

Skriptum zu LGI1 & MAG1 Studienjahr 2014/2015

© Stephan Dreiseitl

Inhaltsverzeichnis

1	Einleitung	3
2	Logik als Sprache der Mathematik	6
2.1	Aussagenlogik	6
2.2	Prädikatenlogik	14
3	Mengen als Bausteine der Mathematik	25
3.1	Mengenbildung: Mengen aus Elementen	26
3.2	Mengenbildung: Mengen aus Mengen	30
3.3	Mengenbildung: Tupel	34
3.4	Prädikate als Mengen	36
3.5	Funktionen als Mengen	45
3.6	Struktur in Mengen	51
3.7	Endliche und unendliche Mengen	54
3.8	Problemspezifikation	57
4	Mengen + Operationen = Algebraische Strukturen	60
4.1	Gruppen	62
4.2	Ringe und Körper	69
4.3	Komplexe Zahlen	72
4.4	Polynome	78
5	Lineare Algebra	83
5.1	Vektorräume	83
5.2	Lineare Abbildungen	89
5.3	Matrizen	93
5.4	Determinanten	101
5.5	Gaußsches Eliminationsverfahren	107
6	Geometrie in \mathbb{R}^n	115
6.1	Länge und Winkel	115
6.2	Projektionen	119
7	Folgen und Reihen	128
7.1	Folgen	128
7.2	Reihen	133
7.3	Potenzreihen	138
7.4	Stetige Funktionen	141

8	Differentialrechnung	146
8.1	Ableitung einer Funktion	146
8.2	Ableitungen und Eigenschaften von Funktionen	151
8.3	Taylorreihen	158
8.4	Die Grenzwertregel von de l'Hôpital	165
9	Integralrechnung	168
9.1	Das Riemann-Integral	168
9.2	Unbestimmte Integrale und Stammfunktionen	170
9.3	Bogenlänge von Kurven	175

Kapitel 1

Einleitung

Die Mathematik dient vielen Bereichen der Naturwissenschaften und Technik als Hilfswissenschaft, mit der die dort auftretenden Probleme in einem strukturierten Formalismus gelöst werden können. Durch diese informelle Beschreibung der Rolle der Mathematik als Hilfswissenschaft werden bereits zwei grundlegende Anforderungen spezifiziert:

- Die Mathematik muss eine präzise Sprache zur Verfügung stellen, mit der konkrete Probleme formal beschrieben werden können.
- Die Mathematik muss ein Kalkül zur Verfügung stellen, mit der aus formalen Problemspezifikationen Lösungen dieser Probleme berechnet werden können. Das Kalkül muss mächtig genug sein, um damit die Richtigkeit der Problemlösungen beweisen zu können.

Wir betrachten ein Beispiel, das die Notwendigkeit dieser Forderungen an einer konkreten Aufgabenstellung erläutert.

Beispiel 1.1 Ein Grundproblem der Bioinformatik ist es, Ähnlichkeiten zweier Nukleotidsequenzen zu bestimmen. Die Motivation dieser Aufgabenstellung ist die Annahme, dass ähnliche Genomsequenzen ähnliche biologische Funktionen haben dürften. Somit kann durch Ähnlichkeitsvergleiche einer noch nicht ausgewerteten Sequenz mit einer Datenbank, in der die Funktion bekannter Sequenzen gespeichert ist, auf die Funktion der unbekannteten Sequenz geschlossen werden.

Die Rolle der Mathematik im Lösungsprozess ist, wie schon oben angedeutet, eine zweifache: Sie muss einerseits eine präzise Sprache zur abstrakten Beschreibung von Nukleotidsequenzen bereitstellen, andererseits eine in der gleichen Sprache gehaltene Problembeschreibung ermöglichen, aus der dann mit formalen Mitteln eine Lösung des Problems ermittelt werden kann. Wir werden in diesem ersten Beispiel noch nicht auf die Problemlösung eingehen, sondern nur die Abbildung einer molekularbiologischen Aufgabenstellung in die Sprache der Mathematik zeigen.

Im konkreten Beispiel ist die Übertragung der Objekte der realen Welt in die Sprache der Mathematik (die *Modellbildung*) nicht schwierig: Ein Teil einer DNA-Doppelhelix kann durch die Sequenz der Nukleotide beschrieben werden, die einen Strang der Doppelhelix ausmachen (da dadurch der andere Strang eindeutig gegeben ist). Da es im Genom nur vier verschiedene Nukleotide gibt (*Adenin*, *Guanin*,

Cytosin und *Thymin*), können diese durch ihre Anfangsbuchstaben repräsentiert werden. Eine Genomsequenz ist somit beispielsweise durch die Zeichenkette **ACGCTT** darstellbar.

In dieser Schreibweise lautet die Aufgabenstellung des Sequenzvergleichs nun, zwei Zeichenketten miteinander zu vergleichen. Intuitiv scheint klar, dass obige Zeichenkette **ACGCTT** ähnlicher zur **ACACTT** als zu **ACGCCA** ist. Wenn man sie zum Vergleich übereinanderschreibt, erhält man

$$\begin{array}{cccccc} \text{A} & \text{C} & \text{G} & \text{C} & \text{T} & \text{T} \\ \text{A} & \text{C} & \text{A} & \text{C} & \text{T} & \text{T} \end{array} \quad \text{und} \quad \begin{array}{cccccc} \text{A} & \text{C} & \text{G} & \text{C} & \text{T} & \text{T} \\ \text{A} & \text{C} & \text{G} & \text{C} & \text{C} & \text{A} \end{array} .$$

Die linken beiden Ketten unterscheiden sich in einem, die rechten beiden in zwei Nukleotiden. Bei einem beliebigen Sequenzvergleich, bei dem Sequenzen unterschiedlicher Länge verglichen werden, ist aber noch überhaupt nicht klar, an welchem Punkt man den Vergleich der beiden Ketten beginnt.

Wir benötigen somit, um das Problem strukturiert betrachten und den Formalismus der Mathematik zur Problemlösung nutzen zu können, eine mathematische Definition der verwendeten Begriffe: Was ist eine Sequenz? Wie vergleicht man zwei Sequenzen? Wann ist eine Übereinstimmung besser als eine andere? Man beachte, dass man durch diese Überlegungen erst das Problem exakt (in der Sprache der Mathematik) spezifizieren kann, dadurch aber einer Lösungsidee noch keinen Schritt näher gekommen ist.

Das Entwickeln einer geeigneten Lösungsidee ist meist ein iterativer Vorgang, der nur bedingt Expertenwissen im konkreten Anwendungsbereich benötigt. Eine gelungene Abstraktion des konkreten Problems in eine mathematische Darstellung zeichnet sich nämlich dadurch aus, dass sämtliche problemrelevanten Größen und Zusammenhänge durch mathematische Objekte repräsentiert werden, die nicht relevanten aber weggelassen werden. Bei der Erarbeitung einer Lösungsidee ist es aber natürlich hilfreich, sich einzelne Ansätze anhand der konkreten Problemstellung veranschaulichen zu können. Die Lösungsidee muss dann in einem geeigneten Rechenkalkül, über den wir noch nicht gesprochen haben, implementiert werden; dadurch erhält man eine Lösung im Modell. Im letzten Schritt wird diese abstrakte Lösung wieder in die Domäne der ursprünglichen Aufgabenstellung zurücktransferiert. Wir gehen im Rahmen dieses Beispiels nicht auf die letztgenannten Punkte ein. \square

Wie wir im letzten Beispiel gesehen haben, erfordert die Lösung eines konkreten Problems mit mathematischen Methoden einen Modellbildungsschritt. Dabei werden alle für die Aufgabenstellung irrelevanten Details weggelassen (etwa chemische Vorgänge innerhalb der Nukleotide), um sich auf die für das Problem wichtigen Informationen konzentrieren zu können. Im Modell, das in unserem Fall in der Sprache der Mathematik formuliert ist, wird dann eine Lösung berechnet, die abschließend aus dem Modell auf die konkrete Ausgangssituation zurücktransferiert wird. Diese Schritte sind in Abbildung 1.1 zusammengefasst.

Vorkenntnisse Im Laufe dieser Vorlesung werden wir uns mit der Sprache und den grundlegenden Konstrukten der Mathematik beschäftigen, um eine Basis für Problemabstraktion und Lösungsschritte im Modell zu schaffen. Dabei werden, bis auf die im Folgenden angeführten Ausnahmen, keinerlei spezielle Mathematik-Kenntnisse vorausgesetzt. Diese Ausnahmen sind:

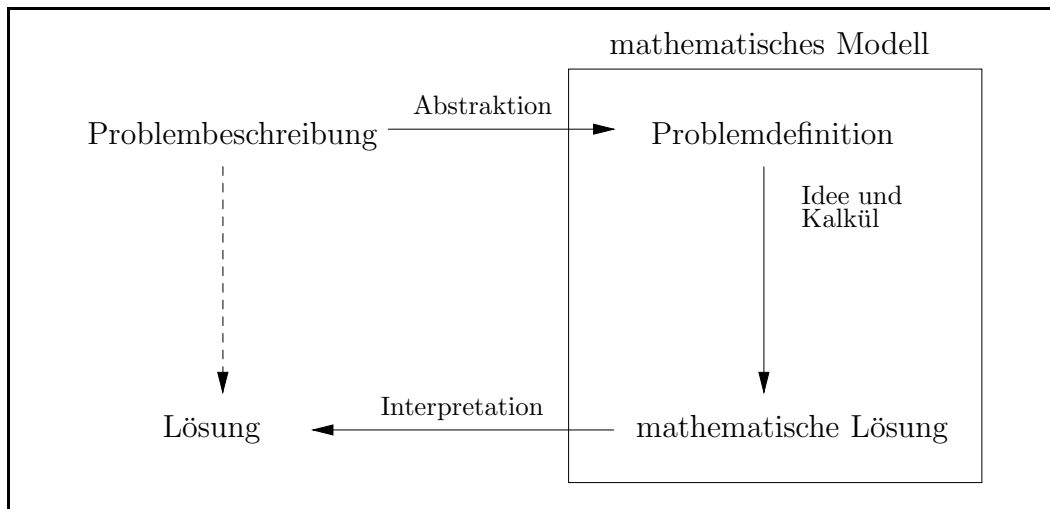


Abbildung 1.1: Problemlösung unter Verwendung eines mathematischen Modells.

- Kenntnis der natürlichen Zahlen \mathbb{N} , ganzen Zahlen \mathbb{Z} , rationalen Zahlen \mathbb{Q} , reellen Zahlen \mathbb{R} , und der darauf definierten Grundrechenarten;
- Kenntnis der “kleiner-als” Beziehung zwischen Zahlen dieser Bereiche;
- Fertigkeit in der Manipulation von Termen, also das Rechnen mit und Umformen von Ausdrücken wie Brüchen oder Potenzen.

Der Grund für die ersten beiden Punkte ist, dass eine saubere Einführung dieser Begriffe zwar möglich und spannend, unter den gegebenen Zeitbeschränkungen aber nicht sinnvoll ist. Der dritte Punkt soll betonen, dass “Rechnen” nicht *Kenntnis* von Rechenregeln bedeutet, sondern die *Fähigkeit*, diese auch richtig anzuwenden.

Notation In diesem Skriptum bezeichne \mathbb{N}_0 die natürlichen Zahlen und 0, und $\mathbb{N}_{\leq n}$ die natürlichen Zahlen kleiner gleich n .

Kapitel 2

Logik als Sprache der Mathematik

In diesem Kapitel werden wir uns mit der Sprache der Mathematik beschäftigen. Dies wird uns erlauben, normalsprachliche Ausdrücke in präzise mathematische Aussagen umzuwandeln, und aus diesen Aussagen Schlüsse zu ziehen. Dazu werden wir sowohl die Struktur (*Syntax*) und die Bedeutung (*Semantik*) von Aussagen betrachten.

2.1 Aussagenlogik

In diesem Abschnitt behandeln wir Aussagen, die entweder wahr oder falsch sind. Wir werden sehen, dass sich solche Aussagen auf unterster Ebene in nicht mehr zerteilbare (*atomare*) Aussagen zerlegen lassen. Beispiele atomarer Aussagen sind etwa

“Eine Vorlesungseinheit ist 45 Minuten lang.” (wahr)

“Hagenberg ist die Hauptstadt von Oberösterreich.” (falsch)

Zusammengesetzte Aussagen bestehen aus atomaren Aussagen, die auf verschiedene Weise miteinander verbunden sein können. Beispiele zusammengesetzter Aussagen sind

“Wenn es regnet, ist die Straße nass.”

“Klaus studiert Software Engineering, aber nicht Mobile Computing.”

Der Wahrheitsgehalt zusammengesetzter Aussagen beruht auf dem Wahrheitswert der atomaren Aussagen, aus denen sie bestehen. Wir werden im Folgenden genau beschreiben, wie atomare Aussagen zu zusammengesetzten Aussagen kombiniert werden können, und wie man den Wahrheitsgehalt zusammengesetzter Aussagen bestimmen kann.

Man beachte, dass nicht alle Sätze des normalen Sprachgebrauchs Aussagen sind. So sind etwa Fragen wie

“Wozu ist das alles gut?”

“Wer entdeckte die Doppelhelixstruktur der DNA?”

oder Befehle wie

“Mach endlich Schluss!”

“Wiederholen Sie dieses Kapitel bis zur nächsten Vorlesung.”

keine Aussagen, da ihnen kein Wahrheitswert zugewiesen werden kann.

Zur Darstellung von Aussagen in einer (für die Mathematik notwendigen) präzisen Struktur repräsentieren wir atomare Aussagen durch Großbuchstaben und Kombinationsmöglichkeiten von Aussagen durch spezielle Symbole. Im Folgenden spezifizieren wir exakt die *Syntax* von Aussagen; damit ist festgelegt, welche mathematischen Zeichenketten Aussagen beschreiben und welche nicht. Die *Semantik*, also die Bedeutung dieser Zeichenketten (entweder wahr oder falsch) wird in einem späteren Schritt definiert.

Definition 2.1 (Aussagen)

Als *Aussagen* bezeichnet man genau jene Zeichenketten, die sich mit den folgenden Regeln bilden lassen:

- (1) Die Zeichen A, B, C, \dots sind (atomare) Aussagen.
- (2) Wenn x und y Aussagen sind, dann sind auch folgende Zeichenketten Aussagen (manchmal auch *wohlgeformte Aussagen* genannt):

$$(\neg x)$$

$$(x \wedge y)$$

$$(x \vee y)$$

$$(x \Rightarrow y)$$

Die Symbole \neg, \wedge, \vee und \Rightarrow , die aus Aussagen zusammengesetzte Aussagen (genannt *Negation, Konjunktion, Disjunktion* bzw. *Implikation*) machen, nennt man *Junktoren*.

Beispiel 2.1 Seien A, B und C atomare Aussagen. Dann sind auch folgende Zeichenketten Aussagen:

$$\begin{aligned} &((A \wedge A) \Rightarrow (\neg B)) \\ &((B \vee (\neg C)) \wedge (A \wedge B)) \\ &((\neg(A \wedge B)) \Rightarrow (\neg(A \vee B))). \end{aligned}$$

Da laut Definition 2.1 nur diejenigen Zeichenketten, die aus den Bildungsregeln zusammengesetzt werden können (und keine anderen!) Aussagen sind, sind folgende Zeichenketten *keine* Aussagen:

$$\begin{aligned} &((\neg A) \wedge B \wedge C \\ &((A \wedge \vee B)) \\ &\wedge B \vee C. \end{aligned}$$

□

Die Struktur von Definition 2.1 nennt man *rekursiv*, da durch Anwendung der Bildungsregeln eine hierarchische Struktur der Zeichenketten erzeugt wird. Man kann bei Aussagen diese hierarchische Struktur herauslesen, indem man die Junktoren auf jeweils oberster Ebene identifiziert.

Beispiel 2.2 Die Aussage $((A \wedge B) \wedge C) \Rightarrow B$ hat folgende hierarchische Struktur:

$$\underbrace{\underbrace{\underbrace{((A \wedge B) \wedge C)}_{\text{Konjunktion}} \Rightarrow B}_{\text{Konjunktion}}}_{\text{Implikation}}$$

Diese Aussage ist also auf oberster Ebene eine Implikation. Ein Beispiel einer Konjunktion ist

$$\underbrace{\underbrace{\underbrace{((\neg A) \vee B) \wedge C}_{\text{Disjunktion}}}_{\text{Negation}}}_{\text{Konjunktion}} \quad \square$$

Um nachzuweisen, dass bestimmte Zeichenketten *keine* wohlgeformten Aussagen sind genügt es nicht, keine hierarchische Zerlegung finden zu können: Man muss vielmehr nachweisen, dass es keine Zerlegung geben kann. Dazu beweist man, dass wohlgeformte Aussagen bestimmte Eigenschaften haben müssen. Wenn eine zu untersuchende Zeichenkette eine dieser Eigenschaften nicht erfüllt, so kann sie auch keine Aussage sein. Solche Eigenschaften von Aussagen sind etwa:

- Es gibt ebensoviele öffnende wie schließende Klammern,
- links und rechts von jedem der Junktoren \wedge und \vee steht eine Aussage,
- eine Aussage endet nie mit einem Junktor.

Man kann beweisen, dass diese Punkte für alle Aussagen gelten. Wir begnügen uns hier damit, auf ihre Plausibilität hinzuweisen, die aus der rekursiven Definition 2.1 herauszulesen ist.

Damit ist klar, warum die Zeichenketten aus Beispiel 2.1 keine Aussagen sind.

Beispiel 2.3 (Fortsetzung von Beispiel 2.1) Die Zeichenkette

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

ist keine Aussage, da sie zwei öffnende, aber nur eine schließende Klammer enthält. Ebenso ist

$$((A \wedge \vee B))$$

keine Aussage, da $\vee B$ (rechts von $A \wedge$) und auch $A \wedge$ (links von $\vee B$) keine Aussagen sind. Mit einem ähnlichen Argument ist auch

$$\wedge B \vee C$$

keine Aussage. □

Die letzten Beispiele haben schon die Grundstruktur mathematischen Schlussfolgerns gezeigt: Um nachzuweisen, dass eine Bedingung erfüllt ist, verwendet man die Definition dieser Bedingung. Da wir Bedingungen als Aussagen formulieren, müssen wir einen Kalkül entwickeln, mit dem man für Aussagen nachrechnen kann, ob sie wahr sind oder nicht. Dazu definieren wir den Begriff des Wahrheitswerts.

Definition 2.2 (Wahrheitswert)

Die beiden einzigen *Wahrheitswerte* sind *wahr* und *falsch*, die als **w** und **f** abgekürzt werden.

Im Gegensatz zur Syntax, mit der die korrekte Form von Aussagen als Zeichenketten beschrieben wird, legt die Semantik die Bedeutung von Aussagen fest. Eine atomare Aussage muss entweder wahr oder falsch sein (aber nicht beides); die Semantik von zusammengesetzten Aussagen lässt sich aus der Semantik der Teilaussagen zusammensetzen. Da sich zusammengesetzte Aussagen über Junktoren aus Teilaussagen bilden lassen genügt es, die Semantik der Junktoren für alle möglichen Wahrheitsbelegungen der Teilausdrücke zu definieren. Eine solche Tabelle von Wahrheitswerten nennt man *Wahrheitstabelle*. Wir verzichten im Folgenden auf die Klammersetzung um Aussagen, die nur einen Junktor enthalten.

Definition 2.3 (Semantik der Negation)

Die Semantik von $\neg x$ ist für die beiden möglichen Wahrheitswerte von x gegeben durch die Tabelle

x	$\neg x$
f	w
w	f

Durch diese Definition wird die Bedeutung der Negation $\neg x$ von x als Verneinung (Gegenteil) von A festgelegt. Die Negation ist der einzige *einstellige* Junktor, da die zusammengesetzte Aussage $\neg x$ nur aus einer Aussage x besteht. Die anderen Junktoren aus Definition 2.1 sind alle *zweistellig*, da sie aus jeweils zwei Aussagen zusammengesetzt sind. Wir geben die Semantik dieser Junktoren in den nächsten Definitionen an. Da es für zwei Aussagen vier verschiedene Möglichkeiten der Zuweisung von Wahrheitswerten gibt, enthalten die Wahrheitstabellen in diesen Definitionen alle vier Zeilen.

Definition 2.4 (Semantik der Konjunktion)

Für zwei Aussagen x und y ist die Semantik der Konjunktion $x \wedge y$ durch folgende Tabelle gegeben:

x	y	$x \wedge y$
f	f	f
f	w	f
w	f	f
w	w	w

Die Konjunktion $x \wedge y$ zweier Aussagen x und y ist somit nur dann wahr, wenn beide Aussagen x und y wahr sind. Der Wahrheitswert der Aussage $x \wedge y$ drückt also die Bedeutung des sprachlichen “und” aus.

Definition 2.5 (Semantik der Disjunktion)

Für zwei Aussagen x und y ist die Semantik der Disjunktion $x \vee y$ durch folgende Tabelle gegeben:

x	y	$x \vee y$
f	f	f
f	w	w
w	f	w
w	w	w

Die Disjunktion $x \vee y$ zweier Aussagen x und y ist also wahr, sobald mindestens eine der beiden Aussagen x und y wahr ist. Der Wahrheitswert von $x \vee y$ ist somit das “oder” der beiden Aussagen x und y , aber nicht im ausschließenden Sinn: $x \vee y$ ist auch dann wahr, wenn sowohl x als auch y wahr sind.

Definition 2.6 (Semantik der Implikation)

Für zwei Aussagen x und y ist die Semantik der Implikation $x \Rightarrow y$ durch folgende Tabelle gegeben:

x	y	$x \Rightarrow y$
f	f	w
f	w	w
w	f	f
w	w	w

Die Aussage x wird die *Prämisse*, die Aussage y die *Konklusion* der Implikation genannt.

Die Definition der Implikation $x \Rightarrow y$ soll die Bedeutung von “aus x folgt y ” widerspiegeln: Wenn x wahr ist, dann ist der Wahrheitswert von $x \Rightarrow y$ von y abhängig: $x \Rightarrow y$ ist für wahre x und y wahr, und für wahres x , aber falsches y falsch. Dies deckt sich mit unserer Erwartung an den Wahrheitswert eines Satzes wie “wenn es regnet, dann ist die Straße nass”. Schwieriger wird es, wenn x falsch ist, da sich die Bedeutung der Implikation dann nicht mehr so leicht durch umgangssprachliche Bedeutung erklären lässt: Wenn es nicht regnet und die Straße nass ist (oder nicht), was ist dann die Bedeutung (wahr oder falsch?) von “wenn es regnet, dann ist die Straße nass”? Wie wir aus Definition 2.6 sehen, legen wir in diesem Fall fest, dass die Implikation wahr ist. Wir halten somit fest, dass die Implikation $x \Rightarrow y$ immer wahr ist, wenn x falsch ist, und sonst gleich dem Wahrheitswert von y ist.

Über Wahrheitstabellen können wir den Wahrheitswert beliebig großer Aussagen bestimmen, da zusammengesetzte Aussagen aus Negationen, Konjunktionen, Disjunktionen oder Implikationen bestehen, und wir die Bedeutung dieser Junktoren soeben definiert haben. Wir betrachten dafür einige Beispiele.

Beispiel 2.4 Die Bedeutung der Aussage $((A \vee B) \Rightarrow (A \wedge (\neg B)))$ ist durch untenstehende Wahrheitstabelle gegeben. Dabei sind zur leichteren Verständlichkeit die Bedeutungen der Teilaussagen ebenfalls aufgelistet.

A	B	$(A \vee B)$	$(A \wedge (\neg B))$	$((A \vee B) \Rightarrow (A \wedge (\neg B)))$
f	f	f	f	w
f	w	w	f	f
w	f	w	w	w
w	w	w	f	f

Wie wir durch den Vergleich der Spalten 2 und 5 dieser Tabelle leicht sehen können, ist der Wahrheitswert von $((A \vee B) \Rightarrow (A \wedge (\neg B)))$ gleich dem Wahrheitswert von $(\neg B)$.

Der Wahrheitswert von $((A \vee B) \wedge C) \wedge (\neg A)$ ist gegeben durch die Tabelle

A	B	C	$(A \vee B)$	$((A \vee B) \wedge C)$	$((A \vee B) \wedge C) \wedge (\neg A)$
f	f	f	f	f	f
f	f	w	f	f	f
f	w	f	w	f	f
f	w	w	w	w	w
w	f	f	w	f	f
w	f	w	w	w	f
w	w	f	w	f	f
w	w	w	w	w	f

Der Wahrheitswert von $((\neg(A \wedge B)) \vee A)$ ist gegeben durch

A	B	$(\neg(A \wedge B))$	$((\neg(A \wedge B)) \vee A)$
f	f	w	w
f	w	w	w
w	f	w	w
w	w	f	w

Aussagen dieser Art, die für jede Belegung der atomaren Aussagen mit Wahrheitswerten immer wahr sind, nennt man *Tautologien*. Das Gegenteil einer Tautologie, also eine Aussage, die für alle Wahrheitswertbelegungen falsch ist, nennt man *Widerspruch*. \square

Definition 2.7 (Gleichwertigkeit von Aussagen)

Zwei Aussagen x und y heißen genau dann *gleichwertig*, wenn ihre Wahrheitswerte für alle Wahrheitswertbelegungen der atomaren Aussagen in x und y gleich sind. In diesem Fall schreiben wir $x \equiv y$ und sagen auch, dass x und y *äquivalent* sind.

Wie wir in Beispiel 2.4 gesehen haben, gilt etwa

$$((A \vee B) \Rightarrow (A \wedge (\neg B))) \equiv (\neg B).$$

Man beachte, dass die Zeichenkette " $x \equiv y$ " *keine* Aussage ist, da sie nicht Definition 2.1 genügt. $x \equiv y$ ist vielmehr eine *Meta-Aussage*, die etwas über die beiden Aussagen x und y aussagt.

Das nächste Beispiel zeigt, dass bei zusammengesetzten Konjunktionen und Disjunktionen zur leichteren Lesbarkeit die Klammern weggelassen werden können, ohne die Semantik der Aussagen zu verändern.

Beispiel 2.5 Es gilt

$$((A \vee B) \vee C) \equiv (A \vee (B \vee C)).$$

Wir rechnen dies über die Wahrheitstabellen nach:

A	B	C	$(A \vee B)$	$((A \vee B) \vee C)$	$(B \vee C)$	$(A \vee (B \vee C))$
f	f	f	f	f	f	f
f	f	w	f	w	w	w
f	w	f	w	w	w	w
f	w	w	w	w	w	w
w	f	f	w	w	f	w
w	f	w	w	w	w	w
w	w	f	w	w	w	w
w	w	w	w	w	w	w

Wir können somit zwischen reinen Disjunktionen die Klammern beliebig setzen; dies gilt auch für längere zusammengesetzte Disjunktionen. Da die Klammersetzung keine Rolle spielt, können wir *alle* Klammern weglassen, und schreiben abkürzend $A \vee B \vee C$ für $((A \vee B) \vee C)$.

Mit einer ähnlichen Herleitung kann man nachrechnen, dass auch bei zusammengesetzten Aussagen, die nur aus Konjunktionen bestehen, die Klammersetzung keine Rolle spielt. Wir kürzen auch hier $((A \wedge B) \wedge C)$ als $A \wedge B \wedge C$ ab. \square

Die im letzten Beispiel nachgerechnete Eigenschaft, Klammern zwischen Disjunktionen beliebig setzen zu können, nennt man *Assoziativität*. Auch die Konjunktion ist assoziativ, da man auch hier die Klammersetzung vertauschen kann. Eine weitere wichtige Eigenschaft von Disjunktionen und Konjunktionen ist die *Kommutativität*: Die Reihenfolge der Operanden kann vertauscht werden, denn es gilt

$$x \vee y \equiv y \vee x \quad \text{und} \quad x \wedge y \equiv y \wedge x.$$

Man kann leicht nachrechnen, dass die Implikation weder assoziativ noch kommutativ ist.

Im Folgenden werden wir alle Klammern weglassen, die zur eindeutigen Lesbarkeit von Aussagen nicht notwendig sind. Wir werden die äußersten Klammern um Aussagen weglassen und festlegen, dass sich die ungeklammerte Negation immer auf die nächste atomare Aussage bezieht. Somit ist die Zeichenkette

$$x \wedge \neg A \quad \text{eine Kurzform für} \quad (x \wedge (\neg A)).$$

Wie wir gesehen haben, lassen sich mit Hilfe von Wahrheitstabellen Äquivalenzen von Aussagen nachweisen. Man kann dann wiederum die Äquivalenzen ausnutzen, um zusammengesetzte Aussagen zu vereinfachen, ohne Wahrheitstabellen konstruieren zu müssen. In Satz 2.2 weiter unten werden die am häufigsten verwendeten Äquivalenzen zusammengefasst. In dieser Liste von Äquivalenzen scheint die Implikation als Junktor nicht auf. Dies ist dadurch begründet, dass sich die Implikation durch Negation und Disjunktion ausdrücken lässt.

Satz 2.1 Seien x und y zwei Aussagen. Dann gilt

$$x \Rightarrow y \equiv \neg x \vee y \quad \text{und} \quad x \Rightarrow y \equiv \neg y \Rightarrow \neg x.$$

Die Implikation kann somit als Kurzform einer speziellen Disjunktion gesehen werden. Wie wir in den nächsten Rechenregeln sehen werden, können auch Disjunktionen durch Konjunktionen und Negationen ausgedrückt werden (und umgekehrt). Man benötigt somit nur Disjunktion und Negation (oder Konjunktion und Negation), um alle Aussagen formulieren zu können.

Satz 2.2 Seien x, y und z Aussagen. Dann gilt:

$\neg(\neg x) \equiv x$		(Involution)
$x \wedge y \equiv y \wedge x$	$x \vee y \equiv y \vee x$	(Kommutativität)
$x \wedge (y \wedge z) \equiv (x \wedge y) \wedge z$	$x \vee (y \vee z) \equiv (x \vee y) \vee z$	(Assoziativität)
$x \vee (y \wedge z) \equiv (x \vee y) \wedge (x \vee z)$	$x \wedge (y \vee z) \equiv (x \wedge y) \vee (x \wedge z)$	(Distributivität)
$x \vee (x \wedge y) \equiv x$	$x \wedge (x \vee y) \equiv x$	(Absorption)
$x \wedge x \equiv x$	$x \vee x \equiv x$	(Idempotenz)
$x \wedge \mathbf{w} \equiv x$	$x \vee \mathbf{f} \equiv x$	(Neutralität)
$x \wedge \neg x \equiv \mathbf{f}$	$x \vee \neg x \equiv \mathbf{w}$	(Komplementarität)
$x \wedge \mathbf{f} \equiv \mathbf{f}$	$x \vee \mathbf{w} \equiv \mathbf{w}$	(Invarianz)
$\neg(x \wedge y) \equiv \neg x \vee \neg y$	$\neg(x \vee y) \equiv \neg x \wedge \neg y$	(DeMorgan Regeln)

Wir illustrieren die Verwendung dieser Regeln anhand eines Beispiels.

Beispiel 2.6 Wir können durch Umformungen unter Verwendung der Rechenregeln aus Satz 2.1 und Satz 2.2 nachrechnen, dass gilt:

$$(A \vee B) \Rightarrow A \equiv B \Rightarrow A.$$

Wir formen die linke Seite dieser Äquivalenz wie unten angegeben um und erhalten dann die rechte Seite. Wir geben in Klammern an, welche der Rechenregeln bei der Umformung angewendet worden ist.

$$\begin{aligned}
 (A \vee B) \Rightarrow A &\stackrel{\text{(Satz 2.1)}}{\equiv} \neg(A \vee B) \vee A \\
 &\stackrel{\text{(DeMorgan)}}{\equiv} (\neg A \wedge \neg B) \vee A \\
 &\stackrel{\text{(Kommutativität)}}{\equiv} A \vee (\neg A \wedge \neg B) \\
 &\stackrel{\text{(Distributivität)}}{\equiv} (A \vee \neg A) \wedge (A \vee \neg B) \\
 &\stackrel{\text{(Komplementarität)}}{\equiv} \mathbf{w} \wedge (A \vee \neg B) \\
 &\stackrel{\text{(Neutralität)}}{\equiv} A \vee \neg B \\
 &\stackrel{\text{(Kommutativität)}}{\equiv} \neg B \vee A \\
 &\stackrel{\text{(Satz 2.1)}}{\equiv} B \Rightarrow A
 \end{aligned}$$

□

Wir haben oben eine Notation kennengelernt, mit der man logische Zusammenhänge zwischen Aussagen ausdrücken kann. Damit haben wir die Möglichkeit, den in Ka-

pitel 1 angesprochenen Modellbildungsschritt vorzunehmen. Da wir das Modell aber meist nur bilden, um damit eine Problemlösung zu vereinfachen, kann die Modellbildung als erster Schritt im Lösungsprozess gesehen werden.

Für den zweiten Schritt, die Problemlösung im Modell, benötigen wir einen *Kalkül*, um die Wahrheit von Aussagen zu bestimmen. Mit Wahrheitstabellen und Rechenregeln stehen uns dafür zwei Möglichkeiten zur Verfügung. Andererseits ist es oft von Vorteil, aus einer gegebenen Menge von wahren Aussagen weitere wahre Aussagen erzeugen zu können – meist sind wir an Aussagen, die in unserem Modell nicht gelten, sowieso nicht interessiert. Faktenwissen über einen Bereich kann meist in Form von Konjunktionen und Implikationen formuliert werden (“was alles gilt und was daraus folgt”). Die Implikationen können verwendet werden, um weitere wahre Aussagen zu erzeugen. Die dabei verwendete Schlussregel wird *modus ponens* genannt.

Satz 2.3 (Modus ponens) Seien x und $x \Rightarrow y$ zwei wahre Aussagen. Dann ist auch y eine wahre Aussage.

Von der Richtigkeit dieser Schlussregel kann man sich durch die Wahrheitstabelle in Definition 2.6 überzeugen. Wir betrachten zwei Beispiele, wie man mittels modus ponens schlussfolgern kann.

Beispiel 2.7 Die Implikation “Wenn es regnet, ist die Straße nass” sei wahr. Wenn wir nun hören, dass es regnet (die Prämisse der Implikation also wahr ist) können wir daraus schließen, dass die Straße nass ist. \square

Beispiel 2.8 Gegeben sei die wahre Aussage

$$A \wedge B \wedge \neg C \wedge (A \Rightarrow D) \wedge (D \Rightarrow E).$$

Aus Definition 2.4 wissen wir, dass dann A , B und $\neg C$ sowie die Implikationen $A \Rightarrow D$ und $D \Rightarrow E$ auch wahre Aussagen sind (eine Konjunktion von Aussagen ist genau dann wahr, wenn alle Aussagen wahr sind). Mit modus ponens folgt zuerst, dass D wahr ist, und mit nochmaliger Anwendung auch, dass E wahr ist. \square

Wir werden diese Art der Schlussfolgerung im weiteren Verlauf noch oft benötigen.

2.2 Prädikatenlogik

Der im letzten Abschnitt besprochene Logikkalkül erlaubt es uns, mittels Aussagen Objekte und deren Zusammenhänge zu beschreiben. Die Ausdruckskraft dieses Kalküls ist jedoch sehr eingeschränkt, da eine Aussage sich immer auf *bestimmte fixe* Objekte bezieht, der Wahrheitswert einer Aussage somit konstant wahr oder falsch ist. In vielen Bereichen benötigt man hingegen Aussagen, die in Abhängigkeit der Objekte, die sie beschreiben, wahr oder falsch sind: Aussagen wie

“ x teilt 28.”

“ y ist der Sohn von Franz.”

sind in Abhängigkeit des Werts der Variablen x bzw. y entweder wahr oder falsch. Die erste Aussage ist somit für die ganzen Zahlen 1, 2, 4, 7, 14 und 28 wahr; zur Überprüfung der Wahrheit der zweiten Aussage benötigen wir mehr Informationen über die Objekte, auf die sich die Aussage bezieht.

Im Folgenden werden wir sehen, wie wir mit Hilfe der *Prädikatenlogik* allgemeinere Zusammenhänge als mit der Aussagenlogik beschreiben können. Dazu benötigen wir einige Grundbegriffe, um mit diesem Logikkalkül formal korrekt operieren zu können. Wir nehmen zuerst an, ein *Universum* von Objekten gegeben zu haben (dieses wird auch *Domäne* genannt). Einzelne Objekte (etwa “Franz” oder π) dieses Universums haben spezielle Bezeichner (*Objektkonstante* genannt). Wenn man Aussagen über beliebige Objekte machen möchte, so verwendet man dazu *Variablen*; diese Aussagen werden dann je nach *Variablenbelegung* wahr oder falsch.

In einem Universum können Objekte – einzeln oder in Kombination – in eindeutigen Zusammenhang mit anderen Objekten stehen; es gibt *Funktionen*, um diesen Zusammenhang auszudrücken:

$$\begin{aligned} \cdot^2 &: 2 \mapsto 4, 3 \mapsto 9 \\ \cdot + \cdot &: (2, 3) \mapsto 5, (3, 9) \mapsto 12 \\ \text{vaterVon}(\cdot) &: \text{Franz} \mapsto \text{Klaus} \end{aligned}$$

Die Namen von Funktionen werden als *Funktionskonstante* bezeichnet. In den obigen Beispielen sind dies 2 , $+$ und *vaterVon*.

Ebenso gibt es *Prädikate*, die Eigenschaften von Objekten, oder den Zusammenhang zwischen Objekten ausdrücken. Für ein gegebenes Prädikat kann man entscheiden, ob bestimmte Objekte diese Eigenschaft haben, oder in dieser Beziehung zueinander stehen. Somit sind Prädikate genau die Eigenschaften von und Zusammenhänge zwischen Objekten, die in Abschnitt 2.1 als atomare Aussagen bezeichnet wurden.

Beispiele von Prädikaten sind etwa:

- “5 ist eine Primzahl.”
- “3 ist kleiner als 4.”
- “Rom ist die Hauptstadt von Italien.”

Analog zu Funktionen werden die Namen von Prädikaten als *Prädikatenkonstante* bezeichnet. Mögliche sprechende Namen der Prädikate in obigem Beispiel sind *istPrim* und *istHauptstadtVon*.

In der Prädikatenlogik ist zu beachten, dass man nur Variablen zulässt, die mit Objekten belegt werden können. Prinzipiell könnten auch Funktions- und Prädikatenvariable zugelassen werden; dies würde zwar die Ausdruckskraft steigern, den logischen Kalkül aber abschwächen. Wir werden uns somit auf die *Prädikatenlogik erster Ordnung* beschränken, deren Grundgerüst wir soeben informell beschrieben haben.

Wir werden in Kapitel 3 sehen, wie sich mathematische Objekte (Zahlen, Funktionen, ...) präzise definieren lassen. In diesem Abschnitt nehmen wir an, dass diese Objekte schon gegeben sind, und wollen nun Aussagen über diese Objekte formulieren können. Als erstes behandeln wir Möglichkeiten, Objekte selbst zu beschreiben.

Definition 2.8 (Terme)

Gegeben sei eine Menge von Variablen, eine Menge von Objektkonstanten, und eine Menge von Funktionskonstanten. Als *Terme* bezeichnet man genau jene Zeichenketten, die sich mit den folgenden Regeln bilden lassen.

- (1) Jede Variable ist ein Term.
- (2) Jede Objektkonstante ist ein Term.
- (3) Wenn f ein Funktionskonstante ist und t_1, t_2, \dots, t_n Terme sind, dann ist auch $f(t_1, t_2, \dots, t_n)$ ein Term.

Wir behandeln einfache Beispiele.

Beispiel 2.9 Sei $\{x, y\}$ die Menge der Variablen, die ganzen Zahlen \mathbb{Z} die Menge der Objektkonstanten, und $\{+, -, *\}$ die Menge der Funktionskonstanten. Dann sind folgende Zeichenketten Terme, die ganze Zahlen beschreiben:

$$\begin{aligned} &+(9, y) \\ &-((2, -3), +(2, x)) \\ &*((3, 2), 8) \end{aligned}$$

□

Die Zusammenhänge zwischen Objekten lassen sich durch atomare Aussagen beschreiben. Diese sind wie folgt definiert.

Definition 2.9 (Atomare Aussagen)

Gegeben sei eine Menge von Prädikatenkonstanten. Als *atomare Aussagen* bezeichnet man Zeichenketten, die von der Form

$$P(t_1, \dots, t_n)$$

sind, wobei P eine Prädikatenkonstante ist und t_1, \dots, t_n Terme sind. Eine atomare Aussage ist entweder wahr oder falsch; wenn zumindest einer der Terme t_1, \dots, t_n Variablen enthält, hängt der Wahrheitswert der atomaren Aussage von der Variablenbelegung ab.

Es folgen wiederum einige einfache Beispiele.

Beispiel 2.10 Seien die Terme wie in Beispiel 2.9. Die Menge der Prädikatenkonstanten enthalte die Prädikate $|$ ("teilt") und $<$ ("kleiner als"). Dann sind die folgenden Zeichenketten atomare Aussagen, deren Wahrheitswert in Klammern angegeben ist.

$$\begin{aligned} &<(33, 5) && \text{(falsch)} \\ &|(3, *(3, 1729)) && \text{(wahr)} \\ &<(3, +(x, *(2, x))) && \text{(hängt vom Wert von } x \text{ ab)} \end{aligned}$$

□

Bei Funktionen und Prädikaten gibt man der Anzahl der Terme, auf die sie angewandt werden, einen Namen.

Definition 2.10 (Stelligkeit)

Die Anzahl der Terme, auf die eine Funktions- oder Prädikatenkonstante angewandt werden kann, wird *Stelligkeit* genannt.

Beispiel 2.11 Die Funktionskonstanten $+$, $-$, $*$, $/$ sind zweistellig; es gibt noch eine weitere einstellige Funktionskonstante $-$. Eigentlich sind die beiden $-$ zwei unterschiedliche Funktionen, zu deren Bezeichnung aber die gleiche Konstante verwendet wird. Weitere einstellige Funktionskonstanten sind $\sqrt{\quad}$ und 2 .

Beispiele von unterschiedlich stelligen Prädikatenkonstanten sind etwa

istGrün(Klee)	(einstellig)	
$< (3, 5)$	(zweistellig)	
istGGT(3,9,15)	(dreistellig).	□

Die Terme und atomaren Aussagen in Beispiel 2.9 und Beispiel 2.10 unterscheiden sich von der “normalen” mathematischen Schreibweise, da laut Definition von Termen und Aussagen das Funktions- bzw. Prädikatensymbol an erster Stelle zu stehen hat. Diese Schreibweise wird *Präfix-Notation* genannt; lesbarer ist meist die gewohnte *Infix-Notation*, die wir im Folgenden bevorzugt verwenden werden. Man beachte, dass bei Präfix-Notation keine Klammern benötigt werden, um die eindeutige Lesbarkeit zu garantieren; bei Infix-Notation müssen möglicherweise noch Klammern eingefügt werden.

Wie schon in Abschnitt 2.1 können atomare Aussagen mittels Junktoren zu Aussagen verbunden werden; diese werden auch *Formeln* genannt. Im folgenden Beispiel untersuchen wir die Bestandteile einer größeren Formel; dieser Vorgang wird *Syntaxanalyse* genannt.

Beispiel 2.12 Gegeben sei die Formel

$$(P(f(x, 12), g(1, 2, 3), 3 + 4) \wedge (x < 3)) \Rightarrow R(h(x), f(3, 1)).$$

Wir bezeichnen mit PK eine Prädikatenkonstante, mit FK eine Funktionskonstante (wobei wir bei beiden noch die Stelligkeit angeben), OK sei eine Objektkonstante, und V eine Variable. Mit diesen Symbolen können wir alle Komponenten obiger Formel identifizieren:

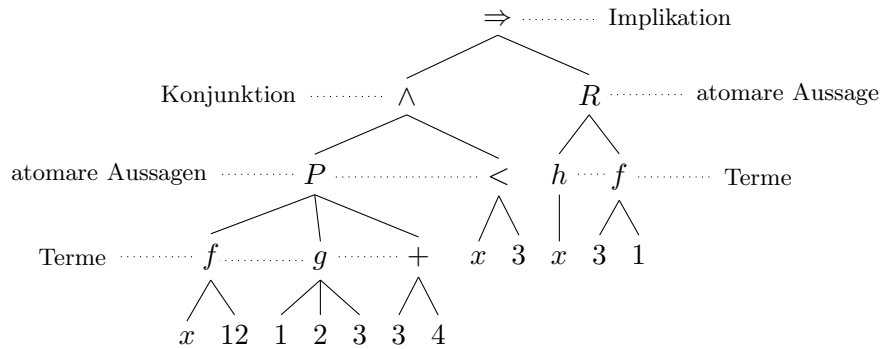
$$\text{PK3 FK2 V OK FK3 OK OK OK OK FK2 OK V PK2 OK} \\ (P (f (x, 12), g (1 , 2 , 3), 3 + 4) \wedge (x < 3))$$

$$\text{PK2 FK1 V FK2 OK OK} \\ \Rightarrow R (h (x), f (3 , 1)).$$

Die *Struktur* dieser Formel ist analog zur Struktur von zusammengesetzten Aussagen in Abschnitt 2.1 zu bestimmen. Wir wechseln hier allerdings zu der in der Informatik

üblichen Darstellung durch einen *Syntaxbaum*, bei der die Wurzel jedes (Teil)Baums vom syntaktisch obersten (äußersten) Symbol gebildet wird.

Wir erhalten für die Formel oben folgenden Syntaxbaum, der inkrementell von unten nach oben aufgebaut wird:



Man beachte, dass der Wahrheitswert dieser Aussage von der Wertebelegung der Variablen x abhängt. □

Alle Variablen, die in einer Aussage vorkommen, werden als *frei* bezeichnet. Die Aussage $x \mid 42$ ist eine Aussage über die freie Variable x ; für bestimmte Belegungen von x ist diese Aussage wahr, für andere falsch. Ebenso nennt man alle in einem Term vorkommenden Variablen frei.

In vielen Fällen möchte man ausdrücken, dass eine Aussage *für alle möglichen Variablenbelegungen* wahr ist. So gilt etwa die Aussage

$$x > 0$$

für alle natürlichen Zahlen. Man möchte diesen Sachverhalt auch in der Sprache der Prädikatenlogik ausdrücken können. Ebenso möchte man oft ausdrücken, dass eine Aussage *für zumindest eine Variablenbelegung* gilt; dies ist etwa bei

$$x \mid 12$$

der Fall. Diese Anforderungen werden in der Prädikatenlogik durch *Quantoren* erfüllt.

Definition 2.11 (Allaussagen, Existenzaussagen)

Sei A ein Aussage mit freier Variable x . Dann nennt man die Zeichenkette

$$(\forall x A)$$

eine *Allaussage*, und die Zeichenkette

$$(\exists x A)$$

eine *Existenzaussage*. Die Symbole \forall (“für alle”) und \exists (“es existiert ein”) werden als *Allquantor* bzw. *Existenzquantor* bezeichnet. Die Variable x wird in A durch die Quantoren *gebunden*.

Obige Definition bestimmt nur die Syntax von All- und Existenzaussage; ihre Semantik ist wiederum ein Wahrheitswert (**w** oder **f**). Wir legen Folgendes fest.

Definition 2.12 (Semantik von All- und Existenzaussagen)

Sei A eine Aussage, in der nur die Variable x frei vorkommt. Dann ist die Allaussage

$$(\forall x A)$$

genau dann wahr, wenn die Aussage A für alle möglichen Belegungen der Variablen x wahr ist. Die Existenzaussage

$$(\exists x A)$$

ist genau dann wahr, wenn es zumindest eine Belegung der Variablen x gibt, die die Aussage A wahr macht.

Wir werden die Klammern um die Quantoraussagen weglassen, wenn klar ist, auf welche Aussage sich der Quantor bezieht. Die Aussage A , in der ein Quantor eine Variable bindet, wird als *Wirkungsbereich* des Quantors bezeichnet.

Man beachte, dass der Name einer gebundenen Variable am Wahrheitswert einer quantifizierten Aussage nichts ändert: So ist die Aussage $\forall x \ x > 0$ genau dann wahr, wenn auch die Aussage $\forall y \ y > 0$ wahr ist. Wenn in einer quantifizierten Aussage noch freie Variablen vorkommen, dann ist auch bei diesen Aussagen der Wahrheitswert von der Variablenbelegung der freien Variablen abhängig. Wir betrachten dazu einige Beispiele.

Beispiel 2.13 Für dieses Beispiel betrachten wir als Universum die ganzen Zahlen. Die Aussage

$$\forall x \exists y \ x < y$$

(“für jede Zahl gibt es eine, die größer ist”) ist dann wahr: Für jede Belegung der Variablen x können wir eine Belegung von y finden, die größer als die von x ist.

Die Aussage

$$\exists x \forall y \ x \mid y$$

(“es gibt eine Zahl, die alle Zahlen teilt”) ist ebenfalls wahr, da dies mit der Belegung $x \leftarrow 1$ für jede beliebige Belegung von y wahr ist.

Falsch ist hingegen die Aussage

$$\exists x \forall y \ y < x.$$

(“es existiert eine Zahl, die größer als alle Zahlen ist”). Wir werden später sehen, wie man auch formal nachrechnen kann, dass diese Aussage falsch ist.

Ebenso falsch ist

$$\forall x \exists y \ y < x$$

(“für jede Zahl gibt es eine, die kleiner ist”). Dies ist für die Belegung $x \leftarrow 1$ und das Universum der natürlichen Zahlen falsch, somit ist auch die Allaussage falsch. \square

Für mathematisch interessante Aussagen benötigt man fast immer eine spezielle zweistellige Prädikatenkonstante, die *Gleichheit*. Dieses Prädikat ist in allen Domänen gegeben und erfüllt folgende Bedingungen.

Definition 2.13 (Gleichheit)

Seien x , y und z Variable, f eine Funktionskonstante, und P eine Prädikatenkonstante. Für die Prädikatenkonstante “=” (*Gleichheit*) gilt:

$$\begin{aligned} x &= x && \text{(Reflexivität)} \\ x = y &\Rightarrow y = x && \text{(Symmetrie)} \\ x = y \wedge y = z &\Rightarrow x = z && \text{(Transitivität)} \\ x = y &\Rightarrow f(\dots, x, \dots) = f(\dots, y, \dots) \\ x = y &\Rightarrow P(\dots, x, \dots) = P(\dots, y, \dots) \end{aligned}$$

Wir werden für $\neg(x = y)$ abkürzend $x \neq y$ schreiben. Die Verwendung des Gleichheitszeichens als Prädikatenkonstante erklärt auch, warum wir für die Schreibweise der Variablenbelegung in Beispiel 2.13 den Pfeil “ \leftarrow ” gewählt haben.

Mit Hilfe der Gleichheit können wir auch interessantere Aussagen formulieren. Diese Aussagen sind wiederum entweder wahr oder falsch; was von beiden zutrifft, weiss man bei manchen allerdings bis heute nicht.

Beispiel 2.14 Die Aussage

$$\forall n \ n \geq 3 \Rightarrow \neg(\exists x \exists y \exists z \ x \neq 0 \wedge y \neq 0 \wedge z \neq 0 \wedge x^n + y^n = z^n)$$

(der *Satz von Fermat*) wurde erst 1995 für die Domäne der ganzen Zahlen als wahr bewiesen.

Der Wahrheitswert der folgenden Aussage (die *Goldbachsche Vermutung*) ist noch immer offen (es wird vermutet, dass sie wahr ist):

$$\forall x \ (\text{gerade}(x) \wedge x > 2) \Rightarrow \exists y \exists z \ \text{prim}(y) \wedge \text{prim}(z) \wedge x = y + z$$

(“jede gerade Zahl größer 2 lässt sich als Summe zweier Primzahlen schreiben”). Hier nehmen wir an, dass die beiden einstelligigen Prädikatenkonstanten *gerade* und *prim* bereits gegeben sind. Wir werden im Folgenden sehen, wie man solche Prädikate aus einfacheren Prädikaten aufbauen kann. \square

Die Übertragung von Zusammenhängen realer Objekte in den Kalkül der Prädikatenlogik erfordert einige Übung; speziell am Beginn ist auf einige Details besonders zu achten.

Beispiel 2.15 Der Satz “die Katze ist grau” ist über eine einstellige Prädikatenkonstante *istGrau* als

$$\text{istGrau}(\text{Katze})$$

zu repräsentieren, wobei *Katze* eine Objektkonstante ist. Falsch ist es, dies als “*Katze = grau*” zu schreiben, da dies aussagt, dass das Objekt *Katze* gleich dem Objekt *grau* ist.

Der Satz “alle Tiroler sind lustig” ist als Allaussage

$$\forall x \text{ istTiroler}(x) \Rightarrow \text{istLustig}(x)$$

zu schreiben. Da durch den Allquantor die gebundene Variable x mit allen Objekten belegt werden kann, die Aussage aber nur für Tiroler gelten soll, muss die Aussage die Form einer Implikation haben. Ein möglicher Fehler wäre, den Satz als $(\forall x \text{ istTiroler}(x) \wedge \text{istLustig}(x))$ zu schreiben. Diese Aussage gibt aber an, dass alle Objekte des Universums Tiroler und lustig sind.

Bei Existenzaussagen ist die Übersetzung meist leichter. Ohne große Schwierigkeiten schreibt man den Satz “es gibt eine grüne Maus” als

$$\exists x \text{ istMaus}(x) \wedge \text{istGrün}(x). \quad \square$$

Die im obigen Beispiel gezeigte genauere Spezifikation von Variablen im Wirkungsbereich eines Quantors ist so weitläufig, dass sich dafür eine eigene Schreibweise eingebürgert hat.

Definition 2.14 (Abkürzende Quantorenschreibweisen)

Sei A eine Aussage und P_x eine Aussage mit freier Variable x . Die Aussage

$$\forall_{P_x} A$$

ist Kurzschreibweise für $\forall x P_x \Rightarrow A$. Ebenso ist

$$\exists_{P_x} A$$

eine Abkürzung für $\exists x P_x \wedge A$.

In dieser Definition betonen wir, dass die Aussage P_x eine freie Variable x enthalten muss (und nur diese!), da wir ja über eine Variable quantifizieren müssen. Im Gegensatz dazu muss die Aussage A die Variable x nicht frei enthalten (sie wird es aber meist tun).

Beispiel 2.16 Wir wählen als Domäne die natürlichen Zahlen. Dann schreiben wir

$$\forall_{i \leq 9} 2^i < 1000$$

für die Aussage $\forall i (i \leq 9) \Rightarrow (2^i < 1000)$. Ebenso ist die Aussage

$$\exists_{x < 4} x | 8$$

eine Kurzschreibweise für die Aussage $\exists x x < 4 \wedge x | 8$. □

Zur weiteren Abkürzung der Schreibweise vereinbaren wir außerdem, dass aufeinanderfolgende Quantoren zusammengezogen werden können. So steht

$$\begin{array}{lll} \forall x, y, z A & \text{für} & \forall x \forall y \forall z A \quad \text{und} \\ \exists x, y, z A & \text{für} & \exists x \exists y \exists z A. \end{array}$$

Über die Negation besteht ein enger Zusammenhang zwischen All- und Existenzquantor; dieser Zusammenhang ist analog zur DeMorgan Regel aus Satz 2.2, durch die verneinte Konjunktion in Disjunktion umgewandelt werden können (und umgekehrt).

Satz 2.4 (DeMorgan) Sei A eine Aussage. Dann gilt

$$\neg \forall x A \equiv \exists x \neg A \quad \text{und} \quad \neg \exists x A \equiv \forall x \neg A.$$

Von der Richtigkeit dieser Äquivalenzen kann man sich mit folgendem Argument überzeugen: $\neg \forall x A$ ist genau dann wahr, wenn nicht für alle Variablenbelegungen von x die Aussage A wahr wird, wenn es also zumindest eine Variablenbelegung gibt, für die A falsch wird. Wenn A aber falsch ist, dann ist $\neg A$ wahr, und somit auch $\exists x \neg A$. Das Argument für die zweite Äquivalenz ist ähnlich.

Ebenso kann man nachrechnen, dass obiger Satz auch gilt, wenn man die abkürzende Schreibweise für erweiterte All- und Existenzaussagen verwendet. Dies ist nicht ganz intuitiv, da die Abkürzung im einen Fall eine Implikation, und im anderen ein Konjunktion bedeutet. Nichtsdestotrotz gilt

$$\neg \forall_{P_x} A \equiv \exists_{P_x} \neg A \quad \text{und} \quad \neg \exists_{P_x} A \equiv \forall_{P_x} \neg A.$$

Nach dieser Einführung in die Grundlagen der Prädikatenlogik können wir uns diesen Kalkül zu Nutze machen, um komplexe Zusammenhänge durch einfachere zu beschreiben. Das eigentliche *Schlussfolgern* in diesem Kalkül werden wir auf die nächsten Kapitel verschieben, wenn wir die Objekte selbst (Zahlen, Funktionen, Prädikate, ...) genauer spezifiziert haben.

Definition 2.15 (Prädikatendefinition)

Eine *Definition* eines Prädikats führt eine neue Prädikatenkonstante ein. Der Wahrheitswert der atomaren Aussage, die durch diese Prädikatenkonstante gegeben ist, wird dabei durch eine andere (meist größere) Aussage bestimmt. Syntaktisch schreiben wir die Definition einer n -wertigen Prädikatenkonstanten P als

$$P(x_1, \dots, x_n) :\Leftrightarrow A_{x_1, \dots, x_n},$$

wobei A_{x_1, \dots, x_n} eine Aussage mit freien Variablen x_1, \dots, x_n ist.

Da wir dieses Werkzeug vorher noch nicht zur Verfügung hatten, haben wir alle Prädikate als gegeben angenommen. Jetzt können wir Eigenschaften wie *gerade* oder *prim* explizit definieren.

Beispiel 2.17 Gerade Zahlen (und nur gerade Zahlen) erfüllen die Eigenschaft, Vielfache von 2 zu sein. Somit können wir definieren:

$$\text{gerade}(x) :\Leftrightarrow \exists_{y \in \mathbb{Z}} x = 2y.$$

Ebenso können wir Primzahlen über ihre Eigenschaften definieren, nur durch 1 und sich selbst teilbar zu sein. Dazu benötigen wir noch den Begriff der Teilbarkeit, den wir auch als zweistelliges Prädikat (diesmal aber in Infix-Notation) definieren können:

$$x \mid y :\Leftrightarrow \exists_{z \in \mathbb{Z}} y = z \cdot x.$$

Damit ist das einstellige Prädikat *prim* gegeben durch:

$$\text{prim}(x) :\Leftrightarrow x \neq 1 \wedge \neg \exists y \ y \neq 1 \wedge y \neq x \wedge y \mid x. \quad \square$$

Als abschließendem Teil dieser Einführung in die Prädikatenlogik wenden wir uns nun detaillierter dem Sprachelement “Quantor” zu. Wir haben in diesem Kapitel bisher Allquantor und Existenzquantor kennengelernt und werden nun sehen, dass diese beiden Symbole Vertreter einer größeren Klasse von Quantoren sind.

Definition 2.16 (Quantor)

Ein *Quantor* ist ein Sprachkonstrukt, das syntaktisch sowohl eine Variable als auch einen Term oder einer Aussage benötigt. Der Quantor bindet diese Variable; je nach Semantik des Quantors ist der Quantorausdruck ein Term oder eine Aussage.

Man kann erkennen, dass All- und Existenzquantor aus Definition 2.11 diese Bedingungen erfüllen. In beiden Fällen werden Variable gebunden, und aus Aussagen All- bzw. Existenzaussagen gebildet.

Ein einfaches Beispiel eines Quantors, der aus einer Variable und einem Term einen Term liefert, ist der *Summenquantor* \sum . Meist ist beim Summenquantor nur die abkürzende Quantorenschreibweise aus Definition 2.14 sinnvoll, da die gebundene Variable einen Wertebereich benötigt. Die Semantik eines Summenquantisierten Terms ist die Summe aller Terme, die sich durch alle gültigen Variablenbelegungen der gebundenen Variablen ergeben. So gelten etwa für das Universum der natürlichen Zahlen:

$$\sum_{i \leq 10} i = 1 + 2 + \dots + 10 = 55 \quad \text{und} \quad \sum_{2 \leq k \leq 6} 2^k = 2^2 + 2^3 + \dots + 2^6 = 124.$$

Die Aussage, die die Werte der gebundenen Variablen einschränkt, wird oftmals auf den Bereich unter- und oberhalb des Summenzeichens aufgeteilt, wie in folgenden Beispielen zu sehen ist:

$$\sum_{m=2}^{20} m^2 = 2^2 + \dots + 20^2 = 2869 \quad \text{und} \quad \sum_{n=3}^7 n^2 + 1 = (3^2 + 1) + \dots + (7^2 + 1) = 140.$$

Analog zum Summenquantor lässt sich der *Produktquantor* \prod definieren. Bei diesem werden die Terme nach erfolgter Variablenbelegung multipliziert statt addiert. Zwei Beispiele seiner Anwendung sind wie folgt:

$$\prod_{\substack{2 \leq k \leq 8 \\ \text{prim}(k)}} k = 2 \cdot 3 \cdot 5 \cdot 7 = 210 \quad \text{und} \quad \prod_{j=2}^5 2^j = 2^2 \cdot \dots \cdot 2^5 = 16384.$$

Wir vereinbaren Folgendes für Summen und Produkte, für die es keine Variablenbelegungen gibt, die die einschränkende Aussage wahr macht (hier ist t ein beliebiger Term):

$$\sum_{\mathbf{f}} t := 0 \quad \text{und} \quad \prod_{\mathbf{f}} t := 1.$$

Neben Summen- und Produktquantor werden öfters noch *Minimum-* bzw. *Maximumquantor* benötigt. Im Gegensatz zu Summen- und Produktquantoren können Minimum- und Maximumquantor sich sowohl auf Terme als auch auf Aussagen beziehen, wodurch sich allerdings auch ihre Semantik ändert.

Der Minimumquantor, angewandt auf einen Term, liefert den kleinsten Wert, den dieser Term annehmen kann. Angewandt auf eine Aussage liefert er die kleinste Variablenbelegung, die diese Aussage wahr macht. In beiden Fällen wird meist eine Aussage den möglichen Wertebereich der gebundenen Variablen einschränken. Analoges gilt für den Maximumquantor.

