LINEARE ALGEBRA I
WS 2009/10

Karl-Hermann Neeb [1]

16. Februar 2010

[1] Dieses Skript ist aus früheren Fassungen von Klaus Keimel und Norbert Schappacher entstanden.

Inhaltsverzeichnis

Kapitel 1

Logik, Mengenlehre und Funktionen

Bevor wir uns den eigentlichen Gegenständen der Linearen Algebra zuwenden, geben wir eine pragmatische Einführung in zwei Disziplinen, die beim systematischen Aufbau der ganzen Mathematik zu den *Grundlagen der Linearen Algebra* gehören: die Logik und die Mengenlehre. Die Mengenlehre ist gewissermaßen die Sprache der Mathematik, die Logik ihre Grammatik. Weder die Logik noch die Mengenlehre können oder sollen im Rahmen dieser Vorlesung systematisch entwickelt werden. Es kommt vielmehr darauf an, sie richtig zu benutzen.

1.1 Aussagen- und Quantorenlogik

Die formale Logik handelt von Aussagen, die nach gewissen Regeln aus bestimmten vorgegebenen Zeichen aufgebaut werden. Wir betrachten eine Aussage als *wohlgeformt*, wenn sie wahr oder falsch ist. In der üblichen Theorie der ganzen Zahlen sind wohlgeformte Aussagen beispielsweise

1) 0 ist eine ganze Zahl

2) $2 + 2 = 5$

3) $a + a = 2a$.

Keine wohlgeformte Aussage hingegen ist etwa $?!a = x+!@$.

Die „Wahrheit“ Regel aus gewissen Grundannahmen der jeweiligen Theorie, den sogenannten *Axiomen*, logisch erschlossen. Die Aussagenlogik regelt, wie sich die Wahrheitswerte bei Verknüpfungen mehrerer Aussagen verhalten.

Definition 1.1.1. [Übliche Junktoren]. Seien p und q Aussagen. Dann lassen sich daraus folgende neue Aussagen bilden:

(i) *Negation:* $\neg p$ (gelesen nicht p oder ‘non p’) ist genau dann wahr, wenn p falsch ist.

(ii) *Konjunktion:* $p \wedge q$ (p und q) ist genau dann wahr, wenn p wahr ist und q wahr sind.

(iii) *Disjunktion:* $p \vee q$ (p oder q) ist genau dann wahr, wenn p wahr ist oder q wahr ist (nicht „entweder - oder“, d.h. $p \vee q$ ist auch dann wahr, wenn sowohl p als auch q wahr ist).

(iv) *Implikation:* $p \Rightarrow q$ (p impliziert q; aus p folgt q) ist definiert als $(\neg p) \vee q$. Die Wahrheit dieser Aussage ist demnach gleichbedeutend mit “Wenn p wahr ist, dann ist auch q wahr“ (Nachweis!).

(v) *Äquivalenz:* $p \iff q$ (p ist äquivalent zu q) genau dann, wenn beide wahr sind oder beide falsch sind. Beachte: $p \iff q$ ist genau dann wahr, wenn $p \Rightarrow q$ und $q \Rightarrow p$ wahr sind.

(vi) *Shefferscher Strich:* [1] $p \mid q$ (p Strich q) genau dann, wenn p und q nicht beide wahr sind. $\square$

Die Wahrheitswerte der oben definierten verknüpften Aussagen sind in der folgenden Wahrheitstafel zusammengefasst:

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \iff q$	$p \mid q$
W	W	F	W	W	W	W	F
W	F	F	F	W	F	F	W
F	W	W	F	W	W	F	W
F	F	W	F	F	W	W	W

Die folgenden Tatsachen, von denen (a) – (d) wichtige **Merkregeln** sind, verifiziert man durch Aufstellen der jeweiligen Wahrheitstafeln.

Bemerkung 1.1.2. (Hintereinanderausführung logischer Operatoren)

(a) Doppelte Verneinung bejaht: $\neg(\neg p) \iff p$. Diese Aussage ist unabhängig von p wahr. Solche Aussagen nennt man *allgemeingültig*.

(b) Negation vertauscht $\wedge$ und $\vee$: (de Morgansche Regeln):

$$\neg(p \vee q) \iff \neg p \wedge \neg q \quad \text{und} \quad \neg(p \wedge q) \iff \neg p \vee \neg q.$$

(c) Wir schreiben **W** bzw. **F** für die Aussage, die immer wahr bzw. immer falsch ist. Dann gilt für jede Aussage p:

$$p \wedge \mathbf{F} \iff \mathbf{F}, \quad p \vee \mathbf{F} \iff p \quad \text{und} \quad p \wedge \mathbf{W} \iff p, \quad p \vee \mathbf{W} \iff \mathbf{W}.$$

(d) Logische Distributivgesetze:

$$p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r) \quad \text{und} \quad p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r).$$

(e) Durch wiederholte Anwendung des Shefferschen Strichs lassen sich alle anderen Junktoren darstellen. $\square$

[1] Nach Henry M. Sheffer (1882–1964), Harvard-Professor. Veröffentlichte 1913 eine Arbeit zur Algebra der Logik, in der er das Resultat 1.1.2(e) — siehe unten — bewies.

Bemerkung 1.1.3. (Regeln für logisches Schließen)

(1) *Direkter Schluss:*

$$(p \wedge (p \Rightarrow q)) \Longrightarrow q$$

Ist p wahr und impliziert p die Aussage q, so ist q wahr. Die Allgemeingültigkeit dieser Aussage verifiziert man wiederum anhand der Tabelle.
Beispiel: Es sei p die Aussage „Es regnet“ und q die Aussage „ Die Straße ist nass“. Dann ist die Implikation $p \Rightarrow q$ eine wahre Aussage: „ Wenn es regnet, dann ist die Straße nass“.

(2) $(\neg q \wedge (p \Rightarrow q)) \Longrightarrow \neg p$
Beispiel: „ Ist die Straße nicht nass, so regnet es nicht.“

(3) *Kontraposition*:

$$(p \Rightarrow q) \Longleftrightarrow (\neg q \Rightarrow \neg p).$$

Beispiel: „ Wenn es regnet, ist die Straße nass“ $\Longleftrightarrow$ „Ist die Straße nicht nass, so regnet es nicht“.

(4) *Schlussketten*: In der Notation logischer Schlüsse verwenden wir zwei Typen von Schlussketten: In einer Schlusskette des Typs

$$p_1 \Longleftrightarrow p_2 \Longleftrightarrow \ldots \Longleftrightarrow p_n$$

verstehen wir das Zeichen $\Longleftrightarrow$ als ein Symbol, das uns signalisiert, dass alle Aussagen $p_1 \Longleftrightarrow p_2$, $p_2 \Longleftrightarrow p_3$, $p_3 \Longleftrightarrow p_4$ usf. wahr sind. Eine solche Kette von Äquivalenzen hat insbesondere $p_1 \Longleftrightarrow p_n$ zur Folge. Zum Beispiel folgt die Gültigkeit von (3) unter Verwendung der doppelten Verneinung (1.1.2(a)) aus folgender Kette von Äquivalenzen:

$$(p \Rightarrow q) \Longleftrightarrow (\neg p) \vee q \Longleftrightarrow q \vee (\neg p) \Longleftrightarrow (\neg q \Rightarrow \neg p).$$

Entsprechend verwenden wir das Symbol $\Rightarrow$. In einer Schlusskette des Typs

$$p_1 \Rightarrow p_2 \Rightarrow \ldots \Rightarrow p_n$$

bedeutet es, dass alle Aussagen $p_1 \Rightarrow p_2$, $p_2 \Rightarrow p_3$, $p_3 \Rightarrow p_4$ usw. wahr sind. Insbesondere gilt dies dann für $p_1 \Rightarrow p_n$.

Daraus ergibt sich ziemlich schnell, dass eine Kette von Äquivalenzen

$$[A] \qquad p_1 \Longleftrightarrow p_2 \Longleftrightarrow \ldots \Longleftrightarrow p_n$$

genau dann wahr ist, wenn dies für die um eins längere Kette von Implikationen

$$[I] \qquad p_1 \Rightarrow p_2 \Rightarrow \ldots \Rightarrow p_n \Rightarrow p_1$$

gilt. Wie wir oben in Definition 1.1.1(v) bemerkt haben ist eine einzelne Äquivalenz formallogisch gleichbedeutend mit zwei Implikationen. Demnach müsste man zum Beweis von [A] a priori $2n-2$ Implikationen nachweisen. In Wirklichkeit reichen aber schon die n Implikationen in [I], wobei man außerdem noch die Reihenfolge der $p_1, p_2, p_3, \ldots$ geeignet vertauschen kann, um den Nachweis von [I] so ökonomisch wie möglich zu gestalten. — Beispiele für dieses Prinzip werden wir in vielen Beweisen dieser Vorlesung sehen. $\square$

Bemerkung 1.1.4. (Formale Struktur mathematischer Sätze bzw. Beweise)
Hat man einen Satz der Gestalt

$$p \Longrightarrow q,$$

so gibt es mehrere Möglichkeiten, ihn zu beweisen.

(1) *direkter Beweis:* Man nimmt an, die *Voraussetzung* p ist wahr und schließt hieraus, dass die *Behauptung* q wahr ist (siehe 1.1.1(iv)).

(2) Für den *indirekten Beweis* gibt es zwei Varianten, die auf den Äquivalenzen

$$(p \Rightarrow q) \Longleftrightarrow (\neg q \Rightarrow \neg p) \Longleftrightarrow \neg(p \wedge \neg q)$$

beruhen.

(a) Man nimmt an, dass q falsch ist und leitet daraus ab, dass p falsch ist.

(b) Die andere Variante besteht darin anzunehmen, dass p wahr ist und q falsch und daraus einen Widerspruch herzuleiten. Hiermit ist die Wahrheit der Aussage $\neg(p \wedge \neg q)$ bewiesen und damit $p \Rightarrow q$. □

Betrachten wir ein erstes Beispiel eines indirekten Beweises. Als *natürliche Zahlen* bezeichnen wir in dieser Vorlesung die Zahlen $1, 2, 3$, usf., d.h. alle Zahlen, die von der 1 ausgehend in endlich vielen Schritten durch Hinzuzählen von Einsen erreicht werden können. Ihre Gesamtheit, also die *Menge der natürlichen Zahlen*, bezeichnen wir mit $\mathbb{N} := \{1, 2, 3, \dots\}$. [2]

Satz 1.1.5. *Ist n eine durch 4 teilbare natürliche Zahl, so ist $n + 3$ keine Quadratzahl.*

Beweis. (Indirekt) Wir nehmen an, $n + 3$ sei eine Quadratzahl, d.h. es gibt eine natürliche Zahl k mit $n + 3 = k^2$.

1. Fall: k ist gerade, d.h. es gibt eine natürliche Zahl m mit $k = 2m$. Dann ist $k^2 = 4m^2$ durch 4 teilbar und folglich $n = k^2 - 3$ nicht durch 4 teilbar.

2. Fall: k ist ungerade, d.h. es gibt eine natürliche Zahl m mit $k = 2m + 1$. (Beachte: k ist sicher größer als 1, mithin mindestens 3.) Dann kommt $k^2 = 4m^2 + 4m + 1$, d.h. $k^2 - 1$ ist durch 4 teilbar, also $n = k^2 - 3 = (k^2 - 1) - 2$ nicht. □

Der eben bewiesene Satz gilt *für alle* natürlichen Zahlen n. Eine übliche Weise, derartige Aussagen in der Logik zu formalisieren, behandelt das Symbol n (nicht als beliebige Konstante, sondern) als *Variable*, d.h. als eine besondere Art Zeichen, über welche man *quantifizieren* kann:

Definition 1.1.6. (Quantoren)
(1) *Der Allquantor:* Sei J eine Menge (vgl. hierzu den nächsten Abschnitt) und

[2] Andere Autoren erkennen auch 0 als eine natürliche Zahl an. Wenn wir die Gesamtheit aller Zahlen $0, 1, 2, 3 \dots$ bezeichnen müssen, tun wir das mit dem Symbol $\mathbb{N}_0$.

$(p_j)_{j\in J}$ eine *Familie von Aussagen*, d.h. für jedes $j \in J$ ist eine Aussage p_j gegeben. Dann ist

$$(\forall j \in J)\, p_j$$

die Aussage, die genau dann wahr ist, wenn p_j für alle $j \in J$ wahr ist.
Beispiele:

(a) $(\forall n \in \mathbb{N})\underbrace{\text{„}n \text{ ist gerade“}}_{p_n}$ ist falsch, da 3 nicht gerade ist.

(b) $(\forall n \in \mathbb{N})\, n \leq n^2$ ist wahr.

(2) *Der Existenzquantor:* Ist $(p_j)_{j\in J}$ eine Familie von Aussagen, so ist

$$(\exists j \in J)\, p_j$$

die Aussage, die genau dann wahr ist, wenn (mindestens) ein $j_0 \in J$ existiert, so dass p_{j_0} wahr ist.
Beispiele:

(a) $(\exists n \in \mathbb{N})$ „ n ist gerade“ ist wahr.

(b) $(\exists n \in \mathbb{N})$ „ n ist Primzahl“ ist ebenfalls wahr.

(3) *Quantor der eindeutigen Existenz:* Die Aussage

$$(\exists!\ j \in J)\, p_j$$

soll bedeuten: Es existiert *genau* ein $j_0 \in J$, so dass p_{j_0} wahr ist. — Mit anderen Worten:

$$(\exists!\ j \in J)\, p_j \iff \Big[(\exists j \in J)\, p_j\Big] \ \wedge\ \Big[(\forall j_1 \in J)(\forall j_2 \in J)\Big([p_{j_1} \wedge p_{j_2}] \Rightarrow j_1 = j_2\Big)\Big].$$

Beispiele:

(a) $(\exists! n \in \mathbb{N})$ „ n ist gerade“ ist falsch.

(b) $(\exists! n \in \mathbb{N})$ $n^3 = 27$ ist wahr. □

Bemerkung 1.1.7. (Merkregeln für den Umgang mit Quantoren)
(a) Die Entsprechungen der *de Morganschen Regeln* für Quantoren sind

$$\neg\Big((\forall j \in J)\, p_j\Big) \iff (\exists j \in J)\, \neg p_j \quad \text{und} \quad \neg\Big((\exists j \in J)\, p_j\Big) \iff (\forall j \in J)\, \neg p_j.$$

(b) Man darf Existenz- und Allquantor im allgemeinen nicht vertauschen:

$$\underbrace{(\forall n \in \mathbb{N})(\exists k \in \mathbb{N})\, n \leq k}_{W} \not\Longleftrightarrow \underbrace{(\exists k \in \mathbb{N})(\forall n \in \mathbb{N})\, n \leq k}_{F}.$$

1.2 Mengenlehre

Dem Begriff der Menge stellen wir uns naiv gegenüber, d.h. wir stellen uns auf den Standpunkt, dass wir eine Menge kennen, wenn uns gesagt wird, welche Elemente sie enthält. Wie kann das aussehen?

Ist M eine Menge, so schreiben wir

$$x \in M$$

für die Aussage, die genau dann wahr ist, wenn x Element der Menge M ist, und

$$x \notin M :\Leftrightarrow \neg(x \in M).$$

Definition 1.2.1. (Beschreibung von Mengen)
(1) (**Aufzählung der Elemente**) Eine Menge kann durch Aufzählung ihrer Elemente beschrieben werden:
$M = \{4, 6, \{1, 2\}\}$, $N = \{+, -, 8\}$.
$\mathbb{N} = \{1, 2, 3, \ldots\}$ — Die Menge der *natürlichen Zahlen*. (Hier und bei den folgenden Beispielen wird eher eine Regel zum immer-weiter-Zählen angegeben, als eine effektive Aufzählung.)
$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$ — Die Menge der natürlichen Zahlen mit der Null.
$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}$ — Die Menge der *ganzen Zahlen*.

Beachte: $\{1, 2, 3, 2\} = \{1, 2, 3\}$. — In der Aufzählung wiederholte Elemente ändern nichts daran, welche Elemente zu der Menge gehören.

Eine andere Möglichkeit besteht in der Beschreibung durch andere Mengen. So kommt man etwa zur folgenden Aufzählung der Menge der *rationalen Zahlen*:
$\mathbb{Q} = \{\frac{p}{q} \ : \ p \in \mathbb{Z} \ \wedge \ q \in \mathbb{N}\}$ — Hier wird die übliche Schreibweise für Brüche als bekannt angenommen. Diese Schreibweise ist definiert durch die Äquivalenz: $\frac{p}{q} = \frac{p'}{q'} \Longleftrightarrow pq' = p'q$. Auch hier werden also alle Elemente mehrfach aufgezählt (z.B. die rationale Zahl $\frac{1}{3}$ ebenfalls als $\frac{2}{6}$, $\frac{3}{9}$, $\frac{4}{12}$, usf. Die negativen rationalen Zahlen freilich werden immerhin nur halb so oft aufgezählt, als wenn wir auch negative ganze Nenner $q \neq 0$ zugelassen hätten.

(2) (**Aussonderung**) Die Elemente einer Menge können durch eine *Aussageform* spezifiziert werden: Zu jedem Element x einer Menge M sei uns eine Aussage $p(x)$ gegeben. Wir nennen das Symbol $p(x)$ dann eine *Aussageform* und x die *freie Variable* in $p(x)$. Wir können hiermit innerhalb der Menge M die Menge

$$N := \{x \in M \ : \ p(x)\}$$

aussondern, die genau diejenigen Elemente x von M enthält, für die die Aussage $p(x)$ wahr ist. □

Beispiele:

a) Die Menge der geraden Zahlen

$$G = \{n \in \mathbb{N} \colon (\exists m \in \mathbb{N})\, n = 2m\} = \{n \in \mathbb{N} \colon n \text{ ist gerade}\}$$

Hier ist $p(n)$ die Aussage „ $(\exists m \in \mathbb{N})\, n = 2m$“ bzw. „ n ist gerade“.

b) Die Menge aller Primzahlen

$$P = \{n \in \mathbb{N} \colon n \text{ ist Primzahl}\}.$$

Hier ist also $p(n)$ die Aussage „ n ist Primzahl“. $\square$

Bemerkung 1.2.2. ACHTUNG.
Die Einschränkung $x \in M$ in Definition 1.2.1(2) ist wesentlich, da sie unerlaubte Konstruktionen wie die folgende, die sogenannte *Russellsche Antinomie*,[3] ausschließt:

$$R \colon = \{x \colon x \notin x\}.$$

Diese Definition führt zu einem Widerspruch, wenn man fragt, ob die Menge R selbst Element von R ist:

- Ist $R \in R$, so folgt aus der definierenden Eigenschaft der Menge R, dass $R \notin R$ ist – Widerspruch; und
- ist $R \notin R$, so gilt die definierende Eigenschaft der Menge R für R, so dass $R \in R$ gilt – Widerspruch!

Diese Art von Konstruktion belegt:

Es gibt keine „ Menge aller Mengen“!

Denn gäbe es die, so könnten wir die Russellsche Menge R durch Aussonderung daraus erhalten, was aber absurd ist.

$\square$

Die logischen Junktoren haben ihre Entsprechungen in Mengenvergleichen und Mengenbildungen, wie man in den folgenden beiden Definitionen sieht:

Definition 1.2.3. (1) $A \subseteq B$ (A ist *Teilmenge von* B) bedeutet $x \in A \Rightarrow x \in B$.
(2) $A = B \colon \iff \big((x \in A) \iff (x \in B)\big)$, d.h. zwei Mengen sind gleich, wenn sie dieselben Elemente enthalten. Man beachte (ebenso wie eine Äquivalenz gleichbedeutend war mit zwei Implikationen), dass:

$$A = B \iff (A \subseteq B) \wedge (B \subseteq A).$$

(3) $\emptyset$: Die *leere Menge*. Sie enthält keine Elemente; die Aussage $x \in \emptyset$ ist immer falsch, d.h. $x \in \emptyset \iff \mathbf{F}$. $\square$

[3] Bertrand Russell (1872–1969), englischer Mathematiker, Logiker und Philosoph.

Definition 1.2.4. (Konstruktion neuer Mengen)
Seien X und Y Mengen.
(i) *Das Komplement von Y in X*, das der logischen Negation entspricht, beschreiben wir durch Aussonderung:

$$X \setminus Y \colon = \{x \in X \colon x \notin Y\}$$

(ii) *Vereinigung zweier Mengen*

$$X \cup Y \colon = \{x \colon x \in X \vee x \in Y\}$$

lässt sich nicht durch Aussonderung innerhalb einer großen Menge beschreiben. dass diese Bildung keine Paradoxien produziert, muss man hier glauben.
(iii) Den *Durchschnitt zweier Mengen* können wir hingegen durch eine regelrechte Aussonderung beschreiben:

$$X \cap Y = \{x \ : \ x \in X \wedge x \in Y\} = \{x \in X \ : \ x \in Y\}$$

(iv) *Beliebige Durchschnitte und Vereinigungen:* Ist $\{A_j \colon j \in J\}$ eine Menge von Mengen, so definieren wir

$$\bigcup_{j \in J} A_j \colon = \{x \colon (\exists j \in J)\ x \in A_j\}.$$

Dann ist $x \in \bigcup_{j \in J} A_j \ \Leftrightarrow \ (\exists j \in J)\ x \in A_j$. Ist $J = \emptyset$, so ist diese Aussage immer falsch und daher gilt in diesem Fall $\bigcup_{j \in J} A_j = \emptyset$. Analog definieren wir für eine nichtleere Menge J:

$$\bigcap_{j \in J} A_j \colon = \{x \colon (\forall j \in J)\ x \in A_j\}.$$

Für endliche Mengen $J = \{1, 2, 3, \ldots, n\}$ schreibt man auch

$$\bigcup_{j=1}^{n} A_j = \bigcup_{j \in J} A_j = A_1 \cup A_2 \cup \ldots \cup A_n$$

bzw.

$$\bigcap_{j=1}^{n} A_j = \bigcap_{j \in J} A_j = A_1 \cap A_2 \cap \ldots \cap A_n.$$

(v) *Das kartesische Produkt zweier Mengen:* Für $x \in X$ und $y \in Y$ nennt man eine geordnete Auflistung (x, y) ein *geordnetes Paar*. Das kann man auch ganz formal als zweielementige Menge so definieren:

$$(x, y) := \{\{x\}, \{x, y\}\}. \tag{‡}$$

Diese Definition garantiert gerade die charakteristische Eigenschaft der geordneten Paare:

$$(x, y) = (x', y') \Longleftrightarrow \Big((x = x') \ \wedge \ (y = y')\Big).$$

Die Menge aller solcher Paare

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

heißt *kartesisches Produkt* der Mengen X und Y.
Folgende Konstruktion ist etwas allgemeiner. Sind $A_1, \ldots, A_n$ Mengen und $a_1 \in A_1, \ldots, a_n \in A_n$, so heißt die geordnete Liste $(a_1, a_2, \ldots, a_n)$ ein *n-Tupel* (2-Tupel sind *Paare*; 3-Tupel werden *Tripel* genannt). Die formale Definition der n-Tupel als Mengen, in Verallgemeinerung zur Formel (‡) für geordnete Paare, ist:

$$(a_1, a_2, a_3, \ldots, a_n) := (a_1, (a_2, \ldots, a_n)), \quad n \geq 3.$$

Man definiert dann

$$A_1 \times \ldots \times A_n \colon = \{(a_1, \ldots, a_n) \colon a_1 \in A_1 \wedge \ldots \wedge a_n \in A_n\}.$$

Für $A_1 = A_2 = \ldots = A_n$ schreibt man auch

$$A^n \colon = A_1 \times \ldots \times A_n = \underbrace{A \times \ldots \times A}_{n \text{ Faktoren}}.$$

Beispiele:

a) $\mathbb{Z}^3 \colon = \{(n, m, k) \colon n, m, k, \in \mathbb{Z}\}$

b) Die Mengen $A \colon = \{\star, \circ\}$ und $B \colon = \{\bullet, 1\}$ liefern

$$A \times B = \{(\star, \bullet), (\star, 1), (\circ, \bullet), (\circ, 1)\}.$$

(vi) *Die Potenzmenge:* Ist A eine Menge, so heißt

$$\mathbf{P}(A) \colon = \{B \colon B \subseteq A\}$$

die *Potenzmenge* von A. Sie enthält alle Teilmengen von A.
Beispiel: Für $A = \{0, 1\}$ ist $\mathbf{P}(A) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$.
Man beachte: Für jede Menge A gilt $\emptyset \subseteq A$, und somit $\emptyset \in \mathbf{P}(A)$. □

1.3 Abbildungen und Funktionen

Mathematiker gebrauchen die Worte ‘Abbildung’ und ‘Funktion’ im Kontext der Mengenlehre meist als gleichbedeutend. Manchmal ist das Wort ‘Funktion’ aber auch für solche Abbildungen reserviert, deren Werte (natürliche, reelle, komplexe, ...) Zahlen sind.

Seien M und N Mengen. Eine naive Definition dessen, was eine Abbildung $\varphi \colon M \to N$ (lese: “φ von M nach N“) ist, ginge etwa so: ‘Eine Abbildung φ von M nach N ist eine Regel, die jedem Element $x \in M$ genau ein Element $\varphi(x)$ der Menge N zuordnet.’ Die Schwachstelle dieser Definition ist, dass das Wort ‘Regel’ keine genau festgelegte Bedeutung hat, so dass die ganze Definition in der Luft hängt. In diesem Abschnitt geben wir eine exakte, mengentheoretische Fassung des Abbildungsbegriffes und betten ihn dabei gleichzeitig in den allgemeineren Rahmen der (zweistelligen) *Relationen* ein.

Relationen, interpretiert man sie inhaltlich, drücken Eigenschaften aus. Eine einstellige Relation r über einer gegebenen Menge M etwa soll erfassen, welche Elemente von M die "Eigenschaft r" haben. Ist etwa M die Menge der Studierenden der TU Darmstadt, so kann r beispielsweise die Eigenschaft bezeichnen, weiblich zu sein. Zumindest mengentheoretisch gesehen ist dieses r ganz und gar dadurch bestimmt, dass ich die Teilmenge der weiblichen Studierenden der TU Darmstadt kenne. So ist eine *einstellige Relation* auf einer Menge M im Rahmen der Mengenlehre nichts anderes als eine Teilmenge von M.

Zweistellige Relationen sind Eigenschaften, die je ein Element zweier gegebener Mengen zueinander in Beziehung setzen. Ist etwa N wie eben die Menge aller derzeitigen Studierenden der TU Darmstadt und M die Menge aller Menschen, die in den letzten 30 Jahren gelebt haben, so ist beispielsweise die Relation „A ist Vater oder Mutter des bzw. der Studierenden B"mengentheoretisch erfasst durch die Teilmenge

$$R = \{(A, B) \in M \times N \ : \ B \text{ ist Kind von } A\} \subseteq M \times N.$$

Diese Überlegung führt zur folgenden ersten Definition:

Definition 1.3.1. (*Erster Anlauf*) Seien M und N Mengen. Eine *(zweistellige) Relation R über M und N* ist eine Teilmenge des kartesischen Produktes $M \times N$. — In einer Formel: "$R \subseteq M \times N$."

Sprechweise: Ist R eine Relation über M und N, so sagt und schreibt man für Elemente $x \in M$ und $y \in N$ statt $(x, y) \in R$ auch "x und y stehen in der Relation R".

ACHTUNG: Die Mengen M und N sollen mit zur Definition von R gehören. Zwei Relationen $R \subseteq M \times N$ und $R' \subseteq M' \times N'$ sollen nicht schon dann gleich sein, wenn R und R' einfach als Mengen gleich sind, d.h. dieselben Elemente (x, y) enthalten; sondern es müssen auch die umgebenden Mengen $M = M'$ und $N = N'$ sein. Die beste Art, dies mathematisch streng zu fassen, ist wie folgt:

Definition 1.3.2. (*Präzisierung.*) Eine *(zweistellige) Relation R* ist ein Tripel (M, N, R), wo M und N Mengen sind und R eine Teilmenge von $M \times N$ ist.

Beispiele 1.3.3. (1) Sei $M = \mathbb{Q}$, $N = \mathbb{Z}$, $R_1 = \{(x, n) \in \mathbb{Q} \times \mathbb{Z} \ : \ x \geq n\}$. Dann ist (M, N, R_1) die Relation "größer oder gleich" zwischen rationalen und ganzen Zahlen.

(2) Seien $M = N = \mathbb{Q}$ und

$$R_2 := \{(x, y) \in \mathbb{Q}^2 \colon x^2 + y^2 = 1\}.$$

Dann ist $(\mathbb{Q}, \mathbb{Q}, R_2)$ die Relation „das Paar (x, y) liegt auf der Einheitskreislinie in der Ebene". Beispiele von Paaren in R_2 sind

$$(0, 1), \Big(\frac{3}{5}, \frac{4}{5}\Big), \Big(\frac{5}{13}, \frac{12}{13}\Big).$$

(3) Seien $M = N = \mathbb{Q}$ und

$$R_3 := \{(x, y) \colon x^2 + y = 1\}.$$

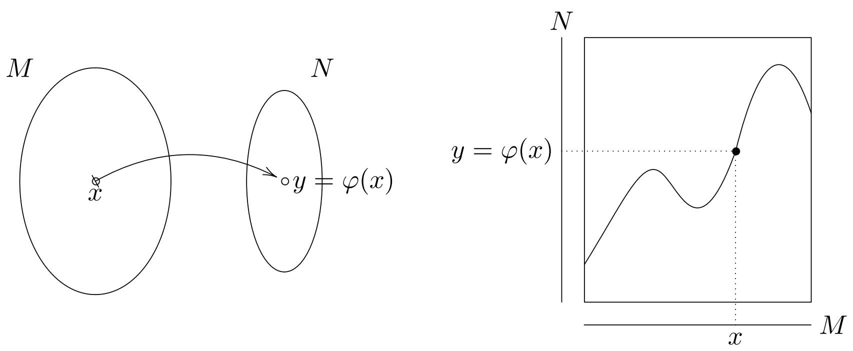
Dann ist $(\mathbb{Q}, \mathbb{Q}, R_3)$ eine Relation.

Definition 1.3.4. Eine *Abbildung (oder auch Funktion) von M nach N*, geschrieben $\varphi \colon M \to N$, ist eine (zweistellige) Relation (M, N, Γ_φ) über M und N, so dass die folgende Bedingung erfüllt ist:

$$(\forall x \in M)(\exists! y \in N)(x, y) \in \Gamma_\varphi.$$

Mit anderen Worten, φ ist eine rechts-eindeutige Relation: Für jedes Element m von M gibt es genau ein $n \in N$, welches durch φ mit m in Relation steht. Dies präzisiert die eingangs angedeutete Idee, dass φ jedem m ein n zuordnet. Die 'Regel', von der dort die Rede war, ist einfach durch den 'Graphen' $\Gamma_\varphi \subset M \times N$ der Funktion ersetzt.

1. Ist φ eine Abbildung und $(x, y) \in \Gamma_\varphi$, so nennt man das wegen der Abbildungseigenschaft eindeutig bestimmte Element $\varphi(x) := y$ den *Wert von φ in* (oder: *an der Stelle*) x. Eine andere Schreibweise hierfür ist: $x \mapsto \varphi(x)$, lies: $x \in M$ wird durch φ das Element $\varphi(x) \in N$ zugeordnet, oder „x geht nach $\varphi(x)$“.



2. Die Menge M heißt der *Definitionsbereich* der Abbildung φ.

3. Die Menge N heißt der *Bildbereich* der Abbildung φ.

4. $\Gamma_\varphi = \{(x, \varphi(x)) \colon x \in M\}$ heißt der *Graph* der Abbildung φ.

5. Die Schreibweise $\varphi \colon M \to N$ besagt, dass $\varphi = (M, N, \Gamma_\varphi)$ eine Abbildung mit Definitionsbereich M und Bildbereich N ist.

Von den oben gegebenen Beispielen für zweistellige Relationen ist R_3 eine Abbildung: $x \mapsto 1 - x^2$, während R_1 und R_2 keine Abbildungen sind.

1.3.5. Eine Abbildung kann auf verschiedene Weisen gegeben werden:

1. *Durch eine Formel*, wie z.B. die folgenden Abbildungen von $\mathbb{Q}$ nach $\mathbb{Q}$:

$$\begin{aligned} x &\mapsto \alpha x && \text{(lineare Abbildungen)} \\ x &\mapsto \alpha x + \beta && \text{(affine Abbildungen)} \\ x &\mapsto p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x + \cdots + \alpha_n x^n && \text{(Polynome)} \end{aligned}$$

2. *Durch eine Beschreibung mit Worten (ggbf. mit zugehörigen Bildern)*, wie z.B. folgende geometrischen Abbildungen der sogenannten euklidischen Ebene in sich selbst:
   - Spiegelung an einer Geraden g
   - orthogonale Projektion auf eine Gerade g
   - Drehungen

3. Häufig liest man auch Beschreibungen wie:
   - Sei $\varphi\colon \mathbb{N}_0 \to \mathbb{N}_0$ die Funktion, die jedem $n \in \mathbb{N}$ den Rest r nach Teilung durch 42 zuordnet.
   - Sei $\varphi\colon \mathbb{Z} \to \{0, 1\}$ definiert durch

$$\varphi(x) = \begin{cases} 1 & \text{für } x \text{ ungerade} \\ 0 & \text{für } x \text{ gerade.} \end{cases}$$

1.3.6. [Der Funktionsbegriff.] Der Funktionsbegriff hat sich in einer jahrhundertelangen Entwicklung stark verändert. Im achtzehnten Jahrhundert, z.B. bei Leonhard Euler (1707–1783), wurden reell- oder komplexwertige Funktionen im wesentlichen mit der sie beschreibenden Formel (wie in 1.3.5.1) identifiziert.

Peter Gustav Lejeune-Dirichlet (1805–1859) befreite den Funktionsbegriff vom Formelausdruck, wodurch eine deutlich allgemeinere Theorie möglich wurde. Von ihm stammt die seinerzeit fast revolutionäre Formulierung, auf die wir in der Einleitung indirekt angespielt haben: „Eine Funktion $y(x)$ ist gegeben, wenn wir irgendeine Regel haben, die jedem x in einer gewissen Punktmenge einen bestimmten Wert y zuordnet.“ Von ihm stammt auch das damals unerhörte Beispiel der folgenden Funktion $\Phi : \mathbb{R} \longrightarrow (0, 1\}$:

$$\Phi(x) = \begin{cases} 1 & \text{falls } x \text{ rational} \\ 0 & \text{falls } x \text{ irrational.} \end{cases}$$

Definitions- und Bildbereich einer Abbildung müssen keineswegs immer Zahlbereiche sein. So können gewisse logische Formeln als input akzeptieren und einen Wahrheitswert als output produzieren.

1.3.7. Fachausdrücke für Abbildungen $\varphi = (M, N, \Gamma_\varphi)$ von M nach N:

1. Definitionsgemäß sind zwei Abbildungen $\varphi_1 = (M_1, N_1, \Gamma_1)$ und $\varphi_2 = (M_2, N_2, \Gamma_2)$ genau dann *gleich*, wenn ihre Definitions- und Bildbereiche, sowie ihre Graphen gleich sind:

$$M_1 = M_2, \quad N_1 = N_2, \quad \Gamma_1 = \Gamma_2.$$

2. Für eine Teilmenge $A \subseteq M$ heißt die Menge

$$\varphi(A)\colon = \{y \in N\colon (\exists x \in A)\, y = \varphi(x)\}$$

das *Bild* von A unter φ.

3. Die Menge $\mathrm{im}(\varphi) := \varphi(M) = \{y \in N\colon (\exists x \in M)\, y = \varphi(x)\}$ (nach engl. "image“) heißt das *Bild der Abbildung* φ.

4. Für jede Teilmenge $B \subseteq N$ heißt die Menge

$$\varphi^{-1}(B)\colon = \{x \in M\colon \varphi(x) \in B\}$$

das *Urbild* B unter φ.

5. Für $B = \{y\} \subseteq N$ schreibt man einfach $\varphi^{-1}(y) := \varphi^{-1}(\{y\})$.

Beispiele 1.3.8. Sei $M = N = \mathbb{Z}$.

1. $\varphi(x) = x^2 + 5$ hat den Graphen $\Gamma_\varphi = \{(x, y) \in \mathbb{Z}^2\colon y = x^2 + 5\}$.

2. $\varphi(x) = x + 1$ hat den Graphen $\Gamma_\varphi = \{(x, y) \in \mathbb{Z}^2\colon y = x + 1\}$.

3. Es gibt keine Funktion $\varphi\colon \mathbb{Z} \to \mathbb{Z}$, deren Graph die Menge

$$\Gamma = \{(x, y) \in \mathbb{Z}^2\colon y^2 = x\}$$

ist; denn $\{(1, 1), (1, -1)\} \subseteq \Gamma$.

4. Abbildungen müssen, wie gesagt, nicht auf Zahlbereichen definiert sein — etwa:

$$\varphi\colon \{\text{ Studierende der TUD}\} \to \mathbb{N}, \quad x \mapsto \text{ Alter von } x.$$

Definition 1.3.9. Seien M und N Mengen.

Die Identität: Für jede Menge hat man die identische Abbildung

$$\mathrm{id}_M = (M, M, \Gamma), \qquad \text{für die} \qquad \Gamma = \{(x, y) \in M \times M : x = y\}$$

die Diagonale im kartesischen Produkt $M \times M$ ist.

Die konstante Abbildung: Für jedes Element $y_0 \in N$ hat man die Abbildung $\varphi\colon M \to N$ mit $\varphi(x) = y_0$ for all $x \in M$. Ihr Graph ist die Menge

$$\Gamma = \{(x, y_0) \colon x \in M\} = X \times \{y_0\}.$$

Diese Abbildung heißt die *konstante Abbildung mit Wert* y_0.

Einschränkungen einer Abbildung: Ist $\varphi\colon M \to N$ eine Abbildung und $A \subseteq M$ eine Teilmenge, so kann man die Abbildung f auf A einschränken und erhält eine Abbildung

$$\varphi|_A \colon A \to N, \quad x \mapsto \varphi(x),$$

welche die *Einschränkung von* φ *auf* A heißt. Ist $\Gamma_\varphi \subseteq M \times N$ der Graph von φ, so ist

$$\Gamma_{\varphi|_A} = \{(x, \varphi(x)) \colon x \in A\} = \Gamma_\varphi \cap (A \times N)$$

der Graph von $\varphi|_A$.

Injektivität, Surjektivität und Bijektivität

Eine Abbildung $\varphi\colon M \to N$ ordnet jedem $x \in M$ genau ein Element $\varphi(x) \in N$ zu. Ist umgekehrt $y \in Y$, so ist $\varphi^{-1}(y) \subseteq M$ die Menge aller Lösungen der Gleichung

$$\varphi(x) = y, \quad x \in M.$$

Offenbar gibt es hier drei sich gegenseitig ausschließende Möglichkeiten:

- $y \notin \mathrm{im}(\varphi)$ und $\varphi^{-1}(y) = \emptyset$. Die Gleichung hat also keine Lösung.
- $\varphi^{-1}(y) = \{x_0\}$ besteht aus genau einem Element. Die Gleichung ist also eindeutig lösbar.
- $\varphi^{-1}(y)$ enthält mehr als ein Element; es gibt verschiedene Lösungen der Gleichung.

Diese Trichotomie wird uns in dieser Vorlesung verschiedentlich begegnen, insbesondere im nächsten Abschnitt bei der Lösung linearer Gleichungssysteme. Hier führt sie uns zur Unterscheidung verschiedener Typen von Abbildungen.

Definition 1.3.10. Eine Abbildung $\varphi\colon M \to N$ heißt

injektiv, falls jedes $y \in N$ höchstens ein Urbild hat. Formallogisch:

$$(\forall x \in M)(\forall x' \in M)\, \varphi(x) = \varphi(x') \Longrightarrow x = x'.$$

Oder in Kontraposition:

$$(\forall x \in M)(\forall x' \in M)\, x \neq x' \Longrightarrow \varphi(x) \neq \varphi(x').$$

surjektiv, falls jedes $y \in N$ mindestens ein Urbild hat. Formallogisch:

$$(\forall y \in N)(\exists x \in M)\, y = \varphi(x),$$

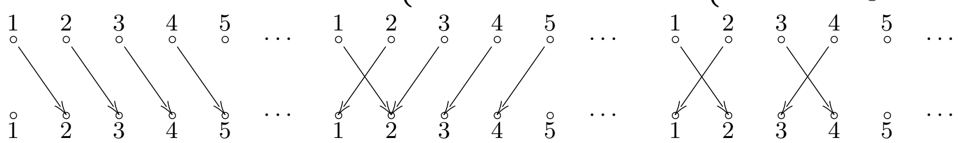
oder auch $\operatorname{im}(\varphi) = N$.

bijektiv, falls jedes $y \in N$ genau ein Urbild hat. Das ist äquivalent dazu, dass φ injektiv und surjektiv ist.

Drückt man diese Eigenschaften durch die Lösungsmengen der Gleichung $\varphi(x) = y$ für gegebenes y aus, so bedeutet die Injektivität, dass es höchstens eine Lösung gibt; die Surjektivität bedeute die Lösbarkeit der Gleichung für jedes y, und die Bijektivität bedeutet die eindeutige Lösbarkeit.

Beispiel 1.3.11. Betrachte die folgenden Abbildungen $\varphi\colon \mathbb{N} \to \mathbb{N}$:

$$n \mapsto n+1 \qquad n \mapsto \begin{cases} n-1 & \text{für } n > 1 \\ 2 & \text{für } n = 1 \end{cases} \qquad n \mapsto \begin{cases} n+1 & \text{für } n \text{ ungerade} \\ n-1 & \text{für } n \text{ gerade} \end{cases}$$



injektiv, nicht surjektiv | surjektiv, nicht injektiv | bijektiv

Seien M und N Mengen. Zu jeder Relation (M, N, R) könne wir die *inverse Relation* wie folgt definieren:

$$R^{-1}\colon = \{(y, x) \in N \times M\colon (x, y) \in R\}.$$

Satz 1.3.12. *Für eine Abbildung $\varphi\colon M \to N$ mit Graph Γ_φ sind die folgenden Bedingungen äquivalent.*

(i) *φ ist bijektiv.*

(ii) *Γ_φ^{-1} ist der Graph einer Abbildung $\psi\colon N \to M$, d.h. $(N, M, \Gamma_\varphi^{-1})$ ist eine Abbildung.*

Beweis. Das Tripel $(N, M, \Gamma_\varphi^{-1})$ ist genau dann eine Funktion, wenn für jedes $y \in N$ genau ein $x \in M$ existiert, so dass $(y, x) \in \Gamma_\varphi^{-1}$. Nach Definition der inversen Relation Γ_φ^{-1} ist dies gleichbedeutend damit, dass es für jedes $y \in Y$ genau ein $x \in M$ gibt mit $(x, y) \in \Gamma_\varphi$, d.h. $\varphi(x) = y$. Nach Definition heißt das, dass φ bijektiv ist. $\square$

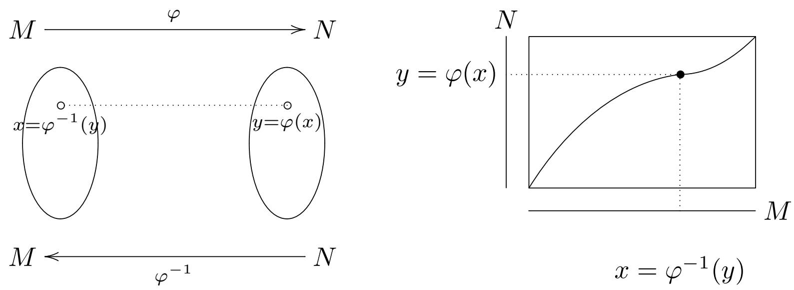
Definition 1.3.13. Erfüllt die Abbildung $\varphi\colon M \to N$ die äquivalenten Bedingungen des Satzes 1.3.12, so schreiben wir

$$\varphi^{-1}\colon = (N, M, \Gamma_\varphi^{-1}) \quad \text{and} \quad \varphi^{-1}\colon N \to M$$

für die Abbildung mit Graph Γ_φ^{-1}. Sie heißt die *Umkehrfunktion von* φ. Sie ist nur definiert, wenn φ bijektiv ist.

Im Fall der Bijektivität ist für $x \in M$ und $y \in N$ die Bedingung $(x, y) \in \Gamma_\varphi$ äquivalent zu $\varphi(x) = y$ und zu $\varphi^{-1}(y) = x$. Insbesondere haben wir

$$\varphi^{-1}(\varphi(x)) = x \quad \text{und} \quad \varphi(\varphi^{-1}(y)) = y.$$



Bemerkung 1.3.14. (a) Die Identität $\mathrm{id}_M\colon M \to M$ ist bijektiv, und $\mathrm{id}_M^{-1} = \mathrm{id}_M$.

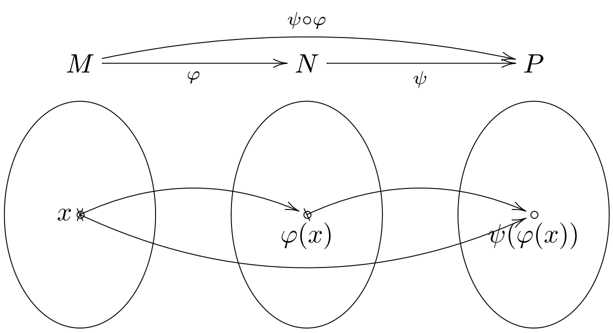
(b) Ist $\varphi\colon M \to N$ bijektiv, so ist $\varphi^{-1}\colon N \to M$ ebenfalls bijektiv, und $(\varphi^{-1})^{-1} = \varphi$.

(c) Die Abbildung $\varphi\colon \mathbb{Q} \to \mathbb{Q}, \varphi(x) = 3x+2$ ist bijektiv, mit Umkehrfunktion $\varphi^{-1} : \mathbb{Q} \to \mathbb{Q}, \varphi^{-1}(y) = \frac{y-2}{3}$.

Definition 1.3.15. Seien $\varphi\colon M \to N$ und $\psi\colon N \to P$ Abbildungen. Dann definieren wir die *Zusammensetzung* oder *Komposition* von ψ und φ als die folgende Abbildung von M nach P:

$$\psi \circ \varphi\colon M \to P, \quad x \mapsto \psi(\varphi(x)).$$

(Prüfe nach, dass dies tatsächlich eine Abbildung ist!).



Bemerkung 1.3.16. (a) Zwei Abbildungen $\varphi\colon M \to N$, $\psi\colon N' \to P$ können nur dann zusammengesetzt (komponiert) werden, wenn $N = N'$ gilt, d.h. wenn der Bildbereich der ersten gleich dem Definitionsbereich der zweiten Abbildung ist.

(b) Zwei Abbildungen einer Menge in sich selbst $\varphi\colon M \to M$, $\psi\colon M \to M$ können immer zusammengesetzt werden. Man kann also sowohl $\psi \circ \varphi$ als auch $\varphi \circ \psi$ bilden. Im allgemeinen gilt dann aber $\psi \circ \varphi \neq \varphi \circ \psi$.

(c) Für

$$\varphi\colon \mathbb{Q} \to \mathbb{Q}, \quad x \mapsto 3x + 1 \quad \text{und} \quad \psi\colon \mathbb{Q} \to \mathbb{Q}, \quad x \mapsto x^2 - 1$$

erhält man zum Beispiel

$$\psi \circ \varphi\colon \mathbb{Q} \to \mathbb{Q}, \quad x \mapsto (3x + 1)^2 - 1 = 9x^2 + 6x$$

aber

$$\varphi \circ \psi\colon \mathbb{Q} \to \mathbb{Q}, \quad x \mapsto 3(x^2 - 1) + 1 = 3x^2 - 2.$$

Die Zusammensetzung von Abbildungen ist also im allgemeinen nicht kommutativ.

Satz 1.3.17. (Regeln für die Zusammensetzung)

(i) *Für jede Abbildung* $\varphi\colon M \to N$ *gilt* $\mathrm{id}_N \circ \varphi = \varphi = \varphi \circ \mathrm{id}_M$.

(ii) *Für beliebige Abbildungen* $\varphi\colon M \to N$, $\psi\colon N \to P$ *und* $\eta\colon P \to Q$ *gilt*

$$\eta \circ (\psi \circ \varphi) = (\eta \circ \psi) \circ \varphi$$

Die Zusammensetzung von Abbildungen ist also assoziativ.

(iii) *Für Abbildungen* $\varphi\colon M \to N$ *und* $\psi\colon N \to P$ *gilt:*

- *Sind* φ *und* ψ *injektiv, so ist auch* $\psi \circ \varphi$ *injektiv.*
- *Sind* φ *und* ψ *surjektiv, so ist auch* $\psi \circ \varphi$ *surjektiv.*
- *Sind* φ *und* ψ *bijektiv, so ist auch* $\psi \circ \varphi$ *bijektiv.*

(iv) *Ist* φ *bijektiv, so ist* $\varphi^{-1} \circ \varphi = \mathrm{id}_M$ *und* $\varphi \circ \varphi^{-1} = \mathrm{id}_N$.

Beweis. (i) Für alle $x \in M$ haben wir

$$\begin{aligned} (\mathrm{id}_N \circ \varphi)(x) &= \mathrm{id}_N(\varphi(x)) = \varphi(x) \\ (\varphi \circ \mathrm{id}_M)(x) &= \varphi(\mathrm{id}_M(x)) = \varphi(x) \end{aligned}$$

Folglich $\mathrm{id}_N \circ \varphi = \varphi$ und $\varphi \circ \mathrm{id}_M = \varphi$.

(ii) Für alle $x \in M$ hat man

$$\begin{aligned} (\eta \circ (\psi \circ \varphi))(x) &= \eta((\psi \circ \varphi)(x)) = \eta(\psi(\varphi(x)) \\ &= (\eta \circ \psi)(\varphi(x)) = ((\eta \circ \psi) \circ \varphi)(x) \end{aligned}$$

Daraus folgt, dass beide Abbildungen gleich sind.

(iii) wird dem Leser als Übung überlassen.

(iv) ist klar, weil $\varphi^{-1}(\varphi(x)) = x$, $\varphi(\varphi^{-1}(y)) = y$ für alle $x \in M$, $y \in N$. $\square$

Satz 1.3.18. (Bijektivitätskriterium) *Eine Abbildung $\varphi\colon M \to N$ ist genau dann bijektiv, wenn es eine Abbildung $\psi\colon N \to M$ gibt, so dass $\psi \circ \varphi = \mathrm{id}_M$ und $\varphi \circ \psi = \mathrm{id}_N$ gilt. In diesem Fall ist $\psi = \varphi^{-1}$.*

Beweis. Eine der beiden behaupteten Implikationen ist leicht: Ist φ bijektiv, so erfüllt φ^{-1} die an ψ gestellten Bedingungen (Satz 1.3.17(iv)); es gibt also ein ψ wie verlangt.

Für die Umkehrung beweisen wir zunächst: Ist $\psi \circ \varphi = \mathrm{id}_M$, so ist φ injektiv. Seien $x_1, x_2 \in M$. Ist $\varphi(x_1) = \varphi(x_2)$, so folgt $(\psi \circ \varphi)(x_1) = \psi(\varphi(x_1)) = \psi(\varphi(x_2)) = (\psi \circ \varphi)(x_2)$ und folglich $x_1 = \mathrm{id}_M(x_1) = \mathrm{id}_M(x_2) = x_2$.

Als nächstes zeigen wir: Ist $\varphi \circ \psi = \mathrm{id}_N$, dann ist φ surjektiv. Ist nämlich $y \in N$ beliebig, so folgt $y = \mathrm{id}_N(y) = (\varphi \circ \psi)(y) = \varphi(\psi(y))$. Also $y = \varphi(x)$ mit $x = \psi(y) \in M$.

Hat man also $\psi \circ \varphi = \mathrm{id}_M$ und $\varphi \circ \psi = \mathrm{id}_N$, so ist φ injektiv und surjektiv, mithin bijektiv. Weiter ist

$$\varphi^{-1} = \varphi^{-1} \circ \mathrm{id}_N = \varphi^{-1} \circ (\varphi \circ \psi) = (\varphi^{-1} \circ \varphi) \circ \psi = \mathrm{id}_M \circ \psi = \psi. \qquad \square$$

Abbildungen endlicher Mengen

Definition 1.3.19. [Endliche Mengen] Eine Menge M heißt *endlich*, wenn sie entweder leer ist, oder wenn für ein geeeignetes $n \in \mathbb{N}$ eine bijektive Abbildung

$$\varphi\colon \{1, 2, \ldots, n\} \to M$$

existiert. Man kann zeigen, dass diese Bedingung die Zahl n eindeutig bestimmt (Übung!). Man definiert dann die *Mächtigkeit von M* oder *Kardinalität* durch

$$|M| := n.$$

Dies ist also die Anzahl der Elemente von M. Ist eine Menge nicht endlich, so heißt sie *unendlich*. In der Analysis-Vorlesung wird auch die Mächtigkeit unendlicher Mengen eingeführt, was eine wichtige Unterscheidung verschiedener Arten unendlicher Mengen erlaubt.

1.3.20. [Abbildungen auf endlichen Mengen] Anstatt beliebiger endlicher Mengen beschränken wir uns auf

$$M = \{1, 2, \ldots, m\} \quad \text{und} \quad N = \{1, 2, \ldots, n\}.$$

Eine Abbildung $\varphi\colon M \to N$ kann durch ihre *Wertetabelle* angegeben werden:

$$\varphi = \begin{pmatrix} 1 & 2 & \ldots & m \\ \varphi(1) & \varphi(2) & \ldots & \varphi(m) \end{pmatrix}$$

wo $\varphi(1), \varphi(2), \ldots, \varphi(m) \in N$. Die Anzahl aller Abbildungen von M nach N beträgt n^m. In der Tat gibt es für jeden Wert $\varphi(1), \varphi(2), \ldots, \varphi(m)$ gerade n Möglichkeiten.

Wir betrachten nun den Fall $M = N$. Sei T_n die Menge aller Abbildungen $\varphi\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$.

Beispiele: [für $n = 4$]

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad \varphi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 2 \end{pmatrix}, \quad \psi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 3 & 1 \end{pmatrix},$$
$$\varphi \circ \psi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 3 & 4 \end{pmatrix}, \quad \psi \circ \varphi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 4 \end{pmatrix}$$

Satz 1.3.21. *Für eine Abbildung $\varphi\colon M \to N$ zwischen endlichen Mengen M und N gleicher Kardinalität sind folgende Bedingungen äquivalent:*

(1) *φ ist injektiv.*

(2) *φ ist surjektiv.*

(3) *φ ist bijektiv.*

Beweis. M habe n Elemente $x_1, \ldots, x_n$. Ist $\varphi\colon M \to M$ injektiv, so sind die Elemente $\varphi(x_1), \ldots, \varphi(x_n)$ paarweise verschieden. Das Bild $\mathrm{im}(\varphi)$ hat also n verschiedene Elemente und kann daher keine echte Teilmenge von N sein. Dies beweist die Surjektivität von φ. Also impliziert (1) die Bedingung (2).

Ist φ surjektiv, so enthält das Bild von φ alle n Elemente von N. Daher müssen die $\varphi(x_1), \varphi(x_2), \ldots, \varphi(x_n)$ paarweise verschieden sein; denn andernfalls hätte man $|\,\mathrm{im}(\varphi)| < n$. Also impliziert auch (2) die Bedingung (1).

(3) ist äquivalent zur Konjunktion: (1) $\wedge$ (2). Damit ist der Beweis beendet. □

Definition 1.3.22. [Permutationen] Eine bijektive Abbildung $\pi\colon M \to M$ einer beliebigen Menge M in sich selbst heißt auch eine *Permutation* von M. Wir schreiben S_M für die Menge aller Permutationen von M.

Betrachten wir jetzt Permutationen von $M = \{1, 2, \ldots, n\}$. Wir bezeichnen mit S_n die Menge aller Permutationen von $\{1, 2, \ldots, n\}$:

$$S_3 = \left\{ \mathrm{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \tau_{12} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \tau_{13} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \right.$$
$$\left. \tau_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \zeta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \zeta^2 = \zeta^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$$

Die Anzahl aller Permutationen von drei Elementen ist 6. Allgemeiner gilt der

Satz 1.3.23. *Auf einer Menge mit n Elementen gibt es $n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdots n$ Permutationen.*

Beweis. Definieren wir eine Permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(n) \end{pmatrix}$$

Schritt für Schritt, so stellen wir fest:

- Für $\pi(1)$ können wir irgendeines der n Elemente $1, 2, \ldots, n$ wählen. Das sind n Möglichkeiten.
- Für $\pi(2)$ können wir irgendeine Zahl $1, 2, \ldots, n$ außer $\pi(1)$ wählen. Das sind $n - 1$ Möglichkeiten.
- Für $\pi(3)$ haben wir die Wahl unter $1, 2, \ldots, n$ außer den zwei verschiedenen Werten $\pi(1)$ und $\pi(2)$. Das sind $n - 2$ Möglichkeiten.
- ...
- Für $\pi(n)$ bleibt nur eine Möglichkeit über.

Alles in allem haben wir $n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 1 = n!$ mögliche Permutationen. □

Bemerkung 1.3.24. Sind π und σ Permutationen, so sind $\pi \circ \sigma$ und π^{-1} ebenfalls Permutationen. Zum Beispiel:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 1 & 5 & 2 & 6 & 4 & 3 \end{pmatrix}, \qquad \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 7 & 6 & 3 & 5 & 1 \end{pmatrix}.$$

Übungen zu Abschnitt 1.3

Übung 1.3.25. Finde Beispiele von Abbildungen $\varphi\colon M \to N$, $\psi\colon N \to M$, so dass $\psi \circ \varphi = \mathrm{id}_M$, aber φ nicht surjektiv ist.

Übung 1.3.26. Betrachte Abbildungen $\varphi\colon M \to N$ und $\psi\colon N \to P$. Beweise:

1. $\psi \circ \varphi$ injektiv $\Rightarrow \varphi$ injektiv.
2. $\psi \circ \varphi$ surjektiv $\Rightarrow \psi$ surjektiv.
3. $\psi \circ \varphi$ bijektiv $\Rightarrow \varphi$ injektiv, ψ surjektiv.
4. ψ, φ injektiv $\Rightarrow \psi \circ \varphi$ injektiv.
5. ψ, φ surjektiv $\Rightarrow \psi \circ \varphi$ surjektiv.

Übung 1.3.27. Seien M und N Mengen. Erläutere, warum die Gesamtheit aller Abbildungen $\varphi\colon M \to N$ eine Menge bildet.
Hinweis: Beschreibe die Menge der Graphen der Abbildungen als Teilmenge von $\mathbf{P}(M \times N)$.

