Notizen zum Vorkurs der Mathematik
Universität Bonn, Sommer 2015

Prof. Dr. Carsten Burstedde

15. Oktober 2015

1 Wahr und falsch

Wir suchen in der Mathematik nach Aussagen, die in einem definierten Kontext als wahr gelten. Wir versuchen, aus einer Aussage eine andere zu folgern, so daß die zweite auch wahr ist, wenn die erste wahr ist. Wir verwenden eine Logik, in der eine Aussage entweder wahr oder falsch sein kann, also nicht beides zugleich und auch nicht etwas anderes außer wahr und falsch: *Tertium non datur*. (Man kann sicher versuchen, eine Logik mit mehr als zwei Zuständen zu erklären, aber das machen wir an dieser Stelle nicht.)

1.1 Wahrheitswerte und Operatoren

Der Einfachheit halber schreiben wir den Wert falsch als 0 und den Wert wahr als 1. Die Menge der beiden Werte sei $W = \{0, 1\}$. Wir definieren zunächst zwei elementare Operationen.

Definition 1.1. Die Negation ist ein unärer Operator (verarbeitet ein Argument), $\neg : W \to W$. Die Konjunktion (der *und*-Operator) ist binär, $\wedge : W \times W \to W$. Die Wertetabellen sind

$$\begin{array}{|cc|}\hline a & \neg a \\ \hline 0 & 1 \\ 1 & 0 \\ \hline\end{array}, \qquad \begin{array}{|ccc|}\hline a & b & a \wedge b \\ \hline 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ \hline\end{array}. \tag{1.1}$$

Eigenschaft 1.2. *Die doppelte Negation führt auf die ursprüngliche Aussage,*

$$\neg\neg a = a. \tag{1.2}$$

(Dies wird in Beweisen benutzt, wenn eine Aussage für alle Werte von $\neg a$ gilt und gefolgert wird, daß sie auch für alle a gilt.) Die Konjunktion erfüllt

$$0 \wedge a = 0, \qquad 1 \wedge a = a, \qquad a \wedge \neg a = 0. \tag{1.3}$$

Beweis. Wir greifen zur ersten wesentlichen Beweistechnik: Wir setzen alle möglichen Werte ein und überprüfen die Aussage anhand der Wertetabelle (Einsetzen und Überprüfen, EuÜ). □

Satz 1.3. *Die Konjunktion ist assoziativ, kommutativ und idempotent. In Formeln gilt also in dieser Reihenfolge*

$$a \wedge (b \wedge c) = (a \wedge b) \wedge c, \tag{1.4a}$$

$$a \wedge b = b \wedge a, \tag{1.4b}$$

$$a \wedge a = a. \tag{1.4c}$$

Beweis. Kommutativität und Idempotenz lassen sich direkt aus (1.1) ablesen. Zum Beweis der Assoziativität nehmen wir zunächst a an, also daß $a = 1$. Durch zweimaliges Anwenden von (1.3) führt das auf $b \wedge c = b \wedge c$, diese Aussage ist immer (d.h. für alle möglichen Werte von b und c) wahr. Nun nehmen wir $\neg a$ an, also $a = 0$, es folgt durch dreimaliges (?!) Anwenden von (1.3) $0 = 0$, was ebenfalls immer wahr ist. Diese Technik kann man mit teile-und-herrsche beschreiben (divide-and-conquer, DaC). □

Im Folgenden werden wir die Assoziativität still ausnutzen und z.B. schreiben $a \wedge b \wedge c$.

Bemerkung 1.4. Einige Aussagen in (1.4) erscheinen „selbstverständlich." Diesen Begriff gibt es jedoch bei uns nicht: Jede einzelne Aussage muß aus den vorhergehenden hergeleitet werden können, sonst ist sie wertlos (und Sie bekommen keine Punkte!). Weiterhin muß in jedem Schritt angegeben werden, *welche* vorhergehenden Aussagen verwendet werden. Darüberhinaus ist es im Sinne der mathematischen Eleganz und Allgemeinheit erstrebenswert, so wenige wie möglich der bisherigen Aussagen zu verwenden. Im Beweis der Assoziativität haben wir beispielsweise die Kommutativität *nicht* benutzt.

Definition 1.5. Wir erklären die Disjunktion (den binären *oder*-Operator) $\vee : W \times W \to W$ durch

$$a \vee b = \neg(\neg a \wedge \neg b) \tag{1.5}$$

und das exklusive oder $\oplus : W \times W \to W$ durch

$$a \oplus b = (a \wedge \neg b) \vee (\neg a \wedge b). \tag{1.6}$$

Die Konjunktion bindet per Konvention stärker als die Disjunktion, d.h. wir könnten anstelle von $(a \wedge b) \vee c$ schreiben $a \wedge b \vee c$, wir werden jedoch in jedem Fall Klammern verwenden.

Eigenschaft 1.6. *Die Wertetabelle für Diskunktion und exklusives oder lautet*

$$\left[\begin{array}{cccc} a & b & a \vee b & a \oplus b \\ \hline 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{array}\right]. \tag{1.7}$$

Weiterhin gelten für die Disjunktion Assoziativität, Kommutativität und Idempotenz und

$$0 \vee a = a, \qquad 1 \vee a = 1, \qquad a \vee \neg a = 1. \tag{1.8}$$

Beweis. Die Wertetabelle und (1.8) lassen sich prinzipiell durch EuÜ beweisen. Dieser Weg ist jedoch auf Dauer unzureichend, da sehr viele Kombinationsmöglichkeiten entstehen können. Wir leiten daher eine Formel formal her (Her):

$$0 \vee a \overset{(1.5)}{=} \neg(\neg 0 \wedge \neg a) \overset{(1.1)}{=} \neg(1 \wedge \neg a) \overset{(1.3)}{=} \neg(\neg a) \overset{(1.2)}{=} a \tag{1.9}$$

Die anderen beiden sollen entsprechend als Übung bewiesen werden. □

Eigenschaft 1.7. *Es gilt*

$$\neg(a \wedge b) = \neg a \vee \neg b, \qquad \neg(a \vee b) = \neg a \wedge \neg b, \tag{1.10a}$$

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \qquad a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \tag{1.10b}$$

$$a \wedge (\neg a \vee b) = a \wedge b, \qquad a \vee (\neg a \wedge b) = a \vee b. \tag{1.10c}$$

Beweis. Zur Übung sollen diese Aussagen ohne Verwendung von Wertetabellen bewiesen werden. □

1.2 Folgerungen

Bisher haben wir uns mit der Verknüpfung von Wahrheitswerten befaßt. Wir können den Formalismus weiter ausbauen und auch das Ziehen von Folgerungen beschreiben.

Definition 1.8. Wir definieren durch die Wahrheitswerte von Voraussetzung und Folgerung

$$(a \Rightarrow b) = \neg(a \wedge \neg b). \tag{1.11a}$$

Wir sagen in diesem Fall, a ist eine hinreichende Bedingung für b, und b ist eine notwendige Bedingung für a. Äquivalent (d.h. gleichbedeutend) sagen wir auch „wenn a, dann b“. Wir treffen dabei keinerlei Aussage für den Fall $\neg a$! Wir definieren weiter

$$(a \Leftarrow b) = (b \Rightarrow a) \quad \text{und} \tag{1.11b}$$

$$(a \Leftrightarrow b) = (a \Rightarrow b) \wedge (b \Rightarrow a). \tag{1.11c}$$

Die letzte Gleichung (1.11c) wird ausgesprochen „b genau dann wenn a“(oder umgekehrt), oder „a ist äquivalent zu b“. Die doppelten Pfeile sind genauso binäre Operatoren wie $\wedge$ und $\vee$, allerdings mit schwächerer Bindung als diese (deswegen brauchen wir oben die Klammern).

Eigenschaft 1.9. *Die Wertetabelle lautet*

$$\left[\begin{array}{cccccc} a & b & a \Rightarrow b & a \Leftarrow b & a \Leftrightarrow b \\ \hline 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{array}\right]. \tag{1.12}$$

Wir sehen also direkt, daß

$$(a \Leftrightarrow b) = \neg(a \oplus b). \tag{1.13a}$$

Es gilt außerdem die Aussage

$$(a \Rightarrow b) = \neg a \vee b = (\neg b \Rightarrow \neg a). \tag{1.13b}$$

Dies ist die Grundlage des Beweises durch Widerspruch: Anstatt aus einer Voraussetzung eine Aussage zu folgern, können wir äquivalent das Gegenteil der Aussage annehmen und nachweisen, daß die Voraussetzung dann nicht erfüllt sein kann.

Beweis. Übung, die Formeln (1.13) sollen hergeleitet werden. □

Wir können nun eine ganz wesentliche Arbeitsweise formalisieren, nämlich das Herleiten einer Aussage aus einer anderen.

Satz 1.10. *Folgerungen können verkettet werden: Es gilt*

$$(a \Rightarrow b) \wedge (b \Rightarrow c) \Rightarrow (a \Rightarrow c). \tag{1.14}$$

Beweis. Wir setzen ein und benutzen sogar zweimal die Kommutativität (wo?!),

$$\begin{aligned} \big((a \Rightarrow b) \wedge (b \Rightarrow c) \Rightarrow (a \Rightarrow c)\big) &\overset{(1.13b)}{=} \neg\big((a \Rightarrow b) \wedge (b \Rightarrow c)\big) \vee (a \Rightarrow c) \\ &\overset{(1.10a)}{=} \neg(a \Rightarrow b) \vee \neg(b \Rightarrow c) \vee (a \Rightarrow c) \overset{(1.13b)}{=} \neg(\neg a \vee b) \vee \neg(\neg b \vee c) \vee (\neg a \vee c) \\ &\overset{(1.10a)}{=} (a \wedge \neg b) \vee (b \wedge \neg c) \vee \neg a \vee c \overset{(1.10c)}{=} \neg a \vee \neg b \vee b \vee c \\ &\overset{(1.8)}{=} \neg a \vee 1 \vee c \overset{(1.8)}{=} 1. \end{aligned} \tag{1.15}$$

Die zu beweisende Aussage stimmt also. □

Wir schreiben die Voraussetzung auch kürzer als

$$a \Rightarrow b \Rightarrow c, \tag{1.16}$$

jedoch wirft das die Frage auf, ob $\Rightarrow$ als binärer Operator assoziativ ist oder eine davon verschiedene Aussage vorliegt. Formal: Gelten eins oder beide der Gleichheitszeichen in

$$((a \Rightarrow b) \Rightarrow c) = (a \Rightarrow (b \Rightarrow c)) = (a \Rightarrow b) \wedge (b \Rightarrow c)?! \tag{1.17}$$

Wir hätten alle bisherigen Aussagen allein über Wahrheitstabellen beweisen können und damit das Instrument der Folgerung neu gefunden. Daß wir es auch herleiten konnten, während wir es schon implizit benutzten, zeigt zumindest, daß unsere Theorie bis hierhin widerspruchsfrei ist.

