Skript zur Vorlesung

Analysis I

Wintersemester 2020/21

Robert Haller-Dintelmann, Karoline Disser

19. Mai 2021

Inhaltsverzeichnis

I. Grundlegende Begriffe 1

1. Mengen 3

2. Logische Symbole 7
2.1. Negation . . . 7
2.2. Quantoren . . . 7
2.3. Implikation und Äquivalenz . . . 8
2.4. Und und Oder . . . 9

3. Abbildungen 11

II. Zahlen 15

4. Reelle Zahlen 17
4.1. Körperaxiome . . . 17
4.2. Anordnung . . . 19
4.3. Vollständigkeit . . . 21
4.4. Betragsfunktion . . . 23

5. Natürliche Zahlen 25
5.1. Induktionsmengen und Definition der natürlichen Zahlen . . . 25
5.2. Induktionsbeweise und Folgerungen . . . 28

6. Rationale Zahlen 33
6.1. Wie verhält sich $\mathbb{Q}$ zu $\mathbb{R}$? . . . 33
6.2. Wurzeln und rationale Exponenten . . . 35

III. Konvergenz 39

7. Folgen 41
7.1. Konvergenz von reellen Folgen . . . 41
7.2. Grenzwertsätze . . . 44
7.3. Bestimmte Divergenz . . . 50

7.4. Monotonie von Folgen . . . 51
7.5. Teilfolgen und Häufungswerte . . . 55
7.6. Beschränkte Folgen . . . 58
7.7. Cauchy-Folgen . . . 63

8. Reihen **67**
8.1. Konvergenzkriterien für Reihen . . . 70
8.2. Absolute Konvergenz und der Riemannsche Umordnungssatz . . . 72
8.3. Kriterien für absolute Konvergenz . . . 79
8.4. Das Cauchyprodukt . . . 84
8.5. Potenzreihen . . . 89

9. *Komplexe Zahlen **95**

IV. Reelle Funktionen **105**

10. Stetigkeit **107**
10.1. Definition von Stetigkeit . . . 107
10.2. Grenzwerte bei Funktionen . . . 111
10.3. Verhalten von Funktionen im Unendlichen . . . 117
10.4. Eigenschaften stetiger Funktionen . . . 118
10.5. Funktionenfolgen . . . 127
10.6. Gleichmäßige Stetigkeit . . . 133

11. Differenzierbarkeit **137**
11.1. Rechenregeln für Ableitungen . . . 139
11.2. Eigenschaften differenzierbarer Funktionen . . . 143
11.3. Die Ableitung von Potenzreihen . . . 149
11.4. Trigonometrische Funktionen . . . 152
11.5. Höhere Ableitungen . . . 164
11.6. Der Satz von Taylor . . . 168

12. Integration **175**
12.1. Das Riemann-Integral . . . 175
12.2. Welche Funktionen sind Riemann-integrierbar? . . . 184
12.3. Eigenschaften integrierbarer Funktionen . . . 192
12.4. Wie bestimme ich ein Integral? . . . 199
12.5. Uneigentliche Integrale . . . 203
12.6. Vermischtes zum Riemann-Integral . . . 211

Tabelle der griechischen Buchstaben **219**

Index **219**

Teil I.

Grundlegende Begriffe

1. Mengen

Bevor wir mit der Analysis anfangen, brauchen wir ein paar Grundbegriffe, um Mathematik sauber aufzuschreiben. Die in diesem Abschnitt eingeführten Konzepte haben also nicht primär etwas mit Analysis zu tun, sondern bilden das Grundvokabular, in dem eigentlich jeder mathematische Text geschrieben ist.
Wir beginnen mit folgender Übereinkunft zur Verwendung des Gleichheitszeichens:

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

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. Dazu dient die folgende, von Georg Cantor gegebene, Definition: „Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die *Elemente* von M genannt werden) zu einem Ganzen.“
Wenn wir Mengen bilden, ist unser Ausgangspunkt immer eine gegebene, unter Umständen sehr große Grundmenge G, aus der Elemente ausgesondert und zu neuen Mengen zusammengefasst werden.
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 jedoch 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:

$$M_1 = \{x \in \mathbb{Z} : 1 \leq x < 6\} \quad \text{oder} \quad M_1 = \{x \in \mathbb{Z} : x(x-6) < 0\}.$$

Allgemein schreibt man

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

Dabei ist G die Grundmenge, aus der die Elemente der Menge M ausgesondert werden sollen und $E(x)$ ist eine *Aussageform*, die durch 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

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

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 zwei Mengen M und N gleich sind, wenn sowohl $M \subseteq N$ als auch $M \supseteq N$ gilt.

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

(a) $M \cup N := \{x \in G : x \in M \text{ oder } x \in N\}$ *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.*

Mit diesen Begriffen können wir nun schon etwas Mathematik betreiben. Wir sammeln die wichtigsten Regeln für obige Mengenoperationen in folgendem Satz.

Satz 1.4. *Es seien A, B und C Mengen. 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 auch $x \in A \cup B$ und $x \in A \cup C$, denn diese Mengen enthalten 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

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

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 möglichen 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 Regel von De Morgan 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 enthalten. 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

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

zeigt man 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 genau $x \in (A \cup B)^c$ bedeutet. □

Definition 1.5. *Ist I eine Menge (Man nennt I in diesem Zusammenhang* Indexmenge*) und ist für jedes $i \in I$ eine Menge M_i gegeben, so ist*

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

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 $\bigcap_{n\in\mathbb{N}} M_n$.

2. Logische Symbole

Für die mathematischen Aussagen hier in der Vorlesung gilt, dass sie entweder wahr oder falsch sind. Ein Zwischending oder eine andere Möglichkeit gibt es nicht (außer das wird explizit dazugesagt). In diesem Kapitel geht es einerseits um vereinfachende Schreibweisen für Ausdrücke, die in Aussagen oft vorkommen (zum Beispiel Quantoren) und andererseits um logische Operationen, mit denen wir gegebene Aussagen zu neuen Aussagen verknüpfen können (Beispiel: Implikation).

2.1. Negation

Jede mathematische Aussage A kann auch verneint werden. Man schreibt dann $\neg A$ und spricht "nicht A". Die Aussage $\neg A$ heißt die *Negation* der Aussage A. Dann gilt natürlich: ist A wahr, dann ist $\neg A$ falsch. Ist A falsch, dann ist $\neg A$ wahr.
Beispiel: Die Aussage „$5 = 2 + 3$“ ist wahr und ihre Negation „$\neg(5 = 2 + 3)$“ ist falsch. Man kann in diesem Fall für die Verneinung auch schreiben: „$5 \neq 2 + 3$“.

2.2. Quantoren

Oft werden in mathematischen Texten die folgenden Symbole verwendet:

- Der *Allquantor* $\forall$ bedeutet „für alle“.

  Beispiel: „$\forall n \in \mathbb{N} : 2n$ ist eine gerade Zahl“ ist eine wahre Aussage.

- Der *Existenzquantor* $\exists$ steht für „es existiert“.

  Beispiel: „$\exists x \in \mathbb{R} : x^2 = -1$“ ist eine falsche Aussage.

Beispiel 2.1. Für ein 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 bekannte Aussage

$$\exists s \in S\ \forall w \in W : w \text{ führt nach } s. \tag{2.1}$$

Übersetzung: Es gibt eine Stadt (meist Rom genannt), zu der alle Wege hinführen. Gesucht ist jetzt die Verneinung dieser Aussage, die also genau dann wahr ist, wenn (2.1) falsch ist und genau dann falsch ist, wenn (2.1) wahr ist.

Es gilt die folgende *Regel zum Verneinen von Aussagen mit Quantoren*:

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. Für den Spezialfall Rom ergibt sich: Es gibt einen Weg, der nicht nach Rom führt.

2.3. 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 solcher zusammengesetzter Aussagen kann man durch eine *Wahrheitstafel* angeben. Für den Wahrheitsgehalt der Aussagen A und B gibt es vier verschiedene Fälle (beide wahr, A wahr und B falsch, A falsch und B wahr sowie beide falsch). Eine Wahrheitstafel gibt nun in jedem dieser Fälle den Wahrheitsgehalt der zusammengesetzten Aussage an. Für Implikation und Äquivalenz ergibt das:

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

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

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

Warnung 2.3. Hüten Sie sich vor dem Umkehrschluss: Wenn die Aussage „$A \implies B$“ wahr ist, kann man daraus nicht folgern, dass auch „$B \implies A$“ stimmt. Dieser Fehlschluss wird auch oft in folgender Version gemacht: Aus „$A \implies B$“ wird gefolgert, dass, wenn A falsch ist, auch B falsch sein muss. Ein Blick in obige Wahrheitstafel zeigt sofort, dass das nicht stimmt. Trotzdem passiert es immer wieder.

Bemerkung 2.4. Eine zur Implikation „$A \implies B$“ äquivalente Aussage ist dagegen die *Kontraposition* „nicht $B \implies$ nicht As“, wie man der folgenden Wahrheitstafel entnehmen kann:

A	B	$A \implies B$	nicht B	nicht A	nicht $B \implies$ nicht A
w	w	w	f	f	w
w	f	f	w	f	f
f	w	w	f	w	w
f	f	w	w	w	w

Im sogenannten *Beweis durch Kontraposition* wird dieser Zusammenhang genutzt.

2.4. Und und Oder

Weitere Möglichkeiten, aus bestehenden mathematischen Aussagen A, B neue Aussagen zu machen, sind die

- *Konjunktion* $A \wedge B$ (A „und “ B), und die
- *Disjunktion* $A \vee B$ (A „oder “ B).

Das funktioniert auf ganz natürliche Weise: $A \wedge B$ ist genau dann wahr, wenn die Aussage A und die Aussage B wahr sind. $A \vee B$ ist genau dann wahr, wenn mindestens eine der beiden Aussagen A oder B wahr ist. Zu beachten ist hier also nur, dass die Disjunktion nicht exklusiv ist: $A \vee B$ ist auch wahr, wenn sowohl A als auch B wahr sind. Das entspricht nicht unbedingt dem alltäglichen Sprachgebrauch.

Übungsaufgabe 2.5. Stellen Sie die Wahrheitstafeln für Konjunktion, Disjunktion, und die Ausdrücke $(\neg A) \vee B$, $\neg(A \wedge B)$ und $\neg(A \vee B)$ auf. Wozu ist $(\neg A) \vee B$ äquivalent?

3. Abbildungen

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

Definition 3.1. *Es seien D und Z Mengen und es sei jedem Element d aus D genau ein Element $f(d)$ in Z zugeordnet. Diese Zuordnung nennt man dann eine* Abbildung *oder auch* Funktion *f und schreibt*

$$f : D \to Z, \quad d \mapsto f(d) \qquad \text{oder} \qquad f : \begin{cases} D \to Z \\ d \mapsto f(d). \end{cases}$$

Dabei heißt D die Definitionsmenge *und Z die* Zielmenge *von f.*
Weiter nennt man ein Element $d \in D$ auch Argument *von f, $f(d)$ das* Bild *von d unter f und ist $z \in Z$, so heißt jedes $d \in D$ mit $f(d) = z$ ein* Urbild *von z.*
Schließlich ist für jedes $A \subseteq D$ das Bild*, auch* Bildmenge *genannt, von A unter f gegeben durch*

$$f(A) := \{f(a) : a \in A\}$$

und für jede Teilmenge B von Z ist

$$f^{-1}(B) := \{d \in D : f(d) \in B\}$$

das Urbild *bzw. die* Urbildmenge *der Menge B.*

Beispiel 3.2. 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 den reellen Zahlen:[1]

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

oder die Wurzelfunktion

$$\sqrt{} : \{x \in \mathbb{R} : x \geq 0\} \to \mathbb{R}, \quad x \mapsto \sqrt{x}.$$

Im Folgenden definieren wir die Verkettung, d. h. die Nacheinanderausführung von Abbildungen.

[1]Rein formal betrachtet können wir diese Beispiele hier noch nicht betrachten, da wir die reellen Zahlen noch nicht eingeführt haben. Es sei also vorübergehend an Ihr Schulwissen appeliert.

Definition 3.3. *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 *oder auch* Komposition *von g mit f.*

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

Definition 3.4. *Es seien D und Z Mengen sowie $f : D \to Z$ eine Funktion. Dann heißt f*

(a) surjektiv*, wenn $f(D) = Z$ gilt.*

(b) injektiv*, wenn für alle $x, y \in D$ mit $f(x) = f(y)$ stets $x = y$ gilt.*

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

Bijektive Abbildungen sind deshalb besonders wichtig, weil sie sich rückgängig machen lassen. Das konkretisieren wir im folgenden Satz.

Satz 3.5. *Es seien D, Z Mengen und $f : D \to Z$ eine Funktion. Dann ist f genau dann bijektiv, wenn es für jedes $b \in Z$ genau ein $a \in D$ gibt, sodass $f(a) = b$ gilt. In diesem Fall existiert eine Abbildung $f^{-1} : Z \to D$ mit*

$$f^{-1}(f(a)) = a \quad \text{für alle } a \in D \quad \text{und} \quad f(f^{-1}(b)) = b \quad \text{für alle } b \in Z.$$

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

Da f surjektiv ist, gibt es zu jedem $b \in Z$ mindestens ein $a \in D$ mit $f(a) = b$. Sind $a_1, a_2 \in D$ mit $f(a_1) = b = f(a_2)$ gegeben, so folgt aus der Injektivität von f sofort $a_1 = a_2$, es kann also nur genau ein solches $a \in D$ geben.

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

Nach Voraussetzung sind alle $b \in Z$ in $f(D)$ enthalten, also ist f surjektiv. Seien nun $a_1, a_2 \in D$ mit $f(a_1) = f(a_2)$ gegeben. Da jedes $b \in Z$ 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 $f^{-1} : Z \to D$ mit $f^{-1}(f(a)) = a$ für alle $a \in D$ und $f(f^{-1}(b)) = b$ für alle $b \in Z$.

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

Definition 3.6. *Es seien D, Z zwei Mengen und $f : D \to Z$ bijektiv. Dann heißt die Abbildung $f^{-1} : Z \to D$ aus Satz 3.5* Umkehrfunktion *von f.*

Beispiel 3.7. Betrachten wir noch einmal die beiden Abbildungen aus Beispiel 3.2, so ist die erste Funktion

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

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

$$\widetilde{f} : \mathbb{R} \to \{x \in \mathbb{R} : x \geq 0\}, \quad x \mapsto x^2,$$

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

$$\widehat{f} : \{x \in \mathbb{R} : x \geq 0\} \to \{x \in \mathbb{R} : x \geq 0\}, \quad x \mapsto x^2.$$

Nun ist $\widehat{f}$ tatsächlich bijektiv 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 3.8. *Seien D, Z Mengen, $f : D \to Z$ eine Funktion und $M \subseteq D$. Dann ist $f|_M$ die* Einschränkung von f auf M*, d. h. $f|_M : M \to Z$ ist gegeben durch $f|_M(x) = f(x)$ für $x \in M$.*

Teil II.

Zahlen

4. Reelle 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 ohne Beweis zu Grunde legen. Aus diesen Axiomen für die reellen Zahlen können wir dann den gesamten weiteren Stoff der Vorlesung herleiten (eventuell von wenigen Ausnahmen abgesehen). In späteren Vorlesungen oder Seminaren werden Sie vielleicht die *Konstruktion* der reellen Zahlen sehen: die Herleitung der Existenz einer Menge, die die Eigenschaften (A1) – (A15) erfüllt, basierend auf noch grundlegenderen Axiomen.
Die Axiome für die reellen Zahlen teilen sich in drei Typen, die in den folgenden Abschnitten eingeführt werden:

1) die *Körperaxiome*, oder wie man mit „+“ und „$\cdot$“ rechnet, (A1) – (A9),
2) die *Anordnungsaxiome*, oder wie man mit „$\leq$“ umgeht, (A10) – (A14), und
3) das *Vollständigkeitsaxiom*, oder warum $\mathbb{R}$ erstmal alle Zahlen enthält, die wir benötigen (A15).

4.1. 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 $0 \in \mathbb{R}$, sodass $a+0 = a$ für alle $a \in \mathbb{R}$ ist (*Nullelement*).

(A3) Für jedes $a \in \mathbb{R}$ gibt es ein $-a \in \mathbb{R}$, sodass $a + (-a) = 0$ 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 $1 \neq 0$, sodass $a \cdot 1 = a$ für alle $a \in \mathbb{R}$ ist (*Einselement*).

(A7) Für jedes $a \in \mathbb{R} \setminus \{0\}$ gibt es ein $a^{-1} \in \mathbb{R}$, sodass $a \cdot a^{-1} = 1$ 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 4.1. Allgemeiner heißt jede Menge, auf der zwei Verknüpfungen definiert sind, die die Eigenschaften (A1) – (A9) haben, ein *Körper*. Insbesondere ist das Tripel $(\mathbb{R}, +, \cdot)$ also ein Körper. Eine Menge, auf der eine Verknüpfung definiert ist, die (A1) – (A4) erfüllt, heißt Abel'sche Gruppe. Das Paar $(\mathbb{R}, +)$ bildet also eine Abel'sche Gruppe, und das Paar $(\mathbb{R} \setminus \{0\}, \cdot)$ ebenfalls (Übung: warum?).

Bemerkung 4.2. Alle bekannten Rechenregeln für „+“ und „·“ lassen sich aus (A1) – (A9) ableiten.

Wir betrachten die folgenden Aussagen als Beispiele:

Satz 4.3. *(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 $a \in \mathbb{R}$ unter Zuhilfenahme von (A9)

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

Daraus folgt nun

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

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

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

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

Definition 4.4. *(a) Den „$\cdot$“ 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)*.*

4.2. Anordnung

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

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

(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 $ac \leq bc$.

Bemerkung 4.5. Allgemeiner heißt eine Menge, auf der es eine Relation gibt, die (A10) – (A12) erfüllt, *total geordnet* und die Relation heißt dann *Ordnungsrelation*. Die Axiome (A13) und (A14) sagen hier etwas über die Verträglichkeit von Addition und Multiplikation mit $\leq$.

Definition 4.6. *Seien $a, b \in \mathbb{R}$. Wir setzen*

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

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

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

Definition 4.7. *Eine reelle Zahl a heißt*

(a) positiv*, falls $a \geq 0$ und* strikt positiv*, wenn $a > 0$ gilt.*

(b) negativ*, falls $a \leq 0$ und* strikt negativ*, wenn $a < 0$ gilt.*

Bemerkung 4.8. Alle Regeln für Ungleichungen lassen sich aus den Axiomen (A1) – (A14) ableiten.

Wir betrachten wieder einige Beispiele.

Satz 4.9. *Seien $a, b, c \in \mathbb{R}$. Dann gilt*

(a) Ist $a \leq b$, so gilt $-a \geq -b$.

(b) Sind $a \leq b$ und $c \leq 0$, so folgt $ac \geq bc$.

(c) $1 > 0$.

Beweis. (a) Nach (A13) mit $c := -a$ bekommen wir aus $a \leq b$ die Ungleichung $0 = a+(-a) \leq b+(-a)$. Wenden wir auf diese Ungleichung dasselbe Axiom mit $c := -b$ an, erhalten wir

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

und sind fertig.

(b) Nach Teil (a) gilt $-c \geq -0 = 0$, also liefert (A14)

$$-ac = a \cdot (-c) \leq b \cdot (-c) = -bc.$$

Eine erneute Anwendung von Teil (a) ergibt dann

$$ac = -(-ac) \geq -(-bc) = bc.$$

(c) Die Annahme $1 \leq 0$ führt nach Teil (b) mit $a := c := 1$ und $b := 0$ sofort auf den Widerspruch $1 = ac \geq bc = 0$. □

Wir verwenden die Anordnung von $\mathbb{R}$, um *Intervalle* von reellen Zahlen zu definieren.

Definition 4.10. *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:

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

4.3. Vollständigkeit

Als Vorbereitung für das 15. Axiom führen wir einige Schreibweisen und Begriffe ein, die bei der Untersuchung von Teilmengen von $\mathbb{R}$ nützlich sind.

Definition 4.11. *Es sei M eine nicht-leere Teilmenge von $\mathbb{R}$.*

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

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

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

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

(c) M heißt beschränkt*, wenn M nach oben und nach unten beschränkt ist.*

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

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

Wir überlegen uns kurz, dass das Supremum und das Infimum einer Menge, wenn es denn existiert, eindeutig bestimmt ist.

Satz 4.12. *Eine Teilmenge von $\mathbb{R}$ hat höchstens ein Supremum und höchstens ein Infimum.*

Beweis. Wir beweisen die Aussage nur für das Supremum, das Argument für das Infimum geht analog.
Es sei M ein Teilmenge von $\mathbb{R}$ mit 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 $C_2 \leq C_1$. Damit ist $C_2 = C_1$. $\square$

Das ermöglicht die nachstehende Definition.

Definition 4.13. *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 4.14. (a) Wir betrachten $M = (0, 1]$. Dann ist M nach oben und nach unten beschränkt, also 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) Nun sei $M = (-\infty, -1)$. 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.

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

Wir können damit die entsprechende Aussage über das Infimum beweisen.

Satz 4.15. *Ist* $M \subseteq \mathbb{R}$*,* $M \neq \emptyset$ *und* M *nach unten beschränkt, so existiert* $\inf M$*.*

Beweis. Wir setzen $\widetilde{M} := \{-x : x \in M\}$. 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 $\widetilde{M}$. Nach Axiom (A15) existiert also $s := \sup \widetilde{M}$. 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 $\widetilde{M}$. Da s das Supremum von $\widetilde{M}$ ist, muss also $s \leq -\sigma$, und damit $\sigma \leq -s$ gelten. Also ist $-s = \inf M$. □

Wir sammeln einige elementare Eigenschaften von Supremum und Infimum.

Satz 4.16. *(a) Ist* $M \subseteq \mathbb{R}$ *nicht-leer und beschränkt, so gilt* $\inf M \leq \sup M$*.*

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

Beweis. Zum Beweis von (a) sei $x \in M$ beliebig gewählt. Man beachte dazu, dass $M \neq \emptyset$ vorausgesetzt ist. Nun gilt $x \geq \inf M$ und $x \leq \sup M$. Also ist $\inf M \leq x \leq \sup M$.
Wir wenden uns dem Beweis von (b) zu. Für jedes $x \in N$ gilt $x \in M$ und damit $x \leq \sup M$ (bzw. $x \geq \inf M$). Also ist N nach oben (bzw. unten) beschränkt und $\sup M$ (bzw. $\inf M$) ist eine obere (bzw. untere) Schranke von N. Das impliziert $\sup N \leq \sup M$ (bzw. $\inf N \geq \inf M$). □

Zum Abschluss dieses Kapitels beweisen wir noch einen Satz, der es oft ermöglicht, die Qualität, dass eine Zahl ein Supremum (oder Infimum) ist, in einem Beweis gewinnbringend umzuformulieren.

Satz 4.17. *Ist $M \subseteq \mathbb{R}$ nicht-leer und C eine obere (bzw. untere) Schranke von M, so ist $C = \sup M$ (bzw. $C = \inf M$) genau dann, wenn für jedes $\varepsilon > 0$ ein $x \in M$ existiert, für das $x > C - \varepsilon$ (bzw. $x < C + \varepsilon$) gilt.*

Beweis. Wir beweisen diese Aussage nur für den „sup"-Fall.

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

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

4.4. Betragsfunktion

Im letzten Abschnitt dieses Kapitels über die reellen Zahlen definieren wir nun die Betragsfunktion, ein fundamentales Hilfsmittel in der gesamten Analysis.

Definition 4.18. *Sei $a \in \mathbb{R}$. Dann ist der* Betrag *von a, symbolisiert durch $|a|$, gegeben durch*

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

Wir zeigen jetzt einige Eigenschaften der Betragsfunktion, die wir aus den Axiomen (A1) – (A15) ableiten können.

Satz 4.19. *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) $|a-b| \geq \big||a|-|b|\big|$ (umgekehrte Dreiecksungleichung).

Beweis. Die Teile (a) bis (d) verbleiben als 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 mit (c) wieder $|a+b| \leq |a|+|b|$ 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. $\square$

Übungsaufgabe 4.20. Beweisen Sie die folgende Aussage: Eine Teilmenge M von $\mathbb{R}$ mit $M \neq \emptyset$ ist genau dann beschränkt, wenn es ein $C > 0$ gibt, sodass $|x| \leq C$ für alle $x \in M$ gilt.

5. Natürliche Zahlen

In diesem Kapitel wird die Menge der natürlichen Zahlen $\mathbb{N}$ definiert und mit deren Hilfe auch die Menge der ganzen Zahlen $\mathbb{Z}$. Wir nutzen dabei aus, dass wir die reellen Zahlen $\mathbb{R}$ schon haben. Die Definition der natürlichen Zahlen ist so gewählt, dass wir das wichtige Beweisverfahren der vollständigen Induktion leicht daraus ableiten können. Als Hilfskonstrukt definieren wir zunächst den Begriff der Induktionsmenge.

5.1. Induktionsmengen und Definition der natürlichen Zahlen

Definition 5.1. *Eine Menge $A \subseteq \mathbb{R}$ heißt* Induktionsmenge*, falls gilt*

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

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

Beispiel 5.2. Beispiele von Induktionsmengen sind $\mathbb{R}$ oder $\{1\} \cup [2, \infty)$.

Satz 5.3. *Ist I eine Indexmenge und ist für jedes $i \in I$ eine Induktionsmenge $A_i \subseteq \mathbb{R}$ gegeben, so ist auch deren Durchschnitt $A := \bigcap_{i \in I} A_i$ eine Induktionsmenge.*

Beweis. Da die Eins in jeder Induktionsmenge liegt, muss sie in jedem A_i liegen und damit auch in deren Durchschnitt A. Damit ist Bedingung (a) aus der Definition einer Induktionsmenge gezeigt.
Für den Nachweis der zweiten Bedingung sei $x \in A$. Nach Definition von A gilt dann $x \in A_i$ für jedes $i \in I$. Da die Mengen A_i alle Induktionsmengen sind, muss dann aber für jedes $i \in I$ auch $x + 1 \in A_i$ gelten. Damit erhalten wir schließlich $x + 1 \in A$ und sind fertig. □

Definition 5.4. *Den Durchschnitt aller Induktionsmengen bezeichnen wir mit $\mathbb{N}$. Das ist die Menge der* natürlichen Zahlen.
Weiter setzen wir

$$\begin{aligned} \mathbb{N}_0 &:= \mathbb{N} \cup \{0\} \quad \textit{und} \\ \mathbb{Z} &:= \mathbb{N}_0 \cup \{-n : n \in \mathbb{N}\} \quad \text{(ganze Zahlen).} \end{aligned}$$

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

Satz 5.5. *(a)* $\mathbb{N}$ *ist eine Induktionsmenge.*

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

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

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

(e) Für alle $x \in \mathbb{R}$ mit $x > 0$ gibt es ein $n \in \mathbb{N}$, sodass $1/n < x$ ist.

Beweis. (a) Das folgt sofort aus Satz 5.3.

(b) 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 $A = \mathbb{N}$.

(c) Wir nehmen an, $\mathbb{N}$ wäre nach oben beschränkt. Nach dem Vollständigkeitsaxiom existiert also das Supremum $s := \sup \mathbb{N}$ in $\mathbb{R}$. Dann liefert Satz 4.17 ein $n \in \mathbb{N}$ mit $n > s - 1$. (Wähle dort $\varepsilon = 1$.) Wir haben damit $n + 1 > s$ und da $\mathbb{N}$ nach (a) eine Induktionsmenge ist, gilt $n + 1 \in \mathbb{N}$. Das heißt wiederum, dass $n + 1 \leq \sup \mathbb{N} = s$ gelten muss, sodass wir bei einem Widerspruch enden.

(d) 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 (c) steht.

(e) Sei $x > 0$. Dann gibt es nach (d) ein $n \in \mathbb{N}$ mit $1/x < n$. Für dieses n gilt $1/n < x$. □

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

Satz 5.6. *(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 $1 \in A$ 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 wegen Satz 5.5 (b) sofort $A = \mathbb{N}$.

(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 5.2 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 $n \in \mathbb{N}$, sodass A_n eine Induktionsmenge ist.

Induktionsschritt (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 $x \in A_{n+1}$. Es können dann zwei Fälle auftreten. Im ersten Fall ist $x \geq n+2$. Dann gilt $x+1 \geq n+3 \geq n+2$, also haben wir sofort $x + 1 \in A_{n+1}$.

Im zweiten Fall ist $x \in \mathbb{N}$ mit $1 \leq x \leq n+1$. Dann 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 $x + 1 \in A_{n+1}$.

Wir haben also gezeigt, dass A auch eine Induktionsmenge ist, also $\mathbb{N} \subseteq A$. Per Definition ist auch $A \subseteq \mathbb{N}$, also zusammen $A = \mathbb{N}$. Die Behauptung über A_n gilt also für alle $n \in \mathbb{N}$.

(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.

□

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

Satz 5.7 (Wohlordnungsprinzip). *Ist $M \neq \emptyset$ eine Teilmenge von $\mathbb{N}$, so existiert $\min M$.*

Beweis. Nach Satz 5.6 (a) ist eins eine untere Schranke von $\mathbb{N}$ und damit auch von M. Also ist

$$1 \in A := \{\nu \in \mathbb{N} : \nu \text{ ist untere Schranke von } M\}.$$

und diese Menge damit nicht leer. Andererseits ist A durch jedes Element von M (beachte $M \neq \emptyset$) nach oben beschränkt, also ist A nach Satz 5.5 (c) eine echte Teilmenge von $\mathbb{N}$ und damit insbesondere keine Induktionsmenge. Da $1 \in A$ gilt, bedeutet dies, dass es ein $\nu_0 \in A$ geben muss, für das $\nu_0 + 1 \notin A$ gilt. Nach der Definition von A sind ν_0 und damit auch $\nu_0 + 1$ in $\mathbb{N}$, ν_0 ist eine untere Schranke von M, aber $\nu_0 + 1$ ist keine. Es gibt also ein $n_0 \in M$ mit $\nu_0 \leq n_0 < \nu_0 + 1$. Da aber alle drei beteiligten Zahlen in dieser Ungleichungskette in $\mathbb{N}$ liegen, gilt nach Satz 5.6 (c) sogar $n_0 = \nu_0$. Damit ist $n_0 = \nu_0$ eine untere Schranke von M, die in M liegt, also das Minimum von M. □

5.2. Induktionsbeweise und Folgerungen

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 5.8. *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.
Induktionsschritt: 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 $n + 1$.
Wir haben also gezeigt, dass die Menge aller n, für die die Aussage richtig ist, eine Induktionsmenge ist. Daraus folgt, dass die Aussage insbesondere für alle $n \in \mathbb{N}$ richtig ist, weil $\mathbb{N}$ ja in jeder Induktionsmenge enthalten ist. □

Bemerkung 5.9. (a) 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 $n \leq N$ 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 5.8 lässt sich damit z. B. so hinschreiben:

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

(b) Analog verwendet man das *Produktzeichen*

$$\prod_{k=n}^{N} a_k := a_n \cdot a_{n+1} \cdot \ldots \cdot a_{N-1} \cdot a_N$$

mit $n, N \in \mathbb{Z}$ und $a_n, a_{n+1}, \ldots, a_N \in \mathbb{R}$ wie oben.

Zum Abschluss dieses Kapitels führen wir noch ein paar wichtige Schreibweisen und Rechenoperationen, wie z. B. das Potenzieren mit ganzzahligen Exponenten, ein und beweisen einige wichtige Ungleichungen und Identitäten.

Definition 5.10. *(a) Für $a \in \mathbb{R}$ und $n \in \mathbb{N}$ schreiben wir für die* Potenzen

$$a^n := \underbrace{a \cdot a \cdot a \cdot \ldots \cdot a}_{n \text{ Faktoren}} = \prod_{k=1}^{n} a.$$

Weiter setzen wir $a^0 := 1$.

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*

$$n! := 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n = \prod_{k=1}^{n} k$$

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 5.11. *(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

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

(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) Wir führen einen Induktionsbeweis.

Induktionsanfang: Für $n = 1$ lautet die behauptete Ungleichung $1 + x \geq 1 + x$ und ist offensichtlich wahr.

Induktionsvoraussetzung: Für ein $n \in \mathbb{N}$ gelte $(1 + x)^n \geq 1 + nx$ für alle $x \geq -1$.

Induktionsschritt: Für $x \geq -1$ gilt $1 + x \geq 0$, also können wir mit Hilfe der Induktionsvoraussetzung folgern

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

Multiplizieren wir nun aus und lassen den (positiven!) Term mit x^2 weg, so erhalten wir

$$(1 + x)^{n+1} \geq 1 + x + nx + nx^2 = 1 + (n + 1)x + nx^2 \geq 1 + (n + 1)x$$

und sind fertig.

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

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

(c) Es gilt

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

Schreibt man beide Summen aus:

$$\begin{aligned}
&= a^{n+1} + a^n b + a^{n-1} b^2 + \cdots + a^2 b^{n-1} + ab^n \\
&\quad - \left(a^n b + a^{n-1} b^2 + \cdots + a^2 b^{n-1} + ab^n + b^{n+1}\right),
\end{aligned}$$

so erkennt man, dass sich die meisten Summanden wegheben. Das kann man auch mit dem Summenzeichen herausarbeiten. Dazu schreiben wir

$$\begin{aligned}(a-b)\sum_{k=0}^{n} a^{n-k}b^k &= \sum_{k=0}^{n} a^{n-k+1}b^k - \sum_{k=0}^{n} a^{n-k}b^{k+1} \\ &= a^{n+1} + \sum_{k=1}^{n} a^{n-k+1}b^k - \sum_{k=0}^{n-1} a^{n-k}b^{k+1} - b^{n+1}. \qquad (5.1)\end{aligned}$$

Wir machen nun einen sogenannten *Indexshift*. In der ersten Summe ersetzen wir $\ell := k-1$. Wenn k von 1 bis n gezählt wird, läuft also ℓ von 0 bis $n-1$ und wir bekommen

$$\begin{aligned}&= a^{n+1} + \sum_{\ell=0}^{n-1} a^{n-\ell}b^{\ell+1} - \sum_{k=0}^{n-1} a^{n-k}b^{k+1} - b^{n+1} \qquad (5.2) \\ &= a^{n+1} - b^{n+1}.\end{aligned}$$

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

(d) Wir verwenden wieder vollständige Induktion.

Induktionsanfang ($n = 1$):

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

Induktionsvoraussetzung: Für ein $n \in \mathbb{N}$ gelte

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

Induktionsschritt: Es gilt

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

Erneut machen wir einen Indexshift. Dieses Mal ersetzen wir in der zweiten Summe $\ell := k+1$. Das liefert

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

Nun können wir die Zählvariable in der zweiten Summe auch wieder k nennen und erhalten

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

Damit können wir die beiden Summen zu einer zusammenfassen und mit Hilfe der Formeln aus (b) folgt

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

Das ist genau die behauptete Formel für $n+1$ statt n. □

Satz 5.12 (Geometrische Summenformel)**.** *Es sei $q \in \mathbb{R} \setminus \{1\}$. Dann gilt für alle $n \in \mathbb{N}$*

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

Beweis. Es gilt für alle $n \in \mathbb{N}$ und $q \neq 1$

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

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

6. Rationale Zahlen

Nun, da wir die natürlichen Zahlen im Griff haben, können wir ohne weitere Umschweife die rationalen Zahlen definieren. Wir diskutieren, wie sich die rationalen Zahlen zu den reellen Zahlen verhalten. Ein weiteres Thema dieses Abschnitts wird dann sein, das Potenzieren auf rationale Exponenten auszuweiten, d. h. Wurzeln einzuführen.

Definition 6.1. *Wir definieren die* rationalen Zahlen $\mathbb{Q}$ *als die Teilmenge der reellen Zahlen mit*

$$\mathbb{Q} := \left\{ \frac{m}{n} \in \mathbb{R} : m \in \mathbb{Z},\ n \in \mathbb{N} \right\}.$$

Bemerkung 6.2. Jede rationale Zahlen lässt sich auf viele unterschiedliche Weisen als Bruch darstellen. Es ist ja zum Beispiel $\frac{1}{2} = \frac{2}{4} = \frac{3}{6} = \dots$. Dass diese Brüche tatsächlich alle die gleiche Zahl sind, ist in Definition 6.1 dadurch kodiert, dass $\mathbb{Q}$ eine Teilmenge von $\mathbb{R}$ ist, wo wir schon wissen, dass Erweitern und Kürzen den Wert nicht verändert.

6.1. Wie verhält sich $\mathbb{Q}$ zu $\mathbb{R}$?

Per Definition ist $\mathbb{Q}$ eine Teilmenge von $\mathbb{R}$. Darüber hinaus lässt sich aber auch noch Einiges sagen. Wir werden dieses Thema hier nicht erschöpfend behandeln, aber einige wichtige Punkte herausarbeiten bzw. diskutieren:

- Die Menge $\mathbb{Q}$ erfüllt zusammen mit den Verknüpfungen $+$ und $\cdot$ ebenfalls die Axiome (A1) – (A14) aus dem letzten Kapitel. Um das plausibel zu machen, können Sie als Übungsaufgabe prüfen, dass man beim Addieren und Multiplizieren von rationalen Zahlen immer wieder eine rationale Zahl herausbekommt.

- Insbesondere ist $\mathbb{Q}$ eine *echte* Teilmenge von $\mathbb{R}$, oder: es gibt irrationale Zahlen. Das war schon in der Antike bekannt. Wir geben mit Satz 6.3 den Euklidischen Beweis dafür, dass $\sqrt{2}$ irrational ist. Weitere wichtige irrationale Zahlen wie π und e lernen wir im Verlauf der Vorlesung noch näher kennen.

- Wir wollen dann natürlich zeigen, dass diese irrationalen Zahlen, also insbesondere $\sqrt{2}$, in $\mathbb{R}$ enthalten sind. Das gelingt mit Hilfe des Vollständigkeitsaxioms (A15). Wir verallgemeinern diese Überlegung dann gleich und definieren das Potenzieren mit rationalen Exponenten in Abschnitt 6.2.

- Auch wenn $\mathbb{R}$ tatsächlich viel mehr Elemente hat als $\mathbb{Q}$, liegt $\mathbb{Q}$ noch *dicht* in $\mathbb{R}$. Das bedeutet, dass jede reelle Zahl durch rationale Zahlen beliebig gut approximiert werden kann. Diese nützliche Eigenschaft beweisen wir in Satz 6.4. Es ist damit dann auch einfach, zu zeigen, dass $\mathbb{Q}$ das Vollständigkeitsaxiom (A15) nicht erfüllt, s. Bemerkung 6.9.

- Die Dezimaldarstellungen der Zahlen in $\mathbb{Q}$ sind genau die, die entweder abbrechen, oder irgendwann periodisch werden (wobei „Abbrechen" der Dezimaldarstellung ja ein Spezialfall von „periodisch werden" ist). Die Menge der reellen Zahlen $\mathbb{R}$ enthält dagegen auch alle irrationalen Zahlen, deren Dezimaldarstellung nie periodisch wird. Das folgt auch sofort aus dem Vollständigkeitsaxiom. Hier im Skript wird die Notation dazu erst viel später eingeführt, aber im Video zu diesem Abschnitt wird das Thema diskutiert.

Wir starten mit einem ersten Beweis, dass irrationale Zahlen keine „irrationale" Erfindung sind – sie tauchen in ganz natürlicher Weise auf.

Satz 6.3. *Es gibt keine rationale Zahl $q \in \mathbb{Q}$ mit $q^2 = 2$.*

Beweis. (Euklid) Wir nehmen an, es gäbe ein $q \in \mathbb{Q}$ mit $q^2 = 2$. Dann wäre auch $(-q)^2 = 2$ und wir können deshalb auch $q > 0$ annehmen. Aus Definition 6.1 und $q > 0$ folgt, dass es natürliche Zahlen $m, n \in \mathbb{N}$ gibt, sodass $q = \frac{m}{n}$ gilt und m, n teilerfremd sind. Dann folgt auch

$$2 = q^2 = \frac{m^2}{n^2}, \text{ also } 2n^2 = m^2,$$

also ist m^2 eine gerade Zahl. Dann ist aber auch m selbst gerade, da 2 ein Faktor in seiner Primfaktorzerlegung sein muss. Es gibt also ein $k \in \mathbb{N}$ mit $m = 2k$ und daraus folgt

$$2 = q^2 = \frac{4k^2}{n^2}, \text{ also } n^2 = 4k^2,$$

und damit, dass auch n gerade ist. Das ist aber ein Widerspruch dazu, dass m, n als teilerfremd vorausgesetzt waren. Wir haben aus der Annahme $q \in \mathbb{Q}$ mit $q^2 = 2$ also eine falsche Aussage abgeleitet und können folgern, dass es keine rationale Zahl gibt, die die Wurzel aus 2 ist. □

In Abschnitt 6.2 unten zeigen wir, dass $\sqrt{2} \in \mathbb{R}$ existiert. Tatsächlich gibt es viel mehr irrationale Zahlen als rationale Zahlen, aber natürlich von beiden unendlich viele. Genauer gesagt kann man zeigen, dass es keine surjektive Funktion von $\mathbb{Q}$ nach $\mathbb{R} \setminus \mathbb{Q}$ gibt. Trotzdem kommen wir mit Hilfe der rationalen Zahlen beliebig nah an die irrationalen Zahlen dran. Das wollen wir im folgenden Satz herausarbeiten.

Satz 6.4. *Für jedes $x \in \mathbb{R}$ und jedes $\varepsilon > 0$ existiert ein $q \in \mathbb{Q}$ mit $|x - q| < \varepsilon$.*

Beweis. **1. Fall $x \geq 0$:** Nach dem Satz von Archimedes 5.5 (e) gibt es ein $m \in \mathbb{N}$ mit $0 < 1/m < \varepsilon$. Weiter hat dank des Wohlordnungsprinzips aus Satz 5.7 die Menge $M := \{k \in \mathbb{N} : k/m > x\}$ ein Minimum $n := \min M$. Wegen $n \in M$ haben wir $n/m > x$ und, da $n-1$ nicht mehr zu M gehört, gilt $n-1/m \leq x$. Das liefert

$$-\varepsilon < 0 \leq x - \frac{n-1}{m} < \frac{n}{m} - \frac{n-1}{m} = \frac{1}{m} < \varepsilon,$$

wir können also $q := n-1/m$ setzen und erhalten $|x - q| < \varepsilon$.
2. Fall $x < 0$: Da $|x| > 0$ ist, gibt es nach den Überlegungen im 1. Fall ein $q \in \mathbb{Q}$ mit $0 \leq |x| - q < \varepsilon$. Nun gilt für $\tilde{q} = -q \in \mathbb{Q}$

$$x - \tilde{q} = -|x| + q = -(|x| - q) \in (-\varepsilon, 0], \quad \text{d. h.} \quad |x - \tilde{q}| < \varepsilon. \qquad \square$$

Diese Eigenschaft bekommt auch einen eigenen Namen.

Definition 6.5. *Eine Menge $M \subseteq \mathbb{R}$ heißt* dicht *in $\mathbb{R}$, wenn für jedes $x \in \mathbb{R}$ und jedes $\varepsilon > 0$ ein $y \in M$ existiert mit $|x - y| < \varepsilon$.*

$\mathbb{Q}$ ist also dicht in $\mathbb{R}$.

6.2. Wurzeln und rationale Exponenten

Zur Definition von Wurzeln benötigen wir zunächst den folgenden Hilfssatz.

Lemma 6.6. *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. Für alle $n \in \mathbb{N}$ haben wir nach Satz 5.11 (c)

$$x^n - y^n = (x - y) \sum_{k=0}^{n-1} x^{n-1-k} y^k.$$

Da $x, y \geq 0$ sind, ist die Summe auf der rechten Seite positiv, also gilt

$$x \leq y \iff x - y \leq 0 \iff x^n - y^n \leq 0 \iff x^n \leq y^n. \qquad \square$$

Wir beweisen nun zuerst die Existenz und Eindeutigkeit einer positiven n-ten Wurzel von positiven reellen Zahlen.

Satz 6.7. *Es sei $a \geq 0$ und $n \in \mathbb{N}$. Dann gibt es genau ein $x \in [0, \infty)$, sodass $x^n = a$ gilt.*

Beweis. Wir zeigen zunächst die Eindeutigkeit. Seien also $x, y \geq 0$ gegeben, sodass $x^n = a = y^n$ gilt. Dann gilt insbesondere $x^n \leq y^n$ und $y^n \leq x^n$ und mit Lemma 6.6 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

$$M_a := \{x \in [0, \infty) : x^n \leq a\}.$$

Dann ist in jedem Fall $0 \in M_a$, also $M_a \neq \emptyset$. Da $n \in \mathbb{N}$ ist, gilt $a \leq na \leq 1+na$ und daher mit der Bernoullischen Ungleichung, vgl. Satz 5.11 (a), auch $a \leq (1+a)^n$. Damit folgern wir nun für alle $x \in M_a$ die Abschätzung $x^n \leq a \leq (1+a)^n$, aus der mit Lemma 6.6 sofort $x \leq 1 + a$ folgt. Also ist M_a nach oben beschränkt. Nach dem Vollständigkeitsaxiom existiert damit $s := \sup M_a$. Wir nehmen an, es wäre $s^n \neq a$ und unterscheiden zwei Fälle.
1. Fall $s^n < a$: Für jedes $m \in \mathbb{N}$ gilt mit der Binomialformel aus Satz 5.11 (d)

$$\begin{aligned}\Big(s + \frac{1}{m}\Big)^n &= \sum_{k=0}^{n} \binom{n}{k} s^{n-k} \frac{1}{m^k} = s^n + \sum_{k=1}^{n} \binom{n}{k} s^{n-k} \underbrace{\frac{1}{m^k}}_{\leq 1/m} \\ &\leq s^n + \frac{1}{m} \sum_{k=1}^{n} \binom{n}{k} s^{n-k} =: s^n + \frac{1}{m} r. \end{aligned} \tag{6.1}$$

Man beachte, dass das so definierte r in jedem Fall größer null ist, wegen Lemma 6.6.
Nach dem Satz von Archimedes (Satz 5.5 (e)) und da nach Annahme $a - s^n > 0$ gilt, existiert ein $m_0 \in \mathbb{N}$ mit $1/m_0 < {}^{a-s^n}/r$. Dann gilt $s^n + r/m_0 < a$, also ist mit Hilfe von (6.1), wobei wir $m = m_0$ wählen,

$$\Big(s + \frac{1}{m_0}\Big)^n \leq s^n + \frac{1}{m_0} r < a.$$

Das liefert uns $s + 1/m_0 \in M_a$, und da s das Supremum dieser Menge war, gilt $s + 1/m_0 \leq s$, d. h. $1/m_0 \leq 0$. 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)

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

Nach dem Satz von Archimedes können wir ein $m \in \mathbb{N}$ mit $1/m \leq s$ finden. Dann gilt $-1/ms \geq -1$, wir können folglich die Bernoullische Ungleichung aus Satz 5.11 (a) anwenden und erhalten

$$\geq s^n \Big(1 - \frac{n}{ms}\Big). \tag{6.2}$$

Wir machen nun bei Bedarf unser m noch einmal größer, damit ${}^1\!/_m \leq s$ und ${}^1\!/_m < {}^{s(s^n-a)}\!/_{ns^n}$ gilt. Da nach Annahme $s^n - a > 0$ ist, haben wir für ein solches m

$$\frac{1}{m} < \frac{s(s^n-a)}{ns^n} \Longrightarrow \frac{ns^n}{ms} < s^n - a \Longrightarrow a < s^n - \frac{ns^n}{ms} = s^n\Big(1 - \frac{n}{ms}\Big)$$

und nehmen wir dies mit (6.2) zusammen, so folgt

$$\Big(s - \frac{1}{m}\Big)^n \geq s^n\Big(1 - \frac{n}{ms}\Big) > a.$$

Da s das Supremum von M_a ist, gibt es nach Satz 4.17 ein $x_0 \in M_a$ mit $x_0 > s - {}^1\!/_m$. Nach Lemma 6.6 gilt dann auch

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

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

Definition 6.8. *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*

$$\sqrt[n]{a} := a^{1/n} := x.$$

Ist $n = 2$, so sagt man einfach Wurzel von *a und schreibt kurz $\sqrt{a}$.*
Außerdem setzen wir für $a > 0$ wieder

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

Bemerkung 6.9. Als kleinen Einschub können wir nun sofort beschränkte Teilmengen von $\mathbb{Q}$ angeben, die kein Supremum in $\mathbb{Q}$ haben: zur Zahl $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$ betrachten wir zum Beispiel die Menge

$$M_2 := \{q \in \mathbb{Q} : q^2 < 2\}.$$

Dann ist M_2 nach oben beschränkt, besitzt aber kein Supremum in $\mathbb{Q}$ (Übung!). Die Menge $\mathbb{Q}$ erfüllt das Axiom (A15) also nicht.

Die Grundidee dieses Approximationsverfahrens verallgemeinern wir jetzt weiter für die Definition von Potenzen mit rationalen Exponenten.
Wir wissen für $a > 0$, was $a^{\pm n}$ und $a^{\pm 1/n}$ ist. Für $q = {}^m\!/_n$ mit $m, n \in \mathbb{N}$ ist die naheliegende Definition $a^q := (\sqrt[n]{a})^m$. Damit das wohldefiniert ist, muss aber zunächst geklärt werden, dass dieser Wert nicht von der speziellen Darstellung von q als Bruch abhängt.

Seien also $m, n, p, r \in \mathbb{N}$, sodass $q = {}^m\!/\!_n = {}^p\!/\!_r$ ist. Dann gilt $mr = np$ und es folgt für jedes $a \geq 0$

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

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

Definition 6.10. *Es sei $a \geq 0$, $q \in \mathbb{Q}$ mit $q > 0$ und $q = {}^m\!/\!_n$ $(m, n \in \mathbb{N})$. Dann setzen wir*

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

Falls $a > 0$ gilt, definieren wir für $q \in \mathbb{Q}$ mit $q < 0$,

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

Übungsaufgabe 6.11. Es gelten die bekannten Rechenregeln für rationale Exponenten, d. h. für alle $a, b \geq 0$ und alle $p, q \in \mathbb{Q}$ gilt

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

Übungsaufgabe 6.12. Überlegen Sie, wie die Exponenten a^x für allgemeine $x \in \mathbb{R}$ definiert werden könnten (passiert später in der Vorlesung).

Teil III.

Konvergenz

7. Folgen

Der Begriff der *Folge* ist für das Folgende fundamental.

Definition 7.1. *Sei X eine 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 $(a_n)_n$ oder manchmal mit $(a_n)_{n=1}^{\infty}$ oder $(a_n)_{n\in\mathbb{N}}$ oder $(a_n)_{n\geq 1}$ oder $(a_1, a_2, a_3, \dots)$ bezeichnet.

Beispiel 7.2. (a) Ist $X = \mathbb{R}$, so spricht man von einer *reellen Folge*. Ein Beispiel einer solchen Folge ist gegeben durch die Vorschrift

$$a_n := \frac{1}{n}, \quad \forall n \in \mathbb{N},$$

bzw.

$$(a_n)_n = \left(\frac{1}{n}\right)_n = \left(1, \frac{1}{2}, \frac{1}{3}, \dots\right).$$

(b) Für $X = \{0, 1\}$ sind $(a_n)_n = (1, 0, 1, 0, 1, 0, \dots)$ oder $(b_n)_n = (1, 1, 1, 1, 1, \dots)$ Beispiele für Folgen in $\{0, 1\}$. Dabei können wir $(a_n)_n$ und $(b_n)_n$ natürlich auch als reelle Folgen auffassen (warum?).

7.1. Konvergenz von reellen Folgen

Wir wollen uns nun dem zentralen Thema der Analysis zuwenden, der mathematisch exakten Behandlung des unendlich Kleinen und unendlich Großen. Dazu führen wir den für alles Weitere zentralen Begriff der *Konvergenz* ein. Insbesondere kann dadurch die *Approximation* von reellen Werten mathematisch exakt gefasst werden.

Definition 7.3. *Eine Folge $(a_n)_n$ in $\mathbb{R}$ heißt* konvergent *gegen $a \in \mathbb{R}$, wenn für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert mit*

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

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

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

Ist eine Folge nicht konvergent, so heißt sie divergent.

Für eine Umformulierung dieser Definition benötigen wir die folgenden Begriffe.

Definition 7.4. *(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 $n \in \mathbb{N}$, *wenn es ein $m \in \mathbb{N}$ gibt, sodass $A(n)$ für alle $n \geq m$ richtig ist.*

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

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0\ |a_n - a| < \varepsilon,$$

oder, genauso gut,

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

bzw. in Worte gefasst:

Für alle $\varepsilon > 0$ gilt $a_n \in U_\varepsilon(a)$ 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. Das formalisieren wir in der folgenden Bemerkung.

Bemerkung 7.5. (a) Sind $(a_n)_n$ und $(b_n)_n$ reelle Folgen mit $a_n = b_n$ für fast alle $n \in \mathbb{N}$, dann ist $(a_n)_n$ konvergent genau dann, wenn $(b_n)_n$ konvergiert, und in diesem Fall gilt $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

(b) Eine konvergente Folge mit Grenzwert null wird auch oft kurz eine *Nullfolge* genannt.

Beispiel 7.6. (a) Sei $a_n := {}^1\!/_n$, für alle $n \in \mathbb{N}$.

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

Beweis. Sei $\varepsilon > 0$. Nach dem Satz von Archimedes (Satz 5.5 (e)) gibt es ein $n_0 \in \mathbb{N}$ mit ${}^1\!/_{n_0} < \varepsilon$. Damit 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 $a_n := {}^{(-1)^n}\!/_n$, für alle $n \in \mathbb{N}$.

Behauptung: $(a_n)_n$ 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 aus (a) übernehmen. $\square$

(c) Sei $a_n := \dfrac{2n^2 - n + 5}{n^2+1}$, für alle $n \in \mathbb{N}$.

