Analysis I

Alexander Grigoryan
Universität Bielefeld

SS 2018

Contents

Chapter 1

Mengen und reelle Zahlen

13.04.18

Das Hauptziel von diesem Kurs Analysis I/II sind *Differentialrechnung* und *Integralrechnung*. Diese sind mathematische Werkzeuge für Untersuchung von *Funktionen*, die in Anwendungen in Physik, Technik und Wirtschaft weit benutzt werden. Davor erforschen wir die Begriffe von Funktionen und Zahlen. Der Begriff von *reeller Zahl* wird in diesem Kurs axiomatisch eingeführt. Aber ganz am Anfang beschäftigen wir uns mit dem Begriff von *Menge*. Der Zweck davon ist, die auf der Mengenlehre basierende Sprache der Mathematik zu lernen.

1.1 Mengen und Operationen auf den Mengen

Elemente von Mengen. In Mathematik arbeitet man mit verschiedenen Objekten. Aus Objekten macht man Mengen.

> Eine *Menge* ist eine Sammlung von anderen Objekten, die selbst als ein Objekt betrachtet wird.

Somit besteht jede Menge M aus bestimmten Objekten, die die *Elemente* von M heißen. Ist x ein Element von M, so schreibt man

$$x \in M$$

("x ist Element von M", "x gehört zu M", "x ist in M", "x liegt in M", "M enthält x"). 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.

Für Mengen benutzt man häufig eine graphische Darstellung. Man zeigt eine Menge als eine Figure auf der Ebene und ihre Elemente – als die Punkte von der Figure.

Eine Menge kann explizit angegeben werden wie folgt. Zum Beispiel, die Menge M, die aus den Elementen (Buchstaben) 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. Zum Beispiel, $a \in M$ während $e \notin M$. Noch ein Beispiel: die Menge $M = \{a\}$ besteht nur aus einem Element a.

Die Elemente dürfen selber 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

$$A \subset B \qquad (\text{oder } A \subseteq B)$$

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

Zum Beispiel, es gilt immer $\emptyset \subset A$.

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

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

wobei das Zeichen $\Rightarrow$ (der Pfeil) 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:

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

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

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

wobei das Zeichen $\wedge$ (der Keil) bedeutet "und".

Beweis. Da $A \subset B$, so gilt

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

und nach $B \subset C$ gilt

$$x \in B \Rightarrow x \in C.$$

Es folgt, dass

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

und deshalb $A \subset C$. ■

[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.

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*:*

$$A \cap B = B \cap A \quad \text{und} \quad A \cup B = B \cup A.$$

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*:*

$$(A \cap B) \cap C = A \cap (B \cap C) \quad \text{und} \quad (A \cup B) \cup C = A \cup (B \cup C)\,.$$

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

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

$$x \in A \cup B \cup C \Leftrightarrow x \in A \lor x \in B \lor x \in C.$$

Behauptung. (Distributivgesetze) *Es gelten die Identitäten*

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C) \tag{1.1}$$

und

$$(A \cup B) \cap C = (A \cap C) \cup (A \cap C) \tag{1.2}$$

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

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

und

$$x \in (A \cup C) \cap (B \cup C) \Leftrightarrow (x \in A \lor x \in C) \land (x \in B \lor x \in C).$$

Es bleibt zu zeigen, dass

$$(x \in A \land x \in B) \lor x \in C \Leftrightarrow (x \in A \lor x \in C) \land (x \in B \lor x \in C). \tag{1.3}$$

Gilt $x \in C$, so sind die beiden Seiten von (1.3) wahr.

Gilt $x \notin C$, so erhalten wir

$$(x \in A \land x \in B) \lor x \in C \Leftrightarrow x \in A \land x \in B$$

und

$$(x \in A \lor x \in C) \land (x \in B \lor x \in C) \Leftrightarrow (x \in A) \land (x \in B),$$

woraus (1.3) folgt. ■

Subtraktion von Mengen.

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

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

Eine äquivalente Definition ist

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

Das Zeichen $\setminus$ (umgekehrter Schrägstrich) heißt “Mengenminus” oder einfach “Minus”.

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 $\mathcal{P}(X)$ (oder 2^X) bezeichnet. d.h.

$$A \in \mathcal{P}(X) \Leftrightarrow A \subset X.$$

In anderer Wörter die Elementen von $\mathcal{P}(X)$ sind die Teilmengen von X. Die Operationen $\cup$, $\cap$, $\setminus$ mit den Elementen von $\mathcal{P}(X)$ ergeben offensichtlich wieder die Elemente von $\mathcal{P}(X)$.

Es gilt immer $\emptyset \in \mathcal{P}(X)$ und $X \in \mathcal{P}(X)$. Betrachten wir ein Beispiel: für $X = \{a, b\}$ gilt

$$\mathcal{P}(X) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}.$$

Komplement. Für die Elementen von $\mathcal{P}(X)$ gibt es noch eine Operation, die "Komplement" heißt.

Definition. Für jede Menge $A \in \mathcal{P}(X)$ 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 $A, B \in \mathcal{P}(X)$:*

$$A \setminus B = A \cap B^c \tag{1.4}$$
$$(A \cap B)^c = A^c \cup B^c \tag{1.5}$$
$$(A \cup B)^c = A^c \cap B^c. \tag{1.6}$$

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 (1.4) folgt. 18.04.18

Um die zweite Identität (1.5) 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

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

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

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

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

Daher

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

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

Analog lässt sich auch die dritte Identität (1.6) beweisen. Alternativ beweist man (1.6) mit Hilfe von (1.5) und der Identität $(A^c)^c = A$ wie folgt:

$$(A^c \cap B^c)^c = (A^c)^c \cup (B^c)^c = A \cup B,$$

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

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

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

Es folgt daraus, dass

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

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

Für symmetrische Differenz gelten die folgenden Identitäten.

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

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

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

(siehe Aufgabe 4).

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 Tripel (x, y, z), indem wir annehmen, dass

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

Dann gilt für kartesisches Produkt das Assoziativgesetz:

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

Darüber hinaus definieren wir kartesisches Produkt $A \times B \times C$ dreier Mengen A, B, C mit

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

und erhalten, dass

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

Kartesisches Produkt erfüllt auch das Distributivgesetz:

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

wobei $\star$ irgendeine Mengenoperation $\cap, \cup, \setminus, \triangle$ bezeichnet (siehe Aufgabe 19).

1.2 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 wie folgt bezeichnet:

$$f : X \to Y$$

oder

$$X \xrightarrow{f} Y.$$

Ist $y \in Y$ dem Element $x \in X$ zugeordnet, so heißt y 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. Unterhalb benutzen wir die folgenden Symbolen (*Quantoren*):

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

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

$\exists!$ "es gibt genau sein".

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

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:

$$\forall x \in X \quad \exists! y \in Y \text{ mit } (x, y) \in G. \tag{1.7}$$

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, so dass $\mathrm{Id}_X(x) = x \ \forall x \in X$. 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 $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ 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, und somit die Zuordnung

$$A \mapsto f^{-1}(A)$$

bestimmt eine Abbildung von $\mathcal{P}(Y)$ nach $\mathcal{P}(X)$. Die Abbildung

$$\begin{aligned} f^{-1} : \mathcal{P}(Y) &\to \mathcal{P}(X) \\ A &\mapsto f^{-1}(A) \end{aligned}$$

heißt die *Urbildabbildung* von f.

Satz 1.2 *Die Urbildabbildung $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ 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. ■

20.04.18

Komposition von Abbildungen. Seien X, Y, Z beliebige Mengen.

Definition. Gegeben seien zwei Abbildungen

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

die *Komposition (Verkettung, zusammengesetzte Abbildung)* von f und g ist eine Abbildung $f \circ g$ von X nach Z die wie folgt definiert ist:

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

In anderen Wörter, wir haben

$$\begin{aligned} f \circ g : X &\to Z \\ x &\mapsto f(g(x)). \end{aligned}$$

Schematisch kann man die Komposition so darstellen:

$$\begin{tikzcd} X \arrow[r, "g"] \arrow[rr, bend right, "f\circ g"'] & Y \arrow[r, "f"] & Z \end{tikzcd}$$

Beispiel. Für jede Abbildung $g : X \to Y$ gilt

$$g \circ \mathrm{Id}_X = g$$

da für jedes $x \in Y$

$$g \circ \mathrm{Id}_X (x) = g (\mathrm{Id}_X (x)) = g (x) .$$

Analog gilt

$$\mathrm{Id}_Y \circ g = g.$$

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 1.3 (Assoziativgesetz für Komposition) *Für je drei Abbildungen*

$$X \overset{h}{\to} Y \overset{g}{\to} Z \overset{f}{\to} U$$

gilt die Identität

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

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:

$$\begin{tikzcd} X \arrow[r, "h"] \arrow[rrr, bend left=50, "(f\circ g)\circ h"] \arrow[rr, bend right, "g\circ h"'] \arrow[rrr, bend right=50, "f\circ(g\circ h)"'] & Y \arrow[r, "g"] \arrow[rr, bend left, "f\circ g"] & Z \arrow[r, "f"] & U \end{tikzcd}$$

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:

$$\begin{tikzcd} X \arrow[r, "h"] \arrow[rrr, bend right, "f\circ g\circ h"'] & Y \arrow[r, "g"] & Z \arrow[r, "f"] & U \end{tikzcd}$$

Satz 1.4 *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}:

$$\begin{tikzcd} \mathcal{P}(X) & \mathcal{P}(Y) \arrow[l, "g^{-1}"'] & \mathcal{P}(Z) \arrow[l, "f^{-1}"'] \arrow[ll, bend left, "g^{-1}\circ f^{-1}"] \end{tikzcd}$$

Somit ist $g^{-1} \circ f^{-1}$ wohldefiniert als eine Abbildung von $\mathcal{P}(Z)$ nach $\mathcal{P}(X)$. Es folgt aus $X \xrightarrow{f \circ g} Z$, dass auch

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

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

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

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

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

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

Diese Kompositionen werden am folgenden Diagramm gezeigt:

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & Y \arrow[r, "g"] \arrow[rr, bend left, "f\circ g"] & X \arrow[r, "f"] & Y \end{tikzcd}$$

Definition. Die Abbildung g heißt *Umkehrabbildung (inverse Abbildung, Umkehrfunktion, inverse Funktion)* von f, falls

$$f \circ g = \mathrm{Id}_Y \quad \text{und} \quad g \circ f = \mathrm{Id}_X .$$

In diesem Fall ist f auch die Umkehrabbildung von g.

Die Umkehrabbildung ist eindeutig bestimmt: gibt es zwei Umkehrabbildungen g_1 und g_2 von f, so erhalten wir

$$g_1 = g_1 \circ \mathrm{Id}_Y = g_1 \circ (f \circ g_2) = (g_1 \circ f) \circ g_2 = \mathrm{Id}_X \circ g_2 = g_2.$$

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

$$\forall y \in Y \quad \exists! x \in X \quad \text{mit} \quad f(x) = y. \tag{1.9}$$

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

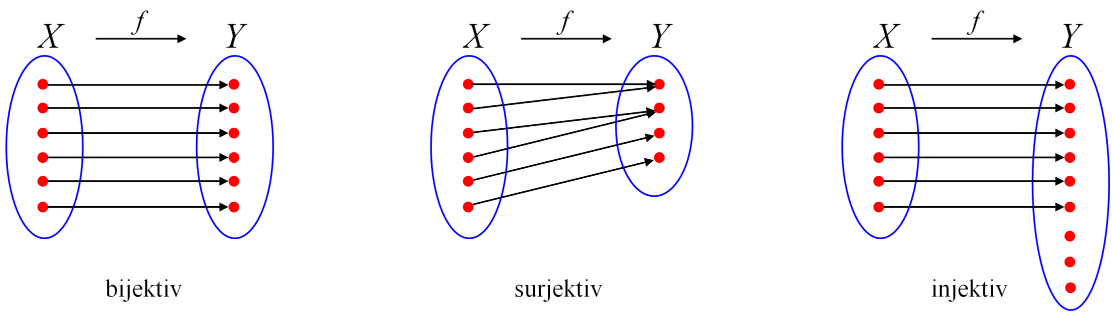
$$\forall y \in Y \quad \exists x \in X \quad \text{mit} \quad f(x) = y.$$

Die Abbildung f heißt *injektiv* falls

$$x_1 \neq x_2 \Rightarrow \; f(x_1) \neq f(x_2).$$

In anderen Wörter, f ist injektiv falls

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



Es folgt, dass eine Abbildung f bijektiv genau dann ist, wenn f surjektiv und injektiv ist. In der Tat, die Existenz von x in (1.9) folgt aus der Surjektivität, und die Eindeutigkeit von x – aus der Injektivität.

Satz 1.5 *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.

$$f\left(g\left(y\right)\right) = y \quad \forall y \in Y \tag{1.10}$$

und

$$g\left(f\left(x\right)\right) = x \;\; \forall x \in X. \tag{1.11}$$

Es folgt aus (1.10), dass $f\left(x\right) = y$ für $x = g\left(y\right)$ erfüllt ist. Somit ist f surjektiv. Zeigen wir jetzt, dass f injektiv ist. Gilt $f\left(x_1\right) = f\left(x_2\right)$, so erhalten wir aus (1.11)

$$x_1 = g\left(f\left(x_1\right)\right) = g\left(f\left(x_2\right)\right) = x_2.$$

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 $f\left(x\right) = y$, so setzen wir $g\left(y\right) = x$. Dann gelten

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

und

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

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 $f^{-1} : \mathcal{P}\left(Y\right) \to \mathcal{P}\left(X\right)$ nicht verwechseln, obwohl sie identisch bezeichnet werden.*

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

Satz 1.6 *Gegeben seien zwei bijektive Abbildungen $X \xrightarrow{g} Y \xrightarrow{f} Z$. Dann ist $f \circ g$ auch bijektiv und die folgende Identität gilt für die Umkehrabbildungen*

$$\left(f \circ g\right)^{-1} = g^{-1} \circ f^{-1}. \tag{1.12}$$

Beweis. Wir haben $X \xleftarrow{g^{-1}} Y \xleftarrow{f^{-1}} Z$ und

$$f \circ g : X \to Z \;\text{ und }\; g^{-1} \circ f^{-1} : Z \to X.$$

Mit Hilfe von dem Satz 1.3 erhalten wir

$$\left(g^{-1} \circ f^{-1}\right) \circ \left(f \circ g\right) = g^{-1} \circ \left(f^{-1} \circ f\right) \circ g = g^{-1} \circ \mathrm{Id}_Y \circ g = g^{-1} \circ g = \mathrm{Id}_X$$

und analog

$$\left(f \circ g\right) \circ \left(g^{-1} \circ f^{-1}\right) = \mathrm{Id}_Z\,.$$

Somit ist $g^{-1} \circ f^{-1}$ die Umkehrabbildung von $f \circ g$. Folglich ist $f \circ g$ bijektiv nach dem Satz 1.5. ■

1.3 Axiomensystem von reellen Zahlen

25.04.18

Hier definieren wir axiomatisch die *Menge von reellen Zahlen*.

Definition. Eine Menge $\mathbb{R}$ heißt die Menge von reellen Zahlen und ihre Elemente heißen reelle Zahlen falls die folgenden vier Gruppen von Axiomen (insgesamt vierzehn Axiome) erfüllt sind.

I. Axiome der Addition. Es gibt eine Abbildung (Operation)

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

namens *Addition* mit den folgenden Eigenschaften.

1. (Das Nullelement) Es existiert eine Zahl $0 \in \mathbb{R}$, so dass für alle $x \in \mathbb{R}$
$$x + 0 = 0 + x = x.$$

2. (Das Negative) Für jedes $x \in \mathbb{R}$ existiert eine Zahl $-x \in \mathbb{R}$ (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
$$x + y = y + x.$$

Die Zahl $x + y$ heißt die *Summe* von x, y.

Jede Menge K, wo die Operation Addition definiert ist und die Axiome 1-3 erfüllt, heißt (additive) *Gruppe*. Soll auch das Axiom 4 erfüllt werden, so heißt die Gruppe K kommutativ.

II. Axiome der Multiplikation. Es gibt eine Abbildung (Operation)

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

namens *Multiplikation* mit den folgenden Eigenschaften.

1. (Das Einheitselement) Es existiert eine Zahl $1 \in \mathbb{R} \setminus \{0\}$, so dass für alle $x \in \mathbb{R}$
$$x \cdot 1 = 1 \cdot x = x.$$

2. (Das Inverse) Für jedes $x \in \mathbb{R} \setminus \{0\}$ existiert eine Zahl $x^{-1} \in \mathbb{R}$ (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
$$(x + y) \cdot z = x \cdot z + y \cdot z.$$

Die Zahl $x \cdot y$ heißt das *Produkt* von x, y. Man schreibt auch xy anstatt $x \cdot y$.

Jede Menge K, 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. Es gibt auch andere Beispiele von Körper, die wir später besprechen.

III. Anordnungsaxiome. Auf $\mathbb{R}$ ist *Ungleichung* $<$ definiert: für je zwei Elemente $x, y \in \mathbb{R}$ ist der Ausdruck $x < y$ eine Aussage, die entweder wahr oder falsch ist (wobei $x < y$ und $y > x$ äquivalent sind). Die Ungleichung erfüllt die folgenden Bedingungen, für alle $x, y, z \in \mathbb{R}$:

1. (Vergleichbarkeit) Es gilt genau eine der folgenden Relationen: $x < y$ oder $y < x$ oder $x = y$.
2. (Transitivität)
$$x < y \ \wedge \ y < z \ \Rightarrow x < z$$
3. (Beziehung zur Addition)
$$x < y \ \Rightarrow x + z < y + z$$
4. (Beziehung zur Multiplikation)
$$x > 0 \ \wedge \ y > 0 \ \Rightarrow x \cdot y > 0.$$

Eine Relation $<$ auf einer Menge K heißt (totale) *Ordnung* falls sie die Anordnungsaxiome 1-2 erfüllt. In diesem Fall heißt die Menge K (total) geordnet.

Die Axiome 3 und 4 etablieren 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.

Wir definieren auch die unechte Ungleichung: $x \leq y$ (oder $y \geq x$) gilt genau dann, when entweder $x < y$ oder $x = y$ gilt, d.h.
$$x \leq y \Leftrightarrow x < y \ \vee \ x = y.$$

IV. Vollständigkeitsaxiom. Seien A, B nichtleere Teilmengen von $\mathbb{R}$ mit der Eigenschaft
$$\forall a \in A \ \ \forall b \in B \ \text{ gilt } a \leq b.$$

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

$$\forall a \in A \ \ \forall b \in B \ \text{ gilt } a \leq c \leq b.$$

Man sagt, dass die Zahl c die Mengen A und B trennt.

Man stellt die reellen Zahlen dar als die Punkte auf einer waagerechten Gerade. Die Punkte am links sind immer kleiner als die Punkte am rechts. Die Vollständigkeitsaxiom bedeutet folgendes: liegt die ganze Menge A links von B, so existiert ein Punkt c zwischen A und B. Man kann es auch so vorstellen, dass die Gerade *keine Lücke* enthält.

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

1.4 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.

[1] *Das Nullelement ist eindeutig bestimmt.*

Seien 0 und $0'$ zwei Nullelemente. Nach Definition erfüllen 0 und $0'$ die folgenden Identitäten, für alle $x \in \mathbb{R}$:

$$x + 0 = 0 + x = x$$

und

$$x + 0' = 0' + x = x.$$

Einsetzen in der ersten Identität $x = 0'$ ergibt

$$0' + 0 = 0' + 0 = 0'$$

und in der zweiten Identität $x = 0$ ergibt

$$0 + 0' = 0' + 0 = 0,$$

woraus $0' = 0$ offensichtlich folgt.

[2] *Das Negative von $x \in \mathbb{R}$ ist eindeutig bestimmt.*

Seien y und z zwei Negative von x. Nach Definition erfüllen y und z die folgenden Identitäten:

$$x + y = y + x = 0$$

und

$$x + z = z + x = 0.$$

Es folgt nach Axiom I.1

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

und nach Axiomen I.3 und I.1

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

woraus $y = z$ folgt.

[3] *Es gelten*

$$-0 = 0$$

und

$$-(-x) = x,$$

für alle $x \in \mathbb{R}$.

Da nach Axiom I.1 gilt $0 + 0 = 0$, so sehen wir, dass 0 die Definition von -0 erfüllt. Nach [2] beschließen wir, dass $-0 = 0$. Bezeichnen wir mit y das Negative von $-x$, d.h. y erfüllt

$$(-x) + y = y + (-x) = 0. \tag{1.13}$$

Da nach Definition von $-x$ gilt

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

so folgt es, dass die Identitäten (1.13) für $y = x$ erfüllt sind. Nach der Eindeutigkeit des Negatives erhalten wir, dass das Negative von $(-x)$ gleich x ist.

[4] *Für jede* $a, b \in \mathbb{R}$ *hat die Gleichung* $x+a = b$ *eine eindeutige Lösung* $x = b+(-a)$.

Die Zahl $x = b + (-a)$ 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

$$\begin{aligned} (x + a) + (-a) &= b + (-a)\,, \\ x + (a + (-a)) &= b + (-a)\,, \\ x &= b + (-a)\,. \end{aligned}$$

Definition. 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. Die Beweise der folgenden Eigenschaften [5] – [8] sind analog zu [1] – [4].

[5] *Das Einheitselement* 1 *ist eindeutig bestimmt.*

[6] *Das Inverse von* $x \in \mathbb{R} \setminus \{0\}$ *ist eindeutig bestimmt.*

[7] *Es gelten* $1^{-1} = 1$ *und* $(x^{-1})^{-1} = x$, *für alle* $x \in \mathbb{R} \setminus \{0\}$.

[8] *Für jede* $a \in \mathbb{R} \setminus \{0\}$ *und* $b \in \mathbb{R}$ *hat die Gleichung* $a \cdot x = b$ *eine eindeutige Lösung* $x = a^{-1} \cdot b$ *(siehe Aufgabe 21).*

Definition. Das Produkt $a^{-1} \cdot b$ 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.* Insbesondere gilt $a^{-1} = 1/a$.

Folgerungen aus dem Distributivgesetz.

[9] $x \cdot 0 = 0 \cdot x = 0$ *für alle* $x \in \mathbb{R}$

Da $0 + 0 = 0$, so erhalten wir aus dem Axiom II.5

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

Setzen wir $a = x \cdot 0$ so dass

$$x \cdot 0 + a = a,$$

und erhalten nach [4], dass

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

[10] $x \cdot y = 0 \Leftrightarrow x = 0 \lor y = 0.$

Ist $x = 0$ oder $y = 0$, so gilt $x \cdot y = 0$ nach [9]. Beweisen wir, dass $x \cdot y = 0$ ergibt $x = 0$ oder $y = 0$. Nehmen wir an, dass $x \neq 0$. Lösen der Gleichung $x \cdot y = 0$ bezüglich y mit Hilfe von [8] ergibt $y = x^{-1} \cdot 0 = 0$, wobei wir [9] benutzt haben.

[11] $(-1) \cdot x = -x$

Mit Hilfe von Axiomen II.1, II.5, I.2 erhalten wir

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

wobei die letzte Identität nach [9] gilt. Somit erfüllt $(-1) \cdot x$ die Definition des Negatives von x, und nach [2] beschließen wir, dass $(-1) \cdot x = -x$.

[12] $(-1) \cdot (-x) = x$

Wir haben nach [11] und [3], dass

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

[13] $-(x + y) = -x - y$ (siehe Aufgabe 21).

[14] $x \cdot (-y) = -(x \cdot y)$

Nach [11] und Axiomen II.3, II.4 erhalten wir

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

[15] $(-x) \cdot (-y) = x \cdot y$. *Insbesondere* $(-1) \cdot (-1) = 1$.

Einsetzen in [14] $(-x)$ statt x ergibt

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

Für $x = y = 1$ erhalten wir

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

[16] *Bezeichnen wir* $2 = 1 + 1$, $3 = 2 + 1$ *und*

$$x^2 = x \cdot x, \quad x^3 = x^2 \cdot x = x \cdot x \cdot x.$$

Dann gilt für alle $x, y \in \mathbb{R}$

$$(x+y)^2 = x^2 + 2 \cdot x \cdot y + y^2$$

und

$$(x+y)^3 = x^3 + 3 \cdot x^2 \cdot y + 3 \cdot x \cdot y^2 + y^3$$

(siehe Aufgabe 26).

Bemerkung. Obwohl wir die Zahl 2 definiert haben, es ist noch nicht klar ob 2 von 0 und 1 abweicht (bemerken wir, dass $1 \neq 0$ nach Axiom II.1 gilt). Die Körperaxiome allein implizieren die Existenz von Zahlen außer 0 und 1 nicht. 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. Dieser Körper with mit $\mathbb{F}_2$ bezeichnet. In $\mathbb{F}_2$ gilt offensichtlich $2 = 0$ (siehe Aufgabe 27).

Dass in $\mathbb{R}$ gilt $2 \neq 0$ ist eine Folgerung von Anordnungsaxiomen unterhalb.

1.5 Folgerungen aus den Anordnungsaxiomen

[17] $x < y \ \wedge \ y \leq z \ \Rightarrow \ x < z$ *und* $x \leq y \ \wedge \ y < z \ \Rightarrow \ x < z$

Ist $y < z$, so folgt die erste Aussage aus dem Axiom III.2. Ist $y = z$ so ist die Implikation trivial. Die zweite Aussage wird analog bewiesen.

[18] $x \leq y \ \wedge \ y \leq z \ \Rightarrow \ x \leq z$

Im Fall $x = y = z$ ist die Implikation trivial. Im Fall $x < y$ oder $y < z$ folgt sie aus [17].

[19] $x \leq y \ \wedge y \leq x \ \Rightarrow \ x = y$

Nach Axiom III.1 gilt immer genau eine von drei Möglichkeiten $x < y$, $x > y$, $x = y$. Da $x < y$ im Widerspruch zu $y \leq x$ steht und $x > y$ – im Widerspruch zu $x \leq y$, so bleibt es nur die Möglichkeit $x = y$.

[20] $x \leq y \ \Rightarrow x + z \leq y + z$

Im Fall $x < y$ folgt aus dem Axiom III.3, dass $x + z < y + z$, im Fall $x = y$ erhalten wir $x + z = y + z$.

[21] $x \leq y \ \wedge \ a < b \ \Rightarrow x + a < y + b$

Mit Hilfe von [20] und Axiom III.3 erhalten wir

$$x + a \leq y + a = a + y < b + y = y + b,$$

woraus $x + a < y + b$ nach [17] folgt.

[22] $x \leq y \ \wedge \ a \leq b \ \Rightarrow x + a \leq y + b$

Im Fall $x < y$ oder $a < b$ erhalten wir aus [21], dass $x + a < y + b$, im Fall $x = y$ und $a = b$ gilt $x + a = y + b$. 27.04.18

Definition. Eine Zahl $x \in \mathbb{R}$ heißt *positiv* falls $x < 0$ und *negativ* falls $x < 0$.

Das Axiom III.1 ergibt, dass jede Zahl $x \in \mathbb{R}$ entweder positiv, oder negativ, oder Null ist.

[23] *Die Summe von positiven Zahlen ist positiv, und die Summe von negativen Zahlen – negativ.*

Falls $x > 0$ und $y > 0$, so erhalten wir nach [21], dass $x + y > 0 + 0 = 0$. Der Fall von negativen x, y ist analog.

[24] *Die folgenden Äquivalenzen gelten:*

$$x < y \ \Leftrightarrow \quad y - x > 0 \ \Leftrightarrow \ -x > -y \tag{1.14}$$

und

$$x \leq y \ \Leftrightarrow \quad y - x \geq 0 \ \Leftrightarrow \ -x \geq -y \ . \tag{1.15}$$

Addieren $(-x)$ zu den beiden Seiten von $x < y$ ergibt nach Axiom III.3

$$x < y \ \Leftrightarrow \ x + (-x) < y + (-x) \ \Leftrightarrow \ 0 < y - x.$$

Analog erhalten wir

$$-x > -y \ \Leftrightarrow \ (-x) + y > (-y) + y \ \Leftrightarrow \ y - x > 0$$

Der Beweis von (1.15) ist analog.

[25] *Für alle $x \in \mathbb{R}$ gelten*

$$x \text{ negativ } \Leftrightarrow -x \text{ positiv}, \tag{1.16}$$

$$x \text{ positiv } \Leftrightarrow -x \text{ negativ}. \tag{1.17}$$

Wir benutzen [24]. Setzen wir in (1.14) $y = 0$ und erhalten

$$x < 0 \Leftrightarrow -x > 0,$$

was äquivalent zu (1.16) ist. Setzen wir in (1.15) $x = 0$ und erhalten

$$0 < y \Leftrightarrow 0 > -y,$$

was äquivalent zu (1.17) ist.

[26] *Sind die Zahlen x und y gleichzeitig positiv oder negativ, so ist $x \cdot y$ positiv. Ist eine Zahl von x, y positiv und andere negativ, so ist $x \cdot y$ negativ.*

Im Fall $x, y > 0$ ergibt das Axiom III.4, dass $x \cdot y > 0$. Im Fall $x, y < 0$ erhalten wir nach [25] dass $-x$ und $-y$ positiv sind, und nach [15] und Axiom III.4

$$x \cdot y = (-x) \cdot (-y) > 0.$$

Im Fall $x > 0$ und $y < 0$, erhalten wir nach [25] dass $-y > 0$ und nach [14] und Axiom III.4 dass

$$-(x \cdot y) = x \cdot (-y) > 0,$$

woraus nach [25] folgt $x \cdot y < 0$. Der Fall $x < 0$ und $y > 0$ ist analog.

[27] *Für alle $x \in \mathbb{R} \setminus \{0\}$ gilt $x^2 > 0$.*

Da $x \neq 0$, so ist x entweder positiv oder negativ (Axiom III.1), und in den beiden Fällen gilt nach [26] dass $x^2 = x \cdot x > 0$.

[28] $1 > 0$ *und* $-1 < 0$.

Nach Axiom II.1 haben wir $1 \neq 0$ und $1 = 1 \cdot 1$. Da $1 \cdot 1 > 0$ nach [27], so folgt es, dass $1 > 0$. Es folgt aus [25], dass $-1 < 0$.

Definieren wir jetzt die Zahlen $2 = 1 + 1$, $3 = 2 + 1$, $4 = 3 + 1$ usw. Es folgt aus $1 > 0$ mit Hilfe von Axiom III.3, dass $2 > 1$, $3 > 2$, $4 > 3$, d.h.

$$0 < 1 < 2 < 3 < 4 < \dots . \tag{1.18}$$

[29] *Ist $x > 0$, so ist $x^{-1} > 0$. Ist $x < 0$ so ist $x^{-1} < 0$.*

Da $x \cdot x^{-1} = 1 > 0$, so gilt $x^{-1} \neq 0$ nach [9]. Nach [26] sind die beiden Zahlen x und x^{-1} gleichzeitig positiv oder negativ, was zu beweisen war.

[30] *Sei $x < y$. Für alle $a > 0$ gilt $a \cdot x < a \cdot y$, und für alle $a < 0$ gilt $a \cdot x > a \cdot y$.*

Die Bedingung $x < y$ ergibt nach [24] dass $y - x > 0$. Falls $a > 0$, so erhalten wir nach Axiomen II.5 und III.4

$$a \cdot y - a \cdot x = a \cdot (y - x) > 0,$$

woraus $a \cdot x < a \cdot y$ folgt. Im Fall $a < 0$ erhalten wir $(-a) > 0$ und somit

$$-(a \cdot x) = (-a) \cdot x < (-a) \cdot y = -(a \cdot y),$$

woraus folgt $-(a \cdot x) < -(a \cdot y)$ und nach [24] auch $a \cdot x > a \cdot y$.

[31] *Falls $0 < x < y$ und $0 < a < b$, so gilt $a \cdot x < b \cdot y$.*

Mit Hilfe von [30] erhalten wir

$$a \cdot x < a \cdot y = y \cdot a < y \cdot b = b \cdot y,$$

woraus $a \cdot x < b \cdot y$ folgt.

[32] *Falls $0 \leq x \leq y$ und $0 \leq a \leq b$ so gilt $a \cdot x \leq b \cdot y$ (*siehe Aufgabe 23*).*

[33] *Für alle $x > y > 0$ gilt $0 < x^{-1} < y^{-1}$.*

Da $y < x$ und x^{-1}, y^{-1} positiv nach [29] sind, so erhalten wir mit Hilfe von [30], dass

$$x^{-1} \cdot y < x^{-1} \cdot x$$

und

$$\left(x^{-1} \cdot y\right) \cdot y^{-1} < \left(x^{-1} \cdot x\right) \cdot y^{-1}. \tag{1.19}$$

Nach Axiomen II.3, II.2, II.1 erhalten wir

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

und
$$\left(x^{-1} \cdot x\right) \cdot y^{-1} = y^{-1}.$$
Somit impliziert (1.19), dass $x^{-1} < y^{-1}$.

Insbesondere erhalten wir aus (1.18) und [32], dass
$$1 > \frac{1}{2} > \frac{1}{3} > \frac{1}{4} > ... > 0.$$

Für jede reelle Zahl $x \in \mathbb{R}$ definieren wir den *Betrag* von x wie folgt:
$$|x| := \begin{cases} x, & x \geq 0, \\ -x, & x < 0. \end{cases}$$

Der Betrag erfüllt die folgenden Eigenschaften (siehe Aufgabe 24):

- $|xy| = |x|\,|y|$ (Multiplikativität). Insbesondere gilt $x^2 = |x|^2$.
- $|x + y| \leq |x| + |y|$ (die Dreiecksungleichung).

1.6 Teilmengen von $\mathbb{R}$ und Folgerungen aus dem Vollständigkeitsaxiom

1.6.1 Intervalle

Für je zwei reelle Zahlen a, b mit $a \leq b$ definieren wir die folgenden *Intervalle*:
$$\begin{array}{rcll} (a,b) & = & \{x \in \mathbb{R} : a < x < b\} & \text{– offenes Intervall,} \\ [a,b] & = & \{x \in \mathbb{R} : a \leq x \leq b\} & \text{– abgeschlossenes Intervall,} \\ (a,b] & = & \{x \in \mathbb{R} : a < x \leq b\} & \text{– halboffenes (linksoffenes) Intervall,} \\ [a,b) & = & \{x \in \mathbb{R} : a \leq x < b\} & \text{– halboffenes (rechtsoffenes) Intervall.} \end{array}$$

Die Zahlen a, b heißen die *Grenzen* des Intervalls.

Satz 1.7 *Jedes Intervall I mit den Grenzen $a < b$ ist nichtleer.*

Beweis. Es reicht zu beweisen, dass (a, b) nichtleer ist. Das Intervall $(0, 1)$ nicht leer ist, da $0 < \frac{1}{2} < 1$ und somit $\frac{1}{2} \in (0, 1)$. Für beliebige $a < b$ zeigen wir, dass die Zahl
$$c := \frac{1}{2} \cdot (a + b)$$
in (a, b) liegt. Es folgt aus $a < b$, dass
$$a + b < b + b = 1 \cdot b + 1 \cdot b = (1 + 1) \cdot b = 2 \cdot b,$$
woraus folgt
$$c = \frac{1}{2} \cdot (a + b) < \frac{1}{2} \cdot (2 \cdot b) = b,$$
d.h. $c < b$. Analog beweist man dass $c > a$ und somit $c \in (a, b)$, woraus folgt $(a, b) \neq \emptyset$. ■

1.6.2 Infimum und Supremum

Sei M eine nichtleere Teilmenge von $\mathbb{R}$.

Definition. Eine Zahl $a \in \mathbb{R}$ heißt *untere Schranke* von M falls gilt $x \geq a$ für alle $x \in M$. Eine Zahl $b \in \mathbb{R}$ heißt *obere Schranke* von M falls gilt $x \leq b$ für alle $x \in M$.

Beispiel. Für jedes Intervall I mit den Grenzen $a < b$ ist a immer eine untere Schranke und b – eine obere Schranke. Jede Zahl $a' < a$ ist auch eine untere Schranke von I, und jede Zahl $b' > b$ – eine obere Schranke von I.

Definition. Sei M eine nichtleere Teilmenge von $\mathbb{R}$. Eine Zahl $a \in \mathbb{R}$ heißt das *Minimum* von M (oder das *minimale Element* von M) und wird mit $\min M$ bezeichnet, falls a das kleinste Element von M ist, d.h. $a \in M$ und $x \geq a$ für alle $x \in M$. In anderen Wörter, $\min M$ ist eine untere Schranke von M die in M liegt.

Eine Zahl $b \in \mathbb{R}$ heißt das *Maximum* von M (oder das *maximale Element* von M) und wird mit $\max M$ bezeichnet, falls b das größte Element von M ist, d.h. $b \in M$ und $x \leq b$ für alle $x \in M$. In anderen Wörter, $\max M$ ist eine obere Schranke von M die in M liegt.

Beispiel. Sei $a < b$. Das abgeschlossene Intervall $I = [a, b]$ hat offensichtlich $\min I = a$ und $\max I = b$.

Zeigen wir, dass das offene Intervall $I = (a, b)$ weder Maximum noch Minimum hat. Existiert $\max I$, so setzen wir $c = \max I$ und bemerken, dass $c \in (a, b)$, insbesondere $c < b$. Das Intervall (c, b) nichtleer ist, sei $d \in (c, b)$. Dann gilt $d \in I$ und $d > c$ so dass c kein Maximum von I ist. Analog beweist man, dass (a, b) kein Minimum hat.

Das rechtsoffene Intervall $I = [a, b)$ hat $\min I = a$ und kein Maximum, und das linksoffene Intervall $I = (a, b]$ hat $\max I = b$ und kein Minimum.

Definition. Sei M eine nichtleere Teilmenge von $\mathbb{R}$. Die größte untere Schranke von M heißt das *Infimum* (oder *untere Grenze*) von M und wird mit $\inf M$ bezeichnet. Die kleinste obere Schranke von M heißt das *Supremum* (oder *obere Grenze*) von M und wird mit $\sup M$ bezeichnet.

Es ist klar aus den obigen Definitionen, dass

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

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

Supremum und Infimum existieren nicht immer. Zum Beispiel, die Menge $M = \mathbb{R}$ 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.

Satz 1.8 *Sei I ein Intervall mit den Grenzen $a < b$. Dann gilt*

$$\inf I = a \quad \textit{und} \quad \sup I = b.$$

Beweis. Beweisen wir, dass $\inf I = a$. Nach definition von I ist a eine untere Schranke von I. Zeigen wir, dass a die größte untere Schranke ist, d.h. für jede andere untere Schranke a' von I gilt $a' \leq a$. Sei a' eine untere Schranke von I mit $a' > a$. Für jede Zahl $x \in (a, b) \subset I$ gilt $a' \leq x$, woraus folgt $a' < b$, d.h.

$$a < a' < b. \tag{1.20}$$

Nach dem Satz 1.7 gibt es $c \in (a, a')$, und nach (1.20) beschließen wir, dass $c \in I$. Aber wir haben $c < a'$, was im Widerspruch zur Annahme steht, dass a' eine untere Schranke von I ist. Die Identität $\sup I = b$ wird analog bewiesen. ∎

2.05.18

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

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

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

$$\forall x \in M \quad \forall y \in N \ \text{ gilt } \ x \leq c \leq y.$$

Da $x \leq c$ für alle $x \in M$, so ist c eine obere Schranke von M und deshalb $c \in N$. Da $c \leq y$ für alle $y \in N$, so ist c das *kleinste* Element von N. Nach Definition erhalten wir $\sup M = c$.

Die Existenz des Infimums wird analog bewiesen. ∎

1.6.3 Quadratwurzel

Satz 1.10 *Für jede nichtnegative Zahl $a \geq 0$ existiert genau eine nichtnegative Zahl $x \in \mathbb{R}$ mit $x^2 = a$.*

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

Beweis. Betrachten wir die Menge

$$M = \left\{ r \in \mathbb{R} : r \geq 0 \text{ und } r^2 \leq a \right\}. \tag{1.21}$$

Diese Menge M ist nicht leer da $0 \in M$. Beweisen wir zunächst die folgende Behauptung.

Behauptung *Jede Zahl b mit $b^2 \geq a$ ist eine obere Schranke von M.*

Ist b keine obere Schranke von M, so gilt $b < r$ für ein $r \in M$, woraus folgt

$$a \leq b^2 < r^2 \leq a,$$

was zum Widerspruch $a < a$ führt.

Es folgt aus dieser Behauptung, dass $b = a + 1$ eine obere Schranke von M ist, da

$$b^2 = (a+1)^2 = a^2 + 2a + 1 > a.$$

Somit ist M nach oben beschränkt. Nach dem Satz 1.9 existiert

$$x = \sup M.$$

Beweisen wir jetzt, dass $x^2 = a$. Dafür reicht es zu zeigen, dass die anderen Fälle $x^2 < a$ und $x^2 > a$ unmöglich sind.

Zuerst nehmen wir an, dass $x^2 > a$. Insbesondere gilt $x > 0$. Wir erhalten einen Widerspruch indem wir zeigen, dass es eine Zahl ε gibt mit

$$0 < \varepsilon < x \tag{1.22}$$

und

$$(x - \varepsilon)^2 > a.$$

Daraus wird folgen, dass $x - \varepsilon$ auch eine obere Schranke von M ist, was im Widerspruch zur Definition von x als die *kleinste* obere Schranke von M steht. Da

$$(x - \varepsilon)^2 = x^2 - 2x\varepsilon + \varepsilon^2 > x^2 - 2x\varepsilon,$$

so reicht es ε so zu wählen, dass ε die folgende Ungleichung erfüllt:

$$x^2 - 2x\varepsilon > a,$$

was äquivalent zu

$$\varepsilon < \frac{x^2 - a}{2x}. \tag{1.23}$$

Da $x^2 > a$, so ist die Zahl $c := \frac{x^2-a}{2x}$ positiv. Nach dem Satz 1.7, es gibt ein $\varepsilon \in (0, c)$ (zum Beispiel, $\varepsilon = \frac{c}{2}$) so dass (1.23) erfüllt ist. Da

$$c < \frac{x^2}{2x} = \frac{x}{2} < x,$$

so gilt auch (1.22). Somit ist der Fall $x^2 > a$ ausgeschlossen.

Jetzt nehmen wir an, dass $x^2 < a$, und zeigen, dass es ein $\varepsilon > 0$ gibt mit

$$(x + \varepsilon)^2 < a$$

und

$$0 < \varepsilon < b, \tag{1.24}$$

wobei ein b mit $b^2 \geq a$ fixiert ist. Daraus wird folgen, dass $x + \varepsilon \in M$ was im Widerspruch zur Definition von x als obere Schranke von M steht. Da

$$(x + \varepsilon)^2 = x^2 + 2x\varepsilon + \varepsilon^2 < x^2 + 2x\varepsilon + b\varepsilon = x^2 + (2x + b)\,\varepsilon,$$

so reicht es ε so zu wählen, dass ε die folgende Ungleichung erfüllt:

$$x^2 + (2x + b)\,\varepsilon < a,$$

was äquivalent zu

$$\varepsilon < \frac{a - x^2}{2x + b}. \tag{1.25}$$

Da $a > x^2$, so ist die Zahl $c := \frac{a-x^2}{2x+b}$ positiv. Nach dem Satz 1.7, es gibt ein $\varepsilon \in (0, c)$, so dass (1.25) erfüllt ist. Die Bedingung (1.24) ist auch erfüllt, da

$$\varepsilon < c < \frac{a}{b} \le \frac{b^2}{b} = b.$$

Somit ist auch der Fall $x^2 < a$ ausgeschlossen, und wir erhalten, dass $x^2 = a$.

Für die Eindeutigkeit der Quadratwurzel zeigen wir, dass für $x, y \ge 0$ gilt die Implikation

$$x^2 = y^2 \Rightarrow x = y.$$

Ist $x \ne y$ so gilt $x < y$ oder $x > y$. Im ersten Fall erhalten wir $x^2 < y^2$ und im zweiten Fall $x^2 > y^2$, so dass $x \ne y$ unmöglich ist. ■

Erwähnen wir die folgenden Eigenschaften von Quadratwurzel (siehe Aufgabe 30).

1. Für alle $0 \le a \le b$ gilt$\sqrt{a} \le \sqrt{b}$.
2. Für jedes $x \in \mathbb{R}$ gilt $\sqrt{x^2} = |x|$.
3. Für alle $a, b \ge 0$ gilt $\sqrt{ab} = \sqrt{a}\sqrt{b}$ und für $b > 0$ gilt $\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}$.
4. Für jedes $a > 0$ hat die Gleichung $x^2 = a$ genau zwei reellen Lösungen: $x = \sqrt{a}$ und $x = -\sqrt{a}$.

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

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

Definition. Die *erweiterte Menge* $\overline{\mathbb{R}}$ von reellen Zahlen ist die geordnete Menge

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

wobei $+\infty$ und $-\infty$ neue Elemente sind, und die Ungleichung $<$ wird auf $\overline{\mathbb{R}}$ wie folgt erweitert: für alle $a \in \mathbb{R}$ gilt immer

$$-\infty < a < +\infty.$$

Der Begriff von Intervall lässt sich 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

$$[a, +\infty) = \left\{x \in \overline{\mathbb{R}} : a \le x < +\infty\right\} = \{x \in \mathbb{R} : x \ge a\}$$

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

Auch gelten $(-\infty,+\infty) = \mathbb{R}$ und $[-\infty,+\infty] = \overline{\mathbb{R}}$. Satz 1.8 gilt auch in diesem Fall.

Die Begriffe von oberen bzw unteren Schranken und somit sup und inf definiert man für nichtleere Teilmengen M von $\overline{\mathbb{R}}$ genau so, wie für Teilmengen von $\mathbb{R}$. Insbesondere ist $+\infty$ immer eine obere Schranke von M und $-\infty$ ist eine untere Schranke von M.

Für die leere Menge $M = \emptyset$ nehmen wir folgendes nach Definition an[2]:

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

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

Beweis. Sei M nichtleer. Ist $+\infty \in M$, so gilt $\sup M = +\infty$. Besteht M aus einem Element $-\infty$, so gilt $\sup M = -\infty$. Sonst ist die Menge $M' = M \setminus \{-\infty\}$ eine nichtleere Teilmenge von $\mathbb{R}$. Ist M' nach oben beschränkt, so existiert $\sup M'$ nach dem Satz 1.9, und dann gilt $\sup M = \sup M'$. Ist M' nach oben unbeschränkt, so ist $+\infty$ die einzige obere Schranke von M', woraus folgt $\sup M = \sup M' = +\infty$. Analog beweist man die Existenz von $\inf M$. ■

[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.

Chapter 2

Ganze Zahlen und vollständige Induktion

4.05.18

2.1 Natürliche Zahlen und Induktionsprinzip

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

- $1 \in \mathbb{N}$
- $x \in \mathbb{N}$ ergibt $x + 1 \in \mathbb{N}$ (die Zahl $x + 1$ der *Nachfolger* von x).

Insbesondere liegen die Zahlen $2, 3, 4$, usw. in $\mathbb{N}$. Allerdings bestimmen these diese zwei Eigenschaften die Menge $\mathbb{N}$ nicht eindeutig. Zum Beispiel, die ganze Menge $\mathbb{R}$ erfüllt sie.

Um eine richtige Definition von $\mathbb{N}$ anzugeben, bezeichnen wir mit $\mathcal{F}$ die Menge (das Mengensystem) von allen Teilmengen S von $\mathbb{R}$ die die folgenden Eigenschaften erfüllen:

- $1 \in S$
- $x \in S$ ergibt $x + 1 \in S$.

Das Mengensystem $\mathcal{F}$ ist nicht leer da $\mathbb{R} \in \mathcal{F}$.

Definition. Die Menge $\mathbb{N}$ wird als der Durchschnitt von allen Mengen $S \in \mathcal{F}$ definiert, d.h.

$$x \in \mathbb{N} \Leftrightarrow x \in S \quad \forall S \in \mathcal{F}. \tag{2.1}$$

Schreibweise für den Durchschnitt:

$$\mathbb{N} = \bigcap_{S \in \mathcal{F}} S. \tag{2.2}$$

Die Elemente von $\mathbb{N}$ heißen *natürliche Zahlen*.

Satz 2.1 *Die Menge* $\mathbb{N}$ *ist Element von dem Mengensystem* $\mathcal{F}$*, d.h.* $1 \in \mathbb{N}$ *und* $x \in \mathbb{N}$ *ergibt* $x+1 \in \mathbb{N}$*. Darüber hinaus ist die Menge* $\mathbb{N}$ *das kleinste (im Sinn von Inklusion) Element von* $\mathcal{F}$*.*

Beweis. Da $1 \in S$ für jedes $S \in \mathcal{F}$, so gilt nach (2.1) dass $1 \in \mathbb{N}$. Jetzt beweisen wir, dass $x \in \mathbb{N}$ ergibt $x+1 \in \mathbb{N}$:

$$\begin{aligned} x \in \mathbb{N} &\Longrightarrow x \in S \quad && \forall S \in \mathcal{F} \\ &\Longrightarrow x+1 \in S && \forall S \in \mathcal{F} \\ &\Longrightarrow x+1 \in \mathbb{N}. \end{aligned}$$

Somit erfüllt $\mathbb{N}$ die Definition von Elementen von $\mathcal{F}$, und $\mathbb{N} \in \mathcal{F}$.

Die Menge $\mathbb{N}$ ist das *kleinste* Element von $\mathcal{F}$ da die Inklusion $\mathbb{N} \subset S$ nach (2.2) für jedes $S \in \mathcal{F}$ gilt. ∎

Beispiel. Das Intervall $[1, +\infty)$ ist offensichtlich Element von $\mathcal{F}$. Es folgt, dass $\mathbb{N} \subset [1, +\infty)$. Insbesondere ist 1 die kleinste natürliche Zahl, d.h. $1 = \min \mathbb{N}$. Da $1 \in \mathbb{N}$, daraus folgt auch $2 = 1 + 1 \in \mathbb{N}$, $3 = 2 + 1 \in \mathbb{N}$ usw.

Um die Eigenschaften von natürlichen Zahlen beweisen zu können, brauchen wir das folgende *Induktionsprinzip*.

Satz 2.2 (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 gilt $1 \in S$, und nach dem Induktionsschritt gilt

$$x \in S \Rightarrow x \in S + 1.$$

Somit $S \in \mathcal{F}$, und nach der Definition von $\mathbb{N}$ erhalten wir $\mathbb{N} \subset S$. Folglich ist $A(n)$ 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 folgenden Eigenschaften von natürliche Zahlen werden mit Hilfe von Induktionsprinzip bewiesen.

Satz 2.3 *Für alle* $n, m \in \mathbb{N}$ *gilt:*

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

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

Beweis. (a) Fixieren wir $m \in \mathbb{N}$ und beweisen per Induction nach n, dass $n+m \in \mathbb{N}$. Dafür bezeichnen wir mit $A(n)$ die Aussage, dass $n+m \in \mathbb{N}$.

(i) Induktionsanfang $A(1)$.

Aussage $A(1)$ bedeutet, dass $1+m \in \mathbb{N}$. Da $m \in \mathbb{N}$, so folgt es nach dem Satz 2.1 dass $m+1 \in \mathbb{N}$, was zu beweisen war.

(ii) Induktionsschritt $A(n) \Rightarrow A(n+1)$.

Die Induktionsvoraussetzung $A(n)$ ist: $n+m \in \mathbb{N}$.
Die Induktionsbehauptung $A(n+1)$ ist: $(n+1)+m \in \mathbb{N}$.
Angenommen, dass $A(n)$ wahr ist, d.h. $n+m \in \mathbb{N}$, so erhalten wir $(n+m)+1 \in \mathbb{N}$, woraus folgt

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

Somit ist die Implikation $A(n) \Rightarrow A(n+1)$ bewiesen. Nach dem Induktionsprinzip beschließen wir, dass $A(n)$ für alle $n \in \mathbb{N}$ gilt, d.h. $n+m \in \mathbb{N}$ für alle $n \in \mathbb{N}$, was zu beweisen war.

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

(i) Induktionsanfang. Für $n=1$ gilt $nm = 1 \cdot m = m \in \mathbb{N}$.

(ii) Induktionsschritt von n nach $n+1$.

Angenommen dass $nm \in \mathbb{N}$, so erhalten wir nach (a)

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

da $m \in \mathbb{N}$ gegeben ist und $nm \in \mathbb{N}$ nach der Induktionsvoraussetzung gilt. ■

Satz 2.4 *Für alle $n, m \in \mathbb{N}$ mit $n > m$ gilt $n - m \in \mathbb{N}$.*

Beweis. Beweisen wir zunächst diese Behauptung für $m = 1$, d.h.

$$\forall n \in \mathbb{N} \text{ mit } n > 1 \text{ gilt } n - 1 \in \mathbb{N}. \tag{2.3}$$

Das letztere ist äquivalent zur folgenden Aussage

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

die wir per Induktion nach n beweisen.

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

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

Somit ist (2.3) bewiesen.
Beweisen wir jetzt per Induktion nach m die neue Aussage

$$B(m) = (\forall n \in \mathbb{N} \text{ mit } n > m \text{ gilt } n - m \in \mathbb{N}).$$

(i) Induktionsanfang $B(1)$ ist genau (2.3), was schon bewiesen wurde.

(ii) Induktionsschritt von m nach $m+1$.

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

$$\forall n \in \mathbb{N} \text{ mit } n > m+1 \text{ gilt } n-(m+1) \in \mathbb{N}.$$

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

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

was beweist, dass $n-(m+1) \in \mathbb{N}$. ■

2.2 Summe und Produkt endlicher Folgen

In diesem Abschnitt benutzen wir das Induktionsprinzip um rigorose Definitionen von Begriffen von Summe und Produkt von mehreren reellen Zahlen anzugeben.

Für jede natürliche Zahl n betrachten wir die Menge

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

von natürlichen Zahlen zwischen 1 und n. Diese Menge wird häufig auch mit $\{1, ..., n\}$ bezeichnet.

Definition. Eine *endliche Folge* ist eine Abbildung $a : \mathcal{E}_n \to \mathbb{R}$. Man bezeichnet den Wert $a(k)$ auch mit a_k und die Folge a mit $\{a_k\}_{k=1}^n$ oder kurz mit $\{a_k\}$.

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

Definition. Definieren wir die Summe $\sum_{k=1}^n a_k$ einer Folge $\{a_k\}$ per Induction nach n wie folgt:

(i) für $n = 1$ setzen wir

$$\sum_{k=1}^{1} a_k = a_1 \ ;$$

(ii) ist $\sum_{k=1}^n a_k$ schon definiert, so setzen wir

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

Nach dem Induktionsprinzip ist die Summe $\sum_{k=1}^n a_k$ für alle $n \in \mathbb{N}$ definiert. Man schreibt auch

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + ... + a_n.$$

Das Zeichen $\sum$ ist ein griechischer Grossbuchstabe namens “sigma” (was sich auf den Buchstaben “S” von “Summe” bezieht).

Zum Beispiel, es gilt

$$\sum_{k=1}^{3} a_k = a_1 + a_2 + a_3.$$

Noch ein Beispiel: sind alle a_k gleich a, so gilt

$$\sum_{k=1}^{n} a_k = na,$$

was man per Induktion nach n beweist.

Definition. Definieren wir das Produkt $\prod_{k=1}^n a_k$ der Folge $\{a_k\}$ per Induktion nach n wie folgt:

$$\prod_{k=1}^{1} a_k = a_1$$

und

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

Man schreibt auch

$$\prod_{k=1}^{n} a_k = a_1 \cdot a_2 \cdot ... \cdot a_n.$$

Das Zeichen $\prod$ ist ein griechischer Grossbuchstabe namens “pi”.

Zum Beispiel, definieren wir die *Fakultät*[1] $n!$ (n mit dem Ausrufezeichen) wie folgt: für $n \in \mathbb{N}$

$$n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot ... \cdot n,$$

und $0! = 1$.

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:

$$a^1 = a \quad \text{und} \quad a^{n+1} = a^n \cdot a \quad \text{für alle } n \in \mathbb{N}. \tag{2.4}$$

Im Ausdruck a^n heißt die Zahl a *Basis* und n – *Exponent.* Man beweist per Induktion die folgenden Identitäten:

$$a^n \cdot a^m = a^{n+m}, \quad (a^n)^m = a^{n \cdot m}, \quad (a \cdot b)^n = a^n \cdot b^n, \tag{2.5}$$

für alle $n, m \in \mathbb{N}$ (siehe Aufgabe 33).

[1] heißt auch *Faktorielle*

Beispiel. Beweisen wir die letzte Identität in (2.5) per Induktion nach n. Induktionsanfang für $n = 1$:

$$(a \cdot b)^1 = a \cdot b = a^1 \cdot b^1.$$

Induktionsschritt von n nach $n+1$:

$$\begin{aligned}(a \cdot b)^{n+1} &= (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a \cdot b) \\ &= (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1}.\end{aligned}$$

Beispiel. Beweisen wir für alle $n \in \mathbb{N}$ die Identität

$$\sum_{k=1}^{n} 2^k = 2^{n+1} - 2, \tag{2.6}$$

d.h.

$$2 + 2^2 + ... + 2^n = 2^{n+1} - 2.$$

Induktionsanfang. Für $n = 1$ ist die linke Seite von (2.6) gleich

$$\sum_{k=1}^{n} 2^k = 2^1 = 2,$$

und die rechte Seite gleich

$$2^{1+1} - 2 = 2^2 - 2 = 2 \cdot 2 - 2 \cdot 1 = 2 \cdot (2-1) = 2,$$

so dass (2.6) gilt.

Induktionsschritt von n nach $n+1$. Angenommen, dass (2.6) für ein $n \in \mathbb{N}$ gilt, so erhalten wir

$$\begin{aligned}\sum_{k=1}^{n+1} 2^k &= \sum_{k=1}^{n} 2^k + 2^{n+1} = \left(2^{n+1} - 2\right) + 2^{n+1} \\ &= 2 \cdot 2^{n+1} - 2 = 2^{n+2} - 2,\end{aligned}$$

d.h. (2.6) auch für $n+1$ anstatt n gilt. Nach dem Induktionsprinzip beschließen wir, dass (2.6) für alle $n \in \mathbb{N}$ gilt.