Wie steht es mit dem Instrument des DaC? Auch dieses hätten wir bisher mit Wahrscheinlichkeitstabellen umgehen und neu finden können.

Satz 1.11. *Es gilt*

$$(a \Rightarrow b) \wedge (\neg a \Rightarrow b) \Rightarrow b. \tag{1.18}$$

Diese bedeutet, daß wir die Aussage b in zwei Schritten folgern können, einmal unter der Annahme a und einmal unter der Annahme $\neg a$. Wenn b die Auswertung von a enthält, vereinfacht dieses Vorgehen möglicherweise jeden der beiden Schritte.

Beweis. Übung: Dies ist eine kurze Rechnung. □

1.3 Quantoren

Wir bezeichnen eine Aussage als wahr, wenn sie unabhängig von den auftretenden Variablen immer richtig ist, und als falsch, wenn es derer eine Kombination gibt, bei der die Aussage nicht richtig ist. Eine heißt hier wie immer in der Mathematik mindestens eine.

Definition 1.12. Die Aussage a enthalte Auswertungen der Aussagen a_i, $i \in I$, wobei I eine Indexmenge ist, die die Aussagen a_i numeriert. Wir identifizieren

$$a(a_i) \text{ ist wahr} = \forall a_i \in W : a(a_i) \tag{1.19a}$$

$$a(a_i) \text{ ist falsch} = \exists a_i \in W : \neg a(a_i). \tag{1.19b}$$

Die sogenannten Quantoren $\forall$ („für alle“) und $\exists$ („es gibt ein“) stehen immer vor der Aussage. Dies hat seinen Grund u.a. darin, daß dann die Verkettung von Negationen einfach zu bewerkstelligen ist.

Eigenschaft 1.13. *Aussagen mit Quantoren werden wie folgt negiert,*

$$\neg\,(\forall a : b) = \exists a : \neg b, \tag{1.20a}$$

$$\neg\,(\exists a : b) = \forall a : \neg b. \tag{1.20b}$$

Beweis. Dies folgt aus der Kombination von (1.19) mit (1.2). □

Verneinte Verkettungen lauten dann (ohne Zwischendoppelpunkte und Klammern) z.B.

$$\neg \forall a\, \exists b\, \forall c : d = \exists a\, \forall b\, \exists c : \neg d. \tag{1.21}$$

Man sagt zu $\neg\exists$ auch „es gibt kein.“

2 Mathematische Grundbegriffe

In diesem Abschnitt führen wir einige wesentliche Begriffe ein, auf denen mehr oder weniger die gesamte Mathematik beruht.

2.1 Mengen und Abbildungen

Im letzten Abschnitt hatten wir schon Mengen und Abbildungen benutzt. Wir verwenden den Begriff einer Menge hier naiv als Ansammlung von Objekten, die gewisse Eigenschaften erfüllen, diese Ansammlung kann die leere Menge $\emptyset = \{\}$ sein. Wir ignorieren beispielsweise Fragestellungen wie die, ob die Menge aller möglichen Mengen sich selbst enthält.

Wir definieren nun Mengen und Beziehungen zwischen Mengen.

Definition 2.1. Eine Menge A' heißt Teilmenge von A, wenn sie ausschließlich Elemente aus A enthält. Es ist vollkommen zulässig, wenn eine oder beide der Mengen leer sind (falls genau eine, welche muß es sein?!). Umgekehrt heißt A Obermenge von A'. In Formeln:

$$(A' \subset A) = (a \in A' \Rightarrow a \in A), \tag{2.1a}$$
$$(A' \not\subset A) = \neg(A' \subset A), \tag{2.1b}$$
$$(A' = A) = (A' \subset A) \wedge (A \subset A'). \tag{2.1c}$$

Definition 2.2. Wir können zwei Mengen miteinander schneiden, vereinigen und die Differenzmenge bilden,

$$A \cap B = \{a : a \in A \wedge a \in B\}, \tag{2.2a}$$
$$A \cup B = \{a : a \in A \vee a \in B\}, \tag{2.2b}$$
$$A \setminus B = \{a \in A : a \notin B\}. \tag{2.2c}$$

Definition 2.3. Es seien A, B Mengen. Eine Vorschrift f, die jedem $a \in A$ genau ein Element $b \in B$ zuordnet, heißt Abbildung oder Funktion, geschrieben als

$$f : A \to B, \quad a \mapsto b = f(a). \tag{2.3}$$

A heißt Definitionsbereich von f, B Wertebereich. Für den Fall $A = B$ definieren wir die Identitätsabbildung

$$\mathrm{id}_A : A \to A, \quad a \mapsto a. \tag{2.4}$$

Wir verallgemeinern die Schreibweise auf Abbildungen von Mengen auf Mengen und bezeichnen mit

$$f(A) = \{b \in B : \exists a \in A : f(a) = b\} \tag{2.5}$$

das Bild von A unter f. Das Urbild einer Teilmenge $B' \subset B$ ist

$$f^{-1}(B') = \{a \in A : \exists b \in B' : f(a) \in B'\}. \tag{2.6}$$

Beispiel 2.4. Hier sind ein paar Beispielmengen:

$$A = \{1, 2, 3, 4, 5, 6\} \tag{2.7a}$$
$$B = A \cup \{7\} \tag{2.7b}$$
$$C = \{a \in A : a \text{ ist gerade}\} = \{2, 4, 6\} \subset A \tag{2.7c}$$
$$D = A \setminus C = \{a \in A : a \text{ ist ungerade}\} = \{1, 3, 5\} \subset A \tag{2.7d}$$
$$E = C \setminus A = \emptyset \tag{2.7e}$$

Wenn wir die Funktion $f : A \to B$, $a \mapsto a + 1$ definieren, erhalten wir:

$$f(C) \not\subset A \tag{2.8a}$$
$$f(D) = C \tag{2.8b}$$
$$f^{-1}(C) = D \tag{2.8c}$$
$$f^{-1}(1) = \emptyset \tag{2.8d}$$

Definition 2.5. Eine Abbildung $f : A \to B$ heißt surjektiv, wenn $f(A) = B$, der Wertebereich also komplett abgedeckt wird. f heißt injektiv, wenn

$$\forall a, a' \in A : f(a) = f(a') \Rightarrow a = a', \tag{2.9}$$

wenn also jedes Element aus B ein höchstens einelementiges Urbild hat. Eine Funktion heißt bijektiv oder eineindeutig, wenn sie injektiv und surjektiv ist.

Bijektivität bedeutet, daß es zu jedem Element in A genau ein Bildelement in B gibt und umgekehrt (in Bezug auf f^{-1}). Beispielsweise gilt

$$f(A) = B \Rightarrow B \subset f(A) \Rightarrow (b \in B \Rightarrow b \in f(A)) \Rightarrow \exists a \in A : b = f(a). \tag{2.10}$$

Beispiel 2.6. Die Funktion f aus Beispiel 2.4 ist injektiv. Sie ist nicht surjektiv wegen (2.8d) und $1 \in B$. Wenn wir B umdefinieren zu $\{2, \ldots, 7\}$, wird f surjektiv.

Abbildungen können hintereinander ausgeführt (komponiert oder verkettet) werden.

Definition 2.7. Seien $f : A \to B'$ und $g : B \to C$ Abbildungen, $B' \subset B$. Dann heißt die Abbildung

$$g \circ f : A \to C, \quad a \mapsto (g \circ f)(a) = g(f(a)) \tag{2.11}$$

Komposition von f und g.

Eigenschaft 2.8. *Es seien $f : A \to B$ und $g : B \to C$ Abbildungen. Dann gilt:*

1. *Falls f und g surjektiv sind, ist es auch $g \circ f$.*
2. *Falls f und g injektiv sind, ist es auch $g \circ f$.*
3. *Falls f und g bijektiv sind, ist es auch $g \circ f$.*

Beweis. Sei $c \in C$ beliebig. Da g surjektiv ist, gibt es ein $b \in B$ mit $g(b) = c$. Da f surjektiv ist, gibt es ein $a \in A$ mit $f(a) = b$. Weil $(g \circ f)(a) = g(f(a)) = g(b) = c$, ist auch $g \circ f : A \to C$ surjektiv.

Seien nun $a, a' \in A$ mit $(g \circ f)(a) = (g \circ f)(a')$, das bedeutet $g(f(a)) = g(f(a'))$. Da g injektiv, folgt $f(a) = f(a')$, und weil f ebenfalls injektiv ist folgt $a = a'$, und wir haben die Injektivität von $g \circ f$ gezeigt.

Wenn f, g bijektiv sind, sind sie surjektiv und injektiv, und daher gilt dies nach den ersten beiden Aussagen auch für $g \circ f$. Diese Funktion ist daher gleichermaßen bijektiv.

Es muß nun noch argumentiert werden, daß die Aussage auch noch stimmt, wenn eine oder mehrere der Mengen A, B, C leer sind (?!). □

Satz 2.9. *Eine Abbildung $f : A \to B$ ist genau dann bijektiv, wenn es eine Abbildung $g : B \to A$ gibt mit $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$.*

Beweis. Wir gehen zunächst davon aus, daß f bijektiv ist. Dies impliziert Surjektivität und daher $\forall b \in B\, \exists a \in A : b = f(a)$. Da auch Injektivität gilt, kann es kein zweites a' mit $f(a) = b = f(a')$ geben, denn daraus würde schon folgen daß $a = a'$. Wir erklären nun die Funktion $g : B \to A$ durch $g(b) = a$. Durch Einsetzen folgt nun $f(g(b)) = b$. Weiter können wir für jedes $a \in A$ ein $b = f(a)$ berechnen und erhalten genauso $g(f(a)) = a$. Damit ist diese Richtung der Aussage komplett.

Wir setzen nun die Existenz der beschriebenen Abbildung g voraus. Um Surjektivität von f zu zeigen, nehmen wir ein beliebiges $b \in B$ und berechnen $a = g(b) \in A$. Wenn wir darauf f anwenden, erhalten wir $f(a) = f(g(b)) = (f \circ g)(b) = \mathrm{id}_B(b) = b$. Zum Zeigen der Injektivität von f nehmen wir an, daß $f(a) = f(a')$ für $a, a' \in A$. Daraus folgt direkt $a = \mathrm{id}_A(a) = g(f(a)) = g(f(a')) = \mathrm{id}_A(a') = a'$.