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

Beweis. Es gilt für alle $n \in \mathbb{N}$

$$|a_n - 2| = \left|\frac{2n^2 - n + 5 - 2n^2 - 2}{n^2+1}\right| = \frac{|3-n|}{n^2+1} \le \frac{|3-n|}{n^2} \le \frac{n+3}{n^2}, \quad (7.1)$$

wobei wir bei der letzten Abschätzung die Dreiecksungleichung angewendet haben. Nun verwenden wir noch, dass für alle $n \in \mathbb{N}$ gilt $n+3 \le n+3n = 4n$ und erhalten damit

$$|a_n - 2| \le \frac{4n}{n^2} = \frac{4}{n}. \quad (7.2)$$

Sei nun $\varepsilon > 0$. Nach dem Satz von Archimedes existiert ein $n_0 \in \mathbb{N}$, sodass ${}^1/_{n_0} < {}^\varepsilon/_4$ gilt. Dann haben wir nach obiger Abschätzung für alle $n \ge n_0$

$$|a_n - 2| \le \frac{4}{n} \le \frac{4}{n_0} < 4 \cdot \frac{\varepsilon}{4} = \varepsilon.$$ $\square$

(d) Sei $a_n := (-1)^n$, für alle $n \in \mathbb{N}$.

Behauptung: Die Folge $(a_n)_n$ 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 $n_0 \in \mathbb{N}$, sodass für jedes $n \ge n_0$ die Ungleichung $|a_n - a| < 1$ gilt. Für $n \ge n_0$ gilt dann aber mit Hilfe der Dreiecksungleichung

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

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

Satz 7.7. *Jede Folge in $\mathbb{R}$ hat höchstens einen Grenzwert.*

Beweis. Wir nehmen an, es gäbe eine Folge $(a_n)_n$ in $\mathbb{R}$ und $a, b \in \mathbb{R}$ mit $a \ne b$, sodass $(a_n)_n$ sowohl gegen a als auch gegen b konvergiert. Wir setzen jetzt $\varepsilon := {}^{|a-b|}/_2 > 0$, sodass insbesondere $U_\varepsilon(a) \cap U_\varepsilon(b) = \emptyset$ gilt. Nach der Definition des Grenzwertes existiert aber ein $n_0 \in \mathbb{N}$, sodass $a_n \in U_\varepsilon(a)$ für alle $n \ge n_0$ gilt. Also gilt $a_n \in U_\varepsilon(b)$ nur für höchstens endlich viele $n \in \mathbb{N}$. Damit kann $(a_n)_n$ nicht gegen b konvergieren. $\square$

Definition 7.8. *Eine reelle Folge* $(a_n)_n$ *heißt* beschränkt*, wenn die Menge der Folgenglieder* $\{a_n \in \mathbb{R} : n \in \mathbb{N}\} = \{a_1, a_2, a_3, \dots\}$ *beschränkt ist. Diese Menge besitzt dann ein Infimum und ein Supremum. Wir schreiben dafür*

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

Beachten Sie, dass eine Folge $(a_n)_n$ genau dann beschränkt ist, wenn ein $C \geq 0$ existiert, sodass $|a_n| \leq C$ für alle $n \in \mathbb{N}$ gilt, vgl. Übungsaufgabe 4.20.

Satz 7.9. *Jede konvergente Folge in* $\mathbb{R}$ *ist beschränkt.*

Beweis. Es sei $(a_n)_n$ eine konvergente Folge in $\mathbb{R}$ und $a := \lim_{n\to\infty} a_n$. Nach Definition der Konvergenz existiert zu $\varepsilon = 1$ ein $n_0 \in \mathbb{N}$, sodass $|a_n - a| < 1$ 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 für diese Indizes gilt

$$|a_n| = |a_n - a + a| \leq |a_n - a| + |a| < 1 + |a| \leq C. \tag{7.3}$$

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

Warnung 7.10. Die Umkehrung von Satz 7.9 ist falsch! Es gibt durchaus beschränkte Folgen, die nicht konvergieren, beispielsweise die Folge $((-1)^n)_n$ aus Beispiel 7.6 (d).

7.2. Grenzwertsätze

In diesem Abschnitt wollen wir die Berechnung von komplizierteren Grenzwerten angehen. Dabei werden wir auch einige wichtige Beispiele von konvergenten Folgen sehen, die im weiteren Verlauf immer wieder benötigt werden.

Satz 7.11. *Sei* $(a_n)_n$ *eine Folge in* $\mathbb{R}$.

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

(b) Ist $a \in \mathbb{R}$ *und* $(\beta_n)_n$ *eine Nullfolge mit* $|a_n - a| \leq \beta_n$ *für fast alle* $n \in \mathbb{N}$*, so ist* $\lim_{n\to\infty} a_n = a$.

Beweis. (a) Der Beweis ist enthalten in Hausaufgabe H 3.3 a)

(b) Nach Voraussetzung gibt es ein $m \in \mathbb{N}$, sodass $|a_n - a| \leq \beta_n$ für alle $n \geq m$ gilt. Sei nun $\varepsilon > 0$. Dann existiert dank der Konvergenz von $(\beta_n)_n$ ein $n_1 \in \mathbb{N}$ mit $0 \leq \beta_n < \varepsilon$ für alle $n \geq n_1$. Wir wählen $n_0 := \max\{m, n_1\}$. Dann gilt für alle $n \geq n_0$ die Abschätzung $|a_n - a| \leq \beta_n < \varepsilon$. □

Die folgenden sogenannten *Grenzwertsätze* enthalten äußerst wichtige Rechenregeln für Grenzwerte von Folgen. Mit ihnen ist es oft möglich, die Frage nach Konvergenz einer komplizierten Folge auf die Untersuchung von mehreren, aber dafür einfacheren Folgen zurückzuführen.

Satz 7.12 (Grenzwertsätze)**.** *Es seien $(a_n)_n$ und $(b_n)_n$ konvergente Folgen mit $a := \lim_{n\to\infty} a_n$ und $b := \lim_{n\to\infty} b_n$. Dann gilt:*

(a) $\lim_{n\to\infty} |a_n| = |a|$.

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

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

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

(e) Ist zusätzlich $a_n \neq 0$ für alle $n \in \mathbb{N}$ und $a \neq 0$, so ist $\lim_{n\to\infty} {}^1\!/_{a_n} = {}^1\!/_a$.

Beweis. (a) Nach der umgekehrten Dreiecksungleichung gilt

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

Da $(a_n)_n$ gegen a konvergiert, konvergiert nach Satz 7.11 (a) die Folge $(\beta_n)_n$ gegen null. Also können wir mit Satz 7.11 (b) folgern, dass $(|a_n|)_n$ gegen $|a|$ konvergiert.

(b)-(d) Übungsaufgabe

(e) Nach Teil (a) konvergiert die Folge $(|a_n|)_n$ gegen $|a|$ und nach Voraussetzung ist $|a| > 0$. Also gibt es zu $\varepsilon := {}^{|a|}\!/_2 > 0$ ein $m \in \mathbb{N}$ mit $||a_n| - |a|| < {}^{|a|}\!/_2$ für alle $n \geq m$. Damit gilt für alle diese n auch

$$|a_n| = |a| - \big(|a| - |a_n|\big) \geq |a| - \big||a| - |a_n|\big| > |a| - \frac{|a|}{2} = \frac{|a|}{2}.$$

Also ist ${}^1\!/_{|a_n|} < {}^2\!/_{|a|}$ für alle $n \geq m$ und wir können für diese n abschätzen:

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \left|\frac{a - a_n}{a_n a}\right| = \frac{|a_n - a|}{|a_n||a|} \leq \frac{2|a_n - a|}{|a|^2} = \frac{2}{a^2}|a_n - a| =: \beta_n. \quad (7.4)$$

Nach (c) und Satz 7.11 (a) gilt nun wieder $\beta_n \to 0$ $(n \to \infty)$ und damit folgt mit Hilfe von Satz 7.11 (b) die Behauptung. □

Das folgende Beispiel zeigt, wie mit Hilfe dieses Satzes schon ein bisschen kompliziertere Grenzwerte angegangen werden können.

Beispiel 7.13. (a) Sei $p \in \mathbb{N}$ fest gewählt und $a_n := 1/n^p$ für $n \in \mathbb{N}$. Wir wissen aus Beispiel 7.6 (a), dass die Folge $(1/n)_n$ gegen null konvergiert. Also gilt mit p-maliger Anwendung von Satz 7.12 (d)

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{1}{n^p} = \lim_{n\to\infty} \Big(\frac{1}{n}\Big)^p = \Big(\lim_{n\to\infty} \frac{1}{n}\Big)^p = 0^p = 0.$$

(b) Wir untersuchen

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

Dazu kürzen wir den Bruch mit der höchsten auftretenden Potenz:

$$a_n = \frac{n^2+2n+3}{n^2+3} = \frac{1+\frac{2}{n}+\frac{3}{n^2}}{1+\frac{3}{n^2}}.$$

Nun gilt $\lim_{n\to\infty} 2/n = 0$ nach Beispiel 7.6 (a) und Satz 7.12 (c) sowie $\lim_{n\to\infty} 3/n^2 = 0$ dank (a) mit $p = 2$ und wieder Satz 7.12 (c). Das bedeutet für die Folge im Zähler mit Hilfe von Satz 7.12 (b)

$$\lim_{n\to\infty} \Big(1+\frac{2}{n}+\frac{3}{n^2}\Big) = \lim_{n\to\infty} 1 + \lim_{n\to\infty} \frac{2}{n} + \lim_{n\to\infty} \frac{3}{n^2} = 1+0+0 = 1$$

und genauso für die Folge im Nenner $\lim_{n\to\infty}(1+3/n^2) = 1$. Damit sind die Voraussetzungen von Satz 7.12 (e) erfüllt und es gilt

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

Dieses Kürzen mit der höchsten auftretenden Potenz ist bei allen Grenzwerten der Form „Polynom in n geteilt durch Polynom in n“ Erfolg versprechend.

Ein weiteres wichtiges Werkzeug für Konvergenzuntersuchungen von reellen Folgen ist der folgende Satz, der zeigt, dass sich die Grenzwertbildung gut mit der Ordnungsrelation auf $\mathbb{R}$ verträgt.

Satz 7.14. *Es seien $(a_n)_n$, $(b_n)_n$ und $(c_n)_n$ reelle Folgen und $(a_n)_n$ und $(b_n)_n$ seien konvergent.*

(a) Ist $a_n \le b_n$ für fast alle $n \in \mathbb{N}$, so folgt $\lim_{n\to\infty} a_n \le \lim_{n\to\infty} b_n$.

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

Beweis. (a) Wir nehmen an, es wäre $\lim_{n\to\infty} a_n =: a > b := \lim_{n\to\infty} b_n$ und setzen $\varepsilon := {}^{a-b}\!/_2 > 0$. Dank der Konvergenz von $(a_n)_n$ und $(b_n)_n$ gibt es ein $n_0 \in \mathbb{N}$, sodass sowohl $b_n \in U_\varepsilon(b) = (b-\varepsilon, b+\varepsilon)$ als auch $a_n \in U_\varepsilon(a) = (a-\varepsilon, a+\varepsilon)$ für alle $n \geq n_0$ 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 $b_n < b+\varepsilon = a-\varepsilon < a_n$ im Widerspruch zur Voraussetzung.

(b) Sei $\varepsilon > 0$. Dann gibt es ein $n_0 \in \mathbb{N}$, sodass für alle $n \geq n_0$ sowohl $a_n \leq c_n \leq b_n$ als auch $|a_n - a| < \varepsilon$ und $|b_n - a| < \varepsilon$ gilt. Hieraus 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 $(c_n)_n$ gegen a. $\square$

Satz 7.15. *Es sei $(a_n)_n$ eine konvergente reelle Folge mit $a_n \geq 0$ für alle $n \in \mathbb{N}$. Dann gilt für alle $p \in \mathbb{N}$*

$$\lim_{n\to\infty} \sqrt[p]{a_n} = \sqrt[p]{\lim_{n\to\infty} a_n}.$$

Beweis. Sei $p \in \mathbb{N}$ beliebig und $a := \lim_{n\to\infty} a_n$. Wir betrachten zunächst den Fall $a = 0$. Sei dazu $\varepsilon > 0$ gegeben. Dann gilt auch $\varepsilon^p > 0$, also gibt es ein $n_0 \in \mathbb{N}$, sodass $a_n < \varepsilon^p$ für alle $n \geq n_0$ gilt. Damit ist nach Lemma 6.6 für diese n auch $\sqrt[p]{a_n} < \varepsilon$, also $|\sqrt[p]{a_n} - 0| < \varepsilon$ und wir sind fertig.

Ab jetzt sei also $a > 0$. Dann gilt für alle $n \in \mathbb{N}$ nach Satz 5.11 (c) (setze dort $n := p - 1$)

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

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 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

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

Setzen wir $c := (\sqrt[p]{a})^{p-1}$, so haben wir für alle $n \in \mathbb{N}$

$$\left|\sqrt[p]{a_n} - \sqrt[p]{a}\right| \leq \frac{1}{c}|a_n - a| =: \beta_n. \tag{7.5}$$

Man beachte, dass $c > 0$ ist, sodass diese Umformung erlaubt ist.
Die Folge $(|a_n - a|)_n$ konvergiert nach Satz 7.11 (a) gegen null. Also ist $(\beta_n)_n$ dank Satz 7.12 (c) eine Nullfolge. Damit folgt $\lim_{n\to\infty} \sqrt[p]{a_n} = \sqrt[p]{a}$ aus Satz 7.11 (b). □

Beispiel 7.16. (a) Wir betrachten

$$a_n := \sqrt{n+1} - \sqrt{n}, \quad n \in \mathbb{N}.$$

Für alle $n \in \mathbb{N}$ gilt mit Hilfe der dritten binomischen Formel

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

Kürzen mit $\sqrt{n}$ und anschließende Anwendung der Grenzwertsätze zusammen mit Satz 7.15 liefert

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

(b) Wir wandeln $(a_n)_n$ leicht ab und betrachten

$$b_n := \sqrt{n}\big(\sqrt{n+1} - \sqrt{n}\big) = \sqrt{n} \cdot a_n, \quad n \in \mathbb{N}.$$

Dann gilt wie oben

$$b_n = \frac{\sqrt{n}}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{\frac{n+1}{n}} + 1} \longrightarrow \frac{1}{1+1} = \frac{1}{2} \quad (n \to \infty).$$

Im folgenden Satz betrachten wir Folgen mit n-ten Wurzeln.

Satz 7.17. *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 $1^n = 1 \leq n$ und damit $1 \leq \sqrt[n]{n}$ nach Lemma 6.6. Also ist $\sqrt[n]{n} = 1 + a_n$ mit einer Folge $(a_n)_n$, für die $a_n \geq 0$ für alle $n \in \mathbb{N}$ gilt. Wir untersuchen nun die Folge $(a_n)_n$ auf Konvergenz und erledigen damit sofort auch $(\sqrt[n]{n})_n$. 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 $\lim_{n\to\infty} {}^2/_{n-1} = 0$ ist, konvergiert nach Satz 7.15 mit $p = 2$ die rechte Seite obiger Ungleichungskette ebenfalls gegen null. Das Sandwich-Theorem 7.14 (b) 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 nach dem Satz von Archimedes 5.5 (c)). Dann gilt nach Lemma 6.6

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

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 hingegen $0 < c < 1$, 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 7.12 (e) ist damit $\lim_{n\to\infty} \sqrt[n]{c} = 1$. □

Einen weiteren wichtigen Grenzwert behandelt der folgende Satz.

Satz 7.18. *Sei $q \in \mathbb{R}$ beliebig. Dann ist die Folge $(q^n)_n$ genau dann konvergent, wenn $q \in (-1, 1]$ ist und es gilt*

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

Beweis. Wir betrachten zunächst die einfachen Fälle $q \in \{-1, 0, 1\}$. Ist $q = 0$, so ist $q^n = 0$ für alle $n \in \mathbb{N}$, die Folge konvergiert also gegen null. Im Falle $q = 1$ findet man genauso wegen $q^n = 1$ für alle $n \in \mathbb{N}$ Konvergenz gegen eins. Für $q = -1$ erhalten wir die schon aus Beispiel 7.6 (d) 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 $(q^n)_n$ nicht beschränkt, denn gäbe es ein $M \geq 0$ mit $|q^n| \leq M$ für alle $n \in \mathbb{N}$, so wäre $n \leq {}^{|q^n|}/_p \leq {}^M/_p$ für alle $n \in \mathbb{N}$, also $\mathbb{N}$ beschränkt. Nach Satz 7.9 kann damit $(q^n)_n$ 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

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

Also ist $|q^n| \leq 1/np$ für alle $n \in \mathbb{N}$ und da $(1/np)_n$ eine Nullfolge ist, konvergiert nach Satz 7.11 (b) die Folge $(q^n)_n$ gegen null. $\square$

7.3. Bestimmte Divergenz

Man ist zuweilen versucht zu sagen, dass eine Folge „gegen unendlich geht". Das ist mit unserem bisherigen Konvergenzbegriff aus gutem Grunde nicht möglich. Wir können aber die folgende Sprechweise einführen.

Definition 7.19. *Eine reelle Folge $(a_n)_n$*

(a) divergiert bestimmt gegen ∞, *falls für jedes $C > 0$ ein $n_0 = n_0(C) \in \mathbb{N}$ existiert, sodass $a_n \geq C$ für alle $n \geq n_0$ gilt.*

In diesem Fall schreiben wir analog zum Grenzwert

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

(b) divergiert bestimmt gegen $-\infty$, *falls für jedes $C > 0$ ein $n_0 = n_0(C) \in \mathbb{N}$ existiert, sodass $a_n \leq -C$ für alle $n \geq n_0$ gilt.*

Auch hier schreiben wir

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

Übungsaufgabe 7.20. (a) Zeigen Sie: Ist $(a_n)_n$ eine bestimmt divergente Folge gegen plus oder minus unendlich, so gilt $a_n \neq 0$ für fast alle $n \in \mathbb{N}$. Außerdem ist die Folge $(1/a_{n+k})_n$ für ein geeigenetes $k \in \mathbb{N}$ eine Nullfolge.

(b) Achtung: ist $(b_n)_n$ eine Nullfolge mit $b_n \neq 0$ für alle $n \in \mathbb{N}$, so ist $(1/b_n)_n$ im Allgemeinen nicht bestimmt divergent. Geben Sie hierzu ein Beispiel an.

(c) Beweisen Sie, dass für jede Nullfolge $(b_n)_n$ mit $b_n \neq 0$ für alle $n \in \mathbb{N}$ die Folge $(1/|b_n|)_n$ bestimmt gegen ∞ divergiert.

(d) Untersuchen Sie, ob oder inwiefern sich die Grenzwertsätze aus dem vorigen Abschnitt auf bestimmt divergente Folgen übertragen lassen.

7.4. Monotonie von Folgen

In diesem Abschnitt definieren wir die nötigen Begriffe, um über das Monotonie-Verhalten von Folgen zu sprechen. Damit werden wir ein neues wichtiges Konvergenzkriterium herleiten, mit dem wir am Ende die Eulersche Zahl e definieren können.

Definition 7.21. *Eine reelle Folge $(a_n)_n$ 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 7.22 (Monotonie-Kriterium). *Eine monotone reelle Folge $(a_n)_n$ ist genau dann konvergent, wenn sie beschränkt ist. In diesem Fall gilt*

$$\lim_{n\to\infty} a_n = \begin{cases} \sup_{n\in\mathbb{N}} a_n, & \text{wenn } a_n \text{ monoton wachsend ist,} \\ \inf_{n\in\mathbb{N}} a_n, & \text{wenn } a_n \text{ monoton fallend ist.} \end{cases}$$

Beweis. Da jede konvergente Folge nach Satz 7.9 beschränkt ist, müssen wir für die behauptete Äquivalenz nur zeigen, dass jede monotone und beschränkte Folge konvergent ist. Dazu betrachten wir den Fall, dass $(a_n)_n$ monoton wächst, der Beweis für monoton fallende Folgen geht dann analog.
Sei $\varepsilon > 0$ und wir setzen $a := \sup_{n\in\mathbb{N}} a_n$. Dann existiert nach Satz 4.17 ein $n_0 \in \mathbb{N}$ mit $a_{n_0} > a - \varepsilon$. Damit gilt für alle $n \geq n_0$ wegen der Monotonie und der Beschränktheit von $(a_n)_n$

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

Das bedeutet $|a_n - a| < \varepsilon$ für alle $n \geq n_0$ und damit genau die Konvergenz von $(a_n)_n$ gegen a. □

Beispiel 7.23. 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 für das n-te Folgenglied gegeben, sondern nur ein Startwert a_1 und dann eine Formel,

wie man aus einem Folgenglied das jeweils nächste berechnen kann. Auch dadurch ist die Folge eindeutig bestimmt.
Hier erhält man für die ersten Folgenglieder

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

Um das Monotoniekriterium anzuwenden, prüfen wir die Voraussetzungen nach, d. h. wir zeigen, dass $(a_n)_n$ 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) $(a_n)_n$ konvergiert und $\lim_{n\to\infty} a_n = 2$.

Beweis. (a) Wir gehen 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 7.22 wissen wir nun, dass $(a_n)_n$ 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 7.12 konvergieren bei dieser Gleichung die Folgen auf beiden Seiten des Gleichheitszeichens. Gehen wir zum Limes über, 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 $a = 2$. Polynomdivision liefert $(a-2)(a^2 + 2a + 3) = 0$ und $a^2 + 2a + 3 = (a + 1)^2 + 2 = 0$ hat keine weiteren reellen Lösungen. Also muss $a = 2$ sein. □

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

$$a_n := \left(1 + \frac{1}{n}\right)^n, \qquad n \in \mathbb{N},$$

$$b_n := \sum_{k=0}^{n} \frac{1}{k!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}, \quad n \in \mathbb{N}.$$

Warnung 7.24. Die Folge $(a_n)_n$ bietet eine gute Gelegenheit, vor einem verbreiteten Fehler bei der Bestimmung von Grenzwerten zu warnen. Man könnte auf die Idee kommen, die Stellen, an denen das n vorkommt, nacheinander zu bearbeiten. Dazu benennen wir z. B. das n im Exponenten in k um und lassen zunächst das n und danach dann das k gegen unendlich laufen. Das liefert

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

Wenn das ein gültiges Verfahren wäre, sollte es aber natürlich auch andersherum funktionieren. Bilden wir zunächst den Grenzwert für k gegen unendlich und dann den für n, so finden wir, dass

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

nicht existiert. Wir werden gleich sehen, dass der Grenzwert von $(a_n)_n$ existiert und nicht 1 ist. Beide Betrachtungen führen also in die Irre.
Das ist ein Beispiel für eine Grundschwierigkeit in der Analysis und führt gleich zu einer weiteren Warnung: Grenzwerte lassen sich im Allgemeinen nicht vertauschen! Das gilt sogar, wenn alle beteiligten Grenzwerte existieren, wie das folgende Beispiel zeigt:

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

Nach diesen Vorbemerkungen wenden wir uns jetzt der Behandlung der beiden oben angegebenen Folgen zu.

Satz 7.25. *Die Folgen $(a_n)_n$ und $(b_n)_n$ konvergieren und* $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

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

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

Die Folge $(b_n)_n$ ist also streng monoton wachsend und es bleibt noch Beschränktheit zu zeigen. Ein Beweis dazu ist auf dem zweiten Übungsblatt, Hausaufgabe 2.3: Insbesondere gilt $b_n \le 3$ für alle $n \in \mathbb{N}$. Also ist mit Hilfe des Monotoniekriteriums aus Satz 7.22 die Folge $(b_n)_n$ konvergent und für $b := \lim_{n\to\infty} b_n$ gilt $b \le 3$.

Wir wenden uns der Folge $(a_n)_n$ zu. Um für diese Folge Monotonie nachzuweisen, betrachten wir 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

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

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 $(a_n)_n$ streng monoton wachsend. Die Beschränktheit dieser Folge spielen wir nun auf die Beschränktheit von $(b_n)_n$ zurück. Nach der Binomialformel gilt

$$\begin{aligned}a_n = \left(1+\frac{1}{n}\right)^{n} &= \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} \\ &= 1+1+\sum_{k=2}^{n}\frac{n!}{k!(n-k)!}\cdot\frac{1}{n^k} \\ &= 1+1+\sum_{k=2}^{n}\frac{1}{k!}\cdot\frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k}.\end{aligned}$$

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

$$= 1+1+\sum_{k=2}^{n}\frac{1}{k!}\frac{n-1}{n}\cdot\frac{n-2}{n}\cdots\frac{n-k+1}{n}. \tag{7.6}$$

Schließlich beobachten wir, dass jeder dieser Faktoren 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 $(a_n)_n$ 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 (7.6)

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

Wir gehen nun in dieser Ungleichung auf beiden Seiten zum Grenzwert $n \to \infty$ über. Man beachte, dass wir es dabei 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 7.12 ist also hier anwendbar.

Da jeder einzelne Klammerausdruck gegen 1 strebt, geht jeder Summand gegen $1/k!$ und damit strebt der gesamte Ausdruck auf der rechten Seite genau gegen $1 + 1 + \sum_{k=2}^{j} 1/k! = b_j$. 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 7.26. *Die Zahl*

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

heißt Eulersche Zahl.

7.5. Teilfolgen und Häufungswerte

Auch divergente Folgen enthalten unter Umständen so etwas wie konvergente Anteile. Die zu deren Beschreibung nützlichen Begriffe wollen wir in diesem Abschnitt herausarbeiten.

Definition 7.27. *Sei $(a_n)_n$ eine Folge, $\varphi : \mathbb{N} \to \mathbb{N}$ eine streng monoton wachsende Funktion und $b_n := a_{\varphi(n)}$, $n \in \mathbb{N}$. Dann heißt die Folge $(b_n)_n$* Teilfolge *von $(a_n)_n$.*

Diese Definition sieht etwas sperrig aus, tut aber genau das, wonach der Name sich anhört: Einen Teil der Folge herauspicken. Man sucht eine gewisse (unendliche) Auswahl von Folgengliedern heraus, ohne deren Reihenfolge zu ändern. Um diesen zweiten Punkt zu gewährleisten, muss φ streng monoton wachsend sein. Wir betrachten zwei Beispiele.

Beispiel 7.28. (a) Für $\varphi(k) := 2k$, $k \in \mathbb{N}$, haben wir $b_k = a_{2k}$, also $(b_n)_n = (a_2, a_4, a_6, a_8, \dots)$ 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 $(b_n)_n = (a_1, a_4, a_9, a_{16}, \dots)$.

Bemerkung 7.29. Zur Notation von Teilfolgen wird oft statt der Funktionsnotation wieder die Indexnotation wie bei Folgen üblich verwendet. Man setzt $n_k := \varphi(k)$ für $k \in \mathbb{N}$ und schreibt für die Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ oder einfach $(a_{n_k})_k$. Diese entsteht also aus $(a_n)_n$, indem die Indizes $n_1 < n_2 < n_3 < \dots$ ausgewählt werden.

Definition 7.30. *Es sei $(a_n)_n$ eine reelle Folge. Eine Zahl $\alpha \in \mathbb{R}$ heißt* Häufungswert *von $(a_n)_n$, wenn für jedes $\varepsilon > 0$ die Menge $\{n \in \mathbb{N} : a_n \in U_\varepsilon(\alpha)\}$ unendlich viele Elemente enthält.*
Weiter definieren wir die Menge aller Häufungswerte der Folge als

$$\mathrm{HW}(a_n) := \big\{\alpha \in \mathbb{R} : \alpha \textit{ ist Häufungswert von } (a_n)_n\big\}.$$

Bemerkung 7.31. Diese Definition sieht unserer Umformulierung der Konvergenz-Definition aus Kapitel 7.1 ähnlich. Wir stellen die beiden gegenüber:

- Es ist $\lim_{n\to\infty} a_n = a$, falls

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

- Es ist $a \in \mathrm{HW}(a_n)$, falls

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

Nun sieht man deutlich, dass die Anforderung an einen Häufungswert schwächer ist als an einen Grenzwert.

Beispiel 7.32. (a) Wir zeigen, dass $(a_n)_n = ((-1)^n)_n$ genau zwei Häufungswerte hat, nämlich 1 und -1.

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 ein Häufungswert. Genauso sieht man durch Betrachtung der Folgenglieder mit ungeradem Index, dass -1 ein Häufungswert ist.

Es bleibt noch zu zeigen, dass alle anderen reellen Zahlen keine Häufungswerte sind. Sei dazu $\beta \in \mathbb{R}$ mit $\beta \neq 1$ und $\beta \neq -1$. Wähle dann $\varepsilon_0 > 0$ so klein, dass $1, -1 \notin U_{\varepsilon_0}(\beta)$ gilt. Das geht, da β sowohl von 1 als auch von -1 einen echt positiven Abstand hat. Dann gilt $a_n \in U_{\varepsilon_0}(\beta)$ für gar kein $n \in \mathbb{N}$. Also kann β kein Häufungswert sein.

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

(c) Wie Sie in den "Grundlagen der Mathematik" sehen oder gesehen haben, ist $\mathbb{Q}$ abzählbar. Es gibt also eine Folge $(a_n)_n$, sodass $\mathbb{Q} = \{a_n \in \mathbb{R} : n \in \mathbb{N}\}$ 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 ein Häufungswert der Folge $(a_n)_n$. Wir haben eine Folge gefunden, für die *jede* reelle Zahl ein Häufungswert ist.

Der folgende Satz befasst sich mit den Zusammenhängen zwischen Teilfolgen, Häufungswerten und Konvergenz.

Satz 7.33. *Es sei $(a_n)_n$ eine reelle Folge. Dann gilt*

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

(b) Ist $(a_n)_n$ konvergent und $(a_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(a_n)_n$, so ist auch $(a_{n_k})_{k\in\mathbb{N}}$ konvergent und es gilt $\lim_{k\to\infty} a_{n_k} = \lim_{n\to\infty} a_n$.

(c) Ist $(a_n)_n$ konvergent, so hat $(a_n)_n$ 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 $(a_n)_n$. Dann existiert 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 $(a_n)_n$ in der $1/2$-Umgebung von α gibt, muss es auch ein $n_2 \in \mathbb{N}$ mit $n_2 > n_1$ geben, sodass $|a_{n_2} - \alpha| < 1/2$ gilt. Genauso finden wir ein $n_3 \in \mathbb{N}$ mit $n_3 > n_2$, sodass $|a_{n_3} - \alpha| < 1/3$ gilt.

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

$$|a_{n_k} - \alpha| \leq \frac{1}{k} \quad \text{für alle } k \in \mathbb{N}. \tag{7.7}$$

Damit ist $(a_{n_k})_{k\in\mathbb{N}}$, eine Teilfolge von $(a_n)_n$, von der noch zu zeigen ist, dass sie gegen α konvergiert. Sei dazu $\varepsilon > 0$. Dann gibt es ein $k_0 \in \mathbb{N}$, sodass $1/k_0 < \varepsilon$ ist und mit (7.7) gilt für alle $k \geq k_0$

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

Wir wenden uns nun der Richtung von rechts nach links zu. Sei also $(a_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(a_n)_n$ mit $a_{n_k} \to \alpha$ $(k \to \infty)$. Zu gegebenem $\varepsilon > 0$ existiert dann ein $k_0 \in \mathbb{N}$, sodass $|a_{n_k} - \alpha| < \varepsilon$ 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. D. h. α ist ein Häufungswert von $(a_n)_n$.

(b) Wir setzen $a := \lim_{n\to\infty} a_n$ und geben ein $\varepsilon > 0$ vor. Dann gibt es ob der Konvergenz von $(a_n)_n$ ein $n_0 \in \mathbb{N}$, sodass $|a_n - a| < \varepsilon$ 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 $|a_{n_k} - a| < \varepsilon$ für alle $k \geq k_0$ folgt.

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

Übungsaufgabe 7.34. Es sei $(a_n)_n$ eine reelle Folge und $\lambda \in \mathbb{R} \setminus \{0\}$. Zeigen Sie

$$\mathrm{HW}(\lambda a_n) = \{\lambda\alpha : \alpha \in \mathrm{HW}(a_n)\}.$$

7.6. Beschränkte Folgen

Beschränkte reelle Folgen sind im Allgemeinen nicht konvergent, aber mit den im letzten Abschnitt eingeführten Werkzeugen „Teilfolge“ und „Häufungswert“ haben wir gute Mittel in der Hand, sie genauer zu untersuchen. Zunächst zeigen wir, dass beschränkte Folgen zumindest immer einen Häufungswert und damit auch eine konvergente Teilfolge besitzen.

Satz 7.35 (Satz von Bolzano-Weierstraß)**.** *Jede beschränkte Folge in $\mathbb{R}$ hat mindestens einen Häufungswert.*

Beweis. Sei $(a_n)_n$ eine beschränkte Folge in $\mathbb{R}$ und $C \geq 0$ so gewählt, dass $|a_n| \leq C$ für alle $n \in \mathbb{N}$ gilt. Wir betrachten die Menge

$$M := \{x \in \mathbb{R} : a_n \leq x \text{ für fast alle } n \in \mathbb{N}\}.$$

Diese ist wegen $C \in M$ nicht-leer. Ist außerdem $x < -C$, so gibt es gar keinen Index n mit $a_n \leq x$, also kann ein solches x nicht in M liegen. Das zeigt uns, dass M durch $-C$ nach unten beschränkt ist. Damit existiert $\alpha := \inf M$ nach Satz 4.15. Wir zeigen nun, dass α ein Häufungswert von $(a_n)_n$ ist.
Sei dazu $\varepsilon > 0$. Dann gibt es nach Satz 4.17 ein $x_\varepsilon \in [\alpha, \alpha + \varepsilon)$, das in M liegt, d. h. für fast alle $n \in \mathbb{N}$ gilt $a_n \leq x_\varepsilon \leq \alpha + \varepsilon$. Damit liegen höchstens endlich viele Folgenglieder oberhalb von $\alpha + \varepsilon$. Andererseits ist $\alpha - \varepsilon$ nicht in M, d. h. es gibt unendlich viele a_n, die oberhalb von $\alpha - \varepsilon$ liegen.
Zusammengenommen bedeutet das, dass unendlich viele Folgenglieder im Intervall $(\alpha - \varepsilon, \alpha + \varepsilon)$ liegen müssen und wir sind fertig. □

Satz 7.33 ermöglicht sofort die folgende Umformulierung des Satzes von Bolzano-Weierstraß.

Korollar 7.36 (Satz von Bolzano-Weierstraß). *Jede beschränkte Folge in $\mathbb{R}$ hat eine konvergente Teilfolge.*

Man kann über die Häufungswerte einer beschränkten Folge sogar noch mehr sagen.

Satz 7.37. *Es sei $(a_n)_n$ eine beschränkte Folge in $\mathbb{R}$.*

(a) Für alle $\alpha \in \mathrm{HW}((a_n)_n)$ gilt $\inf_{n\in\mathbb{N}} a_n \le \alpha \le \sup_{n\in\mathbb{N}} a_n$.

(b) Die Menge $\mathrm{HW}((a_n)_n)$ hat ein Minimum und ein Maximum.

Beweis. (a) Sei $\alpha > \sup_{n\in\mathbb{N}} a_n$. Dann ist $\varepsilon := \alpha - \sup_{n\in\mathbb{N}} a_n > 0$ und für alle $n \in \mathbb{N}$ gilt

$$|\alpha - a_n| = \alpha - a_n \ge \alpha - \sup_{n\in\mathbb{N}} a_n = \varepsilon.$$

Also liegt kein Folgenglied unserer Folge in $U_{\varepsilon/2}(\alpha)$ und damit kann α kein Häufungswert von $(a_n)_n$ sein.

Analog zeigt man, dass jedes $\alpha < \inf_{n\in\mathbb{N}} a_n$ kein Häufungswert ist.

(b) Nach dem Satz von Bolzano-Weierstraß ist $\mathrm{HW}((a_n)_n) \neq \emptyset$ und nach Teil (a) ist diese Menge auch beschränkt. Also hat sie ein Supremum und ein Infimum. Wir zeigen, dass auch $\alpha := \sup \mathrm{HW}((a_n)_n)$ ein Häufungswert von $(a_n)_n$ und damit das Maximum dieser Menge ist. Ein analoges Argument für das Infimum liefert dann die Behauptung.

Sei $\varepsilon > 0$. Nach Satz 4.17 gibt es ein $\beta \in \mathrm{HW}((a_n)_n)$ mit $\alpha - \varepsilon < \beta \le \alpha$. Wählen wir nun ein $\delta \in (0, \varepsilon - (\alpha - \beta))$, so gilt $U_\delta(\beta) \subseteq U_\varepsilon(\alpha)$. (Malen Sie sich ein Bild!) Da β ein Häufungswert der Folge ist, liegen unendlich viele Folgenglieder in $U_\delta(\beta)$ und damit auch in $U_\varepsilon(\alpha)$. Also ist α ein Häufungswert von $(a_n)_n$. $\square$

Das ermöglicht uns die folgende Definition.

Definition 7.38. *Es sei $(a_n)_n$ eine beschränkte Folge in $\mathbb{R}$. Dann heißt*

$$\limsup_{n\to\infty} a_n := \max \mathrm{HW}((a_n)_n) \text{ oberer Limes } \textit{oder} \text{ Limes superior } \textit{von } (a_n)_n,$$
$$\liminf_{n\to\infty} a_n := \min \mathrm{HW}((a_n)_n) \text{ unterer Limes } \textit{oder} \text{ Limes inferior } \textit{von } (a_n)_n.$$

Beispiel 7.39. In Beispiel 7.32 haben wir gesehen, dass die Folge $((-1)^n)_n$ genau die Häufungswerte 1 und -1 hat. Also gilt hier

$$\liminf_{n\to\infty} (-1)^n = -1 \quad \text{und} \quad \limsup_{n\to\infty} (-1)^n = 1.$$

Bemerkung 7.40. (a) Etwas allgemeiner als in Definition 7.38 kann man den Limes superior einer Folge auch definieren, wenn die Folge nur nach oben beschränkt ist und für den Limes inferior reicht Beschränkheit nach unten aus.

(b) Für jede konvergente Folge $(a_n)_n$ ist $\mathrm{HW}((a_n)_n)$ nach Satz 7.33 (c) einelementig. Also ist in diesem Fall

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

Für beschränkte Folgen gilt in Teil (b) sogar die Umkehrung. Das ist ein Teil des folgenden Satzes, der einige unserer Erkenntnisse aus diesem und dem vorigen Kapitel für beschränkte Folgen zusammenfasst.

Satz 7.41. *Es sei $(a_n)_n$ eine beschränkte Folge in $\mathbb{R}$. Dann sind die folgenden Aussagen äquivalent:*

(a) $(a_n)_n$ *ist konvergent.*

(b) $(a_n)_n$ *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)$\Leftrightarrow$(c)“ ergibt sich direkt aus der Definition von Limes inferior und Limes superior. Außerdem ist die Implikation „(a)$\Rightarrow$(b)“ die Aussage von Satz 7.33 (c). Es bleibt uns also nur noch die Implikation „(b)$\Rightarrow$(a)“ zu zeigen. Sei dazu $(a_n)_n$ eine beschränkte Folge mit genau einem Häufungswert α und wir nehmen an, diese sei nicht konvergent. Da dann $(a_n)_n$ insbesondere nicht gegen α konvergiert, gibt es ein $\varepsilon > 0$, für das außerhalb von $U_\varepsilon(\alpha)$ unendlich viele Folgenglieder liegen. Fassen wir diese zu einer Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ zusammen, so ist diese als Teilfolge der beschränkten Folge $(a_n)_n$ ebenfalls beschränkt. Sie besitzt also nach dem Satz von Bolzano-Weierstraß 7.35 eine konvergente Teilfolge $(a_{n_{k_\ell}})_\ell$. Da $(a_{n_{k_\ell}})_\ell$ insbesondere eine Teilfolge der ursprünglichen Folge $(a_n)_n$ ist, ist ihr Grenzwert nach Satz 7.33 (a) ein Häufungswert von $(a_n)_n$. Also muss $\lim_{\ell\to\infty} a_{n_{k_\ell}} = \alpha$ sein. Das ist aber ein Widerspruch dazu, dass $a_{n_{k_\ell}} \notin U_\varepsilon(\alpha)$ für alle $\ell \in \mathbb{N}$ gilt. □

Man beachte, dass eine Folge durchaus einmal größer als der Limes superior oder kleiner als der Limes inferior werden kann. Es gibt dabei aber Grenzen. Das formulieren wir exakter im folgenden Satz.

Satz 7.42. *Ist $(a_n)_n$ eine beschränkte Folge in $\mathbb{R}$ und sind $a, b \in \mathbb{R}$ mit $a < \liminf_{n\to\infty} a_n$ und $b > \limsup_{n\to\infty} a_n$, so gilt $a_n \in (a, b)$ für fast alle $n \in \mathbb{N}$.*

Beweis. Wir nehmen an, es gäbe unendlich viele Folgenglieder, die größer oder gleich b sind. Dann können wir diese als eine Teilfolge von $(a_n)_n$ auffassen. Dank der Beschränktheit der gesamten Folge ist diese Teilfolge ebenfalls beschränkt, sodass sie nach dem Satz von Bolzano-Weierstraß 7.35 wiederum eine konvergente Teilfolge hat. Diese Teilfolge der Teilfolge ist weiterhin eine Teilfolge der Ausgangsfolge $(a_n)_n$ und wir bezeichnen sie mit $(a_{n_k})_k$. Der Grenzwert $\beta := \lim_{k\to\infty} a_{n_k}$ ist zum einen nach Satz 7.33 (a) ein Häufungswert der Folge $(a_n)_n$, zum anderen haben wir durch unsere Konstruktion sichergestellt, dass $a_{n_k} \geq b$ für jedes $k \in \mathbb{N}$ gilt. Damit erhalten wir dank der Monotonie des Grenzwerts, vgl. Satz 7.14 (a),

$$\beta = \lim_{k\to\infty} a_{n_k} \geq b.$$

Das führt aber auf den Widerspruch

$$\limsup_{n\to\infty} a_n = \max \mathrm{HW}((a_n)_n) \geq \beta \geq b > \limsup_{n\to\infty} a_n.$$

Die Annahme, es gäbe unendlich viele Folgenglieder, die kleiner oder gleich a sind, führt man analog zu einem Widerspruch. □

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

Satz 7.43. *Es seien $(a_n)_n$ und $(b_n)_n$ beschränkte Folgen in $\mathbb{R}$. Dann gilt*

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

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

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

(c) $\limsup_{n\to\infty}(a_n + b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$ *und*
$\liminf_{n\to\infty}(a_n + b_n) \geq \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$.

(d) Ist $\lambda \geq 0$, so gilt

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

(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 in (b) und (c) jeweils nur die Aussagen für den Limes superior. Die Beweise von

(d),(e) sind Übungsaufgabe.

(a) Nach der Definition des oberen und unteren Limes gilt

$$\liminf_{n\to\infty} a_n = \inf \mathrm{HW}(a_n) \leq \sup \mathrm{HW}(a_n) = \limsup_{n\to\infty} a_n.$$

(b) Sei $\alpha := \limsup_{n\to\infty} a_n$. Dann ist $\alpha \in \mathrm{HW}((a_n)_n)$. Also gibt es nach Satz 7.33 (a) eine Teilfolge $(a_{n_k})_k$, die gegen α konvergiert. Betrachten wir die analog gebildete Teilfolge $(b_{n_k})_k$ von $(b_n)_n$, so ist diese ebenfalls beschränkt, hat also nach dem Satz von Bolzano-Weierstraß 7.35 wiederum eine konvergente Teilfolge $(b_{n_{k_\ell}})_\ell$, deren Grenzwert β nach Satz 7.33 (a) ein Häufungswert von $(b_n)_n$ ist. Da nach Voraussetzung $a_{n_{k_\ell}} \leq b_{n_{k_\ell}}$ für fast alle $\ell \in \mathbb{N}$ gilt, folgt mit Hilfe von Satz 7.33 (b) und Satz 7.14 (a)

$$\alpha = \lim_{k\to\infty} a_{n_k} = \lim_{\ell\to\infty} a_{n_{k_\ell}} \leq \lim_{\ell\to\infty} b_{n_{k_\ell}} = \beta.$$

Also ist

$$\limsup_{n\to\infty} a_n = \alpha \leq \beta \leq \max \mathrm{HW}((b_n)_n) = \limsup_{n\to\infty} b_n.$$

(c) Sei $\alpha := \limsup_{n\to\infty}(a_n + b_n)$ und $\varepsilon > 0$. Da α ein Häufungspunkt von $(a_n + b_n)_n$ ist, gibt es nach Satz 7.33 (a) eine Teilfolge $(a_{n_k} + b_{n_k})_k$ von $(a_n + b_n)_n$, die gegen α konvergiert. Weiter gilt nach Satz 7.42 für fast alle $k \in \mathbb{N}$

$$a_{n_k} \leq \limsup_{n\to\infty} a_n + \frac{\varepsilon}{2} \quad \text{und} \quad b_{n_k} \leq \limsup_{n\to\infty} b_n + \frac{\varepsilon}{2}.$$

Damit können wir aus der Monotonie des Grenzwertes in Satz 7.14 (a)

$$\begin{aligned}\alpha = \lim_{k\to\infty}(a_{n_k} + b_{n_k}) &\leq \lim_{k\to\infty}\Big(\limsup_{n\to\infty} a_n + \frac{\varepsilon}{2} + \limsup_{n\to\infty} b_n + \frac{\varepsilon}{2}\Big)\\ &= \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n + \varepsilon\end{aligned}$$

folgern. Insbesondere gilt damit für jedes $\ell \in \mathbb{N}$ mit der Wahl $\varepsilon = 1/\ell$

$$\alpha \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n + \frac{1}{\ell}.$$

Die Ausdrücke auf beiden Seiten dieser Ungleichung konvergieren für $\ell \to \infty$, also liefert uns Satz 7.14 (a)

$$\begin{aligned}\limsup_{n\to\infty}(a_n + b_n) = \alpha = \lim_{\ell\to\infty} \alpha &\leq \lim_{\ell\to\infty}\Big(\limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n + \frac{1}{\ell}\Big)\\ &= \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.\end{aligned}$$

□

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

Beispiel 7.44. 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 7.45. Es seien $(a_n)_n$ eine beschränkte und $(b_n)_n$ eine konvergente Folge in $\mathbb{R}$ mit $\lim_{n\to\infty} b_n \geq 0$. Zeigen Sie

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

Gilt für zwei beschränkte Folgen $(a_n)_n$ und $(b_n)_n$ in $\mathbb{R}$ auch

$$\limsup_{n\to\infty}(a_n \cdot b_n) = \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n?$$

7.7. Cauchy-Folgen

Als weitere Anwendung des Satzes von Bolzano-Weierstraß können wir nun das Cauchy-Kriterium beweisen. Sein entscheidender Vorteil ist, dass es die Konvergenz einer reellen Folge nur anhand der Abstände der Folgenglieder untereinander, d. h. insbesondere auch ohne Kenntnis des Grenzwerts, charakterisiert.
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 $(a_n)_n$ eine konvergente Folge in $\mathbb{R}$ mit Grenzwert a. Dann gibt es für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, sodass $|a_n - a| \leq \varepsilon/2$ für alle $n \geq n_0$ gilt. Wählen wir nun einen weiteren beliebigen Index $m \geq n_0$, so gilt mit der Dreiecksungleichung

$$|a_n - a_m| = |a_n - a + a - a_m| \leq |a_n - a| + |a_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{7.8}$$

Diese wichtige Eigenschaft gießen wir nun in eine Definition.

Definition 7.46. *Eine reelle Folge $(a_n)_n$ heißt* Cauchy-Folge, *wenn für jedes $\varepsilon > 0$ ein Index $n_0 \in \mathbb{N}$ existiert, sodass gilt*

$$|a_n - a_m| < \varepsilon \quad \textit{für alle } n, m \geq n_0. \tag{7.9}$$

Satz 7.47 (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 $(a_n)_n$ eine Cauchy-Folge. Wir überlegen uns zunächst, dass diese Folge beschränkt sein muss. Nach der Definition einer Cauchy-Folge gibt es ein $n_0 \in \mathbb{N}$, sodass $|a_n - a_m| < 1$ 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 $n \geq n_0$. Also 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, d. h. es gilt

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

womit die gesamte Folge $(a_n)_n$ beschränkt ist. Nach dem Satz von Bolzano-Weierstraß 7.35 hat die Folge $(a_n)_n$ einen Häufungswert α. Wir zeigen, dass dieser sogar der Grenzwert der Folge ist, und geben dafür ein $\varepsilon > 0$ vor. Da $(a_n)_n$ eine Cauchy-Folge ist, gibt es ein $n_0 \in \mathbb{N}$ mit $|a_n - a_m| < \varepsilon/2$ für alle $n, m \geq n_0$. Außerdem liegen in $U_{\varepsilon/2}(\alpha)$ unendlich viele Folgenglieder von $(a_n)_n$, also gibt es ein $n_* \geq n_0$ mit $|a_{n_*} - \alpha| < \varepsilon/2$. Für alle $n \geq n_0$ gilt damit

$$|a_n - \alpha| \leq |a_n - a_{n_*}| + |a_{n_*} - \alpha| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

und das bedeutet, dass $(a_n)_n$ gegen α konvergiert. □

Bemerkung 7.48. (a) Man kann das Cauchy-Kriterium auch folgendermaßen umformulieren: Eine reelle Folge $(a_n)_n$ ist eine Cauchy-Folge, genau dann wenn

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

(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 $\mathbb{Q}$, die gegen eine irrationale Zahl strebt, klar.

Ein erstes Anwendungsbeispiel für das Cauchy-Kriterium bietet der folgende Satz.

Satz 7.49 (Prinzip der Intervallschachtelung)**.** *Seien I_n, $n \in \mathbb{N}$, abgeschlossene und nicht-leere Intervalle in $\mathbb{R}$ mit $I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$, sodass die Länge von I_n für n gegen unendlich gegen null strebt. Dann ist $\bigcap_{n\in\mathbb{N}} I_n = \{x_*\}$ für ein $x_* \in \mathbb{R}$.*

Beweis. Für jedes $n \in \mathbb{N}$ bezeichnen wir den linken Randpunkt des Intervalls I_n mit a_n und den rechten mit b_n, sodass $I_n = [a_n, b_n]$ ist. Dann gilt nach Voraussetzung $\lim_{n\to\infty}(b_n - a_n) = 0$.

Zunächst überlegen wir uns, dass der Schnitt aller Intervalle höchstens einen Punkt enthalten kann. Sind $x, y \in \bigcap_{n\in\mathbb{N}} I_n$, so sind beide in jedem Intervall I_n, also gilt für alle $n \in \mathbb{N}$

$$|x - y| \leq b_n - a_n \longrightarrow 0 \quad (n \to \infty).$$

Damit ist $|x - y| = 0$, d. h. es gilt $x = y$.
Es bleibt noch auszuschließen, dass der betrachtete Schnitt leer ist. Dazu zeigen wir als Erstes, dass $(a_n)_n$ eine Cauchy-Folge ist.
Sei $\varepsilon > 0$. Da die Länge der Intervalle I_n gegen null konvergiert, gibt es ein $n_0 \in \mathbb{N}$ mit $b_{n_0} - a_{n_0} < \varepsilon$. Wählen wir nun $n, m \geq n_0$, so gilt nach Voraussetzung $a_n \in I_n \subseteq I_{n_0}$ und $a_m \in I_m \subseteq I_{n_0}$. Der Abstand von a_n zu a_m kann also nicht größer werden als die Länge des Intervalls I_{n_0} und das bedeutet

$$|a_n - a_m| \leq b_{n_0} - a_{n_0} < \varepsilon.$$

Folglich ist $(a_n)_n$ eine Cauchy-Folge und damit nach Satz 7.47 konvergent. Wir nennen den Grenzwert a und zeigen, dass $a \in \bigcap_{k\in\mathbb{N}} I_k$ gilt.
Sei $k \in \mathbb{N}$ beliebig. Dann gilt für alle $n \geq k$ nach Voraussetzung $I_n \subseteq I_k$ und deshalb $a_k \leq a_n \leq b_n \leq b_k$. Dank der Konvergenz von $(a_n)_n$ und der Monotonie des Grenzwerts folgt $a_k \leq \lim_{n\to\infty} a_n \leq b_k$, also gilt $a \in I_k$. $\square$

8. Reihen

Will man unendlich viele Zahlen addieren, bekommt man es wieder mit einem Grenzwert zu tun. Das führt auf den Begriff einer *Reihe*, mit dem wir uns in diesem Abschnitt beschäftigen.

Definition 8.1. *Es sei $(a_n)_n$ eine Folge in $\mathbb{R}$.*

*(a) Wir setzen $s_n := \sum_{k=1}^{n} a_k$, $n \in \mathbb{N}$, und nennen s_n die n-*te Partialsumme*. Dann heißt die Folge $(s_n)_n$ der Partialsummen* (unendliche) Reihe*.*

(b) Dass die Reihe $\left(\sum_{k=1}^{n} a_k\right)_n$ konvergiert *bzw.* bestimmt divergiert*, bedeutet also einfach, dass sie als Folge der Partialsummen konvergiert bzw. bestimmt divergiert. Wir schreiben dann auch*

$$\sum_{k=1}^{\infty} a_k := \lim_{n\to\infty} \sum_{k=1}^{n} a_k = \lim_{n\to\infty} s_n$$

für den Grenzwert oder Reihenwert*.*

Diskussion:

(a) Eine Reihe ist eine Folge mit einer speziellen Darstellung. Andererseits kann man auch auch jede reelle Folge als Reihe schreiben (vgl. Übungsblatt 5). Insofern setzen wir in diesem Kapitel nur die Diskussion des Konzepts „Folge“ fort. Wir nutzen im Folgenden immer wieder aus, dass wir zwischen den Perspektiven „Folge“ und „Reihe“ beliebig wählen können.

(b) Sie kennen Reihen vielleicht schon aus der Definition des (Riemann-)Integrals über Ober- und Untersummen. Dort treten sie typischerweise auf und dort werden sie uns auch wieder begegnen. Später sehen wir außerdem über den Begriff der Potenzreihe und mit Hilfe des Satzes von Taylor, sowie bei den trigonometrischen Funktionen, dass sich viele reelle Funktionen gut durch Reihen darstellen und analysieren lassen.

