Analysis 1 *

Martin Brokate **

Inhaltsverzeichnis

---

*Vorlesungsskript, WS 2014/15

**Zentrum Mathematik, TU München

1 Aussagen, Mengen, Induktion, Quantoren

Wir beginnen mit der Frage, wie wir in der Mathematik über etwas reden und was wir erreichen wollen.

Wir gehen davon aus, dass wir die **natürlichen Zahlen**

$$0, 1, 2, 3, 4, \ldots$$

kennen und wissen, wie man sie addiert, subtrahiert, multipliziert und dividiert. (Ob man die Null als natürliche Zahl betrachtet oder nicht, ist reine Verabredungssache. In dieser Vorlesung gehört sie dazu.)

In der Mathematik befassen wir uns mit Aussagen, zum Beispiel mit der Aussage

$$\text{12 ist durch 3 teilbar.} \tag{1.1}$$

Wir wissen, dass (1.1) wahr ist, da $12/3 = 4$ gilt und kein Rest bleibt.

Eine mathematische **Aussage** ist "etwas", was entweder **wahr** oder **falsch** ist. (Das definieren wir nicht genauer. Jedenfalls ist "Servus!" keine mathematische Aussage.)

Wir betrachten als nächstes eine kompliziertere Aussage:

$$\text{Jede durch 6 teilbare natürliche Zahl ist auch durch 3 teilbar.} \tag{1.2}$$

Während in (1.1) nur von zwei ganz bestimmten Zahlen (12 und 3) die Rede ist, handelt (1.2) von "beliebigen" Zahlen. Wir können nun einzelne Zahlen betrachten, die durch 6 teilbar sind, also etwa 6,12,18,24,30. In jedem Fall stellen wir fest, dass sie auch durch 3 teilbar sind. Aber wie wollen wir herausfinden, ob das für jede solche Zahl gilt? Das sind ja unendlich viele. Aber wir werden gleich sehen wie.

Beide Aussagen (1.1) und (1.2) sind Beispiele für mathematische Sätze. Ein **mathematischer Satz** oder kurz **Satz** besteht aus **Voraussetzung** und **Behauptung**.

In (1.1) ist "12 ist durch 3 teilbar" die Behauptung. Die Voraussetzung steht nicht explizit da, sie besteht aus unserem Vorwissen über die Rechenregeln für natürliche Zahlen.

Typischer ist Beispiel (1.2):

$$\begin{aligned} \text{Voraussetzung: } & n \text{ ist eine durch 6 teilbare natürliche Zahl.} \\ \text{Behauptung: } & n \text{ ist durch 3 teilbar.} \end{aligned} \tag{1.3}$$

Sowohl Voraussetzung als auch Behauptung enthalten eine Variable, in diesem Fall ist sie mit "n" bezeichnet. (Unser unausgesprochenes Vorwissen über natürliche Zahlen ist ebenfalls Bestandteil der Voraussetzung.)

Je nachdem, welche Zahl n man in (1.3) betrachtet, können Voraussetzung und Behauptung – für sich genommen – wahr sein oder auch nicht. Ob der Satz (in der Formulierung (1.2) bzw. (1.3)) wahr ist oder nicht, ist eine völlig andere Frage.

Ein mathematischer Satz ist wahr (man sagt auch: **richtig**), falls "aus der Voraussetzung die Behauptung folgt", das heißt, falls die Aussage

$$\text{Immer wenn die Voraussetzung wahr ist, ist auch die Behauptung wahr} \tag{1.4}$$

zutrifft (d.h. wahr ist). Falls (1.4) nicht zutrifft, so ist der Satz falsch.

Um einen Satz als wahr zu erkennen, führt man einen **Beweis**. Beweise können unterschiedliche Struktur haben, wir besprechen zuerst den direkten Beweis. Ein **direkter Beweis** besteht darin, durch "schrittweises logisches Schließen" von der Voraussetzung zur Behauptung zu gelangen. Hier ist ein direkter Beweis für den Satz (1.3):

> Sei n eine beliebige durch 6 teilbare natürliche Zahl. Wir setzen $m = n/6$. Da n nach Voraussetzung durch 6 teilbar ist, ist m ebenfalls eine natürliche Zahl. Es folgt $n = (n/6) \cdot 6 = 6m$ und weiter $n/3 = (6m)/3 = (6/3) \cdot m = 2m$. Da $2m$ eine natürliche Zahl ist, ist auch $n/3$ eine natürliche Zahl. Also ist n durch 3 teilbar.

Der Beweis zeigt, dass (1.4) für unseren Satz zutrifft und dieser somit wahr ist.

Eine Bemerkung: Wir eben haben mit einem 5 Zeilen langen (also aus endlich vielen Zeichen bestehenden) Beweis gesehen, dass eine bestimmte Aussage wahr ist, die sich auf alle (also unendlich viele verschiedene) natürlichen Zahlen bezieht. Solche Situationen kennen wir bereits, beispielsweise die Formel $(a - b)(a + b) = a^2 - b^2$ (Beweis durch Ausmultiplizieren und Zusammenfassen) oder die Aussage, dass die Winkelsumme im Dreieck (von denen es unendlich viele verschiedene gibt) immer 180 Grad beträgt (Beweis mit Hilfsmitteln aus der Geometrie).

Hier ein Beispiel für einen Satz, der falsch ist:

$$\text{Jede durch 6 teilbare natürliche Zahl ist auch durch 5 teilbar.} \tag{1.5}$$

Betrachten wir nämlich die Zahl 6, so ist sie zwar durch 6, aber nicht durch 5 teilbar. Die Aussage (1.4) ist also für diesen Satz falsch, damit auch der Satz. Indem wir ein sogenanntes **Gegenbeispiel** angegeben haben, nämlich die Zahl 6, haben wir also bewiesen, dass der Satz falsch ist. (Ein einziges Gegenbeispiel genügt; dass auch andere Zahlen Gegenbeispiele liefern, etwa 12,18,24,36, ist gleichgültig. Dass es viele Zahlen gibt, die sowohl durch 6 als auch durch 5 teilbar sind, wie etwa 30,60,90, ist ebenfalls irrelevant für die Frage, ob der Satz (1.5) richtig oder falsch ist.)

Noch ein paar Bemerkungen:

- Werden Variablen in Sätzen verwendet, wie etwa "n" in (1.3), so kommt es nicht darauf an, wie man die Variablen nennt. So besagt etwa der mathematische Satz

  $$\begin{aligned} &\text{Voraussetzung: } k \text{ ist eine durch 6 teilbare natürliche Zahl.} \\ &\text{Behauptung: } k \text{ ist durch 3 teilbar.} \end{aligned} \tag{1.6}$$

  dasselbe wie der Satz (1.3).

- Statt "Voraussetzung: X, Behauptung: Y" schreibt man meistens "Sei X, dann gilt Y". Oder so ähnlich.

- Voraussetzung und Behauptung können aus mehreren verknüpften Aussagen bestehen. Beispiel:

> Voraussetzung: n ist eine gerade natürliche Zahl, und n ist durch 5 teilbar.
> Behauptung: n ist durch 10 teilbar und größer als 100.

(Dieser Satz ist falsch, da $n = 20$ die Voraussetzung erfüllt, aber nicht den zweiten Teil der Behauptung. Die Zahl 20 liefert also ein Gegenbeispiel.)

Andere Formen von Beweisen werden wir später besprechen.

Mengen. Die Gegenstände, Begriffe und Objekte, mit denen in der Mathematik hantiert wird, existieren im Denken. Sie können unmittelbare Entsprechungen im Alltagsbereich haben (etwa die Zahl 2), aber auch sehr weit von sinnlichen Erfahrungen entfernt sein. Über das Verhältnis von Mathematik und Realität lässt sich viel sagen und kontrovers diskutieren. Das wollen wir hier nicht tun. (Um Mathematik zu lernen, ist es nicht notwendig, sich mit solchen Fragen zu beschäftigen.)

Mathematische Gegenstände, wie immer man sie auch interpretieren will, werden seit geraumer Zeit in der Sprache der Mengenlehre formuliert. Georg Cantor hat im Jahre 1895 den Begriff einer Menge so definiert:

> Unter einer Menge verstehen wir eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.

Die "Objekte", aus denen sich eine Menge zusammensetzt, heißen **Elemente** der Menge. Ist M eine Menge, so schreiben wir

$$x \in M \qquad (\text{“}x \text{ ist Element von } M\text{”}) \ ,$$

falls x Element von M ist, andernfalls

$$x \notin M \qquad (\text{“}x \text{ ist nicht Element von } M\text{”}) \ .$$

Ein Beispiel ist die Menge aller Wochentage. Sie hat sieben Elemente.

Über die Natur der einzelnen Objekte wird im allgemeinen Mengenbegriff nichts vorausgesetzt. Beispiel: Wir können ohne weiteres die Zahl 2, den Flughafen München, und das Pförtnerhaus in unserem Gebäude als eine Menge mit drei Elementen betrachten. Eine Menge muss lediglich wohldefiniert sein in dem Sinn, dass es sich zu jedem Objekt x eindeutig sagen lässt, ob x zur Menge gehört oder nicht. Normalerweise haben die Elemente einer Menge aber irgend etwas gemeinsam (etwa, Wochentage zu sein, oder Verbindungslinien zwischen zwei Punkten zu sein); sonst hat man kein Motiv, sie zu einer Menge zusammenzufassen.

Mengen kann man auf verschiedene Weise angeben und erhalten. Eine Möglichkeit ist, dass man ihre Elemente einzeln aufschreibt und in geschweifte Klammern setzt. Beispiel: $\{1, 3, 5, 7\}$. Falls es viele Elemente sind, setzt man auch Punkte ein, etwa $\{1, 2, 3, \ldots, 100\}$. Dabei unterstellt man allerdings, dass "klar ist, was gemeint ist". So verfährt man auch gelegentlich bei Mengen mit unendlich vielen Elementen, etwa bei der Menge $\mathbb{N}$ der natürlichen Zahlen

$$\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\} \, ,$$

und der Menge $\mathbb{Z}$ der ganzen Zahlen

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\},$$

oder

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

Eine weitere Möglichkeit, neue Mengen zu erhalten, ist die Bildung von Teilmengen bereits bekannter Mengen. Eine Menge N heißt **Teilmenge** einer Menge M, geschrieben

$$N \subset M \;,$$

wenn jedes Element von N auch Element von M ist; weiterhin nennt man M eine **Obermenge** von N. Beispiele:

$$\{1, 3\} \subset \{0, 1, 2, 3, 4\}, \quad \mathbb{N} \subset \mathbb{Z}, \quad G \subset \mathbb{N},$$

falls wir mit G die Menge aller geraden natürlichen Zahlen bezeichnen.

Teilmengen kann man dadurch angeben, dass man eine sie definierende Eigenschaft formuliert. Beispiel:

$$G = \{n : n \in \mathbb{N}, \text{ es gibt ein } m \in \mathbb{N} \text{ mit } n = 2m\} \tag{1.7}$$

liefert die Menge der geraden Zahlen als Teilmenge von $\mathbb{N}$. Eine solche Beschreibung hat die allgemeine Form

$$N = \{x : x \in M, \ A(x) \text{ ist wahr}\} \tag{1.8}$$

wobei $A(x)$ eine Aussage ist, die die Variable x enthält.

Zwei Mengen M und N heißen **gleich**, geschrieben $M = N$, wenn sie dieselben Elemente enthalten. Das ist gleichbedeutend mit

$$N \subset M \quad \text{und} \quad M \subset N. \tag{1.9}$$

Beispiel:

$$\{1, 3, 5, 7\} = \{1, 5, 7, 3\} = \{1, 3, 5, 7, 5\}.$$

Im rechtsstehenden Ausdruck taucht 5 zweimal auf. Damit ist nicht gemeint, dass die Menge zweimal die 5 enthält, denn alle Elemente müssen "wohlunterschieden" (Cantor) sein, das muss auch in ihrer Bezeichnung zum Ausdruck kommen. Die zweite 5 in $\{1, 3, 5, 7, 5\}$ ist überflüssig. Beschreibt man die Menge aber anders als durch explizite Aufzählung, kann eine solche Situation schon auftreten. Beispiel: Wir beschreiben die rationalen Zahlen $\mathbb{Q}$ durch

$$\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, \ q \neq 0\}. \tag{1.10}$$

Als Ergebnis von "p/q" tauchen alle rationalen Zahlen mehrfach (sogar unendlich oft) auf, die Zahl 3 beispielsweise nicht nur als $3/1$, sondern auch in den "Verkleidungen" $6/2$, $9/3$, $(-3)/(-1)$ … Hier wie im Allgemeinen wäre es sehr unzweckmäßig, würde man nur solche Beschreibungen von Mengen zulassen, in denen jedes Element genau einmal genannt wird.

Es kann sehr wohl passieren, dass eine Beschreibung einer Menge dazu führt, dass diese keine Elemente enthält. Beispiel:

$$N = \{n : n \in \mathbb{N}\,, n = n + 1\}$$

enthält keine Elemente. Es liegt daher nahe, die leere Menge

$$\emptyset\,,$$

welche kein Element enthält, ebenfalls als Menge zuzulassen. Die leere Menge spielt unter den Mengen eine ähnliche Rolle wie die Null bei den Zahlen.

Falls $N \subset M$ und $N \neq M$, so gibt es mindestens ein Element von M, welches nicht in N liegt. Wir sagen dann, dass N eine echte Teilmenge von M ist, und schreiben

$$N \subsetneqq M\,.$$

Indizes und Summenzeichen. Nehmen wir einmal an, wir wollen die ersten zehn Quadratzahlen addieren. Ihre Summe s ist gegeben durch

$$s = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 + 8^2 + 9^2 + 10^2\,.$$

Verwendet man Punkte, sieht es z.B. so aus:

$$s = 1^2 + 2^2 + \cdots + 10^2\,.$$

In der Mathematik verwendet man das Summenzeichen

$$s = \sum_{k=1}^{10} k^2\,. \tag{1.11}$$

Es bedeutet, dass der nach dem Summenzeichen stehende Ausdruck für jedes $k \in \mathbb{N}$ zwischen 1 und 10 (jeweils einschließlich) ausgewertet wird und alle so entstehenden Ausdrücke addiert werden. Die Zahl 1 heißt untere Grenze, die Zahl 10 obere Grenze, und k die Laufvariable der Summe.

In (1.11) taucht eine Variable im Summanden auf. Genausogut können Variable in den Grenzen vorkommen, also etwa

$$s = \sum_{k=m}^{n} k^2\,. \tag{1.12}$$

In diesem Fall wird von m bis n summiert. Der Wert von s hängt also davon ab, wie groß m und n sind.

Es kann auch sein, dass man sich nicht von vorneherein festlegen will, welche Zahlen man addieren will. Weiß man, dass es 10 sind, so kann man schreiben

$$s = a + b + c + d + e + f + g + h + i + j\,,$$

wobei $a, \ldots, j$ Variable sind. Besser ist es aber, mit Indizes zu arbeiten. Wir bezeichnen die 10 einzelnen Variablen nicht mit unterschiedlichen Buchstaben, sondern mit

$$a_1, a_2, a_3, \ldots, a_{10}\,,$$

also mit

$$a_k\,,$$

wobei k Werte zwischen 1 und 10 annimmt. Hier heißt k der **Index** von a_k, und a eine **mit k indizierte** Variable. Die Summe wird dann zu

$$s = \sum_{k=1}^{10} a_k\,.$$

Wie in (1.12) können wir die Grenzen variabel gestalten, etwa

$$s = \sum_{k=m}^{n} a_k\,. \tag{1.13}$$

Formeln dieser Bauart (Variable an unterschiedlichen Stellen sowie Indizes) treten in der Mathematik ständig auf. Am Anfang ist das ungewohnt, da ihr Abstraktionslevel höher ist als der von vielen Dingen des Alltags. Am besten ist es, dem nicht auszuweichen, sondern möglichst schnell sich darauf einzulassen, passiv (sich um Verständnis bemühen) und aktiv (selber damit arbeiten).

Je nachdem welchen Zweck man verfolgt, kann man die Laufvariable im Summenzeichen anpassen. So bedeuten die Ausdrücke

$$\sum_{k=1}^{10} a_k \qquad \text{und} \qquad \sum_{j=1}^{10} a_j$$

dasselbe, ebenso

$$\sum_{k=1}^{10} a_k \qquad \text{und} \qquad \sum_{k=3}^{12} a_{k-2}\,.$$

Im letzteren Fall spricht man von **Indexverschiebung**.

Falls im Ausdruck

$$\sum_{k=m}^{n} a_k$$

die obere Grenze kleiner ist als die untere, also $n < m$, so nennt man ihn die **leere Summe**, sie hat den Wert 0 (eine Konvention, die sich als praktisch erweist).

Völlig analog geht man bei Produkten vor. Statt des Summenzeichens verwendet man das Produktzeichen

$$\prod\,.$$

Das Produkt der ersten 10 natürlichen Zahlen beispielsweise ist gegeben durch

$$\prod_{k=1}^{10} k\,.$$

Man bezeichnet es auch mit 10! ("10 Fakultät"). Für beliebiges $n \in \mathbb{N}$ ist die Fakultät definiert durch

$$n! = \prod_{k=1}^{n} k\,. \tag{1.14}$$

Im Falle $n = 0$ soll gelten
$$0! = 1\,.$$
In der Tat, das passt zur Konvention, dass das **leere Produkt** den Wert 1 hat.

Vollständige Induktion. "Vollständige Induktion" ist eine weitere Methode, um Aussagen zu beweisen, die sich auf alle natürlichen Zahlen beziehen. Das geht folgendermaßen: Sei $\mathrm{A}(n)$ eine Aussage, die $n \in \mathbb{N}$ als Variable enthält.

1. Man nimmt sich ein $n_0 \in \mathbb{N}$ und beweist, dass $\mathrm{A}(n_0)$ wahr ist (**Induktionsverankerung** oder **Induktionsanfang**).
2. Man beweist, dass aus der Gültigkeit von $\mathrm{A}(n)$ auch die Gültigkeit von $\mathrm{A}(n+1)$ folgt (**Induktionsschritt**). Die Aussage $\mathrm{A}(n)$ heißt die **Induktionsannahme**, die Aussage $\mathrm{A}(n+1)$ die **Induktionsbehauptung**.

Das Beweisprinzip der vollständigen Induktion besagt: Hat man diese beiden Schritte erfolgreich ausgeführt, so hat man damit bewiesen

$\mathrm{A}(n)$ gilt für alle $n \in \mathbb{N}$ mit $n \geq n_0$.

Wir wenden das Beweisprinzip der vollständigen Induktion an.

Satz 1.1 *Für alle $n \in \mathbb{N}$ gilt*
$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}\,. \tag{1.15}$$

Beweis: Durch vollständige Induktion. Die Formel (1.15) entspricht der Aussage $\mathrm{A}(n)$. Wir setzen $n_0 = 1$. Induktionsanfang: Es ist
$$\sum_{k=1}^{1} k = 1 = \frac{1(1+1)}{2}\,,$$
also ist $\mathrm{A}(1)$ wahr. Wir führen den Induktionsschritt aus: Sei $\mathrm{A}(n)$ wahr für ein beliebiges gegebenes $n \in \mathbb{N}$, es gelte also
$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}\,. \tag{1.16}$$
Das ist die Induktionsannahme. Die zu beweisende Induktionsbehauptung $\mathrm{A}(n+1)$ ergibt sich, indem wir n in (1.16) durch $n+1$ ersetzen, also
$$\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}\,.$$
Wir rechnen
$$\sum_{k=1}^{n+1} k = \sum_{k=1}^{n} k + (n+1) = \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+2)(n+1)}{2}\,, \tag{1.17}$$

also ist A$(n+1)$ wahr. (Beim zweiten Gleichheitszeichen wurde die Induktionsannahme A(n) verwendet.) Gemäß Beweisprinzip der vollständigen Induktion ist also $A(n)$ wahr für alle $n \geq 1$. Für $n = 0$ gilt (1.15) gemäß der Konvention über die leere Summe. $\square$

Das Zeichen "$\square$" bedeutet: Hier ist der Beweis zu Ende.

Die wesentliche Idee des Beweises steckt in der ersten Gleichheit in (1.17): Die Summe bis $n+1$, über die etwas bewiesen werden muss, wird aufgespalten in die Summe bis n, deren Wert nach Induktionsannahme bereits bekannt ist, und den Term $(n+1)$.

Normalerweise ist der Induktionsschritt deutlich schwieriger als der Induktionsanfang. Trotzdem darf man den Induktionsanfang nicht aus den Augen verlieren: Wir betrachten als Beispiel die Aussage

Für alle $n \geq 1$ gilt $1^n = 0$.

Wir wissen, dass sie falsch ist. Aber der Induktionsschritt funktioniert: Falls wir annehmen, dass $1^n = 0$ gilt, so folgt $1^{n+1} = 1 \cdot 1^n = 1 \cdot 0 = 0$. "Nur" der Induktionsanfang geht nicht, jedenfalls nicht für $n_0 = 1$, denn $1^1 = 1 \neq 0$. Man kann auch kein anderes n_0 nehmen, denn man kann andererseits mit vollständiger Induktion den Satz

Für alle $n \geq 1$ gilt $1^n \neq 0$

beweisen. Denn $1^1 \neq 0$, und aus $1^n \neq 0$ folgt $1^{n+1} = 1 \cdot 1^n = 1^n \neq 0$.

Geordnete Paare. Das Produkt zweier Mengen. Punkte in der Ebene können wir als ein Zahlenpaar (x, y) darstellen. Die Reihenfolge der beiden Zahlen ist wesentlich, die Punkte $(1, 2)$ und $(2, 1)$ beispielsweise sind verschieden. (Im Gegensatz dazu spielt es bei einer Menge keine Rolle, in welcher Reihenfolge die Elemente angegeben werden, es ist $\{1, 2\} = \{2, 1\}$.) Man nennt

$$(x, y)$$

ein **geordnetes Paar**. Die allgemeine Situation ist die folgende:

Seien M und N zwei Mengen. Wir definieren die **Produktmenge** (oder kurz das **Produkt**) $M \times N$ durch

$$M \times N = \{(x, y) : x \in M,\ y \in N\}\,, \tag{1.18}$$

also als Menge aller geordneten Paare, für die die erste Komponente ein Element von M und die zweite Komponente ein Element von N ist. Die Mengen M und N heißen die **Faktoren** von $M \times N$.

Aus der Zahlengerade $\mathbb{R}$ erhalten wir auf diese Weise die Ebene $\mathbb{R} \times \mathbb{R}$, und darin $\mathbb{Z} \times \mathbb{Z}$ als das Gitter der Punkte mit ganzzahligen Koordinaten. (Die mit diesem Beispiel verbundene Vorstellung eines rechten Winkels zwischen den Faktoren ist aber nicht Bestandteil der Definition (1.18)!)

Eine Nebenbemerkung: Wenn man will, kann man den Begriff des geordneten Paares auf den Mengenbegriff zurückführen, indem man setzt

$$(x, y) = \{\{x\}\,,\ \{x, y\}\}\,.$$

Dadurch erreicht man, dass $(x, y) \neq (y, x)$ gilt, falls $x \neq y$.

Verknüpfung von Aussagen. Beim Hantieren mit Aussagen wird eine Reihe von logischen Verknüpfungen verwendet. Für uns sind die folgenden fünf Verknüpfungen wesentlich.

1. Wir beginnen mit der **Negation**. Die Negation von “das Glas ist voll” ist “das Glas ist nicht voll”. Die Negation einer Aussage A ist genau dann wahr, wenn A falsch ist. Die Aussage “das Glas ist leer” ist also *nicht* die Negation von “das Glas ist voll”, denn es ist auch möglich, dass das Glas weder voll noch leer ist. Man bezeichnet die Negation von A mit “$\neg$ A”. Man kann das Verhältnis zwischen A und $\neg$ A übersichtlich mit einer **Wahrheitstafel** ausdrücken. Die Wahrheitstafel für die Negation hat die Form

A	$\neg$ A
W	F
F	W

   Auf der linken Seite stehen die beiden möglichen Wahrheitswerte von A (mit W für wahr und F für falsch), auf der rechten Seite die zugehörigen Wahrheitswerte von $\neg$ A.

2. Als nächstes kommt die **Und-Verknüpfung** (oder **Konjunktion**) zweier Aussagen A und B (“A und B”, “A $\wedge$ B”). Die Aussage

   “das Glas ist voll und der Eimer ist leer”

   ist genau dann wahr, wenn beide Teilaussagen wahr sind. So sind wir das auch im Alltag gewohnt. Die zugehörige Wahrheitstafel ist

A	B	A $\wedge$ B
W	W	W
W	F	F
F	W	F
F	F	F

   Hier hat die linke Seite zwei Spalten. Ihre Zeilen enthalten alle möglichen Kombinationen der Wahrheitswerte von A und B. In der rechten Spalte stehen die zugehörigen Werte von A $\wedge$ B.

   Hier wie auch in allen anderen Fällen gilt: Der Wahrheitswert der durch die Verknüpfung erzeugten Aussage ist durch den Wahrheitswert der an der Verknüpfung beteiligten Aussagen eindeutig festgelegt.

3. Die **Oder-Verknüpfung** (oder **Adjunktion**, geschrieben als “A $\vee$ B”) bedeutet in der Mathematik immer das einschließende “oder”. Sagt man “das Glas ist voll oder der Eimer ist voll”, so kann auch der Fall vorliegen, dass beides voll ist. Die zugehörige Wahrheitstafel ist

A	B	A $\vee$ B
W	W	W
W	F	W
F	W	W
F	F	F

Meint man in einem mathematischen Text “oder, aber nicht beides”, so muß man explizit “entweder – oder” sagen.

4. Die **Implikation** (“Aus A folgt B”, “A impliziert B”, “A $\Rightarrow$ B”) beschreibt das, was in einem mathematischen Satz passiert, denn

   mathematische Sätze sind Aussagen der Form “A $\Rightarrow$ B”.

   Wir fragen uns nun, wann “A $\Rightarrow$ B” gelten soll. Einleuchtend ist, dass “A $\Rightarrow$ B” falsch sein muss, wenn A wahr und B falsch ist (denn B soll ja aus A folgen). Die anderen drei Fälle kann man sich am Beispiel der Aussage

   “für alle Zahlen $n \in \mathbb{N}$ gilt: Aus $n < 3$ folgt $n < 5$”

   klarmachen. Die Wahrheitswerte von “A $\Rightarrow$ B” sollen so festgelegt werden, dass diese Aussage wahr ist, egal welche Zahl man für n einsetzt. Wählen wir $n = 2, 4, 6$, so erhalten wir gerade die drei ausstehenden Fälle. Die Wahrheitstafel hat also die Form

A	B	A $\Rightarrow$ B
W	W	W
W	F	F
F	W	W
F	F	W

   In der dritten und vierten Zeile wird der Wahrheitswert der Aussage “A$\Rightarrow$B” auch für den Fall festgelegt, dass A falsch ist. Das kommt einem zunächst merkwürdig vor, es passt aber dazu, wie man die Implikation normalerweise verwendet: Nehmen wir an, wir wissen, dass “A$\Rightarrow$B” wahr ist. Es sind zwei Fälle möglich:

   - A ist wahr. Dann ist B auch wahr.
   - A ist falsch. Dann können wir nichts darüber sagen, ob B wahr oder falsch ist.

   Das entspricht gerade den drei Zeilen der Wahrheitstafel, in denen rechts W steht.

   Zur Notation: Die Aussage “B $\Leftarrow$ A” bedeutet dasselbe wie “A $\Rightarrow$ B”.

   Ist “A $\Rightarrow$ B” wahr, so sagt man auch: A ist **hinreichend** für B, B ist **notwendig** für A. Sprachlich genauer wäre: Die Gültigkeit von A ist eine hinreichende Bedingung für die Gültigkeit von B.

5. Die **Äquivalenz** zweier Aussagen A und B bedeutet einfach: Wenn wir wissen, dass A gilt, so wissen wir auch, dass B gilt, und umgekehrt. Wir sagen: “A gilt genau dann, wenn B gilt” oder “A und B sind äquivalent”, und schreiben “A $\Leftrightarrow$ B”. Die Wahrheitstafel der Äquivalenz entsteht, indem wir die beiden Spalten für “A$\Rightarrow$B” und “B$\Rightarrow$A mit “und” verknüpfen. Das Ergebnis ist

A	B	A $\Leftrightarrow$ B
W	W	W
W	F	F
F	W	F
F	F	W

Äquivalente Aussagen tauchen beim Umformen von Ausdrücken ständig auf. So gilt etwa für beliebige Zahlen x

$$7x = 3 \qquad \Leftrightarrow \qquad 21x = 9\,.$$

Sie ergeben sich auch beim Umformen von zusammengesetzten Aussagen. Wir suchen etwa die Negation von

$$\text{“das Essen ist heiß und scharf”.} \tag{1.19}$$

Sie lautet

$$\text{“das Essen ist nicht heiß oder nicht scharf”.} \tag{1.20}$$

Dass (1.20) die Negation von (1.19) ist, basiert auf der Regel, dass die Negation einer Aussage der Form $\mathrm{A} \wedge \mathrm{B}$ “dasselbe ist wie” (genauer: äquivalent ist zur) Aussage $(\neg\mathrm{A}) \vee (\neg\mathrm{B})$. Oder, vollständig in Formeln ausgedrückt,

$$\neg(\mathrm{A} \wedge \mathrm{B}) \quad \Leftrightarrow \quad (\neg\mathrm{A}) \vee (\neg\mathrm{B})\,. \tag{1.21}$$

Dass (1.21) wirklich für beliebige Aussagen A und B richtig ist, kann man nachprüfen, indem man die Wahrheitstafeln für die beiden Aussagen $\neg(\mathrm{A}\wedge\mathrm{B})$ und $(\neg\mathrm{A}) \vee (\neg\mathrm{B})$ aufstellt und feststellt, dass sie gleich sind.

Man beachte: Die Aussage

$$\text{“das Essen ist nicht heiß und nicht scharf”} \tag{1.22}$$

(also weder heiß noch scharf) ist *nicht* die Negation von (1.19)!

Die Negation einer durch eine Oder-Verknüpfung entstandenen Aussage erhält man analog, es gilt

$$\neg(\mathrm{A} \vee \mathrm{B}) \quad \Leftrightarrow \quad (\neg\mathrm{A}) \wedge (\neg\mathrm{B})\,. \tag{1.23}$$

Quantoren. Sie liefern weitere Bausteine zur Formulierung mathematischer Aussagen. Es handelt sich dabei um den **Existenzquantor**

$$\exists \qquad (\text{“es gibt”})$$

und den **Allquantor**

$$\forall \qquad (\text{“für alle”})\,.$$

Beispiel: Die Aussage

$$\text{es gibt eine natürliche Zahl, die größer ist als 1000} \tag{1.24}$$

ist wahr. Mit dem Existenzquantor schreibt man sie als

$$\exists\, n \in \mathbb{N} \text{ mit } n > 1000$$

oder noch kürzer als

$$\exists\, n \in \mathbb{N}: \; n > 1000\,.$$

In der anfangs besprochenen Aussage (1.2),

$$\text{Jede durch 6 teilbare natürliche Zahl ist auch durch 3 teilbar} \tag{1.25}$$

taucht der Allquantor auf: Schreiben wir "6 teilt n" (in Zeichen: "$6|n$") statt "n ist durch 6 teilbar", so wird (1.25) zu

$$\forall\, n \in \mathbb{N}:\; 6|n \;\Rightarrow\; 3|n\,.$$

Bei Negation gehen "$\exists$" und "$\forall$" ineinander über. Die Negation von (1.24) ist die (falsche) Aussage

"Für alle $n \in \mathbb{N}$ gilt $n \leq 1000$", "$\forall n \in \mathbb{N}$: $n \leq 1000$" .

Eine rein textliche Formulierung wäre: Alle natürlichen Zahlen sind kleiner als oder gleich 1000.

Zwei Beispiele für eine *falsche* Bildung der Negation von (1.24) sind

"es gibt ein $n \in \mathbb{N}$ mit $n \leq 1000$" ,

"alle natürlichen Zahlen sind größer als 1000" .

Mit "es gibt ein" ist in der Mathematik immer gemeint "es gibt mindestens ein". Will man ausdrücken, dass es auch nicht mehr als eins geben kann, sagt man "es gibt genau ein", Symbol "$\exists!$" oder "$\exists|$". Diese Ausdrucksweise unterscheidet sich etwas von der Umgangssprache, man vergleiche mit

in Garching gibt es Hunde,
in Garching gibt es einen Hund.

Hier schwingen noch Untertöne mit, teils sprecherabhängig, die in der mathematischen Formulierung nicht vorhanden sind.

Aussagen können mehrere Quantoren enthalten. Beispiel:

$$\text{“}\forall n \in \mathbb{N}\ \exists p \in \mathbb{N}\text{: } p > n \text{ und } p \text{ ist Primzahl”} \tag{1.26}$$

In Worten: "Für jede natürliche Zahl n gibt es eine natürliche Zahl p, welche größer als n und Primzahl ist". Die Negation dieser Aussage ist

"$\exists n \in \mathbb{N}\ \forall p \in \mathbb{N}$: $p \leq n$ oder p ist nicht Primzahl",

oder als Text

"es gibt eine natürliche Zahl n, so dass für alle natürlichen Zahlen p gilt, dass p kleiner als oder gleich n ist oder dass p nicht Primzahl ist".

Man könnte die ursprüngliche Aussage kürzer auch so schreiben:

"$\forall n\ \exists p > n$: p ist Primzahl".

Allerdings muss dann zwischen Schreiber und Leser (oder Sprecher und Zuhörer) klar sein, dass für n und p nur natürliche Zahlen in Frage kommen.

Liest man eine Aussage, die zwei Quantoren enthält, so erfasst man ihre Bedeutung oft nicht auf den ersten Blick; immer wieder auch dann nicht, wenn man sich schon länger mit Mathematik beschäftigt, und schon gar nicht, wenn die Aussage mehr als zwei Quantoren enthält. Man muss erstmal in Ruhe nachdenken, was gemeint ist. Auch hier gilt: Verständnis ist wichtiger als Geschwindigkeit.

Treten mehrere Quantoren auf, so ist deren Reihenfolge wesentlich. Die Aussage

> "$\exists p \in \mathbb{N}\ \forall n \in \mathbb{N}$: $p > n$ und p Primzahl"

ist eine andere Aussage als (1.26), die im Gegensatz zu (1.26) falsch ist. Solche Situationen treten auch in der Umgangssprache auf. Beispiel:

> "In jeder deutschen Stadt gibt es einen Bürger, der ein Haus besitzt"

ist eine (wohl wahre) Aussage, während

> "es gibt einen Bürger, der in jeder deutschen Stadt ein Haus besitzt"

eine andere (vermutlich falsche) Aussage ist.

Kontraposition und Widerspruchsbeweis. Der eingangs betrachtete Satz

Jede durch 6 teilbare natürliche Zahl ist auch durch 3 teilbar

hat wie jeder Satz die Form

$$A \Rightarrow B\,. \tag{1.27}$$

Er wurde direkt bewiesen: Es wurden Zwischenbehauptungen $C_1, \ldots, C_n$ gefunden wurden, so dass eine Kette von Implikationen

$$A \Rightarrow C_1\,, \quad C_1 \Rightarrow C_2\,, \quad \ldots\,, \quad C_{n-1} \Rightarrow C_n\,, \quad C_n \Rightarrow B$$

entstand, die jede für sich als wahr erkannt wurden.

Wir betrachten nun als Beispiel den Satz

$$\text{Ist } n^2 \text{ gerade, so ist auch } n \text{ gerade.} \tag{1.28}$$

Er hat die Form $A \Rightarrow B$, wobei A die Aussage "n^2 ist gerade" und B die Aussage "n ist gerade" ist. Wir führen folgenden Beweis:

> Sei n ungerade. Dann ist $n-1$ gerade. Wir setzen $k = (n-1)/2$, dann ist $k \in \mathbb{N}$ und $n - 1 = 2k$, $n = 2k + 1$. Weiter ist $n^2 = (2k+1)^2 = 4k^2 + 4k + 1$, also $n^2 - 1 = 2(2k^2 + 2k)$ gerade und damit n^2 ungerade.

Wir haben also für den Satz

$$\neg B \Rightarrow \neg A \tag{1.29}$$

einen direkten Beweis geführt. Diese Methode, den ursprünglichen Satz (1.27) zu beweisen, nennt man **Kontraposition**.

Man fragt sich: Funktioniert das immer, egal was A und B für Aussagen sind? Die Antwort ist "ja", denn die Aussagen

$$A \Rightarrow B \quad \text{und} \quad \neg B \Rightarrow \neg A$$

sind äquivalent. (Siehe Übungsaufgabe.)

Eine weitere Beweismethode für den Satz "A⇒ B" ist der **Widerspruchsbeweis**, auch **indirekter Beweis** genannt. Dieser läuft darauf hinaus, zu zeigen, dass die Aussagen A und ¬ B nicht beide wahr sein können. (Dann ist es unmöglich, dass A wahr und B falsch ist, und somit ist die Aussage "A⇒ B" wahr, das heißt, der Satz "A⇒ B" ist bewiesen.)

Als Beispiel für einen Widerspruchsbeweis betrachten wir den berühmten Beweis von Euklid für den Satz

$$\sqrt{2} \text{ ist irrational.} \tag{1.30}$$

Mit (1.30) meinen wir: Es gibt keine rationale Zahl x mit $x^2 = 2$.

Die Aussage (1.30) steht für die Behauptung B; die Rolle der Voraussetzung A wird von den bekannten Rechenregeln für Zahlen und Brüche übernommen. Der Beweis geht so:

> Wir nehmen an, das ¬ B gilt. Dann gibt es ein $x \in \mathbb{Q}$ mit $x^2 = 2$. Da dann auch $(-x)^2 = 2$ gilt, können wir $x \geq 0$ annehmen. Seien $p, q \in \mathbb{N}$ mit $x = p/q$. Nach Auskürzen des Bruchs erhalten wir Zahlen $r, s \in \mathbb{N}$ mit $x = r/s$, und für diese Zahlen gilt
>
> r und s sind teilerfremd. (Aussage C)
>
> Es folgt $2s^2 = r^2$, also ist 2 ein Teiler von r^2 und damit auch von r, also gibt es ein $t \in \mathbb{N}$ mit $r = 2t$. Es folgt weiter $2s^2 = 4t^2$, also $s^2 = 2t^2$, also ist 2 auch ein Teiler von s und damit gilt
>
> r und s sind nicht teilerfremd. (Aussage ¬ C)
>
> Wir haben "einen Widerspruch hergestellt". Also ist der behauptete Satz (1.30) wahr. □

Für diejenigen, denen die vor dem Beweis von (1.30) gegebene Erläuterung des Prinzips eines Widerspruchsbeweises nicht formal genug ist: Ein Widerspruchsbeweis zeigt, dass die Implikation

$$A \wedge (\neg B) \quad \Rightarrow \quad C \wedge (\neg C)$$

wahr ist, wobei C eine weitere Aussage ist. Da aber C ∧ (¬ C) immer falsch ist (egal was C ist), muss die Aussage A ∧ (¬ B) falsch sein. Nun zeigt ein Vergleich der Wahrheitstafeln, dass die Aussagen

$$\text{A} \wedge (\neg \text{B}) \quad \text{und} \quad \neg (\text{A} \Rightarrow \text{B})$$

äquivalent sind, somit ist "A⇒ B" wahr.

2 Die reellen Zahlen

Wir stellen uns die reellen Zahlen wie gewohnt als Punkte auf der Zahlengeraden vor. Aber was für mathematische Objekte sind sie, und welche Eigenschaften haben sie?

Geht man davon aus, dass man die natürlichen Zahlen $\mathbb{N}$ als Menge bereits kennt und weiß man, wie in $\mathbb{N}$ Addition und Multiplikation definiert sind, so lassen sich daraus sukzessive die ganzen Zahlen $\mathbb{Z}$, die rationalen Zahlen $\mathbb{Q}$ und die reellen Zahlen $\mathbb{R}$ erhalten. Dieser Konstruktionsprozess ist einerseits ganz interessant, und man kann ihn dazu benutzen, einige grundlegende mathematische Sachverhalte kennenzulernen; andererseits ist er aber auch recht langwierig und für das Verständnis der modernen Analysis nicht so wesentlich. Wir behandeln ihn daher in dieser Vorlesung nicht.

Die Alternative ist, die reellen Zahlen $\mathbb{R}$ axiomatisch als Menge mit gewissen Eigenschaften einzuführen. Dieses Vorgehen geht auf David Hilbert zurück.

Die axiomatisch verlangten Eigenschaften von $\mathbb{R}$ zerfallen in drei Gruppen. Die erste Gruppe enthält die sogenannten Körpereigenschaften: Es gibt eine Addition und eine Multiplikation, die aus zwei Zahlen $x, y \in \mathbb{R}$ deren Summe $x + y$ und Produkt $x \cdot y$ erzeugen, und zwar so, dass $(\mathbb{R}, +, \cdot)$ ein **kommutativer Körper** ist. Das bedeutet, dass Addition und Multiplikation die folgenden Eigenschaften haben.

(1) Für die Addition gilt das Assoziativgesetz

$$(x + y) + z = x + (y + z)\,, \qquad \text{für alle } x, y, z \in \mathbb{R}\,, \tag{2.1}$$

das Kommutativgesetz

$$x + y = y + x\,, \qquad \text{für alle } x, y \in \mathbb{R}\,, \tag{2.2}$$

weiterhin gibt es eine Zahl $0 \in \mathbb{R}$ mit

$$x + 0 = x\,, \qquad \text{für alle } x \in \mathbb{R}\,, \tag{2.3}$$

und zu jedem $x \in \mathbb{R}$ existiert genau ein Element in $\mathbb{R}$, bezeichnet mit $(-x)$, für das gilt

$$x + (-x) = 0\,, \tag{2.4}$$

(das inverse Element bezüglich der Addition).

(2) Für die Multiplikation gilt das Assoziativgesetz

$$(x \cdot y) \cdot z = x \cdot (y \cdot z)\,, \qquad \text{für alle } x, y, z \in \mathbb{R}\,, \tag{2.5}$$

das Kommutativgesetz

$$x \cdot y = y \cdot x\,, \qquad \text{für alle } x, y \in \mathbb{R}\,, \tag{2.6}$$

weiterhin gibt es eine Zahl $1 \in \mathbb{R}$ mit $1 \neq 0$ und

$$1 \cdot x = x\,, \qquad \text{für alle } x \in \mathbb{R}\,, \tag{2.7}$$

und zu jedem $x \in \mathbb{R}$ mit $x \neq 0$ existiert genau ein Element in $\mathbb{R}$, bezeichnet mit x^{-1}, mit

$$x \cdot x^{-1} = 1\,, \tag{2.8}$$

(das inverse Element bezüglich der Multiplikation).

(3) Es gilt das Distributivgesetz

$$x \cdot (y + z) = x \cdot y + x \cdot z\,, \qquad \text{für alle } x, y, z \in \mathbb{R}\,. \tag{2.9}$$

Die zweite Gruppe bezieht sich auf die Ordnungseigenschaften von $\mathbb{R}$. Wir verlangen, dass es eine Teilmenge P von $\mathbb{R}$ (den Positivbereich) gibt mit den Eigenschaften

(4) Für jedes $x \in \mathbb{R}$ ist genau eine der folgenden drei Aussagen wahr:

$$x \in P\,, \quad x = 0\,, \quad -x \in P\,. \tag{2.10}$$

(5) Aus $x, y \in P$ folgt $x + y \in P$ und $x \cdot y \in P$.

Die "Größer-Kleiner-Gleich-Beziehungen" (wir werden sie später als Relationen bezeichnen) definieren wir mit Hilfe der Menge P durch

$$x > y \qquad \text{genau dann, wenn} \qquad x - y \in P\,, \tag{2.11}$$
$$x \geq y \qquad \text{genau dann, wenn} \qquad x > y \text{ oder } x = y\,, \tag{2.12}$$
$$x < y \qquad \text{genau dann, wenn} \qquad y - x \in P\,, \tag{2.13}$$
$$x \leq y \qquad \text{genau dann, wenn} \qquad x < y \text{ oder } x = y\,. \tag{2.14}$$

Mit $\mathbb{R}_+$ bezeichen wir die nichtnegativen reellen Zahlen, $\mathbb{R}_+ = \{x : x \in \mathbb{R},\ x \geq 0\}$.

Die bisher formulierten Eigenschaften (1) – (5) treffen auch auf $\mathbb{Q}$ zu. Es ist daher nicht möglich, allein aus ihnen die Existenz der Zahl $\sqrt{2}$ (das heißt, einer Zahl x mit $x^2 = 2$) zu beweisen, denn wir wissen ja seit Euklid, dass eine solche Zahl nicht rational sein kann. Nun soll $\mathbb{R}$ aber alle Zahlen, mit denen wir rechnen wollen, enthalten, neben den Wurzeln beispielsweise auch π. Zu den bisher verlangten Eigenschaften muss also noch mindestens eine weitere treten, aus denen die Existenz solcher Zahlen folgt. Dafür gibt es verschiedene Möglichkeiten.

In dieser Vorlesung verwenden wir das sogenannte Supremumsaxiom. Wie das genau aussieht, werden wir weiter hinten in diesem Kapitel besprechen.

"Axiomatische Einführung der reellen Zahlen" bedeutet nun, dass wir das folgende Axiomensystem postulieren.

Axiome 2.1 *Es existieren eine Menge $\mathbb{R}$ versehen mit einer Addition und einer Multiplikation, eine Teilmenge P von $\mathbb{R}$ und Elemente $0, 1 \in \mathbb{R}$, welche die Eigenschaften* (1) *bis* (5) *sowie das Supremumsaxiom 2.26 erfüllen.*

Man kann beweisen, dass die reellen Zahlen durch die im Axiomensystem 2.1 genannten Eigenschaften "eindeutig charakterisiert werden". Wir gehen jetzt nicht darauf ein, was das genau bedeutet und wie man das beweist. Gemeint ist dasselbe, wie wenn man sagt, dass das Schachspiel als solches eindeutig charakterisiert ist durch das Schachbrett mit 8 mal 8 Quadraten, die Anfangsstellung und die Zugregeln für die einzelnen Figuren – obwohl Bretter und Figuren unterschiedliche Größe und Form haben können.

Ebenso kann beweisen (was wir auch nicht tun werden), dass die reellen Zahlen

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

“dasselbe sind wie die natürlichen Zahlen $\mathbb{N}$”.

Das weitere Vorgehen ist: Wir besprechen als erstes Eigenschaften von $\mathbb{R}$, die sich nicht auf das Supremumsaxiom beziehen, dann stellen wir das Supremumsaxiom vor und behandeln weitere Eigenschaften von $\mathbb{R}$.

Folgerungen aus den Körpereigenschaften. Wir führen die Subtraktion auf die Addition und die Division auf die Multiplikation zurück, indem wir definieren

$$x - y = x + (-y)\,, \quad \forall\, x, y \in \mathbb{R}\,, \qquad \frac{x}{y} = xy^{-1}\,, \quad \forall\, x, y \in \mathbb{R} \text{ mit } y \neq 0\,. \tag{2.15}$$

Für die vier Grundrechenarten gelten eine Reihe von Rechenregeln, die später ständig verwendet werden. Wir stellen sie zusammen. Man kann sie mit (kurzen) Beweisen aus den Axiomen herleiten. Beispiel: Was ist $-(-x)$? Von der Anschauung der Zahlengerade her ist klar, dass x herauskommen muss, da der Übergang zum Negativen einer Spiegelung am Nullpunkt entspricht. Auf der Grundlage von Axiom 2.1 argumentieren wir folgendermaßen. Gemäß (2.4) ist $-(-x)$ diejenige Zahl z, für die $(-x) + z = 0$ gilt. Das ist aber für $z = x$ der Fall. Also ist $-(-x) = x$.

Hier sind die Rechenregeln.

$$-\,(-x) = x\,, \quad x \cdot 0 = 0\,, \quad (-x)y = -(xy)\,, \quad -(x+y) = -x - y\,, \quad \forall\, x, y \in \mathbb{R}\,, \tag{2.16}$$

$$(x^{-1})^{-1} = x\,, \quad (xy)^{-1} = x^{-1}y^{-1}\,, \quad \forall\, x, y \in \mathbb{R} \setminus \{0\}\,, \tag{2.17}$$

$$xy = 0 \qquad \Leftrightarrow \qquad x = 0 \text{ oder } y = 0\,. \tag{2.18}$$

Wir betrachten als nächstes die Gleichung $a + x = b$, wobei $a, b \in \mathbb{R}$ gegeben sind und $x \in \mathbb{R}$ gesucht wird. Subtrahieren wir a auf beiden Seiten, so ergibt sich $x = b - a$. Da das eine äquivalente Umformung ist (wenn wir wieder auf beiden Seiten a addieren, erhalten wir die ursprüngliche Gleichung), löst diese und keine andere Zahl die gegebene Gleichung. Wir sagen:

$x = b - a$ ist die eindeutige Lösung der Gleichung $a + x = b$.

Analog ergibt sich $x = b/a$ als die eindeutige Lösung der Gleichung $a \cdot x = b$, falls $a, b \in \mathbb{R}$ mit $a \neq 0$ gegeben sind.

Addiert oder multipliziert man mehr als zwei Zahlen, so gilt gemäß Assoziativgesetz beispielsweise

$$(a + b) + c = a + (b + c)\,, \quad (ab)c = a(bc)\,.$$

Man braucht daher keine Klammern zu setzen. Entsprechend sind beliebige endliche Summen und Produkte

$$\sum_{k=m}^{n} x_k\,, \quad \prod_{k=m}^{n} x_k\,, \quad x_k \in \mathbb{R} \text{ für } m \leq k \leq n\,, \tag{2.19}$$

auch ohne das Setzen von Klammern wohldefiniert.

