Anhang A
Logisches Schließen und Mengenlehre

A.1 Aussagenlogik

Aussagen sind (deutsche) Sätze, denen in eindeutiger Weise ein *Wahrheitswert wahr* (kurz: W) oder *falsch* (kurz: F) zugeordnet wird. Diese Zuordnung erfolgt entweder durch *axiomatische* Setzung, wie „0 ist eine natürliche Zahl" oder indem durch verknüpfende Operationen aus Aussagen neue gewonnen werden.

Zwei Aussagen sind *äquivalent*, wenn sie den gleichen Wahrheitswert haben. Dies erzeugt die neue Aussage

$$A \Leftrightarrow B ,$$

gesprochen: „*A ist äquivalent mit B*" oder „*A gilt genau dann, wenn B gilt*" oder „*A ist hinreichend und notwendig für B*". $A \Leftrightarrow B$ ist wahr, wenn A und B denselben Wahrscheinlichkeitswert haben, sonst falsch. Diese Definition kann schematisch durch eine *Wahrheitstafel*, d. h. eine Auflistung der Möglichkeiten an Wahrheitswerten vorgenommen werden, siehe Tabelle A.1. Aus vorhandenen Aussagen können also neue Aussagen gewonnen wer-

A	B	$A \Leftrightarrow B$
W	W	W
W	F	F
F	W	F
F	F	W

Tabelle A.1: Äquivalenz von Aussagen.

den, im Wesentlichen durch zwei Operationen. Das eine ist die *Negation* einer Aussage, d. h. einer Aussage A wird eine Aussage $\neg A$ (in Worten: „*nicht A*") zugeordnet. Diese Aussage $\neg A$ ist wahr, wenn A falsch ist und falsch, wenn A wahr ist. Tabelle A.2 enthält die zugehörige Wahrheitstafel. Im Sinne der Äquivalenz gilt das Prinzip der *doppelten Verneinung*, da sich mit der Wahrheitstafel sofort folgende Äquivalenz verifizieren lässt:

P. Knabner, W. Barth, *Lineare Algebra*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-32186-3, © Springer-Verlag Berlin Heidelberg 2013

$$A \Leftrightarrow \neg\neg A := \neg(\neg A)\ .$$

Mit := wird eine *Definition* bezeichnet, d. h. es wird eine neue Schreibweise, hier $\neg\neg A$ eingeführt, gerade durch den wohldefinierten rechtsstehenden Ausdruck. Weitere Verknüpfungen von jeweils zwei Aussagen, die wiederum eine Aussage ergeben, sind die *Konjunktion (das logische und)*, kurz geschrieben durch $\wedge$, und die *Disjunktion* (*das logische oder*), kurz geschrieben durch $\vee$.

Die neuen Aussagen $A \wedge B$ werden als „A *und* B" bzw. $A \vee B$ als „A *oder* B" bezeichnet. Dabei ist $A \wedge B$ nur dann wahr, wenn sowohl A als auch B wahr ist und in allen anderen Fällen falsch. Die Aussage $A \vee B$ ist nur dann falsch, wenn sowohl A als auch B falsch sind und in allen anderen Fällen wahr. Es handelt sich hier um ein „*nicht ausschließendes oder*" im Gegensatz zu dem „*ausschließenden oder*". Die Definition der neuen Aussagen ist durch folgende Wahrheitstafel zusammengefasst. Anhand von Wahrheitstafeln lassen

A	$\neg A$
W	F
F	W

Tabelle A.2: Negation einer Aussage.

A	B	$A \wedge B$	$A \vee B$
W	W	W	W
W	F	F	W
F	W	F	W
F	F	F	F

Tabelle A.3: Konjunktion und Disjunktion von Aussagen.

sich sofort folgende einfache Beziehungen verifizieren. Die Verknüpfungen $\wedge$ und $\vee$ sind *kommutativ*, d. h. es gilt

$$\begin{aligned} A \wedge B &\Leftrightarrow B \wedge A\ ,\\ A \vee B &\Leftrightarrow B \vee A\ . \end{aligned} \tag{A.1}$$

Außerdem sind diese Verknüpfungen *assoziativ*, d. h. für drei Aussagen A, B, C gilt

$$\begin{aligned} (A \wedge B) \wedge C &\Leftrightarrow A \wedge (B \wedge C)\ ,\\ (A \vee B) \vee C &\Leftrightarrow A \vee (B \vee C)\ . \end{aligned} \tag{A.2}$$

Bei gleichartigen Verknüpfungen ist demnach die Reihenfolge der Ausführung unerheblich, insofern kann man auch für die beiden äquivalenten Aussagen die Kurzschreibweisen

$$A \wedge B \wedge C := (A \wedge B) \wedge C\ ,$$
$$A \vee B \vee C := (A \vee B) \vee C$$

verwenden und ändert bis auf Äquivalenz dadurch nichts an der Aussage. Auf diese Weise ist es auch möglich, beliebig endlich viele Aussagen zu verknüpfen und z. B. von

$$A_1 \wedge A_2 \wedge A_3 \wedge \ldots \wedge A_n \tag{A.3}$$

zu reden.

Für den Zusammenhang zwischen Negation und Konjunktion bzw. Disjunktion verifiziert man analog folgende Äquivalenzen:

$$\begin{aligned} \neg(A \wedge B) &\Leftrightarrow (\neg A) \vee (\neg B)\ , \\ \neg(A \vee B) &\Leftrightarrow (\neg A) \wedge (\neg B)\ . \end{aligned} \tag{A.4}$$

Schließlich gelten als Zusammenhang zwischen $\wedge$ und $\vee$ die *Distributivgesetze*

$$\begin{aligned} A \wedge (B \vee C) &\Leftrightarrow (A \wedge B) \vee (A \wedge C)\ , \\ A \vee (B \wedge C) &\Leftrightarrow (A \vee B) \wedge (A \vee C)\ . \end{aligned} \tag{A.5}$$

Außerdem ergeben sich noch die *Identitätsgesetze* und die *Idempotenzgesetze*. Dabei seien Y eine Aussage, die immer den Wahrheitswert F hat, wie etwa eine Aussage vom Typ $x \neq x$ (z. B. $1 \neq 1$) und Z eine Aussage, die immer den Wahrheitswert W hat, wie eine Aussage vom Typ $x = x$ (z. B. $1 = 1$).

Allgemeiner nennt man eine Aussage Y, die immer ohne jede Voraussetzung falsch ist, einen *Widerspruch*. Ist A eine beliebige Aussage, so ist

$$Y := A \wedge (\neg A)$$

ein Widerspruch. Entsprechend heißt eine Aussage Z, die immer ohne jede Voraussetzung richtig ist, eine *Tautologie*. Ist A eine beliebige Aussage, so ist

$$Z := A \vee (\neg A)$$

eine Tautologie. Das ist die Basis für eine in Überlegungen oft angewandte *Fallunterscheidung*.

Für Konjunktion und Disjunktion gelten:

$$\begin{array}{ll} \textit{Identitätsgesetz} & \\ A \vee Y \Leftrightarrow A\ , & A \wedge Z \Leftrightarrow A\ , \\ A \vee Z \Leftrightarrow Z\ , & A \wedge Y \Leftrightarrow Y\ , \\ \textit{Idempotenzgesetz} & \\ A \vee A \Leftrightarrow A\ , & A \wedge A \Leftrightarrow A\ . \end{array} \tag{A.6}$$

Wir definieren eine weitere Verknüpfung von zwei Aussagen A und B zu einer neuen Aussage $A \Rightarrow B$, gesprochen als „*aus A folgt B*“. Dazu gehört die nachfolgende Wahrheitstafel:

Genaugenommen handelt es sich dabei um keine zusätzliche unabhängige Operation, da

A	B	$A \Rightarrow B$
W	W	W
W	F	F
F	W	W
F	F	W

Tabelle A.4: Implikation „Aus A folgt B“.

man sofort folgende Identität verifizieren kann:

$$(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$$

Ist die Aussage $A \Rightarrow B$ richtig, so sagt man auch, dass A *hinreichend* für B ist bzw. dass B *notwendig* für A ist. Die neue Aussage $A \Rightarrow B$ heißt auch *Implikation*. Wir vergegenwärtigen uns, dass im Allgemeinen aus $A \Rightarrow B$ nicht folgt, dass $B \Rightarrow A$. Das heißt die oft in der Alltagsdiskussion angewandte Überlegung „im Umkehrschluss gilt ...“ ist im Allgemeinen nicht richtig, bzw. bedarf einer gesonderten Untersuchung und Verifikation oder Falsifikation.

Die eben eingeführte Verknüpfung von zwei Aussagen hat folgende leicht nachprüfbare Eigenschaften. Sie ist reflexiv, d. h. es gilt

$$A \Rightarrow A \tag{A.7}$$

und sie ist nicht symmetrisch, d. h. aus $A \Rightarrow B$ folgt nicht $B \Rightarrow A$. Aber sie ist *transitiv*, d. h. für drei Aussagen A, B, C gilt

$$\big((A \Rightarrow B) \wedge (B \Rightarrow C)\big) \Rightarrow (A \Rightarrow C)\,. \tag{A.8}$$

Das ist die Basis der sogenannten *direkten Beweistechnik*.

Die Implikation ist aber nicht assoziativ, da für Aussagen A, B, C

$$(A \Rightarrow B) \Rightarrow C \quad \text{und} \quad A \Rightarrow (B \Rightarrow C)$$

nicht äquivalent sind. Trotzdem wird manchmal für Aussagen $A_1, \ldots, A_n$ die Notation

$$A_1 \Rightarrow A_2 \Rightarrow \ldots \Rightarrow A_n$$

benutzt. Im Gegensatz zu (A.3) wird darunter aber

$$(A_1 \Rightarrow A_2) \wedge (A_2 \Rightarrow A_3) \wedge \ldots \wedge (A_{n-1} \Rightarrow A_n) \tag{A.9}$$

(und damit nach (A.8) als Folgerung $A_1 \Rightarrow A_n$) verstanden. Analoges gilt für die Äquivalenz. Diese ist zwar assoziativ in dem Sinn, dass gilt

$$((A \Leftrightarrow B) \Leftrightarrow C) \;\Leftrightarrow\; (A \Leftrightarrow (B \Leftrightarrow C))\,,$$

für Aussagen $A_1, \ldots, A_n$ ist mit

$$A_1 \Leftrightarrow A_2 \Leftrightarrow \ldots \Leftrightarrow A_n$$

aber nicht eine (beliebige) Klammerung in dieser Aussage gemeint, sondern analog zu (A.9)

$$(A_1 \Leftrightarrow A_2) \wedge (A_2 \Leftrightarrow A_3) \wedge \ldots \wedge (A_{n-1} \Leftrightarrow A_n)\,.$$

Ä Zwar lässt sich bei einer gültigen Implikation die Pfeilrichtung im Allgemeinen nicht umkehren, aber es gilt folgende Äquivalenz, die als *Kontraposition* bezeichnet wird:

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)\,. \tag{A.10}$$

Dies ist die Basis für die *Beweistechnik durch Kontraposition.*

Eine Variante davon ist *der Beweis durch Widerspruch*, bei dem zusätzlich zur Voraussetzung A die Falschheit der Behauptung B, d. h. die Richtigkeit von $\neg B$ angenommen wird. Aus der dann gültigen Aussage $A \wedge \neg B$ muss dann eine immer falsche Aussage, d. h. eine Aussage vom Typ Y, geschlossen werden. Wenn ein solcher Schluss richtig ist, kann das nur bedeuten, dass $A \wedge \neg B$ falsch ist, damit also bei Annahme der Richtigkeit von A die Aussage $\neg B$ falsch und damit B richtig ist, d. h. insgesamt der Schluss $A \Rightarrow B$ richtig ist. Diese Argumentation ist im folgenden Schema noch einmal zusammengestellt:

$$(A \wedge \neg B \Rightarrow Y) \Rightarrow (A \Rightarrow B)\,. \tag{A.11}$$

Schließlich bedeutet die Äquivalenz zweier Aussagen A und B gerade die Gültigkeit beider Implikationen, d. h. von

$$A \Rightarrow B \quad \text{und} \quad B \Rightarrow A\,,$$

d. h. es gilt

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

Dies erklärt auch die Sprechweise „A ist hinreichend und notwendig für B“. Will man eine solche Äquivalenz beweisen, so kann man zum einen eine von A beginnende Schlussfolgerungskette aufbauen und sich bei jedem Schluss vergegenwärtigen, dass er tatsächlich auch umkehrbar ist, oder aber man zeigt tatsächlich in zwei Teilschritten zum einen die Aussage $A \Rightarrow B$ und zum anderen $B \Rightarrow A$. Dabei können dann durchaus verschiedene Argumente, d. h. verschiedene Zwischenschritte, in der Schlussfolgerungskette erzeugt werden.

A.2 Mengenlehre

Genau wie die Begriffe der (Aussagen-) Logik kann man die Bildungen der Mengenlehre, die, wie wir sehen werden, stark verwandt dazu sind, als die Sprache der Mathematik ansehen. Wir lassen im Folgenden den Begriff der Menge tatsächlich undefiniert und legen nur den Umgang mit Mengen genau fest. Eine Einführung in einen axiomatischen Aufbau der Mengenlehre findet man z. B. in DEISER 2002. Dazu gehört, für ein beliebiges Objekt x und eine beliebige Menge A entscheiden zu können, ob x zu dieser Menge gehört oder nicht. Die zugehörige Bezeichnung bzw. Sprechweise ist:

$x \in A$: „*x ist Element der Menge A*" bzw. kurz „*x ist in A*" oder aber

$x \notin A$: „*x ist nicht Element der Menge A*" bzw. kurz „*x ist nicht in A*".

Eine gegebene Menge A definiert eine sogenannte Aussageform $\mathcal{A}(x)$. Dabei verstehen wir unter einer *Aussageform in einer freien Variablen x* etwas, was durch Einsetzen eines konkreten Objektes für x zu einer Aussage wird. Dabei sollte immer klar sein, was eine sinnvolle Grundmenge X ist. Zum Beispiel ist

$$x > 0$$

eine Aussageform, z. B. auf der Menge der reellen Zahlen, $X = \mathbb{R}$. Die sich ergebende Aussage ist wahr für $x = 2$, sie ist falsch für $x = -\pi$. Genau die Objekte, für die eine gegebene Aussageform wahr ist, fassen wir zu einer neuen Menge zusammen (dass dies so geht, ist wesentlich für einen axiomatischen Aufbau der Mengenlehre). Das heißt also, die Aussageform $\mathcal{A}(x)$ erzeugt genau eine Menge A, die definiert ist als

$$A := \{x \in X : \mathcal{A}(x)\} .$$

In Worten: A besteht genau aus den Elementen $x \in X$, für die die Aussage $\mathcal{A}(x)$ gilt, d. h. wahr ist.

Andererseits erzeugt eine gegebene Menge A genau eine Aussageform $\mathcal{A}(x)$ dadurch, dass die Aussageform genau dann wahr ist, wenn $x \in A$,

$$\mathcal{A}(x) \quad \Leftrightarrow \quad x \in A .$$

In Worten: Die Aussage $\mathcal{A}(x)$ ist wahr genau dann, wenn $x \in A$.