Wir haben nun beide Richtungen der Aussage gezeigt. □

Eigenschaft 2.10. *Es seien $f : A \to B$, $g : B \to C$ und $h : C \to D$ Abbildungen. Dann gilt Assoziativität:*

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{2.12}$$

Beweis. Übung. □

Definition 2.11. Eine Abbildung $f : A \to A$ kann mit sich selbst verkettet werden,

$$f^0 = \mathrm{id}_A, \qquad f^k = \underbrace{f \circ \ldots \circ f}_{k\text{-mal } f}. \tag{2.13}$$

Definition 2.12. Es seien A, B Mengen. Wir definieren deren Produktmenge als

$$A \times B = \{(a, b) : a \in A \wedge b \in B\}. \tag{2.14}$$

Wenn $A = B$, schreiben wir auch $A \times A = A^2$. Der Ausdruck (a, b) wird auch (geordnetes) Paar oder Zwei-Tupel genannt. Bei geschweiften Klammern (Mengen) ist die Reihenfolge der Elemente unerheblich, bei runden Klammern (Tupeln) ist sie entscheidend. In offensichtlichen Fällen ersetzen wir im Folgenden die Konjunktion $\wedge$ durch ein einfaches Komma. Binäre Operatoren sind also formal auf Produktmengen definiert.

Das Mengenprodukt ist nicht kommutativ. Wir können jedoch das Produkt von mehr als zwei Mengen (und die entsprechenden Tupel) bilden. (Ist dieses assoziativ?!)

2.2 Natürliche Zahlen

Bisher haben wir keine Zahlen eingeführt, sondern nur Symbole (außer in Beispiel 2.4). Selbst die einfachste Zahlenmenge, die natürlichen Zahlen, muß erst definiert werden. Wir suchen dazu nach einer möglichst reduzierten Liste von Eigenschaften. Die folgenden Axiome beruhen auf G. Peanos Arbeiten.

Definition 2.13. Die natürlichen Zahlen bilden eine Menge $\mathbb{N}_0$, in der ein Element $0 \in \mathbb{N}_0$ ausgezeichnet und eine Selbstabbildung $S : \mathbb{N}_0 \to \mathbb{N}_0$ (der successor oder Nachfolger) definiert ist, so daß

(SI) S ist injektiv,

(SN) $0 \notin S(\mathbb{N}_0)$,

(SV) Wenn eine Teilmenge $M \subset \mathbb{N}_0$ die Null enthält und durch S in sich abgebildet wird, also $S(M) \subset M$, dann ist $M = \mathbb{N}_0$.

Satz 2.14. *Die naiv eingeführten natürlichen Zahlen inklusive Null erfüllen alle obigen Bedingungen mit der Definition* $S(n) = n + 1$.

Beweis. (SN) gilt, da es keine natürliche Zahl n gibt mit $0 = n + 1$. Zu (SI) nehmen wir an, daß $S(n) = S(n')$, also $n + 1 = n' + 1$, was nach den entprechenden Rechenregeln bedeutet, daß $n = n'$. Zu (SV) ist nur zu zeigen daß $\mathbb{N}_0 \subset M$. Dazu überlegen wir, daß mit jedem Element aus M auch dessen Nachfolger in M liegt. Wenn wir bei Null anfangen, sukzessive die Nachfolger zu bilden (zu „zählen“), decken wir nach dem naiven Verständnis die gesamten natürlichen Zahlen ab. □

Wir können die natürlichen Zahlen ohne Null definieren als $\mathbb{N} = S(\mathbb{N}_0)$, die Peano-Null dieser Menge ist dann $1 = S(0)$ aus der ersten. Als Übung kann gezeigt werden, daß $\mathbb{N}$ ebenfalls die Peano-Axiome erfüllt.

Definition 2.15. Die Nachfolgefunktion erlaubt es uns, eine Größenrelation einzuführen,

$$(m \leq n) = \exists k \in \mathbb{N}_0 : n = S^k(m), \tag{2.15a}$$

$$(m < n) = (m \leq n) \wedge (m \neq n), \tag{2.15b}$$

und zum Beispiel die Addition natürlicher Zahlen,

$$m + 0 = m, \tag{2.16a}$$

$$m + S(n) = S(m + n). \tag{2.16b}$$

Satz 2.16. *Die dritte Bedingung (SV) führt auf das Prinzip der vollständigen Induktion: Wenn die Zahl* 0 *eine Eigenschaft* E *hat (Induktionsanfang) und für jede Zahl* n*, welche die Eigenschaft* E *hat, auch der Nachfolger* $S(n)$ *die Eigenschaft* E *hat (Induktionsschluß), dann haben alle natürlichen Zahlen die Eigenschaft* E*.*

Beweis. Wir kürzen „n hat Eigenschaft E“ ab als Aussage $E(n)$. Definiere die Menge

$$M = \{n \in \mathbb{N}_0 : E(n)\} \subset \mathbb{N}_0. \tag{2.17}$$

Wenn also $E(0)$, ist nach (2.17) $0 \in M$. Weiter folgt daraus $(E(n) \Rightarrow E(S(n))) \Rightarrow (n \in M \Rightarrow S(n) \in M)$. Damit sind die Voraussetzungen von (SV) erfüllt und wir erhalten $M = \mathbb{N}_0 \Rightarrow E(\mathbb{N}_0)$, die Induktionsaussage. □

Beispiel 2.17. Eine beliebte Anwendung der vollständigen Induktion ist die Addition aufsteigender Zahlen: Zeige, daß für alle $n \in \mathbb{N}_0$ folgende Aussage $E(n)$ gilt, definiert

$$E(n) = (2(0 + 1 + \ldots + n) = n(n + 1)). \tag{2.18}$$

Beweis. Der Induktionsanfang ist die Rechnung $2 \cdot 0 = 0 \cdot 1 \Rightarrow E(0)$. Im Induktionsschritt nehmen wir $E(n)$ an wie oben. Nun rechnen wir

$$\begin{aligned} 2(0 + 1 + \ldots + n + (n + 1)) &= 2(0 + 1 + \ldots + n) + 2(n + 1) \\ &\overset{E(n)}{=} n(n + 1) + 2(n + 1) = (n + 1)(n + 2) \Rightarrow E(n + 1). \end{aligned} \tag{2.19}$$

Nun folgt aus dem Induktionsprinzip $E(\mathbb{N}_0)$, also (2.18) für alle $n \geq 0$. □

(Wir müssten dazu allerdings zuerst alle Rechenregeln der Addition neu beweisen!)

Nun wissen wir immer noch nicht, was das Nullelement eigentlich ist (alle anderen sind definiert als $S(0)$, $S(S(0))$, etc.). Da wir nur mit den drei Eigenschaften weiterarbeiten, ist dies kein besonderes Hindernis. Trotzdem haben sich Generationen von Mathematikern gefragt, wie man die Definition von „Urelementen“ wie der Null vermeiden kann.

Ein Hinweis ist das Ersetzen von Eigenschaften durch Teilmengen wie in (2.17). Wenn zwei Eigenschaften äquivalent sind, sind die jeweiligen Teilmengen gleich, die Eigenschaften also nicht mehr zu unterscheiden. Wie können wir auch die restlichen Objekte der Mathematik als Mengen definieren?

Definition 2.18. Wir definieren die natürlichen Zahlen nach von Neumann über

$$0 = \emptyset, \quad 1 = \{\emptyset\}, \quad 2 = \{\emptyset, \{\emptyset\}\}, \quad \dots \quad \text{oder allgemein} \quad S(n) = n \cup \{n\}. \tag{2.20}$$

Diese Definition erfüllt bei Verwendung der modernen Mengenlehre die Peano-Axiome.

Übung: Schreibe die Mengen 3 und 4 aus. Beweise durch vollständige Induktion, daß jede Zahl n genau n Elemente enthält, und daß jede Zahl die Menge aller vorhergehenden Zahlen ist.

Geordnete Tupel traten bereits bei der Definiton der Produktmenge (2.14) auf.

Definition 2.19. Wir definieren mengentheoretisch

$$(a, b) = \{\{a\}, \{a, b\}\} \qquad \text{für Zweitupel und} \tag{2.21a}$$

$$(a, b, c) = ((a, b), c) \qquad \text{etc. für 3-, 4-, } \dots, n\text{-Tupel.} \tag{2.21b}$$

Eigenschaft 2.20. *Es gilt mit obiger Definition*

$$(a, b) = (a', b') \qquad \Leftrightarrow \qquad (a = a') \wedge (b = b'). \tag{2.22}$$

Übung: Wie steht es um die Wahrheit von $(a, b, c) = (a, (b, c))$?!

Beweis. In Vorwärtsrichtung folgt zunächst $\{\{a\}, \{a, b\}\} = \{\{a'\}, \{a', b'\}\}$. Nach (2.1c) wissen wir, daß jedes Element in der linken Menge auch in der rechten Menge enthalten ist. Für die einelementige Menge $\{a\}$ kann das nur die einelementige Menge $\{a'\}$ sein, also $a = a'$, und entsprechend folgt $b = b'$. (Zumindest für $a \neq b$. Ansonsten nutzen wir $\{a, a\} = \{a\}$ und es folgt $a = a' = b'$.)

Die Rückwärtsrichtung folgt durch Einsetzen. □

Gleichung (2.22) ist die einzige Eigenschaft von Tupeln, die wir im Weiteren benötigen. Mit der neuen Definition sind Produktmengen (2.14) nun Mengen von Mengen von Mengen. Eine binäre Operation läßt sich als Teilmenge einer Produktmenge schreiben, z.B. kleiner als

$$K = \{(m, n) : m, n \in \mathbb{N}_0, m < n\} \subset \mathbb{N}_0^2, \tag{2.23}$$

und $2 < 3$ bedeutet $(2, 3) \in K$. Mit der von-Neumannschen Definition 2.18 ist $m < n$ übrigens gleichbedeutend mit $m \in n$ (?!), und wir hätten uns (2.15) sparen können.

Als letztes möchten wir nun noch Funktionen als Mengen definieren.

Definition 2.21. Ganz ähnlich wie bisher definieren wir eine Funktion $f : A \to B$ durch ihren Graphen,

$$f = \{(a, f(a)) : a \in A\} \subset A \times B. \tag{2.24}$$

Hier ist zu jedem Objekt a höchstens ein b erlaubt mit $(a, b) \in f$.

Es ist also möglich, sämtliche Objekte, die wir bisher verwendet haben, als Mengen (von Mengen von Mengen...) zu definieren. Da wir mit der leeren Menge beginnen, benötigen wir keine weiteren mathematischen Objekte! Für den Rest unserer Arbeit ist diese Erkenntnis allerdings ohne Bedeutung: Sie funktioniert mit oder ohne die mengentheoretischen Definitionen gleich. Trotzdem wird die Mengentheorie in verschiedensten Gebieten der Mathematik immer wieder wichtig, um dort auftretende logische Konzepte sauber zu erkären.

