Analysis I

Alexander Grigorian
Universität Bielefeld

SS 2014

Contents

Chapter 1

Mengen und Zahlen

Das Hauptziel dieses Kapitels ist die rigorose Entwicklung der Theorie von reellen Zahlen. Obwohl man die reellen Zahlen jeden Tag im wirklichen Leben trifft und benutzt, bedeutet es nicht, dass man wirklich versteht, *was* die reellen Zahlen sind und mit welchen Eigenschaften der reellen Zahlen man in der Mathematik immer rechnen kann.

Theoretisch kann man die ganze Mathematik als eine axiomatische Theorie darstellen. Das bedeutet, dass die Mathematik eine Sammlung von *Aussagen* (*Sätze*) und *Begriffe* ist, wo die neuen Aussagen aus schon bestehenden Aussagen mit Hilfe von logischer Argumentation (*Beweis*) erhalten werden, und die neuen Begriffe durch *Definitionen* erstellt werden.

Es muss allerdings etwas am Anfang der Theorie geben. Die axiomatische Theorie beginnt mit einer Liste von den *Grundbegriffen* und *Axiomen.* Die Axiome sind die Aussagen, die man ohne Beweis akzeptiert und die die Eigenschaften der Grundbegriffe darstellen. Dazu gehören auch die Regeln von logischer Herleitung, die man in den Beweisen benutzen darf.

Die reelle Zahlen gehören zu den Grundlagen der Mathematik, und die Theorie von reellen Zahlen befindet sich sehr nahe zum axiomatischen Anfang der Mathematik. Ganz am Anfang stehen die Mengenlehre und die Mathematische Logik. Im ersten Abschnitt beschäftigen wir uns mit den Begriffen der Mengenlehre, obwohl ohne richtige Axiomatisierung. Der Zweck davon ist, die auf der Mengenlehre basierende Sprache der Mathematik zu lernen. Danach führen wir die Axiome von reellen Zahlen ein und gewinnen aus den Axiomen alle wesentlichen Eigenschaften der reellen Zahlen.

1.1 Grundbegriffe der Mengenlehre

1.1.1 Mengen und Operationen auf den Mengen

Alle Objekte, die man in Mathematik trifft, sind entweder *Mengen* oder *Elemente von Mengen.* In axiomatischer Mengenlehre sind diese die Grundbegriffe, die man nicht definiert. In dieser Einführung benutzen wir keine axiomatische, sondern "naive" Mengenlehre, wo wir uns auf unserem intuitiven Verständnis von Objekten verlassen, das aus unserer Erfahrung stammt.

Der Begründer der Mengenlehre – der deutsche Mathematiker Georg Cantor (1845-1918), erklärte den Begriff von Menge in 1895 wie folgt:

> "Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die *Elemente* von M genannt werden) zu einem Ganzen."

Obwohl sie keine mathematische Definition ist, hilft diese Erklärung die bestimmten Eigenschaften der Mengen zu etablieren. Sobald die Haupteigenschaften der Mengen angegeben sind, kann man die Mathematik, insbesondere Analysis, weiter nur mit rigorosen logischen Methoden entwickeln.

Für Mengen benutzt man häufig eine graphische Darstellung und zeigt sie als Figuren auf der Ebene. Diese sollen nicht als rigorose Methode angenommen werden, insbesondere als der mathematische Begriff von Ebene viel später definiert werden wird. Die graphischen Darstellungen soll man nur als Illustrationen auffassen, die enorm helfen die Beweise bzw Begriffe zu verstehen.

Elemente von Mengen. Jede Menge M besteht aus bestimmten Objekten, die die Elemente von M heißen. Ist x ein Element von M, so schreibt man

$$x \in M$$

("x gehört zu M", "x ist in M", "x liegt in M", "x ist ein Element der Menge M"). Ist x kein Element von M, so schreibt man

$$x \notin M.$$

Es gibt eine Menge, die keine Elemente besitzt. Diese Menge heißt die *leere Menge* und wird mit dem Zeichen $\emptyset$ bezeichnet.

Eine Menge kann explizit angegeben werden wie folgt. Zum Beispiel, die Menge M, die aus Elementen a, b, c, d besteht, bezeichnet man mit

$$M = \{a, b, c, d\}$$

(d.h. alle Elementen von M in den geschwungenen Klammern). Das bedeutet, dass die Elemente von M die Buchstaben a, b, c, d sind, und nichts anderes. Noch ein Beispiel: die Menge $M = \{a\}$ besteht nur aus einem Element a.

Die Elemente dürfen selber die Mengen sein. Zum Beispiel, die Menge $M = \{\emptyset\}$ besteht aus einem Element $\emptyset$.[1]

Teilmengen und Inklusion.

Definition. Menge A heißt *Teilmenge* von Menge B wenn aus $x \in A$ folgt $x \in B$. Man schreibt in diesem Fall

$$A \subset B$$

("A ist eine Teilmenge von B"). Die Beziehung $A \subset B$ zwischen den Mengen A und B heißt *Inklusion.*

Die Aussage "aus $x \in A$ folgt $x \in B$" schreibt man kurz so auf:

$$x \in A \Rightarrow x \in B,$$

[1] Die Menge $\{\emptyset\}$ soll mit der Menge $\emptyset$ nicht verwechselt werden: die erste Menge hat ein Element, während die zweite Menge hat kein Element.

wobei der Pfeil $\Rightarrow$ bedeutet: "impliziert", "ergibt", "aus ... folgt ...".

Zwei Mengen A und B sind gleich genau dann wenn $A \subset B$ und $B \subset A$. In diesem Fall schreibt man

$$A = B$$

("A ist gleich B", "A ist identisch zu B"). Es ist klar, dass $A = B$ genau dann gilt, wenn

$$x \in A \Leftrightarrow x \in B,$$

wobei der Doppelpfeil $\Leftrightarrow$ bedeutet: "genau dann, wenn" oder "äquivalent".

Mit Hilfe von den logischen Symbolen '$\Rightarrow$' und '$\Leftrightarrow$' können wir die Definitionen von Inklusion und Identität von Mengen so umschreiben:

$$\begin{array}{rcl} A & \subset & B \Leftrightarrow (x \in A \Rightarrow x \in B) \\ A & = & B \Leftrightarrow (x \in A \Leftrightarrow x \in B)\,. \end{array}$$

Behauptung. *Die Inklusion von Mengen ist transitiv, d.h.*

$$A \subset B \quad \wedge \quad B \subset C \Rightarrow A \subset C.$$

wobei das Zeichen $\wedge$ bedeutet "und".

Beweis. Wir haben nach $A \subset B$

$$x \in A \Rightarrow x \in B$$

und nach $B \subset C$

$$x \in B \Rightarrow x \in C,$$

woraus folgt

$$x \in A \Rightarrow x \in C$$

und deshalb $A \subset C$. ∎

Durchschnitt und Vereinigung. Jetzt definieren wir einige wichtigen Operationen auf den Mengen. Häufig ist eine Menge M durch eine Eigenschaft E von Elementen angegeben, d.h.

$$x \in M \;\Leftrightarrow\; x \text{ erfüllt } E,$$

was bedeutet: M ist die Menge von den Elementen x mit der Eigenschaft E. In diesem Fall schreibt man auch

$$M = \{x : x \text{ erfüllt } E\} \quad \text{oder} \quad M = \{x \mid x \text{ erfüllt } E\}\,.$$

Definition. Der *Durchschnitt* der Mengen A und B ist die folgende Menge

$$A \cap B = \{x : x \in A \;\wedge\; x \in B\}\,.$$

Die äquivalente Definition ist:

$$x \in A \cap B \;\Leftrightarrow\; x \in A \;\wedge\; x \in B.$$

Die Mengen A und B heißen *disjunkt* wenn $A \cap B = \emptyset$.
Definition. Die *Vereinigung* der Mengen A und B ist die folgende Menge:

$$A \cup B = \{x : x \in A \ \vee \ x \in B\},$$

wobei das Zeichen $\vee$ bedeutet "oder".

Die äquivalente Definition ist:

$$x \in A \cup B \Leftrightarrow x \in A \ \vee \ x \in B$$

Beispiel. Es folgt aus den Definitionen, dass

$$A \cap B \subset A \subset A \cup B$$

und

$$A \cap B \subset B \subset A \cup B.$$

Gelten die Inklusionen $A' \subset A$ und $B' \subset B$, so erhalten wir

$$A' \cap B' \subset A \cap B$$

und

$$A' \cup B' \subset A \cup B.$$

Auch gelten die Identitäten

$$A \cap A = A = A \cup A$$

und

$$A \cap \emptyset = \emptyset, \qquad A \cup \emptyset = A.$$

Die Gesetze von den Operationen $\cap, \cup$.
Behauptung. (Kommutativgesetze) *Die Operationen $\cap$ und $\cup$ sind kommutativ, d.h. die folgenden Identitäten gelten für alle Mengen A, B:*

$$\begin{aligned} A \cap B &= B \cap A \\ A \cup B &= B \cup A. \end{aligned}$$

Das Wort "kommutativ" bedeutet, dass die Operanden A und B vertauschbar sind.
Beweis. Es ist klar, dass

$$x \in A \ \wedge \ x \in B \ \Leftrightarrow \ x \in B \ \wedge \ x \in A$$

woraus die Identität $A \cap B = B \cap A$ folgt. Die zweite Identität beweist man analog. ■

Behauptung. (Assoziativgesetz) *Die Operationen $\cap$ und $\cup$ sind assoziativ, d.h. die folgenden Identitäten gelten für alle Mengen A, B, C:*

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

Das Wort “assoziativ” bedeutet, dass das Ergebnis von zwei Operationen von der Reihenfolge der Operationen unabhängig ist.

Beweis. Nach den Definitionen erhalten wir

$$\begin{aligned} x \in (A \cap B) \cap C &\Leftrightarrow x \in A \cap B \ \wedge \ x \in C \\ &\Leftrightarrow (x \in A \ \wedge \ x \in B) \ \wedge \ x \in C \\ &\Leftrightarrow x \in A \ \wedge \ x \in B \ \wedge \ x \in C. \end{aligned}$$

Gleichfalls erhalten wir

$$\begin{aligned} x \in A \cap (B \cap C) &\Leftrightarrow x \in A \ \wedge \ x \in B \cap C \\ &\Leftrightarrow x \in A \ \wedge \ (x \in B \ \wedge \ x \in C) \\ &\Leftrightarrow x \in A \ \wedge \ x \in B \ \wedge \ x \in C, \end{aligned}$$

woraus die erste Identität folgt. Die zweite Identität beweist man gleichfalls. ■

Man definiert den Durchschnitt der Mengen A, B, C durch

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

wobei das Zeichen “:=” bedeutet “ist definiert durch”, und die Vereinigung dreier Mengen A, B, C durch

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

Es folgt aus dem obigen Beweis, dass

$$\begin{aligned} x &\in A \cap B \cap C \Leftrightarrow x \in A \ \wedge \ x \in B \ \wedge \ x \in C \\ x &\in A \cup B \cup C \Leftrightarrow x \in A \ \vee \ x \in B \ \vee \ x \in C. \end{aligned}$$

Behauptung. (Distributivgesetze) *Es gelten die Identitäten*

$$\begin{aligned} (A \cap B) \cup C &= (A \cup C) \cap (B \cup C) \qquad (1.1) \\ (A \cup B) \cap C &= (A \cap C) \cup (A \cap C) \end{aligned}$$

Das Wort “distributiv” bedeutet, dass C auf A und B distributiert (verteilt) werden kann.

Beweis. Beweisen wir (1.1) (und das zweite Distributivgesetz wird analog bewiesen). We haben nach Definition

$$\begin{aligned} x \in (A \cap B) \cup C &\Leftrightarrow x \in (A \cap B) \ \vee \ x \in C \\ &\Leftrightarrow (x \in A \ \wedge \ x \in B) \ \vee \ x \in C \end{aligned}$$

und

$$x \in (A \cup C) \cap (B \cup C) \Leftrightarrow (x \in A \ \vee \ x \in C) \ \wedge \ (x \in B \ \vee \ x \in C) \, .$$

Es bleibt zu zeigen, dass

$$(x \in A \ \wedge \ x \in B) \ \vee \ x \in C \ \Leftrightarrow \ (x \in A \ \vee \ x \in C) \ \wedge \ (x \in B \ \vee \ x \in C) \, . \qquad (1.2)$$

Gilt $x \in C$, so sind die beiden Seiten von (1.2) wahr.

Gilt $x \notin C$, so erhalten wir

$$(x \in A \ \wedge \ x \in B) \ \vee \ x \in C \ \Leftrightarrow x \in A \ \wedge \ x \in B$$

und

$$(x \in A \ \vee \ x \in C) \ \wedge \ (x \in B \ \vee \ x \in C) \Leftrightarrow (x \in A) \ \wedge \ (x \in B)\,,$$

woraus (1.2) folgt. ■

Subtraktion von Mengen.

Definition. Die Differenzmenge $A \setminus B$ zweier Mengen A, B ist definiert wie folgt:

$$A \setminus B := \{x : x \in A \ \wedge \ x \notin B\}\,.$$

Eine äquivalente Definition ist

$$x \in A \setminus B \Leftrightarrow x \in A \ \wedge \ x \notin B.$$

Das Zeichen \ heißt “Minus” oder “Mengenminus”.

Es folgt, dass $A \setminus B \subset A$, während $A \setminus B$ und B disjunkt sind. Zum Beispiel, $A \setminus A = \emptyset$ und $A \setminus \emptyset = A$.

Potenzmenge. Betrachten wir jetzt nur die Teilmengen einer *Grundmenge* X. Die Menge von allen Teilmengen von X heißt die *Potenzmenge* von X und ist mit 2^X (oder $\mathcal{P}(X)$) bezeichnet. d.h.

$$A \in 2^X \Leftrightarrow A \subset X.$$

In anderer Wörter die Elementen von 2^X sind die Teilmengen von X. Zum Beispiel, es gilt immer $\emptyset \in 2^X$ und $X \in 2^X$. Die Operationen $\cup$, $\cap$, $\setminus$ mit den Elementen von 2^X ergeben offensichtlich wieder die Elemente von 2^X.

Komplement. Für die Elementen von 2^X gibt es noch eine Operation, die “Komplement” heißt.

Definition. Für jede Menge $A \in 2^X$ definieren wir das *Komplement* A^c durch

$$A^c = X \setminus A.$$

Äquivalent haben wir:

$$x \in A^c \Leftrightarrow x \notin A$$

vorausgesetzt dass $x \in X$. Man benutzt für das Komplement A^c auch die Notation $\complement A$ (wobei ‘C’ aus dem englischen Wort “Complement” stammt).

Satz 1.1 *Die folgenden Identitäten gelten für die beliebigen Mengen $A, B \in 2^X$:*

$$\begin{aligned} A \setminus B &= A \cap B^c \\ (A \cap B)^c &= A^c \cup B^c \\ (A \cup B)^c &= A^c \cap B^c. \end{aligned}$$

Die zweite und dritte Identitäten heißen die *Formeln von De Morgan.* Diese Formeln lassen sich als die folgende Regel formulieren: das Komplement der Vereinigung ist der Durchschnitt der Komplementen, und umgekehrt.

Beweis. We haben

$$\begin{aligned} x \in A \setminus B &\Leftrightarrow x \in A \ \wedge \ x \notin B \\ &\Leftrightarrow x \in A \ \wedge \ x \in B^c \\ &\Leftrightarrow x \in A \cap B^c \end{aligned}$$

woraus $A \setminus B = A \cap B^c$ folgt.

Um die zweite Identität zu beweisen, schreiben wir zuerst

$$x \in (A \cap B)^c \Leftrightarrow x \notin A \cap B.$$

Nun brauchen wir die Negation (Verneinung) der Aussage $x \in A \cap B$. Die Negation bezeichnet man mit dem Zeichen $\neg$, so dass

$$x \notin A \cap B \Leftrightarrow \neg \left(x \in A \cap B\right) \Leftrightarrow \neg \left(x \in A \ \wedge \ x \in B\right).$$

Beachten wir, dass für die beliebigen Aussagen $\mathcal{A}, \mathcal{B}$ folgendes gilt:

$$\begin{aligned} \neg \left(\mathcal{A} \text{ und } \mathcal{B}\right) &\Leftrightarrow \neg\mathcal{A} \text{ oder } \neg\mathcal{B} \\ \neg \left(\mathcal{A} \text{ oder } \mathcal{B}\right) &\Leftrightarrow \neg\mathcal{A} \text{ und } \neg\mathcal{B}, \end{aligned}$$

d.h. "und" und "oder" verwandeln sich ineinander unter der Negation.

Daher

$$\begin{aligned} \neg \left(x \in A \ \wedge \ x \in B\right) &\Leftrightarrow x \notin A \ \vee \ x \notin B \\ &\Leftrightarrow x \in A^c \ \vee \ x \in B^c \\ &\Leftrightarrow x \in A^c \cup B^c, \end{aligned}$$

woraus die Identität $(A \cap B)^c = A^c \cup B^c$ folgt.

Gleichfalls wird die dritte Identität bewiesen:

$$\begin{aligned} x \in (A \cup B)^c &\Leftrightarrow x \notin (A \cup B) \\ &\Leftrightarrow \neg \left(x \in A \ \vee \ x \in B\right) \\ &\Leftrightarrow x \notin A \ \wedge \ x \notin B \\ &\Leftrightarrow x \in A^c \ \wedge \ x \in B^c \\ &\Leftrightarrow x \in A^c \cap B^c. \end{aligned}$$

■

Beispiel. Hier zeigen wir, wie die obigen Identitäten benutzt werden können:

$$\begin{aligned} (A \cap B) \setminus C &= (A \cap B) \cap C^c \\ &= A \cap (B \cap C^c) \\ &= A \cap (B \setminus C) \end{aligned}$$

und

$$\begin{aligned} C \setminus (A \cap B) &= C \cap (A \cap B)^c \\ &= C \cap (A^c \cup B^c) \\ &= (C \cap A^c) \cup (C \cap B^c) \\ &= (C \setminus A) \cup (C \setminus B) . \end{aligned}$$

Quantoren. Unterhalb benutzen wir die folgenden Symbolen (*Quantoren*):

$\forall$ bedeutet "für alle", "für jedes",

$\exists$ bedeutet "es existiert", "es gibt mindestens ein", "für mindestens ein".

Das Zeichen $\forall$ stammt aus dem umgedrehten Buchstabe A (Alle) und heißt *Allquantor*. Das Zeichen $\exists$ stammt aus dem umgedrehten E (Existiert) und heißt *Existenzquantor*.

Mengensysteme. Eine Menge, deren Elemente auch Mengen sind, heißt ein *Mengensystem*. Ein Mengensystem wird häufig mit $\{A_i\}_{i\in I}$ bezeichnet, wobei I eine beliebige Menge ist, die Indexmenge heißt. Diese Notation bedeutet, dass zu jedem $i \in I$ eine Menge A_i zugeordnet ist, und das Mengensystem $\{A_i\}_{i\in I}$ aus allen Mengen A_i mit $i \in I$ besteht.

Definition. Man definiert den Durchschnitt des Mengensystems $\{A_i\}_{i\in I}$ durch

$$\bigcap_{i\in I} A_i = \{x : \forall i \in I \;\; x \in A_i\}$$

und die Vereinigung durch

$$\bigcup_{i\in I} A_i = \{x : \exists i \in I \;\; x \in A_i\} .$$

Die äquivalenten Definitionen:

$$x \in \bigcap_{i\in I} A_i \Leftrightarrow \; x \in A_i \text{ für alle } i \in I$$

und

$$x \in \bigcup_{i\in I} A_i \Leftrightarrow \; x \in A_i \text{ für mindestens ein } i \in I.$$

Die Distributivgesetze gelten auch für die Operationen $\bigcap_{i\in I}$ und $\bigcup_{i\in I}$ wie folgt.

Behauptung. *Die Operationen $\bigcap$ und $\bigcup$ erfüllen die folgenden Distributivgesetze:*

$$\left(\bigcap_{i\in I} A_i\right) \bigcup B = \bigcap_{i\in I} (A_i \cup B) \tag{1.3}$$

und

$$\left(\bigcup_{i\in I} A_i\right) \bigcap B = \bigcup_{i\in I} (A_i \cap B) .$$

Beweis. Nach den obigen Definitionen erhalten wir

$$\begin{aligned} x \in \left(\bigcap_{i\in I} A_i\right) \bigcup B \quad &\Leftrightarrow \quad x \in \bigcap_{i\in I} A_i \quad \vee \;\; x \in B \\ &\Leftrightarrow \quad (\ \forall i \in I \quad x \in A_i) \;\; \vee \quad x \in B \end{aligned}$$

und

$$\begin{aligned} x \in \bigcap_{i\in I} (A_i \cup B) \quad &\Leftrightarrow \quad \forall i \in I \quad\; x \in A_i \cup B \\ &\Leftrightarrow \quad \forall i \in I \quad (x \in A_i \;\vee\; x \in B)\,. \end{aligned}$$

Um (1.3) zu beweisen, bleibt es zu zeigen, dass

$$(\ \forall i \in I \quad x \in A_i) \;\vee\; \quad x \in B \quad \Leftrightarrow \forall i \in I \quad (x \in A_i \;\vee\; x \in B)\,. \tag{1.4}$$

Ist x in B, so sind die beiden Seiten von (1.4) wahr. Gilt $x \notin B$, so fällt die Bedingung $x \in B$ aus, und wir erhalten

$$(\ \forall i \in I \quad x \in A_i) \;\vee\; \quad x \in B \;\Leftrightarrow\; \quad \forall i \in I \quad x \in A_i$$

und

$$\forall i \in I \quad (x \in A_i \;\vee\; x \in B) \;\Leftrightarrow \forall i \in I \quad x \in A_i,$$

woraus (1.4) und deshalb auch (1.3) folgen. Die zweite Identität wird analog bewiesen. ■

Satz 1.2 *Die folgenden Identitäten gelten für alle Mengen $A_i \subset X$:*

$$\begin{aligned} \left(\bigcap_{i\in I} A_i\right)^c &= \bigcup_{i\in I} A_i^c \qquad (1.5) \\ \left(\bigcup_{i\in I} A_i\right)^c &= \bigcap_{i\in I} A_i^c \end{aligned}$$

Diese Identitäten heißen aus die Formeln von De Morgan.

Beweis. Zuerst beachten wir folgendes. Wir haben

$$\begin{aligned} x \in \left(\bigcap_{i\in I} A_i\right)^c \quad &\Leftrightarrow \quad x \notin \bigcap_{i\in I} A_i \\ &\Leftrightarrow \quad \neg \left(x \in \bigcap_{i\in I} A_i\right) \\ &\Leftrightarrow \quad \neg\,(\forall i \in I \;\; x \in A_i)\,. \end{aligned}$$

Für jedes $i \in I$ sei $\mathcal{A}_i$ eine Aussage. Nun benutzen wir die Äquivalenz

$\neg$(“für jedes $i \in I$ gilt $\mathcal{A}_i$”) $\Leftrightarrow$ (“es gibt $i \in I$ wenn $\mathcal{A}_i$ gilt nicht),

d.h. mit den logischen Symbolen

$$\neg\,(\forall i \in I \;\; \mathcal{A}_i) \Leftrightarrow (\exists i \in I \;\; \neg\mathcal{A}_i)\,. \tag{1.6}$$

Analog gilt

$$\neg\left(\exists i\in I\quad \mathcal{A}_i\right)\Leftrightarrow\left(\forall i\in I\quad \neg\mathcal{A}_i\right).$$

Man sieht, dass die folgende Regel gilt: bei der Negation verwandeln sich der Allquantor $\forall$ und der Existenzquantor $\exists$ ineinander.

Mit Hilfe von (1.6) erhalten wir

$$\begin{aligned} x\in\left(\bigcap_{i\in I}A_i\right)^c &\Leftrightarrow \left(\exists i\in I\quad x\notin A_i\right)\\ &\Leftrightarrow \left(\exists i\in I\quad x\in A_i^c\right)\\ &\Leftrightarrow x\in\bigcup_{i\in I}A_i^c, \end{aligned}$$

woraus (1.5) folgt. Die zweite Formel von De Morgan wird analog bewiesen. ■

Symmetrische Differenz. Es gibt noch eine interessante Operation auf Mengen, die *symmetrische Differenz* heißt und mit $A\bigtriangleup B$ bezeichnet wird:

$$A\bigtriangleup B=(A\setminus B)\cup(B\setminus A).$$

Es folgt daraus, dass

$$x\in A\bigtriangleup B\Leftrightarrow\ (x\in A\ \wedge\ x\notin B)\ \vee\ (x\in B\ \ \wedge\ x\notin A),$$

d.h. $x\in A\bigtriangleup B$ gilt genau dann, wenn x genau zu einer Menge von A,B gehört.

Für symmetrische Differenz gelten die folgenden Identitäten.

1. Kommutativgesetz:
$$A\bigtriangleup B=B\bigtriangleup A.$$

2. Assoziativgesetz:
$$(A\bigtriangleup B)\bigtriangleup C=A\bigtriangleup(B\bigtriangleup C).$$

3. Distributivgesetz bezüglich $\cap$:
$$A\cap(B\bigtriangleup C)=(A\cap B)\bigtriangleup(A\cap C).$$

Kartesisches Produkt.

Definition. Für je zwei Mengen A,B definieren wir *kartesisches (direktes) Produkt* $A\times B$ der Mengen A,B wie folgt: die Menge $A\times B$ besteht aus allen geordneten *Paaren* (x,y) wobei $x\in A$ und $y\in B$:

$$A\times B=\{(x,y):x\in A,\quad y\in B\}.$$

Das geordnete Paar (x,y) ist ein neues Objekt, das man aus den Elementen von A und B erstellt. Zwei Paaren (x,y) und (x',y') sind gleich genau dann, wenn $x=x'$ und $y=y'$.

Kartesisches Produkt ist nicht kommutativ, aber assoziativ. Beachten wir, dass

$$(A\times B)\times C=\{((x,y),z):x\in A,y\in B,z\in C\}$$

und
$$A \times (B \times C) = \{(x, (y, z)) : x \in A, y \in B, z \in C\}.$$
Nun identifizieren wir die Paaren $((x, y), z)$ und $(x, (y, z))$ miteinander und mit dem geordneten 3-Tupel (x, y, z), indem wir annehmen, dass

$$((x, y), z) = (x, (y, z)) = (x, y, z).$$

Dann gilt für kartesisches Produkt das Assoziativgesetz:

$$(A \times B) \times C = A \times (B \times C) = A \times B \times C,$$

wobei $A \times B \times C$ kartesisches Produkt dreier Mengen A, B, C ist, d.h.

$$A \times B \times C := \{(x, y, z) : x \in A, y \in B, z \in C\}.$$

Kartesisches Produkt erfüllt auch das Distributivgesetz:

$$(A_1 \star A_2) \times B = (A_1 \times B) \star (A_2 \times B)$$

wobei $\star$ eine beliebige Mengenoperation $\cap, \cup, \setminus, \triangle$ bezeichnet.

1.1.2 Relationen

Sei M eine Menge.

Definition. Eine *Relation* R auf M ist eine Teilmenge von $M \times M$.

Man benutzt eine Relation wenn man die bestimmten Paaren (x, y) mit $x \in M$ und $y \in M$ auszeichnen soll. Man kann es so vorstellen, dass zwischen $x, y \in M$ mit $(x, y) \in R$ eine Relation besteht, und deshalb schreibt man xRy statt $(x, y) \in R$ um diese Beziehung zu bezeichnen. Es ist häufig bequemer statt R in der Relation xRy ein Symbol zu benutzen, wie $=, \subset, <, >$, usw.

Beispiel. Wählen wir eine Grundmenge X und definieren eine Relation R auf der Potenzmenge $M = 2^X$ durch
$$(A, B) \in R \Leftrightarrow A = B,$$
Dann bedeutet ARB, dass $A = B$. Diese Relation heißt *Identitätsrelation.*

Beispiel. Definiert wir jetzt eine andere Relation R auf $M = 2^X$ durch

$$(A, B) \in R \Leftrightarrow A \subset B,$$

so dass die Relation ARB bedeutet $A \subset B$. Wie wir es schon kennen, heißt diese Relation *Inklusion.*

In der nächsten Definition bezeichnen wir eine Relation R mit der Schlange $\sim$, d.h. wir schreiben $x \sim y$ falls $(x, y) \in R$ (und $x \not\sim y$ sonst).

Definition. Die Relation $\sim$ heißt eine *Äquivalenzrelation*, wenn sie die folgenden Bedingungen erfüllt:

1. $x \sim x \quad \forall x \in M$ (Reflexivität)

2. $x \sim y \Rightarrow y \sim x$ (Symmetrie)

3. $x \sim y \ \wedge \ y \sim z \Rightarrow x \sim z$ (Transitivität)

Zum Beispiel, die Identitätsrelation erfüllt diese Eigenschaften und deshalb ist Äquivalenzrelation. Die Inklusion ist keine Äquivalenzrelation, da sie nicht symmetrisch ist (obwohl reflexiv und transitiv).

Gegeben ist eine Äquivalenzrelation $\sim$ auf der Menge M und ein Element $x \in M$, bezeichnen wir mit $[x]$ die folgende Teilmenge von M:

$$[x] = \{z \in M : z \sim x\}.$$

Die Teilmenge $[x]$ heißt die *Äquivalenzklasse* von x.

Satz 1.3 *Sei $\sim$ eine Äquivalenzrelation auf M. Dann gilt das folgende.*

(a) *Die Vereinigung von allen Äquivalenzklassen ist M (d.h. $\bigcup_{x\in M} [x] = M$).*

(b) *Je zwei Äquivalenzklassen $[x]$ und $[y]$ sind entweder identisch oder disjunkt.*

(c) *Für je zwei Elemente $x, y \in M$ gilt $x \sim y$ genau dann, wenn x und y in derselber Äquivalenzklasse sind.*

Die Eigenschaften $(a) - (c)$ bedeuten zusammen, dass die verschiedene Äquivalenzklassen eine Zerlegung der M in disjunkte Teilmengen bilden, derart, dass je zwei Elemente von M genau dann äquivalent sind, wenn sie in derselber Teilmenge sind.

Die Menge von verschiedenen Äquivalenzklassen heißt der *Quotientenmenge* von $\sim$ und wird mit $M/\sim$ bezeichnet. Man kann $M/\sim$ als Ersatz von der Operation Division betrachten (ein Menge lässt sich durch eine Äquivalenzrelation "dividieren", wobei das Wort "dividieren" in diesem Fall als "teilen" oder "zerlegen" verstanden werden soll).

Beweis. (a) Nach Reflexivität gilt $x \in [x]$, woraus folgt, dass jedes x in einer Äquivalenzklasse ist und deshalb $\bigcup_{x\in M} [x] = M$.

(b) Seien $[x]$ und $[y]$ nicht disjunkt. Dann gibt es ein element $z \in [x] \cap [y]$. Nach Definition haben wir $z \sim x$ und $z \sim y$. Nach Symmetry und Transitivität erhalten wir $x \sim y$. Dann, für jedes $t \in [x]$, ergeben die Relations $t \sim x$ und $x \sim y$, dass $t \sim y$ und $t \in [y]$. Analog $t \in [y]$ ergibt $t \in [x]$, woraus folgt, dass $[x] = [y]$.

(c) Sind x, y in derselber Äquivalenzklasse $[z]$, so erhalten wir $x \sim z$ und $y \sim z$, daher $x \sim y$. Umgekehrt, gilt $x \sim y$, so erhalten wir $y \in [x]$, was zusammen mit $x \in [x]$ ergibt, dass x, y in derselber Äquivalenzklasse sind. ■

1.1.3 Abbildungen

Definition. Gegeben seien zwei Mengen X, Y. Eine Abbildung (=Funktion) f von X nach Y ist eine Zuordnung (Vorschrift, Regel) $x \mapsto y$, die jedem Element $x \in X$ genau ein Element $y \in Y$ zuordnet.

Die Abbildung wird mit

$$f : X \to Y$$

oder
$$X \xrightarrow{f} Y$$
bezeichnet. Das Element y heißt der *Wert* von f an der Stelle x (oder das *Bild* von x) und wird mit $f(x)$ bezeichnet. Man bezeichnet die Abbildung auch mit
$$\begin{aligned} f : X &\to Y \\ x &\mapsto f(x) \end{aligned}$$
oder mit
$$X \ni x \mapsto f(x) \in Y.$$
Die Menge X heißt der *Definitionsbereich* (oder *Definitionsmenge*) von f, die Menge Y – der *Wertebereich* (oder *Zielmenge*)

Jetzt besprechen wir, was genau eine Zuordnung $x \mapsto y$ bedeutet.

Definition. Eine Zuordnung $x \mapsto y$ (wobei $x \in X, y \in Y$) ist eine Teilmenge G von $X \times Y$ mit der folgenden Bedingung:
$$\forall x \in X \quad \exists! y \in Y \text{ mit } (x, y) \in G, \tag{1.7}$$
wobei $\exists!$ (Existenzquantor mit dem Ausrufezeichen) bedeutet: "es gibt genau ein".

Die Eigenschaft (1.7) erlaubt jedem $x \in X$ genau ein $y \in Y$ zuordnen. Wenn wir die entsprechende Abbildung mit f bezeichnen, dann erhalten wir $y = f(x)$ und somit
$$G = \{(x, f(x)) \in X \times Y : x \in X\}.$$
Die Menge G heißt der *Graph* der Abbildung f. Wie sehen, dass die Begriffe von Abbildung, Zuordnung, Graph praktisch identisch sind, obwohl diese Wörter unterschiedlich benutzt werden.

Beispiel. Die Abbildung $f : X \to X$ mit $f(x) = x$ heißt die *Identitätsabbildung* der Menge X. Man bezeichnet die Identitätsabbildung von X mit Id_X. Der Graph von Id_X besteht aus den Paaren (x, x), die die *Diagonale* von $X \times X$ formen.

Beispiel. Betrachten wir eine beliebige Menge X und die Menge $Y = \{0, 1\}$, die aus den Symbolen $0, 1$ besteht. Sei A eine Teilmenge von X. Definition wie eine Funktion $f : X \to Y$ mit
$$f(x) = \begin{cases} 1, & x \in A, \\ 0, & x \in A^c. \end{cases}$$
Der Graph von f besteht aus den Paaren $(x, 1)$ mit $x \in A$ und $(x, 0)$ mit $x \in A^c$. Diese Funktion f heißt die *charakteristische* Funktion oder *Indikatorfunktion* von A und wird mit $\mathbf{1}_A$ bezeichnet.

Urbild. Jede Abbildung $f : X \to Y$ induziert die Abbildung $f^{-1} : 2^Y \to 2^X$ wie folgt. Für jede Teilmenge $A \subset Y$, definieren wir das *Urbild* $f^{-1}(A)$ von A durch
$$f^{-1}(A) = \{x \in X : f(x) \in A\}.$$
Nach Definition ist $f^{-1}(A)$ eine Teilmenge von X, so dass die Zuordnung $A \mapsto f^{-1}(A)$ eine Abbildung von 2^Y nach 2^X bestimmt. Die Abbildung $f^{-1} : 2^Y \to 2^X$ heißt die *Urbildabbildung* von f.

Satz 1.4 *Die Urbildabbildung* $f^{-1} : 2^Y \to 2^X$ *ist mit den Mengenoperationen* $\cap, \cup. \setminus$ *vertauschbar:*

$$\begin{aligned} f^{-1}(A \cap B) &= f^{-1}(A) \cap f^{-1}(B) \\ f^{-1}(A \cup B) &= f^{-1}(A) \cup f^{-1}(B) \\ f^{-1}(A \setminus B) &= f^{-1}(A) \setminus f^{-1}(B). \end{aligned}$$

Beweis. We haben

$$\begin{aligned} x \in f^{-1}(A \cap B) &\Leftrightarrow f(x) \in A \cap B \\ &\Leftrightarrow f(x) \in A \ \wedge \ f(x) \in B \\ &\Leftrightarrow x \in f^{-1}(A) \ \wedge \ x \in f^{-1}(B) \\ &\Leftrightarrow x \in f^{-1}(A) \cap f^{-1}(B) \end{aligned}$$

woraus die erste Identität folgt. Die anderen Identitäten werden analog bewiesen. ■

Bemerkung. Für jede Teilmenge $A \subset X$ definiert man das Bild $f(A)$ von A wie folgt:

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

so dass f als die *Bildabbildung* von 2^X nach 2^Y betrachtet werden kann. Beachten wir folgendes: die Bildabbildung $f : 2^X \to 2^Y$ ist mit den Mengenoperationen *nicht immer* vertauschbar. Zum Beispiel, betrachten wir die Mengen $X = \{0, 1\}$, $Y = \{2\}$ und die einzige Abbildung $f : X \to Y$. Für die Mengen $A = \{0\}$ und $B = \{1\}$ erhalten wir $f(A) = f(B) = Y$ und deshalb

$$f(A) \cap f(B) = Y,$$

während $A \cap B = \emptyset$ und deshalb

$$f(A \cap B) = \emptyset.$$

Man sieht, dass $f(A) \cap f(B) \neq f(A \cap B)$.

Komposition von Abbildungen.

Definition. Gegeben seien zwei Abbildungen

$$X \xrightarrow{g} Y \xrightarrow{f} Z,$$

definieren wir die *Komposition (Verkettung, zusammengesetzte Abbildung)* $f \circ g$ als eine Abbildung von X nach Z mit

$$(f \circ g)(x) = f(g(x)).$$

Man kann auch schreiben

$$f \circ g : X \to Z, \quad x \mapsto f(g(x)).$$

Schematisch kann man die Verkettung so darstellen:

$$\begin{tikzcd} & Y \arrow[dr, "f"] & \\ X \arrow[ur, "g"] \arrow[rr, "f\circ g"] & & Z \end{tikzcd}$$

Man muss betonen, dass die Komposition $f \circ g$ nur dann wohldefiniert ist, wenn der Wertebereich von g im Definitionsbereich for f liegt. Daraus folgt, dass $g \circ f$ nicht unbedingt wohldefiniert sein soll, sogar wenn $f \circ g$ wohldefiniert ist. Insbesondere kann man nicht erwarten, dass die Komposition kommutativ ist. Aber die Komposition ist immer assoziativ.

Satz 1.5 (Assoziativgesetz für Komposition) *Die folgende Identität*

$$(f \circ g) \circ h = f \circ (g \circ h)$$

gilt für je drei Abbildungen $X \xrightarrow{h} Y \xrightarrow{g} Z \xrightarrow{f} U$.

Beweis. Bemerken wir zunächst, dass die beiden Verkettungen $(f \circ g) \circ h$ und $f \circ (g \circ h)$ von X nach U abbilden, wie man auf den folgenden Diagrammen sieht:

$$\begin{tikzcd} & Y \arrow[dr, "g"] & \\ X \arrow[ur, "h"] \arrow[rr, "g\circ h"] \arrow[dr, "f\circ(g\circ h)"'] & & Z \arrow[dl, "f"] \\ & U & \end{tikzcd} \quad \text{und} \quad \begin{tikzcd} & Y \arrow[dr, "g"] \arrow[d, "f\circ g"] & \\ X \arrow[ur, "h"] \arrow[r, "(f\circ g)\circ h"'] & U & Z \arrow[l, "f"] \end{tikzcd}$$

Für jedes $x \in X$ gilt

$$((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x)))$$

und analog

$$(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x))),$$

woraus die Identität $(f \circ g) \circ h = f \circ (g \circ h)$ folgt. ∎

Das Assoziativgesetz erlaubt uns die Verkettung dreier Abbildungen zu definieren wie folgt:

$$f \circ g \circ h := (f \circ g) \circ h.$$

Schematisch sieht die Komposition $f \circ g \circ h$ so aus:

$$\begin{tikzcd} Y \arrow[r, "g"] & Z \arrow[d, "f"] \\ X \arrow[u, "h"] \arrow[r, "f\circ g\circ h"] & U \end{tikzcd}$$

Satz 1.6 *Gegeben seien zwei Abbildungen* $X \xrightarrow{g} Y \xrightarrow{f} Z$ *. Die folgende Identität gilt für die Urbildabbildungen:*

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}. \tag{1.8}$$

Beweis. Nach Definition haben wir das folgende Diagramm von der Urbildabbildungen f^{-1} und g^{-1}

$$\begin{tikzcd} & 2^Y \arrow[dl, "g^{-1}"'] & \\ 2^X & & 2^Z \arrow[ul, "f^{-1}"'] \arrow[ll, "g^{-1}\circ f^{-1}"] \end{tikzcd}$$

so dass $g^{-1} \circ f^{-1}$ wohldefiniert ist und

$$2^X \overset{g^{-1}\circ\, f^{-1}}{\longleftarrow} 2^Z.$$

Aus $X \overset{f\circ g}{\longrightarrow} Z$ folgt, dass auch

$$2^X \overset{(f\circ g)^{-1}}{\longleftarrow} 2^Z.$$

Jetzt können wir die Abbildungen $(f \circ g)^{-1}$ und $g^{-1} \circ f^{-1}$ vergleichen. Es gilt für jede Teilmenge $A \subset Z$

$$\begin{aligned} x \in (f \circ g)^{-1}(A) &\Leftrightarrow f \circ g(x) \in A \\ &\Leftrightarrow f(g(x)) \in A \\ &\Leftrightarrow g(x) \in f^{-1}(A) \\ &\Leftrightarrow x \in g^{-1}\left(f^{-1}(A)\right) \\ &\Leftrightarrow x \in g^{-1} \circ f^{-1}(A), \end{aligned}$$

woraus $(f \circ g)^{-1}(A) = (g^{-1} \circ f^{-1})(A)$ folgt und somit (1.8). ■

Umkehrabbildung. Gegeben seien zwei Abbildungen $X \overset{f}{\to} Y \overset{g}{\to} X$. Dann sind die beiden Kompositionen $f \circ g$ und $g \circ f$ wohldefiniert und

$$f \circ g : Y \to Y, \quad g \circ f : X \to X.$$

Definition. Die Abbildung g heißt *Umkehrabbildung (inverse Abbildung, Umkehrfunktion, inverse Funktion)* von f, falls $f \circ g = \mathrm{Id}_Y$ und $g \circ f = \mathrm{Id}_X$.

In diesem Fall ist f auch die Umkehrabbildung von g.

Definition. Eine Abbildung $f : X \to Y$ heißt *bijektiv* falls

$$\forall y \in Y \quad \exists! x \in X \quad \text{mit} \quad f(x) = y.$$

Definition. Eine Abbildung $f : X \to Y$ heißt *surjektiv* falls

$$\forall y \in Y \quad \text{existiert mindestens ein } x \in X \quad \text{mit} \quad f(x) = y.$$

Die Abbildung f heißt *injektiv* falls

$$\forall y \in Y \text{ existiert höchstens ein } x \in X \text{ mit } f(x) = y.$$

Es folgt daraus, dass eine Abbildung f bijektiv genau dann ist, wenn f surjektiv und injektiv ist.

Satz 1.7 *Eine Abbildung $f : X \to Y$ hat eine Umkehrabbildung genau dann, wenn f bijektiv ist.*

Beweis. Hat f eine Umkehrabbildung g, so gelten $f \circ g = \mathrm{Id}_Y$ und $g \circ f = \mathrm{Id}_X$, d.h.

$$f(g(y)) = y \quad \forall y \in Y \tag{1.9}$$

und

$$g(f(x)) = x \;\; \forall x \in X. \tag{1.10}$$

Es folgt aus (1.9), dass $f(x) = y$ für mindestens ein x, nämlich für $x = g(y)$, erfüllt ist. Somit ist f surjektiv. Zeigen wir jetzt, dass f injektiv ist. Gilt $f(x) = y$, so erhalten wir aus (1.10)

$$x = g(f(x)) = g(y),$$

so dass x eindeutig bestimmt ist. Somit ist f injektiv und auch bijektiv.

Umgekehrt, ist f bijektiv, so definieren wir die Abbildung $g : Y \to X$ wie folgt: für jedes $y \in Y$ gibt es genau ein $x \in X$ mit $f(x) = y$, so setzen wir $g(y) = x$. Dann gelten

$$(f \circ g)(y) = f(g(y)) = f(x) = y$$

und

$$(g \circ f)(x) = g(f(x)) = g(y) = x,$$

woraus $f \circ g = \mathrm{Id}_Y$ und $g \circ f = \mathrm{Id}_X$ folgen. Damit hat f die Umkehrabbildung. ■

Existiert die Umkehrabbildung von f, so bezeichnet man sie mit f^{-1}, d.h.

$$f \circ f^{-1} = \mathrm{Id}_Y \;\text{ und }\; f^{-1} \circ f = \mathrm{Id}_X .$$

Achtung. *Man soll die Umkehrabbildung $f^{-1} : Y \to X$ mit der Urbildabbildung $f^{-1} : 2^Y \to 2^X$ nicht verwechseln, obwohl sie identisch bezeichnet werden.*

Sei $f : X \to Y$ bijektiv. Dann für jedes $y \in Y$ besteht das Urbild $f^{-1}(\{y\})$ aus einem einzigen Element $x \in X$, d.h. $f^{-1}(\{y\}) = \{x\}$. Nach Definition der Umkehrabbildung f^{-1}, we haben auch $f^{-1}(y) = x$. In diesem Sinn stimmen die Umkehrabbildung und die Urbildabbildung überein. Kombinieren mit dem Satz 1.6 ergibt das folgende.

Korollar 1.8 *Gegeben seien zwei bijektive Abbildungen $X \xrightarrow{g} Y \xrightarrow{f} Z$. Die folgende Identität gilt für die Umkehrabbildungen*

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

1.2 Axiomensystem von reellen Zahlen

Hier definieren wir axiomatisch die *Menge von reellen Zahlen*. Eine Menge $\mathbb{R}$ heißt die Menge von reellen Zahlen und ihre Elemente heißen reelle Zahlen falls die folgenden vier Gruppen von Axiomen (insgesamt sechzehn Axiome) erfüllt sind.

1.2.1 Axiome

Axiome der Addition. Es gibt eine Abbildung (Operation)

$$\begin{aligned} \mathbb{R}\times\mathbb{R} &\to \mathbb{R} \\ (x,y) &\mapsto x+y \end{aligned}$$

die *Addition* heißt und die folgenden Bedingungen erfüllt:

1. (Das Nullelement) Es existiert ein Element $0\in\mathbb{R}$, so dass
$$x+0=0+x=x\ \forall x\in\mathbb{R}.$$
2. (Das Negative) Für jedes $x\in\mathbb{R}$ existiert ein Element $-x\in\mathbb{R}$ (das Negative von x), so dass
$$x+(-x)=(-x)+x=0.$$
3. (Assoziativgesetz für $+$) Für alle $x,y,z\in\mathbb{R}$ gilt
$$(x+y)+z=x+(y+z)\,.$$
4. (Kommutativgesetz für $+$) Für alle $x,y\in\mathbb{R}$ gilt
$$x+y=y+x.$$

Beispiel. Betrachten wir eine Menge M und definieren die Addition auf der Potenzmenge 2^M als symmetrische Differenz Δ. Die Kommutativ- und Assoziativgesetze für Δ kennen wir schon. Das Nullelement ist $\emptyset$ da $A\Delta\emptyset=A$ für alle $A\subset M$, und das Negative ist $-A=A$ da $A\Delta A=\emptyset$.

Eine andere Möglichkeit ist die Addition auf 2^M als Vereinigung $\cup$ zu definieren. Dann gelten Axiome $1,3,4$ aber nicht 2 da $A\cup B$ niemals gleich $\emptyset$ ist, wenn $A\neq\emptyset$.

Axiome der Multiplikation. Es gibt eine Abbildung (Operation)

$$\begin{aligned} \mathbb{R}\times\mathbb{R} &\to \mathbb{R} \\ (x,y) &\mapsto x\cdot y \end{aligned}$$

die *Multiplikation* heißt und die folgenden Bedingungen erfüllt:

1. (Das Einheitselement) Es existiert ein Element $1\in\mathbb{R}\setminus\{0\}$, so dass
$$x\cdot 1=1\cdot x=x\ \forall x\in\mathbb{R}.$$
2. (Das Inverse) Für jedes $x\in\mathbb{R}\setminus\{0\}$ existiert ein Element $x^{-1}\in\mathbb{R}$ (das Inverse von x), so dass
$$x\cdot x^{-1}=x^{-1}\cdot x=1.$$
3. (Assoziativgesetz für $\cdot$) Für alle $x,y,z\in\mathbb{R}$ gilt
$$(x\cdot y)\cdot z=x\cdot(y\cdot z)\,.$$

4. (Kommutativgesetz für $\cdot$) Für alle $x, y \in \mathbb{R}$ gilt
$$x \cdot y = y \cdot x.$$

5. (Distributivgesetz) Für alle $x, y, z \in \mathbb{R}$ gilt
$$(x + y) \cdot z = x \cdot z + y \cdot z.$$

Eine Menge K (statt $\mathbb{R}$), wo die Operationen Addition und Multiplikation definiert sind und die obigen Axiome erfüllen, heißt *Körper.* Die ersten zwei Gruppen von Axiomen heißen *Körperaxiome.* Deshalb ist $\mathbb{R}$ ein Körper. Weitere Beispiele von Körper betrachten wir später.

Beispiel. Betrachten wir die Potenzmenge 2^M mit Multiplikation als Durchschnitt $\cap$ und mit Addition als Δ oder $\cup$. Die Axiome 3-5 gelten nach Eigenschaften von den Mengenoperationen, und Axiom 1 gilt mit dem Einheitselement M, da $A \cap M = A$. Aber das Axiom 2 fällt aus, da $A \cap B$ niemals gleich M ist wenn $A \neq M$.

Anordnungsaxiome. Auf $\mathbb{R}$ ist eine Relation $\leq$ definiert, d.h. für je zwei Elemente $x, y \in \mathbb{R}$ ist $x \leq y$ entweder wahr oder falsch. Diese Relation heißt *Ungleichheit* und erfüllt die folgenden Bedingungen:

1. (Reflexivität) $x \leq x \ \forall x \in \mathbb{R}$.
2. (Antisymmetrie) $x \leq y \ \wedge \ y \leq x \Rightarrow x = y$.
3. (Transitivität) $x \leq y \ \wedge \ y \leq z \Rightarrow x \leq z$
4. (Vergleichbarkeit) Für je zwei Zahlen $x, y \in \mathbb{R}$ gilt $x \leq y \ \vee \ y \leq x$.
5. (Anordnung und Addition) $x \leq y \Rightarrow x + z \leq y + z \ \ \forall z \in \mathbb{R}$
6. (Anordnung und Multiplikation) $0 \leq x \ \wedge \ 0 \leq y \ \Rightarrow \ 0 \leq x \cdot y$.

Erfüllt eine Relation $\leq$ auf einer Menge K die Anordnungsaxiome $1 - 3$, so heißt $\leq$ eine *Ordnung*, und die Menge K heißt geordnet. Erfüllt $\leq$ zusätzlich auf das Axiom 4, so heißt $\leq$ eine *totale Ordnung* und K heißt *total geordnet.* Somit ist $\mathbb{R}$ eine total geordnete Menge. Die Axiome 5 und 6 geben die Beziehung zwischen der Ordnung und den Körperoperationen. Ein Körper K der auch die Anordnungsaxiome erfüllt, heißt *angeordneter Körper.* Somit ist $\mathbb{R}$ ein angeordneter Körper.

Beispiel. Auf der Potenzmenge 2^M definieren wir die Relation $\leq$ als Inklusion $\subset$. Offensichtlich gelten die Axiome $1 - 3$, so dass $\subset$ eine Ordnung ist. Das Axiom 5 gilt, vorausgesetzt, dass Addition von Mengen als $\cup$ definiert wird. Das Axiom 6 gilt für jede Definition von Multiplikation da immer $\emptyset \subset A$. Das Axiom 2 fällt allerdings aus: es kann sein, dass $A \not\subset B$ und $B \not\subset A$. Somit ist Inklusion keine totale Ordnung.

Vollständigkeitsaxiom. Seien X, Y nicht leere Teilmengen von $\mathbb{R}$ mit der Eigenschaft
$$\forall x \in X \ \ \forall y \in Y \ \ \text{gilt } x \leq y.$$

Dann existiert ein $c \in \mathbb{R}$ so dass

$$\forall x \in X \;\; \forall y \in Y \;\; \text{gilt } x \leq c \leq y.$$

Man sagt, dass die Zahl c die Mengen X und Y trennt.

Man stellt die reellen Zahlen vor als die Punkte auf einer waagerechten Gerade (obwohl eine Gerade noch nicht definiert ist!). Die Punkte am links sind immer kleiner als die Punkte am rechts. Die Vollständigkeitsaxiom bedeutet folgendes: liegt die ganze Menge X links von Y, so existiert ein Punkt c zwischen X und Y. Man kann es auch so vorstellen, dass die Gerade keine Lücke enthält.

Beispiel. Betrachten wir die Potenzmenge 2^M mit Relation $\subset$. Wir behaupten, dass das Vollständigkeitsaxiom auf 2^M gilt. Gegeben seien zwei Teilmengen X und Y von 2^M, d.h. zwei Mengensysteme, die aus Teilmengen von M bestehen. Angenommen, dass $A \subset B$ für alle $A \in X$ und $B \in Y$ gilt, definieren wir die Teilmenge $C \subset M$ durch

$$C = \bigcup_{A \in X} A.$$

Zeigen wir, dass C die Mengensysteme X und Y trennt. Nach Definition gilt $A \subset C$ für alle $A \in X$. Die Voraussetzung, dass $A \subset B$ für alle $A \in X$ und $B \in Y$, impliziert, dass $C \subset B$ für alle $B \in Y$, so dass $A \subset C \subset B$ für alle $A \in X$ und $B \in Y$ gilt. Auch die Menge $C' = \bigcap_{B \in Y} B$ trennt X und Y.

Bemerkung. Die Existenz der Menge $\mathbb{R}$, die alle Axiome von reellen Zahlen erfüllt, wird später im Kapitel 2 kurz besprochen.

1.2.2 Folgerungen aus den Körperaxiomen

Folgerungen aus den Axiomen der Addition. Jetzt zeigen wir, wie man aus den Axiomen die üblichen algebraischen Regeln bzw die weiteren Eigenschaften von reellen Zahlen gewinnt.

- *Das Nullelement ist eindeutig bestimmt.*

Sei $0'$ ein anderes Nullelement. Dann haben wir $0' + 0 = 0$ und $0' + 0 = 0'$ woraus $0' = 0$ folgt.

- *Das Negative eines $x \in \mathbb{R}$ ist eindeutig bestimmt.*

Seien y und z zwei Negative von x. Dann erhalten wir

$$y = y + 0 = y + (x + z) = (y + x) + z = 0 + z = z,$$

so that $y = z$.

Daraus folgt, dass

$$-(-x) = x$$

weil $x + (-x) = 0$ und deshalb x die Definition des Negatives von $-x$ erfüllt. Wir sehen auch, dass

$$-0 = 0$$

weil $0 + 0 = 0$.