Insbesondere können wir auch für eine Grundmenge X die folgende Menge definieren:

$$\emptyset_X := \{x \in X : x \neq x\} .$$

Diese Menge hat überhaupt keine Elemente. Man spricht auch von der *leeren Teilmenge von X*.

Mittels der Elementbeziehung können nun auch Beziehungen zwischen Mengen bzw. neue Mengen definiert werden.

Definition A.1

Seien X und Y Mengen.

1) Die Aussage

$$X \subset Y$$

bedeutet (ist also wahr genau dann, wenn):
Jedes Element von X ist auch Element von Y, d. h.

$$x \in X \Rightarrow x \in Y \; .$$

X wird als *Teilmenge* von Y bzw. Y als *Obermenge* von X bezeichnet und dann auch die gleichwertige Bezeichnung

$$Y \supset X$$

verwendet.
Man sagt auch, dass X *in* Y *enthalten ist*.

2) Zwei Mengen X und Y heißen *gleich*,

$$X = Y \; ,$$

wenn

$$(X \subset Y) \wedge (Y \subset X) \; ,$$

d. h. wenn sie genau die gleichen Elemente enthalten.

Gilt $X \subset Y$ und $X \neq Y$ (d. h. $X = Y$ ist falsch und damit gilt nicht $Y \subset X$), so heißt X auch eine *echte Teilmenge* von Y und es wird die Schreibweise

$$X \subsetneq Y$$

benutzt.

Es gibt damit eine Entsprechung zwischen

$$\begin{array}{cc} „\subset“ & „\Rightarrow“ \\ \text{für Mengen} & \text{für Aussagen} \\ „=“ & „\Leftrightarrow“ \\ \text{für Mengen} & \text{für Aussagen} \; . \end{array}$$

Die Aussagen (A.7) und (A.8) bzw. analoge Aussagen für „⇔“ gehen sofort in entsprechende Aussagen für Mengen über. So ist also die Teilmengenbeziehung reflexiv und transitiv, d. h. es gilt

$$X \subset X\,,$$
$$((X \subset Y) \wedge (Y \subset Z)) \quad \Rightarrow \quad (X \subset Z)\,.$$

Der Tatsache, dass aus einer falschen Aussage jede beliebige Aussage folgt, entspricht, dass die leere Menge jede Eigenschaft besitzt, d. h. dass gilt:

Bemerkung A.2 Sei X eine Menge. Sei $\mathcal{A}(x)$ eine Aussageform auf X, dann gilt

$$x \in \emptyset_X \Rightarrow \mathcal{A}(x)\,. \qquad \triangle$$

Damit ergibt sich auch, dass es genau eine leere Menge unabhängig von der Grundmenge gibt:

Bemerkung A.3 Seien X, Y Mengen, dann gilt

$$\emptyset_X = \emptyset_Y. \qquad \triangle$$

Ab sofort wird nur noch die Bezeichnung $\emptyset$ für die leere Menge verwendet. In Analogie zu Konjunktion und Disjunktion werden definiert:

Definition A.4

Seien A, B Teilmengen einer Menge X. Dann heißt

$$A \cap B := \{x \in X : (x \in A) \wedge (x \in B)\}\,,$$

gesprochen „*A geschnitten B*", der *Durchschnitt* von A und B. Ist $A \cap B = \emptyset$, d. h. haben sie keine gemeinsamen Elemente, so heißen A und B *disjunkt*. Die Menge

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

gesprochen „*A vereinigt B*", heißt die *Vereinigung* von A und B.

Aus den entsprechenden Aussagen (A.1), (A.2) (A.5) und (A.6) für Konjunktion und Disjunktion ergeben sich folgende Mengenbeziehungen:

Satz A.5

Seien A, B, C Teilmengen von X, dann gelten folgende Aussagen:

1) $A \cup B = B \cup A, \quad A \cap B = B \cap A.$ **Kommutativität**

2) $A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C.$ **Assoziativität**

3) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$ **Distributivität**

4) $A \cup A = A, \quad A \cap A = A.$ **Idempotenz**

5) $A \cup \emptyset = A, \quad A \cap X = A,$ **Identität**
$A \cup X = X, \quad A \cap \emptyset = \emptyset.$

6) $(A \subset B) \Leftrightarrow (A \cup B = B) \Leftrightarrow (A \cap B = A).$

Wegen der Assoziativität kann auch eindeutig vom Schnitt bzw. der Vereinigung von mehr als zwei Mengen geredet werden, d. h. etwa

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

bzw. für $A_i, i = 1, \ldots, n$, (d. h. $A_1, A_2, \ldots, A_n$) sei z. B.

$$\bigcup_{i=1}^{n} A_i := A_1 \cup A_2 \cup \ldots \cup A_n := \Big(\ldots (A_1 \cup A_2) \cup A_3 \ldots \Big) \cup A_{n-1} \Big) \cup A_n \ . \tag{A.12}$$

Schließlich vereinbaren wir:

Definition A.6

Seien A, B Teilmengen einer Menge X. Dann heißt

$$A \backslash B := \{x \in X : (x \in A) \wedge (x \notin B)\}$$

das (relative) *Komplement von B in A*. Ist die Grundmenge X aus dem Zusammenhang klar, so heißt

$$A^c := X \backslash A$$

das *Komplement* von A (in X).

Aus (A.4), (A.10) bzw. (A.11) folgen:

Satz A.7

Seien A, B Teilmengen einer Menge X. Dann gilt:

1) *Regeln von* DE MORGAN[1]*:*
$$(A \cap B)^c = A^c \cup B^c \ ,$$
$$(A \cup B)^c = A^c \cap B^c \ .$$

2) $A \subset B \Leftrightarrow B^c \subset A^c$.

3) $(A \cap B^c = \emptyset) \Rightarrow (A \subset B)$.

Sei X eine Menge, dann können alle ihre Teilmengen wiederum zu einer Menge zusammengefasst werden, der sogenannten *Potenzmenge* $\mathcal{P}(X)$ oder auch 2^X. Diese letztere

[1] Augustus DE MORGAN ∗27. Juni 1806 in Madurai †18. März 1871 in London

Schreibweise wird sich erst später erschließen. Die Potenzmenge einer Menge ist immer nicht leer, da die leere Menge ein Element von ihr ist. Generell sind folgende Aussagen offensichtlich:

$$\begin{aligned} \emptyset \in \mathcal{P}(X)\,, \quad & X \in \mathcal{P}(X)\,, \\ x \in X \quad & \Leftrightarrow \{x\} \in \mathcal{P}(X)\,, \\ Y \subset X \quad & \Leftrightarrow \; Y \in \mathcal{P}(X)\,. \end{aligned}$$

A.3 Prädikatenlogik

Mathematische Aussagen haben oft die Gestalt von Gleichungen. Definieren wir dazu eine Aussageform in zwei Variablen a und b:

$$\mathcal{A}(a,b) \quad := \quad „(a+b)^2 = a^2 + 2ab + b^2“\,.$$

Dann besagt die binomische Formel, dass die Aussage $\mathcal{A}(a,b)$ richtig ist für alle $a, b \in \mathbb{R}$. Dies ist also eine immer gültige Aussage, zu deren Formulierung wir ein neues logisches Element, den so genannten *All-Quantor* $\forall$, gesprochen „*für alle*“, einführen. Die Aussage lautet dann formal

$$\forall a \in \mathbb{R} \wedge b \in \mathbb{R} : \mathcal{A}(a,b)\,.$$

Die Teilmengenbeziehung

$$A \subset B$$

ist äquivalent zu

$$\forall x \in A : x \in B$$

und damit ist dies aber nicht äquivalent zu

$$A = B\,,$$

da es ja ein $x \in B$ geben kann, für das

$$x \notin A$$

gilt. Um das auszudrücken, ist ein weiterer Quantor nötig, der *Existenz-Quantor* $\exists$, gesprochen „*es gibt*“. Im obigen Fall kann also gelten

$$\forall x \in B : x \in A\,, \tag{A.13}$$

d. h. $B \subset A$ und damit $A = B$, oder (A.13) ist falsch, d. h.

$$\exists x \in B : x \notin A\,.$$

Allgemein ist also für eine Menge A und eine Aussageform $\mathcal{A}(x)$ die Aussage

$$\forall x \in A : \mathcal{A}(x)$$

richtig genau dann, wenn

$$\exists x \in A : \neg\mathcal{A}(x)$$

falsch ist, also

$$\forall x \in A : \mathcal{A}(x) \quad \Leftrightarrow \quad \neg(\exists x \in A : \neg\mathcal{A}(x))$$

und analog

$$\exists x \in A : \mathcal{A}(x) \quad \Leftrightarrow \quad \neg(\forall x \in A : \neg\mathcal{A}(x)) .$$

Es muss deutlich zwischen der *Existenz einer Lösung eines Problems*, d. h. eines Objektes mit einer gewissen Eigenschaft und der Eindeutigkeit einer Lösung eines Problems, d. h. eines Objektes dieser Eigenschaft, unterschieden werden. *Eindeutigkeit* für ein Element in einer Grundmenge X mit der Eigenschaft $\mathcal{A}(x)$ bedeutet dabei

$$\forall x, y \in X : \ \mathcal{A}(x) \wedge \mathcal{A}(y) \Rightarrow x = y .$$

Eindeutigkeit bedeutet also, dass, wenn zwei Objekte vorliegen, die die betreffende Eigenschaft haben, sie notwendigerweise gleich sein müssen. Es bedeutet nicht, dass bei Eindeutigkeit überhaupt ein solches Objekt vorliegt, denn die obige Implikation ist auch dann richtig, wenn die Aussage $\mathcal{A}(x) \wedge \mathcal{A}(y)$ immer falsch ist. Wenn Existenz *und* Eindeutigkeit vorliegt wird dies manchmal auch mit dem Quantor

$\exists!$, gesprochen: „*Es gibt genau ein*“ ,

geschrieben.

Im Allgemeinen wird man aus einer Aussage mit der Struktur „$\forall x \exists y$“ nicht auf die analoge Aussage „$\exists y \forall x$“ schließen können. Die oben festgestellte Verneinungsregel für Aussagen mit Quantoren lässt sich auch auf Aussagenformen mit mehreren Variablen übertragen, sodass etwa in zwei Variablen gilt: Für eine Aussageform $\mathcal{A}(x, y)$ über der Grundmenge $x \in X,\ y \in Y$ gilt:

$$\begin{aligned}
\neg\,(\forall x \in X \quad \exists y \in Y : \mathcal{A}(x, y)) \quad &\Leftrightarrow \quad \exists x \in X \quad \forall y \in Y : \neg\mathcal{A}(x, y) , \\
\neg\,(\exists x \in X \quad \forall y \in Y : \mathcal{A}(x, y)) \quad &\Leftrightarrow \quad \forall x \in X \quad \exists y \in Y : \neg\mathcal{A}(x, y) .
\end{aligned}$$

Dies lässt sich auch für Aussageformen in mehr als zwei Variablen „schematisch“ übertragen, indem jeweils die Quantoren auszutauschen sind und die letztendliche Aussage zu verneinen ist, um die Verneinung der Gesamtaussage zu erhalten. Mit Hilfe der Quantoren können nun auch Schnitte und Vereinigungen für beliebige Mengensysteme definiert werden.

Sei X eine Menge und I eine nicht leere Menge, die zum „Indizieren“, d. h. zum Bezeichnen von Teilmengen von X verwendet werden soll. Liegen z. B. n Teilmengen ($n \in \mathbb{N},\ n \neq 0$) von A vor, können diese mit

$$A_1, \ldots, A_n$$

bezeichnet werden, d. h. $I = \{1, 2, \ldots, n\}$. Um aber etwa die einelementigen Teilmengen $\{n\}$ von $\mathbb{N}$ für alle $n \in \mathbb{N}$ zu indizieren, ist schon $I = \mathbb{N}$ notwendig. Im Folgenden soll zugelassen werden, dass I noch „größer“ ist und man spricht allgemein von einer *Familie von Mengen* oder einem *Mengensystem*

$$\{A_\alpha : \alpha \in I\} .$$

Hierbei wird nur $A_\alpha \subset X$ gefordert, d. h. Mengen dürfen auch mehrfach auftreten oder leer sein.

Definition A.8

Sei $\{A_\alpha : \alpha \in I\}$ ein Mengensystem in X.
Dann heißt

$$\bigcap_{\alpha \in I} A_\alpha := \{x \in X : \forall \alpha \in I : x \in A_\alpha\}$$

der *Durchschnitt* bzw.

$$\bigcup_{\alpha \in I} A_\alpha := \{x \in X : \exists \alpha \in I : x \in A_\alpha\}$$

die *Vereinigung* des Mengensystems.

Für $I = \{1, \ldots, n\}$ stimmen diese Begriffe z. B. für die Vereinigung mit (A.12) überein. Analog zu Satz A.5, A.7 gelten Assoziativitäts- und Distributivitätsgesetze und Regeln von DE MORGAN.

A.4 Produkte von Mengen, Relationen und Abbildungen

Um Beziehungen zwischen Elementen auch von verschiedenen Mengen ausdrücken zu können, definieren wir für zwei Mengen X und Y und zwei Elemente $x \in X$ bzw. $y \in Y$ das *geordnete Paar* (x, y) bzw. 2-*Tupel*. Dies ist ein neues Objekt für das folgender Gleichheitsbegriff gelten soll:

$$(x, y) = (x', y') \quad :\Leftrightarrow \quad (x = x') \wedge (y = y') .$$

Dabei sind $x, x' \in X$, $y, y' \in Y$ beliebige Elemente. (x, y) darf nicht verwechselt werden mit der zweielementigen Menge $\{x, y\}$, für die ja $\{x, y\} = \{y, x\}$ gilt.[2] Die Menge aller geordneten Paare bildet die Menge des *(kartesischen) Produkts* $X \times Y$, d. h.

$$X \times Y := \{(x, y) : x \in X \wedge y \in Y\} .$$

Für $(x, y) \in X \times Y$ heißt x die *erste* bzw. y die *zweite Komponente* von (x, y). Gilt $X = Y$, wird auch die Bezeichnung

$$X^2 := X \times X \tag{A.14}$$

benutzt. Für drei Mengen X, Y, Z kann

$$X \times Y \times Z := (X \times Y) \times Z$$

definiert werden und man erhält einen Raum von *Tripeln* bzw. 3-*Tupeln*. Zur Wohldefinition muss die Identifizierbarkeit von

$$X \times (Y \times Z) \quad \text{und} \quad (X \times Y) \times Z$$

gezeigt werden (Übung). Dies kann für n Mengen $X_1, X_2, \ldots, X_n$ fortgesetzt werden zu

$$X_1 \times \ldots \times X_n := (X_1 \times \ldots \times X_{n-1}) \times X_n ,$$

dem Raum der *n-Tupeln* $(x_1, \ldots, x_n)$. Eine alternative Schreibweise ist

$$\prod_{i=1}^{n} X_i := X_1 \times \ldots \times X_n ,$$

bzw. wenn alle Räume $X_i = X$ sind, dann

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

Seien X und Y nicht leere Mengen, dann wird jede nicht leere Teilmenge R von $X \times Y$ als *(binäre oder zweistellige) Relation* auf X und Y bezeichnet. Statt $(x, y) \in R$ werden auch die Schreibweisen

$$xRy \quad \text{oder} \quad x \underset{R}{\sim} y$$

benutzt.