(c) In vielen Lehrbüchern und anderer mathematischer Literatur wird die Bezeichnung $\sum_{n=1}^{\infty} a_n$ oft gleichzeitig für den Grenzwert und für die Reihe selbst verwendet. Das passiert selbst dann, wenn die Reihe gar nicht konvergiert oder bestimmt divergiert. Diese potentiell verwirrende Notation wollen wir hier vermeiden. Sie sollen aber vorgewarnt sein, dass sie im Allgemeinen oft vorkommt.

Bei der Behandlung von Reihen ist es unpraktisch, dass wir durch unsere Definition einer Folge darauf festgelegt sind, die Summation immer mit eins als Index zu beginnen. Insbesondere ist der Start bei Null oft nützlich. Deshalb erweitern wir unseren Folgenbegriff ein wenig.

Definition 8.2. *Es sei $p \in \mathbb{Z}$ und $X \neq \emptyset$ eine Menge. Eine Abbildung*

$$a : \{p + n : n \in \mathbb{N}_0\} \longrightarrow X$$

nennen wir in Erweiterung von Definition 7.1 ebenfalls Folge in X *und bezeichnen diese mit $(a_n)_{n=p}^{\infty}$. Ist $p = 1$, so schreiben wir weiterhin $(a_n)_n$ für $(a_n)_{n=1}^{\infty}$.*

Bemerkung 8.3. (a) Für eine Folge $(a_n)_{n=p}^{\infty}$ können wir durch $b_n := a_{n+p-1}$, $n \in \mathbb{N}$, eine Folge $(b_n)_n$ konstruieren, die unserer alten Definition entspricht. Alle Erkenntnisse aus den letzten Kapiteln lassen sich damit problemlos auf diesen verallgemeinerten Folgenbegriff übertragen (vgl. auch Übungsblatt 5).

(b) Ist $p \in \mathbb{Z}$ und $(a_n)_{n=p}^{\infty}$ eine Folge, so definiert man für die Reihe über diese Folge entsprechend die Folge der Partialsummen als $s_n := \sum_{k=p}^{n} a_k$ für alle $n \geq p$ und schreibt für die Folge $(s_n)_{n=p}^{\infty}$ wieder $\left(\sum_{k=p}^{n} a_k\right)_n$ und für den Reihenwert $\sum_{k=p}^{\infty} a_k$, wenn er exisitert.

Auch im Folgenden werden 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 8.4. Wir betrachten einige wichtige Reihen, untersuchen diese auf Konvergenz und bestimmen die Reihenwerte.

(a) Wir haben in Kapitel 7.4 bereits eine Reihe kennengelernt, nämlich $(s_n)_n$ mit

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

In Satz 7.25 haben wir gesehen, dass $\lim_{n\to\infty} s_n = \mathrm{e}$ ist. Also ist dieses nach Definition eine konvergente Reihe mit Reihenwert

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

(b) Die *geometrische Reihe* ist für $x \in \mathbb{R}$ gegeben durch

$$\left(\sum_{k=0}^{n} x^k\right)_n \qquad \text{(geometrische Reihe).}$$

In Satz 5.12 haben wir schon gesehen, dass für die Partialsummen dieser Reihe

$$s_n = \sum_{k=0}^{n} x^k = \begin{cases} \dfrac{1 - x^{n+1}}{1 - x}, & \text{falls } x \neq 1, \\ n + 1, & \text{falls } x = 1 \end{cases}$$

gilt.

Für $x = 1$ konvergiert diese Folge offensichtlich nicht und nach Satz 7.18 konvergiert die Folge $(x^{n+1})_n$ genau dann, wenn $x \in (-1, 1]$ ist und dann ist der Grenzwert 0, falls $x \neq 1$ ist. Fazit: Die geometrische Reihe konvergiert, genau dann wenn $|x| < 1$ ist und in diesem Fall gilt

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

(c) Um die Reihe

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

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

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

Also bekommen wir für die zugehörigen Partialsummen für alle $n \geq 2$

$$\begin{aligned} s_n &= \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \Big(\frac{1}{k} - \frac{1}{k+1} \Big) = 1 + \sum_{k=2}^{n} \frac{1}{k} - \sum_{k=1}^{n-1} \frac{1}{k+1} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1} \end{aligned}$$

dank einer freundlichen Teleskopsumme. Für $n \to \infty$ folgt $s_n \to 1$ und damit haben wir auch hier eine konvergente Reihe mit Reihenwert 1, in Formeln

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

(d) Abschließend betrachten wir die *harmonische Reihe*

$$\left(\sum_{k=1}^{n} \frac{1}{k} \right)_n \qquad \text{(harmonische Reihe).}$$

Für diese Reihe gilt

$$s_{2n} = \sum_{k=1}^{2n} \frac{1}{k} = s_n + \sum_{k=n+1}^{2n} \underbrace{\frac{1}{k}}_{\geq 1/2n} \geq s_n + \sum_{k=n+1}^{2n} \frac{1}{2n} = s_n + n \cdot \frac{1}{2n} = s_n + \frac{1}{2}$$

für alle $n \in \mathbb{N}$. Nehmen wir an, die Folge $(s_n)_n$, und damit die harmonische Reihe, wäre konvergent, so existiert $s := \lim_{n\to\infty} s_n$. Da dann aber nach Satz 7.33 (b) auch die Teilfolge $(s_{2n})_n$ gegen s konvergiert, liefert uns obige Ungleichung den Widerspruch $s \geq s + 1/2$. Also ist die harmonische Reihe divergent. Man kann aber bestimmte Divergenz gegen ∞ zeigen (Übungsaufgabe) und erhält in diesem Sinne

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

8.1. Konvergenzkriterien für Reihen

Wir beweisen nun die ersten Aussagen über Konvergenz von Reihen. Dazu betrachten wir zuerst nur einige unserer Konvergenzkriterien für Folgen im Fall der Darstellung als Reihe.

Satz 8.5. *Es sei $(a_n)_n$ eine reelle Folge und $s_n := \sum_{k=1}^n a_k$, $n \in \mathbb{N}$. Dann gilt*

(a) Ist $a_n \geq 0$ für alle $n \in \mathbb{N}$ und die Folge $(s_n)_n$ nach oben beschränkt, dann ist $(s_n)_n$ konvergent. (Monotonie-Kriterium)

(b) Die Reihe $(s_n)_n$ ist genau dann konvergent, wenn für jedes $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert, sodass

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

gilt. (Cauchy-Kriterium)

Beweis. (a) Da alle a_n positiv 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 $(s_n)_n$ ist also monoton wachsend. Da sie nach Voraussetzung auch beschränkt ist, folgt die Behauptung direkt aus dem Monotonie-Kriterium für Folgen, Satz 7.22.

(b) auf Übungsblatt 5. □

Auch die Linearität des Grenzwerts überträgt sich auf Reihen.

Satz 8.6. *Es seien $\left(\sum_{k=1}^n a_k\right)_n$ und $\left(\sum_{k=1}^n b_k\right)_n$ zwei konvergente Reihen in $\mathbb{R}$ sowie $\alpha, \beta \in \mathbb{R}$. Dann konvergiert auch die Reihe $\left(\sum_{k=1}^n (\alpha a_k + \beta b_k)\right)_n$ 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 $n \in \mathbb{N}$

$$r_n = \sum_{k=1}^n (\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^n a_k + \beta \sum_{k=1}^n b_k = \alpha s_n + \beta t_n. \tag{8.1}$$

Also ist

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

nach den Grenzwertsätzen für Folgen in Satz 7.12. □

Außerdem gelten für konvergente Reihen die folgenden Aussagen.

Satz 8.7. *Es sei* $\left(\sum_{k=1}^n a_k\right)_n$ *eine konvergente Reihe in* $\mathbb{R}$*. Dann gilt:*

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

(b) Die Folge gegeben durch $r_\nu := \sum_{k=\nu}^\infty a_k$*,* $\nu \in \mathbb{N}$*, ist eine Nullfolge.*

(c) Die Folge $(a_n)_n$ *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 $m \geq \nu$ betrachten wir die Partialsummen $\sigma_m = \sum_{k=\nu}^m a_k$. 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 $(\sigma_m)_m$ 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 $(r_\nu)_\nu$ eine Nullfolge.

(c) Hierzu beobachten wir, dass $a_n = s_n - s_{n-1}$ für jedes $n \in \mathbb{N}$ gilt. Daraus folgt $\lim_{n\to\infty} a_n = \lim_{n\to\infty}(s_n - s_{n-1}) = s - s = 0$. □

Warnung 8.8. In (c) im obigen Satz gilt die Umkehrung *nicht*, wie das Beispiel der harmonischen Reihe, vgl. Beispiel 8.4 (d), zeigt.

8.2. Absolute Konvergenz und der Riemannsche Umordnungssatz

Im Folgenden betrachten wir insbesondere Reihen, deren Summanden teils positives und teils negatives Vorzeichen haben. Diese haben es leichter zu konvergieren als Reihen, deren Summanden ein konstantes Vorzeichen haben, da sich Beiträge der positiven und der negativen Summanden gegenseitig wegheben können. Dabei können allerdings Überraschungen passieren. Diese guten und schlechten Eigenschaften solcher Reihen wollen wir in diesem Abschnitt beleuchten.
Als Modellfall betrachten wir zuerst sogenannte *alternierende Reihen*, bei denen die Summanden abwechselnd positiv und negativ sind.

Beispiel 8.9. Durch Einbauen von alternierenden Vorzeichen in die harmonischen Reihe, s. Beispiel 8.4 (d), ergibt sich die *alternierende harmonische Reihe* mit Werten,

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

Abbildung 8.1 zeigt die ersten Partialsummen dieser Reihe. Das sieht sehr konvergent aus und wir wollen im Folgenden zeigen, dass die alternierende harmonische Reihe im Gegensatz zur harmonischen Reihe tatsächlich konvergent ist.

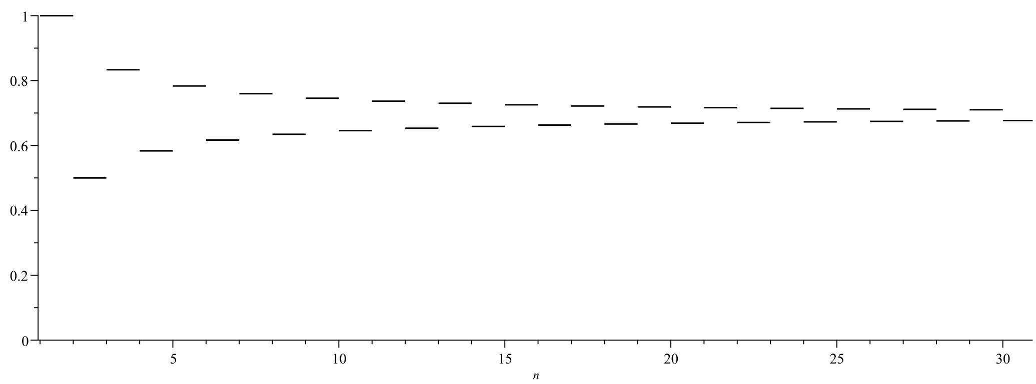

Abbildung 8.1.: Die ersten dreißig Partialsummen der alternierenden harmonischen Reihe.

Wir setzen dazu $a_n := {}^{(-1)^{n+1}}/_{n}$ für jedes $n \in \mathbb{N}$.
Betrachten wir 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 $(s_{2n})_n$ von $(s_n)_n$ ist also monoton wachsend. Analog erkennen wir, dass die Teilfolge $(s_{2n-1})_n$ der ungeraden Folgenglieder monoton fällt. Außerdem gilt für jedes $n \in \mathbb{N}$

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

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

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

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 die Reihe konvergent ist.

Schaut man sich die Argumentation in obigem Beispiel an, so stellt man fest, dass die konkrete Formel für die Folgenglieder a_n gar nicht verwendet wurde, sodass diese Vorgehensweise ohne Mühe auf viele andere alternierende Reihen übertragen werden kann. In diesem Sinne können Sie den Beweis des folgenden Konvergenzkriteriums für Reihen mit alternierenden Vorzeichen selbst führen.

Satz 8.10 (Leibniz-Kriterium)**.** *Es sei $(a_n)_n$ eine monotone Folge und es gelte $\lim_{n\to\infty} a_n = 0$. Dann ist die Reihe $\left(\sum_{k=1}^n (-1)^k a_k\right)_n$ konvergent.*

Als Nächstes wollen wir untersuchen, ob und wie sich das Umordnen von Folgengliedern auf die Konvergenz einer Folge und der zugehörigen Reihe auswirkt. Wir definieren zunächst, was mit Umordnen gemeint ist.

Definition 8.11. *Es sei $(a_n)_n$ eine reelle Folge und $\varphi : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Wir setzen $b_n := a_{\varphi(n)}$ für jedes $n \in \mathbb{N}$. Dann heißt die Folge $(b_n)_n$ eine* Umordnung *von $(a_n)_n$ und die Reihe $(\sum_{k=1}^n b_k)_n$ eine* Umordnung *von $(\sum_{k=1}^n a_k)_n$.*

Bemerkung 8.12. Wegen der Bijektivität der Abbildung φ existiert die Umkehrabbildung φ^{-1} mit $\varphi(\varphi^{-1}(n)) = n$ für jedes $n \in \mathbb{N}$ und sie ist ebenfalls bijektiv. Damit lässt sich jede Umordnung rückgängig machen. Außerdem ist $b_{\varphi^{-1}(n)} = a_{\varphi(\varphi^{-1}(n))} = a_n$, d. h. wann immer $(b_n)_n$ eine Umordnung von $(a_n)_n$ ist, ist auch $(a_n)_n$ eine Umordnung von $(b_n)_n$.

An der Konvergenz von Folgen ändert Umsortieren nichts, wie der folgende Satz zeigt.

Satz 8.13. *Es sei $(a_n)_n$ eine konvergente reelle Folge und $(b_n)_n$ eine Umordnung von $(a_n)_n$. Dann ist auch $(b_n)_n$ konvergent und es gilt $\lim_{n\to\infty} b_n = \lim_{n\to\infty} a_n$.*

Beweis. Es sei $\varphi : \mathbb{N} \to \mathbb{N}$ die zur Umordnung gehörige bijektive Abbildung, d. h. es gilt $b_n = a_{\varphi(n)}$ für alle $n \in \mathbb{N}$. Wir setzen $a := \lim_{n\to\infty} a_n$ und geben ein $\varepsilon > 0$ vor. Dann gibt es ein $n_* \in \mathbb{N}$, sodass $|a_n - a| < \varepsilon$ für alle $n \geq n_*$ gilt.
Nun muss es ein $n_0 \in \mathbb{N}$ geben mit $\varphi(n) \geq n_*$ für alle $n \geq n_0$, denn wäre das nicht der Fall, so gäbe es unendlich viele $n \in \mathbb{N}$ mit $\varphi(n) \in \{1, 2, \dots, n_* - 1\}$ und das stünde im Widerspruch zur Bijektivität von φ. Nehmen wir uns ein solches n_0 her, so haben wir für alle $n \geq n_0$, dass $\varphi(n) \geq n_*$ ist, und damit ist für alle $n \geq n_0$

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

Also konvergiert $(b_n)_n$ ebenfalls gegen a. □

Bei Reihen ist die Situation komplizierter. Wir geben Reihen, deren Konvergenz von der Reihenfolge der Summanden unabhängig ist, eine eigene Bezeichnung.

Definition 8.14. *Eine konvergente Reihe heißt* unbedingt konvergent*, wenn jede ihrer Umordnungen auch konvergent ist und alle Umordnungen denselben Reihenwert besitzen.*
Ist eine konvergente Reihe nicht unbedingt konvergent, so heißt sie bedingt konvergent*.*

Wir werden gleich sehen, dass es bedingt konvergente Reihen gibt. Zuerst definieren und untersuchen wir noch einen dazu ganz verwandten Begriff.

Definition 8.15. *Eine Reihe $\left(\sum_{k=1}^n a_k\right)_n$ heißt* absolut konvergent*, wenn die Reihe $\left(\sum_{k=1}^n |a_k|\right)_n$ konvergiert.*

Es ist wenig überraschend, dass absolute Konvergenz ist ein stärkerer Begriff ist als Konvergenz:

Satz 8.16. *Es sei $\left(\sum_{k=1}^n a_k\right)_n$ eine absolut konvergente Reihe. Dann ist sie konvergent und es gilt die* verallgemeinerte Dreiecksungleichung

$$\Big|\sum_{k=1}^\infty a_k\Big| \leq \sum_{k=1}^\infty |a_k|. \tag{8.2}$$

Beweis. Wir verwenden das Cauchy-Kriterium aus Satz 8.5. Sei dazu $\varepsilon > 0$ gegeben. Dann gibt es dank der absoluten Konvergenz ein $n_0 \in \mathbb{N}$, sodass für alle

$n, m \in \mathbb{N}$ mit $m > n > n_0$ gilt $\sum_{k=n+1}^{m} |a_k| < \varepsilon$. Mit der Dreiecksungleichung gilt dann sofort für alle $m > n > n_0$ auch

$$\Big| \sum_{k=n+1}^{m} a_k \Big| \leq \sum_{k=n+1}^{m} |a_k| < \varepsilon.$$

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 $n \to \infty$

$$\Big|\sum_{k=1}^{\infty} a_k\Big| = \big|\lim_{n\to\infty} s_n\big| = \lim_{n\to\infty} |s_n| \leq \lim_{n\to\infty} \sigma_n = \sum_{k=1}^{\infty} |a_k|. \qquad \square$$

Bemerkung 8.17. (a) Jede absolut konvergente Reihe in $\mathbb{R}$ ist konvergent, aber die Umkehrung gilt nicht, wie das Beispiel der alternierenden harmonischen Reihe zeigt.

(b) Ist $(a_n)_n$ eine Folge mit $a_n \geq 0$ für alle $n \in \mathbb{N}$ und ist die Reihe über $(a_n)_n$ konvergent, so ist sie auch absolut konvergent.

Der nächste Satz zeigt uns, dass absolut konvergente Reihen auch unbedingt konvergent sind. Ihr Konvergenzverhalten ist also gegen Umordnen immun.

Satz 8.18. *Eine absolut konvergente Reihe ist auch unbedingt konvergent.*

Beweis. Sei $\left(\sum_{k=1}^{n} a_k\right)_n$ eine absolut konvergente Reihe und $(b_n)_n$ eine Umordnung von $(a_n)_n$. Wir behandeln zunächst den Spezialfall, dass $a_n \geq 0$ für alle $n \in \mathbb{N}$ gilt und betrachten die Partialsummen $\sigma_n := \sum_{k=1}^{n} b_k$, $n \in \mathbb{N}$. Ist $j \in \mathbb{N}$ vorgegeben, so können wir uns dazu wie im Beweis von Satz 8.13 ein $N \in \mathbb{N}$ verschaffen, für das $\{b_1, b_2, \dots, b_j\} \subseteq \{a_1, a_2, \dots, a_N\}$ gilt. Da alle Folgenglieder von $(a_n)_n$ positiv sind, folgern wir daraus, dass $\sigma_j \leq \sum_{k=1}^{N} a_k$ ist. Wir machen diesen Ausdruck noch größer, indem wir weitere positive Summanden dazu addieren, und erhalten für alle $j \in \mathbb{N}$

$$\sigma_j \leq \sum_{k=1}^{N} a_k \leq \sum_{k=1}^{\infty} a_k =: s. \tag{8.3}$$

Damit ist die Partialsummen-Folge $(\sigma_n)_n$ beschränkt und wegen der Positivität der Summanden konvergiert die Reihe $\left(\sum_{k=1}^{n} b_k\right)_n$ nach dem Monotonie-Kriterium, vgl. Satz 8.5 (a). Somit gilt dank (8.3)

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

Nun ist aber wegen Bemerkung 8.12 auch $\left(\sum_{k=1}^n a_k\right)_n$ eine Umordnung der Reihe über $(b_n)_n$. Man kann damit obigen Beweis mit vertauschten Rollen von $(a_n)_n$ und $(b_n)_n$ noch einmal führen und erhält dann die umgekehrte Ungleichung $s \leq \sum_{n=1}^\infty b_n$. Also sind die beiden Reihenwerte gleich.
Es bleibt noch der Fall einer Folge $(a_n)_n$ zu betrachten, deren Folgenglieder beliebige Vorzeichen haben. Da nach Voraussetzung $\left(\sum_{k=1}^n |a_k|\right)_n$ konvergent ist und alle Summanden dieser Reihe positiv sind, ist nach dem ersten Teil dieses Beweises die Reihe $\left(\sum_{k=1}^n |b_k|\right)_n$ konvergent und es gilt $\sum_{n=1}^\infty |b_n| = \sum_{n=1}^\infty |a_n|$. Also ist die Reihe über $(b_n)_n$ absolut konvergent.
Es bleibt noch die Gleichheit der Reihenwerte für die Reihen ohne Betragsstriche zu zeigen. Dazu setzen wir

$$\tilde{a}_n := a_n + |a_n|, \qquad \tilde{b}_n := b_n + |b_n|, \qquad n \in \mathbb{N}.$$

Dann gilt $\tilde{a}_n \geq 0$ und $\tilde{b}_n \geq 0$ für alle $n \in \mathbb{N}$ und $(\tilde{b}_n)_n$ ist eine Umordnung von $(\tilde{a}_n)_n$ (mit der gleichen Funktion φ). Außerdem ist wegen der absoluten Konvergenz von $\left(\sum_{k=1}^n a_k\right)_n$ die Reihe $\left(\sum_{k=1}^n \tilde{a}_k\right)_n$ konvergent, ja sogar absolut konvergent, denn alle Folgenglieder von $(\tilde{a}_n)_n$ sind positiv. Nun können wir für die Reihe $\left(\sum_{k=1}^n \tilde{b}_k\right)_n$ wieder den schon bewiesenen Fall von oben verwenden und wissen hiermit, dass auch diese Reihe absolut konvergiert und $\sum_{n=1}^\infty \tilde{a}_n = \sum_{n=1}^\infty \tilde{b}_n$ gilt. Damit ergibt sich schließlich mit Satz 8.6,

$$\begin{aligned}\sum_{n=1}^\infty b_n &= \sum_{n=1}^\infty (\tilde{b}_n - |b_n|) = \sum_{n=1}^\infty \tilde{b}_n - \sum_{n=1}^\infty |b_n| = \sum_{n=1}^\infty \tilde{a}_n - \sum_{n=1}^\infty |a_n| \\ &= \sum_{n=1}^\infty (\tilde{a}_n - |a_n|) = \sum_{n=1}^\infty a_n.\end{aligned}$$

□

Wir wollen jetzt zuerst zeigen, dass es Reihen gibt, die nur bedingt konvergent sind, und dann, dass das *genau* diejenigen sind, die nicht absolut konvergieren. Dazu benötigen wir eine Vorüberlegung.

Lemma 8.19. *Es sei $\left(\sum_{k=1}^n a_k\right)_n$ eine konvergente Reihe, die nicht absolut konvergent ist. Setzen wir $a_{n,+} := \max\{a_n, 0\}$ und $a_{n,-} := \min\{a_n, 0\}$ für jedes $n \in \mathbb{N}$, so divergiert $\left(\sum_{k=1}^n a_{k,+}\right)_n$ bestimmt gegen ∞ und $\left(\sum_{k=1}^n a_{k,-}\right)_n$ bestimmt gegen $-\infty$.*

Beweis. Wir beobachten zuerst, dass für alle $n \in \mathbb{N}$ gilt

$$a_n - a_{n,+} = \left\{ \begin{array}{rl} 0, & \text{falls } a_n \geq 0, \\ a_n, & \text{falls } a_n < 0, \end{array} \right\} = a_{n,-} \tag{8.4}$$

und

$$a_{n,+} - a_{n,-} = \left\{ \begin{array}{rl} a_n, & \text{falls } a_n \geq 0, \\ -a_n & \text{falls } a_n < 0, \end{array} \right\} = |a_n|. \tag{8.5}$$

Wir nehmen an, dass die Reihe $\left(\sum_{k=1}^n a_{k,+}\right)_n$ nicht bestimmt gegen ∞ divergiert. Da alle Summanden positiv sind, ist sie monoton wachsend. Die Annahme sagt uns dann, dass diese beschränkt ist, sodass das Monotoniekriterium für Reihen aus Satz 8.5 (a) die Konvergenz dieser Reihe liefert.
Nach Satz 8.6 ist dann wegen der Konvergenz von $\left(\sum_{k=1}^n a_k\right)_n$ und dank (8.4) auch $\left(\sum_{k=1}^n a_{k,-}\right)_n$ konvergent mit

$$\sum_{k=1}^{\infty} a_{k,-} = \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} a_{k,+}. \tag{8.6}$$

Nochmals Satz 8.6 zusammen mit (8.5) liefert die Konvergenz von $\left(\sum_{k=1}^n |a_k|\right)_n$ mit

$$\sum_{k=1}^{\infty} a_{k,+} - \sum_{k=1}^{\infty} a_{k,-} = \sum_{k=1}^{\infty} |a_k|, \tag{8.7}$$

aber das war in der Voraussetzung ja gerade ausgeschlossen. Also divergiert $\left(\sum_{k=1}^n a_{k,+}\right)_n$ bestimmt gegen ∞.
Nimmt man an, dass die Reihe $\left(\sum_{k=1}^n a_{k,-}\right)_n$ nicht bestimmt gegen $-\infty$ divergiert, kommt man analog zu einem Widerspruch. □

Der folgende Satz liefert jetzt das verblüffende Verhalten von nicht absolut konvergenten Reihen unter Umordnung.

Satz 8.20 (Riemannscher Umordnungssatz)**.** *Es sei* $\left(\sum_{k=1}^n a_k\right)_n$ *eine konvergente Reihe, die nicht absolut konvergiert. Dann gibt es für jedes* $S \in \mathbb{R} \cup \{\infty, -\infty\}$ *eine Umordnung* $(b_n)_n$ *von* $(a_n)_n$ *mit* $\sum_{n=1}^{\infty} b_n = S$.

Bemerkung 8.21. (a) Der Satz besagt nicht nur, dass man beim betrachteten Typ von Reihen durch Umordnen den Reihenwert ändern kann, sondern dass man allein durch Umordnen *jeden beliebigen* Reihenwert ansteuern kann, ja sogar eine divergente Reihe erzeugen kann. Außerdem bedeutet dieser Satz, dass es bedingt konvergente Reihen gibt, z. B. die alternierende harmonische Reihe aus Beispiel 8.9.

(b) Der Satz kommt mit der Beweisidee:

Wir wollen die konvergente Reihe $\left(\sum_{k=1}^n a_k\right)_n$ so umordnen, dass als Grenzwert S herauskommt und konzentrieren uns auf den Fall $S \in [0, \infty)$. Wegen Lemma 8.19 muss unsere Reihe sowohl unendlich viele positive als auch unendlich viele strikt negative Summanden haben. Dementsprechend können wir die Teilfolge $(a_{n_k})_{k \in \mathbb{N}}$ aller positiven Summanden und die Teilfolge

$(a_{m_\ell})_{\ell\in\mathbb{N}}$ der strikt negativen Summanden betrachten. Nach Lemma 8.19 divergieren die zugehörigen Reihen bestimmt gegen ∞ bzw. $-\infty$, mit

$$\sum_{k=1}^{\infty} a_{n_k} = \sum_{n=1}^{\infty} a_{n,+} \qquad \text{und} \qquad \sum_{\ell=1}^{\infty} a_{m_\ell} = \sum_{n=1}^{\infty} a_{n,-}.$$

Insbesondere können wir also jede beliebige positive Zahl durch Addition endlich vieler positiver Summanden übertreffen, sowie jede beliebige negative Zahl durch Addition endlich vieler negativer Summanden unterbieten.

Das machen wir uns zu Nutze, um genau den Grenzwert S zu erreichen. Man addiert zunächst die ersten positiven Summanden, solange bis S zum ersten Mal überboten ist. Dann addiert man die ersten negativen Summanden so lange, bis man zum ersten Mal wieder unterhalb von S ist. Wir haben schon endlich viele positive Summanden verbraucht, aber auch die verbliebene Restreihe der noch nicht verwendeten positiven Summanden muss noch bestimmt gegen ∞ divergieren, denn sonst wäre auch die gesamte Reihe $\left(\sum_{k=1}^{n} a_{k,+}\right)_n$ konvergent gewesen. Also haben wir noch beliebig viel positives „Baumaterial" übrig und können wieder solange davon addieren, bis wir zum ersten Mal wieder über S liegen. Auch das ist nach endlich vielen Summanden der Fall.

So wechselt man immer Blöcke von positiven und negativen Summanden ab und pendelt sich dabei immer näher auf S ein. Letzteres liegt daran, dass nach Satz 8.7 (c) die Folge $(a_n)_n$ eine Nullfolge sein muss. Da im Verlauf unserer Umordnerei die Summanden also betragsmäßig immer kleiner werden, wird der Wert, um den wir S in jedem Block über- bzw. unterbieten, immer näher an null rutschen.

Will man die Reihe so umordnen, dass sie bestimmt gegen ∞ divergiert, so behält man die Grundkonstruktion bei. Man summiert aber in jedem Schritt nicht, bis man über bzw. unter S ist, sondern man verschiebt diese Marke im Verlauf der Zeit nach oben. Beispielsweise kann man addieren, bis man über zwei ist, dann abziehen bis unter eins, addieren bis über drei, abziehen bis unter zwei, addieren bis über vier, usw.

Beispiel 8.22. Wir betrachten die Umordnung der alternierenden harmonischen Reihe, die man bekommt, wenn man immer einen positiven und zwei negative Summanden abwechseln lässt, also

$$1 - \frac{1}{2} - \frac{1}{4} + \quad \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \quad \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \dots.$$

Das Schöne an dieser speziellen Umordnung ist, dass man „nachrechnen" kann, wie der Reihenwert dadurch modifiziert wird. Jeder Dreierblock ist von der Form

$$\frac{1}{2k-1} - \frac{1}{2(2k-1)} - \frac{1}{2(2k-1)+2} = \frac{1}{2}\left(\frac{1}{2k-1} - \frac{1}{2k}\right) \tag{8.8}$$

für $k \in \mathbb{N}$. Damit bekommen wir für diese Umordnung

$$\sum_{k=1}^{\infty}\Big(\frac{1}{2k-1}-\frac{1}{2(2k-1)}-\frac{1}{2(2k-1)+2}\Big)=\sum_{k=1}^{\infty}\frac{1}{2}\Big(\frac{1}{2k-1}-\frac{1}{2k}\Big)$$
$$=\frac{1}{2}\Big(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\dots\Big)=\frac{1}{2}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}.$$

Durch diese spezielle Umordnung wird der Wert der alternierenden harmonischen Reihe damit halbiert.

Zum Schluss noch die abstrakte Zusammenfassung der Ergebnisse dieses Abschnitts:

Korollar 8.23. *Eine Reihe ist genau dann unbedingt konvergent, wenn sie absolut konvergent ist.*

Beweis. Jede konvergente, aber nicht absolut konvergente Reihe ist nach dem Riemannschen Umordnungssatz bedingt konvergent. Die Kontraposition dieser Erkenntnis liefert, dass jede unbedingt konvergente Reihe absolut konvergent ist. Dass jede absolut konvergente Reihe unbedingt konvergiert, war der Inhalt von Satz 8.18. □

8.3. Kriterien für absolute Konvergenz

Die Untersuchung einer Reihe auf Konvergenz kann ein verzwicktes Problem darstellen. In diesem Abschnitt wollen wir weitere Kriterien beweisen, die dabei helfen können.

Satz 8.24. *Es seien $(a_n)_n$ und $(b_n)_n$ zwei reelle Folgen.*

(a) Ist $|a_n| \leq b_n$ für fast alle $n \in \mathbb{N}$ und konvergiert die Reihe $\big(\sum_{k=1}^{n} b_k\big)_n$, so ist die Reihe $\big(\sum_{k=1}^{n} a_k\big)_n$ absolut konvergent. (Majorantenkriterium)

Man nennt $(b_n)_n$ dann eine konvergente Majorante *von $(a_n)_n$.*

(b) Ist $a_n \geq b_n \geq 0$ für fast alle $n \in \mathbb{N}$ und divergiert die Reihe $\big(\sum_{k=1}^{n} b_k\big)_n$, so divergiert auch die Reihe $\big(\sum_{k=1}^{n} a_k\big)_n$. (Minorantenkriterium)

Man nennt $(b_n)_n$ dann eine divergente Minorante *von $(a_n)_n$.*

Beweis. (a) Nach Voraussetzung gibt es ein $k \in \mathbb{N}$, sodass $|a_n| \leq b_n$ für alle $n \geq k$ gilt. Seien nun $n, m \in \mathbb{N}$ mit $m > n > k$ gegeben. Dann gilt, da $(b_n)_n$ nach Voraussetzung positive Folgenglieder hat,

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

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 $\left(\sum_{k=1}^{n} a_k\right)_n$ konvergent ist, so wäre $(a_n)_n$ eine konvergente Majorante von $(b_n)_n$ und somit die Reihe $\left(\sum_{k=1}^{n} b_k\right)_n$ konvergent, was ja nicht sein soll. Also divergiert $\left(\sum_{k=1}^{n} a_k\right)_n$. □

Wir wollen diese sehr nützlichen Kriterien an zwei Beispielen erproben.

Beispiel 8.25. (a) Es sei $\alpha \in \mathbb{Q}$ mit $0 < \alpha \leq 1$. Wir betrachten die Reihe

$$\left(\sum_{k=1}^{n} \frac{1}{k^\alpha}\right)_n .$$

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

Sobald die allgemeine Potenz n^x für alle $x \in \mathbb{R}$ definiert ist, verallgemeinert sich auch dieser Beweis sofort und die Aussage bleibt richtig für alle $\alpha \in (0, 1]$.

(b) Nun stellt sich die Frage, wie das Konvergenzverhalten der Reihe aus (a) für Exponenten α aussieht, die größer als eins sind. Um eine Intuition zu bekommen, untersuchen wir die Reihe

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

der Quadratkehrwerte. Dazu zeigen wir mit Hilfe des Majorantenkriteriums, dass die Reihe

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

konvergiert. Dann konvergiert auch $\left(\sum_{k=1}^{n} \frac{1}{k^2}\right)_n$ (Indexverschiebung). Wir beobachten dazu, dass für jedes $n \in \mathbb{N}$ gilt

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

Wir wissen aber seit Beispiel 8.4 (c), dass die Reihe über $(1/k(k+1))$ konvergiert, also haben wir mit dieser Folge eine konvergente Majorante gefunden.

Ein vollständiges Bild liefert der folgende Satz.

Satz 8.26. *Ist $\alpha \in \mathbb{Q}$ mit $\alpha > 1$, so konvergiert die Reihe*

$$\left(\sum_{k=1}^{n} \frac{1}{k^\alpha}\right)_n .$$

Beweis. Da alle Summanden der Reihe positiv sind, reicht es nach dem Monotoniekriterium aus Satz 8.5 (a) aus, die Beschränktheit der Partialsummen $s_n := \sum_{k=1}^{n} 1/k^\alpha$, $n \in \mathbb{N}$, zu zeigen. Zu gegebenem $n \in \mathbb{N}$ sei dazu $j \in \mathbb{N}$ mit $2^j - 1 \geq n$ gewählt. Damit ist

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

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

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

und sind fertig. □

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 8.27 (Wurzelkriterium)**.** *Es sei $(a_n)_n$ eine reelle Folge und*

$$w_n := \sqrt[n]{|a_n|}, \quad n \in \mathbb{N}.$$

(a) Ist die Folge $(w_n)_n$ beschränkt und $\alpha := \limsup_{n\to\infty} w_n$, so ist die Reihe $\left(\sum_{k=1}^{n} a_k\right)_n$ absolut konvergent, wenn $\alpha < 1$ ist, und divergent, falls $\alpha > 1$ gilt.

(b) Ist die Folge $(w_n)_n$ unbeschränkt, so divergiert die Reihe $\left(\sum_{k=1}^{n} a_k\right)_n$.

Bemerkung 8.28. Beachten Sie, dass dieser Satz im Falle $\alpha = 1$ keine Aussage trifft. Das ist kein Versehen, sondern nicht zu ändern, denn mit $\alpha = 1$ gibt es sowohl Folgen, für welche die Reihe konvergiert, als auch Folgen, für die sie divergiert. Spuckt die Überprüfung dieses Kriteriums also $\alpha = 1$ aus, muss man sich etwas Neues einfallen lassen.

Beweis von Satz 8.27. Wir betrachten zunächst den Fall, dass die Folge $(\sqrt[n]{|a_n|})_n$ beschränkt und $\alpha < 1$ ist. Sei $q \in (\alpha, 1)$ fest gewählt. Dann gilt $\sqrt[n]{|a_n|} \leq q$ für fast alle $n \in \mathbb{N}$ nach Satz 7.42, d. h. wir haben $|a_n| \leq q^n$ für fast alle $n \in \mathbb{N}$. Nun ist aber wegen $0 < q < 1$ die geometrische Reihe $(\sum_{k=1}^n q_k)_n$ konvergent, und damit haben wir eine konvergente Majorante für unsere Reihe gefunden. Das sichert die absolute Konvergenz.
Ist $\alpha > 1$ oder die Folge $(\sqrt[n]{|a_n|})_n$ sogar unbeschränkt, dann gibt es unendlich viele $n \in \mathbb{N}$, für die $\sqrt[n]{|a_n|} \geq 1$ gilt. Das heißt aber, dass für diese n auch $|a_n| \geq 1^n = 1$ gilt. Damit konvergiert die Folge $(a_n)_n$ sicher nicht gegen null, was aber nach Satz 8.7 (c) eine notwendige Voraussetzung für die Konvergenz der Reihe ist. Also divergiert die Reihe $(\sum_{k=1}^n a_k)_n$ in diesen Fällen. □

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 8.29. (a) Für die harmonische Reihe 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 ist nach Beispiel 8.4 (d) divergent.

(b) Für die Reihe $(\sum_{k=1}^n 1/k^2)_n$ 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 8.25 (b).

(c) Wir betrachten die Reihe $\left(\sum_{k=1}^n \frac{3^k}{k^2+1}\right)_n$.

Es gilt $1 = \sqrt[n]{1} \leq \sqrt[n]{n^2+1} \leq \sqrt[n]{2n^2} = \sqrt[n]{2}\,\sqrt[n]{n}^2$. Mit Hilfe von Satz 7.17 und dem Sandwich-Theorem aus Satz 7.14 erhalten wir also $\lim_{n\to\infty} \sqrt[n]{n^2+1} = 1$. Das impliziert mit $a_n = \frac{3^n}{n^2+1}$

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \frac{\sqrt[n]{3^n}}{\sqrt[n]{n^2+1}} = \frac{\lim_{n\to\infty} 3}{\lim_{n\to\infty} \sqrt[n]{n^2+1}} = 3.$$

Also ist $\alpha = 3 > 1$ und damit die Reihe divergent.

(d) Wir untersuchen schließlich die Reihe $(\sum_{k=1}^n a_k)_n$ 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 8.30 (Quotientenkriterium). *Es sei $(a_n)_n$ eine reelle Folge mit $a_n \neq 0$ für alle $n \in \mathbb{N}$ und*

$$q_n := \left|\frac{a_{n+1}}{a_n}\right|, \quad n \in \mathbb{N}.$$

Dann gilt:

(a) *Ist die Folge $(q_n)_n$ beschränkt und gilt*

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

so konvergiert die Reihe $\left(\sum_{k=1}^n a_k\right)_n$ absolut.

(b) *Ist $q_n \geq 1$ für fast alle $n \in \mathbb{N}$, so divergiert die Reihe $\left(\sum_{k=1}^n a_k\right)_n$.*

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. Sei also $n_0 \in \mathbb{N}$ so gewählt, dass für alle $n \geq n_0$ gilt $q_n \leq q$. Für all diese n haben wir dann

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

Also ist $|a_{n+1}| \leq q|a_n|$ für alle $n \geq n_0$ und wir bekommen

$$|a_n| \leq q|a_{n-1}| \leq q^2|a_{n-2}| \leq \cdots \leq q^{n-n_0}|a_{n_0}|.$$

Wegen $q^{n-n_0}|a_{n_0}| = \frac{|a_{n_0}|}{q^{n_0}} q^n$ konvergieren die zugehörigen Partialsummen als geometrische Reihe mit Grenzwert

$$\sum_{n=1}^{\infty} q^{n-n_0}|a_{n_0}| = \frac{|a_{n_0}|}{q^{n_0}} \frac{1}{1-q}.$$

Damit haben wir eine konvergente Majorante gefunden und die Reihe über $(a_n)_n$ konvergiert absolut nach dem Majorantenkriterium.

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

Bemerkung 8.31. Ist $(a_n)_n$ eine reelle Folge, für die die Folge $(q_n)_n$ aus dem Quotientenkriterium konvergiert und gilt $\lim_{n\to\infty} q_n > 1$, so ist die Reihe über $(a_n)_n$ divergent. Warum?

Beispiel 8.32. Als Beispielanwendung für das Quotientenkriterium untersuchen wir für jedes $x \in \mathbb{R}$ die Reihe
$$\left(\sum_{k=0}^{n} \frac{x^k}{k!}\right)_n .$$
Für $x = 0$ besteht die Reihe nur aus einem Summanden und ist daher offensichtlich konvergent. Weiterhin ist die Konvergenz für $x = 1$ nach Satz 7.25 schon bekannt. Sei $x \neq 0$. Dann ist jeder Summand von null verschieden, sodass wir das Quotientenkriterium anwenden können. In diesem Fall ist $a_n = {x^n}/{n!}$ und wir 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 $\limsup_{n\to\infty} q_n = 0$, d. h. die Reihe ist absolut konvergent und zwar für jedes beliebige $x \in \mathbb{R}$.

Das Ergebnis des vorhergehenden Beispiels ermöglicht die folgende Definition.

Definition 8.33. *Wir nennen die Reihe* $\left(\sum_{k=0}^{n} {x^k}/{k!}\right)_n$. Exponentialreihe *und definieren die* Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ *durch*
$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}.$$
Im weiteren Verlauf werden wir zeigen, dass der Name dieser Funktion insofern gerechtfertigt ist, als $\mathrm{e}^r = \exp(r)$ für alle $r \in \mathbb{Q}$ gilt. Die Setzung $\mathrm{e}^x := \exp(x)$ für beliebiges $x \in \mathbb{R}$ liefert uns also eine Option (nicht unbedingt die naheliegende), um mit Hilfe des noch zu definierenden Logarithmus allgemeine Potenzen mit reellen Exponenten zu definieren.

8.4. Das Cauchyprodukt

In diesem Abschnitt beschäftigen wir uns mit dem Produkt von Reihen. Dabei werden unsere Überlegungen zu Umordnungen von Reihen ihre Macht entfalten. Sollen zwei Reihen $(\sum_{k=0}^{n} a_k)_n$ und $(\sum_{k=0}^{n} b_k)_n$ multipliziert werden, so erhält man

1) zuerst eine Folge $(S_n)_n$, die aus den Produkten der Partialsummen entsteht:
$$S_n := \left(\sum_{k=0}^{n} a_k\right)\left(\sum_{k=0}^{n} b_k\right) = \sum_{j=0}^{n}\sum_{k=0}^{n} a_j b_k.$$
Insbesondere ist noch nicht ganz klar, inwiefern es sich bei diesem Produkt auch um eine Reihe handelt. Sind beide Reihen konvergent, dann ist es auch $(S_n)_n$ nach Satz 7.12, mit Grenzwert
$$\lim_{n\to\infty} S_n = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big). \tag{8.9}$$

2) Multipliziert man die „unendlichen Summen“ formal aus, so erhält man die Produkte der Form $a_j b_k$ mit $j, k \in \mathbb{N}_0$ je genau einmal. Ordnet man all diese Produkte in einer Folge $(p_n)_{n\geq 0}$ an, wobei jedes einzelne $a_j b_k$ genau einmal vorkommt, so erhält man eine sogenannte *Produktreihe* $\left(\sum_{k=0}^n p_k\right)_n$ von $\left(\sum_{k=0}^n a_k\right)_n$ und $\left(\sum_{k=0}^n b_k\right)_n$. Präziser formuliert bekommt man eine solche Produktreihe, indem man eine bijektive Abbildung $\Phi : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ wählt und $p_{\Phi(j,k)} = a_j b_k$ für alle Paare $j, k \in \mathbb{N}_0$ setzt. Die Funktion Φ kann man oft nicht gut explizit angeben, aber das folgende Beispiel zeigt zwei Möglichkeiten am Bild.

Beispiel 8.34. Wir geben zwei Beispiele für die Anordnungen der Produkte. Dazu schreiben wir uns die zu nummerierenden Summanden in ein Matrix-Schema und nummerieren auf verschiedenen Wegen:

(a) Anordnung nach Diagonalen:

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

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 8.23 können wir hoffen, dass es zumindest für absolut konvergente Reihen egal ist, wie die Summanden im Produkt angeordnet werden. Tatsächlich gilt das folgende Resultat.

Satz 8.35. *Es seien* $\left(\sum_{k=0}^n a_k\right)_n$ *und* $\left(\sum_{k=0}^n b_k\right)_n$ *zwei absolut konvergente Reihen und* $\left(\sum_{k=0}^n p_k\right)_n$ *irgendeine ihrer Produktreihen. Dann konvergiert auch* $\left(\sum_{k=0}^n p_k\right)_n$ *absolut und es gilt*

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

Beweis. In einem ersten Schritt zeigen wir die Behauptung für den speziellen Fall der Anordnung nach Quadraten wie oben in (b). So anzufangen ist logisch, weil das dem Produkt 1), für das wir aus 8.9 schon den Grenzwert kennen, am meisten entspricht, denn es ist ja

$$S_n = \sum_{j=0}^{n}\sum_{k=0}^{n} a_j b_k = \sum_{k=0}^{n^2+2n} p_k,$$

wenn die Anordnung $(p_n)_n$ wie in (b) gewählt ist. Das liegt daran, dass das Quadrat bis zum n-ten Index von a_k und b_k genau die $(n+1)^2 = n^2+2n+1$ Einträge $p_0, p_1, \ldots, p_{n^2+2n}$ hat. Wir setzen $s_n := \sum_{k=0}^n p_k$ und $\sigma_n := \sum_{k=0}^n |p_k|$. Wir wählen ein $n \in \mathbb{N}_0$ fix und betrachten alle Summanden von σ_n. Dann gilt wegen $n \leq n^2+2n$, der Positivität der Summanden und der absoluten Konvergenz der Reihen:

$$\sigma_n = \sum_{k=0}^{n} |p_k| \leq \Big(\sum_{k=0}^{n} |a_k|\Big)\Big(\sum_{k=0}^{n} |b_k|\Big) \leq \Big(\sum_{k=0}^{\infty} |a_k|\Big)\Big(\sum_{k=0}^{\infty} |b_k|\Big).$$

Also ist die Folge $(\sigma_n)_n$ beschränkt. Nach dem Monotoniekriterium für Reihen ist damit die Reihe $\left(\sum_{k=0}^n p_k\right)_n = (s_n)_n$ absolut konvergent und insbesondere auch konvergent. Da $(s_{n^2+2n})_n$ eine konvergente Teilfolge von $(s_n)_n$ ist, hat auch $(s_n)_n$ den Grenzwert

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

Wäre schließlich eine andere Produktreihe von $\left(\sum_{k=0}^n a_k\right)_n$ und $\left(\sum_{k=0}^n b_k\right)_n$ gegeben, so wäre diese eine Umordnung von $\left(\sum_{k=0}^n p_k\right)_n$ und die Behauptung folgt aus Satz 8.23. □

Die Bedeutung von Satz 8.35 liegt darin, dass wir jetzt alternativ aufsummieren können, um Aufschluss über den Grenzwert zu erhalten. Insbesondere hat die Summation über die Diagonalen, etwa wie oben in (a), eine schöne Struktur und einen eigenen Namen:

Definition 8.36. *Es seien zwei Reihen* $\left(\sum_{k=0}^n a_k\right)_n$ *und* $\left(\sum_{k=0}^n b_k\right)_n$ *gegeben. Dann heißt die Reihe*

$$\Big(\sum_{k=0}^{n}\sum_{j=0}^{k} a_j b_{k-j}\Big)_n$$

das Cauchy-Produkt *von* $\left(\sum_{k=0}^n a_k\right)_n$ *und* $\left(\sum_{k=0}^n b_k\right)_n$.

Satz 8.37. *Sind* $\left(\sum_{k=0}^{n} a_k\right)_n$ *und* $\left(\sum_{k=0}^{n} b_k\right)_n$ *zwei absolut konvergente Reihen, so ist das Cauchy-Produkt der beiden ebenfalls absolut konvergent mit Reihenwert*

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n} a_k b_{n-k} = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big).$$

Der Beweis dieses Satzes ergibt sich sofort aus Satz 8.35 und dem folgenden Lemma über stückweises Aufsummieren, weil die Summanden $\sum_{j=0}^{k} a_j b_{k-j}$ des Cauchy-Produkts ja die Summen über die Diagonalen im Bild in (a) sind.

Lemma 8.38. *Es sei* $\left(\sum_{k=0}^{n} a_k\right)_n$ *konvergent und* $(n_1, n_2, n_3, \dots)$ *eine streng monoton wachsende Folge in* $\mathbb{N}_0$*. Setzen wir*

$$b_0 := \sum_{k=0}^{n_1} a_k$$

und für jedes $j \in \mathbb{N}$

$$b_j = \sum_{n_j+1}^{n_{j+1}} a_k,$$

so ist auch die Reihe $\left(\sum_{k=0}^{n} b_k\right)_n$ *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 $j \in \mathbb{N}$

$$\sigma_j = \sum_{k=0}^{n_{j+1}} a_k = s_{n_{j+1}}.$$

Also ist $(\sigma_j)_j$ eine Teilfolge von $(s_n)_n$ und da diese konvergiert, gilt insbesondere auch $\sigma_n \to s$ $(n \to \infty)$, was die Behauptung beweist. □

Wir wollen das Cauchy-Produkt nun verwenden, um einen weiteren Grenzwert zu bestimmen:

Beispiel 8.39. Für jedes $q \in (-1, 1)$ ist die geometrische Reihe $\left(\sum_{k=0}^{\infty} q^k\right)_n$ absolut konvergent, also erhalten wir für diese q auch:

$$\lim_{n\to\infty}\sum_{k=0}^{n}(k+1)q^k = \lim_{n\to\infty}\sum_{k=0}^{n} q^k \sum_{j=0}^{k} 1 = \lim_{n\to\infty}\sum_{k=0}^{n}\sum_{j=0}^{k} q^k = \lim_{n\to\infty}\sum_{k=0}^{n}\sum_{j=0}^{k} q^j q^{k-j}$$
$$\overset{\text{Cauchy-Produkt}}{=} \Big(\sum_{n=0}^{\infty} q^n\Big)\Big(\sum_{n=0}^{\infty} q^n\Big) = \Big(\frac{1}{1-q}\Big)^2$$

Abschließend verwenden wir das Cauchyprodukt, um die folgenden grundlegeden Eigenschaften der Exponentialfunktion zu zeigen.

Satz 8.40. *(a) Es ist* $\exp(0) = 1$ *und* $\exp(1) = \mathrm{e}$.

(b) Für alle $x, y \in \mathbb{R}$ *gilt* $\exp(x)\exp(y) = \exp(x+y)$. (Funktionalgleichung der Exponentialfunktion)

(c) Für alle $x \in \mathbb{R}$ *ist* $\exp(x) > 0$ *und* $\exp(x)^{-1} = \exp(-x)$.

(d) Für alle $r \in \mathbb{Q}$ *gilt* $\exp(r) = \mathrm{e}^r$.

(e) Die Funktion $\exp : \mathbb{R} \to \mathbb{R}$ *ist streng monoton wachsend, d. h. für alle* $x, y \in \mathbb{R}$ *mit* $x < y$ *gilt* $\exp(x) < \exp(y)$.

Beweis. (a) $\exp(0) = 1$ ist klar und $\exp(1) = \mathrm{e}$ ergibt sich aus der Definition der Eulerzahl in Definition 7.26 und Satz 7.25.

(b) Es seien $x, y \in \mathbb{R}$. Dann sind die Werte $\exp(x)$ und $\exp(y)$ der Exponentialfunktion durch absolut konvergente Reihen gegeben. Wir erhalten also mit dem Cauchyprodukt, vgl. Satz 8.37:

$$\exp(x)\exp(y) = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{y^n}{n!}\right) = \sum_{n=0}^{\infty} c_n,$$

wobei

$$c_n = \sum_{k=0}^{n} \frac{x^k}{k!} \cdot \frac{y^{n-k}}{(n-k)!}$$

ist. Mit Hilfe der Binomialformel gilt

$$c_n = \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k y^{n-k} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \frac{1}{n!}(x+y)^n$$

Setzen wir das wieder oben ein, erhalten wir

$$\exp(x)\exp(y) = \sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \frac{1}{n!}(x+y)^n = \exp(x+y).$$

(c) Mit Hilfe der ersten beiden Punkte gilt für alle $x \in \mathbb{R}$

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

Also gilt $\exp(x) \neq 0$ und $\exp(-x) = \exp(x)^{-1}$. Um die Positivität von $\exp(x)$ nachzuweisen, betrachten wir zunächst den Fall $x \geq 0$. Dann gilt

$$\exp(x) = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!} \geq 1 > 0.$$

Ist $x < 0$, so ist $\exp(x) = \exp(-x)^{-1} > 0$, da $-x > 0$ ist.

