Mathematische Grundlagen

Hans-Peter Beise

Wintersemester 2024/2025

Inhaltsverzeichnis

Einleitung

In der gesamten Mathematik wird mit dem grundlegenden Konzept von Abbildungen zwischen Mengen gearbeitet, also einer Vorschrift, die den Elementen einer ersten Menge jeweils ein Element einer zweiten Menge zuordnet. Dieses Vorgehen ist bereits aus der Schulmathematik bekannt, wo man meist Abbildungen (auch Zuordnung oder Funktion genannt) von reellen Zahlen auf reelle Zahlen betrachtet.
Eine solche Abbildung ist beispielsweise durch die Vorschrift $f(x) = x^2$ beschrieben, wobei x, das Argument, eine reelle Zahl, und $f(x)$ der Bildpunkt, wiederum eine reelle Zahl ist. Der sogenannte Graph dieser Abbildung zeigt hier die bekannte Parabel. Durch solche grundlegenden Funktionen, die eine Zahl auf eine Zahl abbilden, lassen sich bereits einige Zusammenhänge der realen Wirklichkeit modellieren:

1. Zusammenhang von Stromverbrauch und Stromrechnung (linear affin)
2. Zusammenhang von Seitenlänge eines Quadrats und Fläche des Quadrats (quadratische Funktion)
3. Zusammenhang von Geschwindigkeit und zurückgelegtem Weg
4. ... .

Das Konzept der Abbildung ist jedoch sehr viel weitreichender anwendbar und taucht quasi überall in der Informatik und Mathematik auf, insbesondere auch dort, reale Zusammenhänge (technische, physikalische, biologische, wirtschaftliche, ...) beschrieben werden:

1. Algorithmen, etwa zur Bestimmung kürzester Wege.
2. Erzeugen von Computergrafiken: z.B. Bewegen von Objekten in einer virtuellen 3D-Welt, Projektion auf einen 2D-Bildschirm, ...
3. Künstliche Intelligenz: z.B. Erkennen von Wörtern im Sprachsignal mittels neuronaler Netze (Deep Learning), ...
4. Medizinische Bildgebung: z.B. Der Zusammenhang von Körperbeschaffenheit und gemessenem Signal (Röntgen) wird durch mathematische Abbildungen modelliert.

5. Komplexere physikalische Vorgänge: z.B. Verformung von Objekten durch Einwirkung von Kraft, Ausbreitung elektromagnetischer Wellen im Raum, ...

6. Statistik: z.B. Gewinnvorhersagen anhand historischer Unternehmensdaten, ...

7. Kryptographie: z.B. Zur Verschlüsselung werden nicht- (oder fast nicht-) umkehrbare Abbildungen angewendet (Public-Key-Kryptographie).

Als Grundlage für das weitere Arbeiten mit mathematischen Werkzeugen und das Vertiefen hin zu komplexeren mathematischen Anwendungen bedarf es somit der Einführung grundlegender Operationen und Strukturen auf Mengen und Abbildungen.

1 Mengen und Abbildungen

Wir starten mit der grundlegenden Definition einer Menge.

Definition 1.1 Eine **Menge** M ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.
Ein solches Objekt x heißt **Element** der Menge M, Schreibweise:

$$x \in M.$$

Ist x nicht Element von M, so schreiben wir

$$x \notin M.$$

Die Menge ohne Elemente heißt die **leere Menge**, Schreibweise:

$$\emptyset \text{ oder } \{\}.$$

Es gibt verschiedene Möglichkeiten der Darstellung von Mengen, etwa die aufzählende Schreibweise

$$M := \{x_1, x_2, x_3, \ldots, x_n\}$$

wobei $x_1, x_2, x_3, \ldots, x_n$ die Elemente von M sind. Oder die beschreibende Schreibweise, also eine Charakterisierung der Elemente. Die beschreibende Schreibweise hat allgemein die Form

$$M := \{x : x \text{ hat die Eigenschaft } E\},$$

wobei E irgendeine „Eigenschaft“ ist. Alternativ schreibt man statt x : auch $x|$.

Beispiel 1.2

$$M = \{1, 2, 3, 4\}$$

$$M = \{x \in \mathbb{N} : 1 \leq x \leq 4\}.$$

Wir setzen die Kenntnis der folgenden üblichen Zahlenmengen voraus.

$$\begin{aligned}
\mathbb{N} &:= \{x : x \text{ natürliche Zahl}\} \\
\mathbb{N}_0 &:= \{x : x \text{ natürliche Zahl oder } x = 0\} \\
\mathbb{Z} &:= \{x : x \text{ ganze Zahl}\} \\
\mathbb{Q} &:= \{x : x \text{ rationale Zahl}\} \\
\mathbb{R} &:= \{x : x \text{ reelle Zahl}\}.
\end{aligned} \tag{1.1}$$

Definition 1.3 Es seien A, B Mengen.

1. A heißt **Teilmenge** von B, falls aus $x \in A$ auch $x \in B$ folgt. Schreibweie:

$$A \subset B \text{ oder auch } A \subseteq B$$

2. A und B heißen **gleich**, falls $A \subset B$ und $B \subset A$ gilt. Schreibweise

$$A = B$$

3. Die Menge

$$B \setminus A := \{x : x \in B \text{ und } x \notin A\}$$

heißt **Differenz** von B und A. Ist $A \subset B$, so heißt

$$A^c := A^{c(B)} := B \setminus A$$

Komplement von A bezüglich B. Die Bezeichnung $A^{c(B)}$ wird jedoch nur verwendet, wenn Uneindeutigkeit bezüglich der Grundmenge B besteht.

4. Die Menge

$$B \cap A := \{x : x \in A \text{ und } x \in B\}$$

heißt **Schnitt** oder **Schnittmenge** von B und A.

5. Die Menge

$$B \cup A := \{x : x \in A \text{ oder } x \in B\}$$

heißt **Vereinigung** von B und A.

Bemerkung 1.4 Komplemente von Mengen treten auch in folgender Form auf:

Es sei A eine Menge, dann gilt für alle $x \in A$, dass $x \notin A^c$.

Hierbei ist nicht direkt klar bezüglich welcher Menge das Komplement gebildet werden soll. Im Allgemeinen wird hier dann von einer Menge M ausgegangen, die die Eigenschaft $A \subset M$ erfüllt, sonst aber nicht weiter beschrieben ist.
Dieses Vorgehen setzt sich fort. So ist beispielsweise mit $A^c \cap B^c \cap C^c$ gemeint $A^{c(M)} \cap B^{c(M)} \cap C^{c(M)}$ für eine Menge M mit den Eigenschaften $A \subset M$, $B \subset M$ und $C \subset M$.

Bemerkung 1.5 Für Mengen $M_1, M_2, \ldots, M_n$ verwendet man die folgende Notation

$$\bigcap_{j=1}^{n} M_j = M_1 \cap M_2 \cap \ldots \cap M_n$$

und

$$\bigcup_{j=1}^{n} M_j = M_1 \cup M_2 \cup \ldots \cup M_n$$

für Schnitt und Vereinigung. Diese Notation kann auch auf unendlich viele Mengen erweitert werden.

Beispiel 1.6

$$\{1,2\} \cup \{2,5\} = \{1,2,5\}$$
$$\{1,2\} \cap \{2,5\} = \{2\}$$
$$\mathbb{Z} \cap \mathbb{N} = \mathbb{N}$$

Das folgende Ergebnis fasst grundlegende Rechenregeln für die oben eingeführten Mengenoperationen zusammen.

Satz 1.7 *Es seien A, B, C Mengen, dann gilt:*

1. *$A \cap B = B \cap A$ und $A \cup B = B \cup A$ (Kommutativgesetzte)*
2. *$A \cap (B \cap C) = (A \cap B) \cap C$ und $A \cup (B \cup C) = (A \cup B) \cup C$ (Assoziativgesetze)*
3. *$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributivgesetze)*

Beweis. Wir zeigen hier nur exemplarisch $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
Dabei meint das Symbol $\Leftrightarrow$ die Äquivalenz zweier Aussagen (Eine formalere Einführung folgt später). Die Klammerung bedeute wie üblich, dass der Ausdruck in der Klammer zuerst ausgewertet wird.
Es gilt:

$$\begin{aligned}
& x \in A \cap (B \cup C) \\
\Leftrightarrow \quad & x \in A \text{ und } x \in (B \cup C) \\
\Leftrightarrow \quad & x \in A \text{ und } (x \in B \text{ oder } x \in C) \\
\Leftrightarrow \quad & (x \in A \text{ und } x \in B) \text{ oder } (x \in A \text{ und } x \in C) \\
\Leftrightarrow \quad & (x \in A \cap B) \text{ oder } (x \in A \cap C) \\
\Leftrightarrow \quad & x \in (A \cap B) \cup (A \cap C).
\end{aligned}$$

Damit ist ein beliebiges Element x in der Menge $A \cap (B \cup C)$ enthalten genau dann, wenn es in der Menge $(A \cap B) \cup (A \cap C)$ enthalten ist, wodurch die Gleichheit dieser Mengen bewiesen ist. □

Satz 1.8 *(De Morgansche Gesetze) Es seien A, B Mengen, dann gilt*

1. $(A \cap B)^c = A^c \cup B^c$ *und*
2. $(A \cup B)^c = A^c \cap B^c$.

Für die üblichen Zahlenmengen aus (1.1) sind höherdimensionale Konstrukte der Form (x_1, x_2), (x_1, x_2, x_3), beispielsweise mit $x_1, x_2, x_3 \in \mathbb{R}$, möglicherweise schon bekannt. In der nächsten Definition führen wir diese sogenannten geordneten Tupel für allgemeine Mengen ein.

Definition 1.9 Es seien M_1, M_2 zwei nichtleere Mengen, dann heißt

$$M_1 \times M_2 := \{(x_1, x_2) : x_1 \in M_1, x_2 \in M_2\}$$

(Menge der geordneten 2-Tupel) das **kartesische Produkt** von M_1 und M_2. Seien allgemeiner mehrere nichtleere Mengen M_1, M_2, $M_3, \ldots, M_n$ gegeben, dann heißt

$$M_1 \times M_2 \times \ldots \times M_n := \{(x_1, x_2, \ldots, x_n) : x_j \in M_j, j = 1, \ldots, n\}$$

das **kartesische Produkt** der Mengen M_1, M_2, $M_3, \ldots, M_n$.
Gilt weiter $M := M_1 = M_2 = \ldots = M_n$, so schreibt man auch

$$M^n = \underbrace{M \times M \times \ldots \times M}_{n-\text{mal}}.$$

Beispiel 1.10

1. Für $\{1, 3\} \times \{3, 4\} = \{(1, 3), (1, 4), (3, 3), (3, 4)\}$.
2. Für eine natürliche Zahl $n \in \mathbb{N}$ ist
$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_j \in \mathbb{R},\ j = 1, \ldots, n\}.$$
3. Für $M = \mathbb{Z} \backslash \mathbb{N}_0$ ist $M^2 = M \times M$ die Menge der geordneten zwei-Tupel negativer ganzer Zahlen.
4. Für $\{1, 3\} \times \{3, 4\} \times \{(\pi, \frac{1}{3})\} = \{(1, 3, (\pi, \frac{1}{3})), (1, 4, (\pi, \frac{1}{3})), (3, 3, (\pi, \frac{1}{3})), (3, 4, (\pi, \frac{1}{3}))\}$. .

Ähnlich wie bei der obigen Einführung von Mengen wollen wir auf eine eher intuitive Definition des zweiten grundlegenden Begriffes der Mathematik zurückgreifen, nämlich den einer Abbildung (oder Funktion).
Dieser Begriff wird später im Zusammenhang mit Relationen nocheinmal formaler definiert.
(Im Zusammenhang mit Abbildungen sind hier die Variablennamen X, Y, Z für Mengen bevorzugt)

Definition 1.11 Es seien X und Y nichtleere Mengen. Eine **Abbildung** (oder **Funktion**) f von X nach Y ist eine „Vorschrift“, die jedem $x \in X$ *genau ein* Element $f(x) \in Y$ zuordnet. Dabei heißen

X der **Definitionsbereich**

und

Y der **Zielbereich**

von f.

Man schreibt

$$f : X \to Y$$

und formuliert die Vorschrift f häufig durch mathematische Formeln oder Algorithmen, an manchen Stellen aber auch durch eine wörtliche Beschreibung. Alternativ schreibt man zur Formulierung der Vorschrift auch

$$x \mapsto f(x).$$

Bemerkung 1.12 An anderer Stelle (z.B. theoretische Informatik) betrachtet man Abbildungen

$$f : X \to Y,$$

die unter Umständen nicht auf der gesamten Menge X ausgewertet werden können. In diesem Fall ist der Definitionsbereich eine Teilmenge von X, nämlich der Bereich wo die Funktion ausgewertet werden kann. Dies kann zum Beispiel dann sinnvoll sein, wenn f einen Algorithmus beschreibt, der für manche Eingabedaten nicht zum Ende kommt und daher hierfür keine definierte Ausgabe erzeugt. Eine solche Eingabe wäre dann ein Element außerhalb des Definitionsbereichs.

Definition 1.13 Für eine Abbildung $f : X \to Y$ heißt die Menge

$$\{(x, f(x)) : x \in X\} \ (\subset X \times Y)$$

der **Graph** der Abbildung f.

Beispiel 1.14 Es sei $f : \{0, 1, 2, 3\} \to \mathbb{R}$ mit $f(x) = x + 2$, dann ist der Graph von f die folgende Menge:

$$\{(0, 2), (1, 3), (2, 4), (3, 5)\}.$$

Definition 1.15 Es $X \subset \mathbb{R}$ und $f : X \to \mathbb{R}$ eine Abbildung. Die Abbildung heißt dann

1. **monoton fallend**, falls $\forall x, y \in X,\ x < y :\ f(x) \geq f(y)$,
2. **streng monoton fallend**, falls $\forall x, y \in X,\ x < y :\ f(x) > f(y)$,
3. **monoton wachsend**, falls $\forall x, y \in X,\ x < y :\ f(x) \leq f(y)$,
4. **streng monoton wachsend**, falls $\forall x, y \in X,\ x < y :\ f(x) < f(y)$.

Allgemeiner heißt f **(streng) monoton**, wenn f (streng) monoton fallend oder (streng) monoton wachsend ist.

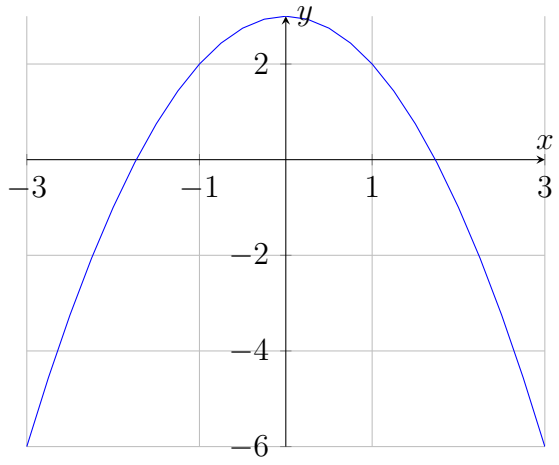

Abbildung 1: Graph der Abbildung $f(x) = -x^2 + 3$

Definition 1.16 Sind $f, g : X \to Y$ zwei Abbildungen, so heißen f und g **gleich**, falls $f(x) = g(x)$ für alle $x \in X$ gilt.

Im Weiteren verwenden wir oft die Schreibweise $(x \in X)$ anstelle von „für alle $x \in X$“.

Beispiel 1.17 Es seien $\mathcal{L} := \{L_1, \ldots, L_n\}$ eine Menge von Buslinien und $\mathcal{H} := \{H_1, \ldots, H_m\}$ eine Menge von Haltestellen, wobei $n, m \in \mathbb{N}$. Die Abbildung $f : \mathcal{L} \to \mathcal{H}^{10}$ ordne jeder Buslinie ein Tupel von 10 Haltestellen zu, welches der Route der Linie entspricht. Somit fährt in diesem Beispiel jede Linie genau 10 Haltestellen an.
Formal ausgedrückt:

$$f(L) = (H_{i_1}, \ldots, H_{i_{10}})$$

wobei L eine Buslinie ist und $i_1, \ldots, i_{10}$ Indizes von Haltestellen sind, die zu dieser Buslinie gehören. Dies könnte konkret wie folgt aussehen:

$$\begin{aligned} f(L_1) &= (H_3, H_{15}, H_{22}, H_{35}, H_{41}, H_{12}, H_{48}, H_7, H_{19}, H_{50}) \\ f(L_2) &= (H_1, H_8, H_{14}, H_{27}, H_{33}, H_{40}, H_5, H_{18}, H_{26}, H_{49}) \\ &\vdots \end{aligned}$$

(Übung) Seien nun weiter $\mathcal{T} := \{T_1, \ldots, T_k\}$ feste Zeiten, an denen Busse and den Haltestellen halten. Wie sähe formal eine Abbildung aus, die jeder Linie die Haltestellen und die entsprechenden Zeiten, an denen die Linie die jeweilige Haltestelle anfährt, zuordnet?

Definition 1.18 Sind X, Y Mengen und ist $f : X \to Y$, so heißt für $B \subset Y$

$$f^{-1}(B) := \{x \in X : f(x) \in B\} \ (\subset X)$$

Urbildmenge von B unter f, und für $A \subset X$

$$f(A) := \{f(x) : x \in A\} = \{y \in Y : y = f(x) \text{ für ein } x \in A\} \ (\subset Y)$$

Bildmenge von A unter f. Speziell heißt

$$W(f) := f(X)$$

Wertebereich von f. Ist $W(f)$ einpunktig, so heißt f **konstant**.

Beispiel 1.19 Es seien $X := Y := \mathbb{N}$, und es sei $f : \mathbb{N} \to \mathbb{N}$ definiert durch

$$f(x) := \begin{cases} x \ , & \text{falls } x \text{ gerade} \\ 2x, & \text{falls } x \text{ ungerade} \end{cases} .$$

Dann gilt

$$f^{-1}\big(\{2,4,6\}\big) = f^{-1}\big(\{1,2,3,4,5,6\}\big) = \{1,2,3,4,6\}$$

und

$$f\big(\{1,2,3\}\big) = \{2,6\}.$$

Außerdem ist $W(f) = \{y \in \mathbb{N} : y \text{ gerade}\}$.

Satz 1.20 *Sind $f : X \to Y$ eine Abbildungen und $B_1, \ldots, B_n \subset Y$. Dann gelten folgende*

$$f^{-1}\left(\bigcap_{k=1}^{n} B_k\right) = \bigcap_{k=1}^{n} f^{-1}\left(B_k\right),$$

$$f^{-1}\left(\bigcup_{k=1}^{n} B_k\right) = \bigcup_{k=1}^{n} f^{-1}\left(B_k\right).$$

Zum Beweis der ersten Aussage beginne man mit

$$x \in f^{-1}\left(\bigcap_{k=1}^{n} B_k\right) \Leftrightarrow \ f(x) \in \bigcap_{k=1}^{n} B_k \ldots$$

und zeigt, dass dies schließlich äquivalent dazu ist, dass x in der rechten Menge enthalten ist (Übung). Die zweite Identität wird analog gezeigt.
Für die Bildmenge gilt diese sogenannte Schnitt- und Vereinigungs-Stabilität nicht!

Definition 1.21 Es seien X, Y Mengen. Eine Abbildung $f : X \to Y$ heißt

1. **surjektiv**, falls für alle $y \in Y$ ein $x \in X$ so existiert, dass $f(x) = y$ (d.h. $W(f) = Y$)
2. **injektiv**, falls gilt: sind $x_1, x_2 \in X$ mit $x_1 \neq x_2$, so ist $f(x_1) \neq f(x_2)$,
3. **bijektiv**, falls f injektiv und surjektiv ist.

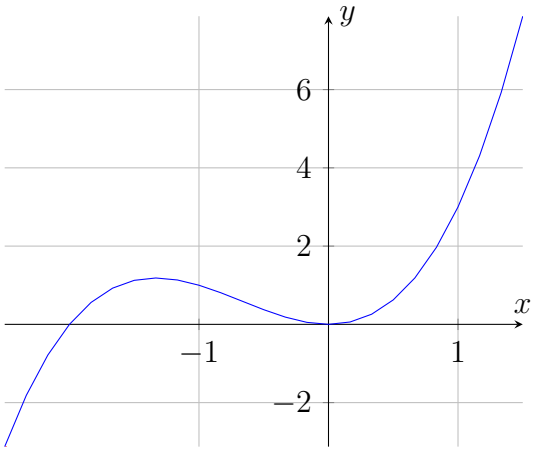

Abbildung 2: Graph der Abbildung $f(x) = x^3 + 2x^2$, f ist nicht injektiv. (Woran kann man das erkennen?)