Definition A.9

Sei R eine Relation auf X und Y, d. h. $\emptyset \neq R \subset X \times Y$. Wenn gilt

[2] Eine formale Definition, die die geforderte Gleichheitsbeziehung erfüllt, ist $(x, y) := \{\{x\}, \{x, y\}\}$.

1) $\forall x \in X \; \exists\, y \in Y : (x, y) \in R$ (Existenz des Bildes),

2) $\forall x \in X \; \forall\, y_1, y_2 \in Y : (x, y_1) \in R \wedge (x, y_2) \in R \;\Rightarrow\; y_1 = y_2$ (Eindeutigkeit des Bildes),

dann heißt R eine *Abbildung von X nach Y*. Das zu $x \in X$ eindeutig existierende $y \in Y$, so dass $(x, y) \in R$, wird (etwa) als $f(x)$ bezeichnet. Man spricht dann von der Abbildung

$$f : X \to Y \quad \text{oder auch} \quad x \in X \mapsto f(x) \in Y\,.$$

Dabei heißt $f(x)$ der *Funktionswert* von f an der Stelle x oder das *Bild* von x unter f, x heißt das *Argument* von f. X heißt der *Definitionsbereich* von f, Y der *Wertebereich* von f. Die Teilmenge von Y

$$\mathrm{Bild}(f) := \{y \in Y : \exists x \in X : y = f(x)\}$$

der „getroffenen" Werte heißt das *Bild* von f.

Zwei Abbildungen f, g heißen *gleich*, wenn die zugehörigen Relationen die gleichen Mengen in den gleichen Grundmengen sind, geschrieben

$$f = g\,.$$

Wichtig ist die Unterscheidung zwischen einem festen Funktionswert $f(x) \in Y$ (oft eine Zahl) und der Abbildung f selbst (nach Definition eine Teilmenge von $X \times Y$). Will man das Argument bei der Abbildung andeuten, so geschieht dies durch

$$f = f(\,.\,)\,.$$

Ein Beispiel für eine Abbildung, die es immer gibt, ist die *Identität* (auf X), d. h.

$$\mathrm{id}_X : X \to X, \quad x \mapsto x\,.$$

Sei $Z \subset X$ nicht leer. Dann kann die *Einbettung* von Z nach X definiert werden als

$$i : Z \to X, \quad x \mapsto x\,.$$

Zu einer beliebigen Abbildung $f : X \to Y$ und $Z \subset X$ nicht leer, kann die *Einschränkung* von f auf Z definiert werden durch

$$f_{|Z} : Z \to Y, \quad x \mapsto f(x)\,.$$

Es gilt:

Satz A.10

Seien $f : X \to Y,\; g : U \to V$ Abbildungen

$$f = g \quad \Leftrightarrow \quad X = U \wedge Y = V \wedge (\forall x \in X : f(x) = g(x)) \ .$$

In Verallgemeinerung von Bild(f) definieren wir:

Definition A.11

Sei $f : X \to Y$ eine Abbildung und $A \subset X, B \subset Y$. Dann heißt

1) $f(A) := \{y \in Y : \exists x \in A : y = f(x)\} = \{f(x) : x \in A\}$ das *Bild von A unter f*,

2) $f^{-1}(B) := \{x \in X : f(x) \in B\}$ das *Urbild von B unter f*.

Es ist also

$$\text{Bild}(f) = f(X) \ .$$

Man beachte, dass $A = \emptyset$ oder $B = \emptyset$ zulässig sind und dann gilt

$$f(\emptyset) = \emptyset \ ,$$
$$f^{-1}(\emptyset) = \emptyset \ .$$

Hierbei *verträgt* sich f^{-1} mit $\subset, \bigcup, \bigcap$ im folgenden Sinn:

$$B_1 \subset B_2 \subset Y \quad \Rightarrow \quad f^{-1}(B_1) \subset f^{-1}(B_2) \tag{A.15}$$

$$B_\alpha \subset Y \ \forall \alpha \in I \quad \Rightarrow$$

$$f^{-1}\left(\bigcup_{\alpha \in I} B_\alpha\right) = \bigcup_{\alpha \in I} f^{-1}(B_\alpha) \tag{A.16}$$

$$f^{-1}\left(\bigcap_{\alpha \in I} B_\alpha\right) = \bigcap_{\alpha \in I} f^{-1}(B_\alpha) \ . \tag{A.17}$$

Für f statt f^{-1} gelten auch die Aussagen (A.15) und (A.16), nicht aber (A.17), statt dessen nur

$$A_\alpha \subset X \ \forall \alpha \in I \Rightarrow f\left(\bigcap_{\alpha \in I} A_\alpha\right) \subset \bigcap_{\alpha \in I} f(A_\alpha) \ .$$

Definition A.12

Seien $f : X \to Y, \ g : Y \to Z$ Abbildungen zwischen nicht leeren Mengen X, Y, Z. Die Abbildung $g \circ f : X \to Z$ wird definiert durch

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

$g \circ f$ („g nach f“, „g verknüpft mit f“ oder „g Kringel f“) heißt die *Komposition* oder *Hintereinanderausführung* von f und g.

Man beachte, dass gemäß der Definitionsgleichung die *zuerst* ausgeführte Abbildung *rechts* in der Bezeichnung steht.

Satz A.13

Seien $f : W \to X$, $g : X \to Y$, $h : Y \to Z$ Abbildungen zwischen nicht leeren Mengen W, X, Y, Z.
Dann sind die Kompositionen

$$(h \circ g) \circ f : W \to Z$$

und

$$h \circ (g \circ f) : W \to Z$$

wohldefiniert und es gilt

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

(Assoziativität der Komposition).

Dies überträgt sich auch auf Kompositionen aus mehr als drei Abbildungen. Wegen Satz A.13 kann kurz

$$h \circ g \circ f$$

bzw.

$$f_n \circ f_{n-1} \circ \ldots \circ f_1$$

geschrieben werden, ohne Missverständnisse befürchten zu müssen.

Jede Relation $R \subset X \times Y$ erzeugt eine *Umkehrrelation* $R^{-1} \subset Y \times X$ durch

$$(y, x) \in R^{-1} :\Leftrightarrow (x, y) \in R\,.$$

Definition A.14

Seien X, Y nicht leere Mengen, $f : X \to Y$ eine Abbildung.
f heißt *surjektiv*, wenn $\mathrm{Bild}(f) = Y$ gilt, d. h. $\forall y \in Y\ \exists\, x \in X : y = f(x)$.
f heißt *injektiv*, wenn gilt:

$$\forall x_1, x_2 \in X : f(x_1) = f(x_2) \quad \Rightarrow \quad x_1 = x_2 \,, \quad \text{d.h.}$$
$$\forall x_1, x_2 \in X : x_1 \neq x_2 \quad \Rightarrow \quad f(x_1) \neq f(x_2) \,.$$

f heißt *bijektiv*, wenn f injektiv und surjektiv ist.
f ist also bijektiv genau dann, wenn die Umkehrrelation R^{-1} eine Abbildung ist. Diese wird mit

$$f^{-1} : Y \to X$$

bezeichnet und heißt *Umkehrabbildung* von f.

Die Umkehrabbildung erfüllt (im Fall ihrer Existenz)

$$f^{-1} \circ f = \mathrm{id}_X \,, \qquad f \circ f^{-1} = \mathrm{id}_Y \,.$$

Genauer gilt:

Satz A.15

Sei $f : X \to Y$ eine Abbildung von X nach Y. Dann ist f bijektiv genau dann, wenn es eine Abbildung $g : Y \to X$ gibt mit:

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

In diesem Fall ist g eindeutig und $g = f^{-1}$.

Die Umkehrabbildung $f^{-1} : Y \to X$ einer bijektiven Abbildung $f : X \to Y$ ist also die durch

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

eindeutig festgelegte Abbildung. Aus der Bijektivität von f folgt auch die Bijektivität von f^{-1} und

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

Weiter gilt

Satz A.16

Seien $f : X \to Y$, $g : Y \to Z$ bijektive Abbildungen. Dann ist auch $g \circ f : X \to Z$ bijektiv und

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

Man beachte die Umkehrung der Reihenfolge in der Komposition, die i. Allg. nicht kommutativ ist.

Sind die natürlichen Zahlen $\mathbb{N}_0$ eingeführt (siehe Definition B.1), kann zwischen endlichen und unendlichen Mengen unterschieden werden.

Definition A.17

Sei M eine Menge. M heißt *endlich*, wenn einer der folgenden Fälle zutrifft:

a) $M = \emptyset$.

b) Es gibt ein $n \in \mathbb{N}$ und eine injektive Abbildung $f : M \to \{1, \ldots, n\}$.

Ist $M \neq \emptyset$ endlich, so gibt es eine eindeutig bestimmte Zahl $n \in \mathbb{N}$, so dass die Abbildung aus Definition A.17, b) bijektiv ist. Sie heißt die *Anzahl* (der Elemente) von M und wird mit

$$\#(M) := n$$

bezeichnet. Im Allgemeinen sind die Begriffe injektiv und surjektiv unabhängig voneinander. Es gilt aber:

Satz A.18

Seien X und Y endliche Mengen mit gleich vielen Elementen, $\#(X) = \#(Y)$. Dann sind für $f : X \to Y$ die folgenden Aussagen äquivalent:

(i) f ist injektiv.

(ii) f ist surjektiv.

(iii) f ist bijektiv.

Bei unendlichen Mengen lässt sich weiter unterscheiden:

Definition A.19

Sei M eine Menge, M heißt *abzählbar unendlich*, wenn es eine bijektive Abbildung

$$f : M \to \mathbb{N}$$

gibt.

Manchmal fasst man auch *endlich* und *abzählbar unendlich* zu *abzählbar* zusammen. Nicht abzählbare Mengen heißen auch *überabzählbar unendlich*. Beispiele für abzählbar unendliche Mengen sind (siehe Anhang B): $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{Q} \times \mathbb{Q}, \ldots$, nicht abzählbar sind $\mathbb{R}, \mathbb{C}, \mathcal{P}(\mathbb{N}), \ldots$

A.5 Äquivalenz- und Ordnungsrelationen

Im Folgenden werden Relationen auf einer nicht leeren Menge X, d. h. $R \subset X \times X$, betrachtet. Neben den als Relationen recht speziellen Abbildungen haben Relationen auch die Aufgabe Elemente einer Menge gemäß bestimmter Kriterien zueinander in Beziehung zu setzen. Dazu sind gewisse Eigenschaften der Relation notwendig.

Definition A.20

Sei R eine Relation auf einer Menge X, d. h. $R \subset X \times X$.
R heißt *reflexiv*, wenn für alle $x \in X$ gilt: xRx.
R heißt *transitiv*, wenn für alle $x, y, z \in X$ gilt: $(xRy) \wedge (yRz) \Rightarrow (xRz)$.
R heißt *symmetrisch*, wenn für alle $x, y \in X$ gilt: $xRy \Rightarrow yRx$.
R heißt *antisymmetrisch*, wenn für alle $x, y \in X$ gilt: $xRy \wedge (yRx) \Rightarrow x = y$.

Ist R reflexiv, transitiv und symmetrisch, dann heißt R *Äquivalenzrelation*.
Ist R reflexiv, transitiv und antisymmetrisch, dann heißt R *Ordnungsrelation*.

Eine Menge X heißt auch R *totalgeordnet* durch R, wenn R eine Ordnungsrelation ist und es gilt:

$$\text{Für alle } x, y \in X : xRy \vee yRx \,.$$

Zur Betonung der fehlenden Totalordnung heißt eine Ordnung manchmal auch *Halbordnung*. Auf jeder Menge X wird eine triviale Äquivalenz- (oder Ordnungs-) relation definiert durch

$$xRy :\Leftrightarrow x = y \,. \tag{A.18}$$

Ein Beispiel für eine Äquivalenzrelation, definiert auf $\mathbb{Z}$, ist

$$x \sim y := n \mid (x - y) \,, \tag{A.19}$$

dabei ist $n \in \mathbb{N}$ fest und $a \mid b$, gesprochen „a teilt b“, ist definiert durch $\exists c \in \mathbb{Z} : b = c \cdot a$.

Sei R eine allgemeine Äquivalenzrelation auf einer Menge X, die im Folgenden als „$\sim$“ geschrieben wird, d. h.

$$x \sim y := xRy \,.$$

Eine Äquivalenzrelation zerlegt X in die Mengen von Elementen, die miteinander in Relation stehen. Dazu sei zu $x \in X$ die *Äquivalenzklasse*

$$[x] := \{y \in X : y \sim x\}$$

definiert. Es gilt also immer:

$$x \in [x] .$$

Äquivalenzklassen zu verschiedenen Elementen haben folgende Beziehung:

Lemma A.21

Sei $\sim$ eine Äquivalenzrelation auf X.
Dann gilt für $x_1, x_2 \in X$:

$$[x_1] = [x_2] \Leftrightarrow x_1 \sim x_2 .$$

Beweis: „$\Rightarrow$“: $x_1 \in [x_1] = [x_2] \Rightarrow x_1 \sim x_2$
„$\Leftarrow$“: $y \in [x_1] \Rightarrow y \sim x_1 \sim x_2 \Rightarrow y \sim x_2 \Rightarrow y \in [x_2]$, also $[x_1] \subset [x_2]$.
Wegen $x_1 \sim x_2 \Rightarrow x_2 \sim x_1$ können x_1 und x_2 getauscht werden und das gleiche Argument ergibt $[x_2] \subset [x_1]$. □

Eine Äquivalenzklasse enthält also genau die miteinander in Relation stehenden Elemente. Jedes $y \in [x]$ heißt *Repräsentant* von $[x]$. Alle Repräsentanten haben die gleiche Äquivalenzklasse $[x]$. Sei

$$X_{/\sim} := \{[x] : x \in X\}$$

die Menge alle Äquivalenzklassen von X bezüglich $\sim$, d. h.

$$X_{/\sim} \subset \mathcal{P}(X) .$$

Es gilt wie angekündigt:

Satz A.22

Sei $\sim$ eine Äquivalenzrelation auf X. Dann ist $X_{/\sim}$ eine Zerlegung von X, d. h. jedes $x \in X$ liegt genau in einem $A \in X_{/\sim}$, nämlich $A = [x]$.