3 Algebraische Strukturen

Die Zahlen, aber auch andere Objekte wie Polynome oder Funktionen lassen sich in Mengen zusammenfassen, für deren Elemente bestimmte Eigenschaften gelten. Einige wichtige stellen wir hier vor.

3.1 Gruppen

Wir formalisieren eine Struktur, die sich vielfach in der Mathematik wiederfindet.

Definition 3.1. Es sei A eine Menge und $\diamond$ eine binäre Operation auf A^2. Dann heißt $(A, \diamond)$ Gruppe, falls

$$\forall a, b, c \in A : (a \diamond b) \diamond c = a \diamond (b \diamond c), \tag{3.GA}$$

$$\exists e \in A \, \forall a \in A : a \diamond e = e \diamond a = a, \tag{3.GE}$$

$$\forall a \in A \, \exists a^{-1} \in A : a \diamond a^{-1} = a^{-1} \diamond a = e. \tag{3.GI}$$

Eine Gruppe heißt abelsch, wenn zusätzlich gilt, daß

$$\forall a, b \in A : a \diamond b = b \diamond a. \tag{3.GK}$$

Gruppen sind also per Definition assoziativ, es gibt ein sogenanntes neutrales oder Einselement e und zu jedem Gruppenelement a ein inverses Element a^{-1}.

Beispiel 3.2. Wir sind dem Gruppenbegriff bereits begegnet. Als Übung berechne zu den folgenden Beispielen, welche Gruppeneigenschaften (nicht) erfüllt sind und identifiziere Eins- und inverse Elemente.

- $(W, \wedge)$ und $(W, \vee)$ sind keine Gruppen.
- $(W, \oplus)$ ist eine abelsche Gruppe.
- Es sei M eine Menge und $A = \{f : M \to M, f \text{ bijektiv}\}$. Dann ist $(A, \circ)$ eine Gruppe. Übung: Gib ein Beispiel einer nichtabelschen Gruppe dieser Form an.
- Die natürlichen Zahlen $\mathbb{N}$, $\mathbb{N}_0$ sind bezüglich der Addition $+$ keine Gruppe. Nur die Null von $\mathbb{N}_0$ wäre geeignet als neutrales Element e.
- Die ganzen Zahlen $\mathbb{Z}$ sind eine abelsche Gruppe bezüglich der Addition mit neutralem Element 0. Das inverse Element zu $z \in \mathbb{Z}$ ist $z^{-1} = -z$. Die Subtraktion ist keine unabhängige Operation: Wir definieren $w - z = w + z^{-1} = w + (-z)$.
- Die reellen Zahlen $\mathbb{R}$ sind eine abelsche Gruppe bezüglich der Addition, geschrieben $(\mathbb{R}, +)$. Ohne die Null gilt dies auch bezüglich der Multiplikation, $(\mathbb{R} \setminus \{0\}, \cdot)$. Was ist das jeweilige neutrale Element?!

Satz 3.3. *Wir beweisen einige Eigenschaften für Gruppenelemente.*

1. *Das neutrale Element e ist eindeutig bestimmt.*
2. *Für jedes $a \in A$ ist sein Inverses a^{-1} eindeutig bestimmt.*
3. *Es gilt $e^{-1} = e$.*
4. *Es gilt $(a^{-1})^{-1} = a$.*
5. *Die Gleichung $a \diamond x = b$ hat die eindeutige Lösung $x = a^{-1} \diamond b$.*
6. *Wenn wir* (3.GI) *ersetzen durch die einseitige Formel $a \diamond a^{-1} = e$, folgt mit den anderen zwei Gleichungen* (3.GA), (3.GE) *wieder die beidseitige Formel* (3.GI).

Beweis. Zur ersten Aussage rechnen wir

$$e \overset{(e' \text{ neutral})}{=} e \diamond e' \overset{(e \text{ neutral})}{=} e'. \tag{3.2}$$

Wir beweisen nun die dritte Aussage. Durch Einsetzen von $a = e$ in (3.GE) erhalten wir $e \diamond e = e$. Dies ist aber gleichzeitig die Definitionsgleichung (3.GI) für e^{-1}, und die Aussage folgt.

Zur vierten Aussage setzen wir a^{-1} als Element a in (3.GI) ein, dies führt auf $a^{-1} \diamond (a^{-1})^{-1} = e$. Verknüpfen wir den Ausdruck von links mit a und nutzen die Assoziativität und die Eigenschaften der Eins und der Inversen, erhalten wir

$$(a \diamond a^{-1}) \diamond (a^{-1})^{-1} = a \diamond e \quad \Rightarrow \quad e \diamond (a^{-1})^{-1} = a \quad \Rightarrow \quad (a^{-1})^{-1} = a. \tag{3.3}$$

Beweise die restlichen Aussagen als Übung. □

3.2 Körper

Wir haben an den Beispielen der Konjunktion/Disjunktion und der reellen Zahlen gesehen, daß es sinnvoll sein kann, zwei verschiedene Operationen miteinander zu verknüpfen wie z.B. Addition und Multiplikation. Um die Addition invertieren zu können, definieren wir negative Zahlen und verwenden das Minuszeichen $-()$ anstelle der Schreibweise $()^{-1}$. Um die Multiplikation invertieren zu können, definieren wir gebrochene Zahlen: $z^{-1} = 1/z$ ist gültig, solange wir die Null (das neutrale Element der Addition) ausnehmen. Das neutrale Element der Multiplikation ist anschaulicherweise $e = 1$, daher auch die Schreibweise e für Einselement. Insgesamt erhalten wir den Zahlenkörper der rationalen Zahlen $\mathbb{Q}$, die Menge aller Brüche mit ganzen Zahlen im Zähler und natürlichen (ohne Null) im Nenner.

Definition 3.4. Ein Körper $\mathbb{K}$ ist definiert durch folgende Eigenschaften:

(KA) $(K, +)$ bildet eine abelsche Gruppe, deren neutrales Element wir mit 0 bezeichnen (Nullelement).

(KM) $(K \setminus \{0\}, \cdot)$ bildet eine abelsche Gruppe mit neutralem Element (Einselement) 1. Das Produkt $x \cdot y$ kürzen wir oft mit xy ab.

(KD) Das Distributivgesetz gilt:

$$(x + y)z = xz + yz. \tag{3.4}$$

Wir schreiben der Multiplikation eine stärkere Bindung zu als der Addition.

Für Körper müssen wir nur Eigenschaften beweisen, die noch nicht aus einer der beiden Gruppeneigenschaften (KA) oder (KM) folgen. Beispielsweise wissen wir dadurch schon, daß 0 und 1 eindeutig sind, und daß $0 \neq 1$ (?!).

Eigenschaft 3.5. *Für alle Elemente a, b eines Körpers gilt*

$$0 \cdot a = 0, \tag{3.5a}$$

$$ab = 0 \quad \Rightarrow \quad a = 0 \vee b = 0. \tag{3.5b}$$

Außerdem folgt für die Zahlen $m = 1 + \ldots + 1$, also m-fache Additionen der Körper-Eins,

$$\underbrace{a + \ldots + a}_{m\text{-mal}} = ma. \tag{3.6}$$

Beweis. Wir rechnen $0 \cdot a \overset{(3.\mathrm{GE})+}{=} (0+0)a \overset{(KD)}{=} 0 \cdot a + 0 \cdot a$. Nun addieren wir auf beiden Seiten $-(0 \cdot a)$ und erhalten die erste Aussage.

Wenn a oder b Null ist, gilt die zweite Behauptung bereits. Wir nehmen nun o.B.d.A. an, daß $ab = 0$ und $a \neq 0$ (Kommutativität). Dann können wir nach Satz 3.3/5. folgern, daß $b = a^{-1} \cdot 0 \overset{(3.\mathrm{GK})\cdot}{=} 0 \cdot a^{-1} \overset{(3.5a)}{=} 0$.

Strenggenommen müßten wir jetzt zuerst das Distributivgesetz für mehr als zwei Summanden durch vollständige Induktion beweisen. Damit folgt dann $a + \ldots + a = 1 \cdot a + \ldots 1 \cdot a = (1 + \ldots + 1)a = ma$. □

Mit den üblichen Rechenregeln sehen wir, daß $\mathbb{Q}$ ein Körper ist, und die Menge der reellen Zahlen $\mathbb{R}$ ebenfalls. Wo liegt nun der Unterschied zwischen beiden?

Schon die alten Griechen haben bewiesen, daß das Wurzelziehen in den rationalen Zahlen nicht immer möglich ist.

Satz 3.6. *Die Gleichung $x^2 = 2$ hat in $\mathbb{Q}$ keine Lösung.*

Beweis. Wir nehmen an, daß $x = p/q$ mit teilerfremden $p, q \in \mathbb{N}$. Es reicht aus, positive Zahlen zu betrachten, da jede Kombination von Negierungen dasselbe Quadrat hat und nach (3.5a) Null im Zähler niemals im Quadrat 2 ergeben kann. Wir folgern $x^2 = p^2/q^2 = 2 \Rightarrow p^2 = 2q^2 \Rightarrow p^2$ ist gerade. Daher muß auch p gerade sein (das Quadrat einer ungeraden Zahl ist ungerade), und es gibt eine natürliche Zahl r mit $p = 2r$. Daraus folgt $(2r)^2 = 2q^2 \Rightarrow q^2 = 2r^2$ und q ist ebenfalls gerade, also haben p und q denselben Teiler 2, was ein Widerspruch zur oben angenommenen Teilerfremdheit ist. □

Etwas allgemeiner läßt sich zeigen, daß keine Primzahl außer 1 eine Wurzel in $\mathbb{Q}$ hat. Wir können jedoch eine Schachtelung aus rationalen Zahlen angeben, die die Wurzel beliebig genau eingrenzt, zum Beispiel

$$\begin{aligned} 1^2 < 3 < 2^2, \\ 1.7^2 < 3 < 1.8^2, \\ 1.73^2 < 3 < 1.74^2, \text{ etc.} \end{aligned} \tag{3.7}$$

Wir definieren den Raum der reellen Zahlen $\mathbb{R}$ daher gerade dadurch, daß das Ziel solcher Schachtelungen mit enthalten ist. Diese Eigenschaft nennt sich intuitiv Vollständigkeit und ist (wie bei $\mathbb{Q}$ gesehen) keineswegs selbstverständlich.

4 Lineare Algebra

Wenn wir von linearen Zusammenhängen sprechen, dann gilt eine Gesetzmäßigkeit unabhängig von Addition oder Skalierung innerhalb eines Systems. Beispielsweise kann man annehmen, daß drei Arbeiter in derselben Zeit dreimal so viel Arbeit erledigen wie einer alleine, oder daß die Gesamtfuttermenge, die zwei Elefanten für sich alleine vertilgen dieselbe ist wie wenn sie im selben Käfig vom selben Futterberg fressen.