- *Die Gleichung $x + a = b$ hat eine eindeutige Lösung $x = b + (-a)$.*

  Der Wert $x = b + (-a)$ erfüllt die Gleichung, weil

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

Andererseits folgt es aus der Gleichung $x + a = b$, dass

$$x = x + 0 = x + (a + (-a)) = (x + a) + (-a) = b + (-a),$$

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

Die Summe $b + (-a)$ wird auch mit $b - a$ bezeichnet und heißt die *Differenz* von b und a. Die Abbildung (Operation) $(a, b) \mapsto b - a$ heißt *Subtraktion*.

Folgerungen aus den Axiomen der Multiplikation.

- *Das Einheitselement ist eindeutig bestimmt.*
- *Das Inverse eines $x \in \mathbb{R} \setminus \{0\}$ ist eindeutig bestimmt. Auch gelten*

$$\left(x^{-1}\right)^{-1} = x \quad \textit{und} \quad 1^{-1} = 1.$$

- *Die Gleichung $x \cdot a = b$ hat genau eine Lösung $x = b \cdot a^{-1}$ vorausgesetzt $a \neq 0$.*

  Die Beweise sind analog zum Fall der Addition. Das Produkt $b \cdot a^{-1}$ heißt der *Quotient* von b und a und wird mit b/a oder $\frac{b}{a}$ bezeichnet. Die Abbildung (Operation) $(a, b) \mapsto b/a$ heißt *Division*.

Folgerungen aus dem Distributivgesetz.

- $x \cdot 0 = 0 \cdot x = 0$ *für alle* $x \in \mathbb{R}$

  Wir haben

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

woraus folgt

$$x \cdot 0 = x \cdot 0 - x \cdot 0 = 0.$$

- $x \cdot y = 0$ *ergibt* $x = 0$ *oder* $y = 0$. *Äquivalent,* $x \neq 0$ *und* $y \neq 0$ *ergeben* $x \cdot y \neq 0$.

  Sei $y \neq 0$. Die Lösung der Gleichung $x \cdot y = 0$ bezüglich x ergibt $x = 0/y = 0 \cdot y^{-1} = 0$.

- $(-1) \cdot x = -x$ *für alle* $x \in \mathbb{R}$

  Es gilt

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0$$

woraus folgt, dass $(-1) \cdot x$ das Negative von x ist.

Es folgt, dass

$$-(x + y) = -x - y$$

(siehe Aufgaben) und

$$(-1) \cdot (-x) = -(-x) = x.$$

- $x \cdot (-y) = -(x \cdot y)$ *und* $(-x) \cdot (-y) = x \cdot y$.

Die erste Identität wird wie folgt bewiesen:

$$x \cdot (-y) = x \cdot ((-1) \cdot y) = (-1) \cdot (x \cdot y) = -(x \cdot y)\,.$$

Einsetzen hier $-x$ statt x ergibt

$$(-x) \cdot (-y) = -((-x) \cdot y) = -(-(x \cdot y)) = x \cdot y.$$

Insbesondere gilt

$$(-1) \cdot (-1) = 1.$$

1.2.3 Folgerungen aus den Anordnungsaxiomen

Die Relation $x \leq y$ wird auch als $y \geq x$ geschrieben. Gilt $x \leq y$ und $x \neq y$, so schreibt man $x < y$ oder $y > x$ (echte Ungleichung).

- *$x < y$ und $y \leq z$ ergeben $x < z$.*

Nach der Transitivität haben wir $x \leq z$. Es bleibt nur zu zeigen, dass $x \neq z$. Nehmen wir das Gegenteil an, dass $x = z$. Dann gelten $x \leq y$ und $y \leq x$, und die Antisymmetrie impliziert, dass $x = y$, was nach $x < y$ nicht möglich ist. Analog beweist man, dass $x \leq y$ und $y < z$ ergeben $x < z$.

- *Für je zwei Zahlen $x, y \in \mathbb{R}$ ist genau eine der Relationen $x < y$, $x = y$, $x > y$ wahr.*

Gilt $x = y$, so gelten $x < y$ und $x > y$ nicht. Nun nehmen wir an, dass $x \neq y$. Nach der Vergleichbarkeit gilt $x \leq y$ oder $x \geq y$. Diese zwei Relationen können nicht gleichzeitig gelten, weil sie sonst implizieren $x = y$. Daraus folgt, dass genau eine der Relationen $x < y$ und $x > y$ gilt.

- *$a \leq b$ und $x \leq y$ implizieren $a + x \leq b + y$.*

Wir haben

$$a + x \leq b + x \quad \text{und} \quad b + x \leq b + y$$

woraus $a + x \leq b + y$ folgt.

- *$a \leq b$ und $x < y$ implizieren $a + x < b + y$.*

Siehe Aufgaben.

- *Die folgenden Ungleichungen sind äquivalent:*

$$x \leq y \quad \Leftrightarrow \quad y - x \geq 0 \quad \Leftrightarrow \; -x \geq -y\ .$$

Addieren $(-x)$ zu den beiden Seiten von $x \leq y$ ergibt

$$x \leq y \quad \Leftrightarrow \quad 0 \leq y - x$$

und Addieren weiter $(-y)$ ergibt

$$0 \leq y - x \quad \Leftrightarrow \quad -y \leq -x.$$

- *Die folgenden Ungleichungen sind äquivalent:*

$$x < y \quad \Leftrightarrow \quad y - x > 0 \quad \Leftrightarrow \; -x > -y\ . \tag{1.11}$$

Der Beweis ist analog zum vorigen Beweis, nur benutzen wir immer $<$ anstelle $\leq$.

1.2.4 Positive und negative Zahlen

Eine Zahl $x \in \mathbb{R}$ heißt *positiv* falls $x > 0$ und *negativ* falls $x < 0$. Anwendung von (1.11) mit $y = 0$ ergibt, dass
$$x \text{ negativ } \Leftrightarrow -x \text{ positiv},$$
und analog
$$x \text{ positiv } \Leftrightarrow -x \text{ negativ}.$$

• *Sind die beiden Zahlen x und y gleichzeitig positiv oder negativ, so ist $x \cdot y$ positiv (insbesondere $x \cdot x > 0$ falls $x \neq 0$). Ist eine von x, y positiv und andere negativ, so ist $x \cdot y$ negativ.*

Falls $x, y > 0$, ergibt das Axiom $x \cdot y \geq 0$. Da $x, y \neq 0$, erhalten wir $x \cdot y \neq 0$, woraus $x \cdot y > 0$ folgt. Falls $x, y < 0$, dann $-x, -y > 0$ und
$$x \cdot y = (-x) \cdot (-y) > 0.$$
Falls $x > 0$ und $y < 0$, dann $-y > 0$,
$$-(x \cdot y) = x \cdot (-y) > 0$$
woraus $x \cdot y < 0$. Das Gleiche gilt wenn $x < 0$ und $y > 0$.

• $1 > 0$ *und* $(-1) < 0$.

Da $1 \neq 0$, gilt $1 \cdot 1 > 0$ nach der vorigen Aussage. Da $1 = 1 \cdot 1$, erhalten wir $1 > 0$. Es folgt, dass $(-1) < 0$.

Behauptung *Sei $x \leq y$. Ist $a \geq 0$, so gilt $a \cdot x \leq a \cdot y$. Ist $a \leq 0$, so gilt $a \cdot x \geq a \cdot y$.*

Die Bedingungen $x \leq y$ und $a \geq 0$ ergeben $y - x \geq 0$ und
$$a \cdot y - a \cdot x = a \cdot (y - x) \geq 0,$$
woraus $a \cdot x \leq a \cdot y$ folgt. Für $a \leq 0$ haben wir $-a \geq 0$ und
$$-(a \cdot x) = (-a) \cdot x \leq (-a) \cdot y = -(a \cdot y),$$
woraus $a \cdot x \geq a \cdot y$ folgt.

• *Sei $x < y$. Ist $a > 0$, so gilt $a \cdot x < a \cdot y$. Ist $a < 0$, so gilt $a \cdot x > a \cdot y$.*

Die Bedingungen $x < y$ und $a > 0$ ergeben $y - x > 0$ und
$$a \cdot y - a \cdot x = a \cdot (y - x) > 0,$$
woraus $a \cdot x < a \cdot y$ folgt. Der Fall $a < 0$ wird wie in dem obigen Beweis behandelt.

• *Ist $x > 0$, so gilt $x^{-1} > 0$. Für alle $x \geq y > 0$ gilt $x^{-1} \leq y^{-1}$. Ist zusätzlich $x > y$, so gilt $x^{-1} < y^{-1}$.*

Da $x \cdot x^{-1} = 1$, ist x^{-1} nicht Null. Wäre $x^{-1} < 0$, so würden wir erhalten
$$1 = x \cdot x^{-1} < 0,$$
was im Widerspruch zu $1 > 0$ steht. Deshalb muss x^{-1} positiv sein.

Für die zweite Aussage bemerken wir, dass $y \leq x$ und $x^{-1} > 0$ ergeben

$$x^{-1} \cdot y \leq x^{-1} \cdot x = 1$$

woraus folgt

$$\left(x^{-1} \cdot y\right) \cdot y^{-1} \leq 1 \cdot y^{-1} = y^{-1}.$$

Da

$$\left(x^{-1} \cdot y\right) \cdot y^{-1} = x^{-1} \cdot \left(y \cdot y^{-1}\right) = x^{-1},$$

so erhalten wir $x^{-1} \leq y^{-1}$. Die echte Ungleichung wird analog bewiesen.

Damit haben wir praktisch alle notwendigen "Bausteine" geschafft, mit denen man weiter die elementare Algebra von reellen Zahlen entwickeln kann. Insbesondere haben wir alle grundlegenden Regeln für Umformen von algebraischen Gleichungen und Ungleichungen bewiesen.

1.2.5 Supremum und Infimum

Sei S eine nicht leere Teilmenge von $\mathbb{R}$.

Definition. Eine reelle Zahl a heißt *obere Schranke* von S falls gilt: $x \leq a$ für alle $x \in S$. Analog heißt a *untere Schranke* von S falls gilt: $x \geq a$ für alle $x \in S$.

Definition. Die kleinste obere Schranke von S heißt das *Supremum* (oder *obere Grenze*) von S und wird mit $\sup S$ bezeichnet. Die größte untere Schranke von S heißt das *Infimum* (oder *untere Grenze*) von S und wird mit $\inf S$ bezeichnet.

Supremum und Infimum existieren nicht immer. Zum Beispiel, die Menge $S = \mathbb{R}$ hat weder Supremum noch Infimum. In diesem Fall gibt es sogar keine obere bzw untere Schranke: für jedes $a \in \mathbb{R}$ ist $x = a + 1$ echt grösser als a, so dass a keine obere Schranke ist, und $x = a - 1$ ist echt kleiner als a, so dass a keine untere Schranke ist. Andererseits hat die Menge $S = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ das Supremum 1 und das Infimum 0 (siehe Satz 1.10 unterhalb).

Es ist klar aus den obigen Definitionen, dass

$$\forall x \in S \qquad \inf S \leq x \leq \sup S,$$

vorausgesetzt, dass $\sup S$ und $\inf S$ existieren.

Definition. Die Menge S heißt *nach oben (bzw unten) beschränkt* falls S eine obere (bzw untere) Schranke besitzt. Die Menge S heißt *beschränkt* falls S nach oben *und* nach unten beschränkt ist. Die Menge S heißt *halbbeschränkt* falls S nach oben *oder* nach unten beschränkt ist.

Satz 1.9 *Sei S eine nicht leere Teilmenge von $\mathbb{R}$. Ist S nach oben beschränkt, so hat S das Supremum. Ist S nach unten beschränkt, so hat S das Infimum. Ist S beschränkt, so hat S das Supremum und das Infimum.*

Beweis. Bezeichnen wir mit T die Menge von oberen Schranken von S. Ist S nach oben beschränkt, so ist T nicht leer. Für alle $x \in S$ und $y \in T$ gilt nach Definition $x \leq y$. Das Vollständigkeitsaxiom ergibt: es existiert die Zahl $c \in \mathbb{R}$, die S und T trennt, d.h.

$$x \leq c \leq y \quad \forall x \in S \quad \forall y \in T.$$

Da $x \leq c$ für alle $x \in S$, ist c eine obere Schranke von S und deshalb $c \in T$. Da $c \leq y$ für alle $y \in T$, ist c das kleinste Element von T. Nach Definition $\sup S = c$. Die Existenz des Infimums wird analog bewiesen. ■

1.2.6 Intervalle

Für je zwei reelle Zahlen a, b mit $a \leq b$ definieren wir die Intervalle wie folgt:

$$\begin{aligned}
(a,b) &= \{x \in \mathbb{R} : a < x < b\} && \text{– offenes Intervall}\\
(a,b] &= \{x \in \mathbb{R} : a < x \leq b\} && \text{– halboffenes (linksoffenes) Intervall}\\
[a,b) &= \{x \in \mathbb{R} : a \leq x < b\} && \text{– halboffenes (rechtsoffenes) Intervall}\\
[a,b] &= \{x \in \mathbb{R} : a \leq x \leq b\} && \text{– abgeschlossenes Intervall}
\end{aligned}$$

Die Zahlen a, b heißen die Grenzen des Intervalles.

Satz 1.10 *Sei S eines von den obigen Intervallen mit $a < b$. Dann ist S nicht leer und*

$$\inf S = a, \qquad \sup S = b.$$

Beweis. Zeigen wir zunächst, dass das Intervall $(0,1)$ nicht leer ist. Setzen wir $2 := 1+1$ und bemerken, dass $2 > 1$ weil $1 > 0$. Daraus folgt $0 < 2^{-1} < 1^{-1}$, d.h. $\frac{1}{2} := 2^{-1} \in (0,1)$. Somit ist das Intervall $(0,1)$ nicht leer. Für beliebige $a < b$ zeigen wir, dass

$$c := \frac{1}{2} \cdot (a+b)$$

ein Element von (a,b) ist. Wir haben

$$a + b < b + b = 2b,$$

woraus folgt

$$\frac{1}{2} \cdot (a+b) < b,$$

d.h. $c < b$. Analog beweist man $c > a$ und somit $c \in S$ und $S \neq \emptyset$.

Beweisen wir jetzt, dass $\inf S = a$. Nach definition von S ist a eine untere Schranke von S. Zeigen wir, dass a die größte untere Schranke ist, d.h. für jede andere untere Schranke a' von S gilt $a' \leq a$. Nehmen wir zunächst das Gegenteil an, dass $a' > a$. Da a' eine untere Schranke von S ist und $c \in S$, erhalten wir $a' \leq c$ und damit $a' < b$, woraus folgt $(a, a') \subset (a,b)$. Setzen wir

$$c' = \frac{1}{2} \cdot (a + a')$$

und bemerken, dass nach dem obigen Argument

$$c' \in (a, a') \subset (a,b) \subset S$$

so dass $c' \in S$. Da a' eine untere Schranke von S ist, so erhalten wir $a' \leq c'$. Andererseits haben wir $c' \in (a, a')$ und somit $c' < a'$. Dieser Widerspruch beschließt den Beweis. Die Identität $\sup S = b$ wird analog bewiesen. ■

Bemerkung. Vor dem Satz 1.10 hatten wir nur die Zahlen 0 und 1 gesehen. Im obigen Beweis haben wir die weiteren Zahlen 2 und $\frac{1}{2}$ bestimmt, die von 0 und 1 abweichen. Die Körperaxiome allein erlauben die Existenz von solchen Zahlen nicht beweisen. Zum Beispiel, betrachten wir die Menge $K = \{0, 1\}$, die aus zwei Elementen 0 und 1 besteht, und definieren Addition und Multiplikation in K mit

$$0 + 0 = 0, \quad 0 + 1 = 1 + 0 = 1, \quad 1 + 1 = 0$$

und

$$0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0, \quad 1 \cdot 1 = 1.$$

Dann werden alle Axiome von Addition und Multiplikation erfüllt, so dass K ein Körper ist, aber in K gilt $2 = 0$. Dass in $\mathbb{R}$ gilt $2 \neq 0$ ist eine Folgerung von Anordnungsaxiomen.

1.2.7 Maximum und Minimum

Definition. Eine Zahl $a \in \mathbb{R}$ heißt das *Maximum* von S (oder das *minimale Element* von S) und wird mit $\max S$ bezeichnet, falls $a \in S$ und $x \leq a$ für alle $x \in S$. Eine Zahl $b \in \mathbb{R}$ heißt das *Minimum* von S (oder das *minimale Element* von S) und wird mit $\min S$ bezeichnet, falls $b \in S$ und $x \geq b$ für alle $x \in S$.

Behauptung. *Existiert* $\max S$*, so existiert auch* $\sup S$ *und gilt* $\max S = \sup S$*. Existiert* $\min S$*, so existiert auch* $\inf S$ *und gilt* $\min S = \inf S$*.*

Beweis. Zunächst bemerken wir, dass $a := \max S$ nach Definition eine obere Schranke von S ist. Beweisen wir, dass a die kleinste obere Schranke ist, d.h. das Supremum. Sei c eine andere obere Schranke. Da $a \in S$, so gilt $a \leq c$, was zu beweisen war. Analog beweist man, dass $\min S = \inf S$ vorausgesetzt, dass $\min S$ existiert. ■

Beispiel. Das abgeschlossene Intervall $S = [a, b]$ mit $a < b$ hat offensichtlich $\max S = a$ und $\max S = b$. Das offene Intervall $S = (a, b)$ hat weder Maximum noch Minimum. In der Tat, existiert $\max S$, so gilt $\max S = \sup S = b$, was in Widerspruch zu $b \notin S$ steht. Analog hat das linksoffene Intervall $[a, b)$ das Minimum a und kein Maximum, und das rechtsoffene Intervall $(a, b]$ das Maximum b und kein Minimum.

1.2.8 Die Zeichen $+\infty$ und $-\infty$

Nach Satz 1.9 existiert $\sup S$ bzw $\inf S$ für jede nach oben bzw nach unten beschränkte nichtleere Menge $S \subset \mathbb{R}$. Definieren wir jetzt $\sup S$ und $\inf S$ für *alle* Mengen $S \subset \mathbb{R}$ mit Hilfe von den Zeichen $+\infty$ und $-\infty$, die die *unendlichen Elemente* heißen. Man definiert die *erweiterte Menge* von reellen Zahlen

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$$

und erweitert die Anordnungsrelation $\leq$ auf $\overline{\mathbb{R}}$ wie folgt: für alle $a \in \overline{\mathbb{R}}$ gilt immer

$$-\infty \leq a \leq +\infty.$$

Die Begriffe von oberen bzw unteren Schranken und somit sup und inf definiert man für nichtleere Teilmengen S von $\overline{\mathbb{R}}$ genau so, wie für Teilmengen von $\mathbb{R}$. Für die leere Menge $S = \emptyset$ nehmen wir folgendes nach Definition an[2]:

$$\sup \emptyset = -\infty, \quad \inf \emptyset = +\infty.$$

Korollar 1.11 *Jede Teilmenge S von $\overline{\mathbb{R}}$ hat $\sup S$ und $\inf S$ (als Elemente von $\overline{\mathbb{R}}$).*

Beweis. Sei S nicht leer. Gilt $+\infty \in S$, so gilt $\sup S = +\infty$. Besteht S aus einem Element $-\infty$, so gilt $\sup S = -\infty$. Sonst ist die Menge $S' = S \setminus \{-\infty\}$ eine nichtleere Teilmenge von $\mathbb{R}$. Ist S' nach oben beschränkt, so existiert $\sup S'$ nach dem Satz 1.9, und dann gilt $\sup S = \sup S'$. Ist S' nicht nach oben beschränkt, so gilt $\sup S = \sup S' = +\infty$. Analog beweist man die Existenz von $\inf S$. ∎

Der Begriff von Intervall lässt sich auch verallgemeinern zu dem Fall, wenn die Grenzen a, b die Elemente von $\overline{\mathbb{R}}$ sind. Zum Beispiel, für $a \in \mathbb{R}$, haben wir

$$\begin{aligned}
[a, +\infty) &= \left\{x \in \overline{\mathbb{R}} : a \leq x < +\infty\right\} = \{x \in \mathbb{R} : x \geq a\} \\
(-\infty, a] &= \left\{x \in \overline{\mathbb{R}} : -\infty < x \leq a\right\} = \{x \in \mathbb{R} : x \leq a\} \\
[a, +\infty] &= \left\{x \in \overline{\mathbb{R}} : a \leq x \leq +\infty\right\} = [a, +\infty) \cup \{+\infty\}.
\end{aligned}$$

Auch gelten $(-\infty, +\infty) = \mathbb{R}$ und $[-\infty, +\infty] = \overline{\mathbb{R}}$. Satz 1.10 gilt auch in diesem Fall.

1.3 Ganze Zahlen und vollständige Induktion

1.3.1 Natürliche Zahlen

In diesem Abschnitt definieren wir den Begriff von natürlichen Zahlen. Man erwartet, dass die Menge $\mathbb{N}$ von natürlichen Zahlen eine Teilmenge von $\mathbb{R}$ ist, die folgenden Eigenschaften erfüllt:

- $1 \in \mathbb{N}$
- $x \in \mathbb{N}$ ergibt $x + 1 \in \mathbb{N}$.

Jedoch bestimmen these diese zwei Eigenschaften die Menge $\mathbb{N}$ nicht eindeutig. Zum Beispiel, die ganze Menge $\mathbb{R}$ erfüllt sie. Um eine vollständiger Definition von $\mathbb{N}$ zu geben, führen wir den folgenden Begriff ein.

Definition. Eine Menge $S \subset \mathbb{R}$ heißt *induktiv* falls $x \in S$ ergibt $x + 1 \in S$.

Zum Beispiel ist die ganze Menge $\mathbb{R}$ induktiv. Auch die Intervalle $(a, +\infty)$ und $[a, +\infty)$ sind induktive Mengen, während die beschränkten Intervalle nicht induktiv ist.

Definition. Die Menge $\mathbb{N}$ ist der Durchschnitt von allen induktiven Mengen die 1 enthalten. Die Elemente von $\mathbb{N}$ heißen *natürliche Zahlen*.

[2] Die Motivation dafür ist wir folgt. Für die leere Menge sind alle Elemente von $\overline{\mathbb{R}}$ die obere und untere Schranken. Deshalb ist $-\infty$ die kleinste obere Schranke und $+\infty$ die größte untere Schranke.

Satz 1.12 *Die Menge* $\mathbb{N}$ *ist induktiv. Folglich ist* $\mathbb{N}$ *die kleinste induktive Menge die* 1 *enthält.*

Das Wort "kleinste" bezieht sich auf Inklusion, d.h. $\mathbb{N} \subset S$ für jede andere induktive Menge S die 1 enthält.

Beweis. Sei $\mathcal{F}$ das Mengensystem von allen induktiven Mengen, die 1 enthalten. Die Familie $\mathcal{F}$ ist nicht leer da $\mathbb{R} \in \mathcal{F}$. Nach Definition haben wir

$$\mathbb{N} = \bigcap_{S \in \mathcal{F}} S. \tag{1.12}$$

Da jedes S induktiv ist, so erhalten wir

$$x \in \mathbb{N} \implies x \in S \quad \forall S \in \mathcal{F} \implies x + 1 \in S \quad \forall S \implies x + 1 \in \mathbb{N},$$

d.h. $\mathbb{N}$ ist induktiv. Da $1 \in S$ für jedes $S \in \mathcal{F}$, daraus folgt, dass $1 \in \mathbb{N}$, insbesondere $\mathbb{N} \in \mathcal{F}$. Die Menge $\mathbb{N}$ ist das kleinste Element von $\mathcal{F}$, da nach (1.12) $\mathbb{N} \subset S$ gilt für jedes $S \in \mathcal{F}$. ■

Für jedes $n \in \mathbb{N}$ heißt die natürliche Zahl $n + 1$ der *Nachfolger* von n.

Beispiel. Das Intervall $[1, +\infty)$ ist eine induktive Menge die 1 enthält. Es folgt, dass $\mathbb{N} \subset [1, +\infty)$. Insbesondere ist 1 die kleinste natürliche Zahl. Da $1 \in \mathbb{N}$, daraus folgt auch $2 := 1 + 1 \in \mathbb{N}$, $3 := 2 + 1 \in \mathbb{N}$ usw. bis $9 := 8 + 1 \in \mathbb{N}$. Die Zahl $zehn = 9 + 1$ wird mit 10 bezeichnet, aber diese Bezeichnung hat mit *Stellenwertsystem* zu tun und wird später betrachtet.

1.3.2 Induktionsprinzip

Häufig braucht man eine Aussage beweisen, die von einem natürlichen Parameter n abhängt. Dafür benutzt man das folgende Induktionsprinzip.

Behauptung. (Induktionsprinzip) *Sei* $A(n)$ *eine von* $n \in \mathbb{N}$ *abhängige Aussage, die die folgenden Bedingungen erfüllt:*

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

(*ii*) *Für jedes* $n \in \mathbb{N}$ *gilt* $A(n) \Rightarrow A(n + 1)$ *(Induktionsschritt).*

Dann ist $A(n)$ *wahr für alle* $n \in \mathbb{N}$.

Beweis. Bezeichnen wir mit S die Menge von allen n, für die $A(n)$ wahr ist. Nach dem Induktionsanfang $1 \in S$, und nach dem Induktionsschritt ist S induktiv. Nach Definition von $\mathbb{N}$ gilt $\mathbb{N} \subset S$, d.h. $A(n)$ ist wahr für alle $n \in \mathbb{N}$. ■

Die Beweismethode, die das Induktionsprinzip benutzt, heißt die *vollständige Induktion.* Im Induktionsschritt nennt man $A(n)$ die *Induktionsvoraussetzung* und $A(n + 1)$ die *Induktionsbehauptung.*

Der folgende Satz wird mit Hilfe von vollständiger Induktion bewiesen.

Satz 1.13 *Seien* n, m *zwei natürliche Zahlen. Dann gilt das folgende.*

(a) $n+m \in \mathbb{N}$

(b) $nm \in \mathbb{N}$

(c) *Ist* $n > m$, *so gilt* $n - m \in \mathbb{N}$.

Beweis. (a) Wir beweisen per Induction nach n, dass $n+m \in \mathbb{N}$. Wählen wir ein $m \in \mathbb{N}$ und bezeichnen with $A(n)$ die Aussage, dass $n+m \in \mathbb{N}$.

- Induktionsanfang. $A(1)$ ist die Aussage, dass $1+m \in \mathbb{N}$. Da $\mathbb{N}$ eine induktive Menge ist und $m \in \mathbb{N}$, daraus folgt, dass $m+1 \in \mathbb{N}$, was zu beweisen war.

- Induktionsschritt. Die Induktionsvoraussetzung $A(n)$ ist: $n+m \in \mathbb{N}$, die Induktionsbehauptung $A(n+1)$ ist: $(n+1)+m \in \mathbb{N}$. Ist $A(n)$ wahr, d.h. $n+m \in \mathbb{N}$, dann auch $(n+m)+1 \in \mathbb{N}$, woraus folgt

$$(n+1)+m = (n+m)+1 \in \mathbb{N}.$$

Damit ist $A(n+1)$ bewiesen. Nach dem Induktionsprinzip beschließen wir, dass $A(n)$ für alle $n \in \mathbb{N}$ gilt.

(b) Beweisen wir per Induktion nach n, dass $nm \in \mathbb{N}$.

- Induktionsanfang. Ist $n = 1$, dann $nm = m \in \mathbb{N}$.

- Induktionsschritt. Ist es schon bekannt, dass $nm \in \mathbb{N}$, dann nach (a)

$$(n+1)\,m = nm + m \in \mathbb{N}$$

da die beiden Zahlen nm und m natürlich sind.

(c) Beweisen wir zunächst diese Behauptung für $m = 1$, d.h.

$$n > 1 \Rightarrow n - 1 \in \mathbb{N}.$$

Das letztere ist äquivalent zu der folgenden Aussage

$$A(n) = (n = 1 \ \vee \ \ n - 1 \in \mathbb{N}),$$

die wir per Induktion nach n beweisen.

- Induktionsanfang. $A(1)$ ist wahr, da $n = 1$.

- Induktionsschritt. $A(n+1)$ ist wahr, da $(n+1) - 1 = n \in \mathbb{N}$.

Beweisen wir per Induktion nach m die neue Aussage $B(m)$:

$$\text{für jedes } n \in \mathbb{N} \text{ mit } n > m \text{ gilt } n - m \in \mathbb{N}.$$

- Induktionsanfang. $B(1)$ bedeutet: für jedes $n > 1$ gilt $n - 1 \in \mathbb{N}$, was aus $A(n)$ folgt.

- Induktionsschritt. Angenommen, dass $B(m)$ wahr ist, beweisen wir $B(m+1)$, d.h.

$$n \in \mathbb{N} \text{ und } n > m+1 \Rightarrow n-(m+1) \in \mathbb{N}.$$

Da $n > m$, so erhalten with nach der Induktionsvoraussetzung, dass $n-m \in \mathbb{N}$. Nach $n > m+1$ erhalten wir $n-m > 1$ und somit nach dem Induktionsanfang $(n-m)-1 \in \mathbb{N}$. Es bleibt nun zu bemerken, dass

$$n-(m+1) = n+(-1)(m+1) = n+(-m)+(-1) = (n-m)-1,$$

woraus $n-(m+1) \in \mathbb{N}$ folgt. ■

1.3.3 Ganze Zahlen

Bezeichnen with mit $\mathbb{Z}$ die Vereinigung von $\{0\}$, $\mathbb{N}$, und von den Negativen von $\mathbb{N}$, d.h.

$$\mathbb{Z} := \{0\} \cup \mathbb{N} \cap (-\mathbb{N}).$$

Die Elemente von $\mathbb{Z}$ heißen die *ganze Zahlen* und $\mathbb{Z}$ heißt die Menge von ganzen Zahlen. Offensichtlich

$$x \in \mathbb{Z} \text{ und } x > 0 \Leftrightarrow x \in \mathbb{N}.$$

Korollar 1.14 *Für jede $x, y \in \mathbb{Z}$ sind $x+y$, $x-y$, xy auch in $\mathbb{Z}$.*

Beweis. Da $x-y = x+(-y)$ und $(-y) \in \mathbb{Z}$, es reicht zu beweisen, dass xy und $x+y$ ganze Zahlen sind. Ist $x=0$ oder $y=0$, so ist die Aussage trivial. Sonst gibt es natürliche Zahlen n, m mit $x = \pm n$ und $y = \pm m$. Betrachten wir, zum Beispiel, den Fall $x = n$ und $y = -m$. Dann $nm \in \mathbb{N}$ und $xy = n(-m) = -nm \in \mathbb{Z}$. Beweisen wir jetzt, dass $x+y = n-m$ ist eine ganze Zahl. Ist $n = m$ dann $n-m = 0 \in \mathbb{Z}$. Ist $n > m$ dann $n-m \in \mathbb{N}$ nach Satz 1.13. Ist $n < m$, dann $m-n \in \mathbb{N}$ und $n-m = -(m-n) \in \mathbb{Z}$.

Die anderen Fälle von Vorzeichen in $x = \pm n$ und $y = \pm m$ werden analog betrachtet. ■

Korollar 1.15 *Für $x, y \in \mathbb{Z}$ ist die Bedingung $x > y$ äquivalent zu $x \geq y+1$. Folglich, für jedes $n \in \mathbb{Z}$ enthält das Intervall $(n, n+1)$ keine ganze Zahl.*

Beweis. Sei $x > y$. Da $x-y \in \mathbb{Z}$ und $x-y > 0$, erhalten wir $x-y \in \mathbb{N}$. Da 1 die kleinste natürliche Zahl ist, folgt $x-y \geq 1$ und damit $x \geq y+1$. Die Implikation in der Rückrichtung ist offensichtlich: $x \geq y+1$ und $y+1 > y$ ergeben $x > y$.

Ist m eine ganze Zahl in $(n, n+1)$, so gilt $m > n$ und somit $m \geq n+1$, was im Widerspruch zu $m < n+1$ steht. ■

Behauptung. (Verallgemeinerung des Induktionsprinzipes) *Sei n_0 eine ganze Zahl und sei $A(n)$ eine von n abhängige Aussage. Angenommen ist das folgende:*

- *$A(n_0)$ ist wahr.*
- *Für jedes $n \in \mathbb{Z}$ mit $n \geq n_0$ gilt $A(n) \Rightarrow A(n+1)$.*

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

1.3.4 Maximum und Minimum der Teilmenge von $\mathbb{Z}$

Satz 1.16 *Sei S eine nicht-leere Teilmenge von $\mathbb{Z}$. Ist S nach oben beschränkt, so existiert* $\max S$. *Ist S nach unten beschränkt, so existiert* $\min S$. *Insbesondere existiert* $\min S$ *für jede nicht-leere Teilmenge von $\mathbb{N}$.*

Beweis. Sie S nach oben beschränkt. Nach Satz 1.9 existiert $\sup S$ als Element von $\mathbb{R}$. Setzen wir $a = \sup S$ und beweisen wir, dass $a \in S$, woraus die Identität $a = \max S$ folgen wird. Nehmen wir das Gegenteil an, dass $a \notin S$. Da $a - 1$ keine obere Schranke von S ist, dann existiert $n \in S$ mit $n > a - 1$. Da a eine obere Schranke von S ist, so haben wir $n \leq a$. Da $a \notin S$ und $n \in S$, erhalten wir $a \neq n$ und somit $n < a$. Deshalb haben wir gezeigt, dass $a - 1 < n < a$, was äquivalent zu

$$n < a < n + 1$$

ist. Nach Korollar 1.15 enthält das Intervall $(n, n + 1)$ keine ganze Zahl, insbesondere kein Element von S. Auch das Intervall $[n + 1, +\infty)$ enthält kein Element von S, da alle Elemente von S von a beschränkt sind. Daraus folgt, dass das Intervall $(n, +\infty)$ kein Element von S enthält und somit n eine obere Schranke von S ist, was im Widerspruch mit $n < a$ steht.

Der Fall von $\min S$ wird analog behandelt.

Die Teilmengen von $\mathbb{N}$ sind immer nach unten von 1 beschränkt, woraus die zweite Aussage folgt. ■

Satz 1.17 (Archimedisches Prinzip) *Für jedes $x \in \mathbb{R}$ existiert genau ein $n \in \mathbb{Z}$ mit*

$$n \leq x < n + 1. \tag{1.13}$$

Die Zahl n ist damit die größte ganze Zahl mit $n \leq x$. Diese Zahl n wird mit $[x]$ (x in den rechteckigen Klammern) bezeichnet und heißt die *Gaußklammer* von x. Der Wert von n heißt auch der *Ganzzahlanteil* von x. Zum Beispiel, $\left[\frac{1}{2}\right] = 0$ und $\left[-\frac{1}{2}\right] = -1$. Für $x \in \mathbb{Z}$ gilt $[x] = x$.

Beweis. Für gegebenes $x \in \mathbb{R}$ betrachten wir die Menge

$$S = \{k \in \mathbb{Z} : k \leq x\}.$$

Zeigen wir zunächst, dass diese Menge nicht leer ist. Nehmen wir das Gegenteil an, dass S leer ist. Dann gilt $k > x$ für alle $k \in \mathbb{Z}$, d.h. x eine untere Schranke von $\mathbb{Z}$ ist. Somit ist $\mathbb{Z}$ nach unten beschränkt und nach Satz 1.16 hat $\mathbb{Z}$ das Minimum. Sei $m = \min \mathbb{Z}$. Aber dann ist $m - 1$ auch in $\mathbb{Z}$, was im Widerspruch mit $m - 1 < \min \mathbb{Z}$ steht.

Da S nicht-leer ist und nach oben mit x beschränkt, erhalten wir nach Satz 1.16, dass S das Maximum hat. Setzen wir $n = \max S$. Da $n \in S$, erhalten wir nach Definition von S, dass $n \leq x$. Da $n + 1 > n = \max S$ und deshalb $n + 1 \notin S$, erhalten wir, dass $n + 1 > x$. Damit erfüllt n die Bedingungen (1.13).

Um die Eindeutigkeit von n zu beweisen, nehmen wir zunächst das Gegenteil an, dass es noch ein $n' \in \mathbb{Z}$ gibt mit

$$n' \leq x < n' + 1.$$

Daraus folgt, dass $n' < n + 1$ und somit nach Korollar 1.15 $n' \leq n$. Analog sieht man, dass $n \leq n'$, woraus $n = n'$ folgt. ■

1.4 Rationale Zahlen

Eine *rationale* Zahl ist ein Quotient von ganzen Zahlen, d.h. eine reelle Zahl der Form $\frac{a}{b}$ mit $a, b \in \mathbb{Z}$, $b \neq 0$. Die Menge von allen rationalen Zahlen wird mit $\mathbb{Q}$ bezeichnet, so dass

$$\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

Offensichtlich gelten die Inklusionen

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

Da $0 < \frac{1}{2} < 1$, so erhalten wir nach Korollar 1.15, dass $\frac{1}{2}$ keine ganze Zahl ist. Da $\frac{1}{2} \in \mathbb{Q}$, so sehen wir, dass die Inklusion $\mathbb{Z} \subset \mathbb{Q}$ echt ist.

Satz 1.18 *Die Menge $\mathbb{Q}$ ist angeordneter Körper.*

Beweis. Wir müssen nur zeigen, dass $\mathbb{Q}$ bezüglich der Operationen Addition und Multiplikation abgeschlossen ist, und auch das Negative und Inverse von rationalen Zahlen auch in $\mathbb{Q}$ liegen. Dann alle Axiome der Addition, Multiplikation und Ordnung gelten automatisch, da $\mathbb{Q}$ Teilmenge von $\mathbb{R}$ ist.

Die obigen Aussagen folgen von den folgende Identitäten, wobei a, b, c, d ganze Zahlen sind:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bd}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$$

und

$$-\frac{a}{b} = \frac{-a}{b}, \qquad \left(\frac{a}{b}\right)^{-1} = \frac{b}{a}.$$

Hier die Nenner b, d sind nicht gleich Null, und in der letzten Identität auch $a \neq 0$. Für Beweis siehe Aufgaben. ∎

Bemerken wir, dass $\mathbb{Q}$ nicht vollständig ist, d.h., die Vollständigkeitsaxiom nicht erfüllt ist. In der Tat impliziert die Vollständigkeitsaxiom axiom die Existenz von quadratischer Wurzel $\sqrt{a}$ für jedes $a \geq 0$, d.h. die Existenz der Zahl x mit $x^2 = a$. Im Körper $\mathbb{Q}$ ist das nicht der Fall: es gibt kein $\sqrt{2}$ in $\mathbb{Q}$. Für weitere Einzelheiten siehe Aufgaben. Insbesondere ist die Inklusion $\mathbb{Q} \subset \mathbb{R}$ echt.

1.5 Endliche Folgen und Mengen

1.5.1 Endliche Folgen

Für je zwei ganze Zahlen n, m mit $n \leq m$, betrachten wie die Menge

$$\{k \in \mathbb{Z} : n \leq k \leq m\} = [n, m] \cap \mathbb{Z}$$

von ganzen Zahlen zwischen n und m und bezeichnen sie kurz mit $\{n, ..., m\}$.

Definition. Eine *endliche Folge* ist eine Abbildung $a : \{n, ..., m\} \to \mathbb{R}$. Man bezeichnet den Wert $a(k)$ auch mit a_k und die Folge a mit $\{a_k\}_{k=n}^m$ oder kurz mit $\{a_k\}$.

Alle Zahlen a_k heißen die *Glieder* oder die *Elemente* der Folge a.

Definieren wir die Summe $\sum_{k=n}^m a_k$ der Folge $\{a_k\}$ per Induction nach m wie folgt:

- ist $m = n$ dann setzen wir $\sum_{k=n}^{m} a_k = a_n$;
- ist $\sum_{k=n}^{m} a_k$ schon definiert, so setzen wir

$$\sum_{k=n}^{m+1} a_k = \left(\sum_{k=n}^{m} a_k \right) + a_{m+1}. \tag{1.14}$$

Man schreibt auch

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

Alle Eigenschaften der Summe $\sum_{k=n}^{m} a_k$ werden per Induktion bewiesen. Zum Beispiel, sind alle a_k gleich a, so gilt

$$\sum_{k=n}^{m} a_k = (m - n + 1)\, a.$$

Analog definiert man das Produkt $\prod_{k=n}^{m} a_k$ der Folge $\{a_k\}$ mit:

- $\prod_{k=n}^{n} a_k = a_n$
- $\prod_{k=n}^{m+1} a_k = \left(\prod_{k=n}^{m} a_k\right) \cdot a_{m+1}$.

Man schreibt auch

$$\prod_{k=n}^{m} a_k = a_n \cdot a_{n+1} \cdot ... \cdot a_m.$$

Betrachten wir die Folge $\{a_k\}_{k=1}^{n}$ mit $a_k = a \in \mathbb{R}$ für alle k und definieren wir die *Potenzen* von a mit

$$a^n = \prod_{k=1}^{n} a = \underbrace{a \cdot a \cdot ... \cdot a}_{n \text{ times}},$$

für alle $n \in \mathbb{N}$. Äquivalent kann man die Potenzen a^n direkt per Induktion definieren:

$$a^1 = a \quad \text{und} \quad a^{n+1} = a^n \cdot a \quad \text{für alle } n \in \mathbb{N}. \tag{1.15}$$

Man beweist per Induktion die folgenden Identitäten:

$$a^n \cdot a^m = a^{n+m}, \quad (a^n)^m = a^{n \cdot m}, \quad (a \cdot b)^n = a^n \cdot b^n. \tag{1.16}$$

Ist $a \neq 0$, so definiert man die Potenzen a^n auch für nicht-positive ganze n wie folgt:

$$a^0 := 1 \quad \text{und} \quad a^n := \left(a^{-1}\right)^{-n} \quad \text{für } n < 0.$$

Die Identitäten (1.16) gelten in diesem Fall auch für alle $n, m \in \mathbb{Z}$.

1.5.2 Endliche Mengen

Definition. Eine nicht-leere Menge X heißt *gleichmächtig* zu einer Menge Y falls es eine bijektive Abbildung $f : X \to Y$ gibt. In diesem Fall schreibt man $X \sim Y$. Die leere Menge $\emptyset$ ist gleichmächtig zu sich selbst, d.h. $\emptyset \sim \emptyset$.

Behauptung. *Die Gleichmächtigkeit ist eine Äquivalenzrelation, d.h. sie erfüllt die folgenden Bedingungen:*

- $X \sim X$
- $X \sim Y \Rightarrow Y \sim X$
- $X \sim Y \ \wedge \ Y \sim Z \Rightarrow X \sim Z.$

Beweis. Der Fall mit leeren Mengen ist trivial, so nehmen wir an, dass alle Mengen X, Y, Z nicht leer sind. Da die Identitätsabbildung $\mathrm{Id}_X : X \to X$ bijektiv ist, haben wir $X \sim X$. Gilt $X \sim Y$, so existiert eine bijektive Abbildung $f : X \to Y$. Dann ist die Umkehrabbildung $f^{-1} : Y \to X$ wohldefiniert und bijektiv, woraus $Y \sim X$ folgt. Existieren die bijektiven Abbildungen $f : X \to Y$ und $g : Y \to Z$, so ist die Verkettung $g \circ f : X \to Z$ bijektiv, woraus $X \sim Z$ folgt. ■

Für jedes $n \in \mathbb{N}$ bezeichnen wir mit $\mathcal{E}_n$ die folgende Menge

$$\mathcal{E}_n := \{1, ..., n\} = \{k \in \mathbb{N} : 1 \leq k \leq n\},$$

und $\mathcal{E}_0 = \emptyset$. Es ist klar, dass $\mathcal{E}_n \subset \mathcal{E}_m$ für $n \leq m$.

Definition. Eine Menge S heißt *endlich* falls $S \sim \mathcal{E}_n$ für ein $n \in \mathbb{N} \cup \{0\}$ ist.

Für $n = 0$ bedeutet $S \sim \mathcal{E}_0$ dass S eine leere Menge ist. Für $n \in \mathbb{N}$ bedeutet $S \sim \mathcal{E}_n$ die Existenz einer Bijektion $f : \mathcal{E}_n \to S$, d.h. man kann die Menge S mit den Zahlen $1, ..., n$ aufzählen.

Definition. Ist $S \sim \mathcal{E}_n$, so sagen wir: die Anzahl von Elementen von S ist n, oder die *Kardinalität* von S ist n. Man bezeichnet die Kardinalität von S mit $\operatorname{card} S$ oder mit $|S|$ (Betrag von S).

Zunächst zeigen wir, dass die Kardinalität wohl-definiert ist, d.h. $S \sim \mathcal{E}_n$ und $S \sim \mathcal{E}_m$ sind mit verschiedenen n, m unmöglich.

Satz 1.19 *Sind n, m natürliche Zahlen mit $n > m$, so gibt es keine injektive Abbildung von $\mathcal{E}_n$ nach $\mathcal{E}_m$. Insbesondere sind $\mathcal{E}_n$ und $\mathcal{E}_m$ gleichmächtig genau dann, wenn $n = m$.*

Die Aussage von Satz 1.19 heißt das *Schubfachprinzip*: sind n Objekten zwischen m Schubfächer verteilt, wobei $n > m$, so gibt es ein Schubfach mit mindestens zwei Objekten. Es gibt viele interessante Anwendungen von diesem Prinzip.

Beweis. Soll eine injektive Abbildung $f : \mathcal{E}_n \to \mathcal{E}_m$ mit $n > m$ existieren, so ergibt die Beschränkung $f|_{\mathcal{E}_{m+1}}$ eine injektive Abbildung von $\mathcal{E}_{m+1}$ nach $\mathcal{E}_m$. Deshalb reicht es zu beweisen, dass es keine injektive Abbildung $\mathcal{E}_{m+1} \to \mathcal{E}_m$ gibt. Diese Aussage beweisen wir per Induktion nach m.

Induktionsanfang für $m = 1$. Es gibt nur eine Abbildung $f : \mathcal{E}_2 \to \mathcal{E}_1$ da $f(k)$ gleich 1 für $k = 1, 2$ sein soll, und diese Abbildung ist nicht injektiv.

Induktionsschritt von m nach $m + 1$. Sei

$$f : \mathcal{E}_{m+2} \to \mathcal{E}_{m+1}$$

eine injektive Abbildung. Setzen with $a = f(m + 2)$. Dann gilt

$$\forall k \in \mathcal{E}_{m+1} \quad f(k) \in \mathcal{E}_{m+1} \setminus \{a\},$$

und deshalb ergibt die Beschränkung von f auf $\mathcal{E}_{m+1}$ eine injektive Abbildung von $\mathcal{E}_{m+1}$ nach $\mathcal{E}_{m+1} \setminus \{a\}$. Zeigen wir, dass

$$\mathcal{E}_{m+1} \setminus \{a\} \sim \mathcal{E}_m.$$

In der Tat ist die folgende Abbildung

$$\begin{aligned} g \quad &: \quad \mathcal{E}_{m+1} \setminus \{a\} \to \mathcal{E}_m \\ g(k) \quad &= \quad \begin{cases} k, & k < a \\ k - 1, & k > a \end{cases} \end{aligned}$$

eine Bijektion. Betrachten wir die Verkettung der folgenden Abbildungen:

$$\mathcal{E}_{m+1} \overset{f}{} \xrightarrow{\text{injektiv}} \mathcal{E}_{m+1} \setminus \{a\} \overset{g}{} \xrightarrow{\text{bijektiv}} \mathcal{E}_m,$$

woraus folgt, dass die Abbildung $g \circ f : \mathcal{E}_{m+1} \to \mathcal{E}_m$ injektiv ist, was im Widerspruch zur Induktionsvoraussetzung steht.

In der zweiten Aussage ist der Fall $n = 0$ oder $m = 0$ offensichtlich. Seien $n, m \in \mathbb{N}$. Gilt $n = m$ so sind $\mathcal{E}_n$ und $\mathcal{E}_m$ offensichtlich gleichmächtig. Gilt $n > m$, so existiert keine injektive Abbildung $\mathcal{E}_n \to \mathcal{E}_m$ und um so mehr keine bijektive Abbildung $\mathcal{E}_n \to \mathcal{E}_m$. Somit sind $\mathcal{E}_n$ und $\mathcal{E}_m$ nicht gleichmächtig. Im Fall $n < m$ sind $\mathcal{E}_n$ und $\mathcal{E}_m$ analog nicht gleichmächtig. ■

Korollar 1.20 *Jede injektive Abbildung $\mathcal{E}_n \to \mathcal{E}_n$ ist immer bijektiv.*

Beweis. Sei $f : \mathcal{E}_n \to \mathcal{E}_n$ eine injektive Abbildung, die nicht surjektiv ist, d.h. es existiert ein $a \in \mathcal{E}_n$ mit $f(k) \neq a$ für alle $k \in \mathcal{E}_n$. Erweitern wir f auf $\mathcal{E}_{n+1}$, indem wir setzen

$$f(n + 1) = a.$$

Dann ist f eine injektive Abbildung von $\mathcal{E}_{n+1}$ nach $\mathcal{E}_n$, was nach Satz 1.19 nicht möglich ist. ■

Man kann auch zeigen, dass jede surjektive Abbildung $\mathcal{E}_n \to \mathcal{E}_n$ immer bijektiv ist.

Im Beweis des nächsten Satzes benutzen wir den Begriff von *disjunkter* Vereinigung.

Definition. Seien A, B zwei Mengen. Die disjunkte Vereinigung $A \sqcup B$ ist gleich die Vereinigung $A \cup B$, vorausgesetzt $A \cap B = \emptyset$. Sonst ist $A \sqcup B$ nicht definiert.

Satz 1.21 (a) *Eine Teilmenge einer endlichen Menge ist endlich.*

(*b*) *Die Vereinigung zweier endlichen Mengen ist endlich.*

(*c*) *Kartesisches Produkt zweier endlichen Mengen ist endlich.*

Beweis. (*a*) Es reicht zu beweisen, dass jede Teilmenge von $\mathcal{E}_n$ endlich ist.

Induktionsanfang. Ist $n = 0$, so ist jede Teilmenge von $\mathcal{E}_0$ leer und somit endlich.

Induktionsschritt. Angenommen sei, dass jede Teilmenge von $\mathcal{E}_n$ endlich ist. Beweisen wir, dass jede Teilmenge $S \subset \mathcal{E}_{n+1}$ endlich ist. Betrachten wir zwei Fälle.

Gilt $S \subset \mathcal{E}_n$, so ist S endlich nach Induktionsvoraussetzung.

Gilt $S \not\subset \mathcal{E}_n$, so enthält S die Zahl $n+1$. Die Menge $S' = S \setminus \{n+1\}$ ist eine Teilmenge von $\mathcal{E}_n$ und somit ist endlich nach Induktionsvoraussetzung, d.h. $S' \sim \mathcal{E}_m$ für ein $m \in \mathbb{N} \cup \{0\}$. Dann erhalten wir

$$S = S' \sqcup \{n+1\} \sim \mathcal{E}_{m+1},$$

was die Endlichkeit von S beweist.

(*b*) Für zwei Mengen A, B gilt

$$A \cup B = (A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A) .$$

Nach Teil (*a*) sind alle Mengen $A \setminus B$, $B \setminus A$, $A \cap B$ endlich. Deshalb reicht es zu beweisen, dass eine disjunkte Vereinigung zweier Mengen endlich ist. Seien A, B zwei disjunkte Mengen mit

$$A \sim \mathcal{E}_n \quad \text{und} \quad B \sim \mathcal{E}_m. \tag{1.17}$$

Betrachten wir die Menge

$$\mathcal{E}'_m = \{n+1, ..., n+m\} .$$

Offensichtlich gilt $\mathcal{E}_m \sim \mathcal{E}'_m$ (mit Bijektion $k \mapsto n+k$) somit $B \sim \mathcal{E}'_m$. Die Mengen $\mathcal{E}_n$ und $\mathcal{E}'_m$ sind disjunkt, woraus folgt, dass

$$A \sqcup B \sim \mathcal{E}_n \sqcup \mathcal{E}'_m.$$

Andererseits haben wir

$$\mathcal{E}_n \sqcup \mathcal{E}'_m = \{1, ..., n\} \cup \{n+1, ..., n+m\} = \mathcal{E}_{n+m},$$

woraus folgt

$$A \sqcup B \sim \mathcal{E}_{n+m}$$

und die Endlichkeit von $A \sqcup B$.

Bemerkung. Es folgt aus dem Beweis von (*b*), dass für disjunkte endlichen Mengen A und B gilt

$$\operatorname{card}(A \sqcup B) = \operatorname{card} A + \operatorname{card} B. \tag{1.18}$$

Für allgemeinen endlichen Mengen A und B gilt

$$\operatorname{card}(A \cup B) = \operatorname{card}(A) + \operatorname{card}(B) - \operatorname{card}(A \cap B) .$$

(c) Wir zeigen per Induktion nach n, dass unter den Bedingungen (1.17),

$$A \times B \sim \mathcal{E}_{nm}. \tag{1.19}$$

Ist $n = 0$, so gilt $A = \emptyset$ und $A \times B = \emptyset$ woraus die Aussage folgt. Für Induktionsschritt mit $A \sim \mathcal{E}_{m+1}$, wählen wir ein Element $x \in A$ und betrachten die Menge $A' = A \setminus \{x\}$, so dass $A' \sim \mathcal{E}_n$. Nach Induktionsvoraussetzung haben wir

$$A' \times B \sim \mathcal{E}_{nm}.$$

Da die Abbildung $b \to (x, b)$ eine Bijektion von B nach $\{x\} \times B$ ist, so gilt

$$\{x\} \times B \sim B \sim \mathcal{E}_m.$$

Da $A = A' \sqcup \{x\}$ und somit

$$A \times B = (A' \times B) \sqcup (\{x\} \times B),$$

so erhalten wir nach (1.18)

$$A \times B \sim \mathcal{E}_{nm+m} = \mathcal{E}_{(n+1)m},$$

was zu beweisen war. ■

Bemerkung. Man schreibt die Relation (1.19) in der Form

$$\operatorname{card}(A \times B) = \operatorname{card}(A) \operatorname{card}(B). \tag{1.20}$$

1.5.3 Funktionen auf endlichen Mengen

Sei M eine nichtleere endliche Menge. Betrachten wir eine Funktion $a : M \to \mathbb{R}$ und bezeichnen $a_k = a(k)$. Die Summe

$$\sum_{k \in M} a_k \tag{1.21}$$

lässt sich per Induktion nach $n = \operatorname{card} M$ definieren, analog zur Summe $\sum_{k=1}^{n} a_k$ (siehe Aufgaben).

Erwähnen wir ohne Beweis die weiteren Eigenschaften der Summe. Sie alle können leicht per Induktion bewiesen werden.

1. Im Fall $M = \mathcal{E}_n = \{1, ..., n\}$ gilt

$$\sum_{k \in \mathcal{E}_n} a_k = \sum_{k=1}^{n} a_k.$$

Um diese zwei Begriffe zu unterscheiden, nennt man $\sum_{k=1}^{n} a_k$ geordnete Summe und $\sum_{k \in \mathcal{E}_n} a_k$ – ungeordnete Summe.