Beweis: Wegen $x \in [x]$ ist $X \subset \bigcup_{x \in X}[x]$, d. h.

$$X = \bigcup_{x \in X}[x] .$$

Seien $x_1, x_2 \in X$, dann gibt es die Möglichkeiten:

1) $[x_1] \cap [x_2] = \emptyset$,

2) $[x_1] \cap [x_2] \neq \emptyset \Rightarrow [x_1] = [x_2]$,

denn aus $y \in [x_1] \cap [x_2]$ folgt: $y \sim x_1, y \sim x_2$, d. h. $x_1 \sim x_2$ und damit die Behauptung nach Lemma A.21. □

Die Äquivalenzklassen für (A.18) sind gerade alle einelementigen Teilmengen, hier hat die Relation nichts zusammengefasst. Die Äquivalenzklassen zu (A.19) sind gerade

$$[0], [1], \ldots, [n-1] .$$

Durch

$$p = p_X : X \to X_{/\sim}, x \mapsto [x]$$

wird allgemein eine surjektive Abbildung definiert, die *Projektion* von X auf $X_{/\sim}$, die nur im Trivialfall von (A.18) auch injektiv ist. Sei $f : X \to Y$ eine Abbildung, die i. Allg. nicht injektiv ist. Um sie injektiv zu machen, müssen alle Urbilder eines $y \in \mathrm{Bild}(f)$ „zusammengefasst" werden. Dies geschieht durch folgende Äquivalenzrelation:

$$x_1 \sim x_2 :\Leftrightarrow f(x_1) = f(x_2) \quad \text{für} \quad x_1, x_2 \in X \tag{A.20}$$

und die zugehörige Äquivalenzklassenzerlegung. Für die Äquivalenzklassen gilt

$$[x] = f^{-1}(\{f(x)\}) \quad \text{für alle} \quad x \in X .$$

Die Abbildung f erzeugt eine injektive Abbildung $\tilde{f}$ von $X_{/\sim}$ nach Y,

$$\tilde{f}([x]) := f(x) . \tag{A.21}$$

Da $[x]$ viele Repräsentanten haben kann, muss die *Wohldefinition* von $\tilde{f}$ geprüft werden, d. h.

$$[x_1] = [x_2] \Rightarrow f(x_1) = f(x_2)$$

gezeigt werden. NachLemma A.21 folgt aber aus $[x_1] = [x_2]$:

$$x_1 \sim x_2 \Leftrightarrow f(x_1) = f(x_2) .$$

Die Abbilding $\tilde{f}$ ist injektiv, da

$$f(x_1) = \tilde{f}([x_1]) = \tilde{f}([x_2]) = f(x_2) \Rightarrow x_1 \sim x_2 \Rightarrow [x_1] = [x_2] .$$

Es gilt:

$$\mathrm{Bild}(\tilde{f}) = \mathrm{Bild}(f)$$

und

$$\tilde{f} \circ p_X = f .$$

Damit wurde gezeigt:

Theorem A.23: Homomorphiesatz

Seien X, Y nicht leere Menge, $f : X \to Y$ eine Abbildung. Für die durch (A.20) definierte Äquivalenzrelation ist die durch (A.21) definierte Abbildung $\tilde{f}$ wohldefiniert und injektiv und das Diagramm

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "p_X", "\text{surjektiv}"'] & & Y \\ & X_{/\sim} \arrow[ur, "\tilde{f}"', "\text{injektiv}"] & \end{tikzcd}$$

ist kommutativ. Insbesondere ist also

$$\tilde{f} : X_{/\sim} \to f(X)$$

eine bijektive Abbildung.

Die Sprechweise, dass ein *Diagramm kommutativ ist*, will sagen, dass alle möglichen Wege in Form von Kompositionen von Abbildungen mit gleichem Definitions- und Wertebereich gleich sind.

Beispiele für *Ordnungsrelationen* sind

$$„\leq\text{“ auf } X = \mathbb{R}\,, \tag{A.22}$$
$$„\subset\text{“ auf } \mathcal{P}(X) \text{ für eine Menge } X\,,$$
$$„\,|\,\text{“ auf } A \subset \mathbb{N}\,. \tag{A.23}$$

Von den obigen Beispielen ist nur (A.22) eine totale Ordnung.

Ist die Grundmenge endlich, kann eine Ordnungsrelation durch ein *Hasse-Diagramm* veranschaulicht werden, was folgendes Beispiel illustriert:

Sei $A = \{1, 2, 5, 10, 20, 30\}$ und die Teilerrelation nach (A.23) definiert. Das zugehörige Hasse-Diagramm ist in Abbildung A.1 dargestellt. Man erkennt, dass z. B. 2 und 5 nicht vergleichbar sind. Ordnungsrelationen werden auch im allgemeinen Fall mit dem bekannten Zeichen geschrieben, d. h.

$$x \leq y :\Leftrightarrow xRy$$

für $x, y \in X$, wobei R eine Ordnungsrelation auf X darstellt. Weitere Beziehungen sind

$$\begin{aligned} x \geq y \quad &:\Leftrightarrow \quad y \leq x\,, \\ x < y \quad &:\Leftrightarrow \quad x \leq y \text{ und } x \neq y\,. \end{aligned} \tag{A.24}$$

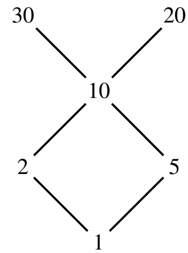

Abb. A.1: Hasse-Diagramm zur Ordnungsrelation (A.23).

Ist $\leq$ eine Ordnungsrelation auf X, so ist auch $\geq$ definiert durch (A.24) eine Ordnungsrelation auf X.

Definition A.24

Sei X eine Menge, $\leq$ eine Ordnungsrelation auf X. Sei $M \subset X$, $a \in X$.

1) a heißt *obere Schranke* von M, wenn $x \leq a$ für alle $x \in M$ gilt.
2) a heißt *Supremum* oder *kleinste obere Schranke* von M, wenn gilt:
   a) a ist obere Schranke von M.
   b) Ist a' eine obere Schranke von M, dann $a \leq a'$.
3) a heißt *maximales Element* von M, wenn gilt:
   a) $a \in M$.
   b) Für alle $x \in M$: $a \leq x \Rightarrow a = x$.
4) a heißt *Maximum* von M, wenn gilt:
   a) $a \in M$.
   b) Für alle $x \in M$: $x \leq a$.

Analog definiert man durch Übergang zu $\geq$ als Ordnungsrelation die Begriffe *untere Schranke*, *Infimum* oder *größte untere Schranke*, *minimales Element* und *Minimum*.

Satz A.25

Sei $\leq$ eine Ordnungsrelation auf einer Menge X, $M \subset X$, nicht leer, dann gilt:

1) Supremum und Maximum von M sind bei Existenz eindeutig und werden mit

$$\sup(M) \quad \text{bzw.} \quad \max(M)$$

bezeichnet.

2) Das Maximum von M ist ein maximales Element, aber i. Allg.nicht umgekehrt.

3) Existiert $\sup(M)$ und $\sup(M) \in M$, dann $\sup(M) = \max(M)$.

4) Ist $\leq$ total, so ist ein maximales Element das Maximum.

Beweis: Übung. □

Das Beispiel aus Abbildung A.1 hat also 30 und 20 als maximales und 1 als minimales Element, aber kein Maximum oder Minimum.

Bemerkungen A.26 Sei $(X, \leq)$ eine geordnete Menge.

1) Typische Schlussweisen mit sup und inf sind z. B.:

a) Seien $M_1, M_2 \subset X$ nicht leer, sei $x \leq y$ für alle $x \in M_1, y \in M_2$, dann gilt

$$\sup(M_1) \leq \inf(M_2) \,,$$

da jedes $y \in M_2$ eine obere Schranke von M_1 ist, also

$$\sup(M_1) \leq y \quad \text{für alle} \quad y \in M_2$$

und damit ist $\sup(M_1)$ *eine untere Schranke von M_2, also* $\sup(M_1) \leq \inf(M_2)$.

b) Seien $M_1 \subset X$ nicht leer und $M_2 := \{f(x) : x \in M_1\}$ für eine Abbildung $f : X \to X$, so dass $f(x) \geq x$, dann gilt

$$\sup(M_1) \leq \sup(M_2) \,.$$

Wegen $x \leq f(x)$ für alle $x \in M_1$ ist $\sup(M_2)$ *eine obere Schranke von M_1, d. h.* $\sup(M_1) \leq \sup(M_2)$.

2) Ist X mit einer Abstandsmessung $d(x, y)$ versehen, z. B. $X = (V, \|\,.\,\|)$ ein normierter $\mathbb{K}$-Vektorraum, $M \subset X$ nicht leer, so heißt $a \in M$ *lokales Maximum* von M, wenn ein $\varepsilon > 0$ existiert, so dass a Maximum von $M \cap \{x \in X : d(x, a \leq \varepsilon\}$ ist. Analog wird *lokales Minimum* und zusammenfassend *lokales Extremum* definiert. △

Anhang B
Zahlenmengen und algebraische Strukturen

B.1 Von den Peano-Axiomen zu den reellen Zahlen

Hier soll das nötige Grundwissen über Zahlenmengen zusammengefasst werden, das spätestens in einem Studiumsvorkurs oder am Anfang einer Vorlesung Analysis behandelt wird, mit einem Blick auf vorliegende Strukturen und mögliche Verallgemeinerungen. Die Darstellung orientiert sich an Amann und Escher 1998.
Der axiomatische Aufbau der reellen Zahlen beginnt mit dem der natürlichen Zahlen.

Definition B.1

Unter den natürlichen Zahlen $\mathbb{N}_0$ verstehen wir eine Menge $\mathbb{N}_0$ mit einem ausgezeichneten Element, genannt 0, und einer Abbildung $.^+ : \mathbb{N}_0 \to \mathbb{N} := \mathbb{N}_0 \setminus \{0\}$, der *Nachfolgerfunktion*, mit den Eigenschaften:

(P1) $n^+ = m^+ \Rightarrow n = m$ für $n, m \in \mathbb{N}_0$.

(P2) Enthält eine Teilmenge N von $\mathbb{N}_0$ das Element 0 und gilt

$$n \in N \Rightarrow n^+ \in N \,,$$

dann ist $N = \mathbb{N}_0$.

(P1) und (P2) heißen auch Peano[1]-Axiome.

Mengentheoretische Begriffe wie Teilmenge, Abbildung etc. werden in Abschnitt A präzisiert.

Bemerkung B.2

1) In manchen Texten wird auch $\mathbb{N}_0$ als $\mathbb{N}$ und $\mathbb{N}$ als $\mathbb{N}^*$ oder $\mathbb{N}^\times$ bezeichnet.

---

[1] Giuseppe Peano ∗27. August 1858 in Spinetta †20. April 1932 in Turin

2) Die natürlichen Zahlen entstehen somit durch sukzessive Anwendung der Nachfolgerfunktion

$$1 := 0^+, 2 := 1^+, 3 := 2^+, \dots$$

und mittels der Nachfolgerfunktion kann eine Addition + und eine Multiplikation · auf $\mathbb{N}_0$ eingeführt werden. △

Das aus $(P1), (P2)$ folgende *Induktionsprinzip* hat folgende Gestalt:

Eine Aussage der Form „Für alle $n \in \mathbb{N}_0 : \mathcal{A}(n)$“ ist richtig, wenn

(a) Induktionsanfang,
(b) Induktionsschluss

nachgewiesen werden können. Dabei sind

(a) *Induktionsanfang*: „Die Aussage gilt für $n = 0 : \mathcal{A}(0)$.“
(b) *Induktionsschluss*: Unter der *Induktionsvoraussetzung*:
„Für $n \in \mathbb{N}$ sei die Aussage $\mathcal{A}(n)$ richtig.“
folgt der *Induktionsschritt* (kurz: $n \to n + 1$):
„Die Aussage gilt für $n + 1 : \mathcal{A}(n + 1)$.“

Es kann auch bei einer anderen festen Zahl $n_0 \in \mathbb{N}_0$ (oder auch $n_0 \in \mathbb{Z}$) „angefangen“ werden.

Satz B.3: Induktionsprinzip

Sei $n_0 \in \mathbb{N}_0$ und für jedes $n \in \mathbb{N}_0, n \geq n_0$ sei $\mathcal{A}(n)$ eine Aussage.
Es gelten:

1) $\mathcal{A}(n_0)$ ist wahr.

2) Für jedes $n \in \mathbb{N}_0$, $n \geq n_0$ gilt: Wenn $\mathcal{A}(n)$ wahr ist, ist auch $\mathcal{A}(n + 1)$ wahr.

Dann ist $\mathcal{A}(n)$ wahr für jedes $n \in \mathbb{N}_0$, $n \geq n_0$.

Beweis: Sei $N := \{n \in \mathbb{N}_0 : \mathcal{A}(n + n_0) \text{ ist wahr}\}$, dann besagt (1), dass $0 \in N$ und (2), dass gilt: $n \in N \Rightarrow n + 1 \in N$, also nach (P2) von Definition B.1: $N = \mathbb{N}_0$. □

Satz B.4: $\mathbb{N}$ mit + und ·

Auf $\mathbb{N}_0$ können in eindeutiger Weise zwei Verknüpfungen, die *Addition* + und die *Multiplikation* · eingeführt werden, so dass gilt:

1) $m + n = n + m$ für alle $m, n \in \mathbb{N}_0$, d. h. + ist *kommutativ*,
$l + (m + n) = (l + m) + n$ für alle $l, m, n \in \mathbb{N}_0$, d. h. + ist *assoziativ*,
$n + 0 = n$ für alle $n \in \mathbb{N}_0$, d. h. 0 ist *neutrales Element* der Addition.

2) Die Multiplikation $\cdot$ ist kommutativ, assoziativ und hat 1 als neutrales Element.

3) Es gilt das *Distributivgesetz*

$$(l + m) \cdot n = l \cdot n + m \cdot n \text{ für alle } l, m, n \in \mathbb{N}_0 \, .$$

4)

a) $0 \cdot n = 0$ und

b) $n^+ = n + 1$ für alle $n \in \mathbb{N}_0$.

5) Seien $m, n \in \mathbb{N}_0$, dann gilt

$$m \cdot n = 0 \Rightarrow m = 0 \text{ oder } n = 0 \, .$$

Beweis: Zum Beispiel AMANN und ESCHER 1998, S. 34 ff. □

Die Eigenschaft 5) heißt auch die *Nullteilerfreiheit* von $\cdot$ auf $\mathbb{N}_0$.

Abstrahieren wir die Eigenschaften von $+$ und $\cdot$, so gilt:

$(\mathbb{N}_0, +, 0)$ und $(\mathbb{N}_0, \cdot, 1)$ sind jeweils
kommutative Halbgruppen mit *neutralem Element*,

wobei

Definition B.5

Sei M eine nicht leere Menge, $*$ eine *Verknüpfung* auf M, d. h. eine Abbildung $* : M \times M \to M$.
$(M, *)$ heißt *Halbgruppe*, wenn $*$ assoziativ ist, e aus M heißt (links- bzw. rechts-)*neutrales Element* bezüglich $*$, wenn

$$e * m = m \quad \text{bzw.} \quad m * e = m \quad \text{für alle } m \in M$$

gilt.

In beiden Fällen fehlen *inverse Elemente*, d. h. ein $m' \in M$, so dass

$$m * m' = m' * m = e \, .$$

Äquivalent damit ist, dass Gleichungen der Art

$$m + x = n \text{ bzw. } m \cdot x = n$$

in $\mathbb{N}_0$ nicht (immer) lösbar sind. Durch Einführung inverser Elemente bezüglich $+$ wird $\mathbb{N}_0$ zu $\mathbb{Z}$ erweitert.