(d) Zunächst ist $\exp(0) = 1 = \mathrm{e}^0$. Sei nun $n \in \mathbb{N}$. Dann gilt mit der Funktionalgleichung

$$\exp(n) = \exp(\underbrace{1 + \cdots + 1}_{n \text{ Stück}}) = (\exp(1))^n = \mathrm{e}^n$$

und genauso

$$\mathrm{e} = \exp(1) = \exp\Big(n \cdot \frac{1}{n}\Big) = \Big(\exp\Big(\frac{1}{n}\Big)\Big)^n.$$

Also ist $\exp(1/n) = \mathrm{e}^{1/n}$. Somit gilt für strikt positive rationale Zahlen $r := m/n$ mit $n, m \in \mathbb{N}$ die Gleichung

$$\exp(r) = \exp\underbrace{\Big(\frac{1}{n} + \cdots + \frac{1}{n}\Big)}_{m \text{ Stück}} = \Big(\exp\Big(\frac{1}{n}\Big)\Big)^m = \big(\mathrm{e}^{1/n}\big)^m = \mathrm{e}^{m/n} = \mathrm{e}^r.$$

Ist $r \in \mathbb{Q}$ kleiner als null, so ist $-r > 0$ und es gilt $\exp(r) = \exp(-r)^{-1} = (\mathrm{e}^{-r})^{-1} = \mathrm{e}^r$.

(e) Es seien $x, y \in \mathbb{R}$ mit $x < y$ gegeben. Dann ist $y - x > 0$ und damit gilt wie im den Beweis von (c) $\exp(y - x) > 1$. Damit ist

$$1 < \exp(y - x) = \exp(y)\exp(-x) = \frac{\exp(y)}{\exp(x)},$$

also $\exp(y) > \exp(x)$. □

Wir definieren die Werte von e^x für nichtrationale Exponenten nun mit Hilfe der Exponentialfunktion.

Definition 8.41. *Für alle $x \in \mathbb{R}$ schreiben wir $\mathrm{e}^x := \exp(x)$.*

8.5. Potenzreihen

Definition 8.42. *Es sei $(a_n)_{n\geq 0}$ eine Folge in $\mathbb{R}$ und $x_0 \in \mathbb{R}$. Eine Reihe der Form*

$$\Big(\sum_{k=0}^{n} a_k (x - x_0)^k\Big)_n$$

heißt Potenzreihe *in $\mathbb{R}$ mit* Entwicklungspunkt x_0.

Hier ist zu beachten, dass wieder auch der Grenzwert, wenn er existiert,

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + \ldots, \qquad x \in \mathbb{R},$$

Potenzreihe genannt wird. Für alle $x \in \mathbb{R}$, für die dieser Grenzwert existiert, wird durch

$$x \mapsto \sum_{n=0}^{\infty} a_n (x - x_0)^n \tag{8.10}$$

die zugehörige reelle Funktion definiert.
Zwei Beispiele von Potenzreihen mit Entwicklungspunkt null sind durch die Exponentialfunktion (Beispiel 8.32) und die geometrische Reihe (Beispiel 8.4 (b)) gegeben. Wir wollen eine allgemeine Theorie solcher Reihen entwickeln. Jede Potenzreihe konvergiert an ihrem Entwicklungspunkt, also für $x = x_0$. Zur weiteren Untersuchung der Konvergenz einer Potenzreihe dient der folgende Satz, der sich aus dem Wurzelkriterium ableitet.

Satz 8.43 (Satz von Hadamard)**.** *Es sei* $\left(\sum_{k=0}^{n} a_k (x - x_0)^k\right)_n$ *eine Potenzreihe. Dann gilt:*

(a) Ist die Folge $(\sqrt[n]{|a_n|})$ *unbeschränkt, so konvergiert die Potenzreihe nur für* $x = x_0$.

(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* $|x - x_0| < 1/\varrho$ *absolut konvergent und für alle* $x \in \mathbb{R}$ *mit* $|x - x_0| > 1/\varrho$ *divergent.*

Beweis. Wir zeigen zunächst (a). Dazu nehmen wir uns ein $x \neq x_0$. Dann ist $\sqrt[n]{|a_n(x - x_0)^n|} = |x - x_0| \sqrt[n]{|a_n|}$ für jedes $n \in \mathbb{N}$ und diese Folge ist nach Voraussetzung unbeschränkt. Damit folgt die Divergenz der Reihe aus dem Wurzelkriterium (Satz 8.27).
Auch (b) und (c) ergeben sich aus dem Wurzelkriterium, denn hier gilt

$$\limsup_{n\to\infty} \sqrt[n]{|a_n(x - x_0)^n|} = \limsup_{n\to\infty} |x - x_0| \sqrt[n]{|a_n|} = |x - x_0| \cdot \varrho.$$

Ist $\varrho = 0$, so ist dieser Wert immer null und damit kleiner als eins und wir bekommen die Aussage in (b). Ist $\varrho > 0$, so ist der betrachtete Limes superior, je nachdem, ob $|x - x_0| < 1/\varrho$ oder $|x - x_0| > 1/\varrho$ gilt, größer oder kleiner als 1, was nach dem Wurzelkriterium genau absolute Konvergenz von Divergenz der Reihe trennt. Das liefert (c). □

Die Zahl $1/\varrho$ aus obigem Satz enthält also wesentliche Informationen zur Konvergenz der Potenzreihe. Sie ist damit eine charakteristische Größe, die einen Namen verdient hat.

Definition 8.44. *Es sei* $\left(\sum_{k=0}^{n} a_k(x-x_0)^k\right)_n$ *eine Potenzreihe und im Falle, dass* $(\sqrt[n]{|a_n|})$ *beschränkt ist,* ϱ *wie in Satz 8.43. Dann heißt die Zahl*

$$r := \begin{cases} 0, & \textit{falls in obigem Satz (a) gilt,} \\ \infty, & \textit{falls in obigem Satz (b) gilt,} \\ \frac{1}{\varrho}, & \textit{falls in obigem Satz (c) gilt,} \end{cases}$$

der Konvergenzradius *der Potenzreihe.*

Die Funktion in 8.10 ist also nach Satz 8.43 nur auf $\{x_0\}$, auf ganz $\mathbb{R}$, oder auf einem Intervall I mit

$$(x_0 - r, x_0 + r) \subseteq I \subseteq [x_0 - r, x_0 + r]$$

definiert. Insbesondere zeigt der Satz von Hadamard auch, dass r der eindeutige *maximale* Wert ist, für den das gilt. Der Konvergenzradius kann oft wie im Satz von Hadamard durch das Wurzelkriterium bestimmt werden, aber alternativ natürlich auch durch andere Argumente, die die Struktur der Reihe nutzen.
Wir betrachten einige Beispiele, die verdeutlichen, dass am Rand des Konvergenzintervalls alles passieren kann. Dazu betrachten wir, der Übersichtlichkeit halber, Reihen mit Entwicklungspunkt null.

Beispiel 8.45. (a) Wir beginnen mit $a_n := 1$, $n \in \mathbb{N}_0$, d. h. der Potenzreihe

$$\left(\sum_{k=0}^{n} x^k\right)_n .$$

Das ist die geometrische Reihe aus Beispiel 8.4 (b) und insofern ist es auch nicht verwunderlich, dass hier wegen $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} 1 = 1$ der Konvergenzradius 1 ist. D. h. diese Potenzreihe konvergiert, wie wir schon wissen, absolut für $|x| < 1$ und divergiert für $|x| > 1$. Über das Konvergenzverhalten am Rand des Konvergenzintervalls gibt der Konvergenzradius keine Auskunft. Aber wir wissen schon, dass diese sowohl für $x = -1$ als auch für $x = 1$ divergiert.

(b) Nun sei $a_n = 1/n$, $n \in \mathbb{N}$, wir betrachten also

$$\left(\sum_{k=1}^{n} \frac{x^k}{k}\right)_n .$$

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, aber nicht absolut konvergent), sodass wir 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

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

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 8.25 (b) konvergente Reihe über ${}^1\!/{n^2}$ und für $x = -1$ die nach dem Leibniz-Kriterium konvergente Reihe $\sum_{n=1}^{\infty} {}^{(-1)^n}\!/{n^2}$. Also ist in diesem Fall das Konvergenzintervall der Potenzreihe $[-1, 1]$ und wir haben Divergenz für alle $x \in \mathbb{R}$ mit $|x| > 1$.

Beispiel 8.46. Es kommt immer mal wieder vor, dass in einer Potenzreihe nicht alle Potenzen vorkommen, wie z. B. in

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

Will man dann den Satz von Hadamard zur Bestimmung des Konvergenzradius anwenden, muss man zunächst dafür sorgen, dass die Reihe die dort behandelte Form hat. Das erreicht man hier durch die Substitution $y := x^3$. Diese liefert die Reihe mit Gliedern

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

Der Konvergenzradius dieser Reihe berechnet sich nach Hadamard als Kehrwert von

$$\limsup_{n\to\infty} \sqrt[n]{\left(\frac{2}{3}\right)^n} = \frac{2}{3},$$

zu ${}^3\!/{2}$. Weiter ist

$$|x|^3 = |x^3| = |y| < \frac{3}{2} \iff |x| < \sqrt[3]{\frac{3}{2}}.$$

Wir bekommen für die ursprünglich betrachtete Reihe also den Konvergenzradius $\sqrt[3]{{}^3\!/{2}}$ heraus.

Bemerkung 8.47. Der Satz von Hadamard kann uns auch noch zu einer anderen, eher unerwarteten Erkenntnis verhelfen. Wir haben in Beispiel 8.32 mit Hilfe des Quotientenkriteriums gesehen, dass der Grenzwert der Exponentialreihe

$\sum_{n=0}^{\infty} x^n/n!$ für alle $x \in \mathbb{R}$ existiert. 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 7.41 und es gilt

$$\lim_{n\to\infty} \frac{1}{\sqrt[n]{n!}} = \limsup_{n\to\infty} \frac{1}{\sqrt[n]{n!}} = 0.$$

Natürlich kann man auch mit dem Quotientenkriterium Konvergenzuntersuchungen bei Potenzreihen anstellen. Wir führen das exemplarisch an einem prominenten Beispiel durch.

Beispiel 8.48. Wir betrachten die Potenzreihe

$$\left(\sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k}\right)_n$$

und untersuchen diese mit dem Quotientenkriterium für Reihen aus Satz 8.30. Für den dort betrachteten Quotienten gilt

$$\left|\frac{\frac{(-1)^{n+1}}{(2n+2)!}x^{2n+2}}{\frac{(-1)^n}{(2n)!}x^{2n}}\right| = \left|\frac{x^{2n+2}}{(2n+2)!} \cdot \frac{(2n)!}{x^{2n}}\right| = \frac{x^2}{(2n+2)(2n+1)},$$

und dieser Ausdruck geht für jedes $x \in \mathbb{R}$ gegen null für $n \to \infty$. Also konvergiert die Potenzreihe $\left(\sum_{k=0}^{n} (-1)^k/(2k)!\, x^{2k}\right)_n$ für jedes $x \in \mathbb{R}$ absolut. Der Konvergenzradius ist damit unendlich.

Was war daran nun prominent? Wie bei der Exponentialfunktion ist durch diese Potenzreihe eine Funktion auf ganz $\mathbb{R}$ gegeben, der wir einen Namen geben.

Definition 8.49. *Für jedes $x \in \mathbb{R}$ nennen wir die Zahl*

$$\cos(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$$

den Cosinus *von x.*

Ganz ähnlich wie in obigem Beispiel kann man für die Reihe in der folgenden Definition absolute Konvergenz für alle $x \in \mathbb{R}$ zeigen.

Definition 8.50. *Für jedes $x \in \mathbb{R}$ nennen wir die Zahl*

$$\sin(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

den Sinus *von x.*

Wir wollen uns nun dem Cauchyprodukt zweier Potenzreihen zuwenden. Wie im Beispiel der Exponentialfunktion können wir damit durch Potenzreihen gegebene Funktionen gegebenenfalls leichter multiplizieren.

Satz 8.51. *Es seien* $\left(\sum_{k=0}^{n} a_k(x-x_0)^k\right)_n$ *und* $\left(\sum_{k=0}^{n} b_k(x-x_0)^k\right)_n$ *Potenzreihen mit demselben Entwicklungspunkt* $x_0 \in \mathbb{R}$ *und zugehörigen Konvergenzradien* $r_a > 0$ *bzw.* $r_b > 0$ *(dabei sind* $r_a = \infty$ *oder* $r_b = \infty$ *zugelassen). Wir setzen*

$$c_n := \sum_{k=0}^{n} a_k b_{n-k}, \ n \in \mathbb{N}_0, \quad \textit{und} \quad R := \min\{r_a, r_b\}.$$

Dann ist das Cauchyprodukt der beiden Potenzreihen gegeben durch die Potenzreihe $\sum_{n=0}^{\infty} c_n(x-x_0)^n$*, deren Konvergenzradius beträgt mindestens* R*, und es gilt für alle* $x \in \mathbb{R}$ *mit* $|x-x_0| < R$

$$\sum_{n=0}^{\infty} c_n(x-x_0)^n = \Big(\sum_{n=0}^{\infty} a_n(x-x_0)^n\Big)\Big(\sum_{n=0}^{\infty} b_n(x-x_0)^n\Big).$$

Der Beweis des Satzes ist auf Übungsblatt 6 gefragt. Dort wird auch mit Hilfe des Cauchyprodukts bewiesen, dass für alle $x \in \mathbb{R}$ die Formel

$$\sin^2(x) + \cos^2(x) = 1$$

gilt.

9. *Komplexe Zahlen

Wir wollen nun noch einmal unseren Zahlenraum erweitern und die komplexen Zahlen einführen. Betrachten wir die bisherigen Zahlenraumerweiterungen, 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 dann in den Algebra-Vorlesungen sehen, dass man in den komplexen Zahlen jede polynomiale Gleichung lösen kann.
Vieles, was wir in Teil III der Vorlesung über und mit Konvergenz von reellen Folgen und Reihen bewiesen haben, überträgt sich ganz analog für komplexe Folgen und Reihen und wird in diesem Kapitel zusammengefasst.

Definition 9.1. *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 9.2. Wir berechnen einige wichtige Produkte:

$$\begin{aligned}
(0,1)^2 = (0,1) \odot (0,1) &= (0\cdot 0 - 1\cdot 1, 0\cdot 1 + 1\cdot 0) = (-1,0)\\
(x,y) \odot (1,0) &= (x\cdot 1 - y\cdot 0, x\cdot 0 + y\cdot 1) = (x,y) \quad \text{für alle } x,y \in \mathbb{R}\\
(y,0) \odot (0,1) &= (y\cdot 0 - 0\cdot 1, y\cdot 1 + 0\cdot 0) = (0,y) \quad \text{für alle } y \in \mathbb{R}.
\end{aligned}$$

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 9.3. (a) Nun können wir auch die Gleichung $z^2 = -1$ lösen, denn es gilt nach der ersten Berechnung in Beispiel 9.2

$$\mathrm{i}^2 = (0, 1) \odot (0, 1) = (-1, 0) = -1,$$

genauso wie übrigens auch $(-\mathrm{i})^2 = (0, -1) \odot (0, -1) = -1$ gilt.

(b) Der Vorteil davon, die komplexe Zahl $z = (x, y) \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ als $x + \mathrm{i}y$ zu schreiben, ist, dass man damit 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 Definition 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 nach obigen Überlegungen mit den komplexen Zahlen wie gewohnt weiter rechnen können.

Satz 9.4. *Die Menge $\mathbb{C}$ mit den eingeführten Verknüpfungen Plus und Mal ist im algebraischen Sinne ein Körper, d. h. sie erfüllt die Axiome (A1) bis (A9) aus Kapitel 4.*

Beweis. Das Axiom (A6) haben wir bereits in Beispiel 9.2 nachgerechnet und auch fast alle anderen Aussagen des Satzes kann man schlicht nachrechnen. Wir behandeln hier nur noch (A7). Sei $z = x + \mathrm{i}y \in \mathbb{C} \setminus \{0\}$ mit $x, y \in \mathbb{R}$. Da $z \neq 0$ ist, muss $x \neq 0$ oder $y \neq 0$ gelten. Insbesondere ist $x^2 + y^2 > 0$. Wir setzen

$$z^{-1} := \frac{x}{x^2 + y^2} + \mathrm{i}\frac{-y}{x^2 + y^2} = \frac{x - \mathrm{i}y}{x^2 + y^2}.$$

Dann gilt unter anderem mit dem Distributivgesetz (A9)

$$zz^{-1} = (x+\mathrm{i}y)\frac{x-\mathrm{i}y}{x^2+y^2} = \frac{x^2+\mathrm{i}yx-\mathrm{i}xy-\mathrm{i}^2y^2}{x^2+y^2} = \frac{x^2+y^2}{x^2+y^2} = 1,$$

wir haben damit also ein multiplikativ-inverses Element zu z angegeben. □

Veranschaulichen kann man sich die komplexen Zahlen am besten in der *komplexen Zahlenebene* (auch *Gaußsche Zahlenebene* genannt), s. Abbildung 9.1.

Definition 9.5. *Ist $z = x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ eine komplexe Zahl, so heißt*

$\overline{z} := x - \mathrm{i}y$ — *die zu z* konjugiert komplexe *Zahl,*

$|z| := \sqrt{x^2+y^2}$ — *der* Betrag *von z.*

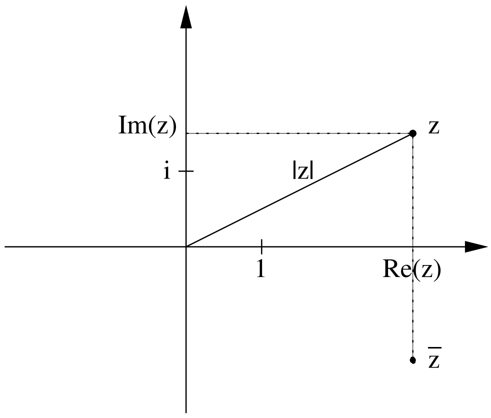

Abbildung 9.1.: Die Gaußsche Zahlenebene

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 9.6. *Sind $z, z_1, z_2 \in \mathbb{C}$, so gilt*

(a) $\overline{z_1+z_2} = \overline{z_1} + \overline{z_2}$ *und* $\overline{z_1z_2} = \overline{z_1}\cdot\overline{z_2}$.

(b) $z + \overline{z} = 2\cdot\mathrm{Re}(z)$ *und* $z - \overline{z} = 2\,\mathrm{i}\cdot\mathrm{Im}(z)$.

Beweis. (a) Es gilt für $z_1 = x_1 + \mathrm{i}y_1$ und $z_2 = x_2 + \mathrm{i}y_2$

$$\begin{aligned}\overline{z_1+z_2} &= \overline{x_1+x_2+\mathrm{i}(y_1+y_2)} = x_1+x_2-\mathrm{i}(y_1+y_2)\\ &= x_1-\mathrm{i}y_1+x_2-\mathrm{i}y_2 = \overline{z_1}+\overline{z_2}\end{aligned}$$

und

$$\begin{aligned}\overline{z_1z_2} &= \overline{(x_1x_2-y_1y_2)+\mathrm{i}(x_1y_2+x_2y_1)} = (x_1x_2-y_1y_2)-\mathrm{i}(x_1y_2+x_2y_1)\\ &= x_1x_2-\mathrm{i}x_1y_2-\mathrm{i}x_2y_1-y_1y_2 = (x_1-\mathrm{i}y_1)(x_2-\mathrm{i}y_2) = \overline{z_1}\cdot\overline{z_2}.\end{aligned}$$

(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 9.7. *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) Es gilt $z\overline{z} = (x + \mathrm{i}y)(x - \mathrm{i}y) = x^2 + \mathrm{i}yx - \mathrm{i}yx - \mathrm{i}^2 y^2 = x^2 + y^2 = (\sqrt{x^2 + y^2})^2 = |z|^2$.

(b) Diese Aussage 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 nur die Lösung $x = y = 0$, also ist in diesem Fall $z = 0$.

(d) Nachrechnen.

(e) Nach (a) und Satz 9.6 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 $\mathrm{Re}(z_1\overline{z_2}) \leq |\mathrm{Re}(z_1\overline{z_2})| \leq |z_1\overline{z_2}| = |z_1||z_2|$. Also ist

$$|z_1 + z_2|^2 \leq |z_1|^2 + |z_2|^2 + 2|z_1||z_2| = (|z_1| + |z_2|)^2.$$

Da alle beteiligten Größen positiv sind, kann man auf beiden Seiten der Ungleichung die Wurzel ziehen und erhält die Behauptung. □

Eine natürliche Frage ist nun: Welche Erkenntnisse, die wir bisher in $\mathbb{R}$ gewonnen haben, gelten auch in $\mathbb{C}$ weiter? Zunächst schauen wir uns aber an, was nicht geht.

Satz 9.8. *Es gibt auf $\mathbb{C}$ keine Relation „$\leq$“, die die Axiome (A10) bis (A14) erfüllt.*

Beweis. Nehmen wir an, es ginge doch, so muss nach (A10) entweder $\mathrm{i} \geq 0$ oder $\mathrm{i} \leq 0$ gelten. Wäre $\mathrm{i} \geq 0$, so ist wegen (A14) auch $-1 = \mathrm{i} \cdot \mathrm{i} \geq 0$ und mit Satz 4.9 (a) folgt $1 \leq 0$, was ein Widerspruch zu (c) desselben Satzes ist.
Nehmen wir dagegen an, dass $\mathrm{i} \leq 0$ ist, so bekommen wir mit Satz 4.9 (b) die Ungleichung $-1 \geq 0$, was zum selben Widerspruch führt. □

Es ist also nicht möglich, zwei komplexe Zahlen der Größe nach zu vergleichen. Man kann aber Beträge von komplexen Zahlen vergleichen, denn das sind reelle Zahlen. Das werden wir uns oft zu Nutze machen. So kann man z. B. unsere Schreibweise $U_\varepsilon(x_0)$ für eine ε-Umgebung sofort nach $\mathbb{C}$ ziehen: Für $\varepsilon > 0$ und $z_0 \in \mathbb{C}$ setzen wir

$$U_\varepsilon(z_0) := \{z \in \mathbb{C} : |z - z_0| < \varepsilon\}.$$

Wir wollen im Rest dieses Abschnitts einen Schnelldurchlauf durch die komplexen Folgen und Reihen machen und die Exponentialfunktion in $\mathbb{C}$ definieren. Wir werden dabei 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 zunächst den Begriff der Konvergenz. Diesen können wir auf Konvergenz in $\mathbb{R}$ zurückspielen, vgl. Satz 7.11 (a).

Definition 9.9. *Es sei $(z_n)_n$ eine Folge in $\mathbb{C}$.*

(a) *Die Folge $(z_n)_n$* konvergiert gegen $z_0 \in \mathbb{C}$, *wenn die reelle Folge $(|z_n - z_0|)_n$ in $\mathbb{R}$ gegen Null konvergiert.*

(b) *Wir nennen $(z_n)_n$* beschränkt, *wenn es ein $C \geq 0$ gibt mit $|z_n| \leq C$ für alle $n \in \mathbb{N}$.*

(c) *Wir nennen $(z_n)_n$* Cauchy-Folge *in $\mathbb{C}$, wenn für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit $|z_n - z_m| < \varepsilon$ für alle $n, m \geq n_0$.*

(d) *Eine Reihe $\left(\sum_{k=1}^n z_k\right)_n$ mit komplexen Summanden heißt* absolut konvergent, *falls die zugehörige (reelle) Reihe über die Beträge $\left(\sum_{k=1}^n |z_k|\right)_n$ konvergiert.*

Satz 9.10. *Es sei $(z_n)_n$ eine Folge in $\mathbb{C}$.*

(a) *Die Folge $(z_n)_n$ konvergiert genau dann gegen $z_0 \in \mathbb{C}$, wenn die beiden Folgen $(\mathrm{Re}(z_n))_n$ und $(\mathrm{Im}(z_n))_n$ 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) *Eine Folge in $\mathbb{C}$ ist genau dann eine Cauchy-Folge, wenn sie konvergiert.* (Cauchy-Kriterium)

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 $(z_n)_n$ 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 7.15.

(b) Dass jede konvergente Folge eine Cauchyfolge ist, zeigt man genauso wie im Kapitel 7.6 in $\mathbb{R}$. Für die umgekehrte Implikation schätzt man mit Hilfe von Satz 9.7 (b) folgendermaßen ab:

$$|\mathrm{Re}(z_n) - \mathrm{Re}(z_m)| = |\mathrm{Re}(z_n - z_m)| \leq |z_n - z_m|, \qquad n, m \in \mathbb{N}.$$

Ist nun $(z_n)_n$ eine Cauchyfolge in $\mathbb{C}$, so liefert diese Abschätzung, dass die reelle Folge $(\mathrm{Re}(z_n))_n$ eine Cauchyfolge in $\mathbb{R}$ ist. Dank Satz 7.47 ist diese also konvergent. Eine analoge Argumentation liefert die Konvergenz von $(\mathrm{Im}(z_n))_n$ und damit liefert Teil (a) die Behauptung. □

Mit Hilfe der Überlegungen aus dem vorhergehenden Satz lassen sich viele Resultate über reelle Folgen ins Komplexe übertragen. Als prominentes Beispiel beweisen wir eine komplexe Version des Satzes von Bolzano-Weierstraß.

Satz 9.11. *Jede beschränkte Folge in $\mathbb{C}$ besitzt eine konvergente Teilfolge.*

Beweis. Sei $(z_n)_n$ eine beschränkte Folge in $\mathbb{C}$ und $C \geq 0$ so, dass $|z_n| \leq C$ für alle $n \in \mathbb{N}$ ist. Es gilt $|\mathrm{Re}(z_n)| \leq |z_n| \leq C$ für alle $n \in \mathbb{N}$, also ist auch die reelle Folge $(\mathrm{Re}(z_n))_n$ beschränkt. Nach dem reellen Satz von Bolzano-Weierstraß in Satz 7.35 hat diese eine konvergente Teilfolge $(\mathrm{Re}(z_{n_k}))_k$. Mit einem analogen Argument ist $(\mathrm{Im}(z_{n_k}))_k$ beschränkt in $\mathbb{R}$ und wir bekommen eine erneute Teilfolge $(z_{n_{k_\ell}})_\ell$, so dass $(\mathrm{Im}(z_{n_{k_\ell}}))_\ell$ konvergent ist. Als Teilfolge der konvergenten Folge $(\mathrm{Re}(z_{n_k}))_k$ ist auch die Folge $(\mathrm{Re}(z_{n_{k_\ell}}))_\ell$ konvergent. Schließlich bedeutet das mit Satz 9.10 (a), dass $(z_{n_{k_\ell}})_\ell$ eine in $\mathbb{C}$ konvergente Folge ist und wir haben unsere konvergente Teilfolge von $(z_n)_n$ gefunden. □

Viele Sätze gelten wortwörtlich wie im Reellen. Dazu gehören auch das Cauchy-Kriterium, die Dreiecksungleichung und die Kriterien für absolute Konvergenz von komplexe Reihen. Das fassen wir in den folgenden beiden Sätzen zusammen.

Satz 9.12. *Es sei $(z_n)_n$ eine komplexe Folge. Dann gilt:*

(a) Die Reihe $\left(\sum_{k=1}^n z_k\right)_n$ konvergiert genau dann, wenn für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$\Big|\sum_{k=n+1}^{m} z_k\Big| < \varepsilon \quad \textit{für alle } m > n \geq n_0.$$

(Cauchy-Kriterium)

(b) Ist $\left(\sum_{k=1}^n z_k\right)_n$ absolut konvergent, so ist die Reihe $\left(\sum_{k=1}^n z_k\right)_n$ konvergent und es gilt

$$\Big|\sum_{n=1}^{\infty} z_n\Big| \leq \sum_{n=1}^{\infty} |z_n|. \quad \text{(verallgemeinerte Dreiecksungleichung)}$$

Satz 9.13. *Es sei (z_n) eine Folge in $\mathbb{C}$ und (a_n) eine reelle Folge.*

(a) Ist $\sum_{n=1}^{\infty} z_n$ konvergent, so gilt $\lim_{n\to\infty} z_n = 0$.

(b) Gilt $|z_n| \leq a_n$ für fast alle $n \in \mathbb{N}$ und ist $\sum_{n=1}^{\infty} a_n$ konvergent, so konvergiert $\sum_{n=1}^{\infty} z_n$ absolut. (Majorantenkriterium)

(c) Ist die Folge $(\sqrt[n]{|z_n|})$ unbeschränkt, so divergiert $\sum_{n=1}^{\infty} z_n$. Ist $(\sqrt[n]{|z_n|})$ beschränkt, so ist die Reihe absolut konvergent, wenn $\limsup_{n\to\infty} \sqrt[n]{|z_n|} < 1$ ist und divergent, wenn $\limsup_{n\to\infty} \sqrt[n]{|z_n|} > 1$ ist. (Wurzelkriterium)

(d) Gilt $z_n \neq 0$ für alle $n \in \mathbb{N}$ und ist die Folge $(|z_{n+1}/z_n|)$ beschränkt und gilt $\limsup_{n\to\infty} |z_{n+1}/z_n| < 1$, so ist die Reihe $\sum_{n=1}^{\infty} z_n$ absolut konvergent. Gilt dagegen $|z_{n+1}/z_n| \geq 1$ für fast alle $n \in \mathbb{N}$, so divergiert die Reihe $\sum_{n=1}^{\infty} z_n$. (Quotientenkriterium)

Die Beweise lassen sich entweder direkt auf die entsprechenden Resultate über reelle Reihen zurückspielen oder direkt aus dem Reellen übertragen. Gleiches gilt für den folgenden Satz über das wie in $\mathbb{R}$ definierte Cauchy-Produkt.

Definition 9.14. *Es seien $\left(\sum_{k=0}^n z_k\right)_n$ und $\left(\sum_{k=0}^n w_k\right)_n$ zwei komplexe Reihen. Dann heißt die Reihe mit den n-ten Gliedern*

$$\sum_{k=0}^{n} \sum_{j=0}^{k} z_j w_{k-j}$$

das Cauchy-Produkt *dieser beiden Reihen.*

Satz 9.15. *Es seien* $\left(\sum_{k=0}^{n} z_k\right)_n$ *und* $\left(\sum_{k=0}^{n} w_k\right)_n$ *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 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. problemlos die Exponentialfunktion für komplexe Zahlen definieren. Tatsächlich kann man leicht feststellen, dass der Satz von Hadamard 8.43 genauso in $\mathbb{C}$ funktioniert.

Beispiel 9.16. (a) Wir betrachten zunächst die komplexe geometrische Reihe, also die Potenzreihe

$$\left(\sum_{k=0}^{n} z_k\right)_n, \qquad z \in \mathbb{C}.$$

Für $z = 1$ ist diese Reihe bekanntermaßen nicht konvergent und wie in Satz 5.12 gilt für alle $z \in \mathbb{C} \setminus \{1\}$ und alle $n \in \mathbb{N}$

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

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 $\left(\sum_{k=0}^{n} |z|^k/k!\right)_n$ für alle $z \in \mathbb{C}$ konvergiert, denn $|z|$ ist ja reell. Also ist die Reihe $\left(\sum_{k=0}^{n} z^k/k!\right)_n$ für alle $z \in \mathbb{C}$ sogar absolut konvergent. Das rechtfertigt die nun folgende Definition.

Definition 9.17. *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 9.18. *(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 $\mathrm{e}^{\mathrm{i}t} = \cos(t) + \mathrm{i}\sin(t)$. (Euler-Formel)

(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 im Beweis von Satz 8.40 (b) berechnet man dies 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

$$\cos(nt) + \mathrm{i}\sin(nt) = \mathrm{e}^{\mathrm{i}nt} = \big(\mathrm{e}^{\mathrm{i}t}\big)^n = \big(\cos(t) + \mathrm{i}\sin(t)\big)^n. \qquad \square$$

Auf gleiche Weise kann man auch den Sinus und den Cosinus für komplexe Argumente definieren, denn auch deren Reihen haben Konvergenzradius unendlich.

Definition 9.19. *Für alle $z \in \mathbb{C}$ definieren wir den* Cosinus *und den* Sinus *durch*

$$\cos(z) := \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k} \qquad \textit{und} \qquad \sin(z) := \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1}.$$

Teil IV.

Reelle Funktionen

10. Stetigkeit

In diesem Abschnitt übertragen wir Begriffe, die wir im Kontext von Folgen und Reihen kennen gelernt haben, auf Funktionen von $D \subseteq \mathbb{R}$ nach $\mathbb{R}$.

10.1. Definition von Stetigkeit

Wir sagen, dass eine Funktion $f\colon \mathbb{R} \supseteq D \to \mathbb{R}$ im Punkt x_0 stetig ist, wenn sich die Konvergenz gegen x_0 jeder Folge $(x_n)_n \subseteq D$ auf die Konvergenz der Funktionswertfolge $(f(x_n))_n$ gegen $f(x_0)$ überträgt:

Definition 10.1. *Es sei $D \subseteq \mathbb{R}$ und $x_0 \in D$. Eine Funktion $f : D \to \mathbb{R}$ heißt* stetig in x_0*, wenn für jede Folge $(x_n)_n$ in D, die gegen x_0 konvergiert, auch die Folge $(f(x_n))_n$ konvergiert und* $\lim_{n\to\infty} f(x_n) = f(x_0)$ *gilt. Ansonsten nennt man* f unstetig *in x_0. .*

Weiter heißt f *stetig auf* D, wenn f in jedem Punkt $x_0 \in D$ stetig ist. Wir setzen

$$\mathrm{C}(D,\mathbb{R}) = \mathrm{C}(D) := \{f : D \to \mathbb{R} : f \text{ stetig auf } D\}.$$

Die Stetigkeit reeller Funktionen kann man tatsächlich genauso gut über das folgende *ε-δ-Kriterium* definieren, das die Konvergenz im Wertebereich (Parameter ε) über die Konvergenz im Definitionsbereich (Parameter δ) quantifiziert:

Satz 10.2. *Es sei $D \subseteq \mathbb{R}$ und $x_0 \in D$. Eine Funkion $f : D \to \mathbb{R}$ ist stetig in x_0 genau dann, wenn gilt:*

$$(\forall \varepsilon > 0)\,(\exists \delta > 0)\,(\forall x \in D) : \, |x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \varepsilon.$$

Beweis. Sei zunächst das ε-δ-Kriterium im Satz für f in x_0 erfüllt, $(x_n)_n$ eine Folge in D, die gegen x_0 konvergiert und $\varepsilon > 0$. Dann existiert nach Voraussetzung ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Da $(x_n)_n$ 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 $(f(x_n))_n$ gegen $f(x_0)$ gezeigt.
Es bleibt die umgekehrte Implikation per Kontraposition zu zeigen. Dazu nehmen wir an, f würde das ε-δ-Kriterium in x_0 nicht erfüllen. Das bedeutet (vgl. Kapitel 2 über Logik): 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 $(x_n)_n$. Wegen $|x_n - x_0| < 1/n$ für alle $n \in \mathbb{N}$ konvergiert diese Folge gegen x_0. Andererseits bleibt die Folge $(f(x_n))_n$ aber immer mindestens ε_0 von $f(x_0)$ entfernt, also konvergiert $(f(x_n))_n$ nicht gegen $f(x_0)$ und f ist per Definition in x_0 nicht stetig. □

Für die Stetigkeit von Funktionen gelten auch Grenzwertsätze:

Satz 10.3. *Es sei $D \subseteq \mathbb{R}$, $\lambda \in \mathbb{R}$ und $f, g : D \to \mathbb{R}$ seien stetig in $x_0 \in D$. Dann gelten die folgenden Aussagen.*

(a) Die Funktionen λf, $f + g$, fg und $|f|$ sind stetig in x_0.

(b) Ist $x_0 \in \widetilde{D} := \{x \in D : g(x) \neq 0\}$, so ist die Funktion $f/g : \widetilde{D} \to \mathbb{R}$ stetig in x_0.

(c) Ist $h : f(D) \to \mathbb{R}$ stetig in $f(x_0)$, so ist $h \circ f : D \to \mathbb{R}$ stetig in x_0.

Beweis. Die Beweise von (a) und (b) ergeben sich direkt aus den entsprechenden Aussagen im Grenzwertsatz 7.12 für reelle Folgen. Wir führen das nur exemplarisch für die Summe $f+g$ aus. Sei $(x_n)_n$ eine Folge in D, die gegen x_0 konvergiert. Da f und g in x_0 stetig sind, konvergieren dann auch die Folgen $(f(x_n))_n$ und $(g(x_n))_n$ und es gilt

$$\lim_{n\to\infty} f(x_n) = f(x_0) \quad \text{und} \quad \lim_{n\to\infty} g(x_n) = g(x_0).$$

Damit ist nach Satz 7.12(b) auch die Folge $(f(x_n) + g(x_n))_n$ konvergent und es gilt

$$\lim_{n\to\infty} \bigl(f(x_n) + g(x_n)\bigr) = \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) = f(x_0) + g(x_0),$$

womit die Stetigkeit von $f + g$ in x_0 bewiesen ist.
Zum Nachweis von (c) sei $(x_n)_n$ eine Folge in D mit $x_n \to x_0$ $(n \to \infty)$. Dann folgt aus der Stetigkeit von f wieder $f(x_n)_n \to f(x_0)$ $(n \to \infty)$. Die Stetigkeit von h in $f(x_0)$ liefert uns dann weiter $h(f(x_n)) \to h(f(x_0))$ $(n \to \infty)$. Also ist die Funktion $h \circ f$ stetig in x_0. □

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. Das liefert uns insbesondere, dass alle Polynome, die Exponentialfunktion, Sinus und Cosinus jeweils auf ganz $\mathbb{R}$ stetige Funktionen sind.

Satz 10.4. *Es sei $\bigl(\sum_{k=0}^{n} a_k x^k\bigr)_n$ eine Potenzreihe mit Konvergenzradius $r > 0$ (dabei ist $r = \infty$ zugelassen). Setzen wir $D := \{x \in \mathbb{R} : |x| < r\}$, bzw. $D := \mathbb{R}$, falls $r = \infty$, so ist die Funktion $f : D \to \mathbb{R}$ mit $f(x) = \sum_{n=0}^{\infty} a_n x^n$, $x \in D$, 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 $x \in \mathbb{R}$ mit $|x| \leq \varrho$, so ist

$$f(x) - f(x_0) = \sum_{n=1}^{\infty} a_n(x^n - x_0^n) \tag{10.1}$$

und für jedes $n \in \mathbb{N}$ ist (vgl. Satz 5.11 (c))

$$\begin{aligned} \left|a_n(x^n - x_0^n)\right| &= \left|a_n(x - x_0)\sum_{k=0}^{n-1} x^k x_0^{n-1-k}\right| \leq |a_n||x - x_0| \sum_{k=0}^{n-1} |x|^k |x_0|^{n-1-k} \\ &\leq |a_n||x - x_0| \sum_{k=0}^{n-1} \varrho^{n-1} = |a_n||x - x_0| n\varrho^{n-1}, \end{aligned} \tag{10.2}$$

da sowohl $|x| \leq \varrho$ als auch $|x_0| \leq \varrho$ gilt.
Wir zeigen als nächstes, dass die Reihe $\left(\sum_{k=1}^{n} |a_k| k\varrho^{k-1}\right)_n$ 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 7.45 verwendet und dabei ausgenutzt, dass $\lim_{n\to\infty} \sqrt[n]{n} = \lim_{n\to\infty} \sqrt[n]{\varrho} = 1$ gilt.
Aus dem Satz von Hadamard folgern wir im Fall $r = \infty$, dass $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 0$ ist und für $r < \infty$ erhalten wir $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 1/r$. Zusammen ist

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|n\varrho^{n-1}} \leq \frac{\varrho}{r} < 1,$$

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 $a_n(x^n - x_0^n)$ aufgrund der Abschätzung in (10.2) und mit Hilfe des Majorantenkriteriums absolut konvergent und wir können mit der verallgemeinerten Dreiecksungleichung abschätzen:

$$\begin{aligned} |f(x) - f(x_0)| &= \left|\sum_{n=1}^{\infty} a_n(x^n - x_0^n)\right| \leq \sum_{n=1}^{\infty} |a_n(x^n - x_0^n)| \\ &\leq \sum_{n=1}^{\infty} |a_n||x - x_0| n\varrho^{n-1} = |x - x_0| s. \end{aligned} \tag{10.3}$$

Um die Stetigkeit in x_0 zu folgern, sei jetzt $\varepsilon > 0$ und $0 < \delta = \min(\varrho - |x_0|, \varepsilon/s)$. Dann folgt für alle $x \in D$ mit $|x - x_0| < \delta$, dass auch $|x| < \delta + |x_0| < \varrho$ gilt und wir deshalb (10.3) anwenden können, sodass auch

$$|f(x) - f(x_0)| \leq s|x - x_0| < s\delta \leq \varepsilon$$

gilt. Wir haben also in jedem $x_0 \in D$ das ε-δ-Kriterium nachgewiesen und so nach Satz 10.2 gezeigt, dass f auf ganz D stetig ist. □

Nachdem wir so schon viele Funktionen kennengelernt haben, die stetig sind, sollten wir uns auch Funktionen ansehen, welche es nicht sind.

Beispiel 10.5. (a) Wir setzen $D = (0, 1]$ und

$$f(x) = \begin{cases} x^2, & 0 < x < \dfrac{1}{2}, \\ 1, \quad \text{falls} & \dfrac{1}{2} \leq x < 1, \\ 7, & x = 1. \end{cases}$$

Aus Satz 10.4 folgt, dass die Funktion f auf den offenen Intervallen, auf denen sie durch eine Potenzreihe darstellbar ist, auch stetig ist. Insbesondere ist sie also als Polynom zweiten Grades stetig auf $(0, \frac{1}{2})$ und als konstante Funktion stetig auf $(\frac{1}{2}, 1)$. Was gilt aber an den Stellen $\frac{1}{2}$ und 1?

- Die Folge $(\frac{1}{2} - \frac{1}{n})_n$ konvergiert gegen $\frac{1}{2}$ und liegt ab $n = 3$ zwischen 0 und $\frac{1}{2}$, also innerhalb von D. Gleichzeitig gilt für die Folge der Funktionswerte

$$\lim_{n\to\infty} f(1/2 - 1/n) = \lim_{n\to\infty} (1/2 - 1/n)^2 = \lim_{n\to\infty} 1/4 - 1/n + 1/n^2 = 1/4 \neq 1 = f(\frac{1}{2}),$$

also ist f nach dem Folgenkriterium aus Definition 10.1 an der Stelle $\frac{1}{2}$ nicht stetig.
- Wir zeigen mit Hilfe von Satz 10.2, dass f in 1 unstetig ist: Wir setzen dazu zum Beispiel $\varepsilon = 1$. Dann gilt für alle $\delta > 0$: ist $x \in D$ und $|x - 1| < \delta$, dann ist wegen $0 \leq f(x) \leq 1$ für alle $x \in D$ jedenfalls

$$|f(x) - f(1)| \geq ||f(x)| - |f(1)|| \geq |1 - 7| = 6 \geq \varepsilon,$$

also das ε-δ-Kriterium verletzt.

(b) Die *Dirichlet-Funktion* $\varphi\colon \mathbb{R} \to \mathbb{R}$ ist gegeben durch

$$\varphi(x) := \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Sie ist in *keinem* Punkt stetig (Übungsaufgabe).

(c) Vorerst noch ein Positiv-Beispiel: Ohne es so zu nennen, haben wir für alle $p \in \mathbb{N}$ die Stetigkeit der Funktion $\sqrt[p]{\cdot}\colon [0,\infty) \to \mathbb{R}$ in Satz 7.15 schon gezeigt. Mit Satz 10.3 folgt daraus auch die Stetigkeit der nicht-negativen rationalen Potenzen $x \mapsto x^{\frac{n}{m}}$ mit $n, m \in \mathbb{N}$.

10.2. Grenzwerte bei Funktionen

Ist $f\colon D \to \mathbb{R}$ eine Funktion und $(x_n)_n \subseteq D$ eine Folge, die gegen $x_0 \in \mathbb{R}$ konvergiert, dann muss der Limes $\lim_{n\to\infty} f(x_n)$ einerseits nicht unbedingt existieren, andererseits kann er existieren, aber einen anderen Wert als $f(x_0)$ annehmen (f ist in x_0 unstetig oder gar nicht erst definiert). Diese Situation wollen wir in diesem Abschnitt genauer auflösen. Zuerst einige Vorberlegungen.

Definition 10.6. *Es sei $D \subseteq \mathbb{R}$ eine Menge. Eine Zahl $x_0 \in \mathbb{R}$ heißt* Häufungspunkt *von D, 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 bekommen, betrachten wir zwei äquivalente Formulierungen dieser Definition. Der Beweis ist eine gute Übungsaufgabe.

Satz 10.7. *Es sei $D \subseteq \mathbb{R}$ und $x_0 \in \mathbb{R}$. Dann sind die folgenden Aussagen äquivalent.*

(a) Die Zahl x_0 ist ein Häufungspunkt von D.

(b) Für alle $\varepsilon > 0$ ist die Menge $D \cap U_\varepsilon(x_0)$ unendlich.

(c) Es existiert eine Folge $(x_n)_n$ in D mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$, die gegen x_0 konvergiert.

Mit Hilfe dieser Umformulierungen kann man sich leicht die folgenden Beispiele veranschaulichen.

Beispiel 10.8. (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.

Hat eine Menge $D \subseteq \mathbb{R}$ einen Häufungspunkt, dann können wir sagen, was es bedeutet, dass eine Funktion $f\colon D \to \mathbb{R}$ an dieser Stelle einen Grenzwert besitzt.

Definition 10.9. *Es sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion sowie $a \in \mathbb{R}$. 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 $(x_n)_n$ in D, die gegen x_0 konvergiert und für die $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, die Folge $(f(x_n))_n$ gegen a konvergiert. Wir schreiben dafür*

$$\lim_{x \to x_0} f(x) = a.$$

Manchmal is es praktisch, zu unterscheiden, ob der Grenzwert nur „von rechts kommend" oder nur „von links kommend" definiert ist.

Definition 10.10. *(a) Ist $D \subseteq \mathbb{R}$ und 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 $(x_n)_n$ in D_+, die gegen x_0 konvergiert, die Folge $(f(x_n))_n$ gegen a konvergiert. Wir schreiben*

$$\lim_{x \to x_0+} f(x) = a.$$

(b) Ist $D \subseteq \mathbb{R}$ und 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 $(x_n)_n$ in D_-, die gegen x_0 konvergiert, die Folge $(f(x_n))_n$ gegen a konvergiert. Wir schreiben*

$$\lim_{x \to x_0-} f(x) = a.$$

Bemerkung 10.11. (a) Eine besondere Betonung beim Lesen dieser Definitionen sollte wie bei der Definition von Stetigkeit 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.

Wichtig ist der Zusammenhang dieser Definitionen zur Stetigkeit:

Übungsaufgabe 10.12. Es sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion sowie $x_0 \in D$ ein Häufungspunkt von D. Dann ist f in x_0 genau dann stetig, wenn $\lim_{x \to x_0} f(x) = f(x_0)$ gilt.

Wir betrachten auch noch einmal das Beispiel 10.5 aus dem vorigen Abschnitt.

Beispiel 10.13. 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 $(x_n)_n$ 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 $(x_n^2)_n$ gegen Null, womit $\lim_{x\to 0} f(x) = 0$ gezeigt ist. Den zweiten Limes bestimmt man analog. Dabei ist zu beachten, dass hier der *Grenzwert* $\lim_{x\to 1} f(x) = 1$ von f an der Stelle 1 nichts mit dem *Wert* $f(1) = 7$ von f an der Stelle 1 zu tun hat, der, anders als beim Häufungspunkt 0, ebenfalls existiert. Wie wir schon gesehen hatten, ist die Funktion in 1 nicht stetig.

- Was ist mit $\lim_{x\to 1/2} f(x)$? Dieser Grenzwert existiert nicht, da die Funktion hier unstetig ist. Das Problem tritt hier insbesondere auf, weil die Annäherung an $1/2$ von links und von rechts im Funktionswert etwas unterschiedliches liefert. Genau dafür haben wir die Begriffe des links- bzw. rechtsseitigen Grenzwertes. Tatsächlich zeigt man wie oben, dass diese existieren und dass gilt
$$\lim_{x\to 1/2+} f(x) = 1 \quad \text{und} \quad \lim_{x\to 1/2-} f(x) = \frac{1}{4}.$$

Wie die Stetigkeit kann auch die Existenz eines Grenzwerts für eine Funktion auf ein ε-δ-Kriterium zurückgeführt werden. Der Beweis des folgenden Satzes funktioniert ganz analog zum Beweis von Satz 10.2.

Satz 10.14. *Es sei $D \subseteq \mathbb{R}$ und x_0 ein Häufungspunkt von D sowie $f : D \to \mathbb{R}$ eine Funktion und $a \in \mathbb{R}$. Dann gilt $\lim_{x\to x_0} f(x) = a$ genau dann, wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $x \in D$ gilt:*

$$|x - x_0| < \delta \quad \Longrightarrow \quad |f(x) - a| < \varepsilon. \tag{10.4}$$

Tatsächlich genügt es, dass die Grenzwerte der Funktionswertfolgen in Definition 10.9 existieren, um zu zeigen, dass sie übereinstimmen.

Satz 10.15. *Es sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und x_0 ein Häufungspunkt von D. Weiter sei für jede Folge $(x_n)_n$ in D, die gegen x_0 konvergiert und für die $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, die Folge $(f(x_n))_n$ konvergent. Dann existiert $\lim_{x\to x_0} f(x)$.*

Beweis. Es seien $(x_n)_n$ und $(y_n)_n$ Folgen in D, welche die Voraussetzungen des Satzes erfüllen. Wir müssen nur zeigen, dass dann

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

gilt. Dazu definieren wir eine Folge $(z_n)_n$ 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 $(f(x_n))_n$ und $(f(y_n))_n$ sind aber Teilfolgen von $(f(z_n))_n$, also konvergieren diese nach Satz 7.33 (b) gegen den selben Grenzwert, d. h.

$$\lim_{n\to\infty} f(x_n) = a = \lim_{n\to\infty} f(y_n). \qquad \square$$

Für Grenzwerte von Funktionen können wir auch wieder Grenzwertsätze formulieren.

Satz 10.16. *Es sei $D \subseteq \mathbb{R}$ und x_0 ein Häufungspunkt von D. Desweiteren seien drei Funktionen $f, g, h : D \to \mathbb{R}$ 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 $b \neq 0$, 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

$$\frac{f}{g} : \big(D \cap U_\delta(x_0)\big) \setminus \{x_0\} \to \mathbb{R} \quad \textit{mit} \quad \frac{f}{g}(x) := \frac{f(x)}{g(x)}$$

definieren. Für diese existiert dann der Limes für x gegen x_0 mit

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)} = \frac{a}{b}.$$

(c) *Gibt es ein $\delta > 0$, so dass für alle $x \in (D \cap U_\delta(x_0)) \setminus \{x_0\}$ die Ungleichung $f(x) \le g(x)$ gilt, so folgt $a \le b$.*

(d) *Gilt $a = b$ und gibt es ein $\delta > 0$ mit*

$$f(x) \le h(x) \le g(x) \quad \text{für alle } x \in (D \cap U_\delta(x_0)) \setminus \{x_0\},$$

so existiert auch der Limes von h für $x \to x_0$ und es gilt $\lim_{x\to x_0} h(x) = a$.

Beweis. Der Beweis dieses Satzes ergibt sich direkt aus den Grenzwertsätzen für Folgen in Satz 7.12. Nur die erste Aussage in (b) wollen wir kurz beweisen.
Es sei $\varepsilon := {}^{|b|}\!/_2 > 0$. Da $\lim_{x\to x_0} g(x) = b$ gilt, existiert dann nach Satz 10.14 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

$$|b| = |b - g(x) + g(x)| \le |b - g(x)| + |g(x)| \le \frac{|b|}{2} + |g(x)|,$$

d. h. $|g(x)| \ge {}^{|b|}\!/_2$. □

Beispiel 10.17. (a) Es gilt

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

Zum Nachweis dieser Aussage überlegen wir uns, dass für $x \neq 0$ gilt

$$\frac{\sin(x)}{x} = \frac{1}{x}\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n+1} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}x^{2n} =: g(x)$$

und die Potenzreihe, die g definiert, hat den Konvergenzradius unendlich (nachrechnen!). 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 $x \mapsto {}^{\sin(x)}\!/_x$.

(b) 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 ∞.

Zum Abschluss dieses Abschnittes können wir aus der Stetigkeit der Funktionen, die durch Potenzreihen gegeben sind, noch einen wichtigen, überraschenden Satz folgern. Er besagt, dass zwei Funktionen, die durch Potenzreihen gegeben sind und die auf einer Nullfolge von Punkten übereinstimmen, schon identisch sein müssen.

Satz 10.18 (Identitätssatz für Potenzreihen)**.** *Gegeben seien zwei Potenzreihen* $\big(\sum_{k=0}^{n} a_k x^k\big)_n$ *und* $\big(\sum_{k=0}^{n} b_k x^k\big)_n$ *mit Konvergenzradien* $r_a > 0$*, bzw.* $r_b > 0$*. Wir setzen* $R := \min\{r_a, r_b\} > 0$ *und*

$$f(x) := \sum_{n=0}^{\infty} a_n x^n, \ |x| < r_a, \quad \textit{und} \quad g(x) := \sum_{n=0}^{\infty} b_n x^n, \ |x| < r_b.$$

Gibt es dann eine Folge $(x_k)_k$ *in* $U_R(0)$ *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 10.4 in 0 stetig ist, also gilt $\lim_{k\to\infty} f(x_k) = f(0) = a_0$. Dasselbe 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 10.4 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 $a_{n+1} = b_{n+1}$. □

10.3. Verhalten von Funktionen im Unendlichen

Wir wollen im Folgenden auch untersuchen, wie sich Funktionen „im Unendlichen", also insbesondere für sehr große (bzw. kleine) x verhalten. Um den dazu erforderlichen „$\lim_{x\to\infty}$" zu definieren, greifen wir auf den Begriff einer bestimmt divergenten Folge aus Definition 7.19 zurück. Dieser ermöglicht es außerdem zu formulieren, dass $f(x)$ für $x \to x_0$ gegen unendlich geht. Wir führen beides in den folgenden zwei Definitionen sauber ein.

Definition 10.19. *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 $(x_n)_n$ in D, die gegen x_0 konvergiert und für die $x_n \neq x_0$ für alle $n \in \mathbb{N}$ gilt, die Folge $(f(x_n))_n$ bestimmt gegen ∞ (bzw. $-\infty$) divergiert.