Die folgenden allgemeinen Eigenschaften der Summe $\sum$ lassen sich per Induktion beweisen:

$$c \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} (c a_k)$$

$$\left(\sum_{k=1}^{n} a_k\right) + \left(\sum_{k=1}^{n} b_k\right) = \sum_{k=1}^{n} (a_k + b_k)$$

$$\left(\sum_{k=1}^{n} a_k\right)\left(\sum_{l=1}^{m} b_l\right) = \sum_{k=1}^{n}\left(\sum_{l=1}^{m} (a_k b_l)\right) = \sum_{l=1}^{m}\left(\sum_{k=1}^{n} (a_k b_l)\right)$$

(siehe Aufgabe 46).

2.3 Ganze Zahlen

9.05.18

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

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

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

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

Viele Eigenschaften von natürlichen Zahlen lassen sich auf ganze Zahlen verallgemeinern.

Operationen in $\mathbb{Z}$

Satz 2.5 *Für alle $x, y \in \mathbb{Z}$ 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 2.4. Ist $n < m$, dann $m - n \in \mathbb{N}$ und $n - m = -(m - n) \in \mathbb{Z}$.

Die anderen Fälle von Vorzeichen von x und y werden analog betrachtet. ■

Korollar 2.6 *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 $x > y$. Dann $x - y > 0$ und nach dem Satz 2.5 gilt $x - y \in \mathbb{Z}$, woraus folgt $x - y \in \mathbb{N}$. Da 1 die kleinste natürliche Zahl ist, so erhalten wir $x - y \geq 1$ und somit $x \geq y + 1$. Die Implikation in der Rückrichtung ist trivial: $x \geq y + 1$ und $y + 1 > y$ ergeben $x > y$.

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

Teilmengen von $\mathbb{Z}$

Satz 2.7 *Sei M eine nichtleere Teilmenge von $\mathbb{Z}$. Ist M nach oben beschränkt, so existiert* $\max M$. *Ist M nach unten beschränkt, so existiert* $\min M$. *Insbesondere existiert* $\min M$ *für jede nichtleere Teilmenge von $\mathbb{N}$.*

Beweis. Sei M nach unten beschränkt. Nach dem Satz 1.9 existiert $a = \inf M \in \mathbb{R}$. Beweisen wir, dass $a \in M$, woraus die Identität $a = \min M$ folgen wird. Nehmen wir das Gegenteil an, dass $a \notin M$. Da $a + 1$ keine untere Schranke von M ist, so existiert ein $n \in M$ mit $n < a + 1$. Da a eine untere Schranke von M ist, so gilt

$n \geq a$. Da $a \notin M$ und $n \in M$, so gilt $a \neq n$ und somit $n > a$, d.h. $a < n < a+1$, woraus folgt

$$n - 1 < a < n.$$

Für jedes $x \in M$ gilt $x \geq a$ und somit $x > n-1$. Aus dem Korollar 2.6 folgt es dass $x \geq n$. Somit ist n eine untere Schranke von M, was unmöglich ist, da a die größte untere Schranke von M ist. Somit beschließen wir, dass $a \in M$. Der Fall von $\max M$ wird analog behandelt.

Jede nichtleere Teilmenge M von $\mathbb{N}$ ist immer nach unten von 1 beschränkt, woraus die Existenz von $\min M$ folgt. ■

Induktionsprinzip mit verschobenem Anfang

Um die Aussagen mit ganzen Zahlen zu beweisen, man benutzt häufig die folgende Verallgemeinerung des Induktionsprinzipes.

Satz 2.8 (Induktionsprinzip mit verschobenem Anfang) *Sei n_0 eine ganze Zahl und sei $A(n)$ eine von $n \in \mathbb{Z}$ abhängige Aussage. Angenommen ist das folgende:*

(i) $A(n_0)$ ist wahr.

(ii) Für jedes $n \in Z$ mit $n \geq n_0$ gilt $A(n) \Rightarrow A(n+1)$.

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

Beweis. Setzen wir $m = n - n_0 + 1$ und bemerken, dass $n \in \mathbb{Z} \Leftrightarrow m \in \mathbb{Z}$ und $n \geq n_0 \Leftrightarrow m \geq 1$. Um die Aussage $A(n)$ für alle ganzen Zahlen $n \geq n_0$ zu beweisen, reicht es die Aussage

$$B(m) := A(m + n_0 - 1)$$

für alle $m \in \mathbb{N}$ zu beweisen. Induktionsanfang: für $m = 1$ die Aussage $B(1) = A(n_0)$ gilt nach (i).

Induktionsschritt: die Implikation

$$B(m) = A(m + n_0 - 1) \implies A(m + n_0) = B(m+1),$$

gilt nach (ii) mit $n = m + n_0 - 1$. ■

Negative Potenzen

Definieren wir jetzt die Potenzen a^n einer Zahl $a \in \mathbb{R} \setminus \{0\}$ für alle $n \in \mathbb{Z}$. Für $n > 0$ wurde a^n schon oberhalb definiert. Setzen wir

$$a^0 := 1 \quad \text{und} \quad a^{-n} := \left(a^{-1}\right)^n \quad \text{für } n \in \mathbb{N},$$

wobei a^{-1} das Inverse von a ist.

Beispiel. Als Beispiel, beweisen wir, dass für $n \in \mathbb{N}$

$$a^{-n} = (a^n)^{-1}.$$

Da nach Definition $a^{-n} = \left(a^{-1}\right)^n$, so reicht es zu beweisen, dass

$$\left(a^{-1}\right)^n = (a^n)^{-1},$$

was äquivalent zu

$$\left(a^{-1}\right)^n \cdot a^n = 1$$

ist. Diese Identität gilt da nach (2.5)

$$\left(a^{-1}\right)^n \cdot a^n = \left(a^{-1} \cdot a\right)^n = 1^n = 1$$

(da $1^n = 1$ beweist man per Induktion nach n).

Die Identitäten (2.5) gelten für alle $n, m \in \mathbb{Z}$:

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

vorausgesetzt $a, b \neq 0$ (siehe Aufgabe 40, 41).

Endliche Folgen

Der Begriff von Folge $\{a_k\}_{k=1}^n$ lässt sich auch verallgemeinern. Für je zwei ganze Zahlen m, n mit $m \leq n$, betrachten wie die Menge

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

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

Definition. Eine *endliche Folge* $\{a_k\}_{k=m}^n$ ist eine Abbildung $a : \{m, ..., n\} \to \mathbb{R}$, wo $a(k) = a_k$.

Die Summe $\sum_{k=m}^n a_k$ der Folge $\{a_k\}$ wird für alle $n \geq m$ per Induction nach n wie folgt definiert:

(*i*) ist $n = m$ dann setzen wir $\sum_{k=m}^m a_k = a_m$;

(*ii*) ist $\sum_{k=m}^n a_k$ schon definiert, so setzen wir

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

Man schreibt auch

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

Man kann beweisen, dass für alle ganzen Zahlen l mit $m \leq l < n$ gilt

$$\sum_{k=m}^n a_k = \sum_{k=m}^l a_k + \sum_{k=l+1}^n a_k \tag{2.7}$$

(siehe Aufgabe 43).

Archimedisches Prinzip und Gaußklammer

Satz 2.9 (Archimedisches Prinzip) *Für jedes $x \in \mathbb{R}$ existiert genau ein $n \in \mathbb{Z}$ mit*

$$n \leq x < n+1, \tag{2.8}$$

d.h. $x \in [n, n+1)$.

Folglich ist $\mathbb{R}$ eine disjunkte Vereinigung von allen Intervallen $[n, n+1)$ mit $n \in \mathbb{Z}$.

Die Zahl n mit (2.8) ist somit 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. Fixieren wir ein $x \in \mathbb{R}$ betrachten 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 2.7 hat $\mathbb{Z}$ das Minimum. Sei $m = \min \mathbb{Z}$. Aber dann ist $m-1$ auch in $\mathbb{Z}$, was im Widerspruch zu $m - 1 < \min \mathbb{Z}$ steht.

Da S nichtleer ist und nach oben von x beschränkt, erhalten wir nach dem Satz 2.7, dass S das Maximum hat. Setzen wir $n = \max S$. Da $n \in S$, so gilt nach Definition von S, dass $n \leq x$. Da $n+1 > n = \max S$ und deshalb $n+1 \notin S$, erhalten wir, dass $n+1 > x$. Somit erfüllt n die Bedingungen (2.8).

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 2.6 $n' \leq n$. Analog sieht man, dass $n \leq n'$, woraus $n = n'$ folgt. ■

Bemerkung. Für den Ganzzahlanteil von x benutzt man auch die Notation $\lfloor x \rfloor$. Mit $\lceil x \rceil$ bezeichnet man die einzige ganze Zahl m mit $m - 1 < x \leq m$. Die Existenz von m beweist man analog oder durch $m = -\lfloor -x \rfloor$.

2.4 Binomischer Lehrsatz

Für alle $a, b \in \mathbb{R}$ gelten die Identitäten

$$(a+b)^2 = a^2 + 2ab + b^2$$

und

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

Hier beweisen wir eine ähnliche Identität für $(a+b)^n$ für beliebiges $n \in \mathbb{N}$. Da die Summe $a+b$ auch Binom heißt, so heißt solche Identität binomische Formel oder binomischer Lehrsatz. 11.05.18

Für ganze Zahlen $n \geq k \geq 0$ definieren wir den *Binomialkoeffizient* $\binom{n}{k}$ ("n über k") durch

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}, \tag{2.9}$$

wobei $n! = 1 \cdot 2 \cdot ... \cdot n$ für $n \in \mathbb{N}$ und $0! = 1$.

Es folgt aus (2.9), dass

$$\binom{n}{k} = \binom{n}{n-k}.$$

Z.B. wir haben

$$\binom{n}{0} = \frac{n!}{0!n!} = 1,$$

$$\binom{n}{1} = \frac{n!}{1!\,(n-1)!} = n,$$

$$\binom{n}{2} = \frac{n!}{2!\,(n-2)!} = \frac{n\,(n-1)}{2}.$$

Satz 2.10 (Binomischer Lehrsatz) *Für alle $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k. \tag{2.10}$$

Beispiel. Es folgt aus (2.10), dass für $n = 4$

$$\begin{aligned} (a+b)^4 &= \sum_{k=0}^{4} \binom{4}{k} a^{4-k} b^k = a^4 + \binom{4}{1} a^3 b + \binom{4}{2} a^2 b^2 + \binom{4}{3} a b^3 + b^4 \\ &= a^4 + 4a^3 b + 6a^2 b^2 + 4ab^3 + b^4 \end{aligned}$$

und für $n = 5$

$$\begin{aligned} (a+b)^5 &= \sum_{k=0}^{5} \binom{5}{k} a^{5-k} b^k = a^5 + \binom{5}{1} a^4 b + \binom{5}{2} a^3 b^2 + \binom{5}{3} a^2 b^3 + \binom{5}{4} a b^4 + b^5 \\ &= a^5 + 5a^4 b + 10a^3 b^2 + 10a^2 b^3 + 5ab^4 + b^5. \end{aligned}$$

Beweis. Wir benutzen die folgende Identität: für alle $n \geq j \geq 1$ gilt

$$\binom{n}{j} + \binom{n}{j-1} = \binom{n+1}{j}. \tag{2.11}$$

(siehe Aufgabe 45). Wir beweisen (2.10) per Induction nach n. Induktionsanfang: für $n = 1$ ist (2.10) äquivalent zu $(a+b)^1 = a + b$.

Induktionsschritt von n nach $n+1$:

$$
\begin{aligned}
(a+b)^{n+1} &= (a+b)\,(a+b)^n \\
&= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k \qquad\qquad \text{(Induktionsvoraussetzung)} \\
&= a\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k + b\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{n-k+1}b^k + \sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^{k+1} \\
&= \sum_{j=0}^{n}\binom{n}{j}a^{n-j+1}b^j + \sum_{j=1}^{n+1}\binom{n}{j-1}a^{n-(j-1)}b^j \qquad \text{(Wechsel } j=k+1\text{)} \\
&= a^{n+1} + \sum_{j=1}^{n}\binom{n}{j}a^{n-j+1}b^j + \sum_{j=1}^{n}\binom{n}{j-1}a^{n-j+1}b^j + b^{n+1} \\
&= a^{n+1} + \sum_{j=1}^{n}\left(\binom{n}{j}+\binom{n}{j-1}\right)a^{n+1-j}b^j + b^{n+1} \\
&= a^{n+1} + \sum_{j=1}^{n}\binom{n+1}{j}a^{(n+1)-j}b^j + b^{n+1} \qquad\qquad (2.11) \\
&= \sum_{j=0}^{n+1}\binom{n+1}{j}a^{(n+1)-j}b^j,
\end{aligned}
$$

was zu beweisen war. ■

2.5 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 2.6, 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 2.11 *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:

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

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 Aufgabe 22. ∎

Bemerken wir, dass die Menge $\mathbb{R} \setminus \mathbb{Q}$ von *irrationalen* Zahlen nicht leer ist, da $\sqrt{2}$ irrational ist (siehe Aufgabe 55). Daraus auch folgt, dass der Körper $\mathbb{Q}$ nicht vollständig ist, da das Vollständigkeitsaxiom die Existenz der Quadratwurzel $\sqrt{a}$ für jedes $a \geq 0$ impliziert.

2.6 Endliche Mengen und Kardinalität

Definition. Seien X, Y zwei nicht-leere Mengen. Die Menge X heißt *gleichmächtig* (oder *äquivalent*) zu Y falls es eine bijektive Abbildung $f : X \to Y$ gibt. In diesem Fall schreibt man $X \sim Y$. Die leere Menge $\emptyset$ ist nach Definition gleichmächtig zu sich selbst, d.h. $\emptyset \sim \emptyset$.

Satz 2.12 *Die Gleichmächtigkeit von Mengen hat die folgenden Eigenschaften:*

- $X \sim X$ *(Reflexivität)*
- $X \sim Y \Rightarrow Y \sim X$ *(Symmetrie)*
- $X \sim Y \ \wedge \ Y \sim Z \Rightarrow X \sim Z$ *(Transitivität).*

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 $f : X \to Y$. Nach dem Satz 1.5 ist die Umkehrabbildung $f^{-1} : Y \to X$ wohldefiniert und bijektiv, woraus $Y \sim X$ folgt.

Existieren die bijektiven Abbildungen $g : X \to Y$ und $f : Y \to Z$, so ist die Verkettung $f \circ g : X \to Z$ bijektiv nach dem Satz 1.6, woraus $X \sim Z$ folgt. ∎

Für jedes $n \in \mathbb{N}$ betrachten wir die Menge

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

Setzen wir auch $\mathcal{E}_0 = \emptyset$. 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 $n \in \mathbb{N} \cup \{0\}$. Gibt es solches n nicht, so heißt S *unendlich*.

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. die Menge S lässt sich mit den Zahlen $1, ..., n$ aufzählen.

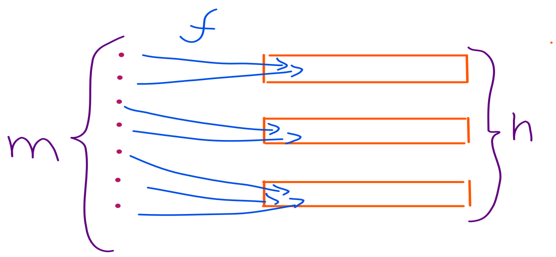
Definition. Gilt $S \sim \mathcal{E}_n$, so sagen wir: die Anzahl von Elementen von S ist n, oder die *Kardinalzahl* von S ist n, oder die *Kardinalität* von S ist n. Man bezeichnet die Kardinalität von S mit $\operatorname{card} S$ oder mit $|S|$ (Betrag von S).

Beispiel. Für die Menge $S = \{a, b\}$ gilt $\operatorname{card} S = 2$, da diese Menge zu $\mathcal{E}_2 = \{1, 2\}$ gleichmächtig ist.

Wir beweisen einige Eigenschaften von endlichen Mengen und ihren Kardinalitäten. Aber zunächst zeigen wir, dass die Kardinalität wohldefiniert ist, d.h. $S \sim \mathcal{E}_n$ und $S \sim \mathcal{E}_m$ mit verschiedenen Zahlen n, m unmöglich ist.

Satz 2.13 *Sind m, n natürliche Zahlen mit $m > n$, so gibt es keine injektive Abbildung $f : \mathcal{E}_m \to \mathcal{E}_n$. Insbesondere sind $\mathcal{E}_m$ und $\mathcal{E}_n$ gleichmächtig genau dann, wenn $m = n$.*

Die Aussage von Satz 2.13 heißt das *Schubfachprinzip*: sind m Objekten zwischen n Schubfächern verteilt, wobei $m > n$, so gibt es ein Schubfach mit mindestens zwei Objekten.



Es gibt viele interessante Anwendungen von diesem Prinzip. Im Beweis benutzen wir den Begriff von *disjunkter* Vereinigung.

Definition. Seien A, B zwei Mengen. Die disjunkte Vereinigung $A \sqcup B$ wird als die Vereinigung $A \cup B$ definiert, vorausgesetzt $A \cap B = \emptyset$. Im Fall $A \cap B \neq \emptyset$ wird $A \sqcup B$ nicht definiert.

Behauptung *Disjunkte Vereinigung hat die folgende Eigenschaft: ist $A \sim A'$ und $B \sim B'$ so gilt*

$$A \sqcup B \sim A' \sqcup B', \tag{2.12}$$

vorausgesetzt dass $A \cap B = \emptyset$ und $A' \cap B' = \emptyset$.

Beweis. Seien $f : A \to A'$ und $g : B \to B'$ bijektive Abbildungen. Definieren wir die Abbildung

$$F : A \sqcup B \to A' \sqcup B'$$

wie folgt:

$$F(x) = \begin{cases} f(x), & \text{falls } x \in A, \\ g(x), & \text{falls } x \in B. \end{cases}$$

Dann ist F bijektiv, woraus (2.12) folgt. ■

Vor dem Beweis des Satzes 2.13 beweisen wir noch ein Lemma.

Lemma 2.14 *Für jedes $n \in \mathbb{N}$ und für jedes $a \in \mathcal{E}_{n+1}$ gilt*

$$\mathcal{E}_{n+1} \setminus \{a\} \sim \mathcal{E}_n. \tag{2.13}$$

Beweis. Induktion nach n. Für $n = 1$ ist (2.13) offensichtlich, dass für $a = 2$

$$\mathcal{E}_2 \setminus \{a\} = \{1, 2\} \setminus \{2\} = \{1\} = \mathcal{E}_1$$

und für $a = 1$

$$\mathcal{E}_2 \setminus \{a\} = \{1, 2\} \setminus \{1\} = \{2\} \sim \{1\} = \mathcal{E}_1.$$

Induktionsschritt von n nach $n+1$. Angenommen sei (2.13), beweisen wir, dass für jedes $a \in \mathcal{E}_{n+2}$

$$\mathcal{E}_{n+2} \setminus \{a\} \sim \mathcal{E}_{n+1}. \tag{2.14}$$

Ist $a = n + 2$, so gilt

$$\mathcal{E}_{n+2} \setminus \{n + 2\} = \mathcal{E}_{n+1} \sim \mathcal{E}_{n+1}.$$

Ist $a < n + 2$, so liegt a in $\mathcal{E}_{n+1}$. Nach Induktionsvoraussetzung haben wir (2.13), woraus mit Hilfe von (2.12) folgt

$$\begin{aligned} \mathcal{E}_{n+2} \setminus \{a\} &= (\mathcal{E}_{n+1} \setminus \{a\}) \sqcup \{n + 2\} \\ &\sim \mathcal{E}_n \sqcup \{n + 1\} = \mathcal{E}_{n+1}. \end{aligned}$$

■

Für den Beweis brauchen wir auch den folgenden Begriff. Sei $f : X \to Y$ eine Abbildung.

Definition. Eine *Einschränkung* (Restriktion) von f auf einer Teilmenge $U \subset X$ ist die Abbildung

$$\begin{aligned} g &: U \to Y \\ g(x) &= f(x). \end{aligned}$$

Schreibweise: $g = f|_U$. In anderen Wörter $f|_U$ hat die gleiche Zuordnung und den gleichen Wertebereich wie f aber den kleineren Definitionsbereich.

Liegt das Bild $f(X)$ in einer Teilmenge $V \subset Y$, so kann man die folgende Abbildung betrachten:

$$\begin{aligned} h &: X \to V \\ h(x) &= f(x). \end{aligned}$$

Für diese Abbildung gibt es keine spezielle Notation, man bezeichnet sie auch mit f mit Erklärung dass der Wertebereich gleich V anstatt Y ist.

Beweis von dem Satz 2.13. Die Bedingung $m > n$ impliziert, dass $m \geq n + 1$ und somit $\mathcal{E}_{n+1} \subset \mathcal{E}_m$. Soll eine injektive Abbildung $f : \mathcal{E}_m \to \mathcal{E}_n$ mit $m > n$ existieren, so ist die Einschränkung $f|_{\mathcal{E}_{n+1}}$ eine injektive Abbildung von $\mathcal{E}_{n+1}$ nach

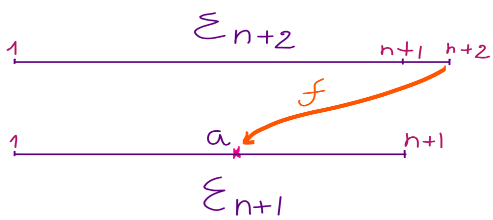
$\mathcal{E}_n$. Deshalb reicht es zu beweisen, dass es keine injektive Abbildung $\mathcal{E}_{n+1} \to \mathcal{E}_n$ gibt. Diese Aussage beweisen wir per Induktion nach n.

Induktionsanfang für $n = 1$. Es gibt nur eine Abbildung $f : \mathcal{E}_2 \to \mathcal{E}_1$ und zwar mit $f(1) = f(2) = 1$. Somit ist f nicht injektiv.

Induktionsschritt von n nach $n+1$. Sei

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

eine injektive Abbildung. Setzen with $a = f(n+2)$.



Da f injektiv ist, so gilt

$$f(k) \neq a \quad \forall k \in \mathcal{E}_{n+1}.$$

Somit ergibt die Einschränkung von f auf $\mathcal{E}_{n+1}$ eine injektive Abbildung

$$h : \mathcal{E}_{n+1} \to \mathcal{E}_{n+1} \setminus \{a\}. \tag{2.15}$$

Nach Lemma 2.14 gilt (2.13), d.h. es gibt eine bijektive Abbildung

$$g : \mathcal{E}_{n+1} \setminus \{a\} \to \mathcal{E}_n. \tag{2.16}$$

Die Komposition zweier injektiven Abbildungen (2.15) und (2.16) ergibt eine injektive Abbildung

$$g \circ h : \mathcal{E}_{n+1} \to \mathcal{E}_n,$$

was im Widerspruch zur Induktionsvoraussetzung steht. Somit ist das Schubfachprinzip bewiesen. 16.05.18

Beweisen wir jetzt, dass

$$\mathcal{E}_m \sim \mathcal{E}_n \Leftrightarrow m = n.$$

Im Fall $n = m$ gilt $\mathcal{E}_n \sim \mathcal{E}_m$. Ist $m > n$ so gibt es keine injektive Abbildung $\mathcal{E}_m \to \mathcal{E}_n$ und somit $\mathcal{E}_m \not\sim \mathcal{E}_n$. Im Fall $n > m$ gibt es keine injektive Abbildung $\mathcal{E}_n \to \mathcal{E}_m$ und wieder $\mathcal{E}_m \not\sim \mathcal{E}_n$. ■

Somit ist die Kardinalität $\operatorname{card} A$ wohldefiniert für jede endliche Menge A. Es folgt auch aus dem Satz 2.13 dass für beliebige endliche Mengen A und B mit $\operatorname{card} A > \operatorname{card} B$ es keine injektive Abbildung $f : A \to B$ gibt.

Korollar 2.15 *Gilt* $\operatorname{card} A = \operatorname{card} B$ *so ist jede injektive Abbildung* $f : A \to B$ *notwendigerweise bijektiv.*

Beweis in Aufgabe 52.

Korollar 2.16 *Die Menge* $\mathbb{N}$ *ist unendlich.*

Beweis. Nehmen wir das Gegenteil an, dass $\mathbb{N}$ endlich ist. Dann existiert eine Bijektion $f : \mathbb{N} \to \mathcal{E}_n$ mit einem $n \in \mathbb{N}$. Die Einschränkung $f|_{\mathcal{E}_{n+1}}$ ist dann eine injektive Abbildung von $\mathcal{E}_{n+1}$ nach $\mathcal{E}_n$, was nach dem Schubfachprinzip unmöglich ist. ■

Satz 2.17 *Jede Teilmenge* S *einer endlichen Menge* M *ist endlich und*

$$\operatorname{card} S \leq \operatorname{card} M.$$

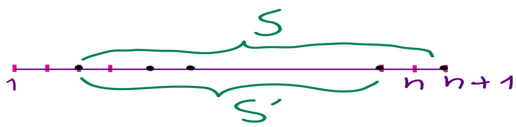
Beweis. Sei $n = \operatorname{card} M$. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $M = \mathcal{E}_n$, und beweisen per Induktion nach n, dass jede Teilmenge S von $\mathcal{E}_n$ endlich ist und $\operatorname{card} S \leq n$.

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

Induktionsschritt von n nach $n+1$. Angenommen sei: jede Teilmenge T von $\mathcal{E}_n$ ist endlich und $\operatorname{card} T \leq n$. Wir beweisen, dass jede Teilmenge $S \subset \mathcal{E}_{n+1}$ endlich ist und $\operatorname{card} S \leq n+1$. Betrachten wir zwei Fälle.

Gilt $S \subset \mathcal{E}_n$, so ist S endlich nach Induktionsvoraussetzung, und $\operatorname{card} S \leq n < n+1$.

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 nach Induktionsvoraussetzung ist S' endlich und $\operatorname{card} S' \leq n$. In anderen Wörter $S' \sim \mathcal{E}_m$ für ein $m \leq n$.



Dann erhalten wir mit Hilfe von (2.12)

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

woraus die Endlichkeit von S folgt, und auch $\operatorname{card} S = m+1 \leq n+1$. ■

Es folgt aus dem Satz 2.17, dass für jede zwei endliche Mengen A, B auch die Mengen $A \cap B$, $A \setminus B$, $B \setminus A$ endlich sind.

Satz 2.18 *Für beliebige endliche Mengen* A, B *ist die Vereinigung* $A \cup B$ *auch endlich. Es gilt auch*

$$\operatorname{card}(A \sqcup B) = \operatorname{card} A + \operatorname{card} B, \tag{2.17}$$

vorausgesetzt dass A *und* B *disjunkt sind.*

Beweis. Die Vereinigung $A \cup B$ lässt sich als eine disjunkte Vereinigung darstellen:

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

(dies folgt aus der Aufgabe 10 mit $X = A \cup B$). Deshalb reicht es zu beweisen, dass eine disjunkte Vereinigung zweier Mengen endlich ist.

Seien jetzt A, B zwei disjunkte endliche Mengen mit $\operatorname{card} A = n$ und $\operatorname{card} B = m$, d.h.

$$A \sim \mathcal{E}_n \quad \text{und} \quad B \sim \mathcal{E}_m. \tag{2.18}$$

Wir beweisen, dass $A \sqcup B$ endlich ist und dass (2.17) gilt. Dafür betrachten wir die Menge

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

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

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

woraus folgt

$$A \sqcup B \sim \mathcal{E}_{n+m}. \tag{2.19}$$

Insbesondere ist $A \sqcup B$ endlich. Die Identität (2.17) folgt aus (2.19). ■

Bemerkung. Es folgt aus (2.17) dass für jede Teilmenge B einer endlichen Menge A gilt

$$\operatorname{card}(A \setminus B) = \operatorname{card} A - \operatorname{card} B,$$

da $A = B \sqcup (A \setminus B)$. Für beliebige endliche Mengen A und B gelten die Identitäten:

$$\begin{gathered} \operatorname{card}(A \cup B) = \operatorname{card} A + \operatorname{card} B - \operatorname{card}(A \cap B), \\ \operatorname{card}(A \times B) = \operatorname{card} A \operatorname{card} B, \\ \operatorname{card} \mathcal{P}(A) = 2^{\operatorname{card} A} \end{gathered} \tag{2.20}$$

(siehe Aufgaben 49, 50, 58).

2.7 * Zahlensystem: q-adische Darstellung natürlicher Zahlen

Ein Zahlensystem ist eine Methode von bequemer Darstellung von Zahlen. Im Alltag benutzt man die dezimale Darstellung. In Rechnern wird ein anderes Zahlensystem angewandt – das Dualsystem. In diesem Abschnitt besprechen wir Darstellung natürlicher Zahlen in einem q-adischen *Zahlensystem*, wobei $q > 1$ eine fixierte natürliche Zahl ist, die die Basis des System heißt.

Bezeichnen wir mit $\mathbb{Z}_+$ die Menge von nichtnegativen ganzen Zahlen, d.h. $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$.

Satz 2.19 *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 $a_n \neq 0$ derart, dass*

$$x = \sum_{k=0}^{n} a_k q^k \qquad (= a_0 + a_1 q + ... + a_n q^n)\,. \tag{2.21}$$

Definition. Die Identität (2.21) heißt die Darstellung von x im q-adischen Zahlensystem (= *das Zahlensystem zur Basis* q). Die Identität (2.21) wird in Kurzform wie folgt geschrieben:

$$x = a_n a_{n-1} ... a_0 \ \text{ oder } \ x = (a_n a_{n-1} ... a_0)_q\,.$$

(mit dem Produkt nicht zu verwechseln).

Die Zahlen $\{0, ..., q-1\}$ heißen q-*adische Ziffern*. Die 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 höchstwertigen Stelle n muß immer positiv sein.

Die gängigsten Basen sind $q = 2$ (Dualsystem), $q = \text{zehn} := 9 + 1$ (Dezimalsystem) und $q = \text{sechzehn} := \text{zehn} + 6$ (Hexadezimalsystem).

Im Dualsystem gibt es nur zwei Ziffern 0 und 1. Die Zahlen $1, 2, 3, 4, 5, 6, 7, 8, 9$ werden im Dualsystem wie folgt dargestellt:

$$1_2,\ 10_2,\ 11_2,\ 100_2,\ 101_2,\ 110_2,\ 111_2,\ 1000_2,\ 1001_2,$$

z.B.

$$110_2 = 0 + 1 \cdot 2 + 1 \cdot 2^2 = 2 + 4 = 6.$$

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

$$C3F_{\text{hex}} = \left(F + 3 \cdot q + C \cdot q^2\right)_{\text{hex}} = \left(15 + 3 \cdot 16 + 12 \cdot 16^2\right)_{\text{dez}} = 3135_{\text{dez}}.$$

Der Satz 2.19 bedeutet, dass jede natürliche Zahl sich im q-adischen System eindeutig darstellen lässt. Für den Beweis brauchen wir das folgende Lemma.

Lemma 2.20 *(*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*

$$x = qy + r. \tag{2.22}$$

Diese Darstellung heißt die *ganzzahlige Division* von x durch q. Die Zahl y heißt der *Quotient* und r heißt der *Rest* der ganzzahligen Division.

Beweis. Für die Existenz setzen wir $y = \left[\frac{x}{q}\right]$ (=die Gaußklammer von $\frac{x}{q}$, siehe Satz 2.9), so dass

$$y \leq \frac{x}{q} < y + 1.$$

Es folgt, dass

$$qy \leq x < qy + q,$$

und somit

$$0 \leq x - qy < q.$$

Setzen wir

$$r := x - qy,$$

so dass (2.22) erfüllt ist. Die Zahl r ist ganz und erfüllt $0 \leq r < q$, woraus folgt

$$r \in \{0, ..., q-1\}.$$

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 $y \neq y'$, z.B. $y > y'$, so gilt $y - y' \geq 1$ und $q(y - y') \geq q$ während $r - r' \leq q - 1$. Somit ist $y = y'$ und dann auch $r = r'$. ■

Im nächsten Beweis benutzen wir die folgende Variante des Induktionsprinzips.

Satz 2.21 (Induktionsprinzip "von $< n$ nach n") *Sei $A(n)$ eine von $n \in \mathbb{N}$ abhängige Aussage, die die folgenden zwei Bedingungen erfüllt:*

(i) $A(1)$ ist wahr;

(ii) für jede natürliche Zahl $n \geq 2$ gilt: ist $A(k)$ wahr für alle $k < n$, so ist $A(n)$ auch wahr.

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

Beweis. Bezeichnen with mit $B(n)$ die Aussage, dass $A(k)$ für alle natürlichen Zahlen k mit $1 \leq k \leq n$ erfüllt ist. Beweisen wir per Induktion nach n, dass $B(n)$ für alle $n \in \mathbb{N}$ erfüllt ist, woraus gleiches auch für $A(n)$ folgen wird.

Induktionsanfang: $B(1)$ ist äquivalent zu $A(1)$ und somit ist wahr nach (i).

Induktionsschritt von n nach $n+1$. Beweisen wir, dass $B(n) \Rightarrow B(n+1)$. Ist $B(n)$ wahr, so ist $A(k)$ wahr für alle $k \leq n$, d.h. für alle $k < n+1$ (siehe Korollar 2.6). Nach (ii) ist dann $A(n+1)$ auch wahr. Somit ist $A(k)$ wahr für alle $k \leq n+1$, was bedeutet, dass $B(n+1)$ wahr ist. Nach dem Induktionsprinzip des Satzes 2.2 beschließen wir, dass $B(n)$ für alle $n \in \mathbb{N}$ wahr ist. ■

Beweis von dem Satz 2.19. Zunächst beweisen wir per Induktion nach $x \in \mathbb{N}$ die Existenz der q-adischen Darstellung (2.21). Bezeichnen wir mit $A(x)$ die Aussage, dass eine q-adische Darstellung von x existiert. Induktionsanfang für $x = 1$: (2.21) ist erfüllt mit $n = 0$ und $a_0 = 1$.

Induktionsschritt von $< x$ nach x bedeutet folgendes: angenommen, dass $A(y)$ für alle $y < x$ gilt, zu beweisen, dass $A(x)$ gilt. So, nehmen wir an, dass es für jedes $y \in \mathbb{N}$ mit $y < x$ eine q-adische Darstellung gibt, und beweisen wir die Existenz der q-adischen Darstellung für x.

Wir benutzen die ganzzahlige Division von x durch q, d.h. die Identität (2.22). 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 (2.22) ergibt eine q-adische Darstellung von x wie folgt:

$$\begin{aligned} x &= r + qy = r + q \sum_{k=0}^{m} b_k q^k \\ &= r + \sum_{k=0}^{m} b_k q^{k+1} \\ &= r + \sum_{l=1}^{m+1} b_{l-1} q^l \\ &= a_0 + \sum_{l=1}^{m+1} a_l q^l = \sum_{l=0}^{n} a_l q^l, \end{aligned}$$

wobei $a_0 = r$, $a_l = b_{l-1}$ für $l \geq 1$, und $n = m + 1$.

Beweisen wir jetzt die Eindeutigkeit der q-adischen Darstellung (2.21), auch per Induktion nach x. Für $x < q$ gibt es nur eine Darstellung $x = a_0$ mit $n = 0$ (ist $n \geq 1$ so gilt $x \geq a_n q^n \geq q$). 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

$$y = a_n q^{n-1} + ... + a_1. \tag{2.23}$$

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 (2.23) von y eindeutig bestimmt nach der Induktionsvoraussetzung, woraus folgt, dass der Wert von n und alle Ziffern $a_1, ..., a_n$ auch eindeutig bestimmt sind. ■

Die q-adische Darstellung von reellen Zahlen (Kommazahlen) wird später besprochen.

2.8 * Schriftliche Addition und Multiplikation

Wir besprechen kurz schriftliche Addition und Multiplikation mit Hilfe von q-adischer Darstellung. Betrachten wir zwei natürliche Zahlen im q-adischen System

$$x = (a_n...a_0)_q = \sum_{k=0}^{n} a_k q^k$$

und

$$y = (b_m...b_0)_q = \sum_{k=0}^{m} b_k q^k.$$

Ist $n > m$ so erweitern wir die Folge $\{b_k\}_{k=0}^m$ zu $\{b_k\}_{k=0}^n$ indem wir $b_k = 0$ für alle $k > m$ setzen. Im Fall $n < m$ erweitern wir analog $\{a_k\}_{k=0}^n$. So, ohne Beschränkung der Allgemeinheit nehmen wir an, dass $n = m$. Die schriftliche Addition basiert auf der offensichtlichen Identität

$$x + y = \sum_{k=0}^{n} a_k q^k + \sum_{k=0}^{n} b_k q^k = \sum_{k=0}^{n} (a_k + b_k) q^k,$$

d.h.

$$x + y = \sum_{k=0}^{n} c_k q^k, \tag{2.24}$$

wobei $c_k = a_k + b_k$. Gilt $c_k \leq q - 1$ für alle k, so erhalten wir schon die q-adische Darstellung von $x + y$:

$$x + y = (c_n ... c_0)_q .$$

Gibt es ein k mit $c_k \geq q$, so setzen wir

$$l = \min \{k : c_k \geq q\} .$$

Ganzzahlige Division von c_l durch q ergibt

$$c_l = dq + r,$$

wobei $r \in \{0, ..., q - 1\}$ and $d \in \mathbb{N}$ (übrigens ist $d = 1$ da $c_l = a_l + b_l < 2q$). Somit gilt

$$c_l q^l = r d^l + d q^{l+1},$$

und man ersetzt zwei Glieder in der Summe (2.24) wie folgt:

$$\begin{array}{rcl} c_l q^l & \text{durch} & r q^l \\ c_{l+1} q^{l+1} & \text{durch} & (c_{l+1} + d) q^{l+1} \end{array}$$

so dass der Wert der Summe bleibt unverändert und der Koeffizient vor q^l schon $\leq q - 1$ ist. Dieses Verfahren heißt *Übertrag* von d von der Stelle l nach der Stelle $l + 1$. Dann holt man Überträge in den höheren Stellen wieder bis alle Koeffizienten $\leq q - 1$ sind.

Für dieses Verfahren braucht man die *Additionstabelle* im Voraus zu erstellen, d.h. alle Summen $a + b$ mit den Ziffern $a, b \in \{0, ..., q - 1\}$ im Voraus zu wissen.

Die schriftliche Multiplikation basiert auf der Identität

$$\begin{aligned} x \cdot y &= \left(\sum_{k=0}^{n} a_k q^k \right) \cdot \left(\sum_{l=0}^{m} b_l q^l \right) \\ &= \sum_{k=0}^{n} \sum_{l=0}^{m} a_k b_l q^{k+l} \end{aligned} \tag{2.25}$$

(siehe Aufgabe 46). Um $a_k b_l$ berechnen zu können braucht man zunächst die *Multiplikationstabelle* für alle Produkte $a \cdot b$ mit den Ziffern $a, b \in \{0, ..., q - 1\}$. Diese Tabelle wird mit Hilfe der Identität

$$ab = \sum_{k=1}^{a} b = \underbrace{b + ... + b}_{a \text{ mal}}$$

erstellt. Die Doppelsumme (2.25) lässt sich wie folgt umformen:

$$x \cdot y = \sum_{j=0}^{N} c_j q^j,$$

wobei $c_j \in \mathbb{Z}_+$. Dann benutzt man Überträge wie oberhalb um alle Koeffizienten $\leq q-1$ zu machen.

2.9 * Alternative Konstruktion von $\mathbb{R}$

Wir haben die Theorie mit den Axiomen von reellen Zahlen angefangen. Hier besprechen wir kurz eine andere Konstruktion von reellen Zahlen, wenn man mit den natürlichen und ganzen Zahlen anfängt und nur danach die reellen Zahlen definiert. Es gibt zwei Ansätze für Einführung von natürlichen Zahlen.

Peano-Axiomen für natürliche Zahlen

In diesem Ansatz wird die Menge $\mathbb{N}$ axiomatisch mit Hilfe von *Peano-Axiomen* definiert.

Eine Menge $\mathbb{N}$ heißt die Menge von natürlichen Zahlen und die Elementen von $\mathbb{N}$ heißen natürliche 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 Teilmenge von $\mathbb{N}$ mit den folgenden Eigenschaften:
   (a) $1 \in M$;
   (b) $n \in M \Rightarrow F(n) \in M$.

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

Die Zahl $F(n)$ heißt der Nachfolger von n und entspricht zu $n+1$. Das 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: $n < m$ gilt genau dann, wenn $m = n + k$ für ein $k \in \mathbb{N}$. Weiter werden die üblichen Eigenschaften von Addition, Multiplikation und Ungleichheit bewiesen.

Mengentheoretisches Modell der natürlichen Zahlen

In diesem Ansatz definiert man die natürlichen Zahlen als Kardinalzahlen von endlichen Mengen. Man sagt, dass eine Menge M endlich ist, falls M zu keiner echten Teilmenge gleichmächtig ist. Die natürlichen Zahlen (inklusive Null) lassen sich als die Kardinalzahlen von endlichen Mengen definieren, d.h. als Äquivalenzklassen von gleichmächtigen Mengen. In anderen Wörter, alle gleichmächtige endliche Mengen stellen eine natürliche Zahl dar.

Man definiert die Zahl 0 als die Kardinalzahl der leeren Menge $\emptyset$. Betrachten wir die Menge $A = \{\emptyset\}$ und definieren die Zahl 1 als die Kardinalzahl von A. Sei eine natürliche Zahl n schon definiert als die Kardinalzahl einer Menge N. Dann definieren wir den Nachfolger $n+1$ als die Kardinalzahl der Menge $M = N \cup \{N\}$, wobei $\{N\}$ die Menge aus einem Element N ist. Zum Beispiel, 2 ist die Kardinalzahl der Menge $B = A \cup \{A\} = \{\emptyset, \{\emptyset\}\}$, die Zahl 3 ist die Kardinalzahl der Menge $C = B \cup \{B\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, usw.

Alternativ kann man Potenzmengen benutzen um höhere natürliche Zahlen zu erhalten. Definieren wir die Menge A als die Potenzmenge der Menge $\emptyset$, d.h. $A = \mathcal{P}(\emptyset) = \{\emptyset\}$ und die Zahl 1 als die Kardinalzahl von A. Sei B die Potenzmenge von A, d.h. $B = \mathcal{P}(A) = \{\emptyset, A\}$. Dann definieren wir 2 als die Kardinalzahl von B. Sei D die Potenzmenge von B, d.h. $D = \mathcal{P}(B) = \{\emptyset, \{\emptyset\}, \{A\}, B\}$. Dann definieren wir 4 als die Kardinalzahl von D, usw. Mit Hilfe von weiteren Potenzmengen können wir beliebig große endliche Mengen konstruieren und somit die Existenz von beliebig großen natürlichen Zahlen zeigen.

Um die Operationen $+$ und $\cdot$ auf den natürlichen Zahlen definieren benutzt man die Identitäten (2.17) und (2.20). Die Ungleichung $<$ wird wie folgt definiert: ist N eine echte Teilmenge einer endlichen Menge M, so ist die Kardinalzahl von N kleiner als die Kardinalzahl von M. Dann beweist man alle notwendigen Eigenschaften von Addition, Multiplikation und Anordnung.

Ganze und rationale Zahlen

Nach der Einführung von $\mathbb{N}$ definiert man die negativen Zahlen, die Null und die Menge $\mathbb{Z}$ als die Vereinigung $\mathbb{N} \cup (-\mathbb{N}) \cup \{0\}$, und auch die Operationen in $\mathbb{Z}$.

Weiter definiert man 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 $n \in \mathbb{Z}$ lässt sich als Element von $\mathbb{Q}$ betrachten 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 $<$ auf $\mathbb{Q}$: die Ungleichung $\frac{a}{b} < \frac{c}{d}$ mit positiven b und d gilt genau dann wenn $ad < bc$. Für negativen Werten von den Nennern gibt es eine offensichtliche Modifikation.

Man zeigt, dass die Operationen $+, \cdot$ und die Relation $<$ auch für Äquivalenzklassen wohldefiniert sind und mit den Operationen $+, \cdot$ bzw mit der Relation $<$ auf $\mathbb{Z}$ kompatibel sind. Die so definierte Menge $\mathbb{Q}$ ist ein angeordneter Körper. Allerdings erfüllt $\mathbb{Q}$ das Vollständigkeitsaxiom nicht.

Reelle Zahlen als Dedekindsche Schnitte

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 gilt auch $y \in S$ für alle $y < x$;
3. S hat kein maximales Element.

Zu jedem $a \in \mathbb{Q}$ entspricht ein Dedekindscher Schnitt

$$S_a = \{x \in \mathbb{Q} : x < a\}.$$

Aber es gibt die Schnitte, die keiner rationalen Zahl entsprechen, zum Beispiel

$$S = \left\{x \in \mathbb{Q} : x \le 0 \quad \vee \quad x^2 < 2\right\}$$

(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 Elemente von $\mathbb{R}$ identifizieren.

Man definiert die Operationen $+, \cdot$ und die Relation $\le$ auf $\mathbb{R}$. Zum Beispiel, für Schnitte S und T ist die Summe der folgende Schnitt:

$$S + T = \{x + y : x \in S, \, y \in T\}.$$

Die Ungleichung $S \le T$ gilt genau dann, wenn

$$\forall x \in S \ \ \exists y \in T \ \ x \le y.$$

Man beweist, dass alle Axiome von $\mathbb{R}$ erfüllt sind, insbesondere das Vollständigkeitsaxiom.

2.10 * Kardinalität unendlicher Mengen

Mächtigkeit von Mengen

Wir haben schon die Gleichmächtigkeit von zwei Mengen definiert.

Definition. Die Mengen X und Y heißen *gleichmächtig* oder *äquivalent*, falls es eine bijektive Abbildung $f : X \to Y$ gibt. In diesem Fall schreibt man $X \sim Y$.

Jede endliche nicht-leere Menge X ist äquivalent zu einer der Mengen $\mathcal{E}_n = \{1, ..., n\}$ mit $n \in \mathbb{N}$. In diesem Fall sagen wir, dass die Anzahl von Elementen von X ist n und schreiben $\operatorname{card} X = n$. Man sagt auch, dass die Kardinalzahl von X ist gleich n.

Die Kardinalzahlen von unendlichen Mengen können als Verallgemeinerung von natürlichen Zahlen betrachtet werden.

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

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

Behauptung. *Die Relation $X \preceq Y$ 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 Ungleichheit zwischen Mengen erfüllt die folgenden Eigenschaften:

1. $\mathcal{E}_n \preceq \mathcal{E}_m \Leftrightarrow n \leq m$ (folgt aus dem Schubfachprinzip)
2. $X \preceq X$ (trivial)
3. $X \preceq Y \ \wedge Y \preceq Z \ \Rightarrow X \preceq Z$ (da Komposition von injektiven Abbildungen wieder injektiv ist)
4. $X \preceq Y \ \wedge \ Y \preceq X \Rightarrow X = Y$
5. für je zwei Mengen X, Y gilt $X \preceq Y \ \vee \ Y \preceq X$.

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

Abzählbare Mengen

Definition. Eine Menge X heißt *abzählbar* falls $X \sim \mathbb{N}$.

In diesem Fall schreibt man

$$\operatorname{card} X = \aleph_0,$$

und sagt, dass die Kardinalzahl von X *Aleph-null* ist 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, wie wir unterhalb sehen.

Der nächste Satz enthält wichtige Eigenschaften von abzählbaren Mengen.

Satz 2.22 (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 $X \preceq \mathbb{N}$, so gilt $X \sim \mathbb{N}$. In der Tat bedeutet $X \preceq \mathbb{N}$, 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 $Y \succeq \mathbb{N}$.

Der Teil (b) ist äquivalent zu $\mathbb{N}\times\mathbb{N} \sim \mathbb{N}$. Zum Vergleich erinnern wir uns daran, dass nach (2.20) $\mathcal{E}_n \times \mathcal{E}_n \sim \mathcal{E}_{n^2}$, insbesondere

$$\operatorname{card} \mathcal{E}_n < \operatorname{card}(\mathcal{E}_n \times \mathcal{E}_n)$$

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 2.7 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

$$f(n) = \min(X \setminus X_n). \tag{2.26}$$

Nach dem 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 (2.26), 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 $\mathbb{N}^2$ aufzählen indem man die Mengen D_k nacheinander aufzählt.

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

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 Aufgabe 70), 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 $X \preceq \mathbb{N}^2$. Nach (b) haben wir $\mathbb{N}^2 \sim \mathbb{N}$ und somit $X \preceq \mathbb{N}$. Da X unendlich ist, so erhalten wir nach (a), dass $X \sim \mathbb{N}$. ■

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 2.23 *Die Mengen $\mathbb{Z}$ und $\mathbb{Q}$ sind abzählbar, d.h. $\mathbb{Q} \sim \mathbb{Z} \sim \mathbb{N}$.*

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

Überabzählbare Mengen

Wir schreiben $X \prec Y$ falls $X \preceq Y$ aber $X \not\sim Y$.

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

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

Satz 2.24 *Die Menge* $\mathbb{R}$ *ist überabzählbar, d.h.* $\mathbb{N} \prec \mathbb{R}$.

Gilt $X \sim \mathbb{R}$ so schreibt man

$$\operatorname{card} X = \mathfrak{c}$$

und sagt, dass die Kardinalzahl von X gleich *Kontinuum* ist.

Da $\mathbb{Q} \sim \mathbb{N} \prec \mathbb{R}$, 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 Aufgabe 55).

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 $I_1 = [x_1 + 1, x_1 + 2]$. 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 $I_{n+1} = [a_n, \frac{x_{n+1}+a_n}{2}]$ oder $I_{n+1} = \left[\frac{b_n+x_{n+1}}{2}, b_n\right]$.

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 2.25 *Seien X, Y zwei Mengen mit $Y \preceq \mathbb{N}$.*
(a) Ist X unendlich, so gilt $X \cup Y \sim X$.
(b) Ist X überabzählbar, so gilt $X \setminus Y \sim X$.

Insbesondere gelten

$$\mathbb{R} \cup Y \sim \mathbb{R} \text{ und } \mathbb{R} \setminus Y \sim \mathbb{R}.$$

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

Beweis. (a) Zunächst bemerken wir, dass

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

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

$$X \sqcup Y \sim X.$$

Jede unendliche Menge erhält eine abzählbare Teilmenge. 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

$$X \sqcup Y = (X_0 \sqcup X_1) \sqcup Y = (X_0 \sqcup Y) \sqcup X_1. \tag{2.27}$$

Da $X_0 = \mathbb{N}$ und $Y \preceq \mathbb{N}$, so erhalten wir nach Satz 2.22, dass

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

was zusammen mit (2.27) 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 $X \cap Y \preceq Y \preceq \mathbb{N}$, so können wir $X \cap Y$ in Y umbenennen und somit voraussetzen, dass $Y \subset X$. Dann haben wir

$$X = (X \setminus Y) \cup Y. \tag{2.28}$$

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

$$(X \setminus Y) \cup Y \sim X \setminus Y,$$

woraus $X \sim X \setminus Y$ folgt. ■

Mit Hilfe von Satz 2.25 kann man beweisen, dass alle Intervalle (a, b) , $[a, b]$, $[a, b)$, $(a, b]$ mit $a < b$ die Kardinalzahl $\mathfrak{c}$ haben.

Algebraische und transzendente Zahlen

Die folgenden Teilmengen von reellen Zahlen sind uns schon bekannt:

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

und wir wissen, dass

$$\mathbb{N} \sim \mathbb{Z} \sim \mathbb{Q} \prec \mathbb{R}.$$

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

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

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_n = 0 \tag{2.29}$$

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. Offensichtlich gelten die Inklusionen $\mathbb{Q} \subset \mathbb{A} \subset \mathbb{R}$ und $\mathbb{Q} \neq \mathbb{A}$. Es möglich zu beweisen, dass die Summe, die Differenz, das Produkt und der Quotient zweier algebraischen Zahlen wieder algebraisch sind, und somit ist $\mathbb{A}$ ein Körper.

Satz 2.26 *Die Menge $\mathbb{A}$ ist abzählbar.*

Beweis. Die Menge von allen Polynomen (2.29) mit rationalen Koeffizienten ist abzählbar nach den Satz 2.22(b). Da jedes Polynom (2.29) höchstens n Nullstellen hat, so folgt es aus dem Satz 2.22(c), dass $\mathbb{A}$ abzählbar ist. ■

Nicht-algebraische reelle Zahlen heißen *transzendent.* Es folgt aus den Sätzen 2.24, 2.25, 2.26 dass die Menge $\mathbb{R} \setminus \mathbb{A}$ von transzendenten Zahlen die Kardinalität Kontinuum hat, insbesondere diese Menge nichtleer ist. Beispiele von transzendenten Zahlen sind die Zahlen π und e, die später definiert werden.

Mächtigkeit von Potenzmenge

Satz 2.27 *Für jede Menge X gilt $X \prec \mathcal{P}(X)$.*

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

1. es existiert eine injektive Abbildung $f : X \to Y$,

2. es existiert keine bijektive Abbildung $f : X \to Y$.

Eine injective Abbildung $f : X \to \mathcal{P}(X)$ kann man wie folgt definieren: $f(x) = \{x\}$, wobei $\{x\}$ eine Teilmenge von X ist, die aus einem Element x besteht.

Sei $f : X \to \mathcal{P}(X)$ eine beliebige Abbildung. Beweisen wir, dass f nicht surjektiv ist (und somit auch nicht bijektiv). Betrachten wir die Menge

$$S = \{x \in X : x \notin f(x)\}. \tag{2.30}$$

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 $\mathcal{P}(X)$. 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 (2.30) $y \notin f(y) = S$, was ist unmöglich.

2. Ist $y \notin S$ so gilt nach (2.30) $y \in f(y) = S$ – auch unmöglich.

Deshalb führt die Bedingung $f(y) = S$ zum Widerspruch, woraus folgt, dass f keine surjektive Abbildung ist. ■

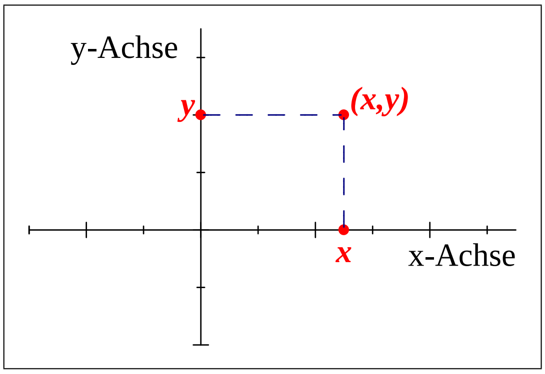
Chapter 3

Komplexe Zahlen

3.1 Die Menge von komplexen Zahlen

Die Produktmenge $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$ heißt die Ebene. Die Elemente von der Ebene sind die Paaren (x, y) von reellen Zahlen, und sie werden auch *Punkte* genannt. Die Zahlen x und y heißen die (kartesischen) *Koordinaten* oder *Komponenten* von dem Punkt $(x, y) \in \mathbb{R}^2$.

Man stellt die Ebene als Produkt von zwei Geraden (=zwei Kopien von $\mathbb{R}$) dar: eine ist waagerecht und heißt die x-Achse, und andere ist senkrecht und heißt die y-Achse.



In der Ebene $\mathbb{R}^2$ werden Addition und Multiplikation definiert wie folgt.

Definition. Die Menge $\mathbb{C}$ von *komplexen Zahlen* ist die Ebene $\mathbb{R}^2$ mit den folgenden Operationen $+$ und $\cdot$:

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

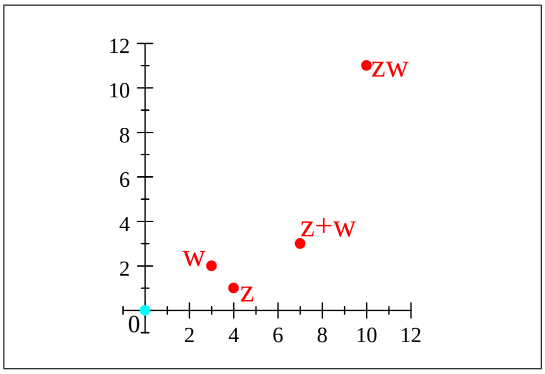
Die Elemente von $\mathbb{C}$ heißen komplexe Zahlen.

Beispiel. Für die komplexen Zahlen $z = (4, 1)$ und $w = (3, 2)$ berechnen wir:

$$z + w = (4, 1) + (3, 2) = (7, 3)$$

und

$$z \cdot w = (4, 1) \cdot (3, 2) = (4 \cdot 3 - 1 \cdot 2,\ 4 \cdot 2 + 1 \cdot 3) = (10, 11)\,.$$



Satz 3.1 *Die Addition von komplexen Zahlen erfüllt alle Axiome von Addition: Kommutativ- und Assoziativgesetze, Existenz von Nullelement und Negative.*

Somit ist $\mathbb{C}$ eine kommutative Gruppe bezüglich Addition.

Beweis. Kommutativ- und Assoziativgesetze für komplexe Zahlen sind die Identitäten

$$z_1 + z_2 = z_2 + z_1 \text{ und } (z_1 + z_2) + z_3 = z_1 + (z_2 + z_3) \quad \forall z_1, z_2, z_3 \in \mathbb{C}.$$

Da Addition komplexer Zahlen komponentenweise definiert ist, so folgen diese Identitäten direkt aus den Axiomen der Addition in $\mathbb{R}$.

Das Nullelement ist $(0, 0)$, da

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

Das Negative von $z = (x, y)$ ist

$$-z := (-x, -y),$$

da $z + (-z) = 0$. ■

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

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

Wir identifizieren[1] jede reelle Zahl x mit der komplexen Zahl $(x, 0)$ und somit betrachten weiter die Menge $\mathbb{R}$ als eine Teilmenge von $\mathbb{C}$. Wichtig ist, dass die Operationen Addition und Multiplikation in $\mathbb{R}$ und $\mathbb{C}$ übereinstimmen.