Satz B.6: $\mathbb{Z}$ mit $+$ und $\cdot$

Es gibt eine (eindeutige „minimale") Erweiterung von $\mathbb{N}_0$ zu $\mathbb{Z}$, den *ganzen Zahlen*, so dass die Verknüpfungen $+$ und $\cdot$ erweitert werden können, unter bleibender Gültigkeit von Satz B.4, 1), 2), 3), 4) a), 5) (für $\mathbb{Z}$ statt $\mathbb{N}_0$).

6) Zu $n \in \mathbb{Z}$ existiert ein eindeutiges $\bar{n} \in \mathbb{Z}$, so dass

$$n + \bar{n} = 0 \, .$$

$\bar{n}$ wird mit $-n$ bezeichnet und heißt das inverse Element von n bezüglich n.

Bezeichnet man die „hinzukommenden" Inversen von $n \in \mathbb{N}$ mit $-n$, so ist also $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup -\mathbb{N} = \{0, 1, -1, 2, -2, \ldots\}$. Damit ist $(\mathbb{Z}, +)$ eine (kommutative) *Gruppe*, wobei:

Definition B.7

Sei M eine nicht leere Menge, $*$ eine Verknüpfung auf M. $(M, *)$ heißt *Gruppe*, wenn $*$ assoziativ ist, ein (links-)neutrales Element e besitzt und zu jedem $m \in M$ ein (links-)*inverses Element* $n \in M$ mit

$$n * m = e$$

existiert.

Bemerkungen B.8

1) Linksneutrale bzw. -inverse Elemente sind auch rechtsneutral bzw. -invers, kurz *neutral* bzw. *invers* (siehe Bemerkungen 3.5, 1), 2)).

2) Neutrale und inverse Elemente sind eindeutig nach Bemerkungen 3.5, daher ist für das inverse Element zu m eine auf m bezogene Bezeichnung wie $-m$ üblich (siehe Bemerkungen 3.5, 3), 4)). Insbesondere ist also

$$-e = e, \; -(-m) = m \text{ für } m \in M \, .$$

3) Wegen der Assoziativität kann auch

$$m * n * l := (m * n) * l = m * (n * l)$$

und damit jede Verknüpfung von endlich vielen Faktoren eindeutig definiert werden (siehe Abschnitt B.2). △

Da $\mathbb{Z}$ auch die assoziative und kommutative Verknüpfung $\cdot$ mit neutralem Element 1 besitzt, die mit + nach Satz B.6 verträglich sind, ist $(\mathbb{Z}, +, \cdot)$ ein kommutativer Ring mit Eins, wobei:

Definition B.9

Eine nicht leere Menge M mit Verknüpfungen +, der *Addition*, und $\cdot$, der *Multiplikation*, heißt *Ring*, wenn gelten:

(R1) $(M, +)$ ist eine kommutative Gruppe. Das neutrale Element wird *Nullelement* genannt.

(R2) $(M, \cdot)$ ist assoziativ.

(R3) Es gelten die *Distributivgesetze*

$$\begin{aligned} (m+n)\cdot l &= m\cdot l + n\cdot l, \\ l\cdot(m+n) &= l\cdot m + l\cdot n \quad \text{für } l, m, n \in M\ . \end{aligned} \tag{B.1}$$

Ein Ring heißt *kommutativ*, wenn $\cdot$ kommutativ ist (dann ist die zweite Bedingung in (B.1) redundant). Hat R ein neutrales Element bezüglich der Multiplikation, so heißt R ein *Ring mit Eins*.

Bemerkungen B.10

1) In (B.1) wird die *Punkt-vor-Strich-Regel* benutzt, in dem Sinn, dass etwa die erste Gleichung zu interpretieren ist als

$$(m+n)\cdot l = (m\cdot l) + (n\cdot l)\ .$$

2) Die eindeutigen neutralen Elemente werden oft mit 0 für die Addition und 1 für die Multiplikation bezeichnet und die Inversen bezüglich der Addition mit $-m$ bezeichnet.

3) $0\cdot m = m\cdot 0 = 0$ für $m \in M$,
$m\cdot(-n) = (-m)\cdot n = -(m\cdot n)$ für $m, n \in M$,
$(-1)\cdot m = -m$ für $m \in M$, falls M ein Einselement besitzt. △

In $(\mathbb{Z}, +, \cdot)$ fehlt noch die Möglichkeit, Gleichungen der Art

$$m\cdot x = n \text{ bzw. } x\cdot m = n \text{ für } m, n \in M, m \neq 0 \tag{B.2}$$

allgemein lösen zu können. Dies ist in der Erweiterung $\mathbb{Q}$ der Fall.

Satz B.11: $\mathbb{Q}$ mit + und ·

Es gibt eine (eindeutige „minimale“) Erweiterung von $(\mathbb{Z}, +, \cdot)$ zu $(\mathbb{Q}, +, \cdot)$, sodass dies ein Körper ist.

Definition B.12

Sei $(K, +, \cdot)$ ein Ring mit Einselement 1. 0 bezeichne das Nullelement. K heißt *Schiefkörper*, wenn gilt:

$$\text{Zu jedem } x \in K\backslash\{0\} \text{ existiert ein } \bar{x} \in K, \text{ so dass } x \cdot \bar{x} = \bar{x} \cdot x = 1.$$

Ist K ein Schiefkörper und die Multiplikation kommutativ, so heißt K *Körper*.

Wenn die Inverse bezüglich der Multiplikation zu $x \in \mathbb{Q}\backslash\{0\}$ mit $\frac{1}{x}$ bzw. $1/x$ bzw. x^{-1} bezeichnet wird und dann die *Brüche*

$$\frac{m}{n} := m \cdot n^{-1} \text{ für } m \in \mathbb{Z}, n \in \mathbb{Z}\backslash\{0\}$$

definiert werden, so ist $\mathbb{Q}$ gerade die Menge der Brüche. Die Elemente von $\mathbb{Q}$ heißen auch *rationale Zahlen*. Die Körpereigenschaften einer zugrundeliegenden „Zahl“ -menge ist ausreichend für einen großen Teil der Linearen Algebra, nicht aber für die Analysis:

Schon seit der Antike ist bekannt, dass die Gleichung

$$x^2 := x \cdot x = 2 \tag{B.3}$$

in $\mathbb{Q}$ nicht lösbar ist, d. h. die Länge der Diagonalen im Einheitsquadrat (siehe Satz von PYTHAGORAS) ist keine rationale Zahl. Dies zeigt, dass die rationalen Zahlen „Lücken“ haben. Es gibt verschiedene äquivalente Konstruktionsprinzipien, diese Lücken zu schließen (siehe *Analysis*).

Es lässt sich zum Beispiel zeigen, dass es eine Folge $r_n \in \mathbb{Q}$ gibt mit $r_n \geq r_0$, für ein $r_0 > 0$, für die gilt:

$$r_n \to 2 \text{ für } n \to \infty \text{ und } x^2 = r_n \text{ hat eine Lösung } x_n \in \mathbb{Q}\,.$$

Dabei bedeutet für eine Folge $(y_n)_n$ in $\mathbb{Q}$ und $y \in \mathbb{R}$, $y_n \to y$ für $n \to \infty$ die Konvergenz in der durch $d(x, y) := |x - y|$ auf $\mathbb{Q} \times \mathbb{Q}$ erzeugten Metrik. Damit ist $(x_n)_n$ eine CAUCHY-Folge in $\mathbb{Q}$ (siehe Definition C.13), die aber keinen Grenzwert in $\mathbb{Q}$ hat, denn für diesen müsste (B.3) gelten. Durch „Hinzunahme“ solcher Grenzwerte entsteht $\mathbb{R}$, die Menge der reellen Zahlen. Man setze dabei:

$$\mathbb{R}^+ := \{x \in \mathbb{R} : x \geq 0\}\,.$$

Für den Begriff der Ordnung(srelation) siehe Definition A.20.

Satz B.13: $\mathbb{R}$ mit $+, \cdot$ und $\leq$

Es gibt eine (eindeutige „minimale“) Körpererweiterung von $(\mathbb{Q}, +, \cdot)$ zu $(\mathbb{R}, +, \cdot)$ mit folgenden weiteren Eigenschaften:
Zu jedem $a \in \mathbb{R}^+$ und $n \in \mathbb{N}$ existiert genau ein

$$x \in \mathbb{R}^+, \text{ so dass: } x^n = a \tag{B.4}$$

(die *positive n-te Wurzel* aus a).
$\mathbb{R}$ ist (topologisch) *vollständig* bezüglich $d(x, y) := |x - y|$, d. h. jede CAUCHY-Folge in $\mathbb{R}$ konvergiert in $\mathbb{R}$.
Auf $\mathbb{R}$ kann eine Ordnung $\leq$ eingeführt werden, die total ist und verträglich mit $+$ und $\cdot$ in folgendem Sinn:

7) Für $x, y \in \mathbb{R}$ gelten

$$x \leq y \Leftrightarrow x + z \leq y + z \text{ für } z \in \mathbb{R} ,$$

8) Für $x, y \in \mathbb{R}$ gelten

$$x \leq y \Leftrightarrow x \cdot z \leq y \cdot z \text{ für } z \in \mathbb{R}, z > 0 .$$

9) Für $\mathbb{N}_0$ gilt in dieser Ordnung

   a) $0 \leq n$ für alle $n \in \mathbb{N}_0$.

   b) Für $m, n \in \mathbb{N}_0$ gelten:

$$\begin{aligned} m \leq n &\Leftrightarrow \text{ Es gibt ein } d \in \mathbb{N}_0 : m + d = n \\ m < n &\Leftrightarrow \text{ Es gibt ein } d \in \mathbb{N} : m + d = n . \end{aligned}$$

$\mathbb{R}$ ist *ordnungsvollständig*, d. h. jede nicht leere, nach oben beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum.

In der Dezimaldarstellung sind die rationalen Zahlen gerade die Zahlen, deren Darstellung ab einer Stelle *periodisch* wird, z. B.

$$x = \frac{1}{3} = 0,333\ldots =: 0,\overline{3} .$$

Dazu gehören auch die Zahlen mit abbrechender Darstellung, z. B.

$$x = 3\frac{1}{8} = 3,125000\ldots = 3,125\overline{0} =: 3,125 .$$

Die *irrationalen*, d. h. die reellen Zahlen, die nicht rational sind, sind gerade die mit nicht-periodischer Dezimaldarstellung.

Beispiele sind neben $\sqrt{2}$ (siehe (B.3)) die Lösungen der Gleichung des *goldenen Schnitts*

$$\frac{1}{a} = \frac{a}{a+1}, \quad a_1 = \frac{1+\sqrt{5}}{2},\ a_2 = \frac{1-\sqrt{5}}{2}\ ,$$

die *Kreiszahl* $\pi = 3,1415926\ldots$ oder die *Eulersche Zahl* $e = 2,71828\ldots$ (siehe *Analysis*). Wichtig ist, dass diese Zahlen nie exakt, sondern nur durch eine Approximation durch eine rationale Zahl angegeben werden können. Diese kann so genau wie gewünscht sein, wenn der entsprechende Aufwand dafür in Kauf genommen wird. So kann $\sqrt{2}$ durch das Iterationsverfahren (*Babylonisches Wurzelziehen*)

$$\begin{aligned} x^{(0)} &\in \mathbb{R} \text{ sei beliebig („nahe" bei } \sqrt{2}) \\ x^{(k+1)} &:= (x^{(k)})/2 + 1/x^{(k)} \end{aligned}$$

beliebig genau berechnet werden. Dies ist genau das NEWTON-Verfahren in $\mathbb{Q}$ (Algorithmus 5 in Abschnitt 7.4), angewandt auf (B.3).

B.2 Schreibweisen und Rechenregeln

Sei $(M, *)$ eine assoziative Verknüpfung, etwa $(\mathbb{R}, +)$, dann kann eindeutig auch die Verknüpfung von endlich vielen Elementen $a_1, \ldots, a_n \in M$ definiert werden, indem eine Klammerung ausgesucht wird, d. h.

$$\begin{aligned} \sum_{i=1}^{1} a_i &:= a_1\ , \\ \sum_{i=1}^{k+1} a_i &:= \sum_{i=1}^{k} a_i + a_{k+1} \qquad \text{für } k = 0, \ldots, n-1\ . \end{aligned}$$

Wird die Verknüpfung als Multiplikation geschrieben, ist statt

$$\sum_{i=1}^{n} a_i \quad \text{der Ausdruck} \quad \prod_{i=1}^{n} a_i$$

üblich. Dabei heißt i der *(Lauf-)Index*, $i = 1$ die *untere (Summations-)Grenze* und $i = n$ die *obere (Summationsgrenze)*. Für das Rechnen mit solchen Summen (Produkten) gilt: Die Bezeichnung des Laufindex ist ohne Bedeutung. Üblich sind Buchstaben wie i, j, k oder $\nu, \mu, \ldots$, also

$$\sum_{i=1}^{n} a_i = \sum_{k=1}^{n} a_k = \sum_{\mu=1}^{n} a_\mu\ .$$

Sind die Summationsgrenzen (hier 1 und n) klar, können sie auch weggelassen werden:

$$\sum_i a_i := \sum\nolimits_i a_i := \sum_{i=1}^{n} a_i \ .$$

Der Summationsbereich kann irgendeine Teilemenge I von $\mathbb{Z}$ (üblicherweise monoton geordnet) sein.

Beispiele B.14

$$\sum_{i=-m}^{n} a_i = a_{-m} + a_{-m+1} + \ldots + a_0 + \ldots + a_n \ .$$

Soll über eine Teilmenge von Indizes summiert werden, die nicht (oder nur umständlich) durch untere und obere Grenzen ausgedrückt werden können, so kann die Einschränkung zusätzlich unter dem Summenzeichen aufgenommen werden, z. B.

$$\sum_{\substack{i=1\\ i\neq k}}^{n} a_i = a_1 + \ldots a_{k-1} + a_{k+1} + \ldots + a_n = \sum_{i=1}^{k-1} a_i + \sum_{i=k+1}^{n} a_i \quad (\text{für ein } k \in \{1, \ldots, n\}) \ .$$

◦

Wird allgemein die Summe

$$\sum_{i\in I} a_i$$

betrachtet, kann durch Transformation von I eine äquivalente Darstellung gewonnen werden: Dazu sei $\varphi : I \to \widetilde{I}$ eine bijektive monotone Abbildung (d. h. $i < j \Rightarrow \varphi(i) < \varphi(j)$), dann gilt

$$\sum_{i\in I} a_i = \sum_{j\in\widetilde{I}} a_{\varphi^{-1}(j)} \ .$$

Beispiel B.15

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

◦

Ist $*$ auch kommutativ, darf auch die Reihenfolge vertauscht werden, d. h. φ darf eine allgemeine Bijektion sein, z. B.

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_{n-i+1} \ .$$

Für eine leere Indexmenge $I = \emptyset$ wird gesetzt

$$\sum_{i \in I} a_i := 0 \, ,$$

also z. B. $\sum_{i=n+1}^{n} a_i = 0$.