Definition 1.22 Es seien X, Y, Z Mengen und $f : X \to Y$ sowie $g : Y \to Z$ Abbildungen. Dann heißt $g \circ f : X \to Z$, definiert durch

$$(g \circ f)(x) := g(f(x)) \qquad (x \in X)$$

Komposition (oder **Hintereinanderausführung** oder **Verkettung**) von g und f.

Beispiel 1.23 Sind $X = Y = Z = \mathbb{N}$ und $f, g : \mathbb{N} \to \mathbb{N}$ definiert durch

$$f(x) := x^2, \quad g(x) := x + 1 \quad (x \in \mathbb{N}) ,$$

so ist $g \circ f : \mathbb{N} \to \mathbb{N}$ gegeben durch

$$(g \circ f)(x) = x^2 + 1 \quad (x \in \mathbb{N}) .$$

Man beachte: Hier ist auch $f \circ g : \mathbb{N} \to \mathbb{N}$ definiert und es gilt

$$(f \circ g)(x) = (x + 1)^2 \quad (x \in \mathbb{N}).$$

Dabei ist $g \circ f \neq f \circ g$ (da etwa $(g \circ f)(1) = 2 \neq 4 = (f \circ g)(1)$).

Satz 1.24 *Es seien X, Y, Z Mengen und $f : X \to Y$ und $g : Y \to Z$ Abbildungen.*

1. *Sind f, g injektiv, dann ist auch die Komposition $g \circ f : X \to Z$ injektiv.*
2. *Sind f, g surjektiv, dann ist auch die Komposition $g \circ f : X \to Z$ surjektiv.*

Beweis. Zu 1.: Es seien $x_1, x_2 \in X$ und $x_1 \neq x_2$. Dann folgt $f(x_1) \neq f(x_2)$ da f injektiv ist. Daraus folgt weiter

$$g \circ f(x_1) = g(f(x_1)) \neq g(f(x_2)) = g \circ f(x_2)$$

da g injektiv ist.
Zu 2.: Es sei $z \in Z$. Da g surjektiv ist gibt es ein $y \in Y$ mit $g(y) = z$, und da f surjektiv ist gibt es ein $x \in X$ mit $f(x) = y$, sodass folgt

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

□

Definition 1.25 Für eine nichtleere Menge X heißt $\mathrm{id}_X : X \to X$, definiert durch $\mathrm{id}(x) := \mathrm{id}_X(x) := x$ $(x \in X)$, die **identische Abbildung** auf X.

Definition 1.26 Es seien X, Y Mengen und es sei $f : X \to Y$ eine Abbildung. Die Abbildung $f^{-1} : Y \to X$ heißt **Umkehrabbildung** von f oder **inverse Abbildung** (bezüglich $\circ$) von f , falls

$$f^{-1} \circ f = \mathrm{id}_X, \text{ und } f \circ f^{-1} = \mathrm{id}_Y$$

gilt.

Beispiel 1.27 Sei $f : [0, \infty) \to [1, \infty)$, $f(x) = x^2 + 1$. Dann gilt $f^{-1} : [1, \infty) \to [0, \infty)$, $f(x) = \sqrt{x - 1}$.

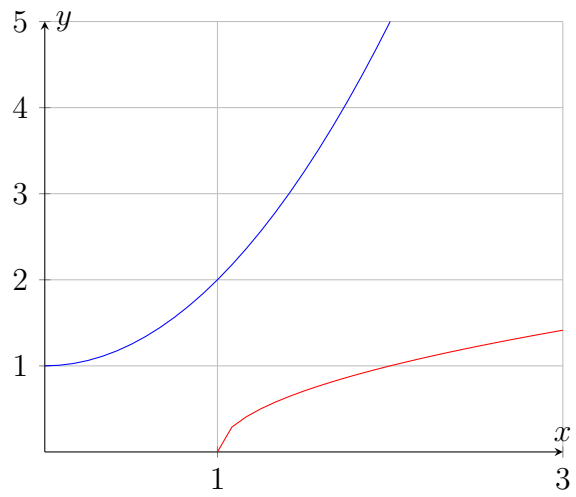

Abbildung 3: Graph der Abbildung $f(x) = x^2 + 1$ (blau), $f^{-1}(x) = \sqrt{x-1}$ (rot)

Satz 1.28 *Es seien X, Y Mengen und es sei $f : X \to Y$ bijektiv. Dann existiert zu jedem $y \in Y$ genau ein $x \in X$ mit $f(x) = y$. Damit ergibt sich die inverse Abbildung durch*

$$f^{-1}(y) := x \qquad (y \in Y) \,,$$

wobei $y = f(x)$.

Für endliche Mengen ist es offensichtlich, dass die Anzahl der Elemente Auskunft über die „Größe “der Menge gibt.

Für eine Menge $M = \{x_1, \ldots, x_n\}$ $(n \in \mathbb{N})$ setzen wir

$$|M| := n.$$

Betrachten wir jedoch zwei Mengen mit unendlich viele Elemente, so ist nicht klar wie man allgemein die „Größe“solcher Mengen vergleichen kann. In der Mathematik gibt es hierzu einige Zugänge, die jeweils in ihrer Hinsicht etwas über die „Größe“aussagen. Wir werden hier nur den elementaren Begriff der Mächtigkeit einführen.

Definition 1.29 Es seien M und N zwei Mengen, dann heißen diese **gleichmächtig**, falls eine bijektive Abbildung $f : M \to N$ (oder äquivalent $f : N \to M$) existiert.

1. Wenn M und $\{1, \ldots, n\}$ für ein $n \in \mathbb{N}$ gleichmächtig sind, dann heißt M **endliche Menge** und man sagt, dass M die **Kardinalität** n hat (gleichbedeutend $|M| = n$).
2. Wenn M und $\mathbb{N}$ gleichmächtig sind, dann heißt M **abzählbar unendlich** oder kurz **abzählbar**.
3. Wenn M endlich oder abzählbar unendlich ist, dann sagt man auch M ist **höchstens abzählbar**.
4. Wenn M nicht höchstens abzählbar ist, dann heißt M **überabzählbar**.

Bemerkung 1.30 Mit Hilfe von Satz 1.24 folgt leicht das Folgende. Sind M_1, M_2, M_3 Mengen so, dass M_1, M_2 gleichmächtig sind und M_2, M_3 gleichmächtig sind, dann sind auch M_1, M_3 gleichmächtig.

Beispiel 1.31 Die Menge der ganzen Zahlen ist abzählbar. Wir definieren dazu die Abbildung $f : \mathbb{N} \to \mathbb{Z}$ wie folgt:

$$f(n) = \begin{cases} \frac{n}{2}, & \text{falls } n \text{ gerade} \\ -\frac{n+1}{2}, & \text{falls } n \text{ ungerade.} \end{cases}$$

Es gilt dann

$$\begin{array}{cc} n & f(n) \\ 1 & -1 \\ 2 & 1 \\ 3 & -2 \\ 4 & 2 \\ 5 & -3 \\ 6 & 3 \\ \vdots & \end{array}$$

Diese Abbildung ordnet jeder natürlichen Zahl n genau eine ganze Zahl zu, und jede ganze Zahl wird von dieser Abbildung getroffen. Mit anderen Worten f ist bijektiv.

Beispiel 1.32 Es sei M die Menge aller $0,1$-Folgen, formaler $M := \{a : a : \mathbb{N} \to \{0,1\}\}$. Dann ist M überabzählbar. Wir betrachten dazu eine beliebige Abbildung $f : \mathbb{N} \to M$, und konstruieren eine neue $0,1$-Folge b, die sich von allen Folgen $f(n)$ unterscheidet. Wir setzen dazu:

$$b(n) := \begin{cases} 1, & \text{falls } (f(n))(n) = 0 \\ 0, & \text{falls } (f(n))(n) = 1. \end{cases}$$

Also gilt $(f(n))(n) \neq b(n)$ und daher $b \neq f(n)$. Dies gilt für alle $n \in \mathbb{N}$ und somit folgt $b \notin f(\mathbb{N})$. Damit ist f nicht surjektiv. Da f beliebig war folgt, dass es keine surjektive Abbildung $f : \mathbb{N} \to M$ geben kann.

Satz 1.33 *Es seien $M_1, \ldots, M_n$ endliche Mengen, dann gilt*

$$|M_1 \times M_2 \times \ldots \times M_n| = \prod_{j=1}^{n} |M_j|.$$

Definition 1.34 Es sei M eine Menge, dann heißt

$$\mathcal{P}(M) := 2^M := \{N : N \subset M\}$$

die **Potenzmenge** von M.

Betrachten wir $M = \{1, \ldots, n\}$ und $N \subset M$, dann definiert

$$\varphi_N : \{1, \ldots, n\} \to \{0, 1\}$$

$$\varphi_N(x) := \begin{cases} 1\ , & \text{falls } x \in N \\ 0, & \text{falls } x \notin N \end{cases}$$

eine eindeutige Zuordnung $N \mapsto (\varphi_N(1), \varphi_N(2), \ldots, \varphi_N(n)) \in \{0, 1\}^n$. Diese Zuordnung definiert eine bijektive Abbildung

$$f : \{N : N \subset M\} \to \{0, 1\}^n,\ N \mapsto (\varphi_N(1), \ldots, \varphi_N(n)),$$

und somit gilt

$$|\{N : N \subset M\}| = |\{0, 1\}^n| = 2^n.$$

Dieser Sachverhalt für endliche Mengen ist der Grund für die Namensgebung in der nächsten Definition.

Satz 1.35 *Es sei M ein Menge mit $|M| = n$ mit $n \in \mathbb{N}_0$, dann gilt*

$$|\mathcal{P}(M)| = 2^n.$$

Zum Abschluss liefern wir noch eine wichtige Definition für Teilmengen der reellen Zahlen.

Definition 1.36 Es sei $M \subset \mathbb{R}$.

Gibt es ein $y \in M$ mit $y \geq x$ für alle $x \in M$, dann heißt y das **Maximum** von M, Schreibweise:

$$y = \max M$$

Gibt es ein $y \in M$ mit $y \leq x$ für alle $x \in M$, dann heißt y das **Minimum** von M, Schreibweise:

$$y = \min M$$

Gibt es ein **kleinstes** $y \in \mathbb{R}$ mit $y \geq x$ für alle $x \in M$, dann heißt y das **Supremum** von M, Schreibweise:

$$y = \sup M$$

Gibt es ein **größtes** $y \in \mathbb{R}$ mit $y \leq x$ für alle $x \in M$, dann heißt y das **Infimum** von M, Schreibweise:

$$y = \inf M$$

Im Weiteren verwenden wir oft die Schreibweise $(x \in X)$ anstelle von „für alle $x \in X$“.

Beispiel 1.37

1. $4 = \max[1, 4],\ 1 = \min[1, 4]$
2. $5 = \sup[1, 5)$, aber die Menge $[1, 5)$ hat kein Maximum.
3. $-1 = \inf(-1, 5)$, aber die Menge $(-1, 5)$ hat kein Minimum.
4. $\sqrt{2} = \inf([\sqrt{2}, 9] \cap \mathbb{Q})$, aber die Menge $[\sqrt{2}, 9] \cap \mathbb{Q}$ hat kein Minimum.

Vollständigkeitsaxiom der reellen Zahlen:

Ist $M \subset \mathbb{R}$ nichtleer und nach oben beschränkt, dann existiert $\sup M \in \mathbb{R}$.

Übung: Warum gilt das Vollständigkeitsaxiom in $\mathbb{Q}$ nicht?

2 Aussagenlogik

In diesem kurzen Kapitel werden wir (u. a. mit Hilfe der Begriffe aus dem ersten Kapitel) die Grundlagen der sogenannten Aussagenlogik einführen.

Einige Beispiele von Aussagen:

a 7 ist kleiner als 10

b 7 ist kleiner als 4

c Mein Auto ist kaputt.

d Ich fahre mit dem Bus.

e Heute räume ich auf.

Die Aussagenlogik bietet einen formalen Rahmen, um mit solchen Aussagen mathematisch zu arbeiten. Anwendungen der Aussagenlogik finden sich zum Beispiel in der Analyse von Hardware und Software.
Vorab sei bemerkt, dass die Aussagenlogik keine semantische Bedeutung der Aussagen berücksichtigt. Aussagen werden als Variablen (oder sogenannte Atome) betrachtet und gemäß der im Folgenden eingeführten Regeln verwendet.
In der Aussagenlogik stellt man sich auf den Standpunkt, dass gewisse Aussagen, hier mit

$$A, B, C, ...$$

bezeichnet, vorliegen, die entweder *wahr* oder *falsch* sein können. Es handelt sich aus formaler Sicht um sogenannte **Aussagenvariablen**, die Werte in der Menge

$$\{\text{wahr, falsch}\}$$

annehmen. Diese Menge werden wir der Kürze halber mit

$$\{w, f\}$$

identifizieren, also

w steht für *wahr*

und

f steht für *falsch*.

Die Aussagenlogik gibt erstmal keine Auskunft über die Gültigkeit elementarer Aussagen, sondern betrachte diese als eine Variable, die entweder wahr oder falsch ist. Vielmehr dient sie als formales Werkzeug allgemeine Aussagen nach gewissen Gesetzen miteinander zu verknüpfen. Dabei werden die folgenden Operation verwendet.

Definition 2.1 Für Aussagen A, B bezeichnet

$$A \wedge B$$

die **Konjunktion** oder **„und“** -Verknüpfung, die durch folgende Wahrheitstabelle (oder Wahrheitstafel) definiert ist

A	B	$A \wedge B$
w	w	w
f	w	f
w	f	f
f	f	f

Definition 2.2 Für Aussagen A, B bezeichnet

$$A \vee B$$

die **Disjunktion** oder **„oder“** -Verknüpfung, die durch folgende Wahrheitswertetabelle definiert ist

A	B	$A \vee B$
w	w	w
f	w	w
w	f	w
f	f	f

Definition 2.3 Für eine Aussage A bezeichnet

$$\neg A$$

die **Negation** von A, also **„nicht“** A, welche durch folgende Wahrheitstabelle definiert ist

A	$\neg A$
w	f
f	w

.

Die oben definierten Verknüpfungen lassen sich auch sehr einfach als Abbildungen ausdrücken, was zeigt, dass der im ersten Abschnitt eingeführte generische Kontext von Mengen und Abbildungen auch Grundlage für die Aussagenlogik ist. Zum Beispiel ist die Konjunktion mathematisch nichts anderes als eine Abbildung

$$\varphi_\wedge : \{w, f\} \times \{w, f\} \to \{w, f\}$$

mit

$$\varphi_\wedge(A, B) := \begin{cases} w \ , & \text{falls } A = w \text{ und } B = w \\ f, & \text{ansonsten} \end{cases} .$$

Definition 2.4 Für Aussagen A, B bezeichnet

$$A \Rightarrow B$$

die **Implikation** oder **„wenn - dann“** -Verknüpfung, die durch folgende Wahrheitstabelle definiert ist

A	B	$A \Rightarrow B$
w	w	w
f	w	w
w	f	f
f	f	w

Diese Verknüpfung ist auch in einfacher Weise durch eine Wahrheitstabelle definiert. Die Namensgebung kann jedoch zu Verwirrungen führen, da der Sprachgebrauch bei Implikation eine Art kausalen Zusammenhang unterstellt. Dies ist hier aber formal nicht wichtig, die Implikation ist eine Verknüpfung, die durch die obige Definition festgelegt ist.

Definition 2.5 Für Aussagen A, B bezeichnet

$$A \Leftrightarrow B$$

die **Äquivalenz** oder **„genau dann, wenn“** -Verknüpfung, die durch folgende Wahrheitstabelle definiert ist

A	B	$A \Leftrightarrow B$
w	w	w
f	w	f
w	f	f
f	f	w

Die oben eingeführten logischen Operationen werden auch *Junktoren* genannt. In der Literatur, die sich vornehmlich auf Logik konzentriert, werden häufig die Symbole $\rightarrow$ und $\leftrightarrow$ anstelle von $\Rightarrow$ und $\Leftrightarrow$ verwendet. Jedoch sind letztere Symbole bereits für den Kontext der Abbildungen reserviert.

Durch Komposition dieser Verknüpfungen kann man nun komplexere Ausdrücke zusammenbauen. Hierbei setzt man folgende Konventionen:

$\wedge$ und $\vee$ werden vor $\Rightarrow$, $\Leftrightarrow$ angewendet,
$\neg$ wird vor $\wedge$ und $\vee$ angewendet,
Ausdrücke in ()-Klammern werden vor allen anderen ausgewertet.

Somit lassen sich für Aussagen A, B, C zum Beispiel (analog zu Termen wie $2 \cdot x + 3 + y$) Ausdrücke wie

$$\begin{aligned} F(A,B,C) &= \neg(A \wedge B) \Rightarrow C \\ F(A,B) &= \neg(A \wedge B) \vee B \\ F(A,B) &= \neg A \Leftrightarrow B \end{aligned} \tag{2.2}$$

erstellen.

Ausdrücke, die aus Aussagenvariablen und den oben eingeführten Verknüpfungen erzeugt werden, wie zum Beispiel in (2.2), werden als *Aussageformeln* bezeichnet. Eine solche Aussageformel ist selbst wieder eine Aussage und kann als Teil weiterer Aussageformeln verwendet werden.

Formaler heißt das:

Definition 2.6 Eine Abbildung $F : \{w, f\}^n \to \{w, f\}$, deren Abbildungsvorschrift sich als Ausdruck logischer Operatoren (Junktoren) darstellen lässt, heißt **Aussageformel**.

Zusammenfassung der logischen Operationen:

A	B	$A \wedge B$	$A \vee B$	$\neg A$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	w	w	f	w	w
f	w	f	w	w	w	f
w	f	f	w	f	f	f
f	f	f	f	w	w	w

Definition 2.7 Zwei Aussageformeln F_1, F_2 heißen **gleichwertig**, wenn für alle möglichen Besetzung von Wahrheitswerten in F_1, F_2 die Aussage

$$F_1 \Leftrightarrow F_2$$

immer wahr ist.

Der folgende Satz führt grundlegende gleichwertige Aussageformeln auf.

Satz 2.8 *Es seien A, B, C Aussagen (Aussagenvariablen), dann gelten (das heißt die Äquivalenzen sind stets wahr)*

1. $A \vee B \Leftrightarrow B \vee A$ *und* $A \wedge B \Leftrightarrow B \wedge A$ *(Kommutativgesetze)*
2. $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$ *und* $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$ *(Assoziativgesetze)*
3. $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$ *und* $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$ *(Distributivgesetze)*
4. $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$ *und* $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$ *(De Morgansche Gesetze)*

Diese letzten Regeln entsprechend den Aussagen aus Satz 1.7 und Satz 1.8 für Mengenoperationen, hierbei entsprechen sich $\cap$ und $\wedge$, $\cup$ und $\vee$, sowie „Komplement“und $\neg$.

Beweis. Wir zeigen hier nur das De Morgansche Gesetz $F_1(A, B) := \neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B =: F_2(A, B)$ mithilfe einer Wahrheitstabelle:

A	B	$A \wedge B$	F_1	$\neg A$	$\neg B$	F_2	$F_1 \Leftrightarrow F_2$
w	w	w	f	f	f	f	w
w	f	f	w	f	w	w	w
f	w	f	w	w	f	w	w
f	f	f	w	w	w	w	w

□

Beispiel 2.9 Wir nehmen an wir haben die Aufgabe einen digitalen Schaltkreis für eine einfache Türsteuerung zu implementieren. Die Tür soll sich öffnen, wenn zwei Bedingungen erfüllt sind: Ein Schalter ist aktiviert und ein bestimmter Zugangscode wurde eingegeben.
Um dieses Problem mit Hilfe der Aussagenlogik zu modellieren, können wir die folgenden Aussagen einführen:

- S: Der Schalter ist aktiviert.
- C: Der Zugangscode wurde eingegeben.

- O: Die Tür wird geöffnet.

Die Bedingungen für das Öffnen der Tür können dann in Form von Aussagenlogik ausgedrückt werden:
- Die Tür wird geöffnet, wenn sowohl der Schalter aktiviert ist als auch der Zugangscode eingegeben wurde. Dies wird durch die Aussage $(S \wedge C)$ repräsentiert. Die Logik lässt sich jedoch **nicht** wie folgt zusammenfassen

$$(S \wedge C) \implies O.$$

Übung: Warum nicht? Wie lautet die korrekte Aussageformel?