Das Kommutativgesetz $a + b = b + a$ für zwei Summanden hat zur Folge, dass bei der Summation endlich vieler Zahlen deren Reihenfolge beliebig vertauscht werden kann, also etwa

$$a + b + c = a + c + b = c + a + b = \dots$$

Dieser Sachverhalt findet im Falle einer **Doppelsumme** häufig Anwendung. Als Beispiel betrachten wir das folgende Tableau von Zahlen

$$\begin{pmatrix} 3 & 5 & 8 \\ 2 & 4 & 6 \end{pmatrix} . \tag{2.20}$$

Wir wollen die Summe aller dieser Zahlen bestimmen. Addieren wir zuerst zeilenweise und anschließend die Zeilensummen, so ergibt sich

$$(3 + 5 + 8) + (2 + 4 + 6) = 16 + 12 = 28 .$$

Dasselbe kommt natürlich heraus, wenn wir spaltenweise vorgehen,

$$(3 + 2) + (5 + 4) + (8 + 6) = 5 + 9 + 14 = 28 .$$

Der allgemeine Fall wird mit einer doppelt indizierten Variablen beschrieben. Wir betrachten $x_{ij} \in \mathbb{R}$, wobei i die Zahlen $1, \dots, m$ und j die Zahlen $1, \dots, n$ durchläuft,

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} . \tag{2.21}$$

Die Summe der i-ten Zeile ist gleich

$$\sum_{j=1}^{n} x_{ij} ,$$

zeilenweises Addieren entspricht der Doppelsumme

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} ,$$

spaltenweises Addieren der Doppelsumme

$$\sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij} .$$

Beide Doppelsummen sind gleich,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij} , \tag{2.22}$$

da die Addition kommutativ ist. Bewiesen wird (2.22) mit vollständiger Induktion.

Produkte behandelt man völlig analog, mit dem Ergebnis

$$\prod_{i=1}^{n}\prod_{j=1}^{m} x_{ij} = \prod_{j=1}^{m}\prod_{i=1}^{n} x_{ij} \,. \tag{2.23}$$

Das Distributivgesetz für endliche Summen lautet

$$\left(\sum_{i=1}^{n} x_i\right) \cdot \left(\sum_{j=1}^{m} y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j \,. \tag{2.24}$$

Auch diese Formel wird mit vollständiger Induktion bewiesen.

Ganzzahlige Potenzen definieren wir durch

$$x^0 = 1\,, \quad x^n = \underbrace{x \cdot x \cdots x}_{n \text{ mal}}, \quad \forall\, x \in \mathbb{R},\, n \in \mathbb{N}\,, \tag{2.25}$$

$$x^{-n} = (x^{-1})^n\,, \quad \forall\, x \in \mathbb{R} \text{ mit } x \neq 0,\, n \in \mathbb{N}\,. \tag{2.26}$$

Aus den Eigenschaften der Multiplikation folgt für alle $x, y \in \mathbb{R}$ und $n, m \in \mathbb{Z}$ (bei negativen Exponenten wird natürlich $x \neq 0$ bzw. $y \neq 0$ vorausgesetzt)

$$x^n x^m = x^{n+m}\,, \quad (x^n)^m = x^{nm}\,, \quad x^n y^n = (xy)^n\,. \tag{2.27}$$

Definition 2.2 (Binomialkoeffizient) *Für $k, n \in \mathbb{N}$ definieren wir den Binomialkoeffizienten*

$$\binom{n}{0} = 1\,, \quad \binom{n}{k} = \prod_{j=1}^{k} \frac{n-j+1}{j} = \frac{n(n-1)\cdots(n-k+1)}{1 \cdot 2 \cdots k}\,, \ \textit{falls } k \geq 1. \tag{2.28}$$

Direkt aus der Definition folgt

$$\binom{n}{k} = 0\,, \text{ falls } k > n\,, \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}, \quad 0 \leq k \leq n\,. \tag{2.29}$$

Lemma 2.3 *Für alle $n, k \in \mathbb{N}$ mit $1 \leq k \leq n$ gilt*

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}\,. \tag{2.30}$$

Beweis: Siehe Übungsaufgabe. □

Satz 2.4 (Binomische Formel) *Für alle $x, y \in \mathbb{R}$ und alle $n \in \mathbb{N}$ gilt*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\,. \tag{2.31}$$

Beweis: Mit vollständiger Induktion über n. Für $n=0$ gilt für alle $x,y \in \mathbb{R}$

$$\sum_{k=0}^{0} \binom{0}{k} x^k y^{0-k} = 1 = (x+y)^0 \,. \tag{2.32}$$

Die Behauptung sei richtig für ein gegebenes $n \in \mathbb{N}$. Wir zeigen ihre Richtigkeit für $n+1$. Seien $x,y \in \mathbb{R}$ beliebig, dann gilt

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

(und nun: Indexverschiebung von k nach $k-1$ im ersten Summenzeichen)

$$\begin{aligned}
&= \binom{n+1}{n+1} x^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] x^k y^{n-k+1} + \binom{n+1}{0} y^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k} \,.
\end{aligned}$$

□

Folgerungen aus den Ordnungseigenschaften. Wir hatten axiomatisch die Existenz einer Teilmenge P von $\mathbb{R}$ gefordert und gemäß (2.11) die Notation "$>$" eingeführt durch $x > y \;\Leftrightarrow\; x - y \in P$, also insbesondere

$$x > 0 \quad \Leftrightarrow \quad x \in P \,. \tag{2.33}$$

Wir erinnern an die Definitionen von "$\geq$", "$<$" und "$\leq$" in (2.12) – (2.14).

Es gilt offensichtlich $x \leq x$ für alle $x \in \mathbb{R}$. (Man sagt, "$\leq$" ist reflexiv.)

Die folgenden Eigenschaften sind anschaulich unmittelbar einleuchtend; wir führen sie auf die Axiome zurück.

Lemma 2.5 *Seien $x,y \in \mathbb{R}$. Dann gilt*

$$x > 0 \text{ und } y \geq 0 \quad \Rightarrow \quad x + y > 0 \,, \tag{2.34}$$

$$x \geq 0 \text{ und } y \geq 0 \quad \Rightarrow \quad x + y \geq 0 \,. \tag{2.35}$$

Beweis: Falls $y > 0$, folgt (2.34) direkt aus Eigenschaft (5), falls $y = 0$, so gilt $x + y = x > 0$. Falls $x = y = 0$, so ist auch $x + y = 0$. □

Lemma 2.6 *Seien $x, x', y, y' \in \mathbb{R}$. Dann gilt*

$$x < y \text{ und } x' \leq y' \quad \Rightarrow \quad x + x' < y + y' \,, \tag{2.36}$$

$$x \leq y \text{ und } x' \leq y' \quad \Rightarrow \quad x + x' \leq y + y' \,. \tag{2.37}$$

Beweis: Folgt aus Lemma 2.5, angewendet auf $y-x$ und $y'-x'$. $\square$

Lemma 2.7 *Seien $x,y \in \mathbb{R}$. Dann gilt (Antisymmetrie)*

$$x \le y \textit{ und } y \le x \quad \Rightarrow \quad x = y\,. \tag{2.38}$$

Beweis: Aus $x \neq y$ folgt $x-y \neq 0$, also $x-y \in P$ oder $y-x \in P$. Damit haben wir die Kontraposition von (2.38) bewiesen, nämlich

$$x \neq y \quad \Rightarrow \quad x < y \text{ oder } y < x\,. \tag{2.39}$$

$\square$

Lemma 2.8 *Seien $x,y,z \in \mathbb{R}$. Dann gilt*

$$x < y \textit{ und } y < z \quad \Rightarrow \quad x < z\,, \tag{2.40}$$

$$x \le y \textit{ und } y < z \quad \Rightarrow \quad x < z\,, \tag{2.41}$$

$$x < y \textit{ und } y \le z \quad \Rightarrow \quad x < z\,, \tag{2.42}$$

$$x \le y \textit{ und } y \le z \quad \Rightarrow \quad x \le z\,. \tag{2.43}$$

Die Eigenschaften (2.40) und (2.43) nennt man Transitivität.

Beweis: Aus $x<y$ und $y<z$ folgt $y-x \in P$ und $z-y \in P$, also auch

$$z - x = (z-y) + (y-x) \in P\,,$$

und damit $x < z$. Die anderen Aussagen ergeben sich aus Betrachtung der Fälle $x = y$ bzw. $y = z$. $\square$

Wir stellen nun eine Verbindung zwischen (Un-)gleichungen und Mengen her. Die Menge

$$\Delta = \{(x,y) : x,y \in \mathbb{R},\, x = y\} = \{(x,x) : x \in \mathbb{R}\} \tag{2.44}$$

stellt die Hauptdiagonale in der Ebene $\mathbb{R} \times \mathbb{R}$ dar, die Menge

$$A = \{(x,y) : x,y \in \mathbb{R},\, x < y\}$$

den Bereich oberhalb der Hauptdiagonale. Beides sind Teilmengen der Produktmenge $\mathbb{R} \times \mathbb{R}$. Für diese Situation gibt es einen allgemeinen Begriff.

Definition 2.9 *Sind M und N Mengen, so heißt jede Teilmenge R von $M \times N$ eine* ***Relation****. Statt $(x,y) \in R$ schreiben wir auch $x\,R\,y$.*

Beispiel: Die Gleichheitsbeziehung definiert für eine beliebigen Menge M (anstelle $\mathbb{R}$ wie in (2.44)) eine Relation R auf der Produktmenge $M \times M$ durch

$$x\,R\,y \quad \Leftrightarrow \quad (x,y) \in \Delta = \{(z,z) : z \in M\} \quad \Leftrightarrow \quad x = y\,.$$

Sind – wie in diesem Beispiel - M und N gleich, so spricht man auch von einer Relation auf M (anstatt auf $M \times M$).

Wir betrachten nun Ungleichheitsbeziehungen als Relationen, die bestimmte Zusatzeigenschaften haben.

Definition 2.10 *Sei M eine Menge. Eine Relation R auf M (also $R \subset M \times M$) heißt* ***Ordnungsrelation****, wenn gilt*

(i) *$x\,R\,x$ für alle $x \in M$ (**Reflexivität**),*

(ii) *aus $x\,R\,y$ und $y\,R\,x$ folgt $x = y$ (**Antisymmetrie**),*

(iii) *aus $x\,R\,y$ und $y\,R\,z$ folgt $x\,R\,z$ (**Transitivität**).*

Die oben bewiesenen Eigenschaften zeigen, dass "$\leq$" und "$\geq$" Ordnungsrelationen auf $\mathbb{R}$ definieren. Durch "$<$" und "$>$" werden keine Relationen definiert, da die Reflexivität nicht gilt.

Für "$\leq$" gilt außerdem:

$$\text{Sind } x, y \in \mathbb{R}, \text{ so gilt } x \leq y \text{ oder } y \leq x. \tag{2.45}$$

Diese Eigenschaft wird von einer Ordnungsrelation gemäß Definition 2.10 nicht verlangt, es kann also für gewisse $x, y \in M$ sehr wohl sein, dass weder $x\,R\,y$ noch $y\,R\,x$ gilt. Im folgenden Beispiel ist das so.

Beispiel 2.11 Auf $M = \mathbb{R} \times \mathbb{R}$ definieren wir eine Relation R durch

$$(x_1, x_2)\ R\ (y_1, y_2) \quad \Leftrightarrow \quad x_1 \leq y_1 \text{ und } x_2 \leq y_2\,.$$

Wir können R als Verallgemeinerung von "$\leq$" auf die Ebene auffassen. R ist eine Ordnungrelation, aber für die beiden Punkte $(1, 2)$ und $(2, 1)$ gilt weder $(1, 2)\,R\,(2, 1)$ noch $(2, 1)\,R\,(1, 2)$. Sie sind sozusagen "durch R nicht vergleichbar".

Wir behandeln weitere Eigenschaften der Ordnungsrelationen.

Lemma 2.12 *Seien $x, y \in \mathbb{R}$. Dann gilt*

(i) *Aus $x \geq 0$ und $y \geq 0$ folgt $xy \geq 0$.*

(ii) *Es ist $xy > 0$ genau dann, wenn*

$$(x > 0 \text{ und } y > 0) \text{ oder } (x < 0 \text{ und } y < 0).$$

Beweis: Zu (i): Sind $x > 0$ und $y > 0$, so folgt $xy > 0$ nach Eigenschaft (5). Ist $x = 0$ oder $y = 0$, so ist auch $xy = 0$. $\square$

Lemma 2.13 *Seien $x, y, z \in \mathbb{R}$. Dann gilt*

$$x < y \text{ und } z > 0 \quad \Rightarrow \quad xz < yz\,, \tag{2.46}$$

$$x \leq y \text{ und } z \geq 0 \quad \Rightarrow \quad xz \leq yz\,, \tag{2.47}$$

$$x < y \text{ und } z < 0 \quad \Rightarrow \quad xz > yz\,, \tag{2.48}$$

$$x \leq y \text{ und } z \leq 0 \quad \Rightarrow \quad xz \geq yz\,. \tag{2.49}$$

Beweis: Nur die erste Ungleichung. Aus $x < y$ folgt $0 < y-x$, also $0 < (y-x)z = yz-xz$, also $xz < yz$. □

Lemma 2.14 *Es gilt $x^2 \geq 0$ für alle $x \in \mathbb{R}$.*

Beweis: Aus Lemma 2.12 folgt: Ist $x \geq 0$, so ist $x^2 = x \cdot x \geq 0$; ist $x < 0$, so ist $-x > 0$ und $x^2 = (-x) \cdot (-x) > 0$. □

Folgerung 2.15 *Es ist $1 > 0$.*

Beweis: $1 = 1^2 \geq 0$, und aus $1 \neq 0$ folgt $1 > 0$. □

Lemma 2.16 *Für alle $x \in \mathbb{R}$ gilt: Aus $x > 0$ folgt $x^{-1} > 0$, aus $x < 0$ folgt $x^{-1} < 0$.*

Beweis: Sei $x > 0$. Dann ist $(x^{-1})^2 > 0$, also $x^{-1} = 1 \cdot x^{-1} = xx^{-1}x^{-1} = x \cdot (x^{-1})^2 > 0$. Der andere Fall wird analog bewiesen. □

Lemma 2.17 *Für alle $x, y \in \mathbb{R}$ gilt: Aus $0 < x < y$ folgt $x^{-1} > y^{-1}$.*

Beweis: Es ist $0 < xy$, also auch $0 < (xy)^{-1} = y^{-1}x^{-1}$. Es folgt weiter

$$y^{-1} = x(x^{-1}y^{-1}) < y(x^{-1}y^{-1}) = x^{-1}\,.$$

□

Lemma 2.18 *Für alle $x, y \in \mathbb{R}$ gilt: Aus $0 \leq x < y$ folgt $0 \leq x^n < y^n$ für alle $n \in \mathbb{N}$ mit $n \geq 1$.*

Beweis: Übungsaufgabe. □

Definition 2.19 (Betrag, Maximum und Minimum zweier Zahlen)
Ist $x \in \mathbb{R}$, so heißt

$$|x| = \begin{cases} x\,, & x \geq 0 \\ -x\,, & x < 0 \end{cases} \tag{2.50}$$

*der **Betrag** von x. Sind $x, y \in \mathbb{R}$, so definieren wir*

$$\max\{x,y\} = \begin{cases} x\,, & x \geq y\,, \\ y\,, & y > x\,, \end{cases} \qquad \min\{x,y\} = \begin{cases} x\,, & x \leq y\,, \\ y\,, & y < x\,. \end{cases} \tag{2.51}$$

Es gilt offensichtlich

$$\min\{x,y\} \leq x \leq \max\{x,y\}\,, \quad \min\{x,y\} \leq y \leq \max\{x,y\}\,, \quad \forall\, x, y \in \mathbb{R}\,, \tag{2.52}$$

$$|x| = \max\{x, -x\} = |-x|\,, \quad \forall\, x \in \mathbb{R}\,. \tag{2.53}$$

Das Maximum endlich vieler reller Zahlen wird per Induktion auf den Fall zweier Zahlen zurückgeführt,

$$\max\{x_1, \ldots, x_n\} = \max\{\max\{x_1, \ldots, x_{n-1}\}, x_n\}\,. \tag{2.54}$$

Analoges gilt für das Minimum. Insbesondere hat jede endliche Menge reeller Zahlen sowohl ein Maximum als auch ein Minimum.

Der Betrag $|x-y|$ der Differenz zweier Zahlen $x, y \in \mathbb{R}$ misst deren Abstand auf der Zahlengerade. So haben etwa die Zahlen 5 und -2 den Abstand $|5-(-2)| = |7| = 7$. Die Menge

$$M = \{x : x \in \mathbb{R}\,,\ |x-4| \le 2\}$$

ist die Menge aller Punkte, die vom Punkt 4 nicht weiter als 2 entfernt sind, das ist gerade das Intervall $[2,6]$ mit dem Mittelpunkt 4.

Lemma 2.20 *Für alle $x, y \in \mathbb{R}$ gilt*

$$|x| \ge 0\,, \quad |x| = 0 \quad \Leftrightarrow \quad x = 0\,, \tag{2.55}$$

$$|xy| = |x| \cdot |y|\,, \tag{2.56}$$

$$|x+y| \le |x| + |y|\,. \tag{2.57}$$

*Die Ungleichung (2.57) heißt **Dreiecksungleichung** [1].*

Beweis: Zu (2.55): Aus der Definition folgt $|x| \ge 0$ und $|0| = 0$. Ist $x \ne 0$, so ist $|x| = \max\{x, -x\} > 0$.
Zu (2.56): Man unterscheidet die verschiedenen Fälle im Vorzeichen von x und y. Ist etwa $x \ge 0$ und $y \le 0$, so ist $xy \le 0$ und

$$|xy| = -(xy) = x(-y) = |x| \cdot |y|\,,$$

analog für die anderen Fälle.
Zu (2.57): Es gilt

$$x + y \le |x| + |y|\,, \quad -(x+y) = (-x) + (-y) \le |x| + |y|\,,$$

also

$$|x+y| = \max\{x+y, -(x+y)\} \le |x| + |y|\,.$$

□

Aus (2.56) erhält man wegen

$$|x| = \left|\frac{x}{y} \cdot y\right| = \left|\frac{x}{y}\right| \cdot |y|$$

die Regel

$$\left|\frac{x}{y}\right| = \frac{|x|}{|y|}\,, \quad \text{für alle } x, y \in \mathbb{R} \text{ mit } y \ne 0\,. \tag{2.58}$$

Satz 2.21 (Bernoullische Ungleichung) *Es gilt*

$$(1+x)^n \ge 1 + nx \tag{2.59}$$

für alle $n \in \mathbb{N}$ und alle $x \in \mathbb{R}$ mit $x \ge -1$.

[1] Warum, werden wir sehen, wenn wir die komplexen Zahlen bzw. den mehrdimensionalen Raum behandeln.

Beweis: Sei $x \geq -1$ beliebig. Wir zeigen mit vollständiger Induktion, dass (2.59) für alle $n \in \mathbb{N}$ gilt. Für $n = 0$ klar. Induktionsschritt $n \to n+1$:

$$\begin{aligned}(1+x)^{n+1} &= (1+x)^n(1+x) \geq (1+nx)(1+x) \\ &\qquad \text{(richtig nach Induktionsvoraussetzung und wegen } x \geq -1) \\ &= 1 + nx + x + nx^2 \geq 1 + (n+1)x\,.\end{aligned}$$

□

Das Supremum. Ist M eine endliche Menge reeller Zahlen, so gibt es immer eine größte und eine kleinste Zahl in M. Sie heißen das **Maximum** bzw. das **Minimum** von M. Ist M eine unendliche Teilmenge von $\mathbb{R}$, so braucht es weder ein Maximum noch ein Minimum von M zu geben. So hat $\mathbb{Z}$ weder Maximum noch Minimum; $\mathbb{N}$ hat ein Minimum, nämlich 0, aber kein Maximum. Das überrascht nicht, da sich $\mathbb{N}$ und $\mathbb{Z}$ "ins Unendliche erstrecken". Bei der Menge

$$M = \left\{\frac{1}{n} : n \in \mathbb{N},\, n > 0\right\} \tag{2.60}$$

liegt eine subtilere Situation vor. Sie hat ein Maximum, nämlich 1, aber kein Minimum, denn zu jedem Element von M gibt es eines, das kleiner ist: Ist $1/n$ ein beliebiges Element von M, so ist

$$\frac{1}{m} \in M\,, \quad \frac{1}{m} < \frac{1}{n} \quad \text{für jedes } m \in \mathbb{N} \text{ mit } m > n.$$

Zu jedem Element von M gibt es also sogar unendlich viele kleinere Elemente von M. Man könnte nun meinen, das Minimum von M sei 0, aber dem ist nicht so, da $0 \notin M$ und das Minimum einer Menge selbst ein Element der Menge sein muss, nach der oben gegebenen Definition.

Definition 2.22 *Sei $M \subset \mathbb{R}$. Ein $x \in \mathbb{R}$ heißt*

- ***obere Schranke** für (oder: von) M, wenn $x \geq y$ für alle $y \in M$,*
- ***untere Schranke** für (oder: von) M, wenn $x \leq y$ für alle $y \in M$.*

*M heißt **nach oben (unten) beschränkt**, falls es eine obere (untere) Schranke für M gibt. M heißt **beschränkt**, falls M nach oben und nach unten beschränkt ist. M heißt **unbeschränkt**, falls M nicht beschränkt ist.*

Jede endliche Menge (das heißt, jede Menge, die nur endlich viele Elemente hat) ist nach oben durch ihr Maximum und nach unten durch ihr Minimum beschränkt.

$\mathbb{N}$ ist nach unten beschränkt, jede Zahl $x \leq 0$ (auch das Minimum 0) ist eine untere Schranke von $\mathbb{N}$. $\mathbb{N}$ ist nicht nach oben beschränkt (das ist anschaulich klar und wird weiter unten in Satz 2.29 bewiesen).

$\mathbb{Z}$ ist weder nach oben noch nach unten beschränkt; wäre x eine untere Schranke für $\mathbb{Z}$, so wäre $-x$ eine obere Schranke für $\mathbb{N}$.

Die Menge M aus (2.60) ist beschränkt. Jedes $x \geq 1$ ist eine obere Schranke, jedes $x \leq 0$ eine untere Schranke von M.

Lemma 2.23 *Eine Teilmenge M von $\mathbb{R}$ ist beschränkt genau dann, wennn es ein $C > 0$ gibt mit $|x| \leq C$ für alle $x \in M$.*

Beweis: Ist a eine untere und b eine obere Schranke für M, so hat $C = \max\{|a|, |b|\}$ die verlangte Eigenschaft. Umgekehrt ist $-C$ untere und C obere Schranke für M, falls $|x| \leq C$ für alle $x \in M$. □

Häufig vorkommende Teilmengen von $\mathbb{R}$ sind die Intervalle. Ein **Intervall** kann unterschiedliche Formen haben. Seien $a, b \in \mathbb{R}$ gegeben. Wir definieren

$$[a,b] = \{x : x \in \mathbb{R},\, a \leq x \leq b\}\,, \tag{2.61}$$
$$(a,b) = \{x : x \in \mathbb{R},\, a < x < b\}\,, \tag{2.62}$$
$$(a,b] = \{x : x \in \mathbb{R},\, a < x \leq b\}\,, \tag{2.63}$$
$$[a,b) = \{x : x \in \mathbb{R},\, a \leq x < b\}\,, \tag{2.64}$$
$$[a,\infty) = \{x : x \in \mathbb{R},\, a \leq x\}\,, \tag{2.65}$$
$$(a,\infty) = \{x : x \in \mathbb{R},\, a < x\}\,, \tag{2.66}$$
$$(-\infty,b] = \{x : x \in \mathbb{R},\, x \leq b\}\,, \tag{2.67}$$
$$(-\infty,b) = \{x : x \in \mathbb{R},\, x \leq b\}\,. \tag{2.68}$$

Intervalle der Form (2.61), (2.65), (2.67) heißen **abgeschlossen**, Intervalle der Form (2.62), (2.66), (2.68) heißen **offen**, die anderen heißen **halboffen**.

Definition 2.24 *Sei M eine Teilmenge von $\mathbb{R}$. Ein $z \in \mathbb{R}$ heißt*

- *Supremum von M, wenn z eine obere Schranke für M ist, und $z \leq x$ gilt für alle oberen Schranken x von M,*
- *Infimum von M, wenn z eine untere Schranke für M ist, und $z \geq x$ gilt für alle unteren Schranken x von M.*

Wir beschäftigen uns zunächst mit dem Supremum.

Lemma 2.25 *Sei $M \subset \mathbb{R}$ mit $M \neq \emptyset$. Dann hat M höchstens ein Supremum.*

Beweis: Sind z_1, z_2 Suprema von M, so gilt $z_1 \leq z_2$ (da z_2 obere Schranke von M) und $z_2 \leq z_1$ (da z_1 obere Schranke von M), also $z_1 = z_2$. □

Es macht daher Sinn, von "dem Supremum" von M zu sprechen, wir bezeichnen es mit

$$\sup M\,, \quad \text{oder} \quad \sup(M)\,. \tag{2.69}$$

Es gilt

$$\sup([1,2]) = 2 = \sup([1,2))\,,$$

da in beiden Fällen jedes $x \geq 2$ (und kein anderes) eine obere Schranke und 2 die kleinste solche Zahl ist. Das Maximum von $[1,2]$ ist ebenfalls 2, aber $[1,2)$ hat kein Maximum. Dieses Beispiel legt nahe, dass eine nach oben beschränkte Teilmenge ein Supremum "haben sollte". In der Tat, das ist das uns noch fehlende Axiom für die reellen Zahlen.

Axiom 2.26 (Supremumsaxiom)

$$\textit{Jede nichtleere nach oben beschränkte Teilmenge von } \mathbb{R} \textit{ besitzt ein Supremum.} \tag{2.70}$$

Diese Eigenschaft wird auch als die Ordnungsvollständigkeit von $\mathbb{R}$ bezeichnet.

Wir betrachten als Beispiel die Menge

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

Sie ist nichtleer, nach oben beschränkt (z.B. 2 ist obere Schranke), und hat daher ein eindeutig bestimmtes Supremum $\sup M \in \mathbb{R}$. (Diese reelle Zahl werden wir später als $\sqrt{2}$ identifizieren.)

Wir fassen zusammen:

$$\text{Das Supremum von } M \text{ ist die kleinste obere Schranke von } M,$$

falls M überhaupt eine obere Schranke hat.

Darüber hinaus vereinbaren wir

$$\sup M = +\infty\,, \quad \text{falls } M \text{ nicht nach oben beschränkt ist,} \tag{2.71}$$

$$\sup \emptyset = -\infty\,. \tag{2.72}$$

Damit haben wir $\sup M$ definiert für alle $M \subset \mathbb{R}$, und

$$\sup M \in \mathbb{R} \cup \{-\infty, +\infty\}\,. \tag{2.73}$$

Nebenbemerkung: Mit "$\cup$" wird die Mengenoperation der Vereinigung bezeichnet,

$$M \cup N = \{x : \; x \in M \text{ oder } x \in N\}\,.$$

Mengenoperationen werden im nächsten Kapitel ausführlicher besprochen.

Wir setzen die auf $\mathbb{R}$ bereits definierte Ordnungsrelation "$\le$" fort auf $\mathbb{R} \cup \{-\infty, +\infty\}$, indem wir setzen

$$-\infty < x < +\infty\,, \quad \text{für alle } x \in \mathbb{R}\,. \tag{2.74}$$

Man prüft leicht nach, dass wir auf diese Weise tatsächlich eine Ordnungsrelation "$\le$" auf $\mathbb{R} \cup \{-\infty, +\infty\}$ erhalten. Es gilt

Lemma 2.27 *Sind $M, N \subset \mathbb{R}$ mit $M \subset N$, so gilt*

$$\sup M \le \sup N\,. \tag{2.75}$$

Beweis: Falls $\sup N \in \mathbb{R}$, so ist $\sup N$ obere Schranke für N und wegen $M \subset N$ auch obere Schranke für M, also gilt (2.75) nach Definition von $\sup M$ (falls $M = \emptyset$, verwenden wir (2.72)). Falls $\sup N = -\infty$, so ist $N = \emptyset$, also auch $M = \emptyset$, also auch $\sup M = -\infty$. Falls $\sup N = +\infty$, ist (2.75) erfüllt, egal was M ist. □

Auch das Infimum einer Menge ist eindeutig bestimmt, falls es existiert, wie man analog zu Lemma 2.25 beweist. Wir bezeichnen es mit

$$\inf M\,, \tag{2.76}$$

und vereinbaren

$$\inf M = -\infty\,, \quad \text{falls } M \text{ nicht nach unten beschränkt ist,} \tag{2.77}$$

$$\inf \emptyset = +\infty\,. \tag{2.78}$$

Analog zu Lemma 2.27 zeigt man, dass für das Infimum gilt

$$\inf M \geq \inf N\,, \quad \text{falls } M \subset N.$$

Ist $M \subset \mathbb{R}$ und $a \in \mathbb{R}$, so definieren wir die Menge $aM \subset \mathbb{R}$ durch

$$aM = \{ax : x \in M\}\,. \tag{2.79}$$

Für $a = -1$ schreiben wir auch $-M$ statt $(-1)M$. Wir vereinbaren weiter

$$-(+\infty) = -\infty\,, \quad -(-\infty) = +\infty\,. \tag{2.80}$$

Satz 2.28 *Sei $M \subset \mathbb{R}$. Dann gilt*

$$\inf M = -\sup(-M)\,. \tag{2.81}$$

Beweis: Ist $M = \emptyset$, so steht auf beiden Seiten $+\infty$. Sei nun $M \neq \emptyset$. Wir behaupten, dass für $x \in \mathbb{R}$ gilt:

$$x \text{ ist untere Schranke von } M \quad \Leftrightarrow \quad -x \text{ ist obere Schranke von } -M\,. \tag{2.82}$$

In der Tat gilt

$$x \leq y \;\; \forall\, y \in M \quad \Leftrightarrow \quad -x \geq -y \;\; \forall\, y \in M \quad \Leftrightarrow \quad -x \geq z \;\; \forall\, z \in -M\,.$$

Wegen (2.82) ist M nach unten beschränkt genau dann wenn $-M$ nach oben beschränkt ist. Falls M nicht nach unten beschränkt ist, steht also $-\infty$ auf beiden Seiten von (2.81). Ist M nach unten beschränkt, so hat $-M$ nach Supremumsaxiom ein Supremum $\sup(-M) \in \mathbb{R}$, und wegen (2.82) ist $-\sup(-M)$ eine untere Schranke von M. Ist x eine beliebige untere Schranke von M, so gilt wegen (2.82) auch $\sup(-M) \leq -x$, also auch $-\sup(-M) \geq x$, und damit folgt (2.81) wegen der Eindeutigkeit des Infimums. □

Satz 2.29 $\mathbb{N}$ *ist nicht nach oben beschränkt, d.h. für alle $x \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$ mit $n > x$.*

Man sagt: "Der Körper $\mathbb{R}$ ist archimedisch angeordnet."

Beweis: Wir nehmen an, dass $\mathbb{N}$ nach oben beschränkt ist. Dann hat $\mathbb{N}$ nach Supremumsaxiom ein Supremum $s = \sup \mathbb{N} \in \mathbb{R}$. Wegen $s-1 < s$ ist $s-1$ keine obere Schranke von $\mathbb{N}$, also gibt es ein $n \in \mathbb{N}$ mit $s - 1 < n$. Es folgt weiter $s < n + 1 \in \mathbb{N}$, also ist auch s keine obere Schranke. Widerspruch. □

Folgerung 2.30 *Für alle $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit*

$$0 < \frac{1}{n} < \varepsilon\,. \tag{2.83}$$

Beweis: Wende Satz 2.29 an mit $x = \varepsilon^{-1}$. $\square$

Aus Folgerung 2.30 können wir schließen, dass

$$\inf \left\{ \frac{1}{n} : n \in \mathbb{N},\, n > 0 \right\} = 0\,.$$

Eine alternative Schreibweise ist

$$\inf_{n\in\mathbb{N}, n>0} \frac{1}{n} = 0\,, \quad \text{oder} \quad \inf_{\substack{n\in\mathbb{N}\\ n>0}} \frac{1}{n} = 0\,.$$

Folgerung 2.31 *Für alle $x, y \in \mathbb{R}$ mit $x > 0$ gibt es ein $n \in \mathbb{N}$ mit $nx > y$.*

Beweis: Falls $y > 0$, wähle $n \in \mathbb{N}$ mit $n > y/x$. Andernfalls $n = 1$. $\square$

Folgerung 2.32 *Für alle $x, y \in \mathbb{R}$ mit $x > 1$ gibt es ein $n \in \mathbb{N}$ mit $x^n > y$.*

Beweis: Übungsaufgabe. $\square$

Alternativ zum Supremumsaxiom kann man die Aussage von Satz 2.29 als Axiom nehmen, man nennt es dann das Archimedesaxiom. Allerdings benötigt man dann noch ein weiteres Axiom, um die Existenz der irrationalen Zahlen zu sichern, denn der Körper $\mathbb{Q}$ erfüllt ebenfalls das Archimedesaxiom.

Oft verwendet wird die folgende Eigenschaft des Supremums.

Satz 2.33 *Sei $M \subset \mathbb{R}$ nichtleer und nach oben beschränkt, sei $s = \sup M$. Dann gibt es zu jedem $\varepsilon > 0$ ein $x \in M$ mit*

$$s - \varepsilon < x \leq s\,. \tag{2.84}$$

Beweis: Übungsaufgabe. $\square$

3 Funktionen

Der Funktionsbegriff. Wir alle kennen die durch

$$f(x) = x^2 \tag{3.1}$$

definierte Funktion. Sie ordnet jeder gegebenen Zahl x die Zahl x^2 zu. Stellen wir sie graphisch in der Ebene dar, indem wir die Punkte $(x, f(x)) = (x, x^2)$ in ein kartesisches (= rechtwinkliges) Koordinatensystem eintragen, so erhalten wir eine Parabel.

Eine andere Funktion wird beispielsweise durch $f(x) = x^3$ definiert.

Eine **Funktion** im mathematischen Sinn hat mehrere Bestandteile, sie finden sich alle wieder in der Notation

$$f : M \to N\,. \tag{3.2}$$

Die Bestandteile sind folgende:

- der **Definitionsbereich** (oder **Definitionsgebiet**), eine Menge M, in der die Argumente der Funktion liegen,
- der **Wertebereich** (oder **Wertevorrat**), eine Menge N, in der die Werte der Funktion liegen,
- die **Funktionsvorschrift**, symbolisiert durch den Zuordnungspfeil "$\to$"; sie gibt an, welchem Argument welcher Wert zugeordnet wird,
- der frei wählbare Name (hier "f").

Statt "Funktion" sagt man auch **Abbildung**, das bedeutet dasselbe.

Eine Funktionsvorschrift (oder **Abbildungsvorschrift**) muss folgende Eigenschaft haben:

$$\text{Jedem } x \in M \text{ wird genau ein Element von } N \text{ zugeordnet.} \tag{3.3}$$

Das einem $x \in M$ eindeutig zugeordnete Element von N wird mit $\boldsymbol{f(x)}$ bezeichnet. Es heißt das **Bild** von x unter f, oder der Wert von f an der Stelle x.

Die Funktionsvorschrift alleine genügt nicht, es müssen auch Definitions- und Wertebereich spezifiziert werden. So kann man beispielsweise mit der Vorschrift $f(x) = x^2$ definieren

$$f : \mathbb{R} \to \mathbb{R}\,, \quad f : \mathbb{N} \to \mathbb{N}\,, \quad f : \mathbb{N} \to \mathbb{R}\,, \tag{3.4}$$

aber nicht $f : \mathbb{R} \to \mathbb{N}$, denn für $1/2 \in \mathbb{R}$ erzeugt die Vorschrft $f(x) = x^2$ den Wert $1/4$, der nicht in $\mathbb{N}$ liegt. Ebensowenig definiert

$$f(x) = \pm x^2$$

eine Funktion $f : \mathbb{R} \to \mathbb{R}$, solange man nicht festlegt, für welches x welches Vorzeichen genommen werden soll.

Manchmal verzichtet man darauf, einer Funktion einen Namen zu geben, man kann auch sagen:

$$\text{Wir definieren eine Funktion von } \mathbb{R} \text{ nach } \mathbb{R} \text{ durch die Vorschrift } x \mapsto x^2. \tag{3.5}$$

Man verwendet den Pfeil “$\mapsto$” für eine auf diese Weise beschriebene elementweise Zuordnung.

“Freie Wählbarkeit” des Namens bedeutet z.B., dass die Beschreibungen

$$\begin{aligned} f : \mathbb{R} \to \mathbb{R}\,, &\quad f(x) = x^2\,, \\ g : \mathbb{R} \to \mathbb{R}\,, &\quad g(x) = x^2\,, \end{aligned}$$

dieselbe Funktion definieren.

Zum Sprachgebrauch: Mit “f” bezeichnen wir die Funktion, mit “$f(x)$” den Wert der Funktion an der Stelle x. Innerhalb der Mathematik ist das der allgemein akzeptierte Standard. Die Sinusfunktion heißt dann “sin”, der Wert der Sinusfunktion an der Stelle x heißt “$\sin(x)$” oder “$\sin x$”. Der Vorteil dieser Sprachregelung ist, dass man zweifelsfrei weiß, ob die Funktion oder der Funktionswert gemeint ist. In anderen Fächern ist der Sprachgebrauch oft anders. Wenn dort von “$\sin x$” gesprochen wird, muss man u.U. dem Zusammenhang entnehmen, ob die Funktion oder der Funktionswert gemeint ist (oder ob der Sprecher da überhaupt einen Unterschied macht ...). Zwischenformen, die auch in der Mathematik vorkommen, können sein

> ”Die Funktion $(4x + 3)^2$ hat die Eigenschaft ...”

statt

> “Die durch $x \mapsto (4x + 3)^2$ definierte Funktion hat die Eigenschaft ...”.

Oder

> ”Die Funktion $f(x) = (4x + 3)^2$ hat die Eigenschaft ...”

statt

> ”Die durch $f(x) = (4x + 3)^2$ definierte Funktion f hat die Eigenschaft ...”

Fängt man an, Mathematik zu lernen, ist es aber besser, die präzisen Formulierungen zu verwenden, bis man sie sicher beherrscht.

Eine Funktionsvorschrift kann eine Fallunterscheidung beinhalten, beispielsweise

$$f : \mathbb{R} \to \mathbb{R}\,, \quad f(x) = \begin{cases} x^2\,, & x \geq 0\,, \\ -1\,, & x < 0\,. \end{cases}$$

Man muss auf die Abgrenzung der Fälle achten. So definiert

$$f(x) = \begin{cases} x^2\,, & x \geq 0\,, \\ -1\,, & x \leq 0\,, \end{cases}$$

keine Funktion $f : \mathbb{R} \to \mathbb{R}$, da $f(0)$ zweimal definiert wird ($f(0) = 0$ und $f(0) = -1$). Die Definition

$$f(x) = \begin{cases} x^2\,, & x \geq 0\,, \\ 0\,, & x \leq 0\,, \end{cases}$$

wäre aber in Ordnung, da $f(0)$ zwar zweimal vorkommt, aber beidesmal auf denselben Wert, nämlich 0, gesetzt wird.

Funktionsvorschriften können neben Formeln auch andere Ausdrücke enthalten. Die Vorschrift

$$f(n) = \text{Anzahl der Ziffern (im Dezimalsystem) des größten Primfaktors von } n$$

liefert eine wohldefinierte Funktion $f : \mathbb{N} \setminus \{0, 1\} \to \mathbb{N}$, aber den Wert $f(n)$ zu berechnen ist für große n nicht so einfach.

Die Vorschrift

$$f(x) = \begin{cases} 1\,, & \text{falls mehr als 10 Hörer dieser Vorlesung einen Hund besitzen,} \\ 0\,, & \text{andernfalls,} \end{cases}$$

definiert eine Funktion $f : \mathbb{R} \to \mathbb{R}$, die entweder konstant 1 oder konstant 0 ist. Ob wir wissen, welcher Fall zutrifft, ist dafür gleichgültig. Vorausgesetzt wird aber, dass für jeden Menschen y eindeutig feststeht, ob y Hörer der Vorlesung ist und ob y einen Hund besitzt.

Will man mit Mathematik die reale Welt beschreiben, so ergeben sich ständig Situationen, bei denen Funktionen betrachtet werden, die nicht durch Formeln definiert sind. Ein Beispiel:

> Wir betrachten ein Auto, dass zum Zeitpunkt $t = 0$ in Schwabing auf die A9 fährt und diese zum Zeitpunkt $t = T$ am Endpunkt am Berliner Ring verlässt. Sei $f(t)$ definiert als die Position des Autos zum Zeitpunkt t (in Autobahn-Kilometern gemessen). Dadurch wird eine Funktion $f : [0, T] \to \mathbb{R}$ definiert. Sie ist nicht durch eine Formel gegeben, ihre Werte (und auch der Wert von T) sind normalerweise nicht bekannt. Man kann nun Annahmen machen und daraus Eigenschaften von f und T herleiten, z.B. aus Annahmen über die Höchstgeschwindigkeit des Autos eine untere Schranke für T berechnen. Oder (wie Google Maps) auf der Basis eines mathematischen Modells und realer Daten die erwartete Fahrzeit T schätzen. Oder ...

Die uns für Funktionen $f : \mathbb{R} \to \mathbb{R}$ bekannte graphische Darstellung in der Zahlenebene hat eine allgemeine Entsprechung. Ist $f : M \to N$ eine Abbildung, so ist der **Graph** von f definiert als die Menge

$$\{(x, f(x)) : x \in M\}\,. \tag{3.6}$$

Sie ist eine Teilmenge der Produktmenge $M \times N$, ihre Elemente sind geordnete Paare (zuerst x, dann $f(x)$).

Gemäß Definition 2.9 ist der Graph einer Funktion also eine Relation. Wir stellen nun die umgekehrte Frage: Welche Relationen können wir als Graphen einer Funktion erhalten?

Sei $R \subset M \times N$ eine Relation, wir suchen eine Funktion $f : M \to N$ mit

$$R = \{(x, f(x)) : x \in M\}\,. \tag{3.7}$$

Dabei müssen wir die Bedingung (3.3) beachten, die wir an jede Funktionsvorschrift stellen. Falls R zwei Elemente (x, y_1) und (x, y_2) enthält mit $y_1 \neq y_2$, so ist "$\subset$" in (3.7) nicht erfüllbar, da es zu gegebenem $x \in M$ nur einen Funktionswert $f(x)$ geben darf. Falls es ein $x \in M$ gibt für welches kein $y \in N$ existiert mit $(x, y) \in R$, so gilt $(x, f(x)) \notin R$ für jede Funktion $f : M \to N$, also ist "$\supset$" in (3.7) nicht erfüllbar. Die Relation R muss also die Eigenschaft

$$\text{zu jedem } x \in M \text{ gibt es genau ein } y \in N \text{ mit } x\,R\,y \tag{3.8}$$

haben. Hat R diese Eigenschaft, so können wir $f : M \to N$ definieren, indem wir jedem $x \in M$ das entsprechende y aus (3.8) zuordnen.

Mit der soeben hergestellten Beziehung kann man den Funktionsbegriff auf den Mengenbegriff zurückführen. "Eine Funktion von M nach N ist eine Teilmenge R der Produktmenge $M \times N$, welche die Eigenschaft (3.8) hat." Die Frage "was ist denn eigentlich eine Funktionsvorschrift?" erledigt sich dann mit der Antwort: Wir ordnen jedem $x \in M$ das entsprechende $y \in N$ aus (3.8) zu.

Eigenschaften von Abbildungen. Sei $f : M \to N$ eine Abbildung. Ist A eine Teilmenge von M, so definieren wir eine Menge $f(A)$ durch

$$f(A) = \{f(x) : x \in A\}\,. \tag{3.9}$$

Sie heißt die **Bildmenge** (oder einfach **Bild**) von A unter f, es gilt $f(A) \subset f(M) \subset N$. Ist $y \in N$, so heißt jedes $x \in M$ mit $f(x) = y$ ein **Urbild** von y.

Zu gegebenem y kann es ein, mehrere oder auch kein Urbild geben. Im Beispiel $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, hat 0 ein Urbild, nämlich 0, jedes $y < 0$ hat kein Urbild. Jedes $y > 0$ hat zwei Urbilder, da mit $x^2 = y$ auch $(-x)^2 = y$ gilt.

Ist $f : M \to N$ und $B \subset N$, so heißt

$$f^{-1}(B) = \{x : x \in M,\ f(x) \in B\} \tag{3.10}$$

die **Urbildmenge** (oder das **Urbild**) von B unter f. (Zum Begriff der Umkehrfunktion kommen wir später.) Im Beispiel $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ gilt

$$f^{-1}([0, 4]) = [-2, 2]\,.$$

Anschaulich ist das klar, ein entsprechender mathematischer Satz wird später bewiesen.

Zwei Abbildungen $f, g : M \to N$ heißen gleich, geschrieben $f = g$, falls

$$f(x) = g(x) \quad \text{für alle } x \in M. \tag{3.11}$$

Als Beispiel betrachten wir $f(x) = x$, $g(x) = |x|$. Falls wir als Definitionsbereich $\mathbb{R}_+$ nehmen, also $f, g : \mathbb{R}_+ \to \mathbb{R}$ betrachten, so gilt $f = g$. Falls wir als Definitionsbereich $\mathbb{R}$ nehmen, gilt $f \neq g$. Mit "$f \neq g$" ist die Negation von (3.11) gemeint, das heißt,

$$\text{es gibt ein } x \in M \text{ mit } f(x) \neq g(x).$$

Beispielsweise gilt $f \neq g$ für die durch $f(x) = x$, $g(x) = -x$ definierten Funktionen $f, g : \mathbb{R} \to \mathbb{R}$, obwohl $f(0) = g(0)$ gilt.

Eine Abbildung $f : M \to N$ heißt

- **surjektiv**, wenn jedes $y \in N$ mindestens ein Urbild hat. Äquivalent dazu ist: $f(M) = N$.
- **injektiv**, wenn jedes $y \in N$ höchstens ein Urbild hat. Äquivalent dazu ist: Sind $x_1, x_2 \in M$ beliebig mit $x_1 \neq x_2$, so muss auch $f(x_1) \neq f(x_2)$ gelten. Ebenfalls äquivalent ist: Sind $x_1, x_2 \in M$ beliebig mit $f(x_1) = f(x_2)$, so muß $x_1 = x_2$ gelten.
- **bijektiv**, wenn sie surjektiv und injektiv ist, das heißt, wenn jedes $y \in N$ genau ein Urbild hat.

Die folgenden Beispiele zeigen, dass Injektivität und Surjektivität einer Abbildung nicht nur von der Funktionsvorschrift, sondern auch von der Wahl von Definitions- und Wertebereich abhängen.

- Wir betrachten die Funktionsvorschrift $f(x) = x + 1$, $f : M \to N$, mit unterschiedlichen Teilmengen M und N von $\mathbb{R}$. In jedem Fall ist f injektiv: Ist $f(x_1) = f(x_2)$, so ist $x_1 + 1 = x_2 + 1$, also auch $x_1 = x_2$. Im Fall $f : \mathbb{N} \to \mathbb{N}$ ist f nicht surjektiv, da 0 kein Urbild in $\mathbb{N}$ hat. Im Fall $f : \mathbb{Z} \to \mathbb{Z}$ ist f surjektiv (also auch bijektiv), da jedes $x \in \mathbb{Z}$ das Urbild $x - 1$ hat. Aus dem gleichen Grund ist $f : \mathbb{R} \to \mathbb{R}$ ebenfalls surjektiv und damit bijektiv.
- Wir betrachten die Funktionsvorschrift $f(x) = x^2$, $f : M \to N$, mit unterschiedlichen Teilmengen M und N von $\mathbb{R}$. Im Fall $f : \mathbb{N} \to \mathbb{N}$ ist f injektiv (aus $x_1 \neq x_2$ folgt $x_1^2 \neq x_2^2$), aber nicht surjektiv (zum Bild von $\mathbb{N}$ gehören nur die Quadratzahlen). In den Fällen $f : \mathbb{Z} \to \mathbb{Z}$ und $f : \mathbb{R} \to \mathbb{R}$ ist f weder injektiv (da $f(-x) = f(x)$) noch surjektiv (da keine negative Zahl im Bild von f liegt). Im Fall $f : \mathbb{R}_+ \to \mathbb{R}_+$ ist f injektiv und surjektiv (letzteres folgt wieder aus einem später bewiesenen Satz), also auch bijektiv.

Die Umkehrfunktion. Wir betrachten nochmals die durch

$$f(x) = x + 1 \tag{3.12}$$

definierte Abbildung $f : \mathbb{R} \to \mathbb{R}$. Wir wollen sie umkehren, das heißt, wir wollen mit einem Wert $f(x)$ beginnen und daraus das ursprüngliche x zurückgewinnen. Für (3.12) lässt sich das direkt ausrechnen. Wir beginnen mit einem beliebigen $y \in \mathbb{R}$ und suchen $x \in \mathbb{R}$ mit $y = f(x)$, also

$$y = f(x) = x + 1\,.$$

Wir sehen,

$$x = y - 1$$

ist das eindeutig bestimmte x, welches $f(x) = y$ erfüllt. Die durch

$$g(y) = y - 1 \tag{3.13}$$

definierte Funktion $g : \mathbb{R} \to \mathbb{R}$ kehrt also f um. Warum hat diese (zugegebenermaßen sehr einfache) Rechnung funktioniert? Weil f bijektiv ist; dadurch wird garantiert, dass jedes $y \in N$ genau ein Urbild $x \in M$ hat.

Allgemein gilt: Ist $f : M \to N$ eine bijektive Abbildung von einer Menge M auf eine Menge N, so können wir durch die Vorschrift

$$y \quad \mapsto \quad \text{“das eindeutig bestimmte } x \in M \text{ mit } f(x) = y\text{”} \tag{3.14}$$

eine Abbildung definieren, sie heißt die **Umkehrfunktion** zu f und wird mit f^{-1} bezeichnet. Es ist dann

$$f^{-1} : N \to M\,.$$

Für Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ erhalten wir den Graphen der Umkehrfunktion f^{-1}, indem wir den Graphen von f an der Diagonale “$x = y$” spiegeln.