Übung 1.3.28.

1. Ist $\varphi\colon M \to \emptyset$ eine Abbildung, so ist $M = \emptyset$.
2. Für jede Menge N ist das Tripel $\varphi = (\emptyset, N, \emptyset)$ eine Abbildung. Ihr Graph Γ ist die leere Menge. Beachte dass $\emptyset = \emptyset \times N$.

Übung 1.3.29. Existiert eine bijektive Abbildung $\varphi\colon \{1, \ldots, n\} \to \{1, \ldots, m\}$, so ist $m = n$. Hinweis: Induktion.

Kapitel 2

Lineare Gleichungssysteme

Mit diesem Kapitel betreten wir das Gebiet der Linearen Algebra. Und zwar nähern wir uns diesem Stoff über die eher algorithmisch, rechnerisch ausgerichtete Theorie der linearen Gleichungssysteme. In Kapitel 4 werden wir dann sehen, wie sich der abstraktere Aspekt der Vektorräume aus den Gleichungssystemen ergibt. Einer der Hauptpunkte dieses Kapitels 2 ist der Gauß'sche Algorithmus zur Lösung linearer Gleichungssysteme. Fast noch wichtiger ist aber die Einführung der Matrizenrechnung, zu der uns die Betrachtung linearer Gleichungssysteme führt.

2.1 Zahlen

In diesem Kapitel werden wir die folgenden Zahlenmengen benutzen, die uns teilweise schon begegnet sind:

$\mathbb{N}$	die Menge der natürlichen Zahlen:	$\{1, 2, \dots\}$
$\mathbb{Z}$	die Menge der ganzen Zahlen:	$\dots, -2, -1, 0, 1, 2, \dots$
$\mathbb{Q}$	die Menge der rationalen Zahlen:	$r = \frac{m}{n}$ mit $m \in \mathbb{Z}$, $0 \neq n \in \mathbb{N}$
$\mathbb{R}$	die Menge der reellen Zahlen:	zum Beispiel $\sqrt{2}$, π, e

Alle diese Mengen können in natürlicher Weise als Teilmengen von $\mathbb{R}$ aufgefasst werden

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}.$$

Auf all diesen Mengen sind die beiden algebraischen Operationen der Addition $+$ und Multiplikation $\cdot$ definiert, welche die folgenden Gesetze für alle $a, b, c \in \mathbb{R}$ erfüllen

Assoziativität	$a + (b + c) = (a + b) + c, \quad a(bc) = (ab)c$
Neutrales Element	$a + 0 = a, \quad a \cdot 1 = a$
Inverse Elemente	$a + (-a) = 0, \quad a \neq 0 \Rightarrow a \cdot \dfrac{1}{a} = 1$
Kommutativität	$a + b = b + a, \quad a \cdot b = b \cdot a$
Distributivität	$a(b + c) = ab + ac$

Was die Arithmetik der rationalen Zahlen angeht, sollten die Rechenregeln für Brüche (und damit auch die genannten Gesetze) aus der Schule bekannt sein. Sobald es um beliebige reelle Zahlen a, b, c geht, wird die Sache etwas komplizierter — und wird in der Analysis-Vorlesung abgehandelt.

Es ist für das Folgende üblich und hilfreich, die Elemente $x = (x_1, \ldots, x_n)$ des n-fachen kartesischen Produktes $\mathbb{R}^n$ als sogenannte *Spaltenvektoren* zu schreiben

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Die Addition und Multiplikation reeller Zahlen überträgt sich unmittelbar, komponentenweise, auf die *Addition* und *Skalarmultiplikation* von Spaltenvektoren:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{und} \quad \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda \cdot x_1 \\ \lambda \cdot x_2 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$$

für $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Demnach definieren also die Addition und die Skalarmultiplikation Abbildungen

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \quad (x, y) \mapsto x + y \quad \text{und} \quad \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, \quad (\lambda, x) \mapsto \lambda x.$$

Aus den entsprechenden Gesetzen für die Addition und Multiplikation reeller Zahlen erhalten wir sofort:

Satz 2.1.1. *Die Addition und die Skalarmultiplikation auf $\mathbb{R}^n$ genügen den folgenden Gesetzen (für alle $x, y, z \in \mathbb{R}^n$ und alle $\lambda, \mu \in \mathbb{R}$):*

Assoziativität $x + (y + z) = (x + y) + z$, $\lambda \cdot (\mu \cdot x) = (\lambda \cdot \mu) \cdot x$

Neutrale Elemente $x + 0 = x, \quad 1 \cdot x = x$.

Inverses Element $x + (-x) = 0$, *wobei* $-x := (-1) \cdot x$.

Kommutativität $x + y = y + x$

Distributivität $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$

2.2 Lineare Gleichungssysteme

Einfache geometrische Probleme, wie z.B. die Bestimmung der Schnittgeraden zweier Ebenen, führen auf lineare Gleichungssysteme. Die Lösung des Problems ist dann auf die Lösung eines solchen Gleichungssystems zurückgeführt. Dieselben Lösungsmethoden lassen sich auch auf ganz andersartig formulierte, außermathematische Probleme anwenden. Diese extreme Vielseitigkeit der Anwendungen macht die Theorie der linearen Gleichungssysteme so wichtig.

Die allgemeine Form eines linearen Gleichungssystems ist:

2.2.1. [Lineares Gleichungssystem]

$$
\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + \ldots + & a_{1j}x_j & + \ldots + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + \ldots + & a_{2j}x_j & + \ldots + & a_{2n}x_n & = & b_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{i1}x_1 & + & a_{i2}x_2 & + \ldots + & a_{ij}x_j & + \ldots + & a_{in}x_n & = & b_i \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + \ldots + & a_{mj}x_j & + \ldots + & a_{mn}x_n & = & b_m
\end{array}
\tag{2.1}
$$

Dabei bezeichnen die a_{ij} (für $i = 1, 2, \ldots, m$ und $j = 1, 2, \ldots, n$) und die b_1 ,b_2, $\ldots$, b_m reelle Zahlen. Gesucht sind Zahlen x_1, x_2, $\ldots$, x_n, die alle m Gleichungen gleichzeitig erfüllen. Ein solches n-Tupel $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ heißt dann eine *Lösung des Gleichungssystems*. Motiviert durch die Matrizenrechnung, die wir weiter unten einführen werden, betrachten wir eine Lösung meistens als Spaltenvektor

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

Ist ein lineares Gleichungssystem gegeben, so treten folgende Fragen auf:

- Ist es lösbar, d.h. gibt es (mindestens) eine Lösung x_1,x_2, $\ldots$, x_n?
- Angenommen, das System ist lösbar, gibt es genau eine Lösung?
- Angenommen, es gibt mehr als eine Lösung, beschreibe die Menge aller Lösungen.
- Wie können die Lösungen praktisch berechnet werden?

In dieser Vorlesung werden die theoretischen Aspekte im Vordergrund stehen, während die rechentechnischen Gesichtspunkte einer Vorlesung über numerische Methoden vorbehalten bleiben.

Lineare Gleichungssysteme wären nicht so schwierig, wenn man immer nur Systeme mit wenigen Gleichungen in wenigen Unbekannten x_j $(j = 1, 2, \ldots, m)$ zu betrachten hätte. In der Praxis treten heute aber Systeme von 10 000 oder mehr Gleichungen in ebenso vielen Unbekannten auf. Da ist eine solide mathematische Theorie dann unerlässlich.

Beispiel 2.2.2. Ein Betrieb verwendet m verschiedene Rohmaterialien und stellt daraus n verschiedene Produkte her. Um eine Einheit des j-ten Produkts herzustellen, braucht der Betrieb a_{ij} Einheiten des i-ten Rohmaterials. Wir stellen die Zahlen a_{ij} in einer Tabelle dar:

	Produkt 1	Prod. 2	......	Prod. j	......	Prod. n
Material 1	a_{11}	a_{12}	......	a_{1j}	......	a_{1n}
Material 2	a_{21}	a_{22}	......	a_{2j}	......	a_{2n}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
Material i	a_{i1}	a_{i2}	......	a_{ij}	......	a_{in}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
Material m	a_{m1}	a_{m2}	......	a_{mj}	......	a_{mn}

Will man dann x_1 Einheiten des Produkts 1, x_2 Einheiten des Produkts 2, ..., x_n Einheiten des Produkts n herstellen, erlauben uns unsere Gleichungen, die Mengen $y_1, y_2, \ldots, y_m$ der verschiedenen benötigten Rohmaterialien $1, 2, \ldots, m$ zu berechnen.

Umgekehrt kann und wird es aber vorkommen, dass die verfügbaren Mengen an Rohmaterial vorgegeben sind — sagen wir, $y_1, y_2, \ldots, y_m$, und die Frage ist dann, wieviele Einheiten $x_1, x_2, \ldots, x_n$ der Produkte $1, 2, \ldots, n$ mit diesem Vorrat hergestellt werden können. Das ist die Frage nach den Lösungen $(x_1, x_2, \ldots, x_n)$ unseres linearen Gleichungssystems aus m Gleichungen in n Unbekannten:

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + \ldots + & a_{1j}x_j & + \ldots + & a_{1n}x_n & = & y_1 \\
a_{21}x_1 & + & a_{22}x_2 & + \ldots + & a_{2j}x_j & + \ldots + & a_{2n}x_n & = & y_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{i1}x_1 & + & a_{i2}x_2 & + \ldots + & a_{ij}x_j & + \ldots + & a_{in}x_n & = & y_i \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + \ldots + & a_{mj}x_j & + \ldots + & a_{mn}x_n & = & y_m
\end{array} \tag{L}$$

2.3 Einfache Fälle linearer Gleichungssysteme

Um schon einmal ein Gefühl für mögliche Lösungsmengen linearer Gleichungssysteme zu bekommen, betrachten wir zunächst einige Beispiele.

Bemerkung 2.3.1. [Einige einfache Fälle]

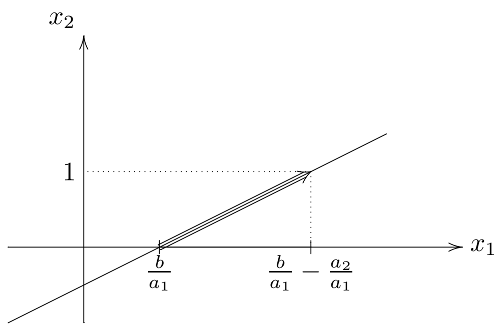
- $m = 1$, $n = 1$:

$$ax_1 = b$$

Drei verschiedene Fälle sind zu unterscheiden:

1. $a \neq 0$: Dann ist $x_1 = \frac{b}{a}$ die einzige Lösung.
2. $a = 0$, $b \neq 0$: Es gibt überhaupt keine Lösung.
3. $a = 0$, $b = 0$: Jedes $x_1 \in \mathbb{R}$ ist eine Lösung.

- $m = 1$, $n = 2$:
$$a_1x_1 + a_2x_2 = b$$
Hier können vier Fälle auftreten:

1. $a_1 \neq 0$: man kann $x_2 = \lambda \in \mathbb{R}$ beliebig wählen und dann x_1 vermöge $x_1 = \frac{b}{a_1} - \frac{a_2}{a_1}\lambda$ berechnen. Die Lösungen sind also
$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{b}{a_1} - \frac{a_2}{a_1}\lambda \\ \lambda \end{pmatrix} = \begin{pmatrix} \frac{b}{a_1} \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \end{pmatrix}$$
Dies ist eine Parameterdarstellung derjenigen Geraden durch den Punkt $(\frac{b}{a_1}, 0)$, deren *Richtung* durch den folgenden Vektor festgelegt ist:
$$v := \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \end{pmatrix}:$$


2. $a_1 = 0$, $a_2 \neq 0$: Wähle $x_1 = \lambda \in \mathbb{R}$ beliebig und berechne $x_2 = \frac{b}{a_2}$
$$x = \begin{pmatrix} \lambda \\ \frac{b}{a_2} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{b}{a_2} \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
Bei variierendem λ beschreibt dies eine horizontale Gerade durch den Punkt $(0, \frac{b}{a_2})$:


3. $a_1 = 0$, $a_2 = 0$, $b \neq 0$: Es gibt keine Lösung.
4. $a_1 = 0$, $a_2 = 0$, $b = 0$: Alle Punkte $\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \in \mathbb{R}^2$ sind Lösungen.

- $m = 2$, $n = 2$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Multiplizieren wir die erste Gleichung mit a_{22} und die zweite mit a_{12}, so erhalten wir

$$\begin{aligned} a_{22}a_{11}x_1 + a_{22}a_{12}x_2 &= a_{22}b_1 \\ a_{12}a_{21}x_1 + a_{12}a_{22}x_2 &= a_{12}b_2, \end{aligned}$$

Subtrahieren wir die zweite Gleichung von der ersten, so ergibt sich

$$(a_{11}a_{22} - a_{12}a_{21})\ x_1 = a_{22}b_1 - a_{12}b_2.$$

Falls demnach die Zahl

$$d := a_{11}a_{22} - a_{12}a_{21}$$

nicht null ist, so erhalten wir die eindeutige Festlegung von x_1:

$$x_1 = \frac{1}{d}\ (a_{22}b_1 - a_{12}b_2).$$

Analog ist auch x_2 im Fall $d \neq 0$ eindeutig bestimmt. Multiplizieren wir nämlich in unserem Gleichungssystem die erste Gleichung mit a_{21} und die zweite mit a_{11}, so erhalten wir

$$\begin{aligned} a_{21}a_{11}x_1 + a_{21}a_{12}x_2 &= a_{21}b_1 \\ a_{11}a_{21}x_1 + a_{11}a_{22}x_2 &= a_{11}b_2, \end{aligned}$$

woraus durch Abziehen der ersten von der zweiten Gleichung folgt:

$$x_2 = \frac{1}{d}\ (a_{11}b_2 - a_{21}b_1).$$

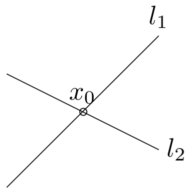
Nehmen wir an, dass in jeder Gleichung einer der Koeffizienten a_{ij} nicht null ist. Dann bilden die Lösungen der ersten Gleichung eine Gerade l_1 in $\mathbb{R}^2$, und die Lösungen der zweiten Gleichung bilden eine Gerade l_2. Die Lösungen des Gleichungssystems sind dann die Schnittpunkte der beiden Geraden l_1, l_2.

Sind a_{11} und a_{21} nicht null, so sind die Richtungen von l_1 und l_2 durch die Vektoren

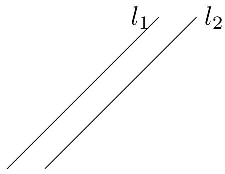
$$v_1 := \begin{pmatrix} -\frac{a_{12}}{a_{11}} \\ 1 \end{pmatrix} \quad \text{bzw.} \quad v_2 := \begin{pmatrix} -\frac{a_{22}}{a_{21}} \\ 1 \end{pmatrix},$$

gegeben, so dass also die beiden Geraden genau dann parallel sind, wenn $v_1 = v_2$ ist. Dies ist äquivalent zu $a_{12}a_{21} = a_{11}a_{22}$, was wiederum zu $d = 0$ äquivalent ist. Die geometrische Konfiguration der beiden Geraden gehört dann zu einem der folgenden drei Fälle:

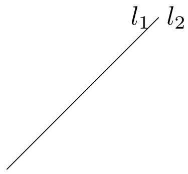
1. l_1 und l_2 schneiden sich in einem Punkt (eindeutige Lösung):



2. Die beiden Geraden sind parallel und voneinander verschieden (keine Lösung):



3. Die beiden Geraden sind identisch (alle Geradenpunkte sind Lösungen):



- $m = 1, n = 3$:

$$a_1x_1 + a_2x_2 + a_3x_3 = b$$

Angenommen einer der Koeffizienten a_i ist $\neq 0$, etwa $a_1 \neq 0$. Dann gilt $x_1 = \frac{1}{a_1}(b - a_2x_2 - a_3x_3)$. Wir können $x_2 = \lambda$ und $x_3 = \mu$ beliebig wählen und erhalten $x_1 = \frac{b}{a_1} - \frac{a_2}{a_1}\lambda - \frac{a_3}{a_1}\mu$.

Mithin sind alle Lösungen der Gleichung gegeben durch

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{b}{a_1} - \frac{a_2}{a_1}\lambda - \frac{a_3}{a_1}\mu \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \frac{b}{a_1} \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -\frac{a_2}{a_1} \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -\frac{a_3}{a_1} \\ 0 \\ 1 \end{pmatrix}$$

Dies ist eine parametrische Darstellung einer Ebene $x = a + \lambda v + \mu w$.

- $m = 3, n = 3$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

Angenommen in jeder Gleichung gibt es einen Koeffizienten $a_{ij} \neq 0$. Wie wir aus obigen Fällen wissen, bildet die Lösungsmenge einer jeden Gleichung eine Ebene im $\mathbb{R}^3$.

Die gemeinsamen Lösungen aller drei Gleichungen sind die Punkte, die im Durchschnitt aller drei Ebenen E_1, E_2, E_3 liegen. Es können mehrere Fälle auftreten, siehe Abbildung 2.1.

1. Die drei Ebenen sind parallel, aber nicht identisch: keine Lösung.
2. Zwei der Ebenen sind parallel, aber nicht die dritte: keine Lösung.
3. Die drei Ebenen sind identisch: die Lösungen sind die Punkte dieser Ebene.
4. Die drei Ebenen schneiden sich in einer Geraden: die Lösungen sind die Punkte dieser Geraden.
5. Die drei Ebenen schneiden sich in einem Punkt: Dieser Punkt ist die eindeutige Lösung.

Zur Berechnung dieser Lösungen können wir den Gauß-Jordan Eliminations-Algorithmus verwenden, dem wir uns als nächstes zuwenden.

2.4 Gauss–Jordan Eliminations-Algorithmus

Der Gauß [1]-Jordan[2]-Algorithmus[3] ist ein systematisches Rechenverfahren zum Lösen linearer Gleichungssysteme, das auch von großem praktischen Interesse ist. Er basiert auf der folgenden
BEOBACHTUNG: Die folgenden drei Operationen transformieren ein gegebenes lineares Gleichungssystem in ein anderes Gleichungssystem mit derselben Lösungsmenge (NACHPRÜFEN !):

(E1) Addition eines skalaren Vielfachen einer Gleichung zu einer anderen Gleichung.

(E2) Vertauschen zweier Gleichungen.

(E3) Multiplikation einer Gleichung mit einem Skalar $\lambda \neq 0$.

Diese drei Operationen heißen *elementare (Zeilen-)Umformungen* eines linearen Gleichungssystems.

Genauer gesagt werden wir **zwei Varianten** des Eliminations-Algorithmus besprechen: **die erste** jetzt gleich — sie reicht zur effektiven Behandlung beliebiger linearer Gleichungssystme im Prinzip aus und wird uns zudem eine genaue

[1] Carl Friedrich Gauß (1777-1855), Deutscher Mathematiker, der in Göttingen gewirkt hat. Hat zu allen Gebieten der Mathematik seiner Zeit bahnbrechende Beiträge geliefert — der untere Seitenrand hier ist zu klein, um davon eine detailliertere Vorstellung zu vermitteln.

[2] Camille Jordan (1838–1922), Französischer Mathematiker, der neben seinen bedeutenden Forschungen auf dem Gebiet der Algebra gegen Ende des 19. Jahrhunderts auch wesentlich die damalige Erneuerung der französischen Analysis-Lehrbücher durch seinen *Cours d'Analyse* mitprägte.

[3] Den Algorithmus, um den es hier geht, haben natürlich auch bedeutende Mathematiker wie Gauß und Jordan immer mal wieder benutzt, und ihre gemeinsame Nennung ist ein Beitrag zur deutsch-französischen Freundschaft. Gauß nannte dieses Rechenverfahren *eliminatio vulgaris*, was nicht unbedingt schmeichelhaft ist. Die wohl früheste Darstellung des Algorithmus stammt aus einem chinesischen Werk, dem *Chui-Chang suan-shu*, oder *Neun Kapitel der mathematischen Kunst*, das allerdings kein systematisches Lehrbuch sondern eine große Aufgabensammlung mit Musterlösungen war. Im achten Kapitel werden lineare Gleichungssysteme muster-gelöst; das Verfahren heißt dort *Fang-Cheng*. Das *Chui-Chang suan-shu* ist wohl um 250 vor Christus, kurz vor der Han-Dynastie, entstanden.

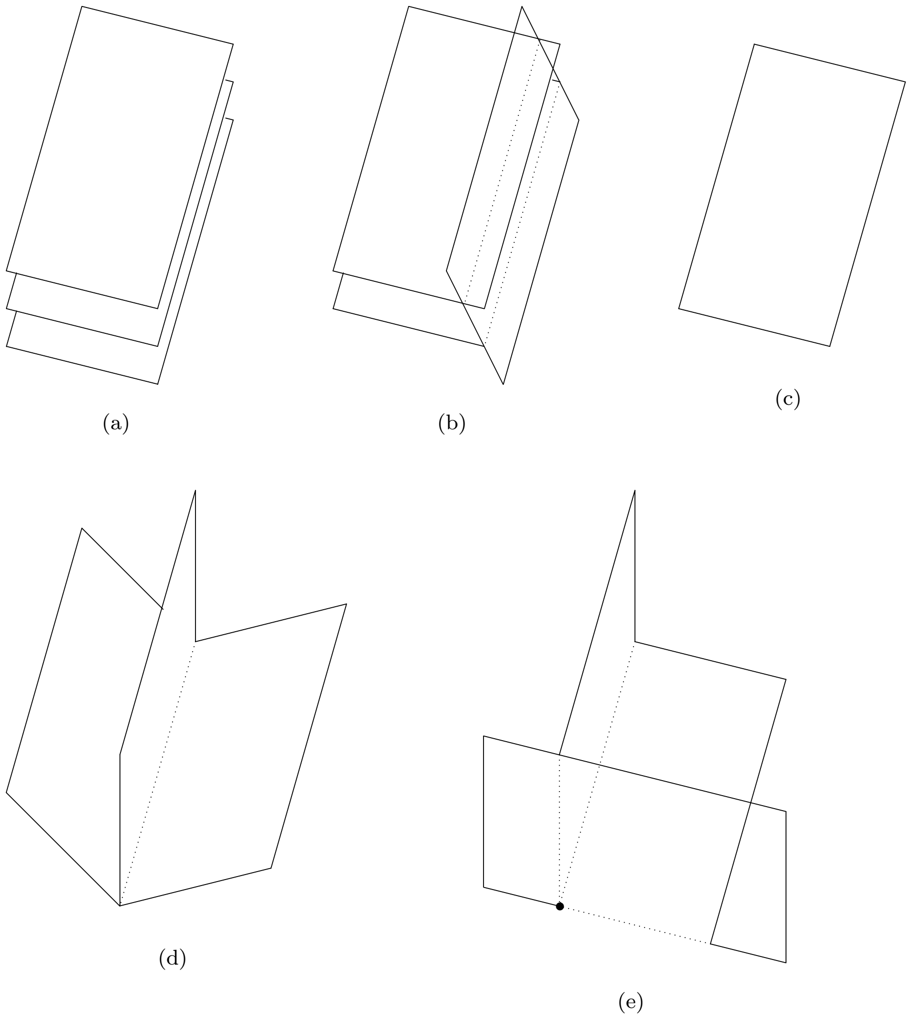


Abbildung 2.1: Mögliche Schnittmengen dreier Ebenen

und vollständige Übersicht über *alle* Lösungen des gegebenen Systems geben —, und **die zweite** in Abschnitt 2.5 unten, wenn wir die effektive Invertierung quadratischer Matrizen besprechen. Ein technischer Unterschied zwischen beiden Varianten ist: *die erste Variante benutzt nur die elementaren Transformauionen (E1) und (E2); bei der zweiten Variante wird i.a. auch (E3) herangezogen.*

Gauß-Algorithmus, erste Variante

Im ERSTEN SCHRITT des Gauß-Algorithmus (dies gilt ebenso für die erste wie die zweite Variante) transformiert man das Originalsystem (2.1) in 2.2.1 durch elementare Umformungen vom Typ (E1) und (E2) in ein System (L') der folgenden Form:

$$\begin{array}{rcrcrcrcl} a'_{11}x_1 + & a'_{12}x_2 & + \ldots + & a'_{1j}x_j & + \ldots + & a'_{1n}x_n & = & b'_1 \\ & a'_{22}x_2 & + \ldots + & a'_{2j}x_j & + \ldots + & a'_{2n}x_n & = & b'_2 \\ & \vdots & & \vdots & & \vdots & & \vdots \\ & a'_{i2}x_2 & + \ldots + & a'_{ij}x_j & + \ldots + & a'_{in}x_n & = & b'_i \\ & \vdots & & \vdots & & \vdots & & \vdots \\ & a'_{m2}x_2 & + \ldots + & a'_{mj}x_j & + \ldots + & a'_{mn}x_n & = & b'_m \end{array} \qquad (\mathrm{L}')$$

Dies kann man auf folgende Weise erreichen:

- Ist $a_{11} \neq 0$, so lasse die erste Gleichung unverändert und subtrahiere

  $\frac{a_{21}}{a_{11}}$-mal die erste Gleichung von der zweiten
  $\frac{a_{31}}{a_{11}}$-mal die erste Gleichung von der dritten
  $\vdots$
  $\vdots$
  $\frac{a_{m1}}{a_{11}}$-mal die erste Gleichung von der letzten Gleichung.

- Ist $a_{11} = 0$, so suche eine Gleichung, in der der erste Koeffizient $a_{i1} \neq 0$; tausche die erste Gleichung mit der i-ten Gleichung und fahre wie im ersten Schritt fort.

- Ist $a_{i1} = 0$ für alle $i = 1, \ldots, m$, dann gibt es nichts zu tun und wir fahren mit dem zweiten Schritt fort.

Im ZWEITEN SCHRITT lassen wir die erste Gleichung unberührt und für die verbleibenden $m-1$ Gleichungen in den verbleibenden $n-1$ Unbekannten fahren wir wie im ersten Schritt fort.

Nach höchstens $m-1$ Schritten enden wir mit einem System von *stufen-*

förmiger Gestalt:

$$
\begin{array}{rcl}
\gamma_{1j_1}x_{j_1} + \ldots + \gamma_{1j_2}x_{j_2} + \ldots + \gamma_{1j_r}x_{j_r} + \ldots + \gamma_{1n}x_n &=& \delta_1 \\
\gamma_{2j_2}x_{j_2} + \ldots + \gamma_{2j_r}x_{j_r} + \ldots + \gamma_{2n}x_n &=& \delta_2 \\
\ddots \qquad\qquad \vdots && \vdots \\
\gamma_{rj_r}x_{j_r} + \ldots + \gamma_{rn}x_n &=& \delta_r \\
0 &=& \delta_{r+1} \\
\vdots && \vdots \\
0 &=& \delta_m
\end{array}
\tag{M}
$$

Hierbei sind die Elemente $\gamma_{1j_1}, \gamma_{2j_2}, \ldots, \gamma_{rj_r}$ alle von Null verschieden; man nennt sie *Pivotelemente*. Die zugehörigen Variablen $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ heißen die *Pivotvariablen*. Im allgemeinen werden nach der Transformation in Stufenform die linken Seiten einiger Gleichungen Null sein.

Diese Stufenform (M) ist das Ziel der ersten Variante des Gauß-Algorithmus. Ist sie erreicht, ist der Durchgang des Algorithmus beendet.

Definition 2.4.1. (*Vorläufige Definition des Rangs.*) Die Anzahl r der Gleichungen in (M), deren linke Seite nicht verschwindet, nennt man den *Rang* des linearen Gleichungssystems (2.1).

AUFGEPASST: Diese Definition ist eigentlich nur dann zulässig, wenn wir nachgewiesen haben, dass diese Anzahl r nicht von der Art und Weise abhängt, in der wir unseren Algorithmus durchgeführt haben. (In jedem Schritt können Zeilenvertauschungen nötig werden, und i.allg. werden verschiedene Zeilenvertauschungen zum Ziel führen. D.h. man kann durch den Gauß–Jordan–Algorithmus auf verschiedenen Wegen zu Matrizen der Form (M) gelangen.) Wir ignorieren diesen Punkt hier aus drei guten Gründen: Erstens reicht es aus, die folgenden Überlegungen für einen festen Durchlauf des Algorithmus anzustellen. Zweitens werden sich gleich Aussagen über die Lösungsmenge in Abhängigkeit von r ergeben, die zeigen, dass r nur von dem System (2.1), und nicht vom Durchlauf des Algorithmus, abhängt. Drittens werden wir bald, bei der Behandlung der Matrizen, eine andere Definition des Ranges an die Stelle der obigen setzen, die diese Schwierigkeiten auflösen wird.

Wenn wir das Gleichungssystem einmal in Stufenform (M) umgeformt haben, ist es leicht zu lösen:

1. Fall. Eine der rechten Seiten $\delta_{r+1}, \ldots, \delta_m$ ist ungleich Null. In diesem Fall widersprechen sich die Gleichungen; das Gleichungssystem hat also keine Lösung.

2. Fall. $\delta_{r+1} = \cdots = \delta_m = 0$. In diesem Fall hat das Gleichungssystem mindestens eine Lösung. Und wir können *alle* Lösungen wie folgt angeben: Die Werte der $n - r$ Variablen x_i, die keine Pivotvariablen sind, können *beliebig* gewählt werden:

$$x_i = \lambda_i \quad \text{für} \quad 1 \leq i \leq n;\ i \neq j_1, \ldots, i \neq j_r.$$

Für jede solche Wahl bestimmt man die Werte der Pivotvariablen durch Rückwärtseinsetzen, indem man von unten her, d.h. mit der folgenden Formel beginnt:

$$x_{j_r} = \frac{1}{\gamma_{rj_r}} \left(\delta_r - \gamma_{r(j_r+1)}\lambda_{j_r+1} - \cdots - \gamma_{rn}\lambda_n\right)$$

Nach r Schritten bekommt man so eine Formel für x_{j_1}.

Beachte, dass die Stufenform in jedem Fall $r \leq n$ impliziert. Damit haben wir (in Bezug auf jede Durchführung des Gauß'schen Algorithmus und den zugehörigen Wert von r) folgenden wichtigen Satz gezeigt, dessen „entweder — oder“ die obige Unterscheidung von Fall 1 und Fall 2 wieder aufnimmt:

Satz 2.4.2. *Sei r der Rang eines linearen Gleichungssystems mit m Gleichungen in n Variablen. Das System hat entweder gar keine Lösung, oder die Werte von $n - r$ Variablen können beliebig gewählt werden.*

Bemerkung 2.4.3. In folgenden Fällen hat das System (2.1) immer mindestens eine Lösung:

1. $r = m$, d.h. der Rang ist gleich der Anzahl Gleichungen: In diesem Fall verschwindet bei keiner Gleichung die linke Seite.

2. Das Gleichungssystem (2.1) ist *homogen*, d.h. $b_1 = b_2 = \cdots = b_m = 0$. Ein homogenes System hat immer eine Lösung, nämlich $x_1 = x_2 = \cdots = x_n = 0$, die sogenannte *triviale* Lösung.

Schließlich können wir die Fälle genau charakterisieren, in denen das System (2.1) eine und nur eine Lösung hat. Um das Kriterium schlichter formulieren zu können, beschränken wir uns auf den Fall $n = m$, in dem es ebensoviele Gleichungen wie Unbekannte gibt:

Satz 2.4.4. *Ein Gleichungssystem mit n linearen Gleichungen in n Unbekannten hat genau dann eine eindeutige Lösung, wenn der Rang $r = n$ ist.*

Beispiel 2.4.5. Betrachten wir ein System mit 5 Gleichungen in 5 Unbekannten. Wir wollen das System für verschiedene rechte Seiten lösen:

1: $b_1 = b_2 = b_3 = b_4 = b_5 = 0$
2: $b_1 = b_2 = b_3 = b_4 = b_5 = 1$
3: $b_1 = b_2 = b_3 = 1, b_4 = 4, b_5 = 2$

	1	2	3	x_1	x_2	x_3	x_4	x_5	b_1	b_2	b_3	b_4	b_5
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = b_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-x_1 - x_2 - 2x_3 + x_4 + x_5 = b_2$	0	1	1	−1	−1	−2	1	1	0	1	0	0	0
$2x_1 + 4x_3 + 2x_4 + x_5 = b_3$	0	1	1	2	0	4	2	1	0	0	1	0	0
$x_1 - 3x_2 + 2x_3 + 6x_4 + 6x_5 = b_4$	0	1	4	1	−3	2	6	6	0	0	0	1	0
$-x_1 - 3x_2 - 2x_3 + 3x_4 + 6x_5 = b_5$	0	1	2	−1	−3	−2	3	6	0	0	0	0	1
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = b_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-2x_2 + 3x_4 + 4x_5 = b_2 + b_1$	0	2	2	0	−2	0	3	4	1	1	0	0	0
$2x_2 - 2x_4 - 5x_5 = b_3 - 2b_1$	0	−1	−1	0	2	0	−2	−5	−2	0	1	0	0
$-2x_2 + 4x_4 + 3x_5 = b_4 - b_1$	0	0	3	0	−2	0	4	3	−1	0	0	1	0
$-4x_2 + 5x_4 + 9x_5 = b_5 + b_1$	0	2	3	0	−4	0	5	9	1	0	0	0	1
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = b_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-2x_2 + 3x_4 + 4x_5 = b_2 + b_1$	0	2	2	0	−2	0	3	4	1	1	0	0	0
$x_4 - x_5 = b_3 + b_2 - b_1$	0	1	1	0	0	0	1	−1	−1	1	1	0	0
$x_4 - x_5 = b_4 - b_2 - 2b_1$	0	−2	1	0	0	0	1	−1	−2	−1	0	1	0
$-x_4 + x_5 = b_5 - 2b_2 - b_1$	0	−2	−1	0	0	0	−1	1	−1	−2	0	0	1
$x_1 - x_2 + 2x_3 + 2x_4 + 3x_5 = b_1$	0	1	1	1	−1	2	2	3	1	0	0	0	0
$-2x_2 + 3x_4 + 4x_5 = b_2 + b_1$	0	2	2	0	−2	0	3	4	1	1	0	0	0
$x_4 - x_5 = b_3 + b_2 - b_1$	0	1	1	0	0	0	1	−1	−1	1	1	0	0
$0 = b_4 - b_3 - 2b_2 - b_1$	0	−3	0	0	0	0	0	0	−1	−2	−1	1	0
$0 = b_5 + b_3 - b_2 - 2b_1$	0	−1	0	0	0	0	0	0	−2	−1	1	0	1

Der rechte Teil obiger Tabelle ist nochmals dargestellt in Abbildung 2.2. Sie zeigt ein Schema, in dem man nur die Koeffizienten der x_i und der b_i darstellt, während die Gleichungen transformiert werden.

x_1	x_2	x_3	x_4	x_5	b_1	b_2	b_3	b_4	b_5
1	−1	2	2	3	1				
−1	−1	−2	1	1		1			
2	0	4	2	1			1		
1	−3	2	6	6				1	
−1	−3	−2	3	6					1
1	−1	2	2	3	1				
0	−2	0	3	4	1	1			
0	2	0	−2	−5	−2		1		
0	−2	0	4	3	−1			1	
0	−4	0	5	9	1				1
1	−1	2	2	3	1				
0	−2	0	3	4	1	1			
0	0	0	1	−1	−1	1	1		
0	0	0	1	−1	−2	−1		1	
0	0	0	−1	1	−1	−2			1
1	−1	2	2	3	1				
0	−2	0	3	4	1	1			
0	0	0	1	−1	−1	1	1		
0	0	0	0	0	−1	−2	−1	1	
0	0	0	0	0	−2	−1	1	0	1

Stellen, die in diesem Schema leer bleiben, sind so zu verstehen, dass sie den Eintrag 0 haben.

Abbildung 2.2: Schema für das Beispiel aus 2.4.5.

Die Lösungen für die drei Fälle können nun wie folgt bestimmt werden:

1. $x_1 = -\frac{3}{2}\lambda - 2\mu$
   $x_2 = \frac{7}{2}\lambda$
   $x_3 = \mu$ (beliebig)
   $x_4 = \lambda$
   $x_5 = \lambda$ (beliebig)
   Die Lösungsmenge ist also:
$$x = \begin{pmatrix} -\frac{3}{2}\lambda - 2\mu \\ \frac{7}{2}\lambda \\ \mu \\ \lambda \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} -\frac{3}{2} \\ \frac{7}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad (\lambda, \mu \text{ beliebig})$$

2. Es gibt keine Lösung; die Umformungen führen zu zwei widersprüchlichen Gleichungen $0 = -3$ und $0 = -1$.

3. Die letzten beiden Gleichungen ergeben $0 = 0$. Wir können $x_5 = \lambda$ und $x_3 = \mu$ beliebig wählen und erhalten die Lösungen:
$$x = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -\frac{3}{2} \\ \frac{7}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad (\lambda, \mu \text{ beliebig})$$

Bemerkungen:

- Der Rang des Systems ist $r = 3$.
- Die Pivotvariablen sind x_1, x_2, x_4.
- x_3, x_5 können beliebig gewählt werden.
- Die Lösungen im homogenen Fall (1) bilden eine Ebene im $\mathbb{R}^5$ durch den Ursprung $0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$.
- Die Menge der Lösungen im inhomogenen Fall kann leer sein (2) oder eine Ebene im $\mathbb{R}^5$ sein, die "parallel" zu der Ebene im homogenen Fall (3) ist.
- Im homogenen Fall (1) bilden die zwei Vektoren
$$v = \begin{pmatrix} -\frac{3}{2} \\ \frac{7}{2} \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad w = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

ein sogenanntes *Fundamentalsystem von Lösungen*: Alle anderen Lösungen sind von der Form $\lambda \cdot v + \mu \cdot w$, für beliebige $\lambda, \mu \in \mathbb{R}$.

2.5 Matrizen

Im vorangehenden Beispiel 2.4.5 haben wir gesehen, dass es zur Lösung von Gleichungssystemen nicht notwendig ist, mit den Gleichungen selbst zu arbeiten: die Tabellen der Koeffizienten und Konstanten reichen aus. Solche Tabellen von Koeffizienten nennt man **Matrizen**.

2.5.1 Rechnen mit Matrizen

2.5.1. [$\mathbf{m} \times \mathbf{n}$-Matrizen mit reellen Einträgen] Eine *reelle $m \times n$-Matrix* ist ein m-tupel von n-Tupeln reeller Zahlen, d.h. ein Element von $(\mathbb{R}^n)^m$. Man schreibt Matrizen übersichtlicherweise als rechteckiges Schema A reeller Zahlen a_{ij} mit m Zeilen und n Spalten. Wir schreiben:

$$A = (a_{ij}) = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2j} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mj} & \ldots & a_{mn} \end{pmatrix}$$

Die a_{ij} heißen die *Einträge* der Matrix. Der Index i bezeichnet die *Zeile* und der Index j die *Spalte*, in der der Eintrag a_{ij} steht. Wir bezeichnen mit $M_{m,n}(\mathbb{R})$, die Menge aller $m \times n$-Matrizen A mit reellen Einträgen.

Für ein lineares Gleichungssystem (2.1) nennt man die $m \times n$-Matrix $A = (a_{ij})$ die *Koeffizientenmatrix* von (2.1).

2.5.2. [Spezielle Matrizen]

- $m \times 1$-Matrizen sind (Spalten-) Vektoren

$$v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \in \mathbb{R}^m.$$

- $1 \times n$-Matrizen sind (Zeilen-) Vektoren mit n Einträgen:

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_n \end{pmatrix}.$$

- Die $m \times n$-Nullmatrix ist die Matrix, deren Einträge alle Null sind:

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

- $n \times n$-Matrizen heißen auch *quadratische* Matrizen:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ii} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{nn} \end{pmatrix}$$

  In einer quadratischen Matrix bilden die Einträge $a_{11}, a_{22}, \dots, a_{nn}$ die sogenannte *(Haupt-)Diagonale* der Matrix. Die Menge aller quadratischen Matrizen mit reellen Koeffizienten, mit n Zeilen und n Spalten bezeichnen wir mit $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$.

- Die $n \times n$-Matrix mit Einsen in der Hauptdiagonalen und Nullen sonst heißt $n \times n$-*Einheitsmatrix* $E = E_n$. Genauer:

$$E_n = (\delta_{ij})_{n \times n}, \text{ mit dem Kronecker}^4\text{-Symbol } \delta_{ij} = \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j \end{cases}$$

$$\text{Zum Beispiel } E_1 = (1), \quad E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

- Unter den quadratischen Matrizen $A = (a_{ij})$ unterscheiden wir

  Diagonalmatrizen, $a_{ij} = 0$ für $i \neq j$,
  obere Dreiecksmatrizen, $a_{ij} = 0$ für $i > j$,
  untere Dreiecksmatrizen, $a_{ij} = 0$ für $i < j$:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \ddots & \vdots \\ 0 & & & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & & & 0 \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

[4] Leopold Kronecker (1823–1891), sehr vielseitiger und produktiver Berliner Mathematiker, der besonders die Algebra, Zahlentheorie und die Theorie der elliptischen Funktionen wesentlich bereichert hat — also für wesentlich mehr in Erinnerung behalten wird als nur für dieses Null-Eins-Symbol.

2.5.3. [Addition von $(m \times n)$-Matrizen und Skalarmultiplikation]
Ähnlich wie in unserer Definition von Vektoren definieren wir die Summe zweier $m \times n$-Matrizen $A = (a_{ij})$ und $B = (b_{ij})$ und die Skalarmultiplikation mit $\lambda \in \mathbb{R}$ Eintrag für Eintrag:

$$A + B := (a_{ij} + b_{ij}), \qquad \lambda A := (\lambda a_{ij}).$$

Genauer:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & \dots & a_{2n} + b_{2n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

$$\lambda \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} := \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix}.$$

Die Rechenregeln aus Satz 2.1.1 sind für Vektoren und daher auch für Matrizen gültig.

2.5.4. [Beispiel]

$$\begin{aligned} 4\begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & -1 \end{pmatrix} - 3\begin{pmatrix} 1 & -1 & 0 \\ 0 & -2 & 1 \end{pmatrix} &= \begin{pmatrix} 8 & 4 & 12 \\ 4 & 0 & -4 \end{pmatrix} + \begin{pmatrix} -3 & 3 & 0 \\ 0 & 6 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 5 & 7 & 12 \\ 4 & 6 & -7 \end{pmatrix} \end{aligned}$$

2.5.5. [Achtung] Die Addition ist nur für Matrizen derselben Größe $m \times n$ definiert.

Die folgende, für uns neuartige Operation erweitert die Nützlichkeit der Matrizen erheblich. Wir lassen sie hier vom Himmel fallen—die Motivation für diese auf den ersten Blick sicher überraschende Definition des Matrizenprodukts wird sich etwas später aus der Theorie der linearen Abbildungen ergeben.

2.5.6. [Matrizenmultiplikation] Das Produkt einer $m \times n$-Matrix $A = (a_{ij})$ und einer $n \times p$-Matrix $B = (b_{jk})$ ist definiert als die $m \times p$-Matrix $C = (c_{ik})$ mit den Einträgen

$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \dots + a_{in} b_{nk},$$

genauer:

$$
\begin{array}{cc}
i \rightarrow & \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}
\end{array}
\cdot
\begin{array}{c}
k \\
\begin{pmatrix} b_{11} & \dots & b_{1k} & \dots & b_{1p} \\ b_{21} & \dots & b_{2k} & \dots & b_{2p} \\ \vdots & & \vdots & & \vdots \\ b_{j1} & \dots & b_{jk} & \dots & b_{jp} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nk} & \dots & b_{np} \end{pmatrix}
\end{array}
=
$$

$$
= \begin{array}{cc}
 & k \\
i \rightarrow & \begin{pmatrix} c_{11} & \dots & c_{1k} & \dots & c_{1p} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \dots & c_{ik} & \dots & c_{ip} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \dots & c_{mk} & \dots & c_{mp} \end{pmatrix}
\end{array}
$$

Merke: $c_{ik} = (i\text{-te Zeile von } A) \cdot (k\text{-te Spalte von } B)$, wobei Spalten als $(n \times 1)$- und Zeilen als $(1 \times m)$-Matrizen behandelt werden.

2.5.7. [Beispiele]

$$
\begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 5 & -3 \end{pmatrix}
$$

$$
\begin{pmatrix} -1 & 1 \\ 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -2 & -2 & 1 \\ 0 & 2 & 3 \\ 3 & 4 & 0 \end{pmatrix}
$$

2.5.8. [Beachte]

- Das Beispiel zeigt, dass im allgemeinen AB verschieden von BA ist, wenn beide Produkte definiert sind.

- Das Produkt zweier Matrizen A und B ist nur definiert, wenn die Anzahl Spalten von A mit der Anzahl Zeilen von B übereinstimmt.

- Quadratische Matrizen derselben Größe können immer multipliziert werden. (Wir werden später aber noch viele Beispiele sehen, wo $AB \neq BA$ auch für zwei quadratische $\times n$-Matrizen A, B gilt.)

2.5.9. [Spezialfälle] Multiplikation einer $m \times n$-Matrix A mit einem Vektor $x \in \mathbb{R}^n$:

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{pmatrix} \quad \text{mit } b_i = \sum_{j=1}^{n} a_{ij} x_j .$$

Das Matrizenprodukt liefert uns also eine kurze Schreibweise für unser lineares Gleichungssystem (2.1):

$$Ax = b \ , \quad \text{wobei } \ b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} .$$

Das Problem, das Gleichungssystem (2.1) zu lösen, kann nun wie folgt ausgedrückt werden: für eine gegebene $m \times n$-Matrix A und einen gegebenen Vektor $b \in \mathbb{R}^m$, finde alle $x \in \mathbb{R}^n$ mit $Ax = b$. Das Gauß-Jordan Eliminations-Verfahren besteht im Wesentlichen darin, die $m \times (n+1)$ Matrix $(A|b)$, d.h. die Matrix A erweitert um b als zusätzliche Spalte, durch elementare Zeilenumformungen in eine Matrix in Stufenform zu bringen.

Satz 2.5.10. (Rechenregeln für Matrixmultiplikation) *Seien A, A' $m \times n$-Matrizen, B, B' $n \times p$-Matrizen und C eine $p \times q$-Matrix sowie λ ein Skalar. Dann gelten die folgenden Regeln:*

$$(AB)C = A(BC) \tag{A1}$$

$$\begin{aligned} (A + A')B &= AB + A'B \\ A(B + B') &= AB + AB' \end{aligned} \tag{A2}$$

$$E_m A = A = A E_n \tag{A3}$$

$$(\lambda A)B = \lambda(AB) = A(\lambda B) \tag{A4}$$

Beweis. Diese Regeln können durch explizites Ausrechnen bewiesen werden. Hierbei ist der Nachweis von (A1) schon am aufwändigsten. Der Eintrag an der Stelle (i, ℓ) der Matrix $(AB)C$ ist

$$\sum_m \Big(\sum_k a_{ik} b_{km} \Big) c_{m\ell} = \sum_{m,k} a_{ik} b_{km} c_{m\ell} = \sum_k a_{ik} \Big(\sum_m b_{km} c_{m\ell} \Big),$$

was auch der Eintrag an der Stelle (i, ℓ) in $A(BC)$ ist.

Für den ersten Teil von (A2) beobachtet man lediglich, dass

$$\sum_j (a_{ij} + a'_{ij}) b_{jk} = \sum_j a_{ij} b_{jk} + \sum_j a'_{ij} b_{jk}$$

ist.

Die Verifikation der weiteren Regeln seien dem geneigten Leser als Übung überlassen. □

Wir werden später sehen, dass es auch unmittelbare begriffliche Gründe für diese Rechenregeln gibt, die man ohne Rechnungen mit vielen Indizes, dafür auf einer höheren Abstraktionsstufe, einsehen kann.

Die Skalarmultiplikation kann übrigens wie folgt durch das Matrizenprodukt ausgedrückt werden:

$$\lambda A = \begin{pmatrix} \lambda\ 0\ 0\ \ldots 0 \\ 0\ \lambda\ 0 \ldots 0 \\ \ldots \\ 0\ 0\ 0 \ldots \lambda \end{pmatrix} \cdot A = (\lambda E_n) \cdot A = A \cdot (\lambda E_n)$$

Das Produkt AB zweier Matrizen A und B ist nur definiert, wenn die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt. Wenn wir uns auf quadratische $n \times n$-Matrizen A und B beschränken, dann ist das Produkt AB immer definiert und das Ergebnis wieder eine $n \times n$-Matrix. Damit haben wir auf der Menge $M_{n,n}(\mathbb{R}) = M_n(\mathbb{R})$ eine Multiplikation, die mit (A1)–(A4) alle üblichen Regeln erfüllt — *außer* der Kommutativität: im allgemeinen ist $AB \neq BA$. Eigenschaft (A3) liest sich wie folgt:

$$AE_n = A = E_n A \quad \text{für alle } A \in M_n(\mathbb{R}).$$

Das bedeutet, dass die Einheitsmatrix E die Rolle eines neutralen Elements bzgl. der Matrizenmultiplikation spielt, vergleichbar der „1“ bei den reellen Zahlen.

2.5.2 Matrizen und lineare Gleichungssysteme

Wir wollen nochmal die Nützlichkeit von Matrizen verdeutlichen. Wir erinnern uns, ein System linearer Gleichungen

$$Ax = b$$

heißt *homogen*, falls $b = 0$. Andernfalls nennt man es *inhomogen*.

Satz 2.5.11. *Sei* $\mathbf{x_s} = \begin{pmatrix} x_{s,1} \\ x_{s,2} \\ x_{s,3} \\ .. \\ .. \\ x_{s,n} \end{pmatrix}$ *eine spezielle Lösung des inhomogenen Systems*

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

Ist $\mathbf{x_h} = \begin{pmatrix} x_{h,1} \\ x_{h,2} \\ .. \\ .. \\ x_{h,n} \end{pmatrix}$ *eine Lösung des homogenen Systems*

$$Ax = 0, \tag{H}$$

so ist $\mathbf{x_h} + \mathbf{x_s}$ *eine Lösung des inhomogenen Systems* (2.1)*. Und umgekehrt, zu jeder Lösung* x *des inhomogenen Systems* (2.1) *gibt es eine Lösung* $\mathbf{x_h}$ *des homogenen Systems (H), so dass* $x = \mathbf{x_s} + \mathbf{x_h}$.

Beweis. Sei $\mathbf{x_s}$ eine spezielle Lösung von (L).

(a) Ist $\mathbf{x_h}$ eine Lösung von (H), dann gilt

$$A\mathbf{x_h} = 0 \quad \text{und} \quad A\mathbf{x_s} = b.$$

Addieren wir beide Gleichungen, erhalten wir $A\mathbf{x_h} + A\mathbf{x_s} = b$. Da $A\mathbf{x_h} + A\mathbf{x_s} = A(\mathbf{x_h} + \mathbf{x_s})$ nach (A2), folgern wir $A(\mathbf{x_h} + \mathbf{x_s}) = b$, d.h. $x := \mathbf{x_h} + \mathbf{x_s}$ ist eine Lösung von (L).

(b) Ist x irgendeine Lösung von (L), dann gilt

$$\begin{aligned} Ax &= b \\ A\mathbf{x_s} &= b. \end{aligned}$$

Subtrahieren wir beide Gleichungen, so erhalten wir $Ax - A\mathbf{x_s} = 0$. Wieder mit (A2) gilt $Ax - A\mathbf{x_s} = A(x - \mathbf{x_s})$. Und damit $A(x - \mathbf{x_s}) = 0$, d.h. $\mathbf{x_h} := x - \mathbf{x_s}$ ist eine Lösung von (H) und folglich $x = \mathbf{x_h} + \mathbf{x_s}$. □

2.5.12. [Merke] Die „allgemeine Lösung“ des inhomogenen Gleichungssystems ist die Summe einer speziellen Lösung des inhomogenen Systems und der „allgemeinen Lösung“ des homogenen Systems.

Definition 2.5.13. [Invertierbare Matrizen] Wir betrachten nur quadratische $n \times n$-Matrizen und setzen zur Abkürzung $E := E_n$.

Eine Matrix $A \in M_n(\mathbb{R})$ heißt *invertierbar* (oder ‘regulär’, oder ‘nicht-singulär’), falls es eine Matrix $B \in M_n(\mathbb{R})$ gibt, so dass

$$AB = E = BA. \tag{I}$$

Andernfalls wird A *singulär* genannt.

Die Menge aller invertierbaren $n \times n$-Matrizen mit reellen Einträgen wird mit $\mathrm{GL}(n, \mathbb{R})$ oder auch $\mathrm{GL}_n(\mathbb{R})$ bezeichnet. („GL” steht dabei für „general linear group“.)

2.5.14. [Beachte] Ist A invertierbar, dann ist die Matrix B mit der Eigenschaft (I) eindeutig bestimmt. Denn wenn B' eine weitere Matrix mit $AB' = E = B'A$ ist, so folgt

$$B = BE = B(AB') = (BA)B' = EB' = B'.$$

Wir nennen B *die Inverse* von A und schreiben $B = A^{-1}$.

Satz 2.5.15. (Eigenschaften invertierbarer Matrizen)

(1) *E ist invertierbar und es gilt $E = E^{-1}$.*

(2) *Falls A invertierbar ist, dann ist auch A^{-1} invertierbar und $(A^{-1})^{-1} = A$.*

(3) *Falls A und B invertierbar sind, dann auch AB und $(AB)^{-1} = B^{-1}A^{-1}$.*

Beweis. (1) $E \cdot E = E = E \cdot E$.
(2) $A^{-1} \cdot A = E = A \cdot A^{-1}$.
(3) $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$ und

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}EB = B^{-1}B = E. \qquad \square$$

Beispiele 2.5.16. (a) Die Nullmatrix 0 ist nicht invertierbar.

(b) Sei $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Dann gilt $AA = E$. Das heißt A ist invertierbar und $A^{-1} = A$.

(c) Falls eine Zeile (oder eine Spalte) von A nur aus Nullen besteht, dann ist A nicht invertierbar.

(Ist zum Beispiel die erste Spalte von A der Nullvektor 0, dann ist die erste Spalte von BA ebenfalls der Nullvektor 0 und damit gilt $BA \neq E$ für jede Matrix B.)

(d) Die 2×2-Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist genau dann invertierbar genau dann, wenn $ad - bc \neq 0$ ist. In diesem Fall ist

$$A^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In der Tat erhalten wir für $C := \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ die Relation $AC = (ad - bc)E$. Ist A invertierbar, so erhalten wir also durch Multiplikation mit A^{-1} die Relation $C = (ad - bc)A^{-1}$. Aus $A \neq 0$ folgt $C \neq 0$ und daher auch $ad - bc \neq 0$. Ist dies umgekehrt der Fall, so zeigt obige Rechnung, dass $(ad-bc)^{-1}C$ ein Inverses von A ist.

Für $n \times n$-Matrizen mit größerer Zeilen- und Spaltenzahl $n > 2$ gibt es zwar immer noch explizite Formeln für die Koeffizienten der inversen Matrix, sofern sie existiert. (Stichwort „Cramersche Regel“, Satz 6.5.11.) Sie sind aber erheblich komplizierter als in dem eben vorgeführten Fall $n = 2$ und für die praktische Rechnung unbrauchbar. Wir fragen daher nach einem Algorithmus, der erstens herausfindet, ob eine gegebene $n \times n$-Matrix A invertierbar ist, und falls ja, die Inverse berechnet. Den Schlüssel zu diesem Problem liefert uns folgender Satz.

Satz 2.5.17. *Seien A und B zwei $n \times n$-Matrizen. B ist genau dann die Inverse von A (d.h. genau dann ist A invertierbar und $B = A^{-1}$), wenn für alle $x, y \in \mathbb{R}^n$ gilt:*

$$Ax = y \iff x = By. \tag{2.5}$$

Beweis. Wir beweisen zuerst, dass (2.5) aus $B = A^{-1}$ folgt. In der Tat erhalten wir aus $Ax = y$ durch Multiplikation von links mit A^{-1} die Relation $x = Ex = A^{-1}(Ax) = By$. Umgekehrt folgt aus $x = By$ durch Multiplikation mit A sofort $Ax = ABy = y$.

Für die umgekehrte Implikation nehmen wir (2.5) an, d.h. wir nehmen an, dass für alle $x, y \in \mathbb{R}^n$ gilt: $Ax = y \iff x = By$. Wir müssen daraus ableiten, dass $B = A^{-1}$ ist. Sei dazu $x \in \mathbb{R}^n$ beliebig. Für $y := Ax$ erhalten wir dann $x = By = BAx$. Umgekehrt erhalten wir für $y \in \mathbb{R}^n$ und $x := By$ die Relation $y = Ax = ABy$. Beide Gleichungen gelten für alle Vektoren x bzw. y in $\mathbb{R}^n$. Daraus folgt $BA = E_n = AB$, nach folgendem Lemma:

Lemma 2.5.18. *Sei $C \in M_n(\mathbb{R})$ eine Matrix, so dass für alle $x \in \mathbb{R}^n$ gilt: $Cx = x$. Dann ist $C = E_n$.*

Beweis. Zum Beweis genügt es, die Identität $Cx = x$ für die folgenden n Vektoren auszunützen:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ .. \\ .. \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ .. \\ .. \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ .. \\ .. \\ 0 \end{pmatrix}, \ldots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ .. \\ .. \\ 1 \end{pmatrix},$$

denn die Gleichung $Ce_i = e_i$ besagt gerade, dass die i-te Spalte der Matrix C gleich der i-ten Zeile der Einheitsmatrix ist. □

□

Dies legt uns folgende Vorgehensweise zur Berechnung der Inversen nahe: Sei $A \in M_n(\mathbb{R})$ gegeben. Betrachte das lineare Gleichungssystem

$$Ax = y. \tag{2.1}$$

Versuche dieses System nach x aufzulösen und versuche x durch y auszudrücken:

$$x = By. \tag{2.2}$$

Wenn der Eliminationsprozess so durchgeführt werden kann, dass das zweite System zum ersten äquivalent ist, dann ist nach obigem Satz A invertierbar und $A^{-1} = B$. — Die Durchführung dieser Idee gibt uns den

Gauß-Algorithmus, zweite Variante

Man führt den Algorithmus der Einfachheit halber gleich simultan für alle Spalten der Einheitsmatrix auf der rechten Seite aus. Anders gesagt, man wendet den Gauß'schen Algorithmus auf die $n \times 2n$ Matrix $(A|E_n)$ an, mit dem Ziel, nicht nur A in Stufenform zu bringen, sondern am Ende eine $n \times 2n$ Matrix der Form $(E_n|B)$ zu erhalten. Insbesondere müssen die Diagonalelemente der Matrix links am Ende gleich 1 sein — dafür ist die Multiplikation gewisser Zeilen mit geeigeneten Skalaren, d.h. die elementare Umformung (E3) i.a. unerlässlich, die wir in der ersten Variante des Algorithmus gar nicht bemüht haben.