2. Ist $M = M_1 \sqcup M_2$, so gilt

$$\sum_{k \in M_1 \sqcup M_2} a_k = \sum_{k \in M_1} a_k + \sum_{k \in M_2} a_k.$$

3. Seien a, b zwei Funktionen auf M. Dann gilt

$$\sum_{k \in M} (a_k + b_k) = \sum_{k \in M} a_k + \sum_{k \in M} b_k.$$

Auch für jedes $c \in \mathbb{R}$

$$\sum_{k \in M} c a_k = c \sum_{k \in M} a_k.$$

4. Ist $a_k \le b_k$ für alle $k \in M$, so gilt

$$\sum_{k \in M} a_k \le \sum_{k \in M} b_k.$$

5. Ist $a_k = 1$ für alle $k \in M$, so gilt

$$\sum_{k \in M} a_k = \operatorname{card} M.$$

6. Seien M, N zwei endliche Mengen. Die folgende Identität gilt für alle Funktionen $a : M \to \mathbb{R}$ und $b : N \to \mathbb{R}$:

$$\left(\sum_{k \in M} a_k \right) \left(\sum_{l \in N} b_l \right) = \sum_{(k,l) \in M \times N} a_k b_l. \tag{1.22}$$

Darüber hinaus gilt für jede Funktion $c : M \times N \to \mathbb{R}$ die Identität

$$\sum_{(k,l) \in M \times N} c_{k,l} = \sum_{k \in M} \left(\sum_{l \in N} c_{k,l} \right).$$

1.6 Komplexe Zahlen

Definition. Die Menge $\mathbb{C}$ von *komplexen Zahlen* its die Menge $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ von allen Paaren (x, y) von reellen Zahlen x, y, mit den folgenden Operationen $+$ und $\cdot$:

- $(x, y) + (x', y') = (x + x', y + y')$
- $(x, y) \cdot (x', y') = (xx' - yy', xy' + yx')$.

Die komplexe Zahl $(0, 0)$ has die Eigenschaft von Nullelement:

$$(x, y) + (0, 0) = (0, 0) + (x, y) = (x, y)$$

und wird mit auch 0 bezeichnet. Die komplexe Zahl $(1, 0)$ has die Eigenschaft von Einheitselement:

$$(x, y) \cdot (1, 0) = (1, 0) \cdot (x, y) = (x, y)$$

und wird auch mit 1 bezeichnet.

Für die komplexen Zahlen der Form $(x,0)$ gelten die Regeln

$$\begin{aligned}(x,0)+(x',0) &= (x+x',0)\\ (x,0)\cdot(x',0) &= (xx',0)\,.\end{aligned}$$

Die Abbildung

$$\begin{aligned}\mathbb{R} &\to \mathbb{C}\\ x &\mapsto (x,0)\end{aligned}$$

ist eine injektive Abbildung, die die Operationen $+$ und $\cdot$ erhält. Deshalb identifiziert man die Menge $\mathbb{R}$ als eine Teilmenge von $\mathbb{C}$, indem man x mit $(x,0)$ identifiziert.

Die komplexe Zahl $(0,1)$ ist besonders wichtig und wird mit i bezeichnet (manchmal wird $(0,1)$ auch mit j bezeichnet). Die Zahl $i=(0,1)$ heißt die *imaginäre Einheit*. Nach Definition haben wir

$$i^2=(0,1)\cdot(0,1)=(-1,0)=-1$$

so that $i^2=-1$.

Bemerken wir auch, dass für alle $a\in\mathbb{R}$ und $(x,y)\in\mathbb{C}$,

$$a\cdot(x,y)=(a,0)\cdot(x,y)=(ax,ay)\,,$$

and analog

$$(x,y)\cdot a=(ax,ay)\,,$$

woraus folgt

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

Somit kann jede komplexe Zahl in der Form $x+yi$ (oder $x+iy$) dargestellt werden. Diese Darstellung heißt die *kartesische* oder *algebraische* Form der komplexen Zahl. Sei $z=x+iy$ eine komplexe Zahl. Dabei wird x als Realteil und y als Imaginärteil von z bezeichnet. Man schreibt:

$$x=\operatorname{Re} z \quad\text{und}\quad y=\operatorname{Im} z.$$

Folglich haben wir

$$z=\operatorname{Re} z+i\operatorname{Im} z.$$

Offensichtlich gilt $z\in\mathbb{R}$ genau dann, wenn $\operatorname{Im} z=0$.

In der kartesischen Form sehen die Rechenregeln von komplexen Zahlen wie folgt aus:

$$(x+iy)+(x'+iy')=(x+x')+i\,(y+y')$$

und

$$(x+iy)\,(x'+iy')=(xx'-yy')+i\,(xy'+x'y)\,.$$

Man addiert und multipliziert die Ausdrücke $x+iy$ und $x'+iy'$ genau so, wie die reellen Zahlen, aber mit der zusätzlichen Regel $i^2=-1$.

Für komplexe Zahlen gibt es eine neue Operation – die *Konjugation*.

Definition. Für jede komplexe Zahl $z=x+iy\in\mathbb{C}$ definieren wir die komplexe *Konjugierte* $\bar z$ durch

$$\bar z=x-iy=\operatorname{Re} z-i\operatorname{Im} z.$$

Offensichtlich gilt $\overline{\overline{z}} = z$ für alle $z \in \mathbb{C}$ und auch $\bar{z} = z$ für alle $z \in \mathbb{R}$.

Erinnern wir uns den Begriff des *Betrages* von reellen Zahlen: für $x \in \mathbb{R}$ setzen wir

$$|x| := \begin{cases} x, \ x \geq 0, \\ -x, \ x < 0. \end{cases}$$

Es folgt, dass Betrag die folgenden Eigenschaften erfüllt:

1. $|x + y| \leq |x| + |y|$ (die Dreiecksungleichung)

2. $|xy| = |x| \, |y|$.

Auch brauchen wir den Begriff der Quadratwurzel.

Behauptung. *Für jede nichtnegative Zahl $a \geq 0$ existiert genau eine nichtnegative Zahl $x \in \mathbb{R}$ mit $x^2 = a$.*

Beweis. Die Existenz wurde in Aufgaben gegeben. Für die Eindeutigkeit nehmen wir an, dass $x^2 = y^2$ für $x, y \geq 0$ und beweisen, dass $x = y$. Im Fall $x \neq y$ gilt entweder $x < y$ oder $x > y$. Im ersten Fall erhalten wir $x^2 < y^2$ und im zweiten Fall $x^2 > y^2$, woraus die Aussage folgt. ■

Definition. Die eindeutige Zahl x mit $x^2 = a$ heißt die *Quadratwurzel* aus a und wird mit $\sqrt{a}$ bezeichnet.

Erwähnen wir die folgenden Eigenschaften von Quadratwurzel.

1. $0 \leq a \leq b$ ergibt $\sqrt{a} \leq \sqrt{b}$, was oberhalb schon beweisen wurde.

2. Für jedes $a \in \mathbb{R}$ gilt $\sqrt{a^2} = |a|$. Ist $a \geq 0$, so gilt es nach Definition. Für $a < 0$ haben wir $\sqrt{a^2} = \sqrt{(-a)^2} = -a = |a|$.

Definition. Für komplexe Zahl $z = x + iy$ definieren wir den Betrag $|z|$ durch

$$|z| = \sqrt{x^2 + y^2} = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}. \tag{1.23}$$

Insbesondere gilt $|z| \geq 0$, und $|z| = 0$ genau dann, wenn $z = 0$.

Für alle $z \in \mathbb{C}$ gelten $|\operatorname{Re} z| \leq |z|$ und $|\operatorname{Im} z| \leq |z|$, wie es aus (1.23) folgt. Ist z reell, d.h. $z = x$, so gilt

$$|z| = \sqrt{x^2} = |x| \, .$$

Wie sehen, dass die Begriffe von Beträgen für reelle Zahlen und für komplexe Zahlen übereinstimmen.

Im nächsten Satz In the next statement versammeln wir einige Eigenschaften der Operationen mit komplexen Zahlen.

Satz 1.22 *Seien z_1, z_2, z_3 beliebige komplexen Zahlen.*

(a) *Es gelten die folgenden Identitäten:*

$$z_1 + z_2 = z_2 + z_1, \quad (z_1 + z_2) + z_3 = z_1 + (z_2 + z_3) \tag{1.24}$$

$$z_1 z_2 = z_2 z_1, \quad (z_1 z_2)\, z_3 = z_1 (z_2 z_3) \tag{1.25}$$

$$(z_1 + z_2)\, z_3 = z_1 z_3 + z_2 z_3 \tag{1.26}$$

(b) *Es gelten auch*

$$z + \bar{z} = 2 \operatorname{Re} z, \quad z\bar{z} = |z|^2 \tag{1.27}$$

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2 \tag{1.28}$$

$$|z_1 z_2| = |z_1|\,|z_2| \quad \textit{und} \quad |z_1 + z_2| \le |z_1| + |z_2|\,. \tag{1.29}$$

(c) *Ist $z_2 \neq 0$, so existiert eine eindeutige Zahl $w \in \mathbb{C}$ mit $z_2 w = z_1$. Die Zahl w wird mit $\frac{z_1}{z_2}$ bezeichnet. Es gelten auch die Identitäten:*

$$\frac{z_1}{z_2} = |z_2|^{-2} z_1 \bar{z}_2 \quad \textit{und} \quad \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}. \tag{1.30}$$

Folglich ist $\mathbb{C}$ ein Körper mit Nullelement $0 = 0 + i0$ und Einheitselement $1 = 1 + i0$.

Die Ungleichung $|z_1 + z_2| \le |z_1| + |z_2|$ (und ihre Folgerung $|z_1 - z_2| \le |z_1| + |z_2|$) heißt die *Dreiecksungleichung*.

Beweis. (a) Die Kommutativgesetze für Addition und Multiplikation, das Assoziativgesetz für Addition, und das Distributivgesetz (die Identitäten (1.24), (1.26) und die erste Identität in (1.25)) sind direkte Folgen der Definition. Beweisen wir das Assoziativgesetz für Multiplikation (die zweite Identität in (1.25)). Bezeichnen wir $z_k = x_k + iy_k$ und erhalten

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i (x_1 y_2 + y_1 x_2)$$

und

$$\begin{aligned} (z_1 z_2)\, z_3 &= ((x_1 x_2 - y_1 y_2) + i (x_1 y_2 + y_1 x_2))\,(x_3 + iy_3) \\ &= (x_1 x_2 - y_1 y_2)\, x_3 - y_3 (x_1 y_2 + y_1 x_2) \\ &\quad + i (x_1 x_2 - y_1 y_2)\, y_3 + (x_1 y_2 + y_1 x_2)\, x_3 \\ &= x_1 x_2 x_3 - y_1 y_2 x_3 - x_1 y_2 y_3 - y_1 x_2 y_3 \\ &\quad + i (x_1 x_2 y_3 - y_1 y_2 y_3 + x_1 y_2 x_3 + y_1 x_2 x_3)\,. \end{aligned} \tag{1.31}$$

Andererseits gilt

$$z_1 (z_2 z_3) = (z_2 z_3)\, z_1,$$

und die ähnliche Entwicklung von $(z_2 z_3)\, z_1$ erhält man aus (1.31) bei der folgenden Permutation von den Indizen

$$\begin{array}{ccc} 1 & 2 & 3 \\ \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1 \end{array}$$

Der Realteil von $(z_1 z_2)\, z_3$ ändert sich bei dieser Permutation wie folgt:

$$\begin{array}{rccccccc} \operatorname{Re}((z_1 z_2)\, z_3) = & x_1 x_2 x_3 & - & y_1 y_2 x_3 & - & x_1 y_2 y_3 & - & y_1 x_2 y_3 \\ & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ \operatorname{Re}((z_2 z_3)\, z_1) = & x_2 x_3 x_1 & - & y_2 y_3 x_1 & - & x_2 y_3 y_1 & - & y_2 x_3 y_1 \end{array}$$

und man sieht in den beiden Zeilen die gleichen Glieder, so dass

$$\operatorname{Re}((z_1 z_2)\, z_3) = \operatorname{Re}((z_2 z_3)\, z_1).$$

Das Gleiche gilt für den Imaginärteil, woraus folgt

$$(z_1 z_2)\, z_3 = (z_2 z_3)\, z_1$$

und somit auch die zweite Identität in (1.25).

(b) Beweisen wir jetzt (1.27). Setzen wir $z = x + iy$. Dann $\bar{z} = x - iy$ und

$$z + \bar{z} = (x + iy) + (x - iy) = 2x = 2 \operatorname{Re} z$$

und

$$z\bar{z} = (x + iy)(x - iy) = \left(x^2 + y^2\right) + i\,(xy - xy) = x^2 + y^2 = |z|^2$$

Die erste Identität in (1.28) ist offensichtlich, die zweite Identität wird wie folgt bewiesen: aus

$$\overline{z_1 z_2} = \overline{(x_1 x_2 - y_1 y_2) + i\,(x_1 y_2 + y_1 x_2)} = (x_1 x_2 - y_1 y_2) - i\,(x_1 y_2 + y_1 x_2)$$

und

$$\bar{z}_1 \bar{z}_2 = (x_1 - i y_1)(x_2 - i y_2) = (x_1 x_2 - y_1 y_2) - i\,(x_1 y_2 + y_1 x_2)$$

erhalten wir $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

The erste Identität in (1.29) ist wie folgt bewiesen:

$$\begin{aligned} |z_1 z_2|^2 &= (x_1 x_2 - y_1 y_2)^2 + (x_1 y_2 + y_1 x_2)^2 \\ &= x_1^2 x_2^2 - 2 x_1 x_2 y_1 y_2 + y_1^2 y_2^2 + x_1^2 y_2^2 + 2 x_1 x_2 y_1 y_2 + y_1^2 x_2^2 \\ &= x_1^2 x_2^2 + y_1^2 y_2^2 + x_1^2 y_2^2 + y_1^2 x_2^2 \\ &= \left(x_1^2 + y_1^2\right)\left(x_2^2 + y_2^2\right) = |z_1|^2 |z_2|^2 . \end{aligned}$$

Alternativ kann man diese Identität wie folgt beweisen:

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

Um die Ungleichung in (1.29) zu beweisen, bemerken wir zunächst, dass

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

wo wir benutzt haben, dass

$$z_1 \bar{z}_2 + z_2 \bar{z}_1 = z_1 \bar{z}_2 + \overline{z_1 \bar{z}_2} = 2 \operatorname{Re}(z_1 \bar{z}_2).$$

Da $|\operatorname{Re} z| \leq |z|$, so erhalten wir

$$|z_1 + z_2|^2 \leq |z_1|^2 + 2\,|z_1 \bar{z}_2| + |z_2|^2 = |z_1|^2 + 2\,|z_1|\,|z_2| + |z_2|^2 = (|z_1| + |z_2|)^2 ,$$

woraus

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

folgt.

(c) Multiplizieren die Gleichung $z_2 w = z_1$ mit $\bar{z}_2$ ergibt

$$|z_2|^2 w = z_1 \bar{z}_2,$$

wobei wir $\bar{z}_2 z_2 = |z_2|^2$ benutzt haben. Da $|z_2| > 0$, können wir diese Gleichung weiter mit $|z_2|^{-2}$ multiplizieren und somit erhalten

$$w = |z_2|^{-2} z_1 \bar{z}_2, \tag{1.32}$$

was die Zahl w eindeutig bestimmt. Umgekehrt, definieren wir w mit (1.32) und erhalten

$$z_2 w = |z_2|^{-2} z_1 \bar{z}_2 z_2 = |z_2|^{-2} z_1 |z_2|^2 = z_1,$$

was sowohl die Existenz von w als auch die erste Identität in (1.30) beweist. Für den Beträge der beiden Seiten von (1.32) erhalten wir

$$\left|\frac{z_1}{z_2}\right| = |w| = |z_2|^{-2} |z_1| |\bar{z}_2| = |z_2|^{-1} |z_1| = \frac{|z_1|}{|z_2|},$$

was die zweite Identität in (1.30) beweist. ■

1.7 Kardinalzahlen

Wir haben schon die Gleichmächtigkeit von zwei Mengen definiert: die Mengen X und Y heißen gleichmächtig, falls es eine bijektive Abbildung $f : X \to Y$ gibt. In diesem Fall schreibt man $X \sim Y$. Diese Relation ist eine Äquivalenzrelation zwischen Mengen. Somit bestimmt jede Menge X eine Äquivalenzklasse, die wird mit $|X|$ bezeichnet. Die Elementen von der Klasse $|X|$ sind alle Mengen, die zu X gleichmächtig sind. Offensichtlich gilt $|X| = |Y|$ genau dann, wenn $X \sim Y$.

Definition. Die Äquivalenzklasse $|X|$ heißt die *Kardinalzahl* (oder *Kardinalität*, *Mächtigkeit*) von X.

Jede endliche nicht-leere Menge X ist äquivalent zu einer der Mengen $\mathcal{E}_n = \{1, ..., n\}$ mit $n \in \mathbb{N}$, d.h. $|X| = |\mathcal{E}_n|$. In diesem Fall sagen wir, dass die Anzahl von Elementen von X ist n und schreiben $\operatorname{card} X = n$. Nach Satz 1.19 ist die Kardinalzahl $|X|$ einer endlichen Menge X eindeutig bestimmt. Häufig identifiziert man $|X|$ mit n und schreibt $|X| = n$. Deshalb sind die Kardinalzahlen von endlichen Mengen äquivalent zu natürlichen Zahlen. In der Tat kann man die natürlichen Zahlen als Kardinalzahlen der endlichen Mengen definieren. In dieser Theorie wird die Endlichkeit der Mengen mit Hilfe von Schubfachprinzip definiert: eine Menge M heißt endlich, falls sie zu keiner ihren echten Teilmenge gleichmächtig ist.

Die Kardinalzahlen von unendlichen Mengen können als Verallgemeinerung von natürlichen Zahlen betrachtet werden. In diesem Abschnitt definieren wir die Ungleichungen zwischen Kardinalzahlen.

Definition. Für je zwei Mengen X, Y schreiben wie $X \leq Y$ falls X zu einer Teilmenge von Y gleichmächtig ist.

Insbesondere gilt $X \leq Y$ fall X eine Teilmenge von Y ist.

Behauptung. *Die Relation $X \leq Y$ gilt genau dann, wenn es eine injektive Abbildung $f : X \to Y$ gibt.*

Beweis. Ist X zu einer Teilmenge $Y' \subset Y$ gleichmächtig, so existiert eine bijektive Abbildung $g : X \to Y'$. Definieren wir $f : X \to Y$ mit $f(x) = g(x)$ und erhalten somit eine injektive Abbildung. Umgekehrt, existiert eine injektive Abbildung $f : X \to Y$, so bezeichnen wir $Y' = f(X)$ und erhalten damit eine bijektive Abbildung $f : X \to Y'$ wobei Y' eine Teilmenge von Y ist. ■

Die Ungleichung $X \leq Y$ zwischen Mengen kann zu Kardinalzahlen erweitert werden wie folgt.

Definition. Man schreibt $|X| \leq |Y|$ falls $X \leq Y$.

Behauptung. *Die Relation "$\leq$" zwischen den Kardinalzahlen ist wohldefiniert.*

Beweis. Sei a, b zwei Kardinalzahlen, d.h. zwei Äquivalenzklassen von Mengen. Nach Definition bedeutet die Ungleichung $a \leq b$ folgendes: wir wählen Mengen $X \in a$ und $Y \in b$ und sagen $a \leq b$ falls $X \leq Y$. Offensichtlich muss man zeigen, dass die Bedingung $X \leq Y$ unabhängig von der Wahl von X und Y ist. Ist sie unabhängig, so sagt man, dass die Ungleichheit "$\leq$" zwischen Äquivalenzklassen wohldefiniert ist.

Seien X' und Y' andere Mengen aus der Klassen a bzw b. Wir müssen zeigen, dass $X \leq Y$ gilt genau dann, wenn $X' \leq Y'$, vorausgesetzt $X \sim X'$ und $Y \sim Y'$. Ist $X \leq Y$, so existieren die Abbildungen wie folgt:

$$\begin{array}{ccc} X & \xrightarrow{h} & Y \\ \downarrow^{f} & & \downarrow^{g} \\ X' & & Y' \end{array}$$

wobei f und g bijektiv und h injektiv sind. Dann ist die Verkettung $g \circ h \circ f^{-1}$ eine injektive Abbildung von X' nach Y', woraus folgt $X' \leq Y'$, was zu beweisen war. ■

Die Ungleichheit zwischen Kardinalzahlen erfüllt die folgenden Eigenschaften:

1. $|\mathcal{E}_n| \leq |\mathcal{E}_m| \Leftrightarrow n \leq m$, was aus dem Schubfachprinzip folgt.

2. $|X| \leq |X|$ (trivial).

3. $|X| \leq |Y| \;\wedge\; |Y| \leq |Z| \;\Rightarrow\; |X| \leq |Z|$ (siehe Aufgaben).

4. $|X| \leq |Y| \;\wedge\; |Y| \leq |X| \;\Rightarrow\; |X| = |Y|$

5. für je zwei Mengen X, Y gilt $|X| \leq |Y| \;\vee\; |Y| \leq |X|$.

Die Eigenschaften 4 und 5 haben die komplizierten Beweise, die wir nicht angeben (und nicht brauchen).

1.7.1 Abzählbare Mengen

Definition. Eine Menge X heißt *abzählbar* falls $X \sim \mathbb{N}$ (d.h. $|X| = |\mathbb{N}|$).

Die Kardinalzahl $|\mathbb{N}|$ wird auch mit $\aleph_0$ bezeichnet und *Aleph-null* genannt (wobei $\aleph$ der erste Buchstabe *Aleph* des hebräischen Alphabetes ist).

Ist X abzählbar, so existiert ein Bijektion $f : \mathbb{N} \to X$. Bezeichnen wir $f(n) = x_n$. Dann lässt sich X als eine *unendliche Folge* darstellen:

$$X = \{x_n : n \in \mathbb{N}\} = \{x_n\}_{n\in\mathbb{N}} = \{x_n\}_{n=1}^{\infty} = \{x_1, x_2, ...\}.$$

Man sagt, dass die Menge X mit natürlichen Zahlen abgezählt ist.

Beispiel. (*Hilberts Hotel*) Stellen wir ein Hotel mit einer abzählbaren Menge von Zimmern vor, die mit allen natürlichen Zahlen durchnummeriert sind. Seien alle Zimmer schon belegt, aber es kommt noch ein Gast an. Man kann doch Platz für den neuen Gast befreien, indem man den Gast aus Zimmer Nr 1 nach Zimmer Nr 2 versetzt, den Gast aus Zimmer Nr 2 nach Zimmer Nr 3, usw. Damit wird das Zimmer Nr 1 frei für den neuen Gast.

Außerdem kann man Platz für abzählbar vielen neuen Gäste machen. Seien die neuen Gäste auch durchnummeriert mit natürlichen Zahlen. Der alte Gast aus dem Zimmer Nr n wir nach Zimmer Nr $2n$ versetzt, und somit werden alle ungeraden Zimmer frei. Der neue Gast mit Nummer m wird dann ins Zimmer Nr $2m - 1$ untergebracht.

Es ist noch interessanter, dass man den Platz für die neuen Gäste aus abzählbar vielen Gruppen je mit abzählbar vielen Gästen befreien kann, wir man unterhalb sieht.

Zunächst beweisen Sie noch eine Erweiterung des Induktionsprinzips.

Behauptung. (Induktionsprinzip) *Sei $A(n)$ eine von $n \in \mathbb{N}$ abhängige Aussage, die die folgenden Bedingungen erfüllt:*

- *$A(1)$ ist wahr;*
- *Für jedes $n \in \mathbb{N}$ gilt*

$$A(k) \text{ ist wahr für alle } k < n \;\Rightarrow A(n) \tag{1.33}$$

Dann ist $A(n)$ wahr für alle $n \in \mathbb{N}$.

Beweis. Bezeichnen wir mit $B(n)$ die folgende Aussage

$$B(n) = (A(k) \text{ ist wahr für alle } k < n).$$

Beweisen wir per Induktion nach n, dass $B(n)$ wahr für alle $n \geq 2$ ist, woraus folgt, dass $A(n)$ wahr für alle $n \in \mathbb{N}$ ist.

Induktionsanfang. $B(2)$ ist wahr, weil $B(2)$ äquivalent zu $A(1)$ ist.

Induktionsschritt. Sei $B(n)$ wahr, beweisen wir, dass $B(n+1)$ wahr ist. Bemerken wir, dass

$$B(n+1) = (A(k) \text{ ist wahr für alle } k < n+1) = B(n) \wedge A(n).$$

Andererseits ist die Voraussetzung (1.33) äquivalent zu

$$B(n) \Rightarrow A(n).$$

Somit erhalten wir $B(n) \Rightarrow B(n+1)$, was zu beweisen war. ■

Der nächste Satz enthält wichtige Eigenschaften von abzählbaren Mengen.

Satz 1.23 (a) *Jede Teilmenge einer abzählbaren Menge ist entweder endlich oder abzählbar.*

(b) *Kartesisches Produkt zweier abzählbaren Mengen ist auch abzählbar.*

(c) *Sei $\{X_n\}_{n\in\mathbb{N}}$ eine Folge von abzählbaren Mengen. Dann ist die Vereinigung $X = \bigcup_{n\in\mathbb{N}} X_n$ auch abzählbar.*

Bemerkung. Der Teil (a) impliziert, dass es keine unendliche Menge X gibt, die kleiner als $\mathbb{N}$ ist: gilt für eine unendliche Menge $|X| \leq |\mathbb{N}|$, so gilt $|X| = |\mathbb{N}|$. In der Tat bedeutet $|X| \leq |\mathbb{N}|$, dass X zu einer Teilmenge von $\mathbb{N}$ gleichmächtig ist. Da diese Teilmenge unendlich sein soll, erhalten wir nach (a), dass sie abzählbar ist. Man kann auch zeigen, dass für jede endliche Menge Y gilt $|Y| \geq |\mathbb{N}|$ (siehe Aufgaben).

Der Teil (b) ist äquivalent zu $|\mathbb{N}\times\mathbb{N}| = |\mathbb{N}|$. Zum Vergleich erinnern wir uns daran, dass $|\mathcal{E}_n \times \mathcal{E}_n| = |\mathcal{E}_{n^2}|$, insbesondere $|\mathcal{E}_n| < |\mathcal{E}_n \times \mathcal{E}_n|$ für alle $n > 1$.

Der Teil (c) ergibt, dass man im Hilberts Hotel auch alle Gäste aus abzählbar vielen Gruppen je mit abzählbar vielen Gästen unterbringen kann.

Beweis. (a) Es reicht zu zeigen, dass jede Teilmenge X von $\mathbb{N}$ entweder endlich oder abzählbar ist. Sei X unendlich. Dann erstellen wir eine Bijektion $f : \mathbb{N} \to X$ woraus die Abzählbarkeit von X folgen wird. Dafür definieren wir $f(n)$ per Induktion nach n.

Induktionsanfang. Nach Satz 1.16 existiert das Minimum von X. Setzen wir $f(1) = \min X$.

Induktionsschritt. Seien $f(k)$ für alle $k < n$ schon definiert. Die Menge

$$X_n = \{f(k) : k < n\} = \{f(k) : k \in \mathcal{E}_{n-1}\}$$

ist eine endliche Teilmenge von X. Deshalb ist die Differenz $X \setminus X_n$ nicht leer, und wir können setzen

$$f(n) = \min(X \setminus X_n). \tag{1.34}$$

Nach dem obigen Induktionsprinzip erhalten wir eine Abbildung $f : \mathbb{N} \to X$. Diese Abbildung ist injektiv, da $f(n) \notin X_n$ und somit $f(n) \neq f(k)$ für alle $k < n$.

Beweisen wir, dass f surjektiv ist, d.h. $f(\mathbb{N}) = X$. Die Menge $f(\mathbb{N})$ ist eine unendliche Teilmenge von $\mathbb{N}$ und deshalb hat keine obere Schranke. Daraus folgt, dass für jedes $x \in X$ existiert ein $n \in \mathbb{N}$ mit $x < f(n)$. Es folgt aus (1.34), dass x nicht in $X \setminus X_n$ liegen kann, und somit $x \in X_n$. Nach Definition von X_n existiert ein $k < n$ mit $f(k) = x$, d.h. $x \in f(\mathbb{N})$.

(b) Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $X = Y = \mathbb{N}$. Dann beweisen wir, dass die Menge

$$\mathbb{N}^2 = \mathbb{N}\times\mathbb{N} = \{(n,m) : n,m \in \mathbb{N}\}$$

abzählbar ist. Für jedes $k \in \mathbb{N}$ betrachten wir die Menge

$$D_k = \{(n,m) : n,m \in \mathbb{N},\ n+m = k\}.$$

Jede Menge D_k ist endlich, da sie $k-1$ Elementen enthält; insbesondere $D_1 = \emptyset$. Auch ist $\mathbb{N}^2$ die disjunkte Vereinigung von allen Mengen D_k, $k \in \mathbb{N}$. Auf dem Diagramm unterhalb sind die Elementen von $\mathbb{N}^2$ in einer Tabelle angeordnet, und die Mengen $D_k, k \geq 2$, sind Diagonalen in dieser Tabelle. Man kann die Menge $\mathbb{N}^2$ abzählen indem man die Mengen D_k nacheinander abzählt.

$$\begin{array}{ccccccccccc}
(1,1)_1 & & (1,2)_3 & & (1,3)_6 & & (1,4)_{10} & & (1,5)_{15} & & ... \\
 & \nearrow & & \nearrow & & \nearrow & & \nearrow & & & \\
(2,1)_2 & & (2,2)_5 & & (2,3)_9 & & (2,4)_{14} & & ... & & \\
 & \nearrow & & \nearrow & & \nearrow & & & & & \\
(3,1)_4 & & (3,2)_8 & & (3,3)_{13} & & ... & & & & \\
 & \nearrow & & \nearrow & & & & & & & \\
(4,1)_7 & & (4,2)_{12} & & ... & & & & & & \\
 & \nearrow & & & & & & & & & \\
(5,1)_{11} & & ... & & & & & & & & \\
... & & & & & & & & & &
\end{array}$$

Dieses Argument heißt die *Diagonal-Abzählung*. Die genaue Definition der Bijektion $f : \mathbb{N}^2 \to \mathbb{N}$ ist wie folgt. Zunächst definieren wir die Folge $\{a_k\}_{k=1}^{\infty}$ wie folgt: a_k ist die Anzahl von Elementen in der Menge $\bigcup_{i=1}^{k} D_i$ (d.h. die ersten k Diagonalen). Da $\operatorname{card} D_k = k-1$, so erhalten wir

$$a_k = \operatorname{card} \bigcup_{i=1}^{k} D_i = \sum_{i=1}^{k} (i-1) = \sum_{j=1}^{k-1} j = \frac{(k-1)\,k}{2}.$$

Für jedes $(n,m) \in \mathbb{N}^2$ gibt es genau ein $k \in \mathbb{N}$ mit $(n,m) \in D_k$, nämlich $k = n+m$. Dann setzen wir

$$f(n,m) := a_{k-1} + m = a_{n+m-1} + m.$$

Offensichtlich ist $f|_{D_k}$ injektiv. Für $(n,m) \in D_k$ nimmt m die Werte $\{1, ..., k-1\}$ an, woraus folgt

$$f(D_k) = \{a_{k-1}+1, ..., a_{k-1}+(k-1)\} = \{a_{k-1}+1, ..., a_k\}$$

wobei wir benutzt haben, dass

$$a_k = a_{k-1} + (k-1)\,.$$

Wir haben dann

$$f(D_k) = \{x \in \mathbb{N} : a_{k-1} < x \leq a_k\} = (a_{k-1}, a_k] \cap \mathbb{N}.$$

Da

$$\bigsqcup_{k \geq 2} (a_{k-1}, a_k] = (0, +\infty)$$

(siehe Aufgaben), so erhalten wir, dass

$$\bigsqcup_{k\geq 2} f(D_k) = \mathbb{N}.$$

Somit ist f eine Bijektion von $\mathbb{N}^2$ nach $\mathbb{N}$, was zu beweisen war.

(c) Jede Menge X_n kann mit natürlichen Zahlen abgezählt werden. Bezeichnen wir mit x_{nm} das m-te Element von X_n, wobei $n, m \in \mathbb{N}$. Definieren wir eine Abbildung $f : \mathbb{N}^2 \to X$ durch

$$f(n, m) = x_{nm}$$

Die Abbildung $f : \mathbb{N}^2 \to X$ ist offensichtlich surjektiv, was impliziert $|X| \leq |\mathbb{N}^2|$ (siehe Aufgaben). Nach (b) haben wir $|\mathbb{N}^2| = |\mathbb{N}|$ und somit $|X| \leq |\mathbb{N}|$. Da X unendlich ist, so erhalten wir nach (a), dass $|X| = |\mathbb{N}|$. ■

Die Vereinigung $X = \bigcup_{n\in\mathbb{N}} X_n$ ist auch abzählbar wenn alle Mengen X_n *höchstens abzählbar* sind (d.h. abzählbar oder endlich) und mindestens eine Menge von X_n abzählbar ist, da man die endlichen Mengen X_n immer in die abzählbaren Mengen umwandeln kann; dann ist X abzählbar als eine unendliche Teilmenge von einer abzählbaren Menge.

Auch die endliche Vereinigung $X = \bigcup_{n=1}^{N} X_n$ ist abzählbar, vorausgesetzt, dass alle Mengen X_n höchstens abzählbar sind und mindestens eine Menge von X_n abzählbar ist, da man die endliche Folge $\{X_n\}_{n=1}^{N}$ immer in die unendliche Folge $\{X_n\}_{n=1}^{\infty}$ fortsetzen kann, indem man setzt $X_n = X_N$ für alle $n > N$.

Korollar 1.24 *Die Mengen $\mathbb{Z}$ und $\mathbb{Q}$ sind abzählbar, d.h.* $|\mathbb{Q}| = |\mathbb{Z}| = |\mathbb{N}|$.

Beweis. Da $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N}) \cup \{0\}$, so erhalten wir $|\mathbb{Z}| = |\mathbb{N}|$. Da

$$\begin{aligned}\mathbb{Q} &= \left\{\frac{n}{m} : n, m \in \mathbb{Z}, m \neq 0\right\} \\ &= \bigcup_{m\in\mathbb{Z}\setminus\{0\}} \left\{\frac{n}{m} : n \in \mathbb{Z}\right\},\end{aligned}$$

so ist $\mathbb{Q}$ eine abzählbare Vereinigung von abzählbaren Mengen. Somit gilt $|\mathbb{Q}| = |\mathbb{N}|$. ■

1.7.2 Überabzählbare Mengen

Wir schreiben $|X| < |Y|$ falls $|X| \leq |Y|$ aber $|X| \neq |Y|$.

Definition. Eine Menge X heißt *überabzählbar* falls $|\mathbb{N}| < |X|$.

Da $|\mathbb{N}| \leq |X|$ für jede unendliche Menge X gilt, so ist eine unendliche Menge X genau dann überabzählbar wenn X nicht abzählbar ist.

Satz 1.25 *Die Menge $\mathbb{R}$ ist überabzählbar, d.h.* $|\mathbb{N}| < |\mathbb{R}|$.

Die Kardinalzahl $|\mathbb{R}|$ heißt das *Kontinuum* und wird mit $\mathfrak{c}$ bezeichnet. Die Menge Da $|\mathbb{Q}| = |\mathbb{N}| < |\mathbb{R}|$, es folgt, dass die Menge $\mathbb{R} \setminus \mathbb{Q}$ von *irrationalen* Zahlen nicht leer ist. Ein anderer Beweis davon: $\sqrt{2}$ ist irrational (siehe Aufgaben).

Beweis. Wir müssen beweisen, dass $\mathbb{R}$ nicht abzählbar ist. Nehmen wir das Gegenteil an, dass $\mathbb{R}$ abzählbar ist, so dass $\mathbb{R} = \{x_n : n \in \mathbb{N}\}$. Konstruieren wir eine Folge $\{I_n\}_{n=1}^{\infty}$ von abgeschlossenen beschränkten Intervallen $I_n = [a_n, b_n]$ so dass

1. $a_n < b_n$
2. $I_{n+1} \subset I_n$
3. $x_n \notin I_n$.

Sei I_1 ein beliebiges Intervall, das x_1 nicht enthält, zum Beispiel $I_1 = [x_1 + 1, x_2 + 2]$. Ist $I_n = [a_n, b_n]$ schon definiert, so wählen wir I_{n+1} als ein Teilintervall von I_n, das x_{n+1} nicht enthält, zum Beispiel wie folgt: im Fall $x_{n+1} \notin I_n$ setzen wir $I_{n+1} = I_n$, und im Fall $x_{n+1} \in I_n$ wählen wir $I_{n+1} = \left[\frac{a_n+x_{n+1}}{2}, b_n\right]$ oder $I_{n+1} = [a_n, \frac{x_{n+1}+b_n}{2}]$.

Dann betrachten wir die Mengen

$$A = \{a_n : n \in \mathbb{N}\} \quad \text{und} \quad B = \{b_m : m \in \mathbb{N}\}.$$

Die Inklusion $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ ergibt, dass für alle $m \geq n$

$$[a_m, b_m] \subset [a_n, b_n].$$

Daraus folgt

$$a_n \leq a_m < b_m \leq b_n.$$

und somit

$$a_n \leq b_m \ \text{ für } n \leq m$$

und

$$a_m \leq b_n \ \text{ für } n \leq m.$$

Umtauschen von m und n ergibt

$$a_n \leq b_m \ \text{ für } n \geq m,$$

so dass

$$a_n \leq b_m \text{ für alle } n, m \in \mathbb{N}.$$

Nach dem Vollständigkeitsaxiom erhalten wir, dass es ein $c \in \mathbb{R}$ gibt, die A und B trennt, d.h. $a_n \leq c \leq b_m$ gilt für alle $n, m \in \mathbb{N}$. Insbesondere gilt $c \in I_n$ für alle n und somit $c \neq x_n$ für alle n, was im Widerspruch zur Annahme $\mathbb{R} = \{x_n : n \in \mathbb{N}\}$ ist. ■

Satz 1.26 *Seien X, Y zwei Mengen mit $|Y| \leq |\mathbb{N}|$.*
(a) Ist X unendlich, so gilt $|X \cup Y| = |X|$.
(b) Ist X überabzählbar, so gilt $|X \setminus Y| = |X|$.

Insbesondere gelten

$$|\mathbb{R} \cup Y| = \mathbb{R} \quad \text{und} \quad |\mathbb{R} \setminus Y| = |\mathbb{R}|.$$

Daraus folgt, dass $|\mathbb{R} \setminus \mathbb{Q}| = |\mathbb{R}|$, d.h. die Menge $\mathbb{R} \setminus \mathbb{Q}$ von irrationalen Zahlen die Kardinalzahl $|\mathbb{R}|$ hat.

Beweis. (a) Zunächst bemerken wir, dass

$$X \cup Y = X \sqcup (Y \setminus X)$$

und $|Y \setminus X| \le |Y| \le |\mathbb{N}|$. Umbenennen wir $Y \setminus X$ in Y. Dann sind X, Y disjunkt, gilt $|Y| \le |\mathbb{N}|$ und wir müssen beweisen, dass

$$|X \sqcup Y| = |X| .$$

Jede unendliche Menge erhält eine abzählbare Teilmenge (siehe Aufgaben). Sei X_0 eine abzählbare Teilmenge von X und sei $X_1 = X \setminus X_0$. Dann gilt

$$X = X_0 \sqcup X_1$$

und somit

$$X \sqcup Y = (X_0 \sqcup X_1) \sqcup Y = (X_0 \sqcup Y) \sqcup X_1. \tag{1.35}$$

Da $|X_0| = |\mathbb{N}|$ und $|Y| \le |\mathbb{N}|$, so erhalten wir nach Satz 1.23, dass

$$X_0 \sqcup Y \sim \mathbb{N} \sim X_0,$$

was zusammen mit (1.35) ergibt

$$X \sqcup Y \sim X_0 \sqcup X_1 = X,$$

was zu beweisen war.

(b) Da $X \setminus Y = X \setminus (X \cap Y)$ und $|X \cap Y| \le |Y| \le |\mathbb{N}|$, so können wir $X \cap Y$ in Y umbenennen und somit voraussetzen, dass $Y \subset X$. Dann haben wir

$$X = (X \setminus Y) \cup Y. \tag{1.36}$$

Die Menge $X \setminus Y$ ist unendlich, da sonst X abzählbar wäre als die Vereinigung von $Y \le \mathbb{N}$ und endlicher Menge $X \setminus Y$). Nach (a) erhalten wir

$$|(X \setminus Y) \cup Y| = |X \setminus Y| ,$$

woraus $|X| = |X \setminus Y|$ folgt. ■

Mit Hilfe von Satz 1.26 kann man beweisen, dass alle Intervalle (a, b) , $[a, b]$, $[a, b)$, $(a, b]$ mit $a < b$ die Kardinalzahl $|\mathbb{R}|$ haben (siehe Aufgaben).

Algebraische Zahlen.

Die folgenden Teilmengen von reellen Zahlen sind uns schon bekannt:

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

deren Kardinalzahlen die folgenden Ungleichungen erfüllen:

$$|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| < |\mathbb{R}| .$$

Betrachten wir noch eine Teilmenge von $\mathbb{R}$.

Definition. Eine Zahl $x \in \mathbb{R}$ heißt *algebraisch*, falls x eine Gleichung

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n = 0$$

erfüllt, wobei $n \in \mathbb{N}$ und alle Koeffizienten a_k rationale Zahlen sind, d.h. x eine Nullstelle eines Polynoms mit rationalen Koeffizienten ist.

Zum Beispiel, jedes $x \in \mathbb{Q}$ ist algebraisch, da es die Gleichung $x + a_1 = 0$ erfüllt mit $a_1 = -x \in \mathbb{Q}$ (und $n = 1$). Auch die Zahl $x = \sqrt{2}$ ist algebraisch da sie die Gleichung $x^2 - 2 = 0$ erfüllt (mit $n = 2$).

Bezeichnen wir mit $\mathbb{A}$ die Mengen von allen algebraischen Zahlen. Es möglich zu beweisen, dass die Summe, die Differenz, das Produkt und das Verhältnis zweier algebraischen Zahlen wieder algebraisch sind, und somit ist $\mathbb{A}$ ein Körper.

Offensichtlich gelten die Inklusionen $\mathbb{Q} \subset \mathbb{A} \subset \mathbb{R}$ und $\mathbb{Q} \neq \mathbb{A}$. Man kann beweisen, dass $|\mathbb{A}| = |\mathbb{N}|$, woraus folgt, dass es nicht-algebraische Zahlen gibt. Die reellen Zahlen, die nicht algebraisch sind, heißen *transzendent*. Ein Beispiel von transzendenter Zahl ist die Zahl π, die später rigoros definiert wird.

Kardinalzahl $\left|2^X\right|$.

Alle uns bisher getroffenen unendlichen Mengen haben die Kardinalzahlen $|\mathbb{N}|$ oder $|\mathbb{R}|$. Man kann fragen, ob es andere Kardinalzahlen gibt, zum Beispiel zwischen $|\mathbb{N}|$ und $|\mathbb{R}|$. Die Vermutung, dass es eine Kardinalzahl zwischen $|\mathbb{N}|$ und $|\mathbb{R}|$ gibt, heißt die *Kontinuumshypothese*. Die Antwort darauf ist erstaunlich: in den Rahmen von der üblichen axiomatischen Theorie von Mengen und Zahlen kann diese Vermutung weder bewiesen noch widerlegt werden. Deshalb kann die Existenz bzw nicht-Existenz der Kardinalzahl zwischen $|\mathbb{N}|$ und $|\mathbb{R}|$ als ein unabhängiges Axiom betrachtet werden.

Im Gegensatz dazu kann die Existenz von Kardinalzahlen $> |\mathbb{R}|$ leicht geantwortet werden.

Satz 1.27 *Für jede Menge X gilt $|X| < \left|2^X\right|$.*

Insbesondere gilt $\left|2^{\mathbb{R}}\right| > |\mathbb{R}|$. Man kann zeigen, dass $|\mathbb{R}| = \left|2^{\mathbb{N}}\right|$, was einen anderen Beweis von der Ungleichung $|\mathbb{R}| > |\mathbb{N}|$ liefert. Satz 1.27 ergibt auch, dass es keine maximale Kardinalzahl gibt.

Beweis. Die Ungleichung $|X| \leq |Y|$ gilt genau dann, wenn es eine injektive Abbildung $f : X \to Y$ gibt, und die Gleichheit $|X| = |Y|$ gilt, wenn es eine bijektive Abbildung $f : X \to Y$ gibt. Um die echte Ungleichung $|X| < |Y|$ zu beweisen, müssen wir folgendes zeigen:

- es existiert eine injektive Abbildung $f : X \to Y$,
- es existiert keine bijektive Abbildung $f : X \to Y$.

Eine injective Abbildung $f : X \to 2^X$ kann man wie folgt definieren: $f(x) = \{x\}$, wobei $\{x\}$ eine Teilmenge von X ist, die aus einem Element x besteht.

Sei $f : X \to 2^X$ eine beliebige Abbildung. Beweisen wir, dass f nicht surjektiv ist (und somit auch nicht bijektiv). Betrachten wir die Menge

$$S = \{x \in X : x \notin f(x)\}. \tag{1.37}$$

Beachten wir, dass $f(x)$ eine Teilmenge von X ist, und die Frage, ob x ein Element von $f(x)$ ist, ist völlig sinnvoll. Die Menge S ist eine Teilmenge von X und deshalb ein Element von 2^X. Zeigen wir, dass die Menge S kein Urbild hat. Nehmen wir das Gegenteil an, dass $f(y) = S$ für ein $y \in X$, und betrachten zwei Fälle.

1. Ist $y \in S$ so gilt nach (1.37) $y \notin f(y) = S$, was ist unmöglich.
2. Ist $y \notin S$ so gilt nach (1.37) $y \in f(y) = S$ – auch unmöglich.

Deshalb führt die Bedingung $f(y) = S$ zum Widerspruch, woraus folgt, dass f keine surjektive Abbildung ist. ■

Zum Abschluss von Kapitel 1. In diesem Kapitel haben wir axiomatisch die Zahlen (und Mengen) eingeführt und ihre grundlegenden Eigenschaften bewiesen. Man kann die Mathematik mit einem (sehr großen) Gebäude vergleichen, wo alle Begriffe und Aussagen wohnen. Dann befindet sich das Material von Kapitel 1 noch im Keller. Das Fundament besteht aus den Axiomen, grundlegenden Begriffen, wie Menge und Zahl, und den Regeln von Herleitung (Mathematische Logik). Die unmittelbaren Folgerungen von Axiomen, die in diesem Kapitel gewonnen wurden liegen im Kellergeschoss wie die Werkzeuge. Im Erdgeschoss befindet sich die Schulmathematik, die jetzt von uns teilweise begründet ist, und in den mehreren Obergeschossen befindet sich die Forschungsmathematik, insbesondere die Uni-Mathematik, die noch gelernt werden muss.

Chapter 2

Folgen und ihre Grenzwerte

2.1 Unendliche Folgen

Eine (*unendliche*) *Folge* von reellen Zahlen ist eine Abbildung $x : \mathbb{N} \to \mathbb{R}$. Der Wert $x(n)$ für $n \in \mathbb{N}$ wird auch mit x_n bezeichnet. Die ganze Folge x wird auch mit $\{x_n\}_{n\in\mathbb{N}}$ oder $\{x_n\}_{n=1}^{\infty}$ oder sogar $\{x_n\}$ bezeichnet. Die Werte x_n werden auch als *Glieder* (oder Folgenglieder) der Folge x genannt.

2.1.1 Der Begriff des Limes

Sei $\{x_n\}_{n=1}^{\infty}$ eine Folge von reellen Zahlen. Wir interessieren uns für das Verhalten der Folgenglieder x_n für großen Werten von n. Verschiedene Folgen können die verschiedenen Verhalten zeigen. Zum Beispiel, in der Folge $\{(-1)^n\}$ wechselt sich 1 mit -1 ab. Hingegen werden die Glieder in der Folge $\left\{\frac{1}{n}\right\}$ immer kleiner und näheren sich der Null für großen Werten des Index n an. Es ist natürlich zu sagen, dass die Folge $\left\{\frac{1}{n}\right\}$ den *Grenzwert* 0 besitzt, während $\{(-1)^n\}$ keinen Grenzwert hat.

Definition. Sei $\{x_n\}_{n=1}^{\infty}$ eine Folge von reellen Zahlen. Die Zahl $a \in \mathbb{R}$ heißt der *Grenzwert* der Folge $\{a_n\}$ genau dann, wenn:

$$\text{für jedes } \varepsilon > 0 \ \text{ existiert } N \in \mathbb{N} \ \text{ so dass für alle } n \geq N \text{ gilt } \ |x_n - a| < \varepsilon.$$

Oder, mit Hilfe von den Quantoren, schreibt man:

$$\boxed{\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N \qquad |x_n - a| < \varepsilon.} \tag{2.1}$$

Besitzt die Folge $\{x_n\}$ einen Grenzwert, so heißt die Folge *konvergent*; sonst heißt die Folge *divergent*. Man sagt auch: die Folge *konvergiert* bzw *divergiert*.

Ist a der Grenzwert von $\{x_n\}$, so benutzt man die folgende Schreibweise:

$$x_n \to a \quad (x_n \text{ konvergiert gegen } a)$$

$$x_n \to a \ \text{ für } \ n \to \infty \qquad (a_n \text{ konvergiert gegen } a \text{ für } n \text{ gegen unendlich}).$$

Um den Grenzwert der Folge $\{x_n\}$ zu bezeichnen, benutzt man die Notation $\lim x_n$, die heißt der *Limes* von x_n. Ist a der Grenzwert von $\{x_n\}$, so schreibt man auch

$$a = \lim x_n$$

oder

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

Die Bedingung (2.1) lässt sich wie folgt umformulieren. Wir sagen, dass eine von $n \in \mathbb{N}$ abhängige Eigenschaft $E(n)$ für *fast alle* n gilt, falls $E(n)$ für alle $n \in \mathbb{N} \setminus S$ gilt, wobei $S \subset \mathbb{N}$ eine endliche Menge ist.

Behauptung. *Die Bedingung* $a = \lim x_n$ *ist äquivalent zur Bedingung:*

$$\boxed{\forall \varepsilon > 0 \quad |x_n - a| < \varepsilon \text{ gilt für fast alle } n.} \tag{2.2}$$

Beweis. Nach Definition ergibt (2.1), dass $|x_n - a| < \varepsilon$ für alle $n \geq N$ für ein (von ε abhängiges) N. Somit gilt (2.2). Umgekehrt, gilt (2.2), so gibt es eine endliche Menge $S \subset \mathbb{N}$ so dass $|x_n - a| < \varepsilon$ für alle $n \in S^c$ gilt. Setzen wir $N = \max S + 1$. Dann $n \geq N$ impliziert $n \in S^c$ und somit (2.1). ■

Die Bedingung $|x_n - a| < \varepsilon$ ist offensichtlich äquivalent zu

$$x_n \in (a - \varepsilon, a + \varepsilon).$$

Das Intervall $(a - \varepsilon, a + \varepsilon)$ heißt die ε-Umgebung von a und wird mit $U_\varepsilon(a)$ bezeichnet. Somit erhalten wir noch eine Umformulierung von dem Begriff des Limes.

Behauptung. *Die Bedingung* $a = \lim x_n$ *gilt genau dann, wenn*

$$\boxed{\forall \varepsilon > 0 \ \ x_n \in U_\varepsilon(a) \ \text{ für fast alle } n,} \tag{2.3}$$

d.h. jede Umgebung $U_\varepsilon(a)$ *von* a *fast alle Folgenglieder* x_n *enthält.*

Daraus erhalten wir folgendes: a ist kein Grenzwert von $\{x_n\}$ genau dann, wenn

$$\exists \varepsilon > 0 \ \text{ so dass die Menge } \{n : x_n \notin U_\varepsilon(a)\} \text{ unendlich ist,} \tag{2.4}$$

wo (2.4) die Negation der Bedingung (2.3) ist.

Bemerkung. Die Definition (2.1) von dem Grenzwert ist einer von den wichtigsten Begriffen in ganzer Mathematik. Der Begriff von Grenzwert wurde intuitiv, ohne genaue Definition, schon von den Begründern von Infinitesimalrechnung Isaac Newton und Wilhelm-Gottfried Leibniz im 17. Jahrhundert benutzt. Man brauchte fast 150 Jahre von weiterer Entwicklung der Analysis um eine rigorose Definition des Grenzwertes vorzulegen. Diese Definition wurde von Augustin Louis Cauchy in ca. 1821 gegeben, was als der Anfang von moderner Analysis betrachtet werden kann.

Beispiel. 1. Zeigen wir, dass die Folge $x_n = \frac{1}{n}$ gegen 0 konvergiert. In der Tat, nach dem Archimedischen Prinzip, für jedes $\varepsilon > 0$ existiert $N \in \mathbb{N}$ mit $N > \frac{1}{\varepsilon}$ woraus folgt $\frac{1}{N} < \varepsilon$. Deshalb für jedes $n \geq N$ gilt $\frac{1}{n} < \varepsilon$, d.h. $\left|\frac{1}{n} - 0\right| < \varepsilon$ und somit

$$\frac{1}{n} \to 0 \ \text{ für } \ n \to \infty.$$

Ebenso beweist man, dass für jedes $c \in \mathbb{R}$

$$\frac{c}{n} \to 0 \text{ für } n \to \infty.$$

2. Betrachten wir die Folge $\{x_n\}$ derart, dass $x_n = a$ für alle $n \geq n_0$. Man kann auch sagen, dass $x_n = a$ für fast alle n gilt. Dann auch $|x_n - a| = 0 < \varepsilon$ für fast alle n gilt, woraus $x_n \to a$ folgt.

3. Die Folge $x_n = n$ divergiert, da für jedes $a \in \mathbb{R}$ außerhalb Intervalles $U_1(a)$ unendliche Menge von Folgenglieder liegt. Ebenso divergiert die Folge $x_n = cn$ für jedes $c \neq 0$.

4. Zeigen wir, dass die Folge $x_n = (-1)^n$ divergiert. Der Wert $a = 1$ ist kein Grenzwert, da es außerhalb $U_1(1)$ unendlich viele Glieder mit dem Wert -1 gibt. Analog ist $a = -1$ kein Grenzwert. Sei $a \neq 1$ und $a \neq -1$. Dann existiert $\varepsilon > 0$ derart, dass die Umgebung $U_\varepsilon(a)$ weder 1 noch -1 enthält, und somit alle Glieder x_n außerhalb $U_\varepsilon(a)$ liegen. Deshalb ist a kein Grenzwert.