Beispiel 2.18 Wir betrachten einige Beispiele, bei denen die Variablen aus \mathbb{N} sind.

$$\min_{2 \leq k \leq 10} 2^k - k^2 = -1, \quad \max_{3 \leq n \leq 6} n^2 = 36.$$

Man beachte, dass beide Ausdrücke syntaktisch nicht eindeutig sind, da Minimum- und Maximumquantoren sich sowohl auf Terme als auch auf Aussagen beziehen können. Mit der Klammerung

$$\left(\min_{2 \leq k \leq 10} 2^k - k^2 \right) = -1$$

wird explizit ausgedrückt, dass -1 der kleinste Wert ist, den der Term $2^k - k^2$ im Bereich 2 bis 10 annehmen kann – der Minimum-Quantor bezieht sich also auf einen Term. Mit der Klammerung

$$\min_{2 \leq k \leq 10} (2^k - k^2 = -1)$$

bezieht sich der Minimum-Quantor auf die Aussage $2^k - k^2 = -1$. Der Wert dieses Quantor-Ausdrucks ist 3, da $k = 3$ im Bereich 2 bis 10 der kleinste Wert ist, der die Aussage $2^k - k^2 = -1$ wahr macht.

Für den Maximum-Quantor gilt Ähnliches: Der Ausdruck $\max_t 2 - t$ ist 1, und $\max_k \text{prim}(k) \wedge \text{gerade}(k)$ ist 2. \square

Mengen als Bausteine der Mathematik

In diesem Kapitel betrachten wir *Mengen* als Grundbausteine mathematischer Objekte. Dabei verzichten wir auf eine genaue Definition des Begriffs “Menge”; die bekannte sprachliche *Beschreibung* stammt von Georg Cantor, dem Begründer der Mengenlehre, aus dem 19. Jahrhundert:

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

Man könnte einen schwierigeren Weg gehen, und Mengen durch *Axiome* beschreiben. Axiome sind spezielle Aussagen, die als wahr postuliert werden. Ein striktes Einhalten dieses Ansatzes würde hier zu weit führen; wir werden dennoch an manchen Stellen auf die Axiome der Mengenlehre zurückgreifen.

Mengen, wie etwa $\{3, \{3, 4\}\}$ sind somit Zusammenfassungen von Elementen, die alle unterscheidbar sein müssen. Wir nehmen somit an, dass es auf dem Universum der Mengen ein zweistelliges Prädikat *istElementVon* gibt (abgekürzt durch \in), über das man für zwei gegebene Mengen x und y feststellen kann, ob x ein Element von y ist oder nicht. Wenn wir die Menge $\{3, \{3, 4\}\}$ mit M bezeichnen, gilt etwa $3 \in M$ und $\{3, 4\} \in M$. Da alle Objekte unseres Universums Mengen sind, müssen somit auch die Objekte 3 und 4 Mengen sein—tatsächlich kann man die natürlichen Zahlen leicht durch Mengen repräsentieren (wir verzichten aber darauf).

Man beachte, dass syntaktisch (für die Sprache der Prädikatenlogik) Mengen *Terme* sind, da sie Objekte bezeichnen. Wir werden in Abschnitt 3.1 sehen, dass Mengenterme syntaktisch nicht streng den Regeln aus Definition 2.8 folgen; wir werden zur Vermeidung weiterer Definitionen darüber hinwegsehen. Durch das \in -Prädikat kann die Elementbeziehung zwischen Mengen entschieden werden. Um zu entscheiden, ob zwei Mengen gleich sind, muss ein Gleichheitsprädikat auf Mengen definiert sein. Diese grundlegende Bestimmung ist ein Beispiel eines Axioms.

Axiom 3.1 (Extensionalitätsaxiom)

Seien x und y zwei Mengen. Dann gilt

$$x = y :\Leftrightarrow \forall z \ z \in x \Leftrightarrow z \in y.$$

Der oben erstmals verwendete Junktoren \Leftrightarrow (“genau dann wenn”) ist dabei eine Abkürzung für zwei Implikationen. Für zwei Aussagen A und B gilt

$$A \Leftrightarrow B \equiv (A \Rightarrow B) \wedge (B \Rightarrow A).$$

$A \Leftrightarrow B$ ist genau dann wahr, wenn A und B immer die gleichen Wahrheitswerte haben. Somit sind zwei Mengen genau dann gleich, wenn sie die gleichen Elemente enthalten.

Wir betrachten einige einfache Beispiele der beiden Prädikate \in und $=$.

Beispiel 3.1 Wir schreiben abkürzend $x \notin y$ für $\neg x \in y$ und $x \neq y$ für $\neg x = y$. Dann gelten die folgenden Aussagen:

$$\begin{aligned} \{x, a, b\} &= \{x, a, b, b, a\} \\ \{2, 5, \{7, 8\}\} &\neq \{2, 5, 7, 8\} \\ 2 &\in \{2, 5, 7, 8\} \\ 7 &\notin \{2, 5, \{7, 8\}\} \\ \Delta &\in \{\circ, *, \diamond, \Delta\}. \end{aligned}$$

□

Obige Beispiele sind intuitiv klar und einfach mit dem Extensionalitätsaxiom nachzurechnen. Was ist aber mit der Gleichheit

$$5 = \{5\}?$$

Intuitiv ist dies falsch, und mit dem Gleichheitsaxiom gilt diese Aussage genau dann, wenn jedes Element in 5 – das ja auch (irgendeine) Menge ist – auch ein Element in $\{5\}$ ist. Letztere Menge hat als einziges Element 5 ; somit läuft die Frage $5 = \{5\}$ darauf hinaus, ob $5 \in 5$ gilt. Diese Frage lässt sich mit unseren bisherigen Überlegungen nicht beantworten; wir bestimmen aber axiomatisch Folgendes, aus dem dann $5 \notin 5$ und daher wie gewünscht auch $5 \neq \{5\}$ folgt.

Axiom 3.2 (Regularitätsaxiom)

Keine Menge kann Element von sich selbst sein, also

$$\forall x \ x \notin x.$$

Wir wenden uns jetzt verschiedenen Möglichkeiten zu, Mengen aus Elementen zusammenzusetzen.

3.1 Mengenbildung: Mengen aus Elementen

Wie wir oben schon gesehen haben, werden Mengen aus Elementen gebildet, indem man die Elemente in Mengenklammern schreibt. Diese Art der Mengenbildung wird *explizit* genannt.

Definition 3.1 (Explizite Mengenbildung)

Seien t_1, \dots, t_n Terme. Dann bezeichnet der Term

$$\{t_1, \dots, t_n\}$$

die Menge der Objekte, die von t_1, \dots, t_n beschrieben werden.

Bei explizit gebildeten Mengen kann man das \in -Prädikat durch die Gleichheit ausdrücken: Es gilt nämlich

$$x \in \{t_1, \dots, t_n\} \Leftrightarrow x = t_1 \vee x = t_2 \vee \dots \vee x = t_n.$$

Beispiel 3.2 Es gilt

$$7 \in \{6, 7, 8, 9\},$$

da $7 = 7$ ist. Ebenso ist

$$3 \notin \{2, \{3, 4\}, 5\},$$

da weder $3 = 2$, $3 = \{3, 4\}$, noch $3 = 5$ gilt. □

Eine explizite Mengenbildung ist für Mengen mit nur endlich vielen Elementen immer möglich, schon bald aber zu aufwändig. Oft ist es komfortabler, die Menge über eine *charakterisierende Eigenschaft* festzulegen.

Definition 3.2 (Mengenbildung durch Aussage)

Sei A eine Aussage. Dann bezeichnet

$$\{x \mid A\}$$

die Menge aller Objekte, für die die Aussage A wahr ist. Die Aussage A wird dabei *charakterisierende Eigenschaft* der Menge genannt; im Normalfall kommt x in A frei vor.

Beispiel 3.3 Die Menge der geraden Zahlen ist

$$\{x \mid \text{gerade}(x)\}.$$

Die Menge aller Primzahlen zwischen 1 und 20 ist

$$\{x \mid \text{prim}(x) \wedge 1 \leq x \leq 20\},$$

wobei die Aussage $1 \leq x \leq 20$ ein dreistelliges Prädikat in Infix-Notation ist, nämlich die Abkürzung für $1 \leq x \wedge x \leq 20$. Wenn wir die Menge der natürlichen Zahlen \mathbb{N} als gegeben annehmen, sind die Quadratzahlen die Menge

$$\{x \mid \exists_{y \in \mathbb{N}} x = y^2\}. \quad \square$$

Bei Konjunktionen in der charakterisierenden Eigenschaft von Mengen, die eine Mengenzugehörigkeit enthalten (meist zu einem Zahlenbereich) hat sich eine abkürzende Schreibweise eingebürgert. So schreiben wir für die Menge

$$\{x \mid x \in \mathbb{N} \wedge x \leq 4\}$$

meist abkürzend

$$\{x \in \mathbb{N} \mid x \leq 4\}.$$

Obige Beispiele zeigen, wie man durch Aussagen leicht endliche und unendliche Mengen angeben kann. Dabei können aber auch grundsätzliche Fragen der Mengenlehre aufgeworfen werden. Die folgenden Mengen können leicht über charakterisierende Eigenschaften angegeben werden:

$$\{x \mid x \neq x\} \quad \text{und} \quad \{x \mid x \notin x\}.$$

Schwieriger ist es, die Mengen hinter diesen Aussagen zu erkennen. Das erste Beispiel ist einfacher. Da die Gleichheit reflexiv ist (also immer $x = x$ gelten muss), kann es kein x mit $x \neq x$ geben, somit enthält die Menge $\{x \mid x \neq x\}$ kein Element. Eine Menge mit dieser Eigenschaft ist durchaus sinnvoll; ihre Existenz kann aber nirgendwo abgeleitet werden, sondern muss mittels Axiom gefordert werden.

Axiom 3.3 (Existenzaxiom)

Es gibt eine Menge ohne Elemente, die wir als *leere Menge* bezeichnen und als $\{\}$ oder \emptyset schreiben.

Die Aussage dieses Axioms kann man auch als Existenzaussage schreiben:

$$\exists x \forall y \ y \notin x.$$

Es gibt noch andere Möglichkeiten, eine Menge mit keinen Elementen zu spezifizieren. Dazu muss nur die charakterisierende Eigenschaft immer falsch sein; die ist etwa auch bei der Menge $\{x \mid x \in 3 \wedge x \notin 3\}$ der Fall. Bedeutet dies nun etwa, dass es unendlich viele leere Mengen gibt? Wir werden nachrechnen, dass dies nicht der Fall ist, dass also alle leeren Mengen gleich sind. Dies ist ein Beispiel für *Schlussfolgern* im Kalkül der Prädikatenlogik.

Beispiel 3.4 Seien $\{x \mid x \neq x\}$ und $\{x \mid x \in 3 \wedge x \notin 3\}$ zwei leere Mengen. Wir zeigen, dass diese beiden Mengen gleich sind, dass also

$$\{x \mid x \neq x\} = \{x \mid x \in 3 \wedge x \notin 3\}$$

gilt. Laut Definition der Gleichheit ist dies der Fall, wenn

$$\forall z \ z \in \{x \mid x \neq x\} \Leftrightarrow z \in \{x \mid x \in 3 \wedge x \notin 3\}.$$

Dies lässt sich mit der Definition des Junktors \Leftrightarrow weiter zerlegen in

$$\forall z \ (z \in \{x \mid x \neq x\} \Rightarrow z \in \{x \mid x \in 3 \wedge x \notin 3\}) \wedge \\ (z \in \{x \mid x \in 3 \wedge x \notin 3\} \Rightarrow z \in \{x \mid x \neq x\}). \quad (*)$$

Wir wollen nun also nachrechnen, dass obige Aussage (eine Allaussage) wahr ist. Eine Allaussage ist genau dann wahr, wenn die Aussage für alle möglichen Belegungen der gebundenen Variablen wahr ist. Da dies in einer unendlich großen Domäne unendlich viele Belegungen sein können, kann man nicht jede Belegung einzeln durchprüfen. Vielmehr rechnet man für den Beweis einer Allaussage mit einer *beliebigen, aber fixen* Belegung. Wenn die Aussage dann für diese beliebige Belegung wahr ist (und damit auch für jede andere!), dann ist die Allaussage bewiesen.

Der erste Schritt im Beweis der Aussage (\star) ist also, z als beliebig aber fix anzunehmen. Dann bleibt die Konjunktion

$$(z \in \{x \mid x \neq x\} \Rightarrow z \in \{x \mid x \in 3 \wedge x \notin 3\}) \wedge \\ (z \in \{x \mid x \in 3 \wedge x \notin 3\} \Rightarrow z \in \{x \mid x \neq x\})$$

zu beweisen. Eine Konjunktion ist genau dann wahr, wenn ihre beiden Teilaussagen wahr sind. Somit müssen wir nachrechnen, dass sowohl

$$z \in \{x \mid x \neq x\} \Rightarrow z \in \{x \mid x \in 3 \wedge x \notin 3\} \quad (\text{i})$$

und

$$z \in \{x \mid x \in 3 \wedge x \notin 3\} \Rightarrow z \in \{x \mid x \neq x\} \quad (\text{ii})$$

wahr sind. Wir beginnen mit der Implikation (i). Eine Implikation ist genau dann wahr, wenn ihre Prämisse falsch oder sowohl die Prämisse als auch die Konklusion wahr sind. Bei der Implikation (i) ist die Prämisse $z \in \{x \mid x \neq x\}$ falsch, da die leere Menge $\{x \mid x \neq x\}$ kein Element enthält; damit ist die Implikation (i) wahr.

Analog dazu ist bei Implikation (ii) ebenfalls die Prämisse $z \in \{x \mid x \in 3 \wedge x \notin 3\}$ falsch und damit die ganze Implikation wahr. Wir haben somit nachgerechnet, dass beide Teile der Aussage (\star) für beliebiges aber fixes z wahr sind.

An dieser Stelle muss darauf hingewiesen werden, dass bei den meisten Fällen von Beweisen von Implikationen die Prämisse *nicht* immer falsch ist. In diesen Fällen geht man dann wie weiter unten gezeigt vor. \square

Die Mengenbildung $\{x \mid x \neq x\}$ führt somit zur (eindeutigen) leeren Menge. Schwieriger ist es bei der Menge

$$\{x \mid x \notin x\}.$$

Wir wir mit dem Regularitätsaxiom 3.2 wissen, erfüllt jede Menge die charakterisierende Eigenschaft $x \notin x$, somit ist $\{x \mid x \notin x\}$ die "Menge aller Mengen." Die Definition dieser Menge birgt aber ein subtiles Problem in sich.

Wenn wir diese Menge mit R bezeichnen, können wir die Frage stellen: Ist $R \in R$? Wenn ja, dann würde dies der charakterisierenden Eigenschaft widersprechen, da ja $R \notin R$ sein müsste. Wenn nein, dann erfüllt R die charakterisierende Eigenschaft, und müsste daher in R sein. Die Frage, ob $R \in R$ gilt, führt somit in jedem Fall zu einem Widerspruch – dies wird das *Russellsche Paradoxon* genannt. Um diesem zu entkommen, muss man die Mengenlehre gänzlich axiomatisch begründen; dies führt zu den Axiomen von Zermelo und Frenkel, mit denen das Russellsche Paradoxon ausgeschlossen wird. In dieser axiomatischen Mengenlehre kann man allerdings nicht genauso wie in Definition 3.2 Mengen nur über charakterisierende Eigenschaften festlegen. Wir beschränken uns hier darauf, aus pragmatischen Gründen weiterhin die sogenannte *naive Mengenlehre* zu betreiben, da alle uns interessierenden

Sachverhalte darin ohne Schwierigkeiten und Paradoxa ausgedrückt werden können. Das Russellsche Paradoxon umgehen wir, indem wir bestimmen, dass die Menge $\{x \mid x \notin x\}$ nicht definiert ist und es daher eine "Menge aller Mengen" nicht geben kann.

Nach diesem kurzen Ausflug in die Grundlagen der Mengenlehre wenden wir uns einer weiteren Möglichkeit zu, Mengen beschreibend anzugeben.

Definition 3.3 (Mengenbildung durch Term und Aussage)

Sei A eine Aussage und t_x ein Term, in dem die Variable x vorkommt. Dann bezeichnet

$$\{t_x \mid A\}$$

die Menge aller durch t_x spezifizierten Objekte, wobei als Variablenbelegungen all jene genommen werden, die die Aussage A erfüllen.

Beispiel 3.5 Die Menge der geraden Zahlen lässt sich, wie wir bereits in Beispiel 3.3 gesehen haben, als $\{x \mid \text{gerade}(x)\}$ schreiben. Wir können diese Menge auch über einen Term und eine Aussage als

$$\{2x \mid x \in \mathbb{N}\}$$

schreiben. Ebenso ist die Menge

$$\{x^2 + x + 41 \mid x \in \mathbb{N} \wedge x < 40\}$$

eine Menge von 39 Primzahlen. □

3.2 Mengenbildung: Mengen aus Mengen

Im letzten Abschnitt haben wir drei verschiedene Ansätze betrachtet, wie Mengen aus Elementen aufzubauen sind. In diesem Abschnitt konstruieren wir Mengen aus den Elementen gegebener Mengen.

Die einfachste Menge, die auf diese Art gebildet werden kann, ist die *Teilmenge*.

Definition 3.4 (Teilmenge)

Sei x eine Menge. Dann bezeichnet man jede Menge y , deren Elemente auch Elemente von x sind, als *Teilmenge* von x und schreibt dafür $y \subseteq x$. Wir definieren also

$$y \subseteq x :\Leftrightarrow \forall z \ z \in y \Rightarrow z \in x.$$

Beispiel 3.6 Sei x die Menge $\{2, 3, \{4, \{5\}\}, 6\}$. Mit der Abkürzung $x \not\subseteq y$ für $\neg x \subseteq y$ gilt dann

$$\begin{aligned} \{2, 3\} &\subseteq x \\ \{\{4, \{5\}\}\} &\subseteq x \\ \{4, \{5\}\} &\not\subseteq x \\ x &\subseteq \{1, 2, 3, \{4, \{5\}\}, 6, 7\}. \end{aligned} \quad \square$$

Ausserdem gilt

$$\forall y \emptyset \subseteq y.$$

Dies kann man durch Einsetzen in die Definition der Teilmenge nachrechnen.

Mit dem Teilmengenbegriff kann man die Gleichheit von Mengen auch anders formulieren.

Satz 3.1 Seien x und y zwei Mengen. Dann gilt

$$x = y \Leftrightarrow (x \subseteq y) \wedge (y \subseteq x).$$

Man kann nachrechnen, dass dieser Satz aus Axiom 3.1 (Extensionalitätsaxiom) und Definition 3.4 (Teilmenge) folgt.

Weitere elementare Möglichkeiten, Mengen aus Mengen zu konstruieren, sind *Vereinigung*, *Durchschnitt*, und *Komplement*. Diese sind weiter unten definiert. Da diese Operationen *Funktionen* sind, die aus einem oder mehreren Objekten weitere Objekte bestimmen, benötigen wir analog zur Schreibweise für Prädikatendefinitionen in Definition 2.15 eine Schreibweise für Funktionsdefinitionen. Wir legen diese Definition allgemeiner für Terme (und damit auch speziell für Funktionen) aus.

Definition 3.5 (Explizite Termdefinition)

Die *explizite Definition* eines Terms führt eine neue Funktionskonstante bzw. Variable oder Variablenbelegung ein. Syntaktisch schreiben wir die Definition einer Variablen x als

$$x := t,$$

wobei t ein Term ist. Für eine n -stellige Funktionskonstante f schreiben wir

$$f(x_1, \dots, x_n) := t_{x_1, \dots, x_n},$$

wobei t_{x_1, \dots, x_n} ein Term mit Variablen x_1, \dots, x_n ist.

Die folgenden Mengenbildungen sind die ersten Anwendungen von Funktionsdefinitionen.

Definition 3.6 (Vereinigung, Durchschnitt, Komplement)

Seien x und y Mengen. Dann ist die *Vereinigung* $x \cup y$ von x und y definiert als

$$x \cup y := \{z \mid z \in x \vee z \in y\},$$

der *Durchschnitt* $x \cap y$ von x und y als

$$x \cap y := \{z \mid z \in x \wedge z \in y\},$$

und das *Komplement* $x \setminus y$ von y in x als

$$x \setminus y := \{z \mid z \in x \wedge z \notin y\}.$$

Man nennt x und y *disjunkt*, wenn $x \cap y = \emptyset$ gilt.

Somit sind in der Vereinigung zweier Mengen alle Elemente, die in einer der beiden Mengen enthalten sind; im Durchschnitt diejenigen, die in beiden Mengen sind; und im Komplement diejenigen der ersten Menge, die nicht in der zweiten sind.

Beispiel 3.7 Seien $x := \{a, b, \{c, d\}, e\}$, $y := \{c, d, e\}$, und $z := \{e, f\}$. Dann gilt

$$\begin{array}{lll} x \cup y = \{a, b, \{c, d\}, e, c, d\} & x \cap y = \{e\} & x \setminus y = \{a, b, \{c, d\}\} \\ x \cup z = \{a, b, \{c, d\}, e, f\} & x \cap z = \{e\} & x \setminus z = \{a, b, \{c, d\}\} \\ y \cup z = \{c, d, e, f\} & y \cap z = \{e\} & y \setminus z = \{c, d\}. \end{array} \quad \square$$

Mit diesen Mengenbildungsfunktionen lassen sich Zusammenhänge zwischen Mengen leicht formulieren und (gegebenfalls schwieriger) nachrechnen. Wir betrachten dazu ein Beispiel.

Beispiel 3.8 Wir wollen nachrechnen, dass die Formel

$$\forall x \forall y \quad x \cup y = (x \cap y) \cup (x \setminus y) \cup (y \setminus x)$$

wahr ist. Wie wir schon wissen, beweist man eine Allaussage, indem man die quantifizierten Variablen beliebig aber fix nimmt. Für beliebig aber fixe x und y bleibt dann eine Mengengleichheit zu beweisen, die wir unter Verwendung von Satz 3.1 in eine Konjunktion zerlegen können, deren Teile

$$x \cup y \subseteq (x \cap y) \cup (x \setminus y) \cup (y \setminus x) \tag{i}$$

und

$$(x \cap y) \cup (x \setminus y) \cup (y \setminus x) \subseteq x \cup y \tag{ii}$$

sind. Wir beschränken uns hier darauf, Teil (i) nachzurechnen. Mit der Definition der Teilmenge in Definition 3.4 gilt diese Aussage genau dann, wenn auch

$$\forall z \quad z \in x \cup y \Rightarrow z \in (x \cap y) \cup (x \setminus y) \cup (y \setminus x)$$

gilt. Wiederum nehmen wir z beliebig aber fix und beweisen

$$z \in x \cup y \Rightarrow z \in (x \cap y) \cup (x \setminus y) \cup (y \setminus x).$$

Für den Beweis einer Implikation muss entweder die Prämisse immer falsch sein (was hier – wie in den meisten Fällen – nicht stimmt), oder die Konklusion muss wahr sein, wenn auch die Prämisse wahr ist. Wir nehmen also an, dass die Prämisse wahr ist (dass also $z \in x \cup y$ gilt) und müssen nun nachrechnen, dass auch

$$z \in (x \cap y) \cup (x \setminus y) \cup (y \setminus x).$$

An diesem Beispiel zeigt sich gut die Struktur eines mathematischen Beweises: Unter Verwendung von Definitionen und Sätzen wird eine Aussage immer weiter zerlegt, bis die kleinsten Teile “einfach” nachzuprüfen sind. Diese Teile sind die *zu zeigenden* Teile eines Beweises. Sie werden bewiesen, indem man auf andere wahre Aussagen zurückgreift: Entweder sind dies schon bewiesene Sätze (das *mathematische Grundwissen*), oder Annahmen, die man im Laufe eines Beweises trifft – wie in diesem Beispiel geschehen.

Wir fassen die bisherigen Umformungen folgendermaßen zusammen. Dabei unterscheiden wir genau zwischen getroffenen Annahmen und noch zu zeigenden Teilen:

$$\begin{array}{ll} \text{wir wissen} & z \in x \cup y \\ \text{zu zeigen} & z \in (x \cap y) \cup (x \setminus y) \cup (y \setminus x). \end{array}$$

Nun müssen wir noch die Mengenbildungsfunktionen mit Definition 3.6 auflösen. Laut Definition gilt

$$(z \in x \cup y) \Leftrightarrow (z \in x \vee z \in y);$$

damit reduziert sich der Beweis auf

$$\begin{array}{ll} \text{wir wissen} & z \in x \vee z \in y \\ \text{zu zeigen} & z \in (x \cap y) \vee z \in (x \setminus y) \vee z \in (y \setminus x). \end{array}$$

Für das Auflösen von Durchschnitt und Komplement geht man ähnlich vor und erhält dann

$$\begin{array}{ll} \text{wir wissen} & z \in x \vee z \in y \\ \text{zu zeigen} & (z \in x \wedge z \in y) \vee (z \in x \wedge z \notin y) \vee (z \in y \wedge z \notin x). \end{array}$$

Wir haben damit den Beweis aus der Mengenlehre in die Prädikatenlogik übersetzt. Aussagen wie in der *zu zeigen* Zeile haben wir in Abschnitt 2.1 schon vereinfacht. Hier formen wir zuerst nur die erste Disjunktion um; mit dem Distributivgesetz erhält man

$$\begin{aligned} (z \in x \wedge z \in y) \vee (z \in x \wedge z \notin y) &\equiv (z \in x) \wedge (z \in y \vee z \notin y) \\ &\equiv (z \in x) \wedge \mathbf{w} \\ &\equiv (z \in x). \end{aligned}$$

Zusammen mit der oben weggelassenen Disjunktion erhalten wir

$$\begin{aligned}(z \in x) \vee (z \in y \wedge z \notin x) &\equiv (z \in x \vee z \in y) \wedge (z \in x \vee z \notin x) \\ &\equiv (z \in x \vee z \in y) \wedge \mathbf{w} \\ &\equiv (z \in x \vee z \in y).\end{aligned}$$

Am Ende dieser Bemühungen ist somit zu beweisen:

$$\begin{array}{ll}\text{wir wissen} & z \in x \vee z \in y \\ \text{zu zeigen} & z \in x \vee z \in y.\end{array}$$

Wir sehen sofort, dass diese Schlussfolgerung stimmt, da beide Aussagen identisch sind. \square

Mit etwas Übung und Erfahrung kann man einige dieser Schritte schneller abarbeiten; die allgemeine Struktur von mathematischen Beweisen ist an diesem Beispiel aber gut abzulesen.

Eine weitere wichtige Mengenbildungsoperation ist die der *Potenzmenge*. Diese ist wie folgt definiert.

Definition 3.7 (Potenzmenge)

Sei x eine Menge. Dann bezeichnen wir mit

$$\text{Pot}(x) := \{y \mid y \subseteq x\}$$

die Menge aller Teilmengen von x .

Beispiel 3.9 Sei $x := \{4, a, \{b\}\}$. Dann ist

$$\text{Pot}(x) = \{\emptyset, \{4\}, \{a\}, \{\{b\}\}, \{4, a\}, \{4, \{b\}\}, \{a, \{b\}\}, \{4, a, \{b\}\}\}. \quad \square$$

Da $x \subseteq x$ ist, gilt immer $x \in \text{Pot}(x)$. Da die leere Menge eine Teilmenge jeder Menge ist, gilt für jede Menge x sowohl $\emptyset \subseteq \text{Pot}(x)$ als auch $\emptyset \in \text{Pot}(x)$.

3.3 Mengenbildung: Tupel

In vielen Anwendungsbereichen möchte man mehrere Objekte zu einem Objekt zusammenfassen, wobei

- mehrere gleiche Objekte zugelassen sein sollen und
- die Reihenfolge der Objekte eine Rolle spielen soll.

Beide Anforderungen werden von Mengen nicht erfüllt, von *Tupeln* hingegen schon. Diese sind wie folgt definiert.

Definition 3.8 (Tupel)

Seien t_1, \dots, t_n Terme. Dann nennt man den Term

$$(t_1, \dots, t_n)$$

ein n -Tupel. Zusätzlich sind n Projektionen auf die einzelnen Komponenten definiert, die in Postfix-Notation geschrieben werden:

$$(t_1, \dots, t_n)_i := t_i.$$

Für kleine n haben n -Tupel spezielle Namen: So bezeichnet man ein 2-Tupel als *Paar* und ein 3-Tupel als *Tripel*.

Man beachte, dass obige Definition das Symbol \cdot_i einmal als Funktionskonstante in Postfix-Notation verwendet (links), und einmal als Teil eines Variablennamens (rechts).

Tupel sind keine neuen Objekte *neben* den Mengen, sondern Kurzschreibweisen für spezielle Mengen – wir gehen auf die spezielle Art, Tupel als Mengen zu schreiben aber nicht näher ein. Somit verlassen wir durch die Definition von Tupeln nicht das Universum der Mengen.

Auch auf Tupeln ist eine Gleichheit definiert; diese ist komponentenweise gegeben.

Definition 3.9 (Gleichheit von Tupeln)

Seien x und y zwei n -Tupel. Dann gilt

$$x = y :\Leftrightarrow \forall_{1 \leq i \leq n} x_i = y_i.$$

Beispiel 3.10 Mit obigen Definitionen gilt:

$$\begin{aligned} (3, 4, 2)_2 &= 4 \\ ((1, 2), (3, 4), 3, 4)_1 &= (1, 2) \\ (7, 8, a) &= (7, 8, a) \\ (7, a, 8) &\neq (7, 8, a). \end{aligned}$$

□

Die Menge aller Tupel, die aus gegebenen Mengen gebildet werden können, nennt man *kartesisches Produkt* dieser Mengen.

Definition 3.10 (Kartesisches Produkt)

Seien x und y Mengen. Dann nennt man die Menge

$$x \times y := \{(a, b) \mid a \in x \wedge b \in y\}$$

das *kartesische Produkt* oder die *Produktmenge* von x und y . Statt $x \times x$ schreibt man abkürzend auch x^2 .

Obige Definition von Mengen von Paaren lässt sich auf beliebige n -Tupel erweitern. Man definiert für Mengen x_1, \dots, x_n :

$$x_1 \times x_2 \times \dots \times x_n := \{(a_1, \dots, a_n) \mid a_1 \in x_1 \wedge \dots \wedge a_n \in x_n\}.$$

Beispiel 3.11 Seien $x := \{2, 4, \{4, 5\}\}$, $y := \{1, 2\}$ und $z := \{a, b\}$. Dann gilt

$$\begin{aligned} x \times y &= \{(2, 1), (2, 2), (4, 1), (4, 2), (\{4, 5\}, 1), (\{4, 5\}, 2)\} \\ x \times z &= \{(2, a), (2, b), (4, a), (4, b), (\{4, 5\}, a), (\{4, 5\}, b)\} \\ y \times z &= \{(1, a), (1, b), (2, a), (2, b)\} \\ x \times y \times z &= \{(2, 1, a), (2, 2, a), (4, 1, a), (4, 2, a), (\{4, 5\}, 1, a), (\{4, 5\}, 2, a), \\ &\quad (2, 1, b), (2, 2, b), (4, 1, b), (4, 2, b), (\{4, 5\}, 1, b), (\{4, 5\}, 2, b)\}. \end{aligned}$$

□

3.4 Prädikate als Mengen

In Abschnitt 3.1 haben wir gesehen, wie wir mittels einer charakterisierenden Eigenschaft (also eines Prädikats) eine Menge bestimmen können: Die Menge umfasst alle Objekte, für die ein Prädikat wahr ist. In diesem Abschnitt gehen wir den umgekehrten Weg: wir untersuchen, wie wir die Elemente einer (speziellen) Menge als all die Objekte betrachten können, für die ein Prädikat wahr ist.

Diese Vorgehensweise erlaubt es uns, einen (bereits in Kapitel 2 erwähnten) Nachteil der Prädikatenlogik als Sprache der Mathematik zu umgehen: In der Prädikatenlogik kann nur über Objekte des Universums quantifiziert werden, nicht aber über Prädikate und Funktionen. Wenn wir nun Prädikate und Funktionen als Objekte (in diesem Fall Mengen) repräsentieren können, umgehen wir diese Einschränkung.

Wir beginnen damit, Prädikate als Mengen darzustellen, da sich die Funktionsdarstellung einfach daraus ergibt. Für einstellige Prädikate ist dies besonders einfach: Für gegebene Menge M seien die Elemente genau die Objekte, für die das Prädikat erfüllt ist. Wir definieren somit eine neue Prädikatenkonstante P über

$$P(x) :\Leftrightarrow x \in M.$$

Dieser Ansatz lässt sich auf n -stellige Prädikate erweitern, da wir jetzt den "Datentyp" Tupel zur Verfügung haben, um mehrere Objekte strukturiert zusammenzufassen. Eine neue n -stellige Prädikatenkonstante ist dann auf genau jenen Tupeln wahr, die Elemente einer gegebenen Menge von Tupeln sind. Wir bezeichnen solche Mengen von Tupeln als *Relationen*.

Definition 3.11 (Relation)

Seien A und B Mengen. Dann bezeichnet man jede Teilmenge R des kartesischen Produkts $A \times B$ als (zweistellige) *Relation*. Für $A=B$ sagt man, dass R eine *Relation auf A* ist.

Man kann diese Definition auf das kartesische Produkt von n Mengen erweitern; in diesem Fall spricht man von n -stelligen Relationen. Relationen sind somit nichts anderes als die Repräsentation von Prädikaten in der Mengenlehre. Wenn wir eine zweistellige Relation $R \subseteq A \times B$ gegeben haben, können wir eine zweistellige Prädikatenkonstante P definieren als

$$P(x, y) :\Leftrightarrow (x, y) \in R.$$

Analog definiert man n -stellige Prädikate durch n -stellige Relationen. Wir betrachten einige Beispiele.

Beispiel 3.12 Gegeben sei die Menge L aller Länder Europas und die Menge H ihrer Hauptstädte. Dann ist durch die Relation

$$R \subseteq H \times L = \{\dots, (\text{Rom, Italien}), (\text{Paris, Frankreich}), \dots\}$$

das Prädikat *istHauptstadtVon* definiert:

$$\text{istHauptstadtVon}(x, y) :\Leftrightarrow (x, y) \in R.$$

Sei $A = \{1, 2\}$, $B = \{3, 4\}$ und $C = \{5, 6\}$. Die Relation R sei definiert als

$$R \subseteq A \times B \times C = \{(1, 4, 5), (1, 4, 6), (2, 3, 5)\}.$$

Dann sind mit der Prädikatenkonstante P , die als

$$P(x, y, z) :\Leftrightarrow (x, y, z) \in R$$

definiert ist, genau folgende Aussagen wahr: $P(1, 4, 5)$, $P(1, 4, 6)$, $P(2, 3, 5)$. \square

Da wir somit Relationen und Prädikate gleichstellen können, schreiben wir Relationssymbole oft in der bisher Prädikatenkonstanten vorbehaltenen Notation. Sei für gegebenes R ein P über $P(x, y) :\Leftrightarrow (x, y) \in R$ definiert; dann schreiben wir oft

$$\begin{aligned} xRy &\Leftrightarrow P(x, y) & (\Leftrightarrow (x, y) \in R) & \quad \text{und} \\ R(x, y) &\Leftrightarrow P(x, y) & (\Leftrightarrow (x, y) \in R). \end{aligned}$$

Beispiel 3.13 Sei R folgende auf \mathbb{N} definierte Relation:

$$R = \{(x, 3x) \mid x \in \mathbb{N}\}.$$

Dann schreiben wir $R(1, 3)$, $R(2, 6)$, \dots oder $1R3$ und $2R6$, ohne dafür ein eigenes Prädikatensymbol einführen zu müssen. Gleiches können wir für eine *teilt*-Relation machen, die wir mit “|” bezeichnen:

$$| = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid \exists_{z \in \mathbb{N}} y = z \cdot x\}.$$

Dann schreiben wir statt $(3, 6) \in |$ oder $|(3, 6)$ einfacher in Infix-Notation $3|6$. \square

Die Definition von Relationen ist nicht einschränkend: *Jede* Teilmenge eines kartesischen Produkts ist eine Relation. Viele Relationen haben aber spezielle Eigenschaften, die sie besonders charakterisieren. Sei etwa R die Relation, die über das $<$ -Prädikat auf $\mathbb{N} \times \mathbb{N}$ definiert ist:

$$R \subseteq \mathbb{N} \times \mathbb{N} = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x < y\}.$$

Aus $(x, y) \in R$ folgt mit unserem Verständnis des $<$ -Prädikats sofort, dass $(y, x) \notin R$ gelten muss (da ja nicht gleichzeitig $x < y$ und $y < x$ gelten kann). Andererseits wissen wir ebenfalls, dass aus $(x, y) \in R$ und $(y, z) \in R$ auch $(x, z) \in R$ folgen muss.

Eigenschaften von Relationen sind von speziellem Interesse, da sie die konkreten Relationen *abstrahieren*. Wenn wir wissen, dass bestimmte Aussagen für Relationen gelten, die bestimmte Eigenschaften erfüllen, dann müssen wir diese Aussagen nicht mehr an konkreten Relationen nachprüfen. Wir werden im Folgenden die wichtigsten Eigenschaften von Relationen untersuchen.

Definition 3.12 (reflexiv, symmetrisch, transitiv)

Sei R eine Relation auf einer Menge M . Dann nennt man R

$$\text{reflexiv} :\Leftrightarrow \forall_{x \in M} (x, x) \in R$$

$$\text{symmetrisch} :\Leftrightarrow \forall_{x, y \in M} (x, y) \in R \Rightarrow (y, x) \in R$$

$$\text{transitiv} :\Leftrightarrow \forall_{x, y, z \in M} (x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R.$$

Wir haben obige Eigenschaften bereits bei der Definition der Gleichheit kennengelernt.

Man beachte, dass es erst durch den Umweg über Relationen (also Mengen) möglich ist, eine Definition wie oben anzuschreiben. *Reflexiv, symmetrisch* und *transitiv* sind nämlich einstellige Prädikatenkonstanten, die eine Aussage über andere Prädikate (bzw. Relationen) machen. Da wir in der Prädikatenlogik nur Aussagen über Objekte machen können, mussten wir zuerst Prädikate als Objekte (nämlich Relationen) repräsentieren.

Eigenschaften von Relationen lassen sich über geeignete Visualisierungen veranschaulichen. Eine mögliche Visualisierung für allgemeine Relationen $R \subseteq A \times B$ ist dadurch gegeben, dass man A und B auf den beiden Achsen eines rechtwinkligen Koordinatensystems aufträgt. Ein Element $(x, y) \in R$ wird dann durch einen Punkt im Koordinatensystem repräsentiert; die Menge all dieser Punkte ist genau die Relation R .

Beispiel 3.14 Gegeben sei die *teilt*-Relation auf \mathbb{N} :

$$| := \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid \exists_{z \in \mathbb{N}} y = z \cdot x\}.$$

Eine mögliche graphische Repräsentation dieser Relation ist in Abbildung 3.1 zu sehen. \square

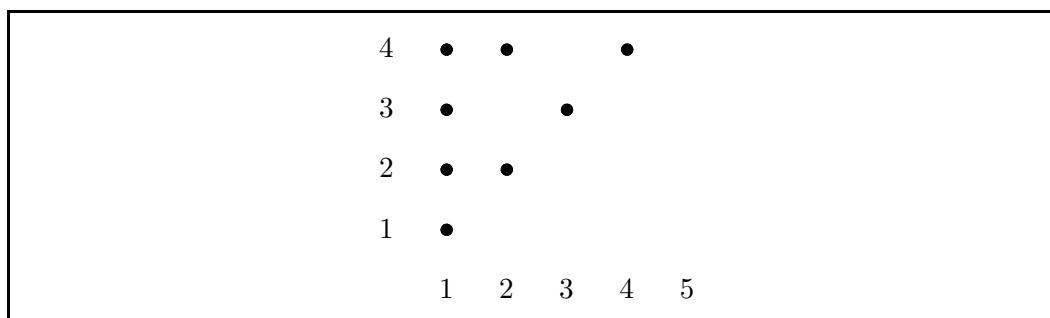


Abbildung 3.1: Graph der Relation aus Beispiel 3.14.

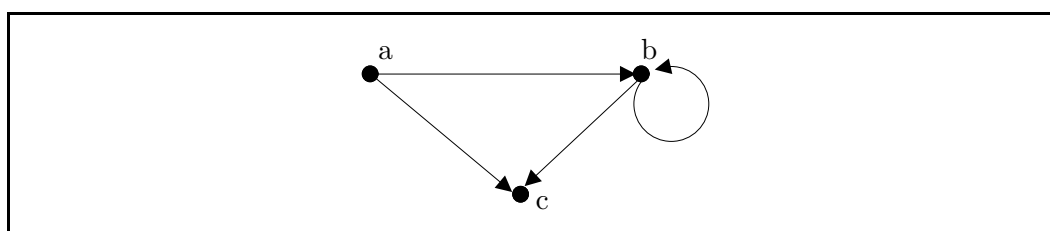


Abbildung 3.2: Graph der Relation aus Beispiel 3.15.

Relationen auf endlichen Mengen kann man mittels Punkten und Pfeilen visualisieren. Dabei gibt es genau dann einen Pfeil von x nach y , wenn das Paar (x, y) in der Relation ist.

Beispiel 3.15 Sei dazu etwa $M := \{a, b, c\}$ und

$$R := \{(a, c), (a, b), (b, b), (b, c)\}.$$

Dann ist diese Relation durch den Graphen repräsentiert, der in Abbildung 3.2 dargestellt ist.

Die Visualisierung erlaubt es uns, schnell die Eigenschaften *reflexiv*, *symmetrisch* und *transitiv* zu überprüfen: Eine Relation ist genau dann reflexiv, wenn es von jedem Punkt einen Pfeil zu sich selbst gibt; sie ist dann symmetrisch, wenn es zu jedem Pfeil einen Pfeil in der Gegenrichtung gibt; und sie ist transitiv, wenn es zu jeden zwei hintereinanderfolgenden Pfeilen einen Pfeil gibt, der diese zwei Pfeile “abkürzt”.

Mit diesen Überlegungen ist ersichtlich, dass obige Relation weder reflexiv noch symmetrisch, dafür aber transitiv ist: Es fehlen etwa der Pfeil von a nach a (daher nicht reflexiv) und die Umkehrung des Pfeils von a nach b (daher nicht symmetrisch). Es ist aber für jede Sequenz von zwei Pfeilen (etwa a nach b und b nach c) ein Pfeil vorhanden, der diese Sequenz abkürzt (hier von a nach c). \square

Beispiel 3.16 Sei $M := \{1, 2, 3\}$ und eine Relation auf M ; diese Relation ist gegeben durch den Graphen, der in Abbildung 3.3 zu sehen ist.

Die dabei dargestellte Relation ist

$$R := \{(1, 2), (3, 3), (2, 2), (2, 1)\}.$$

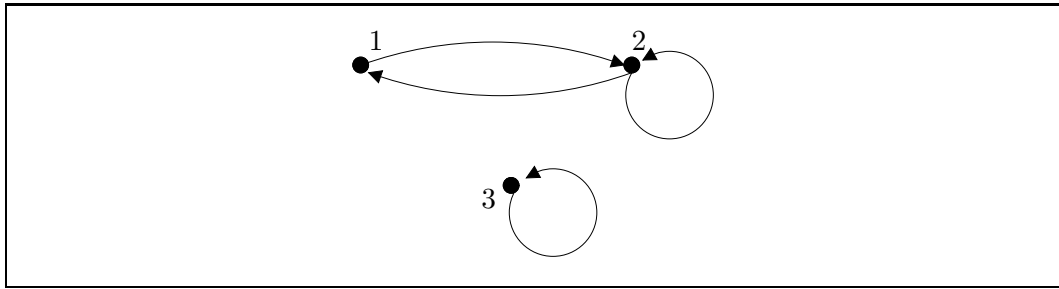


Abbildung 3.3: Graph der Relation aus Beispiel 3.16.

Diese Relation ist nicht reflexiv, da $(1, 1) \notin R$ ist. R ist symmetrisch, da alle Pfeile umkehrbar sind, aber nicht transitiv, weil $(1, 2)$ und $(2, 1)$ in R sind, nicht aber $(1, 1)$. □

Die folgenden Eigenschaften von Relationen sind ebenfalls von Interesse.

Definition 3.13 (irreflexiv, asymmetrisch, antisymmetrisch)

Sei R eine Relation auf einer Menge M . Dann nennt man R

$$irreflexiv \Leftrightarrow \forall_{x \in M} (x, x) \notin R$$

$$asymmetrisch \Leftrightarrow \forall_{x, y \in M} (x, y) \in R \Rightarrow (y, x) \notin R$$

$$antisymmetrisch \Leftrightarrow \forall_{x, y \in M} ((x, y) \in R \wedge (y, x) \in R) \Rightarrow x = y.$$

Beispiel 3.17 Wir betrachten nochmals die Relationen aus den Beispielen 3.14 und 3.16 sowie die \subseteq -Relation auf der Potenzmenge $\text{Pot}(A)$ einer beliebigen Menge A .

$$R_1 := \{(a, c), (a, b), (b, b), (b, c)\}$$

$$R_2 := \{(1, 2), (3, 3), (2, 2), (2, 1)\}$$

$$\subseteq := \{(x, y) \in \text{Pot}(A) \times \text{Pot}(A) \mid x \subseteq y\}$$

Folgende Tabelle fasst zusammen, welche der drei Eigenschaften *irreflexiv*, *asymmetrisch* und *antisymmetrisch* diese Relationen erfüllen:

	\subseteq	R_1	R_2
irreflexiv			
asymmetrisch			
antisymmetrisch	✓	✓	

Wir rechnen die Korrektheit der Einträge in der Teilmengen-Spalte durch: Allgemein gilt für eine Relation R auf M :

R ist nicht irreflexiv

$$\Leftrightarrow \neg \forall_{x \in M} (x, x) \notin R \Leftrightarrow \exists_{x \in M} \neg(x, x) \notin R \Leftrightarrow \exists_{x \in M} (x, x) \in R;$$

dies ist für die Teilmengen-Relation erfüllt, da etwa $A \subseteq A$ gilt. Ebenso gilt

$$\begin{aligned} R \text{ ist nicht asymmetrisch} &\Leftrightarrow \neg \forall_{x,y \in M} (x,y) \in R \Rightarrow (y,x) \notin R \\ &\Leftrightarrow \exists_{x,y \in M} \neg((x,y) \in R \Rightarrow (y,x) \notin R) \\ &\Leftrightarrow \exists_{x,y \in M} (x,y) \in R \wedge \neg(y,x) \notin R \\ &\Leftrightarrow \exists_{x,y \in M} (x,y) \in R \wedge (y,x) \in R. \end{aligned}$$

Diese Eigenschaft ist für die Teilmengen-Relation immer dann erfüllt, wenn $x = y$ ist. Wir sehen: Wenn wir zeigen wollen, dass eine über einen Allquantor definierte Eigenschaft *nicht* erfüllt ist genügt es, ein Gegenbeispiel zu dieser Definition anzugeben (da durch die Negation der Allquantor in einen Existenzquantor umgewandelt wird).

Das *Nachprüfen* einer Definition ist schwieriger. Um zu zeigen, dass die Teilmengen-Relation antisymmetrisch ist, müssen wir also nachrechnen:

$$\subseteq \text{ ist antisymmetrisch} \Leftrightarrow \forall_{x,y \in \mathbb{N}} (x \subseteq y \wedge y \subseteq x) \Rightarrow x = y.$$

Wir wählen x und y beliebig aber fix und müssen nun noch zeigen, dass für diese beliebig aber fixen Variablen gilt:

$$(x \subseteq y \wedge y \subseteq x) \Rightarrow x = y.$$

Dies ist aber genau die Aussage von Satz 3.1 über die Gleichheit zweier Mengen (von rechts nach links gelesen). Damit ist die \subseteq -Relation antisymmetrisch. \square

Im Bereich der Zahlen sind besonders Relationen von Interesse, die mehrere der Eigenschaften aus den Definitionen 3.12 und 3.13 erfüllen. Dabei nehmen die *Äquivalenzrelationen* eine herausragende Stellung ein.

Definition 3.14 (Äquivalenzrelation)

Eine Relation R auf einer Menge A , die sowohl reflexiv, symmetrisch und transitiv ist, nennt man *Äquivalenzrelation*. Zwei Elemente $x, y \in A$ mit $(x, y) \in R$ nennt man *äquivalent*.

Wie wir aus Definition 2.13 wissen, erfüllt die Gleichheitsrelation die Bedingung, reflexiv, symmetrisch und transitiv zu sein, und ist somit eine Äquivalenzrelation. Als Äquivalenzrelation ist die Gleichheit nicht von besonderem Interesse, da ja immer nur ein Objekt zu sich selbst äquivalent ist. Das nächste Beispiel zeigt eine Äquivalenzrelation, bei denen unendlich viele Objekte zueinander äquivalent sind. Dazu benötigen wir noch folgende Aussage über die Division mit Rest.

Satz 3.2 Seien $x, y \in \mathbb{Z}$. Dann gibt es eindeutig bestimmte Zahlen $q, r \in \mathbb{Z}$ mit $0 \leq r < |y|$, sodass gilt

$$x = y \cdot q + r.$$

Wir bezeichnen q als den *Quotienten*, und r als den *Rest* bei der Division von x und y .

Zum Beweis dieses Satzes kann man den (als schon bekannt vorausgesetzten) Algorithmus zur Division mit Rest angeben; wir verzichten hier darauf, gehen aber in Abschnitt 4.2 nochmals darauf ein.

Beispiel 3.18 Wir betrachten folgende Relation auf den ganzen Zahlen:

$$R_7 := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid 7 \mid x - y\}.$$

In dieser Relation sind all jene Zahlen äquivalent, die bei Division durch 7 den gleichen Rest haben. Somit gilt etwa $(13, 27) \in R_7$ oder $(9, -5) \in R_7$. Wir rechnen nach, dass es sich bei R_7 tatsächlich um eine Äquivalenzrelation handelt. Wir behandeln im Folgenden alle Variablen als beliebig aber fix, um so auf die Allquantifizierung verzichten zu können.

reflexiv Zu zeigen ist $(x, x) \in R_7$, wir formen um:

$$\begin{aligned} (x, x) \in R_7 &\Leftrightarrow 7 \mid x - x \\ &\Leftrightarrow 7 \mid 0 \\ &\Leftrightarrow \exists_{z \in \mathbb{Z}} 0 = z \cdot 7, \end{aligned}$$

dies gilt für $z = 0$.

symmetrisch Zu zeigen ist $(x, y) \in R_7 \Rightarrow (y, x) \in R_7$, somit

$$\begin{aligned} ((x, y) \in R_7 \Rightarrow (y, x) \in R_7) &\Leftrightarrow (7 \mid x - y \Rightarrow 7 \mid y - x) \\ &\Leftrightarrow \left(\left(\exists_{s \in \mathbb{Z}} x - y = 7s \right) \Rightarrow \left(\exists_{t \in \mathbb{Z}} y - x = 7t \right) \right). \end{aligned}$$

Diese Implikation ist erfüllt, da das t mit $y - x = 7z$ aus der Konklusion der Implikation existiert: Es ist das Negative der Zahl s mit $x - y = 7s$, deren Existenz wir beim Beweis der Implikation ja annehmen.

transitiv Zu zeigen ist $((x, y) \in R_7 \wedge (y, z) \in R_7 \Rightarrow (x, z) \in R_7)$, Umformen ergibt

$$\begin{aligned} ((x, y) \in R_7 \wedge (y, z) \in R_7 \Rightarrow (x, z) \in R_7) &\Leftrightarrow (7 \mid x - y \wedge 7 \mid y - z \Rightarrow 7 \mid x - z) \\ &\Leftrightarrow \left(\left(\exists_{s \in \mathbb{Z}} x - y = 7s \wedge \left(\exists_{t \in \mathbb{Z}} y - z = 7t \right) \right) \Rightarrow \left(\exists_{u \in \mathbb{Z}} x - z = 7u \right) \right). \end{aligned}$$

Laut Prämisse der zu zeigenden Implikation ist $x - y = 7s$ und $y - z = 7t$. Wenn wir die zweite Gleichung nach y auflösen und in die erste einsetzen, erhalten wir $x - 7t - z = 7s$, also $x - z = 7(s + t)$. Das für den Beweis gesuchte u ist also $s + t$. \square

Zur Bezeichnung von Äquivalenzrelationen verwendet man häufig das Symbol \sim . Als *Äquivalenzklasse* $[x]_{\sim}$ eines Objekts x unter einer Relation \sim bezeichnet man die Menge der Objekte, die zu x äquivalent sind:

$$[x]_{\sim} := \{y \mid x \sim y\}.$$

Wenn die Relation aus dem Kontext klar ist, schreibt man oft $[x]$ statt $[x]_{\sim}$. Für die Relation R_7 aus Beispiel 3.18 gilt somit

$$\begin{aligned} [0] &= \{x \in \mathbb{Z} \mid xR_7 0\} = \{\dots, -14, -7, 0, 7, 14, 21, \dots\} \\ [1] &= \{x \in \mathbb{Z} \mid xR_7 1\} = \{\dots, -13, -6, 1, 8, 15, 22, \dots\} \\ &\vdots \\ [7] &= \{x \in \mathbb{Z} \mid xR_7 7\} = \{\dots, -14, -7, 0, 7, 14, 21, \dots\} = [0] \\ &\vdots \end{aligned}$$

Somit “teilt” die Äquivalenzrelation R_7 die ganzen Zahlen in sieben disjunkte Mengen:

$$\mathbb{Z} = [0] \cup [1] \cup [2] \cup \dots \cup [6].$$

Eine solche Aufteilung einer Menge nennt man *Partition*. Allgemein sind Partitionen wie folgt definiert.

Definition 3.15 (Partition)

Sei M eine Menge. Eine Menge $P \subseteq \text{Pot}(M)$ nennt man *Partition von M* , wenn gilt

- (1) $\forall_{x \in P} x \neq \emptyset$
- (2) $\forall_{x, y \in P} x \neq y \Rightarrow x \cap y = \emptyset$
- (3) $M = \{z \mid \exists_{x \in P} z \in x\}$.

Eine Partition ist somit eine Menge von disjunkten, nichtleeren Teilmengen einer Menge M , die zusammengenommen wieder M ergibt.

Wir haben gesehen, wie man aus einer Äquivalenzrelation eine Partition erhält. Der umgekehrte Schritt, über eine Partition eine Äquivalenzrelation zu definieren, ist ebenso möglich; wir gehen darauf nicht näher ein.

Neben den Äquivalenzrelationen spielen *Ordnungen* eine wichtige Rolle in der Mathematik. Ordnungen sind wie folgt definiert.

Definition 3.16 (Ordnung)

Eine Relation R auf einer Menge A , die sowohl reflexiv, antisymmetrisch und transitiv ist, nennt man (*partielle*) *Ordnung*. Wenn zusätzlich noch $(x, y) \in R$ oder $(y, x) \in R$ für alle $x, y \in A$ gilt, dann nennt man R eine *lineare Ordnung*.

Ordnungen unterscheiden sich von Äquivalenzrelationen also dadurch, dass erste *antisymmetrisch*, zweite aber *symmetrisch* sind. Ordnungen erlauben es uns, Elemente einer Menge miteinander zu vergleichen; bei linearen Ordnungen können zusätzlich alle Elemente miteinander verglichen werden.

Beispiel 3.19 Das Standardbeispiel einer linearen Ordnung ist die \leq -Relation auf den natürlichen Zahlen. Wir können sofort überprüfen, dass die drei geforderten Bedingungen erfüllt sind: Für alle $x \in \mathbb{N}$ gilt $x \leq x$, somit ist \leq reflexiv. Ebenso gilt $x \leq y \wedge y \leq x$ nur dann, wenn $x = y$ ist (Antisymmetrie). Aus $x \leq y \wedge y \leq z$ folgt ebenso (etwa über die Visualisierung auf einem Zahlenstrang), dass dann auch $x \leq z$ sein muss (Transitivität). Ebenfalls anschaulich klar ist, dass sich zwei natürliche Zahlen immer in der Beziehung $x \leq y$ oder $y \leq x$ befinden.

Man beachte: Dieser “Beweis”, dass auf den natürlichen Zahlen eine lineare Ordnung \leq definiert ist, appelliert an unsere Vorstellung der natürlichen Zahlen, und unterscheidet sich dadurch vom genauen Nachrechnen der Bedingungen einer Äquivalenzrelation in Beispiel 3.18. Der Unterschied ist darin begründet, dass wir für die \leq -Relation keine Definition in der Sprache der Prädikatenlogik gegeben haben – dafür würden wir die Repräsentation der natürlichen Zahlen als Mengen benötigen, auf die wir aber verzichten werden.

Eine Ausweitung dieser Definition auf Paare natürlicher Zahlen ergibt ein Beispiel einer Ordnung, die nicht linear ist. Wenn wir die Relation \leq_2 definieren als

$$(x_1, x_2) \leq_2 (y_1, y_2) :\Leftrightarrow x_1 \leq y_1 \wedge x_2 \leq y_2,$$

dann ist diese Relation reflexiv, antisymmetrisch und transitiv. Es sind aber *nicht* alle Zahlenpaare miteinander vergleichbar: So gilt

$$(1, 2) \leq_2 (3, 4), \quad \text{aber weder } (1, 4) \leq_2 (2, 3) \text{ noch } (2, 3) \leq_2 (1, 4). \quad \square$$

Beispiel 3.20 Das Standardbeispiel einer partiellen Ordnung ist die \subseteq -Relation auf der Potenzmenge einer beliebigen Menge A . Wir können die drei Bedingungen einer partiellen Ordnung über die Definition der \subseteq -Relation nachrechnen; seien $x, y, z \in \text{Pot}(A)$ beliebig aber fix.

reflexiv Zu zeigen ist $x \subseteq x$, die Gültigkeit ergibt sich aus der Umformung

$$x \subseteq x \Leftrightarrow \forall s \ s \in x \Rightarrow s \in x.$$

antisymmetrisch Wir haben dies bereits in Beispiel 3.17 nachgerechnet.

transitiv Zu zeigen ist $(x \subseteq y \wedge y \subseteq z) \Rightarrow x \subseteq z$. Über die Definition von \subseteq ist dies äquivalent zu

$$(\forall s \ s \in x \Rightarrow s \in y) \wedge (\forall t \ t \in y \Rightarrow t \in z) \Rightarrow (\forall u \ u \in x \Rightarrow u \in z).$$

Zum Beweis dieser Implikation nehmen wir die Prämisse als wahr an, und zeigen die Konklusion (die Allaussage $\forall u \ u \in x \Rightarrow u \in z$). Sei also u beliebig aber fix, der Beweis reduziert sich dann auf Folgendes:

$$\begin{array}{ll} \text{wir wissen} & (\forall s \ s \in x \Rightarrow s \in y) \wedge (\forall t \ t \in y \Rightarrow t \in z) \\ \text{zu zeigen} & u \in x \Rightarrow u \in z. \end{array}$$

Die zu zeigende Aussage ist wiederum eine Implikation, deren Prämisse wir als wahr annehmen. Damit erhalten wir

$$\begin{array}{ll} \text{wir wissen} & (\forall s \ s \in x \Rightarrow s \in y) \wedge (\forall t \ t \in y \Rightarrow t \in z) \wedge u \in x \\ \text{zu zeigen} & u \in z. \end{array}$$

Die beiden Implikationen im “wir wissen”-Teil sind *für alle Variablenbelegungen* wahr, gelten somit *speziell auch* für das beliebig aber fixe u . Wir erhalten damit die Beweisstruktur

$$\begin{array}{ll} \text{wir wissen} & (u \in x \Rightarrow u \in y) \wedge (u \in y \Rightarrow u \in z) \wedge u \in x \\ \text{zu zeigen} & u \in z. \end{array}$$

Nun wenden wir zweimal die Schlussregel modus ponens an: Aus $u \in x$ und $u \in x \Rightarrow u \in y$ folgt $u \in y$; aus $u \in y$ und $u \in y \Rightarrow u \in z$ folgt $u \in z$. Damit ist der Beweis abgeschlossen, da dies auch die zu zeigende Aussage ist.

Somit haben wir nachgerechnet, dass die \subseteq -Relation eine partielle Ordnung ist. Sie ist aber keine lineare Ordnung, wie das folgende Gegenbeispiel zeigt: Sei dazu A eine beliebige Menge mit mindestens zwei Elementen, die wir als x und y bezeichnen. Dann gilt für die beiden Elemente $\{x\}$ und $\{y\}$ von $\text{Pot}(A)$:

$$\{x\} \not\subseteq \{y\} \wedge \{y\} \not\subseteq \{x\}, \text{ also } \neg(\{x\} \subseteq \{y\} \vee \{y\} \subseteq \{x\}). \quad \square$$

3.5 Funktionen als Mengen

Wie schon bei Relationen ist es in diesem Abschnitt unser Ziel, Funktionen als Objekte der Mengenlehre darzustellen. Wir erinnern uns: Funktionskonstante der Prädikatenlogik weisen einem Objekt ein eindeutig bestimmtes Objekt zu. Eine Funktion lässt sich somit in der Mengenlehre als eine Menge von Paaren beschreiben, wobei die erste Komponente des Paares das Argument der Funktion darstellt, und die zweite Komponente den Wert der Funktion mit diesem Argument. Zu beachten ist, dass bei Funktionen jedes Argument nur einen Wert liefern darf; diese Einschränkung unterscheidet Funktionen von Relationen. Funktionen in der Mengenlehre sind somit nichts anderes als spezielle Relationen; genau sind Funktionen wie folgt definiert.

Definition 3.17 (Funktion)

Seien A und B zwei Mengen. Dann nennt man eine Relation f auf $A \times B$ eine (partielle) *Funktion von A nach B* , wenn gilt:

$$\forall_{x \in A, y_1, y_2 \in B} ((x, y_1) \in f \wedge (x, y_2) \in f) \Rightarrow y_1 = y_2.$$

Man bezeichnet A als *Definitionsbereich* und B als *Bildbereich* von f . Eine Funktion heißt *total*, wenn *jedem* Element aus A ein Element aus B zugewiesen wird. Für eine Funktion von A nach B schreiben wir kurz $f : A \rightarrow B$.