Somit wird die komplexe Zahl $(x, 0)$ einfach mit x bezeichnen. Insbesondere wird das Nullelement $(0, 0)$ mit 0 bezeichnet, und das Einheitselement $(1, 0)$ – einfach mit 1.

Die komplexe Zahl $(0, 1)$ ist besonders wichtig und wird mit i bezeichnet.

Definition. Die Zahl $i = (0, 1)$ heißt die *imaginäre Einheit*.

[1] Logisch bedeutet das, dass wir die Menge $\mathbb{R}$ durch die Menge $\mathbb{R}' = \{(x, 0) : x \in \mathbb{R}\}$ ersetzen und $\mathbb{R}'$ in $\mathbb{R}$ umbenennen.

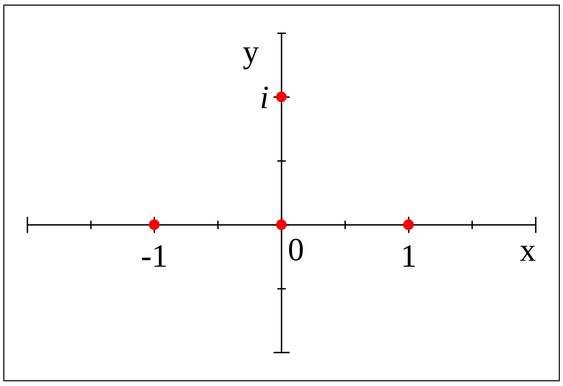
Nach Definition der Multiplikation haben wir

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

so that

$$i^2 = -1.$$

Diese Identität ist eine von Motivationen die Menge von reellen Zahlen zu erweitern, da es in $\mathbb{R}$ keine Zahl x mit $x^2 = -1$ gibt.



18.05.18 H1

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

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

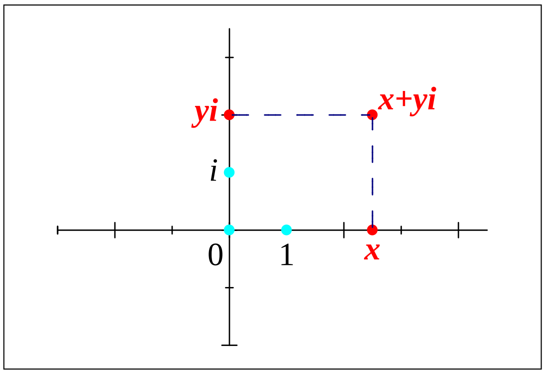
and analog

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

Es folgt, dass

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

d.h. jede komplexe Zahl (x,y) lässt sich in der Form $x+yi$ (oder $x+iy$) darstellen. Diese Darstellung heißt die *algebraische* (oder kartesische) Form der komplexen Zahl. Die komplexen Zahlen der Form yi heißen *imaginär*. Somit sind die komplexen Zahlen die Summen von reellen und imaginären Zahlen.



Für jede komplexe Zahl $z = x + iy$ heißt die Komponente x *Realteil* und y – *Imaginärteil* von z. Man schreibt:

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

so dass
$$z = \operatorname{Re} z + i \operatorname{Im} z.$$
Offensichtlich ist z reell genau dann, wenn $\operatorname{Im} z = 0$ und imaginär – wenn $\operatorname{Re} z = 0$.

In der algebraischen Form sehen die Rechenregeln von komplexen Zahlen wie folgt aus:
$$\boxed{(x + yi) + (x' + y'i) = (x + x') + (y + y')\, i}$$
und
$$\boxed{(x + yi)\,(x' + y'i) = (xx' - yy') + (xy' + x'y)\, i}.$$
Man addiert und multipliziert die Ausdrücke $x + yi$ und $x' + y'i$ genau so, als ob i reell wäre, aber mit der zusätzlichen Regel $i^2 = -1$.

Beispiel. Für $z = 4 + i$ und $w = 3 + 2i$ gilt
$$z + w = (4 + i) + (3 + 2i) = 7 + 3i$$
und
$$zw = (4 + i)\,(3 + 2i) = (4 \cdot 3 - 1 \cdot 2) + \ (4 \cdot 2 + 1 \cdot 3)\, i = 10 + 11i.$$

3.2 Eigenschaften von Multiplikation

Die Menge $\mathbb{C}$ hat offensichtlich das Einheitselement $1 = (1, 0)$, da
$$(x, y) \cdot 1 = 1 \cdot (x, y) = (x, y)\,.$$
Hier beweisen wir weitere Eigenschaften der Multiplikation.

Satz 3.2 *Multiplikation von komplexen Zahlen erfüllt die Kommutativ-, Assoziativ- und Distributivgesetze, d.h. für alle $z_1, z_2, z_3 \in \mathbb{C}$ gelten die folgenden Identitäten:*

(a) $z_1 z_2 = z_2 z_1$ *(Kommutativgesetz)*

(b) $(z_1 z_2)\, z_3 = z_1\, (z_2 z_3)$ *(Assoziativgesetz)*

(c) $(z_1 + z_2)\, z_3 = z_1 z_3 + z_2 z_3$ *(Distributivgesetz)*

Somit gelten in $\mathbb{C}$ alle Axiome von Addition und Multiplikation außer Existenz von Inverse, was später bewiesen wird.

Beweis. Setzen wir $z_k = x_k + iy_k$ wobei x_k, $y_k \in \mathbb{R}$.

(a) Wir haben nach Definition
$$z_1 z_2 = (x_1 + y_1 i)\,(x_2 + y_2 i) = (x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)\, i$$
und
$$z_2 z_1 = (x_2 + y_2 i)\,(x_1 + y_1 i) = (x_2 x_1 - y_2 y_1) + (x_2 y_1 + y_2 x_1)\, i,$$
so dass $z_1 z_2 = z_2 z_1$ nach dem Kommutativgesetz in $\mathbb{R}$.

(c) Wir haben

$$z_1 z_3 = (x_1 x_3 - y_1 y_3) + (x_1 y_3 + y_1 x_3)\, i$$

und

$$z_2 z_3 = (x_2 x_3 - y_2 y_3) + (x_2 y_3 + y_2 x_3)\, i,$$

woraus folgt

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

wobei wir das Distributivgesetz in $\mathbb{R}$ benutzt haben,

(b) Jetzt beweisen das Assoziativgesetz

$$(z_1 z_2)\, z_3 = z_1\, (z_2 z_3)\,. \tag{3.1}$$

Wir haben

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

und somit

$$\begin{aligned} (z_1 z_2)\, z_3 &= ((x_1 x_2 - y_1 y_2) + (x_1 y_2 + y_1 x_2)\, i)\, (x_3 + i y_3) \\ &= (x_1 x_2 - y_1 y_2)\, x_3 - (x_1 y_2 + y_1 x_2)\, y_3 \\ &\quad + ((x_1 x_2 - y_1 y_2)\, y_3 + (x_1 y_2 + y_1 x_2)\, x_3)\, i \\ &= x_1 x_2 x_3 - y_1 y_2 x_3 - x_1 y_2 y_3 - y_1 x_2 y_3 \\ &\quad + (x_1 x_2 y_3 - y_1 y_2 y_3 + x_1 y_2 x_3 + y_1 x_2 x_3)\, i. \end{aligned} \tag{3.2}$$

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 (3.2) bei dem folgenden Wechsel (Permutation) von den Indizen:

$$1 \mapsto 2, \quad 2 \mapsto 3, \quad 3 \mapsto 1.$$

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

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

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

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

Das Gleiche gilt für den Imaginärteil, woraus (3.1) folgt. ■

3.3 Konjugation

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

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

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

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

Satz 3.3 *Für Konjugation gelten die folgenden Identitäten, für alle* $z, w \in \mathbb{C}$*:*

(a) $z + \overline{z} = 2 \operatorname{Re} z$

(b) $z\overline{z} = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2$

(c) $\overline{z + w} = \overline{z} + \overline{w}$ (Additivität)

(d) $\overline{zw} = \overline{z}\,\overline{w}$ (Multiplikativität)

Beweis. $(a) + (b)$ Setzen wir $z = x + iy$. Dann gilt $\overline{z} = x - iy$ und

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

und

$$z\overline{z} = (x + iy)(x - iy) = \left(x^2 + y^2\right) + i(xy - xy) = x^2 + y^2.$$

$(c) + (d)$ Sei $w = u + iv$. Für die Summe haben wir

$$\begin{aligned} \overline{z + w} &= \overline{(x + u) + i(y + v)} = (x + u) - i(y + v) \\ &= (x - iy) + (u - iv) = \overline{z} + \overline{w}. \end{aligned}$$

Für das Produkt gilt

$$zw = (x + iy)(u + iv) = (xu - yv) + i(xv + yu),$$

und

$$\overline{z}\,\overline{w} = (x - iy)(u - iv) = (xu - vy) - i(xv + yu) = \overline{zw}$$

was zu beweisen war. ∎

3.4 Betrag

Definition. Für jede komplexe Zahl $z = x + iy$ definieren wir den Betrag $|z|$ als die einzige nicht negative reelle Zahl mit

$$\boxed{|z|^2 = z\overline{z}},$$

oder, äquivalent,

$$\boxed{|z| = \sqrt{z\overline{z}}}.$$

Nach dem Satz 3.4 ist $z\overline{z}$ eine nicht negative reelle Zahl, so dass $\sqrt{z\overline{z}}$ wohldefiniert ist. Für $z \in \mathbb{R}$ stimmt diese Definition des Betrages mit dem Definition des Betrages für reellen Zahlen überein, da in diesem Fall $\overline{z} = z$.

Nach dem Satz 3.3 gilt für $x = x + iy$ die Identität

$$\boxed{|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2} = \sqrt{x^2 + y^2}}, \tag{3.3}$$

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

Es folgt aus (3.3), dass

$$|\overline{z}| = |z|$$

und

$$|\operatorname{Re} z| \leq |z| \quad \text{und} \quad |\operatorname{Im} z| \leq |z|,$$

Satz 3.4 *Der Betrag hat die folgenden Eigenschaften, für alle $z, w \in \mathbb{C}$:*

(*a*) $|zw| = |z|\,|w|$ (Multiplikativität)

(*b*) $|z + w| \leq |z| + |w|$ (Dreiecksungleichung)

Beweis. (*a*) Nach Multiplikativität der Konjugation erhalten wir

$$|zw|^2 = zw\,\overline{zw} = zw\,\overline{z}\,\overline{w} = (z\overline{z})\,(w\overline{w}) = |z|^2\,|w|^2 = (|z|\,|w|)^2,$$

woraus $|zw| = |z|\,|w|$ folgt.

(*b*) Bemerken wir zunächst, dass

$$\begin{aligned} |z+w|^2 &= (z+w)\,\overline{(z+w)} \\ &= (z+w)\,(\overline{z} + \overline{w}) \\ &= z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} \\ &= |z|^2 + z\overline{w} + \overline{z\overline{w}} + |w|^2 \\ &= |z|^2 + 2\operatorname{Re}(z\overline{w}) + |w|^2. \end{aligned} \tag{3.4}$$

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

$$|z+w|^2 \leq |z|^2 + 2\,|z\overline{w}| + |w|^2 = |z|^2 + 2\,|z|\,|w| + |w|^2 = (|z| + |w|)^2,$$

woraus die Dreiecksungleichung folgt. ∎

Mit Hilfe von Betrag definiert man den Begriff von Abstand auf der Ebene. Für beliebige Punkte $a, b \in \mathbb{R}^2 = \mathbb{C}$ definieren wir den *Abstand* $d\,(a,b)$ zwischen a und b wie folgt:

$$d\,(a,b) = |a - b|.$$

Man kann d als eine Abbildung von $\mathbb{C} \times \mathbb{C}$ nach $\mathbb{R}$ betrachten. Insbesondere gilt $d\,(a,0) = |a|$.

Der Abstand hat die folgenden Eigenschaften:

- $d\,(a,a) = 0$ und $d\,(a,b) > 0$ für $a \neq b$ *(Positivität)*;
- $d\,(a,b) = d\,(b,a)$ *(Symmetrie)*;
- $d\,(a,b) \leq d\,(a,c) + d\,(c,b)$ für alle $a, b, c \in \mathbb{C}$ *(Dreiecksungleichung)*.

3.5 Inverse und Division

Das Inverse von $z \in \mathbb{C}$ ist eine komplexe Zahl z^{-1} mit

$$zz^{-1} = z^{-1}z = 1.$$

Im nächsten Satz beweisen wir die Existenz von Inverse und somit Existenz von Division von komplexen Zahlen.

Satz 3.5 (a) *Jedes $z \in \mathbb{C} \setminus \{0\}$ hat das Inverse wie folgt:*

$$\boxed{z^{-1} = |z|^{-2}\,\overline{z}}. \tag{3.5}$$

Es gilt auch

$$\left|z^{-1}\right| = \frac{1}{|z|}.$$

(b) *Für alle $a, b \in \mathbb{C}$ mit $a \neq 0$ hat die Gleichung $aw = b$ eine eindeutige Lösung $w = a^{-1}b$, was mit $\frac{b}{a}$ bezeichnet wird. Es gelten die Identitäten:*

$$\boxed{\frac{b}{a} = |a|^{-2}\,\bar{a}b \quad \textit{und} \quad \left|\frac{b}{a}\right| = \frac{|b|}{|a|}}. \tag{3.6}$$

Beweis. (a) Wir haben

$$\left(|z|^{-2}\,\overline{z}\right) z = |z|^{-2}\left(\overline{z}z\right) = |z|^{-2}\,|z|^2 = 1,$$

so dass $|z|^{-2}\,\overline{z}$ das Inverse von z ist. Es folgt aus (3.5), dass

$$\left|z^{-1}\right| = |z|^{-2}\,|\overline{z}| = |z|^{-2}\,|z| = |z|^{-1},$$

was zu beweisen war.

(b) Die Gleichung $aw = b$ ist äquivalent zu

$$a^{-1}(aw) = a^{-1}b,$$

woraus $w = a^{-1}b$ folgt. Da $a^{-1} = |a|^{-2}\,\bar{a}$, so erhalten wir

$$w = |a|^{-2}\,\bar{a}b.$$

Auch gilt

$$\left|\frac{b}{a}\right| = \left|a^{-1}b\right| = \left|a^{-1}\right||b| = \frac{|b|}{|a|}.$$

■

Beispiel. Berechnen wir

$$\frac{4-3i}{1+2i}.$$

Man kann die Formel (3.6) benutzen oder, äquivalent, den Nenner und Zähler mit der Konjugierte des Nenners multiplizieren:

$$\frac{4-3i}{1+2i} = \frac{(4-3i)(1-2i)}{(1+2i)(1-2i)} = \frac{-2-11i}{5} = -\frac{2}{5} - \frac{11}{5}i.$$

Bemerkung. Nach den Sätzen 3.1, 3.2 und 5.5 erfüllt $\mathbb{C}$ alle Axiome von Addition und Multiplikation so dass $\mathbb{C}$ ein Körper ist. Alle Eigenschaften von reellen Zahlen, die nur aus den Axiomen von Addition und Multiplikation hergeleitet wurden, gelten auch für komplexe Zahlen: z.B. die Definition und Identitäten von Potenzen z^n mit $z \in \mathbb{C}$ und $n \in \mathbb{Z}$, der binomische Lehrsatz, usw.

3.6 Funktionen und ihre Graphen

23.05.18 H13

Eine Funktion ist eine Abbildung $f : X \to Y$ wobei X und Y Teilmengen von $\mathbb{R}$ oder $\mathbb{C}$ sind. In diesem Abschnitt betrachten wir reellwertige Funktionen $f : I \to \mathbb{R}$ wobei I ein Intervall in $\mathbb{R}$ ist oder eine Vereinigung von Intervallen.

Der *Graph* einer Funktion $f : I \to \mathbb{R}$ ist die folgende Menge:

$$G_f = \left\{(x, f(x)) \in \mathbb{R}^2 : x \in I\right\} = \left\{(x, y) \in \mathbb{R}^2 : x \in I,\ y = f(x)\right\}.$$

In anderen Wörter, G_f ist eine Teilmenge von $\mathbb{R}^2$, die von der Gleichung $y = f(x)$ bestimmt wird.

3.6.1 Gerade und lineare Funktion

Gerade

Definition. *Gerade* ist eine Teilmenge L von $\mathbb{R}^2$ der Form

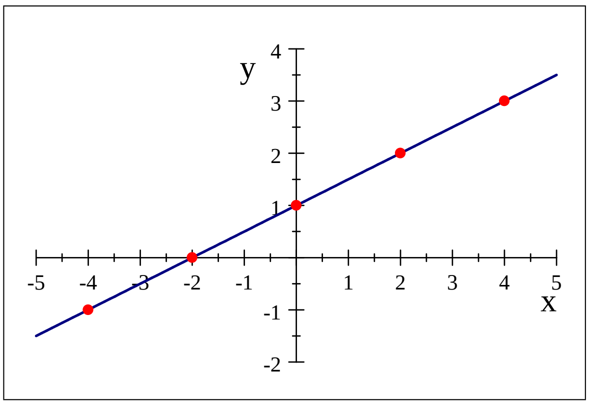
$$L = \left\{(x, y) \in \mathbb{R}^2 : ax + by = c\right\}$$

wobei a, b, c gegebene reelle Zahlen mit $a \neq 0$ oder $b \neq 0$ sind. Man sagt, dass die Gerade L von der Gleichung

$$ax + by = c$$

bestimmt wird.

Z.B. die Gleichung $y = 0$ bestimmt die x-Achse, und die Gleichung $x = 0$ bestimmt die y-Achse, so dass die beiden Achsen Geraden sind. Die Gleichung $x - 2y = -2$ bestimmt die Gerade auf dem Bild unterhalb. Diese Gerade enthält, z.B., die Punkte $(0, 1)$, $(-2, 0)$, $(2, 2)$, $(4, 3)$ usw.



Lineare Funktion

Definition. *Lineare Funktion* ist eine Funktion der Form

$$f(x) = ax + b$$

mit reellen $a, b \in \mathbb{R}$. Der Definitionsbereich von f ist $\mathbb{R}$.

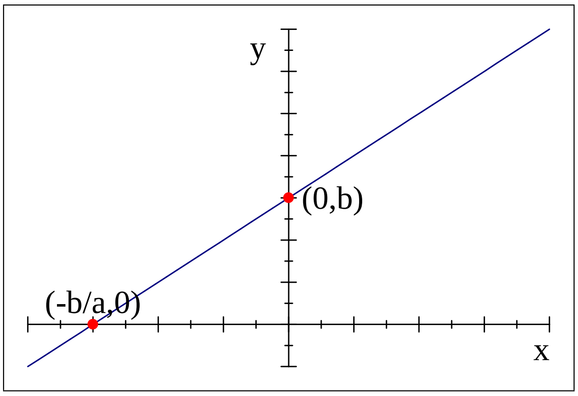
Der Graph G_f hat die Gleichung

$$y = ax + b, \tag{3.7}$$

was offensichtlich eine Gerade ist. Im Fall $a = 0$ ist diese Gerade waagerecht.

Sei $a \neq 0$. Offensichtlich liegen die folgenden zwei Punkten auf der Gerade G_f:

$$(0, b) \quad \text{und} \quad (-\frac{b}{a}, 0).$$



3.6.2 Potenzfunktion

Definition. *Potenzfunktion* is eine Funktion der Form

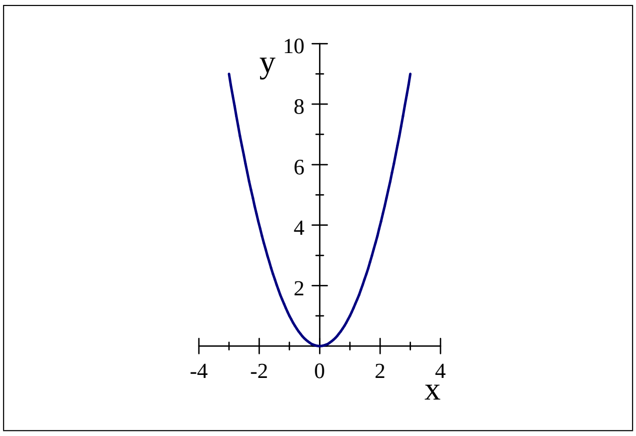
$$f(x) = x^n,$$

wobei n eine nicht Null ganze Zahl ist. Für $n > 0$ hat diese Funktion den Definitionsbereich $\mathbb{R}$, und für $n < 0$ ist der Definitionsbereich $\mathbb{R} \setminus \{0\}$.

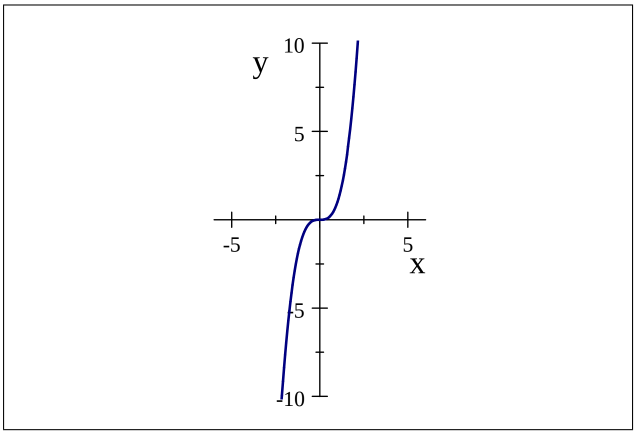
Im Fall $n = 1$ bekommen wir eine lineare Funktion.

Definition. Der Graph der Potenzfunktion mit $n \geq 2$ heißt *Parabel* n-ter Ordnung. Der Graph der Potenzfunktion mit $n < 0$ heißt *Hyperbel* n-ter Ordnung.

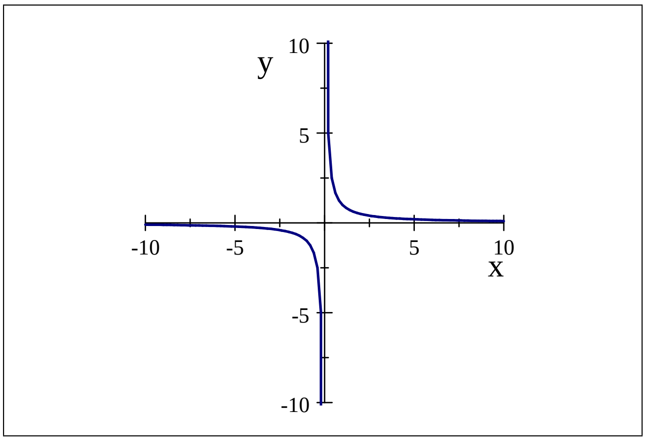
Hier ist die Parabel 2-ter Ordnung, d.h. der Graph der Funktion $y = x^2$:



Hier ist die Parabel 3-ter Ordnung, d.h. der Graph der Funktion $y = x^3$:



Hier ist die Hyperbel 1-ter Ordnung, d.h. der Graph der Funktion $y = \frac{1}{x}$:



3.6.3 Kreis

Definition. Der *Einheitskreis* ist die folgende Menge:

$$\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\} = \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\right\}.$$

Z.B., die Punkte $1, -1, i, -i$ sind Elemente von $\mathbb{S}$. Auch der Punkt $\frac{3}{5} + \frac{4}{5}i$ liegt in $\mathbb{S}$.

Man sagt: die Gleichung von $\mathbb{S}$ ist $x^2 + y^2 = 1$. Die Menge $\mathbb{S}$ ist kein Graph einer Funktion, aber $\mathbb{S}$ lässt sich als Vereinigung von zwei Graphen darstellen. Definieren wir zwei Funktionen:

$$f(x) = \sqrt{1 - x^2}, \quad x \in [-1, 1]$$

und

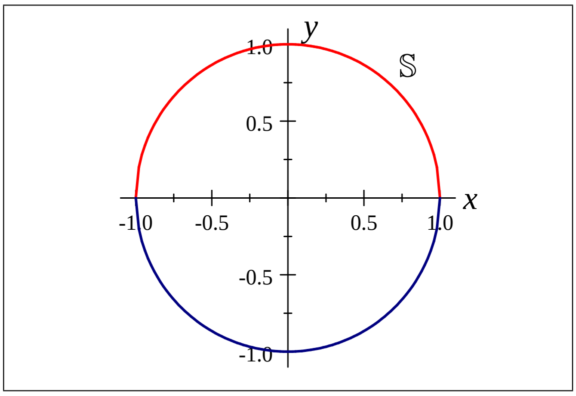
$$g(x) = -\sqrt{1 - x^2}, \quad x \in [-1, 1].$$

Die Gleichung $x^2 + y^2 = 1$ ist äquivalent zu

$$y = \sqrt{1 - x^2} \quad \text{oder} \quad y = -\sqrt{1 - x^2}$$

und $x \in [-1, 1]$, woraus folgt dass

$$\mathbb{S} = G_f \cup G_g.$$



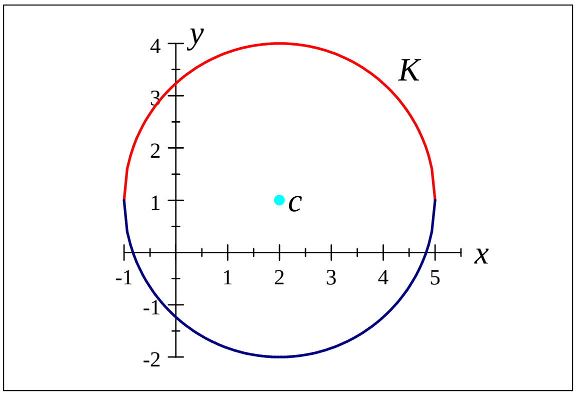
Ein allgemeiner Kreis wird wie folgt definiert.

Definition. Der Kreis von Radius $r > 0$ und mit dem Mittelpunkt $c \in \mathbb{C}$ ist die Menge
$$K_{c,r} = \{z \in \mathbb{C} : |z - c| = r\} .$$

Im Fall $c = 0$ und $r = 1$ erhalten wir den Einheitskreis
$$\mathbb{S} = K_{0,1} = \{z \in \mathbb{C} : |z| = 1\} .$$

Seien $c = a + bi$ und $z = x + yi$. Dann ist die Gleichung $|z - c| = r$ äquivalent zu
$$(x - a)^2 + (y - b)^2 = r^2.$$

Hier ist der Kreis $K = K_{c,r}$ mit $c = (2, 1)$ und $r = 3$.

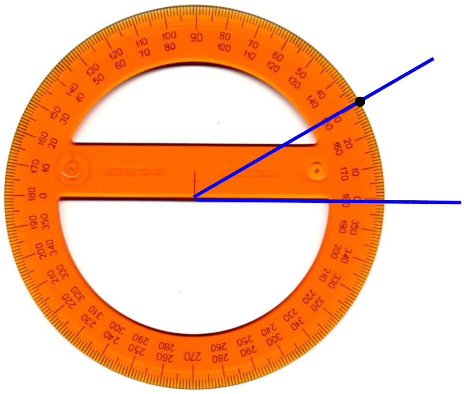


3.7 * Begriff von Winkel und Geometrie der Ebene

Winkel und Winkelfunktionen

Definition. Die Punkte von dem Einheitskreis $\mathbb{S}$ heißen *Polarwinkel* oder einfach *Winkel*.

Motivation für diese Definition von Winkel ist wie folgt. Normalerweise versteht man unter Winkel eine Figur aus zwei Halbgeraden mit gleichem Anfangspunkt. Die Winkel lassen sich messen, z.B. mit einem Winkelmesser.



Man richtet eine Gerade nach der Nullmarke des Gerätes, und dann zeigt die zweite Gerade den Punkt auf dem Kreis, neben denen das Gradmaß geschrieben ist. Deshalb identifizieren wir den Winkel mit einem Punkt auf dem Kreis. Eine Zuordnung zwischen den Punkten des Kreises und reellen Zahlen (Gradmaß) wird später bestimmt werden.

Obwohl $\mathbb{S}$ eine Teilmenge von $\mathbb{C}$ ist, wir betrachten $\mathbb{S}$ unabhängig von $\mathbb{C}$ und sogar benutzen unterschiedliche Notation für die Elemente von $\mathbb{S}$ und $\mathbb{C}$. Insbesondere werden die Elemente von $\mathbb{S}$ mit griechischen Buchstaben bezeichnet und die Elemente von $\mathbb{C}$ – wie üblich mit lateinischen Buchstaben.

Zu jedem Polarwinkel $\alpha \in \mathbb{S}$ entspricht eine komplexe Zahl z_α, die α als Element von $\mathbb{C}$ darstellt. Eigentlich bezeichnen z_α und α den gleichen Punkt in $\mathbb{C}$ aber wir betrachten α als Element von $\mathbb{S}$ und z_α – als Element von $\mathbb{C}$.

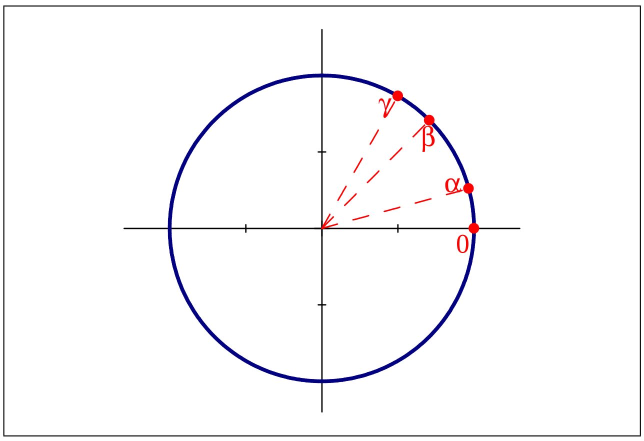
Definition. In $\mathbb{S}$ wird die Operation *Winkeladdition* wie folgt definiert: für $\alpha, \beta \in \mathbb{S}$ definieren wir den Winkel $\alpha + \beta \in \mathbb{S}$ durch die Identität

$$z_{\alpha+\beta} = z_\alpha z_\beta. \tag{3.8}$$

Bemerken wir, dass nach dem Satz 3.4 gilt

$$|z_\alpha z_\beta| = |z_\alpha|\,|z_\beta| = 1,$$

so dass $z_\alpha z_\beta$ wirklich einen Winkel bestimmt, der $\alpha + \beta$ genannt wird. Es folgt aus dem Satz 3.2, dass die Winkeladdition kommutativ und assoziativ ist.



$\alpha + \beta = \gamma$

Definieren wir den *Nullwinkel* $0 \in \mathbb{S}$ mit

$$z_0 = 1.$$

Dann gilt

$$\alpha + 0 = \alpha$$

da

$$z_{\alpha+0} = z_\alpha z_0 = z_\alpha \cdot 1 = z_\alpha.$$

Für jedes $\alpha \in \mathbb{S}$ definieren wir das *Negative* $-\alpha \in \mathbb{S}$ mit

$$z_{-\alpha} := (z_\alpha)^{-1} = \overline{z_\alpha}. \tag{3.9}$$

Nach dem Satz 3.5 gilt $\left|(z_a)^{-1}\right| = |z_\alpha|^{-1} = 1$ so dass $(z_\alpha)^{-1}$ einen Winkel bestimmt. In (3.9) haben wir auch benutzt, dass

$$(z_\alpha)^{-1} = |z_\alpha|^{-2} \overline{z_\alpha} = \overline{z_\alpha}.$$

Es gilt

$$\alpha + (-\alpha) = 0,$$

da

$$z_{\alpha+(-\alpha)} = z_\alpha z_{-\alpha} = z_\alpha (z_\alpha)^{-1} = 1 = z_0.$$

Deswegen erfüllt die Winkeladdition alle Axiome von Addition, und somit ist $\mathbb{S}$ eine additive Gruppe.

Bezeichnen wir mit π den Winkel mit

$$\boxed{z_\pi = -1}.$$

Dann gilt

$$z_{-\pi} = \overline{z_\pi} = \overline{-1} = -1 = z_\pi,$$

so dass $-\pi = \pi$. Für beliebiges $\alpha \in \mathbb{S}$ gilt

$$z_{\pi+\alpha} = z_\pi z_\alpha = -z_\alpha.$$

Definition. Für jedes $\alpha \in \mathbb{S}$ definieren wir $\cos\alpha$ und $\sin\alpha$ mit

$$\boxed{\cos\alpha = \operatorname{Re} z_\alpha, \qquad \sin\alpha = \operatorname{Im} z_\alpha},$$

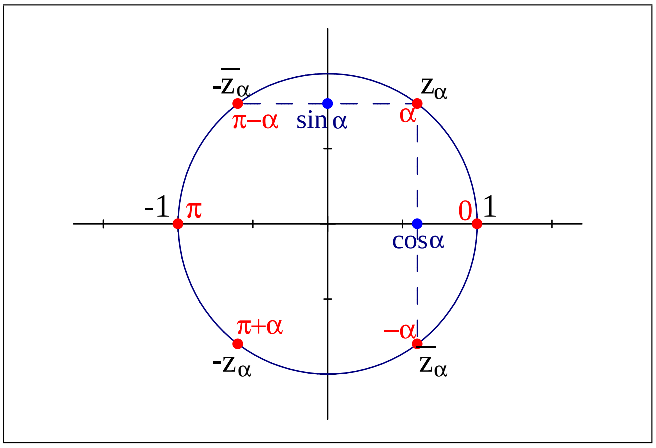
was zur folgenden Identität äquivalent ist:

$$\boxed{z_\alpha = \cos\alpha + i\sin\alpha}. \tag{3.10}$$

Die Abbildung $\cos : \mathbb{S} \to \mathbb{R}$ heißt *Kosinusfunktion* und $\sin : \mathbb{S} \to \mathbb{R}$ – *Sinusfunktion*. Man definiert auch die *Tangensfunktion*

$$\tan\alpha = \frac{\sin\alpha}{\cos\alpha}$$

für alle α mit $\cos\alpha \neq 0$. Die Funktionen cos, sin, tan heißen auch *Winkelfunktionen*.



Z.B., , da $z_0 = 1 = 1 + 0i$ und $z_\pi = -1 = -1 + 0i$, so erhalten wir

$$\cos 0 = 1, \quad \sin 0 = 0, \quad \cos \pi = -1, \quad \sin \pi = 0.$$

Satz 3.6 *Die Winkelfunktionen erfüllen die folgenden Identitäten für alle* $\alpha, \beta \in \mathbb{S}$:

(a) $\cos^2 \alpha + \sin^2 \alpha = 1$

(b) $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$

(c) $\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta$

Die Identitäten (b) und (c) heißen *Additionstheoreme*. Es folgt aus (a), dass

$$-1 \leq \cos \alpha \leq 1, \quad -1 \leq \sin \alpha \leq 1.$$

Beweis. Es folgt aus (3.10), dass

$$\cos^2 \alpha + \sin^2 \alpha = |z_\alpha|^2 = 1,$$

was (a) beweist. Nach (3.8) gilt

$$\begin{aligned} \cos(\alpha + \beta) + i \sin(\alpha + \beta) &= (\cos \alpha + i \sin \alpha)(\cos \beta + i \sin \beta) \\ &= (\cos \alpha \cos \beta - \sin \alpha \sin \beta) + i (\sin \alpha \cos \beta + \cos \alpha \sin \beta), \end{aligned}$$

woraus (b) und (c) folgen, ■

Für jedes $\alpha \in \mathbb{S}$ und $n \in \mathbb{N}$ definieren wir den Winkel $n\alpha \in \mathbb{S}$ per Induktion nach $n \in \mathbb{N}$ wie folgt:

$$1\alpha := \alpha, \qquad (n+1)\alpha := n\alpha + \alpha.$$

Insbesondere gilt:

$$2\alpha = \alpha + \alpha, \quad 3\alpha = 2\alpha + \alpha, \quad \text{usw.}$$

Es folgt, dass[2]

$$z_{n\alpha} = (z_\alpha)^n.$$

[2]Die Potenzen z^n einer komplexen Zahl z werden genau so induktiv definiert wie für reelle Zahlen:

$$z^1 = z \text{ und } z^{n+1} = z^n z.$$

Die Eigenschaften von Potenzen, die mit Ungleichungen nicht verbunden sind, gelten auch für komplexe Zahlen, z.B. der binomische Lehrsatz.

Nach den Eigenschaften der Potenzen (2.5), die auch für die komplexen Zahlen gelten, erhalten wir

$$(n+m)\,\alpha = n\alpha + m\alpha \quad \text{und} \quad n\,(m\alpha) = (nm)\,\alpha. \tag{3.11}$$

Tatsächlich haben wir

$$z_{(n+m)\alpha} = (z_\alpha)^{n+m} = (z_\alpha)^n\,(z_\alpha)^m = z_{n\alpha} z_{m\alpha} = z_{n\alpha+m\alpha}$$

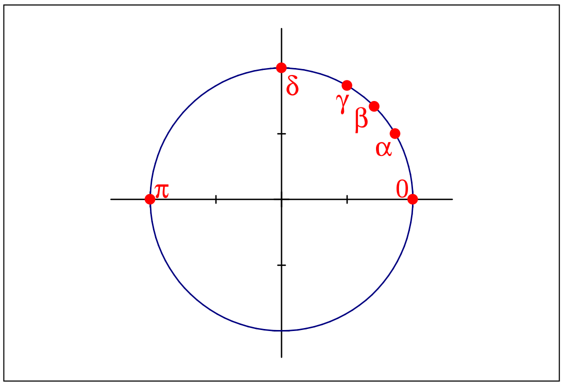
und

$$z_{n(m\alpha)} = (z_{m\alpha})^n = ((z_\alpha)^m)^n = z_\alpha^{nm} = z_{(nm)\alpha},$$

was (3.11) beweist.

Beispiel. Betrachten wir die Winkel $\alpha, \beta, \gamma, \delta$, die wie folgt definiert werden:

- $z_\alpha = \frac{\sqrt{3}}{2} + \frac{1}{2}i$
- $z_\beta = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i$
- $z_\gamma = \frac{1}{2} + \frac{\sqrt{3}}{2}i$
- $z_\delta = i$



Wir haben

$$z_{2\delta} = (z_\delta)^2 = i^2 = -1 = z_\pi,$$

so dass

$$2\delta = \pi.$$

Deshalb wird der Winkel δ auch $\frac{\pi}{2}$ genannt. Es folgt, dass $z_{\frac{\pi}{2}} = i$ und somit

$$\cos\tfrac{\pi}{2} = 0, \quad \sin\tfrac{\pi}{2} = 1.$$

Es gilt

$$z_{2\beta} = (z_\beta)^2 = \left(\tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}}i\right)^2 = i = z_\delta$$

so dass

$$2\beta = \delta \quad \text{und} \quad 4\beta = \pi.$$

Der Winkel β wird $\frac{\pi}{4}$ genannt. Es folgt, dass

$$z_{\frac{\pi}{4}} = \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}}i$$

und somit

$$\cos\tfrac{\pi}{4} = \sin\tfrac{\pi}{4} = \frac{1}{\sqrt{2}}.$$

Es gilt

$$z_{3\gamma} = (z_\gamma)^3 = \left(\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i\right)^3 = -1 = z_\pi$$

so dass
$$3\gamma = \pi.$$
Der Winkel γ wird $\frac{\pi}{3}$ genannt. Es folgt, dass
$$z_{\frac{\pi}{3}} = \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i$$
und somit
$$\cos\tfrac{\pi}{3} = \frac{1}{2} \quad \text{und} \quad \sin\tfrac{\pi}{3} = \tfrac{\sqrt{3}}{2}.$$
Es gilt
$$z_{2\alpha} = (z_\alpha)^2 = \left(\tfrac{\sqrt{3}}{2} + \tfrac{1}{2}i\right)^2 = \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i = z_\gamma,$$
so dass
$$2\alpha = \gamma \ \text{ und } 6\alpha = \pi.$$
Der Winkel α wird $\frac{\pi}{6}$ genannt. Es folgt, dass
$$z_{\frac{\pi}{6}} = \tfrac{\sqrt{3}}{2} + \tfrac{1}{2}i$$
und
$$\cos\tfrac{\pi}{6} = \tfrac{\sqrt{3}}{2} \quad \text{und} \quad \sin\tfrac{\pi}{6} = \tfrac{1}{2}.$$

Beispiel. Betrachten wir den Winkel ε mit
$$z_\varepsilon = \frac{\sqrt{5}+1}{4} + \frac{i}{4}\sqrt{10 - 2\sqrt{5}}.$$
Man kann zeigen, dass diese Zahl den Betrag 1 hat, so dass der Winkel ε wohldefiniert ist, und dass
$$z_{5\varepsilon} = (z_\varepsilon)^5 = -1 = z_\pi$$
(siehe Aufgabe 66) so dass $5\varepsilon = \pi$. Der Winkel ε wird $\frac{\pi}{5}$ genannt.

Es gibt das Verfahren von *Halbieren* des Winkels: für jedes $\alpha \in \mathbb{S} \setminus \{0\}$ existiert genau ein $\beta \in \mathbb{S}$ mit $2\beta = \alpha$ und $\sin\beta > 0$. Der Winkel β wird $\frac{\alpha}{2}$ genannt. Per Induktion erhält man den Winkel $\frac{\alpha}{2^n}$ für jedes $n \in \mathbb{N}$. Z.B. so erhält man die Winkel $\frac{\pi}{2}$ und $\frac{\pi}{4}$ aus π und $\frac{\pi}{6}$ aus $\frac{\pi}{3}$.

Polarkoordinaten

Definition. Für jedes $z \in \mathbb{C} \setminus \{0\}$ definieren wir $\arg z$ als der Winkel α mit
$$z_\alpha = \frac{z}{|z|}.$$
Der Wert $\arg z$ heißt das *Argument* von z oder der *Polarwinkel* von z.

Da $\left|\frac{z}{|z|}\right| = \frac{|z|}{|z|} = 1$, so ist der Winkel α wohldefiniert. Man kann arg betrachten als eine Abbildung $\arg : \mathbb{C} \setminus \{0\} \to \mathbb{S}$. Geometrisch bestimmt $\arg z$ die Richtung von 0 nach z durch ein Element von $\mathbb{S}$.

Satz 3.7 *Für alle $z, w \in \mathbb{C} \setminus \{0\}$ gilt*

$$\arg(zw) = \arg z + \arg w \tag{3.12}$$

und

$$\arg \frac{z}{w} = \arg z - \arg w. \tag{3.13}$$

Beweis. Sei $\alpha = \arg z$, $\beta = \arg w$ und $\gamma = \arg(zw)$. Dann gilt

$$z_\alpha = \frac{z}{|z|} \quad \text{und} \quad z_\beta = \frac{w}{|w|},$$

woraus folgt

$$z_{\alpha+\beta} = z_\alpha z_\beta = \frac{z}{|z|}\frac{w}{|w|} = \frac{zw}{|zw|} = z_\gamma$$

und somit

$$\alpha + \beta = \gamma,$$

was äquivalent zu (3.12). Man erhält (3.13) aus (3.12) wie folgt. Es gilt nach (3.12)

$$\arg \frac{z}{w} + \arg w = \arg\left(\frac{z}{w}w\right) = \arg z.$$

Addieren zu den beiden Seiten von $-\arg w$ ergibt (3.13). ■

Beispiel. Sei a eine reelle Zahl, $a \neq 0$. Im Fall $a > 0$ gilt $\arg a = 0$, da $\frac{a}{|a|} = 1$ und $z_0 = 1$. Im Fall $a < 0$ gilt $\arg a = \pi$, da $\frac{a}{|a|} = -1$ und $z_\pi = -1$. Es folgt aus (3.12), dass

$$\arg(az) = \arg z \quad \text{für } a > 0$$

und

$$\arg(az) = \arg z + \pi \quad \text{für } a < 0.$$

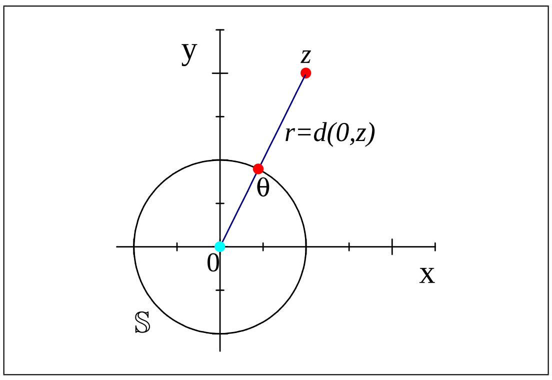
Da $\arg i = \frac{\pi}{2}$, so gilt

$$\arg(iz) = \arg z + \frac{\pi}{2}.$$

Bezeichnen wir $\mathbb{R}_+ = (0, \infty)$. Zu jedem $z \in \mathbb{C} \setminus \{0\}$ entspricht ein Paar (r, θ) mit

$$r := |z| \in \mathbb{R}_+ \quad \text{und} \quad \theta := \arg z \in \mathbb{S}.$$

Die Elemente des Paars (r, θ) heißen *Polarkoordinaten* von z, wobei r *Polarradius* ist und θ – Polarwinkel. Offensichtlich r ist gleich der Abstand zwischen 0 und z, und θ zeigt die *Richtung* von 0 nach z.



Da

$$\frac{z}{|z|} = z_\theta = \cos\theta + i\sin\theta,$$

so erhalten wir die folgende Beziehung zwischen Polarkoordinaten und kartesischen Koordinaten:

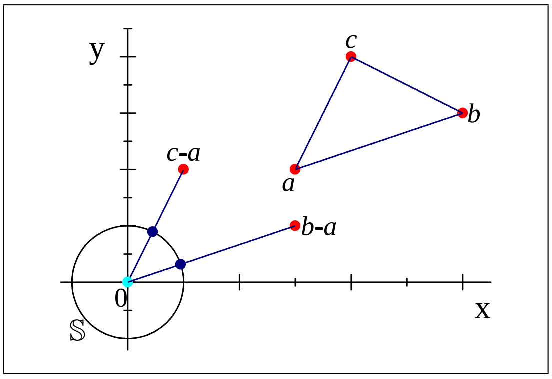
$$z = r(\cos\theta + i\sin\theta) = r\cos\theta + i(r\sin\theta). \tag{3.14}$$

Umgekehrt, für jedes $r \in \mathbb{R}_+$ und jedes $\theta \in \mathbb{S}$ hat die komplexe Zahl (3.14) die Polarkoordinaten (r, θ).

Winkel im Dreieck

Definition. Ein Dreieck Δabc ist eine Folge dreier unterschiedlichen komplexen Zahlen a, b, c. Im Dreieck Δabc definieren wir den Winkel $\angle cab$ an der Ecke a mit

$$\angle cab = \arg(c-a) - \arg(b-a).$$



Analog definiert man den Winkel an der Ecke b

$$\angle abc = \arg(a-b) - \arg(c-b)$$

und den Winkel an der Ecke c:

$$\angle bca = \arg(b-c) - \arg(a-c).$$

Bemerken wir, dass

$$\angle bac = \arg(b-a) - \arg(c-a) = -\angle cab,$$

was bedeutet, dass die Winkel im Dreieck *orientiert* sind, d.h. abhängig von der Anordnung der Ecken a, b, c.

Bemerken wir auch, dass die in der Definition benutzten Folgen cab, abc und bca die *zyklischen* Permutationen von abc sind, d.h. die Teilfolgen von $abcabc$ wie folgt: $ab\underline{cab}c$, $\underline{abc}abc$, $a\underline{bca}bc$.

Satz 3.8 *Im beliebigen Dreieck Δabc bezeichnen wir die Winkel wie folgt*

$$\alpha = \angle cab, \quad \beta = \angle abc, \quad \gamma = \angle bca.$$

Dann gilt

$$\alpha + \beta + \gamma = \pi.$$

Beweis. Nach dem Satz 3.7 haben wir

$$\alpha = \arg \frac{c-a}{b-a}, \quad \beta = \arg \frac{a-b}{c-b}, \quad \gamma = \arg \frac{b-c}{a-c}$$

und

$$\alpha + \beta + \gamma = \arg \left(\frac{c-a}{b-a} \frac{a-b}{c-b} \frac{b-c}{a-c} \right).$$

Bemerken wir:

$$\frac{c-a}{b-a} \frac{a-b}{c-b} \frac{b-c}{a-c} = (-1)^3 = -1,$$

woraus folgt, dass

$$\alpha + \beta + \gamma = \arg(-1) = \pi.$$

■

Lemma 3.9 *Für alle komplexe $z, w \neq 0$ gelten die Identitäten*

$$\operatorname{Re}(z\overline{w}) = \operatorname{Re}(\overline{z}w) = |z|\,|w| \cos\alpha \tag{3.15}$$

und

$$-\operatorname{Im}(z\overline{w}) = \operatorname{Im}(\overline{z}w) = |z|\,|w| \sin\alpha, \tag{3.16}$$

wobei

$$\alpha = \angle w0z = \arg w - \arg z. \tag{3.17}$$

Beweis. Da

$$\overline{z\overline{w}} = \overline{z}w, \tag{3.18}$$

so gilt die erste Identität in (3.15):

$$\operatorname{Re}(z\overline{w}) = \operatorname{Re}(\overline{z}w).$$

Wir haben

$$\alpha = \arg \frac{w}{z}$$

und somit

$$\cos\alpha = \operatorname{Re} \frac{\frac{w}{z}}{\left|\frac{w}{z}\right|} = \operatorname{Re}\left(\frac{w}{z} \frac{|z|}{|w|} \right) = \operatorname{Re}\left(\frac{\overline{z}w}{\overline{z}z} \frac{|z|}{|w|} \right) = \operatorname{Re}\left(\frac{\overline{z}w}{|z|^2} \frac{|z|}{|w|} \right) = \frac{1}{|w|\,|z|} \operatorname{Re}(\overline{z}w), \tag{3.19}$$

woraus (3.15) folgt.

Die erste Identität in (3.16) folgt aus (3.18), die zweite Identität beweist man genau so wie im (3.19), mit Im anstatt Re. ■

Der Wert $|z|\,|w| \cos\alpha$ heißt das *Skalarprodukt* von z und w, und $|z|\,|w| \sin\alpha$ heißt das *Kreuzprodukt* von z, w.

Beispiel. Für $z = 1 - 2i$ und $w = 6 + 3i$ haben wir nach (3.15)

$$\cos\alpha = \frac{\operatorname{Re}(\overline{z}w)}{|z|\,|w|} = \frac{\operatorname{Re}(1+2i)(6+3i)}{|z|\,|w|} = 0$$

und
$$\sin\alpha = \frac{\mathrm{Im}\,(\overline{z}w)}{|z|\,|w|} = \frac{\mathrm{Im}\,(1+2i)\,(6+3i)}{\sqrt{1^2+2^2}\sqrt{6^2+3^2}} = \frac{15}{\sqrt{5}\sqrt{45}} = \frac{15}{\sqrt{225}} = 1.$$
Somit erhalten wir
$$\alpha = \arg i = \frac{\pi}{2}.$$
Es gilt auch
$$d\,(z,w) = |z-w| = |-5-5i| = 5\sqrt{2}.$$

Satz 3.10 (Kosinussatz) *Im beliebigen Dreieck Δabc setzen wir*
$$\mathbf{a} = d\,(b,c)\,, \quad \mathbf{b} = d\,(a,c)\,, \quad \mathbf{c} = d\,(a,b)$$
und
$$\alpha = \angle cab.$$
Dann gilt
$$\mathbf{a}^2 = \mathbf{b}^2 + \mathbf{c}^2 - 2\mathbf{bc}\cos\alpha \tag{3.20}$$

Beweis. Setzen wir $z = b-a$ und $w = c-a$. Dann gilt
$$\begin{aligned} \mathbf{b} &= |c-a| = |w|\,, \\ \mathbf{c} &= |b-a| = |z| \end{aligned}$$
und
$$\mathbf{a} = |b-c| = |(b-a)-(c-a)| = |z-w|\,.$$
Auch haben wir
$$\alpha = \arg\,(c-a) - \arg\,(b-a) = \arg w - \arg z = \angle w0z. \tag{3.21}$$
Somit ist (3.20) äquivalent zu
$$|z-w|^2 = |z|^2 + |w|^2 - 2\,|z|\,|w|\cos\alpha$$
mit α wie im (3.17).

Mit Hilfe von (3.4) und (3.15) erhalten wir
$$\begin{aligned} |z-w|^2 &= |z|^2 - 2\,\mathrm{Re}\,(z\overline{w}) + |w|^2 \\ &= |z|^2 + |w|^2 - 2\,|z|\,|w|\cos\alpha, \end{aligned}$$
was zu beweisen war. ∎

Lemma 3.11 *Für alle komplexe Zahlen z_1, z_2, z_3 mit*
$$z_1 + z_2 + z_3 = 0$$
gilt
$$\mathrm{Im}\,(\overline{z_1}z_2) = \mathrm{Im}\,(\overline{z_2}z_3) = \mathrm{Im}\,(\overline{z_3}z_1)\,. \tag{3.22}$$

Beweis. Es reicht die erste Identität in (3.22) zu beweisen. Da $z_3 = -(z_1 + z_2)$, so erhalten wir

$$\operatorname{Im}(\overline{z_2} z_3) = -\operatorname{Im}(\overline{z_2}(z_1 + z_2)) = -\operatorname{Im}(\overline{z_2} z_1) - \operatorname{Im}(\overline{z_2} z_2).$$

Da $\overline{z_2} z_2$ reell ist, so gilt $\operatorname{Im}(\overline{z_2} z_2) = 0$. Nach (3.16) gilt

$$-\operatorname{Im}(\overline{z_2} z_1) = \operatorname{Im}(\overline{z_1} z_2),$$

woraus die Identität $\operatorname{Im}(\overline{z_1} z_2) = \operatorname{Im}(\overline{z_2} z_3)$ folgt. ■

Satz 3.12 (Sinussatz) *Im beliebigen Dreieck Δabc mit den Seiten*

$$\mathbf{a} = d(b, c), \quad \mathbf{b} = d(a, c), \quad \mathbf{c} = d(a, b)$$

und Winkeln

$$\alpha = \angle cab, \quad \beta = \angle abc, \quad \gamma = \angle bca$$

gilt

$$\frac{\sin \alpha}{\mathbf{a}} = \frac{\sin \beta}{\mathbf{b}} = \frac{\sin \gamma}{\mathbf{c}}. \tag{3.23}$$

Beweis. Setzen wir

$$z_1 = c - b, \quad z_2 = a - c, \quad z_3 = b - a.$$

Nach (3.21) haben wir

$$\alpha = \arg(c - a) - \arg(b - a) = \arg(-z_2) - \arg z_3 = \angle(-z_2)\, 0 z_3$$

und nach Lemma 3.9

$$\sin \alpha = -\frac{\operatorname{Im} \overline{-z_2} z_3}{|z_2|\,|z_3|} = \frac{\operatorname{Im} \overline{z_2} z_3}{\mathbf{bc}},$$

woraus folgt

$$\frac{\sin \alpha}{\mathbf{a}} = \frac{\operatorname{Im} \overline{z_2} z_3}{\mathbf{abc}}.$$

Analog beweist man dass

$$\frac{\sin \beta}{\mathbf{b}} = \frac{\operatorname{Im} \overline{z_3} z_1}{\mathbf{abc}}$$

und

$$\frac{\sin \gamma}{\mathbf{c}} = \frac{\operatorname{Im}(\overline{z_1} z_2)}{\mathbf{abc}}.$$

Da $z_1 + z_2 + z_3 = 0$, so gilt es nach Lemma 3.11

$$\operatorname{Im}(\overline{z_1} z_2) = \operatorname{Im}(\overline{z_2} z_3) = \operatorname{Im}(\overline{z_3} z_1),$$

woraus (3.23) folgt. ■

Transformationen der Ebene

Eine *Transformation* der Ebene ist eine bijektive Selbstabbildung von $\mathbb{C}$. In diesem Abschnitt definieren wir spezielle Transformationen: Rotation und Translation.

Definition. Sei w eine komplexe Zahl. Eine *Translation* der Ebene um w (genannt auch *Parallelverschiebung*) ist die Abbildung

$$\begin{aligned} T_w &: \mathbb{C} \to \mathbb{C} \\ T_w(z) &= z + w. \end{aligned}$$

Wenn man die komplexe Zahl w für Translation T_w benutzt, so wird w auch *Vektor* (oder *Verschiebungsvektor*) genannt. Offensichtlich gilt für alle $u, v \in \mathbb{Z}$ die Identität

$$T_u \circ T_v = T_{u+v}.$$

Die Translation T_w hat die inverse Abbildung $(T_w)^{-1} = T_{-w}$ da

$$T_w \circ T_{-w} = T_{-w} \circ T_w = T_0 = \mathrm{Id}\,.$$

Insbesondere ist T_w bijektiv.

Definition. Für jeden Winkel $\alpha \in \mathbb{S}$ definieren wir die *Rotation* (*Drehung*) R_α der Ebene als die folgende Abbildung:

$$\begin{aligned} R_\alpha &: \mathbb{C} \to \mathbb{C} \\ R_\alpha(z) &= z_\alpha z, \end{aligned}$$

wobei $z_\alpha = \cos\alpha + i\sin\alpha$.

