Lineare Algebra 1

Ein Skriptum zur Vorlesung im
Wintersemester 2013/14

Franz Pauer

6. Auflage

 INNSBRUCK, ÖSTERREICH

Vorwort

Das vorliegende Skriptum soll den Hörerinnen und Hörern der Vorlesung „Lineare Algebra 1“ im Wintersemester 2013/14 das Mitschreiben und *Mitdenken* erleichtern. Diese Vorlesung richtet sich an Studierende der Bachelorstudien Technische Mathematik, Informatik, Physik und Atmosphärenwissenschaften, sowie der Lehramtsstudien in den Unterrichtsfächern Mathematik und Physik. Das Skriptum enthält alle Algorithmen, Definitionen und Sätze der Vorlesung, aber nur wenige Beispiele dazu. In der Vorlesung werden die Algorithmen, Definitionen und Sätze motiviert, der Zusammenhang mit früheren Ergebnissen erläutert und Beispiele dazu besprochen.

Die *Hauptziele dieser Vorlesung* sind:

- *Systeme linearer Gleichungen* zu lösen.
  Die Fragen, ob ein solches System eine Lösung hat, ob sie eindeutig ist, wie die Menge aller Lösungen durch endlich viele Daten beschrieben werden kann und wie diese Daten berechnet werden können, werden vollständig beantwortet. Dazu müssen die Matrizenrechnung und Grundbegriffe der Theorie der Vektorräume eingeführt werden. Das zentrale Rechenverfahren zur Lösung von Systemen linearer Gleichungen ist der Gauss-Algorithmus.
- Mit Hilfe der *Vektorrechnung* Fragen aus der *Geometrie* zu beantworten.
  Die Begriffe „Vektorraum“ und „Skalarprodukt“ bilden die wichtigsten Bausteine für ein mathematisches Modell der Geometrie der Ebene und des Raumes. Mit diesem Modell können viele geometrische Fragen recht einfach beantwortet werden.
- *Eigenwertprobleme* zu lösen.
  Als Hilfsmittel für ihre Lösung werden die Begriffe *Polynomfunktion*, *Permutation* und *Determinante* eingeführt und wichtige Eigenschaften davon besprochen. Diese Begriffe sind auch für andere Anwendungen von großer Bedeutung: Permutationen zum Beispiel für Sortieralgorithmen, Determinanten zum Beispiel für die Integralrechnung.

Im Kapitel 0 werden einige Grundbegriffe der Mathematik eingeführt. Kapitel 5 behandelt lineare Funktionen und ihren Zusammenhang mit Matrizen.

Dieses Skriptum ist eine überarbeitete Fassung von Teilen der Skripten
Arne Dür und Franz Pauer: Lineare Algebra (5. Auflage), 2006.
und
Arne Dür und Franz Pauer: Analytische Geometrie (3. Auflage), 2005.

Am Ende von fünf Kapiteln sind als Lernhilfe einige Fragen angegeben. Antworten und Erläuterungen dazu sind am Ende des Skriptums zu finden. Diese Fragen und Antworten hat Simone Graml zusammengestellt. Die Zeichnung „Translationen“ wurde von Anna Bombasaro, die Zeichnungen in Kapitel 2, §5 und Kapitel 5, §6 wurden von Simone Graml, alle anderen wurden von Roman Liedl angefertigt. Hubert Herdlinger hat das Skriptum kritisch gelesen und mehrere Verbesserungen angeregt. Martin Huber danke ich für die Anregung, die Abschnitte 5 in Kapitel 2 und 6 in Kapitel 5 (Anwendungen der Linearen Algebra in der Elektrotechnik) in das Skriptum aufzunehmen, sowie für die Erlaubnis, dafür seine Materialien in www.tech4math.com zu verwenden.

Die sechste Auflage des Skriptums unterscheidet sich von der fünften (August 2011) nur durch einige Korrekturen und kleinere Änderungen.

Innsbruck, September 2013

Inhaltsverzeichnis

KAPITEL 0

Mengen, Funktionen, Zahlen und Rechenregeln

§1. Mengen und Funktionen

Definitionen setzen Vorwissen voraus. Zum Beispiel setzt die Definition

„Ein Quadrat ist ein gleichseitiges Rechteck“

voraus, dass bekannt ist, was „gleichseitig“ und „Rechteck“ bedeuten. Die Definition

„Eine gerade Zahl ist eine ganze Zahl, die von 2 geteilt wird“

setzt voraus, dass bekannt ist, was „ganze Zahl“ und „teilen“ bedeuten. Für Definitionen wird häufig die folgende Kurzschreibweise verwendet:

zu definierender Begriff := *definierende (schon bekannte) Begriffe* .

Zum Beispiel:

Quadrat := gleichseitiges Rechteck

(in Worten: ein *Quadrat* ist ein gleichseitiges Rechteck)
und

gerade Zahl := ganze Zahl, die von 2 geteilt wird

(in Worten: eine *gerade Zahl* ist eine ganze Zahl, die von 2 geteilt wird).

Der Begriff „Menge“ ist jedoch ein Grundbaustein der Mathematik, der nicht definiert, sondern nur umschrieben wird: Eine *Menge* ist eine Zusammenfassung unterscheidbarer Objekte. Diese heißen *Elemente* der Menge.

Eine Menge kann auf zwei Arten angegeben werden:

(1) durch Anschreiben der Elemente zwischen geschweiften Klammern, zum Beispiel $\{7,3,5,8,1\}$, $\{$Meier, Müller$\}$;
oder

(2) durch ihre Eigenschaften, zum Beispiel
$\{n \mid n$ ganze Zahl, n ist größer als 0 und kleiner als 7$\}$
(Sprechweise: „die Menge aller n, für die gilt: n ist eine ganze Zahl, die größer als 0 und kleiner als 7 ist“ oder „die Menge aller ganzen Zahlen, die größer als 0 und kleiner als 7 sind“).

Bezeichnungen:

$\emptyset := \{\}$	*leere Menge* (Menge ohne Elemente)
$\mathbb{N} := \{0,1,2,3,\dots\}$	Menge der *natürlichen Zahlen*
$\mathbb{Z} := \{0,1,-1,2,-2,\dots\}$	Menge der *ganzen Zahlen*

Ist M eine Menge, so wird

$$e \in M$$

für „e ist ein Element von M“ geschrieben, und analog

$$e \notin M$$

für „e ist kein Element von M“.

Auf logische Probleme, die bei der Einführung des Begriffes „Menge“ auftreten, gehen wir hier nicht ein. Das „Russell'sche Paradoxon“ zeigt, dass man nicht zu sorglos sein darf:

Gilt für $M := \{A \mid A \text{ Menge}, A \notin A\}$ die Beziehung $M \in M$?

Beispiel 1 : $1 \in \mathbb{N}$, $-1 \notin \mathbb{N}$, $1 \notin \emptyset$.

Definition 2 : M und N seien Mengen. M heißt *Teilmenge* von N, in Zeichen

$$M \subset N \quad \text{oder} \quad M \subseteq N\,,$$

wenn jedes Element von M auch Element von N ist.

$$M \not\subset N$$

bedeutet, dass M nicht Teilmenge von N ist. Die Mengen M und N sind *gleich*, in Zeichen

$$M = N\,,$$

wenn $M \subset N$ und $N \subset M$ ist. Falls M und N nicht gleich sind, schreibt man

$$M \neq N\,.$$

Schließlich bedeutet

$$M \subsetneqq N\,,$$

dass $M \subset N$ und $M \neq N$ ist, und man nennt M eine *echte* Teilmenge von N.

Beispiel 3 : Für alle Mengen N ist $N \subseteq N$ und $\emptyset \subseteq N$. Es ist $\{a,b,c\} = \{b,a,c\} = \{c,a,b\}$, beim Anschreiben der Elemente einer Menge kann die Reihenfolge also beliebig gewählt werden.

Definition 4 : M und N seien Mengen. Der *Durchschnitt* von M und N ist die Menge

$$M \cap N := \{a \mid a \in M \text{ und } a \in N\}\,.$$

Die Mengen M und N sind *disjunkt*, wenn ihr Durchschnitt leer ist.
Die *Vereinigung* von M und N ist die Menge

$$M \cup N := \{a \mid a \in M \text{ oder } a \in N\}\,,$$

wobei mit „oder“ das einschließende Oder („und-oder“) und nicht das ausschließende Oder („entweder-oder“) gemeint ist.
Die *(Mengen-)Differenz* von M und N ist die Menge

$$M \setminus N := \{a \mid a \in M \text{ und } a \notin N\} .$$

Beispiel 5 :

$$\{1,2,3\} \cap \{4,3,5\} = \{3\} ,$$

$$\{1,2,3\} \cup \{4,3,5\} = \{1,2,3,4,5\} .$$

$$\{1,2,3\} \setminus \{4,3,5\} = \{1,2\} .$$

Als zweiten Grundbaustein der Mathematik führen wir den Begriff *Funktion* ein.
M und N seien Mengen. Eine *Abbildung* oder *Funktion* von M nach N ist eine Vorschrift, die jedem Element von M genau ein Element von N zuordnet. M heißt dann der *Definitionsbereich* der Funktion, N der *Bildbereich* oder *Wertebereich*. Die Schreibweisen

$$f : M \to N \, , \, m \mapsto f(m) ,$$

oder

$$\begin{array}{rccc} f: & M & \longrightarrow & N \\ & m & \mapsto & f(m) \end{array}$$

bedeuten, dass f eine Funktion von M nach N ist, die dem Element $m \in M$ das Element $f(m) \in N$ zuordnet. Das Element $f(m)$ heißt *Bild* von m (bezüglich f). Ein Element $m \in M$ mit $f(m) = n \in N$ heißt ein *Urbild* von n (bezüglich f).

Beispiel 6 : Die Funktion

$$f : \mathbb{N} \to \mathbb{Z} \, , \, z \mapsto 2z - 3 ,$$

ordnet jeder natürlichen Zahl z die ganze Zahl $2z-3$ zu. Das Bild von 0 bzw. 1 bzw. 2 bezüglich f ist -3 bzw. -1 bzw. 1. Ein Urbild von 5 ist 4. Die Zahl 4 hat kein Urbild bezüglich f.

Definition 7 : Seien $f : M \to N$ und $g : P \to Q$ Funktionen. Dann sind f und g *gleich*, in Zeichen

$$f = g ,$$

wenn gilt: $M = P$, $N = Q$ und für alle $m \in M$ ist $f(m) = g(m)$.

§2. Familien, Tupel, Folgen und kartesisches Produkt

Eine Funktion $f : I \to M$ wird manchmal in der Form

$$(f(i))_{i\in I} \quad \text{oder} \quad (f_i)_{i\in I}$$

geschrieben und als *Familie* von Elementen in M, indiziert durch I, bezeichnet. I heißt dann die *Indexmenge* der Familie $(f_i)_{i\in I}$. Die Familie $(f_i)_{i\in I}$ heißt *endlich*, wenn I endlich ist. Wichtige Spezialfälle sind:

(1) Eine Funktion $x : \{1,2,\dots,n\} \to M$, $i \mapsto x(i) =: x_i$, wird in der Form

$$(x_1,\dots,x_n) = (x_i)_{1\leq i\leq n} = (x_i)_{i\in\{1,\dots,n\}}$$

geschrieben und heißt ein *n-Tupel von Elementen in M*. Das Element x_i heißt dann *i-te Komponente* von $(x_1,\dots,x_n)$.
Die Menge aller n-Tupel von Elementen in M wird mit

$$M^n$$

bezeichnet (sprich „M hoch n“). Für $x, y \in M^n$ gilt

$$x = y$$

genau dann, wenn $x_i = y_i$ für $i = 1,\dots,n$ ist.

In den Spezialfällen $n = 2,3$ nennt man $(x_1,\dots,x_n)$ ein *Paar* bzw. *Tripel*.
Ein Paar (a,b) enthält „mehr Information“ als die Menge $\{a,b\}$. Es ist $\{a,b\} = \{b,a\}$, aber $(a,b) = (b,a)$ nur dann, wenn $a = b$ ist.

(2) Sei $m \in \mathbb{N}$ und $I := \{i \in \mathbb{N} \mid i \geq m\}$. Eine Funktion $x : I \to M$, $i \mapsto x(i) =: x_i$, wird in der Form

$$(x_i)_{i\geq m}$$

geschrieben und heißt eine *Folge* in M.
Die Folge $(x_i)_{i\geq m}$ darf nicht mit der Menge $\{x_i \mid i \geq m\}$ verwechselt werden!

Definition 8 : M und N seien Mengen. Dann heißt

$$M \times N := \{(x,y) \mid x \in M \text{ und } y \in N\} \subseteq (M \cup N)^2$$

das *kartesische Produkt* von M und N.

Definition 9 : Sei $f : M \to N$ eine Funktion. Dann heißt die Menge

$$\mathrm{Graph}(f) := \{(m, f(m)) \mid m \in M\} \subseteq M \times N$$

der *Graph* von f.

Beispiel 10 : Der Graph der Funktion

$$f : \{1,3,4,5\} \longrightarrow \mathbb{N},\ z \longmapsto 3z+1,$$

ist

$$\{(1,4),(3,10),(4,13),(5,16)\} \subseteq \{1,3,4,5\} \times \mathbb{N}.$$

Satz 11 : *Zwei Funktionen von M nach N sind genau dann gleich, wenn ihre Graphen gleich sind.*

Beweis: Es ist zu zeigen:

1. Wenn zwei Funktionen von M nach N gleich sind, dann sind auch ihre Graphen gleich.
2. Wenn die Graphen zweier Funktionen von M nach N gleich sind, dann sind diese zwei Funktionen gleich.

Seien f und g Funktionen von M nach N.
Zu 1): Wenn $f = g$ ist, dann ist $f(m) = g(m)$ für alle $m \in M$. Daher ist

$$\mathrm{Graph}(f) = \{(m, f(m)) \mid m \in M\} =$$

$$= \{(m, g(m)) \mid m \in M\} = \mathrm{Graph}(g).$$

Zu 2): Wenn $\mathrm{Graph}(f) = \mathrm{Graph}(g)$ ist, dann ist für alle $m \in M$ das Paar $(m, f(m))$ ein Element von $\mathrm{Graph}(g)$. In $\mathrm{Graph}(g)$ gibt es genau ein Element, dessen erste Komponente m ist, nämlich $(m, g(m))$. Also ist $f(m) = g(m)$ für alle $m \in M$, somit ist $f = g$.

§3. Ganze Zahlen und rationale Zahlen

Wir setzen die Menge $\mathbb{Z} := \{\ldots,-2,-1,0,1,2,\ldots\}$ der ganzen Zahlen mit der Addition $\mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}$, $(a,b) \longmapsto a+b$, und der Multiplikation $\mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}$, $(a,b) \longmapsto a \cdot b$, als bekannt voraus. Dabei gelten die folgenden Rechenregeln: Sind a,b,c ganze Zahlen, dann ist

- $(a+b)+c = a+(b+c) =: a+b+c$ („Die Addition von ganzen Zahlen ist assoziativ“, das heißt: auf Klammern kann verzichtet werden).
- $0+a = a+0 = a$
- $a+(-a) = (-a)+a = 0$ (dabei ist $-a := (-1) \cdot a$)
- $a+b = b+a$ („Die Addition ist kommutativ“).
- $(a \cdot b) \cdot c = a \cdot (b \cdot c) =: a \cdot b \cdot c$ („Die Multiplikation ist assoziativ“).
- $1 \cdot a = a \cdot 1 = a$
- $a \cdot b = b \cdot a$ („Die Multiplikation ist kommutativ“).
- $(a+b) \cdot c = (a \cdot c)+(b \cdot c) =: a \cdot c + b \cdot c$ („Distributivgesetz“)

Für $a,b,c \in \mathbb{Z}$ mit $c \neq 0$ folgt aus $a \cdot c = b \cdot c$, dass $a = b$ ist. („In $\mathbb{Z}$ kann durch Zahlen $\neq 0$ gekürzt werden“). Insbesondere folgt aus $a \cdot b = 0$,

dass $a = 0$ oder $b = 0$ ist.
Sind $m, n \in \mathbb{Z}$, $m \leq n$ und $a_m, a_{m+1}, \ldots, a_n \in \mathbb{Z}$, dann schreiben wir

$$\sum_{i=m}^{n} a_i$$

für $a_m + a_{m+1} + \ldots + a_n$ und

$$\prod_{i=m}^{n} a_i$$

für $a_m \cdot a_{m+1} \cdot \ldots \cdot a_n$. (Sprechweise: „Die Summe bzw. das Produkt aller a_i mit i von m bis n").