Im Beispiel (3.12) war es möglich, die Umkehrfunktion durch die Formel (3.13) “explizit” anzugeben. Das muss nicht sein, und geht oft nicht.

Betrachten wir $f(x) = x^2$. Als Abbildung $f : \mathbb{R} \to \mathbb{R}$ ist sie nicht bijektiv, daher hat sie keine Umkehrabbildung von $\mathbb{R}$ nach $\mathbb{R}$. Aber als Abbildung $f : \mathbb{R}_+ \to \mathbb{R}_+$ sollte sie bijektiv sein, die Umkehrabbildung $f^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$ sollte gegeben sein durch

$$f^{-1}(y) = \sqrt{y}\,.$$

Gemeint ist die nichtnegative Wurzel.

Allgemein erhalten wir die k-te Wurzel als Umkehrfunktion der k-ten Potenz

$$f(x) = x^k\,, \quad f : \mathbb{R}_+ \to \mathbb{R}_+\,, \tag{3.15}$$

durch den folgenden Satz.

Satz 3.1 *Sei $k \in \mathbb{N}$ mit $k \geq 1$. Dann gibt es zu jedem $y \in \mathbb{R}_+$ genau ein $x \in \mathbb{R}_+$ mit $x^k = y$.*

Beweis: Sei $y \in \mathbb{R}_+$. Es gibt höchstens ein x mit $x^k = y$ wegen Lemma 2.18. Das gesuchte x wollen wir als Supremum erhalten,

$$x = \sup(M)\,, \quad M = \{t : t \geq 0\,,\, t^k \leq y\}\,.$$

Das Supremum von M existiert nach Supremumsaxiom: M ist nichtleer (da $0 \in M$) und nach oben beschränkt (da $1 + y$ eine obere Schranke für M ist). Wir zeigen nun, dass weder $x^k < y$ noch $x^k > y$ richtig sein kann. Zu diesem Zweck betrachten wir ein $\varepsilon > 0$ und schätzen ab mit Hilfe der binomischen Formel

$$(x+\varepsilon)^k = \sum_{j=0}^{k} \binom{k}{j} \varepsilon^j x^{k-j} = x^k + \varepsilon \sum_{j=1}^{k} \binom{k}{j} \varepsilon^{j-1} x^{k-j} \leq x^k + \varepsilon \sum_{j=1}^{k} \binom{k}{j} \varepsilon^{j-1} |x|^{k-j}\,.$$

Ist außerdem $\varepsilon < 1$, so gilt $\varepsilon^{j-1} \leq 1$ und daher

$$(x+\varepsilon)^k \leq x^k + \varepsilon C\,, \quad C = \sum_{j=1}^{k} \binom{k}{j} |x|^{k-j}\,.$$

Wäre nun $x^k < y$, so folgt

$$(x+\varepsilon)^k \leq x^k + \varepsilon C < y\,,$$

falls wir ε so klein wählen, dass $\varepsilon C < y - x^k$. Damit folgt, dass $x + \varepsilon \in M$ und daher x keine obere Schranke ist für M, ein Widerspruch. Eine analoge Rechnung für $0 < \varepsilon < 1$ zeigt, dass

$$(x - \varepsilon)^k \geq x^k - \varepsilon C\,, \quad C = \sum_{j=1}^{k} \binom{k}{j} |x|^{k-j}\,.$$

Wäre $x^k > y$, so wäre $x - \varepsilon$ eine obere Schranke von M, falls ε hinreichend klein ist, im Widerspruch dazu, dass x die kleinste obere Schranke für M ist. Also gilt insgesamt $x^k = y$. □

Es ergibt sich, dass

$$f(x) = x^k\,, \quad f : \mathbb{R}_+ \to \mathbb{R}_+\,,$$

bijektiv ist. Wir definieren

$$\sqrt[k]{y} = f^{-1}(y)\,, \quad f^{-1} : \mathbb{R}_+ \to \mathbb{R}_+\,. \tag{3.16}$$

Der Beweis von Satz 3.1 war relativ aufwendig. Später werden wir sehen, dass die Lage sich erheblich übersichtlicher und einfacher gestaltet, wenn wir über eine allgemeinere Begrifflichkeit verfügen; es ist dann nicht mehr erforderlich, eine explizite Rechnung mit der binomischen Formel durchzuführen.

Satz 3.1 ist ein Beispiel für einen sogenannten Existenz- und Eindeutigkeitssatz. Wenn man die k-te Wurzel einer gegebenen Zahl tatsächlich berechnen will, hilft er allerdings nur wenig.

Da die k-te Wurzel die Umkehrfunktion der k-ten Potenz ist, gilt für $x \geq 0$

$$(\sqrt[k]{x})^k = x\,.$$

Es folgt für $x, y \geq 0$

$$(\sqrt[k]{x}\sqrt[k]{y})^k = (\sqrt[k]{x})^k \cdot (\sqrt[k]{y})^k = xy\,,$$

also gilt für alle $x, y \geq 0$ die Rechenregel

$$\sqrt[k]{xy} = \sqrt[k]{x}\sqrt[k]{y}\,. \tag{3.17}$$

Komposition von Funktionen. Seien f und g definiert durch

$$f(x) = x + 1\,, \quad g(x) = x^2\,. \tag{3.18}$$

Wir wollen f und g "hintereinander ausführen". Sei etwa $x = 2$. Wir setzen 2 in f ein und erhalten $f(2) = 3$. Diesen Wert setzen wir in g ein und erhalten $g(f(2)) = g(3) = 9$. Das Argument von g (ursprünglich mit "x" bezeichnet) wird ersetzt durch die Zahl, die sich durch Auswertung von f ergeben hat.

Dasselbe passiert, wenn wir x als Variable belassen und nicht durch eine konkrete Zahl ersetzen,

$$x \quad \stackrel{f}{\mapsto} \quad x + 1 \quad \stackrel{g}{\mapsto} \quad (x + 1)^2\,.$$

Das Argument von g wird wieder ersetzt durch den Ausdruck, der sich durch Auswertung von f ergeben hat. Die resultierende Funktion ist

$$x \quad \mapsto \quad (x + 1)^2\,. \tag{3.19}$$

Hier kann die Tatsache Verwirrung stiften, dass wir in (3.18) das Argument von f und g beidesmal mit x bezeichnet haben; betrachtet man f und g unabhängig voneinander, so hat man zunächst keinen Grund, verschiedene Buchstaben für die beiden Argumente zu verwenden. Hätten wir statt (3.18) geschrieben

$$f(x) = x + 1\,, \quad g(y) = y^2\,,$$

so hätten wir y durch $x + 1$ ersetzt, mit demselben Ergebnis wie in (3.19).

Die allgemeine Situation ist die folgende. Seien

$$f : M \to N\,, \quad g : N \to P\,, \tag{3.20}$$

Abbildungen zwischen Mengen M, N und P. Die **Komposition** von f und g ist eine Abbildung von M nach P, sie wird bezeichnet mit $g \circ f$ und definiert durch

$$g \circ f : M \to P\,, \qquad (g \circ f)(x) = g(f(x)) \quad \text{für alle } x \in M. \tag{3.21}$$

Zuerst wird f ausgeführt, dann g (im Kontrast dazu, dass beim Lesen von links nach rechts man zuerst den Buchstaben g und dann den Buchstaben f liest).

In Beispiel (3.18) hatten wir stillschweigend $M = N = P = \mathbb{R}$ angenommen; im allgemeinen Fall muss man allerdings auf Definitions- und Wertebereiche achten, denn man kann den Wert $f(x)$ nur dann in g einsetzen, wenn er im Definitionsbereich von g liegt. Das ist in der Situation (3.20) garantiert; um sie herzustellen, muss man u.U. den Definitionsbereich von f geeignet wählen. Sei beispielsweise

$$f(x) = x - 4\,, \quad g(x) = \frac{1}{x^2 - 9}\,.$$

Betrachtet man f für sich genommen, so ist $f : \mathbb{R} \to \mathbb{R}$ eine natürliche Wahl; will man aber die Komposition $g \circ f$ betrachten, so muss man den Definitionsbereich von f so einschränken, dass die Werte $f(x) = \pm 3$ nicht auftreten. Ein natürlicher Kandidat für die Menge M in (3.20) ist $\{x : x \in \mathbb{R}, x \neq 7, x \neq 1\}$ oder eine Teilmenge davon.

Die Reihenfolge bei der Komposition ist wesentlich. Für $f(x) = x + 1$, $g(x) = x^2$ mit $f, g : \mathbb{R} \to \mathbb{R}$ sind sowohl $g \circ f : \mathbb{R} \to \mathbb{R}$ als auch $f \circ g : \mathbb{R} \to \mathbb{R}$ definiert, aber es gilt einerseits, wie wir gesehen haben, $(g \circ f)(x) = (x+1)^2$, und andererseits $(f \circ g)(x) = x^2 + 1$, also $g \circ f \neq f \circ g$.

Sei $f : M \to N$ bijektiv und $f^{-1} : N \to M$ die Umkehrabbildung. Es gilt

$$\begin{aligned} f^{-1}(f(x)) &= x \quad \text{für alle } x \in M, \\ f(f^{-1}(y)) &= y \quad \text{für alle } y \in N. \end{aligned}$$

Das können wir auch schreiben als

$$f^{-1} \circ f = \mathrm{id}_M\,, \quad f \circ f^{-1} = \mathrm{id}_N\,, \tag{3.22}$$

wobei "id" die **identische Abbildung** bezeichnet,

$$\mathrm{id}_M : M \to M\,, \quad \mathrm{id}_M(x) = x \text{ für alle } x \in M.$$

Mengenoperationen. Sind M und N Mengen, so definieren wir

$$M \cup N = \{x : x \in M \text{ oder } x \in N\} \qquad \textbf{(Vereinigung)}\,,$$
$$M \cap N = \{x : x \in M \text{ und } x \in N\} \qquad \textbf{(Durchschnitt)}\,.$$

Falls die Mengen M und N kein gemeinsames Element haben, gilt

$$M \cap N = \emptyset\,.$$

In diesem Fall heißen M und N **disjunkt**.
Beispiele:

$$[0,2] \cup (1,3) = [0,3)\,, \quad [0,2] \cap (1,3) = (1,2]\,, \quad \mathbb{N} \cup \mathbb{Z} = \mathbb{Z}\,, \quad \mathbb{N} \cap \mathbb{Z} = \mathbb{N}\,.$$

Das **Komplement** einer Menge N in einer Menge M ist definiert als

$$M \setminus N = \{x : x \in M \text{ und } x \notin N\}\,.$$

Beispiele:

$$[0,4] \setminus [1,8] = [0,1)\,, \quad [0,4] \setminus [1,2] = [0,1) \cup (2,4]\,, \quad \mathbb{Z} \setminus \mathbb{N} = \{-1,-2,-3,\dots\}\,.$$

Rechenregeln für Mengenoperationen. Seien M, N, P Mengen. Dann gilt: Aus $M \subset N$ und $N \subset P$ folgt $M \subset P$. (Beweis: Ist $x \in M$, so ist $x \in N$ wegen $M \subset N$ und weiter $x \in P$ wegen $N \subset P$.)
Für die Vereinigung gilt das Assoziativgesetz

$$(M \cup N) \cup P = M \cup (N \cup P)\,. \tag{3.23}$$

Beweis: Es gilt

$$\begin{aligned} x \in (M \cup N) \cup P &\Leftrightarrow x \in M \cup N \text{ oder } x \in P \\ &\Leftrightarrow (x \in M \text{ oder } x \in N) \text{ oder } x \in P \\ &\Leftrightarrow x \in M \text{ oder } (x \in N \text{ oder } x \in P) \\ &\Leftrightarrow x \in M \text{ oder } x \in N \cup P \\ &\Leftrightarrow x \in M \cup (N \cup P)\,. \end{aligned}$$

Die zweite und dritte Zeile sind äquivalent, da für die logische Verknüpfung "oder" das Assoziativgesetz gilt. (Wenn man das beweisen will, vergleicht man die Wahrheitstafeln beider Aussagen.) Analog zeigt man das Assoziativgesetz für den Durchschnitt

$$(M \cap N) \cap P = M \cap (N \cap P)\,.$$

Man kann daher die Klammern weglassen und einfach

$$M \cup N \cup P\,, \quad M \cap N \cap P$$

schreiben. Weiter gelten die Kommutativgesetze

$$M \cup N = N \cup M\,, \quad M \cap N = N \cap M\,,$$

und die Distributivgesetze

$$M \cup (N \cap P) = (M \cup N) \cap (M \cup P)\,, \quad M \cap (N \cup P) = (M \cap N) \cup (M \cap P)\,.$$

Für Bild- und Urbildmengen gelten ebenfalls eine Reihe von Rechenregeln. Sei $f : M \to N$ eine Abbildung. Sind $A, B \subset M$, so gilt

$$f(A \cup B) = f(A) \cup f(B)\,, \quad f(A \cap B) \subset f(A) \cap f(B)\,.$$

Wir beweisen die zweite der beiden Regeln. Aus $A \cap B \subset A$ folgt $f(A \cap B) \subset f(A)$. Ebenso erhält man $f(A \cap B) \subset f(B)$. Beide Aussagen zusammengenommen ergeben $f(A \cap B) \subset f(A) \cap f(B)$. Die umgekehrte Inklusion "$\supset$" braucht aber nicht zu gelten. Ist etwa $f(x) = x^2$, $f : \mathbb{R} \to \mathbb{R}$, so ist $f([1,2]) = f([-1,-2]) = [1,4]$, und

$$f([1,2] \cap [-1,-2]) = f(\emptyset) = \emptyset\,, \quad f([1,2]) \cap f([-1,-2]) = [1,4]\,.$$

Sind $C, D \subset N$, so gilt

$$f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)\,, \quad f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)\,.$$

Indizierte Mengen. Wie bei Zahlen kann es auch bei Mengen sinnvoll sein, Indizes zu verwenden. Als Beispiel betrachten wir die Intervalle

$$M_k = [0, \frac{1}{k}]\,, \quad k \in \mathbb{N}\,, \, k \geq 1\,.$$

Für sie gilt

$$M_{k+1} = [0, \frac{1}{k+1}] \subset [0, \frac{1}{k}] = M_k\,,$$

sie sind also ineinander geschachtelt,

$$M_n \subset M_{n-1} \subset \cdots \subset M_2 \subset M_1 = [0,1]\,.$$

Wir bilden den Durchschnitt dieser Mengen mit einer Notation analog zur Summennotation. Wegen deren Schachtelung ergibt sich

$$\bigcap_{k=1}^{n} M_k = M_n = [0, \frac{1}{n}]\,.$$

Wir können auch den Durchschnitt aller solcher Mengen bilden. Er ist definiert als

$$\bigcap_{\substack{k \in \mathbb{N} \\ k > 0}} M_k = \{x : x \in M_k \text{ für alle } k \geq 1\}\,.$$

Das Ergebnis ist

$$\bigcap_{\substack{k \in \mathbb{N} \\ k > 0}} = \{0\}\,,$$

da 0 zu jedem M_k gehört, aber $x \notin M_k$ falls $x > 0$ und $k > 1/x$.

Die Indizes müssen keine natürlichen Zahlen sein. Zu $h \in \mathbb{R}$, $h > 0$ definieren wir beispielsweise

$$M_h = [0, 1+h] \,.$$

Es gilt

$$\bigcap_{\substack{h\in\mathbb{R}\\ h>0}} M_h = \{x : x \in M_h \text{ für alle } h > 0\} = [0,1] \,.$$

Die allgemeine Situation ist die folgende. Sei I eine beliebige Menge. Zu jedem $i \in I$ sei eine Menge M_i gegeben. In diesem Kontext heißt I eine **Indexmenge**. Die "Gesamtheit" der Mengen M_i nennen wir eine **Mengenfamilie** und bezeichnen sie mit $(M_i)_{i\in I}$ oder mit $\{M_i\}_{i\in I}$. Ist beispielsweise $I = \{h : h \in \mathbb{R},\, h > 2\}$ und $M_h = [2, h]$, so ist $(M_h)_{h\in I}$ die Familie aller abgeschlossener beschränkter Intervalle, deren linker Randpunkt gleich 2 ist.

Ist $(M_i)_{i\in I}$ eine Familie von Mengen, so definieren wir

$$\bigcup_{i\in I} M_i = \{x : \text{ es gibt ein } i \in I \text{ mit } x \in M_i\} \,, \tag{3.24}$$

$$\bigcap_{i\in I} M_i = \{x : \text{ für alle } i \in I \text{ gilt } x \in M_i\} \,. \tag{3.25}$$

Ein weiteres Beispiel: Sei $I = \mathbb{N} \setminus \{0\}$, sei für $n \in I$

$$M_n = \{\frac{p}{n} : p \in \mathbb{Z}\} = \{\dots, -\frac{3}{n}, -\frac{2}{n}, -\frac{1}{n}, 0, \frac{1}{n}, \frac{2}{n}, \frac{3}{n}, \dots\} \,.$$

Dann ist

$$\bigcup_{n\in\mathbb{N}} M_n = \mathbb{Q} \,.$$

Für solche Vereinigungen und Durchschnitte nehmen die Distributivgesetze die Form an

$$N \cup \bigcap_{i\in I} M_i = \bigcap_{i\in I} (N \cup M_i) \,, \quad N \cap \bigcup_{i\in I} M_i = \bigcup_{i\in I} (N \cap M_i) \,.$$

Eine Vereinigung $\cup_{i\in I} M_i$ heißt disjunkt, wenn alle beteiligten Mengen paarweise disjunkt sind, das heißt, wenn $M_i \cap M_j = \emptyset$ für alle $i, j \in I$ mit $i \neq j$. Es genügt nicht, dass

$$\bigcap_{i\in I} M_i = \emptyset \,.$$

Beispielsweise ist die Vereinigung der Intervalle $[k, k+2]$ mit $k \in \mathbb{N}$ keine disjunkte Vereinigung, da $[0,2] \cap [1,3] = [1,2] \neq \emptyset$; in diesem Fall gilt

$$\bigcap_{k\in\mathbb{N}} [k, k+2] = \emptyset \,.$$

Mengen von Mengen. Die Elemente einer Menge können ohne weiteres selbst wieder Mengen sein. Beispielsweise hat die mengentheoretische Interpretation des geordneten Paares (x, y) auf die Menge

$$\{\{x\}, \{x, y\}\}$$

geführt, eine Menge mit zwei Elementen, eins davon eine einelementige Menge, das andere eine zweielementige Menge. Ein weiteres Beispiel ist die Menge

$$\{[a,b] : a,b \in \mathbb{R}\,,\ a < b\}$$

aller abgeschlossenen beschränkten Intervalle.

Ist M eine Menge, so heißt die Menge $\mathcal{P}(M)$ aller Teilmengen von M,

$$\mathcal{P}(M) = \{N : N \subset M\} \tag{3.26}$$

die Potenzmenge von M. Wir verabreden, dass $\emptyset \subset M$ gilt für jede Menge M, also $\emptyset \in \mathcal{P}(M)$. (Die Negation von

$$\forall x \in \emptyset \text{ gilt } x \in M$$

liefert die Aussage

$$\exists x \in \emptyset \text{ mit } x \notin M \quad ,$$

welche wir als falsch ansehen wollen.) Die Potenzmengenbildung lässt sich wiederholen: So sind auch

$$\mathcal{P}(\mathcal{P}(M))\,, \quad \mathcal{P}(\mathcal{P}(\mathcal{P}(M)))\,, \quad \ldots$$

Mengen. Beispiel: Ist $[a,b]$ ein beschränktes Intervall, so ist

$$[a,b] \subset \mathbb{R}\,, \quad \text{also} \quad [a,b] \in \mathcal{P}(\mathbb{R})\,.$$

Für die Menge aller solcher Intervalle gilt

$$\{[a,b] : a,b \in \mathbb{R}, a < b\} \subset \mathcal{P}(\mathbb{R})\,, \quad \text{also} \quad \{[a,b] : a,b \in \mathbb{R}, a < b\} \in \mathcal{P}(\mathcal{P}(\mathbb{R}))\,.$$

Der allgemeine Mengenbegriff ermöglicht es einerseits, in ganz unterschiedlichen Situationen (die nichts mit Zahlen und Funktionen zu tun haben brauchen) mathematische Theorien zu entwickeln und zum Einsatz zu bringen; dabei ist es gleichgültig, ob die Akteure das als Mathematik bezeichnen oder nicht.

Es hat sich aber andererseits sehr bald herausgestellt, dass der schrankenlose Umgang mit dem Mengenbegriff zu Widersprüchen führt. ist. So ist es z.B. nicht von vorneherein ausgeschlossen, dass eine Menge sich selbst als Element enthält (obwohl das zugegebenermaßen etwas merkwürdig klingt). Damit hat Bertrand Russell das folgende Beispiel konstruiert:

> Wir nennen eine Menge M normal, falls sie sich nicht selbst als Element enthält, also falls $M \notin M$ gilt. Sei nun
>
> $$\mathcal{M} = \{M : M \text{ ist eine normale Menge}\}\,.$$
>
> Frage: Ist $\mathcal{M}$ normal? Falls ja, so gilt $\mathcal{M} \notin \mathcal{M}$ nach Definition des Begriffs 'normal', aber andererseits $\mathcal{M} \in \mathcal{M}$ nach Definition von $\mathcal{M}$. Falls nein, so gilt $\mathcal{M} \in \mathcal{M}$ nach Definition des Begriffs 'nicht normal', aber andererseits $\mathcal{M} \notin \mathcal{M}$ nach Definition von $\mathcal{M}$.

Der Versuch, solche sogenannte Antinomien (es gibt noch mehr davon) in den Griff zu kriegen, führte zu schwierigen Grundlagenproblemen der Mathematik und hat in der Tat zu Beginn des 20. Jahrhunderts eine Grundlagenkrise der Mathematik ausgelöst. Wir werden uns aber mit diesen Problemen nicht weiter beschäftigen.

Die Inklusion "$\subset$" definiert eine Ordnungsrelation auf der Potenzmenge $\mathcal{P}(M)$ für jede beliebige Menge M. Sie ist reflexiv (da $A \subset A$), transitiv (da $A \subset B \subset C$ impliziert, dass $A \subset C$) und antisymmetrisch (da $A \subset B$ und $B \subset A$ impliziert, dass $A = B$).

Eine Nebenbemerkung: Man kann, wenn man will, die natürlichen Zahlen ebenfalls auf Mengen zurückführen. Man betrachtet sukzessive die Mengen

$$M_0 = \emptyset\,,\ M_1 = \{\emptyset\}\,,\ M_2 = \{\emptyset\,,\ \{\emptyset\}\}\,,\ M_3 = \{\emptyset\,,\ \{\emptyset\}\,,\ \{\emptyset\,,\ \{\emptyset\}\}\,, \ldots$$

als Darstellung der Zahlen $0, 1, 2, \ldots$. Das Bildungsgesetz ist

$$M_{n+1} = M_n \cup \{M_n\}\,.$$

Es gilt $M_n \subset M_{n+1}$, was der Ungleichung $n \leq n + 1$ von Zahlen entspricht. Auf diese Weise gewinnt man alle natürlichen Zahlen aus einer einzigen Menge, nämlich der leeren Menge.

4 Die komplexen Zahlen

Die Gleichung
$$x^2 + 1 = 0 \tag{4.1}$$
hat keine Lösung $x \in \mathbb{R}$. Die allgemeine Lösungsformel für quadratische Gleichungen (oder einfaches Hinsehen) liefert die beiden Lösungen
$$\pm\sqrt{-1}\,.$$
Eine mögliche Reaktion auf dieses Ergebnis ist: Die Wurzel einer negativen Zahl zu ziehen ist Unfug, und (4.1) hat ja eh keine Lösung, was soll das also. Betrachtet man aber beispielsweise die Gleichung $x^3 = 15x + 4$, so hat sie die reelle Lösung $x = 4$. Es erscheint völlig berechtigt, nach einer allgemeinen Lösungsformel zu suchen, die diese Lösung liefert. Damit haben sich Cardano und Bombelli im 16. Jahrhundert beschäftigt. In ihren Rechnungen tauchen immer wieder Wurzeln aus negativen Zahlen auf. Es hat aber mehr als 200 Jahre gedauert, bis sie als vollwertige Zahlen anerkannt wurden. Endgültig geschah das erst zu Beginn des 19. Jahrhundert mit der Interpretation komplexer Zahlen
$$x + iy\,, \quad i = \sqrt{-1}\,, \quad x, y \in \mathbb{R}\,, \tag{4.2}$$
als Punkte (x, y) in der Ebene. Seither sind die komplexen Zahlen zu einem Werkzeug geworden, welches in den verschiedensten Teilgebieten der Mathematik und ihrer Anwendungen umfangreich verwendet wird.

Addition und Multiplikation komplexer Zahlen sind so definiert, dass sie zu (4.2) passen, also
$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)\,,$$
und
$$(x_1 + iy_1) \cdot (x_2 + iy_2) = x_1x_2 + ix_1y_2 + iy_1x_2 + i^2x_2y_2 = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)\,.$$

Definition 4.1 (Komplexe Zahlen)
Wir definieren
$$\mathbb{C} = \{z : z = (x, y),\ x \in \mathbb{R},\ y \in \mathbb{R}\} \tag{4.3}$$
als die Menge aller geordneten Paare reeller Zahlen. Wir definieren eine Addition $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ *durch*
$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \tag{4.4}$$
und eine Multiplikation $\cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ *durch*
$$(x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)\,. \tag{4.5}$$

Im Sinne der Gleichheit von Mengen ist also $\mathbb{C} = \mathbb{R}\times\mathbb{R} = \mathbb{R}^2$. Die Addition in $\mathbb{C}$ entspricht der üblichen Vektoraddition im $\mathbb{R}^2$. Auch die Multiplikation werden wir geometrisch interpretieren, sobald wir die Exponentialfunktion im Komplexen kennengelernt haben. Im Spezialfall $y_1 = 0$ gilt
$$(x_1, 0) \cdot (x_2, y_2) = (x_1x_2, x_1y_2) = x_1(x_2, y_2)\,,$$

das entspricht der Skalarmultiplikation des Vektors $(x_2, y_2) \in \mathbb{R}^2$ mit dem Skalar $x_1 \in \mathbb{R}$.

Die reelle Zahlengerade ist in die komplexe Zahlenebene eingebettet vermittels der Abbildung $j : \mathbb{R} \to \mathbb{C}$,

$$j(x) = (x, 0)\,.$$

Offensichtlich ist j injektiv. Es gilt für alle $x_1, x_2 \in \mathbb{R}$

$$\begin{aligned} j(x_1 + x_2) &= (x_1 + x_2, 0) = (x_1, 0) + (x_2, 0) = j(x_1) + j(x_2)\,, \\ j(x_1 x_2) &= (x_1 x_2, 0) = (x_1, 0) \cdot (x_2, 0) = j(x_1) \cdot j(x_2)\,. \end{aligned}$$

Diese Gleichungen bedeuten, dass es für die Addition und Multiplikation reeller Zahlen gleichgültig ist, ob wir sie als reelle oder als komplexe Zahlen auffassen.

Wir werden im Folgenden den Buchstaben “j” weglassen, d.h. für eine reelle Zahl x werden wir auch dann “x” schreiben, wenn wir sie als komplexe Zahl auffassen.

Man kann direkt nachrechnen, dass gilt:

Satz 4.2 $(\mathbb{C}, +, \cdot)$ *ist ein kommutativer Körper, d.h. die Eigenschaften (2.1) – (2.9) gelten, wenn man überall* $\mathbb{R}$ *durch* $\mathbb{C}$ *ersetzt.* $\square$

Das Nullelement in $\mathbb{C}$ ist $(0, 0)$, das Einselement $(1, 0)$, beide identifizieren wir wie vereinbart mit der reellen 0 und der reellen 1.

Die komplexe Zahl $(0, 1)$ heißt imaginäre Einheit und wird mit i bezeichnet. Ein $z = (x, y) \in \mathbb{C}$ können wir also schreiben als

$$z = (x, y) = x \cdot (1, 0) + y \cdot (0, 1) = x + iy\,.$$

Es ist

$$i^2 = -1\,, \quad i^3 = -i\,, \quad i^4 = 1\,, \quad i^{n+4} = i^n \quad \text{für alle } n \in \mathbb{Z}\,.$$

Definition 4.3 *Sei* $z = x + iy$ *mit* $(x, y) \in \mathbb{R}$. *Wir definieren den Realteil* $\mathrm{Re}\,(z)$ *und den Imaginärteil* $\mathrm{Im}\,(z)$ *durch*

$$\mathrm{Re}\,(z) = x\,, \quad \mathrm{Im}\,(z) = y\,, \tag{4.6}$$

die zu z *konjugiert komplexe Zahl* $\overline{z} \in \mathbb{C}$ *durch*

$$\overline{z} = x - iy\,, \tag{4.7}$$

und den Betrag $|z|$ *von* z *durch*

$$|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}\,. \tag{4.8}$$

Die Konjugation $z \mapsto \overline{z}$ entspricht der Spiegelung an der reellen Achse. Der Betrag $|z|$ einer Zahl $z \in \mathbb{C}$ ist reell und nichtnegativ, er liefert den Abstand des Punktes z vom Nullpunkt in der komplexen Ebene. Der Abstand zweier Punkte $z, w \in \mathbb{C}$ ist gegeben durch $|z - w|$.

Es folgen unmittelbar die Rechenregeln

$$\mathrm{Re}\,(z) = \frac{1}{2}(z + \overline{z})\,, \quad \mathrm{Im}\,(z) = \frac{1}{2i}(z - \overline{z})\,, \quad \forall\, z \in \mathbb{C}\,,$$

sowie

$$\begin{aligned} \overline{\overline{z}} &= z\,, \quad \forall\, z \in \mathbb{C}\,, \\ \overline{z_1 + z_2} &= \overline{z}_1 + \overline{z}_2\,, \quad \forall\, z_1, z_2 \in \mathbb{C}\,, \\ \overline{z_1 z_2} &= \overline{z}_1 \cdot \overline{z}_2\,, \quad \forall\, z_1, z_2 \in \mathbb{C}\,. \end{aligned}$$

Aus (4.8) erhalten wir eine Formel für das multiplikative Inverse. Es ist $|z|^2 = z\overline{z}$, also

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2}\,, \tag{4.9}$$

also, falls $z = x + iy$,

$$z^{-1} = \frac{1}{z} = \left(\frac{x}{x^2 + y^2}\,, \frac{-y}{x^2 + y^2} \right)\,. \tag{4.10}$$

Die Formel (4.9) ist einfacher zu merken als (4.10).
Ist $z = (x, 0) \in \mathbb{C}$, so ist $|z| = \sqrt{x^2} = |x|$. Weiter gilt

$$\mathrm{Re}\,(z) \le |z|\,, \quad \mathrm{Im}\,(z) \le |z|\,, \quad \forall\, z \in \mathbb{C}\,,$$

sowie

$$|\overline{z}| = |z|\,, \quad \forall\, z \in \mathbb{C}\,.$$

Die Eigenschaften der Betragsfunktion aus Lemma 2.20 gelten auch in $\mathbb{C}$:

Lemma 4.4 *Es gilt*

$$|z| \ge 0\,, \quad |z| = 0 \Leftrightarrow z = 0\,, \quad \forall\, z \in \mathbb{C}\,, \tag{4.11}$$

$$|z_1 z_2| = |z_1|\,|z_2|\,, \quad \forall\, z_1, z_2 \in \mathbb{C}\,, \tag{4.12}$$

sowie die Dreiecksungleichung [2]

$$|z_1 + z_2| \le |z_1| + |z_2|\,, \quad \forall\, z_1, z_2 \in \mathbb{C}\,. \tag{4.13}$$

Beweis: (4.11) ist klar wegen $|z| = \sqrt{x^2 + y^2}$ für $z = (x, y)$.
Zu (4.12): Für alle $z_1, z_2 \in \mathbb{C}$ gilt

$$|z_1 z_2|^2 = (z_1 z_2) \cdot (\overline{z_1 z_2}) = z_1 z_2 \overline{z}_1 \overline{z}_2 = z_1 \overline{z}_1 z_2 \overline{z}_2 = |z_1|^2 |z_2|^2\,,$$

also folgt

$$|z_1 z_2| = \sqrt{|z_1|^2 |z_2|^2} = \sqrt{|z_1|^2}\sqrt{|z_2|^2} = |z_1||z_2|\,.$$

Zu (4.13): Für alle $z_1, z_2 \in \mathbb{C}$ gilt

$$\mathrm{Re}\,(z_1 \overline{z}_2) \le |z_1 \overline{z}_2| = |z_1||z_2|\,,$$

also

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1 + z_2}) = z_1 \overline{z}_1 + z_1 \overline{z}_2 + z_2 \overline{z}_1 + z_2 \overline{z}_2 \\ &= |z_1|^2 + 2\mathrm{Re}\,(z_1 \overline{z}_2) + |z_2|^2 \le |z_1|^2 + 2|z_1||z_2| + |z_2|^2 = (|z_1| + |z_2|)^2\,, \end{aligned}$$

also folgt (4.13) wieder durch Wurzelziehen. □

[2] Man kann sich deren geometrische Bedeutung klarmachen, indem man das aus den Punkten 0, z_1 und $z_1 + z_2$ in der komplexen Ebene gebildete Dreieck betrachtet.

5 Folgen

Wir betrachten die durch $f(n) = 2n$ definierte Funktion $f : \mathbb{N} \to \mathbb{N}$. Sie erzeugt die Folge der geraden Zahlen $0, 2, 4, 6, \ldots$. Statt Zahlen können auch andere Objekte auftreten. So ist

$$a\,,\ aa\,,\ aaa\,,\ aaaa\,, \ldots$$

eine Folge, deren n-tes Element aus einer n-fachen Aneinanderreihung des Buchstabens "a" besteht.

Definition 5.1 (Folge)
Sei M Menge. Eine Abbildung von $\mathbb{N}$ nach M heißt Folge in M. Eine Folge in $\mathbb{R}$ heißt auch reelle Folge oder reelle Zahlenfolge.

In der Analysis wird eine Folge in der Regel nicht in der Form

$$f : \mathbb{N} \to M$$

aufgeschrieben, sondern mit

$$(x_n)_{n\in\mathbb{N}} \tag{5.1}$$

oder so ähnlich bezeichnet. Mit (5.1) ist diejenige Folge gemeint, die durch die Abbildung $f : \mathbb{N} \to M$ mit $f(n) = x_n$ definiert wird. Das Element $x_n \in M$ heißt dann das n-te **Glied** (oder **Element**) der Folge $(x_n)_{n\in\mathbb{N}}$.

Oft kommt es vor, dass die Nummerierung der Folgenglieder nicht bei 0 anfängt, sondern bei 1 oder irgendeiner anderen ganzen Zahl. Die Folge

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$

wird durch die Vorschrift $f(n) = 1/n$ erzeugt mit $f : \mathbb{N} \setminus \{0\} \to \mathbb{R}$. Um eine Zahl $n_0 \in \mathbb{Z}$ als Anfang festzulegen, können wir schreiben

$$(x_n)_{n\geq n_0} \tag{5.2}$$

und meinen damit die Abbildung $f : N \to M$, $N = \{n : n \in \mathbb{Z},\, n \geq n_0\}$. In Erweiterung von Definition 5.1 spricht man auch in diesem Fall von Folgen.

Weitere Beispiele für reelle Folgen sind:

$$x_n = c \qquad \forall\, n \in \mathbb{N}\,, \quad c \in \mathbb{R} \text{ fest}, \qquad \text{(konstante Folge)} \tag{5.3}$$

$$x_n = n^2 \qquad \forall\, n \in \mathbb{N}\,, \tag{5.4}$$

$$x_n = x^n \qquad \forall\, n \in \mathbb{N}\,, \quad x \in \mathbb{R} \text{ fest}. \tag{5.5}$$

Wir kommen nun zum Begriff des Grenzwerts, das ist **der** zentrale grundlegende Begriff in der Analysis. Wir behandeln zuerst Grenzwerte von Folgen, später werden wir auch Grenzwerte von Funktionen besprechen.

Wir betrachen nochmals die durch $x_n = 1/n$ definierte Folge

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$$

Je größer n ist, desto näher liegt $x_n = 1/n$ bei 0; einerseits ist $x_n \neq 0$ für alle n, andererseits "strebt x_n gegen 0". Man könnte nun sagen, dass die 0 "nach unendlich vielen Schritten" erreicht wird, aber wie hat man sich das vorzustellen? Darüber ist 2000 Jahre lang debattiert worden.

Wir wechseln die Perspektive. Anstelle die Folgenglieder nacheinander zu durchlaufen, betrachten wir die Situation vom Punkt 0 aus. Legt man ein Intervall $(-\varepsilon, \varepsilon)$ um 0, so liegen die Folgenglieder x_n zunächst außerhalb, aber von einem gewissen Index n_0 an innerhalb des Intervalls. Verkleinert man ε, so bleibt diese Eigenschaft erhalten, falls man n_0 geeignet vergrößert.

Aus dieser Überlegung entsteht der Grenzwertbegriff, so wie er sich im 19. Jahrhundert durchgesetzt hat und seither verwendet wird.

Definition 5.2 (Grenzwert einer Folge)
Sei $(x_n)_{n\in\mathbb{N}}$ eine reelle Folge. Ein $a \in \mathbb{R}$ heißt Grenzwert (oder Limes) von $(x_n)_{n\in\mathbb{N}}$, falls es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt mit

$$|x_n - a| < \varepsilon \qquad \text{für alle } n \in \mathbb{N} \text{ mit } n \geq n_0. \tag{5.6}$$

Falls $(x_n)_{n\in\mathbb{N}}$ einen Grenzwert a hat, so sagt man auch: "(x_n) konvergiert gegen a" oder "(x_n) ist konvergent". Falls $(x_n)_{n\in\mathbb{N}}$ keinen Grenzwert hat, so sagt man: "(x_n) divergiert" oder "(x_n) ist divergent".

Beispiel 5.3

1. Die Folge $x_n = 1/n$ konvergiert gegen 0: Zu $\varepsilon > 0$ wählen wir $n_0 \in \mathbb{N}$ so, dass
$$n_0 > \frac{1}{\varepsilon}\,.$$
Es gilt dann für alle $n \geq n_0$
$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon\,,$$
damit erfüllt 0 die in Definition 5.2 verlangte Bedingung.

2. Die konstante Folge $x_n = c$ konvergiert gegen c: Es gilt $|x_n - c| = 0$ für alle $n \in \mathbb{N}$, wir können daher für jedes $\varepsilon > 0$ die Zahl $n_0 = 0$ wählen, und (5.6) ist erfüllt. In diesem Fall wird der Grenzwert bereits im ersten Schritt exakt (und nicht nur näherungsweise) erreicht.

3. Die Folge $x_n = (-1)^n$ divergiert: Sei $a \in \mathbb{R}$ ein Grenzwert der Folge. Für $\varepsilon = 1$ finden wir daher gemäß Definition ein $n_0 \in \mathbb{N}$ mit $|x_n - a| < 1$ für alle $n \geq n_0$. Dann gilt
$$2 = |x_{n_0+1} - x_{n_0}| \leq |x_{n_0+1} - a| + |x_{n_0} - a| < 1 + 1 = 2\,,$$
ein Widerspruch. Also kann es ein solches a nicht geben.

Eine zu Definition 5.2 äquivalente Definition des Grenzwerts erhält man, wenn man in (5.6) "$<$" durch "$\leq$" ersetzt, es also heißt: Zu jedem $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$ mit

$$|x_n - a| \leq \varepsilon \qquad \text{für alle } n \in \mathbb{N} \text{ mit } n \geq n_0.$$

Satz 5.4 *Jede reelle Folge $(x_n)_{n\in\mathbb{N}}$ hat höchstens einen Grenzwert.*

Beweis: Seien a und b Grenzwerte von (x_n). Es gelte $a \neq b$. Wir definieren

$$\varepsilon = \frac{|b-a|}{2}\,.$$

Dann ist $\varepsilon > 0$. Wähle $n_1, n_2 \in \mathbb{N}$ mit

$$|x_n - a| < \varepsilon\,, \qquad \text{für alle } n \geq n_1\,, \tag{5.7}$$

$$|x_n - b| < \varepsilon\,, \qquad \text{für alle } n \geq n_2\,. \tag{5.8}$$

Dies ist möglich, da a und b Grenzwerte sind. Für ein beliebiges $n \geq \max\{n_1, n_2\}$ folgt nun

$$|a-b| \leq |a - x_n| + |x_n - b| < \varepsilon + \varepsilon = |b-a|\,.$$

Wir haben einen Widerspruch erhalten. Also kann $a \neq b$ nicht richtig sein, und $a = b$ muss gelten. □

Satz 5.4 macht es sinnvoll, von "dem Grenzwert" einer Folge zu sprechen. Ist eine Folge (x_n) konvergent, so bezeichnet

$$\lim_{n\to\infty} x_n \tag{5.9}$$

ihren Grenzwert. Konvergiert (x_n) gegen a, so gilt also

$$a = \lim_{n\to\infty} x_n\,. \tag{5.10}$$

Wir schreiben auch

$$x_n \to a\,.$$

Die Gleichung (5.10) beinhaltet zwei Aussagen: Erstens, der Grenzwert von (x_n) existiert, und zweitens, er ist gleich a.

Lemma 5.5 (Einschließungskriterium) *Seien $(x_n), (y_n)$ Folgen in $\mathbb{R}$ mit $0 \leq |x_n| \leq y_n$ für alle $n \in \mathbb{N}$, es gelte $\lim_{n\to\infty} y_n = 0$. Dann ist auch (x_n) konvergent, und $\lim_{n\to\infty} x_n = 0$.*

Beweis: Übung. □

Definition 5.6 (Beschränkte Folge)
Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt beschränkt, wenn die Menge $\{x_n : n \in \mathbb{N}\}$ beschränkt ist.

Gemäß Lemma 2.23 ist eine Folge $(x_n)_{n\in\mathbb{N}}$ beschränkt genau dann, wenn es ein $C > 0$ gibt mit $|x_n| \leq C$ für alle $n \in \mathbb{N}$.

Satz 5.7 *Jede konvergente reelle Folge ist beschränkt.*

Beweis: Sei (x_n) reelle Folge mit $x_n \to a$. Wähle ein n_0 mit $|x_n - a| < 1$ für alle $n \geq n_0$. Setze

$$C = \max\{|x_0|, \ldots, |x_{n_0-1}|, |a| + 1\},$$

dann gilt $|x_n| \leq C$ für alle $n \in \mathbb{N}$. □

Eine beschränkte Folge braucht nicht konvergent zu sein, wie man am Beispiel $x_n = (-1)^n$ sieht.

Satz 5.8 *Seien $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ reelle Folgen mit $x_n \to a$, $y_n \to b$. Dann ist auch die Folge $(x_n + y_n)_{n\in\mathbb{N}}$ konvergent, und es gilt*

$$\lim_{n\to\infty} (x_n + y_n) = a + b\,. \tag{5.11}$$

Beweis: Sei $\varepsilon > 0$ beliebig. Wir wählen $n_1 \in \mathbb{N}$ mit

$$|x_n - a| < \frac{\varepsilon}{2}, \quad \text{für alle } n \geq n_1\,,$$

sowie $n_2 \in \mathbb{N}$ mit

$$|y_n - b| < \frac{\varepsilon}{2}, \quad \text{für alle } n \geq n_2\,.$$

Dann gilt für alle $n \geq \max\{n_1, n_2\}$

$$|(x_n + y_n) - (a + b)| \leq |x_n - a| + |y_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

□

Die Aussage von Satz 5.8 läßt sich auch als Rechenregel

$$\lim_{n\to\infty} (x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n \tag{5.12}$$

schreiben. Sie ist allerdings nur anwendbar, wenn beide Grenzwerte auf der rechten Seite existieren. Beispiel:

$$\lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = \lim_{n\to\infty} 1 + \lim_{n\to\infty} \frac{1}{n} = 1 + 0 = 1\,.$$

Satz 5.9 *Seien $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ reelle Folgen mit $x_n \to a$, $y_n \to b$. Dann ist auch die Folge $(x_n y_n)_{n\in\mathbb{N}}$ konvergent, und es gilt*

$$\lim_{n\to\infty} (x_n y_n) = ab\,. \tag{5.13}$$

Beweis: Siehe Übung. □

Aus Satz 5.9 folgt: Ist $c \in \mathbb{R}$ und (x_n) reelle Folge mit $x_n \to a$, so gilt $cx_n \to ca$.

Satz 5.10 *Sei $(x_n)_{n\in\mathbb{N}}$ reelle Folge mit $x_n \to a$, $a \neq 0$. Dann gibt es ein $n_0 \in \mathbb{N}$ mit $x_n \neq 0$ für alle $n \geq n_0$, und es gilt*

$$\lim_{n\to\infty} \frac{1}{x_n} = \frac{1}{a}\,. \tag{5.14}$$

Beweis: Siehe Übung. □

Folgerung 5.11 *Seien $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ reelle Folgen mit $x_n \to a$, $y_n \to b$, es gelte $b \neq 0$. Dann ist auch die Folge $(x_n/y_n)_{n\in\mathbb{N}}$ konvergent, und es gilt*

$$\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{a}{b}\,. \tag{5.15}$$

Beweis: Aus Satz 5.9 und Satz 5.10 folgt

$$\frac{x_n}{y_n} = x_n \cdot \frac{1}{y_n} \to a \cdot \frac{1}{b} = \frac{a}{b}\,.$$

□

Die vorstehenden Sätze lassen sich zur Berechnung von Grenzwerten algebraischer Ausdrücke verwenden, z.B. gilt für alle $k \in \mathbb{N}$

$$\lim_{n\to\infty} \frac{1}{n^k} = \lim_{n\to\infty} \prod_{i=1}^{k} \frac{1}{n} = \prod_{i=1}^{k} \lim_{n\to\infty} \frac{1}{n} = 0^k = 0\,,$$

und

$$\frac{18n^3 - 4n^2 + 8}{7n^3 + 4n} = \frac{18 - \frac{4}{n} + \frac{8}{n^3}}{7 + \frac{4}{n^2}} \to \frac{18}{7}\,.$$

Satz 5.12 *Seien $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ reelle Folgen mit $x_n \to a$, $y_n \to b$, es gelte $x_n \leq y_n$ für alle $n \in \mathbb{N}$. Dann ist $a \leq b$.*

Beweis: Wir nehmen an, dass $a > b$. Wir wählen n_0 so, dass

$$|x_n - a| < \frac{a-b}{2}\,, \quad |y_n - b| < \frac{a-b}{2}\,,$$

gilt für alle $n \geq n_0$. Dann gilt für alle solche n, dass

$$\begin{aligned} y_n - x_n &= (y_n - b) + (b - a) + (a - x_n) \\ &< \frac{a-b}{2} + (b-a) + \frac{a-b}{2} \\ &= 0 \end{aligned}$$

im Widerspruch zur Voraussetzung $x_n \leq y_n$. □

Definition 5.13 (Monoton wachsende Folge)
Sei $(x_n)_{n\in\mathbb{N}}$ eine reelle Folge. (x_n) heißt monoton wachsend, wenn $x_n \leq x_{n+1}$ gilt für alle $n \in \mathbb{N}$. (x_n) heißt streng monoton wachsend, wenn $x_n < x_{n+1}$ gilt für alle $n \in \mathbb{N}$. (x_n) heißt (streng) monoton fallend, wenn $(-x_n)$ (streng) monoton wachsend ist. (x_n) heißt (streng) monoton, wenn x_n (streng) monoton wachsend oder (streng) monoton fallend ist.

Zur Notation: Statt
$$\sup\{x_n : n \in \mathbb{N}\}$$
schreiben wir wie bereits im vorigen Kapitel besprochen auch
$$\sup_{n\in\mathbb{N}} x_n \,.$$

Satz 5.14 *Jede nach oben beschränkte monoton wachsende reelle Folge $(x_n)_{n\in\mathbb{N}}$ ist konvergent, und es gilt*
$$\lim_{n\to\infty} x_n = \sup_{n\in\mathbb{N}} x_n \,. \tag{5.16}$$
Jede nach unten beschränkte monoton fallende reelle Folge $(x_n)_{n\in\mathbb{N}}$ ist konvergent, und es gilt
$$\lim_{n\to\infty} x_n = \inf_{n\in\mathbb{N}} x_n \,. \tag{5.17}$$

Beweis: Sei $a = \sup_{n\in\mathbb{N}} x_n$. Sei $\varepsilon > 0$. Gemäß Übungsaufgabe können wir ein $n_0 \in \mathbb{N}$ finden mit
$$a - \varepsilon < x_{n_0} \le a \,.$$
Da (x_n) monoton wachsend ist, gilt für alle $n \ge n_0$
$$a - \varepsilon < x_{n_0} \le x_n \le a \,,$$
also folgt $x_n \to a$ und damit (5.16). Zum Beweis von (5.17) wenden wir (5.16) auf die Folge $(-x_n)$ an und erhalten
$$\lim_{n\to\infty} (-x_n) = \sup_{n\in\mathbb{N}} (-x_n) = -\inf_{n\in\mathbb{N}} x_n$$
woraus (5.17) folgt wegen $\lim(-x_n) = -\lim x_n$. □

Teilfolgen. Oft ist man daran interessiert, zu einer gegebenen Folge, etwa
$$0, 1, 2, 3, 4, 5, 6, \ldots$$
eine Teilfolge, etwa
$$0, 2, 4, 6, 8, \ldots$$
zu betrachten. Formal sieht das so aus: Ist M Menge, $f : \mathbb{N} \to M$ eine Folge, und $g : \mathbb{N} \to \mathbb{N}$ eine streng monoton wachsende Abbildung (d.h. $n < m \Rightarrow g(n) < g(m)$), so beschreibt $f \circ g$ eine Teilfolge von f. In obigem Beispiel ist
$$f(n) = n \,, \quad g(n) = 2n \,, \quad (f \circ g)(n) = 2n \,.$$
Äquivalent dazu ist die folgende Definition, die die in der Analysis übliche Schreibweise mit "Doppelindizes" verwendet.

