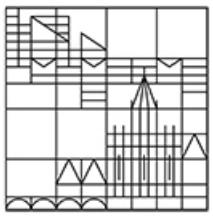
Fachbereich Informatik und
Informationswissenschaft

Universität
Konstanz



Skriptum

zur Vorlesung

Mathematische Grundlagen der Informatik

gehalten in WS 2010/11, WS 2011/12

von

Sven Kosub

7. Februar 2012

Version v1.25

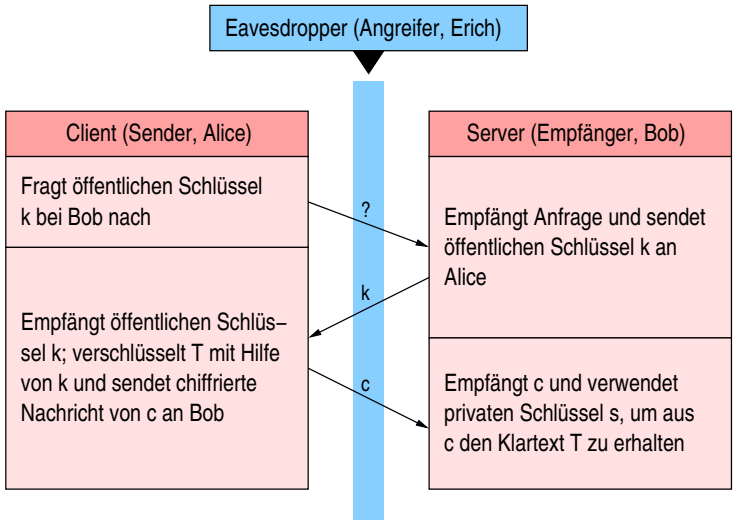
Inhaltsverzeichnis

Prolog

Wir wollen an einem Beispiel aus der Kryptographie für die Informatik typische mathematische Methoden erläutern. Die systematische Einführung erfolgt in den nachfolgenden Kapiteln.

In der Kryptographie unterscheidet man zwischen symmetrischen und asymmetrischen Verschlüsselungsverfahren. Im Gegensatz zu den symmetrischen Verschlüsselungsverfahren, bei denen zur Verschlüsselung und Entschlüsselung geheime (private) Schlüssel verwendet werden, erfolgt bei einem asymmetrischen Verfahren die Verschlüsselung mit einem öffentlich bekannten Schlüssel. Nur für die Entschlüsselung wird ein privater Schlüssel verwendet.

Ein wichtiges asymmetrisches Verschlüsselungsverfahren ist das Diffie-Hellman-Protokoll. Hierbei möchte Alice einen Klartext T sicher vor Erich, der T natürlich erfahren möchte, an Bob schicken. Dazu verfügt Bob über einen öffentlichen Schlüssel k sowie einen privaten Schlüssel s. Die Kommunikation erfolgt dann wie in folgendem Szenario skizziert:



Das Diffie-Hellman-Protokoll ist noch keine vollständige Beschreibung eines Protokolls. Vielmehr ist noch gar nicht sicher, dass sich das Verfahren tatsächlich implementieren lässt. Diese Frage wird durch das berühmte RSA-Verfahren beantwortet, dessen Umsetzung wir sehr stark vereinfacht kurz darstellen:

- öffentlicher Schlüssel ist das Produkt $k = p \cdot q$ zweier großer Primzahlen p und q;

- privater Schlüssel ist das Paar $s = (p, q)$ der Primzahlen, d.h. die Primzahlenzerlegung von k;
- Verschlüsselung von T: Wandle T in eine Zahl t (zum Beispiel unter Verwendung der ASCII-Codes), oder in eine Folge von Zahlen um, so dass für alle Zahlen $t < k$ gilt; setze $c =_{\text{def}} \text{mod}(t^3, k)$;
- Entschlüsselung von c erfolgt mit Hilfe von $s = (p, q)$, die ohne Kenntnis von p und q genauso schwierig ist, wie k in seine Primfaktoren p und q zu zerlegen.

Die Anschauung hinter dem RSA-Verfahren ist wie folgt: Ist $p \cdot q$ klein, dann ist die Zerlegung in p und q einfach, z.B. $111 = 3 \cdot 37$. Für große Primzahl ist es dagegen schwierig auf die entsprechenden Primfaktoren zu kommen. Um einen Eindruck von der Schwierigkeit zu bekommen, bestimme man die beiden Primfaktoren p und q in dem folgenden Produkt:

$$pq = 37852153254637693623290549498896720462797948158601\backslash \\ 27761136816982888921764999850721920649197641542929$$

Für die Sicherheit des RSA-Verfahrens ist eine notwendige Voraussetzung, dass es unendlich viele Primzahlen gibt. Anderenfalls könnten (theoretisch) alle Produkte zweier Primzahlen in einer Datenbank gesammelt und somit aus allen öffentlichen Schlüsseln die privaten bestimmt werden.

Im Folgenden wollen wir uns davon überzeugen, dass es tatsächlich unendlich viele Primzahlen gibt.

Definition 0.1 *Es seien p und q ganze Zahlen.*

1. *Die Zahl p* teilt *q (symbolisch: $p|q$), falls es eine ganze Zahl k gibt mit $q = k \cdot p$.*
2. *Die Zahl p heißt* Primzahl, *falls $p \geq 2$ gilt und unter den natürlichen Zahlen nur 1 und p die Zahl p teilen.*

Die in nachfolgendem Lemma verwendete Methode der Induktion ist zentral für die Informatik und wird in einem eigenen Kapitel ausführlich behandelt werden.

Lemma 0.2 *Zu jeder natürlichen Zahl $n \geq 2$ existiert eine Primzahl p, die n teilt.*

Beweis: (*Induktion*) Wir beweisen das Lemma mittels vollständiger Induktion über n.

- (IA, *Induktionsanfang*): Für $n = 2$ gilt die Aussage mit $p = n$.
- (IS, *Induktionsschritt*): Es sei $n > 2$ eine beliebige natürliche Zahl. Angenommen wir hätten die Aussage bereits für alle $2 \leq k < n$ bewiesen (IV, *Induktionsvoraussetzung*). Wir unterscheiden zwei Fälle für n:
  1. Ist n eine Primzahl, so gilt die Aussage für $p = n$.

2. Ist n keine Primzahl, so gibt es natürliche Zahlen k, ℓ mit $n = k \cdot \ell$ und $2 \leq k, \ell < n$. Nach Induktionsvoraussetzung gibt es somit eine Primzahl p, die k teilt, d.h. $k = p \cdot r$ für ein geeignetes r. Also gilt $n = k \cdot \ell = p \cdot (r \cdot \ell)$. Mithin teilt p auch n.

Damit ist das Lemma bewiesen. ■

Mit Hilfe von Lemma 0.2 kann nun bewiesen werden, dass es unendlich viele Primzahlen gibt. Dazu verwenden wir ein zweites wichtiges Beweisprinzip – den Widerspruchsbeweis.

Theorem 0.3 (Euklid) *Es gibt unendlich viele Primzahlen.*

Beweis: (*Widerspruch*) Angenommen die Aussage ist falsch, d.h., es gibt nur endlich viele Primzahlen $2 \leq p_1 < p_2 < \cdots < p_k$. Wir definieren die Zahl

$$n =_{\text{def}} 1 + \prod_{j=1}^{k} p_j.$$

Wegen $n \geq 2$ folgt aus Lemma 0.2, dass eine Primzahl p_ℓ mit $1 \leq \ell \leq k$ existiert, die n teilt. Auf der anderen Seite gilt jedoch $\text{mod}(n, p_\ell) = 1$. Dies ist ein Widerspruch. Somit ist die Annahme falsch und es gibt unendlich viele Primzahlen. Damit ist das Theorem bewiesen. ■

Logik 1

1.1 Aussagen

1.1.1 Begriff und Beispiele

Definition 1.1 *Eine (mathematische)* Aussage *ist ein sprachlicher Ausdruck (Satz), dem eindeutig einer der Wahrheitswerte* w *(für „wahr“) oder* f *(für „falsch“) zugeordnet werden kann.*

Wir werden Aussagen mit großen Buchstaben bezeichnen und wie folgt beschreiben:

$$X =_{\text{def}} \text{ Beschreibung}$$

Beispiel: Die folgenden Beispiele verdeutlichen die obige Begriffsbildung:

- $A =_{\text{def}}$ „Zu jeder natürlichen Zahl gibt es eine Primzahl, die größer ist“ ist eine wahre Aussage.
- $B =_{\text{def}}$ „Zu jeder natürlichen Zahl gibt es eine Primzahl, die kleiner ist“ ist eine falsche Aussage, da die Zahl 2 ein Gegenbeispiel ist.
- $C =_{\text{def}}$ „Jede gerade Zahl, die größer als 2 ist, ist die Summe zweier Primzahlen“ ist eine Aussage, da der Satz entweder gültig oder nicht gültig ist. Der Wahrheitswert ist noch offen; bei der Aussage handelt es sich um die bekannte Goldbach'sche Vermutung.
- $D =_{\text{def}}$ „Diese Aussage ist falsch“ ist keine Aussage, da kein Wahrheitswert zugeordnet werden kann: Ist D wahr, dann ist D falsch; ist D falsch, dann ist D wahr.

1.1.2 Logische Verknüpfungen

Aussagen können mittels logischer Operationen verknüpft werden. Dabei entstehen wieder Aussagen. Unverknüpfte Aussagen heißen *Elementaraussagen* (oder aussagenlogische Variablen). Verknüpfte Aussagen heißen *zusammengesetzte Aussagen.*

Die wichtigsten logischen Verknüpfungen mit ihren Sprech- und Leseweisen sind wie folgt:

$\neg A$	steht für:	nicht A	(Negation)
$A \wedge B$	steht für:	A und B	(Konjunktion)
$A \vee B$	steht für:	A oder B	(Disjunktion)
$A \rightarrow B$	steht für:	wenn A, dann B	(Implikation)
$A \leftrightarrow B$	steht für:	genau dann A, wenn B	(Äquivalenz)
$A \oplus B$	steht für:	entweder A oder B	(Antivalenz)

Neben $\rightarrow$ und $\leftrightarrow$ werden auch $\Rightarrow$ und $\Leftrightarrow$ für Implikation und Äquivalenz verwendet, wenn wir Aussagen über Aussagen formulieren.

Ähnlich der Addition und Multiplikation („Punktrechnung geht vor Strichrechnung“) gibt es Bindungsregeln bei der Verwendung der logischen Verknüpfungen, um die Klammerungen in zusammengesetzten Ausdrücken wegzulassen. Für die gebräuchlichsten Verknüpfungen $\neg, \wedge$ und $\vee$ vereinbaren wir: „$\neg$ geht vor $\wedge$“ und „$\wedge$ geht vor $\vee$“.

Beispiel: $\neg A \wedge B \vee C$ ist die gleiche Aussage wie $((\neg A) \wedge B) \vee C$.

Um Missverständnissen in komplizierteren Zusammenhängen vorzubeugen, werden wir jedoch auch weiterhin Klammern setzen, wo sie eigentlich nach den Bindungsregeln nicht notwendig wären.

Die Wahrheitswerte der durch logische Verknüpfungen entstandenen zusammengesetzten Aussagen werden durch Wertetabellen definiert.

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftrightarrow B$	$A \oplus B$	–
f	f	w	f	f	w	w	f	w
f	w	w	f	w	w	f	w	w
w	f	f	f	w	f	f	w	w
w	w	f	w	w	w	w	f	f
Funktionsname		NOT	AND	OR	–	–	XOR	NAND

Bei digitalen Schaltungen entsprechen diese Wertetabellen den *booleschen Funktionen*, wobei w mit 1 und f mit 0 identifiziert wird. Die Namen der den Verknüpfungen zugehörigen booleschen Funktionen sind in der untersten Zeile angegeben.

1.1.3 Rechnen mit logischen Verknüpfungen

Definition 1.2 *Zwei Aussagen A und B heißen genau dann (logisch)* äquivalent*, symbolisch $A \equiv B$, wenn $A \leftrightarrow B$ eine wahre Aussage ist, d.h.*

$$A \equiv B \Longleftrightarrow_{\text{def}} \; A \leftrightarrow B \text{ ist wahr.}$$

Logisch äquivalente Aussagen können in zusammengesetzten Aussagen beliebig gegeneinander ausgetauscht werden. Die wichtigsten logischen Äquivalenzen sind in folgendem Theorem zusammengefasst.

Theorem 1.3 *Es seien A, B und C beliebige Aussagen. Dann gilt:*

$$\left.\begin{array}{lcl}(A \wedge B) \wedge C & \equiv & A \wedge (B \wedge C) \\ (A \vee B) \vee C & \equiv & A \vee (B \vee C)\end{array}\right\} \textit{Assoziativgesetze}$$

$$\left.\begin{array}{lcl}A \wedge B & \equiv & B \wedge A \\ A \vee B & \equiv & B \vee A\end{array}\right\} \textit{Kommutativgesetze}$$

$$\left.\begin{array}{lcl}\neg(A \wedge B) & \equiv & (\neg A) \vee (\neg B) \\ \neg(A \vee B) & \equiv & (\neg A) \wedge (\neg B)\end{array}\right\} \text{De Morgan}\textit{'sche Regeln}$$

$$\left.\begin{array}{lcl}(A \wedge B) \vee C & \equiv & (A \vee C) \wedge (B \vee C) \\ (A \vee B) \wedge C & \equiv & (A \wedge C) \vee (B \wedge C)\end{array}\right\} \textit{Distributivgesetze}$$

$$\left.\begin{array}{lcl}A \wedge (\neg A) & \equiv & \mathsf{f} \\ A \vee (\neg A) & \equiv & \mathsf{w}\end{array}\right\} \text{tertium non datur}$$

$$\left.\begin{array}{lcl}A \vee \mathsf{w} & \equiv & \mathsf{w} \\ A \vee \mathsf{f} & \equiv & A \\ A \wedge \mathsf{w} & \equiv & A \\ A \wedge \mathsf{f} & \equiv & \mathsf{f}\end{array}\right\} \textit{Dominanzgesetze}$$

$$\begin{array}{lcll}A \to B & \equiv & (\neg A) \vee B & \textit{Alternative Darstellung der Implikation} \\ & \equiv & (\neg B) \to (\neg A) & \textit{Kontraposition}\end{array}$$

$$A \leftrightarrow B \equiv (A \to B) \wedge (B \to A) \qquad \textit{Alternative Darstellung der Äquivalenz}$$

$$\neg(\neg A) \equiv A \qquad \textit{Doppelte Negation}$$

Beweis: Wir beweisen nur die erste De Morgan'sche Regel $\neg(A \wedge B) \equiv (\neg A) \vee (\neg B)$. Dazu definieren wir zunächst die Hilfsaussagen $H_1 =_{\text{def}} \neg(A \wedge B)$ und $H_2 =_{\text{def}} (\neg A) \vee (\neg B)$. Die Überprüfung der Aussage $H_1 \leftrightarrow H_2$ erfolgt mittels einer Wertetabelle:

A	B	$A \wedge B$	H_1	$\neg A$	$\neg B$	H_2	$H_1 \leftrightarrow H_2$
f	f	f	w	w	w	w	w
f	w	f	w	w	f	w	w
w	f	f	w	f	w	w	w
w	w	w	f	f	f	f	w

Somit ist $H_1 \leftrightarrow H_2$ eine wahre Aussage. Also sind H_1 und H_2 logisch äquivalent. Alle anderen logischen Äquivalenzen können ebenfalls mittels Berechnung der Wertetabellen gezeigt werden. Damit ist das Theorem bewiesen. ■

Mit Hilfe von Theorem 1.3 können Aussagen umgeformt, genauso wie es von der algebraischen Umformung von Gleichungen her bekannt ist.

Beispiel: Zur Demonstration der Anwendung von Theorem 1.3 wollen wir die Aussage

$$C =_{\text{def}} (A \wedge (A \rightarrow B)) \rightarrow B$$

vereinfachen. Wir formen die Aussage wie folgt logisch äquivalent um:

$$\begin{aligned}
C &\equiv (A \wedge (\neg A \vee B)) \rightarrow B && \text{(AD Implikation)}\\
&\equiv ((A \wedge \neg A) \vee (A \wedge B)) \rightarrow B && \text{(Distributivgesetz)}\\
&\equiv (\mathsf{f} \vee (A \wedge B)) \rightarrow B && (\textit{tertium non datur})\\
&\equiv (A \wedge B) \rightarrow B && \text{(Dominanzgesetz)}\\
&\equiv \neg(A \wedge B) \vee B && \text{(AD Implikation)}\\
&\equiv (\neg A \vee \neg B) \vee B && \text{(De Morgan)}\\
&\equiv \neg A \vee (\neg B \vee B) && \text{(Assoziativgesetz)}\\
&\equiv \neg A \vee \mathsf{w} && (\textit{tertium non datur})\\
&\equiv \mathsf{w} && \text{(Dominanzgesetz)}
\end{aligned}$$

Die Aussage C ist also stets wahr unabhängig von den Wahrheitswerten der Aussagen A und B.

1.1.4 Erfüllbarkeit von Aussagen

Es sei A eine zusammengesetzte Aussage mit den aussagenlogischen Variablen $X_1, \ldots, X_n$. Eine *Belegung* I von A ordnet jeder Variablen X_i einen Wahrheitswert $I(X_i)$ zu.

Definition 1.4 *Es sei A eine Aussage.*

1. *A heißt genau dann* erfüllbar*, wenn A wahr für eine mögliche Belegung ist.*
2. *A heißt genau dann* allgemeingültig *(oder* Tautologie*), wenn A wahr für alle möglichen Belegungen ist.*
3. *A heißt genau dann* widerlegbar*, wenn A falsch für eine mögliche Belegung ist.*
4. *A heißt genau dann* unerfüllbar *(oder* Kontradiktion*), wenn A falsch für alle möglichen Wahrheitswerte der Elementaraussagen ist.*

Proposition 1.5 *Es sei A eine Aussage.*

1. *Ist A allgemeingültig, so ist A erfüllbar.*
2. *A ist genau dann erfüllbar, wenn $\neg A$ widerlegbar ist.*
3. *A ist genau dann allgemeingültig, wenn $\neg A$ unerfüllbar.*

Beweis: Die erste Aussage ist offensichtlich. Die zweite und die dritte Aussage folgen aus der Tatsache, dass eine Belegung I, die die Aussage A wahr macht, die Aussage $\neg A$ falsch macht. ■

Beispiele: Folgende Aussagen verdeutlichen die Begriffsbildung:

- Die Aussage $A \vee B$ ist erfüllbar und widerlegbar.
- Die Aussage $((A \to B) \to A) \wedge \neg A$ ist unerfüllbar.
- Die Aussage $(A \wedge (A \to B)) \to B$ ist allgemeingültig.

1.1.5 Normalformen

Im Folgenden betrachten wir Aussagen besonderer Struktur mit den Konnektoren $\neg, \wedge, \vee$. Wir führen zunächst zwei Abkürzungen ein. Für Aussagen $H_1, \ldots, H_n$ definieren wir:

$$\bigwedge_{i=1}^{n} H_i \quad =_{\text{def}} \quad H_1 \wedge H_2 \wedge \ldots \wedge H_n$$

$$\bigvee_{i=1}^{n} H_i \quad =_{\text{def}} \quad H_1 \vee H_2 \vee \ldots \vee H_n$$

Ein *Literal* ist eine Aussage der Form X oder $\neg X$, wobei X eine Aussagenlogische Variable ist.

Definition 1.6 *Eine Aussage A mit den aussagenlogischen Variablen $X_1, \ldots, X_n$ heißt*

1. konjunktive Normalform (KNF, CNF)*, falls für geeignete Zahlen k und ℓ_i sowie Literale L_{ij} gilt:*
$$A = \bigwedge_{i=1}^{k} \bigvee_{j=1}^{\ell_i} L_{ij}$$
2. disjunktive Normalform (DNF)*, falls für geeignete Zahlen k und ℓ_i sowie Literale L_{ij} gilt:*
$$A = \bigvee_{i=1}^{k} \bigwedge_{j=1}^{\ell_i} L_{ij}$$

Beispiele: Wir wollen die Definitionen an einigen Aussagen nachvollziehen.

- Die Aussage $(X_1 \wedge X_2) \vee (\neg X_1 \wedge \neg X_3 \wedge X_4) \vee (X_2 \wedge X_4) \vee X_3$ ist eine disjunktive Normalform mit $k = 4, \ell_1 = 2, \ell_2 = 3, \ell_3 = 2, \ell_4 = 1$ und

$$\begin{array}{lll} L_{11} = X_1, & L_{12} = X_2, & \\ L_{21} = \neg X_1, & L_{22} = \neg X_3, & L_{23} = X_4 \\ L_{31} = X_2, & L_{32} = X_4, & \\ L_{41} = X_3, & & \end{array}$$

- $X_1 \wedge (X_2 \vee X_3)$ ist eine konjunktive Normalform, aber keine disjunktive Normalform.
- $X_1 \vee (X_2 \wedge X_3)$ ist eine disjunktive Normalform, aber keine konjunktive Normalform.
- $X_1 \wedge X_2$ ist eine disjunktive Normalform (mit $k = 2$) und eine konjunktive Normalform (mit $k = 1$).
- $X_1 \wedge (X_2 \vee (X_3 \wedge X_4))$ ist weder eine disjunktive noch eine konjunktive Normalform. Aber es gilt

$$X_1 \wedge (X_2 \vee (X_3 \wedge X_4)) \equiv (X_1 \wedge X_2) \vee (X_1 \wedge X_3 \wedge X_4)),$$

  d.h., die Aussage ist äquivalent zu einer disjunktiven Normalform.

Wir wollen die Einsicht aus dem letzten Beispiel ausdehnen un zeigen, dass jede Aussage äquivalent zu einer konjunktiven und zu einer disjunktiven Normalform ist.

Für eine Aussage X führen wir folgende Schreibweise ein:

$$X^{\mathsf{w}} =_{\text{def}} X, \qquad X^{\mathsf{f}} =_{\text{def}} \neg X$$

Proposition 1.7 *Für eine beliebige Aussage X und einen Wahrheitswert σ gilt*

$$X^\sigma \text{ ist wahr} \iff X \text{ ist } \sigma$$

Beweis: Wir führen eine Fallunterscheidung durch. Ist $\sigma = \mathsf{w}$, so gilt $X^\sigma = X$, d.h., X ist genau dann wahr, wenn X wahr (w) ist. Ist $\sigma = \mathsf{f}$, so gilt $X^\sigma = \neg X$, d.h., $\neg X$ ist genau dann wahr, wenn X falsch (f) ist. Damit ist die Proposition bewiesen. ■

Proposition 1.8 *Es seien $H_1, \ldots H_n$ Aussagen. Dann gilt:*

1. *$\bigwedge_{i=1}^n H_i$ ist genau dann wahr, wenn alle Aussagen H_i wahr sind.*
2. *$\bigvee_{i=1}^n H_i$ ist genau dann wahr, wenn mindestens eine Aussage H_i wahr ist.*

Beweis: Der Induktionsbeweis bleibt zur selbständigen Übung überlassen. ■

Ohne Beweis führen wir den folgenden Satz an, der die Existenz der *kanonischen disjunktiven Normalform* für jede erfüllbare Aussage sichert.

Theorem 1.9 *Für jede erfüllbare Aussage H mit den aussagenlogischen Variablen $X_1, \ldots, X_n$ gilt*

$$H \equiv \bigvee_{\substack{\text{Belegung } I \\ \text{erfüllt } H}} \bigwedge_{i=1}^{n} X_i^{I(X_i)}$$

Beispiele: Für $H =_{\text{def}} X_1 \oplus \neg(X_2 \vee (X_3 \wedge X_1))$ betrachten wir die Wertetabelle. Dabei setzen wir

$$\begin{aligned} H_1 &=_{\text{def}} X_3 \wedge X_1 \\ H_2 &=_{\text{def}} X_2 \vee H_1 \\ H_3 &=_{\text{def}} \neg H_2 \end{aligned}$$

Somit ist $H = X_1 \oplus H_3$. Wir erhalten folgende Wertetabelle:

X_1	X_2	X_3	H_1	H_2	H_3	H
f	f	f	f	f	w	w
f	f	w	f	f	w	w
f	w	f	f	w	f	f
f	w	w	f	w	f	f
w	f	f	f	f	w	f
w	f	w	w	w	f	w
w	w	f	f	w	f	w
w	w	w	w	w	f	w

Nach Theorem 1.9 gilt somit

$$\begin{aligned} H \equiv\ & (\neg X_1 \wedge \neg X_2 \wedge \neg X_3) \vee (\neg X_1 \wedge \neg X_2 \wedge X_3) \vee (X_1 \wedge \neg X_2 \wedge X_3) \\ & \vee\ (X_1 \wedge X_2 \wedge \neg X_3) \vee (X_1 \wedge X_2 \wedge X_3) \end{aligned}$$

Das analoge Theorem gilt auch für die *kanonische konjunktive Normalform* (wiederum ohne Beweis).

Theorem 1.10 *Für jede widerlegbare Aussage H mit den aussagenlogischen Variablen $X_1, \ldots, X_n$ gilt*

$$H \equiv \bigwedge_{\substack{\text{Belegung } I \\ \text{widerlegt } H}} \bigvee_{i=1}^{n} X_i^{\neg I(X_i)}$$

Beispiel: Mit Hilfe obiger Wertetabelle erhalten wir die logische Äquivalenz:

$$X_1 \oplus \neg(X_2 \vee (X_3 \wedge X_1)) \equiv (X_1 \vee \neg X_2 \vee X_3) \wedge (X_1 \vee \neg X_2 \vee \neg X_3) \wedge (\neg X_1 \vee X_2 \vee X_3)$$

1.2 Quantoren

1.2.1 Aussageformen

Definition 1.11 *Eine* Aussageform *über den Universen $U_1, \ldots, U_n$ ist ein Satz $A(x_1, \ldots, x_n)$ mit den freien Variablen $x_1, \ldots, x_n$, der zu einer Aussage wird, wenn jedes x_i durch ein Objekt aus dem Universum U_i ersetzt wird.*

Beispiel: Die Begriffsbildung verdeutlichen wir durch folgende Aussageformen:

- $A(x) =_{\text{def}}$ „x ist eine gerade Zahl“ ist eine Aussageform über den natürlichen Zahlen: $A(2) =$ „2 ist eine gerade Zahl“ ist eine wahre Aussage; $A(3) =$ „3 ist eine gerade Zahl“ ist eine falsche Aussage.
- $B(x, y) =_{\text{def}}$ „Das Wort x ist y Buchstaben lang“ ist eine Aussageform über den Universen U_1 aller Wörter (über einem Alphabet) und U_2 aller natürlichen Zahlen. So ist $B(\text{Konstanz}, 8) =$ „Das Wort Konstanz ist 8 Buchstaben lang“ eine wahre Aussage.
- $C(x) =_{\text{def}}$ „$x < x + 1$“ ist als Aussageform über den natürlichen Zahlen stets wahr unabhängig davon, welche natürliche Zahl n für x eingesetzt wird. Als Aussageform über der Java-Klasse `Integer` gilt dies nicht: $C($`Integer.MAX_VALUE`$)$ ist eine falsche Aussage.

Wenn wir es mit einer Aussageform $A(x_1, \ldots, x_n)$ mit mehreren freien Variablen $x_1, \ldots, x_n$ zu tun haben, die wir alle über dem gleichen Universum $U_1 = U_2 = \ldots = U_n = U$ betrachten, so sprechen wir von einer Aussageform über dem Universum U.

1.2.2 Aussagen mit einem Quantor

Das Einsetzen konkreter Objekte aus einem Universum macht aus einer Aussageform eine Aussage. Eine weitere Möglichkeit dafür ist die *Quantifizierung* von Aussagen mittels Quantoren. Im Unterschied zum konkreten Einsetzen müssen wir dabei die Objekte nicht kennen, deren Einsetzen den Wahrheitswert bestimmt. Wir können nur sagen, dass es solche Objekte gibt oder nicht gibt.

Die beiden wichtigsten Quantoren sind:

- *Existenzquantor* (oder existenzieller Quantor) $\exists$ (manchmal auch $\bigvee$ geschrieben)
- *Allquantor* (oder universeller Quantor) $\forall$ (manchmal auch $\bigwedge$ geschrieben)

Die Quantoren werden wie folgt verwendet, um aus Aussageformen mit einer freien Variablen Aussagen zu machen. Dazu sei $A(x)$ eine Aussageform über dem Universum U.

1. $(\exists x)[A(x)]$ steht für „es gibt ein x, für das $A(x)$ gilt“ und für den Wahrheitswert gilt

   $$(\exists x)[A(x)] \text{ ist wahr } \iff_{\text{def}} \text{ es gibt ein } u \text{ aus } U, \text{ für das } A(u) \text{ wahr ist}$$

2. $(\forall x)[A(x)]$ steht für „für alle x gilt $A(x)$ “ und für den Wahrheitswert gilt

   $$(\forall x)[A(x)] \text{ ist wahr } \iff_{\text{def}} \text{ für alle } u \text{ aus } U \text{ ist } A(u) \text{ wahr}$$

Beispiele: Folgende quantifiziere Aussagen verdeutlichen die Begriffsbildung.

- Für die Aussageform $A(x) =_{\text{def}}$ „x ist eine ungerade Zahl“ über dem Universum der natürlichen Zahlen ist $(\exists x)[A(x)]$ eine wahre Aussage, da $A(3) =$ „3 ist eine ungerade Zahl“ wahr ist, und ist $(\forall x)[A(x)]$ eine falsche Aussage, da $A(2) =$ „2 ist eine ungerade Zahl“ falsch ist.
- Für die Aussageform $C(x) =_{\text{def}}$ „$x < x + 1$ “ über dem Universum der natürlichen Zahlen ist $(\forall x)[C(x)] = (\forall x)[x < x + 1]$ eine wahre Aussage.
- Es sei U ein endliches Universum mit den Objekten $u_1, \ldots, u_n$. Dann gilt:

  $$(\exists x)[A(x)] \equiv A(u_1) \vee A(u_2) \vee \ldots \vee A(u_n) \quad_{\text{def}} = \bigvee_{i=1}^{n} A(u_i)$$

  $$(\forall x)[A(x)] \equiv A(u_1) \wedge A(u_2) \wedge \ldots \wedge A(u_n) \quad_{\text{def}} = \bigwedge_{i=1}^{n} A(u_i)$$

  Der Existenzquantor stellt somit eine endliche oder unendliche Disjunktion und der Allquantor eine endliche oder unendliche Konjunktion dar.

1.2.3 Aussagen mit mehreren Quantoren

Wir erweitern nunmehr die Anwendung von Quantoren auf Aussageformen mit mehr als einer Variablen. Dabei entstehen nicht sofort wieder Aussagen, vielmehr wird pro Anwendung eines Quantors die Anzahl freier Variablen um eine Variable reduziert. Erst wenn alle Variablen durch Quantoren oder Einsetzen konkreter Objekte gebunden sind, können wir der nun entstandenen Aussage einen Wahrheitswert zuordnen.

Es sei $A(x_1, \ldots, x_n)$ eine Aussageform mit n Variablen über den Universen $U_1, \ldots, U_n$. Dann sind $(\exists x_i)[A(x_1, \ldots, x_n)$ und $(\forall x_i)[A(x_1, \ldots, x_n)]$ Aussageformen mit den $n-1$ Variablen $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$.

In $(\exists x_i)[A(x_1, \ldots, x_n)]$ bzw. $(\forall x_i)[A(x_1, \ldots, x_n)]$ heißt $A(x_1, \ldots, x_n)$ der *Wirkungsbereich* des Quantors $\exists x_i$ bzw. $\forall x_i$.

Beispiele: Wir setzen unsere Beispiele für quantifizierte Aussagen fort.

