aus der Stetigkeit von f mit Satz 18.2 sofort $f(x_n) \to f(x_0)$ $(n \to \infty)$. Der selbe Satz zusammen mit der Stetigkeit von g in $f(x_0)$ liefert uns dann $g(f(x_n)) \to g(f(x_0))$ $(n \to \infty)$. Also ist wiederum nach Satz 18.2 die Funktion $g \circ f$ stetig in x_0. $\square$

Wir wollen nun die Stetigkeit einer ganzen Klasse von Funktionen auf einmal zeigen, nämlich all derer, die durch eine Potenzreihe mit positivem Konvergenzradius gegeben sind.

Satz 18.6. *Es sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $r > 0$ (dabei ist $r = \infty$ zugelassen). Setzen wir $D := \{x \in \mathbb{K} : |x| < r\}$, bzw. $D := \mathbb{K}$, falls $r = \infty$, so ist die Funktion $f : D \to \mathbb{K}$, gegeben durch $f(x) = \sum_{n=0}^{\infty} a_n x^n$ für $x \in D$, stetig auf D.*

Beweis. Es sei $x_0 \in D$ beliebig und $\varrho \in \mathbb{R}$ sei so gewählt, dass $|x_0| < \varrho < r$ gilt. Ist $x \in \mathbb{K}$ mit $|x| \leq \varrho$, so ist $f(x) - f(x_0) = \sum_{n=1}^{\infty} a_n(x^n - x_0^n)$ und für jedes $n \in \mathbb{N}$ ist (vgl. Satz 5.2 (c))

$$\begin{aligned}\left|a_n(x^n - x_0^n)\right| &= \left|a_n(x - x_0)(x^{n-1} + x^{n-1}x_0 + \cdots + xx_0^{n-2} + x_0^{n-1})\right| \\ &\leq |a_n||x - x_0|(|x|^{n-1} + |x|^{n-2}|x_0| + \cdots + |x_0|^{n-1}) \\ &\leq |a_n||x - x_0|n\varrho^{n-1},\end{aligned}$$

da sowohl $|x| \leq \varrho$ als auch $|x_0| \leq \varrho$ gilt.
Wir zeigen als nächstes, dass die Reihe $\sum_{n=1}^{\infty} |a_n|n\varrho^{n-1}$ konvergiert. Zur Anwendung des Wurzelkriteriums berechnen wir

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|n\varrho^{n-1}} = \limsup_{n\to\infty} \frac{\sqrt[n]{|a_n|}\,\sqrt[n]{n}\varrho}{\sqrt[n]{\varrho}} = \varrho \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

Dabei haben wir im letzten Schritt Übungsaufgabe 8.6 angewendet und dabei verwendet, dass $\lim_{n\to\infty} \sqrt[n]{n} = \lim_{n\to\infty} \sqrt[n]{\varrho} = 1$ gilt.
Nach dem Satz von Hadamard wissen wir im Fall $r = \infty$ nun $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 0$ und ist $r < \infty$, so gilt $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 1/r$, also ist

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|n\varrho^{n-1}} \leq \varrho/r < 1,$$

die Reihe ist also nach dem Wurzelkriterium konvergent. Wir setzen

$$s := \sum_{n=1}^{\infty} |a_n|n\varrho^{n-1}.$$

Damit ist die Reihe über $a_n(x^n - x_0^n)$ absolut konvergent und wir können mit der verallgemeinerten Dreiecksungleichung abschätzen:

$$\begin{aligned}|f(x) - f(x_0)| &= \left|\sum_{n=1}^{\infty} a_n(x^n - x_0^n)\right| \leq \sum_{n=1}^{\infty} |a_n(x^n - x_0^n)| \\ &\leq \sum_{n=1}^{\infty} |a_n||x - x_0|n\varrho^{n-1} = |x - x_0|s.\end{aligned}$$

Wir haben also die Ungleichungskette

$$0 \leq |f(x) - f(x_0)| \leq s|x - x_0|.$$

Mit dem Sandwich-Theorem für Funktionengrenzwerte (siehe Satz 17.10 (c)) folgt daraus $\lim_{x\to x_0} f(x) = f(x_0)$ und damit ist f nach Übungsaufgabe 18.4 (a) stetig in x_0. Da x_0 beliebig war, ist f stetig auf ganz D. $\square$

Beispiel 18.7. (a) Dank Satz 18.6 wissen wir nun, dass alle Polynome, die Exponentialfunktion, Sinus und Cosinus jeweils auf ganz $\mathbb{C}$, bzw. $\mathbb{R}$, stetige Funktionen sind.

(b) Es gilt
$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1.$$
Zum Nachweis dieser Aussage überlegen wir uns, dass für $x \neq 0$ gilt
$$\begin{aligned}\frac{\sin(x)}{x} &= \frac{1}{x}\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1} = \frac{1}{x}\Big(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots\Big) \\ &= 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \dots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n} =: g(x)\end{aligned}$$
und (wie man leicht nachrechnet) hat die Potenzreihe, die g definiert, den Konvergenzradius unendlich. Also ist diese Funktion in 0 stetig und es gilt, da g mit der von uns untersuchten Funktion für alle $x \neq 0$ übereinstimmt,
$$\lim_{x\to 0} \frac{\sin(x)}{x} = \lim_{x\to 0} g(x) = g(0) = 1.$$
Die Funktion $f : \mathbb{R} \to \mathbb{R}$, gegeben durch
$$f(x) = \begin{cases} \dfrac{\sin(x)}{x}, & \text{falls } x \neq 0, \\ 1, & \text{falls } x = 0, \end{cases}$$
ist also auf ganz $\mathbb{R}$ stetig. Man nennt f die *stetige Fortsetzung* der Funktion $x \mapsto \sin(x)/x$.

(c) Genauso wie eben kann man den Grenzwert
$$\lim_{x\to 0} \frac{E(x) - 1}{x} = 1$$
bestimmen, denn es ist für alle $x \neq 0$
$$\frac{E(x) - 1}{x} = \frac{1}{x}\Big(\sum_{n=0}^{\infty} \frac{x^n}{n!} - 1\Big) = \frac{1}{x}\sum_{n=1}^{\infty} \frac{x^n}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{x^n}{(n+1)!}$$
und auch diese Potenzreihe hat den Konvergenzradius ∞.

19. Eigenschaften stetiger Funktionen

Dieser Abschnitt enthält einen wichtigen Satz über stetige, reellwertige Funktionen nach dem anderen. Wir werden später immer wieder auf die hier entwickelten Ergebnisse zurückgreifen. In allen Sätzen dieses Abschnitts ist nur $\mathbb{K} = \mathbb{R}$ zugelassen. (Überlegen Sie sich, warum!)

Satz 19.1 (Zwischenwertsatz)**.** *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to \mathbb{R}$ sei stetig auf $[a, b]$. Ist y_0 eine Zahl zwischen $f(a)$ und $f(b)$, so gibt es ein $x_0 \in [a, b]$ mit $f(x_0) = y_0$.*

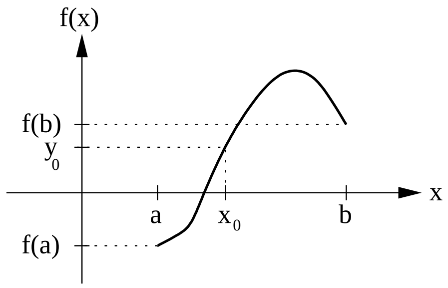

Abbildung 19.1.: Der Zwischenwertsatz

Beweis. Ist $y_0 = f(a)$ oder $y_0 = f(b)$, so sind wir bereits fertig. Wir können also annehmen, dass $y_0 \neq f(a)$ und $y_0 \neq f(b)$ gilt. Weiter nehmen wir an, dass $f(a) < f(b)$ ist, denn im Fall $f(a) = f(b)$ muss $y_0 = f(a) = f(b)$ sein, und den Fall $f(a) > f(b)$ kann man analog behandeln. Wir haben also $f(a) < y_0 < f(b)$. Setze

$$M := \{x \in [a, b] \mid f(x) \leq y_0\}.$$

Dann gilt $a \in M$, also ist $M \neq \emptyset$. Außerdem ist M durch b nach oben beschränkt, d. h. $x_0 := \sup M$ existiert. Daher kann für alle $n \in \mathbb{N}$ die Zahl $x_0 - 1/n$ keine obere Schranke von M sein, also existiert für jedes $n \in \mathbb{N}$ ein $x_n \in M$ mit

$$x_0 - \frac{1}{n} < x_n \leq x_0.$$

Nach dem Sandwich-Theorem konvergiert die Folge (x_n) gegen x_0 und da für jedes Folgenglied $a \leq x_n \leq b$ gilt, haben wir damit auch gleich $x_0 \in [a, b]$.

Nun folgern wir aus der Stetigkeit von f mit Satz 18.2, dass die Folge $(f(x_n))$ konvergiert und $\lim_{n\to\infty} f(x_n) = f(x_0)$ gilt. Da außerdem jedes x_n in M gewählt war, gilt $f(x_n) \leq y_0$ für alle $n \in \mathbb{N}$, also ist im Grenzwert $f(x_0) \leq y_0$. Es bleibt noch die umgekehrte Ungleichung $f(x_0) \geq y_0$ zu zeigen.
Dazu beobachten wir zunächst, dass sogar $x_0 \in [a, b)$ gelten muss, denn $x_0 = b$ kann wegen $f(x_0) \leq y_0$ und $f(b) > y_0$ nicht sein. Nun nehmen wir an, es wäre $f(x_0) < y_0$. Dann ist $\varepsilon := y_0 - f(x_0) > 0$. Nach der Definition der Stetigkeit gibt es zu diesem ε ein $\delta > 0$, so dass

$$|f(z) - f(x_0)| < \varepsilon \quad \text{für alle } z \in [a, b] \cap U_\delta(x_0)$$

gilt. Sei nun ein $z \in [a, b]$ mit $x_0 < z < x_0 + \delta$ gewählt. Das geht, da $x_0 < b$ ist. Dann gilt

$$f(z) - f(x_0) \leq |f(z) - f(x_0)| < \varepsilon = y_0 - f(x_0).$$

Also ist $f(z) < y_0$ und damit $z \in M$. Da x_0 das Supremum von M ist, muss dann $z \leq x_0$ gelten, was ein Widerspruch ist.
Zusammen ist damit $f(x_0) = y_0$. □

Eine wichtige Folgerung aus diesem Satz ist die folgende.

Satz 19.2 (Nullstellensatz von Bolzano)**.** *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f \in C([a, b])$ mit $f(a)f(b) < 0$ gegeben. Dann gibt es ein $x_0 \in (a, b)$ mit $f(x_0) = 0$.*

Beweis. Wir müssen uns nur klarmachen, dass die Voraussetzung $f(a)f(b) < 0$ bedeutet, dass entweder $f(a) > 0$ und $f(b) < 0$ oder $f(a) < 0$ und $f(b) > 0$ ist. Dann folgt die Behauptung sofort aus Satz 19.1. □

Wir können dieses Ergebnis nun anwenden, um das Bild der reellen Exponentialfunktion zu bestimmen.

Beispiel 19.3. Wir zeigen, dass für die Exponentialfunktion gilt:

$$E(\mathbb{R}) = \{E(x) \mid x \in \mathbb{R}\} = (0, \infty).$$

Wir wissen schon (vgl. Satz 14.9 (c)), dass $E(x) > 0$ für alle $x \in \mathbb{R}$ gilt, d. h. es ist $E(\mathbb{R}) \subseteq (0, \infty)$. Es bleibt die umgekehrte Inklusion zu zeigen.
Sei dazu $y_0 \in (0, \infty)$. Wir wissen aus Beispiel 17.15 (b), dass

$$\lim_{x\to-\infty} E(x) = 0 \quad \text{und} \quad \lim_{x\to\infty} E(x) = \infty$$

gilt. Also gibt es ein $a \in \mathbb{R}$, so dass $E(a) < y_0$ gilt und ein $b \in \mathbb{R}$ mit $E(b) > y_0$. Da damit zwangsläufig $E(a) < E(b)$ gilt und die Exponentialfunktion nach Satz 14.9 (e) streng monoton wachsend ist, muss $a < b$ gelten. Da die Exponentialfunktion nach Satz 18.6 auch auf ganz $\mathbb{R}$ stetig ist, sind nun alle Voraussetzungen von Satz 19.1 erfüllt. Es gibt also ein $x_0 \in \mathbb{R}$ mit $E(x_0) = y_0$. Damit ist $y_0 \in E(\mathbb{R})$ und wir sind fertig.

Definition 19.4. *Es sei $D \subseteq \mathbb{K}$. Eine Funktion $f : D \to \mathbb{K}$ heißt* beschränkt*, falls die Menge $f(D)$ beschränkt ist, d. h. falls ein $C \geq 0$ existiert, so dass $|f(x)| \leq C$ für alle $x \in D$ gilt.*

Definition 19.5. *(a) Eine Teilmenge A von $\mathbb{K}$ heißt* abgeschlossen*, wenn für jede Folge (x_n) in A, die in $\mathbb{K}$ konvergiert,* $\lim_{n\to\infty} x_n \in A$ *gilt.*

(b) Eine Teilmenge K von $\mathbb{K}$ heißt kompakt*, wenn sie abgeschlossen und beschränkt ist.*

Warnung 19.6. Die angegebene Definition der Kompaktheit ist für die hier betrachteten Teilmengen von $\mathbb{K}$ sehr einfach und praktisch. Der Kompaktheitsbegriff kommt in vielen Bereichen der Mathematik vor und ist oft sehr wichtig. Im Allgemeinen hat er eine andere, leider deutlich sprödere, Definition, die wir zu Beginn der Analysis II kennenlernen werden. Diese ist im Allgemeinen *nicht* äquivalent zu der hier angegebenen Definition!

Wir können nun den fundamentalen Satz formulieren, der das Verhalten stetiger Funktionen auf kompakten Mengen beschreibt.

Satz 19.7. *Es sei $K \subseteq \mathbb{R}$ kompakt und nicht-leer sowie $f \in C(K)$. Dann gibt es $x_*, x^* \in K$, so dass*

$$f(x_*) \leq f(x) \leq f(x^*) \quad \textit{für alle } x \in K$$

gilt. Insbesondere ist f beschränkt.

Man beachte, dass damit insbesondere $\max f(K) = f(x^*)$ und $\min f(K) = f(x_*)$ existieren.
In Worte gefasst lautet dieser Satz damit:

> Eine stetige Funktion auf einem Kompaktum nimmt ihr Maximum und ihr Minimum auf dem Kompaktum an.

Dass dabei jede der Voraussetzungen zwingend nötig ist, veranschaulichen die folgenden Beispiele.

Beispiel 19.8. (a) Ist $K = [0, 1]$ und

$$f(x) = \begin{cases} x, \text{ falls } x \in (0,1), \\ \dfrac{1}{2}, \text{ falls } x = 0 \text{ oder } x = 1, \end{cases}$$

so ist $f(K) = (0, 1)$, d.h. es gibt keine $x_*, x^* \in [0, 1]$ mit $f(x_*) = 0$ und $f(x^*) = 1$. Wir brauchen also die Stetigkeit von f.

(b) Ist $K = (0, 1]$ und $f(x) = 1/x$ für $x \in K$, so ist f auf K stetig, aber die Menge $f(K)$ ist nicht beschränkt, da $\lim_{x\to 0+} f(x) = \infty$ ist. Wir brauchen also die Abgeschlossenheit von K.

(c) Ist schließlich $K = \mathbb{R}$ und $f(x) = E(x)$, so ist diese Funktion wieder auf ganz K stetig und K ist abgeschlossen, aber f nicht beschränkt. Wir brauchen also die Beschränktheit von K.

Beweis von Satz 19.7. Wir zeigen zunächst, dass unter den Voraussetzungen des Satzes die Funktion f beschränkt ist. Wäre dem nicht so, gäbe es für jedes $n \in \mathbb{N}$ ein $x_n \in K$ mit $|f(x_n)| > n$. Dank der Beschränktheit von K und dem Satz von Bolzano-Weierstraß (Satz 9.9) können wir nun aus der Folge (x_n) eine konvergente Teilfolge (x_{n_k}) auswählen. Da K außerdem abgeschschlossen ist, muss deren Grenzwert $x_0 := \lim_{k\to\infty} x_{n_k}$ ebenfalls in K liegen. Nun nutzen wir die Stetigkeit von f auf K und folgern, dass die Folge $(f(x_{n_k}))$ gegen $f(x_0)$ konvergiert. Insbesondere ist damit die Folge $(f(x_{n_k}))$ beschränkt, was im Widerspruch zur Konstruktion der Folge (x_n) steht, nach der $|f(x_{n_k})| > n_k$ für alle $k \in \mathbb{N}$ gilt.
Da nun $f(K)$ beschränkt ist, existieren zumindest $\sup f(K)$ und $\inf f(K)$. Wir betrachten hier nur $S := \sup f(K)$, die Untersuchung für das Infimum verläuft analog. Zu zeigen ist, dass es ein $x^* \in K$ gibt, so dass $f(x^*) = S$ gilt. Dazu stellen wir zunächst fest, dass für jedes $n \in \mathbb{N}$ die Zahl $S - 1/n$ keine obere Schranke von $f(K)$ sein kann. Also gibt es jeweils ein $y_n \in f(K)$, und damit ein $x_n \in K$ mit $f(x_n) = y_n$, so dass

$$S - \frac{1}{n} < f(x_n) \leq S \quad \text{für alle } n \in \mathbb{N} \tag{19.1}$$

gilt. Die so gewonnene Folge (x_n) enthält wie jede Folge in K eine konvergente Teilfolge (x_{n_k}). Wir setzen

$$x^* := \lim_{k\to\infty} x_{n_k}.$$

Dann gilt wegen der Abgeschlossenheit von K sofort $x^* \in K$ und dank der Stetigkeit von f haben wir $f(x_{n_k}) \to f(x^*)$ für $k \to \infty$. Mit Hilfe von (19.1) und dem Sandwich-Theorem gilt außerdem

$$\lim_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = S.$$

Also ist $f(x^*) = S$. □

Als nächstes wollen wir zeigen, dass sich die Stetigkeit einer Funktion auf ihre Umkehrfunktion überträgt, sofern diese existiert. Wir beobachten dazu, dass für ein Intervall I, insbesondere ist auch $I = \mathbb{R}$ zugelassen, und eine streng monotone Funktion $f : I \to \mathbb{R}$, f auf I injektiv ist (Übungsaufgabe!). Damit ist nach Einschränkung des Wertebereichs die Funktion $f : I \to f(I)$ bijektiv und die Umkehrfunktion f^{-1} existiert in diesem Fall. Wir wissen über die Umkehrfunktion sogar noch mehr.

Lemma 19.9. *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to f(I)$ eine streng monoton wachsende (bzw. fallende) Funktion. Dann ist die Umkehrfunktion f^{-1} ebenfalls streng monoton wachsend (bzw. fallend).*

Beweis. Wir führen den Beweis nur für wachsende Funktionen, die geklammerte Aussage beweist man analog. Es sei also f streng monoton wachsend.
Es seien $y_1, y_2 \in f(I)$ mit $y_1 < y_2$ gegeben. Dann existieren $x_1, x_2 \in I$, so dass $f(x_1) = y_1$ und $f(x_2) = y_2$ gilt. Da f streng monoton wächst und $f(x_1) < f(x_2)$ gilt, muss auch $x_1 < x_2$ sein. Damit ist aber

$$f^{-1}(y_1) = x_1 < x_2 = f^{-1}(y_2),$$

was nichts anderes bedeutet, als dass f^{-1} streng monoton wächst. □

Für den Beweis des nächsten Satzes brauchen wir noch das folgende Lemma, dessen Beweis als Übung stehen bleibt.

Lemma 19.10. *Eine Menge $M \subseteq \mathbb{R}$ ist genau dann ein Intervall, wenn für je zwei Zahlen $a, b \in M$ mit $a < b$ stets $[a, b] \subseteq M$ gilt.*

Satz 19.11. *Es sei $I \subseteq \mathbb{R}$ ein beliebiges Intervall und $f \in C(I)$ sei streng monoton. Dann ist $f(I)$ ebenfalls ein Intervall und es gilt $f^{-1} \in C(f(I))$.*

Beweis. Wir führen den Beweis wieder nur für streng wachsende Funktionen f. Wir verwenden zunächst Lemma 19.10, um zu zeigen, dass $f(I)$ ein Intervall ist. Seien dazu $a, b \in f(I)$ mit $a < b$ und ein $y_0 \in [a, b]$ gegeben. Dann gibt es $\alpha, \beta \in I$, so dass $f(\alpha) = a$ und $f(\beta) = b$ gilt. Wir haben also

$$f(\alpha) = a \leq y_0 \leq b = f(\beta).$$

Da f stetig ist, existiert also nach dem Zwischenwertsatz 19.1 ein $x_0 \in I$ mit $f(x_0) = y_0$. Also ist $y_0 \in f(I)$ und wir erhalten $[a, b] \subseteq f(I)$.
Nun bleibt noch zu zeigen, dass $f^{-1} \in C(f(I))$ gilt. Sei hierzu zunächst I ein kompaktes Intervall, also $I = [\alpha, \beta]$ für irgendwelche $\alpha, \beta \in \mathbb{R}$ mit $\alpha < \beta$. Dann gilt nach dem ersten Teil dieses Beweises, dass $f(I) = [f(\alpha), f(\beta)]$ ist (Beachten Sie die Monotonie-Annahme!). Sei nun $y_0 \in f(I)$ und (y_n) eine Folge in $f(I)$, die gegen y_0 konvergiert. Wir müssen nun zeigen, dass

$$f^{-1}(y_n) \longrightarrow f^{-1}(y_0) \ (n \to \infty)$$

gilt. Wir setzen $x_n := f^{-1}(y_n)$ für jedes $n \in \mathbb{N}$ und $x_0 := f^{-1}(y_0)$. Dann gilt $x_n \in I = [\alpha, \beta]$ für jedes $n \in \mathbb{N}_0$ und wir haben zu zeigen, dass die Folge (x_n) gegen x_0 konvergiert.
Es sei γ ein Häufungswert von (x_n). Da I beschränkt vorausgesetzt ist, hat die Folge mindestens einen solchen. Es existiert also eine Teilfolge (x_{n_k}) von (x_n), die gegen γ konvergiert. Da I abgeschlossen ist, gilt damit auch $\gamma \in I = [\alpha, \beta]$. Wegen

der Stetigkeit von f konvergiert außerdem die Folge $(f(x_{n_k}))$ gegen $f(\gamma)$. Nun gilt aber $f(x_{n_k}) = y_{n_k}$ für alle $k \in \mathbb{N}$ und da (y_n) und damit auch (y_{n_k}) gegen y_0 konvergiert und der Grenzwert einer konvergenten Folge eindeutig ist, muss damit $f(\gamma) = y_0 = f(x_0)$ gelten. Auf Grund der Injektivität von f ist dann $\gamma = x_0$. Die beschränkte Folge (x_n) besitzt also genau einen Häufungswert, nämlich x_0. Damit konvergiert die Folge nach Satz 9.12 und es gilt $\lim_{n\to\infty} x_n = x_0$.
Abschließend bleibt noch der Fall eines beliebigen Intervalls zu behandeln. Dazu beachte man, dass nach Definition eine Funktion auf einer Menge stetig ist, wenn sie in jedem Punkt der Menge stetig ist. Können wir also zeigen, dass $f^{-1} \in C([a,b])$ für jedes Intervall $[a,b] \subseteq f(I)$ ist, so ist f^{-1} auf $f(I)$ stetig, da man um jeden Punkt in $f(I)$ so ein Intervall legen kann.
Seien also $a, b \in f(I)$ mit $a < b$. Dann gibt es wieder $\alpha, \beta \in I$ mit $f(\alpha) = a$ und $f(\beta) = b$ und es gilt $\alpha < \beta$ wegen der strengen Monotonie von f. Also können wir für das Intervall $[\alpha, \beta]$ den obigen Beweis verwenden und erhalten, dass f^{-1} auf $[f(\alpha), f(\beta)] = [a, b]$ stetig ist. □

Auch dieser Satz liefert uns eine neue spannende Information über die Exponentialfunktion, denn von dieser wissen wir, dass sie auf $I = \mathbb{R}$ streng monoton wächst (vgl. Satz 14.9 (e)). Außerdem haben wir in Beispiel 19.3 gesehen, dass $E(\mathbb{R}) = (0, \infty)$ gilt. Damit wissen wir, dass die Umkehrfunktion existiert. Diese bekommt einen eigenen Namen.

Definition 19.12. *Die Funktion*

$$\ln := \log := E^{-1} : (0,\infty) \to \mathbb{R} \quad \textit{mit} \quad \ln(x) := \log(x) := E^{-1}(x),\ x \in (0,\infty),$$

heißt (natürlicher) Logarithmus.

Der Logarithmus hat die folgenden Eigenschaften.

Satz 19.13. *(a) Die Funktion* $\ln$ *ist auf* $(0,\infty)$ *stetig und wächst streng monoton.*

(b) Es gilt $\ln(1) = 0$ *und* $\ln(\mathrm{e}) = 1$.

(c) $\lim\limits_{x\to\infty} \ln(x) = \infty$ *und* $\lim\limits_{x\to 0+} \ln(x) = -\infty$.

(d) Für alle $x, y \in (0,\infty)$ *und* $n \in \mathbb{Z}$ *gilt*

$$\ln(xy) = \ln(x) + \ln(y), \qquad \ln\Big(\frac{x}{y}\Big) = \ln(x) - \ln(y), \qquad \ln(x^n) = n\ln(x).$$

(e) Für alle $x \in (0,\infty)$ *und alle* $r \in \mathbb{Q}$ *gilt*

$$\ln(x^r) = r\ln(x).$$

Beweis. (a) Ergibt sich sofort aus Satz 19.11.

(b) Ergibt sich aus Satz 14.9 (b).

(c) Ergibt sich aus Beispiel 17.15 (b).

(d) Wir setzen $\xi := \ln(x)$ und $\eta := \ln(y)$. Dann gilt nach Satz 14.9 (a)

$$E(\xi + \eta) = E(\xi)E(\eta) = E\big(\ln(x)\big)E\big(\ln(y)\big) = xy.$$

Also ist

$$\ln(xy) = \ln\big(E(\xi + \eta)\big) = \xi + \eta = \ln(x) + \ln(y).$$

Weiter gilt mit Satz 14.9 (c)

$$E\Big(-\ln\Big(\frac{1}{x}\Big)\Big) = \frac{1}{E\big(\ln(1/x)\big)} = \frac{1}{1/x} = x.$$

Also ist

$$-\ln(x) = -\ln\Big(E\Big(-\ln\Big(\frac{1}{x}\Big)\Big)\Big) = \ln\Big(\frac{1}{x}\Big).$$

Damit können wir aus der ersten Formel sofort folgern

$$\ln\Big(\frac{x}{y}\Big) = \ln\Big(x\frac{1}{y}\Big) = \ln(x) + \ln\Big(\frac{1}{y}\Big) = \ln(x) - \ln(y).$$

Es bleibt noch die dritte Formel. Für $n \in \mathbb{N}$ folgt diese sofort aus der ersten und für negative n aus

$$\ln(x^n) = \ln\big((x^{-n})^{-1}\big) = -\ln(x^{-n}) = n\ln(x).$$

(e) Ist $r \in \mathbb{Q}$ und $r = m/n$ mit $m \in \mathbb{Z}$ und $n \in \mathbb{N}$ so gilt mit den in (d) gezeigten Beziehungen

$$\ln(x^r) = \ln\big((x^{1/n})^m\big) = m\ln(x^{1/n}) = nr\ln(x^{1/n}) = r\ln\big((x^{1/n})^n\big) = r\ln(x).$$

□

Sehen wir uns die Aussage in (e) noch einmal an, so folgt daraus insbesondere für alle $a \in (0, \infty)$ und alle $r \in \mathbb{Q}$ die Beziehung $a^r = E(\ln(a^r)) = E(r\ln(a))$. Diese verwenden wir nun um die allgemeine Potenzfunktion zu definieren.

Definition 19.14. *Für alle $a \in (0, \infty)$ und alle $x \in \mathbb{R}$ definieren wir die* allgemeine Potenz *durch*

$$a^x := E\big(x \cdot \ln(a)\big).$$

Man beachte, dass damit insbesondere für $a = \mathrm{e}$ auch die übliche Schreibweise $\mathrm{e}^x = E(x)$ gerechtfertigt ist.

Wir sammeln Eigenschaften dieser Funktion.

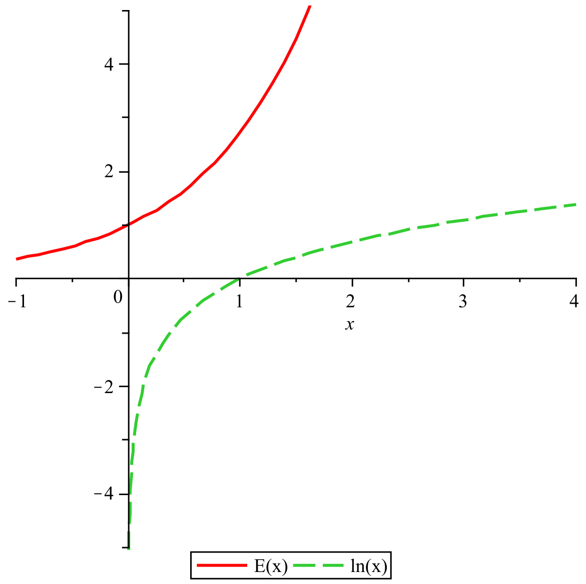

Abbildung 19.2.: Die Graphen der Exponentialfunktion und des Logarithmus

Satz 19.15. *Es sei $a \in (0, \infty)$. Dann ist die Funktion $x \mapsto a^x$ stetig auf $\mathbb{R}$ und es gelten die bekannten Rechenregeln für Potenzen wie beispielsweise*

$$a^{x+y} = a^x a^y, \quad a^{-1} = \frac{1}{a}, \quad (a^x)^y = a^{xy}.$$

Beweis. Wir beobachten zunächst, dass die beiden Funktionen $x \mapsto x \cdot \ln(a)$ und $z \mapsto E(z)$ jeweils auf $\mathbb{R}$ stetig sind, also ist auch die Potenzfunktion als deren Verkettung nach Satz 18.5 stetig.
Die Rechenregeln lassen sich alle direkt aus jenen für die Exponentialfunktion ableiten. Wir behandeln deshalb hier nur beispielhaft

$$a^{x+y} = E\big((x+y)\ln(a)\big) = E\big(x\ln(a) + y\ln(a)\big) = E\big(x\ln(a)\big)E\big(y\ln(a)\big) = a^x a^y.$$

$\square$

20. Funktionenfolgen und -reihen

Wir erinnern zunächst an die Absprache von Seite 97, dass der Buchstabe $\mathbb{K}$ jeweils für $\mathbb{R}$ oder $\mathbb{C}$ steht.
In diesem Abschnitt betrachten wir Folgen, deren Glieder, bzw. Reihen, deren Summanden selbst wieder Funktionen sind. Dafür führen wir den folgenden Begriff ein.

Definition 20.1. *Es sei $D \subseteq \mathbb{K}$ und für jedes $n \in \mathbb{N}$ sei eine Funktion $f_n : D \to \mathbb{K}$ gegeben.*

(a) Wir bezeichnen mit (f_n) die Funktionenfolge $(f_1, f_2, f_3, \dots)$ *und sagen, die Funktionenfolge* konvergiert punktweise, *wenn für jedes $x \in D$ die reelle Zahlenfolge $(f_n(x))$ konvergiert. In diesem Fall heißt die Funktion*

$$f : \begin{cases} D \to \mathbb{K} \\ x \mapsto \lim_{n\to\infty} f_n(x) \end{cases}$$

die Grenzfunktion *von (f_n).*

(b) Wir bezeichnen mit $\sum_{n=1}^{\infty} f_n$ die Funktionenreihe $f_1 + f_2 + f_3 + \dots$. *Die Funktionenreihe* konvergiert punktweise, *wenn für jedes $x \in D$ die Reihe $\sum_{n=1}^{\infty} f_n(x)$ konvergiert. In diesem Fall heißt die Funktion*

$$s : \begin{cases} D \to \mathbb{K} \\ x \mapsto \sum_{n=1}^{\infty} f_n(x) \end{cases}$$

die Summenfunktion.

Beispiel 20.2. (a) Es sei $D = [0, 1]$ und

$$f_n(x) := x^n, \quad x \in [0, 1],$$

für jedes $n \in \mathbb{N}$.

Nach Satz 7.2 gilt $\lim_{n\to\infty} x^n = 0$ für alle $x \in [0, 1)$ und für $x = 1$ erhalten wir für jedes $n \in \mathbb{N}$ den Wert 1, also konvergiert (f_n) punktweise gegen die Funktion $f : [0, 1] \to \mathbb{R}$ mit

$$f(x) := \begin{cases} 0, \text{ falls } x \in [0, 1), \\ 1, \text{ falls } x = 1. \end{cases}$$

(b) Es sei (a_n) eine reelle Folge und für jedes $n \in \mathbb{N}$ die Funktion

$$f_n(x) = a_n x^n$$

gegeben. Dann ist die aus diesen Funktionen gebildete Funktionenreihe genau die durch die Folge (a_n) gegebene Potenzreihe. Sie konvergiert also, falls der Konvergenzradius r der Potenzreihe positiv ist, innerhalb des Intervalls $(-r, r)$ punktweise gegen die Funktion $s : (-r, r) \to \mathbb{R}$ mit

$$s(x) = \sum_{n=1}^{\infty} a_n x^n.$$

Die Potenzreihen sind also ein Spezialfall von Funktionenreihen.

(c) Es sei $D := [0, \infty)$ und für jedes $n \in \mathbb{N}$ sei $f : [0, \infty) \to \mathbb{R}$ gegeben durch

$$f_n(x) := \frac{nx}{1 + n^2 x^2}, \quad x \in [0, \infty).$$

Dann gilt für alle $x \in [0, \infty)$

$$\lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \frac{x/n}{1/n^2 + x^2} = 0,$$

also konvergiert (f_n) in diesem Beispiel auf $[0, \infty)$ punktweise gegen $f = 0$.

Bemerkung 20.3. (a) In Epsilons ausgedrückt bedeutet punktweise Konvergenz einer Funktionenfolge (f_n) auf einer Menge $D \subseteq \mathbb{K}$:

$$\forall \varepsilon > 0 \; \forall x \in D \; \exists n_0 = n_0(\varepsilon, x) \in \mathbb{N} \; \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon. \tag{20.1}$$

Eine entsprechende Aussage lässt sich natürlich auch für Funktionenreihen hinschreiben, wenn man beachtet, dass Konvergenz der Reihe nichts anderes als die Konvergenz der Folge der Partialsummen bedeutet.

(b) Schauen wir uns noch einmal unser drittes Beispiel von oben an, vgl. Abbildung 20.1, so sehen wir rechnerisch sofort ein, dass diese Funktionenfolge punktweise gegen die Nullfunktion konvergiert. Schaut man sich jedoch für jedes $n \in \mathbb{N}$ den größten Abstand des Funktionsgraphen der Funktion f_n von der x-Achse und damit vom Graphen der Nullfunktion an, so weigert sich dieser hartnäckig gegen Null zu streben, sondern bleibt immer konstant $1/2$. Rechnerisch, sieht man das daran, dass für alle $n \in \mathbb{N}$ gilt

$$\left| f_n\Big(\frac{1}{n}\Big) - f\Big(\frac{1}{n}\Big) \right| = \left| \frac{n\frac{1}{n}}{1 + n^2(\frac{1}{n})^2} - 0 \right| = \left| \frac{1}{1+1} \right| = \frac{1}{2}. \tag{20.2}$$

Wir wollen im Folgenden einen weiteren, restriktiveren Konvergenzbegriff einführen, der solch ein Konvergenzverhalten nicht mehr „toleriert“.