Der Zusammenhang wird nichtlinear, wenn die Arbeiter anfangen, gemeinsam zu trödeln oder die Elefanten sich um das Futter streiten. Solche Vorgänge wären uns an dieser Stelle zu kompliziert.

4.1 Lineare Gleichungssysteme

Eine wesentliche Anwendung der linearen Algebra ist das Lösen linearer Gleichungssysteme (LGS).

Beispiel 4.1. Das System besteht aus einer Gleichung

$$3x = 6. \tag{4.1}$$

Dies ist eine Kurzschreibweise für „x löst die Gleichung $3x = 6$.“ Diese Aussage ist zunächst unabhängig von der Existenz eines solchen x. Hier können wir $x = 2$ raten und durch Einsetzen feststellen, daß dieses x die Gleichung löst.

Weiterhin untersuchen wir die Eindeutigkeit der Lösung: Wir nehmen zwei Lösungen x, x' an und folgern von oben nach unten

$$3x = 6 \wedge 3x' = 6 \tag{4.2a}$$

$$\begin{aligned} 3x - 3x' &= 6 - 6 \\ 3(x - x') &= 0 \\ x - x' &= 0 \\ x &= x'. \end{aligned} \tag{4.2b}$$

Wir haben einseitig gefolgert, also aus „x und x' lösen (4.1)“ geschlossen, daß $x = x'$.

Wie können wir die Lösung systematischer bestimmen? Wenn wir (4.1) mit $1/3$ multiplizieren, erhalten wir die Aussage, daß wenn eine Lösung existiert, sie zwangsläufig gleich 2 ist und damit eindeutig. Die obige Probe liefert dann den Nachweis der Existenz.

Beispiel 4.2. Das System habe zwei Unbekannte x und y,

$$x + y = 5, \tag{4.3a}$$

$$2x - y = 1. \tag{4.3b}$$

Wir folgern durch Addition der beiden Gleichungen (4.1) und damit $x = 2$ wie oben beschrieben. Einsetzen hiervon in die untere Gleichung ergibt $y = 3$. Die Lösung $(x, y) = (2.3)$ ist daher eindeutig. Ihre Existenz läßt sich durch die Probe nachweisen.

Beispiel 4.3. Das System enthalte drei Unbekannte x, y und z,

$$x - y + z = 1, \tag{4.4a}$$

$$-x + y - z = 0. \tag{4.4b}$$

Angenommen, es gibt Zahlen, die (4.4) lösen, führt die Addition beider Gleichungen auf $0 = 1$, einen Widerspruch. Das System ist also unlösbar. Wenn wir andererseits die letzte Zeile mit 0 multiplizieren, erhalten wir eine Familie von Lösungen, die sich z.B. als x, y beliebig und $z = 1 - x + y$ schreiben läßt. Diese sind jedoch allesamt keine Lösung des Gesamtsystems (4.4)!

Was ist hier passiert? Wir haben korrekt gefolgert, daß $a \Rightarrow b$ mit

$$a = (x,\, y \text{ und } z \text{ lösen } (4.4))\,, \tag{4.5a}$$
$$b = (\text{gegeben } x \text{ und } y \text{ beliebig, dann ist } z = 1 - x + y)\,. \tag{4.5b}$$

Nun ist $\neg a$ und wir wissen nach (1.13b), daß daraus keine Aussage über den Wahrheitsgehalt von b abgeleitet werden kann. Weiterhin folgt aus $a \Rightarrow b$ nicht $b \Rightarrow a$, so daß b eben keine Garantie für a gibt.

Definition 4.4. Wir möchten unsere Behandlung von LGS nun systematisieren. Ein lineares Gleichungssystem mit M Gleichungen und N Unbekannten $x_1, \ldots, x_N$, auch genannt $M \times N$ System, hat die Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1N}x_N &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2N}x_N &= b_2, \\ &\vdots \\ a_{M1}x_1 + a_{M2}x_2 + \ldots + a_{MN}x_N &= b_M. \end{aligned} \tag{4.6}$$

Die gegebenen Zahlen a_{ij} heißen Koeffizienten, wobei i die Zeilen- und j die Spaltennummer ist. Die b_i heißen die rechten Seiten. Das LGS heißt homogen, wenn alle $b_i = 0$. Die Lösungsmenge enthält alle N-Tupel $(x_1, \ldots, x_N)$, die (4.6) erfüllen.

Um LGS zu lösen, formen wir sie in äquivalente Systeme einfacherer Gestalt um, aus denen wir die Lösung direkt ablesen können.

Satz 4.5. *Die folgenden elementaren Umformungen ändern die Lösungsmenge nicht:*

(EV) Vertauschen zweier Gleichungen,

(EM) Multiplikation einer der Gleichungen mit einer Zahl $\neq 0$,

(EA) Addition einer mit einer beliebigen Zahl multiplizierten Gleichung zu einer anderen.

Wir sagen, sie überführen das LGS in ein dazu äquivalentes.

Beweis. Die Vertauschungsregel (EV) folgt aus der Kommutativität der Konjunktion. Die Multiplikation (EM) von Gleichung i mit $c \neq 0$ führt auf die modifizierte Gleichung

$$a'_{i1}x_1 + a'_{i2}x_2 + \ldots + a'_{iN}x_N = b'_i \qquad \text{mit } a'_{ij} = ca_{ij} \text{ für alle } 1 \leq j \leq N,\ b'_i = cb_i. \tag{4.7}$$

Jede Lösung von (4.6) löst auch das LGS mit ersetzter i-ter Gleichung (4.7). Wenn wir nun eine Lösung des modifizierten Systems annehmen, führen wir (EM) mit $c' = 1/c$ durch, was existiert wegen $c \neq 0$, und erhalten das erste System zurück.

Analog gehen wir für (EA) vor: Wenn die erste Multiplikation mit c erfolgt, können wir die Situation invertieren, indem wir die Regel nochmal anwenden, diesmal mit $-c$. □

Mit dem Wissen um elementare Umformungen können wir Existenz und Eindeutigkeit der Lösung des ersten Beispiels 4.1 in einem zeigen:

$$3x = 6 \quad \Leftrightarrow \quad x = 2 \qquad \text{da } \frac{1}{3} \neq 0 \text{ (EM)}. \tag{4.8}$$

Wir untersuchen nun einige allgemeine Aussagen zur Lösbarkeit.

Satz 4.6. *Ein homogenes LGS mit mehr Unbekannten als Gleichungen, $M < N$, hat beliebig viele Lösungen (hier: mindestens so viele wie Elemente in $\mathbb{K}$).*

Beweis. Wir führen eine vollständige Induktion über M durch. Für der Induktionsanfang $M = 1$, also eine einzige Gleichung, unterscheiden wir zwei Fälle:

1. Alle Koeffizienten sind Null. Dann können wir $x_1, \ldots, x_N$ beliebig wählen.
2. Sei o.B.d.A. $a_{1N} \neq 0$. Wir wählen $x_1, \ldots, x_{N-1}$ beliebig und bestimmen $x_N = -(a_{11}x_1 + \ldots + a_{1,N-1}x_{N-1})/a_{1N}$. Nach (EM) ergibt dies eine Lösung.

Zum Induktionsschritt von $M < N$ auf $M' = M + 1 < N'$ unterscheiden wir ganz ähnlich zwei Fälle.

1. Alle Koeffizienten der $(M+1)$-ten (letzten) Gleichung sind Null. Nach Induktionsvoraussetzung für M, $N = N' > M' > M$ gibt es beliebig viele Lösungen der ersten M Gleichungen, die dann auch die letzte erfüllen.
2. Sei o.B.d.A. $a_{M'N'} \neq 0$. Dann folgern wir mit $N = N' - 1$

$$x_{N'} = -\left(a_{M'1}x_1 + \ldots a_{M'N}x_N\right)/a_{M'N'}. \tag{4.9}$$

Wir setzen dies in die ersten M Gleichungen ein und erhalten das $M \times N$ System

$$\sum_{j=1}^{N} \left(a_{ij} - \frac{a_{M'j}}{a_{M'N'}} a_{iN'}\right) x_j = 0, \quad i = 1, \ldots, M. \tag{4.10}$$

Diesen hat nach Induktionsvoraussetzung beliebig viele Lösungen $x_1, \ldots, x_N$, die wir jeweils mit (4.9) um x'_N ergänzen können. Die Probe zur letzten Gleichung verläuft nun wie oben über (EM). Für die ersten M setzen wir (4.9) ins ursprüngliche System ein und erhalten wieder (4.10), was die angegebene Lösung hat.

□

4.2 Vektorräume

Nachdem wir einen Zahlenbegriff und passende Rechenregeln kennengelernt haben, fragen wir uns, ob wir auch kompliziertere Dinge addieren und in der Größe verändern können. Beispiele sind Pfeile in der Ebene, die wir aneinanderhängen und strecken, oder Geschwindigkeiten, wenn aus einem fahrenden Auto ein Ball geworfen wird.

Ein relativ einfaches Konstrukt sind die geordneten n-Tupel, vgl. auch (2.21).

Definition 4.7. Sei $\mathbb{K}$ ein Körper und $n > 0$. Wir schreiben die n-Tupel mit Elementen aus $x_i \in \mathbb{K}$ als $x = (x_1, \ldots, x_n) \in \mathbb{K} \times \ldots \times \mathbb{K} = \mathbb{K}^n$. Wir definieren zwei Rechenregeln durch elementweise Operationen in $\mathbb{K}$,

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n), \tag{4.11a}$$

$$\lambda \cdot (x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n), \quad \lambda \in \mathbb{K}. \tag{4.11b}$$

Andere Objekte, die wir addieren und skalieren können, sind Funktionen.

Definition 4.8. Seien f, g Funktionen $\mathbb{K}^m \to \mathbb{K}^n$. Wir erklären ihre Addition und Skalierung zu einer jeweils neuen Funktion punktweise und führen sie dadurch erst auf $\mathbb{K}^n$ und mit (4.11) weiter auf den Körper $\mathbb{K}$ zurück,

$$(f+g)(x) = f(x) + g(x), \tag{4.12a}$$
$$(\lambda \cdot f)(x) = \lambda \cdot f(x), \quad \lambda \in \mathbb{K}. \tag{4.12b}$$

Es gibt viele Beispiele für Funktionenmengen, bei denen die Summe zweier Elemente und jedes Vielfache wieder zur Menge gehören (z.B. die stetigen oder die differenzierbaren Funktionen).