- Es seien $A(x) =_{\text{def}}$ „x ist eine ungerade Zahl“ und $B(x, y) =_{\text{def}}$ „$x \cdot y$ ist eine ungerade Zahl“ Aussageformen über dem Universum der natürlichen Zahlen. Dann sind
  - $C_x(y) =_{\text{def}} (\exists x)[A(x) \rightarrow B(x, y)]$ eine Aussageform mit der freien Variable y und
  - $C_y(x) =_{\text{def}} (\forall y)[A(x) \rightarrow B(x, y)]$ eine Aussageform mit der freien Variable x,

  und es gilt beispielsweise:
  - $C_x(3) =_{\text{def}} (\exists x)[A(x) \rightarrow B(x, 3)]$ ist eine wahre Aussage
  - $C_y(3) =_{\text{def}} (\forall y)[A(3) \rightarrow B(3, y)]$ ist eine falsche Aussage, da $A(3)$ zwar wahr aber $B(3, 2)$ falsch ist.

  Für vollständig quantifizierte Aussagen erhalten wir:
  - $(\exists y)(\forall x)[A(x) \rightarrow B(x, y)]$ ist eine wahre Aussage
  - $(\exists x)(\forall y)[A(x) \rightarrow B(x, y)]$ ist eine wahre Aussage
  - $(\forall y)(\forall x)[A(x) \rightarrow B(x, y)]$ ist eine falsche Aussage
  - $(\forall x)(\forall y)[A(x) \rightarrow B(x, y)]$ ist eine falsche Aussage

  In der Aussage $(\exists y)(\forall x)[A(x) \rightarrow B(x, y)]$ ist $A(x) \rightarrow B(x, y)$ Wirkungsbereich von $\forall x$ und $(\forall x)[A(x) \rightarrow B(x, y)]$ der Wirkungsbereich von $\exists y$.
- Für die Aussageform „$x < y$“ über dem Universum der natürlichen Zahlen ist $(\forall x)(\exists y)[x < y]$ (lies: „für alle x gibt es ein y mit $x < y$“) eine wahre Aussage, da $A(x, x+1)$ stets wahr ist, und $(\exists y)(\forall x)[x < y]$ (lies: „es gibt ein y mit $x < y$ für alle x“) eine falsche Aussage, da $A(y, y)$ stets falsch ist. Das letzte Beispiel macht deutlich, dass es bei geschachtelten quantifizierten Aussagen ganz entscheidend auf die Stellung der Existenz- und Allquantoren zueinander ankommt.

Die Namen von Variablen, die zur Quantifizierung verwendet werden, sind nur innerhalb der Wirkungsbereiche der Quantoren relevant: Zum Beispiel ist $(\exists x)(\forall x)[x < x]$ keine korrekte Quantifizierung, da bei der Einsetzung von Objekten nicht klar ist, welches für welches x die Einsetzung erfolgt; $(\exists x)[x < y] \wedge (\forall x)[x < y]$ ist dagegen unmissverständlich, da $\exists x$ und $\forall x$ überschneidungsfreie Wirkungsbereiche besitzen.

1.2.4 Rechnen mit Quantoren

Auch für quantifizierte Aussagen können wir Rechenregeln (d.h. logische Äquivalenzen) angeben. Wir tun dies hier nur auszugsweise, ohne Beweise und deshalb auch nicht in Form eines Theorems:

$$\left.\begin{array}{lcl} (\exists x)[A(x)] \vee (\exists x)[B(x)] & \equiv & (\exists x)[A(x) \vee B(x)] \\ (\forall x)[A(x)] \wedge (\forall x)[B(x)] & \equiv & (\forall x)[A(x) \wedge B(x)] \end{array}\right\} \text{Assoziativität}$$

$$\left.\begin{array}{lcl} (\exists x)(\exists y)[A(x,y)] & \equiv & (\exists y)(\exists x)[A(x,y)] \\ (\forall x)(\forall y)[A(x,y)] & \equiv & (\forall y)(\forall x)[A(x,y)] \end{array}\right\} \text{Kommutativität}$$

$$\left.\begin{array}{lcl} \neg(\exists x)[A(x)] & \equiv & (\forall x)[\neg A(x)] \\ \neg(\forall x)[A(x)] & \equiv & (\exists x)[\neg A(x)] \end{array}\right\} \text{De Morgan'sche Regeln}$$

Die Stichhaltigkeit und Namensgebung der Rechenregeln ist leicht einzusehen, wenn wir endliche Universen für die Aussage zu Grunde legen und endliche Konjunktionen und Disjunktionen betrachten.

Beispiel: Es sei $P(x) =_{\text{def}}$ „x ist eine Primzahl“ eine Aussageform über dem Universum der natürlichen Zahlen. Wir formulieren die Aussage, dass es unendlich viele Primzahlen gibt, wie folgt:

$$A =_{\text{def}} (\forall x)(\exists y)[P(y) \wedge x < y]$$

Die Negation der Aussage ist: „Es gibt endlich viele Primzahlen“. Wir negieren die Aussage A dazu formal:

$$\begin{array}{rcl} \neg A & \equiv & \neg(\forall x)(\exists y)[P(y) \wedge x < y] \\ & \equiv & (\exists x)\Big[\neg(\exists y)[P(y) \wedge x < y]\Big] \\ & \equiv & (\exists x)(\forall y)[\neg(P(y) \wedge x < y)] \\ & \equiv & (\exists x)(\forall y)[\neg P(y) \vee x \geq y] \end{array}$$

Intuitiv ausgedrückt bedeutet dies Aussage: „Es gibt eine größte Primzahl“.

Quantifizierte Aussagen in komplexeren Domänen werden in der Regel schnell unübersichtlich. Deshalb finden sich oft Abkürzungen für häufig benutzte Redewendungen. Wir wollen diesen Abschnitt mit einigen davon beschließen.

1. „Es gibt $x_1, \ldots, x_n$, sodass $A(x_1, \ldots, x_n)$ gilt“: Die zugehörige Abkürzung für die exakte logische Definition ist:
$$(\exists x_1, \ldots, x_n)[A(x_1, \ldots, x_n)] =_{\text{def}} (\exists x_1)(\exists x_2) \cdots (\exists x_n)[A(x_1, \ldots, x_n)]$$
2. „Für alle $x_1, \ldots, x_n$ gilt $A(x_1, \ldots, x_n)$“: Die zugehörige Abkürzung für die exakte logische Definition ist:
$$(\forall x_1, \ldots, x_n)[A(x_1, \ldots, x_n)] =_{\text{def}} (\forall x_1)(\forall x_2) \cdots (\forall x_n)[A(x_1, \ldots, x_n)]$$
3. „Für alle x mit $A(x)$ gilt $B(x)$“: Die zugehörige Abkürzung für die exakte logische Definition ist:
$$(\forall x; A(x))[B(x)] =_{\text{def}} (\forall x)[A(x) \to B(x)]$$
4. „Es gibt ein x mit $A(x)$, sodass $B(x)$ gilt“: Die zugehörige Abkürzung für die exakte logische Definition ist:
$$(\exists x; A(x))[B(x)] =_{\text{def}} (\forall x)[A(x) \wedge B(x)]$$

Die beiden letzten Regeln sind verträglich mit den De Morgan'schen Regeln:

$$\begin{aligned}
\neg(\exists x; A(x))[B(x)] &\equiv \neg(\exists x)[A(x) \wedge B(x)] \\
&\equiv (\forall x)[\neg(A(x) \wedge B(x))] \\
&\equiv (\forall x)[(\neg A(x)) \vee (\neg B(x))] \\
&\equiv (\forall x)[A(x) \to (\neg B(x))] \\
&\equiv (\forall x; A(x))[\neg B(x)]
\end{aligned}$$

Beispiel: Das Pumping-Lemma für reguläre Sprachen ist ein wichtiges Hilfsmittel im Bereich der Automatentheorie und Formaler Sprachen. Die übliche Formulierung als Theorem ist (vgl. z.B. [Wag03, S. 191]):

> „Für jede reguläre Sprache L gibt es ein $n_0 > 0$ mit folgender Eigenschaft: Für jedes z aus L mit $|z| \geq n_0$ gibt es eine Zerlegung $z = uvw$ mit $|uv| \leq n_0$ und $|v| > 0$, sodass $uv^k w$ zu L gehört für alle $k \geq 0$.“

Mit Hilfe unserer Quantorennotationen ist das Theorem wie folgt ausdrückbar:

$$\begin{aligned}
&(\forall L; L \text{ ist regulär})\ (\exists n_0; n_0 > 0)\ (\forall z; z \text{ gehört zu } L \wedge |z| \geq n_0) \\
&\quad (\exists u, v, w; z = uvw \wedge |uv| \leq n_0 \wedge |v| > 0)\ (\forall k; k \geq 0)\ [uv^k w \text{ gehört zu } L]
\end{aligned}$$

Die Handhabung des Theorems (abgesehen vom Wissen um die verwendeten Begriffe und Notationen) bedarf einiger Übung, da die Quantorenstruktur $\forall\exists\forall\exists\forall$ der Aussage vier Wechsel zwischen All- und Existenzquantoren aufweist.

1.3 Beweise

Unter einem Beweis wollen wir eine Folge von allgemeingültigen Implikationen (Regeln) verstehen, die auf allgemeingültigen Anfangsaussagen (Prämissen) basieren und zu der Zielaussage (Folgerung) führen, deren Allgemeingültigkeit damit nachgewiesen wird.

Wichtige Beweisregeln (Implikationen) für den mathematischen Alltagsgebrauch sind:

- *Abtrennungsregel (modus ponens)*: Sind A und $A \to B$ allgemeingültig, so ist B allgemeingültig.

  Korrektheit folgt aus der Allgemeingültigkeit von $(A \wedge (A \to B)) \to B$.

- *Fallunterscheidung*: Sind $A \to B$ und $\neg A \to B$ allgemeingültig, so ist B allgemeingültig.

  Korrektheit folgt aus der Allgemeingültigkeit von $((A \to B) \wedge ((\neg A) \to B)) \to B$.

- *Kettenschluss*: Sind $A \to B$ und $B \to C$ allgemeingültig, so ist $A \to C$ allgemeingültig.

  Korrektheit folgt aus der Allgemeingültigkeit von $((A \to B) \wedge (B \to C)) \to (A \to C)$.

- *Kontraposition*: Ist $A \to B$ allgemeingültig, so ist $(\neg B) \to (\neg A)$ allgemeingültig.

  Korrektheit folgt aus der Allgemeingültigkeit von $(A \to B) \to ((\neg B) \to (\neg A))$.

- *Indirekter Beweis*: Sind $A \to B$ und $A \to \neg B$ allgemeingültig, so ist $\neg A$ allgemeingültig.

  Korrektheit folgt aus der Allgemeingültigkeit von $((A \to B) \wedge (A \to (\neg B))) \to (\neg A)$.

Im Folgenden wollen an dem Beweis der Irrationalität von $\sqrt{2}$ die logische Struktur und das Zusammenspiel der verschiedenen Beweisregeln offenlegen.

Lemma A. *Ist n eine ungerade Zahl, so ist n^2 eine ungerade Zahl.*

Beweis: *(direkt)* Es sei n eine ungerade Zahl, d.h.

$$n = 2\lfloor n/2 \rfloor + 1. \qquad =_{\text{def}} \mathsf{A}$$

A (für eine konkrete Zahl) ist allgemeingültige Prämisse

Wir müssen zeigen:

$$n^2 = 2\left\lfloor n^2/2 \right\rfloor + 1. \qquad =_{\text{def}} \mathsf{Z}$$

Z ist die Zielaussage

Mit $n = 2\lfloor n/2 \rfloor + 1$ gilt:

$$\begin{aligned} n^2 &= (2\lfloor n/2 \rfloor + 1)^2 && =_{\text{def}} \mathsf{B} && \mathsf{A} \to \mathsf{B} \text{ ist allgemeingültig} \\ &= 4\lfloor n/2 \rfloor^2 + 4\lfloor n/2 \rfloor + 1 && =_{\text{def}} \mathsf{C} && \mathsf{B} \to \mathsf{C} \text{ ist allgemeingültig} \\ &= 2\left(2\lfloor n/2 \rfloor^2 + 2\lfloor n/2 \rfloor\right) + 1 && =_{\text{def}} \mathsf{D} && \mathsf{C} \to \mathsf{D} \text{ ist allgemeingültig} \end{aligned}$$

Wir zeigen zunächst die Hilfsaussage:

$$\left\lfloor n^2/2 \right\rfloor = 2\lfloor n/2 \rfloor^2 + 2\lfloor n/2 \rfloor \qquad =_{\text{def}} \mathsf{H}$$

Wegen $n = 2\lfloor n/2 \rfloor + 1$ gilt:

$$\begin{aligned} \lfloor n^2/2 \rfloor &= \left\lfloor (2\lfloor n/2 \rfloor + 1)^2 / 2 \right\rfloor && =_{\text{def}} \mathsf{E} && \mathsf{A} \to \mathsf{E} \text{ ist allgemeingültig} \\ &= \left\lfloor \left(4\lfloor n/2 \rfloor^2 + 4\lfloor n/2 \rfloor + 1\right)/2 \right\rfloor && =_{\text{def}} \mathsf{F} && \mathsf{E} \to \mathsf{F} \text{ ist allgemeingültig} \\ &= \left\lfloor 2\lfloor n/2 \rfloor^2 + 2\lfloor n/2 \rfloor + 1/2 \right\rfloor && =_{\text{def}} \mathsf{G} && \mathsf{F} \to \mathsf{G} \text{ ist allgemeingültig} \\ &= 2\lfloor n/2 \rfloor^2 + 2\lfloor n/2 \rfloor && = \mathsf{H} && \mathsf{G} \to \mathsf{H} \text{ ist allgemeingültig} \end{aligned}$$

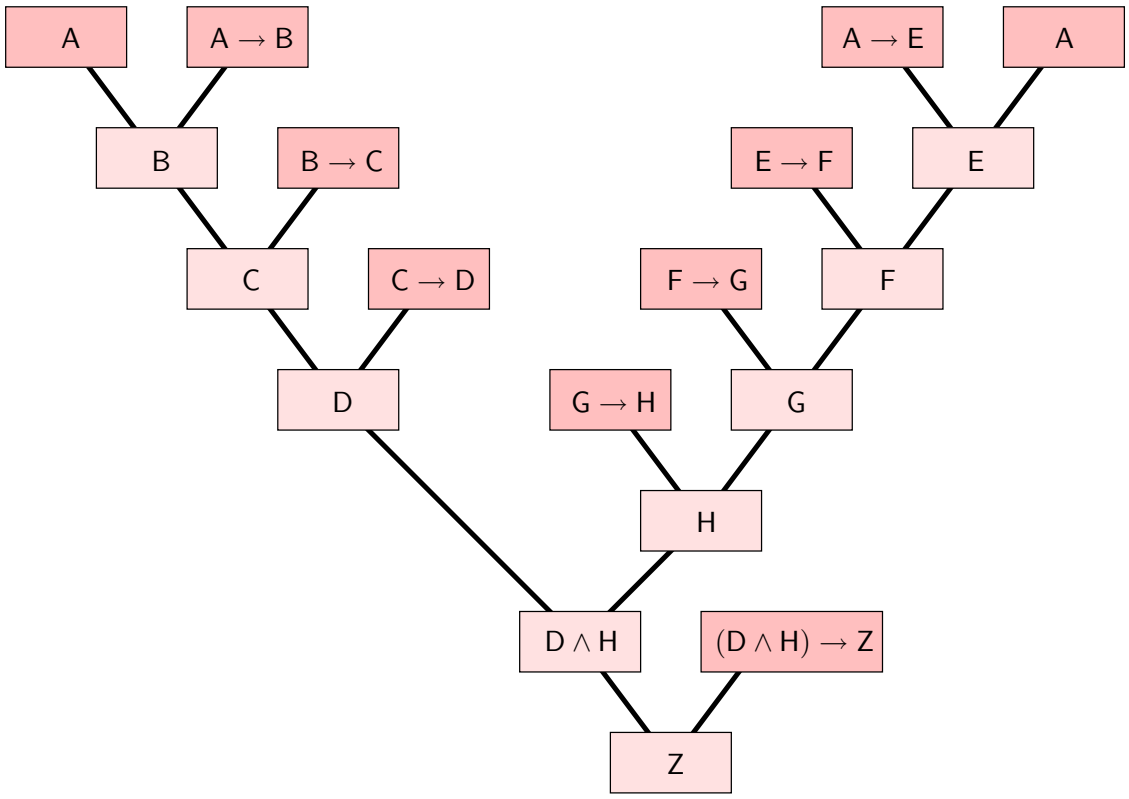
Einsetzen der Hilfsaussage in D ergibt:

$$n^2 = 2\left\lfloor n^2/2 \right\rfloor + 1. \qquad = \mathsf{Z}$$

$(\mathsf{D} \wedge \mathsf{H}) \to \mathsf{Z}$ ist allgemeingültig

d.h. n^2 ist ungerade. ■

Die logische Struktur des Beweises kann schematisch in Form eines Ableitungsbaumes dargestellt werden:



Hierbei sind die heller unterlegten Aussagen (bis auf $\mathsf{D} \wedge \mathsf{H}$) durch Anwendung der Abtrennungsregel aus den beiden darüber liegenden Aussagen abgeleitet worden. Die dunkler unterlegten Aussagen sind per Voraussetzung allgemeingültig (Aussage A) oder durch Anwendung algebraischer Umformungsregeln allgemeingültig.

Durch Kontraposition von Lemma A lässt sich nun direkt Korollar B folgern.

Korollar B. *Ist n^2 eine gerade Zahl, so ist n eine gerade Zahl.*

Beweis: *(Kontraposition)*
Ist n eine ungerade Zahl, $=_{\text{def}} \mathsf{A}$
so ist n^2 eine ungerade Zahl $=_{\text{def}} \mathsf{B}$
(nach Lemma A).

Damit gilt nach Kontraposition:
Ist n^2 eine gerade Zahl, $\equiv \neg\mathsf{B}$
so ist n eine gerade Zahl. $\equiv \neg\mathsf{A}$

Damit ist das Korollar bewiesen. ■

$\mathsf{A} \rightarrow \mathsf{B}$ ist allgemeingültig

$\neg\mathsf{B} \rightarrow \neg\mathsf{A}$ ist allgemeingültig

Mit Hilfe von Korollar B kann die Irrationalität von $\sqrt{2}$ mittels Widerspruchbeweis gezeigt werden.

Theorem C. *$\sqrt{2}$ ist irrational.*

Beweis: *(indirekt)* Wir nehmen an: $\sqrt{2}$ ist eine rationale Zahl, d.h.

$$(\exists p)(\exists q)\Big[\underbrace{\mathrm{ggT}(p,q)=1}_{=_{\text{def}}\ \mathsf{Z}} \ \wedge\ \sqrt{2}=p/q\Big] \quad =_{\text{def}}\ \mathsf{A}$$

$\mathsf{A} \to \mathsf{Z}$ ist allgemeingültig

Dann gilt

$$2q^2 = p^2, \qquad =_{\text{def}}\ \mathsf{B}$$

$\mathsf{A} \to \mathsf{B}$ ist allgemeingültig

d.h. p^2 ist gerade.

Nach Korollar B ist p gerade, d.h.

$$p = 2\lfloor p/2 \rfloor. \qquad =_{\text{def}}\ \mathsf{C}$$

$\mathsf{B} \to \mathsf{C}$ ist allgemeingültig

Wollen zeigen, dass auch q^2 gerade ist, d.h.

$$q^2 = 2\lfloor q^2/2 \rfloor. \qquad =_{\text{def}}\ \mathsf{D}$$

Mit $2q^2 = p^2$ und $p = 2\lfloor p/2 \rfloor$ folgt

$$\begin{aligned} q^2 &= p^2/2 && =_{\text{def}}\ \mathsf{E} && \mathsf{B} \to \mathsf{E} \text{ ist allgemeingültig} \\ &= \left(2\lfloor p/2 \rfloor\right)^2/2 && =_{\text{def}}\ \mathsf{F} && (\mathsf{C} \wedge \mathsf{E}) \to \mathsf{F} \text{ ist allgemeingültig} \\ &= 2\left\lfloor p/2 \right\rfloor^2 && =_{\text{def}}\ \mathsf{G} && \mathsf{F} \to \mathsf{G} \text{ ist allgemeingültig} \end{aligned}$$

und somit

$$\begin{aligned} 2\left\lfloor q^2/2 \right\rfloor &= 2\left\lfloor 2\lfloor p/2 \rfloor^2/2 \right\rfloor && =_{\text{def}}\ \mathsf{H} && \mathsf{G} \to \mathsf{H} \text{ ist allgemeingültig} \\ &= 2\lfloor \lfloor p/2 \rfloor^2 \rfloor && =_{\text{def}}\ \mathsf{I} && \mathsf{H} \to \mathsf{I} \text{ ist allgemeingültig} \\ &= 2\lfloor p/2 \rfloor^2 && =_{\text{def}}\ \mathsf{J} && \mathsf{I} \to \mathsf{J} \text{ ist allgemeingültig} \\ &= q^2 && \equiv\ \mathsf{D} && \mathsf{J} \to \mathsf{D} \text{ ist allgemeingültig} \end{aligned}$$

Nach Korollar B ist q gerade. $=_{\text{def}}\ \mathsf{K}$ $\mathsf{D} \to \mathsf{K}$ ist allgemeingültig

Damit gilt $\mathrm{ggT}(p,q) \geq 2$. $\equiv \neg\mathsf{Z}$ $\mathsf{K} \to \neg\mathsf{Z}$ ist allgemeingültig

$\mathsf{A} \to \neg\mathsf{Z}$ ist allgemeingültig

Dies ist ein Widerspruch, d.h. die Annahme ist falsch und $\sqrt{2}$ ist irrational. $\equiv \neg\mathsf{A}$ $\neg\mathsf{A}$ ist allgemeingültig

Damit ist das Theorem bewiesen. ■

Neben den Beweisregeln, die ganz allgemein für beliebige Aussagen anwendbar sind, gibt es noch eine ganze Menge spezieller Beweisregeln für Aussagen mit bestimmter Quantorenstruktur und bestimmter Universen. Zwei wichtige unter diesen sind die folgenden:

- *Spezialisierung (Substitution)*: Ist $(\forall x)[A(x)]$ allgemeingültig, so ist $A(y)$ allgemeingültig, falls y nicht in einem Wirkungsbereich eines Quantors in $A(x)$ vorkommt.

  Korrektheit folgt aus Allgemeingültigkeit von $(\forall y)\left[(\forall x)[A(x)] \to A(y)\right]$ (mit obiger Einschränkung).

- *Vollständige Induktion*: Es sei $A(n)$ eine Aussageform über dem Universum der natürlichen Zahlen. Sind $A(0)$ und $A(n-1) \to A(n)$ für alle $n > 0$ allgemeingültig, so ist $A(n)$ für alle n allgemeingültig.

Wir wollen die Korrektheit der vollständigen Induktion überprüfen.

Theorem 1.12 *Es sei $A(n)$ eine Aussageform mit der freien Variable n über dem Universum der natürlichen Zahlen. Dann ist die Aussage*

$$\Big(A(0) \wedge (\forall n; n > 0)[A(n-1) \to A(n)]\Big) \to (\forall n)[A(n)]$$

allgemeingültig.

Beweis: *(indirekt)* Es gelte $A(0)$ und $A(n-1) \to A(n)$ für alle $n > 0$. Zum Widerspruch nehmen wir an, dass es ein n gibt, sodass $A(n)$ nicht gilt. Dann gibt es auch eine kleinste natürliche Zahl n_0, für die $A(n_0)$ nicht wahr ist, d.h. es gilt $\neg A(n_0) \wedge (\forall n; n < n_0)[A(n)]$. Wir unterscheiden zwei Fälle für n_0:

- *1. Fall*: Ist $n_0 = 0$, so ist $\neg A(0)$ wahr. Dies steht jedoch im Widerspruch zur Voraussetzung, dass $A(0)$ gilt.

- *2. Fall*: Ist $n_0 > 0$, so ist $\neg A(n_0) \wedge A(n_0 - 1) \equiv \neg(A(n_0 - 1) \to A(n_0))$ wahr. Dies steht jedoch im Widerspruch zur Voraussetzung, dass $A(n-1) \to A(n)$ für alle $n > 0$ gilt, also insbesondere auch $A(n_0 - 1) \to A(n_0)$.

Also ist die Annahme falsch und es gilt $A(n)$ für alle n. Damit ist das Theorem bewiesen. ■

Die logische Struktur des Beweises für Theorem 1.12 ist typisch für einen Widerspruchsbeweis einer Implikation. Wenn wir die Allgemeingültigkeit von $A \to B$ beweisen wollen, so nehmen wir an, dass A aber nicht B gilt. Damit folgt sofort die Allgemeingültigkeit der Aussage $(A \wedge (\neg B)) \to A$. Anschließend müssen wir noch beweisen, dass auch $(A \wedge (\neg B)) \to (\neg A)$ allgemeingültig ist, d.h. wir konstruieren einen Widerspruch zur eigentlichen Prämisse A unserer zu beweisenden Implikation. Nach der Regel vom indirekten Beweis folgt nun, dass $\neg(A \wedge (\neg B)) \equiv A \to B$ allgemeingültig ist.

Mengen 2

In diesem Kapitel beschäftigen wir uns mit den grundlegenden Begriffen der Mengenlehre. Hierbei folgen wir im Wesentlichen der naiven Mengenlehre, wie sie im mathematischen Alltagsgeschäft eines Informatikers ausreichend ist. Unter einer streng mathematischen Sichtweise ist die naive Mengenlehre nicht widerspruchsfrei; jedoch können Widersprüche mit einer gewissen Umsicht vermieden werden.

2.1 Definitionen

Eine *Menge* A besteht aus paarweise verschiedenen Objekten. Damit wird ein mehrfaches Vorkommen von Objekten ignoriert – im Gegensatz z.B. zu Listen als Datenstruktur.

Bei der Beschreibung einer Menge A unterscheiden wir zwei Formen der Darstellung:

- *extensionale Darstellung:* Die in der Menge A enthaltenen Objekte werden aufgezählt (soweit dies möglich ist), wobei die Reihenfolge keine Rolle spielt – auch hier im Gegensatz zu Listen; symbolisch:

  $$A = \{a_1, a_2, \ldots\}$$

- *intensionale Darstellung:* Es werden alle Objekte a selektiert, die aus dem zu einer Aussageform $E(x)$ gehörenden Universum stammen, sodass $E(a)$ eine wahre Aussage ist; symbolisch:

  $$A = \{\ a \mid E(a)\ \}$$

  Mit anderen Worten enthält die Menge A alle Objekte a, die eine gewisse Eigenschaft E erfüllen.

Extensionale Darstellungen sind für die Fälle endlicher Mengen häufig einsichtiger als intensionale Darstellungen, da die Selektion der Objekte bereits ausgeführt vorliegt. Für unendliche Mengen sind extensionale Darstellungen im Allgemeinen nicht mehr möglich.

Beispiele: Die folgenden Darstellung derselben (endlichen) Menge verdeutlichen die unterschiedlichen Beschreibungsaspekte:

- $\{3, 5, 7, 11\} = \{11, 5, 7, 3\}$
- $\{3, 5, 7, 11\} = \{3, 3, 3, 5, 5, 7, 11, 11\}$
- $\{3, 5, 7, 11\} = \{\ a \mid 2 < a < 12 \wedge\ a \text{ ist eine Primzahl}\ \}$

Im Folgenden vereinbaren die Schreib- und Sprechweisen für mengenbezogene Aussagen. Positive Aussagen sind die folgenden:

$a \in A$	steht für:	a ist Element von A
$A \subseteq B$	steht für:	A ist Teilmenge von B
$B \supseteq A$	steht für:	B ist Obermenge von A
$A = B$	steht für:	A und B sind gleich
$A \subset B$	steht für:	A ist echte Teilmenge von B
$B \supset A$	steht für:	B ist echte Obermenge von A

Die zugehörigen negativen Aussagen sind:

$a \notin A$	steht für:	a ist kein Element von A
$A \not\subseteq B$	steht für:	A ist keine Teilmenge von B
$B \not\supseteq A$	steht für:	B ist keine Obermenge von A
$A \neq B$	steht für:	A und B sind verschieden
$A \not\subset B$	steht für:	A ist keine echte Teilmenge von B
$B \not\supset A$	steht für:	B ist keine echte Obermenge von A

Die exakten Bedeutungen der Bezeichnungen werden aussagenlogisch festgelegt. Dazu setzen wir im Folgenden für die verwendeten Aussageformen stets ein Universum voraus.

$a \in A$	$=_{\text{def}}$	a gehört zur Menge A	$a \notin A$	$=_{\text{def}}$	$\neg(a \in A)$
$A \subseteq B$	$=_{\text{def}}$	$(\forall a)[a \in A \to a \in B]$	$A \not\subseteq B$	$=_{\text{def}}$	$\neg(A \subseteq B)$
$B \supseteq A$	$=_{\text{def}}$	$A \subseteq B$	$B \not\supseteq A$	$=_{\text{def}}$	$\neg(B \supseteq A)$
$A = B$	$=_{\text{def}}$	$A \subseteq B \wedge B \subseteq A$	$A \neq B$	$=_{\text{def}}$	$\neg(A = B)$
$A \subset B$	$=_{\text{def}}$	$A \subseteq B \wedge A \neq B$	$A \not\subset B$	$=_{\text{def}}$	$\neg(A \subset B)$
$B \supset A$	$=_{\text{def}}$	$A \subset B$	$B \not\supset A$	$=_{\text{def}}$	$\neg(B \supset A)$

Aussagen über Mengen werden also als Abkürzungen für quantifizierte Aussagen über ihren Elementen eingeführt.

Beispiele: Wir verdeutlichen den Zusammenhang zwischen Aussagen über Mengen und den definierenden quantifizierten Aussagen über den Elementen an Hand zweier Mengenaussagen:

$$\begin{aligned} A \not\subseteq B &\equiv \neg(A \subseteq B) \\ &\equiv \neg(\forall a)[a \in A \to a \in B] \\ &\equiv (\exists a)[\neg(a \in A \to a \in B)] \end{aligned}$$

$$\begin{aligned} &\equiv (\exists a)[\neg(a \notin A \vee a \in B)] \\ &\equiv (\exists a)[a \in A \wedge a \notin B] \end{aligned}$$

$$\begin{aligned} A = B &\equiv A \subseteq B \wedge A \supseteq B \\ &\equiv A \subseteq B \wedge B \subseteq A \\ &\equiv (\forall a)[a \in A \rightarrow a \in B] \wedge (\forall a)[a \in B \rightarrow a \in A] \\ &\equiv (\forall a)[(a \in A \rightarrow a \in B) \wedge (a \in B \rightarrow a \in A)] \\ &\equiv (\forall a)[(a \in A \leftrightarrow a \in B)] \end{aligned}$$

Häufig muss die Gleichheit zweier Mengen A und B, die in intensionaler Darstellung gegeben sind, gezeigt werden. Nach Definition des Wahrheitswertes der Aussage $A = B$ müssen dafür stets zwei Richtungen gezeigt werden. Ein einfaches Beispiel soll dies verdeutlichen.

Beispiel: Es seien die beiden Mengen $A =_{\text{def}} \{\ n \mid n \text{ ist gerade}\ \}$ und $B =_{\text{def}} \{\ n \mid n^2 \text{ ist gerade}\ \}$ als Teilmengen natürlicher Zahlen gegeben. Wir wollen zeigen, dass $A = B$ gilt. Dazu zeigen wir zwei Inklusionen:

- $\subseteq$: Es sei $n \in A$. Dann ist n ist gerade, d.h., es gibt ein $k \in \mathbb{N}$ mit $n = 2k$. Es gilt $n^2 = (2k)^2 = 2(2k^2)$. Somit ist n^2 gerade. Folglich gilt $n \in B$. Damit gilt $A \subseteq B$.
- $\supseteq$: Es sei $n \in B$. Dann ist n^2 gerade. Nach Korollar B (Abschnitt 1.6) ist n gerade. Also gilt $n \in A$. Somit gilt $B \subseteq A$.

Damit ist die Gleichheit der Mengen bewiesen.

Eine ausgezeichnete Menge (in jedem Universum) ist die leere Menge: Eine Menge A heißt *leer* genau dann, wenn A kein Element enthält. Logisch ausgedrückt bedeutet die Bedingung: $(\forall a)[a \notin A]$.