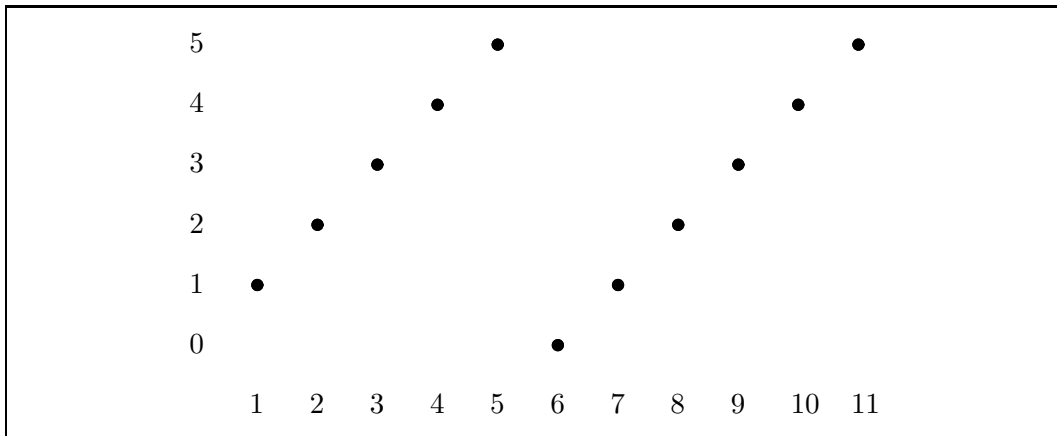


Abbildung 3.4: Graph der Funktion aus Beispiel 3.22.

Beispiel 3.21 Sei $A = \{1, 2, 3, 4\}$ und $B = \{a, b\}$. Dann ist die Menge

$$\{(1, a), (2, a), (3, b)\}$$

eine partielle Funktion von A nach B . Diese Funktion ist nicht total, da dem Element 4 von A kein Element in B zugewiesen wird. Die Menge

$$\{(1, a), (2, a), (3, b), (1, b), (4, a)\}$$

ist keine Funktion, da dem Element 1 zwei verschiedene Werte (a und b) zugewiesen werden. \square

In den meisten Fällen ist es möglich, die Paare einer Funktion über einen Term oder eine Funktionskonstante zu beschreiben. In diesem Fall verwendet man die Schreibweise

$$\begin{aligned} f : A &\rightarrow B \\ x &\mapsto f(x), \end{aligned}$$

für die Funktion $f = \{(x, f(x))\} \subseteq A \times B$. Da diese Schreibweise von der Funktionsrepräsentation als Menge von Paaren abstrahiert, bezeichnet man $\{(x, f(x)) \mid x \in A\}$ auch als *Graph* einer Funktion. Diese Graphen können wie schon in Abschnitt 3.5 für Relationen gezeigt visualisiert werden.

Beispiel 3.22 Wir betrachten für eine natürliche Zahl x den Rest von x bei Division durch 6; mit Satz 3.2 ist dieser Rest eindeutig definiert. Somit ist die Menge $\{(x, r) \mid \exists_{q \in \mathbb{N}} x = q \cdot 6 + r\}$ der Graph der Funktion, die jedem $x \in \mathbb{N}$ seinen Rest bei Division durch 6 zuweist. Dieser Graph ist in Abbildung 3.4 zu sehen. \square

Beispiel 3.23 Die totale Funktion

$$\begin{aligned} f : \mathbb{N} &\rightarrow \mathbb{N} \\ n &\mapsto n^2 \end{aligned}$$

hat den Graphen $\{(1, 1), (2, 4), (3, 9), (4, 16), \dots\}$. Als Relation könnte man diesen Graphen auch als $\{(n, n^2) \mid n \in \mathbb{N}\}$ schreiben. Man beachte, dass die Relation $\{(n^2, n) \mid n \in \mathbb{N}\}$ eine partielle Funktion ist, da nicht jeder Zahl im Definitionsbereich eine Zahl im Bildbereich zugeordnet werden kann.

Ebenso ist die Relation

$$\{(x, x^2) \mid x \in \mathbb{Z}\} = \{\dots, (-1, 1), (0, 0), (1, 1), (2, 4), \dots\}$$

eine Funktion, nicht aber $\{(x^2, x) \mid x \in \mathbb{Z}\}$: In diesem Fall sind sowohl $(1, 1)$ als auch $(1, -1)$ in der Relation; dies ist bei einer Funktion nicht erlaubt. \square

In vielen Fällen ist es von Interesse, wohin eine Menge aus dem Definitionsbereich abgebildet wird, bzw. welche Elemente auf eine Menge im Bildbereich abgebildet werden. Für eine Funktion $f : A \rightarrow B$ und Teilmengen $C \subseteq A$ und $D \subseteq B$ bezeichnet man die Menge

$$f(C) := \{f(x) \mid x \in C\}$$

als *Bild* von C unter f , und die Menge

$$f^{-1}(D) := \{x \in C \mid f(x) \in D\}$$

als *Urbild* von D unter f .

Beispiel 3.24 Mit der Funktion

$$\begin{aligned} f : \mathbb{N} &\rightarrow \mathbb{N} \\ n &\mapsto n^2 \end{aligned}$$

aus Beispiel 3.23 gilt

$$f(\mathbb{N}_{\leq 5}) = \{1, 4, 9, 16, 25\} \quad \text{und} \quad f^{-1}(\mathbb{N}_{\leq 100} \setminus \mathbb{N}_{\leq 70}) = \{9, 10\}. \quad \square$$

Bei Funktionsdefinitionen ist es manchmal nicht möglich, den Wert der Funktion über nur *einen* Term darzustellen. In diesem Fall muss man eine *Fallunterscheidung* vornehmen. Das bekannteste Beispiel einer solchen Funktion ist der Absolutbetrag, der einer ganzen Zahl den Abstand vom Nullpunkt zuweist:

$$\begin{aligned} \text{abs} : \mathbb{Z} &\rightarrow \mathbb{N}_0 \\ x &\mapsto \text{abs}(x) := \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{sonst.} \end{cases} \end{aligned}$$

Fallunterscheidungen sind speziell bei rekursiven Definitionen unumgänglich, um den Rekursionsteil von der Endbedingung zu trennen.

Definition 3.18 (Rekursive Funktionsdefinition)

Eine *rekursive Funktionsdefinition* folgt dem Schema

$$f(x_1, \dots, x_n) := \begin{cases} t_{f, x_1, \dots, x_n} & \text{falls } P(x_1, \dots, x_n) \\ c & \text{sonst,} \end{cases}$$

wobei t_{f, x_1, \dots, x_n} ein Term mit Variablen x_1, \dots, x_n ist, in dem auch f vorkommt, und c eine Konstante ist. Das Prädikat P bestimmt, für welche Werte von x_1, \dots, x_n die Funktion rekursiv aufgerufen wird.

Eine Funktionsdefinition muss einen eindeutigen Wert liefern. Daher muss bei rekursiven Funktionsdefinitionen sichergestellt sein, dass die Rekursion auch terminiert, der zweite Teil der Fallunterscheidung also auch erreicht wird.

Komplizierte Varianten rekursiver Definition sind möglich. Ebenso können auch Prädikatenkonstante oder sogar Quantoren rekursiv definiert werden, wie wir weiter unten in den Beispielen sehen werden.

Wir haben rekursive Definitionen bereits in der Aussagenlogik kennengelernt: Dort war eine Aussage A als Zeichenkette definiert, die entweder wiederum aus (durch Junktoren verbundenen) Aussagen besteht (Rekursionsteil), oder eine atomare Aussage ist (Endbedingung).

Beispiel 3.25 Eine rekursive Funktionsdefinition auf \mathbb{N}_0 ist die *Fakultätsfunktion*, die wie folgt definiert werden kann:

$$\begin{aligned} \cdot! : \mathbb{N}_0 &\rightarrow \mathbb{N} \\ n \mapsto n! &:= \begin{cases} n \cdot (n-1)! & \text{falls } n \geq 1 \\ 1 & \text{sonst.} \end{cases} \end{aligned}$$

Mit dieser Definition gilt

$$4! = 4 \cdot 3! = 4 \cdot 3 \cdot 2! = 4 \cdot 3 \cdot 2 \cdot 1! = 4 \cdot 3 \cdot 2 \cdot 1 \cdot 0! = 4 \cdot 3 \cdot 2 \cdot 1 \cdot 1 = 24.$$

Man erkennt, dass eine nichtrekursive Definition der Fakultätsfunktion

$$n! := \prod_{k=1}^n k$$

ist. Die Definition der Produktquantors \prod ist wiederum rekursiv:

$$\prod_{i=k}^m t_i := \begin{cases} t_m \cdot \prod_{i=k}^{m-1} t_i & \text{falls } m \geq k \\ 1 & \text{sonst,} \end{cases}$$

wobei t_i der Term mit Variable i ist, über den multipliziert wird. Die Definition des Summenquantors \sum erfolgt analog. \square

Beispiel 3.26 Auch Prädikate können rekursiv definiert werden; rekursive Definitionen können auch gegenseitig voneinander abhängen. So könnten etwa die Prädikate *gerade* und *ungerade* auf \mathbb{N}_0 auch folgendermaßen definiert werden:

$$\begin{aligned} \text{gerade}(n) &:\Leftrightarrow \begin{cases} \text{ungerade}(n-1) & \text{falls } n \geq 1 \\ \mathbf{w} & \text{sonst.} \end{cases} \\ \text{ungerade}(n) &:\Leftrightarrow \begin{cases} \text{gerade}(n-1) & \text{falls } n \geq 1 \\ \mathbf{f} & \text{sonst.} \end{cases} \end{aligned} \quad \square$$

Wie bei Relationen gibt es auch bei Funktionen spezielle Eigenschaften, die von besonderem Interesse sind.

Definition 3.19 (injektiv, surjektiv, bijektiv)

Sei $f : A \rightarrow B$ eine totale Funktion. Dann nennt man f

$$\text{injektiv} : \Leftrightarrow \forall_{x_1, x_2 \in A} f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

$$\text{surjektiv} : \Leftrightarrow \forall_{y \in B} \exists_{x \in A} f(x) = y$$

$$\text{bijektiv} : \Leftrightarrow \text{injektiv}(f) \wedge \text{surjektiv}(f).$$

Mit dieser Definition ist eine Funktion injektiv, wenn es für jedes Element y im Wertebereich höchstens ein Element im Definitionsbereich gibt, das auf y abgebildet wird. Eine Funktion ist surjektiv, wenn auf jedes Element des Wertebereichs abgebildet wird. Sind diese beiden Bedingungen erfüllt, so ist die Funktion bijektiv.

Beispiel 3.27 Sei $A := \{1, 2, 3\}$, $B := \{6, 7, 8, 9\}$ und $C := \{10, 11, 12\}$. Auf diesen Mengen seien folgende Funktionen gegeben:

$$f_1 : A \rightarrow C, f_1 := \{(1, 10), (2, 11), (3, 11)\}$$

$$f_2 : A \rightarrow B, f_2 := \{(1, 6), (2, 7), (3, 8)\}$$

$$f_3 : B \rightarrow C, f_3 := \{(6, 10), (7, 11), (8, 12), (9, 10)\}$$

$$f_4 : A \rightarrow C, f_4 := \{(1, 10), (2, 11), (3, 12)\}.$$

Wie man nachprüfen kann, ist f_1 nicht injektiv und nicht surjektiv; f_2 injektiv, aber nicht surjektiv; f_3 surjektiv, aber nicht injektiv; und f_4 injektiv und surjektiv, also bijektiv. \square

Beispiel 3.28 Auf den natürlichen bzw. ganzen Zahlen gelten folgende Eigenschaften der angegebenen Funktionen:

$$d : \mathbb{N} \rightarrow \mathbb{N}, x \mapsto x + 1 \quad \text{injektiv, aber nicht surjektiv}$$

$$e : \mathbb{Z} \rightarrow \mathbb{Z}, x \mapsto x^2 \quad \text{weder injektiv noch surjektiv}$$

$$g : \mathbb{Z} \rightarrow \mathbb{Z}, x \mapsto \text{abs}(x) \quad \text{weder injektiv noch surjektiv}$$

$$h : \mathbb{Z} \rightarrow \mathbb{Z}, x \mapsto x + 1 \quad \text{bijektiv.} \quad \square$$

Eine spezielle Bedeutung kommt Funktionen zu, die auf den natürlichen Zahlen definiert sind. Solche Funktionen nennt man *Folgen*.

Definition 3.20 (Folge)

Sei A eine Menge. Eine Funktion $a : \mathbb{N} \rightarrow A$ nennt man (*unendliche*) *Folge* über A , die Funktionswerte $a(k)$ die *Folglieder* von A . Wenn a nur auf einem endlichen Bereich $\mathbb{N}_{\leq n}$ definiert ist, nennt man a eine *endliche Folge*. Für eine endliche Folge mit Definitionsbereich $\mathbb{N}_{\leq n}$ nennt man n die *Länge* der Folge. Den Ausdruck $a(k)$ kürzt man meist mit a_k ab, die damit bezeichneten Elemente nennt man *Folglieder*.

Für Folgen hat sich eine Schreibweise eingebürgert, die Tupeln sehr ähnlich ist. So schreibt man die Folge

$$a := \{(1, a_1), (2, a_2), (3, a_3), \dots\}$$

meist in der Form

$$a := (a_1, a_2, a_3, \dots),$$

da die Angabe der Werte $1, 2, 3, \dots$ im Definitionsbereich ja redundant ist. Ein weiterer Vorteil dieser Notation ist, dass dann die Schreibweise des Folgenglieds a_i mit der Projektion a_i auf die i -te Komponente eines Tupels a übereinstimmt. Wir unterscheiden damit nicht zwischen Tupeln und endlichen Folgen.

Beispiel 3.29 Gegeben sei die Menge $A = \{7, 3, 8, 9, 5\}$. Die Funktion $a : \mathbb{N}_{\leq 5} \rightarrow A$ mit

$$a := (8, 3, 9, 7, 5)$$

ist eine endliche Folge der Länge 5. Es gilt etwa $a_1 = 8$ und $a_3 = 9$, sowie $\sum_{i=1}^5 a_i = 32$ und $\prod_{i=1}^5 a_i = 7560$. \square

Das Konzept der Folgen kann einfach auf zwei und mehr Dimensionen verallgemeinert werden. Zweidimensionale endliche Folgen nennt man *Matrizen*.

Definition 3.21 (Matrix)

Sei M eine Menge. Eine Funktion $A : \mathbb{N}_{\leq m} \times \mathbb{N}_{\leq n} \rightarrow M$ nennt man eine *Matrix* mit m Zeilen und n Spalten (oder kürzer $m \times n$ -Matrix) über M . Abkürzend schreibt man meist $A_{i,j}$ oder A_{ij} für $A(i, j)$.

Da Matrizen zweidimensionale Objekte sind, werden sie auch in zwei Dimensionen dargestellt. So schreibt man eine $m \times n$ Matrix A als

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}.$$

Beispiel 3.30 Die Matrix

$$A := \begin{pmatrix} -3 & 7 & 0 & 2 \\ 1 & -2 & 1 & 6 \\ -2 & 0 & 4 & 2 \end{pmatrix}$$

ist eine 3×4 -Matrix über den ganzen Zahlen; es gilt etwa $A_{14} = 2$ und $A_{33} = 4$. \square

Zusammenfassend können wir also festhalten, dass Funktionen spezielle Relationen sind, bei denen einem Objekt im Definitionsbereich genau *ein* Objekt im Wertebereich zugewiesen wird. Bei expliziten Funktionsdefinitionen (Definition 3.5) ist diese

Eindeutigkeit dadurch gewährleistet, dass der definierende Term auf der rechten Seite der Definition eindeutig ist und daher für gegebene Variablenbelegung auch nur ein Objekt beschreiben kann.

Manche Funktionen lassen sich aber nicht durch Terme definieren; für diese benötigt man implizite Funktionsdefinitionen.

Definition 3.22 (Implizite Funktionsdefinition)

Bei einer *impliziten Funktionsdefinition* wird eine Funktion über die Eigenschaft desjenigen Objekts definiert, das das Funktionssymbol für gegebenes Eingabobjekte liefern soll:

$$f(x_1, \dots, x_n) := \text{dasjenige } y \text{ mit der Eigenschaft } P_{x_1, \dots, x_n, y},$$

wobei $P_{x_1, \dots, x_n, y}$ ein Prädikatenkonstante mit freien Variablen x_1, \dots, x_n, y ist, die den Zusammenhang zwischen Ein- und Ausgabegrößen definiert.

Man beachte, dass der Ausdruck “dasjenige y mit der Eigenschaft P ” ein Quantor im Sinne der Definition 2.16 ist, da es die Variable y bindet, und aus Variable und einer Aussage einen Term macht.

Beispiel 3.31 Das bekannteste Beispiel einer impliziten Funktionsdefinition ist die Definition der Wurzelfunktion, die man in einem ersten Versuch folgendermaßen definieren könnte:

$$\sqrt{x} := \text{dasjenige } y \text{ mit } y^2 = x.$$

Eine kurze Überlegung zeigt, dass diese Definition aber die Bedingung der Eindeutigkeit verletzt, da etwa $\sqrt{4} = +2$ und $\sqrt{4} = -2$ möglich wären. Die korrekte Definition lautet somit

$$\sqrt{x} := \text{dasjenige } y \text{ mit } y \geq 0 \wedge y^2 = x. \quad \square$$

Beispiel 3.32 Minimum- und Maximumquantoren lassen sich ebenfalls implizit definieren, wie folgendes Beispiel zeigt:

$$\min_{P_i} t_i := \text{dasjenige } t_i \text{ mit } P_i \wedge \forall j P_j \Rightarrow t_j \geq t_i.$$

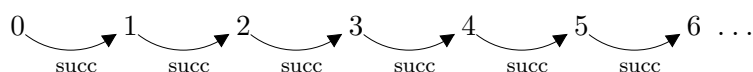
Hier ist i die von Minimum-Quantor gebundene Variable, P_i das Prädikat mit freier Variable i , das die möglichen Werte von i einschränkt, und t_i der Term mit Variable i , dessen Minimum bestimmt werden soll. \square

3.6 Struktur in Mengen

Nach diesen Abschnitten über Funktionen und Relationen beginnen wir nun, Strukturen auf Mengen zu identifizieren. Wir werden in Kapitel 4 sehen, dass Mengen zwar die elementaren Bausteine der Mathematik sind, aber erst durch Operationen auf diesen Mengen den Bausteinen Struktur gegeben werden kann.

In diesem Abschnitt betrachten wir als Vorgeschmack, wie durch einfache Bildungsvorschriften Mengen festgelegt werden, und wie über diese Bildungsvorschriften Aussagen über die Mengen bewiesen werden können. Wir beschränken uns dabei aber auf die natürlichen Zahlen. Die Struktur der natürlichen Zahlen kann über zwei Arten festgelegt werden: Entweder über die Mengenlehre (indem alle Zahlen als Mengen repräsentiert werden), oder axiomatisch über die Eigenschaften, die die natürlichen Zahlen beschreiben. Wir wählen in diesem Abschnitt den zweiten Weg, ohne aber alle Axiome aufzulisten.

Im axiomatischen Ansatz fordert man die Existenz einer ersten (kleinsten) natürlichen Zahl 0. Alle weiteren natürlichen Zahlen erhalten wir durch Anwendung einer Funktion *succ*, die jeder natürlichen Zahl ihren Nachfolger zuweist:



Die Struktur der natürlichen Zahlen ist somit über ihre Nachfolger-Funktion gegeben. Aussagen über die natürlichen Zahlen können unter Verwendung dieser Struktur bewiesen werden: Wenn nämlich eine Aussage P für 0 gilt, und P durch die Nachfolger-Funktion erhalten bleibt (also $P(n) \Rightarrow P(\text{succ}(n))$ gilt), dann gilt $P(n)$ für alle n . Die Gültigkeit dieses Beweisschemas kann zwar anschaulich begründet, nicht aber von anderen Axiomen oder Sätzen hergeleitet werden. Sie muss daher für die natürlichen Zahlen axiomatisch gefordert werden. Da man im Prädikatenkalkül aber nicht "für alle Prädikate P " ausdrücken kann (da ja nur über Objekte quantifiziert werden kann), behilft man sich mit einem *Axiomenschema*: Für jedes Prädikaten-symbol P sei daher die Gültigkeit von

$$(P(0) \wedge \forall_{x \in \mathbb{N}_0} P(x) \Rightarrow P(\text{succ}(x))) \Rightarrow \forall_{x \in \mathbb{N}_0} P(x) \quad (\text{IP})$$

gegeben. Beweise, die diese Struktur der natürlichen Zahlen ausnützen, nennt man *Induktionsbeweise*.

Die Struktur von Induktionsbeweisen ist durch die Struktur des Induktionsprinzips IP festgelegt. Diese Aussage, die wir als wahr annehmen, hat die Form $(A \wedge B) \Rightarrow C$. Zu zeigen ist die Konklusion C dieser Implikation. Wenn wir nun $A \wedge B$ zeigen können, folgt mit modus ponens sofort C . Somit lässt sich unter Verwendung von IP die Aussage $\forall_{x \in \mathbb{N}_0} P(x)$ beweisen, indem man $P(0)$ und $\forall_{x \in \mathbb{N}_0} P(x) \Rightarrow P(\text{succ}(x))$ beweist.

Im Folgenden werden wir die gewohntere Schreibweise $n + 1$ für $\text{succ}(n)$ verwenden.

Beispiel 3.33 Wir beweisen die Aussage

$$\forall_{n \in \mathbb{N}_0} \sum_{i=0}^n 2^i = 2^{n+1} - 1.$$

Da es sich dabei um eine Aussage über alle natürlichen Zahlen handelt, können wir das Induktionsprinzip verwenden. Wir müssen somit die beiden Aussagen

$$\sum_{i=0}^0 2^0 = 2^{0+1} - 1 \quad \text{und} \quad \forall_{x \in \mathbb{N}_0} \sum_{i=0}^x 2^i = 2^{x+1} - 1 \Rightarrow \sum_{i=0}^{x+1} 2^i = 2^{x+2} - 1$$

beweisen. Die erste Aussage wird *Induktionsbeginn* genannt. Die zweite Aussage hat die Form einer allquantifizierten Implikation; nachdem wir die quantifizierte Variable beliebig aber fix gesetzt haben, nehmen wir die Prämisse der Implikation an (*Induktionsannahme*) und zeigen dann die Konklusion (*Induktionsschluss*). Unter Verwendung des Induktionsprinzips IP ist damit die zu zeigende Aussage bewiesen.

Im konkreten Fall sind die drei Komponenten des Beweises wie folgt:

Induktionsbeginn: Für $n = 0$ ist

$$\sum_{i=0}^0 2^0 = 1 = 2^{0+1} - 1.$$

Induktionsannahme: Für beliebig aber fixes n gelte

$$\sum_{i=0}^n 2^i = 2^{n+1} - 1.$$

Induktionsschluss: Zu zeigen ist für beliebig aber fixes n :

$$\sum_{i=0}^{n+1} 2^i = 2^{n+2} - 1,$$

wobei die Induktionsannahme als wahr angenommen wird. An dieser Stelle kann man auf zwei Arten fortfahren: Entweder bringt man durch Umformungen die linke Seite der zu zeigenden Gleichung auf den gleichen Term wie die rechte Seite, oder man formt beide Seiten unabhängig so um, dass auf beiden Seite gleiche Terme stehen.

Wir wählen in diesem Beispiel den ersten Ansatz und formen folgendermaßen um:

$$\begin{aligned} \sum_{i=0}^{n+1} 2^i &= \sum_{i=0}^n 2^i + 2^{n+1} \stackrel{\text{Ind. Ann.}}{=} 2^{n+1} - 1 + 2^{n+1} = 2 \cdot 2^{n+1} - 1 \\ &= 2^{n+2} - 1. \end{aligned}$$

Mit dem Induktionsschluss ist damit auch die zu zeigende Aussage bewiesen. \square

Beispiel 3.34 Wir beweisen mittels Induktion die Aussage

$$\forall_{n \in \mathbb{N}} \sum_{i=1}^n i = \frac{n(n+1)}{2}.$$

Induktionsbeginn: In diesem Fall beginnen wir mit dem Summationsindex 1 und überprüfen somit die Aussage für $n = 1$:

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

Induktionsannahme: Es gelte

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

Induktionsschluss: Zu zeigen ist

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

Umformen der linken Seite unter Verwendung der Induktionsannahme ergibt

$$\begin{aligned} \sum_{i=1}^{n+1} i &= \sum_{i=1}^n i + (n+1) \stackrel{\text{Ind. Ann.}}{=} \frac{n(n+1)}{2} + (n+1) \\ &= \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2}. \quad \square \end{aligned}$$

Das Induktionsprinzip lässt sich auf alle Bereiche anwenden, in denen Objekte durch Anwenden von Bildungsregeln aus einfacheren Objekten erzeugt werden; in diesem Fall spricht man von *struktureller Induktion*. Wir haben schon in Abschnitt 2.1 solche Objekte kennengelernt: Auf syntaktischer Ebene lassen sich alle Aussagen der Aussagenlogik über die in Definition 2.1 gegebenen Bildungsregeln zusammensetzen. Wir werden darauf aber nicht weiter eingehen.

3.7 Endliche und unendliche Mengen

In diesem Abschnitt werden wir uns mit dem Unterschied zwischen *endlichen* und *unendlichen* Mengen beschäftigen. Wir werden sehen, welche Eigenschaften endliche von unendlichen Mengen unterscheiden, und dass es mehrere Arten von unendlichen Mengen gibt.

Zunächst benötigen wir eine mathematische Definition von endlichen bzw. unendlichen Mengen.

Definition 3.23 (endlich, unendlich)

Eine Menge M heißt

$$\text{endlich} : \Leftrightarrow M = \emptyset \vee \exists_{n \in \mathbb{N}} \exists f: \mathbb{N}_{\leq n} \rightarrow M \text{ bijektiv}(f)$$

$$\text{unendlich} : \Leftrightarrow \neg \text{endlich}(M).$$

Für endliche Mengen M heißt das n aus der Bijektion $f: \mathbb{N}_{\leq n} \rightarrow M$ *Mächtigkeit* bzw. *Kardinalität* von M ; man schreibt dafür $|M|$. Wir vereinbaren $|\emptyset| = 0$.

Eine nichtleere Menge M ist somit genau dann endlich, wenn es eine Bijektion zwischen M und einer Teilmenge $\mathbb{N}_{\leq n}$ der natürlichen Zahlen gibt.

Beispiel 3.35 Die Menge $\{a, b, c, d, e, g, h\}$ ist endlich, da es eine Bijektion $f : \mathbb{N}_{\leq 7} \rightarrow M$ gibt; es gilt $|M| = 7$.

Die Menge der natürlichen Zahlen ist unendlich, da wir nachrechnen können, dass sie nicht endlich ist: Es gilt

$$\neg \text{endlich}(\mathbb{N}) \Leftrightarrow \forall_{n \in \mathbb{N}} \forall_{f: \mathbb{N}_{\leq n} \rightarrow \mathbb{N}} \neg \text{bijektiv}(f).$$

Seien also $n \in \mathbb{N}$ und $f : \mathbb{N}_{\leq n} \rightarrow \mathbb{N}$ beliebig aber fix; wir müssen noch $\neg \text{bijektiv}(f)$ zeigen. Wir zeigen, dass f nicht surjektiv sein kann: Sei dazu k die größte Zahl im Bild von $\mathbb{N}_{\leq n}$ unter f . Dann ist $k + 1 \in \mathbb{N}$, aber nicht in $f(\mathbb{N}_{\leq n})$: f ist also nicht surjektiv. \square

Wir werden im Folgenden sehen, dass es mehrere Klassen von unendlichen Mengen gibt. Man kann die Mächtigkeiten unendlicher Mengen über Bijektionen vergleichen.

Definition 3.24 (gleich mächtig)

Zwei Mengen A und B sind *gleich mächtig*, wenn es eine Bijektion zwischen A und B gibt, also

$$\text{gleichmächtig}(A, B) :\Leftrightarrow \exists_{f: A \rightarrow B} \text{bijektiv}(f).$$

Für endliche Mengen A und B gilt sofort

$$\text{gleichmächtig}(A, B) \Leftrightarrow |A| = |B|.$$

Gleichmächtigkeit wird interessant, wenn A und B unendliche Mengen sind. Man unterscheidet zwei Fälle.

Definition 3.25 (abzählbar unendlich, überabzählbar)

Eine unendliche Menge A heißt

$$\begin{aligned} \text{abzählbar unendlich} &:\Leftrightarrow \text{gleichmächtig}(A, \mathbb{N}) \\ \text{überabzählbar} &:\Leftrightarrow \neg \text{gleichmächtig}(A, \mathbb{N}). \end{aligned}$$

Die ganzen Zahlen sind abzählbar unendlich, da man eine Bijektion zu den natürlichen Zahlen angeben kann:

$$f : \mathbb{N} \rightarrow \mathbb{Z} \\ n \mapsto \begin{cases} \frac{n}{2} & \text{falls } n \text{ gerade} \\ -\frac{n-1}{2} & \text{sonst.} \end{cases}$$

Die so definierte Folge ist $f = (0, 1, -1, 2, -2, \dots)$.

Die rationalen Zahlen \mathbb{Q} sind ebenfalls abzählbar unendlich, wie man durch geschicktes Anschreiben und Abzählen nachprüfen kann. Für die reellen Zahlen ist dies nicht mehr möglich.

Satz 3.3 Die reellen Zahlen sind überabzählbar.

Der Beweis dieser Tatsache beruht auf folgendem *Diagonalisierungsargument* von Cantor. Wir werden beweisen, dass das Einheitsintervall

$$I := \{x \in \mathbb{R} \mid 0 < x < 1\}$$

bereits überabzählbar ist. \mathbb{R} ist dann auch überabzählbar, da \mathbb{R} und I gleich mächtig sind: Eine mögliche Bijektion von I nach \mathbb{R} ist durch $x \mapsto \frac{-(x-0.5)}{x(x-1)}$ gegeben.

Cantors Diagonalisierungsbeweis ist ein *Widerspruchsbeweis*. Widerspruchsbeweise basieren auf folgendem Schema: Statt eine Aussage A zu beweisen, beweist man die Implikation $\neg A \Rightarrow \mathbf{f}$. Wie wir aus der Wahrheitstabelle wissen, kann diese Implikation nur wahr sein, wenn $\neg A$ falsch und somit A wahr ist. Zum Beweis dieser Implikation nehmen wir die Prämisse $\neg A$ an und zeigen dann die Konklusion \mathbf{f} , einen Widerspruch.

Das Diagonalisierungsargument von Cantor beruht auf der Tatsache, dass sich jede reelle Zahl in I in Dezimalnotation als unendliche Folge $0.d_1d_2d_3\dots$ schreiben lässt, wobei die d_i die Ziffern 0 bis 9 sind. Sei

$$D := \{0.d_1d_2d_3\dots \mid \forall_{i \geq 1} d_i \in \{0, \dots, 9\}\}$$

die Menge aller Dezimalzahlen. Da sich D und I entsprechen, gibt es eine Bijektion zwischen diesen beiden Mengen. Wir zeigen, dass es keine Bijektion zwischen I und \mathbb{N} geben kann, da es keine Bijektion zwischen D und \mathbb{N} geben kann. Dazu verwenden wir einen Widerspruchsbeweis: Wir nehmen an, dass es eine Bijektion f zwischen \mathbb{N} und D gebe, und leiten daraus einen Widerspruch ab.

Wenn die Bijektion f existiert, können wir das Bild $f(\mathbb{N})$ anschreiben:

$$f(1) = 0.d_{11}d_{12}d_{13}\dots$$

$$f(2) = 0.d_{21}d_{22}d_{23}\dots$$

$$f(3) = 0.d_{31}d_{32}d_{33}\dots$$

Wir zeigen jetzt, dass es ein Element c aus D gibt, das *nicht* in $f(\mathbb{N})$ ist. Die Folge von Dezimalstellen in c sei als

$$c_i := d_{ii} + 1$$

definiert, wobei $9 + 1 = 0$ sein soll. Dann ist die Zahl $0.c_1c_2c_3\dots$ in D , nicht aber in $f(\mathbb{N})$: c unterscheidet sich für alle i an Position i von $f(i)$. Damit ist f nicht surjektiv, ein Widerspruch zu unserer Annahme, dass f bijektiv sei.

Wir fassen zusammen: Die Zahlenmengen \mathbb{N} , \mathbb{Z} und \mathbb{Q} sind gleich mächtig, und \mathbb{R} ist mächtiger als diese Mengen. Es gibt aber noch mächtigere Mengen, wie folgender Satz besagt.

Satz 3.4 Für jede Menge M gilt

$$\neg \text{gleichmächtig}(M, \text{Pot}(M)).$$

Mit diesem Satz kann man Hierarchien von unendlich großen Mengen aufbauen. Am unteren Ende dieser Hierarchie steht \mathbb{N} . Man kann zeigen, dass $\text{Pot}(\mathbb{N})$ und \mathbb{R} gleich mächtig sind – aber gibt es eine Menge, deren Mächtigkeit zwischen \mathbb{N} und \mathbb{R} liegt? Die sogenannte *Kontinuumshypothese* besagt, dass es keine solche Menge gibt, dass also \mathbb{R} die nach \mathbb{N} “nächstgrößere” Menge ist. Die Kontinuumshypothese nimmt in der Mathematik eine Sonderstellung ein, da man beweisen kann, dass weder sie noch ihre Negation aus anderen Sätzen und Axiomen herleitbar ist.

3.8 Problemspezifikation

Eine der wichtigsten Anwendungen einer formal exakten Sprache in der Informatik ist die Möglichkeit, sprachlich präzise über Probleme und deren Lösungen zu sprechen bzw. zu schreiben. Problemspezifikationen werden meist in einer starren formalen Struktur angegeben. Dabei werden Eingabe- und Ausgabegrößen, sowie ihr Zusammenhang spezifiziert.

Wir illustrieren die starre formale Struktur einer Problemspezifikation anhand eines einfachen Beispiels.

Beispiel 3.36 Jede Zahl hat mehrere Teiler; bei Primzahlen sind dies nur 1 und die Zahl selbst. Für gegebene zwei natürliche Zahlen soll die größte Zahl bestimmt werden, die die beiden teilt – diese Zahl ist der *größte gemeinsame Teiler* ggT.

GEGEBEN: a, b

wobei $a, b \in \mathbb{N}$

GESUCHT: g

sodass $g \in \mathbb{N} \wedge g|a \wedge g|b \wedge \forall_{c \in \mathbb{N}_0} (c|a \wedge c|b) \Rightarrow c \leq g$ □

Der GEGEBEN:... wobei ... und GESUCHT:... sodass... Teil des obigen Beispiels ist in allen Problemspezifikationen gleich; allerdings kann die “Typisierung” der Variablen bereits in den GEGEBEN:/GESUCHT:-Zeilen erfolgen. Eine Problemspezifikation gibt somit an, in welchem Zusammenhang gegebene und gesuchte Größen stehen müssen, um die Lösung eines Problems darzustellen.

Zwei wichtige Punkte müssen zu Problemspezifikationen hervorgehoben werden:

- Der Zusammenhang zwischen gegebenen und gesuchten Größen in der “sodass”-Zeile *muss keinen Algorithmus angeben, wie sich die Ausgabe aus der Eingabe berechnen lässt*. Es reicht, wenn mit dem Zusammenhang überprüft werden kann, ob eine Lösung vorliegt.
- Außerhalb der Problemspezifikation muss festgelegt werden, welche Operationen in der Problemspezifikationen erlaubt sind.

Wenn wir das obige Beispiel 3.36 nochmals unter Beachtung dieser beiden Punkte betrachten, dann fällt Folgendes auf: In der “sodass”-Zeile ist tatsächlich nur der Zusammenhang zwischen a , b und g angegeben, und kein Algorithmus, um von a und b auf g zu kommen (obwohl das nicht prinzipiell verboten wäre). Weiters fehlt die Einschränkung der erlaubten Operationen, sodass von der Menge aller bisher bekannten Operationen ausgegangen werden kann.

Warum die Angabe von erlaubten Operationen wichtig ist zeigt sich daran, dass man sonst für jede beliebige Problemspezifikation in der “sodass”-Zeile einfach das Prädikat $istLösungVon(a,b,c,\dots)$ angeben könnte. Die “echte” Lösung wäre dann in der (nicht angegebenen) Definition von $istLösungVon$ versteckt. In Beispiel 3.36 entspricht das einer “sodass”-Zeile, die nur aus $istGGTvon(g,a,b)$ besteht. Wir werden im Folgenden, soweit nicht anders angegeben, alle bisher zur Verfügung stehenden Funktionen und Prädikate als erlaubt annehmen.

Beispiel 3.37 Gegeben sei eine Teilmenge der natürlichen Zahlen. Zu spezifizieren sei das Problem, von dieser Teilmenge das kleinste Element zu bestimmen. Eine mögliche Problemspezifikation ist wie folgt:

GEGEBEN: $M \subseteq \mathbb{N}$

GESUCHT: $a \in \mathbb{N}$

$$\text{sodass } a \in M \wedge \forall_{b \in M} b \geq a.$$

Auch in diesem Beispiel wird keine Berechnungsvorschrift des minimalen Elements a angegeben. Man beachte aber, dass rein durch die Angabe einer Bedingung in der “sodass”-Zeile noch nicht garantiert wird, dass ein gesuchtes Element überhaupt existiert. Wenn man obiges Beispiel in der Domäne der reellen Zahlen überträgt, so ergibt sich (bis auf die Variablentypisierung) die gleiche Spezifikation:

GEGEBEN: $M \subseteq \mathbb{R}$

GESUCHT: $a \in \mathbb{R}$

$$\text{sodass } a \in M \wedge \forall_{b \in M} b \geq a.$$

Wenn man etwa für M die Menge $\{x \in \mathbb{R} \mid 1 < x < 2\}$ wählt, so müsste das minimale Element dieser Menge zwar weiterhin die Bedingung der “sodass”-Zeile erfüllen – es existiert aber nicht, wie folgende Überlegung zeigt: a kann nicht 1 sein, da $1 \notin M$ gilt. Für jedes mögliche $a \in M$ gibt es ein weiteres $b \in M$ mit $1 < b < a$ (etwa $b := (1 + a)/2$), womit a nicht das kleinste Element in M sein kann. \square

Nach Betrachtung einiger Beispiele fällt auf, dass es einen Zusammenhang zwischen Problemspezifikation und impliziter Funktionsdefinition (Definition 3.22) gibt: So kann etwa die “sollte”-Zeile in einer impliziten Definition verwendet werden, um eine Funktion zu definieren, die die gesuchte aus den gegebenen Größen liefert (wenn diese Spezifikation eindeutig ist!). So kann man etwa die Funktion, die das Minimum einer Menge liefert, folgendermaßen definieren:

$$\min : \text{Pot}(A) \rightarrow A$$

$$M \mapsto \min(M) := \text{dasjenige } a \text{ mit } a \in M \wedge \forall_{b \in M} b \geq a.$$

Beispiel 3.38 Ein bekanntes Problem in der Informatik ist das n -Damen Problem. Bei dieser Aufgabenstellung ist eine Anordnung von n Damen auf einem $n \times n$ -Schachbrett gesucht, bei der keine der Damen eine andere angreift. Eine mögliche erste Spezifikation könnte sein:

GEGEBEN: $n \in \mathbb{N}$

GESUCHT: Anordnung von n Damen auf dem Schachbrett der Größe $n \times n$

sodass es keine Attacken zwischen den Damen gibt.

Obwohl diese Spezifikation die formalen Vorgaben erfüllt, ist sie aber noch nicht detailliert genug, um etwa computergestützt überprüft zu werden: Was ist ein “Schachbrett der Größe $n \times n$ ”, was bedeutet “keine Attacken zwischen den Damen”?

Als Datenstruktur für die Positionen von Damen auf einem Schachbrett bietet sich eine Menge von Paaren an, bei denen jeder Eintrag eine Zahl zwischen 1 und n ist:

$$S_n \subseteq \{1, \dots, n\} \times \{1, \dots, n\}.$$

In dieser Menge von Paaren gibt der erste Eintrag eines Paares die Spalte, und der zweite die Zeile einer Dame am Schachbrett an.

Da es sich bei S_n um eine beliebige Teilmenge von $\{1, \dots, n\} \times \{1, \dots, n\}$ handelt, müssen mehrere Bedingungen gestellt werden, um eine gültige Lösung zu spezifizieren:

- (1) S_n darf nur n Elemente enthalten,
- (2) in S_n muss jede Zahl zwischen 1 und n genau einmal an der ersten Position eines Paares auftreten (keine Attacken in einer Spalte),
- (3) in S_n muss jede Zahl zwischen 1 und n genau einmal an der zweiten Position eines Paares auftreten (keine Attacken in einer Zeile),
- (4) Attacken entlang der Diagonalen sind nicht erlaubt.

Man kann erkennen, dass diese Bedingungen unterschiedlich einfach in mathematisch exakte Notation übertragen werden können. So ist dies bei (1) einfach, bei (2) und (3) schwieriger, und bei (4) nicht ohne weitere Überlegungen möglich.

Mögliche Definitionen der Bedingungen (1)–(3), die eine Lösung S_n erfüllen muss, sind:

$$\text{n-elementig}(S_n) :\Leftrightarrow |S_n| = n$$

$$\text{spalten-OK}(S_n) :\Leftrightarrow \forall_{1 \leq k \leq n} \exists_{p \in S_n} p_1 = k \wedge \neg \exists_{q \in S_n} q \neq p \wedge q_1 = k$$

$$\text{zeilen-OK}(S_n) :\Leftrightarrow \forall_{1 \leq k \leq n} \exists_{p \in S_n} p_2 = k \wedge \neg \exists_{q \in S_n} q \neq p \wedge q_2 = k.$$

Für Bedingung (4) benötigt man eine Idee, die wir hier nicht anführen. Wir nehmen an, dass diese Idee in der Definition eines Prädikats *diagonalen-OK* gegeben ist.

Die verfeinerte Problemspezifikation des n -Damen-Problems ist dann wie folgt:

GEGEBEN: $n \in \mathbb{N}$

GESUCHT: $S_n \subseteq \{1, \dots, n\} \times \{1, \dots, n\}$

sodass $\text{n-elementig}(S_n) \wedge \text{spalten-OK}(S_n) \wedge \text{zeilen-OK}(S_n) \wedge \text{diagonalen-OK}(S_n)$.

Wie bei allen bisherigen Beispielen liefert diese Problemspezifikation keinen Algorithmus, mit dem man das Problem lösen könnte, sondern lediglich ein Kriterium, das eine richtige Lösung erfüllen muss. \square

Kapitel 4

Mengen + Operationen = Algebraische Strukturen

Im letzten Kapitel haben wir Mengen hauptsächlich als Behälter von Elementen betrachtet. Durch Operationen können Strukturen auf den Elementen einer Menge definiert werden, und Gemeinsamkeiten und Unterschiede verschiedener Strukturen herausgearbeitet werden.

Wir beginnen mit einer Begriffsbildung: Unter einer *Operation* auf einer Menge M verstehen wir eine zweistellige totale Funktion $f : M \times M \rightarrow M$; manchmal nennt man f auch *Verknüpfung*. Für eine auf M definierte Operation heißt M die *Trägermenge* der Operation. Als *algebraische Struktur* bezeichnen wir eine Menge zusammen mit einer (oder mehreren) darauf definierten Operation(en), die bestimmten Gesetzmäßigkeiten gehorcht. Ein einfaches Beispiel einer Operation ist die Addition $+$ auf den natürlichen Zahlen. Operationen können somit wie andere Funktionen definiert werden. Für endliche (und sehr kleine) Mengen M bietet sich dazu auch die Möglichkeit der *Verknüpfungstabellen*. Diese repräsentieren den Graphen einer zweistelligen Funktion in Matrix-ähnlicher Form.

Beispiel 4.1 Auf der Menge $M = \{a, b, c\}$ sei die Operation \circ definiert als

$$\circ = \{((a, a), c), ((a, b), c), ((a, c), b), ((b, a), b), ((b, b), a), ((b, c), a), ((c, a), b), ((c, b), c), ((c, c), a)\}.$$

Diese Operation lässt sich einfacher über ihre Verknüpfungstabelle anschreiben:

\circ	a	b	c
a	c	c	b
b	b	a	a
c	b	c	a

□

Beispiel 4.2 Die Operation \cap auf der Menge $\text{Pot}(\{0, 1\})$ hat folgende Verknüpfungstabelle:

\cap	\emptyset	$\{0\}$	$\{1\}$	$\{0, 1\}$
\emptyset	\emptyset	\emptyset	\emptyset	\emptyset
$\{0\}$	\emptyset	$\{0\}$	\emptyset	$\{0\}$
$\{1\}$	\emptyset	\emptyset	$\{1\}$	$\{1\}$
$\{0, 1\}$	\emptyset	$\{0\}$	$\{1\}$	$\{0, 1\}$

□

Im Folgenden werden wir exemplarisch eine spezielle Menge mit Verknüpfung betrachten, anhand derer viele wichtige Konzepte einfach erläutert werden können.

Definition 4.1 (Permutation)

Sei M eine endliche Menge. Dann nennt man eine Bijektion von M nach M eine *Permutation*.

Wir werden ausschließlich Permutationen von Teilmengen von \mathbb{N} betrachten, und schreiben dafür

$$P_n := \{f \mid f : \mathbb{N}_{\leq n} \rightarrow \mathbb{N}_{\leq n} \wedge \text{bijektiv}(f)\}.$$

Für Permutationen verwendet man eine abkürzende Notation, bei der man die Zuordnungsvorschriften untereinander schreibt. So schreibt man die Permutation

$$p = \{(1, 3), (2, 1), (3, 4), (4, 2)\}$$

als

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}.$$

Auf Relationen (und damit auch auf Funktionen und Permutationen) kann man die *Hintereinanderausführung* als Verknüpfung definieren.

Definition 4.2 (Hintereinanderausführung)

Seien $R \subseteq A \times B$ und $S \subseteq B \times C$ zwei Relationen. Dann nennt man die Relation

$$S \circ R := \{(a, c) \in A \times C \mid \exists_{b \in B} (a, b) \in R \wedge (b, c) \in S\}$$

die *Hintereinanderausführung* von S nach R .

Beispiel 4.3 Seien R und S folgende Relationen auf der Menge $\{a, b, c, d\}$:

$$\begin{aligned} R &:= \{(a, b), (a, d), (b, c), (b, d)\} \\ S &:= \{(a, c), (a, d), (c, b), (d, c)\} \end{aligned}$$

Dann ist

$$\begin{aligned} R \circ S &= \{(c, c), (c, d)\} \\ S \circ R &= \{(a, c), (b, b), (b, c)\} \\ R \circ R &= \{(a, c), (a, d)\} \\ S \circ S &= \{(a, b), (a, c), (d, b)\}. \end{aligned} \quad \square$$

Beispiel 4.4 Zwei Beispiele für die Hintereinanderausführung zweier Permutationen aus P_5 sind

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix}. \quad \square$$

4.1 Gruppen

Das Interessante an Verknüpfungen ist nun, Regelmässigkeiten in ihnen zu entdecken. Wir werden in diesem Abschnitt die algebraische Struktur *Gruppe* betrachten, deren Rechengesetze wir zuerst anhand von Beispielen vorstellen, bevor wir eine genaue Definition geben.

Wie wir schon in Kapitel 2 gesehen haben, sind viele Verknüpfungen *assoziativ*: Dies ist dann der Fall, wenn für alle Elemente a, b, c einer Menge mit Verknüpfung \circ

$$a \circ (b \circ c) = (a \circ b) \circ c$$

gilt. Da die Klammersetzung keine Rolle spielt, kann sie weggelassen werden. Beispiele assoziativer Verknüpfungen sind etwa \wedge und \vee auf der Menge der Aussagen, \cap und \cup auf Mengen, die Hintereinanderausführung auf Relationen, oder Addition und Multiplikation auf den ganzen Zahlen. Beispiele nicht-assoziativer Verknüpfungen sind die Implikation auf Aussagen, die in Beispiel 4.1 angegebene Operation, oder die Subtraktion auf den ganzen Zahlen. Es gilt

$$\begin{aligned} \mathbf{f} \Rightarrow (\mathbf{f} \Rightarrow \mathbf{f}) &\equiv \mathbf{w}, \text{ aber: } (\mathbf{f} \Rightarrow \mathbf{f}) \Rightarrow \mathbf{f} \equiv \mathbf{f} \\ (a \circ b) \circ c &= a, \text{ aber: } a \circ (b \circ c) = c \quad (\text{aus Beispiel 4.1}) \\ 7 - (5 - 3) &= 5, \text{ aber: } (7 - 5) - 3 = -1. \end{aligned}$$

Es ist nicht unmittelbar einsichtig, dass Hintereinanderausführungen von Relationen assoziativ sind. Zur Einführung in das Rechnen mit Hintereinanderausführungen prüfen wir diese Aussage nach. Seien also A, B, C und D beliebig fixe Mengen, und $R \subseteq A \times B$, $S \subseteq B \times C$ und $T \subseteq C \times D$ drei beliebig fixe Relationen. Zu überprüfen ist die Aussage

$$(T \circ S) \circ R = T \circ (S \circ R).$$

Wir formen zuerst die linke Seite um. Mit der Definition der Verknüpfung von Relationen gilt

$$T \circ S = \{(b, d) \in B \times D \mid \exists_{c \in C} (b, c) \in S \wedge (c, d) \in T\}$$

und

$$\begin{aligned} (T \circ S) \circ R &= \{(a, d) \in A \times D \mid \exists_{b \in B} (a, b) \in R \wedge (b, d) \in (T \circ S)\} \\ &= \{(a, d) \in A \times D \mid \exists_{b \in B} \exists_{c \in C} (a, b) \in R \wedge (b, c) \in S \wedge (c, d) \in T\}. \end{aligned}$$

Durch Umformen der rechten Seite kommen wir zum selben Ergebnis:

$$\begin{aligned} T \circ (S \circ R) &= \{(a, d) \in A \times D \mid \exists_{c \in C} (a, c) \in (S \circ R) \wedge (c, d) \in T\} \\ &= \{(a, d) \in A \times D \mid \exists_{c \in C} \exists_{b \in B} (a, b) \in R \wedge (b, c) \in S \wedge (c, d) \in T\}. \end{aligned}$$

Damit ist nachgeprüft, dass die Hintereinanderausführung von Relationen assoziativ ist. Da dies *für alle* Relationen gilt, gilt es speziell für Funktionen und Permutationen, muss also für diese nicht nochmals nachgeprüft werden.

Bei vielen Operationen ist die Reihenfolge der Operanden vertauschbar; solche Operationen nennt man *kommutativ*. Für kommutative Operationen \circ gilt somit

$$a \circ b = b \circ a$$

für alle Elemente a, b der Trägermenge von \circ . Beispiel kommutativer Verknüpfungen sind Addition und Multiplikation auf den ganzen Zahlen. Nicht-kommutative Operationen sind etwa Subtraktion und Division auf den ganzen Zahlen, oder die Verknüpfung aus Beispiel 4.1.

Beispiel 4.5 Die Hintereinanderausführung auf P_n ist assoziativ, aber nicht kommutativ: Die Assoziativität ist oben bereits allgemeiner für Relationen nachgerechnet worden; für die Nicht-Kommutativität geben wir ein Gegenbeispiel an:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \text{aber}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}. \quad \square$$

Beispiel 4.6 Untenstehende Verknüpfungstabelle gibt eine Operation auf der Menge $\{a, b, c\}$ an, die zwar kommutativ, nicht aber assoziativ ist:

\circ	a	b	c
a	a	b	b
b	b	b	c
c	b	c	c

Die Kommutativität kann man Element für Element überprüfen. Man erkennt dann, dass eine Verknüpfung kommutativ ist, wenn ihre Tabelle symmetrisch ist, wenn also die Einträge links unterhalb der Hauptdiagonale gleich denen rechts oberhalb sind. Dies ist hier der Fall. Als Gegenbeispiel zur Assoziativität genügt etwa

$$(a \circ b) \circ c = c, \quad \text{aber} \quad a \circ (b \circ c) = b. \quad \square$$

Bei manchen Verknüpfungen gibt es Elemente der Trägermenge, die eine besondere Rolle spielen: Sie lassen die Elemente, mit denen sie verknüpft werden, unverändert. Diese speziellen Elemente werden *neutrale Elemente* genannt. Wie wir später sehen werden, kann es bei jeder Verknüpfung maximal ein neutrales Element geben.

Beispiel 4.7 Auf der Menge $\text{Pot}(\{0, 1\})$ mit der Verknüpfung \cap (siehe Beispiel 4.2) ist $\{0, 1\}$ das neutrale Element, da für alle $a \in \text{Pot}(\{0, 1\})$ gilt

$$a \cap \{0, 1\} = \{0, 1\} \cap a = a.$$

Auf der gleichen Menge mit der Verknüpfung \cup ist das neutrale Element die leere Menge, da für alle $a \in \text{Pot}(\{0, 1\})$ gilt

$$a \cup \emptyset = \emptyset \cup a = a. \quad \square$$

Beispiel 4.8 Für die Addition auf den ganzen Zahlen ist 0 das neutrale Element, da für jedes $x \in \mathbb{Z}$ gilt:

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

Für die Multiplikation ist 1 das neutrale Element: Es gilt

$$x \cdot 1 = 1 \cdot x = x. \quad \square$$

Bei Relationen, Funktionen und Permutationen ist für die Hintereinanderausführung die *identische Funktion* id das neutrale Element. Für jede Menge M ist diese Funktion definiert als

$$\begin{aligned} \text{id}_M : M &\rightarrow M \\ x &\mapsto x. \end{aligned}$$

Beispiel 4.9 Für P_n ist das neutrale Element die identische Permutation. So gilt etwa

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}. \end{aligned} \quad \square$$

Neutrale Elemente bezüglich einer Operation nehmen also eine spezielle Rolle in einer Menge ein. Bei manchen Verknüpfungen ist es möglich, zu jedem Element der Trägermenge a ein Element b zu finden, das bei Verknüpfung mit a das neutrale Element ergibt. Ein solches Element b nennt man ein zu a *inverses Element*.

Beispiel 4.10 Auf den ganzen Zahlen mit der Addition ist 0 das neutrale Element und $-a$ das zu a inverse Element, da gilt

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

Für die ganzen Zahlen mit der Multiplikation ist 1 das neutrale Element, es gibt jedoch für keine ganze Zahl x ($x \neq 1$) eine ganze Zahl x^{-1} mit

$$x \cdot x^{-1} = 1. \quad \square$$

Beispiel 4.11 Für jede Permutation gibt es ein inverses Element. Dieses erhält man, indem man die Permutation umdreht (von unten nach oben liest). So gilt etwa

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}. \end{aligned} \quad \square$$

Beispiel 4.12 Die Operation \cap auf $\text{Pot}(\{0, 1\})$, deren Verknüpfungstabelle in Beispiel 4.2 angegeben ist, besitzt das neutrale Element $\{0, 1\}$, aber keine Inverse (außer für $\{0, 1\}$): So gibt es etwa kein Element $x \in \text{Pot}(\{0, 1\})$ mit $x \cap \{0\} = \{0, 1\}$. \square

Wir sind nun in der Lage, die in diesem Abschnitt behandelten Konzepte zur algebraischen Struktur einer Gruppe zusammenzufassen.

Definition 4.3 (Gruppe)

Eine Menge G mit einer Verknüpfung \circ bezeichnet man als *Gruppe* $[G, \circ]$, wenn folgende Bedingungen erfüllt sind:

$$\forall_{a,b,c \in G} (a \circ b) \circ c = a \circ (b \circ c) \quad (\text{ASS})$$

$$\exists_{e \in G} \forall_{a \in G} a \circ e = e \circ a = a \quad (\text{NEU})$$

$$\forall_{a \in G} \exists_{a^{-1} \in G} a \circ a^{-1} = a^{-1} \circ a = e, \quad (\text{INV})$$

wobei e in (INV) das neutrale Element aus Bedingung (NEU) ist. Ist zusätzlich noch die Bedingung

$$\forall_{a,b \in G} a \circ b = b \circ a \quad (\text{KOMM})$$

erfüllt, nennt man $[G, \circ]$ eine *kommutative* (auch: *abelsche*) Gruppe.

Zur Notation einer Gruppe ist zu bemerken, dass je nach Trägermenge das neutrale Element der Gruppe mit 0 , 1 oder e bezeichnet wird. Ebenso wird, in Abhängigkeit der Gruppenoperation, das zu a inverse Element mit $-a$ oder a^{-1} bezeichnet. Oftmals verwendet man die abkürzende Schreibweise

$$a^n := \underbrace{a \circ a \circ \cdots \circ a}_{n\text{-mal}}.$$

Beispiel 4.13 Wie wir aus Beispiel 4.5, Beispiel 4.9 und Beispiel 4.11 sehen können, bildet für jedes $n \in \mathbb{N}$ die Menge der Permutationen P_n mit der Hintereinanderausführung eine Gruppe, die nicht kommutativ ist. \square

Beispiel 4.14 Wie man durch Prüfen der Bedingungen nachrechnen kann, bilden die in Beispiel 4.1 und Beispiel 4.6 angegebenen Verknüpfungstabellen auf der Menge $\{a, b, c\}$ keine Gruppen. Eine Verknüpfung, mit der diese Menge eine Gruppe bildet, ist die folgende:

\circ	a	b	c
a	c	a	b
b	a	b	c
c	b	c	a

Die Assoziativität muss für alle Kombinationen von Elementen einzeln nachgeprüft werden; wir verzichten an dieser Stelle darauf. Das neutrale Element von $[\{a, b, c\}, \circ]$

ist b , und die Inversen sind durch

$$a^{-1} = c, \quad b^{-1} = b, \quad c^{-1} = a$$

gegeben. □

Wir haben in diesem Abschnitt bereits einige Beispiele von Gruppen gesehen: So ist $[\mathbb{Z}, +]$ ebenso eine Gruppe wie $[P_n, \circ]$ für jedes n . Eine interessante Gruppe ist die in Abschnitt 3.4 besprochene Menge der Restklassen. Zur Wiederholung: Für $n \in \mathbb{N}$ sei

$$R_n := \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid n \mid x - y\}$$

die Relation zwischen ganzen Zahlen, die bei Division durch n den gleichen Rest haben. Es gilt etwa $(15, 26) \in R_{11}$ und $(-2, 8) \in R_5$. Mit der Notation

$$[x]_n := \{y \in \mathbb{Z} \mid (x, y) \in R_n\}$$

(die *Restklassen modulo n*) ist etwa

$$\begin{aligned} [4]_5 &= \{\dots, -6, -1, 4, 9, 14, \dots\} \\ [7]_{11} &= \{\dots, -15, -4, 7, 18, 29, \dots\}. \end{aligned}$$

Da es unendlich viele Vertreter für jede Restklassenmenge gibt (es ist etwa $[4]_5 = [-1]_5 = [9]_5$ und $[7]_{11} = [-4]_{11} = [18]_{11}$), verwendet man zur Vereinfachung der Darstellung meist die *Standardvertreter* $[0]_n, [1]_n, \dots, [n-1]_n$.

Über die Standardvertreter definieren wir die Menge

$$\mathbb{Z}_n := \{[0]_n, [1]_n, \dots, [n-1]_n\}$$

und auf dieser Menge eine Operation $+$:

$$\begin{aligned} + : \mathbb{Z}_n \times \mathbb{Z}_n &\rightarrow \mathbb{Z}_n \\ ([x]_n, [y]_n) &\mapsto [x]_n + [y]_n := [x + y]_n. \end{aligned}$$

Man beachte, dass die $+$ Operation auf \mathbb{Z}_n eine andere Operation als das $+$ auf \mathbb{Z} ist, obwohl wir dafür das gleiche Symbol verwenden.

Die so definierte Menge \mathbb{Z}_n mit obiger $+$ Operation bildet eine kommutative Gruppe, wie man unschwer nachprüfen kann: Die Assoziativität und Kommutativität von $+$ auf \mathbb{Z}_n folgt aus der Assoziativität und Kommutativität von $+$ auf \mathbb{Z} ; das neutrale Element ist $[0]_n$, und das zu $[x]_n$ inverse Element ist $[-x]_n = [n-x]_n$.

Beispiel 4.15 Die Verknüpfungstabelle von $[\mathbb{Z}_4, +]$ ist wie folgt, wobei wir zur Vereinfachung der Notation das Subskript weglassen:

$+$	[0]	[1]	[2]	[3]
[0]	[0]	[1]	[2]	[3]
[1]	[1]	[2]	[3]	[0]
[2]	[2]	[3]	[0]	[1]
[3]	[3]	[0]	[1]	[2]

□

Der große Vorteil des Abstrahierens zum Konzept einer Gruppe liegt darin, Rechengesetze nur einmal beweisen zu müssen: Diese können dann in allen Gruppen verwendet werden, ohne sie in jeder konkreten Instanz des Konzepts nochmals nachprüfen zu müssen. Wir fassen im Folgenden einige Rechengesetze in Gruppen zusammen.

Satz 4.1 Sei $[G, \circ]$ eine Gruppe. Dann gelten die folgenden Aussagen:

- (1) Es gibt nur ein neutrales Element in $[G, \circ]$.
- (2) Zu jedem a gibt es nur ein Inverses a^{-1} .
- (3) In Gruppen kann man Gleichungen durch Kürzen umformen: Es gilt für alle $a, b, c \in G$:

$$a \circ b = a \circ c \Rightarrow b = c$$

$$b \circ a = c \circ a \Rightarrow b = c.$$

- (4) $(a^{-1})^{-1} = a$.
- (5) $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.

Wir beweisen nur einige dieser Rechenregeln. Um etwa Aussage (i) nachzuprüfen, nimmt man die Existenz zweier neutraler Elemente e_1 und e_2 an und zeigt, dass diese gleich sein müssen: Da sowohl e_1 als auch e_2 neutral sind, gilt

$$e_1 \circ e_2 = e_1 \quad \text{und} \quad e_1 \circ e_2 = e_2,$$

und somit $e_1 = e_2$.

Die Aussage (iv) rechnet man durch folgende Umformung nach: Es gilt

$$a^{-1} \circ (a^{-1})^{-1} = e.$$

Nach Verknüpfen von links mit a erhält man

$$a \circ (a^{-1} \circ (a^{-1})^{-1}) = a \circ e,$$

mit der Assoziativität kann man die linke Seite anders klammern und erhält

$$\underbrace{(a \circ a^{-1})}_e \circ (a^{-1})^{-1} = \underbrace{a \circ e}_a,$$

und somit wie gefordert

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

Durch obige Rechenregeln können wir Gleichungen in beliebigen Gruppen lösen. Dazu kann man auch negative Exponenten verwenden, diese sind als

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

definiert. Die Gleichheit in obiger Definition ist etwa durch Induktion beweisbar.