Ist $(R, +, \cdot)$ ein Ring, so kann das Distributivgesetz verallgemeinert werden zu

$$\left(\sum_{i=1}^{n} a_i \right) \cdot b = \sum_{i=1}^{n} a_i b \quad \text{bzw.} \quad a \cdot \left(\sum_{i=1}^{n} b_i \right) = \sum_{i=1}^{n} a b_i$$

und damit zu

$$\left(\sum_{i=1}^{n} a_i \right) \cdot \left(\sum_{j=1}^{m} b_j \right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_i \right) b_j = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_i b_j \right) \tag{B.5}$$

bzw.

$$\left(\sum_{i=1}^{n} a_i \right) \cdot \left(\sum_{j=1}^{m} b_j \right) = \sum_{i=1}^{n} a_i \cdot \left(\sum_{j=1}^{m} b_j \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_i b_j \right) \, .$$

Aufgrund dieser Identitäten wird die Notation oft verkürzt zu

$$a \sum_{i=1}^{n} b_i \quad \text{statt} \quad a \cdot \left(\sum_{i=1}^{n} b_i \right) \quad \text{oder} \quad \sum_{i=1}^{n} a_i \sum_{j=1}^{m} b_j \quad \text{statt} \quad \left(\sum_{i=1}^{n} a_i \right) \cdot \left(\sum_{j=1}^{m} b_j \right) \, ,$$

d. h. die Klammerung wird in das Summenzeichen inkorporiert.
Für die identischen Ausdrücke von (B.5) wird auch kurz

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j$$

geschrieben. Wegen der Kommutativität von + kann auch irgendeine andere Anordnung der Summanden gewählt werden und dann bei $n = m$ kurz

$$\sum_{i,j=1}^{n} a_i b_j$$

geschrieben werden. Für beliebige, endliche Indexmengen I und J gilt das allgemeine Distributivgesetz

$$\sum_{i \in I} a_i \sum_{j \in J} b_j = \sum_{\substack{i \in I \\ j \in J}} a_i b_j \, .$$

Liegt auch Kommutativität von + vor, kann folgendermaßen umgeformt werden:

$$\sum_{i=1}^{n} a_i \sum_{j=1}^{m} c_{i,j} b_j = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_i c_{i,j} b_j \right) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_i c_{i,j} b_j \right) = \sum_{j=1}^{m} b_j \sum_{i=1}^{n} c_{i,j} a_i. \tag{B.6}$$

Suggestiv kann man die Umformung mit: „Hineinziehen-Umordnen-Herausziehen" beschreiben. In Matrixschreibweise entspricht dies für $\boldsymbol{a} = (a_i)_i$, $\boldsymbol{b} = (b_j)_j$, $C = (c_{i,j})_{i,j}$:

$$\boldsymbol{a}^t C \boldsymbol{b} = \boldsymbol{b}^t C^t \boldsymbol{a}.$$

Auch in einem nichtkommutativen Ring mit Eins $(\mathrm{R}, +, \cdot, 0, 1)$ gilt:

Seien $a, b \in \mathrm{R}$, so dass $ab = ba$, dann gilt für $n \in \mathbb{N}$ die *binomische Formel*

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

wobei

$$a^0 := 1, \quad a^{n+1} := a^n \cdot a \quad \text{für } n \in \mathbb{N}_0 \tag{B.7}$$

mit den *Binomialkoeffizienten*

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}, \quad n, k \in \mathbb{N}_0,\ k \leq n$$
$$\text{und} \quad 0! := 1, \quad (n+1)! := n!(n+1) \quad \text{für } n \in \mathbb{N}_0.$$

Analog: $(a+b)(a-b) = a^2 - b^2$. Die Variante

$$(a-b) \sum_{k=0}^{n-1} a^k b^{n-1-k} = a^n - b^n$$

braucht keine Vertauschbarkeit von a und b.

B.3 (Formale) Polynome

Auf einen beliebigen Körper K lässt sich als spezielle Teilmenge von $\mathrm{Abb}(K, K)$, den Abbildungen von K nach K, die Menge der *Polynome* einführen.

Definition B.16

Sei $(K, +, \cdot)$ ein Körper. Ein *Polynom* über K (in der Variablen x) ist Abbildung von K nach K und es gibt $a_i \in K, i = 0, \ldots, n$, die Koeffizienten für ein $n \in \mathbb{N}_0$, so dass

$$f(x) = \sum_{i=0}^{n} a_i x^i \text{ für } x \in K \tag{B.8}$$

mit x^i nach (B.7). Die Menge der Polynome wird mit $K[x]$ bezeichnet. Gilt (B.8) und $a_n \neq 0$, so heißt n der *Grad* von f, $n = \text{grad}(f)$ und f *Polynom vom Grad* n. Ist $a_n = 1$, so heißt f *normiert*. Für $f = 0$ wird $\text{grad}(f) = -1$ gesetzt.

$$K_n[x] := \{f \in K[x],\ \text{grad}(f) \leq n\} .$$

Polynome lassen sich nicht nur (punktweise) addieren, wie allgemein in $\text{Abb}(K, K)$, sondern auch multiplizieren.

Satz B.17: Ring und K-Vektorraum $K[x]$

Sei K ein Körper. Auf $K[x]$ wird eine Addition $+$ und eine Multiplikation $\cdot$ eingeführt durch:

Seien $f = \sum_{i=0}^{n} a_i x^i$, $g = \sum_{i=0}^{n} b_i x^i \in K[x]$, dann:

$$\begin{aligned} (f+g)(x) &:= \sum_{i=0}^{n} c_i x^i, c_k := a_k + b_k \text{ für } k = 0, \ldots, n , \\ (f \cdot g)(x) &:= \sum_{i=0}^{2n} c_i x^i, c_k := \sum_{i=0}^{k} a_i b_{k-i} \text{ für } k = 0, \ldots, 2n, \end{aligned} \tag{B.9}$$

wobei $a_i := b_i := 0$ für $i > n$

$$\text{und } \text{grad}(f \cdot g) = \text{grad}(f) \cdot \text{grad}(g), \text{ falls } f \neq 0 \text{ und } g \neq 0 . \tag{B.10}$$

$K[x]$ wird dadurch zu einem kommutativen Ring mit Einselement $f = 1$. Mit der eingeführten Addition und der Skalarmultiplikation

$$(\lambda \cdot f)(x) := \sum_{i=0}^{n} c_i x^i,\ c_i := \lambda a_i \text{ für } \lambda \in K \tag{B.11}$$

wird $K[x]$ zu einem K-Vektorraum und $K_n[x]$ zu einem n-dimensionalen Teilraum mit der *Monombasis* $\varphi_i(x) = x^i$, $i = 0, \ldots, n$.

Beweis: Durch direktes Nachrechnen der Eigenschaften von einem kommutativen Ring mit Eins (siehe Definition B.9). □

Bemerkungen B.18

1) Da Skalarmultiplikation und Multiplikation auf $K[x]$ verträglich sind, ist $(K[x], +, \lambda, \cdot)$ eine K-Algebra nach Definition 3.17.

2) Genauer ist ein Polynom nicht als Abbildung von $K \to K$, sondern als Tupel seiner Koeffizienten definiert, d. h. als

$$\{(a_0, a_1, \ldots) : a_k \neq 0 \text{ für nur endlich viele } k \in \mathbb{N}_0\} \tag{B.12}$$

mit den Verknüpfungen + und · nach (B.9). Für einen endlichen Körper K mit r Elementen, für den also $a^r = a$ für alle $a \in K$ gilt, ist also das Polynom $f(x) := x^r - x$ (als Koeffiziententupel) nicht 0, wohl aber f als Abbildung von K nach K. Damit ist also $f = 0$ als das Tupel $(0, 0, \ldots)$ zu verstehen. △

Wenn ein Polynom nicht als Abbildung, sondern als Element des in (B.12) definierten Raum verstanden wird, spricht man von *formalen Polynomen* und bezeichnet den so definierten Raum mit $K[X]$. Auf diesem Tupelraum wird die Ringstruktur nach (B.9) bzw. die K-Vektorraumstruktur nach (B.9), (B.11) definiert. Statt der Tupelschreibweise

$$(a_0, \ldots, a_n, 0, \ldots)$$

benutzt man

$$\sum_{i=0}^{n} a_i X^i \,, \tag{B.13}$$

was also keine Abbildung und auch nicht den Wert einer Abbildung darstellt. Insbesondere gibt es also eine K-lineare Einbettung $K \to K[X]$ durch $a \mapsto (a, 0, \ldots)$. Setzt man

$$X := \{0, 1, 0, \ldots\}$$

woraus

$$X^i := \{0, \ldots, 0, 1, 0, \ldots\}$$

mit $a_k = 1 \Leftrightarrow k = i$ folgt, entspricht (B.13) genau der angegebenen Linearkombination. Die Abbildung

$$\Phi : K[X] \to K[x] \qquad \sum_{i=0}^{n} a_i X^i \mapsto f \text{ mit } f(x) := \sum_{i=0}^{n} a_i x^i \tag{B.14}$$

ist zwar surjektiv, aber i. Allg. nicht injektiv wie das obige Beispiel zeigt.

In $K[x]$ kann wie in $\mathbb{Z}$ mit Rest dividiert werden. Die folgenden Aussagen gelten alle in $K[X]$. Zur Erhöhung der Anschaulichkeit erfolgt die Formulierung in $K[x]$.

Satz B.19: Teilen mit Rest in $K[x]$

Sei K ein Körper, $f, g \in K[x]$ und $g \neq 0$. Dann gibt es eindeutig bestimmte $h, r \in K[x]$, so dass $f = hg + r$ und $\operatorname{grad}(r) < \operatorname{grad}(g)$. Ist $r = 0$, also $f = h \cdot g$, dann heißt *f durch g teilbar, in Zeichen $f \mid g$.*

Beweis: Seien $f = \sum_{i=0}^{m} a_i x^i$, $g = \sum_{i=0}^{n} b_i x^i$, wobei $m := \text{grad}(f)$, $n := \text{grad}(g)$, d. h. $a_m \neq 0$, $b_n \neq 0$.
Existenz: Falls $m < n$ ist, kann $h = 0$ und $r = f$ gesetzt werden. Für $m \geq n$ wird die Existenz mit vollständiger Induktion über m gezeigt. Für $m = 0$ und $n = 0$ kann $h = a_0/b_0$ und $r = 0$ gewählt werden. Es gelte die Behauptung für m. Sei

$$\tilde{f} := f - \frac{a_m}{b_n} x^{m-n} g \in K_{m-1}[x] \ .$$

Nach Indutktionsvoraussetzung bzw. Vorbemerkung gibt es also $\tilde{h}, r \in K[x]$ mit $\text{grad}(r) < n$, so dass

$$\tilde{f} = \tilde{h}g + r$$

und damit

$$f = \left(\frac{a_m}{b_n} x^{m-n} + \tilde{h}\right) g + r \ .$$

Eindeutigkeit: Ist $f = h_1 g + r_1 = h_2 g + r_2$ mit $\text{grad}(r_i) < n$, dann auch

$$r_1 - r_2 = (h_2 - h_1) g \ .$$

Ist $h_1 \neq h_2$, dann folgt

$$n \leq \text{grad}(g) + \text{grad}(h_2 - h_1) = \text{grad}(r_1 - r_2) < n \ ,$$

also ein Widerspruch. □

Definition B.20

Sei K ein Körper. K heißt *algebraisch abgeschlossen*, wenn für jedes $f \in K[x], \text{grad}(f) \geq 1$ gilt: f hat mindestens eine *Nullstelle*, d. h. ein $c \in K$ so dass $f(c) = 0$.

Daraus folgt:

Satz B.21: Nullstellen von Polynomen

Sei K ein Körper, $f \in K[x], f \neq 0, c \in K$

1) c ist eine Nullstelle von $f \Leftrightarrow f = (x - c)h$ für ein $h \in K[x]$.

2) Ist c eine Nullstelle, von f, dann gibt es ein eindeutiges $m \in \mathbb{N}$, so dass

$$f = (x - c)^m h, \text{ wobei} \tag{B.15}$$

$h \in K[x]$ und $h(c) \neq 0$. m heißt die *Vielfachheit* der Nullstelle c.

3) f hat höchstens endlich viele verschiedene Nullstellen, die Summe ihrer Vielfachheiten ist höchstens $\operatorname{grad}(f)$.

4) Sei K algebraisch abgeschlossen, dann zerfällt f in seine Linearfaktoren:

$$f = a_n \prod_{i=1}^{r} (x - c_i)^{m_i}$$

mit den Nullstellen c_i mit Vielfachheit m_i.

Beweis:

Zu 1): Nach Satz B.19 ist

$$f = (x - c)h + r, \; h, r \in K[x]$$

und $\operatorname{grad}(r) < 1$, also $r = a_0$ und wegen $f(c) = 0$ ist $a_0 = 0$.
Zu 2): Übung.
Zu 3): In (B.15) ist $\operatorname{grad}(h) = \operatorname{grad}(f) - m$ und $b \neq c$ ist Nullstelle von f, genau dann wenn b Nullstelle von h ist. Wiederholte Anwendung von (B.15) führt also zu

$$f = \prod_{i=1}^{r} (x - c_i)^{m_i} g$$

und g hat keine Nullstelle. Dann gilt

$$r \leq \sum_{i=1}^{r} m_i \leq \operatorname{grad}(f) \; .$$

Zu 4): direkt aus 3). □

Bemerkungen B.22

1) In einem unendlichen Körper kann also nur das Polynom $f = 0$ auf einer unendlichen Teilmenge von K verschwinden.

2) Satz B.19 und B.21 gelten auch für formale Polynome p, wenn dann unter einer Nullstelle eine Nullstelle von $\Phi(p) \in K[x]$ verstanden wird.

3) Wegen 1) ist die Abbildung Φ nach (B.14) bijektiv, genau dann wenn K unendlich ist. Für $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ ist also keine Unterscheidung zwischen $K[x]$ und $K[X]$ nötig, für endliche Körper schon. △

Allgemein hat $K[x]$ viele Eigenschaften mit $\mathbb{Z}$ gemeinsam. Insbesondere ist auch $K[X]$ ein *Hauptidealring*, genauer:

Satz B.23

Sei $I \subset K[x], I \neq \emptyset$ ein *Ideal*, d. h. $(I, +)$ sei eine Untergruppe ist und es gelte

$$I \cdot K[x] \subset I, \quad K[x] \cdot I \subset I \ .$$

Dann hat I die Gestalt

$$I =< g >:= K[x] \cdot g = \{hg : h \in K[x]\} \ .$$

Wird g normiert gewählt, so ist es eindeutig.