Offensichtlich gilt für alle $\alpha, \beta \in \mathbb{S}$ die Identität

$$R_\alpha \circ R_\beta = R_{\alpha+\beta}.$$

Die Rotation R_α hat die inverse Abbildung $(R_\alpha)^{-1} = R_{-\alpha}$ da

$$R_\alpha \circ R_{-\alpha} = R_{-\alpha} \circ R_\alpha = R_0 = \mathrm{Id}\,.$$

Für $z = x + iy$ berechnen wir $R_\alpha(z)$ explizit wie folgt:

$$\begin{aligned} R_\alpha(z) &= (\cos\alpha + i\sin\alpha)(x + iy) \\ &= (x\cos\alpha - y\sin\alpha) + i(x\sin\alpha + y\cos\alpha)\,. \end{aligned}$$

Z.B., für $\alpha = \frac{\pi}{2}$ erhalten wir

$$R_{\frac{\pi}{2}}(z) = (-y, x)$$

und für $\alpha = \frac{\pi}{4}$ gilt

$$R_{\frac{\pi}{4}}(z) = \frac{x-y}{\sqrt{2}} + i\frac{x+y}{\sqrt{2}}.$$

Rotation R_α und Translation T_w von der Ebene $\mathbb{R}^2 = \mathbb{C}$ haben die folgende Eigenschaft: unter jeder von diesen Abbildungen bleiben die Abstände zwischen den Punkten in $\mathbb{R}^2$ und die Winkel in den Dreiecken erhalten. Jede Transformation von $\mathbb{R}^2$ mit diesen Eigenschaften heißt *Bewegung*. Man kann beweisen, dass jede Bewegung von $\mathbb{R}^2$ eine Komposition $T_w \circ R_\alpha$ ist. Die Komposition von zwei Bewegungen ist wieder eine Bewegung, und die inverse Abbildung von Bewegung ist auch Bewegung. Somit ist die Menge von allen Bewegungen von $\mathbb{R}^2$ eine (nicht kommutative) Gruppe bezüglich Komposition.

Chapter 4

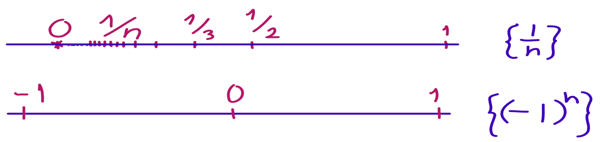
Folgen und ihre Grenzwerte

4.1 Der Begriff des Limes

Definition. 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 als *Glieder* (oder Folgenglieder) der Folge x genannt.

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, die Glieder der Folge $\left\{\frac{1}{n}\right\}$ werden immer kleiner und nähern 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 hat.



Hingegen hat die Folge $\{(-1)^n\}$ keinen Grenzwert, da sie keiner Zahl annähert und zwischen 1 und -1 springt.

Definition. Sei $\{x_n\}_{n=1}^{\infty}$ eine Folge von reellen Zahlen. Eine Zahl $a \in \mathbb{R}$ heißt der *Grenzwert* der Folge $\{x_n\}$ genau dann, wenn die folgende Eigenschaft erfüllt ist:

$$\text{für jedes } \varepsilon > 0 \ \text{ existiert ein } N \in \mathbb{N} \ \text{ so dass für alle } n \geq N \text{ gilt } \ |x_n - a| < \varepsilon.$$

Oder, mit Hilfe von den Quantoren, schreibt man:

$$\boxed{\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N \qquad |x_n - a| < \varepsilon.} \tag{4.1}$$

Hat 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:

$$x_n \to a \quad (x_n \text{ konvergiert gegen } a)$$

$$x_n \to a \quad \text{für} \quad n \to \infty \qquad (a_n \text{ konvergiert gegen } a \text{ für } n \text{ gegen unendlich}).$$

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

$$a = \lim x_n \quad \text{oder } a = \lim_{n\to\infty} x_n.$$

Jetzt besprechen wir einige Umformulierungen der Bedingung (4.1).

Definition. Man sagt, dass eine von $n \in \mathbb{N}$ abhängige Aussage $A(n)$ für *fast alle* n gilt, falls $A(n)$ für alle $n \in \mathbb{N} \setminus S$ gilt, wobei $S \subset \mathbb{N}$ eine endliche Menge ist.

In anderen Wörter, "fast alle" bedeutet: alle außer einer endlichen Menge.

Behauptung. *Die Bedingung* $a = \lim x_n$ *ist äquivalent zu:*

$$\boxed{\forall \varepsilon > 0 \quad \textit{gilt} \ \ |x_n - a| < \varepsilon \ \textit{für fast alle } n.} \tag{4.2}$$

Beweis. Gilt $a = \lim x_n$, so erhalten wir nach (4.1): für jedes $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ so dass

$$|x_n - a| < \varepsilon \text{ gilt für alle } n \geq N,$$

d.h.

$$|x_n - a| < \varepsilon \text{ gilt für alle } n \in \mathbb{N} \setminus \mathcal{E}_{N-1}.$$

Somit gilt auch (4.2). Umgekehrt, gilt (4.2), so für jedes ε existiert eine endliche Menge $S \subset \mathbb{N}$ so dass

$$|x_n - a| < \varepsilon \text{ gilt für alle } n \in \mathbb{N} \setminus S.$$

Da S endlich ist, so existiert $\max S$ (Aufgabe 51). Setzen wir $N = \max S + 1$. Dann

$$n \geq N \Rightarrow n \in \mathbb{N} \setminus S \Rightarrow |x_n - a| < \varepsilon,$$

woraus folgt $a = \lim x_n$. ■

Die Bedingung $|x_n - a| < \varepsilon$ ist offensichtlich äquivalent zu $x_n \in (a - \varepsilon, a + \varepsilon)$.

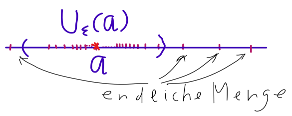
Definition. Das Intervall $(a - \varepsilon, a + \varepsilon)$ mit $\varepsilon > 0$ heißt die ε-*Umgebung* von a und wird mit $U_\varepsilon(a)$ bezeichnet.

Somit erhalten wir noch eine Umformulierung von dem Begriff des Limes.

Behauptung. *Die Bedingung* $a = \lim x_n$ *ist äquivalent zu:*

$$\boxed{\forall \varepsilon > 0 \ \ \textit{gilt } x_n \in U_\varepsilon(a) \ \ \textit{für fast alle } n,} \tag{4.3}$$

d.h. jede Umgebung von a *enthält fast alle Folgenglieder* x_n.



Hier ist noch eine äquivalente Umformulierung von (4.3): $a = \lim x_n$ genau dann, wenn

$$\forall\varepsilon > 0 \text{ ist die Menge } \{n : x_n \notin U_\varepsilon(a)\} \text{ endlich.}$$

Daraus erhalten wir die Negation[1] von $a = \lim x_n$.

Behauptung. *Eine Zahl a ist kein Grenzwert von $\{x_n\}$ genau dann, wenn*

$$\exists\varepsilon > 0 \quad \textit{so dass die Menge } \{n : x_n \notin U_\varepsilon(a)\} \textit{ unendlich ist.} \tag{4.4}$$

Jetzt beweisen wir, dass $\lim x_n$ nicht zwei Werte annehmen kann.

Behauptung. *Existiert der Grenzwert* $\lim x_n$ *so ist* $\lim x_n$ *eindeutig bestimmt.*

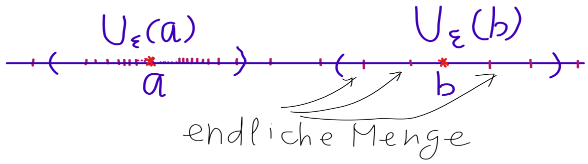
Beweis. Nehmen wir das Gegenteil an, dass es zwei Werte von $\lim x_n$ gibt, d.h. $x_n \to a$ und $x_n \to b$ mit $a \neq b$. Sei $a < b$. Wählen wir ein $\varepsilon > 0$ so dass die Intervalle $(a - \varepsilon, a + \varepsilon)$ und $(b - \varepsilon, b + \varepsilon)$ disjunkt sind. Dafür reicht es zu sichern, dass

$$a + \varepsilon < b - \varepsilon,$$

d.h.

$$\varepsilon < \frac{b - a}{2}.$$

Da $b - a > 0$, so solches ε existiert. Da $a = \lim x_n$, so enthält $U_\varepsilon(a)$ fast alle x_n.



Daraus folgt, dass das Intervall $U_\varepsilon(b)$ nur eine endliche Menge von x_n enthalten kann, was in Widerspruch zum $b = \lim x_n$ steht. ∎

25.05.18 H1

Beispiel. 1. Zeigen wir, dass die Folge $x_n = \frac{1}{n}$ gegen 0 konvergiert. Nach dem Archimedischen Prinzip, für jedes $\varepsilon > 0$ existiert $N \in \mathbb{N}$ mit $N > \frac{1}{\varepsilon}$. Da $\frac{1}{N} < \varepsilon$, so erhalten wir

$$n \geq N \Rightarrow \frac{1}{n} < \varepsilon \Rightarrow \left|\frac{1}{n} - 0\right| < \varepsilon,$$

woraus folgt

$$\frac{1}{n} \to 0 \text{ für } n \to \infty.$$

Ebenso beweist man, dass für jedes $c \in \mathbb{R}$

$$\frac{c}{n} \to 0 \text{ für } n \to \infty.$$

[1] Seien α, β zwei Parameter (Elemente einer beliebigen Menge) und $A = A(\alpha, \beta)$ eine von α, β abhängige Aussage. Dann gilt folgendes:

$$\neg(\forall\alpha\ \exists\beta \text{ s.d. } A(\alpha,\beta) \text{ gilt}) \Leftrightarrow \exists\alpha\ \forall\beta \text{ s.d. } \neg A(\alpha,\beta) \text{ gilt.}$$

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 viel 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.

Bemerkung. Die Konvergenz $x_n \to a$ lässt sich betrachten als eine Annäherung von a mit den Folgenglieder x_n, wobei $\varepsilon > 0$ ein vorgegebener Approximationsfehler ist. Dann bedeutet die Bedingung (4.1), 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.

Bemerkung. Die Definition (4.1) 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.

4.2 Eigenschaften des Limes

Für Beweise von Eigenschaften des Limes benutzen wir die folgende Eigenschaft des Begriffes "für 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

$$n \in S^c \cap T^c = (S \cup T)^c .$$

Nach dem Satz 2.18 ist die Vereinigung $S \cup T$ endlich, woraus folgt, dass $A(n) \wedge B(n)$ für fast alle n gilt. ■

Die weiteren Eigenschaften des Grenzwertes werden in dem folgenden Satz angegeben.

Satz 4.1 *(a) Seien $\{x_n\}$ und $\{y_n\}$ zwei konvergente Folgen mit $x_n \leq y_n$ für fast alle n. Dann gilt*

$$\lim x_n \leq \lim y_n.$$

Insbesondere, gilt $x_n = y_n$ für fast alle n, so gilt auch $\lim x_n = \lim y_n$.

(*b*) *Seien* $\{x_n\}, \{y_n\}, \{z_n\}$ *drei Folgen mit* $x_n \le y_n \le z_n$ *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 beschränkt (d.h.* $|x_n| \le c$ *für eine Zahl* c *und für alle* $n \in \mathbb{N}$*).*

(*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:

$$a - \varepsilon < x_n \le y_n < b + \varepsilon,$$

woraus folgt, dass $a < b + 2\varepsilon$ und

$$a - b < 2\varepsilon. \tag{4.5}$$

Da $\varepsilon > 0$ beliebig ist, so kann die Zahl $a - b$ positiv nicht sein (sonst gilt (4.5) mit $\varepsilon = \frac{1}{2}(a - b)$ nicht), woraus folgt $a - b \le 0$ und somit $a \le b$.

(b) Für jedes $\varepsilon > 0$ gelten die folgenden Bedingungen für fast alle n:

$$a - \varepsilon < x_n \le y_n \le z_n < a + \varepsilon,$$

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 viel Glieder, so dass die Menge davon auch beschränkt ist (jede endliche Teilmenge von $\mathbb{R}$ hat Maximum und Minimum nach Aufgabe 51). 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 $\lim x_n \le \lim y_n$ 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 (Aufgabe 32) haben wir

$$a^n = (1 + c)^n \ge 1 + nc.$$

Da die Folge $\{nc\}_{n=1}^{\infty}$ nach dem Archimedischen Prinzip 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 4.1(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 $-1 < a \leq 1$.

4.3 Rechenregeln

Hier beweisen wir einige Rechenregeln für den Grenzwert.

Satz 4.2 *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:

$$\begin{aligned}
\lim (x_n + y_n) &= \lim x_n + \lim y_n \\
\lim (x_n - y_n) &= \lim x_n - \lim y_n \\
\lim (x_n y_n) &= \lim x_n \lim y_n \\
\lim \frac{x_n}{y_n} &= \frac{\lim x_n}{\lim y_n}
\end{aligned}$$

vorausgesetzt, dass die rechten Seiten sinnvoll sind. Für eine konstante Folge $y_n = c$ erhalten wir, dass

$$\lim (x_n + c) = \lim x_n + c \quad \text{und} \quad \lim (cx_n) = c \lim x_n.$$

Beweis von Satz 4.2. Für jedes $\varepsilon > 0$ gelten

$$|x_n - a| < \varepsilon \text{ und } \ |y_n - b| < \varepsilon \tag{4.6}$$

für fast alle $n \in \mathbb{N}$. Daraus folgt, dass für fast alle n gilt

$$|x_n + y_n - (a+b)| = |x_n - a + y_n - b| \leq |x_n - a| + |y_n - b| < 2\varepsilon =: \varepsilon'. \tag{4.7}$$

Jetzt können wir $\varepsilon' > 0$ beliebig angeben, setzen $\varepsilon = \varepsilon'/2 > 0$ so dass (4.7) für fast alle n gilt, und beschließen, dass für fast alle n

$$|x_n + y_n - (a+b)| < \varepsilon',$$

woraus folgt

$$x_n + y_n \to a + b.$$

Die Konvergenz $x_n - y_n \to a - b$ beweist man analog.

Um $x_n y_n \to ab$ zu beweisen, bemerken wir zuerst dass nach Satz 4.1 die Folge $\{x_n\}$ beschränkt ist, d.h. für eine Zahl $c > 0$ gilt

$$|x_n| \leq c \quad \forall n \in \mathbb{N}.$$

Die Differenz $x_n y_n - ab$ lässt sich für fast alle n wie folgt abschätzen:

$$\begin{aligned} |x_n y_n - ab| &= |x_n y_n - x_n b + x_n b - ab| \\ &\leq |x_n (y_n - b)| + |(x_n - a)\, b| \\ &\leq c\,|y_n - b| + |b|\,|x_n - a| \\ &< c\varepsilon + |b|\,\varepsilon = (c + |b|)\,\varepsilon. \end{aligned}$$

Da $\varepsilon' = (c + |b|)\,\varepsilon$ eine beliebige positive Zahl ist, so erhalten wir $x_n y_n \to ab$.

Für die letzte Behauptung $\frac{x_n}{y_n} \to \frac{a}{b}$ reicht es zu beweisen, dass $\frac{1}{y_n} \to \frac{1}{b}$ da danach gilt

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

Sei $b > 0$ (der Fall $b < 0$ ist analog). Da für fast alle n gilt

$$y_n \in U_{b/2}(b) = \left(\frac{b}{2}, \frac{3b}{2}\right),$$

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 $\varepsilon' = \frac{2\varepsilon}{b^2}$ eine beliebige positive Zahl ist, so folgt es, dass $\frac{1}{y_n} \to \frac{1}{b}$. ∎

Beispiel. Mit Hilfe von dem Satz 4.2 bestimmen wir den Grenzwert der Folge

$$x_n = \frac{an^2 + bn + c}{a'n^2 + b'n + c'}, \tag{4.8}$$

wobei a, b, c, a', b', c' reelle Zahlen sind mit $a' \neq 0$. 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 $\frac{1}{n^2} \to 0$, $\frac{b}{n} \to 0$, $\frac{c}{n^2} \to 0$ und somit

$$\lim\left(a + \frac{b}{n} + \frac{c}{n^2}\right) = a \quad \text{und} \quad \lim\left(a' + \frac{b'}{n} + \frac{c'}{n^2}\right) = a'.$$

Es folgt

$$\lim x_n = \frac{\lim\left(a + \frac{b}{n} + \frac{c}{n^2}\right)}{\lim\left(a' + \frac{b'}{n} + \frac{c'}{n^2}\right)} = \frac{a}{a'}.$$

4.4 Grenzwert in $\overline{\mathbb{R}}$

30.05.18 H13

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:

$$\text{jede Umgebung von } a \text{ enthält fast alle Glieder von } \{x_n\}. \tag{4.9}$$

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 überein. In diesem Fall sagt man, dass x_n gegen a konvergiert, und die Folge $\{x_n\}$ konvergent ist. Ist $a = \pm\infty$, so sagt man, dass dass x_n gegen a *divergiert,* und die Folge $\{x_n\}$ *bestimmt divergent* ist. Hat die Folge keinen Grenzwert in $\overline{\mathbb{R}}$, so sagt man, dass die Folge $\{x_n\}$ *unbestimmt divergent* ist.

Die Bedingung $x_n \to +\infty$ bedeutet:

$$\forall E \in \mathbb{R} \ \text{ gilt } \ x_n > E \text{ für fast alle } n.$$

Die Bedeutung $x_n \to -\infty$ bedeutet:

$$\forall E \in \mathbb{R} \ \text{ gilt } \ x_n < E \text{ für fast alle } n.$$

Beispiel. 1. Die Folge $x_n = n$ divergiert gegen $+\infty$, da die Bedingung $x_n > E$ für alle $n > E$ erfüllt ist und somit für fast alle n. 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. Die Folge $x_n = \sqrt{n}$ auch divergiert gegen $+\infty$, da die Bedingung $x_n > E$ bedeutet, dass $\sqrt{n} > E$, was für positives E äquivalent zu $n > E^2$ ist. Somit gilt $x_n > E$ für fast alle n.

Erweitern wir die Aussagen (a) und (b) des Satzes 4.1 auf die Folgen in $\overline{\mathbb{R}}$.

Satz 4.3 *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 $a > b$.

Behauptung. *Gilt* $a > b$ *so existieren Umgebungen* A *von* a *und* B *von* b *so dass für alle* $x \in A$ *und* $y \in B$ *gilt* $x > y$.

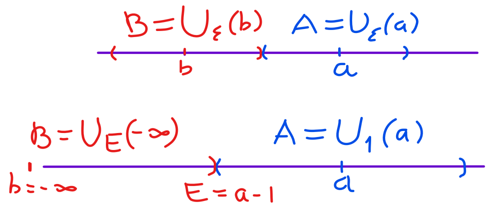
Sind a und b reell so setzen wir $\varepsilon = \frac{a-b}{2}$ und

$$A = U_\varepsilon(a), \quad B = U_\varepsilon(b).$$

Sei a reell und $b = -\infty$. Dann setzen wir

$$A = U_1(a), \quad B = U_E(-\infty)$$

mit $E = a - 1$.



Der Fall $a = +\infty$ und b reell ist analog. Im Fall $a = +\infty$ und $b = -\infty$ setzen wir

$$A = U_0(+\infty), \quad B = U_0(-\infty).$$

So, in allen Fällen haben wir die Umgebungen A von a und B von b so erstellt, dass

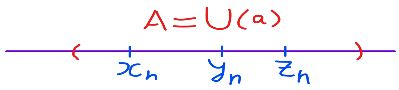
$$x \in A \text{ und } y \in B \Rightarrow x > y. \tag{4.10}$$

Andererseits, für fast alle n gelten

$$x_n \in A, \quad y_n \in B \quad \text{und} \quad x_n \le y_n,$$

was im Widerspruch zum (4.10) steht.

(b) Sei A eine beliebige Umgebung von a.



Nach Voraussetzung liegen fast all x_n und z_n in A, voraus folgt, dass auch fast alle y_n in A liegen, und somit $y_n \to a$. ■

Beispiel. 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 $cn \to +\infty$ und $cn \le a^n < +\infty$ so folgt es, 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 $+\infty$ wie folgt:

$$(+\infty) + a = a + (+\infty) = \begin{cases} +\infty, & -\infty < a \le +\infty \\ \text{unbestimmt}, & a = -\infty. \end{cases}$$

Analog definiert man Addition mit $-\infty$:

$$(-\infty) + a = a + (-\infty) = \begin{cases} -\infty, & -\infty \le a < +\infty \\ \text{unbestimmt}, & a = +\infty. \end{cases}$$

Subtraktion wird auf Addition wie folgt reduziert:

$$a - (+\infty) = a + (-\infty) \quad \text{und} \quad a - (-\infty) = a + (+\infty).$$

Multiplikation mit $+\infty$ wird wie folgt definiert:

$$(+\infty) \cdot a = a \cdot (+\infty) = \begin{cases} +\infty, & 0 < a \le +\infty \\ -\infty, & -\infty \le a < 0 \\ \text{unbestimmt}, & a = 0. \end{cases}$$

und analog mit $-\infty$.

Division durch $+\infty$ wird wie folgt definiert:

$$\frac{a}{+\infty} = \begin{cases} 0, & a \in \mathbb{R} \\ \text{unbestimmt}, & a = +\infty \text{ or } a = -\infty, \end{cases}$$

und analog durch $-\infty$.

Erinnern wir uns, dass auch $\frac{a}{0}$ unbestimmt ist. Somit bleiben die folgenden Operationen unbestimmt:

$$\infty - \infty, \quad 0 \cdot \infty, \quad \frac{\infty}{\infty}, \quad \frac{0}{0}. \tag{4.11}$$

Alle diese Ausdrücke heißen *unbestimmte* Ausdrücke.

Satz 4.4 *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*

$$\lim (x_n + y_n) = a{+}b, \quad \lim (x_n - y_n) = a{-}b, \quad \lim (x_n y_n) = ab, \qquad \lim \frac{x_n}{y_n} = \frac{a}{b}, \tag{4.12}$$

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 wurden die Identitäten (4.12) im Satz 4.2 bewiesen. Sonst muss man verschiedene Möglichkeiten für a und b betrachten und zeigen, dass in allen Fällen die Identitäten (4.12) mit den Definitionen von Operationen mit $\pm\infty$ übereinstimmen.

Betrachten wir den Fall wenn a reell ist und $b = +\infty$. Dann ist die Folge $\{x_n\}$ beschränkt, sei $|x_n| \leq c$ für ein $c > 0$. Jede Umgebung $U_E(+\infty)$ enthält fast alle y_n. Da für $y_n \in U_E(+\infty)$ gilt

$$x_n + y_n > -c + E = E',$$

so enthält $U_{E'}(+\infty)$ fast alle $x_n + y_n$. Da E' beliebig ist, so folgt es, dass

$$x_n + y_n \to +\infty = a + b.$$

Die Differenz bestimmt man analog:

$$x_n - y_n < c - E = E',$$

so dass $U_{E'}(-\infty)$ fast alle $x_n - y_n$ enthält und somit

$$x_n - y_n \to -\infty = a - b.$$

Um $\lim x_n y_n$ zu bestimmen, nehmen wir an, dass $a > 0$ (der Fall von $a < 0$ ist analog). Dann liegen fast alle x_n in $U_{a/2}(a)$ so dass für fast alle n gilt

$$x_n > \frac{a}{2}.$$

Sei E positive. Für fast alle n gilt

$$y_n > E,$$

woraus folgt

$$x_n y_n > \frac{a}{2} E = E'.$$

Da E' beliebige positive Zahl ist, so folgt es

$$x_n y_n \to +\infty = ab.$$

Bestimmen wir jetzt $\lim \frac{x_n}{y_n}$. Für beliebige positive E haben wir $y_n > E$ für fast alle n und $|x_n| < c$, woraus folgt

$$\left|\frac{x_n}{y_n}\right| = \frac{|x_n|}{y_n} < \frac{c}{E} = \varepsilon.$$

Da ε beliebige positive Zahl ist, so erhalten wir

$$\left|\frac{x_n}{y_n}\right| \to 0$$

und somit

$$\frac{x_n}{y_n} \to 0 = \frac{a}{b}.$$

Die anderen Fällen von a und b betrachten man analog. ■

Beispiel. Zeigen wir, warum die Ausdrücke (4.11) unbestimmt sollen sein.

1. 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 eine beliebige reelle Zahl ist. Für die Folge $x_n = n + (-1)^n$ gilt auch $\lim x_n = +\infty$ aber die Differenz $x_n - y_n = (-1)^n$ überhaupt keinen Grenzwert hat. Somit lässt der Ausdruck $\infty - \infty$ sich nicht eindeutig definieren.

2. Sei $x_n = \frac{c}{n}$ und $y_n = n$. Dann gilt $x_n \to 0$ und $y_n \to +\infty$, während $x_n y_n \to c$. Da c beliebig ist, so lässt $0 \cdot \infty$ sich nicht eindeutig definieren.

3. Sei $x_n = cn$ mit $c > 0$ und $y_n = n$ so dass $x_n \to +\infty$ und $y_n \to +\infty$. Dann gilt $\frac{x_n}{y_n} \to c$, so dass der Ausdruck $\frac{\infty}{\infty}$ sich nicht eindeutig definieren lässt.

4. Sei $x_n = \frac{c}{n}$ und $y_n = \frac{1}{n}$ so dass $x_n \to 0$ und $y_n \to 0$. Dann gilt $\frac{x_n}{y_n} \to c$, so dass der Ausdruck $\frac{0}{0}$ sich nicht eindeutig definieren lässt.

01.06.18 H1

Beispiel. Mit Hilfe von dem Satz 4.4 bestimmen wir den folgenden Grenzwert:

$$\lim \frac{3 + \frac{1}{\sqrt{n}}}{2^n - \frac{n+1}{n}}.$$

Da $\sqrt{n} \to +\infty$, so gilt

$$\lim \frac{1}{\sqrt{n}} = \frac{1}{+\infty} = 0$$

und

$$\lim \left(3 + \frac{1}{\sqrt{n}}\right) = 3 + 0 = 3.$$

Da $2^n \to +\infty$ und $\frac{n+1}{n} = 1 + \frac{1}{n} \to 1$, so gilt

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

Somit erhalten wir

$$\lim \frac{3 + \frac{1}{\sqrt{n}}}{2^n - \frac{n+1}{n}} = \frac{3}{+\infty} = 0.$$

Den Grenzwert

$$\lim \frac{n + \sqrt{n}}{n - \sqrt{n}} \tag{4.13}$$

kann man analog nicht bestimmen, da $n + \sqrt{n} \to +\infty$ und auch $n - \sqrt{n} \to +\infty$ (da $\sqrt{n} \le \frac{n}{2}$ für $n \ge 4$ und somit $n - \sqrt{n} \ge n/2$). In diesem Fall erhalten wir den unbestimmten Ausdruck $\frac{\infty}{\infty}$, und man muss andere Methoden einsetzen. Den Grenzwert

$$\lim \left(\sqrt{n+1} - \sqrt{n}\right) \tag{4.14}$$

ist auch ein unbestimmter Ausdruck $\infty - \infty$, so ist der Satz 4.4 in diesem Fall auch nicht benutzbar. Die Grenzwerte (4.13) und (4.14) werden in Aufgabe 72 betrachtet.

4.5 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, so gilt $x_n \geq x_m$ für alle $n \geq m$, was man per Induktion nach $n \geq m$ beweist. 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 4.5 (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. Nach dem Satz 1.11 existiert das Supremum

$$a = \sup\{x_n\} \in \overline{\mathbb{R}}$$

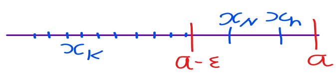
Beweisen wir, dass $x_n \to a$.

Sei zuerst a reell, d.h. die Folge $\{x_n\}$ ist beschränkt. 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

$$x_n \in (a-\varepsilon, a+\varepsilon) \text{ für alle } n \geq N,$$

woraus $x_n \to a$ folgt.

Sei $a = +\infty$, d.h. die Folge $\{x_n\}$ ist unbeschränkt. Sei E beliebige reelle Zahl. Da E keine obere Schranke von $\{x_n\}$ ist, so gibt es ein $N \in \mathbb{N}$ mit $x_N > E$. Da $\{x_n\}$ monoton steigend ist, so erhalten wir, dass

$$x_n > E \text{ für alle } n \geq N,$$

woraus $\lim x_n = +\infty = a$. folgt.

Der Fall von einer monoton fallenden Folge ist analog. ■

Beispiel. Betrachten wir die Folge $\{x_n\}_{n\in\mathbb{N}}$ die wie folgt induktiv definiert ist:

$$x_1 = 1, \quad x_{n+1} = x_n + \frac{1}{x_n^2} \text{ für alle } n \in \mathbb{N}.$$

Die ersten Glieder der Folge sind wie folgt:

$$x_2 = 2, \quad x_3 = \frac{9}{4} = 2,25, \quad x_4 = \frac{793}{324} = 2,447..., \quad x_5 = \frac{532689481}{203747076} = 2,614...$$

usw. Da offensichtlich gilt $x_{n+1} > x_n$, so ist diese Folge 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 $x = +\infty$ und somit $x_n \to +\infty$.

Beispiel. Betrachten wir die Folge

$$x_1 = 2, \quad x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right) \quad \forall n \in \mathbb{N}.$$

Die ersten Glieder der Folge sind wie folgt:

$$x_2 = \frac{5}{4} = 1,2, \quad x_3 = \frac{41}{40} = 1,025, \quad x_4 = \frac{3281}{3280} = 1,0003...$$

usw. Per Induktion beweist man, dass $x_n > 0$ für alle n. Danach bemerken wir, dass für alle $x > 0$ gilt

$$\frac{1}{2}\left(x + \frac{1}{x}\right) \geq 1,$$

da diese Ungleichung äquivalent zu $(x-1)^2 \geq 0$ ist. Daraus folgt, dass $x_n \geq 1$ für alle n und somit

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right) \leq x_n.$$

So ist die Folge $\{x_n\}$ monoton fallend und nach unten von 1 beschränkt. Es folgt: der Grenzwert $x = \lim x_n$ existiert, ist eine positive Zahl ist und erfüllt die Gleichung

$$x = \frac{1}{2}\left(x + \frac{1}{x}\right).$$

Daraus erhalten wir $x^2 = 1$ und somit $x = 1$, so dass $\lim x_n = 1$.

Beispiel. Betrachten wir den Ausdruck:

$$\sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + ...}}}}}$$

Rigoros wird der Wert der unendlichen Folge von Wurzeln als $\lim x_n$ definiert, wobei

$$x_1 = 0, \quad x_{n+1} = \sqrt{1 + x_n}.$$

In der Tat haben wir

$$x_2 = \sqrt{1}, \quad x_3 = \sqrt{1 + \sqrt{1}}, \quad x_3 = \sqrt{1 + \sqrt{1 + \sqrt{1}}},$$

usw. Man kann zeigen, dass die Folge $\{x_n\}$ monoton steigend und beschränkt ist, woraus die Konvergenz folgt (siehe Aufgabe 80). 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 Aufgabe 84). Der Grenzwert

$$e := \lim x_n = 2,718281828459045...$$

ist eine transzendente Zahl, die eine wichtige Rolle in Analysis spielt.

4.6 Intervallschachtelungsprinzip

Fixieren wir eine Grundmenge M und eine nichtleere *Indexmenge* S. Eine *Familie* $\{A_\alpha\}_{\alpha \in S}$ von Teilmengen von M mit der Indexmenge S ist eine Abbildung

$$\begin{array}{l} A : S \to \mathcal{P}(M) \\ S \ni \alpha \mapsto A_\alpha \subset M \end{array}$$

Die Vereinigung $\bigcup\limits_{\alpha \in S} A_\alpha$ und der Durchschnitt $\bigcap\limits_{\alpha \in S} A_\alpha$ der Familie $\{A_\alpha\}$ werden wie folgt definiert:

$$x \in \bigcup_{x \in S} A_\alpha \Leftrightarrow \exists \alpha \in S \text{ mit } x \in A_\alpha$$

$$x \in \bigcap_{\alpha \in S} A_\alpha \Leftrightarrow \forall \alpha \in S \text{ gilt } x \in A_\alpha.$$

Für $S = \mathbb{N}$ heißt die Familie $\{A_n\}_{n \in \mathbb{N}}$ auch Folge.

Definition. Eine Folge $\{I_k\}_{k \in \mathbb{N}}$ von Intervallen heißt *Intervallschachtelung* falls $I_{k+1} \subset I_k$ für alle $k \in \mathbb{N}$.

Satz 4.6 (Intervallschachtelungsprinzip) *Sei $\{I_k\}_{k=1}^{\infty}$ eine Intervallschachtelung wobei alle Intervalle I_k abgeschlossen und beschränkt sind. Dann ist der Durchschnitt $\bigcap_{k\in\mathbb{N}} I_k$ nichtleer.*

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. In der Tat, 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 Prinzip steht.

Auch ist die Beschränktheit der Intervalle wichtig. Zum Beispiel, die Intervallschachtelung von den unbeschränkten Intervallen $I_k = [k, +\infty)$ hat 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

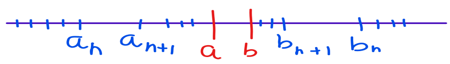
$$a_k \leq a_{k+1} \quad \text{und } b_k \geq b_{k+1}.$$

d.h. die Folge $\{a_k\}$ monoton steigend ist und $\{b_k\}$ – monoton fallend. Nach dem Satz 4.5 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 4.3 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 $[a, b] = \bigcap_{k\in\mathbb{N}} I_k$.

4.7 Überdeckungssatz

Definition. Eine Familie $\{A_\alpha\}_{\alpha\in S}$ heißt *Überdeckung* von einer Menge B falls

$$B \subset \bigcup_{\alpha\in S} A_\alpha.$$

Ist T eine Teilmenge von S, so betrachten wir auch die *Teilfamilie* $\{A_\alpha\}_{\alpha\in T}$. Ist $\{A_\alpha\}_{\alpha\in T}$ auch eine Überdeckung von B, so heißt sie *Teilüberdeckung*.

Hauptsatz 4.7 (Satz von Heine-Borel[2], Überdeckungssatz) *Seien a, b reelle Zahlen mit $a \leq b$ und sei $\{I_\alpha\}_{\alpha \in S}$ eine Überdeckung von $[a, b]$ mit offenen Intervallen I_α. Dann enthält $\{I_\alpha\}_{\alpha \in S}$ eine endliche Teilüberdeckung $\{I_\alpha\}_{\alpha \in T}$ von $[a, b]$.*

06.06.18 H13 Die Endlichkeit von $\{I_\alpha\}_{\alpha \in T}$ bedeutet, dass T eine endliche Teilmenge von S ist.

Definition. Eine Menge $M \subset \mathbb{R}$ heißt *kompakt* falls jede Überdeckung von M mit offenen Intervallen eine endliche Teilüberdeckung enthält.

Offensichtlich ist jede endliche Teilmenge von $\mathbb{R}$ kompakt. Satz 5.6 besagt, dass auch jedes beschränktes abgeschlossenes Intervall $[a, b]$ kompakt ist. Wir wissen, dass jede nichtleere endliche Teilmenge von $\mathbb{R}$ das Maximum und Minimum hat. Man kann beweisen, dass jede nichtleere kompakte Teilmenge von $\mathbb{R}$ auch das Maximum und Minimum hat (siehe Aufgabe 81).

Der Begriff von Kompaktheit ist einer von zentralen Begriffen in Analysis and Topologie. Wir vertiefen und benutzen ihn später.

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. Betrachten wir die offenen Intervalle

$$I_k = (\frac{1}{k}, 2), \quad k \in \mathbb{N}.$$

Die Familie $\{I_k\}$ ist eine Überdeckung von $(0, 1]$, aber sie 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 4.7. Nehmen wir das Gegenteil an, dass eine Überdeckung $\{I_\alpha\}_{\alpha \in S}$ keine endliche Teilüberdeckung von $[a, b]$ enthält. Dann bestimmen wir eine Intervallschachtelung $\{[a_n, b_n]\}_{n=1}^{\infty}$ mit den folgenden Eigenschaften:

- $[a_n, b_n] \subset [a, b]$ für alle $n \in \mathbb{N}$;
- jedes Intervall $[a_n, b_n]$ keine endliche Teilüberdeckung von $\{I_\alpha\}_{\alpha \in S}$ zulässt;
- $b_n - a_n \to 0$ für $n \to \infty$.

[2] Auch als "Satz von Borel-Lebesgue" genannt.

Wir definieren $[a_n, b_n]$ per Induktion nach n. Für Induktionsanfang setzen wir $[a_1, b_1] = [a, b]$.

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 zwei Intervallen $[a_n, c]$, $[c, b_n]$ keine endliche Teilüberdeckung zulässt. In der Tat, gibt es endliche Teilüberdeckungen $\{I_\alpha\}_{\alpha \in T_1}$ von $[a_n, c]$ und $\{I_\alpha\}_{\alpha \in T_2}$ von $[c, b_n]$ so ist $\{I_\alpha\}_{\alpha \in T_1 \cup T_2}$ eine endliche Teilüberdeckung von $[a_n, b_n]$, was nach der Voraussetzung nicht möglich ist.

Bezeichnen wir mit $[a_{n+1}, b_{n+1}]$ das eine der beiden Intervalle $[a_n, c]$, $[c, b_n]$, das keine endliche Teilüberdeckung von $\{I_\alpha\}_{\alpha \in S}$ zulässt.

Offensichtlich haben wir

$$[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$$

und

$$b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2}. \tag{4.15}$$

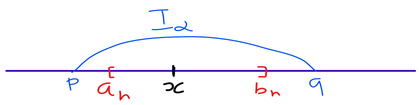
Die Folge $\{[a_n, b_n]\}_{n=1}^{\infty}$ ist somit eine Intervallschachtelung. Es folgt aus (4.15) dass

$$b_n - a_n = \frac{b - a}{2^{n-1}}$$

und somit

$$b_n - a_n \to 0 \text{ für } n \to \infty. \tag{4.16}$$

Nach dem Satz 4.6 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

$$[a_n, b_n] \subset (p, q)\,. \tag{4.17}$$

Da $a_n \le x \le b_n$, so folgt es aus (4.16), dass auch

$$|a_n - x| \to 0 \quad \text{und} \quad |b_n - x| \to 0$$

und somit

$$\lim a_n = \lim b_n = x.$$

Folglich enthält (p, q) fast alle a_n und b_n, woraus (4.17) folgt. Die Bedingung (4.17) bedeutet, dass das Intervall $[a_n, b_n]$ von *einem* Intervall I_α überdeckt ist, was im Widerspruch zur Konstruktion von $[a_n, b_n]$ steht, da $[a_n, b_n]$ keine endliche Teilüberdeckung zulässt. ∎

4.8 Teilfolgen und Satz von Bolzano-Weierstraß

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 $x \circ n$:

$$x_{n_k} = x\left(n\left(k\right)\right) = \left(x \circ n\right)\left(k\right).$$

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. In der Tat, nach Definition von der Konvergenz $x_n \to a$, jede Umgebung $U\left(a\right)$ enthält fast alle Glieder x_n, woraus folgt, dass auch fast alle x_{n_k} in $U\left(a\right)$ 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. Die Umkehrung davon gilt nicht: ein Häufungspunkt soll nicht unbedingt der Grenzwert sein.

Beispiel. Die Folge $x_n = \left(-1\right)^n$ divergiert, obwohl diese Folge zwei Häufungspunkte hat: 1 und -1, da $x_{2k} \to 1$ und $x_{2k+1} \to -1$.

Satz 4.8 *Sei $\{x_n\}$ eine Folge von Elementen von $\overline{\mathbb{R}}$ und $a \in \overline{\mathbb{R}}$. Die folgenden Aussagen sind äquivalent:*

(i) a ist ein Häufungspunkt der Folge $\{x_n\}$;

(ii) jede Umgebung $U\left(a\right)$ von a enthält unendlich viel Glieder der Folge $\{x_n\}$;

(iii) für jede Umgebung $U\left(a\right)$ von a und für jedes $N \in \mathbb{N}$ existiert ein $n > N$ mit $x_n \in U\left(a\right)$.

Beweis. $(i) \Rightarrow (ii)$ Sei a ein Häufungspunkt, d.h. $x_{n_k} \to a$ für eine Teilfolge $\{x_{n_k}\}$. Dann enthält jede Umgebung $U\left(a\right)$ fast alle Glieder der Folge $\{x_{n_k}\}$. Folglich enthält $U\left(a\right)$ unendlich viel Glieder von $\{x_n\}$.

$(ii) \Rightarrow (iii)$ Nehmen wir an, dass jede Umgebung von a unendlich viel Glieder von $\{x_n\}$ enthält. Betrachten wir die Menge

$$S = \{n \in \mathbb{N} : x_n \in U\left(a\right)\}.$$

Da diese Menge unendlich ist, so hat S keine obere Schranke (jede beschränkte Teilmenge von $\mathbb{N}$ ist endlich). Somit ist beliebiges $N \in \mathbb{N}$ keine obere Schranke, woraus folgt, dass es ein $n > N$ gibt mit $x_n \in U\left(a\right)$.

$(iii) \Rightarrow (i)$. Bestimmen wir per Induktion in k eine Teilfolge $\{x_{n_k}\}_{k=1}^{\infty}$ mit $x_{n_k} \to a$. Sei zuerst $a \in \mathbb{R}$. Wählen wir erst $n_1 > 1$ 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 $x_{n_k} \in U_{1/k}(a)$, was immer möglich nach Voraussetzung ist. 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 $x_{n_k} \to a$.

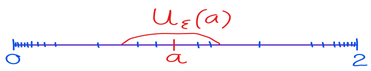
Im Fall $a = \pm\infty$ erfolgt der Beweis analog: man soll nur die Umgebungen $U_{\pm k}(a)$ statt $U_{1/k}(a)$ betrachten. ■

08.06.18 H1

Beispiel. Bestimmen wir die Häufungspunkte der Folge

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

Für gerade n gilt $x_n = \frac{1}{n}$ und für ungerade n gilt $x_n = 2 - \frac{1}{n}$. Da $\frac{1}{n} \to 0$ und $2 - \frac{1}{n} \to 2$, so sind 0 und 2 die Häufungspunkte der Folge. Zeigen wir, dass es keinen anderen Häufungspunkt gibt. Da für alle $n \in \mathbb{N}$ gilt $x_n \in [0,2]$, so gibt es keinen Häufungspunkt außerhalb des Intervalls $[0,2]$. Sei $a \in (0,2)$. Dann gibt es $\varepsilon > 0$ derart, dass die Umgebung $U_\varepsilon(a)$ disjunkt von $U_\varepsilon(0)$ und $U_\varepsilon(2)$ ist.



Die Menge der Glieder x_n in $U_\varepsilon(a)$ mit geraden n ist endlich, und mit ungeraden n auch endlich. Somit ist die Menge von Glieder x_n in $U_\varepsilon(a)$ endlich, und a ist kein Häufungspunkt. Es folgt, dass

$$\liminf x_n = 0 \quad \text{und} \quad \limsup x_n = 2.$$

Für weitere ähnliche Beispiel siehe Aufgabe 87. Es gibt ein interessantes Beispiel der Folge $\{x_n\}$ deren Menge von Häufungspunkten gleich $\overline{\mathbb{R}}$ ist (Aufgabe 93).

Hauptsatz 4.9 (Satz von Bolzano-Weierstraß) *Jede beschränkte Folge von reellen Zahlen besitzt eine konvergente Teilfolge.*

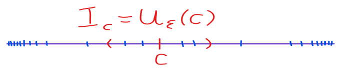
Beweis. Eine Folge $\{x_n\}$ besitzt eine konvergente Teilfolge genau dann, wenn $\{x_n\}$ einen Häufungspunkt in $\mathbb{R}$ hat. Ein $c \in \mathbb{R}$ ist ein Häufungspunkt von $\{x_n\}$ genau dann wenn

$$\forall \varepsilon > 0 \quad U_\varepsilon(c) \text{ enthält unendlich viel Glieder von } \{x_n\} \tag{4.18}$$

(Satz 4.8). Sei $\{x_n\}$ eine beschränkte Folge, d.h. es gibt $a, b \in \mathbb{R}$ derart, dass $a \le x_n \le b$ für alle $n \in \mathbb{N}$. Nehmen wir das Gegenteil an, dass $\{x_n\}$ keinen reellen Häufungspunkt hat. Diese Bedingung ist die Negation von (4.18) für jedes $c \in \mathbb{R}$:

$$\forall c \in \mathbb{R} \ \ \exists \varepsilon > 0 \ \text{ so dass } U_\varepsilon(c) \ \text{ nur endlich viel Glieder von } \{x_n\} \text{ enthält.}$$

Setzen wir $I_c = U_\varepsilon(c)$ für jedes $c \in \mathbb{R}$ (beachten wir, dass ε von c abhängig ist). Die Familie $\{I_c\}_{c \in \mathbb{R}}$ von offenen Intervallen überdeckt offensichtlich die ganze Menge $\mathbb{R}$, insbesondere das Intervall $[a,b]$.



Nach dem Satz 4.7, es gibt eine endliche Teilüberdeckung $\{I_c\}_{c\in T}$ von $[a, b]$. Da jedes I_c nur endlich viel Glieder von $\{x_n\}$ enthält, erhalten wir, dass ihre Vereinigung $\{I_c\}_{c\in T}$ auch nur endlich viel Glieder von $\{x_n\}$ enthält (siehe Aufgabe 91 (c)), was im Widerspruch zur Bedingung steht, dass $[a, b]$ die ganze Folge $\{x_n\}$ enthält. ■

Korollar 4.10 *Jede Folge $\{x_n\}$ von reellen Zahlen besitzt eine Teilfolge, die einen Grenzwert in $\overline{\mathbb{R}}$ hat. Äquivalent ist die Menge von Häufungspunkten von $\{x_n\}$ in $\overline{\mathbb{R}}$ nichtleer.*

Beweis. Jede beschränkte Folge $\{x_n\}$ hat einen Häufungspunkt nach dem Satz 4.9. Ist die Folge $\{x_n\}$ nach oben unbeschränkt, so erhält jede Umgebung $U_E(+\infty) = (E, +\infty]$ von $+\infty$ unendlich viel Glieder x_n, da sonst die endliche Menge von x_n in $(E, +\infty)$ als auch die Menge von x_n in $(-\infty, E)$ nach oben beschränkt sind, was im Widerspruch zur Unbeschränkheit von $\{x_n\}$ steht. Somit ist $+\infty$ ein Häufungspunkt.

Ist die Folge $\{x_n\}$ nach unten unbeschränkt, so ist analog $-\infty$ ein Häufungspunkt. ■

4.9 Cauchy-Folge

Definition. Eine Folge $\{x_n\}$ von reellen Zahlen heißt *Cauchy-Folge* (oder *Fundamentalfolge*) falls

$$\forall\, \varepsilon > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n, m \geq N \ \text{ gilt } \ |x_n - x_m| < \varepsilon. \tag{4.19}$$

Die Bedeutung (4.19) heißt die *Cauchy-Bedingung.* Sie lässt sich äquivalent wie folgt umformulieren:

$$\forall \varepsilon > 0 \text{ gilt } \ |x_n - x_m| < \varepsilon \ \text{ für fast alle } n \text{ und fast alle } m. \tag{4.20}$$

Kurz bezeichnet man die Bedingung (4.19) bzw (4.20) wie folgt:

$$x_n - x_m \to 0 \ \text{ für } \ n, m \to \infty.$$

Hauptsatz 4.11 (Cauchy-Kriterium) *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 4.9 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

$$\forall \varepsilon > 0 \text{ für fast alle } k \text{ gilt } |x_{n_k} - a| < \varepsilon. \tag{4.21}$$

Es folgt aus (4.20), dass für fast alle n und k gilt

$$|x_n - x_{n_k}| < \varepsilon.$$

Zusammen mit (4.21) 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. ■

Beispiel. Betrachten wir den unendlichen *Kettenbruch*:

$$\frac{1}{1 + \frac{1}{1+\frac{1}{1+\ldots}}} \tag{4.22}$$

deren Wert wie folgt definiert ist. Betrachten wir die Folge $\{x_n\}_{n\in\mathbb{N}}$ die induktiv definiert wird wie folgt:

$$x_1 = 1 \quad \text{und} \quad x_{n+1} = \frac{1}{1 + x_n} \quad \text{für } n \in \mathbb{N}.$$

Dann gilt

$$x_2 = \frac{1}{1+1} = \frac{1}{2}, \quad x_3 = \frac{1}{1 + \frac{1}{1+1}} = \frac{1}{1 + \frac{1}{2}} = \frac{2}{3}, \quad x_4 = \frac{1}{1 + \frac{1}{1+\frac{1}{1+1}}} = \frac{1}{1 + \frac{2}{3}} = \frac{3}{5},$$

usw. Man kann beweisen, dass die Folge $\{x_n\}$ die Cauchy-Bedingung erfüllt und somit konvergent ist (siehe Aufgabe 85). Der Grenzwert $x = \lim x_n$ heißt der Wert des Kettenbruches (4.22). Man kann x aus einer Gleichung explizit bestimmen.

Beispiel. Betrachten wir die Folge

$$x_1 = 0, \quad x_{n+1} = \frac{1}{2}\left(1 - x_n^2\right) \quad \text{für } n \in \mathbb{N}, \tag{4.23}$$

so dass $x_2 = \frac{1}{2}$, $x_3 = \frac{3}{8}$, $x_4 = \frac{55}{128}$, $x_5 = \frac{13\,359}{32\,768}$ usw. Man kann beweisen, dass diese Folge die Cauchy-Bedingung erfüllt und somit konvergent ist (siehe Aufgabe 92). Danach bestimmt man $x = \lim x_n$.

4.10 Limes inferior und Limes superior

Sei $\{x_n\}$ eine Folge von reellen Zahlen. Bezeichnen wir mit H die Menge von allen Häufungspunkten in $\overline{\mathbb{R}}$ von $\{x_n\}$. Nach Korollar 4.10 ist die Menge H stets nichtleer. Nach dem Satz 1.11 existieren das Infimum $\inf H \in [-\infty, +\infty)$ und das Supremum $\sup H \in (-\infty, +\infty]$.

Definition. Für die Folge $\{x_n\}$ definieren wir den *Limes inferior* durch

$$\liminf_{n\to\infty} x_n := \inf H$$

und den *Limes superior* durch

$$\limsup_{n\to\infty} x_n := \sup H.$$

Alternative Bezeichnungen:

$$\liminf_{n\to\infty} x_n = \varliminf_{n\to\infty} x_n \quad \text{und} \quad \limsup_{n\to\infty} x_n = \varlimsup_{n\to\infty} x_n.$$

Insbesondere existieren $\liminf x_n$ und $\limsup x_n$ als Elemente von $\overline{\mathbb{R}}$ für beliebige Folge $\{x_n\}$. Da die Menge H nichtleer ist, so gilt immer

$$\liminf x_n \le \limsup x_n.$$

Satz 4.12 *Sei $\{x_n\}$ eine Folge von Elementen von $\overline{\mathbb{R}}$.*
(a) *Der Grenzwert $\lim x_n$ existiert genau dann, wenn*

$$\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n.$$

(b) *Die beiden Werte $\liminf x_n$ und $\limsup x_n$ sind Häufungspunkte der Folge $\{x_n\}$. Folglich ist $\liminf x_n$ der minimale Häufungspunkt von $\{x_n\}$ und $\limsup x_n$ der maximale Häufungspunkt von $\{x_n\}$.*

Der Beweis in Aufgaben 88, 89 (b).

4.11 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 4.13 *Sei $\{z_n\}$ eine komplexwertige Folge.*

(*a*) *Die Konvergenz $z_n \to a$ für ein $a \in \mathbb{C}$ 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*) *(Cauchy-Kriterium) 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 Aufgabe 73). 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. 13.06.18 H13

(*b*) Wie in (*a*) haben wir

$$|z_n - z_m|^2 = |x_n - x_m|^2 + |y_n - y_m|^2$$

woraus folgt, dass

$$|z_n - z_m| \to 0 \quad \Leftrightarrow \quad |x_n - x_m| \to 0 \text{ und } \quad |y_n - y_m| \to 0$$

für $n, m \to \infty$. Somit ist $\{z_n\}$ Cauchy-Folge genau dann, wenn $\{x_n\}$ und $\{y_n\}$ Cauchy-Folgen sind.

(*c*) Nach (*a*), (*b*) und Satz 4.11 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 4.14 (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 4.2 und 4.13 (siehe Aufgabe 90).

Chapter 5

Reihen

5.1 Reellwertige Reihe

Sei $\{a_n\}_{n\in\mathbb{N}}$ eine Folge von reellen Zahlen. Der Ausdruck

$$\sum_{n=1}^{\infty} a_n$$

heißt *Reihe* (unendliche Summe). Die Summe (der Wert) der Reihe wird wie folgt definiert. Für jedes $N \in \mathbb{N}$ definieren wir die *Partialsumme*

$$S_N = \sum_{n=1}^{N} a_n.$$

Definition. 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 gleiches für die Folge $\{S_N\}$ gilt. Ist die Reihe $\sum a_n$ unbestimmt divergent, so ist der Wert von $\sum a_n$ nicht definiert.

Definition. Eine Reihe $\sum_{k=1}^{\infty} a_k$ heißt *nichtnegativ* falls alle Glieder a_k nichtnegative reelle Zahlen sind.

Satz 5.1 *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 $\sum_{n=1}^{\infty} a_n$ konvergiert;*
2. *oder $\sum_{n=1}^{\infty} a_n = +\infty$ 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 4.5. 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

$$S_N = \sum_{n=0}^{N} x^n = 1 + \sum_{n=1}^{N} x^n = 1 + \frac{x^{N+1} - x}{x - 1} = \frac{x^{N+1} - 1}{x - 1} \tag{5.1}$$

(siehe Aufgabe 39). Im Fall $x \in (-1, 1)$ erhalten wir $x^{N+1} \to 0$ für $N \to \infty$, woraus folgt

$$\sum_{n=0}^{\infty} x^n = \lim_{N \to \infty} S_N = \frac{1}{1 - x}. \tag{5.2}$$

Es folgt, dass für jedes $m \in \mathbb{Z}$

$$\sum_{n=m}^{\infty} x^n = x^m \sum_{n=m}^{\infty} x^{n-m} = x^m \sum_{k=0}^{\infty} x^k = \frac{x^m}{1 - x}. \tag{5.3}$$

Im Fall $x > 1$ folgt es aus (5.1), dass

$$S_N = \frac{x^{N+1} - 1}{x - 1} \to +\infty \text{ für } N \to \infty,$$

so dass $\sum_{n=0}^{\infty} x^n = +\infty$. Im Fall $x = 1$ gilt (5.1) 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

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

ist bestimmt divergent, während die Reihe

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

konvergent ist (siehe Aufgabe 96).

Satz 5.2 *Gilt* $0 \leq a_n \leq b_n$ *für alle* $n \in \mathbb{N}$, *so gilt*

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

Beweis. Seien A_N und B_N die Partialsummen von $\sum a_n$ bzw $\sum b_n$. Die Voraussetzung $a_n \leq b_n$ ergibt $A_N \leq B_N$ und somit $\lim A_N \leq \lim B_N$, was zu beweisen war. ∎

5.2 Zahlensystem: q-adische Brüche

Sei $q \geq 2$ eine natürliche Zahl. Betrachten wir eine unendliche Folge$\{a_k\}_{k=1}^{\infty}$ von q-adischen Ziffern (d.h. $a_k \in \{0, ..., q-1\}$) und die Reihe

$$\sum_{k=1}^{\infty} a_k q^{-k} = a_1 q^{-1} + a_2 q^{-2} + ... \,. \tag{5.4}$$

Satz 5.3 *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 5.3. 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 (5.3)

$$\sum_{k=1}^{\infty} a_k q^{-k} \leq (q-1) \sum_{k=1}^{\infty} q^{-k} = (q-1) \frac{q^{-1}}{1-q^{-1}} = 1. \tag{5.5}$$

Somit ist die Reihe $\sum_{k=1}^{\infty} a_k q^{-k}$ konvergent und ihre Summe liegt in $[0,1]$. ∎

Beispiel. Es folgt aus (5.5), dass

$$(0, 111...)_q = \sum_{k=1}^{\infty} q^{-k} = \frac{1}{q-1}. \tag{5.6}$$

5.3 Komplexwertige Reihen

Sei jetzt $\{a_k\}_{k=1}^{\infty}$ eine Folge von komplexen Zahlen.

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 den Eigenschaften des Limes erhalten wir die folgenden Eigenschaften von Reihen.

Satz 5.4 *Für $a_k, b_k, c \in \mathbb{C}$ gelten die Identitäten*

$$\sum_{k=1}^{\infty}(a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k,$$

$$\sum_{k=1}^{\infty} ca_k = c\sum_{k=1}^{\infty} a_k,$$

vorausgesetzt, dass die rechten Seiten bestimmt sind.

Bemerkung. Die letzte Bedingung bedeutet, dass die Reihen $\sum a_k$ und $\sum b_k$ konvergent sind. Im Fall wenn a_k, b_k, c reell sind, kann man auch bestimmt divergente Reihen erlauben, vorausgesetzt, dass die Operationen mit den Unendlichkeiten in den rechten Seiten bestimmt sind.

Beweis. Betrachten wir die Partialsumme

$$S_n = \sum_{k=1}^{n}(a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k = A_n + B_n,$$

wobei A_n und B_n die Partialsummen von $\sum a_k$ bzw $\sum b_k$ sind. Nach dem Satz 4.14 (4.4 im Fall von reellen Reihen) erhalten wir

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty}(A_n + B_n) = \lim_{n\to\infty} A_n + \lim_{n\to\infty} B_n = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k.$$

Die zweite Identität wird analog bewiesen. ■

Satz 5.5 (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. 15.06.18 H1

(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

$$S_n = \sum_{k=0}^{n} x^k = \frac{x^{n+1} - 1}{x - 1} \to \frac{1}{1 - x} \quad \text{für } n \to \infty,$$

weil $|x^{n+1}| = |x|^{n+1} \to 0$. Im Fall $|x| \geq 1$ ist die geometrische Reihe divergent, da $\left|x^k\right| = |x|^k \geq 1$ und somit $\left\{x^k\right\}$ keine Nullfolge ist.

5.4 Majorantenkriterium und absolute Konvergenz

Satz 5.6 (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*

$$|a_k| \leq b_k \quad \textit{für alle } k \in \mathbb{N} \tag{5.7}$$

so ist $\sum_{k=1}^{\infty} a_k$ konvergent und

$$\left| \sum_{k=1}^{\infty} a_k \right| \leq \sum_{k=1}^{\infty} b_k. \tag{5.8}$$

Gilt die Bedingung (5.7), 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. Da $b_k \geq 0$, so existiert $\sum_{k=1}^{\infty} b_k$ immer als Element von $\overline{\mathbb{R}}$ (Satz 5.1). Die Konvergenz von $\sum b_k$ ist äquivalent zu $\sum_{k=1}^{\infty} b_k < \infty$.