Die Konvergenz $x_n \to a$ kann als eine Annäherung von a mit den Folgenglieder x_n betrachtet werden. Betrachten wir $\varepsilon > 0$ als ein vorgegebener Approximationsfehler. Dann bedeutet die Bedingung (2.1), dass jedes Folgenglied x_n mit genügend großem n eine "gute" Annäherung von a ist, nämlich, mit dem Approximationsfehler $< \varepsilon$. Es ist wichtig zu betonen, dass diese Eigenschaft für beliebiges positives ε gelten muss, so dass der Approximationsfehler beliebig klein sein kann.

Beispiel. Betrachten wir eine Folge $\{x_n\}$ von positiven Zahlen, die per Induktion wie folgt definiert wird:

$$x_1 = 1, \quad x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right).$$

Die weiteren Glieder dieser Folge sind

$$\begin{aligned}
x_2 &= \tfrac{3}{2} && = 1.5 \\
x_3 &= \tfrac{17}{12} && = 1.416\,666\,666\,666\,666\,666\,7... \\
x_4 &= \tfrac{577}{408} && = 1.414\,215\,686\,274\,509\,803\,9... \\
x_5 &= \tfrac{665\,857}{470\,832} && = 1.414\,213\,562\,374\,689\,910\,6... \\
x_6 &= \tfrac{886\,731\,088\,897}{627\,013\,566\,048} && = 1.414\,213\,562\,373\,095\,048\,8... \\
x_7 &= \tfrac{1572\,584\,048\,032\,918\,633\,353\,217}{1111\,984\,844\,349\,868\,137\,938\,112} && = 1.414\,213\,562\,373\,095\,048\,8... \\
x_8 &= \tfrac{4946\,041\,176\,255\,201\,878\,775\,086\,487\,573\,351\,061\,418\,968\,498\,177}{3497\,379\,255\,757\,941\,172\,020\,851\,852\,070\,562\,919\,437\,964\,212\,608} && = 1.414\,213\,562\,373\,095\,048\,8...
\end{aligned}$$

Man kann zeigen, dass

$$x_n \to \sqrt{2} \text{ für } n \to \infty,$$

so dass die obigen Glieder x_n die Annäherungen von $\sqrt{2}$ mit den fallenden Fehler sind. In der Tat gilt

$$\sqrt{2} = 1.414\,213\,562\,373\,095\,048\,8...$$

so dass schon x_6 eine sehr gute Annäherung von $\sqrt{2}$ ist.

2.1.2 Eigenschaften des Limes

Für Beweise von Eigenschaften des Limes benutzen wir die folgende Eigenschaft des Begriffes "fast alle n".

Behauptung *Gelten zwei von $n \in \mathbb{N}$ abhängigen Aussagen $A(n)$ und $B(n)$ für fast alle n, so gilt $A(n) \wedge B(n)$ auch für fast alle n.*

Beweis. Nach Definition gibt es endliche Mengen $S, T \subset \mathbb{N}$ so dass $A(n)$ für alle $n \in S^c$ gilt und $B(n)$ für alle $n \in T^c$ gilt. Dann gilt $A(n) \wedge B(n)$ für alle

$$n \in S^c \cap T^c = (S \cup T)^c .$$

Da $S \cup T$ endlich ist, daraus folgt, dass $A(n) \wedge B(n)$ für fast alle n gilt. ■

Die weiteren Eigenschaften des Grenzwertes werden in dem folgenden Satz angegeben.

Satz 2.1 *(a) Seien $\{x_n\}$ und $\{y_n\}$ zwei konvergente Folgen mit $x_n \leq y_n$ für fast alle n, so gilt*

$$\lim x_n \leq \lim y_n.$$

Insbesondere die Bedingung $x_n = y_n$ für fast alle n impliziert $\lim x_n = \lim y_n$. Folglich ist der Grenzwert einer konvergenten Folge eindeutig bestimmt.

(b) Seien $\{x_n\}, \{y_n\}, \{z_n\}$ drei Folgen mit $x_n \leq y_n \leq z_n$ für fast alle n. Gilt

$$\lim x_n = \lim z_n =: a,$$

so ist $\{y_n\}$ auch konvergent und

$$\lim y_n = a.$$

(c) Jede konvergente Folge ist (nach oben und nach unten) beschränkt.

(d) $x_n \to a \Leftrightarrow |x_n - a| \to 0$ (insbesondere $x_n \to 0 \Leftrightarrow |x_n| \to 0$).

Beweis. (a) Bezeichnen wir $a = \lim x_n$ und $b = \lim y_n$. Für jedes $\varepsilon > 0$ gelten die folgenden Bedingungen für fast alle n:

$$a - \varepsilon < x_n \leq y_n < b + \varepsilon,$$

woraus folgt, dass $a < b + 2\varepsilon$ und

$$a - b < 2\varepsilon. \tag{2.5}$$

Da $\varepsilon > 0$ beliebig ist, so kann die Zahl $a - b$ positiv nicht sein (sonst gilt (2.5) mit $\varepsilon = \frac{1}{2}(a - b)$ nicht), woraus folgt $a - b \leq 0$ und somit $a \leq b$.

(b) Für jedes $\varepsilon > 0$ gelten die folgenden Bedingungen für fast alle n:

$$a - \varepsilon < x_n \leq y_n \leq z_n < a + \varepsilon,$$

woraus folgt $y_n \in U_\varepsilon(a)$ und somit $y_n \to a$.

(c) Sei $x_n \to a$. Die Menge von x_n, die in $U_1(a)$ liegen, ist beschränkt, da $U_1(a)$ beschränkt ist. Außerhalb $U_\varepsilon(a)$ liegen nur endlich viele Glieder, so dass die Menge davon auch beschränkt ist. Somit ist die ganze Menge $\{x_n\}_{n=1}^{\infty}$ beschränkt als die Vereinigung zweier beschränkten Mengen.

(d) Die Bedingung $x_n \to a$ bedeutet, dass für jedes $\varepsilon > 0$ die Ungleichung $|x_n - a| < \varepsilon$ für fast alle n erfüllt ist, während $|x_n - a| \to 0$ bedeutet, dass

$$||x_n - a| - 0| < \varepsilon$$

für fast alle n erfüllt ist. Offensichtlich sind die beiden Bedingungen äquivalent. ■

Beachten wir, dass $x_n < y_n$ impliziert $\lim x_n \leq \lim y_n$ aber nicht $\lim x_n < \lim y_n$, was man im Beispiel $x_n = 0, y_n = \frac{1}{n}$ sieht.

Beispiel. Betrachten wir die Folge $x_n = a^n$ mit einem $a \in \mathbb{R}$ und untersuchen wir die Konvergenz von $\{x_n\}$ abhängig von dem Wert von a. Betrachten wir verschiedene Fälle.

1. Sei $a > 1$, d.h. $a = 1 + c$ wobei $c > 0$. Nach Bernoullische Ungleichung haben wir

$$a^n = (1 + c)^n \geq 1 + nc.$$

Da die Folge $\{nc\}_{n=1}^{\infty}$ unbeschränkt ist, so ist die Folge $\{a^n\}_{n=1}^{\infty}$ auch unbeschränkt und somit divergent.

2. Sei $a < -1$. Da $|x_n| = |a|^n$, so erhalten wir, dass die Folge $\{|x_n|\}$ unbeschränkt ist und somit auch $\{x_n\}$. Folglich ist $\{x_n\}$ divergent.

3. Sei $a = -1$. Die Folge $x_n = (-1)^n$ wurde schon betrachtet, und wir wissen, dass sie divergiert.

4. Sei $a = 1$. Dann $x_n = 1$ und diese Folge konvergiert gegen 1.

5. Sei $a = 0$. Dann $\{x_n\}$ ist auch eine konstante Folge, die gegen 0 konvergiert.

6. Sei $0 < a < 1$. Setzen wir $b = \frac{1}{a} > 1$. Wie oberhalb schreiben wir $b = 1 + c$ mit $c > 0$ und erhalten

$$b^n \geq 1 + nc > nc$$

woraus folgt

$$0 < a^n < \frac{1}{nc}.$$

Da $\frac{1}{nc} \to 0$, erhalten wir nach Satz 2.1(b), dass $a_n \to 0$.

7. Sei $-1 < a < 0$. Dann gilt $0 < |a| < 1$ und $|a^n| = |a|^n \to 0$, woraus folgt $a^n \to 0$.

Somit ist die Folge $\{a^n\}_{n\in\mathbb{N}}$ konvergent genau dann, wenn $-1 < a \leq 1$.

2.1.3 Rechenregeln

Hier beweisen wir einige Rechenregeln für den Grenzwert.

Satz 2.2 *Seien $x_n \to a$ und $y_n \to b$. Dann gelten*

$$x_n + y_n \to a + b, \qquad x_n - y_n \to a - b, \qquad x_n y_n \to ab.$$

Falls $y_n \neq 0$ und $b \neq 0$, so gilt auch

$$\frac{x_n}{y_n} \to \frac{a}{b}.$$

Man kann die obigen Regeln mit Hilfe von dem Zeichen lim umschreiben wie folgt:

$$\begin{aligned}\lim(x_n+y_n) &= \lim x_n+\lim y_n\\ \lim(x_n-y_n) &= \lim x_n-\lim y_n\\ \lim(x_ny_n) &= \lim x_n\lim y_n\\ \lim\frac{x_n}{y_n} &= \frac{\lim x_n}{\lim y_n}\end{aligned}$$

vorausgesetzt, dass die rechten Seiten sinnvoll sind. Für die konstante Folge $y_n=a$ erhalten wir, dass

$$\lim(x_n+a)=\lim x_n+a \quad \text{und} \quad \lim(ax_n)=a\lim x_n.$$

Beweis von Satz 2.2. Für jedes $\varepsilon>0$ gelten

$$|x_n-a|<\varepsilon \text{ und } \quad |y_n-b|<\varepsilon \tag{2.6}$$

für fast alle $n\in\mathbb{N}$. Daraus folgt, dass für fast alle n

$$|x_n+y_n-(a+b)|=|x_n-a+y_n-b|\le|x_n-a|+|y_n-b|<2\varepsilon=:\varepsilon'.$$

Da $\varepsilon'>0$ beliebige ist, so erhalten wir

$$x_n+y_n\to a+b.$$

Die Konvergenz $x_n-y_n\to a-b$ beweist man analog.

Um die Produktregel zu beweisen, nehmen wir an $\varepsilon<1$ und beachten, dass (2.6) ergibt

$$|x_n|<|a|+\varepsilon<|a|+1.$$

Die Differenz x_ny_n-ab lässt sich für fast alle n wie folgt abschätzen:

$$\begin{aligned}|x_ny_n-ab| &= |x_ny_n-x_nb+x_nb-ab|\\ &\le |x_n(y_n-b)|+|(x_n-a)b|\\ &\le (|a|+1)|y_n-b|+|b||x_n-a|\\ &\le (|a|+|b|+1)\varepsilon.\end{aligned}$$

Da $(|a|+|b|+1)\varepsilon$ eine beliebige positive Zahl ist, so erhalten wir $x_ny_n\to ab$.

Für die letzte Behauptung $\frac{x_n}{y_n}\to\frac{a}{b}$ reicht es zu beweisen, dass $\frac{1}{y_n}\to\frac{1}{b}$. Sei $b>0$ (für $b<0$ betrachten man die Folge $\{-y_n\}$ mit dem positiven Grenzwert $-b$). Da für fast alle n gilt

$$y_n\in U_{b/2}(b)=(\frac{b}{2},\frac{3b}{2}),$$

so erhalten wir, dass $y_n>\frac{b}{2}$ für fast alle n. Für jedes $\varepsilon>0$ gilt für fast alle n auch

$$|y_n-b|<\varepsilon,$$

woraus folgt, dass für fast alle n

$$\left|\frac{1}{y_n}-\frac{1}{b}\right|=\frac{|y_n-b|}{|y_n|\,b}<\frac{2\varepsilon}{b^2}.$$

Da $\frac{2\varepsilon}{b^2}$ beliebige positive Zahl ist, so folgt es, dass $\frac{1}{y_n} \to \frac{1}{b}$. ■

Beispiel. Betrachten wir die Folge

$$x_n = \frac{an^2 + bn + c}{a'n^2 + b'n + c} \tag{2.7}$$

wobei $a' \neq 0$, und bestimmen ihren Grenzwert. Wir haben

$$x_n = \frac{n^2 \left(a + \frac{b}{n} + \frac{c}{n^2}\right)}{n^2 \left(a' + \frac{b'}{n} + \frac{c}{n^2}\right)} = \frac{a + \frac{b}{n} + \frac{c}{n^2}}{a' + \frac{b'}{n} + \frac{c}{n^2}}.$$

Wir wissen schon, dass $\frac{1}{n} \to 0$. Daraus folgt, dass auch $\frac{1}{n^2} \to 0$, $\frac{b}{n} \to 0$, $\frac{c}{n^2} \to 0$ usw. Nach Satz 2.2 erhalten wir $x_n \to \frac{a}{a'}$.

2.1.4 Grenzwert in $\overline{\mathbb{R}}$

Umgebung in $\overline{\mathbb{R}}$. Hier definieren wir den Begriff des Limes mit den Werten in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

Definition. Für jedes $E \in \mathbb{R}$ definieren wir die Umgebung $U_E(+\infty)$ durch

$$U_E(+\infty) = (E, +\infty] = \{x \in \mathbb{R} : x > E\} \cup \{+\infty\}.$$

Analog definieren wir die Umgebung $U_E(-\infty)$ durch

$$U_E(-\infty) = [-\infty, E) = \{x \in \mathbb{R} : x < E\} \cup \{-\infty\}.$$

Definition. Eine Folge $\{x_n\}$ von Elementen von $\overline{\mathbb{R}}$ hat den Grenzwert $a \in \overline{\mathbb{R}}$ falls:

$$\text{jede Umgebung von } a \text{ enthält fast alle Glieder von } \{x_n\}. \tag{2.8}$$

Man schreibt in diesem Fall

$$\lim x_n = a \quad \text{oder} \quad x_n \to a \ \text{ für } \ n \to \infty$$

Ist a reell, so stimmt diese Definition mit der früheren Definition des Limes. In diesem Fall sagt man, dass x_n gegen a konvergiert. Ist $a = \pm\infty$, so sagt man, dass dass x_n gegen a *divergiert* (weißt man nicht ob a endlich oder unendlich ist, so sagt man doch, dass x_n gegen a konvergiert).

Divergiert die Folge x_n gegen $+\infty$ oder $-\infty$, so sagt man, dass die Folge *bestimmt divergiert.* Hat die Folge keinen Grenzwert in $\overline{\mathbb{R}}$, so sagt man, dass die Folge *unbestimmt divergiert.*

Die Aussagen (a) und (b) des Satzes 2.1 gelten auch in $\overline{\mathbb{R}}$.

Satz 2.1$'$ *Seien $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}}, \{z_n\}_{n\in\mathbb{N}}$ die Folgen von Elementen von $\overline{\mathbb{R}}$.*

(a) *Gilt $x_n \le y_n$ für fast alle n, so gilt $\lim x_n \le \lim y_n$, vorausgesetzt, dass $\lim x_n$ und $\lim y_n$ in $\overline{\mathbb{R}}$ existieren.*

(b) *Gelten $x_n \le y_n \le z_n$ für fast alle n und $\lim x_n = \lim z_n =: a \in \overline{\mathbb{R}}$, so gilt auch $\lim y_n = a$.*

Beweis. (a) Setzen wir $a = \lim x_n$, $b = \lim y_n$ und nehmen das Gegenteil an, dass $a > b$. Dann existieren Umgebungen $U(a)$ und $V(b)$ so dass für alle $x \in U(a)$ und $y \in V(b)$ gilt $x > y$. Andererseits für fast alle n gelten $x_n \in U(a)$, $y_n \in V(b)$ und $x_n \leq y_n$, was zum Widerspruch führt.

(b) Beweis ist analog. ■

Beispiel. 1. Die Folge $x_n = n$ divergiert gegen $+\infty$. Für jedes $E \in \mathbb{R}$ existiert $N \in \mathbb{N}$ mit $N > E$, so dass $\forall n \geq N$ gilt $x_n > E$ d.h. $x_n \in U_E(+\infty)$, woraus folgt $x_n \to +\infty$. Analog divergiert die Folge $x_n = -n$ gegen $-\infty$. Mit gleichem Argument beweist man, dass die Folge $x_n = cn$ gegen $+\infty$ divergiert, falls $c > 0$, und gegen $-\infty$ falls $c < 0$.

2. Betrachten wir die Folge $x_n = a^n$. Wir wissen schon, dass diese Folge konvergiert genau dann, wenn $a \in (-1, 1]$. Zeigen wir, dass falls $a > 1$ dann $a^n \to +\infty$. Schreiben wir $a = 1 + c$ wobei $c > 0$ und erhalten mit Hilfe von Bernoullischer Ungleichung

$$a^n = (1 + c)^n > cn.$$

Da $cn \to +\infty$, daraus folgt, dass auch $a^n \to +\infty$.

Für $a < -1$ wechselt die Folge $x_n = a^n$ immer das Vorzeichen. Somit enthält weder $U_0(+\infty)$ noch $U_0(-\infty)$ fast alle Glieder, und die Folge divergiert unbestimmt.

Operationen mit $+\infty$ und $-\infty$. Für jedes $a \in \overline{\mathbb{R}}$ definieren wir Addition mit $+\infty$ wir folgt:

$$(+\infty) + a = a + (+\infty) = \begin{cases} +\infty, & -\infty < a \leq +\infty \\ \text{unbestimmt}, & a = -\infty. \end{cases}$$

Analog definiert man Addition mit $-\infty$. Die Multiplikation mit $+\infty$ wird wie folgt definiert:

$$(+\infty) \cdot a = a \cdot (+\infty) = \begin{cases} +\infty, & 0 < a \leq +\infty \\ -\infty, & -\infty \leq a < 0 \\ \text{unbestimmt}, & a = 0. \end{cases}$$

und analog mit $-\infty$. Division durch ∞ wird wie folgt definiert:

$$\frac{a}{+\infty} = \begin{cases} 0, & a \in \mathbb{R} \\ \text{unbestimmt}, & a = +\infty \text{ or } a = -\infty. \end{cases}$$

Erinnern wir uns, dass auch $\frac{a}{0}$ unbestimmt ist. Somit bleiben die folgenden Operationen unbestimmt:

$$\infty - \infty, \quad 0 \cdot \infty, \quad \frac{\infty}{\infty}, \quad \frac{0}{0}.$$

Alle diese Ausdrücke heißen die *unbestimmte* Ausdrücke.

Satz 2.3 *Seien* $\{x_n\}$ *und* $\{y_n\}$ *zwei Folgen von reellen Zahlen mit* $\lim x_n = a$ *und* $\lim y_n = b$ *wobei* $a, b \in \overline{\mathbb{R}}$. *Dann gelten*

$$\lim (x_n + y_n) = a + b, \quad \lim (x_n - y_n) = a - b, \quad \lim (x_n y_n) = ab, \quad \lim \frac{x_n}{y_n} = \frac{a}{b}, \tag{2.9}$$

vorausgesetzt, dass die Ausdrücke $a + b$, $a - b$, ab, *bzw* $\frac{a}{b}$ *bestimmt sind (und* $y_n \neq 0$ *im letzten Fall).*

Beweis. Für reellen a, b war das im Satz 2.2 bewiesen. Für unendlichen a, b muss man verschiedene Möglichkeiten betrachten und zeigen, dass in allen Fällen die Identitäten (2.9) mit den Definitionen von Operationen mit $\pm\infty$ übereinstimmen (siehe Aufgaben). Zum Beispiel, $\lim x_n = +\infty$ impliziert, dass $\lim \frac{1}{x_n} = 0$, was der Definition $\frac{1}{+\infty} = 0$ entspricht. ■

Beispiel. Die Folgen $x_n = n + c$ und $y_n = n$ haben den Grenzwert $+\infty$ aber die Differenz $x_n - y_n = c$ hat den Grenzwert c, was beliebige reelle Zahl ist. Für die Folge $x_n = n + (-1)^n$ gilt auch $\lim x_n = +\infty$ aber die Differenz $x_n - y_n = (-1)^n$ keinen Grenzwert hat. Somit lässt der Ausdruck $\infty - \infty$ sich nicht vernünftig definieren. Für den Ausdruck $\frac{\infty}{\infty}$ wird ein Gegenbeispiel von der Folge (2.7) geliefert, und für ein Gegenbeispiel für $0 \cdot \infty$ siehe Aufgaben.

2.1.5 Monotone Folgen

Definition. Eine Folge $\{x_n\}_{n=1}^{\infty}$ von reellen Zahlen heißt *monoton steigend* falls $x_{n+1} \geq x_n$ für alle $n \in \mathbb{N}$, and *monoton fallend* falls $x_{n+1} \leq x_n$ für alle $n \in \mathbb{N}$. Eine Folge heißt monoton, falls sie entweder monoton steigend oder monoton fallend ist.

Beispiel. Die Folge $x_n = n$ ist monoton steigend, die Folge $x_n = \frac{1}{n}$ ist monoton fallend, die konstante Folge $x_n = a$ ist gleichzeitig monoton steigend und fallend, die Folge $x_n = (-1)^n$ ist nicht monoton.

Ist $\{x_n\}$ monoton steigend, dann gilt $x_n \geq x_m$ für alle $n \geq m$. Insbesondere gilt $x_n \geq x_1$ so dass x_1 eine untere Schranke ist, und $\{x_n\}$ nach unten beschränkt ist. Ist $\{x_n\}$ monoton fallend, so gilt $x_n \leq x_m$ für alle $n \geq m$, und $\{x_n\}$ ist nach oben beschränkt.

Hauptsatz 2.4 (Monotoniekriterium) *Sei $\{x_n\}_{n\in\mathbb{N}}$ eine monotone Folge von reellen Zahlen.*
Ist $\{x_n\}$ monoton steigend, so gilt$\lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\}$.
Ist $\{x_n\}$ monoton fallend, so gilt

$$\lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}.$$

Folglich hat jede monotone Folge immer einen Grenzwert in $\overline{\mathbb{R}}$ (endlich oder $\pm\infty$), und sie konvergiert genau dann, wenn sie beschränkt ist.

Beweis. Sei $\{x_n\}$ monoton steigend und beschränkt. Nach dem Satz 1.9 existiert das Supremum

$$a = \sup\{x_n\} \in \mathbb{R}$$

Beweisen wir, dass $x_n \to a$. Für jedes $\varepsilon > 0$ ist die Zahl $a - \varepsilon$ keine obere Schranke von $\{x_n\}$. Deshalb existiert $N \in \mathbb{N}$ mit

$$x_N > a - \varepsilon.$$

Da die Folge $\{x_n\}$ monoton steigend ist, erhalten wir, dass

$$x_n > a - \varepsilon \text{ für alle } n \geq N$$

Da nach Definition von a gilt

$$x_n \leq a,$$

so erhalten wir, dass $x_n \in (a-\varepsilon, a+\varepsilon)$ für alle $n \geq N$, woraus $x_n \to a$ folgt.

Ist $\{x_n\}$ unbeschränkt, so für jedes $E \in \mathbb{R}$ existiert $N \in \mathbb{N}$ mit $x_N > E$. Da $\{x_n\}$ monoton steigend ist, so erhalten wir, dass $x_n > E$ für alle $n \geq N$, woraus $\lim x_n = +\infty$ folgt, was zu bewiesen war.

Der Fall von den monoton fallenden Folge ist analog. ■

Beispiel. Betrachten wir die Folge $\{x_n\}_{n\in\mathbb{N}}$ die wie folgt definiert ist:

$$x_1 = 1, \quad x_{n+1} = x_n + \frac{1}{x_n^2}.$$

Die ersten Glieder der Folge sind wie folgt:

$$1,\ 2,\ \frac{9}{4} = 2,25,\ \frac{793}{324} = 2,447...,\quad \frac{532\,689\,481}{203\,747\,076} = 2,614...,\ \text{usw.}$$

Diese Folge ist offensichtlich monoton steigend und somit hat einen Grenzwert $x = \lim x_n \in [0, +\infty]$. Es folgt, dass x die folgende Gleichung erfüllt:

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

Da es keine reelle Zahl gibt, die diese Gleichung erfüllt, so beschließen wir, dass $x = +\infty$.

Beispiel. Der Wert der unendlichen Folge von Wurzeln

$$\sqrt{1+\sqrt{1+\sqrt{1+\sqrt{1+...}}}}$$

wird als $\lim x_n$ definiert, wobei

$$x_1 = \sqrt{1}, \quad x_{n+1} = \sqrt{1+x_n}.$$

Man kann zeigen, dass die Folge $\{x_n\}$ monoton steigend und beschränkt ist, woraus die Konvergenz folgt (siehe Aufgaben). Der Grenzwert $\lim x_n$ lässt sich danach explizit berechnen.

Beispiel. Die Folge $x_n = \left(1+\frac{1}{n}\right)^n$ ist monoton steigend und beschränkt (siehe Aufgaben). Der Grenzwert

$$e := \lim x_n = 2,718281828459045...$$

ist eine transzendente Zahl, die eine wichtige Rolle in Analysis spielt.

2.1.6 Unendliche Summen (Reihen)

Sei $\{a_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zahlen. Betrachten wir den Ausdruck

$$\sum_{n=1}^{\infty} a_n,$$

der heißt eine *Reihe*. Die Summe (der Wert) der Reihe wird wie folgt definiert. Für jedes $N \in \mathbb{N}$ definieren wir die *Partialsumme*

$$S_N = \sum_{n=1}^{N} a_n,$$

und danach setzen wir

$$\sum_{n=1}^{\infty} a_n = \lim_{N\to\infty} S_N$$

vorausgesetzt, dass der Grenzwert $\lim_{N\to\infty} S_N$ in $\overline{\mathbb{R}}$ existiert.

Man sagt, dass die Reihe $\sum a_n$ konvergent bzw bestimmt divergent bzw unbestimmt divergent, wenn so die Folge $\{S_N\}$ ist. Ist die Reihe $\sum a_n$ unbestimmt divergent, so ist der Wert von $\sum a_n$ nicht definiert.

Analog definiert man den Wert der Reihe

$$\sum_{n=m}^{\infty} a_n,$$

wobei $m \in \mathbb{Z}$. Es folgt, dass für $m' > m$

$$\sum_{n=m}^{\infty} a_n = \sum_{n=m}^{m'-1} a_n + \sum_{n=m'}^{\infty} a_n.$$

Nach den Eigenschaften des Limes erhalten wir die folgenden Eigenschaften von Reihen.

Satz 2.5 (a) *Für $a_n, b_n, c \in \mathbb{R}$ gelten die Identitäten*

$$\sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n,$$

$$\sum_{n=1}^{\infty} c a_n = c \sum_{n=1}^{\infty} a_n,$$

vorausgesetzt, dass die rechten Seiten bestimmt sind.

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

$$\sum_{n=1}^{\infty} a_n \le \sum_{n=1}^{\infty} b_n,$$

vorausgesetzt, dass die beiden Seiten bestimmt sind.

Beweis. (a) Betrachten wir die Partialsumme

$$S_N = \sum_{n=1}^{N} (a_n + b_n) = \sum_{n=1}^{N} a_n + \sum_{n=1}^{N} b_n = A_N + B_N,$$

wobei A_N und B_N die Partialsummen von $\sum a_n$ bzw $\sum b_n$ sind. Nach dem Satz 2.3 erhalten wir

$$\lim_{N\to\infty} S_N = \lim_{N\to\infty} (A_N + B_N) = \lim_{N\to\infty} A_N + \lim_{N\to\infty} B_N = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n.$$

Die zweite Identität wird analog bewiesen.

(b) Die Voraussetzung $a_n \leq b_n$ ergibt $A_N \leq B_N$ und somit $\lim A_N \leq \lim B_N$. ∎

Definition. Eine Reihe $\sum_{k=1}^{\infty} a_k$ heißt *nichtnegativ* falls alle Glieder a_k nichtnegative reelle Zahlen sind.

Satz 2.6 *Für jede nichtnegative Reihe $\sum_{n=1}^{\infty} a_n$ ist ihre Summe immer bestimmt als Element von $[0, +\infty]$. Folglich bestehen es nur zwei Möglichkeiten:*

1. *entweder $\sum_{n=1}^{\infty} a_n < \infty$ und die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert;*
2. *oder $\sum_{k=1}^{\infty} a_k = +\infty$ und die Reihe bestimmt divergiert.*

Beweis. Da $a_n \geq 0$, so ist die Folge $\{S_N\}_{N\in\mathbb{N}}$ von Partialsummen monoton steigend und somit $\lim S_N$ existiert in $\overline{\mathbb{R}}$ nach den Satz 2.4. Ist $\lim S_N$ endlich, so ist die Reihe konvergent, ist $\lim S_N = +\infty$, so ist die Reihe bestimmt divergent. ∎

Beispiel. Betrachten wir die *geometrische* Reihe

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

mit $x \in \mathbb{R}$. Die Partialsumme ist

$$S_N = \sum_{n=0}^{N} x^n = \frac{1 - x^{N+1}}{1 - x} \tag{2.10}$$

(siehe Aufgabe 34). Im Fall $x \in (-1, 1)$ erhalten wir $x^{N+1} \to 0$ für $N \to \infty$, woraus folgt

$$\sum_{n=0}^{\infty} x^n = \lim_{N\to\infty} S_N = \frac{1}{1 - x}. \tag{2.11}$$

Es folgt, dass für jedes $m \in \mathbb{Z}$

$$\sum_{n=m}^{\infty} x^n = x^m \sum_{n=m}^{\infty} x^{n-m} = x^m \sum_{k=0}^{\infty} x^k = \frac{x^m}{1 - x}. \tag{2.12}$$

Im Fall $x > 1$ folgt es aus (2.10), dass

$$S_N = \frac{x^{N+1} - 1}{x - 1} \to +\infty \quad \text{für} \quad N \to \infty,$$

so dass $\sum_{n=0}^{\infty} x^n = +\infty$. Im Fall $x = 1$ gilt (2.10) nicht, aber wir haben $S_N = N$ und somit wieder $\sum_{n=0}^{\infty} x^n = \infty$.

Für $x \leq -1$ hat die Folge $\left\{x^{N+1}\right\}$ keinen Grenzwert, woraus folgt, dass $\sum_{n=0}^{\infty} x^n$ unbestimmt divergiert.

Beispiel. *Die harmonische* Reihe

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

ist bestimmt divergent, während die Reihe

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

konvergent ist (siehe Aufgaben).

2.2 Zahlensystem

2.2.1 q-adische Darstellung von natürlichen Zahlen

Im nächsten Satz führen wir ein q-adisches Zahlensystem ein. Bezeichnen wir mit $\mathbb{Z}_+$ die Menge von nichtnegativen ganzen Zahlen, d.h. $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$.

Satz 2.7 *Sei $q > 1$ eine natürliche Zahl. Für jedes $x \in \mathbb{N}$ existieren genau ein $n \in \mathbb{Z}_+$ und genau eine Folge $\{a_k\}_{k=0}^n$ von ganzen Zahlen $a_k \in \{0, ..., q-1\}$ mit $a_n \neq 0$ und*

$$x = \sum_{k=0}^{n} a_k q^k \quad \left(= a_n q^n + a_{n-1} q^{n-1} + ... + a_1 q + a_0\right). \tag{2.13}$$

Die Identität (2.13) ist die Darstellung von x im q-adischen Zahlensystem (= *das Zahlensystem zur Basis* q). Die übliche symbolische Abkürzung von (2.13) sieht wie folgt aus:

$$x = a_n a_{n-1} ... a_0 \quad \text{oder} \quad x = \left(a_n a_{n-1} ... a_0\right)_q.$$

(nicht mit dem Produkt zu verwechseln). Die Zahlen $\{0, ..., q-1\}$ heißen q-*adische Ziffern*. Jede Zahl a_k heißt der *Ziffernwert* an der Stelle k, und q^k heißt der *Stellenwert* an der Stelle k. Die Ziffer a_n an der höchsten Stelle n muß immer positiv sein.

Die gängigsten Basen sind $q = 2$ (Dualsystem), $q = \text{zehn} := 9 + 1$ (Dezimalsystem) und $q = \text{sechzehn} := 9 + 7$ (Hexadezimalsystem). Im Dualsystem gibt es nur zwei Ziffern 0 und 1, und somit sind die Operationen im Dualsystem sehr einfach. Insbesondere hat die Multiplikationstabelle nur einen Eintrag $1 \cdot 1 = 1$. Die Zahlen $1, 2, 3, 4, 5, 6, 7, 8, 9$ werden im Dualsystem wie folgt dargestellt:

$$1_2,\ 10_2,\ 11_2,\ 100_2,\ 101_2,\ 110_2,\ 111_2,\ 1000_2,\ 1001_2.$$

Die Ziffern im Dezimalsystem sind $0, 1, 2, ..., 9$. Im Hexadezimalsystem sind die Ziffern $0, 1, ..., 9, A, B, C, D, E, F$, wobei die Buchstaben die Ziffern zwischen zehn und fünfzehn bezeichnen. Zum Beispiel, es gilt

$$(C3F)_{\text{hex}} = \left(C \cdot q^2 + 3 \cdot q + F\right)_{\text{hex}} = \left(12 \cdot 16^2 + 3 \cdot 16 + 15\right)_{\text{dez}} = (3135)_{\text{dez}}.$$

Lemma 2.8 *(*Ganzzahlige Division*) Sei $q > 1$ eine natürliche Zahl. Für jedes $x \in \mathbb{Z}$ existiert genau ein $y \in \mathbb{Z}$ und ein $r \in \{0, ..., q-1\}$ mit*

$$x = qy + r. \tag{2.14}$$

Diese Darstellung heißt die *ganzzahlige Division* von x durch q. Die Zahlen y und r heißen der *Quotient* bzw der *Rest* der ganzzahligen Division.

Beweis. Für die Eindeutigkeit nehmen wir an, dass es noch eine solche Darstellung

$$x = qy' + r'$$

gibt. Daraus folgt, dass

$$r - r' = q\,(y - y').$$

Ist $y \neq y'$ so gilt $|y - y'| \geq 1$ und $|q\,(y - y')| \geq q$ während $|r - r'| \leq q - 1$. Somit gilt $y = y'$ und dann auch $r = r'$. Für die Existenz setzen wir $y = \left[\frac{x}{q}\right]$, woraus folgt

$$y \leq \frac{x}{q} < y + 1.$$

Die ganze Zahl $r = x - qy$ liegt dann in $[0, q)$, so dass $r \in \{0, ..., q-1\}$. ∎

Beweis von Satz 2.7. Induktion nach x. Induktionsanfang für $x = 1$: (2.13) ist erfüllt genau dann, wenn $n = 0$ und $a_0 = 1$.

Induktionsschritt. Angenommen ist, dass jedes $y \in \mathbb{N}$ mit $y < x$ eine eindeutige q-adische Darstellung hat. Beweisen wir das Gleiche für x. Für $x \in \mathbb{N}$ bestimmen wir zunächst die Darstellung (2.14). Ist $y = 0$ so ist $x = a_0$ mit $a_0 = r$ die q-adische Darstellung von x. Sonst gilt $1 \leq y < x$ und nach Induktionsvoraussetzung hat y eine q-adische Darstellung

$$y = \sum_{k=0}^{m} b_k q^k.$$

Einsetzen in (2.14) ergibt die q-adische Darstellung von x:

$$x = q \sum_{k=0}^{m} b_k q^k + r = \sum_{k=0}^{m} b_k q^{k+1} + r = \sum_{l=1}^{m+1} b_{l-1} q^l + r = \sum_{l=0}^{m+1} a_l q^l,$$

wobei $a_l = b_{l-1}$ für $l \geq 1$ und $a_0 = r$.

Beweisen wir jetzt die Eindeutigkeit der q-adischen Darstellung (2.13). Für $x < q$ gibt es nur eine Darstellung $x = a_0$ mit $n = 0$. Sei $x \geq q$. Dann $n \geq 1$ und wir haben

$$x = a_n q^n + ... + a_1 q_1 + a_0 = q\left(a_n q^{n-1} + ... + a_1\right) + a_0 = qy + a_0,$$

wobei

$$y = a_n q^{n-1} + ... + a_1. \tag{2.15}$$

Da $x = qy + a_0$ die ganzzahlige Division ist, so sind y und a_0 eindeutig bestimmt. Da $y < x$, so ist die q-adische Darstellung (2.15) von y eindeutig bestimmt nach Induktionsvoraussetzung, woraus folgt, dass der Wert von n und alle Ziffern $a_1, ..., a_n$ auch eindeutig bestimmt sind. ∎

2.2.2 q-adische Brüche

Jetzt definieren wir q-adische Darstellung von den reellen Zahlen. Betrachten wir zunächst eine unendliche Folge$\{a_k\}_{k=1}^{\infty}$ von q-adischen Ziffern (d.h. $a_k \in \{0, ..., q-1\}$) und die Reihe

$$\sum_{k=1}^{\infty} a_k q^{-k} = a_1 q^{-1} + a_2 q^{-2} + \dots \,. \tag{2.16}$$

Satz 2.9 *Für jede Folge $\{a_k\}_{k=1}^{\infty}$ von q-adischen Ziffern ist die Reihe $\sum_{k=1}^{\infty} a_k q^{-k}$ konvergent und ihre Summe liegt in $[0,1]$.*

Definition. Die Reihe $\sum_{k=1}^{\infty} a_k q^{-k}$ heißt q-adischer *Bruch*, und die Summe $x = \sum_{k=1}^{\infty} a_k q^{-k}$ heißt der Wert des Bruches. Die übliche symbolische Abkürzung dieser Identität ist

$$x = (0, a_1 a_2 a_3 ...)_q \quad \text{oder} \quad x = 0, a_1 a_2 a_3 ...$$

Beweis von Satz 2.9. Da $a_k q^{-k} \geq 0$, so ist die Reihe $\sum_{k=1}^{\infty} a_k q^{-k}$ entweder konvergent oder bestimmt divergent. Zeigen wir, dass

$$\sum_{k=1}^{\infty} a_k q^{-k} < \infty.$$

Da $a_n \leq q-1$, so erhalten wir nach (2.12)

$$\sum_{k=1}^{\infty} a_k q^{-k} \leq (q-1) \sum_{k=1}^{\infty} q^{-k} = (q-1) \frac{q^{-1}}{1-q^{-1}} = 1. \tag{2.17}$$

Somit ist die Reihe $\sum_{k=1}^{\infty} a_k q^{-k}$ konvergent und ihre Summe liegt in $[0,1]$. ■

Beispiel. Es folgt aus (2.17), dass

$$(0, 111...)_q = \sum_{k=1}^{\infty} q^{-k} = \frac{1}{q-1}. \tag{2.18}$$

Der nächste Satz ist die Umkehrung von Satz 2.9.

Satz 2.10 *Sei $q > 1$ eine natürliche Zahl. Für jedes $x \in [0,1)$ gibt es einen q-adischen Bruch mit dem Wert x.*

Beweis. Bestimmen wir die Folge $\{a_n\}_{n=1}^{\infty}$ von Ziffern per Induktion nach n, so dass für alle $n \in \mathbb{N}$ gilt

$$\sum_{k=1}^{n} a_k q^{-k} \leq x < \sum_{k=1}^{n} a_k q^{-k} + q^{-n}. \tag{2.19}$$

Induktionsanfang für $n = 1$. Die Ungleichung (2.19) für $n = 1$ ist

$$a_1 q^{-1} \leq x < a_1 q^{-1} + q^{-1}, \tag{2.20}$$

was äquivalent zu

$$a_1 \le qx < a_1 + 1 \tag{2.21}$$

ist, d.h.

$$a_1 = [qx] \, .$$

Da $0 \le x < 1$, so erhalten wir $0 \le qx < q$ und somit $0 \le a_1 < q$. Deshalb ist a_1 eine q-adische Ziffer ist. Damit haben wir gezeigt, dass a_1 existiert.

Induktionsschritt von $< n$ nach n. Seien a_k für alle $k < n$ schon bestimmt. Wir müssen beweisen, dass es eine Ziffer a_n mit (2.19) gibt. Setzen wir

$$y = \sum_{k=1}^{n-1} a_k q^{-k}$$

und umschreiben (2.19) wie folgt:

$$y + a_n q^{-n} \le x < y + a_n q^{-n} + q^{-n}.$$

Diese Ungleichungen sind äquivalent zu

$$a_n \le q^n (x - y) < a_n + 1,$$

was ergibt

$$a_n = [q^n (x - y)] \, .$$

Es bleibt nur zu zeigen, dass a_n eine q-adische Ziffer ist. Nach Induktionsvoraussetzung haben wir

$$y \le x < y + q^{-(n-1)},$$

woraus folgt

$$\begin{aligned} 0 &\le x - y < q^{-n+1}, \\ 0 &\le q^n (x - y) < q, \\ 0 &\le [q^n (x - y)] < q \end{aligned}$$

was zu beweisen war. Es folgt aus (2.19) für $n \to \infty$, dass $x = 0, a_1 a_2 ...$ ■

Mit Hilfe von Sätzen 2.7 und 2.10 kann jedes $x \in \mathbb{R}_+$ im q-adischen Zahlensystem dargestellt werden, wie folgt. In der Tat kann jedes $x \ge 0$ eindeutig in der Form

$$x = a + b$$

zerlegt werden, wobei $a = [x] \in \mathbb{Z}_+$ der Ganzzahlanteil von x ist und $b = x - a \in [0, 1)$ der *Bruchteil* von x. Manchmal benutzt man die Bezeichnung $b = \{x\}$ (x in den geschwungenen Klammern – nicht mit der Menge $\{x\}$ zu verwechseln).

Ist $a = 0$, so gilt $x \in [0, 1)$ und die q-adische Darstellung von x wird von Satz 2.10 gegeben. Sei $a > 0$. Nach Satz 2.7 hat a die q-adische Darstellung

$$a = \sum_{k=0}^{n} a_k q^k = (a_n a_{n-1} ... a_0)_q \ .$$

Nach Satz 2.10 hat b die Darstellung als ein q-adischer Bruch

$$b = \sum_{l=1}^{\infty} b_l q^{-l} = (0, b_1 b_2 b_3 ...)_q \ .$$

Deshalb erhalten wir die Identität

$$x = \sum_{k=0}^{n} a_k q^k + \sum_{l=1}^{\infty} b_l q^{-l},$$

die symbolisch abgekürzt wird wie folgt

$$x = (a_n a_{n-1} ... a_0, b_1 b_2 b_3 ...)_q \ .$$

Der Ausdruck $(a_n a_{n-1} ... a_0, b_1 b_2 b_3 ...)_q$ (wobei a_k und b_l die q-adischen Ziffern sind) heißt eine *q-adische Zahl*. Somit erhalten wir folgendes.

Folgerung. *Jede q-adische Zahl hat einen Wert in $\mathbb{R}_+$. Für jedes $x \in \mathbb{R}_+$ existiert eine q-adische Zahl mit dem Wert x.*

Mit Hilfe von der q-adischen Darstellung der reellen Zahlen können die arithmetischen Operationen mit Zahlen durchgeführt werden, wie schriftliche Addition, Subtraktion, Multiplikation und Division.

Bemerkung. Der Satz 2.10 enthält keine Eindeutigkeitsaussage, da dies nicht immer gilt. Zum Beispiel, es folgt aus (2.18), dass für $a = q - 1$

$$(0, aaa....)_q = \frac{a}{q-1} = 1,$$

wobei 1 auch die q-adische Darstellung $1 = (1)_q$ hat. Ein q-adischer Bruch $(0, b_1 b_2 ...)_q$ heißt *echt* falls die Menge

$$\text{die Menge } \{k \in \mathbb{N} : b_k < q - 1\} \text{ unendlich ist.} \tag{2.22}$$

unendlich ist. Man kann beweisen, dass für jedes $x \in [0, 1)$ es genau einen q-adischen echten Bruch mit dem Wert x gibt (siehe Aufgaben).

Bemerkung. Die Dualsystem kann benutzt werden um zu beweisen, dass $|\mathbb{R}| = 2^{|\mathbb{N}|}$. Dafür betrachten wir die Menge $\mathcal{F}$ von allen Folgen $b = \{b_k\}_{k \in \mathbb{N}}$ mit $b_k \in \{0, 1\}$ (Dualziffern) und die Abbildung

$$\begin{aligned} \varphi &: \mathcal{F} \to [0, 1] \\ \varphi(b) &= (0, b_1 b_2 ...)_2 = \sum_{k=1}^{\infty} b_k 2^{-k}. \end{aligned}$$

Diese Abbildung ist surjektiv nach Satz 2.10, aber nicht injektiv, was wir oberhalb gesehen haben. Bezeichnen wir mit $\mathcal{F}_e$ die Menge von den Folgen $a \in \mathcal{F}$ die echt sind, d.h. die Bedingung (2.22) erfüllen. Nach der obigen Bemerkung ist $\varphi|_{\mathcal{F}_e}$ eine Bijektion von $\mathcal{F}_e$ auf $[0, 1)$, so dass

$$\mathcal{F}_e \sim [0, 1).$$

Man kann beweisen, dass die Differenz $\mathcal{F}\backslash\mathcal{F}_e$ abzählbar ist (siehe Aufgaben). Da $[0,1) \sim \mathbb{R}$ und nach dem Satz 1.26 $\mathcal{F} \sim \mathcal{F}_e$, so erhalten wir

$$\mathcal{F} \sim \mathbb{R}.$$

Andererseits gilt

$$\mathcal{F} \sim 2^{\mathbb{N}},$$

da die folgende Abbildung

$$\begin{aligned} \psi &: \mathcal{F} \to 2^{\mathbb{N}} \\ \psi(b) &= \{k \in \mathbb{N} : b_k = 0\} \end{aligned}$$

bijektiv ist. Somit haben wir, dass $\mathbb{R} \sim \mathcal{F} \sim 2^{\mathbb{N}}$, woraus $|\mathbb{R}| = 2^{|\mathbb{N}|}$ folgt.

2.3 * Existenz und Eindeutigkeit von $\mathbb{R}$ (Skizze)

Die Menge $\mathbb{R}$ von reellen Zahlen wurde von uns axiomatisch eingeführt. Natürlich entsteht die Frage, ob eine Menge $\mathbb{R}$ existiert, die alle Axiomen erfüllt, und ob sie eindeutig bestimmt ist. Die Existenz von $\mathbb{R}$ kann nur in den Rahmen von anderen axiomatischen Theorien bewiesen werden, zum Beispiel in den Rahmen der Mengenlehre.

Konstruieren wir zunächst die Menge $\mathbb{N}$ von natürlichen Zahlen. Dafür definiert man ein endliche Menge als eine Menge die zu keiner echten Teilmenge gleichmächtig ist. Die natürlichen Zahlen (inklusive Null) lassen sich als die Kardinalzahlen von endlichen Mengen definieren.

Zum Beispiel, definieren wir 0 als die Kardinalzahl von $\emptyset$, d.h. $0 := |\emptyset|$. Um 1 zu definieren, benutzen wir Potenzmenge $A = 2^{\emptyset}$, die aus einem Element $\emptyset$ besteht, d.h. $A = \{\emptyset\}$. Dann setzen wir $1 := |A|$. Warum existieren die höheren Zahlen? Betrachten wir die Potenzmenge $B = 2^A = \{\emptyset, A\}$, die aus zwei Elementen besteht, und setzen $2 := |B|$. Dann betrachten wir die Potenzmenge $C = 2^B = \{\emptyset, \{\emptyset\}, \{A\}, B\}$, die aus vier Elementen besteht, usw. Mit Hilfe von weiteren Potenzmengen können wir beliebig große endliche Mengen konstruieren und somit Existenz von beliebig großen natürlichen Zahlen zeigen. Allerdings muss man Operationen in $\mathbb{N}$ definieren und alle notwendigen Eigenschaften beweisen. Danach definiert man die negativen Zahlen und die Operationen in $\mathbb{Z}$.

Es gibt eine andere Theorie, wo die Menge $\mathbb{N}$ axiomatisch mit Hilfe von *Peano-Axiomen* definiert wird. Eine Menge $\mathbb{N}$ heißt die Menge von natürlichen Zahlen und die Elementen von $\mathbb{N}$ heißen die natürlichen Zahlen, falls es eine Abbildung $F : \mathbb{N} \to \mathbb{N}$ gibt, die die folgenden Axiome erfüllt:

1. F ist injektiv (d.h. $F(n) = F(m) \Rightarrow n = m$).
2. Es gibt ein Element $1 \in \mathbb{N}$ das nicht zum Bild von F gehört (d.h. $F(n) \neq 1$ für alle $n \in \mathbb{N}$).
3. Sei M eine Menge mit den folgenden Eigenschaften:
   (a) $1 \in M$
   (b) $n \in M \Rightarrow F(n) \in M$

   Dann enthält M die ganze Menge $\mathbb{N}$.

Die Zahl $F(n)$ heißt der Nachfolger von n und entspricht zu $n+1$. Die dritte Axiom ist genau das Induktionsprinzip.

Man definiert die Addition und Multiplikation von natürlichen Zahlen per Induktion wie folgt:

1. $n + 1 = F(n)$
2. $n + F(m) = F(n + m)$
3. $n \cdot 1 = n$
4. $n \cdot F(m) = (n \cdot m) + m.$

Die Ungleichheit ist wie folgt definiert: $n < m$ genau dann, wenn $m = n + k$ für ein $k \in \mathbb{N}$. Dann werden die üblichen Eigenschaften von Addition, Multiplikation und Ungleichheit bewiesen.

Angenommen, dass die Mengen $\mathbb{N}$ und $\mathbb{Z}$ schon bekannt sind, so gibt es verschiedene Möglichkeiten die ganze Menge $\mathbb{R}$ aufzubauen.

Eine Möglichkeit ist das q-adische Zahlensystem zu benutzen (zum Beispiel, mit $q = 2$) um $\mathbb{R}_+$ als die Menge von q-adischen Zahlen (Folgen)

$$(a_n a_{n-1} ... a_0, b_1 b_2 ...)_q$$

zu definieren, wobei a_k, b_l die q-adischen Ziffern sind. Man muß sowohl die Operationen $+$ und $\cdot$ als auch die Relation $\leq$ mit Hilfe von q-adischen Darstellungen definieren und die Gültigkeit von allen Axiomen von reellen Zahlen beweisen.

Eine andere Möglichkeit ist zunächst die Menge $\mathbb{Q}$ von rationalen Zahlen zu definieren. Man definiert die Brüche $\frac{p}{q}$ als Paaren (p, q) von ganzen Zahlen mit $q \neq 0$. Die zwei Brüche $\frac{p}{q}$ und $\frac{p'}{q'}$ heißen äquivalent, falls $pq' = qp'$. Die Menge von Äquivalenzklassen von Brüchen bezeichnet man mit $\mathbb{Q}$, und die Elementen von $\mathbb{Q}$ heißen die rationalen Zahlen. Jede Zahl $n \in \mathbb{Z}$ kann auch als Element von $\mathbb{Q}$ betrachten werden mit Hilfe der Zuordnung $n \mapsto \frac{n}{1}$.

Man definiert die Summe und das Produkt von Brüchen mit

1. $\frac{a}{b} + \frac{c}{d} = \frac{ad+cb}{bd}$
2. $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$

Auch definiert man die Ungleichheit $\leq$ auf $\mathbb{Q}$: für positive b und d gilt $\frac{a}{b} \leq \frac{c}{d}$ genau dass, wenn $ad \leq bc$. Für negativen Werten von den Nennern gibt es eine offensichtliche Modifikation.

Man zeigt, dass die Operationen $+, \cdot$ und die Relation $\leq$ auch für Äquivalenzklassen wohldefiniert sind und mit den Operationen $+, \cdot$ bzw mit der Relation $\leq$ auf $\mathbb{Z}$ kompatibel sind. Die so definierte Menge $\mathbb{Q}$ ist ein angeordneter Körper. Das einzige Axiom von $\mathbb{R}$, das noch fehlt, ist das Vollständigkeitsaxiom.

Im nächsten Schritt erstellt man die Menge $\mathbb{R}$ aus $\mathbb{Q}$ wie folgt. Eine Teilmenge S von $\mathbb{Q}$ heißt *Dedekindscher Schnitt* falls

1. S ist nicht leer und ist nach oben beschränkt;
2. ist $x \in S$, so ist auch $y \in S$ für alle $y < x$;
3. S hat kein maximales Element.

Zu jedem $a \in \mathbb{Q}$ entspricht ein Schnitt

$$S_a = \{x \in \mathbb{Q} : x < a\}.$$

Aber es gibt die Schnitte, die keiner rationalen Zahl entsprechen, zum Beispiel

$$S = \left\{x \in \mathbb{Q} : x \leq 0 \quad \vee \quad x^2 < 2\right\}$$

(hätten wir $\sqrt{2}$ schon definiert, so könnten wir diesen Schnitt als

$$S = \left\{x \in \mathbb{Q} : x < \sqrt{2}\right\}$$

darstellen).

Die Dedekindschen Schnitten von $\mathbb{Q}$ heißen die reellen Zahlen und die Menge von den reellen Zahlen wird mit $\mathbb{R}$ bezeichnen. Die Zuordnung $a \mapsto S_a$ erlaubt uns die rationalen Zahlen als Elementen von $\mathbb{R}$ identifizieren.

Man definiert die Operationen $+, \cdot$ und die Relation $\leq$ auf $\mathbb{R}$. Zum Beispiel, für Schnitte S und T ist die Summe $S + T$ der folgende Schnitt:

$$S + T = \{x + y : x \in S, y \in T\}.$$

Die Ungleichung $S \leq T$ gilt genau dann, wenn

$$\forall x \in S \ \ \exists y \in T \ \ x \leq y.$$

Man beweist, dass alle Axiome von $\mathbb{R}$ erfüllt sind, insbesondere das Vollständigkeitsaxiom.

Alle Wege von $\mathbb{N}$ zu $\mathbb{R}$ sind ziemlich lang (und auch langweilig) und wir werden keine weiteren Einzelheiten geben.

Zur Eindeutigkeit von $\mathbb{R}$ soll man folgendes bemerken. Es gibt verschiedene Mengen, die die Axiome von $\mathbb{R}$ erfüllen. Alle solchen Mengen heißen die *Modelle* von $\mathbb{R}$. Wir haben schon zwei Modellen von $\mathbb{R}$ gesehen: die Menge von q-adischen Brüchen[1] und die Menge von Dedekindschen Schnitten. Die Eindeutigkeit von $\mathbb{R}$ gilt im folgenden Sinn. Gegeben seien zwei Modellen $\mathbb{R}'$ und $\mathbb{R}''$ von $\mathbb{R}$, dann existiert eine bijektive Abbildung $\varphi : \mathbb{R}' \to \mathbb{R}''$ mit Eigenschaften:

1. $\varphi(x + y) = \varphi(x) + \varphi(y)$
2. $\varphi(xy) = \varphi(x)\varphi(y)$
3. $x \leq y \Leftrightarrow \varphi(x) \leq \varphi(y)$.