Beweis: Sei $g \in I$ mit minimalem nicht negativen Grad. Da I Ideal ist, kann g o. B. d. A. normiert gewählt werden und es gilt: $< g >\subset I$.
Sei $f \in I$, dann kann nach Satz B.19 f durch g geteilt werden, d. h.

$$f = h \cdot g + r \text{ mit } \operatorname{grad}(r) < \operatorname{grad}(g) \ .$$

Wegen $h \cdot g \in I$ und damit $r = f - h \cdot g \in I$ muss $r = 0$ gelten, also $I \subset< g >$, was den Gleichheitsbeweis abschließt. Sei $g' \in I$ auch normiert, so dass gelte

$$I =< g' >$$

dann folgt $g' = pg$ und $g = p'g'$ und $\operatorname{grad}(p) = \operatorname{grad}(p') = 0$, d. h. $p, p' \in K, p, p' \neq 0$, mit der Normierung, also $p = p' = 1$. □

Satz B.24: ggT in $K[x]$

Es seien $f, g \in K[x]$, jeweils nicht das Nullpolynom. Dann gibt es eindeutig bestimmtes normiertes $d \in K[x]$, so dass gilt:

(i) $d \mid f$ und $d \mid g$

(ii) Ist $d' \in K[x]$ und erfüllt:

$$d' \mid f \text{ und } d' \mid g \tag{B.16}$$

dann gilt auch $d' \mid d$.

Definition B.25

Es seien $f, g \in K[x]$. Das eindeutige $d \in K[x]$ nach Satz B.24 heißt *größter gemeinsamer Teiler*, in Zeichen

$$d = \mathrm{ggT}(f, g)\ .$$

Ist $d = 1$, heißen f und g *teilerfremd.*

Beweis (von Satz B.24): Sei $I := \{mf + ng : m, n \in K[x]\}$, dann ist I ein Ideal, zu dem nach Satz B.23 eindeutig ein normiertes $d \in K[x]$ existiert, so dass

$$I =< d >$$

und damit gilt: $d \mid f$ und $d \mid g$, da $f, g \in I$. Sei nun $d' \in K[x]$ mit der Eigenschaft (B.16) und

$$J :=< d' >\ .$$

J ist ein Ideal, das von d' *erzeugte Ideal* und wegen $f, g \in J$ folgt $I \subset J$.
Insbesondere gilt also

$$d \in J \text{ und damit } d' \mid d\ .$$

Zur Eindeutigkeit beachte man, dass für einen weiteren größten gemeinsamen Teiler d' gilt $d' \mid d$ und $d \mid d'$, was bei Normierung $d = d'$ bedeutet. □

Korollar B.26

Ein $p_1, p_2 \in K[x]$, $p_1, p_2 \neq 0$:
Dann gibt es $f_1, f_2 \in K[x]$, so dass

$$f_1 p_1 + f_2 p_2 = \mathrm{ggT}(p_1, p_2)\ .$$

Sind also insbesondere p_1, p_2 teilerfremd, dann

$$f_1 p_1 + f_2 p_2 = 1\ .$$

Beweis: Im Beweis von Satz B.24 wird gezeigt:

$$< \mathrm{ggT}(p_1, p_2) >= \{f_1 p_1 + f_2 p_2 : f_1, f_2 \in K[x]\}$$ □

Bemerkung B.27 Der größte gemeinsame Teiler für p_1, p_2 kann wie in $\mathbb{Z}$ auch mit dem EUKLIDischen Algorithmus beschrieben werden, was auch eine (konstruktive) Beweismethode für Korollar B.26 ergibt. △

Definition B.28

Ein $p \in K[x]$ mit $\mathrm{grad}(p) \geq 1$ heißt *irreduzibel*, wenn aus der Darstellung

$$p = fg \text{ mit } f, g \in K[x]$$

folgt: $f \in K$ oder $g \in K$

Bemerkungen B.29

1) Nach der Gradformel (B.10) ist ein lineares Polynom (grad(p) = 1) immer irreduzibel.

2) Satz B.21, 4) zeigt also, dass für algebraisch abgeschlossenes K sich ein $f \in K[x]$ bis auf einen Faktor als Produkt von normierten irreduziblen Polynomen vom Grad 1 schreiben lässt.

3) Ist K algebraisch abgeschlossen, dann sind äquivalent:

(i) $f, g \in K[x]$ sind teilerfremd.

(ii) f und g haben keine gemeinsame Nullstelle.

Allgemein gilt: Jedes $f \in K[x]$ lässt sich als Produkt aus einem Faktor und normierten irreduziblen Polynomen schreiben, das eindeutig bis auf die Reihenfolge der Faktoren ist. (Dies ist ein Spezialfall von Aussagen der Algebra, siehe Reiffen, Scheja und Vetter 1969.)

4) Seien f_1, f_2 teilerfremde Polynome und es gelte für ein Polynom h: $f_1 \mid h$ und $f_2 \mid h$. Dann gilt auch $f_1 f_2 \mid h$.

Das kann man folgendermaßen einsehen:
Nach Voraussetzung existieren Polynome g_1, g_2, so dass

$$h = g_1 f_1 = g_2 f_2$$

f_1 und f_2 zerfallen nach 3) jeweils in irreduzible Faktoren, die nach Voraussetzung keinen gemeinsamen Faktor haben und also alle in der eindeutigen Zerlegung von h auftreten müssen.

△

Trotz (B.4) ist $\mathbb{R}$ nicht algebraisch abgeschlossen, z. B. hat

$$x^2 = -1 \tag{B.17}$$

keine Lösung in $\mathbb{R}$. In der Linearen Algebra bedeutet dies, dass nicht alle reellen Matrizen reelle Eigenwerte besitzen. Um auch Gleichungen wie (B.17) lösbar zu machen, können auf

$$\mathbb{C} := \mathbb{R} \times \mathbb{R}$$

die Verknüpfungen + und · von $\mathbb{R}$ kompatibel fortgesetzt werden durch:

$$\begin{aligned} (x, y) + (x', y') &:= (x + x', y + y') \\ (x, y) \cdot (x', y') &:= (xx' - yy', xy' + yx') \end{aligned} \qquad \text{für } (x, y), (x', y') \in \mathbb{C}$$

wobei

$$x \in \mathbb{R} \mapsto (x, 0) \in \mathbb{C}$$

die Einbettung ist (siehe Beispiele 3.11, 2)).

Definition B.30

Sei $i := (0, 1) \in \mathbb{C}$, die *imaginäre Einheit*. Für $x, y \in \mathbb{R}$ gilt dann für

$$z := (x, y) = (x, 0) + i(y, 0) =: x + iy : \ z \in \mathbb{C} ,$$

$x = \operatorname{Re} z$ heißt *Realteil*, $y := \operatorname{Im} z$ *Imaginärteil* von z.

$$\overline{z} := x - iy \ (= x + i(-y))$$

heißt *komplex-konjugiert* zu z.

Bemerkungen B.31

1) Es gilt (siehe (3.4), (3.5)) für $z_1, z_2 \in \mathbb{C}$:

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} , \qquad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$$

2) Jedes $f \in \mathbb{R}[x]$ kann auch als $f \in \mathbb{C}[x]$ aufgefasst werden.

3) Nach 1) folgt für $f \in \mathbb{R}[x]$: Hat f die Nullstelle $z \in \mathbb{C}$, so auch die Nullstelle $\overline{z}$. Echt komplexe Nullstellen reeller Polynome treten also in komplex-konjugierten Paaren auf (mit gleicher Vielfachheit).

4) Sei $f \in \mathbb{R}[x]$, seien $x_1, \ldots, x_k \in \mathbb{R}$ die paarweise verschiedenen reellen Nullstellen mit Vielfachheit r_j (die eventuell nicht auftreten), $z_1 = a + ib_1, \ldots z_l = a_l + ib_l$ seien die paarweise verschiedenen echt komplexen Nullstellen mit Vielfachheit s_j (wobei nur entweder z oder $\overline{z}$ aufgenommen worden ist). Dann hat f die folgende (eindeutige) Zerlegung in irreduzible Polynome:

$$f = kp_1^{r_1} \cdots p_k^{r_k} q_1^{q_1} \cdots p_l^{r_l} ,$$

wobei $k \in \mathbb{R}$ und

$$p_i(x) = x - x_i \ , \ i = 1, \ldots, k \quad q_i(x) = (x - a_i)^2 + b_i^2 \ , \ i = 1, \ldots, l.$$

Man beachte dabei die komplexe Zerlegung in Linearfaktoren ausgewertet bei $x \in \mathbb{R}$:

$$(x - z_i)(x - \overline{z_i}) = x^2 - 2a_i x + |z_i|^2 = (x - a_i)^2 + b_i^2$$

5) Zu $z = x + iy$ existiert eindeutig ein $\varphi \in [0, 2\pi)$, so dass

$$\cos(\varphi) = \frac{1}{|z|}x\,, \quad \sin(\varphi) = \frac{1}{|z|}y\,,$$
$$\text{also } z = |z|(\cos\varphi + i\sin\varphi) =: |z|\exp(i\varphi)$$
$$\text{und damit auch } z = |z|\exp(i(\varphi + n2\pi)) \text{ für } n \in \mathbb{Z}\,.$$
△

Satz B.32

Sei $n \in \mathbb{N}$.

1) Die Gleichung $z^n - 1 = 0$ hat in $\mathbb{C}$ n verschiedene Lösungen, die n-ten *Einheitswurzeln*, nämlich ζ_n^k für $k = 0, \ldots, n-1$, wobei

$$\zeta_n^k := \exp\left(i\frac{2\pi k}{n}\right) \text{ definiert ist für } k \in \mathbb{Z}, \tag{B.18}$$

2) Für $j \in \mathbb{Z}$ gilt:

$$\begin{aligned}(\zeta_n^j)^k &= (\zeta_n^{jk}) = (\zeta_n^k)^j \\ \zeta_{nm}^{jm} &= \zeta_n^j \text{ für } m \in \mathbb{Z}, m \neq 0 \\ \overline{\zeta_n^j} &= \zeta_n^{-j}\end{aligned}$$

Beweis: Es ist die Aussage

$$\exp(i\varphi)^n = \exp(in\varphi) \text{ für } n \in \mathbb{N}$$

nötig, die aus den Additionstheoremen für sin und cos hergeleitet werden kann. Dann folgt 1) und die erste Aussage von 2). Der Rest ist klar. □

Hauptsatz B.33

1) $(\mathbb{C}, +, \cdot)$ ist ein Körper.

2) $\mathbb{C}$ ist algebraisch abgeschlossen. (*Fundamentalsatz der Algebra*)

3) $\mathbb{C}$ ist nicht verträglich und total anordenbar.

Beweis: zu 1) Siehe Beispiele 3.11, 2).
zu 2) Sei $f(z) = \sum_{k=0}^n a_k z^k$, $a_n \neq 0$ und $n \geq 1$. Sei $M := \{z \in \mathbb{C} : |f(z)| \leq 1\}$. Wir zeigen:

a) M ist kompakt in $\mathbb{C}$ (siehe Definition C.10), d. h. $|f|$ nimmt sein Infimum auf M an (Satz C.12, 2)), sei $\overline{z}$ die Minimalstelle (d. h. $|f(\overline{z})| = \min\{|f(z)| : z \in M\}$).

b) Ist $z \in M$, dann ist z eine Nullstelle von f oder es gibt ein $\widetilde{z} \in M$, so dass $|f(\widetilde{z})| < |f(z)|$, d. h. notwendig gilt $|f(\overline{z})| = 0$, und damit ist $\overline{z}$ eine Nullstelle.

zu a): M ist abgeschlossen wegen der Stetigkeit von $|f|$ (Satz C.9) und beschränkt, d. h. kompakt (Satz C.11, 2)). Würde nämlich eine Folge $(z_n)_n$ in M existieren, so dass $|z_n| \to \infty$ für $n \to \infty$, dann würde folgen

$$\begin{aligned} |f(z_n)| &\geq |a_n||z_n|^n - \sum_{k=0}^{n-1} |a_k||z_n|^k = |z_n|^n \left(|a_n| - \sum_{k=0}^{n-1} |a_k||z_n|^{k-n} \right) \\ &\geq |z_n|^n \frac{|a_n|}{2} \quad \text{für } n \geq n_0 \text{ für ein } n_0 \in \mathbb{N} \\ &> 1 \quad \text{für } n \geq n_1 \text{ für ein } n_1 \in \mathbb{N} \end{aligned}$$

im Widerspruch zu $z_n \in M$.

zu b): Sei $z \in M$, dann gibt es $b_0, \ldots, b_n \in \mathbb{C}$, so dass $f(z + w) = \sum_{i=0}^n b_i w^i$. Ist $b_0 = 0$, dann ist z Nullstelle von f, sonst sei $k \geq 1$ der erste Index, so dass $b_k \neq 0$. Da $|b_0| \geq |b_k|s^k$ und $|b_k|/2 \geq \sum_{i=k+1}^n |b_i|s^{i-k}$ für $s = 0$ mit $>$ erfüllt sind, gibt es ein (maximales) $s > 0$, so dass die Ungleichungen gelten.
Sei $0 < c \leq 1$ definiert durch $c|b_0| = |b_k|s^k$, dann gibt es (siehe Satz B.32, 1), hier geht $K = \mathbb{C}$ ein) ein $w \in \mathbb{C}$ mit $|w| = s$ und $w^k = -cb_0/b_k$.

Also folgt

$$\begin{aligned} |f(z + w)| &\leq |b_0 + b_k w^k| + \sum_{i=k+1}^{n} |b_i||w|^i = (1 - c)|b_0| + \sum_{i=k+1}^{n} |b_i|s^i \\ &= |b_0| - s^k \left(|b_k| - \sum_{i=k+1}^{n} |b_i|s^{i-k} \right) \leq |f(z)| - \frac{1}{2}|b_k|s^k < |f(z)|. \end{aligned}$$

zu 3) Sei $\leq$ eine verträgliche totale Ordnung auf $\mathbb{C}$. Dann muss (unabhängig von $\mathbb{C}$) gelten

$$a \cdot a \geq 0\,,$$

da aus $a \geq 0$ folgt $a \cdot a \geq 0 \cdot a = 0$ und auch aus $a \leq 0$ folgt $a \cdot a = (-a) \cdot (-a) \geq 0 \cdot (-a) = 0$. Für $a = (0, 1)$ ergibt sich in $\mathbb{C}$ ein Widerspruch. □

Anhang C
Analysis in normierten Räumen

Für Kapitel 7 und darauf aufbauende Abschnitte aus Kapitel 8 ist es nötig, in einem $\mathbb{K}$-Vektorraum auch einen Konvergenzbegriff, definiert durch eine Norm, zu haben. Die zugehörige Analysis in normierten Räumen, die zwischen reeller Analysis und Funktionalanalysis steht, wird im Folgenden in ihrem Nötigsten skizziert. Dabei wird der übliche Kenntnisstand an reeller Analysis eines ersten Semesters vorausgesetzt.

Wie immer ist $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Wenn nicht anders erwähnt, ist im Folgenden $(V, \|\,.\,\|)$ ein normierter Vektorraum (siehe Definition 1.91 oder Definition 7.1).

Definition C.1

Sei $(\boldsymbol{x}_n)_n$ eine Folge in V, $\boldsymbol{x} \in V$.
$(\boldsymbol{x}_n)_n$ konvergiert gegen $\boldsymbol{x}$, in Zeichen $\boldsymbol{x}_n \to \boldsymbol{x}$ für $n \to \infty$, genau dann, wenn

$$\|\boldsymbol{x}_n - \boldsymbol{x}\| \to 0 \quad \text{für} \quad n \to \infty \Leftrightarrow \text{Für alle } \varepsilon > 0 \text{ gibt es ein } n_0 \in \mathbb{N}\,,$$
$$\text{so dass } \|\boldsymbol{x}_n - \boldsymbol{x}\| \leq \varepsilon \text{ für } n \geq n_0\,.$$

Mit $d(\boldsymbol{v}, \boldsymbol{w}) := \|\boldsymbol{v} - \boldsymbol{w}\|$ ist also auf $(V, \|\,.\,\|)$ eine *Metrik* definiert worden, deren erzeugte *Topologie* zugrunde gelegt werden soll.