Beweisprinzipien

Mit Hilfe der Aussagenlogik können typische Prinzipien beim Beweisen mathematischer Sätze aus formaler Perspektive beschrieben werden.

Direkter Beweis

Wir nehmen an, ein mathematischer Satz behauptet die Gültigkeit einer Implikation $B \Rightarrow C$, wobei B und C zwei Aussagen sind. B könnte beispielsweise die Voraussetzungen des Satzes enthalten. Im Sinne der Aussagenlogik bedeutet das, dass $(B \Rightarrow C)$ wahr ist.
Um die Gültigkeit dieser Implikation zu beweisen, genügt es, den Fall zu betrachten, in dem B wahr ist. Denn ansonsten ist die Implikation immer wahr. Wir fordern also $B =$ wahr und leiten daraus eine neue Aussage A_1 her, die aus B folgt, also $(B \Rightarrow A_1)$ im Sinne der Aussagenlogik wahr ist. Eine solche Implikation muss dabei mathematisch direkt einsehbar sein (was dabei als *direkt einsehbar* gilt, ist natürlich eine Frage der Interpretation und hängt insbesondere von der Expertise der Leser des Beweises ab). Auf diese Weise fahren wir fort, bis wir schließlich zu einer Aussage A_n gelangen, von der wir nachvollziehbar behaupten können, dass $(A_n \Rightarrow C)$ wahr ist.

Am Ende haben wir dann gezeigt, dass die folgenden Implikationen wahr sind:

$$B \Rightarrow A_1 \Rightarrow A_2 \Rightarrow \cdots \Rightarrow A_n \Rightarrow C.$$

Es sei nochmals erwähnt, dass dabei die Richtigkeit jeder einzelnen Implikation jeweils direkt einsehbar sein muss.
Hierzu das folgende Beispiel:

Satz: Für ein $n \in \mathbb{N}$ gilt, ist n gerade, dann ist auch n^2 gerade.

In der Sprache der Aussagenlogik bedeutet das:

$$B := (n \text{ ist gerade}) \Rightarrow (n^2 \text{ ist gerade}) =: C.$$

Beweis. Es gilt:

$$B \Rightarrow (\text{es gibt ein } k \in \mathbb{N} \text{ mit } 2k = n) =: A_1.$$

Damit ist $n^2 = 4k^2$, und daher gilt

$$n^2/2 = 2k^2.$$

Das heißt aber

$$A_1 \Rightarrow (n^2/2 \in \mathbb{N}) := A_2.$$

Per Definition gilt weiter:

$$A_2 \Rightarrow C.$$

□

Das folgende Bespiel zeigt einen direkten Beweis zu Satz 1.24 (1)), der hier nochmal kurz formuliert wird:

Satz: (Sind $f : X \to Y$, $g : Y :\to Z$ injektiv) $\Rightarrow$ ($g \circ f : X \to Z$ ist injektiv).

In dem Oberen heißt das dann:

$$B = (f : X \to Y,\ g : Y :\to Z \,\text{sind injektiv}),$$

und

$$C = (g \circ f : X \to Z \text{ ist injektiv}).$$

Der Satz besagt damit kurz: $B \Rightarrow C$ ist wahr.
Beweis. Sei B wahr, dann sind folgende Implikationen wahr:

$$B \Rightarrow A_1^{'} := (\text{ für } x_1, x_2 \in X,\ x_1 \neq x_2 \text{ beliebig gilt } f(x_1) \neq f(x_2))$$

$$B \Rightarrow A_1^{''} := (\text{für, } y_1, y_2 \in Y\ y_1 \neq y_2 \text{ beliebig gilt } g(y_1) \neq g(y_2))$$

und somit ist Folgendes wahr:

$$B \Rightarrow A_1 := (A_1^{'} \wedge A_1^{''}).$$

Weiter gilt damit, dass folgende Implikation wahr ist:

$$A_1 \Rightarrow A_2 := (\text{für } x_1, x_2 \in X,\ x_1 \neq x_2 \text{ beliebig, gilt } g \circ f(x_1) \neq g \circ f(x_2)).$$

Nach Definition ist $A_2 \Leftrightarrow C$ wahr, also insbesondere

$$(A_2 \Rightarrow C) = wahr.$$

Damit haben wir gezeigt, dass

$$B \Rightarrow A_1 \Rightarrow A_2 \Rightarrow C$$

wahr ist. □

Äquivalenzbeweis

Wir nehmen an ein mathematischer Satz behaupte die Gültigkeit einer Äquivalenz $B \Leftrightarrow C$ (B, C zwei Aussagen).
Der Äquivalenzbeweis ist eine Variante des direkten Beweises. Es werden entweder die Implikationen $B \Rightarrow C$ und $C \Rightarrow B$ analog zum Obigen separat bewiesen. Oder man leitet direkt äquivalente Aussagen her so, dass am Ende gilt

$$B \Leftrightarrow A_1 \Leftrightarrow A_2 \Leftrightarrow ... \Leftrightarrow A_n \Leftrightarrow C$$

ist wahr. Dabei muss die Gültigkeit jeder einzelnen Äquivalenz jeweils direkt einsehbar sein.

Beweis durch Widerspruch

Wir nehmen an ein mathematischer Satz behaupte die Gültigkeit einer Aussage A.
A könnte beispielsweise wie oben eine Implikation sein $A = (B \Rightarrow D)$ sein.
Beim Beweis durch Widerspruch nimmt man an, dass $\neg A$ wahr ist und leitet dann, wie im direkten Beweis, neue Aussagen daraus ab, so dass man schließlich bei einer Aussage C landet, die sicher falsch ist:

$$\neg A \Rightarrow A_1 \Rightarrow A_2 \Rightarrow ... \Rightarrow A_n \Rightarrow C.$$

Dabei muss die Gültigkeit jeder einzelnen Implikation jeweils direkt einsehbar sein. Da C sicher falsch ist, kann $\neg A$ nicht wahr sein, denn ansonsten wäre ja die (als wahr erkannte) Implikation $\neg A \Rightarrow C$ falsch. Man erhält somit den Widerspruch zur Annahme ($\neg A =$ wahr). Es muss also $A = wahr$ gelten.

Hierzu das folgende Beispiel:

Satz: Es gibt keine größte Primzahl.

Hier ist

$$A = (\text{es gibt keine größte Primzahl}),$$

und

$$\neg A = (\text{es gibt eine größte Primzahl })$$

Beweis. Wir nehmen an, dass $\neg A = wahr$ gilt. Sei p diese größte Primzahl und seien weiter $1 < p_1 < p_2 < ... < p_n = p$ alle Primzahlen. Dann ist $q = p_1 \cdot p_2 \cdot ... \cdot p_n + 1$ eine Primzahl (dies könnte man genau genommen wiederum detailliert beweisen. Wir betrachten das hier als mathematisch direkt einsehbar). Wenn also $\neg A = wahr$ gilt,

folgt

$$C = (q \leq p) = wahr.$$

Da aber sicher gilt $C = falsch$, haben wir den Widerspruch erzeugt, und es muss somit $\neg A = falsch$ gelten. □

Beweis durch Kontraposition

Beweisen durch Kontraposition beruht auf der folgenden Äquivalenz

$$(B \Rightarrow C) \Leftrightarrow (\neg B \Leftarrow \neg C). \tag{2.3}$$

Wie beim direkten Beweis nehmen wir an ein mathematischer Satz behaupte die Gültigkeit einer Implikation $B \Rightarrow C$ (B, C zwei Aussagen). Man beweist aber stattdessen die Gültigkeit von $\neg B \Leftarrow \neg C$. Dazu geht man wie beim direkten Beweis vor. Man leitet also Implikationen

$$\neg C \Rightarrow A_1 \Rightarrow A_2 \Rightarrow ... \Rightarrow A_n \Rightarrow \neg B$$

her. Dabei muss die Gültigkeit jeder einzelnen Implikation jeweils direkt mathematisch einsehbar sein. Damit gilt dann im Sinne der Aussagenlogik

$$(\neg B \Leftarrow \neg C) = wahr,$$

und mit der obigen Äquivalenz (2.3) ist damit auch $B \Rightarrow C$ wahr.
Hierzu als Beispiel die umgekehrte Implikation des ersten Beispiels zum direkten Beweis:

Satz: Für ein $n \in \mathbb{N}$ gilt, ist n^2 gerade, dann ist auch n gerade.

In der Sprache der Aussagenlogik heißt das

$$B := (n^2 \text{ ist gerade}) \Rightarrow (n \text{ ist gerade}) =: C.$$

Beweis. Es gilt

$$\neg C \Rightarrow (\text{ es gibt ein } k \in \mathbb{N}_0 \text{ mit } 2k + 1 = n) =: A_1$$

Damit ist $n^2 = 4k^2 + 4k + 1$ und daher

$$n^2/2 = 2k^2 + 2k + 1/2,$$

und damit gilt

$$A_1 \Rightarrow (n^2/2 \notin \mathbb{N}) =: A_2.$$

Per Definition gilt weiter

$$A_2 \Rightarrow \neg B.$$

□

3 Relationen

Definition 3.1 Es seien M_1, M_2 nichtleere Mengen und $R \subset M_1 \times M_2$, dann heißt R **Relation** oder auch **binäre Relation**. Für Elemente $(x, y) \in R$ schreibt man auch

$$xRy.$$

Gilt dabei $M_1 = M_2$, so wird R auch **homogene Relation** genannt.

Der Begriff Relation wird an anderer Stelle auch allgemeiner, über den binären Fall hinausgehend, eingeführt. Für unsere Zwecke ist die obige Definition jedoch ausreichend.

Beispiel 3.2 Es sei M eine der Zahlenmengen aus (1.1) und

$$R_< := \{(x, y) \in M \times M : x < y\},$$

dann ist $R_<$ eine Relation. Dies ist die sogenannte **Ordnungsrelation**, die wir nachher in einem allgemeineren Rahmen definieren.

Nachfolgend werden nun einige grundlegende Eigenschaften von Relationen definiert:

Definition 3.3 Es seien M_1, M_2 nichtleere Mengen und $R \subset M_1 \times M_2$ eine Relation. Dann heißt R

1. **linkstotal** wenn gilt: $\forall x \in M_1 \, \exists y \in M_2 : (x, y) \in R$
2. **rechtstotal** wenn gilt: $\forall y \in M_2 \, \exists x \in M_1 : (x, y) \in R$
3. **rechtseindeutig** wenn gilt: $\forall x \in M_1$ und $\forall y, z \in M_2$ gilt $(x, y) \in R \wedge (x, z) \in R \Rightarrow y = z$
4. **linkseindeutig** wenn gilt: $\forall y \in M_2$ und $\forall x, z \in M_1$ gilt $(x, y) \in R \wedge (z, y) \in R \Rightarrow x = z$

Wir können nun den Begriff der **Abbildung** aus Definition 1.11 etwas allgemeiner auf Basis der Relationen definieren:

Definition 3.4 Es seien X und Y nichtleere Mengen und $R_f \subset X \times Y$ eine linkstotale und rechtseindeutige Relation. Dann heißt die Zuordnung

$$f(x) = y \ :\Leftrightarrow \ (x, y) \in R_f$$

Abbildung (Funktion) von X nach Y, Notation $f : X \to Y$. Die Menge R_f heißt **Graph** von f.

Es ergeben sich hieraus folgende Fragen (Übungen)

- Wie würden auf Grundlage von Definition 3.4, der Wertebereich, Urbild und Bildmenge definiert werden?
- Wie würde man mit Hilfe der Begriffe in Definition 3.3 die Injektivität oder Surjektivität einer Abbildung definieren?

Bemerkung 3.5 In einigen Teilgebieten der Informatik und Mathematik, insbesondere im Bereich der *Berechenbarkeitstheorie*, werden partiell definierte Funktionen untersucht. Diese können als rechtseindeutige Relationen charakterisiert werden, die nicht notwendigerweise *linkstotal* im Sinne von Definition 3.4 sein müssen. Die Betrachtung partiell definierter Funktionen ist von zentraler Bedeutung, da es Algorithmen gibt, die für bestimmte Eingaben nicht terminieren und somit kein Ergebnis liefern. In solchen Fällen existiert für einige Eingabewerte, d.h., die linke Komponente im Tupel der Relation, kein korrespondierender Ausgabewert, also keine rechte Komponente im Tupel.

Für homogene Relationen definieren wir das Folgende:

Definition 3.6 Es sei R eine (homogene) Relation auf einer Menge M. Dann heißt R

1. **reflexiv**, wenn für alle $x \in M$ gilt: $(x, x) \in R$,
2. **symmetrisch**, wenn für alle $x, y \in M$ gilt: $(x, y) \in R \Rightarrow (y, x) \in R$,
3. **transitiv**, wenn für alle $x, y, z \in M$ gilt: $(x, y) \in R$ und $(y, z) \in R \Rightarrow (x, z) \in R$,
4. **irreflexiv**, wenn für alle $x \in M$ gilt: $(x, x) \notin R$,
5. **konnex**, wenn für alle $x, y \in M$ gilt: $x \neq y \Rightarrow (x, y) \in R \vee (y, x) \in R$
6. **antisymmetrisch** (oder identitiv), wenn für alle $x, y \in M$ gilt: $(x, y) \in R \wedge (y, x) \in R \Rightarrow x = y$,
7. **asymmetrisch** wenn für alle $x, y \in M$ gilt: $(x, y) \in R \Rightarrow (y, x) \notin R$,
8. **total** wenn für alle $x, y \in M$ gilt: $(x, y) \in R \vee (y, x) \in R$,
9. **trichotom**, wenn für alle $x, y \in M$ genau eine der folgenden Aussagen gilt: $(x, y) \in R$, $(y, x) \in R$. $x = y$.

Satz 3.7 *Es sei R eine (homogene) Relation auf der Menge M. Dann gilt folgendes:*

1. *R ist asymmetrisch $\Leftrightarrow$ R ist antisymmetrisch und irreflexiv,*
2. *R ist trichotom $\Leftrightarrow$ R ist asymmetrisch und konnex,*
3. *R ist total $\Leftrightarrow$ R ist konnex und reflexiv,*
4. *falls zusätzlich $R \neq \emptyset$: R ist symmetrisch $\Rightarrow$ R ist nicht asymmetrisch,*
5. *falls zusätzlich $R \neq \emptyset$: R ist asymmetrisch $\Rightarrow$ R ist nicht symmetrisch,*
6. *falls zusätzlich $M \neq \emptyset$: R ist reflexiv $\Rightarrow$ R ist nicht irreflexiv,*
7. *falls zusätzlich $M \neq \emptyset$: R ist irreflexiv $\Rightarrow$ R ist nicht reflexiv.*

An dieser Stelle führen wir mit Hilfe der oben definierten Begriffe nur zwei spezielle Arten von Relationen ein:

Definition 3.8 Es sei $M \neq \emptyset$ und R eine Relation auf M.

1. Dann heißt R eine **schwache Totalordnung**, wenn sie antisymmetrisch, transitiv, total und reflexiv ist.
2. Dann heißt R eine **starke Totalordnung**, wenn sie trichotom und transitiv ist.

Weitere Kombinationen der Eigenschaften aus Definition 3.6 führen zu weiteren „Arten“ von Relationen. Allen mit *Ordnung* im Namen ist dabei gemein, dass sie transitiv sind.

Wir kommen nun zu einer weiteren wichtigen Relation.

Definition 3.9 Es sei M eine Menge. Eine **Äquivalenzrelation** auf M ist eine Relation $R \subset M \times M$, die reflexiv, symmetrisch und transitiv ist.

Die Idee der Äquivalenzrelation, wie sich bereits in der Namensgebung andeutet, ist es eine irgendwie geartete Äquivalenz zwischen Elementen einer Menge zu definieren.

Beispiel 3.10 Es sei P die Menge aller Computerprogramme, welche Eingaben einer festen Eingabemenge verarbeiten können und dazu eine Ausgabe in einer festen Ausgabemenge liefern. Es sei weiter $R \subset P \times P$ eine Relation definiert durch funktionale Äquivalenz: Zwei Programme p und q sind funktional äquivalent, wenn sie bei gleicher Eingabe immer die selbe Ausgabe liefern. Also $(p, q) \in R$ genau dann, wenn p, q funktional Äquivalent sind. Übung: Verifizieren Sie, dass R eine Äquivalenzrelation ist.

Ein wichtiger Vertreter einer Äquivalenzrelation ist die Folgende.

Bemerkung und Definition 3.11 Es sei $m \in \mathbb{Z} \setminus \{0\}$, dann heißt

$$R_m := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} : \, x - y \text{ ist ohne Rest durch } m \text{ teilbar}\}$$

die **Kongruenz-Relation modulo** m. Wie sich relativ einfach zeigen lässt, ist R_m eine Äquivalenzrelation. Man schreibt für $(x, y) \in R_m$

$$x \equiv y \mod m,$$

und sagt x ist **kongruent** y **modulo** m.

Zu einem gegebenen $x \in \mathbb{Z}$ ist

$$\{y \in \mathbb{Z} : (x, y) \in R_m\} = \{y \in \mathbb{Z} : y = x + k\, m,\, k \in \mathbb{Z}\}. \tag{3.4}$$

Man kann also hier durch Festsetzung eines Elements x mit Hilfe der Äquivalenzrelation die gesamte Menge auf der rechten Seite der vorhergehenden Gleichung ausdrücken. Diese Idee lässt sich auf allgemeine Äquivalenzrelationen erweitern und rechtfertigt den folgenden Begriff:

Definition 3.12 Es sei M eine nichtleere Menge und R ein Äquivalenzrelation auf M. Für ein $x \in M$ setzen wir

$$[x]_R := \{y \in M : (x, y) \in R\}.$$

Die Menge $[x]_R$ wird **Äquivalenzklasse** von x genannt. Für die Äquivalenzklassen bezüglich die Kongruenz-Relation modulo m, R_m ($m \in \mathbb{Z}$) aus Definition 3.11 schreiben wir der Kürze halber auch

$$[x]_m := [x]_{R_m}.$$

Satz 3.13 *Es sei M eine nichtleere Menge und R ein Äquivalenzrelation auf M. Dann gilt für zwei Äquivalenzklassen $[x]_R$, $[y]_R$*

$$\textit{entweder } [x]_R = [y]_R \textit{ oder } [x]_R \cap [y]_R = \emptyset.$$

Beweis: Siehe Übung.

4 Algebraische Strukturen

Indem wir Addition und Multiplikation anwenden, bedienen wir uns bereits seit unserer Schulzeit bestimmter algebraischer Strukturen auf den gewöhnlichen Zahlenmengen (1.1). Hierbei fällt auf, dass die Menge der ganzen Zahlen $\mathbb{Z}$ beispielsweise nicht ausreicht, um eine Gleichung wie

$$2 \cdot x = 1$$

zu lösen. Stattdessen benötigen wir eine reichhaltigere Menge wie $\mathbb{Q}$. Andererseits genügen die ganzen Zahlen für Gleichungen wie

$$2 + x = 0$$

was wiederum auf den natürlichen Zahlen nicht möglich ist. Allerdings lassen sich algebraische Operationen auch weit über die üblichen Zahlenmengen hinaus verallgemeinern. Ein Beispiel hierfür wäre das Addieren zweier Funktionen $f : \mathbb{R} \to \mathbb{R}$ und $g : \mathbb{R} \to \mathbb{R}$.
In der Mathematik strebt man stets danach, solche Strukturen in einem möglichst allgemeinen und natürlichen Rahmen zu definieren. Ergebnisse und Zusammenhänge, die in einem solchen generellen Kontext hergeleitet werden, lassen sich somit auf eine Vielzahl konkreter Anwendungsfälle übertragen. In diesem Kapitel werden wir solche grundlegenden Strukturen einführen und ihre ersten Eigenschaften herleiten.
Wir beginnen mit einer zentralen Struktur auf Mengen, die uns ein wichtiges Anwendungsbeispiel für allgemeinere algebraische Strukturen liefern wird.

Gruppen

Algebraische Strukturen werden typischerweise mittels sogenannter Verknüpfungen definiert.

Es sei M eine Menge, dann nennen wir eine Abbildung der Form

$$* : M \times M \to M$$

auch **Verknüpfung**, und anstelle von $*(x, y)$ schreiben wir dann $x * y$.

Man denke bei einer solchen Verknüpfung zum Beispiel an die Addition $+ : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto x + y$ oder die Multiplikation $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto x \cdot y$. Der obige Begriff der Verknüpfung wird an anderer Stelle auch allgemeiner verwendet.