Daraus folgt, dass auch $\varphi(0') = 0''$, $\varphi(1') = 1''$, $\varphi(-x) = -\varphi(x)$ und $\varphi(x^{-1}) = \varphi(x)^{-1}$ (siehe Aufgaben). Eine Abbildung φ mit diesen Eigenschaften heißt *Isomorphismus*, und die Modellen $\mathbb{R}'$ und $\mathbb{R}''$, für die ein Isomorphismus existiert, heißen *isomorph.* Somit sind je zwei beliebigen Modellen von reellen Zahlen isomorph.

Wir müssen unbedingt betonen, dass der Beweis der Existenz von Isomorphismus das Vollständigkeitsaxiom benötigt. Zum Beispiel, $\mathbb{Q}$ erfüllt alle anderen Axiome von reellen Zahlen, aber $\mathbb{Q}$ und $\mathbb{R}$ sind nicht isomorph.

[1] Für verschiedene Werte von q erhält man allerdings verschiedene Modelle von $\mathbb{R}$.

2.4 Überdeckungssatz

Intervallschachtelungsprinzip. Eine Folge $\{I_k\}_{k\in\mathbb{N}}$ von Intervallen heißt *Intervallschachtelung* falls $I_{k+1} \subset I_k$ für alle $k \in \mathbb{N}$.

Lemma 2.11 (Intervallschachtelungsprinzip) *Sei $\{I_k\}_{k=1}^{\infty}$ eine Intervallschachtelung wobei alle Intervalle I_k abgeschlossen und beschränkt sind. Dann ist der Durchschnitt $\bigcap_{k\in\mathbb{N}} I_k$ nicht leer.*

Bemerkung. Es ist wichtig, dass alle Intervalle abgeschlossen sind. Betrachten wir die folgende Intervallschachtelung: $I_k = (0, \frac{1}{k}]$. Wir behaupten, dass der Durchschnitt $\bigcap_{k\in\mathbb{N}} I_k$ leer ist: gilt $x \in I_k$ für alle $k \in \mathbb{N}$, so gilt $0 < x < \frac{1}{k}$ und somit $x^{-1} \geq k$ für alle $k \in \mathbb{N}$, was im Widerspruch zum Archimedischen Axiom steht. Auch ist die Beschränktheit der Intervalle wichtig: die unbeschränkten Intervalle $I_k = [k, +\infty)$ haben den leeren Durchschnitt.

Beweis. Sei $I_k = [a_k, b_k]$ mit $a_k \leq b_k$. Die Bedingung $I_{k+1} \subset I_k$ ist äquivalent zu

$$a_k \leq a_{k+1} \text{ und } b_k \geq b_{k+1}.$$

d.h. die Folge $\{a_k\}$ monoton steigend ist und $\{b_k\}$ – monoton fallend. Nach dem Satz 2.4 haben wir

$$a := \lim a_k = \sup\{a_k\}$$

und

$$b := \lim b_k = \inf\{b_k\}.$$

Die Bedingung $a_k \leq b_k$ impliziert nach dem Satz 2.1′ dass $a \leq b$. Deshalb gilt für alle k

$$a_k \leq a \leq b \leq b_k,$$

was impliziert, dass das Intervall $[a, b]$ im Durchschnitt $\bigcap_{k\in\mathbb{N}} I_k$ liegt und somit der Durchschnitt nicht leer ist. ■

Bemerkung. Man kann leicht zeigen, dass $[a, b] = \bigcap_{k\in\mathbb{N}} I_k$.

Überdeckungssatz. Betrachten wir eine Familie $\{I_\alpha\}_{\alpha\in S}$ von Intervallen, wobei S eine Indexmenge ist. Gilt für eine Teilmenge M von $\mathbb{R}$

$$M \subset \bigcup_{\alpha\in S} I_\alpha,$$

so nennt man die Familie $\{I_\alpha\}_{\alpha\in S}$ eine *Überdeckung* von M.

Ist T eine Teilmenge von S, so betrachten wir auch die Teilfamilie $\{I_\alpha\}_{\alpha\in T}$. Ist $\{I_\alpha\}_{\alpha\in T}$ auch eine Überdeckung von M, so heißt sie *Teilüberdeckung* von $\{I_\alpha\}_{\alpha\in S}$.

Hauptsatz 2.12 (Satz von Heine-Borel[2], Überdeckungssatz) *Sei $\{I_\alpha\}_{\alpha\in S}$ eine Familie von offenen Intervallen. Für jede Überdeckung $\{I_\alpha\}_{\alpha\in S}$ von einem abgeschlossenen beschränkten Intervall $[a, b]$ mit $a, b \in \mathbb{R}$, $a \leq b$, existiert eine endliche Teilüberdeckung $\{I_\alpha\}_{\alpha\in T}$ von $[a, b]$.*

[2] Auch als "Satz von Borel-Lebesgue" genannt.

Die Endlichkeit von $\{I_\alpha\}_{\alpha\in T}$ bedeutet, dass T eine endliche Menge ist.

Definition. Eine Menge $M \subset \mathbb{R}$ heißt *kompakt* falls jede Überdeckung von M mit offenen Intervallen eine endliche Teilüberdeckung enthält.

Satz 2.21 bedeutet, dass jedes beschränktes abgeschlossenes Intervall $[a, b]$ kompakt ist.

Bemerkung. Der Begriff von Kompaktheit ist einer von zentralen Begriffen in der Mathematik. Wir vertiefen und benutzen ihn später. Jetzt zeigen wir die Beispiele von nichtkompakten Mengen.

Beispiel. Zeigen wir, dass jede unbeschränkte Menge $M \subset \mathbb{R}$ nicht kompakt ist. Dafür betrachten wir die offenen Intervalle $I_k = (-k, k)$ mit $k \in \mathbb{N}$. Offensichtlich ist die Familie $\{I_k\}_{k\in\mathbb{N}}$ eine Überdeckung von M. Soll $\{I_k\}_{k\in T}$ eine endliche Teilfamilie sein, so existiert das Maximum von T, $m = \max T$. Dann gilt $I_k \subset I_m$ für alle $k \in T$ und somit

$$\bigcup_{k\in T} I_k = I_m \not\supset M.$$

Deshalb enthält $\{I_k\}_{k\in\mathbb{N}}$ keine endliche Teilüberdeckung und somit ist M nicht kompakt.

Zeigen wir, dass halboffenes Intervall $M = (0, 1]$ nicht kompakt ist. Dafür betrachten wir die offenen Intervalle

$$I_k = (\frac{1}{k}, 2), \quad k \in \mathbb{N}$$

und beachten, dass $\{I_k\}_{k=1}^\infty$ eine Überdeckung von $(0, 1]$ ist, die keine endliche Teilüberdeckung hat. In der Tat sei T eine endliche Teilmenge der Indexmenge $\mathbb{N}$. Wie in dem obigen Argument setzen wir $m = \max T$ und erhalten, dass

$$\bigcup_{k\in T} I_k = I_m,$$

während ein einziges Intervall I_m keine Überdeckung von $(0, 1]$ ist.

Beweis von Satz 2.12. Nehmen wir das Gegenteil an, dass eine Überdeckung $\{I_\alpha\}_{\alpha\in S}$ keine endliche Teilüberdeckung von $[a, b]$ zulässt. Dann bestimmen wir eine Folge $\{[a_n, b_n]\}_{n=0}^\infty$, $a_n \le b_n$ von Intervallen mit der folgenden Eigenschaften:

(i) jedes Intervall $[a_n, b_n]$ lässt keine endliche Teilüberdeckung von $\{I_\alpha\}_{\alpha\in S}$ zu;

(ii) $[a_0, b_0] = [a, b]$, $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ und

$$b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}. \tag{2.23}$$

Wir definieren $[a_n, b_n]$ per Induktion nach n. Für Induktionsanfang setzen wir $a_0 = a$ und $b_0 = b$.

Für Induktionsschritt nehmen wir an, dass $[a_n, b_n]$ für ein n schon definiert, und bestimmen $[a_{n+1}, b_{n+1}]$ wie folgt. Setzen wir $c = \frac{a_n+b_n}{2}$ und bemerken, dass eines von den Intervallen $[a_n, c]$, $[c, b_n]$ keine endliche Teilüberdeckung zulässt (sonst lässt auch $[a_n, b_n]$ eine endliche Teilüberdeckung). Wählen wir dieses Intervall und bezeichnen es mit $[a_{n+1}, b_{n+1}]$. Offensichtlich gelten (i) und (ii).

Nach Konstruktion ist die Folge $\{[a_n, b_n]\}_{n=0}^{\infty}$ Intervallschachtelung. Nach Lemma 2.11 existiert eine Zahl x, die in allen Intervallen $[a_n, b_n]$ liegt. Da $x \in [a, b]$ und $\{I_\alpha\}_{\alpha \in S}$ eine Überdeckung von $[a, b]$ ist, so gehört x zu einem Intervall I_α mit $\alpha \in S$.

Sei $I_\alpha = (p, q)$ mit $p < q$ und somit $p < x < q$. Wie behaupten, dass es ein $n \in \mathbb{N}$ gibt mit

$$[a_n, b_n] \subset (p, q) . \tag{2.24}$$

Es folgt aus (2.23) dass

$$b_n - a_n = \frac{b - a}{2^n}$$

und somit

$$b_n - a_n \to 0 \text{ für } n \to \infty.$$

Da $a_n \le x \le b_n$, daraus folgt, dass

$$\lim a_n = \lim b_n = x.$$

Somit liegen a_n und b_n in (p, q) für fast alle n, woraus (2.24) folgt. Es bleibt nur zu bemerken, dass die Inklusion 2.24 im Widerspruch zur Bedingung steht, dass $[a_n, b_n]$ keine endliche Teilüberdeckung zulässt. ■

2.5 Existenz des Grenzwertes

2.5.1 Teilfolge

Definition. Sei $\{x_n\}_{n=1}^{\infty}$ eine Folge von Elementen von $\overline{\mathbb{R}}$ und $\{n_k\}_{k=1}^{\infty}$ be eine Folge von natürlichen Zahlen mit $n_k < n_{k+1}$ für alle $k \in \mathbb{N}$ (d.h. die Folge $\{n_k\}$ ist *streng monoton steigend*). Dann die Folge $\{x_{n_k}\}_{k=1}^{\infty}$ heißt eine *Teilfolge* von $\{x_n\}_{n=1}^{\infty}$.

Erinnern wir uns, dass eine Folge $\{x_k\}$ definiert wurde als eine Abbildung $x : \mathbb{N} \to \mathbb{R}$, und die Folge $\{n_k\}$ ist auch eine Abbildung $n : \mathbb{N} \to \mathbb{N}$. Dann ist $\{x_{n_k}\}$ die zusammengesetzte Abbildung $x \circ n$ mit

$$(x \circ n)(k) = x(n(k)) = x_{n_k}.$$

Beispiel. Sei $n_k = 2k$. Dann $x_{n_k} = x_{2k}$ und die Teilfolge $\{x_{n_k}\}$ besteht aus alle Gliedern der Folge $\{x_n\}$ mit geraden n.

Behauptung. *Gilt $x_n \to a \in \overline{\mathbb{R}}$, so gilt auch $x_{n_k} \to a$ für jede Teilfolge $\{n_k\}$.*

Beweis. Nach Definition von der Konvergenz $x_n \to a$, jede Umgebung $U(a)$ enthält fast alle x_n, woraus folgt, dass auch fast alle x_{n_k} in $U(a)$ liegen, d.h. $x_{n_k} \to a$. ■

Definition. Ein $a \in \overline{\mathbb{R}}$ heißt *Häufungspunkt* der Folge $\{x_n\}$ falls es eine Teilfolge $\{x_{n_k}\}$ gibt mit $x_{n_k} \to a$.

Es folgt aus der obigen Behauptung, dass der Grenzwert immer ein Häufungspunkt ist.

Definition. Ein $a \in \overline{\mathbb{R}}$ heißt *Verdichtungspunkt* der Folge $\{x_n\}$ falls jede Umgebung $U(a)$ von a unendlich viele Glieder der Folge $\{x_n\}$ enthält.

Insbesondere ist der Grenzwert immer ein Verdichtungspunkt. Die Umkehrung davon gilt nicht: ein Verdichtungspunkt soll nicht der Grenzwert sein.

Beispiel. Die Folge $x_n = (-1)^n$ divergiert, obwohl diese Folge zwei Häufungspunkte hat: 1 und -1, da $x_{2k} \to 1$ und $x_{2k+1} \to -1$. Die beiden Werten 1 und -1 sind auch Verdichtungspunkte der Folge $\{(-1)^n\}$.

Satz 2.13 *Ein $a \in \overline{\mathbb{R}}$ ist ein Häufungspunkt der Folge $\{x_n\}$ genau dann, wenn a ein Verdichtungspunkt ist.*

Beweis. Sei a ein Häufungspunkt, d.h. $x_{n_k} \to a$ für eine Teilfolge $\{x_{n_k}\}$. Dann enthält jede Umgebung $U(a)$ fast alle Glieder der Folge $\{x_{n_k}\}$. Folglich enthält $U(a)$ unendlich viele Glieder von $\{x_n\}$, was zu beweisen war.

Sei $a \in \mathbb{R}$ ein Verdichtungspunkt. Definieren wir per Induktion in k die Teilfolge $\{x_{n_k}\}_{k=1}^{\infty}$ mit $x_{n_k} \to a$. Wählen wir erst n_1 so dass $x_{n_1} \in U_1(a)$. Ist n_{k-1} schon definiert, so wählen wir $n_k > n_{k-1}$ derart, dass $x_{n_k} \in U_{1/k}(a)$. Das ist immer möglich, da $U_{1/k}(a)$ unendlich viele Glieder enthält, insbesondere die Glieder x_n mit $n > n_{k-1}$. Dann gilt für alle $k \in \mathbb{N}$

$$|x_{n_k} - a| < \frac{1}{k}.$$

Da $\frac{1}{k} \to 0$ für $k \to \infty$, es folgt, dass $|x_{n_k} - a| \to 0$ und somit $x_{n_k} \to a$.

Im Fall $a = \pm$ erfolgt der Beweis analog: man soll nur die Umgebungen $U_k(a)$ statt $U_{1/k}(a)$ betrachten. ■

Hauptsatz 2.14 (Satz von Bolzano-Weierstraß) *Jede beschränkte Folge von reellen Zahlen besitzt eine konvergente Teilfolge.*

Beweis. Sei $\{x_n\}$ eine beschränkte Folge. Dann existieren $a, b \in \mathbb{R}$ derart, dass $a \leq x_n \leq b$, d.h. $x_n \in [a, b]$ für alle $n \in \mathbb{N}$. Nach Satz 2.13 reicht es zu zeigen, dass es einen Verdichtungspunkt der Folge $\{x_n\}$ gibt.

Nehmen wir das Gegenteil an, dass kein $c \in \mathbb{R}$ Verdichtungspunkt ist. Die Negation der Definition des Verdichtungspunktes ergibt folgendes:

$$\forall c \in \mathbb{R} \ \ \exists \varepsilon > 0 \ \text{ so dass } U_\varepsilon(c) \ \text{ nur endlich viele Glieder von } \{x_n\} \text{ enthält.}$$

Bezeichnen wir mit I_c dieses Intervall $U_\varepsilon(c)$ (beachten wir, dass ε von c abhängig ist). Dann überdeckt die Familie $\{I_c\}_{c \in \mathbb{R}}$ von offenen Intervallen die ganze Menge $\mathbb{R}$, insbesondere das Intervall $[a, b]$.

Nach dem Satz 2.12 von Heine-Borel ist $[a, b]$ kompakt, d.h. jede Überdeckung von $[a, b]$ mit offenen Intervallen enthält eine endliche Teilüberdeckung. Daraus folgt, dass es eine endliche Menge $\{I_{c_k}\}_{k=1}^{N}$ von Intervallen gibt, die $[a, b]$ überdecken. Da jedes I_{c_k} nur endliche Menge von Glieder $\{x_n\}$ enthält, erhalten wir, das auch $[a, b]$ nur endliche Menge von Glieder $\{x_n\}$ enthält, was im Widerspruch zur Definition der Folge steht. ■

Korollar 2.15 *Jede Folge von Elementen von Elementen von $\overline{\mathbb{R}}$ besitzt eine Teilfolge, die einen Grenzwert in $\overline{\mathbb{R}}$ hat. Insbesondere ist die Menge von Häufungspunkten bzw Verdichtungspunkten in $\overline{\mathbb{R}}$ nicht leer.*

Beweis. Für beschränkte Folge folgt dies aus dem Satz 2.14, für unbeschränkte Folge siehe Aufgaben. ■

2.5.2 Cauchy-Folge

Definition. Eine Folge $\{x_n\}$ von reellen Zahlen heißt *Cauchy-Folge* (oder *Fundamentalfolge*) falls

$$\forall\ \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geq N \text{ gilt } |x_n - x_m| < \varepsilon. \tag{2.25}$$

Die Bedeutung (2.25) heißt die *Cauchy-Bedingung.* Sie lässt sich äquivalent wie folgt umformulieren:

$$\forall \varepsilon > 0 \text{ für fast alle } n \text{ und fast alle } m \text{ gilt } |x_n - x_m| < \varepsilon. \tag{2.26}$$

Kurz bezeichnet man die Bedingung (2.25) bzw (2.26) wie folgt:

$$x_n - x_m \to 0 \text{ für } n, m \to \infty.$$

Hauptsatz 2.16 (Cauchy-Kriterium von Konvergenz) *Eine Folge $\{x_n\}$ von reellen Zahlen konvergiert genau dann, wenn $\{x_n\}$ eine Cauchy-Folge ist.*

Soll man die Konvergenz einer Folge $\{x_n\}$ beweisen, ohne den Grenzwert zu kennen, so hat die Cauchy-Bedingung einen Vorteil, da sie den Wert des Grenzwertes nicht explizit benutzt.

Beweis. Konvergenz $\Longrightarrow$ Cauchy-Bedingung. Sei $x_n \to a$, d.h.

$$\forall \varepsilon > 0 \text{ für fast alle } n \text{ gilt } |x_n - a| < \varepsilon.$$

Daraus folgt, dass für fast alle n, m

$$|x_n - x_m| = |(x_n - a) - (x_m - a)| \leq |x_n - a| + |x_m - a| < 2\varepsilon.$$

Da 2ε beliebig positiv ist, erhalten wir, dass $\{x_n\}$ die Cauchy-Bedingung erfüllt.

Cauchy-Bedingung $\Longrightarrow$ Konvergenz. Zeigen wir zunächst, dass die Cauchy-Folge $\{x_n\}$ beschränkt ist. Nach Definition für fast alle n, m gilt $|x_n - x_m| \leq 1$. Wählen einen solchen Wert von m. Dann gilt $x_n \in [x_m - 1, x_m + 1]$ für fast alle n, woraus folgt, dass die Folge $\{x_n\}$ beschränkt ist.

Nach dem Satz 2.14 von Bolzano-Weierstraß existiert eine konvergente Teilfolge $\{x_{n_k}\}$. Setzen wir $a = \lim x_{n_k}$ and beweisen wir, dass auch $\lim_{n\to\infty} x_n = a$. Die Konvergenz $x_{n_k} \to a$ bedeutet, dass

$$\forall \varepsilon > 0 \text{ für fast alle } k \text{ gilt } |x_{n_k} - a| < \varepsilon. \tag{2.27}$$

Es folgt aus (2.26), dass für fast alle n und k gilt

$$|x_n - x_{n_k}| < \varepsilon.$$

Zusammen mit (2.27) erhalten wir für fast alle n

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

woraus die Konvergenz $x_n \to a$ folgt. $\blacksquare$

Beispiel. Betrachten wir den unendlichen *Kettenbruch*:

$$\frac{1}{1+\frac{1}{1+\frac{1}{1+\ldots}}} \tag{2.28}$$

deren Wert wie folgt definiert ist. Für jedes $n \in \mathbb{Z}_+$ setzen wir

$$x_n = \frac{1}{1+\frac{1}{1+\frac{1}{1+\ldots\frac{1}{1}}}}$$

wobei die rechte Seite ein endlicher Kettenbruch ist, der n mal das Zeichen '+' enthält. Alternativ kann die Folge $\{x_n\}$ induktiv definiert werden wie folgt: $x_0 = 1$ und

$$x_{n+1} = \frac{1}{1+x_n} \quad \text{für } n \in \mathbb{Z}_+.$$

Man kann zeigen, dass die Folge $\{x_n\}$ die Cauchy-Bedingung erfüllt und somit konvergent ist (siehe Aufgaben). Der Grenzwert $x = \lim x_n$ heißt der Wert des Kettenbruches (2.28). Man kann x explizit bestimmen.

2.6 Limes superior und Limes inferior

Bezeichnen wir mit H die Menge von allen Häufungspunkten in $\overline{\mathbb{R}}$ einer Folge $\{x_n\}$ mit $x_n \in \overline{\mathbb{R}}$. Nach Korollar 2.15 ist die Menge H stets nicht leer. Nach Korollar 1.11 existieren das Supremum $\sup H \in (-\infty, +\infty]$ und das Infimum $\inf H \in [-\infty, +\infty)$.

Definition. Für jede Folge $\{x_n\}$ von Elementen von $\overline{\mathbb{R}}$ definieren wir den *Limes superior* durch

$$\limsup_{n\to\infty} x_n := \sup H$$

und den *Limes inferior* durch

$$\liminf_{n\to\infty} x_n := \inf H.$$

Alternative Bezeichnungen:

$$\limsup_{n\to\infty} x_n = \varlimsup_{n\to\infty} x_n \quad \text{and} \quad \liminf_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n.$$

Satz 2.17 *Sei $\{x_n\}$ eine Folge von Elementen von $\overline{\mathbb{R}}$.*

(*a*) *Die folgenden Identitäten gelten:*

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup_{k\geq n} \{x_k\} \quad \textit{and} \quad \liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf_{k\geq n} \{x_k\}. \tag{2.29}$$

(*b*) *Die beiden Werte* $\limsup x_n$ *und* $\liminf x_n$ *sind Häufungspunkte der Folge* $\{x_n\}$. *Folglich ist* $\limsup x_n$ *der maximale Häufungspunkt der Folge* $\{x_n\}$ *und* $\liminf x_n$ *der minimale Häufungspunkt von* $\{x_n\}$.

Die Folge $\{x_k\}_{k=n}^{\infty}$ heißt *Restfolge* von $\{x_k\}_{k=1}^{\infty}$.

Beweis. Setzen wir

$$y_n = \sup_{k \geq n} \{x_k\} = \sup \{x_n, x_{n+1}, ...\} .$$

Wir müssen zeigen, dass dass $\lim y_n$ existiert und gleich $\sup H$ ist (Punkt (a)) und dass $\lim y_n \in H$, was ergibt $\limsup x_n \in H$ (Punkt (b)). Die Aussagen mit lim inf werden analog bewiesen.

Zunächst vergleichen wir y_n mit $y_{n+1} = \sup \{x_{n+1}, x_{n+2}, ...\}$. Da

$$\{x_{n+1}, x_{n+2}, ...\} \subset \{x_n, x_{n+1}, ..\} ,$$

so gilt für die Suprema die Ungleichung

$$y_{n+1} \leq y_n.$$

Daher ist die Folge $\{y_n\}$ monoton fallend, und somit hat nach Satz 2.4 einen Grenzwert $a \in \overline{\mathbb{R}}$.

Zunächst zeigen wir, dass a eine obere Schranke von H ist. Sei $b \in H$, d.h. b ein Häufungspunkt von $\{x_n\}$ ist. Deshalb ist b auch Häufungspunkt der Restfolge $\{x_n, x_{n+1}, ...\}$. Da alle Glieder dieser Folge mit y_n beschränkt sind, so erhalten wir, dass auch $b \leq y_n$. Die Ungleichung $b \leq y_n$ impliziert für $n \to \infty$, dass $b \leq \lim y_n = a$. Somit $b \leq a$ und a eine obere Schranke von H ist.

Jetzt zeigen wir, dass $a \in H$, d.h. a ein Häufungspunkt von $\{x_n\}$ ist. Es reicht zu zeigen, dass jede Umgebung U von a unendlich viele Glieder der Folge $\{x_n\}$ enthält, d.h. für jede Umgebung U von a und für jedes $N \in \mathbb{N}$ existiert $n \geq N$ mit $x_n \in U$. In der Tat enthält U fast alle y_n, insbesondere existiert $m \geq N$ so dass $y_m \in U$. Es existiert eine Umgebung U' von y_m mit $U' \subset U$. Da $y_m = \sup_{n \geq m} \{x_n\}$, so existiert ein $n \geq m$ mit $x_n \in U'$. Somit für dieses n haben wir $n \geq N$ und $x_n \in U$, was zu beweisen war.

Da $a \in H$ und a eine obere Schranke von H ist, so erhalten wir $a = \sup H = \max H$, was zu beweisen war. ∎

Es folgt aus (2.29) that

$$\inf_{n \in \mathbb{N}} \{x_n\} \leq \liminf_{n \to \infty} x_n \leq \limsup_{n \to \infty} x_n \leq \sup_{n \in \mathbb{N}} \{x_n\} .$$

Auch $\lim x_n$ existiert genau dann, wenn

$$\liminf_{n \to \infty} x_n = \limsup_{n \to \infty} x_n$$

(siehe Aufgaben).

2.7 Komplexwertige Folgen

Definition. Eine Folge $\{z_n\}_{n=1}^{\infty}$ von komplexen Zahlen konvergiert gegen $a \in \mathbb{C}$ falls $|z_n - a| \to 0$ für $n \to \infty$. Man schreibt: $\lim z_n = a$ oder $z_n \to a$ für $n \to \infty$.

Definition. Eine komplexwertige Folge $\{z_n\}$ heißt Cauchy-Folge falls $|z_n - z_m| \to 0$ für $n, m \to \infty$.

Offensichtlich stimmen diese Definitionen mit den entsprechenden Definitionen für reellwertige Folgen überein.

Satz 2.18 *Sei $\{z_n\}$ eine komplexwertige Folge.*

(*a*) *Die Konvergenz $z_n \to a$ mit einem $a \in \mathbb{C}$ gilt genau dann, wenn $\operatorname{Re} z_n \to \operatorname{Re} a$ und $\operatorname{Im} z_n \to \operatorname{Im} a$.*

(*b*) *Die Folge $\{z_n\}$ ist Cauchy-Folge genau dann, wenn $\{\operatorname{Re} z_n\}$ und $\{\operatorname{Im} z_n\}$ Cauchy-Folgen sind.*

(*c*) *Die Folge $\{z_n\}$ ist konvergent genau dann, wenn $\{z_n\}$ Cauchy-Folge ist.*

Beweis. (*a*) Seien $z_n = x_n + iy_n$ und $a = x + iy$ wobei $x_n, y_n, x, y \in \mathbb{R}$. Dann

$$z_n - a = (x_n - x) + i\,(y_n - y)$$

und

$$|z_n - a|^2 = |x_n - x|^2 + |y_n - y|^2\,,$$

woraus folgt, dass

$$|z_n - a|^2 \to 0 \Leftrightarrow |x_n - x|^2 \to 0 \text{ und } |y_n - y|^2 \to 0$$

Bemerken wir, dass für jede Folge $\{r_n\}$ von reellen Zahlen

$$r_n^2 \to 0 \Leftrightarrow r_n \to 0$$

(siehe Aufgaben). Somit erhalten wir, dass

$$|z_n - a| \to 0 \Leftrightarrow |x_n - x| \to 0 \text{ und } |y_n - y| \to 0$$

was zu beweisen war. Teil (*b*) wird analog bewiesen.

(*c*) Nach (*a*), (*b*) und Satz 2.16 erhalten wir

$$\begin{aligned} \{z_n\} \text{ konvergiert} \quad &\Longleftrightarrow \quad \{x_n\} \text{ und } \{y_n\} \text{ konvergieren} \\ &\Longleftrightarrow \quad \{x_n\} \text{ und } \{y_n\} \text{ sind Cauchy-Folgen} \\ &\Longleftrightarrow \quad \{z_n\} \text{ ist Cauchy-Folge.} \end{aligned}$$

■

Satz 2.19 (Rechenregeln) *Seien $\{z_n\}$ und $\{w_n\}$ zwei komplexwertige konvergente Folgen mit $z_n \to a$ und $w_n \to b$. Dann gelten*

$$z_n + w_n \to a + b, \quad z_n - w_n \to a - b, \quad z_n w_n \to ab, \quad \frac{z_n}{w_n} \to \frac{a}{b}$$

wobei im letzten Teil vorausgesetzt ist, dass $w_n \neq 0$ und $b \neq 0$.

Beweis erfolgt mit Hilfe von den Sätzen 2.2 und 2.18 (siehe Aufgaben).

2.8 Komplexwertige Reihen

2.8.1 Allgemeine Konvergenzkriterien

Definition. Eine komplexwertige Reihe $\sum_{k=1}^{\infty} a_k$ heißt konvergent, falls die Folge $\{S_n\}_{n=1}^{\infty}$ von Partialsummen $S_n = \sum_{k=1}^{n} a_k$ konvergent ist. Der Summe (der Wert) der konvergenten Reihe wird wie folgt definiert:

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} S_n.$$

Nach dem Satz 2.18 ist die Reihe $\sum a_k$ konvergent genau dann, wenn die beiden reellwertigen Reihen $\sum \operatorname{Re} a_k$ und $\sum \operatorname{Im} a_k$ konvergieren, und in diesem Fall gilt

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \operatorname{Re} a_k + i \sum_{k=1}^{\infty} \operatorname{Im} a_k.$$

Für komplexwertige Reihen gelten die Rechenregeln

$$\sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n,$$

$$\sum_{n=1}^{\infty} c a_n = c \sum_{n=1}^{\infty} a_n, \quad c \in \mathbb{C},$$

vorausgesetzt, dass die Reihen in den rechten Seiten konvergent sind. Beweis folgt aus dem Satz 2.19.

Satz 2.20 (a) (Restreihe-Kriterium) *Die Reihe* $\sum_{k=1}^{\infty} a_k$ *ist konvergent genau dann, wenn* $\sum_{k=m}^{\infty} a_k$ *konvergent ist, für jedes* $m \in \mathbb{N}$.

(b)(Trivialkriterium) *Ist* $\sum_{k=1}^{\infty} a_k$ *konvergent, so gilt* $\lim a_k = 0$.

Die Reihe $\sum_{k=m}^{\infty} a_k$ heißt die *Restreihe* von $\sum_{k=1}^{\infty} a_k$.

Die Folge $\{a_k\}$ heißt *Nullfolge* falls $\lim a_k = 0$. Dann (b) ergibt folgendes: ist die Folge $\{a_k\}$ keine Nullfolge, so divergiert die Reihe $\sum_{k=1}^{\infty} a_k$.

Beweis. (a) Betrachten wir die Partialsummen $S_n = \sum_{k=1}^{n} a_k$ und $T_n = \sum_{k=m}^{n} a_k$ mit $n > m$. Dann gilt

$$S_n - T_n = \sum_{k=1}^{m-1} a_k =: C,$$

wobei die Konstante C unabhängig von n ist. Daraus folgt, dass S_n konvergiert genau dann, wenn T_n es tut.

(b) Für die Partialsummen gilt

$$S_n - S_{n-1} = \sum_{k=1}^{n} a_k - \sum_{k=1}^{n-1} a_k = a_n.$$

Ist die Reihe $\sum_{k=1}^{n} a_k$ konvergent, so ist die Folge $\{S_n\}$ konvergent. Dann konvergiert auf $\{S_{n-1}\}$ gegen gleichen Grenzwert, woraus folgt

$$\lim a_n = \lim (S_n - S_{n-1}) = \lim S_n - \lim S_{n-1} = 0.$$

∎

Beispiel. Betrachten wir die geometrische Reihe $\sum_{k=0}^{\infty} x^k$ für $x \in \mathbb{C}$. Genauso wie im reellen Fall ist diese Reihe für $|x| < 1$ konvergent, da

$$S_n = \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} \to \frac{1}{1 - x} \quad \text{für } n \to \infty,$$

weil $|x^{n+1}| = |x|^{n+1} \to 0$. Zeigen wir, dass die geometrische Reihe für $|x| \geq 1$ divergent ist. In der Tat ist in diesem Fall die Folge $\left\{x^k\right\}$ keine Nullfolge, da $\left|x^k\right| = |x|^k \geq 1$, woraus folgt, dass die Reihe $\sum_{k=1}^{\infty} x^k$ divergiert.

2.8.2 Majorantenkriterium und absolute Konvergenz

Satz 2.21 (Majorantenkriterium, Vergleichskriterium) *Seien $\sum_{k=1}^{\infty} a_k$ eine komplexwertige Reihe und $\sum_{k=1}^{\infty} b_k$ eine nicht-negative konvergente Reihe. Gilt*

$$|a_k| \leq b_k \quad \textit{für alle } k \in \mathbb{N} \tag{2.30}$$

so ist $\sum_{k=1}^{\infty} a_k$ konvergent und

$$\left|\sum_{k=1}^{\infty} a_k\right| \leq \sum_{k=1}^{\infty} b_k. \tag{2.31}$$

Gilt die Bedingung (2.30), so heißt die Reihe $\sum b_k$ die *Majorante* von $\sum a_k$. Man sagt auch, dass $\sum a_k$ von $\sum b_k$ majorisiert wird.

Beweis. Bezeichnen wir $A_n = \sum_{k=1}^{n} a_k$ und $B_n = \sum_{k=1}^{n} b_k$. Dann gilt es für alle Indizen $n > m$

$$|A_n - A_m| = \left|\sum_{k=m+1}^{n} a_k\right| \leq \sum_{k=m+1}^{n} |a_k| \leq \sum_{k=m+1}^{n} b_k = B_n - B_m. \tag{2.32}$$

Die Folge $\{B_n\}$ konvergiert und somit ist eine Cauchy-Folge, d.h. $B_n - B_m \to 0$ für $n, m \to \infty$. Daraus folgt, dass auch $|A_n - A_m| \to 0$. Somit ist $\{A_n\}$ eine Cauchy-Folge, und die Reihe $\sum a_k$ konvergiert. Die Ungleichung (2.31) folgt aus $|A_n| \leq B_n$, die analog zu (2.32) bewiesen wird. ∎

Definition. Eine komplexwertige Reihe $\sum_{k=1}^{\infty} a_k$ heißt *absolut konvergent* falls

$$\sum_{k=1}^{\infty} |a_k| < \infty.$$

Korollar 2.22 *Ist die Reihe $\sum a_k$ absolut konvergent, so ist sie konvergent. Es gilt auch*

$$\left|\sum_{k=1}^{\infty} a_k\right| \le \sum_{k=1}^{\infty} |a_k|\,. \tag{2.33}$$

Beweis. Setzen wir $b_k = |a_k|$. Die nicht-negative Folge $\sum b_k$ konvergiert nach Voraussetzung und ist Majorante von $\sum a_k$. Nach Satz 2.21 ist $\sum a_k$ konvergent und (2.33) gilt. ■

Beispiel. Betrachten wir die Reihe $\sum_{k=1}^{\infty} \frac{c_k}{k^2}$ wobei $\{c_k\}$ eine beliebige beschränkte Folge von komplexen Zahlen (zum Beispiel, $c_k = i^k$ oder $c_k = (-1)^k$). Wir behaupten, dass diese Reihe absolut konvergiert. Sei C eine obere Schranke von $\{|c_k|\}$. Da

$$\left|\frac{c_k}{k^2}\right| \le \frac{C}{k^2}$$

und die Reihe $\sum \frac{C}{k^2} = C \sum \frac{1}{k^2}$ konvergent ist, erhalten wir nach Satz 2.21 dass $\sum \frac{c_k}{k^2}$ absolut konvergent ist.

Es gibt die Reihen, die konvergent aber nicht absolut konvergent sind, zum Beispiel, die Leibniz-Reihe $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ (siehe Aufgaben).

Für die absolut konvergenten Reihen gelten die Kommutativ- und Assoziativgesetze, die wir ohne Beweis angeben.

Satz. *Sei $\sum a_k$ eine absolut konvergente Reihe von komplexen Zahlen. Sei $\sum b_k$ eine Reihe, die aus $\sum a_k$ durch Vertauschung und Gruppierung der Glieder erhalten wird. Dann ist $\sum b_k$ auch absolut konvergent und*

$$\sum b_k = \sum a_k.$$

Für die nicht absolut konvergente Reihen gilt im Gegenteil folgendes.

Satz. *Sei $\sum a_k$ eine konvergente Reihe von reellen Zahlen, die nicht absolut konvergent ist. Dann für jedes $c \in \overline{\mathbb{R}}$ existiert eine Reihe $\sum b_k$, die aus $\sum a_k$ durch Vertauschung der Glieder erhalten wird und so dass $\sum b_k = c$.*

Zum Beispiel, man kann die Glieder in der Reihe $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ so vertauschen, dass die vertauschte Reihe gegen $+\infty$ divergiert (siehe Aufgaben).

2.8.3 Quotientenkriterium

Satz 2.23 (Quotientenkriterium, d'Alembert-Kriterium) *Sei $\{a_n\}_{n=1}^{\infty}$ eine komplexwertige Folge mit $a_n \ne 0$ für alle n. Gilt*

$$\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1,$$

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

$$\liminf_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1,$$

so ist die Reihe $\sum_{n=1}^{\infty} a_n$ divergent.

Beweis. Bezeichnen wir $r_n = \left|\frac{a_{n+1}}{a_n}\right|$ und $r = \limsup_{n\to\infty} r_n$. Da $r < 1$, so existiert ein $q \in (r, 1)$, d.h. $r < q < 1$. Wir behaupten, dass $r_n < q$ für fast alle n. Gibt es in $[q, +\infty)$ unendlich viele Glieder der Folge $\{r_n\}$, d.h. eine Teilfolge von $\{r_n\}$, dann existiert in $[q, +\infty)$ ein Häufungspunkt der Folge, was nicht möglich ist, da der maximale Häufungspunkt gleich r ist (siehe Satz 2.17).

Deshalb existiert ein $N \in \mathbb{N}$ so dass $r_n < q$ für alle $n \geq N$. Es folgt, dass

$$|a_{n+1}| \leq q\,|a_n| \text{ für alle } n \geq N.$$

Per Induktion erhalten wir, dass

$$|a_n| \leq q^{n-N}\,|a_N| \quad \text{für alle } n \geq N.$$

Da $0 < q < 1$, so erhalten wir

$$\sum_{n=N}^{\infty} |a_n| \leq |a_N| \sum_{n=N}^{\infty} q^{n-N} = |a_N| \sum_{k=0}^{\infty} q^k < \infty,$$

d.h. die Restreihe $\sum_{n=N}^{\infty} |a_n|$ konvergent ist. Somit ist die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent.

Setzen wir $s = \liminf r_n$. Gilt $s > 1$, so gilt $r_n > 1$ for fast alle n, insbesondere für all $n \geq N$ für ein $N \in \mathbb{N}$. Es folgt, dass für all $n \geq N$

$$|a_{n+1}| > |a_n|\,,$$

so dass $a_n \to 0$ unmöglich ist. Nach dem Trivialkriterium ist die Reihe $\sum a_n$ divergent. ■

2.9 Exponentialfunktion und die Zahl e

Definition. Die Reihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ heißt die *Exponentialreihe* und ihre Summe heißt die Exponentialfunktion von $x \in \mathbb{C}$ und wird mit

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{2.34}$$

bezeichnet.

Beweisen wir, dass die Exponentialreihe absolut konvergent für alle $x \in \mathbb{C}$ ist. Setzen wir $a_n = \frac{x^n}{n!}$ und erhalten

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

Somit gilt $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 0$ und die Reihe $\sum a_n$ ist absolut konvergent nach dem Quotientenkriterium von Satz 2.23.

Ist x reell, dann ist $\exp(x)$ auch reell, was man aus (2.34) sieht. Die Zahl $\exp(1)$ wird auch mit e bezeichnet, so dass

$$e=\sum_{n=0}^{\infty}\frac{1}{n!}=1+\frac{1}{1!}+\frac{1}{2!}+...+\frac{1}{n!}+... \,.$$

Man kann berechnen, dass

$$e=2,718281828459045...$$

Es ist bekannt, dass e eine transzendente Zahl ist.

Bemerkung. Eine andere äquivalente Definition von e ist

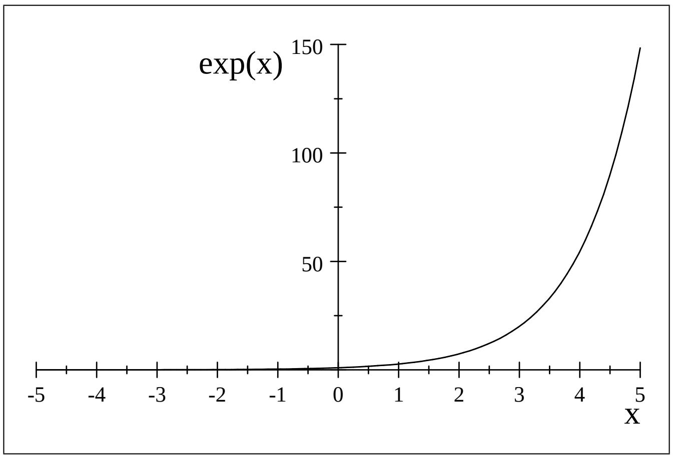
$$e=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

Man kann auch beweisen, dass

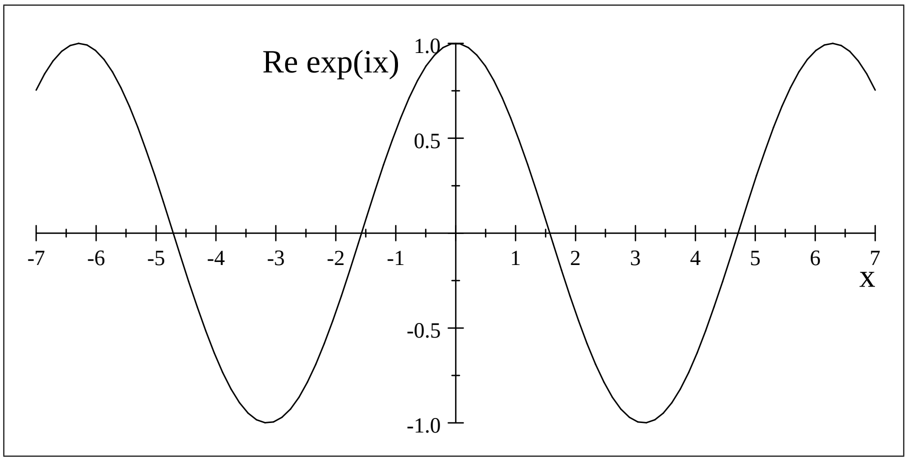
$$\exp(x)=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$$

(siehe Aufgaben).

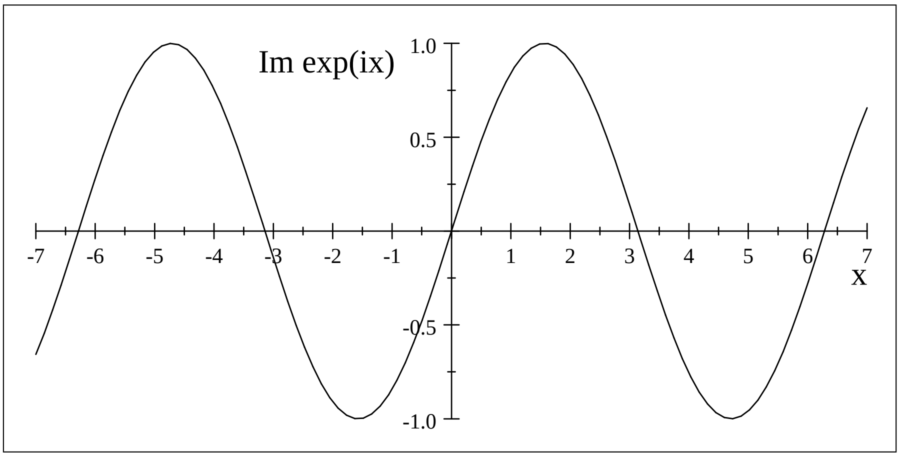
Der Graph der Funktion $f(x)=\exp(x)$ für $x\in\mathbb{R}$ sieht wir folgt aus:



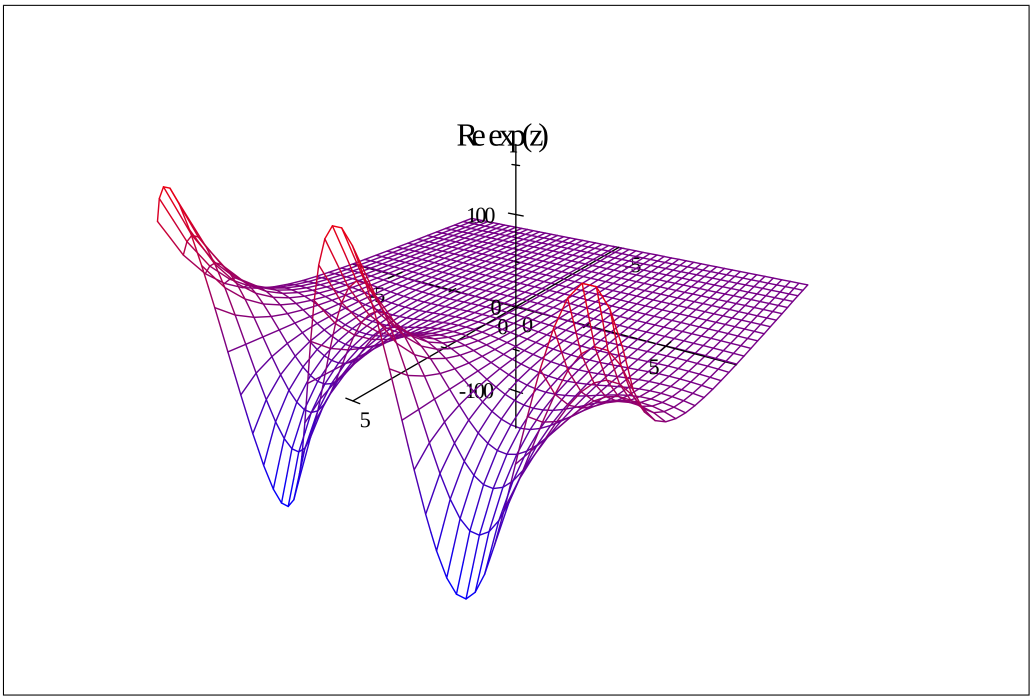
Die Funktion $f(x)=\exp(ix)$ ist komplexwertig sogar für $x\in\mathbb{R}$. Der Graph des Realteiles $\operatorname{Re}\exp(ix)$ ist wie folgt:



Der Graph des Imaginärteiles $\operatorname{Im}\exp(ix)$ ist wie folgt:



Der Graph der Funktion $\operatorname{Re}\exp(z)$ für $z \in \mathbb{C}$ ist wie folgt:



2.10 Cauchy-Produkt zweier Reihen

Erinnern wir uns, dass für das Produkt zweier endlichen Summen die folgende Identität gilt:

$$\left(\sum_{k\in I} a_k\right)\left(\sum_{l\in J} b_l\right) = \sum_{(k,l)\in I\times J} a_k b_l \tag{2.35}$$

wobei I, J die endlichen Indexmengen sind und a_k, b_l reelle oder komplexe Zahlen sind. Wir möchten eine ähnliche Identität für Produkt zweier Reihen erhalten.

Versucht man die Formel (2.35) für unendlichen Indexmengen I und J zu benutzen, so sieht man auf der rechten Seite eine Doppelreihe, was nicht bequem ist. Der Zweck der nächsten Definition ist die Doppelreihe aus (2.35) als eine übliche Reihe darzustellen.

Definition. Gegeben seien zwei komplexwertigen Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$. Das *Cauchy-Produkt* dieser Folgen ist die Folge $\sum_{n=0}^{\infty} c_n$ wobei

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Um die Bedeutung von c_n zu verstehen, betrachten wir die folgende unendliche Tabelle mit allen Summanden $a_k b_l$ für $k, l \in \mathbb{Z}_+$:

$$\begin{array}{ccccccccc}
a_0b_0 & a_0b_1 & a_0b_2 & \dots & \dots & \dots & \dots & \boxed{a_0b_n} & \dots \\
a_1b_0 & a_1b_1 & a_1b_2 & \dots & \dots & \dots & \boxed{a_1b_{n-1}} & \dots & \\
a_2b_0 & a_2b_1 & a_2b_2 & \dots & \dots & \dots & \dots & & \\
\dots & \dots & \dots & \dots & \boxed{a_kb_{n-k}} & \dots & & & \\
\dots & \dots & \dots & \dots & \dots & & & & \\
\dots & \dots & \dots & \dots & & & & & \\
\dots & \boxed{a_{n-1}b_1} & \dots & & & & & \dots & \dots \\
\boxed{a_nb_0} & \dots & & & & & \dots & a_nb_n & \dots \\
\dots & & & & & & \dots & \dots & \dots
\end{array} \tag{2.36}$$

Die Summanden der Form $a_k b_{n-k}$ mit einem vorgegebenen Wert von n liegen auf der n-ten Diagonale der Tabelle (die Einträge in den Rahmen). Deshalb ist c_n gleich die Summe aller Summanden auf der n-ten Diagonale. Die Reihe $\sum_{n=0}^{\infty} c_n$ "enthält" somit alle Summanden $a_k b_l$ aus der Tabelle, und man kann hoffen, dass

$$\sum_{n=0}^{\infty} c_n = \sum_{k,l \in \mathbb{Z}_+} a_k b_l = \sum_{k=0}^{\infty} a_k \sum_{l=0}^{\infty} b_l. \tag{2.37}$$

Aber diese Identitäten gelten nicht immer. Zum Beispiel, die Reihen $\sum a_k$ und $\sum b_k$ mit $a_k = b_k = \frac{(-1)^k}{\sqrt{k+1}}$ sind konvergent, aber ihr Cauchy-Produkt ist divergent (siehe Aufgaben). Im nächsten Satz werden die Bedingungen vorgelegt, die die Identität (2.37) garantieren.

Satz 2.24 (Satz von Mertens) *Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergente Reihen von komplexen Zahlen. Dann ist ihr Cauchy-Produkt $\sum_{n=0}^{\infty} c_n$ auch absolut konvergent und*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{k=0}^{\infty} a_k \right) \left(\sum_{l=0}^{\infty} b_l \right). \tag{2.38}$$

Bemerkung. Man kann die folgende Erweiterung des Satzes 2.24 beweisen: falls die Reihen $\sum a_k$ und $\sum b_k$ konvergent sind und mindestens eine davon absolut konvergent ist, dann ist das Cauchy-Produkt $\sum c_n$ konvergent und erfüllt (2.38). Allerdings ist die Konvergenz von $\sum c_n$ in diesem Fall nicht unbedingt absolut (siehe Aufgaben).

Beweis. Setzen wir

$$A_n = \sum_{k=0}^{n} a_k, \quad B_n = \sum_{l=0}^{n} b_l, \quad C_n = \sum_{m=0}^{n} c_m.$$

Da

$$\sum_{k=0}^{\infty} a_k \sum_{l=0}^{\infty} b_l = \lim A_n \lim B_n = \lim (A_n B_n),$$

müssen wir insbesondere zeigen, dass die Folge $\{C_n\}$ konvergiert und

$$\lim_{n\to\infty} C_n = \lim_{n\to\infty} (A_n B_n) \ . \tag{2.39}$$

Bemerken wir, dass

$$C_n = \sum_{m=0}^{n} c_m = \sum_{m=0}^{n} \sum_{k=0}^{m} a_k b_{m-k} = \sum_{\{k,l\in\mathbb{Z}_+ : k+l\le n\}} a_k b_l,$$

d.h. C_n ist die Summe von allen Gliedern $a_k b_l$ aus dem Dreieck

$$\{k, l \in \mathbb{Z}_+ : k + l \le n\}.$$

Für $A_n B_n$ erhalten wir

$$A_n B_n = \sum_{k=0}^{n} a_k \sum_{l=0}^{n} b_l = \sum_{\{k,l\in\mathbb{Z}_+ : k\le n, l\le n\}} a_k b_l,$$

d.h. $A_n B_n$ ist die Summe von allen Gliedern $a_k b_l$ aus dem Quadrat

$$\{k, l \in \mathbb{Z}_+ : k \le n, l \le n\}.$$

Betrachten wir zunächst der Fall, wenn alle a_k und b_l nicht-negative reelle Zahlen sind. Für Indizen $k, l \in \mathbb{Z}_+$ haben wir die Implikation

$$k + l \le n \ \Rightarrow \ k \le n, l \le n$$

und für $m = [n/2]$,

$$k \le m, l \le m \ \Rightarrow \ k + l \le n.$$

Auf der Tabelle (2.36) bedeutet es, dass das Dreieck $\{k + l \le n\}$ zwischen zwei Rechtecken $\{k, l \le m\}$ und $\{k, l \le n\}$ liegt. Daraus folgt, dass

$$A_m B_m \le C_n \le A_n B_n.$$

Da die beiden Folgen $\{A_n B_n\}$ und $\left\{A_{[n/2]} B_{[n/2]}\right\}$ den gleichen Grenzwert haben, erhalten wir nach Satz 2.1, dass C_n den gleichen Grenzwert hat, d.h. (2.39). Folglich ist die Reihe $\sum_{n=0}^{\infty} c_n$ konvergent. Da $c_n \ge 0$, ist diese Reihe auch absolut konvergent.

Betrachten wir jetzt den allgemeinen Fall von komplexwertigen a_n und b_n. Nach Voraussetzung sind die Reihen $\sum |a_k|$ und $\sum |b_l|$ konvergent. Sei $\sum_{n=0}^{\infty} c_n^*$ das Cauchy-Produkt der Reihen $\sum |a_k|$ und $\sum |b_l|$. Nach dem ersten Teil des Beweises ist die Reihe $\sum c_n^*$ konvergent. Darüber hinaus haben wir

$$|c_n| = \left|\sum_{k=0}^{n} a_k b_{n-k}\right| \le \sum_{k=0}^{n} |a_k| \, |b_{n-k}| = c_n^*.$$

Nach Majorantenkriterium von Satz 2.21 erhalten wir die absolute Konvergenz von $\sum c_n$.