Die *Subtraktion* ist durch $\mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z}$, $(a, b) \longmapsto a - b := a + (-b)$, gegeben.
Das *Vorzeichen* $vz(a)$ einer ganzen Zahl a ist 1, wenn $a \in \mathbb{N}$, und -1, wenn $a \notin \mathbb{N}$ ist. Der *Betrag* $|a|$ einer ganzen Zahl a ist $vz(a) \cdot a$.
Für ganze Zahlen a, b schreiben wir $a \leq b$ genau dann, wenn $b - a \in \mathbb{N}$ ist (Sprechweise: a ist kleiner oder gleich b). Wir schreiben $a < b$ für: $a \leq b$ und $a \neq b$ (Sprechweise: a ist kleiner als b).
Eine ganze Zahl ist *positiv* bzw. *negativ*, wenn sie größer bzw. kleiner als 0 ist. Statt $a \cdot b$ schreibt man oft nur ab.

Es seien a und b ganze Zahlen, wobei $b \neq 0$ ist. Die Aufgabe „Finde eine Zahl z so, dass $b \cdot z = a$ ist" bezeichnen wir als „die Gleichung $b \cdot x = a$". Eine Zahl z mit $b \cdot z = a$ heißt *Lösung* von $b \cdot x = a$. Wenn $|b| \neq 1$ ist, dann hat die Aufgabe $b \cdot x = 1$ in $\mathbb{Z}$ keine Lösung. Um Lösungen zu erhalten, müssen wir „den Zahlenbereich erweitern".

Die Aufgabe $b \cdot x = a$ wird durch das Paar $(a, b) \in \mathbb{Z}^2$ eindeutig beschrieben, also liegt es nahe, die „neuen Zahlen" durch Paare von ganzen Zahlen zu beschreiben. Allerdings sollten für $t \in \mathbb{Z}$, $t \neq 0$, die Gleichungen $b \cdot x = a$ und $t \cdot b \cdot x = t \cdot a$ dieselbe Lösung haben, daher sollen die Zahlenpaare (a, b) und $(t \cdot a, t \cdot b)$ dieselbe „neue Zahl" beschreiben.

Definition 12 : Es seien a und b ganze Zahlen, wobei $b \neq 0$. Dann ist die Menge

$$\frac{a}{b} := \{(c, d) \mid c, d \in \mathbb{Z}, ad = bc, d \neq 0\}$$

die durch den „Zähler" a und den „Nenner" b gegebene *rationale Zahl* oder *Bruchzahl*. (Beachte: Eine rationale Zahl ist durch Vorgabe von Zähler und Nenner eindeutig bestimmt, aber umgekehrt sind Zähler und Nenner durch die rationale Zahl nicht eindeutig bestimmt). Wir schreiben $\mathbb{Q}$ für die Menge der rationalen Zahlen.
Für die Bruchzahl $\frac{a}{1}$ schreiben wir oft nur a und fassen so $\mathbb{Z}$ als Teilmenge von $\mathbb{Q}$ auf. („Jede ganze Zahl ist eine rationale Zahl").

Satz 13 : *Es seien a', b' ganze Zahlen und $b' \neq 0$. Dann sind die Bruchzahlen $\frac{a}{b}$ und $\frac{a'}{b'}$ genau dann gleich, wenn $a \cdot b' = a' \cdot b$ ist.*

Beweis: Wenn $\frac{a}{b} = \frac{a'}{b'}$ ist, dann ist insbesondere $(a', b') \in \frac{a}{b}$, also $a \cdot b' = a' \cdot b$.
Sei umgekehrt $a \cdot b' = a' \cdot b$ und $(c, d) \in \frac{a}{b}$, also $b \cdot c = a \cdot d$. Dann ist zu zeigen, dass $(c, d) \in \frac{a'}{b'}$, also $b' \cdot c = a' \cdot d$ ist.
Es ist

$$b \cdot (b' \cdot c) = (b \cdot c) \cdot b' = (a \cdot d) \cdot b' = d \cdot (a \cdot b') = d \cdot (a' \cdot b) = b \cdot (a' \cdot d) \ ,$$

wegen $b \neq 0$ also auch $b' \cdot c = a' \cdot d$.

Wir werden nun die Rechenoperationen von $\mathbb{Z}$ auf $\mathbb{Q}$ fortsetzen.

Satz 14 : *Die Funktionen*

$$+ : \mathbb{Q} \times \mathbb{Q} \longrightarrow \mathbb{Q} \ , \quad (\frac{a}{b}, \frac{c}{d}) \longmapsto \frac{a}{b} + \frac{c}{d} := \frac{ad + bc}{bd} \ ,$$

und

$$\cdot : \mathbb{Q} \times \mathbb{Q} \longrightarrow \mathbb{Q} \ , \quad (\frac{a}{b}, \frac{c}{d}) \longmapsto \frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd} \ ,$$

sind wohldefiniert. Diese Rechenoperationen in $\mathbb{Q}$ erfüllen die gleichen Rechenregeln wie Addition und Multiplikation in $\mathbb{Z}$. Darüberhinaus hat jedes Element $\frac{a}{b} \in \mathbb{Q} \setminus \{0\}$ ein inverses Element $(\frac{a}{b})^{-1}$ *mit der Eigenschaft*

$$(\frac{a}{b})^{-1} \cdot \frac{a}{b} = 1 \ ,$$

und zwar ist

$$(\frac{a}{b})^{-1} = \frac{b}{a} \ .$$

Die Einschränkungen von $+$ und $\cdot$ auf $\mathbb{Z} \times \mathbb{Z}$ stimmen mit der Addition und der Multiplikation auf $\mathbb{Z}$ überein.

Beweis: Wir müssen zuerst zeigen, dass die Funktionen $+$ und $\cdot$ wohldefiniert sind, das heißt: wenn $\frac{a}{b} = \frac{a'}{b'}$ und $\frac{c}{d} = \frac{c'}{d'}$ ist, dann muss auch

$$\frac{ad + bc}{bd} = \frac{a'd' + b'c'}{b'd'} \quad \text{und} \quad \frac{ac}{bd} = \frac{a'c'}{b'd'}$$

sein.
Aus $a'b = ab'$ und $c'd = cd'$ folgt

$$(ad + bc)b'd' = ab'dd' + bb'cd' = a'bdd' + bb'c'd = bd(a'd' + b'c')$$

und

$$(ac)b'd' = bd(a'c') \ .$$

Die Rechenregeln können leicht nachgeprüft werden.

§4. Zusammengesetzte Aussagen

Wir betrachten Aussagen $A, B, C, \dots$, die nach Vereinbarung entweder wahr oder falsch sind. Mit Hilfe der Worte

„und"	(Zeichen: $\wedge$) ,
„oder"	(Zeichen: $\vee$) ,
„nicht"	(Zeichen: $\neg$) ,
„wenn, dann"	(Zeichen: $\Rightarrow$) ,
„genau dann, wenn"	(Zeichen: $\Leftrightarrow$)

bilden wir *zusammengesetzte* Aussagen, deren „Wahrheitswert" wir durch die folgende Tabelle definieren. Dabei steht w für „wahr" und f für „falsch".

A	B	$A \wedge B$	$A \vee B$	$\neg A$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	w	w	f	w	w
w	f	f	w	f	f	f
f	w	f	w	w	w	f
f	f	f	f	w	w	w

Für $A \Rightarrow B$ verwendet man statt „wenn A, dann B" auch die Sprechweisen „aus A folgt B" oder „A impliziert B".
Man beachte:
A ist genau dann wahr, wenn $\neg A$ falsch ist. Das wird für *indirekte Beweise* verwendet: anstatt zu zeigen, dass eine Aussage A wahr ist, wird gezeigt, dass ihr „Gegenteil" $\neg A$ falsch ist.
In der Mathematik bedeutet das Wort „oder" immer das nicht ausschließende „und-oder" und nicht das ausschließende „entweder-oder".
Ist A falsch, dann ist die Aussage $A \Rightarrow B$ immer wahr („ex falso quodlibet").

§5. Der Induktionsbeweis

Sei m eine natürliche Zahl (meistens 0 oder 1) und sei $(A_m, A_{m+1}, A_{m+2}, \dots)$ eine Folge von Aussagen.

Satz 15 : *Wenn*

(1) A_m *wahr ist und*
(2) *für alle* $n > m$ *aus* A_{n-1} *auch* A_n *folgt,*

dann sind alle Aussagen A_n , $n \geq m$, *wahr.*

Damit erhält man eine Methode zu zeigen, dass alle Aussagen A_n, $n \geq m$, wahr sind. („*Beweis durch vollständige Induktion*"): Es genügt zu zeigen, dass (1) (*„Induktionsanfang"*) und (2) (*„Induktionsschluss"*) richtig sind. Um zu zeigen, dass (2) richtig ist, nimmt man an, dass A_{n-1} wahr ist (*„Induktionsannahme"*) und versucht damit zu zeigen, dass auch A_n wahr ist.

Man könnte im Satz die Annahme (2) auch durch

(2') für alle $n > m$ aus $A_m, A_{m+1}, \ldots, A_{n-1}$ auch A_n folgt,

ersetzen.

Beweis: Wir benutzen die folgende Eigenschaft der natürlichen Zahlen: jede nicht leere Teilmenge von $\mathbb{N}$ hat ein kleinstes Element. Wir führen den Beweis indirekt und nehmen an, dass nicht alle Aussagen A_n , $n \geq m$, wahr sind. Dann ist die Menge

$$M := \{n \in \mathbb{N} \mid n \geq m \text{ und } A_n \text{ ist falsch}\}$$

nicht leer. Daher gibt es eine kleinste Zahl k so, dass $k \geq m$ und A_k falsch ist. Wegen (1) gilt $k \geq m+1$, also $k-1 \geq m$. Weiters muss A_{k-1} wahr sein, weil k die kleinste Zahl in M ist. Aus (2) folgt nun, dass auch A_k wahr ist, was im Widerspruch zur Wahl von k steht. Somit muss unsere Annahme am Anfang des Beweises falsch sein, d.h. alle Aussagen A_n , $n \geq m$, sind wahr.

Satz 16 : *Sei n eine natürliche Zahl. Die Summe der Quadrate aller natürlichen Zahlen von* 1 *bis n ist* $S(n) := \frac{1}{6}n(n+1)(2n+1)$.

Beweis: Induktionsanfang: $S(1) = \frac{1}{6}(2+3+1) = 1 = 1^2$, also ist die Aussage für $n = 1$ wahr.
Induktionsschluss: Wir nehmen an, dass die Summe der Quadrate aller natürlichen Zahlen von 1 bis $n-1$ gleich $S(n-1)$ ist. Die Summe der Quadrate aller natürlichen Zahlen von 1 bis n ist dann

$$S(n-1) + n^2 \ .$$

Wegen

$$S(n-1) + n^2 = \frac{1}{6}(n-1)n(2n-1) + n^2 = S(n)$$

ist die Behauptung richtig.

§6. Zifferndarstellung von Zahlen

Wenn Sie einen Sack mit a Euromünzen haben, die Sie an b Personen verteilen sollen (jede soll gleich viel bekommen), dann werden Sie wahrscheinlich zuerst jeder Person einen Euro geben und diesen Vorgang solange wiederholen, bis im Sack weniger als b Euromünzen sind. Sie haben dann *a mit Rest durch b dividiert*.
Der folgende Satz ist grundlegend für alle Rechenverfahren für ganze Zahlen. Seine Bedeutung liegt darin, dass die drei „Strukturen" $+$, $\cdot$ und $\leq$ zueinander in Beziehung gesetzt werden.

Satz 17 : (Division mit Rest von ganzen Zahlen)
Zu je zwei ganzen Zahlen a und b mit $b \neq 0$ gibt es eindeutig bestimmte ganze Zahlen m und r mit den Eigenschaften

$$a = m \cdot b + r \quad \text{und} \quad 0 \leq r < |b| \ .$$

Die Zahlen m bzw. r heißen ganzzahliger Quotient von a und b *bzw.* Rest von a nach Division durch b. *Die Zahlen m und r können mit dem folgenden Verfahren (*Divisionsalgorithmus*) berechnet werden:*

- *Falls a und b natürliche Zahlen sind:*
  Setze $m := 0$ und $r := a$.
  Solange $r \geq b$ ist, ersetze r durch $r - b$ und m durch $m + 1$.
- *Falls $a < 0$ oder $b < 0$ ist:*
  Berechne wie oben n und s so, dass $|a| = n \cdot |b| + s$ und $0 \leq s < |b|$ ist.
  Wenn $a \geq 0$ ist, dann setze $m := vz(b) \cdot n$ und $r := s$.
  Wenn $a < 0$ und $s > 0$ ist, dann setze $m := -vz(b) \cdot (n+1)$ und $r := |b| - s$.
  Wenn $a < 0$ und $s = 0$ ist, dann setze $m := -vz(b) \cdot n$ und $r := 0$.

Beweis: Wenn a und b natürliche Zahlen sind, dann erhalten wir bei jedem Ersetzen von r durch $r - b$ eine um mindestens 1 kleinere Zahl. Also tritt nach höchstens a Schritten der Fall $r < b$ ein. Somit liefert das obige Verfahren nach endlich vielen Schritten ein Ergebnis m, r. Mit Induktion über $|a|$ ist leicht nachzuprüfen, dass diese Zahlen die angegebenen Bedingungen erfüllen.
Es seien m_1, m_2, r_1, r_2 ganze Zahlen mit $a = m_1 \cdot b + r_1 = m_2 \cdot b + r_2$, $0 \leq r_1, r_2 < |b|$ und o.E.d.A. („ohne Einschränkung der Allgemeinheit“) $r_1 \leq r_2$. Dann ist

$$|b| > r_2 - r_1 = |m_1 - m_2| \cdot |b| \ .$$

Daraus folgt $m_1 = m_2$ und $r_1 = r_2$, also sind der ganzzahlige Quotient von a und b und der Rest von a nach Division durch b eindeutig bestimmt.

Nehmen wir an, Sie kommen mit einem Sack voller Euromünzen in eine Bank und wollen dieses Geld auf ihr Sparbuch einzahlen. Die Anzahl der Euromünzen im Sack ist eine eindeutig bestimmte natürliche Zahl a. Bevor diese Zahl in Ihr Sparbuch eingetragen werden kann, muss der Bankbeamte ihre *Zifferndarstellung* (zur Basis 10) berechnen. Eine Zahl ist also nicht immer schon in Zifferndarstellung gegeben, sondern diese ist eine „Zusatzinformation“ über die Zahl. Wie wird die Zifferndarstellung zur Basis 10 von a ermittelt? Man bildet aus den Euromünzen solange „Zehnerstapel“, bis nur noch weniger als zehn Münzen übrigbleiben, das heißt: a wird mit Rest durch 10 dividiert. Die Anzahl der übriggebliebenen Euromünzen ist dann die „Einerziffer“ von a. Macht man dasselbe nun mit den Zehnerstapeln statt mit den Münzen, dann erhält man die „Zehnerziffer“ von a, usw.