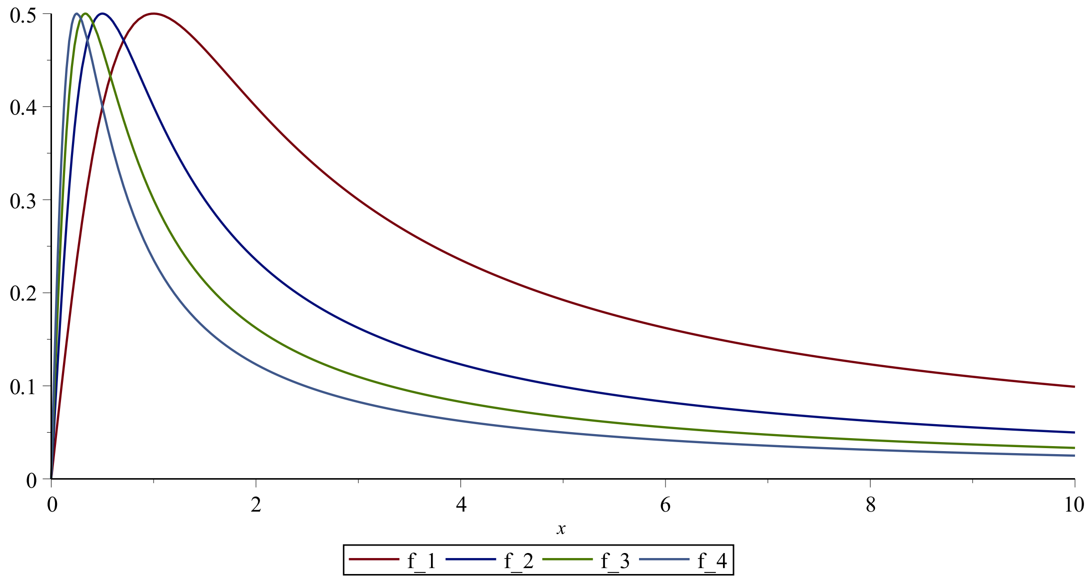

Abbildung 20.1.: Die Graphen der ersten vier Funktionen in der Funktionenfolge aus Beispiel 20.2 (c).

Definition 20.4. *Es sei $D \subseteq \mathbb{K}$ und für jedes $n \in \mathbb{N}$ seien Funktionen $f_n : D \to \mathbb{K}$ gegeben.*

(a) Die Funktionenfolge (f_n) konvergiert gleichmäßig *auf D gegen f, falls für jedes $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert, so dass*

$$|f_n(x) - f(x)| < \varepsilon \text{ für alle } n \geq n_0 \text{ und alle } x \in D$$

gilt.

(b) Die Funktionenreihe $\sum_{n=1}^{\infty} f_n$ konvergiert gleichmäßig *auf D gegen s, genau dann wenn für jedes $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert, so dass*

$$\left|\sum_{k=1}^{n} f_k(x) - s(x)\right| < \varepsilon \text{ für alle } n \geq n_0 \text{ und alle } x \in D$$

gilt.

Bemerkung 20.5. (a) Schreiben wir nun auch die Bedingung für gleichmäßige Konvergenz in Epsilontisch, so erhalten wir

$$\forall \varepsilon > 0 \; \exists n_0 = n_0(\varepsilon) \in \mathbb{N} \; \forall n \geq n_0 \; \forall x \in D : |f_n(x) - f(x)| < \varepsilon.$$

Vergleichen wir mit der entsprechenden Definition der punktweisen Konvergenz aus (20.1), so sehen wir den Unterschied: Der Quantor „$\forall x \in D$“

ist von vorne nach hinten gerutscht. Das macht einen großen Unterschied. Bei punktweiser Konvergenz dürfen wir bei der Auswahl des n_0 sowohl die zugelassene Abweichung von der Grenzfunktion ε als auch den Wert für x einfließen lassen und für verschiedene x unter Umständen verschiedene n_0 wählen, während es bei gleichmäßiger Konvergenz zu jedem ε ein n_0 geben muss, das für alle $x \in D$ das selbe ist. Wir brauchen in diesem Sinne ein universelles oder eben gleichmäßiges n_0, dass für alle $x \in D$ simultan den Abstand $|f_n(x) - f(x)|$ kleiner als ε garantiert.

(b) Obige Überlegung zeigt auch sofort, dass jede Funktionenfolge, die gleichmäßig gegen eine Funktion f konvergiert, insbesondere auch punktweise gegen die selbe Funktion konvergiert: Wenn wir ein universelles n_0 haben, erfüllt dieses die Konvergenzbedingung natürlich auch für jedes $x \in D$ einzeln.

(c) Anschaulich bedeutet gleichmäßige Konvergenz gegen f, dass die Graphen der Funktionen f_n ab einem gewissen n_0 alle ganz in einem ε-Streifen um den Graphen der Funktion f liegen, vgl. Abbildung 20.2.

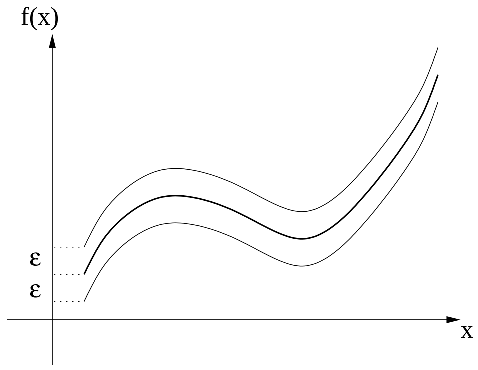

Abbildung 20.2.: Der ε-Streifen um den Graphen der Grenzfunktion f.

Betrachten wir wieder das Beispiel 20.2 (c) von oben, so sehen wir, dass das eben nicht der Fall ist. So verlässt jede Funktion f_n irgendwo den Streifen um die x-Achse mit Breite 1/4.

Beispiel 20.6. Wir betrachten im Lichte der neuen Definition noch einmal (a) und (c) aus Beispiel 20.2. Beide Funktionenfolgen sind nicht gleichmäßig konvergent.
Für das Beispiel

$$f_n(x) = \frac{nx}{1 + n^2x^2}, \quad x \in [0, \infty),$$

folgt das direkt aus (20.2), denn für $\varepsilon := 1/4$ gibt es für jedes $n \in \mathbb{N}$ ein $x \in [0, \infty)$, für das $|f_n(x) - f(x)| \geq \varepsilon$ gilt.

Für
$$f_n(x) = x^n, \quad x \in [0,1],$$
erhält man wegen $1/\sqrt[n]{2} \in (0,1)$ für alle $n \in \mathbb{N}$
$$\left| f_n\Big(\frac{1}{\sqrt[n]{2}}\Big) - f\Big(\frac{1}{\sqrt[n]{2}}\Big) \right| = \left| \Big(\frac{1}{\sqrt[n]{2}}\Big)^n - 0 \right| = \frac{1}{2}$$
auf dem gleichen Wege, dass die Funktionenfolge nicht gleichmäßig konvergiert.

Die Frage der gleichmäßigen Konvergenz hängt manchmal sehr stark vom betrachteten Intervall ab, was nicht weiter verwundert, denn je größer dieses ist, desto mehr x muss ein zu vorgegebenem ε gewähltes n_0 gleichzeitig verarzten. Schauen wir nochmals die obigen Beispiele an, so sehen, wir, dass bei (a) das Problem bei $x = 1$ liegt und bei (c) bei $x = 0$. Halten wir uns von diesen beiden Punkten fern, so können wir tatsächlich gleichmäßige Konvergenz nachweisen.

Beispiel 20.7. (a) Wählen wir ein $\alpha \in (0,1)$, setzen wir $\tilde{D} := [0, \alpha]$ und betrachten nun auf dieser Menge die Funktionenfolge
$$f_n(x) := x^n, \quad x \in \tilde{D},$$
so gilt für alle $x \in \tilde{D}$ die Abschätzung $|f_n(x) - f(x)| = |x^n - 0| = x^n \le \alpha^n$. (Man beachte, dass die Grenzfunktion auf $\tilde{D}$ nun die Nullfunktion ist.) Sei nun $\varepsilon > 0$ gegeben. Dann gibt es ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$, so dass $\alpha^n < \varepsilon$ für alle $n \ge n_0$ ist, da $\lim_{n\to\infty} \alpha^n = 0$ ist. Also gilt
$$|f_n(x) - f(x)| \le \alpha^n < \varepsilon \text{ für alle } n \ge n_0 \text{ und alle } x \in \tilde{D}$$
und das ist genau die Bedingung für gleichmäßige Konvergenz.

Zusammengefasst ist die Funktionenfolge (x^n) also gleichmäßig konvergent auf jedem Intervall der Form $[0, \alpha]$ mit $0 < \alpha < 1$, aber nicht auf $[0,1]$. Auf diesem ist sie aber noch punktweise konvergent.

(b) Für ein $\alpha > 0$ setzen wir nun $\tilde{D} := [\alpha, \infty)$ und betrachten darauf die Funktionenfolge
$$f_n(x) = \frac{nx}{1 + n^2x^2}, \quad x \in \tilde{D}.$$
Dann gilt für jedes $n \in \mathbb{N}$
$$|f_n(x) - f(x)| = \frac{nx}{1 + n^2x^2} \le \frac{nx}{n^2x^2} = \frac{1}{nx} \le \frac{1}{n\alpha}.$$
Für jedes $\varepsilon > 0$ gibt es nun ein $n_0 \in \mathbb{N}$ mit $1/(n\alpha) < \varepsilon$ für alle $n \ge n_0$, also gilt
$$|f_n(x) - f(x)| \le \frac{1}{n\alpha} < \varepsilon \text{ für alle } n \ge n_0 \text{ und alle } x \in \tilde{D}.$$
Zusammenfassend ist diese Funktionenfolge also auf jedem Intervall der Form $[\alpha, \infty)$ für $\alpha > 0$ gleichmäßig konvergent, aber nicht auf $[0, \infty)$.

Beim Nachweis der gleichmäßigen Konvergenz in obigem Beispiel haben wir jeweils das folgende allgemeine Prinzip verwendet.

Satz 20.8. *Es sei $D \subseteq \mathbb{K}$ und (f_n) eine Funktionenfolge auf D sowie $f : D \to \mathbb{K}$ eine Funktion. Gibt es eine Nullfolge (α_n), so dass*

$$|f_n(x) - f(x)| \leq \alpha_n \text{ für fast alle } n \in \mathbb{N} \text{ und alle } x \in D$$

gilt, so konvergiert (f_n) gleichmäßig auf D gegen f.

Der Beweis verläuft genau wie in obigem Beispiel. Führen Sie ihn dennoch zu Übungszwecken aus.

Übungsaufgabe 20.9. Es sei $D \subseteq \mathbb{K}$ und (f_n) eine Funktionenfolge auf D.

(a) Ist (f_n) gleichmäßig konvergent gegen $f : D \to \mathbb{K}$, so konvergiert auch die Folge $(|f_n|)$ gleichmäßig auf D und zwar gegen $|f|$.

(b) Die Funktionenfolge (f_n) konvergiert genau dann gleichmäßig gegen $f : D \to \mathbb{K}$, wenn $\lim_{n\to\infty} \sup_{x\in D} |f_n(x) - f(x)| = 0$ gilt. Ist die Funktion f beschränkt, so gilt in diesem Fall außerdem $\lim_{n\to\infty} \sup_{x\in D} |f_n(x)| = \sup_{x\in D} |f(x)|$.

Satz 20.10 (Majorantenkriterium für Funktionenreihen)**.** *Es seien $D \subseteq \mathbb{K}$ und (f_n) eine Funktionenfolge auf D. Gibt es dann eine Folge (c_n) in $\mathbb{R}$, so dass $\sum_{n=1}^{\infty} c_n$ konvergiert und*

$$|f_n(x)| \leq c_n \text{ für fast alle } n \in \mathbb{N} \text{ und alle } x \in D,$$

so konvergiert die Funktionenreihe $\sum_{n=1}^{\infty} f_n$ auf D gleichmäßig.

Beweis. Es gilt für jedes $x \in D$ und alle ausreichend großen $n \in \mathbb{N}$

$$\Big|\sum_{k=1}^{n} f_k(x) - \sum_{k=1}^{\infty} f_k(x)\Big| = \Big|\sum_{k=n+1}^{\infty} f_k(x)\Big| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} c_k =: \alpha_n.$$

Da die Reihe über die c_n, $n \in \mathbb{N}$, konvergiert, ist die Folge der Reihenreste (α_n) nach Satz 11.6 (b) eine Nullfolge. Damit folgt die Behauptung aus Satz 20.8. □

Wir bestätigen uns nun erneut, dass Potenzreihen etwas Freundliches sind und beweisen, dass diese (als Funktionenreihen aufgefasst) auf kompakten Mengen im Inneren ihres Konvergenzintervalls sogar gleichmäßig gegen ihre Summenfunktion konvergieren.

Satz 20.11. *Es sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit Konvergenzradius $r > 0$ und $[a, b]$ ein Intervall mit $[a, b] \subseteq (-r, r)$. Dann konvergiert die Potenzreihe (als Funktionenreihe aufgefasst) gleichmäßig auf $[a, b]$.*

Beweis. Die Potenzreihe ist eine Funktionenreihe, bei der über die Funktionen $f_n(x) = a_n x^n$, $n \in \mathbb{N}$, summiert wird.
Wir setzen $\varrho := \max\{|a|, |b|\}$. Dann gilt $[a, b] \subseteq [-\varrho, \varrho] \subseteq (-r, r)$, d. h. $|x| \leq \varrho < r$ für alle $x \in [a, b]$. Das bedeutet

$$|f_n(x)| = |a_n x^n| = |a_n||x|^n \leq |a_n|\varrho^n =: c_n.$$

Da $\varrho \in (-r, r)$ gilt, konvergiert die Reihe $\sum_{n=1}^{\infty} c_n$ nach dem Satz von Hadamard. Damit ist diese eine konvergente Majorante und die Behauptung folgt aus Satz 20.10. □

Um zu sehen, dass eine Potenzreihe im Allgemeinen nicht gleichmäßig auf dem vollen Konvergenzintervall konvergiert, kann dass das folgende Beispiel dienen.

Übungsaufgabe 20.12. Die geometrische Reihe $\sum_{n=0}^{\infty} x^n$ ist auf $(-1, 1)$ nicht gleichmäßig konvergent.

Wir zeigen nun, dass gleichmäßige Konvergenz Stetigkeit erhält, eine sehr wichtige Konsequenz dieser Eigenschaft.

Satz 20.13. *Es sei $D \subseteq \mathbb{K}$ und (f_n) sei eine Funktionenfolge (bzw. $\sum_{n=1}^{\infty} f_n$ eine Funktionenreihe) auf D, die auf D gleichmäßig gegen eine Funktion f (bzw. s) konvergiere. Sind die Funktionen f_n für alle $n \in \mathbb{N}$ in einem Punkt $x_0 \in D$ stetig, so ist auch die Grenzfunktion f (bzw. die Summenfunktion s) in x_0 stetig.*

Beweis. Wir führen den Beweis im Falle von Funktionenfolgen.
Es sei $\varepsilon > 0$. Nach Definition der Stetigkeit müssen wir ein $\delta > 0$ finden, so dass gilt

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

Wegen der gleichmäßigen Konvergenz von (f_n) gibt es ein $m \in \mathbb{N}$, so dass

$$|f_m(x) - f(x)| < \frac{\varepsilon}{3} \text{ für alle } x \in D$$

gilt. Weiter ist nach Voraussetzung die Funktion f_m stetig, also gibt es ein $\delta > 0$, so dass

$$|f_m(x) - f_m(x_0)| < \frac{\varepsilon}{3} \text{ für alle } x \in D \text{ mit } |x - x_0| < \delta$$

ist. Kombinieren wir diese Überlegungen, so gilt für alle $x \in D$ mit $|x - x_0| < \delta$ nun

$$\begin{aligned}|f(x) - f(x_0)| &= \big|f(x) - f_m(x) + f_m(x) - f_m(x_0) + f_m(x_0) - f(x_0)\big| \\ &\leq \big|f(x) - f_m(x)\big| + \big|f_m(x) - f_m(x_0)\big| + \big|f_m(x_0) - f(x_0)\big| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

□

Bemerkung 20.14. (a) Wir können Satz 20.13 auch folgendermaßen formulieren: Konvergiert eine Funktionenfolge (f_n) auf D gleichmäßig gegen f und sind alle Funktionen f_n in $x_0 \in D$ stetig, so gilt

$$\lim_{x\to x_0} \lim_{n\to\infty} f_n(x) = \lim_{x\to x_0} f(x) = f(x_0) = \lim_{n\to\infty} f_n(x_0) = \lim_{n\to\infty} \lim_{x\to x_0} f_n(x).$$

Wir haben in diesem Satz also gezeigt, dass man bei gleichmäßig konvergenten Funktionenfolgen den Konvergenz-Limes mit dem Stetigkeits-Limes vertauschen kann. Dieses Vertauschen von Grenzwerten ist im Allgemeinen nicht erlaubt, vgl. Warnung 7.6, und insofern sind Sätze dieser Art, die ein Vertauschen gestatten, sehr wertvoll.

(b) Satz 20.13 gibt außerdem noch ein manchmal sehr brauchbares Kriterium ab, um nachzuweisen, dass eine punktweise konvergente Funktionenfolge oder -reihe nicht gleichmäßig konvergiert. Sind nämlich alle Folgenglieder (bzw. alle Summanden) stetige Funktionen, aber die punktweise Grenzfunktion ist unstetig, so kann die Konvergenz nach diesem Satz nicht gleichmäßig sein.

Zum Abschluss dieses Abschnittes beweisen wir noch den folgenden, überraschenden Satz, dass zwei Funktionen, die durch Potenzreihen gegeben sind und die auf einer Nullfolge von Punkten übereinstimmen, schon identisch sein müssen.

Satz 20.15 (Identitätssatz für Potenzreihen)**.** *Gegeben seien zwei Potenzreihen $\sum_{n=0}^\infty a_n x^n$ und $\sum_{n=0}^\infty b_n x^n$ mit Konvergenzradien $r_1 > 0$, bzw. $r_2 > 0$. Wir setzen $R := \min\{r_1, r_2\} > 0$ und*

$$f(x) := \sum_{n=0}^\infty a_n x^n, \; |x| < r_1, \quad \text{und} \quad g(x) := \sum_{n=0}^\infty b_n x^n, \; |x| < r_2.$$

Gibt es dann eine Folge (x_k) in $\{y \in \mathbb{K} : |y| < R\}$ mit $\lim_{k\to\infty} x_k = 0$, sowie $x_k \neq 0$ und $f(x_k) = g(x_k)$ für alle $k \in \mathbb{N}$, so gilt $a_n = b_n$ für alle $n \in \mathbb{N}_0$, d. h. es ist $f(x) = g(x)$ für alle $|x| < R$.

Beweis. Wir führen den Nachweis, dass $a_n = b_n$ für alle $n \in \mathbb{N}_0$ gilt, induktiv. Für den Induktionsanfang $(n = 0)$ überlegen wir uns, dass f nach Satz 18.6 in 0 stetig ist, also gilt $\lim_{k\to\infty} f(x_k) = f(0) = a_0$. Das selbe folgt für g und da $f(x_k) = g(x_k)$ für alle $k \in \mathbb{N}$ ist, beobachten wir

$$a_0 = f(0) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} g(x_k) = g(0) = b_0.$$

Als Induktionsvoraussetzung gelte im Folgenden $a_j = b_j$ für alle $j \in \{0, \dots, n\}$. Damit gilt für alle $|x| < R$

$$f(x) - g(x) = \sum_{j=n+1}^\infty (a_j - b_j) x^j.$$

Also ergibt sich für alle $k \in \mathbb{N}$

$$\begin{aligned} 0 = f(x_k) - g(x_k) &= \sum_{j=n+1}^{\infty} (a_j - b_j) x_k^j \\ &= (a_{n+1} - b_{n+1}) x_k^{n+1} + (a_{n+2} - b_{n+2}) x_k^{n+2} + \dots \end{aligned}$$

und da $x_k \neq 0$ für alle $k \in \mathbb{N}$ ist, dürfen wir diese Gleichung durch x_k^{n+1} teilen. Das ergibt

$$\sum_{j=0}^{\infty} (a_{n+1+j} - b_{n+1+j}) x_k^j = 0$$

für alle $k \in \mathbb{N}$. Setzen wir

$$\varphi(x) := \sum_{j=0}^{\infty} (a_{n+1+j} - b_{n+1+j}) x^j,$$

so ist das eine Potenzreihe mit Konvergenzradius größer oder gleich $R > 0$ (nachrechnen!) und somit ist φ wieder dank Satz 18.6 stetig in 0. Außerdem gilt $\varphi(x_k) = 0$ für alle $k \in \mathbb{N}$. Daraus folgt

$$a_{n+1} - b_{n+1} = \varphi(0) = \lim_{k \to \infty} \varphi(x_k) = \lim_{k \to \infty} 0 = 0,$$

also $a_{n+1} = b_{n+1}$. $\square$

21. Gleichmäßige Stetigkeit

Wir bekommen es in diesem Abschnitt mit einem ähnlichen Phänomen wie bei der Unterscheidung zwischen punktweiser und gleichmäßiger Konvergenz zu tun. Erinnern wir uns an die Definition der Stetigkeit, so war $f : D \to \mathbb{K}$ in einem Punkt $x_0 \in D$ genau dann stetig, wenn

$$\forall \varepsilon > 0\, \exists \delta = \delta(\varepsilon, x_0) > 0\, \forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - f(x_0)| < \varepsilon.$$

Das zu bestimmende δ darf hierbei außer von ε, von dem es logischerweise abhängen muss, auch von x_0 abhängen. Es liegt also nahe, ähnlich wie bei der Konvergenz von Funktionenfolgen eine „Stetigkeit von höherer Qualität" zu definieren, bei der das δ gleichmäßig in $x_0 \in D$ gewählt werden muss.
Dass das tatsächlich zu einem restriktiveren Begriff führt, zeigt das folgende Beispiel.

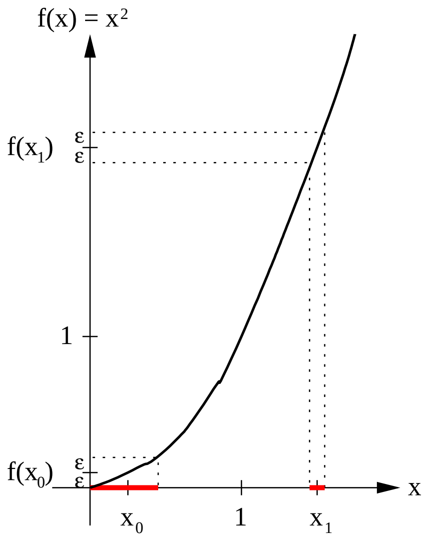

Abbildung 21.1.: Die Abhängigkeit des Stetigkeits-Deltas von x_0

Beispiel 21.1. Es sei $D = [0, \infty)$ und $f(x) = x^2$, $x \in D$, vgl. Abbildung 21.1. Diese Funktion ist offensichtlich stetig, beispielsweise weil sie durch eine Potenzreihe mit Konvergenzradius unendlich dargestellt wird. Also gibt es zu jedem $x_0 > 0$ und jedem $\varepsilon > 0$ ein $\delta > 0$, so dass

$$|x^2 - x_0^2| < \varepsilon$$

für alle $x > 0$ mit $|x - x_0| < \delta$ gilt. Können wir aber dieses δ unabhängig von x_0 wählen? Die Antwort ist Nein, denn wenn wir $x = x_0 + \delta/2$ setzen, so gilt $|x - x_0| = \delta/2 < \delta$, aber damit ist auch

$$\varepsilon > |x^2 - x_0^2| = |x + x_0||x - x_0| = \Big(2x_0 + \frac{\delta}{2}\Big)\frac{\delta}{2} = \delta x_0 + \frac{\delta^2}{4}.$$

Insbesondere ist damit $\delta x_0 < \varepsilon$, d. h.

$$\delta < \frac{\varepsilon}{x_0}.$$

Je größer also das x_0 wird, umso kleiner müssen wir bei gegebenem ε das δ wählen. Anschaulich liegt das daran, dass der Graph der Funktion für große x immer weiter ansteigt, wenn wir also im Bildbereich nur eine Abweichung von ε um das $f(x_0)$ zulassen, wird der verfügbare Platz für das δ auf der x-Achse immer geringer je weiter wir mit dem x_0 nach rechts rutschen.

Wir wollen nun die gleichmäßige Stetigkeit exakt definieren.

Definition 21.2. *Es sei $D \subseteq \mathbb{K}$. Dann heißt $f : D \to \mathbb{K}$* gleichmäßig stetig *auf D, falls für jedes $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert, so dass*

$$|f(x) - f(y)| < \varepsilon \text{ für alle } x, y \in D \text{ mit } |x - y| < \delta \text{ gilt.}$$

Bemerkung 21.3. (a) Es ist klar, dass eine Funktion, die auf einer Menge D gleichmäßig stetig ist, auch auf dieser Menge stetig ist, also zu $C(D)$ gehört. Die Umkehrung gilt i.A. nicht, wie Beispiel 21.1 zeigt.

(b) Wie bei der gleichmäßigen Konvergenz auch, ist die Frage, ob eine stetige Funktion sogar gleichmäßig stetig ist, sehr vom zu Grunde gelegten Definitionsbereich abhängig. Es ist eine Eigenschaft, die der Funktion *auf einer Menge* zukommt. Es ist deshalb im Gegensatz zur Stetigkeit nicht sinnvoll von „gleichmäßiger Stetigkeit in einem Punkt“ zu sprechen.

Wir wollen nun einen Fall behandeln, in dem Stetigkeit und gleichmäßige Stetigkeit tatsächlich zusammenfallen.

Satz 21.4. *Ist $K \subseteq \mathbb{R}$ kompakt und $f \in C(K)$, so ist f gleichmäßig stetig auf K.*

Beweis. Wir nehmen an, f wäre nicht gleichmäßig stetig auf K. Nun müssen wir unsere Gesellenprüfung in elementarer Logik ablegen und die Definition der gleichmäßigen Stetigkeit negieren. Am besten macht man das ganz formal und ohne viel nachzudenken mit den Quantoren. Wir schreiben uns noch einmal hin, was gleichmäßige Stetigkeit bedeutet:

$$\forall \varepsilon > 0\ \exists \delta > 0, \text{ so dass } \forall x, y \in K \text{ mit } |x - y| < \delta \text{ gilt } |f(x) - f(y)| < \varepsilon.$$

Beim Negieren müssen wir aus jedem $\forall$ ein $\exists$ und aus jedem $\exists$ ein $\forall$ machen sowie die Aussage negieren. Das ergibt: f ist auf K nicht gleichmäßig stetig, falls

$$\exists \varepsilon_0 > 0\ \forall \delta > 0\ \exists x = x(\delta) \in K\ \exists y = y(\delta) \in K \text{ mit } |x - y| < \delta,$$
$$\text{so dass } |f(x) - f(y)| \geq \varepsilon_0.$$

Nun machen wir uns klar, was wir da bekommen haben. Nach Annahme gibt es ein $\varepsilon_0 > 0$, so dass für alle $\delta > 0$ etwas gilt. Wir begnügen uns damit, alle δ anzuschauen, die von der Form $1/n$ für ein $n \in \mathbb{N}$ sind. Also gibt es ein $\varepsilon_0 > 0$, so dass für alle $n \in \mathbb{N}$ zwei Zahlen $x_n, y_n \in K$ existieren, für die zum einen

$$|x_n - y_n| < \delta = \frac{1}{n} \quad \text{und zum anderen} \quad |f(x_n) - f(y_n)| \geq \varepsilon_0$$

gilt. Nun ist die Folge (x_n) eine Folge in der kompakten Menge K, d. h. sie ist insbesondere beschränkt. Nach dem Satz von Bolzano-Weierstraß 9.9 besitzt sie damit eine konvergente Teilfolge (x_{n_k}) und es gilt $x_0 := \lim_{k\to\infty} x_{n_k} \in K$, da K abgeschlossen ist. Betrachten wir die Folge (y_{n_k}), so bekommen wir

$$y_{n_k} = x_{n_k} + (y_{n_k} - x_{n_k}).$$

Der erste Summand auf der rechten Seite konvergiert gegen x_0 nach Konstruktion und wegen $|x_n - y_n| < 1/n$ für alle $n \in \mathbb{N}$ konvergiert der zweite gegen Null. Also sind beide Summanden auf der rechten Seite für $k \to \infty$ konvergent, was bedeutet, dass auch die Folge (y_{n_k}) für $k \to \infty$ konvergiert mit $\lim_{k\to\infty} y_{n_k} = x_0$. Da f in x_0 stetig ist, gilt

$$\lim_{k\to\infty} |f(x_{n_k}) - f(y_{n_k})| \leq \lim_{k\to\infty} \big(|f(x_{n_k}) - f(x_0)| + |f(x_0) - f(y_{n_k})|\big) = 0 + 0 = 0,$$

was im Widerspruch zu $|f(x_n) - f(y_n)| \geq \varepsilon_0$ für alle $n \in \mathbb{N}$ steht. Also muss f gleichmäßig stetig in K sein. □

Wir führen noch einen weiteren Stetigkeitsbegriff ein.

Definition 21.5. *Es sei $D \subseteq \mathbb{K}$ und $f : D \to \mathbb{K}$ eine Funktion. Diese heißt* Lipschitz-stetig, *falls eine Konstante $L \geq 0$ existiert, so dass*

$$|f(x) - f(y)| \leq L|x - y|$$

für alle $x, y \in D$ gilt.

Bemerkung 21.6. Man kann sich leicht überlegen, dass der Begriff der Lipschitz-Stetigkeit sogar ein noch stärkerer als der der gleichmäßigen Stetigkeit ist, denn wenn f Lipschitz-stetig und $\varepsilon > 0$ ist, so gilt für jedes $0 < \delta < \varepsilon/L$ sofort

$$|f(x) - f(y)| \leq L|x - y| < L\frac{\varepsilon}{L} = \varepsilon$$

für alle $x, y \in D$ mit $|x - y| < \delta$. Man beachte, dass das δ nur von ε und nicht von x oder y abhängt, also ist f tatsächlich gleichmäßig stetig. Die Umkehrung ist auch hier wieder i.A. falsch, wie das folgende Beispiel zeigt.

Beispiel 21.7. Wir setzen $D = [0, 1]$ und $f(x) = \sqrt{x}$. Dann ist f nach Satz 7.1 stetig und nach Satz 21.4 auch gleichmäßig stetig auf D. Nehmen wir aber an, es gäbe ein $L \geq 0$, so dass für alle $x, y \in D$ gilt

$$|\sqrt{x} - \sqrt{y}| \leq L|x - y|,$$

so folgt für die spezielle Wahl $y = 0$ sofort $\sqrt{x} \leq Lx$, d. h.

$$L \geq \frac{1}{\sqrt{x}} \text{ für alle } x \in (0, 1].$$

Also ist f nicht Lipschitz-stetig.

22. Differenzierbarkeit

Schon aus der Schule werden Sie das Thema dieses Abschnitts kennen. Man möchte das Änderungsverhalten einer Funktion in einem Punkt, d. h. anschaulich gesprochen die Steigung des Funktionsgraphen an dieser Stelle quantitativ fassen. Dazu nähert man die Tangentensteigung mit den bekannten Sekantensteigungen an und kommt auf den Differenzenquotienten. Dessen Grenzwert, die Ableitung, gibt dann die Steigung an. Auch die Differenzierbarkeit einer Funktion ist so im Grunde nichts anderes als ein Grenzwertproblem, das wir mit unseren bisherigen Erkenntnissen behandeln können.
In diesem Abschnitt sei $I \subseteq \mathbb{R}$ immer ein Intervall.

Definition 22.1. *(a) Es sei $x_0 \in I$. Eine Funktion $f : I \to \mathbb{K}$ heißt* differenzierbar *in x_0, wenn der Grenzwert*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

in $\mathbb{K}$ existiert. In diesem Fall heißt dieser Grenzwert die Ableitung von f in x_0 *und wird mit $f'(x_0)$ bezeichnet.*

(b) Eine Funktion $f : I \to \mathbb{K}$ heißt differenzierbar auf I, *falls sie in allen Punkten $x_0 \in I$ differenzierbar ist. In diesem Fall wird durch $x \mapsto f'(x)$ für $x \in I$ eine Funktion $f' : I \to \mathbb{K}$ definiert. Diese Funktion heißt die* Ableitung *oder auch* Ableitungsfunktion von f auf I.

Bemerkung 22.2. Es ist nicht schwer sich klarzumachen, dass der Grenzwert in obiger Definition genau dann existiert, wenn der Grenzwert

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

existiert, und dass dann diese beiden Limites übereinstimmen. Man kann also je nachdem, was in der jeweiligen Situation übersichtlicher erscheint, den einen oder den anderen Grenzwert untersuchen.

Beispiel 22.3. (a) Es sei zunächst $f(x) = c \in \mathbb{K}$ konstant für alle $x \in I$. Dann ist f in I differenzierbar und es gilt $f'(x) = 0$ für alle $x \in I$.

(b) Wir betrachten $I = \mathbb{R}$, $x_0 = 0$ und

$$f(x) = |x|, \quad x \in \mathbb{R}.$$

Dann gilt

$$\frac{f(x)-f(x_0)}{x-x_0}=\frac{|x|}{x}=\begin{cases}1, & \text{für } x>0,\\ -1, & \text{für } x<0.\end{cases}$$

Also existiert der Grenzwert dieses Ausdrucks für $x \to x_0 = 0$ nicht, d. h. f ist in 0 *nicht* differenzierbar. Man beachte, dass f aber in 0 stetig ist.

Wir haben soeben gesehen, dass es stetige Funktionen gibt, die nicht differenzierbar sind. Wir wollen nun zeigen, dass aber umgekehrt jede differenzierbare Funktion notwendigerweise stetig ist.

Satz 22.4. *Es sei $f : I \to \mathbb{K}$ in $x_0 \in I$ differenzierbar. Dann ist f stetig in x_0.*

Beweis. Es gilt

$$\begin{aligned}\lim_{x\to x_0} f(x)-f(x_0) &= \lim_{x\to x_0}\big(f(x)-f(x_0)\big)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}(x-x_0)\\ &= f'(x_0)\cdot 0 = 0.\end{aligned}$$

Damit haben wir $\lim_{x\to x_0} f(x) = f(x_0)$, also ist f in x_0 stetig. □

Wir berechnen beispielhaft noch weitere Ableitungen.