Es bleibt noch die Identität (2.39) zu beweisen, d.h. $A_nB_n - C_n \to 0$. Wir haben

$$A_nB_n - C_n = \sum_{\{k,l:k\leq n,l\leq n\}} a_kb_l - \sum_{\{k,l:k+l\leq n\}} a_kb_l = \sum_{\{k,l:k\leq n.l\leq n.k+l>n\}} a_kb_l.$$

Nach der Dreiecksungleichung gilt

$$|A_nB_n - C_n| \leq \sum_{\{k,l:k\leq n.l\leq n.k+l>n\}} |a_k|\,|b_l|\,.$$

Bezeichnen wir mit A_n^*, B_n^*, C_n^* die Partialsummen der entsprechenden Reihen $\sum |a_k|$, $\sum |b_l|$, $\sum c_m^*$. Dann gilt auch

$$A_n^*B_n^* - C_n^* = \sum_{\{k,l:k\leq n.l\leq n.k+l>n\}} |a_k|\,|b_l|$$

woraus folgt

$$|A_nB_n - C_n| \leq |A_n^*B_n^* - C_n^*|\,.$$

Da nach dem ersten Teil des Beweises gilt $|A_n^*B_n^* - C_n^*| \to 0$, erhalten wir $|A_nB_n - C_n| \to 0$ und somit $A_nB_n - C_n \to 0$, was zu beweisen war. ■

2.11 Eigenschaften der Exponentialfunktion

Satz 2.25 *Für alle $x, y \in \mathbb{C}$ gilt die Identität*

$$\exp(x+y) = \exp(x)\exp(y)\,. \tag{2.40}$$

Beweis. Betrachten wir zwei Reihen

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

und sei $\sum_{n=0}^{\infty} c_n$ das Cauchy-Produkt dieser Reihen, d.h.

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

wo wir den Binomischen Lehrsatz benutzt haben (siehe Aufgaben). Es folgt, dass

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

Nach Satz 2.24 erhalten wir

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

woraus (2.40) folgt. ■

Satz 2.26 *(a) Für jedes $x \in \mathbb{C}$ gilt $\exp(x) \neq 0$.*

(b) Für jedes $x \in \mathbb{R}$ ist $\exp(x)$ reell und positiv.

(c) Für reelle $x > y$ gilt $\exp(x) > \exp(y)$ (d.h. die Funktion $\exp(x)$ ist streng monoton steigend)

(d) Für jedes $k \in \mathbb{Z}$ gilt $\exp(k) = e^k$ wobei $e = \exp(1)$.

Beweis. (a) Die Identität (2.40) ergibt für $y = -x$,

$$\exp(x)\exp(-x) = \exp(0) = 1, \tag{2.41}$$

woraus $\exp(x) \neq 0$ folgt.

(b) Ist $x \in \mathbb{R}$ dann gilt $\exp(x) \in \mathbb{R}$ nach Definition

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{2.42}$$

Ist $x \geq 0$ so ist $\exp(x) > 0$ da alle Glieder in (2.42) nicht negative sind und das Glied mit $k = 0$ gleich 1 ist. Ist $x < 0$ dann $\exp(-x) > 0$ was zusammen mit (2.41) ergibt

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

(c) Wir haben

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

da für $t = x - y > 0$ gilt

$$\exp(t) = 1 + t + \frac{t^2}{2!} + ... > 1.$$

(d) Zunächst beweisen wir per Induktion nach k, dass die Identität

$$\exp(k) = e^k \tag{2.43}$$

für alle $k \in \mathbb{N}$ gilt. Für $k = 1$ gilt (2.43) nach Definition von e. Induktionsschritt: gilt (2.43) für ein $k \in \mathbb{N}$, dann erhalten wir nach (2.40)

$$\exp(k+1) = \exp(k)\exp(1) = e^k e = e^{k+1}.$$

Für $k = 0$ haben wir $\exp(0) = 1 = e^0$. Für negative $k \in \mathbb{Z}$ erhalten wir nach (2.41)

$$\exp(k) = \frac{1}{\exp(-k)} = \frac{1}{e^{-k}} = e^k.$$

■

Die Identität (2.43) motiviert die folgende Definition von der Potenz e^x für alle komplexen Zahlen x:

$$e^x := \exp(x).$$

Zum Beispiel haben wir $e^{1/2} = \sqrt{e}$ da $e^{1/2}e^{1/2} = e^{1/2+1/2} = e$ und somit erfüllt $e^{1/2}$ die Definition von $\sqrt{e}$.

Später definieren wir die Potenz a^x für alle positiven Werten von a and $x \in \mathbb{C}$.

Chapter 3

Funktionen einer reellen Variablen

In diesem Kapitel studieren wir die reellwertigen Funktionen einer reellen Variablen, d.h. die Funktionen $f: J \to \mathbb{R}$ wobei der Definitionsbereich J eine Teilmenge von $\mathbb{R}$ ist. Wir kennen schon viele Beispiele von solchen Funktionen:

1. Exponentialfunktion $f(x) = \exp(x)$ mit dem Definitionsbereich $J = \mathbb{R}$;
2. Potenzfunktion $f(x) = x^n$ mit dem Definitionsbereich $J = \mathbb{R}$ und dem Exponenten $n \in \mathbb{N}$;
3. Potenzfunktion $f(x) = x^n$ mit dem Definitionsbereich $J = (0, +\infty)$ und dem Exponenten $n \in \mathbb{Z}$;
4. Quadratwurzel $f(x) = \sqrt{x}$ mit dem Definitionsbereich $J = [0, +\infty)$;
5. Die n-te Wurzel[1] $f(x) = \sqrt[n]{x}$ mit dem Definitionsbereich $J = [0, +\infty)$ und dem Exponenten $n \in \mathbb{N}$.

Darüber hinaus ist jede Folge $\{a_n\}_{n\in\mathbb{N}}$ auch eine Funktion $f: \mathbb{N} \to \mathbb{R}$ mit $f(n) = a_n$. Für die Folgen haben wir den Begriff des Limes

$$\lim_{\substack{n\to\infty \\ n\in\mathbb{N}}} f(n)$$

schon definiert. In diesem Kapitel fangen wir mit der Definition von Limes

$$\lim_{\substack{x\to a \\ x\in J}} f(x)$$

für allgemeine Funktion $f: J \to \mathbb{R}$ an.

3.1 Grenzwert einer Funktion

Erinnern wir uns an den Begriff der Umgebung: eine Umgebung von $a \in \mathbb{R}$ ist irgendwelches Intervall $(a - \varepsilon, a + \varepsilon)$ mit $\varepsilon > 0$, eine Umgebung von $+\infty$ ist irgendwelches

[1] Die Existenz und Eindeutigkeit von $\sqrt[n]{x}$ für $x \geq 0$ lassen sich genau so beweisen, wie die Existenz und Eindeutigkeit von $\sqrt{x}$.

Intervall $(E, +\infty]$ mit $E \in \mathbb{R}$, und eine Umgebung von $-\infty$ ist irgendwelches Intervall $[-\infty, E)$ mit $E \in \mathbb{R}$.

Definition. Sei J eine Teilmenge von $\mathbb{R}$. Eine Element $a \in \overline{\mathbb{R}}$ heißt *Verdichtungspunkt* von J falls jede Umgebung U von a unendlich viele Elemente von J enthält.

Beispiel. Für $J = \mathbb{N}$ gibt es nur einen Verdichtungspunkt $a = +\infty$. Für $J = \mathbb{Z}$ gibt es zwei Verdichtungspunkte $a = \pm\infty$. Für ein Intervall $J = (a, b)$ die Menge von Verdichtungspunkten ist das abgeschlossene Intervall $[a, b]$.

Definition. Für jede Teilmenge J von $\mathbb{R}$ definieren wir den *Abschluss* $\overline{J}$ als die Vereinigung von J mit der Menge von allen Verdichtungspunkten von J.

Somit ist $\overline{J}$ eine Teilmenge von $\overline{\mathbb{R}}$.

Behauptung. *Es gilt*

$$\overline{J} = \left\{a \in \overline{\mathbb{R}} : U(a) \cap J \neq \emptyset \text{ für jede Umgebung } U(a) \text{ von } a\right\}.$$

Beweis. Offensichtlich gilt $U(a) \cap J \neq \emptyset$ für jedes $a \in J$ und für jedes Verdichtungspunkt a von J, also für alle $a \in \overline{J}$. Für $a \notin \overline{J}$ gibt es eine Umgebung $U(a)$ mit endlich vielen Elementen aus J. Das Reduzieren dieser Umgebung ergibt eine Umgebung $U'(a)$ von a mit leerem Durchschnitt mit J. ■

Beispiel. Die Abschlüsse der folgenden Mengen kann man einfach bestimmen:

$$\overline{(a,b)} = [a,b], \quad \overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}, \quad \overline{\mathbb{Z}} = \mathbb{Z} \cup \{\pm\infty\}, \quad \overline{\mathbb{Q}} = \overline{\mathbb{R}}.$$

Definition. Sei $J \subset \mathbb{R}$ eine nicht-leere Menge und $f : J \to \mathbb{R}$ eine Funktion auf J. Sei $a \in \overline{J}$ and $b \in \overline{\mathbb{R}}$. Wir sagen, dass $f(x)$ den Grenzwert $b \in \overline{\mathbb{R}}$ hat für $x \to a$, $x \in J$ und schreiben

$$\lim_{\substack{x \to a \\ x \in J}} f(x) = b, \tag{3.1}$$

falls für jede Umgebung V von b existiert eine Umgebung U von a mit

$$x \in U \cap J \implies f(x) \in V.$$

Man kann das kurz wie folgt aufschreiben:

$$\boxed{\lim_{\substack{x \to a \\ x \in J}} f(x) = b \text{ falls } \forall V(b) \ \ \exists U(a) \ \ \forall x \in J \cap U(a) \Rightarrow f(x) \in V(b)}$$

wobei $U(a)$ eine Umgebung von a und $V(b)$ eine Umgebung von b. Man schreibt in diesem Fall auch

$$f(x) \to b \ \text{ für } \ x \to a, \ x \in J$$

und sagt: $f(x)$ konvergiert gegen b falls $b \in \mathbb{R}$, und $f(x)$ divergiert gegen b falls $b = \pm\infty$.

Beispiel. Zeigen wir, dass der Limes einer Folge ein spezielle Fall des Limes einer Funktion ist. Sei $\{f_n\}_{n=1}^{\infty}$ eine Folge, d.h. eine Abbildung $f : \mathbb{N} \to \mathbb{R}$ mit $f_n = f(n)$. Da $a = +\infty$

ein Verdichtungspunkt von $\mathbb{N}$ ist und jede Umgebung U von a ein Intervall $(E, +\infty]$ ist, so erhält man nach der obigen Definition folgendes: für jedes $b \in \mathbb{R}$ ist

$$\lim_{\substack{n\to\infty\\ n\in\mathbb{N}}} f_n = b$$

äquivalent zu

$$\forall\varepsilon > 0\ \exists E\ \ \forall n > E \quad |f_n - b| < \varepsilon,$$

d.h. zur Definition des Limes der Folge aus dem Abschnitt 2.1.1.

Bemerkung. Seien jetzt $a, b \in \mathbb{R}$. Dann erhält man die folgende äquivalente Definition des Limes (3.1):

$$\boxed{\lim_{\substack{x\to a\\ x\in J}} f(x) = b \text{ falls } \forall\varepsilon > 0\ \exists\delta > 0\ \ \forall x \in J \quad |x - a| < \delta\ \Rightarrow |f(x) - b| < \varepsilon.} \tag{3.2}$$

Beispiel. Betrachten wir die Funktion

$$f(x) = \frac{x^2 - 1}{x - 1}$$

die auf $J = \mathbb{R} \setminus \{1\}$ definiert ist. Bestimmen wir $\lim_{x\to 1} f(x)$. Da

$$f(x) = \frac{(x+1)(x-1)}{x-1} = x + 1,$$

man kann erwarten, dass

$$\lim_{\substack{x\to 1\\ x\in J}} f(x) = 2.$$

Um das zu beweisen, benutzen wir Definition (3.2):

$$f(x) - 2 = (x+1) - 2 = x - 1$$

und

$$|f(x) - 2| = |x - 1|.$$

Deshalb gilt (3.2) mit $\delta = \varepsilon$.

Beispiel. Sei $f(x) = x^2$, $x \in \mathbb{R}$. Beweisen wir, dass $f(x) \to a^2$ für $x \to a \in \mathbb{R}$. Für jedes $x \in \mathbb{R}$ haben wir

$$\left|f(x) - a^2\right| = \left|x^2 - a^2\right| = |x - a|\,|x + a|.$$

Ist $|x - a| < \delta$ so gilt $x \in (a - \delta, a + \delta)$ woraus folgt

$$|x| \le |a| + |\delta|$$

und

$$|x + a| \le 2\,|a| + |\delta|.$$

Da δ gewählt werden kann, setzen wir voraus, dass $\delta \le 1$, so dass

$$|x + a| \le 2\,|a| + 1.$$

Daraus folgt, dass

$$\left|f(x) - a^2\right| < \delta\,(2\,|a| + 1)$$

und

$$\left|f(x) - a^2\right| < \varepsilon,$$

für das vorgegebene $\varepsilon > 0$, vorausgesetzt

$$\delta \leq \frac{\varepsilon}{2\,|a| + 1}.$$

Deshalb können wir setzen $\delta = \min\left(1, \frac{\varepsilon}{2|a|+1}\right)$.

Beispiel. Sei $f(x) = \exp(x)$, $x \in \mathbb{R}$, und beweisen wir, dass $f(x) \to 1$ für $x \to 0$. Wir haben

$$|\exp(x) - 1| = \left|\sum_{k=1}^{\infty} \frac{x^k}{k!}\right| \leq \sum_{k=1}^{\infty} |x|^k = \frac{|x|}{1 - |x|},$$

für alle $|x| < 1$. Gegeben sei $\varepsilon > 0$, bestimmen wir $0 < \delta \leq 1$ mit

$$|x| < \delta \implies \frac{|x|}{1 - |x|} < \varepsilon. \tag{3.3}$$

Wir können δ weiter beschränken mit $\delta \leq \frac{1}{2}$ so that

$$\frac{|x|}{1 - |x|} \leq \frac{|x|}{1 - \delta} \leq \frac{|x|}{1/2} < 2\delta.$$

Dann (3.3) gilt vorausgesetzt $\delta \leq \varepsilon/2$. Deshalb für $\delta = \min\left(\frac{1}{2}, \frac{\varepsilon}{2}\right)$ erhalten wir aus (3.3)

$$|x| < \delta \implies |\exp(x) - 1| < \varepsilon,$$

was beweist, dass $f(x) \to 1$ für $x \to 0$.

Bemerkung. Für die unendlichen Werte von a bzw b kann man die Definition (3.1) wie folgt umschreiben.

- $a \in \mathbb{R}$ und $b = +\infty$:

$$\boxed{\lim_{\substack{x \to a \\ x \in J}} f(x) = +\infty \text{ falls } \forall E \in \mathbb{R}\ \exists \delta > 0\ \forall x \in J \quad |x - a| < \varepsilon \Rightarrow f(x) > E.}$$

- $a = +\infty$ und $b \in \mathbb{R}$:

$$\boxed{\lim_{\substack{x \to +\infty \\ x \in J}} f(x) = b \text{ falls } \forall \varepsilon > 0\ \exists D \in \mathbb{R}\ \ \forall x \in J \quad x > D \Rightarrow |f(x) - b| < \varepsilon.}$$

- $a = b = +\infty$:

$$\boxed{\lim_{\substack{x \to +\infty \\ x \in J}} f(x) = +\infty \text{ falls} \forall E \in \mathbb{R}\ \ \exists D \in \mathbb{R}\ \ \forall x \in J\ \ x > D \Rightarrow\ f(x) > E.}$$

Der Fall $a = -\infty$ oder $b = -\infty$ ist analog.

Beispiel. Betrachten wir die Funktion $f(x) = \frac{1}{|x|}$, die auf $J = \mathbb{R} \setminus \{0\}$ definiert ist, und zeigen, dass $f(x) \to +\infty$ für $x \to 0$, d.h.

$$\forall E \in \mathbb{R} \ \ \exists \delta > 0 \ \ \forall x \in \mathbb{R} \setminus \{0\} \ \ |x| < \delta \ \Rightarrow \frac{1}{|x|} > E.$$

Das ist offensichtlich erfüllt mit $\delta = \frac{1}{E}$ vorausgesetzt $E > 0$ (und δ ist beliebig falls $E < 0$).

Satz 3.1 *Sei f eine reellwertige Funktion auf einer Menge $J \subset \mathbb{R}$ und sei $a \in \overline{J}$. Dann die folgenden zwei Bedingungen sind äquivalent:*

(i) $\lim_{x \to a} f(x) = b$ *(wobei $b \in \overline{\mathbb{R}}$).*

(ii) Für jede Folge $\{x_n\} \subset J$ mit $\lim_{n \to \infty} x_n = a$ *gilt* $\lim_{n \to \infty} f(x_n) = b$.

Beweis. Betrachten wir den Fall wenn $a, b \in \mathbb{R}$.

$(i) \Rightarrow (ii)$ Die Bedingung (i) bedeutet nach Definition, dass

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \ \forall x \in J \ \ \ |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon. \tag{3.4}$$

Sei $\{x_n\} \subset J$ eine Folge mit $x_n \to a$. Dann gilt für fast alle $n \in \mathbb{N}$, dass

$$|x_n - a| < \delta,$$

woraus folgt, dass für fast alle n gilt

$$|f(x_n) - b| < \varepsilon.$$

Da $\varepsilon > 0$ beliebige ist, so erhalten wir $\lim_{n \to \infty} f(x_n) = b$.

$(ii) \Rightarrow (i)$ Nehmen wir das Gegenteil an, dass $\lim_{x \to a} f(x) = b$ nicht gilt, d.h.

$$\exists \varepsilon > 0 \ \ \forall \delta > 0 \ \ \exists x \in J \text{ mit } |x - a| < \delta \text{ aber } \ |f(x) - b| \geq \varepsilon.$$

Benutzen wir diese Bedingung für die Werte $\delta = \frac{1}{k}$ für alle $k \in \mathbb{N}$. Für jedes $\delta = \frac{1}{k}$ erhalten wir ein $x_k \in J$ mit

$$|x_k - a| < \frac{1}{k} \quad \text{und} \quad |f(x_k) - b| \geq \varepsilon.$$

Dann für $k \to \infty$ gilt $x_k \to a$, aber nicht $f(x_k) \to b$, was im Widerspruch zur Bedingung (ii) steht.

Analog betrachtet man den Fall wenn mindestens eines von a, b unendlich ist. ■

Für zwei beliebige Funktionen $f, g : J \to \mathbb{R}$ definieren wir die arithmetischen Operationen wie folgt:

$$\begin{aligned} (f \pm g)(x) &= f(x) \pm g(x) \\ (fg)(x) &= f(x)\, g(x) \\ \frac{f}{g}(x) &= \frac{f(x)}{g(x)}, \end{aligned}$$

wobei im Fall der Division vorausgesetzt ist, dass $g \neq 0$ in J.

Satz 3.2 *Seien f, g reellwertige Funktionen mit dem Definitionsbereich $J \subset \mathbb{R}$ und sei $c \in \overline{J}$. Angenommen seien die Bedingungen*

$$\lim_{x \to c} f(x) = a \quad \text{und} \quad \lim_{x \to c} g(x) = b,$$

wobei $a, b \in \overline{\mathbb{R}}$.

(a) Es gelten die Identitäten

$$\lim_{x \to c} (f+g) = a+b, \quad \lim_{x \to c} (fg) = ab, \quad \lim_{x \to c} \frac{f}{g} = \frac{a}{b} \tag{3.5}$$

vorausgesetzt, dass die Ausdrücke $a+b$, ab, $\frac{a}{b}$ bestimmt sind (und $g \neq 0$ im Fall der Division).

(b) Gilt $f(x) \leq g(x)$ für alle $x \in J$, so gilt auch $a \leq b$.

(c) Seien f, g, h drei Funktionen auf J mit $f \leq h \leq g$ und

$$\lim_{x \to c} f(x) = \lim_{x \to c} g(x) = a.$$

Dann gilt auch

$$\lim_{x \to c} h(x) = a.$$

Beweis. Nach Satz 3.1 gelten für jede Folge $x_n \to x$, $x_n \in J$ die Bedingungen $f(x_n) \to a$ und $g(x_n) \to b$. Nach Satz 2.3 erhalten wir

$$(f+g)(x_n) \to a+b, \qquad (fg)(x_n) \to ab, \qquad \frac{f}{g}(x_n) \to \frac{a}{b},$$

woraus (3.5) folgt. Analog beweist man (b) und (c) mit Hilfe von Satz 2.1. ∎

3.2 Zusammengesetzte Funktion

Satz 3.3 (Grenzwert der Komposition) *Seien A, B Teilmengen von $\mathbb{R}$ und $f : A \to \mathbb{R}$, $g : B \to \mathbb{R}$ zwei Funktionen Angenommen seien*

$$\lim_{\substack{x \to a \\ x \in A}} f(x) = b \quad \text{und} \quad \lim_{\substack{y \to b \\ y \in B}} g(y) = c,$$

wobei $a \in \overline{A}$, $b \in \overline{B}$, $c \in \overline{\mathbb{R}}$. Ist die zusammengesetzte Funktion $x \mapsto g(f(x))$ auf A definiert (d.h. $f(A) \subset B$), so gilt

$$\lim_{\substack{x \to a \\ x \in A}} g(f(x)) = c. \tag{3.6}$$

Man kann die Identität (3.6) wie folgt umschreiben:

$$\lim_{\substack{x \to a \\ x \in A}} g(f(x)) = \lim_{\substack{y \to b \\ y \in B}} g(y)$$

und betrachten sie, als einen Wechsel von Variable $y = f(x)$ im Limes.

Beweis. Für jede Umgebung W von c existiert eine Umgebung V von b mit

$$y \in V \cap B \implies g(y) \in W. \tag{3.7}$$

Gegeben ist die Umgebung V von b, existiert eine Umgebung U von a mit

$$x \in U \cap A \implies f(x) \in V.$$

Da $f(A) \subset B$, haben wir auch $f(x) \in B$ für alle $x \in A$, so dass

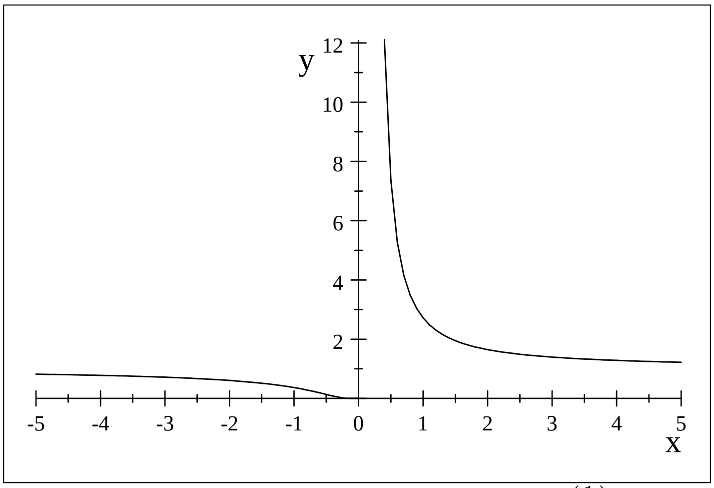
$$x \in U \cap A \implies f(x) \in V \cap B.$$

Setzen wir $y = f(x)$ in (3.7) und erhalten, dass

$$x \in U \cap A \implies g(f(x)) \in W,$$

woraus (3.6) folgt. ■

Beispiel. Betrachten wir die Funktion $f(x) = \exp\left(\frac{1}{x}\right)$ und bestimmen die Grenzwerte davon für $x \to 0$ in zwei Mengen: $x > 0$ und $x < 0$.



Der Graph der Funktion $x \mapsto \exp\left(\frac{1}{x}\right)$

Da

$$\lim_{\substack{x\to 0\\ x>0}} \frac{1}{x} = +\infty,$$

so erhalten wir nach dem Wechsel $y = \frac{1}{x}$, dass

$$\lim_{\substack{x\to 0\\ x>0}} \exp\left(\frac{1}{x}\right) = \lim_{y\to+\infty} \exp(y) = +\infty. \tag{3.8}$$

Da

$$\lim_{\substack{x\to 0\\ x<0}} \frac{1}{x} = -\infty,$$

so erhalten wir

$$\begin{aligned} \lim_{\substack{x\to 0\\ x<0}} \exp\left(\frac{1}{x}\right) &= \lim_{y\to-\infty} \exp(y) \\ &= \lim_{y\to-\infty} \frac{1}{\exp(-y)} = \lim_{z\to+\infty} \frac{1}{\exp(z)} = \frac{1}{+\infty} = 0. \end{aligned}$$

3.3 Stetige Funktionen

Definition. Sei f eine reellwertige Funktion auf einer Menge $J \subset \mathbb{R}$. Die Funktion f heißt *stetig* an der Stelle $a \in J$ falls

$$\lim_{\substack{x \to a \\ x \in J}} f(x) = f(a). \tag{3.9}$$

Sonst ist f *unstetig* an a.

Ist f stetig an alle $a \in J$, so heißt f stetig auf J (oder einfach stetig).

Nach Satz 3.1 ist die Stetigkeit von f an a äquivalent zur folgenden Bedingung: für jede Folge $\{x_n\} \subset J$,

$$x_n \to a \;\Rightarrow f(x_n) \to f(a).$$

Bemerkung Die folgenden Bedingungen sind auch äquivalent zu (3.9):

$$\lim_{\substack{x \to a \\ x \in J \setminus \{a\}}} f(x) = f(a)$$

und

$$\lim_{\substack{x \to a \\ x \in J}} f(x) \text{ existiert.}$$

Beispiel. Es ist offensichtlich, dass die folgenden Funktionen $f(x) = \text{const}$ und $f(x) = x$ stetig auf $\mathbb{R}$ sind.

Zeigen wir, dass die Funktion $f(x) = \exp(x)$ auch stetig auf $\mathbb{R}$ ist. Wir haben schon gesehen, dass

$$\lim_{x \to 0} \exp(x) = 1 = \exp(0),$$

so dass $\exp(x)$ stetig an $a = 0$ ist. Für beliebiges $a \in \mathbb{R}$ haben wir

$$\exp(x) = \exp(x - a + a) = \exp(a)\exp(x - a).$$

Da

$$\lim_{x \to a} \exp(x - a) = \lim_{y \to 0} \exp(y) = 1,$$

es folgt, dass

$$\lim_{x \to a} \exp(x) = \exp(a).$$

Satz 3.4 *Seien f, g zwei reellwertige Funktionen auf einer Menge $J \subset \mathbb{R}$. Sind f und g stetig an der Stelle $a \in J$, so sind auch die Funktionen $f + g$, fg, f/g stetig an a (im Fall f/g vorausgesetzt $g \neq 0$).*

Sind f, g stetig auf J, so sind $f+g$, fg, f/g auch stetig auf J (im Fall f/g vorausgesetzt $g \neq 0$).

Beweis. Die erste Aussage folgt aus dem Satz 3.2. Zum Beispiel für die Summe $f + g$ erhalten wir

$$\lim_{\substack{x \to a \\ x \in J}} (f + g)(x) = \lim_{\substack{x \to a \\ x \in J}} f(x) + \lim_{\substack{x \to a \\ x \in J}} g(x) = f(a) + g(a) = (f + g)(a),$$

woraus die Stetigkeit von $f+g$ an a folgt. Analog behandelt man die Funktionen fg und f/g.

Die zweite Aussage folgt aus der ersten Aussage. ■

Beispiel. Da die Funktionen $f(x)=x$ und $g(x)=\mathrm{const}$ stetig auf $\mathbb{R}$ sind, so folgt es, dass any Funktion $h(x)=cx^n$ stetig auf $\mathbb{R}$ ist, für jedes $n\in\mathbb{N}$ und $c\in\mathbb{R}$. Daraus folgt, dass jedes *Polynom*

$$P(x)=c_0+c_1x+c_2x^2+...+c_nx^n$$

stetig auf $\mathbb{R}$ ist, wobei $c_k\in\mathbb{R}$ and $n\in\mathbb{N}\cup\{0\}$ der Grad des Polynoms heißt.

Betrachten wir eine *rationale* Funktion $R(x)=\frac{P(x)}{Q(x)}$ wobei P und Q zwei Polynome sind. Dann ist R definiert und stetig im Bereich $\{Q\neq 0\}$.

Satz 3.5 *Seien $f:A\to\mathbb{R}$ und $g:B\to\mathbb{R}$ zwei Funktionen, wobei $A,B\subset\mathbb{R}$. Sei die Komposition $g\circ f$ wohldefiniert auf A (d.h. $f(A)\subset B$). Ist f stetig an $a\in A$ und g stetig an $b=f(a)$, so ist $g\circ f$ stetig an a.*

Ist f stetig auf A und g – auf B, so ist $g\circ f$ stetig auf A.

Beweis. Nach Satz 3.3 haben wir

$$\lim_{\substack{x\to a\\x\in A}} g(f(x)) = \lim_{\substack{y\to f(a)\\y\in B}} g(y) = \lim_{\substack{y\to b\\y\in B}} g(y) = g(b) = g(f(a)),$$

woraus die Stetigkeit von $g(f(x))$ an a folgt.

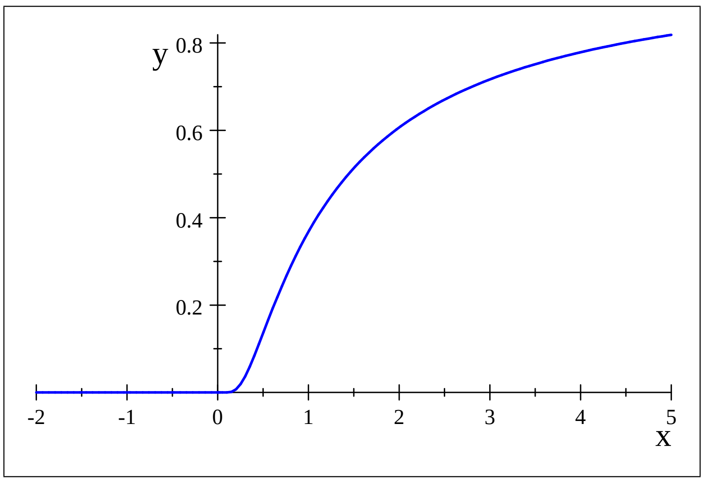
Ist f stetig auf A und g – auf B, so ist g stetig an $b=f(a)$ für jedes $a\in A$, da $f(A)\subset B$. Daraus folgt, dass $g\circ f$ stetig an a für jedes $a\in A$ ist, und somit stetig auf A. ■

Beispiel. Für jede rationale Funktion $f(x)$ ist die Funktion $\exp(f(x))$ stetig im Definitionsbereich von $f(x)$.

Beispiel. Beweisen wir, dass die Funktion

$$f(x)=\begin{cases}\exp\left(-\frac{1}{x}\right), & x>0,\\ 0, & x\leq 0,\end{cases} \tag{3.10}$$

stetig auf $\mathbb{R}$ ist.



Der Graph der Funktion (3.10)

Die Funktion $f(x) = \exp\left(-\frac{1}{x}\right)$ ist stetig im Bereich $\{x > 0\}$ nach Satz 3.5, und $f(x) = 0$ ist stetig im $\{x < 0\}$. Es bleibt zu zeigen, dass f stetig an 0 ist, d.h.

$$\lim_{\substack{x\to 0\\ x\in\mathbb{R}}} f(x) = 0. \tag{3.11}$$

Es folgt aus (3.8), dass

$$\lim_{\substack{x\to 0\\ x\neq 0}} \exp\left(-\frac{1}{x}\right) = 0,$$

d.h., für jedes $\varepsilon > 0$ existiert $\delta > 0$ mit

$$0 < x < \delta, \implies \left|\exp\left(-\frac{1}{x^2}\right)\right| < \varepsilon.$$

Da $f(x) = 0$ für $x \le 0$, es folgt, dass

$$|x| < \delta \implies |f(x)| < \varepsilon,$$

woraus (3.11) folgt.

Beispiel. Die Funktion

$$f(x) = \begin{cases} 1, & x > 0, \\ 0, & x \le 0, \end{cases}$$

ist unstetig an $x = 0$, da für die Folge $x_n = \frac{1}{n} \to 0$ gilt $f(x_n) = 1 \not\to 0 = f(0)$.

3.4 Eigenschaften von stetigen Funktionen

3.4.1 Zwischenwertsatz

Hauptsatz 3.6 (Zwischenwertsatz) *Sei $f(x)$ eine stetige Funktion auf einem geschlossenen Intervall $[a,b]$. Gilt $f(a) < 0$ und $f(b) > 0$, so existiert $x \in (a,b)$ mit $f(x) = 0$.*

Beweis. Definieren wir per Induktion in n eine Folge von Intervallen $\{[a_n,b_n]\}_{n=1}^{\infty}$, $a_n \le b_n$, mit den Eigenschaften:

1. $[a_1,b_1] = [a,b]$;
2. $[a_n,b_n] \subset [a_{n-1},b_{n-1}]$ für $n \ge 2$;
3. $f(a_n) \le 0,\ f(b_n) \ge 0$;
4. $b_n - a_n = \frac{1}{2}(b_{n-1} - a_{n-1})$ für $n \ge 2$.

Induktionsanfang ist offensichtlich. Ist $[a_n,b_n]$ schon bekannt, so betrachten wir $c = \frac{a_n+b_n}{2}$. Gilt $f(c) \le 0$, so setzen wir $[a_{n+1},b_{n+1}] = [c,b_n]$, gilt $f(c) > 0$ so setzen wir $[a_{n+1},b_{n+1}] = [a_n,c]$.

Somit ist $\{[a_n, b_n]\}_{n=1}^{\infty}$ eine Intervallschachtelung. Nach den Intervallschachtelungsprinzip gibt es ein $x \in \bigcap_{n\in\mathbb{N}} [a_n, b_n]$. Da $b_n - a_n \to 0$, so haben wir $a_n \to x$ und $b_n \to x$ für $n \to \infty$. Nach der Stetigkeit der Funktion f gelten

$$\lim_{n\to\infty} f(a_n) = f(x) = \lim_{n\to\infty} f(b_n).$$

Da $f(a_n) \le 0$ und $f(b_n) \ge 0$, es folgt, da $f(x) \le 0$ und $f(x) \ge 0$, somit $f(x) = 0$. ■

Beispiel. Sei $P(x)$ ein Polynom von Grad n mit reellen Koeffizienten, d.h.

$$P(x) = x^n + c_1 x^{n-1} + \dots + c_n,$$

where $c_k \in \mathbb{R}$ and $n \in \mathbb{N}$. Beweisen wir, dass falls n ungerade ist, dann existiert eine reelle Nullstelle von P, d.h. ein $x \in \mathbb{R}$ mit $P(x) = 0$. Nach Aufgaben haben wir

$$\lim_{x\to+\infty} P(x) = +\infty \quad \text{und} \quad \lim_{x\to-\infty} P(x) = -\infty.$$

Somit existiert ein $b \in \mathbb{R}$ mit $P(b) > 0$ und ein $a \in \mathbb{R}$ mit $P(a) < 0$. Nach Satz 3.6 existiert $x \in \mathbb{R}$ mit $P(x) = 0$.

Sei f eine reellwertige Funktion auf einer Menge A. Dann bezeichnen wir

$$\sup_A f := \sup f(A)$$

and

$$\inf_A f := \inf f(A),$$

wobei $f(A) = \{f(x) : x \in A\}$ das Bild von A ist.

Korollar 3.7 *Sei f eine stetige Funktion auf einem Intervall $I \subset \mathbb{R}$. Dann ist das Bild $f(I)$ auch ein Intervall, und die Grenzen von $f(I)$ sind $\inf_I f$ und $\sup_I f$.*

Man kann kurz sagen: stetiges Bild eines Intervalles ist ein Intervall.

Beweis. Nach Definition haben wir

$$f(I) \subset [\inf f, \sup f].$$

Es bleibt zu zeigen, dass

$$f(I) \supset (\inf f, \sup f),$$

d.h.

$$\inf f < t < \sup f \quad \Rightarrow t \in f(I).$$

Nach Definition von sup und inf, es gibt die Zahlen $a, b \in I$ mit

$$\inf f < f(a) < t < f(b) < \sup f.$$

Sei $a < b$ (der Fall $a > b$ ist analog). Betrachten wir auf dem Intervall $[a, b] \subset I$ die Funktion

$$g(x) = f(x) - t$$

die erfüllt

$$g(a) = f(a) - t < 0 \quad \text{und} \quad g(b) = f(b) - t > 0.$$

Nach Satz 3.6 existiert $x \in (a, b)$ mit $g(x) = 0$ woraus $f(x) = t$ und somit $t \in f(I)$ folgt. ■

Beispiel. Beweisen wir, dass

$$\exp(\mathbb{R}) = (0, +\infty). \tag{3.12}$$

Da

$$\lim_{x \to +\infty} \exp(x) = +\infty$$

and

$$\lim_{x \to -\infty} \exp(x) = \lim_{y \to +\infty} \exp(-y) = \lim_{y \to +\infty} \frac{1}{\exp(y)} = \frac{1}{+\infty} = 0,$$

erhalten wir

$$\sup \exp(x) = +\infty \quad \text{and} \quad \inf \exp(x) = 0.$$

Nach Korollar 3.7 ist $\exp(\mathbb{R})$ ein Intervall mit den Grenzen 0 und $+\infty$. Da nach Satz 2.26 gilt $\exp(x) > 0$ für alle $x \in \mathbb{R}$, so erhalten wir (3.12).

3.4.2 Extremwertsatz

Sei f eine reellwertige Funktion auf einer Menge A. Bezeichnen wir

$$\max_A f = \max f(A) \quad \text{und} \quad \min_A f = \min f(A),$$

vorausgesetzt, dass $\max f(A)$ bzw $\min f(A)$ existiert.

Hauptsatz 3.8 (Extremwertsatz oder Satz vom Minimum und Maximum) *Sei f eine stetige Funktion auf einem abgeschlossenen beschränkten Intervall I. Dann existieren die beiden Werten $\max_I f$ und $\min_I f$.*

Beweis. Sei $M = \sup f(I)$. Da Supremum einer Menge ein Verdichtungspunkt der Menge ist (siehe Aufgaben), so gibt es eine Folge $\{a_n\}$ von Elementen von $f(I)$ die gegen M konvergiert (oder divergiert falls $M = +\infty$). Für jedes a_n gibt es $x_n \in I$ mit $f(x_n) = a_n$, so dass $f(x_n) \to M$ für $n \to \infty$. Die Folge $\{x_n\}$ ist beschränkt und somit enthält eine konvergente Teilfolge, zum Beispiel, $x_{n_k} \to x$ mit $x \in [a, b]$. Da f stetig its, erhalten wir

$$f(x_{n_k}) \to f(x) \text{ für } k \to \infty,$$

woraus $f(x) = M$ folgt. Somit liegt M in $f(I)$ und ist das Maximum von $f(I)$. Analog beweist man die Existenz von $\min_I f$. ■

Korollar 3.9 *Sei f eine stetige Funktion auf einem abgeschlossenen beschränkten Intervall I. Dann ist das Bild $f(I)$ auch ein abgeschlossenes beschränktes Intervall. Darüber hinaus gilt*

$$f(I) = \left[\min_I f, \max_I f\right].$$

Kurz: stetiges Bild von kompaktem Intervall ist kompaktes Intervall.

Beweis. Nach Korollar 3.7 ist $f(I)$ ein Intervall mit den Grenzen A und B, wobei

$$A = \inf_I f \quad \text{und} \quad B = \sup_I f.$$

Nach Satz 3.8 haben wir $A, B \in f(I)$ und

$$A = \min_I f \quad \text{und} \quad B = \max_I f,$$

woraus $f(I) = [A, B]$ folgt. ■

3.5 Umkehrfunktion

Nach Satz 1.7, eine Abbildung $f : A \to B$ zwischen zweien beliebigen Mengen A und B hat eine Umkehrabbildung $f^{-1} : B \to A$ genau dann, wenn f bijektiv ist. Dann ist die Umkehrabbildung durch die folgenden Bedingung definiert:

$$y = f(x) \Leftrightarrow x = f^{-1}(y)$$

für alle $x \in A$ und $y \in B$. Sind A, B Teilmengen von $\mathbb{R}$, dann heißt f^{-1} auch die *Umkehrfunktion* oder die *inverse Funktion*.

Beispiel. Betrachten wir die Funktion $f : [0, +\infty) \to [0, +\infty)$, die durch $f(x) = x^2$ definiert ist. Da die Bedingung $y = x^2$ für nicht-negative x und y äquivalent zu $x = \sqrt{y}$ ist, wir sehen, dass die Umkehrfunktion f^{-1} existiert und $f^{-1}(y) = \sqrt{y}$.

Beispiel. Für die Funktion $f : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$, $f(x) = \frac{1}{x}$, die Bedingung $y = \frac{1}{x}$ ist äquivalent zu $x = \frac{1}{y}$. Deshalb f^{-1} existiert und ist gleich f.

Definition. Eine reellwertige Funktion $f(x)$ auf einem Intervall I heißt monoton steigend, falls $x \leq y$ ergibt $f(x) \leq f(y)$. Die Funktion heißt streng monoton steigend, falls $x < y$ ergibt $f(x) < f(y)$.

Analog definiert man die monoton fallende Funktionen.

Beispiel. Die Function $f(x) = \exp(x)$ ist streng monoton steigend auf $\mathbb{R}$ nach Satz 2.26.

Beispiel. Die Function $f(x) = x^n$ its streng monoton steigend auf $[0, +\infty)$ für jedes $n \in \mathbb{N}$, was man per Induktion nach n beweist.

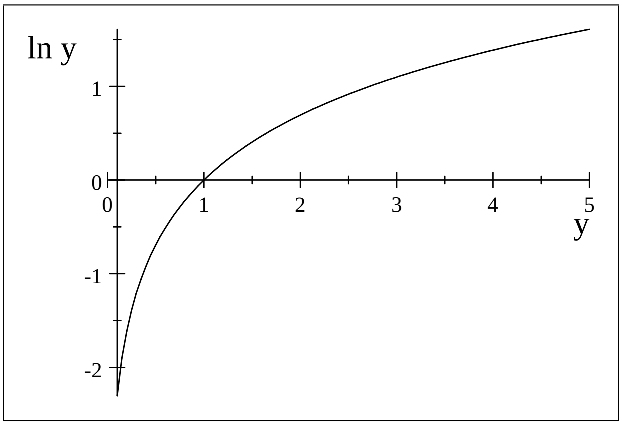
Beispiel. Die Funktion $f(x) = x^n$ ist streng monoton fallend auf $(0, +\infty)$ für jedes $n \in -\mathbb{N}$, was aus $x^n = \frac{1}{x^{|n|}}$ folgt.

Satz 3.10 *Sei f eine streng monotone Funktion auf einem Intervall I. Sei $J = f(I)$. Dann die Umkehrfunktion $f^{-1} : J \to I$ existiert, ist streng monoton und stetig.*

Beispiel. Die Funktion $\exp(x)$ ist streng monoton steigend auf $\mathbb{R}$ und $\exp(\mathbb{R}) = (0, +\infty)$. Deshalb hat sie die Umkehrfunktion auf $(0, +\infty)$, die auch stetig und monoton steigend ist. Diese Funktion heißt *natürlicher Logarithmus* und sie wird mit ln bezeichnet. Somit für jedes $y > 0$ ist $\ln y$ durch die Identität

$$y = \exp(x) \Leftrightarrow x = \ln y$$

definiert. Nach Satz 3.10 ist die Funktion $y \mapsto \ln y$ streng monoton steigend und stetig auf $(0, +\infty)$. Es gelten $\ln 1 = 0$ und $\ln e = 1$.



Der Graph der Funktion $y \mapsto \ln y$

Natürlicher Logarithmus kann benutzt werden um die Potenz a^x für alle $a > 0$ und $x \in \mathbb{R}$ (sogar $x \in \mathbb{C}$) zu definieren:

$$a^x := \exp(x \ln a) . \tag{3.13}$$

Die Funktion $f(x) = a^x$ heißt die Exponentialfunktion zur Basis a. Für $a = e$ erhalten wir $e^x = \exp(x)$, was schon früher definiert wurde.

Die Funktion a^x hat die folgenden Eigenschaften:

1. $a^1 = a$ da $\exp(\ln a) = a$, und $a^0 = 1$ da $\exp(0) = 1$.

2. $a^{x+y} = a^x a^y$ da
$$a^x a^y = \exp(x \ln a) \exp(y \ln a) = \exp((x + y) \ln a) = a^{x+y}.$$

3. Für $n \in \mathbb{N}$ stimmt die neue Definition (3.13) von a^n mit der früheren induktiven Definition (1.15) von a^n überein, da $a^{n+1} = a^n a$.

4. Für $x \in \mathbb{R}$ ist a^x reell, und die Funktion $x \mapsto a^x$ ist stetig auf $\mathbb{R}$.

5. Für alle $x, y \in \mathbb{R}$,
$$(a^x)^y = a^{xy} \tag{3.14}$$
(siehe Aufgaben).

Da $\ln e = 1$, erhalten wir nach (3.13) $e^x = \exp(x)$.

Beispiel. Die Funktion $y = x^n$ (mit $n \in \mathbb{N}$) ist streng monoton steigend auf $[0, +\infty)$ und ihr Bild ist $[0, +\infty)$. Deshalb existiert die Umkehrfunktion auf $[0, +\infty)$ und ist stetig.

Diese Funktion wird mit $x = \sqrt[n]{y}$ bezeichnet. Diese Funktion ist streng monoton steigend und stetig auf $[0, +\infty)$.

Bemerken wir, dass für jedes $y > 0$ gilt

$$\sqrt[n]{y} = y^{1/n}.$$

In der Tat ist der Wert $x = \sqrt[n]{y}$ die einzige positive Lösung der Gleichung $x^n = y$. Andererseits ist diese Gleichung auch von $x = a^{1/n}$ erfüllt, da nach (3.14) gilt $\left(a^{1/n}\right)^n = a^{\frac{1}{n}n} = a$.

Beweis von Satz 3.10. Sei f streng monoton steigend. Dann ist f eine Bijektion von I nach J, und somit existiert die Umkehrfunktion. Beweisen wir, dass f^{-1} streng monotone steigend ist. Seien $y_1 < y_2$ aus J und bezeichnen wir $x_k = f^{-1}(y_k)$ so dass $y_k = f(x_k)$. Zeigen wir, dass $x_1 < x_2$. Gilt $x_1 \geq x_2$, dann erhalten wir nach der Monotonie von f, dass $y_2 = f(x_2) \leq f(x_1) = y_1$, was im Widerspruch zum $y_1 < y_2$ steht.

Beweisen wir, dass f^{-1} stetig ist, d.h. für jedes $b \in J$ und jede Folge $\{y_n\} \subset J$ gilt

$$y_n \to b \Rightarrow f^{-1}(y_n) \to f^{-1}(b).$$

Nehmen wir das Gegenteil an, dass die Folge $x_n = f^{-1}(y_n)$ gegen $a = f^{-1}(b)$ nicht konvergiert. Dann existiert ein $\varepsilon > 0$ so dass es außerhalb $U_\varepsilon(a)$ unendlich viele Glieder der Folge $\{x_n\}$ gibt. Dann enthält eines von Intervallen $(-\infty, a-\varepsilon]$, $[a+\varepsilon, +\infty)$ unendlich viele Glieder der Folge $\{x_n\}$, sei es $(-\infty, a-\varepsilon]$. Dann existiert eine Teilfolge $\{x_{n_k}\}$ mit $x_{n_k} \leq a - \varepsilon$ für alle $k \in \mathbb{N}$ gilt. Da $x_{n_k} \leq a - \varepsilon < a$ und die beiden Werte x_{n_k}, a in I sind, so gilt auch $a - \varepsilon \in I$. Nach der Monotonie von f erhalten wir

$$y_{n_k} = f(x_{n_k}) \leq f(a - \varepsilon),$$

woraus folgt

$$\lim_{k\to\infty} y_{n_k} \leq f(a-\varepsilon) < f(a) = b,$$

was im Widerspruch zum $b = \lim y_n$ steht. ■

Bemerkung. Obwohl man die Stetigkeit von f in diesem Satz nicht braucht, ist f in Anwendungen normalerweise stetig. Daraus folgt, dass das Bild $J = f(I)$ ein Intervall ist und somit die Umkehrfunktion auf einem Intervall definiert ist.

3.6 Trigonometrische Funktionen und die Zahl π

Definition. Für jedes $x \in \mathbb{C}$ definieren wir $\sin x$ und $\cos x$ durch

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} \quad \text{und} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}, \tag{3.15}$$

wobei i die imaginäre Einheit ist.

Bemerken wir, dass sogar für reelle x benutzt die Definition (3.15) von $\sin x$ und $\cos x$ die Werte von $\exp(z)$ im komplexen Bereich.

Es folgt direkt nach (3.15), dass für alle $x \in \mathbb{C}$

$$\boxed{e^{ix} = \cos x + i \sin x}. \tag{3.16}$$

Diese Identität heißt *Eulerformel*.

Einsetzen in (3.15) die Exponentialreihe ergibt die folgenden Potenzreihen für $\sin x$ und $\cos x$, die für alle $x \in \mathbb{C}$ absolut konvergieren:

$$\boxed{\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - ... = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}} \tag{3.17}$$

und

$$\boxed{\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - ... = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}} \tag{3.18}$$

(siehe Aufgaben). Insbesondere sehen wir, dass $\sin x$ und $\cos x$ reell für alle $x \in \mathbb{R}$ sind. Es folgt aus (3.17) und (3.18), dass $\sin x$ eine ungerade Funktion ist, d.h. für all $x \in \mathbb{C}$

$$\sin(-x) = -\sin x,$$

und $\cos x$ eine gerade Funktion ist, d.h.

$$\cos(-x) = \cos x.$$

Es gilt für alle $x \in \mathbb{C}$

$$\boxed{\cos^2 x + \sin^2 x = 1}, \tag{3.19}$$

da nach (3.16)

$$\begin{aligned} \cos^2 x + \sin^2 x &= (\cos x + i \sin x)(\cos x - i \sin x) \\ &= e^{ix} e^{-ix} = e^0 = 1. \end{aligned}$$

Es gelten auch die Identitäten

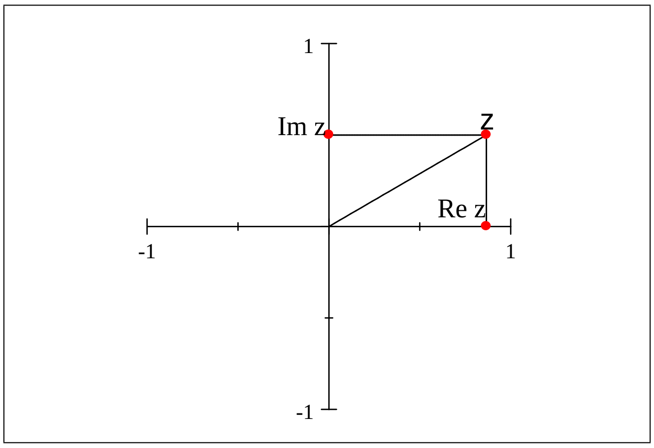
$$\begin{aligned} \cos(x+y) = \cos x \cos y - \sin x \sin y \\ \sin(x+y) = \sin x \cos y + \cos \sin y \end{aligned} \tag{3.20}$$

(siehe Aufgaben).

Betrachten wir jetzt die Funktionen $\sin x$ und $\cos x$ für $x \in \mathbb{R}$. Es folgt aus (3.16), dass für $x \in \mathbb{R}$

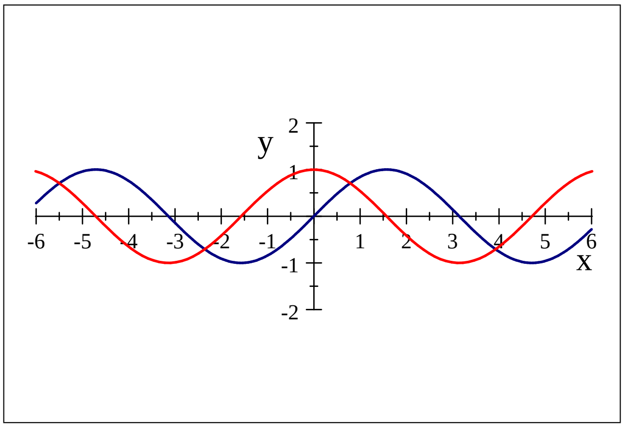
$$\operatorname{Re} \exp(ix) = \cos x \quad \text{und} \quad \operatorname{Im} \exp(ix) = \sin x.$$

In der Tat hat die Zahl $z = e^{ix}$ den Betrag 1, und ihre $\operatorname{Re} z$ und $\operatorname{Im} z$ stellen $\cos x$ und $\sin x$ dar. Der Wert von x kann als der Polarwinkel des Punktes z betrachtet werden.



Die Zahl $z = e^{ix}$ und ihre $\operatorname{Re} z, \operatorname{Im} z$

Es folgt aus (3.15), dass die Funktionen $\sin x$ und $\cos x$ stetig auf $\mathbb{R}$ sind (siehe Aufgaben).

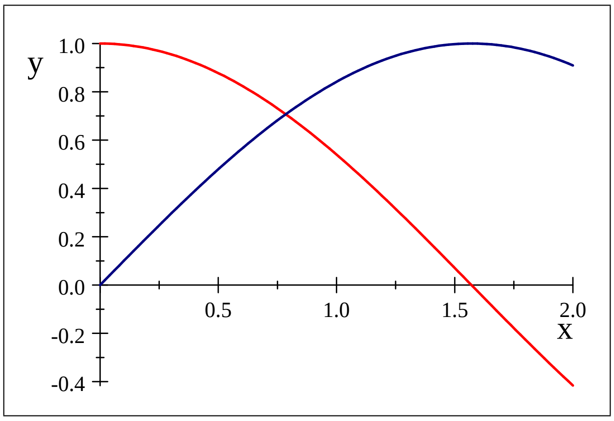


Die Graphe von Funktionen $\cos x$ (rot) und $\sin x$ (blau)

Man sieht aus den Graphen, dass die Funktionen $\sin x$ and $\cos x$ *periodisch* sind, was nicht offensichtlich aus den obigen Identitäten ist. Wir beweisen die Periodizität von $\sin x$ und $\cos x$ unterhalb.

Satz 3.11 (a) *Es existiert ein* $c \in (0, 2)$ *mit* $\cos c = 0$ *und* $\cos x > 0$ *für alle* $0 \leq x < c$.
(b) $\sin x > 0$ *für alle* $x \in (0, 2)$.

Insbesondere ist c die kleinste positive Zahl mit $\cos c = 0$. Es folgt aus Satz 3.11, dass $\sin c > 0$, was zusammen mit (3.19) ergibt $\sin c = 1$.



Die Graphe von Funktionen $\cos x$ (rot) und $\sin x$ (blau) auf dem Intervall $[0, 2]$

Für den Beweise brauchen wir das folgende Lemma.

Lemma 3.12 *(Leibniz-Kriterium) Sei* $\{c_k\}_{k=1}^{\infty}$ *eine monoton fallende Folge von nicht-negativen reellen Zahlen mit* $\lim_{k \to \infty} c_k = 0$. *Sei* $c_0 \geq 0$ *beliebig. Dann ist die Reihe* $\sum_{k=0}^{\infty} (-1)^k c_k$ *konvergent und für die Partialsummen* $S_n = \sum_{k=0}^{n} (-1)^k c_k$ *gelten die Ungleichungen*

$$S_{2m+1} \leq \sum_{k=0}^{\infty} (-1)^k c_k \leq S_{2m}$$

für alle $m \geq 0$.