Beweis. Im Beweis benutzen wir die folgende Verallgemeinerung der Dreiecksungleichung: für beliebige endliche Folge $\{c_k\}_{k=1}^{n}$ von komplexen Zahlen gilt

$$\left| \sum_{k=1}^{n} c_k \right| \leq \sum_{k=1}^{n} |c_k|,$$

was man per Induktion nach n beweist.

Bezeichnen wir

$$A_n = \sum_{k=1}^{n} a_k \quad \text{und} \quad B_n = \sum_{k=1}^{n} b_k.$$

Dann gilt es für alle natürliche Zahlen $n > m$

$$|A_n - A_m| = \left| \sum_{k=m+1}^{n} a_k \right| \leq \sum_{k=m+1}^{n} |a_k| \leq \sum_{k=m+1}^{n} b_k = B_n - B_m. \tag{5.9}$$

Für $m > n$ gilt analog

$$|A_n - A_m| \leq B_m - B_n,$$

und für ale $n, m \in \mathbb{N}$ erhalten wir

$$|A_n - A_m| \leq |B_n - B_m|.$$

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 $\sum_{k=1}^{\infty} a_k$ konvergiert nach dem Satz 4.11.

Um (5.8) zu beweisen, setzen wir

$$A = \lim A_n = \sum_{k=1}^{\infty} a_k \quad \text{und} \quad B = \lim B_n = \sum_{k=1}^{\infty} b_k.$$

Analog zu (5.9) gilt

$$|A_n| = \left|\sum_{k=1}^{n} a_k\right| \leq \sum_{k=1}^{n} |a_k| \leq \sum_{k=1}^{n} b_k = B_n.$$

Wir haben auch

$$|A| = |A - A_n + A_n| \leq |A - A_n| + |A_n| \leq |A - A_n| + B_n$$

woraus folgt

$$|A| \leq \lim |A - A_n| + \lim B_n = 0 + B = B,$$

was zu beweisen war. ■

Definition. Eine komplexwertige Reihe $\sum_{k=1}^{\infty} a_k$ heißt *absolut konvergent* falls

$$\sum_{k=1}^{\infty} |a_k| < \infty.$$

Korollar 5.7 *Ist die Reihe $\sum a_k$ absolut konvergent, so ist sie konvergent. Es gilt auch*

$$\left|\sum_{k=1}^{\infty} a_k\right| \leq \sum_{k=1}^{\infty} |a_k|. \tag{5.10}$$

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 5.6 ist $\sum a_k$ konvergent und (5.10) 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 der Folge $\{|c_k|\}$. Da

$$\left|\frac{c_k}{k^2}\right| \leq \frac{C}{k^2}$$

und die Reihe $\sum \frac{C}{k^2} = C \sum \frac{1}{k^2}$ konvergent ist, so erhalten wir nach dem Satz 5.6 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 $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$.

5.5 Quotientenkriterium

Satz 5.8 (Quotientenkriterium, d'Alembert-Kriterium) *Sei* $\{a_n\}_{n=1}^{\infty}$ *eine komplexwertige Folge mit* $a_n \neq 0$ *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. Nehmen wir das Gegenteil an, dass es in $[q, +\infty)$ unendlich viel Glieder der Folge $\{r_n\}$ gibt. Dann liegt in $[q, +\infty)$ eine Teilfolge von $\{r_n\}$, und nach dem Korollar 4.10 gibt es in $[q, +\infty]$ einen Häufungspunkt der Folge $\{r_n\}$. Allerdings ist das nicht möglich, da alle Häufungspunkte dieser Folge $\leq r$ sind.

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

$$\sum_{n=N}^{\infty} |a_n| \leq |a_N| \sum_{n=N}^{\infty} q^{n-N} = |a_N|\, q^{-N} \sum_{n=N}^{\infty} q^n < \infty,$$

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. ■

5.6 Bedingte Konvergenz

Es gibt die Reihen, die konvergent aber nicht absolut konvergent sind. Sie heißen *bedingt konvergent*. Zum Beispiel, die Leibniz-Reihe $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ ist bedingt konvergent. Für den Beweis brauchen wir den folgenden Satz.

Satz 5.9 *(*Leibniz-Kriterium*) Sei $\{c_k\}_{k=1}^{\infty}$ eine monoton fallende Folge von nicht-negativen reellen Zahlen mit $c_k \to 0$. Dann ist die alternierende Reihe*

$$\sum_{k=1}^{\infty} (-1)^k c_k = -c_1 + c_2 - c_3 + c_4 - ... \tag{5.11}$$

konvergent. Darüber hinaus erfüllen die Partialsummen $S_n = \sum_{k=1}^{n} (-1)^k c_k$ die Ungleichungen

$$S_{2m-1} \le \sum_{k=1}^{\infty} (-1)^k c_k \le S_{2m} \tag{5.12}$$

für alle $m \in \mathbb{N}$.

Bemerken wir, dass im Fall $c_k \not\to 0$ die Reihe (5.11) nach dem Trivialkriterium divergent ist.

Beweis. Beweisen wir zuerst, dass die Folge $\{S_{2m}\}_{m\in\mathbb{N}}$ eine Cauchy-Folge ist. Sei $l > m$. Dann haben wir

$$\begin{aligned} S_{2l} - S_{2m} &= \sum_{k=1}^{2l} (-1)^k c_k - \sum_{k=1}^{2m} (-1)^k c_k \\ &= \sum_{k=2m+1}^{2l} (-1)^k c_k = -(c_{2m+1} - c_{2m+2} + ... + c_{2l-1} - c_{2l}) \\ &= -\sum_{j=m+1}^{l} (c_{2j-1} - c_{2j}) \end{aligned} \tag{5.13}$$

(der rigorose Beweis der letzten Identität erfolgt per Induktion nach $l \ge m+1$). Da $c_{2j-1} \ge c_{2j}$, so folgt es, dass

$$\begin{aligned} 0 &\le S_{2m} - S_{2l} = \sum_{j=m+1}^{l} (c_{2j-1} - c_{2j}) \\ &\le \sum_{j=lm+1}^{l} (c_{2j-1} - c_{2j}) + \sum_{j=m+1}^{l} (c_{2j} - c_{2j+1}) \\ &= \sum_{j=m+1}^{l} c_{2j-1} - \sum_{j=m+1}^{l} c_{2j+1} = (c_{2m+1} + ... + c_{2l-1}) - (c_{2m+3} + ... + c_{2l+1}) \\ &= \sum_{i=m}^{l-1} c_{2i+1} - \sum_{j=m+1}^{l} c_{2j+1} \qquad \text{(Wechsel } i = j-1 \text{ in erster Summe)} \\ &= c_{2m+1} - c_{2l+1} \le c_{2m+1}. \end{aligned}$$

Da $c_{2m+1} \to 0$ für $m \to \infty$, so erhalten wir, dass

$$S_{2m} - S_{2l} \to 0 \quad \text{für } l, m \to \infty,$$

d.h. $\{S_{2m}\}$ ist eine Cauchy-Folge und somit konvergent. Sei

$$a = \lim_{m\to\infty} S_{2m}.$$

Da

$$S_{2m} - S_{2m-1} = (-1)^{2m} c_{2m} \to 0 \ \text{ für } m \to \infty,$$

so erhalten wir, dass auch

$$\lim_{m\to\infty} S_{2m-1} = a,$$

woraus folgt, dass

$$\lim_{n\to\infty} S_n = a,$$

d.h. die Reihe (5.11) konvergent ist, und ihre Summe ist gleich a.

Aus $S_{2m} \geq S_{2l}$ für $l > m$ erhalten wir für $l \to \infty$ dass

$$S_{2m} \geq \lim_{l\to\infty} S_{2l} = a,$$

was äquivalent zu

$$\sum_{k=1}^{\infty} (-1)^k c_k \leq S_{2m}$$

ist. Analog beweist man, dass für alle $l > m$

$$\begin{aligned} S_{2l+1} - S_{2m+1} &= \sum_{k=1}^{2l+1} (-1)^k c_k - \sum_{k=1}^{2m+1} (-1)^k c_k \\ &= \sum_{k=2m+2}^{2l+1} (-1)^k c_k = (c_{2m+2} - c_{2m+3} + ... + c_{2l} - c_{2l+1}) \\ &= \sum_{j=m+1}^{l} (c_{2j} - c_{2j+1}) \geq 0. \end{aligned} \tag{5.14}$$

Aus $S_{2m+1} \leq S_{2l+1}$ erhalten wir für $l \to \infty$ dass

$$S_{2m+1} \leq \lim_{l\to\infty} S_{2l+1} = a,$$

und somit

$$\sum_{k=1}^{\infty} (-1)^k c_k \geq S_{2m+1}.$$

■

Alternativer Beweis. Da

$$S_{2(m+1)} - S_{2m} = (-1)^{2m+2} c_{2m+2} + (-1)^{2m+1} c_{2m+1} = c_{2m+2} - c_{2m+1} \leq 0,$$

so ist die Folge $\{S_{2m}\}$ monoton fallend. Analog haben wir

$$S_{2(m+1)+1} - S_{2m+1} = (-1)^{2m+3} c_{2m+3} + (-1)^{2m+2} c_{2m+2} = -c_{2m+3} + c_{2m+2} \geq 0,$$

so dass die Folge $\{S_{2m+1}\}$ monoton steigend ist. Da

$$S_{2m+1} = S_{2m} + (-1)^{2m+1} c_{2m+1} = S_{2m} - c_{2m+1}, \tag{5.15}$$

so folgt es, dass $S_{2m+1} \leq S_{2m}$. Somit erhalten wir eine Intervallschachtelung $\{[S_{2m+1}, S_{2m}]\}_{m\in\mathbb{N}}$.

Nach dem Intervallschachtelungsprinzip gibt es eine reelle Zahl a mit

$$a \in \bigcap_{m \in \mathbb{N}} [S_{2m+1}, S_{2m}],$$

d.h.

$$S_{2m+1} \le a \le S_{2m} \quad \forall m \in \mathbb{N}. \tag{5.16}$$

Da $c_{2m+1} \to 0$, so folgt es aus (5.15), dass $S_{2m} - S_{2m+1} \to 0$ für $m \to \infty$. Da

$$0 \le S_{2m} - a \le S_{2m} - S_{2m+1}$$

und

$$0 \le a - S_{2m+1} \le S_{2m} - S_{2m+1},$$

so folgt es, dass

$$\lim S_{2m} = \lim S_{2m+1} = a$$

und $\lim S_n = a$. Die Ungleichung (5.12) ist dann äquivalent zu (5.16). ■

Beispiel. Die Leibniz-Reihe $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ ist konvergent nach dem Leibniz-Kriterium, da die Folge $\left\{\frac{1}{n}\right\}$ monoton fallend ist und $\frac{1}{n} \to 0$. Andererseits ist die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ divergent (siehe Aufgabe 101), so dass die Konvergenz von $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ bedingt ist.

5.7 * Zahlensystem: q-adische Darstellung reeller Zahlen

Jetzt definieren wir q-adische Darstellung von den reellen Zahlen. Der nächste Satz ist die Umkehrung von dem Satz 5.3.

Satz 5.10 *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

$$\sum_{k=1}^{n} a_k q^{-k} \le x < \sum_{k=1}^{n} a_k q^{-k} + q^{-n}. \tag{5.17}$$

Induktionsanfang für $n = 1$. Die Ungleichung (5.17) für $n = 1$ ist

$$a_1 q^{-1} \le x < a_1 q^{-1} + q^{-1}, \tag{5.18}$$

was äquivalent zu

$$a_1 \le qx < a_1 + 1 \tag{5.19}$$

ist, d.h.

$$a_1 = [qx],$$

wobei $[\cdot]$ Gaußklammer ist (siehe Satz 2.9). 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 (5.17) gibt. Setzen wir

$$y = \sum_{k=1}^{n-1} a_k q^{-k}$$

und umschreiben (5.17) 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 (5.17) für $n \to \infty$, dass $x = 0, a_1 a_2 \ldots$ ■

Mit Hilfe von Sätzen 2.19 und 5.10 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 5.10 gegeben. Sei $a > 0$. Nach Satz 2.19 hat a die q-adische Darstellung

$$a = \sum_{k=0}^{n} a_k q^k = (a_n a_{n-1} \ldots a_0)_q \ .$$

Nach Satz 5.10 hat b die Darstellung als ein q-adischer Bruch

$$b = \sum_{l=1}^{\infty} b_l q^{-l} = (0, b_1 b_2 b_3 \ldots)_q \ .$$

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 *q-adische Zahl.* Somit erhalten wir folgendes.

Korollar 5.11 *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 5.10 enthält keine Eindeutigkeitsaussage, da dies nicht immer gilt. Zum Beispiel, es folgt aus (5.6), 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

$$\text{die Menge } \{k \in \mathbb{N} : b_k < q - 1\} \text{ unendlich ist.} \tag{5.20}$$

Man kann beweisen, dass für jedes $x \in [0, 1)$ es genau einen q-adischen echten Bruch mit dem Wert x gibt. Eigentlich, die Konstruktion im Beweis des Satzes 5.10 ergibt schon einen echten q-adischen Bruch. Für die Eindeutigkeit vergleicht man zwei echten Brüche $(0, b_1 b_2 ...)_q$ und $(0, b_1' b_2' ...)_q$. Sind die Folgen $\{b_k\}$ und $\{b_k'\}$ nicht identisch, so gibt es ein minimales $m \in \mathbb{N}$ mit $b_m \neq b_m'$, z.B. mit $b_m < b_m'$. Dann zeigt man, dass

$$(0, b_1 b_2 ...)_q < (0, b_1' b_2' ...)_q \, ,$$

woraus die Eindeutigkeit folgt.

Als Beispiel von Anwendung des Dualsystems beweisen wir dass $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$. 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 5.10, 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 (5.20) erfüllen. Dann 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 $\mathcal{F} \setminus \mathcal{F}_e$ abzählbar ist. Da $[0,1) \sim \mathbb{R}$ und nach dem Satz 2.25 $\mathcal{F} \sim \mathcal{F}_e$, so erhalten wir

$$\mathcal{F} \sim \mathbb{R}.$$

Andererseits gilt

$$\mathcal{F} \sim \mathcal{P}(\mathbb{N}),$$

da die folgende Abbildung

$$\begin{aligned} \psi &: \mathcal{F} \to \mathcal{P}(\mathbb{N}) \\ \psi(b) &= \{k \in \mathbb{N} : b_k = 0\} \end{aligned}$$

bijektiv ist. Somit haben wir, dass $\mathbb{R} \sim \mathcal{F} \sim \mathcal{P}(\mathbb{N})$, woraus $\mathbb{R} \sim \mathcal{P}(\mathbb{N})$ folgt

Nach dem Satz 2.27 gilt $\mathbb{N} \prec \mathcal{P}(\mathbb{N})$. Da $\mathcal{P}(\mathbb{N}) \sim \mathbb{R}$, so erhalten wir einen alternativen Beweis der Ungleichung $\mathbb{N} \prec \mathbb{R}$ (siehe Satz 2.24).

5.8 * Kommutativ und Assoziativgesetze für die Reihen

Für die absolut konvergenten Reihen gelten die Kommutativ- und Assoziativgesetze, die wir ohne Beweis angeben.

Satz 5.12 *Sei $\sum_{k=1}^{\infty} a_k$ eine absolut konvergente Reihe von komplexen Zahlen. Sei $\sum_{k=1}^{\infty} b_k$ eine Reihe, die aus $\sum_{k=1}^{\infty} a_k$ durch Vertauschung und Gruppierung der Glieder erhalten wird. Dann ist $\sum_{k=1}^{\infty} b_k$ auch absolut konvergent und*

$$\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} a_k.$$

Für die bedingt konvergenten Reihen gilt im Gegenteil folgendes.

Satz 5.13 *Sei $\sum_{k=1}^{\infty} a_k$ eine bedingt konvergente Reihe von reellen Zahlen. Dann für jedes $c \in \overline{\mathbb{R}}$ es gibt eine Reihe $\sum_{k=1}^{\infty} b_k$, die aus $\sum_{k=1}^{\infty} a_k$ durch Vertauschung der Glieder erhalten wird und so dass $\sum_{k=1}^{\infty} b_k = c$.*

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.

5.9 * Cauchy-Produkt zweier Reihen

Erinnern wir uns, dass für das Produkt zweier endlichen Summen die folgende Identität gilt:

$$\left(\sum_{k\in I} a_k\right)\left(\sum_{l\in J} b_l\right) = \sum_{(k,l)\in I\times J} a_k b_l \tag{5.21}$$

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 (5.21) 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 (5.21) 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

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}. \tag{5.22}$$

Um die Bedeutung von c_n zu verstehen, betrachten wir die folgende unendliche Tabelle mit allen Summanden $a_k b_l$ für $k, l \in \mathbb{Z}_+$:

$$\begin{array}{ccccccccc}
a_0b_0 & a_0b_1 & a_0b_2 & \dots & \dots & \dots & \dots & \boxed{a_0b_n} & \dots \\
a_1b_0 & a_1b_1 & a_1b_2 & \dots & \dots & \dots & \boxed{a_1b_{n-1}} & \dots & \\
a_2b_0 & a_2b_1 & a_2b_2 & \dots & \dots & \dots & \dots & & \\
\dots & \dots & \dots & \dots & \boxed{a_kb_{n-k}} & \dots & & & \\
\dots & \dots & \dots & \dots & \dots & & & & \\
\dots & \dots & \dots & \dots & & & & & \\
\dots & \boxed{a_{n-1}b_1} & \dots & & & & & \dots & \dots \\
\boxed{a_nb_0} & \dots & & & & & \dots & a_nb_n & \dots \\
\dots & & & & & & \dots & \dots & \dots
\end{array} \tag{5.23}$$

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

$$\sum_{n=0}^{\infty} c_n = \sum_{k,l\in\mathbb{Z}_+} a_k b_l = \sum_{k=0}^{\infty} a_k \sum_{l=0}^{\infty} b_l. \tag{5.24}$$

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. Im nächsten Satz werden die Bedingungen vorgelegt, die die Identität (5.24) garantieren.

Satz 5.14 (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*

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{l=0}^{\infty} b_l\right). \tag{5.25}$$

Bemerkung. Man kann die folgende Erweiterung des Satzes 5.14 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 (5.25). Allerdings ist die Konvergenz von $\sum c_n$ in diesem Fall nicht unbedingt absolut.

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

$$\sum_{k=0}^{\infty} a_k \sum_{l=0}^{\infty} b_l = \lim A_n \lim B_n = \lim \left(A_n B_n\right),$$

müssen wir insbesondere zeigen, dass die Folge $\{C_n\}$ konvergiert und

$$\lim_{n\to\infty} C_n = \lim_{n\to\infty} \left(A_n B_n\right). \tag{5.26}$$

Bemerken wir, dass

$$C_n = \sum_{m=0}^{n} c_m = \sum_{m=0}^{n}\sum_{k=0}^{m} a_k b_{m-k} = \sum_{\{k,l\in\mathbb{Z}_+ : k+l\leq n\}} a_k b_l,$$

d.h. C_n ist die Summe von allen Gliedern $a_k b_l$ aus dem Dreieck

$$\{k, l \in \mathbb{Z}_+ : k + l \leq n\}.$$

Für $A_n B_n$ erhalten wir

$$A_n B_n = \sum_{k=0}^{n} a_k \sum_{l=0}^{n} b_l = \sum_{\{k,l\in\mathbb{Z}_+ : k\leq n, l\leq n\}} a_k b_l,$$

d.h. $A_n B_n$ ist die Summe von allen Gliedern $a_k b_l$ aus dem Quadrat

$$\{k, l \in \mathbb{Z}_+ : k \leq n, l \leq n\}.$$

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

$$k + l \leq n \;\Rightarrow\; k \leq n, l \leq n$$

und für $m = [n/2]$,

$$k \leq m, l \leq m \;\Rightarrow\; k + l \leq n.$$

Auf der Tabelle (5.23) 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 4.1, dass C_n den gleichen Grenzwert hat, d.h. (5.26). 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 5.6 erhalten wir die absolute Konvergenz von $\sum c_n$.

Es bleibt noch die Identität (5.26) zu beweisen, d.h. $A_n B_n - C_n \to 0$. Wir haben

$$A_n B_n - C_n = \sum_{\{k,l: k \le n, l \le n\}} a_k b_l - \sum_{\{k,l: k+l \le n\}} a_k b_l = \sum_{\{k,l: k \le n. l \le n. k+l > n\}} a_k b_l.$$

Nach der Dreiecksungleichung gilt

$$|A_n B_n - C_n| \le \sum_{\{k,l: k \le n. l \le n. k+l > n\}} |a_k| \, |b_l| \, .$$

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

$$A_n^* B_n^* - C_n^* = \sum_{\{k,l: k \le n. l \le n. k+l > n\}} |a_k| \, |b_l|$$

woraus folgt

$$|A_n B_n - C_n| \le |A_n^* B_n^* - C_n^*| \, .$$

Da nach dem ersten Teil des Beweises gilt $|A_n^* B_n^* - C_n^*| \to 0$, erhalten wir $|A_n B_n - C_n| \to 0$ und somit $A_n B_n - C_n \to 0$, was zu beweisen war. ■

5.10 * Existenz und Eindeutigkeit von $\mathbb{R}$

Here präsentieren wir kurz noch zwei Wege um die Menge $\mathbb{R}$ zu konstruieren, angenommen dass die Mengen $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ schon bekannt sind (siehe Abschnitt 2.9). Alle Wege von $\mathbb{N}$ zu $\mathbb{R}$ sind ziemlich lang (und auch langweilig) und wir werden keine weiteren Einzelheiten geben.

Reelle Zahlen als q-adische Zahlen

Man benutzt das q-adische Zahlensystem (zum Beispiel, mit $q = 2$) um zunächst $\mathbb{R}_+$ als die Menge von q-adischen Zahlen (Folgen)

$$(b_n b_{n-1} ... b_0, a_1 a_2 ...)_q$$

zu definieren, wobei a_k, b_l die q-adischen Ziffern sind. Man muß sowohl die Operationen $+$ und $\cdot$ als auch die Ungleichung mit Hilfe von q-adischen Darstellungen definieren und die Gültigkeit von allen Axiomen von reellen Zahlen beweisen.

Zum Beispiel, betrachten wie nur die Menge

$$M = \{0, a_1 a_2 ... \mid a_k \text{ sind } q\text{-adische Ziffern}\}$$

und beweisen, dass diese Menge das Vollständigkeitsaxiom erfüllt. Dafür definieren wir zunächst die Ungleichung in M wie folgt: für

$$a = 0, a_1 a_2 ... \quad \text{und} \quad b = 0, b_1 b_2 ...$$

gilt $a \leq b$ falls entweder $a_k = b_k$ für alle $k \in \mathbb{N}$ oder es ein $n \in N$ gibt so dass

$$a_k = b_k \text{ für alle } k < n \ \text{ und } a_n < b_n.$$

Seien X und Y zwei nichtleere Teilmengen von M mit

$$x \leq y \ \ \forall x \in X \ \text{ und } \forall y \in Y.$$

Beweisen wir, dass es ein $c \in M$ gibt mit

$$x \leq c \leq y \quad \forall x \in X \text{ und } \forall y \in Y.$$

Zuerst wählen wir die maximale Zahl von

$$0, x_1 0000....$$

für alle $x \in X$, sei

$$\max_{x \in X} 0, x_1 000... = 0, c_1 000... \tag{5.27}$$

(das Maximum existiert da die Menge von verschiedenen Werten für x_1 endlich ist). Dann setzen wir

$$\max_{x \in X} 0, x_1 x_2 000... = 0, c_1' c_2 000...$$

und zeigen, dass $c_1' = c_2$. Offensichtlich gilt

$$0, c_1 000... \leq 0, c_1' c_2 000...$$

so dass $c_1 \leq c_1'$. Andererseits kann c_1' nicht größer als c_1 sein nach (5.27). Somit gilt $c_1 = c_2$ und

$$\max_{x \in X} 0, x_1 x_2 000... = 0, c_1 c_2 000...$$

Weiter erhalten wir

$$\max_{x \in X} 0, x_1 x_2 x_3 = 0, c_1 c_2 c_3 000...$$

und so weiter. Somit erhalten wir die q-adische Zahl

$$c = 0, c_1 c_2 .... \in M$$

mit $x \le c$ für alle $x \in X$. Auch gilt $c \le y$ für alle y da für jedes n gilt für ein $x \in M$

$$0, c_1 ... c_n 000... = 0, x_1 ... x_n 000... \le x,$$

woraus folgt für alle $y \in M$

$$0, c_1 ... c_n 000... \le y$$

und $c \le y$.

Reelle Zahlen als Cauchy-Folgen

Sei der Körper $\mathbb{Q}$ schon definiert. Wir betrachten wir Folgen $\{x_k\}_{k=1}^{\infty}$ die aus rationalen Zahlen bestehen. Man definiert die Begriffe von Grenzwert und Cauchy-Folge genau so wir im Kapitel 4 aber anstatt $\mathbb{R}$ benutzt man immer $\mathbb{Q}$. Man sagt, dass zwei Cauchy-Folgen $\{x_k\}$ und $\{y_k\}$ äquivalent sind falls $x_k - y_k \to 0$ für $k \to \infty$. Die reellen Zahlen werden als Äquivalenzklassen von Cauchy-Folgen definiert, d.h. jede Cauchy-Folge stellt eine reelle Zahl dar, und die äquivalente Cauchy-Folgen stellen die gleiche Zahl dar. Die Menge von reellen Zahlen wird mit $\mathbb{R}$ bezeichnet.

Wir schreiben $x \sim \{x_k\}$ falls die reelle Zahl x der Cauchy-Folge $\{x_k\}$ entspricht. Jede rationale Zahl a wird mit der Cauchy-Folge $\{a\}$ identifiziert so dass $\mathbb{Q} \subset \mathbb{R}$.

Man definiert die addition in $\mathbb{R}$ wie folgt: für die Zahlen $x \sim \{x_k\}$ und $y \sim \{y_k\}$ setzen wir $x + y \sim \{x_k + y_k\}$ da $\{x_k + y_k\}$ auch Cauchy-Folge ist. Analog definiert man das Produkt und Ungleichung. Man beweist danach dass $\mathbb{R}$ alle Axiome von reellen Zahlen erfüllt inklusive das Vollständigkeitsaxiom.

Eindeutigkeit von $\mathbb{R}$

Es gibt verschiedene Mengen, die die Axiome von $\mathbb{R}$ erfüllen. Alle solchen Mengen heißen die *Modelle* von $\mathbb{R}$. Wir haben schon drei Modellen von $\mathbb{R}$ gesehen: die Menge von Dedekindschen Schnitten, die Menge von q-adischen Brüchen[1] und die Menge von Äquivalenzklassen von Cauchy-Folgen.

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 den Eigenschaften:

1. $\varphi(x + y) = \varphi(x) + \varphi(y)$
2. $\varphi(xy) = \varphi(x)\varphi(y)$
3. $x < y \Leftrightarrow \varphi(x) < \varphi(y)$.

Daraus folgt, dass auch $\varphi(0') = 0''$, $\varphi(1') = 1''$, $\varphi(-x) = -\varphi(x)$ und $\varphi(x^{-1}) = \varphi(x)^{-1}$. Eine Abbildung φ mit diesen Eigenschaften heißt *Isomorphismus*, und die

[1] Für verschiedene Werte von q erhält man allerdings verschiedene Modelle von $\mathbb{R}$.

Modellen $\mathbb{R}'$ und $\mathbb{R}''$, für die ein Isomorphismus existiert, heißen *isomorph*. Somit sind je zwei beliebige 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.

Chapter 6

Exponentialfunktion

20.06.18 H13

6.1 Exponentialreihe und die Zahl e

Definition. Sei $x \in \mathbb{C}$. Die Reihe

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots$$
$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \ldots$$

heißt die *Exponentialreihe.*

Behauptung. *Die Exponentialreihe konvergiert absolut für alle $x \in \mathbb{C}$.*

Beweis. Für $x = 0$ ist die Konvergenz offensichtlich. Sei $x \neq 0$. Setzen wir $a_n = \frac{x^n}{n!}$ und erhalten

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{x^{n+1} n!}{(n+1)! x^n}\right| = \left|\frac{x}{n+1}\right| = \frac{|x|}{n+1} \to 0 \quad \text{für } n \to \infty.$$

Somit gilt

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 0 < 1$$

und die Reihe $\sum a_n$ ist absolut konvergent nach dem Quotientenkriterium des Satzes 5.8. ∎

Definition. Die Summe der Exponentialreihe heißt die Exponentialfunktion von $x \in \mathbb{C}$ und wird wie folgt bezeichnet:

$$\boxed{\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}}. \tag{6.1}$$

Ist x reell, dann ist $\exp(x)$ auch reell, was man aus (6.1) sieht. Zum Beispiel, $\exp(0) = 1$.

Die Zahl $\exp(1)$ wird auch als e bezeichnet. Die Zahl e heißt die *Eulersche Zahl.* Es folgt aus (6.1), dass

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

Man erhält Annäherungen von e indem man die Partialsummen

$$S_N = \sum_{n=0}^{N} \frac{1}{n!}$$

für großen Werten von N berechnet. Zum Beispiel, es gilt

$$S_{10} = \sum_{n=0}^{10} \frac{1}{n!} = \frac{9864101}{3628800} = 2,718281801...$$

$$S_{100} = \sum_{n=0}^{100} \frac{1}{n!} = 2,7182818284590452354...$$

In der Tat gilt

$$e = 2,718\,281828\,4590452354...$$

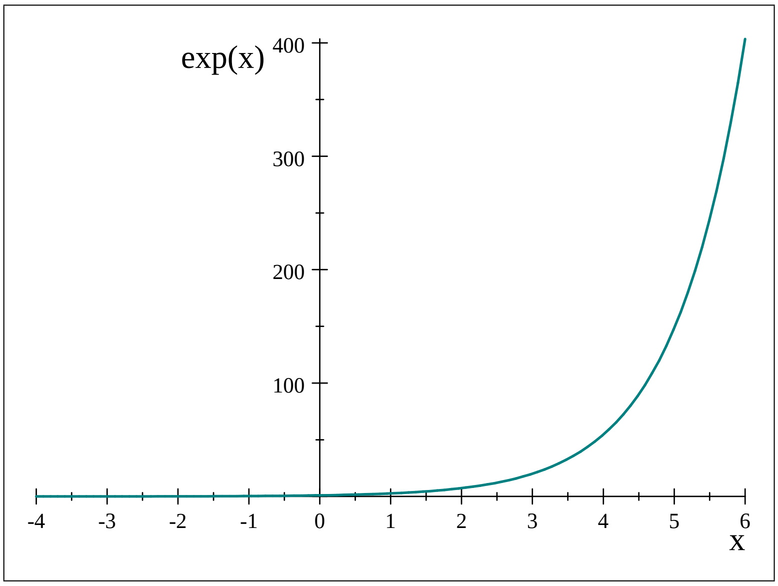
Numerische Berechnung mit Hilfe von (6.1) ergibt folgendes:

$$\begin{aligned}
\exp(2) &= 7,389056098930\,65... \\
\exp(3) &= 20,085536\,9231877... \\
\exp(4) &= 54,598150033144\,2... \\
\exp(5) &= 148,\ 413159102577... \\
\exp(6) &= 403,428793492735...
\end{aligned}$$

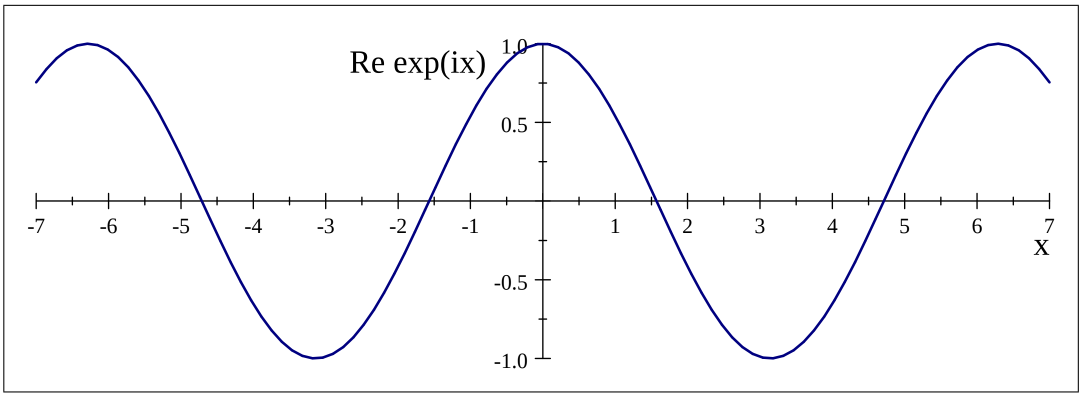
usw. Es folgt aus (6.1), dass die Funktion $\exp(x)$ für $x > 0$ positiv und monoton steigend ist. Später beweisen wir, dass diese Eigenschaften auch für alle $x \in \mathbb{R}$ gelten. Zum Beispiel, es gilt

$$\begin{aligned}
\exp(-1) &= 0,3678794411714 42... \\
\exp(-2) &= 0,135335283236\,613... \\
\exp(-3) &= 0,004978706836786...
\end{aligned}$$

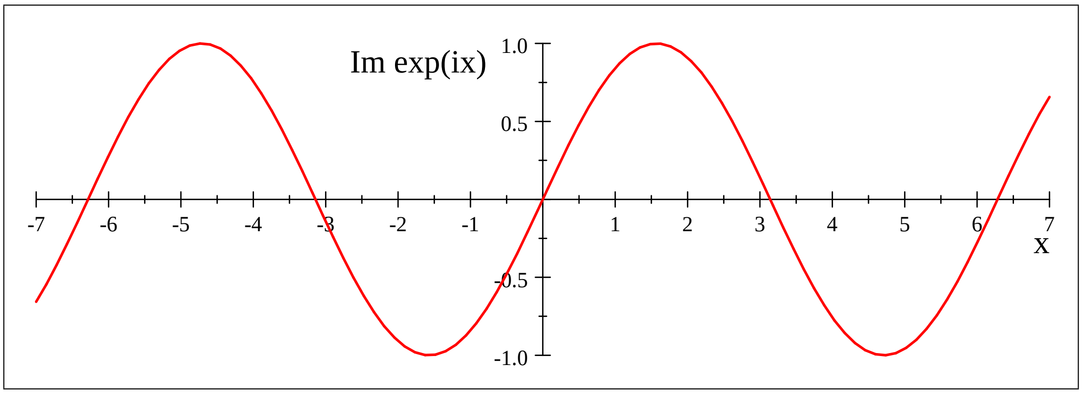
usw. Der Graph der Funktion $f(x) = \exp(x)$ für $x \in \mathbb{R}$ sieht wie folgt aus:



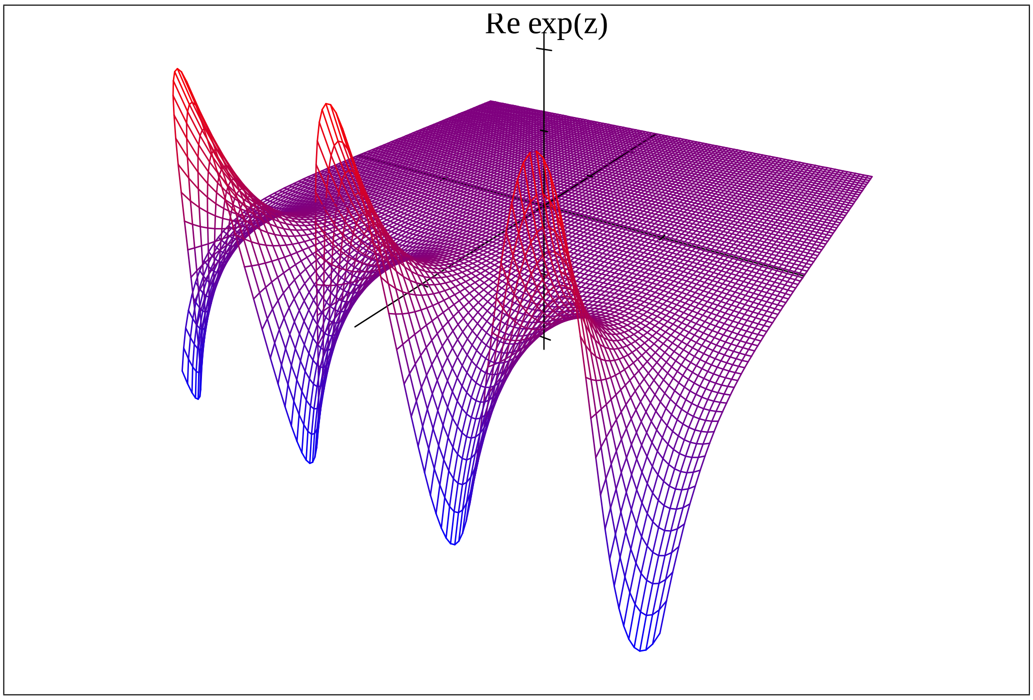
Betrachten wir jetzt die Funktion $f(x) = \exp(ix)$ für $x \in \mathbb{R}$, so dass das Argument der Exponentialfunktion imaginär ist. Der Graph des Realteiles $\operatorname{Re}\exp(ix)$ ist wie folgt:



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:



6.2 Äquivalente Definition der Exponentialfunktion

Satz 6.1 *Für alle $x \in \mathbb{C}$ gilt*

$$\boxed{\exp(x) = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n}. \tag{6.2}$$

Insbesondere erhalten wir die folgende äquivalente Definition der Zahl e:

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

(siehe auch die Aufgabe 84).

Beweis. Setzen wir

$$y_n = \left(1+\frac{x}{n}\right)^n$$

und

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

Da $z_n \to \exp(x)$, so reicht es zu beweisen, dass

$$z_n - y_n \to 0 \quad \text{für } n \to \infty.$$

Nach dem binomischen Lehrsatz (Satz 2.10) gilt

$$y_n = \left(1+\frac{x}{n}\right)^n = \sum_{k=0}^{n}\binom{n}{k}\left(\frac{x}{n}\right)^k = \sum_{k=0}^{n}\frac{n!}{k!\,(n-k)!}\frac{x^k}{n^k} = \sum_{k=0}^{n}\frac{n!}{n^k\,(n-k)!}\frac{x^k}{k!}, \tag{6.3}$$

woraus folgt, dass

$$z_n - y_n = \sum_{k=0}^{n} \left(1 - \frac{n!}{n^k (n-k)!}\right) \frac{x^k}{k!}.$$

Bemerken wir, dass

$$\frac{n!}{n^k (n-k)!} = \frac{1 \cdot 2 \cdot \ldots \cdot n}{n^k (1 \cdot 2 \cdot \ldots \cdot (n-k))} = \frac{\overbrace{(n-k+1) \cdot \ldots \cdot n}^{k \text{ mal}}}{\underbrace{n \cdot n \cdot \ldots \cdot n}_{k \text{ mal}}} = \prod_{j=n-k+1}^{n} \frac{j}{n}. \tag{6.4}$$

Da $\frac{j}{n} \leq 1$, so folgt es , dass

$$\frac{n!}{n^k (n-k)!} \leq 1. \tag{6.5}$$

Für alle $n > N \geq 1$ erhalten wir

$$\begin{aligned} |z_n - y_n| &\leq \sum_{k=0}^{n} \left|1 - \frac{n!}{n^k (n-k)!}\right| \frac{|x|^k}{k!} \\ &= \sum_{k=0}^{N} \left(1 - \frac{n!}{n^k (n-k)!}\right) \frac{|x|^k}{k!} + \sum_{k=N+1}^{n} \left(1 - \frac{n!}{n^k (n-k)!}\right) \frac{|x|^k}{k!} \\ &\leq \sum_{k=0}^{N} \left(1 - \frac{n!}{n^k (n-k)!}\right) \frac{|x|^k}{k!} + \sum_{k=N+1}^{\infty} \frac{|x|^k}{k!}. \end{aligned} \tag{6.6}$$

Die Summe $\sum_{k=N+1}^{\infty} \frac{|x|^k}{k!}$ in (6.6) ist die Restreihe der Exponentialreihe $\sum_{k=0}^{\infty} \frac{|x|^k}{k!}$ und somit konvergiert sie gegen 0 für $N \to \infty$. Fixieren wir ein $\varepsilon > 0$ und wählen ein $N \in \mathbb{N}$ mit

$$\sum_{k=N+1}^{\infty} \frac{|x|^k}{k!} < \varepsilon.$$

Bestimmen wir den Grenzwert der ersten Summe in (6.6) für $n \to \infty$ während N fixiert ist. Nach (6.4) haben wir

$$\frac{n!}{n^k (n-k)!} = \frac{n-k+1}{n} \frac{n-k+2}{n} \cdots \frac{n-1}{n} \frac{n}{n} = \prod_{l=1}^{k} \frac{n-l+1}{n}.$$

Da für jedes $l \leq k$ gilt

$$\lim_{n\to\infty} \frac{n-l+1}{n} = 1,$$

so erhalten wir

$$\lim_{n\to\infty} \frac{n!}{n^k (n-k)!} = \prod_{l=1}^{k} \lim_{n\to\infty} \frac{n-l+1}{n} = 1.$$

Es folgt, dass

$$\lim_{n\to\infty} \sum_{k=0}^{N} \left(1 - \frac{n!}{n^k (n-k)!}\right) \frac{|x|^k}{k!} = \sum_{k=0}^{N} \lim_{n\to\infty} \left(1 - \frac{n!}{n^k (n-k)!}\right) \frac{|x|^k}{k!} = 0.$$

Wir beschließen, dass die erste Summe in (6.6) für fast alle n von ε beschränkt ist. Da die zweite Summe in (6.6) auch von ε beschränkt ist (unabhängig von n), so ist die ganze Summe in (6.6) für fast alle n von 2ε beschränkt. Es folgt, dass $|z_n - y_n| < 2\varepsilon$ für fast alle n, was ergibt $|z_n - y_n| \to 0$. ∎

6.3 Eigenschaften der Exponentialfunktion

Die Haupteigenschaft der Exponentialfunktion ist wie folgt.

Satz 6.2 *Für alle $x, y \in \mathbb{C}$ gilt die Identität*

$$\boxed{\exp(x+y) = \exp(x)\exp(y)}. \tag{6.7}$$

Beweis. Wir benutzen die offensichtliche Folgerung aus (6.3) und (6.5): für $|x| < 1$ gilt

$$\left|\left(1+\frac{x}{n}\right)^n - 1\right| = \left|\sum_{k=1}^{n} \frac{n!}{n^k (n-k)!}\frac{x^k}{k!}\right| \leq \sum_{k=1}^{\infty} |x|^k = \frac{|x|}{1-|x|}. \tag{6.8}$$

Nach dem Satz 6.1 haben wir

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

Weiter benutzen wir die folgende Identität:

$$(1+a)(1+b) = 1+a+b+ab = (1+a+b)\left(1+\frac{ab}{1+a+b}\right). \tag{6.9}$$

Nach (6.9) mit $a = \frac{x}{n}$ und $b = \frac{y}{n}$ erhalten wir

$$\begin{aligned}\left(1+\frac{x}{n}\right)\left(1+\frac{y}{n}\right) &= \left(1+\frac{x}{n}+\frac{y}{n}\right)\left(1+\frac{\frac{x}{n}\frac{y}{n}}{1+\frac{x}{n}+\frac{y}{n}}\right) \\ &= \left(1+\frac{x+y}{n}\right)\left(1+\frac{1}{n}\frac{xy}{n+x+y}\right).\end{aligned}$$

Mit Hilfe von der Identität (6.2) erhalten wir

$$\begin{aligned}\exp(x)\exp(y) &= \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n\left(1+\frac{y}{n}\right)^n \\ &= \lim_{n\to\infty}\left(1+\frac{x+y}{n}\right)^n\left(1+\frac{1}{n}\frac{xy}{n+x+y}\right)^n \\ &= \exp(x+y)\lim_{n\to\infty}\left(1+\frac{1}{n}\frac{xy}{n+x+y}\right)^n.\end{aligned}$$

Es bleibt zu beweisen, dass

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\frac{xy}{n+x+y}\right)^n = 1. \tag{6.10}$$

Bemerken wir, dass für $n \to \infty$

$$\left|\frac{xy}{n+x+y}\right| \to 0,$$

insbesondere gilt $\left|\frac{xy}{n+x+y}\right| < 1$ für fast alle n. Nach (6.8) erhalten wir

$$\left|\left(1+\frac{1}{n}\frac{xy}{n+x+y}\right)^n - 1\right| \le \frac{\left|\frac{xy}{n+x+y}\right|}{1-\left|\frac{xy}{n+x+y}\right|} \to 0 \text{ für } n \to \infty,$$

woraus (6.10) folgt. ■

Satz 6.3 (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 (6.7) ergibt für $y = -x$,

$$\exp(x)\exp(-x) = \exp(0) = 1, \tag{6.11}$$

woraus $\exp(x) \neq 0$ folgt.

(b) Ist $x \in \mathbb{R}$ dann gilt $\exp(x) \in \mathbb{R}$ nach Definition

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{6.12}$$

Ist $x \geq 0$ so ist $\exp(x) > 0$ da alle Glieder in (6.12) nicht negative sind und das Glied mit $k = 0$ gleich 1 ist. Ist $x < 0$ dann $-x > 0$ und $\exp(-x) > 0$ was zusammen mit (6.11) ergibt

$$\exp(x) = \frac{1}{\exp(-x)} > 0.$$

(c) Wir haben

$$\exp(x) = \exp(x-y+y) = \exp(x-y)\exp(y) > \exp(y)$$

da für $t = x - y > 0$ gilt

$$\exp(t) = 1 + t + \frac{t^2}{2!} + ... > 1.$$

(d) Zunächst beweisen wir die Identität

$$\exp(k) = e^k \tag{6.13}$$

für alle $k \in \mathbb{N}$ per Induktion nach k. Für $k = 1$ gilt (6.13) nach Definition von e. Induktionsschritt: gilt (6.13) für ein $k \in \mathbb{N}$, dann erhalten wir nach (6.7)

$$\exp(k+1) = \exp(k)\exp(1) = e^k e = e^{k+1}.$$

Für $k = 0$ haben wir $\exp(0) = 1 = e^0$. Für negative $k \in \mathbb{Z}$ setzen wir $n = -k \in \mathbb{N}$ und erhalten nach (6.11)

$$\exp(k) = \frac{1}{\exp(n)} = \frac{1}{e^n} = e^{-n} = e^k.$$

■

Die Identität (6.13) motiviert die folgende Definition der Potenz e^x für alle komplexen Zahlen x:

$$\boxed{e^x := \exp(x)}.$$

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 $\sqrt{e}$. Analog erfüllt $e^{1/n}$ für jedes $n \in \mathbb{N}$ die Gleichung $\left(e^{1/n}\right)^n = e$, so dass man $e^{1/n}$ als Definition von $\sqrt[n]{e}$ annehmen kann.

Später definieren wir die Potenz a^x für alle reelle $a > 0$ and $x \in \mathbb{C}$. 22.06.18 H1

6.4 Hyperbelfunktionen

Definition. Die Hyperbelfunktionen *Kosinus hyperbolicus* und *Sinus hyperbolicus* werden für alle $x \in \mathbb{C}$ wie folgt definiert:

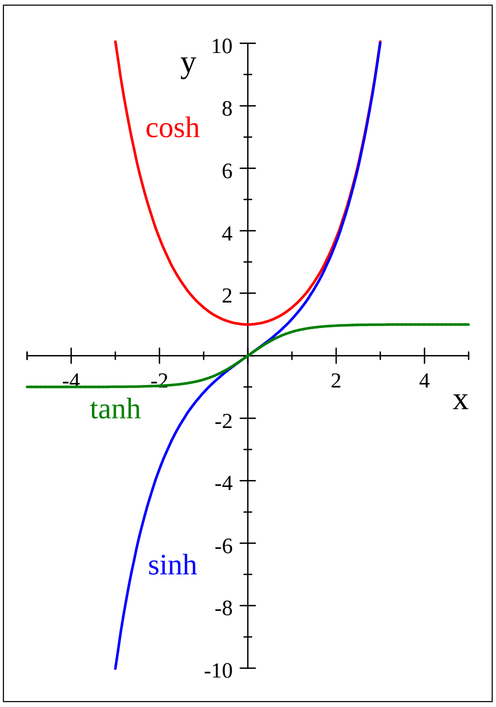
$$\boxed{\cosh x = \frac{e^x + e^{-x}}{2} \quad \text{und} \quad \sinh x = \frac{e^x - e^{-x}}{2}}.$$

Die Funktion $\cosh x$ ist offensichtlich gerade, d.h. $\cosh(-x) = \cosh x$, und $\sinh x$ ist ungerade, d.h. $\sinh(-x) = -\sinh x$. Für $x \in \mathbb{R}$ sind $\cosh x$ und $\sinh x$ offensichtlich reell.

Man definiert auch Tangens hyperbolicus und Kotangens hyperbolicus wie folgt:

$$\tanh x = \frac{\sinh x}{\cosh x} \quad \text{und} \quad \coth x = \frac{\cosh x}{\sinh x} = \frac{1}{\tanh x}.$$

Die Graphen von $\cosh x$, $\sinh x$ und $\tanh x$ für $x \in \mathbb{R}$ sehen wie folgt aus:



Weitere Eigenschaften sind in Aufgaben 109, 113, 114.

6.5 Trigonometrische Funktionen

Definition. Die trigonometrische Funktionen *Kosinus* und *Sinus* werden für alle $x \in \mathbb{C}$ wie folgt definiert:

$$\boxed{\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \text{und} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}}, \tag{6.14}$$

wobei i die imaginäre Einheit ist.

Bemerken wir, dass sogar für reelle x benutzt die Definition (6.14) von $\sin x$ und $\cos x$ die Werte von $\exp(z)$ im komplexen Bereich.

Auch definieren with Tangens und Kotangens mit

$$\tan x = \frac{\sin x}{\cos x} \quad \text{und} \quad \cot x = \frac{\cos x}{\sin x}$$

in den Bereichen, wo die Quotienten wohldefiniert sind.

Es folgt direkt nach (6.14), dass für alle $x \in \mathbb{C}$ gilt

$$\boxed{e^{ix} = \cos x + i \sin x}. \tag{6.15}$$

Diese Identität heißt *Eulerformel*.

Es folgt aus (6.14), dass $\cos x$ eine gerade Funktion ist, d.h. für alle $x \in \mathbb{C}$

$$\cos(-x) = \cos x,$$

und $\sin x$ eine ungerade Funktion ist, d.h. für alle $x \in \mathbb{C}$

$$\sin(-x) = -\sin x.$$

Satz 6.4 *Für alle $x \in \mathbb{C}$ gelten die folgenden Identitäten:*

$$\boxed{\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + ...} \tag{6.16}$$

und

$$\boxed{\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + ...,} \tag{6.17}$$

wobei die beiden Reihen absolut konvergent sind.

Beweis erfolgt nach Einsetzen von der Exponentialreihe in (6.14) (siehe Aufgabe 108). Es folgt aus (6.17) und (6.16), dass $\sin x$ und $\cos x$ reell sind für alle $x \in \mathbb{R}$. Die Identitäten (6.16) und (6.17) lassen sich für numerische Berechnung von $\cos x$ bzw $\sin z$ anwenden. Zum Beispiel,

$$\cos 1 \approx 1 - \frac{1}{2!} + \frac{1}{4!} - \frac{1}{6!} + \frac{1}{8!} - \frac{1}{10!} = 0,54030230...$$

und

$$\sin \frac{1}{2} \approx \frac{1}{2} - \frac{(1/2)^3}{3!} + \frac{(1/2)^5}{5!} - \frac{(1/2)^7}{7!} + \frac{(1/2)^9}{9!} = 0,4794255386...$$

Weitere Eigenschaften sind in Aufgabe 108.

6.6 * Alternativer Beweis der Haupteigenschaft

Wir geben hier einen anderen Beweis des Satzes 6.2 an: für alle $x, y \in \mathbb{C}$ gilt die Identität

$$\exp(x+y) = \exp(x)\exp(y). \tag{6.18}$$

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 (siehe (5.22) und Satz 5.14), 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 auch den Binomischen Lehrsatz benutzt haben. Es folgt, dass

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

Nach dem Satz 5.14 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 (6.18) folgt. ■

Chapter 7

Stetige Funktionen einer reellen Variablen

In diesem Kapitel entwickeln wir Methoden für Untersuchung von reellwertigen Funktionen einer reellen Variablen. Einige Beispiele von solchen Funktionen sind Potenzfunktion $f(x) = x^n$ mit $n \in \mathbb{Z}$, Quadratwurzel $f(x) = \sqrt{x}$, Exponentialfunktion $\exp(x)$, trigonometrische Funktionen und Hyperbelfunktionen.

Wir fangen mit der Definition des Grenzwert einer Funktion an.

7.1 Grenzwert einer Funktion

Erinnern wir uns an den Begriff der Umgebung: eine Umgebung von $a \in \mathbb{R}$ ist irgendwelches Intervall $U_\varepsilon(a) = (a-\varepsilon, a+\varepsilon)$ mit $\varepsilon > 0$, eine Umgebung von $+\infty$ ist irgendwelches Intervall $U_E(+\infty) = (E, +\infty]$ mit $E \in \mathbb{R}$, und eine Umgebung von $-\infty$ ist irgendwelches Intervall $U_E(-\infty) = [-\infty, E)$ mit $E \in \mathbb{R}$.

Definition. Ist $U(a)$ eine Umgebung von $a \in \overline{\mathbb{R}}$, so definieren wir eine *punktierte* Umgebung von $a \in \overline{\mathbb{R}}$ mit

$$\dot{U}(a) = U(a) \setminus \{a\}.$$

Zum Beispiel, für $a \in \mathbb{R}$ gilt

$$\dot{U}_\varepsilon(a) = (a-\varepsilon, a+\varepsilon) \setminus \{a\} = (a-\varepsilon, a) \cup (a, a+\varepsilon),$$

d.h.

$$x \in \dot{U}_\varepsilon(a) \Leftrightarrow 0 < |x-a| < \varepsilon.$$

Für $a = \pm\infty$ gilt

$$\dot{U}_E(+\infty) = (E, +\infty) \quad \text{und} \quad \dot{U}_E(-\infty) = (-\infty, E).$$

Definition. Sei J ein beliebiges Intervall mit den Grenzen $\alpha, \beta \in \overline{\mathbb{R}}$, $\alpha < \beta$. Der *Abschluss* $\overline{J}$ von J ist das Intervall $\overline{J} = [\alpha, \beta]$.

Zum Beispiel, $\overline{(0,1)} = [0,1]$, $\overline{[0,+\infty)} = [0,+\infty]$, $\overline{(-\infty,+\infty)} = [-\infty,+\infty]$ usw.

Definition. Sei $J \subset \mathbb{R}$ ein Intervall und $f : J \to \mathbb{R}$ eine Funktion auf J. Seien $a \in \overline{J}$ and $b \in \overline{\mathbb{R}}$. Die Funktion $f(x)$ hat für $x \to a$ den Grenzwert (Limes) $b \in \overline{\mathbb{R}}$ falls es für jede Umgebung $V(b)$ von b eine Umgebung $U(a)$ von a gibt so dass

$$\forall x \in J \cap \dot{U}(a) \text{ gilt } f(x) \in V(b).$$

Man schreibt in diesem Fall

$$\lim_{x \to a} f(x) = b \tag{7.1}$$

oder

$$f(x) \to b \text{ für } x \to a.$$

Manchmal schreibt man auch

$$\lim_{\substack{x \to a \\ x \in J}} f(x) = b,$$

um den Definitionsbereich von f zu betonen.

Man sagt im Fall $b \in \mathbb{R}$: $f(x)$ konvergiert gegen b für x gegen a, und im Fall $b = \pm\infty$: $f(x)$ divergiert gegen b für x gegen a.

Diese Definition lässt sich kurz wie folgt umformulieren: $\lim\limits_{x \to a} f(x) = b$ falls

$$\boxed{\forall V(b) \ \ \exists U(a) \ \ \forall x \in J \cap \dot{U}(a) \text{ gilt } f(x) \in V(b).} \tag{7.2}$$

Die Bedingung (7.2) bedeutet intuitiv folgendes:

$$x \text{ ist in der Nähe von } a \Rightarrow f(x) \text{ ist in der Nähe von } b.$$

Im Fall $a, b \in \mathbb{R}$ ist die Bedingung (7.2) äquivalent zu folgendes:

$$\boxed{\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in J \text{ mit } 0 < |x - a| < \delta \text{ gilt } |f(x) - b| < \varepsilon.} \tag{7.3}$$

Für die unendlichen Werte von a bzw b erhalten wir analog folgendes:

Bemerkung.

- $a \in \mathbb{R}$ und $b = +\infty$: $\lim_{x \to a} f(x) = +\infty$ falls

$$\boxed{\forall E \in \mathbb{R} \ \ \exists \delta > 0 \ \ \forall x \in J \ \text{ mit } 0 < |x - a| < \varepsilon \ \text{ gilt } f(x) > E.}$$

- $a = +\infty$ und $b \in \mathbb{R}$: $\lim_{x \to +\infty} f(x) = b$ falls

$$\boxed{\forall \varepsilon > 0 \ \ \exists D \in \mathbb{R} \ \ \forall x \in J \ \text{ mit } x > D \ \text{ gilt } |f(x) - b| < \varepsilon.} \tag{7.4}$$

- $a = b = +\infty$: $\lim_{x \to +\infty} f(x) = +\infty$ falls

$$\boxed{\forall E \in \mathbb{R} \ \ \exists D \in \mathbb{R} \ \ \forall x \in J \text{ mit } x > D \text{ gilt } f(x) > E.}$$

Der Fall $a = -\infty$ bzw $b = -\infty$ wird analog betrachtet.