Definition 10.20. *Es sei $D \subseteq \mathbb{R}$ nicht nach oben (bzw. unten) beschränkt, $f : D \to \mathbb{R}$ eine Funktion und $a \in \mathbb{R} \cup \{\infty, -\infty\}$. 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 $(x_n)_n$ in D, die bestimmt gegen ∞ (bzw. $-\infty$) divergiert, $\lim_{n\to\infty} f(x_n) = a$ gilt.

Beispiel 10.21. (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. Da $x > 0$ ist, gilt unter Weglassen aller Summanden bis auf einen:

$$\mathrm{e}^x = \sum_{n=0}^{\infty}\frac{x^n}{n!} \geq \frac{x^{p+1}}{(p+1)!}.$$

Das liefert

$$\frac{\mathrm{e}^x}{x^p} \geq \frac{x}{(p+1)!},\quad \text{d. h.}\quad \lim_{x\to\infty}\frac{\mathrm{e}^x}{x^p}=\infty.$$

Das bedeutet, dass die Exponentialfunktion schneller wächst als jede Potenz. Insbesondere gilt also

$$\lim_{x\to\infty}\mathrm{e}^x=\infty.$$

Wir untersuchen noch das Verhalten für x gegen minus unendlich. Sei dazu $(x_n)_n$ eine bestimmt gegen $-\infty$ divergente Folge. Dann divergiert die Folge $(-x_n)_n$ bestimmt gegen ∞. Da für jedes $n \in \mathbb{N}$

$$\mathrm{e}^{x_n}=\frac{1}{\mathrm{e}^{-x_n}}$$

gilt, und die Folge $(\mathrm{e}^{-x_n})_n$ nach unseren obigen Überlegungen bestimmt nach ∞ divergiert, folgt mit Hilfe der Übungsaufgabe 7.20 sofort

$$\lim_{x\to-\infty}\mathrm{e}^x=0.$$

10.4. 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.

Satz 10.22 (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 $f(x_0) = y_0$.*

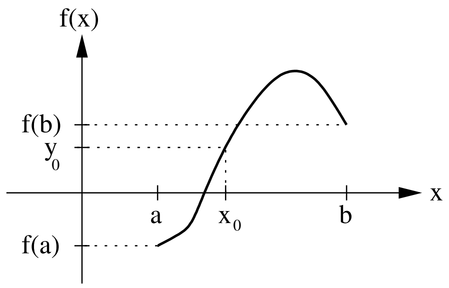

Abbildung 10.1.: Der Zwischenwertsatz

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

$$M := \{x \in [a, b] : f(x) \leq y_0\}.$$

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. In einem nächsten Schritt existiert für jedes $n \in \mathbb{N}$ nach Satz 4.17 ein $x_n \in M$ mit

$$x_0 - \frac{1}{n} < x_n \leq x_0.$$

Die so entstandene Folge $(x_n)_n$ konvergiert nach dem Sandwich-Theorem 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, dass die Folge $(f(x_n))_n$ 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 $f(x_0) = y_0$. $\square$

Eine wichtige Folgerung aus diesem Satz ist die folgende.

Satz 10.23 (Nullstellensatz von Bolzano)**.** *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f \in \mathrm{C}([a, b], \mathbb{R})$ 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 10.22. □

Wir können dieses Ergebnis nun anwenden, um das Bild der reellen Exponentialfunktion zu bestimmen.

Beispiel 10.24. Wir zeigen, dass für die Exponentialfunktion gilt:

$$\exp(\mathbb{R}) = \{\mathrm{e}^x : x \in \mathbb{R}\} = (0, \infty).$$

Wir wissen schon (vgl. Satz 8.40 (c)), dass $\mathrm{e}^x > 0$ für alle $x \in \mathbb{R}$ gilt, d. h. es ist $\exp(\mathbb{R}) \subseteq (0, \infty)$. Es bleibt die umgekehrte Inklusion zu zeigen.
Sei dazu $y_0 \in (0, \infty)$. Wir wissen aus Beispiel 10.21 (b), dass

$$\lim_{x \to -\infty} \mathrm{e}^x = 0 \quad \text{und} \quad \lim_{x \to \infty} \mathrm{e}^x = \infty$$

gilt. Also gibt es ein $a \in \mathbb{R}$, so dass $\mathrm{e}^a < y_0$ gilt und ein $b \in \mathbb{R}$ mit $\mathrm{e}^b > y_0$. Da damit zwangsläufig $\mathrm{e}^a < \mathrm{e}^b$ gilt, muss wegen Satz 8.40 (e) auch $a < b$ gelten. Da die Exponentialfunktion nach Satz 10.4 auch auf ganz $\mathbb{R}$ stetig ist, sind nun alle Voraussetzungen von Satz 10.22 erfüllt. Es gibt also ein $x_0 \in (a, b)$ mit $\mathrm{e}^{x_0} = y_0$. Damit ist $y_0 \in \exp(\mathbb{R})$ und wir sind fertig.

Wir führen nun schnell den folgenden intuitiv klaren Begriff ein.

Definition 10.25. *Es sei $D \subseteq \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ 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.*

x Wir kommen nun zu der Eigenschaft stetiger Funktionen, in bestimmten Situationen immer ein Maximum und ein Minimum zu besitzen. Dazu führen wir zuerst den für die Anlaysis fundamentalen Begriff der *Kompaktheit* ein.

Definition 10.26. *(a) Eine Teilmenge D von $\mathbb{R}$ heißt* abgeschlossen*, wenn für jede Folge $(x_n)_n$ in D, die in $\mathbb{R}$ konvergiert, $\lim_{n\to\infty} x_n \in D$ gilt.*

(b) Eine Teilmenge von $\mathbb{R}$ heißt kompakt*, wenn sie abgeschlossen und beschränkt ist.*

Wir betrachten jetzt stetige Funktionen auf kompakten Mengen.

Satz 10.27. *(Satz von Weierstraß) Es sei $K \subseteq \mathbb{R}$ kompakt und nicht-leer sowie $f : K \to \mathbb{R}$ stetig. 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 und es existieren $\max f(K) = f(x^*)$ und $\min f(K) = f(x_*)$.

Dass dabei jede der Voraussetzungen des Satzes zwingend nötig ist, veranschaulichen die folgenden Beispiele.

Beispiel 10.28. (a) Ist $K = [0, 1]$ und

$$f(x) = \begin{cases} x, \text{ falls } x \in (0,1), \\ \frac{1}{2}, \text{ falls } x = 0 \text{ oder } x = 1, \end{cases}$$

so ist $f(K) = (0, 1)$, d. h. es gibt keine $x_*, x^* \in [0, 1]$ mit $f(x_*) = 0$ und $f(x^*) = 1$. Ohne die Stetigkeit von f kann der Satz also schiefgehen.

(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 nach oben beschränkt, da $\lim_{x \to 0+} f(x) = \infty$ ist. Ohne die Abgeschlossenheit von K kann der Satz also schiefgehen.

(c) Ist schließlich $K = \mathbb{R}$ und $f(x) = \mathrm{e}^x$, so ist diese Funktion wieder auf ganz K stetig und K ist abgeschlossen, aber f nicht beschränkt. Ohne die Beschränktheit von K kann der Satz also ebenfalls schiefgehen.

Beweis von Satz 10.27. 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 7.35) können wir nun aus der Folge $(x_n)_n$ eine konvergente Teilfolge $(x_{n_k})_k$ auswählen. Da K außerdem abgeschlossen 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 $(f(x_{n_k})k$ gegen $f(x_0)$ konvergiert. Insbesondere ist damit die Folge $(f(x_{n_k}))_k$ beschränkt, was im Widerspruch zur Konstruktion der Folge $(x_n)_n$ 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 es nach Satz 4.17 für jedes $n \in \mathbb{N}$ ein $y_n \in f(K)$ gibt mit $S - 1/n < y_n \leq S$. Dazu finden wir jeweils ein $x_n \in K$ mit $f(x_n) = y_n$, so dass zusammengenommen

$$S - \frac{1}{n} < f(x_n) \leq S \quad \text{für alle } n \in \mathbb{N} \tag{10.5}$$

gilt. Die so gewonnene Folge $(x_n)_n$ enthält wie jede Folge in K eine konvergente Teilfolge $(x_{n_k})_k$. 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 (10.5) 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.
Zur Vorbereitung definieren wir dazu erst einmal die ganz natürlichen Begriffe für die Monotonie von allgemeinen reellen Funktionen:

Definition 10.29. *Sei $D \subseteq \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ heißt*

(a) monoton wachsend, *falls für alle $x, y \in D$ mit $x \le y$ gilt, dass $f(x) \le f(y)$ ist.*

(b) monoton fallend, *falls für alle $x, y \in D$ mit $x \le y$ gilt, dass $f(x) \ge f(y)$ ist.*

(c) streng monoton wachsend, *falls für alle $x, y \in D$ mit $x < y$ gilt, dass $f(x) < f(y)$ ist.*

(d) streng monoton fallend, *falls für alle $x, y \in D$ mit $x < y$ gilt, dass $f(x) > f(y)$ ist.*

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

Monotonie ist oft ein praktisches Hilfsmittel, um Bijektivität zu zeigen. Es ist zum Beispiel klar, dass jede streng monotone Funktion $f : D \to \mathbb{R}$ injektiv ist (Übungsaufgabe!). Damit ist nach Einschränkung des Wertebereichs die Funktion $f : D \to f(D)$ bijektiv und die Umkehrfunktion $f^{-1} : f(D) \to D$ existiert in diesem Fall. Wir wissen über f^{-1} sogar noch mehr.

Lemma 10.30. *Sei $D \subseteq \mathbb{R}$ und $f : D \to f(D)$ eine streng monoton wachsende (bzw. fallende) Funktion. Dann ist die Umkehrfunktion $f^{-1} : f(D) \to D$ 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 $y_1, y_2 \in f(D)$ mit $y_1 < y_2$ gegeben. Dann existieren $x_1, x_2 \in D$, 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 10.31. *Eine nicht leere und nicht einelementige 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.*

Jetzt wenden wir uns wieder stetigen Funktionen zu.

Satz 10.32. *Es sei $I \subseteq \mathbb{R}$ ein Intervall (zum Beispiel ist auch $I = \mathbb{R}$ zugelassen) und $f \in \mathrm{C}(I, \mathbb{R})$ sei streng monoton. Dann ist $f(I)$ ebenfalls ein Intervall und es gilt $f^{-1} \in \mathrm{C}(f(I), \mathbb{R})$.*

Beweis. Wir führen den Beweis wieder nur für streng wachsende Funktionen f.
Wir verwenden zunächst Lemma 10.31, um zu zeigen, dass $f(I)$ ein Intervall ist. Wir wissen dabei, dass $f(I)$ nicht leer und nicht einelementig ist, weil I mehr als zwei verschiedene Werte enthält und f streng monoton ist. Seien also $a, b \in f(I)$ 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 10.22 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} auf $f(I)$ stetig ist. Sei dazu $y_0 \in f(I)$ und $(y_n)_n$ eine Folge in $f(I)$, die gegen y_0 konvergiert. Für diese ist zu zeigen, dass die Folge $(x_n)_n := (f^{-1}(y_n))_n$ gegen $x_0 := f^{-1}(y_0)$ konvergiert. Wir nehmen dazu an, das wäre nicht so. Dann gibt es ein $\varepsilon_0 > 0$, für das die Menge $M := \{n \in \mathbb{N} : x_n \notin U_{\varepsilon_0}(x_0)\}$ unendlich viele Elemente besitzt. Das bedeutet, dass es entweder

Fall 1 unendlich viele $n \in \mathbb{N}$, und somit eine Teilfolge $(x_{n_k})_k$ von $(x_n)_n$ gibt, sodass

$$x_{n_k} = f^{-1}(y_{n_k}) \geq f^{-1}(y_0) + \varepsilon_0,$$

gilt, oder, dass es

Fall 2 unendlich viele $n \in \mathbb{N}$, und somit eine Teilfolge $(x_{n_j})_j$ von $(x_n)_n$ gibt, sodass

$$x_{n_j} = f^{-1}(y_{n_j}) \leq f^{-1}(y_0) - \varepsilon_0,$$

gilt, oder, dass sogar beide Teilfolgen existieren.

Im Fall 1 folgt aus der Monotonie von f für alle $k \in \mathbb{N}$ die Abschätzung

$$y_{n_k} = f(f^{-1}(y_{n_k})) \geq f(f^{-1}(y_0) + \varepsilon_0),$$

die sich also auch durch den Grenzwert $k \to \infty$ durchzieht und zusammen mit der strengen Monotonie dann zum Widerspruch führt:

$$y_0 = \lim_{k\to\infty} y_{n_k} \geq f(f^{-1}(y_0) + \varepsilon_0) > f(f^{-1}(y_0)) = y_0.$$

Analog gilt im Fall 2 wegen der Monotonie von f,

$$y_{n_j} = f(f^{-1}(y_{n_k})) \leq f(f^{-1}(y_0 - \varepsilon_0)),$$

für jedes $j \in \mathbb{N}$ und im Grenzwert $j \to \infty$ folgt wegen der strengen Monotonie der Widerspruch

$$y_0 = \lim_{j\to\infty} y_{n_j} \leq f(f^{-1}(y_0 - \varepsilon_0)) < f(f^{-1}(y_0)) = y_0.$$

□

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 8.40 (e)). Außerdem haben wir in Beispiel 10.24 gesehen, dass $\exp(\mathbb{R}) = (0, \infty)$ gilt. Damit wissen wir, dass die Umkehrfunktion existiert. Sie bekommt einen eigenen Namen.

Definition 10.33. *Die Funktion*

$$\ln := \log := \exp^{-1} : (0,\infty) \to \mathbb{R} \quad \textit{mit} \quad \ln(x) := \log(x) := \exp^{-1}(x),\ x \in (0,\infty),$$

heißt (natürlicher) Logarithmus.

Der Logarithmus hat die folgenden Eigenschaften.

Satz 10.34. *(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) $\lim_{x\to\infty} \ln(x) = \infty$ *und* $\lim_{x\to 0+} \ln(x) = -\infty$.

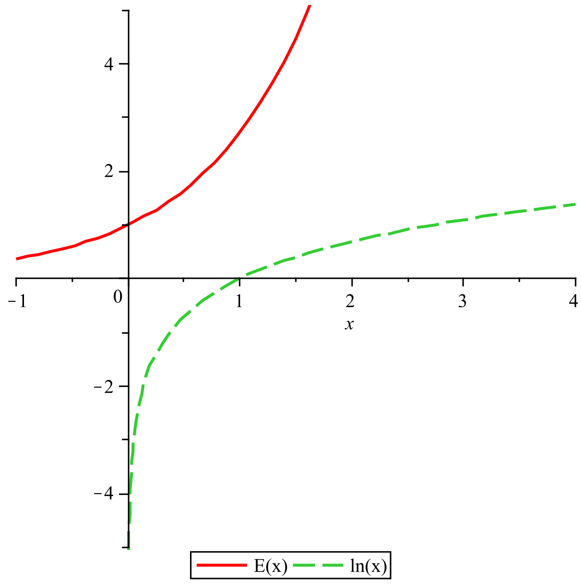

Abbildung 10.2.: Die Graphen der Exponentialfunktion und des Logarithmus

(d) Für alle $x, y \in (0, \infty)$ und alle $r \in \mathbb{Q}$ gilt

$$\ln(xy) = \ln(x) + \ln(y), \qquad \ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y) \quad \textit{und} \quad \ln(x^r) = r\ln(x).$$

Beweis. (a) Ergibt sich sofort aus Satz 10.32.

(b) Ergibt sich aus Satz 8.40 (a).

(c) Ergibt sich aus Beispiel 10.21 (b).

(d) Es gilt nach Satz 8.40 (b)

$$\mathrm{e}^{\ln(x)+\ln(y)} = \mathrm{e}^{\ln(x)}\mathrm{e}^{\ln(y)} = xy.$$

Also ist

$$\ln(xy) = \ln\left(\mathrm{e}^{\ln(x)+\ln(y)}\right) = \ln(x) + \ln(y).$$

Weiter gilt mit Satz 8.40 (c)

$$\mathrm{e}^{-\ln(1/x)} = \frac{1}{\mathrm{e}^{\ln(1/x)}} = \frac{1}{1/x} = x.$$

Also ist

$$\ln(x) = \ln(\mathrm{e}^{-\ln(1/x)}) = -\ln\left(\frac{1}{x}\right). \tag{10.6}$$

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. Mit den Rechenregeln für exp gilt für alle $r \in \mathbb{Q}$,

$$\exp(\ln x^r) = x^r = (\exp(\ln x))^r = \exp(r \ln x).$$

Wegen der Bijektivität von exp folgt aus der Gleichheit $\exp A = \exp B$ für alle $A, B \in \mathbb{R}$ auch die Gleichheit $A = B$, also hier die Behauptung. □

Sehen wir uns die dritte Rechenregel in (d) noch einmal an, so folgt daraus insbesondere für alle $a \in (0,\infty)$ und alle $r \in \mathbb{Q}$ die Beziehung $a^r = \mathrm{e}^{\ln(a^r)} = \mathrm{e}^{r\ln(a)}$. Diese verwenden wir nun um die allgemeine Potenzfunktion zu definieren.

Definition 10.35. *Für alle $a \in (0,\infty)$ und alle $x \in \mathbb{R}$ definieren wir die* allgemeine Potenz *durch*

$$a^x := \mathrm{e}^{x\cdot\ln(a)}.$$

Wir sammeln Eigenschaften dieser Funktion.

Satz 10.36. *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 exp jeweils auf $\mathbb{R}$ stetig sind, also ist auch die Potenzfunktion als deren Verkettung nach Satz 10.3(c) stetig.
Die Rechenregeln lassen sich alle direkt aus jenen für die Exponentialfunktion ableiten. Wir behandeln deshalb hier nur beispielhaft

$$a^{x+y} = \mathrm{e}^{(x+y)\ln(a)} = \mathrm{e}^{x\ln(a)+y\ln(a)} = \mathrm{e}^{x\ln(a)}\mathrm{e}^{y\ln(a)} = a^x a^y. \qquad \square$$

Bemerkung 10.37. Zu Beginn der Vorlesung haben wir die rationalen Potenzen a^q mit $q \in \mathbb{Q}$ definiert und uns gefragt, wie die reellen Potenzen a^{x_0} mit $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ definiert werden können. Dieses Problem ist jetzt mit Hilfe der Exponentialreihe und des Logarithmus auf vielleicht etwas überraschende Weise gelöst. Aus Satz 10.36 folgt, dass die so definierte Funktion

$$a^{\cdot} : \mathbb{R} \to (0,\infty), x \mapsto a^x$$

zur (natürlichen) Definition der rationalen Potenzen passt, insofern sie die *stetige Fortsetzung* (s. Beispiel 10.17) dieser Funktion von $\mathbb{Q}$ auf $\mathbb{R}$ darstellt (Übung).

10.5. Funktionenfolgen

In diesem Abschnitt betrachten wir Folgen, deren Glieder durch Funktionen gegeben sind. Ein spezielles Beispiel für diese Situation haben wir mit den Potenzreihen schon gesehen. Wir beginnen mit den folgenden Begriffen.

Definition 10.38. *Es sei $D \subseteq \mathbb{R}$ und für jedes $n \in \mathbb{N}$ sei eine Funktion $f_n \colon D \to \mathbb{R}$ gegeben. Wir sagen, die* Funktionenfolge $(f_n)_n$ konvergiert punktweise, *wenn für jedes $x \in D$ die Folge $(f_n(x))_n$ in $\mathbb{R}$ konvergiert. In diesem Fall heißt die Funktion*

$$f\colon \begin{cases} D \to \mathbb{R} \\ x \mapsto \lim\limits_{n\to\infty} f_n(x) \end{cases}$$

die Grenzfunktion *von $(f_n)_n$.*

Beispiel 10.39. (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 7.18 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 $(f_n)_n$ punktweise gegen die Grenzfunktion $f : [0, 1] \to \mathbb{R}$ mit

$$f(x) := \begin{cases} 0, \text{ falls } x \in [0, 1), \\ 1, \text{ falls } x = 1. \end{cases}$$

(b) Es sei $(a_n)_n$ eine Folge in $\mathbb{R}$ und für jedes $n \in \mathbb{N}_0$ seien die durch Summation gebildeten Polynome

$$p_n(x) = \sum_{k=0}^{n} a_k x^k$$

gegeben. Sie bilden als Funktionenfolge genau die durch die Folge $(a_n)_n$ gegebene Potenzreihe. Wie wir gesehen haben, konvergiert sie, wenn der Konvergenzradius r der Potenzreihe positiv ist, innerhalb des Intervalls $(-r, r)$ punktweise gegen die Funktion $s : (-r, r) \to \mathbb{R}$ mit

$$s(x) = \sum_{n=1}^{\infty} a_n x^n.$$

(c) Es sei $D := [0, \infty)$ und für jedes $n \in \mathbb{N}$ sei $f_n : [0, \infty) \to \mathbb{R}$ gegeben durch

$$f_n(x) := \frac{nx}{1 + n^2x^2}, \quad x \in [0, \infty).$$

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 $(f_n)_n$ in diesem Beispiel auf $[0, \infty)$ punktweise gegen die konstante Nullfunktion $f = 0$.

Bemerkung 10.40. (a) In Epsilons ausgedrückt bedeutet punktweise Konvergenz einer Funktionenfolge $(f_n)_n$ auf einer Menge $D \subseteq \mathbb{R}$:

$$\forall \varepsilon > 0 \; \forall x \in D \; \exists n_0 = n_0(\varepsilon, x) \in \mathbb{N} \; \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon. \qquad (10.7)$$

(b) Schauen wir uns noch einmal unser drittes Beispiel von oben an, vgl. Abbildung 10.3, 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

$$\left| f_n\Big(\frac{1}{n}\Big) - f\Big(\frac{1}{n}\Big) \right| = \left| \frac{n\frac{1}{n}}{1 + n^2(\frac{1}{n})^2} - 0 \right| = \left| \frac{1}{1+1} \right| = \frac{1}{2}. \qquad (10.8)$$

Wir wollen im Folgenden einen weiteren, restriktiveren Konvergenzbegriff einführen, der solch ein Konvergenzverhalten nicht mehr „toleriert“.

Definition 10.41. *Es sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ und für jedes $n \in \mathbb{N}$ seien Funktionen $f_n : D \to \mathbb{R}$ gegeben.*

(a) Die Funktionenfolge $(f_n)_n$ konvergiert gleichmäßig auf *D gegen f, falls das Kriterium*

$$(\forall \varepsilon > 0) \; (\exists n_0 = n_0(\varepsilon) \in \mathbb{N}) \; (\forall n \geq n_0) \; (\forall x \in D) : |f_n(x) - f(x)| < \varepsilon.$$

erfüllt ist.

(b) Die Funktionenfolge $(f_n)_n$ konvergiert lokal gleichmäßig *auf D gegen f, wenn für jedes $K \subseteq D$ kompakt ihre Einschränkungen auf K gleichmäßig gegen f konvergieren.*

Bemerkung 10.42. (a) Vergleichen wir das Kriterium für gleichmäßige Konvergenz mit der entsprechenden Definition der punktweisen Konvergenz aus (10.7), 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

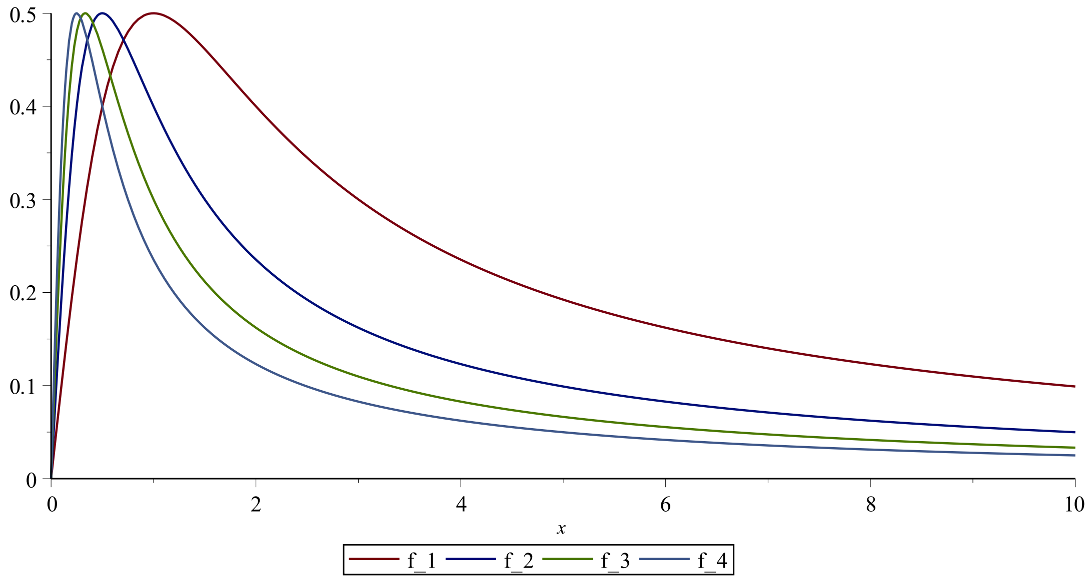

Abbildung 10.3.: Die Graphen der ersten vier Funktionen in der Funktionenfolge aus Beispiel 10.39 (c).

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 10.4.

Betrachten wir wieder das Beispiel 10.39 (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 10.43. Wir betrachten im Lichte der neuen Definition noch einmal (a) und (c) aus Beispiel 10.39. Beide Funktionenfolgen sind nicht gleichmäßig konvergent.

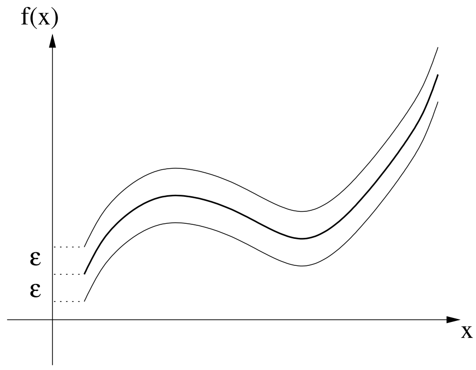

Abbildung 10.4.: Der ε-Streifen um den Graphen der Grenzfunktion f.

Für das Beispiel

$$f_n(x) = \frac{nx}{1+n^2x^2}, \quad x \in [0,\infty),$$

folgt das direkt aus (10.8), 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 10.44. (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 $|f_n(x) - f(x)| = |x^n - 0| = x^n \leq \alpha^n$. (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 $n \geq n_0$ ist, da $\lim_{n\to\infty} \alpha^n = 0$ ist. Also gilt

$$|f_n(x) - f(x)| \leq \alpha^n < \varepsilon \text{ für alle } n \geq n_0 \text{ und alle } x \in \tilde{D}$$

und das ist genau die Bedingung für gleichmäßige Konvergenz.

Zusammengefasst ist die Funktionenfolge $(x^n)_n$ 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.

Jede kompakte Teilmenge von $[0, 1)$ ist in einem Intervall der Form $[0, \alpha]$ mit $\alpha \in (0, 1)$ enthalten. Also haben wir lokal gleichmäßige Konvergenz auf $[0, 1)$. Achtung: Diese Funktionenfolge konvergiert nicht lokal gleichmäßig auf $[0, 1]$! Warum?

(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 $n \in \mathbb{N}$
$$|f_n(x) - f(x)| = \frac{nx}{1 + n^2x^2} \leq \frac{nx}{n^2x^2} = \frac{1}{nx} \leq \frac{1}{n\alpha}.$$
Für jedes $\varepsilon > 0$ gibt es nun ein $n_0 \in \mathbb{N}$ mit $1/(n\alpha) < \varepsilon$ für alle $n \geq n_0$, also gilt
$$|f_n(x) - f(x)| \leq \frac{1}{n\alpha} < \varepsilon \text{ für alle } n \geq n_0 \text{ und alle } x \in \tilde{D}.$$
Zusammenfassend ist diese Funktionenfolge also auf jedem Intervall der Form $[\alpha, \infty)$ für $\alpha > 0$ gleichmäßig konvergent, aber nicht auf $[0, \infty)$. Lokal gleichmäßige Konvergenz haben wir in diesem Fall auf $(0, \infty)$, aber ebenfalls nicht auf $[0, \infty)$.

Übungsaufgabe 10.45. Es sei $D \subseteq \mathbb{R}$ und $(f_n)_n$ eine Funktionenfolge auf D.

(a) Ist $(f_n)_n$ gleichmäßig konvergent gegen $f : D \to \mathbb{R}$, so konvergiert auch die Folge $(|f_n|)_n$ gleichmäßig auf D und zwar gegen $|f|$.

(b) Die Funktionenfolge $(f_n)_n$ konvergiert genau dann gleichmäßig gegen $f : D \to \mathbb{R}$, 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 $\lim_{n\to\infty} \sup_{x\in D} |f_n(x)| = \sup_{x\in D} |f(x)|$.

(c) Sei $f : D \to \mathbb{R}$ eine Funktion. Gibt es eine Nullfolge $(\alpha_n)_n$, 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 $(f_n)_n$ gleichmäßig auf D gegen f.

In diesem Abschnitt bleiben nach zwei Resultate zu erzielen. Zuerst bestätigen wir uns nun erneut, dass Potenzreihen etwas Freundliches sind und beweisen, dass diese (als Folgen von Polynomfunktionen aufgefasst) auf ihrem Konvergenzbereich sogar lokal gleichmäßig konvergieren. Zuletzt beschäftigen wir uns mit dem gleichmäßigen Limes *stetiger* Funktionen.

Satz 10.46. *Es sei* $\left(\sum_{k=1}^{n} a_k x^k\right)_n$ *eine Potenzreihe mit Konvergenzradius* $r > 0$. *Dann konvergiert sie als Folge von Funktionen lokal gleichmäßig auf* $U_r(0)$.

Beweis. Es sei $K \subseteq U_r(0)$ kompakt. Da die Betragsfunktion stetig und K kompakt ist, gibt es nach Satz 10.27 ein $x_0 \in K$ mit $|x_0| = \max\{|x| : x \in K\}$. Weiter ist $K \subseteq U_r(0)$, also bekommen wir

$$\varrho := \max\{|x| : x \in K\} = |x_0| < r.$$

Damit gilt $|x| \leq \varrho < r$ für alle $x \in K$. Für alle $\varepsilon > 0$ wählen wir jetzt $n_0 \in \mathbb{N}$, sodass $\sum_{k=n_0}^{\infty} |a_k|\rho^k < \varepsilon$ gilt. Das ist möglich, da die Potenzreihe ja innerhalb des Konvergenzradius, also auch in ρ, absolut konvergent ist und der zugehörige Reihenrest $\left(\sum_{k=n}^{\infty} |a_k|\rho^k\right)_n$ deshalb gegen 0 konvergieren muss. Es folgt für alle $x \in K$ und $n \geq n_0$,

$$|\sum_{k=1}^{n} a_k x^k - \sum_{k=1}^{\infty} a_k x^k| = |\sum_{k=n}^{\infty} a_k x^k| \leq \sum_{k=n_0}^{\infty} |a_k|\rho^k < \varepsilon,$$

also die gleichmäßige Konvergenz der Potenzreihe gegen ihren punktweisen Grenzwert $x \mapsto \sum_{k=1}^{\infty} a_k x^k$ auf K.

□

Um zu sehen, dass eine Potenzreihe im Allgemeinen nicht gleichmäßig auf dem vollen Konvergenzintervall konvergiert, kann das folgende Beispiel dienen.

Übungsaufgabe 10.47. Die geometrische Reihe $\left(\sum_{k=0}^{n} x^k\right)_n$ ist auf $(-1, 1)$ nicht gleichmäßig konvergent.

Wir zeigen nun, dass gleichmäßige Konvergenz Stetigkeit erhält, eine sehr wichtige Konsequenz dieser Eigenschaft.

Satz 10.48. *Es sei* $D \subseteq \mathbb{R}$ *und* $(f_n)_n$ *sei eine Funktionenfolge, die auf* D *lokal gleichmäßig gegen eine Funktion* f *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 *in* x_0 *stetig.*

Beweis. Sei $(x_k)_k$ eine Folge in D, die gegen x_0 konvergiert. Wir müssen zeigen, dass $\lim_{k\to\infty} f(x_k) = f(x_0)$ gilt.
Sei $\varepsilon > 0$. Wir betrachten $\{x_k : k \in \mathbb{N}_0\}$, d. h. die Menge aller Folgenglieder von $(x_k)_k$ zusammen mit dem Grenzwert. Dank der Konvergenz von $(x_k)_k$ ist das eine abgeschlossene und beschränkte und damit kompakte Teilmenge von D. Wegen der lokal gleichmäßigen Konvergenz von $(f_n)_n$ gibt es also ein $m \in \mathbb{N}$, sodass

$$|f_m(x_k) - f(x_k)| < \frac{\varepsilon}{3} \text{ für alle } k \in \mathbb{N}_0$$

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. Sei nun k_0 so gewählt, dass $|x_k - x_0| < \delta$ für alle $k \geq k_0$ ist. Dann gilt für alle $k \geq k_0$ durch Kombination dieser Überlegungen

$$\begin{aligned} |f(x_k) - f(x_0)| &= \big|f(x_k) - f_m(x_k) + f_m(x_k) - f_m(x_0) + f_m(x_0) - f(x_0)\big| \\ &\leq \big|f(x_k) - f_m(x_k)\big| + \big|f_m(x_k) - f_m(x_0)\big| + \big|f_m(x_0) - f(x_0)\big| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$ □

Bemerkung 10.49. (a) Wir können Satz 10.48 auch folgendermaßen formulieren: Konvergiert eine Funktionenfolge $(f_n)_n$ auf D lokal 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 7.24, und insofern sind Sätze dieser Art, die ein Vertauschen gestatten, sehr wertvoll.

Tatsächlich ist diese Vertauscherei für nur punktweise Konvergenz im Allgemeinen falsch. Beispielsweise gilt für unsere Funktionenfolge aus Beispiel 10.39 (a)

$$\lim_{x \to 1-} \lim_{n \to \infty} x^n = \lim_{x \to 1-} 0 = 0 \neq 1 = \lim_{n \to \infty} 1 = \lim_{n \to \infty} \lim_{x \to 1-} x^n.$$

(b) Satz 10.48 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 (bei Reihen zum Beispiel alle Summanden) stetige Funktionen, aber die punktweise Grenzfunktion ist unstetig, so kann die Konvergenz nach diesem Satz nicht gleichmäßig sein.

10.6. 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 $f : D \to \mathbb{R}$ genau dann stetig, wenn

$$(\forall \varepsilon > 0)(\forall x_0 \in D)(\exists \delta = \delta(\varepsilon, x_0) > 0)(\forall x \in D)\, |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Das zu bestimmende δ darf hierbei außer von ε, von dem es logischerweise (in der Regel) 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.

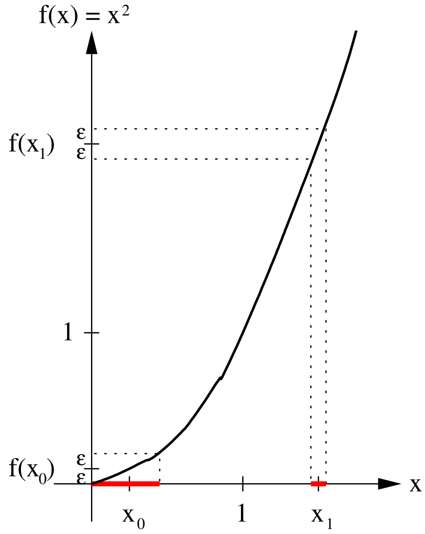

Abbildung 10.5.: Die Abhängigkeit des Stetigkeits-Deltas von x_0

Beispiel 10.50. Es sei $D = [0, \infty)$ und $f(x) = x^2$, $x \in D$, vgl. Abbildung 10.5. 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

$$\varepsilon > |x^2 - x_0^2| = |x + x_0||x - x_0| = \left(2x_0 + \frac{\delta}{2}\right)\frac{\delta}{2} = \delta x_0 + \frac{\delta^2}{4}.$$

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 schneller 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 10.51. *Es sei $D \subseteq \mathbb{R}$. Dann heißt $f : D \to \mathbb{R}$* gleichmäßig stetig *auf D, falls gilt:*

$$(\forall \varepsilon > 0)\,(\exists \delta = \delta(\varepsilon) > 0)\,(\forall x_0, x \in D)\ |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Bemerkung 10.52. (a) Es ist klar, dass eine Funktion, die auf einer Menge D gleichmäßig stetig ist, auch auf dieser Menge stetig ist, also zu $\mathrm{C}(D, \mathbb{R})$ gehört. Die Umkehrung gilt i.A. nicht, wie Beispiel 10.50 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 anders als bei reiner Stetigkeit nicht sinnvoll, von gleichmäßiger Stetigkeit „in einem Punkt" zu sprechen.

Wir wollen nun einen Fall von Mengen behandeln, auf denen Stetigkeit und gleichmäßige Stetigkeit zusammenfallen.

Satz 10.53. *Ist $K \subseteq \mathbb{R}$ kompakt und $f \in \mathrm{C}(K, \mathbb{R})$, 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 wieder 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 $(x_n)_n$ eine Folge in der kompakten Menge K, d. h. sie ist insbesondere beschränkt. Nach dem Satz von Bolzano-Weierstraß 7.35 besitzt sie damit eine konvergente Teilfolge $(x_{n_k})_k$ und es gilt $x_0 := \lim_{k\to\infty} x_{n_k} \in K$, da K abgeschlossen ist. Betrachten wir die Folge $(y_{n_k})_k$, 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 $(y_{n_k})_k$ 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 10.54. *Es sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. Diese heißt* Lipschitz-stetig*, falls eine Konstante $L > 0$ existiert, so dass*

$$|f(x) - f(y)| \leq L|x - y|$$

für alle $x, y \in D$ gilt.

Bemerkung 10.55. 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 10.56. Wir setzen $D = [0, 1]$ und $f(x) = \sqrt{x}$. Dann ist f nach Satz 7.15 stetig und nach Satz 10.53 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.

11. 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.

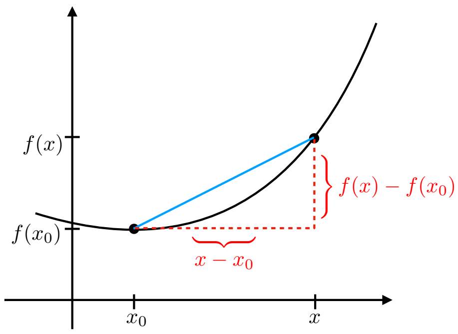

Abbildung 11.1.: Der Differenzenquotient $\frac{f(x)-f(x_0)}{x-x_0}$ entspricht der Steigung der blauen Sekante. Für $x \to x_0$ definiert er die Steigung des Graphen bzw. die Ableitung der Funktion in x_0.

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 11.1. *(a) Es sei $x_0 \in I$. Eine Funktion $f : I \to \mathbb{R}$ heißt* differenzierbar *in x_0, wenn der Grenzwert*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. In diesem Fall heißt dieser Grenzwert die Ableitung von f in x_0 *und wird mit $f'(x_0)$ bezeichnet.*

(b) Eine Funktion $f : I \to \mathbb{R}$ 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 $f' : I \to \mathbb{R}$ definiert. Diese Funktion heißt die Ableitung *oder auch* Ableitungsfunktion von f auf I.

Bemerkung 11.2. 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 11.3. (a) Es sei zunächst $f(x) = c \in \mathbb{R}$ 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 behandeln in diesem Abschnitt noch eine Umformulierung von Differenzierbarkeit. Sie besagt anschaulich, dass eine differenzierbare Funktion affin linear angenähert werden kann (mit Hilfe ihrer Tangente), und dass dabei lokal nur ein kleiner Fehler passiert.

Satz 11.4. *Eine Funktion $f : I \to \mathbb{R}$ ist in $x_0 \in I$ genau dann differenzierbar mit $f'(x_0) = a$, wenn sie sich für $x \in I$ durch*

$$f(x) = f(x_0) + a(x - x_0) + r(x),$$

darstellen lässt, wobei für die Funktion $r : I \to \mathbb{R}$ gilt:

$$\lim_{x \to x_0} \frac{r(x)}{|x - x_0|} = 0. \tag{11.1}$$

Beweis. Der Beweis ist ganz direkt, wenn wir uns klarmachen, wie sich die Funktion r zum Differenzenquotienten verhalten muss. Zu festem $x_0 \in I$ betrachten wir

$$r(x) := f(x) - f(x_0) - a(x - x_0),$$

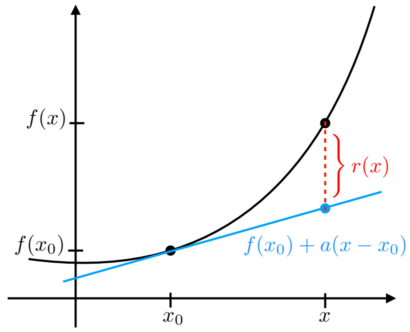

Abbildung 11.2.: Die affin lineare Funktion $x \mapsto f(x_0) + a(x - x_0)$ mit Steigung $a = f'(x_0)$ approximiert die differenzierbare Funktion f nahe x_0. Der Fehler $r(x)$, der dabei im Punkt x passiert, verschwindet und ist stetig in x_0.

dann existiert $a = \lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ genau dann, wenn gilt

$$\frac{r(x)}{|x - x_0|} = \left(\frac{f(x) - f(x_0)}{x - x_0} - a \right) \cdot \frac{x - x_0}{|x - x_0|} \xrightarrow{x \to x_0} 0.$$

$\square$

Die Eigenschaft 11.1 sagt insbesondere, dass der Fehler oder Rest r, der hier bei der affinen Approximation von f bleibt, in x_0 stetig ist und nahe x_0 langsamer als linear anwächst. Zum Beispiel in der Physik wird das sehr oft benutzt: Differenzierbare Funktionen werden durch „starkes Ranzoomen“ oder „lokales Hinschauen“ praktisch linear.
Wir haben in Beispiel 11.3(b) gesehen, dass es stetige Funktionen gibt, die nicht differenzierbar sind. Dass aber umgekehrt jede differenzierbare Funktion notwendigerweise stetig ist, ist jetzt eine direkte Folgerung aus der Darstellung in Satz 11.4.

Satz 11.5. *Es sei $f : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar. Dann ist f stetig in x_0.*

11.1. Rechenregeln für Ableitungen

Eine Ableitung ist nichts anderes als der Grenzwert einer Funktion, nämlich des Differenzenquotienten $x \mapsto \frac{f(x)-f(x_0)}{x-x_0}$. Für die Berechnung von Grenzwerten haben wir schon viele Techniken und Kriterien kennengelernt, die wir jetzt anwenden können. Insbesondere liefern die Grenzwertsätze sofort die Differenzierbarkeit von Summen, Vielfachen und Produkten differenzierbarer Funktionen (vgl. Satz 11.7). Zuerst bestimmen wir beispielhaft die Ableitung von Monomen und der Exponentialfunktion.

Beispiel 11.6. (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 5.11 (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 $(x^n)' = n x^{n-1}$.

(b) Es sei wieder $I = \mathbb{R}$ und jetzt

$$f(x) = \exp(x) = \mathrm{e}^x, \quad x \in \mathbb{R}.$$

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 10.17 (b). Also ist die Exponentialfunktion auf $\mathbb{R}$ differenzierbar und es gilt $\exp'(x) = \mathrm{e}^x = \exp(x)$.

Jetzt kommen wir zum Grenzwertsatz für Ableitungen, der gleichzeitig die Rechenregeln liefert, mit denen die Ableitung als Grenzwert der Differenzenquotienten von Summen, Vielfachen, Produkten und Quotienten von Funktionen bestimmt werden können.

Satz 11.7. *Es seien $f, g : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar und $\alpha, \beta \in \mathbb{R}$. 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 $f/g : J \to \mathbb{R}$ differenzierbar in x_0 und es gilt*

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{\big(g(x_0)\big)^2}. \qquad \text{(Quotientenregel)}$$

Beweis. Die Aussagen (a) und (b) behandeln wir als Übungsaufgaben.
Die Existenz von J in (c) folgt aus der Stetigkeit von g in x_0 (vgl. Satz 11.5) wie in 10.16(b).
Weiter gilt für den Differenzenquotienten

$$\begin{aligned}
\frac{\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}}{x - x_0} &= \frac{1}{g(x)g(x_0)} \cdot \frac{f(x)g(x_0) - f(x_0)g(x)}{x - x_0} \\
&= \frac{1}{g(x)g(x_0)} \cdot \frac{f(x)g(x_0) - f(x_0)g(x_0) + f(x_0)g(x_0) - f(x_0)g(x)}{x - x_0} \\
&= \frac{1}{g(x)g(x_0)} \Big(\frac{f(x) - f(x_0)}{x - x_0} g(x_0) - f(x_0) \frac{g(x) - g(x_0)}{x - x_0} \Big).
\end{aligned}$$

Da g in x_0 differenzierbar ist, ist diese Funktion insbesondere in x_0 stetig (vgl. Satz 11.5), also können wir in obiger Gleichung zum Grenzwert $x \to x_0$ übergehen und erhalten die Behauptung.

□

Es folgt sogleich die Rechenregel für die Verkettung differenzierbarer Funktionen.

Satz 11.8 (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 $f : J \to \mathbb{R}$ sei differenzierbar in $y_0 = g(x_0)$. Dann ist auch die Funktion $f \circ g : I \to \mathbb{R}$ differenzierbar in x_0 und es gilt*

$$(f \circ g)'(x_0) = f'(g(x_0)) \cdot g'(x_0).$$

Beweis. Wir betrachten den Differenzenquotienten $\tilde{f} : J \to \mathbb{R}$ von f auf J:

$$\tilde{f}(y) = \begin{cases} \dfrac{f(y) - f(y_0)}{y - y_0}, & \text{für } y \in J \text{ mit } y \neq y_0, \\ f'(y_0), & \text{für } y = y_0. \end{cases}$$

Dann gilt

$$\tilde{f}(y)(y - y_0) = f(y) - f(y_0) \tag{11.2}$$

und wegen der Differenzierbarkeit von f in y_0 ist

$$\lim_{y \to y_0} \tilde{f}(y) = f'(y_0) = f'(g(x_0)).$$

Nach Satz 11.5 ist g stetig in x_0, also gilt auch

$$\lim_{x \to x_0} \tilde{f}(g(x)) = \tilde{f}(g(x_0)) = f'(g(x_0)).$$

Daher folgt schließlich mit Hilfe von (11.2)

$$\begin{aligned} \frac{f(g(x)) - f(g(x_0))}{x - x_0} &= \frac{\tilde{f}(g(x))\big(g(x) - g(x_0)\big)}{x - x_0} \\ &= \tilde{f}(g(x))\frac{g(x) - g(x_0)}{x - x_0} \overset{x \to x_0}{\longrightarrow} f'(g(x_0)) \cdot g'(x_0). \end{aligned}$$ □

Beispiel 11.9. Es sei $a > 0$ gegeben. Dann betrachten wir auf $I = \mathbb{R}$ die Funktion

$$\varphi(x) := a^x, \quad x \in \mathbb{R}.$$

Per Definition ist $\varphi(x) = \mathrm{e}^{x \ln a}$. 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 11.8 erfüllt und es gilt

$$(a^x)' = f'(g(x))g'(x) = \mathrm{e}^{g(x)} \ln(a) = \mathrm{e}^{x \ln(a)} \ln(a) = a^x \ln(a).$$

Weiter können wir eine allgemeine Rechenregel für die Ableitung der Umkehrfunktion angeben.

Satz 11.10. *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. Für den Differenzenquotienten gilt mit $f^{-1}(y) = x$:

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} \overset{x \to x_0}{\longrightarrow} \frac{1}{f'(x_0)}.$$

Die Behauptung folgt jetzt direkt daraus, dass wegen der Stetigkeit von f^{-1} in y_0 (Satz 10.32) der Grenzübergang $y \to y_0$ die Konvergenz $f^{-1}(y) = x \to x_0 = f^{-1}(y_0)$ nach sich zieht. □

Beispiel 11.11. (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 11.10 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 Potenz mit natürlichem Exponenten aus Beispiel 11.6 (a) verallgemeinert sich also auch auf die allgemeine Potenz, solange $x > 0$.

Insbesondere haben wir im Fall $\alpha = 1/2$

$$(\sqrt{\cdot})'(x) = \frac{1}{2\sqrt{x}}, \quad x > 0.$$

11.2. Eigenschaften differenzierbarer Funktionen

Dieses Kapitel sammelt wichtige Sätze über differenzierbare Funktionen, vergleichbar mit Kapitel 10.4 für stetige Funktionen. Wir beginnen mit einem Satz, der in den verschiedensten Zusammenhängen in- und außerhalb der Mathematik zur Anwendung kommt. Er hilft bei der Bestimmung von Maximal- und Minimalstellen einer Funktion. Wir definieren zunächst genau, was wir damit meinen.

Definition 11.12. *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 lokales Maximum *(bzw.* lokales 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.* lokalen Extremum *in x_0, wenn f dort ein entsprechendes Maximum oder Minimum hat.*

Satz 11.13. *Es seien $I \subseteq \mathbb{R}$ ein offenes Intervall und $f : I \to \mathbb{R}$ differenzierbar in $x_0 \in I$. Hat f in x_0 ein lokales Extremum, so gilt $f'(x_0) = 0$.*

Warnung 11.14. 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 f auf einem Intervall ohne Randpunkte als Definitionsbereich betrachtet wird, ist wesentlich. Ein einfaches Beispiel ist die Funktion $f(x) = x$ auf dem Intervall $[0, 1]$. Diese hat ein globales (und damit insbesondere ein lokales) 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 $U_\delta(0)$ für $\delta > 0$ liegen Punkte mit $f(x) > 0 = f(0)$, z. B. $x = \delta/2$, und mit $f(x) < 0 = f(0)$, z. B. $x = -\delta/2$.

Beweis von Satz 11.13. Wir gehen zunächst davon aus, dass f in x_0 ein lokales 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 lokalen 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 lokales 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 11.15 (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.

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 11.3. Dabei ist die Sekantensteigung zwischen a und b die *mittlere* Steigung, die die Funktion auf diesem Intervall hat. Der *Mittel*wertsatz sagt also, dass differenzierbare Funktionen auf jedem Intervall mindestens einmal ihre mittlere Steigung annehmen müssen.

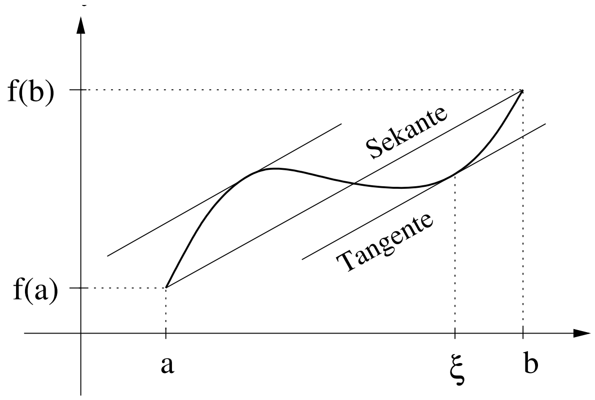

Abbildung 11.3.: Der Mittelwertsatz der Differenzialrechnung

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 $g \in \mathrm{C}([a, b], \mathbb{R})$ 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 10.27 Zahlen $t, s \in [a, b]$, so dass

$$g(t) \le g(x) \le g(s) \quad \text{für alle } x \in [a, b]$$

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 11.13 $g'(t) = 0$ oder $g'(s) = 0$ und der Beweis ist beendet. □

Wir wollen nun einige Folgerungen aus dem Mittelwertsatz ziehen.

Satz 11.16. *(a) 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 $f'(\xi) = 0$.* (Satz von Rolle)

(b) *Es sei $f : I \to \mathbb{R}$ auf dem Intervall I differenzierbar. Dann gilt*

$$\begin{aligned}
&\text{Ist } f' = 0 \text{ auf } I, \text{ so ist } f \text{ auf } I \text{ konstant.}\\
&\text{Ist } f' > 0 \text{ auf } I, \text{ so ist } f \text{ auf } I \text{ streng monoton wachsend.}\\
&\text{Ist } f' < 0 \text{ auf } I, \text{ so ist } f \text{ auf } I \text{ streng monoton fallend.}\\
&\text{Ist } f' \geq 0 \text{ auf } I, \text{ so ist } f \text{ auf } I \text{ monoton wachsend.}\\
&\text{Ist } f' \leq 0 \text{ auf } I, \text{ so ist } f \text{ auf } I \text{ monoton fallend.}
\end{aligned}$$

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.

□

Schon an diesen reichhaltigen Folgerungen sieht man die Stärke des Mittelwertsatzes. Dieser ist aber auch im analytischen Alltag immer wieder ein unverzichtbares Hilfmittel. Eine typische Anwendung zeigt das folgende Beispiel.

Beispiel 11.17. Wir zeigen, dass für alle $a, b \geq 1$ die Ungleichung

$$\left|\sqrt{a} - \sqrt{b}\right| \leq \frac{|a - b|}{2}$$

gilt. Dazu seien $a, b \geq 1$ gegeben. Wir wenden den Mittelwertsatz auf die Wurzelfunktion $f(x) = \sqrt{x}$ an und erhalten mit einem ξ zwischen a und b

$$\left|\sqrt{a} - \sqrt{b}\right| = \left|f(a) - f(b)\right| = \left|f'(\xi)(a - b)\right| = \left|f'(\xi)\right||a - b|.$$

Da a und b größer oder gleich 1 sind, gilt $\xi > 1$, also ist

$$\left|f'(\xi)\right| = \frac{1}{2\sqrt{\xi}} < \frac{1}{2}$$

und wir erhalten

$$\left|\sqrt{a} - \sqrt{b}\right| \leq \frac{1}{2}|a - b| = \frac{|a - b|}{2}.$$

Satz 11.18 (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. Zuerst halten wir fest, dass $g(b) - g(a) \neq 0$ ist, da sich mit $g'(x) \neq 0$ für alle $x \in (a, b)$ sonst ein Widerspruch zum Satz von Rolle ergäbe. Wir betrachten jetzt die Hilfsfunktion

$$h(x) = \big(f(b) - f(a)\big)g(x) - \big(g(b) - g(a)\big)f(x).$$

Dann ist $h \in \mathrm{C}([a, b], \mathbb{R})$, 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. □

Übungsaufgabe 11.19. Es sei $f : I \to \mathbb{R}$ differenzierbar auf einem Intervall $I \subseteq \mathbb{R}$. Dann erfüllt f' die *Zwischenwerteigenschaft*, d. h. sind $u, w \in f'(I)$ mit $u < w$ und $y_0 \in (u, w)$, so gibt es ein $x_0 \in I$ mit $f'(x_0) = y_0$.
Anders formuliert: Für jede differenzierbare Funktion $f : I \to \mathbb{R}$ ist $f'(I)$ ein Intervall (oder ein Punkt).
Beachten Sie bei Ihrem Beweis, dass f' nicht unbedingt stetig vorausgesetzt ist, den Zwischenwertsatz können Sie also in der Tasche lassen.