Wenn der Algorithmus eine Zeile erzeugt, in der alle Koeffizienten der linken Seite Null sind—anders gesagt, wenn der Rang von A kleiner als n ist—, dann ist A nicht invertierbar. Andernfalls aber ist A invertierbar und die rechte Teilmatrix B am Ende ist die inverse von A.

Wir wollen dies an zwei Beispielen illustrieren.

2.5.19. [Berechnung der Inversen]

Beispiel 1. $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ 2 & 1 & 1 \end{pmatrix}$. Nun versuchen wir das System $Ax = y$ nach x aufzulösen:

	A	E
$x_1 + x_2 + x_3 = y_1$	1 1 1	1 0 0
$2x_1 + 2x_2 + 4x_3 = y_2$	2 2 4	0 1 0
$2x_1 + x_2 + x_3 = y_3$	2 1 1	0 0 1
$x_1 + x_2 + x_3 = y_1$	1 1 1	1 0 0
$+2x_3 = -2y_1 + y_2$	0 0 2	−2 1 0
$-x_2 - x_3 = -2y_1 + y_3$	0 −1 −1	−2 0 1
$x_1 + x_2 + x_3 = y_1$	1 1 1	1 0 0
$-x_2 - x_3 = -2y_1 + y_3$	0 −1 −1	−2 0 1
$2x_3 = -2y_1 + y_2$	0 0 2	−2 1 0
$x_1 = -y_1 + y_3$	1 0 0	−1 0 1
$-x_2 - x_3 = -2y_1 + y_3$	0 −1 −1	−2 0 1
$2x_3 = -2y_1 + y_2$	0 0 2	−2 1 0
$x_1 = -y_1 + y_3$	1 0 0	−1 0 1
$x_2 + x_3 = 2y_1 - y_3$	0 1 1	2 0 −1
$x_3 = -y_1 + \frac{1}{2}y_2$	0 0 1	−1 $\frac{1}{2}$ 0
$x_1 = -y_1 + y_3$	1 0 0	−1 0 1
$x_2 = 3y_1 - \frac{1}{2}y_2 - y_3$	0 1 0	3 $-\frac{1}{2}$ −1
$x_3 = -y_1 + \frac{1}{2}y_2$	0 0 1	−1 $\frac{1}{2}$ 0
	E	A^{-1}

Dies zeigt, dass A invertierbar ist und

$$A^{-1} = \begin{pmatrix} -1 & 0 & 1 \\ 3 & -\frac{1}{2} & -1 \\ -1 & \frac{1}{2} & 0 \end{pmatrix}.$$

Beachte, dass für jedes $b \in \mathbb{R}^n$ das System $Ax = b$ eine eindeutige Lösung hat, wenn A invertierbar ist, nämlich $x = A^{-1}b$.

Löse zum Beispiel

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_1 + 2x_2 + 4x_3 &= -1 \\ 2x_1 + x_2 + x_3 &= 2 \end{aligned}$$

d.h.

$$Ax = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ 2 & 1 & 1 \end{pmatrix} x = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Die Lösung ist

$$x = A^{-1}b = \begin{pmatrix} -1 & 0 & 1 \\ 3 & -\frac{1}{2} & -1 \\ -1 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{3}{2} \\ -\frac{3}{2} \end{pmatrix}.$$

Beispiel 2. $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ 1 & 1 & -1 \end{pmatrix}$

Das System $Ax = y$

$$\begin{aligned} x_1 + \ \ x_2 + \ \ x_3 &= y_1 \\ 2x_1 + 2x_2 + 4x_3 &= \quad y_2 \\ x_1 + \ \ x_2 - \ \ x_3 &= \qquad y_3 \end{aligned}$$

wird durch elementare Umformungen übergeführt in

$$\begin{aligned} x_1 + x_2 + \ \ x_3 &= \quad y_1 \\ 2x_3 &= -2y_1 + y_2 \\ 0 &= -3y_1 + y_2 + y_3 \end{aligned}$$

Dieses System kann nicht übergeführt werden in die Form $x = By$. Also ist A *nicht* invertierbar.

A			E		
1	1	1	1	0	0
2	2	4	0	1	0
1	1	−1	0	0	1
1	1	1	1	0	0
0	0	2	−2	1	0
0	0	−2	−1	0	1
1	1	1	1	0	0
0	0	2	−2	1	0
0	0	0	−3	1	1

2.5.20. [Die Transponierte einer Matrix] Die *Transponierte* einer $m \times n$-Matrix

$$A = \begin{array}{r} \\ \\ i \rightarrow \\ \\ \\ \end{array} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

ist definiert als die $n \times m$-Matrix

$$A^T := \begin{array}{cc} & \begin{array}{ccccc} \phantom{a_{11}} & & i & & \phantom{a_{m1}} \end{array} \\ \begin{array}{r} \\ \\ \\ j \rightarrow \\ \\ \\ \end{array} & \begin{pmatrix} a_{11} & \dots & a_{i1} & \dots & a_{m1} \\ a_{12} & \dots & a_{i2} & \dots & a_{m2} \\ \vdots & & \vdots & & \vdots \\ a_{1j} & \dots & a_{ij} & \dots & a_{mj} \\ \vdots & & \vdots & & \vdots \\ a_{1n} & \dots & a_{in} & \dots & a_{mn} \end{pmatrix} \end{array}$$

Die i-te Zeile von A wird die i-te Spalte von A^T, die j-te Spalte von A wird die j-te Zeile A^T.

Satz 2.5.21. (Regeln für die Transponierte) *Für alle $m \times n$-Matrizen A und A' und jede $n \times p$-Matrix B gilt*

$$(A^T)^T = A, \quad (A + A')^T = A^T + (A')^T, \quad (\lambda A)^T = \lambda A^T, \quad (AB)^T = B^T A^T.$$

Beweis. Die ersten drei Regeln sind triviale Konsequenzen der Definition. Der Eintrag an der Stelle (i, k) in AB is $\sum_j a_{ij} b_{jk}$, und das entspricht dem Eintrag an der Stelle (k, i) in $(AB)^\top$. Wir können diese Summe aber auch als $\sum_j b_{jk} a_{ij}$ schreiben, und damit sieht man unmittelbar, dass $(AB)^\top = B^\top A^\top$ ist. $\square$

Beachte, dass die Transponierte eines Produkts von Matrizen das Produkt der transponierten Matrizen ist—freilich in umgekehrter Reihenfolge (ebenso wie die Inverse eines Produkts).

Übung 2.5.22. Zeige, dass für jede invertierbare $n \times n$-Matrix A, auch A^T invertierbar ist und bestimme $(A^T)^{-1}$.

Kapitel 3

Algebraische Strukturen

Nachdem wir die linearen Gleichungssysteme algorithmisch studiert und die Matrizenrechnung definiert und eingeübt haben, holen wir wieder einmal Luft und nehmen einen neuen Standpunkt ein: den der mathematischen, genauer algebraischen *Strukturen*. Dass sich dieser Schub von Abstraktion anschließend wirklich auszahlt, werden wir im nächsten Kapitel sehen.

In diesem Kapitel führen wir kurz und überblicksartig einige Standardbegriffe der Algebra ein: Monoid, Gruppe, Ring, Körper. Der Körperbegriff verallgemeinert die algebraischen Strukturen, die man von den rationalen oder reellen Zahlen her kennt.

So werden wir insbesondere den begrifflichen Rahmen für die Theorie der Vektorräume haben, die uns dann eine Rückkehr zur Theorie der linearen Gleichungssysteme von einem höheren Standpunkt aus erlauben wird.

3.1 Binäre Operationen, Monoide

Wir kennen alle gewisse arithmetische Operationen, z.B. die Addition und Multiplikation ganzer oder rationaler Zahlen. Von solchen Beispielen ausgehend, abstrahieren wir jetzt nur die formalen Eigenschaften solcher Operationen und legen dafür ein eigenes Vokabular an.

Definition 3.1.1. Eine *binäre Operation* auf einer Menge S ist eine Abbildung

$$S \times S \xrightarrow{*} S, \quad (a, b) \mapsto a * b,$$

die also jedem geordneten Paar (a, b) von Elementen von S ein anderes Element $a * b$ von S zuordnet.

Anstelle des neutralen Symbols $*$ wird je nach Zusammenhang eine andere Bezeichnung eingesetzt, etwa

additive Notation: $a + b$ (diese Bezeichnungsweise gebraucht man nur dann, wenn die Operation kommutativ ist, d.h. wenn für alle $a, b \in S$ gilt $a + b = b + a$.)

multiplikative Notation: $a \cdot b$ oder einfach ab

Definition 3.1.2. [Assoziative Operationen – Halbgruppen] Eine Operation $*$ auf S heißt *assoziativ* falls für alle $a, b, c \in S$ gilt

(G1) $a * (b * c) = (a * b) * c.$

Ein Paar $(S, *)$ bestehend aus einer Menge S und einer assoziativen Operation $*$ auf S heißt eine *Halbgruppe*.

Wenn es aus dem Zusammenhang klar ist, welche Operation $*$ auf der Menge gemeint ist, so schreibt man zur Abkürzung die Halbgruppe oft einfach S, anstelle von $(S, *)$. Diesem üblichen Bezeichnungsmißbrauch (in dem dasselbe Zeichen S sowohl für eine nackte Menge, als auch für diese Menge mit ihrer Halbgruppenstruktur steht) darf man sich genau dann anschließen, wenn man ihn als solchen empfindet.

Definition 3.1.3. [Kommutativität] Eine Operation $*$ auf einer Menge S heißt *kommutativ*, wenn für alle $a, b \in S$ gilt

$$a * b = b * a.$$

Eine Halbgruppe mit einer kommutativen Operation heißt eine *kommutative Halbgruppe*.

Beispiele 3.1.4. [Beispiele von Operationen]

1. Die übliche Addition auf $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{R}^n$, $M_{m,n}(\mathbb{R})$ ist eine assoziative und kommutative Operation.

2. Die übliche Multiplikation auf $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $M_n(\mathbb{R})$ ist eine assoziative Operation. Sie ist auch kommutativ auf $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$, aber nicht auf $M_n(\mathbb{R})$ (für $n \geq 2$). Für $n = 2$ hat man z.B. das Gegenbeispiel
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

3. Die Komposition $\circ$ von Abbildungen der Menge $\{1, 2, 3, \ldots, n\}$ in sich selbst ist, wie wir wissen, eine assoziative Operation auf der Menge T_n aller dieser Abbildungen. Sobald $n \geq 2$ ist, ist sie nicht kommutativ. (Übung!).

4. Es gibt 'jede Menge' Beispiele nicht-assoziativer Operationen. So ist etwa die Operation $a * b := a + 2b$ auf $\mathbb{N}$ (weder kommutativ noch) assoziativ; denn
$$(1 * 1) * 1 = 3 * 1 = 5 \neq 7 = 1 * 3 = 1 * (1 * 1).$$

Bemerkung 3.1.5. Ein Produkt von vier Faktoren kann man in fünf verschiedenen Weisen klammern:

$$((a*b)*c)*d, \quad (a*b)*(c*d), \quad a*(b*(c*d)), \quad (a*(b*c))*d, \quad a*((b*c)*d).$$

Ist die Operation $*$ assoziativ, liefern alle diese Ausdrücke das gleiche Element (Beweis?!). Mit vollständiger Induktion (siehe die Analysis-Vorlesung) kann man zeigen, dass auch in einem Produkt mit n Faktoren die Klammerung keine Rolle spielt, falls die Operation assoziativ ist.

In diesem Sinn schreiben wir für assoziative Operationen auch einfach

$$s_1 * \cdots * s_n := \big(\cdots((s_1 * s_2) * s_3) \cdots * s_{n-1}\big) * s_n.$$

Definition 3.1.6. [Potenzen] Ist $(S, *)$ eine Halbgruppe, so können wir die n-te *Potenz* a^n jedes Elementes $a \in S$ für jede natürliche Zahl n wie folgt bilden:

$$a^n := \underbrace{a * a * \cdots * a}_{n \text{ mal}}$$

Es gelten die folgenden Rechenregeln (in die die übliche Addition und Multiplikation natürlicher Zahlen eingehen)

$$a^n * a^m = a^{n+m}, \quad (a^n)^m = a^{n \cdot m} \qquad \text{für alle } \ n, m \in \mathbb{N}.$$

Wird die Operation additiv geschrieben, so heißen die Potenzen *Vielfache* und werden

$$n \cdot a := \underbrace{a + a + \cdots + a}_{n \text{ mal}}$$

geschrieben. Die obigen Regeln lauten dann:

$$n \cdot a + m \cdot a = (n + m) \cdot a, \quad m \cdot (n \cdot a) = (mn) \cdot a.$$

Definition 3.1.7. [Neutrales Element—Monoide] Ein Element $e \in S$ heißt *neutrales Element* (für die Operation $*$ auf S) falls für alle $a \in S$ gilt

(G2) $e * a = a = a * e$.

Ein neutrales Element muss es nicht geben. Wenn es aber existiert, so gibt es davon nur eines: Sind e und e' neutrale Elemente, so folgt $e = e * e' = e'$.

Ein *Monoid* ist ein Tripel $(M, *, e)$, das aus einer Menge M, einer assoziativen Operation $*$ und einem neutralen Element e für diese Operation besteht. Ein Monoid ist demnach eine Halbgruppe zusammen mit einem neutralen Element.

In multiplikativer Schreibweise wird das neutrale Element meistens mit 1 bezeichnet, in additiver Notation mit 0. (Das gilt freilich nicht immer: das neutrale Element der Matrizenmultiplikation in $M_n(\mathbb{R})$ haben wir z.B. im letzten Kapitel mit E_n bezeichnet.)

Beispiel 3.1.8. (a) $(\mathbb{N}_0, +, 0)$, $(\mathbb{N}, \cdot, 1)$ und $(\mathbb{Z}, \cdot, 1)$ sind Monoide.

(b) Sei Σ ein 'Alphabet', d.h. eine Menge von Symbolen. Ein *Wort* über dem Alphabet Σ ist eine endliche Folge $w = a_1 a_2 \dots a_n$ von Symbolen $a_i \in \Sigma$. Die leere Folge, das leere Wort ist auch zugelassen und sei mit e bezeichnet. Auf der

Menge Σ^* aller Wörter über Σ definieren wir die Operation $*$ durch einfaches Hintereinanderschreiben: Für $v = a_1 a_2 \ldots a_n$ und $w = b_1 b_2 \ldots b_m$ setzen wir

$$v * w := a_1 a_2 \ldots a_n b_1 b_2 \ldots b_m.$$

Dann nennt man $(\Sigma^*, *, e)$ das *freie Monoid über dem Alphabet* Σ. — In der Tat ist das Hintereinanderschreiben von Worten offenbar assoziativ, und das leere Wort e ist ein neutrales Element.

Zu jeder mathematischen Struktur gehört eine Klasse von Abbildungen, die 'diese Struktur respektieren'. Man nennt sie die zugehörigen *Morphismen* oder *Homomorphismen*. Wir erläutern dies jetzt für den Fall der Halbgruppen und Monoide.[1]

Definition 3.1.9. [Morphismen] (a) Seien $(S_1, *)$ und $(S_2, \diamond)$ Halbgruppen. Eine Abbildung $f\colon S_1 \to S_2$ heißt *(Halbgruppen-)Homomorphismus*, oder *Morphismus von Halbgruppen*, wenn sie mit den Halbgruppenoperationen im folgenden Sinne verträglich ist:

$$f(s * t) = f(s) \diamond f(t) \quad \text{für alle} \quad s, t \in S_1.$$

Ein Morphismus $f\colon S_1 \to S_2$ heißt *Isomorphismus*, wenn es einen Morphismus $g\colon S_2 \to S_1$ gibt, so dass gilt:

$$g \circ f = \mathrm{id}_{S_1} \quad \text{und} \quad f \circ g = \mathrm{id}_{S_2}\,.$$

Zwei Halbgruppen $(S_1, *)$ und $(S_2, \diamond)$ heißen *isomorph* (in Zeichen: $S_1 \cong S_2$), wenn es einen Isomorphismus $f\colon S_1 \to S_2$ gibt.

Ein *Endomorphismus* einer Halbgruppe S ist ein Morphismus $f\colon S \to S$ der Halbgruppe in sich selbst, und ein Endomorphismus, der gleichzeitig ein Isomorphismus ist, wird *Automorphismus* genannt. Wir bezeichnen die Menge aller Endomorphismen von S mit $\mathrm{End}(S)$, und mit $\mathrm{Aut}(S)$ die Menge aller Automorphismen von S.
(b) Seien $(M_1, *, e_1)$ und $(M_2, \diamond, e_2)$ Monoide. Eine Abbildung $f\colon M_1 \to M_2$ heißt *(Monoid-)Homomorphismus*, oder *Morphismus von Monoiden*, wenn sie ein Halbgruppenhomomorphismus ist, für den außerdem gilt:

$$f(e_1) = e_2.$$

Ein Morphismus $f\colon M_1 \to M_2$ heißt *Isomorphismus*, falls es einen Morphismus $g\colon M_2 \to M_1$ gibt, so dass gilt:

$$g \circ f = \mathrm{id}_{M_1} \quad \text{und} \quad f \circ g = \mathrm{id}_{M_2}\,.$$

[1] Es gibt eine grundlegende mathematische Theorie, die genau auf dieser Parallelität von Strukturen und strukturerhaltenden Abbildungen beruht und sie allgemein präzisiert und verallgemeinert: die *Kategorientheorie*, auch als 'Theorie der Pfeile' (welche 'Abbildungen' bezeichnen) oder als 'abstrakter Unsinn' bekannt. Anstatt der Mengenlehre kann man die ganze Mathematik auch auf die Kategorientheorie gründen. — Eine Kategorie ist am einfachsten als eine Klasse von Objekten definiert, so dass zwischen je zwei Objekten X und Y eine Menge von Morphismen $\mathrm{Mor}(X, Y)$ und eine 'Komposition' von Morphismen so festgelegt ist, dass gewisse Regeln erfüllt sind. . .

Zwei Monoide $(M_1, *, e_1)$ und $(M_2, \diamond, e_2)$ heißen *isomorph* (in Zeichen $M_1 \cong M_2$), wenn es einen Monoidisomorphismus $f\colon M_1 \to M_2$ gibt. Endomorphismen und Automorphismen von Monoiden sind analog zu denen von Halbgruppen definiert — siehe oben.

Isomorphe Halbgruppen oder Monoide sind algebraisch gesehen nur verschiedene Realisierungen ein und derselben Struktur.

Beispiel 3.1.10.

1. Ist $(S, *)$ eine Halbgruppe und $s \in S$, so ist die Potenzierungsabbildung
$$p_s\colon \mathbb{N} \to S, \quad n \mapsto s^n$$
ein Halbgruppenhomomorphismus $(\mathbb{N}, +) \to (S, *)$.

2. Für Elemente m eines Monoids $(M, *, e)$ definieren wir stets $m^0 := e$. Dann ist für jedes $m \in M$ die Potenzierungsabbildung
$$p_s\colon \mathbb{N}_0 \to M, \quad n \mapsto s^n$$
ein Monoidhomomorphismus $(\mathbb{N}_0, +) \to (S, *)$.

3. Sei $\Sigma = \{s\}$ eine Menge mit genau einem Element und Σ^* das freie Monoid über Σ. Dann ist die Potenzierungsabbildung
$$p_s\colon \mathbb{N}_0 \to \Sigma^*, \quad n \mapsto s^n = \underbrace{ssss\cdots s}_{\text{n mal}}$$
ein Monoidisomorphismus.

4. Ist $x \in \mathbb{R}^n$, so ist die Abbildung
$$f_x\colon \mathbb{R} \to \mathbb{R}^n, \quad \lambda \mapsto \lambda x$$
ein Monoidhomomorphismus $(\mathbb{R}, +, 0) \to (\mathbb{R}^n, +, 0)$.

Übungen zu Abschnitt 3.1

Übung 3.1.11.

1. Sei $f\colon M_1 \to M_2$ ein Monoidhomomorphismus. Zeige, dass f genau dann ein Isomorphismus ist, wenn f bijektiv ist.

2. Zeige, dass für beliebige Monoide M_1, M_2 und M_3 die Bedingungen $M_1 \cong M_2$ und $M_2 \cong M_3$ die Isomorphie $M_1 \cong M_3$ implizieren.

Übung 3.1.12. (Die universelle Eigenschaft des freien Monoids.) Sei Σ^* das freie Monoid über dem Alphabet Σ. Sei $(M, *, e)$ ein Monoid und $f\colon \Sigma \to M$ eine Abbildung. Dann gibt es genau einen Monoidmorphismus $f^*\colon \Sigma^* \to M$, der f fortsetzt.

Übung 3.1.13. Sind $(M, *, e_M)$ und $(N, \wedge, e_N)$ Monoide und ist $\varphi\colon M \to N$ ein Isomorphismus von Halbgruppen, so ist φ auch ein Isomorphismus von Monoiden.

3.2 Gruppen

Definition 3.2.1. [Invertierbare Elemente] Sei $(M, *, e)$ ein Monoid. Ein Element $a \in M$ heißt *invertierbar*, falls es eine Element $b \in M$ gibt, so dass gilt

$$a * b = e = b * a.$$

Ein solches Element b heißt dann *Inverses* von a und wird im allgemeinen mit a^{-1} (im Fall der additiven Notation der Operation aber mit $-a$) bezeichnet. Damit diese Bezeichnung sinnvoll ist, müssen wir nachprüfen, dass das Inverse eines invertierbaren Elements eindeutig bestimmt ist: Seien b und b' Inverse von a. Dann kommt

$$b \overset{(G2)}{=} b * e = b * (a * b') \overset{(G1)}{=} (b * a) * b' = e * b' \overset{(G2)}{=} b'$$

Die Notation a^{-1} für das Inverse ist also wohldefiniert.

Wir schreiben $M^\times$ für die Menge aller invertierbaren Elemente in M.

Satz 3.2.2. (Regeln für das Inverse) *Sei $(M, *, e)$ ein Monoid.*

(1) *e ist invertierbar und $e^{-1} = e$.*

(2) *Ist a invertierbar, so ist auch a^{-1} invertierbar und es gilt $(a^{-1})^{-1} = a$.*

(3) *Sind a und b invertierbar, so ist auch $a * b$ invertierbar und es gilt $(a * b)^{-1} = b^{-1} * a^{-1}$.*

Beweis. (1) Da $e * e = e$ ist nach (G2), ist e invertierbar und seinem eigenen Inversen gleich.

(2) Ist a invertierbar, so gilt $a * a^{-1} = e = a^{-1} * a$ — somit ist auch a^{-1} invertierbar mit Inversem a.

(3) Sind a und b invertierbar, so folgt

$$\begin{aligned} (a * b) * (b^{-1} * a^{-1}) &= a * ((b * b^{-1}) * a^{-1}) = a * (e * a^{-1}) = a * a^{-1} = e, \\ (b^{-1} * a^{-1}) * (a * b) &= (b^{-1} * (a^{-1} * a)) * b = (b^{-1} * e) * b = b^{-1} * b = e. \end{aligned}$$

Also ist auch $a * b$ invertierbar, mit Inversem $b^{-1} * a^{-1}$. □

Bemerkung 3.2.3. In additiver Schreibweise ist ein Element a invertierbar, falls es ein Element b gibt mit $a + b = 0 = b + a$. Wie oben schon bemerkt schreibt man dann das Inverse von a als $-a$ anstelle von a^{-1}.

Definition 3.2.4. [Gruppen] Ein Monoid $(G, *, e)$, in dem jedes Element invertierbar ist, heißt eine *Gruppe*.

Demnach ist eine Gruppe also eine Menge G mit einer Operation

$$G \times G \xrightarrow{*} G, \quad (a, b) \mapsto a * b$$

und einem ausgezeichneten Element $e \in G$, so dass die folgenden Eigenschaften erfüllt sind:

(G1) $(\forall a, b, c \in G)\ a * (b * c) = (a * b) * c.$

(G2) $(\forall a \in G)\ a * e = a = e * a.$

(G3) $(\forall a \in G)(\exists b \in G)\ a * b = e = b * a.$

Definition 3.2.5. [Abelsche Gruppen] Eine Halbgruppe (bzw. ein Monoid, bzw. eine Gruppe), deren Operation kommutativ ist, heißt eine kommutative Halbgruppe (bzw. Monoid, bzw. Gruppe). Kommutative Gruppen werden meistens *abelsche Gruppen* genannt.[2]

Die additive Schreibweise benutzt man nur für kommutative Operationen.

Satz 3.2.6. *Seien* $(G_1, *, e_1)$ *und* $(G_2, \diamond, e_2)$ *zwei Gruppen, und sei* $f : G_1 \to G_2$ *eine Abbildung zwischen den zugrundeliegenden Mengen. Dann sind folgende beiden Bedingungen äquivalent.*

(HG) f *ist ein Homomorphismus der Halbgruppen* $(G_1, *), (G_2, \diamond)$.

(GH) f *hat die folgenden drei Eigenschaften:*

- (GH1) $(\forall a, b \in G_1)\ f(a * b) = f(a) \diamond f(b)$.
- (GH2) $f(e_1) = e_2$.
- (GH3) $(\forall a \in G_1)\ f(a^{-1}) = f(a)^{-1}$.

Eine Abbildung, die diesen Bedingungen genügt, heißt *Gruppenhomomorphismus.*

Beweis. Da nach Definition (HG) gleichbedeutend mit (GH1) ist, ist lediglich zu zeigen, dass (HG) auch (GH2) und (GH3) impliziert. Nun ist aber für beliebiges $a \in G_1$:

$$f(a) = f(a * e_1) = f(a) \diamond f(e_1).$$

Wendet man auf beide Seiten dieser Gleichung $f(a)^{-1}\diamond$ an, so kommt

$$e_2 = f(a)^{-1} \diamond f(a) = (f(a)^{-1} \diamond f(a)) \diamond f(e_1) = f(e_1),$$

und somit (GH2).

Weiterhin gilt dann für jedes $a \in G_1$:

$$f(a^{-1}) \diamond f(a) = f(a^{-1} * a) = f(e_1) = e_2$$

und ebenso $f(a) \diamond f(a^{-1}) = f(e_1) = e_2$, womit $f(a^{-1})$ als Inverses von $f(a)$ und also (GH3) nachgewiesen ist. $\square$

[2] Niels Henrik Abel (1802–1829), Norwegischer Mathematiker; lebte zwischen Christiania (dem heutigen Oslo), Berlin und Paris. Er hat als erster einen vollständigen Beweis der Tatsache geliefert, dass es im allgemeinen für algebraische Gleichungen des Grades 5 oder höher *keine* expliziten Lösungsformeln mit beliebigen Wurzelausdrücken geben kann. In diesem Zusammenhang hat er auch in der Tat die Kommutativitätsbedingung für gewisse Permutationsgruppen als erster betrachtet. Außerdem hat er in seinem kurzen Leben Beiträge zur Theorie der Potenzreihen und analytischen Funktionen (sogenannte elliptische und 'abelsche' Integrale und Funktionen) geliefert, ohne die die Mathematik heute wohl anders aussähe.

Beispiele 3.2.7. (a) Beispiele abelscher Gruppen sind: $(\mathbb{Z}, +, 0)$, $(\mathbb{Q}, +, 0)$, $(\mathbb{R}, +, 0)$, $(\mathbb{R}^n, +, 0)$ und $(M_{m,n}(\mathbb{R}), +, \mathbf{0})$.

(b) Für jedes $n \in \mathbb{N}$ ist das Tripel $(\mathrm{GL}_n(\mathbb{R}), \cdot, E_n)$ eine Gruppe, wobei die Operation $\cdot$ das Matrizenprodukt ist. Im Fall $n = 1$ ist diese Gruppe isomorph zu $(\mathbb{R}^\times, \cdot, 1)$ (Übung!) und also abelsch. Aber für $n \geq 2$ ist sie nicht abelsch. Für $n = 2$ folgt das aus Beispiel 3.1.4(2).

Satz 3.2.8. *Sei $(M, *, e)$ ein Monoid und bezeichne mit $M^\times$ die Menge seiner invertierbaren Elemente. Dann macht die Einschränkung der Monoidoperation auf $M^\times \times M^\times$ die Menge $M^\times$ zu einer Gruppe mit neutralem Element e.*

Beweis. Für $a, b \in M^\times$ liegt das Produkt $a * b$ ebenfalls in $M^\times$, gemäß Satz 3.2.2(3). Also liefert $*$ durch Einschränkung tatsächlich eine Operation $M^\times \times M^\times \to M^\times$.

Da die Operation $*$ auf ganz M assoziativ ist, gilt dies auch für die Einschränkung auf $M^\times$, d.h. wir haben die Eigenschaft (G1). Nach Satz 3.2.2(1) liegt $e \in M^\times$ und ist offenbar ein neutrales Element auch für $M^\times$, also haben wir (G2). Schließlich ist jedes $a \in M^\times$ invertierbar in M (nach Definition von $M^\times$), und sein Inverses a^{-1} liegt wieder in $M^\times$ nach Satz 3.2.2(2). Somit ist auch (G3) erfüllt. □

Definition 3.2.9. [Einheitengruppe] Ist M ein Monoid, so heißt die Gruppe $(M^\times, \cdot, e)$ die *Einheitengruppe* von M. Ihre Elemente heißen *Einheiten* von M.

Beispiele 3.2.10.

1. $(\mathbb{Z}, \cdot, 1)$ ist ein Monoid. Seine einzigen invertierbaren Elemente sind $+1$ und -1. Also ist $\mathbb{Z}^\times = \{+1, -1\}$; dies ist eine Gruppe bzgl. der Multiplikation.

2. $(\mathbb{Q}, \cdot, 1)$ ist ein Monoid. Alle von null verschiedenen Elemente sind invertierbar in $\mathbb{Q}$, d.h. $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$ ist die Einheitengruppe von $(\mathbb{Q}, \cdot)$.

3. Sei M eine Menge. Die Menge T_M aller Abbildungen $f\colon M \to M$ ist ein Monoid bzgl. der Hintereinanderausführung $\circ$ von Abbildungen. id_M ist das neutrale Element (siehe Kapitel 1). Die Menge $S_M := T_M^\times$ ist dann gerade die Menge der Bijektionen, oder Permutationen von M. Insbesondere ist S_M eine Gruppe bzgl. der Komposition $\circ$. Diese Gruppe ist nicht kommutativ, falls M mindestens drei verschiedene Elemente enthält. Die Gruppe S_M heißt die *symmetrische Gruppe auf M*. Als wichtigen Spezialfall finden wir für $M = \{1, \ldots, n\}$ die symmetrische Gruppe $S_n := S_M$ wieder, die schon am Ende des ersten Kapitels betrachtet wurde. Sie ist also die Einheitengruppe des Monoids T_n aller Selbstabbildungen von $\{1, \ldots, n\}$.

4. Die Menge $M_n(\mathbb{R})$ aller $n \times n$-Matrizen mit reellen Einträgen ist ein Monoid bzgl. der Matrizenmultiplikation. Das neutrale Element ist die Einheitsmatrix E_n, und die Einheitengruppe ist die Menge $\mathrm{GL}_n(\mathbb{R}) = M_n(\mathbb{R})^\times$ aller invertierbarer reeller $n \times n$-Matrizen, die ebenfalls im 2. Kapitel schon eingeführt wurde. Wir haben schon gesehen, dass diese Gruppe für $n \geq 2$ nicht abelsch ist.

5. Sei $(G, *, e)$ eine Gruppe und $a \in G$. Dann definieren wir für jede ganze Zahl $n \in \mathbb{Z}$:

$$a^n := \begin{cases} a * a * \cdots * a \ (n \text{ mal}) & \text{falls } n > 0 \\ e & \text{falls } n = 0 \\ (a^{-n})^{-1} & \text{falls } n < 0 \end{cases}$$

Dann gilt für *alle* $n \in \mathbb{Z}$ die Formel

$$a^{m+n} = a^m * a^n;$$

oder anders gesagt: *Die Abbildung*

$$\varphi_a : \mathbb{Z} \to G \qquad \varphi_a : n \mapsto a^n$$

ist ein Gruppenhomomorphismus von $(\mathbb{Z}, +, 0)$ *nach* $(G, *, e)$.

Übungen zu Abschnitt 3.2

Übung 3.2.11. Zeige, dass die symmetrische Gruppe S_n genau dann abelsch ist, wenn $n \leq 2$. (Beachte, dass T_2 nicht kommutativ ist !)

Übung 3.2.12. Ein Element f einer Halbgrupep $(S, *)$ heißt *idempotentes Element*, oder *Idempotent*, wenn $f^2 = f * f = f$ gilt. Zeige:

(1) Morphismen von Halbgruppen bilden Idempotente in Idempotente ab.

(2) Ist $(S, *, e)$ eine Gruppe, so ist e das einzige Idempotent in S.

(3) Schließe aus (2), dass jeder Halbgruppenhomomorphismus zwischen Gruppen die neutralen Elemente aufeinander abbildet.

Übung 3.2.13. Beschreibe die Idempotenten in der Halbgruppe T_M der Selbstabbildungen der Menge M.

Können Sie eine Verbindung zu Repräsentantensystemen von Äquivalenzrelationen herstellen?

3.3 Ringe und Körper

Schon in einigen der vorstehenden Beispiele traten Mengen auf, auf denen zwei verschiedene Operationen definiert sind: eine Addition und eine Multiplikation.

Definition 3.3.1. [Ring] Eine Menge R mit zwei (binären) Operationen $+$ und $\cdot$ und einem ausgezeichneten Element 0 (also das Quadrupel $(R, +, \cdot, 0)$) heißt ein *Ring*, falls folgende Bedingungen erfüllt sind:

(R1) $(R, +, 0)$ ist eine abelsche Gruppe,

(R2) $(R, \cdot)$ ist eine Halbgruppe, und

(R3) es gelten die *Distributivgesetze*:

$$(\forall a, b, c \in R) \qquad a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(\forall a, b, c \in R) \qquad (b + c) \cdot a = b \cdot a + c \cdot a$$

Gibt es darüber hinaus auch ein Element $1 \in R$, welches von 0 verschieden ist und $(R, \cdot, 1)$ zu einem Monoid macht, so heißt $(R, +, \cdot, 0, 1)$ ein *Ring mit Eins* oder ein *unitärer Ring*. Ein Ring heißt *kommutativ*, falls (auch) die Multiplikation $\cdot$ auf R kommutativ ist.

Beispiele 3.3.2. Die ganzen Zahlen sind die zugrundeliegende Menge eines kommutativen Rings: $(\mathbb{Z}, +, \cdot, 0, 1)$. Für alle $n \geq 2$ bilden die quadratischen Matrizen einen nicht-kommutativen Ring $(M_n(\mathbb{R}), +, \cdot, \mathbf{0}, E_n)$.

Definition 3.3.3. Seien R und S Ringe mit Eins. Eine Abbildung $f\colon R \to S$ heißt ein *Homomorphismus unitärer Ringe*, wenn für alle $x, y \in R$ gilt:

$$f(x + y) = f(x) + f(y), \quad f(x \cdot y) = f(x) \cdot f(y), \quad \text{und} \quad f(1_R) = 1_S.$$

Ein solcher Homomorphismus f heißt *(Ring-)Isomorphismus*, falls es einen Homomorphismus $g\colon S \to R$ unitärer Ringe gibt mit

$$g \circ f = \mathrm{id}_R \quad \text{und} \quad f \circ g = \mathrm{id}_S\,.$$

Übung 3.3.4. In einem Ring R gilt stets $a \cdot 0 = 0 = 0 \cdot a$, für alle $a \in R$.

Freilich kann es in gewissen Ringen Elemente a, b geben, die beide nicht null sind, für die aber $a \cdot b = 0$ wird. Solche Elemente nennt man *Nullteiler*. Zum Beispiel hat man in $M_2(\mathbb{R})$:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Definition 3.3.5. [Körper] Ein kommutativer Ring $\mathbb{K}$ mit 1, in dem $0 \neq 1$ und $\mathbb{K}^\times = \mathbb{K} \setminus \{0\}$ ist, heißt ein *Körper*.

Da Körper eine besondere Sorte unitärer Ringe sind, ist damit auch definiert, was ein Körperhomomorphismus und Körperisomorphismus ist.

Beispiele 3.3.6. [von Körpern:] $\mathbb{Q}$, $\mathbb{R}$ mit der üblichen Addition und Multiplikation sind Körper. Dagegen ist $\mathbb{Z}$ kein Körper.

Bemerkung 3.3.7. (a) In einem Körper $\mathbb{K}$ bilden die von 0 verschiedenen Elemente eine multiplikative Gruppe $\mathbb{K}^\times = \mathbb{K} \setminus \{0\}$, nach Satz 3.2.8.

(b) In einem Körper kann es keine Nullteiler geben, d.h. aus $ab = 0$ folgt $a = 0$ oder $b = 0$.

In einem Körper hat man die arithmetischen Rechenperationen, die von den reellen Zahlen her vertraut sind: Addition, Subtraktion, Multiplikation und Division durch Elemente, die von 0 verschieden sind. Aber Körper können sehr verschieden von dem Körper der reellen Zahlen sein.

3.3.8. [Zusammenfassung] Wir fassen die definierenden Eigenschaften eines Körpers noch einmal auf einen Blick zusammen.

Ein *Körper* ist eine Menge K, zusammen mit zwei Operationen $+$ und $\cdot$ und zwei ausgezeichneten Elementen $0 \neq 1$, so dass die folgenden Aussagen erfüllt sind:

$$\begin{array}{lcl}
(K1) & (\forall a,b,c \in K) & a+(b+c)=(a+b)+c, \quad a(bc)=(ab)c \\
(K2) & (\forall a \in K) & a+0=a, \quad a\cdot 1=a \\
(K3) & (\forall a \in K)(\exists b \in K) & a+b=0, \qquad a\neq 0 \Rightarrow (\exists c \in K)\ ac=1 \\
(K4) & (\forall a,b \in K) & a+b=b+a, \quad a\cdot b=b\cdot a \\
(K5) & (\forall a,b,c \in K) & a(b+c)=ab+ac
\end{array}$$

Beispiel 3.3.9. [Komplexe Zahlen] Auf $\mathbb{C} := \mathbb{R}^2$ definieren wir eine Addition und eine Multiplikation durch die Formeln

$$(x,y)+(x',y') := (x+x',y+y') \quad \text{und} \quad (x,y)\cdot(x',y') := (xx'-yy',xy'+x'y).$$

Die Elemente von $\mathbb{C}$ heißen *komplexe Zahlen*. Es ist eine elementare Übung, die Körperaxiome für $(\mathbb{C},+,\cdot,(0,0),(1,0))$ nachzuprüfen. Das multiplikativ neutrale Element ist $1 := (1,0)$ und $0 := (0,0)$ ist das additive neutrale Element. Man schreibt außerdem

$$i := (0,1).$$

Dann gilt $i^2 = (-1,0) = -1$, d.h. -1 ist in $\mathbb{C}$ ein Quadrat.

Die reellen Zahlen werden nach $\mathbb{C}$ über die erste Komponente eingebettet. In der Tat gilt:

$$(a,0)+(c,0) := (a+c,0), \quad \text{und} \quad (a,0)\cdot(c,0) := (ac,0)$$

für alle $a,c \in \mathbb{R}$. Die Abbildung

$$\varphi\colon \mathbb{R} \to \mathbb{C}, \quad a \mapsto (a,0)$$

ist also ein injektiver Körperhomomorphismus. In diesem Sinne betrachten wir $\mathbb{R}$ als eine Teilmenge von $\mathbb{C}$, die bzgl. der Einschränkung von Addition und Multiplikation ein Körper ist. Man sagt, $\mathbb{R}$ ist ein *Teilkörper* oder *Unterkörper* von $\mathbb{C}$.

Jede komplexe Zahl (x,y) kann in genau einer Weise als

$$(x,y) = x+iy, \quad x,y \in \mathbb{R}$$

geschrieben werden. Ab jetzt werden wir komplexe Zahlen stets in dieser Form schreiben. Es gilt

$$(x+iy)+(x'+iy') = (x+x')+i(y+y')$$

und

$$(x+iy)(x'+iy') = (xx'-yy')+i(yx'+xy').$$

Beispiel 3.3.10. [Der Körper mit zwei Elementen $\mathbb{Z}/2\mathbb{Z}$] Sei $\mathbb{Z}/2\mathbb{Z} := \{\overline{0}, \overline{1}\}$ und definiere eine Addition $+_2$ und eine Multiplikation $\cdot_2$ auf dieser Menge wie folgt:

$+_2$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$\cdot_2$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

Es ist nicht schwer nachzuprüfen, dass $\mathbb{Z}/2\mathbb{Z}$ auf diese Weise ein Körper wird.

Man kann diesen Körper als eine algebraische Formalisierung des Umgangs mit den Begriffen *gerade* für $\overline{0}$ und *ungerade* für $\overline{1}$ auffassen. Wegen $\overline{1} +_2 \overline{1} = \overline{0}$ hat man in diesem Körper $-\overline{1} = \overline{1}$. Obwohl das etwas gewöhnungsbedürftig erscheinen kann, ist dieser Körper ein grundlegendes Hilfsmittel in der Kodierungstheorie.

Übung 3.3.11. Sei $\mathbb{F} := \{\mathbf{T}, \mathbf{F}\}$ und definiere die Multiplikation und Addition auf $\mathbb{F}$ durch

$$a \cdot b := a \wedge b \quad \text{und} \quad a + b := (a \wedge \neg b) \vee (b \wedge \neg a).$$

Zeige, dass $(\mathbb{F}, +, \cdot, \mathbf{F}, \mathbf{T})$ ein Körper ist und dass die Abbildung

$$f \colon \mathbb{F} \to \mathbb{Z}/2\mathbb{Z}, \quad \mathbf{T} \mapsto \overline{1}, \quad \mathbf{F} \mapsto \overline{0}$$

ein Körperisomorphismus ist.

Beispiel 3.3.12. [Die Ringe $\mathbb{Z}/n\mathbb{Z}$] Sei $n \geq 2$ eine natürliche Zahl. Sei

$$\mathbb{Z}/n\mathbb{Z} := \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}$$

und definiere Operationen $+_n$ und $\cdot_n$ auf $\mathbb{Z}/n\mathbb{Z}$ wie folgt

$$\begin{aligned} \overline{a} +_n \overline{b} \quad &:= \overline{r}\ , \ \text{falls } a+b \text{ bei Division durch } n \text{ den Rest } r \text{ lässt} \\ \overline{a} \cdot_n \overline{b} \quad &:= \overline{s}\ , \ \text{falls } a \cdot b \text{ bei Division durch } n \text{ den Rest } s \text{ lässt.} \end{aligned}$$

Wir erinnern hierzu an die *Division mit Rest*: Für jede natürliche Zahl $n > 0$ und jede ganze Zahl $b \in \mathbb{Z}$ gibt es ganze Zahlen q und r, so dass gilt:

$$b = q \cdot n + r \quad \text{und} \quad 0 \leq r < n.$$

Die Zahlen q und r sind durch n und b eindeutig bestimmt. Die Zahl r heißt der *Rest* von b modulo n.

Beispiele 3.3.13. $n = 3$:

$+_3$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

$\cdot_3$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{1}$

$n = 4$:

$+_4$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

$\cdot_4$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Satz 3.3.14. *Mit den Operationen $+_n$ und $\cdot_n$ bildet die Menge*

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{n-1}\}$$

einen kommutativen Ring; $\overline{0}$ bzw. $\overline{1}$ sind die neutralen Elemente der Addition bzw. Multiplikation.

Hier ein kurzes Programm dessen, was für den Beweis dieses Satzes genau nachzuprüfen ist:

- Die Relation „$a \equiv b \pmod{n}$" [lies: a kongruent b modulo n] auf $\mathbb{Z}$, die zwischen zwei ganzen Zahlen a und b genau dann besteht, wenn a und b den gleichen Rest bei Division durch n lassen, ist eine Äquivalenzrelation. Die Äquivalenzklasse einer ganzen Zahl a: $\{b \in \mathbb{Z} : a \equiv b \pmod{n}\}$, ist gerade die Menge

  $$\overline{a} = \{a + kn : k \in \mathbb{Z}\} = a + k\mathbb{Z}.$$

  Das oben benutzte Zeichen $\overline{a}$ (für $a = 0, 1, 2, \ldots, n-1$) für ein Element von $\mathbb{Z}/n\mathbb{Z}$ steht als Abkürzung für die Äquivalenzklasse von a. Die Menge $\mathbb{Z}/n\mathbb{Z}$ *ist* die Menge der Äquivalenzklassen der Kongruenz modulo n.

- Ist $a \equiv a' \pmod n$ und $b \equiv b' \pmod n$, so ist auch $a+b \equiv a'+b' \pmod n$ und $a \cdot b \equiv a' \cdot b' \pmod n$. Mit anderen Worten: Die übliche Addition bzw. Multiplikation ganzer Zahlen ergeben eine wohldefinierte Addition bzw. Multiplikation auf der Menge der Äquivalenzklassen modulo n. (Man vergröbert also die Operationen ganzer Zahlen lediglich so, dass man alles durch n Teilbare systematisch vernachlässigt.) So ergibt sich die oben definierte Addition $+_n$ bzw. Multiplikation $\cdot_n$ auf $\mathbb{Z}/n\mathbb{Z}$. Um diese Aussage beispielsweise für die Multiplikation nachzuweisen, nehme an, $a \equiv a' \pmod n$ und $b \equiv b' \pmod n$. Dann gibt es also ganze Zahlen $k, \ell \in \mathbb{Z}$, so dass $a' = a + kn; b' = b + \ell n$ ist. Wir erhalten $a'b' = ab + (kb + \ell a + k\ell n)n$, mithin $a'b' \equiv ab \pmod n$ wie behauptet.

- Die Ringeigenschaften für $\mathbb{Z}/n\mathbb{Z}$ ergeben sich dann *automatisch* aus der Tatsache, dass $\mathbb{Z}$ ein Ring ist, 'durch Vergröberung modulo n'. Zum Beispiel ergibt sich die Assoziativität aus

  $$(\overline{a}+\overline{b})+\overline{c} = \overline{a+b}+\overline{c} = \overline{(a+b)+c} = \overline{a+(b+c)} = \overline{a}+\overline{b+c} = \overline{a}+(\overline{b}+\overline{c}).$$

Äquivalenzrelationen und Partitionen

Um die vorstehende Beweisskizze zu verstehen, wiederholen wir hier den Zusammenhang zwischen Äquivalenzrelationen und Partitionen einer Menge. Zunächst diese beiden Begriffe:

Definition 3.3.15. Sei M eine Menge. Eine Relation $\sim \subseteq M \times M$ auf der Menge M heißt *Äquivalenzrelation*, falls sie reflexiv, symmetrisch und transitiv ist, wobei diese drei Eigenschaften wie folgt definiert sind:

(R) (Reflexivität) $(\forall x \in M)\ x \sim x$

(S) (Symmetrie) $(\forall x, y \in M)\ x \sim y \Rightarrow y \sim x$

(T) (Transitivität) $(\forall x, y, z \in M)\ x \sim y \wedge y \sim z \Rightarrow x \sim z$

Definition 3.3.16. Eine Menge $\mathcal{P} \subseteq \mathbf{P}(M)$ von Teilmengen von M heißt *Partition* von M, wenn gilt:

(P1) $M = \bigcup \mathcal{P} = \bigcup \{A\colon A \in \mathcal{P}\}$

(P2) $(\forall A, B \in \mathcal{P})\ A \cap B \neq \emptyset \Rightarrow A = B$.

Anschaulich kann man diese Bedingungen auch so formulieren, dass jedes Element $m \in M$ in genau einer der Mengen liegt, die zu $\mathcal{P}$ gehören. Die Bedingung (P2) garantiert hierbei die Eindeutigkeit.

Wir werden gleich sehen, dass Partitionen und Äquivalenzrelationen letztendlich die gleichen Strukturen beschreiben, allerdings aus verschiedenen Blickwinkeln. Vermittelt wird zwischen beiden durch den Begriff der *Äquivalenzklasse*: Ist $\sim$ eine Äquivalenzrelation auf der Menge M und ist $x \in M$, so nennen wir die Teilmenge

$$\overline{x} := \{y \in M :\ x \sim y\} \subseteq M$$

die *Äquivalenzklasse von x*. Wir schreiben

$$M/\sim := \{\overline{x} \colon x \in M\} \subseteq \mathbb{P}(M)$$

für die Menge der Äquivalenzklassen.

Lemma 3.3.17. *Ist $\sim$ eine Äquivalenzrelation auf M, so bildet die Menge $M/\sim$ der Äquivalenzklassen eine Partition von M.*

Beweis. Dass $\sim$ reflexiv ist, besagt gerade, dass jedes Element $x \in M$ in der Äquivalenzklasse $\overline{x}$ enthalten ist. Also ist $\bigcup M/\sim = M$.

Sind $\overline{x}$ und $\overline{y}$ Äquivalenzklassen, deren Durchschnitt ein Element z enthält, so folgt für jedes $a \in \overline{x}$ aus $a \sim x$, $x \sim z$ und $z \sim y$ wegen der Transitivität von $\sim$ auch $a \sim y$; also $\overline{x} \subseteq \overline{y}$. Vertauschen der Rollen von x und y liefert nun $\overline{y} \subseteq \overline{x}$, also Gleichheit. ☐

Bemerkung 3.3.18. Dieses Lemma zeigt, dass eine Äquivalenzklasse durch jedes beliebige ihrer Elemente eindeutig bestimmt, man sagt: *repräsentiert* ist. Wählen wir aus jeder Äquivalenzklasse der auf M gegebenen Äquivalenzrelation $\sim$ genau ein Element aus, so enthalten wir ein *Repräsentatensystem R* für $\sim$:

$$(\forall x \in M)(\exists! r \in R)\ x \sim r\ ; \qquad (\forall r, r' \in R)\ r \sim r' \Rightarrow r = r'.$$

[Es ist im allgemeinen nicht unproblematisch, 'aus jeder Klasse ein Element auszuwählen', sobald die Menge der Klassen unendlich (und 'größer als $\mathbb{N}$') ist. Um derartige Fragen zu regeln, gibt es in der Mengenlehre das sogenannte *Auswahlaxiom*, dem manche Mathematiker freilich skeptisch gegenüberstehen.]

Satz 3.3.19. *Die Äquivalenzrelationen und die Partitionen auf einer Menge M ensprechen einander bijektiv vermöge der folgenden beiden zueinander inversen Abbildungen:*

$$\sim \mapsto M/\sim, \quad \mathcal{P} \mapsto \sim_{\mathcal{P}}, \quad \sim_{\mathcal{P}} := \{(x, y) \in M \times M : (\exists P \in \mathcal{P}) x, y \in P\}$$

Hat man diese Aussage erst einmal verstanden, ist der Beweis mithilfe des Lemmas eine einfache Übung.

Kehren wir nun zu unserem Besipiel $\mathbb{Z}/n\mathbb{Z}$ zurück: dort geht es um die Äquivalenzrelation auf $\mathbb{Z}$, die für $a, b \in \mathbb{Z}$ genau dann erfüllt ist, wenn $a \equiv b \pmod{n}$ gilt. Ihre Äquivalenzklassen sind die $\overline{a} = \{a + kn : k \in \mathbb{Z}\}$. Ein gebräuchliches Repräsententantensystem für diese Äquivalenzrelation, das wir auch beim ersten Hinschreiben von $\mathbb{Z}/n\mathbb{Z}$ benutzt haben, ist

$$R = \{0, 1, \ldots, n-1\} \subset \mathbb{Z}.$$

Satz 3.3.20. *$\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Körper, wenn n eine Primzahl ist.*

Als Vorüberlegung hierzu studieren wir die (multiplikativ) invertierbaren Elemente, d.h. die Einheitengruppen unserer Ringe:

Bemerkung 3.3.21. [Multiplikativ invertierbare Elemente in $\mathbb{Z}/n\mathbb{Z}$] Wir schreiben $(\mathbb{Z}/n\mathbb{Z})^\times$ für die Menge der (bzgl. der Multiplikation) invertierbaren Elemente des Rings $\mathbb{Z}/n\mathbb{Z}$. Aus Satz 3.2.8 wissen wir dann, dass $(\mathbb{Z}/n\mathbb{Z})^\times$ bzgl. der Multiplikation eine Gruppe bildet.

In den beiden Beispielen oben prüft man nach, dass $(\mathbb{Z}/3\mathbb{Z})^\times = \{\overline{1}, \overline{2}\}$ und $(\mathbb{Z}/4\mathbb{Z})^\times = \{\overline{1}, \overline{3}\}$ ist. Um allgemein die Elemente $a \in (\mathbb{Z}/n\mathbb{Z})^\times$ zu charakterisieren, gehen wir bis auf die Definition der Invertierbarkeit zurück:

$$\begin{aligned} \overline{a} \text{ invertierbar in } \mathbb{Z}/n\mathbb{Z} \quad &\iff (\exists \overline{x} \in \mathbb{Z}/n\mathbb{Z})\ \overline{a} \cdot_n \overline{x} = \overline{1} \\ &\iff (\exists x \in \mathbb{Z})(\exists y \in \mathbb{Z})\ ax = ny + 1 \end{aligned}$$

$\overline{a}$ ist demnach genau dann invertierbar in $\mathbb{Z}/n\mathbb{Z}$, wenn die Gleichung $ax - ny = 1$ mit $x, y \in \mathbb{Z}$ lösbar ist. Das aber ist genau dann der Fall, wenn a und n zueinander teilerfremd sind: $\mathrm{ggT}(a, n) = 1$. Diese letzte Äquivalenz weist man mit Hilfe des *Euklidischen Algorithmus* nach. So nennt man das folgende sukzessive Dividieren mit Rest:

$$\begin{array}{rcll} r_{-1} := a & = & q_1 \cdot n + r_1 & (r_1 < n) \\ r_0 := n & = & q_2 \cdot r_1 + r_2 & (r_2 < r_1) \\ .. & = & \dots & \\ .. & = & \dots & \\ .. & = & \dots & \\ r_{\ell-2} & = & q_\ell \cdot r_{\ell-1} + r_\ell & (r_\ell < r_{\ell-1}) \\ r_{\ell-1} & = & q_{\ell+1} \cdot r_\ell + 0 & \end{array}$$

Dann ist $\mathrm{ggT}(a, n) = \mathrm{ggT}(n, r_1) = \mathrm{ggT}(r_1, r_2) = \dots \mathrm{ggT}(r_{\ell-1}, r_\ell) = r_\ell$ und dieser größte gemeinsame Teiler lässt sich durch Rückeinsetzen der Gleichungen in der Form $xa + yn$, mit $x, y \in \mathbb{Z}$ darstellen. Umgekehrt ist offenbar jeder gemeinsame Teiler d von a und n, der sich als $d = xa + yn$ schreiben lässt, größter gemeinsamer Teiler.

Beispiel 3.3.22. Wir wollen den größten gemeinsamen Teiler von 8 und 19 mit dem euklidischen Algorithmus bestimmen und aus Vielfachen dieser Zahlen kombinieren.

$$\begin{array}{rcl} 19 & = & 2 \cdot 8 + 3 \\ 8 & = & 2 \cdot 3 + 2 \\ 3 & = & 1 \cdot 2 + 1 \\ 2 & = & 2 \cdot 1 + 0 \end{array}$$

Der ggt ist also 1 (das sah man schon vorher...), und durch Rückeinsetzen erhalten wir die Darstellung:

$$\begin{aligned} 1 &= 3 - 1 \cdot 2 = 3 - 1 \cdot (8 - 2 \cdot 3) = 3 - 1 \cdot (8 - 2 \cdot (19 - 2 \cdot 8)) \\ &= (19 - 2 \cdot 8) - (8 - 2 \cdot (19 - 2 \cdot 8)) = 3 \cdot 19 - 7 \cdot 8. \end{aligned}$$

In der Tat ist $57 - 56 = 1$.

Wir haben also bewiesen:

Satz 3.3.23. *Ein Element $a \in \mathbb{Z}/n\mathbb{Z}$ ist genau dann multiplikativ invertierbar (bzgl. $\cdot_n$), wenn a und n teilerfremd sind:*

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{a \in \mathbb{Z}/n\mathbb{Z} : \mathrm{ggT}(a, n) = 1\}.$$

Beispiel 3.3.24. $(\mathbb{Z}/5\mathbb{Z})^\times = \{\overline{1}, \overline{2}, \overline{3}, \overline{4}\}$, $(\mathbb{Z}/12\mathbb{Z})^\times = \{\overline{1}, \overline{5}, \overline{7}, \overline{11}\}$.

Aus dem letzten Satz folgt sofort, dass $\mathbb{Z}/n\mathbb{Z}$ genau dann ein Körper ist (anders gesagt: dass genau dann $(\mathbb{Z}/n\mathbb{Z})^\times = (\mathbb{Z}/n\mathbb{Z}) \setminus \{\overline{0}\}$), wenn alle Zahlen $1, 2, \ldots, n-1$ zu n teilerfremd sind. Das aber ist genau dann der Fall, wenn n Primzahl ist.

Beispiele 3.3.25.

- $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z}$, ..., $\mathbb{Z}/49999\mathbb{Z}$, ... sind Körper.
- Die folgenden Ringe sind keine Körper: $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/6\mathbb{Z}$, ..., $\mathbb{Z}/111\mathbb{Z}$, ...

Übungen zu Abschnitt 3.3

Übung 3.3.26. Sei $\mathbb{K}$ ein Körper, $a \in \mathbb{K}$, und $\mathbb{F} := \mathbb{K}^2 = \{(x, y) \colon x, y \in \mathbb{K}\}$. Auf $\mathbb{F}$ definiere Addition und Multiplikation durch

$$(x, y)+(x', y') := (x+x', y+y') \quad \text{und} \quad (x, y)\cdot(x', y') := (xx'+ayy', xy'+x'y).$$

1. Zeige, dass $(\mathbb{F}, +, \cdot, (0, 0), (1, 0))$ ein Ring mit Eins ist.
2. Für $(x, y) \in \mathbb{F}$ definiere $N(x, y) := x^2 - ay^2$. Dann gilt:
$$N((x, y)(x', y')) = N(x, y)N(x', y').$$
3. $\mathbb{F}^\times = \{(x, y) \in \mathbb{F} \colon N(x, y) \neq 0\}$, und für $(x, y) \in \mathbb{F}^\times$ hat man
$$(x, y)^{-1} = \left(\frac{x}{N(x, y)}, -\frac{y}{N(x, y)}\right).$$
4. $\mathbb{F}$ ist genau dann ein Körper, wenn a kein Quadrat in $\mathbb{K}$ ist, d.h. wenn $a \neq b^2$ für alle $b \in \mathbb{K}$ gilt.
5. Zeige, dass für $\mathbb{K} = \mathbb{R}$ und $a = -1$ der Körper $\mathbb{F}$ isomorph ist zum Körper $\mathbb{C}$ der komplexen Zahlen.
6. Sei $\mathbb{K} = \mathbb{Q}$. Für welche natürlichen Zahlen $a \in \mathbb{N}$ ist $\mathbb{F}$ ein Körper?