Es folgt aus (7.3) und (7.4), dass im Fall $b \in \mathbb{R}$ die folgende Äquivalenz gilt:

$$\lim_{x \to a} f(x) = b \Leftrightarrow \lim_{x \to a} |f(x) - b| = 0,$$

(cf. Satz 4.1(d)).

Satz 7.1 *Sei f eine reellwertige Funktion auf einem Intervall $J \subset \mathbb{R}$. Seien $a \in \overline{J}$ und $b \in \overline{\mathbb{R}}$. Dann die folgenden zwei Bedingungen sind äquivalent:*

(i) $\lim_{x \to a} f(x) = b$.

(ii) *Für jede Folge $\{x_n\}_{n \in \mathbb{N}}$ aus $J \setminus \{a\}$ mit $\lim_{n \to \infty} x_n = a$ gilt $\lim_{n \to \infty} f(x_n) = b$.*

Kurz kann man die Bedingung (ii) so beschreiben:

$$x_n \to a \ \Rightarrow f(x_n) \to b.$$

Beweis. Betrachten wir den Fall wenn $a, b \in \mathbb{R}$.

(i) $\Rightarrow$ (ii) Die Bedingung (i) bedeutet nach Definition, dass

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \ \forall x \in J \text{ mit } \ 0 < |x - a| < \delta \text{ gilt } \ |f(x) - b| < \varepsilon. \tag{7.5}$$

Sei $\{x_n\}$ eine Folge aus $J \setminus \{a\}$ mit $x_n \to a$. Dann gilt für fast alle $n \in \mathbb{N}$, dass

$$0 < |x_n - a| < \delta,$$

woraus folgt, dass für fast alle n

$$|f(x_n) - b| < \varepsilon.$$

Da $\varepsilon > 0$ beliebig 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.

$$\exists \varepsilon > 0 \ \ \forall \delta > 0 \ \ \exists x \in J \text{ mit } 0 < |x - a| < \delta \text{ sodass } \ |f(x) - b| \geq \varepsilon.$$

Benutzen wir diese Bedingung für die Werte $\delta = \frac{1}{k}$ wobei $k \in \mathbb{N}$. Für jedes $\delta = \frac{1}{k}$ erhalten wir ein $x_k \in J$ mit

$$0 < |x_k - a| < \frac{1}{k} \quad \text{und} \quad |f(x_k) - b| \geq \varepsilon.$$

Dann gilt $x_k \to a$ für $k \to \infty$, aber nicht $f(x_k) \to b$, was im Widerspruch zur Bedingung (ii) steht.

Analog beweist man die Äquivalenz (i) $\Leftrightarrow$ (ii) im Fall von unendlichen Werten von a, b. ■

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 7.2 *Seien f, g reellwertige Funktionen auf einem Intervall $J \subset \mathbb{R}$ und sei $a \in \overline{J}$. Angenommen seien die Bedingungen*

$$\lim_{x \to a} f(x) = b \text{ und } \lim_{x \to a} g(x) = c,$$

wobei $b, c \in \overline{\mathbb{R}}$.

(a) Es gelten die Identitäten

$$\lim_{x \to a} (f \pm g) = b \pm c, \quad \lim_{x \to a} (fg) = bc, \quad \lim_{x \to a} \frac{f}{g} = \frac{b}{c} \tag{7.6}$$

vorausgesetzt, dass die Ausdrücke $b + c$, bc, $\frac{b}{c}$ 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 $b \leq c$.

(c) Seien f, g, h drei Funktionen auf J mit $f \leq h \leq g$ und

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = b.$$

Dann gilt auch

$$\lim_{x \to a} h(x) = b.$$

Beweis. Nach dem Satz 7.1 gelten für jede Folge $x_n \to x$, $x_n \in J \setminus \{a\}$ die Bedingungen $f(x_n) \to b$ und $g(x_n) \to c$. Nach dem Satz 4.4 erhalten wir

$$(f \pm g)(x_n) \to b \pm c, \qquad (fg)(x_n) \to bc, \qquad \frac{f}{g}(x_n) \to \frac{b}{c},$$

woraus (7.6) folgt. Analog beweist man (b) und (c) mit Hilfe von dem Satz 4.1. ■

Beispiel. 1. Für die Funktion $f(x) = x$, $x \in \mathbb{R}$, gilt offensichtlich $f(x) \to a$ für $x \to a$ für jedes $a \in \overline{\mathbb{R}}$, da

$$x_n \to a \Rightarrow f(x_n) = x_n \to a.$$

2. Für die Funktion $f(x) = x^2$, $x \in \mathbb{R}$, gilt $f(x) \to a^2$ für $x \to a \in \overline{\mathbb{R}}$ da

$$\lim_{x \to a} f(x) = \lim_{x \to a} x \cdot x = \lim_{x \to a} x \lim_{x \to a} x = a^2.$$

Gleiches gilt für $f(x) = x^n$ mit $n \in \mathbb{N}$:

$$\lim_{x \to a} x^n = a^n.$$

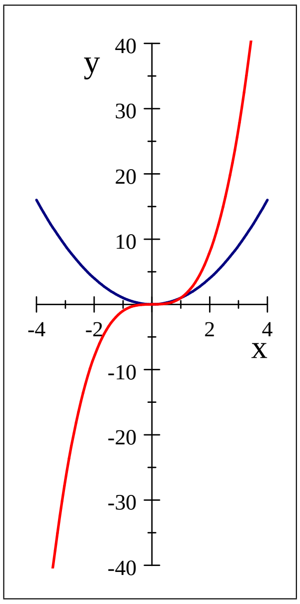
Insbesondere haben wir

$$\lim_{x \to +\infty} x^n = (+\infty)^n = +\infty$$

und

$$\lim_{x \to -\infty} x^n = (-\infty)^n = \begin{cases} +\infty, & n \text{ gerade}, \\ -\infty, & n \text{ ungerade}. \end{cases}$$

Hier sind die Graphen der Funktionen $f(x) = x^2$ und $f(x) = x^3$ wo der Unterschied im $\lim_{x \to -\infty} x^n$ klar ist.



3. Betrachten wir die Funktion $f(x) = x^{-1}$ im Definitionsbereich $J = (0, +\infty)$. Für jedes $a \in (0, +\infty]$ gilt

$$\lim_{x \to a} f(x) = \lim_{x \to a} \frac{1}{x} = \frac{1}{\lim_{x \to a} x} = \frac{1}{a}.$$

Insbesondere haben wir

$$\lim_{x \to +\infty} f(x) = \frac{1}{+\infty} = 0.$$

Für $a = 0$ behaupten wir, dass

$$\lim_{x \to 0} f(x) = +\infty,$$

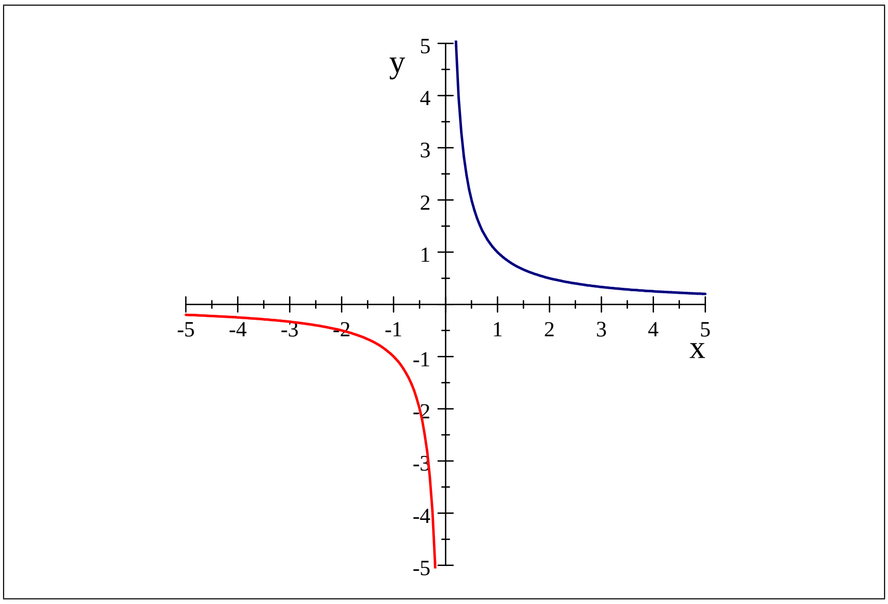
da für jede Folge $\{x_k\}$ aus $(0, +\infty)$ mit $x_k \to 0$ gilt $\frac{1}{x_k} \to +\infty$ (Aufgabe 79) und somit auch $f(x) \to +\infty$.

Für die gleiche Funktion $f(x) = x^{-1}$ auf $J = (-\infty, 0)$ erhalten wir

$$\lim_{x \to 0} f(x) = -\infty$$

da für jede Folge $\{x_k\}$ aus $(-\infty, 0)$ mit $x_k \to 0$ gilt $\frac{1}{x_k} \to -\infty$ und somit $f(x) \to -\infty$.

Hier sind die Graphen de Funktion $f(x) = x^{-1}$ auf $(0, +\infty)$ und $(-\infty, 0)$ wo der Unterschied im $\lim_{x \to 0} \frac{1}{x}$ klar ist.



Beispiel. Betrachten wir die Funktion $\exp(x)$ im Definitionsbereich $\mathbb{R}$. Beweisen wir, dass

$$\exp(x) \to 1 \text{ für } x \to 0. \tag{7.7}$$

Wir haben für alle $|x| < 1$:

$$|\exp(x) - 1| = \left| \sum_{k=1}^{\infty} \frac{x^k}{k!} \right| \le \sum_{k=1}^{\infty} |x|^k = \frac{|x|}{1 - |x|}.$$

Da $\lim_{x\to 0} |x| = 0$, so erhalten wir

$$\lim_{x\to 0} \frac{|x|}{1 - |x|} = 0$$

und somit, nach dem Satz 7.2(c), $|\exp(x) - 1| \to 0$, woraus (7.7) folgt.

Wir haben auch

$$\exp(x) \to +\infty \text{ für } x \to +\infty, \tag{7.8}$$

da für $x > 0$ gilt

$$\exp(x) = 1 + x + \frac{x^2}{2} + ... > x.$$

Beweisen wir, dass

$$\exp(x) \to 0 \text{ für } x \to -\infty. \tag{7.9}$$

Für jede Folge $x_n \to -\infty$ haben wir $y_n = -x_n \to +\infty$ und somit

$$\exp(x_n) = \frac{1}{\exp(-x_n)} = \frac{1}{\exp(y_n)} \to \frac{1}{+\infty} = 0,$$

woraus (7.9) folgt.

7.2 Stetige Funktionen

27.06.18 H13

Definition. Sei f eine reellwertige Funktion auf einem Intervall $J \subset \mathbb{R}$. Die Funktion f heißt *stetig* an der Stelle $a \in J$ falls

$$\lim_{x \to a} f(x) = f(a). \tag{7.10}$$

Sonst heißt f *unstetig* an a. Ist f stetig an alle $a \in J$, so heißt f stetig auf J (oder einfach stetig).

Satz 7.3 *Eine Funktion $f : J \to \mathbb{R}$ ist stetig an $a \in J$ genau dann, wenn für jede Folge $\{x_n\}$ aus J mit $x_n \to a$ gilt $f(x_n) \to f(a)$.*

Beweis. Nach dem Satz 7.1, die Bedingung (7.10) ist äquivalent zu folgendes: für jede Folge $\{x_n\}$ aus $J \setminus \{a\}$ gilt

$$x_n \to a \Rightarrow f(x_n) \to f(a). \tag{7.11}$$

Die Differenz zwischen den Bedingungen der Sätze 7.1 und 7.3 ist dass im ersten Fall die Folge $\{x_n\}$ aus $J \setminus \{a\}$ ist und im zweiten Fall – aus J. So, wir müssen nur beweisen, dass die Gültigkeit von (7.11) für die Folgen aus $J \setminus \{a\}$ das Gleiche für die Folgen aus J ergibt.

Angenommen (7.10), sei $\{x_n\}$ eine Folge aus J mit $x_n \to a$. Beweisen wir, dass $f(x_n) \to f(a)$. Dafür betrachten wir die Menge

$$S = \{n \in \mathbb{N} : x_n = a\}.$$

Ist S endlich, so liegt die Folge $\{x_n\}_{n=N}^{\infty}$ in $J \setminus \{a\}$ für einen $N \in \mathbb{N}$, woraus $f(x_n) \to f(a)$ folgt nach Voraussetzung.

Ist $S^c = \mathbb{N} \setminus S$ endlich, so gilt $x_n = a$ für fast alle n, woraus folgt $f(x_n) = f(a)$ für fast alle n und somit $f(x_n) \to f(a)$.

Seien S und S^c unendlich. So erhalten wir zwei Folgen $\{x_n\}_{n \in S}$ und $\{x_n\}_{n \in S^c}$. Die Folge $\{f(x_n)\}_{n \in S}$ hat den Grenzwert $f(a)$, da alle Glieder davon gleich $f(a)$ sind. Die Folge $\{f(x_n)\}_{n \in S^c}$ hat auch den Grenzwert a, da alle x_n aus dieser Folge in $J \setminus \{a\}$ liegen. Somit hat auch die ganze Folge $\{f(x_n)\}_{n \in \mathbb{N}}$ den Grenzwert $f(a)$.
■

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, d.h. für alle $a \in \mathbb{R}$ gilt

$$\lim_{x \to a} \exp(x) = \exp(a). \tag{7.12}$$

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

$$\exp(x) = \exp(x - a + a) = \exp(a) \exp(x - a).$$

Für jede Folge $x_n \to a$ erhalten wir $x_n - a \to 0$ und somit

$$\lim_{n\to\infty} \exp(x_n - a) = 1,$$

woraus folgt

$$\lim_{n\to\infty} \exp(x_n) = \exp(a) \lim_{n\to\infty} \exp(x_n - a) = \exp(a)$$

und somit (7.12).

Satz 7.4 *Seien f, g zwei reellwertige Funktionen auf einem Intervall $J \subset \mathbb{R}$. Sind f und g stetig an der Stelle $a \in J$, so sind auch die Funktionen $f+g$, $f-g$, fg, f/g stetig an a (im Fall f/g vorausgesetzt $g \neq 0$).*

Sind f, g stetig auf J, so sind $f+g$, $f-g$, fg, f/g auch stetig auf J (im Fall f/g vorausgesetzt $g \neq 0$).

Beweis. Die erste Aussage folgt aus dem Satz 7.2. Zum Beispiel für die Summe $f+g$ erhalten wir

$$\lim_{x\to a} (f+g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = f(a) + g(a) = (f+g)(a),$$

woraus die Stetigkeit von $f+g$ an a folgt. Analog behandelt man die Funktionen $f-g$, fg und f/g.

Die zweite Aussage folgt aus der ersten Aussage. ■

Beispiel. Da die Funktionen $f(x) = x$ und $g(x) = \text{const}$ stetig auf $\mathbb{R}$ sind, so folgt es, dass jede 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*

$$P(x) = \sum_{k=0}^{n} c_k x^k = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n$$

stetig auf $\mathbb{R}$ ist (wobei $c_k \in \mathbb{R}$ und $n \in \mathbb{N} \cup \{0\}$).

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 in jedem Intervall J wo $Q \neq 0$.

7.3 Zusammengesetzte Funktion

Satz 7.5 (Grenzwert der Komposition) *Seien $f : A \to \mathbb{R}$ und $g : B \to \mathbb{R}$ zwei Funktionen die auf den Intervallen A und B definiert sind. Nehmen wir an, dass für einige $a \in \overline{A}$, $b \in \overline{B}$, $c \in \overline{\mathbb{R}}$ gelten*

$$\lim_{x\to a} f(x) = b \quad \textit{und} \quad \lim_{y\to b} g(y) = c.$$

Im Fall $b \in B$ nehmen wir auch an, dass g an b stetig ist, d.h. $g(b) = c$. Ist die zusammengesetzte Funktion $x \mapsto g(f(x))$ auf A definiert (d.h. $f(A) \subset B$), so gilt

$$\lim_{x\to a} g(f(x)) = c. \tag{7.13}$$

Man kann die Identität (7.13) wie folgt umschreiben:

$$\lim_{x\to a} g(f(x)) = \lim_{y\to b} g(y)$$

und betrachten sie als einen Wechsel von Variable $y = f(x)$ im Limes.

Beweis. Sei x_n eine Folge aus $A \setminus \{a\}$ mit $x_n \to a$. Wir müssen beweisen, dass

$$g(f(x_n)) \to c.$$

Betrachten wir die Folge $y_n = f(x_n)$. Nach Voraussetzung gilt $y_n = f(x_n) \in B$ und

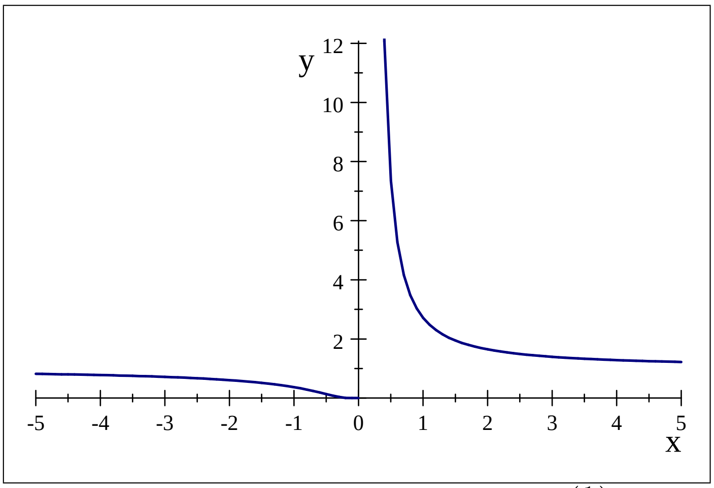
$$y_n = f(x_n) \to b.$$

Beweisen wir, dass

$$g(y_n) \to c.$$

Ist $b \notin B$, so liegt die Folge $\{y_n\}$ in $B \setminus \{b\}$, und $g(y_n) \to c$ gilt nach Voraussetzung. Ist $b \in B$, so ist g stetig an b, und $g(y_n) \to g(b) = c$ gilt nach dem Satz 7.3. ■

Beispiel. Betrachten wir die Funktion $f(x) = \exp\left(\frac{1}{x}\right)$ und bestimmen die Grenzwerte davon für $x \to 0$ in zwei Intervallen: $x \in (0, +\infty)$ und $x \in (-\infty, 0)$.



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

$$\lim_{\substack{x\to 0\\ x>0}} \exp\left(\frac{1}{x}\right) = \lim_{y\to+\infty} \exp(y) = +\infty \tag{7.14}$$

(siehe (7.8)). Da

$$\lim_{\substack{x\to 0\\ x<0}} \frac{1}{x} = -\infty,$$

so erhalten wir

$$\lim_{\substack{x\to 0\\ x<0}} \exp\left(\frac{1}{x}\right) = \lim_{y\to-\infty} \exp(y) = 0$$

(siehe (7.9)).

Satz 7.6 *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. Da $f(x) \to b$ für $x \to a$ und $g(y) \to g(b)$ für $y \to b$, so erhalten wir nach dem Satz 7.5

$$\lim_{x \to a} g(f(x)) = \lim_{y \to b} g(y) = g(b) = g(f(a)),$$

woraus die Stetigkeit von $g(f(x))$ an a folgt.

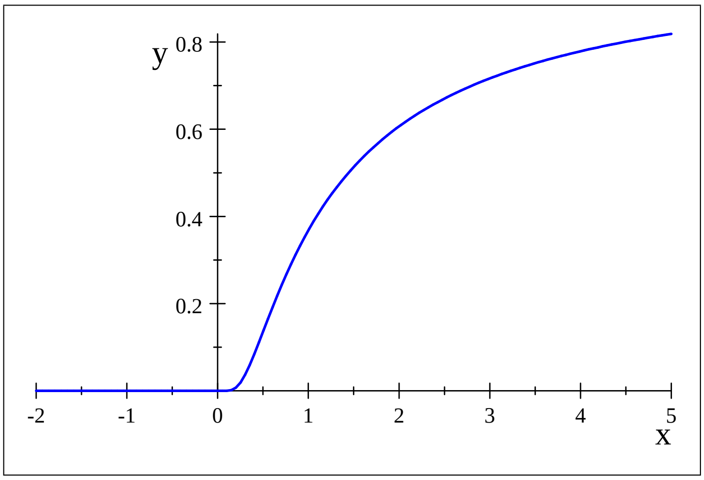
Ist f stetig auf A und g – auf B, so ist g stetig an $b = f(a)$ für jedes $a \in A$, da $f(A) \subset B$. Daraus folgt, dass $g \circ f$ stetig an a für jedes $a \in A$ ist, und somit stetig auf A. ■

Beispiel. Für jede rationale Funktion $R(x)$ sind die Funktionen $\exp(R(x))$ und $R(\exp(x))$ stetig im Definitionsbereich.

Beispiel. Beweisen wir, dass die Funktion

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x}\right), & x > 0, \\ 0, & x \leq 0, \end{cases} \tag{7.15}$$

stetig auf $\mathbb{R}$ ist.



Der Graph der Funktion (7.15)

Die Funktion $f(x) = \exp\left(-\frac{1}{x}\right)$ ist stetig an alle Stellen $a > 0$ nach dem Satz 7.6, und $f(x) = 0$ ist offensichtlich stetig an alle Stellen $a < 0$. Es bleibt zu zeigen, dass f stetig an 0 ist, d.h.

$$\lim_{x \to 0} f(x) = 0.$$

Dafür reicht es zu zeigen, dass für beliebige Folge $\{x_n\}$ aus $\mathbb{R}$ mit $x_n \to 0$ gilt

$$f(x_n) \to 0 \quad \text{für } n \to \infty. \tag{7.16}$$

Ist diese Folge positiv, d.h. $x_n > 0$ für alle n, so gilt nach (7.14) dass

$$\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} \exp\left(-\frac{1}{x_n}\right) = \lim_{n \to \infty} \frac{1}{\exp\left(\frac{1}{x_n}\right)} = \frac{1}{+\infty} = 0.$$

Ist diese Folge nichtpositiv, d.h. $x_n \leq 0$ für alle n, so erhalten wir nach (7.15)

$$\lim_{n\to\infty} f(x_n) = \lim 0 = 0.$$

Für allgemeine Folge $\{x_n\}$ mit $x_n \to 0$ betrachten wir die Menge

$$S = \{n \in \mathbb{N} : x_n \leq 0\}$$

und das Komplement $S^c = \mathbb{N} \setminus S$. Ist S oder S^c endlich, so liegt die Folge $\{x_n\}_{n=N}^{\infty}$ für ein $N \in \mathbb{N}$ entweder in $(0, +\infty)$ oder in $(-\infty, 0]$, so dass wir (7.16) wie oberhalb erhalten. Sind S und S^c unendlich, so gilt nach den obigen Fällen

$$\lim_{n\in S} f(x_n) = 0 = \lim_{n\in S^c} f(x_n),$$

woraus (7.16) folgt.

Beispiel. Die Funktion

$$f(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

ist unstetig an $a = 0$, da für die Folge $x_n = \frac{1}{n} \to 0$ gilt $f(x_n) \to 1 \neq f(0)$.

7.4 Zwischenwertsatz

Hauptsatz 7.7 *(*Zwischenwertsatz*) Sei $f(x)$ eine stetige Funktion auf einem geschlossenen Intervall $[a, b]$, wobei $a, b \in \mathbb{R}$, $a < b$. 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 $\{[a_n, b_n]\}_{n=1}^{\infty}$, $a_n \leq b_n$, mit den Eigenschaften:

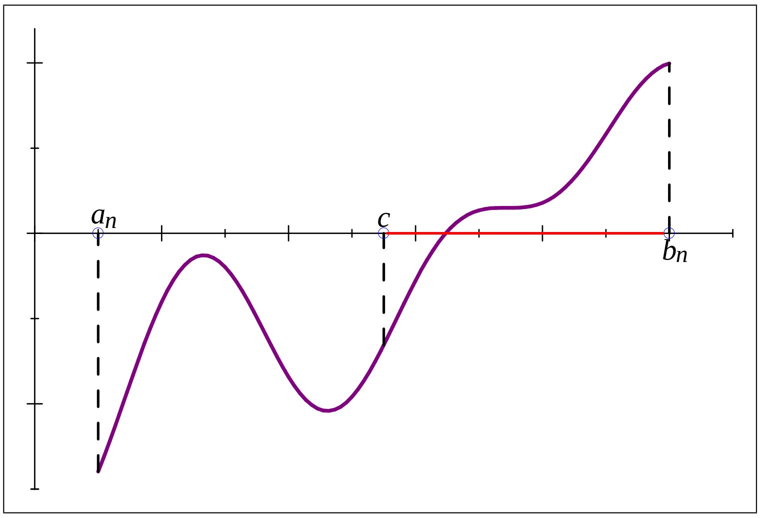
1. $[a_1, b_1] = [a, b]$;

2. $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ für alle $n \in \mathbb{N}$;

3. $b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n)$ für alle $n \in \mathbb{N}$;

4. $f(a_n) \leq 0$, $f(b_n) \geq 0$ für alle $n \in \mathbb{N}$.

Induktionsanfang ist offensichtlich: $[a_1, b_1] = [a, b]$. Induktionsschritt von n nach $n + 1$. Ist $[a_n, b_n]$ schon bekannt, so betrachten wir den Mittelpunkt $c = \frac{a_n+b_n}{2}$. Gilt $f(c) \leq 0$, so setzen wir $[a_{n+1}, b_{n+1}] = [c, b_n]$. Gilt $f(c) > 0$ so setzen wir $[a_{n+1}, b_{n+1}] = [a_n, c]$. In den beiden Fällen erhalten wir alle Eigenschaften 2-4.



Somit ist $\{[a_n, b_n]\}_{n=1}^{\infty}$ eine Intervallschachtelung. Nach den Intervallschachtelungsprinzip gibt es ein

$$x \in \bigcap_{n \in \mathbb{N}} [a_n, b_n],$$

d.h. $a_n \leq x \leq b_n$ für alle n. Da $b_n - a_n \to 0$, so haben wir auch $a_n \to x$ und $b_n \to x$ für $n \to \infty$. Nach der Stetigkeit der Funktion f gilt

$$\lim_{n\to\infty} f(a_n) = f(x) = \lim_{n\to\infty} f(b_n).$$

Da $f(a_n) \leq 0$ und $f(b_n) \geq 0$, es folgt, da $f(x) \leq 0$ und $f(x) \geq 0$, somit $f(x) = 0$. Es bleibt nur zu bemerken, dass $x \in (a, b)$ da $f(a)$ und $f(b)$ nicht Null sind. ∎ 29.06.18 H1

Beispiel. Sei $P(x)$ ein Polynom von Grad n mit reellen Koeffizienten der Form

$$P(x) = x^n + c_1 x^{n-1} + \ldots + c_n,$$

wobei $c_k \in \mathbb{R}$ und $n \in \mathbb{N}$. Beweisen wir die folgende Aussage: ist n ungerade, so existiert eine reelle Nullstelle von P, d.h. ein $x \in \mathbb{R}$ mit $P(x) = 0$. Bemerken wir zuerst, dass

$$\lim_{x\to+\infty} P(x) = \lim_{x\to+\infty} x^n \left(1 + \frac{c_1}{x} + \ldots + \frac{c_n}{x^n}\right) = (+\infty)^n \cdot 1 = +\infty$$

und

$$\lim_{x\to-\infty} P(x) = \lim_{x\to-\infty} x^n \left(1 + \frac{c_1}{x} + \ldots + \frac{c_n}{x^n}\right) = (-\infty)^n \cdot 1 = -\infty.$$

Somit existiert ein $a \in \mathbb{R}$ mit $P(a) < 0$ und ein $b \in \mathbb{R}$ mit $P(b) > 0$. Da P stetig ist, so existiert nach dem Satz 7.7 ein $x \in \mathbb{R}$ mit $P(x) = 0$.

Mit Hilfe des Satzes 7.7 beschreiben wir für jede stetige Funktion f auf J die Bildmenge

$$f(J) = \{f(x) : x \in J\}.$$

Setzen wir

$$\sup f = \sup_J f := \sup f(J)$$

und

$$\inf f = \inf_J f := \inf f(J).$$

Satz 7.8 *Sei $f : J \to \mathbb{R}$ eine stetige Funktion auf einem Intervall $J \subset \mathbb{R}$. Dann ist das Bild $f(J)$ ein Intervall und zwar mit den Grenzen* $\inf f$ *und* $\sup f$.

Man kann kurz sagen: stetiges Bild eines Intervalles ist ein Intervall.

Beweis. Nach Definition haben wir

$$f(J) \subset [\inf f, \sup f].$$

Wir werden beweisen, dass

$$f(J) \supset (\inf f, \sup f), \tag{7.17}$$

was implizieren wird, dass $f(J)$ ein Intervall mit den Grenzen $\inf f$, $\sup f$ ist.

Um die Inklusion (7.17) zu beweisen, zeigen wir, dass jedes $t \in (\inf f, \sup f)$ in $f(J)$ liegt, d.h. $t = f(x)$ für ein $x \in J$. Da $t > \inf f(J)$ so ist t keine untere Schranke von $f(J)$. Somit existiert ein $\alpha \in f(J)$ mit $\alpha < t$. Da $t < \sup f(J)$, so ist t keine obere Schranke von $f(J)$, und so existiert ein $\beta \in f(J)$ mit $\beta > t$. Da α und β Elemente von $f(J)$ sind, so existieren Zahlen $a, b \in J$ mit $f(a) = \alpha$ und $f(b) = \beta$, so dass

$$f(a) < t < f(b).$$

Nehmen wir an, dass $a < b$ (der Fall $a > b$ ist analog). Betrachten wir auf dem Intervall $[a, b] \subset J$ die Funktion

$$g(x) = f(x) - t,$$

die stetig in J ist und erfüllt

$$g(a) = f(a) - t < 0 \quad \text{und} \quad g(b) = f(b) - t > 0.$$

Nach dem Satz 7.7 existiert ein $x \in (a, b)$ mit $g(x) = 0$ woraus $f(x) = t$ folgt. ■

Beispiel. Beweisen wir, dass

$$\exp(\mathbb{R}) = (0, +\infty). \tag{7.18}$$

Da

$$\lim_{x \to +\infty} \exp(x) = +\infty$$

und

$$\lim_{x \to -\infty} \exp(x) = 0,$$

so erhalten wir

$$\sup \exp = +\infty \quad \text{und} \quad \inf \exp = 0.$$

Nach dem Satz 7.8 ist $\exp(\mathbb{R})$ ein Intervall mit den Grenzen 0 und $+\infty$. Da nach dem Satz 6.3 gilt $\exp(x) > 0$ für alle $x \in \mathbb{R}$, so erhalten wir (7.18).

7.5 Extremwertsatz

Sei f eine reellwertige Funktion auf einem Intervall J. Bezeichnen wir

$$\max f = \max_J f = \max f(J) \quad \text{und} \quad \min f = \min_J f = \min f(J),$$

vorausgesetzt, dass $\max f(J)$ bzw $\min f(J)$ existiert.

Hauptsatz 7.9 (Extremwertsatz oder Satz vom Minimum und Maximum) *Sei f eine stetige Funktion auf einem abgeschlossenen beschränkten Intervall J. Dann existieren auf J die beiden Werte* $\max f$ *und* $\min f$.

Beweis. Sei $J = [a, b]$. Setzen wir

$$c := \sup f = \sup f(J) \in \overline{\mathbb{R}}.$$

Dann gibt es eine Folge $\{c_n\}_{n\in\mathbb{N}}$ von Elementen von $f(J)$ mit

$$c_n \to c \text{ für } n \to \infty.$$

In der Tat, im Fall $c \in \mathbb{R}$ wählen wir c_n als beliebiges Element von $f(J)$ in $U_{1/n}(c)$ (c_n existiert da sonst ist $c - \frac{1}{n}$ eine obere Schranke von $f(J)$ was im Widerspruch zur Definition von c steht). Im Fall $c = +\infty$ wählen wir c_n als beliebiges Element von $f(J)$ in $U_n(c)$.

Da $c_n \in f(J)$, so gibt es ein Vorbild von c_n, d.h. ein $x_n \in J$ mit $f(x_n) = c_n$, so dass

$$f(x_n) \to c \text{ für } n \to \infty.$$

Die Folge $\{x_n\}$ ist beschränkt und somit enthält nach dem Satz 4.9 von Bolzano-Weierstraß eine konvergente Teilfolge $\{x_{n_k}\}$. Sei $x = \lim_{k\to\infty} x_{n_k}$. Da $x_n \in [a, b]$, so folgt es, dass auch $x \in [a, b]$. Da f stetig ist, so erhalten wir, dass

$$f(x_{n_k}) \to f(x) \text{ für } k \to \infty,$$

woraus $f(x) = c$ folgt. Somit liegt c in $f(J)$ und das Maximum von $f(J)$ existiert und ist gleich c. Analog beweist man die Existenz des Minimums von $f(J)$. ■

Korollar 7.10 *Sei f eine stetige Funktion auf einem abgeschlossenen beschränkten Intervall J. Dann ist das Bild $f(J)$ auch ein abgeschlossenes beschränktes Intervall. Darüber hinaus gilt*

$$f(J) = [\min f, \max f].$$

Kurz: stetiges Bild eines kompakten Intervalls ist kompaktes Intervall.

Beweis. Nach dem Satz 7.9 existieren

$$\alpha = \min f(J) \quad \text{und} \quad \beta = \max f(J).$$

Nach den Satz 7.8 ist $f(J)$ ein Intervall mit den Grenzen α und β. Da α und β Elemente von $f(J)$ sind, so gilt $f(J) = [\alpha, \beta]$, was zu beweisen war. ■

7.6 Monotone Funktionen und inverse Funktion

Ist $f : A \to B$ eine Abbildung zwischen zwei beliebigen Mengen A und B, so wird die Umkehrabbildung $f^{-1} : B \to A$ durch die folgende Bedingung definiert:

$$y = f(x) \Leftrightarrow x = f^{-1}(y) \quad \forall x \in A \text{ und } \forall y \in B.$$

Nach dem Satz 1.5, die Umkehrabbildung existiert genau dann, wenn f bijektiv ist. Sind A, B Teilmengen von $\mathbb{R}$, so heißt f^{-1} auch die *inverse Funktion.*

Beispiel. Betrachten wir die Funktion

$$\begin{aligned} f &: [0, +\infty) \to [0, +\infty) \\ f(x) &= x^2 \end{aligned}$$

Da die Bedingung $y = x^2$ für nichtnegative x und y äquivalent zu $x = \sqrt{y}$ ist, so sehen wir, dass die inverse Funktion f^{-1} existiert und $f^{-1}(y) = \sqrt{y}$.

Beispiel. Für die Funktion

$$\begin{aligned} f &: (0, +\infty) \to (0, +\infty) \\ f(x) &= \frac{1}{x} \end{aligned}$$

ist die Bedingung $y = \frac{1}{x}$ äquivalent zu $x = \frac{1}{y}$. Deshalb erhalten wir $f^{-1}(y) = \frac{1}{y}$.

Hier geben wir eine hinreichende Bedingung für Existenz der inversen Funktion an.

Definition. Eine reellwertige Funktion f auf einem Intervall J heißt *monoton steigend*, falls für alle $x \le y$ gilt $f(x) \le f(y)$. Die Funktion heißt *streng* monoton steigend, falls $x < y$ ergibt $f(x) < f(y)$. Analog definiert man (streng) *monoton fallende* Funktionen.

Beispiel. Die Function $f(x) = \exp(x)$ ist streng monoton steigend auf $\mathbb{R}$ nach Satz 6.3.

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 $(0, +\infty)$ da $x^{-n} = \frac{1}{x^n}$.

Hauptsatz 7.11 *(*Existenz der inversen Funktion*) Sei f eine streng monotone steigende (bzw fallende) stetige Funktion auf einem Intervall J. Dann ist die Bildmenge $I = f(J)$ ein Intervall und die inverse Funktion $f^{-1} : I \to J$ existiert, ist streng monoton steigend (bzw fallend) und stetig.*

Beispiel. Die Funktion $f(x) = x^n$ (mit $n \in \mathbb{N}$) ist stetig und streng monoton steigend auf $[0, +\infty)$. Da ihre Bildmenge ist $[0, +\infty)$, so existiert die inverse Funktion f^{-1} auf $[0, +\infty)$. Diese Funktion heißt die n-te Wurzel und wird mit $f^{-1}(y) = \sqrt[n]{y}$ bezeichnet. Somit ist die Funktion $\sqrt[n]{y}$ stetig und streng monoton steigend auf $[0, +\infty)$. In der Tat ist der Wert $x = \sqrt[n]{y}$ die einzige nichtnegative Lösung der Gleichung $x^n = y$.

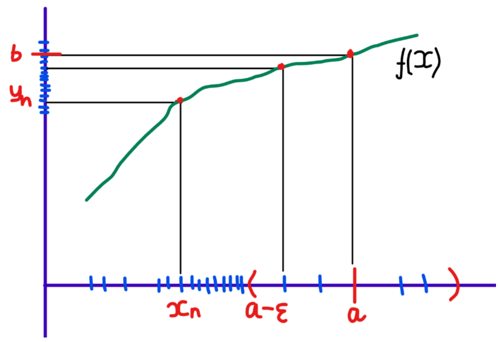
Beweis. Ist f stetig, so ist das Bild $I = f(J)$ ein Intervall nach dem Satz 7.8. Sei f streng monoton steigend (der Fall von einer fallenden Funktion ist analog). Dann ist die Abbildung $f : J \to I$ surjektiv nach der Definition von I und injektiv nach der streng Monotonie. Somit ist f bijektiv und es gibt die inverse Funktion $f^{-1} : I \to J$.

Beweisen wir, dass f^{-1} streng monotone steigend ist. Seien $y_1 < y_2$ aus I 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 $b \in I$ und für jede Folge $\{y_n\}_{n\in\mathbb{N}}$ aus I gilt

$$y_n \to b \Rightarrow f^{-1}(y_n) \to f^{-1}(b).$$

Nehmen wir das Gegenteil an, dass $a = f^{-1}(b)$ kein Grenzwert der Folge $x_n = f^{-1}(y_n)$ ist. Dann existiert ein $\varepsilon > 0$ so dass es außerhalb $U_\varepsilon(a)$ unendlich viel Glieder der Folge $\{x_n\}$ gibt. Somit enthält eines von Intervallen $(-\infty, a - \varepsilon]$, $[a + \varepsilon, +\infty)$ unendlich viel Glieder der Folge $\{x_n\}$, sei es $(-\infty, a - \varepsilon]$.



Dann existiert eine Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ mit $x_{n_k} \leq a - \varepsilon$ für alle $k \in \mathbb{N}$. Da $x_{n_k} \leq a - \varepsilon < a$ und die beiden Werte x_{n_k}, a in J sind, so gilt auch $a - \varepsilon \in J$. 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 $\lim y_n = b$ steht. ■

7.7 Logarithmische Funktion

04.07.18 H13

Hier wenden wir den Satz 7.11 auf die Exponentialfunktion an. Die Exponentialfunktion e^x ist stetig, streng monoton steigend auf $\mathbb{R}$, und die Bildmenge von e^x ist $(0, +\infty)$. Nach dem Satz 7.11 hat e^x die inverse Funktion mit dem Definitionsbereich $(0, +\infty)$ und Bildmenge $\mathbb{R}$.

Definition. Die inverse Funktion von e^x heißt *natürlicher Logarithmus* und wird mit ln bezeichnet.

Die folgende Äquivalenz ist somit für alle $x \in \mathbb{R}$ und $y \in (0, +\infty)$ erfüllt:

$$y = e^x \Leftrightarrow x = \ln y. \tag{7.19}$$

Nach dem Satz 7.11 ist die Funktion $y \mapsto \ln y$ stetig und streng monoton steigend auf $(0, +\infty)$.

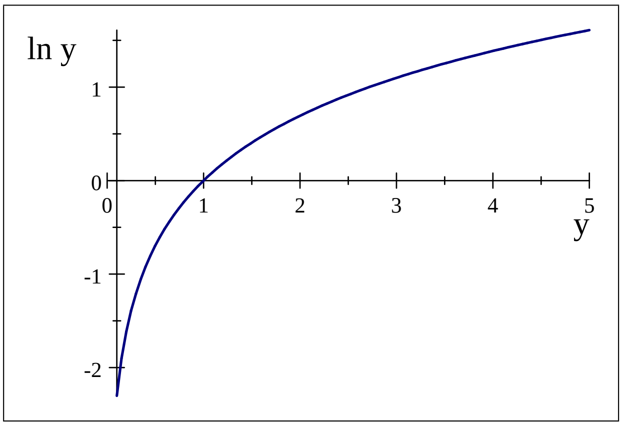
Es folgt aus (7.19) dass für alle $x \in \mathbb{R}$ gilt

$$\ln e^x = x$$

und für alle $y \in (0, +\infty)$ gilt

$$e^{\ln y} = y.$$

Insbesondere gelten $\ln 1 = 0$ und $\ln e = 1$.



Der Graph der Funktion $y \mapsto \ln y$

Die Funktion ln erfüllt für alle $u, v > 0$ die Identität

$$\boxed{\ln(uv) = \ln u + \ln v}, \tag{7.20}$$

da

$$uv = e^{\ln u} e^{\ln v} = e^{\ln u + \ln v}.$$

Insbesondere gilt es

$$\ln \frac{1}{u} = -\ln u.$$

Beispiel. Für $x \approx 0$ gilt es nach Aufgabe 106

$$e^x \approx 1 + x,$$

woraus folgt

$$\ln(1 + x) \approx x.$$

Z.B., $\ln 1,1 \approx 0,1$. In der Tat gilt $\ln 1,1 = 0,09531...$

Berechnen wir ungefähr $\ln 10$. Nach (7.20) haben wir

$$\ln 10 = \ln \frac{10}{e^2} + \ln e^2 = \ln \frac{10}{e^2} + 2.$$

Da
$$\frac{10}{e^2} = 1,3...$$
so erhalten wir
$$\ln 10 = \ln(1 + 0,3...) + 2 \approx 2,3.$$
In der Tat gilt $\ln 10 = 2,302585...$

Definition. Für jede reelle Zahl $a > 0$ und jedes $x \in \mathbb{C}$ definieren wir die Potenz a^x wie folgt:
$$\boxed{a^x := e^{x \ln a}}. \tag{7.21}$$
Die Funktion $f(x) = a^x$ heißt die *Exponentialfunktion* zur Basis a.

Die Funktion a^x hat die folgenden Eigenschaften:

1. $a^0 = 1$ da $e^0 = 1$ und $a^1 = a$ da $e^{\ln a} = a$.

2. $a^{x+y} = a^x a^y$ für alle $x, y \in \mathbb{C}$ da
$$a^x a^y = e^{x \ln a} e^{y \ln a} = e^{(x+y) \ln a} = a^{x+y}. \tag{7.22}$$
Es folgt, dass $a^{-x} = \frac{1}{a^x}$.

3. Für $x = n \in \mathbb{Z}$ stimmt die neue Definition (7.21) von a^n mit der früheren Definition von a^n für $n \in \mathbb{Z}$ überein, was man genauso beweist, wie den Satz 6.3(d).

4. Für $x \in \mathbb{R}$ ist die Funktion a^x reell, positive und stetig (als Komposition zweier stetigen Funktionen).

5. Für $x \in \mathbb{R}$ gilt $\ln a^x = x \ln a$, was aus (7.21) folgt.

6. Für alle $x, y \in \mathbb{R}$ gilt
$$(a^x)^y = a^{xy}, \tag{7.23}$$
da
$$(a^x)^y = e^{y \ln a^x} = e^{y(x \ln a)} = e^{(xy) \ln a} = a^{xy}.$$

Es folgt aus (7.23), dass für jedes $n \in \mathbb{N}$
$$\left(a^{1/n}\right)^n = a,$$
was ergibt $a^{1/n} = \sqrt[n]{a}$.

Im Fall $a = 1$ folgt es aus (7.21) dass $1^x \equiv 1$. Sei $a \neq 1$ (und $a > 0$ wie immer). Bestimmen wir die inverse Funktion von a^x, $x \in \mathbb{R}$. Die Gleichung $y = a^x$ für $x \in \mathbb{R}$ und $y \in (0, +\infty)$ ist äquivalent zu $y = e^{x \ln a}$ und somit zu $\ln y = x \ln a$, d.h.
$$y = a^x \Leftrightarrow x = \frac{\ln y}{\ln a}.$$
Somit ist $\frac{\ln y}{\ln a}$ die inverse Funktion von a^x.

Definition. Die inverse Funktion von a^x heißt *der Logarithmus zur Basis a* und wird mit $\log_a y$ bezeichnet, d.h.

$$\log_a y = \frac{\ln y}{\ln a}. \tag{7.24}$$

Die Funktion $\log_a y$ hat den Definitionsbereich $(0, +\infty)$ und die Bildmenge $\mathbb{R}$, wie $\ln y$.

Es folgt aus (7.20) und (7.24), dass für alle $u, v > 0$

$$\log_a (uv) = \log_a u + \log_a v$$

und für alle $u, a, b > 0$, $a, b \neq 1$,

$$\log_a u = \frac{\log_b u}{\log_b a}.$$

7.8 Die Zahl π und inverse trigonometrische Funktionen

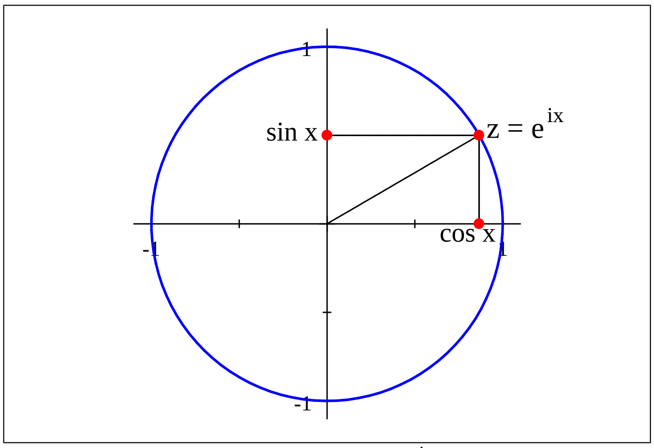
Hier bestimmen wir die inversen Funktionen zu $\sin x$ und $\cos x$ für reelles x. Die Eulerformel

$$e^{ix} = \cos x + i \sin x$$

ergibt für $x \in \mathbb{R}$ dass

$$\cos x = \operatorname{Re} e^{ix} \quad \text{und} \quad \sin x = \operatorname{Im} e^{ix}.$$

Die Zahl $z = e^{ix}$ hat den Betrag 1 (siehe Aufgabe 107) und somit liegt auf dem Einheitskreis $\mathbb{S}$; ihre $\operatorname{Re} z$ und $\operatorname{Im} z$ stellen $\cos x$ bzw $\sin x$ dar:



Die Zahl $z = e^{ix}$

Behauptung. *Die Funktionen $\cos x$ und $\sin x$ sind stetig auf $\mathbb{R}$.*

Beweis. Sei $\{x_n\}_{n \in \mathbb{N}}$ eine konvergente Folge von reellen Zahlen mit $x_n \to a \in \mathbb{R}$. Dann haben wir $ix_n \to ia$ und somit auch

$$e^{ix_n} \to e^{ia}$$

(siehe Aufgabe 106(d)), woraus folgt nach dem Satz 4.13 dass

$$\cos x_n = \operatorname{Re} e^{ix_n} \to \operatorname{Re} e^{ia} = \cos a$$

und analog

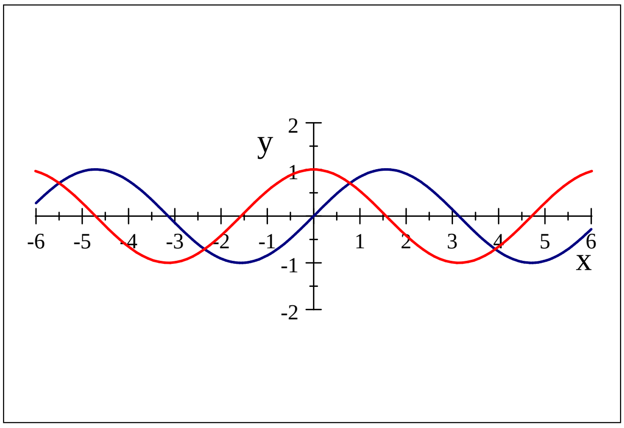
$$\sin x_n = \operatorname{Im} e^{ix_n} \to \operatorname{Im} e^{ia} = \sin a,$$

so dass $\cos x$ und $\sin x$ stetig sind. ■

Da für alle x gilt

$$\cos^2 x + \sin^2 x = 1 \tag{7.25}$$

(Aufgabe 108), so erhalten wir, dass $|\cos x| \leq 1$ und $|\sin x| \leq 1$ für alle $x \in \mathbb{R}$. Auch ist es klar, dass $\cos 0 = 1$ und $\sin 0 = 0$.



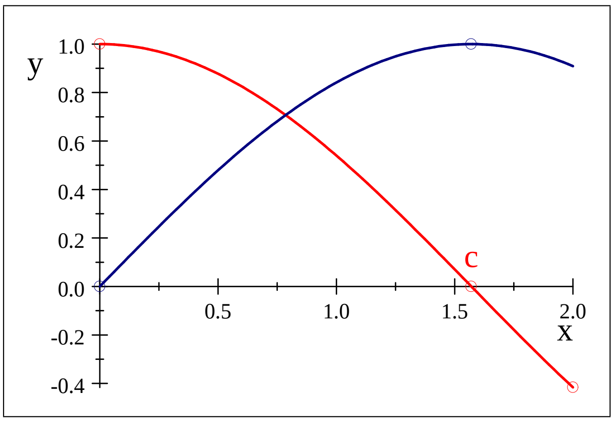
Die Graphe von Funktionen $\cos x$ (rot) und $\sin x$ (blau)

Man sieht aus den Graphen, dass die Funktionen $\cos x$ and $\sin x$ *periodisch* sind, was nicht offensichtlich aus der Definition ist. Wir beweisen die Periodizität von $\cos x$ und $\sin x$ unterhalb, und danach bestimmen die Intervalle wo diese Funktionen monoton sind.

Satz 7.12 (a) *Es gilt* $\sin x > 0$ *für alle* $x \in (0, 2)$

(b) *Die Funktion* $\cos x$ *hat im Intervall* $(0, 2)$ *genau eine Nullstelle* c. *Es gilt* $\cos x > 0$ *für alle* $0 \leq x < c$ *und* $\cos x < 0$ *für alle* $c < x < 2$. *Darüber hinaus gilt* $c > \frac{3}{2}$.

Bemerkung. Nach (b) ist c die kleinste positive Nullstelle von $\cos x$. Es folgt aus (7.25) und (a) dass $\sin c = 1$.



Die Graphe von Funktionen $\cos x$ (rot) und $\sin x$ (blau) auf dem Intervall $[0, 2]$

Beweis von Satz 7.12. Wir benutzen den Satz 5.9 (Leibniz-Kriterium) in der folgenden Form. Sei $\{a_n\}_{n=0}^{\infty}$ eine Folge von positiven reellen Zahlen, so dass $\{a_n\}_{n=1}^{\infty}$ monoton fallend ist und $a_n \to 0$ für $n \to \infty$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$ und für die Summe

$$S = \sum_{n=0}^{\infty} (-1)^n a_n = a_0 - a_1 + a_2 - ...$$

gilt für alle $m \in \mathbb{N}$

$$S_{2m} \leq S \leq S_{2m-1} \tag{7.26}$$

wobei $S_n = \sum_{k=0}^{n} (-1)^k a_k$ die Partialsumme ist. Im Satz 5.9 betrachtet man die Reihe $\sum_{n=1}^{\infty} (-1)^n a_n$ mit $n \geq 1$, während jetzt $n \geq 0$. Trotzdem gilt die Aussage (5.12) des Satzes 5.9 da zu S und S_n dieselbe Zahl a_0 addiert wird.

(a) Nach dem Satz 6.4 (und Aufgabe 108) gilt für alle $x \in \mathbb{C}$

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - ... = a_0 - a_1 + a_2 - ... \tag{7.27}$$

wobei $a_n = \frac{x^{2n+1}}{(2n+1)!}$. Für $x \in (0, 2)$ und jedes $n \geq 0$ haben wir $a_n > 0$ und

$$\frac{a_{n+1}}{a_n} = \frac{x^{2n+3}}{(2n+3)!} \Big/ \frac{x^{2n+1}}{(2n+1)!} = \frac{x^2}{(2n+2)(2n+3)} \leq \frac{4}{2 \cdot 3} < 1,$$

so dass $a_{n+1} < a_n$ und die Folge $\{a_n\}_{n=0}^{\infty}$ monoton fallend ist.

Nach (7.26) gilt für die Partialsumme $S_n = \sum_{k=0}^{n} (-1)^k a_k$ die Ungleichung

$$\sin x \geq S_{2m-1} \text{ für alle } m \in \mathbb{N}.$$

Insbesondere für $m = 1$ und $x \in (0, 2)$ erhalten wir

$$\sin x \geq S_1 = x - \frac{x^3}{6} = x\left(1 - \frac{x^2}{6}\right) > 0,$$

was zu beweisen war.

(b) Da $\cos 0 = 1 > 0$, so reicht es zu beweisen, dass $\cos 2 < 0$, woraus die Existenz der Nullstelle von $\cos x$ im $(0, 2)$ aus dem Zwischenwertsatz (Satz 7.7) folgt. Da nach dem Satz 6.4

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \tag{7.28}$$

so haben wir

$$\cos 2 = \sum_{n=0}^{\infty} (-1)^n \frac{2^{2n}}{(2n)!} = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \frac{2^6}{6!} - ... = b_0 - b_1 + b_2 - b_3 + ... \tag{7.29}$$

wobei $b_n = \frac{2^{2n}}{(2n)!}$. Die Folge $\{b_n\}_{n=1}^{\infty}$ ist monoton fallend da für $n \geq 1$

$$\frac{b_{n+1}}{b_n} = \frac{2^{2n+2}}{(2n+2)!} \Big/ \frac{2^{2n}}{(2n)!} = \frac{2^2}{(2n+1)(2n+2)} < 1.$$

Nach (7.26) gilt für die Partialsumme $S_n = \sum_{k=0}^{n} (-1)^k b_k$ die Ungleichung

$$\cos 2 \le S_{2m} \quad \text{für ale } m \in \mathbb{N}.$$

Für $m = 1$ erhalten wir

$$\cos 2 \le S_2 = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = 1 - 2 + \frac{16}{24} = -\frac{1}{3} < 0,$$

woraus $\cos 2 < 0$ folgt. Nach dem Satz 7.7 beschließen wir, dass $\cos x = 0$ für ein $x \in (0, 2)$.

Beweisen wir jetzt die Eindeutigkeit der Nullstelle von $\cos x$ in $(0, 2)$. Dafür benutzen wir Additionstheorem

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y \tag{7.30}$$

(siehe Aufgabe 108). Seien x, y zwei verschiedene Nullstellen von cos in $(0, 2)$. Sei $x > y$. Da $\cos x = \cos y = 0$, so erhalten wir nach (7.30)

$$\sin(x - y) = \sin x \cos y - \cos x \sin y = 0,$$

was nicht möglich ist, da $0 < x - y < 2$ und nach (a) gilt $\sin(x - y) > 0$.

Bezeichnen wir mit c die einzige Nullstelle von cos in $(0, 2)$. Dann gilt $\cos x > 0$ für alle $x \in [0, c)$ da im Fall $\cos x \le 0$ es noch eine Nullstelle von cos in $[0, x]$ gibt. Genau so gilt $\cos x < 0$ für alle $x \in (c, 2)$.

Es bleibt zu beweisen, dass $c > \frac{3}{2}$. Dafür reicht es zu beweisen, dass $\cos \frac{3}{2} > 0$. In der Tat, mit Hilfe von (7.26) und (7.28) erhält man, dass

$$\cos \frac{3}{2} \ge S_3 = 1 - \frac{(3/2)^2}{2!} + \frac{(3/2)^4}{4!} - \frac{(3/2)^6}{6!} = \frac{359}{5120} > 0.$$

■

Definition. Setzen wir $\pi := 2c$ wobei c die kleinste positive Nullstelle von $\cos x$ ist, die nach dem Satz 7.12 existiert.

Es folgt aus dem Satz 7.12, dass $3 < \pi < 4$. Numerische Berechnung ergibt

$$\pi = 3,14159265358979...$$

Es ist bekannt, dass π eine irrationale und sogar transzendente Zahl ist. 06.07.18 H1

Es folgt aus dem Satz 7.12, dass $\cos \frac{\pi}{2} = 0$, $\sin \frac{\pi}{2} = 1$ und dass $\cos x$ und $\sin x$ positiv im Intervall $\left(0, \frac{\pi}{2}\right)$ sind.

Satz 7.13 (a) *Es gelten die Identitäten*

$$\boxed{e^{\frac{\pi}{2}i} = i, \quad e^{\pi i} = -1, \quad e^{2\pi i} = 1.}$$

(b) *Die Funktion e^z ist $2\pi i$ periodisch auf $\mathbb{C}$, d.h.*

$$e^{z+2\pi i} = e^z \ \forall z \in \mathbb{C}. \tag{7.31}$$

(c) *Die Funktionen $\sin x$ und $\cos x$ sind 2π periodisch, d.h.*

$$\sin(x + 2\pi) = \sin x \ \ \textit{und} \ \ \cos(x + 2\pi) = \cos x \ \ \forall x \in \mathbb{C}.$$

Beweis. (a) Nach der Eulerformel erhalten wir

$$e^{\frac{\pi}{2}i} = \cos\tfrac{\pi}{2} + i\sin\tfrac{\pi}{2} = i.$$

Es folgt, dass

$$e^{\pi i} = e^{\frac{\pi}{2}i}e^{\frac{\pi}{2}i} = i \cdot i = -1,$$

und

$$e^{2\pi i} = e^{\pi i}e^{\pi i} = (-1)^2 = 1. \tag{7.32}$$

(b) Mit Hilfe von (7.32) und der Haupteigenschaft der Exponentialfunktion erhalten wir

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z.$$

(c) Mit Hilfe von (6.14) und (7.31) erhalten wir

$$\begin{aligned} \sin(x+2\pi) &= \frac{1}{2i}\left(e^{ix+2\pi i} - e^{-ix-2\pi i}\right) \\ &= \frac{1}{2i}\left(e^{ix} - e^{-ix}\right) = \sin x, \end{aligned}$$

und dasselbe Argument funktioniert für $\cos x$. ■

Für die nächsten Beispiele benutzen wir die Additionstheoreme:

$$\begin{aligned} \cos(x+y) &= \cos x \cos y - \sin x \sin y, \\ \sin(x+y) &= \sin x \cos y + \cos x \sin y, \end{aligned} \tag{7.33}$$

die für alle $x, y \in \mathbb{C}$ gelten (Aufgabe 108).