Beispiel 22.5. (a) Es sei $I = \mathbb{R}$ und $n \in \mathbb{N}$. Wir betrachten

$$f(x) = x^n, \quad x \in \mathbb{R}.$$

Dann gilt nach Satz 5.2 (c) für alle $x, x_0 \in \mathbb{R}$ mit $x \neq x_0$

$$\frac{f(x)-f(x_0)}{x-x_0}=\frac{x^n-x_0^n}{x-x_0}=\sum_{k=0}^{n-1} x^k x_0^{n-1-k}.$$

Also ist

$$\begin{aligned}\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} &= \lim_{x\to x_0}\sum_{k=0}^{n-1} x^k x_0^{n-1-k}=\sum_{k=0}^{n-1}\lim_{x\to x_0} x^k x_0^{n-1-k}\\ &= \sum_{k=0}^{n-1} x_0^{n-1} = n x_0^{n-1}.\end{aligned}$$

Damit ist f auf $\mathbb{R}$ differenzierbar und es gilt $(x^n)' = nx^{n-1}$.

(b) Es sei wieder $I = \mathbb{R}$ und jetzt

$$f(x) = E(x) = \mathrm{e}^x, \quad x \in \mathbb{R}.$$

Dann gilt

$$\begin{aligned}\frac{f(x_0+h)-f(x_0)}{h} &= \frac{\mathrm{e}^{x_0+h}-\mathrm{e}^{x_0}}{h} = \frac{\mathrm{e}^{x_0}\mathrm{e}^{h}-\mathrm{e}^{x_0}}{h} \\ &= \mathrm{e}^{x_0}\frac{\mathrm{e}^{h}-1}{h} \longrightarrow \mathrm{e}^{x_0} \ (h\to 0)\end{aligned}$$

mit Hilfe von Beispiel 18.7 (c). Also ist die Exponentialfunktion auf $\mathbb{R}$ differenzierbar und es gilt $E'(x) = \mathrm{e}^x = E(x)$.

Um kompliziertere Ableitungen berechnen zu können, brauchen wir Rechenregeln. Einen ersten Satz wollen wir jetzt beweisen.

Satz 22.6. *Es seien $f, g : I \to \mathbb{K}$ in $x_0 \in I$ differenzierbar und $\alpha, \beta \in \mathbb{K}$. Dann gilt*

(a) $\alpha f + \beta g$ ist in x_0 differenzierbar und

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0). \qquad \text{(Linearität)}$$

(b) fg ist differenzierbar in x_0 und

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0). \qquad \text{(Produktregel)}$$

(c) Ist $g(x_0) \neq 0$, so existiert ein Intervall $J \subseteq I$ mit $x_0 \in J$ und $g(x) \neq 0$ für alle $x \in J$. Außerdem ist die Funktion $f/g : J \to \mathbb{K}$ differenzierbar in x_0 und es gilt

$$\Big(\frac{f}{g}\Big)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{\big(g(x_0)\big)^2}. \qquad \text{(Quotientenregel)}$$

(d) Im Falle $\mathbb{K} = \mathbb{C}$ sind auch $\overline{f} : I \to \mathbb{C}$ sowie $\mathrm{Re} f, \mathrm{Im} f : I \to \mathbb{R}$ differenzierbar in x_0 mit $(\overline{f})'(x_0) = \overline{f'(x_0)}$ sowie $(\mathrm{Re} f)'(x_0) = \mathrm{Re}(f'(x_0))$ und $(\mathrm{Im} f)'(x_0) = \mathrm{Im}(f'(x_0))$.

Beweis. Die Aussagen (a) und (b) behandeln wir als Übungsaufgaben.
Zum Beweis von (c) müssen wir zuerst die Existenz von J begründen. Da g in x_0 stetig ist, gibt es ein $\delta > 0$, so dass $|g(x) - g(x_0)| < |g(x_0)|/2$ für alle $x \in (x_0 - \delta, x_0 + \delta) =: J$ gilt. Für diese x ist dann $|g(x)| > |g(x_0)|/2 > 0$, also insbesondere $g(x) \neq 0$.
Weiter gilt

$$\begin{aligned}\frac{\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}}{x - x_0} &= \frac{1}{g(x)g(x_0)} \cdot \frac{f(x)g(x_0) - f(x_0)g(x)}{x - x_0} \\ &= \frac{1}{g(x)g(x_0)} \cdot \frac{f(x)g(x_0) - f(x_0)g(x_0) + f(x_0)g(x_0) - f(x_0)g(x)}{x - x_0} \\ &= \frac{1}{g(x)g(x_0)}\Big(\frac{f(x) - f(x_0)}{x - x_0}g(x_0) - \frac{g(x) - g(x_0)}{x - x_0}f(x_0)\Big).\end{aligned}$$

Da g in x_0 differenzierbar ist, ist diese Funktion insbesondere in x_0 stetig (vgl. Satz 22.4), also können wir in obiger Gleichung zum Grenzwert $x \to x_0$ übergehen und erhalten die Behauptung.
Es bleibt noch (d) zu zeigen. Nach den Rechenregeln für die komplexe Konjugation, vgl. Satz 16.6, und Dank Übungsaufgabe 18.4 (c), gilt für alle $x_0 \in I$

$$\begin{aligned}(\overline{f})'(x_0) &= \lim_{x\to x_0} \frac{\overline{f(x)} - \overline{f(x_0)}}{x - x_0} = \lim_{x\to x_0} \frac{\overline{f(x) - f(x_0)}}{\overline{x - x_0}} = \lim_{x\to x_0} \overline{\left(\frac{f(x) - f(x_0)}{x - x_0}\right)} \\ &= \overline{\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}} = \overline{f'(x_0)}.\end{aligned}$$

Die Differenzierbarkeit von $\mathrm{Re} f$ und $\mathrm{Im} f$ folgt jetzt aus $\mathrm{Re} f = (f + \overline{f})/2$ und $\mathrm{Im} f = (f - \overline{f})/2$, in Zusammenarbeit mit (a). □

Es folgt sogleich die Rechenregel für die Verkettung differenzierbarer Funktionen.

Satz 22.7 (Kettenregel)**.** *Es seien $I, J \subseteq \mathbb{R}$ Intervalle und $g : I \to \mathbb{R}$ sei differenzierbar in $x_0 \in I$. Weiter gelte $g(I) \subseteq J$ und die Funktion $f : J \to \mathbb{K}$ sei differenzierbar in $y_0 = g(x_0)$. Dann ist auch die Funktion $f \circ g : I \to \mathbb{K}$ differenzierbar in x_0 und es gilt*

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

Beweis. Wir betrachten die Hilfsfunktion $\tilde{f} : J \to \mathbb{K}$ mit

$$\tilde{f}(y) = \begin{cases} \dfrac{f(y) - f(y_0)}{y - y_0}, & \text{für } y \in J \text{ mit } y \neq y_0, \\ f'(y_0), & \text{für } y = y_0. \end{cases}$$

Dann gilt

$$\tilde{f}(y)(y - y_0) = f(y) - f(y_0) \tag{22.1}$$

für alle $y \in J$ (auch für y_0!). Da f in y_0 differenzierbar ist, haben wir nun

$$\lim_{y\to y_0} \tilde{f}(y) = \tilde{f}(y_0) = f'(y_0) = f'(g(x_0)),$$

insbesondere ist $\tilde{f}$ stetig in y_0. Nach Satz 22.4 ist g stetig in x_0, und da die Verkettung von stetigen Funktionen wieder stetig ist, sehen wir damit

$$\lim_{x\to x_0} \tilde{f}(g(x)) = \tilde{f}(g(x_0)) = f'(g(x_0)).$$

Daher folgt schließlich mit Hilfe von (22.1)

$$\begin{aligned}\frac{f(g(x)) - f(g(x_0))}{x - x_0} &= \frac{\tilde{f}(g(x))\big(g(x) - g(x_0)\big)}{x - x_0} \\ &= \tilde{f}(g(x))\frac{g(x) - g(x_0)}{x - x_0} \longrightarrow f'(g(x_0)) \cdot g'(x_0) \ (x \to x_0).\end{aligned}$$ □

Beispiel 22.8. Es sei $a > 0$ gegeben. Dann betrachten wir auf $I = \mathbb{R}$ die Funktion

$$\varphi(x) := a^x, \quad x \in \mathbb{R}.$$

Dann ist nach Definition $\varphi(x) = \mathrm{e}^{x \ln a}$. Um die Kettenregel anzuwenden setzen wir $f(y) := \mathrm{e}^y$ und $g(x) := x \ln(a)$. Dann ist $\varphi = f \circ g$. Da sowohl f als auch g auf ganz $\mathbb{R}$ differenzierbar sind und f auf ganz $\mathbb{R}$ definiert ist, sind die Voraussetzungen von Satz 22.7 erfüllt und es gilt

$$(a^x)' = f'(g(x))g'(x) = \mathrm{e}^{g(x)} \ln(a) = \mathrm{e}^{x \ln(a)} \ln(a) = a^x \ln(a).$$

Wir können sogar eine allgemeine Rechenregel für die Ableitung der Umkehrfunktion angeben.

Satz 22.9. *Es sei $f : I \to \mathbb{R}$ stetig, streng monoton und in $x_0 \in I$ differenzierbar mit $f'(x_0) \neq 0$. Dann existiert die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$, diese ist differenzierbar in $y_0 = f(x_0)$ und es gilt*

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

Beweis. Die Existenz der Umkehrfunktion folgt sofort aus der strengen Monotonie von f.
Zu gegebenem $h \neq 0$ setzen wir

$$k := f^{-1}(y_0 + h) - f^{-1}(y_0) = f^{-1}(y_0 + h) - x_0.$$

Da f^{-1} nach Satz 19.11 in y_0 stetig ist, folgt aus $h \to 0$ sofort $k \to 0$. Außerdem ist $x_0 + k = f^{-1}(y_0 + h)$, d. h. $f(x_0 + k) = y_0 + h$ und wir erhalten

$$h = f(x_0 + k) - f(x_0).$$

Nun ist damit

$$\begin{aligned}\frac{f^{-1}(y_0 + h) - f^{-1}(y_0)}{h} &= \frac{f^{-1}\big(f(x_0 + k)\big) - x_0}{f(x_0 + k) - f(x_0)} = \frac{x_0 + k - x_0}{f(x_0 + k) - f(x_0)} \\ &= \frac{k}{f(x_0 + k) - f(x_0)} \longrightarrow \frac{1}{f'(x_0)} \quad (h \to 0),\end{aligned}$$

da $f'(x_0) \neq 0$ gilt. □

Bemerkung 22.10. Man beachte, dass die Voraussetzung $f'(x_0) \neq 0$ notwendig ist. Als Beispiel diene hierzu $I = [0, \infty)$ und $f(x) = x^2$. Dann ist $f'(x) = 2x$ und somit $f'(0) = 0$. Tatsächlich ist die Umkehrfunktion $f^{-1}(x) = \sqrt{x}$ in $x_0 = 0$ nicht differenzierbar, denn es gilt

$$\frac{\sqrt{x} - \sqrt{0}}{x - 0} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}} \longrightarrow \infty \quad (x \to 0+).$$

Beispiel 22.11. (a) Wir bestimmen die Ableitung des Logarithmus als Umkehrfunktion der Exponentialfunktion. Sei dazu $I = \mathbb{R}$ und $f(x) = \mathrm{e}^x$ auf I. Dann ist $f^{-1}(x) = \ln(x)$ für alle $x \in (0, \infty)$ und mit Satz 22.9 gilt für $y = f(x)$ die Beziehung

$$(\ln)'(y) = (f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{\mathrm{e}^x} = \frac{1}{\mathrm{e}^{\ln(y)}} = \frac{1}{y}, \quad y \in (0, \infty).$$

(b) Damit bekommen wir dank der Kettenregel für jede auf I differenzierbare Funktion $f : I \to \mathbb{R}$, die $f(x) > 0$ für alle $x \in I$ erfüllt,

$$(\ln \circ f)'(x) = \frac{1}{f(x)} f'(x) = \frac{f'(x)}{f(x)}, \quad x \in I.$$

Man nennt das die *logarithmische Ableitung* von f.

(c) Für $x > 0$, $\alpha \in \mathbb{R}$ und $f(x) := x^\alpha = \mathrm{e}^{\alpha \ln(x)}$ erhalten wir

$$f'(x) = \mathrm{e}^{\alpha \ln(x)} (\alpha \ln(x))' = x^\alpha \frac{\alpha}{x} = \alpha x^{\alpha - 1}.$$

Die Ableitungsregel für die ganzzahlige Potenz aus Beispiel 22.5 (a) verallgemeinert sich also auch auf die allgemeine Potenz, solange $x > 0$.

Insbesondere haben wir im Fall $\alpha = 1/2$

$$(\sqrt{\cdot})'(x) = \frac{1}{2\sqrt{x}}, \quad x > 0.$$

Wir kommen nun zu dem Satz, der im Zusammenhang mit Ableitungen in den verschiedensten Wissenschaften wahrscheinlich am häufigsten verwendet wird. Er ermöglicht die Bestimmung von Maximal- und Minimalstellen einer Funktion. Wir definieren zunächst genau was wir damit meinen.

Definition 22.12. *Es sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion.*

(a) Man sagt, dass f in $x_0 \in D$ ein globales Maximum *(bzw.* globales Minimum*) hat, falls $f(x) \le f(x_0)$ (bzw. $f(x) \ge f(x_0)$) für alle $x \in D$ gilt.*

(b) f hat in $x_0 \in D$ ein relatives Maximum *(bzw.* relatives Minimum*), falls ein $\delta > 0$ existiert, so dass $f(x) \le f(x_0)$ (bzw. $f(x) \ge f(x_0)$) für alle $x \in D \cap U_\delta(x_0)$ gilt.*

(c) Allgemein spricht man von einem globalen *bzw.* relativen Extremum *in x_0, wenn f dort ein entsprechendes Maximum oder Minimum hat.*

Bemerkung 22.13. Statt „relatives“ Extremum/Maximum/Minimum ist auch die Bezeichnung *lokales Extremum/Maximum/Minimum* üblich.

Satz 22.14. *Es sei $f : I \to \mathbb{R}$ differenzierbar in $x_0 \in I$. Ist x_0 kein Randpunkt von I und hat f in x_0 ein relatives Extremum, so gilt $f'(x_0) = 0$.*

Warnung 22.15. Da dieser Satz so oft verwendet wird, wird er auch gerne falsch verwendet. Darum hier (aus vielfach gegebenem Anlass) zwei Warnungen.

(a) Die Voraussetzung, dass x_0 kein Randpunkt ist, ist wesentlich. Ein einfaches Beispiel ist die Funktion $f(x) = x$ auf dem Intervall $[0, 1]$. Diese hat ein relatives Minimum in $x_0 = 0$, aber $f'(0) = 1$.

(b) Die Umkehrung gilt nicht! Das sieht man sofort an dem Beispiel $f(x) = x^3$ auf $I = \mathbb{R}$. Dann ist nämlich $f'(x) = 3x^2$, also $f'(0) = 0$, aber diese Funktion hat in 0 kein Extremum, denn in jeder Umgebung $U_\varepsilon(0)$ für $\varepsilon > 0$ liegen Punkte mit $f(x) > 0 = f(0)$, z. B. $x = \varepsilon/2$, und mit $f(x) < 0 = f(0)$, z. B. $x = -\varepsilon/2$.

Beweis von Satz 22.14. Wir gehen zunächst davon aus, dass f in x_0 ein relatives Maximum hat. Dann existiert ein $\delta > 0$, so dass gleichzeitig $U_\delta(x_0) \subseteq I$ und $f(x) \leq f(x_0)$ für alle $x \in U_\delta(x_0)$ gilt. Die erste Bedingung können wir erfüllen, weil x_0 kein Randpunkt von I ist, die zweite ist genau die Definition des relativen Maximums. Sei nun $x \in U_\delta(x_0)$ aber $x \neq x_0$. Dann ist

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \leq 0, \text{ falls } x > x_0, \\ \geq 0, \text{ falls } x < x_0. \end{cases}$$

Da f außerdem in x_0 differenzierbar ist, muss damit gelten

$$f'(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0 \quad \text{und} \quad f'(x_0) = \lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

Also ist $f'(x_0) = 0$.
Wir widmen uns nun dem Fall, dass f ein relatives Minimum in x_0 hat. Dann hat die Funktion $-f$ in x_0 ein relatives Maximum, denn es gibt ein $\delta > 0$, so dass für alle $x \in U_\delta(x_0)$ die Ungleichung $f(x) \geq f(x_0)$ erfüllt ist. Also ist für alle diese x auch $-f(x) \leq -f(x_0)$. Nach dem ersten Teil des Beweises gilt also $f'(x_0) = -(-f)'(x_0) = -0 = 0$. □

Satz 22.16 (Mittelwertsatz der Differenzialrechnung)**.** *Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to \mathbb{R}$ sei stetig auf $[a, b]$ und differenzierbar in (a, b). Dann gibt es ein $\xi \in (a, b)$, so dass*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi), \quad \textit{bzw. gleichbedeutend} \quad f(b) - f(a) = f'(\xi)(b - a)$$

gilt.

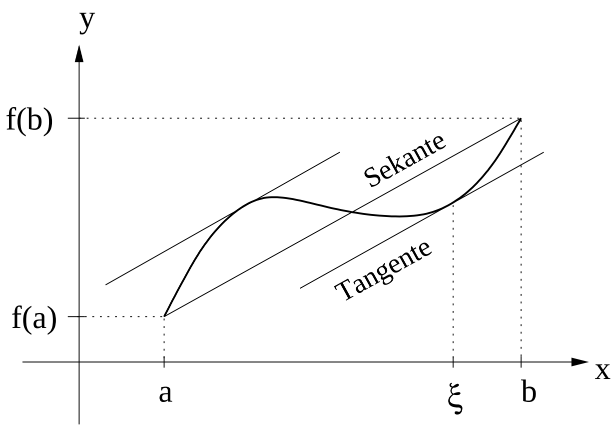

Abbildung 22.1.: Der Mittelwertsatz der Differenzialrechnung

Anschaulich bedeutet dieser Satz, dass die Sekantensteigung der Funktion, die man anhand der beiden Punkte a und b erhält, irgendwann dazwischen tatsächlich als Tangentensteigung angenommen wird, vgl. Abbildung 22.1. Man kann sich das verdeutlichen, indem man versucht, eine differenzierbare Funktion zu zeichnen, für die das nicht gilt, was (hoffentlich) nicht klappen wird.

Beweis. Wir betrachten die Hilfsfunktion

$$g(x) := f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a), \quad x \in [a, b].$$

Dann ist offensichtlich auch $g \in C([a, b])$ und g ist differenzierbar auf (a, b) mit

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}, \quad x \in (a, b).$$

Außerdem ist $g(a) = g(b) = 0$. Können wir nun zeigen, dass es ein $\xi \in (a, b)$ gibt, für das $g'(\xi) = 0$ gilt, so haben wir damit $f'(\xi) = (f(b) - f(a))/(b - a)$ und sind fertig.
Wir beobachten zunächst, dass im Fall $g(x) = 0$ für alle $x \in [a, b]$ nichts mehr zu tun ist, denn dann ist insbesondere $g'(x) = 0$ für alle $x \in (a, b)$. Es sei also g nicht konstant. Da g eine stetige Funktion auf der kompakten Menge $[a, b]$ ist, gibt es nach Satz 19.7 Zahlen $t, s \in [a, b]$, so dass

$$g(t) \leq g(x) \leq g(s) \quad \text{für alle } x \in [a, b]$$

gilt. Wäre nun sowohl $t \in \{a, b\}$, als auch $s \in \{a, b\}$, so wäre $g(s) = g(t) = 0$ und damit wieder $g(x) = 0$ für alle $x \in [a, b]$, was wir gerade ausgeschlossen haben. Es gilt also $t \in (a, b)$ oder $s \in (a, b)$, d. h. eins der beiden ist ein innerer Punkt von $[a, b]$. Weiterhin hat g dort ein relatives Extremum. Also gilt nach Satz 22.14 $g'(t) = 0$ oder $g'(s) = 0$ und der Beweis ist beendet. $\square$

Warnung 22.17. Die Verallgemeinerung des Mittelwertsatzes auf komplexwertige Funktionen ist im Allgemeinen falsch! Hier ist Vorsicht angebracht, auch im Hinblick auf alles, was im Folgenden mit Hilfe des Mittelwertsatzes bewiesen wird. Manchmal kann man sich allerdings durch die getrennte Behandlung von Real- und Imaginärteil einer komplexwertigen Funktion behelfen, vgl. den Beweis von Satz 22.18 (c).

Wir wollen nun einige Folgerungen aus dem Mittelwertsatz ziehen.

Satz 22.18. *(a)* (**Satz von Rolle**) *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to \mathbb{R}$ sei stetig. Ist f auf (a, b) differenzierbar und gilt $f(a) = f(b)$, so gibt es ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.*

(b) Es sei $f : I \to \mathbb{R}$ auf dem Intervall I differenzierbar. Dann gilt

$$\begin{aligned}
&\textit{Ist } f' = 0 \textit{ auf } I, \textit{ so ist } f \textit{ auf } I \textit{ konstant.}\\
&\textit{Ist } f' > 0 \textit{ auf } I, \textit{ so ist } f \textit{ auf } I \textit{ streng monoton wachsend.}\\
&\textit{Ist } f' < 0 \textit{ auf } I, \textit{ so ist } f \textit{ auf } I \textit{ streng monoton fallend.}\\
&\textit{Ist } f' \geq 0 \textit{ auf } I, \textit{ so ist } f \textit{ auf } I \textit{ monoton wachsend.}\\
&\textit{Ist } f' \leq 0 \textit{ auf } I, \textit{ so ist } f \textit{ auf } I \textit{ monoton fallend.}
\end{aligned}$$

(c) Sind $f, g : I \to \mathbb{K}$ auf I differenzierbare Funktionen und gilt $f' = g'$ auf I, so gibt es eine Konstante $c \in \mathbb{K}$, so dass $f(x) = g(x) + c$ für alle $x \in I$ gilt.

Beweis. (a) folgt direkt aus dem Mittelwertsatz.

(b) Es seien $a, b \in I$ mit $a < b$. Dann gibt es nach dem Mittelwertsatz ein $\xi \in (a, b)$ mit $f(b) - f(a) = (b - a)f'(\xi)$. Ist die Ableitung von f nun konstant Null auf I, so muss also $f(a) = f(b)$ gelten. Da a und b in I beliebig waren, ist f auf I konstant,

Weiter ist der Ausdruck $b - a$ immer positiv, also ergibt sich das Vorzeichen von $f(b) - f(a)$ direkt aus dem Vorzeichen von $f'(\xi)$. Daraus kann man die 4 restlichen Behauptungen sofort ablesen.

(c) Mit $f' = g'$ gelten auch $(\operatorname{Re} f)' = \operatorname{Re} f' = \operatorname{Re} g' = (\operatorname{Re} g)'$ und $(\operatorname{Im} f)' = \operatorname{Im} f' = \operatorname{Im} g' = (\operatorname{Im} g)'$, vgl. Satz 22.6 (d). Wir setzen $h := \operatorname{Re} f - \operatorname{Re} g$. Dann ist $h' = \operatorname{Re} f' - \operatorname{Re} g' = 0$ auf I, d. h. h ist konstant nach (b). Genauso erhält man, dass $k = \operatorname{Im} f - \operatorname{Im} g$ konstant ist. Setzt man $c = h(x_0) + \mathrm{i}k(x_0)$ für ein $x_0 \in I$, so gilt damit für alle $x \in I$

$$\begin{aligned}
f(x) &= \operatorname{Re} f(x) + \mathrm{i}\operatorname{Im} f(x) = \operatorname{Re} g(x) + h(x) + \mathrm{i}\big(\operatorname{Im} g(x) + k(x)\big)\\
&= g(x) + h(x) + \mathrm{i}k(x) = g(x) + h(x_0) + \mathrm{i}k(x_0) = g(x) + c. \qquad \square
\end{aligned}$$

Satz 22.19 (verallgemeinerter Mittelwertsatz)**.** *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f, g : [a, b] \to \mathbb{R}$ seien stetig sowie differenzierbar auf (a, b). Ist $g'(x) \neq 0$ für alle $x \in (a, b)$, so gibt es ein $\xi \in (a, b)$ mit*

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

Beweis. Wir betrachten die Hilfsfunktion

$$h(x) = \big(f(b) - f(a)\big)g(x) - \big(g(b) - g(a)\big)f(x).$$

Dann ist $h \in C([a, b])$, differenzierbar in (a, b) und es gilt

$$\begin{aligned} h(a) &= f(b)g(a) - f(a)g(a) - g(b)f(a) + g(a)f(a) \\ &= f(b)g(b) - f(a)g(b) - g(b)f(b) + g(a)f(b) = h(b). \end{aligned}$$

Also gibt es nach dem Satz von Rolle ein $\xi \in (a, b)$ mit

$$0 = h'(\xi) = \big(f(b) - f(a)\big)g'(\xi) - \big(g(b) - g(a)\big)f'(\xi),$$

woraus wegen $g'(x) \neq 0$ für alle $x \in I$ die Behauptung folgt. $\square$

Die Differenzierbarkeit gibt uns unter Anderem ein starkes Hilfsmittel zur Bestimmung von Grenzwerten bei Quotienten von Funktionen in die Hand, das wir zum Abschluss dieses Abschnitts beweisen wollen.

Satz 22.20 (Satz von de l'Hospital)**.** *Es sei (a, b) ein offenes Intervall in $\mathbb{R}$ (dabei ist hier $a = -\infty$ oder $b = \infty$ zugelassen) und $f, g : (a, b) \to \mathbb{R}$ seien differenzierbar auf (a, b) mit $g'(x) \neq 0$ für alle $x \in (a, b)$. Gilt dann*

(I) $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$ *oder*

(II) $\lim_{x\to a} g(x) = \pm\infty$

und existiert der Grenzwert

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

(hierbei ist wieder $L = \pm\infty$ zugelassen), dann gilt

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

Die Aussage dieses Satzes bleibt richtig, wenn man überall a durch b ersetzt.

Warnung 22.21. Dieser Satz hat viele Voraussetzungen und diese sind wirklich alle nötig! Im Eifer des Gefechts gegen einen hartnäckigen Grenzwert wird hier gerne die eine oder andere vergessen. Besonderer Beliebtheit erfreut es sich, nicht nachzuprüfen, ob es sich wirklich um einen sogenannten „uneigentlichen" Grenzwert der Form $0/0$ oder $\pm\infty/\pm\infty$ handelt. Nur solche kann dieser Satz behandeln!

Beweis von Satz 22.20. Wir betrachten den Fall (I) und gehen zunächst davon aus, dass $a \in \mathbb{R}$ gilt. Dann setzen wir f und g durch $f(a) := g(a) := 0$ stetig fort und erhalten so $f, g \in C([a, b])$, denn die Funktionen sind ja auf (a, b) differenzierbar und damit insbesondere stetig. Ist nun $x \in (a, b)$, so gibt es nach dem verallgemeinerten Mittelwertsatz (Satz 22.19) ein $\xi = \xi(x) \in (a, x)$ mit

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

Strebt nun $x \to a$, so muss zwangsläufig auch $\xi \to a$ gehen, also folgt

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

Wir können uns also dem Fall $a = -\infty$ zuwenden. Dann substituieren wir $t = 1/x$. Das führt dazu, dass der Grenzübergang $x \to a = -\infty$ in $t \to 0-$ übergeht. Wir setzen

$$\tilde{f}(t) = f(1/t) \quad \text{und} \quad \tilde{g}(t) = g(1/t).$$

Dann gilt

$$\tilde{f}'(t) = f'(1/t)\Big(-\frac{1}{t^2}\Big) \quad \text{und} \quad \tilde{g}'(t) = g'(1/t)\Big(-\frac{1}{t^2}\Big)$$

und somit gilt auch

$$\lim_{t\to 0-} \frac{\tilde{f}'(t)}{\tilde{g}'(t)} = \lim_{t\to 0-} \frac{f'(1/t)(-t^2)}{g'(1/t)(-t^2)} = \lim_{t\to 0-} \frac{f'(1/t)}{g'(1/t)} = \lim_{x\to -\infty} \frac{f'(x)}{g'(x)} = L.$$

Wir können also das oben schon Bewiesene anwenden und erhalten

$$\lim_{x\to -\infty} \frac{f(x)}{g(x)} = \lim_{t\to 0-} \frac{f(1/t)}{g(1/t)} = \lim_{t\to 0-} \frac{\tilde{f}(t)}{\tilde{g}(t)} = \lim_{t\to 0-} \frac{\tilde{f}'(t)}{\tilde{g}'(t)} = L.$$

Der Fall (II) und die Ersetzung von a durch b bleiben als Übungsaufgabe stehen. □

Beispiel 22.22. (a) Es seien $a, b > 0$. Wir wollen den Grenzwert

$$\lim_{x\to 0+} \frac{a^x - b^x}{x}$$

untersuchen. Wir betrachten das Intervall $(0,1)$ und die Funktionen $f(x) = a^x - b^x$ und $g(x) = x$ auf $(0,1)$. Diese sind dort beide differenzierbar und es gilt $g'(x) = 1 \neq 0$ für alle $x \in (0,1)$. Außerdem ist

$$\lim_{x\to 0+} f(x) = \lim_{x\to 0+} (a^x - b^x) = \lim_{x\to 0+} a^x - \lim_{x\to 0+} b^x = 1 - 1 = 0 = \lim_{x\to 0+} g(x).$$

Wir können also den Satz von de l'Hospital anwenden und erhalten

$$\lim_{x\to 0+} \frac{a^x - b^x}{x} = \lim_{x\to 0+} \frac{a^x \ln(a) - b^x \ln(b)}{1} = \ln(a) - \ln(b),$$

was wohl nur sehr schwer zu erraten gewesen wäre.

(b) Ebenso kann man zeigen:

$$\lim_{x\to\infty} \frac{\ln(x)}{x} = \lim_{x\to\infty} \frac{\frac{1}{x}}{1} = 0.$$

In diesem Fall hat man es mit einem uneigentlichen Grenzwert der Form ∞/∞ zu tun.

(c) Eine kleine Umformung führt dazu, dass man mit der Regel von de l'Hospital auch Grenzwerte der Form $0\cdot\infty$ behandeln kann. Das geht exemplarisch so:

$$\lim_{x\to 0+} x\ln(x) = \lim_{x\to 0+} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x\to 0+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x\to 0+} (-x) = 0.$$

Man beachte, dass hier u.a. wegen $\lim_{x\to 0+} \ln(x) = -\infty$ und $\lim_{x\to 0+} 1/x = \infty$ die Anwendung des Satzes gerechtfertigt war.

Dieser Grenzwert ermöglicht uns nun zusammen mit der Stetigkeit der Exponentialfunktion noch die Berechnung von

$$\lim_{x\to 0+} x^x = \lim_{x\to 0+} \mathrm{e}^{x\ln(x)} = \mathrm{e}^{\lim_{x\to 0+} x\ln(x)} = \mathrm{e}^0 = 1.$$

23. Ableitung von Potenzreihen und trigonometrische Funktionen

Wir wollen uns in diesem Abschnitt zunächst mit der Ableitung von Funktionen, die durch Potenzreihen gegeben sind, beschäftigen. Wie nicht anders zu erwarten stellt sich heraus, dass diese wieder besonders schön sind.

Satz 23.1. *Es sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit positivem Konvergenzradius r und wir setzen $I := (-r, r)$. Dann gilt:*

(a) Die Potenzreihe

$$g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

hat ebenfalls den Konvergenzradius r.

(b) Die Funktion f ist auf I differenzierbar und es gilt

$$f'(x) = g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \text{ für alle } x \in I.$$

Beweis. Da

$$\limsup_{n\to\infty} \sqrt[n]{|n a_n|} = \limsup_{n\to\infty} \left(\sqrt[n]{n} \sqrt[n]{|a_n|} \right) = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$$

gilt, folgt die Aussage in (a) sofort aus dem Satz von Hadamard.
Zum Beweis der Aussage in (b) bezeichnen wir für jedes $n \in \mathbb{N}$ mit $s_n := \sum_{k=0}^{n} a_k x^k$ die n-te Partialsumme und mit $R_n := \sum_{k=n+1}^{\infty} a_k x^k$ den n-ten Reihenrest. Nun sei $x_0 \in I$ fest. Dann wählen wir ein $0 < \varrho < r$, so dass $x_0 \in [-\varrho, \varrho] \subseteq I$ gilt. Es folgt für alle $x \in [-\varrho, \varrho]$ mit $x \neq x_0$ und alle $n \in \mathbb{N}$

$$\begin{aligned} & \Big| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \Big| \\ = & \Big| \frac{s_n(x) + R_n(x) - s_n(x_0) - R_n(x_0)}{x - x_0} - s_n'(x_0) + s_n'(x_0) - g(x_0) \Big| \\ \leq & \Big| \frac{s_n(x) - s_n(x_0)}{x - x_0} - s_n'(x_0) \Big| + \Big| \frac{R_n(x) - R_n(x_0)}{x - x_0} \Big| + |s_n'(x_0) - g(x_0)|. \qquad (23.1) \end{aligned}$$

Sei nun $\varepsilon > 0$. Unser Ziel ist es, abhängig von ε ein $\delta > 0$ zu finden, so dass für alle $x \in U_\delta(x_0)$ der Ausdruck auf der linken Seite der oberen Ungleichung kleiner als ε wird. Dazu untersuchen wir die 3 Summanden auf der rechten Seite jeweils einzeln und versuchen sie jeweils kleiner als $\varepsilon/3$ abzuschätzen.
Zunächst gilt

$$s_n'(x) = \sum_{k=1}^{n} k a_k x^{k-1},$$

also gilt nach (a) $\lim_{n\to\infty} s_n'(x) = g(x)$ für alle $x \in I$. Insbesondere gibt es also ein $n_1 \in \mathbb{N}$, so dass

$$|s_n'(x_0) - g(x_0)| < \frac{\varepsilon}{3}$$

für alle $n \geq n_1$ gilt.
Für den zweiten Summanden gilt nach Satz 5.2 (c)

$$\begin{aligned}
\Big|\frac{R_n(x) - R_n(x_0)}{x - x_0}\Big| &= \Big|\sum_{k=n+1}^{\infty} \frac{a_k(x^k - x_0^k)}{x - x_0}\Big| \leq \sum_{k=n+1}^{\infty} |a_k| \Big|\frac{x^k - x_0^k}{x - x_0}\Big| \\
&= \sum_{k=n+1}^{\infty} |a_k| |x^{k-1} + x^{k-2}x_0 + \cdots + x x_0^{k-2} + x_0^{k-1}| \\
&\leq \sum_{k=n+1}^{\infty} |a_k| \big(|x^{k-1}| + |x^{k-2}||x_0| + \cdots + |x||x_0^{k-2}| + |x_0^{k-1}|\big) \\
&\leq \sum_{k=n+1}^{\infty} |a_k| \big(\varrho^{k-1} + \varrho^{k-2}\varrho + \cdots + \varrho\varrho^{k-2} + \varrho^{k-1}\big) \\
&= \sum_{k=n+1}^{\infty} k|a_k|\varrho^{k-1} =: c_n.
\end{aligned}$$

Wir haben in (a) gesehen, dass die Potenzreihe $\sum_{n=1}^{\infty} n|a_n|x^{n-1}$ auch den Konvergenzradius r hat. Da $\varrho \in (-r, r)$ gilt, kovergiert die Reihe $\sum_{n=1}^{\infty} n|a_n|\varrho^{n-1}$ und damit folgt $\lim_{n\to\infty} c_n = 0$, da es sich dabei um die Folge der Reihenreste handelt, vgl. Satz 11.6 (b). Also können wir ein $n_2 \in \mathbb{N}$ wählen, so dass

$$\Big|\frac{R_n(x) - R_n(x_0)}{x - x_0}\Big| < \frac{\varepsilon}{3}$$

für alle $n \geq n_2$ gilt. Sei nun $n_0 = \max\{n_1, n_2\}$. Da s_n für jedes $n \in \mathbb{N}$ als Polynom differenzierbar ist, gibt es nun ein $\delta > 0$, so dass für alle $x \in U_\delta(x_0)$ mit $x \neq x_0$ gilt

$$\Big|\frac{s_{n_0}(x) - s_{n_0}(x_0)}{x - x_0} - s_{n_0}'(x_0)\Big| < \frac{\varepsilon}{3}.$$

Wiederholen wir die Abschätzung aus (23.1) für $n = n_0$, und nehmen wir das soeben gewählte δ, so ist jeder der 3 Summanden auf der rechten Seite von (23.1) kleiner als $\varepsilon/3$ und die Behauptung ist bewiesen. $\square$

Die besondere Stärke dieses Satzes besteht darin, dass er uns nicht nur abstrakt die Differenzierbarkeit einer durch eine Potenzreihe gegebenen Funktion sichert, sondern dass er uns auch gleich sagt, wie wir diese, oder wenigstens eine Potenzreihe dieser Ableitungsfunktion, bekommen können: Nämlich auf die denkbar einfachste Weise, wir dürfen jeden einzelnen Summanden „unter dem Summenzeichen“ differenzieren. Da sowohl die Summation als auch die Differenziation einen Grenzübergang darstellen, haben wir hier also wieder ein Beispiel für einen Satz, der das Vertauschen zweier Grenzprozesse gestattet.