Beweis. Die Konvergenz der Reihe wurde in Aufgabe 94 bewiesen, sei

$$S = \sum_{k=0}^{\infty} (-1)^k c_k = \lim_{n\to\infty} S_n.$$

Wir haben

$$\begin{aligned} S_{2m+2} - S_{2m} &= (-1)^{2m+2} c_{2m+2} + (-1)^{2m+1} c_{2m+1} \\ &= c_{2m+2} - c_{2m+1} \le 0, \end{aligned}$$

woraus folgt, dass die Folge $\{S_{2m}\}_{m=0}^{\infty}$ monoton fallend ist, woraus $S \le S_{2m}$ folgt. Analog ist die Folge $\{S_{2m+1}\}_{m=0}^{\infty}$ monotone steigend, was $S \ge S_{2m+1}$ ergibt. ■

Beweis von Satz 3.11. (a) Bemerken wir, dass $\cos 0 = 1 > 0$. Zeigen wir, dass $\cos 2 < 0$. Wir haben

$$\cos 2 = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \frac{2^6}{6!} - ... = a_0 - a_1 + a_2 - a_3 + ...$$

wobei $a_n = \frac{2^{2n}}{(2n)!}$. Die Folge $\{a_n\}_{n\ge1}$ ist monoton fallend da

$$\frac{a_{n+1}}{a_n} = \frac{2^2}{(2n+1)(2n+2)} \le 1.$$

Sei $S_n = \sum_{k=0}^{n} (-1)^k a_k$ die Partialsummen der Reihe. Nach Lemma 3.12 gilt es für jedes $m \ge 0$,

$$\cos 2 = \lim_{n\to\infty} S_n \le S_{2m}.$$

Insbesondere gilt

$$\cos 2 \le S_2 = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = -\frac{1}{3} < 0,$$

woraus $\cos 2 < 0$ folgt. Nach Satz 3.6 erhalten wir, dass $\cos x = 0$ für ein $x \in (0,2)$.

Betrachten wir die Menge

$$N = \{x \in (0,2) : \cos x = 0\},$$

die nicht-leer ist, und setzen

$$c := \inf N.$$

Zeigen wir, dass $c \in N$. Es existiert eine Folge $\{x_n\}_{n=1}^{\infty} \subset N$ mit $x_n \to c$ für $n \to \infty$, woraus folgt nach der Stetigkeit von $\cos x$, dass

$$\cos c = \lim_{n\to\infty} \cos x_n = 0,$$

und somit $c \in N$. Beachten wir, dass $c > 0$ weil $\cos 0 > 0$, und $c < 2$ weil $\cos 2 < 0$.

Zeigen wir, dass $\cos x > 0$ für $0 < x < c$. Da c eine untere Schranke von N ist, ergibt die Bedingung $x < c$ dass $x \notin N$ und somit $\cos x \ne 0$. Ist $\cos x < 0$, so erhalten wir nach Satz 3.6 die Existenz von $0 < y < x$ mit $\cos y = 0$ was nicht möglich wegen $y < c$ ist. Daher gilt $\cos x > 0$, was zu beweisen war.

(b) Wir haben

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - ... = b_0 - b_1 + b_2 - ...$$

wobei $b_n = \frac{x^{2n+1}}{(2n+1)!}$. Für $x \in (0,2)$ und jedes $n \geq 0$ haben wir

$$\frac{b_{n+1}}{b_n} = \frac{x^2}{(2n+2)(2n+3)} < \frac{4}{(2n+2)(2n+3)} < 1,$$

woraus folgt, dass die Folge $\{b_n\}_{n=0}^{\infty}$ monoton fallend ist.

Betrachten wir die Partialsummen $S_n = \sum_{k=0}^{n} (-1)^k b_k$. Nach Lemma 3.12 gilt

$$\sin x = \lim_{n\to\infty} S_n \geq S_{2m+1} \text{ für alle } m \geq 0.$$

Insbesondere gilt für $x \in (0,2)$

$$\sin x \geq S_1 = x - \frac{x^3}{6} = x\left(1 - \frac{x^2}{6}\right) > 0,$$

was zu beweisen war. ■

Definition. Setzen wir $\pi := 2c$ wobei c die kleinste positive Nullstelle von $\cos x$ ist, die nach dem Satz 3.11 existiert.

Es folgt aus Satz 3.11, dass $0 < \pi < 4$. Man kann beweisen, dass $c > \frac{3}{2}$ und somit $\pi > 3$ (siehe Aufgaben). Numerische Berechnung ergibt

$$\pi = 3.14159265358979...$$

Es ist bekannt, dass π eine transzendente Zahl ist. Derzeitig ist π mit über 6 Milliarden Dezimalstellen berechnet worden.

Satz 3.13 (a) *Es gelten die Identitäten*

$$\exp\left(\frac{\pi}{2}i\right) = i, \quad \exp(\pi i) = -1, \qquad \exp(2\pi i) = 1.$$

(b) *Die Funktion* $\exp(z)$ *ist* $2\pi i$ *periodisch auf* $\mathbb{C}$, *d.h.*

$$\exp(z + 2\pi i) = \exp(z) \ \forall z \in \mathbb{C}. \tag{3.21}$$

(c) *Die Funktionen* $\sin x$ *und* $\cos x$ *sind* 2π *periodisch, d.h.*

$$\sin(x + 2\pi) = \sin x \ \textit{ und } \ \cos(x + 2\pi) = \cos x \ \textit{ für alle } x \in \mathbb{C}.$$

Beweis. (a) Benutzen wir die Bezeichnung $c = \pi/2$ wie in Satz 3.11. Da $\cos c = 0$ und somit nach (3.19) $\sin c = 1$, erhalten wir nach der Eulerformel

$$\exp\left(\frac{\pi}{2}i\right) = \exp(ci) = \cos c + i \sin c = i.$$

Es folgt, dass

$$\exp(\pi i) = \exp(ci)\exp(ci) = i \cdot i = -1,$$

und

$$\exp(2\pi i) = \exp(\pi i)\exp(\pi i) = (-1)^2 = 1. \tag{3.22}$$

(b) Mit Hilfe von (3.22) und der Haupteigenschaft der Exponentialfunktion erhalten wir

$$\exp(z + 2\pi i) = \exp(z)\exp(2\pi i) = \exp(z).$$

(c) Mit Hilfe von (3.15) und (3.21) erhalten wir

$$\begin{aligned}\sin(x + 2\pi) &= \frac{1}{2i}\left(\exp(ix + 2\pi i) - \exp(-ix - 2\pi i)\right)\\ &= \frac{1}{2i}\left(\exp(ix) - \exp(-ix)\right) = \sin x,\end{aligned}$$

und das gleiche Argument funktioniert für $\cos x$. ■

Da $\cos\frac{\pi}{2} = 0$ und $\sin\frac{\pi}{2} = 1$, so erhalten wir aus (3.20)

$$\sin\left(\frac{\pi}{2} - x\right) = \cos x \quad\text{und}\quad \cos\left(\frac{\pi}{2} - x\right) = \sin x.$$

Es folgt aus $\exp(\pi i) = -1$, dass $\cos\pi = -1$, $\sin\pi = 0$ und somit aus (3.20)

$$\sin(\pi - x) = \sin x \quad\text{und}\quad \cos(\pi - x) = -\cos x.$$

Die anderen Identitäten für trigonometrischen Funktionen lassen sich analog beweisen. Weitere Untersuchung der trigonometrischen Funktionen führen wir unterhalb mit Hilfe von Differentialrechnung durch.

3.7 Differentialrechnung

3.7.1 Definition von Ableitung

Definition. Eine reellwertige Funktion f auf einem Intervall I heißt *differenzierbar* an der Stelle $x \in I$ falls der folgende Grenzwert existiert:

$$\lim_{\substack{y\to x\\ y\in I\setminus\{x\}}} \frac{f(y) - f(x)}{y - x}.$$

Der Wert des Grenzwertes heißt die *Ableitung* (Differentialquotient) von f an der Stelle $x \in I$ und wird mit $f'(x)$ bezeichnet.

Ist f an allen Stellen $x \in I$ differenzierbar, so sagt man, dass f auf dem Intervall I differenzierbar. In diesem Fall wird die Ableitung f' als eine Funktion auf dem Intervall I betrachtet.

Einsetzen $h = y - x$ ergibt die äquivalente Definition der Ableitung:

$$f'(x) = \lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(x+h) - f(x)}{h}. \tag{3.23}$$

Der Quotient $\frac{f(x+h)-f(x)}{h}$ heißt *Differenzenquotient*.

Beispiel. Berechnen wir die Ableitungen der folgenden Funktionen.

1.$f(x) = \text{const}$. Dann $f(y) - f(x) = 0$ und $f'(x) = 0$ für alle x. Man schreibt

$$(\text{const})' = 0.$$

2. $f(x) = x$. Dann

$$\frac{f(y) - f(x)}{y - x} = \frac{y - x}{y - x} = 1$$

woraus folgt $f'(x) = 1$ für alle $x \in \mathbb{R}$. Daher gilt

$$(x)' = 1.$$

3. $f(x) = x^2$. Dann

$$f'(x) = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{2hx + h^2}{h} = 2x,$$

daher

$$\left(x^2\right)' = 2x.$$

4. $f(x) = \frac{1}{x}$. Dann gilt

$$f'(x) = \lim_{h \to 0} \frac{1}{h} \left(\frac{1}{x+h} - \frac{1}{x} \right) = \lim_{h \to 0} \left(-\frac{1}{(x+h)\,x} \right) = -\frac{1}{x^2},$$

so dass

$$\left(\frac{1}{x} \right)' = -\frac{1}{x^2}.$$

Analog zeigt man, dass, für alle $n \in \mathbb{Z}$,

$$\boxed{(x^n)' = nx^{n-1}}$$

(siehe Aufgaben).

5. $f(x) = \exp(x)$. Dann gilt

$$f'(x) = \lim_{y \to x} \frac{\exp(y) - \exp(x)}{y - x} = \exp(x),$$

(siehe Aufgaben), daher

$$\boxed{(\exp(x))' = \exp(x)}.$$

6. $f(x) = \sin x$. Dann

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{\sin(x+h) - \sin x}{h} \\ &= \lim_{h \to 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \lim_{h \to 0} \sin x \frac{\cos h - 1}{h} + \lim_{h \to 0} \cos x \frac{\sin h}{h}. \end{aligned}$$

Nach Aufgaben haben wir

$$\lim_{h\to 0}\frac{\cos h-1}{h}=0 \text{ and } \lim_{h\to 0}\frac{\sin h}{h}=1,$$

daher

$$\boxed{(\sin x)'=\cos x}$$

Analog beweist man

$$\boxed{(\cos x)'=-\sin x}$$

(siehe Aufgaben)

Satz 3.14 *Ist die Funktion f differenzierbar an x, so ist f stetig an x.*

Beweis. Wir haben

$$\lim_{\substack{y\to x\\ y\in I\setminus\{x\}}}(f(y)-f(x))=\lim_{\substack{y\to x\\ y\in I\setminus\{x\}}}\frac{f(y)-f(x)}{y-x}(y-x)=f'(x)\lim_{\substack{y\to x\\ y\in I\setminus\{x\}}}(y-x)=0,$$

woraus folgt

$$\lim_{\substack{y\to x\\ y\in I\setminus\{x\}}}f(y)=f(x).$$

Somit ist f stetig an x. ■

Beispiel. Betrachten wir die Funktion

$$f(x)=\begin{cases}1, & x>0\\ 0, & x\le 0\end{cases}$$

Da $f=\text{const}$ auf $(0,+\infty)$ und $(-\infty,0)$, erhalten wir, dass $f'(x)=0$ für alle $x\neq 0$. An $x=0$ ist die Funktion f unstetig, da

$$\lim_{\substack{x\to 0\\ x>0}}f(x)=1\neq 0=f(0),$$

woraus folgt nach Satz 3.14, dass f nicht differenzierbar an 0 ist.

Beispiel. Die Funktion $f(x)=|x|$ ist stetig auf $\mathbb{R}$, aber an der Stelle $x=0$ ist sie nicht differenzierbar (siehe Aufgaben).

3.7.2 Physikalische Bedeutung der Ableitung

Sei $f(t)$ die Position eines Körpers an der Zahlengerade am Zeitpunkt t. Dann gilt für den Differenzenquotient

$$\begin{aligned}\frac{f(t+h)-f(t)}{h} &= \frac{\text{Verschiebung von Körper}}{\text{Zeitintervall}}\\ &= \text{durchschnittliche Geschwindigkeit im } [t,t+h]\end{aligned}$$

Für $h\to 0$ erhalten wir, dass die Ableitung $f'(t)$ gleich *unmittelbare* Geschwindigkeit von Körper am Zeitpunkt t ist.

Zum Beispiel, im Auto wird $f'(t)$ am Tachometer an jedem Zeitpunkt t gezeigt.

3.7.3 Geometrische Bedeutung der Ableitung

In $\mathbb{R}^2$ definieren wir eine *Gerade* als der Graph einer linearen Funktion $y = Ax + B$ mit $A, B \in \mathbb{R}$. Diese Definition ist allerdings nicht vollständig da sie die senkrechten Geraden $x = \text{const}$ nicht bedeckt. Trotzdem ist diese Definition für uns ausreichend. Die Funktion $y = Ax + B$ heißt die Gleichung der Gerade, und der Koeffizient A heißt die *Steigung* der Gerade.

Jede Gerade, die durch einen Punkt (x_0, y_0) geht, hat die Gleichung

$$Y = A(X - x_0) + y_0$$

wobei wir (X, Y) als Koordinaten in $\mathbb{R}^2$ benutzen. Geht die Gerade auch durch (x_1, y_1), so erhalten wir für die Steigung

$$A = \frac{y_1 - y_0}{x_1 - x_0}.$$

Betrachten wir den Graph einer Funktion $y = f(x)$. Die Gerade durch zwei Punkte $(x, f(x))$ und $(x + h, f(x + h))$ auf dem Graph heißt die *Sekante*. Für $x_0 = x$, $y_0 = f(x)$, $x_1 = x + h$, und $y_1 = f(x + h)$ erhalten wir die Steigung der Sekante

$$A = A(h) = \frac{f(x + h) - f(x)}{h}.$$

Für $h \to 0$ erhalten wir $A(h) \to f'(x)$, während die Sekante gegen die *Tangente* am Punkt $(x, f(x))$ konvergiert. Deshalb ist die Ableitung $f'(x)$ gleich die Steigung der Tangente, und die Tangente hat die Gleichung

$$\boxed{Y = f'(x)(X - x) + f(x)}. \tag{3.24}$$

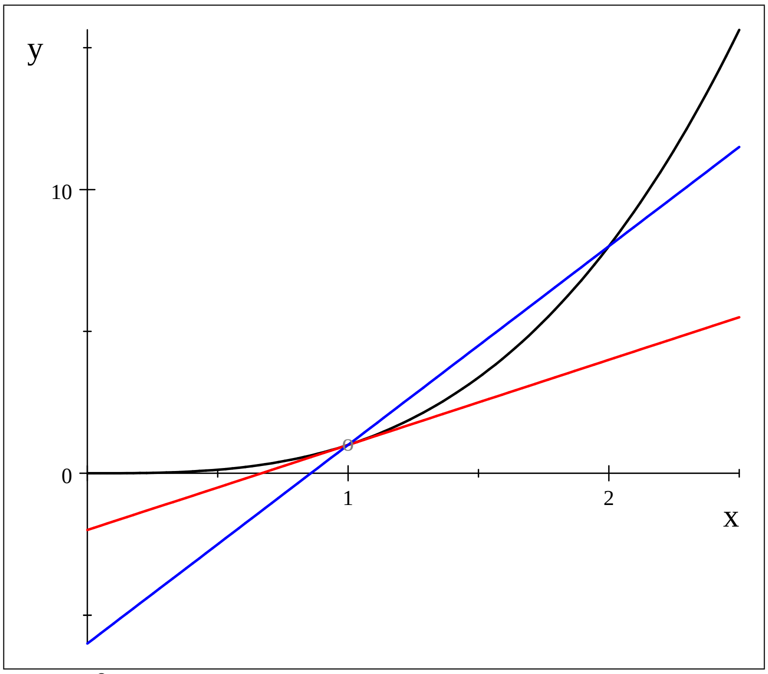
Beispiel. Für $f(x) = x^3$ erhalten wir aus (3.24) die Gleichung der Tangente

$$Y = 2x^2(X - x) + x^3.$$

Zum Beispiel, für $x = 1$ ergibt sie

$$Y = 3(X - 1) + 1 = 3X - 2$$

(siehe das Bild).



Die Funktion $f(x) = x^3$ (schwarz), ihre Tangente an $(1,1)$ (rot), und ihre Sekante durch $(1,1)$ und $(2,8)$ (blau)

3.7.4 Landau-Symbol o

Die Ableitung kann benutzt werden um die Werte von $f(x+h)$ für kleine h ungefähr berechnen, angenommen, dass $f(x)$ und $f'(x)$ gegeben sind. Es folgt aus (3.23) dass für kleinen Werte von h

$$f(x+h) \approx f(x) + f'(x)\,h. \tag{3.25}$$

Die genaue Bedeutung von dieser Relation wird unterhalb erklärt.

Definition. Seien $f(x)$ und $g(x)$ zwei Funktionen auf einem interval $I \subset \mathbb{R}$, und sei $a \in \overline{I}$. Man schreibt

$$f(x) = o(g(x)) \text{ für } x \to a \tag{3.26}$$

($f(x)$ ist klein o von $g(x)$, $f(x)$ ist vernachlässigbar klein gegenüber $g(x)$) falls

$$\lim_{\substack{x \to a \\ x \in I \setminus \{a\}}} \frac{f(x)}{g(x)} = 0.$$

Das Symbol o heißt das *Landau-Symbol.*

Zum Beispiel, es gilt

$$x^2 = o(x) \text{ für } x \to 0$$

während

$$x = o\left(x^2\right) \text{ für } x \to +\infty.$$

Lemma 3.15 *Existiert $f'(x)$, so gilt*

$$f(x+h) = f(x) + f'(x)\,h + o(h) \quad \textit{für} \quad h \to 0. \tag{3.27}$$

Die Relation (3.27) kann als eine rigorose Version von (3.25) betrachtet werden. Der Term $o(h)$ stellt den Approximationsfehler in (3.25) dar.

Beweis. Wir müssen beweisen, dass

$$f(x+h) - f(x) - f'(x)\,h = o(h) \quad \text{für} \quad h \to 0.$$

Wir haben

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - f'(x)\,h}{h} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} - f'(x) = 0,$$

was zu beweisen war. ■

Beispiel. Sei $f(x) = \sin x$. Nach (3.27) haben wir

$$\sin(x+h) = \sin x + h\cos x + o(h).$$

Insbesondere für $x = 0$ erhalten wir

$$\sin h = h + o(h). \tag{3.28}$$

Diese Relation kann auch mit Hilfe von der Potenzreihe von sin bewiesen werden. In der Tat gilt für $h = 0,1$

$$\sin 0,1 = 0,09983\,341\,664\,682\,82...$$

so dass (3.28) eine gute Approximation für $\sin h$ liefert.

3.7.5 Differential

In der Definition (3.23) der Ableitung bezeichnet man häufig h mit dx und nennt dx das *Differential* der Variable x. Die Relation (3.27) kann man wie folgt umschreiben:

$$f(x+dx) - f(x) = f'(x)\,dx + o(dx) \text{ für } dx \to 0. \tag{3.29}$$

Definition. Der Ausdruck $f'(x)\,dx$ in der rechten Seite von (3.29) heißt das *Differential* der Funktion f und wird mit $df(x)$ bezeichnet, so dass

$$df(x) = f'(x)\,dx. \tag{3.30}$$

Insbesondere gilt

$$f'(x) = \frac{df(x)}{dx}.$$

Der Ausdruck $\frac{df}{dx}$ wird häufig statt f' für Bezeichnung der Ableitung benutzt.

Das Differential $df(x)$ ist für jedes x eine lineare Funktion von dx. By (3.29) haben wir

$$f(x+dx) - f(x) = df(x) + o(dx),$$

so dass das Differential ein linearer dominanter Teil der Differenz der Funktion ist.

Beispiel. Benutzen wir die oberhalb berechneten Ableitungen und erhalten aus (3.30)

$$\begin{aligned} dx^n &= nx^{n-1}dx \\ d\exp(x) &= \exp(x)\,dx \\ d\sin x &= \cos x\ dx \\ d\cos x &= -\sin x\ dx. \end{aligned}$$

3.8 Rechenregeln für Ableitung

Satz 3.16 *Seien f und g zwei Funktionen auf einem Intervall $I \subset \mathbb{R}$, die differenzierbar in $x \in I$ sind. Dann sind die Funktionen $f+g$, fg, $\frac{f}{g}$ auch differenzierbar in x (im Fall von f/g vorausgesetzt $g \neq 0$) und die folgenden Identität sind erfüllt:*

(a) Summenregel:

$$\boxed{(f+g)'(x) = f'(x) + g'(x).} \tag{3.31}$$

(b) Produktregel oder Leibnizregel:

$$\boxed{(fg)'(x) = f'(x)\,g(x) + f(x)\,g'(x)\,.} \tag{3.32}$$

(c) Quotientenregel:

$$\boxed{\left(\frac{f}{g}\right)'(x) = \frac{f'(x)\,g(x) - f(x)\,g'(x)}{g^2(x)}.} \tag{3.33}$$

Es folgt aus (3.32), dass für jedes $c \in \mathbb{R}$

$$\boxed{(cf)' = cf'.}$$

Beispiel. Bestimmen wir die Ableitung der Funktion

$$\tan x = \frac{\sin x}{\cos x}.$$

Nach der Quotientenregel haben wir

$$(\tan x)' = \frac{(\sin x)'\cos x - \sin x\,(\cos x)'}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

Satz 3.17 (Kettenregel) *Seien f eine Funktion auf einem Intervall A und g eine Funktion auf einem Intervall B, so dass die Verkettung $g \circ f$ definiert ist (d.h. $f(A) \subset B$). Sei f differenzierbar in $x \in A$ und g differenzierbar in $y = f(x) \in B$. Dann ist $g \circ f$ differenzierbar in x und es gilt*

$$\boxed{(g \circ f)'(x) = g'(y)\,f'(x)}. \tag{3.34}$$

Man kann auch schreiben

$$(g(f(x)))' = g'(f(x))\,f'(x)\,.$$

Beispiel. Betrachten wir die Exponentialfunktion zur Basis $a >$, d.h. die Funktion $F(x) = a^x$ für $x \in \mathbb{R}$. Nach Definition haben wir

$$F(x) = \exp(x \ln a)\,.$$

Wir stellen F als eine Verkettung dar:

$$F(x) = \exp(y) \quad \text{mit} \quad y = x \ln a.$$

Nach der Kettenregel erhalten wir

$$F'(x) = (\exp(y))' (x \ln a)' = \exp(y) \ln a = a^x \ln a,$$

d.h.

$$\boxed{(a^x)' = a^x \ln a.}$$

Beispiel. Die Funktion $F(x) = \exp(\cos 2x)$ ist die Verkettung von drei Funktionen:

$$F(x) = \exp(z) \quad \text{mit} \quad z = \cos y \quad \text{und} \quad y = 2x.$$

Zwei Anwendungen von der Kettenregel ergeben

$$F'(x) = (\exp(z))' (\cos y)' (2x)' = -2 \exp(z) \sin y = -2 \exp(\cos(2x)) \sin 2x.$$

Satz 3.18 (Ableitung der Umkehrfunktion) *Sei f eine streng monotone, stetige Funktion auf einem Intervall I, so dass die Umkehrfunktion f^{-1} auf dem Intervall $J = f(I)$ definiert ist. Ist f differenzierbar in $x \in I$ und $f'(x) \neq 0$, so ist f^{-1} differenzierbar in $y = f(x)$ und es gilt*

$$\boxed{(f^{-1})'(y) = \frac{1}{f'(x)}}. \tag{3.35}$$

Bemerkung. Die Tangente zum Graph der Funktion f am Punkt (x, y) hat die Gleichung

$$Y - y = A(X - x),$$

wobei $A = f'(x)$ die Steigung ist. Für die Umkehrfunktion werden die Rollen von X und Y vertauscht, so dass die Gleichung der Tangente zum f^{-1} ist

$$X - x = \frac{1}{A}(Y - y),$$

vorausgesetzt $A \neq 0$. Somit ist die Steigung der Tangente gleich $\frac{1}{A}$, was die Identität (3.35) erklärt .

Da $y = f(x)$ äquivalent zu $x = f^{-1}(y)$ ist, so können wir (3.35) wie folgt umschreiben:

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(f^{-1}(y))}$$

und

$$\left(f^{-1}\right)'(f(x)) = \frac{1}{f'(x)}.$$

Beispiel. Sei $f(x) = \exp(x)$ auf $I = \mathbb{R}$. Die Umkehrfunktion von $x \mapsto \exp(x)$ ist die Funktion $y \mapsto \ln y$ auf $(0, +\infty)$. Nach (3.35) erhalten wir in $y = \exp(x)$

$$(\ln y)' = \frac{1}{(\exp(x))'} = \frac{1}{\exp(x)},$$

woraus folgt

$$\boxed{(\ln y)' = \frac{1}{y}}.$$

Beispiel. Betrachten wir die Potenzfunktion $f(x) = x^a$ wobei $x > 0$ und $a \in \mathbb{R}$. Nach Definition haben wir

$$f(x) = \exp(a \ln x),$$

so dass f eine Verkettung ist:

$$f(x) = \exp(y) \quad \text{mit} \quad y = a \ln x.$$

Nach der Kettenregel erhalten wir

$$(x^a)' = (\exp(y))' (a \ln x)' = \exp(y) \frac{a}{x} = x^a \frac{a}{x} = a x^{a-1},$$

d.h.

$$\boxed{(x^a)' = a x^{a-1}}.$$

Diese Identität für $a \in \mathbb{Z}$ haben wir schon früher gesehen. Zum Beispiel, es gilt

$$\left(\sqrt{x}\right)' = \left(x^{1/2}\right)' = \frac{1}{2} x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}$$

und

$$\left(\frac{1}{x}\right)' = \left(x^{-1}\right)' = -x^{-2} = -\frac{1}{x^2}.$$

Beweis von Satz 3.16. (a) Nach Definition von Ableitung und Satz 3.2 erhalten wir

$$\begin{aligned}
(f+g)'(x) &= \lim_{y \to x} \frac{(f(y) + g(y)) - (f(x) + g(x))}{y - x} \\
&= \lim_{y \to x} \frac{f(y) - f(x)}{y - x} + \lim_{y \to x} \frac{g(y) - g(x)}{y - x} \\
&= f'(x) + g'(x).
\end{aligned}$$

(b) Mit gleichem Argument und mit Hilfe von Stetigkeit von f an der Stelle x, die nach Satz 3.14 gilt, erhalten wir

$$\begin{aligned}
(fg)'(x) &= \lim_{y \to x} \frac{f(y) g(y) - f(x) g(x)}{y - x} \\
&= \lim_{y \to x} \frac{f(y) g(y) - f(y) g(x)}{y - x} + \lim_{y \to x} \frac{f(y) g(x) - f(x) g(x)}{y - x} \\
&= \lim_{y \to x} f(y) \lim_{y \to x} \frac{g(y) - g(x)}{y - x} + g(x) \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \\
&= f(x) g'(x) + g(x) f'(x),
\end{aligned}$$

was zu beweisen war.

(c) Berechnen wir zunächst $\left(\frac{1}{g}\right)'$:

$$\begin{aligned}\left(\frac{1}{g}\right)'(x) &= \lim_{y\to x}\frac{1}{y-x}\left(\frac{1}{g(y)}-\frac{1}{g(x)}\right)\\ &= \lim_{y\to x}\frac{1}{y-x}\left(\frac{g(x)-g(y)}{g(y)\,g(x)}\right)\\ &= \lim_{y\to x}\frac{g(x)-g(y)}{y-x}\lim_{y\to x}\frac{1}{g(y)\,g(x)}\\ &= \frac{-g'(x)}{g^2(x)},\end{aligned}$$

d.h.

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g^2(x)}, \tag{3.36}$$

was ein spezielle Fall von (3.33) für $f = 1$ ist.

Für allgemeine Funktion f erhalten wir mit Hilfe von (3.32) und (3.36):

$$\begin{aligned}\left(\frac{f}{g}\right)' &= \left(f\frac{1}{g}\right)' = f'\left(\frac{1}{g}\right) + f\left(\frac{1}{g}\right)'\\ &= \frac{f'}{g} - \frac{fg'}{g^2} = \frac{f'g - fg'}{g^2}.\end{aligned}$$

■

Bemerkung. Für das Produkt von n Funktionen $f_1, ..., f_n$ gilt die Regel

$$(f_1...f_n)' = f_1'f_2...f_n + f_1f_2'f_3...f_n + ...f_1...f_k'...f_n + ...f_1...f_{n-1}f_n',$$

die man per Induktion nach n beweisen kann.

Das folgende Lemma brauchen wir für die Beweise von Sätzen 3.17 und 3.18

Lemma 3.19 *Sei f eine Funktion auf Intervall I die in einem $x \in I$ differenzierbar ist. Dann existiert eine Funktion F auf I mit den folgenden Eigenschaften:*

(i) $f(y) - f(x) = F(y)(y-x)$ *für alle* $y \in I$

(ii) $F(x) = f'(x)$

(iii) F *ist stetig in* x.

Bemerkung. Dieses Lemma impliziert Lemma 3.15 da

$$\begin{aligned}f(x+h) - f(x) &= F(x+h)\,h\\ &= f'(x)\,h + (F(x+h) - f'(x))\,h\\ &= f'(x)\,h + o(h),\end{aligned}$$

wobei wir benutzt haben, dass $F(x+h) - f'(x) \to 0$ für $h \to 0$.

Beweis. Die Eigenschaften (i) und (ii) zwingen die folgende Identität für F:

$$F(y) = \begin{cases} \dfrac{f(y) - f(x)}{y - x}, & y \neq x, \\ f'(x), & y = x. \end{cases} \tag{3.37}$$

So, definieren wir F mit (3.37). Dann (i) gilt für $y \neq x$ nach Definition, und (i) gilt für $y = x$, da die beiden Seiten von (i) verschwinden. Die Identität (ii) gilt auch nach Definition, und (iii) gilt, da

$$\lim_{\substack{y \to x \\ y \in I \setminus \{x\}}} F(y) = \lim_{\substack{y \to x \\ y \in I \setminus \{x\}}} \frac{f(y) - f(x)}{y - x} = f'(x) = F(x).$$

■

Die Funktion F aus Lemma 3.19 heißt die *Verhältnisfunktion* von f an der Stelle x. In der Tat ist $F(y)$ gleich die Steigung der Sekante der Graph von f durch $(x, f(x))$ und $(y, f(y))$ falls $y \neq x$, und die Steigung der Tangente an der Stelle $(x, f(x))$ für $y = x$. Nach Lemma 3.19 ist die Verhältnisfunktion stetig in x.

Beweis von Satz 3.17. Sei F die Verhältnisfunktion von f in x und G – die Verhältnisfunktion von g in y. Dann haben wir

$$\begin{aligned} f(u) - f(x) &= F(u)(u - x) \quad \forall\, u \in A \\ g(v) - g(y) &= G(v)(v - y) \quad \forall v \in B, \end{aligned} \tag{3.38}$$

Einsetzen in (3.38) $v = f(u)$ und $y = f(x)$ ergibt

$$\begin{aligned} (g \circ f)(u) - (g \circ f)(x) &= g(f(u)) - g(f(x)) \\ &= G(f(u))(f(u) - f(x)) \\ &= G(f(u)) F(u)(u - x), \end{aligned}$$

woraus folgt

$$(g \circ f)'(x) = \lim_{u \to x} \frac{(g \circ f)(u) - (g \circ f)(x)}{u - x} = \lim_{u \to x} G(f(u)) F(u).$$

Da $G(v)$ stetig in $y = f(x)$ und $f(u)$ stetig in x ist, so erhalten wir nach dem Satz 3.5, dass die Verkettung $G(f(u))$ stetig in x ist. Da $F(u)$ auch stetig in x ist, so gilt

$$(g \circ f)'(x) = G(f(x)) F(x) = G(y) F(x) = g'(y) f'(x),$$

was zu beweisen war. ■

Beweis von Satz 3.18. Das Bild $J = f(I)$ ist ein Intervall nach Korollar 3.7, und the Umkehrfunktion f^{-1} existiert und ist stetig auf J nach Satz 3.10. Sei F die Verhältnisfunktion von f in x, so dass für alle $u \in I$

$$f(u) - f(x) = F(u)(u - x). \tag{3.39}$$

Setzen wir $v = f(u)$, $y = f(x)$ und erhalten aus (3.39), dass für alle $v \in J$

$$v - y = F\left(f^{-1}(v)\right)\left(f^{-1}(v) - f^{-1}(y)\right).$$

Daraus folgt, dass

$$\left(f'\right)'(y) = \lim_{\substack{v\to y\\ v\in J\setminus\{y\}}} \frac{f^{-1}(v) - f^{-1}(y)}{v-y} = \lim_{\substack{v\to y\\ v\in J\setminus\{y\}}} \frac{1}{F\left(f^{-1}(v)\right)}.$$

Da die Funktion f^{-1} stetig auf J ist und F stetig in $x = f^{-1}(y)$ ist, so erhalten wir, dass

$$\lim_{\substack{v\to y\\ v\in J}} F\left(f^{-1}(v)\right) = F\left(f^{-1}(y)\right) = F(x) = f'(x),$$

woraus folgt

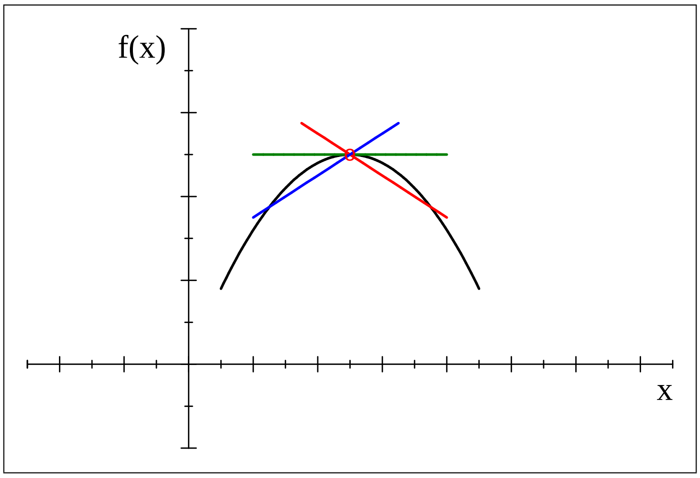
$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)},$$

was zu beweisen war. ∎

3.9 Sätze von Fermat, Rolle und Lagrange (Mittelwertsatz)

Satz 3.20 (Satz von Fermat) *Sei f eine Funktion auf einem offenen Intervall I und sei $x \in I$ eine Maximumstelle von f, d.h. $f(x) = \max_I f$. Ist f in x differenzierbar, so gilt $f'(x) = 0$. Das gleiche gilt für eine Minimumstelle.*

Die geometrische Bedeutung von dieser Aussage ist wie folgt: die Steigung der Tangente an der Maximumstelle von f ist 0, d.h. die Tangente an dieser Stelle waagerecht ist.



Die rechte Sekante (rot), die linke Sekante (blau), und die Tangente (grün), die an der Maximumstelle waagerecht ist.

Diese Aussage gilt nur wenn die Maximumstelle in einem offenen Intervall liegt, und gilt nicht für die Grenzen des Intervalles.

Beweis. Da $f(y) \le f(x)$ für alle $y \in I$, so erhalten wir

$$f'(x) = \lim_{\substack{y\in I\\ y>x}} \frac{f(y)-f(x)}{y-x} \le 0 \quad \text{und} \quad f'(x) = \lim_{\substack{y\in I\\ y<x}} \frac{f(y)-f(x)}{y-x} \ge 0,$$

woraus folgt $f'(x) = 0$.

Der Fall von Minimumstelle wird analog behandelt. ∎

Korollar 3.21 *Sei f eine stetige Funktion auf einem Intervall $[a, b]$ mit $a < b$, die auf (a, b) differenzierbar ist. Ist x eine Maximumstelle bzw Minimumstelle, so gilt entweder $f'(x) = 0$ oder $x = a$ oder $x = b$.*

Beweis. Gilt $x \in (a, b)$ so gilt $f'(x) = 0$ nach Satz 3.20. Sonst $x = a$ oder $x = b$. ■

Die Menge

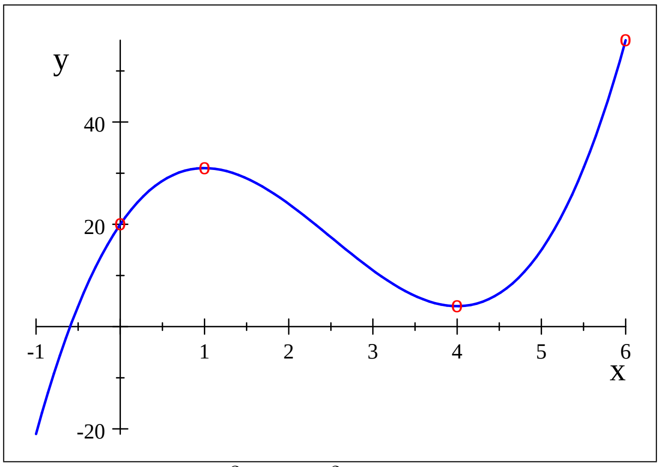
$$K_f = \{x \in I : f'(x) = 0\} \cup \{a, b\}$$

heißt die *kritische* Menge von f. Nach Korollar 3.21 liegen die Maximum- und Minimumstellen von f in K_f.

Beispiel. Bestimmen wir das Maximum und Minimum der Funktion

$$f(x) = 2x^3 - 15x^2 + 24x + 20$$

auf dem Intervall $[0, 6]$.



Der Graph der Funktion $f(x) = 2x^3 - 15x^2 + 24x + 20$ und die kritischen Stellen von $f(x)$

Die Ableitung

$$f'(x) = 6x^2 - 30x + 24 = 6(x-1)(x-4)$$

hat die Nullstellen $x_1 = 1$ und $x_2 = 4$. Somit erhalten wir die kritische Menge

$$K_f = \{0, 1, 4, 6\}.$$

Berechnung den Werten an den kritischen Stellen ergibt:

$$f(0) = 20, \quad f(1) = 31, \quad f(4) = 4, \quad f(6) = 56.$$

Deshalb $\max_{[0,6]} f = 56$ wird an der Stelle $x = 6$ angenommen, und $\min_{[0,6]} f = 4$ wird an $x = 4$ angenommen.

Satz 3.22 (Satz von Rolle) *Sei f eine stetige Funktion auf einem Intervall $[a, b]$ mit $a < b$, die auf (a, b) differenzierbar ist. Gilt $f(a) = f(b)$ so existiert eine Stelle $c \in (a, b)$ mit $f'(c) = 0$.*

Beweis. Nach Satz 3.8 (Extremwertsatz) existieren $\max_{[a,b]} f$ und $\min_{[a,b]} f$. Sei c_1 eine Maximumstelle von f und c_2 - eine Minimumstelle. Liegt eine von diesen Stellen im (a,b) so verschwindet f' an dieser Stelle nach Satz 3.20. Seien c_1 und c_2 die Grenzen von $[a,b]$. Da $f(a) = f(b)$, es folgt, dass der Wert von $f(a)$ ist gleichzeitig das Maximum und Minimum von f, so dass f eine Konstantefunktion auf $[a,b]$ ist. Dann gilt $f'(c) = 0$ für jedes $c \in (a,b)$. ■

Hauptsatz 3.23 (Mittelwertsatz von Lagrange) *Sei f eine stetige Funktion auf einem Intervall $[a,b]$ mit $a < b$, die auf (a,b) differenzierbar ist. Dann existiert ein $c \in (a,b)$ mit*

$$f(b) - f(a) = f'(c)(b-a).$$

Beweis. Betrachten wir die Funktion

$$h(x) := f(x) - \frac{f(b) - f(a)}{b-a}(x-a).$$

Die Funktion h ist offensichtlich stetig auf $[a,b]$ und differenzierbar auf (a,b), und es gilt

$$h(b) = h(a) = f(a).$$

Nach Satz 3.22, es gibt ein $c \in (a,b)$ mit $h'(c) = 0$. Da

$$h' = f' - \frac{f(b) - f(a)}{b-a},$$

es folgt daraus, dass

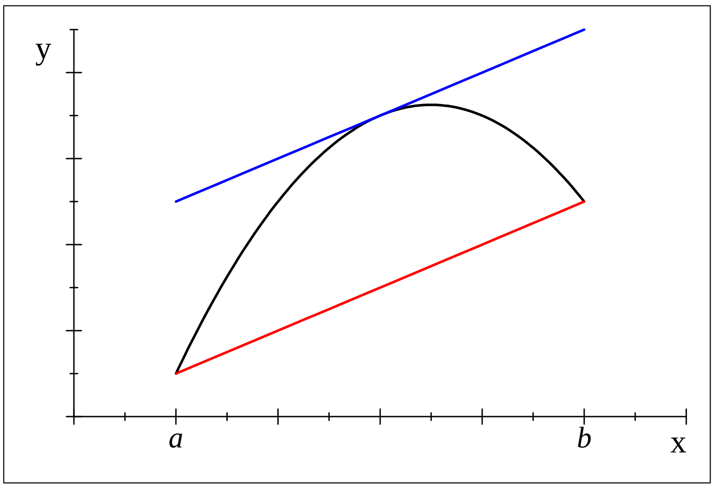
$$f'(c) = \frac{f(b) - f(a)}{b-a},$$

was zu beweisen war. ■

Bemerkung. Die Zahl

$$A = \frac{f(b) - f(a)}{b-a}$$

ist die Steigung der Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$. Somit besagt Satz 3.23, dass die Steigung dieser Sekante gleich die Steigung der Tangente an einer Mittelstelle $c \in (a,b)$ ist.



Die Tangente an $(c, f(c))$ (blau) und die Sekante durch $(a, f(a)), (b, f(b))$ (rot) haben die gleichen Steigungen

3.10 Untersuchung von Funktion mit Hilfe von Ableitung

Satz 3.24 (Konstantentest) *Sei f eine differenzierbare Funktion auf einem Intervall I. Dann gilt $f = \mathrm{const}$ auf I genau dann wenn $f'(x) = 0$ für alle $x \in I$.*

Beweis. Gilt $f = \mathrm{const}$, so gilt auch $f'(x) = 0$ für alle $x \in I$. Für die Rückrichtung beweisen wir, dass $f(a) = f(b)$ für alle verschiedene $a, b \in I$. Nach Satz 3.23 erhalten wir, für ein $c \in (a, b)$,

$$f(a) - f(b) = f'(c)(a - b).$$

Da $f'(c) = 0$, so erhalten wir $f(a) = f(b)$. ■

Beispiel. Bestimmen wir alle Funktionen f mit $f'(x) = x$ für alle $x \in \mathbb{R}$. Bemerken wir zunächst, dass die Funktion $f(x) = \frac{x^2}{2}$ diese Gleichung erfüllt. Gibt es weitere Funktionen, die diese Gleichung erfüllen? Ist f eine andere Funktion mit $f' = x$, so folgt daraus $f' = \left(\frac{x^2}{2}\right)'$ und daher

$$\left(f - \frac{x^2}{2}\right)' = 0 \text{ auf } \mathbb{R}.$$

Nach Satz 3.24 folgt $f - \frac{x^2}{2} = C$ für eine Konstante C. Deshalb erhalten wir

$$f(x) = \frac{x^2}{2} + C,$$

was alle Funktionen darstellt, die die Gleichung $f' = x$ erfüllen.

Für jedes Intervall I mit den Grenzpunkten $a < b$ bezeichnen wir mit I^0 das *offene* Intervall (a, b). Offensichtlich gilt immer $I^0 \subset I$. Jedes $x \in I^0$ heißt *innerer Punkt* von I.

Satz 3.25 (Monotonie Test) *Sei I ein Intervall. Sei f eine stetige Funktion auf I, die auf I^0 differenzierbar ist. Gilt $f'(x) \geq 0$ für alle $x \in I^0$, so ist f monoton steigend auf I. Gilt $f'(x) > 0$ für alle $x \in I^0$, so ist f streng monoton steigend auf I.*

Ähnliche Aussagen gelten für fallende Monotonie: $f' \leq 0$ auf I^0 ergibt, dass f monoton fallend auf I ist, und $f' < 0$ auf I^0 ergibt, dass f streng monoton fallend auf I ist.

Beweis. Betrachten wir zwei Werte $a < b$ in I. Die Funktion f ist stetig in $[a, b]$ und differenzierbar in $(a, b) \subset I^0$. Nach Satz 3.23 existiert $c \in (a, b)$ mit

$$f(b) - f(a) = f'(c)(b - a).$$

Somit impliziert die Voraussetzung $f' \geq 0$ auf I^0, dass $f(b) \geq f(a)$, d.h. f monoton steigend ist. Gilt $f' > 0$ auf I^0, so erhalten wir $f(b) > f(a)$, d.h. f streng monoton steigend ist. Die anderen Aussagen werden analog bewiesen. ■

Bemerkung. Ist eine differenzierbare Funktion f auf einem Intervall I monoton steigend, so gilt $f'(x) \geq 0$ für alle $x \in I$ da

$$f'(x) = \lim_{\substack{y \to x \\ y \in I,\ y > x}} \frac{f(y) - f(x)}{y - x} \geq 0.$$

Allerdings die Bedingung, dass f streng monoton steigend ist, impliziert die echte Ungleichung $f'(x) > 0$ nicht, wie man im Beispiel der Funktion $f(x) = x^3$ sieht.

Beispiel. Betrachten wir die Funktion $\sin x$ auf $[0, \pi]$. Die Ableitung $\cos x$ verschwindet an $x = \pi/2$, aber für $x \in [0, \pi/2)$ ist $\cos x$ positiv. Für $x \in (\pi/2, \pi]$ haben wir $\cos x = -\cos(\pi - x) < 0$. Wir beschließen, dass $\sin x$ auf $[0, \pi/2]$ streng monoton steigend ist und auf $[\pi/2, \pi]$ streng monoton fallend. Die Maximumstelle von $\sin x$ auf $[0, \pi]$ ist somit $x = \frac{\pi}{2}$ wo $\sin \frac{\pi}{2} = 1$, und die Minimumstellen sind $x = 0$ und $x = \pi/2$, wo $\sin x$ verschwindet.

Mit Hilfe von der Identität $\sin x = -\sin(\pi + x)$ zeigt man, dass $\sin x$ auf $\left[\pi, \frac{3\pi}{2}\right]$ streng monoton fallend ist, und auf $\left[\frac{3\pi}{2}, 2\pi\right]$ – streng monoton steigend. Die Minimumstelle von $\sin x$ auf $[\pi, 2\pi]$ ist $x = \frac{3\pi}{2}$ wo $\sin x = -1$.

Satz 3.25 kann benutzt werden um Ungleichungen zu beweisen.

Satz 3.26 (Vergleichstest)

(a) *Seien f und g zwei stetige Funktionen auf einem Intervall $[a, b)$, $a < b$, die auf (a, b) differenzierbar sind. Angenommen seien die Bedingungen:*

1. $f(a) \leq g(a)$
2. $f'(x) \leq g'(x)$ *für alle* $x \in (a, b)$.

Dann gilt $f(x) \leq g(x)$ für alle $x \in (a, b)$. Darüber hinaus, die Bedingung $f'(x) < g'(x)$ für alle $x \in (a, b)$ ergibt $f(x) < g(x)$ für alle $x \in (a, b)$.

(b) *Seien f und g zwei stetige Funktionen auf einem Intervall $(a, b]$, $a < b$, die auf (a, b) differenzierbar sind. Angenommen seien die Bedingungen:*

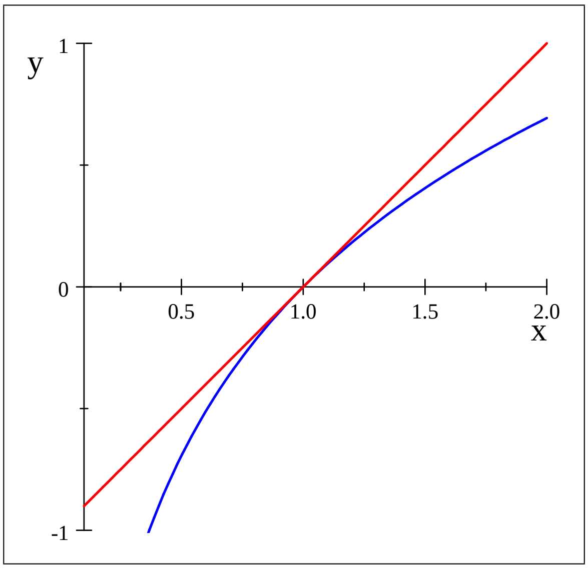
1. $f(b) \leq g(b)$
2. $f'(x) \geq g'(x)$ *für alle* $x \in (a, b)$.

   Dann gilt $f(x) \leq g(x)$ für alle $x \in (a, b)$. Darüber hinaus, die Bedingung $f'(x) > g'(x)$ für alle $x \in (a, b)$ ergibt $f(x) < g(x)$ für alle $x \in (a, b)$.

Beispiel. Beweisen wir die Ungleichung

$$\ln x \leq x - 1 \tag{3.40}$$

für alle $x > 0$. Die Graphen dieser zwei Funktion werden auf dem folgenden Bild gezeigt:



Die Funktionen $y = \ln x$ (blau) und $y = x - 1$ (rot)

Für $x = 1$ sind die beiden Seiten von (3.40) gleich 0. Für $x > 1$ haben wir

$$(\ln x)' = \frac{1}{x} < 1 = (x - 1)' .$$

Anwendung des Vergleichstests von Satz 3.26(a) auf Interval $[1, +\infty)$ ergibt, dass $\ln x < x - 1$ für alle $x > 1$.

Für $0 < x < 1$ haben wir

$$(\ln x)' = \frac{1}{x} > 1 = (x - 1)' .$$

Anwendung des Vergleichstests von Korollar 3.26(b) auf Interval $(0, 1]$ ergibt $\ln x < x - 1$ für alle $x \in (0, 1)$. Somit ist (3.40) erfüllt für alle $x > 0$.

Beweis von Satz 3.26. (a) Definieren wir eine neue Funktion $h = g - f$ und beachten, dass $h(a) \geq 0$ und $h' \geq 0$ auf (a, b). Nach Satz 3.25 ist h monoton steigend auf $[a, b)$. Daraus folgt, dass für alle $x \in (a, b)$,

$$h(x) \geq h(a) \geq 0,$$

und somit $f(x) \leq g(x)$. Im Fall von der echten Ungleichung $f'(x) < g'(x)$ erhalten wir, dass $h' > 0$ und somit h streng monoton steigend ist, woraus $h(x) > 0$ und $f(x) < g(x)$ folgen.

(b) In diesem Fall ist die Funktion h monoton fallend und $h(b) \geq 0$, daher

$$h(x) \geq h(b) \geq 0$$

für alle $x \in (a, b)$, was ergibt $f(x) \leq g(x)$. Gilt die echte Ungleichung $f'(x) < g'(x)$, so ist h streng monoton fallend, woraus folgt $h(x) > 0$ und somit $f(x) < g(x)$. ■

Satz 3.25 impliziert auch die folgende Eigenschaft von Umkehrfunktionen.

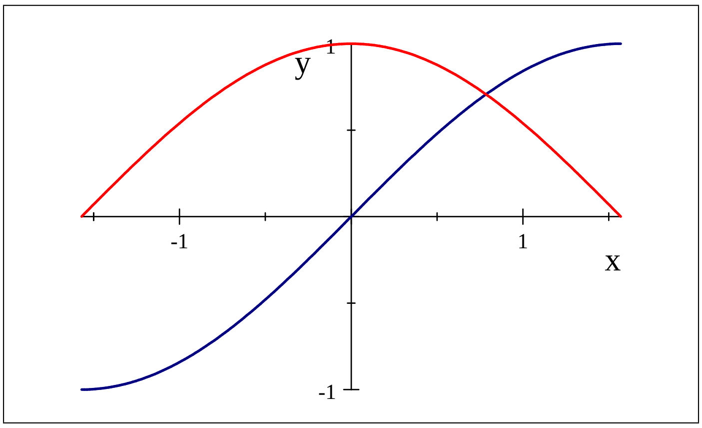
Satz 3.27 (Satz über inverse Funktion) *Sei f eine differenzierbare Funktion auf einem Intervall I mit $f'(x) > 0$ für alle $x \in I$ (oder $f'(x) < 0$ für alle $x \in I$). Dann existiert* *die Umkehrfunktion f^{-1} auf dem Intervall $J = f(I)$, f^{-1} ist differenzierbar auf J und es gilt für alle $y \in J$*

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)}, \tag{3.41}$$

wobei $x = f^{-1}(y)$.

Beweis. Nach Satz 3.25 ist f streng monoton auf I. Nach Satz 3.10 existiert die Umkehrfunktion f^{-1} auf J. Nach Satz 3.18 ist f^{-1}differenzierbar und erfüllt (3.41). ■

Beispiel. Betrachten wir die Funktion $f(x) = \sin x$ auf $I = (-\pi/2, \pi/2)$. Da $f'(x) = \cos x$ und $\cos x > 0$ auf I, so erhalten wir, dass f^{-1} auf $J = (-1, 1)$ existiert und differenzierbar ist.

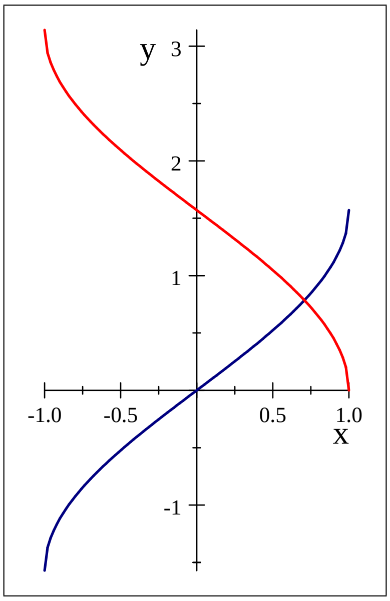


Die Graphe von Funktionen $\sin x$ (blau) und $\cos x$ (rot) auf $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

Die Umkehrfunktion von sin wird mit arcsin bezeichnen und heißt *Arkussinus*. Es folgt aus (3.35) dass für alle $y \in (-1, 1)$

$$(\arcsin y)' = \frac{1}{(\sin x)'} = \frac{1}{\cos x} = \frac{1}{\sqrt{1 - y^2}}.$$

In der Tat ist der Definitionsbereich von $\arcsin y$ das Intervall $[-1, 1]$ und der Wertebereich ist $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, da nach den Satz 3.25 $\sin x$ auf $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ streng monoton steigend ist. Allerdings ist $\arcsin y$ in $y = \pm 1$ nicht differenzierbar.



arcsin (blau) und arccos (rot)

Analog betrachten wir $f(x) = \cos x$ auf $(0, \pi)$. Dann gilt $f' = -\sin x < 0$ auf $(0, \pi)$, woraus folgt, dass f^{-1} auf $(-1, 1)$ existiert und differenzierbar ist. Die Funktion f^{-1} heißt *Arkuscosinus* und wird mit arccos bezeichnet. Für jedes $y \in (-1, 1)$ erhalten wir

$$(\arccos y)' = \frac{1}{(\cos x)'} = -\frac{1}{\sin x} = -\frac{1}{\sqrt{1-y^2}}.$$

Der vollständige Definitionsbereich von $\arccos y$ ist $[-1, 1]$ und der Wertebereich ist $[0, \pi]$ (siehe Aufgaben).