Definition 5.15 (Teilfolge)
Sei M Menge, $(x_n)_{n\in\mathbb{N}}$ eine Folge in M. Ist $(n_k)_{k\in\mathbb{N}}$ eine streng monoton wachsende Folge in $\mathbb{N}$, so heißt die Folge $(x_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$.

Im obigen Beispiel ist $n_k = 2k$. Als weiteres Beispiel betrachten wir

$$x_n = \frac{1}{n}\,, \quad n_k = 3k\,, \quad x_{n_k} = \frac{1}{3k}\,.$$

Diese Teilfolge enthält jedes dritte Element der ursprünglichen Folge.

Jede Folge ist Teilfolge von sich selbst (setze $n_k = k$). Für jede streng monoton wachsende Folge (n_k) in $\mathbb{N}$ gilt

$$k \le n_k\,. \tag{5.18}$$

Beweis mit Induktion: $0 \le n_0$, und $k \le n_k \Rightarrow k+1 \le n_k + 1 \le n_{k+1}$. (Aus $n > m$ in $\mathbb{N}$ folgt $n \ge m+1$.)

Satz 5.16 *Sei $(x_n)_{n\in\mathbb{N}}$ eine konvergente reelle Folge. Dann ist jede Teilfolge (x_{n_k}) konvergent, und es gilt*

$$\lim_{k\to\infty} x_{n_k} = \lim_{n\to\infty} x_n\,. \tag{5.19}$$

Beweis: Sei $x_n \to a$, sei $\varepsilon > 0$ beliebig. Wähle $n_0 \in \mathbb{N}$ mit $|x_n - a| < \varepsilon$ für alle $n \ge n_0$, dann gilt wegen $n_k \ge k$ auch

$$|x_{n_k} - a| < \varepsilon\,, \qquad \text{für alle } k \ge n_0\,.$$

□

Satz 5.17 *Jede reelle Folge $(x_n)_{n\in\mathbb{N}}$ hat eine monotone Teilfolge.*

Beweis: Wir definieren

$$M = \{n : n \in \mathbb{N},\, x_n \ge x_k \text{ für alle } k \ge n\}\,.$$

Fall 1: M ist endlich. Sei m das größte Element von M (falls $M = \emptyset$, setzen wir $m = -1$.). Wir setzen $n_0 = m+1$. Ist n_k bereits definiert, so wählen wir ein $n_{k+1} > n_k$ mit $x_{n_k} < x_{n_{k+1}}$. (Ein solches n_{k+1} gibt es, da $n_k \notin M$.) Die hierdurch definierte Folge $(x_{n_k})_{k\in\mathbb{N}}$ ist streng monoton wachsend.
Fall 2: M ist unendlich. Wir wählen $n_0 \in M$ beliebig. Ist $n_k \in M$ bereits definiert, so wählen wir ein $n_{k+1} \in M$ mit $n_{k+1} > n_k$. Nach Definition von M gilt $x_{n_k} \ge x_{n_{k+1}}$. Die hierdurch definierte Folge $(x_{n_k})_{k\in\mathbb{N}}$ ist monoton fallend. □

Satz 5.18 (Bolzano-Weierstraß)
Jede beschränkte reelle Folge $(x_n)_{n\in\mathbb{N}}$ besitzt eine konvergente Teilfolge.

Beweis: Nach Satz 5.17 hat (x_n) eine monotone Teilfolge (x_{n_k}). Als Teilfolge einer beschränkten Folge ist sie ebenfalls beschränkt. Nach Satz 5.14 ist (x_{n_k}) konvergent. □

Der Satz von Bolzano-Weierstraß ist ein grundlegender Satz der Analysis. Er besagt: Können wir eine Folge nach oben und unten abschätzen, so folgt die Existenz eines Grenzwerts; zugegeben, nur für eine Teilfolge, die wir nicht kennen, aber das braucht nicht zu stören.

Definition 5.19 (Häufungspunkt)
Sei $(x_n)_{n\in\mathbb{N}}$ eine reelle Folge. Ein $a \in \mathbb{R}$ heißt Häufungspunkt von (x_n), wenn es eine Teilfolge (x_{n_k}) von (x_n) gibt mit $x_{n_k} \to a$.

Aus Satz 5.18 folgt also, dass jede beschränkte Folge mindestens einen Häufungspunkt besitzt. Eine konvergente Folge hat wegen (5.19) genau einen Häufungspunkt, nämlich ihren Grenzwert. Die divergente Folge

$$x_n = (-1)^n$$

besitzt die beiden Häufungspunkte

$$1 = \lim_{k\to\infty} x_{2k}\,, \quad -1 = \lim_{k\to\infty} x_{2k+1}\,.$$

Definition 5.20 (Cauchyfolge)
Eine reelle Folge $(x_n)_{n\in\mathbb{N}}$ heißt Cauchyfolge, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt mit

$$|x_n - x_m| < \varepsilon\,, \quad \text{für alle } n, m \geq n_0\,. \tag{5.20}$$

Satz 5.21 *Sei $(x_n)_{n\in\mathbb{N}}$ eine konvergente reelle Folge. Dann ist (x_n) eine Cauchyfolge.*

Beweis: Sei $x_n \to a$, sei $\varepsilon > 0$ beliebig. Wir wählen $n_0 \in \mathbb{N}$ mit

$$|x_n - a| < \frac{\varepsilon}{2}$$

für alle $n \geq n_0$. Dann gilt für alle $n, m \geq n_0$

$$|x_n - x_m| \leq |x_n - a| + |x_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,,$$

also ist (x_n) Cauchyfolge. □

Satz 5.22 *Sei $(x_n)_{n\in\mathbb{N}}$ eine reelle Cauchyfolge. Dann ist (x_n) konvergent.*

Beweis: Wir wählen ein $n_0 \in \mathbb{N}$ mit $|x_n - x_{n_0}| < 1$ für alle $n \geq n_0$ und setzen

$$C = \max\{|x_0|, \ldots, |x_{n_0-1}|, |x_{n_0}| + 1\}\,.$$

Dann gilt $|x_n| \leq C$ für alle $n \in \mathbb{N}$, also ist (x_n) beschränkt. Sei (x_{n_k}) eine konvergente Teilfolge – eine solche existiert nach Satz 5.18 – sei $x_{n_k} \to a$. Sei nun $\varepsilon > 0$ beliebig. Wir wählen $m_1, m_2 \in \mathbb{N}$ mit

$$|x_{n_k} - a| < \frac{\varepsilon}{2}\,, \quad \text{für alle } k \geq m_1\,,$$

$$|x_n - x_m| < \frac{\varepsilon}{2}\,, \quad \text{für alle } n, m \geq m_2\,.$$

Wir setzen $m_3 = \max\{m_1, m_2\}$. Dann gilt für alle $k \geq m_3$ wegen $n_k \geq k$

$$|x_k - a| \leq |x_k - x_{n_k}| + |x_{n_k} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,,$$

also folgt $x_n \to a$. □

Um die Konvergenz einer reellen Folge (x_n) nachzuprüfen, genügt es also festzustellen, dass (x_n) eine Cauchyfolge ist. Den Grenzwert brauchen wir dabei nicht zu kennen.

Uneigentliche Konvergenz. Die Folgen

$$0, 1, 2, 3, 4, 5, \ldots \tag{5.21}$$

und

$$0, 1, 0, 2, 0, 3, 0, 4, 0, 5, 0, \ldots \tag{5.22}$$

sind beide divergent und nach oben unbeschränkt. Im Falle der Folge (5.21) möchte man aber irgendwie ausdrücken, dass sie "gegen $+\infty$ strebt".

Definition 5.23 (Uneigentliche Konvergenz)
Eine reelle Folge $(x_n)_{n\in\mathbb{N}}$ heißt uneigentlich konvergent gegen $+\infty$, wenn es zu jedem $C > 0$ ein $n_0 \in \mathbb{N}$ gibt mit

$$x_n \geq C\,, \qquad \textit{für alle } n \geq n_0\,. \tag{5.23}$$

Eine reelle Folge $(x_n)_{n\in\mathbb{N}}$ heißt uneigentlich konvergent gegen $-\infty$, wenn $(-x_n)$ uneigentlich konvergent gegen $+\infty$ ist. Eine reelle Folge $(x_n)_{n\in\mathbb{N}}$ heißt uneigentlich konvergent, wenn sie uneigentlich konvergent gegen $+\infty$ oder uneigentlich konvergent gegen $-\infty$ ist.

Für die uneigentliche Konvergenz einer Folge (x_n) verwendet man in der Analysis ebenfalls die Notation

$$\lim_{n\to\infty} x_n = +\infty\,, \quad \text{bzw.} \qquad \lim_{n\to\infty} x_n = -\infty\,.$$

Satz 5.24 *Jede monotone reelle Folge $(x_n)_{n\in\mathbb{N}}$ ist entweder konvergent oder uneigentlich konvergent.*

Beweis: Sei (x_n) monoton wachsend. Ist (x_n) nach oben beschränkt, so ist sie nach Satz 5.14 konvergent. Sei nun (x_n) nach oben unbeschränkt, sei $C > 0$ beliebig. Wir wählen ein $n_0 \in \mathbb{N}$ mit $x_{n_0} \geq C$, dann gilt $x_n \geq x_{n_0} \geq C$ für alle $n \geq n_0$, also ist (x_n) uneigentlich konvergent gegen $+\infty$. Ist (x_n) monoton fallend, so wenden wir das eben Bewiesene auf die Folge $(-x_n)$ an. □

Beispiel 5.25
Wir untersuchen die Konvergenz der durch

$$x_n = x^n \tag{5.24}$$

definierten Folge (x_n), wobei $x \in \mathbb{R}$ fest ist. Es sind verschiedene Fälle zu unterscheiden.
Fall 1: $x > 1$. Wegen $x^{n+1} = x \cdot x^n > 1 \cdot x^n = x^n$ ist (x_n) streng monoton wachsend, wegen Folgerung 2.32 ist (x_n) unbeschränkt, wegen Satz 5.24 ist (x_n) uneigentlich konvergent gegen $+\infty$.

Fall 2: $0 < x < 1$. Wegen $0 < x^{n+1} = x \cdot x^n < 1 \cdot x^n = x^n$ ist (x_n) streng monoton fallend und nach unten beschränkt, also ist (x_n) konvergent und

$$\lim_{n\to\infty} x_n = \inf_{n\in\mathbb{N}} x_n \,.$$

Zu jedem $\varepsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit $x^n < \varepsilon$ (wähle n mit $(x^{-1})^n > \varepsilon^{-1}$), also ist

$$0 \le \inf_{n\in\mathbb{N}} x_n \le 0 \,,$$

und damit folgt $x_n \to 0$.
Fall 3: $-1 < x < 0$. Nach Fall 2 gilt $|x_n| \to 0$, also auch $x_n \to 0$ nach Übungsaufgabe.
Fall 4: $x < -1$. Dann ist $x_n = (-1)^n |x|^n$ divergent.
Fall 5: $x = 0$ oder $x = 1$. (x_n) ist konstant.
Fall 6: $x = -1$. $x_n = (-1)^n$ ist divergent.

Limes superior und Limes inferior. Sei $(x_n)_{n\in\mathbb{N}}$ eine beliebige reelle Folge. Ist (x_n) nach oben beschränkt, so ist

$$y_n = \sup_{k\ge n} x_k \in \mathbb{R} \tag{5.25}$$

für alle $n \in \mathbb{N}$. Wegen

$$y_n = \sup_{k\ge n} x_k = \max\{x_n, \sup_{k\ge n+1} x_k\} = \max\{x_n, y_{n+1}\} \tag{5.26}$$

ist (y_n) eine monoton fallende Folge, also nach Satz 5.24 entweder konvergent oder uneigentlich konvergent gegen $-\infty$. Wir definieren den Limes superior der Folge (x_n) durch

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \lim_{n\to\infty} \sup_{k\ge n} x_k \,. \tag{5.27}$$

Beispiel: Für $x_n = (-1)^n$ gilt $y_n = 1$ für alle $n \in \mathbb{N}$, also

$$\limsup_{n\to\infty} (-1)^n = \lim_{n\to\infty} 1 = 1 \,.$$

Ist (x_n) nach oben unbeschränkt, so gilt $y_n = +\infty$ für alle $n \in \mathbb{N}$; wir verabreden in diesem Fall, dass

$$\limsup_{n\to\infty} x_n = +\infty \,. \tag{5.28}$$

Auf diese Weise erreichen wir, dass für jede beliebige reelle Folge (x_n) der Limes superior definiert ist, und

$$\limsup_{n\to\infty} x_n \in \mathbb{R} \cup \{-\infty, +\infty\} \,.$$

Auf analoge Weise sehen wir, dass, falls (x_n) nach unten beschränkt ist, durch

$$z_n = \inf_{k\ge n} x_k$$

eine monoton wachsende reelle Folge (z_n) definiert ist. Wir definieren den Limes inferior von (x_n) durch

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} z_n = \lim_{n\to\infty} \inf_{k\ge n} x_k \,. \tag{5.29}$$

Mit der Konvention $\liminf_{n\to\infty} x_n = -\infty$, falls (x_n) nach unten unbeschränkt ist, erhalten wir ebenfalls

$$\liminf_{n\to\infty} x_n \in \mathbb{R} \cup \{-\infty, +\infty\}$$

für jede beliebige reelle Folge (x_n).

Für das betrachtete Beispiel $x_n = (-1)^n$ gilt $z_n = -1$ für alle $n \in \mathbb{N}$, also

$$\liminf_{n\to\infty} (-1)^n = -1 \,.$$

Die Folge

$$1, -1, 0, 1, -1, 0, 1, -1, 0, \ldots$$

hat die drei Häufungspunkte -1, 0 und 1; ihr Limes superior ist 1, ihr Limes inferior ist -1.

Die Folge

$$0, 1, 0, 1, 2, 0, 1, 2, 3, 0, 1, 2, 3, 4, 0, 1, 2, 3, 4, 5 \ldots$$

hat unendlich viele Häufungspunkte (jede natürliche Zahl ist Häufungspunkt), ihr Limes superior ist $+\infty$, ihr Limes inferior ist 0.

Folgen komplexer Zahlen. Entsprechend unserer allgemeinen Definition 5.1 definiert jede Abbildung von $\mathbb{N}$ nach $\mathbb{C}$ eine Folge von komplexen Zahlen.

Definition 5.26 *Eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ heißt konvergent gegen $z \in \mathbb{C}$, falls es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt mit*

$$|z_n - z| < \varepsilon \,, \qquad \textit{für alle } n \geq n_0 \,. \tag{5.30}$$

Wir schreiben wieder

$$\lim_{n\to\infty} z_n = z \,, \quad z_n \to z \,.$$

Definition 5.26 sieht formal genauso aus wie im Reellen; der Unterschied besteht darin, dass der Betrag im Komplexen genommen wird.

Konvergenz in $\mathbb{C}$ kann auf Konvergenz in $\mathbb{R}$ zurückgeführt werden.

Lemma 5.27 *Sei $(z_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$, sei $z \in \mathbb{C}$. Dann gilt*

$$\lim_{n\to\infty} z_n = z \qquad \Leftrightarrow \qquad \lim_{n\to\infty} |z_n - z| = 0 \,. \tag{5.31}$$

Beweis: Übung. □

Es kommt also darauf an, dass die Folge der Abstände zum Grenzwert gegen 0 konvergiert. Als Beispiel betrachten wir

$$z_n = \frac{1}{n} i^n \,.$$

Es gilt $z_n \to 0$, da $|z_n - 0| = |z_n| = 1/n \to 0$. Zeichnet man z_n in der komplexen Ebene ein, so sieht man, dass (z_n) “spiralförmig” gegen 0 konvergiert.

Der folgende Satz liefert eine weitere Methode, Konvergenz im Komplexen auf Konvergenz im Reellen zurückzuführen.

Satz 5.28 *Sei $(z_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$. Dann sind äquivalent:*
(1) Die Folge (z_n) ist konvergent (in $\mathbb{C}$).
(2) Die Folgen $(\operatorname{Re} z_n)$ und $(\operatorname{Im} z_n)$ sind konvergent (in $\mathbb{R}$).
Ist (z_n) konvergent, so gilt

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} \operatorname{Re} z_n + i \lim_{n\to\infty} \operatorname{Im} z_n\,. \tag{5.32}$$

Beweis: (1) $\Rightarrow$ (2): Es ist

$$0 \le |\operatorname{Re} z_n - \operatorname{Re} z| \le |z_n - z|\,, \quad 0 \le |\operatorname{Im} z_n - \operatorname{Im} z| \le |z_n - z|\,.$$

Hieraus und aus den Lemmata 5.27, 5.5 schließen wir

$$\begin{aligned} z_n \to z \quad &\Rightarrow \quad |z_n - z| \to 0 \\ &\Rightarrow \quad |\operatorname{Re} z_n - \operatorname{Re} z| \to 0 \quad \text{und} \quad |\operatorname{Im} z_n - \operatorname{Im} z| \to 0 \end{aligned} \tag{5.33}$$

und weiter

$$\operatorname{Re} z_n \to \operatorname{Re} z\,, \quad \operatorname{Im} z_n \to \operatorname{Im} z\,. \tag{5.34}$$

Hieraus folgt (5.32) wegen

$$\lim_{n\to\infty} \operatorname{Re} z_n + i \lim_{n\to\infty} \operatorname{Im} z_n = \operatorname{Re} z + i \operatorname{Im} z = z = \lim_{n\to\infty} z_n\,.$$

(2) $\Rightarrow$ (1): Sei $\operatorname{Re} z_n \to a$, $\operatorname{Im} z_n \to b$, setze $z = a + ib$. Dann ist

$$\begin{aligned} 0 \le |z_n - z| &= |(\operatorname{Re} z_n + i \operatorname{Im} z_n) - (a + i\,b)| \\ &\le |\operatorname{Re} z_n - a| + |\operatorname{Im} z_n - b| \quad \to \quad 0\,, \end{aligned} \tag{5.35}$$

also $|z_n - z| \to 0$ und daher $z_n \to z$. □

Beispiel:

$$z_n = \frac{1}{n} + i\frac{n+1}{n} \to 0 + i \cdot 1 = i\,.$$

Satz 5.29 *Seien $(z_n)_{n\in\mathbb{N}}$ und $(w_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{C}$. Dann sind auch (z_n+w_n) und $(z_n w_n)$ konvergent, und es gilt*

$$\lim_{n\to\infty} (z_n + w_n) = \lim_{n\to\infty} z_n + \lim_{n\to\infty} w_n\,, \tag{5.36}$$

$$\lim_{n\to\infty} (z_n w_n) = (\lim_{n\to\infty} z_n) \cdot (\lim_{n\to\infty} w_n)\,. \tag{5.37}$$

Ist außerdem $\lim_{n\to\infty} w_n \neq 0$, so ist auch (z_n/w_n) konvergent, und

$$\lim_{n\to\infty} \frac{z_n}{w_n} = \frac{\lim_{n\to\infty} z_n}{\lim_{n\to\infty} w_n}\,. \tag{5.38}$$

Außerdem ist auch $(\overline{z}_n)$ konvergent, und

$$\lim_{n\to\infty} \overline{z}_n = \overline{\lim_{n\to\infty} z_n}\,. \tag{5.39}$$

Beweis: Für (5.36) – (5.38) lassen sich die Beweise für den reellen Fall wörtlich übertragen. Ist $z_n \to z$, so folgt (5.39) aus

$$\overline{z}_n = \operatorname{Re} z_n - i \operatorname{Im} z_n \to \operatorname{Re} z - i \operatorname{Im} z = \overline{z}\,.$$

□

Definition 5.30 *Eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ heißt Cauchyfolge, wenn für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit*

$$|z_n - z_m| < \varepsilon\,, \qquad \textit{für alle } n, m \geq n_0\,. \tag{5.40}$$

Lemma 5.31 *Eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ ist Cauchyfolge genau dann, wenn die Folgen $(\operatorname{Re} z_n)_{n\in\mathbb{N}}$ und $(\operatorname{Im} z_n)_{n\in\mathbb{N}}$ reelle Cauchyfolgen sind.*

Beweis: Analog zum Beweis von Satz 5.28. □

Satz 5.32 *Eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ ist konvergent genau dann, wenn sie Cauchyfolge ist.*

Beweis: Nach den Sätzen 5.21 und 5.22 sind die reellen Folgen $(\operatorname{Re} z_n)$ und $(\operatorname{Im} z_n)$ genau dann konvergent, wenn sie Cauchyfolgen sind. Die Behauptung folgt nun aus Satz 5.28. □

6 Reihen

Aus einer Folge, etwa

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

bilden wir durch Summation eine neue Folge

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

also

$$1, \frac{3}{2}, \frac{7}{4}, \frac{15}{8}, \frac{31}{16}, \dots$$

Definition 6.1 (Reihe)
Sei $n_0 \in \mathbb{Z}$, $(a_n)_{n \geq n_0}$ eine Folge in $\mathbb{C}$. Wir definieren

$$s_n = \sum_{k=n_0}^{n} a_k \,, \quad n \geq n_0 \,. \tag{6.1}$$

*Die Folge $(s_n)_{n \geq n_0}$ heißt **unendliche Reihe** oder einfach **Reihe**, s_n heißt die **n-te Partialsumme**. Wir bezeichnen die Reihe mit*

$$\sum_{k=n_0}^{\infty} a_k \,. \tag{6.2}$$

Die Reihe (6.2) heißt konvergent, falls die Folge (s_n) konvergent ist, andernfalls heißt sie divergent. Der Grenzwert der Reihe ist definiert als der Grenzwert der Folge (s_n). Wir bezeichnen ihn ebenfalls mit

$$\sum_{k=n_0}^{\infty} a_k \,. \tag{6.3}$$

Für $z \in \mathbb{C}$ definieren wir

$$z^0 = 1 \,.$$

Beispiel 6.2 (Geometrische Reihe)
Für $z \in \mathbb{C}$ betrachten wir die Reihe

$$\sum_{k=0}^{\infty} z^k \,. \tag{6.4}$$

Für die Partialsummen s_n gilt

$$(1-z)s_n = (1-z)\sum_{k=0}^{n} z^k = (1-z) + (z-z^2) + \ldots + (z^n - z^{n+1}) = 1 - z^{n+1} \,,$$

also

$$s_n = \frac{1-z^{n+1}}{1-z} \,. \tag{6.5}$$

Für $|z| < 1$ gilt $|z|^{n+1} \to 0$, also auch $z^{n+1} \to 0$ und damit $s_n \to 0$. Daher ist die Reihe (6.4) konvergent für $|z| < 1$, und es gilt

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}\,. \tag{6.6}$$

□

Beispielsweise ist

$$\sum_{k=0}^{\infty} \left(\frac{1}{3}\right)^k = \frac{1}{1-\frac{1}{3}} = \frac{3}{2}\,.$$

Für die Frage, ob eine Reihe konvergiert, ist (genauso wie bei einer Folge) das "Verhalten am Anfang" gleichgültig: Für beliebiges $m \geq n_0$ gilt, dass

$$\sum_{k=n_0}^{\infty} a_k$$

konvergent ist genau dann, wenn

$$\sum_{k=m}^{\infty} a_k$$

konvergent ist. (Beweis: Übung.) Der Grenzwert der Reihe ändert sich aber, wenn wir bei $m > n_0$ anfangen. Beispiel:

$$\frac{3}{2} = \sum_{k=0}^{\infty} \left(\frac{1}{3}\right)^k = 1 + \frac{1}{3} + \sum_{k=2}^{\infty} \left(\frac{1}{3}\right)^k, \quad \sum_{k=2}^{\infty} \left(\frac{1}{3}\right)^k = \frac{1}{6}\,.$$

Satz 6.3 *Seien $\sum_{k=n_0}^{\infty} a_k$, $\sum_{k=n_0}^{\infty} b_k$ konvergente Reihen in $\mathbb{C}$, sei $c \in \mathbb{C}$. Dann sind auch die Reihen $\sum_{k=n_0}^{\infty}(a_k + b_k)$ und $\sum_{k=n_0}^{\infty} c a_k$ konvergent, und für die Grenzwerte gilt*

$$\sum_{k=n_0}^{\infty} (a_k + b_k) = \sum_{k=n_0}^{\infty} a_k + \sum_{k=n_0}^{\infty} b_k\,, \quad \sum_{k=n_0}^{\infty} c a_k = c \sum_{k=n_0}^{\infty} a_k\,. \tag{6.7}$$

Beweis: Folgt aus Satz 5.29, angewandt auf die Folgen der Partialsummen

$$s_n = \sum_{k=n_0}^{n} a_k\,, \quad \sigma_n = \sum_{k=n_0}^{n} b_k\,.$$

□

Beispiel: Umwandlung einer rationalen Zahl von periodischer Dezimaldarstellung in eine Darstellung als Bruch, etwa

$$\begin{aligned} 0.01\overline{7} &= 10^{-2} + 7 \cdot 10^{-3} + 7 \cdot 10^{-4} + \ldots = 10^{-2} + \sum_{k=0}^{\infty} 7 \cdot 10^{-3} \cdot \left(\frac{1}{10}\right)^k \\ &= \frac{1}{100} + \frac{7}{1000} \cdot \frac{1}{1-\frac{1}{10}} = \frac{4}{225}\,. \end{aligned}$$

Satz 6.4 (Cauchy-Kriterium) *Sei $\sum_{k=n_0}^{\infty} a_k$ eine Reihe in $\mathbb{C}$. Dann sind äquivalent:*
(1) $\sum_{k=n_0}^{\infty} a_k$ ist konvergent.
(2) Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass

$$\left| \sum_{k=m}^{n} a_k \right| < \varepsilon \qquad \textit{für alle } n \geq m \geq N\,. \tag{6.8}$$

Beweis: Für die Partialsummen gilt nach Definition

$$s_n - s_{m-1} = \sum_{k=m}^{n} a_k\,. \tag{6.9}$$

Die Bedingung (2) ist also äquivalent dazu, dass (s_n) eine Cauchyfolge in $\mathbb{C}$ ist. Nach Satz 5.32 ist das äquivalent dazu, dass (s_n) konvergiert. □

Beispiel 6.5 (Harmonische Reihe)
Die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k} \tag{6.10}$$

heißt die harmonische Reihe. Sie ist divergent: Es gilt nämlich für alle $m \geq 1$

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

Für $\varepsilon = \frac{1}{2}$ läßt sich also kein N finden, so dass die Bedingung (6.8) erfüllt ist, also ist (6.10) divergent.

Folgerung 6.6 *Für jede konvergente Reihe $\sum_{k=n_0}^{\infty} a_k$ in $\mathbb{C}$ gilt*

$$\lim_{k\to\infty} a_k = 0\,. \tag{6.11}$$

Beweis: Nach Satz 6.4 finden wir zu jedem $\varepsilon > 0$ ein N mit

$$|a_m| = \left| \sum_{k=m}^{m} a_k \right| < \varepsilon \qquad \text{für alle } m \geq N\,,$$

also gilt $|a_k| \to 0$ und damit auch $a_k \to 0$. □

Die Umkehrung von Folgerung 6.6 gilt nicht, d.h. aus Bedingung (6.11) folgt nicht, dass die Reihe $\sum_{k=n_0}^{\infty} a_k$ konvergent ist. (Beispiel: Die harmonische Reihe.)

Absolute Konvergenz.

Definition 6.7
Eine Reihe $\sum_{k=n_0}^{\infty} a_k$ in $\mathbb{C}$ heißt ***absolut konvergent****, falls die Reihe $\sum_{k=n_0}^{\infty} |a_k|$ konvergent ist.*

Gilt $a_k \geq 0$ für alle k, so sind Konvergenz und absolute Konvergenz offenbar äquivalent.

Beispiel: Die geometrische Reihe

$$\sum_{k=0}^{\infty} z^k$$

ist absolut konvergent für $|z| < 1$ (da dann $\big||z|\big| = |z| < 1$), und es gilt

$$\sum_{k=0}^{\infty} |z|^k = \frac{1}{1-|z|}\,.$$

Satz 6.8 *Ist eine Reihe $\sum_{k=n_0}^{\infty} a_k$ in $\mathbb{C}$ absolut konvergent, so ist sie auch konvergent, und*

$$\left|\sum_{k=n_0}^{\infty} a_k\right| \leq \sum_{k=n_0}^{\infty} |a_k|\,. \tag{6.12}$$

Beweis: Sei $\sum_{k=n_0}^{\infty} a_k$ in $\mathbb{C}$ absolut konvergent. Dann gibt es nach Satz 6.4 zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\left|\sum_{k=m}^{n} |a_k|\right| < \varepsilon \qquad \text{für alle } n, m \geq N\,,$$

also gilt auch

$$\left|\sum_{k=m}^{n} a_k\right| \leq \sum_{k=m}^{n} |a_k| = \left|\sum_{k=m}^{n} |a_k|\right| < \varepsilon \qquad \text{für alle } n, m \geq N\,,$$

also erfüllt auch $\sum_{k=n_0}^{\infty} a_k$ das Cauchy-Kriterium und ist nach Satz 6.4 konvergent. Da

$$\left|\sum_{k=n_0}^{n} a_k\right| \leq \sum_{k=n_0}^{n} |a_k|$$

für alle $n \geq n_0$ gilt, folgt (6.12) mit Grenzübergang $n \to \infty$. □

Die Umkehrung von Satz 6.8 gilt nicht: Die alternierende harmonische Reihe

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

ist nicht absolut konvergent (da die harmonische Reihe divergiert), aber, wie wir später sehen werden, konvergent.

Satz 6.9 *Sei $\sum_{k=n_0}^{\infty} a_k$ eine Reihe in $\mathbb{C}$. Dann sind äquivalent:*
(1) $\sum_{k=n_0}^{\infty} a_k$ ist absolut konvergent.
(2) Die reelle Folge (s_n) der Partialsummen von $\sum_{k=n_0}^{\infty} |a_k|$,

$$s_n = \sum_{k=n_0}^{n} |a_k|\,, \quad n \geq n_0\,,$$

ist beschränkt.

Beweis: (1) $\Rightarrow$ (2): (s_n) ist konvergent, also beschränkt.
(2) $\Rightarrow$ (1): (s_n) ist beschränkt und monoton wachsend, also konvergent. $\square$

Reihen in $\mathbb{C}$ können wir in Real- und Imaginärteil wie folgt zerlegen. Wir betrachten

$$\sum_{k=n_0}^{\infty} a_k \,,$$

und die aus den Real- und Imaginärteilen $x_k = \operatorname{Re} a_k$, $y_k = \operatorname{Im} a_k$, gebildeten Reihen

$$\sum_{k=n_0}^{\infty} x_k \,, \quad \sum_{k=n_0}^{\infty} y_k \,.$$

Satz 6.10 *Sei $\sum_{k=n_0}^{\infty} a_k$ eine Reihe in $\mathbb{C}$, seien $x_k = \operatorname{Re} a_k$, $y_k = \operatorname{Im} a_k$. Dann gilt:*
(1) $\sum_{k=n_0}^{\infty} a_k$ ist konvergent genau dann, wenn $\sum_{k=n_0}^{\infty} x_k$ und $\sum_{k=n_0}^{\infty} y_k$ konvergent sind. In diesem Fall gilt

$$\sum_{k=n_0}^{\infty} a_k = \sum_{k=n_0}^{\infty} x_k + i \sum_{k=n_0}^{\infty} y_k \,. \tag{6.13}$$

(2) $\sum_{k=n_0}^{\infty} a_k$ ist absolut konvergent genau dann, wenn $\sum_{k=n_0}^{\infty} x_k$ und $\sum_{k=n_0}^{\infty} y_k$ absolut konvergent sind.

Beweis: Für die Partialsummen

$$s_n = \sum_{k=n_0}^{n} a_k \,, \quad r_n = \sum_{k=n_0}^{n} x_k \,, \quad p_n = \sum_{k=n_0}^{n} y_k \,,$$

gilt $s_n = r_n + ip_n$. Es folgt (1), da (s_n) genau dann konvergiert, wenn (r_n) und (p_n) konvergieren. Weiter folgt (2) mit Satz 6.9, da (s_n) genau dann beschränkt ist, wenn (r_n) und (p_n) beschränkt sind. $\square$

Als Beispiel betrachten wir die geometrische Reihe

$$\sum_{k=0}^{\infty} \big(\frac{i}{2}\big)^k = 1 + \frac{i}{2} - \frac{1}{4} - \frac{i}{8} + \frac{1}{16} \cdots = \frac{1}{1 - \frac{i}{2}} = \frac{4}{5}\big(1 + \frac{i}{2}\big) \,.$$

Die Reihe der Realteile ist

$$1 + 0 - \frac{1}{4} + 0 + \frac{1}{16} \cdots = \sum_{j=0}^{\infty} \big(\frac{i}{2}\big)^{2j} = \sum_{j=0}^{\infty} \big(-\frac{1}{4}\big)^j = \frac{1}{1 + \frac{1}{4}} = \frac{4}{5} \,.$$

Die Reihe der Imaginärteile ist

$$0 + \frac{1}{2} + 0 - \frac{1}{8} \cdots = \frac{1}{2} \sum_{j=0}^{\infty} \big(-\frac{1}{4}\big)^j = \frac{2}{5} \,.$$

Mit folgender Definition und Satz kann man die Konvergenz gegebener Reihen auf die Konvergenz bereits bekannter Reihen zurückführen.

Definition 6.11 (Majorante)
Sei $\sum_{k=n_0}^{\infty} a_k$ eine Reihe in $\mathbb{C}$. Eine reelle Reihe $\sum_{k=n_0}^{\infty} b_k$ mit $|a_k| \leq b_k$ für alle $k \geq n_0$ heißt Majorante der Reihe $\sum_{k=n_0}^{\infty} a_k$.

Satz 6.12 (Majorantenkriterium)
Sei $\sum_{k=n_0}^{\infty} a_k$ eine Reihe in $\mathbb{C}$, welche eine konvergente Majorante $\sum_{k=n_0}^{\infty} b_k$ besitzt. Dann ist $\sum_{k=n_0}^{\infty} a_k$ absolut konvergent.

Beweis: Da $b_k \geq 0$ für alle k, ist $\sum_{k=n_0}^{\infty} b_k$ absolut konvergent, also gibt es nach Satz 6.9 ein $C > 0$ mit

$$C \geq \sum_{k=n_0}^{n} b_k \geq \sum_{k=n_0}^{n} |a_k|$$

für alle $n \geq n_0$, also ist $\sum_{k=n_0}^{\infty} a_k$ absolut konvergent nach Satz 6.9. $\square$

Das Majorantenkriterium liefert allerdings keinen Hinweis darauf, wie man den Grenzwert der Reihe $\sum b_k$ berechnet.

Beispiel 6.13
(1) Wir betrachten

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \frac{k^3 + k}{k^3 - 2} z^k, \quad |z| < 1\,. \tag{6.14}$$

Für die durch

$$y_k = \frac{k^3 + k}{k^3 - 2}$$

definierte Folge gilt $y_k \to 1$, also ist sie beschränkt, also gibt es $C > 0$ mit

$$|a_k| \leq C|z|^k$$

für alle k, und $b_k = C|z|^k$ definiert eine konvergente Majorante der Reihe (6.14).
(2) Wir betrachten

$$\sum_{k=1}^{\infty} \frac{1}{k^2}\,. \tag{6.15}$$

Die Reihe

$$\sum_{k=2}^{\infty} \frac{1}{(k-1)k} \tag{6.16}$$

ist eine Majorante von

$$\sum_{k=2}^{\infty} \frac{1}{k^2}\,, \tag{6.17}$$

wir behaupten, dass sie sogar eine konvergente Majorante ist. Für die Partialsummen von (6.16) gilt nämlich

$$s_n = \frac{n-1}{n}\,. \tag{6.18}$$

Beweis von (6.18) mit Induktion: $s_2 = \frac{1}{2}$,

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

Wegen $s_n \to 1$ ist (6.16) konvergent, also auch (6.17) und damit auch (6.15).

(3) Die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^n} \tag{6.19}$$

ist für festes $n \geq 2$ ebenfalls konvergent, da sie die konvergente Majorante (6.15) hat. □

Folgerung 6.14 (Divergente Minorante) *Seien $\sum_{k=n_0}^{\infty} a_k$ und $\sum_{k=n_0}^{\infty} b_k$ Reihen in $\mathbb{R}$ mit $0 \leq a_k \leq b_k$ für alle $k \geq n_0$. Ist $\sum_{k=n_0}^{\infty} a_k$ divergent, so ist auch $\sum_{k=n_0}^{\infty} b_k$ divergent.*

Beweis: Ist $\sum_{k=n_0}^{\infty} b_k$ konvergent, so ist nach Satz 6.12 auch $\sum_{k=n_0}^{\infty} a_k$ konvergent. □

In der Situation von Folgerung 6.14 heißt $\sum_{k=n_0}^{\infty} a_k$ eine divergente Minorante von $\sum_{k=n_0}^{\infty} b_k$.

Satz 6.15 (Quotientenkriterium) *Sei $\sum_{k=n_0}^{\infty} a_k$ eine Reihe in $\mathbb{C}$. Es gebe ein $q \in \mathbb{R}$ mit $q < 1$ und*

$$\frac{|a_{k+1}|}{|a_k|} \leq q\,, \qquad \textit{für alle } k \geq n_0\,. \tag{6.20}$$

Dann ist $\sum_{k=n_0}^{\infty} a_k$ absolut konvergent.

Beweis: Folgt aus dem Majorantenkriterium, da $|a_k| \leq |a_{n_0}| q^{k-n_0}$ für $k \geq n_0$ und daher die Reihe

$$\sum_{k=n_0}^{\infty} |a_{n_0}| q^{k-n_0}$$

eine konvergente Majorante von $\sum_{k=n_0}^{\infty} |a_k|$ ist. □

Es genügt **nicht**, dass

$$\frac{|a_{k+1}|}{|a_k|} < 1\,, \qquad \text{für alle } k \geq n_0\,.$$

Das erkennt man am Beispiel der harmonischen Reihe $a_n = 1/n$: Für sie gilt

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

aber sie ist divergent. In der Tat gibt es kein $q < 1$, für das (6.20) erfüllt ist.

Beispiel 6.16 (Exponentialreihe)
Wir betrachten die sogenannte Exponentialreihe

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}\,, \tag{6.21}$$

wobei $z \in \mathbb{C}$ eine feste komplexe Zahl ist. Mit $a_k = z^k/k!$ gilt

$$\frac{|a_{k+1}|}{|a_k|} = \frac{|z|^{k+1}}{(k+1)!} \cdot \frac{k!}{|z|^k} = \frac{|z|}{k+1}\,,$$

also
$$\frac{|a_{k+1}|}{|a_k|} \le \frac{1}{2}\,, \qquad \text{für alle } k \ge 2|z| - 1\,,$$
also ist (6.21) absolut konvergent.

Die Exponentialreihe werden wir später zur Definition der Exponentialfunktion verwenden.

Ein weiteres Konvergenzkriterium, das Wurzelkriterium, basiert ebenfalls auf dem Vergleich mit der geometrischen Reihe.

Satz 6.17 (Wurzelkriterium)
Sei $\sum_{k=n_0}^{\infty} a_k$ Reihe in $\mathbb{C}$, sei
$$w = \limsup_{k\to\infty} \sqrt[k]{|a_k|}\,. \tag{6.22}$$
Ist $w < 1$, so konvergiert die Reihe absolut; ist $w > 1$, so divergiert sie.

Beweis: Ist $w > 1$, so gibt es unendlich viele k mit $\sqrt[k]{|a_k|} > 1$, also auch $|a_k| > 1$ für diese k. Daher konvergiert (a_k) nicht gegen Null, und gemäß Folgerung 6.6 ist die Reihe divergent. Sei nun $w < 1$. Wähle q mit $w < q < 1$, dann gibt es nach Definition des Limes superior ein N mit $\sqrt[k]{|a_k|} \le q$ für alle $k \ge N$. Es folgt $|a_k| \le q^k$ für alle $k \ge N$. Die geometrische Reihe $\sum_{k\ge N} q^k$ ist also eine konvergente Majorante der Reihe $\sum_{k\ge N} |a_k|$, also ist $\sum_{k\ge n_0} a_k$ absolut konvergent. □

Im Fall $w = 1$ kann sowohl Konvergenz als auch Divergenz vorliegen (das heißt, es gibt konvergente und divergente Reihen mit $w = 1$).

Alternierende Reihen. Eine reelle Reihe $\sum_{k=n_0}^{\infty} a_k$ heißt **alternierend**, wenn die Folgenglieder a_k abwechselnd positives und negatives Vorzeichen haben. So ist etwa
$$1 - 1 + 1 - 1 + 1 - 1 + \dots$$
eine alternierende Reihe. Sie ist divergent, da ihre Partialsummen abwechselnd 1 und 0 sind. Bei der alternierenden harmonischen Reihe
$$1 - \frac{1}{2} + \frac{1}{3} - \dots$$
hofft man auf Konvergenz. Das ist tatsächlich der Fall; der folgende Satz garantiert das für jede alternierende Reihe, deren Folgenglieder gegen 0 konvergieren.

Satz 6.18 (Leibnizkriterium)
Sei $(a_n)_{n\ge n_0}$ eine monoton fallende Folge in $\mathbb{R}$ mit $a_n \to 0$. Dann konvergiert die Reihe
$$\sum_{k=n_0}^{\infty} (-1)^k a_k\,, \tag{6.23}$$
und es gilt für alle $n \ge n_0$
$$\left| \sum_{k=n_0}^{\infty} (-1)^k a_k - s_n \right| = \left| \sum_{k=n_0}^{\infty} (-1)^k a_k - \sum_{k=n_0}^{n} (-1)^k a_k \right| \le a_{n+1}\,. \tag{6.24}$$

Beweis: Sei zunächst $n_0 = 0$. Für $k \geq 1$ ist $a_{2k} - a_{2k-1} \leq 0$, also

$$s_{2k} = s_{2k-2} - a_{2k-1} + a_{2k} \leq s_{2k-2}\,,$$

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

also ist $(s_{2k})_{k\in\mathbb{N}}$ monoton fallend, $(s_{2k-1})_{k\in\mathbb{N}}$ monoton wachsend. Es gilt

$$s_1 \leq s_{2k-1} = s_{2k} - a_{2k} \leq s_{2k} \leq s_0 \tag{6.25}$$

für alle k, und weiter für alle $m \geq k$

$$s_1 \leq s_{2k-1} \leq s_{2m-1} \leq s_{2m} \leq s_{2k} \leq s_0\,. \tag{6.26}$$

Also sind beide Teilfolgen (s_{2k}) und (s_{2k-1}) beschränkt, also konvergent nach Satz 5.14, und

$$s_{2k-1} \leq \underbrace{\lim_{m\to\infty} s_{2m-1}}_{=:b} \leq s_{2k}\,, \quad s_{2k-1} \leq \underbrace{\lim_{m\to\infty} s_{2m}}_{=:c} \leq s_{2k} \tag{6.27}$$

gilt für alle $k \in \mathbb{N}$. Aus (6.27) folgt

$$0 \leq |c-b| \leq a_{2k}\,,$$

und wegen $a_{2k} \to 0$ gilt $c = b$. Da für alle n entweder $s_n \leq b \leq s_{n+1}$ oder $s_n \geq b \geq s_{n+1}$ gilt, folgt

$$0 \leq |b - s_n| \leq |s_n - s_{n+1}| = a_{n+1} \to 0\,,$$

also $|b - s_n| \to 0$ wegen Lemma 5.5. Der allgemeine Fall $n_0 \in \mathbb{Z}$ kann durch Umnummerieren und Multiplikation mit -1 darauf zurückgeführt werden. □

Beispiel: Die alternierende harmonische Reihe

$$\sum_{k=1}^{\infty} (-1)^{k-1}\frac{1}{k} \tag{6.28}$$

ist nach dem Leibniz-Kriterium konvergent. (Ihr Grenzwert ist, wie sich später herausstellen wird, die Zahl $\ln 2$.) Sie ist aber nicht absolut konvergent, wie wir bereits gesehen haben.

Das Produkt von Reihen. Wir beginnen mit dem Beispiel des Produkts zweier Polynome

$$(1-2x+3x^2)(2+x) = (2-4x+6x^2) + (x-2x^2+3x^3) = 2-3x+4x^2+3x^3\,.$$

Das Ergebnis entsteht durch Ausmultiplizieren und Zusammenfassen der Terme mit gleichen Exponenten. Sind nun p und q zwei beliebige Polynome,

$$p(x) = \sum_{k=0}^{n} \alpha_k x^k\,, \quad q(x) = \sum_{k=0}^{m} \beta_k x^k\,, \tag{6.29}$$

so hat deren Produkt die Form

$$(pq)(x) = \sum_{k=0}^{n+m} \gamma_k x^k\,, \quad \gamma_k = \sum_{j=0}^{k} \alpha_{k-j}\beta_j\,. \tag{6.30}$$

Eine Reihe der Form

$$\sum_{k=0}^{\infty} \alpha_k x^k \tag{6.31}$$

heißt **Potenzreihe**. Mit ihnen werden wir uns später ausführlich beschäftigen. Hier geht es zunächst darum, die Formel (6.30) auf Potenzreihen zu übertragen, so dass die oberen Grenzen n, m und $n+m$ in (6.29) und (6.30) durch ∞ ersetzt werden können. Das führt auf das sogenannte Cauchy-Produkt von Reihen. Es wird für allgemeine Reihen, nicht nur für Potenzreihen definiert.

Satz 6.19 (Cauchy-Produkt)
Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergente Reihen in $\mathbb{C}$, sei für $k \in \mathbb{N}$

$$c_k = \sum_{j=0}^{k} a_{k-j} b_j \,. \tag{6.32}$$

Dann ist auch die Reihe $\sum_{k=0}^{\infty} c_k$ absolut konvergent, und es gilt

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right) . \tag{6.33}$$

Beweis: Wir setzen

$$s_n = \sum_{k=0}^{n} c_k \,, \tag{6.34}$$

$$s_n^* = \left(\sum_{k=0}^{n} a_k \right) \cdot \left(\sum_{k=0}^{n} b_k \right) , \quad p_n^* = \left(\sum_{k=0}^{n} |a_k| \right) \cdot \left(\sum_{k=0}^{n} |b_k| \right) . \tag{6.35}$$

Nach Satz 5.9 gilt

$$\lim_{n\to\infty} s_n^* = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right) , \quad \lim_{n\to\infty} p_n^* = \left(\sum_{k=0}^{\infty} |a_k| \right) \cdot \left(\sum_{k=0}^{\infty} |b_k| \right) .$$

Es genügt daher zu zeigen, dass

$$\lim_{n\to\infty} (s_n - s_n^*) = 0 \,. \tag{6.36}$$

Sei $\varepsilon > 0$ beliebig. Wir wählen $m \in \mathbb{N}$ mit

$$|p_n^* - p_m^*| < \varepsilon \,, \qquad \text{für alle } n \geq m \,. \tag{6.37}$$

Sei $n \geq m$, dann gilt

$$|p_n^* - p_m^*| = \sum_{(i,j)\in\Gamma_n} |a_i b_j| \,, \tag{6.38}$$

mit

$$\Gamma_n = \{(i,j) : 0 \leq i,j \leq n,\ i > m \text{ oder } j > m\} \,. \tag{6.39}$$

Weiter ist
$$s_n^* = \sum_{i,j=0}^{n} a_i b_j \,, \quad s_n = \sum_{\substack{i,j=0\\ i+j\le n}}^{n} a_i b_j \,, \tag{6.40}$$
also
$$|s_n^* - s_n| \le \sum_{(i,j)\in\Delta_n} |a_i b_j| \,, \tag{6.41}$$
mit
$$\Delta_n = \{(i,j) : 0 \le i,j \le n,\, i+j > n\} \,. \tag{6.42}$$
Für $n \ge 2m$ gilt $\Delta_n \subset \Gamma_n$, da gilt
$$(i,j) \notin \Gamma_n \quad \Rightarrow \quad i+j \le 2m \le n \quad \Rightarrow \quad (i,j) \notin \Delta_n \,.$$
Es folgt
$$|s_n^* - s_n| \le \sum_{(i,j)\in\Delta_n} |a_i b_j| \le \sum_{(i,j)\in\Gamma_n} |a_i b_j| = |p_n^* - p_m^*| < \varepsilon \,,$$
Damit ist gezeigt, dass (6.36) gilt. Die absolute Konvergenz von $\sum_{k=0}^\infty c_k$ folgt wegen
$$\sum_{k=0}^{n} |c_k| \le \sum_{\substack{i,j=0\\ i+j\le n}}^{n} |a_i b_j| \le p_n^* \le \lim_{l\to\infty} p_l^*$$
aus Satz 6.9. □

Wir wenden den vorangehenden Satz an auf das Produkt zweier Exponentialreihen.

Satz 6.20 *Es gilt für alle $z, w \in \mathbb{C}$*
$$\left(\sum_{k=0}^{\infty} \frac{z^k}{k!}\right) \cdot \left(\sum_{k=0}^{\infty} \frac{w^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{(z+w)^k}{k!} \,. \tag{6.43}$$

Beweis: Gemäß Beispiel 6.16 ist die Exponentialreihe absolut konvergent. Wir können daher Satz 6.19 anwenden auf
$$a_k = \frac{z^k}{k!} \,, \quad b_k = \frac{w^k}{k!} \,, \quad c_k = \sum_{j=0}^{k} a_{k-j} b_j \,.$$
Wir rechnen
$$\begin{aligned}\left(\sum_{k=0}^{\infty} \frac{z^k}{k!}\right) \cdot \left(\sum_{k=0}^{\infty} \frac{w^k}{k!}\right) &= \left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{k=0}^{\infty} b_k\right) = \sum_{k=0}^{\infty} c_k = \sum_{k=0}^{\infty}\sum_{j=0}^{k} a_{k-j} b_j \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{k} \frac{z^{k-j}}{(k-j)!}\frac{w^j}{j!} = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} z^{k-j} w^j \\ &= \sum_{k=0}^{\infty} \frac{(z+w)^k}{k!} \,.\end{aligned}$$
Beim letzten Gleichheitszeichen wurde die binomische Formel verwendet (Satz 2.4, dieser gilt auch in $\mathbb{C}$ mit demselben Beweis wie in $\mathbb{R}$). □

Der vorangehende Satz liefert eine wesentliche Grundlage für das Rechnen mit der Exponentialfunktion.