Der verallgemeinerte Mittelwertsatz gibt uns ein starkes Hilfsmittel zur Bestimmung von Grenzwerten bei Quotienten von Funktionen, das wir zum Abschluss dieses Abschnitts beweisen wollen.

Satz 11.20 (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

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

oder hat dieser im Sinne von bestimmter Divergenz einen Wert ∞ oder $-\infty$, 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.

Beweis. 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 wählen ein $\tilde{b} \in (a, b)$. Dann gilt $f, g \in \mathrm{C}([a, \tilde{b}], \mathbb{R})$, denn die Funktionen sind ja auf (a, b) differenzierbar und damit insbesondere stetig. Ist nun $x \in (a, \tilde{b})$, so gibt es nach dem verallgemeinerten Mittelwertsatz (Satz 11.18) 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

$$\tilde{f}(t) := f({}^1\!/\!_t) \quad \text{und} \quad \tilde{g}(t) := g({}^1\!/\!_t), \quad t \in (-{}^1\!/\!_{|b|}, 0).$$

Dann gilt für alle $t \in (-{}^1\!/\!_{|b|}, 0)$

$$\tilde{f}'(t) = f'\Big(\frac{1}{t}\Big) \cdot \Big(-\frac{1}{t^2}\Big) \quad \text{und} \quad \tilde{g}'(t) = g'\Big(\frac{1}{t}\Big) \cdot \Big(-\frac{1}{t^2}\Big)$$

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. $\square$

Beispiel 11.21. (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:

$$\lim_{x\to\infty} \frac{\ln(x)}{x} = \lim_{x\to\infty} \frac{1/x}{1} = 0.$$

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:

$$\lim_{x\to 0+} x\ln(x) = \lim_{x\to 0+} \frac{\ln(x)}{1/x} = \lim_{x\to 0+} \frac{1/x}{-1/x^2} = \lim_{x\to 0+} (-x) = 0.$$

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

$$\lim_{x\to 0+} x^x = \lim_{x\to 0+} \mathrm{e}^{x\ln(x)} = \mathrm{e}^{\lim_{x\to 0+} x\ln(x)} = \mathrm{e}^0 = 1.$$

Warnung 11.22. 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! Die besondere Gemeinheit ist, dass man bei einer nicht gerechtfertigten Anwendung von l'Hospital keine Warnmeldung zurückbekommt, sondern ein schönes Ergebnis, das aber meistens leider einfach falsch ist. Hier ist ein typisches Beispiel: Einfach mit Anwendung der Grenzwertsätze ergibt sich, dass

$$\lim_{x\to 0} \frac{\mathrm{e}^x - 1}{x-1} = \frac{\lim_{x\to 0}(\mathrm{e}^x - 1)}{\lim_{x\to 0}(x-1)} = \frac{0}{-1} = 0$$

ist. l'Hospital ist hier nicht anwendbar, da weder Fall (I) noch Fall (II) aus dem Satz vorliegt. Wendet man ihn aber trotzdem an, bekommt man

$$\lim_{x\to 0} \frac{\mathrm{e}^x - 1}{x-1} \overset{???}{=} \lim_{x\to 0} \frac{\mathrm{e}^x}{1} = \mathrm{e}^0 = 1,$$

also ein falsches Ergebnis.

11.3. Die Ableitung von Potenzreihen

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 11.23. *Es sei $f(x) = \sum_{n=0}^{\infty} a_n(x-p_0)^n$ eine Potenzreihe mit Entwicklungspunkt $p_0 \in \mathbb{R}$ und strikt positivem Konvergenzradius r. Wir setzen $I := (p_0 - r, p_0 + r)$. Dann gilt:*

(a) Die Potenzreihe

$$g(x) = \sum_{n=1}^{\infty} na_n(x-p_0)^{n-1}$$

hat ebenfalls den Konvergenzradius r.

(b) Die Funktion f ist auf I differenzierbar und es gilt

$$f'(x) = g(x) = \sum_{n=1}^{\infty} na_n(x-p_0)^{n-1} \text{ für alle } x \in I.$$

Beweis. Da mit Hilfe von Übungsaufgabe 7.45

$$\limsup_{n\to\infty} \sqrt[n]{|na_n|} = \limsup_{n\to\infty}\left(\sqrt[n]{n}\,\sqrt[n]{|a_n|}\right) = \lim_{n\to\infty} \sqrt[n]{n} \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$$

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 $s_n := \sum_{k=0}^{n} a_k(x-p_0)^k$ die n-te Partialsumme und mit $R_n := \sum_{k=n+1}^{\infty} a_k(x-p_0)^k$ den n-ten Reihenrest. Nun sei $x_0 \in I$ fest. Dann wählen wir ein $0 < \varrho < r$, so dass $x_0 \in [p_0 - \varrho, p_0 + \varrho] \subseteq I$ gilt. Es folgt für alle $x \in [p_0 - \varrho, p_0 + \varrho]$ mit $x \neq x_0$ und alle $n \in \mathbb{N}$

$$\begin{aligned}
&\left|\frac{f(x)-f(x_0)}{x-x_0} - g(x_0)\right| \\
&= \left|\frac{s_n(x)+R_n(x)-s_n(x_0)-R_n(x_0)}{x-x_0} - s_n'(x_0) + s_n'(x_0) - g(x_0)\right| \\
&\leq \left|\frac{s_n(x)-s_n(x_0)}{x-x_0} - s_n'(x_0)\right| + \left|\frac{R_n(x)-R_n(x_0)}{x-x_0}\right| + |s_n'(x_0) - g(x_0)|. \qquad (11.3)
\end{aligned}$$

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

$$s_n'(x) = \sum_{k=1}^{n} ka_k(x-p_0)^{k-1},$$

also existiert 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.
Wir wenden uns dem zweiten Summanden zu. Für alle $x \in [p_0 - \varrho, p_0 + \varrho]$ mit $x \neq x_0$ und alle $n \in \mathbb{N}$ gilt nach der Dreiecksungleichung und mit Hilfe von Satz 5.11 (c)

$$\begin{aligned}
\Big|\frac{R_n(x) - R_n(x_0)}{x - x_0}\Big| &= \Big|\sum_{k=n+1}^{\infty} \frac{a_k\big((x-p_0)^k - (x_0-p_0)^k\big)}{x - x_0}\Big| \\
&= \Big|\sum_{k=n+1}^{\infty} \frac{a_k}{x - x_0}(x - p_0 - x_0 + p_0)\sum_{j=0}^{k-1}(x-p_0)^j(x_0-p_0)^{k-1-j}\Big| \\
&\leq \sum_{k=n+1}^{\infty} |a_k| \sum_{j=0}^{k-1} |x - p_0|^j |x_0 - p_0|^{k-1-j}.
\end{aligned}$$

Verwenden wir nun, dass sowohl $x - p_0$ als auch $x_0 - p_0$ im Betrag unterhalb ϱ bleiben, ergibt sich als weitere Abschätzung

$$\begin{aligned}
&\leq \sum_{k=n+1}^{\infty} |a_k| \sum_{j=0}^{k-1} \varrho^j \varrho^{k-1-j} = \sum_{k=n+1}^{\infty} |a_k| \sum_{j=0}^{k-1} \varrho^{k-1} \\
&= \sum_{k=n+1}^{\infty} k|a_k|\varrho^{k-1} =: c_n.
\end{aligned}$$

Wir haben in (a) gesehen, dass die Potenzreihe $\sum_{n=1}^{\infty} n|a_n|(x-p_0)^{n-1}$ auch den Konvergenzradius r hat. Da $\varrho \in (-r, r)$ gilt, konvergiert 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 8.7 (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 (11.3) für $n = n_0$, und nehmen wir das soeben gewählte δ, so ist jeder der 3 Summanden auf der rechten Seite von (11.3) kleiner als $\varepsilon/3$ und die Behauptung ist bewiesen. □

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 11.24. 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 11.16 (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.

11.4. Trigonometrische Funktionen

Wir hatten im Abschnitt 8.5 über Potenzreihen die trigonometrischen Funktionen Sinus und Cosinus definiert als

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad \text{und} \quad \cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Geometrisch kennen wir $\sin(\alpha)$ und $\cos(\alpha)$ als die Längen der Katheten des rechtwinkligen Dreiecks, das im Einheitskreis zum Winkel α (im Bogenmaß) abgetragen werden kann, s. Abbildung 11.4. Unter anderem dadurch, dass wir den *trigonometrischen Pythagoras*, $\sin^2(\alpha) + \cos^2(\alpha) = 1$, mit Hilfe der Potenzreihendarstellung beweisen, überzeugen wir uns, dass die Potenzreihen dem geometrischen Bild entsprechen. In diesem Abschnitt beweisen wir noch weitere Eigenschaften von Sinus und Cosinus und definieren die Zahl π und weitere *trigonometrische Funktionen*. Wir definieren auch die zugehörigen Umkehrfunktionen und bestimmen die Ableitungen.

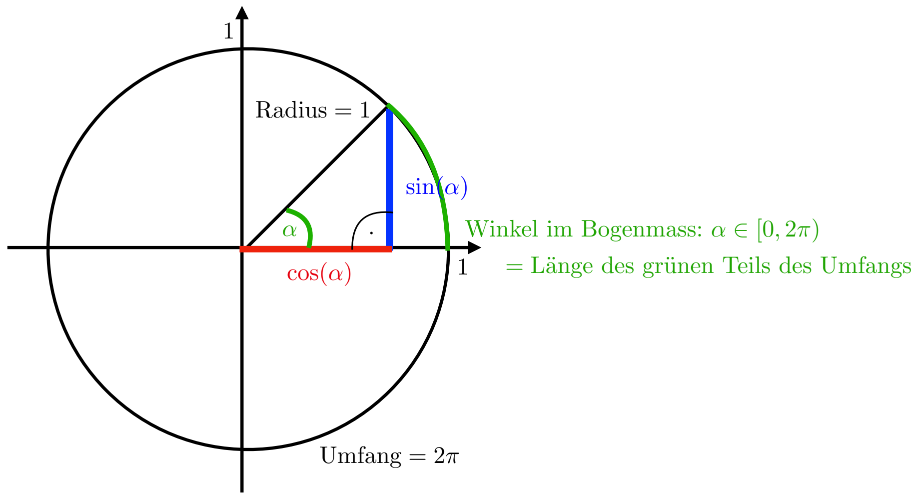

Abbildung 11.4.: Cosinus und Sinus sind die Längen der Katheten zum Winkel α im Einheitskreis (oder: die x- und y-Abschnitte). Die Größe des Winkels ist im Bogenmaß die Länge des entsprechenden Anteils des Umfangs im Einheitskreis. Den trigonometrischen Pythagoras, $\sin^2(\alpha) + \cos^2(\alpha) = 1$, sehen wir hier geometrisch.

Mit unserem Satz 11.23 über die Differenzierbarkeit von Potenzreihen können wir direkt die Ableitung des Sinus berechnen 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 11.25. *Sinus und Cosinus sind auf $\mathbb{R}$ differenzierbar und es gilt*

$$\sin'(x) = \cos(x) \quad \textit{und} \quad \cos'(x) = -\sin(x), \qquad x \in \mathbb{R}.$$

Weitere Eigenschaften von Cosinus und Sinus:

Satz 11.26. *Für alle $x, y \in \mathbb{R}$ gelten die folgenden Aussagen:*

(a) $\sin^2(x) + \cos^2(x) = 1.$ (trigonometrischer Pythagoras)

(b) $|\sin(x)| \leq 1$ *und* $|\cos(x)| \leq 1.$

(c) $\sin(-x) = -\sin(x)$ *und* $\cos(-x) = \cos(x).$

(d) Additionstheoreme*:*

$$\begin{aligned}\sin(x + y) &= \sin(x)\cos(y) + \sin(y)\cos(x),\\ \cos(x + y) &= \cos(x)\cos(y) - \sin(x)\sin(y).\end{aligned}$$

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) Für alle $x \in \mathbb{R}$ gilt $|\sin(x)| = \sqrt{\sin^2(x)} \leq \sqrt{\sin^2(x) + \cos^2(x)} = 1$. Genauso rechnet man für den Cosinus.

(c) 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.

(d) Das erste Additionstheorem wurde mit Hilfe des Cauchyprodukts für Potenzreihen in der Zusatzaufgabe von Übungsblatt 7 bewiesen . Das zweite kann analog bewiesen werden. *Alternativ ist der Beweis deutlich kürzer, wenn wir die Euler-Formel, Satz 9.18 (c), aus dem Komplexen nutzen, mit der gilt

$$\cos(x + y) + \mathrm{i}\sin(x + y) = \mathrm{e}^{i(x+y)}.$$

Nach Teil (a) des selben Satzes haben wir für die komplexe Exponentialfunktion die gewohnten Potenzrechenregeln, also gilt mit rückwärtiger Anwendung der Euler-Formel

$$\begin{aligned}\cos(x + y) + \mathrm{i}\sin(x + y) = \mathrm{e}^{\mathrm{i}x}\mathrm{e}^{\mathrm{i}y} &= \big(\cos(x) + \mathrm{i}\sin(x)\big)\big(\cos(y) + \mathrm{i}\sin(y)\big)\\ &= \cos(x)\cos(y) - \sin(x)\sin(y)\\ &\qquad + \mathrm{i}\big(\cos(x)\sin(y) + \sin(x)\cos(y)\big).\end{aligned}$$

Vergleicht man links und rechts dieser Gleichung jeweils den Real- und den Imaginärteil, so ergeben sich gleich beide Additionstheoreme. □

Die Eigenschaften aus Teil (c) des obigen Satzes kommen immer wieder mal vor und bekommen deshalb einen eigenen Namen.

Definition 11.27. *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 $f(-x) = -f(x)$.*

Wir wollen jetzt die Zahl π mit Hilfe der Cosinus-Reihe definieren. Geometrisch gibt sie das für jeden Kreis identische Verhältnis des Umfangs U zum Durchmesser D an. Im Einheitskreis gilt also

$$\pi = \frac{U}{D} = \frac{U}{2}. \tag{11.4}$$

Dazu zuerst zwei Vorüberlegungen.

Satz 11.28. *(a) Für alle $x \in (0, 2)$ gilt $\sin(x) > x - x^3/3! > 0$.*

(b) Die Funktion cos *hat eine kleinste strikt positive Nullstelle ξ_0.*

Beweis. (a) Sei $x \in (0, 2)$. Dann gilt für alle $n \in \mathbb{N}$

$$x^2 < 4 \leq 4n^2 \leq 4n^2 + 2n = 2n(2n + 1)$$

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

$$\sin(x) = \underbrace{\left(x - \frac{x^3}{3!}\right)}_{>0} + \underbrace{\left(\frac{x^5}{5!} - \frac{x^7}{7!}\right)}_{>0} + \underbrace{\left(\frac{x^9}{9!} - \frac{x^{11}}{11!}\right)}_{>0} + \cdots > x - \frac{x^3}{3!} > 0.$$

(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 11.26 (b)

$$\begin{aligned} \cos(2) &= \cos(1+1) = \cos^2(1) - \sin^2(1) = \cos^2(1) + \sin^2(1) - 2\sin^2(1) \\ &= 1 - 2\sin^2(1) < 1 - 2 \cdot \frac{5^2}{6^2} = 1 - \frac{50}{36} < 0. \end{aligned}$$

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, gibt es nach Satz 4.17 für jedes $n \in \mathbb{N}$ ein $\xi_n \in M$ mit $\xi_0 \leq \xi_n \leq \xi_0 + {}^1\!/\!_n$. Insbesondere konvergiert diese Folge $(\xi_n)_n$ gegen ξ_0. Nun ist aber die Cosinus-Funktion stetig, d. h.

$$\cos(\xi_0) = \lim_{n\to\infty} \cos(\xi_n) = \lim_{n\to\infty} 0 = 0.$$

Also ist $\xi_0 = \min M$ und schließlich ist ξ_0 strikt positiv, da $\xi_0 \geq 0$ ist und $\xi_0 = 0$ wegen $\cos(0) = 1$ nicht sein kann. $\square$

In Abbildung 11.4 sehen wir, dass die kleinste Nullstelle ξ_0 des Cosinus gleich dem Viertel des Umfangs U des Einheitskreises sein sollte. Wegen (11.4) gilt außerdem $\frac{U}{4} = \frac{\pi}{2}$. Daraus ergibt sich die folgende Definition.

Definition 11.29. *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 etwa $3,14159\ldots$ beträgt.
Wir haben nun also nach Definition $\cos(\pi/2) = 0$ und dank $\pi/2 \in (0, 2)$ und Satz 11.28 (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, die sich alternativ geometrisch am Einheitskreis ablesen lassen.

Satz 11.30. *Für alle $x \in \mathbb{R}$ gilt*

$$\begin{aligned} \sin(x + \pi/2) &= \cos(x), & \cos(x + \pi/2) &= -\sin(x), \\ \sin(x + \pi) &= -\sin(x), & \cos(x + \pi) &= -\cos(x), \\ \sin(x + 2\pi) &= \sin(x), & \cos(x + 2\pi) &= \cos(x). \end{aligned}$$

Auch diese Eigenschaft von Sinus und Cosinus, dass die Funktionen bei einer geeigneten Verschiebung im Argument unverändert bleiben, bekommt einen Namen.

Definition 11.31. *Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt* periodisch *mit* Periode $L \in \mathbb{R}$ *oder kurz* L-periodisch, *wenn für alle $x \in \mathbb{R}$ gilt $f(x + L) = f(x)$.*

Damit sind Cosinus und Sinus beide 2π-periodisch.

Satz 11.32. *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 11.28 sofort $\eta \geq \pi/2$. Wir setzen $\tilde{\eta} := \pi - \eta$. Dann ist $\tilde{\eta} \in [0, \pi/2]$. Außerdem ist nach Satz 11.30 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 11.33. *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 11.30. 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 die Gaußklammer kennt, kann $k = \lfloor x/\pi \rfloor$ 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 11.32 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. $x = k\pi$. □

Wir können nun weitere trigonometrische Funktionen definieren.

Definition 11.34. *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.*

Tangens und Cotangens haben auch eine geometrische Interpretation. Sie geben die Länge der dem Winkel im Einheitskreis entsprechenden Tangentenabschnitte an, s. Abbildung 11.5.

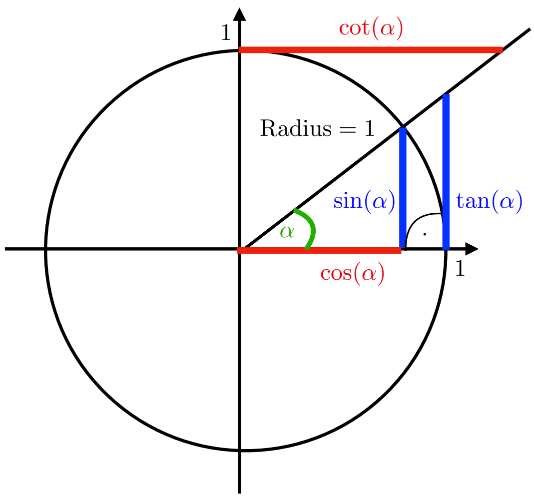

Abbildung 11.5.: Tangens und Cotangens beziehen sich auf die Tangenten am Einheitskreis. Mit den Strahlensätzen $\frac{\tan\alpha}{\sin\alpha} = \frac{1}{\cos\alpha}$ und $\frac{\cos\alpha}{\cot\alpha} = \frac{\sin\alpha}{1}$ ergeben sich die Definitionen geometrisch.

Für die Ableitung des Tangens gilt mit der Quotientenregel,

$$\tan'(x) = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = 1 + \tan^2(x) > 0$$

für alle x im Definitionsbereich. 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 11.35. *Die Umkehrfunktion des Tangens auf* $(-\pi/2, \pi/2)$

$$\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$$

heißt Arcustangens.

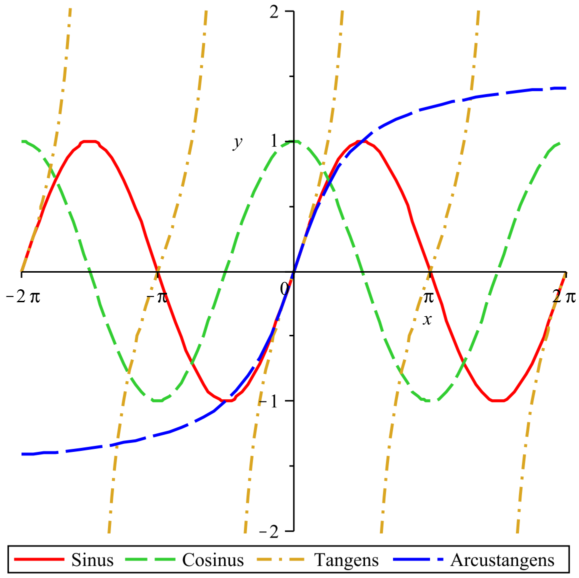

Abbildung 11.6.: Die Graphen von Sinus, Cosinus, Tangens und Arcustangens

Die Ableitung des Arcustangens können wir nun nach der Formel für die Ableitung der Umkehrfunktion (Satz 11.10) 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. Das rechtfertigt die folgende Definition.

Definition 11.36.

$$\begin{aligned} \arcsin &:= \sin^{-1} : [-1, 1] \to [-\pi/2, \pi/2] \quad \text{(Arcussinus)}, \\ \arccos &:= \cos^{-1} : [-1, 1] \to [0, \pi] \quad \text{(Arcuscosinus)}. \end{aligned}$$

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 11.37. *Für alle $x \in \mathbb{R}$ definieren wir*

$$\begin{aligned} \cosh(x) &:= \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x}) && \text{(Cosinus hyperbolicus)},\\ \sinh(x) &:= \frac{1}{2}(\mathrm{e}^x - \mathrm{e}^{-x}) && \text{(Sinus hyperbolicus)},\\ \tanh(x) &:= \frac{\sinh(x)}{\cosh(x)} && \text{(Tangens hyperbolicus)}. \end{aligned}$$

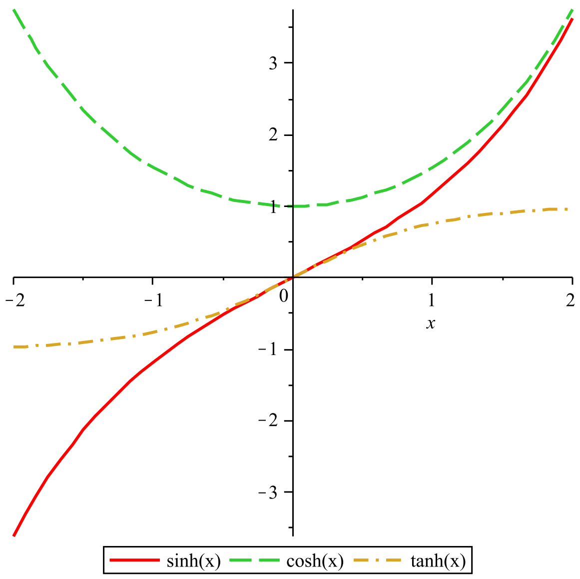

Abbildung 11.7.: Die Graphen von Sinus, Cosinus und Tangens hyperbolicus

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

$$\mathrm{Artanh}'(x) = \frac{1}{1 - \tanh^2(\mathrm{Artanh}(x))} = \frac{1}{1 - x^2}, \quad |x| < 1.$$

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	$(-1,1)$	$\mathbb{R}$	$\frac{1}{1-x^2}$

*Ergänzung: komplexe Polarkoordinaten und komplexer Logarithmus

Sinus und Cosinus wurden in Kapitel 9 im Komplexen analog über ihre Potenzreichen definiert. Wir können ihre Eigenschaften nutzen, um nun die komplexen Zahlen noch genauer zu verstehen. Insbesondere lässt sich beschreiben, was die komplexe Multiplikation anschaulich tut. Dazu erinnern wir uns erneut an die Euler-Formel

$$\mathrm{e}^{it} = \cos(t) + \mathrm{i}\sin(t) \qquad \text{für alle } t \in \mathbb{R}$$

aus Satz 8.40(c). Diese liefert uns im Zusammenspiel mit dem trigonometrischen Pythagoras, vgl. Satz 11.26 (b), die folgenden Erkenntnisse.

Satz 11.38. *Für alle $t \in \mathbb{R}$ und alle $z \in \mathbb{C}$ gelten:*

(a) $|\mathrm{e}^{\mathrm{i}t}| = 1$ *und* $|\mathrm{e}^{z}| = \mathrm{e}^{\mathrm{Re}(z)}$.

(b) $\mathrm{e}^{z+2\pi\mathrm{i}} = \mathrm{e}^{z}$*, d. h.* $\exp : \mathbb{C} \to \mathbb{C}$ *ist* $2\pi\mathrm{i}$*-periodisch.*

Beweis. Alle drei Gleichheiten rechnet man mit den oben schon angegebenen Zutaten direkt nach:

$$\begin{aligned} |\mathrm{e}^{\mathrm{i}t}| &= \big|\cos(t) + \mathrm{i}\sin(t)\big| = \sqrt{\cos^2(t) + \sin^2(t)} = \sqrt{1} = 1, \\ |\mathrm{e}^{z}| &= \big|\mathrm{e}^{\mathrm{Re}(z)+\mathrm{i}\mathrm{Im}(z)}\big| = |\mathrm{e}^{\mathrm{Re}(z)}||\mathrm{e}^{\mathrm{i}\mathrm{Im}(z)}| = \mathrm{e}^{\mathrm{Re}(z)} \cdot 1 = \mathrm{e}^{\mathrm{Re}(z)}, \\ \mathrm{e}^{z+2\pi\mathrm{i}} &= \mathrm{e}^{z}\mathrm{e}^{2\pi\mathrm{i}} = \mathrm{e}^{z}\big(\cos(2\pi) + \mathrm{i}\sin(2\pi)\big) = \mathrm{e}^{z}(1+0) = \mathrm{e}^{z}. \end{aligned}$$ □

Neben der Darstellung einer komplexen Zahl als $x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$, die dem kartesischen Koordinatensystem in der komplexen Zahlenebene entspricht, gibt es eine zweite Möglichkeit der Darstellung, die in Polarkoordinaten. Dazu beobachten wir zunächst anschaulich geometrisch, vgl. Abbildung 11.4, 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 11.39. *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 $|u| = |\mathrm{Re}(w)| \le 1$. 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

$$\cos(\phi - \psi) = \mathrm{Re}\big(\mathrm{e}^{\mathrm{i}(\phi-\psi)}\big) = 1 \quad \text{und} \quad \sin(\phi - \psi) = \mathrm{Im}\big(\mathrm{e}^{\mathrm{i}(\phi-\psi)}\big) = 0$$

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. □

Definition 11.40. *Es sei $z \in \mathbb{C} \setminus \{0\}$. Die nach Satz 11.39 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 11.41. *Es sei $w \in \mathbb{C}$. Jede Zahl $z \in \mathbb{C}$ mit $\mathrm{e}^z = w$ heißt ein* Logarithmus *von 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 9.18 (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 11.39 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.

11.5. Höhere Ableitungen

Wir wollen in diesem Abschnitt das Konzept der zweiten, dritten und allgemein n-ten Ableitung einer Funktion einführen.

Definition 11.42. *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f : I \to \mathbb{R}$ sei differenzierbar auf I.*

(a) *Die Funktion f heißt in x_0* zweimal differenzierbar, *falls die Funktion $f' : I \to \mathbb{R}$ 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 11.43. (a) Ist $f(x) = \sin(x)$, so gilt

$$f'(x) = \cos(x), \qquad f''(x) = -\sin(x) = -f(x),$$
$$f'''(x) = -\cos(x), \quad f^{(4)}(x) = \sin(x) = f(x).$$

(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 11.3 (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 11.23 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

$$f^{(k)}(x_0) = a_k k(k-1)\cdots 2\cdot 1 = k!a_k.$$

Andersrum gedacht zeigt sie uns insbesondere: ist eine Funktion f durch eine Potenzreihe darstellbar, dann lassen sich die Koeffizienten a_k dieser Reihe durch die Ableitungen von f im Entwicklungspunkt bestimmen:

$$a_k = \frac{f^{(k)}(x_0)}{k!}.$$

(d) Wir betrachten auf $I = [0, \infty)$ die Funktion

$$f(x) := \begin{cases} x^{3/2}\sin(1/x), & \text{falls } x > 0 \\ 0, & \text{falls } x = 0, \end{cases}$$

vgl. Abbildung 11.8.

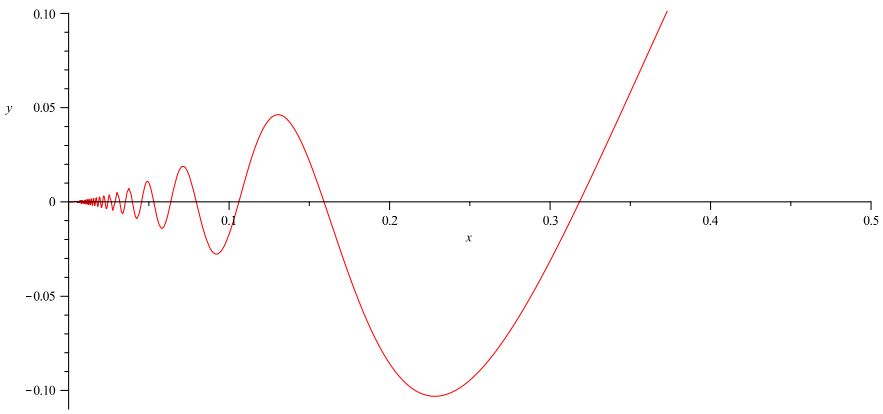

Abbildung 11.8.: Der Graph des „Flattersinus'" aus Beispiel 11.43 (d)

Dann ist f in allen $x > 0$ offensichtlich differenzierbar und es gilt

$$\begin{aligned} f'(x) &= \frac{3}{2}\sqrt{x}\sin\Big(\frac{1}{x}\Big) + x^{3/2}\cos\Big(\frac{1}{x}\Big)\Big(-\frac{1}{x^2}\Big) \\ &= \frac{3}{2}\sqrt{x}\sin\Big(\frac{1}{x}\Big) - \frac{1}{\sqrt{x}}\cos\Big(\frac{1}{x}\Big). \end{aligned}$$

Um f auf Differenzierbarkeit in 0 zu untersuchen, betrachten wir den Differenzenquotienten

$$\begin{aligned}\lim_{x\to 0+}\left|\frac{f(x)-f(0)}{x-0}\right| &= \lim_{x\to 0+}\left|\frac{x^{3/2}\sin(1/x)}{x}\right| = \lim_{x\to 0+}\left|\sqrt{x}\sin\Big(\frac{1}{x}\Big)\right| \\ &\le \lim_{x\to 0+}\sqrt{x} = 0.\end{aligned}$$

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. Darüber hinaus ist sie „gegen 0 “ unbeschränkt und stark oszillierend, also in gewisser Weise so schlimm, wie es überhaupt nur geht. Um das zu sehen, betrachten wir die Nullfolge $x_n := 1/n\pi$, $n\in\mathbb{N}$. Dann gilt für jedes $n\in\mathbb{N}$

$$\begin{aligned}f'(x_n) &= \frac{3}{2}\frac{1}{\sqrt{n\pi}}\sin(n\pi) - \sqrt{n\pi}\cos(n\pi) = -\sqrt{n\pi}\cos(n\pi) \\ &= -(-1)^n\sqrt{n\pi} = (-1)^{n+1}\sqrt{n\pi}.\end{aligned}$$

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 11.44. *Es sei $I\subseteq\mathbb{R}$ ein Intervall und $f : I\to\mathbb{R}$ 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.*

Wir setzen $\mathrm{C}^n(I,\mathbb{R}) := \{f : I\to\mathbb{R} : f$ n-mal stetig differenzierbar auf $I\}$.

(b) Wir setzen

$$\mathrm{C}^0(I,\mathbb{R}) := \mathrm{C}(I,\mathbb{R}) \quad \textit{und} \quad \mathrm{C}^\infty(I,\mathbb{R}) := \bigcap_{n\in\mathbb{N}_0}\mathrm{C}^n(I,\mathbb{R})$$

und nennen eine Funktion $f\in\mathrm{C}^\infty(I,\mathbb{R})$ beliebig oft differenzierbar.

Bemerkung 11.45. (a) Da die Differenzierbarkeit insbesondere Stetigkeit impliziert, sind für eine Funktion $f\in\mathrm{C}^n(I,\mathbb{R})$ 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. $f\in\mathrm{C}^\infty(I)$ heißt aber eben nur, dass $f^{(n)}$ für jedes $n\in\mathbb{N}$ auf I existiert und dort stetig ist.

11.6. Der Satz von Taylor

Dieser Abschnitt stellt mit dem Satz von Taylor einen fundamental wichtigen Satz der Analysis vor, der sowohl in abstrakten als auch in ganz angewandten Zusammenhängen immer wieder gebraucht wird. Es geht dabei darum, Funktionen möglichst gut durch Polynome anzunähern.
Konkreter betrachten wir eine n-mal differenzierbare Funktion f und suchen ein Polynom p vom maximalen Grad n, sodass in einem Punkt x_0,

$$f(x_0) = p(x_0), \quad f'(x_0) = p'(x_0), \quad \dots \quad f^{(n)}(x_0) = p^{(n)}(x_0),$$

gilt. Ist dabei p von der Form

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 \cdots + a_n(x - x_0)^n,$$

dann folgt induktiv wie in Beispiel 11.43 (c), dass die Koeffizienten durch

$$a_0 = f(x_0), \quad a_1 = f'(x_0), \quad \dots \quad a_n = \frac{f^{(n)}(x_0)}{n!},$$

gegeben sind. Das Polynom ist also eindeutig bestimmt und es bekommt einen eigenen Namen.

Definition 11.46. *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$, $n \geq 1$ und $f \in \mathrm{C}^n(I, \mathbb{R})$. Dann heißt das Polynom*

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

das n-te Taylorpolynom *von f in x_0.*

Nun stellt sich sofort die

Frage 1: Liefert das Taylorpolynom eine gute Näherung der Funktion f?
Um diese Frage zu beantworten, untersuchen wir das *Restglied*

$$R_n(x, x_0) = f(x) - T_n(x, x_0),$$

das den Fehler angibt. Der Satz von Taylor ist eine Aussage über die Struktur von R_n und erlaubt uns insbesondere, den Fehler abzuschätzen, wenn wir die $n+1$-te Ableitung von f gut genug kennen.

Satz 11.47 (Satz von Taylor)**.** *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x, x_0 \in I$ und $f : I \to \mathbb{R}$ sei $(n+1)$-mal differenzierbar auf I. 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 11.48. (a) Im Fall $n = 0$ ist dieser Satz genau der Mittelwertsatz (vgl. Satz 11.15).

(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 (und damit in R_n) steckt sozusagen die Komplexität der Funktion f.

Beweis von Satz 11.47. 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 $g \in \mathrm{C}([x_0, x], \mathbb{R})$ 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 Satz von Rolle 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 betrachten jetzt den Spezialfall einer beliebig oft differenzierbaren Funktion $f \in C^\infty(I, \mathbb{R})$. Wir können dann für ein $x_0 \in I$ *alle* Taylor-Koeffizienten $\frac{f^{(n)}(x_0)}{n!}$ ausrechnen und erhalten damit eine ganze (Potenz-)reihe von Taylorpolynomen. Sie bekommt zuerst einen eigenen Namen.

Definition 11.49. *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $x_0 \in I$ und $f \in C^\infty(I, \mathbb{R})$. Die Potenzreihe*

$$\left(\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k\right)_n = \big(T_n(x, x_0)\big)_n$$

heißt dann die Taylorreihe von f um x_0.

Nun stellt sich sofort die

Frage 2: Konvergiert die Taylorreihe und wenn ja, gilt dann auch 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 (leider nur) ein entschiedenes *manchmal*.
Ist einerseits f durch eine Potenzreihe gegeben (wie ja zum Beispiel exp, sin, ...), dann entspricht diese mit Beispiel 11.43 (c) auch automatisch der Taylorreihe (wenn der Entwicklungspunkt übereinstimmt) und die Antwort ist dann *ja*.
Andererseits zeigt das folgende Beispiel, dass nicht jede beliebig oft differenzierbare Funktion lokal durch eine Potenzreihe darstellbar ist. Für diesen Fall wäre die Antwort dann *nein*:

Beispiel 11.50. Wir wählen $I = \mathbb{R}$, $x_0 = 0$ und

$$f(x) := \begin{cases} e^{-1/x^2}, \text{ falls } x \neq 0, \\ 0, \text{ falls } x = 0. \end{cases}$$

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 für alle $x \neq 0$

$$\frac{f(x) - f(0)}{x - 0} = \frac{e^{-1/x^2}}{x} = \frac{1/x}{e^{1/x^2}}.$$

Da der Nenner dieses Ausdrucks für x gegen null bestimmt nach unendlich divergiert, ist die Bedingung (II) für die Regel von de l'Hospital aus Satz 11.20 erfüllt. Dieser Satz liefert dann

$$\begin{aligned} f'(0) &= \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{1/x}{e^{1/c^2}} = \lim_{x\to 0} \frac{-1/x^2}{e^{1/x^2} \cdot \left(-2/x^3\right)} \\ &= \lim_{x\to 0} \frac{x}{2} \cdot e^{-1/x^2} = 0 \cdot 0 = 0. \end{aligned}$$

Mittels Induktion kann man sogar zeigen, dass f in null beliebig oft differenzierbar ist und $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$ gilt. Also konvergiert in diesem Fall die Taylorreihe, allerdings nicht gegen die richtigen Werte:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} 0 \cdot x^n = 0 \neq f(x) \quad \text{für alle } x \neq 0.$$

Sie stellt also in keiner auch noch so kleinen Umgebung des Nullpunktes die Funktion f dar. Weiterhin gilt: Wäre f nahe dem Nullpunkt durch eine andere Potenzreihe darstellbar, müsste diese der Taylorreihe entsprechen. Das ist also auch nicht möglich.

Der folgende Satz liefert jetzt 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 11.50 nicht passieren können. Es handelt sich natürlich um eine Folgerung aus dem Satz von Taylor.

Satz 11.51. *Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f \in \mathrm{C}^\infty(I, \mathbb{R})$ 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 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 nur 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. $\square$

Wir wollen nun den Satz von Taylor in ganz verschiedenen konkreten Situationen anwenden. Zunächst einmal kann man ihn dazu verwenden, den einen oder anderen Reihenwert zu bestimmen. Wir betrachten hierzu das folgende Beispiel.

Beispiel 11.52. 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 $x = 1$):

$$\begin{aligned} \ln(2) = f(1) &= f(0) + \sum_{k=1}^{n} \frac{f^{(k)}(0)}{k!} 1^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} 1^{n+1} \qquad (11.5) \\ &= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} + \frac{(-1)^n}{(1+\xi)^{n+1}(n+1)} =: \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} + c_n \end{aligned}$$

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 (11.5) 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 8.9) (Hinweis: Satz 11.51 war hier nicht direkt anwendbar - warum?).
Um etwaigen Mäkeleien zuvorzukommen: Wer meint, dass das aber viel Aufwand für so einen mickrigen Reihenwert war, hat noch nie selbst versucht, einen (diesen) Reihenwert zu bestimmen.

Abschließend sei darauf hingewiesen, dass wir damit unser Ergebnis aus Beispiel 11.24 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 kommen nun zu einer weiteren, vielleicht eher überraschenden, Anwendung des Satzes von Taylor.

Satz 11.53. $\mathrm{e} \notin \mathbb{Q}$.

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 $n \geq 2$ 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

$$0 < \frac{\mathrm{e}^\xi}{n+1} < \frac{\mathrm{e}}{n+1} \leq \frac{\mathrm{e}}{3} < 1. \qquad \square$$

Schließlich wenden wir uns noch einmal dem Thema *Extremwerte* zu. Wir haben in Satz 11.13 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 relatives Extremum ist, brauchen wir genauere Hilfsmittel. Auch hier hilft der Satz von Taylor.

Satz 11.54. *Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $x_0 \in I$ und $f \in \mathrm{C}^n(I, \mathbb{R})$ 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 lokales Extremum, ist n gerade, so liegt in x_0 ein lokales Extremum vor, und zwar falls $f^{(n)}(x_0) > 0$ ein lokales Minimum und falls $f^{(n)}(x_0) < 0$ ein lokales 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 $x \in U_\delta(x_0)$ dasselbe 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 $f^{(n)}(\xi)$ dasselbe 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 dasselbe für $f^{(n)}(\xi)$ und damit ist

$$c(x) \begin{cases} \gtrless 0, & \text{falls } x \in (x_0, x_0 + \delta), \\ \lessgtr 0, & \text{falls } x \in (x_0 - \delta, x_0), \end{cases}$$

da Potenzieren mit einem ungeraden Exponenten das Vorzeichen erhält. Daraus folgt aber mit $f(x) = f(x_0) + c(x)$

$$f(x) \begin{cases} \gtrless f(x_0), & \text{falls } x \in (x_0, x_0 + \delta), \\ \lessgtr f(x_0), & \text{falls } x \in (x_0 - \delta, x_0). \end{cases}$$

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. □

12. Integration

12.1. Das Riemann-Integral

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 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, die von der x-Achse, den beiden Geraden $x = a$ und $x = b$ und dem Graphen der Funktion eingeschlossen wird.

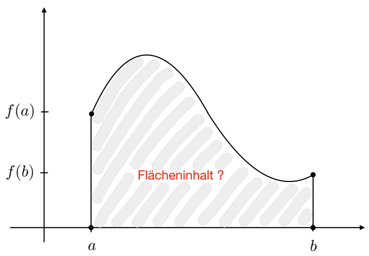

Abbildung 12.1.: gesuchter Flächeninhalt

Die grundlegende Idee der Integration nach Riemann ist es, die Fläche unter dem Graphen durch die Summation der Flächeninhalte von Rechtecken zu approximieren, die parallel zu den Koordinatenachsen liegen, wie im folgenden Bild skizziert:

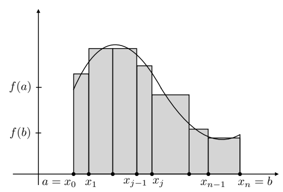

Abbildung 12.2.: Eine Approximation des Flächeninhalts durch Rechtecke über einer Zerlegung $Z = \{a = x_0, x_1, \ldots, x_{n-1}, x_n = b\}$ von $[a, b]$.

Wenn wir eine Weile über das Problem nachdenken, stellen wir fest, dass es ja eigentlich auch gar nicht viele andere Möglichkeiten gibt, einen unbekannten

Flächeninhalt zu bestimmen, als ihn durch *bekannte* Flächeninhalte anzunähern. Außerdem fällt uns auf, dass uns außer denen von Rechtecken gar nicht sooo viele Flächeninhalte bekannt sind. Dass sich zum Beispiel Kreise oder Dreiecke allerdings weniger gut eignen, um das Problem anzugehen, ist am Bild aber eigentlich auch schon klar (oder?). Die Riemannsche Idee erscheint also ganz natürlich.
Zu einer Approximation wie im Bild 12.4 gehört eine *Zerlegung* des Intervalls $[a, b]$ in Teilstücke. Als Analytiker fällt es uns jetzt leicht, zu sagen, dass die tatsächliche Fläche unter der Funktion durch den *Grenzwert* der *approximierenden Flächen* unter immer *feiner* werdenden Zerlegungen gegeben ist.
Die im obigen Absatz kursiv gedruckten Begriffe machen wir uns jetzt mathematisch präzise.

Definition 12.1. *Es seien $a, b \in \mathbb{R}$ mit $a < b$.*

(a) Eine Menge $Z := \{x_0, x_1, \dots, x_n\} \subseteq [a, b]$ heißt Zerlegung *des Intervalls $[a, b]$ genau dann, wenn gilt*

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

(b) Zu einer gegebenen Zerlegeung $Z = \{x_0, \dots, x_n\}$ des Intervalls $I := [a, b]$ schreiben wir $I_j := [x_{j-1}, x_j]$ für die dadurch entstehenden n Teilstücke von I und $|I_j| = x_j - x_{j-1}$ für ihre Länge.

(c) Den Wert

$$\mu := \mu_Z := \max_{j \in \{1,\dots,n\}} |I_j|$$

bezeichnen wir als die Feinheit *einer gegebenen Zerlegung Z von $[a, b]$.*

Beispiel 12.2. Auf dem Intervall $[a, b]$ ist zu $n \in \mathbb{N}$ die *äquidistante Zerlegung* $Z_n = \{x_0, \dots, x_n\}$ definiert durch

$$x_0 := a, \dots, x_k := a + k \cdot \frac{b-a}{n}, \dots, x_n = a + n\frac{b-a}{n} = b.$$

Da jedes Teilstück $I_j = [x_{j-1}, x_j]$ die Länge $|I_j| = \frac{b-a}{n}$ hat, ist $\frac{b-a}{n}$ auch die Feinheit der Zerlegung Z_n.

Wir können jetzt zu jeder gegebenen Zerlegung Z von $[a, b]$ eine Approximation des Flächeninhalts unter einer gegebenen Funktion $f\colon [a, b] \to \mathbb{R}$ definieren, wenn wir uns noch entscheiden, wo die Funktion in den Teilstücken I_j jeweils ausgewertet werden soll:

Definition 12.3. *Sei $f\colon [a, b] \to \mathbb{R}$ eine gegebene Funktion und Z eine Zerlegung von $[a, b]$.*

(a) In jedem I_j wählen wir einen Punkt ξ_j und setzen $\xi := (\xi_1, \xi_2, \dots \xi_n)$. Ein solches ξ nennen wir einen zu Z passenden Zwischenvektor.

(b) *Die zu einer Zerlegung Z und einem passenden Zwischenvektor ξ gegebene Approximation*

$$S_f(Z,\xi) := \sum_{j=1}^{n} f(\xi_j)|I_j|$$

des Flächeninhalts unter f heißt Riemannsche Summe*.*

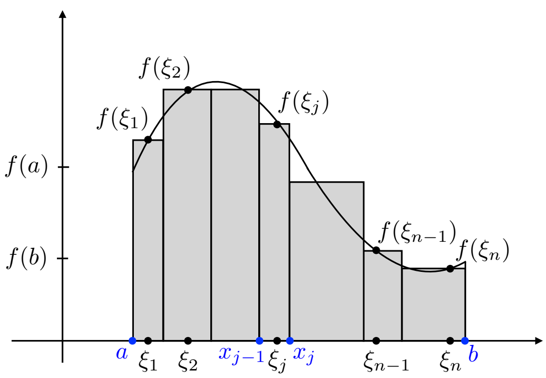

Abbildung 12.3.: Eine Darstellung der Auswertungspunkte ξ_j eines Zwischenvektors ξ zu einer Zerlegung Z von $[a,b]$ und der zugehörigen Funktionswerte $f(\xi_j)$. Die Riemannsche Summe $S_f(Z,\xi) = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$ entspricht dem grau schraffierten Flächeninhalt.

Beispiel 12.4. Es sei $a < b$ und $f\colon [a,b] \to \mathbb{R}$ eine gegebene Funktion.

(a) Zu einer gegebenen Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a,b]$ können wir die Auswertung von f am linken Rand, rechten Rand oder in der Mitte der I_j durch die Wahl der Zwischenvektoren

$$\begin{aligned}
\xi^l &:= (x_0, x_1, \ldots, x_{n-1}),\\
\xi^r &:= (x_1, x_2, \ldots, x_n), \text{ oder}\\
\xi^m &:= (\frac{x_0+x_1}{2}, \frac{x_1+x_2}{2}, \ldots, \frac{x_{n-1}+x_n}{2}),
\end{aligned}$$

realisieren.

(b) Ist zum Beispiel $f(x) = x$ auf $I = [0,1]$ gegeben und Z_n die äquidistante Zerlegung von $[0,1]$ mit Feinheit $\frac{1}{n}$, dann berechnen sich die Riemann-

Summen zu den entsprechenden Zwischenvektoren durch

$$S_f(Z_n, \xi^l) = \sum_{j=1}^{n} \frac{f(\xi_j^l)}{n} = \sum_{j=1}^{n} \frac{j-1}{n^2} = \frac{n(n-1)}{2n^2},$$
$$S_f(Z_n, \xi^r) = \sum_{j=1}^{n} \frac{f(\xi_j^r)}{n} = \sum_{j=1}^{n} \frac{j}{n^2} = \frac{n(n+1)}{2n^2},$$
$$S_f(Z_n, \xi^m) = \sum_{j=1}^{n} \frac{f(\xi_j^m)}{n} = \sum_{j=1}^{n} \frac{2j-1}{2n^2} = \frac{2n^2}{4n^2} = \frac{1}{2},$$

und alle liefern eine etwas andere Approximationsfolge für den tatsächlichen Flächeninhalt $\frac{1}{2}$ unter dem Graphen von f über $[0, 1]$ (wobei die Letzte hier eindeutig die Beste ist).

Damit eine gegebene Funktion *Riemann-integrierbar* ist, fordern wir jetzt, dass die Riemannschen Summen konvergieren, wenn wir nur die Zerlegungen beliebig fein wählen. Der Grenzwert ist dann das *Riemann-Integral*, bzw. der Flächeninhalt, wenn wir es mit Vorzeichen erstmal nicht so genau nehmen.

Definition 12.5. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to \mathbb{R}$ sei eine beschränkte Funktion. Dann heißt f* Riemann-integrierbar*, wenn es ein $S \in \mathbb{R}$ gibt, so dass für alle $\varepsilon > 0$ eine Feinheit $\delta > 0$ existiert, sodass für alle Zerlegungen Z mit Feinheit $\mu < \delta$ und alle passenden Zwischenvektoren ξ gilt:*

$$|S_f(Z, \xi) - S| < \varepsilon.$$

In diesem Fall schreiben wir $f \in R([a, b])$und nennen

$$\int_a^b f(x)\,\mathrm{d}x := S$$

das (Riemann-)Integral *von f auf $[a, b]$.*

Bemerkung 12.6. Diskussion und Beispiele:

- Diese Definition wirkt vielleicht etwas sperrig, denn tatsächlich möchte man sich ja nicht immer mit allen möglichen Zerlegungen und allen möglichen Zwischenvektoren rumschlagen, um zu zeigen, dass eine Funktion integrierbar ist. Wir untersuchen gleich noch eine äquivalente Definition des Integrals über *Ober-* und *Untersummen*, die vielleicht eleganter wirkt und die wir in der Theorie oft verwenden.

- Gleichzeitig ist die obige Definition unglaublich praktisch, wenn man schon weiß, dass eine Funktion integrierbar ist, und nur noch das Integral bestimmen will: sie sagt ja gerade, dass dann jede Approximationsfolge, die man sich ausdenken kann, gegen den richtigen Wert konvergiert, solange die Zerlegungen nur beliebig fein werden. Eine solche Strategie wäre zum Beispiel: Wähle zu jedem $n \in \mathbb{N}$ die äquidistanten Zerlegungen Z_n von $I = [a, b]$ der Feinheit $\frac{b-a}{n}$, jeweils mit Auswertung ξ^r immer am rechten Rand der Teilstücke. Dann folgt aus der Riemann-Integrierbarkeit von f:

$$\lim_{n\to\infty} S_f(Z_n, \xi^r) = S.$$

Ist ganz konkret $f : \mathbb{R} \to \mathbb{R}$ konstant mit Wert c, dann folgern wir so direkt

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \sum_{j=1}^{n} f(\xi_j^r)\frac{b-a}{n} = n \cdot c \cdot \frac{b-a}{n} = c \cdot (b-a).$$

Gilt $f(x) = x$ für $x \in \mathbb{R}$, dann können wir

$$\int_0^1 f(x)\,\mathrm{d}x = \lim_{n\to\infty} S_f(Z_n, \xi^m) = \frac{1}{2}$$

direkt aus Beispiel 12.4 folgern. Da diese beiden Flächeninhalte auch einfacher durch ein geometrisches Argument zu haben gewesen wären, war das ein guter erster Check dafür, dass das Riemann-Integral das Richtige liefert. Klar ist jetzt: das Integral ist ein *Grenzwert* – und mit denen kennen wir uns aus!

- Was die obige Defintion insbesondere sperrig macht, sind die beiden Allquantoren über den (ausreichend feinen) Zerlegungen und über den Zwischenvektoren bzw. Auswertungspunkten. Die sind aber gleichzeitig auch sinnvoll bis notwendig, denn sonst dürfte das Integral ja von der Wahl der Approximation über verschiedene Zerlegungen oder Zwischenvektoren abhängen. Dazu ein Beispiel:
  Die *Dirichletsche Sprungfunktion*

$$f(x) = \begin{cases} 1, \text{ falls } x \in [a, b] \cap \mathbb{Q} \\ 0, \text{ falls } x \in [a, b] \text{ und } x \notin \mathbb{Q}, \end{cases}$$

  ist nicht Riemann-integrierbar. Ist nämlich Z eine beliebige Zerlegung von $[a, b]$, dann liegen in jedem dadurch bestimmten Teilstück I_j von $[a, b]$ unendliche viele rationale und unendlich viele irrationale Zahlen. Inbesondere können wir also einerseits einen passenden Zwischenvektor ξ^{rat} aus rationalen Zahlen wählen und erhalten $S_f(Z, \xi^{\text{rat}}) = b - a > 0$, oder wir wählen andererseits einen Zwischenvektor ξ^{irr}, der nur aus irrationalen Zahlen besteht, und erhalten $S_f(Z, \xi^{\text{irr}}) = 0$. Da Z mit beliebiger Feinheit gegeben war, kann das Riemann-Kriterium nicht erfüllt werden.