Übung 3.3.27. Sei M eine Menge und $(R, +, \cdot, 0, 1)$ ein Ring mit Eins. Dann bildet die Menge R^M aller Funktionen $f \colon M \to R$ einen Ring bzgl.

$$(f + g)(x) := f(x) + g(x), \quad (f \cdot g)(x) := f(x) \cdot g(x),$$

wo die konstante Funktion $\mathbf{0}(x) := 0$ das additive neutrale Element und die konstante Funktion $\mathbf{1}(x) := 1$ das multiplikative neutrale Element ist.

Übung 3.3.28. Sei $(R, +, \cdot, 0, 1)$ ein Ring mit Eins und $n \in \mathbb{N}$. Dann ist die Menge $M_n(R)$ aller $(n \times n)$-Matrizen mit Einträgen in R ein Ring mit Eins bzgl. der üblichen Addition und dem Matrizenprodukt.

Übung 3.3.29. (Darstellung komplexer Zahlen als Matrizen.) Wir definieren eine Abbildung

$$\varphi\colon \mathbb{C} \to M_2(\mathbb{R}), \quad \varphi(x+iy) := \begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

Zeige, dass φ ein injektiver Homomorphismus unitärer Ringe ist.

Übung 3.3.30. (Invertierbarkeit von (2×2)-Matrizen) Sei $\mathbb{K}$ ein Körper und

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{K}).$$

Wir definieren die *Determinante von A* als

$$\det(A) := ad - bc.$$

1. Für $A, B \in M_2(\mathbb{K})$ gilt $\det(AB) = \det(A)\det(B)$. Die Abbildung
$$\det\colon (M_2(\mathbb{K}), \cdot, E_2) \to (\mathbb{K}, \cdot, 1)$$
ist sogar ein Morphismus von Monoiden.

2. Die Matrix A ist genau dann invertierbar, wenn $\det(A) \neq 0$ ist, und in diesem Fall gilt
$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Übung 3.3.31. (Quaternionen) Betrachte die Menge

$$\mathbb{H} := \left\{ \begin{pmatrix} z & -\overline{w} \\ w & \overline{z} \end{pmatrix} \in M_2(\mathbb{C})\colon z, w \in \mathbb{C} \right\}$$

komplexer Matrizen.

1. Zeige, dass $\mathbb{H}$ ein Unterring von $M_2(\mathbb{C})$ ist, d.h. dass diese Teilmenge unter Addition und Multiplikation abgeschlossen ist, und dass sie die Einheitsmatrix 1 enthält.

2. Zeige, dass die Elemente
$$\mathbf{1} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{i} := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{j} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{k} := \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$$
von $\mathbb{H}$ folgende Gleichungen erfüllen:
$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1}$$
sowie
$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}.$$

3. $\mathbb{H}^\times = \mathbb{H} \setminus \{0\}$.

4. Welche Körperaxiome sind für $\mathbb{H}$ *nicht* erfüllt?

Kapitel 4

Vektorräume

Den Begriff des Vektorraums kann man als eine Abstraktion und Verallgemeinerung der Struktur auffassen, die wir auf der Menge $\mathbb{R}^n$ kennengelernt haben: mit einer Addition und einer Skalarmultiplikation. Dabei abstrahieren und verallgemeinern wir ganz im Sinne des letzten Kapitels: Einerseits werden anstelle des Grundkörpers $\mathbb{R}$ der reellen Zahlen beliebige Körper zugelassen. Andererseits lösen wir uns von der konkreten Gestalt der Elemente als n-Tupel und fordern nur noch formale Regeln, denen die Addition und Skalarmultiplikation gehorchen müssen. Der Vorteil dieser abstrakten Begriffsbildung ist enorm: Vektorräume sind aus keinem Gebiet der Mathematik mehr wegzudenken. Wichtige Beispiele für Vektorräume haben als Grundmengen gewisse Klassen von Funktionen; häufig ist es vorteilhaft, gerade auch für die Anwendungen einer mathematischer Theorie, als Skalarbereich den Körper der komplexen Zahlen zu haben; andere Vektorräume, die z.B. in der Kodierungstheorie wichtig sind, sind Vektorräume über dem Körper mit zwei Elementen.

4.1 Der Begriff des Vektorraums

Definition 4.1.1. Ein *Vektorraum über dem Körper* $\mathbb{K}$ (oder ein $\mathbb{K}$-*Vektorraum*) ist eine Menge V, zusammen mit

- einem ausgezeichneten Element $0 \in V$
- einer *Addition* $(v, w) \mapsto v + w \colon V \times V \to V$
- einer *Skalarmultiplikation* $(\lambda, v) \mapsto \lambda v \colon \mathbb{K} \times V \to V$

so dass die folgenden Bedingungen erfüllt sind

(V1) $(\forall u, v, w \in V)\ u + (v + w) = (u + v) + w.$

(V2) $(\forall v \in V)\ v + 0 = v.$

(V3) $(\forall v \in V)\ v + (-1)v = 0.$

(V4) $(\forall v, w \in V)\ v + w = w + v.$

(V5) $(\forall v, w \in V)\ (\forall \lambda \in \mathbb{K})\ \lambda(v + w) = \lambda v + \lambda w.$

(V6) $(\forall v \in V)\ (\forall \lambda, \mu \in \mathbb{K})\ (\lambda + \mu)v = \lambda v + \mu v.$

(V7) $(\forall v \in V)(\forall \lambda, \mu \in \mathbb{K})\ \lambda(\mu v) = (\lambda\mu)v.$

(V8) $(\forall v \in V)\ 1v = v.$

(V1)–(V4) besagen gerade, dass $(V, +, 0)$ eine abelsche Gruppe ist; das additive Inverse von $v \in V$ ist $-v = (-1)v$. Elemente eines Vektorraums werden auch *Vektoren* genannt.

Beachte, dass wir im allgemeinen nicht typographisch zwischen der 0 des Körpers $\mathbb{K}$ und der $0 \in V$ unterscheiden — obwohl das natürlich in der Regel ganz verschiedene Elemente, ein Skalar, bzw. ein Vektor, sind.

Satz 4.1.2. (Abgeleitete Regeln) *In jedem $\mathbb{K}$-Vektorraum V gilt:*

(V9) $(\forall u, v, w \in V)\ u + w = v + w \Rightarrow u = v.$

(V10) $(\forall v \in V)\ 0v = 0, \qquad (\forall \lambda \in \mathbb{K})\ \lambda \cdot 0 = 0.$

(V11) $(\forall v \in V)\ (\forall \lambda \in \mathbb{K})\ \lambda v = 0 \Rightarrow \lambda = 0$ *or* $v = 0.$

Beweis. (V9): Angenommen $u + w = v + w$, dann folgt

$$u = u + (w - w) = (u + w) - w = (v + w) - w = v + (w - w) = v,$$

und also $u = v$.

(V10): Es gilt $0 + 0v = 0v = (0 + 0)v = 0v + 0v$. Aus (V9) folgt dann, dass $0 = 0v$. Für die zweite Behauptung beachten wir, dass

$$\lambda \cdot 0 = \lambda \cdot (0 + 0) = \lambda \cdot 0 + \lambda \cdot 0$$

gilt, und wenden (V9) an, um $\lambda \cdot 0 = 0$ zu erhalten.

(V11): Angenommen $\lambda v = 0$. Falls $\lambda = 0$, haben wir nichts zu beweisen. Ist aber $\lambda \neq 0$, so folgt aus (V10):

$$0 = \lambda^{-1}0 = \lambda^{-1} \cdot (\lambda \cdot v) = (\lambda\lambda^{-1})v = 1 \cdot v = v.$$

Mithin $v = 0$. □

Beispiel 4.1.3. Wie in Kapitel 2 werden wir die Elemente

$$x = (x_1, \ldots, x_n) \in \mathbb{K}^n$$

meistens als *Spaltenvektoren* schreiben:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n.$$

Die Addition und Multiplikation in $\mathbb{K}$ gibt uns sofort durch komponentenweise Anwendung die *Addition* und *Skalarmultiplikation* auf $\mathbb{K}^n$:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{und} \quad \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \lambda \cdot x_1 \\ \lambda \cdot x_2 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$$

für $x, y \in \mathbb{K}^n$ und $\lambda \in \mathbb{K}$. Mit diesen Operationen wird $\mathbb{K}^n$ zu einem $\mathbb{K}$-Vektorraum. (Das nachzuprüfen ist eine einfache Übung; vgl. Satz 2.1.1 für den exemplarischen Fall $\mathbb{K} = \mathbb{R}$.)

Beispiel 4.1.4. Analog zu Abschnitt 2.5 definieren wir eine $m \times n$-Matrix mit Koeffizienten aus $\mathbb{K}$ als ein Element von $(\mathbb{K}^n)^m$ und schreiben sie in rechteckiger Anordnung mit m Zeilen und n Spalten:

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

wobei jetzt für alle i, j die Einträge $a_{ij} \in \mathbb{K}$ sind. Die Menge aller $m \times n$-Matrizen mit Koeffizienten aus $\mathbb{K}$ bezeichnen wir mit $M_{m,n}(\mathbb{K})$.

Die Summe $A + B$ zweier $m \times n$-Matrizen A, B und die Multiplikation einer Matrix A mit einem Skalar $\lambda \in \mathbb{K}$ werden ebenso definiert wie wir es in 2.5.3 für reelle Matrizen gemacht haben. So wird auch $M_{m,n}(\mathbb{K})$ zu einem $\mathbb{K}$-Vektorraum, wie man rasch nachprüft.

Beispiel 4.1.5. Sei X eine beliebige Menge. Schreibe $\mathcal{F}(X, \mathbb{K})$ für die Menge aller Funktionen $f \colon X \to \mathbb{K}$. Sei 0 die konstante Funktion mit Wert $0 \in \mathbb{K}$; und für $f, g \in \mathcal{F}(X, \mathbb{K})$ und $\lambda \in \mathbb{K}$ definiere $f + g$ und λf durch die Regeln

$$(f + g)(x) := f(x) + g(x), \quad (\lambda f)(x) := \lambda \cdot f(x) \quad \text{für alle } x \in X.$$

Auf diese Weise wird $\mathcal{F}(X, \mathbb{K})$ zu einem $\mathbb{K}$-Vektorraum (Übung).

Ist etwa X ein Intervall in $\mathbb{R}$ und $\mathbb{K} = \mathbb{R}$, so wird die Menge aller Funktionen $f \colon X \to \mathbb{R}$ mit der üblichen Addition und Multiplikation mit Konstanten λ ein reeller Vektorraum.

4.2 Lineare Abbildungen

Getreu dem im letzten Kapitel erklärten Prinzip wenden wir uns jetzt, nach der Definition der Vektorraumstruktur, den Abbildungen zu, die diese algebraische Struktur erhalten. Dies sind die sogenannten *linearen Abbildungen*.

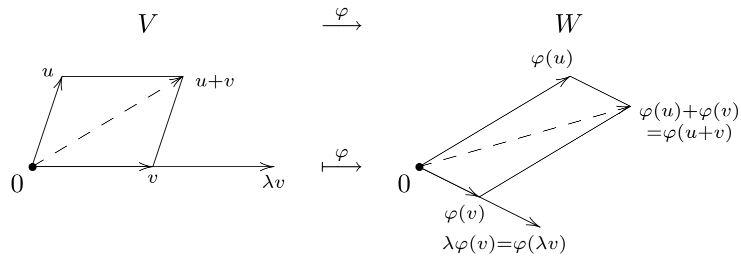
Definition 4.2.1. Seien V und W $\mathbb{K}$-Vektorräume. Eine Abbildung $\varphi \colon V \to W$ heißt *lineare Abbildung*, falls sie die folgenden beiden Eigenschaften besitzt:

(L1) $\varphi(u+v) = \varphi(u) + \varphi(v)$ für alle $u, v \in V$.

(L2) $\varphi(\lambda v) = \lambda\varphi(v)$ für alle $v \in V,\ \lambda \in \mathbb{K}$.

Man kann (L1) und (L2) in eine einzige Aussage vereinigen:

$$(\forall u, v \in V)(\forall \lambda, \mu \in \mathbb{K})\ \ \varphi(\lambda u + \mu v) = \lambda\varphi(u) + \mu\varphi(v).$$



Eine lineare Abbildung $\varphi\colon V \to W$ kann man auch als *Vektorraumhomomorphismus* bezeichnen. Von einem *Vektorraumisomorphismus* spricht man, wenn es eine lineare Abbildung $\psi\colon W \to V$ gibt, so dass $\varphi \circ \psi = \mathrm{id}_W$ und $\psi \circ \varphi = \mathrm{id}_V$. Eine lineare Abbildung $\varphi\colon V \to V$ von einem Vektorraum V in sich selbst heißt ein *Endomorphismus* von V, und an ein Isomorphismus $\varphi\colon V \to V$ heißt ein *Automorphismus*. Die Menge aller linearen Abbildungen $\varphi\colon V \to W$ wird mit $\mathrm{Hom}(V, W)$ bezeichnet, die Menge aller Endomorphismen von V mit $\mathrm{End}(V)$, und die Menge aller Automorphismen von V mit $\mathrm{Aut}(V)$.

Bemerkung 4.2.2. Sei $\varphi\colon V \to W$ eine lineare Abbildung. Dann gilt:

1. $\varphi(0) = 0$
2. $\varphi(-v) = -\varphi(v)$
3. $\varphi(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = \lambda_1\varphi(v_1) + \lambda_2\varphi(v_2) + \cdots + \lambda_n\varphi(v_n)$

In der Tat folgt (1) aus (L2) durch die Wahl $\lambda = 0$; (2) folgt ebenfalls aus (L2) mit $\lambda = -1$. (3) schließlich erhält man durch wiederholte Anwendung von (L1) und (L2), mit einem Induktionsbeweis.

Beispiele 4.2.3. (Lineare Abbildungen)

1. Wir beginnen mit einem Beispiel, das in gewissem Sinne den allgemeinen Fall einer linearen Abbildung illustriert: Sei $A \in M_{m,n}(\mathbb{K})$ eine $(m \times n)$-Matrix. Dann wird für jedes $p \in \mathbb{N}$ durch das Matrizenprodukt wie folgt eine lineare Abbildung definiert:

   $$L_A\colon M_{n,p}(\mathbb{K}) \to M_{m,p}(\mathbb{K}), \quad X \mapsto AX\ ;$$

   denn für $X, X' \in M_{n,p}(\mathbb{K})$ und $\lambda, \lambda' \in \mathbb{K}$ gilt in der Tat

   $$L_A(\lambda X + \lambda' X') = A(\lambda X + \lambda' X') = \lambda AX + \lambda' AX' = \lambda L_A(X) + \lambda' L_A(X').$$

2. Der Fall $p = 1$ des ersten Beispiels: Sei $A \in M_{m,n}(\mathbb{K})$ wie zuvor. Dann ist

$$L_A \colon \mathbb{K}^n \to \mathbb{K}^m, \quad x \mapsto Ax$$

eine lineare Abbildung. Beachte, dass die Bilder dieser Abbildung wie die linke Seite eines linearen Gleichungssystems aussehen!

3. Der Fall $m = p = 1$ des ersten Beispiels: Seien $a_1, \dots, a_n \in \mathbb{K}$. Wenden wir dann das erste beispiel mit $p = 1$ auf den Fall des Zeilenvektors $a = (a_1, \dots, a_n)$ an, so ergibt sich die lineare Abbildung

$$L_a \colon \mathbb{K}^n \to \mathbb{K}, \quad x \mapsto \sum_{j=1}^{n} a_j x_j.$$

4. Sei $V := \mathcal{F}(X, \mathbb{K})$ die Menge der $\mathbb{K}$-wertigen Funktionen auf einer Menge X. Dies ist ein $\mathbb{K}$-Vektorraum (Beispiel 4.1.5). Auf V definieren wir außerdem eine Multiplikation durch

$$(f \cdot g)(x) := f(x) \cdot g(x) \quad \text{für} \quad x \in X.$$

Dann ist für jedes $f \in V$ die Abbildung

$$\varphi_f \colon V \to V, \quad g \mapsto f \cdot g$$

linear (Übung).

5. Ist V ein Vektorraum und $\lambda \in \mathbb{K}$, dann ist die Skalarmultiplikationsabbildung

$$L_\lambda \colon V \to V, \quad v \mapsto \lambda v$$

linear. Das folgt unmittelbar aus (V5) und (V7).

Satz 4.2.4. (Zusammensetzung (Komposition) linearer Abbildungen)

(1) *Sind $\psi \colon U \to V$ und $\varphi \colon V \to W$ linear, so auch $\varphi \circ \psi \colon U \to W$.*

(2) *Die Identität $\mathrm{id}_V \colon V \to V$ ist linear.*

(3) *Ist $\varphi \colon V \to W$ eine bijektive lineare Abbildung, so gilt das gleiche auch für die Umkehrabbildung $\varphi^{-1} \colon W \to V$.*

Beweis. (1) Für $\lambda, \mu \in \mathbb{K}$ und $u_1, u_2 \in U$ haben wir

$$\begin{aligned}
(\varphi \circ \psi)(\lambda u_1 + \mu u_2) &= \varphi\big(\psi(\lambda u_1 + \mu u_2)\big) \\
&= \varphi\big(\lambda \psi(u_1) + \mu \psi(u_2)\big) && \text{wegen der Linearität von } \psi \\
&= \lambda \varphi\big(\psi(u_1)\big) + \mu \varphi\big(\psi(u_2)\big) && \text{wegen der Linearität von } \varphi \\
&= \lambda (\varphi \circ \psi)(u_1) + \mu (\varphi \circ \psi)(u_2)
\end{aligned}$$

(2) Die Linearität der Identität ist offensichtlich.

(3) Wir müssen nur zeigen dass φ^{-1} linear ist. Seien dazu $w_1, w_2, \in W$ und $\lambda, \mu \in \mathbb{K}$. Setze $v_1 = \varphi^{-1}(w_1)$ und $v_2 = \varphi^{-1}(w_2)$. Dann ist

$$\varphi(\lambda v_1 + \mu v_2) = \lambda\varphi(v_1) + \mu\varphi(v_2) = \lambda w_1 + \mu w_2,$$

also

$$\varphi^{-1}(\lambda w_1 + \mu w_2) = \lambda v_1 + \mu v_2 = \lambda\varphi^{-1}(w_1) + \mu\varphi^{-1}(w_2). \qquad \square$$

Übung 4.2.5. Sei $(A, +)$ eine abelsche Gruppe.

1. Zeige, dass die Menge $\mathrm{End}(A, +)$ von Gruppenendomorphismen ein Ring mit Eins bzgl. der folgenden Operationen ist:

   $$(\varphi + \psi)(v) := \varphi(v) + \psi(v) \quad \text{und} \quad (\varphi \cdot \psi)(v) := \varphi(\psi(v)), \quad v \in V.$$

   Das neutrale Element der Addition ist dabei die konstante Abbildung mit Wert 0, und das neutrale Element der Multiplikation ist id_A.

2. Ist V ein $\mathbb{K}$-Vektorraum, so ist die Skalarmultiplikationsabbildung

   $$S\colon \mathbb{K} \to \mathrm{End}(V, +), \quad S(\lambda)(v) := \lambda v, \quad v \in V$$

   ein Morphismus von Ringen mit Eins.

3. Sei $(V, +)$ eine abelsche Gruppe und $\mathbb{K}$ ein Körper. Sei $S\colon \mathbb{K} \to \mathrm{End}(V, +)$ ein Morphismus von Ringen mit Eins. Dann definiert die Skalarmultiplikation

   $$\mathbb{K} \times V \to V, \quad (\lambda, v) \mapsto S(\lambda)(v)$$

   eine $\mathbb{K}$-Vektorraumstruktur auf V.

   Somit entsprechen die $\mathbb{K}$-Vektorraumstrukturen auf einer abelschen Gruppe $(V, +)$ genau den Morphismen von Ringen mit Eins $\mathbb{K} \to \mathrm{End}(V, +)$.[1]

Übung 4.2.6. Ist V ein Vektorraum, so ist $\mathrm{End}(V)$ ein Ring mit Eins und $\mathrm{Aut}(V) = \mathrm{End}(V)^\times$ ist die zugehörige Einheitengruppe. Insbesondere ist $\mathrm{Aut}(V)$ eine Gruppe bzgl. der Komposition mit neutralem Element id_V.

4.3 Untervektorräume

In diesem Abschnitt bezeichnet V einen Vektorraum über einem Körper $\mathbb{K}$.

Definition 4.3.1. Eine Teilmenge $U \subseteq V$ heißt *linearer Unterraum* oder *Untervektorraum* von V, falls sie die folgenden Eigenschaften hat:

(U1) $0 \in U$.

[1] Es gibt einen Begriff in der Algebra, der den Begriff des Vektorraums über einem Körper auf eine Situation verallgemeinert, in der die Skalare Elemente eines Ringes sind. Ist R ein Ring mit Eins und $(V, +)$ eine abelsche Gruppe, so sagt man, dass ein Morphismus von Ringen mit Eins $S\colon R \to \mathrm{End}(V, +)$ auf $(V, +)$ eine *R-Modul*-Struktur definiert. In diesem Sinne verallgemeinern Moduln über einem Ring Vektorräume über einem Körper.

(U2) $v, w \in U \Rightarrow v + w \in U$.

(U3) $v \in U,\ \lambda \in \mathbb{K} \Rightarrow \lambda v \in U$.

Bemerkung 4.3.2. Die Eigenschaften (U2) und (U3) können auch äquivalent in eine Aussage komprimiert werden:

$$(\forall v, w \in U)\ (\forall \lambda, \mu \in \mathbb{K}) \quad \lambda v + \mu w \in U.$$

Wiederholte Anwendung dieser Eigenschaft ergibt: Ist U ein Untervektorraum von V, so gilt

$$v_1, v_2, \ldots, v_n \in U, \quad \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K} \quad \Rightarrow \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in U.$$

Beispiele 4.3.3. Sei V ein $\mathbb{K}$-Vektorraum.

1. $\{0\}$ und V sind Untervektorräume von V.

2. Für jedes Element $v \in V$ ist die Menge $\mathbb{K}v = \{\lambda v : \lambda \in \mathbb{K}\}$ ein Untervektorraum von V; denn

   (U0) $0 = 0v \in \mathbb{K}v$.

   (U1) $\lambda v + \mu v = (\lambda + \mu)v$ für $\lambda, \mu \in \mathbb{K}$.

   (U2) $\lambda(\mu v) = (\lambda\mu)v$ für $\lambda, \mu \in \mathbb{K}$, d.h. $\lambda(\mathbb{K}v) \subseteq \mathbb{K}v$.

   Ist $v = 0$, so ist $\mathbb{K}v = \{0\}$. Ist $v \neq 0$, so heißt der Untervektorraum $\mathbb{K}v$ die *von v in V erzeugte Gerade.*

3. Die Untervektorräume von $\mathbb{R}^2$ sind: $\{0\}$; die Geraden $\mathbb{R}\begin{pmatrix} x \\ y \end{pmatrix} \subset \mathbb{R}^2$ im Sinne von Beispiel 2, für alle $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ in $\mathbb{R}^2$; sowie $\mathbb{R}^2$ selbst. (Einen formalen Beweis dafür können wir erst führen, wenn uns der Begriff der Dimension zur Verfügung steht.)

4. Eine *reelle Polynomfunktion vom Grade n* ist eine Funktion $p \colon \mathbb{R} \to \mathbb{R}$, deren Werte durch einen Ausdruck der Form

   $$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, \quad x \in \mathbb{R}$$

   gegeben werden, wobei die Koeffizienten $a_0, a_1, \ldots, a_n$ sämtlich reelle Zahlen sind, und $a_n \neq 0$ ist.

   Folgende Teilmengen des reellen Vektorraums $\mathcal{F}(\mathbb{R}, \mathbb{R})$ aller Funktionen $f \colon \mathbb{R} \to \mathbb{R}$ sind Untervektorräume:

   - $\mathcal{P}(\mathbb{R})$: die Menge aller reellen Polynomfunktionen
   - $\mathcal{P}_n(\mathbb{R})$: die Menge aller reellen Polynomfunktionen vom Grad $\leq n$.

Allgemeiner: Für irgendeinen Körper $\mathbb{K}$ heißt eine Funktion $p\colon \mathbb{K} \to \mathbb{K}$ eine Polynomfunktion, wenn die Werte $p(x)$ für alle $x \in \mathbb{K}$ durch einen Ausdruck der Form

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

gegeben sind, wobei $a_0, a_1, \ldots, a_n \in \mathbb{K}$ und $n \in \mathbb{N}_0$ ist. Die Polynomfunktionen bilden einen Untervektorraum des $\mathbb{K}$-Vektorraums $\mathcal{F}(\mathbb{K}, \mathbb{K})$.

Satz 4.3.4. *Ist U Untervektorraum eines $\mathbb{K}$-Vektorraums V, so ist U selbst ein $\mathbb{K}$-Vektorraum, bzgl. der Einschränkungen der Addition und der Skalarmultiplikation.*

Beweis. Definitionsgemäß gilt $0 \in U$, und die Einschränkung der Addition von $V \times V$ auf $U \times U$ liefert wegen (U2) eine Operation

$$(v, w) \mapsto v + w\colon U \times U \to U.$$

Ebenso liefert wegen (U3) die Einschränkung der Skalarmultiplikation auf Elemente $v \in U$ eine Skalarmultiplikation $(\lambda, v) \mapsto \lambda v\colon \mathbb{K} \times U \to U$. Die Axiome (V1)–(V8) gelten natürlich auch für Elemente von U. Daher ist U ein $\mathbb{K}$-Vektorraum. □

Bemerkung 4.3.5. Um zu zeigen, dass eine gegebene Struktur U ein Vektorraum ist, müssen im Prinzip die Axiome (V1)–(V8) nachgeprüft werden. Aber häufig ist es viel einfacher zu zeigen, dass U Untervektorraum eines schon bekannten Vektorraums V ist — dann ist auch U ein Vektorraum, nach dem eben bewiesenen Satz.

Die Mengen $\mathcal{P}(\mathbb{R})$ und $\mathcal{P}_n(\mathbb{R})$ reeller Polynome zum Beispiel bilden $\mathbb{R}$-Vektorräume, weil sie nach 4.3.3(d) Untervektorräume des Vektorraums $\mathcal{F}(\mathbb{R}, \mathbb{R})$ aller Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ sind.

Satz 4.3.6. *Ist $(U_i)_{i\in I}$ eine Familie von Untervektorräumen von V, so ist ihr Durchschnitt*

$$U := \bigcap_{i\in I} U_i = \{v \in V\colon (\forall i \in I)\ v \in U_i\}$$

ebenfalls ein Untervektorraum von V.

Beweis. **(U1)** Da jedes U_i ein Untervektorraum ist, gilt: $(\forall i \in I)\ 0 \in U_i$. Also $0 \in U$.

(U2) Seien $v, w \in U$. Dann ist also $v, w \in U_i$ für alle $i \in I$. Da die U_i Untervektorräume sind, folgt $v + w \in U_i$ für alle i, somit $v + w \in U$.

(U3) Sei $v \in U$ und $\lambda \in \mathbb{K}$. Dann ist $v \in U_i$ für alle $i \in I$. Da die U_i Untervektorräume sind, gilt $\lambda v \in U_i$ für alle i; mithin $\lambda v \in U$. □

Übung 4.3.7. Gilt auch die folgende Aussage? — „Sind U_1 und U_2 Untervektorräume von V, so ist $U_1 \cup U_2$ ebenfalls ein Untervektorraum.“

Definition 4.3.8. Sei V ein Vektorraum und $B \subseteq V$ eine Teilmenge von V. Nach dem vorangehenden Satz ist der Durchschnitt

$$\operatorname{span}(B) := \bigcap\{U \subseteq V : B \subseteq U, U \text{ Untervektorraum von } V\}$$

aller Unterräume, die B enthalten, wieder ein Untervektorraum von V. Er heißt *der von B aufgespannte* (oder: *erzeugte*) *Untervektorraum von V*. Er ist nach Definition Untervektorraum aller B enthaltenden Untervektorräume, und ist also in diesem sehr starken Sinne 'der kleinste' B enthaltende Untervektorraum von V. Ein Teil der Arbeit dieses Kapitels wird darin bestehen, ihn explizit zu beschreiben. Einen ersten Vorgeschmack gibt das folgende Beispiel.

Beispiele 4.3.9.

1. Zunächst erhalten wir für jeden Vektor $v \in V$ und $B = \{v\}$, dass $\operatorname{span}(\{v\}) = \mathbb{K}v$ (Beispiel 4.3.3(a)).

2. Für eine zweielementige Teilmenge $B = \{v, w\} \subset V$ muss $\operatorname{span}(B)$, da es ein Untervektorraum ist, alle Elemente der Form $\lambda v + \mu w$, für beliebige $\lambda, \mu \in \mathbb{K}$ enthalten. Andererseits ist die Menge $\{\lambda v + \mu w : \lambda, \mu \in \mathbb{K}\}$, wie man in Verallgemeinerung von Beispiel 4.3.3(a) oben zeigen kann, ein Untervektorraum von V, gehört also zur Familie, über die bei der Bildung von $\operatorname{span}(B)$ geschnitten wird. Folglich ist

   $$\operatorname{span}(B) = \{\lambda v + \mu w : \lambda, \mu \in \mathbb{K}\}.$$

   Sind v und w weder beide 0, noch in einer Geraden enthalten [man sagt dann: *sind v und w linear unabhängig*], so ist dieser Untervektorraum die von v und w erzeugte *Ebene* in V.

3. In $\mathbb{R}^3$ gibt es folgende Untervektorräume: $\{0\}$; die Geraden $\mathbb{R}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ im Sinne von Beispiel 4.3.3.2, für alle $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ in $\mathbb{R}^3$; die durch irgend zwei linear unabhängige $v, w \in \mathbb{R}^3$ erzeugten Ebenen $\operatorname{span}(\{v, w\})$; sowie $\mathbb{R}^3$.

4. Für jede Teilmenge B eines $\mathbb{K}$-Vektorraums V gilt

   $$\operatorname{span}\big(\operatorname{span}(B)\big) = \operatorname{span}(B).$$

   Dies folgt sofort aus der Definition: auf beiden Seiten steht der Schnitt über *dieselbe* Menge von Untervektorräumen.

Definition 4.3.10. [Summe von Untervektorräumen] Für Untervektorräume U_1,U_2,$\ldots$, U_n von V definiert man die *Summe* wie folgt:

$$U_1 + U_2 + \cdots + U_n := \{u_1 + u_2 + \cdots + u_n : u_j \in U_j, j = 1, \ldots, n\}$$

Satz 4.3.11. *Die Summe $U_1+U_2+\cdots+U_n$ von Untervektorräumen $U_1, U_2, \ldots, U_n$ von V ist ein Untervektorraum von V.*

Wir überlassen den direkten Beweis dieses Satzes dem Leser zur Übung. — Man kann diesen Beweis übrigens auch dadurch führen, dass man

$$U_1 + U_2 + \cdots + U_n = \operatorname{span}(B)$$

nachweist, wobei $B = U_1 \cup \cdots \cup U_n$ ist. — Vgl. Satz 4.5.2 unten. □

Beispiel 4.3.12. Wir erinnern an die Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

im $\mathbb{R}^3$ und betrachten die beiden Ebenen.

$$U_1 = \{\lambda_1 e_1 + \lambda_2 e_2 : \lambda_1, \lambda_2 \in \mathbb{R}\} = \mathbb{R}e_1 + \mathbb{R}e_2 = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}$$

und

$$U_2 = \{\lambda_1 e_1 + \lambda_3 e_3 : \lambda_1, \lambda_3 \in \mathbb{R}\} = \mathbb{R}e_1 + \mathbb{R}e_3 = \left\{ \begin{pmatrix} x \\ 0 \\ z \end{pmatrix} : x, z \in \mathbb{R} \right\}.$$

Dann ist der Schnitt

$$U_1 \cap U_2 = \{\lambda_1 e_1 : \lambda_1 \in \mathbb{R}\} = \mathbb{R}e_1 = \left\{ \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}$$

eine Gerade und die Summe $U_1 + U_2 = \mathbb{R}^3$.

Das Element $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3$ lässt sich auf mehr als eine Weise als Summe $v = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$ darstellen:

$$v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \underbrace{e_1 + e_2}_{\in U_1} + \underbrace{e_3}_{\in U_2} = \underbrace{e_2}_{\in U_1} + \underbrace{e_1 + e_3}_{\in U_2}.$$

Definition 4.3.13. Seien $U_1, U_2, \ldots, U_n$ Untervektorräume von V. Man sagt, dass V die *direkte Summe* von $U_1, U_2, \ldots, U_n$ ist, und schreibt

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_n,$$

falls jedes Element $v \in V$ eine *eindeutige* Darstellung der Form

$$v = u_1 + u_2 + \cdots + u_n \qquad \text{mit} \qquad u_i \in U_i \;\; \text{für} \;\; i = 1, 2, \ldots, n$$

besitzt.

Satz 4.3.14. *Seien U_1 und U_2 Untervektorräume von V. Es gilt genau dann $V = U_1 \oplus U_2$, wenn folgende zwei Bedingungen erfüllt sind:*

(1) $V = U_1 + U_2$,

(2) $U_1 \cap U_2 = \{0\}$.

Beweis. "$\Rightarrow$": Angenommen $V = U_1 \oplus U_2$, so folgt insbesondere $V = U_1 + U_2$. Um auch (2) nachzuprüfen, sei $v \in U_1 \cap U_2$. Dann hat v zwei Darstellungen:

$$v = \underbrace{v}_{\in U_1} + \underbrace{0}_{\in U_2} = \underbrace{0}_{\in U_1} + \underbrace{v}_{\in U_2}$$

als Summe eines Elements von U_1 und eines Elements von U_2. Aus der Eindeutigkeit der Darstellung folgt dann also $v = 0$, mithin $U_1 \cap U_2 = \{0\}$.

"$\Leftarrow$": Gelten umgekehrt (1) und (2), so lässt sich wegen (2) jedes $v \in V$ als

$$v = u_1 + u_2 \quad \text{mit } u_1 \in U_1,\ u_2 \in U_2$$

schreiben. Um die Eindeutigkeit der Darstellung nachzuweisen, nehme an, man hätte

$$v = u_1' + u_2' \quad \text{mit } u_1' \in U_1,\ u_2' \in U_2$$

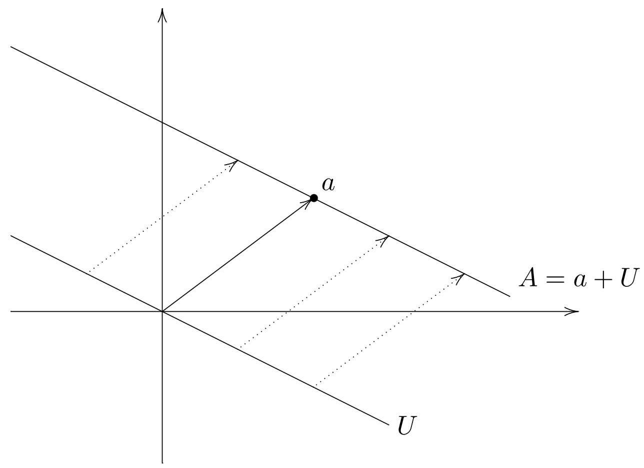
Durch Abziehen bekommt man $0 = (u_1 - u_1') + (u_2 - u_2')$, d.h. $u_1 - u_1' = u_2' - u_2$. Da U_1 und U_2 Untervektorräume sind, gilt $u_1 - u_1' \in U_1$ und $u_2' - u_2 \in U_2$, folglich

$$u_1 - u_1' = u_2' - u_2 \in U_1 \cap U_2$$

Aber $U_1 \cap U_2 = \{0\}$ nach (4.3.14(ii)). Daraus folgt $u_1 - u_1' = 0 = u_2' - u_2$, und somit $u_1 = u_1'$ und $u_2 = u_2'$. Die Darstellung von v als Summe eines Elements $u_1 \in U_1$ und eines Elements $u_2 \in U_2$ ist also eindeutig. □

Definition 4.3.15. [Affine Unterräume] Eine Teilmenge A eines $\mathbb{K}$-Vektorraums V heißt *affiner Unterraum* von V, wenn es einen Untervektorraum U von V und ein Element $a \in V$ gibt, so dass $A = a + U = \{a + u : u \in U\}$. — Mit anderen Worten: Affine Unterräume sind nichts anderes als 'parallelverschobene', oder 'translatierte' Untervektorräume.

Bemerkung 4.3.16. (1) Ein Untervektorraum enthält stets die 0; ein affiner Teilraum enthält sie genau dann, wenn er schon selbst ein Untervektorraum ist. (Übung.)



(2) Zwei affine Unterräume zum selben Untervektorraum: $a + U$ und $b + U$ (für $a, b \in V$) sind genau dann gleich, wenn $a - b \in U$:

$$a + U = b + U \;\Leftrightarrow\; a - b \in U \;\Leftrightarrow\; b - a \in U.$$

(3) Für jedes $b \in a + U$ ist $a + U = b + U$, was sofort aus (2) und $b - a \in U$ folgt.

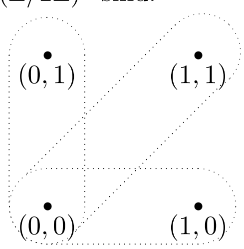
(4) Aus (2) folgt sofort, dass durch $a \sim b :\Leftrightarrow a - b \in U$ eine Äquivalenzrelation auf V definiert wird.

Beispiele 4.3.17.

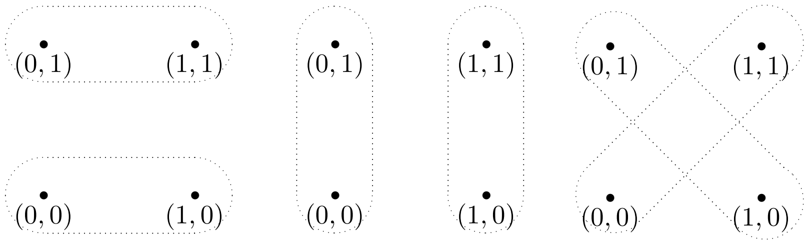
1. Affine Unterräume von $\mathbb{R}^2$ sind Punkte: $\{a\} = a + \{0\}$, beliebige Geraden (durch $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ oder nicht), und $\mathbb{R}^2$ selbst.

2. Affine Unterräume von $\mathbb{R}^3$ sind Punkte $\{a\}$, beliebige Geraden, beliebige Ebenen, sowie $\mathbb{R}^3$ selbst.

3. Untervektorräume von $(\mathbb{Z}/2\mathbb{Z})^2$ sind:



Die entsprechenden affinen Unterräume sind daher:



Satz 4.3.18. (Lineare Abbildungen und Unterräume) *Sei $\varphi\colon V \to W$ eine lineare Abbildung.*

(1) *Ist U ein Untervektorraum von V, so ist sein Bild $\varphi(U)$ ein Untervektorraum von W.*

(2) *Ist U ein Untervektorraum von W, so ist sein Urbild*
$$\varphi^{-1}(U) := \{v \in V \mid \varphi(v) \in U\}$$
ein Untervektorraum von V.

(3) *Ist A ein affiner Unterraum von V, so ist sein Bild $\varphi(A)$ ein affiner Unterraum von W.*

(4) *Ist B ein affiner Unterraum von W, so ist sein Urbild $\varphi^{-1}(B)$ entweder ein affiner Unterraum von V oder die leere Menge. Das Urbild von B ist genau dann leer, wenn $B \cap \varphi(V) = \emptyset$.*

Beweis. (1) Sei $v_1, v_2 \in \varphi(U)$. Dann gibt es $u_1, u_2 \in U$ mit $\varphi(u_1) = v_1$ und $\varphi(u_2) = v_2$. Für $\lambda, \mu \in \mathbb{K}$ haben wir dann $\lambda u_1 + \mu u_2 \in U$ und daher
$$\lambda v_1 + \mu v_2 = \lambda\varphi(u_1) + \mu\varphi(u_2) = \varphi(\lambda u_1 + \mu u_2) \in \varphi(U).$$

(2) (U1) folgt aus $\varphi(0) = 0 \in U$, weil U Untervektorraum ist.

Um (U2) einzusehen, seien $u, v \in \varphi^{-1}(U)$; dann ist $\varphi(u) \in U$ und $\varphi(v) \in U$. Da U Untervektorraum ist, erfüllt er (U1) und so kommt $\varphi(u)+\varphi(v) \in U$. Aber $\varphi(u) + \varphi(v) = \varphi(u + v)$ nach (L1). Mithin $u + v \in \varphi^{-1}(U)$.

Nun zu (U3). Ist $v \in \varphi^{-1}(U)$, so $\varphi(v) \in U$. Als Untervektorraum erfüllt U (U3) und daher ist $\lambda\varphi(v) \in U$ für alle $\lambda \in \mathbb{K}$. Da $\lambda\varphi(v) = \varphi(\lambda v)$ nach (L2), schließen wir: $\lambda v \in \varphi^{-1}(U)$.

(3) Sei A affiner Unterraum von V. Dann gibt es einen Untervektorraum U und ein Element $a \in V$ so dass $A = a + U$. Folglich ist $\varphi(A) = \varphi(a + U) = \varphi(a) + \varphi(U)$ nach (L1). Gemäß (1) ist $\varphi(U)$ ein Untervektorraum von W. Also ist $\varphi(A)$ ein affiner Unterraum von W.

(4) Schreibe $B = b + Y$. Sei $\varphi^{-1}(B) \neq \emptyset$ und wähle $a \in V$ mit $\varphi(a) \in B$. Gemäß Bemerkung 4.3.16(3) ist dann $B = \varphi(a) + Y$. Daher ist $\varphi(x) \in B$ äquivalent zu
$$\varphi(x - a) = \varphi(x) - \varphi(a) \in B - \varphi(a) = Y,$$
also
$$\varphi^{-1}(B) = a + \varphi^{-1}(Y).$$
□

Korollar 4.3.19. *Ist $\varphi\colon V \to W$ eine lineare Abbildung und $b \in W$, so ist die Menge*

$$\{x \in V\colon \varphi(x) = b\} = \varphi^{-1}(b)$$

aller Lösungen der linearen Gleichung $\varphi(x) = b$ entweder leer oder ein affiner Unterraum von V.

Beweis. Wende (4) des vorstehenden Satzes auf den affinen Unterraum $\{b\}$ von W an. $\square$

Beispiel 4.3.20. Die Lösungsmenge eines beliebigen linearen Gleichungssystems

$$Ax = b \qquad [A \in M_{m,n}(\mathbb{K}),\ b \in \mathbb{K}^m]$$

ist stets entweder leer oder ein affiner Unterraum von V. Denn wir wissen aus Beispiel 4.2.3(2), dass die Abbildung $\mathbb{K}^n \longrightarrow \mathbb{K}^m, x \mapsto Ax$, linear ist.

Übung 4.3.21. Finde alle Untervektorräume und alle affinen Unterräume von $(\mathbb{Z}/2\mathbb{Z})^3$.

Übung 4.3.22. Eine Teilmenge A eines Vektorraums V ist genau dann ein affiner Unterraum, wenn A nicht leer ist und für $a, b \in A$ und $\lambda \in \mathbb{K}$ gilt:

$$\lambda a + (1 - \lambda) b \in A.$$

Was besagt dies geometrisch?

4.4 Direkte Produkte

Definition 4.4.1. Gegeben $\mathbb{K}$-Vektorräume $V_1, V_2, \ldots, V_n$, definieren wir auf dem kartesischen Produkt

$$V := V_1 \times V_2 \times \cdots \times V_n$$

eine Addition und eine Skalarmultiplikation komponentenweise, d.h. durch die Regeln:

$$(v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n) := (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$$

und

$$\lambda \cdot (v_1, v_2, \ldots, v_n) := (\lambda v_1, \lambda v_2, \ldots, \lambda v_n).$$

Man prüft leicht nach, dass die Axiome (V1)–(V8) gelten, so dass

$$V_1 \times V_2 \times \cdots \times V_n$$

ein Vektorraum ist. Er heißt das *direkte Produkt* der Vektorräume $V_1, V_2, \ldots, V_n$.

Als Spezialfall dieser Definition beachte man den uns schon wohlbekannten Vektorraum:

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{n \text{ mal}}.$$

Bemerkung 4.4.2. Direkte Produkte und direkte Summen hängen eng zusammen: Ist $V = V_1 \oplus \ldots \oplus V_n$ eine direkte Summe, so ist die Summationsabbildung

$$S\colon V_1 \times \ldots \times V_n \to V, \quad (v_1, \ldots, v_n) \mapsto v_1 + \ldots + v_n$$

ein Isomorphismus von Vektorräumen.

Umgekehrt ist für Vektorräume $V_1, \ldots, V_n$ das Produkt

$$V := V_1 \times \ldots \times V_n$$

eine direkte Summe von Unterräumen

$$U_i := \{(v_1, \ldots, v_n)\colon j \neq i \;\Rightarrow\; v_j = 0\} = \{\underbrace{(0, \ldots, 0)}_{j-1 \text{ mal}}\} \times V_i \times \{\underbrace{(0, \ldots, 0)}_{n-j \text{ mal}}\},$$

welche isomorph zu den Vektorräumen V_i sind. Die Abbildung

$$\eta_i\colon V_i \to U_i, \quad v \mapsto (\underbrace{0, \ldots, 0}_{j-1 \text{ mal}}, v, \underbrace{0, \ldots, 0}_{n-j \text{ mal}})$$

ist ein Vektorraumisomorphismus. Jedes Element von V kann offenbar auf genau eine Weise als

$$(v_1, \ldots, v_n) = \sum_{j=1}^{n} \eta_j(v_j) = (v_1, 0, \ldots, 0) + \ldots + (0, \ldots, 0, v_n)$$

geschrieben werden, wo die einzelnen Summanden Elemente der Unterräume U_j sind.

4.5 Linearkombinationen, lineare Abhängigkeit, Basen

In diesem Anschnitt bezeichnet V stets einen Vektorraum über dem Körper $\mathbb{K}$, und $v_1, v_2, \ldots, v_m$ bezeichnen Elemente von V.

Definition 4.5.1. [Linearkombinationen] Jedes Element der Form

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m = \sum_{j=1}^{m} \lambda_j v_j,$$

wo $\lambda_1, \lambda_2, \ldots, \lambda_m$ beliebige Skalare aus $\mathbb{K}$ sind, heißt *Linearkombination* von $v_1, v_2, \ldots, v_m$. Die Menge aller dieser Linearkombinationen schreiben wir als

$$\operatorname{span}(v_1, v_2, \ldots, v_m) := \{\lambda_1 v_1 + \cdots + \lambda_m v_m\colon \lambda_1, \ldots, \lambda_m \in \mathbb{K}\} = \mathbb{K}v_1 + \ldots + \mathbb{K}v_m.$$

Aus Beispiel 4.3.3(2) wissen wir, dass für jedes $j \in \{1, \ldots, m\}$ die Teilmenge $\mathbb{K}v_j$ von V ein Untervektorraum ist. Also ist $\operatorname{span}(v_1, \ldots, v_m)$ als Summe von Untervektorräumen selbst ein Untervektorraum nach Satz 4.3.11.

Korollar 4.5.2. $\operatorname{span}(v_1, v_2, \ldots, v_m)$ *ist der kleinste Untervektorraum von* V*, der* v_1*,* v_2*,. . . ,* v_m *enthält. Mit anderen Worten* (siehe Definition 4.3.8)*:*

$$\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(\{v_1, \ldots, v_m\}).$$

Beweis. Zunächst enthält $\operatorname{span}(v_1, v_2, \ldots, v_m)$ alle v_i $(i = 1, \ldots, n)$; denn

$$v_i = 0 \cdot v_1 + \cdots + 0 \cdot v_{i-1} + 1 \cdot v_i + 0 \cdot v_{i+1} + \cdots + 0 \cdot v_m.$$

Es bleibt nur zu bemerken, dass jeder Untervektorraum von V, der v_1,. . . ,v_m enthält, auch sämtliche Linearkombinationen der v_i enthalten muss. □

Aus diesem Korollar folgt sofort weiter:

Korollar 4.5.3. *Sind* $w_1, w_2, \ldots, w_k \in \operatorname{span}(v_1, v_2, \ldots, v_m)$*, so ist*

$$\operatorname{span}(w_1, w_2, \ldots, w_k) \subseteq \operatorname{span}(v_1, v_2, \ldots, v_m).$$

Satz 4.5.4. *Ist* $B \subseteq V$ *eine Teilmenge, so ist*

$$\operatorname{span}(B) = \bigcup_{E \subseteq B \text{ endlich}} \operatorname{span}(E) = \{\lambda_1 b_1 + \ldots + \lambda_n b_n : n \in \mathbb{N}, b_i \in B, \lambda_i \in \mathbb{K}\}.$$

Beweis. Jede (insbesondere jede endliche) Teilmenge $E \subseteq B$ ist in $\operatorname{span}(B)$ enthalten. Also enthält $\operatorname{span}(B)$ als Untervektorraum auch $\operatorname{span}(E)$. Daraus folgt die Inklusion

$$F := \bigcup_{E \subseteq B \text{ endlich}} \operatorname{span}(E) \subseteq \operatorname{span}(B).$$

Die Menge F enthält alle endlichen Teilmengen von B, insbesondere enthält sie B.
Wir zeigen jetzt, dass F ein Untervektorraum von V ist:
(U1): $0 \in F$; denn $0 \in \operatorname{span}(E)$ für jedes endliche $E \subset B$.
(U2/3): Seien $u, v \in F$ und $\lambda, \mu \in \mathbb{K}$. Dann gibt es endliche Teilmengen $E_1 \subseteq B$ und $E_2 \subseteq B$ mit $u \in \operatorname{span}(E_1)$ und $v \in \operatorname{span}(E_2)$. Somit $u, v \in \operatorname{span}(E_1 \cup E_2)$, woraus $\lambda u + \mu v \in \operatorname{span}(E_1 \cup E_2) \subseteq F$ folgt. Also ist F wie behauptet ein Untervektorraum von V, welcher B enthält.

Damit folgt $\operatorname{span}(B) \subseteq F$, und daraus die behauptete Gleichheit. □

Definition 4.5.5. (1) Man sagt, U sei von v_1, v_2,. . . ,v_m *aufgespannt* oder *erzeugt*, falls

$$U = \operatorname{span}(v_1, v_2, \ldots, v_m)\ ,$$

d.h. falls jedes Element $v \in U$ als Linearkombination von $v_1, v_2, \ldots, v_m$ geschrieben werden kann. Dann nennt man auch v_1, v_2,. . . ,v_m ein *Erzeugendensystem* von U.

(2) Allgemeiner sagt man, dass ein Untervektorraum U *von* V *von einer Teilmenge* B *erzeugt* wird, wenn $U = \operatorname{span}(B)$ gilt. Die Menge B heißt dann ein *Erzeugendensystem* von U. Der Untervektorraum U besteht in diesem Fall, wie wir gesehen haben, aus sämtlichen Linearkombinationen endlich vieler Elemente von B.

Beispiele 4.5.6.

1. $\mathbb{K}^n$ wird erzeugt durch

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \dots \qquad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

   Aber es gibt sehr viele andere Erzeugendensysteme von $\mathbb{K}^n$. Zum Beispiel wird $\mathbb{R}^2$ auch durch $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ aufgespannt.

2. Der Vektorraum $\mathcal{P}_n(\mathbb{R})$ der reellen Polynome vom Grade $\leq n$ wird von p_0, p_1,..., p_n aufgespannt, wobei

$$p_j(x) = x^j \quad \text{für alle } x \in \mathbb{R} \ .$$

Definition 4.5.7. [Lineare Unabhängigkeit]

1. Die Elemente $v_1, v_2, \dots, v_m$ des Vektorraums V heißen *linear unabhängig*, falls keines der v_i als Linearkombination der anderen v_j $(j \neq i)$ dargestellt werden kann. Andernfalls heißen sie *linear abhängig*.

2. Allgemeiner heißt eine Teilmenge $B \subseteq V$ *linear unabhängig*, falls für alle $b \in B$ gilt: $b \notin \mathrm{span}(B \setminus \{b\})$. In diesem Sinne sind also n verschiedene Elemente $v_1, \dots, v_n$ von V genau dann linear unabhängig, wenn die Menge $\{v_1, \dots, v_n\}$ linear unabhängig ist.

Satz 4.5.8. (Kriterium für lineare Unabhängigkeit) *Elemente* $v_1, v_2, \dots, v_m \in V$ *sind genau dann linear unabhängig, wenn für alle Skalare* $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ *die folgende Implikation gilt:*

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \cdots = \lambda_m = 0.$$

Beweis. "$\Leftarrow$": Seien $v_1, v_2, \dots, v_m$ linear abhängig. Dann gibt es ein i, so dass v_i Linearkombination der v_j $(j \neq i)$ ist:

$$v_i = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1} + \mu_{i+1} v_{i+1} + \cdots + \mu_m v_m$$

für geeignete Skalare μ_j $(j \neq i)$. Dann ist aber

$$0 = \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1} - v_i + \mu_{i+1} v_{i+1} + \cdots + \mu_m v_m,$$

wo nicht alle Koeffizienten 0 sind; denn der i-te Koeffizient ist -1.

"$\Rightarrow$": Sei andererseits $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ mit Koeffizienten λ_i, die nicht alle 0 sind. Ist etwa $\lambda_1 \neq 0$, so folgt

$$v_1 = -\lambda_1^{-1}(\lambda_2 v_2 + \cdots + \lambda_m v_m) = -\frac{\lambda_2}{\lambda_1} v_2 - \cdots - \frac{\lambda_m}{\lambda_1} v_m,$$

d.h. v_1 ist Linearkombination von $v_2, \ldots, v_m$, und folglich sind $v_1, v_2, \ldots, v_m$ linear abhängig. □

Korollar 4.5.9. (Kriteriun für lineare Unabhängigkeit von Teilmengen) *Eine Teilmenge $B \subseteq V$ ist genau dann linear unabhängig, wenn jede endliche Teilmenge von B linear unabhängig ist.*

[Man sagt auch: die Eigenschaft, linear unabhängig zu sein, sei 'von endlichem Typ' oder 'von endlichem Charakter'.]

Beweis. Nehmen wir zuerst B als linear unabhängig an. Sei $E \subseteq B$ eine endliche Teilmenge. Ist E nicht linear unabhängig, so gibt es ein Element $b \in E$ mit $b \in \operatorname{span}(E \setminus \{b\})$. Also ist auch B nicht linear unabhängig.

Ist umgekehrt $B \subseteq V$ eine linear abhängige Teilmenge, so gibt es ein Element $b \in B$ mit $b \in \operatorname{span}(B \setminus \{b\})$. Dann gibt es also Elemente $b_1, \ldots, b_m \in B \setminus \{b\}$ mit $b \in \operatorname{span}(b_1, \ldots, b_m)$ (Satz 4.5.4). Die endliche Teilmenge $E = \{b, b_1, \ldots, b_m\} \subseteq B$ ist also ebenfalls linear abhängig. □

Beispiele 4.5.10.

1. $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sind linear unabhängig in $\mathbb{R}^2$; denn keiner dieser Spaltenvektoren ist ein skalares Vielfaches des anderen. Ebenso sind $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ linear unabhängig in $\mathbb{R}^2$.

2. $e_1, e_2, \ldots, e_n$ sind linear unabhängig in $\mathbb{K}^n$.

3. $p_0, p_1, \ldots, p_n$ sind linear unabhängig in $\mathcal{P}_n(\mathbb{R})$.

   Diese Tatsache verdient, hier im Skriptum bewiesen zu werden. Zuerst zur Erinnerung: $p_i : \mathbb{R} \to \mathbb{R}$ ist die Funktion $x \mapsto x^i$. Für die lineare Unabhängigkeit der p_i $(i = 0, \ldots, n)$ ist nach 4.5.10 zu zeigen, dass für $a_0, \ldots, a_n \in \mathbb{R}$ aus der Annahme, dass $p = \sum_{i=0}^n a_i p_i$ die konstante Nullfunktion ist, die Gleichungen $a_0 = a_1 = \cdots = a_n = 0$ folgen. Wäre jetzt irgendein a_i, für ein $i > 0$, nicht 0, so gäbe es auch ein größtes ν, mit $0 < \nu \leq n$, so dass $a_\nu \neq 0$; d.h. p wäre eine Polynomfunktion vom Grade $\nu \geq 1$. Nun gilt aber das

   Lemma 4.5.11. *Ist $\mathbb{K}$ ein Körper, so besitzt eine Polynomfunktion $p \colon \mathbb{K} \to \mathbb{K}$ vom Grad $\leq n$ höchstens n verschiedene Nullstellen.*

   Beweis. Wir schreiben p as $p(x) = \sum_{k=0}^n a_k x^k$ mit $a_n \neq 0$ und zeigen die Behauptung durch Induktion nach dem Grad n des Polynoms. Ist der Grad 1, so ist die Behauptung trivial, denn aus $p(x) = a_0 + a_1 x$ folgt, dass p nur die Nullstelle $x_0 = -a_0/a_1$ besitzt.

Sei $n > 1$ der Grad von p, also o.B.d.A. $a_n \neq 0$. Ist x_0 eine Nullstelle von p, so erhalten wir mit

$$\begin{aligned} p(x) &= \sum_{k=0}^{n} a_k (x - x_0 + x_0)^k = \sum_{k=0}^{n} a_k \sum_{j=0}^{k} \binom{k}{j} x_0^{k-j} (x - x_0)^j \\ &= \sum_{j=0}^{n} \Big(\sum_{k=j}^{n} a_k \binom{k}{j} x_0^{k-j} \Big) (x - x_0)^j \end{aligned}$$

eine Darstellung von p als

$$p(x) = \sum_{j=0}^{n} b_j (x - x_0)^j.$$

Nun ist

$$0 = p(x_0) = \sum_{j=0}^{n} b_j (x_0 - x_0)^j = b_0.$$

Also ist $p(x) = (x - x_0) \sum_{j=0}^{n-1} b_{j+1} (x - x_0)^j$. Da der zweite Faktor nach Induktionsvoraussetzung höchstens $n - 1$ Nullstellen besitzt, sehen wir, dass p höchstens n Nullstellen hat. $\square$

Unser p hat jedoch unendlich viele Nullstellen — also mehr als ν, für jedes ν —; denn *jede* reelle Zahl wird von der Nullfunktion auf Null abgebildet, und es gibt unendlich viele reelle Zahlen. Das ist ein Widerspruch zur Annahme, dass ein Koeffizient $a_i \neq 0$ sein kann. Alle a_i sind also Null, wie zu beweisen war.