Satz 18 : (Darstellung von Zahlen durch Ziffern)
Es seien a und b natürliche Zahlen mit $a \neq 0$ und $b \geq 2$. Dann gibt es eindeutig bestimmte natürliche Zahlen $n, z_0, z_1, \dots, z_n$ so, dass

$$z_n \neq 0,\ 0 \leq z_0, z_1, \dots, z_n < b$$

und

$$a = z_n b^n + z_{n-1} b^{n-1} + \ldots + z_1 b^1 + z_0 = \sum_{i=0}^{n} z_i b^i$$

ist.
Wenn b fest gewählt ist, dann ist a durch die Zahlen $n, z_0, z_1, \dots, z_n$ eindeutig bestimmt. Man wählt Zeichen für die Zahlen von 0 bis $b-1$ und schreibt dann

$$z_n z_{n-1} \cdots z_0 \quad \textit{statt} \quad \sum_{i=0}^{n} z_i b^i \quad .$$

Die Zahlen $z_0, z_1, \dots, z_n$ heißen Ziffern von a zur Basis b *(für b=2 bzw. 10: „Binärziffern“ bzw. „Dezimalziffern“).*

Die Ziffern z_i von $a \neq 0$ zur Basis b können mit dem folgenden Verfahren berechnet werden:

- *Setze $i := 0$.*
- *Solange a nicht 0 ist: Die i-te Ziffer z_i ist der Rest von a nach Division durch b. Ersetze a durch den ganzzahligen Quotienten von a und b. Ersetze i durch $i+1$.*

Beweis: Induktion über a:
Wenn $a < b$ ist, ist $n = 0$ und $z_0 = a$.
Für $a \geq b$ seien m bzw. r der ganzzahlige Quotient von a und b bzw. der Rest von a nach Division durch b. Wegen $b > 1$ ist $m < a$ und wegen $a \geq b$ ist $m > 0$, also gibt es nach Induktionsannahme eindeutig bestimmte Zahlen $k, y_0, y_1, \dots, y_k$ so, dass $y_k \neq 0$, $0 \leq y_0, y_1, \dots, y_k < b$ und

$$m = y_k b^k + y_{k-1} b^{k-1} + \ldots + y_1 b^1 + y_0$$

ist. Dann ist

$$a = m \cdot b + r = y_k b^{k+1} + y_{k-1} b^k + \ldots + y_1 b^2 + y_0 b + r\ ,$$

und $y_k, \dots, y_0, r$ sind die Ziffern von a. Wegen der Eindeutigkeit von m und r folgt aus der Induktionsannahme die Eindeutigkeit der Ziffern von a zur Basis b.

Wird für die Zifferndarstellung einer Zahl die Basis b gewählt, dann können alle Zahlen durch Aneinanderreihen von b verschiedenen Symbolen angeschrieben werden. Eine kleine Basis (zum Beispiel 2) hat den Vorteil, dass man nur wenige Symbole braucht und dass das „kleine Einmaleins“ sehr einfach ist. Allerdings braucht man dann für größere Zahlen sehr viele Ziffern.

Beispiel 19 : Die Zifferndarstellung des heurigen Jahres zur Basis zehn ist 2013, zur Basis zwei aber 11111011101.
Denn: $2013 = 1006 \cdot 2 + 1, 1006 = 503 \cdot 2 + 0, 503 = 251 \cdot 2 + 1, \ \ldots$.

Definition 20 : Es seien $v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n)$ zwei verschiedene n-Tupel von ganzen Zahlen und j die kleinste Zahl in $\{1, \ldots, n\}$ mit der Eigenschaft, dass $v_j \neq w_j$ ist.
Dann ist v genau dann *lexikographisch kleiner* als w, wenn $v_j < w_j$ ist (Schreibweise: $v <_{lex} w$).

Beispiel 21 : $(1,2,3,4) <_{lex} (1,2,4,3) <_{lex} (2,-7,-3,-5)$

Satz 22 : (Vergleich von zwei Zahlen, die durch Ziffern dargestellt sind) *Es seien b, x, y positive natürliche Zahlen, $b \geq 2$ und*

$$x_k, x_{k-1}, \ldots, x_0 \quad \textit{bzw.} \quad y_\ell, y_{\ell-1}, \ldots, y_0$$

die Ziffern von x bzw. y bezüglich b.
Dann ist x genau dann kleiner als y, wenn

$$k < \ell \quad \textit{oder} \quad (k = \ell \textit{ und } (x_k, x_{k-1}, \ldots, x_0) <_{lex} (y_\ell, y_{\ell-1}, \ldots, y_0)) \textit{ ist.}$$

Beweis: Wenn $k < \ell$ ist, dann ist

$$x = \sum_{i=0}^{k} x_i b^i \leq \sum_{i=0}^{k} (b-1) b^i = \sum_{i=1}^{k+1} b^i - \sum_{i=0}^{k} b^i = b^{k+1} - 1 < b^{k+1} \leq y \,.$$

Es seien $k = \ell$ und j die größte Zahl mit der Eigenschaft, dass $x_j \neq y_j$ ist.
Wenn $x_j < y_j$ ist, dann ist

$$\sum_{i=0}^{j} x_i b^i \leq x_j b^j + (b^j - 1) < (x_j + 1) b^j \leq y_j b^j \leq \sum_{i=0}^{j} y_i b^i$$

und

$$x = \sum_{i=j+1}^{k} x_i b^i + \sum_{i=0}^{j} x_i b^i < \sum_{i=j+1}^{k} x_i b^i + \sum_{i=0}^{j} y_i b^i = y \,.$$

Wenn zur Darstellung einer Zahl am Computer 32 bits (also 32 Binärziffern) zur Verfügung stehen, dann können in der *Zweierkomplementdarstellung* die Zahlen in

$$\{-2^{31} = -2147483648, \ldots, -1, 0, 1, \ldots,\ 2^{31} - 1 = 2147483647\}$$

(also insgesamt 2^{32} Zahlen) dargestellt werden.
Ist a eine natürliche Zahl in diesem Zahlenbereich, dann wird a durch

$$0\ a_{30}\ a_{29}\ \ldots\ a_1\ a_0$$

dargestellt, wobei $a_{30}\ a_{29}\ \ldots\ a_1\ a_0$ die Ziffern von a zur Basis 2 sind.
Ist a eine negative Zahl in diesem Zahlenbereich, dann wird a durch

$$1\ a_{30}\ a_{29}\ \ldots\ a_1\ a_0$$

dargestellt, wobei $1\ a_{30}\ a_{29}\ \ldots\ a_1\ a_0$ die Ziffern von $a + 2^{32}$ zur Basis 2 sind.

Satz 23: (Zifferndarstellung von rationalen Zahlen)
Es seien b, c, d, p positive ganze Zahlen mit $b \geq 2$. Dann gibt es eindeutig bestimmte natürliche Zahlen $n, z_n, \ldots, z_0, z_{-1}, \ldots, z_{-p}$ so, dass

$$z_n \neq 0 \text{ oder } n = 0, \qquad 0 \leq z_n, \ldots, z_0, z_{-1}, \ldots, z_{-p} < b$$

und

$$0 \leq \frac{c}{d} - (z_n b^n + z_{n-1} b^{n-1} + \ldots + z_1 b^1 + z_0 + z_{-1} b^{-1} + \ldots + z_{-p} b^{-p}) < b^{-p}$$

ist. Ist b fest gewählt, schreibt man

$$z_n z_{n-1} \ldots z_0 . z_{-1} z_{-2} \ldots z_{-p} \quad \text{statt} \quad \sum_{i=-p}^{n} z_i b^i \ .$$

Die Zahlen $z_n, \ldots, z_0, z_{-1}, \ldots, z_{-p}$ heißen Ziffern von $\frac{c}{d}$ zur Basis b. *Die Ziffern z_i von $\frac{c}{d}$ zur Basis b können wie folgt berechnet werden:*

- *Berechne (mit Satz 18) die Ziffern $y_0, \ldots, y_k$ zur Basis b des ganzzahligen Quotienten m von $c \cdot b^p$ und d .*
- *Setze $z_i := y_{i+p}$, $-p \leq i \leq k - p =: n$.*

Beweis: Sei r der Rest von $c \cdot b^p$ nach Division durch d. Wegen $c \cdot b^p = m \cdot d + r$ ist dann

$$\frac{c \cdot b^p}{d \cdot b^p} = \frac{m \cdot d}{d \cdot b^p} + \frac{r}{d \cdot b^p}\ ,$$

also

$$\frac{c}{d} = \frac{m}{b^p} + \frac{r}{d} \cdot b^{-p}\ .$$

Wegen $0 \leq r < d$ ist

$$\frac{r}{d} < 1 \quad \text{und} \quad \frac{c}{d} - \frac{m}{b^p} < b^{-p}\ .$$

Rationale Zahlen können also „beliebig genau“ durch Zahlen der Form $z_n z_{n-1} \dots z_0 . z_{-1} z_{-2} \dots z_{-p}$ angenähert werden, aber es gibt rationale Zahlen, die für alle p von $z_n z_{n-1} \dots z_0 . z_{-1} z_{-2} \dots z_{-p}$ verschieden sind.

Eine rationale Zahl

$$z_0 . z_{-1} z_{-2} \dots z_{-p} \, Ee \; := \; z_0 . z_{-1} z_{-2} \dots z_{-p} \cdot b^e$$

mit $b \geq 2$ und $z_0 \neq 0$ ist in *Exponentialform* oder *Gleitkommaform* zur Basis b dargestellt. Die Zahlen e und $z_0 . z_{-1} z_{-2} \dots z_{-p}$ heißen *Exponent* und *Mantisse*.

Am Computer kann eine Zahl durch die Ziffern des Exponenten und der Mantisse zur Basis 2 dargestellt werden. Die Anzahl dieser Ziffern ist durch eine vorgegebene Zahl beschränkt. Die so am Computer verfügbaren Zahlen heißen *Maschinenzahlen*. Es gibt nur endlich viele Maschinenzahlen, alle Maschinenzahlen sind rationale Zahlen.

Beim Rechnen mit so dargestellten Zahlen gibt es im allgemeinen keine exakten Ergebnisse, sondern Rundungsfehler. Bei Rechenverfahren muss daher darauf geachtet werden, dass sich die Fehler nicht akkumulieren. Fehlerabschätzungen sind erforderlich.

Beispiel 24 : Die Zahl 0.1 (Dezimaldarstellung) auf der Tastatur wird vom Computer in Binärdarstellung 0.00011001100110011001100... umgewandelt und zum Beispiel als

$$1.1001100110011001100110011001100110011001100 \, E-4$$

gespeichert. Also ergibt schon die Eingabe von 0.1 einen Rundungsfehler!

Will man mit rationalen Zahlen am Computer exakt rechnen, kann man $\frac{a}{b}$ als Zahlenpaar (a,b) eingeben. Dann müssen für Zahlenpaare die Rechenoperationen

$$(a,b)+(c,d) := (ad+bc,bd) \quad \text{und} \quad (a,b)\cdot(c,d) := (ac,bd)$$

definiert werden.

§7. Gruppen, Ringe und Körper

Definition 25 : Sei G eine Menge und $* : G \times G \to G$ eine Funktion. Für Elemente $a,b \in G$ schreiben wir statt $*(a,b)$ kurz $a*b$. Das Paar $(G,*)$ heißt eine *Gruppe*, wenn die folgenden drei Bedingungen („Gruppen-Axiome“) erfüllt sind:

(1) Für alle Elemente $a,b,c \in G$ ist
$a*(b*c) = (a*b)*c =: a*b*c$ (*Assoziativgesetz*).

(2) Es gibt ein Element $e \in G$ so, dass für alle $a \in G$ gilt :
$a*e = e*a = a$ (e heißt dann *neutrales Element* in G).

(3) Für alle Elemente $a \in G$ gibt es ein $b \in G$ so, dass $a * b = b * a = e$ ist (b heißt dann zu a *inverses Element* und wird mit a^{-1} bezeichnet).

Eine Gruppe $(G, *)$ heißt *kommutativ* oder *abelsch*, wenn zusätzlich gilt:

(4) Für alle $a, b \in G$ ist $a * b = b * a$ (*Kommutativgesetz*).

Ist $(G, *)$ eine Gruppe, dann wird die Funktion $*$ als *Gruppenverknüpfung*, *Multiplikation* oder, wenn $(G, *)$ abelsch ist, als *Addition* bezeichnet. Wenn aus dem Zusammenhang ersichtlich ist, welche Verknüpfung auf G betrachtet wird, schreibt man statt $(G, *)$ kürzer G.

Beispiel 26 : $(\mathbb{Z}, +)$, $(\{1, -1\}, \cdot)$, $(\mathbb{Q}, +)$ und $(\mathbb{Q} \setminus \{0\}, \cdot)$ sind kommutative Gruppen.

Satz 27 : *Seien $(G, *)$ eine Gruppe und $a, b, c \in G$. Dann gilt:*

(1) *Es gibt genau ein neutrales Element in G.*
(2) *Zu jedem Element in G gibt es genau ein inverses Element in G.*
(3) *Es ist* $(a * b)^{-1} = b^{-1} * a^{-1}$.
(4) *Aus* $a * b = a * c$ *oder* $b * a = c * a$ *folgt* $b = c$ *(„In einer Gruppe kann gekürzt werden“).*

Beweis:

(1) Seien e und e' neutrale Elemente in G. Dann ist $e' = e * e'$ und $e = e * e'$, also $e = e'$.
(2) Seien b und b' zu a inverse Elemente. Dann ist $b = e * b = (b' * a) * b = b' * (a * b) = b' * e = b'$.
(3) Es ist $(a * b) * (b^{-1} * a^{-1}) = a * (b * (b^{-1} * a^{-1})) =$
$= a * ((b * b^{-1}) * a^{-1}) = a * (e * a^{-1}) = a * a^{-1} = e$ und
$(b^{-1} * a^{-1}) * (a * b) = b^{-1} * (a^{-1} * (a * b)) = b^{-1} * ((a^{-1} * a) * b) =$
$b^{-1} * (e * b) = b^{-1} * b = e$.

(4) Aus $a * b = a * c$ folgt $b = a^{-1} * a * b = a^{-1} * a * c = c$.

Definition 28 : Seien R eine Menge und $+ : R \times R \to R$ sowie $\cdot : R \times R \to R$ Funktionen. Wir schreiben statt $+(a, b)$ kurz „$a + b$“ und statt $\cdot(a, b)$ kurz „$a \cdot b$“ oder „ab“. Das Tripel $(R, +, \cdot)$ heißt ein *Ring*, wenn die folgenden Bedingungen („Ring-Axiome“) erfüllt sind:

(1) $(R, +)$ ist eine abelsche Gruppe.
(2) Für alle $a, b, c \in R$ ist $(ab)c = a(bc)$ (*Assoziativgesetz*).
(3) Es gibt ein Element $e \in R$ so, dass für alle $a \in R$ gilt : $ea = ae = a$ (e heisst dann *Einselement* und wird mit 1_R bezeichnet).

(4) Für alle $a,b,c \in R$ ist $a(b+c) = (ab)+(ac)$ und
$(a+b)c = (ac)+(bc)$ (*Distributivgesetz*).

Ein Ring $(R,+,\cdot)$ heißt *kommutativ*, wenn zusätzlich gilt:

(5) Für alle $a,b \in R$ ist $ab = ba$ (*Kommutativgesetz*).

Ist $(R,+,\cdot)$ ein Ring, dann heißt $+$ die *Addition* und $\cdot$ die *Multiplikation* des Ringes. Das neutrale Element von $(R,+)$ heißt *Nullelement* und wird 0_R geschrieben. Das zu $a \in R$ bezüglich $+$ inverse Element wird mit $-a$ bezeichnet. Die *Subtraktion* ist dann definiert durch

$$a - b := a + (-b).$$

Um Klammern einzusparen, wird verabredet, dass die Multiplikation immer vor der Addition ausgeführt wird, ausgenommen bei gegenteiliger Klammerung. Zum Beispiel wird $(ab)+c$ abgekürzt als $ab+c$.

Wenn aus dem Zusammenhang ersichtlich ist, welche Addition und Multiplikation auf der Menge R betrachtet werden, so schreibt man statt $(R,+,\cdot)$ kurz R.

Beispiel 29: $(\mathbb{Z},+,\cdot)$ und $(\mathbb{Q},+,\cdot)$ sind kommutative Ringe.

Definition 30: Ein Element a eines Ringes R mit Einselement 1_R ist *invertierbar*, wenn es ein Element $b \in R$ mit

$$ab = 1_R = ba$$

gibt. Das Element b heißt dann zu a (bezüglich $\cdot$) *inverses Element* und wird mit a^{-1} bezeichnet.

Satz 31: *Die Menge aller invertierbaren Elemente eines Ringes R ist mit der Multiplikation von R eine Gruppe. Das Einselement von R ist das neutrale Element dieser Gruppe. Für invertierbare Elemente $a,b \in R$ ist*

$$(ab)^{-1} = b^{-1}a^{-1} \quad \textit{und} \quad (a^{-1})^{-1} = a\,.$$

Beweis: Es ist

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = 1 = b^{-1}(a^{-1}a)b = (b^{-1}a^{-1})(ab)\,.$$

Definition 32: Sei $(R,+,\cdot)$ ein kommutativer Ring mit mindestens zwei Elementen. R heißt ein *Körper*, wenn jedes Element von $R \setminus \{0\}$ invertierbar ist. Die *Division* in R ist dann durch

$$a/b := ab^{-1}$$

definiert.