7 Stetige Funktionen, Zwischenwertsatz

In diesem Abschnitt betrachten wir Funktionen

$$f : D \to \mathbb{K}\,, \quad D \subset \mathbb{K}\,. \tag{7.1}$$

Hier steht $\mathbb{K}$ für $\mathbb{R}$ oder für $\mathbb{C}$. Das bedeutet, (7.1) ist eine Kurzform dafür, dass wir sowohl Funktionen

$$f : D \to \mathbb{R}\,, \quad D \subset \mathbb{R}\,,$$

als auch Funktionen

$$f : D \to \mathbb{C}\,, \quad D \subset \mathbb{C}\,,$$

betrachten.

Typische Definitionsgebiete $D \subset \mathbb{R}$ sind Intervalle.

Definition 7.1 (Abschluss)
Sei $D \subset \mathbb{R}$. Wir definieren den Abschluss $\overline{D}$ von D durch

$$\overline{D} = \{a : a \in \mathbb{R},\ \textit{es gibt eine Folge } (x_n) \textit{ in } D \textit{ mit } x_n \to a\}\,. \tag{7.2}$$

Es gilt $D \subset \overline{D}$, da jede konstante Folge $x_n = a \in D$ den Grenzwert a hat.

Beispiel 7.2
1. In $\mathbb{R}$:

$$[a, b] = \overline{(a, b)} = \overline{(a, b]} = \overline{[a, b)} = \overline{[a, b]}\,.$$

2. In $\mathbb{C}$: Für $r > 0$ ist

$$B_r = \{z : z \in \mathbb{C}\,,\ |z| < r\}$$

die Kreisscheibe um 0 mit Radius r ohne ihren Rand (genannt: die offene Kreisscheibe um 0). Es gilt (Übung)

$$\overline{B}_r = \{z : z \in \mathbb{C}\,,\ |z| \le r\}\,.$$

3. In $\mathbb{R}$: Es gilt $\overline{\mathbb{Q}} = \mathbb{R}$. Ist nämlich $a \in \mathbb{R}$, $a > 0$, so setzen wir für $n \ge 1$

$$x_n = \frac{m}{n}\,, \quad m = \sup\left\{k : \frac{k}{n} \le a\,,\ k \in \mathbb{N}\right\}.$$

Es ist dann $x_n \in \mathbb{Q}$ und

$$x_n \le a < x_n + \frac{1}{n}$$

und daher $x_n \to a$ sowie $-x_n \to -a$ für $n \to \infty$. □

Definition 7.3 (Grenzwert einer Funktion)
Sei $D \subset \mathbb{K}$, $f : D \to \mathbb{K}$ und $a \in \overline{D}$. Die Zahl $c \in \mathbb{K}$ heißt Grenzwert von f an der Stelle a (oder: in a), falls

$$\lim_{n\to\infty} f(x_n) = c \tag{7.3}$$

gilt für jede Folge (x_n) in D mit $x_n \to a$.

Da Grenzwerte von Folgen eindeutig bestimmt sind, kann es keine zwei verschiedenen Werte von c geben, welche die Bedingung in Definition 7.3 erfüllen. Also:

Lemma 7.4 *Sei $D \subset \mathbb{K}$, $f : D \to \mathbb{K}$ und $a \in \overline{D}$. Dann hat f höchstens einen Grenzwert in a.* □

Notation 7.5
Ist c der Grenzwert von $f : D \to \mathbb{R}$ in a, so schreiben wir

$$\lim_{x\to a} f(x) = c\,, \quad \text{oder auch} \quad \lim_{\substack{x\to a\\ x\in D}} f(x) = c\,,$$

oder auch

$$f(x) \to c \quad \text{für} \quad x \to a\,.$$

Beispiel 7.6
(1) Jede Funktion $f : D \to \mathbb{K}$, die konstant gleich c ist, hat in jedem Punkt $a \in \overline{D}$ den Grenzwert c.
(2) Die Funktion $f : D \to \mathbb{K}$, $f(x) = x$, hat in jedem Punkt $a \in \overline{D}$ den Grenzwert a.
(3) Die durch

$$f(x) = \begin{cases} 1\,, & x > 0\,, \\ 0\,, & x \le 0\,, \end{cases}$$

definierte Funktion $f : \mathbb{R} \to \mathbb{R}$ hat in Punkten $a < 0$ den Grenzwert 0, in Punkten $a > 0$ den Grenzwert 1, aber im Punkt $a = 0$ keinen Grenzwert, da $f(x_n) = 0$ gilt für Folgen $x_n \to 0$, die aus negativen Zahlen bestehen, und $f(x_n) = 1$ gilt für solche Folgen, die aus positiven Zahlen bestehen.
(4) Die durch

$$f(x) = \begin{cases} 1\,, & x \in \mathbb{Q}\,, \\ 0\,, & x \notin \mathbb{Q}\,, \end{cases}$$

definierte Funktion $f : \mathbb{R} \to \mathbb{R}$ hat in keinem Punkt $a \in \mathbb{R}$ einen Grenzwert, da es für jedes a sowohl Folgen $x_n \to a$ gibt, die in $\mathbb{Q}$ verlaufen, als auch Folgen $x_n \to a$, die in $\mathbb{R} \setminus \mathbb{Q}$ verlaufen.

Definition 7.7 (Stetigkeit)
*Sei $D \subset \mathbb{K}$, $f : D \to \mathbb{K}$. Ist $a \in D$, so heißt f **stetig in a**, falls*

$$\lim_{x\to a} f(x) = f(a)\,. \tag{7.4}$$

*f heißt **stetig** (oder deutlicher: stetig in D), falls f in jedem Punkt $a \in D$ stetig ist.*

Eine Funktion $f : D \to \mathbb{R}$ ist also genau dann stetig in $a \in D$, falls

$$\lim_{n\to\infty} f(x_n) = f(a) \tag{7.5}$$

gilt für alle Folgen (x_n) in D mit $x_n \to a$. f ist genau dann stetig in (oder: auf) D, falls

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right) \tag{7.6}$$

gilt für jede Folge (x_n) in D, welche gegen einen Grenzwert in D konvergiert.

Beispiel 7.8
(1) Jede konstante Funktion ist stetig in $\mathbb{K}$.
(2) Die Identität, also die durch $f(x) = x$ definierte Funktion, ist stetig in $\mathbb{K}$.
(3) Die durch

$$f(x) = \begin{cases} 1\,, & x > 0\,, \\ 0\,, & x \le 0\,, \end{cases}$$

definierte Funktion $f : \mathbb{R} \to \mathbb{R}$ ist nicht stetig in 0, aber stetig in allen anderen Punkten $a \neq 0$.
(4) Die durch

$$f(x) = \begin{cases} 1\,, & x \in \mathbb{Q}\,, \\ 0\,, & x \notin \mathbb{Q}\,, \end{cases}$$

definierte Funktion $f : \mathbb{R} \to \mathbb{R}$ ist in keinem Punkt $a \in \mathbb{R}$ stetig. □

Algebraische Operationen mit Funktionen. Sind $f, g : D \to \mathbb{K}$ und $\lambda \in \mathbb{K}$, so sind

$$f + g, f \cdot g, \lambda f : D \to \mathbb{K}$$

definiert durch

$$(f+g)(x) = f(x) + g(x)\,, \quad (f \cdot g)(x) = f(x) \cdot g(x)\,, \quad (\lambda f)(x) = \lambda f(x)\,,$$

und

$$\frac{f}{g} : D \setminus \{x : g(x) = 0\} \to \mathbb{K}$$

durch

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}\,.$$

Satz 7.9 *Seien $D \subset \mathbb{K}$, $f, g : D \to \mathbb{K}$, $a \in D$ und $\lambda \in \mathbb{K}$. Sind f und g stetig in a, so sind auch $f+g$, fg und λf stetig in a; ist $g(a) \neq 0$, so ist auch f/g stetig in a.*

Beweis: Folgt aus den entsprechenden Sätzen für Grenzwerte von Folgen. Ist etwa (x_n) Folge in D mit $x_n \to a$, so gilt

$$(fg)(a) = f(a)g(a) = \left(\lim_{n\to\infty} f(x_n)\right) \cdot \left(\lim_{n\to\infty} g(x_n)\right) = \lim_{n\to\infty} (f(x_n)g(x_n)) = \lim_{n\to\infty} (fg)(x_n)\,.$$

□

Folgerung 7.10 *Alle rationalen Funktionen, d.h. alle Funktionen der Form*

$$f(x) = \frac{p(x)}{q(x)}\,,$$

wobei p und q Polynome sind, sind auf ihrem Definitionsbereich $\{x : x \in \mathbb{K}, q(x) \neq 0\}$ stetig.

Beweis: Folgt aus Satz 7.9 und der Stetigkeit der Identität sowie der konstanten Funktionen. □

Satz 7.11 *Seien $D, E \subset \mathbb{K}$, $a \in D$, $f : D \to \mathbb{K}$, $g : E \to \mathbb{K}$ und $f(D) \subset E$. Sind f stetig in a und g stetig in $f(a)$, so ist auch $g \circ f$ stetig in a.*

Beweis: Sei (x_n) eine beliebige Folge in D mit $x_n \to a$. Da f stetig ist in a, gilt

$$\lim_{n\to\infty} f(x_n) = f(a)\,.$$

Da $f(x_n) \in E$ für alle n, und da g stetig ist in $f(a)$, folgt weiter

$$\lim_{n\to\infty} g(f(x_n)) = g(f(a))\,.$$

□

Satz 7.12 (Zwischenwertsatz)
Sei $f : [a,b] \to \mathbb{R}$ stetig, es gelte $f(a)f(b) < 0$. Dann gibt es ein $x \in (a,b)$ mit $f(x) = 0$.

Beweis: Sei $f(a) < 0$. (Andernfalls betrachten wir $-f$.) Wir setzen

$$G = \{t : t \in [a,b],\ f(t) < 0\}\,.$$

Es ist $a \in G$, also $\emptyset \neq G \subset [a,b]$. Wir setzen

$$x = \sup G\,. \tag{7.7}$$

Wir wählen eine Folge (x_n) in G mit $x_n \to x$, dann ist

$$f(x_n) < 0 \text{ für alle } n \in \mathbb{N}\,, \quad \lim_{n\to\infty} f(x_n) = f(x)\,,$$

also $f(x) \le 0$. Zum Beweis der umgekehrten Ungleichung setzen wir $x_n = x + \frac{1}{n}$. Da $x < b$ wegen $f(b) > 0$, gilt $x_n \in [a,b]$ für hinreichend große n. Es ist $x_n \notin G$ für alle n nach Definition von x, also $f(x_n) \ge 0$ für alle n. Wegen $x_n \to x$ folgt $f(x) \ge 0$. □

Folgerung 7.13 *Sei $f : [a,b] \to \mathbb{R}$ stetig, sei $y \in \mathbb{R}$ mit $f(a) \le y \le f(b)$ oder $f(a) \ge y \ge f(b)$. Dann gibt es ein $x \in [a,b]$ mit $f(x) = y$.*

Beweis: Wir wenden Satz 7.12 auf die durch $g(x) = f(x) - y$ definierte Funktion $g : [a,b] \to \mathbb{R}$ an. □

Der nächste Satz zeigt, dass strikte Ungleichungen "$f(a) > y$" erhalten bleiben, wenn f stetig ist und man a nur wenig abändert.

Satz 7.14 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$, $a \in D$. Ist f stetig in a und gilt $f(a) > y$ (bzw. $f(a) < y$), so gibt es ein $\delta > 0$ mit $f(x) > y$ (bzw. $f(x) < y$) für alle $x \in D$ mit $|x - a| < \delta$.*

Beweis: Sei $f(a) > y$ (andernfalls betrachte $-f$). Falls es ein solches δ nicht gibt, können wir eine Folge (x_n) in D wählen mit $|x_n - a| \leq \frac{1}{n}$ und $f(x_n) \leq y$. Da f stetig ist und $x_n \to a$, folgt $f(x_n) \to f(a)$ und damit $f(a) \leq y$, im Widerspruch zur Voraussetzung. □

Seien $(x, f(x))$ und $(y, f(y))$ zwei Punkte auf dem Graphen einer Funktion $f : D \to \mathbb{R}$, $D \subset \mathbb{R}$. Die Gerade durch sie heißt **Sekante**. Ihre Steigung ist gegeben durch

$$\frac{f(x) - f(y)}{x - y}.$$

Definition 7.15 (Lipschitzstetigkeit)
Sei $D \subset \mathbb{K}$. Eine Funktion $f : D \to \mathbb{K}$ heißt lipschitzstetig (in D), falls es ein $L > 0$ gibt mit

$$|f(x) - f(y)| \leq L|x - y|, \quad \text{für alle } x, y \in D. \tag{7.8}$$

Ein solches L heißt Lipschitzkonstante für f in D.

Die Bedingung (7.8) bedeutet, dass alle Sekanten für f eine Steigung haben, deren Betrag kleiner oder gleich L ist.

Die Betragsfunktion $f : \mathbb{K} \to \mathbb{R}$, $f(x) = |x|$ ist lipschitzstetig, da

$$|f(x) - f(y)| = ||x| - |y|| \leq |x - y|$$

(umgekehrte Dreiecksungleichung) gilt für alle $x, y \in \mathbb{K}$. Die Zahl 1 ist eine Lipschitzkonstante für die Betragsfunktion, und zwar die kleinstmögliche.

Satz 7.16 *Jede in $D \subset \mathbb{K}$ lipschitzstetige Funktion ist stetig in D.*

Beweis: : Übungsaufgabe. □

Uneigentliche Grenzwerte von Funktionen. Wir wollen Grenzwerte von Funktionen betrachten, bei denen die Argumente oder die Werte gegen $\pm\infty$ uneigentlich konvergieren. Dabei treten eine Reihe verschiedener Fälle auf.

Definition 7.17 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$, $a \in \overline{D}$. Wir sagen, dass*

$$\lim_{x \to a} f(x) = \infty, \quad \text{bzw.} \quad \lim_{x \to a} f(x) = -\infty, \tag{7.9}$$

falls

$$\lim_{n \to \infty} f(x_n) = \infty, \quad \text{bzw.} \quad \lim_{n \to \infty} f(x_n) = -\infty$$

gilt für jede Folge (x_n) in D mit $x_n \to a$.

Wie gehabt schreiben wir

$$\lim_{\substack{x \to a \\ x \in D}} f(x) = \infty,$$

falls wir D explizit angeben wollen.

Beispiel 7.18
Wir betrachten
$$f(x) = \frac{1}{x}\,.$$
Mit $D = (0, \infty)$ gilt
$$\lim_{\substack{x \to 0 \\ x \in D}} \frac{1}{x} = \infty\,,$$
mit $D = (-\infty, 0)$ gilt
$$\lim_{\substack{x \to 0 \\ x \in D}} \frac{1}{x} = -\infty\,,$$
mit $D = \mathbb{R} \setminus \{0\}$ gilt
$$\lim_{\substack{x \to 0 \\ x \in D}} \frac{1}{x} \qquad \text{existiert nicht.}$$

Wir betrachten nun den Fall, dass die Argumente uneigentlich gegen ∞ konvergieren.

Definition 7.19 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$, es gelte $D \cap (M, \infty) \neq \emptyset$ für alle $M \in \mathbb{R}$. (Diese Voraussetzung stellt sicher, dass es eine Folge (x_n) in D gibt mit $x_n \to \infty$.)*

(i) *Wir sagen, dass für $c \in \mathbb{R}$ gilt*
$$\lim_{x \to \infty} f(x) = c\,, \tag{7.10}$$
falls
$$\lim_{n \to \infty} f(x_n) = c$$
gilt für jede Folge (x_n) in D mit $x_n \to \infty$.

(ii) *Wir sagen, dass gilt*
$$\lim_{x \to \infty} f(x) = \infty\,, \tag{7.11}$$
falls
$$\lim_{n \to \infty} f(x_n) = \infty$$
gilt für jede Folge (x_n) in D mit $x_n \to \infty$.

Im Falle (7.11) sind beide Grenzwerte ($x_n \to \infty$ und $f(x_n) \to \infty$) uneigentliche Grenzwerte.

Analog werden definiert
$$\lim_{x \to -\infty} f(x) = c\,, \quad \lim_{x \to \infty} f(x) = -\infty\,, \quad \lim_{x \to -\infty} f(x) = \infty\,, \quad \lim_{x \to -\infty} f(x) = -\infty\,.$$

Beispiel 7.20
(1) Mit $D = (0, \infty)$ gilt
$$\lim_{x \to \infty} \frac{1}{x} = 0\,, \quad \lim_{x \to \infty} x = \infty\,.$$

(2) Allgemeiner: Ist p ein Polynom der Form

$$p(x) = x^n + \sum_{k=0}^{n-1} a_k x^k \,,$$

so gilt (Übungsaufgabe)

$$\lim_{x\to\infty} p(x) = \infty \,, \qquad \lim_{x\to-\infty} p(x) = \begin{cases} \infty \,, & \text{falls } n \text{ gerade,} \\ -\infty \,, & \text{falls } n \text{ ungerade.} \end{cases}$$

Teil (1) des vorangehenden Beispiels ist auch ein Beispiel für den folgenden allgemeinen Sachverhalt.

Satz 7.21 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$, es gelte $D \cap (M, \infty) \neq \emptyset$ für alle $M \in \mathbb{R}$. Gilt*

$$\lim_{x\to\infty} f(x) = \infty \,, \tag{7.12}$$

so gilt auch

$$\lim_{x\to\infty} \frac{1}{f(x)} = 0 \,. \tag{7.13}$$

Beweis: Sei (x_n) eine Folge in D mit $x_n \to \infty$. Nach Voraussetzung gilt $f(x_n) \to \infty$. Sei nun $\varepsilon > 0$. Wir wählen $n_0 \in \mathbb{N}$ mit $f(x_n) \geq 1/\varepsilon$ für alle $n \geq n_0$. Es folgt

$$\frac{1}{f(x_n)} \leq \varepsilon \,, \quad \text{für alle } n \geq n_0.$$

Da ε beliebig war, folgt $1/f(x_n) \to 0$. Da auch die Folge (x_n) beliebig war, ist (7.13) gezeigt. □

Monotone Funktionen.

Definition 7.22 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$. f heißt **monoton wachsend**, falls $f(x) \leq f(\tilde{x})$ für alle $x, \tilde{x} \in D$ mit $x \leq \tilde{x}$. f heißt **streng monoton wachsend**, falls $f(x) < f(\tilde{x})$ für alle $x, \tilde{x} \in D$ mit $x < \tilde{x}$. f heißt **(streng) monoton fallend**, falls $-f$ (streng) monoton wachsend ist. f heißt **(streng) monoton**, falls f (streng) monoton wachsend oder (streng) monoton fallend ist.*

Satz 7.23 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ streng monoton wachsend. Dann ist $f : D \to f(D)$ bijektiv, und $f^{-1} : f(D) \to D$ ist streng monoton wachsend.*

Beweis: Sind $x, \tilde{x} \in D$ mit $x \neq \tilde{x}$, so ist $x < \tilde{x}$ oder $\tilde{x} < x$. Im ersten Fall ist $f(x) < f(\tilde{x})$, andernfalls ist $f(\tilde{x}) < f(x)$, also $f(x) \neq f(\tilde{x})$. Also ist f injektiv und damit $f : D \to f(D)$ bijektiv. Seien nun $y, \tilde{y} \in f(D)$. Ist $f^{-1}(y) \geq f^{-1}(\tilde{y})$, so ist auch

$$y = f(f^{-1}(y)) \geq f(f^{-1}(\tilde{y})) = \tilde{y} \,.$$

Aus $y < \tilde{y}$ folgt daher $f^{-1}(y) < f^{-1}(\tilde{y})$. □

Folgerung 7.24
Satz 7.23 gilt auch, wenn "streng monoton wachsend" überall durch "streng monoton fallend" ersetzt wird.

Beweis: Nach Satz 7.23 ist $-f : D \to -f(D)$ bijektiv und $(-f)^{-1} : -f(D) \to D$ streng monoton wachsend. Sind $y, \tilde{y} \in f(D)$ mit $y < \tilde{y}$, so ist $-y > -\tilde{y}$ und

$$f^{-1}(y) = (-f)^{-1}(-y) > (-f)^{-1}(-\tilde{y}) = f^{-1}(\tilde{y}) .$$

Also ist f^{-1} streng monoton fallend. □

In Kapitel 3 hatten wir die k-te Wurzel als Umkehrfunktion der k-ten Potenz

$$f : \mathbb{R}_+ \to \mathbb{R}_+ , \quad f(x) = x^k$$

erhalten. Da f streng monoton wachsend ist, existiert die Umkehrfunktion $f^{-1} : f(\mathbb{R}_+) \to \mathbb{R}_+$. Damit die k-te Wurzel tatsächlich auf ganz $\mathbb{R}_+$ definiert ist, musste gezeigt werden, dass $f(\mathbb{R}_+) = \mathbb{R}_+$ gilt, dass also jedes $y \geq 0$ als $y = x^k$ erhalten werden kann. Das war die Hauptschwierigkeit. Der Zwischenwertsatz ermöglicht es nun, für stetige Funktionen f auf explizite Konstruktionen wie im Beweis von Satz 3.1 zu verzichten. Zusätzlich erhalten wir, dass auch die Umkehrfunktion stetig ist.

Satz 7.25 *Sei $f : [a, b] \to \mathbb{R}$ streng monoton wachsend und stetig. Dann ist*

$$f([a, b]) = [f(a), f(b)] , \tag{7.14}$$

und

$$f^{-1} : [f(a), f(b)] \to [a, b] \tag{7.15}$$

ist ebenfalls streng monoton wachsend und stetig.

Beweis: Aus $a \leq x \leq b$ folgt $f(a) \leq f(x) \leq f(b)$; ist $y \in [f(a), f(b)]$, so gibt es wegen Folgerung 7.13 ein $x \in [a, b]$ mit $f(x) = y$. Damit ist (7.14) bewiesen. Nach Satz 7.23 ist f^{-1} streng monoton wachsend. Wir nehmen nun an, f^{-1} sei nicht stetig auf $[f(a), f(b)]$. Dann gibt es ein $y \in [f(a), f(b)]$, in dem f^{-1} unstetig ist; es gibt dann eine Folge (y_n) in $[f(a), f(b)]$ mit $y_n \to y$, aber $f^{-1}(y_n) \not\to f^{-1}(y)$. Also gibt es ein $\varepsilon > 0$ und eine Teilfolge (y_{n_k}) mit

$$|f^{-1}(y_{n_k}) - f^{-1}(y)| \geq \varepsilon , \quad \text{für alle } k \in \mathbb{N} . \tag{7.16}$$

Da $(f^{-1}(y_{n_k}))_{k \in \mathbb{N}}$ eine Folge in $[a, b]$ und also beschränkt ist, hat sie eine konvergente Teilfolge (eine "Teilteilfolge" der ursprünglichen Folge (y_n)), sei

$$f^{-1}(y_{n_{k_l}}) \to x , \quad x \in [a, b] .$$

Es gilt, da f stetig ist,

$$f(x) = f(\lim_{l \to \infty} f^{-1}(y_{n_{k_l}})) = \lim_{l \to \infty} f(f^{-1}(y_{n_{k_l}})) = \lim_{l \to \infty} y_{n_{k_l}} = y ,$$

also

$$f^{-1}(y_{n_{k_l}}) \to x = f^{-1}(y)$$

im Widerspruch zu (7.16). Also war die Annahme falsch, das heißt, f^{-1} ist stetig. □

Folgerung 7.26
Satz 7.25 gilt auch, wenn "streng monoton wachsend" überall durch "streng monoton fallend" ersetzt wird.

Beweis: Folgt aus Folgerung 7.24, da aus der Stetigkeit von $(-f)^{-1}$ auch die Stetigkeit von f^{-1} folgt. □

Wir wenden den Satz an auf die stetige und streng monoton wachsende Funktion

$$f : [0, M] \to [0, M^k]\,, \quad f(x) = x^k\,, \quad k \geq 1\,,$$

mit einem festen $M > 0$. Aus (7.14) folgt, dass $f([0, M]) = [0, M^k]$ gilt und damit $f(\mathbb{R}_+) = \mathbb{R}_+$, da wir M beliebig groß wählen können und $M^k \to \infty$ für $M \to \infty$ gilt. Die Umkehrfunktion

$$f^{-1}(y) = \sqrt[k]{y}$$

ist eine Funktion

$$f^{-1} : \mathbb{R}_+ \to \mathbb{R}_+\,,$$

die auf $\mathbb{R}_+$ stetig ist, da sie für beliebiges $M > 0$ gemäß des vorangehenden Satzes in jedem Punkt $y \in [0, M^k]$ stetig ist.

Die Wurzelfunktion ist ein Beispiel für eine Funktion, die auf $[0, 1]$ stetig, aber nicht lipschitzstetig ist. Für $x > 0$ gilt nämlich

$$|x - 0| = x = \sqrt{x} \cdot \sqrt{x} = \sqrt{x} \cdot |\sqrt{x} - \sqrt{0}|\,,$$

also

$$|\sqrt{x} - \sqrt{0}| = \frac{1}{\sqrt{x}} \cdot |x - 0|\,.$$

Da $1/\sqrt{x} \to \infty$ für $x \to 0$, kann es keine Zahl L geben mit

$$|\sqrt{x} - \sqrt{0}| \leq L|x - 0|\,, \quad \text{für alle } x > 0.$$

8 Exponentialfunktion, Trigonometrische Funktionen, Logarithmus

Die reelle Exponentialfunktion

$$\exp : \mathbb{R} \to \mathbb{R} \tag{8.1}$$

beschreibt Wachstumsprozesse. Wir betrachten als Beispiel eine Population, die sich pro Zeiteinheit um den Faktor x vergrößert. Ein diskretes Modell dafür ist folgendes. Sei $p_0 > 0$ die Anfangsgröße der Population zum Zeitpunkt $t_0 = 0$, sei $t_1 = t_0 + h$ mit der Zeitschrittweite $h > 0$. Wir setzen

$$p_1 = p_0 + hxp_0 = (1 + hx)p_0 \,.$$

Die weiteren Zeitschritte verlaufen genauso,

$$p_{k+1} = (1 + hx)p_k \,, \quad t_{k+1} = t_k + h \,.$$

Wir sind an der Größe der Population zum Endzeitpunkt $T = 1$ interessiert. Ist $h = 1/n$, so sind dafür n Schritte nötig, und

$$p_n = \Big(1 + \frac{x}{n}\Big)p_{n-1} = \cdots = \Big(1 + \frac{x}{n}\Big)^n p_0 \,.$$

Das Verhältnis p_n/p_0 von End- und Anfangsgröße hängt vom Wachstumsfaktor x und der Zeitschrittweite $h = 1/n$ ab und ist gegeben durch

$$f_n(x) = \Big(1 + \frac{x}{n}\Big)^n \,. \tag{8.2}$$

Ist $x > 0$, so wächst die Population, ist $x < 0$, so schrumpft die Population. (Im Fall $x < 0$ macht das Modell nur Sinn, wenn $n > |x|$ ist; nur dann ist $p_k > 0$ für alle k.) Hier ein Zahlenbeispiel: Sei $x = 12$. Für $n = 2$ ist $h = 1/2$, $1 + hx = 7$ und $f_2(x) = 7^2 = 49$. Analog erhält man

$$f_4(12) = 4^4 = 256 \,, \quad f_{12}(12) = 2^{12} = 4096 \,, \quad f_{48}(12) = \Big(\frac{5}{4}\Big)^{48} \approx 44841.55 \,.$$

Aus diesem diskreten Wachstumsmodell erhält man ein kontinuierliches Wachstumsmodell, indem man den Grenzübergang $h \to 0$ bzw. $n \to \infty$ durchführt. Das Verhältnis von End- und Anfangsgröße hängt dann nur noch vom Wachstumsfaktor x ab und ist gegeben durch

$$\lim_{n\to\infty} \Big(1 + \frac{x}{n}\Big)^n \,, \tag{8.3}$$

falls dieser Grenzwert existiert. Das ist, wie wir später sehen werden, tatsächlich für beliebiges $x \in \mathbb{R}$ der Fall; im obigen Beispiel ist er ungefähr gleich 162754.79. Er stellt eine Möglichkeit dar, die Exponentialfunktion durch

$$\exp(x) = \lim_{n\to\infty} \Big(1 + \frac{x}{n}\Big)^n \tag{8.4}$$

zu definieren. Eine andere Möglichkeit (und das ist die, die wir hier verwenden) ist die Definition durch die Exponentialreihe

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \,. \tag{8.5}$$

Wir werden später sehen, dass beide Definitionen zum gleichen Ergebnis führen, das heißt, dass

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

gilt.

Gemäß Beispiel 6.16 konvergiert die Exponentialreihe absolut, und zwar auch für komplexe Argumente. Diese sind ebenfalls von Interesse, da mit ihnen Schwingungsphänomene beschrieben werden können.

Definition 8.1 (Exponentialfunktion)
Die durch

$$\exp(z) = \sum_{k=0}^{\infty}\frac{z^k}{k!} \tag{8.6}$$

definierte Funktion

$$\exp : \mathbb{C} \to \mathbb{C}$$

heißt Exponentialfunktion. Wir definieren die Zahl e ("Eulersche Zahl") durch

$$e = \exp(1)\,. \tag{8.7}$$

□

Aus der Definition folgt unmittelbar, dass $\exp(z) \in \mathbb{R}$, falls $z \in \mathbb{R}$, und

$$\exp(0) = 1\,.$$

Es ist

$$e = 2.718281828459045235\ldots$$

In Satz 6.20 haben wir das Produkt zweier Exponentialreihen berechnet,

$$\sum_{k=0}^{\infty}\frac{(z+w)^k}{k!} = \left(\sum_{k=0}^{\infty}\frac{z^k}{k!}\right)\cdot\left(\sum_{k=0}^{\infty}\frac{w^k}{k!}\right), \qquad z,w\in\mathbb{C}\,.$$

Das ist nichts anderes als die **Funktionalgleichung der Exponentialfunktion**

$$\exp(z+w) = \exp(z)\exp(w)\,, \tag{8.8}$$

welche für alle $z, w \in \mathbb{C}$ gilt. Aus ihr erhalten wir unmittelbar eine Reihe von Aussagen über die Exponentialfunktion. Es gilt

$$\exp(-z) = \frac{1}{\exp(z)}\,, \qquad \text{für alle } z\in\mathbb{C}\,, \tag{8.9}$$

da

$$\exp(z)\exp(-z) = \exp(z-z) = \exp(0) = 1\,.$$

Insbesondere ist

$$\exp(z) \neq 0\,, \qquad \text{für alle } z\in\mathbb{C}\,. \tag{8.10}$$

Weiter gilt

$$\exp(\overline{z}) = \overline{\exp(z)}\,, \qquad \text{für alle } z \in \mathbb{C}\,, \tag{8.11}$$

da (wir können Konjugation und Grenzwerte gemäß (5.39) vertauschen)

$$\exp(\overline{z}) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{\overline{z}^k}{k!} = \lim_{n\to\infty} \overline{\sum_{k=0}^{n} \frac{z^k}{k!}} = \overline{\exp(z)}\,.$$

Die Exponentialfunktion im Reellen. Für Argumente $x \in \mathbb{R}$ sind alle Glieder der Exponentialreihe

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

reell, also ist auch $\exp(x) \in \mathbb{R}$. Für $n \in \mathbb{N}$ gilt

$$\exp(n) = \exp\left(\sum_{k=1}^{n} 1\right) = \prod_{k=1}^{n} \exp(1) = e^n\,,$$

und mit (8.9) folgt

$$\exp(n) = e^n\,, \qquad \text{für alle } n \in \mathbb{Z}\,. \tag{8.12}$$

Analog gilt

$$\exp(1) = \exp\Big(\sum_{k=1}^{n} \frac{1}{n}\Big) = \prod_{k=1}^{n} \exp\Big(\frac{1}{n}\Big) = \Big(\exp\Big(\frac{1}{n}\Big)\Big)^n\,,$$

also

$$\exp\left(\frac{1}{n}\right) = \sqrt[n]{e}\,, \quad \text{für alle } n \in \mathbb{N}. \tag{8.13}$$

Aus der Exponentialreihe erkennt man, dass $\exp(x) > \exp(0) = 1$ für $x > 0$. Da $\exp(-x) = 1/\exp(x)$, gilt

$$\begin{aligned} \exp(x) &> 1\,, \quad \text{falls } x > 0, \\ 0 < \exp(x) &< 1\,, \quad \text{falls } x < 0, \end{aligned} \tag{8.14}$$

und insgesamt

$$\exp(x) > 0\,, \qquad \text{für alle } x \in \mathbb{R}\,. \tag{8.15}$$

Hieraus folgt, dass die reelle Exponentialfunktion auf $\mathbb{R}$ streng monoton wachsend ist, denn $x < \tilde{x}$ impliziert

$$\exp(x) = \exp(x - \tilde{x})\exp(\tilde{x}) < \exp(\tilde{x})\,.$$

Weiterhin gilt

$$\lim_{x\to\infty} \exp(x) = \infty\,, \tag{8.16}$$

da $\exp(x) \geq x$ gemäß Exponentialreihe und $x \to \infty$ für $x \to \infty$. Da außerdem $\exp(x) \geq x^{k+1}/(k+1)!$, gilt sogar

$$\lim_{x\to\infty} \frac{\exp(x)}{x^k} = \infty\,, \quad \text{für alle } k \in \mathbb{N}, \tag{8.17}$$

das heißt, die Exponentialfunktion wächst schneller als jede Potenz von x für $x \to \infty$. Gemäß Satz 7.21 folgt

$$\lim_{x\to\infty} \frac{x^k}{\exp(x)} = 0\,, \quad \text{für alle } k \in \mathbb{N}. \tag{8.18}$$

Da $\exp(-x) = 1/\exp(x)$, folgt weiter

$$0 = \lim_{x\to-\infty} |x|^k \exp(x) = \lim_{x\to-\infty} x^k \exp(x)\,, \quad \text{für alle } k \in \mathbb{N}. \tag{8.19}$$

Das Wachstum der Exponentialfunktion ist sehr schnell, beispielsweise ist $\exp(100)$ ungefähr gleich $2.69 \cdot 10^{43}$ und $\exp(-100)$ ungefähr gleich $3.72 \cdot 10^{-44}$. Letzteres ist ziemlich verblüffend, wenn man direkt die Definition

$$\exp(-100) = \sum_{k=0}^{\infty} (-1)^k \frac{100^k}{k!}$$

betrachtet, denn die einzelnen Summanden wachsen zunächst sehr stark an.

Die Exponentialfunktion im Komplexen. Trigonometrische Funktionen. Wegen

$$\exp(n) = e^n\,, \quad n \in \mathbb{Z}\,,$$

können wir (für die Basis e) die Exponentialfunktion als Verallgemeinerung der Potenzfunktion auffassen.

Notation 8.2 *Wir vereinbaren*

$$e^z = \exp(z)\,, \quad z \in \mathbb{C}\,. \tag{8.20}$$

Die Funktionalgleichung (8.8) wird dann zu

$$e^{z+w} = e^z \cdot e^w\,, \quad z, w \in \mathbb{C}\,. \tag{8.21}$$

Für $x \in \mathbb{R}$ betrachten wir nun die komplexe Zahl e^{ix}. Es gilt $\overline{e^{ix}} = e^{\overline{ix}} = e^{-ix}$, also

$$|e^{ix}|^2 = e^{ix}\overline{e^{ix}} = e^{ix}e^{-ix} = e^{ix-ix} = e^0 = 1\,, \tag{8.22}$$

also

$$|e^{ix}| = 1\,, \tag{8.23}$$

das heißt, e^{ix} liegt auf dem Einheitskreis in der komplexen Ebene.

Definition 8.3 (Sinus und Cosinus)
Wir definieren die Funktionen $\cos, \sin : \mathbb{R} \to \mathbb{R}$ *durch*

$$\cos x = \operatorname{Re} e^{ix}\,, \quad \sin x = \operatorname{Im} e^{ix}\,. \tag{8.24}$$

Äquivalent zu (8.24) ist die **Eulergleichung**

$$e^{ix} = \cos x + i\,\sin x\,, \quad x \in \mathbb{R}\,. \tag{8.25}$$

Unmittelbar aus der Definition und aus den elementaren Eigenschaften komplexer Zahlen folgt für alle $x \in \mathbb{R}$

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}), \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}), \tag{8.26}$$

$$\cos(-x) = \cos x, \quad \sin(-x) = -\sin x, \tag{8.27}$$

$$\cos^2 x + \sin^2 x = 1, \tag{8.28}$$

$$\cos 0 = 1, \quad \sin 0 = 0. \tag{8.29}$$

Bemerkung 8.4 (Geometrische Definition von Sinus und Cosinus)
Die geometrische Definition von Sinus und Cosinus geht so: Ist z ein Punkt auf dem Einheitskreis, welcher mit der reellen Achse den Winkel x bildet, so ist

$$\cos x = \operatorname{Re} z, \quad \sin x = \operatorname{Im} z.$$

Für den Winkel x sind zwei Maßeinheiten gebräuchlich. Zum einen in Grad, wobei 360 Grad dem Vollkreis entspricht, ein rechter Winkel also 90 Grad hat. Zum anderen im **Bogenmaß**, das heißt, x ist die Länge des Kreisbogens vom Punkt 1 nach z, durchlaufen im **mathematisch positiven Sinn** (gegen den Uhrzeiger). Im Bogenmaß hat der Vollkreis den Winkel 2π, ein rechter Winkel hat den Winkel $\pi/2$. Hierbei ist die Zahl π geometrisch definiert als das Verhältnis von Kreisumfang zu Kreisdurchmesser.

Die analytische Definition in 8.3 und die geometrische Definition von Sinus und Cosinus führen zum selben Ergebnis, falls die Zahl $z = e^{ix}$ dem Winkel x entspricht, das heißt, falls die Länge des Kreisbogens von 1 nach e^{ix} gerade x beträgt. Das werden wir durch die folgende Grenzwertbetrachtung plausibel machen. (Die exakte Definition der Länge einer Kurve werden wir im zweiten Semester behandeln.)

Sei $z = e^{ix}$ mit $x > 0$ ein Punkt auf dem Einheitskreis. Zusammen mit den Punkten 0 und 1 bildet er ein Dreieck, die 1 und z verbindende Seite ist eine Sehne des Einheitskreises. Wir nennen sie S_0, sie hat die Länge

$$s_0 = |e^{ix} - 1|.$$

Ihr Mittelpunkt ist

$$w = \frac{1}{2}(e^{ix} + 1).$$

Die Gerade durch 0 und w ist eine Winkelhalbierende des Dreiecks. Es gilt

$$w = e^{ix/2}e^{-ix/2} \cdot \frac{1}{2}(e^{ix} + 1) = e^{ix/2} \cdot \frac{1}{2}(e^{ix/2} + e^{-ix/2}) = e^{ix/2}\operatorname{Re}\left(e^{ix/2}\right).$$

Also ist $e^{ix/2}$ ein skalares Vielfaches von w und liegt daher ebenfalls auf der Winkelhalbierenden. Damit teilt $e^{ix/2}$ den Kreisbogen von 1 nach e^{ix} in zwei gleiche Teile, und die Sehnen von 1 nach $e^{ix/2}$ und von $e^{ix/2}$ nach e^{ix} sind ebenfalls gleichlang. Letzteres erkennt man auch unmittelbar an der Rechnung

$$|e^{ix} - e^{ix/2}| = |e^{ix/2}||e^{ix/2} - 1| = |e^{ix/2} - 1|.$$

Sei nun S_1 der Streckenzug, der aus diesen beiden Sehnen gebildet wird. Seine Länge ist

$$s_1 = 2|e^{ix/2} - 1| \,.$$

Eine weitere Halbierung der beiden Sehnen führt auf einen Streckenzug S_2, welcher die Punkte 1, $e^{ix/4}$, $e^{ix/2}$, $e^{3ix/4}$ und e^{ix} verbindet. Seine Länge ergibt sich analog zu

$$s_2 = 4|e^{ix/4} - 1| \,.$$

Nach k Halbierungen ergibt sich der Streckenzug S_k mit der Länge

$$s_k = 2^k|e^{2^{-k}ix} - 1| = x \cdot \left| \frac{e^{2^{-k}ix} - 1}{2^{-k}x} \right| . \tag{8.30}$$

Wir werden später in (9.7) sehen, dass der durch den Betrag rechts gegebene Ausdruck gegen 1 konvergiert. Es folgt

$$\lim_{k\to\infty} s_k = x \,,$$

das heißt, die Längen der den Kreisbogen approximierenden Streckenzüge S_k konvergieren gegen x. □

Aus der Eulerformel erhalten wir unmittelbar Rechenregeln für Sinus und Cosinus.

Lemma 8.5 *Für alle $x, y \in \mathbb{R}$ gilt*

$$\begin{aligned} \cos(x + y) &= \cos(x)\cos(y) - \sin(x)\sin(y)\,, \\ \sin(x + y) &= \sin(x)\cos(y) + \sin(y)\cos(x)\,, \end{aligned} \tag{8.31}$$

Beweis: Es ist

$$e^{i(x+y)} = e^{ix}e^{iy} \,,$$

also

$$\begin{aligned} \cos(x + y) + i\,\sin(x + y) &= (\cos x + i\,\sin x)(\cos y + i\,\sin y) \\ &= \cos(x)\cos(y) - \sin(x)\sin(y) + i(\sin(x)\cos(y) + \cos(x)\sin(y))\,. \end{aligned}$$

Vergleich der Real- und Imaginärteile liefert die Behauptung. □

Für $x = y$ werden die Formeln (8.31) zu

$$\cos(2x) = \cos^2(x) - \sin^2(x)\,, \quad \sin(2x) = 2\sin(x)\cos(x)\,. \tag{8.32}$$

Wir wollen klären, dass exp, sin und cos stetige Funktionen sind.

Lemma 8.6 *Es gilt*

$$\lim_{z\to 0} e^z = 1\,. \tag{8.33}$$

Der Grenzwert ist als Grenzwert in $\mathbb{C}$ zu verstehen.

Beweis: Es ist

$$e^z - 1 = \sum_{k=1}^{\infty} \frac{z^k}{k!} = z \sum_{k=1}^{\infty} \frac{z^{k-1}}{k!} \,.$$

Für $|z| < 1$ folgt, da $k! \geq 2^{k-1}$,

$$0 \leq |e^z - 1| \leq |z| \sum_{k=1}^{\infty} \frac{1}{k!} \leq |z| \sum_{k=1}^{\infty} 2^{1-k} = 2|z| \quad \to \quad 0 \qquad \text{für } z \to 0,$$

und damit die Behauptung. □

Satz 8.7 *Die Exponentialfunktion ist stetig in $\mathbb{C}$, Sinus und Cosinus sind stetig in $\mathbb{R}$.*

Beweis: Wir betrachten $\exp : \mathbb{C} \to \mathbb{C}$. Für jedes $a \in \mathbb{C}$ gilt

$$|e^z - e^a| = |e^a(e^{z-a} - 1)| = |e^a| \cdot |e^{z-a} - 1| \quad \to \quad 0 \qquad \text{für } z \to a$$

nach Lemma 8.6, und daher

$$\lim_{z \to a} e^z = e^a \,.$$

Die Exponentialfunktion ist also stetig in $\mathbb{C}$. Der Sinus $\sin : \mathbb{R} \to \mathbb{R}$ entsteht als Komposition stetiger Funktionen,

$$x \;\mapsto\; ix \;\mapsto\; e^{ix} \;\mapsto\; \operatorname{Im} e^{ix}$$

und ist daher stetig, analog der Cosinus. □

Die in Bemerkung 8.4 erläuterte geometrische Interpretation besagte, dass e^{ix} der Endpunkt des in 1 beginnenden Kreisbogens der Länge x ist. Lässt man x wachsen, so durchläuft e^{ix} den Einheitskreis, Sinus und Cosinus oszillieren periodisch zwischen 1 und -1. Wir machen uns nun dieses Verhalten auch anhand der analytischen Definition 8.3 klar und geben dabei eine alternative Definition der Zahl π.

Zunächst betrachten wir Reihenentwicklungen für Sinus und Cosinus. Es ist für $x \in \mathbb{R}$

$$\begin{aligned} e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} &= 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots \\ &= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \cdots \end{aligned} \tag{8.34}$$

Gemäß Satz 6.10 sind die Reihen der Real- und Imaginärteile ebenfalls absolut konvergent, und zwar gegen Real- bzw. Imaginärteil der Summe. Es folgt

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \tag{8.35}$$

und

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \,. \tag{8.36}$$

Diese Reihenentwicklungen sind "unanschaulich" (das heißt, ohne Verbindung zur geometrischen Vorstellung am Einheitskreis), aber für die mathematische Behandlung von Sinus und Cosinus sehr nützlich.

Beide Reihen sind alternierend. Da sie konvergent sind, konvergieren die Folgen ihrer Glieder gegen 0. Für die Sinusreihe gilt, falls $0 < x \leq 1$, dass ihre Glieder (ohne Vorzeichen) eine monoton fallende Folge bilden. Wie beim Leibnizkriterium erläutert wird ihr Grenzwert von den Partialsummen eingeschachtelt. Insbesondere gilt

$$0 < x - \frac{x^3}{3!} = x\big(1 - \frac{x^2}{2}\big) \leq \sin(x)\,, \quad \text{falls } 0 < x \leq 1. \tag{8.37}$$

Daran erkennen wir, dass die Funktion $x \mapsto e^{ix}$ den Einheitskreis gegen den Uhrzeigersinn durchläuft, wenn x von 0 ausgehend monoton wächst.

Wir betrachten die Cosinusreihe für $0 < x \leq 2$. Ihre Glieder (ohne Vorzeichen) sind ab dem zweiten Glied $x^2/2!$ monoton fallend, da für $k \geq 1$ der Quotient zweier aufeinanderfolgender Glieder

$$\frac{x^{2k+2}}{(2k+2)!} \cdot \frac{(2k)!}{x^{2k}} = \frac{x^2}{(2k+2)(2k+1)} \leq \frac{4}{4 \cdot 3} = \frac{1}{3}$$

erfüllt. Die Einschachtelung des Grenzwerts gemäß Leibnizkriterium impliziert nun

$$1 - \frac{x^2}{2!} \leq \cos(x) \leq 1 - \frac{x^2}{2!} + \frac{x^4}{4!}\,.$$

Für $x = 2$ bedeutet das

$$-1 \leq \cos(2) \leq -1 + \frac{2}{3} = -\frac{1}{3} < 0\,. \tag{8.38}$$

Da $\cos(0) = 1$ und der Cosinus stetig ist, muss es nach dem Zwischenwertsatz eine Nullstelle $x \in [0, 2]$ geben. Für sie gilt $\sin(x) = 1$ und daher

$$e^{ix} = 0 + i \cdot 1 = i\,.$$

Das entspricht dem oberen Punkt auf dem Einheitskreis. Lassen wir x von 0 ausgehend monoton wachsen, so durchläuft e^{ix} den Einheitskreis stetig im mathematisch positiven Sinn. Der obere Punkt i des Einheitskreises wird erreicht, wenn der Cosinus zum erstenmal 0 wird. Der zugehörige Kreisbogen ist ein Viertel des Einheitskreises. Im Bogenmaß hat der zugehörige rechte Winkel den Wert $\pi/2$ gemäß der geometrischen Definition von π.