3.11 Die trigonometrische Form von komplexen Zahlen

Gegeben sei $a \in \mathbb{R}$, $a \neq 0$, betrachten wir die folgende Relation für reelle x, y:

$$x = y \mod a \text{ genau dann wenn } x - y = ka \text{ für ein } k \in \mathbb{Z} \tag{3.42}$$

(und sagen: x gleich y modulo a). Offensichtlich ist (3.42) die Äquivalenzrelation. Die Äquivalenzklassen heißen *Restklassen* modulo a, und die Restklasse von x wird mit $x \bmod a$ bezeichnet. Die Menge von allen Restklassen $\bmod a$ wird mit $\mathbb{T}_a$ bezeichnet.

Man definiert die Addition in $\mathbb{T}_a$ wie folgt

$$x \bmod a + y \bmod a := (x + y) \bmod a,$$

was offensichtlich wohldefiniert ist. Dann $\mathbb{T}_a$ mit der Operation $+$ ist eine Gruppe, wobei das Nullelement ist $0 \bmod a$ und das Negative von $x \bmod a$ ist $(-x) \bmod a$.

Definition. Die Elementen von $\mathbb{T}_{2\pi}$ (d.h. die Restklassen modulo 2π) heißen *Winkel.*

In anderen Wörter, ein Winkel ist die Restklasse $x \bmod 2\pi$, d.h. die Menge $\{x + 2\pi k : k \in \mathbb{Z}\}$, für ein $x \in \mathbb{R}$. Wir schreiben kurz $\mathbb{T} = \mathbb{T}_{2\pi}$. Für jeden Winkel $\varphi \in \mathbb{T}$, $\varphi = x \bmod 2\pi$, setzen wir $\sin \varphi := \sin x$. Dann ist $\sin \varphi$ wohldefiniert, da $\sin x$ 2π-periodisch ist und somit $\sin y = \sin x$ für alle $y = x \bmod 2\pi$. Analog werden alle andere trigonometrischen Funktionen auf $\mathbb{T}$ definiert.

Satz 3.28 *Jede komplexe Zahl $z \neq 0$ lässt sich in der folgenden Form eindeutig darstellen:*

$$z = r(\cos\varphi + i\sin\varphi), \tag{3.43}$$

wobei $r > 0$ und φ ein Winkel ist.

Der Ausdruck $r(\cos\varphi + i\sin\varphi)$ heißt *trigonometrische* Form von z, im Gegensatz zur *algebraischen* Form $z = \operatorname{Re} z + i\operatorname{Im} z$. Mit Hilfe von Eulerformel (3.16) lässt sich die Identität (3.43) wie folgt umschreiben:

$$z = re^{i\varphi}.$$

Der Ausdruck $re^{i\varphi}$ heißt *exponentielle Form* von z. Die Zahl r heißt *Polarradius* von z und der Winkel φ heißt *Polarwinkel* von z. Zusammen r und φ heißen die *Polarkoordinaten* for z.

We werden unterhalb sehen, dass $r = |z|$. Die Polarwinkel wird mit $\arg z$ bezeichnet und heißt auch *Argument* von z. Somit gilt die Darstellung

$$z = |z|\exp(i\arg z).$$

Beweis. Gilt (3.43) so erhalten wir

$$|z| = |r|\,|\cos\varphi + i\sin\varphi| = r\sqrt{\cos^2\varphi + \sin^2\varphi} = r.$$

so dass $r = |z|$. Insbesondere ist der Polarradius von z eindeutig bestimmt. Vergleichen mit der algebraischen Darstellung $z = x + iy$ ergibt

$$x = r\cos\varphi \quad \text{und} \quad y = r\sin\varphi$$

was äquivalent zu

$$\cos\varphi = \frac{x}{r} \quad \text{und} \quad \sin\varphi = \frac{y}{r} \tag{3.44}$$

ist. Zeigen wir, dass φ von (3.44) eindeutig bestimmt ist. Der Winkel φ wird von einer reellen Zahl aus $(-\pi, \pi]$ dargestellt, die auch mit φ bezeichnet wird. Dann erhalten wir aus (3.44) folgendes. Ist $y \geq 0$, dann $\sin\varphi \geq 0$ und somit $\varphi \in [0, \pi]$. Dann ist die Gleichung $\cos\varphi = \frac{x}{r}$ äquivalent zu $\varphi = \arccos\frac{x}{r}$. Ist $y < 0$, dann haben wir

$$\cos(-\varphi) = \frac{x}{r} \quad \text{und} \quad \sin(-\varphi) = -\frac{y}{x}$$

woraus folgt $-\varphi = \arccos\frac{x}{r}$. In den beiden Fällen haben wir

$$\varphi = \begin{cases} \arccos\frac{x}{r}, & y \geq 0 \\ -\arccos\frac{x}{r}, & y < 0 \end{cases} \tag{3.45}$$

was beweist die Eindeutigkeit von Polarwinkel.

Für Existenz von (r, φ) setzen wir $r = |z|$ und definieren φ durch (3.45), woraus (3.44) folgt. ∎

Wir wissen schon die folgenden Regeln für Beträge von Produkt bzw Quotient:

$$|z_1 z_2| = |z_1|\,|z_2| \quad \text{und} \quad \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}.$$

Die Argumente von Produkt bzw Quotient lassen sich auch schön bestimmen.

Satz 3.29 *Für alle $z_1, z_2 \in \mathbb{C} \setminus \{0\}$ gelten die folgenden Identitäten*

$$\arg(z_1 z_2) = \arg z_1 + \arg z_2 \tag{3.46}$$
$$\arg \frac{z_1}{z_2} = \arg z_1 - \arg z_2. \tag{3.47}$$

Beweis. Seien $z_1 = r_1 e^{i\varphi_1}$ und $z_2 = r_2 e^{i\varphi_2}$. Nach (2.40) erhalten wir

$$z_1 z_2 = r_1 e^{i\varphi_1} r_2 e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1+\varphi_2)}.$$

Daraus folgt $|z_1 z_2| = r_1 r_2 = |z_1| |z_2|$ (cf. Satz 1.22) und

$$\arg(z_1 z_2) = \varphi_1 + \varphi_2 = \arg z_1 + \arg z_2,$$

woraus (3.46) folgt. Mit Hilfe von (3.46) wir haben

$$\arg z_1 = \arg\left(\frac{z_1}{z_2} z_2\right) = \arg \frac{z_1}{z_2} + \arg z_2,$$

woraus (3.47) follows. ■

3.12 Unbestimmte Ausdrücke und Regel von l'Hôspital

Nach der Rechenregel für lim gilt die Identität

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$$

vorausgesetzt, dass die beiden Grenzwerte auf der rechten Seite existieren und ihrer Quotient bestimmt ist. Allerdings ist das nicht der Fall, wenn die beiden Grenzwerte gleich 0 oder gleich $\pm\infty$ sind. Man spricht in diesem Fall von der *unbestimmten Ausdrücken* $\frac{0}{0}$ bzw $\frac{\infty}{\infty}$. Häufig können diese Ausdrücke mit Hilfe von den folgenden Satz gelöst werden.

Hauptsatz 3.30 (Regel von l'Hôspital) *Seien f und g zwei differenzierbare Funktionen auf einem offenen Intervall $I \subset \mathbb{R}$. Sei $a \in \overline{\mathbb{R}}$ eine Grenze von I. Angenommen sei, dass $g \neq 0$ und $g' \neq 0$ auf I und*

(*a*) *entweder*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0, \tag{3.48}$$

(*b*) *oder*

$$\lim_{x\to a} f(x) = \pm\infty \quad \textit{und} \quad \lim_{x\to a} g(x) = \pm\infty. \tag{3.49}$$

Gilt

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = b \in \overline{\mathbb{R}}$$

so gilt auch

$$\lim_{x\to a} \frac{f(x)}{g(x)} = b. \tag{3.50}$$

Die Regel ist dann sehr einfach: um die unbestimmte Ausdrücke $\frac{0}{0}$ bzw $\frac{\infty}{\infty}$ zu lösen, sollen die Funktionen f und g in $\frac{f}{g}$ durch ihre Ableitungen ersetzt werden.

Die Voraussetzung $\lim_{x\to a} f(x) = \pm\infty$ in (3.49) ist für die Gültigkeit des Satzes nicht notwendig (da sie im Beweis nicht benutzt wird), aber in Anwendungen ist diese Bedingung normalerweise der Fall.

Beispiel. 1. Bestimmen wir $\lim_{x\to 0}\frac{\sin x}{x}$. Das ist unbestimmter Ausdruck der Form $\frac{0}{0}$. Nach Satz 3.30 erhalten wir in den beiden Intervallen $(0,+\infty)$ und $(-\infty,0)$,

$$\lim_{x\to 0}\frac{\sin x}{x} = \lim_{x\to 0}\frac{(\sin x)'}{(x)'} = \lim_{x\to 0}\frac{\cos x}{1} = 1. \tag{3.51}$$

Um rigorous zu sein, man soll die Gültigkeit der Gleichheiten in (3.51) rückwärts beweisen: da $\lim_{x\to 0}\frac{(\sin x)'}{(x)'}$ existiert und gleich 1 ist, so gilt nach Satz 3.30 auch $\lim_{x\to 0}\frac{\sin x}{x} = 1$. Diese Bemerkung gilt auch für alle Anwendungen von der Regel von l'Hôspital.

2. Die Regel von l'Hôspital gilt nur wenn $\lim\frac{f(x)}{g(x)}$ unbestimmter Ausdruck ist. Betrachten wir, zum Beispiel, $\lim_{x\to 1}\frac{x^2}{x}$, dass kein unbestimmter Ausdruck ist und offensichtlich gleich 1 ist. Andererseits gilt

$$\lim_{x\to 1}\frac{(x^2)'}{(x)'} = \lim_{x\to 1}\frac{2x}{1} = 2$$

und somit

$$\lim_{x\to 1}\frac{x^2}{x} \neq \lim_{x\to 1}\frac{(x^2)'}{(x)'}.$$

3. Bestimmen wir $\lim_{x\to+\infty}\frac{\exp(x)}{x^2}$, was unbestimmter Ausdruck der Form $\frac{\infty}{\infty}$ ist. Versuchen wir die Regel von l'Hôspital benutzen:

$$\lim_{x\to+\infty}\frac{\exp(x)}{x^2} = \lim_{x\to+\infty}\frac{(\exp(x))'}{(x^2)'} = \lim_{x\to+\infty}\frac{\exp(x)}{2x}, \tag{3.52}$$

and erhalten, dass der Grenzwert in der rechten Seite auch unbestimmter Ausdruck der Form $\frac{\infty}{\infty}$ ist. Benutzen wir die Regel von l'Hôspital zu diesem Grenzwert und erhalten

$$\lim_{x\to+\infty}\frac{\exp(x)}{2x} = \lim_{x\to+\infty}\frac{(\exp(x))'}{(2x)'} = \lim_{x\to+\infty}\frac{\exp(x)}{2} = +\infty. \tag{3.53}$$

Somit gilt nach zwei Anwendungen von der Regel von l'Hôspital

$$\lim_{x\to+\infty}\frac{\exp(x)}{x^2} = +\infty.$$

Analog beweist man per Induktion nach n, dass

$$\lim_{x\to+\infty}\frac{\exp(x)}{x^n} = +\infty \tag{3.54}$$

für alle $n\in\mathbb{N}$. Das gleiche Ergebnis kann man auch mit Hilfe von der Exponentialreihe erhalten.

4. Bestimmen wir $\lim_{x\to 0} x \ln x$ im Bereich $x > 0$. Da $x \to 0$ und $\ln x \to -\infty$, so haben wir unbestimmten Ausdruck der Form $0 \cdot \infty$. Um ihn zu lösen, stellen wir den Grenzwert in der folgenden Form dar:

$$\lim_{x\to 0} \frac{\ln x}{1/x},$$

was unbestimmter Ausdruck der Form $\frac{\infty}{\infty}$ ist. Nach der Regel von l'Hôspital erhalten wir

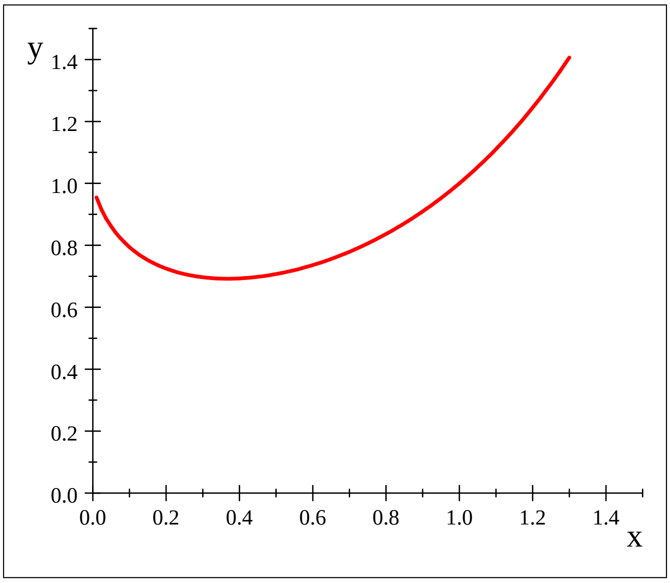
$$\lim_{x\to 0} \frac{\ln x}{1/x} = \lim_{x\to 0} \frac{(\ln x)'}{(1/x)'} = -\lim_{x\to 0} \frac{1/x}{1/x^2} = -\lim_{x\to 0} x = 0.$$

5. Bestimmen wir $\lim_{x\to 0} x^x$ im Bereich $x > 0$. Das ist unbestimmter Ausdruck der Form 0^0. Um ihn zu lösen, nehmen wir den Logarithmus von der Funktion x^x:

$$\lim_{x\to 0} \ln x^x = \lim_{x\to 0} x \ln x = 0,$$

wo wir das vorige Beispiel benutzt haben. Nach Satz 3.3 erhalten wir

$$\lim_{x\to 0} x^x = \lim_{x\to 0} \exp(\ln x^x) = \lim_{y\to 0} \exp(y) = 1.$$



Funktion $f(x) = x^x$

6. Bestimmen wir

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

was unbestimmter Ausdruck der Form 1^∞ ist. Der Logarithmus der gegebenen Funktion ist

$$x \ln\left(1 + \frac{1}{x}\right),$$

was unbestimmter Ausdruck der Form $\infty \cdot 0$ ist. Dann erhalten wir mit Hilfe von Wechsel $y = \frac{1}{x}$ und Regel von l'Hôspital

$$\lim_{x\to +\infty} x \ln\left(1 + \frac{1}{x}\right) = \lim_{y\to 0} \frac{\ln(1+y)}{y} = \lim_{y\to 0} \frac{(\ln(1+y))'}{(y)'} = \lim_{y\to 0} \frac{\frac{1}{1+y}}{1} = 1$$

woraus folgt

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

Für den Beweis des Satzes 3.30 brauchen wir eine Erweiterung des Mittelwertsatzes.

Hauptsatz 3.31 *(*Erweiterter Mittelwertsatz von Cauchy) *Seien f, g stetige Funktionen auf einem Intervall $[a, b]$, $a < b$, die auf (a, b) differenzierbar sind. Dann existiert ein $c \in (a, b)$ mit*

$$g'(c)\,(f(b)-f(a)) = f'(c)\,(g(b)-g(a)). \tag{3.55}$$

Der Mittelwertsatz von Satz 3.23 (Mittelwertsatz von Lagrange) ist ein spezieller Fall von Satz 3.31 für $g(x) = x$ da in diesem Fall (3.55) wird

$$f(b)-f(a) = f'(c)\,(b-a).$$

Beweis. Betrachten wir die Funktion

$$h(x) = f(x)\,(g(b)-g(a)) - g(x)\,(f(b)-f(a)).$$

Offensichtlich gilt

$$h(b)-h(a) = (f(b)-f(a))\,(g(b)-g(a)) - (g(b)-g(a))\,(f(b)-f(a)) = 0$$

so dass

$$h(a) = h(b).$$

Nach Satz 3.22 (Satz von Rolle) existiert ein $c \in (a, b)$ mit $h'(c) = 0$. Da

$$h'(x) = f'(x)\,(g(b)-g(a)) - g'(x)\,(f(b)-f(a)),$$

so erhalten wir (3.55). ■

Beweis von Satz 3.30. (a) Zunächst betrachten wir den Fall $a \in \mathbb{R}$. Definieren wir die Funktionen f und g an der Stelle a indem wir $f(a) = g(a) = 0$ setzen. Dann sind die Funktionen f und g stetig auf dem Intervall $I \cup \{a\}$. Nach dem Mittelwertsatz von Cauchy (Satz 3.31) gilt folgendes: für jedes $x \in I$ existiert ein $c \in (x, a)$ mit

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(c)}{g'(c)}.$$

Da

$$\lim_{x\to a}\frac{f'(c)}{g'(c)} = \lim_{c\to a}\frac{f'(c)}{g'(c)} = b,$$

es folgt

$$\lim_{x\to a}\frac{f(x)}{g(x)} = b.$$

Sei jetzt $a = +\infty$. Dann sind die Funktionen $F(t) = f\left(\frac{1}{t}\right)$ und $G(t) = g\left(\frac{1}{t}\right)$ auf einem Intervall $(0, \delta)$ mit $\delta > 0$ definiert. Anwendung von dem obigen Fall der Regel von l'Hôspital ergibt

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)} = \lim_{t\to0}\frac{F(t)}{G(t)} = \lim_{t\to0}\frac{F'(t)}{G'(t)} = \lim_{t\to0}\frac{f'\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right)}{g'\left(\frac{1}{t}\right)\left(-\frac{t}{t^2}\right)} = \lim_{t\to0}\frac{f'\left(\frac{1}{t}\right)}{g'\left(\frac{1}{t}\right)} = \lim_{x\to+\infty}\frac{f(x)}{g(x)} = b.$$

Der Fall $a = -\infty$ wird analog bewiesen.

(b) Sei a die rechte Grenze von I, und nehmen wir auch an, dass $b \in \mathbb{R}$ (die Fälle wenn a die linke Grenze von I ist bzw $b = \pm\infty$ werden analog behandelt). Der Mittelwertsatz von Cauchy ergibt folgendes: für alle $y < x$ aus I existiert ein $c \in (y, x)$ mit

$$f'(c)\,(g(x) - g(y)) = g'(c)\,(f(x) - f(y)).$$

Nach Voraussetzung gilt $g'(c) \neq 0$. Auch gilt $g(x) \neq g(y)$, weil anderenfalls nach Satz von Rolle (Satz 3.22) die Ableitung g' eine Nullstelle hat. Dividieren durch $g'(c)\,(g(x) - g(y))$ ergibt

$$\frac{f'(c)}{g'(c)} = \frac{f(x) - f(y)}{g(x) - g(y)} = \frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}},$$

woraus folgt

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)}. \tag{3.56}$$

Nach Voraussetzung gilt

$$\lim_{z \to a} \frac{f'(z)}{g'(z)} = b,$$

d.h. für jedes $\varepsilon > 0$ existiert eine Umgebung U von a mit

$$\left|\frac{f'(z)}{g'(z)} - b\right| < \varepsilon \quad \text{für alle } z \in I \cap U.$$

Für $x, y \in U \cap I$ gilt auch $c \in U \cap I$ woraus folgt

$$b - \varepsilon < \frac{f'(c)}{g'(c)} < b + \varepsilon.$$

Betrachten wir eine Folge $\{x_k\}$ in (y, a) mit $x_k \to a$ für $k \to \infty$ und setzen in (3.56) $x = x_k$. Da $g(x_k) \to \pm\infty$, erhalten wir, dass

$$\frac{g(y)}{g(x_k)} \to 0 \quad \text{und} \quad \frac{f(y)}{g(x_k)} \to 0.$$

Daraus folgt, dass

$$\limsup_{k \to \infty} \frac{f(x_k)}{g(x_k)} \le b + \varepsilon$$

und

$$\liminf_{k \to \infty} \frac{f(x_k)}{g(x_k)} \ge b - \varepsilon.$$

Da diese Ungleichungen für alle $\varepsilon > 0$ gelten, so erhalten wir

$$\limsup_{k \to \infty} \frac{f(x_k)}{g(x_k)} = \liminf_{k \to \infty} \frac{f(x_k)}{g(x_k)} = b,$$

woraus folgt

$$\lim_{x \to a} \frac{f(x)}{g(x)} = b.$$

■

3.13 Höhere Ableitungen

Definition. Sei f eine Funktion auf einem Intervall I. Die Ableitung $f^{(n)}$ der Ordnung $n \in \mathbb{Z}_+$ (oder die n-te Ableitung) wird per Induktion wie folgt definiert:

$$f^{(0)} = f \text{ und } f^{(n)} = \left(f^{(n-1)}\right)' \text{ für jedes } n \geq 1,$$

vorausgesetzt, dass die entsprechenden Ableitungen existieren auf I.

Zum Beispiel, wir haben

$$\begin{aligned} f^{(1)} &= f' \\ f^{(2)} &= (f')' =: f'' \\ f^{(3)} &= (f'')' =: f''' \\ f^{(4)} &= (f''')' =: f^{IV} \end{aligned}$$

usw.

Definition. Die Funktion f heißt n fach differenzierbar auf I falls $f^{(n)}$ existiert auf I (insbesondere müssen auch $f^{(k)}$ existieren für alle $k \leq n$). Die Funktion f heißt unendlich oft differenzierbar auf I, falls $f^{(n)}$ existiert auf I für alle $n \in \mathbb{N}$.

Beispiel. 1. Sei $f = \exp(x)$. Dann $f' = \exp(x)$ und per Induktion erhalten wir, dass

$$(\exp(x))^{(n)} = \exp(x)$$

für alle $n \in \mathbb{N}$. Insbesondere ist $\exp(x)$ unendlich oft differenzierbar.

2. Sei $f = \sin x$. Dann

$$f' = \cos x, \; f'' = -\sin x, \; f''' = -\cos x, \; f^{IV} = \sin x,$$

woraus folgt

$$(\sin x)^{(n)} = \begin{cases} \sin x, & n = 0 \bmod 4 \\ \cos x, & n = 1 \bmod 4 \\ -\sin x, & n = 2 \bmod 4 \\ -\cos x, & n = 3 \bmod 4 \end{cases}$$

3. Sei $f(x) = x^a$ wobei $a \in \mathbb{R}$ und $x > 0$. Dann

$$f' = ax^{a-1}, \quad f'' = a(a-1)x^{a-2}, \quad \text{usw.}$$

Per Induktion erhalten wir

$$(x^a)^{(n)} = a(a-1)\ldots(a-n+1)x^{a-n}.$$

4. Sei $f(x) = x^k$ wobei $k \in \mathbb{N}$ und $x \in \mathbb{R}$. Dann gilt

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

Für $n = k$ erhalten wir

$$\left(x^k\right)^{(k)} = k! = \text{const},$$

woraus folgt, dass $\left(x^k\right)^{(k+1)} \equiv 0$ und $\left(x^k\right)^{(n)} \equiv 0$ für alle $n > k$.

Alle Funktionen in den obigen Beispielen sind unendlich oft differenzierbar.

3.13.1 Taylorformel

Ist f eine differenzierbare Funktion auf einem Intervall I, so gilt nach Lemma 3.15 für jedes $a \in I$
$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \text{ für } x \to a. \tag{3.57}$$
Diese Identität lässt sich betrachten als eine Approximation der Funktion $f(x)$ in der Nähe von a mit der linearen Funktion

$$T(x) = f(a) + f'(a)(x-a),$$

die die Gleichung der Tangente darstellt. In diesem Abschnitt beweisen wir die *Taylorformel*, die eine Approximation der Funktion f mit Hilfe von Polynomen liefert.

Motivation. Betrachten wir ein Polynom

$$f(x) = c_0 + c_1 x + ... + c_n x^n = \sum_{k=0}^{n} c_k x^k, \tag{3.58}$$

mit reellen Koeffizienten c_k und mit $n \in \mathbb{Z}_+$. Der maximale Wert von k mit $c_k \neq 0$ heißt der Grad von f und wird mit $\deg f$ bezeichnet. Offensichtlich gilt $\deg f \le n$.

Die Ableitung von Polynom f ist offensichtlich wieder ein Polynom, und $\deg f' = \deg f - 1$ falls $\deg f \ge 1$. Ist $\deg f = 0$, so gilt $f = \text{const}$ und $f' = 0$.

Lemma 3.32 *Für jedes Polynom f von Grad $\le n$ gilt für alle $a, x \in \mathbb{R}$,*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k. \tag{3.59}$$

Beweis. Induktionsanfang für $n = 0$. Ist $\deg f = 0$ so ist $f = \text{const}$, und die Identität (3.59) ist offensichtlich erfüllt.

Induktionsschritt von $n-1$ nach n. Gilt $\deg f \le n$, so gilt $\deg f' \le n-1$. Nach der Induktionsvoraussetzung haben wir

$$f'(x) = f'(a) + \frac{f''(a)}{1!}(x-a) + \frac{f'''(a)}{2!}(x-a)^2 + ... + \frac{f^{(n)}(a)}{(n-1)!}(x-a)^{n-1}. \tag{3.60}$$

Bezeichnen mit $g(x)$ die rechte Seite von (3.59). Ableiten von g ergibt

$$g'(x) = f'(a) + \frac{f''(a)}{1!}(x-a) + \frac{f'''(a)}{2!}(x-a)^2 + ... + \frac{f^{(n)}(a)}{(n-1)!}(x-a)^{n-1},$$

was zusammen mit (3.60) impliziert die Identität $f'(x) = g'(x)$. Daher $(f-g)' \equiv 0$, und nach Konstantentest (Satz 3.24) $f - g = \text{const}$. Da nach (3.59) $g(a) = f(a)$, so erhalten wir dass $\text{const} = 0$, woraus $f(x) = g(x)$ für alle x folgt. ∎

Bezeichnen wir $x - a = b$ und schreiben (3.59) wie folgt um:

$$f(a+b) = f(a) + \frac{f'(a)}{1!}b + \frac{f''(a)}{2!}b^2 + ... + \frac{f^{(n)}(a)}{n!}b^n.$$

Für $f(x) = x^n$ erhalten wir binomischen Lehrsatz:

$$(a+b)^n = a^n + \frac{na^{n-1}b}{1!} + \frac{n(n-1)a^{n-2}b^2}{2!} + ... + \frac{n!}{n!}b^n = \sum_{l=0}^{n} \binom{n}{l} a^{n-l} b^l.$$

Hauptsatz 3.33 (Taylorformel mit der Restgliedform nach Peano) *Sei $f(x)$ eine n-fach differenzierbare Funktion auf einem Intervall I, wobei $n \in \mathbb{N}$. Dann gilt für jedes $a \in I$*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n + o((x-a)^n) \quad \text{für } x \to a. \tag{3.61}$$

Gilt für $c_0, c_1, ..., c_n \in \mathbb{R}$

$$f(x) = c_0 + c_1(x-a) + ... + c_n(x-a)^n + o((x-a)^n) \quad \text{für } x \to a \tag{3.62}$$

so gilt dann $c_k = \frac{f^{(k)}(a)}{k!}$ für alle $k = 0, 1, ..., n$.

Das Polynom

$$T_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k, \tag{3.63}$$

heißt *Taylor-Polynom* von f der Ordnung n an der Stelle a. Die vollständige Notation von dem Taylor-Polynom ist $T_{n,f}(x;a)$, aber häufig schreibt man $T_n(x)$ wenn es klar ist was f und a sind.

Es folgt aus (3.61), dass

$$f(x) - T_n(x) = o((x-a)^n) \quad \text{für } x \to a. \tag{3.64}$$

Darüber hinaus ist $T_n(x)$ das einzige Polynom des Grades $\leq n$, das (3.64) erfüllt.

Die Differenz $f(x) - T_n(x)$ heißt der *Restglied* der Taylorformel. Die Darstellung des Restgliedes in der Form (3.64) heißt die *Restgliedform nach Peano*.

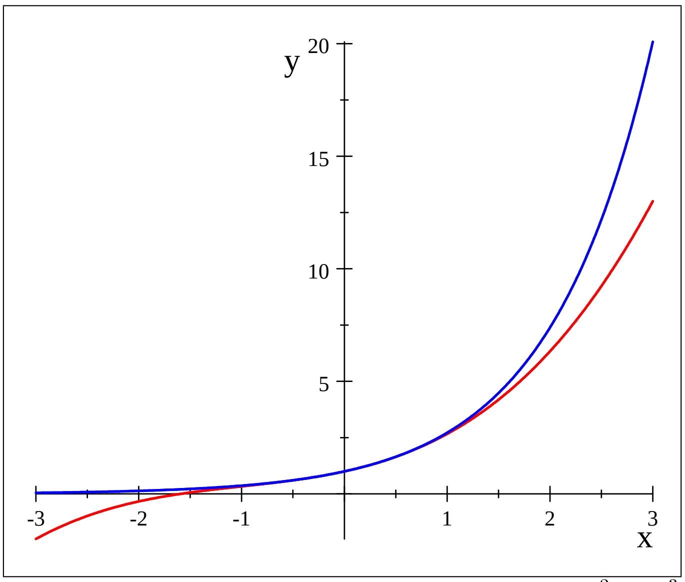
Beispiel. Sei $f(x) = \exp(x)$. Da $f^{(n)}(a) = \exp(a)$ für alle n, erhalten wir aus (3.61)

$$\exp(x) = \exp(a)\left(1 + \frac{(x-a)}{1!} + \frac{(x-a)^2}{2!} + ... + \frac{(x-a)^n}{n!} + o((x-a)^n)\right).$$

Insbesondere für $a = 0$ erhalten wir

$$\exp(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ... + \frac{x^n}{n!} + o(x^n).$$

In diesem Fall ist das Taylor-Polynom $T_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ... + \frac{x^n}{n!}$ identisch zur Partialsumme der Exponentialreihe.



Funktionen $\exp(x)$ (blau) und $T_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$ (rot)

Beispiel. Sei $f(x) = x^p$, wobei $x > 0$ und $p \in \mathbb{R}$. Für $h = x - a$ erhalten wir aus (3.61)

$$(a+h)^p = a^p + \binom{p}{1} a^{p-1} h + \binom{p}{2} a^{p-2} h^2 + ... + \binom{p}{n} a^{p-n} h^n + o(h^n), \tag{3.65}$$

für $h \to 0$, wobei

$$\binom{p}{n} := \frac{p(p-1)...(p-n+1)}{n!}.$$

Zum Beispiel für $p = \frac{1}{2}$ und $n = 1$ haben wir

$$\sqrt{a+h} = \sqrt{a} + \frac{1}{2} \frac{h}{\sqrt{a}} + o(h) \tag{3.66}$$

and für $n = 2$

$$\sqrt{a+h} = \sqrt{a} + \frac{1}{2} \frac{h}{\sqrt{a}} - \frac{1}{8} \frac{h^2}{(\sqrt{a})^3} + o(h^2). \tag{3.67}$$

Für $a = 25$ und $h = 1$ ergibt (3.66)

$$\sqrt{26} = \sqrt{25+1} \approx 5 + \frac{1}{2} \frac{1}{5} = 5,1$$

und (3.67)

$$\sqrt{26} \approx 5 + \frac{1}{2} \frac{1}{5} - \frac{1}{8} \frac{1}{125} = 5,099$$

In der Tat gilt es

$$\sqrt{26} = 5,099019\,51359278... \,.$$

Beweis von Satz 3.33. Beweisen wir (3.64) per Induktion nach n. Induktionsanfang: für $n = 1$ haben wir

$$T_1(x) = f(a) + f'(a)(x-a),$$

und (3.64) gilt nach Lemma 3.15 (oder direkt nach Definition der Ableitung).

Induktionsschritt von $n-1$ nach n. Leiten wir das Taylor-Polynom $T_n(x) = T_{n,f}(x;a)$ ab and erhalten folgendes:

$$\begin{aligned} T_n'(x) &= \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} k (x-a)^{k-1} = \sum_{k=1}^{n} \frac{f^{(k)}(a)}{(k-1)!} (x-a)^{k-1} \\ &= \sum_{l=0}^{n-1} \frac{(f')^{(l)}(a)}{l!} (x-a)^l, \end{aligned}$$

wobei $l = k - 1$. Somit sehen wir, dass $T_n'(x)$ gleich das Taylor-Polynom der Ordnung $n-1$ der Funktion f' ist, d.h.

$$T_{n,f}'(x;a) = T_{(n-1),f'}(x;a). \tag{3.68}$$

Somit erhalten wir nach Induktionsvoraussetzung

$$f'(x) - T_n'(x) = o\left((x-a)^{n-1}\right) \text{ für } x \to a,$$

d.h.

$$\lim_{x\to a}\frac{f'(x)-T_n'(x)}{(x-a)^{n-1}}=0.$$

Nach der Regel von l'Hôspital gilt

$$\lim_{x\to a}\frac{f(x)-T_n(x)}{(x-a)^{n}}=\lim_{x\to a}\frac{f'(x)-T_n'(x)}{n(x-a)^{n-1}}=0,$$

woraus (3.64) folgt.

Für die zweite Aussage bezeichnen wir

$$b_k=\frac{f^{(k)}(a)}{k!}-c_k$$

und bemerken, dass nach (3.61) und (3.62)

$$b_0+b_1(x-a)+...+b_n(x-a)^n=o((x-a)^n)\quad\text{für } x\to a. \tag{3.69}$$

Beweisen wir, dass $b_k=0$ für alle $k=0,...,n$. Nehmen wir das Gegenteil an, dass $b_k\neq 0$ für einige k. Wählen wir den minimalen Wert von k mit $b_k\neq 0$. Da $|(x-a)^n|\leq\left|(x-a)^k\right|$ für $x\to a$, so erhalten wir aus (3.69), dass

$$\frac{b_k(x-a)^k+b_{k+1}(x-a)^{k+1}+...+b_n(x-a)^n}{(x-a)^k}\to 0\text{ für } x\to a,$$

woraus folgt

$$b_k+b_{k+1}(x-a)+...+b_n(x-a)^{n-k}\to 0\text{ für } x\to a.$$

Da der Limes der linken Seite hier gleich b_k ist, so erhalten wir $b_k=0$, was im Widerspruch zur Annahme steht. ■

Beispiel. Bestimmen wir die Taylor-Polynome für $\sin x$ an $a=0$. Wir haben die folgende Reihe für $\sin x$:

$$\sin x=\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}=x-\frac{x^3}{3!}+\frac{x^5}{5!}-...$$

Zeigen wir, dass das Taylor-Polynom $T_{2n+1}(x)$ von $\sin x$ ist gleich die entsprechende Partialsumme:

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

In der Tat gilt es für $|x|\leq 1$

$$\begin{aligned}|\sin x-S_{2n+1}(x)| &= \left|\sum_{k=n+1}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!}\right|\\ &\leq |x|^{2n+3}\sum_{k=n+1}^{\infty}\frac{1}{(2k+1)!}\\ &\leq \text{const}\,|x|^{2n+3}\end{aligned}$$

so dass

$$\sin x - S_{2n+1}(x) = o\left(x^{2n+2}\right) \quad \text{für } x \to 0.$$

Nach dem Eindeutigkeit-Teil des Satzes 3.33 erhalten wir $T_{2n+1} = S_{2n+1}$ und somit

$$\boxed{\sin x = x - \tfrac{x^3}{3!} + \tfrac{x^5}{5!} - ... + (-1)^n \tfrac{x^{2n+1}}{(2n+1)!} + o\left(x^{2n+2}\right)} \quad \text{für } x \to 0.$$

Analog bestimmt man die Taylor-Polynome von $\cos x$. Alternativ kann man die Taylor-Polynome von $\cos x$ als die Ableitungen von Taylor-Polynomen von $\sin x$ erhalten:

$$T_{2n,\cos}(x) = T_{2n+1,\sin}(x)' = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!}$$

woraus folgt

$$\boxed{\cos x = 1 - \tfrac{x^2}{2!} + \tfrac{x}{4!} - ... + (-1)^n \tfrac{x^{2n}}{(2n)!} + o\left(x^{2n+1}\right)} \quad \text{für } x \to 0.$$

Wir erwähnen auch die folgenden Taylor-Formeln die aus dem Satz 3.33 folgen:

$$\boxed{\ln(1+x) = x - \tfrac{x^2}{2} + \tfrac{x^3}{3} - ... + (-1)^{n-1} \tfrac{x^n}{n} + o(x^n)} \quad \text{für } x \to 0$$

$$\boxed{\arctan x = x - \tfrac{x^3}{3} + \tfrac{x^5}{5} - ... + (-1)^n \tfrac{x^{2n+1}}{2n+1} + o\left(x^{2n+2}\right)} \quad \text{für } x \to 0$$

(siehe Aufgaben).

Wir werden unterhalb die Taylorformel verbessern, indem wir eine bessere Abschätzung des Restgliedes $f(x) - T_n(x)$ beweisen.

Hauptsatz 3.34 (Taylor-Formel mit der Restgliedform nach Lagrange) *Sei $f(x)$ eine $(n+1)$-fach differenzierbare Funktion auf einem Intervall I, wobei $n \in \mathbb{Z}_+$. Dann, für alle $a, x \in I$, $x \neq a$, existiert ein $c \in (a, x)$ mit*

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{n+1}(c)}{(n+1)!}(x-a)^{n+1}. \tag{3.70}$$

Für $n = 0$ wird (3.70) wir folgt

$$f(x) = f(a) + f'(c)(x-a),$$

was nicht anderes als Mittelwertsatz von Lagrange (Satz 3.23) ist.

Die Identität (3.70) ergibt die folgende Darstellung des Restgliedes:

$$f(x) - T_n(x) = \frac{f^{n+1}(c)}{(n+1)!}(x-a)^{n+1}, \tag{3.71}$$

die die *Restgliedform nach Lagrange* heißt.

Beweis von dem Satz 3.34. Betrachten wir eine Hilfsfunktion

$$F(t) = f(x) - \left(f(t) + \frac{f'(t)}{1!}(x-t) + \frac{f''(t)}{2!}(x-t)^2 + ... + \frac{f^{(n)}(t)}{n!}(x-t)^n\right) \tag{3.72}$$

wobei der Ausdruck in den Klammern das Taylor-Polynom von f an der Stelle t ist, d.h. $T_{n,f}(x;t)$. Wir betrachten $x, a \in I$ als gegebene Werte, und t als eine Variable im Intervall $[a, x]$. Bemerken wir zunächst folgendes:

$$F(x) = 0 \quad \text{und} \quad F(a) = f(x) - T_n(x),$$

wobei $T_n(x) = T_{n,f}(x;a)$. Leiten wir jetzt jedes Glied in (3.72) in t ab:

$$\begin{aligned}\left(\frac{f^{(k)}(t)}{k!}(x-t)^k\right)' &= -\frac{f^{(k)}(t)}{k!}k(x-t)^{k-1} + \frac{f^{(k+1)}(t)}{k!}(x-t)^k \\ &= -\frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} + \frac{f^{(k+1)}(t)}{k!}(x-t)^k.\end{aligned}$$

Daher erhalten wir

$$F'(t) = -f'(t) + \frac{f'(t)}{0!} - \frac{f''(t)}{1!}(x-t) + \frac{f''(t)}{1!}(x-t) - \frac{f'''(t)}{2!}(x-t)^2 - \ldots - \frac{f^{(n+1)}(t)}{n!}(x-t)^n$$

Alle Glieder in diesem Ausdruck lassen sich wegkürzen, außer des letzten Glied. Somit erhalten wir

$$F'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n. \tag{3.73}$$

Betrachten wir auch die Funktion

$$G(t) = (x-t)^{n+1}$$

für $t \in [a, x]$. Für diese Funktion haben wir

$$G(x) = 0, \qquad G(a) = (x-a)^{n+1}$$

und

$$G'(t) = -(n+1)(x-t)^{n+1}.$$

Anwendung des Mittelwertsatzes von Cauchy 3.31 zu den Funktionen F und G ergibt folgendes: es existiert ein $c \in (a, x)$ mit

$$G'(c)(F(x) - F(a)) = F'(c)(G(x) - G(a)).$$

Einsetzen von den Werten von $F, G, F', G;$ ergibt

$$(n+1)(x-c)^n(f(x) - T_n(x)) = \frac{f^{(n+1)}(c)}{n!}(x-c)^n(x-a)^{n+1}.$$

Da $c \in (a, x)$ und somit $c \neq x$, so dividieren durch $(n+1)(x-c)^n$ ergibt (3.70). ■

Beispiel. Betrachten wir die Funktion $f(x) = \sin x$ und ihre Taylor-Polynome an 0

$$T_3(x) = T_4(x) = x - \frac{x^3}{3!}$$

Nach Satz 3.34 haben wir

$$\sin x - T_4(x) = \frac{f^V(c)}{5!}x^5$$

für ein $c \in (0,x)$. Da $f^V(c) = \cos c$ und $|\cos c| \le 1$, so erhalten wir die Abschätzung des Approximationsfehlers:

$$|\sin x - T_4(x)| \le \frac{x^5}{120}.$$

Zum Beispiel, für $x = 0.1$ erhalten wir

$$\sin 0.1 \approx T_4(0.1) = 0.1 - \frac{0.001}{6} = 0.0998333... ,$$

und der Approximationsfehler ist kleiner gleich

$$\frac{0.1^5}{120} = 8.333... \times 10^{-8} < 10^{-7}.$$

3.13.2 Lokale Extrema

Definition. Sei f eine Funktion auf einem offenen Intervall I und sei $a \in I$. Man sagt, dass a eine *lokale Maximumstelle* von f ist (bzw f hat an der Stelle a ein *lokales Maximum*) falls eine Umgebung $U \subset I$ von a existiert, so dass a eine Maximumstelle von f in U ist, d.h.

$$f(a) = \max_U f(x) \ .$$

Analog definiert man lokale Minimumstelle von f.

Man sagt, dass a eine *lokale Extremumstelle* von f ist, falls a lokale Maximum- oder Minimumstelle ist.

Satz 3.35 (a) (*Notwendige Bedingung für locales Extremum*) *Sei f eine differenzierbare Funktion auf einem offenen Intervall I. Ist $a \in I$ eine Extremumstelle von f, so gilt $f'(a) = 0$.*

(b) (*Hinreichende Bedingung für locales Extremum) Sei f eine 2-fach differenzierbare Funktion auf einem offenen Intervall I. Sei $f'(a) = 0$ für ein $a \in I$. Gilt $f''(a) > 0$ so ist a eine lokale Minimumstelle von f. Gilt $f''(a) < 0$, so ist a eine lokale Maximumstelle von f.*

Bemerkung. Die Nullstellen von der Ableitung f' heißen auch die kritischen Punkte von Funktion f. Nach Satz 3.35 sind die lokalen Extremumstellen auch die kritischen Punkte, aber die Umkehrung gilt nicht immer. Zum Beispiel, die Funktion $f(x) = x^3$ hat nur einen kritischen Punkt $x = 0$, der keine lokale Extremumstelle ist.

Beweis. (a) Ist a eine lokale Maximumstelle, dass ist a eine Maximumstelle von f in einer Umgebung U von a. Nach Satz von Fermat (Satz 3.20) gilt $f'(a) = 0$. Gleiches gilt für eine lokale Minimumstelle.

(b) Nach der Taylor-Formel von Satz 3.33 haben wir

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + R(x), \tag{3.74}$$

wobei $R(x) = o\left((x-a)^2\right)$ für $x \to a$, d.h.

$$\lim_{x \to a} \frac{R(x)}{(x-a)^2} = 0.$$

Ist $f''(a) > 0$, dann gibt es eine Umgebung $U \subset I$ von a mit

$$\left|\frac{R(x)}{(x-a)^2}\right| < \frac{f''(a)}{4} \quad \text{für alle } x \in U \setminus \{a\},$$

woraus folgt

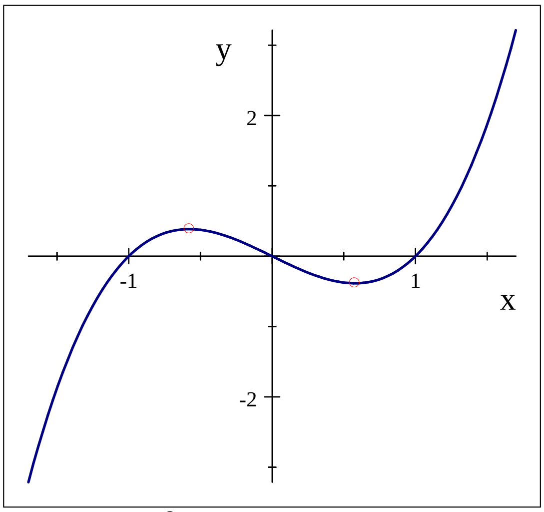
$$R(x) \geq -\frac{f''(a)}{4}(x-a)^2 \text{ für alle } x \in U.$$

Da $f'(a) = 0$, so erhalten wir aus (3.74), dass für alle $x \in U$,

$$\begin{aligned} f(x) &\geq f(a) + \frac{f''(a)}{2}(x-a)^2 - \frac{f''(a)}{4}(x-a)^2 \\ &\geq f(a) + \frac{f''(a)}{4}(x-a)^2 \\ &\geq f(a). \end{aligned}$$

Somit ist a eine Minimumstelle von f in U. Der Fall $f''(a) < 0$ wird analog behandelt. ■

Beispiel. Betrachten wir die Funktion $f(x) = x^3 - x$. Dann hat die Ableitung $f'(x) = 3x^2 - 1$ die Nullstellen $x_1 = \frac{1}{\sqrt{3}}$ und $x_2 = -\frac{1}{\sqrt{3}}$. Da $f''(x) = 6x$, so erhalten wir $f''(x_1) > 0$ und $f''(x_2) < 0$. Deshalb ist x_1 eine lokale Minimumstelle und x_2 – eine lokale Maximumstelle.



Die Funktion $f(x) = x^3 - x$ hat zwei lokale Extremumstellen.

Bemerkung. Hat eine Funktion f auf einem Intervall I keine Extremumstelle, so ist f auf I streng monoton (anderenfalls hätte f zwei stellen $a < b$ in I mit $f(a) = f(b)$, und nach Extremwertsatz existierte eine Extremumstelle in (a, b)). Insbesondere ist f immer monoton zwischen zwei aufeinander folgenden Extremumstellen.

3.13.3 Konvexe und konkave Funktionen

Definition. Eine Funktion $f : I \to \mathbb{R}$ auf einem Intervall I heißt *konvex* falls für alle $a, b \in I$ und $t \in (0,1)$ gilt

$$f((1-t)a + tb) \leq (1-t) f(a) + tf(b). \tag{3.75}$$

Die Funktion f heißt *konkav* auf I falls für alle $a, b \in I$ und $t \in (0,1)$ gilt

$$f((1-t)a + tb) \geq (1-t) f(a) + tf(b). \tag{3.76}$$

Diese Definition hat die folgende geometrische Bedeutung. Bezeichnen wir

$$x = (1-t)a + tb \tag{3.77}$$

und beachten, dass

$$t \in (0,1) \Leftrightarrow x \in (a,b).$$

Es folgt aus (3.77)

$$t = \frac{x-a}{b-a},$$

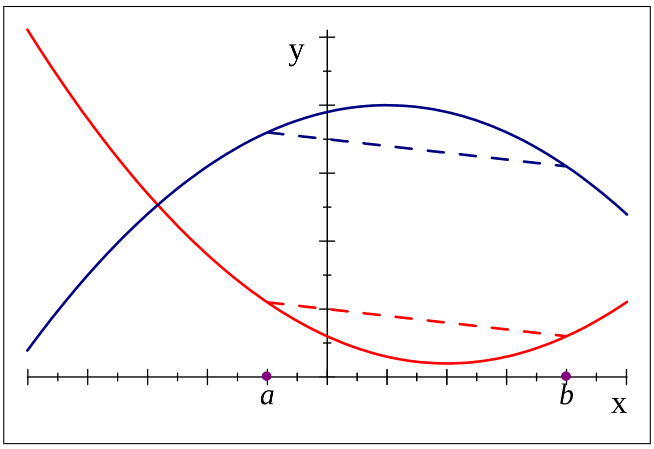
und (3.75) ist äquivalent zu

$$f(x) \leq f(a) + \frac{f(b) - f(a)}{b-a}(x-a).$$

Da die Funktion

$$x \mapsto f(a) + \frac{f(b) - f(a)}{b-a}(x-a)$$

die Sekante durch $(a, f(a))$ und $(b, f(b))$ bestimmt, so ist die Bedingung (3.75) von Konvexität äquivalent zur Bedingung, dass der Graph der Funktion f auf (a,b) unter der Sekante liegt, für alle $a, b \in I$. Analog ist die Funktion f konkav, falls der Graph von f auf (a,b) über der Sekante liegt.



konvex - rot, konkav - blau

Satz 3.36 *Sei f eine 2-fach differenzierbare Funktion auf einem offenen Intervall $I \subset \mathbb{R}$.*

(*a*) *Funktion f ist konvex auf I genau dann, wenn $f'' \geq 0$ auf I*

(*b*) *Funktion f ist konkav auf I genau dann, wenn $f'' \leq 0$ auf I*

Beweis. (a) Sei $f'' \geq 0$ auf I und beweisen wir, dass f konvex ist, d.h. für alle $a, b \in I$ und $t \in [0,1]$ gilt

$$f(x) \leq (1-t) f(a) + t f(b)$$

mit $x = (1-t)a + tb$. Nach dem Satz 3.34 haben wir

$$f(a) = f(x) + f'(x)(a-x) + f''(\alpha) \frac{(a-x)^2}{2}$$

und

$$f(b) = f(x) + f'(x)(b-x) + f''(\beta) \frac{(b-x)^2}{2}$$

für $\alpha, \beta \in (a,b)$. Da $f''(\alpha)$ und $f''(\beta)$ nichtnegativ sind, so erhalten wir daraus

$$\begin{aligned}(1-t) f(a) + t f(b) &\geq (1-t) f(x) + (1-t) f'(x)(x-a) \\ &\quad + t f(x) + t f'(x)(x-b) \\ &= f(x) + f'(x)(x-(1-t)a - tb) \\ &= f(x),\end{aligned}$$

was to beweisen war.

Ist f konvex, so beweisen wir, dass $f'' \geq 0$ auf I. Nach dem Satz 3.33 haben wir für jedes $x \in I$ und $h \to 0$

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2} + o(h^2)$$

und

$$f(x-h) = f(x) - f'(x)h + f''(x)\frac{h^2}{2} + o(h^2)$$

woraus folgt

$$\frac{f(x+h) + f(x-h)}{2} = f(x) + f''(x)\frac{h^2}{2} + o(h^2).$$

Nach Voraussetzung gilt

$$f(x) \leq \frac{f(x+h) + f(x-h)}{2},$$

woraus folgt, dass

$$f''(x)\frac{h^2}{2} + o(h^2) \geq 0$$

und somit

$$f''(x)\frac{1}{2} + o(1) \geq 0.$$

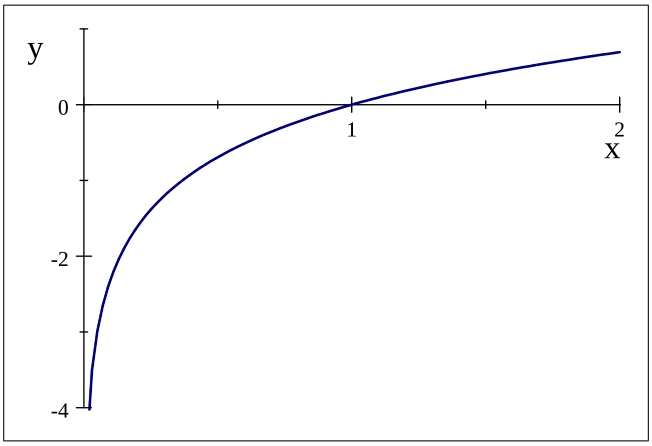
Für $h \to 0$ erhalten wir $f''(x) \geq 0$.

(b) Diese Aussage folgt aus (a) da f konkav genau dann ist, wenn $-f$ konvex. ■

Beispiel. Betrachten wir die Funktion $f(x) = \ln x$ für $x > 0$. Da

$$(\ln x)'' = \left(\frac{1}{x}\right)' = -\frac{1}{x^2} < 0,$$

so erhalten wir nach Satz 3.36, dass $\ln x$ konkav ist.



$\ln x$ ist konkav

Nach (3.76) gilt die folgende Ungleichung

$$\ln\left((1-t)\,a+tb\right) \geq (1-t)\ln a + t\ln b \tag{3.78}$$

für alle $a, b > 0$ und $t \in (0,1)$. Bezeichnen wit $p = \frac{1}{1-t}$ und $q = \frac{1}{t}$, so dass $p, q > 1$ und

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{3.79}$$

Es folgt aus (3.78), dass

$$\ln\left(\frac{a}{p} + \frac{b}{q}\right) \geq \frac{1}{p}\ln a + \frac{1}{q}\ln b = \ln\left(a^{1/p} b^{1/q}\right). \tag{3.80}$$

Für $x = a^{1/p}$, $y = b^{1/q}$ erhalten wir aus (3.80) die *Young-Ungleichung*

$$\frac{x^p}{p} + \frac{y^q}{q} \geq xy,$$

die für alle $x, y \geq 0$ and $p, q > 1$ mit (3.79) gilt.

Beispiel. Zeigen wir: ist f eine positive konkave Funktion, so ist $\frac{1}{f}$ konvex. Wir haben

$$\left(\frac{1}{f}\right)' = \frac{-f'}{f^2}$$

und

$$\left(\frac{1}{f}\right)'' = -\left(\frac{f'}{f^2}\right)' = -\frac{f'' f^2 - (f')^2}{f^4} = -\frac{f''}{f^2} + \frac{(f')^2}{f^4}.$$

Da $f'' \leq 0$, daraus folgt $\left(\frac{1}{f}\right)'' \geq 0$ und somit die Konvexität von f.

Die Umkehrung gilt nicht: die Funktion x^2 ist konvex, und die Funktion $\frac{1}{x^2}$ auf $(0, +\infty)$ ist auch konvex.