Für eine alternative Perspektive auf das Riemann-Integral betrachten wir jetzt die folgenden Rechteckskonstruktionen, die den gesuchten Flächeninhalt von oben bzw. unten einschachteln:

Definition 12.7. *Es seien $a, b \in \mathbb{R}$ mit $a < b$, $Z = \{x_0, \ldots, x_n\}$ eine Zerlegung von $[a, b]$ und $f : [a, b] \to \mathbb{R}$ beschränkt.*

(a) Wir bezeichnen mit $m_j := \inf f(I_j)$ bzw. $M_j := \sup f(I_j)$ das Infimum bzw. Supremum von f über den I_j.

(b) Der Wert

$$U_f(Z) := \sum_{j=1}^{n} m_j |I_j|$$

heißt dann die Untersumme *von f zu Z und*

(c) der Wert

$$O_f(Z) := \sum_{j=1}^{n} M_j |I_j|$$

heißt die Obersumme *von f zu Z.*

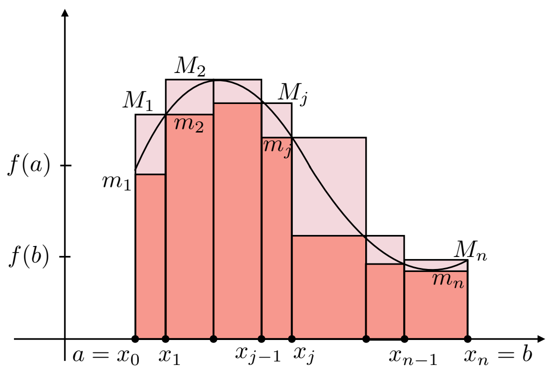

Abbildung 12.4.: Ober- und Untersumme schachteln den Graphen von f und damit den tatsächlichen Flächeninhalt von oben bzw. von unten ein. Das Darboux-Kriterium besagt, dass f integrierbar ist, wenn Zerlegungen gewählt werden können, sodass die Differenz von Ober- und Untersumme – also der helle Bereich – beliebig klein wird.

Offensichtlich ist wegen $m_j \leq M_j$ jeder Summand der Untersumme kleiner oder gleich dem entsprechenden Summanden der Obersumme, d. h. es gilt immer

$$U_f(Z) \leq O_f(Z). \tag{12.1}$$

Das überträgt sich zu der allgemeinen Abschätzung

$$\underline{S}_f := \sup_{\{Z:Z \text{ Zerlegung von } I\}} U_f(Z) \leq \inf_{\{Z:Z \text{ Zerlegung von } I\}} O_f(Z) =: \overline{S}_f$$

von „größtmöglicher“ Untersumme gegen „kleinstmögliche“ Obersumme (der Beweis ist eine Übungsaufgabe).Wegen (12.1) existieren dabei die Werte $\underline{S}_f$ und $\overline{S}_f$ zu jeder beschränkten Funktion f. Sie werden auch das *obere Integral* bzw. das *untere Integral* von f genannt.
Die Wahrheit - das Riemann-Integral, bzw. der gesuchte Flächeninhalt, - liegt also irgendwo dazwischen. Der folgende Satz sagt, dass f genau dann Riemann-integrierbar ist, wenn sich Ober- und Untersumme beliebig nahe kommen.

Satz 12.8. *Sei $a < b$ und $f\colon [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Dann sind äquivalent:*

(a) f ist Riemann-integrierbar, und,

(b) für alle $\varepsilon > 0$ existiert eine Zerlegung Z_ε von $[a,b]$ mit

$$O_f(Z_\varepsilon) - U_f(Z_\varepsilon) < \varepsilon.$$

Beweis. Da die Funktion f ja fest vorgegeben ist, verzichten wir im Folgenden der Übersichtlichkeit halber immer wieder auf das angestellte f bei der Bezeichnung der Ober-, Unter- und Riemann-Summen $O(Z), U(Z), S(Z,\xi)$.
Zuerst zeigen wir *(a) $\Rightarrow$ (b)*:
Nach Voraussetzung gibt es zu jedem $\varepsilon > 0$ ein Z, sodass für alle Zwischenvektoren ξ gilt:

$$S - \frac{\varepsilon}{4} < S(Z,\xi) < S + \frac{\varepsilon}{4}. \tag{12.2}$$

Durch geschickte Wahl der Zwischenvektoren können wir diese Abschätzung mit kleinen Abstrichen auf die Ober- und Untersumme über Z übertragen: zu jedem $M_j = \sup_{x \in I_j} f(x_j)$ existiert per Definition ein ξ_j^O, sodass $f(\xi_j^O) > M_j - \frac{\varepsilon}{4n}$ gilt, wobei $n \in \mathbb{N}$ die Zahl der Teilstücke der Zerlegung Z ist. Mit dieser Wahl für ξ^O auf jedem Teilstück erhalten wir

$$S(Z,\xi^O) = \sum_{j=1}^{n} |I_j| f(\xi_j^O) \geq \sum_{j=1}^{n} |I_j|(M_j - \frac{\varepsilon}{4n}) = O(Z) - \frac{\varepsilon}{4}.$$

Ein analoges Argument liefert einen Zwischenvektor ξ^U mit

$$S(Z,\xi^U) \leq U(Z) + \frac{\varepsilon}{4}.$$

Zusammen ergibt sich mit (12.2):

$$O(Z) - U(Z) \leq S(Z,\xi^O) - S(Z,\xi^U) + \frac{\varepsilon}{2} < \varepsilon.$$

Jetzt zur umgekehrten Richtung *(b) $\Rightarrow$ (a)*:
Nach Voraussetzung existiert zu jedem $\varepsilon > 0$ eine Zerlegung Z_ε von $[a, b]$, sodass

$$O(Z_\varepsilon) - U(Z_\varepsilon) < \frac{\varepsilon}{3} \tag{12.3}$$

gilt. Sei jetzt $n \in \mathbb{N}$ die Zahl der Teilstücke von Z_ε und sei

$$M := \sup_{x \in I} f(I) - \inf_{x \in I} f(I) > 0$$

(dieser Wert existiert, da f ja beschränkt ist). Wir zeigen jetzt, dass dann

$$O(Z) - U(Z) < \varepsilon, \qquad \text{für alle } Z \text{ mit Feinheit } \mu_Z < \delta := \frac{\varepsilon}{3nM} \tag{12.4}$$

gilt. Als Kandidaten für das Riemann-Integral wählen wir das untere Integral $S := \underline{S}_f$. Wegen der elementaren Abschätzung $U(Z) \leq S(Z, \xi) \leq O(Z)$ und wegen $U(Z) \leq S \leq O(Z)$ folgt dann auch

$$|S(Z, \xi) - S| \leq O(Z) - U(Z) < \varepsilon,$$

für alle diese Zerlegungen Z und beliebige passende Zwischenvektoren ξ, also die Behauptung.
Jetzt also zum Beweis von (12.4): wir definieren zu gegebenem Z mit Feinheit $\mu_Z < \delta$ und Z_ε die *gemeinsame Verfeinerung* $\tilde{Z} := Z \cup Z_\varepsilon$ von Z und Z_ε. Es handelt sich wieder um eine Zerlegung von $[a, b]$, die bis zu n der Teilstücke I_j von Z in jeweils $1 \leq n_j \leq n$ Teilstücke I_{j_i} zerlegt, indem alle in I_j liegende Elemente aus Z_ε hinzugenommen werden. Insofern ist $\tilde{Z}$ natürlich *feiner* als Z selbst und als Z_ε (oder genauso fein). Wir zerlegen jetzt die abzuschätzende Differenz,

$$O(Z) - U(Z) = [O(Z) - O(\tilde{Z})] + [O(\tilde{Z}) - U(\tilde{Z})] + [U(\tilde{Z}) - U(Z)], \tag{12.5}$$

und machen ein $\frac{\varepsilon}{3}$-Argument: um den ersten Term rechts abzuschätzen, halten wir fest, dass wir beim Verfeinern eines Teilstücks I_j von Z in der Obersumme vom Summanden $|I_j| M_j$ zu einem Summanden der Form $\sum_{i=1}^{n_j \leq n} |\tilde{I}_{j_i}| M_{j_i}$ übergehen, wobei sich die Längen der Teilstücke summieren, $\sum_{i=1}^{n_j} |\tilde{I}_{j_i}| = |I_j|$, und es gilt

$$M_j - M_{j_i} = \sup_{x \in I_j} f(x) - \sup_{x \in \tilde{I}_{j_i}} f(x) \leq M.$$

Wir haben also für jedes Teilstück I_j von Z, in dem sich durch die Verfeinerung etwas ändert, die Abschätzung

$$|I_j| M_j - \sum_{i=1}^{n_j} |\tilde{I}_{j_i}| M_{j_i} = (|I_j| - \sum_{i=1}^{n_j} |\tilde{I}_{j_i}|) M_j + \sum_{i=1}^{n_j} |\tilde{I}_{j_i}| (M_j - M_{j_i}) \leq |I_j| M < \delta M.$$

Da es höchstens n Teilstücke von Z gibt, auf denen sich durch die Verfeinerung etwas ändert, folgern wir daraus

$$O(Z) - O(\tilde{Z}) < n\delta M = \frac{\varepsilon}{3}.$$

Mit einem ganz analogen Argument schätzen wir die Untersummen im dritten Term rechts in (12.5) zu

$$U(\tilde{Z}) - U(Z) < \frac{\varepsilon}{3}$$

ab. Es bleibt nur noch, den mittleren Ausdruck rechts in (12.5) zu untersuchen: wegen der (trivialen) Abschätzungen

$$M_{j_i} = \sup_{x \in \tilde{I}^{\varepsilon}_{j_i}} f(x) \leq \sup_{x \in I_j} f(x) = M_j,$$
$$m_{j_i} = \inf_{x \in \tilde{I}^{\varepsilon}_{j_i}} f(x) \geq \inf_{x \in I_j} f(x) = m_j,$$

auf den Teilstücken I^{ε}_j, auf denen sich bei der Verfeinerung von Z_ε zu $\tilde{Z}$ etwas ändert, folgt mit (12.3) auch

$$O(\tilde{Z}) - U(\tilde{Z}) \leq O(Z_\varepsilon) - U(Z_\varepsilon) < \frac{\varepsilon}{3},$$

und damit die Behauptung. □

Bemerkung 12.9. Die Bedingung in *(b)* in Satz 12.8 liefert uns ein alternatives Kriterium für die Integrierbarkeit beschränkter Funktionen. Wir nennen es *Darboux-Kriterium.* Da dafür nur *eine* geeignete Folge von Zerlegungen untersucht/gefunden werden muss, ist es meistens praktischer als das Riemann-Kriterium, um eine Aussage aus der Theorie der Riemann-Integrale zu beweisen (das sehen wir in den folgenden Abschnitten). Um eine Ober- oder Untersumme zu einer gegebenen Zerlegung zu berechnen, ist aber typischerweise schon relativ viel Aufwand nötig – insbesondere müssen ja über jedem Teilstück sup und inf der Funktion gefunden werden. Zum Ausrechnen von Integralen eignet sich das Darboux-Kriterium also weniger gut. So hat jede Sichtweise ihre Vor- und Nachteile und das Gute an obigem Satz ist, dass wir jetzt beliebig zwischen Darbouxscher und Riemannscher Perspektive wechseln können.

Bemerkung 12.10. Reelle Integrale unterscheiden sich vor allem dadurch von Flächeninhalten, dass sie auch negative Vorzeichen haben können. Zum Beispiel zeigt ja eine schnelle Rechnung analog zum Beispiel in 12.6:

$$\int_{-1}^{0} x \, \mathrm{d}x = \lim_{n \to \infty} \sum_{j=1}^{n} \frac{-1 + \frac{j}{n}}{n} = -1 + \lim_{n \to \infty} \frac{n(n+1)}{2n^2} = -\frac{1}{2}.$$

Funktionen, die unterhalb der x-Achse liegen, bekommen also ein negatives Integral. (Wir verwenden in diesem Absatz ohne Beweis, dass die Funktion f mit $f(x) = x$ auf $[-1, 1]$ Riemann-integrierbar ist. Mehr dazu im nächsten Abschnitt.) Bestimmen wir das Integral zu einer Funktion, die mal positiv und mal negativ ist, dann können sich die so *signierten* Flächeninhalte wegheben. Zum Beispiel gilt ja:

$$\int_{-1}^{1} x \, \mathrm{d}x = 0.$$

Diese Effekte sind einfach immer zu beachten, wenn Flächeninhalte durch Integrale bestimmt werden. In diesem Zusammenhang treffen wir noch die folgende Vereinbarung: Ist $a < b$ und f über $[a, b]$ Riemann-integrierbar, dann definieren wir:

$$\int_{b}^{a} f(x) \, \mathrm{d}x := -\int_{a}^{b} f(x) \, \mathrm{d}x \quad \text{und} \quad \int_{a}^{a} f(x) \, \mathrm{d}x := 0.$$

12.2. Welche Funktionen sind Riemann-integrierbar?

Da es sich beim Integral um einen Grenzwert handelt, sind wir nicht überrascht, dass *Grenzwertsätze* gelten:

Satz 12.11. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f, g \in R([a, b])$, also Riemann-integrierbar. Dann gilt*

(a) (Monotonie) *Ist $f \leq g$ auf $[a, b]$, so gilt*

$$\int_{a}^{b} f(x) \, \mathrm{d}x \leq \int_{a}^{b} g(x) \, \mathrm{d}x.$$

(b) (Homogenität) *Ist $\alpha \in \mathbb{R}$ gegeben, so gilt $\alpha f \in R([a, b])$ und*

$$\int_{a}^{b} \alpha f(x) \, \mathrm{d}x = \alpha \int_{a}^{b} f(x) \, \mathrm{d}x.$$

(c) (Additivität) Es gilt auch $f + g \in R([a, b])$ mit

$$\int_{a}^{b} \big(f(x) + g(x)\big) \, \mathrm{d}x = \int_{a}^{b} f(x) \, \mathrm{d}x + \int_{a}^{b} g(x) \, \mathrm{d}x.$$

Beweis. (a) Ist $(Z_n)_n$ eine Folge von Zerlegungen von $[a, b]$ mit Feinheiten $\mu_n \to 0$ für $n \to \infty$ (die gibt es immer, zum Beispiel die äquidistanten Zerlegungen

Z_n mit Feinheit $\frac{b-a}{n}$) und sind ξ^{Z_n} dazu passende Zwischenvektoren, dann folgt aus $g \geq f$ die Ungleichung

$$S_g(Z_n, \xi^{Z_n}) - S_f(Z_n, \xi^{Z_n}) \geq 0.$$

Also folgt aus dem Riemann-Kriterium mit der Linearität und der Monotonie des Grenzwerts:

$$\begin{aligned}\int_a^b g(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x &= \lim_{n\to\infty} S_g(Z_n, \xi^{Z_n}) - \lim_{n\to\infty} S_f(Z_n, \xi^{Z_n}) \\ &= \lim_{n\to\infty} \Big(S_g(Z_n, \xi^{Z_n}) - S_f(Z_n, \xi^{Z_n})\Big) \geq 0,\end{aligned}$$

und damit die Behauptung.

(b) Wir folgern auch die Homogenität direkt mit dem Riemann-Kriterium: zu jedem $\varepsilon > 0$ und für alle Zerlegungen Z mit geeigneter Feinheit $\mu_Z < \delta$ und alle passenden Zwischenvektoren gilt nach Voraussetzung

$$\left|\alpha \int_a^b f(x)\,\mathrm{d}x - S_{\alpha f}(Z, \xi)\right| = |\alpha| \left|\int_a^b f(x)\,\mathrm{d}x - S_f(Z, \xi)\right| < |\alpha|\varepsilon,$$

und daraus folgt die Riemann-Integrierbarkeit von αf mit dem behaupteten Integral.

(c) auch hier funktioniert der Beweis direkt mit dem Riemann-Kriterium und verbleibt als Übungsaufgabe.

□

Aus der Homogenität und Additivität ergibt sich insbesondere auch die *Linearität* des Integrals. Weiter können wir auch Beträge, Produkte und Quotienten Riemann-integrierbarer Funktionen behandeln:

Satz 12.12. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ gegeben und f, g auf $[a, b]$ Riemann-integrierbar. Dann gelten*

(a) $|f|$ ist Riemann-integrierbar auf $[a, b]$ und es ist

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x. \qquad \text{Dreiecksungleichung für Integrale}$$

(b) Ist $|f(x)| \leq K$ für alle $x \in [a, b]$, so gilt die Standardabschätzung

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq K(b-a).$$

(c) $f \cdot g$ ist Riemann-integrierbar auf $[a, b]$.

(d) Ist $g(x) \neq 0$ für alle $x \in [a,b]$ und $1/g$ auf $[a,b]$ beschränkt, dann ist auch f/g auf $[a,b]$ Riemann-integrierbar.

Beweis. (a) Übungsaufgabe mit Hilfe von Satz 12.19.

(b) Aus (a) folgt direkt mit Hilfe des Beispiels in 12.6:

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x \leq \int_a^b K\,\mathrm{d}x = K(b-a).$$

(c) Überraschenderweise genügt es hierzu, zu zeigen, dass für jede Riemann-integrierbare Funktion φ die Funktion φ^2 Riemann-integrierbar ist. Denn wenn wir das gezeigt haben, können wir folgendermaßen argumentieren: Nach Satz 12.11 ((c)) sind die Funktionen $f+g$ und $f-g$ Riemann-integrierbar, also damit auch $(f+g)^2$ und $(f-g)^2$, womit schließlich wieder nach Satz 12.11 ((c)) auch

$$fg = \frac{1}{4}\big((f+g)^2 - (f-g)^2\big)$$

Riemann integrierbar ist.

Gehen wir also die Riemann-Integrierbarkeit des Quadrats mit Hilfe des Darboux-Kriteriums an. Wir setzen $D := \varphi([a,b])$. Da φ als integrierbare Funktion beschränkt sein muss, gibt es ein $K \geq 0$, so dass $D \subseteq [-K,K]$ gilt. Nach Voraussetzung gibt es zu jedem $\varepsilon > 0$ eine Zerlegung Z von $[a,b]$, sodass $O_\varphi(Z) - U_\varphi(Z) < \varepsilon$ gilt. Wegen der Monotonie der Funktion $x \mapsto x^2$ für $x \geq 0$ gilt außerdem

$$\begin{aligned}\sup_{x\in I_j} \varphi^2(x) &= \Big(\sup_{x\in I_j} \varphi(x)\Big)^2 =: M_j^2, \text{ und}\\ \inf_{x\in I_j} \varphi^2(x) &= \Big(\inf_{x\in I_j} |\varphi(x)|\Big)^2 =: |m_j|^2.\end{aligned}$$

Es folgt mit den üblichen Bezeichnungen:

$$\begin{aligned}O_{\varphi^2}(Z) - U_{\varphi^2}(Z) &= \sum_{j=1}^n |I_j|(M_j^2 - |m_j|^2) = \sum_{j=1}^n |I_j|(M_j + |m_j|)(M_j - |m_j|)\\ &\leq \sum_{j=1}^n |I_j| 2K(M_j - m_j) = 2K\big(O_\varphi(Z) - U_\varphi(Z)\big) < 2K\varepsilon,\end{aligned}$$

woraus die Riemann-Integrierbarkeit von φ^2 folgt.

(d) Übungsaufgabe mit Hilfe von (c) und Satz 12.19.

□

Verkettungen von Riemann-integrierbaren Funktionen sind leider etwas problematischer als zum Beispiel Verkettungen von stetigen oder differenzierbaren Funktionen (Erinnerung: was gilt da?). Wir behandeln sie am Ende dieses Abschnitts.

Nachdem wir mit obigen Sätzen gesehen haben, wie wir aus integrierbaren Funktionen neue Riemann-integrierbare Funktionen bauen können, wollen wir uns jetzt „Baumaterial“ beschaffen: die folgenden beiden Sätze zeigen, dass sowohl monotone Funktionen als auch stetige Funktionen einfach immer integrierbar sind. Man beachte dabei, dass monotone und stetige Funktionen auf kompakten Intervallen automatisch beschränkt sind.

Satz 12.13. *Sind $a, b \in \mathbb{R}$ mit $a < b$ und ist $f : [a, b] \to \mathbb{R}$ eine monotone Funktion, dann ist f auf $[a, b]$ Riemann-integrierbar.*

Beweis. Wir führen den Beweis im Fall, dass f monoton wachsend ist. Für monoton fallende Funktionen geht die Argumentation analog. Wir verwenden das Darboux-Kriterium. Wir verwenden dazu die äquidistante Zerlegung $Z_n = \{x_0, \dots, x_n\}$ von $[a, b]$ mit

$$x_j := a + j\frac{b-a}{n}, \quad j = 0, \dots, n.$$

Wegen der Monotonie von f gilt sofort $m_j = f(x_{j-1})$ und $M_j = f(x_j)$. Damit haben wir für die Differenz von Ober- und Untersumme

$$O(Z_n) - U(Z_n) = \sum_{j=1}^{n} (M_j - m_j)|I_j| = \frac{b-a}{n} \sum_{j=1}^{n} \big(f(x_j) - f(x_{j-1})\big).$$

Diese Summe ist nun eine Teleskopsumme, die sich verkürzt auf

$$O(Z_n) - U(Z_n) = \frac{b-a}{n}(f(b) - f(a)) =: d_n \longrightarrow 0 \quad (n \to \infty)$$

Für jedes $\varepsilon > 0$ gibt es also ein $n \in \mathbb{N}$ mit $d_n < \varepsilon$. Für die zugehörige Zerlegung Z_n gilt dann $O_f(Z_n) - U_f(Z_n) < \varepsilon$, also ist das Darboux-Kriterium erfüllt und f ist nach Satz 12.8 Riemann-integrierbar. □

Satz 12.14. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ gegeben. Dann ist jede auf $[a, b]$ stetige Funktion dort auch Riemann-integrierbar. Anders ausgedrückt: $C([a, b]) \subseteq R([a, b])$.*

Beweis. Sei f auf $[a, b]$ stetig. Dann ist f dank der Kompaktheit des Intervalls nach Satz 10.27 automatisch beschränkt. Es macht also Sinn, über Integrierbarkeit nachzudenken. Sei dazu ein $\varepsilon > 0$ gegeben. Wegen der Kompaktheit des Intervalls ist f außerdem sogar gleichmäßig stetig (vgl. Satz 10.53). Es gibt also ein $\delta > 0$, so dass

$$|f(x) - f(y)| < \varepsilon \qquad \text{für alle } x, y \in [a, b] \text{ mit } |x - y| < \delta.$$

Sei nun eine Zerlegung $Z = \{x_0, \ldots, x_n\}$ von $[a, b]$ so gewählt, dass $|I_j| < \delta$ für alle $j \in \{1, \ldots, n\}$ gilt. Fixieren wir nun vorläufig ein $j \in \mathbb{N}$, so ist f eine stetige Funktion auf I_j, nimmt dort also nach Satz 10.27 ihr Maximum und Minimum an, d. h. es gibt $\xi, \eta \in I_j$ mit $m_j = f(\xi)$ und $M_j = f(\eta)$ und da beide in I_j liegen, gilt auf jeden Fall $|\xi - \eta| \leq |I_j| < \delta$. Nun garantiert uns die gleichmäßige Stetigkeit von f damit aber

$$M_j - m_j = |M_j - m_j| = |f(\eta) - f(\xi)| < \varepsilon.$$

Also ist

$$O_f(Z) - U_f(Z) = \sum_{j=1}^{n} (M_j - m_j)|I_j| < \varepsilon \sum_{j=1}^{n} |I_j| = \varepsilon(b - a),$$

und damit das Darboux-Kriterium erfüllt. □

Wir wollen nun zeigen, dass man die Voraussetzung der Stetigkeit noch ein wenig abschwächen kann, was letztlich nur unserer Intuition entspricht, das Funktionen stückweise integriert werden können. Dazu das folgende Lemma:

Lemma 12.15. *Es seien $a, b, c \in \mathbb{R}$ mit $a < c < b$ und $f : [a, b] \to \mathbb{R}$ sei beschränkt. Dann ist f auf $[a, b]$ genau dann Riemann-integrierbar, wenn f auf $[a, c]$ und $[c, b]$ Riemann-integrierbar ist und in diesem Falle gilt*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

Beweis. Wir zeigen zuerst, dass sich die Integrierbarkeit von ganzem Intervall auf die Teile und andersrum überträgt. Danach beweisen wir, dass die Integrale übereinstimmen.
Zuerst zu „$\Leftarrow$“:
Wir verwenden das Darboux-Kriterium. Sei $\varepsilon > 0$. Nach Voraussetzung existieren dann Zerlegungen $Z_a = \{a, x_1, \ldots, x_{k-1}, c =: x_k\}$ von $[a, c]$ und $Z_b = \{c, x_{k+1}, \ldots, x_{n-1}, x_n := b\}$ von $[c, b]$, sodass $O_a(Z_a) - U_a(Z_a) < \frac{\varepsilon}{2}$ und $O_b(Z_b) - U_b(Z_b) < \frac{\varepsilon}{2}$ gelten. Damit ist aber auch $Z := Z_a \cup Z_b = \{x_0, \ldots, x_n\}$ eine Zerlegung von $[a, b]$ und es folgt direkt

$$O(Z) - U(Z) = O(Z_a) + O(Z_b) - U(Z_a) - U(Z_b) < \varepsilon.$$

Jetzt zu „$\Rightarrow$“:
Wir verwenden wieder das Darboux-Kriterium. Nach Voraussetzung gibt es zu jedem $\varepsilon > 0$ eine Zerlegung Z von $[a, b]$, sodass $O(Z) - U(Z) < \varepsilon$ gilt. Wir unterscheiden zwei Fälle: ist erstens c bereits in $Z = \{a = x_0, \ldots, x_k = c, \ldots, x_n = b\}$ enthalten, dann sind $Z_a := \{x_0, \ldots, x_k\}$ und $Z_b := \{x_k, \ldots, x_n\}$ Zerlegungen von $[a, c]$ bzw. $[c, b]$ und es gilt

$$O(Z_a) - U(Z_a) + O(Z_b) - U(Z_b) = O(Z) - U(Z) < \varepsilon.$$

Aus der Positivität der beiden Differenzen auf der linken Seite folgt daraus auch die Behauptung. Wäre zweitens c noch nicht in Z enthalten, dann erhalten wir mit

$$\tilde{Z} := Z \cup \{c\}$$

eine Zerlegung $\tilde{Z}$ von $[a,b]$, die c enthält. Wie im ersten Fall können wir jetzt $\tilde{Z}$ in zwei Zerlegungen $\tilde{Z}_a, \tilde{Z}_b$ von $[a,c], [c,b]$ zerteilen und erhalten

$$O(\tilde{Z}_a) - U(\tilde{Z}_a) + O(\tilde{Z}_b) - U(\tilde{Z}_b) = O(\tilde{Z}) - U(\tilde{Z}).$$

Da es sich bei $\tilde{Z}$ außerdem um eine Verfeinerung von Z handelt, gelten die Abschätzungen $O(\tilde{Z}) \leq O(Z)$ und $U(\tilde{Z}) \geq U(Z)$ (Übung!), also auch

$$O(\tilde{Z}) - U(\tilde{Z}) \leq O(Z) - U(Z) < \varepsilon,$$

und damit die Behauptung.
Jetzt zur Gleichheit der Integrale:
Sei $(Z^n)_n$ eine Folge von Zerlegungen von $[a,b]$, sodass die Feinheiten $\mu_{Z^n} \to 0$ für $n \to \infty$ erfüllen. Wir wählen die Z^n außerdem so, dass sie alle den Punkt c enthalten, sich also wie oben in die Zerlegungen Z_a^n und Z_b^n von $[a,c]$ bzw. $[c,b]$ unterteilen lassen. Das geht, weil wir den Punkt c ja immer hinzufügen können und dadurch die Feinheit höchstens verkleinern. Dann gilt für alle Zwischenvektoren ξ^n passend zu Z^n wegen der Linearität des Grenzwerts und der Riemann-Integrierbarkeit auf beiden Seiten der Gleichung im Lemma:

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x &= \lim_{n\to\infty} S_f(Z^n,\xi^n) = \lim_{n\to\infty}\left(S_f(Z_a^n,\xi_a^n) + S_f(Z_b^n,\xi_b^n)\right)\\ &= \lim_{n\to\infty} S_f(Z_a^n,\xi_a^n) + \lim_{n\to\infty} S_f(Z_b^n,\xi_b^n) = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x,\end{aligned}$$

also die Gleichheit der Integrale. □

Satz 12.16. *Es seien $a,b \in \mathbb{R}$ mit $a < b$ gegeben und $f : [a,b] \to \mathbb{R}$ sei eine beschränkte Funktion, die nur in endlich vielen Punkten aus $[a,b]$ nicht stetig ist. Dann ist f Riemann-integrierbar auf $[a,b]$.*

Beweis. Im gesamten Beweis seien $m := \inf f([a,b]) \leq M := \sup f([a,b])$ die größte untere bzw. kleinste obere Schranke an f. Dann gilt auch $|f(x)| \leq K := \max(|m|,|M|)$ für alle $x \in [a,b]$.
Wir betrachten zuerst den Speziallfall, dass f nur in genau einem Punkt $y \in (a,b)$ unstetig ist. Es sei dazu ein $\varepsilon > 0$ vorgegeben. Dann wählen wir ein $c \in (a,y)$ und ein $d \in (y,b)$ so, dass $2K(d-c) < \varepsilon/3$ ist. Nun ist $f|_{[a,c]}$, also die Einschränkung von f auf das Intervall $[a,c]$ eine dort stetige Funktion, die nach Satz 12.14 integrierbar ist. Es gibt dazu also eine Zerlegung Z_a von $[a,c]$ mit

$$O(Z_a) - U(Z_a) < \frac{\varepsilon}{3}.$$

Mit der gleichen Argumentation gibt es für $f|_{[d,b]}$ eine Zerlegung Z_b von $[d, b]$, so dass
$$O(Z_b) - U(Z_b) < \frac{\varepsilon}{3}.$$
gilt. Nun ist auch hier $Z := Z_a \cup Z_b$ eine Zerlegung von $[a, b]$. Diese hat eben nur keinen Zerlegungspunkt in der kritischen Region zwischen c und d. Für diese Zerlegung gilt dann

$$\begin{aligned} O(Z) - U(Z) &\leq O(Z_a) + M(d-c) + O(Z_b) - \big[U(Z_a) + m(d-c) + U(Z_b)\big] \\ &= O(Z_a) - U(Z_a) + (M-m)(d-c) + O(Z_b) - U(Z_b) \\ &\leq \frac{\varepsilon}{3} + 2K(d-c) + \frac{\varepsilon}{3} < \varepsilon, \end{aligned}$$

und wir erhalten in diesem Fall Riemann-Integrabilität von f.
Als zweiten Fall beschäftigen wir uns damit, dass die (weiterhin einzige) Unstetigkeitsstelle von f am Rand des Intervalls, also in a (oder b) liegt. In diesem Fall wählen wir nur ein c (oder d) aus $[a, b]$ mit $2K(c-a) < \varepsilon/2$ (oder $2K(b-d) < \varepsilon/2$) und argumentieren genau so wie oben mit der Zerlegung $Z := Z_a \cup \{a\}$ (oder $Z := Z_b \cup \{b\}$).
Wir wenden uns also nun dem allgemeinen Fall von endlich vielen Unstetigkeitsstellen zu. Wir nummerieren diese als $y_1, y_2, \ldots, y_p \in [a, b]$ mit $y_1 < y_2 < \cdots < y_p$ durch und wählen Zahlen $\eta_1, \eta_2, \ldots, \eta_{p-1}$, so dass

$$y_1 < \eta_1 < y_2 < \eta_2 < y_3 < \cdots < \eta_{p-1} < y_p$$

gilt. Dann ist nach den bisherigen Überlegungen in diesem Beweis f auf jedem der Intervalle $[a, \eta_1]$, $[\eta_j, \eta_{j+1}$ für jedes $j = 1, \ldots p-2$ und $[\eta_{p-1}, b]$ integrierbar, denn in jedem dieser Intervalle liegt genau eine Unstetigkeitsstelle. Die Behauptung folgt nun aus Lemma 12.15 über die Möglichkeit, Integrale zu zerteilen. □

Wir kommen nun zu einer wichtigen Folgerung dieses Satzes, die man etwas flapsig so beschreiben kann: Das Integral sieht Mengen, die aus endlich vielen Punkten bestehen, nicht.

Satz 12.17. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und eine Riemann-integrierbare Funktion $f : [a, b] \to \mathbb{R}$, sowie eine beschränkte Funktion $g : [a, b] \to \mathbb{R}$ gegeben. Ist dann die Menge $P := \{x \in [a, b] : f(x) \neq g(x)\}$ endlich, so ist auch $g \in R([a, b])$ und es gilt*
$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x.$$
Insbesondere sind die Integrale gleich, falls $f(x) = g(x)$ für alle $x \in (a, b)$ gilt.

Beweis. Wir setzen $h := f - g$ und wie im letzten Beweis $P = \{y_1, y_2, \ldots, y_p\}$. Dann gilt $h(x) = 0$ für alle $x \in [a, b] \setminus P$. Also ist h nur genau in den Punkten

von P unstetig, und damit nach Satz 12.16 integrierbar. Damit ist aber sofort auch $g = f - h$ integrierbar (vgl. Satz 12.11). Wir müssen also nur noch zeigen, dass $\int_a^b h(x)\,\mathrm{d}x = 0$ gilt.
Um Integrale auszurechnen, eignet sich das Riemann-Kriterium: Wir wählen eine Folge von Zerlegungen $(Z_n)_n$ mit Feinheiten $\mu_{Z_n} \to 0$ für $n \to \infty$ Da h nur in endlich vielen Punkten $\neq 0$ ist, gibt es zu jeder Zerlegung Z_n von $[a, b]$ auch einen Zwischenvektor ξ^{Z_n}, der h nur an Nullstellen auswertet. Dadurch haben dann auch alle Riemann-Summen der Folge den Wert Null. Aus dem Riemann-Kriterium folgt also

$$\int_a^b h(x)\,\mathrm{d}x = \lim_{n\to\infty} S(Z_n, \xi^{Z_n}) = \lim_{n\to\infty} 0 = 0.$$

□

Wir wollen uns zum Schluss dieses Abschnitts noch mit der Integrierbarkeit der Verkettung zweier Riemann-integrierbarer Funktionen auseinandersetzen. Es gibt dabei ein negatives und ein positives Resultat:

- Beispiel 12.18 unten zeigt, dass die Verkettung von Riemann-integrierbaren Funktionen im Allgemeinen nicht Riemann-integrierbar ist.
- Satz 12.19 besagt, dass die Verkettung eine Lipschitz-stetigen Funktion mit einer Riemann-integrierbaren Funktion wieder Riemann-integrierbar ist.

Beispiel 12.18. Auf dem Intervall $[0, 1]$ sei

$$f(x) = \begin{cases} 1, \text{ falls } x = 0 \\ 0, \text{ falls } x \in [0,1] \text{ und } x \notin \mathbb{Q}, \\ 1/q, \text{ falls } x \in (0,1] \cap \mathbb{Q} \text{ und } x = p/q \text{ mit } p, q \in \mathbb{N} \text{ teilerfrei.} \end{cases}$$

Auf Übungsblatt 12 wird gezeigt, dass f auf $[0, 1]$ Riemann-integrierbar ist. Insbesondere haben wir hier ein Beispiel für eine Funktion, die Riemann-integrierbar ist, aber an mehr als endlich vielen Stellen unstetig. Satz 12.16 liefert also ein hinreichendes, aber kein notwendiges Kriterium für Integrierbarkeit.
Wir betrachten weiter die Funktion g mit

$$g(x) = \begin{cases} 0, \text{ falls } x = 0, \\ 1, \text{ falls } x \in (0,1]. \end{cases}$$

Dann ist g monoton wachsend, also auch Riemann-integrierbar. Für die Verkettung von g und f gilt aber

$$(g \circ f)(x) = g(f(x)) = \begin{cases} 1, \text{ falls } x \in [0,1] \cap \mathbb{Q}, \\ 0, \text{ falls } x \in [0,1] \text{ und } x \notin \mathbb{Q} \end{cases}$$

und dies ist die Dirichletsche Sprungfunktion, die nach dem Beispiel in 12.6 nicht Riemann-integrierbar ist.

Es stellt sich also die Frage, wie die Voraussetzungen verschärft werden können, damit die Verkettung zweier Funktionen integrierbar wird. Eine mögliche Antwort gibt der folgende Satz. Der Beweis ist Übungsaufgabe.

Satz 12.19. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f \in R([a, b])$ gegeben. Bezeichnen wir mit $D := f([a, b])$ das Bild von f und ist $h : D \to \mathbb{R}$ eine auf D Lipschitz-stetige Funktion, so gilt auch $h \circ f \in R([a, b])$.*

12.3. Eigenschaften integrierbarer Funktionen

Eine sehr wichtige und nützliche Eigenschaft der Integration ist, dass sie die Differenziation oft nur rückgängig macht, und umgekehrt. Um das präzise zu machen, benötigen wir noch den folgenden Begriff:

Definition 12.20. *Es seien $I \subseteq \mathbb{R}$ ein Intervall und $g, G : I \to \mathbb{R}$ Funktionen. Dann heißt G eine* Stammfunktion *von g, falls G auf I differenzierbar ist und $G' = g$ auf I gilt.*

Nach Satz 11.16 sind Stammfunktionen bis auf eine Konstante eindeutig bestimmt, d. h. sind G und H Stammfunktionen von g auf I, so gibt es eine Konstante $c \in \mathbb{R}$, so dass $G = H + c$ auf I gilt. Wir können nun das Hauptresultat dieses Abschnitts beweisen.

Satz 12.21. *[Hauptsatz der Differenzial- und Integralrechnung] Es seien $a, b \in \mathbb{R}$ mit $a < b$ gegeben. Ist $f \in R([a, b])$ und $F \in C[a, b]$, so dass F auf (a, b) eine Stammfunktion von f ist, so gilt*

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) =: F\Big|_a^b.$$

Beweis. Da wir ein Integral berechnen wollen, eignet sich das Riemann-Kriterium. Wir halten zunächst fest, dass es zu jeder Zerlegung Z von $[a, b]$ nach dem Mittelwertsatz für jedes Teilstück I_j ein $\xi_j \in (x_{j-1}, x_j)$ gibt mit

$$F(x_j) - F(x_{j-1}) = F'(\xi_j)(x_j - x_{j-1}) = f(\xi_j)(x_j - x_{j-1}) = f(\xi_j)|I_j|.$$

Ist also $(Z^n)_n$ eine Folge von Zerlegungen von $[a, b]$ mit je $n \in \mathbb{N}$ Teilstücken und Feinheiten $\mu_{Z^n} \to 0$ für $n \to \infty$, können wir dazu wie oben jeweils den Zwischenvektor $\xi^n = (\xi_1^n, \dots, \xi_n^n)$ wählen und erhalten

$$S_f(Z^n, \xi^n) = \sum_{j=1}^{n} f(\xi_j^n)|I_j^n| = \sum_{j=1}^{n} F(x_j^n) - F(x_{j-1}^n) = F(b) - F(a),$$

weil es sich um eine Teleskopsumme handelt. Wir haben also eine konstante Approximationsfolge und folgern aus der Riemann-Integrierbarkeit von f,

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} S_f(Z^n, \xi^n) = F(b) - F(a).$$

□

Beispiel 12.22. (a) Es seien $0 < a < b$ und

$$f(x) = \frac{1}{x}, \quad x \in [a, b].$$

Dann ist f als stetige Funktion Riemann-integrierbar. Außerdem ist $F(x) = \ln(x)$ eine Stammfunktion von f, denn $F'(x) = f(x)$ für alle $x \in [a, b]$. Damit haben wir mit dem Hauptsatz

$$\int_a^b \frac{1}{x}\, \mathrm{d}x = \ln(x)\Big|_a^b = \ln(b) - \ln(a).$$

Die Berechnung des Integrals ist jetzt also viel einfacher geworden als vielleicht „zu Fuß“ über die Riemann-Summen.

(b) Auf dem Intervall $[0, \pi]$ betrachten wir die Funktion $f(x) = \cos(x)$. Aufgrund ihrer Stetigkeit ist sie Riemann-integrierbar. Eine Stammfunktion ist $F(x) = \sin(x)$, also gilt

$$\int_0^\pi \cos(x)\, \mathrm{d}x = \sin(x)\Big|_0^\pi = 0 - 0 = 0.$$

Im Zusammenhang mit dem Hauptsatz der Differenzial- und Integralrechnung sind noch zwei Warnungen angebracht.

Warnung 12.23. (a) Es gibt Funktionen, die Stammfunktionen besitzen, aber nicht integrierbar sind. (Das ist kein Widerspruch zum Hauptsatz, denn dort haben wir ja die Integrierbarkeit von f vorausgesetzt!) 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 11.43 (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 Riemann-integrierbar sein, aber F ist natürlich eine Stammfunktion von f, denn so haben wir f ja gerade konstruiert.

(b) Andersherum gibt es auch integrierbare 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 monoton wachsend und damit Riemann-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.

Ist f auf dem Intervall $[a,b]$ integrierbar und wählen wir ein beliebiges $x \in [a,b]$, so ist f nach Lemma 12.15 auch auf $[a,x]$ integrierbar. Für jedes $x \in [a,b]$ können wir also die Abbildung

$$x \mapsto \int_a^x f(t)\,\mathrm{d}t.$$

betrachten. Mit den Eigenschaften der so gegebenen Funktion, die nicht notwendigerweise eine Stammfunktion ist, beschäftigt sich der folgende Satz.

Satz 12.24. *[Zweiter Hauptsatz der Differenzial- und Integralrechnung] Es seien $a, b \in \mathbb{R}$ mit $a < b$ gegeben und $f : [a,b] \to \mathbb{R}$ sei auf $[a,b]$ integrierbar. Dann gelten für die Funktion $F : [a,b] \to \mathbb{R}$, gegeben durch*

$$F(x) := \int_a^x f(t)\,\mathrm{d}t, \qquad x \in [a,b]$$

die folgenden Aussagen:

(a) F ist Lipschitz-stetig auf $[a,b]$ und es gilt insbesondere $F \in C([a,b])$.

(b) Ist f stetig in $x_0 \in [a,b]$, so ist F in x_0 differenzierbar und es gilt $F'(x_0) = f(x_0)$.

(c) Ist $f \in C([a,b])$, so ist $F \in C^1([a,b])$ und für alle $x \in [a,b]$ gilt $F'(x) = f(x)$, d. h. F ist eine Stammfunktion von f auf $[a,b]$.

Beweis. (a) Da f als integrierbare Funktion auf jeden Fall beschränkt ist, gibt es eine Zahl $L \geq 0$, so dass $|f(x)| \leq L$ für alle $x \in [a,b]$ gilt. Seien nun $x, y \in$

$[a,b]$. Gilt dabei $x < y$, so haben wir unter Zuhilfenahme von Lemma 12.15

$$\begin{aligned} |F(y) - F(x)| &= \left| \int_a^y f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t \right| \\ &= \left| \int_a^x f(t)\,\mathrm{d}t + \int_x^y f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t \right| = \left| \int_x^y f(t)\,\mathrm{d}t \right| \\ &\leq \int_x^y |f(t)|\,\mathrm{d}t \leq \int_x^y L\,\mathrm{d}t = L(y-x) = L|y-x|. \end{aligned}$$

Für den Fall $x > y$ kann man analog armguentieren, indem man die Rollen von x und y vertauscht.

(b) Wir betrachten zunächst den Fall, dass $x_0 \in [a,b)$ ist. Sei dann ein $h > 0$ mit $x_0 + h \leq b$ gegeben. Damit gilt wie im Beweis des vorherigen Punkts

$$\frac{F(x_0+h) - F(x_0)}{h} = \frac{1}{h} \int_{x_0}^{x_0+h} f(t)\,\mathrm{d}t$$

Um nun die Konvergenz dieses Differenzenquotienten für $h \to 0+$ gegen $f(x_0)$ zu beweisen, beobachten wir zunächst, dass gilt

$$\frac{1}{h} \int_{x_0}^{x_0+h} f(x_0)\,\mathrm{d}t = \frac{f(x_0)}{h} \int_{x_0}^{x_0+h} \mathrm{d}t = \frac{f(x_0)}{h}(x_0 + h - x_0) = f(x_0).$$

Also ist

$$\begin{aligned} \left| \frac{F(x_0+h) - F(x_0)}{h} - f(x_0) \right| &= \left| \frac{1}{h} \int_{x_0}^{x_0+h} f(t)\,\mathrm{d}t - \frac{1}{h} \int_{x_0}^{x_0+h} f(x_0)\,\mathrm{d}t \right| \\ &= \frac{1}{h} \left| \int_{x_0}^{x_0+h} \big(f(t) - f(x_0)\big)\,\mathrm{d}t \right| \\ &\leq \frac{1}{h} \int_{x_0}^{x_0+h} \big|f(t) - f(x_0)\big|\,\mathrm{d}t \end{aligned}$$

Sein nun ein $\varepsilon > 0$ gegeben. Da f stetig in x_0 ist, gibt es dann ein $\delta > 0$, so dass $x_0 + \delta \leq b$ ist und $|f(t) - f(x_0)| < \varepsilon$ ist für alle $t \in [x_0, x_0 + \delta]$. Betrachten wir ab nun nur noch solche $h > 0$, die kleiner als δ sind, so haben wir damit $|f(t) - f(x_0)| \leq \varepsilon$ und also auch

$$\left| \frac{F(x_0+h) - F(x_0)}{h} - f(x_0) \right| \leq \varepsilon$$

für alle $h \in (0, \delta)$. Also gilt

$$\lim_{h \to 0+} \frac{F(x_0+h) - F(x_0)}{h} = f(x_0).$$

Mit einer analogen Betrachtung kann man für alle $x \in (a,b]$ zeigen, dass

$$\lim_{h\to 0-} \frac{F(x_0+h)-F(x_0)}{h} = f(x_0)$$

ist. Also haben wir nun im Inneren des Intervalls, also für alle $x_0 \in (a,b)$ auch den vollen Grenzwert

$$\lim_{h\to 0} \frac{F(x_0+h)-F(x_0)}{h} = f(x_0).$$

(c) folgt aus (a) und (b).

□

Weiter geht es mit wichtigen Sätzen über integrierbare Funktionen:

Satz 12.25 (Mittelwertsätze für Integrale)**.** *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und zwei Funktionen $f, g \in R([a,b])$ gegeben, so dass g auf $[a,b]$ keinen Vorzeichenwechsel hat, d. h. es gilt $g(x) \geq 0$ für alle $x \in [a,b]$ oder $g(x) \leq 0$ für alle $x \in [a,b]$. Weiter bezeichnen wir mit $m := \inf f([a,b])$ und $M := \sup f([a,b])$. Dann gilt*

(a) (1. Mittelwertsatz) *Es gibt ein $\mu \in [m, M]$, so dass gilt*

$$\int_a^b f(x)\,\mathrm{d}x = \mu(b-a).$$

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

(b) (2. Mittelwertsatz) *Es gibt ein $\mu \in [m, M]$, so dass gilt*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = \mu \int_a^b g(x)\,\mathrm{d}x.$$

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

Beweis. Da der Teil (a) direkt aus der Aussage in (b) mit $g = 1$ folgt, beweisen wir nur diesen zweiten Teil. Teil (a) ist anschaulich auch völlig klar: Es gibt ein Rechteck der Höhe $m \leq \mu \leq M$ und Breite $b - a$, das den gleichen Flächeninhalt hat wie die Fläche unter f. Oder anders gesagt: es gibt den mittleren Wert μ der Funktion f über $[a,b]$, s. Bild:
Jetzt zum Beweis der Aussage in (b) im Fall, dass $g(x) \geq 0$ für alle $x \in [a,b]$ gilt (der Beweis für negatives g verläuft analog): Da $m \leq f(x) \leq M$ für alle $x \in [a,b]$ gilt, haben wir dank der Positivität von g auch $mg(x) \leq g(x)f(x) \leq Mg(x)$ für alle diese x. Also gilt wegen der Monotonie des Integrals

$$m\int_a^b g(x)\,\mathrm{d}x \leq \int_a^b f(x)g(x)\,\mathrm{d}x \leq M\int_a^b g(x)\,\mathrm{d}x.$$

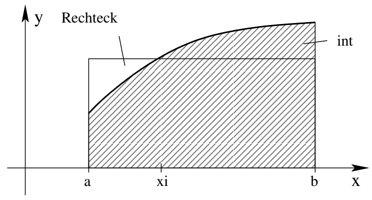

Abbildung 12.5.: Zum 1. Mittelwertsatz der Integralrechnung.

Ist nun das Integral von g über $[a, b]$ Null, so folgt aus dieser Ungleichungskette sofort, dass auch das Integral über das Produkt von f und g Null sein muss und die Aussage des Satzes ist für jedes $\mu \in \mathbb{R}$ richtig. Ist der Wert des Integrals von g nicht Null, so setzen wir

$$\mu := \frac{\int_a^b f(x)g(x)\,\mathrm{d}x}{\int_a^b g(x)\,\mathrm{d}x}.$$

Damit gilt auf jeden Fall die im Satz behauptete Gleichheit. Da g nicht-negativ ist, kann auch das Integral von g auf $[a, b]$ nicht negativ sein, es muss hier also sogar strikt positiv sein, so dass wir aus unserer Ungleichungskette nach Division durch dieses Integral auch noch bekommen, dass $\mu \in [m, M]$ liegt.
Damit bleibt noch die Zusatzaussage für stetige Funktionen f zu beweisen. Ist $f \in C([a, b])$, so existieren dank der Kompaktheit von $[a, b]$ Zahlen $x_1, x_2 \in [a, b]$ mit $m = f(x_1)$ und $M = f(x_2)$. Da $\mu \in [m, M]$ ist, gibt es also nach dem Zwischenwertsatz (vgl. Satz 10.22) ein $\xi \in [a, b]$ mit $f(\xi) = \mu$. □

Im Rest dieses Abschnitts wollen wir uns mit der Frage beschäftigen, ob man bei Funktionenfolgen und -reihen den Grenzübergang mit der Integration vertauschen darf, ob also das Integral der Grenzfunktion der Grenzwert der Integrale der Funktionenfolgenglieder sind. Da es sich bei der Integration ja auch um einen Grenzwertprozess handelt ist das also mal wieder eine Frage nach der Vertauschbarkeit zweier Grenzübergänge und damit steht zu befürchten, dass es eine eher knifflige Frage ist. Tatsächlich ist das gewünschte im Allgemeinen falsch, wie das folgende Beispiel zeigt:

Beispiel 12.26. Wir betrachten auf dem Intervall $[0, 1]$ die Funktionenfolge der folgenden „Hütchenfunktionen“:

$$f_n(x) = \begin{cases} n^2 x, \text{ falls } x \in [0, 1/n], \\ -n^2 x + 2n, \text{ falls } x \in (1/n, 2/n], \\ 0, \text{ falls } x \in (2/n, 1], \end{cases}$$

wobei $n \geq 2$ ist. Es ist dann eine leichte Übungsaufgabe zu zeigen, dass die Funktionenfolge $(f_n)_n$ punktweise aber auf $[0, 1]$ nicht gleichmäßig gegen die Nullfunktion konvergiert. Außerdem ist f_n für jedes $n \geq 2$ Riemann-integrierbar und es gilt, wie man sich auch leicht durch Berechnung der Dreiecksfläche klar macht,

$$\int_0^1 f_n(x)\,\mathrm{d}x = 1 \quad \text{für alle } n \geq 2.$$

Damit haben wir die Nichtvertauschbarkeit der Grenzprozesse, denn es ist

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = 1 \neq 0 = \int_0^1 f(x)\,\mathrm{d}x = \int_0^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x.$$

Auch in diesem Zusammenhang zeigt der Begriff der gleichmäßigen Konvergenz seine Berechtigung, denn das Problem im obigen Beispiel war tatsächlich das Fehlen derselben, wie man am folgenden Satz sieht.

Satz 12.27. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $(f_n)_n$ sei eine Funktionenfolge, so dass $f_n \in R([a,b])$ für alle $n \in \mathbb{N}$ gilt und $(f_n)_n$ auf $[a,b]$ gleichmäßig gegen eine Funktion $f : [a,b] \to \mathbb{R}$ konvergiert. Dann ist auch f Riemann-integrierbar und es gilt*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Beweis. Sei $\varepsilon > 0$. Dann gibt es Dank der gleichmäßigen Konvergenz von $(f_n)_n$ ein $n_0 \in \mathbb{N}$, so dass

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{b-a} \quad \text{für alle } n \geq n_0 \text{ und alle } x \in [a,b]$$

gilt. Lösen wir den Betrag auf, so ergibt das für alle $n \geq n_0$

$$-\frac{\varepsilon}{b-a} + f_n(x) \leq f(x) \leq f_n(x) + \frac{\varepsilon}{b-a} \qquad \text{für alle } x \in [a,b]. \tag{12.6}$$

Da jedes f_n Riemann-integrierbar ist, ist insbesondere f_{n_0} beschränkt, also ist wegen der rechten Ungleichung auch f beschränkt. Weiterhin existiert eine Zerlegung Z_ε, sodass $O_{f_{n_0}}(Z_\varepsilon) - U_{f_{n_0}}(Z_\varepsilon) < \varepsilon$ gilt. Gleichzeitig folgen die Ungleichungen

$$O_f(Z_\varepsilon) \leq O_{f_{n_0}}(Z_\varepsilon) + \varepsilon \quad \text{und} \quad U_{f_{n_0}}(Z_\varepsilon) - \varepsilon \leq U_f(Z_\varepsilon)$$

aus (12.6). Zusammen ergibt sich

$$O_f(Z_\varepsilon) - U_f(Z_\varepsilon) \leq O_{f_{n_0}}(Z_\varepsilon) - U_{f_{n_0}}(Z_\varepsilon) + 2\varepsilon < 3\varepsilon,$$

und damit die Riemann-Integrierbarkeit von f aus dem Darboux-Kriterium. Um die Gleichheit der Integrale im Grenzwert zu sehen, folgern wir aus (12.6) mit der Monotonie des Integrals, dass für alle $n \geq n_0$

$$\int_a^b f_n(x)\,\mathrm{d}x - \varepsilon \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b f_n(x)\,\mathrm{d}x + \varepsilon,$$

und damit auch

$$\left| \int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x \right| < 2\varepsilon,$$

erfüllt ist. Die Folge $\left(\int_a^b f_n(x)\,\mathrm{d}x\right)_n$ konvergiert also gegen $\int_a^b f(x)\,\mathrm{d}x$. □

12.4. Wie bestimme ich ein Integral?

Um Integrale zu approximieren, liefert das Riemann-Kriterium schon eine sehr gute Grundlage. Um den Wert eines Integrals exakt zu bestimmen, 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. Stattdessen müssen wir uns mit Rechenregeln und Tricks begnügen, mit denen wir das Problem der Integration einer Funktion auf das entsprechende Problem für eine andere Funktion zurückspielen können, 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 sich aber nicht in einer geschlossenen Form angeben lässt.
Zusammengefasst:

Differenzieren ist Handwerk, Integrieren ist Kunst.