Definition 4.1 Es sei G eine Menge und $* : G \times G :\to G$ eine Verknüpfung auf G. Dann heißt $(G, *)$ **Gruppe**, wenn gilt:

1. Für alle $x, y, z \in G$ gilt $(x * y) * z = x * (y * z)$ (Assoziativgesetz).
2. Es existiert ein **neutrales Element** $e \in G$ mit $e * x = x$ für alle $x \in G$.
3. Zu jedem $x \in G$ gibt es ein **inverses Element** $x^{-1} \in G$ mit $x^{-1} * x = e$.

Weiter heißt $(G, *)$ **abelsche Gruppe** oder **kommutative Gruppe**, wenn zusätzlich gilt:

4. $x * y = y * x$ für alle $x, y \in G$.

Beispiel 4.2
1. $(\mathbb{Z}, +)$ ist eine abelsche Gruppe.
2. $(\mathbb{R} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.
3. $(\mathbb{N}, +)$, $(\mathbb{N}_0, +)$ sind keine Gruppen.

Satz 4.3 *Es sei $(G, *)$ eine Gruppe, dann gilt:*

1. *Für jedes $x \in G$ gilt $x * x^{-1} = x^{-1} * x$.*
2. *Es gibt genau ein neutrales Element e und es gilt*

$$e * x = x * e = x$$

   für alle $x \in G$.

3. *Für jedes $x \in G$ gibt es genau ein inverses Element x^{-1} und es gilt*

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

Beweis. Im folgenden werden ausschließlich die Axiome in Definition 4.1 und die jeweils bereits bewiesenen Punkte angewendet. Zu (1):

$$\begin{aligned} e &= (x^{-1})^{-1} * x^{-1} = (x^{-1})^{-1} * e * x^{-1} = (x^{-1})^{-1} * (x^{-1} * x) * x^{-1} \\ &= ((x^{-1})^{-1} * x^{-1}) * x * x^{-1} = e * x * x^{-1} = x * x^{-1}. \end{aligned} \tag{4.5}$$

Zu (2): Es gilt mit $e = x * x^{-1}$

$$e * x = (x * x^{-1}) * x = x * (x^{-1} * x) = x * e.$$

Sei nun $\tilde{e}$ ebenfalls ein neutrales Element, d.h. es gilt $\tilde{e} * x = x \; \forall x \in G$. Dann gilt

$$\tilde{e} = \tilde{e} * e = e.$$

Zu (3): Zu einem beliebigen $x \in G$ sei $b \in G$ mit $b * x = e$. Dann gilt

$$b = b * e = b * (x * x^{-1}) = (b * x) * x^{-1} = e * x^{-1} = x^{-1}$$

womit die Eindeutigkeit von x^{-1} gezeigt ist. Da nach (1) gilt $e = x * x^{-1}$, muss wegen der gerade gezeigten Eindeutigkeit des Inversen $x = (x^{-1})^{-1}$ gelten.

□

Folgerung 4.4 Es sei $(G, *)$ eine Gruppe. Dann gilt:

1. für alle $x, y \in G$ gilt $(x * y)^{-1} = y^{-1} * x^{-1}$
2. die Gruppe ist genau dann kommutativ, wenn für alle $x, y \in G$ gilt $x * y * x^{-1} * y^{-1} = e$.

Beweis: Übung.

Definition 4.5 Eine bijektive Abbildung

$$\sigma : \{1, ..., n\} \to \{1, ..., n\}, \ (n \in \mathbb{N})$$

heißt **Permutation** (von $\{1, ..., n\}$). Weiter setzen wir

$$S_n := \{\sigma : \{1, ..., n\} \to \{1, ..., n\} : \sigma \text{ ist bijektiv}\},$$

die Menge aller Permutationen (von $\{1, ..., n\}$).

Für $\sigma \in S_n$ hat sich die folgende definierende Schreibweise etabliert:

$$\sigma = \begin{pmatrix} 1 & 2 & ... & n \\ \sigma(1) & \sigma(2) & ... & \sigma(n) \end{pmatrix}.$$

Satz 4.6 *Mit $\circ$, der Komposition von Abbildungen, ist $(S_n, \circ)$ eine Gruppe.*

Ringe und Körper

Auf einer Gruppe hat man nur eine Verknüpfung. Es sind jedoch auf vielen Fällen in natürlicher Weise zwei Verknüpfungen geben (zum Beispiel Addition und Multiplikation), die miteinander in Beziehung stehen. Eine wichtige algebraische Struktur, die diesen Fall berücksichtigt, wird in der folgenden Definition eingeführt.

Definition 4.7 Es sei R eine Menge und $\oplus : R \times R \to R$ und $\odot : R \times R \to R$ zwei Verknüpfungen. Dann ist $(R, \oplus, \odot)$ ein **Ring**, wenn gilt:

1. $(R, \oplus)$ ist eine abelsche Gruppe.
2. Für alle $x, y, z \in R$ gilt $x \odot (y \odot z) = (x \odot y) \odot z$ (Assoziativität).
3. Für alle $x, y, z \in R$ gelten $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$ und $(x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$ (Distributivgesetze).

Weiter heißt $(R, \oplus, \odot)$ **kommutativer Ring**, wenn zusätzlich gilt

4. $x \odot y = y \odot x$ für alle $x, y \in R$.

Ein wichtiges Beispiel eines Rings ist auf den Äquivalenzklassen der in (3.11) eingeführten Kongruenz-Relation definiert.

Bemerkung und Definition 4.8 Es sei $m \in \mathbb{Z} \setminus \{0\}$ und R_m die Kongruenz-Relation modulo m (vgl. 3.11). Dann heißt

$$\mathbb{Z}_m := \{[x]_m : x \in \mathbb{Z}\}$$

die Menge der **Restklassen** modulo m. Nach (3.4) ist

$$[x]_m := \{y \in \mathbb{Z} : y = x + k\,m,\ k \in \mathbb{Z}\}.$$

Auf $\mathbb{Z}_m$ definieren wir zwei Verknüpfungen wie folgt. Für $x, y \in \mathbb{Z}$ sei die Addition der zugehörigen Äquivalenzklassen definiert durch

$$[x]_m + [y]_m = [x + y]_m$$

und die Multiplikation sei definiert durch

$$[x]_m \cdot [y]_m = [x \cdot y]_m\,.$$

Damit ist $(\mathbb{Z}_m, +, \cdot)$ ein Ring und $\mathbb{Z}_m$ und wird als **Restklassenring** modulo m bezeichnet.

Wir führen nun die Struktur des Körpers ein, eine Struktur, die algebraisch das leistet was man von reellen oder rationalen Zahlen gewohnt ist.

Definition 4.9 Es sei K eine Menge mit mindestens zwei Elementen und $\oplus : K \times K \to K$ und $\odot : K \times K \to K$ zwei Verknüpfungen (genannt Addition und Multiplikation). Dann ist $(K, \oplus, \odot)$ ein **Körper**, wenn gilt:

1. $(K, \oplus)$ ist eine abelsche Gruppe. Dabei sei das neutrale Element bezüglich $\oplus$ mit 0 bezeichnet und für ein $x \in K$ sei das inverse Element (bezüglich $\oplus$) mit $-x$ bezeichnet.
2. Es gelten
   - $(x \odot y) \odot z = x \odot (y \odot z)$ für alle $x, y, z \in K$,
   - es gibt ein neutrales Element $1 \in \mathbb{K} \setminus \{0\}$ bezüglich $\odot$[a],
   - zu jedem $x \in \mathbb{K} \setminus \{0\}$ existiert ein $x^{-1} \in K$ so, dass $x^{-1} \odot x = 1$ [b],
   - für alle $x, y \in K$ ist $x \odot y = y \odot x$
3. Für alle $x, y, z \in K$ gelten $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$ und $(x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$ (Distributivgesetze).

[a] 1 ist hier ein Symbol (und nicht notwendigerweise die numerische Eins).
[b] „hoch -1“ist eine symbolische Schreibweise

Beispiel 4.10
1. $(\mathbb{R}, +, \cdot)$ ist ein Körper.
2. $(\mathbb{Q}, +, \cdot)$ ist ein Körper.
3. $(\mathbb{Z}, +, \cdot)$ ist kein Körper.

Ringstrukturen treten häufig im Fall von Mengen von Abbildungen auf. Hierbei kann häufig keine Körperstruktur definiert werden da inverse Element bezüglich einer Multiplikation nicht existieren. Allgemein ist ein Ring häufig nicht zu einem Körper erweiterbar, weil es nicht möglich ist zu *teilen*. Dazu das folgende Beispiel.

Beispiel 4.11 Es sei $X := \{f \text{ ist Abbildung} : f : [0,1] \to \mathbb{R}\}$ mit der (punktweisen) Addition

$$+ : X \times X \to X,\ (f+g)(x) = f(x) + g(x),\ x \in [0,1]$$

und der (punktweisen) Multiplikation

$$\cdot : X \times X \to X,\ (f \cdot g)(x) = f(x) \cdot g(x),\ x \in [0,1].$$

Damit ist $(X, +, \cdot)$ ein Ring mit neutralem Element bezüglich $+$ gegeben durch

$$f_0(x) : [0,1] \to \mathbb{R},\ f_0(x) = 0$$

und neutralem Element bezüglich $\cdot$ gegeben durch

$$f_1 : [0,1] \to \mathbb{R},\ f_1(x) = 1.$$

Man sieht jedoch leicht, dass $(X, +, \cdot)$ kein Körper ist, weil zum Beispiel jede Funktion, die eine Nullstelle hat ($x \in [0,1]$ mit $f(x) = 0$) kein inverses Element besitzt. Denn dies wäre ja eine Funktion $g : [0,1] \to \mathbb{R}$ mit $f(x)g(x) = 1$ für alle $x \in [0,1]$, was aber an der Nullstellen nicht möglich ist.

Bemerkung 4.12 Allgemein können wir an dieser Stelle folgendes definierten. Seien X, Y Mengen und

$$f, g : X \to Y$$

Abbildungen sowie $*$ eine Verknüpfung auf Y. Dann ist die **punktweise Verknüpfung** $f * g : X \to Y$ definiert durch

$$f * g(x) := f(x) * g(x), \quad \forall x \in X.$$

Man beachte, dass diese Definition etwa die Verknüpfungen der Abbildungen in Beispiel 4.11 einschließt.

Satz 4.13 *Es sei $(K, \oplus, \odot)$ ein Körper. Dann gelten:*

1. *Für alle $x \in K$ ist $x \odot 0 = 0$*
2. *Für alle $x, y \in K$ mit $x \odot y = 0$ folgt $x = 0$ oder $y = 0$*
3. *Für alle $x, y \in K$ gilt $-(x \odot y) = -x \odot y = x \odot (-y)$*

Beweis. Zu 1. : Es gilt

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

und somit

$$0 = (x \odot 0) - (x \odot 0) = (x \odot 0) + (x \odot 0) - (x \odot 0) = (x \odot 0).$$

Zu 2. : Gilt $y = 0$ so sind wir fertig. Für den Falle $y \neq 0$ gilt mit 1.:

$$x = x \odot e = x \odot (y \odot y^{-1}) = 0 \odot y^{-1} = 0.$$

Zu 3. : Übung. □

Folgerung 4.14 Nach Satz 4.13 (2) besagt Definition 4.9 (2)´, dass $(K \setminus \{0\}, \odot)$ eine abelsche Gruppe ist.

Bevor das nächste Ergebnis formuliert werden kann, müssen wir kurz den Begriff der Primzahl definieren.

Definition 4.15 Eine natürliche Zahl $p > 1$ heißt **Primzahl** wenn das Folgende gilt: Für $m \in \mathbb{N}$ ist $p \cdot m^{-1}$ ist eine natürliche Zahl genau dann, wenn $m = 1$ oder $m = p$. Die Menge der Primzahlen sei mit $\mathbb{P}$ bezeichnet.

Satz 4.16 *Für $m \in \mathbb{N}$ ist der Restklassenring $\mathbb{Z}_m$ (Definition 3.11) ein Körper, genau dann, wenn m eine Primzahl ist.*

Die Körper $\mathbb{Z}_p$ (p eine Primzahl) sind zentrale Vertreter der sogenannten endlichen Körper.
Beweisskizze Der wesentliche Punkt ist hierbei die Existenz inverser Element bzgl. der Multiplikation.
Gilt $m \notin \mathbb{P}$, so gibt es natürliche Zahlen $m > p, q > 1$ mit $pq = m$. Damit folgt

$$[p]_m \cdot [q]_m = [m]_m = [0]_m.$$

Wäre nun $\mathbb{Z}_m$ ein Körper, so würde nach Satz 4.13 2. gelten, dass $[p]_m = 0$ oder $[q]_m = 0$. Dies kann aber wegen $m > p, q > 0$ nicht sein.
Gilt $m \in \mathbb{P}$, so gilt für jedes $k \in \{1, ..., m-1\}$

$$\mathrm{ggT}(k, m) = 1$$

(ggT meint größter gemeinsamer Teiler). Mit dem erweiterten euklidischen Algorithmus (vgl. Anhang) folgt die Existenz von $l_1, l_2 \in \mathbb{Z}$ mit

$$l_1 m + l_2 k = 1.$$

Also gilt

$$[1]_m = [l_1 m]_m + [l_2 k]_m = [0]_m + [l_2]_m \, [k]_m = [l_2]_m \, [k]_m,$$

sodass $[l_2]_m$ multiplikative Inverses zu , $[k]_m$ ist. Die übrigen Körpereigenschaften folgen nun leicht.

Beispiel 4.17 Bei der Übertragung/Speicherung binärer Codes kann es aus unterschiedlichen Gründen zu Übertragungsfehlern kommen. Eines der ersten Verfahren zur Vermeidung von Fehlern bei der Übertragung solcher Codes geht auf Richard W. Hamming (1950) zurück, und ist daher unter dem Namen **Hamming-Code** bekannt. Angenommen, wir haben eine Nachricht, die wir als Binärcode der Länge vier übertragen möchten, etwa 1101, dann lässt sich die (7,4)-Hamming Codierung wie folgt darstellen

$$H_{7,4} : \mathbb{Z}_2^4 \to \mathbb{Z}_2^7,$$

$$H((s_1, s_2, s_2, s_4)) := \begin{pmatrix} [1]_2 & [0]_2 & [0]_2 & [0]_2 \\ [0]_2 & [1]_2 & [0]_2 & [0]_2 \\ [0]_2 & [0]_2 & [1]_2 & [0]_2 \\ [0]_2 & [0]_2 & [0]_2 & [1]_2 \\ [1]_2 & [1]_2 & [1]_2 & [0]_2 \\ [0]_2 & [1]_2 & [1]_2 & [1]_2 \\ [1]_2 & [0]_2 & [1]_2 & [1]_2 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix}$$

Dabei ist das Matrix-Vektor-Produkt wie in Lineare Algebra[1] definiert, wobei die Addition und Multiplikation bezüglich des Körpers $(\mathbb{Z}_2, +, \cdot)$ zu verstehen sind.
Die genauere Funktionsweise solcher Code-Korrekturverfahren wird in späteren Modulen eingehend besprochen.

Der Vollständigkeit zum Beweis von Satz 4.16 halber wird nachfolgend der erweiterte Euklidische Algorithmus formuliert:

[1] Diejenigen, die die Veranstaltung LA noch nicht besucht haben, werden dieses Beispiel später nachvollziehen können.

Homomorphismen

Im Verlauf dieser Vorlesung wurde bereits mehrfach auf die zentrale Bedeutung des Konzepts der Abbildung eingegangen. In diesem kurzen Unterabschnitt wollen wir nun die „natürlichen“Abbildungen auf den oben definierten algebraischen Strukturen einführen. Der Begriff „natürlich“bedeutet dabei, dass die Abbildungen die Struktur des Raums erhalten.

Definition 4.18 1. Es seien $(G, *)$ und $(H, \odot)$ zwei Gruppen und

$$\varphi : G \to H$$

eine Abbildung, so dass

$$\varphi(x * y) = \varphi(x) \odot \varphi(y) \text{ für alle } x, y \in G \text{ gilt.}$$

Dann ist φ ein **(Gruppen-)Homomorphismus**.

2. Es seien $(R, \oplus, *)$ und $(S, \times, \odot)$ zwei Ringe und

$$\varphi : R \to S$$

eine Abbildung, so dass

$$\varphi(x \oplus y) = \varphi(x) \times \varphi(y), \varphi(x * y) = \varphi(x) \odot \varphi(y), \text{ für alle } x, y \in R \text{ gilt.}$$

Dann ist φ ein **(Ring-)Homomorphismus**.

3. Es seien $(K, \oplus, *)$ und $(L, \times, \odot)$ zwei Körper und

$$\varphi : K \to L$$

eine Abbildung, so dass

$$\varphi(x \oplus y) = \varphi(x) \times \varphi(y), \varphi(x * y) = \varphi(x) \odot \varphi(y), \text{ für alle } x, y \in K.$$

Dann ist φ ein **(Körper-)Homomorphismus**.

Ein bijektiver Homomorphismus heißt **Isomorphismus**.

Satz 4.19 *Es seien* $(G, *)$ *und* $(H, \odot)$ *zwei Gruppen und*

$$\varphi : G \to H$$

ein Homomorphismus. Sei 0 *jeweils das neutrale Element auf* G *und* H*, und zu* $a \in G$ *oder* $a \in H$ *sei* $-a$ *jeweils das inverse Element. Dann gilt*

1. $\varphi(0) = 0$
2. $\varphi(-a) = -\varphi(a)$ *für alle* $a \in G$.

5 Beweisen mit vollständiger Induktion

Die vollständige Induktion ist ein Beweisverfahren, mit dem sich Gleichungen, Ungleichungen oder allgemeine Sätze beweisen lassen, die von einem n abhängen.
Zum Einstieg betrachten wir Folgendes: Es sei $n \in \mathbb{N}$. Dann gilt die

$$\sum_{\nu=1}^{n} \nu = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2} \quad \text{Gaußsche Summenformel.} \tag{5.6}$$

Für $n = 1$ oder $n = 2$ überzeugt man sich leicht von der Richtigkeit dieser Gleichung.
Die Idee der Induktion ist es, nun wie folgt zu schließen:

$$\text{Angenommen, es gilt für ein } n \in \mathbb{N} : \sum_{\nu=1}^{n} \nu = \frac{n(n+1)}{2},$$

$$\text{dann folgt: } \sum_{\nu=1}^{n+1} \nu = \frac{(n+1)(n+2)}{2}.$$

Gelingt es, dies zu zeigen, dann folgt mit der Verifikation von (5.6) für $n = 1$ durch sukzessives Anwenden der oberen Implikation, dass (5.6) allgemein für $n \in \mathbb{N}$ gilt.
Diese Idee lässt sich nun in ein allgemeines Beweisverfahren überführen, das als *vollständige Induktion*, oder kurz *Induktion*, bekannt ist.
Ziel ist es, die Gültigkeit einer Aussage $A(n)$ für alle $n \in J = \{n_0, n_1, n_2, \ldots\} \subset \mathbb{N}_0$ zu zeigen.

Vollständige Induktion

Induktionsanfang: Man zeigt die Gültigkeit von $A(n_0)$.

Induktionsannahme: Man nimmt an, dass $A(n_k)$ für ein allgemeines $n_k \in J$ gilt.

Induktionsschritt, $n_k \to n_{k+1}$: Man zeigt, dass unter der Induktionsannahme für n_k (also $A(n_k)$ gilt), folgt, dass auch $A(n_{k+1})$ gilt.

Aus Induktionsanfang und Induktionsschritt folgt dann, dass $A(n_k)$ für alle k gilt. Mit anderen Worten, die folgenden Implikationen sind als wahr erkannt:

$$A(n_0) \implies A(n_1) \implies \ldots \implies A(n_k),$$

und weil $A(n_0)$ wahr ist (Induktionsanfang), muss dann auch $A(n_k)$ für jedes k wahr sein.

Da der Induktionsanfang meist trivial ist und die Induktionsannahme ein Formalismus ist, steckt die „Arbeit“ im Induktionsschritt.
Im Folgenden werden nun einige wichtige Summenformeln und Ungleichungen eingeführt. Die zugehörigen Beweise werden dabei mittels vollständiger Induktion geführt.