Beispiel 4.16 Man löse die Gleichung

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}^2 \circ x = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}^{-3}$$

für $x \in P_4$. Erweitern von links liefert

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}^{-2} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}^2}_{\text{id}_{P_4}} \circ x = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}^{-2} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}^{-3}$$

und damit

$$\begin{aligned} x &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}^{-2} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}^{-3} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}. \end{aligned} \quad \square$$

Teilmengen von Gruppen, die bezüglich der Gruppenoperation abgeschlossen sind, bezeichnet man als *Untergruppen*.

Definition 4.4 (Untergruppe)

Sei $[G, \circ]$ eine Gruppe. Eine nichtleere Teilmenge $U \subseteq G$ mit der Verknüpfung \circ nennt man *Untergruppe*, wenn die beiden Bedingungen

$$(1) \quad \forall_{u_1, u_2 \in U} u_1 \circ u_2 \in U$$

$$(2) \quad \forall_{u \in U} u^{-1} \in U$$

erfüllt sind.

Aus dieser Definition folgt sofort, dass in jeder Untergruppe U das neutrale Element der Gruppe enthalten ist: Für $u \in U$ muss auch $u \circ u^{-1} = e$ in U sein.

Beispiel 4.17 Die Menge der geraden Zahlen bildet in $[\mathbb{Z}, +]$ eine Untergruppe, da die Addition und Inversenbildung auf dieser Menge abgeschlossen ist. Die Menge der ungeraden Zahlen bildet hingegen keine Untergruppe, da die Summe zweier ungerader Zahlen gerade ist (und auch das neutrale Element 0 nicht in dieser Teilmenge ist). \square

Beispiel 4.18 Die beiden Elemente

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

bilden in P_3 eine Untergruppe. \square

4.2 Ringe und Körper

Auf vielen Mengen ist mehr als eine Operation definiert, so z.B. auf \mathbb{Z} und \mathbb{Z}_n die Addition und Multiplikation, und auf $\text{Pot}(M)$ Durchschnitt und Vereinigung. Wir werden in diesem Abschnitt diejenigen Mengen und Verknüpfungen untersuchen, die Erweiterungen des Konzepts der Gruppe aus Abschnitt 4.1 sind.

Im Folgenden bildet die Trägermenge mit der ersten Verknüpfung eine kommutative Gruppe. An die zweite Verknüpfung werden wechselnde Anforderungen gestellt, zumindest aber muss sie assoziativ und mit der ersten Verknüpfung in einem gewissen Sinn verträglich sein. Diese Verträglichkeit wird *Distributivität* genannt und ist in der folgenden Definition der algebraischen Struktur *Ring* genauer spezifiziert.

Definition 4.5 (Ring)

Sei M eine Menge mit zwei Verknüpfungen $+$ und \cdot . Die Struktur $[M, +, \cdot]$ heißt *Ring*, wenn folgende Bedingungen erfüllt sind:

- (1) $[M, +]$ ist eine kommutative Gruppe.
- (2) Die Operation \cdot ist assoziativ auf M .
- (3) Die Operationen $+$ und \cdot erfüllen die *Distributivitätsgesetze*:

$$\begin{aligned} \forall_{a,b,c \in M} \quad a \cdot (b + c) &= (a \cdot b) + (a \cdot c) \\ \forall_{a,b,c \in M} \quad (a + b) \cdot c &= (a \cdot c) + (b \cdot c). \end{aligned}$$

Zu dieser Definition gibt es zwei gebräuchliche Erweiterungen: Wenn \cdot auf M kommutativ ist, nennt man die Struktur einen *kommutativen Ring*; wenn \cdot in M ein neutrales Element besitzt, nennt man die Struktur einen *Ring mit Einselement*. Auch bei Ringen, die keine Zahlenmengen sind, bezeichnet man das neutrale Element bezüglich $+$ als Nullelement 0 und das neutrale Element bezüglich \cdot als Einselement 1 .

Das klassische Beispiel eines Rings mit Einselement bilden die ganzen Zahlen.

Satz 4.2 Die Struktur $[\mathbb{Z}, +, \cdot]$ ist ein kommutativer Ring mit Einselement.

Ohne genaue Definition von $+$ und \cdot können wir diesen Satz nicht beweisen; wir verweisen aber auf die schon bekannt vorausgesetzten Rechenregeln für ganze Zahlen, die die Assoziativität und die Distributivitätsgesetze umfassen. Das Einselement ist die Zahl 1 .

Ebenso wie bei Gruppen gibt es bei Ringen Rechenregeln, die allgemein bewiesen werden können und dann für alle Ringe gelten. Wir erwähnen exemplarisch die Regel

$$a \cdot 0 = 0 \cdot a = 0.$$

Diese Gleichheit gilt, da

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

ist; die Gleichung

$$a \cdot 0 = a \cdot 0 + a \cdot 0$$

wird dann durch Addition des Inversen von $a \cdot 0$ auf die Form $0 = a \cdot 0$ gebracht.

Ein Beispiel eines Rings über einer endlichen Menge bilden die *Restklassenringe*, deren additive Gruppenstruktur wir bereits in Abschnitt 4.1 kennengelernt haben.

Beispiel 4.19 Die Struktur $[\mathbb{Z}_4, +, \cdot]$ ist ein kommutativer Ring mit Einselement. Die Verknüpfungstabelle von $[\mathbb{Z}_4, +]$ ist in Beispiel 4.15 angegeben; die Verknüpfungstabelle von $[\mathbb{Z}_4, \cdot]$ ist

\cdot	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

Man erkennt, dass [1] das Einselement ist; die Assoziativität dieser Tabelle folgt aus der Assoziativität der Multiplikation auf \mathbb{Z} ; gleiches gilt für die Distributivität. \square

Wir werden nach der folgenden Definition der Struktur *Körper* noch weitere Ringe kennenlernen. Körper ergeben sich aus kommutativen Ringen, wenn man zusätzlich zum Einselement noch die Existenz von Inversen fordert.

Definition 4.6 (Körper)

Sei $[M, +, \cdot]$ ein kommutativer Ring mit Einselement. Dann nennt man $[M, +, \cdot]$ einen *Körper*, wenn $[M \setminus \{0\}, \cdot]$ eine kommutative Gruppe ist.

Ein Körper ist somit eine Menge, die mit zwei Operationen eine kommutative Gruppe bildet, und deren Operationen über die Distributivitätsgesetze kompatibel sind. Zu beachten ist, dass für die zweite Operation das Nullelement ausgenommen werden muss, da dafür keine Inversen existieren.

Die Standardbeispiele von Körpern sind die rationalen Zahlen \mathbb{Q} und die reellen Zahlen \mathbb{R} , die wir als bekannt voraussetzen. Beide Mengen können aus den ganzen Zahlen hergeleitet werden; dies erfordert für die reellen Zahlen Konzepte, über die wir noch nicht verfügen.

Wir skizzieren daher kurz die Idee der Herleitung von \mathbb{Q} aus \mathbb{Z} . Sei dafür die Äquivalenzrelation \sim auf der Menge $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ definiert als

$$(a, b) \sim (c, d) :\Leftrightarrow a \cdot d = b \cdot c.$$

Es ist nicht schwer nachzuprüfen, dass \sim tatsächlich eine Äquivalenzrelation ist. Wenn wir dann noch für $b \neq 0$ die Schreibweise $\frac{a}{b}$ als die Menge aller Äquivalenzklassen von (a, b)

$$\frac{a}{b} := [(a, b)]_{\sim} = \{(c, d) \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\} \mid (a, b) \sim (c, d)\}$$

eingeführen, können wir die rationalen Zahlen \mathbb{Q} als

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

definieren. Auf dieser Menge legen wir folgende Addition fest:

$$\begin{aligned} + : \mathbb{Q} \times \mathbb{Q} &\rightarrow \mathbb{Q} \\ \left(\frac{a}{b}, \frac{c}{d} \right) &\mapsto \frac{a}{b} + \frac{c}{d} := \frac{a \cdot d + b \cdot c}{b \cdot d}. \end{aligned}$$

Mit dieser Addition ist \mathbb{Q} eine kommutative Gruppe; das neutrale Element ist $\frac{0}{1}$, das zu $\frac{a}{b}$ inverse Element ist $\frac{-a}{b}$:

$$\frac{a}{b} + \frac{0}{1} = \frac{a \cdot 1 + b \cdot 0}{b \cdot 1} = \frac{a}{b}, \quad \text{und} \quad \frac{a}{b} + \frac{-a}{b} = \frac{a \cdot b + b \cdot (-a)}{b \cdot b} = \frac{0}{b^2} = \frac{0}{1}.$$

Die letzte Gleichheit gilt, da $\frac{0}{b^2}$ und $\frac{0}{1}$ zwei verschiedene Vertreter derselben Äquivalenzklasse sind.

Auf \mathbb{Q} definieren wir die Multiplikation als

$$\begin{aligned} \cdot : \mathbb{Q} \times \mathbb{Q} &\rightarrow \mathbb{Q} \\ \left(\frac{a}{b}, \frac{c}{d} \right) &\mapsto \frac{a}{b} \cdot \frac{c}{d} := \frac{a \cdot c}{b \cdot d}. \end{aligned}$$

Die Menge \mathbb{Q} ist ohne das Nullelement der Addition (also ohne $\frac{0}{1}$) mit obiger Multiplikation ebenfalls eine kommutative Gruppe: Das neutrale Element ist $\frac{1}{1}$, das zu $\frac{a}{b}$ inverse Element ist $\frac{b}{a}$:

$$\frac{a}{b} \cdot \frac{1}{1} = \frac{a \cdot 1}{b \cdot 1} = \frac{a}{b}, \quad \text{und} \quad \frac{a}{b} \cdot \frac{b}{a} = \frac{a \cdot b}{b \cdot a} = \frac{1}{1};$$

wobei die letzte Gleichheit mit dem gleichen Argument wie bei der Überprüfung der Addition gilt.

Wir verzichten auf Überprüfung von Assoziativität, Kommutativität, und Distributivität, und halten fest:

Satz 4.3 Die Strukturen $[\mathbb{Q}, +, \cdot]$ und $[\mathbb{R}, +, \cdot]$ sind Körper.

Wir werden in Abschnitt 4.3 ein weiteres Beispiel einer unendlichen Zahlenmenge mit Körperstruktur kennenlernen, und in Abschnitt 4.4 ein wichtiges Beispiel eines kommutativen Rings mit Einselement. Zuvor betrachten wir noch einen Körper, der nur endlich viele Elemente enthält.

Wie wir bereits wissen, bilden die Restklassen \mathbb{Z}_n für alle n einen kommutativen Ring mit Einselement; somit unterscheiden sich diese Strukturen nur mehr durch die fehlende Inversenbildung bezüglich der Multiplikation von Körpern. Wir untersuchen die multiplikativen Inversen anhand zweier Beispiele.

Beispiel 4.20 Die Verknüpfungstabellen von $[\mathbb{Z}_4, \cdot]$ und $[\mathbb{Z}_5, \cdot]$ sind:

\cdot	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

\cdot	[0]	[1]	[2]	[3]	[4]
[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]
[2]	[0]	[2]	[4]	[1]	[3]
[3]	[0]	[3]	[1]	[4]	[2]
[4]	[0]	[4]	[3]	[2]	[1]

Man erkennt, dass in $[\mathbb{Z}_4, \cdot]$ $[1]^{-1} = [1]$ und $[3]^{-1} = [3]$ gilt, es aber für 2 kein multiplikatives Inverses gibt. Somit ist $[\mathbb{Z}_4, \cdot]$ kein Körper. Im Gegensatz dazu gibt es in $[\mathbb{Z}_5, \cdot]$ für alle Elemente außer [0] ein Inverses, das man aus der Tabelle ablesen kann.

Der Unterschied zwischen \mathbb{Z}_4 und \mathbb{Z}_5 liegt darin, dass $4 = 2 \cdot 2$ ist, 4 also ein Vielfaches von 2 ist. Damit gibt es kein $[x] \in \mathbb{Z}_4$, sodass $[2] \cdot [x] = [2 \cdot x]$ nach Division durch 4 Rest 1 hat. Genau dieses x wäre aber das multiplikative Inverse von [2] in \mathbb{Z}_4 . In \mathbb{Z}_5 kann diese Situation nicht auftreten, da 5 eine Primzahl ist und somit keine Teiler hat. \square

Wir fassen diese Erkenntnis in folgendem Satz zusammen.

Satz 4.4 Die Struktur $[\mathbb{Z}_n, +, \cdot]$ ist genau dann ein Körper, wenn n eine Primzahl ist.

Wir wenden uns nun einem weiteren Körper mit unendlich vielen Elementen zu.

4.3 Komplexe Zahlen

Die Erweiterung der Zahlenbereiche von \mathbb{N} über \mathbb{Z} zu \mathbb{Q} und \mathbb{R} war dadurch motiviert, dass man in manchen Bereichen bestimmte Gleichungen nicht lösen konnte. So hat etwa

$$x + 1 = 0$$

in \mathbb{N} keine Lösung, wohl aber in \mathbb{Z} ; ebenso ist

$$2 \cdot x = 1$$

nicht in \mathbb{Z} , sondern erst in \mathbb{Q} lösbar. Als man erkannte, dass sich die Gleichung

$$x^2 = 2$$

nicht über \mathbb{Q} lösen lässt, motivierte dies die Einführung der reellen Zahlen. Der letzte Schritt, den wir in diesem Abschnitt vollziehen wollen, ist durch die Unlösbarkeit der Gleichung

$$x^2 + 1 = 0$$

in \mathbb{R} gegeben. Mit der Einführung komplexer Zahlen \mathbb{C} erhalten wir dann eine lineare Ordnung auf den Zahlenbereichen:

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

So wie rationale Zahlen sich als Paare von ganzen Zahlen darstellen lassen, so kann man komplexe Zahlen als Paare von reellen Zahlen auffassen. Wir werden nach dieser Definition zwei weitere Darstellungsarten komplexer Zahlen kennenlernen, mit denen sich leichter rechnen lässt.

Wir definieren die komplexen Zahlen \mathbb{C} als $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ mit folgenden Verknüpfungen:

$$\begin{aligned} + : \mathbb{C} \times \mathbb{C} &\rightarrow \mathbb{C} \\ ((a, b), (c, d)) &\mapsto (a, b) + (c, d) := (a + c, b + d) \\ \\ \cdot : \mathbb{C} \times \mathbb{C} &\rightarrow \mathbb{C} \\ ((a, b), (c, d)) &\mapsto (a, b) \cdot (c, d) := (a \cdot c - b \cdot d, a \cdot d + b \cdot c) \end{aligned}$$

Wie man erkennen kann, ist $[\mathbb{C}, +]$ eine kommutative Gruppe mit $(0, 0)$ als neutralem Element und $(-a, -b)$ dem zu (a, b) Inversen. Nicht sofort einsichtig ist, dass auch $[\mathbb{C} \setminus \{(0, 0)\}, \cdot]$ eine kommutative Gruppe ist: Das Einselement ist $(1, 0)$, da gilt:

$$(a, b) \cdot (1, 0) = (a \cdot 1 - b \cdot 0, a \cdot 0 + b \cdot 1) = (a, b).$$

Das zu (a, b) multiplikative Inverse ist $(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$:

$$\begin{aligned} (a, b) \cdot \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right) \\ = \left(a \cdot \frac{a}{a^2+b^2} - b \cdot \frac{-b}{a^2+b^2}, a \cdot \frac{-b}{a^2+b^2} + b \cdot \frac{a}{a^2+b^2} \right) \\ = (1, 0). \end{aligned}$$

Wir prüfen Assoziativität, Kommutativität und Distributivität nicht nach, sondern halten fest.

Satz 4.5 Die Struktur $[\mathbb{C}, +, \cdot]$ ist ein Körper.

Da die Multiplikationsregel nicht einfach zu merken ist, und auch nicht durch die Lösung der Gleichung $x^2 + 1 = 0$ motiviert scheint, verwenden wir im Folgenden eine etwas einleuchtendere Zahlendarstellung. Wir schreiben

$$a + i b := (a, b)$$

wobei wir i als Symbol mit der Eigenschaft $i^2 = -1$ betrachten. Man nennt $a \in \mathbb{R}$ den *Realteil*, und $b \in \mathbb{R}$ den *Imaginärteil* der komplexen Zahl $a + i b$. Wenn wir zwei komplexe Zahlen in dieser Schreibweise multiplizieren (also als Summe reeller Zahlen, unter Berücksichtigung von $i^2 = -1$), erhalten wir obige Definition der Multiplikation:

$$\begin{aligned} (a + i b) \cdot (c + i d) &= a \cdot c + i(b \cdot c) + i(a \cdot d) + i^2 b \cdot d \\ &= a \cdot c - b \cdot d + i(b \cdot c + a \cdot d). \end{aligned}$$

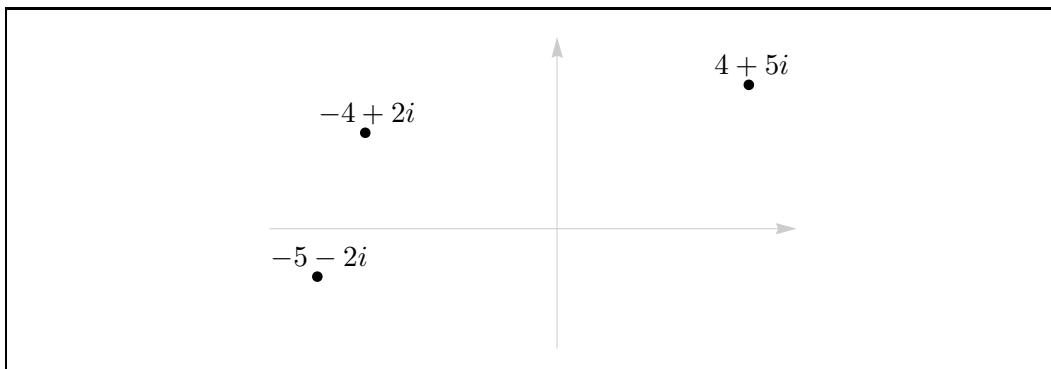


Abbildung 4.1: Beispiele komplexer Punkte in der komplexen Zahlenebene.

Da komplexe Zahlen aus zwei reellen Komponenten bestehen, visualisiert man sie in der *komplexen Zahlenebene*. Dabei ist die x -Achse die *reelle Achse*, und die y -Achse die *imaginäre Achse*. Abbildung 4.1 zeigt die Lage dreier komplexer Zahlen in der komplexen Zahlenebene.

Beispiel 4.21 Seien $z_1 = 3 + 4i$ und $z_2 = 5 - 2i$. Dann ist

$$z_1 + z_2 = 8 + 2i$$

$$z_1 - z_2 = -2 + 6i$$

$$z_1 \cdot z_2 = 23 + 14i$$

$$\frac{z_1}{z_2} = \frac{(3 + 4i) \cdot (5 + 2i)}{(5 - 2i)(5 + 2i)} = \frac{7 + 26i}{29}. \quad \square$$

Die im letzten Beispiel bei der Division auftretende komplexe Zahl $a - ib$ wird die zu $a + ib$ *konjugiert komplexe Zahl* genannt. Allgemein dividiert man komplexe Zahlen, indem man mit der zum Nenner konjugiert komplexen Zahl erweitert, um so den Nenner reell zu machen.

Neben der Darstellung komplexer Zahlen als $a + ib$, die *kartesische Koordinatendarstellung* genannt wird, können komplexe Zahlen auch eindeutig durch ihren Abstand (Radius) r vom Ursprung und Winkel φ von der x -Achse bestimmt werden. Diese Form der Darstellung, die zur Unterscheidung von kartesischen Koordinaten oft in eckigen Klammern als $[r, \varphi]$ geschrieben wird, nennt man *Polarkoordinaten*.

Aus der einfachen Visualisierung in Abbildung 4.2 kann man die Umrechnungsformeln zwischen kartesischen und Polarkoordinaten herauslesen. Das Rechteck mit Seitenlängen a , b und r ist rechtwinklig, somit gilt

$$r = \sqrt{a^2 + b^2}$$

$$\varphi = \arctan \frac{b}{a}.$$

Bei der Berechnung des Winkels muss man die Vorzeichen von a und b berücksichtigen, da $\arctan \frac{b}{a} = \arctan \frac{-b}{-a}$ und $\arctan \frac{-b}{a} = \arctan \frac{b}{-a}$ ist.

Umgekehrt kann man aus gegebenen Polarkoordinaten $[r, \varphi]$ leicht die kartesischen Koordinaten $a + ib$ ausrechnen:

$$a = r \cos \varphi$$

$$b = r \sin \varphi.$$

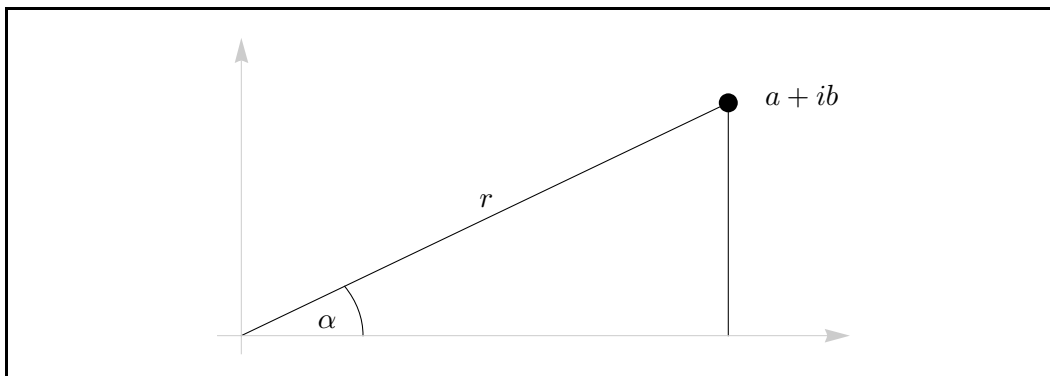


Abbildung 4.2: Zusammenhang zwischen kartesischen Koordinaten und Polarkoordinaten.

Mit diesen Formeln erhält man die für viele Rechnungen vorteilhafte Darstellung einer komplexen Zahl $z = a + ib$ als

$$z = r(\cos \varphi + i \sin \varphi).$$

Beispiel 4.22 Die komplexe Zahl $1 + 3i$ hat die Polarkoordinaten

$$r = \sqrt{1^2 + 3^2} = \sqrt{10}$$

$$\varphi = \arctan \frac{3}{1} = 1.249 = 71.5651^\circ.$$

Die komplexe Zahl mit Polarkoordinaten $[4, \frac{\pi}{6}]$ hat die kartesischen Koordinaten

$$a = 4 \cos \frac{\pi}{6} = 4 \frac{\sqrt{3}}{2} = 2\sqrt{3}$$

$$b = 4 \sin \frac{\pi}{6} = 4 \frac{1}{2} = 2. \quad \square$$

Das Umrechnen auf Polarkoordinaten bringt auf den ersten Blick keinen Vorteil mit sich. Tatsächlich bietet diese Darstellung aber für das Multiplizieren und Potenzieren zweier komplexer Zahlen gegenüber dem Ausmultiplizieren bei kartesischer Darstellung einiges an Rechenersparnis. Es gilt nämlich

$$(r_1(\cos \varphi_1 + i \sin \varphi_1)) \cdot (r_2(\cos \varphi_2 + i \sin \varphi_2))$$

$$= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2));$$

diese Formel kann man unter Verwendung trigonometrischer Rechenregeln nachprüfen. Die Multiplikation zweier komplexer Zahlen hat somit eine einfache geometrische Veranschaulichung: Für das Produkt zweier Zahlen z_1 und z_2 ist der Radius das Produkt der Radien von z_1 und z_2 , und der Winkel die Summe der Winkel von z_1 und z_2 .

Beispiel 4.23 Zur Berechnung von $(1 - 2i)^{10}$ ist es sinnvoll, zuerst auf Polarkoordinaten umzurechnen, und dann in dieser Darstellung zu potenzieren. Es ist $r = \sqrt{5}$ und $\varphi = \arctan(-2) = -1.10715$. Somit ist

$$(1 - 2i) = \sqrt{5}(\cos(-1.10715) + i \sin(-1.10715))$$

und

$$\begin{aligned} (1 - 2i)^{10} &= \sqrt{5}^{10} (\cos(-11.0715) + i \sin(-11.0715)) \\ &= 3125(0.07584 + 0.99712i) \\ &= 237 + 3116i. \end{aligned} \quad \square$$

Für die Addition zweier komplexer Zahlen in Polarkoordinaten gibt es keine einfache Formel; man rechnet daher zuerst in kartesische Koordinaten um.

Beispiel 4.24 Die Summe der beiden komplexen Zahlen mit Polarkoordinaten $[3, \frac{\pi}{2}]$ und $[2, \frac{\pi}{3}]$ ist nach Umrechnen in kartesische Koordinaten

$$\begin{aligned} 3\left(\cos \frac{\pi}{2} + i \sin \frac{\pi}{2}\right) + 2\left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3}\right) &= 3i + 2\left(\frac{1}{2} + i \frac{\sqrt{3}}{2}\right) \\ &= 1 + (3 + \sqrt{3})i. \end{aligned} \quad \square$$

Zusammenfassend kann man also sagen, dass sich die Darstellung in kartesischen Koordinaten besser für die Addition, und die Darstellung in Polarkoordinaten besser für die Multiplikation eignet.

Der Vorteil der Polarkoordinatendarstellung ist bei der Umkehrung des Potenzierens, also dem *Wurzelziehen*, besonders evident. Wir überlegen uns: Da für $z := r(\cos \varphi + i \sin \varphi)$ die n -te Potenz einfach

$$z^n = r^n (\cos(n\varphi) + i \sin(n\varphi)),$$

ist, ist in Umkehrung die n -te Wurzel von z durch eine Zahl

$$\sqrt[n]{z} = s(\cos \psi + i \sin \psi)$$

gegeben. Für die beiden Koordinaten $[s, \psi]$ gilt

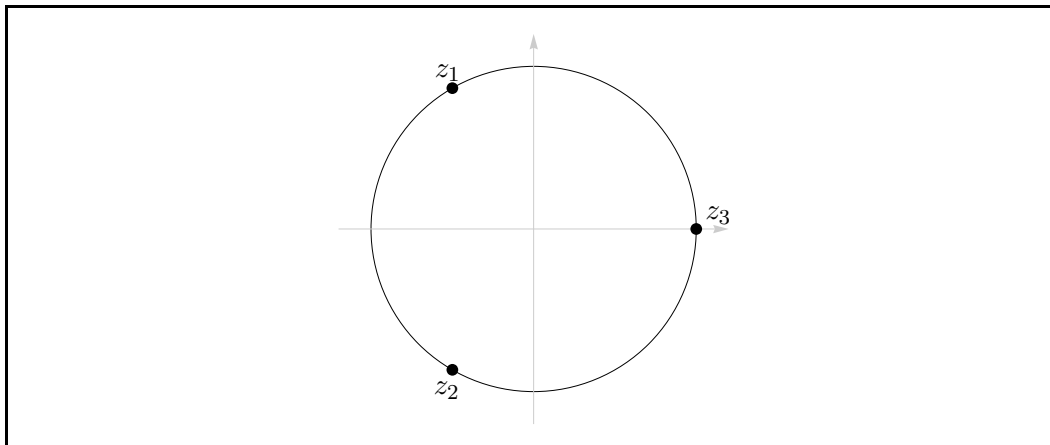
$$s^n = r \quad \text{und} \quad n\psi = \varphi.$$

Aus der ersten Bedingung ergibt sich sofort $s = \sqrt[n]{r}$. Bei der zweiten muss man berücksichtigen, dass Winkel nur als Reste bei Division durch 2π eindeutig sind, dass also für alle Winkel φ

$$\varphi = \varphi + k2\pi,$$

gilt, wobei $k \in \mathbb{Z}$ beliebig ist. Somit ist die Lösung der Gleichung

$$n\psi = \varphi = \varphi + k2\pi$$

Abbildung 4.3: Die Lösungen der Gleichung $z^3 = 1$ aus Beispiel 4.25.

durch

$$\psi = \frac{\varphi}{n} + k \frac{2\pi}{n}$$

gegeben. Man kann sich überlegen, dass diese Formel nur für $k = 0, \dots, n - 1$ verschiedene Werte liefert, da für $k = n$ der gleiche Wert wie für $k = 0$ erhalten wird.

Zur Erläuterung dieser Herleitung betrachten wir ein einfaches Beispiel.

Beispiel 4.25 Gesucht sind alle Lösungen der Gleichung

$$z^3 = 1,$$

die in den reellen Zahlen ja nur die Lösung $z = 1$ hat. In Polarkoordinaten ausgedrückt gilt

$$z^3 = 1(\cos 0 + i \sin 0) = 1(\cos 2\pi + i \sin 2\pi).$$

Mit obigen Überlegungen erhalten wir für $k = 0, 1, 2$ drei verschiedene Lösungen:

$$\begin{aligned} z_1 &= 1 \left(\cos \left(\frac{2\pi}{3} + 0 \frac{2\pi}{3} \right) + i \sin \left(\frac{2\pi}{3} + 0 \frac{2\pi}{3} \right) \right) = -\frac{1}{2} + \frac{\sqrt{3}}{2} i \\ z_2 &= 1 \left(\cos \left(\frac{2\pi}{3} + 1 \frac{2\pi}{3} \right) + i \sin \left(\frac{2\pi}{3} + 1 \frac{2\pi}{3} \right) \right) = -\frac{1}{2} - \frac{\sqrt{3}}{2} i \\ z_3 &= 1 \left(\cos \left(\frac{2\pi}{3} + 2 \frac{2\pi}{3} \right) + i \sin \left(\frac{2\pi}{3} + 2 \frac{2\pi}{3} \right) \right) = 1. \end{aligned}$$

Da sich der Radius 1 durch Wurzelziehen nicht verändert, liegen alle Lösungen am Einheitskreis. Die Lage der drei Lösungen ist in Abbildung 4.3 dargestellt. \square

Wie wir in Verallgemeinerung dieses Beispiels sehen können, besitzt die Gleichung $z^n = 1$ über \mathbb{C} n verschiedene Lösungen, die *Einheitswurzeln* genannt werden. Für die Wurzeln anderer Zahlen ergibt sich ein ähnliches Bild: die komplexen Lösungen von $z^n = a$ liegen immer gleichmäßig auf einem Kreis in der komplexen Zahlenebene verteilt.

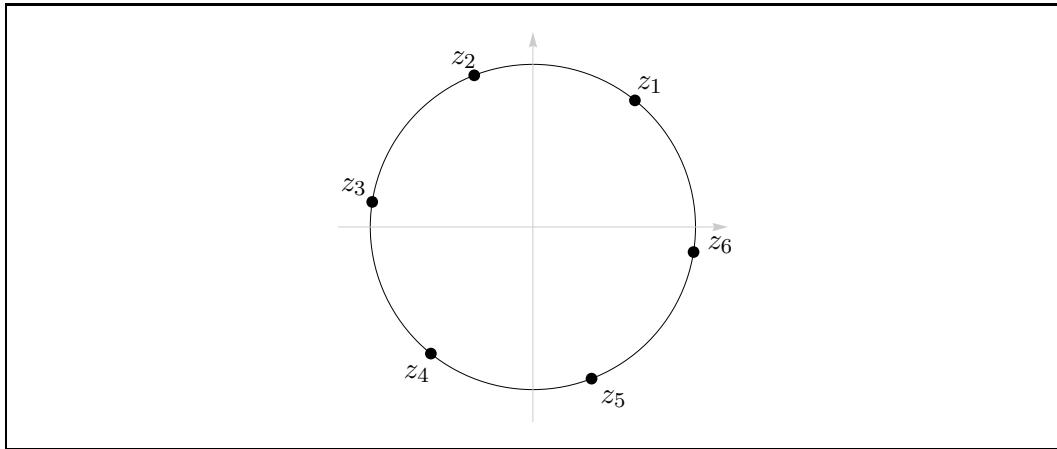


Abbildung 4.4: Die Lösungen der Gleichung $z^6 = 3 - 4i$ aus Beispiel 4.26.

Beispiel 4.26 Zu bestimmen seien alle Lösungen der Gleichung

$$z^6 = 3 - 4i.$$

Die Polarkoordinaten von $3 - 4i$ sind $[5, -0.9273] = [5, 5.3559]$. Damit sind die sechs Lösungen

$$\begin{aligned} z_1 &= \sqrt[6]{5}(\cos(0.8926) + i \sin(0.8926)) = 0.8204 + 1.0183 i \\ z_2 &= \sqrt[6]{5}(\cos(1.9398) + i \sin(1.9398)) = -0.4717 + 1.2196 i \\ z_3 &= \sqrt[6]{5}(\cos(2.9870) + i \sin(2.9870)) = -1.2921 + 0.2013 i \\ z_4 &= \sqrt[6]{5}(\cos(4.0342) + i \sin(4.0342)) = -0.8204 - 1.0183 i \\ z_5 &= \sqrt[6]{5}(\cos(5.0814) + i \sin(5.0814)) = 0.4717 - 1.2196 i \\ z_6 &= \sqrt[6]{5}(\cos(6.1287) + i \sin(6.1287)) = 1.2921 - 0.2013 i. \end{aligned}$$

Eine Graphik dieser sechs Lösungen ist in Abbildung 4.4 zu sehen. □

4.4 Polynome

Nach der genaueren Betrachtung des speziellen Körpers der komplexen Zahlen in Abschnitt 4.3 wenden wir uns nun *Polynomen* zu, einem Beispiel eines kommutativen Rings mit Einselement.

Definition 4.7 (Polynome)

Sei K ein Körper und $a_0, \dots, a_n \in K$ mit $a_n \neq 0$. Dann nennt man den Ausdruck

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

ein *Polynom* über K , und n den *Grad* des Polynoms. Die Zahlen a_0, \dots, a_n heißen *Koeffizienten* des Polynoms; a_n heißt *Leitkoeffizient*. Die Menge aller Polynome über K wird als $K[x]$ bezeichnet.

Die Addition auf $K[x]$ ist (hier für $n > m$) als

$$(a_n x^n + \cdots + a_1 x + a_0) + (b_m x^m + \cdots + b_1 x + b_0) = \\ a_n x^n + \cdots + (a_m + b_m) x^m + \cdots + (a_1 + b_1) x + (a_0 + b_0),$$

definiert, das Aufsummieren der Koeffizienten bei gleichen Potenzen von x . Mit dieser Addition bildet $K[x]$ eine kommutative Gruppe; das neutrale Element ist das Nullpolynom 0; das zu $a_n x^n + \cdots + a_1 x + a_0$ inverse Polynom ist $-a_n x^n - \cdots - a_1 x - a_0$.

Die Multiplikation auf $K[x]$ wird durch stückweises Ausmultiplizieren definiert:

$$(a_n x^n + \cdots + a_1 x + a_0) \cdot (b_m x^m + \cdots + b_1 x + b_0) = \\ a_n b_m x^{n+m} + \cdots + (a_0 b_1 + a_1 b_0) x + a_0 b_0.$$

Das neutrale Element bezüglich der Multiplikation ist das Einspolynom 1. Mit diesen beiden Operationen bildet $K[x]$ einen kommutativen Ring mit Einselement.

Ein bereits erwähnter Vorteil des Klassifizierens von Mengen und Operationen anhand ihrer algebraischen Struktur ist es, Rechenregeln und Algorithmen für Strukturen herzuleiten. Diese Rechenregeln und Algorithmen sind dann für alle Vertreter dieser Strukturen gültig bzw. anwendbar. Wir betrachten als Beispiel für dieses Vorgehen die Division mit Rest, die für ganze Zahlen bereits in Satz 3.2 erwähnt wurde. Anstatt den Algorithmus formal anzugeben, betrachten wir ein Beispiel mit ganzen Zahlen.

Wir bestimmen die Zahlen q und r so, dass $541 = 37 \cdot q + r$ ist (mit $0 \leq r < 37$):

$$\begin{array}{r} 541: 37 = 14 \\ \underline{-37} \\ 171 \\ \underline{-148} \\ 23 \end{array}$$

An dieser Stelle wird der Algorithmus abgebrochen, da die Bedingung $0 \leq r < 37$ erfüllt ist. Die gesuchten Zahlen sind somit $q = 14$ und $r = 23$. Die bei diesem Algorithmus auftretenden Rechenoperationen waren Multiplikation und Subtraktion (also Inversenbildung bezüglich der Addition). Diese Operationen sind in jedem Ring möglich, sodass sich der gleiche Algorithmus auch auf Polynome anwenden lässt.

Dabei ist zu beachten, dass wir für die Abbruchbedingung eine $<$ -Relation auf $K[x]$ definieren müssen; diese ist durch

$$\forall_{P, Q \in K[x]} P < Q :\Leftrightarrow \text{grad}(P) < \text{grad}(Q)$$

gegeben. Jetzt können wir auch in $K[x]$ mit Rest dividieren, wie das folgende Beispiel zeigt.

Beispiel 4.27 Man bestimme zu $P_1 = 4x^3 + 2x^2 - 3x + 4$ und $P_2 = 2x^2 + 1$ aus $\mathbb{Q}[x]$ zwei Polynome Q und R , für die $P_1 = Q \cdot P_2 + R$ und $\text{grad}(R) < \text{grad}(P_2)$ gilt. Man erhält mit dem Divisionsalgorithmus die Rechenschritte

$$\begin{array}{r}
4x^3 + 2x^2 - 3x + 4 : 2x^2 + 1 = 2x + 1 \\
\underline{\pm 4x^3 \qquad \pm 2x} \\
+ 2x^2 - 5x + 4 \\
\underline{\pm 2x^2 \qquad \pm 1} \\
-5x + 3
\end{array}$$

und als Ergebnis $Q = 2x + 1$ und $R = -5x + 3$. \square

Die komplexen Zahlen wurden eingeführt, um Gleichungen der Art $x^2 + 1 = 0$ lösen zu können. Tatsächlich lassen sich *alle* polynomialen Gleichungen (zumindest theoretisch) über \mathbb{C} lösen. Dies besagt auch der folgende Satz, der eine bedeutende Stellung in der Mathematik einnimmt, mit unseren Methoden aber nicht zu beweisen ist.

Satz 4.6 (Fundamentalsatz der Algebra) Jede Gleichung der Form

$$a_n x^n + \dots + a_1 x + a_0 = 0$$

mit $n \geq 1$, $a_n, \dots, a_0 \in \mathbb{C}$ und $a_n \neq 0$ hat über \mathbb{C} mindestens eine Lösung.

Eine Lösung einer solchen Gleichung nennt man auch *Nullstelle* der durch die linke Seite der Gleichung definierten Funktion. So hat die Funktion $P(x) := x^2 - x - 6$ die Nullstellen $\xi = -2$ und $\xi = 3$, da $P(-2) = P(3) = 0$ ist.

Für Polynomfunktionen gibt es in Spezialfällen Formeln, mit denen die Nullstellen bestimmt werden können; diese Spezialfälle sind schnell aufgezählt. Seien dafür $a_0, a_1, a_2 \in \mathbb{C}$. Für $P(x) := a_1 x + a_0$ ist

$$\xi_0 = -\frac{a_0}{a_1}$$

die einzige Nullstelle, und für $P(x) = a_2 x^2 + a_1 x + a_0$ sind

$$\xi_0 = \frac{-a_1 - \sqrt{a_1^2 - 4a_2 a_0}}{2a_2} \quad \text{und} \quad \xi_1 = \frac{-a_1 + \sqrt{a_1^2 - 4a_2 a_0}}{2a_2}$$

die einzigen Nullstellen. Für Polynome 3. und 4. Grades gibt es ähnliche (wenn auch kompliziertere) Formeln. Es kann aber gezeigt werden, dass es für höhergradige Polynomfunktionen keine geschlossene Lösungsformel für die Nullstellenfindung geben kann. Die Lösungsformeln liefern für Polynomfunktionen 3. Grades drei, für solche 4. Grades vier mögliche Nullstellen, die allerdings nicht alle verschieden sein müssen. Folgender Satz besagt, dass auch allgemein ein Zusammenhang zwischen Grad der Polynomfunktion und Anzahl der Nullstellen besteht.

Satz 4.7 Sei $P := a_n x^n + \dots + a_1 x + a_0$ ein Polynom über \mathbb{C} mit Grad n . Dann hat P höchstens n Nullstellen in \mathbb{C} .

Zur Begründung dieses Satzes verwenden wir das Ergebnis, dass man Nullstellen aus einem Polynom “herausdividieren” kann.

Satz 4.8 Sei P ein Polynom über \mathbb{C} mit $\text{grad}(P) \geq 1$ und $\xi \in \mathbb{C}$ eine Nullstelle der durch P gegebenen Polynomfunktion. Dann gibt es ein Polynom S mit $\text{grad}(S) = \text{grad}(P) - 1$ über \mathbb{C} , für das gilt:

$$P = (x - \xi) \cdot S.$$

Wenn ξ eine Nullstelle von P ist, so nennt man $(x - \xi)$ einen *Linearfaktor* von P .

Man kann Satz 4.8 leicht beweisen: Wie wir bereits wissen, kann man durch Polynomdivision schreiben:

$$P = (x - \xi) \cdot S + R,$$

wobei $\text{grad}(R) < \text{grad}(x - \xi) = 1$ ist. Somit ist $\text{grad}(R) = 0$, R also eine Konstante. Den Wert dieser Konstanten erhält man, indem man auf beiden Seiten der Gleichung für x den Wert ξ einsetzt:

$$P(\xi) = (\xi - \xi) \cdot S(\xi) + R.$$

Aus $P(\xi) = 0$ folgt dann $R = 0$, womit Satz 4.8 bewiesen ist. Das Polynom S kann man mit dem Algorithmus der Polynomdivision bestimmen, der in Beispiel 4.27 vorgestellt wurde.

Da das Herausdividieren von Nullstellen für ein Polynom vom Grad n höchstens n -mal möglich ist, ergibt sich Satz 4.7 direkt aus Satz 4.6 und Satz 4.8.

Beispiel 4.28 Durch Skizzieren oder Probieren “errät” man die Nullstelle $\xi = 3$ der Polynomfunktion

$$P(x) = x^3 - 7x - 6.$$

Division durch $(x - 3)$ liefert

$$\begin{array}{r} x^3 \quad -7x - 6: x - 3 = x^2 + 3x + 2 \\ \underline{\pm x^3 \mp 3x^2} \\ \quad + 3x^2 - 7x \\ \quad \underline{\pm 3x^2 \mp 9x} \\ \qquad \quad 2x - 6 \\ \qquad \quad \underline{\pm 2x \mp 6} \\ \qquad \qquad \qquad 0 \end{array}$$

Das gesuchte S aus Satz 4.8 ist also $x^2 + 3x + 2$, dessen Nullstellen mit der Formel für quadratische Gleichungen bei $\xi = -1$ und $\xi = -2$ liegen. Durch wiederholtes Anwenden von Satz 4.8 können wir $P(x)$ als Produkt seiner Linearfaktoren schreiben; man erhält

$$P(x) = (x - 3) \cdot (x + 1) \cdot (x + 2). \quad \square$$

Die im letzten Beispiel gezeigte Struktur eines Polynoms als Produkt von Linearfaktoren gilt allgemein.

Satz 4.9 Sei $P = a_n x^n + \dots + a_1 x + a_0$ ein Polynom über \mathbb{C} vom Grad n . Dann lässt sich P schreiben als

$$P = a_n (x - \xi_1)^{m_1} \dots (x - \xi_k)^{m_k},$$

wobei $\xi_i \neq \xi_j$ für $i \neq j$ die verschiedenen komplexen Nullstellen der Polynomfunktion $P(x)$ sind. Die Zahl m_i nennt man *Multiplizität* der Nullstelle ξ_i ; es gilt $m_1 + \dots + m_k = n$.

Beispiel 4.29 Die Polynomfunktion

$$P(x) = x^6 - 2x^5 - 8x^4 + 14x^3 + 11x^2 - 28x + 12$$

hat die einfache Nullstelle 3, die zweifache Nullstelle -2 und die dreifache Nullstelle 1. Somit kann man $P(x)$ schreiben als

$$P(x) = (x - 3) \cdot (x - 2)^2 \cdot (x - 1)^3. \quad \square$$

Kapitel 5

Lineare Algebra

In diesem Kapitel werden wir die Grundbegriffe der algebraischen Strukturen aus Kapitel 4 aufgreifen, um eine weitere algebraische Struktur (die des Vektorraums) zu definieren, die sich aber durch eine Verknüpfung mit einem Körperelement von allen bisher untersuchten Strukturen unterscheidet. Aufbauend auf dem Begriff des Vektorraums werden wir eine umfassende Theorie der linearen Funktionen, Matrizen, und linearen Gleichungssysteme erarbeiten.

5.1 Vektorräume

Die Menge M^n aller n -stelligen Tupel von Elementen aus M bildet eine kommutative Gruppe $[M^n, +]$, wenn $[M, +]$ bereits eine kommutative Gruppe ist: Die Addition zweier Elemente (x_1, \dots, x_n) und (y_1, \dots, y_n) aus M^n ist definiert als

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n),$$

mit neutralem Element $(0, \dots, 0)$ und Inversen

$$-(x_1, \dots, x_n) := (-x_1, \dots, -x_n).$$

Man sieht, dass sich alle Eigenschaften von $[M^n, +]$ komponentenweise auf die Eigenschaften der zugrundeliegenden Gruppe $[M, +]$ zurückführen lassen; dies gilt auch für die Kommutativität. Beispiele dieser Struktur sind etwa \mathbb{R}^2 und \mathbb{R}^3 , die Mengen aller Paare bzw. Tripel reeller Zahlen.

Im Folgenden werden wir Mengen betrachten, die neben der Gruppenstruktur noch eine Multiplikation mit einer reellen Zahl erlauben. Auf der Menge \mathbb{R}^n der n -Tupel reeller Zahlen kann man etwa die Multiplikation mit einer reellen Zahl λ definieren als

$$\lambda \cdot (x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

Aus diesem konkreten Beispiel einer kommutativen Gruppe, auf der zusätzlich eine Multiplikation mit einem *Skalar* (einer reellen Zahl) definiert ist, abstrahiert man das Konzept eines *Vektorraums*, das wie folgt definiert ist.

Definition 5.1 (Vektorraum)

Eine Menge V mit einer Verknüpfung $+$ nennt man *Vektorraum* (über den reellen Zahlen), wenn $[V, +]$ eine kommutative Gruppe ist und zwischen \mathbb{R} und V eine weitere Operation $\cdot : \mathbb{R} \times V \rightarrow V$ definiert ist (genannt *Skalarmultiplikation*). Die Skalarmultiplikation muss mit $[V, +]$ und der Addition und Multiplikation in \mathbb{R} kompatibel sein, d.h. für alle $v, w \in V, \lambda, \mu \in \mathbb{R}$ muss gelten:

- (1) $1 \cdot v = v$
- (2) $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$
- (3) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$
- (4) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$

Als *Vektor* bezeichnet man ein Element eines Vektorraums.

Im Folgenden werden wir abkürzend λv statt $\lambda \cdot v$ schreiben, wenn sich daraus keine Zweideutigkeiten ergeben. Zusätzlich schreiben wir für das neutrale Element der Gruppe $[V, +]$ einfach 0 , obwohl Vektoren meist Tupel von Zahlen sein werden. Diese Unterscheidung ist etwa in Definition 5.5 wichtig.

Vektorräume sind also Mengen, die zur Beschreibung von Vektoren im \mathbb{R}^n (wie aus der Schulmathematik bekannt) geeignet sind. Wir werden im Folgenden bis zur Notwendigkeit einer Unterscheidung in Abschnitt 5.3 reelle Vektoren sowohl als Spalten oder Zeilen schreiben. Geometrisch veranschaulicht kann man Vektoren also als Pfeile oder Punkte betrachten, die man addieren oder mit einem konstanten Faktor dehnen oder stauchen kann.

Es sei noch erwähnt, dass auch ganz andere Mengen der Vektorraumdefinition genügen; wir betrachten dazu zwei Beispiele.

Beispiel 5.1 Sei \mathcal{C} die Menge aller stetigen Funktionen von \mathbb{R} nach \mathbb{R} . Dann ist \mathcal{C} auch ein Vektorraum über \mathbb{R} , wenn man die Operationen $+$ und \cdot aus Definition 5.1 geeignet definiert: Mit

$$(f + g)(x) := f(x) + g(x)$$

und

$$(c \cdot f)(x) := cf(x)$$

für beliebige $f, g \in \mathcal{C}$ und $c, x \in \mathbb{R}$ ist $[\mathcal{C}, +]$ eine kommutative Gruppe und auch ein Vektorraum über \mathbb{R} , da die Vektorraumoperationen $+$ und \cdot auf die entsprechenden Operationen bei reellen Zahlen zurückgeführt werden. In \mathbb{R} sind die Bedingungen aus Definition 5.1 alle erfüllt. \square

Beispiel 5.2 Man kann einfach nachrechnen, dass die Menge $\mathbb{R}[x]$ aller reellen Polynome einen Vektorraum über \mathbb{R} bildet. Wie wir bereits wissen, ist $\mathbb{R}[x]$ eine Gruppe

mit der Addition: Für $P = a_n x^n + \dots + a_1 x + a_0$, $Q = b_m x^m + \dots + b_1 x + b_0 \in \mathbb{R}[x]$, $\lambda \in R$ und $n > m$ ist ist

$$P + Q = a_n x^n + \dots + (a_m + b_m) x^m + \dots + (a_1 + b_1) x + (a_0 + b_0)$$

und

$$\lambda \cdot P = \lambda a_n x^n + \dots + \lambda a_1 x + \lambda a_0. \quad \square$$

Auch Teilmengen von Vektorräumen können die Vektorraumaxiome erfüllen. Wir werden in Abschnitt 5.5 sehen, dass spezielle Untervektorräume beim Lösen linearer Gleichungssysteme eine wichtige Rolle spielen. Allgemein sind Unterräume wie folgt definiert.

Definition 5.2 (Untervektorraum)

Sei V ein Vektorraum über \mathbb{R} . Eine Teilmenge $U \subseteq V$ heißt *Untervektorraum*, wenn für beliebige $u_1, u_2 \in U$ und $\lambda \in \mathbb{R}$ gilt:

- (1) $u_1 + u_2 \in U$
- (2) $\lambda u_1 \in U$

Damit sind Untervektorräume also bezüglich der gegebenen Operationen abgeschlossene Teilmengen eines Vektorraums.

Beispiel 5.3 Die von den Vektoren $(1, 1, 1)$ und $(0, 0, 1)$ durch den Ursprung aufgespannte Ebene E des \mathbb{R}^3 bildet einen Untervektorraum von \mathbb{R}^3 . Zum Nachrechnen muss man sich überlegen, dass der Begriff des ‘‘Aufspannens’’ von E folgendermaßen zu formalisieren ist:

$$E = \{(x, y, z) \in \mathbb{R}^3 \mid \exists \lambda, \mu \in \mathbb{R} \ (x, y, z) = \lambda(1, 1, 1) + \mu(0, 0, 1)\}.$$

Wir müssen nun die Behauptungen

$$\forall_{u, v \in E} u + v \in E \quad \text{und} \quad \forall_{u \in E} \forall_{\lambda \in K} \lambda u \in E$$

nachrechnen. Zur Erinnerung an die formal korrekte Vorgehensweise gehen wir in diesem Beispiel langsam vor: Zu zeigen sind (in etwas anderer Schreibweise) die Implikationen

$$\forall_{u, v} (u, v \in E) \Rightarrow (u + v \in E) \quad \text{und} \quad \forall_{u, \lambda} (u \in E \wedge \lambda \in K) \Rightarrow (\lambda u \in E).$$

Wir nehmen also u, v, λ beliebig aber fix und teilen die zu zeigenden Implikationen in einen ‘‘wir wissen’’ und einen ‘‘zu zeigen’’ Teil auf. Für die erste Implikation erhalten wir

$$\begin{array}{ll} \text{wir wissen} & u, v \in E \\ \text{zu zeigen} & u + v \in E \end{array}$$

Dies kann man mit der Definition von E umformulieren und erhält dann

$$\begin{array}{l} \text{wir wissen} \quad \exists_{\lambda_1, \mu_1 \in \mathbb{R}} u = \lambda_1(1, 1, 1) + \mu_1(0, 0, 1) \quad \wedge \\ \quad \quad \quad \exists_{\lambda_2, \mu_2 \in \mathbb{R}} v = \lambda_2(1, 1, 1) + \mu_2(0, 0, 1) \\ \text{zu zeigen} \quad \exists_{\lambda_3, \mu_3 \in \mathbb{R}} u + v = \lambda_3(1, 1, 1) + \mu_3(0, 0, 1) \end{array}$$

Zusammenzählen der beiden Gleichungen aus dem “wir wissen” Teil ergibt

$$u + v = (\lambda_1 + \lambda_2)(1, 1, 1) + (\mu_1 + \mu_2)(0, 0, 1);$$

die gesuchten λ_3 und μ_3 sind damit $\lambda_3 = \lambda_1 + \lambda_2$ und $\mu_3 = \mu_1 + \mu_2$. Der Beweis der Implikation $\forall_{u, \lambda} (u \in E \wedge \lambda \in K) \Rightarrow (\lambda u \in E)$ verläuft ähnlich. \square

Da in Definition 5.2 $\lambda u_1 \in U$ für alle λ (also auch $\lambda = 0!$) gefordert wird sieht man sofort, dass ein Unterraum durch den Ursprung verlaufen muss. Somit sind etwa beliebige Geraden im \mathbb{R}^2 *keine* Unterräume, wenn sie nicht durch den Nullpunkt gehen. Für solcherart “verschobene” Unterräume gibt es einen eigenen Begriff, der ebenfalls bei der Lösung von linearen Gleichungssystemen auftritt.

Definition 5.3 (Mannigfaltigkeit)

Sei U ein Unterraum des Vektorraums V über \mathbb{R} , und $p \in V$ ein Vektor in V . Dann nennt man die Menge

$$p + U := \{p + u \mid u \in U\}$$

eine *lineare Mannigfaltigkeit* in V .

Vom Beispiel 5.3 ausgehend kann man sich fragen, welche Untervektorräume man aus einer Menge von Vektoren bilden kann. Dafür ist die nächste Begriffsbildung nützlich.

Definition 5.4 (Lineare Hülle)

Sei V ein Vektorraum über \mathbb{R} und $v_1, \dots, v_n \in V$. Dann nennt man

$$L(\{v_1, \dots, v_n\}) := \{\lambda_1 v_1 + \dots + \lambda_n v_n \mid \lambda_1, \dots, \lambda_n \in K\}$$

die *lineare Hülle* von v_1, \dots, v_n . Den in dieser Definition auftretenden Ausdruck

$$\lambda_1 v_1 + \dots + \lambda_n v_n$$

nennt man *Linearkombination* der Vektoren v_1, \dots, v_n .

Die lineare Hülle einer Menge von Vektoren W gibt somit an, welche Vektoren durch Linearkombinationen von W gebildet werden können. Wie man leicht nachprüfen kann, bildet die lineare Hülle einer Menge von Vektoren einen Untervektorraum.

Aus der geometrischen Anschauung heraus sieht man etwa, dass zwei nicht-parallele Vektoren im \mathbb{R}^3 eine Ebene aufspannen. Man kann aber auch lineare Hüllen von Vektoren in \mathcal{C} finden.

Beispiel 5.4 Sei \mathcal{C} wie in Beispiel 5.1. Die Vektoren f_0, \dots, f_n seien definiert durch

$$\begin{aligned} f_i &: \mathbb{R} \rightarrow \mathbb{R} \\ x &\mapsto x^i \end{aligned}$$

Dann ist $L(\{f_0, \dots, f_n\})$ der Untervektorraum aller Polynomfunktionen bis zum Grad n von \mathbb{R} nach \mathbb{R} . \square

Die lineare Hülle von Vektoren gibt also denjenigen Raum an, der von diesen Vektoren "aufgespannt" wird. Umgekehrt stellt sich aber die Frage, ob eine Menge von Vektoren, die einen Raum aufspannt, auch die kleinste solche Menge ist. So ist etwa

$$L(\{(1, 1, 1), (0, 0, 1)\}) = L(\{(1, 1, 1), (0, 0, 1), (1, 1, 0)\}),$$

da der dritte Vektor $(1, 1, 0)$ bereits in der von den ersten beiden aufgespannten Ebene liegt (und somit den Raum nicht weiter vergrößert). Dies führt zum wichtigen Begriff der *linearen Unabhängigkeit*.

Definition 5.5 (Lineare Unabhängigkeit)

Sei V ein Vektorraum über \mathbb{R} . Eine Menge von Vektoren $v_1, \dots, v_n \in V$ heißt *linear unabhängig*, wenn gilt

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0 \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_n = 0.$$

Eine Menge von Vektoren heißt *linear abhängig*, wenn sie nicht linear unabhängig ist.

Beispiel 5.5 Gegeben seien die Vektoren $(1, 2, 1), (0, 1, 1), (0, 1, 0)$ des \mathbb{R}^3 . Um zu überprüfen, ob diese drei Vektoren linear abhängig oder unabhängig sind, muss man die Bedingung

$$\lambda_1(1, 2, 1) + \lambda_2(0, 1, 1) + \lambda_3(0, 1, 0) = (0, 0, 0)$$

darauf überprüfen, ob nur die Lösung $\lambda_1 = \lambda_2 = \lambda_3 = 0$ möglich ist. Es ergibt sich somit das Gleichungssystem

$$\begin{aligned} \lambda_1 &= 0 \\ 2\lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ \lambda_1 + \lambda_2 &= 0 \end{aligned}$$

woraus $\lambda_1 = \lambda_2 = \lambda_3 = 0$ folgt. Somit sind die drei Vektoren linear unabhängig. \square

Mit den Begriffen der linearen Hülle und der linearen Unabhängigkeit haben wir nun die Möglichkeit, genau die Teilmengen eines Vektorraums auszuzeichnen, die einen Vektorraum aufspannen. Wir suchen also für einen gegebenen Vektorraum V

eine Menge U mit $L(U) = V$, die zudem noch linear unabhängig ist. Trivialerweise gibt es eine Menge U , die die erste Bedingung erfüllt, nämlich V selbst – man nennt übrigens jede Menge $M \subseteq V$ mit $L(M) = V$ ein *Erzeugendensystem* von V . Die Bedingung der linearen Unabhängigkeit schränkt die Wahl von U schon mehr ein. Man kann allerdings auch dafür nachrechnen, dass es für ein linear abhängiges $U \subseteq V$ mit $L(U) = V$ auch eine *echte* Teilmenge $W \subset U$ mit $L(W) = V$ gibt. Man kann also aus einer linear abhängigen Menge solange Vektoren entfernen, bis diese linear unabhängig wird (und dabei immer noch Erzeugendensystem bleibt). Mengen mit diesen Eigenschaften bezeichnet man als *Basis* eines Vektorraums.

Definition 5.6 (Basis)

Sei V ein Vektorraum über \mathbb{R} und $B \subseteq V$. Wenn B die Bedingungen

- (i) $L(B) = V$
- (ii) B ist linear unabhängig

erfüllt, dann nennt man B eine *Basis* von V .

Wir werden im Folgenden nur Vektorräume mit Basen betrachten, die endlich viele Elemente enthalten (dies ist etwa bei dem Funktions-Vektorraum in Beispiel 5.1 nicht der Fall). Für gegebenen Vektorraum V ist eine Basis nicht eindeutig bestimmt, wohl aber die Anzahl der Elemente in der Basis.

Definition 5.7 (Dimension)

Sei V ein Vektorraum mit endlicher Basis B . Dann nennt man die eindeutig bestimmte Anzahl der Elemente in B die *Dimension* von V , und schreibt dafür $\dim(V)$. Für lineare Mannigfaltigkeiten $M = p + U$ (mit $p \in V$ und U ein Unterraum von V) ist die Dimension von M als die Dimension von U definiert.

Beispiel 5.6 Die bekanntesten Basen von \mathbb{R}^2 , \mathbb{R}^3 und allgemeiner \mathbb{R}^n sind die *Standardbasen* $E_n = (e_1, \dots, e_n)$, deren Komponenten durch

$$e_{ik} = \begin{cases} 1 & \text{wenn } i = k \\ 0 & \text{sonst} \end{cases}$$

definiert sind. Es ist einsichtig, dass die so definierten E_n eine Basis bilden, da sie linear unabhängig sind und \mathbb{R}^n aufspannen. Somit ist jetzt auch formal festgelegt, dass \mathbb{R}^n ein n -dimensionaler Vektorraum ist. \square

Beispiel 5.7 Die von den Vektoren $v_1 = (1, 1, 0)$ und $v_2 = (0, 1, 1)$ aufgespannte Teilmenge des \mathbb{R}^3 ist ein zweidimensionaler Unterraum von \mathbb{R}^3 : Da die lineare Hülle von Vektoren immer einen Unterraum bildet, und v_1 und v_2 linear unabhängig sind, hat dieser Unterraum die Dimension 2. \square

Beispiel 5.8 Die Menge $\{(1, 1), (0, 2)\}$ ist eine alternative Basis des \mathbb{R}^2 . Es ist sofort ersichtlich, dass diese beiden Vektoren linear unabhängig sind. Außerdem lässt sich jeder Vektor des \mathbb{R}^2 als Linearkombination

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{y-x}{2} \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

schreiben; somit ist $L(\{(1, 1), (0, 2)\}) = \mathbb{R}^2$.

Die Vektorenmenge $B = \{(1, 2, 1), (0, 1, 1), (0, 1, 0)\}$ aus Beispiel 5.5 ist eine Basis des \mathbb{R}^3 , da sie linear unabhängig (siehe oben) und wegen

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + (z-x) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + (y-x-z) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

auch ein Erzeugendensystem von \mathbb{R}^3 ist. \square

Wenn eine Menge von Vektoren eine Basis eines Vektorraumes bildet, kann man *jedes* Element dieses Vektorraumes *eindeutig* als Linearkombination von Basisvektoren anschreiben – das folgt sofort aus den Eigenschaften einer Basis, Erzeugendensystem und linear unabhängig zu sein. Die Skalare, mit denen bei dieser eindeutigen Darstellung die Basiselemente multipliziert werden, nennt man *Koordinaten* bezüglich dieser Basis.

Definition 5.8 (Koordinaten)

Sei $B = \{b_1, \dots, b_n\} \subseteq V$ eine Basis des Vektorraums V über \mathbb{R} . Wir schreiben der Eindeutigkeit der Reihenfolge wegen B als *geordnete Basis*, d.h. als Tupel (b_1, \dots, b_n) . Dann nennt man für jedes $v \in V$ das eindeutig bestimmte n -Tupel $(\lambda_1, \dots, \lambda_n) \in K^n$ mit

$$v = \lambda_1 b_1 + \dots + \lambda_n b_n$$

die *Koordinaten* von v bezüglich (b_1, \dots, b_n) .