Beispiel. Da $\cos\frac{\pi}{2} = 0$ und $\sin\frac{\pi}{2} = 1$, so erhalten wir aus (7.33) dass

$$\cos\left(x + \tfrac{\pi}{2}\right) = -\sin x \quad \text{und} \quad \sin\left(x + \tfrac{\pi}{2}\right) = \cos x.$$

Ersetzen x durch $-x$ ergibt

$$\cos\left(\tfrac{\pi}{2} - x\right) = \sin x \quad \text{und} \quad \sin\left(\tfrac{\pi}{2} - x\right) = \cos x.$$

Es folgt aus $e^{\pi i} = -1$, dass $\cos\pi = -1$, $\sin\pi = 0$ und somit aus (7.33)

$$\cos(x+\pi) = -\cos x \quad \text{und} \quad \sin(x+\pi) = -\sin x\,.$$

Beispiel. Berechnen wir $\sin\frac{\pi}{4}$. Es gilt

$$\sin\tfrac{\pi}{4} = \cos\left(\tfrac{\pi}{2} - \tfrac{\pi}{4}\right) = \cos\tfrac{\pi}{4},$$

sodass $x = \sin\frac{\pi}{4}$ die Gleichung $x^2 + x^2 = 1$ erfüllt. Es folgt, dass $x = \pm\frac{\sqrt{2}}{2}$. Das Vorzeichen muss $+$ sein da $\sin\frac{\pi}{4} > 0$. Somit erhalten wir

$$\sin\tfrac{\pi}{4} = \cos\tfrac{\pi}{4} = \frac{\sqrt{2}}{2}.$$

Beispiel. Berechnen wir $\sin\frac{\pi}{6}$. Wir haben

$$\sin\tfrac{\pi}{6} = \cos\left(\tfrac{\pi}{2}-\tfrac{\pi}{6}\right) = \cos\tfrac{\pi}{3} = \cos 2\tfrac{\pi}{6} = \cos^2\tfrac{\pi}{6} - \sin^2\tfrac{\pi}{6} = 1 - 2\sin^2\tfrac{\pi}{6}.$$

Somit erfüllt $x = \sin\frac{\pi}{6}$ die Gleichung $2x^2 + x - 1 = 0$, woraus folgt $x = -1$ oder $x = \frac{1}{2}$. Da $\sin\frac{\pi}{6} > 0$, so erhalten wir $\sin\frac{\pi}{6} = \frac{1}{2}$ und somit auch

$$\cos\tfrac{\pi}{6} = \sqrt{1 - \left(\tfrac{1}{2}\right)^2} = \tfrac{\sqrt{3}}{2}.$$

Es folgt, dass

$$\cos\tfrac{\pi}{3} = \frac{1}{2} \quad \text{und} \quad \sin\tfrac{\pi}{3} = \tfrac{\sqrt{3}}{2}.$$

$$e^{0i} = e^{2\pi i} = 1$$

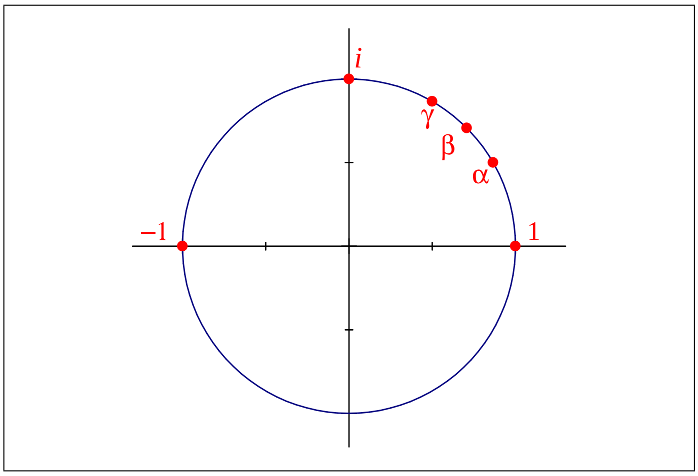
$$e^{\frac{\pi}{6}i} = \tfrac{\sqrt{3}}{2} + \tfrac{1}{2}i = \alpha$$

$$e^{\frac{\pi}{4}i} = \tfrac{1}{\sqrt{2}} + \tfrac{1}{\sqrt{2}}i = \beta$$

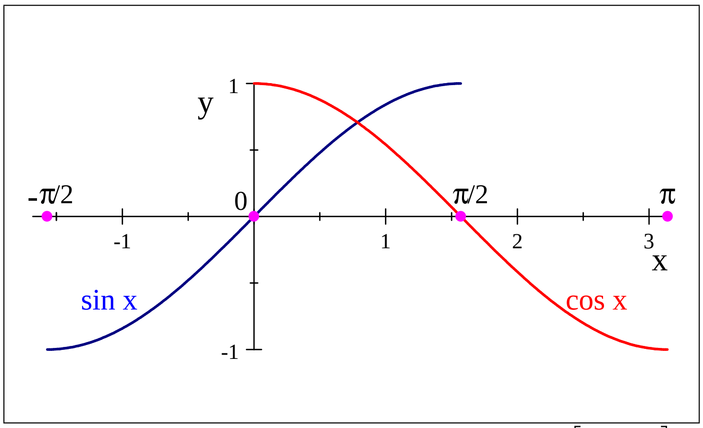
$$e^{\frac{\pi}{3}i} = \tfrac{1}{2} + \tfrac{\sqrt{3}}{2}i = \gamma$$

$$e^{\frac{\pi}{2}i} = i$$

$$e^{\pi i} = -1$$



Satz 7.14 (*a*) *Die Funktion* $\sin x$ *ist streng monoton steigend auf* $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
(*b*) *Die Funktion* $\cos x$ *ist streng monoton fallend auf* $[0, \pi]$.



Funktionen $\cos x$ auf $[0, \pi]$ (rot) und $\sin x$ auf $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ (blau).

Beweis. (*a*) Wir haben $\sin\frac{\pi}{2} = 1$, $\sin 0 = 0$ und $\sin\left(-\frac{\pi}{2}\right) = -1$. Für $x \in (0, \frac{\pi}{2})$ gilt $0 < \sin x < 1$ und für $x \in \left(-\frac{\pi}{2}, 0\right)$ gilt $\sin x = -\sin(-x)$ und somit $-1 < \sin x < 0$. So reicht es zu beweisen, dass $\sin x$ streng monoton steigend auf $(0, \frac{\pi}{2})$ ist, d.h. für alle $x, y \in \left(0, \frac{\pi}{2}\right)$

$$x > y \Rightarrow \sin x > \sin y.$$

Setzen wir $y = x - t$ wobei $0 < t < x$, insbesondere $0 < t < \frac{\pi}{2}$. Wir haben

$$\begin{aligned}
\sin x - \sin y &= \sin x - \sin(x - t) \\
&= \sin x - (\sin x \cos t - \cos x \sin t) \\
&= \sin x (1 - \cos t) + \cos x \sin t > 0,
\end{aligned}$$

da alle Zahlen $\sin x$, $1 - \cos t$, $\cos x$, $\sin t$ positiv sind.

(b) Diese Aussage folgt aus (a) und aus der Identität $\cos x = -\sin\left(x - \frac{\pi}{2}\right)$, da

$$x \in [0, \pi] \Leftrightarrow x - \frac{\pi}{2} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

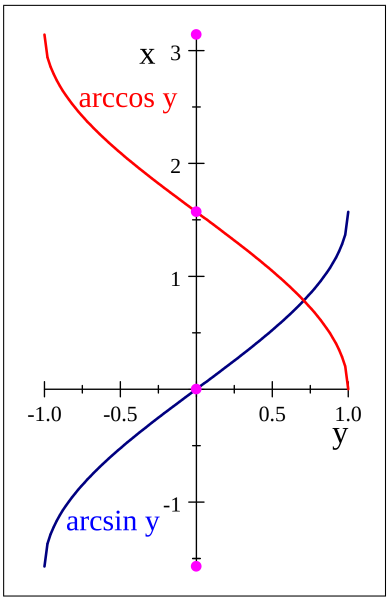
■

Betrachten wir die Funktion $f(x) = \cos x$ auf dem Intervall $J = [0, \pi]$. Diese Funktion ist stetig und streng monoton fallend, so dass die inverse Funktion f^{-1} existiert mit dem Definitionsbereich $I = f(J) = [-1, 1]$ und der Bildmenge $[0, \pi]$. Die Funktion f^{-1} heißt *Arkuscosinus* und wird mit arccos bezeichnet. Somit gilt

$$y = \cos x \Leftrightarrow x = \arccos y \quad \forall x \in [0, \pi], \ y \in [-1, 1].$$

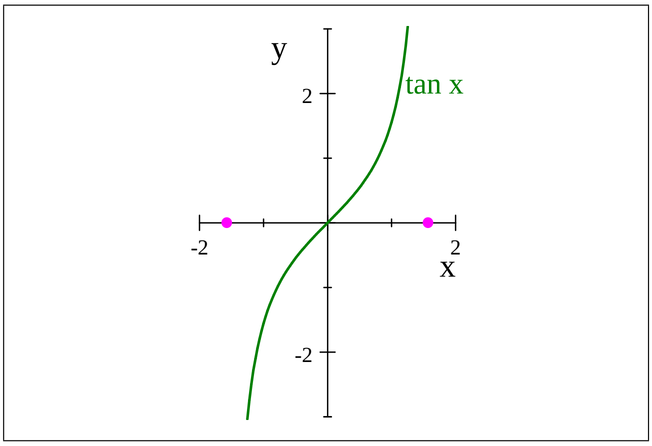
Analog betrachten wir $f(x) = \sin x$ auf $J = [-\pi/2, \pi/2]$. Da $f(x)$ auf diesem Intervall stetig und streng monoton steigend ist und $I = f(J) = [-1, 1]$, so existiert die inverse Funktion f^{-1} mit dem Definitionsbereich $[-1, 1]$ und der Bildmenge $[-\pi/2, \pi/2]$. Die inverse Funktion von sin heißt *Arkussinus* wird mit arcsin bezeichnet. Somit gilt

$$y = \sin x \Leftrightarrow x = \arcsin y \quad \forall x \in [-\pi/2, \pi/2], \ y \in [-1, 1].$$

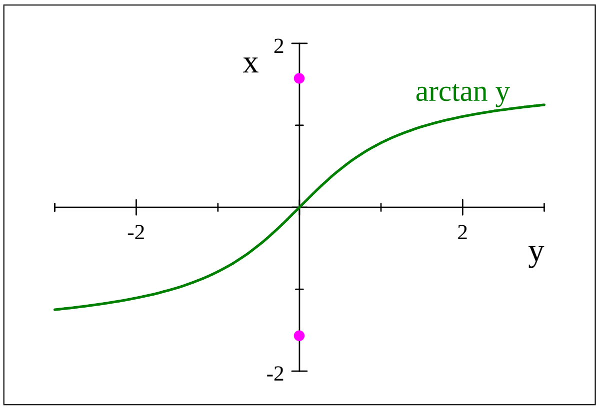


Funktionen arcsin (blau) und arccos (rot)

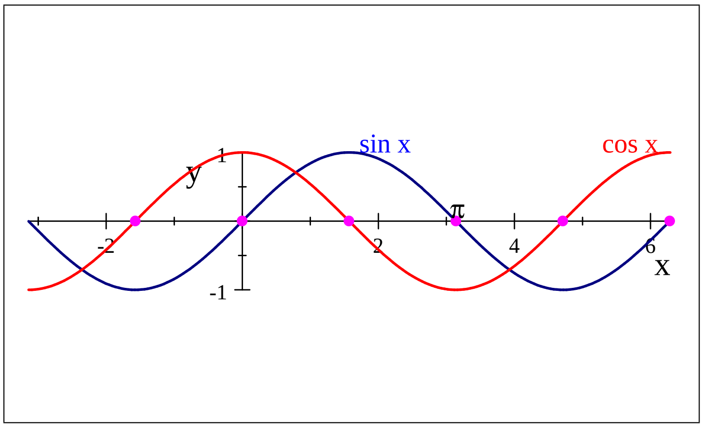
Man kann beweisen, dass die Funktion $f(x) = \tan x = \frac{\sin x}{\cos x}$ im Intervall $(-\pi/2, \pi/2)$ streng monoton steigend ist und die Bildmenge von tan gleich $(-\infty, +\infty)$ ist (Aufgabe 115).



Somit existiert die inverse Funktion von $\tan x$. Sie heißt *Arkustangens* und wird mit arctan bezeichnet. Der Definitionsbereich von arctan ist $(-\infty, +\infty)$ und die Bildmenge ist $(-\pi/2, \pi/2)$.



Bemerkung. Da $\sin x > 0$ auf $(0, \pi/2)$ und $\sin x < 0$ auf $(-\pi/2, 0)$ so folgt es aus $\sin(x + \pi) = -\sin x$, dass $\sin x > 0$ auf $(0, \pi)$ und somit $\sin x < 0$ auf $(-\pi, 0)$. Analog folgt es aus $\cos(x + \pi) = -\cos x$ dass $\cos x > 0$ auf $(-\pi/2, \pi/2)$ und $\cos x < 0$ auf $(\pi/2, 3\pi/2)$.



$\cos x$ (rot) und $\sin x$ (blau) auf $[-\pi, 2\pi]$

7.9 Trigonometrische Form komplexer Zahlen

Satz 7.15 *Jede komplexe Zahl $z \neq 0$ lässt sich in der folgenden Form eindeutig darstellen:*

$$z = re^{i\theta} = r(\cos\theta + i\sin\theta), \tag{7.34}$$

wobei $r \in (0, +\infty)$ und $\theta \in (-\pi, \pi]$.

Wir werden im Beweis sehen, dass $r = |z|$, d.h. r der Betrag von z ist. Die Zahl r heißt auch der *Polarradius* von z. Die Zahl θ heißt der *Polarwinkel* von z; auch heißt sie das *Argument* von z und wird $\arg z$ bezeichnet. Zusammen r und θ heißen die *Polarkoordinaten* for z.

Der Ausdruck $r(\cos\theta + i\sin\theta)$ heißt die *trigonometrische* Form von z, im Gegensatz zur *algebraischen* Form $z = x + yi$. Der Ausdruck $re^{i\theta}$ heißt die *exponentielle Form* von z.

Obwohl θ im Intervall $(-\pi, \pi]$ eindeutig bestimmt ist, häufig erlaubt man θ (und $\arg z$) beliebige Werte aus $\mathbb{R}$ annehmen. Da $\cos\theta$ und $\sin\theta$ 2π-periodisch sind, so lässt sich θ in (7.34) durch $\theta + 2\pi k$ für jedes $k \in \mathbb{Z}$ ersetzen. So, für $\theta \in \mathbb{R}$ fehlt die Eindeutigkeit der Darstellung (7.34), aber trotzdem ist θ *bis zur additiven Konstante* $2\pi k$ mit $k \in \mathbb{Z}$ bestimmt.

Beweis. Sei $z = x + yi$. Wir betrachten die Identität (7.34) als eine Gleichung mit unbekannten r und θ. Aus (7.34) erhalten wir $|z| = |r|\left|e^{i\theta}\right| = r$ so dass

$$r = |z| = \sqrt{x^2 + y^2}.$$

Vergleichen von (7.34) mit der algebraischen Darstellung $z = x + iy$ ergibt uns zwei Gleichungen für θ:

$$x = r\cos\theta \quad \text{und} \quad y = r\sin\theta, \tag{7.35}$$

d.h.

$$\cos\theta = \frac{x}{r} \quad \text{und} \quad \sin\theta = \frac{y}{r}. \tag{7.36}$$

Beweisen wir, dass die Gleichungen (7.36) eine eindeutige Lösung $\theta \in (-\pi, \pi]$ haben.

Fall 1: sei $y \geq 0$. Dann gilt $\sin\theta \geq 0$ und somit $\theta \in [0, \pi]$. Die Gleichung $\cos\theta = \frac{x}{r}$ im Bereich $\theta \in [0, \pi]$ ist äquivalent zu $\theta = \arccos\frac{x}{r}$ (bemerken wir, dass $\frac{x}{r} \in [-1, 1]$ so dass $\frac{x}{r}$ im Definitionsbereich von arccos liegt). Insbesondere ist θ eindeutig bestimmt. Für $\theta = \arccos\frac{x}{r}$ gilt offensichtlich $\cos\theta = \frac{x}{r}$ und

$$\sin\theta = \sqrt{1 - \cos^2\theta} = \sqrt{1 - \frac{x^2}{r^2}} = \frac{1}{r}\sqrt{r^2 - x^2} = \frac{y}{r},$$

da $\sin\theta$ und y nicht negative sind. Somit erfüllt $\theta = \arccos\frac{x}{r}$ die beiden Gleichungen (7.36).

Fall 2: sei $y < 0$. Die Gleichungen (7.36) sind äquivalent zu

$$\cos(-\theta) = \frac{x}{r} \quad \text{und} \quad \sin(-\theta) = -\frac{y}{r} > 0,$$

woraus folgt $-\theta \in (0, \pi)$ und somit $\theta = -\arccos\frac{x}{r}$. Dieses θ erfüllt die beiden Gleichungen in (7.36) wie im Fall 1. ∎

Aus dem Beweis erhalten wir die folgenden Formeln für die Polarkoordinaten

$$r = \sqrt{x^2 + y^2} \tag{7.37}$$

und

$$\theta = \begin{cases} \arccos \frac{x}{r}, & y \geq 0, \\ -\arccos \frac{x}{r}, & y < 0. \end{cases} \tag{7.38}$$

Umgekehrt, aus den Polarkoordinaten r, θ erhält man die *Kartesischen* Koordinaten x, y mit Hilfe von (7.35).

Beispiel. Für $z = 1 + i$ haben wir nach (7.37) und (7.38) $r = |z| = \sqrt{2}$ und

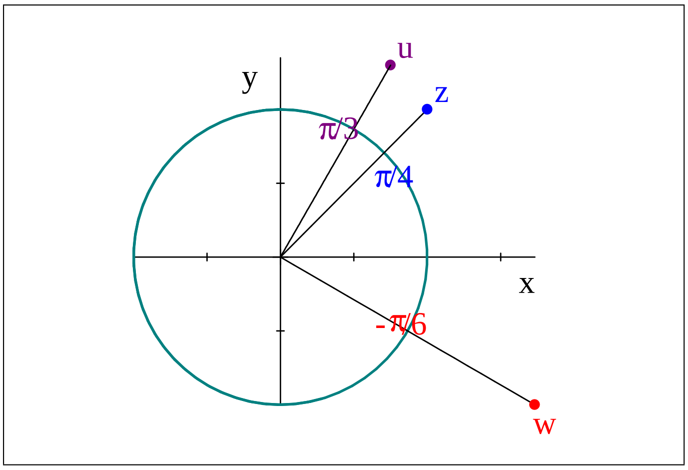
$$\theta = \arg z = \arccos \tfrac{1}{\sqrt{2}} = \tfrac{\pi}{4}$$

so dass $1 + i = \sqrt{2}e^{i\frac{\pi}{4}}$.

Für $w = \sqrt{3} - i$ haben wir $r = |w| = 2$ und

$$\theta = \arg w = -\arccos \tfrac{\sqrt{3}}{2} = -\tfrac{\pi}{6}$$

so dass $\sqrt{3} - i = 2e^{-\frac{\pi}{6}i}$.



Für $u = \frac{3}{2}e^{i\frac{\pi}{3}}$ haben wir $x = \frac{3}{2}\cos\frac{\pi}{3} = \frac{3}{2} \cdot \frac{1}{2} = \frac{3}{4}$ und $y = \frac{3}{2}\sin\frac{\pi}{3} = \frac{3}{2} \cdot \frac{\sqrt{3}}{2} = \frac{3\sqrt{3}}{4}$.

7.10 * Numerische Berechnung von π

Man kann die erste positive Nullstelle c von $\cos x$ (und somit die Zahl π) numerisch mit Hilfe von einem Taschenrechner mit Funktion cos berechnen. Wir wissen, dass $c \in (a, b)$ wobei $a = 1,5$ und $b = 2$. Hier führen wir sechs Schritte des Verfahrens im Beweis des Satzes 7.7 durch und erhalten die folgende Intervallschachtelung.

k	a_k	$\frac{a_k+b_k}{2}$	b_k
1	$1,5$ $\cos 1,5 \approx 0,07 > 0$	$1,75$ $\cos 1,75 \approx -0.178 < 0$	2 $\cos 2 \approx -0,41 < 0$
2	$1,5$ $\cos 1,5 \approx 0,07 > 0$	$1,625$ $\cos 1,625 \approx -0,054 < 0$	$1,75$ $\cos 1,75 \approx -0.178 < 0$
3	$1,5$ $\cos 1,5 \approx 0,07 > 0$	$1,562\,5$ $\cos 1,562\,5 \approx 0,0082 > 0$	$1,625$ $\cos 1,625 \approx -0,054 < 0$
4	$1,562\,5$ $\cos 1,562\,5 \approx 0,0082 > 0$	$1,59375$ $\cos 1,593\,75 \approx -0.0229 < 0$	$1,625$ $\cos 1,625 \approx -0,054 < 0$
5	$1,562\,5$ $\cos 1,562\,5 \approx 0,0082 > 0$	$1,578\,125$ $\cos 1,578\,125 \approx -0.0073 < 0$	$1,59375$ $\cos 1,593\,75 \approx -0.0229 < 0$
6	$1,562\,5$ $\cos 1,562\,5 \approx 0,0082 > 0$	$1,570\,312\,5$ $\cos 1,578\,125 \approx 0.00048 > 0$	$1,578\,125$ $\cos 1,578\,125 \approx -0.0073 < 0$

Somit $c \approx 1,57$ und $\pi \approx 3,14$.

Chapter 8

Differentialrechnung

8.1 Ableitung

Sei $f : J \to \mathbb{R}$ eine Funktion auf einem Intervall J.

Definition. Die *Ableitung* der Funktion f an einer Stelle $x \in J$ ist der Grenzwert

$$f'(x) := \lim_{y \to x} \frac{f(y) - f(x)}{y - x},$$

vorausgesetzt, dass er existiert. Ist der Grenzwert endlich, so heißt f *differenzierbar* an der Stelle $x \in J$. Ist f an allen Stellen $x \in J$ differenzierbar, so sagt man, dass f auf dem Intervall J differenzierbar ist. In diesem Fall ist die Ableitung f' auch eine Funktion auf J.

Einsetzen $h = y - x$ ergibt eine äquivalente Definition der Ableitung:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}. \tag{8.1}$$

Der Ausdruck $\frac{f(x+h)-f(x)}{h}$ 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

$$\boxed{(\text{const})' = 0}.$$

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

$$f'(x) = \lim_{y \to x} \frac{y^2 - x^2}{y - x} = \lim_{y \to x} (y + x) = 2x,$$

daher

$$\left(x^2\right)' = 2x.$$

4. $f(x) = \frac{1}{x}$. Dann gilt

$$f'(x) = \lim_{y\to x} \frac{1}{y-x}\left(\frac{1}{y} - \frac{1}{x}\right) = -\lim_{y\to x} \frac{1}{yx} = -\frac{1}{x^2},$$

so dass

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}.$$

Es gilt auch für alle $n \in \mathbb{Z}$,

$$\boxed{\left(x^n\right)' = nx^{n-1}}$$

was später bewiesen wird.

5. $f(x) = e^x$. Dann gilt

$$f'(x) = \lim_{h\to 0} \frac{e^{x+h} - e^x}{h} = e^x \lim_{h\to 0} \frac{e^h - 1}{h} = e^x,$$

da

$$\lim_{h\to 0} \frac{e^h - 1}{h} = 1. \tag{8.2}$$

In der Tat, für $|h| < 1$ gilt

$$\left|e^h - 1 - h\right| = \left|\frac{h^2}{2!} + \frac{h^3}{3!} + ...\right| \le |h|^2 + |h|^3 + ... = \frac{|h|^2}{1-|h|}$$

und somit

$$\left|\frac{e^h - 1}{h} - 1\right| = \left|\frac{e^h - 1 - h}{h}\right| \le \frac{|h|}{1-|h|} \to 0 \text{ für } h \to 0,$$

woraus (8.2) folgt. Somit haben wir

$$\boxed{\left(e^x\right)' = e^x}.$$

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 Aufgabe 111 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
$$\boxed{(\cos x)' = -\sin x}.$$

11.07.18 H13

Satz 8.1 *Ist die Funktion f differenzierbar an x, so ist f stetig an x.*

Beweis. Wir haben
$$\lim_{y\to x} \left(f\left(y\right) - f\left(x\right)\right) = \lim_{y\to x} \frac{f\left(y\right) - f\left(x\right)}{y - x} \left(y - x\right) = f'\left(x\right) \lim_{y\to x} \left(y - x\right) = 0,$$
woraus folgt
$$\lim_{y\to x} f\left(y\right) = f\left(x\right).$$
Somit ist f stetig an x. ■

Beispiel. Betrachten wir die Funktion
$$f\left(x\right) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0, \end{cases}$$
die an $x = 0$ unstetig ist. Folglich ist f an $x = 0$ nicht differenzierbar.

Beispiel. Allerdings reicht die Stetigkeit für die Differenzierbarkeit nicht. Die Funktion $f\left(x\right) = |x|$ ist stetig auf $\mathbb{R}$ (Aufgabe 112) aber an der Stelle $x = 0$ ist sie nicht differenzierbar. Für jedes $y > 0$ haben wir $|y| = y$ und somit
$$\frac{f\left(y\right) - f\left(0\right)}{y - 0} = 1,$$
während für $y < 0$ gilt $|y| = -y$ und somit
$$\frac{f\left(y\right) - f\left(0\right)}{y - 0} = -1.$$
Es folgt, dass
$$\lim_{y\to 0} \frac{f\left(y\right) - f\left(0\right)}{y - 0}$$
nicht existiert und f an $x = 0$ nicht differenzierbar ist.

8.2 Physikalische und geometrische Bedeutung der Ableitung

Sei $f(t)$ die Position eines bewegenden Körpers an der Zahlengerade am Zeitpunkt t. Dann gilt für den Differenzenquotient

$$\begin{aligned}\frac{f(t+h)-f(t)}{h} &= \frac{\text{Verschiebung des Körpers}}{\text{Zeitintervall}}\\ &= \text{durchschnittliche Geschwindigkeit im } [t,t+h]\end{aligned}$$

Für $h \to 0$ erhalten wir, dass die Ableitung $f'(t)$ gleich *unmittelbare* Geschwindigkeit vom Körper am Zeitpunkt t ist. Zum Beispiel, im Auto wird $f'(t)$ am Tachometer an jedem Zeitpunkt t gezeigt.

Das Gleiche gilt für beliebige zeitabhängige Variable $f(t)$: die Ableitung ist immer die unmittelbare Geschwindigkeit von Änderung der Variable. Analysis lösst damit zwei verbundene Probleme:

1. gegeben sei $f(t)$, wie berechnet man $f'(t)$ (Differentialrechnung);

2. gegeben sei $f'(t)$, wie berechnet man $f(t)$ (Integralrechnung).

Eine andere Bedeutung der Ableitung gibt es in Geometrie. Betrachten wir eine Gerade G in $\mathbb{R}^2$ mit der Gleichung

$$y = Ax + B,$$

wobei $A, B \in \mathbb{R}$. Geht G durch einen Punkte (a,b), so kann man die Gleichung wie folgt umschreiben

$$y = A(x-a) + b.$$

Geht G durch noch einen anderen Punkt (a_1, b_1), so erhalten wir die Gleichung

$$b_1 = A(a_1 - a) + b,$$

woraus die folgende Identität für die *Steigung* A folgt:

$$A = \frac{b_1 - b}{a_1 - a}.$$

Betrachten wir jetzt den Graph einer Funktion $y = f(x)$ auf einem Intervall. Die Gerade durch zwei Punkte $(a, f(a))$ und $(a+h, f(a+h))$ auf dem Graph heißt die *Sekante*. Die Steigung der Sekante ist somit gleich

$$A = A(h) = \frac{f(a+h) - f(a)}{h},$$

und die Gleichung der Sekante ist

$$y = A(h)(x-a) + f(a).$$

Für $h \to 0$ erhalten wir $A(h) \to f'(a)$. Die Gerade mit der Gleichung

$$\boxed{y = f'(a)(x-a) + f(a)} \tag{8.3}$$

heißt die *Tangente* zum Graph von $f(x)$ am Punkt $(a, f(a))$. Somit konvergiert die Sekante gegen die Tangente für $h \to 0$.

So, die Ableitung $f'(a)$ ist gleich die Steigung der Tangente des Graphes an $(a, f(a))$.

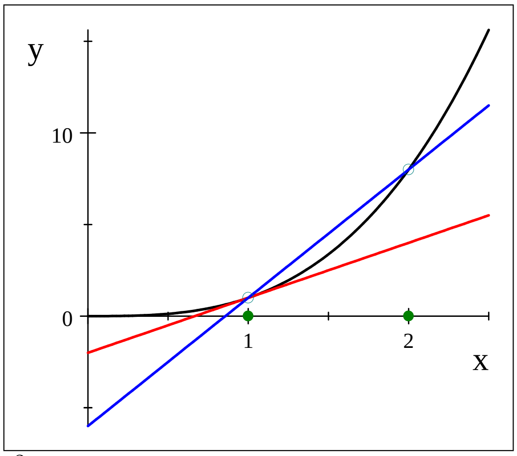
Beispiel. Für $f(x) = x^3$ erhalten wir aus (8.3) die Gleichung der Tangente an (a, a^3) :

$$y = 2a^2(x-a) + a^3.$$

Zum Beispiel, für $a = 1$ 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)

8.3 Rechenregeln für Ableitung

Satz 8.2 *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:

$$\boxed{(f+g)'(x) = f'(x) + g'(x).} \tag{8.4}$$

(b) Produktregel oder Leibnizregel:

$$\boxed{(fg)'(x) = f'(x)\,g(x) + f(x)\,g'(x)\,.} \tag{8.5}$$

(c) Quotientenregel:

$$\boxed{\left(\frac{f}{g}\right)'(x) = \frac{f'(x)\,g(x) - f(x)\,g'(x)}{g^2(x)}}. \tag{8.6}$$

Es folgt aus (8.5) mit $g(x) \equiv c \in \mathbb{R}$ dass

$$\boxed{(cf)' = cf'},$$

und aus (8.6) mit $f(x) \equiv 1$ dass

$$\boxed{\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}} \tag{8.7}$$

Beispiel. Beweisen wir, dass für $n \in \mathbb{Z}$ gilt

$$(x^n)' = nx^{n-1}.$$

Für $n \geq 0$ beweisen wir per Induktion nach n. Für $n = 0$ das ist trivial da $1' = 0$. Induktionsschritt von n nach $n+1$: nach der Produktregel gilt

$$\left(x^{n+1}\right)' = (x \cdot x^n)' = (x)'\,x^n + x\,(x^n)' = x^n + nx^n = (n+1)\,x^n.$$

Sei $n < 0$ so dass $n = -m$ für $m \in \mathbb{N}$ und somit nach der Quotientenregel

$$(x^n)' = \left(\frac{1}{x^m}\right)' = -\frac{(x^m)'}{x^{2m}} = -\frac{mx^{m-1}}{x^{2m}} = -\frac{m}{x^{m+1}} = nx^{n-1}.$$

Beispiel. Bestimmen wir die Ableitung der Funktion

$$\tan x = \frac{\sin x}{\cos x}.$$

Nach der Quotientenregel haben wir

$$(\tan x)' = \frac{(\sin x)'\cos x - \sin x\,(\cos x)'}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$$

d.h.

$$\boxed{(\tan x)' = \frac{1}{\cos^2 x}}$$

Analog beweist man, dass für

$$\cot x = \frac{\cos x}{\sin x}$$

gilt

$$\boxed{(\cot x)' = -\frac{1}{\sin^2 x}}.$$

Beweis von Satz 8.2. (a) Nach Definition von Ableitung und Satz 7.2 erhalten wir

$$\begin{aligned}(f+g)'(x) &= \lim_{y\to x}\frac{(f+g)(y)-(f+g)(x)}{y-x}\\ &= \lim_{y\to x}\frac{f(y)-f(x)}{y-x}+\lim_{y\to x}\frac{g(y)-g(x)}{y-x}\\ &= f'(x)+g'(x).\end{aligned}$$

(b) Mit gleichem Argument und mit Hilfe von Stetigkeit von f an der Stelle x, die nach Satz 8.1 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 $\left(\frac{1}{g}\right)'$:

$$\begin{aligned}\left(\frac{1}{g}\right)'(x) &= \lim_{y\to x}\frac{1}{y-x}\left(\frac{1}{g(y)}-\frac{1}{g(x)}\right)\\ &= \lim_{y\to x}\frac{1}{y-x}\left(\frac{g(x)-g(y)}{g(y)g(x)}\right)\\ &= \lim_{y\to x}\frac{g(x)-g(y)}{y-x}\lim_{y\to x}\frac{1}{g(y)g(x)}\\ &= -g'(x)\frac{1}{g^2(x)},\end{aligned}$$

d.h.

$$\left(\frac{1}{g}\right)'(x)=-\frac{g'(x)}{g^2(x)}. \tag{8.8}$$

Für allgemeine Funktion f erhalten wir mit Hilfe von Produktregel (8.5) und (8.8):

$$\begin{aligned}\left(\frac{f}{g}\right)' &= \left(f\frac{1}{g}\right)'=f'\left(\frac{1}{g}\right)+f\left(\frac{1}{g}\right)'\\ &= \frac{f'}{g}-\frac{fg'}{g^2}=\frac{f'g-fg'}{g^2}.\end{aligned}$$

■

Bemerkung. Für das Produkt dreier Funktionen f, g, h gilt die Regel

$$(fgh)'=f'gh+fg'h+fgh',$$

da nach (8.5)

$$\begin{aligned}(fgh)' &= ((fg)h)' = (fg)'h + (fg)h' \\ &= (f'g + fg')h + fgh' = f'gh + fg'h + fgh'.\end{aligned}$$

Analog gilt für das Produkt von n Funktionen $f_1, ..., f_n$ 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.

8.4 Kettenregel

Satz 8.3 (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*

$$\boxed{(g \circ f)'(x) = g'(y) f'(x)} = g'(f(x)) f'(x). \tag{8.9}$$

Man wenden die Kettenregel (8.9) an um die Ableitungen von komplizierteren Funktionen zu berechnen. Man stellt eine gegebene Funktion $F(x)$ in der Form von Verkettung

$$F(x) = g(y) \text{ mit } y = f(x)$$

dar, und dann gilt

$$F'(x) = g'(y) f'(x) = g'(f(x)) f'(x).$$

Beispiel. Betrachten wir die Funktion $F(x) = e^{cx}$ wobei $c \in \mathbb{R}$. Wir stellen F als eine Verkettung dar:

$$F(x) = e^y \text{ mit } y = cx.$$

Nach der Kettenregel erhalten wir

$$F'(x) = (e^y)'(cx)' = e^y c = ce^{cx}.$$

Insbesondere, für $a > 0$, haben wir $a^x = e^{x \ln a}$ und somit

$$\boxed{(a^x)' = a^x \ln a.}$$

Die Kettenregel lässt sich für Komposition von mehreren Funktionen verallgemeinern wir folgt. Sei

$$F = h \circ g \circ f,$$

d.h.

$$F(x) = h(z) \text{ mit } z = g(y) \text{ und } y = f(x).$$

Dann gilt

$$\boxed{F'(x) = h'(z)\, g'(y)\, f'(x)} = h'(g(f(x)))\, g'(f(x))\, f'(x), \tag{8.10}$$

da $F = h \circ (g \circ f)$, d.h.

$$F(x) = h(z) \quad \text{mit } z = g \circ f,$$

woraus folgt

$$F'(x) = h'(z)\,(g \circ f)'(x) = h'(z)\, g'(y)\, f'(x).$$

Ähnliche Regel gilt für Verkettung von mehreren Funktionen.

Beispiel. Die Funktion $F(x) = e^{\sin x^2}$ ist die Verkettung dreier Funktionen:

$$F(x) = e^z \quad \text{mit} \quad z = \sin y \quad \text{und} \quad y = x^2.$$

Somit erhalten wir

$$F'(x) = (e^z)'\,(\sin y)'\,\left(x^2\right)' = e^z\,(\cos y)\,(2x) = 2xe^{\sin x^2}\cos x^2.$$

Das folgende Lemma brauchen wir für den Beweis des Satzes 8.3.

Lemma 8.4 *Sei f eine Funktion auf Intervall J die in einem $a \in J$ differenzierbar ist. Dann existiert eine Funktion Φ auf J mit den folgenden Eigenschaften:*

(i) $f(x) - f(a) = \Phi(x)(x-a)$ *für alle $x \in J$;*

(ii) $\Phi(a) = f'(a)$

(iii) *Φ ist stetig in a.*

Beweis. Die Eigenschaften (i) und (ii) zwingen die folgende Identität für Φ:

$$\Phi(x) = \begin{cases} \dfrac{f(x) - f(a)}{x-a}, & x \neq a, \\ f'(a), & x = a. \end{cases} \tag{8.11}$$

So, definieren wir Φ mit (8.11). Dann (i) gilt für $x \neq a$ nach Definition, und für $x = a$, da die beiden Seiten von (i) verschwinden. Die Identität (ii) gilt auch nach Definition, und (iii) gilt, da

$$\lim_{x\to a} \Phi(x) = \lim_{x\to a} \frac{f(x) - f(a)}{x-a} = f'(a) = \Phi(a).$$

■

Die Funktion Φ aus Lemma 8.4 heißt die *Verhältnisfunktion* von f an der Stelle a. In der Tat ist $\Phi(x)$ gleich die Steigung der Sekante der Graph von f durch $(a, f(a))$ und $(x, f(x))$ falls $x \neq a$, und die Steigung der Tangente an der Stelle $(a, f(a))$ für $x = a$.

Beweis von Satz 8.3. Im Beweis wir wechseln die Notation: sei f differenzierbar in $a \in A$ und g – in $b = f(a) \in B$. Setzen wir $F = g \circ f$ und beweisen, dass

$$F'(a) = g'(b) f'(a).$$

Sei Φ die Verhältnisfunktion von f in a und Ψ – die Verhältnisfunktion von g in b. Dann haben wir

$$\begin{aligned} f(x) - f(a) &= \Phi(x)(x-a) \quad \forall\, x \in A \\ g(y) - g(b) &= \Psi(y)(y-b) \quad \forall y \in B, \end{aligned} \tag{8.12}$$

Einsetzen in (8.12) $y = f(x)$ und $b = f(a)$ ergibt

$$\begin{aligned} F(x) - F(a) &= g(f(x)) - g(f(a)) \\ &= \Psi(f(x))(f(x) - f(a)) \\ &= \Psi(f(x))\Phi(x)(x-a), \end{aligned}$$

woraus folgt

$$\frac{F(x) - F(a)}{x-a} = \Psi(f(x))\Phi(x)$$

und

$$F'(a) = \lim_{x\to a} \frac{F(x) - F(a)}{x-a} = \lim_{x\to a} \Psi(f(x))\Phi(x).$$

Da $f(x)$ stetig in a ist und $\Psi(y)$ stetig in $b = f(a)$, so erhalten wir nach dem Satz 7.6, dass die Verkettung $\Psi(f(x))$ stetig in a ist. Da $\Phi(x)$ auch stetig in a ist, so gilt

$$F'(a) = \Psi(f(a))\Phi(a) = \Psi(b)\Phi(a) = g'(b) f'(a),$$

was zu beweisen war. ■

8.5 Ableitung der inversen Funktion

Satz 8.5 (Ableitung der inversen Funktion) *Sei f eine stetige streng monotone Funktion auf einem Intervall J, so dass die inverse Funktion f^{-1} auf dem Intervall $I = f(J)$ definiert ist. Ist f differenzierbar in $x \in J$ und $f'(x) \neq 0$, so ist f^{-1} differenzierbar in $y = f(x)$ und es gilt*

$$\boxed{\left(f^{-1}\right)'(y) = \frac{1}{f'(x)}}. \tag{8.13}$$

Da $y = f(x)$ äquivalent zu $x = f^{-1}(y)$ ist, so können wir (8.13) wie folgt umschreiben:

$$\left(f^{-1}\right)'(y) = \frac{1}{f'\left(f^{-1}(y)\right)}$$

und

$$\left(f^{-1}\right)'(f(x)) = \frac{1}{f'(x)}.$$

Beispiel. Sei $f(x) = e^x$ auf $J = \mathbb{R}$. Die inverse Funktion von e^x ist $\ln y$ auf $I = (0, +\infty)$. Nach (8.13) erhalten wir für $y = e^x$

$$(\ln y)' = \frac{1}{(e^x)'} = \frac{1}{e^x} = \frac{1}{y},$$

d.h.

$$\boxed{(\ln y)' = \frac{1}{y}}.$$

Beispiel. Betrachten wir $f(x) = \sin x$ auf $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ so dass $f^{-1}(y) = \arcsin y$ wobei $y \in [-1, 1]$. Es folgt, dass für $y = \sin x$ gilt

$$(\arcsin y)' = \left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{\cos x}.$$

Da auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ ist $\cos x > 0$ (insbesondere $\cos x \neq 0$) und

$$\cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2},$$

so folgt es, dass für $y \in (-1, 1)$

$$\boxed{(\arcsin y)' = \frac{1}{\sqrt{1 - y^2}}}.$$

Analog beweist man, dass

$$\boxed{(\arccos y)' = -\frac{1}{\sqrt{1 - y^2}}}.$$

13.07.18 H1

Beweis von dem Satz 8.5. Die inverse Funktion f^{-1} existiert und ist stetig auf I nach dem Satz 7.11. Um (8.13) zu beweisen, wir wechseln die Notation: sei f differenzierbar an $a \in J$ und $f'(a) \neq 0$. Beweisen wir, dass f^{-1} differenzierbar in $b = f(a)$ ist und

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

Sei Φ die Verhältnisfunktion von f in a, so dass für alle $x \in J$

$$f(x) - f(a) = \Phi(x)(x - a). \tag{8.14}$$

Setzen wir in (8.14) $y = f(x)$, $b = f(a)$ ein und erhalten, dass für alle $y \in I$

$$y - b = \Phi\left(f^{-1}(y)\right)\left(f^{-1}(y) - f^{-1}(b)\right)$$

und

$$\frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \frac{1}{\Phi\left(f^{-1}(y)\right)}.$$

Da die Funktion f^{-1} stetig auf J ist und Φ stetig in $a = f^{-1}(b)$ ist, so erhalten wir, dass

$$\lim_{y\to b} \Phi\left(f^{-1}(y)\right) = \Phi\left(f^{-1}(b)\right) = \Phi(a) = f'(a),$$

woraus folgt

$$\left(f^{-1}\right)'(b) = \lim_{y\to b} \frac{f^{-1}(y) - f^{-1}(b)}{y-b} = \lim_{y\to b} \frac{1}{\Phi\left(f^{-1}(y)\right)} = \frac{1}{f'(a)},$$

was zu beweisen war. ∎

Bemerkung. Die Tangente zum Graph $y = f(x)$ der Funktion f am Punkt (a, b) hat die Gleichung

$$y - b = A(x-a),$$

wobei $A = f'(a)$ die Steigung ist. Für die inverse Funktion werden die Rollen von x und y vertauscht, so dass der Graph von f^{-1} die Gleichung $x = f^{-1}(y)$ hat und die Gleichung der Tangente zum f^{-1} ist

$$x - a = \frac{1}{A}(y-b),$$

vorausgesetzt $A \neq 0$. Somit ist die Steigung der Tangente gleich $\frac{1}{A}$, was die Identität (8.13) erklärt .

8.6 Weitere Beispiele

Beispiel. Für $f(x) = \tan x$ auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ gibt es die inverse Funktion $f^{-1}(y) = \arctan y$ im Definitionsbereich $y \in (-\infty, +\infty)$ (Aufgabe 115). Somit gilt es für $y = \tan x$

$$(\arctan y)' = \left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \cos^2 x = \frac{1}{1+\tan^2 x} = \frac{1}{1+y^2},$$

so dass

$$\boxed{(\arctan y)' = \frac{1}{1+y^2}}.$$

Beispiel. Betrachten wir die Funktion

$$f(x) = \sinh x = \frac{e^x - e^{-x}}{2}$$

auf $J = \mathbb{R}$. Da die beiden Funktion e^x und $-e^{-x}$ streng monoton steigend sind, so ist $\sinh x$ auch streng monoton steigend auf $\mathbb{R}$. Somit existiert die inverse Funktion $\sinh^{-1} y$ im Definitionsbereich $I = \sinh(\mathbb{R}) = \mathbb{R}$ (da $\sinh x \to +\infty$ für $x \to +\infty$ und $\sinh x \to -\infty$ für $x \to -\infty$).

Für die inverse Funktion gilt

$$\left(\sinh^{-1} y\right)' = \frac{1}{\cosh x},$$

wobei $\sinh x = y$. Da

$$\cosh^2 x - \sinh^2 x = 1$$

(Aufgabe 109) und somit

$$\cosh x = \sqrt{1 + \sinh^2 x} = \sqrt{1 + y^2},$$

so erhalten wir, dass

$$\left(\sinh^{-1} y\right)' = \frac{1}{\sqrt{1 + y^2}}.$$

Beispiel. Betrachten wir die Potenzfunktion

$$f(x) = x^a$$

wobei $x > 0$ und $a \in \mathbb{R}$. Nach Definition haben wir

$$f(x) = e^{a \ln x},$$

so dass f eine Verkettung ist:

$$f(x) = e^y \quad \text{mit} \quad y = a \ln x.$$

Nach der Kettenregel erhalten wir

$$(x^a)' = (e^y)' (a \ln x)' = e^y \frac{a}{x} = x^a \frac{a}{x} = a x^{a-1},$$

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 für $n \in \mathbb{N}$

$$\left(\sqrt[n]{x}\right)' = \left(x^{1/n}\right)' = \frac{1}{n} x^{\frac{1}{n} - 1} = \frac{1}{n} x^{-\frac{n-1}{n}}.$$

Beispiel. Für die Funktion $f(x) = \sqrt{x^2 + 1}$ haben wir

$$f(x) = \sqrt{y} \text{ mit } y = x^2 + 1$$

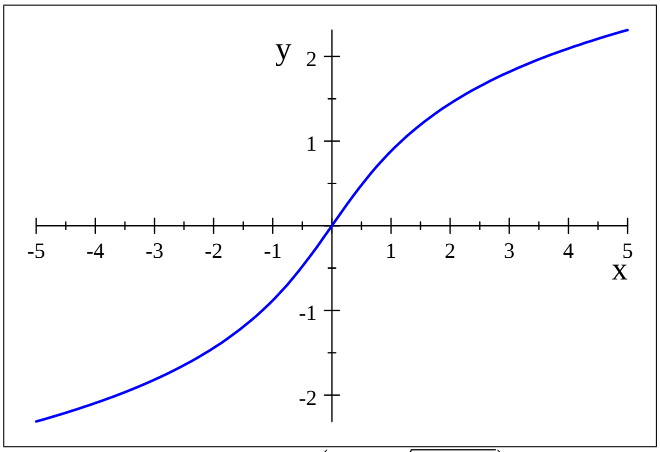
und somit

$$f'(x) = (\sqrt{y})' \left(x^2 + 1\right)' = \frac{1}{2\sqrt{y}} 2x = \frac{x}{\sqrt{x^2 + 1}}.$$

Beispiel. Leiten wir die folgende Funktion ab:

$$f(x) = \ln\left(x + \sqrt{x^2 + 1}\right)$$

im Definitionsbereich $J = \mathbb{R}$ (da $x + \sqrt{x^2 + 1} > x + |x| \geq 0$).



Funktion $\ln\left(x+\sqrt{x^2+1}\right)$

Da

$$f(x) = \ln y \quad \text{mit} \quad y = x+\sqrt{x^2+1},$$

so erhalten wir

$$\begin{aligned} f'(x) &= (\ln y)'\left(x+\sqrt{x^2+1}\right)' = \frac{1}{y}\left(1+\frac{x}{\sqrt{x^2+1}}\right) \\ &= \frac{1}{x+\sqrt{x^2+1}}\frac{x+\sqrt{x^2+1}}{\sqrt{x^2+1}} = \frac{1}{\sqrt{x^2+1}}. \end{aligned}$$

Somit gilt

$$\boxed{\left(\ln\left(x+\sqrt{x^2+1}\right)\right)' = \frac{1}{\sqrt{x^2+1}}}.$$

Wir haben oberhalb gesehen, dass auch

$$\left(\sinh^{-1} x\right)' = \frac{1}{\sqrt{x^2+1}}.$$

Die Übereinstimmung von Ableitungen von $\sinh^{-1} x$ und $\ln\left(x+\sqrt{x^2+1}\right)$ ist kein Zufall, da diese zwei Funktionen identisch gleich sind (siehe Aufgabe 114).

Beispiel. (*Logarithmische Ableitung*) Sei f eine differenzierbare Funktion auf einem Intervall J und sei $f(x) > 0$ für alle $x \in J$. Dann ist auch die Funktion $\ln f(x)$ wohldefiniert, und mit Hilfe von der Substitution $y = f(x)$

$$(\ln f(x))' = (\ln y)' f'(x) = \frac{1}{y} f'(x) = \frac{f'(x)}{f(x)},$$

so dass

$$\boxed{(\ln f(x))' = \frac{f'(x)}{f(x)}}.$$

Die Funktion $(\ln f(x))'$ heißt die *logarithmische Ableitung* der Funktion f. Häufig ist es einfacher $(\ln f)'$ zu bestimmen als f'. Dann bestimmt man f' durch

$$f'(x) = f(x)(\ln f(x))'.$$

Zum Beispiel, bestimmen wir die Ableitung der Funktion

$$f(x) = x^x$$

für $x \in (0, +\infty)$. Wir haben

$$\ln f(x) = x \ln x$$

und somit

$$(\ln f(x))' = (x \ln x)' = (x)' \ln x + x (\ln x)' = \ln x + \frac{x}{x} = \ln x + 1,$$

woraus folgt

$$(x^x)' = x^x (\ln x + 1).$$

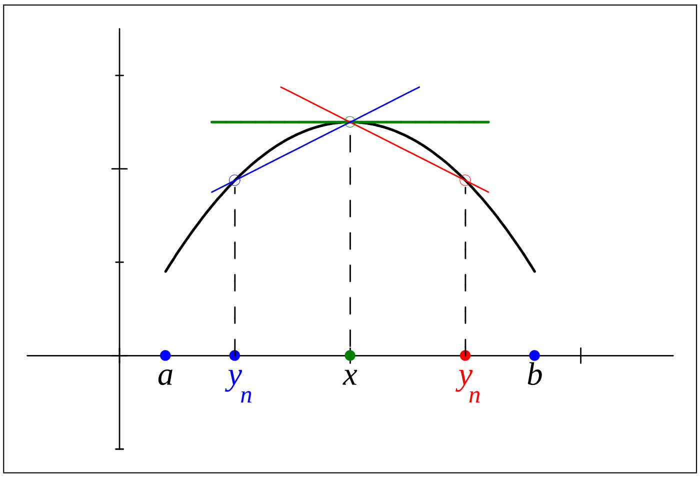
8.7 Sätze von Fermat, Rolle und Lagrange

Hauptsatz 8.6 (Satz von Fermat) *Sei f eine Funktion auf einem offenen Intervall (a,b) und sei $x \in (a,b)$ eine Maximumstelle (bzw Minimumstelle) von f. Ist f in x differenzierbar, so gilt $f'(x) = 0$.*

Die geometrische Bedeutung dieser Aussage ist wie folgt: die Steigung der Tangente an der Maximumstelle (Minimumstelle) von f ist 0, d.h. die Tangente an dieser Stelle waagerecht ist.

Beweis. Für jede Folge $\{y_n\}$ aus (a,b) mit $y_n \to x$ haben wir

$$f'(x) = \lim_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x}.$$



Die rechte Sekante (rot), die linke Sekante (blau), und die Tangente (grün), die an der Maximumstelle waagerecht ist.

Da x eine Maximumstelle ist, so gilt $f(y_n) \leq f(x)$ für alle n. Somit erhalten wir für jede Folge $\{y_n\}$ aus (x,b), dass

$$f'(x) = \lim_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} \leq 0$$

und für jede Folge $\{y_n\}$ aus (a, x), dass

$$f'(x) = \lim_{n\to\infty} \frac{f(y_n) - f(x)}{y_n - x} \geq 0$$

woraus folgt $f'(x) = 0$.

Der Fall von Minimumstelle wird analog behandelt. $\blacksquare$

Korollar 8.7 *Sei f eine stetige Funktion auf einem kompakten Intervall $[a, b]$ mit $a < b$, die auf (a, b) differenzierbar ist. Ist x eine Maximumstelle oder Minimumstelle von f auf $[a, b]$, so gilt entweder $f'(x) = 0$ oder $x = a$ oder $x = b$.*

Bemerken wir, dass die Maximum- und Minimumstellen von f auf $[a, b]$ existieren nach dem Satz 7.9.

Beweis. Gilt $x \in (a, b)$ so gilt $f'(x) = 0$ nach dem Satz 8.6. Sonst $x = a$ oder $x = b$. $\blacksquare$

Die Menge

$$K_f = \{x \in (a, b) : f'(x) = 0\} \cup \{a, b\}$$

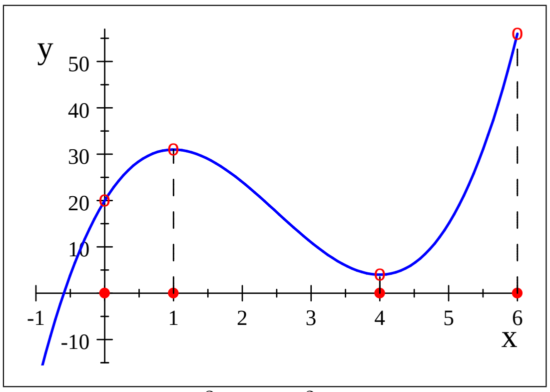
heißt die *kritische* Menge von f. Nach Korollar 8.7 liegen die Maximum- und Minimumstellen von f in K_f. Es folgt, dass

$$\max_{[a,b]} f = \max_{K_f} f \quad \text{und} \quad \min_{[a,b]} f = \min_{K_f} f.$$

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 von 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 8.8 (Satz von Rolle) *Sei f eine stetige Funktion auf einem kompakten 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 7.9 (Extremwertsatz) existieren $\max_{[a,b]} f$ und $\min_{[a,b]} f$. Sei c_1 eine Maximumstelle von f und c_2 – eine Minimumstelle von f. Liegt eines von c_1, c_2 im (a, b) so verschwindet f' an dieser Stelle nach dem Satz 8.6. 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 8.9 (Mittelwertsatz von Lagrange) *Sei f eine stetige Funktion auf einem kompakten Intervall $[a, b]$ mit $a < b$, die auf (a, b) differenzierbar ist. Dann existiert ein $c \in (a, b)$ mit*

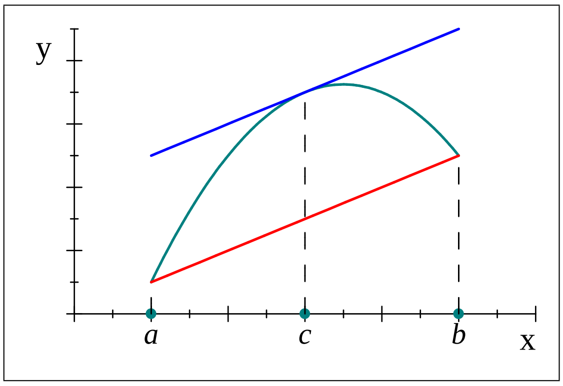
$$f(b) - f(a) = f'(c)(b - a). \tag{8.15}$$

Bemerkung. Im Fall $f(a) = f(b)$ ergibt (8.15) dass $f'(c) = 0$, was den Satz 8.8 impliziert.

Bemerkung. Die Identität (8.15) ist äquivalent zu

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

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 8.9 folgendes: die Steigung dieser Sekante ist gleich die Steigung der Tangente an einer Mittelstelle $c \in (a, b)$, d.h. es gibt eine Tangente parallel zur Sekante.