Definition 8.8 (Analytische Definition von π)
Wir definieren

$$\pi = 2 \cdot \inf\{x : x \geq 0,\, \cos x = 0\}\,. \tag{8.39}$$

Damit wird erreicht, dass die analytische Definition den "richtigen" Wert von π liefert. Nach dem Gesagten gilt

$$\cos\frac{\pi}{2} = 0\,, \tag{8.40}$$

also

$$\sin\frac{\pi}{2} = 1\,. \tag{8.41}$$

Aus (8.40) und (8.41) folgt

$$e^{i\frac{\pi}{2}} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i\,, \tag{8.42}$$

$$e^{i\pi} = \left(e^{i\frac{\pi}{2}}\right)^2 = -1 = \cos\pi + i\sin\pi\,, \tag{8.43}$$

also

$$\cos\pi = -1\,, \quad \sin\pi = 0\,, \tag{8.44}$$

und weiter

$$e^{i\frac{3\pi}{2}} = -i\,, \quad e^{2\pi i} = 1\,. \tag{8.45}$$

Aus Lemma 8.5 folgt für $x \in \mathbb{R}$

$$\cos(x+\pi) = \cos(x)\cos(\pi) - \sin(x)\sin(\pi) = -\cos x\,, \tag{8.46}$$

analog

$$\sin(x+\pi) = -\sin x\,, \tag{8.47}$$

und weiter die Periodizitätseigenschaften

$$\cos(x+2\pi) = \cos x\,, \quad \sin(x+2\pi) = \sin x\,. \tag{8.48}$$

Wegen

$$e^{i(\pi/2-x)} = e^{i\pi/2}e^{-ix} = i(\cos(-x) + i\sin(-x)) = \sin x + i\cos x$$

gilt

$$\cos x = \sin(\frac{\pi}{2} - x)\,, \quad \sin x = \cos(\frac{\pi}{2} - x)\,. \tag{8.49}$$

Berücksichtigen wir außerdem (8.46) und (8.47), so sehen wir, dass die Funktionen $\cos, \sin : \mathbb{R} \to \mathbb{R}$ bereits durch ihre Werte auf $[0, \frac{\pi}{2}]$ festgelegt sind. Insbesondere sind ihre Nullstellen gegeben durch

$$\sin x = 0 \quad \Leftrightarrow \quad x = k\pi\,,\ k \in \mathbb{Z}\,, \qquad \cos x = 0 \quad \Leftrightarrow \quad x = \frac{\pi}{2} + k\pi\,,\ k \in \mathbb{Z}\,. \tag{8.50}$$

Definition 8.9 (Tangens und Cotangens)
Wir definieren die Funktionen

$$\tan : \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi : k \in \mathbb{Z}\} \to \mathbb{R}\,, \quad \cot : \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\} \to \mathbb{R}\,, \tag{8.51}$$

durch

$$\tan x = \frac{\sin x}{\cos x}\,, \quad \cot x = \frac{\cos x}{\sin x}\,. \tag{8.52}$$

Es gilt wegen (8.46) und (8.47)

$$\tan(x+\pi) = \tan x\,, \quad \cot(x+\pi) = \cot x\,. \tag{8.53}$$

Wir können den Tangens geometrisch am Einheitskreis interpretieren. Verlängern wir zu gegebenem $x > 0$ die Gerade durch 0 und e^{ix} bis zu ihrem Schnittpunkt p mit der Senkrechten durch den Punkt 1, so beträgt der Abstand von 1 und p gerade $\tan x$. Ist $x < 0$, so liegt P unterhalb der reellen Achse, und $\tan x$ ist gleich $-|1-p|$. An diesem Bild erkennen wir, dass der Tangens auf dem Intervall $(-\pi/2, \pi/2)$ streng monoton wachsend ist; wir werden das später (wie in dieser Vorlesung üblich) mit Mitteln der Analysis beweisen, und zwar durch Differenzieren. Weiterhin gilt

$$\lim_{x \to -\pi/2} \tan x = -\infty\,, \quad \lim_{x \to \pi/2} \tan x = \infty\,,$$

da der Cosinus Nullstellen in $\pm\pi/2$ hat und $\sin(\pm\pi/2) = \pm 1$ gilt. Analog gilt, dass der Cotangens auf $(0, \pi)$ streng monoton fallend ist mit

$$\lim_{x \to -\pi/2} \tan x = \infty\,, \quad \lim_{x \to \pi/2} \tan x = -\infty\,.$$

Hyperbolischer Sinus und Cosinus. Sie sind definiert durch

$$\sinh(x) = \frac{e^x - e^{-x}}{2}\,, \quad \cosh(x) = \frac{e^x + e^{-x}}{2}\,. \tag{8.54}$$

Polarkoordinaten. Zur Darstellung von Punkten der Ebene sind Polarkoordinaten eine Alternative zu den rechtwinkligen kartesischen Koordinaten. Sei $z = x + iy$, $z \neq 0$, ein Punkt in der komplexen Ebene. Die von 0 ausgehende Halbgerade durch z schneidet den Einheitskreis in einem Punkt $e^{i\varphi}$, wobei φ der im Bogenmaß gemessene Winkel zwischen der Halbgerade und der positiven reellen Achse ist. Es gilt dann

$$z = re^{i\varphi} \tag{8.55}$$

mit $r = |z|$. Die Zahlen r und φ heißen die Polarkoordinaten von $z = re^{i\varphi}$. Wegen

$$e^{i\varphi} = e^{i(\varphi + 2\pi k)}\,, \quad k \in \mathbb{Z}\,,$$

ist der Winkel zunächst nur bis auf ganzzahlige Vielfache von 2π eindeutig bestimmt. Die Darstellung wird eindeutig, wenn wir verlangen, dass $\varphi \in [0, 2\pi)$ gelten muss. So werden wir in dieser Vorlesung verfahren. Eine andere übliche Normierung ist $\varphi \in (-\pi, \pi]$.

Beispiele:

Punkt	1	$1+i$	i	-1	$-i$
Kartesisch	$(1,0)$	$(1,1)$	$(0,1)$	$(-1,0)$	$(0,-1)$
Polar	$(1,0)$	$(\sqrt{2}, \pi/4)$	$(1, \pi/2)$	$(1, -\pi)$	$(1, -3\pi/2)$

Gemäß (8.55) erhalten wir die kartesischen Koordinaten aus den Polarkoordinaten durch

$$\begin{aligned} x &= r\cos\varphi\,, \\ y &= r\sin\varphi\,. \end{aligned} \tag{8.56}$$

Lösen wir (8.56) nach r und φ auf, so erhalten wir

$$r = \sqrt{x^2 + y^2}\,, \tag{8.57}$$

und aus

$$\frac{y}{x} = \frac{r\sin\varphi}{r\cos\varphi} = \tan\varphi$$

ergibt sich

$$\varphi = \tan^{-1}\left(\frac{y}{x}\right). \tag{8.58}$$

(Die Umkehrfunktion des Tangens – den Arcustangens – werden wir später behandeln.)

Mit Hilfe der Polarkoordinaten können wir die Multiplikation zweier komplexer Zahlen geometrisch veranschaulichen. Ist

$$z = re^{i\varphi}, \quad w = se^{i\psi},$$

so gilt

$$z \cdot w = rse^{i\varphi}e^{i\psi} = rse^{i(\varphi+\psi)}.$$

Es werden also die Längen multipliziert und die Winkel addiert.

Der Logarithmus. Wir haben bereits gesehen, das die reelle Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ stetig und streng monoton wachsend ist. Gemäß Satz 7.25 gilt

$$\exp([-x, x]) = [\exp(-x), \exp(x)], \quad \text{für alle } x > 0.$$

Wegen

$$\lim_{x\to-\infty} e^x = 0, \quad \lim_{x\to\infty} e^x = \infty, \tag{8.59}$$

gilt $\exp(\mathbb{R}) = (0, \infty)$, und aus der strengen Monotonie folgt:

$$\exp : \mathbb{R} \to (0, \infty) \quad \text{ist bijektiv.} \tag{8.60}$$

Definition 8.10 (Logarithmus)
Wir definieren den Logarithmus

$$\ln : (0, \infty) \to \mathbb{R}$$

als die Umkehrfunktion von $\exp : \mathbb{R} \to (0, \infty)$.

Aus Satz 7.25 folgt unmittelbar:

Satz 8.11 *Der Logarithmus* $\ln : (0, \infty) \to \mathbb{R}$ *ist stetig, streng monoton wachsend und bijektiv.* □

Aus (8.59) folgt

$$\lim_{\substack{x\to 0\\ x>0}} \ln x = -\infty, \quad \lim_{x\to\infty} \ln x = \infty. \tag{8.61}$$

Die Rechenregeln für den Logarithmus erhalten wir aus der Funktionalgleichung für die Exponentialfunktion. Wir erinnern an die Definition

$$x^k = \prod_{i=1}^{k} x, \quad x^0 = 1, \quad x^{-k} = \prod_{i=1}^{k} x^{-1}, \quad \text{falls } k \in \mathbb{N},\, x \in \mathbb{R}.$$

Satz 8.12 *Für den Logarithmus* $\ln : (0, \infty) \to \mathbb{R}$ *gilt*

$$\ln(1) = 0\,, \quad \ln(e) = 1\,, \tag{8.62}$$

$$\ln(xy) = \ln x + \ln y\,, \quad \textit{für alle } x, y > 0, \tag{8.63}$$

$$\ln x^k = k \ln x\,, \quad \textit{für alle } k \in \mathbb{Z},\ x > 0, \tag{8.64}$$

Beweis: Da $e^0 = 1$ und $e^1 = e$, folgt (8.62). Weiter gilt für alle $x, y \in \mathbb{R}$

$$\begin{aligned} \exp(\ln x + \ln y) &= \exp(\ln x) \cdot \exp(\ln y) = xy\,, \\ \ln(xy) &= \ln(\exp(\ln x + \ln y)) = \ln x + \ln y\,. \end{aligned}$$

Wir beweisen (8.64). Zunächst ist $\ln x^0 = \ln 1 = 0 = 0 \ln x$. Für $k \geq 1$ gilt

$$\exp(k \ln x) = \exp\Big(\sum_{i=1}^{k} \ln x\Big) = \prod_{i=1}^{k} \exp(\ln x) = x^k\,,$$

also

$$k \ln x = \ln(\exp(k \ln x)) = \ln x^k\,.$$

Weiter gilt

$$0 = \ln 1 = \ln(x x^{-1}) = \ln x + \ln x^{-1}\,,$$

also

$$\ln\left(\frac{1}{x}\right) = -\ln x\,,$$

also für $k \geq 1$

$$\ln x^{-k} = \ln\left(\left(\frac{1}{x}\right)^k\right) = k \ln\left(\frac{1}{x}\right) = -k \ln x\,.$$

□

Die allgemeine Potenzfunktion. Bisher kennen wir die Potenzen

$$x^a$$

für ganzzahlige Exponenten a und beliebige $x \in \mathbb{R}$. Schränken wir uns ein auf $x > 0$, so können wir x^a für beliebige $a \in \mathbb{R}$ so definieren, dass die üblichen Rechenregeln gültig bleiben.

Definition 8.13 (Allgemeine Potenzfunktion)
Für $a \in \mathbb{R}$, $x > 0$ definieren wir die a-te Potenz von x durch

$$x^a = \exp(a \ln x)\,. \tag{8.65}$$

Satz 8.14 *Sei $a \in \mathbb{R}$. Die durch*

$$f(x) = x^a$$

definierte Funktion $f : (0, \infty) \to (0, \infty)$ ist stetig, für $a \neq 0$ bijektiv, für $a > 0$ streng monoton wachsend und für $a < 0$ streng monoton fallend.

Beweis: f entsteht als Komposition

$$(0,\infty) \xrightarrow{\ln} \mathbb{R} \xrightarrow{p_a} \mathbb{R} \xrightarrow{\exp} (0,\infty)\,, \tag{8.66}$$

wobei p_a definiert ist durch $p_a(y) = ay$. Alle drei Abbildungen sind mit den in (8.66) genannten Definitions- und Wertebereichen stetig und bijektiv, dasselbe gilt für deren Komposition. Exponentialfunktion und Logarithmus sind streng monoton wachsend, die Abbildung p_a ist streng monoton wachsend für $a > 0$ und streng monoton fallend für $a < 0$. Daraus folgt die Behauptung über die Monotonie von f. $\square$

Für die allgemeine Potenzfunktion gelten die üblichen Rechenregeln. Seien $a, b \in \mathbb{R}$ und $x, y > 0$. Es gelten

$$\begin{aligned}
x^{a+b} &= \exp((a+b)\ln x) = \exp(a\ln x)\cdot\exp(b\ln x) = x^a x^b\,,\\
(x^a)^b &= \exp(b\ln x^a) = \exp(b\ln(\exp(a\ln x))) = \exp(ba\ln x) = x^{ab}\,,\\
(xy)^a &= \exp(a\ln(xy)) = \exp(a(\ln x + \ln y)) = \exp(a\ln x)\cdot\exp(a\ln y) = x^a y^a\,.
\end{aligned}$$

Für $n \in \mathbb{N}$ mit $n \geq 1$ folgt hieraus

$$\left(x^{1/n}\right)^n = x^{(1/n)\cdot n} = x^1 = x\,,$$

also

$$x^{\frac{1}{n}} = \sqrt[n]{x}\,. \tag{8.67}$$

Würde man (8.67) als Definition von $x^{1/n}$ verwenden anstatt (8.65), so könnte man vermittels

$$x^{m/n} = (\sqrt[n]{x})^m$$

die Potenz x^a für rationale Zahlen $a = m/n$ definieren, ohne die Exponentialfunktion und den Logarithmus zu verwenden.

9 Differenzierbarkeit

Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}$ und $a \in D$ gegeben. Wir wollen f in der Nähe von a durch eine Gerade approximieren. Einen naheliegenden Kandidaten für eine gute Approximation stellt die Tangente an den Graphen von f durch den Punkt $(a, f(a))$ dar. Wir erhalten die Tangente geometrisch aus der Sekante an f durch die Punkte $(a, f(a))$ und $(a+h, f(a+h))$, indem wir den Grenzübergang $h \to 0$ vornehmen. Die Sekante kann beschrieben werden durch die Geradengleichung

$$g_h(x) = f(a) + \frac{f(a+h) - f(a)}{h}(x-a)\,, \tag{9.1}$$

Wir betrachten den **Differenzenquotienten**

$$d_a(h) = \frac{f(a+h) - f(a)}{h}\,, \tag{9.2}$$

welcher die Steigung der Sekante g_h angibt, und wollen den Differenzialquotienten (welcher die Steigung der Tangente angibt) als den Grenzwert

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h) - f(a)}{h} \tag{9.3}$$

erhalten. Damit dieser Grenzwert Sinn macht, muss $a + h$ in D liegen, falls h klein ist. Das ist dann der Fall, wenn a ein **innerer Punkt** von D ist, das heißt, wenn es ein $\delta > 0$ gibt mit

$$(a - \delta, a + \delta) \subset D\,.$$

Die Menge aller inneren Punkte von D heißt das **Innere** von D, bezeichnet mit $\operatorname{int}(D)$. Ist D ein Intervall, so gehören nur dessen Randpunkte nicht zu $\operatorname{int}(D)$.

Eine Teilmenge D von $\mathbb{R}$ heißt **offen**, falls jeder Punkt von D ein innerer Punkt von D ist, also $\operatorname{int}(D) = D$. Jedes offene Intervall ist offen, und auch jede Vereinigung von offenen Intervallen ist offen. So ist beispielsweise $\mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$ offen.

Definition 9.1 (Differenzierbare Funktion)
Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$, sei $a \in \operatorname{int}(D)$. Wir sagen, dass f differenzierbar in $a \in D$ ist, falls der Grenzwert

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h) - f(a)}{h} \tag{9.4}$$

existiert. In diesem Fall definieren wir die Ableitung $f'(a)$ von f in a durch

$$f'(a) = \lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h) - f(a)}{h}\,. \tag{9.5}$$

Ist D offen, so heißt f in D differenzierbar, falls f in jedem Punkt $a \in D$ differenzierbar ist.

Statt (9.4) schreibt man oft

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \tag{9.6}$$

und setzt stillschweigend voraus, dass $h = 0$ ausgeschlossen ist. So werden wir es im Folgenden handhaben.

Beispiel 9.2

1. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$, $c \in \mathbb{R}$ fest. Für alle $a \in \mathbb{R}$ gilt
$$\frac{f(a+h) - f(a)}{h} = \frac{c-c}{h} = 0\,,$$
also
$$f'(a) = 0\,.$$

2. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x$. Für alle $a \in \mathbb{R}$ gilt
$$\frac{f(a+h) - f(a)}{h} = \frac{a+h-a}{h} = 1\,,$$
also
$$f'(a) = 1\,.$$

3. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Für alle $a \in \mathbb{R}$ gilt
$$\frac{f(a+h) - f(a)}{h} = \frac{(a+h)^2 - a^2}{h} = \frac{2ah + h^2}{h} = 2a + h\,,$$
also
$$f'(a) = 2a\,.$$

4. $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = 1/x$. Für alle $a \in \mathbb{R} \setminus \{0\}$ gilt
$$\frac{f(a+h) - f(a)}{h} = \frac{\frac{1}{a+h} - \frac{1}{a}}{h} = \frac{\frac{a-(a+h)}{(a+h)a}}{h} = \frac{\frac{-h}{(a+h)a}}{h} = -\frac{1}{(a+h)a}\,,$$
also
$$f'(a) = -\frac{1}{a^2}\,.$$

Wir betrachten nun die Exponentialfunktion.

Lemma 9.3 *Im Komplexen gilt*
$$\lim_{z \to 0} \frac{e^z - 1}{z} = 1\,. \tag{9.7}$$

Beweis: : Übung. □

Wir beschränken uns nun auf reelle Argumente. Aus (9.7) folgt unmittelbar, dass die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ in 0 differenzierbar ist mit
$$f'(0) = 1\,.$$

Die Ableitung in einem beliebigen Punkt $a \in \mathbb{R}$ wird darauf zurückgeführt. Es ist
$$\frac{f(a+h) - f(a)}{h} = \frac{e^{a+h} - e^a}{h} = e^a \frac{e^h - 1}{h}\,.$$

Aus (9.7) folgt nun
$$f'(a) = e^a\,. \tag{9.8}$$
Für den Sinus und den Cosinus kann man ähnlich argumentieren. Wir beginnen mit der Rechnung
$$\frac{e^{ix}-1}{ix} = \frac{\cos x + i\sin x - 1}{ix} = \frac{1}{i}\frac{\cos x - 1}{x} + \frac{\sin x}{x} = \frac{\sin x}{x} + i\frac{1-\cos x}{x}\,. \tag{9.9}$$

Lemma 9.4 *Sinus und Cosinus sind in 0 differenzierbar, und es gilt*
$$\sin'(0) = \lim_{h\to 0}\frac{\sin h}{h} = 1\,, \quad \cos'(0) = \lim_{h\to 0}\frac{\cos h - 1}{h} = 0\,. \tag{9.10}$$

Beweis: Aus Lemma 9.3 folgt
$$\lim_{x\to 0}\frac{e^{ix}-1}{ix} = 1\,.$$
(Der Grenzwert “$x \to 0$” ist als reeller Grenzwert gemeint, es werden also Folgen $z_n = ix_n \to 0$ entlang der imaginären Achse betrachtet. Das ist ein Spezialfall des Grenzwerts (9.7), in dem beliebige Folgen $z_n \to 0$ im Komplexen betrachtet werden.) Es folgt
$$\mathrm{Re}\left(\frac{e^{ix}-1}{ix}\right) \to \mathrm{Re}\,(1) = 1\,, \quad \mathrm{Im}\left(\frac{e^{ix}-1}{ix}\right) \to \mathrm{Im}\,(1) = 0\,.$$
Real- und Imaginärteil sind in (9.9) berechnet worden, daraus ergeben sich die Behauptungen. □

Die Ableitungen von Sinus und Cosinus in beliebigen Punkten $a \in \mathbb{R}$ werden mit Hilfe der Additionstheoreme auf die Ableitungen in 0 zurückgeführt. Für den Sinus gilt
$$\begin{aligned}\frac{\sin(a+h)-\sin(a)}{h} &= \frac{\sin a\cos h + \cos a \sin h - \sin a}{h} \\ &= \sin(a)\frac{\cos h - 1}{h} + \cos(a)\frac{\sin h}{h}\,.\end{aligned}$$
Aus (9.10) folgt nun
$$\sin'(a) = \cos a\,.$$
Analog gilt für den Cosinus
$$\begin{aligned}\frac{\cos(a+h)-\cos(a)}{h} &= \frac{\cos a\cos h - \sin a \sin h - \cos a}{h} \\ &= \cos(a)\frac{\cos h - 1}{h} - \sin(a)\frac{\sin h}{h}\,,\end{aligned}$$
also wegen (9.10)
$$\cos'(a) = -\sin a\,.$$
Als nächstes betrachten wir die Betragsfunktion $f(x) = |x|$. Sie ist differenzierbar in allen Punkten $a \neq 0$, und
$$f'(a) = \begin{cases} 1\,, & a > 0\,, \\ -1\,, & a < 0\,. \end{cases} \tag{9.11}$$

In $a = 0$ ist f nicht differenzierbar, denn für die Folge

$$h_n = (-1)^n \frac{1}{n}$$

gilt

$$\frac{|0 + h_n| - |0|}{h_n} = (-1)^n \,,$$

also existiert der in Definition 9.1 verlangte Grenzwert nicht.

Satz 9.5 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$, a innerer Punkt von D. Dann gilt*

(1) *Ist f differenzierbar in a und setzen wir*

$$r(h) = f(a+h) - f(a) - f'(a)h \,, \tag{9.12}$$

so gilt

$$\lim_{h \to 0} \frac{r(h)}{h} = 0 \,. \tag{9.13}$$

(2) *Ist $c \in \mathbb{R}$ und gilt (9.13) für die durch*

$$r(h) = f(a+h) - f(a) - ch \tag{9.14}$$

definierte Funktion r, so ist f differenzierbar in a und $f'(a) = c$.

Beweis: Ist r durch (9.14) definiert, so gilt

$$\frac{r(h)}{h} = \frac{f(a+h) - f(a)}{h} - c \,. \tag{9.15}$$

Da der Limes auf der rechten Seite existiert und gleich Null ist genau dann, wenn f in a differenzierbar ist und $f'(a) = c$ gilt, folgen beide Behauptungen. $\square$

Wir können die Aussage von Satz 9.5 folgendermaßen interpretieren: Die Differenz

$$r(x-a) = f(x) - [f(a) + f'(a)(x-a)]$$

zwischen den Werten der Funktion f und ihrer *Linearisierung* geht für $x \to a$ schneller gegen 0 als die Differenz $x - a$. Für die Differenz

$$f(x) - [f(a) + c(x-a)]$$

ist dies aber nicht der Fall, wenn $c \neq f'(a)$. Wir können also die Linearisierung

$$g(x) = f(a) + f'(a)(x-a) \tag{9.16}$$

interpretieren als die beste affin-lineare Approximation von f in der Nähe von a. Die Gleichung (9.16) ist nichts anderes als die Gleichung der Tangente an den Graphen von f im Punkt $(a, f(a))$.

Satz 9.6 *Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$, $a \in \mathrm{int}\,(D)$. Dann gilt: Ist f differenzierbar in a, so ist f stetig in a.*

Beweis: Nach Satz 9.5 gilt

$$\lim_{h\to 0} \frac{r(h)}{h} = 0\,,$$

also auch

$$\lim_{h\to 0} r(h) = 0\,.$$

Grenzübergang $h \to 0$ auf beiden Seiten von

$$f(a) + f'(a)h + r(h) = f(a+h)$$

liefert also

$$f(a) = \lim_{h\to 0} f(a+h) = \lim_{x\to a} f(x)\,.$$

□

Wir befassen uns nun mit den Rechenregeln für Ableitungen.

Satz 9.7 *Seien $D \subset \mathbb{R}$, $f, g : D \to \mathbb{R}$, a ein innerer Punkt von D, $\lambda \in \mathbb{R}$. Dann gilt: Sind f und g differenzierbar in a, so sind auch $f+g$, λf und fg differenzierbar in a, und es gelten*

$$(f+g)'(a) = f'(a) + g'(a)\,, \quad (\lambda f)'(a) = \lambda f'(a)\,, \tag{9.17}$$

sowie die Produktregel

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a)\,. \tag{9.18}$$

Ist außerdem $g(a) \neq 0$, so ist auch f/g in a differenzierbar, und es gilt die Quotientenregel

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{(g(a))^2}\,. \tag{9.19}$$

Beweis: Wegen

$$\frac{(f+g)(a+h) - (f+g)(a)}{h} = \frac{f(a+h) - f(a)}{h} + \frac{g(a+h) - g(a)}{h}\,,$$

$$\frac{(\lambda f)(a+h) - (\lambda f)(a)}{h} = \lambda \frac{f(a+h) - f(a)}{h}\,,$$

folgt (9.17) aus den entsprechenden Grenzwertsätzen für Folgen. Zum Beweis der Produktregel betrachten wir

$$\frac{(fg)(a+h) - (fg)(a)}{h} = f(a+h)\frac{g(a+h) - g(a)}{h} + g(a)\frac{f(a+h) - f(a)}{h}\,. \tag{9.20}$$

Da $\lim_{h\to 0} f(a+h) = f(a)$ nach Satz 9.6, folgt (9.18) durch Grenzübergang $h \to 0$ in (9.20). Zum Beweis der Quotientenregel bemerken wir zunächst, dass wegen Satz 9.6 und Folgerung 7.14 für ein hinreichend kleines $\delta > 0$ gilt

$$g(x) \neq 0$$

für alle $x \in D$ mit $|x - a| < \delta$. Wir betrachten nun den Spezialfall $f = 1$. Ist $|h| < \delta$ und $a + h \in D$, so gilt

$$\frac{1}{h}\left(\frac{1}{g(a+h)} - \frac{1}{g(a)}\right) = -\frac{1}{g(a+h)g(a)} \cdot \frac{g(a+h) - g(a)}{h}\,. \tag{9.21}$$

Grenzübergang $h \to 0$ liefert

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{(g(a))^2}\,. \tag{9.22}$$

Ist nun f beliebig, so folgt aus (9.22) und der Produktregel

$$\left(\frac{f}{g}\right)'(a) = \left(f \cdot \frac{1}{g}\right)'(a) = f'(a)\frac{1}{g(a)} + f(a)\frac{-g'(a)}{(g(a))^2} = \frac{g(a)f'(a) - f(a)g'(a)}{(g(a))^2}\,.$$

□

Beispiel 9.8

1. (Summenregel) Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3\sin x - 4e^x$. Dann gilt für jedes $a \in \mathbb{R}$

$$f'(a) = 3\cos a - 4e^a\,.$$

   In $a = 2$ gilt

$$f'(2) = 3\cos 2 - 4e^2\,.$$

2. (Produktregel) Sei $f_n : \mathbb{R} \to \mathbb{R}$, $f_n(x) = x^n$, $n \in \mathbb{N} \setminus \{0\}$. Dann gilt

$$f_n'(a) = na^{n-1}\,. \tag{9.23}$$

   Induktion. $n = 1$ klar, $n \to n+1$:

$$f_{n+1}'(a) = (f_1 f_n)'(a) = f_1'(a)f_n(a) + f_1(a)f_n'(a) = 1 \cdot a^n + a \cdot na^{n-1} = (n+1)a^n\,.$$

3. (Quotientenregel) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^{-n}$, $n \in \mathbb{N} \setminus \{0\}$. Dann gilt

$$f'(a) = -na^{-n-1}\,.$$

   Wir wenden (9.23) und die Quotientenregel auf die Funktion $x \mapsto 1/x^n$ an. Es folgt

$$f'(a) = \frac{-na^{n-1}}{a^{2n}} = -na^{-n-1}\,.$$

4. (Quotientenregel) $f(x) = \tan x = \frac{\sin x}{\cos x}$. Es gilt

$$f'(a) = \frac{\cos a \cos a - \sin a \cdot (-\sin a)}{\cos^2 a} = \frac{1}{\cos^2 a}\,.$$

   In $a = \pi$ gilt $\tan(\pi) = \frac{0}{-1} = 0$ und $\tan'(\pi) = \frac{1}{(-1)^2} = 1$.

□

Die Ableitung einer zusammengesetzten Funktion ergibt sich aus den einzelnen Ableitungen gemäß der Kettenregel.

Satz 9.9 (Kettenregel)
Seien $D, E \subset \mathbb{R}$, $f : D \to \mathbb{R}$, $f(D) \subset E$, $g : E \to \mathbb{R}$, $a \in \operatorname{int}(D)$ und $f(a) \in \operatorname{int}(E)$. Sind f in a und g in $f(a)$ differenzierbar, so ist auch $g \circ f$ in a differenzierbar, und es gilt

$$(g \circ f)'(a) = g'(f(a))f'(a)\,. \tag{9.24}$$

Beweis: Das naheliegende Argument,

$$\frac{g(f(x)) - g(f(a))}{x - a} = \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \cdot \frac{f(x) - f(a)}{x - a}$$

zu betrachten und in den beiden Brüchen auf der rechten Seite einzeln zum Grenzwert überzugehen, bereitet Schwierigkeiten, wenn $f(x) = f(a)$ ist. Wir umgehen dieses Problem, indem wir definieren

$$d(y) = \begin{cases} \frac{g(y)-g(f(a))}{y-f(a)}\,, & y \neq f(a)\,, \\ g'(f(a))\,, & y = f(a)\,. \end{cases}$$

Es gilt dann für alle $x \in D$ mit $x \neq a$

$$\frac{g(f(x)) - g(f(a))}{x - a} = d(f(x))\frac{f(x) - f(a)}{x - a}\,. \tag{9.25}$$

Nach Definition von d gilt, da g in $f(a)$ differenzierbar ist,

$$\lim_{y \to f(a)} d(y) = g'(f(a))\,,$$

also folgt, da f stetig ist in a,

$$\lim_{x \to a} d(f(x)) = g'(f(a))\,,$$

also folgt die Behauptung durch Grenzübergang $x \to a$ in (9.25). □

Beispiel 9.10

1. Wir betrachten $f, g : \mathbb{R} \to \mathbb{R}$, $f(x) = x^4$, $g(x) = \sin x$, und suchen die Ableitung von $g \circ f : \mathbb{R} \to \mathbb{R}$, $(g \circ f)(x) = \sin(x^4)$ in $a = 2$. Es ist

   $$f'(x) = 4x^3\,, \quad g'(x) = \cos x\,,$$

   also

   $$f'(a) = 4a^3\,, \quad f'(2) = 32\,, \quad g'(f(a)) = g'(a^4) = \cos(a^4)\,, \quad g'(f(2)) = \cos 16\,,$$

   also

   $$(g \circ f)(2) = g'(f(2)) \cdot f'(2) = 32 \cos 16\,.$$

2. Wir betrachten $f : (0, \infty) \to \mathbb{R}$, $f(x) = x^c$, $c > 0$ fest. Es ist

   $$f(x) = \exp(c \ln x)\,,$$

   also

   $$f'(x) = \exp(c \ln x) \cdot \frac{c}{x} = x^c \frac{c}{x} = cx^{c-1}\,.$$

3. Wir betrachten $f:(0,\infty)\to\mathbb{R}$, $f(x)=c^x$, $c>0$ fest. Es ist
$$f(x)=\exp(x\ln c)\,,$$
also
$$f'(x)=\exp(x\ln c)\cdot\ln c=c^x\ln c\,.$$

$\square$

Die Ableitung einer Umkehrfunktion lässt sich auf die Ableitung der ursprünglichen Funktion zurückführen. Das erkennt man, indem man die Tangentensteigungen am Graphen von Funktion und Umkehrfunktion betrachtet. Die zugehörige Formel kann man aus der Kettenregel erhalten. Wenn wir annehmen, dass f und f^{-1} differenzierbar sind, so folgt aus
$$f(f^{-1}(y))=y$$
durch Differenzieren beider Seiten
$$f'(f^{-1}(y))\cdot(f^{-1})'(y)=1\,,$$
und damit
$$(f^{-1})'(y)=\frac{1}{f'(f^{-1}(y))}\,. \tag{9.26}$$
Analysiert man die Situation durch den folgenden Satz genauer, so stellt es sich heraus, dass man nicht vorauszusetzen braucht, dass f^{-1} differenzierbar ist; das ergibt sich bereits aus der Differenzierbarkeit von f.

Satz 9.11 *Seien $[a,b]\subset\mathbb{R}$, $f:[a,b]\to\mathbb{R}$ stetig und streng monoton, sei f differenzierbar in $x\in(a,b)$ mit $f'(x)\neq 0$. Dann ist die Umkehrfunktion $f^{-1}:f([a,b])\to\mathbb{R}$ in $y=f(x)$ differenzierbar, und es gilt (9.26).*

Beweis: Sei (h_n) eine beliebige Folge mit $h_n\to 0$, $h_n\neq 0$ für alle n. Der Punkt $y=f(x)$ liegt im Innern des Intervalls $f([a,b])$ (sonst wäre x eine Maximal- bzw. Minimalstelle von f und daher, wie wir später in Satz 10.5 sehen werden, $f'(x)=0$). Also gilt $y+h_n\in f([a,b])$ für hinreichend große $n\in\mathbb{N}$; wir können annehmen, es gilt für alle $n\in\mathbb{N}$. Wir setzen
$$x_n=f^{-1}(y+h_n)\,.$$
Da f^{-1} stetig und streng monoton ist nach Satz 7.25, gilt
$$x_n=f^{-1}(y+h_n)\to f^{-1}(y)=x\,,\quad x_n\neq x\quad\forall\, n\in\mathbb{N}\,,$$
und
$$\frac{f^{-1}(y+h_n)-f^{-1}(y)}{h_n}=\frac{x_n-x}{f(x_n)-f(x)}=\frac{1}{\frac{f(x_n)-f(x)}{x_n-x}}\,,$$
also
$$\lim_{n\to\infty}\frac{f^{-1}(y+h_n)-f^{-1}(y)}{h_n}=\frac{1}{f'(x)}\,.$$

$\square$

Beispiel 9.12

1. Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$. Für $x \neq 0$ gilt $f'(x) = 3x^2 \neq 0$, und Satz 9.11 ist anwendbar. Es ist $f(\mathbb{R}) = \mathbb{R}$, und
$$f^{-1}(y) = \begin{cases} \sqrt[3]{y}\,, & y > 0\,, \\ -\sqrt[3]{-y}\,, & y < 0\,, \\ 0\,, & y = 0\,. \end{cases}$$
Gemäß (9.26) folgt für $y > 0$
$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} = \frac{1}{f'(\sqrt[3]{y})} = \frac{1}{3(\sqrt[3]{y})^2} = \frac{1}{3}y^{-2/3}\,.$$
In $y = 0$ ist f^{-1} nicht differenzierbar: Für $h_n \to 0$ mit $h_n \neq 0$ gilt
$$\frac{f^{-1}(0 + h_n) - f^{-1}(0)}{h_n} = \frac{\sqrt[3]{|h_n|}}{|h_n|} = \frac{1}{(\sqrt[3]{|h_n|})^2}\,,$$
und der rechtsstehende Bruch konvergiert nicht gegen eine reelle Zahl (sondern uneigentlich gegen ∞).

2. Für $\ln : (0, \infty) \to \mathbb{R}$ folgt aus Satz 9.11
$$(\ln)'(y) = \frac{1}{(\exp)'(\ln y)} = \frac{1}{\exp(\ln y)} = \frac{1}{y}\,. \tag{9.27}$$

Folgerung 9.13 *Es gilt für alle $x \in \mathbb{R}$*
$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x\,. \tag{9.28}$$

Beweis: Für $x = 0$ sind beide Seiten gleich 1. Für $x \neq 0$ gilt
$$\ln\left[\left(1 + \frac{x}{n}\right)^n\right] = n \ln\left(1 + \frac{x}{n}\right) = x \cdot \frac{\ln\left(1 + \frac{x}{n}\right) - \ln 1}{\frac{x}{n}}\,,$$
also
$$\lim_{n\to\infty} \ln\left[\left(1 + \frac{x}{n}\right)^n\right] = x \cdot (\ln)'(1) = x\,,$$
also
$$\begin{aligned} e^x = \exp(x) &= \exp\left(\lim_{n\to\infty} \ln\left[\left(1 + \frac{x}{n}\right)^n\right]\right) = \lim_{n\to\infty} \exp\left(\ln\left[\left(1 + \frac{x}{n}\right)^n\right]\right) \\ &= \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n\,. \end{aligned}$$

□

10 Eigenschaften stetiger und differenzierbarer Funktionen

Ist $D \subset \mathbb{K}$, so bezeichnen wir mit $C(D;\mathbb{K})$ die Menge aller auf D definierten und stetigen Funktionen, also

$$C(D;\mathbb{K}) = \{f|\ f : D \to \mathbb{K},\ f \text{ stetig auf } D\}\,. \tag{10.1}$$

Im Fall $\mathbb{K} = \mathbb{R}$ schreiben wir auch einfach $C(D)$.

Wir wissen bereits, dass Summen und skalare Vielfache von stetigen Funktionen ebenfalls stetig sind. Es folgt:

Lemma 10.1 *Sei $D \subset \mathbb{K}$. Dann ist $C(D;\mathbb{K})$ ein Vektorraum über $\mathbb{K}$.* □

Ein Vektorraum, dessen Elemente Funktionen sind, heißt **Funktionenraum**. Der damit verbundene Abstraktionsschritt (wir fassen Funktionen als "Punkte" in einem Vektorraum auf) ist etwa seit dem Beginn des 20. Jahrhunderts eine für die Analysis grundlegende Sichtweise. (In den Lehrveranstaltungen des ersten Studienjahrs steht sie eher im Hintergrund, wird aber später zunehmend wichtig.)

Definition 10.2 (Beschränkte Funktion)
*Eine Funktion $f : D \to \mathbb{R}$ heißt **nach oben beschränkt**, wenn es ein $M \in \mathbb{R}$ gibt mit $f(x) \le M$ für alle $x \in D$. f heißt **nach unten beschränkt**, wenn $-f$ nach oben beschränkt ist. f heißt **beschränkt**, wenn f nach oben und nach unten beschränkt ist.*

Ob eine Funktion beschränkt ist oder nicht, hängt nicht nur von der Funktionsvorschrift, sondern auch von der Wahl des Definitionsbereichs ab. So ist etwa die durch $f(x) = x$ definierte Funktion auf $D = [0,1]$ beschränkt (durch 0 nach unten, durch 1 nach oben), aber auf $D = \mathbb{R}$ nicht, ebensowenig auf $D = [0,\infty)$.

Die durch $f(x) = 1/x$ definierte Funktion ist auf $D = (0,\infty)$ nach oben unbeschränkt, da $f(x) \to \infty$ für $x \to 0$, aber nach unten durch 0 beschränkt, da $f(x) > 0$ für alle $x > 0$. Dasselbe gilt für $D = (0,1]$.

Die Sinusfunktion ist beschränkt auf $\mathbb{R}$, da ihre Werte zwischen -1 und 1 liegen.

Satz 10.3 *Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann ist f beschränkt und nimmt auf $[a,b]$ ein Maximum und ein Minimum an, das heißt, es gibt $p, q \in [a,b]$ mit*

$$f(p) = \max_{x\in[a,b]} f(x)\,, \quad f(q) = \min_{x\in[a,b]} f(x)\,. \tag{10.2}$$

Solche Punkte p und q heißen Maximalstelle bzw. Minimalstelle von f auf $[a,b]$.

Beweis: Wir nehmen zunächst an, f sei nach oben unbeschränkt. Dann gibt es eine Folge (x_n) in $[a,b]$ mit $f(x_n) \ge n$ für alle $n \in \mathbb{N}$. Nach dem Satz von Bolzano-Weierstraß hat (x_n) eine konvergente Teilfolge, sei etwa $x_{n_k} \to c \in [a,b]$. Da f stetig ist, gilt $f(x_{n_k}) \to f(c)$, im Widerspruch zu $f(x_{n_k}) \ge n_k$. Also ist f nach oben beschränkt. Sei nun

$$M = \sup_{x\in[a,b]} f(x)\,.$$

Zu $n \geq 1$ wählen wir nun ein $x_n \in [a,b]$ mit

$$M - \frac{1}{n} \leq f(x_n) \leq M\,.$$

Dann gilt $f(x_n) \to M$. Sei wieder (x_{n_k}) eine konvergente Teilfolge, sei $x_{n_k} \to p$. Dann gilt $f(x_{n_k}) \to f(p)$, also $f(p) = M$ und damit

$$f(p) = \max_{x \in [a,b]} f(x)\,.$$

Anwendung des eben Bewiesenen auf $-f$ ergibt, dass f auch nach unten beschränkt ist und das Minimum annimmt. □

Die vor dem Satz gegebenen Beispiele zeigen, dass man weder auf die Beschränktheit noch auf die Abgeschlossenheit des Definitionsbereichs $[a,b]$ als Voraussetzung verzichten kann.

Maximal- und Minimalstellen brauchen nicht eindeutig bestimmt sein, beispielsweise gilt für konstante Funktionen, dass jedes $p \in [a,b]$ sowohl Maximalstelle als auch Minimalstelle ist.

Definition 10.4 (Lokales Extremum)
Sei $D \subset \mathbb{R}$, $f : D \to \mathbb{R}$. Ein $x \in D$ heißt ***lokales Maximum*** *(bzw. lokales Minimum) von f in D, falls es ein $\varepsilon > 0$ gibt mit*

$$f(x) = \max_{\substack{y \in D \\ |y-x|<\varepsilon}} f(y) \tag{10.3}$$

bzw.

$$f(x) = \min_{\substack{y \in D \\ |y-x|<\varepsilon}} f(y)\,. \tag{10.4}$$

Ein $x \in D$ heißt ***lokales Extremum*** *von f in D, wenn x lokales Maximum oder lokales Minimum von f in D ist.*

Satz 10.5 *Sei $a < b$, $f : (a,b) \to \mathbb{R}$. Ist $x \in (a,b)$ lokales Extremum von f in (a,b), und ist f differenzierbar in x, so gilt*

$$f'(x) = 0\,. \tag{10.5}$$

Beweis: Sei x lokales Maximum, sei ε so gewählt, dass (10.3) gilt mit $D = (a,b)$ und dass $(x-\varepsilon, x+\varepsilon) \subset (a,b)$. Dann gilt für alle $n > \varepsilon^{-1}$

$$\frac{f(x+\frac{1}{n}) - f(x)}{\frac{1}{n}} \leq 0\,,$$
$$\frac{f(x-\frac{1}{n}) - f(x)}{-\frac{1}{n}} \geq 0\,.$$

Grenzübergang $n \to \infty$ liefert $f'(x) \leq 0$ und $f'(x) \geq 0$, also $f'(x) = 0$. Analog für ein lokales Minimum. □

Die Umkehrung von Satz 10.5 gilt nicht. Beispiel: Für $f(x) = x^3$ gilt $f'(0) = 0$, aber 0 ist kein lokales Extremum.

Satz 10.6 (Satz von Rolle)
Seien $a < b$, $f \in C[a,b]$, f differenzierbar in (a,b). Ist $f(a) = f(b)$, so gibt es ein $x \in (a,b)$ mit $f'(x) = 0$.

Beweis: Nach Satz 10.3 hat f in $[a,b]$ eine Maximalstelle p und eine Minimalstelle q. Ist f nicht konstant (andernfalls ist die Behauptung trivialerweise richtig), so gilt $f(p) > f(q)$, und p oder q muss in (a,b) liegen. Wählen wir dieses als x, so gilt $f'(x) = 0$ nach Satz 10.5. □

Satz 10.7 (Verallgemeinerter Mittelwertsatz)
Seien $f, g \in C[a,b]$, f, g differenzierbar in (a,b), sei $g'(x) \neq 0$ für alle $x \in (a,b)$. Dann ist $g(a) \neq g(b)$, und es gibt ein $\xi \in (a,b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)} . \tag{10.6}$$

Beweis: Aus $g(a) = g(b)$ folgt $g'(x) = 0$ für mindestens ein $x \in (a,b)$ nach Satz 10.6, im Widerspruch zur Voraussetzung. Es ist also $g(a) \neq g(b)$. Wir definieren $F : [a,b] \to \mathbb{R}$ durch

$$F(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)) . \tag{10.7}$$

Es ist $F(a) = f(a) = F(b)$. Nach dem Satz von Rolle gibt es ein $\xi \in (a,b)$ mit

$$0 = F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi) ,$$

woraus Satz 10.7 folgt. □

Folgerung 10.8 (Mittelwertsatz) *Sei $f \in C[a,b]$, f differenzierbar in (a,b). Dann gibt es ein $\xi \in (a,b)$ mit*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi) . \tag{10.8}$$

Beweis: Anwendung von Satz 10.7 mit $g(x) = x$. □

Folgerung 10.9 *Sei $f \in C[a,b]$, f differenzierbar in (a,b), sei*

$$m = \inf_{\xi \in (a,b)} f'(\xi) , \quad M = \sup_{\xi \in (a,b)} f'(\xi) . \tag{10.9}$$

Dann gilt

$$m(y - x) \leq f(y) - f(x) \leq M(y - x) \tag{10.10}$$

für alle $x, y \in [a,b]$ mit $x \leq y$.

Beweis: Anwendung von Satz 10.8 auf das Teilintervall $[x,y]$ von $[a,b]$. □

Folgerung 10.10 *Sei $f \in C[a,b]$, f differenzierbar in (a,b), sei $f'(x) = 0$ für alle $x \in (a,b)$. Dann ist f konstant.*

Beweis: Folgt aus (10.10) wegen $m = M = 0$. □

Mit Hilfe der Ableitung kann man oft feststellen, ob Funktionen monoton sind.

Satz 10.11 *Sei $f \in C[a,b]$, f differenzierbar in (a,b). Dann gilt:*

$$f'(x) \geq 0 \text{ für alle } x \in (a,b) \qquad \Rightarrow \qquad f \text{ ist monoton wachsend}$$

$$f'(x) > 0 \text{ für alle } x \in (a,b) \qquad \Rightarrow \qquad f \text{ ist streng monoton wachsend}$$

$$f'(x) \leq 0 \text{ für alle } x \in (a,b) \qquad \Rightarrow \qquad f \text{ ist monoton fallend}$$

$$f'(x) < 0 \text{ für alle } x \in (a,b) \qquad \Rightarrow \qquad f \text{ ist streng monoton fallend}$$

Beweis: Nur die erste Behauptung. Sei $x < y$, dann existiert nach Satz 10.8 ein $\xi \in (a,b)$ mit

$$f(x) - f(y) = f'(\xi)(x-y) \leq 0\,.$$

□

Umkehrfunktionen der trigonometrischen Funktionen. Wir betrachten zunächst den Sinus.

$$\sin(-\frac{\pi}{2}) = -1\,, \quad \sin(\frac{\pi}{2}) = 1\,, \quad \sin'(x) = \cos(x) > 0 \quad \text{für alle } x \in (-\frac{\pi}{2}, \frac{\pi}{2})\,.$$

Also ist $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1,1]$ stetig, streng monoton wachsend (wegen Satz 10.11) und bijektiv. Die Umkehrfunktion heißt **Arcus-Sinus**,

$$\arcsin : [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]\,, \tag{10.11}$$

und ist ebenfalls stetig und streng monoton wachsend. Die Ableitung des Arcus-Sinus für $x \in (-1,1)$ errechnet sich als

$$(\arcsin)'(x) = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-\sin^2(\arcsin x)}} = \frac{1}{\sqrt{1-x^2}}\,. \tag{10.12}$$

Analog erhält man, dass $\cos : [0,\pi] \to [-1,1]$ streng monoton fallend ist; die Umkehrfunktion heißt **Arcus-Cosinus**,

$$\arccos : [-1,1] \to [0,\pi]\,, \tag{10.13}$$

sie ist ebenfalls stetig und streng monoton fallend. Ihre Ableitung in $(-1,1)$ ist

$$(\arccos)'(x) = -\frac{1}{\sin(\arccos(x))} = -\frac{1}{\sqrt{1-\cos^2(\arccos(x))}} = -\frac{1}{\sqrt{1-x^2}}\,. \tag{10.14}$$

Wegen

$$(\tan)'(x) = \frac{1}{\cos^2 x} > 0\,, \quad \text{für alle } x \in (-\frac{\pi}{2}, \frac{\pi}{2})$$

erhalten wir, dass $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ stetig, streng monoton wachsend und bijektiv ist. Die Umkehrfunktion

$$\arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2}) \tag{10.15}$$

heißt **Arcus-Tangens**, und wegen

$$\frac{1}{\cos^2 y} = \frac{\sin^2 y + \cos^2 y}{\cos^2 y} = \tan^2 y + 1$$

gilt

$$(\arctan)'(x) = \cos^2(\arctan x) = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}\,. \tag{10.16}$$

Höhere Ableitungen. Ist $f : (a,b) \to \mathbb{R}$ differenzierbar auf (a,b), so können wir in jedem Punkt $x \in (a,b)$ die Ableitung f' betrachten. Dadurch erhalten wir eine Funktion

$$f' : (a,b) \to \mathbb{R}\,. \tag{10.17}$$

Ist $f' : (a,b) \to \mathbb{R}$ ebenfalls differenzierbar auf (a,b), so können wir deren Ableitung $(f')' : (a,b) \to \mathbb{R}$ betrachten. Sie heißt die **zweite Ableitung** von f auf (a,b), wir schreiben f'' statt $(f')'$. Diesen Prozess können wir fortsetzen und die dritte, vierte usw. Ableitung von f betrachten. Allgemein ist die ***n*-te Ableitung** $f^{(n)}$ einer Funktion $f : (a,b) \to \mathbb{R}$ induktiv definiert als die Ableitung der $(n-1)$-ten Ableitung $f^{(n-1)}$ von f, sofern sie existiert.

Als Beispiel betrachten wir

$$f(x) = x^2\,, \quad f'(x) = 2x\,, \quad f''(x) = 2\,, \quad f^{(3)}(x) = 0\,,$$

alle weiteren Ableitungen sind ebenfalls gleich Null. Weitere Beispiele sind

$$f(x) = e^x\,, \quad f'(x) = e^x\,, \quad f^{(n)}(x) = e^x \quad \text{für alle } n,$$

sowie

$$f(x) = \sin x\,, \quad f'(x) = \cos x\,, \quad f''(x) = -\sin x\,, \quad f'''(x) = -\cos x\,, \quad f^{(4)}(x) = \sin x\,.$$

Eine auf (a,b) differenzierbare Funktion f heißt **stetig differenzierbar** auf (a,b), falls die Ableitung $f' : (a,b) \to \mathbb{R}$ stetig ist.

Existiert die zweite Ableitung $f'' : (a,b) \to \mathbb{R}$, so ist f' stetig, da jede differenzierbare Funktion stetig ist. Entsprechend folgt, dass $f^{(k)} : (a,b) \to \mathbb{R}$ stetig ist für alle $k < n$, falls $f^{(n)} : (a,b) \to \mathbb{R}$ existiert.

Mit Hilfe der zweiten Ableitung kann man lokale Extrema zusätzlich charakterisieren.

Satz 10.12 *Sei $f : (a,b) \to \mathbb{R}$ zweimal stetig differenzierbar, sei $x \in (a,b)$.*
(i) Ist x ein lokales Minimum, so gilt $f''(x) \geq 0$.
(ii) Gilt $f'(x) = 0$ und $f''(x) > 0$, so ist x ein striktes lokales Minimum.

"Striktes lokales Minimum" bedeutet, dass $f(x+h) > f(x)$ gilt, falls $|h|$ hinreichend klein ist, $h \neq 0$.

Beweis: Zu (i): Sei $\varepsilon > 0$ so gewählt, dass $f(x+h) \geq f(x)$ für alle h mit $|h| < \varepsilon$ gilt. Sei $h_n \to 0$, $0 < h_n < \varepsilon$ für alle n. Gemäß Mittelwertsatz gibt es $t_n \in (x, x+h_n)$ mit

$$0 \leq f(x+h_n) - f(x) = f'(t_n)h_n = (f'(t_n) - f'(x))h_n \,,$$

da $f'(x) = 0$. Gemäß Mittelwertsatz, angewendet auf f', gibt es $s_n \in (x, x+t_n)$ mit

$$0 \leq f''(s_n)(t_n - x)h_n \,.$$

Es folgt $f''(s_n) \geq 0$ für alle n. Da $s_n \to x$ wegen $h_n \to 0$, folgt $f''(x) \geq 0$.
Zu (ii): Gemäß Satz **??** gibt es ein $\delta > 0$, so dass $f''(t) > 0$ für alle $t \in (x-\delta, x+\delta)$. Die Funktion f' ist daher streng monoton wachsend auf $(x-\delta, x+\delta)$, und es folgt $f'(t) > f'(x) = 0$ für $t > x$ und $f'(t) < f'(x) = 0$ für $t < x$. Ist nun $h \neq 0$ mit $|h| < \delta$, so gilt nach Mittelwertsatz

$$f(x+h) - f(x) = f'(t)h$$

mit einem geeigneten t zwischen x und $x+h$. Da $f'(t)h > 0$ gilt für alle $h \neq 0$, $|h| < \delta$, folgt die Behauptung. □

Die Regel von de l'Hospital. Sie ist manchmal nützlich bei der Berechnung von Grenzwerten.