Bemerkung 5.1 Wie wir später noch sehen werden, ist es sinnvoll für $x \in \mathbb{R}$ das Folgende zu setzen:

$$x^0 := 1.$$

Satz 5.2 *(Geometrische Summenformel) Es sei $q \in \mathbb{R}$, $q \neq 1$ und $n \in \mathbb{N}_0$, dann gilt*

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

Beweis.
Induktionsanfang: Für $n = 0$ ist

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

Induktionsannahme: Für ein $n \in \mathbb{N}_0$ gilt

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

Induktionsschritt $n \to n + 1$**:** Unter der Induktionsannahme gilt

$$\sum_{\nu=0}^{n+1} q^\nu = \frac{q^{n+1} - 1}{q - 1} + q^{n+1} = \frac{(q^{n+1} - 1) + q^{n+1}(q - 1)}{q - 1} = \frac{q^{n+2} - 1}{q - 1}.$$

□

Satz 5.3 *(Bernoullische Ungleichung) Es sei $a \in \mathbb{R}$ mit $a \geq -1$ und $n \in \mathbb{N}$, dann gilt*

$$(1 + a)^n \geq 1 + na.$$

Beweis.
Induktionsanfang: Für $n = 1$ ist

$$(1+a)^1 \geq 1+a.$$

Induktionsannahme: Für ein $n \in \mathbb{N}$, gilt

$$(1+a)^n \geq 1+na$$

Induktionsschritt: Unter der Induktionsannahme gilt

$$\begin{aligned}(1+a)^{n+1} = &\ (1+a)^n(1+a) \geq (1+na)(1+a)\\ = &\ 1+(n+1)a+na^2 \geq 1+(n+1)a\end{aligned}$$

□

Bevor wir die nächsten Ergebnisse formulieren können muss noch Folgendes eingeführt werden.

Definition 5.4 Für $n \in \mathbb{N}_0$ definiert man n-**Fakultät** durch

$$n! := \prod_{\nu=1}^{n} \nu\ .$$

und

$$0! := 1.$$

Beispiel 5.5 Es gilt also etwa $6! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720$.

Definition 5.6 Für $n, \nu \in \mathbb{N}_0$ gilt

$$\binom{n}{\nu} = \begin{cases} \dfrac{n!}{\nu!(n-\nu)!} = \binom{n}{n-\nu}, & \text{falls } \ \nu \leq n \\ 0, & \text{falls } \ \nu > n \end{cases}.$$

Beispiel 5.7 Es gilt also

$$\binom{7}{5} = \frac{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3}{5!} = 21\ .$$

In einem kombinatorischen Experiment, bei dem Stichproben aus einer Urne mit n nummerierten Kugeln gezogen werden, entspricht

$$\binom{n}{\nu}$$

der Anzahl von Möglichkeiten **ungeordnete Stichproben** bestehend aus ν Kugeln zu ziehen.
Dabei bedeutet **ungeordnet**, dass beispielsweise die Ziehungen 1,8,4 und 8,4,1 als gleich betrachtet werden. Hingegen wird bei **geordneten Stichproben** die Reihenfolge der Ziehung berücksichtigt, sodass 1,8,4 und 8,4,1 als unterschiedliche Ziehungen betrachtet werden.
Die folgende Tabelle fasst die Anzahl der möglichen Ziehungen von ν Kugeln aus n nummerierten Kugeln in unterschiedlichen Urnenexperimenten zusammen.

mit Zurücklegen	**geordnet**	**Anzahl Möglichkeiten**
ja	ja	n^ν
nein	ja	$\frac{n!}{(n-\nu)!}$
ja	nein	$\binom{n+\nu-1}{\nu}$
nein	nein	$\binom{n}{\nu}$

Wir stellen einige Eigenschaften der Binomialkoeffizienten zusammen.
Ordnet man die Binomialkoeffizienten $\binom{n}{\nu}$ in einem dreieckigen Schema an, wobei in der n-ten Zeile die Koeffizienten $\binom{n}{0}, \ldots, \binom{n}{n}$ stehen, so entsteht das sogenannte *Pascalsche Dreieck*:

$$\begin{array}{ccccccccc}
\binom{0}{0} & & & & & & & & \\
\binom{1}{0} & & \binom{1}{1} & & & & & & \\
\binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & \\
\vdots & & \vdots & & \vdots & & & & \\
\vdots & & & \vdots & & & \vdots & & \\
\binom{n}{0} & \binom{n}{1} & \cdots & \binom{n}{\nu-1} & \binom{n}{\nu} & \cdots & \binom{n}{n} & & \\
\binom{n+1}{0} & \binom{n+1}{1} & \cdots & \binom{n+1}{\nu} & \cdots & \binom{n+1}{n+1} & & &
\end{array}$$

Die ersten Zeilen berechnen sich nach Definition 5.6 zu

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
1 & 5 & & 10 & & 10 & & 5 & 1 \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}
$$

Im Pascalschen Dreieck erkennt man leicht folgende Identität für $n > \nu > 0$:

$$\boxed{\binom{n}{\nu} = \binom{n-1}{\nu} + \binom{n-1}{\nu-1}} \tag{5.7}$$

Der Beweis dazu dient als Übung. Wir beweisen damit nun folgendes Resultat.

Satz 5.8 *(binomischer Satz) Für alle $a, b \in \mathbb{R}$ und alle $n \in \mathbb{N}_0$ gilt*

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

Beweis.
Induktionsanfang: Für $n = 0$ gilt $(a+b)^0 = 1 = \sum_{\nu=0}^{0} \binom{0}{\nu} a^\nu b^{0-\nu}$.
Induktionsannahme: Für ein $n \in \mathbb{N}_0$ gelte $(a+b)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} a^\nu b^{n-\nu}$.
Induktionsschritt: Unter der Induktionsannahme gilt durch Anwendung von Satz

5.8

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

Also gilt die Behauptung dann auch für $n+1$. □

Beispiel 5.9 Für $n = 6$ gilt

$$\begin{aligned}
(a+b)^6 &= \sum_{\nu=0}^{6}\binom{6}{\nu}a^\nu b^{6-\nu} \\
&= 1\cdot b^6 + 6\cdot ab^5 + 15a^2b^4 + 20a^3b^3 + 15a^4b^2 + 6a^5b + 1\cdot a^6 .
\end{aligned}$$

Bemerkung 5.10 Als Spezialfälle aus dem binomischen Satz ergeben sich interessante Beziehungen für das Pascalsche Dreieck: Für $a = 1, b = 1$ ergibt sich

$$2^n = (1+1)^n = \sum_{\nu=0}^{n}\binom{n}{\nu}1^\nu 1^{n-\nu} = \sum_{\nu=0}^{n}\binom{n}{\nu} ,$$

d. h. die Summe der Binomialkoeffizienten in der n-ten Zeile des Pascalschen Dreiecks ergibt stets 2^n. Für $a = -1, b = 1$ ergibt sich für $n \in \mathbb{N}$

$$0 = 0^n = \big((-1)+1\big)^n = \sum_{\nu=0}^{n}\binom{n}{\nu}(-1)^\nu ,$$

d. h. versieht man die Binomialkoeffizienten in der n-ten Zeile jeweils abwechselnd mit dem Vorzeichen + und −, so erhält man als Summe 0. Für $n = 6$ gilt etwa

$$1 + 6 + 15 + 20 + 15 + 6 + 1 = 64 = 2^6$$

und

$$1 - 6 + 15 - 20 + 15 - 6 + 1 = 0 \; .$$

6 Komplexe Zahlen

Wir haben oben gesehen, dass das Lösen der Gleichung

$$4 \cdot x = 1$$

einer Erweiterung der Gruppe $(\mathbb{Z}, +)$ zum Körper $(\mathbb{Q}, +, \cdot)$ bedurfte. Weiter bedurfte das Lösen der Gleichung

$$x^2 = 2$$

einer Erweiterung von $(\mathbb{Q}, +, \cdot)$ nach $(\mathbb{R}, +, \cdot)$.
In diesem Abschnitt wollen wir $(\mathbb{R}, +, \cdot)$ in einer Weise erweitern, dass die Gleichung

$$x^2 = -1 \tag{6.8}$$

lösbar ist. Dies führt zum Körper der **komplexen Zahlen**. Dazu definieren wir das Folgende:

Definition 6.1 Die **imaginäre Einheit** i ist definiert als die Lösung der Gleichung (6.8), d.h. es gilt

$$i^2 = -1.$$

Eine **komplexe** Zahl ist ein Element (x, iy) mit $x, y \in \mathbb{R}$, und wir setzen dafür $(x, iy) =: x + iy$. Für $z := x + iy$ heißen:

1. $\mathrm{Re}(z) := x$ der **Realteil** von z,
2. $\mathrm{Im}(z) := y$ der **Imaginärteil** von z.

Wir setzten

$$\mathbb{C} := \{z = x + iy : x \in \mathbb{R},\, y \in \mathbb{R}\},$$

die Menge der komplexen Zahlen.

Wie zu Beginn des Abschnitts bemerkt, wollen wir auf $\mathbb{C}$ die algebraische Struktur eine Körpers definieren, und benötigen daher entsprechende algebraische Operationen $+$ und $\cdot$.

Definition 6.2 Für $z = x + iy$ und $w = u + iv$ aus $\mathbb{C}$ definieren wir

1. $z + w := x + u + i(y + v)$ **Addition**,
2. $z\,w := z \cdot w := x\,u + i(x\,v + u\,y) - v\,y$ **Multiplikation**.

Damit gilt:

Satz 6.3 *$(\mathbb{C}, +, \cdot)$ ist ein Ring mit neutralem Element $0 + i0$ bezüglich der Addition $+$, und $1 + i0$ als neutralem Element bezüglich der Multiplikation $\cdot$.*

Der Kürze halber schreiben wir auch

$$x + i0 = x, \quad 0 + iy = iy, \quad 0 + i0 = 0.$$

Die Erweiterung zum Körper ergibt sich nun dadurch, dass man für alle $z \in \mathbb{C} \setminus \{0\}$ ein inverses Element bezüglich der Multiplikation finden kann:

Satz und Definition 6.4 Es sei $z = x + iy \in \mathbb{C} \setminus \{0\}$, und

$$z^{-1} := \frac{1}{z} = \frac{x}{x^2 + y^2} + i\frac{-y}{x^2 + y^2},$$

dann gilt

$$z \cdot z^{-1} = 1.$$

Somit ist z^{-1} das inverse Element von z bezüglich $\cdot$. Damit ist $(\mathbb{C}, +, \cdot)$ ein Körper.

Beweis. Es gilt

$$\begin{aligned} zz^{-1} &= (x + iy)\left(\frac{x}{x^2 + y^2} + i\frac{-y}{x^2 + y^2}\right) \\ &= \frac{x^2}{x^2 + y^2} + \frac{-xy}{x^2 + y^2} - \frac{-xy}{x^2 + y^2} + \frac{y^2}{x^2 + y^2} \\ &= \frac{x^2 + y^2}{x^2 + y^2} = 1. \end{aligned} \tag{6.9}$$

Dass damit nun weiter sämtliche Körperaxiome erfüllt sind kann leicht nachgerechnet werden (Übung). □

Für $z, w \in \mathbb{C}$, $w \neq 0$ schreiben wir auch

$$z\, w^{-1} = z\, \frac{1}{w} = \frac{z}{w}.$$

Da $(\mathbb{C}, +, \cdot)$ ein Körper ist, und weiterhin die Konvention *Punkt- vor Strichrechnung* gilt, können die Regeln der Bruchrechnung unmittelbar auf das Rechnen mit komplexen Zahlen übertragen werden. Damit können wir mit den komplexen Zahlen rechnen wie in $\mathbb{R}$ gewohnt und es gelten beispielsweise:

$$\begin{aligned} \frac{1}{z} + \frac{1}{w} &= \tfrac{z+w}{zw} && (z, w \neq 0), \\ \frac{1}{z} \cdot \frac{1}{w} &= \tfrac{1}{zw} && (z, w \neq 0), \\ \frac{z+w}{w} &= \tfrac{z}{w} + 1 && (w \neq 0) \\ \frac{1}{\frac{1}{w}} &= w && (w \neq 0) \end{aligned} \tag{6.10}$$

Definition 6.5 Es sei $z = x + iy$ eine komplexe Zahl.

1. Die komplexe Zahl $\overline{z} := x - iy$ heißt zu z **konjugiert komplex**.
2. Die Zahl $|z| := \sqrt{x^2 + y^2} \in [0, \infty)$ heißt **Betrag** von z.

Bemerkung 6.6 Für $z, w \in \mathbb{C}$ ergebe sich leicht (Übung)

$$\overline{z + w} = \overline{z} + \overline{w}, \qquad \overline{zw} = \overline{z} \cdot \overline{w}, \qquad \overline{(\overline{z})} = z,$$

sowie

$$\mathrm{Re}(z) = \frac{1}{2}\,(z + \overline{z}) \quad \text{und} \quad \mathrm{Im}(z) = \frac{1}{2i}\,(z - \overline{z})\,.$$

Für den Betrag gelten folgende Rechenregeln:

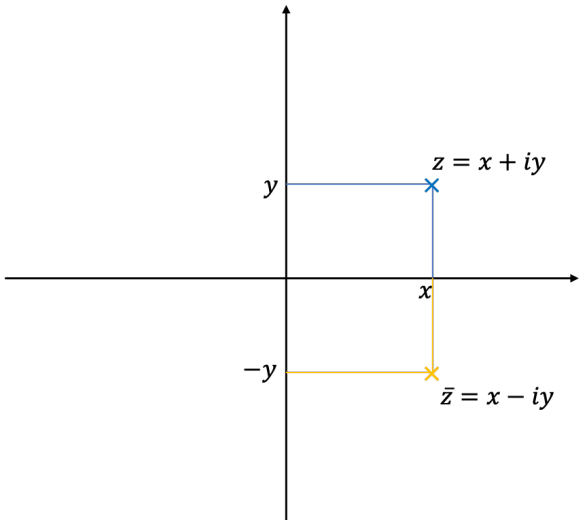

Abbildung 4: Gaußsche Zahlenebene (oder komplexe Ebene)

Satz 6.7 *Für $z, w \in \mathbb{C}$ gelten*

1. $|z| \geq 0$, *und dabei ist* $|z| = 0$ *genau dann, wenn* $z = 0$ *ist,*
2. $|z| = |\overline{z}|\,, \quad |z| = |-z|\,, \quad |\mathrm{Re}\, z| \leq |z|\,, \quad |\mathrm{Im}\, z| \leq |z|,$
3. $|z|^2 = z\overline{z}$ *und* $1/z = \overline{z}/|z|^2$, *falls* $z \neq 0$,
4. $|zw| = |z|\,|w|$ *und* $|z+w|^2 = |z|^2 + 2\mathrm{Re}(z\overline{w}) + |w|^2$,
5. *(Dreiecksungleichung)* $|z \pm w| \leq |z| + |w|$.

Beweis. 1., 2. und 3. dienen als Übung.
Zu 4. Es gilt nach 3.

$$|zw|^2 = (zw)(\overline{zw}) = (z\overline{z})(w\overline{w}) = |z|^2|w|^2 = (|z||w|)^2\,.$$

Durch Wurzelziehen folgt die erste Behauptung. Weiter gilt

$$|z+w|^2 = (z+w)(\overline{z}+\overline{w}) = z\overline{z} + z\overline{w} + \overline{z\overline{w}} + w\overline{w} = |z|^2 + 2\,\mathrm{Re}\,(z\overline{w}) + |w|^2\,.$$

Zu 5. Nach 2. und 4. ist

$$|z+w|^2 \leq |z|^2 + 2|z\overline{w}| + |w|^2 = |z|^2 + 2|z||w| + |w|^2 = (|z| + |w|)^2$$

Durch Wurzelziehen folgt die Behauptung für $z + w$. Damit erhält man dann auch

$$|z - w| \leq |z| + |-w| = |z| + |w| \, .$$

□

Beispiel 6.8 Für $z = 3 - i$ gilt

$$|z| = \sqrt{9+1} = \sqrt{10}, \qquad \overline{z} = 3 - i(-1) = 3 + i$$

und

$$z\overline{z} = (3-i)(3+i) = 9 + 1 \left(= |z|^2\right) .$$

Wir kommen nun zu einem zentralen Ergebnis der Mathematik. Wie zu Beginn des Abschnitts erwähnt, haben die bisherigen Erweiterungen von $\mathbb{Z}$ über $\mathbb{Q}$ nach $\mathbb{R}$ und schließlich nach $\mathbb{C}$ uns ermöglicht Gleichungen der Form

$$4 \cdot x = 1, \quad x^2 = 2, \quad x^2 = -1$$

zu lösen. Wir werden nun sehen, dass die komplexen Zahlen ein natürliches Ende dieser Erweiterungen darstellt, denn hierin sind sämtliche (Linear-)Kombinationen solcher Gleichungen stets lösbar.
Dazu zunächst folgende Definition

Definition 6.9 Eine **Polynomfunktion** (oder kurz **Polynom**) ist eine Funktion $P : \mathbb{C} \to \mathbb{C}$ der Form

$$P(z) = \sum_{\nu=0}^{d} a_\nu z^\nu$$

mit $a_0, \ldots, a_d \in \mathbb{C}$. Ist $a_d \neq 0$, so heißt $\deg(P) := d$ der **Grad** von P, und $a_0, \ldots, a_d$ sind die **Koeffizienten** von P. Ein $z \in \mathbb{C}$ heißt **Nullstelle** von P, falls $P(z) = 0$ gilt.

In der oberen Skizze sehen wir, dass das Polynom $P(x) = x^2 - 1$ genau zwei Nullstellen an ± 1 hat, und somit die Gleichung

$$0 = x^2 - 1$$

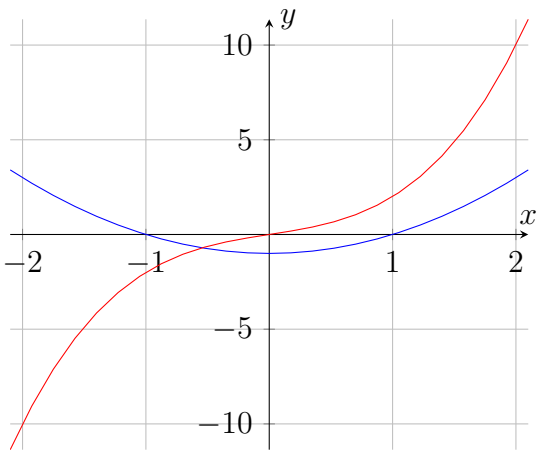

Abbildung 5: Graphen der Polynome $P(x) = x^2 - 1$ (blau), und $P(x) = x^3 + x$ (rot), jeweils als Funktion der reellen Zahlen.

genau zwei Lösungen in $\mathbb{R}$ hat. Hingegen hat das Polynom $P(x) = x^3 + x$ nur eine reelle Nullstelle an 0 und damit hat

$$0 = x^3 + x$$

nur eine reelle Lösung. Beachtet man jedoch, dass

$$P(x) = x^3 + x = x(x - i)(x + i)$$

gilt, so sieht man, dass die letzte Gleichung an $\pm i$ noch zwei weitere Lösungen in $\mathbb{C}$ hat. Diese Beobachtung lässt sich weitreichend verallgemeinern:

Satz 6.10 *(Fundamentalsatz der Algebra und Faktorisierungssatz) Es sei*

$$P(z) = \sum_{\nu=0}^{d} a_\nu z^\nu$$

ein Polynom vom Grad $d > 0$. Dann existieren $z_1, ..., z_k \in \mathbb{C}$ und zugehörige $\alpha_1, ..., \alpha_k \in \mathbb{N}$, so dass

$$P(z) = a_d(z - z_1)^{\alpha_1} (z - z_2)^{\alpha_2} \cdot ... \cdot (z - z_k)^{\alpha_k}$$

gilt. Dabei gilt $\alpha_1 + \alpha_2 + ... + \alpha_k = d$.

Definition 6.11 Die $\alpha_1, \alpha_2, ..., \alpha_k$ in Satz 6.10 heißen **Vielfachheiten** der Nullstellen $z_1, z_2, ..., z_k$.

Beispiel 6.12 Es sei $P(z) = z^2 - c$ mit $c \geq 0$ fest. Dann gilt $P(z) = (z - \sqrt{c})(z + \sqrt{c})$. Somit sind für $c >$, $\sqrt{c}$ und $-\sqrt{c}$ zwei Nullstellen mit jeweils Vielfachheit 1, welche für $c \to 0$ zu einer Nullstelle der Vielfachheit 2 werden.