Beispiel 4.9. Die Polynome von höchstens Grad $L \geq 0$ über einem Körper $\mathbb{K}$ sind die Funktionen $p : \mathbb{K} \to \mathbb{K}$,

$$x \mapsto p(x) = \sum_{i=0}^{L} a_i x^i. \tag{4.13}$$

Hier sind die Koeffizienten $a_i \in \mathbb{K}$ beliebig. Die Potenz x^i wird erklärt als i-fache Multiplikation, vgl. (3.6). Da ein Polynom als Funktion nur Elemente von $\mathbb{K}$ miteinander verknüpft, gelten die Eigenschaften (4.12). Wir wissen außerdem, daß $f+g$ wieder ein Polynom vom Grad höchstens L ist.

Wir fassen nun die bisher angeführten Punkte zusammen.

Definition 4.10. Sei $\mathbb{K}$ ein fest gewählter Körper. Seine Elemente werden auch Skalare genannt. Eine Menge V heißt $\mathbb{K}$-Vektorraum oder Vektorraum über $\mathbb{K}$, wenn für sie zwei Operationen definiert sind, eine Addition und eine Skalarmultiplikation,

$$+ : V \times V \to V, \quad (u, v) \mapsto u + v \text{ und} \tag{4.14a}$$
$$\cdot : \mathbb{K} \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v, \tag{4.14b}$$

so daß folgende Regeln gelten: $(V, +)$ ist eine kommutative Gruppe und

$$(\lambda + \mu)u = \lambda u + \mu u, \tag{4.15a}$$
$$(\lambda\mu)u = \lambda(\mu u), \tag{4.15b}$$
$$1 \cdot u = u, \tag{4.15c}$$
$$\lambda(u + v) = \lambda u + \lambda v. \tag{4.15d}$$

Die Elemente von V heißen auch Vektoren. Das Neutralelement der Addition wird als 0 geschrieben, obwohl es i.A. verschieden ist von der Körper-Null, und Nullvektor genannt.

Beispiel 4.11. Wir kennen bereits einige Vektorräume.

- Die Körper $\mathbb{Q}$ und $\mathbb{R}$ sind Vektorräume über sich selbst, jeweils mit derselben Null.
- Die n-Tupel nach Definition 4.7 bilden den $\mathbb{K}$-Vektorraum $\mathbb{K}^n$ mit $0 = (0, \ldots, 0)$.
- Eine Menge von Funktionen M wie in Definition 4.8 bildet einen Vektorraum, wenn die Summe von je zwei Funktionen und das Ergebnis jeder Skalarmultiplikation wieder eine Funktion in M sind.

  Übung: Definiere eine Menge von Funktionen, die kein Vektorraum ist!

- Die Polynome über $\mathbb{K}$ vom Grad höchstens L bilden einen $\mathbb{K}$-Vektorraum.
- Wenn eine einelementige Menge $\{\epsilon\}$ ein Vektorraum sein soll, gibt es nur eine Möglichkeit, die Grundoperationen zu definieren: $\epsilon + \epsilon = \epsilon$, $\lambda \cdot \epsilon = \epsilon$, mit Nullelement ϵ. Daraus folgen dann alle weiteren Vektorraumaxiome (?!). Dieser Raum wird als Nullraum bezeichnet.

Satz 4.12. *Für einen $\mathbb{K}$-Vektorraum V, $\lambda \in \mathbb{K}$ und $u \in V$ gelten die Regeln*

$$0 \cdot u = 0, \tag{4.16a}$$
$$\lambda \cdot 0 = 0, \tag{4.16b}$$
$$\lambda \cdot u = 0 \quad \Rightarrow \quad \lambda = 0 \vee u = 0, \tag{4.16c}$$
$$(-\lambda)u = \lambda(-u) = -(\lambda u), \quad \textit{geschrieben als } -\lambda u. \tag{4.16d}$$

Beweis. Setzen wir in (4.15a) $\lambda = \mu = 0$, so folgt $0 \cdot u \overset{(3.\text{GI})}{=} (0+0) \cdot u \overset{(4.15\text{a})}{=} 0 \cdot u + 0 \cdot u$ und nach Subtraktion von $0 \cdot u$ die erste Behauptung. Die zweite folgt analog aus (4.15d). Zur dritten Aussage nehmen wir das Gegenteil an, nämlich daß $\lambda \cdot u = 0$ obwohl $\lambda \neq 0$ und $u \neq 0$. Multiplizieren wir von links mit λ^{-1}, folgt

$$0 \overset{(4.16\text{b})}{=} \lambda^{-1} \cdot 0 = \lambda^{-1}(\lambda u) \overset{(4.15\text{b})}{=} (\lambda^{-1}\lambda)u \overset{(3.\text{GI})}{=} 1 \cdot u \overset{(4.15\text{c})}{=} u, \tag{4.17}$$

ein Widerspruch zur Annahme. Die letzte Regel folgt durch Einsetzen von $\mu = -\lambda$ bzw. $v = -u$ in (4.15) (Übung!). □

Durch vollständige Induktion folgen analoge Behauptungen für Kombinationen beliebig (aber endlich) vieler Terme.

4.3 Lineare Unabhängigkeit

Ein Vektorraum wird auch linearer Raum genannt. Diese Bezeichnung kommt daher, daß wir Vektoren linear kombinieren, das heißt anschaulich, Pfeile in der Länge verändern und aneinanderhängen können. Wir versuchen, die minimale Anzahl von Pfeilen zu finden, aus denen sich alle anderen kombinieren lassen. Das muß den Nullvektor mit einschließen: Mit Vektoren $a_j \in V$, $j = 1, \ldots, M$, untersuchen wir die Gleichung

$$\lambda_1 a_1 + \ldots + \lambda_M a_M = 0. \tag{4.18}$$

Definition 4.13. Ein Vektorsystem $(a_j)_{j=1}^M$ in V heißt

1. linear abhängig, falls Elemente $\lambda_j \in \mathbb{K}$ existieren, die (4.18) erfüllen und nicht alle gleich 0 sind (dies nennt man auch eine nichttriviale Linearkombination),
2. linear unabhängig andernfalls, d.h. falls aus (4.18) stets folgt $\lambda_1 = \ldots = \lambda_M = 0$.

Folgerung 4.14. *Das Vektorsystem $a_1 = 0$ ist immer linear abhängig. Die lineare (Un-) Abhängigkeit eines Systems gilt unabhängig von der Reihenfolge seiner Vektoren.*

Beweis. Wir verwenden die Körpereins als nicht-null-Koeffizienten, $1 \cdot a_1 = 1 \cdot 0 \overset{(4.16\text{b})}{=} 0$. Die zweite Aussage folgt aus der Kommutativität der Addition. □

Beispiel 4.15. In $\mathbb{K}^n$ definieren wir die Einheitsvektoren

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0), \\ e_2 &= (0, 1, 0, \ldots), \\ &\ldots \\ e_n &= (0, \ldots, 0, 1). \end{aligned} \tag{4.19}$$

Diese sind nach den Rechenvorschriften (4.11) linear unabhängig. Die Hinzunahme jedes beliebigen weiteren Vektors würde dieses System linear abhängig machen.

Beispiel 4.16. Finde zu den folgenden Vektoren im $\mathbb{R}^3$ jeweils einen Vektor a_3, der das System linear abhängig bzw. linear unabhängig macht!

$$a_1 = (4, 2, -1), \quad a_2 = (0, 1, 3). \tag{4.20}$$

Satz 4.17. *Die lineare Abhängigkeit eines Vektorsystems bleibt bei Hinzunahme weiterer Vektoren erhalten: Ist $a_1, \ldots, a_M$ linear abhängig, dann ist es auch $a_1, \ldots, a_M, a_{M+1}, \ldots, a_N$ für $M < N$.*

Die lineare Unabhängigkeit eines Vektorsystems bleibt bei Wegnahme von Vektoren erhalten: Ist $a_1, \ldots, a_M$ linear unabhängig, dann ist es auch $a_1, \ldots, a_L$ mit $L < M$.

Beweis. Nach der ersten Voraussetzung existieren $\lambda_1, \ldots, \lambda_M$, die nicht alle Null sind und (4.18) erfüllen. Mit den neuen Koeffizienten $\lambda_{M+1} = \ldots \lambda_N = 0$ gilt $\sum_{j=1}^{N} \lambda_j a_j = 0$.

Wäre in der zweiten Aussage nun das verkürzte System linear abhängig, so gälte dies nach der ersten auch für das ursprüngliche System, ein Widerspruch. □

Nach Folgerung 4.14 kommt es nicht darauf an, an welcher Stelle im System wir neue Vektoren einfügen oder enthaltene herausnehmen.

Satz 4.18. *Ein Vektorsystem $a_j \in V$ ist genau dann linear unabhängig, wenn jede Linearkombination von $a_1, \ldots, a_M$ eindeutig bestimmte Koeffizienten hat, wenn also*

$$\sum_{j=1}^{M} \beta_j a_j = \sum_{j=1}^{M} \gamma_j a_j \quad \Rightarrow \quad \beta_j = \gamma_j \text{ für alle } j = 1, \ldots, M. \tag{4.21}$$

Beweis. Setzen wir lineare Unabhängigkeit voraus, so folgt aus der linken Formel nach $\sum_{j=1}^{M} (\beta_j - \gamma_j) a_j = 0$ direkt $\beta_j - \gamma_j = 0$.

Wenn umgekehrt die Koeffizienten eindeutig sind, nehmen wir (4.18) an und folgern aus $\sum_{j=1}^{M} \lambda_j a_j = 0 = \sum_{j=1}^{M} 0 \cdot a_j$ daß $\lambda_j = 0$ für alle j. □

Satz 4.19. *Das Vektorsystem $a_1, \ldots, a_M, a_{M+1}$ in V sei linear abhängig, das Teilsystem ohne a_{M+1} jedoch linear unabhängig. Dann ist a_{M+1} eine Linearkombination von $a_1, \ldots, a_M$.*

Beweis. Nach Voraussetzung existieren Skalare $\lambda_1, \ldots, \lambda_{M+1}$, die nicht alle 0 sind, so daß $\sum_{j=1}^{M} \lambda_j a_j + \lambda_{M+1} a_{M+1} = 0$. Wäre nun $\lambda_{M+1} = 0$, so folgt (4.18) und wir schließen aus der linearen Unabhängigkeit der ersten M Vektoren $\lambda_j = 0$, einen Widerspruch. Da deswegen $\lambda_{M+1} \neq 0$, errechnen wir die Linearkombination $a_{M+1} = \sum_{j=1}^{M} (-\lambda_j / \lambda_{M+1}) \cdot a_j$. □

Satz 4.20. *Gegeben seien zwei Vektorsysteme $a_1, \ldots, a_M$ und $b_1, \ldots, b_{M+1}$ in V. Jedes b_i sei eine Linearkombination der a_j. Dann ist (b_i) linear abhängig.*

Beweis. Wir wissen nach Voraussetzung, daß $M(M+1)$ Koeffizienten c_{ij} existieren mit $b_i = \sum_{j=1}^{M} c_{ij} a_j$ für alle $1 \leq i \leq M+1$. Untersuchen wir Linearkombinationen der Form

$$\sum_{i=1}^{M+1} \mu_i b_i = \sum_{i=1}^{M+1} \sum_{j=1}^{M} \mu_i c_{ij} a_j = \sum_{j=1}^{M} \left(\sum_{i=1}^{M+1} \mu_i c_{ij} \right) a_j = 0. \tag{4.22}$$

Wir wissen aus Satz 4.6, daß sich die M Gleichungen $\sum_{i=1}^{M+1} c_{ij} \mu_i = 0$ nichttrivial lösen lassen. Jede solche Lösung liefert nach (4.22) eine Null-Kombination der b_i. □

4.4 Dimension und Basis

Wenn ein Vektorraum nicht der Nullraum ist, so enthält er mindestens ein Element zusätzlich zur (eindeutigen) Null. Dieses Element ist ein linear unabhängiges System der Länge eins aufgrund von Eigenschaft (4.16c). Wenn die Länge aller möglichen linear unabhängigen Systeme nach oben beschränkt ist, bezeichnen wir diese maximale Länge als Dimension des Vektorraums.