Unsere bisherige Schreibweise von reellen Vektoren als n -Tupel von Zahlen kann somit als Koordinatenschreibweise bezüglich der Standardbasis E_n des \mathbb{R}^n aufgefasst werden.

Beispiel 5.9 Wie wir in Beispiel 5.8 gesehen haben, ist $B_1 = ((1, 1), (0, 2))$ eine geordnete Basis des \mathbb{R}^2 , und $B_2 = ((1, 2, 1), (0, 1, 1), (0, 1, 0))$ eine geordnete Basis des \mathbb{R}^3 . Die Koordinaten eines Vektors $(x, y) \in \mathbb{R}^2$ bezüglich B_1 sind $(x, \frac{y-x}{2})$, die Koordinaten eines Vektors $(x, y, z) \in \mathbb{R}^3$ bezüglich B_2 sind $(x, z-x, y-x-z)$. \square

5.2 Lineare Abbildungen

In diesem Abschnitt betrachten wir *lineare Funktionen* (auch *lineare Abbildungen* genannt). Diese Abbildungen zwischen zwei Vektorräumen sind mit den Operationen der Vektorräume kompatibel, d.h., sie bilden Vektorraumelemente so aufeinander ab, dass die Operationen in den Vektorräumen erhalten bleiben. Genauer ist dies wie folgt definiert.

Definition 5.9 (Lineare Abbildung)

Seien V und W zwei Vektorräume über K mit den Vektorraumoperationen $(+, \cdot)$ bzw. (\oplus, \cdot) . Eine Funktion $f : V \rightarrow W$ nennt man *lineare Abbildung*, wenn für alle $\lambda \in K$ und $v_1, v_2 \in V$ folgende Bedingungen erfüllt sind

- (i) $f(\lambda \cdot v) = \lambda \cdot f(v)$
- (ii) $f(v_1 + v_2) = f(v_1) \oplus f(v_2)$

Bei linearen Abbildungen spielt es also keine Rolle, ob man zuerst zwei Elemente in V verknüpft und dann nach W abbildet, oder die Elemente einzeln abbildet und dann in W verknüpft. Wir werden in Zukunft wieder die "Standardschreibweise" mit $+$ und Weglassen des Multiplikationszeichens verwenden.

Man kann nachprüfen, dass für lineare Abbildungen f immer $f(0) = 0$ gilt. Bei nicht-injektiven Abbildungen werden allerdings auch noch andere Vektoren auf 0 abgebildet. Ebenso ist bei nicht-surjektiven Abbildungen $f : V \rightarrow W$ die Menge $f(V)$ nicht ganz W . Man führt daher folgende Bezeichnungen ein.

Definition 5.10 (Bild und Kern)

Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann nennt man

$$f(V) := \{w \in W \mid \text{es gibt ein } v \in V \text{ mit } f(v) = w\}$$

das *Bild* von f und schreibt auch $\text{bild}(f)$. Der *Kern* von f

$$\text{kern}(f) := \{v \in V \mid f(v) = 0\}$$

ist die Menge aller Elemente von V , die von f auf 0 abgebildet werden.

Diese beiden Teilmengen des Bild- bzw. Definitionsbereichs einer linearen Abbildung sind abgeschlossen, wie der folgende Satz besagt.

Satz 5.1 Sei $f : V \rightarrow W$ eine lineare Abbildung. Dann sind $\text{bild}(f)$ und $\text{kern}(f)$ Untervektorräume von W bzw. V .

Dieser Satz ist leicht nachzuprüfen: Seien etwa $v_1, v_2 \in \text{kern}(f)$. Da f linear ist, folgt aus $f(v_1) = 0$ und $f(v_2) = 0$ auch $f(v_1 + v_2) = 0$ und $f(\lambda v_1) = 0$ für beliebiges λ . Analoges gilt für zwei Vektoren $w_1, w_2 \in \text{bild}(f)$.

Beispiel 5.10 Wir betrachten drei Abbildungen zwischen reellen Vektorräumen:

$$\begin{array}{lll}
 f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 & g : \mathbb{R}^2 \rightarrow \mathbb{R}^2 & h : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\
 \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + y \\ 3x - 4y \end{pmatrix} & \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 3x + 2y - 2 \\ x + 4 \end{pmatrix} & \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 + 2y \\ x \cdot y \end{pmatrix}
 \end{array}$$

Man kann nachrechnen, dass f eine lineare Funktion ist, g und h aber nicht: Es gilt

$$\begin{aligned} f\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} &= \begin{pmatrix} x_1 + x_2 + y_1 + y_2 \\ 3(x_1 + x_2) - 4(y_1 + y_2) \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ 3x_1 - 4y_1 \end{pmatrix} + \begin{pmatrix} x_2 + y_2 \\ 3x_2 - 4y_2 \end{pmatrix} \\ &= f\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + f\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \end{aligned}$$

und

$$f\begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} = \begin{pmatrix} \lambda x + \lambda y \\ 3\lambda x - 4\lambda y \end{pmatrix} = \begin{pmatrix} \lambda(x + y) \\ \lambda(3x - 4y) \end{pmatrix} = \lambda f\begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Abbildung g kann nicht linear sein, da

$$g\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

ist, bei einer linearen Abbildung aber immer der Nullpunkt auf den Nullpunkt abgebildet werden muss. Dies ist bei h zwar der Fall, trotzdem ist diese Abbildung ebenfalls nicht linear: es gilt etwa

$$h\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 20 \\ 8 \end{pmatrix},$$

für eine lineare Abbildung müsste aber

$$h\begin{pmatrix} 4 \\ 2 \end{pmatrix} = h\begin{pmatrix} 2 \\ 1 \end{pmatrix} + h\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix} + \begin{pmatrix} 6 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 4 \end{pmatrix}$$

gelten. □

Aus diesem Beispiel sieht man, dass die linearen Abbildungen zwischen reellen Vektorräumen genau die Funktionen sind, die sich folgendermaßen anschreiben lassen:

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{R}^m \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &\mapsto x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}. \end{aligned}$$

Diese Schreibweise ist bei den Funktionen aus Beispiel 5.10 nur für die lineare Funktion f als

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ 3x - 4y \end{pmatrix} = x \begin{pmatrix} 1 \\ 3 \end{pmatrix} + y \begin{pmatrix} 1 \\ -4 \end{pmatrix}$$

möglich, nicht aber für g und h . Später werden wir Matrizen verwenden, um lineare Funktionen kompakt anschreiben zu können.

Im Folgenden werden meist lineare Abbildungen zwischen reellen Vektorräumen behandelt. Zur Verdeutlichung der Begriffe ist es aber illustrativ, auch andere Beispiele zu untersuchen.

Beispiel 5.11 Sei M die Menge aller Polynome 3. Grades. Nach Beispiel 5.2 ist M dann auch ein Vektorraum. Die Funktion

$$f : \mathbb{R}^2 \rightarrow M$$

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto ax^3 + 4x^2 + bx + 2$$

ist *keine* lineare Abbildung, da der Nullpunkt in \mathbb{R}^2 nicht auf den Nullpunkt in M (das Nullpolynom) abgebildet wird. Andererseits kann man nachrechnen, dass

$$g : \mathbb{R}^2 \rightarrow M$$

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto ax^3 + 3(a+b)x^2 + bx$$

die Bedingungen einer linearen Funktion erfüllt: So ist

$$\begin{aligned} g\begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} &= \lambda ax^3 + 3(\lambda a + \lambda b)x^2 + \lambda bx \\ &= \lambda(ax^3 + 3(a+b)x^2 + bx) \\ &= \lambda g\begin{pmatrix} a \\ b \end{pmatrix} \end{aligned}$$

und

$$\begin{aligned} g\left(\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix}\right) &= (a+c)x^3 + 3((a+c) + (b+d))x^2 + (b+d)x \\ &= ax^3 + 3(a+b)x^2 + bx + cx^3 + 3(c+d)x^2 + dx \\ &= g\begin{pmatrix} a \\ b \end{pmatrix} + g\begin{pmatrix} c \\ d \end{pmatrix}. \end{aligned} \quad \square$$

Wie wir bereits wissen, kann man jedes Element eines Vektorraums eindeutig als Linearkombination von Basisvektoren darstellen. Für lineare Abbildungen f , die ja Vektorraumoperationen erhalten, ergibt sich unmittelbar mit einem Vektor v und einer Basis (b_1, \dots, b_n) der Zusammenhang

$$f(v) = f(\lambda_1 b_1 + \dots + \lambda_n b_n) = \lambda_1 f(b_1) + \dots + \lambda_n f(b_n),$$

also folgender Satz.

Satz 5.2 Eine lineare Abbildung ist durch die Bilder ihrer Basisvektoren eindeutig gegeben.

Beispiel 5.12 Sei $B = ((3, 1), (2, 3))$ eine geordnete Basis des \mathbb{R}^2 und f die lineare Abbildung, die durch

$$f\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad f\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

gegeben ist. Die Koordinaten eines beliebigen Vektors (x, y) bezüglich B sind aus der Gleichung

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{7}(3x - 2y) \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \frac{1}{7}(3y - x) \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

abzulesen. Somit ist

$$\begin{aligned} f \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{7}(3x - 2y)f \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \frac{1}{7}(3y - x)f \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ &= \frac{1}{7}(3x - 2y) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{7}(3y - x) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{7} \begin{pmatrix} 2x + y \\ -x + 3y \end{pmatrix}. \end{aligned}$$

□

5.3 Matrizen

Da die Definition einer linearen Abbildung

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$$

nur von den Faktoren a_{11}, \dots, a_{mn} abhängt und das wiederholte Anschreiben der Koeffizienten x_1, \dots, x_n eigentlich redundant ist, kürzt man derartige Ausdrücke durch *Matrizen* ab.

Definition 5.11 (Matrix)

Seien a_{11}, \dots, a_{mn} $m \times n$ reelle Zahlen. Eine Anordnung dieser Zahlen in einem Schema

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

nennt man eine *Matrix* mit m Zeilen und n Spalten, oder kurz $m \times n$ Matrix, und schreibt auch $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ oder noch kürzer $A = (a_{ij})$, wenn die Dimensionen

bekannt sind. Für $m = n$ nennt man die Matrix *quadratisch*. Wir bezeichnen die Menge aller reellen $m \times n$ Matrizen mit $\mathcal{M}_{m \times n}$.

Die Spalten

$$a_{-1} = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, a_{-2} = \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, a_{-n} = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

einer Matrix $A = (a_{ij})$ nennt man *Spaltenvektoren* von A , die Zeilen

$$\begin{aligned} a_{1-} &= (a_{11} \ a_{12} \ \dots \ a_{1n}) \\ &\vdots \\ a_{m-} &= (a_{m1} \ a_{m2} \ \dots \ a_{mn}) \end{aligned}$$

Zeilenvektoren von A .

Es lassen sich natürlich auch Zahlen aus anderen Zahlenbereichen in Matrixform anordnen; wir werden hier nur reelle Matrizen behandeln. Auf der Menge der Matrizen müssen nun geeignete Operationen definiert werden, um sie als abkürzende Schreibweise verwenden zu können. Dabei gilt der folgende Satz.

Satz 5.3 Seien $A, B \in \mathcal{M}_{m \times n}$ zwei reelle Matrizen, und $c \in \mathbb{R}$ ein Skalar. Dann bildet $\mathcal{M}_{m \times n}$ mit den Operationen

$$A + B := \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & \dots & a_{2n} + b_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}, \quad cA := \begin{pmatrix} ca_{11} & \dots & ca_{1n} \\ ca_{21} & \dots & ca_{2n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \dots & ca_{mn} \end{pmatrix}$$

einen reellen Vektorraum.

Der Beweis dieses Satzes ist einfach, da alle Operationen komponentenweise definiert sind und alle Axiome für die reellen Zahlen gelten (und damit auch für Matrizen). Es gibt aber noch eine weitere interessante Rechenoperation auf der Menge der Matrizen.

Definition 5.12 (Matrixmultiplikation)

Sei A eine $m \times n$ Matrix und B eine $n \times k$ Matrix. Dann ist die Matrix

$$A \cdot B := \begin{pmatrix} \sum_{j=1}^n a_{1j}b_{j1} & \dots & \sum_{j=1}^n a_{1j}b_{jk} \\ \sum_{j=1}^n a_{2j}b_{j1} & \dots & \sum_{j=1}^n a_{2j}b_{jk} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^n a_{mj}b_{j1} & \dots & \sum_{j=1}^n a_{mj}b_{jk} \end{pmatrix}$$

eine $m \times k$ Matrix.

Weiters ist es sinnvoll, bei einer Matrix Spalten und Zeilen vertauschen zu können. Dies ist durch *Transposition* möglich. Für eine $m \times n$ Matrix A ist $A^T = (a^T_{ij})$, definiert über die Bedingung

$$a^T_{ij} := a_{ji}$$

eine $n \times m$ Matrix, die zu A *transponierte Matrix*.

Die Matrizenmultiplikation ist etwas komplizierter definiert als die komponentenweise Addition und Skalarmultiplikation (mit gutem Grund, wie wir in Satz 5.8

sehen werden): in der i -ten Zeile und j -ten Spalte de Produktes $A \cdot B$ steht die Summe der komponentenweisen Produkte der i -ten Zeile von A mit der j -ten Spalte von B . Speziell bei der Multiplikation mit Spaltenvektoren wird das Operationszeichen “ \cdot ” oft weggelassen.

Beispiel 5.13 Seien

$$A = \begin{pmatrix} 2 & 3 & -1 \\ 1 & -2 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -3 & 0 \\ 2 & 1 & -2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{pmatrix}$$

drei Matrizen, sowie $d = 2$. Dann ist

$$A + C = \begin{pmatrix} 3 & 0 & -1 \\ 3 & -1 & 2 \end{pmatrix}, \quad dA = \begin{pmatrix} 4 & 6 & -2 \\ 2 & -4 & 8 \end{pmatrix},$$

$$A \cdot B = \begin{pmatrix} 5 & -7 \\ 9 & 18 \end{pmatrix}, \quad A^T = \begin{pmatrix} 2 & 1 \\ 3 & -2 \\ -1 & 4 \end{pmatrix}.$$

□

Der Zusammenhang zwischen Matrizen und linearen Abbildungen ist sehr einfach und aus der Definition von Matrizen und Matrizenmultiplikation direkt ablesbar. Um der Bedeutung dieses Zusammenhanges Genüge zu tun, formulieren wir ihn als folgenden Satz.

Satz 5.4 Eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix}$$

kann in Matrixschreibweise kürzer und übersichtlicher als

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

dargestellt werden.

Natürlich erfüllt auch eine durch eine Matrix definierte Abbildung wiederum die Linearitätsbedingungen aus Definition 5.9, ist also wiederum eine lineare Abbildung. Dies folgt aus

$$A \cdot (x + y) = A \cdot x + A \cdot y \quad \text{und} \quad A \cdot (\lambda x) = \lambda A \cdot x,$$

wie man durch Nachrechnen für beliebige $m \times n$ Matrizen A , $n \times 1$ Spaltenvektoren x und y sowie Skalare λ überprüfen kann.

Matrizen besitzen ähnliche algebraische Strukturen wie Polynome: So bilden sie bezüglich der Operation $+$ eine kommutative Gruppe und mit der Skalarmultiplikation einen Vektorraum. Dies ist einfach nachzuprüfen, da durch die komponentenweise

Art der Operationen alle Bedingungen auf Operationen mit Skalaren zurückgeführt werden können. Die Skalare besitzen bei uns stets genügend "Struktur" (sind also selbst Gruppen-, Vektorraum- oder Körperelemente), um die Bedingungen zu erfüllen.

Eine interessante Teilmenge der Menge aller Matrizen ist jene, die auch bezüglich der Matrizenmultiplikation eine Gruppenstruktur aufweist.

Definition 5.13 (Reguläre Matrizen)

Die Teilmenge der quadratischen Matrizen, die bezüglich der Matrizenoperation \cdot eine Gruppe bildet, nennt man die Gruppe der *regulären Matrizen*. Dabei bezeichne A^{-1} das zur Matrix A inverse Element, und I_n das für $n \times n$ Matrizen neutrale Element (die *Einheitsmatrix*). Wenn A^{-1} existiert, nennt man A *invertierbar*.

Man kann nachrechnen, dass I_n die Struktur

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

haben muss. Die Gültigkeit des Assoziativitätsgesetzes ist durch Rückführung auf die Skalaroperationen auch einfach nachzuprüfen. Man kann nicht sofort erkennen, ob eine quadratische Matrix regulär ist oder nicht; ein Kriterium für die Regularität einer Matrix werden wir in Abschnitt 5.4 behandeln.

Doch nun wieder zurück zu linearen Abbildungen. Der Zusammenhang zwischen Matrizen und linearen Abbildungen wird im nächsten wichtigen Satz dargestellt, der nur auf den ersten Blick kompliziert aussieht.

Satz 5.5 Seien V und W zwei Vektorräume mit Basen (v_1, \dots, v_n) und (w_1, \dots, w_m) , A eine $m \times n$ Matrix und $f : V \rightarrow W$ die durch $x \mapsto Ax$ gegebene lineare Funktion.

Dann ist die i -te Spalte von A der Koordinatenvektor des Bildes von v_i bezüglich der Basis (w_1, \dots, w_m) .

Dieser Satz ist somit eine Weiterführung von Satz 5.2 (eine lineare Abbildung ist durch die Bilder der Basisvektoren eindeutig bestimmt).

Wir betrachten ein geometrisches Beispiel, um diesen Zusammenhang klar zu machen.

Beispiel 5.14 Gesucht ist die lineare Funktion, die in \mathbb{R}^2 eine Rotation um den Ursprung um 30° gegen den Uhrzeigersinn beschreibt. Wir wählen die Standardbasis E_2 und brauchen uns mit Satz 5.5 nur zu überlegen, wohin die beiden Basisvektoren abgebildet werden. Mit einer einfachen Skizze kann man erkennen, dass

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos 30^\circ \\ \sin 30^\circ \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin 30^\circ \\ \cos 30^\circ \end{pmatrix},$$

somit ist die gesuchte lineare Abbildung durch

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

gegeben. □

An dieser Stelle muss erwähnt werden, dass bei der Matrizendarstellung einer linearen Abbildung eigentlich mit *Koordinaten* gerechnet wird; dies ist aber oftmals nicht unmittelbar ersichtlich, da die Koordinaten eines Vektors (x_1, \dots, x_n) bezüglich der Standardbasis wegen

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

ebenfalls wieder (x_1, \dots, x_n) sind, also vom *Vektor* (x_1, \dots, x_n) nicht unterscheidbar.

Wir halten fest: Bei einer linearen Abbildung, die durch $y = A \cdot x$ oder genauer

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

gegeben ist, muss Folgendes beachtet werden:

- die Matrix A ist die Abbildungsmatrix einer linearen Funktion bezüglich einer Basis D im Definitionsbereich und einer Basis B im Bildbereich,
- x bezeichnet Koordinaten bezüglich der Basis D ,
- y bezeichnet Koordinaten bezüglich der Basis B .

Diese Überlegungen haben keine praktischen Konsequenzen, wenn wir immer mit den Standardbasen rechnen, sonst aber sehr wohl, wie das nächste Beispiel zeigen wird.

Beispiel 5.15 Wir betrachten die lineare Abbildung

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 3x + 2y \\ -x + y \end{pmatrix}.$$

Als erstes wählen wir sowohl im Definitions- wie im Bildbereich die Standardbasis. Mit Satz 5.5 sind die Spalten der Matrix die Koordinaten der Bilder der Basisvektoren: Wir berechnen somit

$$f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} \quad \text{und} \quad f \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

bezüglich der Standardbasis im Bildbereich sind die Koordinaten dieser Vektoren gleich den Vektoren selbst, und die Matrix ist

$$A_1 = \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix}.$$

Man kann erkennen, dass diese Matrix bereits direkt aus der Definition der Funktion f ablesbar ist.

Wenn wir nun im Bildbereich eine andere Basis wählen, etwa $B = ((1, 2), (0, 3))$, dann ergibt sich daraus eine andere Matrixrepräsentation der linearen Funktion: Wir benötigen dafür die Koordinaten der Bilder der Basisvektoren bezüglich B . Diese erhalten wir durch Lösen der Gleichungen

$$\begin{pmatrix} 3 \\ -1 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu_1 \begin{pmatrix} 0 \\ 3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \lambda_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

als $\lambda_1 = 3$, $\mu_1 = -\frac{7}{3}$, $\lambda_2 = 2$, $\mu_2 = -1$. Somit ist die Matrix von f bezüglich der Standardbasis im Definitions- und B im Bildbereich

$$A_2 = \begin{pmatrix} 3 & 2 \\ -\frac{7}{3} & -1 \end{pmatrix}.$$

Wir können nun noch überprüfen, dass diese Matrizen tatsächlich dieselbe Funktion repräsentieren: Wie wir durch Einsetzen in die Definition von f sehen, ist

$$f \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix},$$

dies ergibt sich auch durch Matrizenmultiplikation mit A_1 :

$$\begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}.$$

Dabei ist das Resultat $(5, 5)$ die Koordinaten bezüglich der Basis im Bildbereich (der Standardbasis), und damit auch wieder der Vektor $(5, 5)$.

Bei der Matrizenmultiplikation mit A_2 erhalten wir

$$A_2 \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -\frac{7}{3} & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ -\frac{5}{3} \end{pmatrix}.$$

Da das Ergebnis Koordinaten bezüglich B (der Basis im Bildbereich) darstellt, ergibt sich erst durch Zurückrechnen der Koordinatenschreibweise dasselbe Resultat wie oben:

$$5 \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \frac{5}{3} \begin{pmatrix} 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}.$$

In ähnlicher Weise könnte man noch eine andere Basis für den Definitionsbereich festlegen, und erhielte dadurch eine nochmals andere Matrizenrepräsentation der linearen Abbildung. \square

Bei manchen Problemstellungen ist es nötig, Koordinaten zwischen verschiedenen Basen umzurechnen; eine solche Umrechnung nennt man *Basistransformation*.

Seien dazu $A = (a_{-1}, \dots, a_{-n})$ und $B = (b_{-1}, \dots, b_{-n})$ zwei $n \times n$ Matrizen, deren Spalten a_{-i} bzw. b_{-i} Basen eines n -dimensionalen Vektorraums V sind. Die Umrechnung von Koordinaten bezüglich A in Koordinaten bezüglich B kann man sich folgendermassen überlegen: Der Vektor $v \in V$ hat die uns bekannte Koordinatendarstellung

$$v = \lambda_1 a_{-1} + \dots + \lambda_n a_{-n} = A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix},$$

wir suchen die Koordinaten (ξ_1, \dots, ξ_n) bezüglich B :

$$v = \xi_1 b_{-1} + \dots + \xi_n b_{-n} = B \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Somit gilt

$$A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = B \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},$$

woraus wir die gesuchten Koordinaten durch das Lösen dieser Gleichung ausrechnen können.

Satz 5.6 Seien A und B zwei Matrizen, deren Spalten geordnete Basen darstellen, und $(\lambda_1, \dots, \lambda_n)$ die Koordinaten eines Vektors v bezüglich A . Dann sind die Koordinaten (ξ_1, \dots, ξ_n) von v bezüglich B gegeben durch

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = B^{-1} \cdot A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Bis jetzt haben wir noch keine Methoden kennengelernt, Inverse von Matrizen zu berechnen. Wir werden aber in Abschnitt 5.5 sehen, dass das Berechnen von Inversen eng mit dem Lösen von Gleichungssystemen verwandt ist.

Nach diesen Ausführungen über den Zusammenhang zwischen Basen, linearen Abbildungen und Matrizen wollen wir jetzt Matrizen selbst genauer untersuchen.

Wenn eine Matrix nicht invertierbar ist (oder etwa auch nicht quadratisch ist), so kann man doch ein Maß dafür angeben, wieviel "Struktur" (in einem noch zu erläuternden Sinn) in der Matrix steckt. Mit Satz 5.5 wissen wir, dass eine Matrix nichts anderes als die Bilder der Basisvektoren einer linearen Abbildung ist. Wenn jetzt $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine solche lineare Abbildung ist, kann man sich fragen, welche Dimension der Bildbereich $\text{bild}(f)$ hat (der ja mit Satz 5.1 Vektorraumstruktur hat). Diese Vektorraumdimension ist genau das gesuchte Strukturmaß für Matrizen: es gibt an, wie groß der Raum ist, den die Bilder einer linearen Abbildung aufspannen. Dies lässt sich über den Zusammenhang mit Matrizen auch einfach berechnen: als Anzahl linear unabhängiger Spalten (bzw. Zeilen) einer Matrix.

Definition 5.14 (Rang)

Sei A eine $m \times n$ Matrix. Dann bezeichnet man die maximale Anzahl linear unabhängiger Spalten von A als *Rang* von A und schreibt dafür $\text{rg}(A)$. Wenn A quadratisch ist mit $n = m = \text{rg}(A)$, so hat A *maximalen* (oder *vollen*) Rang.

Wie wir später sehen werden, ist die maximale Anzahl linear unabhängiger Spalten gleich der maximalen Anzahl linear unabhängiger Zeilen.

Beispiel 5.16 Seien

$$A = \begin{pmatrix} 2 & 1 & -3 \\ 0 & 1 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

zwei Matrizen. Dann ist $\text{rg}(A) = 2$, weil die ersten beiden Spalten linear unabhängig sind und in \mathbb{R}^2 maximal zwei Vektoren linear unabhängig sind (oder anders: Da A nur zwei Zeilen hat, kann $\text{rg}(A)$ maximal 2 sein). Bei B sind wiederum die ersten beiden Spalten linear unabhängig, man sieht aber sofort, dass $a_{-3} = 2a_{-1} + 2a_{-2}$ ist. Somit ist auch $\text{rg}(B) = 2$. \square

Wir werden erst in Abschnitt 5.4 die schrittweisen Matrizenumformungen genauer behandeln, die die systematische Berechnung des Rangs einer Matrix ermöglichen. Der Zusammenhang von Rang und Invertierbarkeit einer Matrix ist wie folgt.

Satz 5.7 Sei A eine $n \times n$ Matrix. Dann ist A genau dann invertierbar, wenn $\text{rg}(A) = n$ ist.

Dieser Satz ist anschaulich klar: Eine lineare Abbildung zwischen zwei gleichdimensionalen Vektorräumen ist genau dann invertierbar, wenn die Bilder der Basisvektoren den gesamten Zielraum aufspannen – was wiederum genau dann der Fall ist, wenn alle Spaltenvektoren von A linear unabhängig sind, der Rang von A also maximal ist.

Nachdem nun einige Zusammenhänge zwischen linearen Abbildungen und Matrizen klar erarbeitet wurden, werden wir kurz den Zusammenhang zwischen Hintereinanderausführung von Funktionen und Matrizenmultiplikation behandeln.

Beispiel 5.17 Gegeben seien die zwei linearen Funktionen

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3 \qquad g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2 & 1 \\ -1 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Man kann durch Einsetzen leicht nachrechnen, dass $g \circ f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ebenfalls wieder eine lineare Funktion ist: So ist

$$\begin{aligned} g \circ f : \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & -2 \end{pmatrix} \cdot f \begin{pmatrix} x \\ y \end{pmatrix} \\ &\mapsto \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & -2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &\mapsto \begin{pmatrix} 2 & 5 \\ -5 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \end{aligned}$$

die Matrix der Hintereinanderausführung ist also nichts anderes als das Produkt der Matrizen von f und g . Zu beachten ist, dass obige Umformung auf der Assoziativität der Matrizenmultiplikation beruht. \square

Der in diesem Beispiel nachgerechnete Zusammenhang zwischen der Hintereinanderausführung von Funktionen und Matrizenmultiplikation gilt auch allgemein, wie der folgende Satz besagt.

Satz 5.8 Seien f und g zwei lineare Abbildungen, die durch die Matrizen A bzw. B definiert sind. Dann gilt für die Matrix C der Hintereinanderausführung $g \circ f$ von g nach f :

$$C = B \cdot A.$$

5.4 Determinanten

Wir betrachten in diesem Abschnitt ein wichtiges Konzept, dessen Nützlichkeit sich über den Zusammenhang mit Gleichungssystemen motivieren lässt. Dafür definieren wir zunächst den Begriff eines *linearen Gleichungssystems*.

Definition 5.15 (Lineares Gleichungssystem)

Als *lineares Gleichungssystem* mit n Unbekannten x_1, \dots, x_n bezeichnet man ein System von Gleichungen der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

Dabei nennt man die Zahlen $a_{ij} \in K$ für $1 \leq i \leq m, 1 \leq j \leq n$ und beliebigen Körper K die *Koeffizienten* des Gleichungssystems. Eine *Lösung* des Gleichungssystems ist ein n -Tupel (x_1, \dots, x_n) , das jede Gleichung des Systems erfüllt. Wenn $b_1 = \dots = b_m = 0$ ist, so nennt man das Gleichungssystem *homogen*, andernfalls *inhomogen*. Der Vektor $(0, \dots, 0)$, der ja Lösung aller homogenen Gleichungssysteme ist, heißt *triviale Lösung*; alle anderen Lösungen eines homogenen Gleichungssystems heißen *nichttrivial*.

Das Gleichungssystem aus obiger Definition kann man mit Matrizen auch übersichtlicher als

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

schreiben; noch kürzer ist die Schreibweise

$$A \cdot x = b,$$

wobei A eine $m \times n$ Matrix, x ein $n \times 1$ und b ein $m \times 1$ Spaltenvektor ist. Für den Spezialfall, dass A quadratisch und noch dazu regulär (also invertierbar) ist, kann man beide Seiten der Gleichung mit A^{-1} multiplizieren und erhält damit eine eindeutige Lösung

$$x = A^{-1} \cdot b.$$

Wir werden im Folgenden ein Kriterium entwickeln, mit dessen Hilfe man die Invertierbarkeit einer Matrix überprüfen kann; ein Verfahren zur Invertierung solcher Matrizen behandeln wir erst in Abschnitt 5.5. Dazu konstruieren wir eine Funktion, die einer reellen Matrix A eine Zahl zuordnet, die das Volumen des von den Spalten von A aufgespannten Parallelepipeds angibt (ein *Parallelepiped* ist die Verallgemeinerung eines Parallelogramms auf höhere Dimensionen). Für ein solches Volumensmaß sollen die geometrisch einleuchtenden Bedingungen gelten:

- Wenn eine Seite um den Faktor λ gestreckt wird, vergrößert sich das Volumen auch um den Faktor λ ; wenn sich eine Seite aus zwei Vektoren v und w zusammensetzt, so ist das Volumen des Parallelepipeds gleich der Summe der Volumen der kleineren Parallelepipede mit v und w als Seiten,
- wenn zwei Seiten des Parallelepipeds zusammenfallen, verschwindet das Volumen.
- Um eine Größenskala festzulegen, sei das Volumen des Einheitswürfels eins.

Diese Bedingungen führen zu folgender Definition.

Definition 5.16 (Determinante)

Eine Funktion $\det : \mathcal{M}_{n \times n} \rightarrow \mathbb{R}$, die die Bedingungen

- (i) $\det(a_{-1}, \dots, \lambda a_{-i}, \dots, a_{-n}) = \lambda \det(a_{-1}, \dots, a_{-i}, \dots, a_{-n})$
- (ii) $\det(a_{-1}, \dots, v + w, \dots, a_{-n}) = \det(a_{-1}, \dots, v, \dots, a_{-n}) + \det(a_{-1}, \dots, w, \dots, a_{-n})$
- (iii) $\det(a_{-1}, \dots, v, \dots, v, \dots, a_{-n}) = 0$
- (iv) $\det(I_n) = 1$

für beliebige $n \times 1$ Spaltenvektoren v und w sowie Skalare λ erfüllt, nennt man *Determinante*. Statt $\det(A)$ schreibt man auch oft $|A|$.

Eine Determinante ist somit eine Funktion, die in jeder Spalte linear ist und dann null wird, wenn zwei Spalten gleich sind. Man kann zeigen, dass es für jedes n nur eine Funktion gibt, die Definition 5.16 erfüllt, und nennt für eine Matrix A die somit eindeutig bestimmte Zahl $\det(A)$ die *Determinante* von A . Ebenso gilt, dass dieselbe eindeutige Funktion gegeben ist, wenn alle Bedingungen in obiger Definition auf Zeilen und nicht auf Spalten definiert sind.

Der Zusammenhang zwischen Determinante und Invertierbarkeit einer Matrix ergibt sich aus Bedingungen (i) bis (iii) in obiger Definition: eine Matrix, die nicht vollen Rang hat, kann mittels Umformungen wie in (i) und (ii) auf die Form gebracht werden, dass zwei Spalten gleich sind, und hat damit wegen (iii) Determinante null.

Umgekehrt kann man für eine Matrix mit maximalem Rang durch Umformungen zeigen, dass dann die Determinante ungleich null ist. Zusammengefasst ergibt dies folgenden Satz.

Satz 5.9 Sei A eine $n \times n$ Matrix. Dann gilt

$$\det(A) \neq 0 \quad \text{genau dann wenn} \quad A^{-1} \text{ existiert.}$$

Zur Berechnung der Determinante bringt man eine Matrix durch geeignete Spalten- und Zeilenoperationen auf eine spezielle Form, die *obere Dreiecksform* genannt wird (siehe unten). Diese Operationen sind im nächsten Satz beschrieben.

Satz 5.10 Sei $A = (a_{-1}, \dots, a_{-n})$ eine $n \times n$ Matrix. Dann gilt für beliebige $i \neq j$ mit $1 \leq i, j \leq n$ und $\lambda \in \mathbb{R}$:

$$\begin{aligned} \det(a_{-1}, \dots, a_{-i} + \lambda a_{-j}, \dots, a_{-n}) &= \det(a_{-1}, \dots, a_{-i}, \dots, a_{-n}) \\ \det(a_{-1}, \dots, a_{-i}, \dots, a_{-j}, \dots, a_{-n}) &= -\det(a_{-1}, \dots, a_{-j}, \dots, a_{-i}, \dots, a_{-n}) \end{aligned}$$

Analoges gilt für Zeilenumformungen.

Die Determinante einer Matrix ändert sich also nicht, wenn man zu einer Spalte (Zeile) das Vielfache einer anderen Spalte (Zeile) hinzuzählt; sie ändert ihr Vorzeichen, wenn man zwei Spalten (Zeilen) vertauscht. Man spricht deshalb auch von der Determinante als *gerichtetes Volumen*, da die Reihenfolge der Spalten (Zeilen) in der Matrix eine Rolle spielt. Wir werden auf diese Orientierungseigenschaft nicht näher eingehen.

Definition 5.17 (Obere Dreiecksform)

Eine $n \times n$ Matrix der Form

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

nennt man eine *obere Dreiecksmatrix*.

Eine obere Dreiecksmatrix hat also unter der *Hauptdiagonale* a_{11}, \dots, a_{nn} nur Nullen. Eine solche Form ist für das Berechnen von Determinanten praktisch.

Satz 5.11 Für eine Matrix in oberer Dreiecksform gilt

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} = a_{11} a_{22} \dots a_{nn}$$

Dies kann man leicht mit der Definition von Determinanten nachrechnen. Somit berechnet man Determinanten konkret, indem man sie auf obere Dreiecksform bringt und dann die Diagonalelemente multipliziert. Wenn eines der Diagonalelemente null ist, so ist auch die Determinante null und die Matrix somit nicht invertierbar.

Nach soviel Theorie nun einige Beispiele.

Beispiel 5.18 Die Determinante einer allgemeinen 2×2 Matrix mit $a \neq 0$ ist

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ 0 & d - \frac{c}{a}b \end{vmatrix} = ad - bc$$

Somit ist eine 2×2 Matrix genau dann invertierbar, wenn $ad - bc \neq 0$ ist. Die in Beispiel 5.14 beschriebene Rotationsmatrix

$$\begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$$

ist als Rotation natürlich *volumserhaltend*, verändert also das Volumen des Einheitsquadrats nicht. Dies ist auch numerisch durch $\sin^2 x + \cos^2 x = 1$ gegeben. \square

Beispiel 5.19 Durch Zeilenumformungen erhält man

$$\begin{vmatrix} 1 & 2 & -1 \\ 3 & 0 & 2 \\ 2 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & -6 & 5 \\ 0 & -3 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & -6 & 5 \\ 0 & 0 & \frac{1}{2} \end{vmatrix} = -3 \quad \square$$

Man kann die Determinante von 3×3 Matrizen durch die Regel von Sarrus einfach berechnen.

Satz 5.12 (Satz von Sarrus) Die Determinante einer 3×3 Matrix (a_{ij}) kann man sich mit dem Anordnungsschema

$$\begin{pmatrix} \swarrow & \swarrow & \swarrow \\ a_{11} & a_{12} & a_{13} \\ \swarrow & \swarrow & \swarrow \\ a_{21} & a_{22} & a_{23} \\ \swarrow & \swarrow & \swarrow \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix}$$

als Summe der Hauptdiagonalprodukte minus Summe der Nebendiagonalelemente leicht merken:

$$(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}) - (a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12})$$

Beispiel 5.20 Mit dem Satz von Sarrus ist

$$\begin{vmatrix} 1 & 2 & 3 \\ -1 & -3 & -2 \\ 0 & 1 & 2 \end{vmatrix} = (1 \cdot (-3) \cdot 2 + 3 \cdot (-1) \cdot 1) - (1 \cdot (-2) \cdot 1 + 2 \cdot (-1) \cdot 2) \\ = -3 \quad \square$$

Nach obigen *Rechenregeln* für Determinanten stellt sich die Frage, wie diese Funktion eigentlich *definiert* ist. Dazu benötigen wir noch folgenden Begriff.

Definition 5.18 (Streichungsmatrix)

Sei A eine $n \times n$ Matrix. Dann nennt man die $(n - 1) \times (n - 1)$ Matrix A_{ij}^S , die man aus A durch Streichen der i -ten Zeile und j -ten Spalte erhält, eine *Streichungsmatrix* von A . Die Determinante einer Streichungsmatrix nennt man *Streichungsdeterminante*.

Folgender Satz gibt eine rekursive Definition einer Funktion, die die Determinanteneigenschaften erfüllt.

Satz 5.13 Für eine $n \times n$ Matrix A definieren wir rekursiv über n :

$$\det(A) := \begin{cases} \sum_{k=1}^n (-1)^{k+1} a_{1k} \det(A_{1k}^S) & \text{falls } n > 1 \\ a_{11} & \text{sonst.} \end{cases}$$

Dann erfüllt die so definierte Funktion die Determinantenbedingungen aus Definition 5.16.

Das Nachrechnen dieser Behauptung ist etwas aufwändig. Wir beschränken uns daher auf eine Erläuterung dieses Satzes: Er besagt, dass man die Determinantenberechnung rekursiv zerlegen kann, bis man bei einer 1×1 Matrix ankommt, deren Determinante gleich dem einzigen Matricelement ist. Den Zerlegungsschritt von n auf $n - 1$ nennt man *Entwickeln der Matrix*. In Satz 5.13 wird als Definition das Entwickeln nach der ersten Zeile angegeben; man kann aber auch nach jeder Zeile oder Spalte entwickeln. Die entsprechende Definition für Entwicklung nach Zeile i bzw. nach Spalte j lautet

$$\det(A) = \sum_{k=1}^n (-1)^{i+k} a_{ik} \det(A_{ik}^S) \quad \text{bzw.} \quad \det(A) = \sum_{k=1}^n (-1)^{k+j} a_{kj} \det(A_{kj}^S).$$

Man sieht, dass der Vorzeichenfaktor $(-1)^{i+k}$ bzw. $(-1)^{k+j}$ schachbrettartig positiv oder negativ ist.

Wir werden auf den Beweis der allgemeinen Eindeutigkeit von Determinantenberechnungen verzichten. Wenn es aber egal ist, ob wir eine Matrix nach Zeilen oder Spalten entwickeln, dann kommt es nur auf die Zahlen in der Matrix an und nicht, ob diese in den Zeilen oder Spalten angeordnet sind (ob wir also die Matrix oder die transponierte Matrix betrachten). Der folgende Satz fasst dies zusammen.

Satz 5.14 Sei A eine $n \times n$ Matrix. Dann gilt

$$\det(A) = \det(A^T).$$

Beispiel 5.21 Wir demonstrieren das Entwickeln einer Matrix am Beispiel einer beliebigen 3×3 Matrix. Dann ist

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Wenn man dies auflöst, erhält man (natürlich) das gleiche Ergebnis wie in Satz 5.12. Ebenso erhält man das gleiche Ergebnis, wenn man nach einer anderen Zeile oder Spalte entwickelt. \square

Beispiel 5.22 Die Determinante folgender 4×4 Matrix berechnet man am einfachsten durch Entwickeln nach der 2. Zeile. Es ist

$$\begin{vmatrix} 1 & 2 & -3 & 5 \\ 0 & 12 & 0 & 1 \\ 1 & 0 & -1 & 2 \\ -1 & 2 & 2 & 1 \end{vmatrix} = 12 \begin{vmatrix} 1 & -3 & 5 \\ 1 & -1 & 2 \\ -1 & 2 & 1 \end{vmatrix} + 1 \begin{vmatrix} 1 & 2 & -3 \\ 1 & 0 & -1 \\ -1 & 2 & 2 \end{vmatrix} \\ = 12 \cdot 9 + 1 \cdot (-6) = 102. \quad \square$$

Die folgende wichtige Eigenschaft von Determinanten kann man mit der Veranschaulichung von Determinanten als Volumsmaß beweisen.

Satz 5.15 Seien A, B zwei $n \times n$ Matrizen. Dann gilt

$$\det(A \cdot B) = \det(A) \det(B),$$

und somit für reguläres A

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Ebenso gilt folgender Satz über die Berechnung von Matrixinversen, der wegen der auftretenden Streichungsdeterminanten und der damit verbundenen aufwändigen Rechenschritte nur von theoretischem Interesse ist.

Satz 5.16 Sei A eine reguläre $n \times n$ Matrix. Dann gilt für die Matrix $B = (b_{ij})$ mit $b_{ij} = (-1)^{i+j} \det(A_{ji}^S)$

$$A^{-1} = \frac{1}{\det(A)} B.$$

Diesen Satz kann man sinnvollerweise nur für kleine Matrizen anwenden.

Beispiel 5.23 Die Inverse einer allgemeinen 2×2 Matrix ist mit Satz 5.16

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad \square$$

5.5 Gaußsches Eliminationsverfahren

Nachdem wir uns nun recht eingehend mit Determinanten und ihrem Zusammenhang mit den Lösungen linearer Gleichungssysteme beschäftigt haben, wenden wir uns nun dem Fall zu, dass die Koeffizientenmatrix des Gleichungssystems nicht quadratisch ist. Wir werden sehen, wie diese Matrix durch *Äquivalenzumformungen* auf eine Gestalt gebracht werden kann, bei der die Lösungen (keine, eine, oder unendlich viele) einfach abgelesen werden können.

Zuerst allerdings noch eine Überlegung, die für das Lösen von linearen Gleichungssystemen (abseits vom rein mechanischen Anwenden vorgegebener Rechenschritte) recht hilfreich ist. Wenn wir das lineare Gleichungssystem

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

in der Form

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

schreibt sieht man sofort, dass eine Lösung genau dann existiert, wenn der Vektor b in der linearen Hülle der Spaltenvektoren (a_{-1}, \dots, a_{-m}) liegt. Man schreibt

$$A|b = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right)$$

und nennt A die *Koeffizientenmatrix*, sowie $A|b$ die *erweiterte Matrix* des Gleichungssystems. Es gilt der folgende Satz.

Satz 5.17 Ein lineares Gleichungssystem $Ax = b$ besitzt genau dann eine Lösung, wenn $\text{rg}(A|b) = \text{rg}(A)$ ist.

Dieser Satz bietet jedoch keine “Abkürzung” um herauszufinden, ob ein gegebenes Gleichungssystem eine Lösung hat, da zur Rangberechnung die gleichen Operationen wie zur Lösungsberechnung notwendig sind. Diese Operationen sind wie folgt:

- (1) Addieren einer Zeile zu einer anderen Zeile,
- (2) Multiplizieren einer Zeile mit einem Faktor $\lambda \neq 0$,
- (3) Vertauschen von Spalten.

Das Vertauschen von Zeilen ist durch die beiden ersten Operationen realisierbar und deswegen nicht separat aufgeführt.

Satz 5.18 Die *elementaren Zeilenumformungen* (i) und (ii) oben sind *Äquivalenzumformungen*. Das bedeutet, dass diese Umformungen die Lösung eines Gleichungssystems nicht verändern. Die Umformung (iii) ändert nur die Position von Variablen in der Lösung, muss damit aber gesondert berücksichtigt werden.

Dieser Satz kann leicht nachgeprüft werden: Seien etwa

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$$

und

$$a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j$$

zwei Zeilen eines Gleichungssystems, und (x_1, \dots, x_n) eine Lösung. Dann ist (x_1, \dots, x_n) auch eine Lösung des Systems mit den Zeilen

$$\lambda a_{i1}x_1 + \lambda a_{i2}x_2 + \cdots + \lambda a_{in}x_n = \lambda b_i$$

$$\vdots$$

$$(a_{i1} + a_{j1})x_1 + (a_{i2} + a_{j2})x_2 + \cdots + (a_{in} + a_{jn})x_n = b_i + b_j$$

Dieselbe Argumentation gilt natürlich auch in umgekehrter Richtung. Ebensovienig wie die Lösungen wird der Rang eines Gleichungssystems durch Umformungen beeinflusst. Der Begriff der elementaren Spaltenumformungen ist dabei völlig analog zu dem der elementaren Zeilenumformungen definiert.

Satz 5.19 Der Rang einer Matrix bleibt bei elementaren Zeilen- und Spaltenumformungen unverändert.

Wir werden nun ein lineares Gleichungssystem dadurch lösen, dass wir es durch Äquivalenzumformungen in obere Dreiecksform bringen. In dieser Form lässt sich einerseits erkennen, ob das System lösbar ist, andererseits kann man in diesem Fall die Lösungen leicht ablesen. Den Algorithmus zum systematischen Anwenden der Umformungen nennt man *Gaußsches Eliminationsverfahren*. Wir betrachten zunächst einige Beispiele, um den Algorithmus anschaulicher zu machen. Ziel ist jeweils das Umformen in obere Dreiecksform und anschließendes Bestimmen der Lösung(en).

Beispiel 5.24 Das lineare Gleichungssystem

$$\begin{pmatrix} -1 & 0 & 2 \\ 2 & -1 & 1 \\ -3 & 1 & 2 \end{pmatrix} \cdot x = \begin{pmatrix} 8 \\ 3 \\ 5 \end{pmatrix}$$

formt man um, indem durch Addition geeigneter Vielfacher der ersten Zeile zunächst die Komponenten a_{21} und a_{31} eliminiert, und dann ein Vielfaches der zweiten Zeile zur dritten addiert, um a_{32} zu eliminieren. Es ergibt sich

$$\left(\begin{array}{ccc|c} -1 & 0 & 2 & 8 \\ 2 & -1 & 1 & 3 \\ -3 & 1 & 2 & 5 \end{array} \right) = \left(\begin{array}{ccc|c} -1 & 0 & 2 & 8 \\ 0 & -1 & 5 & 19 \\ 0 & 1 & -4 & -19 \end{array} \right) = \left(\begin{array}{ccc|c} -1 & 0 & 2 & 8 \\ 0 & -1 & 5 & 19 \\ 0 & 0 & 1 & 0 \end{array} \right),$$

somit $x_3 = 0, x_2 = -19$ und $x_1 = -8$. □

Beispiel 5.25 Wir beginnen direkt mit der erweiterten Matrix des Gleichungssystems und führen zu oben analoge Umformungen durch:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -1 & -2 & 1 & 2 \\ 1 & -1 & 5 & 0 \end{array} \right) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & 2 & 3 \\ 0 & -2 & 4 & -1 \end{array} \right) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & 2 & 3 \\ 0 & 0 & 0 & -7 \end{array} \right)$$

Wegen $0 \cdot x_3 = -7$ besitzt dieses Gleichungssystem keine Lösung. \square

Beispiel 5.26

$$\left(\begin{array}{ccc|c} 3 & 2 & 1 & -1 \\ 3 & 2 & 0 & 1 \\ -6 & -4 & -1 & 0 \end{array} \right) = \left(\begin{array}{ccc|c} 3 & 2 & 1 & -1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & -2 \end{array} \right) = \left(\begin{array}{ccc|c} 3 & 1 & 2 & -1 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

Im letzten Schritt haben wir die 2. und 3. Spalte vertauscht, um auf obere Dreiecksform zu kommen – dies muss bei der Lösungsangabe berücksichtigt werden! Die letzte Zeile $0 = 0$ ist keine Einschränkung; wir können x_3 frei wählen (etwa $x_3 = \lambda$) und erhalten so: $x_2 = -2$ und $x_1 = \frac{1}{3}(1 - 2\lambda)$. Nach Rückvertauschen von x_2 und x_3 ist die Lösungsmenge die Gerade

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid \text{es gibt ein } \lambda \in \mathbb{R}, \text{ sodass } \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ 0 \\ -2 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix} \right\} \quad \square$$

Wir können nun die einzelnen Schritte des Gaußschen Eliminationsverfahrens genauer angeben. Der angegebene Algorithmus ist für einfache Lesbarkeit ausgelegt (z.B. sind die `gotos` nicht nötig) und ist nicht notwendigerweise die effizienteste Computerimplementierung. Sei dafür $Ax = b$ ein lineares Gleichungssystem mit Koeffizientenmatrix $A \in \mathcal{M}_{m \times n}$ und $s = 1$ ein Iterationszähler.

```
while ( s ≤ m ) {
  (Pivot) // bringe nicht-null Element auf Position (s, s)
  if (ass ≠ 0)
    goto (Elimination);
  else if (exists i ∈ {s+1, ..., m} s.t. ais ≠ 0) {
    // wenn in Spalte nicht-null Element
    // vertausche mit dieser Zeile
    swap (ai-, as-);
    goto (Elimination);
  }
  else if (exists (i, j) ∈ {s, ..., m} × {s+1, ..., n} s.t.
    aij ≠ 0) {
    // wenn in Matrix unter Position (s, s) nicht-null Element
    // vertausche mit dieser Zeile und Spalte
    swap (ai-, as-);
    swap (a-j, a-s);
    goto (Elimination);
  }
  else break; // springe zu (Solution)
```

```

( Elimination ) // eliminiere Elemente in Spalte s unter Zeile s
  if ( s ≠ m ) {
    for ( i = s + 1, i ≤ m, i ++ ) {
      λ = ais/ass;
      ai- = ai- - λas-;
    }
    s = s + 1;
  }
} // ( while )
( Solution ) // bestimme Lösungen aus Dreiecksmatrix
  if ( (amm ≠ 0) && (m = n) ) {
    // eindeutig lösbare quadratische Matrix
    // Rückwärtseinsetzen und gegebenenfalls Variablenvertauschen
    // Siehe Beispiel 5.24
  }
  if ( (s ≤ m) &&
    !(exists (i, j) ∈ {s, ..., m} × {1, ..., n} s.t. aij ≠ 0) &&
    (exists i ∈ {s, ..., m} s.t. bi ≠ 0) ) {
    // nicht lösbar, da Widerspruch
    // Siehe Beispiel 5.25
  }
  if ( ((s ≤ m) &&
    !(exists (i, j) ∈ {s, ..., m} × {1, ..., n} s.t. aij ≠ 0) &&
    !(exists i ∈ {s, ..., m} s.t. bi ≠ 0) ) ||
    ((amm ≠ 0) && (m < n))) {
    // mehrdeutig lösbar
    // Siehe Beispiel 5.26 und unten
  }
}

```

Zusammenfassend kann man somit das Gaußsche Eliminationsverfahren als eine zweistufige Methode beschreiben: Im ersten Teil wird die erweiterte Matrix des Gleichungssystems auf obere Dreiecksform gebracht, aus der dann im zweiten Schritt die Lösung(en) ermittelt werden. Die beiden Fälle von keiner bzw. genau einer Lösung sind nicht weiter interessant. Die dritte Möglichkeit von unendlich vielen Lösungen sollte aber noch genauer betrachtet werden.

Nehmen wir dazu an, dass das Gleichungssystem durch Umformungen auf die Gestalt

$$\left(\begin{array}{ccccccc|c} a'_{11} & a'_{12} & \cdots & a'_{1r} & a'_{1(r+1)} & \cdots & a'_{1n} & b'_1 \\ 0 & a'_{22} & \cdots & a'_{2r} & a'_{2(r+1)} & \cdots & a'_{2n} & b'_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a'_{rr} & a'_{r(r+1)} & \cdots & a'_{rn} & b'_r \\ 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 & b'_{r+1} \\ \vdots & & & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 & b'_m \end{array} \right)$$

gebracht wurde; dabei ist $r = \text{rg}(A)$, da ja laut Satz 5.19 der Rang einer Matrix durch elementare Umformungen nicht verändert wird. Für den Fall $b'_{r+1} = \cdots = b'_m = 0$ und $r = n$ waren $m - n$ der Gleichungen Linearkombinationen der anderen n Gleichungen.

chungen, und das System ist eindeutig lösbar. Wenn bei diesem System unendlich viele Lösungen auftreten sollen, muss also $b'_{r+1} = \dots = b'_m = 0$ und $r < n$ gefordert werden.

Um die weitere Notation zu vereinfachen, nehmen wir nun an, dass obige Matrix durch elementare Zeilenoperationen so vereinfacht wurde, dass nur mehr Einsen in der Hauptdiagonale auftreten und dass auch (soweit wie möglich) Elemente über der Hauptdiagonale eliminiert wurden. Laut Annahme ist also $b'_{r+1} = \dots = b'_m = 0$, und wir schreiben statt der erweiterten Matrix wiederum das volle Gleichungssystem, um die Lösungen für x_1, \dots, x_n besser ablesen zu können. Mit den sich daraus ergebenden anderen Zahlenwerten erhalten wir

$$\begin{pmatrix} 1 & 0 & \dots & 0 & a''_{1(r+1)} & \dots & a''_{1n} \\ 0 & 1 & & 0 & a''_{2(r+1)} & \dots & a''_{2n} \\ \vdots & \ddots & \ddots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & a''_{r(r+1)} & \dots & a''_{rn} \\ 0 & \dots & \dots & \dots & 0 & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_r \\ x'_{r+1} \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} b''_1 \\ b''_2 \\ \vdots \\ b''_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (*)$$

Hier schreiben wir x'_1, \dots, x'_n , da sich die Reihenfolge der ursprünglichen Unbekannten x_1, \dots, x_n durch Spaltenvertauschungen geändert haben kann. Da das Gleichungssystem Rang r und $n > r$ Variablen hat, können $n - r$ Variablen (etwa x'_{r+1}, \dots, x'_n) frei gewählt werden, wie wir dies in Beispiel 5.26 getan haben. Wir setzen etwa

$$\lambda_1 = x'_{r+1}, \lambda_2 = x'_{r+2}, \dots, \lambda_{n-r} = x'_n$$

und können dann die einzelnen Gleichungen nach den x_i auflösen. Somit ist die gesuchte Lösung

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_r \\ x'_{r+1} \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} b''_1 - (\lambda_1 a''_{1(r+1)} + \dots + \lambda_{n-r} a''_{1n}) \\ b''_2 - (\lambda_1 a''_{2(r+1)} + \dots + \lambda_{n-r} a''_{2n}) \\ \vdots \\ b''_r - (\lambda_1 a''_{r(r+1)} + \dots + \lambda_{n-r} a''_{rn}) \\ \lambda_1 \\ \vdots \\ \lambda_{n-r} \end{pmatrix}$$

für beliebige Skalare $\lambda_1, \dots, \lambda_{n-r}$. Anders geschrieben ist die Struktur der Lösung besser zu erkennen:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_r \\ x'_{r+1} \\ x'_{r+2} \\ x'_{r+3} \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} b''_1 \\ b''_2 \\ \vdots \\ b''_r \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -a''_{1(r+1)} \\ -a''_{2(r+1)} \\ \vdots \\ -a''_{r(r+1)} \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -a''_{1(r+2)} \\ -a''_{2(r+2)} \\ \vdots \\ -a''_{r(r+2)} \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + \lambda_{n-r} \begin{pmatrix} -a''_{1n} \\ -a''_{2n} \\ \vdots \\ -a''_{rn} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Gegebenenfalls müssen in dieser Lösung noch die Spaltenvertauschungen rückgängig gemacht werden.

Da die Vektoren mit Koeffizienten a''_{ij} auf der rechten Seite sicher linear unabhängig sind (durch den Null/Eins Koordinatenblock), ist die Lösung somit eine lineare Mannigfaltigkeit der Dimension $n - r$. Es gilt somit allgemein folgender Satz.

Satz 5.20 Sei $L \neq \emptyset$ die Lösungsmenge des $m \times n$ Gleichungssystems $Ax = b$ mit Koeffizienten in K . Dann ist L eine lineare Mannigfaltigkeit in K^n , und es gilt

$$\dim(L) = n - \operatorname{rg}(A).$$

Beispiel 5.27 Wir lösen das lineare Gleichungssystem

$$\begin{pmatrix} 2 & -1 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ 3 & -3 & 2 & 3 \\ 0 & -3 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -3 \end{pmatrix}$$

Die erweiterte Koeffizientenmatrix dieses Systems wird folgendermaßen umgeformt (nach Vertauschen der ersten und zweiten Zeile)

$$\begin{pmatrix} 1 & 1 & 0 & 1 & | & 2 \\ 2 & -1 & 1 & 0 & | & 1 \\ 3 & -3 & 2 & 3 & | & 0 \\ 0 & -3 & 1 & 0 & | & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 & 1 & | & 2 \\ 0 & -3 & 1 & 0 & | & -3 \\ 0 & -6 & 2 & 0 & | & -6 \\ 0 & -3 & 1 & 0 & | & -3 \end{pmatrix} \rightarrow$$

$$\begin{pmatrix} 1 & 1 & 0 & 1 & | & 2 \\ 0 & 1 & -\frac{1}{3} & 0 & | & 1 \\ 0 & 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & \frac{1}{3} & 1 & | & 1 \\ 0 & 1 & -\frac{1}{3} & 0 & | & 1 \\ 0 & 0 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$$

Damit hat das System die Form (\star) ; der Rang der Koeffizientenmatrix ist 2. Demnach ist die Lösung die folgende zweidimensionale Mannigfaltigkeit des \mathbb{R}^4 :

$$L = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 \mid \text{es gibt } \lambda_1, \lambda_2 \in \mathbb{R}, \text{ sodass} \right.$$

$$\left. \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \square$$

Nachdem wir nun das Lösen von allgemeinen Gleichungssystemen ziemlich genau behandelt haben, kommen wir noch zu einer weiteren Anwendung der hier präsentierten Matrizenumformungen: Das Invertieren von Matrizen. Dieses Problem passt folgendermaßen in den Kontext des Gleichungslösens: Die Inverse $B = (b_{-1}, \dots, b_{-n})$

einer $n \times n$ Matrix A erfüllt die Gleichung $A \cdot B = I_n$. Man muss somit Vektoren b_{-1}, \dots, b_{-n} finden, die die n Gleichungen

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} b_{-1} = e_{-1} \quad \dots \quad \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} b_{-n} = e_{-n}$$

lösen. Hier bezeichnet e_{-i} die i -te Spalte der Einheitsmatrix I_n . Da der Lösungsaufwand darin besteht, diese Gleichungssysteme durch Umformungen auf die Gestalt (\star) zu bringen, und dieser Vorgang für jedes der n Systeme gleich ist, sollte man diese Systeme alle simultan lösen, um so die notwendigen Umformungen nur einmal durchzuführen. Daraus ergibt sich folgender Satz.

Satz 5.21 (Matrizeninvertierung) Seien A, B zwei reguläre Matrizen. Die Umformung des erweiterten Gleichungssystems

$$\left(\begin{array}{ccc|c} a_{11} & \dots & a_{nn} & I_n \\ \vdots & \ddots & \vdots & \\ a_{n1} & \dots & a_{nn} & \end{array} \right) \longrightarrow \left(\begin{array}{c|ccc} I_n & b_{11} & \dots & b_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & \dots & \dots & b_{nn} \end{array} \right)$$

ist immer möglich, und es gilt $A^{-1} = B$.

Beispiel 5.28 Es gilt

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & 1 \\ \frac{2}{3} & -\frac{1}{3} & 0 \end{pmatrix},$$

wie man aus folgenden Transformationen sehen kann:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & -3 & -2 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \end{array} \right) \xrightarrow{\text{Zeilenvertauschen}}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 0 & -3 & -2 & 1 & 0 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & 0 \end{array} \right) \rightarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 1 & 0 & -\frac{1}{3} & -\frac{1}{3} & 1 \\ 0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & 0 \end{array} \right) \quad \square$$

Zum Abschluss dieser Überlegungen über das Lösen linearer Gleichungssysteme betrachten wir diese Systeme kurz aus der Richtung linearer Funktionen. Durch die Zuordnungsvorschrift $f(x) = Ax$ ist ja eindeutig eine lineare Funktion definiert; den Kern von f bilden all jene Vektoren x , für die $Ax = 0$ gilt (Um diesen Kern zu bestimmen, muss man das homogene Gleichungssystem $Ax = 0$ lösen). Wenn man nun bereits *eine* konkrete Lösung x_0 des ursprünglichen Gleichungssystems kennt,

so kann man beliebige Vektoren v aus $\text{kern}(f)$ dazuzählen (für die ja $Av = 0$ gilt), und erhält so weitere Lösungen. Sei dazu $v \in \text{kern}(f)$; da x_0 eine Lösung von $Ax = b$ ist, erhalten wir

$$A(x_0 + v) = Ax_0 + Av = b + 0 = b.$$

Dieser Zusammenhang kann wie folgt formuliert werden.

Satz 5.22 Sei $Ax = b$ ein lineares Gleichungssystem mit einer konkreten Lösung x_0 , und f die durch die Matrix A gegebene lineare Funktion. Dann ist die Lösungsmenge M des Gleichungssystems die lineare Mannigfaltigkeit

$$M = x_0 + \text{kern}(f).$$

Dieser Satz besagt also, dass man alle Lösungen eines Gleichungssystems erhält, indem man zu *einer konkreten* Lösung beliebige Elemente der Lösungsmenge des dazugehörigen homogenen Gleichungssystems dazuzählt. Somit ist der Lösungsraum eines linearen Gleichungssystems $Ax = b$ "so groß" (dimensionsmäßig) wie der Raum, der von den Vektoren aufgespannt wird, die durch A auf null abgebildet werden (wie auch schon Satz 5.20 ausgesagt hat).