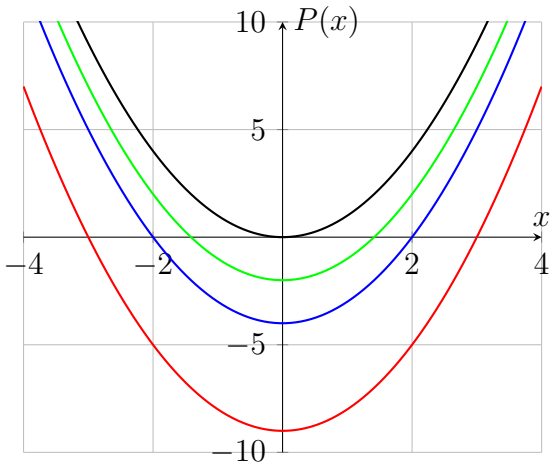

Abbildung 6: Graphen zu $P(z) = z^2 - c$ als Funktion $\mathbb{R} \to \mathbb{R}$ für $c = 9, 4, 2, 0$ (rot, blau, grün, schwarz)

Beispiel 6.13

1. Das Polynom $P(z) = z^2 + 9 = (z + i3)(z - i3)$ hat einfache Nullstellen an $3i$ und $-3i$

2. Das Polynom $P(z) = z^3 + z^2 + 4z + 4$ hat nur eine einfache reelle Nullstelle an $z = -1$. Nach dem dem Fundamentalsatz der Algebra (Satz 6.10) muss P noch zwei weitere Nullstellen (inklusive Vielfachheit) haben. Es gilt

$$P(z) = (z + 1)(z + 2i)(z - 2i),$$

woran man erkennt, dass P die beiden komplexen Nullstellen an $z = 2i$ und $z = -2i$ hat.

3. Das Polynom $P(z) = z^3 + iz^2 + 4z + i4 = (z+i)(z-2i)(z+2i)$ hat ausschließlich komplexe Nullstellen an $z = -i$, $z = i$ und $z = -i$.

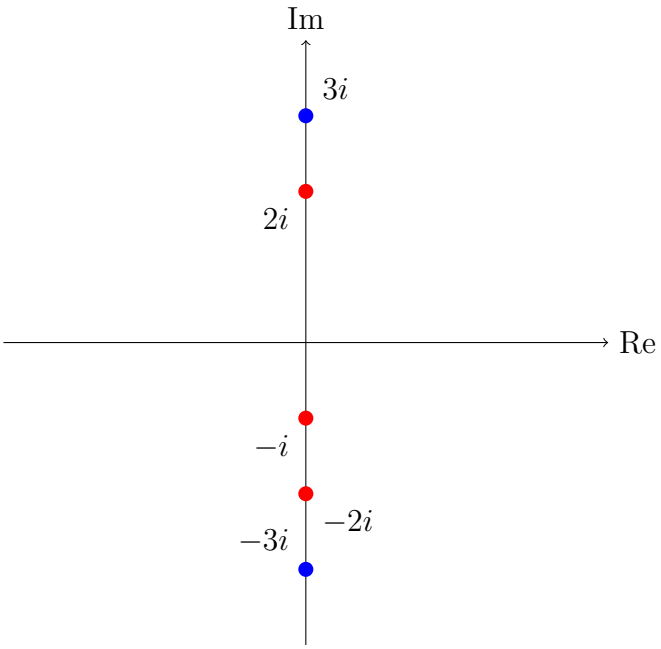

Abbildung 7: Nullenstellen der Polynome aus Beispiel 6.13, (blau) $P(z) = z^2 + 9 = (z + i3)(z - i3)$, (rot) $P(z) = z^3 + iz^2 + 4z + i4 = (z+i)(z-2i)(z+2i)$

In dem oberen Beispiel beobachtet man, dass die Nullstellen von Polynomen mit reellen Koeffizient stets symmetrisch bezüglich der reellen Achse in $\mathbb{C}$ liegen. Dies gilt auch im Allgemeinen:

Satz 6.14 *Es sei*

$$P(z) = \sum_{\nu=0}^{d} a_\nu z^\nu$$

ein Polynom mit $a_\nu \in \mathbb{R}$ für $\nu = 0, ..., d$. Dann gilt $P(z) = 0$ genau dann, wenn $P(\overline{z}) = 0$ gilt.

Beweis. Durch mehrfache Anwendung von Bemerkung 6.6 und wegen $\overline{a_\nu} = a_\nu$ $\nu =$

$0, ..., d$ ergibt sich:

$$\begin{aligned}
0 &= P(z) \\
\Leftrightarrow 0 &= \overline{P(z)} \\
\Leftrightarrow 0 &= \overline{\sum_{\nu=0}^{d} a_\nu z^\nu} \\
\Leftrightarrow 0 &= \sum_{\nu=0}^{d} \overline{a_\nu z^\nu} \\
\Leftrightarrow 0 &= \sum_{\nu=0}^{d} \overline{a_\nu}\,\overline{z^\nu} \\
\Leftrightarrow 0 &= \sum_{\nu=0}^{d} a_\nu \overline{z}^\nu \\
\Leftrightarrow 0 &= P(\overline{z})
\end{aligned}$$

□

7 Folgen und Reihen

Folgen und deren Grenzwerte sind für uns aus folgenden Gründen essenziell:

1. Viele Verfahren und Algorithmen erzeugen Folgen (von Ausgaben), und der Grenzwert der Folge ist in diesen Fällen typischerweise das angestrebte Ergebnis.

2. Mittels Folgen werden wir Grenzwerte für allgemeine Abbildungen definieren, auf Basis derer die gesamte Differential- und Integralrechnung (in einer und mehrerer Veränderlichen) beruhen. Diese wiederum sind wesentliche und unverzichtbare Werkzeuge der Physik, Optimierung, maschinelles Lernen/KI und vielen weiteren Bereichen.

Konvergenz und Grenzwerte von Folgen

Definition 7.1 Sei $N \subset \mathbb{N}_0$ abzählbar unendlich, so nennen wir eine Abbildung

$$a : N \to \mathbb{R}^d \text{ (oder } \mathbb{C}^d)$$

eine **Folge**. Für die Auswertung der Folge schreibt man typischerweise

$$a_n := a(n) \qquad (n \in N)$$

und die a_n heißen dann **Folgeglieder** und N heißt **Indexmenge**. Anstatt a schreiben wir :

$$(a_n)_{n\in N} := a \text{ oder kurz } (a_n) := (a_n)_{n\in N}.$$

Ist $J \subset N$ abzählbar unendlich, dann heißt

$$(a_n)_{n\in J}$$

eine **Teilfolge** von $(a_n)_{n\in N}$.

Beispiel 7.2 1. $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n = \frac{1}{n}$$

2. $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n = n^2$$

3. $(a_n)_{n\in\mathbb{N}}$ mit
$$a_n = (-1)^n + i^n.$$

4. $(a_n)_{n\in\mathbb{N}}$ mit
$$a_n = \begin{pmatrix} 1/n \\ (1+1/n)^n \end{pmatrix}$$

Bemerkung 7.3 Folgen sind nach dem Obigen spezielle Abbildungen (oder Funktionen) mit einem abzählbar unendlichen Definitionsbereich in $\mathbb{N}_0$. Somit sind alle diesbezüglichen bisherigen Definitionen und Resultate hierauf anwendbar.

Da wir an dieser Stelle mit Elementen in $\mathbb{R}^d$ (oder $\mathbb{C}^d$) arbeiten, definieren wir an dieser Stelle kurz die darauf üblichen algebraischen Operation solcher Elemente (vgl. lineare Algebra).

Definition 7.4 Für $x = (x_1, ..., x_d)$, $y = (y_1, ..., y_d)$ in $\mathbb{R}^d$ (oder $\mathbb{C}^d$) und $\lambda \in \mathbb{R}$ (oder in $\mathbb{C}$) ist

1. $\lambda x = (\lambda x_1, ..., \lambda x_d)$,
2. $x \pm y = (x_1 \pm y_1, ..., x_d \pm y_d)$,
3. und
$$\|x\| = \sqrt{\sum_{j=1}^{d} |x_j|^2}$$
die Euklidische Norm (kurz Norm) von x.

Bemerkung 7.5 Es seien $x = (x_1, ..., x_d)$, $y = (y_1, ..., y_d)$ in $\mathbb{R}^d$ (oder $\mathbb{C}^d$) .

1. $\|x - y\|$ ist der **euklidische Abstand** zwischen x, y und stimmt für $d = 1, 2, 3$ mit dem überein, was man anschaulich mit Abstand meint.
2. Für $d = 1$ ist $\|x\| = |x|$.
3. $\|x\| \geq 0$ wobei nur für $0 = x_1 = x_2 = ... = x_d$ gilt $\|x\| = 0$.

4. $\|x \pm y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)

Definition 7.6 Es sei $(a_n)_{n\in N}$ eine Folge. Dann heißt $c \in \mathbb{R}^d$ (oder $c \in \mathbb{C}^d$) **Grenzwert** von (a_n), wenn gilt

$$\forall \varepsilon > 0\, \exists n_\varepsilon \in N,\ \forall n \geq n_\varepsilon,\, n \in N :\ \|c - a_n\| < \varepsilon.$$

(*Textuell:* Für alle $\varepsilon > 0$ existiert ein $n_\varepsilon \in N$ so, dass für alle $n \in N$ mit $n \geq n_\varepsilon$ gilt $\|a_n - c\| < \varepsilon$). Wir schreiben

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

und sagen $(a_n)_{n\in N}$ **strebt** gegen c. Existiert ein solches c, dann heißt die Folge **konvergent**. Gilt dabei $c = 0$, so heißt die Folge auch **Nullfolge**.

Bemerkung 7.7 Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$. Wir sagen (a_n) strebt gegen $+\infty$ $(-\infty)$, falls

$$\forall_{m\in\mathbb{N}}, \exists_{n_m\in N}, \forall_{n\geq n_m, n\in N} : a_n > m\, (a_n < -m).$$

(*Textuell:* Für jedes $m \in \mathbb{N}$ existiert ein $n_m \in N$ so, dass für alle $n \in N$ mit $n \geq$ gilt $a_n > m$ $(a_n < -m)$)
Wir schreiben dafür:

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

Folgen, die gegen $+\infty$ oder $-\infty$ streben gelten jedoch **nicht als konvergent** in $\mathbb{R}$. Das liegt daran, dass $+\infty$, $-\infty$ keine Elemente der reellen Zahlen sind.

Satz 7.8 *Es sei (a_n) eine Folge in $\mathbb{R}^d$ $(\mathbb{C}^d)$ und für jedes n sei $a_n = (a_n^{(1)}, ..., a_n^{(d)})$. Dann gilt für $c = (c^{(1)}, ..., c^{(d)})$*

$$\lim_{n\to\infty} a_n = c$$

genau dann, wenn

$$\forall\, j = 1, ..., d :\quad \lim_{n\to\infty} a_n^{(j)} = c^{(j)}.$$

Wir reduzieren uns im Folgenden meist auf den Falls $d = 1$. Wegen dem oberen Ergebnis können die zugehörigen Resultate und Definitionen problemlos auf den Fall $d > 1$ ausgedehnt werden.
Die Analyse der Konvergenz und Grenzwerte von Folgen erfolgt in vielen Fällen durch die Untersuchung einzelner Subterme. Betrachtet wir beispielsweise

$$a_n = \left(1 + \frac{1}{n}\right)^2 = 1 + \frac{2}{n} + \frac{1}{n^2},$$

dann sieht man unmittelbar, dass die drei Summanden rechts jeweils gegen 1, 0 und 0 konvergieren, und man sieht daher ein, dass a_n gegen 1 strebt. Wie wir nun zeigen werden, ist eine solche Argumentation auch allgemeiner Möglich,.
Dazu definieren wir zunächst für eine beliebige Menge $X \subset \mathbb{R}$ (oder $X \subset \mathbb{C}$) und Abbildungen $f : X \to \mathbb{R}$ (oder $\mathbb{C}$), $g : X \to \mathbb{R}$ (oder $\mathbb{C}$) die **punktweise Addition**

$$(f + g)(x) := f(x) + g(x) \qquad (x \in X),$$

und die **punktweise Multiplikation**

$$(f \cdot g)(x) := f(x) \cdot g(x) \qquad (x \in X).$$

Falls g nullstellenfrei ist (also $g(x) \neq 0$ für alle $x \in X$), dann ist die Funktion $1/g : X \to \mathbb{R}$ argumentweise durch

$$(1/g)(x) := 1/g(x) \qquad (x \in X)$$

definiert, und damit auch $f/g := f \cdot (1/g)$ für $f : X \to \mathbb{R}$ (oder $\mathbb{C}$).
Der folgende Satz zeigt, dass Grenzwertbildung mit den algebraischen Operationen in $\mathbb{R}$ oder $\mathbb{C}$ verträglich ist.

Satz 7.9 *Es seien $(a_n)_{n\in N}$, $(b_n)_{n\in N}$ Folgen mit*

$$\lim_{n\to\infty} a_n = c, \ \lim_{n\to\infty} b_n = g$$

Dann gilt

1. $\lim_{n\to\infty} (a_n + b_n) = c + g.$
2. $\lim_{n\to\infty} (a_n \cdot b_n) = c \cdot g.$
3. *Ist $a_n \neq 0$ für alle $n \in N$ und ist $c \neq 0$, so folgt $\lim_{n\to\infty} (b_n/a_n) = g/c.$*

Beweis. Teil (1.) dient als Übung. Wir beweisen nur (2.). (3.) folgt dann in analoger Weise. Für ein beliebiges $\varepsilon > 0$ muss gezeigt werden, dass $|a_n b_n - cg| < \varepsilon$ wenn n genügend groß ist. Da $(a_n)_{n\in N}$ konvergiert, gibt es ein $M > 0$ sodass $|a_n| < M$ für alle n gilt (Übung). Wir nehmen zunächst $g \neq 0$ an. Da $a_n \to c$ und $b_n \to g$ für $n \to \infty$ gilt, gibt es $n_a \in N$ und $n_b \in N$ so dass

$$n \geq n_a \;\Rightarrow\; |a_n - c| < \frac{\varepsilon}{2|g|}$$

und

$$n \geq n_b \;\Rightarrow\; |b_n - g| < \frac{\varepsilon}{2M}.$$

Für $n \geq n_\varepsilon := \max\{n_a, n_b\}$ erhalten wir somit durch Anwendung der Dreiecksungleichung

$$\begin{aligned} |a_n b_n - cg| &= |a_n b_n - a_n g + a_n g - cg| \\ &\leq |a_n|\,|b_n - g| + |g|\,|a_n - c| \\ &\leq M\,|b_n - g| + |g|\,|a_n - c| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Falls $g = 0$ gilt, verschwindet jeweils der rechte Summand in der zweiten und dritten Zeile in der oberen Gleichung, und somit folgt die Behauptung auch für diesen Fall. □

Beispiel 7.10 1. $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n = \frac{1}{n^k},\ (k \in \mathbb{Z})$$

2. $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n = (-1)^n + 2$$

Satz 7.11 *Für eine Folge $(z_n)_{n\in N}$ in $\mathbb{C}$ gilt*

$$\lim_{n\to\infty} z_n = z \in \mathbb{C}$$

genau dann, wenn gilt

$$\lim_{n\to\infty} \mathrm{Re}(z_n) = \mathrm{Re}(z) \text{ und } \lim_{n\to\infty} \mathrm{Im}(z_n) = \mathrm{Im}(z).$$

Satz 7.12 *(Hauptsatz über monotone Folgen) Jede monotone und beschränkte*[a] *Folge* $(a_n)_{n\in\mathbb{N}}$ *in* $\mathbb{R}$ *ist konvergent.*

[a] d.h. $\forall n : |a_n| \leq M$ für eine Schranke $M > 0$

Man kann damit nun Folgende zeigen:

Satz 7.13 (Satz von Bolzano-Weierstraß für Folgen) *Jede beschränkte Folge reeller Zahlen besitzt eine konvergente Teilfolge.*

Definition 7.14 Eine Folge (a_n) reeller Zahlen heißt *Cauchy-Folge*, wenn gilt:

$$\forall \epsilon > 0, \exists N \in \mathbb{N} : \forall m, n \geq N, |a_m - a_n| < \epsilon.$$

Satz 7.15 ***Cauchy-Konvergenzkriterium:*** *Eine Folge* (a_n) *konvergiert in* $\mathbb{R}$ *genau dann, wenn sie eine Cauchy-Folge ist.*

Beweis.
Notwendigkeit: Angenommen, (a_n) konvergiert gegen $c \in \mathbb{R}$. Sei zu einem beliebigen $\varepsilon > 0$ nun $n_\varepsilon \in \mathbb{N}$ so, dass $|a_n - c| < \epsilon/2$ für alle $n \geq n_\varepsilon$. Daraus folgt, dass $|a_m - a_n| \leq |a_m - c| + |c - a_n| < \epsilon$ für alle $m, n \geq n_\varepsilon$. Dies zeigt, dass (a_n) eine Cauchy-Folge ist.
Hinreichend: Jede Cauchy-Folge in $\mathbb{R}$ ist beschränkt und besitzt somit nach Bolzano-Weierstraß, vgl. Satz 7.13, eine konvergente Teilfolge. Sei (a_{n_k}) eine solche Teilfolge, und gelte $\lim_{k\to\infty} a_{n_k} = c$. Es sei weiter $\varepsilon > 0$ beliebig, und dazu $n_{\varepsilon/2} \in \mathbb{N}$ so groß, dass

$$n, n_k \geq_{\varepsilon/2} : \Rightarrow \ |a_n - a_{n_k}| < \frac{\varepsilon}{2} \ \wedge \ |a_{n_k} - c| < \frac{\varepsilon}{2}.$$

Damit gilt dann sofort für alle $n \geq n_{\varepsilon/2}$ (mit beliebigem $n_k \geq n_{\varepsilon/2}$):

$$|a_n - c| \leq |a_n - a_{n_k}| + |a_{n_k} - c| < \varepsilon.$$

□

Funktionsgrenzwerte und Stetigkeit

Mit Hilfe des Grenzwerts für Folgen lässt sich der Grenzwert für allgemeine Abbildung definieren.

Definition 7.16 Es sei $X \subset \mathbb{R}$ (oder $\subset \mathbb{C}$) und $f : X \to \mathbb{R}$ (oder $f : X \to \mathbb{C}$) sowie $z \in \mathbb{R}$ (oder $\in \mathbb{C}$).

1. Existiert ein $c \in \mathbb{R}$ (oder $\in \mathbb{C}$) so, dass für **jede** Folge (x_n) in X mit $\lim\limits_{n\to\infty} x_n = z$ und $x_n \neq z$, gilt

$$\lim_{n\to\infty} f(x_n) = c,$$

   so heißt c **Grenzwert** von f an der Stelle z und wir schreiben

$$\lim_{x\to z} f(x) = c.$$

2. Gilt in (1) außerdem $z \in X$ und $f(z) = c$, dann heißt f **stetig** an der Stelle z.

Bemerkung 7.17

1. In Definition 7.16 (1) ist die Existenz einer Folge (x_n) mit $x_n \in X$ für alle Folgenglieder vorausgesetzt. Existiert eine solche Folge nicht, so kann es an der fraglichen Stelle auch keinen Grenzwert geben.

2. Falls in Definition 7.16 (2) keine Folge (x_n) mit $x_n \in X$ und $\lim\limits_{n\to\infty} x_n = z$ existiert, dann heißt f dennoch stetig an z. In diesem Fall ist z ein sogenannter **isolierter Punkt** von X.

3. In Definition 7.16 heißt f **stetig auf der Menge** $M \subset X$, falls f stetig an jeder Stelle $a \in M$ ist. Ist $M = X$, so heißt f kurz **stetig**.

Satz 7.18 *Es sei $X \subset \mathbb{R}$ (oder $\subset \mathbb{C}$) und $f : X \to \mathbb{R}$. Dann ist f stetig an $x \in X$ genau dann, wenn gilt*

$$\forall \varepsilon > 0\, \exists \delta > 0 : y \in X, |y - x| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

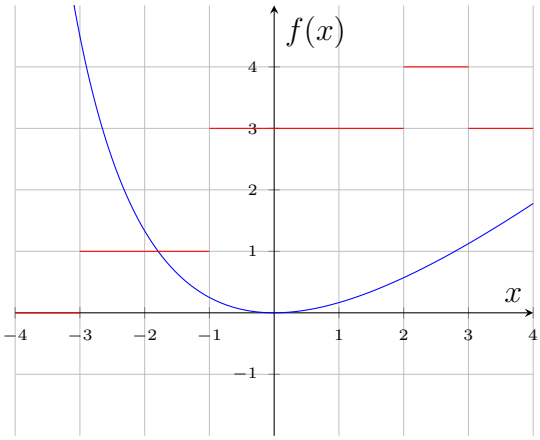

Abbildung 8: Graphen der Funktion $f(x) = \frac{x^2}{x+5}$ (stetig) und einer Treppenfunktion (unstetig)

Beweis. Es gelte

$$\forall \varepsilon > 0 \, \exists \delta > 0 : \, y \in X, \, |y - x| < \delta \; \Rightarrow \; |f(x) - f(y)| < \varepsilon.$$

Um zu zeigen, dass dann (2) in Definition 7.16 gilt, sei (x_n) eine beliebige Folge in X, die gegen x konvergiert und so, dass $x \neq x_n$ für alle n gilt. Sei weiter $\varepsilon > 0$ beliebig. Dazu wählen wir ein zugehöriges (gemäß der obigen Eigenschaft) $\delta > 0$. Ist nun $m \in \mathbb{N}$ so groß, dass $|x_n - x| < \delta$ für alle $n \geq m$ gilt, dann folgt $|f(x_n) - f(x)| < \varepsilon$. Da $\varepsilon > 0$ beliebig war gilt also

$$\lim_{y \to x} f(y) = f(x).$$

Die umgekehrt Implikation dient als Übung. □

Das folgende Beispiel verdeutlicht den Unterschied zwischen der Existenz eines Funktionsgrenzwert und der Stetigkeit.