Definition 4.21. Der Vektorraum V hat die endliche Dimension M, wenn gilt:

1. V enthält ein linear unabhängiges Vektorsystem der Länge M.
2. Jedes Vektorsystem in V der Länge $M+1$ ist linear abhängig.

Man schreibt auch

$$\dim V = M < \infty. \tag{4.23}$$

Für den Nullraum definieren wir gesondert ein leeres Vektorsystem und $M = 0$.

Ist die zweite Aussage erfüllt, so ist nach Satz 4.17 jedes Vektorsystem mit Länge größer $M+1$ erst recht linear abhängig.

Satz 4.22. *Sei V von der Dimension M und $a_1, \ldots, a_M$ ein linear unabhängiges Vektorsystem in V. Dann kann jeder Vektor $u \in V$ als Linearkombination der a_j mit eindeutig bestimmten Koeffizienten $\lambda_j \in \mathbb{K}$ geschrieben werden.*

Beweis. Per definitionem ist das Vektorsystem $(a_1, \ldots, a_M, u)$ linear abhängig, also folgt nach Satz 4.19, daß u eine Linearkombination der a_j ist. Nach Satz 4.18 sind deren Koeffizienten eindeutig bestimmt. □

Satz 4.23. *Sei $a_1, \ldots, a_M$ ein Vektorsystem in V mit der Eigenschaft, daß jeder Vektor in V als deren Linearkombination mit eindeutigen Koeffizienten dargestellt werden kann. Dann sind die a_j linear unabhängig, und V hat die Dimension M.*

Beweis. Die lineare Unabhängigkeit folgt aus Satz 4.18. Da je $M+1$ Vektoren aus V als Linearkombinationen von $a_1, \ldots, a_M$ geschrieben werden können, also nach Satz 4.19 linear abhängig sind, ist auch die zweite Eigenschaft von Definition 4.21 bestätigt. □

Definition 4.24. Ein Vektorsystem $a_1, \ldots, a_M$ in V ist eine Basis von V, wenn jeder Vektor $u \in V$ als Linearkombination der a_j,

$$u = \sum_{j=1}^{M} \lambda_j a_j \tag{4.24}$$

mit eindeutig bestimmten Koeffizienten λ_j, dargestellt werden kann. In diesem Fall heißt λ_j die j-te Koordinate von u bezüglich (oder in) der Basis (a_j).

Die gerade eingeführten Koordinaten hängen von der Wahl der Basis ab (man denke alleine an eine Umordnung oder Skalierung der Basisvektoren). Nur bezüglich der Einheitsbasis (4.19) des $\mathbb{K}^n$ sind die Elemente eines Tupel-Vektors identisch mit seinen Koordinaten! Koordinaten sind in der Praxis unumgänglich, wohingegen mathematische Darstellungen allgemeiner sind, wenn sie nicht auf Koordinaten zurückgreifen.

Definition 4.25. Schreiben wir Koordinaten $\lambda = (\lambda_i)$ als Elemente eines Tupels hintereinander, so nennen wir dies einen Zeilenvektor. Schreiben wir seine Elemente untereinander,

$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_M \end{pmatrix}, \tag{4.25}$$

sprechen wir von einem Spaltenvektor. Analog zu Tupeln bilden die Spaltenvektoren fester Länge einen $\mathbb{K}$-Vektorraum, in dem wir die Einheitsbasis gleichfalls als Spalten schreiben.

Satz 4.26. *Ein Vektorraum V hat genau dann die endliche Dimension M, wenn in V eine Basis der Länge M existiert. In diesem Fall hat jede Basis die Länge M.*

Beweis. Die Aussage folgt aus den obigen Sätzen, ebenso wie die Tatsache, daß die Dimension eindeutig ist. □

Beispiel 4.27. Nicht jeder Vektorraum ist endlich dimensional: Sei M eine nichtleere Menge und V der Vektorraum aller Abbildungen $f : M \to \mathbb{K}$. Definiere für jedes $p' \in M$ ein $f_{p'} \in V$ durch

$$f_{p'}(p) = \begin{cases} 1 & \text{für } p = p' \\ 0 & \text{für } p \neq p'. \end{cases} \tag{4.26}$$

Es gibt so viele linear unabhängige Elemente in V wie Elemente in M, diese Anzahl kann unbegrenzt sein (Beweis: Übung).

4.5 Lineare Abbildungen

Wir kommen nun zu dem Konzept, das der linearen Algebra ihren Namen gibt.

Definition 4.28. Seien V und W zwei $\mathbb{K}$-Vektorräume. Eine Abbildung $\mathcal{A} : V \to W$ heißt linear, wenn für alle $u, v \in V$ und $\lambda \in \mathbb{K}$ gilt:

$$\mathcal{A}(u + v) = \mathcal{A}(u) + \mathcal{A}(v), \tag{4.27a}$$

$$\mathcal{A}(\lambda \cdot u) = \lambda \cdot \mathcal{A}(u). \tag{4.27b}$$

Man nennt V den Definitionsraum und W den Zielraum von $\mathcal{A}$. Die zweite Eigenschaft nennt man Homogenität von $\mathcal{A}$. Im Falle $W = \mathbb{K}$ heißt $\mathcal{A}$ auch Linearform oder lineares Funktional über V. Äquivalent könnte man auch die Formel nutzen

$$\mathcal{A}(\lambda u + \mu v) = \lambda \mathcal{A}(u) + \mu \mathcal{A}(v). \tag{4.28}$$

Eigenschaft 4.29. *Für lineare Abbildungen $\mathcal{A}$ gilt*

$$\mathcal{A}(0) = 0, \qquad \mathcal{A}(-u) = -\mathcal{A}(u). \tag{4.29}$$

Außerdem ist jede Verkettung zweier linearer Abbildungen linear.

Beweis. Wir setzen in (4.27b) nacheinander $\lambda = 0$, $\lambda = -1$. (Verkettung: Übung.) □

Satz 4.30. *Definieren wir die Menge*

$$\operatorname{Kern} \mathcal{A} = \{u \in V : \mathcal{A}(u) = 0\}, \tag{4.30}$$

so ist $\operatorname{Kern} \mathcal{A} \subset V$ *wieder ein Vektorraum mit derselben Null wie* $\mathcal{A}$ *(ein Untervektorraum). Es gilt weiterhin*

$$\operatorname{Kern} \mathcal{A} = \{0\} \quad \Leftrightarrow \quad \mathcal{A} \text{ ist injektiv.} \tag{4.31}$$

Beweis. Wir zeigen die Abgeschlossenheit von $\operatorname{Kern} \mathcal{A}$ unter den Vektorraumoperationen. Seien dazu $u, v \in \operatorname{Kern} \mathcal{A}$ mit $\mathcal{A}(u) = \mathcal{A}(v) = 0$. Nach (4.28) folgt $\mathcal{A}(\lambda u + \mu v) = \lambda \mathcal{A}(u) + \mu \mathcal{A}(v) = \lambda \cdot 0 + \mu \cdot 0 = 0$, also $\lambda u + \mu v \in \operatorname{Kern} \mathcal{A}$. (Injektivitätsaussage: Übung.) □

Satz 4.31. *Je zwei Vektorräume* U, V *gleicher Dimension* $M < \infty$ *sind zueinander isomorph, d.h. es gibt eine bijektive lineare Abbildung* $f : U \to V$*. Die Umkehrabbildung* g*, definiert durch Satz 2.9, ist ebenfalls linear.*

Beweis. Übung. □

4.6 Matrixdarstellung linearer Abbildungen

Wir beschäftigen uns nun mit tabellarischen Darstellungen von linearen Abbildungen zwischen Vektorräumen. Diese Tabellen nennen wir Matrizen. Wir werden sehen, daß sich über den Matrizen eine multiplikative Struktur aufbauen läßt.

Seien U, V Vektorräume und $\mathcal{B} : U \to V$ eine lineare Abbildung. Gegeben die Basen $(u_k)_{k=1}^N$ von U und $(v_j)_{j=1}^M$ von V, so können wir MN Koeffizienten identifizieren, die die Abbildung $\mathcal{B}$ eindeutig festlegen.