Einen kreativen Einsatz dieses Satzes zeigt das folgende Beispiel.

Beispiel 23.2. Wir betrachten die durch eine Potenzreihe gegebene Funktion

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

Wie man leicht nachrechnet, hat diese den Konvergenzradius 1. Für $|x| < 1$ gilt nun nach dem obigen Satz

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

und (welch Glücksfall) diese Reihe ist eine geometrische, wir können also den Reihenwert angeben und erhalten

$$f'(x) = \frac{1}{1+x} = \big(\ln(1+x)\big)'.$$

Nach Satz 22.18 (b) gibt es also ein $c \in \mathbb{R}$, so dass $f(x) = \ln(1+x) + c$ für alle $x \in (-1,1)$ gilt. Nun gilt aber $f(0) = 0$, genauso wie $\ln(1+0) = \ln(1) = 0$ ist, also muss $c = 0$ sein. Das liefert

$$\ln(1+x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1}, \quad x \in (-1,1).$$

Wir konnten also mit Hilfe des obigen Satzes eine Potenzreihendarstellung einer Funktion angeben, für die wir bis jetzt keine solche hatten, bzw. andersherum formuliert, hat obiger Satz uns geholfen, einen zunächst schwierig aussehenden Reihenwert zu berechnen.

Wir hatten im Abschnitt 14 über Potenzreihen die trigonometrischen Funktionen Sinus und Cosinus definiert als

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad \text{und} \quad \cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Mit unserem Satz können wir nun die Ableitungen dieser Funktionen bestimmen und erhalten

$$\sin'(x) = \sum_{n=0}^{\infty}(-1)^n(2n+1)\frac{x^{2n}}{(2n+1)!} = \sum_{n=0}^{\infty}(-1)^n\frac{x^{2n}}{(2n)!} = \cos(x).$$

Genauso berechnet man die Ableitung des Cosinus und erhält zusammengefasst das folgende Resultat.

Satz 23.3. *Sinus und Cosinus sind auf* $\mathbb{R}$ *differenzierbar und es gilt*

$$\sin'(x) = \cos(x) \quad \textit{und} \quad \cos'(x) = -\sin(x), \qquad x \in \mathbb{R}.$$

Wir wollen noch weitere Eigenschaften von Cosinus und Sinus beweisen.

Satz 23.4. *Für alle* $x, y \in \mathbb{R}$ *gelten die folgenden Aussagen:*

(a) *Es ist* $|\sin(x)| \leq 1$ *und* $|\cos(x)| \leq 1$ *und es gilt der* trigonometrische Pythagoras
$$\sin^2(x) + \cos^2(x) = 1 \quad \textit{für alle } x \in \mathbb{R}.$$

(b) $\sin(-x) = -\sin(x) \quad \textit{und} \quad \cos(-x) = \cos(x).$

(c) Additionstheoreme*:*

$$\sin(x+y) = \sin(x)\cos(y) + \sin(y)\cos(x),$$
$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y).$$

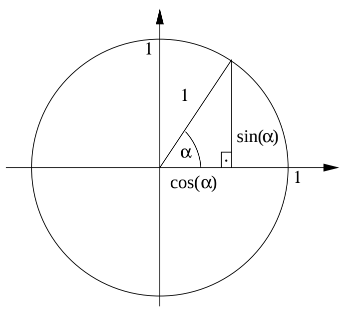

Abbildung 23.1.: Trigonometrischer Pythagoras

Beweis. (a) Wir setzen $g(x) := \sin^2(x) + \cos^2(x)$ für $x \in \mathbb{R}$. Dann gilt

$$g'(x) = 2\sin(x)\cos(x) + 2\cos(x)\big(-\sin(x)\big) = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Also ist g wieder eine konstante Funktion und da $g(0) = \sin^2(0) + \cos^2(0) = 0 + 1 = 1$ ist, gilt $g(x) = 1$ für alle $x \in \mathbb{R}$.

(b) Diese Beziehungen folgen direkt aus der Potenzreihendarstellung, denn in der Potenzreihe des Sinus tauchen nur ungerade x-Potenzen und in der des Cosinus nur gerade x-Potenzen auf.

(c) Der Beweis des ersten Additionstheorems ist nicht schwer aber mühselig. Man rechnet mit den Potenzreihen und dem Cauchyprodukt ähnlich wie wir es mit der Exponentialfunktion in Beispiel 13.11 gemacht haben.

Hat man dann das erste Additionstheorem gezeigt, bekommt man das zweite, indem man im ersten y festhält und die Gleichung dann nach x differenziert. Das ergibt

$$\begin{aligned}\cos(x+y) = \sin'(x+y) &= \cos(x)\cos(y) + \sin(y)\big(-\sin(x)\big)\\ &= \cos(x)\cos(y) - \sin(y)\sin(x). \qquad \square\end{aligned}$$

Die Eigenschaften aus Teil (b) des obigen Satzes sind sehr wichtig und bekommen deshalb einen eigenen Namen.

Definition 23.5. *Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt*

(a) gerade, *wenn für alle $x \in \mathbb{R}$ gilt $f(-x) = f(x)$.*

(b) ungerade, *wenn für alle $x \in \mathbb{R}$ gilt $f(-x) = -f(x)$.*

Satz 23.6. *(a) Für alle $x \in (0,2)$ gilt $\sin(x) > x - x^3/3! > 0$.*

(b) Die Funktion $\cos$ hat eine kleinste positive Nullstelle ξ_0.

Beweis. (a) Sei $x \in (0,2)$. Dann gilt für alle $n \in \mathbb{N}$

$$x^2 < 4 \le 4n^2 \le 4n^2 + 2n = 2n(2n+1)$$

und deshalb $x^2/(2n(2n+1)) < 1$. Multiplizieren wir diese Ungleichung nun mit $x^{2n-1}/(2n-1)! > 0$ durch, so erhalten wir

$$\frac{x^{2n+1}}{(2n+1)!} < \frac{x^{2n-1}}{(2n-1)!} \quad \text{bzw.} \quad \frac{x^{2n-1}}{(2n-1)!} - \frac{x^{2n+1}}{(2n+1)!} > 0$$

für jedes $n \in \mathbb{N}$. Damit ist

$$\sin(x) = \underbrace{\Big(x - \frac{x^3}{3!}\Big)}_{>0} + \underbrace{\Big(\frac{x^5}{5!} - \frac{x^7}{7!}\Big)}_{>0} + \underbrace{\Big(\frac{x^9}{9!} - \frac{x^{11}}{11!}\Big)}_{>0} + \cdots > x - \frac{x^3}{3!} > 0.$$

(b) Wir beobachten zunächst, dass aus (a)

$$\sin(1) > 1 - \frac{1}{3!} = 1 - \frac{1}{6} = \frac{5}{6}$$

folgt. Damit gilt mit dem Additionstheorem und mit Satz 23.4 (a)

$$\cos(2) = \cos(1+1) = \cos^2(1) - \sin^2(1) = \cos^2(1) + \sin^2(1) - 2\sin^2(1)$$
$$= 1 - 2\sin^2(1) < 1 - 2 \cdot \frac{5^2}{6^2} = 1 - \frac{50}{36} < 0.$$

Da außerdem $\cos(0) = 1 > 0$ gilt und cos eine stetige Funktion ist, gibt es nach dem Zwischenwertsatz ein $\xi \in (0,2)$ mit $\cos(\xi) = 0$.

Nun haben wir also eine Nullstelle, sind aber noch nicht fertig, denn wir wollen ja zeigen, dass es eine *kleinste* positive Nullstelle gibt. Dazu setzen wir $M := \{\eta > 0 : \cos(\eta) = 0\}$. Nach dem oben Gezeigten wissen wir, dass $M \neq \emptyset$ ist. Da M außerdem durch Null nach unten beschränkt ist, existiert also $\xi_0 := \inf M$. Zu zeigen ist noch, dass ξ_0 selbst eine Nullstelle des Cosinus ist, dass also $\xi_0 \in M$ und damit das Minimum von M ist. Da ξ_0 das Infimum ist, ist auf jeden Fall für jedes $n \in \mathbb{N}$ die Zahl $\xi_0 + 1/n$ keine untere Schranke von M, d. h. für jedes $n \in \mathbb{N}$ gibt es ein $\xi_n \in M$ mit $\xi_0 \leq \xi_n \leq \xi_0 + 1/n$. Insbesondere konvergiert also die Folge ξ_n gegen ξ_0. Nun ist aber die Cosinus-Funktion stetig, d. h.

$$\cos(\xi_0) = \cos(\lim_{n\to\infty} \xi_n) = \lim_{n\to\infty} \cos(\xi_n) = \lim_{n\to\infty} 0 = 0.$$

Also ist $\xi_0 = \min M$ und schließlich ist ξ_0 positiv, da $\xi_0 \geq 0$ ist und $\xi_0 = 0$ wegen $\cos(0) = 1$ nicht sein kann. □

Das Ergebnis des vorstehenden Satzes gibt Anlass zu folgender Definition.

Definition 23.7. *Die Zahl*

$$\pi := 2\xi_0$$

heißt Pi.

Da nach obigem Satz $\xi_0 \in (0,2)$ liegt, wissen wir bereits, dass π echt zwischen 0 und 4 liegt, tatsächlich ist π eine irrationale Zahl, deren Wert ungefähr $3,1415\ldots$ beträgt.
Wir haben nun also nach Definition $\cos(\pi/2) = 0$ und dank $\pi/2 \in (0,2)$ und Satz 23.6 (a) außerdem $\sin(\pi/2) > 0$. Wir können diesen Wert sogar ausrechnen, denn nach dem trigonometrischen Pythagoras gilt

$$1 = \sin^2(\pi/2) + \cos^2(\pi/2) = \sin^2(\pi/2), \quad \text{also} \quad \sin(\pi/2) = 1,$$

da der Wert nicht negativ und somit nicht -1 sein kann. Kombiniert man nun dieses Wissen mit den Additionstheoremen, so sieht man schnell ein, dass die folgenden Identitäten gelten.

Satz 23.8. *Für alle $x \in \mathbb{R}$ gilt*

$$\begin{aligned} \sin(x+\pi/2) &= \cos(x), & \cos(x+\pi/2) &= -\sin(x), \\ \sin(x+\pi) &= -\sin(x), & \cos(x+\pi) &= -\cos(x), \\ \sin(x+2\pi) &= \sin(x), & \cos(x+2\pi) &= \cos(x). \end{aligned}$$

Für die Identitäten der letzten Zeile sagt man auch, dass Cosinus und Sinus *periodische* Funktionen der *Periode* 2π, oder kurz 2π-periodische Funktionen sind.

Satz 23.9. *Der Cosinus hat in $[0,\pi]$ genau eine Nullstelle, nämlich $\pi/2$.*

Beweis. Es sei $\eta \in [0,\pi]$ eine Nullstelle des Cosinus. Dann gilt nach Satz 23.6 sofort $\eta \geq \pi/2$. Wir setzen $\tilde{\eta} := \pi - \eta$. Dann ist $\tilde{\eta} \in [0,\pi/2]$. Außerdem ist nach Satz 23.8 und da der Cosinus eine gerade Funktion ist

$$\cos(\tilde{\eta}) = \cos(-\eta+\pi) = -\cos(-\eta) = -\cos(\eta) = 0.$$

Somit muss aber $\tilde{\eta} = \pi/2$ und damit auch $\eta = \pi/2$ sein. □

Nun haben wir das gesamte Rüstzeug zusammen, um sämtliche Nullstellen von Sinus und Cosinus zu bestimmen und das sind ganz schön viele.

Satz 23.10. *Es gilt*

(a) $\cos(x) = 0 \iff x = k\pi + \dfrac{\pi}{2}$ *für ein $k \in \mathbb{Z}$.*

(b) $\sin(x) = 0 \iff x = k\pi$ *für ein $k \in \mathbb{Z}$.*

Beweis. Die Beweisrichtung von rechts nach links ergibt sich in beiden Fällen sofort aus Satz 23.8. Wir beweisen also jeweils nur noch von links nach rechts.

(a) Sei $x \in \mathbb{R}$ mit $\cos(x) = 0$. Dann gibt es ein $k \in \mathbb{Z}$, so dass $k\pi \leq x \leq (k+1)\pi$ gilt. (Wer sich noch an die Gaußklammer erinnert, kann $k = [x/\pi]$ nehmen.) Setze nun $\eta := x - k\pi$. Dann gilt $\eta \in [0,\pi]$ und

$$\cos(\eta) = \cos(x-k\pi) = \cos(x)\cos(k\pi) - \sin(x)\sin(k\pi) = 0,$$

da $\cos(x) = 0$ und $\sin(k\pi) = 0$ gilt. Nach Satz 23.9 ist damit $\eta = x - k\pi = \pi/2$, d. h. $x = k\pi + \pi/2$.

(b) Für die entsprechende Aussage für den Sinus müssen wir die Arbeit nicht wiederholen, denn falls $\sin(x) = 0$ ist, so haben wir $\cos(x+\pi/2) = 0$. Also gibt es nach dem soeben Bewiesenen ein $k \in \mathbb{Z}$ mit

$$x + \frac{\pi}{2} = k\pi + \frac{\pi}{2},$$

d. h. $x = k\pi$. □

Mit Hilfe von Sinus und Cosinus können wir nun noch die komplexen Zahlen etwas genauer verstehen. Insbesondere lässt sich beschreiben, was die komplexe Multiplikation anschaulich tut.
Dazu erinnern wir uns zunächst an den Zusammenhang

$$\mathrm{e}^z = \mathrm{e}^x \cos(y) + \mathrm{i}\mathrm{e}^x \sin(y), \qquad \text{für } z = x + \mathrm{i}y$$

aus Satz 16.16 (d). Setzt man speziell eine rein imaginäre Zahl $z = \mathrm{i}t$ mit $t \in \mathbb{R}$ ein, so sieht man

$$\mathrm{e}^{\mathrm{i}t} = \mathrm{e}^0 \cos(t) + \mathrm{i}\mathrm{e}^0 \sin(t) = \cos(t) + \mathrm{i}\sin(t).$$

Insbesondere gilt mit Hilfe des trigonometrischen Pythagoras, vgl. Satz 23.4 (a),

$$|\mathrm{e}^{\mathrm{i}t}| = 1 \quad \text{für alle } t \in \mathbb{R},$$

und das liefert wiederum für den Betrag der Exponentialfunktion

$$|\mathrm{e}^z| = |\mathrm{e}^x||\mathrm{e}^{\mathrm{i}y}| = \mathrm{e}^x = \mathrm{e}^{\mathrm{Re}(z)},$$

ein Zusammenhang, den man auch oft mit Gewinn verwenden kann.
Setzen wir nun noch speziell $t = 2\pi$, so folgt

$$\mathrm{e}^{2\pi\mathrm{i}} = \cos(2\pi) + \mathrm{i}\sin(2\pi) = 1$$

und damit allgmein für alle $z \in \mathbb{C}$

$$\mathrm{e}^{z+2\pi\mathrm{i}} = \mathrm{e}^z.$$

Die Exponentialfunktion ist also im komplexen eine periodische Funktion mit Periode $2\pi\mathrm{i}$.

Neben der Darstellung einer komplexen Zahl als $x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$, die dem karthesischen Koordinatensystem in der komplexen Zahlenebene entspricht, gibt es eine zweite Möglichkeit der Darstellung, die in Polarkoordinaten. Dazu beobachten wir zunächst anschaulich geometrisch, vgl. Abbildung 16.1, dass für jede komplexe Zahl z gilt

$$\mathrm{Re}(z) = |z|\cos(\varphi) \qquad \mathrm{Im}(z) = |z|\sin(\varphi),$$

wobei φ den Winkel zwischen der positiven reellen Achse und der Verbindungslinie zwischen z und dem Ursprung bezeichnet. Also ist

$$z = |z|\big(\cos(\varphi) + i\sin(\varphi)\big) = |z|\mathrm{e}^{\mathrm{i}\varphi}.$$

Diese geometrische Überlegung können wir auch analytisch untermauern.

Satz 23.11. *Es sei $z \in \mathbb{C}\backslash\{0\}$. Dann gibt es genau ein $\varphi \in (-\pi, \pi]$ mit $z = |z|\mathrm{e}^{\mathrm{i}\varphi}$.*

Beweis. Wir setzen $w := z/|z|$ und wählen $u, v \in \mathbb{R}$ so, dass $w = u + \mathrm{i}v$. Wegen $|w| = 1$ gilt dann $|u| = |\mathrm{Re}(w)| \leq 1$. Nach dem Zwischenwertsatz gibt es dann wegen $\cos(0) = 1$ und $\cos(\pi) = -1$ ein $\alpha \in [0, \pi]$ mit $u = \cos(\alpha)$.
Nun ist $u^2 + v^2 = |w|^2 = 1$, also haben wir $v^2 = 1 - \cos^2(\alpha) = \sin^2(\alpha)$, womit $v = \sin(\alpha)$ oder $v = -\sin(\alpha)$ gelten muss. Ist $v = \sin(\alpha)$, so setzen wir $\varphi := \alpha$, denn dann gilt

$$z = |z|w = |z|(u + \mathrm{i}v) = |z|\big(\cos(\varphi) + \mathrm{i}\sin(\varphi)\big) = |z|\mathrm{e}^{\mathrm{i}\varphi},$$

wobei $\varphi \in [0, \pi]$ ist.
Ist dagegen $v = -\sin(\alpha)$, so setzen wir $\varphi := -\alpha$, denn dann gilt, da der Cosinus gerade und der Sinus ungerade ist

$$z = |z|(u + \mathrm{i}v) = |z|\big(\cos(-\varphi) - \mathrm{i}\sin(-\varphi)\big) = |z|\big(\cos(\varphi) + \mathrm{i}\sin(\varphi)\big) = |z|\mathrm{e}^{\mathrm{i}\varphi}.$$

In diesem zweiten Fall ist dabei $\varphi \in [-\pi, 0]$. Wir haben jetzt also immer ein $\varphi \in [-\pi, \pi]$ gefunden, wollten aber eigentlich den Fall $\varphi = -\pi$ ausschließen. Das ist kein Problem, denn es gilt $\cos(\pi) = \cos(-\pi)$ und $\sin(\pi) = \sin(-\pi)$, also können wir falls $\varphi = -\pi$ ist, auch $\varphi = \pi$ nehmen.
Schließlich müssen wir noch die behauptete Eindeutigkeit beweisen. Wir nehmen also an, wir hätten zwei Winkel $\phi, \psi \in (-\pi, \pi]$ mit $z = |z|\mathrm{e}^{\mathrm{i}\phi} = |z|\mathrm{e}^{\mathrm{i}\psi}$. Dann ist $\mathrm{e}^{\mathrm{i}\phi} = \mathrm{e}^{\mathrm{i}\psi}$, d. h. wir haben $\mathrm{e}^{\mathrm{i}(\phi-\psi)} = 1$. Damit muss

$$\cos(\phi - \psi) = \mathrm{Re}\big(\mathrm{e}^{\mathrm{i}(\phi-\psi)}\big) = 1 \quad \text{und} \quad \sin(\phi - \psi) = \mathrm{Im}\big(\mathrm{e}^{\mathrm{i}(\pi-\psi)}\big) = 0$$

sein. Diese Konstellation tritt nur genau an den Stellen $\phi - \psi = 2\pi k$ für alle $k \in \mathbb{Z}$ auf. Da aber sowohl ϕ als auch ψ in $(-\pi, \pi]$ liegen, gilt $|\phi - \psi| < 2\pi$. Damit kommt nur der Fall $k = 0$ in Frage, was aber $\phi - \psi = 0$ und damit $\phi = \psi$ impliziert. $\square$

Definition 23.12. *Es sei $z \in \mathbb{C} \setminus \{0\}$. Die nach Satz 23.11 existierende Zahl $\varphi \in (-\pi, \pi]$ heißt das* Argument *von z und wird mit $\arg(z)$ bezeichnet.*

Anschaulich gibt das Argument von $z \in \mathbb{C} \setminus \{0\}$ den Winkel an, in dem diese Zahl in der Gaußschen Zahlenebene zur positiven reellen Achse steht.
Wir haben damit für alle $z \in \mathbb{C}$, die nicht Null sind, mit $z = |z|\mathrm{e}^{\mathrm{i}\arg(z)}$ eine weitere Möglichkeit neben $z = \mathrm{Re}(z) + \mathrm{i}\mathrm{Im}(z)$ diese durch reelle Größen auszudrücken. Umgekehrt erhalten wir für $r \in (0, \infty)$ und $\phi \in (-\pi, \pi]$ mit $r\mathrm{e}^{\mathrm{i}\phi}$ genau alle komplexen Zahlen außer der Null. Als Faustregel kann man sagen, dass diese Polardarstellung immer dann zu bevorzugen ist, wenn komplexe Zahlen multipliziert oder dividiert werden müssen, denn für zwei komplexe Zahlen $z = r\mathrm{e}^{\mathrm{i}\phi}$ und $w = s\mathrm{e}^{\mathrm{i}\psi}$ gilt

$$zw = rs\mathrm{e}^{\mathrm{i}\phi}\mathrm{e}^{i\psi} = rs\mathrm{e}^{\mathrm{i}(\phi+\psi)} \quad \text{und} \quad \frac{z}{w} = \frac{r\mathrm{e}^{\mathrm{i}\phi}}{s\mathrm{e}^{\mathrm{i}\psi}} = \frac{r}{s}\mathrm{e}^{\mathrm{i}(\phi-\psi)},$$

d. h. man muss zum Multiplizieren (Dividieren) nur die Beträge multiplizieren (dividieren) und die Argumente addieren (subtrahieren).

Nun, da wir eine Exponentialfunktion haben, stellt sich die Frage nach einem komplexen Logarithmus. Dies ist nicht so einfach wie im reellen, denn durch die Periodizität der Exponentialfunktion wird der Logarithmus im Komplexen mehrdeutig. Zuerst definieren wir die Logarithmen aber ganz abstrakt.

Definition 23.13. *Es sei $w \in \mathbb{C}$. Jede Zahl $z \in \mathbb{C}$ mit $\mathrm{e}^z = w$ heißt ein* Logarithmus *von z und man schreibt $z = \log(w)$.*

Nun wollen wir natürlich wissen, wie man einen solchen Logarithmus finden kann. Wir suchen also für ein vorgegebenes $w \in \mathbb{C} \setminus \{0\}$ (einen Logarithmus für Null wird es natürlich auch im komplexen nicht geben, da ja $\mathrm{e}^z \neq 0$ für alle $z \in \mathbb{C}$ gilt, vgl. Satz 16.16 (b)) also alle Lösungen $z \in \mathbb{C}$ der Gleichung $\mathrm{e}^z = w$. Sei dazu $r := |w|$ und $\phi := \arg(w) \in (-\pi, \pi]$. Weiter setzen wir z in der Form $x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ an. Dann soll also gelten

$$\mathrm{e}^z = \mathrm{e}^{x+\mathrm{i}y} = \mathrm{e}^x \mathrm{e}^{\mathrm{i}y} = r\mathrm{e}^{\mathrm{i}\phi}.$$

Also müssen insbesondere die beiden letzten Ausdrücke in dieser Gleichungskette den selben Betrag haben, was wegen $|\mathrm{e}^{\mathrm{i}y}| = |\mathrm{e}^{\mathrm{i}\phi}| = 1$, sofort $r = \mathrm{e}^x$, also

$$\mathrm{Re}(z) = x = \ln(r) = \ln(|w|)$$

liefert. Es bleibt also noch y zu bestimmen. Da nun $\mathrm{e}^x = r$ gilt, muss auch $\mathrm{e}^{\mathrm{i}y} = \mathrm{e}^{\mathrm{i}\phi}$ gelten, d. h. wir haben $\mathrm{e}^{\mathrm{i}(y-\phi)} = 1$. Wie am Schluss des Beweises von Satz 23.11 folgt daraus $y - \phi = 2k\pi$ für ein $k \in \mathbb{Z}$. Als Lösungen unserer Gleichung kommen also nur solche Zahlen $z = x + \mathrm{i}y$ mit $x = \ln(|w|)$ und

$$y = \phi + 2k\pi = \arg(w) + 2k\pi \quad \text{für ein } k \in \mathbb{Z}$$

in Betracht. Dass alle diese Zahlen tatsächlich Logarithmen von w sind, rechnet man schließlich einfach nach:

$$\mathrm{e}^{\ln(|w|)+\mathrm{i}(\arg(w)+2k\pi)} = \mathrm{e}^{\ln(|w|)}\mathrm{e}^{\mathrm{i}\arg(w)}\mathrm{e}^{\mathrm{i}2k\pi} = |w|\mathrm{e}^{\mathrm{i}\arg(w)} \cdot 1 = w.$$

Da der komplexe Logarithmus mehrdeutig, und damit gar keine Funktion mehr ist, tritt das gleiche Phänomen auch beim Versuch auf eine allgemeine Potenzfunktion in $\mathbb{C}$ zu definieren. Wir wollen das hier nicht weiter vertiefen, sondern nur vorwarnen, dass in diesem Zusammenhang größte Vorsicht angeraten ist.

Wir können nun weitere trigonometrische Funktionen definieren.

Definition 23.14. *Für $x \in \mathbb{R} \setminus \{k\pi + \pi/2 : k \in \mathbb{Z}\}$ heißt die Funktion*

$$\tan(x) := \frac{\sin(x)}{\cos(x)}$$

der Tangens *und für* $x \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$

$$\cot(x) := \frac{\cos(x)}{\sin(x)}$$

der Cotangens *von* x.

Nach der Quotientenregel gilt

$$\tan'(x) = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = 1 + \tan^2(x) > 0$$

für alle x im Definitionsbereich des Tangens. Also ist der Tangens insbesondere auf dem Intervall $(-\pi/2, \pi/2)$ streng monoton wachsend und da

$$\lim_{x \to -\pi/2+} \tan(x) = -\infty \quad \text{und} \quad \lim_{x \to \pi/2-} \tan(x) = \infty$$

gilt, ist $\tan((-\pi/2, \pi/2)) = \mathbb{R}$. Somit existiert die Umkehrfunktion $\tan^{-1} : \mathbb{R} \to (-\pi/2, \pi/2)$. Diese bekommt wieder einen Namen.

Definition 23.15. *Die Umkehrfunktion des Tangens auf* $(-\pi/2, \pi/2)$

$$\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$$

heißt Arcustangens.

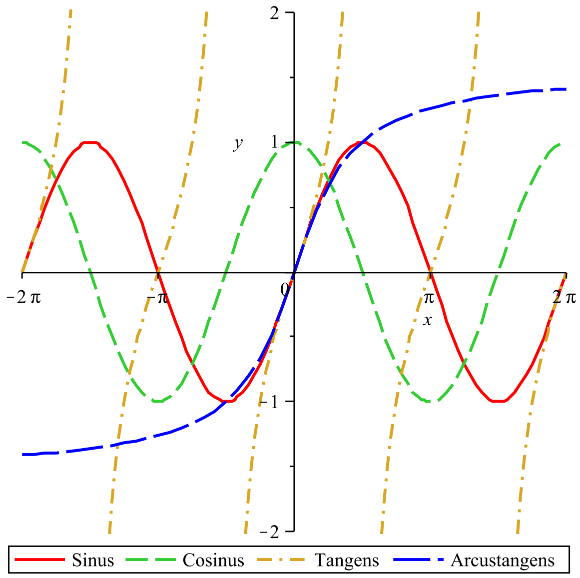

Abbildung 23.2.: Die Graphen von Sinus, Cosinus, Tangens und Arcustangens

Die Ableitung des Arcustangens können wir nun nach der Formel für die Ableitung der Umkehrfunktion (Satz 22.9) bestimmen:

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

Überraschenderweise erhalten wir als Ableitung eine gebrochen-rationale Funktion und nichts Trigonometrisches.
In der gleichen Weise kann man auch Umkehrfunktionen von Sinus und Cosinus definieren, wenn man sich im Definitionsbereich einschränkt. So sind

$$\sin : [-\pi/2, \pi/2] \to [-1,1] \quad \text{und} \quad \cos : [0,\pi] \to [-1,1]$$

bijektiv. Wir nennen die Umkehrfunktionen dieser Einschränkungen

$$\begin{aligned} \arcsin : [-1,1] &\to [-\pi/2, \pi/2] \quad \textit{(Arcussinus)}, \\ \arccos : [-1,1] &\to [0,\pi] \quad \textit{(Arcuscosinus)}. \end{aligned}$$

Die Ableitungen berechnen sich wieder mit Hilfe der Ableitungsregel für die Umkehrfunktion zu

$$\arcsin'(x) = \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-\sin^2(\arcsin(x))}} = \frac{1}{\sqrt{1-x^2}}$$

und genauso

$$\arccos'(x) = -\frac{1}{\sqrt{1-x^2}}$$

für alle $|x| < 1$.
Man beachte, dass die Auswahl des Bereichs, in dem man diese Funktionen invertiert, willkürlich ist. Üblicherweise werden zwar die hier gewählten Intervalle verwendet, aber welcher Bereich gewählt wurde, sollte bei Verwendung der Arcusfunktionen am besten immer dazugesagt werden. Hier ist Wachsamkeit angesagt!

Zum Abschluss definieren wir noch die hyperbolischen Funktionen.

Definition 23.16. *Für alle $x \in \mathbb{K}$ definieren wir*

$$\begin{aligned} \cosh(x) &:= \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x}) && \text{(Cosinus hyperbolicus)}, \\ \sinh(x) &:= \frac{1}{2}(\mathrm{e}^x - \mathrm{e}^{-x}) && \text{(Sinus hyperbolicus)}, \\ \tanh(x) &:= \frac{\sinh(x)}{\cosh(x)} && \text{(Tangens hyperbolicus)}. \end{aligned}$$

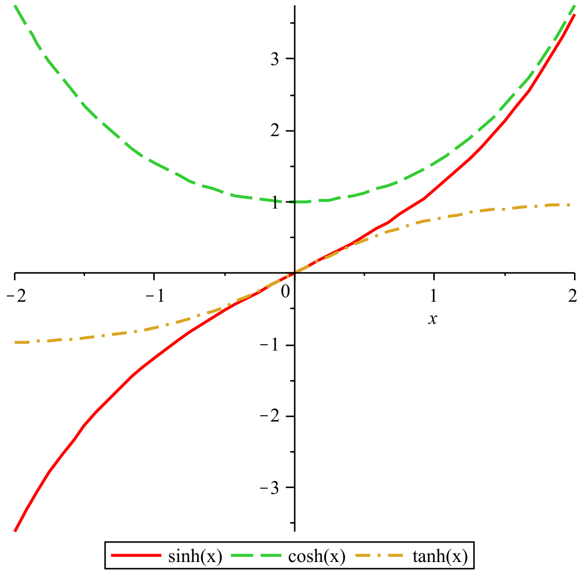

Abbildung 23.3.: Die Graphen von Sinus, Cosinus und Tangens hyperbolicus

Wir berechnen auch hier die Ableitungen:

$$\cosh'(x) = \frac{1}{2}(\mathrm{e}^x - \mathrm{e}^{-x}) = \sinh(x) \quad \text{und} \quad \sinh'(x) = \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x}) = \cosh(x),$$

sowie

$$\tanh'(x) = \frac{\cosh^2(x) - \sinh^2(x)}{\cosh^2(x)} = 1 - \tanh^2(x) > 0$$

für alle $x \in \mathbb{R}$, und weiterhin die Grenzwerte

$$\lim_{x\to\infty} \tanh(x) = \lim_{x\to\infty} \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{\mathrm{e}^x + \mathrm{e}^{-x}} = \lim_{x\to\infty} \frac{\mathrm{e}^x}{\mathrm{e}^x} \cdot \frac{1 - \mathrm{e}^{-2x}}{1 + \mathrm{e}^{-2x}} = \frac{1+0}{1-0} = 1.$$

$$\lim_{x\to-\infty} \tanh(x) = \lim_{x\to-\infty} \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{\mathrm{e}^x + \mathrm{e}^{-x}} = \lim_{x\to-\infty} \frac{\mathrm{e}^{2x-1}}{\mathrm{e}^{2x} + 1} = \frac{0-1}{0+1} = -1.$$

Damit ist der Tangens hyperbolicus streng motonon wachsend auf $\mathbb{R}$ und es gilt $\tanh(\mathbb{R}) = (-1, 1)$. Also existiert auch hier die Umkehrfunktion

$$\mathrm{Artanh} : (-1, 1) \to \mathbb{R} \quad \textit{(Areatangens hyperbolicus)}.$$

Die Ableitung ergibt sich hier zu

$$\mathrm{Artanh}'(x) = \frac{1}{1 - \tanh^2(\mathrm{Artanh}(x)))} = \frac{1}{1 - x^2}, \quad |x| < 1.$$

Abschließend wollen wir noch erwähnen, dass für die hyperbolischen Funktionen der Zusammenhang

$$\cosh^2(t) - \sinh^2(t) = 1$$

gilt, der als Pendant zum trigonometrischen Pythagoras gesehen werden kann. Man verifiziert diesen durch einfaches Nachrechnen.