4. Die Zeilen $w_1, w_2, \ldots, w_r$ einer Matrix in Stufenform, welche nicht identisch 0 sind, sind linear unabhängig :

$$\begin{pmatrix}
0 & \ldots & 0 & a_{1j_1} & \ldots & a_{1j_2} & \ldots & a_{1j_r} & \ldots & a_{1n} \\
0 & \ldots & \ldots & \ldots & 0 & a_{2j_2} & \ldots & a_{2j_r} & \ldots & a_{2n} \\
\vdots & & & & & & \ddots & \vdots & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & a_{rj_r} & \ldots & a_{rn} \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \ldots & 0 \\
\vdots & & & & & & & \vdots & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \ldots & 0
\end{pmatrix}$$

In der Tat: Heißen die ersten r Zeilen $v_1, \ldots, v_r$ und nehmen wir an, dass für gewisse Skalare $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ gilt:

$$v := \sum_{i=1}^{r} \lambda_i v_i = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = (0, 0, \ldots, 0),$$

so ist die j_1-te Komponente von v gleich $\lambda_1 a_{1j_1} = 0$, woraus wegen $a_{1j_1} \neq 0$ folgt: $\lambda_1 = 0$. Weiter ist die j_2-te Komponente von v gleich $\lambda_1 a_{1j_2} + \lambda_2 a_{2j_2} = 0$, woraus wegen $\lambda_1 = 0$ und $a_{2j_2} \neq 0$ folgt: $\lambda_2 = 0$. Und so schließt man induktiv fort, bis man $\lambda_1 = \cdots = \lambda_{r-1} = 0$ weiß und die Gleichung für die j_r-te Komponente von v erhält:

$$\lambda_1 a_{1j_r} + \lambda_2 a_{2j_r} + \cdots + \lambda_r a_{rj_r} = 0,$$

woraus dann auch $\lambda_r = 0$ folgt, womit die lineare Unabhängigkeit bewiesen ist.

Satz 4.5.12. *Sei V ein Vektorraum. Für eine Teilmenge $B \subseteq V$ sind folgende Bedingungen äquivalent:*

(1) *B ist ein minimales Erzeugendensystem von V, d.h. $V = \operatorname{span}(B)$, aber keine echte Teilmenge von B hat diese Eigenschaft,*

(2) *B ist maximal linear unabhängig, d.h. B ist linear unabhängig, aber für jedes $v \in V \setminus B$ ist die Menge $B \cup \{v\}$ linear abhängig.*

Beweis. (1) $\Rightarrow$ (2): Sei B ein minimales Erzeugendensystem von V. Wir zeigen zuerst, dass B linear unabhängig ist. Wäre das nicht der Fall, so gäbe es ein $b_0 \in B$ mit $b_0 \in \operatorname{span}(B \setminus \{b_0\})$. Da B ganz V erzeugt, müsste schon $B \setminus \{b_0\}$ ganz V erzeugen, im Widerspruch zu (1). Also folgt aus (1), dass B linear unabhängig ist.

Um einzusehen, dass B auch maximal linear unabhängig ist, sei $v \in V \setminus B$ und $B' := B \cup \{v\}$. Dann ist B' is linear abhängig, weil

$$v \in V = \operatorname{span}(B) = \operatorname{span}(B' \setminus \{v\}).$$

(2) $\Rightarrow$ (1): Sei nun $B \subseteq V$ eine maximal linear unabhängige Teilmenge. Zuerst zeigen wir, dass B ganz V erzeugt: Gegeben $v \in V$, ist $B \cup \{v\}$ linear abhängig. Dann gibt es also eine endliche Teilmenge $E \subseteq B \cup \{v\}$, die linear abhängig ist (nach Korollar 4.5.9). Da jede endliche Teilmenge von B linear unabhängig ist, muss $v \in E$ sein: $E = \{v, b_1, \ldots, b_m\}$ für gewisse paarweise verschiedene Elemente $b_1, \ldots, b_m \in B$. Dann gibt es also $\lambda_0, \ldots, \lambda_m \in \mathbb{K}$, nicht alle null, so dass

$$0 = \lambda v + \lambda_1 b_1 + \ldots + \lambda_m b_m.$$

Aus der linearen Unabhängigkeit von $b_1, \ldots, b_m$ folgt $\lambda \neq 0$, und so erhalten wir

$$v = -\lambda^{-1}(\lambda_1 b_1 + \ldots + \lambda_m b_m) \in \operatorname{span}(B).$$

Damit ist B ein Erzeugendensystem.

Dass B minimal mit dieser Eigenschaft ist, folgt aus der linearen Unabhängigkeit: für jedes $b \in B$ gilt $b \notin \operatorname{span}(B \setminus \{b\})$. $\square$

Definition 4.5.13. Eine Teilmenge B eines Vektorraums V heißt *Basis von V*, falls sie V erzeugt und linear unabhängig ist.

Aus der linearen Unabhängigkeit einer Basis B folgt, dass sie ein minimales Erzeugendensystem von V ist. Umgekehrt ist nach Satz 4.5.12 jedes minimale Erzeugendensystem von V linear unabhängig und also eine Basis.

Genauso sieht man, dass die Basen von V genau die maximal linear unabhängigen Teilmengen von V sind.

Bemerkung 4.5.14. (a) Ein System $v_1, v_2, \dots, v_m$ von m Elementen des Vektorraums V bildet genau dann eine Basis von V, wenn es linear unabhängig ist und wenn jedes $v \in V$ als Linearkombination von $v_1, \dots, v_m$ geschrieben werden kann.

(b) Allgemeiner, ist eine Teilmenge $B \subseteq V$ genau dann eine Basis, wenn B linear unabhängig ist und sich jedes Element $v \in V$ in der Form

$$v = \sum_{b \in B} \lambda_b b$$

darstellen lässt, wobei nur endlich viele λ_b von 0 verschieden sind, die Summe also endlich ist.

Beispiele 4.5.15.

1. $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ist eine Basis von $\mathbb{R}^2$. Ebenso bilden aber auch $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ eine Basis von $\mathbb{R}^2$.
2. $e_1, e_2, \dots, e_n$ ist eine Basis von $\mathbb{K}^n$.
3. $p_0, p_1, \dots, p_n$ bilden eine Basis des reellen Vektorraums $\mathcal{P}_n(\mathbb{R})$.

4.5.1 Wie findet man eine Basis?

Wir stellen jetzt einen Algorithmus vor, der eine Basis eines durch Erzeugende gegebenen Untervektorraums $U = \operatorname{span}(v_1, v_2, \dots, v_m)$ von $\mathbb{K}^n$ liefert. Dazu benötigen wir zunächst einige Vorbereitungen.

4.5.16. [Vorbereitende Überlegungen] Seien $v_1, v_2, \dots, v_m$ und $w_1, w_2, \dots, w_r$ Elemente eines $\mathbb{K}$-Vektorraums V.

1. *Hat man*
$$w_1, w_2, \dots, w_r \in \operatorname{span}(v_1, v_2, \dots, v_m)$$
sowie
$$v_1, v_2, \dots, v_m \in \operatorname{span}(w_1, w_2, \dots, w_r),$$
dann ist
$$\operatorname{span}(w_1, w_2, \dots, w_r) = \operatorname{span}(v_1, v_2, \dots, v_m).$$
In Worten: lassen sich alle w_1, w_2, …, w_r als Linearkombinationen der v_1, v_2, …, v_m schreiben und umgekehrt, so erzeugen diese beiden endlichen Systeme denselben Untervektorraum.

2. Aus (1) ersehen wir, dass die folgenden *elementaren Umformungen* eines Systems von Vektoren $v_1, v_2, \dots, v_m$ den von ihm erzeugten Untervektorraum nicht verändern:

   (E1) Ersetze v_k durch $v_k + \lambda v_i$ für ein $i \neq k$ und ein $\lambda \in \mathbb{K}$. Zum Beispiel:
   $$\operatorname{span}(v_1, v_2, \dots, v_m) = \operatorname{span}(v_1, v_2 + \lambda v_1, \dots, v_m)$$

   (E2) Beliebige Vertauschungen der Vektoren. — Zum Beispiel:
   $$\operatorname{span}(v_1, v_2, \dots, v_m) = \operatorname{span}(v_2, v_1, \dots, v_m)$$

   (E3) Multiplikation eines der v_i mit einem Skalar $\lambda \neq 0$. — Zum Beispiel:
   $$\operatorname{span}(v_1, v_2, \dots, v_m) = \operatorname{span}(\lambda v_1, v_2, \dots, v_m) \quad \text{mit} \quad \lambda \neq 0.$$

4.5.17. [Algorithmus] Der folgende Algorithmus liefert eine Basis des von einem System $v_1, v_2, \dots, v_m$ von Elementen von $\mathbb{K}^n$ erzeugten Untervektorraums $U = \operatorname{span}(v_1, v_2, \dots, v_m) \subseteq \mathbb{K}^n$.

Bilde die $m \times n$-Matrix A deren *Zeilen* (!) die Vektoren $v_1, v_2, \dots, v_m$ sind. Wende den Gauß-Algorithmus in seiner ersten Version an, um A mittels elementarer Zeilenumformungen in eine Matrix B in Stufenform zu bringen. Die Zeilen $w_1, w_2, \dots, w_r$ von B, die nicht identisch Null sind, sind zufolge Beispiel 4.5.10(4) linear unabhängig und erzeugen nach 4.5.16(2) denselben Untervektorraum wie $v_1, v_2, \dots, v_m$. Also bilden $w_1, w_2, \dots, w_r$ eine Basis von $\operatorname{span}(v_1, v_2, \dots, v_m)$.

Beispiel 4.5.18. Betrachte in $\mathbb{R}^4$ den Untervektorraum U, der von

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \\ -1 \end{pmatrix} \qquad v_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \qquad v_4 = \begin{pmatrix} 2 \\ 6 \\ 2 \\ -1 \end{pmatrix} \qquad v_5 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ -3 \end{pmatrix}$$

aufgespannt wird. In Tafel 4.2 wird der in 4.5.17 beschriebene Algorithmus auf dieses Erzeugendensystem angewandt und liefert die folgende Basis für U:

$$w_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \qquad w_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -3 \end{pmatrix} \qquad w_3 = \begin{pmatrix} 0 \\ 0 \\ 4 \\ -9 \end{pmatrix}.$$

Wir haben noch mehr getan: Gleichzeitig mit der Umformung der Matrix A haben wir rechts daneben auch die 5×5-Einheitsmatrix entsprechend umgeformt und erhalten daraus eine gewisse 5×5-Matrix C. Von den Zeilen von C können wir ablesen, wie die w_1, w_2, w_3 von den v_j abhängen:

$$w_1 = v_1, \quad w_2 = -2v_1 + v_2, \quad w_3 = -5v_1 + 3v_2 + v_3.$$

$$
\begin{array}{cccc|c}
a_{11} & a_{12} & \dots & a_{1n} & v_1 \\
a_{21} & a_{22} & \dots & a_{2n} & v_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
a_{m1} & a_{m2} & \dots & a_{mn} & v_m
\end{array}
$$

$$
\Big\downarrow \text{ Gauß Algorithmus}
$$

$$
\begin{array}{ccccccccccc|c}
0 & \dots & 0 & b_{1j_1} & \dots & b_{1j_2} & \dots & b_{1j_r} & \dots & b_{1n} & & w_1 \\
0 & \dots & \dots & \dots & 0 & b_{2j_2} & \dots & b_{2j_r} & \dots & b_{2n} & & w_2 \\
\vdots & & & & & & \ddots & \vdots & & \vdots & & \vdots \\
0 & \dots & \dots & \dots & \dots & 0 & & b_{rj_r} & \dots & b_{rn} & & w_r \\
\hline
0 & \dots & \dots & \dots & \dots & \dots & & 0 & \dots & 0 & & \\
\vdots & & & & & & & \vdots & & \vdots & & \\
0 & \dots & \dots & \dots & \dots & \dots & & 0 & \dots & 0 & &
\end{array}
$$

$$
b_{1j_1} \neq 0, b_{2j_2} \neq 0, \dots, b_{rj_r} \neq 0
$$

Abbildung 4.1: Schema für den Algorithmus 4.5.17

Von den letzten beiden Zeilen von C lesen wir ab, wie v_4 und v_5 von v_1, v_2, v_3 abhängen:

$$-v_1 - v_2 - v_3 + v_4 = 0, \quad \text{d.h. } v_4 = v_1 + v_2 + v_3$$

$$v_1 - v_2 - v_3 + v_5 = 0, \quad \text{d.h. } v_5 = -v_1 + v_2 + v_3.$$

Beachte, dass v_1, v_2, v_3 ebenfalls eine Basis von U bilden (Warum?).

Beispiel 4.5.19. In $(\mathbb{Z}/2\mathbb{Z})^4$ betrachten wir den Untervektorraum U, der von den Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \qquad v_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad v_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

aufgespannt wird. In Tafel 4.3 wird der Algorithmus aus 4.5.17 durchgeführt, und man erhält die folgende Basis für U:

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad w_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \qquad w_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

4.5.2 Koordinaten bezüglich einer Basis

Satz 4.5.20. *Sei $B = \{v_1, v_2, \ldots, v_n\}$ eine Basis von V. Für jedes $v \in V$ gibt es eindeutig bestimmte Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, so dass*

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \ .$$

Beweis. Gegeben $v \in V$, finden wir $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, so dass

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

ist; denn V wird von $v_1, v_2, \ldots, v_n$ erzeugt. Um die Eindeutigkeit der Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n$ zu zeigen, nehmen wir an, es gebe eine zweite Darstellung

$$v = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n$$

mit $\mu_1, \mu_2, \ldots, \mu_n \in \mathbb{K}$. Ziehen wir beide Darstellungen voneinander ab, so ergibt sich:

$$0 = (\lambda_1 - \mu_1)v_1 + (\lambda_2 - \mu_2)v_2 + \cdots + (\lambda_n - \mu_n)v_n$$

Da $v_1, v_2, \ldots, v_n$ linear unabhängig sind, schließen wir aus dem Kriterium 4.5.10, dass $\lambda_1 - \mu_1 = \lambda_2 - \mu_2 = \cdots = \lambda_n - \mu_n = 0$, d.h. $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$, ..., $\lambda_n = \mu_n$. $\square$

1	2	0	1	v_1						1	0	0	0	0
2	3	1	-1		v_2					0	1	0	0	0
-1	1	1	-1			v_3				0	0	1	0	0
2	6	2	-1				v_4			0	0	0	1	0
0	2	2	-3					v_5		0	0	0	0	1
1	2	0	1	v_1						1	0	0	0	0
0	-1	1	-3	$-2v_1$	$+v_2$					-2	1	0	0	0
0	3	1	0	v_1		$+v_3$				1	0	1	0	0
0	2	2	-3	$-2v_1$			$+v_4$			-2	0	0	1	0
0	2	2	-3					v_5		0	0	0	0	1
1	2	0	1	v_1						1	0	0	0	0
0	-1	1	-3	$-2v_1$	$+v_2$					-2	1	0	0	0
0	0	4	-9	$-5v_1$	$+3v_2$	$+v_3$				-5	3	1	0	0
0	0	4	-9	$-6v_1$	$+2v_2$		$+v_4$			-6	2	0	1	0
0	0	4	-9	$-4v_1$	$+2v_2$			$+v_5$		-4	2	0	0	1
1	2	0	1	v_1					w_1	1	0	0	0	0
0	-1	1	-3	$-2v_1$	$+v_2$				w_2	-2	1	0	0	0
0	0	4	-9	$-5v_1$	$+3v_2$	$+v_3$			w_3	-5	3	1	0	0
0	0	0	0	$-v_1$	$-v_2$	$-v_3$	$+v_4$			-1	-1	-1	1	0
0	0	0	0	$+v_1$	$-v_2$	$-v_3$		$+v_5$		1	-1	-1	0	1

Abbildung 4.2: Bestimmung einer Basis des Untervektorraums aus Beispiel 4.5.18

1	0	0	1	v_1						1	0	0	0	0
0	1	1	1		v_2					0	1	0	0	0
1	1	1	1			v_3				0	0	1	0	0
0	0	0	1				v_4			0	0	0	1	0
0	0	0	1					v_5		0	0	0	0	1
1	0	0	1	v_1						1	0	0	0	0
0	1	1	1		v_2					0	1	0	0	0
0	1	1	0	v_1		$+v_3$				1	0	1	0	0
0	0	0	1				v_4			0	0	0	1	0
0	0	0	1					v_5		0	0	0	0	1
1	0	0	1	v_1						1	0	0	0	0
0	1	1	1		v_2					0	1	0	0	0
0	0	0	1	v_1	$+v_2$	$+v_3$				1	1	1	0	0
0	0	0	1				v_4			0	0	0	1	0
0	0	0	1					v_5		0	0	0	0	1
1	0	0	1	v_1					w_1	1	0	0	0	0
0	1	1	1		v_2				w_2	0	1	0	0	0
0	0	0	1	v_1	$+v_2$	$+v_3$			w_3	1	1	1	0	0
0	0	0	0	v_1	$+v_2$	$+v_3$	$+v_4$			1	1	1	1	0
0	0	0	0	v_1	$+v_2$	$+v_3$		$+v_5$		1	1	1	0	1

Abbildung 4.3: Bestimmung einer Basis des Untervektorraums aus Beispiel 4.5.19

Definition 4.5.21. Ist die endliche Menge $\{v_1, \ldots, v_n\} \subseteq V$ eine Basis von V, die aus n Vektoren besteht, und halten wir eine bestimmte Anordnung der $v_1, \ldots, v_n$ fest, so heißt das n-Tupel $B := (v_1, \ldots, v_n)$ eine *angeordnete Basis von* V.

Es ist oft sinnvoll, zwischen Basen von V (die einfach Teilmengen von V sind) und angeordneten Basen zu unterscheiden.

Sei $B = (v_1, \ldots, v_n)$ eine angeordnete Basis von V. Dann heißen die (eindeutig bestimmten) Skalare $x_1, x_2, \ldots, x_n$, so dass

$$v = x_1 v_1 + x_2 v_2 + \ldots x_n v_n$$

die *Koordinaten* von v bezgl. der angeordneten Basis B. Diese Koordinaten stellen wir durch den Spaltenvektor

$$[v]_B := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

dar.

Bemerkung 4.5.22. Nach dem vorhergehenden Satz ist die Abbildung

$$\iota_B \ : \ \mathbb{K}^n \to V, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$$

bijektiv.

Diese Abbildung ι_B ist linear, weil für $\lambda \in \mathbb{K}$ und $x, y \in \mathbb{K}^n$ gilt:

$$\iota_B(x + y) = \sum_{i=1}^n (x_i + y_i) v_i = \sum_{i=1}^n x_i v_i + \sum_{i=1}^n y_i v_i = \iota_B(x) + \iota_B(y)$$

und

$$\iota_B(\lambda x) = \sum_{i=1}^n \lambda x_i v_i = \lambda \sum_{i=1}^n x_i v_i = \lambda \iota_B(x).$$

Da ι_B linear und bijektiv ist, ist sie ein Vektorraumisomorphismus zwischen $\mathbb{K}^n$ und V. Die Umkehrabbildung $\kappa_B \colon V \to \mathbb{K}^n$ ordnet jedem $v \in V$ das Bild $[v]_B$ zu, das aus den Koordinaten von v bzgl. der Basis B besteht.

In diesem Sinne bedeutet die Wahl einer angeordneten Basis eines Vektorraums V nichts anderes, als: „ein lineares Koordinatensystem auf V einzuführen."

Wählen wir eine andere Basis $B' = (w_1, \ldots, w_m)$ anstelle von $B = (v_1, \ldots, v_n)$, so lässt sich jeder dieser neuen Basisvektoren w_j als Linearkombination

$$w_j = \sum_{i=1}^n a_{ij} v_i$$

mit $a_{ij} \in \mathbb{K}$ schreiben. Daher ist

$$v = \sum_{j=1}^{m} x'_j w_j = \sum_{j=1}^{m} \sum_{i=1}^{n} x'_j a_{ij} v_i.$$

Die Koordinaten $x'_1, \ldots, x'_m$ bzgl. B' und die Koordinaten $x_1, \ldots, x_n$ bzgl. B hängen also wie folgt zusammen:

$$x_i = \sum_{j=1}^{m} a_{ij} x'_j.$$

Mithilfe des Matrixproduktes lässt sich dies so schreiben:

$$x = Ax', \quad [v]_B = A[v]_{B'} \quad \text{wobei} \quad A = (a_{ij}) \in M_{n,m}(\mathbb{K}).$$

Aus der vorstehenden Bemerkung erhalten wir insbesondere:

Satz 4.5.23. *Jeder n-dimensionale $\mathbb{K}$-Vektorraum V ist isomorph zu $\mathbb{K}^n$.*

Allerdings gibt es viele verschiedene Isomorphismen zwischen V und $\mathbb{K}^n$: einen für jede Auswahl einer angeordneten Basis von V. Es kann also je nach Argumentationszusammenhang durchaus ungeschickt sein, alle n-dimensionalen V einfach irgendwie mit $\mathbb{K}^n$ zu identifizieren.

Beispiel 4.5.24. Der Vektorraum $\mathcal{P}_n(\mathbb{R})$ aller reellen Polynomfunktionen vom Grad $\leq n$ ist isomorph zu $\mathbb{R}^{n+1}$.

Um einen Isomorphismus wirklich anzugeben, wählen wir eine angeordnete Basis von $\mathcal{P}_n(\mathbb{R})$, etwa wie gehabt: $B = (p_0, p_1, \ldots, p_n)$, wo $p_j(x) = x^j$. Der zugehörige Isomorphismus ι_B sieht dann so aus:

$$\begin{aligned} \iota_B \colon \mathbb{R}^{n+1} &\to \mathcal{P}_n(\mathbb{R}) \\ \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} &\mapsto a_0 + a_1 p_1 + \ldots + a_n p_n \ : \ (x \mapsto \sum_{j=0}^{n} a_j x^j). \end{aligned}$$

Beispiele 4.5.25. (a)Sei $B = (e_1, e_2, \ldots, e_n)$ die übliche Standardbasis von $\mathbb{K}^n$. Für

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$$

gilt

$$\begin{aligned} x &= x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \\ &= x_1 \; e_1 + x_2 \; e_2 + \ldots + x_n \; e_n \end{aligned}$$

Also sind die Koordinaten von x bzgl. der Standardbasis $e_1, e_2, \dots, e_n$ gerade die $x_1, x_2, \dots, x_n$. Anders gesagt: die Abbildung $\varphi_B \colon \mathbb{K}^n \to \mathbb{K}^n$ ist in diesem Fall die Identität $\mathrm{id}_{\mathbb{K}^n}$.

(b) Ein Vektorraum hat i.a. sehr viele verschiedene Basen. Im $\mathbb{R}^2$ kann man z.B. betrachten:

- die Standardbasis $B\colon e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$;
- die Basis $B'\colon v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Für $v = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ sind die Koordinaten bzgl. der Standardbasis B die Skalare 3 und -1. Die Koordinaten von v bzgl. der Basis B' dagegen sind $[v]_{B'} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$; denn

$$v_1 + 2v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix} = v.$$

Allgemeiner seien für einen Vektor $v \in \mathbb{R}^2$

- $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ die Koordinaten bzgl. der Standardbasis;
- $[v]_{B'} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$ die Koordinaten bzgl. der Basis B'.

Welche Beziehung besteht zwischen x_1, x_2 und x'_1, x'_2 ?

$$[v]_{B'} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \qquad \text{bedeutet: } v = x'_1 v_1 + x'_2 v_2$$

$$\text{d.h.} \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x'_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x'_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\text{d.h.} \qquad \left\{ \begin{array}{l} x_1 = x'_1 + x'_2 \\ x_2 = x'_1 - x'_2 \end{array} \right. \quad \text{d.h.} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$$

Wir haben also die Koordinaten x_1, x_2 bzgl. der Standardbasis in Abhängigkeit von den Koordinate x'_1, x'_2 bzgl. der Basis B' ausgedrückt. Löst man diese beiden Gleichungen nach x'_1, x'_2, ergibt sich

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

$$x'_1 = \frac{1}{2}(x_1 + x_2), \quad x'_2 = \frac{1}{2}(x_1 - x_2).$$

4.6 Sätze über Basen; der Dimensionsbegriff

In diesem Abschnitt bezeichnet V sets einen Vektorraum über einem beliebigen Körper $\mathbb{K}$.

Wie wir gesehen haben, hat V im allgemeinen viele verschiedene Basen. Nun kann man beweisen, dass je zwei Basen B, B' von V die gleiche Mächtigkeit haben. Wenn es also eine Basis gibt, die aus endlich vielen Elementen besteht, so werden alle anderen Basen dieselbe endliche Anzahl von Elementen haben. Diese Anzahl ist die entscheidende einem Vekttorraum zugeordnete *Invariante.* Sie wird die *Dimension* von V genannt. Die Invarianz der Anzahl der Elemente einer Basis ist eines der grundlegenden Resultate der Theorie der Vektorräume über einem Körper. Für ihren Beweis ist der folgende Satz ein wichtiger erster Schritt:

Satz 4.6.1. *Jede Menge von $(m+1)$ Vektoren $w_1, w_2, \ldots, w_m, w_{m+1}$, die in einem von m Vektoren $v_1, v_2, \ldots, v_m$ erzeugten Untervektorraum von V liegen, sind linear abhängig.*

Beweis. Durch Induktion über m.

$m = 1$: Sei $w_1, w_2 \in \operatorname{span}(v_1)$. Dann gibt es Skalare λ_1 und λ_2 so dass $w_1 = \lambda_1 v_1$ und $w_2 = \lambda_2 v_1$. Aus $\lambda_2 w_1 - \lambda_1 w_2 = 0$ folgt dann, dass w_1 und w_2 linear abhängig sind.

Nach Induktionsannahme setzen wir jetzt voraus, dass je m Vektoren in einem von $(m-1)$ Vektoren erzeugten Untervektorraum linear abhängig sind, und wollen zeigen, dass dieselbe Aussage mit m statt $m-1$ gilt.

Seien also $w_1, w_2, \ldots, w_m, w_{m+1} \in \operatorname{span}(v_1, v_2, \ldots, v_m)$. Schreiben wir zunächst die w_i als Linearkombinationen der v_j:

$$\begin{aligned}
w_1 &= \alpha_{11}v_1 + \alpha_{12}v_2 + \cdots + \alpha_{1m}v_m \\
w_2 &= \alpha_{21}v_1 + \alpha_{22}v_2 + \cdots + \alpha_{2m}v_m \\
&\vdots \\
w_{m+1} &= \alpha_{m+1,1}v_1 + \alpha_{m+1,2}v_2 + \cdots + \alpha_{m+1,m}v_m.
\end{aligned}$$

Erster Fall: Alle Koeffizienten in der ersten Spalte auf der rechten Seite sind Null, d.h. $\alpha_{11} = \alpha_{21} = \cdots = \alpha_{m+1,1} = 0$. Dann sind die m Vektoren $w_2, \ldots, w_{m+1}$ in dem von den $(m-1)$ Vektoren $v_2, \ldots, v_m$ erzeugten Untervektorraum enthalten, und nach Induktionsvoraussetzung sind $w_2, \ldots, w_{m+1}$ linear abhängig. Also sind $w_1, w_2, \ldots, w_{m+1}$ erst recht linear abhängig.

Zweiter Fall: Mindestens einer der Koeffizienten α_{i1} der ersten Spalte ist von Null verschieden. Nach eventuellem Umnummerieren der w_i's können wir annehmen, dass $\alpha_{11} \neq 0$ ist. Wie im Gauß'schen Algorithmus bilden wir dann

$$\begin{aligned}
w_2' &:= w_2 - \frac{\alpha_{21}}{\alpha_{11}} w_1 \\
&= 0 \cdot v_1 + \big(\alpha_{22} - \frac{\alpha_{21}}{\alpha_{11}} \cdot \alpha_{12}\big) v_2 + \cdots + \big(\alpha_{2m} - \frac{\alpha_{21}}{\alpha_{11}} \cdot \alpha_{1m}\big) v_m \\
&\vdots \\
w_{m+1}' &:= w_{m+1} - \frac{\alpha_{m+1,1}}{\alpha_{11}} w_1 \\
&= 0 \cdot v_1 + \big(\alpha_{m+1,2} - \frac{\alpha_{m+1,1}}{\alpha_{11}} \cdot \alpha_{12}\big) v_2 + \dots \\
&\qquad \cdots + \big(\alpha_{m+1,m} - \frac{\alpha_{m+1,1}}{\alpha_{11}} \cdot \alpha_{1m}\big) v_m.
\end{aligned}$$

So erhalten wir m Vektoren $w_2', \ldots, w_{m+1}'$, die in dem von den $(m-1)$ Vektoren $v_2, \ldots, v_m$ erzeugten Untervektorraum enthalten sind. Nach Induktionsvoraussetzung sind die Vektoren $w_2', \ldots, w_{m+1}'$ linear abhängig. Es gibt also Skalare $\mu_2, \ldots, \mu_{m+1}$, die nicht sämtlich Null sind, so dass

$$\mu_2 w_2' + \mu_3 w_3' + \cdots + \mu_{m+1} w_{m+1}' = 0.$$

Wegen $w_i' = w_i - \frac{\alpha_{i1}}{\alpha_{11}} w_1$, erhalten wir

$$\mu_2\big(w_2 - \frac{\alpha_{21}}{\alpha_{11}} w_1\big) + \cdots + \mu_{m+1}\big(w_{m+1} - \frac{\alpha_{m+1,1}}{\alpha_{11}} w_1\big) = 0,$$

und mithin

$$\left(-\sum_{i=2}^{n+1} \mu_i \frac{\alpha_{i1}}{\alpha_{11}}\right) w_1 + \mu_2 w_2 + \cdots + \mu_{m+1} w_{m+1} = 0.$$

Das ist eine Linearkombination von $w_1, w_2, \ldots, w_{m+1}$, welche Null ist, ohne dass alle Koeffizienten Null sind; denn nicht alle $\mu_2, \ldots, \mu_{m+1}$ sind 0. Also sind $w_1, w_2, \ldots, w_{m+1}$ linear abhängig.

□

Korollar 4.6.2. *Sind $w_1, \ldots, w_m$ linear unabhängig in einem von n Elementen erzeugten Vektorraum V, so ist $m \leq n$.*

Theorem 4.6.3. *Je zwei Basen eines Vektorraums haben die gleiche Mächtigkeit (d.h. die gleiche Anzahl von Elementen, falls sie aus endlich vielen Elementen bestehen).*

Beweis. Wir behandeln nur den Fall, wo V von endlich vielen Elementen erzeugt werden kann.

Seien $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$ zwei Basen von V. Nach Korollar 4.6.2 impliziert die Bedingung $w_1, \ldots, w_n \in \operatorname{span}(v_1, \ldots, v_m)$, dass $n \leq m$. Ebenso erhalten wir aber $m \leq n$, und somit $m = n$. □

Definition 4.6.4. [Dimension] Sei V ein $\mathbb{K}$-Vektorraum. Hat V eine Basis aus n Vektoren $v_1, v_2, \ldots, v_n$, so definiert man die *Dimension* von V durch

$$\dim V = n.$$

Hat V keine endliche Basis, so begnügen wir uns damit, einfach V als *unendlich dimensionalen Vektorraum* zu bezeichnen und

$$\dim V = \infty$$

zu schreiben.

Beispiele 4.6.5. (a) $\dim\{0\} = 0$.
(b) $\dim \mathbb{K}^n = n$.
(c) $\dim \mathcal{P}_n(\mathbb{R}) = n + 1$ (Beispiel 4.5.10).
(d) $\dim \mathcal{P}(\mathbb{R}) = \infty$ (wegen (c)) und $\dim \mathcal{F}(\mathbb{R}, \mathbb{R}) = \infty$.
(e) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$, der $\mathbb{Q}$-Vektorraum $\mathbb{R}$ ist nicht endlichdimensional, da er nicht abzählbar ist.

Lemma 4.6.6. *Seien $B_1, B_2 \subseteq V$ endliche Teilmengen, so dass gilt:*

(1) *B_1 ist linear unabhängig, und*

(2) *$B_1 \cup B_2$ erzeugt V.*

Dann gibt es eine Teilmenge $B_2' \subseteq B_2$, so dass $B := B_1 \cup B_2'$ eine Basis von V ist.

Beweis. Wähle $B_2' \subseteq B_2$ derart, dass $B_1 \cup B_2'$ linear unabhängig und die Anzahl der Elemente von B_2' maximal ist. Dann ist für jedes $b \in B_2 \setminus B_2'$ die Menge $B_1 \cup B_2' \cup \{b\}$ linear abhängig. Schreibe $B_1 = \{v_1, \ldots, v_n\}$ und $B_2' = \{w_1, \ldots, w_k\}$. Daher gibt es also $\lambda, \lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_k \in \mathbb{K}$, nicht alle Null, mit

$$\lambda b + \lambda_1 v_1 + \ldots + \lambda_n v_n + \mu_1 w_1 + \ldots + \mu_k w_k = 0.$$

Da $B_1 \cup B_2'$ linear unabhängig ist, muss $\lambda \neq 0$ sein, also

$$b = -\lambda^{-1}(\lambda_1 v_1 + \ldots + \lambda_n v_n + \mu_1 w_1 + \ldots + \mu_k w_k) \in \operatorname{span}(B_1 \cup B_2').$$

Da $b \in B_2 \setminus B_2'$ beliebig gewählt war, sieht man, dass

$$B_2 \setminus B_2' \subseteq \operatorname{span}(B_1 \cup B_2'),$$

und damit, dass

$$B_1 \cup B_2 \subseteq \operatorname{span}(B_1 \cup B_2').$$

Das bedeutet, dass die Menge $B_1 \cup B_2'$ ganz V erzeugt, und somit eine Basis ist; denn sie war auch linear unabhängig. □

Betrachten wir jetzt einige Spezialfälle dieses wichtigen Lemmas. Zunächst die Situation, wo $B_1 = \emptyset$. Das bedeutet, dass kein v vorkommt. Dann besagt 4.6.6:

Korollar 4.6.7. (Basisauswahlsatz) *Jedes endliche Erzeugendensystem B eines Vektorraums V enthält eine Basis.*

Beweis. Wende das vorangehende Lemma mit $B_1 = \emptyset$ und $B_2 := B$ an. □

Bemerkung 4.6.8. Dieses Korollar besagt insbesondere, dass *in jedem Vektorraum, der von endlich vielen Elementen erzeugt werden kann, eine Basis existiert.* In der Tat gilt ganz allgemein der

SATZ: *Jeder Vektorraum* (auch ein solcher, der nicht endlich erzeugt ist) *besitzt eine Basis.*

. . . , zumindest gilt dieser Satz dann, wenn man mit einer hinreichend starken Mengenlehre an die Sache herangeht, in der z.B. das sogenannte *Auswahlaxiom* gilt. Darüber will ich mich aber hier jetzt nicht verbreiten.

Korollar 4.6.9. *Ist der Vektorraum V von endlich vielen Vektoren erzeugt, so ist V endlich dimensional, und es ist $\dim V \leq |B|$, für jedes Erzeugendensystem B von V.*

Korollar 4.6.10. (Steinitzscher Austauschsatz) *Sei B eine endliche Basis von V und C eine endliche linear unabhängige Teilmenge von V. Dann gibt es eine Teilmenge $B' \subseteq B$, so dass $B' \cup C$ eine Basis von V ist. (Diese neue Basis von V entsteht also durch Austausch von $B \setminus B'$ durch C.)*

Beweis. Wende Lemma 4.6.6 auf $B_1 := C$ und $B_2 := B$ an. □

Korollar 4.6.11. (Basisergänzungssatz) *Sind $v_1, \ldots, v_m$ linear unabhängig im Vektorraum V der Dimension n, so gibt es Vektoren $v_{m+1}, \ldots, v_n$, so dass $v_1, \ldots, v_m, v_{m+1}, \ldots, v_n$ eine Basis von V ist.*

Beweis. Nach Korollar 4.6.10, angewendet auf eine beliebige Basis $B = \{w_1, \ldots, w_n\}$ und auf $C = \{v_1, \ldots, v_m\}$, wählen wir $n - m$ Vektoren in B so aus, dass sie zusammen mit $v_1, \ldots, v_m$ eine Basis von V bilden. □

Satz 4.6.12. *Für ein System $v_1, v_2, \ldots, v_n$ von n Vektoren eines Vektorraum V der Dimension n sind folgende Eigenschaften äquivalent:*

(1) *$v_1, \ldots, v_n$ sind linear unabhängig.*

(2) *$v_1, \ldots, v_n$ erzeugen V.*

(3) *$v_1, \ldots, v_n$ bilden eine Basis von V.*

Beweis. Nach Definition impliziert (3) die Bedingungen (1) und (2).

(1) $\Rightarrow$ (3): Aus Korollar 4.6.11 folgt, dass sich die Vektoren zu einer Basis ergänzen lassen. Da aber jede Basis von V genau n Elemente enthält, ist keine Ergänzung notwendig und die Vektoren sind bereits eine Basis.

(2) $\Rightarrow$ (3): Wegen Korollar 4.6.7 enthalten $v_1, \ldots, v_n$ eine Basis. Da aber jede Basis von V aus n Elementen besteht, sind $v_1, \ldots, v_n$ bereits eine Basis. □

Wir betrachten jetzt die Dimensionen von Untervektorräumen. Zunächst gilt:

Satz 4.6.13. *Seit V endlich dimensional. Dann gilt* $\dim U \leq \dim V$ *für jeden Untervektorraum U von V. Gleichheit gilt genau dann, wenn $U = V$.*

Beweis. Sei $\dim V = n$. Für jedes System $u_1, \ldots, u_m$ linear unabhängiger Vektoren gilt $m \leq n$ nach Satz 4.6.1. Daher existiert auch eine maximale endliche linear unabhängige Teilmenge $B \subseteq U$. Nach der Folgerung aus Korollar 4.6.11 ist B eine Basis von U. Also ist auch U endlich dimensional und $\dim U \leq \dim V$. Falls $\dim U = \dim V$ gilt, so hat U eine Basis $u_1, \ldots, u_n$ aus n Elementen. Da diese n Vektoren linear unabhängig in V sind, bilden sie auch eine Basis von V nach Satz 4.6.12, also ist $U = V$. □

Satz 4.6.14. *Seien U_1 und U_2 Untervektorräume von V. Dann gilt:*

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2.$$

Beweis. Ist $\dim U_1 = \infty$ oder $\dim U_2 = \infty$, so ist auch $\dim(U_1 + U_2) = \infty$ nach Satz 4.6.13, und die Gleichung ist trivialerweise erfüllt. Nehmen wir also an, dass $\dim U_1 = m_1 < \infty$ und $\dim U_2 = m_2 < \infty$ gilt. Dann ist auch $U_1 \cap U_2$ endlich dimensional (Satz 4.6.13). Sei B_0 eine Basis von $U_1 \cap U_2$. Nach dem Basisergänzungssatz finden wir eine Basis B_1 von U_1, die B_0 enthält, und eine Basis B_2 von U_2, die B_0 enthält. Dann erzeugt $B := B_1 \cup (B_2 \setminus B_0)$ den Untervektorraum $U_1 + U_2$. Wir behaupten, dass B linear unabhängig ist. Schreibe also

$$B_0 = \{v_1, \ldots, v_k\}, \quad B_1 = \{v_1, \ldots, v_n\}, \quad B_2 = \{v_1, \ldots, v_k, w_1, \ldots, w_m\}.$$

Ist B nicht linear unabhängig, so gibt es $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m \in \mathbb{K}$, nicht alle Null, so dass

$$\lambda_1 v_1 + \ldots + \lambda_n v_n + \mu_1 w_1 + \ldots + \mu_m w_m = 0.$$

Dann ist

$$\mu_1 w_1 + \ldots + \mu_m w_m = -\lambda_1 v_1 - \ldots - \lambda_n v_n \in U_1 \cap U_2 = \operatorname{span}\{v_1, \ldots, v_k\}.$$

Die Tatsache, dass B_1 eine Basis von U_1 ist, hat

$$\lambda_{k+1} = \ldots = \lambda_n = 0$$

zur Folge. Nun ist aber B_2 eine Basis von U_2; also folgt $\mu_1 = \ldots = \mu_m = 0 = \lambda_1 = \ldots = \lambda_k$.

Da $B_1 \cup (B_2 \setminus B_0)$ eine Basis von $U_1 + U_2$ ist, gilt

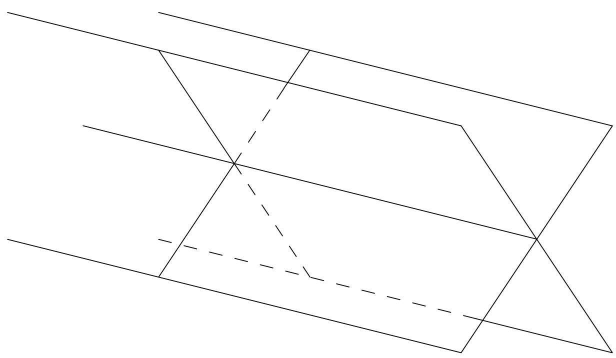
$$\dim(U_1+U_2)+\dim(U_1\cap U_2)-\dim U_1-\dim U_2 = n+m+k-n-(k+m) = 0. \quad \square$$

Bemerkung 4.6.15. (a) Ist $U \subseteq V$ ein eindimensionaler Unterraum, so existiert ein Vektor $0 \neq u \in U$ und dann ist $U = \mathbb{K}u$, da $\{u\}$ bereits eine Basis ist.

(b) Ist $U \subseteq V$ ein zweidimensionaler Unterraum, so existiert ein Vektor $0 \neq u \in U$ und wegen $\mathbb{K}u \neq U$ ein zweiter Vektor $v \in U \setminus \mathbb{K}u$. Dann ist $\{u, v\}$ linear unabhängig, also eine Basis von U. Insbesondere ist $U = \mathbb{K}u + \mathbb{K}v$ eine Ebene.

(c) In einem 3-dimensionalen Vektorraum V erfüllen zwei beliebige voneinander verschiedene Untervektorräume U_1 und U_2 der Dimension 2 stets: $U_1 + U_2 = V$. Denn jede Basis von U_1, zusammen mit einem Element aus $U_2 \setminus U_1$, ist eine linear unabhängige Menge, die drei Elemente enthält und somit eine Basis von V (Satz 4.6.12). Folglich schneiden sich U_1 und U_2 in einem Untervektorraum der Dimension 1; denn

$$\dim(U_1 \cap U_2) = \dim U_1 + \dim U_2 - \dim(U_1 + U_2) = 2 + 2 - 3 = 1$$



(d) In jedem 4-dimensionalen Vektorraum V schneiden sich je zwei verschiedene Untervektorräume U_1 und U_2 der Dimension 3 in einem Untervektorraum der Dimension 2. Wie oben ergibt sich dann $U_1 + U_2 = V$, und daher

$$\dim(U_1 \cap U_2) = \dim U_1 + \dim U_2 - \dim(U_1 + U_2) = 3 + 3 - 4 = 2$$

(e) Ist $V = U_1 \oplus U_2$, so hat man $\dim V = \dim U_1 + \dim U_2$.

In der Tat folgt aus $V = U_1 \oplus U_2$, dass $V = U_1 + U_2$ und $U_1 \cap U_2 = \{0\}$, also

$$\dim(V) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2,$$

da $\dim(U_1 \cap U_2) = \dim(\{0\}) = 0$.

Kapitel 5

Lineare Abbildungen und Matrizen

Die Lineare Algebra hat zwei Gesichter: ein konkretes und ein abstraktes. Im Kapitel 2 haben wir beim Studium der linearen Gleichungssysteme mit reellen Koeffizienten und aller ihrer Lösungen mithilfe des Gauß-Jordanschen Eliminationsalgorithmus die konkrete Seite kennengelernt. Im Kapitel 4 hingegen wurde es dann abstrakt, als wir Vektorräume über einem beliebigen Körper $\mathbb{K}$ und lineare Abbildungen zwischen ihnen eingeführt und zu studieren begonnen haben.

Die Abstraktion wird aber nicht um ihrer selbst willen getrieben. Vielmehr liegt in der Kombination konkreter und abstrakter Aspekte die spezifische Stärke und universelle Anwendbarkeit der Linearen Algebra begründet. Die Vermittlung der beiden Aspekte erfordert eine gewisse Übersetzungsarbeit, die häufig damit beginnt, dass Ergebnisse der abstrakten Theorie der Vektorräume auf den Spezialfall von „konkreten" Vektorräumen wie $\mathbb{K}^n$ oder $M_{n,m}(\mathbb{K})$ angewendet wird: schon im letzten Kapitel haben wir gesehen, wie die Auswahl einer (angeordneten) Basis B in einem beliebigen (endlich dimensionalen) $\mathbb{K}$-Vektorraum V einem beliebigen Element $v \in V$ seine Koordinaten $[v]_B$ bzgl. B zuordnet; d.h. die Auswahl einer Basis leistet eine Identifikation des („abstrakten") Vektorraums V mit dem „konkreten" $\mathbb{K}^n$ (für $n = |B|$).

In diesem Kapitel setzen wir diesen Gedankengang fort, indem wir auch lineare Abblidungen in ähnlicher Weise „koordinatisieren". Genauer werden wir lineare Abbildungen zwischen endlich dimensionalen $\mathbb{K}$-Vektorräumen (nach Auswahl von Basen) durch Matrizen beschreiben. Damit wird dann auch die Interpretation dessen, was im Gauß–Jordan Algorithmus passiert, im Rahmen der Theorie der Vektorräume, d.h. die volle Ausnutzung der abstrakten Theorie für die kalkulatorischen Aspekte der Linearen Algebra, möglich.

5.1 Lineare Abbildungen

Wir erinnern daran (Kapitel 4), dass eine lineare Abbildung $\varphi\colon V \to W$ zwischen $\mathbb{K}$-Vektorräumen V und W eine Abbildung $V \to W$ mit folgenden Eigenschaften ist:

(L1) $\varphi(u+v) = \varphi(u) + \varphi(v)$ für alle $u, v \in V$.

(L2) $\varphi(\lambda v) = \lambda\varphi(v)$ für alle $v \in V,\ \lambda \in \mathbb{K}$.

Die Eigenschaften (L1) und (L2) können auch in eine einzige zusammengefasst werden:

(L) $\varphi(\lambda u + \mu v) = \lambda\varphi(u) + \mu\varphi(v)$ für alle $u, v \in V,\ \lambda, \mu \in \mathbb{K}$.

Definition 5.1.1. [Bild und Kern] Für jede lineare Abbildung $\varphi\colon V \to W$ definieren wir ihr *Bild* durch

$$\operatorname{im}\varphi := \varphi(V) = \{\varphi(v) \mid v \in V\} \subseteq W$$

und ihren *Kern* durch

$$\ker\varphi := \varphi^{-1}(0) = \{v \in V \mid \varphi(v) = 0\} \subseteq V$$

Nach Satz 4.3.18 sind $\operatorname{im}\varphi$ und $\ker\varphi$ Untervektorräume von W bzw. V.

Beispiel 5.1.2. Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, \ldots, v_n \in V$. Betrachte die lineare Abbildung

$$\varphi \ :\ \mathbb{K}^n \to V, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^{n} x_i v_i.$$

Dann ist

$$\operatorname{im}(\varphi) = \operatorname{span}(v_1, \ldots, v_n)$$

und

$$\ker(\varphi) = \{x \in \mathbb{K}^n \colon x_1 v_1 + \ldots + x_n v_n = 0\}.$$

Insbesondere ist φ genau dann injektiv, wenn $v_1, \ldots, v_n$ linear unabhängig sind.

Satz 5.1.3. (Injektivitätskriterium für lineare Abbildungen) *Eine lineare Abbildung $\varphi\colon V \to W$ ist genau dann injektiv, wenn $\ker\varphi = \{0\}$ ist.*

Beweis. Der Kern des Beweises besteht in der folgenden Äquivalenz:

$$(*) \qquad \varphi(v) = \varphi(w) \quad \Longleftrightarrow \quad v - w \in \ker\varphi.$$

Denn wegen der Linearität von φ ist die Gleichheit $\varphi(v) = \varphi(w)$ äquivalent zu

$$0 = \varphi(v) - \varphi(w) = \varphi(v - w),$$

was wiederum bedeutet, dass $v - w \in \ker \varphi$ ist.

Damit können wir die im Satz behauptete Äquivalenz ableiten. Ist φ injektiv, so enthält $\ker \varphi = \varphi^{-1}(0)$ höchstens ein Element, stimmt also mit $\{0\}$ überein. Ist umgekehrt $\ker \varphi = \{0\}$, so zeigt obige Äquivalenz, dass $\varphi(v) = \varphi(w)$ zu $v = w$ äquivalent ist, also ist φ injektiv.

Den folgenden Satz sprechen wir auch für unendlich dimensionale Vektorräume aus. Dafür setzen wir

$$n + \infty := \infty + n := \infty \quad \text{für} \quad n \in \mathbb{N}_0 \cup \{\infty\}.$$

Satz 5.1.4. (Dimensionssatz) *Für jede lineare Abbildung $\varphi\colon V \to W$ gilt*

$$\dim(\operatorname{im} \varphi) + \dim(\ker \varphi) = \dim V.$$

Beweis. Ist $\dim V = \infty$, so müssen wir zeigen dass $\operatorname{im} \varphi$ oder $\ker \varphi$ unendlich dimensional ist. Per Kontraposition genügt es also zu zeigen, dass, wenn sowohl $\operatorname{im} \varphi$ als auch $\ker \varphi$ endlich dimensional sind, die Summe ihrer Dimensionen gleich $\dim V$ ist.

Nach Satz 4.3.18 ist $\ker \varphi$ ein Untervektorraum von V. Sei $k := \dim(\ker \varphi)$ und $v_1, \ldots, v_k$ eine Basis von $\ker \varphi$. Sei weiter $m := \dim(\operatorname{im} \varphi)$ und $w_1, \ldots, w_m$ eine Basis von $\operatorname{im} \varphi$. Wähle Vektoren $v_{k+1}, \ldots, v_{k+m} \in V$, so dass $\varphi(v_{k+j}) = w_j$ gilt für $j = 1, \ldots, m$. Wir behaupten, dass

$$B := \{v_1, \ldots, v_{k+m}\}$$

eine Basis von V ist. Haben wir das gezeigt, ist der Beweis beendet; denn dann ergibt sich

$$\dim V = k + m = \dim(\ker \varphi) + \dim(\operatorname{im} \varphi).$$

Zunächst zeigen wir, dass B den Vektorraum V erzeugt: Sei $v \in V$ beliebig. Dann ist $\varphi(v) \in \operatorname{im}(\varphi)$ und es gibt $\mu_1, \ldots, \mu_m \in \mathbb{K}$ mit

$$\begin{aligned}\varphi(v) &= \mu_1 w_1 + \ldots + \mu_m w_m = \mu_1 \varphi(v_{k+1}) + \ldots + \mu_m \varphi(v_{k+m}) \\ &= \varphi(\mu_1 v_{k+1} + \ldots + \mu_m v_{k+m}).\end{aligned}$$

Folglich

$$v - \mu_1 v_{k+1} + \ldots + \mu_m v_{k+m} \in \ker \varphi = \operatorname{span}(v_1, \ldots, v_k),$$

zusammen also

$$v \in \operatorname{span}(v_1, \ldots, v_{k+m}) = \operatorname{span}(B).$$

Also erzeugt B ganz V.

Wir zeigen jetzt, dass B linear unabhängig ist. Seien dazu $\lambda_1, \ldots, \lambda_{k+m} \in \mathbb{K}$ gegeben mit

$$\sum_{j=1}^{k+m} \lambda_j v_j = 0.$$

Dann ist

$$0 = \varphi(\sum_{j=1}^{k+m} \lambda_j v_j) = \sum_{j=1}^{k+m} \lambda_j \varphi(v_j) = \lambda_{k+1} w_1 + \ldots + \lambda_{k+m} w_m.$$

Aus der linearen Unabhängigkeit der $w_1, \ldots, w_m$ folgt

$$\lambda_{k+1} = \cdots = \lambda_{k+m} = 0.$$

Damit ist $\sum_{j=1}^{k} \lambda_j v_j = 0$, was aber wegen der linearen Unabhängigkeit der $v_1, \ldots, v_k$ wiederum $\lambda_1 = \cdots = \lambda_k = 0$ zur Folge hat.

Wir haben also bewiesen, dass B ein linear unabhängiges Erzeugendensystem ist, also eine Basis von V ist. □

Als Folgerung aus dem Dimensionssatz ergibt sich:

Satz 5.1.5. *Für eine lineare Abbildung $\varphi\colon V \to W$ zwischen endlich dimensionalen Vektorräumen V und W gleicher Dimension sind folgende Eigenschaften äquivalent:*

(1) *φ ist injektiv*

(2) *φ ist surjektiv*

(3) *φ ist bijektiv*

Beweis. Sei $n = \dim V = \dim W$. Der Dimensionssatz 5.1.4 besagt

$$\dim(\operatorname{im}\varphi) + \dim(\ker\varphi) = n.$$

Daher

$$\begin{aligned}
\varphi \text{ injektiv} &\iff \ker\varphi = \{0\} && \text{nach 5.1.3}\\
&\iff \dim(\ker\varphi) = 0 &&\\
&\iff \dim(\operatorname{im}\varphi) = n && \text{nach 5.1.4}\\
&\iff \operatorname{im}\varphi = W && \text{nach 4.6.12}\\
&\iff \varphi \text{ surjektiv} &&
\end{aligned}$$

Daraus folgt sofort die Äquivalenz der drei Eigenschaften. □

Der obige Satz ist eine Variation des Themas, das wir auch schon in Satz 1.3.21 kennengelernt haben. Dort haben wir gesehen dass eine Abbildung $\varphi\colon X \to Y$ zwischen endlichen Mengen mit gleich vielen Elementen genau dann bijektiv ist, wenn sie injektiv oder surjektiv ist.

Bemerkung 5.1.6. Für einen unendlich dimensionalen Vektorraum V gibt es stets injektive lineare Abbildungen $\varphi\colon V \to V$, die nicht surjektiv sind, und surjektive lineare Abbildungen $V \to V$, die nicht injektiv sind. Ist zum

Beispiel $V = \mathcal{P}(\mathbb{R})$ der Vektorraum aller reeller Polynomfunktionen und ist $\varphi\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ die Abbildung

$$p(x) \mapsto x \cdot p(x),$$

so ist diese Abbildung linear und injektiv, aber nicht surjektiv. Denn jede Polynomfunktion q in $\operatorname{im}\varphi$ erfüllt $q(0) = 0$, so dass etwa $1 \notin \operatorname{im}\varphi$.

Definition 5.1.7. [Rang einer linearen Abbildung $\varphi\colon V \to W$] Der *Rang* einer linearen Abbildung wird als die Dimension des Bildes definiert:

$$\operatorname{rank}\varphi := \dim(\operatorname{im}\varphi) = \dim(\varphi(V)).$$

Ist $V = W$ endlich dimensional, so gilt nach dem Dimensionssatz:

$$\operatorname{rank}\varphi = \dim V - \dim(\ker\varphi).$$

Im nächsten Satz untersuchen wir, wie die Eigenschaften einer linearen Abbildung in den Bildern der Vektoren einer Basis reflektiert werden.

Satz 5.1.8. *Sei $v_1, \dots, v_n$ eine Basis von V und $\varphi : V \to W$ eine lineare Abbildung. Dann gilt:*

(1) *Die Vektoren $\varphi(v_1), \dots, \varphi(v_n)$ erzeugen* $\operatorname{im}\varphi$.

(2) *Der Rang von φ ist gleich der größten Anzahl linear unabhängiger Vektoren unter den $\varphi(v_1), \dots, \varphi(v_n)$.*

(3) *φ ist surjektiv $\iff \operatorname{rank}\varphi = \dim W < \infty$.*

(4) *φ ist injektiv $\iff \varphi(v_1), \dots, \varphi(v_n)$ sind linear unabhängig.*

(5) *φ ist bijektiv $\iff \varphi(v_1), \dots, \varphi(v_n)$ bilden eine Basis von W.*

Beweis. (1) Sei $w \in \operatorname{im}\varphi = \varphi(V)$. Dann gibt es ein $v \in V$ so dass $w = \varphi(v)$. Da $v_1, \dots, v_n$ eine Basis von V ist, kann der Vektor v als Linearkombination

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad \lambda_i \in \mathbb{K}$$

dargestellt werden. Da φ linear ist, folgt

$$w = \varphi(v) = \lambda_1\varphi(v_1) + \dots + \lambda_n\varphi(v_n).$$

(2) $\operatorname{rank}\varphi = \dim(\operatorname{im}\varphi)$ ist die Anzahl der Elemente in einer Basis von $\operatorname{im}\varphi$. Nach dem Basisauswahlsatz 4.6.7 ist jedes maximal linear unabhängige Teilsystem eines Erzeugendensystems eine Basis von $\operatorname{im}\varphi$.

(3) Da jeder echte Unterraum von W echt kleiner Dimension hat, ist $\operatorname{rank}\varphi = \dim(\operatorname{im}(\varphi)) = \dim W$ gleichbedeutend zu $\operatorname{im}(\varphi) = W$, also zur Surjektivität von φ.

(4) Nach 5.1.3 ist φ genau dann injektiv, wenn $\ker\varphi = \varphi^{-1}(0)$ nur aus der Null besteht. Da die v_i eine Basis von V bilden, besteht $\varphi^{-1}(0)$ aus allen Linearkombinationen $\sum_i \lambda_i v_i$ mit $0 = \varphi(\sum_i \lambda_i v_i) = \sum_i \lambda_i \varphi(v_i)$. Dass diese Gleichung nur die eine Lösung $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$ besitzt ist aber gleichbedeutend mit der linearen Unabhängigkeit von $\varphi(v_1), \dots, \varphi(v_n)$ (Satz 4.5.8).