Beispiel 7.19 1. Sei $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x & \text{falls } x \neq 2 \\ 1 & \text{falls } x = 2 \end{cases}.$$

Dann gilt $\lim_{x \to 2} f(x) = 2$. Aber f ist an der Stelle 2 nicht stetig (unstetig) da $\lim_{x \to 2} f(x) \neq f(2) = 1$.

2. Sei $f : \mathbb{N} \to \mathbb{R}$ (also eine Folge). Dann kann an keinem Elementen des Definitionsbereichs eine Grenzwertuntersuchung erfolgen (warum ist das so?), der Grenzwert ist hier also an keiner Stelle definiert. Dennoch gilt eine solche Funktion als stetig an allen $n \in \mathbb{N}$, da jedes Element ein isolierter Punkt des Definitionsbereichs ist.

3. Sei $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ mit

$$f((x,y)) = \frac{x^2 \cdot y^2}{x^2 + y^2}.$$[2]

Dann gilt $\lim\limits_{(x,y)\to(0,0)} f((x,y)) = 0$. Es kann aber an $(0,0)$ nicht von Stetigkeit der Funktion f gesprochen werden, da diese Stelle nicht zum Definitionsbereich gehört. Mann könnte jedoch f stetig zu einer Funktion $\tilde{f} : \mathbb{R}^2 \to \mathbb{R}$ fortsetzen, in dem man $f((x,y)) =: \tilde{f}((x,y))$ für $(x,y) \neq (0,0)$ und $0 =: \tilde{f}((0,0))$ setzt.

Aus Satz 7.9 und Definition 7.16) ergibt sich:

Satz 7.20 *Es sei $X \subset \mathbb{R}$ (oder $\subset \mathbb{C}$) und $f, h : X \to \mathbb{R}$ (oder $f : X \to \mathbb{C}$) sowie $z \in \mathbb{R}$ (oder $\mathbb{C}$) mit $\lim\limits_{x\to z} f(x) = c$ und $\lim\limits_{x\to z} h(x) = g$. Dann gilt:*

1. $\lim\limits_{x\to z}(f + h) = c + g$.
2. $\lim\limits_{x\to z}(f \cdot h) = c \cdot g$.
3. *Ist f nullstellenfrei und ist $c \neq 0$, so folgt $\lim\limits_{x\to z}(h/f) = g/c$.*

Ist weiter $z \in X$ und f, h beide stetig an z, so folgt in den drei Fällen jeweils die Stetigkeit von $f + h$, $f \cdot h$ und h/f an z.

Rekursive Folgen

In der Mathematik bezeichnet eine **rekursive Folge** eine Folge, bei der jedes Element durch eine oder mehrere vorherige Elemente definiert ist. Diese Definitionen können explizit oder implizit sein. Eine rekursive Folge kann durch eine oder mehrere Startwerte und eine rekursive Beziehung definiert werden. Formaler definieren wir

[2] Bei solchen Funktionen wird meist auf die (formal richtige) doppelte Klammerung $f((x,y))$ verzichtet und man schreibt kurz $f(x,y)$

Definition 7.21 Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt **rekursiv**, falls es eine Abbildung $\varphi : \mathbb{R}^k \to \mathbb{R}$ (oder $\varphi : \mathbb{C}^k \to \mathbb{C}$) und festgesetzte $\xi_1, ..., \xi_k \in \mathbb{R}$ (oder $\in \mathbb{C}$) (**Rekursionsanfang** genannt) so gibt, dass

$$a_n = \begin{cases} \varphi(a_{n-1}, a_{n-2}, ..., a_{n-k}), & \text{falls } n > k \\ \xi_n, & \text{falls } n \leq k \end{cases}.$$

Für allgemeine Indexmengen $N \subset \mathbb{N}_0$ werden rekursive Folgen analog definiert.

Beispiel 7.22 Ein Beispiel für eine rekursive Folge ist die **Fibonacci-Folge**, die wie folgt definiert ist:

$$\begin{aligned} F_0 &= 0 \\ F_1 &= 1 \\ F_n &= F_{n-1} + F_{n-2} \quad \text{für } n \geq 2 \end{aligned}$$

Hier ist F_n das n-te Fibonacci-Zahl für $n \in \mathbb{N}$.
Die ersten paar Fibonacci-Zahlen lauten:

$$\begin{aligned} F_0 &= 0, \\ F_1 &= 1, \\ F_2 &= F_1 + F_0 = 1 + 0 = 1, \\ F_3 &= F_2 + F_1 = 1 + 1 = 2, \\ F_4 &= F_3 + F_2 = 2 + 1 = 3, \\ F_5 &= F_4 + F_3 = 3 + 2 = 5, \\ &\vdots \end{aligned}$$

Die Fibonacci-Folge ist eine der bekanntesten rekursiven Folgen.

Beispiel 7.23 1.

$$a_n = \begin{cases} (1 + a_{n-1})/a_{n-1}, & \text{falls } n > 1 \\ 1, & \text{falls } n = 1 \end{cases}.$$

Hier sind $a_1 = 1$, $a_2 = 2$, $a_3 = \frac{3}{2}$, $a_4 = \frac{5}{3}$, ...

2.
$$a_n = \begin{cases} (1 + a_{n-1})/a_{n-2}, & \text{falls } \ n > 2 \\ n, & \text{falls } n \leq 2 \end{cases}.$$

Hier sind $a_1 = 1$, $a_2 = 2$, $a_3 = 3$, $a_4 = \frac{4}{3}, ...$

3.
$$a_n = \begin{cases} \frac{a_{n-1}}{2} + 1, & \text{falls } \ n > 2 \\ \xi, & \text{falls } n = 2 \end{cases}.$$

Für konvergente rekursive Folgen ist folgendes Ergebnis zur Bestimmung des konkreten Grenzwerts sehr nützlich.

Satz 7.24 *Es sei $X \subset \mathbb{R}^k$ (oder $X \subset \mathbb{C}$) und (a_n) eine durch eine stetige Abbildung $\varphi : X \to \mathbb{R}$ (oder $\varphi : X \to \mathbb{C}$) rekursive definiert Folge*

$$a_n = \begin{cases} \varphi(a_{n-1}, ..., a_{n-k}), & \textit{falls } \ n > k \\ \xi_k, & \textit{falls } n \leq k \end{cases},$$

und es gelte $\lim\limits_{n\to\infty} a_n = c$. Sei weiter

$$C = \underbrace{(c, ..., c)}_{k-\text{mal}},$$

und gelte $C \in X$, dann folgt

$$\varphi(C) = c.$$

Insbesondere gilt für $k = 1$ die Fixpunktgleichung:

$$\varphi(c) = c.$$

Beweis. Laut Annahme gilt $(a_{n-1}, ..., a_{n-k}) \to C$ für $n \to \infty$. Weiter folgt wegen der Stetigkeit von φ

$$\varphi((a_{n-1}, ..., a_{n-k})) \to \varphi(C) \ (n \to \infty).$$

Weil (a_n) rekursiv definiert ist und gegen c konvergiert, gilt weiter auch

$$\varphi((a_{n-1}, ..., a_{n-k})) = a_n \to c, \ (n \to \infty).$$

Wegen der Eindeutigkeit des Folgengrenzwerts ergibt sich die Behauptung. □

Beispiel 7.25 Ein wichtige Familie von Folgen sind die **geometrischen Folgen** $(a_n) = (q^n)$ für $q \in \mathbb{R}$, die sich rekursiv durch $\varphi : \mathbb{R} \to \mathbb{R}$ mit $\varphi(x) = qx$ und $a_0 = 1$ ergeben. Es gilt

1. Für $|q| > 1$ ist (q^n) unbeschränkt.
2. Für $|q| < 1$ ist (q^n) eine Nullfolge, also $q^n \to 0$ $(n \to \infty)$.

(Denn: 1. Hier ist $|q| = 1 + \delta$ für ein $\delta > 0$. Mit der Bernoullischen Ungleichung gilt

$$|q^n| = |q|^n = (1+\delta)^n \geq 1 + n\delta > n\delta \qquad (n \in \mathbb{N}).$$

Also ist (q^n) unbeschränkt.
2. Für $q = 0$ ist die Behauptung klar. Es sei also $0 < |q| < 1$. Dann ist $1/|q| = 1 + \delta$ mit einem $\delta > 0$ und wie in 1.

$$|q^n| < \frac{1}{n\delta} \qquad (n \in \mathbb{N}).$$

Aus $1/n \to 0$ folgt $q^n \to 0$ $(n \to \infty)$.)

Beispiel 7.26 (*Heron-Verfahren; Babylonisches Wurzelziehen*)
Es sei $c > 0$ gegeben und $\varphi : (0, \infty) \to (0, \infty)$ definiert durch

$$\varphi(x) := \frac{1}{2}\left(x + \frac{c}{x}\right) \quad (x > 0)\ .$$

Dann gilt $\sqrt{c} = \sqrt{x \cdot c/x} \leq \varphi(x)$ (geometrisches Mittel ≤ arithmetisches Mittel). Wir betrachten mit einem beliebigen Startwert $a_0 > 0$ die Folge (a_n) in $(0, \infty)$ mit

$$a_{n+1} := \varphi(a_n) = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) \qquad (n \in \mathbb{N}_0)\ .$$

Mithilfe des Hauptsatzes über monotone Folgen kann man zeigen, dass (a_n) konvergent ist mit

$$\lim_{n\to\infty} a_n = \sqrt{c},$$

d. h. die Folgeglieder a_n sind Approximationen (also Näherungen) für $\sqrt{c}$. Dabei sind im Falle $c \in \mathbb{Q}$ und $a_0 \in \mathbb{Q}$ die a_n stets rationale Zahlen.
Wie sieht es dabei mit dem Fehler aus, wenn man a_n statt $\sqrt{c}$ verwendet? Wir schätzen den Fehler nach oben ab. Dazu sei

$$\varepsilon_n = \frac{a_n - \sqrt{c}}{\sqrt{c}} = \frac{a_n}{\sqrt{c}} - 1 \geq 0 \quad (n \in \mathbb{N})$$

der relative Fehler. Dann gilt

$$1 + \varepsilon_{n+1} = \frac{1}{\sqrt{c}} a_{n+1} = \frac{1}{2}\left(\frac{a_n}{\sqrt{c}} + \frac{\sqrt{c}}{a_n}\right) = \frac{1}{2}\left(1 + \varepsilon_n + \frac{1}{1+\varepsilon_n}\right),$$

also

$$\varepsilon_{n+1} = \frac{1}{2}\frac{\varepsilon_n^2}{1+\varepsilon_n} \leq \frac{1}{2}\varepsilon_n^2 .$$

Hat man nach n Schritten für a_n einen Fehler $\varepsilon_n \leq 10^{-m}$, so ist der Fehler ε_{n+1} im nächsten Schritt $\leq \frac{1}{2}(10^{-m})^2 = \frac{1}{2}10^{-2m}$; die Anzahl der exakt.

Reihen

Definition 7.27 Es sei $(a_n)_{n\geq m}$, $m \in \mathbb{N}_0$ eine Folge. Die Folge $(s_n)_{n\geq m}$ der **Partial**- oder **Teilsummen**

$$s_n := \sum_{\nu=m}^{n} a_\nu =: a_m + \cdots + a_n \qquad (n \geq m)$$

heißt die (**mit** (a_n) **gebildete**) **Reihe**. Die a_ν heißen dann **Reihenglieder**.
Ist die Folge (s_n) konvergent, so heißt $\lim\limits_{n\to\infty} s_n$ der **Reihenwert** und man schreibt

$$\sum_{\nu=m}^{\infty} a_\nu := \lim_{n\to\infty} s_n.$$

Die Reihe heißt **absolut konvergent**, wenn die Reihe

$$\sum_{\nu=m}^{\infty} |a_\nu|$$

konvergiert.

Bemerkung 7.28 Traditionell wird neben dem Reihenwert auch die Teilsummenfolge (s_n) mit $\sum_{\nu=m}^{\infty} a_\nu$ bezeichnet. Das ist ganz praktisch, weil man dann kurz von Konvergenz oder Divergenz von $\sum_{\nu=m}^{\infty} a_\nu$ sprechen kann. Man beachte aber, dass das Symbol $\sum_{\nu=m}^{\infty} a_\nu$ damit zwei Bedeutungen hat: Erstens steht es für die Folge (s_n) der Teilsummen und zweitens (im Falle der Konvergenz!) für deren Grenzwert.

Beispiel 7.29 (**geometrische Reihe**) Es sei $a_n = q^n$ für ein $q \in \mathbb{R}$, $|q| < 1$. Dann ist $\sum_{\nu=0}^{\infty} q^\nu$ konvergent mit

$$\sum_{\nu=0}^{\infty} q^\nu = \lim_{n\to\infty} \sum_{\nu=0}^{n} q^\nu = \lim_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q} \,.$$

Für $q = 1/2$ ergibt sich etwa $\sum_{\nu=0}^{\infty} 1/2^\nu = 2$.

Bemerkung 7.30 Durch Anwendung von Satz 7.9 ergibt sich leicht: Sind $\sum_{\nu=m}^{\infty} a_\nu$ und $\sum_{\nu=m}^{\infty} b_\nu$ konvergente Reihen und ist $\lambda \in \mathbb{R}$, so sind auch

$$\sum_{\nu=m}^{\infty} (a_\nu + b_\nu) \text{ und } \sum_{\nu=m}^{\infty} \lambda a_\nu$$

konvergent mit

$$\sum_{\nu=m}^{\infty} (a_\nu + b_\nu) = \sum_{\nu=m}^{\infty} a_\nu + \sum_{\nu=m}^{\infty} b_\nu \quad \text{und} \quad \sum_{\nu=m}^{\infty} \lambda a_\nu = \lambda \sum_{\nu=m}^{\infty} a_\nu.$$

Beispiel 7.31 Mit Bemerkung 7.30 und Beispiel 7.29 ist

$$\sum_{\nu=0}^{\infty} \frac{2\cdot 3^\nu + 4}{5^\nu} = 2\cdot \sum_{\nu=0}^{\infty} \frac{3^\nu}{5^\nu} + 4\cdot \sum_{\nu=0}^{\infty} \frac{1}{5^\nu} = 2\cdot \frac{1}{1-3/5} + 4\cdot \frac{1}{1-1/5} = 10 \,.$$

Insbesondere erhält man aus Bemerkung 7.30 auch: Ist $k > m$, so ist $\sum_{\nu=k}^{\infty} a_\nu$ genau dann konvergent, wenn $\sum_{\nu=m}^{\infty} a_\nu$ konvergiert, und in diesem Fall ist

$$\sum_{\nu=m}^{\infty} a_\nu = \sum_{\nu=m}^{k-1} a_\nu + \sum_{\nu=k}^{\infty} a_\nu \,.$$

Für Konvergenzuntersuchungen ist es also unwichtig, wie die untere Summationsgrenze aussieht.

Satz 7.32 *Ist* $\sum_{\nu=m}^{\infty} a_\nu$ *konvergent, so gilt*

$$a_n \to 0 \qquad (n \to \infty).$$

Für absolut konvergente Reihen lässt sich folgendes zeigen

Satz 7.33 *(Cauchy-Produktformel) Es seien* $\sum_{k=0}^{\infty} a_n$, $\sum_{k=0}^{\infty} b_n$ *absolut konvergent. Dann konvergiert das Produkt der Reihen (Produkt der Partialsummen) und der Wert (Grenzwert) der Reihe ergibt sich durch das Cauchy-Produkt:*

$$\sum_{n=0}^{\infty} \sum_{\nu=0}^{n} a_\nu \, b_{n-\nu}.$$

Satz 7.34 *(Majorantenkriterium)*
Es seien $(a_n)_{n\geq m}$ *und* $(b_n)_{n\geq k}$ *Folgen mit*

$$0 \leq |a_n| \leq b_n$$

für n genügend groß. Man nennt dann $\sum_{\nu=k}^{\infty} b_\nu$ *eine konvergente* ***Majorante*** *(von* $\sum_{\nu=m}^{\infty} a_\nu$*). Ist* $\sum_{\nu=k}^{\infty} b_\nu$ *konvergent, so ist auch* $\sum_{\nu=m}^{\infty} a_\nu$ *konvergent (sogar absolut konvergent). Durch Kontraposition ergibt sich in der obigen Situation: Ist die Reihe über* $(a_n)_{n\geq m}$ *divergent, dann ist auch die Reihe über* $(b_n)_{n\geq k}$ *divergent. Man nennt dann* $\sum_{\nu=m}^{\infty} a_\nu$ *eine divergente* ***Minorante***

Beweis. Angenommen, die Reihe $\sum_{\nu=k}^{\infty} b_\nu$ ist konvergent. Dann existiert zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt:

$$\left| \sum_{\nu=N+1}^{n} b_\nu \right| < \epsilon,$$

da die Partialsummen eine Cauchy-Folge bilden. Aufgrund der Ungleichung $0 \leq |a_n| \leq$

b_n folgt, dass für alle $n \geq N$:

$$\left| \sum_{\nu=N+1}^{n} a_\nu \right| \leq \left| \sum_{\nu=N+1}^{n} b_\nu \right| < \epsilon.$$

Das zeigt, dass auch die Partialsummen von $\sum_{\nu=m}^{\infty} a_\nu$ eine Cauchy-Folge bilden, was die Konvergenz von $\sum_{\nu=m}^{\infty} a_\nu$ impliziert. Da $0 \leq |a_n|$ für alle n, folgt daraus die absolute Konvergenz.
Die Kontraposition dieses Beweises zeigt, dass, wenn $\sum_{\nu=m}^{\infty} a_\nu$ divergent ist, dann kann die Reihe $\sum_{\nu=k}^{\infty} b_\nu$ nicht konvergent sein, da andernfalls $\sum_{\nu=m}^{\infty} a_\nu$ nach obigem Beweis auch konvergent sein müsste. □

Beispiel 7.35 (allgemeine harmonische Reihen) Es sei $d \in \mathbb{N}$. Dann gilt

$$\sum_{\nu=1}^{\infty} \frac{1}{\nu^d} \begin{cases} \text{divergent, falls } d = 1 \\ \text{konvergent, falls } d > 1 \end{cases}.$$

Für den Fall $d = 1$ betrachten wir

$$s_n = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 1/2} + \underbrace{\frac{1}{5} + ... + \frac{1}{8}}_{\geq 1/2} + \underbrace{\frac{1}{9} + ... + \frac{1}{16}}_{\geq 1/2} + ...$$

und folgern damit

$$s_{2^m} \geq 1 + \frac{m}{2} \to \infty \quad (m \to \infty).$$

Somit divergiert die Reihe.
Für $d > 1$ ist

$$\frac{1}{\nu^d} \leq \frac{1}{\nu^2} \leq \frac{1}{(\nu - 1)\nu} =: b_\nu \ .$$

Weiter ist $\sum_{\nu=2}^{n} b_\nu = 1 - 1/n$ (Übung), also $\sum_{\nu=2}^{\infty} b_\nu = 1$. Damit ist $\sum_{\nu=2}^{\infty} b_\nu$ eine konvergente Majorante. Mit Satz 7.34 folgt die Behauptung.

Beispiel 7.36 (Münzstapel) Betrachten wir das Problem, Münzen so zu stapeln, dass jede Münze über die Kante der darunterliegenden Münze hinausragt, ohne dass der Stapel umfällt. Überraschenderweise lässt sich zeigen, dass es möglich ist, einen beliebig

großen Überhang zu erzielen, indem die Münzen in einer spezifischen Weise angeordnet werden.
Sei d der Durchmesser einer Münze und U_n der maximale Überhang, der mit n Münzen erreicht werden kann. Es zeigt sich, dass der Überhang proportional zur Summe der inversen natürlichen Zahlen ist, also zur harmonischen Reihe:

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

Dies bedeutet, dass der Überhang mit jeder hinzugefügten Münze wächst, jedoch der Zuwachs mit zunehmender Anzahl von Münzen immer kleiner wird. Theoretisch kann so ein unbegrenzter Überhang erreicht werden.