In diesem Abschnitt haben wir nun so viele neue Funktionen eingeführt, dass wir diese noch einmal alle in einer Tabelle zusammenfassen wollen:

Name	Symbol	Def.-bereich	Bild	Ableitung
Sinus	sin	$\mathbb{R}$	$[-1,1]$	cos
Cosinus	cos	$\mathbb{R}$	$[-1,1]$	$-\sin$
Tangens	tan	$\mathbb{R} \setminus \{k\pi + \pi/2\}$	$\mathbb{R}$	$\frac{1}{\cos^2} = 1 + \tan^2$
Cotangens	cot	$\mathbb{R} \setminus \{k\pi\}$	$\mathbb{R}$	$-\frac{1}{\sin^2} = -1 - \cot^2$
Arcussinus	arcsin	$[-1,1]$	$[-\frac{\pi}{2}, \frac{\pi}{2}]$	$\frac{1}{\sqrt{1-x^2}}$
Arcuscosinus	arccos	$[-1,1]$	$[0,\pi]$	$-\frac{1}{\sqrt{1-x^2}}$
Arcustangens	arctan	$\mathbb{R}$	$(-\frac{\pi}{2}, \frac{\pi}{2})$	$\frac{1}{1+x^2}$
Sinus hyperbolicus	sinh	$\mathbb{R}$	$\mathbb{R}$	cosh
Cosinus hyp.	cosh	$\mathbb{R}$	$[1,\infty)$	sinh
Tangens hyp.	tanh	$\mathbb{R}$	$(-1,1)$	$\frac{1}{\cosh^2} = 1 - \tanh^2$
Areasinus hyp.	Arsinh	$\mathbb{R}$	$\mathbb{R}$	$\frac{1}{\sqrt{x^2+1}}$
Areacosinus hyp.	Arcosh	$[1,\infty)$	$[0,\infty)$	$\frac{1}{\sqrt{x^2-1}}$
Areatangens hyp.	Artanh	$(-1,1)$	$\mathbb{R}$	$\frac{1}{1-x^2}$

24. Höhere Ableitungen und der Satz von Taylor

Wir wollen in diesem Abschnitt zunächst das Konzept der zweiten, dritten und allgemein n-ten Ableitung einer Funktion einführen und dann mit dem Satz von Taylor einen fundamental wichtigen Satz der Analysis beweisen, der sowohl in abstrakten als auch in ganz angewandten Zusammenhängen immer wieder gebraucht wird. Es geht dabei darum komplizierte Funktionen durch möglichst angepasste Polynome zu nähern. Vor allem die PhysikerInnen werden mit diesem Satz sehr häufig zu tun bekommen.

Definition 24.1. *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{K}$ sei differenzierbar auf I.*

(a) Die Funktion f heißt in x_0 zweimal differenzierbar, *falls die Funktion $f' : I \to \mathbb{K}$ in x_0 wiederum differenzierbar ist. In diesem Fall heißt $f''(x_0) := (f')'(x_0)$ die* zweite Ableitung *von f in x_0.*

(b) Ist f in jedem Punkt $x \in I$ zweimal differenzierbar, so heißt f zweimal differenzierbar auf *I und die Funktion $x \mapsto f''(x)$ ist die* zweite Ableitung *von f* auf *I.*

(c) Für $n \geq 3$ beliebig definieren wir rekursiv: Die Funktion f heißt in x_0 (bzw. auf I) n mal differenzierbar, *falls sie $(n-1)$ mal differenzierbar ist und die Funktion $f^{(n-1)}$ in x_0 (bzw. auf I) wieder differenzierbar ist. In diesem Fall heißt $f^{(n)}(x_0) = (f^{(n-1)})'(x_0)$ die n-*te Ableitung *von f in x_0, bzw. $x \mapsto f^{(n)}(x)$ die n-te Ableitung von f auf I.*

Häufig ist es praktisch die Funktion selbst als ihre nullte Ableitung aufzufassen, also

$$f^{(0)} := f.$$

Beispiel 24.2. (a) Ist $f(x) = \sin(x)$, so gilt

$$\begin{aligned} f'(x) &= \cos(x), & f''(x) &= -\sin(x), \\ f'''(x) &= -\cos(x), & f^{(4)}(x) &= \sin(x) = f(x). \end{aligned}$$

(b) Betrachten wir auf $\mathbb{R}$ die Funktion

$$f(x) = \begin{cases} x^2, \text{ falls } x \geq 0 \\ -x^2, \text{ falls } x < 0, \end{cases}$$

so ist f auf ganz $\mathbb{R}$ differenzierbar mit $f'(x) = 2|x|$, $x \in \mathbb{R}$ (nachrechnen!), aber da die Betragsfunktion in Null nicht differenzierbar ist (vgl. Beispiel 22.3 (b)), ist diese Funktion in Null nicht mehr differenzierbar, d. h. f ist in $x_0 = 0$ nicht zweimal differenzierbar.

(c) Es sei $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$, d. h. f sei durch eine Potenzreihe gegeben, von der wir annehmen wollen, dass der Konvergenzradius $r > 0$ ist. Wir setzen wieder $I := (x_0 - r, x_0 + r)$. In Satz 23.1 haben wir gesehen, dass f dann auf I differenzierbar ist mit

$$f'(x) = \sum_{n=1}^{\infty} a_n n(x - x_0)^{n-1}, \quad x \in I,$$

und dass dies wieder eine Potenzreihe mit Konvergenzradius r ist. Also ist nach nochmaliger Anwendung dieses Satzes f sogar zweimal auf I differenzierbar mit

$$f''(x) = \sum_{n=2}^{\infty} a_n n(n-1)(x - x_0)^{n-2}, \quad x \in I.$$

Durch weitere Iteration dieses Arguments (Formalisten mögen eine saubere Induktion führen), ist dann f auf I beliebig oft differenzierbar und es gilt für jedes $k \in \mathbb{N}$

$$f^{(k)}(x) = \sum_{n=k}^{\infty} a_n n(n-1) \cdots (n-(k-1))(x - x_0)^{n-k}, \quad x \in I.$$

Setzt man speziell $x = x_0$ ein, so erhält man die für das Weitere wichtige Beziehung

$$f^{(k)}(x_0) = a_k k(k-1) \cdots 2 \cdot 1 = k! a_k.$$

Diese verrät uns insbesondere die Gestalt der Koeffizienten a_k für Funktionen, die durch eine Potenzreihe dargestellt werden:

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

(d) Wir betrachten auf $I = [0, \infty)$ die Funktion

$$f(x) := \begin{cases} x^{3/2} \sin(1/x), & \text{falls } x > 0 \\ 0, & \text{falls } x = 0. \end{cases}$$

Dann ist f in allen $x > 0$ offensichtlich differenzierbar und es gilt

$$\begin{aligned} f'(x) &= \frac{3}{2}\sqrt{x}\sin(1/x) + x^{3/2}\cos(1/x)\left(-\frac{1}{x^2}\right) \\ &= \frac{3}{2}\sqrt{x}\sin(1/x) - \frac{1}{\sqrt{x}}\cos(1/x). \end{aligned}$$

Um f auf Differenzierbarkeit in 0 zu untersuchen, betrachten wir den Differenzenquotienten

$$\begin{aligned} \lim_{x\to 0}\left|\frac{f(x)-f(0)}{x-0}\right| &= \lim_{x\to 0}\left|\frac{x^{3/2}\sin(1/x)}{x}\right| = \lim_{x\to 0}\left|\sqrt{x}\sin(1/x)\right| \\ &\leq \lim_{x\to 0}\sqrt{x} = 0. \end{aligned}$$

Da der Grenzwert selbst über etwas Positives gebildet wird, muss er auch größer oder gleich Null und damit gleich Null sein. Somit ist f auf ganz $[0,\infty)$ differenzierbar, wobei $f'(0) = 0$ gilt.

Anhand dieses Beispiels sieht man nun, dass etwas Abstruses passieren kann. Die Funktion $f' : [0,\infty)$ ist nämlich in Null nicht nur nicht noch einmal differenzierbar, sie ist nicht einmal mehr stetig, so dass wir über Differenzierbarkeit erst gar nicht mehr nachzudenken brauchen. Um das zu sehen, betrachten wir die Folge $x_n := 1/(n\pi)$, $n \in \mathbb{N}$. Dann gilt für jedes $n \in \mathbb{N}$

$$\begin{aligned} f'(x_n) &= \frac{3}{2}\frac{1}{\sqrt{n\pi}}\sin(n\pi) - \sqrt{n\pi}\cos(n\pi) = -\sqrt{n\pi}\cos(n\pi) \\ &= -(-1)^n\sqrt{n\pi} = (-1)^{n+1}\sqrt{n\pi}. \end{aligned}$$

Also ist die Funktion f' auf dem kompakten Intervall $[0,1]$ nicht beschränkt, sie kann also nach Satz 19.7 nicht stetig sein.

Das schlechte Differenzierbarkeitsverhalten der Funktion in (d) des obigen Beispiels nehmen wir zum Anlass, um einen stärkeren Differenzierbarkeitsbegriff zu formulieren, der so „hässliche“ Funktionen ausschließt.

Definition 24.3. *Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{K}$ eine Funktion.*

(a) Für $n \in \mathbb{N}$ heißt die Funktion f n-mal stetig differenzierbar *auf I, falls sie n-mal differenzierbar auf I ist und die Funktion $f^{(n)}$ auf I stetig ist.*

In diesem Fall schreiben wir $f \in C^n(I)$.

(b) Wir setzen

$$C^0(I) := C(I) \quad \text{und} \quad C^\infty(I) := \bigcap_{n\in\mathbb{N}_0} C^n(I)$$

und nennen eine Funktion $f \in C^\infty(I)$ beliebig oft differenzierbar.

Bemerkung 24.4. (a) Da die Differenzierbarkeit insbesondere Stetigkeit impliziert, sind für eine Funktion $f \in C^n(I)$ alle Ableitungen $f', f'', \dots, f^{(n)}$ stetige Funktionen auf I.

(b) Oft hört man statt „beliebig oft" auch die Bezeichnung „unendlich oft" differenzierbar. Diese ist etwas unglücklich, denn sie hört sich so an, als existiere auch eine „unendlichste" Ableitung. $f \in C^\infty(I)$ heißt aber eben nur, dass $f^{(n)}$ für jedes $n \in \mathbb{N}$ auf I existiert und dort stetig ist.

Wir haben in Beispiel 24.2 (c) gesehen, dass eine Funktion, die auf einem Intervall durch eine Potenzreihe $\sum_{n=0}^\infty a_n(x-x_0)^n$ gegeben ist, immer Koeffizienten a_n der Form

$$a_n = \frac{f^{(n)}(x_0)}{n!} \quad \text{für alle } n \in \mathbb{N}_0$$

aufweist. Haben wir umgekehrt eine Funktion aus $C^\infty(I)$ vorliegen, können wir für ein $x_0 \in I$ obige Koeffizienten ausrechnen und die dadurch gegebene Potenzreihe betrachten. Diese bekommt zunächst einen Namen.

Definition 24.5. *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f \in C^\infty(I)$. Die Potenzreihe*

$$\sum_{n=0}^\infty \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

heißt dann die Taylorreihe von f um x_0.

Nun stellt sich sofort die

Frage: Gilt in der Situation obiger Definition nun in einer Umgebung von x_0

$$f(x) = \sum_{n=0}^\infty \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \ ?$$

Die **Anwort** ist ein entschiedenes *manchmal.* Wir betrachten dazu das folgende Beispiel.

Beispiel 24.6. Wir wählen $I = \mathbb{R}$, $x_0 = 0$ und

$$f(x) := \begin{cases} \mathrm{e}^{-1/x^2}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0. \end{cases}$$

Dann ist f offensichtlich in jedem Punkt $x \neq 0$ differenzierbar, aber wir sind ja gerade an $x_0 = 0$ interessiert. Um Differenzierbarkeit in Null zu untersuchen, müssen wir über den Differenzenquotienten gehen. Es ist mit $t := 1/x$

$$\frac{f(x) - f(0)}{x - 0} = \frac{\mathrm{e}^{-1/x^2}}{x} = \frac{\mathrm{e}^{-t^2}}{1/t} = \frac{t}{\mathrm{e}^{t^2}}.$$

Dieser Ausdruck strebt für $x \to 0$, d. h. $t \to \pm\infty$, gegen 0. Also ist f in Null differenzierbar und es gilt $f'(0) = 0$. Mittels Induktion kann man nun zeigen, dass f in 0 sogar beliebig oft differenzierbar ist und $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$ gilt. Also ist in diesem Fall die Taylorreihe

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} 0 \cdot x^n = 0 \neq f(x) \quad \text{für alle } x \neq 0.$$

Da die Taylorreihe nach einem vielversprechenden Mittel aussieht, um Potenzreihenentwicklungen von Funktionen auszurechnen, und da diese ein unverzichtbares Hilfsmittel der Analysis, der Physik, der Ingenieurwissenschaften und vieler anderer Bereiche sind, hätten wir gerne ein Kriterium, wann die Taylorreihe brav zu ihrer Funktion passt. Ein solches folgt aus dem folgenden Satz.

Satz 24.7 (Satz von Taylor)**.** *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x, x_0 \in I$ und $f \in C^n(I)$ für ein $n \in \mathbb{N}_0$ eine reellwertige Funktion, so dass $f^{(n+1)}$ auf I existiert. Dann gibt es ein ξ zwischen x und x_0, so dass gilt*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}.$$

Bemerkung 24.8. (a) Im Fall $n = 0$ ist dieser Satz genau der Mittelwertsatz (vgl. Satz 22.16).

(b) Der Wert von ξ hängt natürlich jeweils von x_0, x und n ab und ist im Allgemeinen nicht zu bestimmen. Das wäre auch sehr erstaunlich, denn dann wäre ja die Berechnung von f auf den Schwierigkeitsgrad eines Polynoms zurückgeführt, was dann für einfach nur $(n+1)$-mal differenzierbare Funktionen doch ein bisschen zu simpel wäre. Im ξ steckt sozusagen die Komplexität der Funktion f.

Beweis von Satz 24.7. Wir betrachten ohne Beschränkung der Allgemeinheit den Fall $x_0 < x$ und setzen

$$\varrho := \frac{(n+1)!}{(x - x_0)^{n+1}} \Big(f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \Big).$$

Dann gilt

$$f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = \varrho \frac{(x - x_0)^{n+1}}{(n+1)!}$$

und unsere Aufgabe ist es ein $\xi \in (x_0, x)$ zu finden, so dass $\varrho = f^{(n+1)}(\xi)$ ist. Dazu schreiben wir die letzte Gleichung so um, dass auf der rechten Seite Null steht und definieren dann die Hilfsfunktion

$$g(t) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x - t)^k - \varrho \frac{(x - t)^{n+1}}{(n+1)!}, \quad t \in [x_0, x].$$

Dann ist nach den Voraussetzungen $g \in C([x_0, x])$ und da in g höchstens die n-te Ableitung von f auftaucht, ist g sogar noch einmal differenzierbar auf $[x_0, x]$. Außerdem gilt direkt $g(x) = f(x) - f(x) = 0$ und so wie wir ϱ gewählt haben gilt auch $g(x_0) = 0$. Nach dem Mittelwertsatz gibt es also ein $\xi \in (x_0, x)$, so dass

$$g'(\xi) = \frac{g(x) - g(x_0)}{x - x_0} = 0$$

gilt.
Andererseits ist (nachrechnen!)

$$g'(t) = \varrho \frac{(x-t)^n}{n!} - f^{(n+1)}(t) \frac{(x-t)^n}{n!},$$

womit

$$0 = g'(\xi) = \varrho \frac{(x-\xi)^n}{n!} - f^{(n+1)}(\xi) \frac{(x-\xi)^n}{n!}$$

und schließlich $\varrho = f^{(n+1)}(\xi)$ folgt. □

Wir wollen diesen Satz nun in konkreten Situationen anwenden. Zunächst einmal kann man den Satz dazu verwenden, den einen oder anderen Reihenwert zu bestimmen. Wir betrachten hierzu das folgende Beispiel.

Beispiel 24.9. Wir untersuchen auf $(-1, \infty)$ die Funktion $f(x) = \ln(1 + x)$ um $x_0 = 0$. Um den Satz anwenden zu können, müssen wir die ersten $n+1$, also alle, Ableitungen von f ausrechnen. Wir finden für $x > -1$

$$f'(x) = \frac{1}{1+x}, \quad f''(x) = \frac{-1}{(1+x)^2}, \quad f'''(x) = \frac{2}{(1+x)^3}, \quad f^{(4)}(x) = \frac{-2 \cdot 3}{(1+x)^4}$$

und per Induktion allgemein

$$f^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{(1+x)^k}, \quad k \in \mathbb{N}.$$

Insbesondere ist also für alle $k \in \mathbb{N}$

$$\frac{f^{(k)}(0)}{k!} = \frac{(-1)^{k-1}}{k}.$$

Nach dem Satz von Taylor gilt nun (mit $x = 1$):

$$\begin{aligned} \ln(2) = f(1) &= f(0) + \sum_{k=1}^{n} \frac{f^{(k)}(0)}{k!} 1^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} 1^{n+1} \qquad (24.1) \\ &= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} + \frac{(-1)^n}{(1+\xi)^{n+1}(n+1)} =: \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} + c_n \end{aligned}$$

mit einem $\xi \in (0,1)$. Unabhängig davon was das ξ nun genau ist, haben wir in jedem Fall $1+\xi > 1$ und damit auch $(1+\xi)^{n+1} > 1$. Deshalb gilt

$$|c_n| = \left|\frac{(-1)^n}{(1+\xi)^{n+1}(n+1)}\right| \le \frac{1}{n+1}, \quad n \in \mathbb{N}.$$

Also gilt $\lim_{n\to\infty} c_n = 0$. Damit können wir in (24.1) n gegen unendlich streben lassen und erhalten mit

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} = \ln(2)$$

den Reihenwert der alternierenden harmonischen Reihe (vgl. Beispiel 11.11).
Um etwaigen Mäkeleien zuvorzukommen: Wer meint, dass das aber viel Aufwand für so einen mickrigen Reihenwert war, hat noch nie selbst versucht einen Reihenwert zu bestimmen.
Abschließend sei darauf hingewiesen, dass wir damit unser Ergebnis aus Beispiel 23.2 insofern verbessert haben, als wir nun

$$\ln(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} \quad \text{für alle } x \in (-1,1]$$

und damit auch an einem Randpunkt wissen. Am anderen Randpunkt $x = -1$ ist die Reihe eine harmonische Reihe und somit divergent.

Wir haben hier schon gesehen, dass man mit dem Satz von Taylor auch andere Dinge als Potenzreihen berechnen kann, z. B. eben den obigen Reihenwert. Wir kommen nun zu einer weiteren, eher überraschenden, Anwendung des Satzes von Taylor.

Satz 24.10. $\mathrm{e} \notin \mathbb{Q}$.

Beweis. Wir wissen schon, dass $2 < \mathrm{e} < 3$ gilt. Nehmen wir nun an, es gäbe $n, m \in \mathbb{N}$ mit $\mathrm{e} = m/n$, so muss $n \ge 2$ sein, denn sonst wäre $\mathrm{e} \in \mathbb{N}$. Mit dem so gewählten n, $f(x) = \mathrm{e}^x$, $x = 1$ und $x_0 = 0$ wenden wir nun den Satz von Taylor an. Dieser liefert ein $\xi \in (0,1)$ mit

$$\frac{m}{n} = \mathrm{e} = f(1) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} + \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Nun sind die Ableitungen der Exponentialfunktion zum Glück nicht schwer zu bestimmen. Wir erhalten also

$$\frac{m}{n} = \sum_{k=0}^{n} \frac{1}{k!} + \frac{\mathrm{e}^{\xi}}{(n+1)!}$$

und nach Multiplikation dieser Gleichung mit $n!$

$$\underbrace{m(n-1)!}_{\in\mathbb{N}} = \underbrace{n! + n! + \frac{n!}{2!} + \cdots + 1}_{\in\mathbb{N}} + \frac{\mathrm{e}^\xi}{n+1}.$$

Da der Ausdruck $\mathrm{e}^\xi/(n+1)$ positiv ist, bleibt ihm damit nichts anderes übrig als selbst zu $\mathbb{N}$ zu gehören. Damit haben wir aber einen Widerspruch, denn wegen $n \geq 2$ und $\xi \in (0,1)$ gilt

$$0 < \frac{\mathrm{e}^\xi}{n+1} < \frac{\mathrm{e}}{n+1} \leq \frac{\mathrm{e}}{3} < 1. \qquad \square$$

Was uns immer noch fehlt, ist ein Kriterium, wann eine beliebig oft differenzierbare Funktion in einer Umgebung des Entwicklungspunktes durch ihre Taylorreihe dargestellt wird, wann also solch unangenehme Dinge wie in Beispiel 24.6 nicht passieren können. Hier ist eines.

Satz 24.11. *Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f \in C^\infty(I)$ eine Funktion. Weiter existiere eine Konstante $C \geq 0$, so dass für alle $n \in \mathbb{N}$ und alle $x \in I$ gilt*

$$\left|\frac{f^{(n)}(x)}{n!}\right| \leq C^n.$$

Dann gilt für jedes $x_0 \in I$ die Identität

$$f(x) = \sum_{n=0}^\infty \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \quad \textit{für alle } x \in J := I \cap (x_0 - 1/C, x_0 + 1/C).$$

Beweis. Wir müssen zunächst sicherstellen, dass die obige Potenzreihe überhaupt für alle $x \in J$ konvergiert. Das folgt direkt aus der Voraussetzung und dem Satz von Hadamard, denn

$$\sqrt[n]{\left|\frac{f^{(n)}(x_0)}{n!}\right|} \leq C, \quad \text{d. h. } \limsup_{n\in\mathbb{N}} \sqrt[n]{\left|\frac{f^{(n)}(x_0)}{n!}\right|} \leq C$$

und damit ist der Konvergenzradius größer oder gleich $1/C$.
Sei nun $x \in J$ beliebig gewählt. Nach dem Satz von Taylor mit $b = x$ und $a = x_0$ gibt es für jedes $n \in \mathbb{N}$ ein ξ_n zwischen x_0 und x, so dass

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(\xi_n)}{(n+1)!}(x-x_0)^{n+1}.$$

Uns bleibt nun wie im Beispiel 24.9 zu zeigen, dass der letzte Summand für n gegen unendlich gegen Null strebt. Dazu schätzen wir mit der Voraussetzung ab:

$$\left|\frac{f^{(n+1)}(\xi_n)}{(n+1)!}(x-x_0)^{n+1}\right| \leq C^{n+1}|x-x_0|^{n+1} = \big(C|x-x_0|\big)^{n+1}.$$

Da aber $x \in J$ ist, gilt $|x - x_0| < 1/C$, bzw. $C|x-x_0| < 1$, so dass obiger Ausdruck tatsächlich gegen Null geht, wenn n nach unendlich strebt. $\square$

Definition 24.12. *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$, $n \geq 1$ und $f \in C^n(I)$. Dann heißt das Polynom*

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

n-tes Taylorpolynom *von f in x_0.*

Bemerkung 24.13. (a) Das n-te Taylorpolynom hat immer höchstens Grad n, dieser kann aber auch kleiner sein. Es ist das eindeutig bestimmte Polynom p vom Grade kleiner oder gleich n, für das $p^{(k)}(x_0) = f^{(k)}(x_0)$ für alle $k \in \{0, 1, \dots, n\}$ gilt.

(b) Wir können mit dieser Definition den Satz von Taylor (Satz 24.7) folgendermaßen umformulieren:

Sei $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f \in C^n(I)$, so dass $f^{(n)}$ auf I differenzierbar ist, dann gibt es zu jedem $x \in I$ ein $\xi = \xi(x)$ zwischen x_0 und x mit

$$f(x) = T_n(x, x_0) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}.$$

(c) Die Differenz zwischen der Funktion und ihrem Taylorpolynom, also der Wert

$$\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

wird oft als *Restglied* bezeichnet.

Wir haben in Satz 22.14 gesehen, dass eine differenzierbare Funktion, die im Inneren eines Intervalls ein relatives Extremum hat, dort eine verschwindende Ableitung haben muss. Allerdings ist dies kein hinreichendes Kriterium für das Vorliegen eines Extremums, d. h. es kann sein, dass die Ableitung Null ist, ohne dass an dieser Stelle tatsächlich ein relatives Extremum vorliegen muss. Um wirklich nachzuweisen, dass eine solche kritische Stelle ein Extremum ist, brauchen wir genauere Hilfsmittel und auch hier hilft der Satz von Taylor.

Satz 24.14. *Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $x_0 \in I$ und $f \in C^n(I)$ für ein $n \geq 2$. Weiter gelte $f'(x_0) = f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0$, aber $f^{(n)}(x_0) \neq 0$. Ist nun n ungerade so hat f in x_0 kein Extremum, ist n gerade, so liegt in x_0 ein Extremum vor, und zwar falls $f^{(n)}(x_0) > 0$ ein Minimum und falls $f^{(n)}(x_0) < 0$ ein Maximum.*

Beweis. Da $f^{(n)}$ in x_0 stetig und I ein offenes Intervall ist, gibt es ein $\delta > 0$ mit $U_\delta(x_0) \subseteq I$, so dass $f^{(n)}(x)$ für alle $x \in U_\delta(x_0)$ das selbe Vorzeichen wie $f^{(n)}(x_0)$

hat. Sei nun $x \in U_\delta(x_0) \setminus \{x_0\}$ gewählt. Dann gibt es nach dem Satz von Taylor ein ξ zwischen x und x_0 mit

$$f(x) = T_{n-1}(x, x_0) + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

Nach Voraussetzung sind aber die ersten $n-1$ Ableitungen von f in x_0 alle Null, also bleibt vom $(n-1)$-ten Taylorpolynom nur der Summand nullter Ordnung übrig, d. h. es gilt $T_{n-1}(x, x_0) = f(x_0)$ und damit

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

Da ξ zwischen x_0 und x liegt, liegt es auch in $U_\delta(x_0)$, also hat $f^{(n)}(\xi)$ das selbe Vorzeichen wie $f^{(n)}(x_0)$. Mit diesen Überlegungen können wir nun die verschiedenen Fälle der Behauptung nacheinander untersuchen.
Ist n ungerade und $f^{(n)}(x_0) \gtrless 0$, so gilt das selbe für $f^{(n)}(\xi)$ und damit ist

$$c(x) \begin{cases} \gtrless 0, & \text{falls } x \in (x_0, x_0 + \delta), \\ \lesseqgtr 0, & \text{falls } x \in (x_0 - \delta, x_0), \end{cases}$$

da Potenzieren mit einem ungeraden Exponenten das Vorzeichen erhält. Daraus folgt aber mit $f(x) = f(x_0) + c(x)$

$$f(x) \begin{cases} \gtrless f(x_0), & \text{falls } x \in (x_0, x_0 + \delta), \\ \lesseqgtr f(x_0), & \text{falls } x \in (x_0 - \delta, x_0). \end{cases}$$

Damit kann f in x_0 kein Extremum haben, denn die Funktionswerte von f sind auf der einen Seite von x_0 kleiner und auf der anderen Seite von x_0 größer als in x_0.
Ist dagegen n gerade, so lässt das Potenzieren mit n das Vorzeichen verschwinden. Also gilt für $f^{(n)}(x_0) \gtrless 0$ wegen $f^{(n)}(\xi) \gtrless 0$ dann $c(x) \gtrless 0$ unabhängig davon auf welcher Seite von x_0 das x liegt. Das liefert schließlich $f(x) \gtrless f(x_0)$ für alle $x \in U_\delta(x_0)$ und damit die Behauptung. $\square$

25. Das Regelintegral

Wir haben nun zunächst mal unsere Betrachtungen zur Differenziation abgeschlossen und wollen uns einem auf den ersten Blick ganz anderen Problem zuwenden, der Berechnung von Flächeninhalten von krummlinig begrenzten Flächen. Wir werden jedoch feststellen, dass sich uns dabei ein sehr überraschender Zusammenhang zur Differenziation offenbart.
Wir betrachten das Problem der Flächenberechnung unter einem Funktionsgraphen, d. h. für $a, b \in \mathbb{R}$ mit $a < b$ und eine gegebene beschränkte Funktion $f : [a, b] \to \mathbb{R}$ wollen wir den Flächeninhalt der Fläche berechnen, die von der x-Achse, den beiden Geraden $x = a$ und $x = b$ und dem Graphen der Funktion eingeschlossen wird.[1]

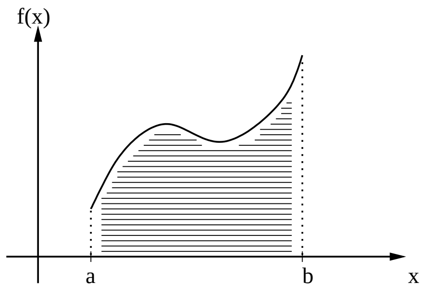

Abbildung 25.1.: Zu bestimmende Fläche unter dem Funktionsgraphen

Wir werden dazu zunächst ganz einfache Funktionen, sogenannte Treppenfunktionen, betrachten, für die sich der Flächeninhalt elementar-geometrisch bestimmen lässt. Den Übergang zu allgemeineren Funktionen wird dann wieder durch einen Grenzwertprozess gewährleistet, bei dem wir unser bisher gesammeltes Wissen über Funktionenfolgen gewinnbringend verwenden können.
Dabei werden wir den oben schon angedeuteten Zusammenhang zur Differenzialrechnung entdecken, und so schließlich tatsächlich in der Lage sein, den Flächeninhalt unter Umständen exakt angeben zu können.
Wie schon vorher verwenden wir wieder den Buchstaben $\mathbb{K}$, wann immer man sowohl $\mathbb{R}$ als auch $\mathbb{C}$ einsetzen kann. Außerdem seien in diesem gesamten Kapitel $a, b \in \mathbb{R}$ mit $a < b$ gegeben und wir bezeichnen mit I das abgeschlossene Intervall $[a, b]$.

[1] Genaugenommen geht es um die signierte Fläche, d. h. es sollen Flächen oberhalb der x-Achse positiv und Flächen unterhalb der x-Achse negativ zählen.

Definition 25.1. *(a) Eine endliche Menge* $Z := \{x_0, x_1, \ldots, x_n\} \subseteq I$ *heißt* Zerlegung *des Intervalls* I *genau dann, wenn gilt*

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

(b) Sind Z *und* $\tilde{Z}$ *zwei Zerlegungen des Intervalls* I*, so heißt* $\tilde{Z}$ *eine* Verfeinerung *von* Z*, falls* $Z \subseteq \tilde{Z}$ *gilt.*

(c) Sind Z_1 *und* Z_2 *zwei Zerlegungen von* I*, so ist auch* $Z := Z_1 \cup Z_2$ *eine Zerlegung dieses Intervalls (warum?) und heißt die* gemeinsame Verfeinerung *von* Z_1 *und* Z_2*.*

(d) Eine Funktion $f : I \to \mathbb{K}$ *heißt* Treppenfunktion*, falls es eine Zerlegung* $Z = \{x_0, \ldots, x_n\}$ *von* I *gibt, so dass* $f|_{(x_{j-1}, x_j)}$ *für jedes* $j \in \{1, \ldots, n\}$ *konstant ist.*

Eine solche Zerlegung Z *heißt* zu f passend.

(e) Die Menge aller Treppenfunktionen auf I *bezeichnen wir mit* $\mathcal{T}(I)$*.*

(f) Ist $A \subseteq \mathbb{K}$*, so heißt* $\mathbf{1}_A : \mathbb{K} \to \mathbb{K}$ *mit*

$$\mathbf{1}_A(x) := \begin{cases} 1, & \textit{falls } x \in A, \\ 0, & \textit{falls } x \notin A \end{cases}$$

charakteristische Funktion *von* A*.*

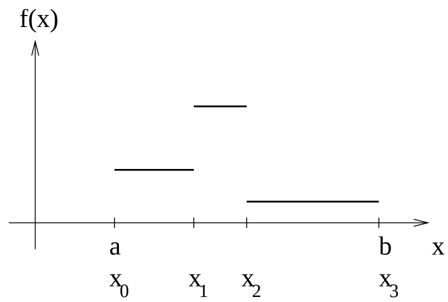

Abbildung 25.2.: Beispiel einer Treppenfunktion

Bemerkung 25.2. (a) Man beachte, dass bei der Definition einer Treppenfunktion die Werte von f an den Zerlegungsstellen $x_0, \ldots, x_n$ keinen Einschränkungen unterliegen. Hier kann f irgendwelche Werte annehmen.

(b) Ist $A \subseteq I$ eine endliche Vereinigung von Intervallen, so ist $\mathbf{1}_A$ eine Treppenfunktion.

(c) Ist $f : I \to \mathbb{K}$ eine Treppenfunktion, so können wir mit einer zu f passenden Zerlegung $\{x_0, \dots, x_n\}$ und geeigneten $c_1, \dots, c_n \in \mathbb{K}$ diese schreiben als

$$f(x) = \sum_{j=1}^{n} c_j \mathbf{1}_{(x_{j-1}, x_j)}(x) \quad \text{für alle } x \in I \setminus \{x_0, \dots, x_n\}.$$

Die Fläche unter dem Graph einer Treppenfunktion ist leicht zu bestimmen, es sind nur einige Rechteckflächen zu bestimmen. Das führt auf die folgende Definition.

Definition 25.3. *Sei $f \in \mathcal{T}(I)$ mit einer Darstellung $f(x) = \sum_{j=1}^{n} c_j \mathbf{1}_{(x_{j-1}, x_j)}(x)$ für eine passende Zerlegung $Z = \{x_0, \dots, x_n\}$. Dann heißt*

$$\int^{Z} f := \sum_{j=1}^{n} c_j (x_j - x_{j-1})$$

das Integral von f bezüglich Z*.*

Damit aus dieser Definition eine sinnvolle Setzung des Integrals von f werden kann, müssen wir nun natürlich als erstes sicherstellen, dass der Wert des Integrals für jede zu f passende Zerlegung der selbe ist. Das ist der Inhalt des ersten Lemmas in diesem Abschnitt.

Lemma 25.4. *Ist $f \in \mathcal{T}(I)$ eine Treppenfunktion und sind Z_1 und Z_2 zwei zu f passende Zerlegungen von I, so gilt*

$$\int^{Z_1} f = \int^{Z_2} f.$$

Beweis. Wir betrachten zunächst den Spezialfall, dass Z_2 eine Verfeinerung von Z_1 um genau einen Punkt ist, d. h. es gilt

$$Z_1 = \{x_0, \dots, x_n\} \quad \text{und} \quad Z_2 = \{x_0, \dots x_{k-1}, y, x_k, \dots, x_n\}$$

für ein $k \in \{1, \dots, n\}$ und ein $y \in (x_{k-1}, x_k)$. Dann gilt, da schon Z_1 zu f passend war, dass die jeweils konstanten Werte von $f|_{(x_{k-1}, y)}$, $f|_{(y, x_k)}$ und $f|_{(x_{k-1}, x_k)}$ alle gleich sind. Also erhalten wir mit der vorstehenden Definition

$$\begin{aligned}\int^{Z_1} f &= \sum_{j=1}^{n} c_j (x_j - x_{j-1}) = \sum_{j=1}^{k-1} c_j (x_j - x_{j-1}) + c_k (x_k - x_{k-1}) + \sum_{j=k+1}^{n} c_j (x_j - x_{j-1}) \\ &= \sum_{j=1}^{k-1} c_j (x_j - x_{j-1}) + c_k (x_k - y) + c_k (y - x_{k-1}) + \sum_{j=k+1}^{n} c_j (x_j - x_{j-1}) = \int^{Z_2} f,\end{aligned}$$

d. h. die Behauptung in diesem Spezialfall.
Ist Z_2 eine beliebige Verfeinerung von Z_1, so können wir einfach das obere Argument mehrfach anwenden, und so immer noch einen Punkt mehr dazu packen. Ganz formal sauber würde man dazu eine Induktion machen.
Wir wenden uns dem allgemeinen Fall zu, dass Z_1 und Z_2 beide zu f passen, aber keine der beiden eine Verfeinerung der anderen ist. In diesem Fall betrachten wir die gemeinsame Verfeinerung $Z := Z_1 \cup Z_2$. Auch diese ist dann zu f passend und wir erhalten mit obigen Überlegungen

$$\int^{Z_1} f = \int^{Z} f = \int^{Z_2} f$$

und sind fertig. $\square$

Definition 25.5. *Sei $f : I \to \mathbb{K}$ eine Treppenfunktion und Z eine zu f passende Zerlegung. Dann heißt*

$$\int_I f := \int^Z f$$

das Integral von f über I.