Proposition 2.1 *Es gibt nur eine leere Menge (in jedem Universum).*

Beweis: *(Kontraposition)* Wir wollen zeigen: Sind A und B leere Mengen, so gilt $A = B$. Dafür zeigen wir: Gilt $A \neq B$, so ist A nicht leer oder B nicht leer. Es gilt:

$$\begin{aligned} A \neq B &\equiv A \not\subseteq B \vee B \not\subseteq A \\ &\equiv (\exists a)[a \in A \wedge a \notin B] \vee (\exists a)[a \in B \wedge a \notin A] \\ &\equiv (\exists a) \underbrace{[(a \in A \wedge a \notin B) \vee (a \in B \wedge a \notin A)]}_{=_{\text{def}}\ D(a)} \end{aligned}$$

Es sei x ein Objekt im Universum, so dass $D(x)$ eine wahre Aussage ist. Dann gilt $x \in A$ oder $x \in B$. Also ist A oder B nicht leer. ∎

Damit ist gerechtfertigt, dass ein eigenes Symbol $\emptyset$ für die Bezeichnung der leeren Menge eingeführt wird.

$\|A\|$ (oder auch: $|A|, \#A$) ist die Anzahl der Elemente von A bzw. die *Kardinalität* von A. Die Kardinalität der leeren Menge ist also stets 0. Ist $\|A\| < \infty$, so heißt A *endliche* Menge, sonst *unendliche* Menge. Mengen mit nur einem Element werden *Einermengen* genannt.

2.2 Mengenoperationen

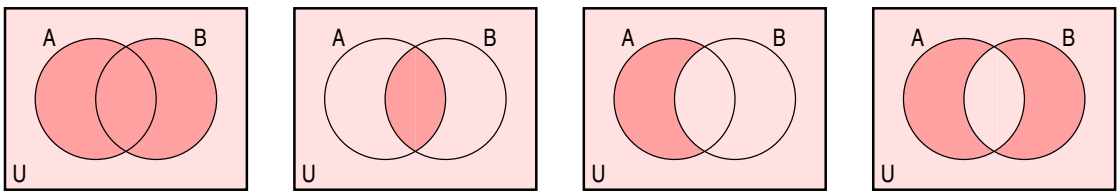
Wir definieren die folgenden Operationen, die aus zwei Mengen A und B eines Universums U wieder eine Menge desselben Universums U formen:

Vereinigung:	$A \cup B =_{\text{def}} \{ x \mid x \in A \vee x \in B \}$
Durchschnitt:	$A \cap B =_{\text{def}} \{ x \mid x \in A \wedge x \in B \}$
Differenz:	$A \setminus B =_{\text{def}} \{ x \mid x \in A \wedge x \notin B \}$
symmetrische Differenz:	$A \triangle B =_{\text{def}} (A \setminus B) \cup (B \setminus A)$

Eine besondere Differenzoperation ist die Komplementierung einer Menge A:

Komplement:	$\overline{A} =_{\text{def}} U \setminus A$

Üblicherweise werden Mengenoperationen zur Veranschaulichung durch die aus der Schule bekannten Venn-Diagramme dargestellt. Die vier obigen Operationen auf zwei Mengen lassen sich wie folgt visualisieren:



$A \cup B$ $A \cap B$ $A \setminus B$ $A \triangle B$

Dabei sind die dunkler dargestellten Punktmengen immer das Ergebnis der jeweiligen Mengenoperationen auf den durch die Kreis A und B eingefassten Punktmengen. Diese Darstellungsformen sind zwar illustrativ; sie sind jedoch *keinesfalls* ausreichend für Beweise.

Beispiele: Es seien $A = \{2, 3, 5, 7, 11\}$ und $B = \{2, 3, 4, 5, 6\}$. Dann gilt:

- $A \cup B = \{2, 3, 4, 5, 6, 7, 11\}$
- $A \cap B = \{2, 3, 5\}$

- $A \setminus B = \{7, 11\}$
- $B \setminus A = \{4, 6\}$
- $A \triangle B = \{4, 6, 7, 11\}$
- $(A \setminus B) \cap B = \{7, 11\} \cap \{2, 3, 4, 5, 6\} = \emptyset$

Zwei Mengen A und B heißen *disjunkt* genau dann, wenn $A \cap B = \emptyset$ gilt.

2.3 Verallgemeinerung von Vereinigung und Durchschnitt

In einigen Fällen werden auch Verallgemeinerungen von Vereinigung und Durchschnitt auf eine beliebige, auch unendliche, Anzahl von Mengen betrachtet.

Dazu betrachten wir Teilmengen eines Universums U. Weiterhin sei I eine beliebige Menge (Indexmenge). Für jedes $i \in I$ sei eine Menge $A_i \subseteq U$ gegeben. Dann sind Vereinigung und Durchschnitt aller A_i definiert als:

$$\bigcup_{i \in I} A_i \quad =_{\text{def}} \quad \{\ a \mid (\exists i \in I)[a \in A_i]\ \}$$

$$\bigcap_{i \in I} A_i \quad =_{\text{def}} \quad \{\ a \mid (\forall i \in I)[a \in A_i]\ \}$$

Für $I = \mathbb{N}$ schreiben wir auch $\bigcup_{i=0}^{\infty} A_i$ bzw. $\bigcap_{i=0}^{\infty} A_i$.

Beispiele: Folgende Beispiele und Spezialfälle sollen die Wirkungsweise von allgemeiner Vereinigung und Durchschnitt demonstrieren:

- Es seien $U = \mathbb{R}$, $I = \mathbb{N}_+$ und

$$A_k =_{\text{def}} \left\{ x \;\middle|\; |x^2 - 1| \leq \frac{1}{k} \right\}$$

Dann gilt:

$$\begin{aligned} \bigcup_{k \in I} A_k \;=\; \bigcup_{k=1}^{\infty} A_k \;&=\; \left\{ x \;\middle|\; |x^2 - 1| \leq 1 \right\} \\ &=\; \left\{ x \;\middle|\; -\sqrt{2} \leq x \leq \sqrt{2} \right\} \;\overset{\text{def}}{=}\; \left[-\sqrt{2}, \sqrt{2}\right] \end{aligned}$$

$$\bigcap_{k \in I} A_k \;=\; \bigcap_{k=1}^{\infty} A_k \;=\; \{-1, 1\}$$

- Es gilt stets $\bigcup_{i \in \emptyset} A_i = \emptyset$.
- Es gilt stets $\bigcap_{i \in \emptyset} A_i = M$.

2.4 Potenzmenge

Eine Operation von einem anderen Typ als Vereinigung, Durchschnitt und Differenzen ist die Potenzierung einer Menge: Die *Potenzmenge* einer Menge A ist definiert als

$$\mathcal{P}(A) =_{\text{def}} \{ X \mid X \subseteq A \}$$

Für die Potenzmenge von A gelten folgenden Aussagen:

Proposition 2.2 *Es sei A eine beiliebige Menge.*

1. $X \in \mathcal{P}(A) \Longleftrightarrow X \subseteq A$
2. $\emptyset, A \in \mathcal{P}(A)$
3. *Ist A endlich, so gilt* $\|\mathcal{P}(A)\| = 2^{\|A\|}$.

Beweis: Die erste beiden Aussagen folgen direkt aus der Definition. Die dritte Aussage werden wir im Kapitel über Kombinatorik beweisen. ■

Die Elemente der Potenzmenge sind also Mengen aus dem Universum $\mathcal{P}(A)$. Der letzte Sachverhalt lässt die mitunter auch verwendete Bezeichnung 2^A für die Potenzmenge von A plausibel erscheinen.

Beispiele: Folgende Mengen verdeutlichen die Potenzmengenkonstruktion.

- $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}$
- $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$
- $\mathcal{P}(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$
- $\mathcal{P}(\emptyset) = \{\emptyset\}$
- $\mathcal{P}(\mathcal{P}(\emptyset)) = \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$

Die Teilmengen der Potenzmenge heißen *Mengenfamilien*.

2.5 Partitionen

Eine wichtige Mengenfamilie bilden Partitionen oder Zerlegungen eines Universums.

Definition 2.3 *Eine Mengenfamilie $\mathcal{F} \subseteq \mathcal{P}(A)$ heißt (ungeordnete)* Partition von A*, falls folgende Bedingungen erfüllt sind:*

1. $B \cap C = \emptyset$ *für alle Mengen* $B, C \in \mathcal{F}$ *mit* $B \neq C$
2. $\bigcup_{B \in \mathcal{F}} B = A$

Die Mengen $B \in \mathcal{F}$ heißen Komponenten *der Partition.*

Beispiele: Folgende Familien verdeutlichen das Konzept von Partitionen.

- $\{\mathbb{R}_{<0}, \{0\}, \mathbb{R}_{>0}\}$ ist eine Partition von $\mathbb{R}$ mit drei Komponenten.
- $\{\ \{x\} \mid x \in \mathbb{R}\ \}$ ist eine Partition von $\mathbb{R}$ mit unendlich vielen Komponenten. (Genauer gesagt besteht die Partition aus überabzählbar vielen Komponenten.)
- Für jede Menge $A \subseteq U$ ist $\{A, \overline{A}\}$ eine Partition von U mit zwei Komponenten. Tatsächlich ist für jede Menge $U \neq \emptyset$ die Partition $\{A, \overline{A}\}$ die einzige Partition von U, die die Menge A als Komponente besitzt. Dazu muss nur gezeigt werden, dass für zwei Partitionen $\{A, B\}$ und $\{A, \overline{A}\}$ stets $B = \overline{A}$ gilt. Dies ist leicht einzusehen, da einerseits $A \cap B = \emptyset$ äquivalent zu $A \subseteq \overline{B}$ und andererseits $A \cup B = U$ einmal äquivalent zu $\overline{A} \cap \overline{B} = \emptyset$ und somit auch zu $\overline{B} \subseteq A$ ist. Damit folgt $B = \overline{A}$.

Im letzten Beispiel haben wir die Gleichheit zweier Partitionen dadurch gezeigt, dass wir die Gleichheit aller einzelnen Komponenten nachgeweisen haben ($A = A$ und $B = \overline{A}$). Damit haben wir uns zuviel Arbeit gemacht. Es hätte genügt Inklusionen der Komponenten zu zeigen ($A \subseteq A$ und $B \subseteq \overline{A}$). Die Gleichheiten der Komponenten folgen mittels des Hauber'schen Theorems.

Theorem 2.4 (Hauber) *Es seien $\{A_1, \ldots, A_n\}$ und $\{B_1, \ldots, B_n\}$ zwei Partitionen von U. Gilt $A_i \subseteq B_i$ für alle $i \in \{1, \ldots, n\}$, so gilt $B_i \subseteq A_i$ (und mithin $A_i = B_i$) für alle $i \in \{1, \ldots, n\}$.*

Beweis: (*Induktion*) Wir beweisen den Satz mittels vollständiger Induktion über die Anzahl n der Komponenten der Partitionen.

- *Induktionsanfang*: Wir führen den Induktionsanfang für $n = 1$ und $n = 2$ durch. Für $n = 1$ gilt $A_1 = B_1 = U$ wegen der zweiten Eigenschaft von Partitionen. Für $n = 2$ seien $\{A_1, A_2\}$ und $\{B_1, B_2\}$ zwei Partitionen von U mit $A_1 \subseteq B_1$ und $A_2 \subseteq B_2$. Wegen $A_2 = \overline{A_1}$ und $B_2 = \overline{B_1}$ gilt $B_2 = \overline{B_1} \subseteq \overline{A_1} = A_2$ sowie $B_1 = \overline{B_2} \subseteq \overline{A_2} = A_1$.

- *Induktionsschritt*: Es sei $n > 1$. Es seien $\{A_1, A_2, \ldots, A_n\}$ und $\{B_1, B_2, \ldots, B_n\}$ Partitionen von U mit $A_i \subseteq B_i$ für alle $i \in \{1, 2, \ldots, n\}$. Wir betrachten die beiden Mengen

$$A' =_{\text{def}} A_2 \cup \cdots \cup A_n, \qquad B' =_{\text{def}} B_2 \cup \cdots \cup B_n.$$

  Wegen $A_1 \cap A_i = \emptyset$ für alle $i \in \{2, \ldots, n\}$ gilt $A_1 \cap A' = \emptyset$. Andererseits gilt offensichtlich $A_1 \cup A' = U$. Somit ist $\{A_1, A'\}$ eine Partition von U. Mit einem analogen Argument erhalten wir, dass $\{B_1, B'\}$ ebenfalls eine Partition von U ist. Wegen $A_i \subseteq B_i$ für alle $i \in \{1, 2, \ldots, n\}$ folgt $A_1 \subseteq B_1$ und $A' \subseteq B'$. Nach Induktionsvoraussetzung (für $n = 2$) gilt somit $B_1 \subseteq A_1$ (mithin $A_1 = B_1$) sowie $B' \subseteq A'$ (mithin $A' = B'$).

  Wir müssen noch zeigen, dass $B_i \subseteq A_i$ für alle $i \in \{2, \ldots, n\}$ gilt. Da $A_i \cap A_j = \emptyset$ für alle $i, j \in \{2, \ldots, n\}$ mit $i \neq j$ gilt, ist die Mengenfamilie $\{A_2, \ldots, A_n\}$ eine Partition von A'. Wegen $A' = B'$ folgt mit dem gleichen Argument, dass $\{B_2, \ldots, B_n\}$ eine Partition von A' ist. Außerdem gilt weiterhin $A_i \subseteq B_i$ für alle $i \in \{2, \ldots, n\}$. Nach Induktionsvoraussetzung (für $n - 1$ sowie A') folgt $B_i \subseteq A_i$ für alle $i \in \{2, \ldots, n\}$.

  Insgesamt haben wir also gezeigt, dass $B_i \subseteq A_i$ für alle $i \in \{1, 2, \ldots, n\}$ gilt.

Damit ist der Satz bewiesen. ■

Relationen 3

Relationen beschreiben die Beziehungen zwischen Mengen und sind somit der eigentliche Gegenstand der Mathematik.

3.1 Kreuzprodukt

Es seien $A_1, \ldots, A_n$ beliebige Mengen. Das *Kreuzprodukt* (oder kartesisches Produkt) von $A_1, \ldots, A_n$ ist definiert als:

$$A_1 \times \cdots \times A_n =_{\text{def}} \{\ (a_1, \ldots, a_n) \mid \text{für alle } i \in \{1, \ldots, n\} \text{ gilt } a_i \in A_i\ \}$$

Die Elemente von $A_1 \times \cdots \times A_n$ heißen *n-Tupel* (*Paare* für $n = 2$, *Tripel* für $n = 3$, *Quadrupel* für $n = 4$).

Im Gegensatz zu Mengen sind Tupel geordnet (und damit eine Formalisierung von Listen): Für zwei n-Tupel $(a_1, \ldots, a_n)$ und $(a'_1, \ldots, a'_n)$ gilt

$$(a_1, \ldots, a_n) = (a'_1, \ldots, a'_n) \iff \text{für alle } i \in \{1, \ldots, n\} \text{ gilt } a_i = a'_i$$

Sind alle Mengen gleich, so schreibt man:

$$A^n =_{\text{def}} \underbrace{A \times \cdots \times A}_{n\text{-mal}}$$

Beispiele: Folgende Mengen verdeutlichen die Kreuzproduktkonstruktion.

- Mit $A = \{1, 2, 3\}$ und $B = \{a, b\}$ gilt

$$A \times B = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$$

- Mit $A = \{5, 7\}$ und $n = 3$ gilt

$$\begin{aligned} A^3 &= \{5, 7\} \times \{5, 7\} \times \{5, 7\} \\ &= \{(5, 5, 5), (5, 5, 7), (5, 7, 5), (5, 7, 7), \\ &\qquad (7, 5, 5), (7, 5, 7), (7, 7, 5), (7, 7, 7)\} \end{aligned}$$

- $\emptyset \times A = \emptyset$ (*beachte:* Die rechte leere Menge ist die Menge in der kein Paar enthalten ist)
- $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ beschreibt den dreidimensionalen Raum

Es seien $A_1, \ldots, A_n$ beliebige Mengen. Eine Menge $R \subseteq A_1 \times \cdots \times A_n$ heißt *n-stellige Relation*.

Beispiel: Eine relationale Datenbank ist eine Sammlung von Tabellen mit einer gewissen Struktur. Eine Tabelle wiederum ist eine (extensionale Darstellung einer) Relation. Beispielsweise sei folgender Auszug einer Tabelle gegeben:

Vorname	Name	Geburtsdatum
Max	Mustermann	07.07.1977
Erika	Mustermann	12.09.1945
John	Smith	05.05.1955
Johanna	König-Hock	27.03.1921
Lyudmila	Dyakovska	02.04.1976
Peter	Draisaitl	07.12.1965
Thomas	Holtzmann	01.04.1927
Weiwei	Ai	28.08.1957
Robert	Palfrader	11.11.1968
Nikon	Jevtic	03.06.1993
Leslie	Valiant	28.03.1949
⋮	⋮	⋮

Wir fassen die Tabelle als Teilmenge eines Kreuzproduktes auf. Dazu seien:

$$\begin{aligned} A_1 &=_{\text{def}} \text{Menge aller Vornamen in der Tabelle} \\ A_2 &=_{\text{def}} \text{Menge aller Namen in der Tabelle} \\ A_3 &=_{\text{def}} \text{Menge aller Geburtsdaten in der Tabelle} \end{aligned}$$

Dann ist $(\texttt{Max}, \texttt{Mustermann}, \texttt{07.07.1977}) \in A_1 \times A_2 \times A_3$ und die Menge aller Zeilen der Tabelle ist eine Relation $R \subseteq A_1 \times A_2 \times A_3$.

Eine Relation $R \subseteq A_1 \times A_2$ heißt *binäre Relation*. Gilt $A_1 = A_2 = A$, so sprechen wir von einer *binären Relation auf A*.

Binäre Relationen R werde auch in Infix-Notation geschrieben:

$$xRy \iff_{\text{def}} (x, y) \in R$$

Der Ausdruck „xRy“ steht dabei für die Leseweise: „x steht in Relation R zu y.“

Beispiele: Wir betrachten binäre Relationen über der Menge $A = \mathbb{N}$.

- $R_1 =_{\text{def}} \mathbb{N} \times \mathbb{N}$
- $R_2 =_{\text{def}} \{(0,0), (2,3), (5,1), (5,3)\}$

- $R_3 =_{\text{def}} \{\ (n_1, n_2) \mid n_1 \leq n_2\ \} = \{(0,0),(0,1),(1,1),(0,2),\ldots\}$
- $R_4 =_{\text{def}} \{\ (n_1, n_2) \mid n_1 \text{ teilt } n_2\ \} = \{(1,2),(2,4),(2,6),(7,0),\ldots\}$
- $R_5 =_{\text{def}} \{\ (n_1, n_2) \mid 2 \text{ teilt } |n_1 - n_2|\ \} = \{(0,2),(2,2),(1,1),(3,1),\ldots\}$
- $R_6 =_{\text{def}} \{\ (n_1, n_2) \mid 2n_1 = n_2\ \} = \{(0,0),(1,2),(2,4),(3,6),\ldots\}$

Die Relationen R_3 und R_4 sind *Ordnungsrelationen*. Relation R_5 ist eine *Äquivalenzrelation*. Relation R_6 ist eine *Funktion*.

In den folgenden Abschnitten wenden wir uns den im Beispiel erwähnten Relationentypen systematisch zu.

3.2 Funktionen

3.2.1 Totalität und Eindeutigkeit

In diesem Abschnitt führen wir Begriffe ein, die Funktionen, oder synomym Abbildungen, als spezielle Relationen zwischen Mengen von Argumenten und Mengen von Werten charakterisieren.

Definition 3.1 *Eine binäre Relation $R \subseteq A \times B$ heißt*

1.	linkstotal	$\iff_{\text{def}}$	$(\forall x \in A)(\exists y \in B)[(x,y) \in R]$
2.	rechtseindeutig	$\iff_{\text{def}}$	$(\forall x \in A)(\forall y,z \in B)[((x,y) \in R \wedge (x,z) \in R) \rightarrow y = z]$
3.	rechtstotal	$\iff_{\text{def}}$	$(\forall y \in B)(\exists x \in A)[(x,y) \in R]$
4.	linkseindeutig	$\iff_{\text{def}}$	$(\forall x,y \in A)(\forall z \in B)[((x,z) \in R \wedge (y,z) \in R) \rightarrow x = y]$

Beispiele: Wir überprüfen die Eigenschaften für die folgenden endlichen Relationen über den Mengen $A = \{1, 2, 3\}$ und $B = \{1, 2, 3, 4\}$:

Relation	linkstotal	rechtseindeutig	rechtstotal	linkseindeutig
$\{ (1,1), (1,2), (2,2) \}$ 				
$\{ (1,1), (1,2), (2,2), (3,3) \}$ 	X			
$\{ (1,1), (2,1) \}$ 		X		
$\{ (1,1), (1,2), (1,3), (1,4), (2,4) \}$ 			X	
$\{ (1,1), (1,2) \}$ 				X
$\{ (1,1), (1,2), (2,2), (3,3), (3,4) \}$ 	X		X	
$\{ (1,1), (2,2), (3,2) \}$ 	X	X		
$\{ (1,1), (1,2), (1,3), (1,4) \}$ 			X	X
$\{ (1,1), (2,2), (3,3) \}$ 	X	X		X

Definition 3.2 *Es sei $R \subseteq A \times B$ eine binäre Relation.*

1. *R heißt (totale)* Funktion*, falls R linkstotal und rechtseindeutig ist.*
2. *R heißt* partielle Funktion*, falls R rechtseindeutig ist.*

Beispiele: Wir diskutieren an folgenden Relationen die Funktionenbegriffe.

- Die Relation $R =_{\text{def}} \{ (1,1), (2,2), (3,2) \} \subseteq \{1, 2, 3\} \times \{1, 2, 3, 4\}$ ist eine Funktion.

- Die Relation $R =_{\text{def}} \{ (1,1),(2,1) \} \subseteq \{1,2,3\} \times \{1,2,3,4\}$ ist eine partielle Funktion. Fassen wir R jedoch als Teilmenge von $\{1,2\} \times \{1,2,3,4\}$ auf, so ist R eine Funktion.
- Die Relation $R =_{\text{def}} \{ (x,y) \mid y = |x| \} \subseteq \mathbb{Z} \times \mathbb{N}$ ist eine Funktion.
- Die Relation $R =_{\text{def}} \{ (y,x) \mid y = |x| \} \subseteq \mathbb{N} \times \mathbb{Z}$ ist keine Funktion.
- Die folgende Methode einer in Java implementierten Klasse

```
int gcd(int x, int y) {
    if (y==0) return x;
    if (y>x) return gcd(y,x);
    return gcd(y,x%y);
}
```

  ist eine partielle Funktion als Teilmenge von $\texttt{int}^2 \times \texttt{int}$, wobei wir `gcd` als Relation $\{ (x,y,z) \mid z = \texttt{gcd}(x,y) \}$ auffassen.

  In Java gilt $\text{mod}(-1,-2) = -1$, d.h. $(-1)\ \%\ (-2)$ wird zu -1 ausgewertet. Damit wird beim Methodenaufruf $\texttt{gcd}(-1,-2)$ erst rekursiv $\texttt{gcd}(-2,-1)$ und dann wieder $\texttt{gcd}(-1,-2)$ aufgerufen. Somit terminiert $\texttt{gcd}(-1,-2)$ nicht, und es gibt folglich kein $z \in \texttt{int}$ mit $(-1,-2,z) \in \texttt{gcd}$. Die Methode ist also nicht linkstotal. Die Rechtseindeutigkeit ist gegeben (wenn der verwendete Java-Compiler und die verwendete *Java Virtual Machine* korrekt sind).

Bisher haben wir Funktionen als binäre Relationen $R \subseteq A \times B$ eingeführt und damit streng genommmen lediglich einstellige Funktionen definiert. Dies ist jedoch keine inhaltliche Einschränkung, da die Mengen A und B hinreichend kompliziert werden können. Dennoch vereinbaren wir Folgendes: Es sei $R \subseteq A_1 \times \cdots \times A_n$ eine n-stellige Relation. Dann heißt R eine *k-stellige Funktion*, falls die binäre Relation $R \subseteq B \times C$ mit $B = A_1 \times \cdots \times A_k$ und $C = A_{k+1} \times \cdots \times A_n$ eine Funktion ist. Die Begriffsbildung für partielle Funktionen überträgt sich entsprechend.

Für Funktionen werden üblicherweise eigene Schreibweisen verwendet (wie im letzten der obigen Beispiele):

- Funktionen werden häufig klein geschrieben: $f \subseteq A \times B$.
- Statt $f \subseteq A \times B$ schreiben wir auch $f : A \to B$; statt $(a,b) \in f$ schreiben wir auch $f(a) = b$.
- Kompakt notieren wir eine Funktion als $f : A \to B : a \mapsto f(a)$; für den dritten Fall in den obigen Beispielen schreiben wir also z.B. $f : \mathbb{Z} \to \mathbb{N} : x \mapsto |x|$.

3.2.2 Bild- und Urbildmengen

Wichtige Begriffe zur Beschreibung der Eigenschaften von Funktionen sind die Bild- und Urbildmengen.

Definition 3.3 *Es seien $f : A \to B$ eine Funktion, $A_0 \subseteq A$ und $B_0 \subseteq B$.*

1. *Die Menge $f(A_0) \subseteq B$ ist definiert als*

$$f(A_0) =_{\text{def}} \{\ b \mid (\exists a \in A_0)[f(a) = b]\ \} \qquad _{\text{def}} = \{\ f(a) \mid a \in A_0\ \}$$

   und heißt Bild(menge) *von A_0 unter f. Die Elemente von $f(A_0)$ heißen* Bilder *von A_0 unter f.*

2. *Die Menge $f^{-1}(B_0) \subseteq A$ ist definiert als*

$$f^{-1}(B_0) =_{\text{def}} \{\ a \mid (\exists b \in B_0)[f(a) = b]\ \} \quad _{\text{def}} = \{\ a \mid f(a) \in B_0\ \}$$

   und heißt Urbild(menge) *von B_0 unter f. Die Elemente von $f^{-1}(B_0)$ heißen* Urbilder *von B_0 unter f.*

Beispiele: Wir verdeutlichen Bilder und Urbilder exemplarisch.

- Es sei die Funktion $f =_{\text{def}} \{\ (1,1),(2,2),(3,2)\ \} \subseteq \{1,2,3\} \times \{1,2,3,4\}$ gegeben. Unter anderem können folgende Bildmengen gebildet werden:

$$\begin{aligned} f(\{1\}) &= \{1\} \\ f(\{1,2\}) &= \{1,2\} \\ f(\{1,2,3\}) &= \{1,2\} \end{aligned}$$

  Beispiele für Urbildmengen sind unter anderem:

$$\begin{aligned} f^{-1}(\{1\}) &= \{1\} \\ f^{-1}(\{1,2\}) &= \{1,2,3\} \\ f^{-1}(\{3\}) &= \emptyset \end{aligned}$$

- Es sei die Funktion $f : \mathbb{Z} \to \mathbb{N} : x \mapsto |x|$ gegeben. Die Bildmenge das ganzzahligen Intervalls $[-1,1]$ unter f ist:

$$f([-1,1]) = f(\{-1,0,1\}) = \{0,1\}$$

  Die Urbildmengen zu $\{2\}$ und $[2,4]$ unter f sind wie folgt:

$$\begin{aligned} f^{-1}(\{2\}) &= \{-2,2\} \\ f^{-1}([2,4]) &= \{-4,-,3,-2,2,3,4\} \end{aligned}$$

Proposition 3.4 *Es seien A und B endliche Mengen und $f : A \to B$ eine Funktion. Dann gilt:*

$$\|A\| = \sum_{b \in B} \left\| f^{-1}(\{b\}) \right\|$$

Beweis: Da f eine Funktion ist, bildet die Mengenfamilie $\{\ f^{-1}(\{b\}) \mid b \in B\ \}$ eine Partition von A. Damit folgt

$$\|A\| = \sum_{b \in B} \|f^{-1}(\{b\})\|$$

und die Proposition ist bewiesen. ■

3.2.3 Injektivität, Surjektivität und Bijektivität

Funktionen werden danach klassifiziert, welche Eigenschaften sie zusätzlich zur Linkstotalität und Rechtseindeutigkeit erfüllen.

Definition 3.5 *Eine Funktion $f : A \to B$ heißt*

1. surjektiv $\iff_{\text{def}}$ *f ist rechtstotal*
2. injektiv $\iff_{\text{def}}$ *f ist linkseindeutig*
3. bijektiv $\iff_{\text{def}}$ *f ist rechtstotal und linkseindeutig*

Beispiele: Folgende Funktionen verdeutlichen die Begriffsbildung.

- Die Funktion $f : \mathbb{Z} \to \mathbb{N} : x \mapsto |x|$ ist surjektiv, aber nicht injektiv.
- Die Funktion $f : \mathbb{N} \to \mathbb{Z} : x \mapsto x^3$ ist injektiv, aber nicht surjektiv.
- Die Funktion $f : \mathbb{Z} \to \mathbb{Z} : x \mapsto -x$ ist bijektiv.

Das folgende Lemma ergibt sich unmittelbar aus den Definition der Funktioneneigenschaften. Der Beweis bleibt dem Leser zur Übung überlassen.

Lemma 3.6 *Es sei $f : A \to B$ eine Funktion. Dann gilt:*

1. *f ist surjektiv* $\iff (\forall b \in B)\left[\ \left\|f^{-1}(\{b\})\right\| \geq 1\ \right]$
2. *f ist injektiv* $\iff (\forall b \in B)\left[\ \left\|f^{-1}(\{b\})\right\| \leq 1\ \right]$
3. *f ist bijektiv* $\iff (\forall b \in B)\left[\ \left\|f^{-1}(\{b\})\right\| = 1\ \right]$

Während das vorangegangene Lemma eine Charaktiersierung der Eigenschaften für eine konkrete Funktion angibt, stellt Theorem 3.7 eine Beziehung zwischen Mengen mit Hilfe von Funktioneneigenschaften her. Das durch das Theorem beschriebene Abzählprinzip ist eine fundamentale Technik beim Lösen kombinatorischer Fragestellungen.

Theorem 3.7 *Es seien A und B nicht-leere, endliche Mengen. Dann gilt:*

1. *Es gibt eine surjektive Funktion $f : A \to B$* $\iff$ $\|A\| \geq \|B\|$
2. *Es gibt eine injektive Funktion $f : A \to B$* $\iff$ $\|A\| \leq \|B\|$
3. *Es gibt eine bijektive Funktion $f : A \to B$* $\iff$ $\|A\| = \|B\|$

Beweis: Wir beweisen die Äquivalenzen im Block.

($\Leftarrow$): Es seien $A = \{a_1, \ldots, a_n\}$ und $B = \{b_1, \ldots, b_m\}$ endliche Mengen. Wir definieren eine Funktion $f : A \to B$ wie folgt für $a_i \in A$:

$$f(a_i) =_{\text{def}} \begin{cases} b_i & \text{falls } i \leq m \\ b_1 & \text{falls } i > m \end{cases}$$

Dann gelten folgende Aussage in Abhängigkeit von A und B:

1. Ist $\|A\| \geq \|B\|$, d.h. $n \geq m$, so ist f surjektiv
2. Ist $\|A\| \leq \|B\|$, d.h. $n \leq m$, so ist f injektiv
3. Ist $\|A\| = \|B\|$, d.h. $n = m$, so ist f bijektiv

($\Rightarrow$): Es sei $f : A \to B$ eine Funktion. Dann gelten folgende Aussagen:

1. Ist f surjektiv, so gilt nach Proposition 3.4 und Lemma 3.6.1:

$$\|A\| = \sum_{b \in B} \left\| f^{-1}(\{b\}) \right\| \geq \sum_{b \in B} 1 = \|B\|$$

2. Ist f injektiv, so gilt nach Proposition 3.4 und Lemma 3.6.2:

$$\|A\| = \sum_{b \in B} \left\| f^{-1}(\{b\}) \right\| \leq \sum_{b \in B} 1 = \|B\|$$

3. Ist f bijektiv, so gilt nach Proposition 3.4 und Lemma 3.6.3:

$$\|A\| = \sum_{b \in B} \left\| f^{-1}(\{b\}) \right\| = \sum_{b \in B} 1 = \|B\|$$

Damit ist das Theorem bewiesen ∎

Theorem 3.8 *Es seien A und B endliche Mengen mit $\|A\| = \|B\| > 0$. Dann sind folgende Aussagen äquivalent:*

1. *f ist surjektiv*
2. *f ist injektiv*
3. *f ist bijektiv*

Die logische Struktur des Theorems besagt, dass entweder alle Aussagen gelten oder keine.