Kapitel 6

Geometrie in \mathbb{R}^n

In diesem Kapitel werden wir das Grundgerüst der linearen Algebra um den Begriff des *Skalarprodukts* erweitern. Wir werden sehen, wie man mit wenig mathematischem Aufwand über die Verwendung von *Projektionen* erstaunliche Ergebnisse erreichen kann. Als Einführung betrachten wir im ersten Abschnitt die *Länge* von und den *Winkel* zwischen Vektoren.

6.1 Länge und Winkel

In reellen Vektorräumen macht es oftmals Sinn, die *Länge* eines Vektors bestimmen zu können. Diese ist wie folgt definiert.

Definition 6.1 (Skalarprodukt, Länge)

Seien $v, w \in \mathbb{R}^n$ zwei reelle Vektoren. Dann nennt man die Funktion

$$\begin{aligned} \cdot : \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R} \\ (v, w) &\mapsto \sum_{i=1}^n v_i w_i \end{aligned}$$

das *Skalarprodukt* von v und w . Oft lässt man das Funktionszeichen weg und schreibt nur vw und v^2 für $v \cdot w$ bzw. $v \cdot v$. Die *Länge* eines Vektors ist mit Hilfe des Skalarprodukts definiert als

$$\|v\| = \sqrt{v^2}.$$

Für das Skalarprodukt macht es keinen Unterschied, ob reelle Vektoren als Zeilen oder Spalten geschrieben werden. Bis zur Notwendigkeit einer Unterscheidung (die bei der Multiplikation mit Matrizen auftritt) werden wir daher beide Schreibweisen als äquivalent betrachten.

Beispiel 6.1 Das Skalarprodukt der beiden Vektoren $v = (1, 2, -1, 4)$ und $w = (-1, 0, 2, 1)$ ist

$$v \cdot w = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} = 1.$$

Die Längen der beiden Vektoren sind

$$\|v\| = \left\| \begin{pmatrix} 1 \\ 2 \\ -1 \\ 4 \end{pmatrix} \right\| = \sqrt{22} \quad \text{und} \quad \|w\| = \left\| \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \right\| = \sqrt{6}. \quad \square$$

Mit dem Skalarprodukt kann auf den Vektoren (fast) wie mit der Multiplikation auf den reellen Zahlen gerechnet werden. So gelten etwa folgende Rechenregeln, die nicht schwer nachzuprüfen sind:

$$\begin{aligned} v \cdot w &= w \cdot v \\ v^2 &= \|v\|^2 \\ u \cdot (v + w) &= u \cdot v + u \cdot w. \end{aligned}$$

Mit der Länge eines Vektors haben wir auch die Möglichkeit, den Abstand zwischen zwei Vektoren (Punkten) zu definieren. Mathematisch fordert man, dass eine solche *Metrik* folgende Axiome erfüllt.

Definition 6.2 (Metrik)

Sei V ein Vektorraum. Dann nennt man eine Funktion $d : V \times V \rightarrow \mathbb{R}$, für die die drei Bedingungen

$$d(v, w) \geq 0 \quad \text{mit } d(v, w) = 0 \text{ genau dann wenn } v = w \quad (\text{NNG})$$

$$d(v, w) = d(w, v) \quad (\text{SYM})$$

$$d(v, w) \leq d(v, u) + d(u, w) \quad (\text{TRI})$$

für alle $u, v, w \in V$ gelten, eine *Metrik* auf V .

Geometrisch motiviert definieren wir den Abstand zweier Vektoren v und w im \mathbb{R} durch

$$d(v, w) = \|v - w\|.$$

Für uns ist nun interessant, ob die damit definierte Abstandsfunktion auch eine Metrik im Sinne obiger Definition ist. Dies lässt sich nachprüfen: So gilt für die Bedingung (NNG)

$$\|v - w\| = \sqrt{(v_1 - w_1)^2 + \cdots + (v_n - w_n)^2} \geq 0.$$

Diese Länge ist genau dann null, wenn alle Summanden null sind – und dies ist nur bei $v = w$ der Fall. Die Symmetriebedingung (SYM) folgt ebenso direkt aus der Definition der Abstandsfunktion.

Die letzte Bedingung (TRI) ist etwas schwieriger zu beweisen. Diese Bedingung, auch *Dreiecksungleichung* genannt besagt, dass der kürzeste Abstand zwischen zwei Punkten eine Gerade ist: Der Umweg über einen weiteren Punkt kann die Distanz nur länger machen (siehe auch Abbildung 6.1). Zum Beweis rechnen wir mit den Quadraten der Abstände. Dies ist korrekt, da die Ungleichung (TRI) genau dann gilt,

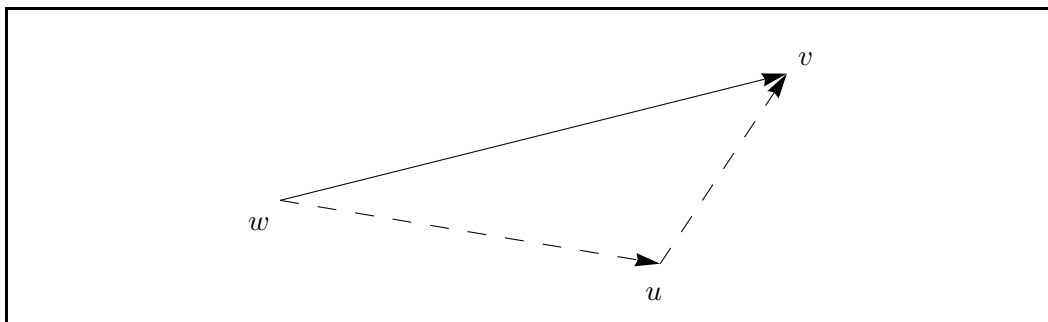


Abbildung 6.1: Illustration der Dreiecksungleichung: Der kürzeste Weg zwischen w und v ist die gerade Linie $v - w$.

wenn sie auch für die Quadrate der beiden Seiten gilt. Wir formen folgendermaßen um:

$$\begin{aligned}
 \|v - w\|^2 &= \|v - u + u - w\|^2 \\
 &= ((v - u) + (u - w)) \cdot ((v - u) + (u - w)) \\
 &= (v - u)^2 + 2(v - u) \cdot (u - w) + (u - w)^2 \\
 &\leq (v - u)^2 + 2\|v - u\|\|u - w\| + (u - w)^2 && \text{(CS)} \\
 &\leq \|v - u\|^2 + 2\|v - u\|\|u - w\| + \|u - w\|^2 \\
 &\leq (\|v - u\| + \|u - w\|)^2
 \end{aligned}$$

und somit auch für die Wurzel

$$\|v - w\| \leq \|v - u\| + \|u - w\|.$$

In dieser Herleitung ist an der mit (CS) gekennzeichneten Zeile noch ein Umformungsschritt, von dessen Richtigkeit wir uns erst genauer überzeugen müssen. Dieser Schritt folgt unmittelbar aus dem nächsten Satz, den wir nicht beweisen werden.

Satz 6.1 (Cauchy-Schwartz'sche Ungleichung) Für zwei Vektoren v und w in \mathbb{R}^n gilt

$$|v \cdot w| \leq \|v\|\|w\|.$$

Gleichheit tritt genau dann ein, wenn v und w linear abhängig sind.

Da natürlich $v \cdot w \leq |v \cdot w|$ ist, gilt also auch

$$v \cdot w \leq \|v\|\|w\|,$$

wie wir für Umformung (CS) benötigen.

Mit Satz 6.1 gilt somit auch die Dreiecksungleichung, und damit insgesamt folgender Satz.

Satz 6.2 Die durch das Skalarprodukt auf \mathbb{R}^n gegebene Länge ist eine Metrik auf \mathbb{R} .

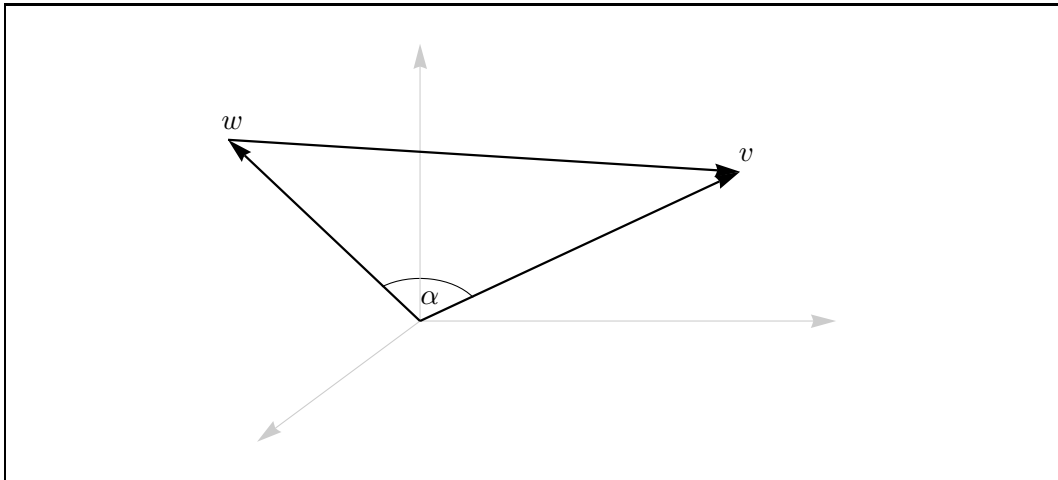


Abbildung 6.2: Illustration zur Herleitung des Winkels zwischen zwei Vektoren im \mathbb{R}^n .

Mit den Definitionen über Skalarprodukt und Länge von Vektoren kann man auch den *Winkel* zwischen zwei Vektoren im \mathbb{R}^n definieren. Wir verallgemeinern dazu unsere geometrische Vorstellung aus dem \mathbb{R}^3 . Mit dem Kosinussatz gilt für das in Abbildung 6.2 dargestellte Dreieck die Gleichung

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos \alpha,$$

wobei α der Winkel zwischen v und w ist. Diese Gleichung ergibt ausmultipliziert unter Verwendung von $\|u\|^2 = u^2$

$$v^2 - 2v \cdot w + w^2 = v^2 + w^2 - 2\|v\|\|w\| \cos \alpha.$$

Wir *definieren* nun den Winkel zwischen zwei Vektoren im \mathbb{R}^n durch Umformen dieser Gleichung.

Definition 6.3 (Winkel zwischen Vektoren)

Seien v und w zwei Vektoren im \mathbb{R}^n . Dann ist der Kosinus des Winkels α zwischen v und w definiert als

$$\cos \alpha = \frac{v \cdot w}{\|v\|\|w\|}.$$

Hier muss man nun nachprüfen, ob diese Definition überhaupt mathematisch korrekt ist: Der Kosinus eines Winkels muss immer zwischen -1 und 1 liegen. Mit Satz 6.1 gilt das auch für obige Definition.

Beispiel 6.2 Der Winkel α zwischen der x -Achse und der Diagonalen im \mathbb{R}^2 ist

durch die Gleichung

$$\cos(\alpha) = \frac{\begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| \left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\|} = \frac{1}{\sqrt{2}}$$

als $\alpha = \arccos\left(\frac{\sqrt{2}}{2}\right) = \frac{\pi}{4} = 45^\circ$ zu berechnen. \square

Beispiel 6.3 Der Winkel α zwischen den Vektoren $(1, 2, 0)$ und $(0, -1, 3)$ im \mathbb{R}^3 ist wegen

$$\cos(\alpha) = \frac{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right\| \left\| \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \right\|} = \frac{2}{\sqrt{5}\sqrt{10}}$$

gleich $\alpha = \arccos\left(\frac{\sqrt{2}}{5}\right) = 1.284 = 73.57^\circ$. \square

6.2 Projektionen

Wir werden in diesem Abschnitt sehen, wie man Vektoren aufeinander projizieren kann, und welche theoretischen Schlussfolgerungen sich daraus ziehen lassen. Dafür müssen wir uns zuerst überlegen, wie man sinnvoll definieren kann, dass zwei Vektoren aufeinander senkrecht stehen. Dafür ist der im letzten Abschnitt eingeführte Begriff des Winkels zwischen Vektoren hilfreich. Da der Kosinus bei einem Winkel von 90° null wird, definiert man den Begriff *senkrecht* für Vektoren wie folgt.

Definition 6.4 (Normalvektor)

Sei $v \in \mathbb{R}^n$. Einen Vektor $w \in \mathbb{R}^n$ mit

$$v \cdot w = 0$$

nennt man *Normalvektor* von v . Man sagt auch, dass v und w *senkrecht*, *normal* oder *orthogonal* aufeinander stehen und schreibt dafür $v \perp w$.

Man kann im folgenden Beispiel sehen, dass die Begriffe *Skalarprodukt* und damit auch *Normalität* in vielen Vektorräumen sinnvoll anwendbar sind.

Beispiel 6.4 Für gegebenes n sein \mathcal{P}_n der Vektorraum aller Polynomfunktionen, deren Grad höchstens n ist. Wir definieren ein Skalarprodukt auf \mathcal{P}_n durch

$$f \cdot g = \int_{-1}^1 f(x)g(x)dx.$$

Dann stehen die sechs Vektoren

$$P_0 = 1$$

$$P_1 = x$$

$$P_2 = x^2 - \frac{1}{3}$$

$$P_3 = x^3 - \frac{3}{5}x \quad P_4 = x^4 - \frac{6}{7}x^2 + \frac{3}{35} \quad P_5 = x^5 - \frac{10}{9}x^3 + \frac{5}{21}x$$

(genannt *Legendre-Polynome*) senkrecht aufeinander und bilden eine Basis von \mathcal{P}_6 . Die zweite Behauptung ist leicht nachzuprüfen, da die Vektoren linear unabhängig sind und \mathcal{P}_6 Dimension 6 hat. Für die erste Behauptung überprüfen wir die Normalität nur eines von 15 Paaren von Polynomen, etwa P_2 und P_4 :

$$\begin{aligned} \int_{-1}^1 \left(x^2 - \frac{1}{3}\right) \left(x^4 - \frac{6}{7}x^2 + \frac{3}{35}\right) dx &= \int_{-1}^1 \left(x^6 - \frac{25}{21}x^4 + \frac{13}{35}x^2 - \frac{1}{35}\right) dx \\ &= \frac{1}{7}x^7 - \frac{5}{21}x^5 + \frac{13}{1-5}x^3 - \frac{1}{35}x \Big|_{-1}^1 \\ &= 0. \end{aligned} \quad \square$$

Bei reellen Vektorräumen ist die Überprüfung der Orthogonalität von Vektoren einfacher.

Beispiel 6.5 Von den drei Vektoren $a = (-1, 2, 1, -1)$, $b = (2, 5, 0, 8)$ und $c = (-7, -2, 2, 3)$ stehen sowohl a und b als auch b und c senkrecht aufeinander, nicht aber a und c , wie man aus unterer Rechnung leicht sieht.

$$\begin{pmatrix} -1 \\ 2 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \\ 0 \\ 8 \end{pmatrix} = 0, \quad \begin{pmatrix} 2 \\ 5 \\ 0 \\ 8 \end{pmatrix} \cdot \begin{pmatrix} -7 \\ -2 \\ 2 \\ 3 \end{pmatrix} = 0, \quad \text{aber} \quad \begin{pmatrix} -1 \\ 2 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -7 \\ -2 \\ 2 \\ 3 \end{pmatrix} = 2. \quad \square$$

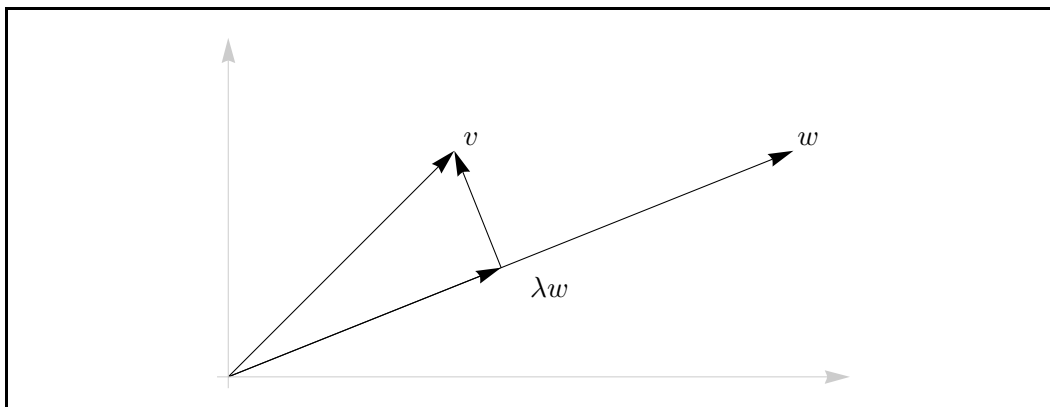
Mit dem Begriff der Normalität zweier Vektoren lässt sich auch einer der bekanntesten Sätze der Geometrie formulieren.

Satz 6.3 (Pythagoras) Seien $v, w \in V$ zwei Vektoren mit $v \perp w$. Dann gilt

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Beispiel 6.6 Im letzten Beispiel haben wir gezeigt, dass die beiden Vektoren $(-1, 2, 1, -1)$ und $(2, 5, 0, 8)$ und die beiden Vektoren $(2, 5, 0, 8)$ und $(-7, -2, 2, 3)$ jeweils senkrecht aufeinander stehen. Mit dem Satz von Pythagoras ist das Längenquadrat der Summe der Vektoren gleich der Summe der Längenquadrate der Vektoren, wie man anhand dieses Beispiels nachprüfen kann. Es ist

$$\left\| \begin{pmatrix} -1 \\ 2 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ 5 \\ 0 \\ 8 \end{pmatrix} \right\|^2 = 100 = \left\| \begin{pmatrix} -1 \\ 2 \\ 1 \\ -1 \end{pmatrix} \right\|^2 + \left\| \begin{pmatrix} 2 \\ 5 \\ 0 \\ 8 \end{pmatrix} \right\|^2 = 7 + 93$$

Abbildung 6.3: Illustration zur Projektion des Vektors v auf den Vektor w .

und

$$\left\| \begin{pmatrix} 2 \\ 5 \\ 0 \\ 8 \end{pmatrix} + \begin{pmatrix} -7 \\ -2 \\ 2 \\ 3 \end{pmatrix} \right\|^2 = 159 = \left\| \begin{pmatrix} 2 \\ 5 \\ 0 \\ 8 \end{pmatrix} \right\|^2 + \left\| \begin{pmatrix} -7 \\ -2 \\ 2 \\ 3 \end{pmatrix} \right\|^2 = 93 + 66. \quad \square$$

Da wir die Länge von Vektoren berechnen können ist es auch möglich, einen Vektor auf einen anderen zu *projizieren*. Für gegebene Vektoren v und w möchte man denjenigen Abschnitt auf w finden, der sich durch senkrechte Projektion von v auf w ergibt. Die Problemstellung ist in Abbildung 6.3 erläutert. Gesucht ist derjenige Skalarfaktor λ , für den der Vektor $v - \lambda w$ senkrecht auf w steht. Dies lässt sich so ausdrücken:

$$(v - \lambda w) \cdot w = 0 \quad \Leftrightarrow \quad \lambda = \frac{v \cdot w}{\|w\|^2}.$$

Wir verwenden diese Herleitung für eine allgemeine Definition der orthogonalen Projektion.

Definition 6.5 (Projektion)

Seien v und w zwei Vektoren. Dann bezeichnet man den Vektor

$$\text{proj}_w(v) = \frac{v \cdot w}{\|w\|^2} w$$

als *orthogonale Projektion* von v auf w .

Man kann an einem einfachen Beispiel überprüfen, dass diese Formel mit unserer Intuition von Projektion übereinstimmt.

Beispiel 6.7 Die Projektion eines beliebigen Vektors $v = (v_1, v_2, v_3) \in \mathbb{R}^3$ auf die

y -Achse $\{w \in \mathbb{R}^3 \mid w = \lambda(0, 1, 0)\}$ ergibt den Vektor

$$\text{proj}_{y\text{-Achse}}(v) = \frac{\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}{\left\| \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\|^2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ v_2 \\ 0 \end{pmatrix}.$$

Dies deckt sich mit unserer Erwartung, dass eine Projektion auf die y -Achse dem Abschnitt entspricht, der durch die y -Koordinate des Punkts gegeben ist. \square

Beim Arbeiten mit Projektionen lohnt es sich oft, für den umgebenden Vektorraum spezielle Basen zu konstruieren, mit denen sich dann besonders einfach rechnen lässt. Diese Basen, bestehend aus Vektoren, die senkrecht aufeinander stehen, sind wie folgt definiert.

Definition 6.6 (Orthonormalbasen)

Eine Basis $\{b_1, \dots, b_n\}$ eines Vektorraums V mit den Eigenschaften

$$\begin{aligned} b_i &\perp b_j && \text{für alle } i \neq j \\ \|b_i\| &= 1 && \text{für alle } i \end{aligned}$$

nennt man *Orthonormalbasis (ONB)* von V .

Beispiel 6.8 Das einfachste Beispiel einer Orthonormalbasis ist die Standardbasis des \mathbb{R}^n , wie man sofort nachprüfen kann. \square

Beispiel 6.9 Aus Beispiel 5.14 wissen wir, dass die Rotation um den Winkel φ relativ zum Ursprung im \mathbb{R}^2 durch die lineare Abbildung

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

ausgedrückt wird. Wenn wir die Spalten dieser Rotationsmatrix als Vektoren im \mathbb{R}^2 auffassen, so formen sie eine Orthonormalbasis des \mathbb{R}^2 . Dies lässt sich numerisch aus

$$\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \cdot \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} = -\cos \varphi \sin \varphi + \sin \varphi \cos \varphi = 0$$

und

$$\left\| \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \right\| = \left\| \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \right\| = \sqrt{\sin^2 \varphi + \cos^2 \varphi} = 1$$

für beliebige Winkel φ sofort überprüfen. \square

Im letzten Beispiel haben wir nur die beiden Bedingungen einer Orthonormalbasis aus Definition 6.6 nachgeprüft und nicht, ob die Vektoren überhaupt eine Basis bilden. Die Eigenschaften der linearen Unabhängigkeit und Erzeugendensystems folgen aber aus den Orthonormalitätsbedingungen, wie man unschwer nachrechnen kann.

Orthonormalbasen haben gegenüber "herkömmlichen" Basen den großen Vorteil, dass man mit ihnen leichter rechnen kann. So muss man zur Berechnung der Koordinaten eines Vektors bezüglich einer Orthonormalbasis keine aufwändigen Berechnungen durchführen, wie dies bei anderen Koordinatentransformationen der Fall ist. Zur Erinnerung: Ein Koordinatenvektor $(\lambda_1, \dots, \lambda_n)$, der Koordinaten bezüglich einer Basis repräsentiert, deren Spalten die Matrix A bilden, wird durch folgende Formel in einen Koordinatenvektor (μ_1, \dots, μ_n) bezüglich der Basis umgerechnet, deren Spalten durch die Matrix B gebildet werden:

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = B^{-1} \cdot A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Eine Vereinfachung ergibt sich, wenn der zu transformierende Vektor bezüglich der Standardbasis gegeben ist, da dann die Matrix A in obiger Gleichung die Einheitsmatrix ist.

Koordinaten bezüglich Orthonormalbasen können leichter berechnet werden. Eine einfache Überlegung zeigt, warum das so ist. Sei dazu (b_1, \dots, b_n) eine geordnete Orthonormalbasis, bezüglich derer der Vektor v die Koordinaten $\lambda_1, \dots, \lambda_n$ hat. Somit ist

$$v = \lambda_1 b_1 + \dots + \lambda_n b_n,$$

und wir wollen die λ_i bestimmen. Zuerst müssen wir uns überlegen, dass für beliebige Vektoren u, v und w gilt: aus $u = v$ folgt $u \cdot w = v \cdot w$ (sofort einsichtig). Wir multiplizieren dann die Koordinatendarstellung oben auf beiden Seiten mit beliebigem b_i und erhalten

$$b_i \cdot v = \lambda_1 \underbrace{b_i \cdot b_1}_{=0} + \dots + \lambda_i \underbrace{b_i \cdot b_i}_{=1} + \dots + \lambda_n \underbrace{b_i \cdot b_n}_{=0}.$$

Somit ergibt sich folgendes einfache Resultat.

Satz 6.4 Sei $B = (b_1, \dots, b_n)$ eine Orthonormalbasis. Dann ist die i -te Koordinate λ_i eines Vektors v bezüglich B gegeben durch

$$\lambda_i = b_i \cdot v.$$

Der obige Satz lässt sich auch geometrisch interpretieren: Wegen $\|b_i\| = 1$ ist $\text{proj}_{b_i}(v) = b_i \cdot v b_i$, die Koordinate $b_i \cdot v$ von v bezüglich b_i kann man also auch als Projektion interpretieren. Über Projektionen ist es somit auch möglich, Koordinaten von Vektoren bezüglich *Orthogonalbasen* zu berechnen (also Basen aus orthogonalen Vektoren, die aber nicht Länge 1 haben). Der Zusammenhang zwischen Koordinatenberechnung und Projektion ist auch in Abbildung 6.4 dargestellt.

Beispiel 6.10 Die drei Vektoren $\frac{1}{\sqrt{3}}(1, 1, 1)$, $\frac{1}{\sqrt{6}}(-1, 2, -1)$ und $\frac{1}{\sqrt{2}}(-1, 0, 1)$ bilden eine Orthonormalbasis des \mathbb{R}^3 , wie man leicht nachprüfen kann. Die Koordinaten

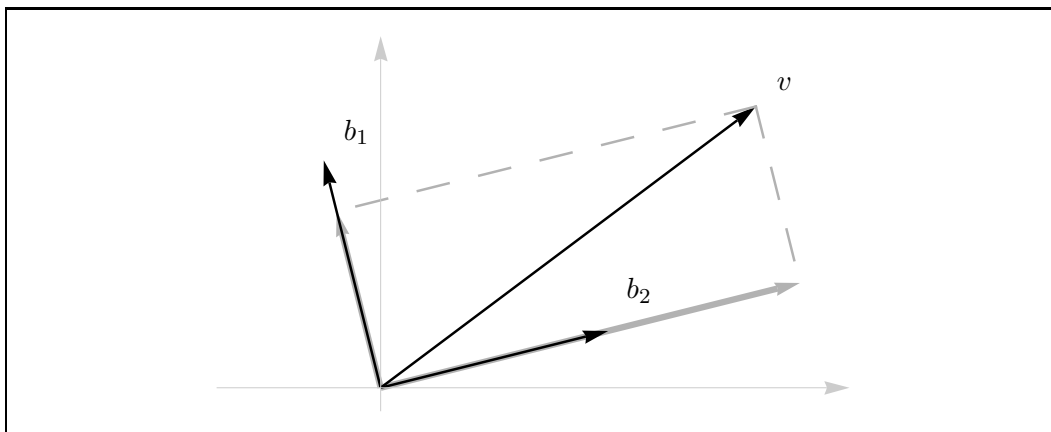


Abbildung 6.4: Der Vektor v lässt sich als Summe der beiden grauen Vektoren schreiben, die Projektionen auf die Orthonormalbasis $\{b_1, b_2\}$ sind.

des Punktes $(-1, 0, 2)$ bezüglich dieser Basis sind

$$\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{3}}, \quad \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = -\frac{1}{\sqrt{6}},$$

$$\frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = \frac{3}{\sqrt{2}}.$$

Die Berechnungen in diesem Beispiel können mittels Matrizenmultiplikation kürzer geschrieben werden. Wenn wir die drei Basiselemente nämlich in den *Zeilen* einer Matrix anordnen (und nicht wie sonst üblich in den *Spalten*), kann man obige Rechnungen schreiben als

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} \\ \frac{3}{\sqrt{2}} \end{pmatrix}. \quad \square$$

Orthonormalbasen haben also den großen Vorteil, dass Koordinatenberechnungen und Koordinatentransformationen leichter durchzuführen sind als mit normalen Basen. Die wenigsten Mengen von linear unabhängigen Vektoren bestehen aber aus paarweise orthogonalen Vektoren; um die Vorteile von Orthonormalbasen nützen zu können, müssen diese Basen erst orthogonalisiert und dann normiert werden. Es ist durch eine einfache geometrische Überlegung einsichtig, dass dies für alle Mengen von Vektoren möglich ist.

Wie man sich anhand einer Skizze veranschaulichen kann (etwa Abbildung 6.3 auf Seite 121), steht die Verbindungslinie zwischen dem zu projizierenden Vektor v und dem projizierten Vektor $\text{proj}_w(v)$ immer senkrecht auf den Vektor w , auf den projiziert wird. Es gilt also allgemein

$$v - \text{proj}_w(v) \perp w.$$

Dabei ist wichtig, dass sich die lineare Hülle von v und w nicht ändert, wenn man v durch $v - \text{proj}_w(v)$ ersetzt.

Beispiel 6.11 Mit dieser Einsicht können wir die beiden Vektoren $(1, 2, 3)$ und $(1, 0, 1)$ orthogonalisieren, indem wir etwa den Vektor $(1, 2, 3)$ so umformen, dass er orthogonal zu $(1, 0, 1)$ steht. Es ist

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \text{proj}_{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \frac{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\|^2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}.$$

Es ist sofort ersichtlich, dass die Vektoren $(1, 0, 1)$ und $(-1, 2, 1)$ senkrecht aufeinander stehen. \square

Diese Methode kann man auf mehrere Vektoren verallgemeinern: Man iteriert durch die Menge der Vektoren und zieht von jedem Vektor v die Projektionen auf alle vorherig orthogonalisierten Vektoren ab, um v auf diese Vektoren senkrecht zu stellen. Dieser Verfahren wird *Gram-Schmidt'sches Orthogonalisierungsverfahren* genannt. Die so erzeugte Menge von Vektoren spannt dabei den gleichen Vektorraum wie die ursprüngliche Menge auf.

Satz 6.5 (Gram-Schmidt) Seien b_1, \dots, b_n eine Basis von V . Dann ist die Menge von Vektoren

$$\begin{aligned} v_1 &= b_1 \\ v_2 &= b_2 - \text{proj}_{v_1}(b_2) \\ v_3 &= b_3 - \text{proj}_{v_1}(b_3) - \text{proj}_{v_2}(b_3) \\ &\vdots \\ v_n &= b_n - \text{proj}_{v_1}(b_n) - \text{proj}_{v_2}(b_n) - \dots - \text{proj}_{v_{n-1}}(b_n) \end{aligned}$$

eine orthogonale Basis von V .

Um aus der so erzeugten Basis eine Orthonormalbasis zu machen, muss man die einzelnen Vektoren nur mehr normalisieren. Wir illustrieren dies an einem Beispiel.

Beispiel 6.12 Die drei Vektoren $(1, 0, -1)$, $(2, 0, 1)$ und $(1, 1, 1)$ bilden eine Basis des \mathbb{R}^3 . Um daraus eine Orthonormalbasis zu erzeugen, wenden wir das Gram-Schmidt'sche Verfahren an und erhalten so die Vektoren

$$\begin{aligned} v_1 &= \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ v_2 &= \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} - \text{proj}_{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} \end{aligned}$$

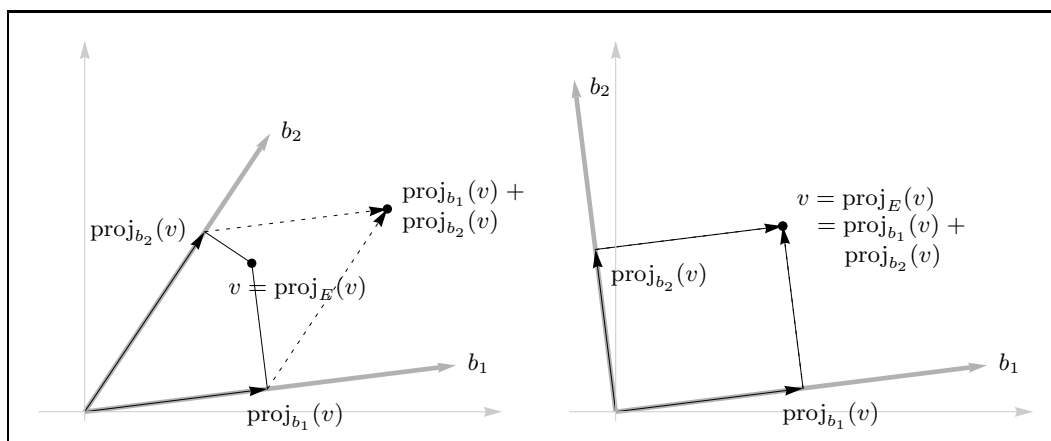


Abbildung 6.5: Illustration der Projektion von v auf die Ebene E , die von b_1 und b_2 aufgespannt wird. Die Ansichten sind von oben entlang der Projektion von v auf E , sodass v und $\text{proj}_E(v)$ in den Abbildungen identisch sind. Links sind b_1 und b_2 nicht orthogonal, sodass $\text{proj}_E(v) \neq \text{proj}_{b_1}(v) + \text{proj}_{b_2}(v)$ ist. Wenn b_1 und b_2 orthogonal sind, dann gilt die Aussage von Satz 6.6.

$$v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \text{proj}_{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \text{proj}_{\frac{1}{2} \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Durch Nachrechnen kann man sich überzeugen, dass diese Vektoren senkrecht aufeinander stehen. Mit Normalisieren erhalten wir die Vektoren $\frac{1}{\sqrt{2}}(1, 0, -1)$, $\frac{1}{\sqrt{2}}(1, 0, 1)$ und $(0, 1, 0)$, die eine Orthonormalbasis des \mathbb{R}^3 bilden. \square

Ein weiterer wichtiger Zusammenhang zwischen v , w , und $\text{proj}_w(v)$ ist die Tatsache, dass $\text{proj}_w(v)$ derjenige Punkt von w ist, der v am nächsten ist; auch dies ist aus einer Skizze wie in Abbildung 6.3 auf Seite 121 zu erkennen. Man kann diesen Zusammenhang verwenden, um mittels Projektion denjenigen Punkt eines Unterraums zu bestimmen, der einem gegebenen Punkt am nächsten liegt. Dazu muss man zuerst überlegen, wie man auf einen Unterraum projizieren kann.

Im einfachsten Fall ist $E = L(\{b_1, b_2\})$ eine Ebene, die von den beiden Vektoren b_1 und b_2 aufgespannt wird. Wir bezeichnen mit $\text{proj}_E(v)$ den Projektionspunkt von v auf E . Wie man aus Abbildung 6.5 (links) erkennen kann genügt es nicht, v auf b_1 und b_2 zu projizieren, und die Ergebnisse zu addieren: Das Resultat ist im Allgemeinen nicht der gewünschte Projektionspunkt. Erst wenn die beiden Vektoren b_1 und b_2 senkrecht aufeinander stehen (wie in Abbildung 6.5 rechts zu sehen), erhält man den Projektionspunkt durch Addition der Projektionen auf b_1 und b_2 .

Die Verallgemeinerung dieser Beobachtung auf beliebige Unterräume ist im nächsten Satz zusammengefasst.

Satz 6.6 Sei U ein Unterraum eines Vektorraums V , $v \in V$, sowie $\{b_1, \dots, b_n\}$ eine orthogonale Basis von U . Dann ist derjenige Punkt von U , der v am nächsten liegt, gegeben durch

$$\text{proj}_U(v) = \text{proj}_{b_1}(v) + \dots + \text{proj}_{b_n}(v).$$

Wenn die Basis $\{b_1, \dots, b_n\}$ von U im letzten Satz sogar eine ONB ist, vereinfacht sich die Projektion auf U zu $\text{proj}_U(v) = \sum_{k=1}^n b_k \cdot v \cdot b_k$. Aus diesen Formeln sieht man, dass die Projektion eines Punktes v ausserhalb eines Unterraums U auf U der Koordinatenberechnung von v bezüglich einer orthogonalen Basis von U entspricht (obwohl dieser Punkt gar nicht in U ist).

Beispiel 6.13 Sei E die Ebene, die von $b_1 = (-1, 0, 1)$ und $b_2 = (2, 1, 0)$ aufgespannt wird, sowie $v = (2, 1, -2)$ ein Punkt, der nicht auf der Ebene liegt. Da b_1 und b_2 nicht senkrecht stehen, liefert

$$p' = \text{proj}_{b_1}(v) + \text{proj}_{b_2}(v) = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ -2 \end{pmatrix}$$

noch nicht das richtige Ergebnis. Das erkennt man etwa daran, dass der Abstand $\|v - p'\| = 2$ noch nicht der kürzeste Abstand von v zu E ist. Wenn man nämlich b_1 und b_2 zuerst orthogonalisiert, erhält man $v_1 = b_1 = (-1, 0, 1)$ und $v_2 = b_2 - \text{proj}_{b_1}(b_2) = (1, 1, 1)$, und damit

$$v' = \text{proj}_E(v) = \text{proj}_{v_1}(v) + \text{proj}_{v_2}(v) = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 7 \\ 1 \\ -5 \end{pmatrix}.$$

Der Abstand von v zu v' ist $\|v - v'\| = \sqrt{\frac{2}{3}}$ und damit kürzer als der Abstand von v zu p' . \square

Kapitel 7

Folgen und Reihen

Mit diesem Kapitel werden wir einen Schwenk durchführen, weg von der algebraisch strukturierten Welt der Vektorräume, hin zu Funktionen natürlicher und reeller Zahlen. Wir werden bald sehen, dass man aus so einfachen Bausteinen wie Funktionen natürlicher Zahlen sehr viele Konzepte der Mathematik bis hin zu Differential- und Integralrechnung aufbauen kann.

7.1 Folgen

Wie wir bereits in Abschnitt 3.5 gesehen haben, sind Folgen nichts anderes als Funktionen mit \mathbb{N} oder $\mathbb{N}_{\leq n}$ als Definitionsbereich. Im ersten Fall spricht man von *unendlichen Folgen*, im zweiten Fall von *endlichen Folgen*. Wir werden ab jetzt stets die reellen Zahlen als Wertebereich verwenden; diese Folgen nennt man *reelle Folgen*. Wir werden für eine Funktion $a : \mathbb{N} \rightarrow \mathbb{R}$ meist a_n schreiben, um schon in der Notation herauszustreichen, dass es sich bei der Funktion a um eine Folge handelt.

Spezialformen von Folgen sind *arithmetische* und *geometrische* Folgen, die sich beide leicht rekursiv definieren lassen.

Beispiel 7.1 Eine Funktion $a : \mathbb{N} \rightarrow \mathbb{R}$ heißt *arithmetische Folge*, wenn ein Startwert a_1 sowie eine Zahl d existieren, mit denen sich die Folgenglieder schreiben lassen als

$$a_n := \begin{cases} a_{n-1} + d & \text{für } n > 1 \\ a_1 & \text{sonst.} \end{cases}$$

Bei einer arithmetischen Folge unterscheiden sich zwei Folgenglieder somit um einen konstanten additiven Wert; wenn wir die Rekursion auflösen, ergibt sich

$$a_n = a_1 + (n - 1)d.$$

Seien $a_1 = 2$ und $d = 3$ die Werte, die eindeutig eine arithmetische Folge definieren. Die ersten paar Folgenglieder sind

$$(2, 5, 8, 11, 14, \dots)$$

Man sieht sofort, dass sich die gleiche Zahlenfolge aus der Gleichung $a_n = 2 + 3(n - 1)$ ergibt. \square

Im Gegensatz zu arithmetischen Folgen, bei denen sich die Folgenglieder um einen konstanten *additiven Term* unterscheiden, ist dieser Unterschied bei geometrischen Folgen ein konstanter *multiplikativer Faktor*.

Beispiel 7.2 Eine Funktion $b : \mathbb{N} \rightarrow \mathbb{R}$ heißt *geometrische Folge*, wenn ein Startwert b_1 sowie eine Zahl q existieren, mit denen sich die Folgenglieder schreiben lassen als

$$b_n := \begin{cases} b_{n-1} \cdot q & \text{für } n > 1 \\ b_1 & \text{sonst.} \end{cases}$$

Im Fall der geometrischen Folgen liefert die Auflösung der Rekursion das Bildungsgesetz

$$b_n = b_1 q^{n-1}.$$

Gegeben sei die geometrische Folge, deren erste Glieder

$$\left(-8, 4, -2, 1, -\frac{1}{2}, \frac{1}{4}, \dots \right)$$

sind. Man kann erkennen, dass $b_1 = -8$ und $q = -\frac{1}{2}$ sind. Dieselbe Folge erhält man aus der Definition $b_n = -8 \left(-\frac{1}{2}\right)^{n-1}$. \square

In unseren weiteren Überlegungen werden wir hauptsächlich unendliche Folgen betrachten. Dabei ist das Verhalten dieser Folgen von Interesse, wenn der Index n der Folge gegen Unendlich geht.

Beispiel 7.3 Gegeben sei die Folge $a_n = \frac{1}{n}$. Man erkennt, dass sich die Folgenglieder für große Werte von n immer mehr 0 nähern.

Im Gegensatz dazu tendiert die Folge $b_n = 3n$ gegen unendlich, und die Folge $c_n = (-1)^n$ alterniert zwischen -1 und $+1$. \square

Wie man aus diesen Beispielen erkennen kann, zeigen unterschiedliche Folgen unterschiedliches Verhalten für große Indizes. Uns interessiert speziell die Situation, bei der Folgen gegen einen reellen Wert tendieren.

Definition 7.1 (Grenzwert einer Folge)

Sei $a : \mathbb{N} \rightarrow \mathbb{R}$ eine reelle Folge. Diese Folge heißt *konvergent*, wenn ein $\alpha \in \mathbb{R}$ existiert mit der Eigenschaft

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N \quad |a_n - \alpha| < \varepsilon.$$

Das $\alpha \in \mathbb{R}$ mit dieser Eigenschaft heißt *Grenzwert* $\lim_{n \rightarrow \infty} a_n$ der Folge a_n . Eine Folge mit Grenzwert 0 heißt *Nullfolge*. Eine Folge ohne Grenzwert heißt *divergent*.

Diese Definition bedeutet, dass für konvergente Folgen ab einem bestimmten Index N , der von gewähltem (beliebigem) $\varepsilon > 0$ abhängt, alle weiteren Folgenglieder (mit Index $n \geq N$) innerhalb eines " ε -Schlauchs" um α liegen. Dies ist in Abbildung 7.1 illustriert.

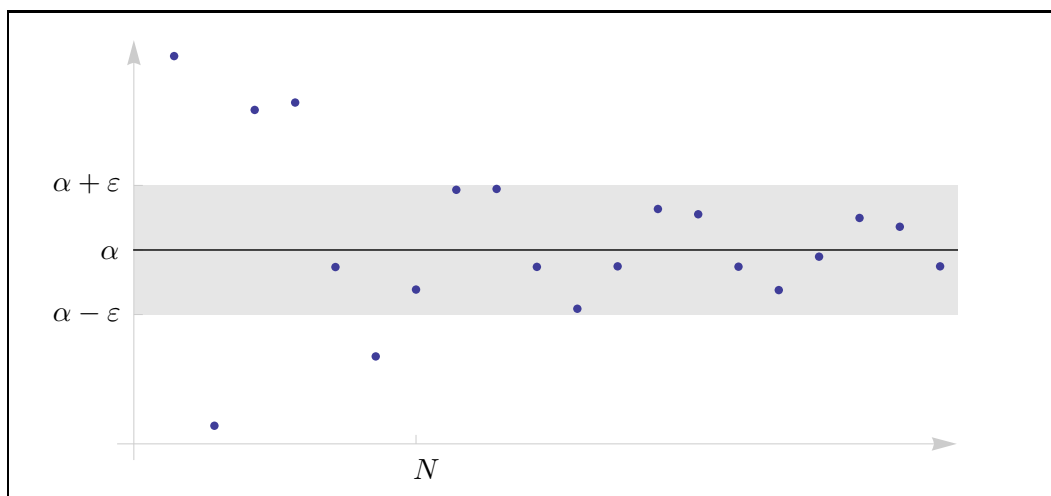


Abbildung 7.1: Illustration zum ε -Schlauch. Ab Index N befinden sich alle Folgenglieder näher als ε bei α .

Das Überprüfen der Konvergenz von Folgen erfordert einiges Umformen, und Erfahrung im Nachrechnen von Definitionen. Wir werden diese Umformungen anhand eines einfachen Beispiels durchführen.

Beispiel 7.4 Wir haben in Beispiel 7.3 behauptet, dass $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ sei. Um dies nachzuprüfen, setzen wir in Definition 7.1 ein. Zu zeigen ist somit

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N \quad \left| \frac{1}{n} - 0 \right| < \varepsilon.$$

Wie wir wissen, ist diese Quantorenschreibweise nur eine Abkürzung für Implikationen und Konjunktionen; man kann obigen Ausdruck also in Langform schreiben als

$$\forall \varepsilon \quad \varepsilon > 0 \Rightarrow \left(\exists N \quad N \in \mathbb{N} \wedge \left(\forall n \quad n \geq N \Rightarrow \left| \frac{1}{n} \right| < \varepsilon \right) \right).$$

Die ersten beiden Schritte im Überprüfen der Gültigkeit dieses Ausdrucks sind die Elimination des äußersten Allquantors, und der obersten Implikation. Sei dazu also ε beliebig aber fix; Auflösen der Implikation erweitert die Wissensbasis:

wir wissen $\varepsilon > 0$

zu zeigen $\exists N \quad N \in \mathbb{N} \wedge \left(\forall n \quad n \geq N \Rightarrow \left| \frac{1}{n} \right| < \varepsilon \right).$

An dieser Stelle benötigen wir eine Idee, was das geforderte $N \in \mathbb{N}$ sein könnte. Da dieses N natürlich auch von der Folge abhängt, diese aber bis jetzt noch nicht in die Rechnung eingegangen ist, werden wir so tun, als hätten wir das gewünschte N bereits gefunden. Wir werden erst später sehen, wie dieses N wirklich auszusehen hat.

wir wissen $\varepsilon > 0, N = (\text{füllen wir später ein})$

zu zeigen $N \in \mathbb{N} \wedge \left(\forall n \quad n \geq N \Rightarrow \left| \frac{1}{n} \right| < \varepsilon \right).$

Zu zeigen ist also eine Konjunktion von zwei Aussagen. Erstere wird durch die (zu erfolgende) richtige Wahl von N sowieso erfüllt sein, bleibt noch zweite:

$$\begin{array}{ll} \text{wir wissen} & \varepsilon > 0, N = (\text{füllen wir später ein}) \\ \text{zu zeigen} & \forall n \ n \geq N \Rightarrow \left| \frac{1}{n} \right| < \varepsilon. \end{array}$$

Hier gehen wir wie gehabt vor, mit Eliminieren des Allquantors und Auflösen der Implikation. Man erhält mit n beliebig aber fix:

$$\begin{array}{ll} \text{wir wissen} & \varepsilon > 0, N = (\text{füllen wir später ein}), n \geq N \\ \text{zu zeigen} & \left| \frac{1}{n} \right| < \varepsilon. \end{array}$$

An dieser Stelle kann man nichts mehr formal umformen; man muss sich die korrekte Wahl von N überlegen. Für ein besseres Verständnis der Problematik setzen wir konkrete Werte für $\varepsilon > 0$ ein, z.B. $\varepsilon = \frac{1}{10}$. Dann sieht man, dass ab $N = 11$ die weiteren Folgenglieder $\frac{1}{11}, \frac{1}{12}, \frac{1}{13}, \dots$ sind und somit die zu zeigende Aussage erfüllen.

Für beliebig aber fixes $\varepsilon > 0$ erhält man das allgemeine N also als $\frac{1}{\varepsilon} + 1$; da dies nicht für alle $\varepsilon > 0$ eine ganze Zahl ist, nehmen wir noch die nächstgrößere ganze Zahl durch Verwendung der *ceiling*-Funktion, setzen also

$$N := \left\lceil \frac{1}{\varepsilon} + 1 \right\rceil.$$

Damit können wir den Beweis abschließen; wir verwenden noch, dass $\frac{1}{n}$ immer positiv ist, wir im “zu zeigen” Teil also den Absolutbetrag weglassen können:

$$\begin{array}{ll} \text{wir wissen} & \varepsilon > 0, N = \left\lceil \frac{1}{\varepsilon} + 1 \right\rceil, n \geq N \\ \text{zu zeigen} & \frac{1}{n} < \varepsilon. \end{array}$$

Wir formen den “wir wissen” Teil um:

$$n \geq \left\lceil \frac{1}{\varepsilon} + 1 \right\rceil \Leftrightarrow n > \frac{1}{\varepsilon} \Leftrightarrow \frac{1}{n} < \varepsilon,$$

und erhalten die zu zeigende Aussage. \square

Das direkte Nachrechnen von Grenzwerten ist somit ziemlich aufwändig. Man kann aber Grenzwerte von elementaren Folgen sowie Eigenschaften von Grenzwerten verwenden, um viele Grenzwertberechnungen auf einfachere Formen zu reduzieren.

Zuvor sei noch darauf hingewiesen, dass es Spezialfälle von divergenten Folgen a gibt, für die wir (etwas schlampig) ebenfalls die Grenzwert-Notation $\lim_{n \rightarrow \infty} a_n$ verwenden. Für Folgen mit

$$\forall_{S>0} \exists_{N \in \mathbb{N}} \forall_{n \geq N} a_n > S,$$

schreiben wir $\lim_{n \rightarrow \infty} a_n = \infty$, für solche mit

$$\forall_{S>0} \exists_{N \in \mathbb{N}} \forall_{n \geq N} a_n < -S$$

schreiben wir $\lim_{n \rightarrow \infty} a_n = -\infty$.

Für das Rechnen mit Grenzwerten gelten mehrere einfache Rechenregeln, die im nächsten Satz zusammengefasst sind.

Satz 7.1 (Rechenregeln für Grenzwerte) Seien $a, b : \mathbb{N} \rightarrow \mathbb{R}$ zwei konvergente reelle Folgen mit Grenzwerten $\lim_{n \rightarrow \infty} a_n = \alpha$ und $\lim_{n \rightarrow \infty} b_n = \beta$. Dann gelten folgende Aussagen für Kompositionen der Folgen a und b :

- (1) $\lim_{n \rightarrow \infty} a_n + b_n = \alpha + \beta$
- (2) $\lim_{n \rightarrow \infty} a_n - b_n = \alpha - \beta$
- (3) $\lim_{n \rightarrow \infty} a_n \cdot b_n = \alpha \cdot \beta$
- (4) $\lim_{n \rightarrow \infty} a_n / b_n = \alpha / \beta$, falls $\beta \neq 0$.
- (5) $\lim_{n \rightarrow \infty} a_n^c = \alpha^c$ für $c \in \mathbb{R}$, falls $\alpha > 0$.

Diese Rechenregeln können alle über Nachrechnen der Definition 7.1 bewiesen werden; dies erfordert aber ähnliche (und noch kompliziertere) Umformungen wie die, die in Beispiel 7.4 verwendet wurden. Wir verzichten an dieser Stelle darauf.

Man beachte, dass diese Rechenregeln nur für *konvergente* Folgen gelten. Gemeinsam mit den Grenzwerten einfacher Folgen, die im nächsten Satz (ohne Beweis) angegeben sind, können Grenzwerte von einer großen Menge von Folgen berechnet werden.

Satz 7.2 (Grenzwerte elementarer Folgen) Es gelten die folgenden Grenzwerte:

- (1) $\lim_{n \rightarrow \infty} q^n = \begin{cases} 0 & \text{für } |q| < 1 \\ 1 & \text{für } q = 1 \\ \infty & \text{für } q > 1 \\ \text{divergent} & \text{für } q < -1 \end{cases}$
- (2) $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$
- (3) $\lim_{n \rightarrow \infty} \sqrt[n]{a} = 1$, falls $a > 0$
- (4) $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$.

Beispiel 7.5 Dieses Beispiel verwendet mehrere Rechenregeln aus den Sätzen 7.1 und 7.2 sowie die Tatsache, dass für $c \in \mathbb{R}$ $\lim_{n \rightarrow \infty} c = c$ gilt:

$$\lim_{n \rightarrow \infty} \frac{5}{3^n} + \sqrt[n]{4n} = \lim_{n \rightarrow \infty} \frac{5}{3^n} + \lim_{n \rightarrow \infty} \sqrt[n]{4n}$$

$$\begin{aligned}
&= \lim_{n \rightarrow \infty} 5 \cdot \lim_{n \rightarrow \infty} \left(\frac{1}{3}\right)^n + \lim_{n \rightarrow \infty} \sqrt[n]{4} \cdot \lim_{n \rightarrow \infty} \sqrt[n]{n} \\
&= 5 \cdot 0 + 1 \cdot 1 = 1. \quad \square
\end{aligned}$$

Beispiel 7.6 Mit Satz 7.1, Teil (4) und Satz 7.2, ebenfalls Teil (4) gilt:

$$\lim_{n \rightarrow \infty} \left(\frac{n-1}{n-3}\right)^n = \frac{n^n \left(1 - \frac{1}{n}\right)^n}{n^n \left(1 - \frac{3}{n}\right)^n} = \frac{e^{-1}}{e^{-3}} = e^2. \quad \square$$

Beispiel 7.7 Für das Berechnen der Grenzwerte von Brüchen von Polynomen (sogenannten *rationalen Funktionen*) kürzt man durch die höchste Potenz des Nenners und kann damit viele Terme ignorieren, die Nullfolgen sind:

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{n^4 - n^3 + 2n^2}{3n^5 - 7n^4 + n + 1} &= \lim_{n \rightarrow \infty} \frac{\frac{1}{n} - \frac{1}{n^2} + \frac{2}{n^3}}{3 - \frac{7}{n} + \frac{1}{n^4} + \frac{1}{n^5}} \\
&= \frac{\lim_{n \rightarrow \infty} \frac{1}{n} - \frac{1}{n^2} + \frac{2}{n^3}}{\lim_{n \rightarrow \infty} 3 - \frac{7}{n} + \frac{1}{n^4} + \frac{1}{n^5}} = \frac{0}{3} = 0.
\end{aligned}$$

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{n^5 + 4n^4 + 2n^3 + 3}{3n^5 + 2n^3 + n^2} &= \lim_{n \rightarrow \infty} \frac{1 + \frac{4}{n} + \frac{2}{n^2} + \frac{3}{n^5}}{3 + \frac{2}{n^2} + \frac{1}{n^3}} \\
&= \frac{\lim_{n \rightarrow \infty} 1 + \frac{4}{n} + \frac{2}{n^2} + \frac{3}{n^5}}{\lim_{n \rightarrow \infty} 3 + \frac{2}{n^2} + \frac{1}{n^3}} = \frac{1}{3}.
\end{aligned}$$

Hier erkennt man, dass der Nenner gegen eine reelle Zahl konvergiert, der Zähler aber gegen ∞ , der gesamte Bruch somit auch gegen ∞ :

$$\lim_{n \rightarrow \infty} \frac{3n^6 - 5n^5 + 7n - 2}{-2n^4 - 3n^2 + n} = \lim_{n \rightarrow \infty} \frac{3n^2 - 5n + \frac{7}{n^3} - \frac{2}{n^4}}{-2 - \frac{3}{n^2} + \frac{1}{n^4}} = \infty.$$

Aus diesen drei Fällen kann man somit eine einfache Regel für das Konvergenzverhalten rationaler Funktionen ablesen: Wenn der Grad des Polynoms im Zähler größer ist als der Grad des Polynoms im Nenner, geht die Folge gegen $\pm\infty$. Bei gleichen Graden konvergiert die Folge gegen den Quotienten der Leitkoeffizienten, und bei höherem Grad im Nenner konvergiert sie gegen 0. \square

7.2 Reihen

In diesem Abschnitt werden wir die Grenzwerte von speziellen Folgen betrachten, die als Summen von Folgen definiert sind. Dazu benötigen wir den Begriff der *Partialsumme*.

Definition 7.2 (Partialsomme)

Sei $a : \mathbb{N} \rightarrow \mathbb{R}$ eine reelle Folge. Dann bezeichnet man

$$s_n := \sum_{k=1}^n a_k = a_1 + a_2 + \cdots + a_n$$

als Folge der *Partialsommen* der Folge a_n .

Für spezielle Klassen von Folgen lassen sich Partialsommen direkt berechnen.

Beispiel 7.8 (Fortsetzung von Beispiel 7.1) Für eine arithmetische Folge $a_n = a_1 + (n-1)d$ erhält man mit den Rechenregeln für das Manipulieren von Summen die Umformung

$$\begin{aligned} \sum_{k=1}^n a_k &= \sum_{k=1}^n a_1 + (k-1)d = \sum_{k=1}^n a_1 + d \sum_{k=1}^n (k-1) \\ &= n \cdot a_1 + d \frac{n(n-1)}{2}, \end{aligned}$$

wobei die letzte Gleichheit die Summationsformel aus Beispiel 3.34 verwendet. Man kann auch ohne Formel durch Aufsummieren nachrechnen, dass für die arithmetische Folge mit $a_1 = 2$ und $d = 3$ die Folge der Partialsommen $(2, 7, 15, 26, 40, \dots)$ ist. \square

Beispiel 7.9 (Fortsetzung von Beispiel 7.2) Für Partialsommen, die aus geometrischen Folgen gebildet werden, benötigt man folgende Summationsformel für Potenzen, die sich mit Induktion beweisen lässt:

$$\sum_{k=1}^n q^k = q \cdot \frac{1-q^n}{1-q} \quad \text{oder äquivalent} \quad \sum_{k=0}^n q^k = \frac{1-q^{n+1}}{1-q}.$$

Damit gilt für Partialsommen geometrischer Folgen:

$$\sum_{k=1}^n b_1 q^{k-1} = b_1 \sum_{k=0}^{n-1} q^k = b_1 \frac{1-q^n}{1-q}.$$

Für $b_1 = -8$ und $q = -\frac{1}{2}$ erhält man somit als Folge der Partialsommen die Zahlen $(-8, -4, -6, -5, -\frac{11}{2}, -\frac{21}{4}, \dots)$. \square

Wenn man die letzten beiden Beispiele vergleicht, so erkennt man, dass die Partialsommenfolge der arithmetischen Folge gegen ∞ geht, und die der geometrischen Folge gegen einen Wert um -5 zu konvergieren scheint. Es wird im Folgenden von speziellem Interesse sein, Grenzwerte von Partialsommenfolgen zu bestimmen, da dadurch etwa neue interessante Funktionen definiert werden können.

Definition 7.3 (Reihe)

Für gegebene reelle Folge a_n nennt man $\sum_{n=1}^{\infty} a_n$ eine (unendliche) Reihe. Wenn die Partialsommenfolge $\sum_{k=1}^n a_k$ von a_n konvergiert, so heißt die Reihe $\sum_{n=1}^{\infty} a_n$ *konvergent*, und es gilt

$$\sum_{n=1}^{\infty} a_n = \lim_{n \rightarrow \infty} \sum_{k=1}^n a_k;$$

andernfalls nennt man die Reihe *divergent*.

Beispiel 7.10 (Fortsetzung von Beispiel 7.9) Für die geometrische Folge mit $b_1 = -8$ und $q = -\frac{1}{2}$ ist $b_n = -8 \cdot (-\frac{1}{2})^{n-1}$; man erhält die Reihe $-8 \cdot \sum_{n=1}^{\infty} (-\frac{1}{2})^{n-1}$. Wegen der numerischen Werte der Partialsummenfolge aus Beispiel 7.9 vermuten wir, dass diese Reihe konvergiert. Da $\sum_{k=0}^n q^k = \frac{1-q^{n+1}}{1-q}$ ist, gilt

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

Der Term auf der rechten Seite dieser Gleichung konvergiert, da $|-\frac{1}{2}| < 1$ ist; mit Teil 1 von Satz 7.2 gilt

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \left(-\frac{1}{2}\right)^{k-1} = \lim_{n \rightarrow \infty} \frac{1 - (-\frac{1}{2})^n}{1 - (-\frac{1}{2})} = \frac{1}{\frac{3}{2}} = \frac{2}{3}.$$

Somit ist die ursprüngliche Reihe konvergent mit

$$-8 \cdot \sum_{n=1}^{\infty} \left(-\frac{1}{2}\right)^{k-1} = -8 \cdot \lim_{n \rightarrow \infty} \sum_{k=1}^n \left(-\frac{1}{2}\right)^{k-1} = -8 \cdot \frac{2}{3} = -\frac{16}{3}. \quad \square$$

Reihen der Gestalt $\sum_{n=1}^{\infty} q^n$, wie sie im letzten Beispiel vorgekommen sind, werden *geometrische Reihen* genannt.

Satz 7.3 Geometrische Reihen $\sum_{n=1}^{\infty} q^n$ sind für $|q| < 1$ konvergent mit

$$\sum_{n=1}^{\infty} q^n = q \cdot \frac{1}{1-q} \quad \text{oder äquivalent} \quad \sum_{n=0}^{\infty} q^n = \frac{1}{1-q}.$$

Bei allgemeinen Reihen kann man nicht so leicht erkennen, ob sie konvergent oder divergent sind. Es ist aber unmittelbar einsichtig, dass nur Reihen von Nullfolgen überhaupt konvergent sein können: Wenn a_n keine Nullfolge ist, wird die Folge der Partialsummen immer größer und kann somit sicher nicht gegen einen endlichen Wert konvergieren.

Das Konvergenzverhalten vieler Reihen kann mit Hilfe der folgenden Konvergenzkriterien überprüft werden. Diese Kriterien liefern nur Aussagen über die Konvergenz bzw. Divergenz einer Reihe, können aber nicht verwendet werden, um den Grenzwert einer Reihe zu berechnen.

Satz 7.4 (Konvergenzkriterien für Reihen)

Majorantenkriterium Seien $a, b : \mathbb{N} \rightarrow \mathbb{R}^+$ zwei positive reelle Folgen, mit $b_n \leq a_n$ ab einem Index $N \in \mathbb{N}$. Wenn $\sum_{n=1}^{\infty} a_n$ konvergiert, dann konvergiert auch $\sum_{n=1}^{\infty} b_n$.

Minorantenkriterium Seien $a, b : \mathbb{N} \rightarrow \mathbb{R}^+$ zwei positive reelle Folgen, mit $b_n \leq a_n$ ab einem Index $N \in \mathbb{N}$. Wenn $\sum_{n=1}^{\infty} b_n$ divergiert, dann divergiert auch $\sum_{n=1}^{\infty} a_n$.