Die Tangente an $(c, f(c))$ (blau) und die Sekante durch $(a, f(a)), (b, f(b))$ (rot) haben die gleichen Steigungen

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

$$h(a) = f(a) - \frac{f(b) - f(a)}{b - a}(a - a) = f(a)$$

und

$$h(b) = f(b) - \frac{f(b) - f(a)}{b - a}(b - a) = f(a),$$

d.h. $h(a) = h(b)$. Nach dem Satz 8.8, es gibt ein $c \in (a, b)$ mit $h'(c) = 0$. Da

$$h'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

es folgt daraus, dass

$$f'(c) = \frac{f(b) - f(a)}{b - a},$$

was zu beweisen war. ∎

8.8 Untersuchung von Funktion mit Hilfe von f'

8.8.1 Konstantentest

Satz 8.10 (Konstantentest) *Sei f eine differenzierbare Funktion auf einem Intervall J. Dann gilt $f = \text{const}$ auf J genau dann wenn $f'(x) = 0$ für alle $x \in J$.*

Beweis. Gilt $f = \text{const}$, so gilt auch $f'(x) = 0$ für alle $x \in J$. Für die Rückrichtung beweisen wir, dass $f(a) = f(b)$ für alle verschiedene $a, b \in J$. Nach Satz 8.9 erhalten wir, für ein $c \in (a, b)$,

$$f(a) - f(b) = f'(c)(a - b).$$

Da $f'(c) = 0$, so erhalten wir $f(a) = f(b)$. ∎

Der Konstantentest wird häufig in der folgenden Situation benutzt. Gilt

$$f'(x) = g'(x) \quad \forall x \in J,$$

so folgt es, dass

$$f(x) = g(x) + \text{const} \ \forall x \in J,$$

da für die Funktion $f - g$ gilt $(f - g)' = 0$ und somit $f - g = \text{const}$ auf J.

Beispiel. Bestimmen wir alle Funktionen f mit $f'(x) = x$ für alle $x \in \mathbb{R}$. Bemerken wir zunächst, dass die Funktion $g(x) = \frac{x^2}{2}$ die Gleichung $g' = x$ erfüllt. Es folgt, dass für beliebige Funktion f mit $f' = x$ gilt

$$f(x) = \frac{x^2}{2} + C,$$

wobei C eine beliebige Konstante ist. Somit lässt sich die Funktion $f(x)$ durch ihre Ableitung wiederherstellen.

8.8.2 Monotonietest

Satz 8.11 (Monotonietest) *Sei J ein beliebiges Intervall in $\mathbb{R}$ mit den Grenzen $a, b \in \overline{\mathbb{R}}$, $a < b$. Sei f eine stetige Funktion auf J, die auf (a, b) differenzierbar ist. Gilt $f'(x) \geq 0$ für alle $x \in (a, b)$, so ist f monoton steigend auf J; gilt $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton steigend auf J.*

Gilt $f'(x) \leq 0$ für alle $x \in (a, b)$ so ist f monoton fallend auf J; gilt $f'(x) < 0$ für alle $x \in (a, b)$, so ist f streng monoton fallend auf J.

Beweis. Betrachten wir zwei beliebige Werte $x < y$ in J. Die Funktion f ist stetig im $[x, y]$ und differenzierbar im $(x, y) \subset (a, b)$. Nach dem Satz 8.9 existiert ein $c \in (x, y)$ mit

$$f(y) - f(x) = f'(c)(y - x).$$

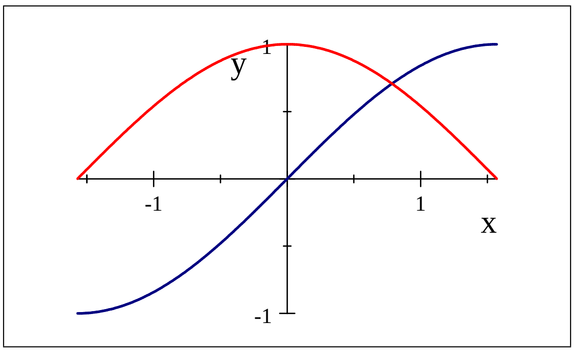
Gilt $f' \geq 0$ auf (a, b), so gilt $f'(c) \geq 0$ und somit $f(y) \geq f(x)$, d.h. f monoton steigend ist. Gilt $f' > 0$ auf (a, b), so haben wir $f'(c) > 0$ und somit $f(y) > f(x)$, d.h. f streng monoton steigend ist. Die Aussagen mit $f' \leq 0$ und $f' < 0$ werden analog bewiesen. ■

Bemerkung. Ist f monoton steigend auf J, so gilt für alle $x \in (a, b)$

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \geq 0.$$

Somit ist die Bedingung $f' \geq 0$ auf (a, b) äquivalent zur Bedingung, dass f monoton steigend auf J ist. 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. Für die Funktion $f(x) = \sin x$ auf $J = [-\pi/2, \pi/2]$ gilt $f'(x) = \cos x > 0$ für alle $x \in (-\pi/2, \pi/2)$. Es folgt, dass $\sin x$ streng monoton steigend auf $[-\pi/2, \pi/2]$ ist, was wir schon nach dem Satz 7.14 wissen.



Die Graphe von Funktionen $\sin x$ (blau) und $\cos x$ (rot) auf $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

18.07.18 H13

Beispiel. Beweisen wir die *Ungleichung vom arithmetischen und geometrischen Mittel.* Für zwei nichtnegative reelle Zahlen a, b gilt

$$\frac{a + b}{2} \geq \sqrt{ab},$$

da
$$\frac{a+b}{2} - \sqrt{ab} = \frac{1}{2}\left(a+b-2\sqrt{ab}\right) = \frac{1}{2}\left(\sqrt{a}-\sqrt{b}\right)^2 \geq 0.$$
Beweisen wir, dass für drei nichtnegative reelle Zahlen a, b, c gilt
$$\frac{a+b+c}{3} \geq \sqrt[3]{abc}. \tag{8.16}$$
Setzen wir $x = \sqrt[3]{c}$ und schreiben (8.16) wie folgt um:
$$x^3 + a + b \geq 3x\sqrt[3]{ab},$$
oder, mit $d = \sqrt[3]{ab}$,
$$x^3 - 3dx + a + b \geq 0.$$
Es reicht somit zu beweisen, dass die Funktion
$$f(x) = x^3 - 3dx + a + b$$
auf $[0, +\infty)$ nicht-negative ist. Da
$$f'(x) = 3x^2 - 3d = 3\left(x^2 - d\right),$$
so erhalten wir den kritischen Punkt $x = \sqrt{d}$, und somit
$$K_f = \left\{0, \sqrt{d}, +\infty\right\}.$$
Auf dem Intervall $\left(0, \sqrt{d}\right)$ gilt $f' < 0$ so dass f streng monoton fallend ist. Auf dem Intervall $\left(\sqrt{d}, +\infty\right)$ gilt $f' > 0$ so dass f streng monoton steigend ist. Somit ist $x = \sqrt{d}$ eine Minimumstelle von f.

Wir haben
$$f\left(\sqrt{d}\right) = \left(\sqrt{d}\right)^3 - 3d\sqrt{d} + a + b = a + b - 2\sqrt{ab} \geq 0,$$
da $d\sqrt{d} = \left(\sqrt{d}\right)^3 = \sqrt{d^3} = \sqrt{ab}$. Daraus folgt $f(x) \geq 0$ für alle $x \geq 0$, was zu beweisen war.

Analog kann man per Induktion nach $n \in \mathbb{N}$ beweisen, dass für beliebige Folge $\{a_k\}_{k=1}^n$ von nichtnegativen reellen Zahlen gilt
$$\frac{a_1 + a_2 + ... + a_n}{n} \geq \sqrt[n]{a_1 a_2 ... a_n}.$$

8.9 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 lassen sich diese Ausdrücke mit Hilfe von dem folgenden Satz lösen.

Hauptsatz 8.12 (Regel von l'Hôspital) *Seien f und g differenzierbare Funktionen auf einem Intervall $J \subset \mathbb{R}$. Angenommen sei, dass für ein $a \in \overline{J}$*
(a) entweder

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0, \tag{8.17}$$

(b) oder

$$\lim_{x\to a} f(x) = \pm\infty \quad \textit{und} \quad \lim_{x\to a} g(x) = \pm\infty. \tag{8.18}$$

Gilt

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = b \in \overline{\mathbb{R}}, \tag{8.19}$$

so gilt auch

$$\lim_{x\to a} \frac{f(x)}{g(x)} = b, \tag{8.20}$$

vorausgesetzt dass $g \neq 0$ und $g' \neq 0$ auf $J \setminus \{a\}$.

Die Regel ist somit sehr einfach: um die unbestimmte Ausdrücke $\frac{0}{0}$ bzw $\frac{\infty}{\infty}$ zu lösen, sollen die Funktionen f und g im Bruch $\frac{f}{g}$ durch ihre Ableitungen ersetzt werden.

Beispiel. 1. Bestimmen wir

$$\lim_{x\to 0} \frac{\ln(1+x)}{\sin x}.$$

Das ist unbestimmter Ausdruck der Form $\frac{0}{0}$. Nach dem Satz 8.12 mit $J = (-1, 1)$ und $a = 0$ erhalten wir

$$\lim_{x\to 0} \frac{\ln(1+x)}{\sin x} = \lim_{x\to 0} \frac{(\ln(1+x))'}{(\sin x)'} = \lim_{x\to 0} \frac{1/(1+x)}{\cos x} = 1. \tag{8.21}$$

Um rigorous zu sein, man soll die Gültigkeit der Gleichheiten in (8.21) rückwärts beweisen: da $\lim_{x\to 0} \frac{(\ln(1+x))'}{(\sin x)'}$ existiert und gleich 1 ist, so gilt nach dem Satz 8.12 auch

$$\lim_{x\to 0} \frac{\ln(1+x)}{\sin x} = 1.$$

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

$$\lim_{x\to +\infty} \frac{e^x}{x^2},$$

was unbestimmter Ausdruck der Form $\frac{\infty}{\infty}$ ist. Versuchen wir den Satz 8.12 mit $J = (0, +\infty)$ und $a = +\infty$ zu benutzen:

$$\lim_{x\to +\infty} \frac{e^x}{x^2} = \lim_{x\to +\infty} \frac{(e^x)'}{(x^2)'} = \lim_{x\to +\infty} \frac{e^x}{2x}, \tag{8.22}$$

und erhalten, dass der Grenzwert in der rechten Seite wieder unbestimmter Ausdruck der Form $\frac{\infty}{\infty}$ ist. Benutzen wir wieder den Satz 8.12 und erhalten

$$\lim_{x\to +\infty} \frac{e^x}{2x} = \lim_{x\to +\infty} \frac{(e^x)'}{(2x)'} = \lim_{x\to +\infty} \frac{e^x}{2} = +\infty. \tag{8.23}$$

Somit erhalten wir nach zwei Anwendungen von der Regel von l'Hôspital, dass

$$\lim_{x\to +\infty} \frac{e^x}{x^2} = +\infty.$$

Analog beweist man per Induktion nach n, dass

$$\lim_{x\to +\infty} \frac{e^x}{x^n} = +\infty \tag{8.24}$$

für alle $n \in \mathbb{N}$. Das gleiche Ergebnis kann man auch mit Hilfe von der Exponentialreihe erhalten (Aufgabe 117).

4. Bestimmen wir

$$\lim_{x\to 0} x \ln x$$

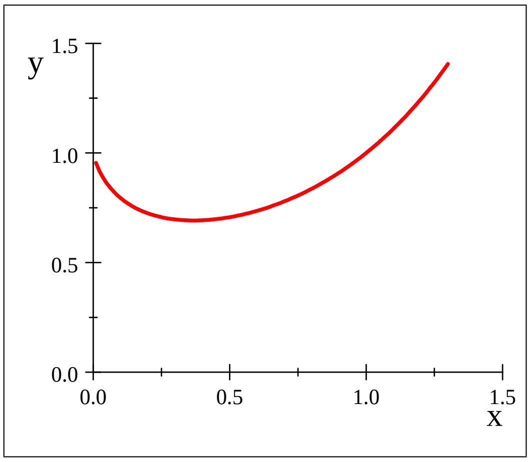
wobei $J = (0, +\infty)$. Da $\ln x \to -\infty$ für $x \to 0$, so haben wir unbestimmten Ausdruck der Form $0 \cdot \infty$. Um ihn zu lösen, stellen wir den Grenzwert in der Form von Quotient dar:

$$\lim_{x\to 0} \frac{\ln x}{1/x},$$

was unbestimmter Ausdruck der Form $\frac{\infty}{\infty}$ ist. Nach der Regel von l'Hôspital erhalten wir

$$\lim_{x\to 0} x \ln x = \lim_{x\to 0} \frac{\ln x}{1/x} = \lim_{x\to 0} \frac{(\ln x)'}{(1/x)'} = -\lim_{x\to 0} \frac{1/x}{1/x^2} = -\lim_{x\to 0} x = 0. \tag{8.25}$$

5. Bestimmen wir
$$\lim_{x\to 0} x^x,$$
wobei $J = (0, +\infty)$. Das ist unbestimmter Ausdruck der Form 0^0. Um ihn zu lösen, betrachten wir den Logarithmus der Funktion x^x:
$$\lim_{x\to 0} \ln x^x = \lim_{x\to 0} x \ln x = 0,$$
wo wir (8.25) benutzt haben. Nach dem Satz 7.5 erhalten wir
$$\lim_{x\to 0} x^x = \lim_{x\to 0} e^{x\ln x} = \lim_{y\to 0} e^y = 1.$$



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. Sei $J = (0, +\infty)$. 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
$$\lim_{x\to+\infty} \left(1 + \frac{1}{x}\right)^x = \lim_{x\to+\infty} e^{x\ln\left(1+\frac{1}{x}\right)} = \lim_{y\to 1} e^y = e.$$

Für den Beweis des Satzes 8.12 brauchen wir die folgende Erweiterung des Mittelwertsatzes von Lagrange.

Hauptsatz 8.13 (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*
$$g'(c)(f(b) - f(a)) = f'(c)(g(b) - g(a)). \tag{8.26}$$

Der Mittelwertsatz von Lagrange (Satz 8.9) ist ein spezieller Fall des Satzes 8.13 für $g(x) = x$ da in diesem Fall (8.26) 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 dem Satz von Rolle (Satz 8.8) 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 (8.26). ■

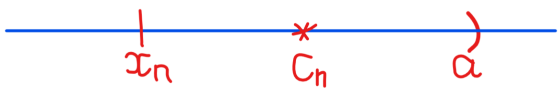
Beweis von Satz 8.12(a). Zunächst betrachten wir den Fall $a \in \mathbb{R}$ (im Gegenteil zum $a = \pm\infty$). Ist $a \in J$ so haben wir

$$f(a) = g(a) = 0.$$

Ist $a \notin J$, so erweitern wir die Funktionen f und g auf der Stelle a (falls $a \notin J$) indem wir setzen $f(a) = g(a) = 0$. Dann sind die Funktionen f und g stetig auf dem Intervall $J \cup \{a\}$.

Sei $\{x_n\}_{n\in\mathbb{N}}$ eine beliebige Folge aus $J \setminus \{a\}$ mit $x_n \to a$. Für jedes n ist f stetig in $[x_n, a]$ und differenzierbar in (x_n, a). Nach dem Mittelwertsatz von Cauchy (Satz 8.13) existiert ein $c_n \in (x_n, a)$ mit

$$\frac{f(x_n)}{g(x_n)} = \frac{f(x_n) - f_n(a)}{g(x_n) - g_n(a)} = \frac{f'(c_n)}{g'(c_n)}.$$



Da $c_n \to a$, so erhalten wir nach (8.19)

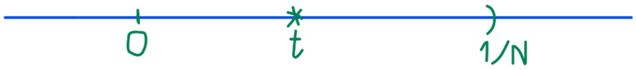
$$\lim_{n\to\infty} \frac{f'(c_n)}{g'(c_n)} = b,$$

woraus folgt

$$\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = b,$$

und somit (8.20).

Sei jetzt $a = +\infty$. Dann ist $+\infty$ eine Grenze von J. Fixieren wir ein $N \in J$ so dass $J \supset (N, +\infty)$. Betrachten wir die Funktionen $F(t) = f\left(\frac{1}{t}\right)$ und $G(t) = g\left(\frac{1}{t}\right)$ auf dem Intervall $(0, 1/N)$.



Wir haben

$$\lim_{t\to 0} F(t) = \lim_{x\to+\infty} f(x) = 0 \quad \text{und} \quad \lim_{t\to 0} G(t) = \lim_{x\to+\infty} g(x) = 0.$$

Auch gilt

$$\lim_{t\to 0} \frac{F'(t)}{G'(t)} = \lim_{t\to 0} \frac{f'\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right)}{g'\left(\frac{1}{t}\right)\left(-\frac{t}{t^2}\right)} = \lim_{t\to 0} \frac{f'\left(\frac{1}{t}\right)}{g'\left(\frac{1}{t}\right)} = \lim_{x\to+\infty} \frac{f'(x)}{g'(x)} = b.$$

Anwendung von dem obigen Fall der Regel von l'Hôspital ergibt

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = \lim_{t\to 0} \frac{F(t)}{G(t)} = b.$$

Der Fall $a = -\infty$ wird analog bewiesen. ■

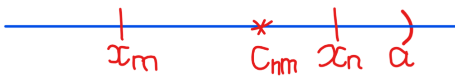
*** Beweis von Satz 8.12**(b)**.**

Nehmen wir an, dass $b \in \mathbb{R}$ (der Fall $b = \pm\infty$ ist ähnlich). Wir müssen beweisen, dass für jede Folge $\{x_n\}_{n\in\mathbb{N}}$ aus $J \setminus \{a\}$ mit $x_n \to a$ gilt

$$\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = b.$$

Wir beweisen dies im Fall wenn alle $x_n < a$ (der Fall mit $x_n > a$ ist analog). Für alle $m, n \in \mathbb{N}$ mit verschiedenen x_n, x_m ist f im $[x_m, x_n]$ differenzierbar. Nach dem Mittelwertsatz von Cauchy existiert ein $c_{nm} \in (x_m, x_n)$ mit

$$f'(c_{nm})\,(g(x_n) - g(x_m)) = g'(c_{nm})\,(f(x_n) - f(x_m)).$$



Nach Voraussetzung gilt $g'(c_{nm}) \neq 0$. Auch gilt $g(x_n) \neq g(x_m)$, da anderenfalls nach dem Satz von Rolle (Satz 8.8) die Ableitung g' eine Nullstelle im $J \setminus \{a\}$ hat. Dividieren durch $g'(c_{nm})\,(g(x_n) - g(x_m))$ ergibt

$$\frac{f'(c_{nm})}{g'(c_{nm})} = \frac{f(x_n) - f(x_m)}{g(x_n) - g(x_m)} = \frac{\frac{f(x_n)}{g(x_n)} - \frac{f(x_m)}{g(x_n)}}{1 - \frac{g(x_m)}{g(x_n)}},$$

woraus folgt

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(c_{nm})}{g'(c_{nm})}\left(1 - \frac{g(x_m)}{g(x_n)}\right) + \frac{f(x_m)}{g(x_n)}. \tag{8.27}$$

Nach Voraussetzung (8.19) gilt

$$\lim_{x\to a}\frac{f'(x)}{g'(x)}=b,$$

d.h. für jedes $\varepsilon>0$ existiert eine Umgebung U von a mit

$$\left|\frac{f'(x)}{g'(x)}-b\right|<\varepsilon \text{ für alle } x\in J\cap U.$$

Für fast alle n,m gilt $x_n,x_m\in U\cap J$ woraus folgt, dass auch $c_{nm}\in U\cap J$ und somit

$$b-\varepsilon<\frac{f'(c_{nm})}{g'(c_{nm})}<b+\varepsilon.$$

Für fixiertes m erhalten wir für $n\to\infty$ dass $x_n\to a$ und $g(x_n)\to+\infty$, so dass

$$\frac{g(x_m)}{g(x_n)}\to 0 \quad \text{und} \quad \frac{f(x_m)}{g(x_n)}\to 0.$$

Es folgt aus (8.27), dass

$$\limsup_{n\to\infty}\frac{f(x_n)}{g(x_n)}\le b+\varepsilon$$

und

$$\liminf_{n\to\infty}\frac{f(x_n)}{g(x_n)}\ge b-\varepsilon.$$

Da diese Ungleichungen für alle $\varepsilon>0$ gelten, so erhalten wir

$$\limsup_{n\to\infty}\frac{f(x_n)}{g(x_n)}=\liminf_{n\to\infty}\frac{f(x_n)}{g(x_n)}=b,$$

und somit

$$\lim_{n\to\infty}\frac{f(x_n)}{g(x_n)}=b.$$

■

8.10 Landau-Symbol

Definition. Seien $f(x)$ und $g(x)$ zwei Funktionen auf einem Interval $J\subset\mathbb{R}$, und sei $a\in\overline{J}$. Man schreibt

$$f(x)=o(g(x)) \text{ für } x\to a \tag{8.28}$$

($f(x)$ ist klein o von $g(x)$, $f(x)$ ist vernachlässigbar klein gegenüber $g(x)$) falls

$$\lim_{x\to a}\frac{f(x)}{g(x)}=0$$

(vorausgesetzt $g(x)\neq 0$ in $J\setminus\{a\}$). Das Symbol o heißt das *Landau-Symbol.*

Zum Beispiel, es gilt

$$x^2=o(x) \text{ für } x\to 0$$

während

$$x = o\left(x^2\right) \text{ für } x \to +\infty.$$

Auch gilt $x^2 = o\left(e^x\right)$ für $x \to +\infty$.

Beispiel. Beweisen wir folgendes: ist f differenzierbar in $a \in J$, so gilt

$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \quad \text{für } x \to a. \tag{8.29}$$

In der Tat haben wir

$$\frac{f(x) - f(a) - f'(a)(x-a)}{x-a} = \frac{f(x) - f(a)}{x-a} - f'(a) \to 0 \text{ für } x \to a,$$

woraus folgt

$$f(x) - f(a) - f'(a)(x-a) = o(x-a),$$

was äquivalent zu (8.29). Das Glied $o(x-a)$ in (8.30) heißt das *Restglied*. Es lässt sich betrachten als der Approximationsfehler der ungefähren Gleichheit

$$f(x) \approx f(a) + f'(a)(x-a),$$

die für x in der Nähe von a gilt.

8.11 Zweite Ableitung und Taylorformel

Sei f eine differenzierbare Funktion auf einem Intervall J, so dass die Ableitung f' auch eine Funktion auf J ist.

Definition. Ist f' in $a \in J$ differenzierbar, so heißt f *2-fach* differenzierbar in a. Die Ableitung $(f')'(a)$ heißt die *zweite Ableitung* von f in a und wird mit $f''(a)$ bezeichnet. Ist f 2-fach differenzierbar in jedem $a \in J$, so heißt f 2-fach differenzierbar in J.

Beispiel. Wir haben

$$\begin{aligned}
(e^x)'' &= (e^x)' = e^x \\
(\sin x)'' &= (\cos x)' = -\sin x \\
(\cos x)'' &= -(\sin x)' = -\cos x \\
(\ln x)'' &= \left(\frac{1}{x}\right)' = -\frac{1}{x^2} \\
(x^a)'' &= \left(ax^{a-1}\right)' = a(a-1)x^{a-2}.
\end{aligned}$$

Hauptsatz 8.14 (Taylorformel 2er Ordnung mit der Restgliedform nach Peano) *Sei f differenzierbar in J und 2-fach differenzierbar in $a \in J$. Dann gilt die asymptotische Identität*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + o\left((x-a)^2\right) \quad \textit{für } x \to a. \tag{8.30}$$

Die Identität (8.29) lässt sich als eine ungefähre Gleichheit betrachten:

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2, \tag{8.31}$$

und das Restglied $o\left((x-a)^2\right)$ ist der Approximationsfehler.

Beweis. Wir brauchen zu beweisen, dass

$$\lim_{x\to a} \frac{f(x) - \left(f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2\right)}{(x-a)^2} = 0,$$

was äquivalent zu

$$\lim_{x\to a} \frac{f(x) - (f(a) + f'(a)(x-a))}{(x-a)^2} = \frac{f''(a)}{2}.$$

Da der Nenner und der Zähler für $x \to a$ verschwinden und wir den unbestimmten Ausdruck $\frac{0}{0}$ bekommen, so erhalten wir nach der Regel von l'Hôspital

$$\lim_{x\to a} \frac{f(x) - (f(a) + f'(a)(x-a))}{(x-a)^2} = \lim_{x\to a} \frac{f'(x) - f'(a)}{2(x-a)} = \frac{1}{2} f''(a),$$

was zu beweisen war. ■

Beispiel. Für $f(x) = \ln x$ erhalten wir

$$(\ln x)' = \frac{1}{x}, \quad (\ln x)'' = -\frac{1}{x^2}$$

und somit

$$\ln x = \ln a + \frac{1}{a}(x-a) - \frac{1}{2a^2}(x-a)^2 + o\left((x-a)^2\right) \text{ für } x \to a.$$

Für $a = 1$ und $x = 1,2$ erhalten wir

$$\ln 1,2 \approx \ln 1 + 0,2 - \frac{1}{2} \cdot 0,2^2 = 0,18$$

während $\ln 1,2 = 0,182\,321556\,793955...$

Beispiel. Für $f(x) = \sqrt{x}$ haben wir

$$f'(x) = \frac{1}{2} x^{-1/2} \quad \text{und} \quad f''(x) = -\frac{1}{4} x^{-3/2}.$$

Für jedes $a > 0$ erhalten wir

$$\sqrt{x} = \sqrt{a} + \frac{1}{2\sqrt{a}}(x-a) - \frac{1}{8\sqrt{a}^3}(x-a)^2 + o\left((x-a)^2\right) \text{ für } x \to a.$$

Zum Beispiel, für $a = 25$ und $x = 26$ gilt

$$\sqrt{26} \approx 5 + \frac{1}{2 \cdot 5} - \frac{1}{8 \cdot 125} = 5,099$$

während $\sqrt{26} = 5,099019...$

Der folgende Satz ist eine Verallgemeinerung des Satzes 8.9.

Hauptsatz 8.15 (Taylorformel 2er Ordnung mit der Restgliedform nach Lagrange) *Sei f 2-fach differenzierbar auf $[a,b]$. Dann gibt es ein $c \in (a,b)$ so dass*

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(c)}{2}(b-a)^2.$$

20.07.18 H6

Beweis. Betrachten wir die Funktionen

$$F(x) = f(x) + f'(x)(b-x)$$

und

$$G(x) = (x-b)^2.$$

Nach dem Mittelwertsatz von Cauchy (Satz 8.13), es gibt ein $c \in (a,b)$ mit

$$G'(c)(F(b) - F(a)) = F'(c)(G(b) - G(a)).$$

Da

$$\begin{aligned} F'(x) = f'(x) + f'(x)(b-x)' + f''(x)(b-x) &= f''(x)(b-x) \\ F(a) - F(a) &= f(a) + f'(a)(b-a) - f(b) \\ G'(x) &= 2(x-b) \\ G(b) - G(a) &= -(a-b)^2, \end{aligned}$$

so erhalten wir

$$2(c-b)(f(b) - f(a) - f'(a)(b-a)) = -f''(c)(b-c)(a-b)^2.$$

Es folgt, dass

$$f(b) - f(a) - f'(a)(b-a) = \frac{1}{2}f''(c)(b-a)^2,$$

was zu beweisen war. ■

8.12 Lokale Extrema

Definition. Sei f eine Funktion auf einem offenen Intervall J und sei $a \in J$. Man sagt, dass a eine *lokale Maximumstelle* von f ist (bzw f hat an der Stelle a ein *lokales Maximum*) falls es eine Umgebung $U \subset J$ von a gibt, so dass a eine Maximumstelle von f in U ist. 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 8.16 (*a*) (*Notwendige Bedingung für locales Extremum*) *Sei f eine differenzierbare Funktion auf einem offenen Intervall J. Ist $a \in J$ 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 J. Sei $f'(a) = 0$ für ein $a \in J$. 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 der Ableitung f' sind die kritischen Punkte der Funktion f. Nach dem Satz 8.16 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 dem Satz von Fermat (Satz 8.6) gilt $f'(a) = 0$. Gleiches gilt für eine lokale Minimumstelle.

(b) Nach dem Satz 8.14 haben wir

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + R(x), \tag{8.32}$$

wobei $R(x) = o\left((x-a)^2\right)$ für $x \to a$, d.h.

$$\lim_{x\to a} \frac{R(x)}{(x-a)^2} = 0.$$

Dann für jedes $\varepsilon > 0$ gibt es eine Umgebung $U \subset J$ von a mit

$$\left|\frac{R(x)}{(x-a)^2}\right| < \varepsilon \text{ für alle } x \in U \setminus \{a\} =: U',$$

woraus folgt

$$-\varepsilon(x-a)^2 < R(x) < \varepsilon(x-a)^2 \text{ für alle } x \in U'. \tag{8.33}$$

Sei $f''(a) > 0$. Da $f'(a) = 0$, so erhalten wir aus (8.32) und (8.33), dass für alle $x \in U \setminus \{a\}$,

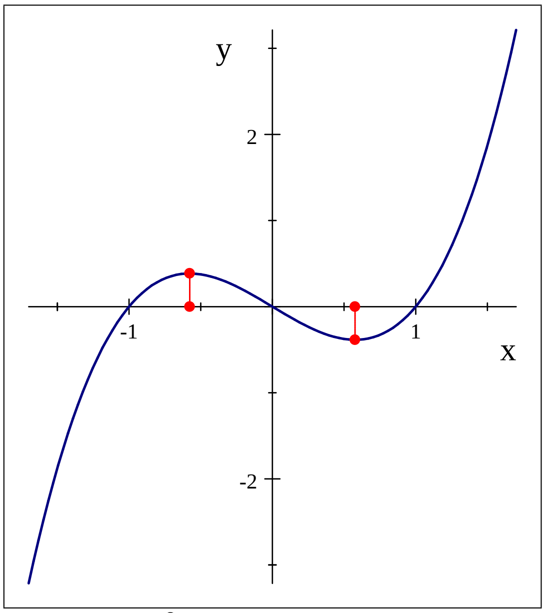
$$\begin{aligned} f(x) &> f(a) + \frac{f''(a)}{2}(x-a)^2 - \varepsilon(x-a)^2 \\ &= f(a) + \left(\frac{f''(a)}{2} - \varepsilon\right)(x-a)^2. \end{aligned}$$

Wählen wir ein $\varepsilon < \frac{1}{2}f''(a)$ und erhalten, dass

$$f(x) > f(a) \quad \text{für alle } x \in U',$$

so dass a eine Minimumstelle von f in U ist. 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.

8.13 Konvexe und konkave Funktionen

Definition. Eine Funktion $f : J \to \mathbb{R}$ auf einem Intervall J heißt *konvex* falls für alle $a, b \in J$ und $t \in (0,1)$ gilt

$$f((1-t)a + tb) \leq (1-t) f(a) + t f(b). \tag{8.34}$$

Die Funktion f heißt *konkav* auf J falls für alle $a, b \in J$ und $t \in (0,1)$ gilt

$$f((1-t)a + tb) \geq (1-t) f(a) + t f(b). \tag{8.35}$$

Diese Definition hat die folgende geometrische Bedeutung. Bezeichnen wir

$$x = (1-t)a + tb \tag{8.36}$$

und beachten, dass

$$t \in (0,1) \Leftrightarrow x \in (a,b).$$

Es folgt aus (8.36)

$$t = \frac{x-a}{b-a},$$

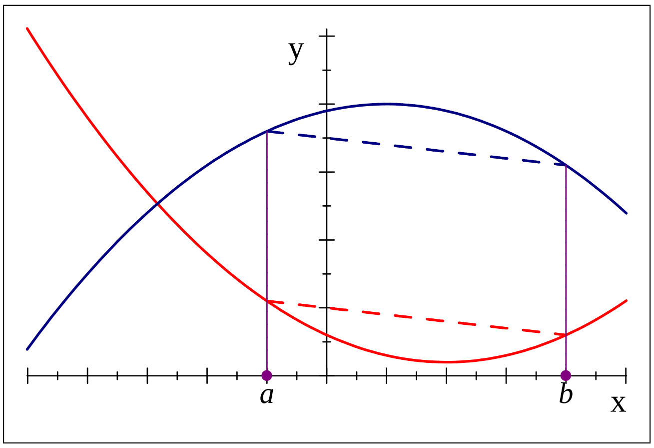
und (8.34) ist äquivalent zu

$$f(x) \leq f(a) + \frac{f(b) - f(a)}{b-a}(x-a),$$

für alle $x \in (a,b)$. Da die Funktion

$$g(x) = f(a) + \frac{f(b) - f(a)}{b-a}(x-a)$$

die Sekante durch $(a, f(a))$ und $(b, f(b))$ bestimmt, so ist die Bedingung (8.34) von Konvexität äquivalent zur Bedingung, dass der Graph der Funktion f auf (a, b) unter der Sekante liegt, für alle $a, b \in J$. Analog ist die Funktion f konkav, wenn der Graph von f auf (a, b) über der Sekante liegt, auch für alle $a, b \in J$.



konvex - rot, konkav - blau

Satz 8.17 *Sei f eine 2-fach differenzierbare Funktion auf einem offenen Intervall $J \subset \mathbb{R}$.*

(a) *Funktion f ist konvex auf J genau dann, wenn $f'' \geq 0$ auf J.*

(b) *Funktion f ist konkav auf J genau dann, wenn $f'' \leq 0$ auf J.*

Beweis. (a) Sei $f'' \geq 0$ auf J. Wir beweisen, dass f konvex ist, d.h. für alle $a, b \in J$ und $t \in (a, b)$ gilt

$$f(x) \leq (1-t) f(a) + t f(b),$$

wobei

$$x = (1-t) a + tb.$$

Nach dem Satz 8.15 haben wir

$$f(a) = f(x) + f'(x)(a-x) + f''(c_1)\frac{(a-x)^2}{2}$$

und

$$f(b) = f(x) + f'(x)(b-x) + f''(c_2)\frac{(b-x)^2}{2}$$

wobei $c_1, c_2 \in (a, b)$. Da $f''(c_1) \geq 0$ und $f''(c_2) \geq 0$, so folgt es

$$\begin{aligned} f(a) &\geq f(x) + f'(x)(a-x) \\ f(b) &\geq f(x) + f'(x)(b-x) \end{aligned}$$

und somit

$$\begin{aligned} (1-t) f(a) + t f(b) \geq\ & (1-t) f(x) + t f(x) \\ & + f'(x)((1-t)(a-x) + t(b-x)) = f(x), \end{aligned}$$

da
$$(1-t)(a-x)+t(b-x)=(1-t)a+tb-x=0.$$

Sei jetzt f konvex auf J. Wir beweisen, dass $f''(x)\geq 0$ für alle $x\in J$. Nach der Konvexität von f gilt für jedes $x\in J$

$$f(x)=f\left(\tfrac{1}{2}(x+h)+\tfrac{1}{2}(x-h)\right)\leq\frac{f(x+h)+f(x-h)}{2}, \tag{8.37}$$

wobei h so klein ist dass $x+h$ und $x-h$ in J liegen. Nach dem Satz 8.14 haben wir

$$f(x+h)=f(x)+f'(x)h+f''(x)\frac{h^2}{2}+o\left(h^2\right)\quad \text{für } h\to 0$$

und

$$f(x-h)=f(x)-f'(x)h+f''(x)\frac{h^2}{2}+o\left(h^2\right)\quad \text{für } h\to 0.$$

Es folgt, dass

$$\frac{f(x+h)+f(x-h)}{2}=f(x)+f''(x)\frac{h^2}{2}+o\left(h^2\right).$$

Vergleichen mit (8.37) ergibt

$$f''(x)\frac{h^2}{2}+o\left(h^2\right)\geq 0,$$

woraus folgt

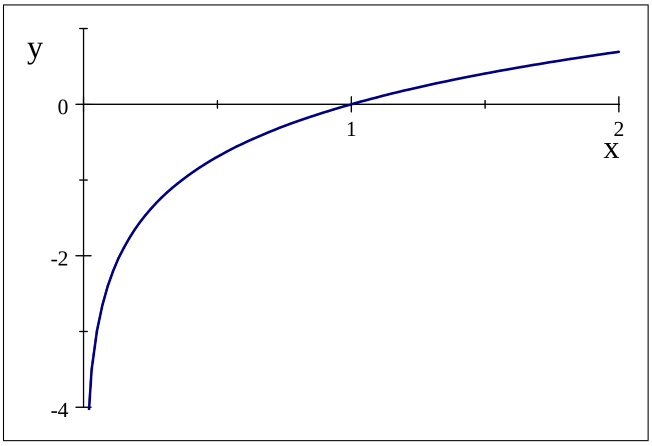
$$\frac{f''(x)}{2}=\lim_{h\to 0}\frac{f''(x)\frac{h^2}{2}+o\left(h^2\right)}{h^2}\geq 0.$$

(b) Diese Aussage folgt aus (a) da f konkav genau dann ist, wenn $-f$ konvex, und $f''\leq 0$ äquivalent zu $(-f)''\geq 0$. ■

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 8.17, dass $\ln x$ konkav auf $(0,+\infty)$ ist.



$\ln x$ ist konkav

Nach (8.35) gilt die folgende Ungleichung

$$\ln\left(\left(1-t\right)a+tb\right)\geq\left(1-t\right)\ln a+t\ln b \tag{8.38}$$

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

$$\frac{1}{p}+\frac{1}{q}=1. \tag{8.39}$$

Die Zahlen $p, q > 1$ mit (8.39) heißen *konjugierte Hölder-Exponenten.* Es folgt aus (8.38), dass

$$\ln\left(\frac{a}{p}+\frac{b}{q}\right)\geq\frac{1}{p}\ln a+\frac{1}{q}\ln b=\ln\left(a^{1/p}b^{1/q}\right). \tag{8.40}$$

Für $x = a^{1/p}$, $y = b^{1/q}$ erhalten wir aus (8.40) die *Young-Ungleichung*

$$\frac{x^p}{p}+\frac{y^q}{q}\geq xy,$$

die für alle konjugierte Hölder-Exponenten p, q und alle $x, y > 0$ gilt.

Beispiel. Sei $f(x) = x^p$ auf $J = (0, +\infty)$, wobei $p \in \mathbb{R}$, $p \neq 0, 1$. Da

$$f''(x)=p(p-1)x^{p-2},$$

so erhalten wir folgendes:

- für $0 < p < 1$ gilt $f''(x) < 0$ für alle $x > 0$;
- für $p < 0$ oder $p > 1$ gilt $f''(x) > 0$ für alle $x > 0$.

Somit ist die Funktion $f(x) = x^p$ konkav wenn $0 < p < 1$ und konvex wenn $p < 0$ oder $p > 1$.

Es folgt, dass für $p > 1$ und alle $a, b > 0$

$$\left(\frac{a+b}{2}\right)^p\leq\frac{a^p+b^p}{2},$$

d.h.

$$\frac{a+b}{2}\leq\left(\frac{a^p+b^p}{2}\right)^{1/p}.$$

Diese ist die Ungleichung vom arithmetischen Mittel und *Hölder-Mittel zur Stufe* p.

8.14 Untersuchung von Funktion mit Hilfe von f' und f''

Sei f eine stetige Funktion auf einem offenen Intervall $J = (a, b)$ mit den Grenzen $a, b \in \overline{\mathbb{R}}$, $a < b$. Nehmen wir an, dass f auf (a, b) 2-fach differenzierbar ist und dass f'' auf (a, b) stetig ist (obwohl die letzte Voraussetzung unwesentlich ist).

Schritt 1. Man bestimmt die Ableitung $f'(x)$ und die kritische Menge der Funktion $f(x)$:

$$K_f = \{x \in (a,b) : f'(x) = 0\} \cup \{a,b\}.$$

Nehmen wir an, dass die Menge K_f endlich ist. Seien $x_0, x_1, ..., x_n$ alle Punkte von K_f in steigender Reihenfolge, d.h. $x_k < x_{k+1}$ für alle $k = 0, ..., n-1$, insbesondere $x_0 = a$ und $x_n = b$.

Schritt 2. Man bestimmt das Vorzeichen von f' in jedem Intervall (x_k, x_{k+1}). Da $f'(x)$ in (x_k, x_{k+1}) nicht verschwindet, so gibt es zwei Möglichkeiten:

- entweder $f' > 0$ auf (x_k, x_{k+1}) so dass f streng monoton steigend auf (x_k, x_{k+1}) ist;
- oder $f' < 0$ auf (x_k, x_{k+1}) so dass f streng monoton fallend auf (x_k, x_{k+1}) ist.

Schritt 3. Man bestimmt alle Werte $f(x_k)$. Für $x_0 = a$ erweitern wir f auf a mit

$$f(a) := \lim_{x \to a} f(x) \in \overline{\mathbb{R}}$$

(der Grenzwert existiert, da f in (a, x_1) monoton ist). Analog erweitern wir f auf $x_n = b$.

Schritt 4. Man bestimmt die zweite Ableitung f'' und alle Werte $f''(x_k)$ für $x_k \in (a,b)$. Ist $f''(x_k) > 0$ so ist x_k eine lokale Minimumstelle, und im Fall $f''(x_k) < 0$ – eine lokale Maximumstelle. Alternative kann man die Monotonie von f in (x_{k-1}, x_k) und (x_k, x_{k+1}) benutzen um zu bestimmen ob x_k eine lokale Extremumstelle ist (was funktioniert auch im Fall $f''(x_k) = 0$).

Schritt 5. Man bestimmt die kritische Menge von f':

$$K_{f'} = \{x \in (a,b) : f''(x) = 0\} \cup \{a,b\}.$$

Angenommen, dass $K_{f'}$ endlich ist, bezeichnen wir mit $y_0, ..., y_m$ alle Punkte von $K_{f'}$ in steigender Reihenfolge; insbesondere $y_0 = a$ und $y_m = b$.

Schritt 6. Man bestimmt das Vorzeichen von f'' in jedem Intervall (y_j, y_{j+1}). Da f'' in jedem Intervall (y_j, y_{j+1}) nicht verschwindet, so gibt es zwei Möglichkeiten:

- entweder $f'' > 0$ auf (y_j, y_{j+1}) und somit f auf (y_j, y_{j+1}) konvex ist;
- oder $f'' < 0$ auf (y_j, y_{j+1}) und somit f auf (y_j, y_{j+1}) konkav ist.

Definition. Jeder Punkt y_j zwischen den Intervallen (y_{j-1}, y_j) und (y_j, y_{j+1}) wo f Konvexität nach Konkavität (oder umgekehrt) wechselt, heißt *Wendepunkt* von f.

Somit bestimmt man alle Wendepunkte und die Werte von f daran.

Schritt 7. Man skizziert den Graph von f auf jedem Intervall (x_k, x_{k+1}) wo f monoton zwischen den Werten $f(x_k)$ und $f(x_{k+1})$ ist. Auf jedem Intervall (y_j, y_{j+1}) soll der Graph konkav bzw konvex sein. Somit erhält man den Graph von f auf J.

Beispiel. Untersuchen wir die Funktion

$$f(x) = \frac{\ln x}{x}$$

auf $J = (0, +\infty)$.

Schritt 1. Wir haben

$$f'(x) = \frac{(\ln x)' x - (\ln x)(x)'}{x^2} = \frac{1 - \ln x}{x^2}.$$

Die Gleichung $f'(x) = 0$ ergibt $\ln x = 1$ und somit $x = e$. Deshalb haben wir

$$K_f = \{0, e, +\infty\}.$$

Schritt 2. Für $x \in (0, e)$ gilt $f'(x) > 0$ und für $x \in (e, +\infty)$ gilt $f'(x) < 0$. Somit ist die Funktion $f(x)$ streng monoton steigend in $(0, e)$ und streng monoton fallend in $(e, +\infty)$.

Schritt 3. Weiter haben wir $f(e) = \frac{1}{e} \approx 0,37$,

$$f(+\infty) := \lim_{x \to +\infty} \frac{\ln x}{x} = 0$$

(Aufgabe 117) und

$$f(0) := \lim_{x \to 0} \frac{\ln x}{x} = \lim_{x \to 0} \ln x \cdot \lim_{x \to 0} \frac{1}{x} = (-\infty) \cdot (+\infty) = -\infty.$$

Schritt 4. Wir haben

$$f'' = \left(\frac{1 - \ln x}{x^2}\right)' = \frac{-\frac{1}{x} \cdot x^2 - 2x(1 - \ln x)}{x^4} = \frac{2 \ln x - 3}{x^3}.$$

Insbesondere $f''(e) = \frac{2-3}{e^2} < 0$ so dass e eine lokale Maximumstelle ist. In der Tat ist e sogar die Maximumstelle von f auf $(0, +\infty)$, da $f(x)$ streng monoton steigend in $(0, e)$ und streng monoton fallend in $(e, +\infty)$ ist.

Schritt 5. Bestimmen wir die kritische Menge von f'. Die Gleichung $f'' = 0$ ergibt

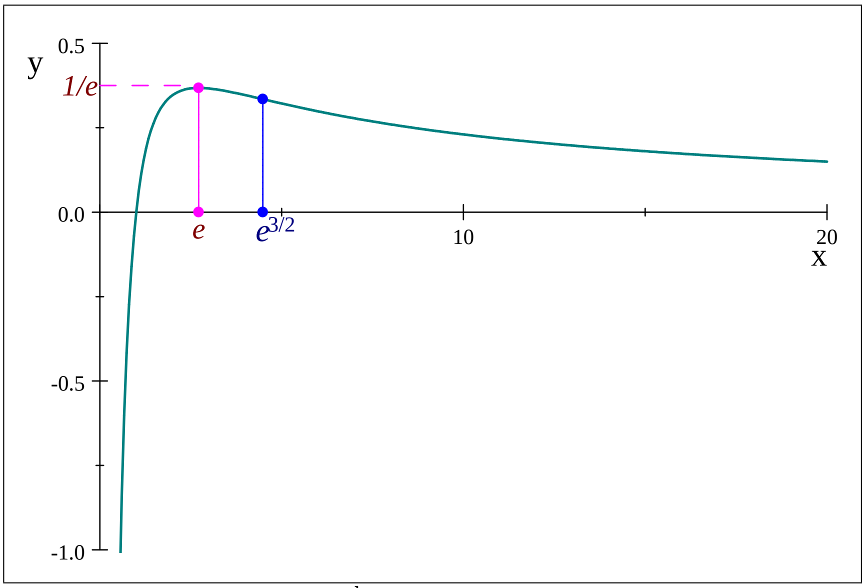
$$2 \ln x - 3 = 0$$

d.h. $x = e^{3/2} \approx 4,48$. Somit

$$K_{f'} = \left(0, e^{3/2}, +\infty\right).$$

Schritt 6. Im Intervall $\left(0, e^{3/2}\right)$ gilt $2 \ln x < 3$ und $f''(x) < 0$; somit ist die Funktion f auf $\left(0, e^{3/2}\right)$ konkav. Im $\left(e^{3/2}, +\infty\right)$ gilt $f''(x) > 0$ und somit ist f konvex. Der Punkt $e^{3/2}$ ist ein Wendepunkt von f, und $f\left(e^{3/2}\right) = \frac{3}{2e^{3/2}} \approx 0,33$.

Schritt 7. Der Graph der Funktion $f(x) = \frac{\ln}{x}$ sieht wie folgt aus:



Der Graph der Funktion $f(x) = \frac{\ln x}{x}$ mit der Maximumstelle und Wendepunkt

Beispiel. Untersuchen wir die Funktion

$$f(x) = 2x^3 - 3x^2 - 36x + 10$$

auf $(-\infty, +\infty)$.

Schritt 1. Wir haben

$$f'(x) = 6x^2 - 6x - 36.$$

Die Gleichung $f'(x) = 0$ ergibt zwei Nullstellen $x_1 = -2$ und $x_2 = 3$, so dass

$$K_f = \{-\infty, -2, 3, +\infty\}.$$

Schritt 2. Da

$$f'(x) = 6(x+2)(x-3),$$

so ist $f'(x)$ positive auf $(-\infty, -2)$, negative auf $(-2, 3)$ und wieder positive auf $(3, +\infty)$. Somit ist f auf $(-\infty, -2)$ und $(3, +\infty)$ streng monoton steigend, und auf $(-2, 3)$ streng monoton fallend.

Schritt 3. Wir haben

$$f(-2) = 54, \quad f(3) = -71, \quad f(+\infty) = +\infty \quad \text{und} \quad f(-\infty) = -\infty.$$

Schritt 4. Wir haben

$$f'' = 12x - 6$$

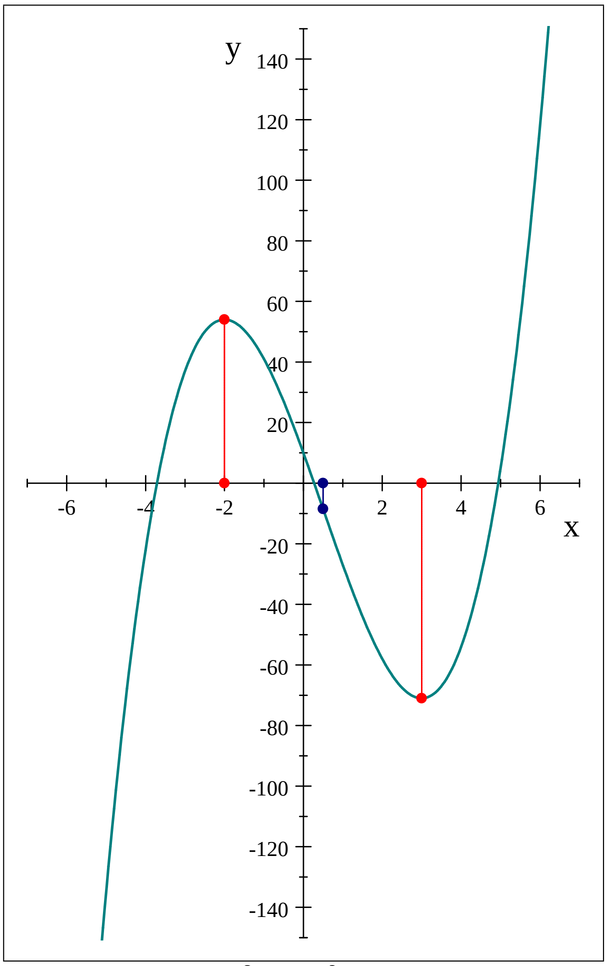
und $f''(-2) < 0$ und $f''(3) > 0$. Somit ist -2 eine lokale Maximumstelle und 3 ist eine lokale Minimumstelle.

Schritt 5. Die Gleichung $f''(x) = 0$ ergibt $x = \frac{1}{2}$. Somit

$$K_{f'} = \left\{-\infty, \tfrac{1}{2}, +\infty\right\}.$$

Schritt 6. Auf $\left(-\infty, \frac{1}{2}\right)$ gilt $f'' < 0$ so dass f konkav ist. Auf $\left(\frac{1}{2}, +\infty\right)$ gilt $f'' > 0$ so dass f konvex ist. Der Punkt $x = \frac{1}{2}$ ist somit ein Wendepunkt, und $f\left(\frac{1}{2}\right) = -8,5$.

Schritt 7. Somit erhalten wir den folgenden Graph.



Der Graph der Funktion $f(x) = 2x^3 - 3x^2 - 36x + 10$, die lokalen Extrema und Wendepunkt

8.15 * Vergleichstest und Ungleichungen

Satz 8.11 kann benutzt werden um Ungleichungen zu beweisen.

Satz 8.18 (Vergleichstest)

(*a*) *Seien f und g zwei stetige Funktionen auf einem Intervall $[c, b)$, $c < b$, die auf (c, b) differenzierbar sind. Angenommen seien die Bedingungen:*

(*i*) $f(c) \leq g(c)$

(*ii*) $f'(x) \leq g'(x)$ *für alle* $x \in (c, b)$.

Dann gilt $f(x) \leq g(x)$ für alle $x \in (c, b)$. Gilt $f'(x) < g'(x)$ für alle $x \in (c, b)$, so gilt auch $f(x) < g(x)$ für alle $x \in (c, b)$.

(*b*) *Seien f und g zwei stetige Funktionen auf einem Intervall $(a, c]$, $a < c$, die auf (a, c) differenzierbar sind. Angenommen seien die Bedingungen:*

(*i*) $f(c) \leq g(c)$

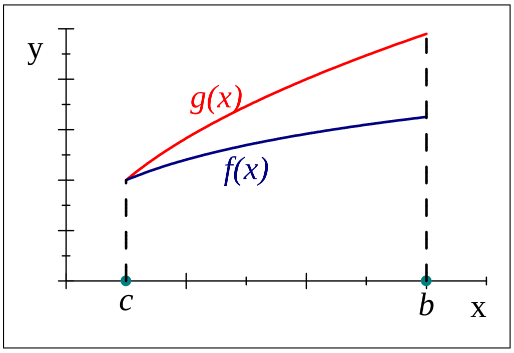
(ii) $f'(x) \geq g'(x)$ *für alle* $x \in (a,c)$.

Dann gilt $f(x) \leq g(x)$ *für alle* $x \in (a,c)$. *Gilt* $f'(x) > g'(x)$ *für alle* $x \in (a,c)$, *so gilt auch* $f(x) < g(x)$ *für alle* $x \in (a,c)$.

Beweis. (a) Betrachten wir eine neue Funktion $h = g - f$ und beachten, dass $h(c) \geq 0$ und $h' \geq 0$ auf (c,b). Nach dem Satz 8.11 ist h monoton steigend auf $[c,b)$. Daraus folgt, dass für alle $x \in (c,b)$,

$$h(x) \geq h(c) \geq 0,$$

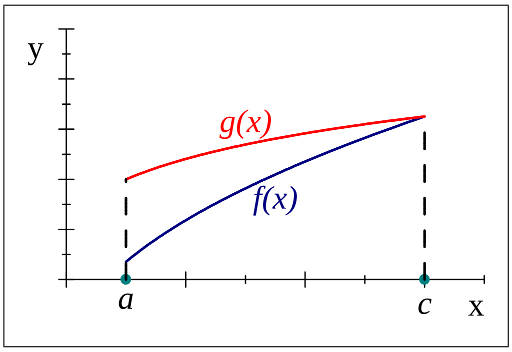
und somit $f(x) \leq g(x)$. Im Fall von der echten Ungleichung $f'(x) < g'(x)$ erhalten wir, dass $h' > 0$ und somit h streng monoton steigend ist, woraus $h(x) > 0$ und $f(x) < g(x)$ folgen.



(b) In diesem Fall ist die Funktion h monoton fallend und $h(c) \geq 0$, daher

$$h(x) \geq h(c) \geq 0$$

für alle $x \in (a,c)$, 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 $f(x) < g(x)$.

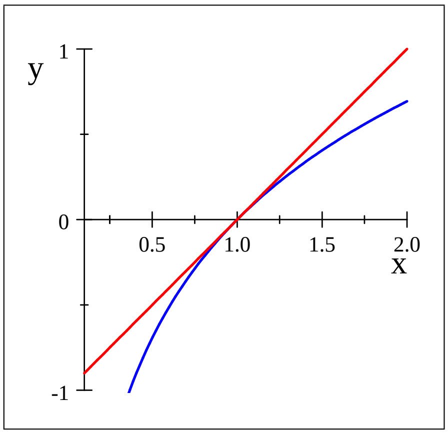


■

Beispiel. Beweisen wir die Ungleichung

$$\ln x \leq x - 1 \tag{8.41}$$

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 (8.41) gleich 0. Für $x > 1$ haben wir

$$(\ln x)' = \frac{1}{x} < 1 = (x - 1)'.$$

Anwendung des Vergleichstests von Satz 8.18(a) auf Interval $[1, +\infty)$ ergibt, dass $\ln x < x - 1$ für alle $x > 1$.

Für $0 < x < 1$ haben wir

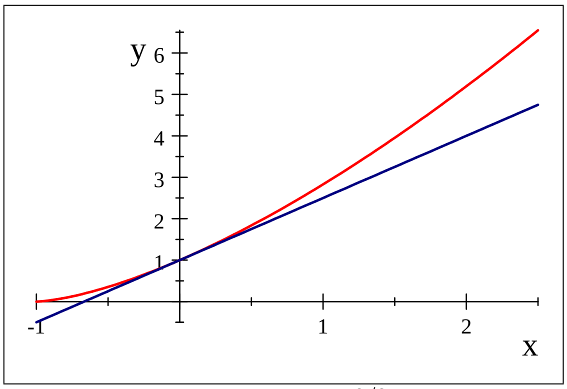
$$(\ln x)' = \frac{1}{x} > 1 = (x - 1)'.$$

Anwendung des Vergleichstests von Korollar 8.18(b) auf Interval $(0, 1]$ ergibt $\ln x < x - 1$ für alle $x \in (0, 1)$. Somit gilt (8.41) für alle $x > 0$.

Beispiel. Beweisen wir, dass für alle $a > 1$ und $x > -1$ gilt

$$(1 + x)^a \geq 1 + ax. \tag{8.42}$$

Das ist eine Verallgemeinerung der Bernoulli-Ungleichung, die wir für $a \in \mathbb{N}$ wissen.



Die Graphen von Funktionen $(1 + x)^{3/2}$ (rot) und $1 + \frac{3}{2}x$ (blue)

Für $x = 0$ sind die beiden Seiten von (8.42) gleich 1. Für $x > 0$ gilt

$$((1 + x)^a)' = a\,(1 + x)^{a-1} > 1 = (1 + ax)'$$

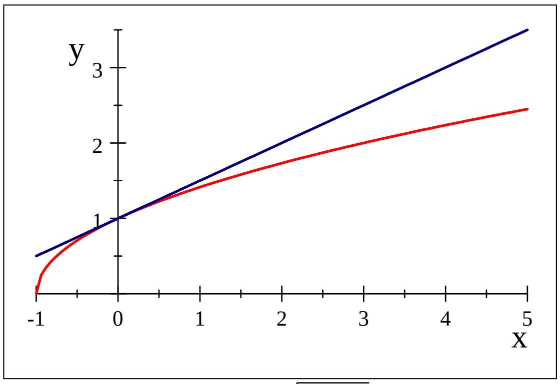
und für $-1 < x < 0$ gilt

$$((1 + x)^a)' = a\,(1 + x)^{a-1} < 1 = (1 + ax)'.$$

Nach dem Satz 5.6 beschließen wir, dass (8.42) für alle $x > -1$ gilt.

Analog beweist man, dass für $0 < a < 1$ die umgekehrte Ungleichung gilt

$$(1+x)^a \leq 1 + ax.$$



Die Graphen von Funktionen $\sqrt{1+x}$ (rot) und $1 + \frac{1}{2}x$ (blue)