Beweis: Wir zeigen die paarweise Äquivalenz aller Aussagen über einzelne Implikationen.

- (3) $\Rightarrow$ (1): Ist f bijektiv, so ist f surjektiv (nach Definition).
- (3) $\Rightarrow$ (2): Ist f bijektiv, so ist f injektiv (nach Definition).
- (1) $\Rightarrow$ (3): Es sei f surjektiv, d.h. für alle $b \in B$ gilt $\|f^{-1}(\{b\})\| \geq 1$ (nach Lemma 3.6.1). Dann gilt nach Proposition 3.4 und der Voraussetzung:

  $$\|A\| = \sum_{b \in B} \|f^{-1}(\{b\})\| \geq \|B\| = \|A\|$$

  Somit gilt $\|f^{-1}(\{b\})\| = 1$ für alle $b \in B$. Folglich ist f bijektiv (nach Lemma 3.6.3).
- (2) $\Rightarrow$ (3): Es sei f injektiv, d.h. für alle $b \in B$ gilt $\|f^{-1}(\{b\})\| \leq 1$ (nach Lemma 3.6.2). Dann gilt nach Proposition 3.4 und der Voraussetzung:

  $$\|A\| = \sum_{b \in B} \|f^{-1}(\{b\})\| \leq \|B\| = \|A\|$$

  Somit gilt $\|f^{-1}(\{b\})\| = 1$ für alle $b \in B$. Folglich ist f bijektiv (nach Lemma 3.6.3).

Damit ist das Theorem bewiesen. ■

3.2.4 Invertierbarkeit

Für eine Relation $R \subseteq A \times B$ definieren wir die *Umkehrrelation* $R^{-1} \subseteq B \times A$ wie folgt:

$$R^{-1} =_{\text{def}} \{\ (y, x) \mid (x, y) \in R\ \}$$

Die folgende Proposition ist einfach an Hand der Definitionen einzusehen.

Proposition 3.9 *Es sei R eine binäre Relation. Dann gelten die folgenden Aussagen:*

1. *R ist linkstotal* $\iff$ *R^{-1} ist rechtstotal*
2. *R ist rechtseindeutig* $\iff$ *R^{-1} ist linkseindeutig*
3. *R ist rechtstotal* $\iff$ *R^{-1} ist linkstotal*
4. *R ist linkseindeutig* $\iff$ *R^{-1} ist rechtseindeutig*

Korollar 3.10 *Ist f eine bijektive Funktion, so ist die Umkehrrelation f^{-1} eine bijektive Funktion.*

Definition 3.11 *Eine Funktion f heißt* invertierbar *(umkehrbar), falls die Umkehrrelation f^{-1} eine Funktion ist.*

Korollar 3.12 *Eine Funktion f ist genau dann invertierbar, wenn f bijektiv ist.*

3.2.5 Hintereinanderausführung

Eine wichtige Operation auf Funktionen ist die *Hintereinanderausführung* (oder auch *Verkettung*, *Superposition* oder *Komposition* in anderen Zusammenhängen): Für Funktionen $f : A \to B$ und $g : B \to C$ definieren wir die Funktion $g \circ f : A \to C$ wie folgt für alle $x \in A$:

$$(g \circ f)(x) =_{\text{def}} g(f(x))$$

Beispiele: Wir beleuchten im Folgenden Aspekte der Hintereinanderausführung exemplarisch.

- Für die beiden Funktionen $f : \mathbb{N} \times \mathbb{N} : x \mapsto x^2$ und $g : \mathbb{N} \times \mathbb{N} : x \mapsto 2^x$ gilt

$$\begin{aligned}(g \circ f)(x) &= g(f(x)) = g(x^2) = 2^{(x^2)} = 2^{x^2}\\ (f \circ g)(x) &= f(g(x)) = f(2^x) = (2^x)^2 = 2^{2x}\end{aligned}$$

  Mithin gilt $g \circ f \neq f \circ g$, denn wir erhalten $(g \circ f)(3) = 2^9 = 512$ und $(f \circ g)(3) = 2^6 = 64$.

- Wodurch unterscheiden sich Klassen- und Instanzenmethoden in Java (ohne Nebeneffekte) mathematisch? Zur Veranschaulichung sei dazu eine Methode `method` einerseits als Klassenmethode

```
public static int method (int x, int y)
```

und andererseits als Instanzenmethode

```
public int method (int x, int y)
```

deklariert. Im ersten Fall beschreibt die Methode eine Funktion

$$\texttt{method} : \texttt{int} \times \texttt{int} \to \texttt{int}.$$

Im zweiten Fall dagegen wird eine Funktion

$$\texttt{method} : S \times \texttt{int} \times \texttt{int} \to \texttt{int}$$

beschrieben, wobei S für die Menge der verfügbaren Speicheradressen steht. Bei der Instanziierung eines Objektes `obj` aus der entsprechenden Klasse ordnet die *Java Virtual Machine* eine Adresse $s(\texttt{obj}) \in S$ zu, d.h. die Instanzenmethode wird dann zu einer Funktion

$$\texttt{obj.method} : \texttt{int} \times \texttt{int} \to \texttt{int} : (x, y) \mapsto \texttt{method}(s(\texttt{obj}), x, y)$$

als Hintereinanderausführung der Funktionen s und `method`.

Proposition 3.13 *Es seien $f : A \to B$ und $g : B \to C$ beliebige Funktionen.*

1. *Sind f und g injektiv, so ist $g \circ f$ injektiv.*
2. *Sind f und g surjektiv, so ist $g \circ f$ surjektiv.*
3. *Sind f und g bijektiv, so ist $g \circ f$ bijektiv.*

Beweis: Wir zeigen die Aussagen einzeln.

1. Es seien f und g injektive Funktionen. Wir müssen zeigen, dass $g \circ f$ linkseindeutig ist. Dazu seien $x, y \in A$ beliebig mit $(g \circ f)(x) = (g \circ f)(y) \in C$. Da g injektiv ist, folgt aus $g(f(x)) = g(f(y))$ die Gleichheit $f(x) = f(y)$. Da auch f injektiv ist, folgt aus $f(x) = f(y)$ wiederum die Gleichheit $x = y$. Mithin ist $g \circ f$ linkseindeutig und also injektiv.

2. Es seien f und g surjektive Funktionen. Wir müssen zeigen, dass $g \circ f$ rechtstotal ist. Es sei $x \in C$ beliebig. Da g surjektiv ist, gibt es ein $y \in B$ mit $y \in g^{-1}(\{x\}) \subseteq B$, d.h. $g(y) = x$. Da auch f surjektiv ist, gibt es ein $z \in A$ mit $z \in f^{-1}(\{y\}) \subseteq A$, d.h. $f(z) = y$. Insgesamt erhalten wir also

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

   Somit gilt $\|(g \circ f)^{-1}(\{x\})\| \geq 1$ für alle $x \in C$. Mithin ist $g \circ f$ surjektiv (nach Lemma 3.6.1).

3. Direkte Folgerung aus der ersten und der zweiten Aussage dieser Proposition.

Damit ist die Proposition bewiesen. ■

Für eine Menge A heißt die Funktion $\mathrm{id}_A : A \to A : x \mapsto x$ *Identitätsfunktion* von A.

Proposition 3.14 *Es sei $f : A \to B$ eine bijektive Funktion. Dann gilt $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$.*

Beweis: Es genügt $f^{-1} \circ f = \mathrm{id}_A$ zu zeigen (da wir f und f^{-1} vertauschen können). Es gilt $f^{-1} \circ f : A \to A$ wegen $f : A \to B$ und $f^{-1} : B \to A$. Außerdem gilt $f^{-1}(\{f(x)\}) = \{x\}$, da f bijektiv ist. Somit gilt $f^{-1}(f(x)) = x$ für alle $x \in A$, d.h. $f^{-1} \circ f = \mathrm{id}_A$. Damit ist die Proposition bewiesen. ■

3.3 Äquivalenzrelationen

3.3.1 Reflexivität, Transitivität und Symmetrie

Definition 3.15 *Eine binäre Relation $R \subseteq A \times A$ heißt*

1.	reflexiv	$\iff_{\mathrm{def}}$	$(\forall a \in A)[(a,a) \in R]$
2.	transitiv	$\iff_{\mathrm{def}}$	$(\forall a,b,c \in A)[((a,b) \in R \wedge (b,c) \in R) \to (a,c) \in R]$
3.	symmetrisch	$\iff_{\mathrm{def}}$	$(\forall a,b \in A)[(a,b) \in R \to (b,a) \in R]$
4.	Äquivalenzrelation	$\iff_{\mathrm{def}}$	*R ist reflexiv, transitiv und symmetrisch*

Bei einer Äquivalenzrelation R verwenden wir statt $(a, b) \in R$ die Infix-Schreibweise $a \sim_R b$ (oder $a \approx_R b$, $a \equiv_R b$).

Beispiele: Wir überprüfen die Eigenschaften für die folgenden endlichen Relationen über der Menge $A = \{0, 1, 2\}$:

Relation	reflexiv	transitiv	symmetrisch	Äquivalenzrelation
$\{ (0,1), (1,0), (0,2), (2,0) \}$			X	
$\{ (0,0), (0,1), (1,0), (1,1), (0,2), (2,0), (2,2) \}$	X		X	
$\{ (0,0), (0,1), (1,0), (1,1), (2,2) \}$	X	X	X	X
$\{ (0,0), (0,1), (1,0), (1,1) \}$		X	X	
$\{ (0,0), (0,1), (1,1), (2,2) \}$	X	X		

Wir wollen die Intuitivität des Äquivalenzrelationenbegriff an komplexeren Relationen verdeutlichen.

Beispiele: Folgende Beispiele sind typisch für die Bildung von Äquivalenzrelationen.

- Es seien $A =_{\text{def}}$ Menge aller (logischen) Aussagen und

$$R =_{\text{def}} \{ (H, H') \mid H \leftrightarrow H' \text{ ist eine Tautologie } \} \subseteq A \times A.$$

  Dann ist R eine Äquivalenzrelation, denn es gelten folgende Aussagen (z.B. mittels Überprüfung durch Wertetabellen):
  - R ist reflexiv: $H \leftrightarrow H$ ist eine Tautologie für alle Aussagen H
  - R ist transitiv: Sind $H \leftrightarrow H'$ und $H' \leftrightarrow H''$ Tautologien, so ist auch $H \leftrightarrow H''$ eine Tautologie (wegen doppelter Anwendung der Kettenschlussregel)

  - R ist symmetrisch: Ist $H \leftrightarrow H'$ eine Tautologie, so ist auch $H' \leftrightarrow H$ eine Tautologie.

- Es sei $f : A \to B$ eine beliebige Funktion (mit Argumenten aus A und Funktionswerten in B). Dann ist die Relation

$$R_f =_{\text{def}} \{\ (x, y) \mid f(x) = f(y)\ \} \subseteq A \times A$$

ganz offensichtlich eine Äquivalenzrelation. Zum Beispiel ergeben sich für spezielle Funktionen folgende Äquivalenzrelationen:

  - Auf der Menge $A =_{\text{def}} \mathbb{Z}$ sei die Funktion $f_n(x) = \text{mod}(x, n)$ mit Funktionswerten in der Menge $\{0, 1, \ldots, n-1\}$ definiert. Dann schreiben wir auch $x \equiv y \bmod n$ für $(x, y) \in R_{f_n}$ und sagen „x ist kongruent y modulo n“.
  - Auf der Menge A aller Wörter eines Wörterbuches (wobei alle Wörter nur aus Kleinbuchstaben bestehen und keine Umlaute enthalten) sei f als Funktion definiert, die jedes Wort auf den ersten Buchstaben abbildet.

3.3.2 Äquivalenzklassen

Definition 3.16 *Es seien $R \subseteq A \times A$ eine Äquivalenzrelation und $x \in A$ ein beliebiges Element. Dann heißt die Menge*

$$[x]_R =_{\text{def}} \{\ y \mid (x, y) \in R\ \} \subseteq A$$

Äquivalenzklasse *von x. Wir nennen x* Repräsentant *der Äquivalenzklasse.*

Beispiel: Wir betrachten die Kongruenz „ $\equiv \bmod 8$ “ auf den ganzen Zahlen. Dann gilt:

$$\begin{aligned} [13]_{\equiv} &= \{\ y \mid y \equiv 13 \bmod 8\ \} \\ &= \{\ y \mid \text{mod}(y - 13, 8) = 0\ \} \\ &= \{\ldots, -11, -3, 5, 13, 21 \ldots\ \} \\ &= [5]_{\equiv} \end{aligned}$$

Proposition 3.17 *Es seien $R \subseteq A \times A$ eine Äquivalenzrelation und $x, y \in A$. Dann gilt:*

1. *Ist $(x, y) \in R$, so gilt $[x]_R = [y]_R$.*
2. *Ist $(x, y) \notin R$, so sind $[x]_R$ und $[y]_R$ disjunkt.*

Beweis: Wir beweisen die Aussagen einzeln.

1. Es gelte $(x, y) \in R$, d.h. $y \in [x]_R$. Wegen der Transitivität von R gilt $(x, z) \in R$ für alle $z \in [y]_R$ (d.h. $(y, z) \in R$). Somit gilt $[y]_R \subseteq [x]_R$. Wegen der Symmetrie von R gilt $(y, x) \in R$. Somit können wir analog auch $[x]_R \subseteq [y]_R$ zeigen. Mithin gilt $[x]_R = [y]_R$.

2. Wir zeigen die Kontraposition der Aussage. Dazu gelte $[x]_R \cap [y]_R \neq \emptyset$. Dann gibt es ein $z \in A$ mit $z \in [x]_R$ und $z \in [y]_R$ bzw. $(x, z) \in R$ und $(y, z) \in R$. Wegen der Symmetrie von R gilt $(z, y) \in R$. Wegen der Transitivität gilt somit $(x, y) \in R$.

Damit ist die Proposition bewiesen. ■

3.3.3 Repräsentantensysteme

Definition 3.18 *Es sei $R \subseteq A \times A$ eine Äquivalenzrelation. Eine Menge $K \subseteq A$ heißt* Repräsentantensystem *von R, falls folgende Bedingungen erfüllt sind:*

1. *Für alle $k_1, k_2 \in K$ mit $k_1 \neq k_2$ gilt $(k_1, k_2) \notin R$*
2. $A = \bigcup_{k \in K} [k]_R$

Beispiel: Wir betrachten die Kongruenz „ $\equiv \mod 8$ “ auf den ganzen Zahlen.

- $\{0, 1, 2, 3, 4, 5, 6, 7\}$ ist ein Repräsentantensystem
- $\{8, 1, 2, 19, -4, 13, 6, 7\}$ ist ebenfalls ein Repräsentantensystem

Die zu einem Repräsentantensystem gehörenden Äquivalenzklassen bilden eine Partition der Grundmenge.

Korollar 3.19 *Es seien $R \subseteq A \times A$ eine Äquivalenzrelation und $K \subseteq A$ ein Repräsentantensystem von R. Dann bilden die Äquivalenzklassen (der Elemente) von K eine Partition von A.*

Beweis: Wegen $(k_1, k_2) \notin R$ für $k_1, k_2 \in K$ mit $k_1 \neq k_2$ (die erste Eigenschaft eines Repäsentantensystems) folgt aus Proposition 3.17:

$$[k_1]_R \cap [k_2]_R = \emptyset$$

Aus der zweiten Eigenschaft eines Repräsentantensystem folgt für K weiterhin

$$\bigcup_{k \in K} [k]_R = A.$$

Somit ist die Mengenfamilie $\{ [k]_R \mid k \in K \}$ eine Partition von A. Damit ist das Korollar bewiesen. ■

Proposition 3.20 *Es sei $\mathcal{F} \subseteq \mathcal{P}(A)$ eine Partition von A. Dann ist die Relation $R \subseteq A \times A$ mit*

$$(x, y) \in R \iff_{\text{def}} (\exists X \in \mathcal{F})[x \in X \wedge y \in X]$$

eine Äquivalenzrelation.

Beweis: Wir überprüfen die Eigenschaften von Äquivalenzrelationen:

- R ist reflexiv: Für jedes $x \in A$ gibt es ein $X \in \mathcal{F}$ mit $x \in X$, da $\mathcal{F}$ eine Partition ist. Somit gilt $(x, x) \in R$.
- R ist transitiv: Es seien $(x, y) \in R$ und $(y, z) \in R$. Dann gibt es $X_1, X_2 \in \mathcal{F}$ mit $x, y \in X_1$ sowie $y, z \in X_2$. Mithin gilt $y \in X_1 \cap X_2$. Also sind X_1 und X_2 nicht disjunkt. Da $\mathcal{F}$ eine Partition ist, gilt folglich $X_1 = X_2$. Somit gilt $x, z \in X_1$. Es folgt $(x, z) \in R$.
- R ist symmetrisch: Ist $(x, y) \in R$, so gilt $x, y \in X$ für ein geeignetes $X \in \mathcal{F}$. Also gilt auch $(y, x) \in R$.

Damit ist die Proposition bewiesen ■

3.4 Ordnungsrelationen

3.4.1 Antisymmetrie und Linearität

Ordnungsrelationen extrahieren den mathematischen Gehalt von natürlichen Ordnungen, wie sie beispielsweise beim Sortieren benötigt werden. Dafür werden die folgenden zusätzlichen Begriffe definiert.

Definition 3.21 *Eine binäre Relation $R \subseteq A \times A$ heißt*

1. antisymmetrisch $\iff_{\text{def}} (\forall a, b \in A)[((a, b) \in R \wedge (b, a) \in R) \rightarrow a = b]$
2. linear $\iff_{\text{def}} (\forall a, b \in A)[a \neq b \rightarrow ((a, b) \in R \vee (b, a) \in R)]$

Die Eigenschaft der Antisymmetrie wird anschaulicher, wenn für alle $a, b \in A$ die Kontraposition

$$a \neq b \rightarrow ((a, b) \notin R \vee (b, a) \notin R)$$

betrachtet wird. Mit anderen Worten darf für verschiedene Elemente a und b höchstens eines der Paare (a, b) oder (b, a) zu R gehören. Zu beachten ist weiterhin, dass die Eigenschaft der Antisymmetrie nicht die Negation der Symmetrie ist, wie sie im Kapitel über Äquivalenzrelationen eingeführt wurde.

Beispiele: Wir überprüfen die Eigenschaften für die folgenden endlichen Relationen über der Menge $A = \{0, 1, 2\}$:

Relation	reflexiv	transitiv	antisymmetrisch	total
$\{ (0,0), (0,1), (0,2), (1,1), (1,2), (2,2) \}$	X	X	X	X
$\{ (0,0), (0,1), (2,0), (1,1), (1,2), (2,2) \}$	X		X	X
$\{ (0,0), (0,1), (2,0), (0,2), (1,1), (1,2), (2,2) \}$	X			X
$\{ (0,0), (0,1), (2,0), (0,2), (1,1), (2,2) \}$	X			
$\{ (0,1), (2,0), (0,2), (1,1), (2,2) \}$				
$\{ (0,1), (1,2), (0,2) \}$		X	X	X
$\{ (0,0), (0,1), (1,0), (1,1), (2,2) \}$	X	X		
$\{ (0,0), (0,1), (0,2), (1,1), (2,2) \}$	X	X	X	
$\{ (0,1), (1,2), (2,1), (2,0) \}$				X

Durch die Analyse der obigen Beispiele bekommt man ein technisches Gefühl für die Definitionen. Im Folgenden wollen wir auch die Intuitivität von Definition 3.21 durch weitere Beispiele verdeutlichen.

Beispiele: Die folgenden Beispiele repräsentieren im Allgemeinen unendliche Relationen.

- Wir betrachten die Relation $R_1 =_{\text{def}} \{ (m,n) \mid m \leq n \} \subseteq \mathbb{N}^2$. Zur Erinnerung halten wir fest: $m \leq n \Leftrightarrow (\exists c \in \mathbb{N})[n = m + c]$. Dann besitzt R alle Eigenschaften von Definition 3.21:
  - R_1 ist reflexiv, denn für alle $n \in \mathbb{N}$ gilt $n = n + 0$ bzw. $n \leq n$.
  - R_1 ist transitiv, denn gilt $k \leq m$ und $m \leq n$, so gibt es $c_1, c_2 \in \mathbb{N}$ mit $m = k + c_1$ sowie $n = m + c_2$ und es gilt $n = k + (c_2 + c_1)$ bzw. $k \leq n$.
  - R_1 ist antisymmetrisch, denn gilt $m \leq n$ und $n \leq m$, so gibt es $c_1, c_2 \in \mathbb{N}$ mit $n = m + c_1$ sowie $m = n + c_2$ und mit $n = n + c_1 + c_2$ folgt $c_1 = c_2 = 0$ und mithin $n = m$.
  - R_1 ist total, denn $n - m \in \mathbb{N}$ oder $m - n \in \mathbb{N}$.
- Wir betrachten die Relation $R_2 =_{\text{def}} \{ (m,n) \mid m \text{ teilt } n \}$. Auch hier halten wir zur Erinnerung fest: m teilt $n \Leftrightarrow (\exists c \in \mathbb{N})[n = c \cdot m]$. Für R_2 gelten folgende Aussagen:
  - R_2 ist reflexiv, denn für alle $n \in \mathbb{N}$ gilt $n = 1 \cdot n$ bzw. n teilt n.
  - R_2 ist transitiv, denn teilt k die Zahl m und teilt m die Zahl n, so gibt es $c_1, c_2 \in \mathbb{N}$ mit $m = c_1 \cdot k$ sowie $n = c_2 \cdot m$ und es gilt $n = (c_2 \cdot c_1) \cdot k$ bzw. k teilt n.
  - R_2 ist antisymmetrisch, denn teilt m die Zahl n und teilt n die Zahl m, so gibt es $c_1, c_2 \in \mathbb{N}$ mit $n = c_1 \cdot m$ sowie $m = c_2 \cdot n$ und mit $n = c_1 \cdot c_2 \cdot n$ folgt $c_1 = c_2 = 1$ und mithin $m = n$.
  - R_2 ist nicht total, denn weder teilt 2 die Zahl 3 noch teilt 3 die Zahl 2.
- Wir betrachten die Relation $R_3 =_{\text{def}} \{ (A,B) \mid A \subseteq B \} \subseteq \mathcal{P}(X)^2$ für eine Grundmenge X. Für R gelten folgende Eigenschaften:
  - R_3 ist reflexiv, denn es gilt $A \subseteq A$ für alle $A \subseteq X$.
  - R_3 ist transitiv, denn gilt $A \subseteq B$, d.h. $(\forall a \in A)[a \in B]$, und gilt $B \subseteq C$, d.h. $(\forall a \in B)[a \in C]$, so gilt nach dem Kettenschluss auch $(\forall a \in A)[a \in C]$, d.h. $A \subseteq C$.
  - R_3 ist antisymmetrisch, denn mit $A \subseteq B$ und $B \subseteq A$ gilt $A = B$.
  - R_3 ist nicht total, falls $\|X\| \geq 2$: Es seien $a, b \in X$ mit $a \neq b$, dann gilt $\{a\} \cap \{b\} = \emptyset$.

Definition 3.22 *Es sei $R \subseteq A \times A$ eine binäre Relation über A.*

1. *R heißt* Halbordnung *(oder* partielle Ordnung*), falls R reflexiv, transitiv und antisymmetrisch ist.*
2. *R heißt* Ordnung *(oder* totale Ordnung*), falls R eine Halbordnung und zusätzlich total ist.*
3. *Ist R eine Halbordnung, so heißt das Paar (A, R)* halbgeordnete *(oder* partiell geordnete*)* Menge*.*
4. *Ist R eine Ordnung, so heißt das Paar (A, R)* geordnete *(oder* total geordnete*)* Menge*.*

Beispiele (Fortsetzung): Für die drei Relationen aus obigem Beispiel gilt:

- R_1 ist eine Ordnung; wir schreiben die geordnete Menge als $(\mathbb{N}, \leq)$.
- R_2 ist eine Halbordnung; wir schreiben die halbgeordnete Menge als $(\mathbb{N}, |)$.
- R_3 ist eine Halbordnung für jede Grundmenge X; wir schreiben die halbgeordnete Menge als $(\mathcal{P}(X), \subseteq)$.

3.4.2 HASSE-Diagramme

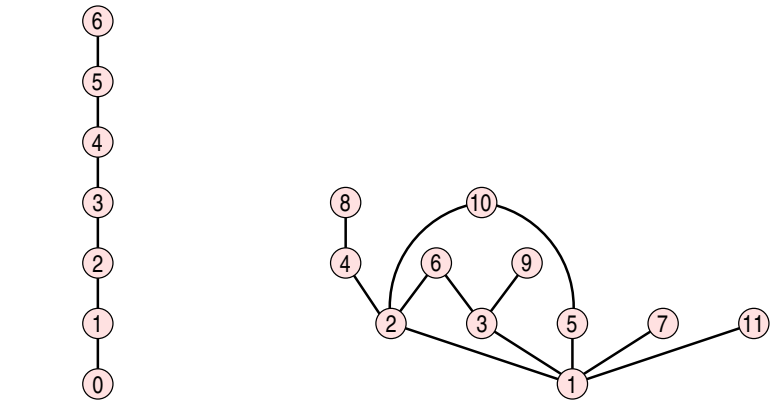
Endliche Halbordnungen lassen sich durch HASSE-Diagramme graphisch darstellen. Diese Diagramme sind wie folgt für eine halbgeordnete Menge (A, R) definiert:

- Elemente der Grundmenge A werden durch Punkte (Knoten) in der Ebene dargestellt
- Ist $(x, y) \in R$ für $x \neq y$, so wird der Knoten y oberhalb von Knoten x gezeichnet
- Genau dann, wenn $(x, y) \in R$ für $x \neq y$ gilt und es kein $z \notin \{x, y\}$ mit $(x, z) \in R$ und $(z, y) \in R$ gibt, werden x und y durch eine Linie (Kante) verbunden

Bei dieser Darstellungsform werden gerade alle Paare einer Halbordnung nicht mit dargestellt, deren Zugehörigkeit zur Relation sich wegen der Transitivität sowieso aus den

anderen Paaren ergeben würde. Eine derart vollständig reduzierte Relation heißt auch *transitive Reduktion* einer Halbordnung.

Beispiele: Die folgende Abbildung zeigt HASSE-Diagramme für die endlichen, halbgeordneten Mengen $\left(\{0,1,\ldots,6\}^2,\leq\right)$ und $\left(\{1,\ldots,11\}^2,|\right)$:



$\left(\{0,1,\ldots,6\}^2,\leq\right)$ $\left(\{1,\ldots,11\}^2,|\right)$

Im Folgenden verwenden wir $x \leq_R y$ für $(x,y) \in R$, falls R eine Halbordnung ist.

3.4.3 Minima und Maxima

Definition 3.23 *Es seien $R \subseteq A \times A$ eine Halbordnung und $K \subseteq A$.*

1. *Ein Element $a \in K$ heißt* Minimum *(bzw.* Maximum*) von K, falls $a \leq_R b$ (bzw. $a \geq_R b$) für alle $b \in K$ gilt.*
2. *Ein Element $a \in A$ heißt* untere Schranke *(bzw.* obere Schranke*) von K, falls $a \leq_R b$ (bzw. $a \geq_R b$) für alle $b \in K$ gilt.*
3. *Ein Element $a \in A$ heißt* Infimum *(bzw.* Supremum*) von K, falls a eine untere Schranke (bzw. obere Schranke) von K und $a \geq_R b$ (bzw. $a \leq_R b$) für alle unteren Schranken (bzw. oberen Schranken) von K gilt.*

Der Unterschied zwischen einer unteren Schranke von K und einem Minimum von K liegt darin, dass die untere Schranke nicht zur Menge K gehören muss, was für das Minimum verlangt ist. Gleiches gilt natürlich auch für obere Schranken von K und einem Maximum von K. Darüber hinaus sind Minima, Maxima, Infima und Suprema stets eindeutig, falls sie überhaupt existieren.

Proposition 3.24 *Es seien $R \subseteq A \times A$ eine Halbordnung und $K \subseteq A$. Existiert das Minimum (Maximum, Infimum, Supremum) von K, so ist es eindeutig.*

Beweis: (*nur für das Minimum*) Es seien $a, a' \in K$ Minima von K. Dann gilt $a \leq_R a'$, da a ein Minimum von K ist, und es gilt $a' \leq_R a$, da a' ein Minimum von K ist. Wegen der Antisymmetrie von $\leq_R$ gilt $a = a'$. Damit ist die Proposition bewiesen. ■

Die Eindeutigkeit dieser Elemente ermöglicht uns spezielle Notationen einzuführen:

$\min(K)$	steht für das Minimum von K
$\max(K)$	steht für das Maximum von K
$\inf(K)$	steht für das Infimum von K
$\sup(K)$	steht für das Supremum von K

Anschaulich ist das Infimum die größte untere Schranke und das Supremum die kleinste obere Schranke. Im Allgemeinen müssen Minimum, Maximum, Infimum und Supremum nicht existieren.

Beispiele: Folgende Beispiele verdeutlichen die Begriffsbildungen.

- $\min(\emptyset)$ und $\max(\emptyset)$ existieren für keine Halbordnung.
- Es sei $A =_{\text{def}} \mathbb{Q}$ und $R =_{\text{def}} \{ (m, n) \mid m \leq n \}$. Für die Mengen

$$\begin{aligned} K_+ &=_{\text{def}} \{ x \mid 0 < x \} \subseteq A \\ K_- &=_{\text{def}} \{ x \mid x < 0 \} \subseteq A \end{aligned}$$

  gelten die folgenden Aussagen:
  - $\min(K_+)$ und $\min(K_-)$ existieren nicht
  - $\max(K_+)$ und $\max(K_-)$ existieren nicht
  - Die Menge der unteren Schranken von K_+ ist $K_- \cup \{0\}$
  - Die Menge der unteren Schranken von K_- ist $\emptyset$
  - Die Menge der oberen Schranken von K_+ ist $\emptyset$
  - Die Menge der oberen Schranken von K_- ist $K_+ \cup \{0\}$
  - $\inf(K_+) = \max(K_- \cup \{0\}) = 0$
  - $\inf(K_-)$ existiert nicht
  - $\sup(K_+)$ existiert nicht
  - $\sup(K_-) = \min(K_+ \cup \{0\}) = 0$
- Wir setzen das vorangehende Beispiel fort. Bei veränderter Grundmenge $A =_{\text{def}} \mathbb{Q} \setminus \{0\}$ sowie unverändertem R, K_+ und K_- gelten die folgenden Aussagen:
  - Die Menge der unteren Schranken von K_+ ist K_-
  - $\inf(K_+)$ existiert nicht, da K_- kein Maximum besitzt
- Es seien $A =_{\text{def}} \{0, 1, \ldots, 10\}$ und $R =_{\text{def}} \{ (m, n) \mid m \leq n \} \subseteq A \times A$. Dann gelten folgende Aussagen:
  - $\inf(\emptyset) = 10$
  - $\sup(\emptyset) = 0$

3.4.4 Minimale und maximale Elemente

Definition 3.25 *Es seien $R \subseteq A \times A$ eine Halbordnung und $K \subseteq A$. Ein Element $a \in K$ heißt* minimal *(bzw.* maximal*) in K, falls für alle $b \in K$ gilt: Ist $b \leq_R a$ (bzw. $b \geq_R a$), so ist $a = b$.*

Beispiel: Es seien $A =_{\text{def}} \mathbb{N}$ und $R =_{\text{def}} \{ (m, n) \mid m \text{ teilt } n \} \subseteq A \times A$. Für

$$K_1 =_{\text{def}} \mathbb{N} \quad \text{und} \quad K_2 =_{\text{def}} \mathbb{N} \setminus \{1\}$$

gelten die Aussagen:

- Die Menge der minimalen Elemente von K_1 ist $\{1\}$
- Die Menge der minimalen Elemente von K_2 ist die Menge der Primzahlen

Proposition 3.26 *Es seien $R \subseteq A \times A$ eine Ordnung und $K \subseteq A$. Ist $a \in K$ minimal (bzw. maximal) in K, so ist a ein Minimum (bzw. Maximum) von K.*

Beweis: (*nur für die Minimalität*) Es sei $a \in K$ ein minimales Element. Für $b \in K$ gilt $a \leq_R b$ oder $b \leq_R a$ wegen der Totalität von R. Gilt $b \leq_R a$, so folgt $a = b$ (bzw. $a \leq_R b$) wegen der Minimalität von a. Somit gilt in jedem Fall $a \leq_R b$ für alle $b \in K$. Somit ist a das Minimum von K. Damit ist die Proposition bewiesen. ■

3.5 Graphen*

Endliche binäre Relationen können durch Graphen beschrieben werden.