Bemerkung 25.6. Für das Integral gibt es mehrere synonyme Schreibweisen. Es ist

$$\int_I f =: \int_a^b f =: \int_a^b f(x)\,\mathrm{d}x =: \int_I f(x)\,\mathrm{d}x.$$

Wir wollen in den folgenden Sätzen ein paar erste Eigenschaften des Integrals für Treppenfunktionen beweisen.

Satz 25.7. *Seien $f, g \in \mathcal{T}(I)$ und $\lambda, \mu \in \mathbb{K}$. Dann sind auch $\lambda f + \mu g$, $|f|$, $\operatorname{Re} f$, $\operatorname{Im} f$ und $\overline{f}$ Treppenfunktionen auf I und es gilt*

(a) $\displaystyle\int_I (\lambda f + \mu g) = \lambda \int_I f + \mu \int_I g$ (Linearität des Integrals),

(b) $\displaystyle\left|\int_I f\right| \leq \int_I |f|$ (Dreiecksungleichung).

(c) $\displaystyle\int_I \operatorname{Re} f = \operatorname{Re}\Big(\int_I f\Big)$, $\displaystyle\int_I \operatorname{Im} f = \operatorname{Im}\Big(\int_I f\Big)$, *und* $\displaystyle\int_I \overline{f} = \overline{\int_I f}$.

Beweis. Übung.

Satz 25.8. *Sei $f : I \to \mathbb{K}$ eine Treppenfunktion. Dann gilt*

$$\left|\int_I f\right| \leq (b-a) \cdot \sup_{x \in I} |f(x)|. \qquad (\text{„Standardabschätzung“})$$

Beweis. Sei $Z = \{x_0, \dots, x_n\}$ eine zu f passende Zerlegung von I und $f(x) = \sum_{j=1}^n c_j \mathbf{1}_{(x_{j-1},x_j)}(x)$ für geeigente $c_1, \dots, c_n \in \mathbb{K}$. Dann gilt nach der Defintion des Integrals für Treppenfunktionen

$$\Big|\int_I f\Big| = \Big|\int^Z f\Big| = \Big|\sum_{j=1}^n c_j(x_j - x_{j-1})\Big|.$$

Also erhalten wir mit der Dreiecksungleichung für Summen und dank einer freundlichen Teleskopsumme

$$\begin{aligned}\Big|\int_I f\Big| &\leq \sum_{j=1}^n |c_j||x_j - x_{j-1}| \leq \sum_{j=1}^n \max\{|c_1|, \dots |c_n|\} \cdot (x_j - x_{j-1}) \\ &\leq \sup_{x\in I} |f(x)| \sum_{j=1}^n (x_j - x_{j-1}) = \sup_{x\in I} |f(x)|(x_n - x_0) = \sup_{x\in I} |f(x)|(b-a). \square\end{aligned}$$

Satz 25.9. *Seien $f, g : I \to \mathbb{R}$ Treppenfunktionen mit $f(x) \leq g(x)$ für alle $x \in I$. Dann gilt $\int_I f \leq \int_I g$.*

Beweis. Seien Z eine zu f passende und W eine zu g passende Zerlegung von I. Dann passt die gemeinsame Verfeinerung $Z \cup W =: \{x_0, \dots x_n\}$ sowohl zu f als auch zu g und es sei

$$f(x) = \sum_{j=1}^n c_j \mathbf{1}_{(x_{j-1},x_j)}(x), \quad g(x) = \sum_{j=1}^n d_j \mathbf{1}_{(x_{j-1},x_j)}(x) \quad \text{für alle } x \in I \backslash \{x_0, \dots, x_n\}$$

für geeignete $c_1, \dots, c_n, d_1, \dots, d_n \in \mathbb{R}$. Nach Voraussetzung gilt dann $c_j \leq d_j$ für alle $j = 1, \dots, n$ und wir haben damit

$$\int_I f = \int^{Z\cup W} f = \sum_{j=1}^n c_j(x_j - x_{j-1}) \leq \sum_{j=1}^n d_j(x_j - x_{j-1}) = \int^{Z\cup W} g = \int_I g. \qquad \square$$

Nun ist das Leben im Allgemeinen nicht stückweise konstant und irgenwann werden Treppenfunktionen auch langweilig. Wir wollen uns also nun einer größeren Funktionenklasse zuwenden.

Definition 25.10. *Eine Funktion $f : I \to \mathbb{K}$ heißt* sprungstetig *auf I, falls für alle $x_0 \in [a, b)$ der rechtsseitige Grenzwert $\lim_{x\to x_0+} f(x)$ und für alle $x_0 \in (a, b]$ der linksseitige Grenzwert $\lim_{x\to x_0-} f(x)$ existiert.*
Die Menge aller sprungstetigen Funktionen auf I bezeichnen wir mit $\mathcal{S}(I)$.

Um einem häufigen Missverständnis gleich vorzubeugen: Nichts, aber auch gar nichts in dieser Definition fordert, dass der links- und der rechtsseitige Grenzwert in obiger Definition gleich sein müssen. Damit eine Funktion sprungstetig ist, müssen die beiden nur existieren.

Lemma 25.11. *Jede sprungstetige Funktion $f : I \to \mathbb{K}$ ist beschränkt.*

Beweis. Übung.

Das folgende Lemma liefert einen großen Zoo von Beispielen sprungstetiger Funktionen.

Lemma 25.12. *Alle*

(a) Treppenfunktionen $f : I \to \mathbb{K}$

(b) stetigen Funktionen $f : I \to \mathbb{K}$ und

(c) monotonen Funktionen $f : I \to \mathbb{R}$

sind sprungstetig.

Beweis. Für Treppenfunktionen und stetige Funktionen ist die Behauptung klar. Wir wenden uns also (c) zu und betrachten im Folgenden nur den Fall einer monoton wachsenden Funktion f. Für monoton fallende Funktionen kann man entweder einen analogen Beweis führen oder $-f$ betrachten.
Sei also $x_0 \in (a, b]$. Dann setzen wir

$$\alpha := \sup\{f(x) : x \in [a, x_0)\}.$$

Sei nun (x_n) eine Folge in $[a, x_0)$ mit $x_n \to x_0$ $(n \to \infty)$. Wir wollen nun zeigen, dass $\lim_{n\to\infty} f(x_n) = \alpha$ ist. Dann folgt $\lim_{x\to x_0-} f(x) = \alpha$ und dieser Grenzwert existiert damit.
Zunächst beobachten wir, dass nach Defintion von α auf jeden Fall $f(x_n) \leq \alpha$ für alle $n \in \mathbb{N}$ gilt. Die Folge $(f(x_n))$ ist also nach oben beschränkt. Ein erster Reflex ist nun auf das Monotoniekriterium loszugehen, schließlich ist ja f monoton! Aber das funktioniert nicht, denn wir wissen nichts über Monotonie von (x_n). Wir müssen uns also etwas anderes einfallen lassen.
Sei also $\varepsilon > 0$ gegeben. Nach der Definition von α ist dann $\alpha - \varepsilon$ keine obere Schranke von $\{f(x) : x \in [a, x_0)\}$ mehr, d. h. es gibt ein $y \in [a, x_0)$ mit $f(y) > \alpha - \varepsilon$. Nun gilt $\lim_{n\to\infty} x_n = x_0$, also existiert ein $n_0 \in \mathbb{N}$ mit $x_n \geq y$ für alle $n \geq n_0$. Für all diese n haben wir nun schlussendlich mit Hilfe der Monotonie von f

$$f(x_n) \geq f(y) > \alpha - \varepsilon, \quad \text{d. h.} \quad 0 \leq \alpha - f(x_n) < \varepsilon.$$

Das liefert $\lim_{n\to\infty} f(x_n) = \alpha$.
Zur Behandlung der rechtsseitigen Grenzwerte, wählt man $x_0 \in [a, b)$. Man definiert

$$\beta := \inf\{f(x) : x \in (x_0, b]\}$$

und zeigt dann mit einem analogen Argument $\lim_{x\to x_0+} f(x) = \beta$. □

Was haben nun sprungstetige Funktionen mit unseren Integralen zu tun? Die Verbindung schafft das folgende für unser Integral fundamentale Theorem, der Approximationssatz für sprungstetige Funktionen.

Theorem 25.13. *Eine Funktion $f : I \to \mathbb{K}$ ist genau dann sprungstetig, wenn es eine Folge (f_n) von Treppenfunktionen auf I gibt, die gleichmäßig auf I gegen f konvergiert.*

Beweis. **„$\Rightarrow$"** [2] Wir führen einen Beweis durch Widerspruch. Dazu nehmen wir an, die zu beweisende Aussage wäre falsch, d. h. es gibt eine sprungstetige Funktion $f \in \mathcal{S}(I)$, die sich nicht gleichmäßig durch Treppenfunktionen approximieren lässt und zeigen, dass das auf einen Widerspruch führt.

In einem **ersten Schritt** überlegen wir uns, dass es unter dieser Annahme ein $\varepsilon_0 > 0$ geben muss, so dass für alle Treppenfunktionen $g \in \mathcal{T}(I)$ gilt

$$\sup_{x \in I} |g(x) - f(x)| > \varepsilon_0.$$

Nehmen wir an, das wäre nicht so, so gibt es für jedes $\varepsilon > 0$ ein $g_\varepsilon \in \mathcal{T}(I)$ mit $\sup_{x \in I} |g_\varepsilon(x) - f(x)| \leq \varepsilon$. Nimmt man speziell für jedes $n \in \mathbb{N}$ die Wahl $\varepsilon = 1/n$, so erhalten wir eine Folge $h_n = g_{1/n}$ von Treppenfunktionen auf I, so dass $\sup_{x \in I} |h_n(x) - f(x)| \leq 1/n$ ist. Also gilt

$$\big|h_n(x) - f(x)\big| \leq \frac{1}{n} \text{ für alle } n \in \mathbb{N} \text{ und alle } x \in I.$$

Nach Satz 20.8 konvergiert also die Folge von Treppenfunktionen (h_n) gleichmäßig gegen f, aber gerade das ist ja nicht möglich. Das postulierte $\varepsilon_0 > 0$ existiert also und wir bewahren dieses für die weiteren Schritte des Beweises sicher auf.

Im ebenfalls vorbereitenden **zweiten Schritt** betrachten wir ein beliebiges abgeschlossenes Teilintervall $J = [c, d]$ von I, wobei natürlich $c < d$ gelten soll sowie die beiden Teilintervalle $J^- := [c, (c + d)/2]$ und $J^+ := [(c + d)/2, d]$, d. h. die „linke und rechte Hälfte" von J.

Für $\varepsilon > 0$ nennen wir ein $\varphi \in \mathcal{S}(J)$ *auf J ε-biestig*, falls für jede Treppenfunktion $g : J \to \mathbb{K}$ gilt, dass $\sup_{x \in J} |g(x) - \varphi(x)| > \varepsilon$ ist. Das Ziel ist nun zu zeigen, dass für jedes $\varphi \in \mathcal{S}(J)$, das auf J ε-biestig ist, auch zumindest eine der beiden Einschränkungen $\varphi|_{J^-}$ und $\varphi|_{J^+}$ ε-biestig auf J^- oder J^+ sein muss. Auch hier gehen wir indirekt vor und zeigen die Kontraposition, d. h. wir beweisen: Ist $\varphi \in \mathcal{S}(J)$ so, dass φ_{J^-} und $\varphi|_{J^+}$ nicht ε-biestig auf J^- bzw. J^+ sind, so ist auch φ nicht ε-biestig auf J.

[2] Ich danke Herrn Moritz Egert für die Idee zu diesem schönen Beweis, der mit den Mitteln dieser Vorlesung darstellbar ist.

Also los: Da $\varphi|_{J^-}$ und $\varphi|_{J^+}$ nicht ε-biestig sind, gibt es Treppenfunktionen $g_- \in \mathcal{T}(J^-)$ und $g_+ \in \mathcal{T}(J^+)$ mit $\sup_{x\in J^-} |g_-(x) - \varphi(x)| \le \varepsilon$ und $\sup_{x\in J^+} |g_+(x) - \varphi(x)| \le \varepsilon$. Wir betrachten nun die Funktion $g : J \to \mathbb{K}$, die durch

$$g(x) = \begin{cases} g_-(x), & c \le x < (c+d)/2 \\ \varphi(x), \quad \text{falls} & x = (c+d)/2 \\ g_+(x), & (c+d)/2 < x \le d \end{cases}$$

gegeben ist. Diese ist wieder eine Treppenfunktion auf J (warum?) und es gilt dank $J = J^- \cup J^+$

$$\begin{aligned} \sup_{x\in J} \big|g(x) - \varphi(x)\big| &= \max\big\{ \sup_{x\in J^-} \big|g(x) - \varphi(x)\big|, \sup_{x\in J^+} \big|g(x) - \varphi(x)\big|\big\} \\ &= \max\big\{ \sup_{x\in J^-} \big|g_-(x) - \varphi(x)\big|, \sup_{x\in J^+} \big|g_+(x) - \varphi(x)\big|\big\} \le \varepsilon. \end{aligned}$$

Also ist auch φ nicht ε-biestig auf J und Schritt 2 ist erledigt.

Im **dritten Schritt** können wir nun den eigentlichen Beweis führen. Sei also $f \in \mathcal{S}(I)$ nicht gleichmäßig durch Treppenfunktionen approximierbar. Dann haben wir im ersten Schritt gezeigt, dass es ein $\varepsilon_0 > 0$ gibt, für das f auf I ε_0-biestig ist. Wir wissen nun aus Schritt 2, dass eine der beiden Einschränkungen $f|_{I^+}$ oder $f|_{I^-}$ ebenfalls auf I^+ bzw. I^- ε_0-biestig sein muss. Eine Hälfte von I, für die das der Fall ist, nennen wir I_1.

Nun können wir dieses Verfahren iterieren und erhalten eine Folge $I = I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$ von abgeschlossenen Teilintervallen von I, deren Länge gegen Null schrumpft und auf denen die Einschränkung von f jeweils ε_0-biestig ist.

Nach Übungsaufgabe 9.13 gilt also $\bigcap_{n\in\mathbb{N}} I_n = \{x_\infty\}$ mit einem $x_\infty \in I$. Der weitere Beweis geht davon aus, dass x_∞ weder mit a noch b übereinstimmt. Die (einfachen) Modifikationen, die sonst nötig werden, bleiben den Lesenden als Übung.

Da f sprungstetig auf I ist, existieren die beiden Grenzwerte $f(x_\infty-) := \lim_{x\to x_\infty-} f(x)$ und $f(x_\infty+) := \lim_{x\to x_\infty+} f(x)$. Zu unserem ε_0 gibt es daher $a_\infty, b_\infty \in I$ mit $a_\infty < x_\infty < b_\infty$ sowie $|f(x) - f(x_\infty-)| < \varepsilon_0$ für alle $x \in [a_\infty, x_\infty]$ und $|f(x) - f(x_\infty+)| < \varepsilon_0$ für alle $x \in [x_\infty, b_\infty]$.

Wir betrachten nun die Treppenfunktion

$$g(x) = \begin{cases} f(x_\infty-), & a_\infty \le x < x_\infty \\ f(x_\infty), \quad \text{falls} & x = x_\infty \\ f(x_\infty+), & x_\infty < x \le b_\infty \end{cases}$$

auf $[a_\infty, b_\infty]$. Für diese gilt dort nach den obigen Erkenntnissen

$$\sup_{x\in[a_\infty,b_\infty]} \big|g(x) - f(x)\big| \le \varepsilon_0.$$

Sei nun $n_0 \in \mathbb{N}$ so, dass $I_{n_0} \subseteq [a_\infty, b_\infty]$ gilt. Das geht, da die Länge der Intervalle I_n gegen Null schrumpft und x_∞ in jedem I_n enthalten ist. Dann ist g auch eine Treppenfunktion auf I_{n_0}, für die gilt

$$\sup_{x \in I_{n_0}} \big| g(x) - f(x) \big| \leq \varepsilon_0.$$

Also ist f nicht ε_0-biestig auf I_{n_0}, womit wir bei einem Widerspruch wären.

„$\Leftarrow$" Seien nun $f_n : I \to \mathbb{K}$ für jedes $n \in \mathbb{N}$ Treppenfunktionen und die Folge (f_n) konvergiere gleichmäßig auf I gegen eine Funktion $f : I \to \mathbb{K}$. Wir müssen nun zeigen, dass f sprungstetig ist.

Seien dazu $x_0 \in (a, b]$ und eine Folge (x_j) in $[a, x_0)$ mit $x_j \to x_0$ $(j \to \infty)$ gegeben. Sei außerdem $\varepsilon > 0$. Dank der gleichmäßigen Konvgergenz von (f_n) existiert dann ein $n_0 \in \mathbb{N}$, so dass

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2} \quad \text{für alle } n \geq n_0 \text{ und alle } x \in I$$

gilt. Wir wählen für das Folgende ein $n \geq n_0$ fest. Nun ist f_n nach Voraussetzung eine Treppenfunktion, also gibt es ein $\alpha \in [a, x_0)$, so dass $f_n|_{(\alpha, x_0)}$ konstant ist. Da weiterhin (x_j) von links gegen x_0 konvergiert, gibt es ein $j_0 \in \mathbb{N}$ mit $x_j \in (\alpha, x_0)$ für alle $j \geq j_0$. Das bedeutet, dass $f_n(x_j) = f_n(x_k)$ für alle Wahlen von $j, k \geq j_0$ gilt. Also haben wir für alle $j, k \geq j_0$ mit der Dreiecksungleichung

$$\begin{aligned} \big| f(x_j) - f(x_k) \big| &= \big| f(x_j) - f_n(x_j) + f_n(x_k) - f(x_k) \big| \\ &\leq \big| f(x_j) - f_n(x_j) \big| + \big| f(x_k) - f_n(x_k) \big| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

D.h. $(f(x_j))$ ist eine Cauchy-Folge und damit konvergent.

Wir haben damit gezeigt, dass für jede Folge (x_j) die von links gegen x_0 konvergiert, der Grenzwert $\lim_{j \to \infty} f(x_j)$ existiert. Nach Satz 17.9 existiert damit der Grenzwert $\lim_{x \to x_0-} f(x)$.

Der Beweis zur Existenz von $\lim_{x \to x_0+} f(x)$ für alle $x_0 \in [a, b)$ geht analog. □

Wir können dieses Approximations-Resultat auch mit Hilfe von Funktionenreihen statt Funktionenfolgen schreiben. Das bringt per se zwar keine neue Erkenntnis, denn Reihen sind ja auch nur Folgen, aber je nach Situation kann man die folgende Formulierung des obigen Theorems leichter anwenden.

Korollar 25.14. *Eine Funktion $f : I \to \mathbb{K}$ ist genau dann sprungstetig, wenn es eine Folge (f_n) von Treppenfunktionen auf I gibt mit $f = \sum_{n=1}^\infty f_n$ mit gleichmäßiger und absoluter Konvergenz.*

Beweis. **„⇒“** Nach Theorem 25.13 gibt es für jedes $k \in \mathbb{N}$ eine Treppenfunktion $\varphi_k : I \to \mathbb{K}$ mit

$$\left|f(x) - \varphi_k(x)\right| \leq \frac{1}{2^k} \quad \text{für alle } x \in I.$$

Wir setzen nun $f_1 := \varphi_1$ und $f_k := \varphi_k - \varphi_{k-1}$ für jedes $k \geq 2$. Dann gilt für alle $n \in \mathbb{N}$ und alle $x \in I$

$$\begin{aligned} \left|\sum_{k=1}^{n} f_k(x) - f(x)\right| &= \left|\varphi_1(x) + \sum_{k=2}^{n}\big(\varphi_k(x) - \varphi_{k-1}(x)\big) - f(x)\right| \\ &= \left|\varphi_1(x) - \varphi_1(x) + \varphi_n(x) - f(x)\right| = \left|\varphi_n(x) - f(x)\right| \\ &\leq \frac{1}{2^n}. \end{aligned}$$

Also haben wir $f = \sum_{k=1}^{\infty} f_k$ mit gleichmäßiger Konvergenz. Für den Nachweis, dass die Reihe auch absolut konvergiert, überlegen wir uns, dass dank der Approximationseigenschaften von φ_k für alle $k \geq 2$ und alle $x \in I$ gilt

$$\begin{aligned} \left|f_k(x)\right| &= \left|\varphi_k(x) - \varphi_{k-1}(x)\right| \leq \left|\varphi_k(x) - f(x)\right| + \left|f(x) - \varphi_{k-1}(x)\right| \\ &\leq \frac{1}{2^k} + \frac{1}{2^{k-1}} = \frac{3}{2^k}. \end{aligned}$$

Da $\sum_{k=0}^{\infty} \frac{3}{2^k}$ als geometrische Reihe konvergiert, haben wir damit eine konvergente Majorante für $(|f_k(x)|)$ gefunden, d. h. wir haben auch absolute Konvergenz.

„⇐“ Wir setzen $\varphi_k := \sum_{n=1}^{k} f_n$ für jedes $k \in \mathbb{N}$. Dann ist φ_k nach Satz 25.7 eine Treppenfunktion und es gilt für alle $x \in I$ und alle $k \in \mathbb{N}$

$$|f(x) - \varphi_k(x)| = \left|f(x) - \sum_{n=1}^{k} f_n(x)\right| = \left|\sum_{n=k+1}^{\infty} f_n(x)\right| \leq \sum_{n=k+1}^{\infty} \left|f_n(x)\right|.$$

Der Reihenrest im letzten Ausdruck in dieser Ungleichungskette konvergiert nun gleichmäßig in $x \in I$ gegen 0, da nach Voraussetzung die Reihe $\sum_{n=1}^{\infty} f_n$ gleichmäßig und absolut konvergiert.

Also ist (φ_k) eine Folge von Treppenfunktionen auf I, die gleichmäßig auf I gegen f konvergiert. Theorem 25.13 impliziert dann, dass f sprungstetig ist. □

Die Idee für das weitere Vorgehen ist nun vorgezeichnet. Wir können jede sprungstetige Funktion f gleichmäßig durch Treppenfunktionen annähern und von jeder Treppenfunktion das Integral bestimmen. Wir zeigen nun also, dass die Integrale dieser annähernden Treppenfunktionen konvergiereren und definieren dann den Grenzwert als Integral von f. Dabei müssen wir natürlich sicherstellen, dass der so definierte Wert des Integrals nicht von der speziellen Wahl der approximierenden Folge von Treppenfunktionen abhängt.

Theorem 25.15. *Sei $f \in \mathcal{S}(I)$. Dann existiert für jede Folge (f_n) von Treppenfunktionen in I, die gleichmäßig auf I gegen f konvergieren, der Grenzwert $\lim_{n\to\infty} \int_I f_n$ und dieser ist für alle solchen Folgen der selbe.*

Beweis. Es sei $\alpha_n := \int_I f_n$ und $\varepsilon > 0$. Dank der gleichmäßigen Konvergenz von (f_n) gegen f existiert dann ein $n_0 \in \mathbb{N}$ mit

$$\bigl|f_n(x) - f(x)\bigr| \leq \frac{\varepsilon}{2(b-a)} \quad \text{für alle } n \geq n_0 \text{ und alle } x \in I.$$

Für alle Wahlen von $n, m \geq n_0$ gilt damit mit ein wenig Unterstützung der Standardabschätzung aus Satz 25.8

$$\begin{aligned}
|\alpha_n - \alpha_m| = \Bigl|\int_I (f_n - f_m)\Bigr| &\leq (b-a) \sup_{x\in I} \bigl|f_n(x) - f_m(x)\bigr| \\
&\leq (b-a)\Bigl[\sup_{x\in I}\bigl(|f_n(x) - f(x)| + |f(x) - f_m(x)|\bigr)\Bigr] \\
&\leq (b-a)\Bigl[\sup_{x\in I} |f_n(x) - f(x)| + \sup_{x\in I} |f(x) - f_m(x)|\Bigr] \\
&\leq (b-a)\Bigl[\frac{\varepsilon}{2(b-a)} + \frac{\varepsilon}{2(b-a)}\Bigr] = \varepsilon.
\end{aligned}$$

Also ist (α_n) eine Cauchy-Folge und damit schon mal konvergent.
Sei nun (g_n) eine weitere Folge von Treppenfunktionen auf I, die gleichmäßig auf I gegen f konvergiert. Wir betrachten die Folge $(h_n) := (f_1, g_1, f_2, g_2, f_3, g_3, \dots)$. Dann ist auch (h_n) eine Folge von Treppenfunktionen auf I und diese konvergiert auch gleichmäßig auf I gegen f (Übungsaufgabe!).
Nach dem ersten Teil des Beweises ist dann die Folge $(\int_I h_n)$ konvergent. Damit haben aber insbesondere die beiden Teilfolgen dieser Folge $(\int_I f_n)$ und $(\int_I g_n)$ den selben Grenzwert und wir sind fertig. □

Das vorstehende Theorem rechtfertigt nun also die folgende Defintion.

Definition 25.16. *Sei $f : I \to \mathbb{K}$ sprungstetig und (f_n) eine Folge von Treppenfunktionen auf I, die gleichmäßig auf I gegen f konvergiert. Dann heißt*

$$\int_I f := \int_a^b f := \int_I f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x := \lim_{n\to\infty} \int_I f_n$$

das Integral von f auf I.

Damit haben wir unseren Integralbegriff nun auf eine große Klasse von Funktionen ausgeweitet. Man beachte, dass wir damit dank Lemma 25.12 insbesondere alle stetigen und alle monotonen Funktionen auf I integrieren können. Nun brauchen wir natürlich ein Beispiel einer so nicht integrierbaren Funktion. Zunächst sind natürlich alle unbeschränkten Funktionen zu nennen, die wir bisher nicht behandeln können. Aber gibt es auch beschränkte Funktionen, die sich unserem Integral widersetzen? Ja, hier ist so eine:

Beispiel 25.17. Es sei $f : [0,1] \to \mathbb{R}$ die sogenannte *Dirichletsche Sprungfunktion*, gegeben durch

$$f(x) = \begin{cases} 1, \text{ falls } x \in [0,1] \cap \mathbb{Q} \\ 0, \text{ falls } x \in [0,1] \text{ und } x \notin \mathbb{Q}. \end{cases}$$

Da man jeden rationalen Punkt in $[0,1]$ durch eine irrationale Folge und jeden irrationalen Punkt durch eine rationale Folge annähern kann, ist diese Funktion in keinem $x \in [0,1]$ stetig. Dass sie dann nie und nimmer nach unserem Integralbegriff integrierbar sein kann, folgt aus dem nächsten Satz.

In der Vorlesung Maß- und Intgrationstheorie werden Sie mit dem Lebesgue-Integral einen deutlich mächtigeren Integralbegriff kennen lernen, mit dem man sogar dem Integral über die Dirichletsche Sprungfunktion einen sinnvoll definierten Wert zuweisen kann. Das Lebesgue-Integral basiert auf der gleichen Idee wie das Regelintegral, der Approximation durch Treppenfunktionen. Es weitet die Betrachtungen jedoch auf *punktweise* Grenzwerte von Treppenfunktionen aus, was deutlich mehr Funktionen integrierbar macht. Diese Verallgemeinerung wirft jedoch einige knifflige Probleme auf, die eine genauere Beschäftigung mit der Frage nötig machen, wie man das Volumen von Mengen messen kann. Wir stellen die Behandlung dieses Integralbegriffs, der die Grundlage der modernen Analysis darstellt, deshalb erst einmal zurück. Im Moment reicht uns das Regelintegral vollständig aus.

Satz 25.18. *Jede sprungstetige Funktion auf I besitzt höchstens abzählbar viele Unstetigkeitsstellen. Insbesondere gilt das damit für alle monotonen Funktionen.*

Beweis. Sei $f \in \mathcal{S}(I)$. Dann existiert nach Korollar 25.14 eine Folge (f_n) von Treppenfunktionen auf I mit $f = \sum_{n=1}^{\infty} f_n$ und die Konvergenz der Reihe ist absolut und gleichmäßig in I. Für jedes $n \in \mathbb{N}$ sei

$$M_n := \{x \in I : f_n \text{ unstetig in } x\}.$$

Da jedes f_n eine Treppenfunktion ist, sind alle Mengen M_n endlich. Also ist $M := \bigcup_{n=1}^{\infty} M_n$ höchstens abzählbar. Wir zeigen, dass f in allen Punkten x, die nicht in M liegen, stetig ist. Sei also $x \notin M$. Dann ist f_n stetig in x für alle $n \in \mathbb{N}$. Da die Reihe über f_n gleichmäßig konvergiert ist dann nach Satz 20.13 auch f stetig in x. □

26. Eigenschaften des Integrals

Wir wollen nun Rechenregeln und Eigenschaften unseres Integrals sammeln und insbesondere den Hauptsatz der Differenzial- und Integralrechnung beweisen, der uns die Möglichkeit eröffnet, verschiedenste Integrale auch konkret zu bestimmen. Zunächst übertragen wir die Eigenschaften des Integrals über Treppenfunktionen, indem wir die sprungstetigen Funktionen durch Treppenfunktionen approximieren und dann die gleichmäßige Konvergenz ausnutzen.
Auch in diesem Kapitel seien wieder durchgehend $a, b \in \mathbb{R}$ mit $a < b$ und $I := [a, b]$ und wir verwenden wieder den Buchstaben $\mathbb{K}$ für $\mathbb{R}$ oder $\mathbb{C}$.

Satz 26.1. *Es seien $f, g \in \mathcal{S}(I)$ und $\alpha, \beta \in \mathbb{K}$. Dann sind auch $\alpha f + \beta g$, $|f|$, $\operatorname{Re} f$, $\operatorname{Im} f$ sowie $\overline{f}$ sprungstetig und es gilt*

(a) $\displaystyle\int_I (\alpha f + \beta g) = \alpha \int_I f + \beta \int_I g.$ (Linearität des Integrals)

(b) $\displaystyle\Big|\int_a^b f\Big| \le \int_a^b |f| \le (b-a) \sup_{x \in I} |f(x)|.$
(Dreiecksungleichung und Standardabschätzung)

(c) $\displaystyle\operatorname{Re}\Big(\int_I f\Big) = \int_I \operatorname{Re} f, \quad \operatorname{Im}\Big(\int_I f\Big) = \int_I \operatorname{Im} f,$ *und* $\displaystyle\overline{\int_I f} = \int_I \overline{f}.$

(d) Ist $\mathbb{K} = \mathbb{R}$ und gilt $f(x) \le g(x)$ für alle $x \in I$, so ist auch
$\displaystyle\int_I f \le \int_I g.$ (Monotonie des Integrals)

Beweis. (a) Wir wählen Folgen (f_n) und (g_n) von Treppenfunktionen auf I, so dass (f_n) gleichmäßig gegen f und (g_n) gleichmäßig gegen g konvergiert. Dann konvergiert auch die Funktionenfolge $(\alpha f_n + \beta g_n)$ gleichmäßig gegen $\alpha f + \beta g$ und für jedes $n \in \mathbb{N}$ ist $\alpha f_n + \beta g_n$ eine Treppenfunktion auf I. Also ist nach dem Approximationssatz für sprungstetige Funktionen 25.13 auch $\alpha f + \beta g$ sprungstetig.

Weiter haben wir nach Defintion des Integrals und da das Integral für Treppenfunktionen nach Satz 25.7 linear ist

$$\int_I (\alpha f + \beta g) = \lim_{n \to \infty} \int_I (\alpha f_n + \beta g_n) = \lim_{n \to \infty} \Big(\alpha \int_I f_n + \beta \int_I g_n\Big) = \alpha \int_I f + \beta \int_I g.$$

(b) Wir wählen eine Folge (f_n) von Treppenfunktionen auf I, die gleichmäßig auf I gegen f konvergiert. Nach Übungsaufgabe 20.9 (a) konvergiert dann auch $(|f_n|)$ gleichmäßig auf I gegen $|f|$. Da auch die Funktionen $|f_n|$, $n \in \mathbb{N}$, Treppenfunktionen sind, ist damit also auch $|f|$ sprungstetig und wir haben

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

Weiter gilt mit Hilfe der Dreiecksungleichung

$$\Big|\int_I f\Big| = \Big|\lim_{n\to\infty} \int_I f_n\Big| = \lim_{n\to\infty} \Big|\int_I f_n\Big| \le \lim_{n\to\infty} \int_I |f_n| = \int_I |f|.$$

Weiter bekommen wir mit der Standardabschätzung aus Satz 25.8 und Übungsaufgabe 20.9 (b)

$$\int_a^b |f| = \lim_{n\to\infty} \int_a^b |f_n| \le \lim_{n\to\infty} (b-a) \sup_{x\in I} |f_n(x)| = (b-a) \sup_{x\in I} |f(x)|.$$

(c) Übungsaufgabe.

(d) Seien wieder $f_n, g_n \in \mathcal{T}(I)$, $n \in \mathbb{N}$, so dass (f_n) gleichmäßig gegen f und (g_n) gleichmäßig gegen g konvergiert. Dann konvergieren nach Übungsaufgabe 20.9 (b) die Folgen

$$\alpha_n := \sup_{y\in I} |f(y) - f_n(y)| \qquad \text{und} \qquad \beta_n := \sup_{y\in I} |g(y) - g_n(y)|, \qquad n \in \mathbb{N},$$

gegen Null.

Wir betrachten nun die Funktionenfolgen

$$\varphi_n(x) := f_n(x) - \alpha_n \qquad \text{und} \qquad \psi_n(x) := g_n(x) + \beta_n, \qquad x \in I,\ n \in \mathbb{N}.$$

Dann sind auch φ_n und ψ_n für jedes $n \in \mathbb{N}$ Treppenfunktionen auf I. Wir zeigen, dass auch (φ_n) gleichmäßig gegen f konvergiert. Für alle $x \in I$ gilt

$$|\varphi_n(x) - f(x)| = |f_n(x) - \alpha_n - f(x)| \le |f_n(x) - f(x)| + |\alpha_n|.$$

Also ist

$$\sup_{x\in I} |\varphi_n(x) - f(x)| \le |\alpha_n| + \sup_{x\in I} |f_n(x) - f(x)|$$

und die gewünschte Konvergenzaussage folgt aus der gleichmäßigen Konvergenz von (f_n) und wieder Übungsaufgabe 20.9 (b). Eine analoge Überlegung zeigt schließlich, dass auch (ψ_n) gleichmäßig auf I gegen g konvergiert.