Quotientenkriterium Sei $a : \mathbb{N} \rightarrow \mathbb{R}$ eine reelle Folge, und

$$q = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|.$$

Für $q < 1$ ist $\sum_{n=1}^{\infty} a_n$ konvergent; für $q > 1$ divergent; für $q = 1$ ist keine Aussage möglich.

Wurzelkriterium Sei $a : \mathbb{N} \rightarrow \mathbb{R}$ eine reelle Folge, und

$$q = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}.$$

Für $q < 1$ ist $\sum_{n=1}^{\infty} a_n$ konvergent; für $q > 1$ divergent; für $q = 1$ ist keine Aussage möglich.

Während die Gültigkeit der Majoranten- und Minorantenkriterien plausibel erscheint, ist dies bei Quotienten- und Wurzelkriterium nicht der Fall. Wir werden an dieser Stelle auf einen Beweis dieser Kriterien verzichten und einige Beispiele betrachten, die die Anwendung dieses Satzes demonstrieren.

Beispiel 7.11 Wir betrachten das Konvergenzverhalten der Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots$$

Für diese Reihe konstruieren wir eine Majorante, indem wir die einzelnen Terme der Folge $b_n = \frac{1}{n^2}$ mit denen einer anderen Folge a_n vergleichen:

$$\begin{array}{cccccccccc} b_n = & 1, & \frac{1}{4}, & \frac{1}{9}, & \frac{1}{16}, & \frac{1}{25}, & \frac{1}{36}, & \frac{1}{49}, & \frac{1}{64}, & \frac{1}{81}, & \dots \\ a_n = & 1, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{16}, & \frac{1}{16}, & \frac{1}{16}, & \frac{1}{16}, & \frac{1}{64}, & \frac{1}{64}, & \dots \end{array}$$

Die Reihe $\sum_{n=1}^{\infty} a_n$ kann man aber folgendermaßen evaluieren:

$$\begin{aligned} \sum_{n=1}^{\infty} a_n &= 1 + \underbrace{\frac{1}{4} + \frac{1}{4}}_{\frac{1}{2}} + \underbrace{\frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16}}_{\frac{1}{4}} \\ &\quad + \underbrace{\frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64} + \frac{1}{64}}_{\frac{1}{8}} + \dots \\ &= \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{1 - \frac{1}{2}} = 2. \end{aligned}$$

Da a_n eine Majorante von $\frac{1}{n^2}$ ist, und $\sum_{n=1}^{\infty} a_n$ konvergiert, konvergiert auch $\sum_{n=1}^{\infty} \frac{1}{n^2}$. \square

Beispiel 7.12 Die *harmonische Reihe* $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert, wenn auch sehr langsam: Man erhält die Partialsummen

$$\sum_{n=1}^{100} \frac{1}{n} = 5.187, \quad \sum_{n=1}^{10\,000} \frac{1}{n} = 9.788, \quad \sum_{n=1}^{1\,000\,000} \frac{1}{n} = 14.393.$$

Zum Beweis der Divergenz der harmonischen Reihe kann man eine Minorante finden, deren Reihe divergent ist. Wir gehen dazu ähnlich wie im letzten Beispiel vor. Für $a_n = \frac{1}{n}$ sei b_n gegeben durch

$$\begin{aligned} a_n &= 1, & \frac{1}{2}, & \frac{1}{3}, & \frac{1}{4}, & \frac{1}{5}, & \frac{1}{6}, & \frac{1}{7}, & \frac{1}{8}, & \frac{1}{9}, & \dots \\ b_n &= 1, & \frac{1}{2}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{8}, & \frac{1}{8}, & \frac{1}{8}, & \frac{1}{8}, & \frac{1}{16}, & \dots \end{aligned}$$

Die Partialsummen von b_n sind

$$\sum_{n=1}^{\infty} b_n = 1 + \frac{1}{2} + \underbrace{\frac{1}{4} + \frac{1}{4}}_{\frac{1}{2}} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}}_{\frac{1}{2}} + \dots,$$

somit ist $\sum_{n=1}^{\infty} b_n$ divergent und mit dem Minorantenkriterium auch die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$. \square

Die Ergebnisse der beiden letzten Beispiele können folgendermaßen verallgemeinert werden.

Satz 7.5 Sei $\alpha \in \mathbb{R}$. Die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$$

ist für $\alpha = 1$ divergent, für $\alpha > 1$ aber konvergent.

Wir betrachten noch Beispiele zu Quotienten- und Wurzelkriterium.

Beispiel 7.13 Gegeben sei die Reihe

$$\sum_{n=1}^{\infty} \frac{2^n}{n!}.$$

Wir können die Konvergenz dieser Reihe mit dem Quotientenkriterium untersuchen.

Es ist

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \frac{2^{n+1}}{(n+1)!} \cdot \frac{n!}{2^n} = \lim_{n \rightarrow \infty} \frac{2}{(n+1)} = 0.$$

Somit ist $\sum_{n=1}^{\infty} \frac{2^n}{n!}$ konvergent. \square

Beispiel 7.14 Zur Untersuchung der Konvergenz der Reihe

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

kann man das Wurzelkriterium verwenden. Es gilt

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} \sqrt[n]{\left| \frac{2n+1}{n+2} \right|^n} = \lim_{n \rightarrow \infty} \frac{2n+1}{n+2} = 2.$$

Die gegebene Reihe ist also divergent. \square

7.3 Potenzreihen

Mit den Überlegungen im letzten Abschnitt ist es uns möglich, Reihen auf Konvergenz hin zu untersuchen. In diesem Abschnitt werden wir sehen, dass man über konvergente Reihen Funktionen definieren kann. Dazu benötigen wir den Begriff der Potenzreihe.

Definition 7.4 (Potenzreihe)

Für gegebene Folge $a : \mathbb{N}_0 \rightarrow \mathbb{R}$ nennt man den Ausdruck $\sum_{n=0}^{\infty} a_n x^n$ eine *Potenzreihe*. Für $x \in \mathbb{R}$ ist durch die Abbildung

$$x \mapsto \sum_{n=0}^{\infty} a_n x^n$$

eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ für diejenigen Werte x definiert, für die Potenzreihe konvergiert.

Potenzreihen beginnen traditionell mit Summationsindex 0; wir werden diese Konvention hier übernehmen. Wie aus der Definition ersichtlich, ist die Frage der Konvergenz von Potenzreihen von zentraler Bedeutung, da nur konvergente Potenzreihen eine Funktion definieren. Da die Konvergenz vom Wert von x abhängt, stellt sich unmittelbar die Frage, für welche Werte von x die Potenzreihe konvergiert. Die Konvergenzkriterien aus Satz 7.4 werden uns helfen, diese Frage zu beantworten. Für Spezialfälle kann man sich auch mit anderen Resultaten behelfen, wie das nächste Beispiel zeigen wird.

Beispiel 7.15 Für die konstante Folge $a_n = 1$ konvergiert die Potenzreihe $\sum_{n=0}^{\infty} x^n$ (und zwar gegen $\frac{1}{1-x}$) genau dann, wenn $|x| < 1$ gilt. Dies folgt unmittelbar aus Satz 7.3. \square

Zur Untersuchung des Konvergenzverhaltens von Potenzreihen sind speziell das Quotienten- und das Wurzelkriterium von Interesse. Wir betrachten zuerst das Quotientenkriterium; angewandt auf Potenzreihen besagt dies, dass $\sum_{n=0}^{\infty} a_n x^n$ konvergent ist, wenn

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| < 1 \quad \Leftrightarrow \quad \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| |x| < 1$$

$$\Leftrightarrow |x| < \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right|.$$

Eine ähnliche Überlegung kann für das Wurzelkriterium durchgeführt werden:

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt[n]{|a_n x^n|} < 1 &\Leftrightarrow \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} |x| < 1 \\ &\Leftrightarrow |x| < \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{|a_n|}}. \end{aligned}$$

Damit kann der Wertebereich für x , innerhalb dessen die Potenzreihe konvergiert, durch Grenzwerte berechnet werden. Man definiert dafür den Begriff des Konvergenzradius.

Definition 7.5 (Konvergenzradius)

Für die Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ nennt man die Zahl

$$r := \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{|a_n|}}$$

den *Konvergenzradius* der Potenzreihe.

Der Name ‘‘Konvergenzradius’’ bezieht sich darauf, dass Potenzreihen auch für komplexe Zahlen definiert sind; in diesem Fall ist der Konvergenzradius wirklich ein Radius. Für $x \in \mathbb{R}$ erhält man durch den Konvergenzradius die Größe des Intervalls, innerhalb dessen die Potenzreihe konvergiert. Für Werte von x außerhalb des Intervalls divergiert die Potenzreihe. Für die Intervallgrenzen selbst, also für $x = \pm r$, muss das Konvergenzverhalten eigens untersucht werden.

Die beiden in Definition 7.5 angeführten Möglichkeiten zur Berechnung der Konvergenzradius liefern das gleiche Ergebnis; allerdings wird, je nach Potenzreihe, eine der beiden Berechnungen einfacher durchzuführen sein als die andere.

Beispiel 7.16 Zu bestimmen sei der Konvergenzradius der Potenzreihe

$$\sum_{n=0}^{\infty} \frac{3^n}{n!} x^n.$$

Man erhält für den Konvergenzradius den Wert

$$\lim_{n \rightarrow \infty} \left| \frac{3^n}{n!} \frac{(n+1)!}{3^{n+1}} \right| = \lim_{n \rightarrow \infty} \left| \frac{(n+1)}{3} \right| = \infty.$$

Diese Potenzreihe konvergiert somit für alle Werte von x . □

Beispiel 7.17 Für die Potenzreihe

$$\sum_{n=0}^{\infty} \frac{n}{2^{n+1}} x^n.$$

ergibt sich folgender Konvergenzradius:

$$\lim_{n \rightarrow \infty} \left| \frac{n}{2^{n+1}} \frac{2^{n+2}}{(n+1)} \right| = \lim_{n \rightarrow \infty} \left| \frac{2n}{n+1} \right| = 2. \quad \square$$

Beispiel 7.18 Für den Konvergenzradius der Potenzreihe

$$\sum_{n=0}^{\infty} \frac{n^n}{2^{n+1}} x^n.$$

kann man das Wurzelkriterium anwenden:

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{\left| \frac{n^n}{2^{n+1}} \right|}} = \lim_{n \rightarrow \infty} \sqrt[n]{\left| \frac{2^n \cdot 2}{n^n} \right|} = \lim_{n \rightarrow \infty} \frac{2}{n} \sqrt[n]{2} = 0.$$

Damit konvergiert diese Potenzreihe nur für $x = 0$. □

Die wichtigsten reellen Funktionen können über Potenzreihen definiert werden. Der Grund, warum diese Potenzreihen genau so definiert sind, wird erst in Abschnitt 8.3 klar werden.

Definition 7.6 (Elementare Funktionen)

Bekannte reelle Funktionen sind über Potenzreihen folgendermaßen definiert:

- $e^x := 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \dots = \sum_{n=0}^{\infty} \frac{x^n}{n!}$
- $\sin(x) := x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$
- $\cos(x) := 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$
- $\log(x+1) := x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} \dots = \sum_{n=1}^{\infty} (-1)^{(n+1)} \frac{x^n}{n}$

Man kann unschwer nachrechnen, dass die Exponential-, Sinus- sowie Kosinusfunktionen für alle $x \in \mathbb{R}$ konvergieren, der Konvergenzradius der Logarithmusfunktion (hier ist der natürliche Logarithmus definiert) aber nur 1 ist.

Durch die Möglichkeit, reelle Funktionen über Potenzreihen zu definieren, kann der Definitionsbereich dieser Funktionen ohne weitere Überlegungen auf die komplexen Zahlen ausgeweitet werden. Die Konvergenzradien verändern sich für komplexe Funktionsargumente nicht. Ein Beispiel für den Sinn der Erweiterung des Definitionsbereichs von Funktionen liefert folgender Satz.

Satz 7.6 (Euler'sche Formel) Für $x \in \mathbb{R}$ gilt

$$e^{ix} = \cos(x) + i \sin(x).$$

Diese Formel folgt durch direktes Einsetzen der Definitionen 7.6 in die Formel. Aus der Euler'schen Formel ergibt sich auch eine Gleichung, die die wichtigsten Konstanten der Mathematik vereint:

$$e^{i\pi} + 1 = 0.$$

Weitere Beispiele von Potenzreihen, sowie Anwendungen, werden wir in Abschnitt 8.3 kennenlernen.

7.4 Stetige Funktionen

Im Folgenden werden wir Grenzwerte von reellen Funktionen betrachten; diese Überlegungen sind für die Einführung der Differentialrechnung in Kapitel 8 wichtig. Wir werden erkennen, dass durch das Zurückführen der Grenzwerte von *Funktionen* auf Grenzwerte von *Folgen* die Regeln für das Rechnen mit Grenzwerten von Folgen auch auf den Fall von reellen Funktionen übertragen werden können.

In diesem Abschnitt werden wir das Verhalten von reellen Funktionen f untersuchen, wenn sich Werte im Definitionsbereich von f immer näher einem bestimmten Punkt x_0 nähern. Interessant wird die Situation erst dann, wenn f in x_0 nicht definiert ist. Wir wissen auch bereits, wie man ein "immer näher an x_0 " formalisieren kann: nämlich über eine Folge, die gegen x_0 konvergiert. Dementsprechend definiert man den Begriff des Grenzwertes von Funktionen.

Definition 7.7 (Grenzwert von Funktionen)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion mit Definitionsbereich D , und $x_0 \in \mathbb{R}$. Sei $a : \mathbb{N} \rightarrow D$ eine beliebige reelle Folge, die gegen x_0 konvergiert. Wenn auch die Folge $f(a_n)$ konvergiert, so nennt man den Grenzwert dieser Folge den Grenzwert $\lim_{x \rightarrow x_0} f(x)$ von f für $x \rightarrow x_0$, man definiert also

$$\lim_{x \rightarrow x_0} f(x) := \lim_{n \rightarrow \infty} f(a_n).$$

Man beachte, dass es mit dieser Definition nicht ausreicht, dass der Grenzwert von $f(a_n)$ für eine bestimmte (konkrete) Folge a_n konvergiert; diese Konvergenz muss vielmehr *für alle* Folgen gegeben sein, die gegen x_0 konvergieren. Beim Nachrechnen von Beispielen werden wir uns aber dennoch damit begnügen, das Konvergenzverhalten von $f(a_n)$ für *konkrete* Folgen a_n zu untersuchen, die gegen x_0 konvergieren.

Wie bei reellen Folgen erlauben wir auch die Grenzwert-Notation für Funktionen, die gegen $\pm\infty$ streben, schreiben also

$$\lim_{x \rightarrow x_0} f(x) = \pm\infty \quad \text{für} \quad \lim_{n \rightarrow \infty} f(a_n) = \pm\infty.$$

Ebenso kann man für x_0 auch die Werte $\pm\infty$ zulassen; in diesem Fall betrachtet man Folgen a_n mit $\lim_{n \rightarrow \infty} a_n = \pm\infty$.

Wir werden ebenso sehen, dass man in manchen Situationen beim Berechnen von Grenzwerten mit Hilfe von Definition 7.7 zwischen Folgen a_n unterschieden werden muss, die sich x_0 *von links* bzw. *von rechts* annähern. Den Grenzwert von

Folgen, die sich x_0 von links nähern, nennt man *linksseitigen Grenzwert*, und schreibt dafür $\lim_{x \rightarrow x_0^-} f(x)$. Analog dazu schreibt man für den *rechtsseitigen Grenzwert* $\lim_{x \rightarrow x_0^+} f(x)$. Der Grenzwert von f in x_0 existiert dann klarerweise nur, wenn links- und rechtsseitiger Grenzwert identisch sind (da der Grenzwert ja für jede beliebige Folge gleich sein muss).

All diese unterschiedlichen Situationen werden in den nächsten Beispielen untersucht und erläutert werden.

Beispiel 7.19 Für $f(x) = x^2$ sei $\lim_{x \rightarrow 2} f(x)$ zu berechnen. Eine Folge a_n mit $\lim_{n \rightarrow \infty} a_n = 2$ ist etwa $a_n = 2 + \frac{1}{n}$. Mit dieser Folge gilt

$$\lim_{x \rightarrow 2} f(x) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} \left(2 + \frac{1}{n}\right)^2 = \lim_{n \rightarrow \infty} 4 + \frac{2}{n} + \frac{1}{n^2} = 4.$$

Das Resultat ist identisch, wenn man etwa $a_n = 2 - \frac{1}{n}$ verwendet. □

Beispiel 7.20 Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ definiert als

$$f(x) = \begin{cases} 2x + 3 & \text{falls } x \neq 1 \\ 2 & \text{falls } x = 1, \end{cases}$$

gesucht sei $\lim_{x \rightarrow 1} f(x)$. Sei $a_n = 1 + \frac{1}{n^2}$ eine Folge, die gegen 1 konvergiert. Dann ist

$$\lim_{x \rightarrow 1} f(x) = \lim_{n \rightarrow \infty} f(a_n) \stackrel{(*)}{=} \lim_{n \rightarrow \infty} 2\left(1 + \frac{1}{n^2}\right) + 3 = \lim_{n \rightarrow \infty} 5 + \frac{2}{n^2} = 5.$$

Hier gilt die Umformung (*), da $1 + \frac{1}{n^2}$ für alle $n \in \mathbb{N}$ ungleich 1 ist. Das gleiche Ergebnis erhält man, wenn man sich $x_0 = 1$ von links nähert. In diesem Fall ist also der Grenzwert der Funktion in x_0 nicht identisch mit $f(x_0)$. □

Beispiel 7.21 Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ definiert als

$$f(x) = \begin{cases} x + 1 & \text{falls } x \geq 3 \\ -x - 2 & \text{falls } x < 3, \end{cases}$$

zu berechnen sei $\lim_{x \rightarrow 3} f(x)$. Da die Werte von f unterschiedlich sind, je nachdem, ob man sich von links oder rechts dem Wert $x_0 = 3$ nähert, berechnen wir hier sowohl den links-, als auch den rechtsseitigen Grenzwert. Für den linksseitigen Grenzwert verwenden wir $a_n = 3 - \frac{1}{n}$. Damit gilt

$$\lim_{x \rightarrow 3^-} f(x) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} -\left(3 - \frac{1}{n}\right) - 2 = -5,$$

da $3 - \frac{1}{n} < 3$ ist. Für den rechtsseitigen Grenzwert, mit $a_n = 3 + \frac{1}{n}$, erhalten wir aber

$$\lim_{x \rightarrow 3^+} f(x) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} \left(3 + \frac{1}{n}\right) + 1 = 4.$$

Sind links- und rechtsseitiger Grenzwert sind also unterschiedlich; damit existiert $\lim_{x \rightarrow 3} f(x)$ nicht. □

Beispiel 7.22 Für $f(x) = \frac{1}{x}$ ist $\lim_{x \rightarrow 0} f(x)$ gesucht. Man beachte, dass $f(x)$ in $x_0 = 0$ gar nicht definiert ist; dies ist für die Grenzwertberechnung aber nicht notwendig. Für $a_n = \frac{1}{n}$ erhält man

$$\lim_{x \rightarrow 0^+} f(x) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} \frac{1}{\frac{1}{n}} = \infty;$$

für $a_n = -\frac{1}{n}$ aber

$$\lim_{x \rightarrow 0^-} f(x) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} -\frac{1}{\frac{1}{n}} = -\infty.$$

Der Grenzwert $\lim_{x \rightarrow 0} f(x)$ existiert also nicht. \square

Beispiel 7.23 Wir betrachten $f(x) = \frac{1}{x^2}$ und berechnen wiederum $\lim_{x \rightarrow 0} f(x)$. Mit $a_n = \frac{1}{n}$ erhalten wir

$$\lim_{x \rightarrow 0^+} f(x) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} \frac{1}{\left(\frac{1}{n}\right)^2} = \infty,$$

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

$$\lim_{x \rightarrow 0^-} f(x) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} \frac{1}{\left(-\frac{1}{n}\right)^2} = \infty.$$

Bei diesem Beispiel sind links- und rechtsseitiger Grenzwert identisch, und es gilt $\lim_{x \rightarrow 0} f(x) = \infty$. \square

Eine wichtige Klasse von Funktionen werden durch diejenigen Funktionen gebildet, bei denen der Grenzwert an einer Stelle gleich dem Funktionswert an dieser Stelle ist.

Definition 7.8 (Stetigkeit)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion mit Definitionsbereich D , und $x_0 \in D$. Dann heißt f in x_0 *stetig*, wenn gilt

$$\lim_{x \rightarrow x_0} f(x) = f(x_0),$$

sonst heißt f in x_0 *unstetig*. Wenn f in ganz D stetig ist, dann nennt man f eine *stetige Funktion*.

Einsetzen in die Definition des Grenzwerts von Funktionen liefert für stetige Funktionen und Folgen a_n mit $\lim_{n \rightarrow \infty} a_n = x_0$ die Aussage

$$\lim_{x \rightarrow x_0} f(x) = \lim_{n \rightarrow \infty} f(a_n) = f(x_0) = f\left(\lim_{n \rightarrow \infty} a_n\right).$$

Bei stetigen Funktionen kann man also Funktionsanwendung und Grenzwertbildung vertauschen.

Wenn wir nochmals die letzten Beispiele untersuchen, so erkennt man:

- Die Funktion in Beispiel 7.19 ist in $x_0 = 2$ stetig (wie auch an allen anderen Stellen).

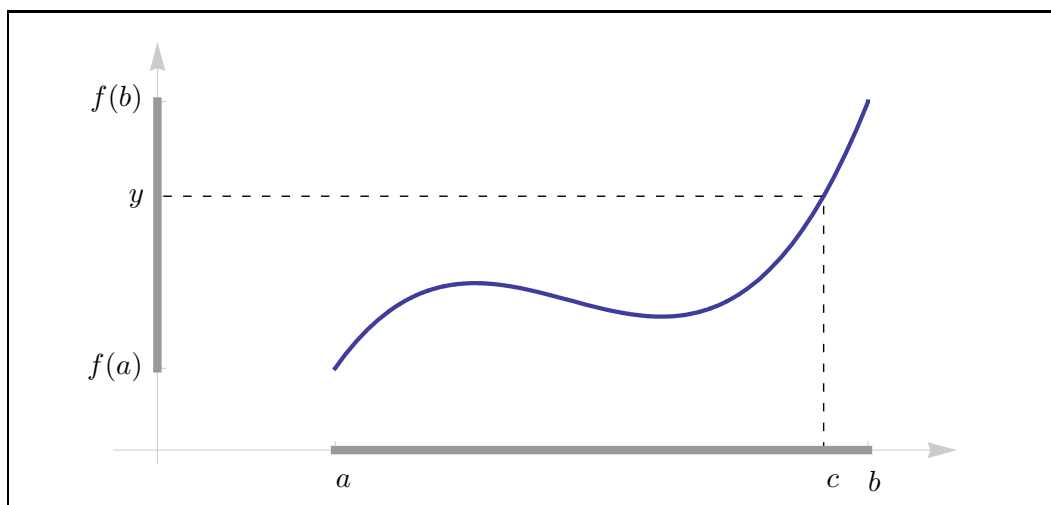


Abbildung 7.2: Illustration zum Zwischenwertsatz.

- Die Funktion in Beispiel 7.20 ist in $x_0 = 1$ unstetig, da sich Grenzwert und Funktionswert an dieser Stelle unterscheiden. Die Funktion ist aber an allen anderen Stellen stetig.
- Die Funktion in Beispiel 7.21 ist in $x_0 = 3$ unstetig, da der Grenzwert an dieser Stelle nicht existiert. An allen anderen Stellen ist diese Funktion aber stetig.
- Für die Funktionen aus Beispielen 7.22 und 7.23 stellt sich die Frage nach Stetigkeit in $x_0 = 0$ nicht, da Funktionen nur in ihrem Definitionsbereich stetig sein können. Dies ist bei beiden Funktionen der Fall, somit sind beides stetige Funktionen.

Stetige Funktionen sind somit genau diejenigen Funktionen, deren Graph in ihren Definitionsbereichen ohne Absetzen gezeichnet werden kann. Formal wird dieser Sachverhalt durch den nächsten Satz ausgedrückt.

Satz 7.7 (Zwischenwertsatz) Sei f eine Funktion, die auf einem Intervall $[a, b]$ stetig ist. Dann gibt es zu jedem Wert y zwischen $f(a)$ und $f(b)$ mindestens ein c zwischen a und b mit $f(c) = y$.

Eine Illustration des Zwischenwertsatzes ist in Abbildung 7.2 zu sehen.

Wie wir in den Beispielen dieses Abschnitts gesehen haben ist es etwas aufwändig, die Stetigkeit von Funktionen direkt über die Definition nachzuprüfen. Hier helfen die beiden folgenden Resultate.

Satz 7.8 Seien $f, g : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ stetige Funktionen. Dann sind auch, soweit definiert,

$$f \cdot g, \quad \frac{f}{g}, \quad f + g, \quad f - g, \quad f \circ g$$

stetige Funktionen.

Beispiel 7.24 Wie man unschwer nachprüfen kann, sind $f(x) = 3x + 1$ und $g(x) = x^2 - x$ auf ganz \mathbb{R} stetige Funktionen. Damit sind auch

$$(3x + 1) \cdot (x^2 - x), \quad (3x + 1) + (x^2 - x), \quad (3x + 1) - (x^2 - x), \quad 3(x^2 - x) + 1$$

auf ganz \mathbb{R} stetige Funktionen. Die Funktion $\frac{3x+1}{x^2-x}$ ist in ihrem Definitionsbereich $\mathbb{R} \setminus \{0, 1\}$ ebenfalls überall stetig. \square

Somit sind beliebige *endliche* Kombinationen von stetigen Funktionen wieder stetig. Es ist nicht unmittelbar klar, dass dies auch bei *unendlichen* Kombinationen der Fall ist, wie sie etwa bei der Definition von Potenzreihen in Satz 7.6 auftreten. Unter speziellen Bedingungen an Folgen von Funktionen, die die Partialsummen der Potenzreihen erfüllen, sind auch solche unendlichen Kombinationen stetig. Dies ist im nächsten Satz zusammengefasst.

Satz 7.9 Polynomfunktionen, rationale Funktionen, Wurzelfunktionen, sowie die Exponential-, Logarithmus-, Sinus- und Kosinusfunktion sind stetige Funktionen.

Damit sind eigentlich alle Funktionen, die wir als Terme ohne Fallunterscheidungen anschreiben können, stetige Funktionen.

Beispiel 7.25 Mit obigen Sätzen sind folgende Funktionen auf ganz \mathbb{R} stetig:

$$4x^2 - \sqrt{3}x + \pi \sin(x^2), \quad \cos(x) \sin(x + e^{2x}), \quad -5x/(x^2 + 1).$$

Folgende Funktionen sind ebenfalls stetig, aber nicht auf ganz \mathbb{R} definiert:

$$\log(x^2) : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, \quad \frac{e^{-x}}{x^2 - 1} : \mathbb{R} \setminus \{-1, 1\} \rightarrow \mathbb{R}. \quad \square$$

Kapitel 8

Differentialrechnung

In diesem Kapitel beschäftigen wir uns mit Ableitungen von Funktionen sowie deren Anwendungen. Wir beginnen mit einer Einläuterung des Begriffs der Ableitung.

8.1 Ableitung einer Funktion

In vielen technischen und naturwissenschaftlichen Anwendungen ist es wichtig, Veränderungen einer Kenngröße in Abhängigkeit einer anderen Größe (etwa der Zeit) zu untersuchen.

Beispiel 8.1 Ein Autofahrer benötigt für die Strecke Linz-München (270km) 2 Stunden und 15 Minuten. Die Durchschnittsgeschwindigkeit betrug somit $270/2.25 = 120\text{km/h}$.

Durch kleinere Beobachtungszeiträume kann man immer genauere Schätzungen für die Geschwindigkeit zu einem bestimmten Zeitpunkt erhalten. Für eine 5km lange gerade Strecke kurz vor München benötigte er nur 2 Minuten. Auf dieser Strecke betrug seine Durchschnittsgeschwindigkeit also $5/(2/60) = 5 * 60/2 = 150\text{km/h}$. Auf dieser Strecke legte er innerhalb von 10 Sekunden eine Strecke von einem halben Kilometer zurück und erreichte so eine Geschwindigkeit von $0.5/(10/3600) = 180\text{km/h}$. \square

Beispiel 8.2 Ein Anlagefonds, der am 1.7.2005 um 10 000€ gekauft wurde, ist am 31.12.2009 nur mehr 7 800€ Wert. Der durchschnittliche Verlust pro Jahr ist $2200\text{€}/4.5 = 488.9\text{€}$, oder knapp 5% des ursprünglichen Werts.

Wenn der Fonds am 23.10.2008 9 430€ Wert war, und 24 Stunden später nur mehr 9 420€, dann beträgt der aktuelle jährliche Verlust $10/(1/365) = 3650\text{€}$. \square

Das allgemeine Konzept hinter diesen Beispielen ist das des Differenzenquotienten, mit dem der Begriff der Ableitung eng verwandt ist.

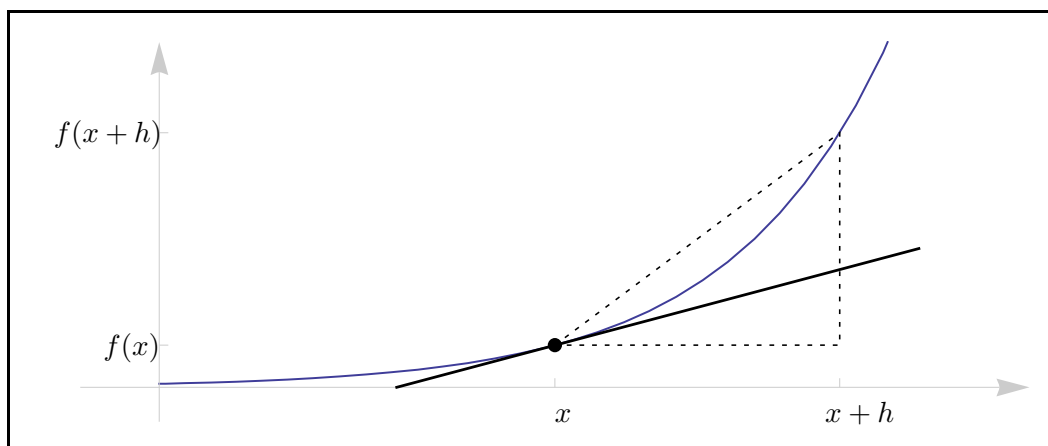


Abbildung 8.1: Illustration zu Differenzenquotienten und Ableitung: Der Differenzenquotient ist die Steigung im gestrichelten Dreieck; die Ableitung von f an der Stelle x ist die Steigung der Tangente an f im Punkt $(x, f(x))$.

Definition 8.1 (Differenzenquotient, Ableitung)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion. Dann bezeichnet man für $x, x+h \in D$ den Term $\frac{f(x+h)-f(x)}{h}$ als *Differenzenquotienten* von f , und seinen Grenzwert

$$f'(x) := \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

(falls dieser existiert) als *Ableitung (Differentialquotient)* von f in x . Als alternative Schreibweise für $f'(x)$ wird auch $df(x)/dx$ verwendet. Falls $f'(x)$ existiert, heißt f in x *differenzierbar*. Eine differenzierbare Funktion ist eine, die in allen Punkten ihres Definitionsbereichs differenzierbar ist.

Die Ableitung einer Funktion entspricht somit der Steigung der *Sekante* an eine Kurve (Gerade, die durch zwei Punkte einer Kurve verläuft), wenn die Sekante zu einer Tangente wird, indem die x -Koordinaten der beiden Punkte immer näher zusammenrücken. Diese Situation ist in Abbildung 8.1 graphisch dargestellt.

Mit Definition der Ableitung als Grenzwert von Differenzenquotienten kann man Ableitungen direkt ausrechnen, wie dies in den nächsten Beispielen demonstriert wird.

Beispiel 8.3 Zu bestimmen sei die Ableitung der Funktion $f(x) = x^2$. Es gilt

$$\begin{aligned} f'(x) &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \rightarrow 0} \frac{x^2 + 2xh + h^2 - x^2}{h} \\ &= \lim_{h \rightarrow 0} 2x + h = 2x. \end{aligned} \quad \square$$

Beispiel 8.4 Seien f und g zwei beliebige Funktionen. Dann gilt für die Ableitung der Summe dieser Funktionen

$$\begin{aligned}(f + g)'(x) &= \lim_{h \rightarrow 0} \frac{(f + g)(x + h) - (f + g)(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x + h) + g(x + h) - f(x) - g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} + \lim_{h \rightarrow 0} \frac{g(x + h) - g(x)}{h} \\ &= f'(x) + g'(x). \quad \square\end{aligned}$$

Der Zusammenhang zwischen Stetigkeit und Differenzierbarkeit ist wie folgt: Jede differenzierbare Funktion ist stetig (wir werden das nicht beweisen), aber nicht jede stetige Funktion ist differenzierbar. Dies ist aus folgendem Beispiel ersichtlich.

Beispiel 8.5 Für die Ableitung der Betragsfunktion betrachten wir links- und rechtsseitige Grenzwerte im Punkt 0. Es gilt

$$\lim_{h \rightarrow 0^+} \frac{|0 + h| - |0|}{h} = \lim_{h \rightarrow 0^+} \frac{0 + h - 0}{h} = 1$$

und

$$\lim_{h \rightarrow 0^-} \frac{|0 + h| - |0|}{h} = \lim_{h \rightarrow 0^-} \frac{-(0 + h) - 0}{h} = -1.$$

Da die beiden Ergebnisse nicht identisch sind, existiert der Grenzwert nicht, und die Ableitung ist an dieser Stelle nicht definiert. \square

Somit ist Differenzierbarkeit ein stärkeres Kriterium als Stetigkeit: Eine stetige Funktion hat keine Sprungstellen; eine differenzierbare Funktion darf zusätzlich auch keine "Knicke" haben, also Stellen, an denen die Funktion nicht glatt verläuft.

Durch direktes Nachrechnen des Grenzwertes können die Ableitungen bestimmt werden, die im nächsten Satz zusammengefasst sind. Für die Funktionen, die über die Potenzreihen in Definition 7.6 gegeben sind, ergeben sich die Ableitungen durch stückweises Differenzieren der Terme in den Summen. Ein Satz (auf den wir nicht näher eingehen) besagt, dass dies auch bei den unendlichen Summen der Potenzreihen erlaubt ist.

Satz 8.1 Es gelten die folgenden Ableitungsregeln, wobei alle Ableitungen nach x erfolgen, und $c \in \mathbb{R}$ eine Konstante ist:

$$\begin{aligned}c' &= 0 & (x^c)' &= c x^{c-1} \\ (e^x)' &= e^x & \log'(x) &= \frac{1}{x} \\ \sin'(x) &= \cos(x) & \cos'(x) &= -\sin(x).\end{aligned}$$

Mit diesem Resultat können wir für *einzelne* Funktionen die Ableitungen berechnen, nicht aber für Summen, Produkte, oder Hintereinanderausführungen von Funktionen. Diese Ergebnisse sind im nächsten Satz zusammengefasst, dessen Resultate sich wieder durch Einsetzen in Definition 8.1 nachrechnen lassen.

Satz 8.2 (Rechenregeln für Ableitungen) Seien f und g zwei differenzierbare Funktionen, und $\alpha, \beta \in \mathbb{R}$ zwei Konstante. Dann gelten folgende Rechenregeln:

$$(\alpha f + \beta g)'(x) = \alpha f'(x) + \beta g'(x) \quad (\text{Linearität der Ableitung})$$

$$(f \cdot g)'(x) = f'(x) \cdot g(x) + g'(x) \cdot f(x) \quad (\text{Produktregel})$$

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x) \cdot g(x) - g'(x) \cdot f(x)}{g^2(x)} \quad (\text{Quotientenregel})$$

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x) \quad (\text{Kettenregel})$$

Die bei der Kettenregel auftretenden Terme $f'(g(x))$ und $g'(x)$ werden *äußere* bzw. *innere* Ableitung genannt.

Beispiel 8.6 Wir berechnen einige Ableitungen.

$$\frac{d}{dx} \sin(x^2 + 2 \log(x)) = \cos(x^2 + 2 \log(x)) \cdot \left(2x + \frac{2}{x}\right)$$

$$\begin{aligned} \frac{d}{dx} \exp(\cos(x^2) \cdot \sin(x)) = \\ \exp(\cos(x^2) \cdot \sin(x)) \cdot \left(-\sin(x^2) \cdot 2x \cdot \sin(x) + \cos(x^2) \cdot \cos(x)\right) \quad \square \end{aligned}$$

Beispiel 8.7 Die Ableitungsregel für Exponentialfunktionen a^x für beliebige positive Basis $a \in \mathbb{R}^+$ erhält man als

$$\begin{aligned} (a^x)' &= \exp'(\log(a^x)) = \exp'(x \log(a)) = \exp(x \log(a)) \log(a) \\ &= a^x \log(a). \end{aligned}$$

Analog dazu gilt für Ableitungen der Logarithmusfunktion $\log_a(x)$ zu beliebiger Basis $a \in \mathbb{R}^+, a \neq 1$ unter Verwendung von $\log_a(x) = \frac{\log(x)}{\log(a)}$ (folgt aus der Definition des Logarithmus):

$$\log'_a(x) = \left(\frac{\log(x)}{\log(a)}\right)' = \frac{1}{\log(a)} \log'(x) = \frac{1}{x \log(a)}. \quad \square$$

Beispiel 8.8 Wir rechnen die Gültigkeit der Quotientenregel durch Verwendung der Produkt- und Kettenregeln durch. Mit der Regel über das Ableiten von Potenzen gilt

$$\left(\frac{f}{g}\right)'(x) = (f \cdot g^{-1})'(x) = f'(x) \cdot g^{-1}(x) + (g^{-1})'(x) \cdot f(x)$$

$$\begin{aligned}
&= f'(x) \cdot g^{-1}(x) - g^{-2}(x) \cdot g'(x) \cdot f(x) \\
&= \frac{f'(x) \cdot g(x) - g'(x) \cdot f(x)}{g^2(x)}. \quad \square
\end{aligned}$$

Beispiel 8.9 Der Beweis der Kettenregel ist etwas subtil. Wir überprüfen anhand eines Beispiels, dass diese Regel das korrekte Ergebnis liefert. Dazu setzen wir direkt in die Definition der Ableitung ein und formen um:

$$\begin{aligned}
(e^{2x})' &= \lim_{h \rightarrow 0} \frac{e^{2(x+h)} - e^{2x}}{h} = \lim_{h \rightarrow 0} \frac{e^{2(x+h)} - e^{2x}}{h} \cdot \frac{2(x+h) - 2x}{2(x+h) - 2x} \\
&= \lim_{h \rightarrow 0} \frac{e^{2x+2h} - e^{2x}}{2(x+h) - 2x} \cdot \frac{2(x+h) - 2x}{h} \\
&= \lim_{h \rightarrow 0} \frac{e^{2x+2h} - e^{2x}}{2h} \cdot 2 = 2 \lim_{h \rightarrow 0} e^{2x} \frac{e^{2h} - 1}{2h} = 2e^{2x} \lim_{h \rightarrow 0} \frac{e^{2h} - 1}{2h} \\
&= 2e^{2x}.
\end{aligned}$$

An dieser Stelle sind zwei Bemerkungen angebracht: Die Umformung im letzten Schritt, die auf $\lim_{h \rightarrow 0} \frac{e^{2h} - 1}{2h} = 1$ beruht, können wir momentan noch nicht begründen. Wir werden uns mit Grenzwerten dieser Art in Abschnitt 8.4 beschäftigen.

Der oben angewandte Trick des Erweiterns um einen geeigneten Term, mit Hilfe dessen sich innere und äußere Ableitung aufspalten lassen, ist nicht immer möglich (da im Allgemeinen nicht garantiert werden kann, dass nicht mit $\frac{0}{0}$ erweitert wird). \square

Wie man aus Definition 8.1 und obigen Beispielen erkennen kann, ist die Ableitung einer Funktion wiederum eine Funktion, für die man (wenn möglich) auch wieder die Ableitung berechnen kann. Diese Überlegung führt zu folgender Definition.

Definition 8.2 (Höhere Ableitungen)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine in D differenzierbare Funktion mit Ableitung f' . Wenn $f'' := (f')'$ definiert ist, dann nennt man diese Funktion die *zweite Ableitung* von f . Analog dazu definiert man (falls sie existieren) die *höheren Ableitungen* von f als

$$f^{(n)} := \begin{cases} (f^{(n-1)})' & \text{für } n > 1 \\ f' & \text{für } n = 1, \end{cases}$$

wobei meist $f^{(3)}$ noch als f''' geschrieben wird.

Beispiel 8.10 Für $f(x) = \cos(x^2)$ gilt

$$\begin{aligned}
f'(x) &= -\sin(x^2) 2x \\
f''(x) &= -\cos(x^2) 4x - 2\sin(x^2) \\
f'''(x) &= \sin(x^2) 8x^3 - \cos(x^2) 12x \\
f^{(4)}(x) &= \cos(x^2) 16x^4 - 12\cos(x^2) + \sin(x^2) 48x^2. \quad \square
\end{aligned}$$

Mit Hilfe des Ableitungsbegriffs lassen sich viele interessante technische und naturwissenschaftliche Aufgabenstellung lösen lassen, wie wir in den nächsten Abschnitten sehen werden.

8.2 Ableitungen und Eigenschaften von Funktionen

Der Graph einer differenzierbaren Funktion weist verschiedene Charakteristika auf (Hochpunkte, Tiefpunkte, steigende/fallende Intervalle, . . .), die sich besonders leicht über die Ableitungen der Funktion charakterisieren lassen. Wir benötigen dafür exakte Definitionen dieser Begriffe.

Definition 8.3 (Monotonie)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion und $[a, b] \subseteq D$. Dann nennt man f auf $[a, b]$

$$\textit{monoton wachsend} \quad :\Leftrightarrow \quad \forall_{x_1, x_2 \in [a, b]} \quad x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2)$$

$$\textit{monoton fallend} \quad :\Leftrightarrow \quad \forall_{x_1, x_2 \in [a, b]} \quad x_1 < x_2 \Rightarrow f(x_1) \geq f(x_2)$$

Wenn sogar stärker $f(x_1) < f(x_2)$ bzw. $f(x_1) > f(x_2)$ gilt, dann nennt man die Funktionen *streng* monoton wachsend bzw. fallend.

Beispiel 8.11 Die Exponential- und Logarithmusfunktionen sind in ihren Definitionsbereichen streng monoton wachsend. Die Funktion $f(x) = x^2$ ist im Intervall $(-\infty, 0]$ streng monoton fallend, im Intervall $[0, \infty)$ streng monoton wachsend.

Die Kosinusfunktion ist für beliebiges $k \in \mathbb{N}$ in den Intervallen $[2k\pi, (2k+1)\pi]$ streng monoton fallend, in den Intervallen $[(2k+1)\pi, (2k+2)\pi]$ streng monoton wachsend. Die Sinusfunktion ist in den Intervallen $[2k\pi + \frac{\pi}{2}, (2k+1)\pi + \frac{\pi}{2}]$ streng monoton fallend und in den Intervallen $[(2k+1)\pi + \frac{\pi}{2}, (2k+2)\pi + \frac{\pi}{2}]$ streng monoton wachsend. \square

Das Monotonieverhalten von Funktionen hängt direkt mit deren Ableitungen zusammen. Wie die meisten der Aussagen in diesem Abschnitt ist auch dieser Zusammenhang, der in Satz 8.4 formuliert wird, graphisch einleuchtend. Wir werden dennoch exemplarisch aufzeigen, wie Zusammenhänge dieser Art bewiesen werden können. Dazu benötigen wir ein fundamentales Resultat der Differentialrechnung. Wir werden dieses Resultat selbst nicht beweisen, sondern auch hier auf die graphische Veranschaulichung in Abbildung 8.2 verweisen.

Satz 8.3 (Mittelwertsatz) Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine auf dem Intervall $[a, b] \subseteq D$ differenzierbare Funktion. Dann gibt es ein $c \in [a, b]$ mit der Eigenschaft

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Der Mittelwertsatz sagt somit aus, dass die Steigung einer differenzierbaren Funktion in einem Intervall zumindest an einem Punkt gleich der Steigung der Sekante durch die Endpunkte dieses Intervalls ist. Dieser Satz kann nun verwendet werden, um den Zusammenhang zwischen Monotonieverhalten und Ableitungen von Funktionen zu beweisen.

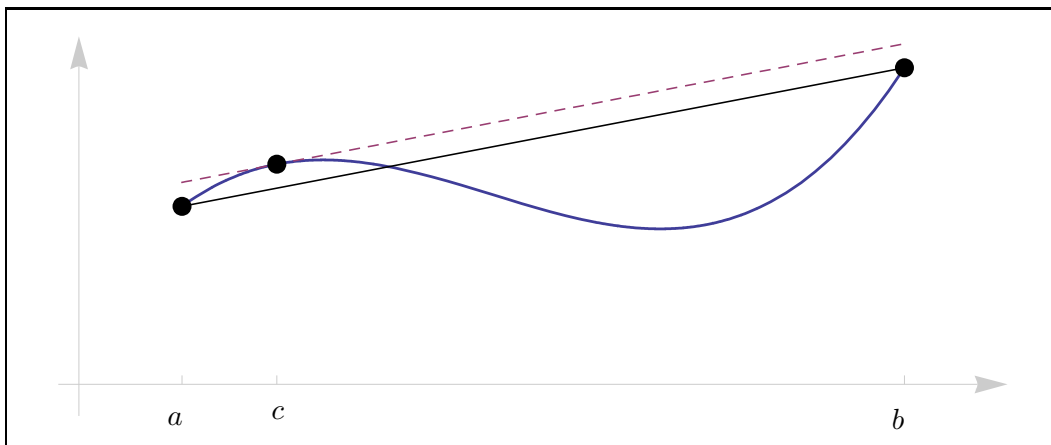


Abbildung 8.2: Illustration zum Mittelwertsatz.

Satz 8.4 Für eine differenzierbare Funktion $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ und Intervall $[a, b] \subseteq D$ gilt:

$$f \text{ in } [a, b] \text{ monoton wachsend} \Leftrightarrow \forall_{x \in [a, b]} f'(x) \geq 0$$

$$f \text{ in } [a, b] \text{ monoton fallend} \Leftrightarrow \forall_{x \in [a, b]} f'(x) \leq 0.$$

Für streng monotone Funktionen gilt, dass die Ableitungen echt größer bzw. kleiner null sind.

Wir beweisen exemplarisch die erste Aussage, und hier als ersten Teil die Implikation

$$f \text{ in } [a, b] \text{ monoton wachsend} \Rightarrow \forall_{x \in [a, b]} f'(x) \geq 0.$$

Wir nehmen die linke Seite dieser Implikation als wahr an, und zeigen damit die rechte Seite. Da f in $[a, b]$ monoton wachsend ist, gilt mit Definition 8.3 für alle $x_1, x_2 \in [a, b]$ die Implikation $x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2)$. Seien also $x_1, x_2 \in [a, b]$ beliebig aber fix, mit der Eigenschaft $x_1 < x_2$. Dann folgt mit modus ponens (Satz 2.3) auch $f(x_1) \leq f(x_2)$.

Wir wissen somit $x_2 - x_1 > 0$ und $f(x_2) - f(x_1) \geq 0$. Mit dem Mittelwertsatz folgt dann, dass es ein $c \in [x_1, x_2]$ geben muss mit der Eigenschaft

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0.$$

Da wir $x_1 < x_2$ in $[a, b]$ beliebig wählen können, folgt daraus $f'(x) \geq 0$ für alle $x \in [a, b]$.

Die Implikation in die andere Richtung beweist man ähnlich: Zu zeigen ist

$$\forall_{x \in [a, b]} f'(x) \geq 0 \Rightarrow f \text{ in } [a, b] \text{ monoton wachsend.}$$

Wir müssen somit unter der Annahme von $f'(x) \geq 0$ auf $[a, b]$ die Implikation

$$\forall_{x_1, x_2 \in [a, b]} x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2)$$

zeigen. Sei also $x_1, x_2 \in [a, b]$ beliebig aber fix mit der Eigenschaft $x_1 < x_2$. Zu zeigen ist $f(x_1) \leq f(x_2)$, also $f(x_2) - f(x_1) \geq 0$. Einsetzen in den Mittelwertsatz garantiert die Existenz eines $c \in [x_1, x_2]$ mit

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0,$$

woraus wegen $x_2 - x_1 > 0$ sofort wie gefordert $f(x_2) - f(x_1) \geq 0$ folgt.

Beispiel 8.12 (Fortsetzung von Beispiel 8.11) Wie wir bereits gesehen haben, sind die Exponential- und Logarithmusfunktionen streng monoton wachsend. Dies ist mit obigem Satz auch aus den Ableitungen ersichtlich, da $\forall x \exp'(x) = \exp(x) > 0$ und $\forall x > 0 \log'(x) = \frac{1}{x} > 0$ gilt. \square

Die Punkte, an denen sich das Monotonieverhalten von Funktionen ändert, sind die *Extremwerte* dieser Funktion. Dazu benötigen wir noch den Begriff der ε -Umgebung $U_\varepsilon(x_0)$, der für $x_0 \in \mathbb{R}$ definiert ist als

$$U_\varepsilon(x_0) = \{x \in \mathbb{R} \mid |x - x_0| < \varepsilon\},$$

also das Intervall all jener Punkte, die von x_0 weniger als ε entfernt sind.

Definition 8.4 (Lokale Minima und Maxima)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$. Dann heißt ein $x_0 \in D$

$$\text{lokales Minimum von } f := \Leftrightarrow \exists_{\varepsilon > 0} \forall_{x \in U_\varepsilon(x_0)} f(x_0) \leq f(x)$$

$$\text{lokales Maximum von } f := \Leftrightarrow \exists_{\varepsilon > 0} \forall_{x \in U_\varepsilon(x_0)} f(x_0) \geq f(x).$$

Die lokalen Minima und Maxima einer Funktion nennt man die *Extremwerte* bzw. *Extrema* dieser Funktion.

Globale Minima und Maxima sind solche, bei denen die Epsilon-Umgebung dieser Punkte der gesamte Definitionsbereich der Funktion ist.

Satz 8.5 Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine in D differenzierbare Funktion. Wenn f in x einen Extremwert hat, dann gilt $f'(x) = 0$.

Man beachte, dass die Umkehrung dieses Satzes *nicht* gilt, dass also aus $f'(x) = 0$ nicht darauf geschlossen werden kann, dass f an dieser Stelle ein Extremum hat.

Beispiel 8.13 Der Graph der Funktion

$$f(x) = 3x^5 - 31x^4 + 118x^3 - 204x^2 + 152x - 27$$

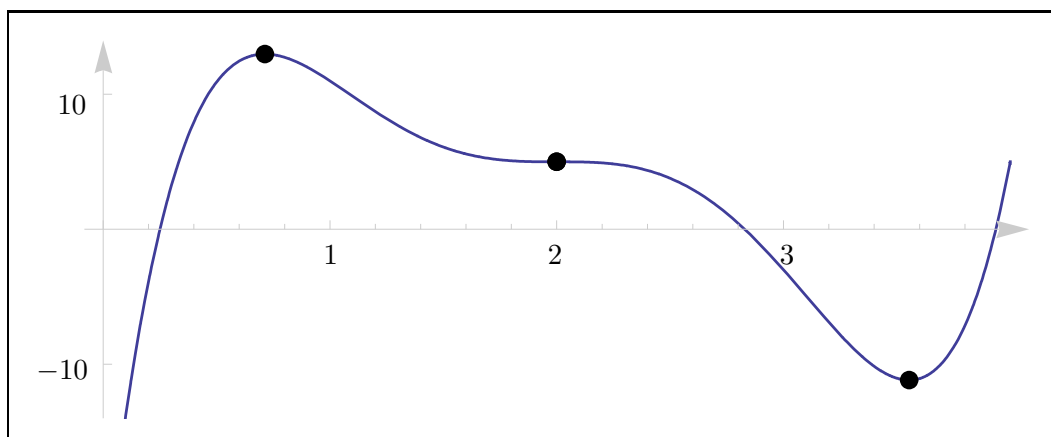


Abbildung 8.3: Graph der Funktion aus Beispiel 8.13. Die Punkte $(x, f(x))$ mit $f'(x) = 0$ sind am Graphen markiert.

ist in Abbildung 8.3 zu sehen. Für diese Funktion ist

$$f'(x) = 15x^4 - 124x^3 + 354x^2 - 408x + 152.$$

Man kann nachrechnen, dass $f'(x)$ in $\xi_1 = 2$ eine doppelte Nullstelle und in $\xi_2 = 0.71285$ und $\xi_3 = 3.55382$ jeweils einfache Nullstellen hat. Wie man aus dem Graphen erkennen kann, befindet sich in ξ_2 ein lokales Maximum, sowie in ξ_3 ein lokales Minimum. Obwohl $f'(\xi_1) = 0$ ist, ist an dieser Stelle kein Extremum. \square

Die etwas unbefriedigende Situation, dass man aus $f'(x) = 0$ noch nicht auf die Existenz eines Extremums in x schließen kann, lässt sich unter Verwendung der zweiten Ableitung $f''(x)$ beheben. Wir holen dazu etwas weiter aus und betrachten zuerst das Krümmungsverhalten von reellen Funktionen.

Definition 8.5 (Konkave und konvexe Funktionen)

Eine Funktion $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ heißt in einem Intervall $[a, b] \subseteq D$

$$\text{konkav} :\Leftrightarrow \forall_{\lambda \in [0,1]} f(\lambda a + (1 - \lambda)b) \geq \lambda f(a) + (1 - \lambda)f(b)$$

$$\text{konvex} :\Leftrightarrow \forall_{\lambda \in [0,1]} f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b).$$

Bei auf $[a, b]$ konkaven Funktionen verläuft der Graph also immer *über* der Sekante zwischen a und b , bei konvexen Funktionen immer *unter* der Sekante.

In konkaven Bereichen wird eine Funktion auch als *nach unten offen* bezeichnet, in konvexen Bereichen *nach oben offen*. Diese beiden Eigenschaften von Funktionen werden zusammenfassend auch als ihr *Krümmungsverhalten* bezeichnet. Dieses Krümmungsverhalten lässt sich über die zweiten Ableitungen von Funktionen charakterisieren. Der folgende Satz fasst die zu Satz 8.4 über das Monotonieverhalten von Funktionen analogen Aussagen zusammen.

Satz 8.6 Für eine zweimal differenzierbare Funktion $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ und Intervall $[a, b] \subseteq D$ gilt:

$$f \text{ in } [a, b] \text{ konkav} :\Leftrightarrow \forall_{x \in [a, b]} f''(x) \leq 0$$

$$f \text{ in } [a, b] \text{ konvex} :\Leftrightarrow \forall_{x \in [a, b]} f''(x) \geq 0.$$

Beispiel 8.14 Wegen $\forall x \exp''(x) = \exp(x) > 0$ ist die Exponentialfunktion auf ganz \mathbb{R} konvex. Im Gegensatz dazu ist die Logarithmusfunktion wegen $\log''(x) = -\frac{1}{x^2} < 0$ auf ganz \mathbb{R}^+ konkav.

Da für die Sinusfunktion $\sin''(x) = -\sin(x)$ gilt, ist die Sinusfunktion in den Intervallen $[2k\pi, (2k+1)\pi]$ konkav, und in den Intervallen $[(2k+1)\pi, (2k+2)\pi]$ konvex. Für die Kosinusfunktion gilt wegen $\cos''(x) = -\cos''(x)$, dass sie in den Intervallen $[2k\pi + \frac{\pi}{2}, (2k+1)\pi + \frac{\pi}{2}]$ konvex, und in den Intervallen $[(2k+1)\pi + \frac{\pi}{2}, (2k+2)\pi + \frac{\pi}{2}]$ konkav ist. \square

Die Punkte, an denen sich das Krümmungsverhalten einer Funktion ändert, werden als Wendepunkte dieser Funktion bezeichnet.

Definition 8.6 (Wendepunkt)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine in D differenzierbare Funktion. Ein Extremum von f' wird als *Wendepunkt* von f bezeichnet.

In Analogie zu Satz 8.5 über die notwendige Bedingung für die Existenz von Extremwerten gilt folgender Satz.

Satz 8.7 Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine in D differenzierbare Funktion. Wenn f in x einen Wendepunkt hat, dann gilt $f''(x) = 0$.

Wie schon bei den Extremwerten gilt die Umkehrung dieses Satz *nicht*. Dies zeigt folgendes Beispiel.

Beispiel 8.15 Die Funktion

$$f(x) = x^7 - 12x^6 + 36x^5 + 64x^4 - 384x^3 + 1024x + 500$$

mit der zweiten Ableitung

$$f''(x) = 42x^5 - 360x^4 + 720x^3 + 768x^2 - 2304x$$

ist in Abbildung 8.4 dargestellt. Die Nullstellen der zweiten Ableitung liegen bei $\xi_1 = -1.58784, \xi_2 = 0, \xi_3 = 2.15927$, sowie $\xi_4 = 4$ (doppelte Nullstelle). Obwohl $f''(4) = 0$ ist, liegt an dieser Stelle kein Wendepunkt vor, da dieser Punkt kein Extremwert von $f'(x)$ ist. Wir werden bald sehen, wie man dies aus dem obigen Datenmaterial erkennen kann. \square

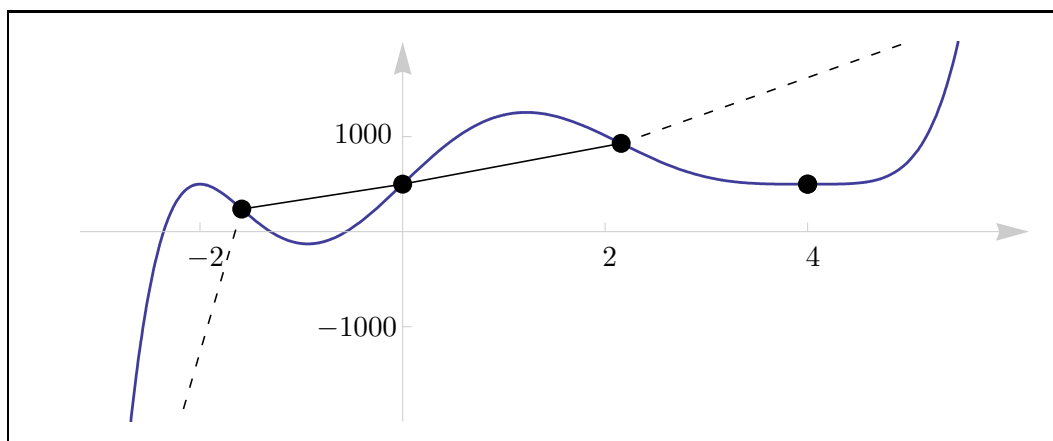


Abbildung 8.4: Graph der Funktion aus Beispiel 8.15. Die Punkte $(x, f(x))$ mit $f''(x) = 0$ sind am Graphen markiert. Zur leichteren Sichtbarkeit des Krümmungsverhaltens sind die Sekanten durch die Wendepunkte eingezeichnet. Links von diesen Punkten ist die Funktion konkav, rechts davon konvex.

Nach diesen Ausführungen über das Krümmungsverhalten und die Wendepunkte von Funktionen kehren wir nun wieder zur Frage zurück, die durch Satz 8.5 und das darauffolgende Beispiel 8.13 aufgeworfen wurde: Wie lassen sich Extremwerte von Funktionen charakterisieren? Der nächste Satz gibt nach der notwendigen Bedingung von Satz 8.5 auch *hinreichende* Bedingungen an.

Satz 8.8 Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine in D zweimal differenzierbare Funktion, sowie $x_0 \in D$. Dann gilt

$$f'(x_0) = 0 \wedge f''(x_0) > 0 \Rightarrow x_0 \text{ lokales Minimum von } f$$

$$f'(x_0) = 0 \wedge f''(x_0) < 0 \Rightarrow x_0 \text{ lokales Maximum von } f.$$

Wenn für einen Punkt x_0 sowohl $f'(x_0) = 0$ als auch $f''(x_0) = 0$ gilt, dann ist obiger Satz nicht anwendbar, und x_0 kann, muss aber kein Extremwert von f sein.

Diese erste Situation liegt etwa beim Punkt $x_0 = 0$ der Funktion $f(x) = x^4$ vor. Hier gilt $f'(x_0) = f''(x_0) = 0$, und f hat in x_0 ein Minimum. Die zweite Situation liegt beim Punkt $\xi_1 = 2$ in Beispiel 8.13 vor. Auch hier ist $f'(\xi_1) = f''(\xi_1) = 0$, f hat in ξ_1 jedoch *kein* Extremum. Die gleiche Situation ergibt sich auch im Punkt 0 der Funktion $f(x) = x^3$.

Analog zur den Kriterien für die Existenz von Extremwerten lässt sich die Existenz von Wendepunkten durch Bedingungen an die *dritte* Ableitung in diesen Punkten charakterisieren. Aus der Kombination von Definition 8.6 und Satz 8.8 folgt sofort, dass f in x_0 dann einen Wendepunkt hat, wenn $f''(x_0) = 0$ und $f'''(x_0) \neq 0$ sind.

Beispiel 8.16 (Fortsetzung von Beispiel 8.15) Für die Punkte $\xi_1 = -1.58784$, $\xi_2 = 0$, $\xi_3 = 2.15927$, $\xi_4 = 4$ in Abbildung 8.4 gilt, dass f'' an diesen Stellen null

wird. Wegen

$$f'''(x) = 210x^4 - 1440x^3 + 2160x^2 + 1536x - 2304$$

gilt $f'''(\xi_1) = 7802.61$, $f'''(\xi_2) = -2304$, $f'''(\xi_3) = 1151.42$, und $f'''(\xi_4) = 0$. Somit besitzt f in ξ_1 , ξ_2 und ξ_3 Wendepunkte, in ξ_4 aber nicht. \square

Als Anwendungen dieser Überlegungen betrachten wir zwei Beispiele von *Extremwertaufgaben*.

Beispiel 8.17 Zu Verpackungszwecken soll ein Quader entworfen werden, der mit minimaler Oberfläche ein Volumen von einem Liter fassen kann. Wie lange sollen die Seiten dieses Quaders sein?