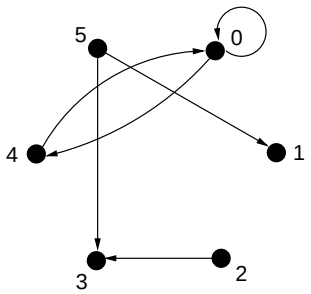
Definition 3.27 *Ein (gerichteter)* Graph $G = (V, E)$ *ist ein Paar bestehend aus einer Knotenmenge V und einer Kantenmenge $E \subseteq V \times V$. Die Elemente von V heißen* Knoten*; die Elemente von E heißen* Kanten*.*

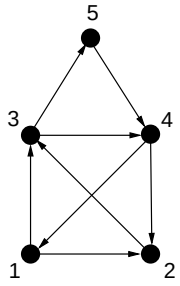
Für einen Graphen $G = (V, E)$ ist E eine binäre Relation auf V. Für eine binäre Relation R auf A ist (A, R) ein Graph.

Einen Graphen $G = (V, E)$ können wir wie folgt visualisieren:

- V sind Punkte in der Ebene.
- Für eine Kante $e = (v_i, v_k) \in E$ zeichnen wir einen Pfeil von v_i nach v_k.

Beispiel: Folgende Abbildungen repräsentieren Graphen:





Der linke Graph entspricht der Relation $\{(0,0),(0,4),(2,3),(4,0),(5,1),(5,3)\}$ auf $\{0,1,2,3,4,5\}$. Der rechte Graph $H = (V,E)$ ist ein Beispiel für einen Graphen, der keine Schleifen (Kanten von Knoten zu sich selbst) sowie keine zwei Kanten zwischen Paaren von Knoten enthält.

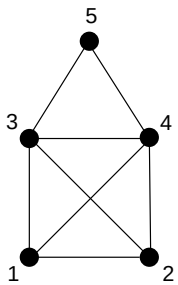
Definition 3.28 *Es sei $G = (V,E)$ ein Graph.*

1. *Ein* Weg *der Länge k in G ist eine Folge $(v_0, v_1, \ldots, v_k)$ mit $v_i \in V$ und $(v_{i-1}, v_i) \in E$ für alle $i \in \{1, \ldots, k\}$.*
2. *Ein* Kreis *(der Länge k) in G ist ein Weg $(v_0, v_1, \ldots, v_k)$ in G mit $v_0 = v_k$.*

Beispiel: Der Graph $H = (V,E)$ aus obigem Beispiel enthält einen Weg $(1,2,3,4,1,3,5,4,2)$ der Länge 8, der jede Kante genau einmal benutzt. Ein solcher Weg heißt Euler-*Weg*. Außerdem enthält H Kreise, z.B. $(1,3,4,1)$.

Neben gerichteten Graphen werden auch ungerichtete Graphen betrachtet. Eine *ungerichtete Kante* e zwischen Knoten u und v wird mit $e = \{u,v\}$ beschrieben, d.h. es muss folgende Symmetrieeigenschaft gelten: $(u,v) \in E \rightarrow (v,u) \in E$. Gilt dies für alle $u,v \in V$, so heißt $G = (V,E)$ *ungerichteter Graph*.

Beispiel: Der obige Graph $H = (V,E)$ kann als ungerichteter Graph folgendermaßen visualisiert werden:



Definition 3.29 *Es sei $G = (V, E)$ ein ungerichteter Graph.*

- *G heißt* zusammenhängend*, falls es für alle $u, v \in V$ einen Weg $(w_0, \ldots, w_k)$ in G mit $u = w_0$ und $w_k = v$ gibt.*
- *G heißt* Baum*, falls G zusammenhängend ist und keine Kreise der Länge $k \geq 1$ enthält.*

Beispiel: Von den beiden folgende Graphen ist der linke nicht zusammenhängend und der rechte ein Baum:

Induktion 4

4.1 Vollständige Induktion

Die vollständige Induktion ist ein Beweisprinzip, um eine mit dem Allquantor versehene Aussage über dem Universum der natürlichen Zahlen zu beweisen. Die Korrektheit des Beweisprinzips haben wir im Kapitel 1 bewiesen. Die allgemeinste Form der vollständigen Induktion ist Gegenstand von Theorem 4.1, das ohne Beweis angeben.

Theorem 4.1 *Es seien $A(n)$ eine Aussageform mit der freien Variable n über dem Universum der natürlichen Zahlen und $n_0 \in \mathbb{N}$. Dann ist die Aussage*

$$\Big((\forall n; n \leq n_0)[A(n)] \ \wedge \ (\forall n; n > n_0)\,[(\forall m; m < n)[A(m)] \to A(n)] \Big) \to (\forall n)[A(n)]$$

allgemeingültig.

Die Anwendung von Theorem 4.1 als Beweisverfahren sieht wie folgt aus:

- *Induktionsanfang* (IA): Zeige $A(0), A(1), \ldots, A(n_0)$.
- *Induktionsschritt* (IS): Zeige $A(n)$ für ein allgemeines $n > n_0$ unter der Annahme (*Induktionsvoraussetzung*, IV), dass $A(m)$ für alle $m < n$ gilt.

Im Folgenden diskutieren wir einige Beispiele für Induktionsbeweise. (Für weitere Beispiele sei auf das Skriptum zum „Brückenkurs Mathematik“ verwiesen.)

Proposition 4.A *Für alle $n \in \mathbb{N}$ gilt $\log_2(n+1) \leq n$.*

Beweis: (*Induktion über n*)

- *Induktionsanfang*: Für $n = 0$ gilt $\log_2(0+1) = 0 \leq 0$.
- *Induktionsschritt*: Für $n > 0$ gilt

$$\begin{aligned} \log_2(n+1) &\leq \log_2(2n) && (\text{da } n \geq 1) \\ &= 1 + \log_2 n && (\text{nach } \textit{Induktionsvoraussetzung}) \\ &\leq 1 + (n-1) && (\text{da } n \geq 1) \\ &= n \end{aligned}$$

Damit ist die Proposition bewiesen ■

Es ist jedoch nicht immer so einfach ein geeignetes n_0 zu wählen, um eine Aussage mittels Induktion zu beweisen.

Beispiel: Wir betrachten folgende Fragestellung: Gegeben $k \in \mathbb{N}$, für welches $n_0 \in \mathbb{N}$ gilt die Aussage: *Für alle* $n \in \mathbb{N}$ *mit* $n \geq n_0$ *gilt* $n \leq 2^{n-k}$?

Zunächst versuchen wir n_0 so zu bestimmen, dass der Induktionsschritt funktioniert:

- *Induktionsschritt*: Für $n > n_0$ gilt

$$\begin{aligned} n &= (n-1)+1 \\ &\leq 2^{n-1-k}+1 && \text{(nach } \textit{Induktionsvoraussetzung}\text{)} \\ &\leq 2^{n-1-k}+2^{n_0-k} && \text{(falls } n_0 \text{ die Ungleichung } 1 \leq 2^{n_0-k} \text{ erfüllt)} \\ &\leq 2^{n-1-k}+2^{n-1-k} && \text{(da } n-1 \geq n_0\text{)} \\ &= 2^{n-k} \end{aligned}$$

Der Induktionsschritt kann also für jedes $n_0 \geq k$ durchgeführt werden. Setzen wir allerdings $n_0 = k$, so gilt im Induktionsanfang $k \leq 2^{k-k}$ lediglich für $k \in \{0, 1\}$. Wir müssen also *ein* n_0 so finden, dass

$$2 \leq k \leq n_0 \leq 2^{n_0-k}$$

gilt. Durch Experimentieren erhält man z.B. als Ansatz $n_0 = k^2$. In der Tat gelingt der Induktionsanfang mit dieser Wahl.

- *Induktionsanfang*: Für $n = n_0 = k^2$ gilt zunächst

$$\begin{aligned} 2\log_2 k &\leq k \cdot \log_2 k && \text{(da } k \geq 2\text{)} \\ &\leq k \cdot (k-1) && \text{(nach Proposition 4.B)} \\ &= k^2 - k \end{aligned}$$

  Damit gilt $k^2 = 2\log_2 k \leq 2^{k^2-k}$.

Wir können die Aussage also für alle $n \geq k^2$ beweisen.

Proposition 4.B *Es sei* $k \in \mathbb{N}$. *Dann gilt* $n \leq 2^{n-k}$ *für alle* $n \in \mathbb{N}$ *mit* $n \geq k^2$.

Der Induktionsbeweis ergibt sich aus obigem Beispiel.

An einem weiteren Beispiel wollen wir auch auf die Fallstricke hinweisen, die zu fehlerhaften Induktionsbeweisen führen können.

Beispiel: Wir zeigen folgende, empirisch ganz offensichtlich falsche Aussage:

> *Alle Personen in einer Menge X von $\|X\| = n > 0$ Personen sind gleich groß.*

Wir beweisen diese Aussage mittels vollständiger Induktion über n.

- *Induktionsanfang*: Für $n = 1$ ist die Aussage offensichtlich wahr.
- *Induktionsschritt*: Für $n > 1$ sei X eine beliebige Menge mit $\|X\| = n$ Personen. Dann zerlegen wir X in Mengen Y und Z mit folgenden Eigenschaften:

  1. $X = Y \cup Z$
  2. $\|Y \cap Z\| = 1$
  3. $\|Y\| < \|X\|$
  4. $\|Z\| < \|X\|$

  Nach Induktionsvoraussetzung folgt aus der Eigenschaft $\|Y\| < \|X\|$, dass alle Personen in Y gleich groß sind. Hierbei ist zu beachten, dass $1 \leq \|Y\| \leq n - 1$ gilt; wir dürfen also die Induktionsvoraussetzung anwenden. Gleichfalls folgt nach Induktionsvoraussetzung aus der Eigenschaft $\|Z\| < \|X\|$, dass alle Personen in Z gleich groß sind. Da $Y \cap Z \neq \emptyset$ gilt, gibt es eine Person in beiden Mengen Y und Z, zu der alle Personen der beiden Mengen gleich groß sind. Wegen $Y \cup Z = X$ sind alle Personen in X gleich groß.

Damit wäre die Aussage bewiesen, hätten wir nicht einen Fehler gemacht. Wo liegt er? (Die Beantwortung der Frage sei zur Übung überlassen.)

4.2 Strukturelle Induktion

Wir wollen das Induktionsprinzip zu einer Beweismethode über induktiv definierten Mengen erweitern. Dabei führen wir nur den Spezialfall mittels einer Operation aus.

Definition 4.2 *Es sei $f : A^n \to A$ eine n-stellige Funktion (Operation). Dann sind die Abbildungen $\Gamma_f : \mathcal{P}(A) \to \mathcal{P}(A)$ sowie $\Gamma_f^k : \mathcal{P}(A) \to \mathcal{P}(A)$ für alle $k \in \mathbb{N}$ wie folgt definiert (für $B \subseteq A$):*

$$\begin{aligned}
\Gamma_f^0(B) &=_{\text{def}} B \\
\Gamma_f^k(B) &=_{\text{def}} \Gamma_f^{k-1}(B) \cup \left\{ f(a_1, \ldots, a_n) \mid a_1, \ldots, a_n \in \Gamma_f^{k-1}(B) \right\} \\
\Gamma_f(B) &=_{\text{def}} \bigcup_{k=0}^{\infty} \Gamma_f^k(B)
\end{aligned}$$

Die Menge $\Gamma_f(B)$ heißt Abschluss *von B unter f.*

Anschaulich gehören zur Menge $\Gamma_f(B)$ alle diejenigen Elemente von A, die sich durch eine endliche Anzahl von Hintereinanderausführungen der Operation f aus den Elementen der Menge B konstruieren lassen.

Beispiele: Wir wollen Definition 4.2 an Beispielen nachvollziehen.

- Wenn wir die Operation $\text{inc} : \mathbb{N} \to \mathbb{N} : n \mapsto n + 1$ betrachten, so gilt:

$$\begin{aligned}
\Gamma_{\text{inc}}^0(\{0\}) &= \{0\} \\
\Gamma_{\text{inc}}^1(\{0\}) &= \{0\} \cup \{1\} = \{0, 1\} \\
\Gamma_{\text{inc}}^2(\{0\}) &= \{0, 1\} \cup \{1, 2\} = \{0, 1, 2\} \\
\Gamma_{\text{inc}}^3(\{0\}) &= \{0, 1, 2\} \cup \{1, 2, 3\} = \{0, 1, 2, 3\} \\
&\vdots \\
\Gamma_{\text{inc}}^k(\{0\}) &= \{0, 1, \ldots, k\} \\
&\vdots \\
\Gamma_{\text{inc}}(\{0\}) &= \mathbb{N}
\end{aligned}$$

  Der Beweis der Gleichheit $\Gamma_{\text{inc}}^k(\{0\}) = \{0, 1, \ldots, k\}$ für alle $k \in \mathbb{N}$ kann mittels vollständiger Induktion über k geführt werden.

- Wenn wir die Operation $+ : \mathbb{N}^2 \to \mathbb{N} : (n, m) \mapsto n + m$ betrachten, so gilt:

$$\begin{aligned}
\Gamma_{\text{inc}}^0(\{1\}) &= \{1\} \\
\Gamma_{\text{inc}}^1(\{1\}) &= \{1\} \cup \{2\} = \{1, 2\}
\end{aligned}$$

$$
\begin{aligned}
\Gamma^2_{\text{inc}}(\{1\}) &= \{1,2\} \cup \{2,3,4\} = \{1,2,3,4\} \\
\Gamma^3_{\text{inc}}(\{1\}) &= \{1,2,3,4\} \cup \{2,3,4,5,6,7,8\} = \{1,2,3,4,5,6,7,8\} \\
&\vdots \\
\Gamma^k_{\text{inc}}(\{1\}) &= \{1,2,\ldots,2^k\} \\
&\vdots \\
\Gamma_{\text{inc}}(\{1\}) &= \mathbb{N}_+
\end{aligned}
$$

Auch hier kann der Beweis der Gleichheit $\Gamma^k_{\text{inc}}(\{1\}) = \{1,2,\ldots,2^k\}$ für alle $k \in \mathbb{N}$ mittels vollständiger Induktion über k geführt werden.

Definition 4.3 *Eine Menge $B \subseteq A$ heißt* endlich erzeugt *(oder* induktiv definiert*) aus $B_0 \subseteq A$, falls es eine Funktion $f : A^n \to A$ gibt mit $B = \Gamma_f(B_0)$.*

Für endlich erzeugte Mengen können wir das Beweisprinzip der strukturelle Induktion formal angeben (ohne Beweis).

Theorem 4.4 *Es sei B eine endlich erzeugte Menge aus B_0 mittels f. Es sei $A(x)$ eine Aussageform mit der freien Variablen x über dem Universum B. Dann ist die Aussage*

$$
\Big((\forall x \in B_0)[A(x)] \ \wedge \ (\forall k; k > 0)\left[(\forall x \in \Gamma^{k-1}_f(B_0))[A(x)] \to (\forall x \in \Gamma^k_f(B_0))[A(x)]\right]\Big) \to (\forall x \in B)[A(x)]
$$

allgemeingültig.

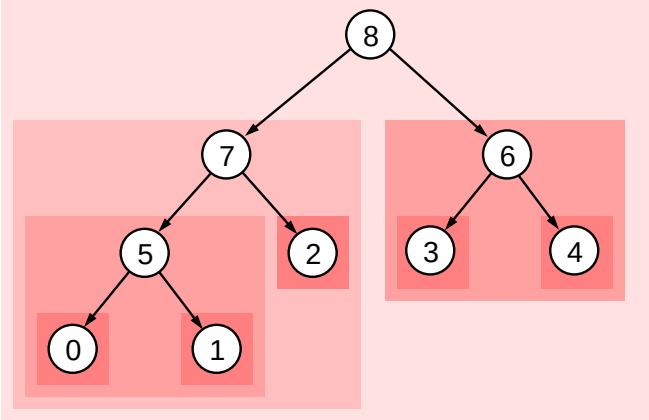
Die Anwendung dieser recht komplizierten Formulierung als Beweisprinzip ist wie folgt:

- *Induktionsanfang* (IA): Zeige $A(x)$ für alle $x \in B_0$.
- *Induktionsschritt* (IS): Zeige $A(x)$ für ein allgemeines $x \in B \setminus B_0$ unter der Annahme (*Induktionsvoraussetzung*, IV), dass $A(y_1), \ldots, A(y_n)$ für $x = f(y_1, \ldots, y_n)$ gilt. (Hier ist unter technischen Gesichtspunkten darauf zu achten, dass $y_1, \ldots, y_n$ einfacher sind, d.h. ist $x \in \Gamma^k_f(B_0)$, so muss $y_1, \ldots, y_n \in \Gamma^{k-1}_f(B_0)$ gelten.)

Zum Abschluss wollen wir uns ein Beispiel für die in der Informatik typische konstruktive Vorgehensweise anschauen.

Eine wichtige Datenstruktur in der Informatik sind Binärbäume als Verallgemeinerung von Listen. In einer Liste hat jedes Element bis auf das letzte genau einen Nachfolger und jedes Element bis auf das erste genau einen Vorgänger. Verlangt man nur die Eigenschaft das jedes Element bis auf eines genau einen Vorgänger besitzt (und Kreise ausgeschlossen

werden), gelangt man zu Bäumen. Eine Sonderklasse von Bäumen sind volle, gewurzelte Binärbäume. Ein Beispiel ist der folgende Baum:



Die Menge aller vollen, gewurzelten Binärbäume kann man wie folgt induktiv definieren. Bäume sind Tripel (V, E, r), wobei V für die Menge der Knoten, $E \subseteq V^2$ für die Menge der Kanten sowie $r \in V$ für die Wurzel stehen. Wir geben nun eine Menge B_0 und eine Operation f an:

$$\begin{aligned} B_0 &=_{\text{def}} \{\ (\{r\}, \emptyset, r) \mid r \in \mathbb{N}\ \} \\ f((V_1, E_1, r_1), (V_2, E_2, r_2)) &=_{\text{def}} (V_1 \cup V_2 \cup \{r\}, E_1 \cup E_2 \cup \{(r, r_1), (r, r_2)\}, r), \\ &\qquad \text{wobei } V_1 \cap V_2 = \emptyset \text{ sowie } r \notin V_1 \cup V_2 \text{ gilt} \end{aligned}$$

Die Menge der vollen, gewurzelten Binärbäume ist dann gerade die Menge $\Gamma_f(B_0)$.

Bevor wir unseren zu beweisende Eigenschaften formulieren, führen wir noch zwei Begriffe ein. Es sei $T = (V, E, r)$ ein voller, gewurzelter Binärbaum. Ein Element $v \in V$ heißt *Blatt* (bzw. *Blattknoten*), falls es kein $u \in V$ mit $(v, u) \in E$ gibt; sonst heißt v *innerer Knoten*.

Proposition 4.F *Für einen vollen, gewurzelten Binärbaum T seien n_T die Anzahl innerer Knoten und m_T die Anzahl der Blätter. Dann gilt stets $n_T = m_T - 1$.*

Beweis: (*Induktion über den Aufbau der Bäume*)

- *Induktionsanfang*: Ist $T = (\{r\}, \emptyset, r)$, so gilt $n_T = 0$ und $m_T = 1$.
- *Induktionsschritt*: Es sei $T = f(T_1, T_2)$ für geeignete Bäume $T_1 = (V_1, E_1, r_1)$ und $T_2 = (V_2, E_2, r_2)$. Dann gilt insbesondere, dass die Blätter bzw. inneren Knoten von

T_1 und T_2 auch Blätter bzw. innere Knoten von T sind, da in T nur die Paare (r, r_1) und (r, r_2) hinzukommen. Mithin gilt:

$$\begin{aligned} n_T &= n_{T_1} + n_{T_2} + 1 && (r \text{ ist ein innerer Knoten von } T) \\ &= (m_{T_1} - 1) + (m_{T_2} - 1) + 1 && (\text{nach } \textit{Induktionsvoraussetzung}) \\ &= (m_{T_1} + m_{T_2}) - 1 \\ &= m_T - 1 \end{aligned}$$

Damit ist die Proposition bewiesen. ■

4.3 Transitive Hülle*

Es sei $R \subseteq A \times A$ eine binäre Relation auf A. Der *transitive Join* $\bowtie$ von R ist definiert als

$$R \bowtie R =_{\text{def}} \{\ (a, c) \mid (\exists b \in A)[(a, b) \in R \wedge (b, c) \in R]\ \}$$

Außerdem definieren wir auf Elementebene $(a, b) \bowtie (b, c) =_{\text{def}} (a, c)$.

Definition 4.5 *Es sei $R \subseteq A \times A$ eine binäre Relation.*

1. *Die* transitive Hülle R^+ *von R ist definiert durch*

$$R^+ =_{\text{def}} \Gamma_{\bowtie}(R).$$

2. *Die* reflexive und transitive Hülle R^* *von R ist definiert durch*

$$R^* =_{\text{def}} R^+ \cup \{\ (a, a) \mid a \in A\ \}.$$

Proposition 4.6 *Es seien $R \subseteq A \times A$ eine binäre Relation und $x, y \in A$. Dann gilt:*

$$\begin{aligned} &(x, y) \in R^+ \\ &\quad \Longleftrightarrow \quad (\exists k \in \mathbb{N}_+)(\exists x_0, \ldots, x_k \in A)\,[x_0 = x \wedge x_k = y \wedge (\forall i \in \{1, \ldots, k\})[(x_{i-1}, x_i) \in R]] \\ &(x, y) \in R^* \\ &\quad \Longleftrightarrow \quad (\exists k \in \mathbb{N})(\exists x_0, \ldots, x_k \in A)\,[x_0 = x \wedge x_k = y \wedge (\forall i \in \{1, \ldots, k\})[(x_{i-1}, x_i) \in R]] \end{aligned}$$

Für eine beliebige binäre Relation $R \subseteq A \times A$ definieren wir die Relation $\sim_{R^*}$ wie folgt:

$$x \sim_{R^*} y \Longleftrightarrow_{\text{def}} (x, y) \in R^* \wedge (y, x) \in R^*$$

Proposition 4.7 *Für jede Relation $R \subseteq A \times A$ ist $\sim_{R^*}$ eine Äquivalenzrelation.*

Beweis: Übungsaufgabe. ■

Für $R \subseteq A \times A$ und $\sim_{R^*}$ definieren wir auf der Menge $\{\ [x]_{\sim_{R^*}} \mid x \in A\ \}$ der Äquivalenzklassen die Relation $\leq_{R^*}$ vermöge

$$[x]_{\sim_{R^*}} \leq_{R^*} [y]_{\sim_{R^*}} \iff_{\text{def}} (x,y) \in R^*$$

Damit diese Definition sinnvoll ist, muss gezeigt werden, dass die Relation $\leq_{R^*}$ unabhängig von der Wahl der Repräsentanten ist.

Proposition 4.8 *Für jede Relation $R \subseteq A \times A$ und alle $x, y \in A$ gilt:*

$$[x]_{\sim_{R^*}} \leq_{R^*} [y]_{\sim_{R^*}} \iff (\forall x' \in [x]_{\sim_{R^*}})(\forall y' \in [y]_{\sim_{R^*}})[(x',y') \in R^*]$$

Beweis: Die Richtung ($\Leftarrow$) ist offensichtlich. Für die Richtung ($\Rightarrow$) seien $x', y' \in A$ mit $x' \in [x]_{\sim_{R^*}}$ und $y' \in [y]_{\sim_{R^*}}$ gegeben. Somit gilt $(x', x) \in R^*$ und $(y, y') \in R^*$. Wegen $[x]_{\sim_{R^*}} \leq_{R^*} [y]_{\sim_{R^*}}$ gilt $(x,y) \in R^*$. Mit der Transitivität von R^* folgt daraus zunächst $(x', y) \in R^*$ und weiter $(x', y') \in R^*$. ■

Proposition 4.9 *Für jede Relation $R \subseteq A \times A$ ist $\left(\{[x]_{\sim_{R^*}} | x \in A\}, \leq_{R^*}\right)$ eine halbgeordnete Menge.*

Beweis: Wir überprüfen die Halbordnungseigenschaften von $\leq_{R^*}$:

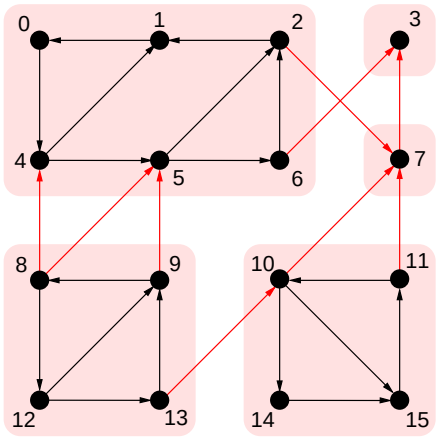
1. *Reflexivität*: Wegen $(x,x) \in R^*$ gilt $[x]_{\sim_{R^*}} \leq_{R^*} [x]_{\sim_{R^*}}$ für alle $x \in A$.

2. *Transitivität*: Es gelte $[x]_{\sim_{R^*}} \leq_{R^*} [y]_{\sim_{R^*}}$ und $[y]_{\sim_{R^*}} \leq_{R^*} [z]_{\sim_{R^*}}$. Mit anderen Worten gilt $(x,y) \in R^*$ und $(y,z) \in R^*$. Mithin gilt $(x,z) \in R^*$ wegen der Transitivität von R^*. Es folgt $[x]_{\sim_{R^*}} \leq_{R^*} [y]_{\sim_{R^*}}$.

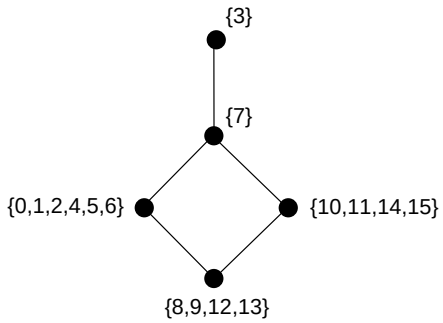
3. *Antisymmetrie*: Es gelte $[x]_{\sim_{R^*}} \leq_{R^*} [y]_{\sim_{R^*}}$ und $[y]_{\sim_{R^*}} \leq_{R^*} [x]_{\sim_{R^*}}$. Folglich gilt $(x,y) \in R^*$ und $(y,x) \in R^*$. Somit gilt $x \sim_{R^*} y$ bzw. $[x]_{\sim_{R^*}} = [y]_{\sim_{R^*}}$.

Damit ist die Proposition bewiesen. ■

Beispiel: Wir betrachten eine Relation R über $\{0, 1, \ldots, 15\}$ wie durch den folgenden gerichteten Graphen beschrieben:



R^* kann durch den gleichen Graphen beschrieben werden, wenn wir der Vereinbarung folgen, dass Schleifen weggelassen werden und auch Erreichbarkeitkeit von Knoten über Wege, denn es gilt: $(x, y) \in R^* \iff$ es gibt einen Weg von x nach y. Die farbig hinterlegten Bereiche entsprechen gerade den Äquivalenzklassen hinsichtlich der Erreichbarkeit im Graphen. Die roten Pfeile geben die Halbordnung auf den Äquivalenzklassen an. Das zugehörige HASSE-Diagramm sieht dannwie folgt aus:



4.4 Mächtigkeit von Mengen*

Es seien A und B endliche Mengen. Wir wissen bereits, dass genau dann $\|A\| = \|B\|$ gilt, wenn es eine bijektive Funktion $f : A \to B$ gibt. Für unendliche Mengen erheben wir dies zur Definition.

Definition 4.10 *Zwei Mengen A und B heißen* gleichmächtig, *falls es eine bijektive Funktion $f : A \to B$ gibt.*

Beispiel: $\mathbb{N}$ und $\mathbb{Z}$ sind gleichmächtig, denn die Funktion

$$f : \mathbb{N} \to \mathbb{Z} : n \mapsto \begin{cases} n/2 & \text{falls } n \text{ gerade ist} \\ -(n+1)/2 & \text{sonst} \end{cases}$$

ist bijektiv mit der Umkehrfunktion

$$f^{-1} : \mathbb{Z} \to \mathbb{N} : z \mapsto \begin{cases} 2z & \text{falls } z \geq 0 \\ -2z - 1 & \text{sonst} \end{cases}$$

Definition 4.11 *Eine Menge A heißt*

1. abzählbar*, falls eine surjektive Funktion $f : \mathbb{N} \to A$ existiert;*
2. abzählbar unendlich*, falls eine bijektive Funktion Funktion $f : \mathbb{N} \to A$ existiert;*
3. überabzählbar*, falls A nicht abzählbar ist.*

Beispiele:

1. Jede endliche Menge ist abzählbar.
2. $\mathbb{Z}$ ist abzählbar (unendlich).
3. $\mathbb{Q}$ ist abzählbar (siehe Übungsblatt).
4. $\mathbb{R}$ ist überabzählbar, denn bereits das Intervall $[0, 1) = \{x \in \mathbb{R} | 0 \leq x < 1\}$ ist überabzählbar. Um dies einzusehen, betrachten wir die Dezimaldarstellung von $x \in [0, 1)$:

   $$x = 0, d_0 d_1 d_2 \ldots = \sum_{n=0}^{\infty} d_n \cdot 10^{-(n+1)},$$

   wobei $d_n \in \{0, 1, \ldots, 9\}$ gilt für alle $n \in \mathbb{N}$. (Zu beachten ist, dass wir für die Eindeutigkeit der Folge für x voraussetzen, dass für alle $k \in \mathbb{N}$ ein $n > k$ existiert mit $d_n \neq 9$. Z.B. schließen wir für $0, 50000 \ldots = 0, 49999 \ldots$ die letztere aus.)
   Angenommen $[0, 1)$ wäre abzählbar. Dann gibt es eine bijektive Funktion $f : \mathbb{N} \to [0, 1)$, d.h.

   $$f(k) = 0, d_{k,0} d_{k,1} d_{k,2} \ldots = \sum_{n=0}^{\infty} d_{k,n} \cdot 10^{-(n+1)}$$

   mit $d_{k,n} \in \{0, 1, \ldots, 9\}$. Wir betrachten nun die folgende reelle Zahl $x \in [0, 1)$:

   $$x =_{\text{def}} 0, b_0 b_1 b_2 \ldots \text{ mit } b_n =_{\text{def}} \begin{cases} 1 & \text{falls } d_{n,n} \text{ gerade ist} \\ 2 & \text{falls } d_{n,n} \text{ ungerade ist} \end{cases}$$

Wären z.B. die ersten fünf Werte von f die Brüche $1/2, 1/3, \ldots, 1/6$, so ergäbe sich

$$\begin{aligned} f(0) &= 0,\underline{5}0000\ldots && \text{mit } d_{0,0} = 5 \\ f(1) &= 0,3\underline{3}333\ldots && \text{mit } d_{1,1} = 3 \\ f(2) &= 0,25\underline{0}00\ldots && \text{mit } d_{2,2} = 0 \\ f(3) &= 0,200\underline{0}0\ldots && \text{mit } d_{3,3} = 0 \\ f(4) &= 0,16666\underline{6}\ldots && \text{mit } d_{4,4} = 6 \end{aligned}$$

und folglich $x = 0,22111\ldots$.