Beispiel 33 : $(\mathbb{Q},+,\cdot)$ ist ein Körper. Der Ring $(\mathbb{Z},+,\cdot)$ der ganzen Zahlen ist kein Körper.

Als Merkregel für diese Definitionen gilt:
In einem Ring kann addiert, subtrahiert und multipliziert werden. In einem Körper kann zusätzlich noch durch Elemente ungleich Null dividiert werden. Die Ring-Axiome sind den Rechenregeln für ganze Zahlen nachgebildet.

Satz 34 : *Die Menge* $\{0,1\}$ *mit der Addition*

$$0+0=0,\ 0+1=1+0=1,\ 1+1=0$$

und der Multiplikation

$$0\cdot 0=0\cdot 1=1\cdot 0=0,\ 1\cdot 1=1$$

ist ein Körper mit Nullelement 0 und Einselement 1.
Dieser Körper heißt binärer Körper *und wird mit* $\mathbb{Z}_2$ *(sprich „Z modulo 2“) bezeichnet.*

Beweis: Übung.

Satz 35 : *Seien* $(R,+,\cdot)$ *ein Ring und* $a,b,c \in R$*. Dann gilt:*

(1) *Aus* $a+b=a+c$ *folgt* $b=c$.
(2) $0_R \cdot a = a\cdot 0_R = 0_R$
(3) $(-a)\cdot b = a\cdot(-b) = -(a\cdot b)$
(4) $(-a)\cdot(-b) = a\cdot b$.

Wenn R ein Körper und $a \neq 0$ *ist, dann gilt:*

(5) *Aus* $a\cdot b = a\cdot c$ *folgt* $b=c$.

Beweis:

(1) folgt nach Satz 27 durch Kürzen in der Gruppe $(R,+)$.
(2) Aus $0_R + 0_R\cdot a = 0_R\cdot a = (0_R+0_R)\cdot a = 0_R\cdot a + 0_R\cdot a$ folgt nach (1), dass $0_R = 0_R\cdot a$ ist. Analog beweist man die zweite Behauptung.

(3) Aus $(-a)\cdot b + a\cdot b = (-a+a)\cdot b = 0_R\cdot b = 0_R$ folgt $(-a)\cdot b = -(a\cdot b)$. Analog beweist man die zweite Behauptung.
(4) Nach (3) ist $(-a)\cdot(-b) = -(a\cdot(-b)) = -(-(a\cdot b)) = a\cdot b$.
(5) Aus $a\cdot b = a\cdot c$ erhält man $a^{-1}\cdot a\cdot b = a^{-1}\cdot a\cdot c$ und schließlich $b=c$.

§8. Rechnen mit Summen und Produkten

Definition 36: Seien $(R,+,\cdot)$ ein Ring, n eine positive ganze Zahl und $a_1,a_2,\ldots,a_n \in R$. Dann sei

$$\sum_{i=1}^{n} a_i := a_1 + a_2 + \ldots + a_n := (\ldots((a_1+a_2)+a_3)+\ldots+a_{n-1})+a_n\,.$$

Sprechweise: Die Summe aller a_i mit i von 1 bis n. Die Elemente $a_1,a_2,\ldots,a_n$ heißen *Summanden* von $\sum_{i=1}^{n} a_i$.
Für $n=1$ ist

$$\sum_{i=1}^{1} a_i = a_1\,.$$

Es sei

$$\prod_{i=1}^{n} a_i := a_1 \cdot a_2 \cdot \ldots \cdot a_n := (\ldots((a_1 \cdot a_2)\cdot a_3)\cdot \ldots \cdot a_{n-1})\cdot a_n\,.$$

Sprechweise: Das Produkt aller a_i mit i von 1 bis n. Die Elemente $a_1,a_2,\ldots,a_n$ heißen *Faktoren* von $\prod_{i=1}^{n} a_i$.
Für $n=1$ ist

$$\prod_{i=1}^{1} a_i = a_1\,.$$

Durch Induktion über n kann man zeigen, dass es beim Bilden einer Summe oder eines Produktes nicht auf die Reihenfolge des Ausführens der Rechenoperationen $+$ bzw. $\cdot$ ankommt, die Klammern können also weggelassen werden („Allgemeines Assoziativgesetz“).

Definition 37: Eine Funktion $f: M \to N$ heißt *bijektiv*, wenn jedes Element von N genau ein Urbild hat.
Eine Menge M heißt *endlich*, wenn sie leer ist oder es eine positive ganze Zahl n und eine bijektive Funktion $f:\{1,\ldots,n\} \to M$ gibt. In diesem Fall nennt man

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

die *Anzahl der Elemente* von M. Wenn M leer ist, sei

$$\#(\emptyset) := 0\,.$$

(„Die leere Menge hat 0 Elemente“).
M heißt *unendlich*, wenn M nicht endlich ist.

Satz 38: *(*Allgemeines Kommutativgesetz*) Seien $(R,+,\cdot)$ ein kommutativer Ring, n eine positive ganze Zahl und $a_1, a_2, \ldots, a_n \in R$ und f eine bijektive Funktion von $\{1,\ldots,n\}$ nach $\{1,\ldots,n\}$. Dann gilt*

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a_{f(i)},$$

und

$$\prod_{i=1}^{n} a_i = \prod_{i=1}^{n} a_{f(i)},$$

d.h. auf die Reihenfolge der Summanden bzw. Faktoren kommt es bei der Berechnung von $\sum_{i=1}^{n} a_i$ bzw. $\prod_{i=1}^{n} a_i$ nicht an, sie können beliebig umgeordnet werden. werden.

Beweis: Induktion über n.

Definition 39: Sei I eine endliche Menge, $(R,+,\cdot)$ ein kommutativer Ring und $(a_i)_{i\in I}$ eine Familie von Elementen in R. Wenn I nicht leer ist, sei

$$\sum_{i\in I} a_i := \sum_{i=1}^{n} a_{f(i)} \in R$$

und

$$\prod_{i\in I} a_i := \prod_{i=1}^{n} a_{f(i)} \in R,$$

wobei f eine bijektive Funktion von $\{1,2,\ldots,n := \#(I)\}$ nach I ist.
Nach Satz 38 hängt $\sum_{i\in I} a_i$ nicht von der Wahl der bijektiven Funktion f ab.
Wenn I leer ist, vereinbaren wir

$$\sum_{i\in\emptyset} a_i := 0_R$$

und

$$\prod_{i\in\emptyset} a_i := 1_R.$$

Beispiel 40: Sei $I := \{(0,2),(1,1),(2,0)\}$, $(R,+,\cdot) := (\mathbb{Z},+,\cdot)$ und $a_{(0,2)} := 3, a_{(1,1)} := 4, a_{(2,0)} := 1$. Sei f die bijektive Funktion von $\{1,2,3\}$ nach I mit $f(1) = (1,1), f(2) = (0,2), f(3) = (2,0)$. Dann ist

$$\sum_{i\in I} a_i = a_{(1,1)} + a_{(0,2)} + a_{(2,0)} = 4+3+1 = 8.$$

und

$$\prod_{i\in I} a_i = a_{(1,1)} \cdot a_{(0,2)} \cdot a_{(2,0)} = 4\cdot 3\cdot 1 = 12.$$

Beispiel 41 : Seien $m,n \in \mathbb{Z}$, $m < n$, $I := \{m, m+1, \ldots, n\}$ und $a_m, a_{m+1}, \ldots, a_n \in R$. Die Funktion

$$f : \{1, \ldots, n-m+1\} \longrightarrow I\,,\ j \longmapsto j+m-1\,,$$

ist bijektiv, also ist

$$\sum_{i \in I} a_i = \sum_{i=1}^{n-m+1} a_{f(i)} = a_m + a_{m+1} + \ldots + a_n\,.$$

In diesem Fall schreibt man oft

$$\sum_{i=m}^{n} a_i \quad \text{statt} \quad \sum_{i \in I} a_i$$

und analog

$$\prod_{i=m}^{n} a_i \quad \text{statt} \quad \prod_{i \in I} a_i\,.$$

Satz 42 : *(*Allgemeines Distributivgesetz*)*
Seien $(R,+,\cdot)$ ein Ring, $m,n \in \mathbb{N}$ und $a_1, \ldots, a_m, b_1, \ldots, b_n \in R$. Dann gilt

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

Beweis: Durch Induktion über n bzw. m zeigt man für $a \in R$ bzw. b_j, $1 \leq j \leq n$,

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

Sei $a := \sum_{i=1}^{m} a_i \in R$. Dann ist

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

Seien I und J endliche Mengen, R ein Ring und $(a_i)_{i \in I}$, $(b_j)_{j \in J}$ Familien in R. Mit der Schreibweise von Definition 39 erhält man aus Satz 42:

$$\left(\sum_{i \in I} a_i\right) \cdot \left(\sum_{j \in J} b_j\right) = \sum_{i \in I} \left(\sum_{j \in J} a_i \cdot b_j\right) = \sum_{j \in J} \left(\sum_{i \in I} a_i \cdot b_j\right) =$$

$$= \sum_{(i,j) \in I \times J} a_i \cdot b_j =: \sum_{i \in I,\, j \in J} a_i \cdot b_j\,.$$

Beispiel 43 : Berechne die Summe aller Produkte $i \cdot j$ von Zahlen i in $\{1, 2, \ldots, m\}$ mit Zahlen j in $\{1, 2, \ldots, n\}$! Nach Satz 38 ist

$$\sum_{i=1}^{m} i = 1 + 2 + \ldots + (m-1) + m =$$

$$= (1+m) + (2 + (m-1)) + (3 + (m-2)) + \ldots = \frac{m(m+1)}{2} .$$

Nach Satz 42 ist daher

$$\sum_{i=1}^{m} \sum_{j=1}^{n} i \cdot j = \left(\sum_{i=1}^{m} i \right) \cdot \left(\sum_{j=1}^{n} j \right) = \frac{m \cdot (m+1) \cdot n \cdot (n+1)}{4} .$$

§9. Fragen

1. Es sind vier Funktionen gegeben.
   $f : \mathbb{N} \longrightarrow \mathbb{Z}, z \longmapsto 16 - 4z^2$
   $g : \mathbb{N} \longrightarrow \mathbb{Q}, z \longmapsto -4(z^2 - 4)$
   $h : \mathbb{N} \longrightarrow \mathbb{Z}, z \longmapsto (4 - 2z)(4 + 2z)$
   i sei die Funktion von $\mathbb{N}$ nach $\mathbb{Z}$, deren Graph $\{(z, 16 - 4z^2) \mid z \in \mathbb{N}\}$ ist.
   Welche der folgenden Aussagen sind wahr?
   (a) Die Funktionen f und g sind gleich.
   (b) Die Funktionen f und h sind gleich.
   (c) Die Funktionen f und i sind gleich.

2. Es seien $A := \{1, 2, 3, 5, 6, 7, 9, 10\}$, $B := \{2, 5, 7, 10\}$ und $C := \{1, 3, 6, 9\}$.
   Welche der folgenden Aussagen sind wahr?

   (a) $B \cap C = A$
   (b) $(2 \in A) \wedge (2 \notin B)$
   (c) $(C \subseteq A) \vee (B \not\subseteq A)$
   (d) $B \cup C = A$

3. Induktionsbeweis
   Sei $1 \leq m \leq n$ eine natürliche Zahl und sei $(A_1, A_2, ..., A_n)$ eine Folge von Aussagen.
   Welche der folgenden Aussagen sind wahr?
   (a) Ist die Aussage A_m wahr und gilt für alle $i \in \{1, \ldots, n-1\}$: wenn A_i wahr ist, dann auch A_{i+1}, dann sind alle Aussagen $A_1, \ldots, A_n$ wahr.
   (b) Wenn A_1 falsch ist, aber alle anderen Aussagen wahr sind, dann gilt für alle $i \in \{1, \ldots, n-1\}$: wenn A_i wahr ist, dann auch A_{i+1}.
   (c) Wenn eine der Aussagen $A_1, \ldots, A_n$ falsch ist, dann ist A_1 falsch oder es gibt einen Index i mit $1 \leq i \leq n-1$ so, dass A_i wahr und A_{i+1} falsch ist.

4. Division mit Rest
   Berechne den ganzzahligen Quotient und den Rest von a nach Division duch b!
   (a) $a = 12, b = -5$
   (b) $a = -12, b = 5$
   (c) $a = -12, b = -5$

5. Zifferndarstellung von rationalen Zahlen
   Es seien x bzw. y die ganze Zahl, deren Zifferndarstellung zur Basis zehn 7 bzw. 25 ist.
   (a) Berechne die Zifferndarstellung von x und y zur Basis 2.
   (b) Berechne die Zifferndarstellung der rationalen Zahl $\frac{x}{y}$ zur Basis zwei mit vier Ziffern mit negativem Index.
   (c) Berechne die Zifferndarstellung der rationalen Zahl $\frac{x}{y}$ zur Basis zehn mit vier Ziffern mit negativem Index.

6. Gruppen, Ringe und Körper
   Seien $(R, +, \cdot)$ ein Ring und a, b, c Elemente von R. Welche der folgenden Aussagen sind wahr?
   (a) $(R, +)$ ist eine kommutative Gruppe.
   (b) Wenn $a \cdot c = b \cdot c$ ist, dann ist $a = b$.
   (c) $(R, +, \cdot)$ ist ein Körper, wenn alle Elemente von R invertierbar sind.

7. Summen
   Es seien $a_1, \ldots, a_n, b_1, \ldots, b_n$ Elemente eines Ringes. Welche der folgenden Aussagen sind wahr?
   (a)
   $$\sum_{i=1}^{n} a_i \cdot \sum_{j=1}^{n} b_i = \sum_{k=1}^{n} a_k b_k$$
   (b)
   $$\sum_{i=1}^{n} a_i \cdot \sum_{j=1}^{n} b_j = \sum_{k=1}^{n} a_k \cdot \sum_{\ell=1}^{n} b_\ell$$

KAPITEL 1

Matrizenrechnung

In diesem Kapitel werden mit m,n,p,q immer positive ganze Zahlen und mit K ein Körper (zum Beispiel $\mathbb{Q}$ oder $\mathbb{Z}_2$) bezeichnet.

§1. Matrizen

Definition 44 : Eine *$m \times n$-Matrix mit Koeffizienten in K* (oder eine *$m \times n$-Matrix über K*) ist eine Funktion

$$A : \{1,2,\ldots,m\} \times \{1,2,\ldots,n\} \to K\,,\ (i,j) \mapsto A(i,j)\,.$$

Statt $A(i,j)$ schreiben wir kurz A_{ij}. Eine Matrix wird üblicherweise als Familie

$$A = (A_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1n} \\ A_{21} & A_{22} & \ldots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \ldots & A_{mn} \end{pmatrix}$$

geschrieben. A_{ij} heißt der ij-te *Koeffizient* von A (oder *Koeffizient* in der i-ten *Zeile* und j-ten *Spalte* von A). Als Kurzschreibweise wird

$$A = (A_{ij})_{i,j}$$

verwendet. Anstelle von „Matrix-Koeffizienten“ spricht man auch von „Matrix-Einträgen“. Eine $m \times 1$-Matrix heißt eine *m-Spalte*,
eine $1 \times n$-Matrix eine *n-Zeile*. Die n-Zeile

$$A_{i-} := (A_{i1}, A_{i2}, \ldots, A_{in})$$

heißt i-te *Zeile* von A, die m-Spalte

$$A_{-j} := \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{pmatrix}$$

j-te *Spalte* von A. Die Zahl i bzw. j heißt *Zeilenindex* bzw. *Spaltenindex* des Matrix-Eintrages A_{ij}.
Die Menge aller $m \times n$-Matrizen mit Koeffizienten in K wird mit

$$K^{m \times n}$$

bezeichnet. Matrizen mit Koeffizienten in $\mathbb{Z}$ oder $\mathbb{Q}$ werden kurz *ganzzahlige* oder *rationale* Matrizen genannt.

1×1-Matrizen mit Koeffizienten in K werden üblicherweise mit den entsprechenden Elementen von K identifiziert, d.h.

$$K^{1\times 1} = K \quad , \quad (A_{ij})_{\substack{1\le i\le 1\\ 1\le j\le 1}} = A_{11}\,.$$

$1 \times n$-Matrizen mit Koeffizienten in K werden üblicherweise mit den entsprechenden n-Tupeln in K^n identifiziert, d.h. die Funktionen

$$A : \{1\} \times \{1,\dots,n\} \longrightarrow K,\ (1,i) \longmapsto A(1,i)\,,$$

und

$$\{1,\dots,n\} \longrightarrow K,\ i \longmapsto A(1,i)\,,$$

werden als gleich aufgefasst, obwohl ihre Definitionsbereiche ($\{1\} \times \{1,\dots,n\}$ und $\{1,\dots,n\}$) nicht gleich sind.