Satz 7.37 *(Wurzelkriterium, Quotientenkriterium)*
Es sei $\sum_{\nu=k}^{\infty} a_\nu$ *eine Reihe.*

1) *(Wurzelkriterium) Gibt es ein* $0 \leq q < 1$ *und ein* $m \in \mathbb{N}$ *so, dass*

$$|a_\nu|^{\frac{1}{\nu}} \leq q, \text{ für alle } \nu \geq m > 0,$$

dann konvergiert $\sum_{\nu=k}^{\infty} a_\nu$ *(ist sogar absolut konvergent). Gilt umgekehrt* $q > 1$ *und*

$$|a_\nu|^{\frac{1}{\nu}} \geq q, \text{ für alle } \nu \geq m > 0,$$

so ist die Reihe divergent.

2) *(Quotientenkriterium) Gibt es ein* $0 \leq q < 1$ *und ein* $m \in \mathbb{N}$ *so, dass*

$$\left| \frac{a_{\nu+1}}{a_\nu} \right| \leq q, \text{ für alle } \nu \geq m, \text{ falls } a_\nu \neq 0,$$

dann konvergiert $\sum_{\nu=k}^{\infty} a_\nu$ *(ist sogar absolut konvergent). Gilt umgekehrt* $q > 1$ *und*

$$\left| \frac{a_{\nu+1}}{a_\nu} \right| \geq q, , \text{ für alle } \nu \geq m > 0,$$

so ist die Reihe divergent.

Beweis. Wurzelkriterium:
Angenommen, es gibt ein $0 \leq q < 1$ und ein $m \in \mathbb{N}$, so dass $|a_\nu|^{\frac{1}{\nu}} \leq q$ für alle $\nu \geq m > 0$. Wir wollen zeigen, dass $\sum_{\nu=k}^{\infty} a_\nu$ konvergiert.
Da $|a_\nu|^{\frac{1}{\nu}} \leq q < 1$ für alle $\nu \geq m > 0$, folgt $|a_\nu| \leq q^\nu < 1$ für alle $\nu \geq N$. Weil die Reihe

$$\sum_{\nu=m}^{\infty} q^\nu$$

konvergiert, ist eine konvergente Majorante gefunden und die Behauptung folgt mit atz 7.34.
Der umgekehrte Fall lässt sich analog beweisen, indem man zeigt, dass $\sum_{\nu=k}^{\infty} a_\nu$ divergiert, falls $|a_\nu|^{\frac{1}{\nu}} \geq q > 1$ für alle ν gilt.
Quotientenkriterium: Übung.

□

8 Exponentialfunktion, Logarithmus

Satz 8.1 *Für jedes $z \in \mathbb{C}$ ist die Reihe $\sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!}$ absolut konvergent. Dies Folgt beispielsweise sofort mit dem Quotientenkriterium.*

Damit können wir nun die Exponentialfunktion definieren:

Definition 8.2 Die Abbildung $\exp : \mathbb{C} \to \mathbb{C}$, gegeben durch

$$\exp(z) := \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \qquad (z \in \mathbb{C}) \ ,$$

heißt **Exponentialfunktion**. Nach Definition ist $\exp(0) = 1$.

Satz 8.3 *Es gilt:*

1. exp *ist stetig,*
2. exp *ist streng monoton wachsend und damit injektiv als Funktion von $\mathbb{R}$ nach $\mathbb{R}$,*
3. $\exp(\mathbb{R}) = (0, \infty)$.

Wir wollen nun Eigenschaften der Exponentialfunktion einführen, die von fundamentaler Bedeutung für die Mathematik sind.

Satz 8.4 *Für $z, y \in \mathbb{C}$ und $x \in \mathbb{R}$ gilt*

1. $\exp(z + y) = \exp(z) \cdot \exp(y)$,
2. $\exp(-z) = 1/\exp(z)$,
3. $|\exp(ix)| = 1$.

Beweis.

1. Hier folgt mit der Cauchy-Produktformel (Satz 7.33),

$$\exp(z)\,\exp(y) = \sum_{\nu=0}^{\infty} \frac{z^\nu}{\nu!} \cdot \sum_{\nu=0}^{\infty} \frac{y^\nu}{\nu!} = \sum_{n=0}^{\infty} \sum_{\nu=0}^{n} \frac{z^\nu}{\nu!} \frac{y^{n-\nu}}{(n-\nu)!}.$$

Erweitert man nun den innersten Term mit $n!$, klammert $1/n!$ aus der inneren Summe aus, und wendet anschließend den binomischen Satz (5.8) an, so folgt:

$$\exp(z)\,\exp(y) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\nu=0}^{n} \binom{n}{\nu} z^\nu\, y^{n-y} = \sum_{n=0}^{\infty} \frac{(z+y)^n}{n!} = \exp(z+y).$$

2. ergibt sich wegen der Eindeutigkeit der Inversen, da nach 1. gilt:

$$\begin{aligned} \exp(-z) &= \frac{1}{\exp(z)} \\ \Leftrightarrow \quad \exp(-z) \cdot \exp(z) &= 1 \\ \Leftrightarrow \quad \exp(-z)\,\exp(z) = \exp(0) &= 1. \end{aligned}$$

3. folgt mit 6.7.3 (sowie 6.6 verbunden mit 8.2) und 1.

$$|\exp(ix)|^2 = \exp(ix)\,\overline{\exp(ix)} = \exp(ix)\,\exp(-ix) = \exp(ix - ix) = 1 = 1^2$$

$$\Leftrightarrow |\exp(ix)| = 1$$

□

Bemerkung 8.5 Aus Satz 8.4 folgt

$$\exp(mz) = (\exp(z))^m \qquad (z \in \mathbb{C},\, m \in \mathbb{Z}).$$

Die Zahl $e := \exp(1)$ heißt **Eulersche Zahl**. Es gilt damit $\exp(m) = e^m$ für alle $m \in \mathbb{Z}$ und deshalb ist die Schreibweise e^z statt $\exp(z)$ für allgemeines $z \in \mathbb{C}$ konsistent mit der Definition ganzzahliger Potenzen. Wir werden diese im Weiteren meist verwenden.

Satz 8.6 *Es gilt*

$$\lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n = e^z, \quad (z \in \mathbb{C})$$

Satz 8.4 (3) zeigt, dass die Exponentialfunktion für Werte der imaginären Achse Werte auf dem Einheitskreis $\{z \in \mathbb{C} : |z| = 1\}$ der komplexen Ebene annimmt (im Gegensatz zum Verhalten auf der reellen Achse). Der folgend Satz beschreibt das Verhalten der Exponentialfunktion für Argumente der imaginären Achse, also $z = ix$ mit $x \in \mathbb{R}$.

Satz 8.7 *Es gelten:*

1. $e^{i\pi} = -1$ *(eulersche Identität)*
2. $e^{ix} = e^{ix+2\pi ik}$, *für alle* $k \in \mathbb{Z}$ *und alle* $x \in \mathbb{R}$ *(*$2\pi-$*Periodizität).*

Zusammen mit der Stetigkeit zeigt Satz 8.7, dass $\exp(ix)$ für $x \in (-\pi, \pi]$ den gesamten Einheitskreis $\{z \in \mathbb{C} : |z| = 1\}$ der komplexen Ebene abläuft.

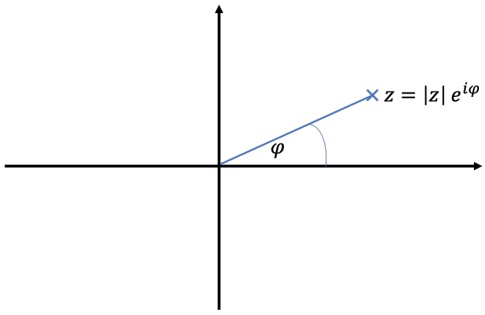

Abbildung 9: Polarkoordinaten in der komplexen Ebene

Bemerkung und Definition 8.8 Für jedes $z \in \mathbb{C}$, $z \neq 0$, gibt es ein eindeutiges $\varphi \in (-\pi, \pi]$ so, dass

$$z = |z|e^{i\varphi}.$$

Diese Darstellung heißt **Polarkoordinaten-Darstellung** von z und φ heißt **Argument** von z. Wegen Satz 8.7 (2) gilt dann weiter

$$z = |z|e^{i\varphi+2\pi ik}$$

für alle $k \in \mathbb{Z}$.

Beispiel 8.9

$$\begin{aligned} 1 &= e^{i0} \\ i &= e^{i\pi/2} \\ -i &= e^{-i\pi/2} \\ -1 &= e^{i\pi} \\ z &= 1+i = |z|\, e^{i\pi/4} = \sqrt{2}\, e^{i\pi/4} \end{aligned}$$

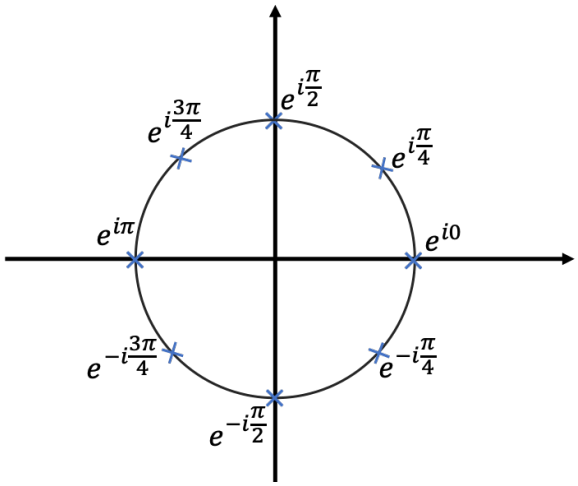

Abbildung 10: Polarkoordinaten auf dem Einheitskreis der komplexen Ebene

Das Verhalten von $e^{i\varphi}$ auf dem Einheitskreis der komplexen Ebene liefert einen schönen analytischen Zugang zur Definition trigonometrischer Funktionen.

Bemerkung und Definition 8.10 Die **Sinusfunktion** $\sin : \mathbb{C} \to \mathbb{C}$ und **Kosinusfunktion** $\cos : \mathbb{C} \to \mathbb{C}$ sind definiert durch

$$\sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz})$$

und

$$\cos(z) = \frac{1}{2}(e^{iz} + e^{-iz}).$$

Damit gilt für $x \in \mathbb{R}$ gilt damit:

$$\sin(x) = \mathrm{Im}(e^{ix})$$

und

$$\cos(x) = \mathrm{Re}(e^{ix})$$

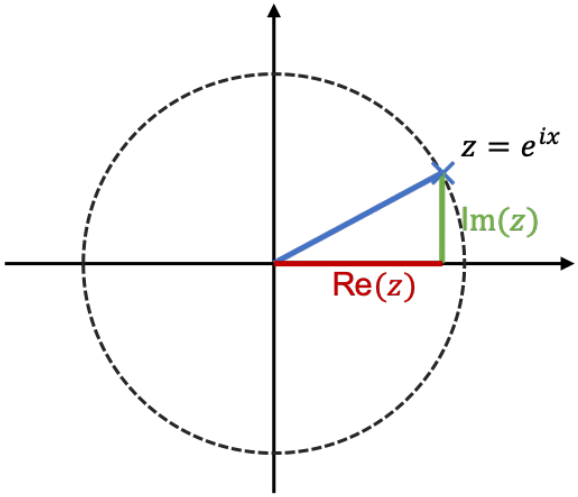

Abbildung 11: Einheitsdreieck in der komplexen Ebene

Wichtige Eigenschaften von Sinus und Kosinus als Funktionen auf $\mathbb{R}$ sind im folgenden Ergebnis zusammen gefasst:

Satz 8.11 *Für Sinus und Kosinus gilt das Folgende:*

1. $|\sin(x)| \leq 1$ *und* $|\cos(x)| \leq 1$ *für alle* $x \in \mathbb{R}$,
2. $\sin(x) = \sin(x + 2\pi k)$ *und* $\cos(x) = \cos(x + 2\pi k)$, *für alle* $k \in \mathbb{Z}$ *und alle* $x \in \mathbb{R}$ *($2\pi-$Periodizität),*
3. $\sin(x) = -\sin(-x)$ *und* $\cos(x) = \cos(-x)$, *für und alle* $x \in \mathbb{R}$
4. $\sin(x + \pi/2) = \cos(x)$ *für und alle* $x \in \mathbb{R}$,
5. $\sin(k\pi) = 0$ *und* $\cos(k\pi + \pi/2) = 0$ *für und alle* $k \in \mathbb{Z}$.

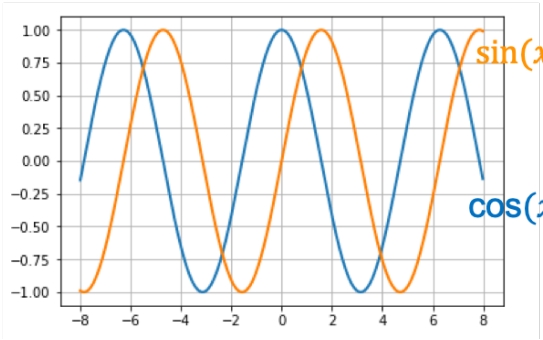

Abbildung 12: Sinus und Kosinus als Funktion von $\mathbb{R}$ nach $[-1, 1]$

Bemerkung 8.12 Genau wie bei der Definition der Exponentialfunktion (8.2) lassen sich die Sinus und Kosinusfunktion als Grenzwert einer Reihe darstellen:

$$\cos(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k}$$

und

$$\sin(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1}.$$

für all $z \in \mathbb{C}$.

Satz und Definition 8.13 Zu exp existiert die Umkehrfunktion auf $\mathbb{R}$ und diese ist dort stetig und streng wachsend. Diese Funktion nennnen wir (natürliche) **Logarithmusfunktion** und schreiben dafür ln oder auch log. Aus den entsprechenden Eigenschaften der Exponentialfunktion ergibt sich leicht:

1. Für alle $s, t > 0$ ist $\ln(st) = \ln(s) + \ln(t)$.
2. Für alle $t > 0$ und alle $m \in \mathbb{Z}$ ist $\ln(t^m) = m \ln(t)$.

Für $a > 0$ und $m \in \mathbb{Z}$ ist

$$a^m = e^{\ln(a^m)} = e^{m \ln a}$$

nach S.u.D.8.13.2.

Definition 8.14 Für $a > 0$ und $x \in \mathbb{R}$ setzen wir:

$$a^x := \exp(x \cdot \ln a) = e^{x \cdot \ln a} .$$

Aus den Rechenregeln für ln und exp erhält man $a^{1/n} = \sqrt[n]{a}$ für $a > 0$ und die folgenden Potenzgesetze.

Satz 8.15 *Es seien $a, b > 0$ und $x, y, c \in \mathbb{R}$. Dann gilt*

$$a^x a^y = a^{x+y}, \quad a^x b^x = (ab)^x \quad \textit{und} \quad (a^c)^x = a^{cx}.$$

Beweis. Es gilt

$$a^z a^w = e^{z \ln a} e^{w \ln a} = e^{z \ln a + w \ln a} = e^{(z+w) \ln a} = a^{z+w}$$

und

$$a^z b^z = e^{z \ln a} e^{z \ln b} = e^{z(\ln a + \ln b)} = e^{z \ln(ab)} = (ab)^z.$$

Für $c \in \mathbb{R}$ ist $a^c = e^{c \ln a} > 0$ und damit

$$(a^c)^z = e^{z \ln(e^{c \ln a})} = e^{cz \ln a} = a^{cz} .$$

□

Satz und Definition 8.16 Es sei $a > 0$, dann definiert $f(x) = a^x$ eine bijektive Funktion $f : \mathbb{R} \to (0, \infty)$. Dabei gilt für $f^{-1} : (0, \infty) \to \mathbb{R}$,

$$f^{-1}(x) = \frac{\ln(x)}{\ln(a)}.$$

Die Funktion f^{-1} heißt **Logarithmus zur Basis** a, und wir schreiben $\log_a(x) := f^{-1}(x)$. Wie in (8.13) gelten:

1. Für alle $s, t > 0$ ist $\log_a(st) = \log_a(s) + \log_a(t)$.
2. Für alle $t > 0$ und alle $m \in \mathbb{Z}$ ist $\log_a(t^m) = m \log_a(t)$.

A Intervall-Notation

Für $a \leq b$, $a, b \in \mathbb{R}$:

- $[a, b] := \{x \in \mathbb{R} : a \leq x \text{ und } x \leq b\}$ (abgeschlossenes Intervall)
- $(a, b) := \{x \in \mathbb{R} : a < x \text{ und } x < b\}$ (offenes Intervall)
- $[a, b) := \{x \in \mathbb{R} : a \leq x \text{ und } x < b\}$ (rechts halboffenes Intervall)
- $(a, b] := \{x \in \mathbb{R} : a < x \text{ und } x \leq b\}$ (links halboffenes Intervall)

Für unbegrenzte Intervalle schreiben wir auch

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

B Quantoren

Dieser Abschnitt dient der informellen Einführung der **Quantoren**.
In dieser Vorlesung werden die folgenden **Quantoren** zur vereinfachten Formulierungen verwenden

- $\forall$ Bedeutung: „für alle“
- $\exists$ Bedeutung: „es existiert“
- $\neg\exists$ Bedeutung: „es existiert nicht“oder auch „es existiert kein“

Ist nun $A(x)$ eine Aussage, die von einer Variablen x abhängt, so kann mittels Quantoren das Folgende beschrieben werden:

- $\forall x : A(x)$, „für alle x gilt $A(x)$ “.
- $\exists x : A(x)$, „es existiert ein x für das $A(x)$ gilt “.

Die Quantoren werden häufig mit Bedingungen versehen. Typische Ausdrücke mit Quantoren sind beispielsweise die Folgenden:

- $\forall x \in \mathbb{R} \setminus \{0\} : x^2 > 0$ „für alle reellen Zahlen x ungleich 0 gilt $x^2 > 0$“. Man findet auch folgende Schreibweise: $x^2 > 0, \forall x \in \mathbb{R} \setminus \{0\}$.
- $\neg\exists x \in \mathbb{R} : x^2 = -1$ „es gibt kein reelles x, so dass $x^2 = -1$ gilt.“.
- $\forall(a, b \in \mathbb{R},\ a < b)\exists x \in \mathbb{R} : a < x < b$, „für alle reellen Zahlen a, b mit $a < b$ gibt es ein reelles x mit $a < x < b$.“. Die Bedingungen des äußeren Quantors (weiter links stehend) beschreiben den Gültigkeitsbereich der inneren Quantoren.
- $\exists x \in \mathbb{R}\ \forall y \in \mathbb{R} : x \cdot y = 0$, „es gibt ein reelles x, so dass für alle reellen Zahlen y gilt $x \cdot y = 0$“

Eine formale Einführung der Quantoren liefert die sogenannte **Prädikatenlogik**, die eine Erweiterung der Aussagenlogik ist. Wir werden hier die Prädikatenlogik nicht behandeln.

C Erweiterter Euklidischer Algorithmus

Gegeben seien zwei nicht-negative ganze Zahlen a und b, wobei $a \geq b$.
Wir möchten den größten gemeinsamen Teiler (ggT) von a und b sowie $s, t \in \mathbb{Z}$ finden, sodass $as + bt = \mathrm{ggT}(a, b)$ gilt.

1. Initialisiere die Variablen $r_0 = a$, $s_0 = 1$,$t_0 = 0$, $r_1 = b$, $s_1 = 0$,$t_1 = 1$, $k = 1$
2. Wiederhole folgende Schritte, solange $r_k \neq 0$:

$$q_k = \left\lfloor \frac{r_{k-1}}{r_k} \right\rfloor \tag{C.11}$$
$$r_{k+1} = r_{k-1} - q_k r_k \tag{C.12}$$
$$s_{k+1} = s_{k-1} - q_k s_k \tag{C.13}$$
$$t_{k+1} = t_{k-1} - q_k t_k \tag{C.14}$$
$$k \leftarrow k + 1 \tag{C.15}$$

3. Der ggT von a und b ist der Wert von r_{k-1} nach Abschluss der Schleife. Die Werte von s_k und t_k entsprechen s und t wie oben.($\lfloor\rfloor$ bedeutet ganzzahliges abrunden)