Nach Definition von x gilt nun aber stets $b_n \neq d_{n,n}$ für alle $n \in \mathbb{N}$. Damit gilt $x \neq f(n)$ für alle $n \in \mathbb{N}$. Folglich kann f nicht surjektiv also auch nicht bijektiv sein. Somit ist $[0, 1)$ überabzählbar.

Analysis 5

In diesem Kapitel wollen wir eine Auswahl grundlegender Begriffe der reellen Analysis einführen mit dem Ziel, das Konzept der Reihenentwicklung von Funktionen in Potenzreihen zu entwickeln.

5.1 Konvergenz von Folgen

Eine (reelle) Folge ist eine Abbildung $a : \mathbb{N} \to \mathbb{R}$. Wenn wir die natürliche Anordnung von $\mathbb{N}$ zugrunde legen, so kann a durch die Folge der Funktionswerte $a(0), a(1), a(2), \ldots$ beschrieben werden. Allgemein hat sich dafür die Schreibweise $(a_0, a_1, a_2, \ldots)$ oder kompakt $(a_n)_{n\in\mathbb{N}}$ etabliert.

Der zentrale Begriff der Analysis ist die Konvergenz.

Definition 5.1 *Es seien $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen und $c \in \mathbb{R}$.*

1. *Die Folge $(a_n)_{n\in\mathbb{N}}$* konvergiert gegen *c, falls folgende Aussage wahr ist:*

$$(\forall \varepsilon > 0)\ (\exists n_0 \in \mathbb{N})\ (\forall n \geq n_0)\ [\ |a_n - c| < \varepsilon\ ]$$

   Falls $(a_n)_{n\in\mathbb{N}}$ gegen c konvergiert, so heißt c Grenzwert *von $(a_n)_{n\in\mathbb{N}}$ und wir schreiben $\lim\limits_{n\to\infty} a_n = c$.*

2. *Die Folge $(a_n)_{n\in\mathbb{N}}$ heißt* konvergent*, falls ein Grenzwert für $(a_n)_{n\in\mathbb{N}}$ existiert. Anderenfalls heißt die Folge* divergent*.*

Die Annäherung einer Folge an einen Grenzwert wird gemäß dieser Definition so verstanden, dass in *jeder* beliebig kleinen Umgebung des Grenzwertes alle Folgenglieder bis auf eine endliche Menge von Ausnahmen (nämlich höchstens diejenigen Folgenglieder mit einem Index kleiner als n_0) zu finden sind.

Die logische Struktur der Definitionsbildung ist einigermaßen komplex und bedarf Trainings in der Handhabung. Zur Übung negiere man die formale Definition einer konvergenten Folge mit allen in der vollständigen Definition enthaltenen Quantoren.

Bevor wir Beispiele geben, überzeugen wir uns von der Wohldefiniertheit des Grenzwertes.

Proposition 5.2 *Der Grenzwert einer konvergenten Folge ist eindeutig.*

Beweis: Wir führen einen Widerspruchsbeweis. Es sei $(a_n)_{n\in\mathbb{N}}$ eine beliebige konvergente Folge. Es seien $c_1, c_2 \in \mathbb{R}$ Grenzwerte von $(a_n)_{n\in\mathbb{N}}$. Angenommen $c_1 \neq c_2$. Dann gibt es für $\varepsilon =_{\text{def}} \frac{1}{2} \cdot |c_1 - c_2| > 0$ natürliche Zahlen $n_0^{(1)}$ und $n_0^{(2)}$ mit

- $|a_n - c_1| < \varepsilon$ für alle $n \geq n_0^{(1)}$ und
- $|a_n - c_2| < \varepsilon$ für alle $n \geq n_0^{(2)}$.

Somit gilt für $N =_{\text{def}} \max\left\{n_0^{(1)}, n_0^{(2)}\right\}$

$$|c_1 - c_2| = |c_1 - a_N + a_N - c_2| \leq |c_1 - a_N| + |a_N - c_2| < 2\varepsilon = |c_1 - c_2|.$$

Dies ist jedoch ein Widerspruch. Damit ist die Annahme $c_1 \neq c_2$ falsch und die Proposition ist bewiesen. ∎

Beispiele: Wir führen einige Beispiele für die Definition 5.1 an:

- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} \frac{1}{n+1}$ konvergiert gegen 0. Dies ist wie folgt einzusehen: Für $\varepsilon > 0$ definiere $n_0 =_{\text{def}} \lfloor \varepsilon^{-1} \rfloor$. Dann gilt für alle $n \geq n_0$:

$$|a_n - 0| = \left|\frac{1}{n+1} - 0\right| = \frac{1}{n+1} \leq \frac{1}{n_0+1} = \frac{1}{\lfloor \varepsilon^{-1} \rfloor + 1} \leq \frac{1}{\varepsilon^{-1}} = \varepsilon$$

- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} 2$ konvergiert gegen 2. Dies ist wie folgt einzusehen: Für $\varepsilon > 0$ definiere $n_0 =_{\text{def}} 0$. Dann gilt $|a_n - 2| = 0 < \varepsilon$ für $n \geq n_0$.
- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} \lfloor \sqrt{2} \cdot 10^n \rfloor \cdot 10^{-n}$ konvergiert gegen $\sqrt{2}$. Dies ist wie folgt einzusehen: Für $\varepsilon > 0$ definiere $n_0 =_{\text{def}} -\lfloor \log_{10} \varepsilon \rfloor$. Dann gilt für alle $n \geq n_0$:

$$\begin{aligned} |a_n - \sqrt{2}| &= \left|\frac{\lfloor \sqrt{2} \cdot 10^n \rfloor}{10^n} - \sqrt{2}\right| = \frac{\sqrt{2} \cdot 10^n - \lfloor \sqrt{2} \cdot 10^n \rfloor}{10^n} \\ &< 10^{-n} \leq 10^{\lfloor \log_{10} \varepsilon \rfloor} \leq \varepsilon \end{aligned}$$

- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} (-1)^n$ ist nicht konvergent. Angenommen $\lim_{n\to\infty} a_n = c$. Dann gibt es für $\varepsilon = \frac{1}{3}$ ein $n_0 \in \mathbb{N}$, sodass $|a_n - c| < \frac{1}{3}$ für alle $n \geq n_0$ gilt. Dann gilt aber auch für $n \geq n_0$:

$$2 = |a_n - a_{n+1}| = |a_n - c - a_{n+1} + c| \leq |a_n - c| + |a_{n+1} - c| < \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

  Dies ist jedoch ein Widerspruch. Somit existiert kein Grenzwert.
- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_0 =_{\text{def}} 1$ und $a_n =_{\text{def}} \frac{n^{(-1)^n}}{n}$ ist nicht konvergent. Die Folge beginnt mit $1, 1, 1, \frac{1}{9}, 1, \frac{1}{25}, 1, \frac{1}{49}, \ldots$ Die Begründung für die Nichtkonvergenz folgt analog zum vorangegangenen Beispiel aus $|a_n - a_{n+1}| \geq \frac{8}{9}$ für alle $n \geq 2$. ∎

Das dritte Beispiel ist bemerkenswert: Obwohl alle Folgenglieder rationale Zahlen sind, liegt der Grenzwert $\sqrt{2}$ der Folge nicht im Bereich der rationalen Zahlen. Nicht jede Menge ist also abgeschlossen unter Grenzwertbildung. Die Menge der reellen Zahlen kann gerade als die Menge aller Grenzwerte konvergenter rationaler Folgen konstruiert werden. Damit kann gezeigt werden, dass die reellen Zahlen im Gegensatz zu den rationalen Zahlen auch abgeschlossen unter Grenzwertbildung sind.

Das folgende Lemma gibt Regeln zur Bestimmung von Grenzwerten an.

Lemma 5.3 *Es seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ reelle konvergente Folgen. Dann gilt:*

1. *Gilt $a_n \leq b_n$ für alle $n \geq n_0$ und geeignetes $n_0 \in \mathbb{N}$, so gilt $\lim\limits_{n\to\infty} a_n \leq \lim\limits_{n\to\infty} b_n$*
2. *Gilt $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = x$ und gilt $a_n \leq b_n \leq c_n$ für alle $n \geq n_0$ und geeignetes $n_0 \in \mathbb{N}$, so gilt $\lim\limits_{n\to\infty} b_n = x$*
3. $\lim\limits_{n\to\infty} (a_n + b_n) = \left(\lim\limits_{n\to\infty} a_n\right) + \left(\lim\limits_{n\to\infty} b_n\right)$
4. $\lim\limits_{n\to\infty} (a_n \cdot b_n) = \left(\lim\limits_{n\to\infty} a_n\right) \cdot \left(\lim\limits_{n\to\infty} b_n\right)$
5. $\lim\limits_{n\to\infty} \left(\dfrac{a_n}{b_n}\right) = \dfrac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n}$, *falls* $\lim\limits_{n\to\infty} b_n \neq 0$ *und* $b_n \neq 0$ *für alle* $n \in \mathbb{N}$

Beweis: Wir beweisen die Aussagen einzeln:

1. Es gelte $\lim\limits_{n\to\infty} a_n = x$ und $\lim\limits_{n\to\infty} b_n = y$ sowie $a_n \leq b_n$ für alle $n \geq n_0$ und ein geeignetes $n_0 \in \mathbb{N}$. Angenommen es gilt $x > y$. Dann gibt es für $\varepsilon = \frac{1}{2}(x - y) > 0$ natürliche Zahlen $n_0^{(1)}$ und $n_0^{(2)}$ mit

   - $|a_n - x| < \varepsilon$ und folglich $a_n > x - \varepsilon$ für alle $n \geq n_0^{(1)}$
   - $|b_n - y| < \varepsilon$ und folglich $b_n < y + \varepsilon$ für alle $n \geq n_0^{(2)}$

   Damit gilt für alle $n \geq \max\left\{n_0, n_0^{(1)}, n_0^{(2)}\right\}$

   $$a_n - b_n > (x - \varepsilon) - (y + \varepsilon) = x - y - 2\varepsilon = 0,$$

   d.h. $a_n > b_n$. Dies ist ein Widerspruch. Mithin gilt $x \leq y$.

2. folgt aus der ersten Aussage.

3. Es gelte $\lim\limits_{n\to\infty} a_n = x$ und $\lim\limits_{n\to\infty} b_n = y$. Dann gibt es für jedes $\varepsilon > 0$ natürliche Zahlen $n_0^{(1)}$ und $n_0^{(2)}$ mit

   - $|a_n - x| < \dfrac{\varepsilon}{2}$ für alle $n \geq n_0^{(1)}$

- $|b_n - y| < \frac{\varepsilon}{2}$ für alle $n \geq n_0^{(2)}$

Für $n \geq \max\left\{n_0^{(1)}, n_0^{(2)}\right\}$ gilt somit auch

$$a_n + b_n - (x + y) = |a_n - x + b_n - y| \leq |a_n - x| + |b_n - y| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Folglich konvergiert die Folge $(a_n + b_n)_{n\in\mathbb{N}}$ gegen $x + y$.

4. Es gelte $\lim\limits_{n\to\infty} a_n = x$ und $\lim\limits_{n\to\infty} b_n = y$. Zunächst halten wir fest, dass $(a_n)_{n\in\mathbb{N}}$ beschränkt ist (siehe Übungsblatt 9). Es gibt also ein $s > 0$ mit $|a_n| \leq s$ für alle $n \in \mathbb{N}$. Weiterhin existieren zu jedem $\varepsilon > 0$ geeignete natürliche Zahlen $n_0^{(1)}$ und $n_0^{(2)}$ mit

   - $|a_n - x| < \frac{\varepsilon}{2} \cdot (\max\{|y|, 1\})^{-1}$ für alle $n \geq n_0^{(1)}$
   - $|b_n - y| < \frac{\varepsilon}{2} \cdot s^{-1}$ für alle $n \geq n_0^{(2)}$

   Für $n \geq \max\left\{n_0^{(1)}, n_0^{(2)}\right\}$ gilt somit

$$\begin{aligned} |a_n b_n - xy| &= |a_n b_n - a_n y + a_n y - xy| \\ &= |a_n(b_n - y) + (a_n - x)y| \\ &\leq |a_n| \cdot |b_n - y| + |y| \cdot |a_n - x| \\ &< s \cdot \left(\frac{\varepsilon}{2} \cdot s^{-1}\right) + |y| \cdot \left(\frac{\varepsilon}{2} \cdot (\max\{|y|, 1\})^{-1}\right) \\ &\leq \varepsilon \end{aligned}$$

   Mithin konvergiert die Folge $(a_n b_n)_{n\in\mathbb{N}}$ gegen xy.

5. Wegen der vierten Aussage genügt es, $\lim\limits_{n\to\infty} \left(b_n^{-1}\right) = \left(\lim\limits_{n\to\infty} b_n\right)^{-1}$ unter den angegebenen Voraussetzungen zu zeigen. Es gelte also $\lim\limits_{n\to\infty} b_n = y \neq 0$ sowie $b_n \neq 0$ für alle $n \in \mathbb{N}$. Wegen $b_n \neq 0$ ist b_n^{-1} stets definiert. Wegen $y \neq 0$ und da $(b_n)_{n\in\mathbb{N}}$ gegen y konvergiert, gibt es eine natürliche Zahl n_0, sodass für alle $n \geq n_0$

$$|y| = |y - b_n + b_n| \leq |y - b_n| + |b_n| \leq \frac{|y|}{2} + |b_n|$$

   also $|b_n| \geq \frac{|y|}{2}$ gilt. Weiterhin gibt es für alle $\varepsilon > 0$ ein $n_0' \in \mathbb{N}$ mit $|b_n - y| < \frac{\varepsilon}{2} \cdot |y|^2$. Für $n \geq \max\{n_0, n_0'\}$ folgt somit

$$\left|\frac{1}{b_n} - \frac{1}{y}\right| = \frac{|y - b_n|}{|b_n| \cdot |y|} \leq \frac{2 \cdot |b_n - y|}{|y|^2} < \varepsilon$$

   Mithin konvergiert $(b_n^{-1})_{n\in\mathbb{N}}$ gegen y^{-1}.

Damit ist das Lemma bewiesen. ■

Beispiele: Folgende Grenzwerte werden mit Hilfe von Lemma 5.3 bestimmt:

- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} \frac{2n^2+17n}{3n^2+5}$ konvergiert gegen $\frac{2}{3}$. Denn es gilt:

$$\lim_{n\to\infty} \frac{2n^2+17}{3n^2+5} = \lim_{n\to\infty} \frac{n^2\cdot(2+\frac{17}{n})}{n^2\cdot(3+\frac{5}{n^2})} = \frac{\lim\limits_{n\to\infty} 2+\frac{17}{n}}{\lim\limits_{n\to\infty} 3+\frac{5}{n^2}} = \frac{2}{3}$$

- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} \frac{n^2}{2^n}$ konvergiert gegen 0. Wegen $2^n \geq n^3$ für $n \geq 10$ gilt nämlich

$$0 \leq \lim_{n\to\infty} \frac{n^2}{2^n} \leq \lim_{n\to\infty} \frac{n^2}{n^3} = \lim_{n\to\infty} \frac{1}{n} = 0.$$

- Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} \sqrt[n]{n}$ konvergiert gegen 1. Zum Nachweis betrachten wir zunächst die Hilfsfolge $(d_n)_{n\in\mathbb{N}}$ mit $d_n =_{\text{def}} (1+\sqrt{2/n})^n$. Dann gilt für $n \geq 2$ nach dem Binomialtheorem

$$d_n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} \left(\sqrt{2/n}\right)^k \geq 1 + \binom{n}{1}\sqrt{2/n} + \binom{n}{2}\frac{2}{n} = n + \sqrt{2n} \geq n.$$

  Wir definieren Folgen $(b_n)_{n\in\mathbb{N}}$ mit $b_n =_{\text{def}} 1$ und $(c_n)_{n\in\mathbb{N}}$ mit $c_n =_{\text{def}} \sqrt[n]{d_n}$. Für $n \geq 2$ gilt

$$b_n \leq a_n \leq c_n = \sqrt[n]{d_n} = \sqrt[n]{\left(1+\sqrt{2/n}\right)^n} = 1+\sqrt{2/n}.$$

  Wir erhalten $1 = \lim\limits_{n\to\infty} b_n = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = 1$.

5.2 Konvergenz von Reihen

Zu einer Folge $(a_n)_{n\in\mathbb{N}}$ definieren wir die *Reihe* $(s_n)_{n\in\mathbb{N}}$ wie folgt für $n \in \mathbb{N}$:

$$s_n =_{\text{def}} \sum_{k=0}^{n} a_k$$

Die Folgenglieder a_k heißen *Koeffizienten*. Falls $(s_n)_{n\in\mathbb{N}}$ konvergiert, schreiben wir

$$\sum_{k=0}^{\infty} a_k =_{\text{def}} \lim_{n\to\infty} s_n.$$

Üblicherweise wird die unendliche Summe $\sum_{k=0}^{\infty} a_k$ (also ein Grenzwert) mit der Reihe $(s_n)_{n\in\mathbb{N}}$ identifiziert.

Beispiele: Wir geben einige Reihen und ihre Grenzwerte an, falls sie existieren:

- Zur Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} q^n$ für $0 \leq q < 1$ gehört die geometrische Reihe $\sum_{n=0}^{\infty} q^n$. Um den Grenzwert zu berechnen, definieren wir $s_n =_{\text{def}} \sum_{k=0}^{n} q^k$. Wir wissen bereits, dass mittels vollständiger Induktion gezeigt werden kann:

  $$s_n = \frac{1-q^{n+1}}{1-q} \qquad \text{für } n \in \mathbb{N}$$

  Wegen $\lim_{n\to\infty}\left(1-q^{n+1}\right) = 1$ und $\lim_{n\to\infty}(1-q) = 1-q$ gilt

  $$\sum_{n=0}^{\infty} q^n = \lim_{n\to\infty} s_n = \frac{\lim_{n\to\infty}\left(1-q^{n+1}\right)}{\lim_{n\to\infty}(1-q)} = \frac{1}{1-q}.$$

- $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$, denn es gilt:

  $$\begin{aligned} s_n &=_{\text{def}} \sum_{k=1}^{n+1} \frac{1}{k(k+1)} = \sum_{k=1}^{n+1} \frac{(k+1)-k}{k(k+1)} = \sum_{k=1}^{n+1} \left(\frac{1}{k} - \frac{1}{k+1}\right) \\ &= \left(\sum_{k=1}^{n+1} \frac{1}{k}\right) - \left(\sum_{k=2}^{n+2} \frac{1}{k}\right) = 1 - \frac{1}{n+2} \end{aligned}$$

  Also folgt $\lim_{n\to\infty} s_n = 1$.

- $\sum_{n=1}^{\infty} \frac{1}{n}$ existiert nicht. Dies ist wie folgt einzusehen: Es sei $s_n =_{\text{def}} \sum_{k=1}^{n+1} \frac{1}{k}$. Angenommen s_n konvergiert gegen c. Dann gibt es für $\varepsilon = \frac{1}{2}$ eine natürliche Zahl n_0 mit $|s_n - c| = c - s_n < \frac{1}{2}$. Hierbei ist zu beachten, dass s_n monoton wachsend ist, d.h. $s_n \leq s_m$ für $n \leq m$. Mithin gilt für $m \geq n \geq n_0$ auch $s_m - s_n < \frac{1}{2}$. Insbesondere erhalten wir für $m \geq 2n+1 \geq n \geq n_0$

  $$\begin{aligned} \frac{1}{2} > s_m - s_n &= \sum_{k=n+2}^{m+1} \frac{1}{k} \geq \frac{1}{m+1} \cdot (m+1-(n+2)+1) \\ &= 1 - \frac{n+1}{m+1} \geq 1 - \frac{n+1}{2n+2} = \frac{1}{2} \end{aligned}$$

  Dies ist ein Widerspruch. Somit gibt es keinen Grenzwert für die Reihe.

Die zentrale Frage für Reihen ist: Welche Kriterien müssen die Koeffizienten erfüllen, damit die Reihe konvergiert? Um dieser Frage nachzugehen, führen wir einen verschärften Konvergenzbegriff für Reihen ein.

Definition 5.4 *Die Reihe* $\sum_{n=0}^{\infty} a_n$ *heißt* absolut konvergent*, falls* $\sum_{n=0}^{\infty} |a_n|$ *konvergent ist.*

Klarerweise ist jede absolut konvergente Reihe auch konvergent. Die Umkehrung gilt nicht.

Beispiel: $\sum_{n=0}^{\infty} (-1)^n \frac{1}{n}$ ist konvergent aber nicht absolut konvergent.

Folgende hinreichende Kriterien können für die absolute Konvergenz von Reihen aufgestellt werden.

Lemma 5.5 *Es seien* $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *Folgen.*

1. *Ist* $\sum_{n=0}^{\infty} b_n$ *absolut konvergent und gilt* $|a_n| \leq |b_n|$ *für alle* $n \geq n_0$ *und ein geeignetes* n_0*, so ist* $\sum_{n=0}^{\infty} a_n$ *absolut konvergent (Majoranten-Kriterium).*

2. *Gibt es* $0 \leq q < 1$ *und* $n_0 \in \mathbb{N}$ *mit* $\sqrt[n]{|a_n|} \leq q$ *für alle* $n \geq n_0$*, so ist* $\sum_{n=0}^{\infty} a_n$ *absolut konvergent (Wurzel-Kriterium).*

3. *Gibt es* $0 \leq q < 1$ *und* $n_0 \in \mathbb{N}$ *mit* $|a_{n+1}| \leq q \cdot |a_n|$ *für alle* $n \geq n_0$*, so ist* $\sum_{n=0}^{\infty} a_n$ *absolut konvergent (Quotienten-Kriterium).*

Beweis: Wir zeigen die Aussagen einzeln (und nehmen zur Vereinfachung $n_0 = 0$ an):

1. Es seien $s_n =_{\text{def}} \sum_{k=0}^{n} |a_k|$ und $x =_{\text{def}} \sum_{k=0}^{\infty} |b_k|$. Dann gilt $s_n \leq x$. Weiterhin gilt $s_n \leq s_m$ für alle $n \leq m$, d.h. $(s_n)_{n\in\mathbb{N}}$ ist monoton wachsend. Wir definieren
$$s =_{\text{def}} \sup\{\ s_n \mid n \in \mathbb{N}\ \}.$$
Wegen $s_n \leq x$ für alle $n \in \mathbb{N}$ existiert s als reelle Zahl und es gilt $\lim_{n\to\infty} s_n = s$. Letzteres ist wie folgt einzusehen: Für $\varepsilon > 0$ gibt es ein n_0 mit $s_{n_0} > s - \varepsilon$ nach Definition von s als Supremum für die Menge aller Folgenglieder s_n. Da s_n monoton wachsend ist, folgt $s \geq s_n > s - \varepsilon$ für alle $n \geq n_0$. Mithin gilt $|s_n - s| < \varepsilon$ für alle $n \geq n_0$. Somit konvergiert s_n gegen s.

2. Mit $\sqrt[n]{|a_n|} \leq q$ gilt $|a_n| \leq q^n$ für alle n. Folglich ist $\sum_{n=0}^{\infty} |a_n|$ konvergent nach dem Majorantenkriterium, denn es gilt

$$\sum_{n=0}^{\infty} |a_n| \;\leq\; \sum_{n=0}^{\infty} q^n \;=\; \frac{1}{1-q} \qquad \text{für } 0 \leq q < 1$$

3. Mit $|a_{n+1}| \leq q \cdot |a_n|$ gilt $|a_n| \leq q^n \cdot |a_0|$ für alle n. Folglich ist $\sum_{n=0}^{\infty} |a_n|$ konvergent nach dem Majorantenkriterium, denn es gilt

$$\sum_{n=0}^{\infty} |a_n| \;\leq\; \sum_{n=0}^{\infty} q^n \cdot |a_0| \;=\; \frac{|a_0|}{1-q} \qquad \text{für } 0 \leq q < 1$$

Damit ist das Lemma bewiesen. ■

5.3 Oberer und unterer Grenzwert

Definition 5.6 *Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen.*

1. *Konvergiert die Folge $(a\uparrow_n)_{n\in\mathbb{N}}$ mit $a\uparrow_n =_{\text{def}} \sup\{a_m | m \geq n\}$, so heißt*

$$\limsup_{n\to\infty} a_n =_{\text{def}} \lim a\uparrow_n$$

   oberer Grenzwert *von* $(a_n)_{n\in\mathbb{N}}$.

2. *Konvergiert die Folge $(a\downarrow_n)_{n\in\mathbb{N}}$ mit $a\downarrow_n =_{\text{def}} \inf\{a_m | m \geq n\}$, so heißt*

$$\liminf_{n\to\infty} a_n =_{\text{def}} \lim a\downarrow_n$$

   unterer Grenzwert *von* $(a_n)_{n\in\mathbb{N}}$.

Beispiele:

1. Für die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} (-1)^n$ gilt

$$\inf\,\{a_m | m \geq n\} = \inf\,\{-1, 1\} = \min\,\{-1, 1\} = -1$$

   für alle $n \in \mathbb{N}$. Folglich erhalten wir $\liminf_{n\to\infty} (-1)^n = -1$.
   Analog ergibt sich $\limsup_{n\to\infty} (-1)^n = 1$.

2. Für die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} \dfrac{(-1)^n}{n+1}$ gilt

$$\inf\{a_m | m \geq n\} = \begin{cases} -\dfrac{1}{n+2} & \text{falls } n \text{ gerade ist} \\ -\dfrac{1}{n+1} & \text{falls } n \text{ ungerade ist} \end{cases}$$

für alle $n \in \mathbb{N}$. Somit ergibt sich $0 \geq \liminf\limits_{n\to\infty} a_n \geq \lim\limits_{n\to\infty} -\dfrac{1}{n+1} = 0$.

Analog erhalten wir $0 \leq \limsup\limits_{n\to\infty} a_n \leq \lim\limits_{n\to\infty} \dfrac{1}{n+1} = 0$.

3. Für die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n =_{\text{def}} (-1)^n \cdot n$ existieren weder $\limsup\limits_{n\to\infty} a_n$ noch $\liminf\limits_{n\to\infty} a_n$.

4. Ohne Beweis erwähnen wir $\limsup\limits_{n\to\infty} \sin n = 1$ und $\liminf\limits_{n\to\infty} \sin n = -1$.

Proposition 5.7 *Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen.*

1. $-\infty \leq \liminf\limits_{n\to\infty} a_n \leq \limsup\limits_{n\to\infty} \leq +\infty$.
2. *Die Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert genau dann gegen c, wenn $\liminf\limits_{n\to\infty} a_n = \limsup\limits_{n\to\infty} a_n = c$.*
3. *Gibt es $c, d \in \mathbb{R}$ mit $c \leq a_n \leq d$ für alle $n \geq n_0$ und ein geeignetes $n_0 \in \mathbb{N}$, so exitsieren $\limsup_{n\to\infty} a_n$ sowie $\liminf_{n\to infty} a_n$ und es gilt*

$$c \leq \liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n \leq d.$$

Beweis: Die erste und die zweite Aussage sind einfache Übungsaufgaben zur Vertiefung der Begriffsbildungen.

Beim Beweis der dritten Aussage beschränken wir uns auf den Fall des oberen Grenzwertes. Es gelte $c \leq a_n \leq d$ für alle $n \geq n_0$ mit geeignetem $n_0 \in \mathbb{N}$. Da d somit eine obere Schranke von $\{a_n | n \geq n_0\}$ ist, existiert $a\uparrow_n = \sup\{a_m | m \geq n\}$ stets und es gilt $a\uparrow_n \leq d$. Außerdem gilt stets $a\uparrow_n \geq a\uparrow_{n+1}$. Wegen $a\uparrow_n \geq a_n \geq c$ für alle $n \geq n_0$ existiert $\inf\{a\uparrow_n | n \geq n_0\}$. Wir definieren $s =_{\text{def}} \lim\limits_{n\to\infty} \inf\{a\uparrow_n | n \geq n_0\}$. Es gilt $s = \lim\limits_{n\to\infty} a\uparrow_n = \limsup\limits_{n\to\infty} a_n$. Die ist wie folgt einzusehen: Für jedes $\varepsilon > 0$ gibt es ein $n' \geq n_0$ mit $a\uparrow_{n'} \leq s + \varepsilon$, da anderenfalls s nicht die kleinste obere Schranken wäre. Da die Folge der $a\uparrow_n$ monoton fallend ist, folgt somit $s \leq a\uparrow_n \leq s + \varepsilon$ für alle $n \geq n'$. Mithin gilt $|a\uparrow_n - s| \leq \varepsilon$ für alle $n \geq n'$. ∎

Proposition 5.8 *Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen.*

1. $\liminf\limits_{n\to\infty} (-a_n) = -\limsup\limits_{n\to\infty} a_n$.
2. *Gilt $a_n \neq 0$ für alle $n \in \mathbb{N}$ und gilt $\limsup\limits_{n\to\infty} a_n \neq 0$, so gilt*

$$\liminf_{n\to\infty} a_n^{-1} = \left(\limsup_{n\to\infty} a_n\right)^{-1}.$$

5.4 Asymptotik von Folgen und Funktionen

Im Folgenden betrachten wir nur Funktionen $f : \mathbb{N} \to \mathbb{R}_{>0}$ bzw. Folgen $(f(n))_{n\in\mathbb{N}}$ mit $f(n) > 0$. Die Begriffsbildungen in diesem Abschnitte können jedoch auf beliebige Funktionen ausgedehnt werden.

Definition 5.9 *Es seien $f, g \in \mathbb{N} \to \mathbb{R}_{>0}$ Funktionen.*

1. $f(n) \in O(g(n)) \iff_{\text{def}} (\exists c > 0)(\exists n_0)(\forall n \geq n_0)\,[f(n) \leq c \cdot g(n)]$
2. $f(n) \in \Omega(g(n)) \iff_{\text{def}} (\exists c > 0)(\exists n_0)(\forall n \geq n_0)\,[f(n) \geq c \cdot g(n)]$
3. $f(n) \in \Theta(g(n)) \iff_{\text{def}} f(n) \in O(g(n)) \cap \Omega(g(n))$.
4. $f(n) \in o(g(n)) \iff_{\text{def}} (\forall c > 0)(\exists n_0)(\forall n \geq n_0)\,[f(n) \leq c \cdot g(n)]$
5. $f(n) \in \omega(g(n)) \iff_{\text{def}} (\forall c > 0)(\exists n_0)(\forall n \geq n_0)\,[f(n) \geq c \cdot g(n)]$

Für die einzelnen Symbole verwenden wir folgende Sprechweisen, die die zugehörigen Anschauungen widerspiegeln:

$f(n) \in O(g(n))$	entspricht:	„$f(n)$ wächst höchstens so schnell wie $g(n)$ “
$f(n) \in \Omega(g(n))$	entspricht:	„$f(n)$ wächst mindestens so schnell wie $g(n)$ “
$f(n) \in \Theta(g(n))$	entspricht:	„$f(n)$ wächst genauso schnell wie $g(n)$ “
$f(n) \in o(g(n))$	entspricht:	„$f(n)$ wächst langsamer als $g(n)$ “
$f(n) \in \omega(g(n))$	entspricht:	„$f(n)$ wächst schneller als $g(n)$ “

Bemerkungen:

1. Die übliche Verwendung der O-Notation ist $f(n) = O(g(n))$. Diese Schreibweise ist aber mit Vorsicht gebrauchen. Beispielsweise wird man häufig in Laufzeitabschätzungen von Algorithmen argumentative Ketten der Form $f(n) = O(n) = O(n^2)$ lesen. Klarerweise gilt aber natürlich nicht die Gleichheit der letzten beiden *Mengen von Funktionen*.