Definition 45 : Seien $A,B \in K^{m\times n}$ und $r,s \in K$. Dann heißt

$$A+B := (A_{ij}+B_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} = \begin{pmatrix} A_{11}+B_{11} & \dots & A_{1n}+B_{1n} \\ \vdots & & \vdots \\ A_{m1}+B_{m1} & \dots & A_{mn}+B_{mn} \end{pmatrix} \in K^{m\times n}$$

die *Summe* von A und B, und

$$r\cdot A := (rA_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} = \begin{pmatrix} rA_{11} & \dots & rA_{1n} \\ \vdots & & \vdots \\ rA_{m1} & \dots & rA_{mn} \end{pmatrix} \in K^{m\times n}$$

heißt das *r-fache von A*. Wir schreiben im Folgenden statt „$r\cdot A$“ kurz „rA“. Weiters vereinbaren wir, dass $rA+sB$ immer $(rA)+(sB)$ bedeutet (d.h. rA und sB sind zuerst zu berechnen, dann die Summe von rA und sB).

Satz 46 :

(1) $(K^{m\times n},+)$ *ist eine kommutative Gruppe, wobei das neutrale Element die* $m\times n$-Nullmatrix

$$0 = \begin{pmatrix} 0_K & \dots & 0_K \\ \vdots & & \vdots \\ 0_K & \dots & 0_K \end{pmatrix} \in K^{m\times n}$$

und das zu $A \in K^{m\times n}$ *inverse Element*

$$-A = (-A_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} \in K^{m\times n}$$

ist.

(2) *Für $r,s \in K$ und $A,B \in K^{m\times n}$ ist*
$$(r+s)A = rA + sA$$
und
$$r(A+B) = rA + rB.$$
(3) *Für $r,s \in K$ und $A \in K^{m\times n}$ ist*
$$(rs)A = r(sA) \ \text{ und } \ 1_K A = A\,.$$

Beweis:

(1) Wir zeigen nur die Assoziativität, die anderen Eigenschaften einer Gruppe werden analog bewiesen. Für $A,B,C \in K^{m\times n}$ ist
$(A+B)+C = (A_{ij}+B_{ij})_{i,j} + (C_{i,j})_{i,j} = ((A_{ij}+B_{ij})+C_{ij})_{i,j} =$
$= (A_{ij}+(B_{ij}+C_{ij}))_{i,j} = (A_{ij})_{i,j} + (B_{ij}+C_{ij})_{i,j} = A+(B+C)$
aufgrund des Assoziativgesetzes für die Addition in K.

(2) , (3) Übung.

Definition 47 : Für eine n-Zeile A und eine n-Spalte B heißt
$$A \cdot B := (A_1,\dots,A_n)\cdot \begin{pmatrix} B_1 \\ \vdots \\ B_n \end{pmatrix} := A_1B_1 + \cdots + A_nB_n = \sum_{i=1}^{n} A_iB_i$$
das *Produkt* von A und B (sprich „A mal B“). Oft wird statt $A\cdot B$ nur AB geschrieben.

Beispiel 48 : Ein Korb voller Waren werde durch die Zeile
$$S := (S_1,\dots,S_n) \in \mathbb{Q}^{1\times n}$$
beschrieben, dabei ist S_i die Stückzahl der Ware i im Korb. Sei
$$P := \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix} \in \mathbb{Q}^{n\times 1},$$
dabei ist P_i der Preis der Ware i in Euro. Dann ist
$$SP = \sum_{i=1}^{n} S_iP_i \in \mathbb{Q}$$
der Gesamtpreis (in Euro) für die Waren im Korb.

Definition 49 : Für $A \in K^{m\times n}$ und $B \in K^{n\times p}$ heißt

$$A \cdot B := (A_{i-} \cdot B_{-j})_{\substack{1\le i\le m\\ 1\le j\le p}} := (\sum_{k=1}^{n} A_{ik}B_{kj})_{\substack{1\le i\le m\\ 1\le j\le p}} \in K^{m\times p}$$

das *Produkt* von A und B (sprich „A mal B“). Oft wird statt $A \cdot B$ nur AB geschrieben. Zur Berechnung des Koeffizienten

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj}$$

werden die Koeffizienten in der i-ten Zeile von A der Reihe nach mit den entsprechenden Koeffizienten in der j-ten Spalte von B multipliziert und anschließend alle diese Produkte addiert.

Im Spezialfall $m = 1$ und $p = 1$, d.h. A ist eine n-Zeile und B eine n-Spalte, stimmen die Definitionen 47 und 49 überein.

Beispiel 50 : Die Waren $1,\dots,m$ werden aus Rohstoffen $1,\dots,n$ hergestellt, die von Lieferanten $1,\dots,p$ bezogen werden. Für die Erzeugung der Ware i werden Q_{ij} Einheiten des Rohstoffes j benötigt. Der Preis einer Einheit des Rohstoffes j beim Lieferanten k beträgt P_{jk}. Setzt man $Q := (Q_{ij})_{i,j} \in \mathbb{Q}^{m\times n}$ und $P := (P_{jk})_{j,k} \in \mathbb{Q}^{n\times p}$, dann ist

$$(QP)_{ik} = \sum_{j=1}^{n} Q_{ij}P_{jk}$$

der Gesamtpreis der Rohstoffe für Ware i beim Lieferanten k. Sollen jeweils S_i Stück der Ware i produziert werden und setzt man $S = (S_1,\dots,S_m) \in \mathbb{Q}^{1\times n}$, dann ist $(SQ)_{1j}$ die Anzahl der insgesamt benötigten Einheiten von Rohstoff j und

$$((SQ)P)_{1k}$$

ist der Preis bei Lieferant k für alle benötigten Rohstoffe.

Satz 51 : *Die Matrizenmultiplikation ist assoziativ, d.h. für Matrizen $A \in K^{m\times n}$, $B \in K^{n\times p}$ und $C \in K^{p\times q}$ gilt*

$$(AB)C = A(BC).$$

Beweis: Da AB eine $m \times p$-Matrix und BC eine $n \times q$-Matrix ist, sind sowohl $(AB)C$ als auch $A(BC)$ $m \times q$-Matrizen. Für $1 \le i \le m$ und $1 \le j \le q$

ist

$$\begin{aligned}((AB)C)_{ij} &= \sum_{k=1}^{p}(AB)_{ik}C_{kj} = \sum_{k=1}^{p}\left(\sum_{\ell=1}^{n}A_{i\ell}B_{\ell k}\right)C_{kj} \\ &= \sum_{k=1}^{p}\sum_{\ell=1}^{n}A_{i\ell}B_{\ell k}C_{kj} = \sum_{\ell=1}^{n}\sum_{k=1}^{p}A_{i\ell}B_{\ell k}C_{kj} \\ &= \sum_{\ell=1}^{n}A_{i\ell}\left(\sum_{k=1}^{p}B_{\ell k}C_{kj}\right) = \sum_{\ell=1}^{n}A_{i\ell}(BC)_{\ell j} \\ &= (A(BC))_{ij}\,.\end{aligned}$$

Definition 52 : Für Elemente i, j einer beliebigen Indexmenge ist das *Kronecker-Delta* in K

$$\delta_{ij} := \begin{cases} 1_K & \text{falls } i = j \text{ ist,} \\ 0_K & \text{falls } i \neq j \text{ ist.} \end{cases}$$

Die Matrix

$$I_n := (\delta_{ij})_{\substack{1\le i\le n\\1\le j\le n}} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \in K^{n\times n}$$

heißt $n \times n$*-Einheitsmatrix*.

Satz 53 : *Für eine beliebige Matrix $A \in K^{m\times n}$ ist*

$$I_m A = A \quad \textit{und} \quad AI_n = A\,.$$

Beweis: Für $1 \le i \le m$ und $1 \le j \le n$ ist $(I_mA)_{ij} = \sum_{k=1}^{m}\delta_{ik}A_{kj} = A_{ij}$, also $I_mA = A$. Analog beweist man $AI_n = A$.

Satz 54 : *Für $A, B \in K^{m\times n}$ und $C \in K^{n\times p}$ gilt*

$$(A+B)C = AC + BC.$$

Für $A \in K^{m\times n}$ und $B, C \in K^{n\times p}$ gilt

$$A(B+C) = AB + AC.$$

Für $A \in K^{m\times n}$, $B \in K^{n\times p}$ und $r \in K$ gilt

$$r(AB) = (rA)B = A(rB).$$

Beweis: Übung.

Satz 55: $(K^{n\times n},+,\cdot)$ *ist ein Ring mit Einselement I_n. Wegen*

$$\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},$$

aber

$$\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},$$

ist $K^{n\times n}$ im Allgemeinen nicht kommutativ.

Beweis: Die Behauptung folgt aus Satz 46, Satz 51, Satz 53 und Satz 54.

Definition 56: Eine Matrix $A \in K^{n\times n}$ heißt *invertierbar*, wenn es eine Matrix $B \in K^{n\times n}$ gibt mit

$$AB = I_n \quad \text{und} \quad BA = I_n\,.$$

In diesem Fall nennt man B die zu A *inverse* Matrix und schreibt

$$B = A^{-1}.$$

Sei

$$\mathrm{GL}_n(K) := \{A \in K^{n\times n} \mid A \text{ invertierbar}\}.$$

Nach Satz 31 ist $(\mathrm{GL}_n(K),\cdot)$ eine Gruppe, sie heißt *allgemeine lineare Gruppe* (auf Englisch „general linear group“).

§2. Elementare Umformungen

Definition 57: Seien $1 \le k \le m$ und $1 \le \ell \le n$. Dann heißt die Matrix $E_{k\ell} \in K^{m\times n}$ mit Koeffizienten

$$(E_{k\ell})_{ij} := \delta_{ik}\delta_{j\ell} = \begin{cases} 1_K & \text{falls } i = k \text{ und } j = \ell \text{ ist,}\\ 0_K & \text{falls } i \neq k \text{ oder } j \neq \ell \text{ ist,}\end{cases}$$

eine *Standard-Matrix* von $K^{m\times n}$. Zum Beispiel sind die Standard-Matrizen von $\mathbb{Q}^{2\times 2}$

$$E_{11} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\ E_{12} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\ E_{21} = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix} \text{ und } E_{22} = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}.$$

Im Spezialfall $m = 1$ (oder $n = 1$) schreibt man statt $E_{1\ell}$ (bzw. E_{k1}) kurz e_ℓ (bzw. e_k). Zum Beispiel sind die Standard-Zeilen von $\mathbb{Q}^{1\times 3}$

$$e_1 = (1,0,0)\,,\ e_2 = (0,1,0) \text{ und } e_3 = (0,0,1)$$

und die Standard-Spalten von $\mathbb{Q}^{2\times 1}$ sind

$$e_1 = \begin{pmatrix}1\\ 0\end{pmatrix} \text{ und } e_2 = \begin{pmatrix}0\\ 1\end{pmatrix}.$$

Satz 58: *Für $E_{k\ell} \in K^{m\times m}$, $A \in K^{m\times n}$ und $1 \le i \le m$ ist*

$$(E_{k\ell}A)_{i-} = \begin{cases} A_{\ell-} & \text{falls } i = k \text{ ist}, \\ 0 & \text{falls } i \neq k \text{ ist} \end{cases} .$$

Beweis: Sei $1 \le j \le n$. Dann ist

$$(E_{k\ell}A)_{ij} = \sum_{p=1}^{m} (E_{k\ell})_{ip}A_{pj} = \sum_{p=1}^{m} \delta_{ki}\delta_{\ell p}A_{pj} =$$

$$= \delta_{ki}A_{\ell j} = \begin{cases} A_{\ell j} & \text{falls } i = k \text{ ist}, \\ 0 & \text{falls } i \neq k \text{ ist} \end{cases} .$$

Definition 59: Die folgenden Matrizen heißen *Elementarmatrizen* in $K^{n\times n}$:

Typ 1: $I_n + rE_{k\ell}$, wobei $r \in K$, $1 \le k,\ell \le n$ und $k \neq \ell$ ist,
Typ 2: $I_n - E_{kk} - E_{\ell\ell} + E_{k\ell} + E_{\ell k}$, wobei $1 \le k,\ell \le n$ und $k \neq \ell$ ist,
Typ 3: $I_n + (t-1)E_{kk}$, wobei $1 \le k \le n$ und $t \in K$ invertierbar ist.

Zum Beispiel sind

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} , \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$$

Elementarmatrizen vom Typ 1, Typ 2 bzw. Typ 3.

Satz 60: *Sei $A \in K^{m\times n}$ und seien $P \in K^{m\times m}$ sowie $Q \in K^{n\times n}$ Elementarmatrizen. Dann erhält man PA aus A, indem man*

Typ 1: *zur k-ten Zeile von A das r-fache der ℓ-ten Zeile addiert,*
Typ 2: *die k-te und ℓ-te Zeile von A vertauscht,*
Typ 3: *die k-te Zeile von A mit t multipliziert.*

Diese Umformungen der Matrix A heißen elementare Zeilenumformungen.
Analog erhält man AQ aus A, indem man

Typ 1: *zur ℓ-ten Spalte von A das r-fache der k-ten Spalte addiert,*
Typ 2: *die k-te und ℓ-te Spalte von A vertauscht,*
Typ 3: *die k-te Spalte von A mit t multipliziert.*

Diese Umformungen der Matrix A heißen elementare Spaltenumformungen.

Beweis: Für $P = I_n + rE_{k\ell}$, wobei $r \in K$ und $k \neq \ell$, ist
$PA = (I_n + rE_{k\ell})A = A + rE_{k\ell}A$. Nach Satz 58 ist

$$(rE_{k\ell}A)_{i-} = \begin{cases} rA_{\ell-} & \text{falls } i = k \text{ ist}, \\ 0 & \text{falls } i \neq k \text{ ist}. \end{cases}$$

Die anderen Fälle beweist man analog.

Wenn eine elementare Zeilenumformung von $A \in K^{m\times n}$ durchgeführt wird, entspricht das der Multiplikation (von links) einer gewissen Elementarmatrix $P \in K^{m\times m}$ mit A. Wegen $P \cdot I_m = P$ erhält man diese Elementarmatrix, indem man diese elementare Umformung an der Einheitsmatrix I_m durchführt. Zum Beispiel ist die Elementarmatrix in $\mathbb{Q}^{2\times 2}$, die der elementaren Umformung „addiere zur ersten Zeile einer Matrix deren 4-fache zweite Zeile“ entspricht, jene Matrix, die man erhält, indem man zur ersten Zeile der Einheitsmatrix in $\mathbb{Q}^{2\times 2}$ ihre 4-fache zweite Zeile addiert:

$$\begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}.$$

Satz 61 : *Elementarmatrizen sind invertierbar, genauer gilt:*

Typ 1: $(I_n + rE_{k\ell})^{-1} = I_n - rE_{k\ell}$

Typ 2: $(I_n - E_{kk} - E_{\ell\ell} + E_{k\ell} + E_{\ell k})^{-1} = I_n - E_{kk} - E_{\ell\ell} + E_{k\ell} + E_{\ell k}$

Typ 3: $(I_n + (t-1)E_{kk})^{-1} = I_n + (t^{-1} - 1)E_{kk}\,.$

Somit können alle elementaren Zeilen- oder Spaltenumformungen einer beliebigen Matrix durch elementare Zeilen- oder Spaltenumformungen wieder rückgängig gemacht werden.

Beweis: Die Matrix

$$(I_n - rE_{k\ell})(I_n + rE_{k\ell}) = (I_n - rE_{k\ell})[(I_n + rE_{k\ell})I_n]$$

erhält man aus I_n, indem man zuerst zur k-ten Zeile das r-fache der ℓ-ten Zeile addiert und anschließend das r-fache der ℓ-ten Zeile subtrahiert. Daher ist

$$(I_n - rE_{k\ell})(I_n + rE_{k\ell}) = I_n.$$

Die anderen Fälle beweist man analog. Ist P eine Elementarmatrix, so bekommt man A aus $B := PA$ wieder zurück, indem man B von links mit P^{-1} multipliziert.