Satz 10.13 (Satz von de l'Hospital) *Seien $f, g \in C[a,b]$, f, g differenzierbar in (a,b), sei $x \in (a,b)$ mit $f(x) = g(x) = 0$, sei $g'(\xi) \neq 0$ für alle $\xi \in (a,b)$ mit $\xi \neq x$. Dann gilt: Falls der Grenzwert*

$$\lim_{\xi \to x} \frac{f'(\xi)}{g'(\xi)} \tag{10.18}$$

existiert, so existiert auch der Grenzwert

$$\lim_{\xi \to x} \frac{f(\xi)}{g(\xi)} \,, \tag{10.19}$$

und

$$\lim_{\xi \to x} \frac{f(\xi)}{g(\xi)} = \lim_{\xi \to x} \frac{f'(\xi)}{g'(\xi)} \,. \tag{10.20}$$

Beweis: Sei (x_n) eine beliebige Folge in (a,b) mit $x_n \to x$ und $x_n \neq x$ für alle n. Wie im Beweis von Satz 10.7 folgt $g(x_n) \neq 0$ für alle n. Nach Satz 10.7 gibt es für alle n ein ξ_n zwischen x und x_n mit

$$\frac{f(x_n)}{g(x_n)} = \frac{f(x_n) - f(x)}{g(x_n) - g(x)} = \frac{f'(\xi_n)}{g'(\xi_n)} \,. \tag{10.21}$$

Aus $x_n \to x$ folgt $\xi_n \to x$ und damit

$$\lim_{n \to \infty} \frac{f(x_n)}{g(x_n)} = \lim_{n \to \infty} \frac{f'(\xi_n)}{g'(\xi_n)} = \lim_{\xi \to x} \frac{f'(\xi)}{g'(\xi)} \,.$$

Da die Folge (x_n) beliebig war, folgt die Behauptung. □

Beispiel 10.14

1.
$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = \frac{\cos 0}{1} = 1\,. \tag{10.22}$$

2. Wir berechnen
$$\lim_{x\to 0} \left(\frac{1}{\sin x} - \frac{1}{x} \right) . \tag{10.23}$$
Es ist
$$\frac{1}{\sin x} - \frac{1}{x} = \frac{f(x)}{g(x)} \tag{10.24}$$
mit
$$f(x) = x - \sin x\,, \quad g(x) = x \sin x\,, \quad f(0) = g(0) = 0\,. \tag{10.25}$$
Es gilt
$$\frac{f'(x)}{g'(x)} = \frac{1 - \cos x}{\sin x + x \cos x}\,, \quad f'(0) = g'(0) = 0\,, \tag{10.26}$$
und weiter
$$\frac{f''(x)}{g''(x)} = \frac{\sin x}{2 \cos x - x \sin x}\,, \quad f''(0) = 0\,, \quad g''(0) = 2\,. \tag{10.27}$$
Wir können Satz 10.13 zweimal anwenden und erhalten
$$0 = \lim_{x\to 0} \frac{f''(x)}{g''(x)} = \lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{f(x)}{g(x)}\,. \tag{10.28}$$

Folgen von Funktionen. Folgen sind allgemein definiert als Abbildungen von $\mathbb{N}$ in irgendeine Menge M; bisher haben wir lediglich Zahlenfolgen betrachtet mit $M = \mathbb{N}$, $M = \mathbb{R}$ oder $M = \mathbb{C}$. In der Analysis werden auch **Funktionenfolgen** betrachtet, das sind Folgen, deren Elemente Funktionen sind.

Definieren wir beispielsweise $f_n : \mathbb{R} \to \mathbb{R}$,
$$f_n(x) = x^n\,, \tag{10.29}$$
so erhalten wir eine Folge $(f_n)_{n\in\mathbb{N}}$, deren n-tes Element das n-te Monom ist. Für jedes fest gewählte $x \in \mathbb{R}$ entsteht daraus eine Zahlenfolge $(f_n(x))_{n\in\mathbb{N}}$. Deren Konvergenz haben wir bereits untersucht, es gilt
$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0\,, & |x| < 1\,, \\ 1\,, & x = 1\,. \end{cases} \tag{10.30}$$
Für $x \le -1$ ist sie divergent. Für $x > 1$ ist sie ebenfalls divergent, aber uneigentlich konvergent gegen ∞.

Definition 10.15 (Punktweise Konvergenz)
Sei D Menge, sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : D \to \mathbb{K}$ mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Die Folge (f_n) heißt ***punktweise konvergent*** *gegen die Funktion $f : D \to \mathbb{K}$, falls*
$$\lim_{n\to\infty} f_n(x) = f(x)\,, \qquad \text{für alle } x \in D. \tag{10.31}$$

Die Form der Menge D ist völlig gleichgültig, da die Bedingung (10.31) keine Beziehung zwischen unterschiedlichen Werten des Arguments x herstellt.

Für das Beispiel (10.29) zeigt (10.30), dass die Folge (f_n) auf $D = [0,1]$ punktweise gegen die durch

$$f(x) = \begin{cases} 0\,, & |x| < 1\,, \\ 1\,, & x = 1\,, \end{cases} \tag{10.32}$$

definierte Funktion $f : [0,1] \to \mathbb{R}$ konvergiert.

Als weiteres Beispiel betrachten wir

$$f_n(x) = x^2 + \frac{x}{n}\,, \quad f_n : [0,1] \to \mathbb{R}\,. \tag{10.33}$$

Setzen wir $f(x) = x^2$, so gilt

$$|f_n(x) - f(x)| = \left|\frac{x}{n}\right| \le \frac{1}{n}\,, \quad \text{für alle } x \in [0,1]. \tag{10.34}$$

Also konvergiert f_n punktweise gegen f auf $[0,1]$.

Auf der rechten Seite von (10.34) kommt kein x vor, die Konvergenz von $f_n(x)$ gegen $f(x)$ ist in gewisser Weise "gleichmäßig" hinsichtlich der Wahl von x.

Definition 10.16 (Gleichmäßige Konvergenz)
Sei D Menge, sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : D \to \mathbb{K}$ mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Die Folge (f_n) heißt ***gleichmäßig konvergent*** *gegen die Funktion $f : D \to \mathbb{K}$, falls zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass gilt*

$$|f_n(x) - f(x)| \le \varepsilon \quad \textit{für alle } x \in D \textit{ und alle } n \ge n_0. \tag{10.35}$$

Gleichbedeutend mit (10.35) ist die Bedingung

$$\sup_{x\in D} |f_n(x) - f(x)| \le \varepsilon \quad \text{für alle } n \ge n_0. \tag{10.36}$$

Hieraus erkennen wir, dass die Folge (f_n) aus (10.33) wegen (10.34) gleichmäßig gegen $f(x) = x^2$ konvergiert.

Lemma 10.17 *Jede gleichmäßig gegen $f : D \to \mathbb{K}$ konvergente Folge (f_n) von Funktionen $f_n : D \to \mathbb{K}$ ist auch punktweise gegen f konvergent.*

Beweis: Das ist eine unmittelbare Folge der Definition. □

Umgekehrt braucht eine punktweise konvergente Folge aber nicht gleichmäßig konvergent zu sein. Wir betrachten nochmals das Beispiel $f_n(x) = x^n$ mit

$$\lim_{n\to\infty} f_n(x) = f(x) = \begin{cases} 0\,, & |x| < 1\,, \\ 1\,, & x = 1\,. \end{cases} \tag{10.37}$$

Wir wählen $x_n \in [0,1]$ so, dass $f_n(x_n) = \frac{1}{2}$, das ist möglich nach dem Zwischenwertsatz. Es gilt dann

$$|f_n(x_n) - f(x_n)| = \frac{1}{2}\,, \quad \text{für alle } n \in \mathbb{N}.$$

Die Bedingung (10.36) ist daher für $\varepsilon < 1/2$ nicht erfüllbar, die Konvergenz von f_n gegen f ist nicht gleichmäßig.

In diesem Beispiel beobachten wir außerdem, dass einerseits alle Funktionen f_n auf $[0, 1]$ stetig sind, andererseits die "Grenzfunktion" f aber nicht (f ist im Punkt 1 unstetig). Durch den punktweisen Grenzübergang haben wir die Stetigkeit verloren. Das kann bei der gleichmäßigen Konvergenz nicht passieren.

Satz 10.18 *Sei $D \subset \mathbb{K}$, sei (f_n) Folge stetiger Funktionen $f_n : D \to \mathbb{K}$, sei $f : D \to \mathbb{K}$, und es gelte $f_n \to f$ gleichmäßig. Dann ist f auf D stetig.*

Gleichmäßige Grenzwerte stetiger Funktionen sind also stetig.

Beweis: Seien $a \in D$ und $x_k \to a$ beliebig, wir wollen zeigen, dass $f(x_k) \to f(a)$ gilt. Sei $\varepsilon > 0$. Da (f_n) gleichmäßig gegen f konvergiert, gibt es ein n mit

$$|f(x) - f_n(x)| \le \varepsilon\,, \quad \text{für alle } x \in D.$$

Aus der Dreiecksungleichung folgt

$$\begin{aligned} |f(x_k) - f(a)| &\le |f(x_k) - f_n(x_k)| + |f_n(x_k) - f_n(a)| + |f_n(a) - f(a)| \\ &\le |f_n(x_k) - f_n(a)| + 2\varepsilon\,. \end{aligned} \tag{10.38}$$

Da f_n stetig ist, gilt $f_n(x_k) \to f_n(a)$ für $k \to \infty$, es gibt also ein $k_0 \in \mathbb{N}$ mit

$$|f_n(x_k) - f_n(a)| \le \varepsilon\,, \quad \text{für alle } k \ge k_0. \tag{10.39}$$

Setzen wir (10.38) und (10.39) zusammen, so ergibt sich

$$|f(x_k) - f(a)| \le 3\varepsilon\,, \quad \text{für alle } k \ge k_0.$$

Es folgt $f(x_k) \to f(a)$ und damit die Behauptung. □

In Quantorenschreibweise drückt sich der Unterschied von punktweiser und gleichmäßiger Konvergenz folgendermaßen aus:

$$f_n \to f \text{ punktweise:} \quad \forall \varepsilon > 0 \quad \forall x \in D \quad \exists n_0 \in \mathbb{N} \quad \forall n \ge n_0 : \quad |f_n(x) - f(x)| < \varepsilon\,,$$

$$f_n \to f \text{ gleichmäßig:} \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall x \in D \quad \forall n \ge n_0 : \quad |f_n(x) - f(x)| < \varepsilon\,.$$

Wir betrachten nun die gleichmäßige Konvergenz aus einem anderen Blickwinkel. Zu einer beliebigen Menge D definieren wir den Raum aller beschränkten Funktionen auf D mit Werten in $\mathbb{K}$ durch

$$B(D; \mathbb{K}) = \{f \,|\, f : D \to \mathbb{K},\ f \text{ ist beschränkt}\}\,. \tag{10.40}$$

Lemma 10.19
$B(D; \mathbb{K})$ ist ein Vektorraum.

Beweis: Summe und skalare Vielfache von beschränkten Funktionen sind ebenfalls beschränkt, daher ist $B(D; \mathbb{K})$ ein Unterraum des Vektorraums aller Funktionen von D nach $\mathbb{K}$. □

Definition 10.20 (Supremumsnorm)
Sei $f : D \to \mathbb{K}$ beschränkt. Wir definieren die ***Supremumsnorm*** *von f auf D durch*

$$\|f\|_\infty = \sup_{x \in D} |f(x)| \,. \tag{10.41}$$

Beispiel 10.21

1. $f(x) = \sin x$ mit $D = \mathbb{R}$:
$$\|f\|_\infty = \sup_{x \in \mathbb{R}} |\sin x| = 1 \,.$$
2. Wir betrachten $f(x) = x^3$. Mit $D = [0,1]$:
$$\|f\|_\infty = \sup_{x \in [0,1]} |x^3| = 1 \,.$$
Mit $D = [-2,1]$ oder $D = (-2,1]$:
$$\|f\|_\infty = \sup_{x \in [-2,1]} |x^3| = \sup_{x \in (-2,1]} |x^3| = 8 \,.$$
3. Sei $f(x) = 0$ für $x < 0$, $f(x) = 1$ für $x \geq 0$. Mit $D = [-1,0)$ gilt $\|f\|_\infty = 0$, mit $D = [-1,0]$ gilt $\|f\|_\infty = 1$.

Falls f stetig und D ein abgeschlossenes und beschränktes Intervall ist, so hat $|f|$ ein Maximum auf D, und

$$\|f\|_\infty = \sup_{x \in D} |f(x)| = \max_{x \in D} |f(x)| \,.$$

Satz 10.22
Sei D Menge, seien $f, g : D \to \mathbb{K}$ beschränkte Funktionen. Dann gilt

$$\|f\|_\infty = 0 \qquad \Leftrightarrow \qquad f = 0 \,, \tag{10.42}$$

$$\|\lambda f\|_\infty = |\lambda| \, \|f\|_\infty \qquad \textit{für alle } \lambda \in \mathbb{K}, \tag{10.43}$$

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty \,. \tag{10.44}$$

Beweis: Es gilt

$$f \neq 0 \quad \Leftrightarrow \quad \exists x \in D \text{ mit } |f(x)| > 0 \quad \Leftrightarrow \quad \|f\|_\infty > 0 \,,$$
$$\|\lambda f\|_\infty = \sup_{x \in D} |\lambda f(x)| = |\lambda| \sup_{x \in D} |f(x)| = |\lambda| \, \|f\|_\infty \,,$$
$$|(f+g)(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty \,, \qquad \text{für alle } x \in D,$$

also

$$\|f + g\|_\infty = \sup_{x \in D} |f(x) + g(x)| \leq \|f\|_\infty + \|g\|_\infty \,.$$

□

Ersetzt man in (10.42) – (10.44) die Supremumsnorm einer Funktion durch den Betrag einer Zahl, so erhalten wir die bekannten Eigenschaften des Betrags in $\mathbb{R}$ oder $\mathbb{C}$. In der Tat, diese Eigenschaften werden uns später auf den Begriff der Norm führen. Dieser verallgemeinert den Begriff der Länge eines Vektors vom Vektorraum $\mathbb{R}^n$ auf Funktionenräume.

Satz 10.23 *Sei $D \subset \mathbb{K}$, sei (f_n) Folge beschränkter Funktionen $f_n : D \to \mathbb{K}$, sei $f : D \to \mathbb{K}$ beschränkt. Es gilt: f_n konvergiert gleichmäßig gegen f genau dann, wenn*

$$\lim_{n\to\infty} \|f_n - f\|_\infty = 0\,. \tag{10.45}$$

Beweis: Für $n \in \mathbb{N}$ und $\varepsilon > 0$ gilt

$$|f_n(x) - f(x)| \le \varepsilon\,, \qquad \text{für alle } x \in D, \tag{10.46}$$

genau dann, wenn

$$\|f_n - f\|_\infty = \sup_{x\in D} |f_n(x) - f(x)| \le \varepsilon\,. \tag{10.47}$$

Gleichmäßige Konvergenz $f_n \to f$ ist äquivalent zu

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \text{so dass (10.46) gilt für alle } n \ge n_0,$$

Konvergenz $\|f_n - f\|_\infty \to 0$ ist äquivalent zu

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \text{so dass (10.47) gilt für alle } n \ge n_0.$$

□

11 Das Integral

Seien $a, b \in \mathbb{R}$, $a < b$, sei $f : [a, b] \to \mathbb{R}$. Wir wollen eine Zahl $I(f)$ – das Integral von f – definieren, welche für nichtnegative Funktionen der anschaulichen Vorstellung "Flächeninhalt unterhalb des Graphen von f" entspricht. Die zugehörige Abbildung I sollte auf einer möglichst großen Klasse $\mathcal{F}$ von Funktionen definiert sein, und

$$I : \mathcal{F} \to \mathbb{R}$$

sollte möglichst viele "gute mathematische Eigenschaften" haben. Der Versuch, diesen beiden Forderungen gerecht zu werden, hat in den letzten 150 Jahren zu einer ganzen Reihe unterschiedlicher Definitionen des Integrals geführt. In der heutigen Analysis wird überwiegend das sogenannte Lebesgue-Integral verwendet.

In dieser einführenden Vorlesung befassen wir uns mit einem etwas einfacheren Integralbegriff.

Im folgenden ist $[a, b]$ immer ein abgeschlossenes Intervall im $\mathbb{R}$ mit $a, b \in \mathbb{R}$, $a < b$.

Definition 11.1 (Zerlegung)
*Eine endliche Menge $Z = \{x_0, \ldots, x_n\}$ mit $n \geq 1$ und $a = x_0 < x_1 < \ldots < x_n = b$ heißt **Zerlegung** von $[a, b]$. Die Intervalle (x_{i-1}, x_i), $1 \leq i \leq n$, heißen Teilintervalle von Z. Eine Zerlegung Z' heißt **Verfeinerung** der Zerlegung Z, wenn $Z' \supset Z$.*

Beispielsweise ist $\{0, 1/3, 1/2, 1\}$ eine Zerlegung von $[0, 1]$, ebenso

$$\left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}, \quad \text{also} \quad x_i = \frac{i}{n}\,. \tag{11.1}$$

Zerlegungen wie (11.1), für die $x_{i+1} - x_i$ konstant ist, heißen **äquidistant**.

Definition 11.2 (Treppenfunktion)
Eine Funktion $\varphi : [a, b] \to \mathbb{R}$ heißt Treppenfunktion auf $[a, b]$, wenn es eine Zerlegung Z gibt, so dass φ auf allen Teilintervallen von Z konstant ist, d.h. für alle i gibt es $c_i \in \mathbb{R}$ mit $\varphi(x) = c_i$ für alle $x \in (x_{i-1}, x_i)$. Eine solche Zerlegung Z heißt zugehörig zu φ. (Über das Verhalten von φ in den Teilpunkten x_i wird nichts vorausgesetzt.) Wir setzen

$$T[a, b] = \{\varphi |\ \varphi : [a, b] \to \mathbb{R}\ \textit{Treppenfunktion}\}\,. \tag{11.2}$$

Lemma 11.3 *Summe, Produkt und skalare Vielfache von Treppenfunktionen sind ebenfalls Treppenfunktionen. $T[a, b]$ ist ein Vektorraum.*

Beweis: Sind $\varphi, \psi \in T[a, b]$ mit zugehörigen Zerlegungen Z_φ und Z_ψ, so ist $Z_\varphi \cup Z_\psi$ zugehörige Zerlegung zu $\alpha\varphi + \beta\psi$ für alle $\alpha, \beta \in \mathbb{R}$, sowie zu $\varphi\psi$. Somit ist $T[a, b]$ ein Unterraum des Vektorraums aller Abbildungen von $[a, b]$ nach $\mathbb{R}$. □

Ist $\varphi : [a, b] \to \mathbb{R}$ Treppenfunktion, so wollen wir das Integral von φ durch

$$I(\varphi) = \sum_{i=1}^{n} c_i(x_i - x_{i-1}) \tag{11.3}$$

definieren, wobei $J_i = (x_{i-1}, x_i)$ Teilintervalle einer zu φ zugehörigen Zerlegung Z sind und φ auf J_i den konstanten Wert c_i hat. Falls alle c_i nichtnegativ sind, entspricht das gerade dem Flächeninhalt der zwischen der x-Achse und dem Graphen von φ liegenden Menge.

Beispiel: $\varphi : [0,4] \to \mathbb{R}$,

$$\varphi(x) = \begin{cases} 2\,, & 0 \le x < 1\,, \\ 1\,, & 1 \le x < 3\,, \\ 4\,, & 3 \le x \le 4\,, \end{cases}$$
$$I(\varphi) = 2 \cdot (1-0) + 1 \cdot (3-1) + 4 \cdot (4-3) = 8\,.$$

Lemma 11.4 *Ist $\varphi \in T[a,b]$, und sind $Z = \{x_0, \ldots, x_n\}$ und $Z' = \{x'_0, \ldots, x'_m\}$ zu φ zugehörige Zerlegungen mit $\varphi = c_i$ auf (x_{i-1}, x_i) und $\varphi = c'_i$ auf (x'_{i-1}, x'_i), so ist*

$$\sum_{i=1}^{n} c_i(x_i - x_{i-1}) = \sum_{i=1}^{m} c'_i(x'_i - x'_{i-1})\,. \tag{11.4}$$

Beweis: Wir betrachten zunächst den Spezialfall, dass Z' aus Z durch Hinzunahme eines Teilpunktes $x \in (x_{k-1}, x_k)$ entsteht. Dann ist $m = n+1$ und

$$\begin{aligned} \sum_{i=1}^{m} c'_i(x'_i - x'_{i-1}) &= \sum_{\substack{i=1 \\ i \ne k}}^{n} c_i(x_i - x_{i-1}) + c_k(x - x_{k-1}) + c_k(x_k - x) \\ &= \sum_{i=1}^{n} c_i(x_i - x_{i-1})\,. \end{aligned}$$

Der allgemeine Fall wird darauf zurückgeführt: Jede Verfeinerung Z' von Z entsteht, indem wir von Z ausgehend endlich oft einen einzelnen Teilpunkt hinzunehmen. Somit gilt (11.4), falls Z' eine Verfeinerung von Z ist. Sind nun Z und Z' beliebige Zerlegungen, so besitzen sie $Z \cup Z'$ als gemeinsame Verfeinerung, und (11.4) gilt. □

Lemma 11.4 zeigt, dass der Wert der rechten Seite von (11.3) nicht davon abhängt, welche der zu φ zugehörigen Zerlegungen Z zugrundegelegt wird. Die folgende Definition ist daher sinnvoll.

Definition 11.5 (Integral einer Treppenfunktion)
Ist $\varphi \in T[a,b]$, so definieren wir das Integral $I(\varphi)$ von φ über $[a,b]$ durch (11.3). Statt $I(\varphi)$ schreiben wir auch

$$\int_a^b \varphi(x)\,dx\,. \tag{11.5}$$

Lemma 11.6 *Für alle $\varphi, \psi \in T[a,b]$ und alle $\alpha, \beta \in \mathbb{R}$ gilt*

$$\int_a^b (\alpha\varphi + \beta\psi)(x)\,dx = \alpha \int_a^b \varphi(x)\,dx + \beta \int_a^b \psi(x)\,dx\,, \tag{11.6}$$

$$\left| \int_a^b \varphi(x)\,dx \right| \le \int_a^b |\varphi(x)|\,dx \le (b-a)\|\varphi\|_\infty\,. \tag{11.7}$$

Ist $\varphi \le \psi$ (d.h. $\varphi(x) \le \psi(x)$ für alle $x \in [a,b]$), so gilt

$$\int_a^b \varphi(x)\,dx \le \int_a^b \psi(x)\,dx\,. \tag{11.8}$$

Beweis: Die Aussagen (11.6) und (11.8) werden mit Definition 11.5 auf die entsprechenden Eigenschaften von endlichen Summen zurückgeführt. (11.7) folgt aus (11.8). □

Definition 11.7 (Regelfunktion)
Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt Regelfunktion, wenn es eine Folge (φ_n) von Treppenfunktionen in $T[a,b]$ gibt, welche gleichmäßig gegen f konvergiert. Wir definieren

$$R[a,b] = \{f \mid f : [a,b] \to \mathbb{R},\ f \text{ ist Regelfunktion}\}\,. \tag{11.9}$$

Regelfunktionen brauchen nicht stetig zu sein (jede Treppenfunktion ist eine Regelfunktion), aber jede Regelfunktion f ist beschränkt: Ist φ Treppenfunktion mit $\|f-\varphi\|_\infty \le 1$, so gilt

$$|f(x)| \le |f(x) - \varphi(x)| + |\varphi(x)| \le 1 + \|\varphi\|_\infty$$

für alle $x \in [a,b]$. (Jede Treppenfunktion ist beschränkt, da sie nur endlich viele verschiedene Werte annimmt.)

Lemma 11.8 *Eine Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann eine Regelfunktion, wenn es zu jedem $\varepsilon > 0$ eine Treppenfunktion φ gibt mit*

$$\|f - \varphi\|_\infty \le \varepsilon\,.$$

Beweis: Das ist eine unmittelbare Folge von Definition 11.7. □

Lemma 11.9 *Sei $f \in R[a,b]$, seien (φ_n) und (ψ_n) Folgen in $T[a,b]$ mit $\varphi_n \to f$ und $\psi_n \to f$ gleichmäßig. Dann gilt*

$$\lim_{n\to\infty} \int_a^b \varphi_n(x)\,dx = \lim_{n\to\infty} \int_a^b \psi_n(x)\,dx\,. \tag{11.10}$$

Beweis: Für alle $n, m \in \mathbb{N}$ gilt nach Lemma 11.6

$$\begin{aligned} \left| \int_a^b \varphi_n(x)\,dx - \int_a^b \varphi_m(x)\,dx \right| &\le (b-a)\|\varphi_n - \varphi_m\|_\infty \\ &\le (b-a)(\|f-\varphi_n\|_\infty + \|f - \varphi_m\|_\infty)\,. \end{aligned} \tag{11.11}$$

Die durch

$$I_n = \int_a^b \varphi_n(x)\,dx$$

definierte Folge ist also eine Cauchyfolge in $\mathbb{R}$, also konvergent. Dasselbe gilt für

$$J_n = \int_a^b \psi_n(x)\,dx$$

Wie in (11.11) folgt

$$\left| \int_a^b \varphi_n(x)\, dx - \int_a^b \psi_n(x)\, dx \right| \le (b-a)(\|f-\varphi_n\|_\infty + \|f-\psi_n\|_\infty) \quad \to \quad 0$$

für $n \to \infty$, also gilt (11.10). □

Lemma 11.9 zeigt, dass die folgende Definition sinnvoll ist.

Definition 11.10 (Integral einer Regelfunktion)
Sei $f \in R[a,b]$. Ist (φ_n) eine gleichmäßig gegen f konvergente Folge von Treppenfunktionen, so definieren wir das Integral von f über $[a,b]$ durch

$$\int_a^b f(x)\, dx = \lim_{n\to\infty} \int_a^b \varphi_n(x)\, dx\,. \tag{11.12}$$

Lemma 11.11 *$R[a,b]$ ist ein Vektorraum. Für alle $f,g \in R[a,b]$ und alle $\alpha,\beta \in \mathbb{R}$ gilt*

$$\int_a^b (\alpha f + \beta g)(x)\, dx = \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx\,, \tag{11.13}$$

$$\left| \int_a^b f(x)\, dx \right| \le \int_a^b |f(x)|\, dx \le (b-a)\|f\|_\infty\,. \tag{11.14}$$

Ist $f \le g$, so gilt

$$\int_a^b f(x)\, dx \le \int_a^b g(x)\, dx\,. \tag{11.15}$$

Das Produkt fg zweier Regelfunktionen $f,g : [a,b] \to \mathbb{R}$ ist ebenfalls eine Regelfunktion.

Beweis: Seien (φ_n), (ψ_n) Folgen in $T[a,b]$ mit $\varphi_n \to f$ und $\psi_n \to g$ gleichmäßig. Dann gilt

$$\|(\alpha\varphi_n + \beta\psi_n) - (\alpha f + \beta g)\|_\infty \le \alpha\|\varphi_n - f\|_\infty + \beta\|\psi_n - g\|_\infty\,,$$

also gilt $\alpha\varphi_n + \beta\psi_n \to \alpha f + \beta g$ gleichmäßig. Daher sind Summen und skalare Vielfache von Regelfunktionen wiederum Regelfunktionen, und $R[a,b]$ ist ein Vektorraum. Analog zeigt man, dass $\varphi_n\psi_n \to fg$ gleichmäßig. Weiter gilt

$$\begin{aligned}
\int_a^b (\alpha f + \beta g)(x)\, dx &= \lim_{n\to\infty} \int_a^b (\alpha\varphi_n + \beta\psi_n)(x)\, dx \\
&= \alpha \lim_{n\to\infty} \int_a^b \varphi_n(x)\, dx + \beta \lim_{n\to\infty} \int_a^b \psi_n(x)\, dx \\
&= \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx\,.
\end{aligned}$$

Zum Beweis von (11.15) definieren wir

$$\tilde\varphi_n = \varphi_n - \|f - \varphi_n\|_\infty\,, \quad \tilde\psi_n = \psi_n + \|g - \psi_n\|_\infty\,.$$

Dann gilt $\tilde\varphi_n \le f$ und $g \le \tilde\psi_n$ für alle $n \in \mathbb{N}$, sowie $\tilde\varphi_n \to f$ und $\tilde\psi_n \to g$ gleichmäßig, also

$$\int_a^b \tilde\varphi_n(x)\, dx \le \int_a^b \tilde\psi_n(x)\, dx\,,$$

und Grenzübergang $n \to \infty$ liefert (11.15). (11.14) folgt aus (11.15). □

Bemerkung 11.12
Der größte Teil der Aussagen von Lemma 11.11 lässt sich zusammenfassen in dem Satz: Das Integral $I : (R[a,b], \|\cdot\|_\infty) \to \mathbb{R}$ ist eine lineare, stetige und monotone Abbildung. Die Linearität steht in (11.13). Aus (11.14) folgt, dass $f_n \to f$ bezüglich $\|\cdot\|_\infty$ die Konvergenz $I(f_n) \to I(f)$ zur Folge hat. Die Monotonie von I in (11.15) bezieht sich auf die durch den punktweisen Vergleich definierte Ordnungsrelation in $R[a,b]$.

Satz 11.13 *Jede stetige Funktion $f : [a,b] \to \mathbb{R}$ ist eine Regelfunktion. (In Kurzform: $C[a,b] \subset R[a,b]$.)*

Beweis: Sei $f \in C[a,b]$, sei $n \in \mathbb{N}$. Wir betrachten die äquidistante Zerlegung Z_n von $[a,b]$, definiert durch

$$h = \frac{b-a}{n}, \quad x_j = a + jh, \quad Z_n = \{x_0, \ldots, x_n\}.$$

Wir definieren die Treppenfunktion φ_n durch $\varphi_n(b) = f(b)$ und

$$\varphi_n(x) = f(x_j), \quad \text{falls } x \in [x_j, x_{j+1}), \quad j \in \{0, \ldots, n-1\}.$$

Für $x \in [x_j, x_{j+1})$ gilt $\varphi_n(x) = \varphi_n(x_j) = f(x_j)$, also

$$|f(x) - \varphi_n(x)| = |f(x) - f(x_j)|, \quad |x - x_j| \le \frac{b-a}{n}. \tag{11.16}$$

Wir nehmen an, f sei keine Regelfunktion. Dann gibt es gemäß Lemma 11.8 ein $\varepsilon > 0$ mit $\|f - \varphi\|_\infty > \varepsilon$ für alle Treppenfunktionen φ. Es gibt also $t_n \in [a,b]$ mit

$$|f(t_n) - \varphi_n(t_n)| \ge \varepsilon.$$

Sei nun $\tilde{x}_n$ der linke Endpunkt desjenigen Teilintervalls von Z_n, zu dem t_n gehört. Gemäß (11.16) gilt

$$\varepsilon \le |f(t_n) - \varphi_n(t_n)| = |f(t_n) - f(\tilde{x}_n)|, \quad |t_n - \tilde{x}_n| \le \frac{b-a}{n}.$$

Nach dem Satz von Bolzano-Weierstraß gibt es eine Teilfolge $\{t_{n_k}\}$ mit $t_{n_k} \to t \in [a,b]$. Es folgt $\tilde{x}_{n_k} \to t$ und weiter

$$\varepsilon \le |f(t_{n_k}) - f(\tilde{x}_{n_k})| \;\to\; |f(t) - f(t)| = 0$$

falls $k \to \infty$, ein Widerspruch. Also ist f eine Regelfunktion. $\square$

Lemma 11.14 *Sei $a < b < c$, sei $f \in R[a,c]$. Dann sind $f|[a,b] \in R[a,b]$ und $f|[b,c] \in R[b,c]$, und es gilt*

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx. \tag{11.17}$$

Beweis: Sei (φ_n) Folge in $T[a,c]$ mit $\varphi_n \to f$ gleichmäßig, dann gilt $\varphi_n|[a,b] \in T[a,b]$ und $\varphi_n|[a,b] \to f|[a,b]$ gleichmäßig. Also ist $f|[a,b] \in R[a,b]$. Ebenso für $[b,c]$. Ist Z_n zugehörige Zerlegung zu $\varphi_n|[a,b]$ und Z_n' zugehörige Zerlegung zu $\varphi_n|[b,c]$, so ist $Z_n \cup Z_n'$ zugehörige Zerlegung zu φ_n, und aus Definition 11.5 folgt unmittelbar

$$\int_a^c \varphi_n(x)\,dx = \int_a^b (\varphi_n|[a,b])(x)\,dx + \int_b^c (\varphi_n|[b,c])(x)\,dx\,.$$

Grenzübergang $n \to \infty$ liefert die Behauptung. □

Definition 11.15 *Ist $f \in R[a,b]$, so definieren wir*

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx\,, \quad \textit{sowie} \quad \int_c^c f(x)\,dx = 0 \tag{11.18}$$

für alle $c \in [a,b]$.

Die in Lemma 11.11 und Lemma 11.14 formulierten Eigenschaften lassen sich entsprechend auf den Fall $b < a$ übertragen, insbesondere gilt (11.17) für beliebige $a, b, c \in \mathbb{R}$.

Satz 11.16 (Mittelwertsatz der Integralrechnung)
Seien $f \in C[a,b]$, $g \in R[a,b]$ mit $g \geq 0$. Dann gibt es ein $\xi \in [a,b]$ mit

$$\int_a^b f(x)g(x)\,dx = f(\xi)\int_a^b g(x)\,dx\,. \tag{11.19}$$

Beweis: Das Produkt fg ist gemäß Lemma 11.11 und Satz 11.13 eine Regelfunktion, somit ist das Integral auf der linken Seite von (11.19) definiert. Mit

$$m = \min_{x\in[a,b]} f(x)\,, \quad M = \max_{x\in[a,b]} f(x)\,,$$

gilt wegen $g \geq 0$

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

Wähle $y \in [m,M]$ mit

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

Wähle gemäß Zwischenwertsatz ein $\xi \in [a,b]$ mit $f(\xi) = y$, dann gilt (11.19). □

Folgerung 11.17 *Sei $f \in C[a,b]$. Dann gibt es ein $\xi \in [a,b]$ mit*

$$\int_a^b f(x)\,dx = (b-a)f(\xi)\,. \tag{11.20}$$

Beweis: Anwendung von Satz 11.16 mit $g = 1$. □

Satz 11.18 *Seien $f \in C[a,b]$, $c \in [a,b]$. Wir definieren $F : [a,b] \to \mathbb{R}$ durch*

$$F(x) = \int_c^x f(t)\,dt\,. \tag{11.21}$$

Ist $x \in (a,b)$, so ist F differenzierbar in x und

$$F'(x) = f(x)\,. \tag{11.22}$$

Beweis: Nach Satz 11.13 ist $f \in R[a,b]$, also integrierbar. Sei $h \in \mathbb{R}$ mit $h \neq 0$ und $x + h \in [a,b]$. Dann ist

$$\begin{aligned}\frac{F(x+h) - F(x)}{h} - f(x) &= \frac{1}{h}\left(\int_c^{x+h} f(t)\,dt - \int_c^x f(t)\,dt\right) - f(x) \\ &= \frac{1}{h}\int_x^{x+h} f(t)\,dt - f(x) = \frac{1}{h} f(\xi)h - f(x) = f(\xi) - f(x)\end{aligned} \tag{11.23}$$

für ein geeignetes $\xi \in [x, x+h]$, gemäß Mittelwertsatz der Integralrechnung (mit $g = 1$). Sei nun $h_n \to 0$, dann $\xi_n \to x$ für die zugehörige Folge (ξ_n), und damit

$$\frac{F(x+h) - F(x)}{h} - f(x) = f(\xi_n) - f(x) \;\to\; 0$$

für $n \to \infty$. Also ist F differenzierbar in x, und $F'(x) = f(x)$. □

Definition 11.19 (Stammfunktion)
Seien $f, F : [a,b] \to \mathbb{R}$, sei F stetig auf $[a,b]$ und differenzierbar auf (a,b). Gilt $F'(x) = f(x)$ für alle $x \in (a,b)$, so heißt F Stammfunktion von f.

Satz 11.20 *Ist $f \in C[a,b]$, so ist die durch (11.21) definierte Funktion F eine Stammfunktion von f.*

Beweis: Nach (11.13) gilt

$$|F(y) - F(x)| = \left|\int_x^y f(t)\,dt\right| \le \|f\|_\infty |y - x|\,, \quad \text{für alle } x, y \in [a,b].$$

Also ist F stetig (sogar lipschitzstetig) auf $[a,b]$. Aus Satz 11.18 folgt, dass F Stammfunktion von f ist. □

Lemma 11.21 *Seien $F, G \in C[a,b]$, sei F Stammfunktion von $f : [a,b] \to \mathbb{R}$. Dann ist G genau dann Stammfunktion von f, wenn $G - F$ konstant ist.*

Beweis: Ist $G - F$ konstant, so ist mit F auch G auf (a,b) differenzierbar, und $G' = (G - F)' + F' = F' = f$. Ist umgekehrt G Stammfunktion von f, so gilt $G' = f = F'$, also $(G - F)' = 0$ und damit $G - F$ konstant nach Folgerung 10.10. □

Satz 11.22 (Hauptsatz der Differential- und Integralrechnung)
Sei $f \in C[a,b]$, sei $F : [a,b] \to \mathbb{R}$ Stammfunktion von f. Dann gilt

$$\int_a^b f(x)\,dx = F(b) - F(a)\,. \tag{11.24}$$

Beweis: Nach Satz 11.18 wird durch

$$F_a(x) = \int_a^x f(t)\,dt$$

eine in $[a,b]$ stetige und in (a,b) differenzierbare Funktion $F_a : [a,b] \to \mathbb{R}$ mit $F_a'(x) = f(x)$, also eine Stammfunktion von f definiert, und es gilt

$$\int_a^b f(t)\,dt = F_a(b) = F_a(b) - F_a(a)\,.$$

Aus Lemma 11.21 folgt

$$F(b) - F_a(b) = F(a) - F_a(a)\,,$$

also gilt (11.24). □

Wir schreiben auch

$$\int_a^b f(x)\,dx = F(b) - F(a) = F\Big|_a^b = F(x)\Big|_{x=a}^{x=b}\,.$$

Bemerkung 11.23
Oft werden die Aussagen von Satz 11.18, Lemma 11.21 und Satz 11.22 zusammengenommen als "Hauptsatz der Differential- und Integralrechnung" bezeichnet. □

Mit dem Hauptsatz erhalten wir aus jeder Formel für die Differentiation eine Formel für die Integration: Ist F gegeben und $f = F'$, so können wir mit dem Hauptsatz das Integral von f berechnen.

Beispiel 11.24
1. Ist $F(x) = x^4$, also $F'(x) = 4x^3$, so ist

$$\int_1^3 x^3\,dt = \frac{1}{4}x^4\Big|_{x=1}^{x=3} = \frac{1}{4}3^4 - \frac{1}{4}1^4 = 20\,.$$

2. Es gilt

$$(\ln)'(x) = \frac{1}{x}\,, \quad x > 0\,,$$

also

$$\int_1^2 \frac{1}{x}\,dx = \ln 2 - \ln 1 = \ln 2\,.$$

Ebenso erhalten wir aus jeder Rechenregel für die Differentiation eine Rechenregel für die Integration.

Satz 11.25 (Substitutionsregel)
Sei $I \subset \mathbb{R}$ Intervall, $f : I \to \mathbb{R}$ stetig. Sei $g \in C[a,b]$, stetig differenzierbar in (a,b), es gelte $g([a,b]) \subset I$. Dann gilt

$$\int_a^b f(g(t))g'(t)\,dt = \int_{g(a)}^{g(b)} f(x)\,dx\,. \tag{11.25}$$

Beweis: Sei F Stammfunktion von f auf dem Intervall $J = g([a,b])$. Dann ist $F \circ g : [a,b] \to \mathbb{R}$ stetig und in (a,b) differenzierbar. Aus der Kettenregel folgt

$$(F \circ g)'(t) = F'(g(t))g'(t) = f(g(t))g'(t)\,, \quad t \in (a,b)\,.$$

Zweimalige Anwendung des Hauptsatzes liefert nun

$$\int_a^b f(g(t))g'(t)\,dt = (F \circ g)(b) - (F \circ g)(a) = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(x)\,dx\,.$$

□

Beispiel 11.26

1. Wir betrachten
$$\int_0^{2\pi} 2t\cos(t^2)\,dt\,.$$
Wir setzen $f(x) = \cos x$, $g(t) = t^2$. Es ist $g'(t) = 2t$, $a = 0$, $b = 2\pi$, und wir erhalten
$$\int_0^{2\pi} 2t\cos(t^2)\,dt = \int_0^{(2\pi)^2} \cos x\,dx = \sin x\Big|_{x=0}^{x=4\pi^2} = \sin(4\pi^2)\,.$$

2. $$\int_a^b f(t+c)\,dt = \int_{a+c}^{b+c} f(x)\,dx\,,$$
falls $f : [a+c, b+c] \to \mathbb{R}$ stetig ist, $c \in \mathbb{R}$. (Substitution $g(t) = t + c$.)

3. $$\int_a^b f(ct)\,dt = \frac{1}{c}\int_{ac}^{bc} f(x)\,dx\,,$$
falls $f : [ac, bc] \to \mathbb{R}$ stetig ist, $c \neq 0$. (Substitution $g(t) = ct$.)

4. Ist g wie im Satz vorausgesetzt, es gelte $g(t) > 0$ für alle $t \in [a,b]$. Anwendung der Substitutionsregel mit $f(x) = 1/x$ liefert
$$\int_a^b \frac{g'(t)}{g(t)}\,dt = \ln(g(t))\Big|_{t=a}^{t=b}\,.$$
Für $g(t) = \cos t$ ergibt sich, falls $[a,b] \subset (-\pi/2, \pi/2)$,
$$\int_a^b \frac{-\sin t}{\cos t}\,dt = \ln(\cos t)\Big|_{t=a}^{t=b}\,,$$
also
$$\int_a^b \tan t\,dt = -\ln(\cos t)\Big|_{t=a}^{t=b}\,.$$

5. Die Substitution
$$g(t) = 2\arctan t\,, \quad g'(t) = \frac{2}{1+t^2}\,,$$
führt vermittels der Identität
$$\sin x = \frac{2\tan\frac{x}{2}}{1+\tan^2\frac{x}{2}}$$
auf
$$\sin(g(t)) = \frac{2t}{1+t^2}\,.$$
Hieraus ergibt sich zum Beispiel
$$\begin{aligned}\int_{g(a)}^{g(b)} \frac{1}{\sin x}\,dx &= \int_a^b \frac{1}{\sin(g(t))} g'(t)\,dt = \int_a^b \frac{1+t^2}{2t}\frac{2}{1+t^2}\,dt = \int_a^b \frac{1}{t}\,dt \\ &= \ln(b) - \ln(a)\,,\end{aligned}$$
falls etwa $0 < a < b$, und damit
$$\int_c^d \frac{1}{\sin x}\,dx = \ln\left(\tan\frac{x}{2}\right)\Big|_{x=c}^{x=d}, \quad 0 < c < d < \pi\,.$$

Satz 11.27 (Partielle Integration)
Seien $f, g \in C[a,b]$, stetig differenzierbar in (a,b). Dann gilt
$$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big|_{x=a}^{x=b} - \int_a^b f'(x)g(x)\,dx\,. \tag{11.26}$$

Beweis: Wir definieren $F : [a,b] \to \mathbb{R}$ durch $F(x) = f(x)g(x)$. Dann folgt aus der Produktregel
$$F'(x) = f'(x)g(x) + f(x)g'(x)\,,$$
also ist F Stammfunktion von $f'g + fg'$, und nach dem Hauptsatz gilt
$$\int_a^b f'(x)g(x) + f(x)g'(x)\,dx = F(x)\Big|_{x=a}^{x=b},$$
woraus (11.26) folgt. □

Beispiel 11.28

1. Wir berechnen
$$\int_a^b xe^x\,dx\,.$$
Wir setzen $f(x) = x$ und wählen ein g mit $g'(x) = e^x$, nämlich $g(x) = e^x$. Es ist dann
$$\int_a^b xe^x = xe^x\Big|_{x=a}^{x=b} - \int_a^b 1\cdot e^x\,dx = xe^x\Big|_{x=a}^{x=b} - e^x\Big|_{x=a}^{x=b} = (x-1)e^x\Big|_{x=a}^{x=b} = (b-1)e^b - (a-1)e^a\,.$$

2. Für $0 < a < b$ gilt

$$\int_a^b \ln x\, dx = \int_a^b 1 \cdot \ln x\, dx = x \ln x \Big|_{x=a}^{x=b} - \int_a^b 1\, dx = (x \ln x - x)\Big|_{x=a}^{x=b},$$

und

$$G(x) = x \ln(x) - x$$

ist eine Stammfunktion des Logarithmus.

3. Wir finden eine Rekursionsformel für

$$I_m = \int_a^b (\sin x)^m\, dx\,. \tag{11.27}$$

Es ist

$$I_0 = b - a\,, \quad I_1 = \cos a - \cos b\,.$$

Für $m \geq 2$ gilt

$$\begin{aligned} I_m &= \int_a^b (\sin x)^m\, dx = -\int_a^b (\sin x)^{(m-1)} (\cos)'(x)\, dx \\ &= -(\sin x)^{(m-1)} \cos x \Big|_{x=a}^{x=b} + (m-1) \int_a^b (\sin x)^{(m-2)} (\cos x)^2\, dx \\ &= -(\sin x)^{(m-1)} \cos x \Big|_{x=a}^{x=b} + (m-1) \int_a^b (\sin x)^{(m-2)} (1 - \sin^2 x)\, dx \\ &= -(\sin x)^{(m-1)} \cos x \Big|_{x=a}^{x=b} + (1-m) I_m + (m-1) I_{m-2}\,, \end{aligned} \tag{11.28}$$

also

$$I_m = \frac{m-1}{m} I_{m-2} - \frac{1}{m} (\sin x)^{(m-1)} \cos x \Big|_{x=a}^{x=b}.$$

Satz 11.29 *Sei $f \in C[a,b]$, $f \geq 0$, $f \neq 0$. Dann ist*

$$\int_a^b f(x)\, dx > 0\,. \tag{11.29}$$

Beweis: Sei $t \in [a,b]$ mit $f(t) > 0$. Wir setzen $\varepsilon = \frac{f(t)}{2}$, also $f(t) - \varepsilon > 0$. Gemäß Satz 7.14 gibt es ein $\delta > 0$ mit $f(x) - \varepsilon > 0$ für alle $x \in [a,b]$ mit $|x - t| < \delta$. Wir definieren $\varphi : [a,b] \to \mathbb{R}$ durch

$$\varphi(x) = \begin{cases} \varepsilon\,, & |x - t| < \delta\,, \\ 0\,, & \text{sonst.} \end{cases}$$

Dann gilt $\varphi(x) \leq f(x)$ für alle $x \in [a,b]$ und also

$$\int_a^b f(x)\, dx \geq \int_a^b \varphi(x)\, dx \geq \delta\varepsilon > 0\,.$$

□

Satz 11.30 *Sei $f : [a,b] \to \mathbb{R}$. Dann sind äquivalent:*

(1) $f \in R[a,b]$.

(2) *Für alle $x \in [a,b)$ existiert*

$$\lim_{\substack{\xi \to x \\ \xi > x}} f(\xi)\,, \tag{11.30}$$

und für alle $x \in (a,b]$ existiert

$$\lim_{\substack{\xi \to x \\ \xi < x}} f(\xi)\,. \tag{11.31}$$

Beweis:
"(1) $\Rightarrow$ (2)": Übungsaufgabe.
"(2) $\Rightarrow$ (1)": Es gelte (2). Wir nehmen an, f sei keine Regelfunktion. Dann gibt es ein $\varepsilon > 0$, so dass

$$\|f - \varphi\|_\infty > \varepsilon\,, \quad \text{für alle } \varphi \in T[a,b]. \tag{11.32}$$

Wir definieren eine Folge von Intervallen $I_n = [a_n, b_n]$, indem wir $I_0 = [a,b]$ setzen und, falls I_n bereits konstruiert ist,

$$I_{n+1} = \left[a_n, \frac{a_n + b_n}{2}\right] \quad \text{oder } I_{n+1} = \left[\frac{a_n + b_n}{2}, b_n\right] \tag{11.33}$$

setzen. Die Wahl wird dabei so getroffen, dass für alle $n \in \mathbb{N}$

$$\|f|I_n - \varphi\|_\infty > \varepsilon\,, \quad \text{für alle } \varphi \in T(I_n) \tag{11.34}$$

gilt. Das ist möglich: Für $n = 0$ gilt (11.34) wegen (11.32); gäbe es für beide Wahlmöglichkeiten von I_{n+1} ein $\varphi \in T(I_{n+1})$ mit $\|f|I_{n+1} - \varphi\|_\infty \le \varepsilon$, so gäbe es auch ein $\varphi \in T(I_n)$ mit $\|f|I_n - \varphi\|_\infty \le \varepsilon$. Es ist dann

$$\bigcap_{n \in \mathbb{N}} I_n = \{x\}\,, \quad x = \sup_{n \in \mathbb{N}} a_n = \inf_{n \in \mathbb{N}} b_n\,.$$

Wir setzen (falls $a < x < b$, andernfalls wird nur einer der Grenzwerte betrachtet)

$$y_l = \lim_{\substack{\xi \to x \\ \xi < x}} f(\xi)\,, \quad y_r = \lim_{\substack{\xi \to x \\ \xi > x}} f(\xi)\,,$$

und wählen $n \in \mathbb{N}$ so, dass

$|f(t) - y_l| \le \varepsilon$ für alle $t \in [a_n, x)$,
$|f(t) - y_r| \le \varepsilon$ für alle $t \in (x, b_n]$.