Zwei Techniken, die man immer parat haben sollte, sind *partielle Integration* und *Substitution*, die auf der Produkt- bzw. der Kettenregel fürs Differenzieren beruhen. Sie sind, zusammen mit einer Reihe von Beispielen, das Thema dieses Abschnitts.

Satz 12.28 (Partielle Integration). *Es seien $I := [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall und $f, g : I \to \mathbb{R}$ stetig differenzierbare Funktionen. Dann gilt*

$$\int_a^b f'g = fg \,\Big|_a^b - \int_a^b fg'.$$

Beweis. Zunächst einmal existieren alle in der Behauptung auftretenden Integrale, denn nach Voraussetzung sind $f'g$ und fg' stetige Funktionen und damit auch Riemann-integrierbar auf I.
Nach der Produktregel gilt nun

$$(fg)' = f'g + fg',$$

also ist fg eine Stammfunktion von $f'g + fg'$. Aus dem ersten Hauptsatz der Differenzial- und Integralrechnung, Satz 12.21, folgt also

$$\int_a^b (f'g + fg') = \int_a^b (fg)' = fg \,\Big|_a^b, \quad \text{d. h.} \quad \int_a^b f'g = fg \,\Big|_a^b - \int_a^b fg'.$$

□

Beispiel 12.29. (a) Wir betrachten das Integral

$$\int_0^1 x \sinh(x)\,\mathrm{d}x,$$

d. h. wir wenden unseren Satz mit $g(x) = x$ und $f'(x) = \sinh(x)$ auf dem Intervall $[0,1]$ an. Dann ist $f(x) = \cosh(x)$ eine mögliche Wahl für die Funktion f und wir erhalten mit partieller Integration:

$$\begin{aligned}\int_0^1 x \sinh(x)\,\mathrm{d}x &= x\cosh(x)\Big|_{x=0}^{x=1} - \int_0^1 \cosh(x)\,\mathrm{d}x \\ &= \cosh(1) - 0 - \Big(\sinh(x)\Big|_{x=0}^{x=1}\Big) = \cosh(1) - \sinh(1) \\ &= \frac{\mathrm{e} + {}^1\!/\!_{\mathrm{e}}}{2} - \frac{\mathrm{e} - {}^1\!/\!_{\mathrm{e}}}{2} = \frac{1}{\mathrm{e}}.\end{aligned}$$

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 $g(x) = \sinh(x)$ und $f'(x) = x$ genommen hätten, wären wir bei

$$\int_0^1 x\sinh(x)\,\mathrm{d}x = \frac{1}{2}x^2\sinh(x)\Big|_{x=0}^{x=1} - \int_0^1 \frac{1}{2}x^2\cosh(x)\,\mathrm{d}x$$

gelandet. Diese Umformung ist natürlich auch richtig, aber von dem nun entstandenen Integral wissen wir erst recht nicht, wie wir es berechnen sollen.

(b) Manchmal muss man sich die zweite Funktion zur partiellen Integration erst künstlich schaffen:

$$\begin{aligned}\int_1^2 \ln(x)\,\mathrm{d}x &= \int_1^2 1\cdot\ln(x)\,\mathrm{d}x = x\ln(x)\Big|_{x=1}^{x=2} - \int_1^2 x\frac{1}{x}\,\mathrm{d}x \\ &= \big(x\ln(x) - x\big)\Big|_{x=1}^{x=2} = 2\cdot\ln(2) - 2 - 0 + 1 = 2\cdot\ln(2) - 1,\end{aligned}$$

wobei wir $g(x) = \ln(x)$ und $f'(x) = 1$ gewählt haben.

(c) Oft genug braucht man noch weitere Tricks. Bekannte Eigenschaften der zu untersuchenden Funktionen, wie zum Beispiel der trigonometrische Pythagoras im Fall von Sinus und Cosinus, können helfen. Wir wollen

$$A := \int_0^{\pi/2} \sin^2(x)\,\mathrm{d}x$$

bestimmen. Dazu wählen wir $f'(x) = g(x) = \sin(x)$ und berechnen

$$= -\cos(x)\sin(x)\Big|_{x=0}^{x=\pi/2} - \int_0^{\pi/2} \cos(x)(-\cos(x))\,\mathrm{d}x = \int_0^{\pi/2} \cos^2(x)\,\mathrm{d}x.$$

Wenden wir nun mit $f'(x) = g(x) = \cos(x)$ noch einmal partielle Integration an, so erhalten wir

$$= \sin(x)\cos(x)\Big|_{x=0}^{x=\pi/2} - \int_0^{\pi/2} \sin(x)(-\sin(x))\,\mathrm{d}x = \int_0^{\pi/2} \sin^2(x)\,\mathrm{d}x = A$$

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

$$A = \int_0^{\pi/2} \cos^2(x)\,\mathrm{d}x = \int_0^{\pi/2} (1-\sin^2(x))\,\mathrm{d}x = \frac{\pi}{2} - \int_0^{\pi/2} \sin^2(x)\,\mathrm{d}x = \frac{\pi}{2} - A,$$

woraus $2A = \pi/2$ und schließlich

$$A = \int_0^{\pi/2} \sin^2(x)\,\mathrm{d}x = \frac{\pi}{4}$$

folgt.

Satz 12.30 (Substitution(sregel))**.** *Es seien $[a,b] \subseteq \mathbb{R}$ und $[c,d] \subseteq \mathbb{R}$ kompakte Intervalle, sowie $f \in \mathrm{C}([a,b],\mathbb{R})$ und $g \in \mathrm{C}^1([c,d],\mathbb{R})$ 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 dem Zweiten Hauptsatz 12.24(c) 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 12.31. 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:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = g'(t) \quad \text{„} \Rightarrow \mathrm{d}x = g'(t)\,\mathrm{d}t.\text{“}$$

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 12.32. (a) Wir berechnen das Integral

$$\int_1^2 \frac{x}{\sqrt{2+x^2}}\,\mathrm{d}x$$

mit unserer Schmierrechnungsmethode. Dazu setzen wir $2+x^2=t$. Wegen $x\in[1,2]$ können wir das auflösen zu $x=\sqrt{t-2}$, d. h. wir wenden die Substitutionsregel mit $g(t):=\sqrt{t-2}$ an. Weiter ist bei der Anwendung des Satzes $c=3$ und $d=6$, 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 $\frac{\mathrm{d}x}{\mathrm{d}t}={}^1\!/_{2\sqrt{t-2}}={}^1\!/_{2x}$, also „$2x\mathrm{d}x=\mathrm{d}t$“. Damit haben wir

$$\int_1^2 \frac{x}{\sqrt{2+x^2}}\,\mathrm{d}x=\frac{1}{2}\int_1^2\frac{1}{\sqrt{2+x^2}}\cdot 2x\mathrm{d}x=\frac{1}{2}\int_3^6\frac{1}{\sqrt{t}}\,\mathrm{d}t=\int_3^6\frac{1}{2\sqrt{t}}\,\mathrm{d}t$$
$$=\sqrt{t}\Big|_{t=3}^{t=6}=\sqrt{6}-\sqrt{3}.$$

(b) Als zweites Beispiel wollen wir das Integral

$$A:=\int_0^1\sqrt{1-x^2}\,\mathrm{d}x$$

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 ${}^{\mathrm{d}x}\!/_{\mathrm{d}t}=\cos'(t)=-\sin(t)$ 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 12.29 (c) tatsächlich

$$A=\int_{\pi/2}^0\sin(t)(-\sin(t))\,\mathrm{d}t=\int_0^{\pi/2}\sin^2(t)\,\mathrm{d}t=\frac{\pi}{4}.$$

(c) Schließlich bestimmen wir

$$H(b) = \int_0^b \frac{t+1}{t^2+2t+1}\,\mathrm{d}t$$

für alle $b > 0$. Wir setzen dazu $g(t) = t^2 + 2t + 1$ und $f(x) = \frac{1}{x}$ und stellen fest, dass dann mit der Substitutionsregel gilt:

$$\begin{aligned} H(b) &= \int_0^b \frac{\frac{1}{2}g'(t)}{g(t)}\,\mathrm{d}t = \frac{1}{2}\int_0^b f(g(t))\cdot g'(t)\,\mathrm{d}t \\ &= \frac{1}{2}\int_{g(0)=1}^{g(b)} \frac{1}{x}\,\mathrm{d}x = \frac{1}{2}\ln(x)\Big|_1^{g(b)} = \frac{1}{2}\ln(b^2+2b+1). \end{aligned}$$

12.5. Uneigentliche Integrale

Bisher können wir Integrale nur über kompakte Intervalle und beschränkte Funktionen bilden. Wir wollen unser mächtiges Werkzeug des Grenzübergangs jetzt auch hier verwenden, um allgemeinere Integrale zuzulassen.
In diesem Abschnitt seien stets $a, b \in \mathbb{R}$ und $\alpha \in \mathbb{R} \cup \{-\infty\}$ sowie $\beta \in \mathbb{R} \cup \{\infty\}$.

Definition 12.33. *Es sei $f : [a, \beta) \to \mathbb{R}$ (bzw. $f : (\alpha, b] \to \mathbb{R}$) Riemann-integrierbar 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*

$$\lim_{t\to\beta-}\int_a^t f \qquad \Big(\text{bzw.}\quad \lim_{t\to\alpha+}\int_t^b f\Big)$$

existiert. In diesem Fall heißt der Grenzwert

$$\int_a^\beta f := \lim_{t\to\beta-}\int_a^t f \qquad \Big(\text{bzw.}\quad \int_\alpha^b f := \lim_{t\to\alpha+}\int_t^b f\Big)$$

das uneigentliche Integral *von f über (a, β) (bzw. (α, b)).*

Beispiel 12.34. (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 $\lim_{x\to1-} 1/\sqrt{1-x^2} = \infty$ nicht beschränkt. Für jedes $t \in (0, 1)$ ist sie aber stetig auf dem Intervall $[0, t]$, also dort integrierbar. Wir haben damit im Sinne der obigen Defintion den Fall $a = 0$ und $\beta = 1$. Dann ist für jedes $t \in (0, 1)$

$$\int_0^t \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \arcsin(x)\Big|_{x=0}^{x=t} = \arcsin(t)$$

und wegen $\lim_{t\to 1-}\arcsin(t) = \pi/2$ existiert das uneigentliche Integral

$$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{t\to 1-}\int_0^t \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{t\to 1-}\arcsin(t) = \frac{\pi}{2}.$$

(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

$$\int_0^t \frac{1}{1+x^2}\,\mathrm{d}x = \arctan(x)\Big|_{x=0}^{x=t} = \arctan(t) \longrightarrow \frac{\pi}{2} \quad (t\to\infty),$$

also existiert auch dieses uneigentliche Integral mit dem Wert

$$\int_0^\infty \frac{1}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2}.$$

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 $s = 1$

$$\int_1^t \frac{1}{x}\,\mathrm{d}x = \ln(x)\Big|_{x=1}^{x=t} = \ln(t),$$

also existiert das uneigentliche Integral in diesem Fall wegen $\lim_{t\to\infty}\ln(t) = \infty$ nicht.

Für $s \neq 1$ ist

$$\int_1^t \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{1-s}x^{1-s}\Big|_{x=1}^{x=t} = \frac{1}{1-s}(t^{1-s}-1).$$

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 existiert, wenn $s < 1$ ist. In diesem Fall gilt

$$\int_0^1 \frac{1}{x^s}\,\mathrm{d}x = \frac{1}{1-s}.$$

Bisher haben wir nur uneigentliche Integrale betrachtet, die an einer Grenze uneigentlich sind. Natürlich will man auch den Fall behandeln können, dass es an beiden Intervallgrenzen Probleme gibt. Dazu müssen wir unsere Definition nur ein bisschen modifizieren.

Definition 12.35. *Es sei $f : (\alpha, \beta) \to \mathbb{R}$ Riemann-integrierbar 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 12.33 existieren. In diesem Fall sagen wir, dass das uneigentliche Integral

$$\int_\alpha^\beta f := \int_\alpha^c f + \int_c^\beta f$$

existiert bzw. konvergiert.

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 12.36. Definition 12.35 ist von der Wahl von $c \in (\alpha, \beta)$ unabhängig.

Warnung 12.37. Das uneigentliche Integral $\int_{-\infty}^{\infty} f$ ist *nicht* definiert durch $\lim_{t\to\infty} \int_{-t}^{t} f$, sondern durch

$$\int_{-\infty}^{\infty} f = \lim_{t\to-\infty} \int_t^0 f + \lim_{s\to\infty} \int_0^s f.$$

Das ist ein wesentlicher Unterschied, wie man an dem Beispiel $\int_{-\infty}^{\infty} x\,\mathrm{d}x$ sieht. Dieser Wert existiert offensichtlich nicht, denn sowohl $\int_{-\infty}^{0} x\,\mathrm{d}x$, als auch $\int_0^\infty x\,\mathrm{d}x$ sind divergent. Gleichzeitig ist aber für jedes $t > 0$

$$\int_{-t}^{t} x\,\mathrm{d}x = 0,$$

sodass der ganz oben angegebene „gleichzeitige“ Limes existiert und den Wert Null hat. Trotzdem ergibt 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

$$\lim_{t\to\infty} \int_{-t}^{t+1} x \,\mathrm{d}x = \lim_{t\to\infty} \frac{1}{2}(t+1)^2 - \frac{1}{2}t^2 = \lim_{t\to\infty}\Big(t + \frac{1}{2}\Big) = \infty,$$

sieht das Ergebnis schon anders aus.
Also merke: Ein doppelt uneigentliches Integral konvergiert nur dann, wenn es an beiden Integrationsgrenzen unabhängig voneinander konvergiert.

Beispiel 12.38. (a) Es existiert mit Hilfe von Beispiel 12.34 (b) das doppelt uneigentliche Integral

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} \,\mathrm{d}x = \pi.$$

(b) Sei $s > 0$. Kombiniert man (c) und (d) aus Beispiel 12.34, so sieht man, dass das doppelt uneigentliche Integral

$$\int_0^\infty \frac{1}{x^s} \,\mathrm{d}x$$

genau dann existiert, wenn $s > 1$ und $s < 1$ gilt, also nie.

Die folgenden Sätze und Definitionen formulieren wir der Übersichtlichkeit halber nur für uneigentliche Integrale der Form $\int_a^\beta f(x) \,\mathrm{d}x$. Entsprechende Sätze und Defintionen gelten auch für die anderen Arten uneigentlicher Integrale.
Im nächsten Satz geht es um einen ersten präzisen Zusammenhang zwischen der Konvergenz von Reihen und von Integralen.

Satz 12.39 (Vergleichssatz Integral-Reihe)**.** *Ist $f\colon [1,\infty) \to \mathbb{R}_+$ monoton fallend und Riemann-integrierbar auf $[1,\beta]$ für alle $\beta > 1$, dann gilt:*

$$\sum_{n=1}^{\infty} f(n) \text{ existiert} \quad \Leftrightarrow \quad \int_1^\infty f(x) \,\mathrm{d}x \text{ existiert.}$$

Beweis. Zuerst zur Richtung „$\Rightarrow$“: Für $k \geq 2$ und $x \in [k-1, k]$ ist nach Voraussetzung $f(k) \leq f(x) \leq f(k-1)$, also auch $f(k) \leq \int_{k-1}^k f(x) \,\mathrm{d}x \leq f(k-1)$ und damit

$$\sum_{k=2}^{n} f(k) \leq \int_1^n f(x) \,\mathrm{d}x \leq \sum_{k=1}^{n-1} f(k)$$

für $n \geq 2$. Wegen der Positivität von f ist die Folge $\left(\int_1^n f(x)\,\mathrm{d}x\right)_n$ monoton und wegen der Folgerung

$$\int_1^n f(x)\,\mathrm{d}x \leq \sum_{k=1}^{n-1} f(k) \leq \sum_{k=1}^{\infty} f(k)$$

aus der zweiten Abschätzung oben auch beschränkt und damit konvergent (warum ist das ausreichend?).
Jetzt zur umgekehrten Richtung „$\Leftarrow$“: Wie oben und aufgrund der Monotonie und Positivität von f folgt

$$\sum_{k=2}^{n} f(k) \leq \int_1^n f(x)\,\mathrm{d}x \leq \int_1^\infty f(x)\,\mathrm{d}x,$$

sodass die Folge $\left(\sum_{k=1}^n f(k)\right)_n$ unter der Voraussetzung der Existenz des uneigentlichen Integrals monoton wachsend und beschränkt ist, also konvergent. □

Beispiel 12.40. Zu gegebenem $\alpha \in \mathbb{R}_+$ betrachten wir die Funktion $f\colon (1,\infty) \to \mathbb{R}_+$ mit $f(x) = \frac{1}{x(\ln x)^\alpha}$. Es ist $F(x) = \frac{(\ln x)^{1-\alpha}}{1-\alpha}$ eine Stammfunktion von f für $\alpha \in \mathbb{R}_+ \setminus \{1\}$ und $F(x) = \ln(\ln x)$ eine Stammfunktion von $x \mapsto \frac{1}{x \ln x}$ (beides nachrechnen!). Damit lässt sich das uneigentliche Integral $\int_2^\infty f(x)\,\mathrm{d}x$ bestimmen und wir sehen, dass es genau dann existiert, wenn $\alpha > 1$ ist (auch nachprüfen!). Weiter ist f wegen

$$f'(x) = -\frac{1 + \frac{\alpha}{\ln x}}{x^2(\ln x)^\alpha}$$

auf $(1,\infty)$ monoton fallend, sodass die Voraussetzungen von Satz 12.39 erfüllt sind. Es folgt, dass der Reihenwert

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^\alpha}$$

genau dann existiert, wenn $\alpha > 1$ ist.

Bemerkung 12.41. Unter den Voraussetzungen von Satz 12.39 betrachten wir die Folge $(a_n)_n$ gegeben durch

$$a_n := \sum_{k=1}^{n} f(k) - \int_1^{n+1} f(x)\,\mathrm{d}x.$$

Der Beweis des Satzes zeigt, dass die Folge monoton wachsend und aufgrund der Abschätzung $0 \leq a_n \leq f(1) - f(n+1) \leq f(1)$ für alle $n \in \mathbb{N}$ beschränkt ist.

Die Folge ist also konvergiert und der Grenzwert erfüllt $0 \leq \lim_{n\to\infty} a_n \leq f(1)$. Speziell für die Funktion $f(x) = \frac{1}{x}$ ergibt sich

$$\lim_{n\to\infty}(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n) =: \gamma \leq 1,$$

was bedeutet, dass sich die harmonische Reihe in etwa wie die Folge $n \mapsto \ln n$ verhält. Der Grenzwert γ heißt *Euler-Konstante.*

Grob gesagt funktioniert die Theorie zur Existenz von uneigentlichen Integralen irgendwie so ähnlich wie die Theorie zur Konvergenz von Reihen. Bei Reihen war dabei ja insbesondere wichtig, ob die Konvergenz unabhängig von den Vorzeichen der Summanden auftritt, oder durch einen Wechsel des Vorzeichens bedingt wird – mit anderen Worten: ob es sich um *absolute* Konvergenz handelt (Stichwort: Umordnen). Auch bei den uneigentlichen Integralen füren wir deshalb die entsprechende Unterscheidung ein.

Definition 12.42. *Das uneigentliche Integral $\int_a^\beta f$ heißt* absolut konvergent*, wenn das uneigentliche Integral $\int_a^\beta |f|$ existiert.*

Parallel zu unseren Überlegungen bei absolut konvergenten Reihen gilt hier:

Satz 12.43 (Dreiecksungleichung für uneigentliche Integrale)**.** *Ist das uneigentliche Integral $\int_a^\beta f$ absolut konvergent, so ist es auch konvergent und es gilt*

$$\Big|\int_a^\beta f\Big| \leq \int_a^\beta |f|.$$

Beweis. Sei $\varepsilon > 0$ und sei $(t_n)_n \subseteq (a, \beta)$ eine Folge mit $\lim_{n\to\infty} t_n = \beta$. Da das Integral absolut konvergiert, ist $\left(\int_a^{t_n} |f(x)|\,\mathrm{d}x\right)_n$ eine Cauchyfolge, es ex. also ein $n_0 \in \mathbb{N}$, sodass

$$\Big|\int_a^{t_n} |f(x)|\,\mathrm{d}x - \int_a^{t_m} |f(x)|\,\mathrm{d}x\Big| = \Big|\int_{t_n}^{t_m} |f(x)|\,\mathrm{d}x\Big| < \varepsilon$$

für alle $n, m \geq n_0$ erfüllt ist. Daraus folgt mit der Dreiecksungleichung für (gewöhnliche, nicht uneigentliche) Integrale aber direkt

$$\Big|\int_a^{t_n} f(x)\,\mathrm{d}x - \int_a^{t_m} f(x)\,\mathrm{d}x\Big| = \Big|\int_{t_n}^{t_m} f(x)\,\mathrm{d}x\Big| \leq \int_{t_n}^{t_m} |f(x)|\,\mathrm{d}x < \varepsilon,$$

also, dass auch $\left(\int_a^{t_n} f(x)\,\mathrm{d}x\right)_n$ eine Cauchyfolge, und deshalb auch konvergent ist. Per Definition existiert damit das uneigentliche Integral. Die Dreiecksungleichung ergibt sich jetzt aus der Stetigkeit der Betragsfunktion und wieder aus der Dreiecksungleichung für nicht uneigentliche Integrale:

$$\Big|\int_a^\beta f\Big| = \Big|\lim_{t\to\beta_-}\int_a^t f\Big| = \lim_{t\to\beta_-}\Big|\int_a^t f\Big| \leq \lim_{t\to\beta_-}\int_a^t |f| = \int_a^\beta |f|.$$

□

Für die Untersuchung von Reihen auf absolute Konvergenz waren Majoranten- und Minorantenkriterium die vielleicht wichtigsten Hilfsmittel. Sie lassen sich auch auf die absolute Konvergenz uneigentlicher Integrale übertragen. Der Beweis verbleibt dabei als Übungsaufgabe:

Satz 12.44 (Majoranten-/Minorantenkriterium). *(a) Ist $|f(x)| \leq g(x)$ für alle $x \in [a, \beta)$ und ist $\int_a^\beta g$ konvergent, dann ist $\int_a^\beta f$ absolut konvergent.*

(b) Ist $f(x) \geq g(x) \geq 0$ für alle $x \in [a, \beta)$ und ist g nicht uneigentlich integrierbar, dann ist auch f nicht uneigentlich integrierbar.

Beispiel 12.45. (a) Wir können jetzt relativ schnell beweisen, dass

$$\int_0^\infty \frac{\sin(x)}{x} \, \mathrm{d}x$$

ein konvergentes, aber nicht absolut konvergentes uneigentliches Integral ist. Dazu stellen wir zunächst fest, dass dieses, obwohl es auf den ersten Blick doppelt uneigentlich aussieht, eigentlich nur einfach uneigentlich für x gegen unendlich ist, denn nach Beispiel 10.17 (a) können wir die Funktion $x \mapsto {}^{\sin(x)}/_x$ durch den Wert 1 stetig nach $x = 0$ fortsetzen. Mit Hilfe von partieller Integration erhalten wir für jedes $t > 1$:

$$\int_1^t \frac{\sin(x)}{x} \, \mathrm{d}x = \cos 1 - \frac{\cos t}{t} - \int_1^t \frac{\cos x}{x^2} \, \mathrm{d}x.$$

Das uneigentliche Integral $\int_1^t \frac{\cos x}{x^2} \, \mathrm{d}x$ existiert, weil es die konvergente Majorante $\int_1^t \frac{1}{x^2} \, \mathrm{d}x$ besitzt. Also existiert auch der Grenzwert

$$\lim_{t \to \infty} \int_1^t \frac{\sin x}{x} \, \mathrm{d}x.$$

Andererseits haben wir für alle $k \in \mathbb{N}$ die Abschätzung

$$\int_{(k-1)\pi}^{k\pi} \frac{|\sin(x)|}{x} \, \mathrm{d}x \geq \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin(x)| \, \mathrm{d}x = \frac{2}{k\pi},$$

die sich zu der Abschätzung

$$\int_0^{n\pi} \frac{|\sin(x)|}{x} \, \mathrm{d}x \geq \frac{2}{\pi} \sum_{k=1}^n \frac{1}{k},$$

aufsummieren läßt. Aus der Divergenz der harmonischen Reihe können wir also schließen, dass das Integral nicht absolut konvergiert.

(b) 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 12.34 (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

$$\int_1^\infty f(x)\,\mathrm{d}x \leq \int_1^\infty g(x)\,\mathrm{d}x = \int_1^\infty \frac{1}{x^{3/2}}\,\mathrm{d}x = \frac{1}{3/2-1} = 2.$$

(c) 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 12.34 (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 12.46. 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 einem typischen Fehler gewarnt.

Warnung 12.47. Wir haben in Beispiel 12.34 (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.

12.6. Vermischtes zum Riemann-Integral

Unbestimmte Integrale

Im Verlauf der letzten Abschnitte, insbesondere bei der Berechnung von Integralen, hat sich aus unserer ursprünglichen Motivation, nämlich der der Flächenberechnung, zunehmend auch die abstrakte Fragestellung nach der Bestimmung einer Stammfunktion ergeben. Damit hängt der Begriff des *unbestimmten Integrals* zusammen, den wir hier noch einführen, weil er in der Literatur sehr gebräuchlich ist. Dazu gehört eine dicke Warnung: Die zugehörige Notation ist zwar manchmal praktisch, aber auch sehr ungenau und irreführend!

Definition 12.48. *Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion. Besitzt f auf I eine Stammfunktion F, so schreibt man auch*

$$\begin{aligned}\int f = \int f(x)\,\mathrm{d}x &:= \{G\colon I \to \mathbb{R} : G' = f\}\\ &= \{G\colon I \to \mathbb{R} : \exists c \in \mathbb{R}\ \forall x \in I\ \ G(x) = F(x) + c\},\end{aligned}$$

für das sogenannte unbestimmte Integral*, also für die* Menge *aller Stammfunktionen von f.*

Diskussion und Beispiele:

(a) Oft sieht man eine Schreibweise der Form $\displaystyle\int \mathrm{e}^x\,\mathrm{d}x = \mathrm{e}^x + c, \quad c \in \mathbb{R}.$

Diese Zeile ist formal unsauber, denn es fehlen rechts vom Gleichheitszeichen die Mengenklammern. Außerdem steht links eine Menge von Funktionen, rechts ist $e^x + c$ zu gegebenen $x, c \in \mathbb{R}$ eine Zahl. Wenn man diese (manchmal praktische) Notation verwendet, muss einem klar sein, was hier gerade passiert und wie sie richtig zu interpretieren sind. Nur wenn man genau weiß, dass man gerade schlampt, darf man in der Mathematik schlampen!
Diese Schreibweise findet sich oft in den „Integraltafeln“, also in Nachschlagewerken für Stammfunktionen von verschiedenen Funktionentypen.

(b) Wir erinnern uns, dass $F\colon I \to \mathbb{R}$ eine Stammfunktion von $f\colon I \to \mathbb{R}$ sein kann, auch wenn f nicht integrierbar ist, vgl. den Flattersinus in Warnung 12.23 (a). Das ist vielleicht das größte Problem dieser suggestiven Schreibweise: das unbestimmte Integral $\int f$ ist manchmal definiert, obwohl f nicht integrierbar ist!
*Für Interessierte: es kann sogar noch deutlich schlimmer kommen – die sogenannte *Volterra-Funktion* ist zwar differenzierbar auf $[0,1]$ und insofern eine Stammfunktion, ihre Ableitung ist aber nicht Riemann-integrierbar, obwohl sie beschränkt ist.

(c) Auch wenn die Notation also schrecklich ist, zum Schluß noch ein Beispiel eines unbestimmten Integrals:

$$\int \sqrt{1-x^2}\,\mathrm{d}x.$$

Für die Berechnung des Werts $\frac{\pi}{4}$ für das bestimmte Integral auf $[0,1]$ hatten wir hier geschickt getrickst (vgl. Beispiel 12.32 (b)). Diese Technik wollen wir jetzt zur Bestimmung der Menge aller Stammfunktion leicht anpassen. Dabei ist zu beachten, dass wir ja nur eine Stammfunktion brauchen und alle anderen dann durch Addieren einer Konstanten (Funktion) erhalten. Da hier der Integrand stetig ist, können wir den zweiten Hauptsatz 12.24 anwenden und erhalten, dass die Funktion

$$F(r) := \int_0^r \sqrt{1-x^2}\,\mathrm{d}x,$$

definiert für $r \in [0,1]$, eine Stammfunktion sein muss. Mit der Substitution $x = \cos(t)$ wie in 12.32, partieller Integration wie in Beispiel 12.29 (c) und dem Satz von Pythagoras ergibt sich

$$\begin{aligned} F(r) = \int_{\arccos(r)}^{\pi/2} \sin^2(t)\,\mathrm{d}t &= -\sin\cos\Big|_{\arccos(r)}^{\pi/2} + \int_{\arccos(r)}^{\pi/2} \cos^2(t)\,\mathrm{d}t, \\ &= -\sin\cos\Big|_{\arccos(r)}^{\pi/2} + \pi/2 - \arccos(r) - \int_{\arccos(r)}^{\pi/2} \sin^2(t)\,\mathrm{d}t, \\ &= -\sin\cos\Big|_{\arccos(r)}^{\pi/2} + \pi/2 - \arccos(r) - F(r), \end{aligned}$$

woraus

$$F(r) = \frac{2r\sin(\arccos(r)) - 2\arccos(r) + \pi}{4}$$

folgt. Wieder aus dem Satz von Pythagoras ergibt sich

$$\sin^2(\arccos(r)) = 1 - \cos^2(\arccos(r)) = 1 - r^2,$$

sodass wir $\sin(\arccos(r))$ zu $\sqrt{1-r^2}$ umschreiben können. Das unbestimmte Integral in der „doppelt geschlampten“ Schreibweise ist also gegeben durch

$$\int \sqrt{1-x^2}\,\mathrm{d}x = F + c = \frac{1}{2}\big(x\sqrt{1-x^2} - \arccos(x)\big) + c.$$

Satz von Taylor mit Integralrest

Wir wollen in diesem Abschnitt 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 die Idee bei der Anwendung dieses Satzes meist die Abschätzung des Restgliedes ist, ist es einfach praktisch, möglichst viele verschiedene Darstellungen dafür zu haben.

Satz 12.49 (Satz von Taylor (mit Integralrestglied))**.** *Es seien $I \subseteq \mathbb{R}$ ein Intervall, $f \in \mathrm{C}^{n+1}(I, \mathbb{R})$ für ein $n \in \mathbb{N}_0$ und $x, x_0 \in I$. Dann gilt*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t) \, \mathrm{d}t.$$

Beweis. Wir führen einen Induktionsbeweis. Für $n = 0$ ist die Behauptung

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t) \, \mathrm{d}t,$$

was gerade der Hauptsatz der Differenzial- und Integralrechnung ist. Gilt die Formel nun für ein $n \in \mathbb{N}_0$, so haben wir für $f \in \mathrm{C}^{n+2}(I, \mathbb{R})$ nach Induktionsvoraussetzung

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t) \, \mathrm{d}t.$$

Mit partieller Integration folgt daraus

$$\begin{aligned}
&= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \left(\frac{1}{n+1} \frac{-(x-t)^{n+1}}{n!} f^{(n+1)}(t) \right)\Big|_{t=x_0}^{t=x} \\
&\qquad + \int_{x_0}^{x} \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t) \, \mathrm{d}t \\
&= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1} + \int_{x_0}^{x} \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t) \, \mathrm{d}t \\
&= \sum_{k=0}^{n+1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \int_{x_0}^{x} \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t) \, \mathrm{d}t.
\end{aligned}$$

□

*Die Γ-Funktion

Wir starten mit dem folgenden Resultat über ein spezielles uneigentliches Integral:

Satz 12.50. *Es sei $x > 0$. Dann existiert das doppelt uneigentliche Integral*

$$\Gamma(x) := \int_0^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$$

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 12.46 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 $\int_0^c {}^1\!/\!_{t^{1-x}} \, \mathrm{d}t$ existiert, existiert nach dem Majorantenkriterium auch $\int_0^c \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$, und damit auch $\int_0^1 \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t$.
Für das Integral von Eins bis ∞ vergleichen wir mit ${}^1\!/\!_{t^2}$ 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. □

Das soeben behandelte Integral ist wichtig genug, dass es einen Namen verdient hat.

Definition 12.51. *Die nach Satz 12.50 durch* $\Gamma : (0, \infty) \to \mathbb{R}$ *mit*

$$\Gamma(x) := \int_0^\infty \mathrm{e}^{-t} t^{x-1} \, \mathrm{d}t, \quad x > 0,$$

gegebene Funktion heißt Gamma-Funktion.

Die Gamma-Funktion wurde von Leonard Euler eingeführt, um die Fakultätsfunktion $n \mapsto n!$ sinnvoll auf die reellen Zahlen zu erweitern. Es gilt jedenfalls:

Satz 12.52. *Für alle $x > 0$ gilt $\Gamma(x + 1) = x\Gamma(x)$.*

Beweis. Es seien $0 < \alpha < \beta$. Dann gilt mit partieller Integration

$$\int_\alpha^\beta \mathrm{e}^{-t} t^x \,\mathrm{d}t = -\mathrm{e}^{-t} t^x \Big|_\alpha^\beta - \int_\alpha^\beta (-e^{-t}) x t^{x-1} \,\mathrm{d}t = \frac{1}{\mathrm{e}^\alpha}\alpha^x - \frac{1}{\mathrm{e}^\beta}\beta^x + x \int_\alpha^\beta \mathrm{e}^{-t} t^{x-1} \,\mathrm{d}t.$$

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

$$\int_1^\beta \mathrm{e}^{-t} t^x \,\mathrm{d}t = \frac{1}{\mathrm{e}} - \frac{1}{\mathrm{e}^\beta}\beta^x + x \int_1^\beta \mathrm{e}^{-t} t^{x-1} \,\mathrm{d}t,$$

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

$$\begin{aligned}\Gamma(x+1) &= \int_0^\infty \mathrm{e}^{-t} t^x \,\mathrm{d}t = \int_0^1 \mathrm{e}^{-t} t^x \,\mathrm{d}t + \int_1^\infty \mathrm{e}^{-t} t^x \,\mathrm{d}t \\ &= -\frac{1}{\mathrm{e}} + x \int_0^1 \mathrm{e}^{-t} t^{x-1} \,\mathrm{d}t + \frac{1}{\mathrm{e}} + x \int_1^\infty \mathrm{e}^{-t} t^{x-1} \,\mathrm{d}t = x \int_0^\infty \mathrm{e}^{-t} t^{x-1} \,\mathrm{d}t \\ &= x\Gamma(x).\end{aligned}$$ □

... und für die natürlichen Zahlen ergibt sich entsprechend:

Korollar 12.53. *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 12.52

$$\Gamma(n+2) = \Gamma((n+1)+1) = (n+1)\Gamma(n+1) = (n+1)n! = (n+1)!.$$ □

*Das Lebesgue-Kriterium für Riemann-Integrierbarkeit

Mit Satz 12.16 hatten wir gezeigt, dass eine Funktion Riemann-integrierbar ist, wenn sie an nur endlich vielen Stellen unstetig ist. In Beispiel 12.18 hatten wir gesehen, dass es auch Riemann-integrierbare Funktionen gibt, die an mehr als endlich vielen Stellen unstetig (und nicht monoton) sind. In diesem Abschnitt lernen wir einen Satz kennen, der ein gleichzeitig hinreichendes und notwendiges Kriterium an die Menge der Stetigkeitsstellen einer Funktion liefert, unter dem sie Riemann-integrierbar ist. Das ist schon aus theoretischer Sicht eine sehr interessante Ergänzung zum Riemannschen und zum Darbouxschen Kriterium. Für den Beweis des Satzes fehlen uns leider noch die Techniken, deshalb ist dieser Abschnitt nur als kleiner Ausblick auf die Analysis 2 gedacht, in der wir uns noch einmal intensiv mit alternativen Blickwinkeln auf den Integralbegriff beschäftigen.
Zur Vorbereitung des Kriteriums brauchen wir noch die folgende Definition.

Definition 12.54. *Eine Menge $N \subseteq \mathbb{R}$ heißt* Nullmenge*, wenn es zu jedem $\varepsilon > 0$ eine Folge von Intervallen $(I_n)_n$ gibt, die N überdecken und in der Summe höchstens Länge ε haben, also*

$$N \subseteq \bigcup_{n\in\mathbb{N}} I_n, \quad \textit{und} \quad \sum_{n=1}^{\infty} |I_n| < \varepsilon$$

erfüllen.

Beispiel 12.55. (a) Das Bild $\{a_n \in \mathbb{R} : n \in \mathbb{N}\}$ jeder reellen Folge $(a_n)_n$ ist eine Nullmenge. In anderen Worten: endliche und abzählbar unendliche Teilmengen von $\mathbb{R}$ sind Nullmengen.

(b) In Übungsaufgabe 11.3 wurde gezeigt, dass die *Cantor-Menge* eine Nullmenge ist. Man kann auch zeigen, dass sie überabzählbar, also insbesondere nicht von der Form der Mengen in (a) ist.

(c) Die (abzählbare) Vereinigung über jede Folge von Nullmengen ist wieder eine Nullmenge.

Jetzt können wir den entscheidenden Satz formulieren.

Satz 12.56 (Lebesgue-Kriterium für Riemann-Integrierbarkeit)**.** *Es gilt $f \in R([a,b])$ genau dann, wenn f beschränkt ist und die Menge der Unstetigkeitsstellen von f auf $[a,b]$ eine Nullmenge ist.*

Für die Theorie des Riemann-Integrals ist dieser Satz sehr wertvoll – er bietet eine ganz neue Perspektive. Wir halten hier nur noch die folgende Verbesserung von Satz 12.19 über die Integrierbarkeit der Verkettung von integrierbaren Funktionen fest, weil sie zum einen eine direkte Folgerung des Lebesgue-Kriteriums ist (der Beweis ist eine Übungsaufgabe), zum anderen sehr nützlich ist.

Satz 12.57. *Es seien $a, b \in \mathbb{R}$ mit $a < b$ und eine Funktion $f \in R([a, b])$ gegeben. Ist $h : f([a, b]) \to \mathbb{R}$ eine stetige Funktion, dann gilt auch $h \circ f \in R([a, b])$.*

— ⋆⋆⋆ —

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

Abbildung, 11
abgeschlossene Menge, 120
abgeschlossenes Intervall, 20
Ableitung, 137
  logarithmische, 143
Ableitungsfunktion, 138
absolute Konvergenz, 74
  für uneigentliche Integrale, 208
  in $\mathbb{C}$, 99
Addition, 17
  in $\mathbb{C}$, 95
Additionstheoreme, 154
allgemeine Potenz, 126
Allquantor, 7
alternierende harmonische Reihe, 72
alternierende Reihen, 72
Antisymmetrie, 19
Äquivalenz von Aussagen, 8
Archimedes, Satz von, 26
Arcuscosinus, 159
Arcussinus, 159
Arcustangens, 159
Areacosinus hyperbolicus, 161
Areasinus hyperbolicus, 161
Areatangens hyperbolicus, 161
Argument einer Funktion, 11
Argument einer komplexen Zahl, 163
Assoziativgesetz, 17
Aussage, 3
Aussageform, 3
Axiom, 17

bedingt konvergente Reihe, 74
Bernoullische Ungleichung, 29
beschränkte
  Funktion, 120
beschränkte
  Folge, 44
    in $\mathbb{C}$, 99
  Menge, 21
bestimmt divergente Folge, 50
Betrag, 23
  in $\mathbb{C}$, 97
Beweis
  durch Kontraposition, 9
  durch vollständige Induktion, 26
bijektiv, 12
Bild, 11
Bildmenge, 11
Binomialformel, 29
Binomialkoeffizienten, 29
Bolzano, Nullstellensatz von, 120
Bolzano-Weierstraß, Satz von, 58
  in $\mathbb{C}$, 100

$\mathrm{C}(D, \mathbb{K})$, 107
Cauchy-Folge, 63
  in $\mathbb{C}$, 99
Cauchy-Kriterium
  für Folgen, 63
    in $\mathbb{C}$, 100
  für Reihen, 70
    in $\mathbb{C}$, 101
Cauchy-Produkt, 86
  in $\mathbb{C}$, 101
$\mathrm{C}^n(I, \mathbb{R})$, 167
Cosinus, 93
  hyperbolicus, 160

in $\mathbb{C}$, 104
Cotangens, 158
$C^\infty(I,\mathbb{R})$, 167

de l' Hospital, Satz von, 147
De Moivre, Formel von, 103
De Morgan, Regeln von, 5
Definitionsmenge, 11
dichte Teilmenge, 35
Differenzierbarkeit, 137
der Umkehrfunktion, 142
n-malige, 165
stetige, 167
zweimalige, 165
Dirichletsche Sprungfunktion, 179
Distributivgesetz, 18
divergente
Folge, 41
Minorante, 79
Divergenz, bestimmte, 50
Division, 19
Dreiecksungleichung, 24
für uneigentliche Integrale, 208
für Integrale, 185
in $\mathbb{C}$, 98
umgekehrte, 24
verallgemeinerte, 74
in $\mathbb{C}$, 101
Durchschnitt von Mengen, 4

e, 55
Eins, 17
Einschränkung einer Funktion, 13
Element, 3
Entwicklungspunkt, 89
ε-Umgebung, 42
in $\mathbb{C}$, 99
Euler-Formel, 103
Eulersche Zahl, 55
Existenzquantor, 7
Exponentialfunktion, 84, 88
Funktionalgleichung, 88
in $\mathbb{C}$, 103
Exponentialreihe, 84
Extremum
globales, 143
lokales, 143

Fakultät, 29
Folge, 41
beschränkte, 44
in $\mathbb{C}$, 99
bestimmt divergente, 50
Cauchy-, 63
in $\mathbb{C}$, 99
divergente, 41
geometrische, 49
konvergente, 41
in $\mathbb{C}$, 99
monoton fallende, 51
monoton wachsende, 51
monotone, 51
Null-, 42
rekursiv definierte, 51
streng monoton fallende, 51
streng monoton wachsende, 51
streng monotone, 51
Teil-, 55
umgeordnete, 73
Folgenglied, 41
Folgenstetigkeit, 107
Fortsetzung
stetige, 115
Funktion, 11
beliebig oft differenzierbare, 167
beschränkte, 120
bijektive, 12
differenzierbare, 137
gerade, 155
gleichmäßig stetige, 135
injektive, 12
Lipschitz-stetige, 136
monoton fallende, 122
monoton wachsende, 122
monotone, 122
n mal differenzierbare, 165

periodische, 157
Stamm-, 192
stetig differenzierbare, 167
stetige, 107
streng monoton fallende, 122
streng monoton wachsende, 122
streng monotone, 122
surjektive, 12
Umkehr-, 13
uneigentlich integrierbare, 203, 205
ungerade, 155
unstetige, 107
zweimal differenzierbare, 165
Funktionalgleichung der Exponentialfunktion, 88
Funktionenfolge, 127
gleichmäßig konvergente, 128
lokal gleichmäßig konvergente, 128
punktweise konvergente, 127
Funktionenreihe
lokal gleichmäßig konvergente, 128

Gamma-Funktion, 214
ganze Zahlen, 25
Gaußsche Zahlenebene, 97
geometrische
Folge, 49
Reihe, 68
in $\mathbb{C}$, 102
Summenformel, 32
gerade Funktion, 155
Gleichheitszeichen, 3
gleichmäßige Konvergenz, 128
gleichmäßige Konvergenz
lokal, 128
gleichmäßige Stetigkeit, 135
globales Maximum/Minimum, 143
globlaes Extremum, 143
Grenzfunktion, 127
Grenzwert, 41
einer Funktion, 112
linksseitig, 112
rechtsseitig, 112
Vertauschung, 53
Grenzwertsätze
für Funktionen, 114
Grenzwertsätze, 45
für Reihen, 70

Hadamard, Satz von, 90, 102
Häufungspunkt einer Menge, 111
Häufungswert einer Folge, 56
halboffenes Intervall, 20
harmonische Reihe, 69
Hauptsatz d. Diff.- u. Integr.-Rechn., 192
Hauptsatz d. Diff.- u. Integr.-Rechn.
zweiter, 194
Hospital, Satz von de l', 147

i, 95
Identitätssatz für Potenzreihen, 116
imaginäre Einheit, 95
Imaginärteil, 95
Implikation, 8
Indexmenge, 6
Indexshift, 31
Induktion, vollständige, 26
Induktionsmenge, 25
Infimum einer Menge, 21
injektiv, 12
Inklusion, 4
Integral, 178
unbestimmtes, 211
uneigentliches, 203, 205
Integration, partielle, 199
Integrierbarkeit, uneigentliche, 203, 205
Intervall, 20
abgeschlossenes, 20
halboffenes, 20
offenes, 20
Intervallschachtelung, Prinzip der, 64

$\mathbb{K}$, 99
kartesisches Produkt, 4
Kettenregel, 141
Kommutativgesetz, 17

kompakte Menge, 120
Komplement einer Menge, 4
komplexe Zahlen, 95
Komposition, 12
Konjugation, 97
konjugiert komplexe Zahl, 97
Kontraposition, 9
konvergente
  Folge, 41
    in $\mathbb{C}$, 99
  Majorante, 79
Konvergenz, 41
  absolute, 74
    in $\mathbb{C}$, 99
  bedingte, 74
  gleichmäßige, 128
  lokal gleichmäßige, 128
  punktweise, 127
  unbedingte, 74
Konvergenzradius, 91, 102

leere Menge, 4
Leibniz-Kriterium, 73
Limes, 41
  inferior, 59
  superior, 59
linksseitiger Grenzwert, 112
Lipschitz-Stetigkeit, 136
logarithmische Ableitung, 143
Logarithmus
  in $\mathbb{C}$, 164
  natürlicher, 124
  Reihenentwicklung, 152
lokal gleichmäßige Konvergenz, 128
lokales Extremum, 143
lokales Maximum/Minimum, 143

Majorantenkriterium, 79
  für uneigentliche Integrale, 209
  in $\mathbb{C}$, 101
Maximum
  einer Funktion, 143
  globales, 143
  lokales, 143
Maximum einer Menge, 21
Menge, 3
  abgeschlossene, 120
  beschränkte, 21
  Bild-, 11
  Definitions-, 11
  dichte, 35
  Element einer, 3
  Index-, 6
  kompakte, 120
  Komplement einer, 4
  nach oben beschränkte, 21
  nach unten beschränkte, 21
  Ober-, 4
  Teil-, 4
  Urbild-, 11
  wohlgeordnete, 27
  Ziel-, 11
Mengendifferenz, 4
Minimum
  einer Funktion, 143
  globales, 143
  lokales, 143
Minimum einer Menge, 22
Minorantenkriterium, 79
  für uneigentliche Integrale, 209
Mittelwertsatz, 144
  für Integrale, 196
  verallgemeinerter, 146
monoton fallende
  Funktion, 122
monoton fallende Folge, 51
monoton wachsende
  Funktion, 122
monoton wachsende Folge, 51
monotone
  Funktion, 122
monotone Folge, 51
Monotonie-Kriterium, 51
  für Reihen, 70
Multiplikation, 17
  in $\mathbb{C}$, 95

$\mathbb{N}$, 25
n-te Ableitung, 165
natürlicher Logarithmus, 124
natürliche Zahlen, 25
negative reelle Zahlen, 19
  strikt, 19
Null, 17
Nullfolge, 42
Nullstellensatz von Bolzano, 120

obere Schranke, 21
oberer Limes, 59
Obermenge, 4
Obersumme, 180
offenes Intervall, 20

Partialsumme, 67
partielle Integration, 199
periodische Funktion, 157
π, 156
Poduktzeichen, 28
Polarkoordinaten, 162
positive reelle Zahlen, 19
  strikt, 19
Potenz
  allgemeine, 126
  ganzzahliger Exponent, 29
  rationaler Exponent, 38
Potenzfunktion, 126
Potenzreihe, 89
  Entwicklungspunkt einer, 89
  Konvergenzradius einer, 91
Prinzip der Intervallschachtelung, 64
Produkt, kartesisches, 4
Produktregel, 140
Produktreihe, 85
punktweise Konvergenz, 127

$\mathbb{Q}$, 33
Quantor, 7
Quantoren-Umklappprinzip, 8
Quotientenkriterium, 83
  in $\mathbb{C}$, 101
Quotientenregel, 141

rationale Zahlen, 33
Realteil, 95
rechtsseitiger Grenzwert, 112
reelle Zahlen, 17
  negative, 19
  positive, 19
  strikt negative, 19
  strikt positive, 19
Reihe, 67
  absolut konvergente, 74
    in $\mathbb{C}$, 99
  alternierende, 72
  alternierende harmonische, 72
  bedingt konvergente, 74
  Exponential-, 84
  geometrische, 68
    in $\mathbb{C}$, 102
  harmonische, 69
  Potenz-, 89
  umgeordnete, 73
  unbedingt konvergente, 74
rekursiv definierte Folge, 51
Restglied, 168
Riemann-Integral, 178
Riemannsche Summe, 177
Riemannscher Umordnungssatz, 77
Ringschluss, 8
Rolle, Satz von, 145

Sandwich-Theorem, 46
Satz
  Haupt-, 192
    zweiter, 194
  Mittelwert-, 144
    für Integrale, 196
    verallgemeinerter, 146
  Riemannscher Umordnungs-, 77
  von Archimedes, 26
  von Bolzano, Nullstellen-, 120
  von Bolzano-Weierstraß, 58
    in $\mathbb{C}$, 100
  von de l'Hospital, 147
  von Hadamard, 90, 102

von Rolle, 145
von Taylor, 168
mit Integralrestglied, 213
Zwischenwert-, 118
Schnitt von Mengen, 4
Schranke
obere, 21
untere, 21
Sinus, 93
hyperbolicus, 160
in $\mathbb{C}$, 104
Stammfunktion, 192
stetige Differenzierbarkeit, 167
stetige Fortsetzung, 115
Stetigkeit, 107
gleichmäßige, 135
Lipschitz-, 136
streng monoton fallende Funktion, 122
streng monoton fallende Folge, 51
streng monoton wachsende Funktion, 122
streng monoton wachsende Folge, 51
streng monotone Funktion, 122
streng monotone Folge, 51
strikt negative reelle Zahlen, 19
strikt positive reelle Zahlen, 19
Substitutionsregel, 201
Subtraktion, 19
Summe
Partial-, 67
Summenzeichen, 28
Supremum einer Menge, 21
surjektiv, 12

Tangens, 158
hyperbolicus, 160
Taylor, Satz von, 168
mit Integralrestglied, 213
Taylorpolynom, 168
Taylorreihe, 170
Teilfolge, 55
Teilmenge, 4
Teleskopsumme, 31
Totalordnung, 19
Transitivität, 19
trigonometrische Funktionen, 93, 152
Additionstheoreme, 154
trigonometrischer Pythagoras, 154

umgekehrte Dreiecksungleichung, 24
Umkehrfunktion, 13
Differenzierbarkeit der, 142
Umordnung, 73
unbedingt konvergente Reihe, 74
unbestimmtes Integral, 211
uneigentlich integrierbar, 203, 205
uneigentliches Integral, 203, 205
absolut konvergentes, 208
ungerade Funktion, 155
Ungleichung
Bernoullische, 29
Dreiecks-, 24, 185, 208
in $\mathbb{C}$, 98
umgekehrte Dreiecks-, 24
verallgemeinerte Dreiecks-, 74
in $\mathbb{C}$, 101
Unstetigkeit, 107
untere Schranke, 21
unterer Limes, 59
Untersumme, 180
Urbild, 11
Urbildmenge, 11

verallgemeinerte Dreiecksungleichung, 74
verallgemeinerter Mittelwertsatz, 146
Vereinigung von Mengen, 4
Vergleichssatz
fur uneigentliche Integrale und Reihen, 206
Verkettung von Funktionen, 12
Verneinen von Aussagen, 8
Vertauschen von Grenzwerten, 53
vollständige Induktion, 26

Vollständigkeitsaxiom, 21

Wahrheitstafel, 8
wohlgeordnete Menge, 27
Wohlordnungsprinzip, 27
Wurzel, 37
    n-te, 37
Wurzelkriterium, 81
    in $\mathbb{C}$, 101

$\mathbb{Z}$, 25
Zahlen
    ganze, 25
    komplexe, 95
    konjugiert komplexe, 97
    natürliche, 25
    rationale, 33
    reelle, 17
Zerlegung eines Intervalls, 176
Zielmenge, 11
zweite Ableitung, 165
Zwischenvektor, 176
Zwischenwerteigenschaft, 147
Zwischenwertsatz, 118