§3. Fragen

Im Folgenden seien K ein Körper und $K^{m\times n}$ die Menge aller $m\times n$-Matrizen mit Koeffizienten in K.

1. Multiplikation von Matrizen
   Sei $A\in K^{m\times n}$, $B\in K^{n\times p}$ und $A\cdot B$ das Produkt der Matrizen A und B. Der Koeffizient $(A\cdot B)_{ij}$ der Matrix $A\cdot B$ in der i-ten Zeile und j-ten Spalte ist
   (a)
   $$A_{ij}\cdot B_{ij}$$
   (b)
   $$\sum_{k=1}^{n} A_{ik}\cdot B_{kj}$$
   (c)
   $$\sum_{k=1}^{n} A_{ki}\cdot B_{jk}$$

2. Rechenregeln für Matrizen
   Welche der folgenden Aussagen sind falsch?
   (a) Die Matrizenmultiplikation ist kommutativ,
   d.h. für alle Matrizen $A,B\in K^{n\times n}$ ist $A\cdot B=B\cdot A$.
   (b) Die Matrizenmultiplikation ist assoziativ,
   d.h. für alle Matrizen $A,B,C\in K^{n\times n}$ ist $(A\cdot B)\cdot C=A\cdot(B\cdot C)$.
   (c) Die Matrizenmultiplikation ist bezüglich der Addition distributiv, d.h. für alle Matrizen $A,B,C\in K^{n\times n}$ ist
   $A\cdot(B+C)=A\cdot B+A\cdot C$ und $(A+B)\cdot C=A\cdot C+B\cdot C$.

3. Elementare Zeilenumformungen
   Es sei $P\in\mathbb{Q}^{m\times m}$ eine Elementarmatrix, $A\in\mathbb{Q}^{m\times n}$ und $P\cdot A$ die Matrix, die man aus A erhält, indem man die 3-fache zweite Zeile von A zur ersten Zeile addiert.
   Dann ist P die Matrix, die man aus der Einheitsmatrix I_m erhält, indem man
   (a) die 3-fache zweite Spalte von I_m zur ersten addiert.
   (b) die 3-fache zweite Zeile von I_m von der ersten subtrahiert.
   (c) die 3-fache zweite Zeile von I_m zur ersten addiert.
   (d) die 3-fache zweite Spalte von I_m von der ersten subtrahiert.

4. Invertierbare Matrizen
   Welche der folgenden Aussagen sind wahr?
   (a) Elementarmatrizen sind invertierbar.
   (b) Für alle invertierbaren Matrizen A und B ist ihr Produkt $A\cdot B$ invertierbar und es ist $(A\cdot B)^{-1}=B^{-1}\cdot A^{-1}$.
   (c) Für alle invertierbaren Matrizen A und B ist ihr Produkt $A\cdot B$ invertierbar und es ist $(A\cdot B)^{-1}=A^{-1}\cdot B^{-1}$.

KAPITEL 2

Systeme linearer Gleichungen

In diesem Kapitel werden mit m,n immer positive ganze Zahlen und mit K ein Körper (zum Beispiel $\mathbb{Q}$ oder $\mathbb{Z}_2$) bezeichnet.

§1. Systeme linearer Gleichungen

Definition 62 : Ein *System linearer Gleichungen mit Koeffizienten in K* ist eine Aufgabe:

- Gegeben sind eine Matrix $A \in K^{m\times n}$ und eine Spalte $b \in K^{m\times 1}$.
- Gesucht ist eine „gute Beschreibung“ der Menge

  $$L(A,b) := \{x \mid x \in K^{n\times 1} \text{ mit } Ax = b\}$$

  aller n-Spalten x (mit Koeffizienten in K), für die $Ax = b$ ist.

Die Menge $L(A,b)$ heißt *Lösungsmenge* des durch A und b gegebenen Systems linearer Gleichungen. Ihre Elemente heißen *Lösungen* dieses Systems.

Das durch A und b gegebene System linearer Gleichungen heißt *homogen*, wenn b die Nullspalte ist, ansonsten *inhomogen*.

Ohne Matrizen kann man das so formulieren: Gegeben sind Elemente $A_{ij} \in K$ und $b_i \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$. Gesucht ist eine „gute Beschreibung“ der Menge aller n-Tupel $(x_1,\ldots,x_n)$ mit Komponenten in K, für die

$$\begin{array}{rcl} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n & = & b_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n & = & b_m \end{array}$$

ist.

Das durch $A \in K^{m\times n}$ und $b \in K^{m\times 1}$ gegebene System linearer Gleichungen wird kurz mit „(A,b)“ oder „$Ax = b$“ bezeichnet. Die Zahl m heißt die *Anzahl der Gleichungen*, die Zahl n die *Anzahl der Unbekannten*.

Wenn 0_n bzw. 0_m die Nullspalte in $K^{n\times 1}$ bzw. $K^{m\times 1}$ ist, dann ist $A0_n = 0_m$, also hat ein homogenes System linearer Gleichungen immer mindestens eine Lösung, nämlich die Nullspalte. Hingegen gibt es inhomogene Systeme linearer Gleichungen ohne Lösung, zum Beispiel

$$\begin{array}{rcl} x_1 + x_2 & = & 0 \\ 2x_1 + 2x_2 & = & 1 \end{array} \quad .$$

Beispiel 63: Es soll eine Legierung aus b_i Gramm der Metalle M_i, $1 \leq i \leq m$, hergestellt werden. Zur Verfügung stehen Legierungen $L_1, \ldots, L_n$ der Metalle $M_1, \ldots, M_m$, wobei 1 Gramm der Legierung L_j jeweils A_{ij} Gramm des Metalls M_i enthält. Wieviele Gramm von $L_1, \ldots, L_n$ müssen für die gewünschte Legierung verschmolzen werden? Wir nehmen dabei an, dass beim Verschmelzen der Legierungen die Masse erhalten bleibt.

Das Verschmelzen von $x_1, \ldots, x_n$ Gramm der Legierungen $L_1, \ldots, L_n$ ergibt dann eine Legierung mit jeweils

$$\sum_{j=1}^{n} A_{ij} x_j$$

Gramm des Metalls M_i. Gesucht ist somit $L(A,b)$.

Definition 62 ist noch unvollständig, weil wir noch nicht präzise vereinbart haben, was eine „gute Beschreibung“ von $L(A,b)$ ist. Um diese Definition zu vervollständigen, brauchen wir ein paar Vorüberlegungen. Grundlegend dafür sind zwei einfache Beobachtungen, die wir in den Sätzen 64 und 65 formulieren:

Satz 64: *Sei (A,b) ein System linearer Gleichungen mit Koeffizienten in K und z irgendeine Lösung davon. (Wir nehmen also insbesondere an, dass $L(A,b)$ nicht leer ist). Dann ist*

$$L(A,b) = \{z + v \mid v \in L(A,0)\}.$$

Beweis: Sei $v \in L(A,0)$. Dann ist $A(z+v) = Az + Av = b + 0 = b$, also $z + v \in L(A,b)$.

Sei $y \in L(A,b)$. Dann ist $A(y-z) = Ay - Az = b - b = 0$, also $y - z \in L(A,0)$ und $y = z + (y-z) \in \{z + v \mid v \in L(A,0)\}$.

Um die Lösungsmenge eines inhomogenen Systems linearer Gleichungen (A,b) zu beschreiben, genügt es somit, nur *eine* Lösung zu finden und die Lösungsmenge des homogenen linearen Gleichungssystems $(A,0)$ zu beschreiben.

Satz 65: *Seien $A \in K^{m \times n}$, $r,s \in K$ und $v,w \in L(A,0)$.*
Dann ist auch $rv + sw \in L(A,0)$.

Beweis: $A(rv + sw) = r(Av) + s(Aw) = 0 + 0 = 0.$

Für die Lösungsmenge L eines Systems linearer Gleichungen gilt genau eine der folgenden drei Aussagen:

(1) L ist leer („es gibt keine Lösung").

(2) L enthält genau ein Element („eine Lösung existiert und ist eindeutig bestimmt").

(3) L enthält mindestens zwei Elemente.

Im Fall (2) wird L durch die Angabe der eindeutig bestimmten Spalte, die Lösung des gegebenen Systems linearer Gleichungen ist, gut beschrieben.

Im Fall (3) ist das aber schwieriger. Wenn K unendlich ist (z.B. $K = \mathbb{Q}$), dann enthält L im Fall (3) nach Satz 64 und Satz 65 unendlich viele Elemente. Wie können wir L dann durch *endlich viele Daten* beschreiben? Um diese Frage zu beantworten, führen wir im nächsten Abschnitt die Begriffe „Vektorraum" und „Basis" ein.

§2. Vektorräume

Beim Rechnen mit $m \times n$-Matrizen mit Koeffizienten in einem Körper K haben wir zwei Rechenoperationen kennengelernt:

(1) Die Addition von zwei $m \times n$-Matrizen,
(2) Die „Multiplikation" eines Elementes von K (eines „Skalars") mit einer $m \times n$-Matrix.

Im Satz 46 wurden die dafür geltenden Rechenregeln angegeben.

Definition 66 : Sei V eine Menge und seien

$$+ : V \times V \to V \quad \text{sowie} \quad \cdot : K \times V \to V$$

Funktionen. Wir schreiben statt „$+(v,w)$" kurz „$v+w$" und statt „$\cdot(r,v)$" kurz „$r \cdot v$ oder nur „rv". Das Tripel $(V,+,\cdot)$ ist ein *Vektorraum über K*, wenn die folgenden drei Bedingungen erfüllt sind:

(1) $(V,+)$ ist eine abelsche Gruppe.
(2) Für alle $r,s \in K$ und für alle $v,w \in V$ ist
$r(v+w) = (rv)+(rw)$ und $(r+s)v = (rv)+(sv)$.
(3) Für alle $r,s \in K$ und für alle $v \in V$ ist $(rs)v = r(sv)$ und $1_K v = v$.

Ist $(V,+,\cdot)$ ein Vektorraum über K, dann heißen die Elemente von V „*Vektoren*", die Elemente von K „*Skalare*", $+$ „*Addition*" und $\cdot$ „*Skalarmultiplikation*". Statt $(V,+,\cdot)$ wird oft nur V geschrieben. Das neutrale Element von $(V,+)$ wird mit 0_V bezeichnet und heißt der *Nullvektor*. Ein Vektor v heißt *ein skalares Vielfaches* eines Vektors w, wenn es ein Element $r \in K$ gibt mit $v = rw$. In diesem Fall sagt man auch, daß v das *r-fache von w* ist.

Man beachte, dass der Begriff „Vektor" erst nach dem Begriff „Vektorraum" eingeführt werden kann, so wie der Begriff „Tiroler" erst nach dem Begriff „Tirol" eingeführt werden kann.

Die Eigenschaften von Vektoren können kurz so beschrieben werden: Vektoren können miteinander addiert und mit Skalaren multipliziert werden. Dabei gelten gewisse Rechenregeln.

Ein Beispiel aus der Physik: alle Kräfte, die in einem vorgegebenen Punkt angreifen, bilden einen Vektorraum, weil sie addiert und mit Zahlen multipiziert werden können und dabei obige Rechenregeln gelten. Daher sind solche Kräfte Vektoren.

Die Addition und die Skalarmultiplikation von $m \times n$-Matrizen mit Koeffizienten in einem Körper erfüllen die Rechenregeln eines Vektorraums. Daher: „Matrizen sind Vektoren“.

Die Addition und die (Skalar-)Multiplikation von rationalen Zahlen erfüllen die Rechenregeln eines Vektorraums. Daher: „Rationale Zahlen sind Vektoren“.

Satz 67 : *Die Menge*

$$K^n = \{(a_1, \dots, a_n) \mid a_1, \dots, a_n \in K\}$$

aller n-Tupel in K mit der komponentenweisen Addition

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) := (a_1 + b_1, \dots, a_n + b_n)$$

und der komponentenweisen Skalarmultiplikation

$$r(a_1, \dots, a_n) := (ra_1, \dots, ra_n)$$

ist ein Vektorraum über K und heißt Standard-Vektorraum über K der Dimension n. *In diesem Vektorraum ist* $0_V = (0, \dots, 0)$ *und* $-(a_1, \dots, a_n) = (-a_1, \dots, -a_n)$.

Beweis: Übung.

Beispiel 68 : Ein Kaufhaus bietet n Waren an. Sei U_{ij} die Anzahl der am i-ten Tag verkauften Einheiten der j-ten Ware. Dann gibt

$$U_{i-} = (U_{i1}, \dots, U_{in}) \in K^n$$

den Umsatz am i-ten Tag an,

$$U_{1-} + \cdots + U_{k-} = \left(\sum_{i=1}^{k} U_{i1}, \dots, \sum_{i=1}^{k} U_{in} \right)$$

ist der Umsatz vom ersten bis zum k-ten Tag, und

$$\frac{1}{k} (U_{1-} + \cdots + U_{k-}) \in K^n$$

ist der durchschnittliche Tagesumsatz während der ersten k Tage.

Satz 69: *Seien V ein Vektorraum über K, $r \in K$ und $v \in V$. Dann gilt:*

(1) *Es ist $rv = 0$ genau dann, wenn $r = 0$ oder $v = 0$ ist.*
(2) $(-r)v = r(-v) = -(rv)$.

Beweis:

(1) Aus
$$0_V + 0_K v = 0_K v = (0_K + 0_K)v = 0_K v + 0_K v$$
folgt
$$0_V + 0_K v - 0_K v = 0_K v + 0_K v - 0_K v$$
und daher $0_K v = 0_V$. Ebenso folgt aus
$$0_V + r0_V = r0_V = r(0_V + 0_V) = r0_V + r0_V,$$
dass $r0_V = 0_V$ ist. Wenn umgekehrt $rv = 0$ aber $r \neq 0$ ist, dann ist r invertierbar, weil K ein Körper ist, und
$$v = 1_K v = (r^{-1}r)v = r^{-1}(rv) = r^{-1}0_V = 0_V .$$
(2) Wegen $(rv) + (-r)v = [r + (-r)]v = 0_K v = 0_V$ ist $-(rv) = (-r)v$. Wegen $(rv) + (r(-v)) = r[v + (-v)] = r0_V = 0_V$ ist $-(rv) = r(-v)$.

Definition 70: Sei V ein Vektorraum über K. Eine Teilmenge U von V heißt *Untervektorraum* von V, wenn die folgenden drei Bedingungen erfüllt sind:

(1) $0_V \in U$
(2) Sind zwei Vektoren u, v Elemente von U, dann auch ihre Summe $u + v$.
(3) Ist ein Vektor v Element von U, dann auch alle seine skalaren Vielfachen rv, $r \in K$.

Man schreibt dann
$$U \leq_K V \quad \text{oder kurz} \quad U \leq V.$$

Beispiel 71: Sei V ein Vektorraum über einem Körper K. Dann sind $\{0\}$, V und $Kv := \{cv \mid c \in K\}$ (für jeden Vektor $v \in V$) Untervektorräume von V.

Satz 72: *Seien $(V, +, \cdot)$ ein Vektorraum über K und W ein Untervektorraum von V. Dann ist W mit $W \times W \longrightarrow W,\ (u, w) \longmapsto u + w$, und $K \times W \longrightarrow W,\ (c, w) \longmapsto c \cdot w$, selbst ein Vektorraum über K.*

Beweis: Übung.

Satz 73 : *Die Lösungsmenge eines homogenen Systems linearer Gleichungen mit der komponentenweisen Addition und Skalarmultiplikation ist ein Vektorraum. (Anders formuliert: Die Nullspalte ist eine Lösung, die Summe zweier Lösungen ist wieder eine Lösung und alle skalaren Vielfachen von Lösungen sind wieder Lösungen).*

Beweis: Folgt aus Satz 65.