2. Offensichtlich gelten die folgenden Zusammenhänge für Funktionen $f, g : \mathbb{N} \to \mathbb{R}_{>0}$:

$$\begin{aligned} f(n) \in \Omega(g(n)) &\iff g(n) \in O(f(n)) \\ f(n) \in \omega(g(n)) &\iff g(n) \in o(f(n)) \end{aligned}$$

Beispiele:

1. Es gilt $n + 9996 \in O(n^2)$ mit $c = 1$ und $n_0 = 1$ (oder $c = 2$ und $n_0 = 71$).
2. Es gilt $100n^5+200n^4+n^3+1.000.357n^2+7n+33 \in O(n^5)$ mit $c = 1.000.698$ und $n_0 = 1$.
3. Es gilt $\log^{k+1} n \in \Omega(\log^k n)$ mit $c = 1$ und $n_0 = 1$.
4. Es gilt $n^k \in o(n^{k+1})$, denn für $c > 0$, $n_0 =_{\text{def}} \left\lceil \frac{1}{c} \right\rceil$ und alle $n \geq n_0$ folgt:

$$n^k = c \cdot \frac{1}{c} \cdot n^k \leq c \cdot n_0 \cdot n^k \leq c \cdot n^{k+1}$$

5. Es gilt $\log n^k \in \Theta(\log n^{k+1})$ für alle $k \in \mathbb{N}_+$.
6. Es gilt $3^n \in \omega(2^n)$, denn für $c > 0$, $n_0 =_{\text{def}} \max\{0, \lceil \log_{3/2} c \rceil\}$ und alle $n \geq n_0$ folgt:

$$3^n = \left(\frac{3}{2}\right)^n \cdot 2^n \geq \left(\frac{3}{2}\right)^{n_0} \cdot 2^n \geq \left(\frac{3}{2}\right)^{\log_{3/2} c} \cdot 2^n \geq c \cdot 2^n$$

7. Die Funktionen $\max\{1, (-1)^n \cdot n^3\}$ und n^2 sind unvergleichbar bezüglich der fünf Notationen.

Proposition 5.10 *Es seien $f, g : \mathbb{N} \to \mathbb{R}_{>0}$ Funktionen. Dann gilt:*

1. $f(n) \in O(g(n)) \iff 0 \leq \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty.$
2. $f(n) \in \Omega(g(n)) \iff 0 \leq \limsup_{n\to\infty} \frac{g(n)}{f(n)} < \infty.$
3. $f(n) \in \Theta(g(n)) \iff 0 < \liminf_{n\to\infty} \frac{f(n)}{g(n)} \leq \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty.$
4. $f(n) \in o(g(n)) \iff \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0.$
5. $f(n) \in \omega(g(n)) \iff \lim_{n\to\infty} \frac{g(n)}{f(n)} = 0.$

Beweis: Wir zeigen die Aussagen einzeln:

1. Auch hier zeigen wir beide Richtungen separat.

   ($\Leftarrow$): Es gelte $0 \leq \limsup_{n\to\infty} f(n)/g(n) = c < \infty$. Dann definieren wir eine Folge $(a_n)_{n\in\mathbb{N}}$ bestehend aus $a_n =_{\text{def}} \sup\{\, f(n)/g(n) \mid m \geq n \,\}$. Da $(a_n)_{n\in\mathbb{N}}$ gegen c konvergiert, gibt es für $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $|a_n - c| < \varepsilon$ für alle $n \geq n_0$ und folglich $a_n < c + \varepsilon$ für alle $n \geq n_0$. Somit folgt $f(n)/g(n) \leq a_n < c + \varepsilon$ bzw. $f(n) \leq (c+\varepsilon) \cdot g(n)$ für alle $n \geq n_0$. Mithin gilt $f(n) \in O(g(n))$.

   ($\Rightarrow$): Es gelte $f(n) \in O(g(n))$. Es gibt also $c > 0$ und $n_0 \in \mathbb{N}$ mit $0 \leq f(n) \leq c \cdot g(n)$ für alle $n \geq n_0$. Folglich gilt $0 \leq f(n)/g(n) \leq c$ für alle $n \geq n_0$. Nach Proposition 5.7 existiert $\limsup_{n\to\infty} f(n)/g(n)$ und es gilt

   $$0 \leq \limsup_{n\to\infty} \frac{f(n)}{g(n)} \leq c < \infty.$$

2. Folgt aus der Äquivalenz $f(n) \in \Omega(g(n)) \Leftrightarrow g(n) \in O(f(n))$ und der ersten Aussage.

3. Wieder zeigen wir beide Richtungen separat.

   ($\Leftarrow$): Es gelte $f(n) \in \Theta(g(n))$, d.h., $f(n) \in O(g(n)) \cap \Omega(g(n))$. Folglich gilt:

   $$0 \leq \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty, \qquad 0 \leq \limsup_{n\to\infty} \frac{g(n)}{f(n)} = c < \infty$$

   Nach Proposition 5.8 erhalten wir

   $$\liminf_{n\to\infty} \frac{f(n)}{g(n)} = \left(\limsup_{n\to\infty} \frac{g(n)}{f(n)}\right)^{-1} = \frac{1}{c} > 0$$

   und mithin

   $$0 < \liminf_{n\to\infty} \frac{f(n)}{g(n)} \leq \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty$$

   ($\Rightarrow$): Gilt $0 < \liminf_{n\to\infty} \frac{f(n)}{g(n)} \leq \limsup_{n\to\infty} \frac{f(n)}{g(n)} < \infty$, so folgt einerseits $f(n) \in O(g(n))$ direkt. Andererseits folgt wieder wegen

   $$0 < \liminf_{n\to\infty} \frac{f(n)}{g(n)} = \left(\limsup_{n\to\infty} \frac{f(n)}{g(n)}\right)^{-1}$$

   auch $\limsup_{n\to\infty} \frac{g(n)}{f(n)} < \infty$, d.h., $f(n) \in \Omega(g(n))$. Folglich gilt $f(n) \in \Theta(g(n))$.

4. Übungsaufgabe.

5. Folgt aus der Äquivalenz $f(n) \in \omega(g(n)) \Leftrightarrow g(n) \in o(f(n))$ und der vierten Aussage.

Damit ist die Proposition bewiesen. ■

5.5 Potenzreihen

Definition 5.11 *Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Eine Reihe der Form*

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n$$

mit $x, x_0 \in \mathbb{R}$ heißt Potenzreihe *(um den Enwicklungspunkt x_0).*

Theorem 5.12 *Zu jeder Potenzreihe $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ existiert eine Konvergenzradius R mit $0 \le R \le \infty$ und*

1. $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ *ist absolut konvergent, falls $|x-x_0| < R$,*
2. $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ *ist divergent, falls $|x-x_0| > R$.*

Eine allgemeine Aussage für den Konvergenzradius, d.h. $|x-x_0| = R$, ist nicht möglich.

Lemma 5.13 *Es sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit $a_n \neq 0$ für alle $n \in \mathbb{N}$.*

1. *Konvergiert (oder divergiert bestimmt) $\frac{|a_n|}{|a_{n+1}|}$, so gilt $\lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = R$ (Quotienten-Kriterium).*
2. *Konvergiert (oder divergiert bestimmt) $\frac{1}{\sqrt[n]{|a_n|}}$, so gilt $\lim_{n\to\infty} \frac{1}{\sqrt[n]{|a_n|}} = R$ (Wurzel-Kriterium).*

Hierbei ist die bestimmte Divergenz wie folgt definiert: Eine Folge $(a_n)_{n\in\mathbb{N}}$ divergiert bestimmt *gegen $+\infty$, falls gilt*

$$(\forall c > 0)(\exists n_0 \in \mathbb{N})(\forall n \ge n_0)[a_n > c].$$

Gilt die Aussage mit $c < 0$ statt $c > 0$, so divergiert die Folge bestimmt gegen $-\infty$.

Beispiel: Die Konvergenzradien für folgende Potenzreihen (um den Entwicklungspunkt $x_0 = 0$) können mittels des Wurzel-Kriteriums aus Lemma 5.13 bestimmt werden:

- $\sum_{n=0}^{\infty} x^n$ besitzt den Konvergenzradius $R = 1$ mit Divergenz für $|x| = 1$

- $\sum_{n=0}^{\infty} \frac{x^n}{n}$ besitzt den Konvergenzradius $R = 1$ mit Divergenz für $x = 1$ und Konvergenz für $x = -1$
- $\sum_{n=0}^{\infty} \frac{x^n}{n^2}$ besitzt den Konvergenzradius $R = 1$ mit Konvergenz für $|x| = 1$

Lemma 5.14 *Es sei* $f(x) =_{\text{def}} \sum_{n=0}^{\infty} a_n \cdot x^n$ *eine Potenzreihe (um den Entwicklungspunkt* $x_0 = 0$*) mit Konvergenzradius* $R > 0$*, wobei* $a_n \neq 0$ *für mindestens ein* $n \in \mathbb{N}$*. Dann gibt es ein* $r \in \mathbb{R}$ *mit* $0 < r < R$*, so dass* $f(x)$ *in der Menge* $U_r =_{\text{def}} \{\ x \mid x \in \mathbb{R}, |x| < r\ \}$ *nur endlich viele Nullstellen besitzt.*

Beweis: Bevor wir die Aussage beweisen, leiten wir eine Ungleichung her, auf der der eigentliche Beweis basiert. Es sei $N =_{\text{def}} \min\{\ n \mid a_n \neq 0\ \} < \infty$. Dann gilt für jedes r mit $0 < r < R$ und jedes $x \in U_r$:

$$\begin{aligned}
\left|f(x) - a_N \cdot x^N\right| &= \left|\left(a_N \cdot x^N + \sum_{n=N+1}^{\infty} a_n \cdot x^n\right) - a_N \cdot x^N\right| = \left|\sum_{n=N+1}^{\infty} a_n \cdot x^n\right| \\
&\leq \sum_{n=N+1}^{\infty} |a_n| \cdot |x|^n = \sum_{n=N+1}^{\infty} |a_n| \cdot \underbrace{|x| \cdot \cdots \cdot |x|}_{n-(N+1)} \cdot \underbrace{|x| \cdot \cdots \cdot |x|}_{N+1} \\
&\leq \sum_{n=N+1}^{\infty} |a_n| \cdot \underbrace{r \cdot \cdots \cdot r}_{n-(N+1)} \cdot \underbrace{|x| \cdot \cdots \cdot |x|}_{N+1} \\
&= |x|^{N+1} \cdot \underbrace{\sum_{n=N+1}^{\infty} |a_n| \cdot r^{n-(N+1)}}_{s =_{\text{def}}} \\
&= s \cdot |x|^{N+1}
\end{aligned}$$

Hierbei ist zweierlei zu beachten: Einmal hängt s nicht mehr von x ab und zweitens ist s tatsächlich definiert, da wegen $r < R$ die Reihe in jedem Fall konvergiert.

Wir führen nun einen Widerspruchsbeweis, um die Aussage zu beweisen. Dazu fixieren wir zunächst ein $r \in \mathbb{R}$ mit $0 < r < R$. Wir nehmen an, dass f in jeder Umgebung $U_{r/k}$ für alle $k \in \mathbb{N}_+$ unendlich viele Nullstellen besitzt. Wir betrachten eine beliebige Folge $(x_k)_{k \in \mathbb{N}_+}$, wobei x_k gerade eine Nullstelle von f in $U_{r/k} \setminus \{0\}$ ist. Mit obiger Abschätzung gilt damit für alle $k \in \mathbb{N}_+$:

$$s \cdot |x_k|^{N+1} \geq \left|f(x_k) - a_N \cdot x_k^N\right| = |a_N| \cdot |x_k|^N$$

Folglich ist $|a_N| \leq s \cdot |x_k|$ für alle $k \in \mathbb{N}_+$. Nun gilt aber:

$$0 \leq |a_N| \leq \lim_{k \to \infty} s \cdot |x_k| \leq s \cdot \lim_{k \to \infty} \frac{r}{k} = 0$$

Dies ist jedoch ein Widerspruch zu $a_N \neq 0$. Folglich gibt es ein $k \in \mathbb{N}_+$, sodass f in $U_{r/k}$ nur endlich viele Nullstellen besitzt. Damit ist das Lemma bewiesen. ■

Theorem 5.15 *Es seien* $f_a(x) =_{\text{def}} \sum_{n=0}^{\infty} a_n \cdot x^n$ *und* $f_b(x) =_{\text{def}} \sum_{n=0}^{\infty} b_n \cdot x^n$ *zwei Potenzreihen (um* $x_0 = 0$*) mit den Konvergenzradien* $R_a \geq R_b > 0$*. Gibt es eine reelle Zahl* r *mit* $0 < r < R_b$ *und* $f_a(x) = f_b(x)$ *für alle* x *mit* $|x| < r$*, so gilt* $a_n = b_n$ *für alle* $n \in \mathbb{N}$*.*

Beweis: Wir betrachten die Funktion $f =_{\text{def}} f_a - f_b$ mit der zugehörigen Potenzreihe $\sum_{n=0}^{\infty}(a_n - b_n)x^n$. Dann gilt $f(x) = 0$ für alle x mit $|x| < r$. Für alle $0 < r^* < r$ existieren also unendlich viele Nullstellen von f in den Mengen U_{r^*}. Nach Lemma 5.14 ist dies für eine Potenzreihe mit wenigstens einem von 0 verschiedenen Koeffizienten nicht möglich. Mithin muss $a_n - b_n = 0$ bzw. $a_n = b_n$ für alle $n \in \mathbb{N}$ gelten. Damit ist das Theorem bewiesen. ■

Wie können wir nun Potenzreihen für gegebenen Funktionen bestimmen? Eine Antwort darauf gibt die folgende Definition: Für eine beliebig oft differenzierbare Funktion f heißt

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} \cdot (x - x_0)^n$$

TAYLOR-*Reihe* von f am Punkt x_0. Hierbei bezeichnet wie üblich $f^{(n)}(x_0)$ den Wert der n-ten Ableitung von f nach x an der Stelle x_0.

Beispiele: Einige wichtige TAYLOR-Reihen für Funktionen (ohne Begründung der Konvergenzradien) sind wie folgt:

- $e^x = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot x^n$ für alle $x \in \mathbb{R}$ wegen $(e^x)' = e^x$
- $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ für alle $-1 < x < 1$ wegen $\left(\frac{1}{1-x}\right)^{(n)} = \frac{n!}{(1-x)^{n+1}}$
- $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \cdot x^n$ für $-1 < x \leq 1$ wegen $\ln(1+0) = 0$ und $(\ln(1+x))^{(n)} = \frac{(-1)^{n+1}(n-1)!}{(1+x)^n}$. Ein interessanter Spezialfall dieser Potenzreihe ergibt sich für $x = 1$:

  $$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \cdot 1^n = -\ln(1+1) = -\ln 2 = -0,6931\ldots$$

Theorem 5.16 *Jede Potenzreihe* $f(x) =_{\text{def}} \sum_{n=0}^{\infty} a_n(x-x_0)^n$ *mit dem Konvergenzradius* $R > 0$ *ist für alle* x *mit* $|x - x_0| < R$ *gleich ihrer* TAYLOR*-Reihe.*

Beweis: Wir beweisen die Aussage lediglich für $x_0 = 0$. Zunächst halten wir fest, dass die Ableitung der Potenzreihe von $f(x)$

$$\left(\sum_{n=0}^{\infty} a_n \cdot x^n\right)' = \sum_{n=0}^{\infty} (a_n \cdot x^n)' = \sum_{n=1}^{\infty} n a_n \cdot x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} \cdot x^n$$

den gleichen Konvergenzradius R wie die Potenzreihe von $f(x)$ besitzt. Damit können wir $f(x)$ innerhalb des Konvergenzradius der Potenzreihe beliebig oft differenzieren. Wir erhalten im Einzelnen folgende Ableitungen:

$$\begin{array}{rcccccccccccc}
f(x) &=& a_0 &+& a_1 x &+& a_2 x^2 &+& a_3 x^3 &+& \cdots &+& a_n x^n &+& \cdots \\
f'(x) &=& & & a_1 &+& 2a_2 x &+& 3a_3 x^2 &+& \cdots &+& n a_n x^{n-1} &+& \cdots \\
f''(x) &=& & & & & 2a_2 &+& 6a_3 x &+& \cdots &+& n(n-1) a_n x^{n-2} &+& \cdots \\
f'''(x) &=& & & & & & & 6a_3 &+& \cdots &+& n(n-1)(n-2) a_n x^{n-3} &+& \cdots \\
\vdots & & & & & & & & & & & & & & \\
f^{(n)}(x) &=& & & & & & & & & & & n!\, a_n &+& \cdots \\
\vdots & & & & & & & & & & & & & &
\end{array}$$

Insbesondere gilt $f^{(n)}(0) = n! \cdot a_n$ für alle $n \in \mathbb{N}$. Durch Umstellung nach a_n ergibt sich das Theorem. ∎

Eine TAYLOR-Reihe direkt zu bestimmen, ist mitunter sehr aufwändig. Daher wollen wir zum Abschluss dieses Kapitels noch einige „Tricks“ und Methoden zur Bestimmung von Potenzreihen exemplarisch vorführen.

Summen und Produkte von Potenzreihen. Wir verwenden folgende Regeln für die Summe bzw. das Produkt zweier Potenzreihen:

$$\begin{aligned}
\left(\sum_{n=0}^{\infty} a_n \cdot x^n\right) + \left(\sum_{n=0}^{\infty} b_n \cdot x^n\right) &= \sum_{n=0}^{\infty} (a_n + b_n) \cdot x^n \\
\left(\sum_{n=0}^{\infty} a_n \cdot x^n\right) \cdot \left(\sum_{n=0}^{\infty} b_n \cdot x^n\right) &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k \cdot b_{n-k}\right) \cdot x^n
\end{aligned}$$

Die zweite Formel wird *Konvolution* zweier Potenzreihen genannt.

Beispiel: Wir wollen die Potenzreihe von $(1-x)^{-2}$ bestimmen. Dazu verwenden wir die geometrische Reihe und erhalten mittels Konvolution

$$\frac{1}{(1-x)^2} = \left(\sum_{n=0}^{\infty} x^n\right)^2 = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} 1\right) \cdot x^n = \sum_{n=0}^{\infty} (n+1) \cdot x^n$$

Durch einfache Summenbildung können wir daraus folgenden Zusammenhang zwischen Potenzreihe und Funktion herleiten:

$$\begin{aligned}\sum_{n=0}^{\infty} n \cdot x^n &= \sum_{n=0}^{\infty}(n+1-1)\cdot x^n &= \left(\sum_{n=0}^{\infty}(n+1)\cdot x^n\right) - \left(\sum_{n=0}^{\infty} x^n\right) \\ &&= \frac{1}{(1-x)^2} - \frac{1}{1-x} = \frac{x}{(1-x)^2}\end{aligned}$$

Alternativ kann dieser Zusammenhang auch mittels Ausklammern und Indexverschiebung gezeigt werden:

$$\sum_{n=0}^{\infty} n \cdot x^n = \sum_{n=1}^{\infty} n \cdot x^n = x \cdot \sum_{n=1}^{\infty} n \cdot x^{n-1} = x \cdot \sum_{n=0}^{\infty}(n+1)\cdot x^n = \frac{x}{(1-x)^2}$$

Substitution. Eine wichtige Methode, um neue aus bereits bekannten Potenzreihen abzuleiten, ist die Substitutionsmethode.

Beispiel: Wir wollen die Potenzreihe von $f(x) =_{\text{def}} \left(1+x^2\right)^{-1}$ bestimmen. Mit der Substitution $z =_{\text{def}} -x^2$ besteht die Aufgabe nun darin, die Potenzreihe der Funktion $h(z) =_{\text{def}} (1-z)^{-1}$ zu bestimmen. Somit gilt also:

$$\frac{1}{1+x^2} = \frac{1}{1-z} = \sum_{n=0}^{\infty} z^n = \sum_{n=0}^{\infty}\left(-x^2\right)^n = \sum_{n=0}^{\infty}(-1)^n \cdot x^{2n}$$

Differenzieren und Integrieren. Eine (differenzierbare und integrierbare) Funktion $f(x)$ ist gleich der ersten Ableitung ihrer Stammfunktion bzw. gleich der Stammfunktion der ersten Ableitung. Somit können wir die gleichen Operationen für die zugehörigen Potenzreihen durchführen und erhalten eine Potenzreihe für die Funktion f.

Beispiel:

1. Angenommen wir möchten wissen, welche Funktion $f(x)$ zur Potenzreihe $\sum_{n=0}^{\infty}(n+1)\cdot x^n$ gehört (ohne die Antwort bereits zu kennen). Dann könnten wir zunächst die Stammfunktion $F(x)$ durch koeffizientenweises Integrieren bestimmen:

$$\begin{aligned}F(x) &= \int f(x)dx = \int\left(\sum_{n=0}^{\infty}(n+1)\cdot x^n\right)dx = \sum_{n=0}^{\infty}\int(n+1)\cdot x^n dx \\ &= \sum_{n=0}^{\infty} x^{n+1} = x\cdot\sum_{n=0}^{\infty} x^n = \frac{x}{(1-x)}\end{aligned}$$

Wenn wir nun $F(x)$ wieder differenzieren, erhalten wir:

$$f(x) = F'(x) = \left(\frac{x}{1-x}\right)' = \frac{1 \cdot (1-x) - x \cdot (-1)}{(1-x)^2} = \frac{1}{(1-x)^2}$$

2. Wir wollen die Potenzreihe von $f(x) =_{\text{def}} \arctan x$ bestimmen. Wir bilden zunächst die Ableitung von $f(x)$ mittels der Ableitungsregel für inverse Funktionen $\left(h^{-1}\right)'(x) = \frac{1}{h'(h^{-1}(x))}$. In unserem Fall müssen wir also zusätzlich die erste Ableitung der Tangensfunktion bestimmen:

$$(\tan x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\sin^2 x + \cos^2 x}{\cos^2 x} = 1 + \tan^2 x$$

Somit folgt:

$$f'(x) = (\arctan x)' = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1+x^2}$$

Mithin ergibt sich (aus obigem Beispiel):

$$f'(x) = \sum_{n=0}^{\infty} (-1)^n x^{2n}$$

Durch Integration erhalten wir schließlich die Potenzreihe:

$$\begin{aligned} f(x) &= \int f'(x)dx = \int \left(\sum_{n=0}^{\infty} (-1)^n x^{2n}\right) dx = \sum_{n=0}^{\infty} (-1)^n \int x^{2n} dx \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \cdot x^{2n+1} \end{aligned}$$

Inbesondere ergibt sich für $x = 1$ eine einfache Reihe zur (approximativen) Berechnung der Zahl π:

$$\frac{\pi}{4} = \arctan 1 = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \cdots$$

Potenzreihen haben vielfältige Anwendungen in der Numerik. Funktionen ohne geschlossene Darstellung können vielfach durch Potenzreihen approximativ behandelt werden.

Beispiel: Das Gauss*'sche Fehlerintegral* dient der Beschreibung der Normalverteilung und ist definiert als:

$$\Phi(x) =_{\text{def}} \frac{2}{\sqrt{\pi}} \int_{\infty}^{x} e^{-t^2} dt$$

Das verwendete Integral kann nicht geschlossen als Stammfunktion dargestellt werden. Durch Verwendung der Potenzreihe

$$e^{-t^2} = \sum_{n=0}^{\infty} \frac{1}{n!} \cdot \left(-t^2\right)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot t^{2n}$$

erhalten eine Darstellung von $\Phi(x)$ als Potenzreihe:

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \cdot x^{2n} \right) dx = \frac{2}{\sqrt{\pi}} \cdot \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \cdot (2n+1)} \cdot x^{2n+1}$$

Lineare Algebra 6

In diesem Kapitel geben wir eine kurze Einführung in die Grundbegriffe der linearen Algebra.

6.1 Lineare Räume

Definition 6.1 *Ein* linearer Raum *ist eine Menge V mit zwei Funktionen (Operationen) $+ : V \times V \to V$ (Vektoraddition) und $\cdot : \mathbb{R} \times V \to V$ (skalare Multiplikation), sodass folgende Eigenschaften erfüllt sind:*

1. *Für alle $u, v, w \in V$ gilt $(u + v) + w = u + (v + w)$*
2. *Es gibt ein $e \in V$, sodass $e + v = v + e = v$ für alle $v \in V$ gilt.*
3. *Für alle $v \in V$ gibt es ein $u \in V$ mit $v + u = u + v = e$.*
4. *Für alle $u, v \in V$ gilt $u + v = v + u$.*
5. *Für alle $a, b \in \mathbb{R}$ und $v \in V$ gilt $(a + b) \cdot v = a \cdot v + b \cdot v$*
6. *Für alle $a \in \mathbb{R}$ und $v, w \in V$ gilt $a \cdot (v + w) = a \cdot v + a \cdot w$*
7. *Für alle $a, b \in \mathbb{R}$ und $v \in V$ gilt $(a \cdot b) \cdot v = a \cdot (b \cdot v)$*
8. *Für alle $v \in V$ gilt $1 \cdot v = v$*

Die Elemente von V heißen Vektoren.

Die ersten drei Eigenschaften der Definition etablieren $(V, +)$ als Gruppe. Zusammen mit der vierten Eigenschaft wird $(V, +)$ zur abelschen Gruppe.

Beispiele: Die folgenden Menge mit den geeigneten zugehörigen Operationen bilden lineare Räume:

- $\mathbb{R}^n$ für $n \in \mathbb{N}_+$ mit komponentenweiser Addition und Multiplikation mit Konstanten:

$$v + w = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} \text{ sowie } a \cdot v = \begin{pmatrix} av_1 \\ av_2 \\ \vdots \\ av_n \end{pmatrix}$$

- Die Menge aller Polynome $\sum_{i=0}^{n} a_i \cdot x^i$ vom Grad $\leq n$, $a_i \in \mathbb{R}$ mit der üblichen Addition und Multiplikation von Konstanten:

$$\left(\sum_{i=0}^{n} a_i \cdot x^i\right) + \left(\sum_{i=0}^{n} b_i \cdot x^i\right) = \sum_{i=0}^{n} (a_i + b_i) \cdot x^i \qquad \text{sowie}$$
$$c \sum_{i=0}^{n} a_i \cdot x^i = \sum_{i=0}^{n} (c \cdot a_i) \cdot x^i$$

- Die Menge aller konvergenten Folgen mit komponentenweise Addition und Multiplikation mit Konstanten
- Die Menge aller (formalen) Potenzreihen

Definition 6.2 *Es seien V ein linearer Raum und $\emptyset \neq W \subseteq V$. Dann heißt W* linearer Unterraum *(von V) , falls für alle $v, w \in V$ und $a \in \mathbb{R}$ die folgenden Bedingungen erfüllt sind:*

1. *Sind $v, w \in W$, so ist $v + w \in W$ (Abgeschlossenheit unter $+$)*
2. *Ist $v \in W$, so ist $a \cdot v \in W$ (Abgeschlossenheit unter $\cdot$)*

Die Benennung einer Teilmenge W von V, die die beiden obigen Eigenschaften erfüllt, als Unterraum ist plausibel.

Proposition 6.3 *Jeder lineare Unterraum eines linearen Raumes ist ein linearer Raum.*

Beweis: Offensichtlich durch Überprüfung der Eigenschaften eines linearen Raumes. ■

Beispiele: Folgende Mengen bilden Unterräume in den entsprechenden linearen Räumen:

- $\mathbb{R}^2 \times \{0\}$ ist ein Unterraum von $\mathbb{R}^3$, denn es gilt

$$\underbrace{\begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix}}_{\in \mathbb{R}^2 \times \{0\}} + \underbrace{\begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix}}_{\in \mathbb{R}^2 \times \{0\}} = \underbrace{\begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ 0 \end{pmatrix}}_{\in \mathbb{R}^2 \times \{0\}} \quad \text{sowie} \quad a \cdot \underbrace{\begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix}}_{\in \mathbb{R}^2 \times \{0\}} = \underbrace{\begin{pmatrix} av_1 \\ av_2 \\ 0 \end{pmatrix}}_{\in \mathbb{R}^2 \times \{0\}}$$

- Die Menge der Polynome vom Grad m ist ein Unterraum in der Menge der Polynome vom Grad $n \geq m$, denn

$$\left(\sum_{i=0}^{m} a_i \cdot x^i\right) + \left(\sum_{i=0}^{m} b_i \cdot x^i\right) = \sum_{i=0}^{m} (a_i + b_i) \cdot x^i \qquad \text{und}$$
$$c \sum_{i=0}^{m} a_i \cdot x^i = \sum_{i=0}^{m} (c \cdot a_i) \cdot x^i$$

sind wieder Polynome vom Grad m.

Definition 6.4 *Es seien V ein linearer Raum und $v_1, \ldots, v_n \in V$ Vektoren. Der Vektor $w \in V$ heißt* Linearkombination *von $v_1, \ldots, v_n$, falls $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ existieren mit*

$$w = \lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n.$$

Die Menge aller Linearkombinationen von $v_1, \ldots, v_n$ heißt lineare Hülle *(engl.* span*) von $v_1, \ldots, v_n$ und wird mit* $\text{span}\{v_1, \ldots, v_n\}$ *bezeichnet.*

Proposition 6.5 *Es seien V ein linearer Raum und $v_1, \ldots, v_n \in V$ Vektoren. Dann ist die lineare Hülle von $v_1, \ldots, v_n$ ein linearer Unterraum von V.*

Beweis: Wir müssen zeigen, dass $\text{span}\{v_1, \ldots, v_n\}$ abgeschlossen unter Addition und skalarer Multiplikation ist. Es seien $w, w' \in V$ Vektoren mit $w = \lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n$ und $w' = \lambda'_1 \cdot v_1 + \cdots + \lambda'_n \cdot v_n$. Dann gilt

$$\begin{aligned} w + w' &= (\lambda_1 + \lambda'_1) \cdot v_1 + \cdots + (\lambda_n + \lambda'_n) \cdot v_n && \in \text{span}\{v_1, \ldots, v_n\} \\ c \cdot w &= (c\lambda_1) \cdot v_1 + \cdots + (c\lambda_n) \cdot v_n && \in \text{span}\{v_1, \ldots, v_n\} \end{aligned}$$

Damit ist die Proposition bewiesen. ∎

Definition 6.6 *Es sei V ein linearer Raum. Eine Menge von Vektoren $v_1, \ldots, v_n \in V$ heißt* Erzeugendensystem *von V, falls $\text{span}\{v_1, \ldots, v_n\} = V$ gilt.*

Beispiele: Die folgenden Vektoren bilden Erzeugendensysteme in den jeweiligen linearen Räumen, d.h. ihre lineare Hülle spannt immer gerade den gesamten Raum auf:

- Für $V = \mathbb{R}^2$ und die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{und} \quad v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

gilt $\text{span}\{v_1, v_2, v_3\} = \mathbb{R}^2$, denn für $x_1, x_2 \in \mathbb{R}$ gilt:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- Für $V = \mathbb{R}^2$ und die Vektoren

$$v_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

gilt $\text{span}\{v_1, v_2\} = \mathbb{R}^2$, denn für $x_1, x_2 \in \mathbb{R}$ gilt:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left(\frac{x_2}{2} - \frac{x_1}{2}\right) \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \left(\frac{x_1}{2} + \frac{x_2}{2}\right) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Definition 6.7 *Es seien V ein linearer Raum und $v_1, \ldots, v_n \in V$ Vektoren. Dann heißen* $v_1, \ldots, v_n$ linear unabhängig*, falls für alle $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ gilt:*

Ist $\lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n = 0$, so sind $\lambda_1 = \cdots = \lambda_n = 0$

Hierbei steht 0 für das neutrale Element von V (Nullvektor).

Mit anderen Worten: Sind die Vektoren $v_1, \ldots, v_n$ linear abhängig, so gibt es einen Vektor v_i, der in der linearen Hülle von $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n\}$ liegt. Dies kann man sehr leicht einsehen: Angenommen die Gleichung $\lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n = 0$ ist auch mit $\lambda_1 \neq 0$ möglich, dann folgt

$$v_1 = \left(-\frac{\lambda_2}{\lambda_1}\right) \cdot v_2 + \cdots + \left(-\frac{\lambda_n}{\lambda_1}\right) \cdot v_n,$$

d.h. v_1 ist eine Linearkombination von $v_2, \ldots, v_n$. Das Argument läßt sich natürlich auf jeden Vektor v_i mit $\lambda_i \neq 0$ verallgemeinern.

Beispiele: Folgende Vektoren verdeutlichen die Definition der linearen Unabhängigkeit im Raum $V = \mathbb{R}^2$:

- Die Vektoren $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $v_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sind linear abhängig, denn es gilt:

$$1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- Die Vektoren $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sind linear unabhängig, denn aus

$$\lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

folgt $\lambda_1 + \lambda_2 = 0$ und $\lambda_2 = 0$. Damit gilt $\lambda_1 = \lambda_2 = 0$.