Die Formel für das Volumen eines Quaders mit drei Seitenlängen a , b und c ist $V = abc$, für die Oberfläche $O = 2ab + 2ac + 2bc$. Für $V = 1$ ist $c = 1/(ab)$; daraus ergibt sich die Oberfläche als Funktion von a für konstantes b

$$O(a) = 2\left(ab + \frac{1}{a} + \frac{1}{b}\right).$$

Man erhält aus $O'(a) = 2b - \frac{2}{a^2} = 0$ den Zusammenhang $a = \frac{1}{\sqrt{b}}$. Da diese Überlegung ebenso für O als Funktion von b gemacht werden kann, erhält man außerdem $b = \frac{1}{\sqrt{a}}$. Das nichtlineare Gleichungssystem

$$a = \frac{1}{\sqrt{b}}, \quad b = \frac{1}{\sqrt{a}} \quad (8.1)$$

hat die eindeutige Lösung $a = b = 1$, und damit auch $c = 1$. Wegen $O''(a) = \frac{4}{a^3}$ ist $O''(1) = 4 > 0$; der Würfel mit Seitenlänge 10cm ist somit der Quader mit minimaler Oberfläche. \square

Beispiel 8.18 Gesucht ist für gegebene zwei Vektoren v und w derjenige Punkt auf w , der von v minimalen Abstand hat. Ein beliebiger Punkt auf w ist durch λw mit skalarem λ gegeben; der Abstand dieser Punkte als Funktion von λ ist

$$d(\lambda) = \|v - \lambda w\| = \sqrt{(v - \lambda w) \cdot (v - \lambda w)}.$$

Nullsetzen der Ableitung $d'(\lambda)$ liefert

$$d'(\lambda) = \frac{1}{2} \frac{-2v \cdot w + 2\lambda w \cdot w}{\sqrt{(v - \lambda w) \cdot (v - \lambda w)}} = 0,$$

woraus durch Umformen wiederum der Projektions-Anteil wie in Definition 6.5 ersieht:

$$\lambda = \frac{v \cdot w}{w \cdot w}. \quad \square$$

8.3 Taylorreihen

In diesem Abschnitt werden wir das Konzept der Potenzreihen verallgemeinern und zeigen, dass Funktionen damit lokal sehr gut approximiert werden können.

Zum Einstieg in diese Problematik betrachten wir die beste lineare Approximation an eine Funktion. Um die Notation aus Definition 8.1 mit der dieses Abschnitts konsistent zu halten, ersetzen wir in der dortigen Definition der Ableitung jedes x durch x_0 und anschließend h durch $x - x_0$. Man erhält dann

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Für $x_0 \approx x$ fällt das Weglassen des Grenzwerts nicht sehr ins Gewicht; Auflösen nach $f(x)$ liefert

$$f(x) \approx \underbrace{f(x_0) + (x - x_0)f'(x_0)}_{=: t(x)}.$$

Für gegebenes (konstantes) x_0 ist durch die rechte Seite dieser Approximation eine lineare Funktion $t(x)$ definiert: die Tangente an f in x_0 . Wie man sofort nachrechnen kann, gilt

$$t(x_0) = f(x_0) \quad \text{und} \quad t'(x_0) = f'(x_0),$$

die Funktionswerte und Ableitungen sind in x_0 für f und t gleich.

Man erhält bessere Approximationen \tilde{f} an f , wenn man zusätzlich fordert, dass auch höhere Ableitungen von \tilde{f} und f in x_0 gleich sind. Wenn wir \tilde{f} auf die Klasse der Polynomfunktionen beschränken, lassen sich die zugehörigen Koeffizienten über ein lineares Gleichungssystem bestimmen. In Analogie zur linearen Situation wählen wir den Ansatz

$$\tilde{f}(x) = a(x - x_0)^2 + b(x - x_0) + c,$$

woraus sich wegen $\tilde{f}(x_0) = f(x_0)$ sofort $c = f(x_0)$ ergibt. Weiters ist

$$\tilde{f}'(x) = 2a(x - x_0) + b,$$

woraus mit der Bedingung $\tilde{f}'(x_0) = f'(x_0)$ sofort $b = f'(x_0)$ folgt. Mit $\tilde{f}''(x_0) = f''(x_0)$ folgt schließlich $a = \frac{f''(x_0)}{2}$, sodass die beste quadratische Approximation an f in x_0 gegeben ist durch

$$\tilde{f}(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

Die hier gezeigte Herleitung lässt sich beliebig für höhergradige Polynome fortsetzen; dies führt zum Begriff der *Taylorpolynome*.

Definition 8.7 (Taylorpolynome)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, die in x_0 mindestens n -mal differenzierbar ist. Dann nennt man

$$T_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

das *Taylorpolynom n -ter Ordnung* von f mit *Entwicklungspunkt* x_0 .

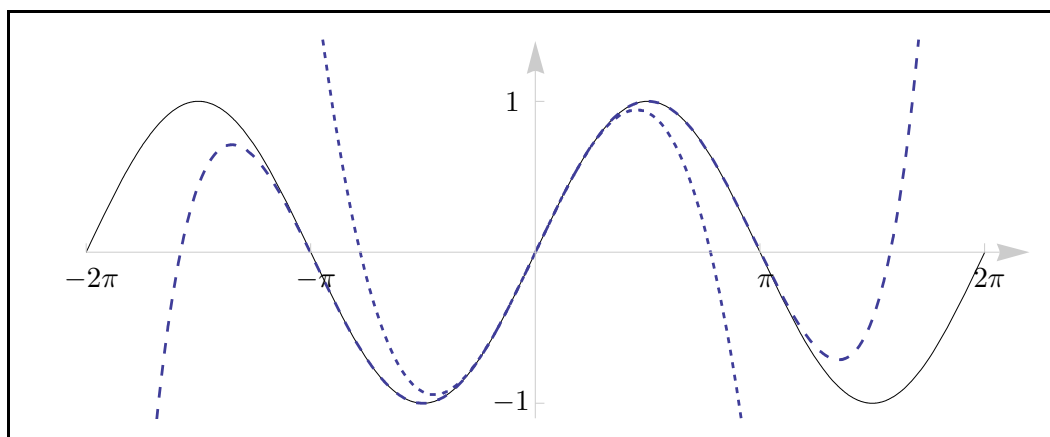


Abbildung 8.5: Die Sinusfunktion sowie die Taylorpolynome dritter (gepunktet) und neunter Ordnung (gestrichelt) mit Entwicklungspunkt $x_0 = 0$.

Beispiel 8.19 Das Taylorpolynom n -ter Ordnung für die Sinus-Funktion mit Entwicklungspunkt $x_0 = 0$ ist

$$T_n(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + (-1)^{(n-1)/2} \frac{x^n}{n!}.$$

Man beachte, dass in diesem Taylorpolynom wegen $\sin(0) = 0$ alle geraden Potenzen wegfallen.

In Abbildung 8.5 sind die Sinusfunktion sowie die Approximationen dritter und neunter Ordnung dargestellt. \square

Der interessanteste Aspekt von Taylorpolynomen ist allerdings nicht, dass man damit Funktionen approximieren kann (lokal sogar beliebig genau), sondern dass

- Funktionen sich über unendliche Taylorpolynome definieren lassen, und
- der Fehler in der Approximation der Taylorpolynome abgeschätzt werden kann.

Wir werden diese beiden Punkte nun näher erläutern.

Definition 8.8 (Taylorreihen)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, die in x_0 beliebig oft differenzierbar ist. Dann nennt man

$$\begin{aligned} T(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \dots \\ &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \end{aligned}$$

die *Taylorreihe* von f mit *Entwicklungspunkt* x_0 .

Wie man unschwer erkennen kann, sind Taylorreihen eine Verallgemeinerung der Potenzreihen aus Abschnitt 7.3; bei Potenzreihen ist der Entwicklungspunkt immer 0. Man sieht auch, dass die Potenzreihen aus Definition 7.6 eigentlich Taylorreihen sind.

Beispiel 8.20 Wegen $(e^x)' = e^x$ und $e^0 = 1$ ist die Taylorreihenentwicklung von e^x im Entwicklungspunkt $x_0 = 0$ gegeben durch

$$T(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Wie wir bereits wissen, gilt hier $T(x) = e^x$. □

Die obige Definition und das obige Beispiel werfen die Frage auf, unter welchen Umständen $T(x) = f(x)$ gilt, die Taylorreihenentwicklung einer Funktion also mit der Funktion identisch ist. Wir müssen dafür den Begriff des Konvergenzradius für Potenzreihen aus Definition 7.5 für Taylorreihen mit beliebigem Entwicklungspunkt verallgemeinern. Die Taylorreihe $T(x)$ mit Entwicklungspunkt x_0 hat den Konvergenzradius r , wenn gilt

$$\forall x \in (x_0 - r, x_0 + r) \quad T(x) \text{ konvergiert.}$$

Zur Berechnung des Konvergenzradius gilt mit einer ähnlichen Überlegung wie der, die zu Definition 7.5 geführt hat (mit $a_n = f^{(n)}(x_0)/n!$)

$$r = \lim_{n \rightarrow \infty} (n+1) \left| \frac{f^{(n)}(x_0)}{f^{(n+1)}(x_0)} \right| \quad \text{bzw.} \quad r = \lim_{n \rightarrow \infty} \sqrt[n]{\frac{n!}{|f^{(n)}(x_0)|}}.$$

Innerhalb des Konvergenzradius sind eine Funktion und ihre Taylorreihenentwicklung identisch, wie der folgende Satz besagt.

Satz 8.9 Sei $T(x)$ die Taylorreihenentwicklung von f im Punkt x_0 , und r der Konvergenzradius von $T(x)$. Dann gilt

$$\forall x \in (x_0 - r, x_0 + r) \quad T(x) = f(x).$$

Beispiel 8.21 Wegen $(e^x)^{(n)} = e^x$ ist für $x_0 = 0$ der Konvergenzradius der Exponentialfunktion

$$r = \lim_{n \rightarrow \infty} (n+1) \frac{1}{1} = \infty.$$

Für die Taylorreihenentwicklung der Funktion \sqrt{x} im Entwicklungspunkt $a > 0$ gilt

$$T(x) = \sqrt{a} + \frac{1}{2\sqrt{a}}(x-a) - \frac{1}{8a^{3/2}}(x-a)^2 + \frac{1}{16a^{5/2}}(x-a)^3 + \cdots + (-1)^{(n+1)} \frac{1 \cdot 3 \cdots (2n-3)}{n! 2^n a^{(2n-1)/2}} (x-a)^n,$$

sodass für den Konvergenzradius gilt

$$r = \lim_{n \rightarrow \infty} (n+1) \frac{1 \cdot 3 \cdots (2n-3)}{2^n a^{(2n-1)/2}} \bigg/ \frac{1 \cdot 3 \cdots (2n-1)}{2^{n+1} a^{(2n+1)/2}}$$

$$= \lim_{n \rightarrow \infty} (n+1) \frac{2a}{2n-1} = 2a \lim_{n \rightarrow \infty} \frac{n+1}{2n-1} = a.$$

Dieses Ergebnis ist auch anschaulich klar: Da \sqrt{x} nur für $x \geq 0$ definiert ist, kann für Entwicklungspunkt a der Konvergenzradius gar nicht größer als a sein. \square

Man kann beweisen, dass man Taylorreihen innerhalb ihres Konvergenzradius gliedweise addieren, multiplizieren, und auch differenzieren bzw. integrieren kann. Damit kann man etwa direkt die Korrektheit der Ableitungsregeln aus Satz 8.1 beweisen, oder auch die Rechenregel $\sin^2(x) + \cos^2(x) = 1$.

Beispiel 8.22 Wir können nachrechnen, dass $\sin(x)$ ein Beispiel einer *ungeraden Funktion* ist: Für solche Funktionen gilt $f(-x) = -f(x)$. Wegen

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

gilt

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

Im Gegensatz dazu sind *gerade Funktionen* diejenigen Funktionen f , für die $f(-x) = f(x)$ gilt. Der Kosinus ist eine gerade Funktion, wie man aus der Taylorreihenentwicklung

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

erkennen kann. Damit gilt

$$\cos(-x) = \sum_{n=0}^{\infty} (-1)^n \frac{(-x)^{2n}}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos(x). \quad \square$$

Wie vorher bereits angedeutet ist ein großer Vorteil der Taylorreihenentwicklung der, dass man den Fehler in einer endlichen Approximation (über ein Taylorpolynom) abschätzen kann. Formal ist dies im folgenden Satz ausgedrückt.

Satz 8.10 Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, die in x_0 beliebig oft differenzierbar ist. Dann gilt

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + R_n(x),$$

wobei das *Restglied*

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{(n+1)}$$

den Fehler zwischen endlicher und unendlicher Approximation angibt. Dabei ist ξ eine Zahl zwischen x_0 und x .

Aus diesem Satz kann man erkennen, dass die Differenz $T(x) - T_n(x)$ durch einen Term $R_n(x)$ angegeben werden kann, der die gleiche Struktur wie die Terme von Taylorpolynomen hat. Den *exakten* Fehler kann man nicht angeben, da man den Wert von ξ nicht kennt. Man kann den Fehler aber *beschränken*, indem man für ξ denjenigen Wert einsetzt, der das Restglied maximiert. Dann ist der echte Fehler sicher nicht größer als der so berechnete Wert. Wir illustrieren dies anhand von Beispielen.

Beispiel 8.23 In Abbildung 8.5 sind zwei Approximationen der Sinusfunktion durch Taylorpolynome zu sehen. Für die Approximation dritter Ordnung gilt

$$\sin(x) = x - \frac{x^3}{3!} + R_3(x),$$

für die Approximation neunter Ordnung

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} + R_9(x).$$

Da n immer ungerade ist, ist die $(n+1)$ -te Ableitung von $\sin(x)$ wiederum $\pm \sin(x)$, und das Restglied ist

$$R_n(x) = (-1)^{\frac{n+1}{2}} \frac{\sin(\xi)}{(n+1)!} x^{(n+1)}.$$

Wir verwenden Taylorpolynome unterschiedlicher Ordnung, um die Sinusfunktion an der Stelle $x = 3$ zu approximieren. Das Restglied gibt an, wie groß der Fehler dieser Approximation maximal ist. Da ein Fehler immer positiv ist, sind wir nur am Absolutbetrag von $R_n(x)$ interessiert. Wegen $|\sin(\xi)| \leq 1$ gilt

$$|R_n(3)| \leq \frac{1}{(n+1)!} 3^{n+1}.$$

Die Abschätzungen und die tatsächlichen Fehler (als Absolutbetrag) sind für $n = 3, 5, 7$ und 9 in folgender Tabelle zusammengefasst.

	$n = 3$	$n = 5$	$n = 7$	$n = 9$
Abschätzung	3.375	1.0125	0.1627	0.01627
tatsächlicher Fehler	1.641	0.3839	0.05005	0.004192

Die Abschätzung ist in diesem Beispiel für ungerade n etwa um den Faktor 3 zu hoch; dies liegt daran, dass bei der Taylorreihenentwicklung des Sinus jeder zweite Term wegfällt. Da der Konvergenzradius des Sinus unendlich ist, wird die Approximation für höhere n immer besser. Dies gilt unabhängig von dem in diesem Beispiel gewählten Wert von $x = 3$. \square

Das folgende Beispiel zeigt, welchen Einfluss der Konvergenzradius auf die Approximationen durch Taylorpolynome hat.

Beispiel 8.24 Wir gehen wie im letzten Beispiel für den natürlichen Logarithmus vor, den wir um den Punkt $x_0 = 1$ entwickeln. Allgemein gilt

$$\log^{(n)}(x) = (-1)^{n+1} \frac{(n-1)!}{x^n},$$

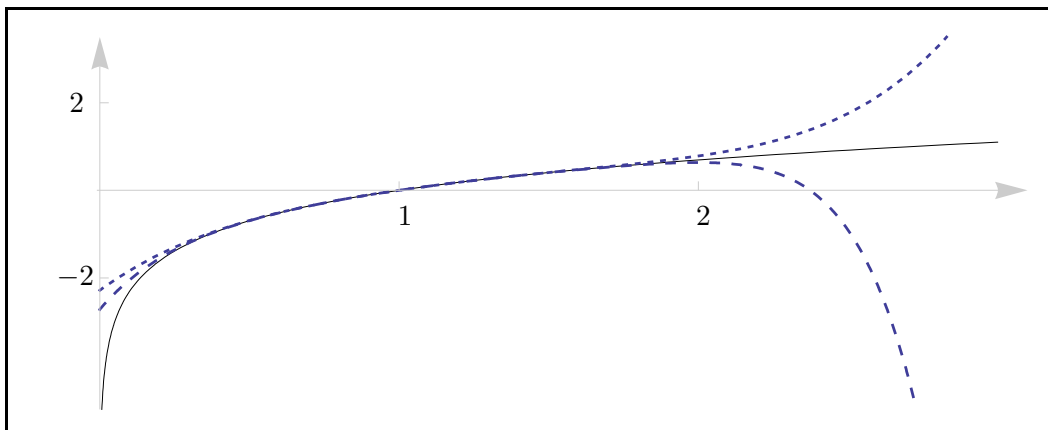


Abbildung 8.6: Der natürliche Logarithmus sowie die Taylorpolynome fünfter (gepunktet) und achter Ordnung (gestrichelt) mit Entwicklungspunkt $x_0 = 1$.

und somit ist die Entwicklung des Logarithmus um 1 durch die Reihe

$$\log(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \dots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(x-1)^n}{n}$$

gegeben (vergleiche auch Definition 7.6 mit der Substitution $x \rightarrow x+1$). Die Taylorpolynome fünfter und achter Ordnung sind zusammen mit dem Logarithmus in Abbildung 8.6 zu sehen. Man erkennt, dass im Gegensatz zum Sinus der Konvergenzradius endlich ist (nämlich 1), und die Approximation außerhalb des Konvergenzradius immer schlechter wird.

Numerisch kann man mit dem Restglied wiederum eine Abschätzung des Approximationsfehlers machen. Für den Logarithmus, entwickelt um $x_0 = 1$, ist das Restglied

$$R_n(x) = \frac{(-1)^n}{(n+1)\xi^{(n+1)}}(x-1)^{(n+1)}.$$

Wir evaluieren die Taylorpolynome $T_n(x)$ für $x = 1.9$, also knapp innerhalb des Konvergenzradius. Da $\xi \in [1, 1.9]$ ist und $|R_n(x)|$ immer kleiner wird, je größere Werte wir für ξ einsetzen, wird das Restglied für $\xi = 1$ maximiert. Man erhält die Abschätzung

$$|R_n(1.9)| \leq \frac{1}{(n+1)1^{(n+1)}}(1.9-1)^{(n+1)} \quad (8.2)$$

Eine Fehlertabelle für $x = 1.9$ und $n = 5, 6, 7, 8$ ist unten angeführt.

	$n = 5$	$n = 6$	$n = 7$	$n = 8$
Abschätzung	0.08857	0.06833	0.05381	0.04305
tatsächlicher Fehler	0.05022	0.03835	0.02997	0.02383

Innerhalb des Konvergenzradius wird die Approximation also mit höhergradigen Taylorpolynomen immer besser. Ausserhalb des Konvergenzradius wird die Approximation allerdings schlechter. Wir schätzen den Fehler für $x = 2.5$ ab. Dann ist $\xi \in [1, 2.5]$, und $|R_n(2.5)|$ wird wiederum durch $\xi = 1$ maximiert. Wir erhalten folgende Werte:

	$n = 5$	$n = 6$	$n = 7$	$n = 8$
Abschätzung	1.898	2.4409	3.204	4.271
tatsächlicher Fehler	0.8368	1.062	1.379	1.824

Außerhalb des Konvergenzradius wächst der Fehler somit mit dem Grad des Taylorpolynoms; die Abschätzung durch das Restglied ist aber weiterhin gültig. \square

Beispiel 8.25 Wir können die Taylorreihenentwicklung auch verwenden, um eine numerische Approximation an $\pi \approx 3.141592654$ zu generieren. Wie man sich durch eine Skizze überlegen kann, sind bei $45^\circ = \frac{\pi}{4}$ die Werte der Sinus- und der Kosinusfunktion identisch. Somit gilt

$$\frac{\sin\left(\frac{\pi}{4}\right)}{\cos\left(\frac{\pi}{4}\right)} = \tan\left(\frac{\pi}{4}\right) = 1,$$

sodass für die Umkehrfunktion arctan des Tangens $\arctan(1) = \frac{\pi}{4}$ gilt. Weiters gilt (ohne Beweis)

$$\arctan'(x) = \frac{1}{1+x^2},$$

woraus nach einigem Umrechnen die Taylorreihenentwicklung von arctan mit Entwicklungspunkt 0 folgt:

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$$

Einsetzen von 1 in diese Formel liefert

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots\right)$$

Diese Formel konvergiert nur sehr langsam gegen π , wie man auch aus dem Restglied dieser Reihenentwicklung erkennen kann. Ebenfalls ohne Beweis gilt, dass der Term $|\arctan^{(n+1)}(\xi)|$ für gerades n und für ξ zwischen Entwicklungspunkt 0 und Evaluationspunkt 1 bei $\xi = 0$ maximiert wird, sodass

$$|R_n(1)| \leq \frac{1}{n+1}$$

gilt. Ein Vergleich von Abschätzung $\frac{4}{n+1}$ und echtem Fehler für verschiedene Werte von n liefert folgende Werte:

	$n = 10$	$n = 100$	$n = 1000$	$n = 10000$
Abschätzung	0.3636	0.039	0.0039	0.0004
tatsächlicher Fehler	0.198	0.02	0.002	0.0002

Man kann erkennen, dass diese Abschätzung für jede weitere Stelle Genauigkeit eine Zehnerpotenz mehr Terme benötigt. Im Gegensatz dazu erhält man mit der Formel

$$\frac{1}{\pi} = \frac{2\sqrt{2}}{9801} \sum_{k=1}^{\infty} \frac{(4k)!(1103 + 26390k)}{(k!)^4 396^{4k}}$$

(die Basis der besten heute bekannten Approximationen) mit jedem weiteren Term *acht* weitere Stellen Genauigkeit. \square

8.4 Die Grenzwertregel von de l'Hôpital

Grenzwerte, die zu $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ evaluieren (wie etwa $\lim_{h \rightarrow 0} \frac{e^{2h}-1}{2h}$ aus Beispiel 8.9), können mit den bis jetzt bekannten Grenzwertregeln nicht berechnet werden.

Man kann etwa plausibel argumentieren, dass $\lim_{n \rightarrow \infty} \frac{2^n}{n^2} = \infty$ ist, da Exponentialfunktionen schneller wachsen als Potenzfunktionen. In diesem Abschnitt werden wir eine Methode erarbeiten, wie Argumente dieser Art formalisiert werden können.

Als einfachsten Fall betrachten wir die Situation $\frac{0}{0}$, also den Grenzwert $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$ des Quotienten zweier Funktionen f, g mit $f(x_0) = 0$ und $g(x_0) = 0$. Wir nehmen an, dass f und g in x_0 differenzierbar sind (dies ist bei allen Beispielen in diesem Abschnitt der Fall). Wir begründen die Plausibilität des folgenden allgemeinen Satzes, indem wir diesen Spezialfall nachprüfen und dabei von hinten anfangen:

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} &= \frac{f'(x_0)}{g'(x_0)} = \frac{\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}}{\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0}} \\ &= \frac{\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}}{\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0}} = \frac{\lim_{x \rightarrow x_0} f(x) - f(x_0)}{\lim_{x \rightarrow x_0} g(x) - g(x_0)} = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)} \\ &= \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}. \end{aligned}$$

Der Grenzwert eines Bruches $\frac{0}{0}$ ist somit gleich dem Grenzwert des Bruches aus den Ableitungen von Zähler und Nenner des ursprünglichen Bruches.

Die Situation $\frac{\infty}{\infty}$ ist schwieriger zu behandeln. Wir verzichten hier auf eine genauere Überlegung und fassen zusammen.

Satz 8.11 (Regel von de l'Hôpital) Seien $f, g : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ zwei Funktionen mit

- $\lim_{x \rightarrow x_0} f(x) = 0$ und $\lim_{x \rightarrow x_0} g(x) = 0$, oder
- $\lim_{x \rightarrow x_0} f(x) = \infty$ und $\lim_{x \rightarrow x_0} g(x) = \infty$,

und $x_0 \in \mathbb{R}$ oder $x_0 = \pm\infty$. Dann gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

wenn dieser Grenzwert existiert.

Die Situationen $\frac{-\infty}{\infty}$, $\frac{\infty}{-\infty}$ bzw. $\frac{-\infty}{-\infty}$ können durch Multiplizieren der jeweiligen Funktion mit (-1) auf die Form $\frac{\infty}{\infty}$ gebracht werden, sodass auch in diesen Fällen obiger Satz angewandt werden kann.

Beispiel 8.26 Wir betrachten das bereits angesprochene Beispiel $\lim_{h \rightarrow 0} \frac{e^{2h}-1}{2h}$. Mit obigem Satz gilt

$$\lim_{h \rightarrow 0} \frac{e^{2h}-1}{2h} = \lim_{h \rightarrow 0} \frac{2e^{2h}}{2} = \lim_{h \rightarrow 0} e^{2h} = 1.$$

Der Grenzwert $\lim_{x \rightarrow 0} \frac{\sin(x)}{x}$ evaluiert zu

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = 1. \quad \square$$

Beispiel 8.27 Grenzwerte $\frac{\infty}{\infty}$ sind ebenso leicht zu evaluieren. So gilt etwa

$$\lim_{x \rightarrow \infty} \frac{\log(x)}{e^x} = \lim_{x \rightarrow \infty} \frac{1}{x e^x} = 0.$$

Das ebenfalls oben angesprochene Beispiel $\lim_{n \rightarrow \infty} \frac{2^n}{n^2}$ ergibt mit der Ableitungsregel $(2^n)' = \log(2)2^n$ und dem zweimaligen Anwenden der Regel von de l'Hôpital

$$\lim_{n \rightarrow \infty} \frac{2^n}{n^2} = \lim_{n \rightarrow \infty} \frac{\log(2)2^n}{2n} = \lim_{n \rightarrow \infty} \frac{\log^2(2)2^n}{2} = \infty. \quad \square$$

Beispiel 8.28 Dieses Beispiel zeigt, dass die Forderung nach der Existenz des umgeformten Grenzwerts in obigem Satz notwendig ist:

$$\lim_{x \rightarrow \infty} \frac{x + \sin(x)}{x} = \lim_{x \rightarrow \infty} \frac{1 + \cos(x)}{1} = 1 + \lim_{x \rightarrow \infty} \cos(x);$$

dieser Grenzwert existiert nicht. Direktes Umformen des ursprünglichen Grenzwerts, zusammen mit der Überlegung, dass $\lim_{x \rightarrow \infty} \frac{\sin(x)}{x} = 0$ ist, ergibt die richtige Lösung

$$\lim_{x \rightarrow \infty} \frac{x + \sin(x)}{x} = \lim_{x \rightarrow \infty} \left(1 + \frac{\sin(x)}{x} \right) = 1. \quad \square$$

Mit einigen Rechenricks und Umformungsregeln lassen sich auch Grenzwerte wie $0 \cdot \infty$ oder ∞^0 (oder ähnliche) auf die Form $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ bringen, wo dann wieder die Regel von de l'Hôpital anwendbar ist.

Beispiel 8.29 Der Grenzwert $\lim_{x \rightarrow 0} x \log(x)$ ist von der Form $0 \cdot (-\infty)$. Mit $x = 1/(1/x)$ erhält man einen Ausdruck der Form $\frac{-\infty}{\infty}$, bei dem dann noch das Vorzeichen angepasst werden muss. Umformen liefert

$$\lim_{x \rightarrow 0} x \log(x) = - \lim_{x \rightarrow 0} \frac{-\log(x)}{1/x} = - \lim_{x \rightarrow 0} \frac{-1/x}{-1/x^2} = - \lim_{x \rightarrow 0} x = 0.$$

Mit der Rechenregel $a^b = \exp(\log(a^b)) = \exp(b \log(a))$ und dem letzten Resultat lässt sich auch der Grenzwert $\lim_{x \rightarrow 0} x^x$ evaluieren. Es gilt

$$\lim_{x \rightarrow 0} x^x = \lim_{x \rightarrow 0} e^{x \log(x)} = e^{\lim_{x \rightarrow 0} x \log(x)} = e^0 = 1.$$

Zur Erinnerung: Bei stetigen Funktionen darf die Grenzwertbildung mit dem Funktionsaufruf vertauscht werden. \square

Beispiel 8.30 Zu evaluieren sei $\lim_{x \rightarrow \infty} \log(x)^{1/x}$, also ein Ausdruck der Form ∞^0 . Wir verwenden den gleichen Trick wie im letzten Beispiel und formen ähnlich um:

$$\begin{aligned} \lim_{x \rightarrow \infty} \log(x)^{1/x} &= \lim_{x \rightarrow \infty} e^{1/x \log(\log(x))} = e^{\lim_{x \rightarrow \infty} \log(\log(x))/x} \\ &= e^{\lim_{x \rightarrow \infty} 1/(x \log(x))} = e^0 = 1. \end{aligned} \quad \square$$

In manchen Fällen kann man mit der Regel von de l'Hôpital sogar Ausdrücke der Form $\infty - \infty$ berechnen, wie das folgende Beispiel zeigt.

Beispiel 8.31 Zu berechnen sei $\lim_{x \rightarrow 0} \frac{1}{x} - \frac{1}{e^x - 1}$, also ein Grenzwert der Form $\infty - \infty$. Nach Umwandeln in einen Bruch kann man wie gewohnt umformen:

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{1}{x} - \frac{1}{e^x - 1} &= \lim_{x \rightarrow 0} \frac{e^x - 1 - x}{(e^x - 1)x} = \lim_{x \rightarrow 0} \frac{e^x - 1}{x e^x + e^x - 1} = \lim_{x \rightarrow 0} \frac{e^x}{2e^x + x e^x} \\ &= \lim_{x \rightarrow 0} \frac{1}{2 + x} = \frac{1}{2}. \quad \square \end{aligned}$$

Integralrechnung

Das in diesem Kapitel behandelte Problem ist das der Berechnung von Längen, Flächen und Volumina. Wir werden nach einigen einleitenden Überlegungen erkennen, dass diese Problemstellung zu einer Operation führt, die als Umkehrung der Ableitung von Funktionen betrachtet werden kann.

9.1 Das Riemann-Integral

Die elementarste Aufgabenstellung der Integralrechnung umfasst die Berechnung der Fläche unter einer Funktion in einem bestimmten Intervall. Die grundlegende Idee zur Lösung dieser Aufgabenstellung ist die Approximation der gesuchten Fläche durch immer schmalere Rechtecke. Dabei unterscheidet man *Untersummen* (wenn die Oberkanten der Rechtecke die Funktion von unten approximieren) und *Obersummen* von Rechtecken (wenn die Oberkanten die Funktion von oben approximieren). Diese beiden Situationen sind graphisch in Abbildung 9.1 dargestellt.

Wenn die Breite der Rechtecke gegen Null geht, so nähert sich sowohl die Untersumme als auch die Obersumme der Fläche unter der Funktion an. Der allgemeine Fall ist etwas schwierig zu behandeln, da der größte bzw. kleinste Wert in einem Rechteck nicht nur am Rand, sondern irgendwo in der Mitte des Rechtecks angenommen werden kann, die beiden Summen also schwierig zu bestimmen sind.

Weiters ist anzumerken, dass mit diesem Ansatz die Fläche unter der Funktion in Bereichen, in denen die Funktion negativ ist, auch negativ ist. Wenn man am Absolutbetrag der Fläche interessiert ist, muss man somit die Bereiche zwischen den Nullstellen getrennt betrachten.

Wir betrachten als Beispiel der Berechnung von Unter- und Obersummen den Spezialfall einer monoton wachsenden Funktion (bei der der kleinste Wert in einem Intervall immer am linken Rand angenommen wird, der größte Wert immer am rechten Rand).

Beispiel 9.1 Zu bestimmen sei die Fläche unter der Funktion $f(x) = x^2$ im Intervall $[0, b]$. Wir unterteilen dieses Intervall in n Teilstücke der Länge $h = b/n$. Die Untersumme ist dann

$$U_n(f, 0, b) = \sum_{k=1}^n h f(0 + (k-1)h) = \frac{b}{n} \sum_{k=1}^n \left((k-1) \frac{b}{n} \right)^2,$$

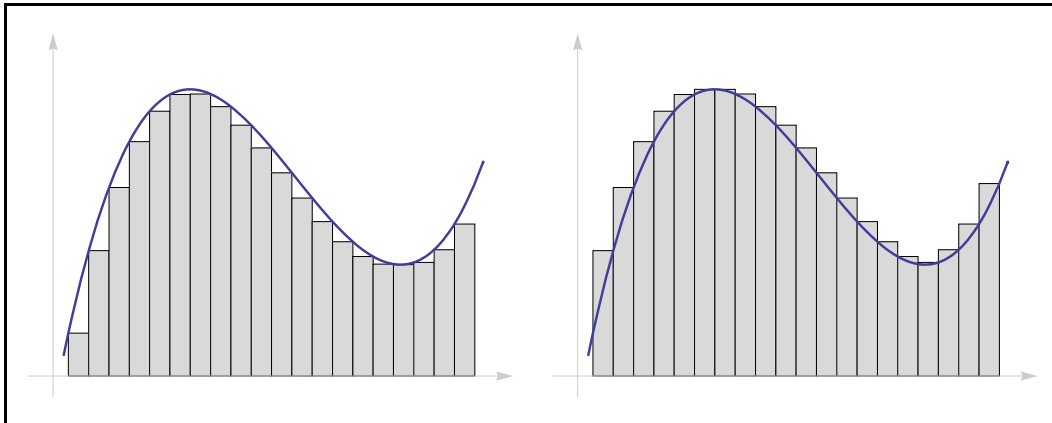


Abbildung 9.1: Approximation der Fläche unter einer Funktion durch eine Untersumme (links) bzw. Obersumme (rechts) von Rechtecksflächen.

die Obersumme

$$O_n(f, 0, b) = \sum_{k=1}^n h f(0 + kh) = \frac{b}{n} \sum_{k=1}^n \left(k \frac{b}{n}\right)^2.$$

Der Grenzwert der Untersumme $U_n(f, 0, b)$ ist unter Verwendung der Summenformel $\sum_{k=1}^n k^2 = n(n+1)(2n+1)/6$

$$\begin{aligned} \lim_{n \rightarrow \infty} U_n(f, 0, b) &= \lim_{n \rightarrow \infty} \frac{b}{n} \sum_{k=1}^n \left((k-1) \frac{b}{n}\right)^2 = \lim_{n \rightarrow \infty} \frac{b^3}{n^3} \sum_{k=1}^n (k-1)^2 \\ &= \lim_{n \rightarrow \infty} \frac{b^3}{n^3} \frac{n(n-1)(2n-1)}{6} = \frac{b^3}{6} \lim_{n \rightarrow \infty} \frac{n(n-1)(2n-1)}{n^3} \\ &= \frac{b^3}{3}; \end{aligned}$$

eine analoge Umformung ergibt für die Obersumme

$$\lim_{n \rightarrow \infty} O_n(f, 0, b) = \lim_{n \rightarrow \infty} \frac{b}{n} \sum_{k=1}^n \left(k \frac{b}{n}\right)^2 = \frac{b^3}{6} \lim_{n \rightarrow \infty} \frac{n(n+1)(2n+1)}{n^3} = \frac{b^3}{3}. \quad \square$$

Man kann zeigen, dass für bestimmte Klassen von Funktionen die Unter- und Obersummen immer gegen den gleichen Wert konvergieren; wir beschränken uns hier auf stetige Funktionen.

Definition 9.1 (Riemann-Integral)

Sei $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, die auf einem Intervall $[a, b]$ stetig ist. Dann ist $\lim_{n \rightarrow \infty} U_n(f, a, b) = \lim_{n \rightarrow \infty} O_n(f, a, b)$, und man nennt

$$\int_a^b f(x) dx := \lim_{n \rightarrow \infty} U_n(f, a, b) = \lim_{n \rightarrow \infty} O_n(f, a, b)$$

das (Riemann-)Integral bzw. das bestimmte Integral von f zwischen a und b . Die Funktion f heißt in diesem Intervall *integrierbar*.

Die Notation $\int f(x)dx$ für Integrale orientiert sich an der Herleitung mittels Unter- und Obersummen: Das Symbol dx bezeichnet einen konzeptionell unendlich kleinen Abschnitt auf der x -Achse, und $f(x)$ die dazugehörige Höhe auf der y -Achse. Durch das Produkt $f(x)dx$ ist somit der Inhalt eines (unendlich) schmalen Rechteckstreifens gegeben. Da nicht nur abzählbar unendlich viele, sondern sogar überabzählbar unendlich viele solche Streifen aufaddiert werden, wird dafür das Integral-Symbol \int als Summenzeichen verwendet.

Die Berechnung der Flächen über den Grenzwert von Unter- oder Oberfunktionen ist aufwändig. Wir werden sehen, wie man unter Verwendung von Stammfunktionen Integrale leichter evaluieren kann. Dazu benötigen wir noch einige Eigenschaften von Integralen, die sich durch Nachrechnen mit Unter- bzw. Obersummen beweisen lassen.

Satz 9.1 Seien $f, g : \mathbb{R} \rightarrow \mathbb{R}$ zwei auf einem Intervall $[a, b]$ integrierbare Funktionen, $\alpha \in \mathbb{R}$ und $c \in [a, b]$. Dann gelten die folgenden Rechenregeln

$$\begin{aligned}\int_a^b f(x) + g(x)dx &= \int_a^b f(x)dx + \int_a^b g(x)dx \\ \int_a^b \alpha f(x)dx &= \alpha \int_a^b f(x)dx \\ \int_a^c f(x)dx + \int_c^b f(x)dx &= \int_a^b f(x)dx \\ \forall_{x \in [a, b]} f(x) \leq g(x) &\Rightarrow \int_a^b f(x)dx \leq \int_a^b g(x)dx\end{aligned}$$

Die ersten beiden Gleichungen besagen, dass das Integral eine lineare Funktion auf dem Vektorraum der integrierbaren Funktionen ist.

9.2 Unbestimmte Integrale und Stammfunktionen

Als Hinleitung zum Begriff der Stammfunktionen betrachten wir den Flächeninhalt unter einer Funktion in Abhängigkeit der oberen Integrationsgrenze. Sei dazu f auf einem Intervall I stetig; wir definieren für $\alpha \in I$ eine Funktion

$$F_\alpha(x) := \int_\alpha^x f(t)dt.$$

Wie man sich mit der dritten Aussage von Satz 9.1 überlegen kann, gilt dann für $a, b \in I$ die Gleichung

$$\int_a^b f(t)dt = F_\alpha(b) - F_\alpha(a)$$

für beliebiges α . Wie wir in Beispiel 9.1 gesehen haben, ist

$$\int_0^x t^2 dt = \frac{x^3}{3},$$

und für beliebiges $\alpha > 0$ gilt

$$\int_{\alpha}^x t^2 dt = \frac{x^3}{3} - \underbrace{\int_0^{\alpha} t^2 dt}_C.$$

Der zweite Term C hängt nicht von x ab. Dieser Term kürzt sich beim Evaluieren von Integralen auf einem Intervall $[a, b]$ immer weg, sodass man unabhängig von α

$$\int_a^b t^2 dt = \frac{b^3}{3} - \frac{a^3}{3}$$

erhält. Es stellt sich nun die Frage, wie man im allgemeinen Fall zu gegebener Funktion f eine Funktion F finden kann, die die gewünschte Vereinfachung bei der Evaluation von Integralen bringt. Wie wir sehen werden, erfüllen *Stammfunktionen*, die wie folgt definiert sind, diese Rolle.

Definition 9.2 (Stammfunktion, unbestimmtes Integral)

Sei $F : \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, die auf $[a, b]$ differenzierbar ist. Dann heißt F *Stammfunktion* einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, wenn $F' = f$ gilt. Die Menge aller Stammfunktionen von f wird als *unbestimmtes Integral* $\int f$ bezeichnet.

Eine Stammfunktion ist nur bis auf einen konstanten additiven Term eindeutig, wie man sich anhand von Beispielen leicht überlegen kann. Die beiden Begriffe “Stammfunktion” und “unbestimmtes Integral” werden daher auch oft synonym verwendet.

Beispiel 9.2 Die Funktionen $\int x^2 dx = \frac{x^3}{3} + C$ sind für beliebiges $C \in \mathbb{R}$ Stammfunktionen der Funktion x^2 . Die Funktionen $\int e^x dx = e^x + C$ sind Stammfunktionen von e^x . \square

Oft bezeichnet man die Stammfunktion $F + C$ mit $C = 0$ als *die* Stammfunktion von f . Stammfunktionen sind nun genau die Funktionen, die man für die Evaluierung von Integralen benötigt. Bevor wir dies beweisen können, benötigen wir einen weiteren Mittelwertsatz.

Satz 9.2 (Mittelwertsatz der Integralrechnung) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine auf dem Intervall $[a, b]$ stetige Funktion. Dann gibt es ein $c \in [a, b]$ mit

$$\int_a^b f(x) dx = f(c)(b - a).$$

Dieser Mittelwert besagt also, dass die Fläche unter einer Funktion gleich dem Flächeninhalt eines Rechtecks ist, dessen Höhe der Funktionswert an einer Stelle c im Integrationsbereich ist. Graphisch ist diese Situation in Abbildung 9.2 dargestellt.

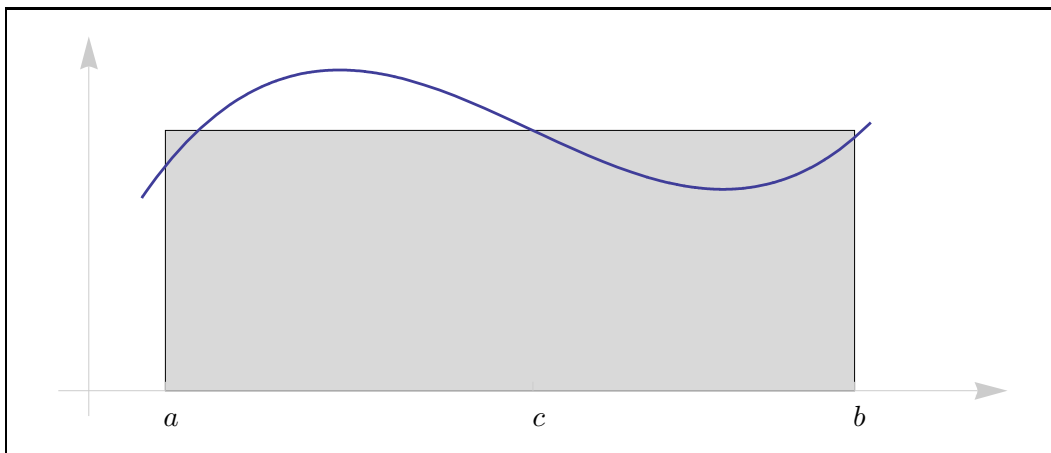


Abbildung 9.2: Illustration zum Mittelwertsatz der Integralrechnung.

Da die Begründung dieses Satzes ziemlich einfach ist, wollen wir sie hier nachvollziehen. Weil f auf $[a, b]$ stetig ist, gibt es einen kleinsten und größten Wert, den f in diesem Intervall annimmt:

$$\alpha = \min_{x \in [a, b]} f(x), \quad \beta = \max_{x \in [a, b]} f(x).$$

Mit dem Zwischenwertsatz (Satz 7.7) nimmt f auch *jeden* Wert zwischen α und β an. Wenn wir die Ungleichung $\alpha \leq f(x) \leq \beta$ integrieren, so folgt wegen $\int_a^b 1 dx = (b - a)$

$$\alpha(b - a) \leq \int_a^b f(x) dx \leq \beta(b - a).$$

Da f jeden Wert zwischen α und β annimmt (speziell also auch den Wert $\frac{1}{b-a} \int_a^b f(x) dx$), gibt es zumindest einen Punkt $c \in [a, b]$ mit der gewünschten Eigenschaft

$$f(c)(b - a) = \int_a^b f(x) dx.$$

Satz 9.3 (Hauptsatz der Differential- und Integralrechnung) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine auf $[a, b]$ stetige Funktion. Dann ist die Funktion

$$F_\alpha(x) := \int_\alpha^x f(t) dt$$

für beliebiges $\alpha \in [a, b]$ eine Stammfunktion von f , und es gilt

$$\int_a^b f(x) dx = F(b) - F(a)$$

für beliebige Stammfunktion F .

Zur Begründung dieses Satzes müssen wir also nachrechnen, dass F_α eine Stamm-

funktion von f ist, dass also $F'_\alpha = f$ gilt. Mit der Definition der Ableitung ist

$$\begin{aligned} F'_\alpha(x) &= \lim_{h \rightarrow 0} \frac{F_\alpha(x+h) - F_\alpha(x)}{h} = \lim_{h \rightarrow 0} \frac{\int_\alpha^{x+h} f(t) dt - \int_\alpha^x f(t) dt}{h} \\ &= \lim_{h \rightarrow 0} \frac{\int_x^{x+h} f(t) dt}{h} \stackrel{(*)}{=} \lim_{h \rightarrow 0} \frac{hf(c_{[x, x+h]})}{h} = f(x). \end{aligned}$$

Bei der mit (*) markierten Umformung wird Satz 9.2 angewandt: Es gibt ein $c_{[x, x+h]}$ im Intervall $[x, x+h]$ mit der geforderten Eigenschaft. Für $h \rightarrow 0$ wird das Intervall zum Punkt x , woraus die Behauptung des Hauptsatzes folgt.

Die zweite Aussage des Hauptsatzes ist einfacher: Für beliebiges α ist F_α eine Stammfunktion von f , und es gilt

$$F_\alpha(b) - F_\alpha(a) = \int_\alpha^b f(x) dx - \int_\alpha^a f(x) dx = \int_a^b f(x) dx,$$

wie gefordert.

Zur Notation sei angemerkt, dass das Berechnen eines Integrals mit Integrationsgrenzen a und b und Stammfunktion F oftmals kurz als

$$F \Big|_a^b := F(b) - F(a)$$

geschrieben wird.

Wir können nun eine Liste von Stammfunktionen angeben; die Richtigkeit dieser Rechenregeln kann sofort durch Differenzieren überprüft werden.

Satz 9.4 Es gelten die folgenden Regeln zum Bestimmen von Stammfunktionen, wobei $a, C \in \mathbb{R}$ Konstante sind:

$$\begin{aligned} \int a \, dx &= ax + C & \int x^a \, dx &= \frac{x^{a+1}}{a+1} + C, \quad a \neq -1 \\ \int e^x \, dx &= e^x + C & \int \frac{1}{x} \, dx &= \log(x) + C \\ \int \sin(x) \, dx &= -\cos(x) + C & \int \cos(x) \, dx &= \sin(x) + C. \end{aligned}$$

Da das Integrieren also die zum Differenzieren inverse Operation ist, können weitere Rechenregeln für das Integrieren durch das Invertieren der Rechenregeln für das Differenzieren angegeben werden. Im Speziellen werden wir die Umkehrungen der in Satz 8.2 angegebenen Produkt- und Kettenregeln für Ableitungen betrachten.

Satz 9.5 Seien $f, g : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ differenzierbare Funktionen. Dann nennt man die Integrationsregel

$$\int_a^b f(x)g'(x) dx = f(x)g(x) \Big|_a^b - \int_a^b f'(x)g(x) dx$$

partielles Integrieren, und die Integrationsregel

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(t)dt \quad \text{mit } t = g(x) \text{ und } dt/dx = g'(x)$$

die *Substitutionsregel*.

Beispiel 9.3 Wir betrachten zwei Beispiele zum partiellen Integrieren, wobei wir uns auf das Bestimmen von Stammfunktionen beschränken. Es gilt mit $f(x) = x$ und $g'(x) = \sin(x)$:

$$\int x \sin(x)dx = -x \cos(x) - \int -\cos(x)dx = -x \cos(x) + \sin(x) + C.$$

Mit den Teilen $f(x) = \sin(x)$ und $g'(x) = \cos(x)$ gilt

$$\int \sin(x) \cos(x)dx = \sin^2(x) - \int \cos(x) \sin(x)dx,$$

woraus durch Auflösen dieser Gleichung folgt:

$$\int \sin(x) \cos(x)dx = \frac{1}{2} \sin^2(x) + C. \quad \square$$

Beispiel 9.4 Mit partieller Integration kann auch eine Stammfunktion für $\log(x)$ gefunden werden. Wir schreiben $\log(x) = \log(x) \cdot 1$ und verwenden $f(x) = \log(x)$ und $g'(x) = 1$. Damit ist

$$\int \log(x)dx = \int \log(x) \cdot 1dx = \log(x)x - \int \frac{1}{x}x dx = x(\log(x) - 1) + C. \quad \square$$

Das Anwenden der Substitutionsregel ist meist nur dann möglich, wenn die innere Ableitung einer zu integrierenden Funktion (bis auf einen konstanten Faktor) ebenfalls im Integranden vorkommt. Zu beachten ist auch, dass die Integralgrenzen durch die Substitution angepasst werden müssen. Wir betrachten auch hier einfache Beispiele.

Beispiel 9.5 Zu bestimmen sei das Integral $\int_2^5 x \sin(x^2)dx$. Mit der Substitution $t = x^2$ ist $t' = dt/dx = 2x$ bzw. $dx = dt/2x$, und es gilt

$$\int_2^5 x \sin(x^2)dx = \int_{2^2}^{5^2} x \sin(t) \frac{dt}{2x} = \frac{1}{2} \int_4^{25} \sin(t)dt = -\frac{1}{2} \cos(x) \Big|_4^{25} = -0.8224.$$

Das Integral $\int \frac{\log(x)}{x}dx$ ist ebenfalls über die Substitutionsregel lösbar, obwohl man hier nicht nach dem Schema vorgehen kann, t auf die innere Ableitung zu setzen. Wir substituieren stattdessen $t = \log(x)$, woraus sich $dt/dx = \frac{1}{x}$ und $dx = x dt$ ergibt. Man erhält

$$\int \frac{\log(x)}{x}dx = \int \frac{t}{x}x dt = \frac{t^2}{2} + C = \frac{1}{2} \log^2(x) + C. \quad \square$$

Bei der Integration von Funktionen kann man sich überlegen, dass es durch die Rechenregeln für Ableitungen in Kapitel 8 für jede *noch so komplexe Funktion* möglich ist, die Ableitung dieser Funktion zu berechnen. Wie man aber aus den Sätzen in diesem Kapitel erkennen kann, müssen für das Integrieren die Funktionen in speziellen Formen vorliegen. So können für eine Reihe einfacher Funktionen, wie etwa

$$e^{-x^2}, \quad \sin(x^2), \quad \sin(\cos(x)), \quad e^x \log(x), \quad \dots$$

keine Stammfunktionen bestimmt werden, da diese Funktionen nicht den gültigen Schemata entsprechen. Die Auswertung dieser Integrale ist somit nur über numerische Approximationen möglich. Eine mögliche numerische Approximation ist über die Taylorreihenentwicklung, da Taylorreihen innerhalb ihres Konvergenzradius gliedweise integrierbar sind.

Beispiel 9.6 Zu bestimmen sei $\int_0^2 e^{-x^2} dx$. Da der Konvergenzradius von e^x unendlich ist, kann man für beliebigen Integrationsbereich den Integranden durch seine Taylorreihenentwicklung ersetzen. Es ist

$$e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots,$$

somit

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \dots,$$

und

$$\int_0^2 e^{-x^2} dx = x \Big|_0^2 - \frac{x^3}{3} \Big|_0^2 + \frac{x^5}{5 \cdot 2!} \Big|_0^2 - \frac{x^7}{7 \cdot 3!} \Big|_0^2 + \dots + C. \quad (9.1)$$

Die ersten vier Terme liefern den numerischen Wert -0.5143 , die ersten sieben Terme 1.1798 , und erst die ersten 17 Summanden liefern den auf fünf Stellen genauen Wert 0.88208 . \square

9.3 Bogenlänge von Kurven

Als Anwendung der Integralrechnung betrachten wir die Problemstellung, von gegebener Funktion die Länge des Graphen dieser Funktion zu berechnen. Mit Hilfe einer Skizze (wie in Abbildung 9.3) kann man sich überlegen, dass man eine Kurve durch viele kleine Tangenten an die Kurve approximieren kann. Wenn die Länge jeder einzelnen Tangente gegen Null geht, konvergiert die Summe dieser Tangenten gegen die gesuchte Bogenlänge. Da das Streckenstück der Tangente die Hypotenuse eines rechtwinkligen Dreiecks darstellt, dessen Katheten dx und $f'(x)dx$ sind, ist die Länge dieses Tangentenstücks gerade

$$\sqrt{dx^2 + (f'(x)dx)^2} = \sqrt{1 + (f'(x))^2} dx.$$

Das Aufsummieren über diese Teilstücke liefert das Ergebnis, das im folgenden Satz zusammengefasst ist.

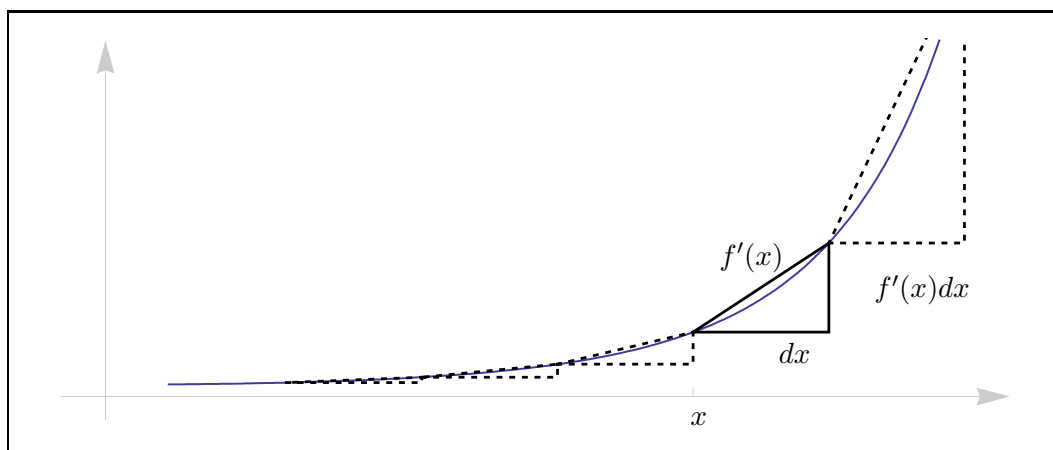


Abbildung 9.3: Illustration zum Berechnung der Bogenlänge einer Funktion.

Satz 9.6 Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ auf $[a, b]$ differenzierbar. Dann ist die Länge des Graphen von f (genannt *Bogenlänge*) zwischen a und b gegeben durch

$$\int_a^b \sqrt{1 + (f'(x))^2} dx.$$

Obwohl diese Formel für alle Funktionen die Länge des Graphen angibt, erhält man in nur wenigen Fällen Integranden, die eine einfache Stammfunktion haben. So erhält man die Länge des Graphen der Sinus-Funktion zwischen 0 und 2π mit obiger Formel als

$$\int_0^{2\pi} \sqrt{1 + \cos^2(x)} dx;$$

dieser Ausdruck ist aber nicht direkt über eine Stammfunktion evaluierbar (numerische Integration ergibt ein Resultat von 7.6404).

Beispiel 9.7 Aus der Kreisgleichung $x^2 + y^2 = 1$ für den Einheitskreis folgt, dass die Funktion $f(x) = \sqrt{1 - x^2}$ im Definitionsbereich $[-1, 1]$ den oberen Einheits-Halbkreis beschreibt. Wir rechnen nun nach, dass die Länge dieser Kurve wie erwartet π ist.

Für $f(x) = \sqrt{1 - x^2}$ ist $f'(x) = -x(1 - x^2)^{-1/2}$; Einsetzen in die Formel aus Satz 9.6 liefert

$$\int_{-1}^1 \sqrt{1 + \frac{x^2}{1 - x^2}} dx = \int_{-1}^1 \sqrt{\frac{1}{1 - x^2}} dx.$$

Mit der Substitution $\sin(t) = x$ (motiviert durch $\cos^2(t) = 1 - \sin^2(t)$) ist $dx/dt = \cos(t)$. Die angepassten Integrationsgrenzen müssen $\sin(t) = \pm 1$ erfüllen, was bei $\pm \frac{\pi}{2}$ der Fall ist. Somit erhält man

$$\int_{-1}^1 \sqrt{\frac{1}{1 - x^2}} dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{\frac{1}{1 - \sin^2(t)}} \cos(t) dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{\cos(t)} \cos(t) dt = \pi. \quad \square$$

Mit der Formel aus Satz 9.6 kann man nur die Länge von *Funktionen* berechnen, nicht aber die von *beliebigen Kurven*. Ein Beispiel einer solchen beliebigen Kurve ist etwa durch den Einheitskreis gegeben, der nicht durch eine Funktion beschrieben werden kann, weil jeder x -Wert (außer ± 1) genau zwei y -Werte hat. Diese Situation kann über parametrische Kurven ausgedrückt werden, bei denen x - und y -Werte unabhängig voneinander von einer dritten Größe abhängen.

Definition 9.3 (Parametrische Kurven)

Sei $[a, b] \subseteq \mathbb{R}$ ein Intervall und $x, y : [a, b] \rightarrow \mathbb{R}$ zwei reelle Funktionen. Dann bezeichnet man die vektorwertige Funktion $[a, b] \rightarrow \mathbb{R}^2$ mit

$$t \mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

als *parametrische Kurve*.

Obige Definition kann auch auf dreidimensionale Kurven erweitert werden, wenn man noch eine dritte Funktion $z(t)$ zulässt. Wir werden uns hier aber auf den zweidimensionalen Fall beschränken.

Beispiel 9.8 Der Einheitskreis lässt sich mit $x(t) = \cos(t)$, $y(t) = \sin(t)$ für $t \in [0, 2\pi]$ als parametrische Kurve darstellen. Daraus ergibt sich eine Ellipse, indem man die x -Koordinate um den Faktor a , die y -Koordinate um den Faktor b streckt, und so die Parametrisierung $(a \cos(t), b \sin(t))$ erhält. \square

Wie man sich in Verallgemeinerung von Abbildung 9.3 überlegen kann, ist für parametrische Kurven die Bogenlänge gegeben durch

$$\int_a^b \sqrt{(x'(t))^2 + (y'(t))^2} dx.$$

Beispiel 9.9 Der Kreisumfang ist mit der Parametrisierung aus Beispiel 9.8 gegeben durch

$$\int_0^{2\pi} \sqrt{\sin^2(t) + \cos^2(t)} dt = \int_0^{2\pi} 1 dt = 2\pi.$$

Ebenso gilt für den Umfang der Ellipse

$$\int_0^{2\pi} \sqrt{a^2 \sin^2(t) + b^2 \cos^2(t)} dt;$$

diese Funktion lässt sich aber nicht in geschlossener Form integrieren. \square

Beispiel 9.10 Eine Spirale ist nichts anderes als ein Kreis, bei dem der Radius immer größer bzw. kleiner wird. Eine mögliche Parametrisierung einer Spirale ist die *logarithmische Spirale*

$$t \mapsto \begin{pmatrix} \alpha e^{\beta t} \cos(t) \\ \alpha e^{\beta t} \sin(t) \end{pmatrix}.$$

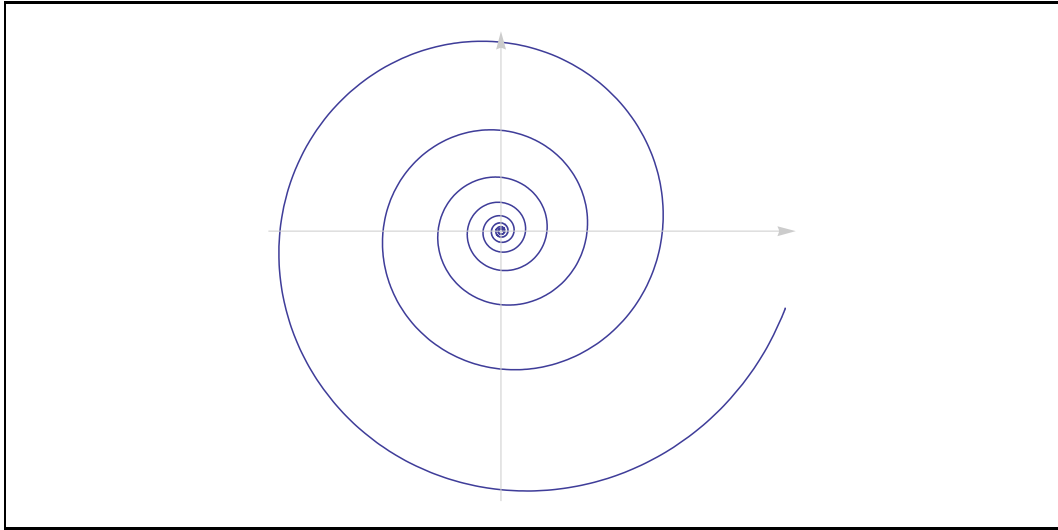


Abbildung 9.4: Logarithmische Spirale aus Beispiel 9.10.

Bei dieser Spirale bildet die Folge der Radien (pro Windung) eine geometrische Folge. Logarithmische Spiralen haben dadurch die fraktale Eigenschaft, bei jeder Vergrößerung gleich auszusehen. Viele Wachstumsformen in der Natur (Muscheln, Pflanzen,...) zeigen dieses logarithmische Spiralmuster.

Für die Spirale in Abbildung 9.4 wurden $\alpha = 1$ und $\beta = 0.1$ gewählt, der Parametrisierungsbereich war $[0, 20\pi]$. Wegen

$$\begin{aligned} (x'(t))^2 + (y'(t))^2 &= \left(\alpha \beta e^{\beta t} \cos(t) - \alpha e^{\beta t} \sin(t)\right)^2 + \\ &\quad \left(\alpha e^{\beta t} \cos(t) + \alpha \beta e^{\beta t} \sin(t)\right)^2 \\ &= \alpha^2(1 + \beta^2)e^{2\beta t} \end{aligned}$$

ist die Länge dieser Spirale

$$\int_a^b \sqrt{\alpha^2(1 + \beta^2)e^{2\beta t}} dx = \alpha \sqrt{1 + \beta^2} \int_a^b e^{\beta t} dt = \frac{\alpha}{\beta} \sqrt{1 + \beta^2} (e^{b\beta} - e^{a\beta}).$$

Für die konkreten Werte $a = 0$, $b = 20\pi$, $\alpha = 1$ und $\beta = 0.1$ erhält man eine Bogenlänge von 5371.57. \square