§3. Erzeugendensysteme, lineare Unabhängigkeit und Basen

Definition 74 : Seien V ein Vektorraum über K und $v_1,\ldots,v_n$ Vektoren in V. Ein Vektor $w \in V$ heißt eine *Linearkombination* von $v_1,\ldots,v_n$, wenn es Elemente $c_1,\ldots,c_n$ in K gibt, sodass

$$w = c_1v_1 + c_2v_2 + \ldots + c_nv_n \ (= \sum_{i=1}^{n} c_iv_i)$$

ist.
Die Elemente $c_1,\ldots,c_n$ heißen *Koeffizienten* von w bezüglich $v_1,\ldots,v_n$.
Die Menge

$$\{\sum_{i=1}^{n} c_iv_i \,|\, c_1,\ldots,c_n \in K\}$$

aller Linearkombinationen von $v_1,\ldots,v_n$ ist ein Untervektorraum von V (nachprüfen), enthält $v_1,\ldots,v_n$ und heißt der *von* $v_1,\ldots,v_n$ *erzeugte Untervektorraum von* V. Er wird mit

$${}_K\langle v_1,\ldots,v_n\rangle \qquad \text{oder} \qquad \sum_{i=1}^{n} Kv_i$$

bezeichnet.

Definition 75 : Sei $V \neq \{0\}$ ein Vektorraum über K. Ein n-Tupel $(v_1,\ldots,v_n)$ von Vektoren in V heißt genau dann ein *Erzeugendensystem* von V bzw. *linear unabhängig* in V bzw. eine *Basis* von V, wenn jeder Vektor in V auf mindestens eine bzw. höchstens eine bzw. genau eine Weise als Linearkombination von $(v_1,\ldots,v_n)$ geschrieben werden kann.
Wir schreiben *linear abhängig* anstatt *nicht linear unabhängig*.

Satz 76 : *Sei* $A \in K^{m\times n}$ *und* $y \in K^{n\times 1}$.

(1) *Die Spalte* Ay *ist eine Linearkombination der Spalten* $A_{-1},\ldots,A_{-n}$ *mit Koeffizienten* $y_1,\ldots,y_n$*, das heißt*

$$Ay = y_1A_{-1} + \ldots y_nA_{-n} \ .$$

(2) $L(A,0)$ *enthält genau dann nur ein Element (und zwar* $0 \in K^{n\times 1}$*), wenn das n-Tupel der Spalten* $(A_{-1},\ldots,A_{-n})$ *linear unabhängig ist.*

(3) *Für $b \in K^{m\times 1}$ hat das lineare Gleichungssystem (A,b) genau dann eine Lösung, wenn b ein Element des von den Spalten von A erzeugten Untervektorraums von $K^{m\times 1}$ ist.*
(4) *Das n-Tupel $(A_{-1},\dots,A_{-n})$ der Spalten von A ist genau dann eine Basis von $K^{m\times 1}$, wenn für jede Spalte $b \in K^{m\times 1}$ das System linearer Gleichungen (A,b) genau eine Lösung hat.*

Beweis: (1) ist leicht nachzurechnen, (2), (3) und (4) folgen aus (1).

Satz 77 : *Sei $V \neq \{0\}$ ein Vektorraum über K und $(v_1,\dots,v_n)$ ein n-Tupel von Vektoren in V.*

(1) *Eine Basis von V ist ein linear unabhängiges Erzeugendensystem.*
(2) *Das n-Tupel $(v_1,\dots,v_n)$ von Vektoren ist genau dann ein Erzeugendensystem von V, wenn*

$$_K\langle v_1,\dots,v_n\rangle = V$$

ist.
(3) *Das n-Tupel $(v_1,\dots,v_n)$ von Vektoren ist genau dann linear unabhängig, wenn für jedes n-Tupel $(c_1,\dots,c_n) \in K^n$ aus*

$$\sum_{i=1}^{n} c_i v_i = 0$$

folgt, dass

$$c_1 = c_2 = \ldots = c_n = 0$$

ist.

Beweis:

(1) und (2) folgen aus der Definition der Begriffe Erzeugendensystem, linear unabhängig und Basis.
(3) Wenn sich jeder Vektor aus V auf höchstens eine Weise als Linearkombination von $(v_1,\dots,v_n)$ schreiben lässt, dann folgt aus

$$\sum_{i=1}^{n} c_i v_i = 0_V = \sum_{i=1}^{n} 0_K v_i$$

auf Grund der Eindeutigkeit $c_i = 0_K$ für $1 \leq i \leq n$. Sei umgekehrt $(v_1,\dots,v_n)$ linear unabhängig in V und $w \in V$ so, dass es ein n-Tupel $(c_1,\dots,c_n) \in K^n$ mit

$$w = \sum_{i=1}^{n} c_i v_i$$

gibt. Falls $(d_1,\dots,d_n) \in K^n$ eine weiteres n-Tupel mit

$$w = \sum_{i=1}^{n} d_i v_i$$

ist, erhält man

$$0_V = w - w = \sum_{i=1}^{n} c_i v_i - \sum_{i=1}^{n} d_i v_i = \sum_{i=1}^{n} (c_i - d_i) v_i \,.$$

Nach Annahme folgt $c_i - d_i = 0_K$ für $1 \leq i \leq n$, also $c_i = d_i$ für $1 \leq i \leq n$.

Beispiel 78 : Sei V ein Vektorraum über K. Ein einzelner Vektor $v \in V$ ist linear unabhängig genau dann, wenn $v \neq 0$ ist, weil aus $rv = 0$ nach Satz 69 $r = 0$ oder $v = 0$ folgt.

Satz 79 : *Das $m \cdot n$-Tupel*

$$(E_{11},\dots,E_{1n},E_{21},\dots,E_{2n},\dots,E_{m1},\dots,E_{mn})$$

der Standard-Matrizen (siehe Definition 57) ist eine Basis von $K^{m\times n}$ und heißt die Standardbasis *von $K^{m\times n}$.*

Insbesondere ist das n-Tupel $(e_1,\dots,e_n)$ der Standard-Zeilen (siehe Definition 57) eine Basis von K^n und heißt die Standardbasis *von K^n.*

Beweis: Da für eine beliebige Matrix $A \in K^{m\times n}$

$$A = \sum_{\substack{1\leq k\leq m\\ 1\leq \ell\leq n}} A_{k\ell} E_{k\ell}$$

ist, ist das $m \cdot n$-Tupel $(E_{11},\dots,E_{1n},E_{21},\dots,E_{2n},\dots,E_{m1},\dots,E_{mn})$ ein Erzeugendensystem von $K^{m\times n}$. Um die lineare Unabhängigkeit zu zeigen, nehmen wir an, dass für gewisse $c_{k\ell} \in K$

$$\sum_{\substack{1\leq k\leq m\\ 1\leq \ell\leq n}} c_{k\ell} E_{k\ell} = 0$$

ist. Dann folgt für alle $1 \leq i \leq m$ und $1 \leq j \leq n$

$$0 = \sum_{\substack{1\leq k\leq m\\ 1\leq \ell\leq n}} c_{k\ell} (E_{k\ell})_{ij} = c_{ij} \,,$$

was zu zeigen war.

Nach Einführung des Begriffes „Basis“ kann die Definition eines Systems linearer Gleichungen (A,b) präzisiert werden. Eine „gute Beschreibung von $L(A,b)$“ finden heißt:

„Berechne irgendeine Lösung von (A,b) und eine Basis des Lösungsraums von $(A,0)$“.

Man beachte, dass damit nur *endlich viele Daten* zu berechnen sind, diese bestimmen aber vollständig die (möglicherweise unendliche) Lösungsmenge $L(A,b)$.

§4. Der Gauss-Algorithmus

Wir werden nun eine Methode zum Lösen eines Systems linearer Gleichungen (A,b) kennenlernen. Zunächst betrachten wir einen Spezialfall, in dem die Matrix A eine besonders einfache Gestalt hat. In diesem Fall können wir die Lösungsmenge direkt anschreiben (Satz 83). Danach werden wir den allgemeinen Fall auf den einfachen Fall zurückführen, indem wir die Daten (A,b) schrittweise so verändern, dass die geänderten Daten (A',b') schließlich die einfache Gestalt haben und

$$L(A,b) = L(A',b')$$

ist (Satz 86 und Satz 87).
Den Übergang von den Daten (A,b) zu (A',b') nennt man „die Gleichungen umformen“. Gilt dabei $L(A,b) = L(A',b')$, dann sind die Umformungen „erlaubt“.

Definition 80 : Eine Matrix $A \in K^{m\times n}$ hat *Stufenform*, wenn die folgenden vier Bedingungen erfüllt sind:

(1) Ist $A_{i-} = 0$, dann sind auch $A_{(i+1)-} = \cdots = A_{m-} = 0$.
(2) Der (von links gelesen) erste Koeffizient $\neq 0$ in jeder Zeile heißt *Pivot* und ist 1.
(3) Der Pivot in der Zeile $i+1$ steht rechts vom Pivot in der Zeile i.
(4) Der Pivot jeder Zeile ist der einzige Koeffizient $\neq 0$ in seiner Spalte.

Eine Matrix in Stufenform hat die Gestalt

$$\begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 & * & \dots \\ \vdots \end{pmatrix},$$

wobei die Sterne für beliebige Elemente von K stehen.

Definition 81 : Der *Spaltenraum* bzw. *Zeilenraum* einer Matrix mit m Zeilen und n Spalten ist der Untervektorraum von $K^{m\times 1}$ bzw. $K^{1\times n}$, der von den Spalten bzw. Zeilen dieser Matrix erzeugt wird.

Beispiel 82: $A \in K^{m\times n}$ sei eine Matrix in Stufenform mit r Pivots. Dann ist das r-Tupel $(e_1,\dots,e_r)$ der ersten r Standard-Spalten in $K^{m\times 1}$ eine Basis des Spaltenraums von A.
Wenn $p_1,\dots,p_r$ die Spaltenindizes der Pivots sind, dann ist

$$A_{-p_1} = e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \end{pmatrix}, \dots, A_{-p_r} = e_r = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix}$$

und

$$A_{-\ell} = \sum_{i=1}^{r} A_{i\ell} e_i = \sum_{i=1}^{r} A_{i\ell} A_{-p_i}\,.$$

Satz 83: *Sei (A,b) ein System linearer Gleichungen über K mit $A \in K^{m\times n}$ in Stufenform. Sei r die Anzahl der Pivots, $p_1 < p_2 < \dots < p_r$ die Spaltenindizes der Pivots und $\{q_1,\dots,q_{n-r}\} := \{1,\dots,n\} \setminus \{p_1,\dots,p_r\}$. Dann gilt:*

(1) *$L(A,b)$ ist genau dann nicht leer, wenn für alle $i > r$ gilt: $b_i = 0$. In diesem Fall ist*

$$z := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_1 \\ 0 \\ \vdots \\ 0 \\ b_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in L(A,b),$$

wobei die Elemente $b_1,\dots,b_r$ in den Zeilen mit Index $p_1,\dots,p_r$ stehen.

(2) *Sei*

$$w_j := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -A_{1q_j} \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ -A_{rq_j} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$1 \leq j \leq n-r$, wobei die Elemente $-A_{1q_j}, \ldots, -A_{rq_j}$ in den Zeilen mit Index $p_1, \ldots, p_r$ sowie 1 in der Zeile mit Index q_j stehen.
Dann ist $(w_1, \ldots, w_{n-r})$ eine K-Basis von $L(A, 0)$.

Beweis: Seien $e_1, \ldots, e_m$ die Standardspalten von $K^{m \times 1}$. Dann ist

$$A_{-p_1} = e_1, \ldots, A_{-p_r} = e_r\,.$$

(1) Wenn $L(A,b)$ nicht leer ist, dann gibt es ein y mit $Ay = b$, und es folgt $b_i = A_{i-}y = 0$ für $i > r$. Sei umgekehrt $b_i = 0$ für $i > r$. Sei $z \in K^{n \times 1}$. Nach Satz 76 ist $Az = b$ genau dann, wenn

$$\sum_{i=1}^{r} b_i e_i = b = \sum_{i=1}^{n} z_i A_{-i} = \sum_{i=1}^{r} z_{p_i} A_{-p_i} + \sum_{i=1}^{n-r} z_{q_i} A_{-q_i} =$$

$$= \sum_{i=1}^{r} z_{p_i} e_i + \sum_{i=1}^{n-r} z_{q_i} A_{-q_i}$$

ist. Wählt man daher $z_{q_i} := 0$ für $1 \leq i \leq n-r$ und $z_{p_i} := b_i$, $1 \leq i \leq r$, dann ist $Az = b$.

(2) Für $1 \leq j \leq n$ ist

$$A_{-j} = \sum_{i=1}^{r} A_{ij} e_i = \sum_{i=1}^{r} A_{ij} A_{-p_i}\,,$$

also

$$A_{-j} - \sum_{i=1}^{r} A_{ij} A_{-p_i} = 0\,.$$

Nach Satz 76 ist also $w_j \in L(A,0)$.

Seien $t_1,\ldots,t_{n-r} \in K$. Für $1 \leq j \leq n-r$ ist dann die q_j-te Komponente der Spalte $\sum_{i=1}^{n-r} t_i w_i$ gleich t_j. Wenn die Linearkombination gleich 0 ist, müssen auch alle t_j, $1 \leq j \leq n-r$ gleich 0 sein. Daher ist $(w_1,\ldots,w_{n-r})$ linear unabhängig.

Es bleibt noch zu zeigen, dass jedes $v \in L(A,0)$ als Linearkombination von $(w_1,\ldots,w_{n-r})$ geschrieben werden kann. Wir zeigen dazu, dass

$$v = \sum_{j=1}^{n-r} v_{q_j} w_j$$

ist. Für $1 \leq i \leq n-r$ ist

$$\Big(\sum_{j=1}^{n-r} v_{q_j} w_j\Big)_{q_i} = \sum_{j=1}^{n-r} v_{q_j} \delta_{ij} = v_{q_i}\,,$$

es genügt also zu zeigen, dass $(\sum_{j=1}^{n-r} v_{q_j} w_j)_{p_i} = v_{p_i}$ ist, $1 \leq i \leq r$. Aus $Av = 0$ folgt für alle $i \leq r$:

$$0 = A_{i-}v = \sum_{j=1}^{n} A_{ij} v_j = \sum_{\ell=1}^{r} A_{ip_\ell} v_{p_\ell} + \sum_{j=1}^{n-r} A_{iq_j} v_{q_j} = v_{p_i} + \sum_{j=1}^{n-r} A_{iq_j} v_{q_j}\,.$$

Somit gilt für $1 \leq i \leq r$

$$v_{p_i} = \sum_{j=1}^{n-r} (-A_{iq_j}) v_{q_j} = \sum_{j=1}^{n-r} (w_j)_{p_i} v_{q_j} = \Big(\sum_{j=1}^{n-r} v_{q_j} w_j\Big)_{p_i}\,.$$

Beispiel 84 : Der Lösungsmenge der linearen Gleichung

$$((1\ a_2\ a_3\ \ldots\ a_n),(b_1))$$

(oder, anders geschrieben, $x_1 + a_2x_2 + a_3x_3 + \ldots + a_nx_n = b_1$) ist

$$\left\{ \begin{pmatrix} b_1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} -a_2 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -a_3 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + c_{n-1} \begin{pmatrix} -a_n \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \,\middle|\, c_1,\ldots,c_{n-1} \in K \right\}.$$

Unser nächstes Ziel ist es, beliebige Systeme linearer Gleichungen zu lösen. Ein erster Schritt dazu ist der folgende Satz.

Satz 85 : *Sei (A,b) ein System linearer Gleichungen mit $A \in K^{m\times n}$. Dann gilt für alle $P \in \mathrm{GL}_m(K)$*

$$L(PA,Pb) = L(A,b).$$

Beweis: Ist $y \in L(A,b)$, dann ist $Ay = b$, also auch $PAy = Pb$ und $y \in L(PA,Pb)$. Daher ist $L(A,b)$ eine Teilmenge von $L(PA,Pb)$.
Ist $y \in L(PA,Pb)$, dann ist $PAy = Pb$, also auch $Ay = P^{-1}PAy = P^{-1}Pb = b$ und $y \in L(A,b)$. Daher ist $L(PA,Pb)$ eine Teilmenge von $L(A,b)$.

Satz 85 legt nahe, zu einem gegebenen System linearer Gleichungen (A,b) eine invertierbare Matrix P zu suchen, sodass das System (PA,Pb) Stufenform hat. Dann kann die Lösungsmenge mit Satz 83 bestimmt werden.