Als nächsten Schritt beobachten wir, dass für alle $n \in \mathbb{N}$ und alle $x \in I$

$$\begin{aligned}\varphi_n(x) = f_n(x) - \alpha_n + f(x) - f(x) &\leq \underbrace{\big|f_n(x) - f(x)\big|}_{\leq \alpha_n} - \alpha_n + f(x) \\ &\leq f(x) \leq g(x) = g(x) - g_n(x) + g_n(x) = g(x) - g_n(x) + \psi_n(x) - \beta_n \\ &\leq \underbrace{\big|g(x) - g_n(x)\big|}_{\leq \beta_n} - \beta_n + \psi_n(x) \leq \psi_n(x)\end{aligned}$$

gilt. Also haben wir mit Hilfe von Satz 25.9 schlussendlich

$$\int_I f = \lim_{n\to\infty} \int_I \varphi_n \leq \lim_{n\to\infty} \int_I \psi_n = \int_I g. \qquad \square$$

Definition 26.2. *Es seien $f \in \mathcal{S}(I)$ und $c, d \in I$ mit $c < d$. Dann definieren wir*

$$\int_c^c f(x)\,\mathrm{d}x := 0 \qquad \textit{und} \qquad \int_d^c f(x)\,\mathrm{d}x := -\int_c^d f(x)\,\mathrm{d}x = -\int_{[c,d]} f.$$

Lemma 26.3. *Seien $f : I \to \mathbb{K}$ eine sprungstetige Funktion und $c \in I$. Dann gilt*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Beweis. Die Aussage ist richtig für Treppenfunktionen (man wähle eine zu f passende Zerlegung von I, die c enthält). Wählen wir nun eine Folge (f_n) von Treppenfunktionen auf I, die auf I gleichmäßig gegen f konvergiert, so sind auch $f_n|_{[a,c]}$ und $f_n|_{[c,b]}$ für jedes $n \in \mathbb{N}$ Treppenfunktionen auf $[a,c]$ bzw. $[c,b]$ und es gilt

$$\begin{aligned}\sup_{x\in[a,c]} |f_n(x) - f(x)| &\leq \sup_{x\in I} |f_n(x) - f(x)| \qquad \text{und} \\ \sup_{x\in[c,b]} |f_n(x) - f(x)| &\leq \sup_{x\in I} |f_n(x) - f(x)|.\end{aligned}$$

Also konvergieren dank Übungsaufgabe 20.9 (b) auch $(f_n|_{[a,c]})$ bzw. $(f_n|_{[c,b]})$ gleichmäßig auf $[a,c]$ bzw. $[c,b]$ gegen $f|_{[a,c]}$ bzw. $f|_{[c,b]}$. Das liefert

$$\int_I f = \lim_{n\to\infty} \int_I f_n = \lim_{n\to\infty} \Big(\int_{[a,c]} f_n + \int_{[c,b]} f_n \Big) = \int_{[a,c]} f + \int_{[c,b]} f. \qquad \square$$

Lemma 26.4. *Es sei $f : I \to \mathbb{R}$ eine sprungstetige Funktion mit $f(x) \geq 0$ für alle $x \in I$. Ist f an einer Stelle $c \in I$ stetig mit $f(c) > 0$, so gilt $\int_I f > 0$.*

Beweis. Wir führen den Beweis für den Fall $c \in (a,b)$. Die einfachen Modifikationen der Argumentation in den Fällen $c = a$ und $c = b$ bleiben als Übungsaufgabe stehen.

Da f in c stetig ist und $c \in (a,b)$ liegt, existiert ein $\delta > 0$ mit $[c-\delta, c+\delta] \subseteq (a,b)$ und

$$f(x) \geq \frac{1}{2} f(c) > 0 \quad \text{für alle } x \in [c-\delta, c+\delta].$$

Weiter ist f auf ganz I nicht-negativ, also haben wir dank der Monotonie des Integrals $\int_{[a,c-\delta]} f \geq 0$ und $\int_{[c+\delta,b]} f \geq 0$. Damit erhalten wir

$$\int_I f = \int_a^{c-\delta} f + \int_{c-\delta}^{c+\delta} f + \int_{c+\delta}^{b} f \geq \int_{c-\delta}^{c+\delta} f \geq \int_{c-\delta}^{c+\delta} \frac{1}{2} f(c) = 2\delta \frac{1}{2} f(c) > 0. \qquad \square$$

Mit der Vorarbeit aus diesem Lemma können wir nun den Mittelwertsatz der Integralrechnung beweisen.

Satz 26.5 (Mittelwertsatz der Integralrechnung)**.** *Seien $f, \varphi : I \to \mathbb{R}$ stetig und es gelte $\varphi(x) \geq 0$ für alle $x \in I$. Dann existiert ein $\xi \in I$ mit*

$$\int_a^b f(x)\varphi(x) \, \mathrm{d}x = f(\xi) \int_a^b \varphi(x) \, \mathrm{d}x.$$

Bemerkung 26.6. Bevor wir diesen Satz beweisen, lohnt es sich den wichtigen Spezialfall $\varphi = 1$ zu betrachten. Er lautet dann: Es existiert ein $\xi \in I$ mit

$$\int_a^b f(x) \, \mathrm{d}x = f(\xi)(b-a).$$

In dieser Formulierung hat der Satz auch eine geometrisch-anschauliche Bedeutung. Er besagt, dass es einen Funktionswert $f(\xi)$ mit $\xi \in [a,b]$ gibt, so dass das Rechteck mit den Seitenlängen $f(\xi)$ und $b-a$ den gleichen Flächeninhalt hat wie die Fläche unter dem Graphen von f zwischen a und b, vgl. Abbildung 26.1.
Wie schon in mehreren ähnlichen Fällen vorher (Mittelwertsatz der Differenzialrechnung, Satz von Taylor) ist auch hier der genaue Wert von ξ meist nicht bestimmbar, oder seine Bestimmung zumindest genau so schwierig wie die Bestimmung des Integrals von f.

Beweis von Satz 26.5. Wir bemerken zunächst, dass die Aussage des Satzes für den Fall, dass φ konstant Null ist, offensichtlich richtig ist. Folglich konzentrieren wir uns auf den Fall, dass es ein $x_0 \in I$ mit $\varphi(x_0) \neq 0$ gibt. Da φ nicht-negativ ist, muss dann $\varphi(x_0) > 0$ gelten und wir erhalten $\int_I \varphi > 0$ aus Lemma 26.4.
Da I kompakt und f stetig ist, nimmt f auf I sein Minimum und Maximum an, wir können also

$$m := \min_{x \in I} f(x) \qquad \text{und} \qquad M := \max_{x \in I} f(x)$$

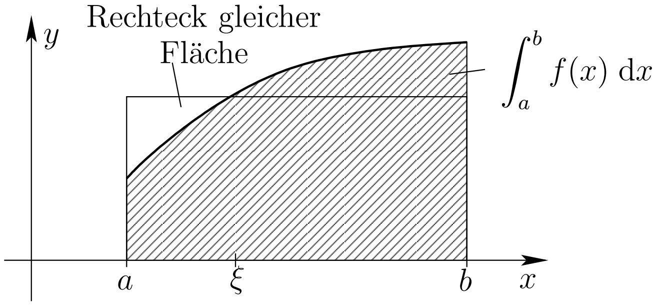

Abbildung 26.1.: Der Mittelwertsatz der Integralrechnung

setzen. Dann gilt, eingedenk der Positivität von φ, die Ungleichungskette $m\varphi(x) \leq f(x)\varphi(x) \leq M\varphi(x)$ für alle $x \in I$. Also erhalten wir

$$m \int_I \varphi(x)\,\mathrm{d}x = \int_I m\varphi(x)\,\mathrm{d}x \leq \int_I f(x)\varphi(x)\,\mathrm{d}x \leq M \int_I \varphi(x)\,\mathrm{d}x$$

und damit ist wegen $\int_I \varphi \neq 0$

$$m \leq \frac{\int_I f\varphi}{\int_I \varphi} \leq M.$$

Nun ist f nach Voraussetzung eine stetige Funktion, also erkennen wir an dieser Ungleichungskette, dass es nach dem Zwischenwertsatz ein $\xi \in I$ geben muss mit

$$f(\xi) = \frac{\int_I f\varphi}{\int_I \varphi}.$$

Multiplikation dieser Gleichung mit $\int_I \varphi$ liefert die Behauptung. □

Satz 26.7. *Es sei $f \in \mathcal{S}(I)$. Für $c \in [a,b]$ definieren wir die Funktion $F_c : I \to \mathbb{K}$ durch $F_c(x) := \int_c^x f$, $x \in I$. Dann gelten die folgenden Aussagen.*

(a) F_c ist Lipschitz-stetig, also insbesondere stetig auf I.

(b) Ist f an einer Stelle $d \in I$ stetig, so ist F_c in d differenzierbar und es gilt $F_c'(d) = f(d)$.

Beweis. (a) Es seien $x, y \in I$. Dann gilt nach der Definition von F_c

$$F_c(x) - F_c(y) = \int_c^x f - \int_c^y f = \int_y^x f.$$

Also können wir den Betrag mit Hilfe der Standardabschätzung durch

$$\left|F_c(x) - F_c(y)\right| = \left|\int_x^y f\right| \leq |x-y| \sup_{s\in[x,y]} |f(s)| \leq \sup_{s\in I} |f(s)| \cdot |x-y|$$

abschätzen. Mit $L := \sup_{s\in I} |f(s)|$ (Man bedenke, dass f als sprungstetige Funktion auf I insbesondere beschränkt ist, vgl. Lemma 25.11) folgt damit die Behauptung.

(b) Für jedes $h \in \mathbb{R} \setminus \{0\}$ mit $d + h \in I$ gilt

$$\frac{F_c(d+h) - F_c(d)}{h} = \frac{1}{h}\Big(\int_c^{d+h} f - \int_c^d f\Big) = \frac{1}{h}\int_d^{d+h} f.$$

Weiter gilt $f(d) = \frac{1}{h} \cdot \int_d^{d+h} f(d)\, \mathrm{d}s$. Also haben wir, wieder mit Hilfe der Standardabschätzung

$$\begin{aligned}&\Big|\frac{F_c(d+h) - F_c(d)}{h} - f(d)\Big| = \Big|\frac{1}{h}\int_d^{d+h} f(s)\, \mathrm{d}s - \frac{1}{h}\int_d^{d+h} f(d)\, \mathrm{d}s\Big| \\ &= \frac{1}{|h|}\Big|\int_d^{d+h} \big(f(s) - f(d)\big)\, \mathrm{d}s\Big| \leq \sup_{s\in[d-|h|,d+|h|]} \big|f(s) - f(d)\big|.\end{aligned}$$

Da f in d stetig ist, geht nun dieses Supremum für $h \to 0$ gegen Null (warum?). Damit ist gezeigt, dass F_c in d differenzierbar ist mit $F_c'(d) = f(d)$. □

Mit diesen Vorarbeiten können wir nun den Hauptsatz der Differenzial- und Integralrechnung beweisen. Dieser verknüpft auf verblüffend einfache Weise die Integral- mit der Differenzialrechung und ermöglicht so die explizite Berechnung von vielen Integralen, indem er unsere Erkenntnisse über die Differentiation zur Integralberechnung nutzbar macht.

Theorem 26.8 (Hauptsatz der Differenzial- und Integralrechnung)**.**
Sei $f : I \to \mathbb{K}$ eine stetige Funktion. Dann gelten die folgenden Aussagen.

(a) Sei $c \in I$ fest und $F_c(x) := \int_c^x f(s)\, \mathrm{d}s$ für jedes $x \in I$. Dann ist F_c differenzierbar auf I und $F_c'(x) = f(x)$ für alle $x \in I$.

(b) Ist $\Phi : I \to \mathbb{K}$ differenzierbar mit $\Phi'(x) = f(x)$ für jedes $x \in I$, dann gilt

$$\Phi(x) = \Phi(c) + \int_c^x f(s)\, \mathrm{d}s \qquad \text{für alle } c, x \in I.$$

Beweis. Die Hauptarbeit ist schon erledigt, wir müssen nur noch alles zusammensetzen.

(a) Ist bereits mit Satz 26.7 (b) erledigt.

(b) Sei F_a wie in (a) mit $c = a$. Dann gilt mit Hilfe von (a) und der Voraussetzung für jedes $x \in I$

$$(F_a - \Phi)'(x) = F_a'(x) - \Phi'(x) = f(x) - f(x) = 0.$$

Also gibt es eine Konstante $\alpha \in \mathbb{K}$ mit $F_a(x) = \Phi(x) + \alpha$. Damit erhalten wir schließlich für jede Wahl von c und x aus I

$$\begin{aligned}\int_c^x f(s)\,\mathrm{d}s &= \int_a^x f(s)\,\mathrm{d}s - \int_a^c f(s)\,\mathrm{d}s = F_a(x) - F_a(c) \\ &= \Phi(x) - \alpha - \Phi(c) + \alpha = \Phi(x) - \Phi(c),\end{aligned}$$

woraus durch Umstellen der Gleichung die Behauptung folgt. □

Nach Teil (b) des Hauptsatzes können wir den Wert eines Integrals über f leicht bestimmen, wenn wir eine Funktion Φ finden, für die $\Phi' = f$ gilt. Damit ist das Problem der Integration darauf zurück geführt den Vorgang der Differentiation umzukehren. Das ist leider leichter gesagt als getan, hilft aber schon oft weiter.

Definition 26.9. *Sei $f : I \to \mathbb{K}$ eine Funktion. Jede differenzierbare Funktion $F : I \to \mathbb{K}$ mit $F'(x) = f(x)$ für alle $x \in I$ heißt eine* Stammfunktion *von f.*

Mit diesem Begriff formulieren wir den Hauptsatz noch einmal leicht um.

Korollar 26.10. *(a) Ist $f : I \to \mathbb{K}$ stetig, so besitzt f eine Stammfunktion F auf I und es gilt*

$$\int_y^x f(s)\,\mathrm{d}s = F(x) - F(y) =: F(s)\Big|_{s=y}^{s=x} \qquad \textit{für alle } x, y \in I.$$

(b) Sind $F, \hat{F} : I \to \mathbb{K}$ Stammfunktionen einer Funktion $f : I \to \mathbb{K}$, so existiert eine Konstante $c \in \mathbb{K}$ mit $F(x) = \hat{F}(x) + c$ für alle $x \in I$.

Beweis. (a) Nach dem Hauptsatz 26.8 (a) ist die Funktion $F : I \to \mathbb{K}$ mit $F(x) := \int_a^x f(s)\,\mathrm{d}s$ eine Stammfunktion von f und aus Teil (b) des selben Satzes folgt

$$\int_y^x f(s)\,\mathrm{d}s = F(x) - F(y).$$

(b) Nach Satz 26.8 (b) muss für die beiden Stammfunktionen F und $\hat{F}$ von f für jedes $x \in I$ gelten

$$F(x) = F(a) + \int_a^x f(s)\,\mathrm{d}s \qquad \text{und} \qquad \hat{F}(x) = \hat{F}(a) + \int_a^x f(s)\,\mathrm{d}s.$$

Also ist $F(x) - \hat{F}(x) = F(a) - \hat{F}(a) =: c$ konstant auf I. □

Warnung 26.11. (a) Es gibt Funktionen, die eine Stammfunktion besitzen, aber nicht sprungstetig sind. Zur Konstruktion eines Beispiels betrachten wir auf $[0,1]$ die Funktion

$$F(x) = \begin{cases} x^{3/2}\sin(1/x), \text{ falls } x \in (0,1], \\ 0, \text{ falls } x = 0. \end{cases}$$

Von dieser haben wir in Beispiel 24.2 (d) gezeigt, dass sie auf $[0,1]$ differenzierbar ist, aber dass die Funktion $f := F'$ auf $[0,1]$ nicht beschränkt ist. Damit kann f auf $[0,1]$ nicht sprungstetig sein, aber F ist natürlich eine Stammfunktion von f.

(b) Andersherum gibt es auch sprungstetige Funktionen, die keine Stammfunktion besitzen. Als Beispiel dient uns hier auf $[-1,1]$ die Funktion

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

Diese ist als Treppenfunktion integrierbar. Wir nehmen nun an, es gäbe auf dem Intervall $[-1,1]$ eine Stammfunktion F von f. Dann gilt $F'(x) = f(x) = -1$ für alle $x \in [-1,0)$, also gibt es eine Konstante $c_1 \in \mathbb{R}$, so dass auf diesem Intervall $F(x) = -x + c_1$ gilt. Genauso gibt es eine Konstante $c_2 \in \mathbb{R}$, so dass für alle $x \in [0,1]$ die Identität $F(x) = x + c_2$ gilt. Da F als Stammfunktion in 0 differenzierbar sein muss, ist sie dort insbesondere stetig. Es gilt also

$$c_1 = \lim_{x\to 0-} F(x) = \lim_{x\to 0} F(x) = \lim_{x\to 0+} F(x) = c_2.$$

Also ist F für alle $x \in [-1,1]$ bestimmt als $F(x) = |x| + c$ mit einem $c \in \mathbb{R}$ und das kann nicht sein, denn die Betragsfunktion ist bekanntermaßen in Null nicht differenzierbar und damit kann es F auch nicht sein und wir haben einen Widerspruch.

Wir haben unser Integral auf den sprungstetigen Funktionen bekommen, indem wir es auf Treppenfunktionen definiert haben und dann gleichmäßige Grenzwerte von Treppenfunktionen betrachtet haben. Da wir nun für jede sprungstetige Funktion ein Integral haben, könnte man auf die Idee kommen, den selben Trick noch einmal zu machen und gleichmäßige Limiten von sprungstetigen Funktionen betrachten, in der Hoffnung, so für eine noch größere Klasse von Funktionen einen Integralbegriff definieren zu können.
Der nachfolgende Satz zeigt, dass das leider nichts wird, denn es stellt sich heraus, dass eine Funktion, die gleichmäßig durch sprungstetige Funktionen approximiert werden kann, selbst schon sprungstetig sein muss, wir bekommen also durch so ein Vorgehen keine neuen Funktionen mehr dazu.

Dahinter steckt ein allgemeines Prinzip. Spätestens nach dem Besuch einer Vorlesung in Topologie oder Funktionalanalysis im weiteren Studium werden Sie obige Hoffnung als naiv erkennen, aber im Moment spricht noch nichts gegen sie.

Satz 26.12. *Für jedes $n \in \mathbb{N}$ sei $f_n : I \to \mathbb{K}$ eine sprungstetige Funktion, so dass die Funktionenfolge (f_n) gleichmäßig auf I gegen eine Funktion f konvergiert. Dann ist f sprungstetig und es gilt*

$$\lim_{n\to\infty} \int_I f_n = \int_I \lim_{n\to\infty} f_n = \int_I f.$$

Bemerkung 26.13. (a) Die Bedeutung dieses Satzes geht weit über die Zerstörung der obigen Hoffnung hinaus. Wir haben es hier wieder mit einem Resultat zu tun, das uns das Vertauschen von zwei Grenzwertprozessen erlaubt, nämlich die Integration und den Grenzwert der Funktionenfolge. Auch hier zeigt sich wieder wie nützlich der Begriff der gleichmäßigen Konvergenz ist. Tatsächlich ist ein entsprechender Satz für nur punktweise konvergente Funktionenfolgen falsch.

(b) Beachten Sie, dass nach obigem Satz für gleichmäßig konvergente Funktionenreihen von sprungstetigen Funktionen

$$\sum_{n=0}^{\infty} \int_I f_n = \int_I \sum_{n=0}^{\infty} f_n$$

gilt.

Übungsaufgabe 26.14. Zeigen Sie durch ein Beispiel, dass Satz 26.12 für eine nur punktweise konvergente Funktionenfolge im Allgemeinen falsch ist.

Beweis von Satz 26.12. Für jedes $n \in \mathbb{N}$ ist f_n nach Voraussetzung sprungstetig, also gibt es jeweils eine Treppenfunktion φ_n auf I, für die $|\varphi_n(x) - f_n(x)| < 1/n$ für alle $x \in I$ gilt. Auf diese Weise erhalten wir eine Funktionenfolge (φ_n) von Treppenfunktionen auf I, von der wir nun zeigen wollen, dass sie ebenfalls gleichmäßig gegen f konvergiert.

Sei dazu $\varepsilon > 0$. Wir wählen nun ein $n_0 \in \mathbb{N}$ so groß, dass erstens $n_0 > 2/\varepsilon$ gilt und zweitens $|f_n(x) - f(x)| < \varepsilon/2$ ist für alle $n \geq n_0$ und alle $x \in I$. Letzteres geht, da (f_n) gleichmäßig gegen f konvergiert. Dann erhalten wir für alle $x \in I$ und alle $n \geq n_0$

$$\big|\varphi_n(x) - f(x)\big| \leq \big|\varphi_n(x) - f_n(x)\big| + \big|f_n(x) - f(x)\big| < \frac{1}{n} + \frac{\varepsilon}{2} \leq \varepsilon.$$

Somit ist die gleichmäßige Konvergenz von (φ_n) gegen f auf I gezeigt. Damit folgt sofort, dass f sprungstetig ist, also existiert $\int_I f$. Mit diesem Wissen haben wir nun gewonnen, denn mit Hilfe der Standardabschätzung gilt

$$\Big|\int_I f_n - \int_I f\Big| = \Big|\int_I (f_n - f)\Big| \leq (b-a) \sup_{x\in I} \big|f_n(x) - f(x)\big|$$

und letzterer Ausdruck geht nach Übungsaufgabe 20.9 (b) gegen Null. Also gilt

$$\lim_{n\to\infty} \int_I f_n = \int_I f$$

und wir sind fertig. $\square$

Als Anwendung dieses Ergebnisses wollen wir abschließend einen Satz beweisen, der thematisch ins Kapitel 20 gehört, der sich allerdings schön mit Hilfe der Integralrechnung beweisen lässt. Deshalb liefern wir ihn hier nach.

Satz 26.15. *Es sei (f_n) eine Funktionenfolge auf $I = [a, b]$ mit $f_n \in C^1(I)$ für alle $n \in \mathbb{N}$. Konvergiert die Funktionenfolge $(f_n')_{n\in\mathbb{N}}$ gleichmäßig auf I gegen eine Funktion $g : I \to \mathbb{K}$ und ist die Zahlenfolge $(f_n(x_0))_{n\in\mathbb{N}}$ für ein $x_0 \in I$ konvergent, so konvergiert auch die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ gleichmäßig auf I gegen eine Funktion $f \in C^1(I)$ und es gilt $f' = g$, d. h.*

$$\big(\lim_{n\to\infty} f_n\big)' = \lim_{n\to\infty} f_n'.$$

Beweis. Wir setzen $c := \lim_{n\to\infty} f_n(x_0)$. Nach dem Hauptsatz gilt nun für alle $x \in I$ und alle $n \in \mathbb{N}$

$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(t)\, \mathrm{d}t.$$

Da (f_n') auf dem Intervall zwischen x_0 und x eine gleichmäßig konvergente Funktionenfolge ist und die Funktionen f_n' für jedes $n \in \mathbb{N}$ stetig und damit insbesondere sprungstetig sind, gilt nach Satz 26.12

$$\lim_{n\to\infty} \int_{x_0}^{x} f_n'(t)\, \mathrm{d}t = \int_{x_0}^{x} g(t)\, \mathrm{d}t.$$

Also ist (f_n) auf I punktweise konvergent und für die Grenzfunktion f gilt

$$f(x) = \lim_{n\to\infty} f_n(x) = c + \int_{x_0}^{x} g(t)\, \mathrm{d}t$$

für jedes $x \in I$.
Weiter ist nach dem Hauptsatz die Abbildung $x \mapsto \int_{x_0}^{x} g(t)\, \mathrm{d}t$ differenzierbar und für die Ableitung gilt

$$\Big(\int_{x_0}^{x} g(t)\, \mathrm{d}t\Big)' = g(x).$$

Da g ein gleichmäßiger Limes der nach Voraussetzung stetigen Funktionen f_n, $n \in \mathbb{N}$ ist, ist auch g eine stetige Funktion. Das bedeutet weiter, dass $f' = g$ auf I

stetig ist, d. h. $f \in C^1(I)$. Es bleibt noch die gleichmäßige Konvergenz von (f_n) auf I zu zeigen. Dazu beobachten wir für jedes $x \in I$

$$\begin{aligned} |f_n(x) - f(x)| &= \Big| \int_{x_0}^{x} f_n'(t)\,\mathrm{d}t - \int_{x_0}^{x} g(t)\,\mathrm{d}t + f_n(x_0) - c \Big| \\ &\leq \Big| \int_{x_0}^{x} \big(f_n'(t) - g(t)\big)\,\mathrm{d}t \Big| + |f_n(x_0) - c| \\ &\leq \int_{x_0}^{x} |f_n'(t) - g(t)|\,\mathrm{d}t + |f_n(x_0) - c| \\ &\leq \int_{a}^{b} |f_n'(t) - g(t)|\,\mathrm{d}t + |f_n(x_0) - c|. \end{aligned}$$

Sein nun $\varepsilon > 0$ vorgegeben. Dann existiert zum einen dank der gleichmäßigen Konvergenz von (f_n') ein $n_1 \in \mathbb{N}$, so dass für alle $n \geq n_1$ und alle $t \in I$

$$\big|f_n'(t) - g(t)\big| < \frac{\varepsilon}{2(b-a)}$$

gilt. Zum anderen gibt es ein $n_2 \in \mathbb{N}$ mit $|f_n(x_0) - c| < \varepsilon/2$ für alle $n \geq n_2$. Wählen wir nun $n_0 := \max\{n_1, n_2\}$, so gilt für alle $n \geq n_0$ mit der Abschätzung von oben

$$\big|f_n(x) - f(x)\big| \leq \int_a^b \big|f_n'(t) - g(t)\big|\,\mathrm{d}t + |f_n(x_0) - c| \leq (b-a)\frac{\varepsilon}{2(b-a)} + \frac{\varepsilon}{2} = \varepsilon.$$

Da wir n_0 unabhängig von x wählen konnten, ist damit die Gleichmäßigkeit der Konvergenz bewiesen. □

27. Integrationsregeln

Zur Integration von Funktionen ist das Auffinden von Stammfunktionen von zentraler Bedeutung. Leider gibt es dazu nicht wie bei der Differenziation einen kompletten Satz von Regeln, mit dessen Hilfe, genug Zeit und Konzentration vorausgesetzt, im Prinzip jede Funktion differenziert werden kann. Statt dessen müssen wir uns mit Rechenregeln begnügen, die meist das Problem der Integration einer Funktion auf das entsprechende Problem für eine andere Funktion zurückspielen, die dann hoffentlich einfacher ist.
Das liegt nicht daran, dass uns im Moment noch starke mathematische Hilfsmittel fehlen, sondern ist ein prinzipielles Problem. Es gibt einfache stetige (sogar beliebig oft differenzierbare) Funktionen, die nach dem Hauptsatz eine Stammfunktion haben, die aber nicht in einer geschlossenen Form angebbar ist.
Zusammengefasst ist dies in dem Spruch:

Differenzieren ist Handwerk, Integrieren ist Kunst.

Wir wollen uns dieser Kunst nun nähern, indem wir aus den bekannten Differenziationsregeln, Rechenregeln für Integrale ableiten. Wir beginnen mit der Produktregel.
Auch in diesem Kapitel steht wieder $\mathbb{K}$ für die mögliche Wahl von $\mathbb{R}$ oder $\mathbb{C}$.

Satz 27.1 (Partielle Integration)**.** *Es seien $I := [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall und $f, g : I \to \mathbb{K}$ stetig differenzierbare Funktionen. Dann gilt*

$$\int_a^b f'g = f(b)g(b) - f(a)g(a) - \int_a^b fg' = fg\,\Big|_a^b - \int_a^b fg'.$$

Beweis. Zunächst einmal existieren alle in der Behauptung auftretenden Integrale, denn nach Voraussetzung sind $f'g$ und fg' stetige Funktionen und damit auch sprungstetig auf I.
Nach der Produktregel gilt nun

$$(fg)' = f'g + fg'.$$

Also haben wir mit dem Hauptsatz der Differenzial- und Integralrechnung, Theorem 26.8,

$$\int_a^b \big(f'g + fg'\big) = \int_a^b (fg)' = fg\,\Big|_a^b, \quad \text{d. h.} \quad \int_a^b f'g = fg\,\Big|_a^b - \int_a^b fg'$$

und damit die Behauptung. □

Beispiel 27.2. (a) Wir betrachten das Integral

$$\int_0^1 x\mathrm{e}^x \,\mathrm{d}x,$$

d. h. wir wenden unseren Satz mit $g(x) = x$ und $f'(x) = \mathrm{e}^x$ auf dem Intervall $[0,1]$ an. Dann ist $f(x) = \mathrm{e}^x$ eine mögliche Wahl für die Funktion f und wir erhalten mit partieller Integration:

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

(b) Die Wahl von f und g kann für den Erfolg einer Anwendung dieser Regel sehr entscheidend sein. Wenn wir beispielsweise im Integral aus (a) umgekehrt $g(x) = \mathrm{e}^x$ und $f'(x) = x$ genommen hätten, wären wir bei

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

gelandet. Diese Umformung ist natürlich auch richtig, aber von dem nun entstandenen Integral weiß man erst recht nicht, wie man es berechnen soll.

(c) Manchmal muss man sich die zweite Funktion zur partiellen Integration erst künstlich schaffen:

$$\begin{aligned}\int_1^2 \ln(x) \,\mathrm{d}x = \int_1^2 1 \cdot \ln(x) \,\mathrm{d}x &= x\ln(x)\Big|_1^2 - \int_1^2 x\frac{1}{x} \,\mathrm{d}x = \big(x\ln(x) - x\big)\Big|_1^2 \\ &= 2 \cdot \ln(2) - 2 - 0 + 1 = 2 \cdot \ln(2) - 1,\end{aligned}$$

wobei wir $g(x) = \ln(x)$ und $f'(x) = 1$ gewählt haben.

(d) Wir wollen

$$I := \int_0^{\pi/2} \sin^2(x) \,\mathrm{d}x$$

bestimmen. Dazu wählen wir $f'(x) = g(x) = \sin(x)$ und berechnen

$$I = -\cos(x)\sin(x)\Big|_0^{\pi/2} - \int_0^{\pi/2} \cos(x)(-\cos(x)) \,\mathrm{d}x = \int_0^{\pi/2} \cos^2(x) \,\mathrm{d}x.$$

Wenden wir nun mit $f'(x) = g(x) = \cos(x)$ noch einmal partielle Integration an, so erhalten wir

$$I = \sin(x)\cos(x)\Big|_0^{\pi/2} - \int_0^{\pi/2} \sin(x)(-\sin(x)) \,\mathrm{d}x = \int_0^{\pi/2} \sin^2(x) \,\mathrm{d}x = I$$

und damit außer der Gewissheit, dass wir uns unterwegs nicht verrechnet haben, nichts neues. Wir müssen also einen anderen Weg suchen: Mit dem Ergebnis unserer ersten partiellen Integration und dem trigonometrischen Pythagoras $\sin^2(x) + \cos^2(x) = 1$, finden wir

$$I = \int_0^{\pi/2} \cos^2(x)\,\mathrm{d}x = \int_0^{\pi/2} (1-\sin^2(x))\,\mathrm{d}x = \frac{\pi}{2} - \int_0^{\pi/2} \sin^2(x)\,\mathrm{d}x = \frac{\pi}{2} - I,$$

woraus $2I = \pi/2$ und schließlich

$$I = \int_0^{\pi/2} \sin^2(x)\,\mathrm{d}x = \frac{\pi}{4}$$

folgt.

Die zweite wichtige Integrationsregel ergibt sich aus der Kettenregel der Differenzialrechnung.

Satz 27.3 (Substitutionsregel)**.** *Es seien $[a, b] \subseteq \mathbb{R}$ und $[c, d] \subseteq \mathbb{R}$ kompakte Intervalle, sowie $f \in C([a, b])$ und $g \in C^1([c, d])$ mit $g([c, d]) \subseteq [a, b]$. Dann ist*

$$\int_c^d f(g(t)) \cdot g'(t)\,\mathrm{d}t = \int_{g(c)}^{g(d)} f(x)\,\mathrm{d}x.$$

Beweis. Nach Korollar 26.10 besitzt f auf $[a, b]$ eine Stammfunktion F. Wir betrachten die Funktion $H := F \circ g$ auf $[c, d]$. Dann gilt für alle $t \in [c, d]$ nach der Kettenregel

$$H'(t) = F'(g(t)) \cdot g'(t) = f(g(t)) \cdot g'(t).$$

Also können wir mit zweimaliger Anwendung des Hauptsatzes folgern:

$$\int_c^d f(g(t)) \cdot g'(t)\,\mathrm{d}t = H(d) - H(c) = F(g(d)) - F(g(c)) = \int_{g(c)}^{g(d)} f(x)\,\mathrm{d}x. \quad \square$$

Bemerkung 27.4. Häufig behilft man sich bei der Anwendung der Substitutionsregel einer intuitiven, aber nicht rigorosen Schreibweise. Diese leitet sich aus der alternativen Notation $\frac{\mathrm{d}y}{\mathrm{d}x}$ (gesprochen „dy nach dx“) statt y' für eine differenzierbare Funktion y ab. Man fasst dann in der Substitutionsregel die Setzung $x = g(t)$ so auf, als sei x eine Funktion von t und rechnet mit den Differenzialen dx und dy wie gewohnt:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = g'(t) \quad „ \Rightarrow \mathrm{d}x = g'(t)\,\mathrm{d}t. “$$

Dabei erhält man genau die in der Substitionsformel stehende Ersetzung von dx durch $g'(t)\mathrm{d}t$.

Dieser Formalismus ist sehr übersichtlich und praktisch, es sollte dabei aber nicht in Vergessenheit geraten, dass das *keine* saubere Mathematik ist.

Beispiel 27.5. (a) Wir berechnen das Integral

$$\int_1^2 \frac{\mathrm{e}^{2x}+1}{\mathrm{e}^x}\,\mathrm{d}x$$

mit unserer Schmierrechnungsmethode. Dazu setzen wir $x = \ln(t)$, d. h. wir wenden die Substitutionsregel mit $g(t) = \ln(t)$ an. Weiter ist bei der Anwendung des Satzes $c = \mathrm{e}$ und $d = \mathrm{e}^2$, denn dann ist $g(c) = 1$ und $g(d) = 2$. Die natürliche Wahl für $[a, b]$ ist $[1, 2]$, aber auch $[a, b] = [-3, 15]$ ist in Ordnung. Nun wenden wir die Substitutionsregel an. Es ist $\frac{\mathrm{d}x}{\mathrm{d}t} = 1/t$, also „$\mathrm{d}x = \mathrm{d}t/t$“. Damit haben wir

$$\begin{aligned}\int_1^2 \frac{\mathrm{e}^{2x}+1}{\mathrm{e}^x}\,\mathrm{d}x &= \int_{\mathrm{e}}^{\mathrm{e}^2} \frac{t^2+1}{t}\cdot\frac{\mathrm{d}t}{t} = \int_{\mathrm{e}}^{\mathrm{e}^2}\Big(1+\frac{1}{t^2}\Big)\,\mathrm{d}t = t - \frac{1}{t}\Big|_{\mathrm{e}}^{\mathrm{e}^2}\\ &= \mathrm{e}^2 - \mathrm{e}^{-2} - (\mathrm{e} - \mathrm{e}^{-1}).\end{aligned}$$

(b) Als zweites Beispiel wollen wir das Integral

$$I := \int_0^1 \sqrt{1-x^2}\,\mathrm{d}x$$

bestimmen. Dieses hat auch eine anschauliche Bedeutung, denn der Graph der Funktion $\sqrt{1-x^2}$ ist für $x \in [0, 1]$ der Viertelkreisbogen des Kreises mit Radius 1 um 0 zwischen den Punkten $(1, 0)$ und $(0, 1)$. Wir bestimmen mit diesem Integral also die Fläche dieses Viertelkreises, es sollte also, bitteschön, $\pi/4$ herauskommen.