Für $\varphi \in T(I_n)$, definiert durch $\varphi(x) = f(x)$ und

$$\varphi(t) = \begin{cases} y_l\,, & t < x\,, \\ y_r\,, & t > x\,, \end{cases}$$

gilt dann $\|f|I_n - \varphi\|_\infty \le \varepsilon$ im Widerspruch zu (11.34). $\square$

Aus Satz 11.30 folgt beispielsweise, dass die durch

$$f(x) = \begin{cases} 1\,, & x \in \mathbb{Q}\,, \\ 0\,, & x \notin \mathbb{Q}\,, \end{cases}$$

keine Regelfunktion ist.

Folgerung 11.31 *Sei $f : [a,b] \to \mathbb{R}$ monoton. Dann ist f Regelfunktion.*

Beweis: Nach Übungsaufgabe erfüllt jede monotone Funktion die Bedingung (2) aus Satz 11.30. $\square$

Satz 11.32 *Sei (f_n) Folge in $R[a,b]$, sei $f_n \to f$ gleichmäßig, $f \in R[a,b]$. Dann gilt*

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx\,. \tag{11.35}$$

Beweis: Es gilt

$$\begin{aligned} 0 \;&\le\; \left| \int_a^b f(x)\,dx - \int_a^b f_n(x)\,dx \right| \le \int_a^b |f(x) - f_n(x)|\,dx \le \int_a^b \|f - f_n\|_\infty \,dx \\ &=\; (b-a)\|f - f_n\|_\infty \to 0\,. \end{aligned}$$

$\square$

Bemerkung 11.33

1. Die Voraussetzung “$f \in R[a,b]$” in Satz 11.32 ist überflüssig, da sie bereits aus der gleichmäßigen Konvergenz $f_n \to f$ und $f_n \in R[a,b]$ folgt. (Man zeigt, dass f die in Satz 11.30 verlangten rechts- und linksseitigen Grenzwerte besitzt.)

2. Ersetzt man in Satz 11.32 die gleichmäßige Konvergenz durch die punktweise, so gilt (11.35) im allgemeinen nicht. Beispiel:

   $$f_n : [0,1] \to \mathbb{R}\,, \quad f_n(x) = \begin{cases} n^2 x\,, & x \in [0, \frac{1}{n}]\,, \\ 2n - n^2 x\,, & x \in [\frac{1}{n}, \frac{2}{n}]\,, \\ 0\,, & \text{sonst.} \end{cases}$$

   Es gilt $f_n \to 0$ punktweise in $[0,1]$, aber

   $$\int_0^1 f_n(x)\,dx = 1$$

   für alle $n \in \mathbb{N}$.

Satz 11.34 *Sei (f_n) Folge in $C[a,b]$, seien alle f_n differenzierbar auf (a,b) und f_n' stetig fortsetzbar auf $[a,b]$. Seien f_n' gleichmäßig konvergent gegen eine Funktion $g : [a,b] \to \mathbb{R}$. Es gebe ein $x_0 \in [a,b]$, so dass $(f_n(x_0))$ konvergent ist. Dann gibt es ein $f \in C[a,b]$ mit $f_n \to f$ gleichmäßig und $f' = g$ in (a,b).*

Beweis: Zunächst ist g stetig nach Satz 10.18. Aus dem Hauptsatz folgt

$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(t)\, dt$$

für alle $x \in [a,b]$. Wir setzen

$$c = \lim_{n\to\infty} f_n(x_0)\,, \quad f(x) = c + \int_{x_0}^{x} g(t)\, dt\,.$$

Dann ist $f \in C[a,b]$ und $f' = g$ in (a,b) nach dem Hauptsatz, und es gilt für alle $x \in [a,b]$

$$\begin{aligned} |f(x) - f_n(x)| &= \left| c + \int_{x_0}^{x} g(t)\, dt - f_n(x_0) - \int_{x_0}^{x} f_n'(t)\, dt \right| \\ &\le |c - f_n(x_0)| + \int_{x_0}^{x} |g(t) - f_n'(t)|\, dt \\ &\le |c - f_n(x_0)| + |x_0 - x|\, \|g - f_n'\|_\infty\,, \end{aligned}$$

also

$$0 \le \|f - f_n\|_\infty \le |c - f_n(x_0)| + (b-a)\|g - f_n'\|_\infty\,,$$

woraus die Behauptung folgt. □

Definition 11.35 (Uneigentliche Integrale)
Sei $f : [a,\infty) \to \mathbb{R}$ mit $f \in R[a,M]$ für alle $M > a$. Wir definieren

$$\int_a^\infty f(x)\, dx = \lim_{M\to\infty} \int_a^M f(x)\, dx\,, \tag{11.36}$$

falls der Grenzwert auf der rechten Seite existiert.

Analog wird

$$\int_{-\infty}^{a} f(x)\, dx = \lim_{M\to-\infty} \int_M^a f(x)\, dx$$

definiert. Solche Integrale heißen uneigentliche Integrale.

Beispiel 11.36
Wir betrachten

$$\int_1^\infty \frac{1}{x^\alpha}\, dx\,, \quad \alpha > 1\,.$$

Es ist

$$\int_1^M \frac{1}{x^\alpha}\, dx = \frac{1}{1-\alpha} x^{1-\alpha}\Big|_{x=1}^{x=M} = \frac{M^{1-\alpha} - 1}{1-\alpha}\,,$$

also

$$\int_1^\infty \frac{1}{x^\alpha}\, dx = \lim_{M\to\infty} \frac{M^{1-\alpha} - 1}{1-\alpha} = \frac{1}{\alpha - 1}\,.$$

Für $\alpha = 1$ gilt

$$\int_1^M \frac{1}{x}\, dx = \ln(M) \to \infty\,, \quad \text{falls } M \to \infty\,,$$

also existiert $\int_1^\infty \frac{1}{x}\, dx$ nicht.

Definition 11.37 *Sei $f : (a, b] \to \mathbb{R}$, $f \in R[a + \varepsilon, b]$ für alle $\varepsilon > 0$. Wir definieren*

$$\int_a^b f(x)\,dx = \lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} \int_{a+\varepsilon}^b f(x)\,dx\,, \tag{11.37}$$

falls der Grenzwert auf der rechten Seite existiert.

Beispiel 11.38
Wir betrachten

$$\int_0^1 \frac{1}{x^\alpha}\,dx\,, \quad \alpha < 1\,.$$

Es ist

$$\int_\varepsilon^1 \frac{1}{x^\alpha}\,dx = \frac{1}{1-\alpha} x^{1-\alpha}\Big|_{x=\varepsilon}^{x=1} = \frac{1-\varepsilon^{1-\alpha}}{1-\alpha}\,,$$

also

$$\int_0^1 \frac{1}{x^\alpha}\,dx = \lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} \frac{1-\varepsilon^{1-\alpha}}{1-\alpha} = \frac{1}{1-\alpha}\,.$$

Für $\alpha = 1$ gilt

$$\int_\varepsilon^1 \frac{1}{x}\,dx = -\ln(\varepsilon) \to \infty\,, \quad \text{falls } \varepsilon \to 0\,,$$

also existiert $\int_0^1 \frac{1}{x}\,dx$ nicht.

Die eben beschriebenen Situationen können an beiden Integrationsgrenzen auftreten. Ist etwa $f : \mathbb{R} \to \mathbb{R}$ und $f \in R[a, b]$ für alle Intervalle $[a, b]$, so definieren wir

$$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^c f(x)\,dx + \int_c^\infty f(x)\,dx\,, \tag{11.38}$$

falls für ein $c \in \mathbb{R}$ beide Integrale auf der rechten Seite existieren. (Sie existieren dann für alle $c \in \mathbb{R}$, und die Summe hängt nicht von der Wahl von c ab.)

Beispiel:

$$\begin{aligned}
\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx &= \int_{-\infty}^0 \frac{1}{1+x^2}\,dx + \int_0^\infty \frac{1}{1+x^2}\,dx \\
&= \lim_{M\to-\infty} (-\arctan M) + \lim_{M\to\infty} \arctan M \\
&= \frac{\pi}{2} + \frac{\pi}{2} = \pi\,.
\end{aligned} \tag{11.39}$$

Vorsicht! Die Existenz von

$$\lim_{M\to\infty} \int_{-M}^M f(x)\,dx$$

ist **nicht** hinreichend für die Existenz von $\int_{-\infty}^\infty f(x)\,dx$ im Sinne von (11.38), so ist etwa

$$\int_{-M}^M x\,dx = 0$$

für alle M, aber

$$\int_0^M x\,dx = \frac{1}{2}M^2 \to \infty \quad \text{für } M \to \infty.$$

Satz 11.39 (Integralvergleichskriterium)
Sei $f : [1, \infty) \to \mathbb{R}$ monoton fallend, $f \geq 0$. Dann gilt

$$\int_1^\infty f(x)\,dx \quad \textit{existiert} \qquad \Leftrightarrow \qquad \sum_{k=1}^\infty f(k) \quad \textit{ist konvergent}. \tag{11.40}$$

Beweis: Nach Folgerung 11.31 ist $f \in R[1, M]$ für alle $M > 1$. Für alle $k \in \mathbb{N}$ und alle $x \in [k, k+1]$ gilt

$$f(k) \geq f(x) \geq f(k+1)\,,$$

also auch

$$f(k) \geq \int_k^{k+1} f(x)\,dx \geq f(k+1)$$

für alle k, also gilt für alle $n \in \mathbb{N}$, $n \geq 2$,

$$\sum_{k=1}^{n-1} f(k) \geq \int_1^n f(x)\,dx \geq \sum_{k=2}^n f(k)\,. \tag{11.41}$$

Aus der rechten Ungleichung folgt: Existiert

$$\lim_{n\to\infty} \int_1^n f(x)\,dx\,,$$

so ist wegen $f \geq 0$ die Zahl

$$\lim_{n\to\infty} \int_1^n f(x)\,dx = \sup_{n\in\mathbb{N}} \int_1^n f(x)\,dx$$

eine obere Schranke für $\sum_{k=2}^n f(k)$, also ist die Reihe konvergent. Aus der linken Ungleichung in (11.41) folgt: Ist die Reihe konvergent, so ist die monoton wachsende Folge

$$a_n = \int_1^n f(x)\,dx$$

beschränkt, also konvergent, und wegen

$$a_n \leq \int_1^M f(x)\,dx \leq a_{n+1} \quad \text{für alle } M \in [n, n+1]$$

folgt auch die Existenz von $\int_1^\infty f(x)\,dx$ im Sinne von Definition 11.35. □

Beispiel 11.40
Wir betrachten für $s > 1$ die Funktion

$$f : [1, \infty) \to \mathbb{R}\,, \quad f(x) = \frac{1}{x^s}\,.$$

Nach Beispiel 11.36 existiert $\int_1^\infty f(x)\,dx$, also ist nach Satz 11.39 die Reihe

$$\sum_{k=1}^\infty \frac{1}{k^s}$$

konvergent für $s > 1$. Die Funktion $\zeta : (1, \infty) \to \mathbb{R}$,

$$\zeta(s) = \sum_{k=1}^\infty \frac{1}{k^s}\,,$$

heißt die Riemannsche Zetafunktion.

12 Potenzreihen, Taylorreihen

Wir haben bereits Beispiele für Potenzreihen kennengelernt, nämlich die Exponentialreihe

$$\sum_{k=0}^{\infty} \frac{z^k}{k!}$$

und die daraus entstehenden Reihen für Sinus und Cosinus.

Eine allgemeine Potenzreihe hat die Form

$$f(z) = \sum_{k=0}^{\infty} c_k (z-a)^k \,, \tag{12.1}$$

mit dem **Entwicklungspunkt** $a \in \mathbb{C}$ und Koeffizienten $c_k \in \mathbb{C}$. Es stellt sich die Frage: Für welche $z \in \mathbb{C}$ ist (12.1) konvergent, oder anders formuliert, was ist der Definitionsbereich der durch

$$f(z) = \sum_{k=0}^{\infty} c_k (z-a)^k \tag{12.2}$$

definierten Funktion? Die Antwort ist: Eine Kreisscheibe.

Notation 12.1 *Sei $a \in \mathbb{C}$, $r > 0$. Dann heißt*

$$B(a,r) = \{z : z \in \mathbb{C},\ |z-a| < r\}$$

die offene Kreisscheibe um a mit Radius r, und

$$K(a,r) = \{z : z \in \mathbb{C},\ |z-a| \le r\}$$

die abgeschlossene Kreisscheibe um a mit Radius r.

Definition 12.2 (Konvergenzradius)
Die Zahl

$$r = \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|c_k|}} \tag{12.3}$$

heißt der Konvergenzradius der Potenzreihe (12.1).

Der folgende Satz rechtfertigt die Bezeichnung 'Konvergenzradius'.

Satz 12.3 *Sei (c_k) Folge in $\mathbb{C}$, $a \in \mathbb{C}$, sei r die in (12.3) definierte Zahl. Die Potenzreihe $\sum_{k=0}^{\infty} c_k (z-a)^k$*

konvergiert absolut, falls $|z-a| < r$,
divergiert, falls $|z-a| > r$.

Beweis: Der Beweis verwendet das Wurzelkriterium aus Satz 6.17 für die Konvergenz einer Reihe

$$\sum_{k=n_0}^{\infty} a_k \,.$$

Es besagt: Ist die durch

$$w = \limsup_{k\to\infty} \sqrt[k]{|a_k|}$$

definierte Zahl kleiner als 1, so konvergiert die Reihe absolut; ist sie größer als 1, so divergiert sie. Wir wenden diesen Satz an mit $a_k = c_k(z-a)^k$. Es gilt $\sqrt[k]{|a_k|} = \sqrt[k]{|c_k|}\cdot|z-a|$, also

$$w = \limsup_{k\to\infty} \sqrt[k]{|a_k|} = |z-a| \cdot \limsup_{k\to\infty} \sqrt[k]{|c_k|} \,,$$

also

$$w < 1 \qquad \Leftrightarrow \qquad |z-a| < \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|c_k|}} \,, \tag{12.4}$$

$$w > 1 \qquad \Leftrightarrow \qquad |z-a| > \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|c_k|}} \,, \tag{12.5}$$

also folgen die Behauptungen. □

Wir stellen uns nun die Frage, ob diese Konvergenz gleichmäßig ist.

Satz 12.4 (Weierstraß-Kriterium)
Sei $D \subset \mathbb{K}$, sei $(f_k)_{k\in\mathbb{N}}$ eine Folge von Funktionen in $B(D;\mathbb{K})$, es gelte

$$\sum_{k=0}^{\infty} \|f_k\|_\infty < \infty \,. \tag{12.6}$$

Dann ist für alle $x \in D$ die Reihe

$$\sum_{k=0}^{\infty} f_k(x) \tag{12.7}$$

absolut konvergent, und die Partialsummen $s_n : D \to \mathbb{K}$,

$$s_n = \sum_{k=0}^{n} f_k \,, \tag{12.8}$$

konvergieren gleichmäßig gegen $f : D \to \mathbb{K}$,

$$f(x) = \sum_{k=0}^{\infty} f_k(x) \,. \tag{12.9}$$

Beweis: Für alle $x \in D$ gilt $|f_k(x)| \le \|f_k\|_\infty$, also auch

$$\sum_{k=0}^{n} |f_k(x)| \le \sum_{k=0}^{n} \|f_k\|_\infty \le \sum_{k=0}^{\infty} \|f_k\|_\infty < \infty \,,$$

also ist (12.7) absolut konvergent für alle $x \in D$. Sei nun $\varepsilon > 0$. Aus (12.6) folgt, daß es ein $n_0 \in \mathbb{N}$ gibt mit

$$\sum_{k=n_0+1}^{\infty} \|f_k\|_\infty < \varepsilon\,.$$

Dann gilt für alle $n \geq n_0$ und alle $x \in D$

$$|s_n(x) - f(x)| = \left|\sum_{k=n+1}^{\infty} f_k(x)\right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} \|f_k\|_\infty < \varepsilon\,.$$

Also konvergiert s_n gleichmäßig gegen f. □

Beispiel: Für $f_k : \mathbb{R} \to \mathbb{R}$,

$$f_k(x) = \frac{\cos kx}{k^2}\,,$$

gilt

$$\|f_k\|_\infty \leq \frac{1}{k^2}\,,$$

also

$$\sum_{k=1}^{\infty} \|f_k\|_\infty \leq \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty\,,$$

also ist Satz 12.4 anwendbar und damit die Reihe

$$\sum_{k=1}^{\infty} \frac{\cos kx}{k^2}$$

absolut und in $\mathbb{R}$ gleichmäßig konvergent.

Satz 12.5 *Sei (c_k) Folge in $\mathbb{C}$, sei $a \in \mathbb{C}$, sei*

$$\sum_{k=0}^{\infty} c_k (z-a)^k$$

eine Potenzreihe mit Konvergenzradius $r > 0$. Dann konvergieren für jedes $\rho < r$ die Partialsummen

$$s_n(z) = \sum_{k=0}^{n} c_k (z-a)^k \tag{12.10}$$

auf $K(a,\rho)$ gleichmäßig gegen

$$f(z) = \sum_{k=0}^{\infty} c_k (z-a)^k\,. \tag{12.11}$$

Die Funktion f ist stetig auf $B(a,r)$.

Beweis: Für $z \in \mathbb{C}$ mit $|z-a| = \rho$ gilt

$$\sum_{k=0}^{\infty} |c_k| \rho^k = \sum_{k=0}^{\infty} |c_k|\; |z-a|^k < \infty\,, \tag{12.12}$$

da die Potenzreihe nach Satz 12.3 dort absolut konvergiert. Wir setzen

$$f_k(z) = c_k(z-a)^k \,,$$

dann gilt für alle z mit $|z-a| \le \rho$

$$|f_k(z)| \le |c_k|\rho^k \,,$$

also wegen (12.12)

$$\sum_{k=0}^{\infty} \|f_k\|_\infty \le \sum_{k=0}^{\infty} |c_k|\rho^k < \infty \,,$$

wobei die Supremumsnorm auf $D = K(a,\rho)$ betrachtet wird. Aus dem Weierstraß-Kriterium Satz 12.4 folgt, dass $s_n \to f$ gleichmäßig auf $K(a,\rho)$. Da alle s_n stetig sind, ist nach Satz 10.18 auch f stetig auf $K(a,\rho)$ und damit auch auf

$$B(a,r) = \bigcup_{0<\rho<r} K(a,\rho) \,.$$

□

Als nächstes wollen wir zeigen, dass eine Potenzreihe im Innern ihres Konvergenzintervalls differenzierbar ist und dass wir ihre Ableitung durch gliedweises Differenzieren erhalten können.

Satz 12.6 *Sei* (c_k) *Folge in* $\mathbb{R}$, $a \in \mathbb{R}$, *sei*

$$\sum_{k=0}^{\infty} c_k(x-a)^k \tag{12.13}$$

Potenzreihe mit Konvergenzradius $r > 0$. *Dann ist die durch*

$$f(x) = \sum_{k=0}^{\infty} c_k(x-a)^k \tag{12.14}$$

definierte Funktion $f : (a-r, a+r) \to \mathbb{R}$ *stetig differenzierbar in* $(a-r, a+r)$, *und es gilt*

$$f'(x) = \sum_{k=1}^{\infty} kc_k(x-a)^{k-1} \,. \tag{12.15}$$

Diese Potenzreihe hat ebenfalls den Konvergenzradius r, *sie konvergiert absolut in* $(a-r, a+r)$, *und für jedes* $\rho < r$ *konvergieren die Partialsummen gleichmäßig auf* $[a-\rho, a+\rho]$.

Beweis: Variante 1: Wir definieren für $k \ge 1$

$$g_k(x) = kc_k(x-a)^{k-1} \,.$$

Ist $\rho < r$, so gilt für die Supremumsnorm auf $[a-\rho, a+\rho]$

$$\|g_k\|_\infty \le k|c_k|\rho^{k-1} \,. \tag{12.16}$$

Sei x gewählt mit $\rho := |x-a| < r$, Sei weiter $\tilde{x}$ gewählt mit $\rho < |\tilde{x}-a| < r$. Es gibt $M > 0$ mit

$$|c_k|\,|\tilde{x}-a|^k \le M\,,$$

da die Reihe (12.13) in $\tilde{x}$ absolut konvergiert. Aus (12.16) folgt für alle x mit $|x-a| \le \rho$

$$|g_k(x)| \le k|c_k|\rho^{k-1} = \frac{k}{\rho}\left(\frac{\rho}{|\tilde{x}-a|}\right)^k |c_k|\,|\tilde{x}-a|^k \le \frac{k}{\rho}\left(\frac{\rho}{|\tilde{x}-a|}\right)^k M\,.$$

Es folgt weiter

$$\sum_{k=1}^{\infty} |g_k(x)| \le \sum_{k=1}^{\infty} \frac{M}{\rho} k \left(\frac{\rho}{|\tilde{x}-a|}\right)^k < \infty\,,$$

da die rechts stehende Reihe nach dem Quotientenkriterium konvergent ist. Die Reihe

$$\sum_{k=1}^{\infty} k c_k (x-a)^{k-1} \tag{12.17}$$

ist also für jedes x mit $|x-a| < r$ konvergent und hat daher Konvergenzradius $\ge r$. Ein analoges Argument zeigt die Divergenz für $|x-a| > r$, also ist der Konvergenzradius von (12.17) gleich r. Aus Satz 12.5 folgt die absolute Konvergenz auf $(a-r, a+r)$ und die gleichmäßige Konvergenz auf $[a-\rho, a+\rho]$ für jedes $\rho < r$. Wir wenden nun Satz 11.34 auf die Partialsummen der Reihen in (12.14) und (12.15) an und erhalten, dass f differenzierbar und f' durch (12.15) gegeben ist, und zwar auf jedem solchen Intervall $[a-\rho, a+\rho]$ und damit auf $(a-r, a+r)$.

Variante 2: Wir zeigen, dass für $0 < |x-a| < r$ und

$$\tilde{c_k} = \frac{k}{x-a} c_k$$

gilt, dass

$$\frac{1}{\limsup_{k\to\infty} \sqrt[k]{|\tilde{c}_k|}} = \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|c_k|}} = r\,. \tag{12.18}$$

Aus (12.18) folgt, dass die Potenzreihe (12.17) ebenfalls den Konvergenzradius r hat, und der Beweis geht weiter wie bei Variante 1. Um (12.18) zu beweisen, verwendet man

$$\lim_{k\to\infty} \sqrt[k]{|x-a|} = 1\,, \quad \lim_{k\to\infty} \sqrt[k]{k} = 1\,,$$

letzteres gilt, da $\ln(\sqrt[k]{k}) = (\ln k)/k \to 0$ für $k \to \infty$ und daher

$$\sqrt[k]{k} = \exp(\ln(\sqrt[k]{k})) \;\to\; \exp(0) = 1\,.$$

Es ergibt sich dann

$$\limsup_{k\to\infty} \sqrt[k]{|\tilde{c}_k|} = \frac{1}{\lim_{k\to\infty} \sqrt[k]{|x-a|}} \cdot \big(\lim_{k\to\infty} \sqrt[k]{k}\big) \cdot \limsup_{k\to\infty} \sqrt[k]{|c_k|} = \limsup_{k\to\infty} \sqrt[k]{|c_k|}\,.$$

□

Beispiel 12.7
Die die Funktion

$$f(x) = \sum_{k=0}^{\infty} x^k$$

definierende Potenzreihe ist konvergent für $|x| < 1$, divergent für $|x| > 1$, hat also Konvergenzradius $r = 1$. Für $|x| < 1$ gilt

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

Aus Satz 12.6 folgt nun

$$f'(x) = \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}\,.$$

Hieraus erhalten wir die Formel

$$\sum_{k=1}^{\infty} kx^k = x\sum_{k=1}^{\infty} kx^{k-1} = \frac{x}{(1-x)^2}\,.$$

□

Notation 12.8 *Sei I offenes Intervall in $\mathbb{R}$. Wir setzen*

$$C^n(I) = \{f | f : I \to \mathbb{R},\ f \text{ ist } n\text{-mal stetig differenzierbar}\}\,,$$
$$C^\infty(I) = \bigcap_{n=1}^{\infty} C^n(I)\,.$$

Die Elemente von $C^\infty(I)$ sind also diejenigen Funktionen, die auf I beliebig oft differenzierbar sind. Man sagt auch: Sie sind "unendlich oft differenzierbar".

Satz 12.9 *Sei (c_k) Folge in $\mathbb{R}$, $a \in \mathbb{R}$,*

$$\sum_{k=0}^{\infty} c_k (x-a)^k \tag{12.19}$$

Potenzreihe mit Konvergenzradius $r > 0$. Dann ist die durch

$$f(x) = \sum_{k=0}^{\infty} c_k (x-a)^k \tag{12.20}$$

definierte Funktion $f : (a-r, a+r) \to \mathbb{R}$ unendlich oft differenzierbar in $(a-r, a+r)$, und es gilt

$$c_k = \frac{1}{k!} f^{(k)}(a) \tag{12.21}$$

für alle $k \in \mathbb{N}$.

Beweis: Wir zeigen mit Induktion über k: f ist k-mal differenzierbar in $(a-r, a+r)$, und $f^{(k)}$ ist gegeben durch die in $(a-r, a+r)$ konvergente Potenzreihe

$$f^{(k)}(x) = \sum_{m=k}^{\infty} m(m-1)\cdots(m-k+1)c_m(x-a)^{m-k} . \tag{12.22}$$

Für $k = 1$ gilt (12.22) nach Satz 12.6. Gilt die Behauptung für k, so können wir Satz 12.6 auf $f^{(k)}$ anwenden und erhalten die Formel (12.22) für $k+1$ durch gliedweises Differenzieren. Die Formel (12.21) folgt, indem wir $x = a$ in (12.22) setzen. □

Für Potenzreihen

$$f(x) = \sum_{k=0}^{\infty} c_k(x-a)^k \tag{12.23}$$

mit Konvergenzradius $r > 0$ gilt also

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(x-a)^k , \quad x \in (a-r, a+r) . \tag{12.24}$$

Als Beispiel stellen wir den Logarithmus um den Punkt 1 als Potenzreihe dar. Zu diesem Zweck betrachten wir die Funktion

$$f(x) = \ln(1+x)$$

zum Entwicklungspunkt $a = 0$. Diese Funktion ist beliebig oft differenzierbar,

$$f'(x) = \frac{1}{1+x} , \quad f''(x) = -\frac{1}{(1+x)^2} , \quad f'''(x) = \frac{2}{(1+x)^3} ,$$

allgemein

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

Gemäß (12.21) erhalten wir für die Koeffizienten der zugehörigen Potenzreihe

$$c_k = (-1)^{(k+1)} \frac{1}{k} ,$$

also wegen $f(0) = \ln(1) = 0$

$$\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{(k+1)} \frac{x^k}{k} . \tag{12.25}$$

Die Potenzreihe hat Konvergenzradius 1, da $\sqrt[k]{k} \to 1$ für $k \to \infty$. Sie divergiert für $x = -1$ (harmonische Reihe) und konvergiert für $x = 1$ (alternierende harmonische Reihe).

Taylorreihen. Ausgangspunkt ist die Formel (12.24), welche es (jedenfalls für Potenzreihen) erlaubt, den Wert $f(x)$ einer Funktion durch die Werte von f und ihren Ableitungen im Punkt a darzustellen.

Definition 12.10 (Taylorreihe)
Sei $a \in \mathbb{R}$, sei $I \subset \mathbb{R}$ offenes Intervall mit $a \in I$, sei $f \in C^\infty(I)$. Die Reihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{12.26}$$

heißt die Taylorreihe von f mit Entwicklungspunkt a. Falls sie den Konvergenzradius $r > 0$ hat, so definiert

$$T_f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{12.27}$$

eine Funktion $T_f : (a-r, a+r) \to \mathbb{R}$.

Bemerkung 12.11

1. Der Konvergenzradius der Potenzreihe (12.26) kann 0 sein. (In diesem Fall können wir $T_f(x)$ lediglich für $x = a$ definieren, $T_f(a) = f(a)$.)

2. Ist f durch eine Potenzreihe definiert, so gilt nach Satz 12.9

$$f = T_f \, .$$

   So ist etwa

$$\exp(x) = T_{\exp}(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \, .$$

3. Für $f \in C^\infty(I)$ ist es möglich, daß $f \neq T_f$. Beispiel:

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x}\right) \, , & x > 0 \, , \\ 0 \, , & x \leq 0 \, . \end{cases}$$

   Es ist $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}$ (siehe Übungsaufgabe), also ist $T_f = 0$ im Entwicklungspunkt $a = 0$.

□

Die Approximation einer Funktion f durch die Partialsummen ihrer Taylorreihe

$$T_f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_{n+1}(x) \tag{12.28}$$

mit einem Restglied R_{n+1} ist ein wichtiges Werkzeug der Analysis. Sie ist bereits dann nützlich und sinnvoll, wenn nur die ersten n bzw. $n+1$ Ableitungen von f existieren.

Definition 12.12 (Taylorpolynom)
Sei $I \subset \mathbb{R}$ offenes Intervall, $f \in C^n(I)$ und $a \in I$. Dann heißt das durch

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{12.29}$$

definierte Polynom das n-te Taylorpolynom von f in a.

Satz 12.13 (Taylorentwicklung)
Sei $I \subset \mathbb{R}$ offenes Intervall, $f \in C^{n+1}(I)$ und $a \in I$. Dann gilt für alle $x \in I$

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_{n+1}(x) \tag{12.30}$$

mit

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt\,. \tag{12.31}$$

Beweis: Induktion über n. Für $n = 0$ wird die Behauptung zu

$$f(x) = f(a) + \int_a^x f'(t)\,dt$$

was nach dem Hauptsatz richtig ist. Zum Beweis von $n-1 \to n$ sei

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + R_n(x)\,, \tag{12.32}$$

$$R_n(x) = \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)\,dt\,. \tag{12.33}$$

Es gilt dann mit partieller Integration

$$\begin{aligned} R_n(x) &= \frac{1}{(n-1)!}\left[-\frac{1}{n}(x-t)^n f^{(n)}(t)\Big|_{t=a}^{t=x} - \int_a^x -\frac{1}{n}(x-t)^n f^{(n+1)}(t)\,dt\right] \\ &= \frac{1}{n!}(x-a)^n f^{(n)}(a) + \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt\,. \end{aligned}$$

Einsetzen in (12.32) liefert die Behauptung. □

Satz 12.14 *Seien die Voraussetzungen von Satz 12.13 erfüllt. Dann gibt es zu jedem $x \in I$ ein ξ zwischen a und x mit*

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}\,. \tag{12.34}$$

Beweis: Da $f^{(n+1)}$ stetig ist und die durch

$$g(t) = \frac{(x-t)^n}{n!}$$

definierte Funktion zwischen a und x nicht das Vorzeichen wechselt, können wir den Mittelwertsatz der Integralrechnung (Satz 11.16) anwenden und erhalten, daß es ein ξ zwischen a und x gibt mit

$$\begin{aligned} R_{n+1}(x) &= \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\,dt = \frac{1}{n!} f^{(n+1)}(\xi)\int_a^x (x-t)^n\,dt \\ &= \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}\,. \end{aligned}$$

□

Folgerung 12.15 *Seien die Voraussetzungen von Satz 12.13 erfüllt. Gilt dann zusätzlich $f^{(n+1)} = 0$ in I, so ist f ein Polynom vom Grad $\leq n$.*

Beweis: In diesem Fall ist $R_{n+1} = 0$ in I. □

Folgerung 12.16 *Seien die Voraussetzungen von Satz 12.13 erfüllt. Dann gibt es eine Funktion $\eta : I \to \mathbb{R}$ mit*

$$f(x) = \sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \eta(x)(x-a)^{n+1} \tag{12.35}$$

und

$$\lim_{x\to a} \eta(x) = 0\,. \tag{12.36}$$

Beweis: Nach Satz 12.14 finden wir zu jedem $x \in I$ ein $\xi(x)$ mit $|\xi(x) - a| \leq |x - a|$ und

$$f(x) - \sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!}(x-a)^k = \frac{f^{(n+1)}(\xi(x)) - f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1}\,.$$

Wir definieren

$$\eta(x) = \frac{f^{(n+1)}(\xi(x)) - f^{(n+1)}(a)}{(n+1)!}\,,$$

dann gilt (12.36), da $f^{(n+1)}$ stetig ist und $\lim_{x\to a} \xi(x) = a$. □

Ist also $f \in C^n(I)$ und T_n das n-te Taylorpolynom von f in a, so geht der Fehler $|f(x) - T_n(x)|$ für $x \to a$ schneller gegen 0 als $(x-a)^n$.

Definition 12.17 (Klein-O und Groß-O)
Sei $\delta > 0$, $I = (-\delta, \delta)$ und $f, g : I \to \mathbb{R}$ Funktionen. Wir sagen

$$f(h) = O(g(h))\,, \tag{12.37}$$

falls es ein $C > 0$ gibt mit

$$|f(h)| \leq C|g(h)|\,, \quad \text{für alle } h \in I. \tag{12.38}$$

Wir sagen

$$f(h) = o(g(h))\,, \tag{12.39}$$

falls $g(h) \neq 0$ für alle $h \neq 0$ und

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(h)}{g(h)} = 0\,. \tag{12.40}$$

Beispiel 12.18
Sei T_n das n-te Taylorpolynom von f in $a \in I$, I Intervall, sei

$$r(h) = f(a+h) - T_n(a+h)\,.$$

Dann gilt nach Folgerung 12.16

$$r(h) = o(h^n)\,, \quad \text{falls } f \in C^n(I),$$

und

$$r(h) = O(h^{n+1})\,, \quad \text{falls } f \in C^{n+1}(I),$$

da für eine geeignete Zwischenstelle ξ

$$r(h) = R_{n+1}(a+h) = \frac{f^{(n+1)}(\xi)}{(n+1)!} h^{n+1}\,,$$

also für $|h| \le h_0$, $h_0 > 0$ fest,

$$|r(h)| \le C|h|^{n+1}\,, \quad C = \frac{1}{(n+1)!} \sup_{|\xi - a| \le |h_0|} |f^{(n+1)}(\xi)|\,.$$

Beispiel 12.19

1. Die Taylorreihe des Sinus mit Entwicklungspunkt $a = 0$ ist

$$T_{\sin}(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}\,, \tag{12.41}$$

und für das Restglied folgt, da alle Ableitungen des Sinus durch 1 beschränkt sind, aus Satz 12.14

$$|R_n(x)| \le \frac{|x|^n}{n!}\,, \tag{12.42}$$

also gilt für alle $x \in \mathbb{R}$

$$\lim_{n\to\infty} R_n(x) = 0\,,$$

und damit auch

$$\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \tag{12.43}$$

für alle $x \in \mathbb{R}$. Schreiben wir wieder h statt x (da $a = 0$, macht das keinen Unterschied), so ist

$$\begin{aligned} \sin h &= h + O(h^3) \\ &= h - \frac{h^3}{3!} + O(h^5) \\ &= h - \frac{h^3}{3!} + \frac{h^5}{5!} + O(h^7)\,. \end{aligned}$$

2. Analog folgt für den Cosinus mit Entwicklungspunkt $a = 0$

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \tag{12.44}$$

für alle $x \in \mathbb{R}$.

13 Unendliche Mengen

Zwei endliche Mengen sind “gleich groß”, wenn sie gleichviele Elemente haben. Eine Menge von 3 Straßenbahnen ist in diesem Sinn genauso groß wie eine Menge von 3 Äpfeln, obwohl eine Straßenbahn deutlich größer als ein Apfel ist.

Wir wollen eine Definition haben, mit der man auch unendliche Mengen hinsichtlich ihrer “Größe” vergleichen kann.

Definition 13.1 *Zwei Mengen M und N heißen gleichmächtig, wenn es eine bijektive Abbildung $f : M \to N$ gibt.*

Diese Definition macht Sinn, wenn zwei endliche Mengen genau dann gleichmächtig sind, wenn sie gleichviele Elemente haben. Das ergibt sich im folgenden Lemma samt anschließender Definition.

Lemma 13.2 *Seien $m, n \in \mathbb{N}$. Die Mengen $\{0, \ldots, m\}$ und $\{0, \ldots, n\}$ sind gleichmächtig genau dann, wenn $m = n$.*

Beweis: Ist $m = n$, so können wir für f die Identität wählen. Sei nun $m \neq n$. Wir zeigen mit Induktion über n die Behauptung:

> Ist $m > n$, so gibt es keine bijektive Abbildung $f : \{0, \ldots, m\} \to \{0, \ldots, n\}$.

Induktionsanfang $n = 0$: Die einzige Abbildung $f : \{0, \ldots, m\} \to \{0\}$ ist gegeben durch $f(k) = 0$ für alle k, und diese Abbildung ist für $m > 0$ nicht bijektiv.
Induktionsschritt $n - 1 \to n$: Es genügt zu zeigen:

> Ist $m > n$ und $f : \{0, \ldots, m\} \to \{0, \ldots, n\}$ bijektiv, so gibt es eine bijektive Abbildung $g : \{0, \ldots, m - 1\} \to \{0, \ldots, n - 1\}$.

Daraus folgt, dass es ein solches f nicht geben kann, da es nach Induktionsannahme ein solches g nicht geben kann.

Die Abbildung g wird aus f wie folgt konstruiert. Zunächst ist die Restriktion

$$\tilde{f} : \{0, \ldots, m - 1\} \to \{0, \ldots, n\} \setminus \{f(m)\}$$

von f auf $\{0, \ldots, m - 1\}$ bijektiv. Wir definieren eine Abbildung

$$h : \{0, \ldots, n\} \setminus \{f(m)\} \to \{0, \ldots, n - 1\}$$

durch

$$h(i) = \begin{cases} i\,, & \text{falls } i < f(m)\,, \\ i - 1\,, & \text{falls } i > f(m)\,. \end{cases}$$

Dann ist h ebenfalls bijektiv, und $g = h \circ \tilde{f}$ ist die gesuchte Bijektion. □

Die Anzahl der Elemente einer Menge M erhalten wir dadurch, dass wir sie in irgendeiner Reihenfolge abzählen. “Abzählen” bedeutet aber gerade, dass wir eine bijektive Abbildung von $\{0, \ldots, n - 1\}$ nach M definieren.

Definition 13.3 *Sei M Menge. Ist $n \geq 1$ und sind M und $\{0, \ldots, n-1\}$ gleichmächtig, so definieren wir die Anzahl der Elemente von M als n, geschrieben*

$$\#M = n\,.$$

Wir setzen $\#\emptyset = 0$. M heißt endlich, wenn es ein $n \in \mathbb{N}$ gibt mit $\#M = n$, andernfalls heißt M unendlich, und wir schreiben $\#M = \infty$.

Definition 13.4 (Abzählbare Menge)
Eine nichtleere Menge M heißt abzählbar, wenn es eine surjektive Abbildung $\tau : \mathbb{N} \to M$ gibt.

“M ist abzählbar” bedeutet also, dass es eine Folge $(x_n)_{n\in\mathbb{N}}$ in M gibt, in der jedes Element von M mindestens einmal auftaucht. Beispiele: Jede endliche nichtleere Menge ist abzählbar, $\mathbb{N}$ ist abzählbar. $\mathbb{Z}$ ist abzählbar: Die Abbildung $\tau : \mathbb{N} \to \mathbb{Z}$,

$$\tau(n) = \begin{cases} \frac{n}{2}\,, & n \text{ gerade }\,, \\ -\frac{n+1}{2}\,, & n \text{ ungerade }\,, \end{cases}$$

liefert die Folge $0, -1, 1, -2, 2, \ldots$ und ist surjektiv. Sie ist sogar bijektiv, $\mathbb{N}$ und $\mathbb{Z}$ sind also gleichmächtig. Andererseits ist $\mathbb{N}$ eine echte Teilmenge von $\mathbb{Z}$. (Dass eine unendliche Menge ihre “Mächtigkeit” nicht zu ändern braucht, wenn man Elemente aus ihr entfernt, klingt zunächst merkwürdig, ist aber so.)

Lemma 13.5 *Seien M, N Mengen und $f : M \to N$ Abbildung. Ist M abzählbar und f surjektiv, so ist auch N abzählbar.*

Beweis: Ist $\tau : \mathbb{N} \to M$ surjektiv, so ist auch $f \circ \tau : \mathbb{N} \to N$ surjektiv. □

Folgerung 13.6 *Sei M abzählbar, $N \subset M$. Dann ist N abzählbar.*

Beweis: Wähle als $f : M \to N$ irgendeine Fortsetzung der Identität auf N. □

Satz 13.7 $\mathbb{N} \times \mathbb{N}$ *ist abzählbar.*

Beweis: In der Folge

$$(0,0), (0,1), (1,0), (0,2), (1,1), (2,0), (0,3) \ldots$$

taucht jedes Element von $\mathbb{N} \times \mathbb{N}$ mindestens einmal (sogar genau einmal) auf. In Formeln können wir die entsprechende Abbildung $\tau : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ beschreiben durch

$$\tau(k) = (k - s_n, s_{n+1} - k - 1)\,, \quad \text{falls } s_n \leq k < s_{n+1}\,,$$

wobei

$$s_n = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}\,, \quad s_0 = 0\,.$$

Man kann nachrechnen, dass τ surjektiv ist. □

Folgerung 13.8 *Sind M und N abzählbare Mengen, so ist auch $M \times N$ abzählbar.*

Beweis: Seien $\tau_M : \mathbb{N} \to M$ und $\tau_N : \mathbb{N} \to N$ surjektiv. Dann ist auch

$$\tau : \mathbb{N} \times \mathbb{N} \to M \times N\,, \quad \tau(j,k) = (\tau_M(j), \tau_N(k))$$

surjektiv, also $M \times N$ abzählbar wegen Folgerung 13.5 und Satz 13.7. □

Folgerung 13.9 $\mathbb{Q}$ *ist abzählbar.*

Beweis: Die Abbildung

$$\tau : \mathbb{Z} \times \mathbb{N} \to \mathbb{Q}\,, \quad \tau(k,n) = \frac{k}{n+1}\,,$$

ist surjektiv. □

Folgerung 13.10 *Sei $(M_n)_{n\in\mathbb{N}}$ eine Folge von abzählbaren Mengen. Dann ist auch*

$$\bigcup_{n\in\mathbb{N}} M_n$$

abzählbar.

Beweis: Sei $\tau_n : \mathbb{N} \to M_n$ surjektiv für alle $n \in \mathbb{N}$. Dann ist auch

$$\tau : \mathbb{N} \times \mathbb{N} \to \bigcup_{n\in\mathbb{N}} M_n\,, \quad \tau(n,k) = \tau_n(k)\,,$$

surjektiv. □

Definition 13.11 (Überabzählbare Menge)
Eine unendliche Menge M heißt überabzählbar, wenn sie nicht abzählbar ist.

Wir wollen beweisen, dass $\mathbb{R}$ überabzählbar ist. Das ist ein berühmtes, auf Georg Cantor zurückgehendes Resultat aus den Anfangszeiten der Mengenlehre.

Wegen Folgerung 13.6 genügt es zu zeigen, dass

$$[0,1) = \{x : x \in \mathbb{R},\, 0 \le x < 1\}$$

überabzählbar ist. Zu diesem Zweck betrachten wir für $x \in [0,1)$ die Dezimalbruchentwicklung

$$x = \sum_{k=1}^{\infty} a_k \cdot 10^{-k}\,, \quad a_k \in \{0,\dots,9\}\,. \tag{13.1}$$

Zu gegebenem x werden die a_k folgendermaßen konstruiert: Sei $s_0 = 0$. Sind $a_1,\dots,a_{n-1}$ und damit

$$s_{n-1} = \sum_{k=1}^{n-1} a_k \cdot 10^{-k} \tag{13.2}$$

bereits konstruiert, so wählen wir $a_n \in \{0,\dots,9\}$ so, dass

$$s_{n-1} + a_n 10^{-n} \le x < s_{n-1} + (a_n+1)10^{-n}\,. \tag{13.3}$$

Satz 13.12 *Die Menge $[0,1)$ ist überabzählbar.*

Beweis: Wir zeigen, dass es keine surjektive Abbildung $\tau : \mathbb{N} \to [0,1)$ gibt. Sei $\tau : \mathbb{N} \to [0,1)$ eine beliebige Abbildung. Sei

$$\tau(j) = \sum_{k=1}^{\infty} a_{jk} \cdot 10^{-k}\,, \quad j \in \mathbb{N}\,,$$

die gemäß (13.1) – (13.3) konstruierte Dezimalbruchentwicklung von $\tau(j)$. Wir definieren ein $x \in [0,1)$ durch

$$x = \sum_{k=1}^{\infty} b_k \cdot 10^{-k}\,, \tag{13.4}$$

wobei wir $b_k \in \{0, \dots, 8\}$ so wählen, dass $b_k \neq a_{kk}$ gilt für alle $k \in \mathbb{N}$. Man prüft nach, dass wegen

$$0 \leq \sum_{k=n+1}^{\infty} b_k \cdot 10^{-k} < 10^{-n}$$

(hier geht $b_k \leq 8$ ein) die Darstellung (13.4) mit der gemäß (13.1) – (13.3) konstruierten Dezimalbruchentwicklung von x übereinstimmt. Es folgt, dass $x \neq \tau(j)$ gilt für alle $j \in \mathbb{N}$. Also ist τ nicht surjektiv. □

Folgerung 13.13 *Sei M eine abzählbare Teilmenge von $\mathbb{R}$. Dann ist $\mathbb{R} \setminus M$ überabzählbar. Insbesondere ist die Menge aller irrationalen Zahlen überabzählbar.*

Beweis: Wäre $\mathbb{R} \setminus M$ abzählbar, so wäre auch $\mathbb{R}$ abzählbar nach Satz 13.10, im Widerspruch zu Satz 13.12. □

Bemerkung 13.14 (Kontinuumshypothese)
Nach dem eben Bewiesenen sind $\mathbb{N}$ und $\mathbb{R}$ nicht gleichmächtig. Man kann sich fragen, ob es "zwischen $\mathbb{N}$ und $\mathbb{R}$ noch etwas gibt", d.h. ob es eine Menge M und surjektive Abbildungen

$$f : \mathbb{R} \to M\,, \quad g : M \to \mathbb{N}$$

gibt, so dass M weder zu $\mathbb{N}$ noch zu $\mathbb{R}$ gleichmächtig ist. Die sogenannte Kontinuumshypothese besagt, dass es eine solche Menge M nicht gibt. Es ist mit Methoden der Mathematischen Logik bewiesen worden, dass die Kontinuumshypothese von einer Reihe üblicher Axiomensysteme der Mengenlehre unabhängig ist, das heißt, dass sowohl die Annahme, sie sei wahr, als auch die Annahme, sie sei falsch, mit diesen Axiomensystemen verträglich ist.

Wir kehren zurück zu abzählbaren Teilmengen von $\mathbb{R}$. Die Menge $\mathbb{Q}$ können wir auch beschreiben als die Menge aller reellen Zahlen, die sich als Nullstellen einer linearen Gleichung

$$qx + p = 0$$

mit $p, q \in \mathbb{Z}$ schreiben lassen.

Definition 13.15 (Algebraische Zahl, transzendente Zahl)
Eine Zahl $x \in \mathbb{R}$ heißt algebraisch, wenn es ein von Null verschiedenes Polynom p mit ganzzahligen Koeffizienten gibt, so dass $p(x) = 0$. Eine Zahl $x \in \mathbb{R}$ heißt transzendent, wenn x nicht algebraisch ist.

Beispiel: $x = \sqrt{2}$ ist algebraisch, da $x^2 - 2 = 0$.

Satz 13.16 *Die Menge der algebraischen Zahlen ist abzählbar.*

Beweis: Die Menge G aller Polynome, deren Koeffizienten alle ganzzahlig sind, ist abzählbar. (Beweis: Übungsaufgabe.) Da jedes Polynom $p \neq 0$ nur endlich viele Nullstellen hat, ist die Menge N aller solcher Nullstellen abzählbar; betrachte

$$\tau : \mathbb{N} \times G \to N\,,$$

wobei $\tau(k, p)$ als die k-te Nullstelle von p definiert ist, falls die Anzahl der Nullstellen von p größer oder gleich k ist, und $\tau(k, p) = 0$ andernfalls. □

Bemerkung 13.17
Die Zahlen e und π sind transzendent. Aus letzterem folgt (wie man erkennt, wenn man sich näher mit Algebra beschäftigt), dass die sogenannte "Quadratur des Kreises" ein unlösbares Problem ist. (Das Problem besteht darin, zu einem gegebenen Kreis nur unter Verwendung von Zirkel und Lineal ein Quadrat gleicher Fläche zu konstruieren.) Die Transzendenz von π ist im 19. Jahrhundert (Ferdinand von Lindemann, 1882) bewiesen worden. □

Eine reelle Zahl x heißt **berechenbar**, falls es einen Algorithmus gibt, welcher nacheinander die Ziffern der Dezimalbruchentwicklung von x berechnet. Damit das eine mathematische Definition wird, muss zunächst definiert werden, was ein Algorithmus ist; oder äquivalent, es muss die Menge "aller" Algorithmen definiert werden. Das hat Alan Turing in den 30er Jahren mit dem nach ihm benannten Begriff der "Turingmaschine" getan. Dabei ergibt es sich, dass die Menge aller in diesem Sinne erlaubter Algorithmen abzählbar ist. Daraus folgt, dass es nur abzählbar viele Zahlen gibt, die berechenbar sind. Zu ihnen gehören beispielsweise auch e und π.