Definition 6.8 *Es seien V ein linearer Raum und $\{v_1, \ldots, v_n\}$ ein Erzeugendensystem von V. Dann heißt die Menge $\{v_1, \ldots, v_n\}$* Basis *von V, falls $v_1, \ldots, v_n$ linear unabhängig sind. Die Anzahl der Vektoren einer Basis heißt* Dimension *von V und wird mit* $\dim(V)$ *bezeichnet.*

Zur Wohldefiniertheit der Dimension merken wir ohne Beweis an, dass jede (endliche) Basis von V die gleiche Anzahl von Vektoren besitzt.

Beispiel: Als Vektorraum betrachten wir den $\mathbb{R}^2$.

- Die Menge der Vektoren $\begin{pmatrix}1\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\end{pmatrix}$ und $\begin{pmatrix}1\\1\end{pmatrix}$ ist ein linear abhängiges Erzeugendensystem und damit keine Basis
- Die Menge der Vektoren $\begin{pmatrix}1\\0\end{pmatrix}$ und $\begin{pmatrix}0\\1\end{pmatrix}$ ist ein linear unabhängiges Erzeugendensystem und damit eine Basis. Mithin ist $\dim(\mathbb{R}^2) = 2$.
- Die Menge der Vektoren $\begin{pmatrix}1\\0\end{pmatrix}$ und $\begin{pmatrix}1\\1\end{pmatrix}$ sowie die Menge der Vektoren $\begin{pmatrix}0\\1\end{pmatrix}$ und $\begin{pmatrix}1\\1\end{pmatrix}$ sind ebenfalls Basen von $\mathbb{R}^2$.

Im Folgenden interessieren wir uns für eine spezielle Klasse von Basen gegebener Vektorräume. Dazu führen wir den Begriff des Skalarproduktes ein, welches eine Abstraktion des Winkelkonzeptes zwischen Vektoren darstellt. Da der Schwerpunkt in diesem Kapitel nicht auf der geometrischen Interpretation linearalgebraischer Konzepte liegt, gehen wir auf die Winkelinterpretation nicht näher ein.

Definition 6.9 *Es seien V ein linearer Raum. Eine Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ heißt (euklidisches)* Skalarprodukt*, falls für alle $u, v, w \in V$ und $a \in \mathbb{R}$ gilt:*

1. *$\langle v, w\rangle = \langle w, v\rangle$ (Symmetrie)*
2. *$\langle a \cdot v + w, u\rangle = a \cdot \langle v, u\rangle + \langle w, u\rangle$ (Linearität)*
3. *$\langle v, v\rangle \geq 0$ und $\langle v, v\rangle = 0 \Longleftrightarrow v = 0$ (positive Definitheit)*

Ein linearer Raum mit einem Skalarprodukt heißt euklidischer Raum*. Vektoren v und w in einem euklidischen Raum mit $\langle v, w\rangle = 0$ heißen* orthogonal*.*

Proposition 6.10 *Es sei V ein euklidischer Raum. Sind zwei Vektoren $v, w \in V \setminus \{0\}$ orthogonal zueinander, so sind sie linear unabhängig.*

Beweis: Wir beweisen die Kontraposition der Aussage. Es seien also $v, w \neq 0$ linear abhängige Vektoren, d.h. es gibt $\lambda_1, \lambda_2 \neq 0$ mit $\lambda_1 \cdot v + \lambda_2 \cdot w = 0$. (Eigentlich dürfen wir nur voraussetzen, dass nur eines der λ's verschieden von 0 ist, da wir aber nur zwei Vektoren betrachten, ist das andere λ somit auch stets verschieden von 0.) Damit gilt also $v = (-\lambda_2/\lambda_1) \cdot w$ und wir erhalten

$$\langle v, w\rangle \quad = \quad \langle(-\lambda_2/\lambda_1) \cdot w, w\rangle$$

$$\begin{aligned} &= -\frac{\lambda_2}{\lambda_1} \cdot \langle w, w \rangle && \text{(Linearität von } \langle \cdot, \cdot \rangle\text{)} \\ &\neq 0 && (\lambda_1, \lambda_2 \neq 0 \text{ und } \langle w, w \rangle > 0 \text{ wegen } w \neq 0) \end{aligned}$$

Damit sind v und w nicht orthogonal und die Proposition ist bewiesen. ■

Beispiel: Im $\mathbb{R}^3$ ist für Vektoren $v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ und $w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$ das Standardskalarprodukt definiert als:

$$\langle v, w \rangle \ =_{\text{def}} \ v_1 w_1 + v_2 w_2 + v_3 w_3$$

Die Überprüfung der Axiome für Skalarprodukte ist eine Übungsaufgabe.

Mit Hilfe des Skalarproduktes kann in einem euklidischen Raum die *Norm* (Länge) $\|v\|$ eines Vektors $v \in V$ definiert werden:

$$\|v\| =_{\text{def}} \sqrt{\langle v, v \rangle}$$

Definition 6.11 *Es sei V ein euklidischer Raum. $B = \{v_1, \ldots, v_n\}$ sei eine Basis von V.*

1. *B heißt* Orthogonalbasis *von V, falls $\langle v_j, v_k \rangle = 0$ für alle j, k mit $j \neq k$ gilt.*
2. *B heißt* Orthonormalbasis *von V, falls B eine Orthogonalbasis ist und $\|v\| = 1$ für alle $v \in B$ gilt.*

Beispiele: Wir betrachten wieder den $\mathbb{R}^3$ mit dem Standardskalarprodukt.

- $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ ist eine Basis aber keine Orthogonalbasis
- $\left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ ist eine Orthogonalbasis
- $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$ ist eine Orthogonalbasis

Keine der Basen ist eine Orthonormalbasis.

6.2 Lineare Abbildungen

In diesem Abschnitt interessieren wir uns für bestimmte Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^m$. Eine solche Abbildung f nennen wir *linear*, falls wir die Funktion in der Form $f(x) = A \cdot x$ schreiben können, wobei $A \in \mathbb{R}^{m \times n}$ eine Matrix und $x \in \mathbb{R}^n$ ein Vektor (eine einspaltige Matrix) ist. Das Produkt $A \cdot B$ zweier Matrizen $A \in \mathbb{R}^{\ell \times m}$ und $B \in \mathbb{R}^{m \times n}$ ist definiert als die Matrix $C \in \mathbb{R}^{\ell \times n}$ mit den Einträgen

$$c_{ij} =_{\text{def}} \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Eine typische Anwendung linearer Abbildungen ist die *Koordinatentransformation*. Wir betrachten dabei vereinfachend den $\mathbb{R}^n$ und die Vektoren $w, v_1, \ldots, v_n \in \mathbb{R}^n$. Ist w eine Linearkombination von $v_1, \ldots, v_n$, d.h. gilt $w = \lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n$, so lässt sich dies wie folgt ausdrücken:

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_{11} & v_{21} & \ldots & v_{n1} \\ v_{12} & v_{22} & \ldots & v_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ v_{1n} & v_{2n} & \ldots & v_{nn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Die Werte $\lambda_1, \ldots, \lambda_n$ heißen *Koordinaten* von w bezüglich der Vektoren $v_1, \ldots, v_n$.

Damit stellt sich als prinzipielle Frage, wie zu gegebenen Vektoren $a_1, \ldots, a_n \in \mathbb{R}^n$ für beliebige $w \in \mathbb{R}^n$ die Koordinaten bestimmt werden können. Wir formulieren dieses Problem wie folgt in ein Problem für Matrizen um. Ausgehend von der Gleichheit $w = A \cdot u$, wobei A die wie oben aus den Vektoren $a_1, \ldots, a_n$ gebildete Matrix und u ein Vektor der Koordinaten sind, bestimmen wir, falls dies möglich ist, eine Matrix $B \in \mathbb{R}^{n \times n}$ mit der Eigenschaft $B \cdot A = I$, wobei $I \in \mathbb{R}^{n \times n}$ die *Einheitsmatrix* im $\mathbb{R}^n$ ist:

$$I = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix}$$

Damit gilt dann unter Ausnutzung der Assoziativität der Matrizenmultiplikation:

$$B \cdot w = B \cdot (A \cdot u) = (B \cdot A) \cdot u = I \cdot u = u$$

Mit der Kenntnis der Matrix B, die in einem gewissen Sinne die zu A *inverse* Matrix darstellt, hätten wir das Problem der Koordinationberechnung gelöst.

Beispiel: Wir betrachten den linearen Raum $\mathbb{R}^2$. Zunächst wollen wir die inverse Matrix von $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ bestimmen, d.h. wir wollen eine Gleichung

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot w = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot u$$

in eine Gleichung

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot w = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot u$$

für geeignete b_{11}, b_{12}, b_{21} und b_{22} umwandeln. Dazu schreiben wir die Matrizen nebeneinander und versuchen durch GAUSS-Elimination die Einheitsmatrix von der linekn auf die rechten Seite zu bringen:

$$\left(\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \end{array}\right)$$

$$\left(\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ -1 & 1 & 0 & 2 \end{array}\right) \qquad \text{(Ziehe Zeile (1) von Zeile (2) ab)}$$

$$\left(\begin{array}{cc|cc} 1 & 0 & 1 & -1 \\ -1/2 & 1/2 & 0 & 1 \end{array}\right) \qquad \text{(Multipliziere Zeile (2) mit 1/2)}$$

$$\left(\begin{array}{cc|cc} 1/2 & 1/2 & 1 & 0 \\ -1/2 & 1/2 & 0 & 1 \end{array}\right) \qquad \text{(Addiere Zeile (2) zu Zeile (1))}$$

Damit gilt $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Zur Überprüfung rechnen wir nach:

$$\frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Als zweites Beispiel wollen wir einsehen, dass eine inverse Matrix nicht immer existieren muss. Dazu betrachten wir die Matrix $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Zur Berechnung der inversen Matrix müssen geeignete reelle Zahlen $a, b, c, d \in \mathbb{R}$ existieren mit

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Insbesondere muss also $a + 2b = 1$ sowie $2a + 4b = 0$ gelten, was nicht möglich ist. Mithin gibt es keine inverse Matrix.

Definition 6.12 *Es sei $A \in \mathbb{R}^{m \times n}$ für $m, n \in \mathbb{N}_+$ eine Matrix.*

1. *A heißt* quadratisch*, falls $m = n$ gilt.*
2. *$A^T \in \mathbb{R}^{n \times m}$ heißt die zu A* transponierte *Matrix, falls $a_{jk} = (a^T)_{kj}$ für alle $j \in \{1, \ldots, m\}$ und $k \in \{1, \ldots, n\}$ gilt.*
3. *A heißt* invertierbar*, falls A quadratisch ist und eine Matrix A^{-1} existiert mit $A \cdot A^{-1} = A^{-1} \cdot A = I$.*
4. *A heißt* symmetrisch*, falls $A = A^T$ gilt.*
5. *A heißt* orthogonal*, falls $A^{-1} = A^T$ gilt.*

Beispiel: Die Matrix $\frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ ist orthogonal, denn:

$$\frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Definition 6.13 *Die* Determinante *einer Matrix $A \in \mathbb{R}^{n \times n}$ ist definiert durch*

$$\det(A) =_{\text{def}} \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot \prod_{i=1}^{n} a_{i,\pi(i)}.$$

Hierbei stehen:

- *S_n für die Menge der Permutationen (Bijektionen) $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$,*
- $\operatorname{sgn}(\pi) =_{\text{def}} (-1)^{\|F(\pi)\|}$ *für das* Vorzeichen *von π mit*
- *$F(\pi) =_{\text{def}} \{ (j,k) \mid j < k \wedge \pi(j) > \pi(k) \}$, d.h., $F(\pi)$ ist gerade die Menge der* Fehlstände *der Permutation π.*

Beispiele: Die Determinante einer 2×2-Matrix ergibt sich wie folgt:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \underbrace{(+1) \cdot a_{11}a_{22}}_{\pi=\binom{12}{12}} + \underbrace{(-1) \cdot a_{12}a_{21}}_{\pi=\binom{12}{21}}$$

Für eine 3×3-Matrix erhalten wir als Determinante:

$$\begin{aligned} \det & \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= \underbrace{(+1) \cdot a_{11}a_{22}a_{33}}_{\pi=\binom{123}{123}} + \underbrace{(+1) \cdot a_{12}a_{23}a_{31}}_{\pi=\binom{123}{231}} + \underbrace{(+1) \cdot a_{13}a_{21}a_{32}}_{\pi=\binom{123}{312}} + \\ &\quad + \underbrace{(-1) \cdot a_{11}a_{23}a_{32}}_{\pi=\binom{123}{132}} + \underbrace{(-1) \cdot a_{13}a_{22}a_{31}}_{\pi=\binom{123}{321}} + \underbrace{(-1) \cdot a_{12}a_{21}a_{33}}_{\pi=\binom{123}{213}} \end{aligned}$$

Im Allgemeinen sind $n!$ Produkte von Matrixeinträgen zu bestimmen. Das folgende Theorem gibt einen wichtigen Spezialfall von Matrizen an, für die die Determinante sehr einfach zu berechnen ist.

Theorem 6.14 *Es sei $A \in \mathbb{R}^{n\times n}$ eine Matrix in oberer Dreiecksform, d.h.*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \dots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \dots & 0 & a_{nn} \end{pmatrix}$$

bzw. $a_{jk} = 0$ für alle j, k mit $j > k$ Dann gilt:

$$\det(A) = \prod_{i=1}^{n} a_{ii}$$

Beweis: Es sei $\pi \in S_n$ eine Permutation mit $\pi(j) < j$ für ein $j \in \{1, \dots, n\}$. Dann gilt $a_{j,\pi(j)} = 0$ wegen der oberen Dreiecksform der Matrix A. Somit gilt

$$\prod_{i=1}^{n} a_{i,\pi(i)} = 0.$$

Die einzige Permutation, die obige Eigenschaft nicht besitzt, ist die Identität $\pi = \mathrm{id}_n$. Insgesamt folgt damit

$$\det(A) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) \cdot \prod_{i=1}^{n} a_{i,\pi(i)} = \prod_{i=1}^{n} a_{ii}$$

und das Theorem ist bewiesen. ∎

Aus dem Beweis wird deutlich, dass das Theorem 6.14 auch für Matrizen in unterer Dreiecksform gilt. Weiterhin gibt uns Theorem 6.14 die Gauss-Elimination als Verfahren an die Hand, um die Determinante einer Matrix schneller als gemäß der Definition berechnen zu können. Folgende Regeln sind dabei zu beachten:

- Entsteht eine Matrix A' aus A durch Addieren des x-fachen von Zeile (k) zu Zeile (j) mit $j \neq k$, so gilt:
$$\det(A') = \det(A)$$
- Entsteht eine Matrix A' aus A durch Vertauschen von Zeile (k) und Zeile (j) mit $j \neq k$, so gilt:
$$\det(A') = -\det(A)$$

- Entsteht eine Matrix A' aus A durch Multiplikation von Zeile (j) mit $x \neq 0$, so gilt:

$$\det(A') = x \cdot \det(A)$$

Theorem 6.15 *Es seien $A, B \in \mathbb{R}^{n\times n}$. Dann gilt:*

1. *A ist invertierbar $\Longleftrightarrow \det(A) \neq 0$*
2. *$\det(A \cdot B) = \det(A) \cdot \det(B)$*
3. *$\det(A^{-1}) = \det(A)^{-1}$, falls A invertierbar ist*

Beweis: (*Nur dritte Aussage*) Es sei A eine invertierbare Matrix. Nach Theorem 6.14 und der zweiten Aussage gilt

$$1 = \det(I) = \det(A^{-1} \cdot A) = \det(A^{-1}) \cdot \det(A).$$

Nach der ersten Aussage ist $\det(A) \neq 0$ und wir können die beiden äußeren Ausdrücke durch $\det(A)$ teilen. Mithin gilt $\det(A^{-1}) = \det(A)^{-1}$ und die dritte Aussage des Theorems ist bewiesen. ■

6.3 Eigenwerte und Eigenvektoren

Betrachten wir eine lineare Abbildung $f : V \to V : x \mapsto Ax$ im linearen Raum V, so hängt die Matrix A von der Wahl der Basis in V ab.

Beispiel: Zunächst betrachten wir im linearen Raum $V = \mathbb{R}^2$ die Basis

$$\mathcal{A} =_{\text{def}} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

sowie die Abbildung $f : V \to V : x \mapsto M_{\mathcal{A}} \cdot x$ mit der Matrix $M_{\mathcal{A}} \in \mathbb{R}^{2\times 2}$

$$M_{\mathcal{A}} =_{\text{def}} \frac{1}{2} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

Wählen wir nunmehr im $\mathbb{R}^2$ die Basis

$$\mathcal{B} =_{\text{def}} \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

Dann ist die lineare Abbildung f von oben jetzt gegeben durch die Zuordnung $y \mapsto M_{\mathcal{B}} \cdot y$ mit der Matrix

$$M_{\mathcal{B}} =_{\text{def}} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Wieso ist das der Fall?

Wir wissen bereits aus dem letzten Abschnitt: Ist $x \in V$ ein Vektor bezüglich der Basis $\mathcal{A}$, so entspricht x dem Vektor

$$y =_{\text{def}} \underbrace{\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}^{-1}}_{B=_{\text{def}}} \cdot x = \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot x$$

bezüglich der Basis $\mathcal{B}$, wobei die Spalten der Matrix B gerade aus den Vektoren der Basis $\mathcal{B}$ bestehen. Somit muss die Gleichung

$$M_{\mathcal{A}} \cdot x = B \cdot \left(M_{\mathcal{B}} \cdot \left(B^{-1} \cdot x\right)\right) = \left(B \cdot M_{\mathcal{B}} \cdot B^{-1}\right) \cdot x$$

für alle $x \in V$ gelten. Dies ist jedoch äquivalent zu $M_{\mathcal{A}} = B \cdot M_{\mathcal{B}} \cdot B^{-1}$ bzw. $B^{-1} \cdot M_{\mathcal{A}} \cdot B = M_{\mathcal{B}}$. Mithin ergibt sich:

$$\begin{aligned} M_{\mathcal{B}} &= \frac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \frac{1}{2} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{4} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & -2 \\ 4 & 2 \end{pmatrix} \\ &= \frac{1}{4} \cdot \begin{pmatrix} 8 & 0 \\ 0 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

Die im Beispiel angegebene Matrix hat eine besonders einfache Struktur, da lediglich auf der Diagonalen der Matrix Werte, die verschieden von 0 sind, auftreten. Im Folgenden beschäftigen wir uns mit einem Verfahren–der *Hauptachsentransformation*–, mit dem wir zu einer gegebenen Matrix eine solche Diagonalmatrix sowie die zugehörige Basis bestimmen können. Zur Vereinfachung werden wir uns auf den Fall symmetrischer Matrizen beschränken. Von grundlegender Bedeutung für dieses Verfahren sind die Begriffe *Eigenwert* und *Eigenvektor*.

Definition 6.16 *Es sei V ein linearer Raum mit $\dim(V) = n$ bezüglich einer beliebigen Basis. Weiterhin sei $f : V \to V : x \mapsto A \cdot x$ eine lineare Abbildung. Ein Vektor $v \in V \setminus \{0\}$ heißt* Eigenvektor *zum* Eigenwert $\lambda \in \mathbb{R}$*, falls gilt:*

$$A \cdot v = \lambda \cdot v$$

Beispiel: Wir betrachten wiederum den linearen Raum $V = \mathbb{R}^2$ mit der Basis

$$\mathcal{A} =_{\text{def}} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

sowie die Abbildung $f : V \to V : x \mapsto A \cdot x$ mit der Matrix $A \in \mathbb{R}^{2\times 2}$

$$A =_{\text{def}} \frac{1}{2} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

Dann ist

- $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist Eigenvektor zum Eigenwert 2. Zur Überprüfung:

$$\frac{1}{2} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

- $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ist Eigenvektor zum Eigenwert 1. Zur Überprüfung:

$$\frac{1}{2} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Definition 6.17 *Es sei V ein linearer Raum mit $\dim(V) = n$ bezüglich einer beliebigen Basis. Weiterhin seien $f : V \to V : x \mapsto A \cdot x$ eine lineare Abbildung sowie $\lambda \in \mathbb{R}$ ein Eigenwert von A. Dann heißt*

$$E_\lambda(A) =_{\text{def}} \{\ v \in V \mid A \cdot v = \lambda \cdot v\ \}$$

der Eigenraum *von λ.*

Proposition 6.18 *Der Eigenraum $E_\lambda(A)$ ist ein Unterraum von V.*

Beweis: Es seien $v_1, v_2 \in E_\lambda(A)$ zwei Vektoren sowie $a \in \mathbb{R}$ ein Skalar. Dann gilt

$$A \cdot (v_1 + v_2) = A \cdot v_1 + A.v_2 = \lambda \cdot v_1 + \lambda \cdot v_2 = \lambda \cdot (v_1 + v_2)$$

Mithin gilt $v_1 + v_2 \in E_\lambda(A)$. Somit ist $E_\lambda(A)$ abgeschlossen unter Addition. Weiterhin erhalten wir

$$A \cdot (a \cdot v_1) = a \cdot (A \cdot v_1) = a \cdot \lambda \cdot v_1 = \lambda \cdot (a \cdot v_1)$$

Mithin gilt $a \cdot v_1 \in E_\lambda(A)$. Somit ist $E_\lambda(A)$ auch abgeschlossen unter Multiplikation mit Skalaren. Folglich ist $E_\lambda(A)$ ein linearer Unterraum von V. ■

Bemerkungen: Wir führen einige ergänzende Anmerkungen an:

- Jeder Vektor $v \in E_\lambda(A) \setminus \{0\}$ ist Eigenvektor zu λ.
- Der Nullvektor 0 erfüllt stets $A \cdot 0 = \lambda \cdot 0$ und wird deshalb als Eigenvektor ausgeschlossen.

- $E_{\lambda_1}(A) \cap E_{\lambda_2}(A) = \{0\}$ für Eigenwerte λ_1, λ_2 mit $\lambda_1 \neq \lambda_2$.

Definition 6.19 *Es sei V ein linearer Raum mit $\dim(V) = n$ bezüglich einer beliebigen Basis. Weiterhin seien $f : V \to V : x \mapsto A \cdot x$ eine lineare Abbildung sowie $\lambda \in \mathbb{R}$ ein Eigenwert von A. Dann heißt $\dim(E_\lambda(A))$ die* geometrische Vielfachheit *von λ.*

Die Frage ist nunmehr: *Wie bestimmen wir die Eigenwerte einer Matrix?* Dazu betrachten wir folgende Herleitung (in Form einer Folge von Äquivalenzen). Gegeben sei ein Matrix $A \in \mathbb{R}^{n \times n}$. Dann gilt:

$$\begin{aligned}
\lambda \text{ ist Eigenwert von } A &\iff \text{es gibt ein } v \in V \setminus \{0\} \text{ mit } A \cdot v = \lambda \cdot v \\
&\iff \text{es gibt ein } v \in V \setminus \{0\} \text{ mit } A \cdot v - \lambda \cdot v = 0 \\
&\iff \text{es gibt ein } v \in V \setminus \{0\} \text{ mit } (A - \lambda I) \cdot v = 0 \\
&\iff \det(A - \lambda I) = 0
\end{aligned}$$

Die letzte Äquivalenz folgt aus der Charakterisierung der Existenz einer inversen Matrix zu $A - \lambda I$. Könnten wir nämlich die Matrix $A - \lambda I$ invertieren, so wäre die einzige Lösung der Gleichung $(A - \lambda I) \cdot v = 0$ der Nullvektor $v = (A - \lambda I)^{-1} \cdot 0 = 0$. Diesen hatten wir aber gerade ausgeschlossen.

Definition 6.20 *Es sei $A \in \mathbb{R}^{n \times n}$ eine Matrix. Dann ist*

$$p_A(x) =_{\text{def}} \det(A - xI) = \det \begin{pmatrix} a_{11} - x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - x & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - x \end{pmatrix}$$

ein Polynom vom Grad n, dessen Nullstellen $\lambda \in \mathbb{R}$ die Eigenwerte von A sind. Das Polynom p_A heißt charakteristisches Polynom *von A. Die Vielfachheit einer Nullstelle heißt* algebraische Vielfachheit*.*

Beispiel: Wir greifen wieder auf den linearen Raum $V = \mathbb{R}^2$ sowie die Matrix

$$A =_{\text{def}} \frac{1}{2} \cdot \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

zurück. Dann gilt:

$$\begin{aligned}
\det(A - xI) &= \det \begin{pmatrix} 3/2 - x & 1/2 \\ 1/2 & 3/2 - x \end{pmatrix} \\
&= \left(\frac{3}{2} - x\right) \cdot \left(\frac{3}{2} - x\right) - \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{9}{4} - 3x + x^2 - \frac{1}{4} \\
&= x^2 - 3x + 2 \\
&= (x - 2)(x - 1)
\end{aligned}$$

Die Matrix hat also das charakteristische Polynom $p_A(x) = x^2 - 3x + 2$ (in expandierter Form) bzw. $p_A(x) = (x-2)(x-1)$ (in Nullstellenform). Damit sind die Eigenwerte der Matrix die reellen Zahlen 1 und 2 jeweils mit algebraischer Vielfachheit 1.

Im Folgenden schränken wir uns auf den einfachen Fall symmetrischer Matrizen ein. Dafür gibt es vor allem zwei vereinfachende Gründe:

1. Für die Eigenwerte einer beliebigen Matrix gilt, dass die geometrische Vielfachheiten höchstens so groß wie die algebraische Vielfachheiten (aber auch kleiner) sein können. Bei symmetrischen Matrizen gilt stets die Gleichheit.

2. Ein Polynom vom Grad n hat maximal n reelle Nullstellen (Vielfachheiten mitgezählt) und in der Menge $\mathbb{C}$ der komplexen Zahlen genau n Nullstellen. Im Allgemeinen kann ein charakteristisches Polynom somit auch komplexe Nullstellen und die Matrix damit komplexe Eigenwerte besitzen. Bei symmetrischen Matrizen treten dagegen keine komplexen Eigenwerte auf.

Theorem 6.21 (Hauptachsentransformation für symmetrische Matrizen) *Es seien V ein euklidischer Raum mit $\dim(V) = n$ und $A \in \mathbb{R}^{n\times n}$. Dann gilt:*

1. *Alle Eigenwerte $\lambda_1, \lambda_2, \ldots, \lambda_n$ sind reelle Zahlen.*

2. *Es gibt eine Orthonormalbasis $\mathcal{B} =_{\text{def}} \{b_1, b_2, \ldots, b_n\}$ von V, die aus Eigenvektoren $b_1, b_2, \ldots, b_n$ (zu den Eigenwerten $\lambda_1, \lambda_2, \ldots, \lambda_n$) besteht.*

3. *$A = Q \cdot D \cdot Q^T$ mit Q als Matrix mit den Spaltenvektoren $b_1, b_2, \ldots, b_n$ sowie*

$$D =_{\text{def}} \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Wir geben das Theorem ohne Beweis an und führen stattdessen zur Verdeutlichung die Hauptachsentransformation für ein Beispiel an.

Beispiel: Gegeben sei die Matrix

$$A =_{\text{def}} \begin{pmatrix} 3/2 & 1/2 & 0 \\ 1/2 & 3/2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Wir führen die Hauptachsentransformation in Schritten durch.

1. *Bestimmung des charakteristischen Polynoms sowie der Eigenwerte* λ_1, λ_2 *und* λ_3:

$$\begin{aligned} p_A(x) &= \det \begin{pmatrix} 3/2-x & 1/2 & 0 \\ 1/2 & 3/2-x & 0 \\ 0 & 0 & 2-x \end{pmatrix} \\ &= \left(\frac{3}{2}-x\right)\cdot\left(\frac{3}{2}-x\right)\cdot(2-x) - \frac{1}{2}\cdot\frac{1}{2}\cdot(2-x) \\ &= (2-x)\cdot\left[\left(\frac{3}{2}-x\right)^2 - \frac{1}{4}\right] \\ &= -(x-2)(x-2)(x-1) \end{aligned}$$

   Damit sind die Eigenwerte $\lambda_1 = 2, \lambda_2 = 2$ und $\lambda_3 = 1$. Der Eigenwert 2 tritt mit Vielfachheit 2 und der Eigenwert 1 mit Vielfachheit 1 auf.

2. *Bestimmung der Eigenräume* $E_1(A)$ *und* $E_2(A)$: $E_2(A)$ besteht aus allen Vektoren v mit

$$(A - 2\cdot I)\cdot v = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}\cdot v = 0,$$

   d.h. $E_2(A) = \text{span}\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$.

   $E_1(A)$ besteht aus allen Vektoren v mit

$$(A - 1\cdot I)\cdot v = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\cdot v = 0,$$

   d.h. $E_1(A) = \text{span}\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}\right\}$.

3. *Bestimmung der Orthonormalbasis*: Für die Eigenvektoren

$$v_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}, \quad v_2 = \begin{pmatrix}0\\0\\1\end{pmatrix}, \quad v_3 = \begin{pmatrix}-1\\1\\0\end{pmatrix}$$

   gilt bereits $\langle v_1, v_2\rangle = 0$, $\langle v_1, v_3\rangle = 0$ und $\langle v_2, v_3\rangle = 0$. Somit ist $\{v_1, v_2, v_3\}$ eine Orthogonalbasis. Um eine Orthonormalbasis $\{b_1, b_2, b_3\}$ zu bekommen, müssen die Vektoren so normiert werden, dass $\langle b_i, b_i\rangle = 1$ gilt. Dazu

wählen wir den Ansatz $b_i = a_i \cdot v_i$, wobei $a_i \in \mathbb{R}$ gilt. Aus den Eigenschaften des Skalarproduktes erhalten wir:

$$1 = \langle a_i \cdot v_i, a_i \cdot v_i \rangle = a_i \cdot \langle v_i, a_i \cdot v_i \rangle = a_i \cdot \langle a_i \cdot v_i, v_i \rangle = a_i^2 \cdot \langle v_i, v_i \rangle$$

Damit ergibt sich $a_i = \langle v_i, v_i \rangle^{-\frac{1}{2}}$ bzw. in unserem konrekten Falle

$$a_1 = \frac{1}{\sqrt{2}}, \quad a_2 = 1, \quad a_3 = \frac{1}{\sqrt{2}}$$

Die Orthnormalbasis $\{b_1, b_2, b_3\}$ besteht daher aus den Eigenvektoren

$$b_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad b_3 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

4. *Bestimmung der Matrizen D und Q*:

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad Q = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix}$$

Damit ist die Hauptachsentransformation abgeschlossen. Wir können nun zur Überprüfung der Korrektheit übergehen. Zunächst halten wir fest, dass Q eine Orthogonalmatrix ist, denn es gilt $Q^{-1} = Q^T$:

$$Q \cdot Q^T = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Ebenfalls leicht nachzurechnen ist die Gleichung $Q^T \cdot Q = I$. Wir überprüfen noch die Identität $A = Q \cdot D \cdot Q^T$ rechnerisch:

$$\begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 2 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} = \begin{pmatrix} 3/2 & 1/2 & 0 \\ 1/2 & 3/2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Literaturverzeichnis

[MM06] Bernd Kreußler und Gerhard Pfister. *Mathematik für Informatiker.*. Springer-Verlag, Berlin, 2009.

[MM06] Christoph Meinel und Martin Mundhenk. *Mathematische Grundlagen der Informatik. Mathematisches Denken und Beweisen. Eine Einführung.* 3., überarbeitete und erweiterte Auflage. B. G. Teubner Verlag, Wiesbaden, 2006.

[Ste07] Angelika Steger. *Diskrete Strukturen. Band 1: Kombinatorik-Graphentheorie-Algebra.* 2. Auflage. Springer-Verlag, Berlin, 2007.

[Wag03] Klaus W. Wagner. *Theoretische Informatik. Eine kompakte Einführung.* 2. überarbeitete Auflage. Springer-Verlag, Berlin, 2003.

[WHK04] Manfred Wolff, Peter Hauck und Wolfgang Küchlin. *Mathematik für Informatik und Bioinformatik.* Springer-Verlag, Berlin, 2004.