(5) Folgt sofort aus (1) und (4). $\square$

Satz 5.1.9. *Seien V, W Vektorräume, $v_1, \ldots, v_n$ eine Basis von V und $w_1, \ldots, w_n$ beliebige Vektoren in W. Dann gibt es genau eine lineare Abbildung $\varphi\colon V \to W$, so dass*

$$\varphi(v_1) = w_1,\ \varphi(v_2) = w_2,\ \ldots, \varphi(v_n) = w_n.$$

Beweis. **Existenz:** Jedes $v \in V$ hat eine Darstellung als Linearkombination

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n,$$

deren Koeffizienten (Koordinaten) $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ durch v eindeutig bestimmt sind. Also liefert die Festsetzung

$$\varphi(v) := \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n$$

eine wohldefinierte Abbildung von V nach W. Für sie gilt $\varphi(v_j) = w_j$ für $j = 1, \ldots, n$; denn

$$v_j = 0 \cdot v_1 + \cdots + 0 \cdot v_{j-1} + 1 \cdot v_j + 0 \cdot v_{j+1} + \cdots + 0 \cdot v_n.$$

Die Linearität von φ folgt leicht aus der Definition. Also gibt es eine lineare Abbildung $\varphi\colon V \to W$ mit der gewünschten Eigenschaft.

Um die **Eindeutigkeit** von φ nachzuweisen, nehmen wir an, $\psi : V \to W$ sei eine andere lineare Abbildung mit $\psi(v_j) = w_j$ für alle $j = 1, \ldots, n$. Für jedes

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n \in V$$

gilt dann

$$\psi(v) = \lambda_1 \psi(v_1) + \cdots + \lambda_n \psi(v_n) = \lambda_1 w_1 + \cdots + \lambda_n w_n = \varphi(v),$$

also $\psi = \varphi$. □

Übung 5.1.10. Beweise die folgende Verallgemeinerung von Satz 5.1.9: Sei B eine Basis des Vektorraums V; sei W ein Vektorraum und sei $f\colon B \to W$ eine Abbildung zwischen den Mengen B und W. Dann gibt es genau eine *lineare* Abbildung $\varphi\colon V \to W$ mit $\varphi|_B = f$.

5.2 Quotientenräume und kanonische Faktorisierung

Wir haben schon in Satz 4.3.18 gesehen, dass der Kern einer linearen Abbildung $\varphi\colon V \to W$ immer ein Untervektorraum von V ist. Man kann sich nun die Frage stellen, ob in der Tat auch jeder Untervektorraum ein Kern ist. Das führt auf die Konstruktion des sogenannten Quotientenraums.

Definition 5.2.1. Sei V ein Vektorraum und $U \subseteq V$ ein Untervektorraum. Wir betrachten die Menge

$$V/U := \{v + U \colon v \in V\}$$

aller affiner Unterräume der Gestalt $v + U$.

Auf dieser Menge definieren wir eine Addition und eine Skalarmultiplikation durch

$$(v + U) + (w + U) := v + w + U \quad \text{und} \quad \lambda(v + U) := \lambda v + U.$$

Um zu zeigen, dass dies sinnvolle Definitionen sind, haben wir zu verifizieren, dass die rechte Seite jeweils nicht vom Repräsentanten v bzw. w des affinen Unterraums abhängen: Ist $v + U = v' + U$ und $w + U = w' + U$, so sind $v - v', w - w' \in U$ und daher

$$v' + w' + U = (v + w) + (v' - v) + (w' - w) + U = v + w + U,$$

da $x + U = U$ für jedes Element $x \in U$ gilt. Analog sieht man

$$\lambda v' + U = \lambda v + \lambda(v' - v) + U = \lambda v + U$$

ein. Wie wir gleich sehen werden, erhalten wir so eine Vektorraumstruktur auf V/U bzgl. der die Abbildung

$$q \colon V \to V/U, \quad v \mapsto v + U$$

linear und surjektiv wird. Wir nennen V/U den *Quotientenvektorraum* bzw. den *Quotienten von V nach U*.

Satz 5.2.2. *V/U ist mit der oben beschriebenen Addition und Skalarmultiplikation ein Vektorraum und $q \colon V \to V/U, v \mapsto v + U$ ist eine surjektive lineare Abbildung mit* $\ker q = U$.

Beweis. Der Nachweis der Vektorraumaxiome für V/U ist trivial, wenn man beachtet, dass die Abbildung q surjektiv ist und den Bedingungen

$$q(v + w) = q(v) + q(w) \quad \text{und} \quad q(\lambda v) = \lambda q(v), \quad v, w \in V, \lambda \in \mathbb{K}, \tag{5.1}$$

genügt. Das ist aber in der Definition der Addition und der Skalarmultiplikation enthalten.

Z.B. verifiziert man die Assoziativität der Addition wie folgt. Für $x, y, z \in V/U$ existieren $a, b, c \in V$ mit $x = q(a)$, $y = q(b)$ und $z = q(c)$. Dann ist

$$\begin{aligned}(x + y) + z &= q(a + b) + q(c) = q((a + b) + c) = q(a + (b + c)) = q(a) + q(b + c) \\ &= x + (y + z).\end{aligned}$$

Analog verifiziert man (V2)-(V8) durch Zurückführung auf die entsprechenden Eigenschaften von V.

Ist V/U einmal als Vektorraum erkannt, so folgt die Linearität von q aus (5.1). $\square$

Beispiel 5.2.3. Sei $V = \mathbb{K}^3$ und $U := \mathbb{K}\begin{pmatrix}1\\1\\1\end{pmatrix}$. Wenden wir den Dimensionssatz auf die Quotientenabbildung $q\colon V \to V/U$ an, so erhalten wir sofort

$$\dim(V/U) = \dim(V) - \dim(U) = 3 - 1 = 2.$$

Der Raum V/U aller Geraden $x + U$ in $\mathbb{K}^3$ mit der Richtung U ist also zweidimensional.

Wir suchen einen Basis. Zunächst ist klar, dass die Vektoren $b_i := q(e_i)$, $i = 1, 2, 3$, den Raum V/U aufspannen. Sie sind aber nicht linear unabhängig, wenn wegen $e_1 + e_2 + e_3 \in U = \ker q$ ist $b_1 + b_2 + b_3 = 0$, also $b_3 = -b_1 - b_2$. Daher ist die zweielementige Menge $\{b_1, b_2\}$ erzeugend und folglich eine Basis.

Satz 5.2.4. (Homomorphiesatz) *Sei $\varphi\colon V \to W$ eine lineare Abbildung und $U \subseteq V$ ein Untervektorraum sowie $q\colon V \to V/U$ die Quotientenabbildung. Genau dann existiert eine lineare Abbildung $\widetilde{\varphi}\colon V/U \to W$ mit $\widetilde{\varphi} \circ q = \varphi$, wenn $U \subseteq \ker\varphi$ gilt. In diesem Fall ist die lineare Abbildung $\widetilde{\varphi}$ durch $\widetilde{\varphi} \circ q = \varphi$ eindeutig bestimmt.*

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "q"] & W \arrow[d, "\mathrm{id}_W"] \\ V/U \arrow[r, "\widetilde{\varphi}"] & W \end{tikzcd}$$

Beweis. Existiert $\widetilde{\varphi}$, so ist

$$\ker\varphi = \ker(\widetilde{\varphi} \circ q) \supseteq \ker q = U.$$

Ist umgekehrt die Bedingung $U \subseteq \ker\varphi$ erfüllt, so ist

$$\widetilde{\varphi}\colon V/U \to W, \quad \widetilde{\varphi}(x + U) := \varphi(x)$$

wohldefiniert, denn für $u \in U$ ist

$$\varphi(x + u) = \varphi(x) + \varphi(u) = \varphi(x).$$

Wegen

$$\widetilde{\varphi}\big((x{+}U){+}(y{+}U)\big) = \widetilde{\varphi}\big(x{+}y{+}U\big) = \varphi(x{+}y) = \varphi(x){+}\varphi(y) = \widetilde{\varphi}(x{+}U){+}\widetilde{\varphi}(y{+}U)$$

und

$$\widetilde{\varphi}(\lambda(x + U)) = \widetilde{\varphi}(\lambda x + U) = \varphi(\lambda x) = \lambda\varphi(x) = \lambda\widetilde{\varphi}(x + U)$$

ist $\widetilde{\varphi}$ linear.

Dass $\widetilde{\varphi}$ durch $\widetilde{\varphi} \circ q = \varphi$ eindeutig bestimmt ist, folgt aus der Surjektivität von q. □

Satz 5.2.5. (Kanonische Faktorisierung) *Sei $\varphi\colon V \to W$ eine lineare Abbildung und $q\colon V \to V/\ker\varphi$ die Quotientenabbildung. Dann definiert*

$$\overline{\varphi}\colon V/\ker\varphi \to \operatorname{im}(\varphi), \quad \overline{\varphi}(x + \ker\varphi) := \varphi(x)$$

einen Isomorphismus von Vektorräumen.

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ \big\downarrow q & & \big\uparrow \\ V/\ker(\varphi) & \xrightarrow{\overline{\varphi}} & \mathrm{im}(\varphi) \end{array}$$

Beweis. Aus dem Homomorphiesatz erhalten wir eine lineare Abbildung

$$\overline{\varphi}\colon V/\ker q \to \operatorname{im}\varphi, \quad x + \ker\varphi \mapsto \varphi(x).$$

Es ist klar, dass $\overline{\varphi}$ surjektiv ist. Um einzusehen, dass $\overline{\varphi}$ auch injektiv ist, nehmen wir an, dass $0 = \overline{\varphi}(x + \ker\varphi) = \varphi(x)$ ist. Dann ist $x \in \ker\varphi$, und somit $x + \ker\varphi = \ker\varphi$ das Nullelement in $V/\ker\varphi$. Daher ist $\overline{\varphi}$ injektiv, mithin ein Isomorphismus von Vektorräumen. □

5.3 Operationen auf linearen Abbildungen

In diesem Abschnitt bezeichnen U, V, W Vektorräume über dem selben Körper $\mathbb{K}$.

Definition 5.3.1. [Hom(V, W) als Vektorraum] Wir erinnern daran, dass mit $\mathrm{Hom}(V, W)$ die Menge aller linearer Abbildungen von V nach W bezeichnet wird. Für $\varphi, \psi \in \mathrm{Hom}(V, W)$ und $\lambda \in \mathbb{K}$, definiere Abbildungen $\varphi + \psi$ und $\lambda\varphi$ von V nach W durch

$$\begin{aligned} (\varphi + \psi)(v) &:= \varphi(v) + \psi(v) \\ (\lambda\varphi)(v) &:= \lambda \cdot \varphi(v) \in V \ . \end{aligned}$$

für alle $v \in V$.

Lemma 5.3.2. *$\varphi + \psi$ und $\lambda\varphi$ sind linear, d.h. sie sind selbst Elemente von* $\mathrm{Hom}(V, W)$.

Beweis. Die Abbildung $\varphi + \psi$ ist linear; denn

$$\begin{aligned} (\varphi + \psi)(u + v) &= \varphi(u + v) + \psi(u + v) && \text{nach Definition von } \varphi + \psi \\ &= \varphi(u) + \varphi(v) + \psi(u) + \psi(v) && \text{nach (L1)} \\ &= \varphi(u) + \psi(u) + \varphi(v) + \psi(v) \\ &= (\varphi + \psi)(u) + (\varphi + \psi)(v) && \text{nach Definition von } \varphi + \psi \end{aligned}$$

und

$$\begin{aligned} (\varphi + \psi)(\mu v) &= \varphi(\mu v) + \psi(\mu v) && \text{nach Definition von } \varphi + \psi \\ &= \mu \cdot \varphi(v) + \mu \cdot \psi(v) && \text{nach (L2)} \\ &= \mu \cdot \big(\varphi(v) + \psi(v)\big) \\ &= \mu \cdot (\varphi + \psi)(v) && \text{nach Definition von } \varphi + \psi \end{aligned}$$

Ganz ähnlich verifiziert man die Linearität von $\lambda\varphi$. □

Satz 5.3.3. *Mit der eben definierten Addition und Skalarmultiplikation ist die Menge* $\mathrm{Hom}(V, W)$ *aller linearen Abbildungen* $\varphi\colon V \to W$ *ein* $\mathbb{K}$*-Vektorraum.*

Zum Beweis muss man für die eben definierten Operationen die Vektorraumaxiome (V1)–(V8) nachprüfen. Wir überlassen das dem Leser als Übung.

Satz 5.3.4. (Regeln für die Komposition linearer Abbildungen) *Für je zwei lineare Abbildungen* $\psi, \psi'\colon U \to V$, $\varphi, \varphi'\colon V \to W$ *und alle* $\lambda \in \mathbb{K}$ *gilt:*

$$\text{Die Komposition linearer Abbildungen ist assoziativ.} \tag{A1}$$

$$\mathrm{id}_W \circ \varphi = \varphi = \varphi \circ \mathrm{id}_V \tag{A2}$$

$$\varphi \circ (\psi + \psi') = \varphi \circ \psi + \varphi \circ \psi', \quad (\varphi + \varphi') \circ \psi = \varphi \circ \psi + \varphi' \circ \psi \tag{A3}$$

$$(\lambda\varphi) \circ \psi = \lambda \cdot (\varphi \circ \psi) = \varphi \circ (\lambda\psi) \tag{A4}$$

Beweis. Die Regeln (A1) und (A2) gelten allgemein für beliebige Abbildungen zwischen Mengen (Satz 1.3.17).

(A3): Für alle $u \in U$ gilt

$$\begin{aligned}
\big(\varphi \circ (\psi + \psi')\big)(u) &= \varphi\big((\psi + \psi')(u)\big) && \text{nach Definition von } \circ \\
&= \varphi\big(\psi(u) + \psi'(u)\big) && \text{nach Definition von } \psi + \psi' \\
&= \varphi\big(\psi(u)\big) + \varphi\big(\psi'(u)\big) && \text{wegen der Linearität von } \varphi \\
&= (\varphi \circ \psi)(u) + (\varphi \circ \psi')(u) \\
&= (\varphi \circ \psi + \varphi \circ \psi')(u).
\end{aligned}$$

Mithin $\varphi \circ (\psi + \psi') = \varphi \circ \psi + \varphi \circ \psi'$.

Den zweiten Teil von (A3) und (A4) beweist man ähnlich. □

Definition 5.3.5. Ein Vektorraum $\mathcal{A}$, versehen mit einer binären Operation $(a, b) \mapsto ab$ heißt *Algebra*, wenn diese Operation folgende Eigenschaften hat (für alle $a, b, c \in \mathcal{A}$):

(1) $(ab)c = a(bc)$ (Assoziativität).

(2) $a(b + c) = ab + ac, (a + b)c = ac + bc$ (Distributivität)

(3) $(\lambda a)b = a(\lambda b) = \lambda \cdot ab$ für $\lambda \in \mathbb{K}$.

Existiert zusätzlich ein neutrales Element **1** für die Multiplikation von $\mathcal{A}$, so sprechen wir von einer Algebra mit Eins.

Definition 5.3.6. [$\mathrm{End}(V) = \mathrm{Hom}(V, V)$ als $\mathbb{K}$-Algebra] Setzen wir $V = W$ in Definition 5.3.1, so erhalten wir auf dem der Menge $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ der linearen Abbildungen $\varphi\colon V \to V$ von V in sich selbst zuerst eine $\mathbb{K}$-Vektorraum-Struktur vermöge der dort definierten Addition und Skalarmultiplikation. Da aber die Komposition linearer Abbildungen nach 4.2.4(1) linear ist, erhalten wir noch eine weitere Operation $\circ$ auf $\mathrm{End}(V)$:

$$\begin{aligned}
\mathrm{End}(V) \times \mathrm{End}(V) &\to \mathrm{End}(V) \\
(\varphi, \psi) &\mapsto \varphi \circ \psi,
\end{aligned}$$

die wir als Multiplikation auffassen können; denn nach 5.3.4 hat sie die folgenden Eigenschaften für alle $\varphi, \varphi', \psi, \psi', \eta \in \mathrm{End}(V)$:

$$\varphi \circ (\psi \circ \eta) = (\varphi \circ \psi) \circ \eta \tag{A1}$$

$$\mathrm{id}_V \circ \varphi = \varphi = \varphi \circ \mathrm{id}_V \tag{A2}$$

$$\varphi \circ (\psi + \psi') = \varphi \circ \psi + \varphi \circ \psi', \quad (\varphi + \varphi') \circ \psi = \varphi \circ \psi + \varphi' \circ \psi \tag{A3}$$

$$(\lambda \cdot \varphi) \circ \psi = \lambda \cdot (\varphi \circ \psi) = \varphi \circ (\lambda \psi) \tag{A4}$$

Wegem $\mathrm{id}_V \in \mathrm{End}(V)$ gemäß 4.2.4(2), ersehen wir aus (A1) und (A2), dass $\circ$ eine assoziative binäre Operation auf $\mathrm{End}(V)$ mit id_V als neutralem Element ist. Gemäß (A3) ist $\circ$ distributiv bzgl. der Addition. Also ist $\mathrm{End}(V)$ ein Ring bzgl. $+$ und $\circ$. Und außerdem gilt noch (A4). Alles in allem ist $\mathrm{End}(V)$ sei eine $\mathbb{K}$-Algebra.

Bemerkung 5.3.7. Mit $\mathrm{Aut}(V)$ haben wir die Menge aller Automorphismen des Vektorraums V bezeichnet. Anders gesagt (vgl. die Übungen aus Abschnitt 4.1):

$$\mathrm{Aut}(V) = \mathrm{End}(V)^\times$$

ist die Einheitengruppe des Monoids $(\mathrm{End}(V), \circ, \mathrm{id}_V)$. (Die Gruppenoperation ist die Komposition, das neutrale Element ist id_V.) Diese Gruppe wird auch häufig als $\mathrm{GL}(V)$ (für *general linear group of* V) bezeichnet.

Da die Umkehrabbildung φ^{-1} eines bijektiven Endomorphismus $\varphi \in \mathrm{End}(V)$ automatisch linear ist (nach 4.2.4(3)) ist sie auch ein Automorphismus. Daher ist $\mathrm{Aut}(V)$ die Menge aller bijektiver linearer Endomorphismen von V.

5.4 Matrizen mit Einträgen aus $\mathbb{K}$

Dem Lernenden sollte dieser Abschnitt großenteils entbehrlich vorkommen: Wir stellen in ihm nur fest, dass alles, was wir im Kapitel 2 über Matrizen mit reellen Koeffizienten gesagt haben, und was sich mit dem Begriff des $\mathbb{R}$-Vektorraums etwas stromlinienförmiger ausdrücken lässt, auch gilt, wenn man den Körper $\mathbb{R}$ durch einen beliebigen Körper $\mathbb{K}$ ersetzt. An den Beweisen ändert sich nichts — so ist der Körperbegriff gerade gemacht. Der logischen Vollständigkeit deuten wir hier die einzelnen Tatsachen so an, wie wir sie im Folgenden benutzen werden.

5.4.1. [$M_{m,n}(\mathbb{K})$ als $\mathbb{K}$-Vektorraum] Wie schon im Beispiel 4.1.4 gesehen ist $M_{m,n}(\mathbb{K})$ ein $\mathbb{K}$-Vektorraum vermöge der Operationen

$$A + B = (a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij})$$
$$\lambda A = \lambda(a_{ij}) := (\lambda a_{ij}).$$

Um die Dimension von $M_{m,n}(\mathbb{K})$ zu bestimmen, betrachten wir die Matrizen

$$E_{ij} := \begin{array}{cc} & \begin{array}{ccccc} & & j & & \\ & & \downarrow & & \end{array} \\ i \rightarrow & \begin{pmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{pmatrix} \end{array}$$

die genau einen Eintrag $\neq 0$ haben: die 1 am Schnittpunkt der i-ten Zeile mit der j-ten Spalte. Diese speziellen Matrizen E_{ij}, für $i = 1, \dots, m$ und $j = 1, \dots, n$, bilden eine Basis des $\mathbb{K}$-Vektorraums $M_{m,n}(\mathbb{K})$.

In der Tat kann nämlich jede Matrix $A = (a_{ij})$ in eindeutiger Weise als Linearkombination

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}$$

geschrieben werden, wo die Koeffizienten der Kombination gerade die Einträge der Matrix A sind. Da es mn viele Matrizen E_{ij} gibt, haben wir gezeigt:

$$\dim M_{m,n}(\mathbb{K}) = m \cdot n.$$

Satz 5.4.2. (Regeln für das Matrizenprodukt) *Seien A, A' zwei $m \times n$-Matrizen, und B, B' zwei $n \times p$-Matrizen, C eine $p \times q$-Matrix mit Einträgen aus $\mathbb{K}$, und $\lambda \in \mathbb{K}$. Dann gilt*

$$A(BC) = (AB)C \tag{A1}$$

$$E_m A = A = A E_n \tag{A2}$$

$$(A + A')B = AB + A'B, \quad A(B + B') = AB + AB' \tag{A3}$$

$$(\lambda A)B = \lambda(AB) = A(\lambda B) \tag{A4}$$

Wir kennen all diese Regeln und ihre Beweise von reellen Matrizen!

Diese Regeln sollten mit denen für die Zusammensetzung linearer Abbildungen verglichen werden (5.3.4). Die Ähnlichkleit kommt nicht von ungefähr — wir werden den Zusammenhang zwischen linearen Abbildungen und Matrizen im nächsten Abschnitt systematisch entwickeln.

Vorher aber sei daran erinnert, dass man zwei $n \times n$-Matrizen A und B stets miteinander multiplizieren kann und als Ergebnis $A \cdot B$ wieder eine $n \times n$-Matrix erhält. Auf der Menge $M_n(\mathbb{K})$ aller $n \times n$-Matrizen mit Einträgen aus $\mathbb{K}$ definiert das Matrizenprodukt also eine binäre Operation. Aus dem obigen Satz erhalten wir insbesondere:

Satz 5.4.3. *$M_n(\mathbb{K})$ ist eine $\mathbb{K}$-Algebra mit neutralem Element E_n für die Multiplikation.*

Bemerkung 5.4.4. Es ist lehrreich, das Matrizenprodukt in den Basen E_{ij} von $M_{m,n}(\mathbb{K})$ und $E_{k\ell} \in M_{n,p}(\mathbb{K})$ zu beschreiben. Das Produkt zweier Basiselemente ist:

$$E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell},$$

wobei $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ wieder das Kronecker-Symbol bezeichnet.

Mit der Distributivität des Produktes kann man aus dieser Multiplikationstabelle auf den Basen das Produkt zweier beliebiger Matrizen ausrechnen: Für $A = \sum_{i,j} a_{ij}E_{ij}$ und $B = \sum_{k,\ell} b_{k\ell}E_{k\ell}$ erhalten wir

$$\begin{aligned} AB &= \sum_{i,j} a_{ij}E_{ij} \sum_{k,\ell} b_{k\ell}E_{k\ell} = \sum_{i,j,k,\ell} a_{ij}b_{k\ell}E_{ij}E_{k\ell} \\ &= \sum_{i,j,k,l} a_{ij}b_{k\ell}\delta_{jk}E_{i\ell} = \sum_{i,j,l} a_{ij}b_{j\ell}E_{i\ell} = \sum_{i,l}\Big(\sum_j a_{ij}b_{j\ell}\Big)E_{i\ell}. \end{aligned}$$

Definition 5.4.5. [Die Gruppe $\mathrm{GL}_n(\mathbb{K})$ der invertierbaren Matrizen] Bezüglich der Matrizenmultiplikation ist $M_n(\mathbb{K})$ ein Monoid mit neutralem Element E_n. Seine Einheitengruppe $M_n(\mathbb{K})^\times$ (vgl. Kapitel 3) schreiben wir $\mathrm{GL}_n(\mathbb{K})$ oder $\mathrm{GL}(n,\mathbb{K})$. Sie besteht aus allen invertierbaren Matrizen, d.h. aus allen Matrizen $A \in M_n(\mathbb{K})$, für die eine Matrix $B \in M_n(\mathbb{K})$ existiert, so dass

$$A \cdot B = E_n = B \cdot A.$$

Die inverse Matrix B ist, wie wir wissen, eindeutig durch A bestimmt und man schreibt sie $B = A^{-1}$.

Um in der Praxis zu entscheiden, ob eine gegebene Matrix A invertierbar ist, benutzt man den (erweiterten) Gauß-Jordan Algorithmus eben so, wie das in Kapitel 2 beschrieben wurde.

5.5 Lineare Abbildungen $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ und Matrizen

In diesem Abschnitt beziehen wir uns stets auf die Standardbasis

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \dots, \qquad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

des $\mathbb{K}^n$ (für verschiedene n).

5.5.1. [Von Matrizen zu linearen Abbildungen] Sei $A = (a_{ij})$ eine $m \times n$-Matrix mit Koeffizienten $a_{ij} \in \mathbb{K}$. Jeder Spaltenvektor $x \in \mathbb{K}^n$ kann als $n \times 1$-Matrix aufgefasst werden. Also ist das Matrizenprodukt Ax wohldefiniert: Ax ist eine $m \times 1$-Matrix oder, mit anderen Worten, ein Spaltenvektor $Ax \in \mathbb{K}^m$:

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1+ & \dots & +a_{1n}x_n \\ a_{21}x_1+ & \dots & +a_{2n}x_n \\ & \vdots & \\ a_{m1}x_1+ & \dots & +a_{mn}x_n \end{pmatrix}$$

So haben wir eine lineare Abbildung $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$ durch die Vorschrift $\varphi_A(x) = Ax$ für alle $x \in \mathbb{K}^n$ definiert — wir kennen diese Abbildung schon aus Beispiel 4.2.3(2)).

BEACHTE:

$$\varphi_A(e_j) = Ae_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = j\text{-te Spalte von } A.$$

Anders gesagt, die Bilder der Standardbasisvektoren $e_1, e_2, \dots, e_n$ sind gerade die Spalten $s_1 = Ae_1, s_2 = Ae_2, \dots, s_n = Ae_n$ der Matrix A:

$$A = \begin{array}{c} \begin{array}{cccc} \overbrace{\phantom{a_{11}}}^{s_1} & \overbrace{\phantom{a_{12}}}^{s_2} & \dots & \overbrace{\phantom{a_{1n}}}^{s_n} \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \end{array}$$

5.5.2. [Das Bild von φ_A] Definitionsgemäß ist

$$\operatorname{im} \varphi_A = \varphi_A(\mathbb{K}^n) = \{b \in \mathbb{K}^m : (\exists x \in \mathbb{K}^n)\ Ax = b\}$$

Das besagt, dass das Bild von φ_A aus den Vektoren $b \in \mathbb{K}^m$ besteht, für die das lineare Gleichungssystem

$$Ax = b$$

eine Lösung hat. Wir wissen aus Satz 4.3.18, dass das Bild von φ_A ein Untervektorraum von $\mathbb{K}^m$ ist.

Satz 5.5.3. *Das Bild von φ_A ist der von den Spaltenvektoren $s_1, s_2, \dots s_n$ von A aufgespannte Untervektorraum von $\mathbb{K}^m$.*

Beweis. Nach Satz 5.1.8(1) wissen wir, dass $\operatorname{im} \varphi_A$ von den Bildern $\varphi_A(e_1), \dots, \varphi_A(e_n)$ der Standardbasisvektoren erzeugt wird. Der Satz folgt dann aus der vorangehenden Diskussion. $\square$

Den Rang einer linearen Abbildung φ hatten wir als die Dimension von $\operatorname{im} \varphi$ definiert. Der Rang von φ_A ist also die Dimension des von den Spaltenvektoren $s_1, s_2, \dots, s_n$ von A aufgespannten Untervektorraums.

5.5.4. [Rang einer Matrix $A \in M_{m,n}(\mathbb{K})$] In Analogie zum Fall der linearen Abbildungen definieren wir jetzt den Rang einer Matrix A als die Dimension des von den Spaltenvektoren $s_1, s_2, \ldots, s_n$ von A aufgespannten Untervektorraums:

$$\operatorname{rank} A := \dim(\operatorname{span}(s_1, s_2, \ldots, s_n)).$$

Nach der vorangehenden Bemerkung ist also

$$\operatorname{rank} A = \operatorname{rank} \varphi_A.$$

Wir kennen eine praktische Methode, um eine Basis des Bildes von φ_A und somit auch den eben definierten Rang r einer gegebenen Matrix A zu berechnen: Man wende den Gauß-Jordan Algorithmus auf die Spaltenvektoren $s_1, \ldots, s_n$ von A an (vgl. Abschnitt 4.5), also auf die transponierte Matrix $A^\top$.

Beispiel 5.5.5. Sei

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & -4 \end{pmatrix} \in M_3(\mathbb{K}).$$

Wir wenden den Gauß-Jordan Algorithmus an (siehe die Tabelle in 5.1). Die

1	2	−1	s_1
0	1	2	s_2
2	3	−4	s_3
1	2	−1	
0	1	2	
0	−1	−2	
1	2	−1	
0	1	2	
0	0	0	

Abbildung 5.1: Gauß-Jordan Algorithmus für die Matrix in 5.5.5

Rechnung zeigt, dass $\operatorname{rank} A = 2$ ist; die Vektoren $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$ bilden eine Basis des von den Spalten von A aufgespannten Untervektorraums von $\mathbb{K}^3$. Also ist $\operatorname{rank} \varphi_A = 2$; das Bild von φ_A hat die beiden obigen Vektoren als Basis.

Definition 5.5.6. [Spaltenrang, Zeilenrang, algorithmischer Rang] Sei $A \in M_{m,n}(\mathbb{K})$. In 5.5.5 haben wir eben den Rang von A als die Dimension des von den Spalten $s_1, \ldots, s_n$ von A aufgespannten Untervektorraums von $\mathbb{K}^m$ definiert. Man nennt dies auch den *Spaltenrang* von A.

Andererseits können wir auch den von den Zeilen $z_1, \ldots, z_m \in \mathbb{K}^n$ von A aufgespannten Untervektorraum von $\mathbb{K}^n$ betrachten. Seine Dimension heißt der *Zeilenrang* von A.

Bemerkung 5.5.7. Schließlich haben wir vor langer Zeit in Definition 2.4.1, sozusagen experimentell, durch Anwendung des Gauß-Jordan Algorithmus, den *Rang* eines linearen Gleichungssystems in einer Weise definiert, die sich ohne weiteres auf jede Matrix übertragen lässt: Der *algorithmische Rang* einer Matrix $A \in M_{m,n}(\mathbb{K})$ ist die Anzahl der Zeilen, die nach einer Umformung von A in Stufenform durch den Gauß-Jordan-Algorithmus nicht gänzlich null sind. Oder anders gesagt: Der *algorithmische Rang* von A ist die Anzahl der Pivotelemente.

In Beispiel 4.5.10(4) und Theorem 5.5.9 haben wir festgestellt, dass dieser algorithmische Rang von A nichts anderes als der Zeilenrang von A ist. Der springende Punkt dabei ist, dass der von den Zeilenvektoren erzeugte Untervektorraum von $\mathbb{K}^n$ sich unter elementaren Umformungen nicht ändert. Insbesondere ist damit auch erledigt, was uns in Definition 2.4.1 noch Anlass zu einer Warnung war: dass der algorithmische Rang nicht von der Art und Weise abhängt, wie der Eliminationsalgorithmus durchgeführt wird. Es gibt also schon einmal keine drei wirklich verschiedenen Rangbegriffe für eine Matrix. Es bleibt aber die Frage, wie sich Zeilen- und Spaltenrang im allgemeinen zueinander verhalten.

Beispiel 5.5.8. Sei

$$A = \begin{pmatrix} 1 & 2 & 1 & -1 \\ 2 & 0 & 1 & 3 \\ -1 & 2 & 0 & -4 \end{pmatrix} \in M_{3,4}(\mathbb{R})$$

Wir berechnen den Zeilenrang in dem Schema links, und rechts den Spaltenrang — siehe Tabelle 5.2.

$$\begin{array}{rrrrc} 1 & 2 & 1 & -1 & z_1 \\ 2 & 0 & 1 & 3 & z_2 \\ -1 & 2 & 0 & -4 & z_3 \\ \hline 1 & 2 & 1 & -1 & \\ 0 & -4 & -1 & 5 & \\ 0 & 4 & 1 & -5 & \\ \hline 1 & 2 & 1 & -1 & \\ 0 & -4 & -1 & 5 & \\ 0 & 0 & 0 & 0 & \\ \hline \end{array}
\qquad
\begin{array}{rrrc} 1 & 2 & -1 & s_1 \\ 2 & 0 & 2 & s_2 \\ 1 & 1 & 0 & s_3 \\ -1 & 3 & -4 & s_4 \\ \hline 1 & 2 & -1 & \\ 0 & -4 & 4 & \\ 0 & -1 & 1 & \\ 0 & 5 & -5 & \\ \hline 1 & 2 & -1 & \\ 0 & -4 & 4 & \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \\ \hline \end{array}$$

Zeilenrang von $A = 2$ Spaltenrang von $A = 2$

Abbildung 5.2: Zeilen- und Spaltenrang der Matrix A in 5.5.8

In diesem Beispiel sind Zeilen- und Spaltenrang gleich — und dies, obwohl die Zeilen im $\mathbb{R}^4$, die Spalten aber im $\mathbb{R}^3$ leben. Das ist kein Zufall:

Theorem 5.5.9. *Für jede Matrix $A \in M_{m,n}(\mathbb{K})$ ist Zeilenrang gleich Spaltenrang.*

Dieser Satz, der es uns ab jetzt ermöglicht, einfach vom *Rang einer Matrix* zu reden, liegt etwas tiefer. Einen begrifflichen, nicht rechnerischen Beweis des Theorems werden wir in einem späteren Abschnitt geben.

Beweis. Der Beweis, den wir hier wählen, basiert auf dem Gauß-Jordan Algorithmus. Sei $A \in M_{m,n}(\mathbb{K})$. Überlegen wir uns zunächst (was auch außerhalb dieses Beweises gut zu wissen ist!) die folgende *Behauptung: Jede elementare Zeilenoperation, wie wir sie im Eliminationsalgorithmus benutzen, formt die Matrix A in eine Matrix $A' = LA$ um, wo $L \in \mathrm{GL}_m(\mathbb{K})$ eine geeignete* invertierbare *Matrix ist.* In der Tat finden wir für die verschiedenen elementaren Operationen die folgenden Matrizen.

(E1): Addition des Vielfachen der i-ten Zeile zur j-ten Zeile (für $i \neq j$).
In diesem Fall ist
$$L = E_m + \lambda E_{ji},$$
und aus $E_{ji}^2 = 0$ folgt
$$L^{-1} = E_m - \lambda E_{ji},$$
d.h. L liegt gewiss in $\mathrm{GL}_m(\mathbb{K})$.

(E2): Vertauschen der i-ten und der j-ten Zeile.
Hier ist $L = (\ell_{kl})$ die folgende Permutationsmatrix:
$$L = L_\sigma := \sum_{i=1}^{n} E_{\sigma(i),i},$$
wobei $\sigma \in S_n$ die Transposition von i und j ist. Die Formel $L_\sigma^{-1} = L_{\sigma^{-1}}$ (Übung) besagt insbesondere, dass L invertierbar ist.

Damit ist unsere obige Behauptung bewiesen. Sei nun A' die Matrix in Stufenform, die ein gegebener Durchlauf des Algorithmus nach endlich vielen, sagen wir k, elementaren Umformungen geliefert hat. Dann ist insgesamt $A' = LA$, wo $L = L_k L_{k-1} \cdots L_1$ das Produkt der zu den einzelnen elementaren Umformungen gehörigen Matrizen L_j ist.

Sei r die Anzahl der Zeilen in der Stufenmatrix A', die nicht 0 sind. Nach Beispiel 4.5.10(4) und 4.5.17 ist r der Zeilenrang von A', welcher seinerseits gleich dem Zeilenrang der Ausgangsmatrix A ist, da sich der Zeilenrang bei elementaren Zeilenumformungen nicht ändert.

Sei s der Spaltenrang von A'. Das ist, wie wir wissen, die Dimension von
$$\mathrm{im}(\varphi_{A'}) = \{b \in \mathbb{K}^m \colon (\exists x \in \mathbb{K}^n)\ A'x = b\}.$$
Aus der speziellen Struktur von A' folgt, dass das Gleichungssystem $A'x = b$ genau dann eine Lösung $x \in \mathbb{K}^m$ besitzt, wenn $b_{r+1} = \cdots = b_m = 0$ ist. Folglich ist $\mathrm{im}(\varphi_{A'}) = \mathrm{span}(e_1, \ldots, e_r)$, wobei e_i die Standardbasisvektoren im $\mathbb{K}^m$ bezeichnen.

Also ist $r = s$. Wir müssen also nur noch zeigen, das s auch der Spaltenrang der Ausgangsmatrix A ist. Nun ist aber $\varphi_{A'} = \varphi_L \circ \varphi_A$, wie man sich sofort überlegt (vgl. Satz 5.5.17). Daraus folgt

$$\operatorname{im} \varphi_{A'} = \varphi_L(\operatorname{im} \varphi_A).$$

Andererseits ist φ_L ebenso wie L invertierbar, also ein Isomorphismus von $\mathbb{K}^m$. Also haben wir

$$\dim(\operatorname{im} \varphi_A) = \dim(\operatorname{im} \varphi_{A'}),$$

d.h. A und A' haben denselben Spaltenrang $s = r$.

Damit ist das Theorem vollständig bewiesen. □

Satz 5.5.10. *Der Rang einer Matrix A ist gleich der Maximalzahl linear unabhängiger Spalten, und ebenso gleich der Maximalzahl linear unabhängiger Zeilen von A.*

Beweis. Kombiniere Satz 5.1.8(2) mit dem vorangehenden Theorem. □

5.5.11. [Injektivität, Surjektivität, Bijektivität] Sei $A \in M_{m,n}(\mathbb{K})$. Schreibe $s_1, \ldots, s_n$ für die Spalten von A, so dass $Ae_1 = s_1, \ldots, Ae_n = s_n$ ist.

Sei $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$ die durch A definierte lineare Abbildung. Dann gilt:

(1) Das Surjektivitätskriterium:

$$\varphi_A \text{ is surjektiv} \iff \operatorname{rank} A = m$$

$$\iff \text{die Zeilen von } A \text{ sind linear unabhängig.}$$

Die erste Äquivalenz folgt aus Satz 5.1.8(3) und aus der Tatsache, dass $\operatorname{rank} A = \operatorname{rank} \varphi_A$. Die zweite folgt aus dem obigen Theorem: da es überhaupt nur m Zeilen in A gibt, sind sie genau dann linear unabhängig, wenn $\operatorname{rank} A = m$ ist.

(2) Das Injektivitätskriterium:

$$\varphi_A \text{ is injektiv} \iff \operatorname{rank} A = n$$

$$\iff \text{die Spalten von } A \text{ sind linear unabhängig.}$$

(Die Äquivalenz der ersten und dritten Aussage folgt aus Satz 5.1.8(4). Der Rest ergibt sich aus Satz 5.5.10.)

(3) φ_A ist bijektiv, d.h. φ_A ist ein Isomorphismus genau dann, wenn $n = m$ und $\operatorname{rank} A = n$. Das folgt aus (1) und (2).

(4) Für eine quadratische Matrix $A \in M_n(\mathbb{K})$ gilt:

$$\begin{aligned} \varphi_A : \mathbb{K}^n \to \mathbb{K}^n \text{ ist bijektiv} &\iff \operatorname{rank} A = n \\ &\iff \text{die Spalten von } A \text{ sind linear unabhängig} \\ &\iff \text{die Zeilen von } A \text{ sind linear unabhängig} \end{aligned}$$

5.5.12. [Der Kern von $\varphi_A\colon \mathbb{K}^n \to \mathbb{K}^m$] Definitionsgemäß ist

$$\ker \varphi_A = \{x \in \mathbb{K}^n : \varphi_A(x) = 0\} = \{x \in \mathbb{K}^n : Ax = 0\}.$$

Dieser Kern von φ_A besteht also gerade aus den Lösungen $x \in \mathbb{K}^n$ des linearen Gleichungssystems

$$Ax = 0. \tag{H}$$

Nach 5.1.1 ist $\ker \varphi_A$ ein Untervektorraum von $\mathbb{K}^n$. Also ist die Lösungsmenge von (H) ein Untervektorraum von $\mathbb{K}^n$. Seine Dimension ist nach 5.1.7:

$$\dim(\ker \varphi_A) = n - \operatorname{rank} \varphi_A = n - \operatorname{rank} A.$$

Sei $q := n - \operatorname{rank} A$. Dann besteht jede Basis von $\ker \varphi_A$ aus q Vektoren. Jede solche Basis $v_1, \ldots, v_q$ von $\ker \varphi_A$ heißt ein *Fundamentalsystem* von Lösungen des Systems (H). Die allgemeine Lösung von (H) ist dann eine beliebige Linearkombination

$$x_H = \lambda_1 v_1 + \ldots + \lambda_q v_q$$

mit $\lambda_1, \ldots, \lambda_q \in \mathbb{K}$. Das Problem der Lösung eines homogenen linearen Gleichungssystems ist also darauf zurückgeführt, q linear unabhängige Lösungen $v_1, \ldots, v_q$ zu finden (wobei $q = n - \operatorname{rank} A$ ist).

Praktisch lässt sich ein Fundamentalsystem für (H) wie folgt bestimmen:

Bringe das System durch elementare Zeilenumformungen auf Stufenform. Seien $x_{i_1}, \ldots, x_{i_q}$ die Nicht-Pivot Variablen. Ihre Anzahl q ist gerade $n - \operatorname{rank} A$. Um v_1 zu finden, wähle

$$x_{i_j} = \delta_{jk} \quad \text{für} \quad k = 1, \ldots, q = n - \operatorname{rank} A.$$

und bestimme die Lösungen $v_1, \ldots, v_q$.

Beispiel 5.5.13. Betrachte das folgende lineare Gleichungssystem:

$$\begin{array}{rcrcrcrcl}
x_1 & + & 2x_2 & - & x_3 & + & x_4 & = & 0 \\
2x_1 & & & + & x_3 & - & 2x_4 & = & 0 \\
-x_1 & + & 2x_2 & - & 2x_3 & + & 3x_4 & = & 0
\end{array}
\qquad
\begin{array}{rrrr}
1 & 2 & -1 & 1 \\
2 & 0 & 1 & -2 \\
-1 & 2 & -2 & 3
\end{array}$$

$$\begin{array}{rrrr}
1 & 2 & -1 & 1 \\
0 & -4 & 3 & -4 \\
0 & 4 & -3 & 4
\end{array}$$

$$\begin{array}{rcrcrcrcl}
x_1 & + & 2x_2 & - & x_3 & + & x_4 & = & 0 \\
 & - & 4x_2 & + & 3x_3 & - & 3x_4 & = & 0 \\
 & & & & & & 0 & = & 0
\end{array}
\qquad
\begin{array}{rrrr}
1 & 2 & -1 & 1 \\
0 & -4 & 3 & -4 \\
0 & 0 & 0 & 0
\end{array}$$

- Pivot-Variable: x_1, x_2
- Nich-Pivot Variable: x_3, x_4.

Die Wahl $x_3 = 1$, $x_4 = 0$ ergibt

$$v_1 = \begin{pmatrix} -\frac{1}{2} \\ \frac{3}{4} \\ 1 \\ 0 \end{pmatrix}$$

während die Wahl $x_3 = 0$, $x_4 = 1$ liefert:

$$v_2 = \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{4} \\ 0 \\ 1 \end{pmatrix}$$

Die Vektoren v_1, v_2 bilden ein Fundamentalsystem von Lösungen des homogenen Systems.

Definition 5.5.14. (Matrix einer linearen Abbildung bzgl. der Standardbasen) Sei $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ eine lineare Abbildung. Wir bilden die $m \times n$-Matrix, deren Spalten die Bilder $\varphi(e_1), \ldots, \varphi(e_n)$ der Standardbasisvektoren $e_1, \ldots, e_n$ in $\mathbb{K}^n$ sind:

$$[\varphi] := A := \begin{pmatrix} | & | & & | \\ | & | & & | \\ \varphi(e_1) & \varphi(e_2) & \ldots & \varphi(e_n) \\ | & | & & | \\ | & | & & | \end{pmatrix}$$

Diese Matrix $[\varphi] = A$ heißt *die Matrix von* φ *bzgl. der Standardbasen* von $\mathbb{K}^n$ bzw. $\mathbb{K}^m$.

Dann gilt für alle $x \in \mathbb{K}^n$ die Gleichung $\varphi(x) = Ax$, oder mit anderen Worten:

$$\varphi = \varphi_A.$$

In der Tat ist nämlich $Ae_j = j$-te Spalte von $A = \varphi(e_j)$ für alle $j = 1, \ldots, n$, also $\varphi_A(e_j) = \varphi(e_j)$ für alle $j = 1 \ldots, n$. Wenn aber zwei lineare Abbildungen auf einer Basis übereinstimmen, so stimmen sie auf allen Elementen überein (Satz 5.1.9).

Wir können also jede lineare Abbildung $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ durch eine eindeutige $m \times n$-Matrix $[\varphi] = A$ beschreiben. Was wir über die Beziehungen zwischen Eigenschaften von φ_A und Eigenschaften von A ausgesagt haben, gilt ebenso für die Beziehungen zwischen Eigenschaften von linearen Abbildungen φ und Eigenschaften der ihnen entsprechenden Matrizen $[\varphi] = A$.

Wir schießen diesen Gedankengang mit folgender Beobachtung ab:

Satz 5.5.15. *Ordnet man jeder $m \times n$-Matrix A die lineare Abbildung $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$ zu, so liefert das einen linearen Isomorphismus*

$$\Phi\colon M_{m,n}(\mathbb{K}) \to \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m), \quad A \mapsto \varphi_A.$$

Die Umkehrabbildung ordnet jeder linearen Abbildung $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ ihre Matrix $[\varphi]$ bzgl. der Standardbasen zu.

Beweis. Wir wissen schon, dass die Abbildung Φ bijektiv ist. Es bleibt nur zu zeigen, dass sie auch linear ist, d.h.:

$$\varphi_{A+B} = \varphi_A + \varphi_B, \quad \varphi_{\lambda A} = \lambda\varphi_A$$

für alle $A, B \in M_{m,n}(\mathbb{K})$ und alle $\lambda \in \mathbb{K}$. Hier ist der Beweis dafür:

$$\varphi_{A+B}(x) = (A+B)x = Ax + Bx = \varphi_A(x) + \varphi_B(x) = (\varphi_A + \varphi_B)(x)$$

und

$$\varphi_{\lambda A}(x) = (\lambda A)x = \lambda(Ax) = \lambda\varphi_A(x) = (\lambda\varphi_A)(x)$$

für alle $x \in \mathbb{K}^n$. $\square$

Aus 4.2.4(3) folgt, dass auch die Umkehrabbildung

$$\mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m) \to M_{m,n}(\mathbb{K}), \qquad \varphi \mapsto [\varphi]$$

linear ist. Das bedeutet

$$[\varphi + \psi] = [\varphi] + [\psi], \qquad [\lambda\varphi] = \lambda[\varphi]$$

für alle $\varphi, \psi \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ und alle $\lambda \in \mathbb{K}$.

Anders gesagt: Ist A die zu φ gehörige Matrix und B die zu ψ gehörige Matrix, so ist die Matrix von $\varphi + \psi$ gleich $A + B$ und die Matrix von $\lambda\varphi$ ist λA.

Korollar 5.5.16. $\dim \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m) = n \cdot m$.

Beweis. In der Tat ist $\dim(M_{m,n}(\mathbb{K})) = nm$ gemäß 5.4.1 und isomorphe Vektorräume haben dieselbe Dimension (Übung). $\square$

Satz 5.5.17. (1) *Für $A \in M_{m,n}(\mathbb{K})$ und $B \in M_{n,p}(\mathbb{K})$ gilt*

$$\varphi_A \circ \varphi_B = \varphi_{AB}.$$

(2) *Für lineare Abbildungen $\psi\colon \mathbb{K}^p \to \mathbb{K}^n$ und $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ gilt*

$$[\varphi \circ \psi] = [\varphi][\psi].$$

Beweis. (1) $(\varphi_A \circ \varphi_B)(x) = \varphi_A(\varphi_B(x)) = \varphi_A(Bx) = A(Bx) = (AB)x = \varphi_{AB}(x)$ für alle $x \in \mathbb{K}^n$; daher $\varphi_A \circ \varphi_B = \varphi_{AB}$.

(2) $[\varphi \circ \psi]x = (\varphi \circ \psi)(x) = \varphi(\psi(x)) = [\varphi]([\psi]x) = ([\varphi][\psi])x$ für alle $x \in \mathbb{K}^p$, und daher $[\varphi \circ \psi] = [\varphi][\psi]$. $\square$

Beispiel 5.5.18. Für viele Anwendungen ist es nützlich, sich lineare Abbildungen $\mathbb{R}^2 \to \mathbb{R}^2$ und $\mathbb{R}^3 \to \mathbb{R}^3$ auch geometrisch vorzustellen. Im Kontext der sogenannten euklidischen Vektorräume wird das im zweiten Semester systematisch zum Thema gemacht; wir wollen aber schon hier einige Begriffe und Beispiele einführen.

Das *Skalarprodukt* auf $\mathbb{R}^n$ wird wie folgt definiert:

$$\langle x, y\rangle := \sum_{i=1}^{n} x_i y_i.$$

Identifiziert man $\mathbb{R}^n$ mit dem Vektorraum $M_{n,1}(\mathbb{R})$ aller $(n \times 1)$-Matrizen (Spaltenvektoren), so kann man das Skalarprodukt als Matrizenprodukt schreiben:

$$\langle x, y\rangle = x^\top y = y^\top x.$$

Für $x \neq 0$ ergibt sich:

$$\langle x, x\rangle = \sum_{i=1}^{n} x_i^2 > 0.$$

Die nicht-negative reelle Zahl

$$\|x\| := \sqrt{\langle x, x\rangle}$$

heißt die *Länge von x*. Der einzige Vektor der Länge 0 ist der Nullvektor.

Zwei Vektoren $x, y \in \mathbb{R}^n$ heißen *orthogonal*, falls $\langle x, y\rangle = 0$ ist.

(a) Für $n = 2$ hat man

$$\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \rangle = -x_1 x_2 + x_1 x_2 = 0.$$

Ist $x \neq 0$, so sind die Vektoren x und

$$\widetilde{x} := \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

linear unabhängig (denn $\widetilde{x}$ ist kein Vielfaches von x), und wir haben die direkte Summenzerlegung

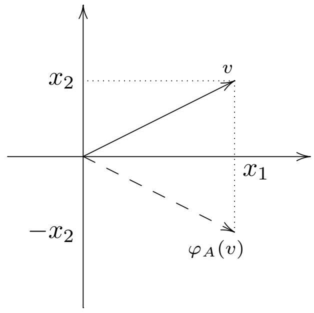
$$\mathbb{R}^2 = \mathbb{R}x \oplus \mathbb{R}\widetilde{x}$$

von $\mathbb{R}^2$ in zwei zueinander orthogonale Geraden. Die *orthogonale Spiegelung* an der Gerade $\mathbb{R}x$ ist definiert als die eindeutig bestimmte lineare Abbildung $r_x \colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $r_x(x) = x$ und $r_x(\widetilde{x}) = -\widetilde{x}$ (vgl. Satz 5.1.5) — man stelle sich diese Abbildung geometrisch in der Ebene vor!

Für $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ erhalten wir $\widetilde{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und die Matrix

$$[r_x] = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

Für die Matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ gilt $\varphi_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$. Die zugehörige Abbildung φ_A ist also die Spiegelung an der x_1-Achse $\mathbb{R}e_1$.



(b) Eine andere interessante Klasse linearer Abbildungen $\mathbb{R}^2 \to \mathbb{R}^2$ sind *Drehungen um den Ursprung*. Die Matrix der Drehung um den Winkel $\alpha \in \mathbb{R}$ ist:

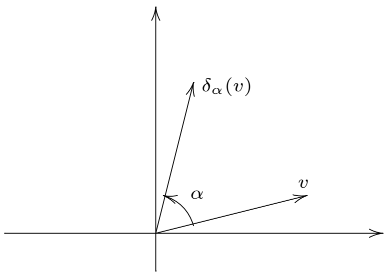
$$\delta_\alpha := \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

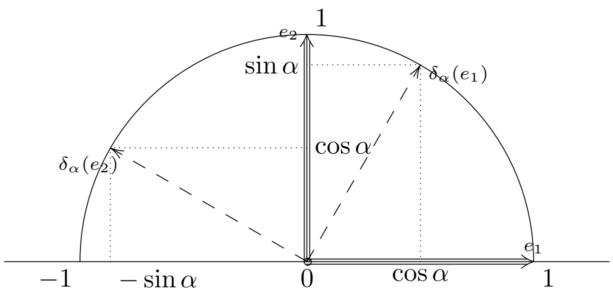
wobei wir für die reellen Funktionen Kosinus und Sinus auf die Analysis Vorlesung verweisen müssen. Aus den Additionstheoremen der trigonometrischen Funktionen folgt:

$$\delta_\alpha\delta_\beta = \delta_{\alpha+\beta} \quad \text{und} \quad \delta_\alpha\delta_{-\alpha} = \delta_0 = \mathrm{id}_{\mathbb{R}^2}\,.$$

Insbesondere ist $\delta_\alpha \in \mathrm{GL}_2(\mathbb{R})$ mit $\delta_\alpha^{-1} = \delta_{-\alpha}$. Es gilt

$$\delta_\alpha(e_1) = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \quad \text{und} \quad \delta_\alpha(e_2) = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}.$$





Als Spezialfall betrachten wir etwa $\alpha := \frac{\pi}{3} = 60°$. Dann ist $\cos\frac{\pi}{3} = \frac{1}{2}$,

$\sin\frac{\pi}{3} = \frac{1}{2}\sqrt{3}$. Also ist die Matrix der Drehung um 60°:

$$\delta_\alpha = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$$

Das Bild des Vektors $a = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ unter der Drehung um 60° ist also

$$\delta_\alpha(a) = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1-\sqrt{3} \\ 1+\sqrt{3} \end{pmatrix}$$

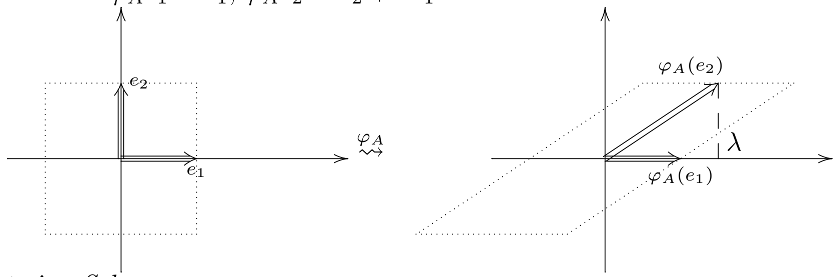
(c) Sei jetzt $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die lineare Abbildung, die man durch Drehung (um den Ursprung) um den Winkel α und nachfolgender Spiegelung an der Gerade $\mathbb{R}x$, wo $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ist, erhält. Dann ist

$$\varphi = r_x \circ \varphi_\alpha$$

wobei φ_α die Drehung und r_x die Spiegelung bezeichnet. Die zu φ gehörige Matrix (bzgl. der Standardbasis) erhält man durch Multiplikation der Matrix von r_x mit der Matrix von δ_α, d.h.:

$$[\varphi] = [r_x][\varphi_\alpha] = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} -\sin\alpha & -\cos\alpha \\ -\cos\alpha & \sin\alpha \end{pmatrix}.$$

(d) Sei $A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. Dann ist $\varphi_A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + \lambda x_2 \\ x_2 \end{pmatrix}$. Insbesondere hat man $\varphi_A e_1 = e_1$, $\varphi_A e_2 = e_2 + \lambda e_1$.



φ_A ist eine *Scherung*.

□

Wir nehmen jetzt noch einmal systematisch auf, was wir schon in 5.5.11(4) erkannt hatten:

Satz 5.5.19. (Isomorphismen und invertierbare Matrizen)

(1) *Eine $n \times n$-Matrix A ist genau dann invertierbar, wenn die lineare Abbildung $\varphi_A\colon \mathbb{K}^n \to \mathbb{K}^n$ ein Isomorphismus ist. In diesem Fall gilt:*

$$\varphi_{A^{-1}} = (\varphi_A)^{-1}$$

(2) *Eine lineare Abbildung* $\varphi\colon \mathbb{K}^n \to \mathbb{K}^m$ *ist genau dann ein Isomorphismus, wenn* $n = m$ *und wenn die zu* φ *gehörige Matrix* $[\varphi]$ *invertierbar ist. In diesem Fall gilt* $[\varphi^{-1}] = [\varphi]^{-1}$.

Beweis. (1) Ist A invertierbar, so haben wir $AA^{-1} = E_n = A^{-1}A$. Daher ist

$$\varphi_A \circ \varphi_{A^{-1}} = \varphi_{AA^{-1}} = \varphi_{E_n} = \mathrm{id}_{\mathbb{K}^n} = \varphi_{E_n} = \varphi_{A^{-1}A} = \varphi_{A^{-1}} \circ \varphi_A.$$

Daraus folgt, dass φ_A bijektiv und $\varphi_{A^{-1}}$ sein Inverses ist.

Ist umgekehrt φ_A ein Isomorphismus und $B := [\varphi_A^{-1}]$ die Matrix der inversen Abbildung, so ergibt sich

$$AB = [\varphi_A][\varphi_A^{-1}] = [\varphi_A \circ \varphi_A^{-1}] = [\mathrm{id}] = E_n$$

und ebenso $BA = E_n$. Also ist A invertierbar mit $A^{-1} = [\varphi_A^{-1}]$, so dass auch $\varphi_{A^{-1}} = \varphi_A^{-1}$ gilt.

(2) Ist φ ein Isomorphismus, so folgt $n = m$ nach 5.5.11(3). Wir wenden nun (1) auf $A = [\varphi]$ und $\varphi = \varphi_A$ an. □

Beispiel 5.5.20. Die Abbildung $\varphi = r_x \circ \varphi_\alpha$ aus 5.5.18(c) ist bijektiv. Die Matrix von φ^{-1} ist: $[\varphi^{-1}] = [\varphi]^{-1} = [r_x \circ \varphi_\alpha]^{-1} = ([r_x][\varphi_\alpha])^{-1} = [\varphi_\alpha]^{-1}[r_x]^{-1} =$

$$= \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} -\sin\alpha & -\cos\alpha \\ -\cos\alpha & \sin\alpha \end{pmatrix} = [\varphi].$$

Insbesondere ist also $\varphi^{-1} = \varphi$ und $\varphi^2 = \mathrm{id}_{\mathbb{R}^2}$.