Wir substituieren $x = \cos(t)$. Dann gilt z. B. $x = 0$ für $t = \pi/2$ und $x = 1$ für $t = 0$. Wir wählen also $c = \pi/2$ und $d = 0$. Die Schmierrechnung gibt uns wegen $\frac{\mathrm{d}x}{\mathrm{d}t} = \cos'(t) = -\sin(t)$ die Ersetzung $\mathrm{d}x = -\sin(t)\,\mathrm{d}t$. Nun gilt für alle $t \in [0, \pi/2]$

$$\sqrt{1-x^2} = \sqrt{1-\cos^2(t)} = \sqrt{\sin^2(t)} = |\sin(t)| = \sin(t).$$

Setzen wir das nun alles zusammen, ergibt sich mit Beispiel 27.2 (d) tatsächlich

$$I = \int_{\frac{\pi}{2}}^0 \sin(t)(-\sin(t))\,\mathrm{d}t = \int_0^{\frac{\pi}{2}} \sin^2(t)\,\mathrm{d}t = \frac{\pi}{4}.$$

Im Verlauf der letzten Abschnitte hat sich aus unserer ursprünglichen Motivation für die Integralrechnung, nämlich die der Flächenberechnung, zunehmend die abstrakte Fragestellung nach Bestimmung einer Stammfunktion ergeben. Der Integralkalkül kann natürlich auch von Anfang an so motiviert werden. Dann gelangt man zum unbestimmten Integral, das wir nun einführen wollen.

Definition 27.6. *Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{K}$. Besitzt f auf I eine Stammfunktion, so schreibt man für diese auch das sogenannte* unbestimmte Integral

$$\int f \qquad \textit{oder} \qquad \int f(x)\, \mathrm{d}x.$$

Man beachte dabei, dass nun das Symbol $\int f$ eine *Funktion* bezeichnet, während das bestimmte Integral $\int_a^b f$ für vorgegebene $a, b \in \mathbb{R}$ eine *Zahl* ist.

Beispiel 27.7. (a) $\displaystyle\int \mathrm{e}^x \,\mathrm{d}x = \mathrm{e}^x + c, \quad c \in \mathbb{R}.$

(b) $\displaystyle\int \sin(x)\,\mathrm{d}x = -\cos(x) + c, \quad c \in \mathbb{R}.$

(c) Man beachte, dass das unbestimmte Integral auch für Funktionen Sinn machen kann, die überhaupt nicht integrierbar sind. So gilt zum Beispiel für die Funktion

$$f(x) = \begin{cases} \frac{3}{2}\sqrt{x}\sin(1/x) - \frac{1}{\sqrt{x}}\cos(1/x), & \text{falls } x > 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

wegen Beispiel 24.2 (d)

$$\int f(x)\,\mathrm{d}x = \left.\begin{cases} x^{3/2} \cdot \sin(1/x), & \text{falls } x > 0, \\ 0, & \text{falls } x = 0, \end{cases}\right\} + c, \quad c \in \mathbb{R}$$

aber f ist für jedes $b > 0$ auf dem Intervall $[0, b]$ nicht integrierbar (vgl. Warnung 26.11 (a)).

Wir können nun unsere Rechenregeln für Integrale auch für die unbestimmten Integrale formulieren. Zuvor ist jedoch noch die folgende Beoachtung interessant.

Bemerkung 27.8. Auf einem kompakten Intervall $[a, b] \subseteq \mathbb{R}$ besitzt nach Korollar 26.10 jede stetige Funktion eine Stammfunktion. Dies gilt tatsächlich sogar für jedes beliebige Intervall $I \subseteq \mathbb{R}$. Dazu wählen wir ein beliebiges $\xi \in I^\circ$ aus und setzen für jedes $x \in I$

$$F(x) = \int_\xi^x f(t)\,\mathrm{d}t.$$

Zum Nachweis, dass nun F tatsächlich eine Stammfunktion von f ist, unterscheiden wir die Fälle $x \geq \xi$ und $x < \xi$. Ist $x \geq \xi$, so wählen wir ein $b \geq x$ (Im Falle $x = \xi$, bitte nicht auch noch $b = x$, was sich dank unserer Wahl von ξ im Inneren des Intervalls zum Glück vermeiden lässt). Nun ist nach dem Hauptsatz die Funktion F auf dem Intervall $[\xi, b]$ eine Stammfunktion von f, also ist $F'(x) = f(x)$ für all diese x.
Der Fall $x < \xi$ geht analog.

Im Lichte dieser Bemerkung können wir also die folgenden Betrachtungen auf beliebigen Intervallen in $\mathbb{R}$ anstellen.

Satz 27.9 (partielle Integration)**.** *Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f, g : I \to \mathbb{K}$ seien differenzierbare Funktionen auf I. Besitzt dann die Funktion fg' auf I eine Stammfunktion, so besitzt auch $f'g$ dort eine Stammfunktion und es gilt*

$$\int f'g = fg - \int fg'$$

Beweis. Es gilt nach der Produktregel $(fg)' = f'g + fg'$, also ist $f'g = (fg)' - fg'$. Weiter besitzt fg' nach Voraussetzung eine Stammfunktion F und fg ist natürlich eine Stammfunktion von $(fg)'$. Also ist $fg - F$ eine Stammfunktion von $f'g$, was nichts anderes bedeutet als

$$\int f'g = fg - F + c = fg - \int fg'. \qquad \square$$

Beispiel 27.10. (a) $\int \ln(x)\,\mathrm{d}x = \int 1 \cdot \ln(x)\,\mathrm{d}x = x \cdot \ln(x) - \int x \cdot \frac{1}{x}\,\mathrm{d}x = x \cdot \ln(x) - \int 1\,\mathrm{d}x = x \cdot \ln(x) - x + c, \quad c \in \mathbb{R}.$

(b) $\int x\mathrm{e}^x\,\mathrm{d}x = x\mathrm{e}^x - \int \mathrm{e}^x\,\mathrm{d}x = x\mathrm{e}^x - \mathrm{e}^x + c, \quad c \in \mathbb{R}.$

Satz 27.11 (Substitutionsregeln für unbestimmte Integrale)**.**
Es seien $I, J \subseteq \mathbb{R}$ Intervalle und $f \in C(I)$, sowie $g \in C^1(J)$ seien Funktionen mit $g(J) = I$. Dann gilt

(a) $\int f(g(t)) \cdot g'(t)\,\mathrm{d}t = \int f(x)\,\mathrm{d}x\Big|_{x=g(t)}$ *auf J.*

(b) Ist $g'(t) \neq 0$ für alle $t \in J$, so gilt

$$\int f(x)\,\mathrm{d}x = \int f(g(t)) \cdot g'(t)\,\mathrm{d}t\Big|_{t=g^{-1}(x)}.$$

Bemerkung 27.12. (a) Die Notation $|_{x=g(t)}$ bedeutet, dass man zunächst den davor stehenden Ausdruck bestimmt, und dann überall $g(t)$ für x einsetzt.

(b) Man beachte, dass die Voraussetzung $g'(t) \neq 0$ für alle $t \in J$ in Teil (b) des Satzes, dank der Stetigkeit der Ableitung impliziert, dass f' entweder überall strikt positiv oder überall strikt negativ ist. Auf jeden Fall ist also g streng monoton auf I und damit existiert die Umkehrfunktion g^{-1} auf I, die ja in der dann folgenden Aussage auch verwendet wird.

Beweis. Nach Bemerkung 27.8 hat f auf I eine Stammfunktion F. Damit setzen wir $H := F \circ g$. Man beachte, dass damit für alle $t \in J$

$$H(t) = F(g(t)) = \int f(x)\,\mathrm{d}x\Big|_{x=g(t)}$$

gilt. Außerdem gilt nach der Kettenregel für alle $t \in J$

$$H'(t) = F'(g(t)) \cdot g'(t) = f(g(t)) \cdot g'(t) \tag{27.1}$$

woraus bereits die erste Aussage folgt. Für die zweite Aussage machen wir uns klar, dass

$$H \circ g^{-1} = F \circ g \circ g^{-1} = F$$

gilt. Also ist wieder mit Hilfe von (27.1)

$$\begin{aligned}\int f(x)\,\mathrm{d}x &= F(x) + c = H(g^{-1}(x)) + c = H(t)\big|_{t=g^{-1}(x)} + c \\ &= \int f(g(t)) \cdot g'(t)\,\mathrm{d}t\Big|_{t=g^{-1}(x)}.\end{aligned}$$ □

Beispiel 27.13. (a) Wir berechnen auf dem Intervall $(0, \infty)$

$$h(t) := \int \frac{t+5}{t^2+10t+4}\,\mathrm{d}t = \frac{1}{2}\int \frac{2t+10}{t^2+10t+4}\,\mathrm{d}t.$$

Dazu setzen wir $f(x) = 1/x$ und $g(t) = t^2 + 10t + 4$. Dann ist

$$h(t) = \frac{1}{2}\int \frac{g'(t)}{g(t)}\,\mathrm{d}t = \frac{1}{2}\int f(g(t)) \cdot g'(t)\,\mathrm{d}t.$$

Also haben wir mit der Substitutionsregel

$$h(t) = \frac{1}{2}\int f(x)\,\mathrm{d}x\Big|_{x=g(t)} = \frac{1}{2} \cdot \ln(x)\big|_{x=t^2+10t+4} = \frac{1}{2} \cdot \ln(t^2 + 10t + 4).$$

(b) Auf dem Intervall $(0, 1)$ betrachten wir das unbestimmte Integral (vgl. Beispiel 27.5 (b))

$$\int \sqrt{1-x^2}\,\mathrm{d}x.$$

An diesem Beispiel werden wir sehen, wie die verschiedenen Integrationsregeln zusammenwirken können, denn die Stammfunktion hiervon wird sich als einigermaßen anstrengend zu berechnen erweisen.

Wir setzen zunächst $x = g(t) = \sin(t)$ für $t \in J := (0, \pi/2)$. Dann ist $\frac{\mathrm{d}x}{\mathrm{d}t} = \cos(t)$, also $\mathrm{d}x = \cos(t)\mathrm{d}t$ und damit haben wir

$$\int \sqrt{1-x^2}\,\mathrm{d}x = \int \sqrt{\cos^2(t)}\cos(t)\,\mathrm{d}t\Big|_{t=\arcsin(x)}.$$

Da $t \in (0, \pi/2)$ ist, ist das Vorzeichen des Cosinus hier immer positiv, so dass tatsächlich $\sqrt{\cos^2(t)} = \cos(t)$ gilt. Das nun auftretende Integral über $\cos^2(t)$ berechnen wir wieder wie in Beispiel 27.2 (d) mit partieller Integration:

$$\begin{aligned}\int \cos^2(t)\,\mathrm{d}t &= \sin(t)\cos(t) + \int \sin^2(t) = \sin(t)\cos(t) + \int (1-\cos^2(t))\,\mathrm{d}t \\ &= \sin(t)\cos(t) + t - \int \cos^2(t)\,\mathrm{d}t,\end{aligned}$$

woraus mit der Rücksubstitution $t = \arcsin(x)$ (man beachte, dass unser Intervall $(0,1)$ im Definitionsbereich des Arcussinus liegt) folgt:

$$\begin{aligned}2\int \sqrt{1-x^2}\,\mathrm{d}x = 2\int \cos^2(t)\,\mathrm{d}t &= \sin(t)\cos(t) + t + c \\ &= \sin(\arcsin(x))\cos(\arcsin(x)) + \arcsin(x) + c \\ &= x\cos(\arcsin(x)) + \arcsin(x) + c.\end{aligned}$$

Das können wir noch ein wenig vereinfachen, denn wegen $\cos^2(y)+\sin^2(y) = 1$ gilt (mit $y = \arcsin(x)$)

$$\cos^2(\arcsin(x)) = 1-\sin^2(\arcsin(x)) = 1-x^2, \text{ also } \cos(\arcsin(x)) = \sqrt{1-x^2}.$$

Damit haben wir nun das Endergebnis

$$\int \sqrt{1-x^2}\,\mathrm{d}x = \frac{1}{2}\big(x\sqrt{1-x^2} + \arcsin(x)\big) + c \quad \text{auf } (0,1).$$

Zum Abschluss dieses Abschnitts wollen wir nun einen weiteren Beweis für den Satz von Taylor mit Methoden aus der Integrationstheorie angeben. Auf diese Weise bekommen wir als neue Erkenntnis eine alternative Darstellung des Restgliedes mit Hilfe eines Integrals. Da der schwierigste Teil bei der Anwendung dieses Satzes meist die Abschätzung des Restgliedes ist, ist es natürlich praktisch möglichst viele Darstellungen für dieses zu haben.

Satz 27.14 (Satz von Taylor (mit Integralrestglied))**.** *Es seien $I \subseteq \mathbb{R}$ ein Intervall, $f \in C^{n+1}(I)$ für ein $n \in \mathbb{N}$ und $a, b \in I$. Dann gilt*

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

Beweis. Wir führen einen Induktionsbeweis. Für $n = 0$ ist die Behauptung

$$f(b) = f(a) + \int_a^b f'(x)\,\mathrm{d}x,$$

was gerade der Hauptsatz der Differenzial- und Integralrechnung ist. Gilt die Formel nun für ein $n \in \mathbb{N}$, so haben wir für $f \in C^{n+2}(I)$ nach Induktionsvoraussetzung

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

Mit partieller Integration folgt daraus

$$\begin{aligned} f(b) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \Big(\frac{1}{n+1} \frac{-(b-x)^{n+1}}{n!} f^{(n+1)}(x)\Big)\Big|_a^b \\ &\qquad + \int_a^b \frac{(b-x)^{n+1}}{(n+1)!} f^{(n+2)}(x) \,\mathrm{d}x \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{f^{(n+1)}}{(n+1)!}(b-a)^{n+1} + \int_a^b \frac{(b-x)^{n+1}}{(n+1)!} f^{(n+2)}(x) \,\mathrm{d}x \\ &= \sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!}(b-a)^k + \int_a^b \frac{(b-x)^{n+1}}{(n+1)!} f^{(n+2)}(x) \,\mathrm{d}x. \end{aligned}$$

□

28. Uneigentliche Integrale

Bisher können wir Integrale nur über kompakte Intervalle und sprungstetige, d. h. insbesondere beschränkte Funktionen bilden. Wir wollen unser mächtiges Werkzeug des Grenzübergangs jetzt auch hier verwenden, um etwas allgemeinere Integrale zuzulassen.
In diesem Abschnitt seien stets $a, b \in \mathbb{R}$ und $\alpha, \beta \in \mathbb{R} \cup \{-\infty, \infty\}$.

Definition 28.1. *Es sei $f : [a, \beta) \to \mathbb{K}$ (bzw. $f : (\alpha, b] \to \mathbb{K}$) sprungstetig auf dem Intervall $[a, t]$ (bzw. $[t, b]$) für jedes $t \in (a, \beta)$ (bzw. $t \in (\alpha, b)$). Dann heißt f* uneigentlich integrierbar, *wenn der Grenzwert*

$$\lim_{t \to \beta-} \int_a^t f \qquad \Big(\textit{bzw.} \quad \lim_{t \to \alpha+} \int_t^b f\Big)$$

existiert. In diesem Fall heißt das uneigentliche Integral

$$\int_a^\beta f := \lim_{t \to \beta-} \int_a^t f \qquad \Big(\textit{bzw.} \quad \int_\alpha^b f := \lim_{t \to \alpha+} \int_t^b f\Big)$$

konvergent. *Sonst nennt man es* divergent.

Beispiel 28.2. (a) Wir betrachten

$$\int_0^1 \frac{1}{\sqrt{1 - x^2}} \, \mathrm{d}x.$$

Das ist ein uneigentliches Integral, denn die Funktion $1/\sqrt{1 - x^2}$ ist auf $[0, 1]$ wegen $\lim_{x \to 1-} 1/\sqrt{1 - x^2} = \infty$ nicht beschränkt. Für jedes $t \in (0, 1)$ ist sie aber stetig auf dem Intervall $[0, t]$, also dort insbesondere sprungstetig. Wir haben damit im Sinne der obigen Defintion den Fall $a = 0$ und $\beta = 1$. Dann ist für jedes $t \in (0, 1)$

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

und wegen $\lim_{t \to 1-} \arcsin(t) = \pi/2$ ist das uneigentliche Integral konvergent und wir haben

$$\int_0^1 \frac{1}{\sqrt{1 - x^2}} \, \mathrm{d}x = \lim_{t \to 1-} \int_0^t \frac{1}{\sqrt{1 - x^2}} \, \mathrm{d}x = \lim_{t \to 1-} \arcsin(t) = \frac{\pi}{2}.$$

(b) Während im ersten Beispiel die Funktion unbeschränkt war, schauen wir uns nun eine Integration über ein unbeschränktes Intervall an:

$$\int_0^\infty \frac{1}{1+x^2}\,\mathrm{d}x,$$

es ist also $a = 0$ und $\beta = \infty$. Für $t \in (0, \infty)$ gilt nun

$$\int_0^t \frac{1}{1+x^2}\,\mathrm{d}x = \arctan(x)\Big|_0^t = \arctan(t) \longrightarrow \frac{\pi}{2} \quad (t \to \infty),$$

also ist auch dieses uneigentliche Integral konvergent und es ist

$$\int_0^\infty \frac{1}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2}.$$

Genauso sieht man

$$\int_{-\infty}^0 \frac{1}{1+x^2} = \frac{\pi}{2}.$$

(c) Es sei $s > 0$. Wann ist die Funktion $1/x^s$ auf dem Intervall $[1, \infty)$ uneigentlich integrierbar? Für $t \in (1, \infty)$ gilt für $s = 1$

$$\int_1^t \frac{1}{x}\,\mathrm{d}x = \ln(x)\Big|_1^t = \ln(t),$$

also ist das uneigentliche Integral in diesem Fall wegen $\lim_{t\to\infty} \ln(t) = \infty$ divergent.

Für $s \neq 1$ ist

$$\int_1^t \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{1-s}x^{1-s}\Big|_1^t = \frac{1}{1-s}(t^{1-s} - 1).$$

Der Grenzwert dieses Ausdrucks existiert nun genau für $s > 1$ und es ist in diesem Fall

$$\int_1^\infty \frac{1}{x^s}\,\mathrm{d}x = \lim_{t\to\infty} \frac{1}{1-s}(t^{1-s} - 1) = -\frac{1}{1-s} = \frac{1}{s-1}.$$

(d) Genauso wie im vorherigen Beispiel kann man zeigen, dass das uneigentliche Integral

$$\int_0^1 \frac{1}{x^s}\,\mathrm{d}x$$

genau dann konvergiert, wenn $s < 1$ ist. In diesem Fall gilt

$$\int_0^1 \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{s-1}.$$

Bisher haben wir nur uneigentliche Integrale betrachtet, die an einer Grenze uneigentlich sind. Natürlich will man auch den Fall behandeln, dass es an beiden Intervallgrenzen Probleme gibt, man spricht dann oft von einem doppelt uneigentlichen Integral. Dazu müssen wir unsere Definition modifizieren.

Definition 28.3. *Es sei $f : (\alpha, \beta) \to \mathbb{K}$ sprungstetig auf jedem Intervall $[\xi, \eta] \subseteq (\alpha, \beta)$. Dann heißt f auf (α, β)* uneigentlich integrierbar*, wenn es ein $c \in (\alpha, \beta)$ gibt, so dass die beiden uneigentlichen Integrale*

$$\int_\alpha^c f \quad \textit{und} \quad \int_c^\beta f$$

im Sinne von Defintion 28.1 konvergieren. In diesem Fall heißt das uneigentliche Integral

$$\int_\alpha^\beta f := \int_\alpha^c f + \int_c^\beta f$$

konvergent.

Natürlich muss man, damit diese Definition Sinn macht, zeigen, dass der so erhaltene Wert für das uneigentliche Integral nicht von der speziellen Wahl von c abhängt:

Übungsaufgabe 28.4. Definition 28.3 ist von der Wahl von $c \in (\alpha, \beta)$ unabhängig.

Beispiel 28.5. (a) Es ist mit Hilfe von Beispiel 28.2 (b) das doppelt uneigentliche Integral

$$\int_{-\infty}^\infty \frac{1}{1+x^2} \, \mathrm{d}x$$

konvergent und gleich π.

(b) Sei $s > 0$. Kombiniert man (c) und (d) aus Beispiel 28.2, so sieht man, dass das doppelt uneigentliche Integral

$$\int_0^\infty \frac{1}{x^s} \, \mathrm{d}x$$

genau dann konvergiert, wenn $s > 1$ und $s < 1$ gilt, d. h. es ist immer divergent.

Die folgenden Sätze und Definitionen formulieren wir der Übersichtlichkeit halber nur für uneigentliche Integrale der Form $\int_a^\beta f(x) \, \mathrm{d}x$. Dabei sei stets vorausgesetzt, dass f für jedes $t \in (\alpha, \beta)$ auf $[a, t]$ sprungstetig ist. Entsprechende Sätze und Defintionen gelten auch für die anderen Arten uneigentlicher Integrale, wobei bei doppelt uneigentlichen Integralen immer darauf geachtet werden muss, dass an beiden Grenzen unabhängig voneinander Konvergenz vorliegt.

Auf Beweise der nächsten Sätze verzichten wir weitgehend, da diese den Beweisen der entsprechenden Aussagen für Reihen nachgebildet werden können. Dieses für den einen oder anderen Beweis zu tun, wird als Übung aber sehr empfohlen.
Im Folgenden seien jeweils $f, g : [a, \beta) \to \mathbb{K}$ Funktionen, die für jedes $t \in (a, \beta)$ auf dem Intervall $[a, t]$ sprungstetig sind.

Satz 28.6 (Cauchy-Kriterium). *Das uneigentliche Integral $\int_a^\beta f$ ist genau dann konvergent, wenn es für jedes $\varepsilon > 0$ ein $c = c(\varepsilon) \in (a, \beta)$ gibt, so dass*

$$\left| \int_u^v f \right| < \varepsilon \quad \text{für alle } u, v \in (c, \beta)$$

gilt.

Definition 28.7. *Das uneigentliche Integral $\int_a^\beta f$ heißt* absolut konvergent*, wenn das uneigentliche Integral $\int_a^\beta |f|$ konvergent ist.*

Satz 28.8 (Dreiecksungleichung für uneigentliche Integrale). *Ist das uneigentliche Integral $\int_a^\beta f$ absolut konvergent, so ist es auch konvergent und es gilt*

$$\left| \int_a^\beta f \right| \leq \int_a^\beta |f|.$$

Satz 28.9 (Majoranten-/Minorantenkriterium). *Es sei $\mathbb{K} = \mathbb{R}$.*

(a) Ist $|f(x)| \leq g(x)$ für alle $x \in [a, \beta)$ und ist das uneigentliche Integral $\int_a^\beta g$ konvergent, so konvergiert $\int_a^\beta f$ absolut.

(b) Ist $f(x) \geq g(x) \geq 0$ für alle $x \in [a, \beta)$ und ist das uneigentliche Integral $\int_a^\beta g$ divergent, so ist auch $\int_a^\beta f$ divergent.

Beispiel 28.10. (a) Wir untersuchen

$$\int_1^\infty \frac{x}{\sqrt{1 + x^5}} \, \mathrm{d}x =: \int_1^\infty f(x) \, \mathrm{d}x$$

auf Konvergenz. Wegen

$$|f(x)| = \frac{x}{\sqrt{1 + x^5}} \leq \frac{x}{\sqrt{x^5}} = \frac{1}{x^{3/2}} =: g(x)$$

und da nach Beispiel 28.2 (c) das uneigentliche Integral $\int_1^\infty x^{-3/2} \, \mathrm{d}x$ konvergiert, ist das untersuchte uneigentliche Integral nach dem Majorantenkriterium absolut konvergent.

Einen genauen Wert für das Integral können wir, wie beim Majorantenkriterium üblich, nicht angeben, aber das ist auch meist nicht nötig, denn wir haben ja mit Hilfe der Dreiecksungleichung die Abschätzung

$$\int_1^\infty f(x) \, \mathrm{d}x \leq \int_1^\infty g(x) \, \mathrm{d}x = \int_1^\infty \frac{1}{x^{3/2}} \, \mathrm{d}x = \frac{1}{3/2 - 1} = 2.$$

(b) Wir untersuchen noch das uneigentliche Integral

$$\int_1^\infty \frac{x}{x^2+7x+10}\,\mathrm{d}x =: \int_1^\infty f(x)\,\mathrm{d}x.$$

Vergleichen wollen wir die Funktion f mit der Funktion $g(x) := 1/x$ für große x. Dazu bemerken wir zunächst

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{x^2}{x^2+7x+10} = 1.$$

Also gibt es ein $c > 1$, so dass $f(x)/g(x) \geq 1/2$ für alle $x \geq c$ gilt, was wiederum $f(x) \geq g(x)/2 = 1/(2x) > 0$ für alle diese x bedeutet. Da nun das uneigentliche Integral $\int_c^\infty 1/(2x)\,\mathrm{d}x$ nach Beispiel 28.2 (c) divergent ist, divergiert nach dem Minorantenkriterium auch das uneigentliche Integral $\int_c^\infty f(x)\,\mathrm{d}x$, und damit divergiert auch das Ausgangsintegral.

Bemerkung 28.11. Das im letzten Beispiel verwendete Verfahren ist ziemlich universell einsetzbar. Allgemein folgt für zwei Funktionen f und g aus der Beziehung $\lim_{x\to\beta} f(x)/g(x) = L > 0$ die Ungleichungskette

$$\frac{1}{2}L \leq \frac{f(x)}{g(x)} \leq \frac{3}{2}L \quad \text{für alle } x \in [c,\beta)$$

für ein nahe genug bei β gewähltes c. Daraus lässt sich dann immer wie oben eine Abschätzung für das Majoranten- bzw. das Minorantenkriterium bekommen. Qualitativ gesprochen bedeutet die Existenz eines endlichen Grenzwertes von $f(x)/g(x)$, wenn x gegen die Problemstelle läuft, dass f und g das gleiche Verhalten an der Problemstelle haben.

Bevor wir dieses Kapitel abschließen sei noch einmal vor zwei typischen Fehlern gewarnt.

Warnung 28.12. (a) Das uneigentliche Integral $\int_{-\infty}^\infty f$ ist *nicht* definiert durch $\lim_{t\to\infty} \int_{-t}^t f$, sondern

$$\int_{-\infty}^\infty f = \lim_{t\to-\infty} \int_t^0 f + \lim_{s\to\infty} \int_0^s f.$$

Das ist ein wesentlicher Unterschied, wie man an dem Beispiel $\int_{-\infty}^\infty x\,\mathrm{d}x$ sieht. Dieses ist offensichtlich divergent, denn sowohl $\int_{-\infty}^0 x\,\mathrm{d}x$, als auch $\int_0^\infty x\,\mathrm{d}x$ sind divergent, aber für jedes $t > 0$ gilt

$$\int_{-t}^t x\,\mathrm{d}x = 0.$$

Der oben angegebene Limes existiert hier also und ist Null. Trotzdem macht es keinen Sinn, dadurch das uneigentliche Integral zu definieren, denn dass sich die positiven und negativen Beiträge hier gerade aufheben, liegt daran, dass wir die Grenzwerte in Richtung ∞ und $-\infty$ genau gleich schnell laufen lassen. Bildet man z.B. den genau so sinnvollen (bzw. nicht sinnvollen) Grenzwert

$$\lim_{t\to\infty} \int_{-t}^{t+1} x \, \mathrm{d}x = \lim_{t\to\infty} \frac{1}{2}(t+1)^2 - \frac{1}{2}t^2 = \lim_{t\to\infty} \left(t + \frac{1}{2}\right) = \infty,$$

sieht das Ergebnis schon anders aus.

Also merke: Ein doppelt uneigentliches Integral konvergiert nur dann, wenn es an beiden Integrationsgrenzen unabhängig voneinander konvergiert.

(b) Wir haben in Beispiel 28.2 (d) bemerkt, dass das uneigentliche Integral $\int_0^1 1/\sqrt{x} \, \mathrm{d}x$ konvergiert, aber $\int_0^1 1/x \, \mathrm{d}x$ divergiert. Daran sieht man, dass man im Allgmeinen nicht schließen kann, dass mit f auch sofort f^2 uneigentlich integrierbar ist! Das wird trotzdem immer wieder gerne versucht. Es gilt also allgemein *nicht*, dass das Produkt uneigentlich integrierbarer Funktionen wieder uneigentlich integrierbar ist.

Übungsaufgabe 28.13. Das uneigentliche Integral $\int_a^\beta f$ ist genau dann konvergent, wenn es ein $c \in [a, \beta)$ gibt, so dass $\int_c^\beta f$ konvergent ist. In diesem Falle gilt

$$\int_a^\beta f = \int_a^c f + \int_c^\beta f.$$

29. Die Γ-Funktion

Satz 29.1. *Es sei $x > 0$. Dann ist das doppelt uneigentliche Integral*

$$\int_0^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$$

konvergent.

Beweis. Wir untersuchen zunächst das Integral von Null bis Eins. Dazu beobachten wir, dass

$$\lim_{t \to 0+} \frac{\mathrm{e}^{-t} t^{x-1}}{\frac{1}{t^{1-x}}} = \lim_{t \to 0+} \mathrm{e}^{-t} = 1$$

ist. Also gibt es wie in Bemerkung 28.11 ein $c \in (0, 1)$ mit

$$0 \leq \mathrm{e}^{-t} t^{x-1} \leq \frac{3}{2} \frac{1}{t^{1-x}} \quad \text{für alle } t \in (0, c).$$

Da außerdem für alle $x > 0$ das uneigentliche Integral $\int_0^c 1/t^{1-x} \, \mathrm{d}t$ konvergiert, ist nach dem Majorantenkriterium auch $\int_0^c \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$, und damit nach Übungsaufgabe 28.13 auch $\int_0^1 \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$ konvergent.
Für das Integral von Eins bis ∞ vergleichen wir mit $1/x^2$ und erhalten

$$\lim_{t \to \infty} \frac{\mathrm{e}^{-t} t^{x-1}}{\frac{1}{t^2}} = \lim_{t \to \infty} \frac{t^{x+1}}{\mathrm{e}^t} = 0.$$

Also gibt es wieder ein $c > 1$ mit

$$0 \leq \mathrm{e}^{-t} t^{x-1} \leq \frac{1}{2} \frac{1}{t^2} \quad \text{für alle } t \geq c$$

und da $\int_c^\infty 1/t^2 \, \mathrm{d}t$ konvergent ist, konvergiert damit nach dem Majorantenkriterium auch wieder $\int_c^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$ und somit auch $\int_1^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$. Also sind beide Teile des doppelt uneigentlichen Integrals konvergent, d. h. es konvergiert auch als ganzes. $\square$

Das soeben behandelte Integral ist wichtig genug, dass es einen Namen verdient hat.

Definition 29.2. *Die nach Satz 29.1 durch* $\Gamma : (0, \infty) \to \mathbb{R}$ *mit*

$$\Gamma(x) := \int_0^\infty \mathrm{e}^{-t} t^{x-1}, \quad x > 0,$$

gegebene Funktion heißt Gamma-Funktion*.*

Satz 29.3. *Für alle* $x > 0$ *gilt* $\Gamma(x+1) = x\Gamma(x)$*.*

Beweis. Es seien $0 < \alpha < \beta$. Dann gilt mit partieller Integration

$$\int_\alpha^\beta \mathrm{e}^{-t} t^x \, \mathrm{d}t = -\mathrm{e}^{-t} t^x \Big|_\alpha^\beta - \int_\alpha^\beta (-\mathrm{e}^{-t}) x t^{x-1} \, \mathrm{d}t = \frac{1}{\mathrm{e}^\alpha} \alpha^x - \frac{1}{\mathrm{e}^\beta} \beta^x + x \int_\alpha^\beta \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t.$$

Setzen wir speziell $\beta = 1$, so erhalten wir

$$\int_\alpha^1 \mathrm{e}^{-t} t^x \, \mathrm{d}t = \frac{1}{\mathrm{e}^\alpha} \alpha^x - \frac{1}{\mathrm{e}} + x \int_\alpha^1 \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t.$$

und mit $\alpha \to 0$ also

$$\int_0^1 \mathrm{e}^{-t} t^x \, \mathrm{d}t = -\frac{1}{\mathrm{e}} + x \int_0^1 \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t.$$

Machen wir die gleichen Überlegungen mit der speziellen Wahl $\alpha = 1$ und dem Grenzübergang $\beta \to \infty$, so erhalten wir

$$\int_1^\beta \mathrm{e}^{-t} t^x \, \mathrm{d}t = \frac{1}{\mathrm{e}} + x \int_1^\beta \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t,$$

bzw.

$$\int_1^\infty \mathrm{e}^{-t} t^x \, \mathrm{d}t = \frac{1}{\mathrm{e}} + x \int_1^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t.$$

Zusammengenommen bedeutet das

$$\begin{aligned}
\Gamma(x+1) &= \int_0^\infty \mathrm{e}^{-t} t^x \, \mathrm{d}t = \int_0^1 \mathrm{e}^{-t} t^x \, \mathrm{d}t + \int_1^\infty \mathrm{e}^{-t} t^x \, \mathrm{d}t \\
&= -\frac{1}{\mathrm{e}} + x \int_0^1 \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t + \frac{1}{\mathrm{e}} + x \int_1^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t = x \int_0^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t \\
&= x\Gamma(x).
\end{aligned}$$

□

Aus diesem Resultat lässt sich nun relativ schnell folgern, dass die Gamma-Funktion eine Erweiterung der Fakultät auf die reellen Zahlen darstellt.

Korollar 29.4. *Für alle* $n \in \mathbb{N}_0$ *gilt*

$$\Gamma(n+1) = n!$$

Beweis. Wir führen einen Induktionsbeweis. Für $n = 0$ gilt

$$\Gamma(1) = \int_0^\infty \mathrm{e}^{-t} t^0 \,\mathrm{d}t = \lim_{s\to\infty} \int_0^s \mathrm{e}^{-t} \,\mathrm{d}t = \lim_{s\to\infty} -\mathrm{e}^{-t}\Big|_0^s = \lim_{s\to\infty} 1 - \mathrm{e}^{-s} = 1 = 0!.$$

Also haben wir den Induktionsanfang erledigt. Gilt nun $\Gamma(n+1) = n!$ für ein $n \in \mathbb{N}_0$, so haben wir nach Satz 29.3

$$\Gamma(n+2) = \Gamma((n+1)+1) = (n+1)\Gamma(n+1) = (n+1)n! = (n+1)!. \qquad \square$$

Tabelle der griechischen Buchstaben

groß	klein	Name
A	α	Alpha
B	β	Beta
Γ	γ	Gamma
Δ	δ	Delta
E	ϵ, ε	Epsilon
Z	ζ	Zeta
H	η	Eta
Θ	θ, ϑ	Theta
I	ι	Iota
K	κ, $\varkappa$	Kappa
Λ	λ	Lambda
M	μ	My
N	ν	Ny
Ξ	ξ	Xi
O	o	Omikron
Π	π, ϖ	Pi
P	ρ, ϱ	Rho
Σ	σ, ς	Sigma
T	τ	Tau
Y	υ	Ypsilon
Φ	ϕ, φ	Phi
X	χ	Chi
Ψ	ψ	Psi
Ω	ω	Omega

Index