Wir wissen, daß jeder Vektor $u \in U$ die folgende Form hat,

$$u = \sum_{k=1}^{N} \nu_k u_k \tag{4.32}$$

mit eindeutigen Koeffizienten ν_k. Das Bild jedes Basisvektors von U liegt in V und hat deswegen eine Darstellung in dessen Basis mit eindeutigen Koeffizienten b_{jk},

$$\mathcal{B}u_k = \sum_{j=1}^{M} b_{jk} v_j \tag{4.33}$$

Wir kombinieren weiter unter Ausnutzung der Linearität von $\mathcal{B}$,

$$\mathcal{B}u \overset{(4.32)}{=} \sum_k \nu_k \mathcal{B}u_k \overset{(4.33)}{=} \sum_k \nu_k \sum_j b_{jk} v_j = \sum_j \left(\sum_k b_{jk} \nu_k \right) v_j. \tag{4.34}$$

Benennen wir die Koordinaten von $Bu \in V$ bezüglich der Basis (v_j) mit μ_j, so folgt aus deren Eindeutigkeit und (4.34), daß

$$\mu_j = \sum_{k=1}^{N} b_{jk} \nu_k. \tag{4.35}$$

Wir können mit (4.35) die Wirkung von $\mathcal{B}$ auf jeden beliebigen Vektor berechnen, wenn wir uns allein im Koordinatenraum aufhalten. Wir ersetzen also Rechnungen in V äquivalent

durch Rechnungen im $\mathbb{K}^M$. Der Übersichtlichkeit halber ordnen wir die Koeffizienten in einer $M \times N$ Matrix B an,

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1N} \\ b_{21} & b_{22} & \dots & b_{2N} \\ \dots & \dots & \dots & \dots \\ b_{M1} & \dots & \dots & b_{MN} \end{pmatrix}. \tag{4.36}$$

Interpretieren wir $\mu = (\mu_j)$ und $\nu = (\nu_k)$ als Spaltenvektoren, kürzen wir (4.35) ab zu

$$\mu = B\nu. \tag{4.37}$$

Bemerkung 4.32. Das allgemeine $M \times N$ LGS (4.6) hat die Form (4.35), daher können wir es mit Spaltenvektoren x der Länge N und b der Länge M schreiben als

$$Ax = b. \tag{4.38}$$

Satz 4.33. *Erklären wir Addition und Skalarmultiplikation von Matrizen fester Größe elementweise, so wird die Menge dieser Matrizen zu einem Vektorraum.*

Die binäre Abbildung $(B, \nu) \mapsto B\nu$, auch Matrix-Vektor-Produkt genannt, ist bilinear, d.h. linear in jedem der beiden Argumente, solange man das jeweils andere festhält.

Beweis. Die erste Aussage folgt aus den Rechenregeln über $\mathbb{K}$. Zur zweiten siehe unten. □

Wir untersuchen nun Verkettungen von linearen Abbildungen. Sei dazu W ein weiterer Vektorraum und $(w_i)_{i=1}^L$ eine Basis. Die lineare Abbildung $\mathcal{A} : V \to W$ habe bezüglich der angegebenen Basen die $L \times M$ (warum?!) Matrixdarstellung $A = (a_{ij})$.

Satz 4.34. *Die Verkettung $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ hat die $L \times N$ Matrixdarstellung $C = (c_{ik})$ mit*

$$c_{ik} = \sum_{j=1}^{M} a_{ij} b_{jk}. \tag{4.39}$$

Wir schreiben dies kürzer als $C = AB$. Das Matrixprodukt $(A, B) \mapsto AB$ ist bilinear und assoziativ, aber nicht kommutativ.

Beweis. Übung. Warum folgt hieraus der zweite Teil von Satz 4.33? □

Bemerkung 4.35. Zeilenvektoren sind nichts anderes als Matrizen mit $M = 1$ und Spaltenvektoren Matrizen mit $N = 1$. Zahlen in $\mathbb{K}$ sind 1×1 Matrizen und damit automatisch auch einelementige Zeilen- und Spaltenvektoren.

Bemerkung 4.36. Wenn B nur eine Spalte hat, gilt dies auch für C. Wir können uns dann den Spaltenindex für B und C sparen und erhalten das vorher beschriebene Matrix-Vektor-Produkt (4.35) von Spaltenvektoren als Spezialfall des Matrixproduktes. Spaltenvektoren von rechts werden wieder auf Spaltenvektoren abgebildet (durchaus verschiedener Länge).

Analog sehen wir, daß wenn A nur eine Zeile hat, dies auch für C gilt. Zeilenvektoren von links werden so auf Zeilenvektoren abgebildet (ebenfalls i.A. verschiedener Länge).

Um die Matrixmultiplikation (4.39) praktisch durchzuführen, benutzen wir

Folgerung 4.37. *Seien b_j N Spaltenvektoren der Länge M. Für eine gegebene $L \times M$ Matrix A besteht die Matrix $C = (c_1, \dots, c_N)$ aus genausovielen Spalten wie $B = (b_1, \dots, b_N)$, jeweils der Länge L, die sich einzeln als Matrix-Vektor-Produkt errechnen lassen:*

$$c_j = Ab_j, \quad 1 \le j \le N. \tag{4.40}$$

Zu guter Letzt beschreiben wir die Äquivalenz der Zeilen- und Spaltenperspektive.

Definition 4.38. Sei B eine beliebige $M \times N$ Matrix. Dann ist die transponierte Matrix B^T erklärt als diejenige $N \times M$ Matrix mit

$$B^T = (b^T_{jk}), \quad 1 \leq j \leq N, 1 \leq k \leq M, \quad b^T_{jk} = b_{kj}. \tag{4.41}$$

Satz 4.39. *Für das Produkt von Matrizen passender Größe gilt*

$$(AB)^T = B^T A^T. \tag{4.42}$$

Beweis. Dies folgt durch Einsetzen in (4.39) und Nachrechnen. □

4.7 Basiswechsel und inverse Matrizen

Seien $(u_k), (v_j)$ zwei Basen eines M-dimensionalen Vektorraumes V. Wie können wir die Koordinaten bezüglich einer Basis in die zur anderen umrechnen? Wir gehen vor wie in Abschnitt 4.6 und spezialisieren zu $U = V, \mathcal{B} = \mathrm{id}_V$, denn der Vektor ändert sich nicht (nur seine Koordinaten!). Die Darstellung (4.33) ist auch in diesem Zusammenhang eindeutig gegeben und definiert uns die Abbildung der Koordinaten nach (4.35).

Folgerung 4.40. *Es lassen sich die Basisvektoren (u_k) aus den Vektoren (v_j) einer zweiten Basis linear kombinieren gemäß*

$$u_k = \sum_{j=1}^{M} b_{jk} v_j. \tag{4.43}$$

Zur Vermeidung von Indizes bilden wir abstrakte Spaltenvektoren $\tilde{u}$, $\tilde{v}$ aus den jeweiligen Basisvektoren (zu was für einem Vektorraum gehören diese?!) und erweitern die Definition des Matrixproduktes darauf, so daß wir äquivalent schreiben

$$\tilde{u} = B^T \tilde{v}. \tag{4.44}$$

Damit folgt weiter die Äquivalenz der beiden Koordinatendarstellungen,

$$\nu^T \tilde{u} = \mu^T \tilde{v} \quad \Leftrightarrow \quad \mu = B\nu. \tag{4.45}$$

Vertauschte Rollen für (u_k) und (v_j) liefern die Matrix $C = (c_{kj})$ und mit (4.43)

$$v_j = \sum_{k=1}^{M} c_{kj} u_k = \sum_{k=1}^{M} c_{kj} \sum_{i=1}^{M} b_{ik} v_i = \sum_{i=1}^{M} \left(\sum_{k=1}^{M} b_{ik} c_{kj} \right) v_i. \tag{4.46}$$

Wir erhalten aus der linearen Unabhängigkeit der v_i die Beziehung $BC = \mathbb{1}$ mit der (grundsätzlich quadratischen) Einheitsmatrix

$$\mathbb{1} = (\delta_{ij}), \quad \delta_{ij} = \begin{cases} 1 & \text{wenn } i = j, \\ 0 & \text{sonst.} \end{cases} \tag{4.47}$$

Wenn wir den Basiswechsel komplett mit vertauschten Rollen von (u_k) und (v_j) wiederholen, sehen wir, daß insgesamt $CB = BC = \mathbb{1}$ und schreiben $B = C^{-1}$, $C = B^{-1}$. Nun folgen

$$(BC)^{-1} = C^{-1} B^{-1}, \tag{4.48a}$$

$$(B^{-1})^T = (B^T)^{-1}, \tag{4.48b}$$

so daß wir letzteres einfacher schreiben als B^{-T}.

Satz 4.41. *Die Menge der $M \times M$ Matrizen, die einen Basiswechsel beschreiben, bilden bezüglich des Matrixproduktes eine (für $M > 1$ nichtabelsche) Gruppe, genannt* $\mathrm{GL}(M)$. *Wir nennen diese Matrizen invertierbar.*

Beweis. Durch Einsetzen in (4.39) verifizieren wir, daß sogar für beliebige Matrizen A gilt

$$\mathbb{1}A = A, \qquad A\mathbb{1} = A. \tag{4.49}$$

Die Einheitsmatrix $\mathbb{1}$ übernimmt so die Rolle des neutralen Elements. Die Existenz des inversen Elements haben wir oben nachgewiesen. Es ist nur noch zu zeigen, daß jede Matrix, die eine Inverse besitzt, einen Basiswechsel darstellt. Wähle also eine beliebige Basis $\tilde{z}$ und invertierbare Matrix A. Wir müssen zeigen, daß $A^T\tilde{z}$ auch eine Basis ist. Nehmen wir eine Linearkombination mit Koeffizienten $\lambda = (\lambda_i)$ an, rechnen wir

$$0 = \lambda^T \left(A^T\tilde{z}\right) = (A\lambda)^T\tilde{z} \quad \Rightarrow \quad A\lambda = 0. \tag{4.50}$$

Die letzte Gleichung multiplizieren wir von links mit A^{-1} und erhalten $\lambda = 0$. □

Satz 4.42. *Die folgenden Aussagen über $M \times M$ Matrizen sind äquivalent:*

1. *A ist invertierbar.*
2. *Die Spalten von A sind linear unabhängig.*
3. *Die Zeilen von A sind linear unabhängig.*

Beweis. Da die Gleichung $A\lambda = 0$ eine Linearkombination von Spalten beschreibt, führt nach (4.50) die erste Aussage auf die zweite.

Die zweite besagt gerade, daß $A^T\tilde{z}$ eine Basis ist, also A einen Basiswechsel darstellt und damit invertierbar ist.

Die dritte Aussage ist äquivalent zu den ersten beiden, da die Transposition (also die Umordnung von Spalten zu Zeilen und umgekehrt) nach (4.48b) nichts an der Invertierbarkeit einer Matrix ändert. □

Bemerkung 4.43. Wegen der Isomorphie zweier Vektorräume gleicher Dimension sind die Invertierbarkeit einer Matrix und die Einträge einer Inversen unabhängig vom angesetzten Vektorraum. Es ist also zulässig, in den obigen Sätzen 4.41 und 4.42 den Vektorraum nicht zu spezifizieren. Die Invertierbarkeit von Matrizen läß sich tatsächlich ganz ohne den Rückgriff auf Vektorräume definieren.

Wir sehen nun aus (4.38), daß ein lineares Gleichungssystem mit $A \in \mathrm{GL}(M)$ die eindeutige Lösung hat $x = A^{-1}b$.

Danksagung

Teile dieser Notizen orientieren sich an einer Handschrift von Prof. Dr. Sven Beuchler. Weiterhin waren die Bücher „Zahlen“ von Ebbinghaus und anderen und „Einführung in die lineare Algebra“ von Walter sehr nützlich. Ich bedanke mich bei Katharina Hofer für das Korrekturlesen.