Satz 5.5.21. *Seien* A *und* B *zwei* $n \times n$*-Matrizen mit* $AB = E_n$*. Dann sind sowohl* A *als auch* B *invertierbar und es ist* $B = A^{-1}$.

Beweis. Zunächst ist $\varphi_A \circ \varphi_B = \varphi_{AB} = \varphi_{E_n} = \mathrm{id}_{\mathbb{K}^n}\colon \mathbb{K}^n \to \mathbb{K}^n$. Daraus schließen wir, dass φ_A surjektiv ist. Eine lineare Abbildung eines endlich dimensionalen Vektorraums in sich selbst ist aber genau dann surjektiv, wenn sie bijektiv ist (5.1.5). Also ist φ_A bijektiv. Nach 5.5.11(4) ist A dann invertierbar. Multipliziert man die Gleichung $AB = E_n$ mit A^{-1} von links, erhält man $B = A^{-1}$, und B ist ebenfalls invertierbar. □

Bemerkung 5.5.22. In einem beliebigen Monoid folgt aus $ab = e$ nicht schon, dass a invertierbar ist; dafür muss auch $ba = e$ mit demselben b gelten. Für $n \times n$-Matrizen reicht aber schon eine der beiden Identitäten.

Betrachten wir z.B. die lineare Abbildung

$$\varphi\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R}), \quad \varphi(p)(x) := xp(x).$$

Wir haben schon gesehen, dass φ eine injektive lineare Abbildung ist, die *nicht surjektiv* ist. Da die Monome p_n mit $p_n(x) = x^n$, $n \in \mathbb{N}_0$, eine Basis von $\mathcal{P}(\mathbb{R})$ bilden, gibt es genau eine lineare Abbildung

$$\psi\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$$

mit

$$\psi(p_n) = \begin{cases} p_{n-1} & \text{if } n \geq 1 \\ 0 & \text{if } n = 0. \end{cases}$$

Wegen $\varphi(p_n) = p_{n+1}$ sieht man sofort ein, dass $\psi \circ \varphi = \mathrm{id}_{\mathcal{P}(\mathbb{R})}$, aber $p_0 \in \ker(\varphi \circ \psi)$, so dass also sicher $\varphi \circ \psi \neq \mathrm{id}_{\mathcal{P}(\mathbb{R})}$ ist. Also hat die Abbildung φ zwar ein Linksinverses, aber kein Rechtsinverses.

Übung 5.5.23. Sei $\varphi\colon V \to W$ ein Vektorraumisomorphismus. Zeige:

1. Eine Teilmenge B erzeugt V genau dann, wenn $\varphi(B)$ den Raum W erzeugt.

2. Eine Teilmenge $B \subseteq V$ ist genau dann linear unabhängig, wenn $\varphi(B)$ linear unabhängig ist.

3. Eine Teilmenge $B \subseteq V$ ist genau dann eine Basis, wenn $\varphi(B)$ eine Basis von W ist.

4. $\dim V = \dim W$.

Die nächste Übung gibt einen begrifflicheren Beweis des Satzes 5.4.18.

Übung 5.5.24. Sei $\varphi\colon M_1 \to M_2$ ein Monoidhomomorphismus und schreibe $M_i^\times$ für die Einheitengruppe von M_i (für $i = 1, 2$). Zeige:

1. Es gilt $\varphi(M_1^\times) \subseteq M_2^\times$ und die Einschränkung

$$\varphi|_{M_1^\times} : M_1^\times \to M_2^\times$$

ist ein Gruppenhomomorphismus. Insbesondere gilt:

$$\varphi(m^{-1}) = \varphi(m)^{-1} \quad \text{für} \quad m \in M_1^\times.$$

2. Ist φ ein Isomorphismus, so gilt $\varphi(M_1^\times) = M_2^\times$ und $m \in M_1$ ist genau dann invertierbar, wenn $\varphi(m)$ invertierbar ist.

Übung 5.5.25. Sei M ein Monoid mit neutralem Element e, und sei $a \in M$. Ist

$$ab = e \quad \text{und} \quad b'a = e,$$

so folgt $b = b'$. Mit anderen Worten: hat a ein Rechtsinverses b und ein Linksinverses b', so stimmen beide überein und a is invertierbar mit $a^{-1} = b$.

Übung 5.5.26. Gib eine geometrische Beschreibung der Abbildung $\varphi_A \in \mathrm{End}(\mathbb{R}^2)$, die durch die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

gegeben ist.

5.6 Lineare Abbildungen und ihre Matrizen

In diesem Abschnitt seien V ein $\mathbb{K}$-Vektorraum der Dimension n mit einer angeordneten Basis $B = (v_1, v_2, \ldots, v_n)$ und W ein $\mathbb{K}$-Vektorraum der Dimension m mit einer angeordneten Basis $C = (w_1, w_2, \ldots, w_m)$.

Aus Abschnitt 4.5 wissen wir, dass jedes Element $v \in V$ eindeutig in der Form

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \quad \text{mit } \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$$

dargestellt werden kann. Die durch v eindeutig bestimmten Skalare $\lambda_1, \lambda_2, \ldots, \lambda_n$ heißen die *Koordinaten* von v bzgl. der angeordneten Basis B, und wir schreiben

$$[v]_B := \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n$$

für den Koordinatenvektor von v bzgl. der angeordneten Basis B. *Im Folgenden werden wir der Einfachheit halber das Wort „angeordnet“ weglassen, und $(v_1, \ldots, v_n)$ schlicht als Basis bezeichnen. Die Anordnung steckt aber in der Nummerierung der Basiselements $v_1, \ldots, v_n$.* Die Abbildung

$$\kappa_B \colon V \to \mathbb{K}^n, \quad v \mapsto [v]_B$$

die jedem $v \in V$ seinen Koordinatenvektor $[v]_B$ zuordnet, ist ein Vektorraumisomorphismus, wie wir in 4.5 gesehen haben. Beachte, dass für die Koordinaten der Basisvektoren v_j gilt:

$$[v_j]_B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_j$$

Wir kommen nun zu der zentralen Technik der Linearen Algebra, der Beziehung zwischen Matrizen und linearen Abbildungen.

$\boxed{\text{Die Matrix einer linearen Abbildung}}$

Definition 5.6.1. Sei $\varphi \colon V \to W$ eine lineare Abbildung. Wir bestimmen die Koordinaten der Bilder $\varphi(v_j)$ der Basisvektoren $v_1, \ldots, v_n$ bzgl. der Basis $C = (w_1, \ldots, w_m)$ von W:

$$\varphi(v_j) = a_{1j} w_1 + \cdots + a_{mj} w_m,$$

und wir bilden die $m \times n$-Matrix A, deren j-te Spalte gerade die Koordinaten von $\varphi(v_j)$ bzgl. C angibt:

$$A = [\varphi]_C^B := \begin{array}{c} \begin{array}{ccc} [\varphi(v_1)]_C & \qquad & [\varphi(v_n)]_C \end{array} \\ \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \end{array}$$

Die j-te Spalte von A ist also der Koordinatenvektor $[\varphi(v_j)]_C$. Die Matrix $A = [\varphi]_C^B$ heißt *die Matrix von φ bzgl. der Basen B und C*.

Satz 5.6.2. *$[\varphi]_C^B[v]_B = [\varphi(v)]_C$ für alle $v \in V$.*

Das bedeutet: Die Matrix von φ stellt φ in dem Sinne dar, dass die Koordinaten von $\varphi(v)$ als Matrizenprodukt der Matrix von φ mit den Koordinaten von v erhalten werden.

Oder noch anders gesagt: Sei $A = [\varphi]_C^B$ die zu φ bzgl. B und C gehörige Matrix. Sind $[v]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ die Koordinaten eines Vektors $v \in V$, so sind die Koordinaten $[\varphi(v)]_C = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ von $\varphi(v)$ bzgl. C durch die Gleichung:

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

gegeben.

Das folgende Diagramm fasst diese Beziehung zusammen. Seine obere Hälfte gehört zur abstrakten Seite der Linearen Algebra, die untere zur konkreten Seite.

$$\begin{tikzcd}
v \arrow[rrr, mapsto] \arrow[ddd, mapsto] & & & \varphi(v) \arrow[ddd, mapsto] \\
& V \arrow[r, "\varphi"] \arrow[d, "\kappa_B"'] & W \arrow[d, "\kappa_C"] & \\
& \mathbb{K}^n \arrow[r, "\varphi_A", "A=[\varphi]_C^B"'] & \mathbb{K}^m & \\
{[v]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}} \arrow[rrr, mapsto] & & & {A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}}
\end{tikzcd}$$

Beweis. Da die Abbildungen $v \mapsto [\varphi]^B_C[v]_B$ und $v \mapsto [\varphi(v)]_C, V \to \mathbb{K}^n$ beide linear sind, genügt es nachzuprüfen, dass sie auf den Basisvektoren $v_1, \ldots, v_n$ übereinstimmen. Weil $[v_j]_B = e_j$ ist, gilt

$$[\varphi]^B_C[v_j]_B = [\varphi]^B_C e_j = j.\text{ Spalte von } [\varphi]^B_C = [\varphi(v_j)]_C,$$

womit die Behauptung bewiesen ist. □

5.6.3. [Spezialfall: Die Matrix eines Endomorphismus] Ist $\varphi\colon V \to V$ ein Endomorphismus, so wählt man üblicherweise dieselbe Basis $B = C$ auf beiden Seiten, anstatt für ein und denselben Vektorraum zwei verschiedene Basen gleichzeitig zu betrachten.

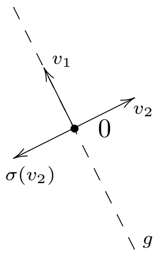
Die Matrix $A = [\varphi]^B_B$ heißt dann *die Matrix von φ bzgl. der Basis B*. Sie ist eine (quadratische!) $n \times n$-Matrix. Gemäß dem allgemeinen Rezept von 5.6.1, wird sie wie folgt gebildet: Man stelle die Bildvektoren $\varphi(v_j)$ der Basis $v_1, \ldots, v_n$ durch ihre Koordinaten bzgl. derselben Basis dar:

$$\varphi(v_j) = a_{1j}v_1 + \ldots + a_{nj}v_n.$$

Die Koeffizienten $a_{1j}, \ldots, a_{nj}$ bilden die j-te Spalte von A. Wir schreiben einfach $[\varphi]_B$ für $[\varphi]^B_B$.

Beispiele 5.6.4.

(a) Sei $\sigma\colon \mathbb{R}^2 \to \mathbb{R}^2$ die (orthogonale) Spiegelung an der Gerade g durch 0. Um σ durch eine Matrix darzustellen, wählen wir eine Basis v_1, v_2, die dieser besonderen Abbildung angepasst ist: Sei v_1 ein Vektor, der in der Richtung der Gerade g liegt, und v_2 ein zu g orthogonaler Vektor. Dann ist



$$\sigma(v_1) = v_1 = 1 \cdot v_1 + 0v_2$$
$$\sigma(v_2) = -v_2 = 0 \cdot v_1 + (-1)v_2$$

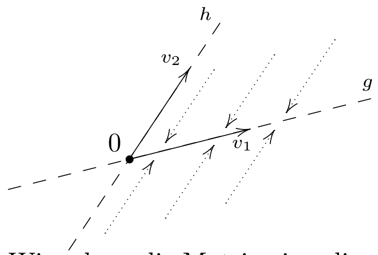
Die Matrix von σ bzgl. der Basis $B = (v_1, v_2)$ ist

$$[\sigma]_B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(b) Seien g und h zwei Geraden im $\mathbb{R}^2$, die sich in 0 schneiden. Sei $\pi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Parallelprojektion auf die Gerade g in Richtung von h. Wieder wählen wir eine der Situation angepasste Basis, um π durch eine Matrix darzustellen: $B = (v_1, v_2)$, wobei $g = \mathbb{R}v_1$ und $h = \mathbb{R}v_2$ ist. Dann ist

$$\pi(v_1) = v_1 = 1 \cdot v_1 + 0 \cdot v_2$$
$$\pi(v_2) = 0 = 0 \cdot v_1 + 0 \cdot v_2$$

und die Matrix von π bzgl. der Basis B ist $[\pi]_B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.



Wir sehen: die Matrix einer linearen Abbildung kann durch geeignete Basiswahl sehr einfach werden.

Satz 5.6.5. *Seien V und W endlich dimensionale Vektorräume und $\varphi\colon V \to W$ eine lineare Abbildung. Wir behaupten: es gibt eine Basis B von V und eine Basis C von W, so dass die Matrix von φ bzgl. B und C folgende Form hat:*

$$[\varphi]_C^B = \left(\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & & \vdots \\ 0 & \cdots & 0 & 1 & \vdots & & & \vdots \\ \hline 0 & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & 0 \\ \vdots & & & & \vdots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & 0 \end{array}\right)$$

Die Größe der Einheitsmatrix in der linken oberen Ecke ist $r = \operatorname{rank}\varphi$.

Beweis. Wie im Beweis der Dimensionsformel 5.1.4 wählen wir eine Basis $v_{r+1}, \ldots, v_n$ von $\ker\varphi$ und eine Basis $w_1, \ldots, w_r$ von $\operatorname{im}(\varphi)$. Beachte, dass $r = \operatorname{rank}\varphi = \dim(\operatorname{im}(\varphi)))$. Und weiter wählen wir $v_1, \ldots, v_r \in V$ mit $\varphi(v_j) = w_j$ für $j = 1, \ldots, r$. Dann folgt mit denselben Schlüssen wie im Beweis der Dimensionsformel, dass $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$ eine Basis B von V ist. Nach dem Basiserweiterungssatz finden wir Vektoren $w_{r+1}, \ldots, w_m$, so dass $w_1, \ldots, w_m$ eine Basis von W ist. Da

$$\varphi(v_1) = w_1, \ldots, \varphi(v_r) = w_r, \quad \text{und} \quad \varphi(v_{r+1}) = 0, \ldots, \varphi(v_n) = 0,$$

hat die Matrix $[\varphi]_C^B$ von φ bzgl. B und C die gewünschte Form. □

Jeder linearen Abbildung $\varphi\colon V \to W$ haben wir ihre Matrix $[\varphi]_C^B$ bzgl. der Basen B und C zugeordnet: So erhalten wir, für feste Basen B und C, eine Abbildung

$$\Phi\colon \operatorname{Hom}(V, W) \to M_{m,n}(\mathbb{K}), \quad \varphi \mapsto [\varphi]_C^B = [\kappa_C \circ \varphi \circ \kappa_B^{-1}].$$

Wie in Abschnitt 5.5 sehen wir, dass dies ein linearer Isomorphismus ist, der die Multiplikation respektiert:

Satz 5.6.6.

(1) *Die Abbildung* $\Phi\colon \operatorname{Hom}(V,W) \to M_{m,n}(\mathbb{K})$ *ist bijektiv.*

(2) *Die Abbildung* Φ *ist auch linear, und daher ein Vektorraumisomorphismus: Für alle* $\varphi, \psi \in \operatorname{Hom}(V,W)$ *und* $\lambda \in \mathbb{K}$, *gilt*

$$[\varphi + \psi]_C^B = [\varphi]_C^B + [\psi]_C^B \qquad\qquad [\lambda\varphi]_C^B = \lambda[\varphi]_C^B$$

(3) *Da die Vektorräume* $\operatorname{Hom}(V,W)$ *und* $M_{m,n}(\mathbb{K})$ *isomorph sind, haben sie dieselbe Dimension:*

$$\dim\big(\operatorname{Hom}(V,W)\big) = m \cdot n = \dim V \cdot \dim W.$$

Satz 5.6.7. *Seien* U, V *und* W *Vektorräume mit den Basen* $A = (u_1, \dots, u_p)$, $B = (v_1, \dots, v_n)$, *und* $C = (w_1, \dots, w_m)$ *sowie* $\psi\colon U \to V$ *und* $\varphi\colon V \to W$ *lineare Abbildungen. Dann gilt*

$$[\varphi \circ \psi]_C^A = [\varphi]_C^B [\psi]_B^A$$

Dies ist im folgenden Diagramm zusammengefasst:

$$\begin{tikzcd}
U \arrow[r, "\psi"] \arrow[d, "\kappa_A"'] & V \arrow[r, "\varphi"] \arrow[d, "\kappa_B"'] & W \arrow[d, "\kappa_C"'] \\
\mathbb{K}^p \arrow[r, "{[\psi]_B^A}"'] & \mathbb{K}^n \arrow[r, "{[\varphi]_C^B}"'] & \mathbb{K}^m
\end{tikzcd}$$

Beweis. Aus

$$[\varphi(\psi(u_i))]_C = [\varphi]_C^B[\psi(u_i)]_B = [\varphi]_C^B[\psi]_B^A[u_i]_A = [\varphi]_C^B[\psi]_B^A e_i$$

folgt sofort, dass $[\varphi \circ \psi]_C^A = [\varphi]_C^B[\psi]_B^A$ gilt. □

Satz 5.6.8. *Eine lineare Abbildung* $\varphi\colon V \to W$ *ist genau dann bijektiv (und also ein linearer Isomorphismus), wenn* $\dim V = \dim W$ *und die Matrix* $[\varphi]_C^B$ *invertierbar ist. Ist dem so, dann gilt*

$$[\varphi^{-1}]_B^C = \big([\varphi]_C^B\big)^{-1}$$

Beweis. Für $A = [\varphi]_C^B$ gilt $\varphi = \kappa_C^{-1} \circ \varphi_A \circ \kappa_B$, so dass φ genau dann ein Isomorphismus ist, wenn dies für φ_A der Fall ist. Wir wissen aber schon, dass dies gleichbedeutend zur Invertierbarkeit von A ist. In diesem Fall folgt aus $\varphi^{-1} = \kappa_B^{-1} \circ \varphi_A^{-1} \circ \kappa_C$ sofort $[\varphi^{-1}]_B^C = A^{-1}$. □

5.7 Basiswechsel

In diesem Abschnitt betrachten wir zwei Basen in einem Vektorraum. Unser Ziel ist es zu verstehen, wie sich die Koordinaten eines Vektors und die Matrix einer linearen Abbildung beim Übergang von einer Basis zur anderen ändern.

Definition 5.7.1. [Koordinatenwechsel] Seien

$$B = (v_1, \dots, v_n), \qquad B' = (v'_1, \dots, v'_n)$$

zwei Basen in einem n-dimensionalen Vektorraum V über dem Körper $\mathbb{K}$. Für ein beliebiges $v \in V$ seien $x_1, \dots, x_n$ die Koordinaten bzgl. der Basis B und $x'_1, \dots, x'_n$ die Koordinaten bzgl. B', d.h.:

$$[v]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad [v]_{B'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Die *Übergangsmatrix* $S := (s_{ij})$ erhält man wie folgt: Bestimme die Koordinaten der neuen Basisvektoren v'_j bzgl. der Basis B:

$$v'_j = s_{1j} v_1 + \ldots + s_{nj} v_n$$

Diese Koordinaten bilden die j-te Spalte von S:

$$S = \begin{array}{c} \begin{matrix} [v'_1]_B & \qquad & [v'_j]_B & \qquad & [v'_n]_B \end{matrix} \\ \begin{pmatrix} s_{11} & \cdots & s_{1j} & \cdots & s_{1n} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ s_{n1} & \cdots & s_{nj} & \cdots & s_{nn} \end{pmatrix} \end{array} = [\mathrm{id}]_B^{B'}.$$

Da die $v'_1, \dots, v'_n$ linear unabhängig sind, sind die Spalten von S linear unabhängig. Folglich ist S invertierbar (Satz 5.6.8). Wir behaupten, dass

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = S \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = S^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

oder kürzer

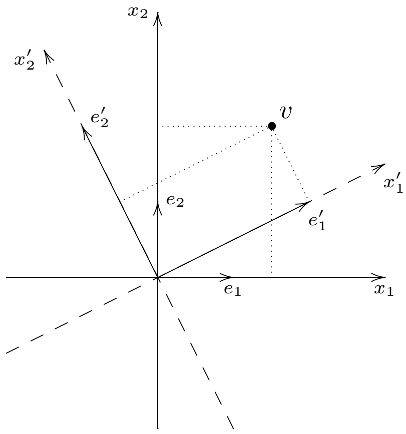
$$[v]_B = S[v]_{B'}, \qquad [v]_{B'} = S^{-1} \cdot [v]_B.$$

In der Tat folgt aus $S = [\mathrm{id}_V]_B^{B'}$ sofort $S[v]_{B'} = [\mathrm{id}_V]_B^{B'}[v]_{B'} = [v]_B$ aus Satz 5.6.2.

$$\begin{tikzcd} V_{B'} \arrow[r, "\mathrm{id}_V"] \arrow[d, "\kappa_{B'}"'] & V_B \arrow[d, "\kappa_B"] \\ \mathbb{K}^n \arrow[r, "\varphi_S"'] & \mathbb{K}^n \end{tikzcd}$$

Beispiel 5.7.2. Betrachte die Standardbasis B: $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und die Basis B': $e_1' = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $e_2' = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

Dann ist die Übergangsmatrix $S = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$.
Und ihre Inverse ist (nachprüfen!) $S^{-1} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$. Also



$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}, \qquad \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

d.h.,

$$\begin{aligned} x_1 &= 2x_1' - x_2', \qquad & x_1' &= \frac{1}{5}(2x_1 + x_2), \\ x_2 &= x_1' + 2x_2', \qquad & x_2' &= \frac{1}{5}(-x_1 + 2x_2). \end{aligned}$$

Satz 5.7.3. (Transformationsformel für Matrizen) *Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen $\mathbb{K}$-Vektorräumen V und W. Seien $B = (v_1, \ldots, v_n)$ und $B' = (v_1', \ldots, v_n')$ zwei Basen von V, und sei $S = [\mathrm{id}_V]_B^{B'}$ die Übergangsmatrix wie in 5.7.1. Seien $C = (w_1, \ldots, w_m)$ und $C' = (w_1', \ldots, w_m')$ zwei Basen von W, und $R = [\mathrm{id}_W]_C^{C'}$ die zugehörige Übergangsmatrix. Dann besteht zwischen der Matrix $A := [\varphi]_C^B$ von φ bzgl. der Basen B und C, und der Matrix $A' := [\varphi]_{C'}^{B'}$ von φ bzgl. der Basen B' und C' folgende Beziehung*

$$\boxed{A' = R^{-1}AS, \quad A = RA'S^{-1}.}$$

Beweis. Mit Satz 5.6.7 erhalten wir

$$A' = [\varphi]_{C'}^{B'} = [\mathrm{id}_W \circ \varphi \circ \mathrm{id}_V]_{C'}^{B'} = [\mathrm{id}_W]_{C'}^{C} \cdot [\varphi]_C^B \cdot [\mathrm{id}_V]_B^{B'} = R^{-1} \cdot A \cdot S.$$

$$\begin{tikzcd} V \arrow[r, "\mathrm{id}_V"'] \arrow[rrr, bend left, "\varphi"] \arrow[d, "\kappa_{B'}"'] & V \arrow[r, "\varphi"'] \arrow[d, "\kappa_B"'] & W \arrow[r, "\mathrm{id}_W"'] \arrow[d, "\kappa_C"'] & W \arrow[d, "\kappa_{C'}"'] \\ \mathbb{K}^n \arrow[r, "S"] \arrow[rrr, bend right, "A'"'] & \mathbb{K}^n \arrow[r, "A"] & \mathbb{K}^m \arrow[r, "R^{-1}"] & \mathbb{K}^m \end{tikzcd}$$

□

Satz 5.7.4. (Transformation der Matrix eines Endomorphismus) *Sei $\varphi\colon V \to V$ ein Endomorphismus von V. Seien B und B' zwei Basen von V, und sei A die Matrix von φ bzgl. B, und A' die Matrix von φ bzgl. B', wie in 5.6.3. Dann gilt*

$$A' = S^{-1}AS \quad \textit{und} \quad A = SA'S^{-1},$$

wobei $S = [\mathrm{id}_V]_B^{B'}$ *die Übergangsmatrix gemäß 5.7.1 ist.*

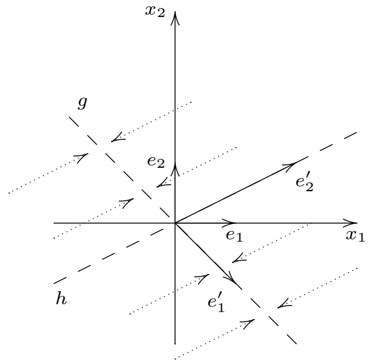
Beweis. Setze $V = W$, $B = C$; dann ergibt sich $R = S$ aus Satz 5.7.3. □

5.7.5. [Beispiele]

(a) In $\mathbb{R}^2$ betrachte die zwei Geraden g und h, die durch die Gleichungen

$$g\colon x_1 + x_2 = 0$$
$$h\colon x_1 - 2x_2 = 0$$

gegeben werden. Sei $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ die Projektion auf g längs der Geraden h. **Problem:** Finde die Matrix A von φ bzgl. der Standardbasis $B = (e_1, e_2)$.



Wie wir in Beispiel 5.6.4(b) gesehen haben, ergibt sich bei Wahl einer geeigneten Basis $B' = (e_1', e_2')$ (mit e_1' in Richtung von g und e_2' in Richtung von h) die Matrix von φ bzgl. B' zu

$$A' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Im vorliegenden Fall wählen wir etwa $e_1' = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $e_2' = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Nach der Transformationsformel von 5.7.4 erhalten wir: $A = SA'S^{-1}$, wobei

$$S = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$$

die Übergangsmatrix ist. Wir berechnen:

$$S^{-1} = \frac{1}{3} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}.$$

Mithin

$$A = SA'S^{-1} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$$
$$= \frac{1}{3} \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}.$$

(b) In $\mathbb{R}^2$ betrachten wir die Basen B, die aus $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ besteht, und B', die aus $e_1' = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $e_2' = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ besteht, wie im vorigen Beispiel. Betrachten wir den Endomorphismus $\varphi = \varphi_A \colon \mathbb{R}^2 \to \mathbb{R}^2$, der durch die Matrix

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

bzgl. der Standardbasis B gegeben ist.

Problem: Finde die Matrix A' von φ bzgl. der Basis B'.

Nach 5.7.4 ist $A' = S^{-1}AS$, wobei S und S^{-1} wie in 5.7.5(a) sind. Also

$$\begin{aligned} A' &= \frac{1}{3}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \frac{1}{3}\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 12 \\ -3 & 6 \end{pmatrix} \\ &= \frac{1}{9}\begin{pmatrix} 9 & 0 \\ 0 & 18 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}. \end{aligned}$$

Da A' eine Diagonalmatrix ist, sieht man, wie die Abbildung φ auf $\mathbb{R}^2$ operiert:

$$\varphi(e_1') = e_1', \ \varphi(e_2') = 2e_2'.$$

Die Basis B' ist der Operation von φ viel besser angepasst als die Basis B. Im Sommersemster werden wir lernen, wie man eine Basis B' finden kann, bzgl. der ein gegebener Endomorphismus φ eine einfache Gestalt bekommt.

5.8 Lineare Gleichungssysteme

Wie in Kapitel 2 betrachten wir jetzt ein System von m linearen Gleichungen in n Unbekannten $x_1, \ldots, x_n$:

$$\begin{array}{ccccccccc} a_{11}x_1 &+& a_{12}x_2 &+\ldots+& a_{1j}x_j &+\ldots+& a_{1n}x_n &=& b_1 \\ a_{21}x_1 &+& a_{22}x_2 &+\ldots+& a_{2j}x_j &+\ldots+& a_{2n}x_n &=& b_2 \\ \vdots && \vdots && \vdots && \vdots && \vdots \\ a_{i1}x_1 &+& a_{i2}x_2 &+\ldots+& a_{ij}x_j &+\ldots+& a_{in}x_n &=& b_i \\ \vdots && \vdots && \vdots && \vdots && \vdots \\ a_{m1}x_1 &+& a_{m2}x_2 &+\ldots+& a_{mj}x_j &+\ldots+& a_{mn}x_n &=& b_m. \end{array} \tag{L}$$

Die Koeffizienten a_{ij} und die Konstanten b_i auf der rechten Seite mögen jetzt dem Körper $\mathbb{K}$ angehören.

Betrachten wir die Koeffizientenmatrix A und die Vektoren x, b:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(\mathbb{K}), \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Dieses System (L) kann in der Form

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

geschrieben werden. Ist $b \neq 0$, heißt es *inhomogen*; das System

$$Ax = 0 \tag{H}$$

heißt das zu (L) gehörige *homogene* System.

Aufgabe ist es, die Menge aller Lösungen von (L) zu bestimmen, d.h. die Menge aller $x \in \mathbb{K}^n$, für die (L) erfüllt ist, zu finden und explizit zu beschreiben.

5.8.1. [Lösungen des homogenenen Systems (H)] Wie wir in 5.5.12 gesehen haben, ist die Menge aller Lösungen von (H) ein Untervektorraum U von $\mathbb{K}^n$ der Dimension $k = n - r$, wobei $r := \operatorname{rank} A$ ist.

Sei $v_1, \ldots, v_k$ eine Basis dieses Untervektorraums U der Lösungen von (H). Wir nennen $v_1, \ldots, v_k$ auch ein *System von Fundamentallösungen* von (H). Die „allgemeine“ Lösung von (H) ist dann

$$x_H = \lambda_1 v_1 + \cdots \lambda_k v_k,$$

mit beliebigen $\lambda_i \in \mathbb{K}$.

5.8.2. [Existenz von Lösungen für (L)] Das homogene System (H) hat stets mindestens eine Lösung, nämlich $x = 0$. Man nennt sie die *triviale* Lösung. Ein inhomogenes System (L) kann dagegen durchaus unlösbar sein. Aus 5.5.2 wissen wir, dass

$$Ax = b \tag{L}$$

genau dann mindestens eine Lösung x hat, wenn b in dem von den Spalten $s_1, \ldots, s_n$ von A aufgespannten Untervektorraum des $\mathbb{K}^n$ liegt, d.h. genau dann, wenn

$$\operatorname{rank} A \quad = \quad \operatorname{rank}(A \mid b)$$

wo $(A \mid b)$ die $m \times (n+1)$-Matrix ist, die man aus A durch Hinzufügen von b als letzter Spalte erhält:

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right)$$

5.8.3. [Die Lösungen des inhomogenen Systems (L)] Angenommen, (L) hat wenigstens eine Lösung $x_S \in \mathbb{K}^n$ (eine *spezielle* Lösung, wie man sagt). Dann bildet die Menge *aller* Lösungen einen affinen Unterraum von $\mathbb{K}^n$, nämlich $x_S + U$, wobei U die Lösungsmenge des zugehörigen homogenen Systems (H) ist. Die Dimension dieses affinen Unterraums ist $k = n - r$ mit $r = \operatorname{rank} A$. In der Tat folgt nämlich aus $Ax_S = b$, dass $Ax = b$ zu $A(x - x_S) = 0$ äquivalent ist, und damit zur Tatsache, dass $x_H := x - x_S$ Lösung des homogenen Systems ist.

Die allgemeine Lösung von (L) kann also als

$$\boxed{x = x_S + x_H = x_S + \lambda_1 v_1 + \ldots + \lambda_k v_k, \quad \lambda_i \in \mathbb{K},}$$

geschrieben werden, wobei x_S eine spezielle Lösung von (L) und $v_1, \ldots, v_k$ ein System von Fundamentallösungen von (H) ist.

Beschränken wir uns nun auf den Fall $n = m$. Dann ist also A eine quadratische Matrix.

Satz 5.8.4. (Eindeutigkeit der Lösungen) *Für eine $n \times n$-Matrix A sind die folgenden Bedingungen äquivalent:*

(1) *Es existiert ein $b \in \mathbb{K}^n$, so dass $Ax = b$ genau eine Lösung x_0 besitzt.*

(2) *A ist invertierbar.*

(3) $\operatorname{rank} A = n$ *(d.h., die Spalten von A sind linear unabhängig.)*

Beweis. Die Äquivalenz von (2) und (3) folgt aus 5.5.11(4) (Surjektivitätskriterium).

(1) $\Rightarrow$ (3): Hat $Ax = b$ genau eine Lösung x_0, so ist die Lösungsmenge von $Ax = b$ null-dimensional; nach 5.8.3 bedeutet dies $0 = k = n - r$, d.h. $n = r = \operatorname{rank} A$.

(2) $\Rightarrow$ (1): Ist A invertierbar, so folgt aus $Ax = b$ auch $x = A^{-1}b$, d.h. $A^{-1}b$ ist die einzige Lösung von $Ax = b$. □

Korollar 5.8.5. *Für eine $n \times n$-Matrix A sind die folgenden Eigenschaften äquivalent:*

(1) *Das homogene System $Ax = 0$ hat eine Lösung $x \neq 0$ (eine solche Lösung nennt man* nichttrivial*)*

(2) *A ist nicht invertierbar.*

(3) $\operatorname{rank} A < n$ *(d.h. die Spalten von A sind linear abhängig.)*

Beweis. Die Eigenschaft (1) bedeutet, dass für kein $b \in \mathbb{K}^n$ die Lösung von $Ax = b$ eindeutig ist. Das Korollar folgt also aus dem vorhergehenden Satz durch Negation aller Bedingungen. □

Bemerkung 5.8.6. Auf den letzten zwei Seiten haben wir die grundsätzliche Beschreibung der Lösungsmenge eines linearen Gleichungssystems zusammengefasst. Wir wiederholen hier noch einmal die Methode, um solche Systeme explizit zu lösen:

Gegeben ein lineares Gleichungssystem (L) $Ax = b$, fragen wir:

- Hat es überhaupt eine Lösung? Falls ja, müssen wir eine spezielle Lösung x_S finden.

- Finde ein System von Fundamentallösungen $v_1, \dots, v_k$ des zugehörigen homogenen Gleichungssystems (H) $Ax = 0$, wobei $k = n - \operatorname{rank}(A)$.

Wie in Kapitel 2, werden beide Probleme durch den Gauß–Jordanschen Eliminationsalgorithmus gelöst:
Bilde die erweiterte Matrix

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right).$$

Durch elementaren Zeilentransformationen wird diese Matrix in Stufenform gebracht:

$$\left(\begin{array}{cccccccc|c} c_{1j_1} & * & * & * & \dots & \dots & \dots & * & \delta_1 \\ & & c_{2j_2} & * & \dots & \dots & \dots & * & \delta_2 \\ & & & \ddots & & & & \vdots & \vdots \\ & & & & c_{rj_r} & * & \dots & * & \delta_r \\ & 0 & & & & & & & \delta_{r+1} \\ & & & & & & & & \vdots \\ & & & & & & & & \delta_m \end{array}\right),$$

wobei $j_1 < j_2 < \ldots < j_r$ und $c_{ij_i} \neq 0$ für $i = 1, \dots, r$.

Erster Fall: Einer der Skalare $\delta_{r+1}, \dots, \delta_m$ ist $\neq 0$. Dann ist das System (L) $Ax = b$ unlösbar (weil widersprüchlich).

Zweiter Fall: $\delta_{r+1} = \ldots = \delta_m = 0$. Dann besitzt das System (L) $Ax = b$ Lösungen. Eine spezielle Lösung x_S finden wir z.B. dadurch, dass wir alle nicht-Pivot-Variablen $x_i = 0$ setzen ($i \neq j_1, \dots, j_r$).

Nunmehr betrachte das zugehörige homogene System (H) $Ax = 0$. Um ein System von Fundamentallösungen $v_1, \dots, v_k$ hierfür zu finden, betrachtet man das auf Stufenform umgeformte Gleichungssystem, mit $\delta_1 = \ldots =$

$\delta_r = 0$ und löst dieses für die k verschiedenen Fälle, die sich dadurch ergeben, dass man für die nicht-Pivot-Variablen x_i $(i \neq j_1, \ldots, j_r)$ nacheinander einsetzt:

$$\begin{array}{ccccc} 1, & 0, & 0, & \ldots, & 0 \\ 0, & 1, & 0, & \ldots, & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0, & \ldots & \ldots & 0, & 1 \end{array}.$$

Beispiel 5.8.7. Für $\mathbb{K} = \mathbb{R}$ betrachte das System

$$\begin{array}{rcrcrcrcrcr} x_1 & + & 2x_2 & - & x_3 & + & 2x_4 & + & 3x_5 & = & 1 \\ 2x_1 & + & 4x_2 & - & 2x_3 & + & 5x_4 & + & 5x_5 & = & 3 \\ -x_1 & - & 2x_2 & - & x_3 & + & x_4 & + & x_5 & = & 1 \\ 2x_1 & + & 4x_2 & & & + & 2x_4 & + & x_5 & = & 1 \\ 3x_1 & + & 6x_2 & - & 3x_3 & + & 7x_4 & + & 8x_5 & = & 4 \end{array}$$

Gauß–Jordan Elimination liefert: $r = 3$, $k = 5 - 3 = 2$.

1. Das System ist lösbar.
2. Die Pivot-Variablen sind x_1, x_3, x_4. Für eine spezielle Lösung x_S setze $x_2 = x_5 = 0$ im inhomogenen System in Stufenform. Das führt auf
$$x_S = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{pmatrix}.$$
3. Für ein System von Fundamentallösungen v_1, v_2 des homogenen Systems setze
$$x_2 = 1, x_5 = 0: \qquad v_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
$$x_2 = 0, x_5 = 1: \qquad v_2 = \begin{pmatrix} -\frac{3}{2} \\ 0 \\ \frac{7}{2} \\ 1 \\ 1 \end{pmatrix}.$$

$$
\begin{array}{|rrrrr|r|}
\hline
1 & 2 & -1 & 2 & 3 & 1 \\
2 & 4 & -2 & 5 & 5 & 3 \\
-1 & -2 & -1 & 1 & 1 & 1 \\
2 & 4 & 0 & 2 & 1 & 1 \\
3 & 6 & -3 & 7 & 8 & 4 \\
\hline
1 & 2 & -1 & 2 & 3 & 1 \\
0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & -2 & 3 & 4 & 2 \\
0 & 0 & 2 & -2 & -5 & -1 \\
0 & 0 & 0 & 1 & -1 & 1 \\
\hline
1 & 2 & -1 & 2 & 3 & 1 \\
0 & 0 & -2 & 3 & 4 & 2 \\
0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & 2 & -2 & -5 & -1 \\
0 & 0 & 0 & 1 & -1 & 1 \\
\hline
1 & 2 & -1 & 2 & 3 & 1 \\
0 & 0 & -2 & 3 & 4 & 2 \\
0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & 0 & 1 & -1 & 1 \\
\hline
1 & 2 & -1 & 2 & 3 & 1 \\
0 & 0 & -2 & 3 & 4 & 2 \\
0 & 0 & 0 & 1 & -1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\hline
\end{array}
$$

Die allgemeine Lösung ist also von der Form:

$$
x = x_S + \lambda_1 v_1 + \lambda_2 v_2 = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -\frac{3}{2} \\ 0 \\ \frac{7}{2} \\ 1 \\ 1 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R}.
$$

Kapitel 6

Spur und Determinante

In diesem Kapitel führen wir zwei Abbildungen ein, die jeder quadratischen Matrix, und von da aus letztlich auch jedem Endomorphismus eines endlich dimensionalen Vektorraums, einen Skalar zuordnen: die Spur und die Determinante. Dass diese beiden numerischen Invarianten von Matrizen und Endomorphismen tatsächlich eng zusammenhängen, werden wir im zweiten Semester sehen: es handelt sich nämlich bei Spur und Determinante um den höchsten, beziehungsweise den konstanten Koeffizienten des *charakteristischen Polynoms* der Matrix oder des Endomorphismus. Dieser Zusammenhang ist aber aus der Definition nicht offensichtlich, und so zerfällt dieses Kapitel 6 in zwei unabhängige Teile: einen kurzen ersten Abschnitt über die Spur einerseits, und die späteren Abschnitte andererseits, in denen die Determinante eingeführt wird. Die Einführung der Determinante benötigt ihrerseits den Begriff der Signatur einer Permutation, der in Abschnitt 6.2 kurz abgehandelt wird.

Wir arbeiten im ganzen Kapitel wie gewohnt über einem beliebigen Körper $\mathbb{K}$. Die Spur ist eine $\mathbb{K}$-*lineare* Funktion

$$\mathrm{tr}\colon M_n(\mathbb{K}) \to \mathbb{K},$$

die *invariant unter Konjugation* ist, d.h. für jede invertierbare Matrix $T \in \mathrm{GL}_n(\mathbb{K})$ und jede Matrix $A \in M_n(\mathbb{K})$ gilt:

$$\mathrm{tr}(TAT^{-1}) = \mathrm{tr}(A).$$

Diese Invarianz erlaubt es, jedem Endomorphismus eines endlich dimensionalen $\mathbb{K}$-Vektorraums eine wohldefinierte Spur zuzuordnen. Der Begriff der Spur einer quadratischen Matrix hat übrigens auch eine wichtige geometrische Interpretation in der Theorie der dynamischen Systeme, als infinitesimale Volumenänderung. Außerdem ist er grundlegend für die sogenannte Darstellungstheorie von Gruppen. Diese Perspektiven können wir aber im Rahmen dieser Vorlesung nicht entwickeln.

Die Determinante hingegen ist eine *multiplikative* und, ebenso wie die Spur, *konjugationsinvariante* Funktion

$$\det\colon M_n(\mathbb{K}) \to \mathbb{K},$$

d.h. es gilt

$$\det(AB) = \det(A)\det(B).$$

Eine der am häufigsten benutzten Eigenschaften der Determinante ist die, dass eine quadratische Matrix genau dann invertierbar ist, wenn ihre Determinante nicht null ist. Eine anderes mitunter nützliches Ergebnis (das freilich für große Matrizen numerisch nicht gut brauchbar ist) ist die als *Cramersche Regel* bekannte explizite Formel für die Lösungen eindeutig lösbarer quadratischer Gleichungssysteme, und damit auch für die Inverse einer invertierbaren Matrix. *Last but not least*, besitzt die Determinante für $\mathbb{K} = \mathbb{R}$ die geometrische Interpretation als Volumen des von n Vektoren im $\mathbb{R}^n$ aufgespannten Parallelepipeds.

6.1 Die Spur

Definition 6.1.1. [Spur einer quadratischen Matrix] Sei $A = (a_{ij}) \in M_n(\mathbb{K})$. Die Summe der Diagonalelemente der Matrix: $\sum_{i=1}^n a_{ii}$ heißt die *Spur* [auf Englisch *trace*] von A. Oder in Formeln:

$$\operatorname{tr} A = \operatorname{tr}\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} := a_{11} + a_{22} + \ldots + a_{nn} \in \mathbb{K}$$

Beispiele 6.1.2. (a) $\operatorname{tr} E_n = n$.

(b) $\operatorname{tr}\begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 + 2\cos\gamma.$

Satz 6.1.3. *Für beliebige $n \times n$-Matrizen A, B und $\lambda \in \mathbb{K}$ gilt:*

$$\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B \tag{T1}$$

$$\operatorname{tr}(\lambda A) = \lambda \operatorname{tr} A \tag{T2}$$

$$\operatorname{tr}(AB) = \operatorname{tr}(BA). \tag{T3}$$

Beweis. Die ersten beiden Eigenschaften sind offensichtlich, da die linearen Operationen auf Matrizen koeffizientenweise definiert sind. Beachte, dass sie besagen: *die Spur ist eine lineare Abbildung* $M_n(\mathbb{K}) \to \mathbb{K}$. Für den Beweis von (T3) seien $A = (a_{ij})$, $B = (b_{ij})$ und $C := AB$, $C' := BA$. Dann ist

$$\operatorname{tr}(AB) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji}$$

$$\operatorname{tr}(BA) = \sum_{i=1}^n c'_{ii} = \sum_{i=1}^n \sum_{j=1}^n b_{ij} a_{ji} = \sum_{j=1}^n \sum_{i=1}^n a_{ji} b_{ij} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji} = \operatorname{tr}(AB)$$

$\square$

Bemerkung 6.1.4. Beachte, dass im Allgemeinen $\operatorname{tr}(AB) \neq \operatorname{tr} A \cdot \operatorname{tr} B$ gilt. Ist z.B. $A = B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, so gilt $\operatorname{tr} A = \operatorname{tr} B = 0$ und somit auch $\operatorname{tr} A \cdot \operatorname{tr} B = 0$, aber $\operatorname{tr}(A \cdot B) = \operatorname{tr} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2$.

Satz 6.1.5. *Für jeder invertierbare $n \times n$-Matrix S und jede $n \times n$-Matrix A, gilt*

$$\boxed{\operatorname{tr}(S^{-1}AS) = \operatorname{tr} A}$$

Beweis. $\operatorname{tr}(S^{-1}AS) = \operatorname{tr}((S^{-1}A)S) \overset{(T3)}{=} \operatorname{tr}(S(S^{-1}A)) = \operatorname{tr}((S^{-1}S)A) = \operatorname{tr} A$. $\square$

Definition 6.1.6. [Spur eines Endomorphismus] Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $\varphi : V \to V$ eine lineare Abbildung. Betrachte die Matrix $[\varphi]_B$ von φ bzgl. einer Basis B von V und die Matrix $[\varphi]_{B'}$ von φ bzgl. einer anderen Basis B' von V.

Nach (5.6.3) gibt es eine invertierbare Matrix $S = [\mathrm{id}]_B^{B'}$ (die Übergangsmatrix) so dass gilt $[\varphi]_{B'} = S^{-1}[\varphi]_B S$. Nach 6.1.5 schließen wir, dass

$$\operatorname{tr}[\varphi]_{B'} = \operatorname{tr}[\varphi]_B,$$

d.h., die Spur der Matrix ist unter Basiswechsel invariant, oder anders gesagt: die Spur der Matrix $[\varphi]_B$ hängt nur vom Endomorphismus φ und nicht von der herangezogenen Basis B ab.

So könne wir also die *Spur von* φ als Spur der φ bzgl. einer beliebigen Basis zugeordneten Matrix B von V definieren:

$$\operatorname{tr} \varphi := \operatorname{tr}([\varphi]_B).$$

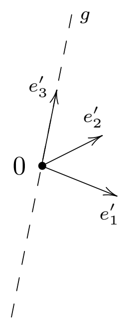
Beispiele 6.1.7. (a) $\operatorname{tr}(\mathrm{id}_V) = n$ $(= \dim V)$

Denn die Matrix von id_V (sogar bzgl. irgendeiner Basis von V) ist E_n und $\operatorname{tr} E_n = n$.

(b) In $\mathbb{R}^3$ sei δ die Drehung um die Achse g durch den Ursprung, um den Winkel γ.

Wri behaupten

$$\operatorname{tr} \delta = 1 + 2\cos\gamma$$



In der Tat könne wir etwa folgende Basis B' von $\mathbb{R}^3$ benutzen:
Wähle zwei zueinander orthogonale Vektoren e_1', e_2', die beide auch orthogonal zur Gerade g sind, und wähle e_3' in Richtung g. Weiterhin normalisiere e_1', e_2', e_3' (durch Multiplikation mit geeigneten reellen Skalaren) so, dass jeder der dieser Vektoren Länge 1 hat (Beispiel 5.5.18).

Bezüglich dieser Basis B' ist der Abbildung δ die Matrix

$$[\delta]_{B'} = A' = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

zugeordnet. Also ist $\operatorname{tr}\delta = \operatorname{tr}A' = 1 + 2\cos\gamma$.

Ist uns nur die Matrix $A = (a_{ij})$ von δ bzgl. der Standardbasis B bekannt, nicht aber der Drehwinkel γ, so können wir damit $\cos\gamma$ finden:

$$\cos\gamma = \frac{1}{2}(\operatorname{tr}\delta - 1) = \frac{1}{2}(\operatorname{tr}A - 1) = \frac{1}{2}(a_{11} + a_{22} + a_{33} - 1).$$

6.2 Die Signatur einer Permutation

Seit Beispiel 3.2.10 wissen wir, dass eine *Permutation* der n-elementigen Menge $\{1, 2, \ldots, n\}$ eine bijektive Abbildung

$$\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$$

ist. Wir schreiben $\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$. Mit S_n bezeichnen wir die Menge aller Permutationen σ der n-elementigen Menge $\{1, 2, \ldots, n\}$ und mit T_n die Menge aller Selbstabbildungen dieser Menge.

Für eine ganze Zahl k definieren wir

$$\operatorname{sgn}(k) := \begin{cases} 1 & \text{für } k > 0 \\ 0 & \text{für } k = 0 \\ -1 & \text{für } k < 0. \end{cases}$$

Für ein n-Tupel $j = (j_1, \ldots, j_n) \in \mathbb{Z}^n$ setzen wir nun

$$\operatorname{sgn}(j) := \prod_{p<q} \operatorname{sgn}(j_q - j_p),$$

wobei sich das Produkt über alle Paare $(p, q) \in \{1, \ldots, n\}^2$ mit $p < q$ erstreckt.

Definition 6.2.1. (a) Für eine Abbildung $\sigma \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, also $\sigma \in T_n$, setzen wir

$$\operatorname{sgn}(\sigma) := \operatorname{sgn}(\sigma(1), \ldots, \sigma(n)) = \prod_{i<j} \operatorname{sgn}(\sigma(j) - \sigma(i)).$$

(b) Sei $\sigma \in S_n$. Jedes Ziffernpaar (i, j) mit $i < j$ und $\sigma(i) > \sigma(j)$ nennt man eine *Inversion* von σ. Wir schreiben

$$I(\sigma) = |\{(i, j) \in \{1, \ldots, n\}^2 \colon i < j, \sigma(i) > \sigma(j)\}|$$

für die Anzahl der Inversionen von σ.

In der Formel für die Signatur erhalten wir für jede Inversion einen Faktor -1 und für alle anderen Paare einen Faktor 1, also

$$\operatorname{sgn}(\sigma) = (-1)^{I(\sigma)}.$$

Ist $I(\sigma)$ gerade, so heißt σ eine *gerade* Permutation; ist $I(\sigma)$ ungerade, so heißt σ eine *ungerade* Permutation. Dann gilt also

$$\operatorname{sgn}(\sigma) = \begin{cases} +1 & \text{falls } \sigma \text{ eine gerade Permutation ist;} \\ -1 & \text{falls } \sigma \text{ eine ungerade Permutation ist.} \end{cases}$$

Lemma 6.2.2. *Die Signatur*

$$\operatorname{sgn}\colon T_n \to (\{0, 1, -1\}, \cdot)$$

is ein Monoidhomomorphismus und

$$S_n = \{\sigma \in T_n \colon \operatorname{sgn}(\sigma) \neq 0\},$$

d.h., eine Selbstabbildung σ von $\{1, \dots, n\}$ ist genau dann invertierbar, wenn $\operatorname{sgn}(\sigma) \neq 0$ ist. Insbesondere ist

$$\operatorname{sgn}\colon S_n \to (\{\pm 1\}, \cdot)$$

ein Gruppenhomomorphismus.

Beweis. Man sieht sofort, dass $\operatorname{sgn}(\sigma)$ genau dann von 0 verschieden ist, wenn σ injektiv ist. Wegen Satz 1.3.21 ist das äquivalent zur Bijektivität, also $\sigma \in S_n$.

Ist $\pi\colon \{1, \dots, n\} \to \{1, \dots, n\}$ nicht injektiv bzw. bijektiv, so ist

$$\operatorname{sgn}(\sigma)\operatorname{sgn}(\pi) = 0 = \prod_{i<j} \operatorname{sgn}(\sigma(\pi(j)) - \sigma(\pi(i))) = \operatorname{sgn}(\sigma\pi),$$

da mindestens ein Faktor auf der rechten Seite verschwindet. Ist $\pi \in S_n$, so erhalten wir

$$\begin{aligned}
&\operatorname{sgn}(\sigma\pi) = \prod_{i<j} \operatorname{sgn}(\sigma(\pi(j)) - \sigma(\pi(i))) \\
&= \prod_{\substack{i<j \\ \pi(i)<\pi(j)}} \operatorname{sgn}(\sigma(\pi(j)) - \sigma(\pi(i))) \cdot \prod_{\substack{i<j \\ \pi(j)<\pi(i)}} \operatorname{sgn}(\sigma(\pi(j)) - \sigma(\pi(i))) \\
&= \prod_{\substack{i<j \\ \pi(i)<\pi(j)}} \operatorname{sgn}(\sigma(\pi(j)) - \sigma(\pi(i))) \cdot \operatorname{sgn}(\pi) \cdot \prod_{\substack{i<j \\ \pi(j)<\pi(i)}} \operatorname{sgn}(\sigma(\pi(i)) - \sigma(\pi(j))) \\
&= \prod_{\substack{i<j \\ \pi(i)<\pi(j)}} \operatorname{sgn}(\sigma(\pi(j)) - \sigma(\pi(i))) \cdot \operatorname{sgn}(\pi) \cdot \prod_{\substack{i>j \\ \pi(i)<\pi(j)}} \operatorname{sgn}(\sigma(\pi(j)) - \sigma(\pi(i))) \\
&= \prod_{k<\ell} \operatorname{sgn}(\sigma(\ell) - \sigma(k)) \cdot \operatorname{sgn}(\pi) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi).
\end{aligned}$$

□

Bemerkung 6.2.3. (a) Da sgn auf S_n ein Gruppenhomomorphismus ist und jedes Element der Gruppe $(\{\pm 1\}, \cdot)$ seinem eigenen Inversen gleich ist, gilt

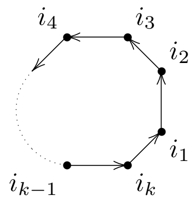
$$\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)^{-1} = \operatorname{sgn}(\sigma)$$

für jedes $\sigma \in S_n$. Insbesondere ist also σ genau dann gerade (bzw. ungerade), wenn σ^{-1} gerade (bzw. ungerade) ist.

(b) Das Produkt zweier gerader, ebenso wie das Produkt zweier ungerader Permutationen ist gerade. Das Produkt einer geraden und einer ungeraden Permutation, in irgendeiner Reihenfolge, ist ungerade.

Zykelzerlegung und Signatur

Definition 6.2.4. Eine Permutation $\sigma \in S_n$ heißt *Zyklus der Länge* k, oder einfach k-*Zyklus*, fall es Ziffern $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$ gibt, welche zyklisch permutiert werden, während σ alle anderen Ziffern festlässt. In Formeln:



$$\begin{aligned}&\sigma(i_1) = i_2,\ \sigma(i_2) = i_3,\ \ldots \sigma(i_{k-1}) = i_k,\\ &\sigma(i_k) = i_1,\\ &\sigma(j) = j \text{ für alle } j \notin \{i_1, i_2, \ldots, i_k\}.\end{aligned}$$

Wir schreiben

$$\sigma = (i_1\ i_2\ \cdots\ i_k)$$

für den Zyklus, der genau die $i_1, \ldots, i_k$ zyklisch permutiert.

Es gibt nur einen 1-Zyklus, die Identität.

Ein Zyklus der Länge 2 heißt *Transposition*. Wir schreiben mitunter auch $\tau_{ij} := (i\ j)$ für die Transposition, die i und j vertauscht, und alle anderen Ziffern festhält.

Beispiel 6.2.5. In S_5 ist das Element

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

ein Zyklus der Länge 5, und

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

ist ein Zyklus der Länge 3. Die Identität

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = (1) = (2) = \cdots = (5)$$

ist ein Zyklus der Länge 1.

Satz 6.2.6. *Jede Permutation $\sigma \in S_n$ $(n \in \mathbb{N})$ ist Produkt von $\leq n-1$ Transpositionen.*

Beweis. Für jedes $n \in \mathbb{N}$ identifizieren wir die Gruppe S_{n-1} mit der Untergruppe

$$\{\sigma \in S_n \colon \sigma(n) = n\}$$

von S_n. Das ermöglicht uns den Beweis des Satzes durch vollständige Induktion über n. Für $n = 1$ ist nichts zu beweisen, da $S_1 = \{(1)\}$ nur aus einem 1-Zyklus besteht.

Nehmen wir jetzt an, die Behauptung gelte für S_{n-1}. Sei $\sigma \in S_n$ gegeben. Gilt $\sigma(n) = n$, dann liegt σ in S_{n-1}, und die Induktionsvoraussetzung garantiert uns, dass σ ein Produkt von höchstens $n-1$ Transpositionen ist.

Ist aber $\sigma(n) \neq n$, so lässt die Permutation $\tau_{n,\sigma(n)}\sigma$ die Ziffer n fest, und ist daher nach Induktionsvoraussetzung ein Produkt von maximal $n-2$ Transpositionen:

$$\tau_{n,\sigma(n)}\sigma = \tau_1 \cdots \tau_k.$$

Dann folgt aber

$$\sigma = \tau_{n,\sigma(n)}\tau_{n,\sigma(n)}\sigma = \tau_{n,\sigma(n)}\tau_1 \cdots \tau_k,$$

so dass also auch σ ein Produkt von Transpositionen ist. □

Beispiel 6.2.7. Ein Zyklus $\sigma = (i_1\ i_2\ \cdots\ i_k)$ der Länge k kann als Produkt von $(k-1)$ Transpositionen dargestellt werden:

$$\sigma = (i_1\ i_2\ \cdots\ i_k) = (i_1\ i_2)(i_2\ i_3)\cdots(i_{k-1}\ i_k).$$

Für den Zyklus (1 2 4 5) zum Beispiel gilt:

$$(1\ 2\ 4\ 5) = (1\ 2)(2\ 4)(4\ 5).$$

Die Beschreibung einer Permutation als Produkt von Transpositionen ist ganz und gar nicht eindeutig. Zum Beispiel gilt offenbar auch:

$$(1\ 2\ 4\ 5) = (2\ 4\ 5\ 1) = (2\ 4)(4\ 5)(5\ 1).$$

Beispiel 6.2.8. (a) Die Identität hat keine Inversionen, d.h. $I(\mathrm{id}) = 0$ und somit $\mathrm{sgn}(\mathrm{id}) = 1$.
(b) Jede Transposition is ungerade, denn für $i < j$ ist die menge der Inversionen der Transposition $\tau_{ij} = (i\ j)$:

$$(i, i+1), (i, i+2), \ldots, (i, j), (i+1, j), (i+2, j), \ldots, (j-1, j),$$

womit

$$I((i\ j)) = (j-i) + (j-i-1) = 2(j-i) - 1$$

folgt, was eine ungerade Zahl ist.

Bemerkung 6.2.9. Da jede Permutation σ ein Produkt von Transpositionen ist:

$$\sigma = \tau_1 \cdots \tau_k,$$

folgt aus Lemma 6.2.2:

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1) \cdots \operatorname{sgn}(\tau_k) = (-1)^k.$$

Eine Permutation ist also genau dann gerade (bzw. ungerade), wenn sie Produkt einer geraden (bzw. ungeraden) Anzahl von Transpositionen ist. So ergibt sich auch die Tatsache, dass —bei aller Nicht-Eindeutigkeit der Darstellung als Produkt von Transpositionen— eine Permutation nie gleichzeitig Produkt einer geraden und einer ungeraden Anzahl von Transpositionen sein kann.