Definition C.2

Sei $\varepsilon > 0, \boldsymbol{v}_0 \in V$, dann:

$$B_\varepsilon(\boldsymbol{v}_0) := \{\boldsymbol{v} \in V : \|\boldsymbol{v} - \boldsymbol{v}_0\| < \varepsilon\}$$

bzw.

$$\overline{B}_\varepsilon(\boldsymbol{v}_0) := \{\boldsymbol{v} \in V : \|\boldsymbol{v} - \boldsymbol{v}_0\| \leq \varepsilon\}\,,$$

bzw.

$$\partial B_\varepsilon(\boldsymbol{v}_0) := \{\boldsymbol{v} \in V : \|\boldsymbol{v} - \boldsymbol{v}_0\| = \varepsilon\}$$

heißen die *offene* bzw. *abgeschlossene Kugel* bzw. der *Rand der Kugel* um $\boldsymbol{v}_0$ mit Radius ε.
Sei $M \subset V$. M heißt *offen*, wenn für alle $\boldsymbol{v}_0 \in V$ ein $\varepsilon = \varepsilon(\boldsymbol{v}_0) > 0$ existiert, so dass

$$\overline{B}_\varepsilon(\boldsymbol{v}_0) \subset M.$$

M heißt *abgeschlossen*, wenn $V \backslash M$ offen ist.

Daraus folgt sofort:

Satz C.3

Sei $\varepsilon > 0, \boldsymbol{v}_0 \in V$.

1) $\emptyset$ und V sind offen.

2) Die Vereinigung beliebig vieler [endlich vieler] offener [abgeschlossener] Mengen ist offen [abgeschlossen].

3) Der Schnitt endlich vieler [beliebig vieler] offener [abgeschlossener] Mengen ist offen [abgeschlossen].

4) $B_\varepsilon(\boldsymbol{v}_0)$ ist offen, $\overline{B}_\varepsilon(\boldsymbol{v}_0)$ und $\partial B_\varepsilon(\boldsymbol{v}_0)$ sind abgeschlossen.

Beweis: Siehe z. B. AMANN und ESCHER 1998, S. 246, 247. □

Es ist auch eine Charkterisierung über Folgenkonvergenz möglich:

Satz C.4

Sei $M \subset V$. M ist genau dann abgeschlossen, wenn für jede Folge $(\boldsymbol{v}_n)_n$ mit $\boldsymbol{v}_n \in M$ für alle $n \in \mathbb{N}$ gilt:

$$\boldsymbol{v} := \lim_{n\to\infty} \boldsymbol{v}_n \text{ existiert } \Rightarrow \boldsymbol{v} \in M\,.$$

Beweis: Siehe z. B. AMANN und ESCHER 1998, S. 247. □

Definition C.5

Sei $M \subset V, \boldsymbol{v} \in M$.

1) $\boldsymbol{v}$ gehört zum *Inneren* von M, in Zeichen $\boldsymbol{v} \in \mathrm{int}(M)$, genau dann, wenn ein $\varepsilon > 0$ existiert, so dass $\overline{B}_\varepsilon(\boldsymbol{v}) \subset M$.

2) Der *Abschluss* von M, in Zeichen $\mathrm{cl}(M)$, wird definiert durch

$$\mathrm{cl}(M) := V \setminus \mathrm{int}(V \setminus M) .$$

3) Gilt $\mathrm{cl}(M) = V$, so heißt M *dicht* in V.

Die Charakterisierung über Folgenkonvergenz lautet:

Satz C.6

Sei $M \subset V, \boldsymbol{v}_0 \in M$.
$\boldsymbol{v}_0 \in \mathrm{cl}(M) \Leftrightarrow$ Es gibt eine Folge $(\boldsymbol{v}_n)_n$ mit $\boldsymbol{v}_n \in M$ für alle $n \in \mathbb{N}$ und

$$\lim_{n\to\infty} \boldsymbol{v}_n = \boldsymbol{v}_0 .$$

Beweis: Siehe z. B. Amann und Escher 1998, S. 246-250. □

Seien $(V, \|\,.\,\|_V), (W, \|\,.\,\|_W)$ normierte Räume, wobei die unterscheidende Indizierung im Folgenden weggelassen wird. Stetigkeit kann über ein aus der reellen Analysis als äquivalent bekanntes Kriterium definiert werden:

Definition C.7

Sei $f : V \to W, \boldsymbol{v}_0 \in V$.
f heißt *stetig* in $\boldsymbol{v}_0$ genau dann, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass aus $\|\boldsymbol{v} - \boldsymbol{v}_0\| \leq \delta$ folgt: $\|f(\boldsymbol{v}) - f(\boldsymbol{v}_0)\| \leq \varepsilon$, bzw. dazu äquivalent:
Wenn für jede Folge $(\boldsymbol{v}_n)_n$ in V mit $\lim_{n\to\infty} \boldsymbol{v}_n = \boldsymbol{v}_0$ gilt

$$\lim_{n\to\infty} f(\boldsymbol{v}_n) = f(\boldsymbol{v}_0) .$$

f heißt *stetig* bzw. *stetig auf* V, wenn f stetig von v_0 ist für jedes $v_0 \in V$.

Beweis (der Äquivalenz): Siehe z. B. Amann und Escher 1998, S. 236. □

Satz C.8

Seien U, V, W normierte Räume, $f : U \to V$ und $g : V \to W$ Abbildungen.
Ist f in $\boldsymbol{u}_0 \in U$ stetig und g in $f(\boldsymbol{u}_0)$ stetig, dann ist auch $g \circ f$ in $\boldsymbol{u}_0$ stetig.

Beweis: Siehe z. B. AMANN und ESCHER 1998, S. 239. □

Satz C.9

Sei $f : V \to W$.
f ist stetig auf V (d. h. für alle $\boldsymbol{v}_0 \in V$) genau dann, wenn

$$f^{-1}(M) \text{ offen ist für beliebige offene } M \subset W$$

bzw. äquivalent

$$f^{-1}(M) \text{ abgeschlossen ist für beliebige abgeschlossene } M \subset W\,.$$

Beweis: siehe z. B. AMANN und ESCHER 1998, S.253. □

Definition C.10

Sei $(V, \|\,.\,\|)$ ein normierter $\mathbb{K}$-Vektorraum, $K \subset V$.

1) K heißt *beschränkt*, wenn ein $k > 0$ existiert, so dass $K \subset \overline{B}_k(\mathbf{0})$ bzw. $\|\boldsymbol{v}\| \leq k$ für alle $\boldsymbol{v} \in K$ gilt.

2) K heißt *kompakt*, wenn aus jeder offenen Überdeckung, d. h. offenen $U_i \subset V, i \in I$ mit $K \subset \bigcup\{U_i : i \in I\}$ eine endliche Teilüberdeckung ausgewählt werden kann, d. h. es gibt ein endliches $J \subset I$ mit $K \subset \bigcup\{U_i : i \in J\}$.

Von den benötigten Aussagen hat nur die folgende einen nicht elementaren Beweis:

Satz C.11

Sei $(V, \|\,.\,\|)$ ein normierter $\mathbb{K}$-Vektorraum.

1) Ist $K \subset V$ kompakt, dann ist K abgeschlossen und beschränkt.

2) Ist $(V, \|\,.\,\|) = (\mathbb{K}^n, \|\,.\,\|_2)$, dann gilt in 1) auch die Umkehrung (Satz von HEINE-BOREL).

3) Gilt in 1) die Umkehrung, so ist V endlichdimensional.

4) $K \subset V$ ist kompakt genau dann, wenn für jede Folge $(\boldsymbol{v}_n)_n$ in V gilt: Es gibt eine Teilfolge $(\boldsymbol{v}_{n_k})_k$ und ein $\boldsymbol{v}_0 \in K$, so dass

$$\boldsymbol{v}_{n_k} \to \boldsymbol{v}_0 \quad \text{für} \quad k \to \infty\,.$$

Beweis: Zu 1): Siehe z. B. AMANN und ESCHER 1998, S. 264.
Zu 2): Siehe z. B. AMANN und ESCHER 1998, S. 266.
Zu 4): Siehe z. B. AMANN und ESCHER 1998, S. 266.
Zu 3): Aus der Umkehrung von 1) folgt, dass die abgeschlossene Einheitskugel kompakt ist. Dann folgt die Aussage mit ALT 2006, Satz 2.9. □

Daraus ergibt sich sofort:

Satz C.12

Seien V, W normierte Räume, $K \subset V$ und $f : K \to W$ stetig.

1) Ist K kompakt, so ist auch $f(K)$ kompakt.

2) Ist $W = \mathbb{R}$ und K kompakt, so nimmt f sein Infimum bzw. Supremum auf K an.

Beweis: Zu 1): Siehe AMANN und ESCHER 1998, S. 267.
Zu 2): unmittelbar aus 1). □

Kurz gilt also:

$$\begin{gathered}\text{kompakt} = \text{abgeschlossen} + \text{beschränkt}\\ \Leftrightarrow \\ V \text{ endlichdimensional}\,.\end{gathered}$$

Definition C.13

Sei $(V, \|\,.\,\|)$ normierter Raum,

1) $(\boldsymbol{v}_n)_n$ sei Folge in V. $(\boldsymbol{v}_n)_n$ heißt *CAUCHY-Folge*, wenn für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert, so dass

$$\|\boldsymbol{v}_n - \boldsymbol{v}_m\| \leq \varepsilon \quad \text{für} \quad n, m \geq n_0 .$$

2) $(V, \|\,.\,\|)$ heißt *vollständig* oder BANACH-Raum, wenn zu jeder CAUCHY-Folge $(\boldsymbol{v}_n)_n$ in V ein $\boldsymbol{v}_0 \in V$ existiert, so dass

$$\boldsymbol{v}_n \to \boldsymbol{v}_0 \quad \text{für} \quad n \to \infty\,.$$

$(\mathbb{R}, |\,.\,|)$ und darauf aufbauend $(\mathbb{C}, |\,.\,|)$ bzw. $(\mathbb{K}^n, \|\,.\,\|_2)$ zeichnen sich gegenüber z. B. $(\mathbb{Q}, |\,.\,|)$ dadurch aus, dass sie vollständig sind.

Literaturverzeichnis

ALT, H. W. (2006). *Lineare Funktionalanalysis.* 5. Aufl. Berlin: Springer.

AMANN, H. (1995). *Gewöhnliche Differentialgleichungen.* Berlin: De Gruyter.

AMANN, H. und J. ESCHER (1998). *Analysis I.* 1. Aufl. Basel, Boston, Berlin: Birkhäuser.

— (1999). *Analysis II.* Berlin: Birkhäuser.

BEN-ISRAEL, A. und T. N. E. GREVILLE (2003). *Generalized Inverses: Theory and Applications.* 2. Aufl. Berlin: Springer.

BERMAN, A. und R. J. PLEMMONS (1994). *Nonnegative Matrices in the Mathematical Sciences.* Philadelphia: SIAM.

BUNSE, W. und J. BUNSE-GERSTNER (1985). *Numerische lineare Algebra.* Stuttgart: Teubner.

CHVATAL, V. (1983). *Linear Programming.* New York: W. H. Freeman Publication.

COOLEY, J. W. und J. W. TUKEY (1965). „An Algorithm for the Machine Calculation of the Complex Fourier Series". In: *Math. Comp.* 19, S. 297–301.

DANTZIG, G. (1966). *Lineare Programmierung und Erweiterungen.* Berlin: Springer.

DEISER, O. (2002). *Einführung in die Mengenlehre.* Berlin, Heidelberg, New York: Springer.

DEMMEL, J. W. (1997). *Applied Numerical Linear Algebra.* Philadelphia: SIAM.

DEUFLHARD, P. (2006). *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms.* Berlin, Heidelberg, New York: Springer.

DEUFLHARD, P. und A. HOHMANN (1991). *Numerische Mathematik: Eine algorithmisch orientierte Einführung.* Berlin, New York: de Gruyter.

DONGARRA, J. und F. SULLIVAN (2000). „Guest Editors' Introduction: The Top 10 Algorithms". In: *Computing in Science and Engineering* 2.1, S. 22–23.

ECK, C., H. GARCKE und P. KNABNER (2011). *Mathematische Modellierung.* 2. Aufl. Berlin: Springer.

ELAYDI, S. (2005). *An Introduction to Difference Equations.* 3. Aufl. Berlin: Springer.

FISCHER, G. (1978). *Analytische Geometrie.* Reinbek bei Hamburg: Rowohlt.

FORSTER, O. (2008). *Analysis 1.* 9. Aufl. Wiesbaden: Vieweg.

GOLUB, G. H. und C. F. VAN LOAN (1996). *Matrix Computations.* Johns Hopkins University Press.

HACKBUSCH, W. (1991). *Iterative Lösung großer schwachbesetzter Gleichungssysteme.* Stuttgart: Teubner.

HIGHAM, N. J. (1996). *Accuracy and Stability of Numerical Algorithms.* Philadelphia: SIAM.

HUPPERT, B. (1990). *Angewandte lineare Algebra.* Berlin: de Gruyter.

HUPPERT, B. und W. WILLEMS (2006). *Lineare Algebra.* 1. Aufl. Wiesbaden: Teubner.

JARRE, F. und J. STOER (2004). *Optimierung.* Berlin, Heidelberg: Springer.

JECH, T. J. (1973). *The Axiom of Choice.* Amsterdam: North Holland.

KNABNER, P. und L. ANGERMANN (2000). *Numerik partieller Differentialgleichungen.* Berlin: Springer.

LAX, P. (2007). *Linear Algebra and its Applications.* Hoboken, N.J.: Wiley-Interscience.

NEWMAN, M. E. J. (2010). *Networks: An Introduction.* Oxford: Oxford University Press.

Putzer, E. J. (1966). „Avoiding the Jordan Canonical form in the Discussion of Linear Systems with Constant Coefficients“. In: *American Mathematical Monthly* 73.1, S. 2–7.

Reiffen, H.-J., G. Scheja und U. Vetter (1969). *Algebra*. Mannheim: Bibliographisches Institut.

Saad, Y. (2003). *Iterative Methods for Sparse Linear Sytems*. Philadelphia: SIAM.

— (2011). *Numerical Methods for Large Eigenvalue Problems*. 2nd. Philadelphia: SIAM.

Schumann, J. (1968). *Input-Output-Analysen*. Berlin: Springer.

Strang, G. (2003). *Lineare Algebra*. Berlin, Heidelberg: Springer.

Tutte, W. T. (2001). *Graph Theory*. Cambridge: Cambridge University Press.

Watkins, D. S. (2007). *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods*. Philadelphia: SIAM.

Wolsey, L. A. (1998). *Integer Programming*. New York: Wiley-Interscience Publication.

Sachverzeichnis