Übungen zu Abschnitt 6.2

Übung 6.2.10. (Inversionen)

(a) Zeige, dass die Zahl $I(\sigma)$ von Inversionen einer Permutation $\sigma \in S_n$ höchstens $n(n-1)/2$ ist.

(b) Finden Sie eine Permutation, für die diese Maximalzahl angenommen wird und zeigen Sie, dass sie eindeutig ist.

(c) Ist $\tau = (i, i+1)$ eine Transposition benachbarter Elemente, so gilt für jedes Element $\sigma \in S_n$

$$I(\sigma\tau) = I(\sigma) + 1 \quad \text{oder} \quad I(\sigma\tau) = I(\sigma) - 1.$$

Wovon hängt es ab welcher Fall eintritt?

6.3 Alternierende multilineare Abbildungen

Definition 6.3.1. Seien V und W Vektorräume. Eine Funktion $\Delta\colon V^n \to W$ heißt

(A1) *n-linear* oder *multilinear*, wenn für jedes $j \in \{1, \ldots, n\}$ und feste Elemente $v_i \in V$, $i \neq j$, die Abbildung

$$V \to W, \quad v \mapsto \Delta(v_1, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_n)$$

linear ist. Mit anderen Worten, für jedes j gilt

$$\Delta(\ldots, \lambda v_j + \mu v_j', \ldots) = \lambda\Delta(\ldots, v_j, \ldots) + \mu\Delta(\ldots, v_j', \ldots).$$

Für $n = 2$ spricht man von *bilinearen* Abbildungen und für $n = 3$ von *trilinearen* Abbildungen.

(A2) *alternierend*, wenn aus $v_i = v_j$ für zwei Indizes $i \neq j$ schon $\Delta(v_1, \ldots, v_n) = 0$ folgt.

Beispiel 6.3.2. (a) Die Multiplikationsabbildung

$$\mu\colon \mathbb{K}^n \to \mathbb{K}, \quad (x_1, \ldots, x_n) \mapsto x_1 \cdots x_n$$

ist n-linear. Sie ist nicht alternierend, da $1^n = 1 \neq 0$ ist.

(b) Für $p_1, \ldots, p_{n+1} \in \mathbb{N}$ ist die Multiplikationsabbilding

$$\mu\colon M_{p_1,p_2}(\mathbb{K}) \times \cdots \times M_{p_n,p_{n+1}}(\mathbb{K}) \to M_{p_1,p_{n+1}}(\mathbb{K}), \quad (A_1, \ldots, A_n) \mapsto A_1 \cdots A_n$$

eine n-lineare Abbildung. Für $p_1 = p_2 = \cdots = p_{n+1}$ erhalten wir die Abbildung unter (a).

(c) Für $n \in \mathbb{N}$ ist die Kommutatorklammer

$$\Delta\colon M_n(\mathbb{K})^2 \to M_n(\mathbb{K}), \quad (A, B) \mapsto AB - BA$$

eine bilineare alternierende Abbildung.

Es ist zunächst nicht klar, ob es außer der Nullabbildung noch andere alternierende n-lineare Abbildungen gibt. Wir werden aber bald sehen, dass dies für $V = \mathbb{K}^n$ der Fall ist und diese Abbildungen durch den Wert $\Delta(e_1, \ldots, e_n)$ auf den Basiselementen eindeutig bestimmt sind.

Lemma 6.3.3. *Für jede alternierende n-lineare Abbildung $\Delta\colon V^n \to W$ gilt:*

(1) *Für $i \neq j$ und $\lambda \in \mathbb{K}$ ist*

$$\Delta(v_1, \ldots, v_{j-1}, v_j + \lambda v_i, v_{j+1}, \ldots, v_n) = \Delta(v_1, \ldots, v_n).$$

(2) *Vertauscht man v_i und v_j, so multipliziert sich der Wert $\Delta(v_1, \ldots, v_n)$ mit -1, d.h. für $i < j$ ist*

$$\Delta(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) = -\Delta(v_1, \ldots, v_n).$$

(3) *Für jedes $\sigma \in T_n$ gilt*

$$\Delta(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \mathrm{sgn}(\sigma)\Delta(v_1, \ldots, v_n).$$

(4) *Für $w_j = \sum_{i=1}^n a_{ij} v_i$, $j = 1, \ldots, n$, gilt*

$$\Delta(w_1, \cdots, w_n) = \Big(\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}\Big)\Delta(v_1, \ldots, v_n).$$

Beweis. (1) Für $i < j$ gilt:

$$\begin{aligned} &\Delta(\ldots, v_i, \ldots, v_j + \lambda v_i, \ldots) = \\ &\overset{\text{(A1)}}{=} \Delta(\ldots, v_i, \ldots, v_j, \ldots) + \Delta(\ldots, v_i, \ldots, \lambda v_i, \ldots) \\ &\overset{\text{(A1)}}{=} \Delta(\ldots, v_i, \ldots, v_j, \ldots) + \lambda \cdot \underbrace{\Delta(\ldots, v_i, \ldots, v_i, \ldots)}_{= 0 \text{ wegen (A2)}}. \end{aligned}$$

(2) Für $i < j$ gilt:

$$\begin{aligned}
&\Delta(\ldots, v_i, \ldots, v_j, \ldots) + \Delta(\ldots, v_j, \ldots, v_i, \ldots) \\
\overset{(1)}{=}\ &\Delta(\ldots, v_i + v_j, \ldots, v_j, \ldots) + \Delta(\ldots, v_j + v_i, \ldots, v_i, \ldots) \\
\overset{(A1)}{=}\ &\Delta(\ldots, v_i + v_j, \ldots, v_j + v_i, \ldots) \overset{(A2)}{=} 0,
\end{aligned}$$

also $\Delta(\ldots, v_i, \ldots, v_j, \ldots) = -\Delta(\ldots, v_j, \ldots, v_i, \ldots)$.

(3) Ist $\sigma \notin S_n$, so ist σ nicht injektiv und aus (A2) folgt sofort

$$\Delta(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) \overset{(A2)}{=} 0 = \operatorname{sgn}(\sigma)\Delta(v_1, \ldots, v_n).$$

Sei $\sigma \in S_n$. Ist σ eine Transposition, so ergibt sich (3) aus (2). Ein beliebiges $\sigma \in S_n$ schreiben wir als Produkt von Transpositionen $\sigma = \tau_1 \cdots \tau_k$ (Lemma 6.1.3), und schließen mit vollständiger Induktion über die Anzahl k: Setze $\pi := \tau_1 \cdots \tau_{k-1}$, so dass $\sigma = \pi\tau_k$ ist. Nach Induktionsvoraussetzung und (2) haben wir dann:

$$\begin{aligned}
\Delta(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) &= (-1)\Delta(v_{\pi(1)}, \ldots, v_{\pi(n)}) \\
&= (-1)\operatorname{sgn}(\pi)\Delta(v_1, \ldots, v_n) = \operatorname{sgn}(\sigma)\Delta(v_1, \ldots, v_n).
\end{aligned}$$

(4) Aus der n-Linearität von Δ erhalten wir mit (3) unmittelbar

$$\begin{aligned}
\Delta(w_1, \cdots, w_n) &= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{j_1,1} \cdots a_{j_n,n} \Delta(v_{j_1}, \cdots, v_{j_n}) \\
&= \sum_{\sigma \in T_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \Delta(v_{\sigma(1)}, \cdots, v_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \operatorname{sgn}(\sigma)\Delta(v_1, \ldots, v_n),
\end{aligned}$$

da $\operatorname{sgn}(\sigma) = 0$ für alle $\sigma \in T_n \setminus S_n$ gilt (Lemma 6.2.2). □

6.4 Die Determinante einer $n \times n$-Matrix

Definition 6.4.1. Die Abbildung

$$\det\colon M_n(\mathbb{K}) \to \mathbb{K}, \quad \det A := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \ldots \cdot a_{\sigma(n),n}$$

heißt *Determinante*. Die obige Formel für die Determiannte einer Matrix $A \in M_n(\mathbb{K})$ heißt die *Leibnizsche Formel*. Wir benutzen auch folgende Schreibweise für die Determinante einer Matrix:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} := \det(A).$$

Wir betrachten $n \times n$-Matrizen A mit Einträgen a_{ij} aus einem beliebigen Körper $\mathbb{K}$ und bezeichnen die Spaltenvektoren von A mit $s_1, s_2, \dots, s_n \in \mathbb{K}^n$, d.h.

$$A = (s_1, s_2, \dots, s_n)$$

in dem Sinne, dass

$$A = \begin{array}{cccc} \overbrace{\phantom{a_{11}}}^{s_1} & \overbrace{\phantom{a_{12}}}^{s_2} & \dots & \overbrace{\phantom{a_{11}}}^{s_n} \\ \end{array}$$
$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{11} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Theorem 6.4.2. (Hauptsatz über Determinanten)

(1) *Die Determinante ist n-linear und alternierend als Funktion der Spalten, d.h., als Abbildung*

$$M_n(\mathbb{K}) \cong (\mathbb{K}^n)^n \to \mathbb{K},$$
$$(s_1, \dots, s_n) \mapsto \det(s_1, \dots, s_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) s_{1,\sigma(1)} \cdots s_{n,\sigma(n)}.$$

(2) *Ist $\Delta\colon (\mathbb{K}^n)^n \to \mathbb{K}$ eine n-lineare alternierende Abbildung, so ist*

$$\Delta(s_1, \dots, s_n) = \Delta(e_1, \dots, e_n) \det(A) \quad \textit{für} \quad A = (s_1, \dots, s_n).$$

(3) *Die Determinante ist die eindeutig bestimmte alternierende n-lineare Abbildung $\Delta\colon M_n(\mathbb{K}) \cong (\mathbb{K}^n)^n \to \mathbb{K}$, welche $\Delta(E_n) = 1$ erfüllt.*

Beweis. (1) Wir müssen zeigen, dass det (A1) und (A2) erfüllt. Für den Nachweis von (A1) rechnen wir für $\lambda, \mu \in \mathbb{K}$:

$$\begin{aligned}
&\det(s_1, \dots, s_{j-1}, \lambda s_j + \mu s'_j, s_{j+1}, \dots, s_n) \\
&= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(j-1),j-1} \cdot (\lambda a_{\sigma(j),j} + \mu a'_{\sigma(j),j}) \cdot a_{\sigma(j+1),j+1} \cdots a_{\sigma(n),n} \\
&= \lambda \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(j-1),j-1} \cdot a_{\sigma(j),j} \cdot a_{\sigma(j+1),j+1} \cdots a_{\sigma(n),n} \\
&+ \mu \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(j-1),j-1} \cdot a'_{\sigma(j),j} \cdot a_{\sigma(j+1),j+1} \cdots a_{\sigma(n),n} \\
&= \lambda \det(s_1, \dots, s_{j-1}, s_j, s_{j+1}, \dots, s_n) + \mu \det(s_1, \dots, s_{j-1}, s'_j, s_{j+1}, \dots, s_n).
\end{aligned}$$

Für den Nachweis von (A2) nehmen wir an, dass die beiden Spalten s_i und s_j $(i < j)$ gleich sind. Für $\sigma' := \sigma \circ (i, j)$ erhalten wir dann

$$\begin{aligned}
\prod_{k=1}^{n} a_{\sigma(k),k} &= a_{\sigma(i),i} a_{\sigma(j),j} \prod_{k \neq i,j}^{n} a_{\sigma(k),k} = a_{\sigma(i),j} a_{\sigma(j),i} \prod_{k \neq i,j}^{n} a_{\sigma(k),k} \\
&= a_{\sigma'(j),j} a_{\sigma'(i),i} \prod_{k \neq i,j}^{n} a_{\sigma'(k),k} = \prod_{k=1}^{n} a_{\sigma'(k),k}.
\end{aligned}$$

Nun ist $\operatorname{sgn}(\sigma') = \operatorname{sgn}(\sigma) \operatorname{sgn}((i,j)) = -\operatorname{sgn}(\sigma)$ und $\sigma'' = \sigma \circ (i,j)^2 = \sigma$. Durchläuft σ alle Permutationen negativer Signatur, so durchläuft σ' daher alle Permutationen positiver Signatur. Mit dieser Vorüberlegung ergibt sich

$$\begin{aligned}
\det(A) = \det(s_1, \dots, s_n) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{\sigma(k),k} \\
&= \sum_{\operatorname{sgn}(\sigma)=1} \prod_{k=1}^{n} a_{\sigma(k),k} - \sum_{\operatorname{sgn}(\sigma)=-1} \prod_{k=1}^{n} a_{\sigma(k),k} \\
&= \sum_{\operatorname{sgn}(\sigma)=1} \prod_{k=1}^{n} a_{\sigma(k),k} - \sum_{\operatorname{sgn}(\sigma')=1} \prod_{k=1}^{n} a_{\sigma'(k),k} = 0.
\end{aligned}$$

(2) Jede Spalte s_j von A kann in der Standardbasis des $\mathbb{K}^n$ geschrieben werden:

$$s_j = \sum_{i=1}^{n} a_{ij} e_i.$$

Mit Lemma 6.3.3(4) erhalten wir daher $\Delta(s_1, \dots, s_n) = \det(A) \Delta(e_1, \dots, e_n)$.

(3) Ist $A = E_n$ und $\prod_{i=1}^{n} a_{\sigma(i),i} \neq 0$, so ist $\sigma(i) = i$ für jedes i, also $\sigma = \mathrm{id}$. Damit ist $\det(E_n) = a_{11} \cdots a_{nn} = 1$.

Ist umgekehrt $\Delta \colon M_n(\mathbb{K}) \to \mathbb{K}$ eine alternierende n-lineare Abbildung mit $\Delta(E_n) = 1$, so erhalten wir mit (2), dass $\Delta = \det$ ist. □

Die Leibniz-Formel gestattet manche Einsicht in die Determinante einer Matrix, ist aber für nicht sehr kleine Werte von n i.a. numerisch zu aufwendig, um für praktische Berechnungen brauchbar zu sein. So wächst alleine schon die Anzahl der Summanden außerordentlich schnell; denn es gibt $n!$ Permutationen von n Ziffern:

$$2! = 2, \quad 3! = 6, \quad 4! = 24, \quad 5! = 120, \quad 6! = 720, \quad 7! = 5040, \quad 8! = 40320,$$

und

$$10! = 3,6288 \times 10^6, \quad 20! = 204329 \times 10^{18}, \quad 100! = 9,3326 \times 10^{157}.$$

Um die Leibniz-Formel zu verstehen, betrachten wir die Fälle $n = 2$ und $n = 3$.

- Ist $n = 2$, so haben wir die symmetrische Gruppe $S_2 = \{\mathrm{id}, (1\ 2)\}$ zu betrachten, und es ergibt sich die Formel, die wir schon seit langem kennen:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

- Ist $n = 3$, so geht es um $S_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$, mit

$$\operatorname{sgn}(\mathbf{1}) = \operatorname{sgn}((1\ 2\ 3)) = \operatorname{sgn}((1\ 3\ 2)) = 1,$$

$$\operatorname{sgn}((1\ 2)) = \operatorname{sgn}((1\ 3)) = \operatorname{sgn}((2\ 3)) = -1.$$

Also finden wir die sogenannte *Regel von Sarrus*, die man sich am besten durch aufsteigende und absteigende Diagonalen in der erweiterten Matrix $(s_1, s_2, s_3, s_1, s_2)$ veranschaulicht. Es gibt genau drei aufsteigende Diagonalen, die positiv in die Determinante eingehen und drei absteigende, die negativ eingehen:

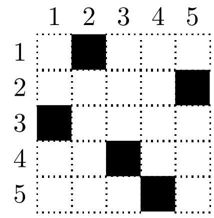
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$
$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23}$$

Die Regel von Sarrus gilt nur in Dimension 3; verallgemeinert sich also z.B. nicht auf 4×4-Matrizen.

Bemerkung 6.4.3. In jedem einzelnen Produkt

$$a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \dots \cdot a_{\sigma(n),n}$$

kommt der j-te Faktor aus der j-te Spalte der Matrix A, und welchen wir nehmen, d.h. sein Zeilenindex, hängt von der Permutation σ ab. So wählen wir (wegen der Bijektivität von σ) verschiedene Zeilenindizes für verschiedene Spalten.

Für $n = 5$ können wir uns z.B. $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 4 & 5 & 2 \end{pmatrix}$ als 

veranschaulichen.

Lemma 6.4.4. *Ist $A \in M_n(\mathbb{K})$ eine obere Dreiecksmatrix, d.h. gilt $a_{ij} = 0$ für $i > j$, so hat man*

$$\det(A) = \begin{vmatrix} a_{11} & * & \dots & * \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$$

Insbesondere gilt

$$\det(E_n) = 1.$$

Beweis. Sei $\sigma \in S_n$ eine Permutation. Falls $\sigma(i) > i$, so ist $a_{\sigma(i),i} = 0$. Also trägt eine Permutation σ nur dann zur Summe

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

bei, wenn $\sigma(i) \leq i$ für alle i gilt. Darus folgt aber per Induktion $\sigma(1) = 1$, $\sigma(2) = 2$ usf., so dass $\sigma = \mathrm{id}$ ist und die ganze Determinante auf

$$\det(A) = \operatorname{sgn}(\mathrm{id}) a_{11} \cdots a_{nn} = a_{11} \cdots a_{nn}$$

zusammenschnurrt. □

6.5 Eigenschaften der Determinante

In diesem Abschnitt lernen wir insbesondere effektive Methoden zur Berechnung der Determinante einer Matrix kennen.

Lemma 6.5.1. $\det(A^\top) = \det(A)$.

Beweis. Die Transposition einer Matrix vertauscht die Rollen von Zeilen und Spalten. Daher gilt:

$$\Delta(A^\top) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot \ldots \cdot a_{n,\sigma(n)}$$

Aber

$$a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdot \dots \cdot a_{n,\sigma(n)} = a_{\sigma^{-1}(1),1} \cdot a_{\sigma^{-1}(2),2} \cdot \dots \cdot a_{\sigma^{-1}(n),n}.$$

Denn auf beiden Seiten haben wir in der Tat dieselben Faktoren, nur in anderer Reihenfolge. Folglich ist

$$\det(A^\top) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma^{-1}(1),1} \cdot a_{\sigma^{-1}(2),2} \cdot \dots \cdot a_{\sigma^{-1}(n),n}$$

Nun glt aber nach 6.2.2 $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$, also erhalten wir

$$\det(A^\top) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) \cdot a_{\sigma^{-1}(1),1} \cdot a_{\sigma^{-1}(2),2} \cdot \dots \cdot a_{\sigma^{-1}(n),n}.$$

Durchläuft nun σ alle Permutationen, so tut das auch σ^{-1}. Folglich ist

$$\det(A^\top) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot \dots \cdot a_{\sigma(n),n} = \det(A),$$

womit das Lemma bewiesen ist. □

Bemerkung 6.5.2. (a) Da die Zeilen von A gerade die Spalten von $A^\top$ sind, folgt aus dem vorstehenden Lemma, dass die Determinante auch eine alternierenden n-linear Abbildung auf dem Zeilenraum $(\mathbb{K}^n)^n \to \mathbb{K}$ definiert. Insbesondere ist Lemma 6.3.3 daher auch auf die Determinante als Funktion der Zeilen anwendbar.

(b) So wie die Spur nicht daran denkt, multiplikativ zu sein, denkt die Determinante nicht daran, additiv zu sein, d.h. im allgemeinen ist $\det(A+B) \neq \det A + \det B$; zum Beispiel

$$1 = \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \det\Bigg(\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{A} + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{B}\Bigg), \text{ aber}$$
$$0 = \det\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{A} + \det\underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{B}.$$

(c) $\det(\lambda A) = \lambda^n \det A$ für alle $\lambda \in \mathbb{K}$ und alle $n \times n$-Matrizen A.

6.5.3. [Eigenschaften der Determinante] Bisher haben wir die folgenden Eigenschaften der Determinante kennengelernt:

(D1) $\det(A)$ hängt linear von jeder Spalte von A ab; anders gesagt:

$$\begin{vmatrix} a_{11}+a'_{11} & a_{12} & \dots & a_{1n} \\ a_{21}+a'_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1}+a'_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a'_{11} & a_{12} & \dots & a_{1n} \\ a'_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a'_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$\begin{vmatrix} \lambda a_{11} & a_{12} & \dots & a_{1n} \\ \lambda a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \lambda \cdot \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

und das gleiche gilt für die Zeilen anstatt der Spalten.

(D2) Sind zwei Spalten/Zeilen von A gleich, so ist $\det A = 0$.

(D3) $\det(E_n) = 1$.

(D4) $\det(A^\top) = \det(A)$, wobei $A^\top$ die Transponierte der Matrix A ist.

(D5) Hat A eine Spalte, die nur aus Nullen besteht, so ist $\det A = 0$ (wegen (D1)).

(D6) Addiert man ein Vielfaches einer Spalte/Zeile zu einer anderen Spalte/Zeile, so lässt das die Determinante unverändert.

(D7) Permutiert man die Spalten von A gemäßt der Permutation σ, so ändert das die Determinante um das Vorzeichen $\operatorname{sgn}(\sigma)$.

(D8) Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente.

Satz 6.5.4. (D9) $\det(A \cdot B) = \det A \cdot \det B$.

(D10) *Ist A is invertierbar, so gilt* $\det(A^{-1}) = \dfrac{1}{\det A}$.

(D11) $\det A \neq 0 \Longleftrightarrow A$ *ist invertierbar.*

(D12) *Für Blockdiagonalmatrizen mit $n = p + q$, $A \in M_p(\mathbb{K})$, $C \in M_{p,q}(\mathbb{K})$ und $B \in M_q(\mathbb{K})$ gilt:*

$$\det \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array} \right) = \det A \cdot \det B.$$

Beweis. (D9) Betrachte für eine feste Matrix A die Abbildung

$$\Delta\colon M_n(\mathbb{K}) \cong (\mathbb{K}^n)^n \to \mathbb{K}, \quad B \mapsto \det(AB).$$

Sind $t_1, \dots, t_n$ die Spalten von B, so sind $At_1, \dots, At_n$ die Spalten von AB. Für jedes j hängt die j-te Spalte von AB linear von der j-ten Spalte von B ab, ist aber unabhängig von den anderen Spalten von B. Daraus folgt, dass Δ eine n-lineare Abbildung ist.

Weiterhin folgt aus $t_i = t_j$ auch $At_i = At_j$. Hat also B zwei gleiche Spalten, so auch AB, und daher $\Delta(B) = \det(AB) = 0$.

Also ist Δ alternierend n-linear und mit Hauptsatz 6.4.2(3) ergibt sich daher

$$\det(AB) = \Delta(B) = \Delta(E_n)\det(B) = \det(A)\det(B).$$

(D10) Ist A invertierbar, so folgt aus (D9) sofort

$$1 = \det(E_n) = \det(A)\det(A^{-1}),$$

also $0 \neq \det(A)$ und $\det(A^{-1}) = \det(A)^{-1}$.

(D11) Aus (D10) folgt, dass $\det(A) \neq 0$ für jede invertierbare Matrix A gilt. Ist A nicht invertierbar, so sind die Spalten $s_1, \dots, s_n$ von A linear abhängig (Korollar 5.8.5). Dann existiert ein j und Skalare $\lambda_i \in \mathbb{K}$, $i \neq j$, mit

$$s_j = \sum_{i \neq j} \lambda_i s_i.$$

Damit ergibt sich aus der Linearität von det in der j.Spalte und (D2):

$$\det(A) = \det(s_1, \dots, s_n) = \sum_{i \neq j} \lambda_i \det(s_1, \dots, s_{j-1}, s_i, s_{j+1}, \dots, s_n) = 0,$$

da in jedem Summanden jeweils zwei Spalten übereinstimmen.

(D12) Zuerst beweisen wir den Spezialfall:

$$\det \left(\begin{array}{c|c} E_p & C \\ \hline 0 & B \end{array}\right) = \det B,$$

indem wir erst einmal zeigen, dass $\det \left(\begin{array}{c|c} E_p & C \\ \hline 0 & B \end{array}\right) = \det \left(\begin{array}{c|c} E_p & 0 \\ \hline 0 & B \end{array}\right)$ ist.

In der Tat brauchen wir nur geeignete skalare Vielfache der ersten p Spalten von den letzten q Spalten abzuziehen. Dabei ändert sich die Determinant nicht (D6). Weiter definieren wir

$$\Delta(B) := \det \left(\begin{array}{c|c} E_p & 0 \\ \hline 0 & B \end{array}\right) \qquad \text{für } B \in M_q(\mathbb{K}).$$

Man prüft leicht nach, dass $\Delta(B)$ eine alternierende n-lineare Abbildung auf $M_q(\mathbb{K})$ ist. Aus dem Hauptsatz 6.3.3 folgt also $\Delta(B) = \Delta(E_q)\det(B) = \det(B)$. Ebenso ergibt sich

$$\det \left(\begin{array}{c|c} A & 0 \\ \hline 0 & E_q \end{array}\right) = \det A$$

Dann folgt aber der allgemeine Fall aus

$$\left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right) = \left(\begin{array}{c|c} E_p & C \\ \hline 0 & B \end{array}\right) \cdot \left(\begin{array}{c|c} A & 0 \\ \hline 0 & E_q \end{array}\right)$$

nach den Regeln für die Multiplikation von Blockmatrizen. Also nach (D9)

$$\det \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right) = \det \left(\begin{array}{c|c} A & 0 \\ \hline 0 & E_q \end{array}\right) \det \left(\begin{array}{c|c} E_p & C \\ \hline 0 & B \end{array}\right) = \det A \cdot \det B.$$

□

6.5.5. [Eine effiziente Methode zur Berechnung von Determinanten]

Durch elementare Zeilen- (oder Spalten-) Unformungen lässt sich jede Matrix A auf Stufenform bringen (Gauß–Jordan Algorithmus). Aus (D4), (D6) und (D7) wissen wir, dass diese Umformungen höchstens das Vorzeichen der Determinante ändern. Nach (D8) können wir die Determinante einer Matrix in Stufenform einfach berechnen. Zusammen ergibt sich also eine recht effiziente Methode zur Bestimmung der Determinante einer quadratischen Matrix.

Beispiel 6.5.6.

$$\begin{vmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 2 & 3 \\ 2 & 1 & 0 & 3 \\ 3 & 1 & 1 & 1 \end{vmatrix} \overset{(D7)}{=} - \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 2 & 1 & 0 & 3 \\ 3 & 1 & 1 & 1 \end{vmatrix} \overset{(D6)}{=} - \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & -4 & -3 \\ 0 & 1 & -5 & -8 \end{vmatrix} \overset{(D12)}{=} - \begin{vmatrix} 1 & 2 & 3 \\ 1 & -4 & -3 \\ 1 & -5 & -8 \end{vmatrix}$$

$$\overset{(D6)}{=} - \begin{vmatrix} 1 & 2 & 3 \\ 0 & -6 & -6 \\ 0 & -7 & -11 \end{vmatrix} \overset{(D12)}{=} - \begin{vmatrix} -6 & -6 \\ -7 & -11 \end{vmatrix} \overset{(D1)}{=} 6 \cdot \begin{vmatrix} 1 & 1 \\ -7 & -11 \end{vmatrix} = 6(-11+7) = -24.$$

Satz 6.5.7. [Entwicklung nach einer Zeile oder Spalte] *Sei $A = (a_{ij})$ eine $n \times n$-Matrix. Wir bezeichnen mit A_{ij} die $(n-1) \times (n-1)$-Matrix, die aus A durch Weglassen der i-ten Zeile und der j-ten Spalte entsteht:*

$$A_{ij} := \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{pmatrix} \in M_{n-1}(\mathbb{K}).$$

Dann gilt folgende Entwicklung:

$$\boxed{\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})} \qquad \textit{(Entwicklung nach der j-ten Spalte)}$$

and entsprechend

$$\boxed{\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})} \qquad \textit{(Entwicklung nach der i-ten Zeile)}$$

Die jeweils anzuwendenden Vorzeichen $(-1)^{i+j}$ lassen sich leicht aus der Schachbrettregel entnehmen:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

Beweis. Im Hinblick auf (D4) genügt es, die Entwicklung nach einer Spalte zu beweisen; denn $(A^\top)_{ij} = A_{ji}^\top$, womit sich die Zeilenentwicklungen von $\det(A)$ aus den Spaltenentwicklungen von $\det(A^\top)$ ergeben.

Bezeichnet man mit $s_1, \dots, s_n$ die Spalten von A, so ist

$$s_j = \sum_{i=1}^{n} a_{ij} e_i,$$

und aus (D1) erhalten wir

$$\begin{aligned}\det A = \det(s_1,\dots,s_n) &= \sum_{i=1}^{n} a_{ij} \det(s_1,\dots,s_{j-1},e_i,s_{j+1},\dots,s_n)\\ &= \sum_{i=1}^{n} (-1)^{j-1} a_{ij} \det(e_i,s_1,\dots,s_{j-1},s_{j+1},\dots,s_n),\end{aligned}$$

wobei wir die zyklische Vertauschnung $(j\ 1\ 2\ \cdots\ j-1)$ auf die Spalten angewendet haben, deren Signum $(-1)^{j-1}$ ist. Vertauschen wir entsprechend die ersten i Zeilen der erhaltenen Matrix ergibt wegen (D12):

$$\det(e_i,s_1,\dots,s_{j-1},s_{j+1},\dots,s_n) = (-1)^{i-1} \det \begin{pmatrix} 1 & * \\ 0 & A_{ij} \end{pmatrix} = (-1)^{i-1} \det(A_{ij}).$$

Alles in allem erhalten wir die Entwicklungsformel

$$\det A = \sum_{i=1}^{n} (-1)^{j-1} a_{ij} (-1)^{i-1} \det(A_{ij}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \qquad \square$$

Beispiel 6.5.8. Enthält eine Zeile oder Spalte viele Nullen, so erlaubt die Entwicklung nach dieser Zeile bzw. Spalte eine besonders effektive Berechnung der Determinante. Für die folgende Matrix z.B. entwickeln wir nach der dritten Zeile und dann nach der ersten Spalte, usf.:

$$\begin{aligned}\begin{vmatrix} 1&0&1&0&1\\ 0&1&1&1&1\\ 1&0&1&0&0\\ 0&1&1&1&0\\ 1&0&1&1&0 \end{vmatrix} &= \begin{vmatrix} 0&1&0&1\\ 1&1&1&1\\ 1&1&1&0\\ 0&1&1&0 \end{vmatrix} + \begin{vmatrix} 1&0&0&1\\ 0&1&1&1\\ 0&1&1&0\\ 1&0&1&0 \end{vmatrix}\\ &= -\begin{vmatrix} 1&0&1\\ 1&1&0\\ 1&1&0 \end{vmatrix} + \begin{vmatrix} 1&0&1\\ 1&1&1\\ 1&1&0 \end{vmatrix} + \begin{vmatrix} 1&1&1\\ 1&1&0\\ 0&1&0 \end{vmatrix} - \begin{vmatrix} 0&0&1\\ 1&1&1\\ 1&1&0 \end{vmatrix}\\ &= -\begin{vmatrix} 1&1\\ 1&1 \end{vmatrix} + (1-1-1) + \begin{vmatrix} 1&1\\ 0&1 \end{vmatrix} - \begin{vmatrix} 1&1\\ 1&1 \end{vmatrix}\\ &= 0-1+1+0 = 0.\end{aligned}$$

Definition 6.5.9. [Die adjungierte Matrix] Betrachten wir die $(n-1)\times(n-1)$-Matrizen A_{ij} wie in 6.5.7, und definieren folgende Skalare:

$$a'_{ij} := (-1)^{i+j} \det(A_{ij}), \qquad i=1,\dots,n;\ j=1,\dots,n.$$

Diese Skalare heißen die *Kofaktoren* von A. Man fasst sie als Einträge einer neuen Matrix zusammen

$$\mathbf{Ad}A := (a'_{ij}) = \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ a'_{21} & a'_{22} & \dots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ a'_{n1} & a'_{n2} & \dots & a'_{nn} \end{pmatrix} \in M_n(\mathbb{K}).$$

und nennt diese neue Matrix die *adjungierte Matrix* of A.

Satz 6.5.10.

$$\boxed{(\mathbf{Ad}A)^\top \cdot A = \det A \cdot E_n}$$

Beweis. Zuerst halten wir fest:

$$\begin{aligned} & (j\text{-te Zeile von } (\mathbf{Ad}A)^\top) \cdot (j\text{-te Spalte von } A) \\ = \; & (j\text{-te Spalte von } \mathbf{Ad}A)^\top \cdot (j\text{-te Spalte von } A) \\ = \; & \sum_{i=1}^n a'_{ij} \cdot a_{ij} = \sum_{i=1}^n (-1)^{i+j} \det(A_{ij}) a_{ij} \overset{6.5.7}{=} \det A. \end{aligned}$$

Auf der Diagonalen der Produktmatrix $(\mathbf{Ad}A)^\top A$ steht also stets der Skalar $\det A$. Alle anderen Einträge, außerhalb der Diagonalen, sind aber Null, da für $j \neq k$ gilt:

$$\begin{aligned} & (j\text{-te Zeile von } (\mathbf{Ad}A)^\top) \cdot (k\text{-te Spalte von } A) \\ = \; & (j\text{-te Spalte von } \mathbf{Ad}A) \cdot (k\text{-te Spalte von } A) \\ = \; & \sum_{i=1}^n a'_{ij} a_{ik} = \sum_{i=1}^n (-1)^{i+j} \det(A_{ij}) a_{ik} = \sum_{i=1}^n (-1)^{i+j} a_{ik} \det(A_{ij}) \\ \overset{!}{=} \; & 0. \end{aligned}$$

Für den letzten Schritt betrachten wir die Matrix $\widetilde{A}$, die man aus A dadurch erhält, dass man die j-te Spalte von A durch die k-te Spalte ersetzt. In diesem speziellen Fall stimmen diese beiden Spalten überein, also ist $\det(\widetilde{A}) = 0$. □

Satz 6.5.11. [Die Cramersche Regel, erste Version: die Inverse einer invertierbaren Matrix] *Für eine invertierbare $n \times n$-Matrix A gilt*

$$\boxed{A^{-1} = \frac{1}{\det A} (\mathbf{Ad}A)^\top .}$$

Beweis. Ist A invertierbar, so ist $\det A \neq 0$ und aus Satz 6.4.12 folgt

$$\frac{1}{\det A} (\mathbf{Ad}A)^\top \cdot A = E_n.$$

Mit A^{-1} von rechts multipliziert ergibt sich die Behauptung. □

Diese Fomel für Inverse einer Matrix ist vor allem für theoretische Zwecke wichtig. Z.B. zeigt sie, dass die Einträge der inversen Matrix A^{-1} „rationale Funktionen“ der Einträge von A sind. Für praktische Berechnungen ist der Gauß-Jordan Algorithmus effizienter.

Beispiel 6.5.12. (a) Für $n = 2$ gilt

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{mit} \quad \mathbf{Ad}A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

Ist A invertierbar, so gilt

$$A^{-1} = \frac{1}{\det A}(\mathbf{Ad}A)^\top = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

wie wir schon in Beispiel 2.5.16 gesehen haben.

(b) Für $n = 3$ und $\lambda \in \mathbb{K}$ betrachte

$$A = \begin{pmatrix} \lambda & 0 & 1 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix} \quad \text{mit} \quad \mathbf{Ad}A = \begin{pmatrix} \lambda^2 & -\lambda & 1 \\ 1 & \lambda^2 & -\lambda \\ -\lambda & 1 & \lambda^2 \end{pmatrix}.$$

Eine Anwendung der Regel von Sarrus liefert $\det A = \lambda^3 + 1$. Für alle λ mit $\lambda^3 \neq -1$ gilt $\det A \neq 0$. Also ist in diesen Fällen A invertierbar und

$$A^{-1} = \frac{1}{\lambda^3 + 1}\begin{pmatrix} \lambda^2 & 1 & -\lambda \\ -\lambda & \lambda^2 & 1 \\ 1 & -\lambda & \lambda^2 \end{pmatrix}$$

Bemerkung 6.5.13. [Cramersche Regel zur Lösung lin. Systeme] Sei A eine $n \times n$-Matrix mit den Spalten $s_1, \ldots, s_n$ und sei $b \in \mathbb{K}^n$. Sei $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ eine Lösung des inhomogenen Systems n linearer Gleichungen in n Unbekannten:

$$Ax = b, \qquad \text{d.h.} \qquad \sum_{i=1}^{n} x_i s_i = b. \tag{L}$$

Sei $A_i[b]$ die Matrix, die aus A dadurch entsteht, dass anstelle der i-ten Spalte der Spaltenvektor b eingesetzt wird:

$$A_i[b] := \begin{array}{c} \begin{matrix} \phantom{a_{11}} & & i & & \phantom{a_{1n}} \\ & & \downarrow & & \end{matrix} \\ \begin{pmatrix} a_{11} & \ldots & b_1 & \ldots & a_{1n} \\ a_{21} & \ldots & b_2 & \ldots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & b_n & \ldots & a_{nn} \end{pmatrix} \end{array}$$

Dann folgt aus (L) und aus der Linearität der Determinante in der i-ten Spalte, dass

$$\det A_i[b] = \sum_{k=1}^{n} x_k \det(s_1, \ldots, s_{i-1}, s_k, s_{i+1}, \ldots, s_n),$$

wobei der Spaltenvektor s_k als i-te Spalte in A eingefügt ist. Für $i \neq k$ ist die Determinante in der Summe Null, weil zwei Spalten gleich sind, die i-te und die k-te. Also bleibt von der obigen Summe nur der Summand mit $i = k$ über und wir erhalten:

$$\det A_i[b] = x_i \det(A).$$

Nun gibt es zwei Fälle: entweder ist $\det(A) = 0$ und die Formel besagt nur, dass auch $\det A_i[b] = 0$ ist — eine Information, über deren Wert man streiten kann. Ist dagegen $\det(A) \neq 0$, anders gesagt: *ist die Matrix A invertierbar*, so erhalten wir eine Formel, nämlich

$$x_i = \frac{\det A_i[b]}{\det(A)}$$

für die i-te Komponenten der, wie wir wissen *einzigen* Lösung x des Systems (L) in diesem Fall.

Um diese Formeln für alle $i = 1, \ldots, n$ zusammenzusetzen, können wir auch so argumentieren: Da wegen der Invertierbarkeit von A das Gleichungssystem (L) genau eine Lösung

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

besitzt, nämlich $x = A^{-1}b$, erhalten wir aus 6.5.11: $x = \frac{1}{\det A}(\mathbf{Ad}A)^\top b$, d.h.

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det A} \begin{pmatrix} a'_{11} & a'_{21} & \ldots & a'_{n1} \\ a'_{12} & a'_{22} & \ldots & a'_{n2} \\ \vdots & \vdots & & \vdots \\ a'_{1n} & a'_{2n} & \ldots & a'_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{\det(A)} (\operatorname{Ad} A)^\top b.$$

Denn die Determinante von $A_i[b]$ berechnet sich durch Entwicklung nach der i-ten Spalte zu:

$$\det A_i[b] = \sum_{j=1}^{n} (-1)^{i+j} b_j \det(A_{ji}).$$

Beispiel 6.5.14. Das Gleichungssystem $\left\{ \begin{array}{rrrr} \lambda x_1 & & + & x_3 = 1 \\ x_1 & +\lambda x_2 & & = 1 \\ & x_2 & + & \lambda x_3 = 1 \end{array} \right\}$ hat die Lösung

$$x_1 = \frac{\begin{vmatrix} 1 & 0 & 1 \\ 1 & \lambda & 0 \\ 1 & 1 & \lambda \end{vmatrix}}{\lambda^3 + 1}, \qquad x_2 = \frac{\begin{vmatrix} \lambda & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & \lambda \end{vmatrix}}{\lambda^3 + 1}, \qquad x_3 = \frac{\begin{vmatrix} \lambda & 0 & 1 \\ 1 & \lambda & 1 \\ 0 & 1 & 1 \end{vmatrix}}{\lambda^3 + 1}.$$

Daraus ergibt sich

$$x_1 = \frac{\lambda^2 - \lambda + 1}{\lambda^3 + 1}, \qquad x_2 = \frac{\lambda^2 - \lambda + 1}{\lambda^3 + 1}, \qquad x_3 = \frac{\lambda^2 - \lambda + 1}{\lambda^3 + 1},$$

$$x_1 = \frac{1}{\lambda + 1}, \qquad x_2 = \frac{1}{\lambda + 1}, \qquad x_3 = \frac{1}{\lambda + 1}$$

(Vergleiche dies mit Beispiel (b) in 6.5.11.)

6.6 Die Determinante eines Endomorphismus

Wir machen jetzt denselben Schritt von Matrizen zu Endomorphismen, den wir bereits von der Spur her kennen. Sei dazu V ein endlich dimensionaler $\mathbb{K}$-Vektorraum und $\varphi\colon V \to V$ ein Endomoprhismus von V. Sei $A = [\varphi]_B$ die Matrix von φ bzgl. der Basis B von V und sei $A' = [\varphi]_{B'}$ die Matrix von φ bzgl. einer anderen Basis B' von V. Aus der Trasnformationsformel bei Basiswechsel wissen wir, dass es eine invertierbare Matrix $S = [\mathrm{id}_V]_B^{B'}$ gibt (die Übergangsmatrix), so dass gilt:

$$A' = S^{-1}AS.$$

Aus (D9/10) in 6.5.3 folgt

$$\begin{aligned}\det(A') &= \det(S^{-1}AS) = \det(S^{-1}) \cdot \det A \cdot \det S = (\det S)^{-1} \cdot \det A \cdot \det S \\ &= \det(A)\end{aligned}$$

Die Determinante der Matrix eines Endomorphismus φ ist also von der hinzugezogenen Basis unabhängig. Also können wir wie im Fall der Spur auch die Determinante eines Endomorphismus definieren:

Definition 6.6.1. Die Determinante eines Endomorphismus $\varphi : V \to V$ eines endlich dimensionalen Vektorraums V ist die Determinante irgendeiner (und damit jeder) Matrix $[\varphi]_B$, die φ bzgl. einer Basis B von V zugeordnet ist:

$$\det \varphi := \det([\varphi]_B).$$

Die Eigenschaften (D3), (D9), (D10), (D11) in 6.5.3 ergeben:

Satz 6.6.2. *Für* $\varphi, \psi \in \mathrm{End}(V)$ *und* $\dim V < \infty$ *gilt*

(i) $\det(\mathrm{id}_V) = 1$

(ii) $\det \varphi \neq 0 \iff \varphi$ *ist invertierbar, und in diesem Fall ist* $\det(\varphi^{-1}) = \frac{1}{\det \varphi}$.

(iii) $\det(\varphi \circ \psi) = \det \varphi \cdot \det \psi$.

Beispiel 6.6.3. (a) Sei δ die Drehung im $\mathbb{R}^3$ um eine Gerade g durch 0. Dann gibt es, wie wir wissen, eine Basis des $\mathbb{R}^3$, so dass die Matrix von δ bzgl. dieser Basis die Form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

hat. Also ist $\det \delta = \cos^2 \alpha + \sin^2 \alpha = 1$.

(b) Sei σ die Spiegelung des $\mathbb{R}^3$ an einer Ebene. In Bezug auf eine geeignete Basis hat die zugehörige Matrix die Form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Also ist $\det \sigma = -1$.

6.7 Die Determinante als „orientierter Flächeninhalt" im $\mathbb{R}^2$

In diesem Abschnitt diskutieren wir die geometrische Bedeutung des Absolutbetrages der Determinante reeller (2×2)-Matrizen als Fläche eines ebenen Parallelogramms. Wir erinnern uns an das Skalarprodukt zweier Vektoren x, y im $\mathbb{R}^2$:

$$\langle x, y\rangle = x_1y_1 + x_2y_2,$$

sowie an die Länge eines Vektors:

$$\|x\| := \sqrt{x_1^2 + x_2^2}$$

(Beispiel 5.5.18).

Definition 6.7.1. Den *Winkel zwischen zwei Vektoren $x \neq 0$ und $y \neq 0$* in der Ebene $\mathbb{R}^2$ definieren wir als die eindeutig bestimmte reelle Zahl $\gamma \in [0, 2\pi[$ (dass sie eindeutig bestimmt ist, wird aus der Analysisvorlesung als bekannt vorausgesetzt), so dass y aus x durch die Drehung δ_γ (gegen den Uhrzeigersinn) um den Ursprung hervorgeht, d.h.,

$$y = \lambda \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} x, \quad \lambda > 0, \gamma \in [0, 2\pi[.$$

Mit anderen Worten:

$$y_1 = (\cos\gamma)x_1 - (\sin\gamma)x_2, \quad y_2 = (\sin\gamma)x_1 + (\cos\gamma)x_2.$$

Für den Vektor $\widetilde{x} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$, der auf x senkrecht steht, erhalten wir die Darstellung

$$y = \lambda\cos(\gamma)x + \lambda\sin(\gamma)\widetilde{x}. \tag{6.1}$$

Wir nenne das Paar (x, y) *positiv orientiert*, falls $\gamma \in]0, \pi[$, und *negativ orientiert*, falls $\gamma \in]\pi, 2\pi[$ ist.

Definition 6.7.2. [Determinante eines Vektorpaares im $\mathbb{R}^2$]

Für jedes Paar von zwei Vektoren $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ und $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ im $\mathbb{R}^2$ definieren wir die Determinante durch

$$\det(u, v) := \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} = u_1v_2 - u_2v_1 = \langle \widetilde{u}, v\rangle.$$

Bemerkung 6.7.3. Da sich die Länge eines Vektors unter Drehungen nicht ändert, ist

$$\|y\| = \lambda\|x\|.$$

Darüber hinaus finden wir wegen $\langle x, \widetilde{x}\rangle = 0$:

$$\langle x, y\rangle = \langle x, \lambda\cos(\gamma)x\rangle = \lambda\cos(\gamma)\|x\|^2 = \cos(\gamma)\|x\|\|y\|.$$

Für Vektoren, die nicht null sind, erhalten wir also den Kosinus des von ihnen eingeschlossenen Winkels aus dem Skalarprodukt vermöge der Formel

$$\cos\gamma = \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

Für die Determinante erhalten wir analog

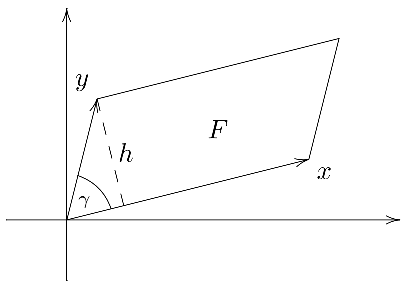
$$\det(x, y) = \langle\widetilde{x}, y\rangle = \langle\widetilde{x}, \lambda\sin(\gamma)\widetilde{x}\rangle = \lambda\|\widetilde{x}\|^2\sin(\gamma) = \lambda\|x\|^2\sin(\gamma) = \|x\|\|y\|\sin(\gamma).$$

Insbesondere ist (x, y) positiv orientiert, wenn $\det(x, y) > 0$ ist und negativ orientiert, wenn $\det(x, y) < 0$ ist. In beiden Fällen sind x, y linear unabhängig, also eine Basis von $\mathbb{R}^2$. Darüber hinaus ist $\det(x, y) = 0$ äquivalent zu $\gamma \in \{0, \pi\}$, falls beide Vektoren von 0 verschieden sind. Wir sehen also ganz allgemein, dass $\det(x, y) = 0$ dazu äquivalent ist, dass x, y linear abhängig sind.

6.7.4. [Interpretation als orientierter Flächeninhalt]

Sei γ der Winkel zwischen x und y mit $0 \leq \gamma < 2\pi$.
Für die Fläche F des von x und y aufgespannten Parallelogramms gilt:

$$\begin{aligned}\operatorname{area}(F) &= (\text{Länge der Basis}) \cdot (\text{Höhe}\, h)\\ &= \|x\| \cdot \|y\| \cdot |\sin\gamma| = |\det(x, y)|\end{aligned}$$



Wir erhalten daher folgenden Satz:

Theorem 6.7.5. *Für $x, y \in \mathbb{R}^2$ gilt:*

(a) $|\det(x, y)|$ *ist die Fläche des von x und y aufgespannten Parallelogramms.*

(b)
$$\det(x, y)\begin{cases} > 0 \iff x, y & \textit{lin. unabh. pos. orientiert}\\ = 0 \iff x, y & \textit{lin. abbh}\\ < 0 \iff x, y & \textit{lin. unabh., neg. orientiert}\end{cases}$$

Bemerkung 6.7.6. Jetzt können wir die Regeln (D1) und (D2) für die Determinante geometrisch deuten.

(D1) sagt aus, dass bei festgehaltenem v der Wert von $\det(u, v)$ linear von u, und bei festgehaltenem u linear von v abhängt. Um die erste dieser beiden

Eigenschaften geometrisch zu deuten, sei $v \neq 0$ und $\pi\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Projektion auf die Gerade $\mathbb{R}v'$ mit $\ker \pi = \mathbb{R}v$, wo $v' = (v_2, -v_1)^\top$. Diese lineare Abbildung heißt die *orthogonale Projektion* auf $\mathbb{R}v'$. Da $\ker \pi = \mathbb{R}v$ und $\det(\lambda v, v) = 0$ ist, gilt

$$\det(u, v) = \det(\pi(u), v).$$

Da die Vektoren $\pi(u)$ und v orthogonal zueinander stehen, ist die Fläche des von $\pi(u)$ und v aufgespannten Parallelogramms gleich $\|\pi(u)\| \cdot \|v\|$. Da $\pi(u_1), \pi(u_2) \in \mathbb{R}^+ v'$ sind, entspricht die Gleichheit

$$\det(u_1 + u_2, v) = \det(u_1, v) + \det(u_2, v)$$

der Tatsache, dass die Fläche des von $\pi(u_1) + \pi(u_2) = \pi(u_1 + u_2)$ und v aufgespannten Parallelogramms durch Aufaddieren der Flächen der zu u_1 und u_2 gehörigen Parallelogramme entsteht.

Die Relation

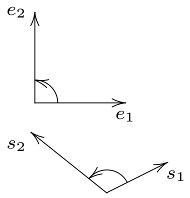
$$\det(\lambda u, v) = \lambda \det(u, v)$$

für $\lambda > 0$ entspricht der geometrischen Tatsache, dass sich die Fläche des Parallelogramms mit λ multipliziert, wenn *eine* Seite um diesen Faktor gestreckt oder gestaucht wird.

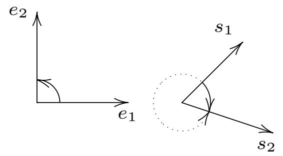
(D2) besagt einfach, dass $\det(u, u) = 0$ ist für alle $u \in \mathbb{R}^2$. In der Tat hat ein entartetes, in einer Gerade liegendes Parallelogramm keine Fläche.

6.7.7. [Geometrische Deutung der Determinante einer linearen Abbildung $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$] Sei $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ eine lineare Abbildung. Sei $s_1 := \varphi(e_1)$ und $s_2 := \varphi(e_2)$, und sei $A = [\varphi]$ die 2×2-Matrix mit den Spalten s_1, s_2, d.h. die φ bzgl. der Standardbasis zugeordnete Matrix: $\varphi(x) = Ax$ für alle $x \in \mathbb{R}^2$.

Ist $\det A > 0$, so erhält φ die Orientierung:



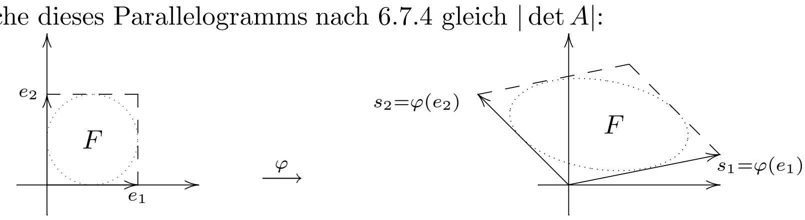
Ist $\det A < 0$, so kehrt φ die Orientierung um:



$$A = \begin{array}{c} \begin{array}{cc} \overbrace{\phantom{a_{11}}}^{s_1} & \overbrace{\phantom{a_{12}}}^{s_2} \end{array} \\ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \end{array}$$

Typische Beispiele für orientierungserhaltende Abbildungen sind Drehungen. Typische Beispiele für orientierungsumkehrende Abbildungen sind Spiegelungen an einer Gerade im $\mathbb{R}^2$.

Weiterhin ist das Bild des von e_1 und e_2 aufgespannten Einheitsquadrats das von $s_1 = \varphi(e_1)$ und $s_2 = \varphi(e_2)$ aufgespannte Parallelogramm. Also ist die Fläche dieses Parallelogramms nach 6.7.4 gleich $|\det A|$:



Das Bild des in das Einheitsquadrat einbeschriebenen Kreises ist eine Ellipse im besagten Parallelogramm. Ist F die Fläche dieses Kreises, so berechnet sich die Fläche F' der Ellipse durch:

$$F' = |\det A| \cdot F.$$

Diese letzte Formel gilt ganz allgemein für die Fläche F' des Bildes unter einer linearen Abbildung φ einer Menge $M \subseteq \mathbb{R}^2$, deren Fläche man als F messen kann. Die Theorie solcher allgemeineren Flächeninhalte wird in beliebigen Dimensionen in der Vorlesung Analysis III entwickelt.

6.8 Die Determinante als „orientiertes Volumen“ im $\mathbb{R}^3$

In diesem Abschnitt entwickeln wir das dreidimensionale Analogon des vorangehenden Abschnitts. Wir erinnern uns an Skalarprodukt und Länge von Vektoren im $\mathbb{R}^3$:

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + x_3y_3, \quad \|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Darüber hinaus führen wir das dem 3-dimensionalen Raum spezifische Vektor-Kreuzprodukt ein:

$$a \times b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ -(a_1b_3 - a_3b_1) \\ a_1b_2 - a_2b_1 \end{pmatrix} = \begin{pmatrix} + \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \\ - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \\ + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \end{pmatrix}$$

Für einen dritten Vektor $c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \in \mathbb{R}^3$ definieren wir das *Spatprodukt* als $\langle a \times b, c\rangle$.

6.8.1. [Die Determinante eines Tripels von Vektoren im $\mathbb{R}^3$]
Für ein Tripel

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \quad c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \in \mathbb{R}^3$$

gilt

$$\begin{aligned}\langle a \times b, c \rangle &= \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} \cdot c_1 - \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix} \cdot c_2 + \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \cdot c_3 \\ &= a_2b_3c_1 + a_3b_1c_2 + a_1b_2c_3 - a_3b_2c_1 - a_1b_3c_2 - a_2b_1c_3 \\ &= \det(a, b, c).\end{aligned}$$

6.8.2. [Geometrische Deutung als dreidimensionales Volumen] Aus der vorstehenden Deutung der Determinante durch das Spatprodukt dreier Vektoren und aus (D2) folgt sofort

$$0 = \langle a \times b, a \rangle = \langle a \times b, b \rangle.$$

Also steht $a \times b$ senkrecht auf a und b. Außerdem folgt aus $a \times a = 0$, dass $a \times b$ nur von der orthogonalen Projektion von b auf die zu a orthogonale Ebene $a^\perp$ abhängt. Die Länge dieser Projektion beträgt $\|b\| \sin \alpha$, wo $\alpha \in [0, \pi[$ der Winkel zwischen a und b ist (wie in der Ebene). Aus alldem ist es nicht schwer, zu zeigen, dass

$$\|a \times b\| = \|a\| \cdot \|b\| \cdot \sin \alpha,$$

was auch mit der Fläche des von a und b aufgespannten Parallelogramms $[0, 1]a+[0, 1]b$ übereinstimmt.

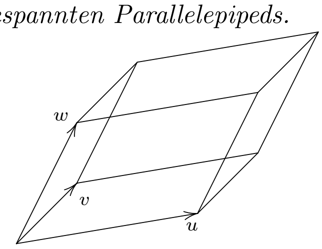
Ist c ein dritter Vektor, so erhalten wir

$$|\det(a, b, c)| = |\langle a \times b, c \rangle| = \|a \times b\| \cdot \|c\| \cdot |\cos \beta|,$$

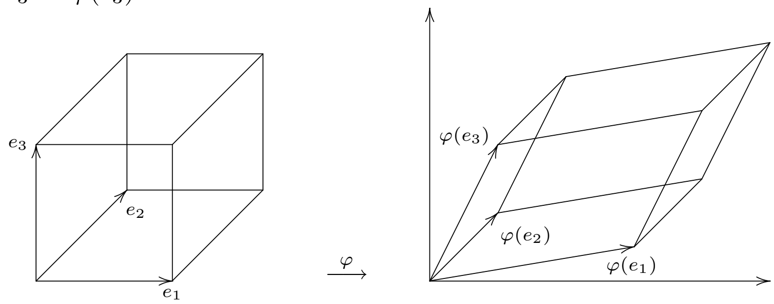
wo β der Winkel zwischen c und $a \times b$ ist. Da $\|c\| \cdot \cos \beta$ auch als Länge der Projektion von c auf die Gerade $a \times b$ orthogonal zur Ebene von a und b aufgefasst werden kann, stimmt sie mit der Höhe h des von a, b und c über der Ebene von a und b aufgespannten Parallelepipeds $P := [0, 1]a + [0, 1]b + [0, 1]c$ überein. Alles in allem erhalten wir

$$|\det(a, b, c)| = \text{area}(F) \cdot h = \text{vol}(P).$$

Satz 6.8.3. $|\det(u, v, w)|$ *ist das Volumen* V *des von den Vektoren* u, v, w *aufgespannten Parallelepipeds.*



6.8.4. [Geometrische Deutung der Determinante einer linearen Abbildung] Gegeben eine lineare Abbildung $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$, betrachten wir die ihr bzgl. der Standardbasis zugeordnete Matrix A. Ihre Spalten sind $s_1 := \varphi(e_1)$, $s_2 := \varphi(e_2)$, $s_3 := \varphi(e_3)$.



Der Einheitswürfel (mit Volumen 1) wird auf das von den Spalten $s_1 = \varphi(e_1)$, $s_2 = \varphi(e_2)$, $s_3 = \varphi(e_3)$ aufgespannte Parallelepiped abgebildet. Sein Volumen ist $|\det(s_1, s_2, s_3)| = |\det A|$.

Allgemeiner gilt: Ist M eine 'meßbare' Teilmenge in $\mathbb{R}^3$ mit Volumen V, so hat ihr Bild unter der linearen Abbildung φ das Volumen

$$V' = |\det \varphi| \cdot V.$$