Satz 86 : *Jede Matrix $A \in K^{m\times n}$ kann durch elementare Zeilenumformungen auf Stufenform gebracht werden, genauer gesagt gibt es eine Matrix $P \in \mathrm{GL}_m(K)$, die Produkt von höchstens mn Elementarmatrizen ist, sodass PA Stufenform hat. Diese Matrix PA kann mit dem folgenden Algorithmus berechnet werden* (Gauss-Elimination)*:*

(1) *Setze $C := A$, $i := 1$ und $j := 1$.*
(2) *Falls $C_{ij} \neq 0$ ist, gehe zu Schritt 4.*
 Ansonsten suche ein $k \in \{i+1,\dots,m\}$ mit $C_{kj} \neq 0$.
(3) *Falls es kein $k \in \{i+1,\dots,m\}$ mit $C_{kj} \neq 0$ gibt, erhöhe j um 1 und wiederhole Schritt 2.*
 Ansonsten vertausche die i-te und k-te Zeile von C und nenne die neue Matrix wieder C. (Dann ist $C_{ij} \neq 0$).
(4) *Multipliziere die i-te Zeile von C mit C_{ij}^{-1} und nenne die neue Matrix wieder C. (Dann ist $C_{ij} = 1$).*
(5) *Für $\ell \in \{1,\dots,m\} \setminus \{i\}$ mit $C_{\ell j} \neq 0$ subtrahiere das $C_{\ell j}$-fache der i-ten Zeile von der ℓ-ten Zeile und nenne die neue Matrix wieder C. (Dann ist $C_{-j} = e_i$).*
(6) *Erhöhe i und j um 1 und gehe zu Schritt 2.*

Der Algorithmus wird abgebrochen, sobald $i > m$ oder $j > n$ ist. Dann hat die Matrix C Stufenform.

Beweis: Der Algorithmus bricht nach höchstens n Durchläufen ab, weil in jedem Durchlauf j um 1 erhöht wird. Seien $P_1,\dots,P_s$ die Elementarmatrizen zu den im Algorithmus der Reihe nach durchgeführten elementaren Zeilenumformungen. Da pro Durchlauf höchstens m elementare Zeilenumformungen durchgeführt werden, ist $s \leq nm$. Nach Satz 60 erhält man am Ende des Algorithmus die Matrix

$$(P_s \dots (P_2(P_1A))\dots) = (P_s \dots P_2P_1)A,$$

und nach Satz 61 ist

$$P := P_s \dots P_2P_1 \in \mathrm{GL}_m(K).$$

Schließlich hat die Matrix PA Stufenform, weil in jedem Durchlauf die ersten $j-1$ Spalten nicht mehr verändert werden und in der j-ten Spalte entsprechend umgeformt wird.

Der Beweis zeigt, dass man die Matrix P erhält, indem man die elementaren Zeilenumformungen nicht nur auf A, sondern auch auf die Einheitsmatrix I_m anwendet:

$$(A|I_m) \to (P_1A|P_1I_m) \to \cdots \to (P_s \dots P_1A|P_s \dots P_1I_m) = (PA|P).$$

Satz 87 : *Ein System linearer Gleichungen (A,b) über dem Körper K kann gelöst werden, indem man die Matrix A durch Gauss-Elimination auf Stufenform bringt und b „mittransformiert". Dazu führt man die elementaren Zeilenumformungen an der „erweiterten Matrix" $(A|b)$ aus und erhält $(PA|Pb)$. Dann ist*

$$L(A,b) = L(PA,Pb)$$

und $L(A,b)$ kann nach Satz 83 berechnet werden.
Insbesondere gibt es genau dann genau eine Lösung, wenn

$$PA = \begin{pmatrix} I_n \\ 0 \end{pmatrix} \qquad \textit{und} \qquad Pb = \begin{pmatrix} c \\ 0 \end{pmatrix}$$

mit $c \in K^{n\times 1}$ ist. Die eindeutige Lösung ist dann

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

Wenn A invertierbar ist, dann hat (A,b) die eindeutige Lösung

$$A^{-1}b.$$

Beweis: Die Behauptung folgt aus Satz 85 und Satz 83.

Satz 88 : *Ein homogenes System linearer Gleichungen $(A,0)$ über dem Körper K hat immer die* triviale Lösung 0. *Wenn $A \in K^{m\times n}$ mit $m < n$ ist, d.h. es sind weniger Gleichungen als Unbekannte vorhanden, dann gibt es auch eine Lösung, die nicht trivial ist.*

Beweis: Wenn $m < n$ ist, dann kann $PA \in K^{m\times n}$ nicht die Form

$$\begin{pmatrix} I_n \\ 0 \end{pmatrix}$$

haben, und nach Satz 87 gibt es mehr als eine Lösung.

Satz 89: *Sei $A \in K^{m\times m}$. Auf die folgende Weise kann überprüft werden, ob A invertierbar ist und, wenn ja, die zu A inverse Matrix A^{-1} berechnet werden:*

Bringe A durch Gauss-Elimination auf Stufenform, wobei die elementaren Zeilenoperationen auch auf die Einheitsmatrix angewendet werden:

$$(A|I_m) \to (PA|P).$$

Dann ist A invertierbar genau dann, wenn $PA = I_m$ ist. In diesem Fall ist $A^{-1} = P$.
Insbesondere kann jede invertierbare Matrix als Produkt von Elementarmatrizen geschrieben werden.

Beweis: Wenn A invertierbar ist, dann sind $A, P \in \mathrm{GL}_m(K)$ und somit auch $B := PA \in \mathrm{GL}_m(K)$. Da B Stufenform hat, ist $B = I_m$ oder $B_{m-} = 0$. Aus $B_{m-} = 0$ würde aber

$$0 = B_{m-}(B^{-1})_{-m} = (BB^{-1})_{mm} = 1,$$

folgen, also muss $B = I_m$ sein. Umgekehrt folgt aus $PA = I_m$ auch
$AP = P^{-1}(PA)P = I_m$ und somit $P = A^{-1}$.

Beispiel 90: Die Matrix

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

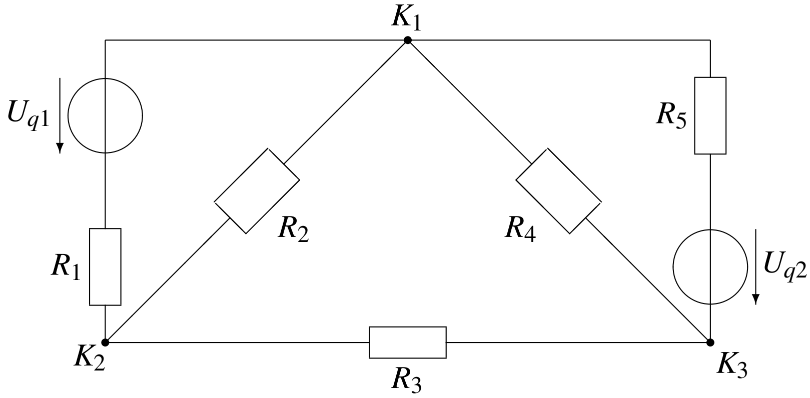
ist genau dann invertierbar, wenn $ad - bc \neq 0$ ist. In diesem Fall ist

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

§5. Kirchhoff'sche Gesetze und Systeme linearer Gleichungen

Mit den Methoden von §4 kann die folgende Aufgabe aus der Elektrotechnik gelöst werden.

In der folgenden Schaltung sind die Spannungen U_{q1} und U_{q2}, sowie die Widerstände R_1, R_2, R_3, R_4, R_5 bekannt. Gesucht sind die Ströme
I_1, I_2, I_3, I_4, I_5 durch die Widerstände R_1, R_2, R_3, R_4, R_5.
Die Spannung wird in Volt (V), die Stromstärke in Ampère (A) und der Widerstand in Ohm (Ω) gemessen.



Wir beschreiben diese Aufgabe als System linearer Gleichungen. Für die Modellierung brauchen wir die folgenden physikalischen Gesetze:

Ohm'sches Gesetz

$$R \cdot I = U$$

(Widerstand mal Stromstärke ist Spannung)

Kirchhoff'sche Gesetze

Die Kirchhoff'schen Gesetze beschreiben jeweils den Zusammenhang zwischen mehreren elektrischen Strömen bzw. zwischen mehreren elektrischen Spannungen in einem elektrischen Netzwerk.

Werden mehrere Leitungen in einem Punkt verbunden, nennt man diesen einen *Knoten* der Schaltung (im Bild oben sind K_1, K_2 und K_3 Knoten). Leitungen zwischen zwei Knoten nennt man *Zweige* der Schaltung.
Für jeden Zweig wird eine Richtung vorgegeben, d.h. eine Reihenfolge seiner zwei Knoten.
Die Spannung und die Stromstärke haben positives Vorzeichen, wenn die (angenommene) Richtung des Stroms mit der Richtung des Zweiges übereinstimmt, sonst negatives.
Ein Strom durch einen Zweig heißt in *a zufließend*, wenn a der zweite Knoten des Zweiges ist, und von *a abfließend*, wenn a der erste Knoten des Zweiges ist.

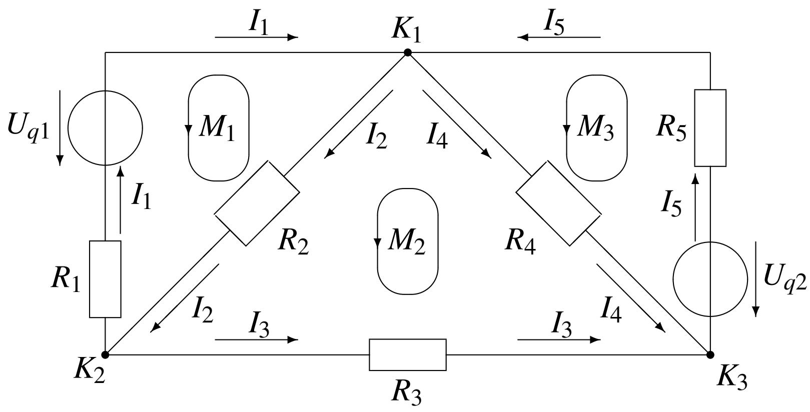
1. Kirchhoff'sches Gesetz (Knotensatz):
   Die Summe der zufließenden Ströme in einem Knoten ist gleich der Summe der abfließenden Ströme.

Eine endliche Menge $M \neq \{\}$ von Zweigen mit der folgenden Eigenschaft heißt *Masche*: Ist $\ell \in M$ und ist a ein Knoten von ℓ, dann gibt es einen eindeutig bestimmten Zweig $k \in M \setminus \{\ell\}$, sodass a auch ein Knoten von k ist.
Auf die folgende Weise legen wir eine *Umlaufrichtung* der Masche fest: Es sei M eine Masche und $\{a_1, \ldots, a_n\}$ die Menge aller Knoten dieser Masche. Wir setzen $a_{n+1} := a_1$. Wir können die Indizes so wählen, dass die Zweige der Masche zwischen den Knoten a_i und a_{i+1}, $1 \leq i \leq n+1$, verlaufen und

dass a_1 der erste Knoten des Zweiges zwischen a_1 und a_2 ist.
Dann ist der Zweig zwischen a_i und a_{i+1} *in der Umlaufrichtung* der Masche, wenn a_i sein erster Knoten ist, sonst *gegen die Umlaufrichtung*.

2. Kirchhoff'sches Gesetz (Maschensatz):
   Die Summe der Spannungen von Zweigen einer Masche in der Umlaufrichtung ist gleich der Summe der Spannungen von Zweigen einer Masche gegen die Umlaufrichtung.

Wir wählen Stromrichtungen von I_j, für $j = 1, \dots, 5$ und Richtungen für die Quellspannungen, somit auch Richtungen der Zweige.



Aus den Kirchhoff'schen Gesetzen erhalten wir die folgenden Bedingungen für die Zahlen $I_1, \dots, I_n$:

Knotengleichungen:

$$\begin{array}{rrcl}
K_1: & I_1 + I_5 & = & I_2 + I_4 \\
K_2: & I_2 & = & I_1 + I_3 \\
K_3: & I_3 + I_4 & = & I_5
\end{array}$$

Maschengleichungen:

$$\begin{array}{rccc}
M_1 & : & I_1 \cdot R_1 + I_2 \cdot R_2 & = & U_{q1} \\
M_2 & : & I_2 \cdot R_2 + I_3 \cdot R_3 & = & I_4 \cdot R_4 \\
M_3 & : & I_4 \cdot R_4 + I_5 \cdot R_5 & = & U_{q2} \\
M_4 & : & I_1 \cdot R_1 + I_4 \cdot R_4 & = & I_3 \cdot R_3 + U_{q1} \\
M_5 & : & I_2 \cdot R_2 + I_3 \cdot R_3 + I_5 \cdot R_5 & = & U_{q2} \\
M_6 & : & I_1 \cdot R_1 + U_{q2} & = & U_{q1} + I_3 \cdot R_3 + I_5 \cdot R_5
\end{array}$$

Die Bedingungen K_1, M_4, M_5, M_6 folgen aus K_2, K_3, M_1, M_2, M_3, also können wir sie weglassen.

Man erhält nun das System linearer Gleichungen

$$\begin{pmatrix} -1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ R_1 & R_2 & 0 & 0 & 0 \\ 0 & R_2 & R_3 & -R_4 & 0 \\ 0 & 0 & 0 & R_4 & R_5 \end{pmatrix} \cdot \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ U_{q1} \\ 0 \\ U_{q2} \end{pmatrix}.$$

Wir berechnen I_1, I_2, I_3, I_4, I_5 für $U_{q1} = 18.6\,V$, $U_{q2} = 13.2\,V$ und $R_1 = 1.5\,k\Omega$, $R_2 = 680\,\Omega$, $R_3 = 470\,\Omega$, $R_4 = 520\,\Omega$, $R_5 = 710\,\Omega$:

$$\begin{pmatrix} -1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ 1500 & 680 & 0 & 0 & 0 \\ 0 & 680 & 470 & -520 & 0 \\ 0 & 0 & 0 & 520 & 710 \end{pmatrix} \cdot \begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 18.6 \\ 0 \\ 13.2 \end{pmatrix},$$

die (eindeutig bestimmte) Lösung ist

$$\begin{aligned} I_1 &= 0.00859\,A &&= 8.59\,mA \\ I_2 &= 0.00841\,A &&= 8.41\,mA \\ I_3 &= -0.00018\,A &&= -0.18\,mA \\ I_4 &= 0.01083\,A &&= 10.83\,mA \\ I_5 &= 0.01066\,A &&= 10.66\,mA \end{aligned}$$

Das negative Vorzeichen von I_3, bedeutet, dass dieser Strom gegen die gewählte Richtung des Zweiges fließt.

§6. Dimension

Satz 91 : *Seien V ein Vektorraum über K, $(u_1, \ldots, u_m)$ ein linear unabhängiges m-Tupel von Vektoren in V und $v \in V$ mit*

$$v \notin {}_K\langle u_1, \ldots, u_m \rangle .$$

Dann ist auch das $m+1$-Tupel $(u_1, \ldots, u_m, v)$ linear unabhängig.

Beweis: Seien $c_1, \ldots, c_m \in K$ und $d \in K$ mit

$$\sum_{j=1}^{m} c_j u_j + dv = 0 .$$

Wenn $d \neq 0$ ist, wäre

$$v = -\sum_{j=1}^{m} d^{-1} c_j u_j \in {}_K\langle u_1, \ldots, u_m \rangle$$

im Widerspruch zur Voraussetzung. Somit ist $d = 0$ und, da $(u_1, \ldots, u_m)$ linear unabhängig ist, auch $c_j = 0$ für alle $1 \leq j \leq m$.

Definition 92 : Ein Vektorraum V heißt *endlich-erzeugt*, wenn es ein endliches Erzeugendensystem von V gibt.

Satz 93 : *Sei $\{0\} \neq V$ ein endlich-erzeugter Vektorraum über K und $(w_1, \ldots, w_n)$ ein Erzeugendensystem von V.*
Dann gilt:

(1) *Jedes linear unabhängige m-Tupel $(u_1, \ldots, u_m)$ von Vektoren in V kann durch Vektoren aus $\{w_1, \ldots, w_n\}$ zu einer Basis von V ergänzt werden, d.h. es gibt Indizes $j_1, \ldots, j_k \in \{1, \ldots, n\}$, sodass $(u_1, \ldots, u_m, w_{j_1}, \ldots, w_{j_k})$ eine Basis von V ist.*
(2) *Aus $(w_1, \ldots, w_n)$ kann eine Basis ausgesondert werden, d.h. es gibt eine Teilmenge $\{i_1, i_2, \ldots i_\ell\}$ von $\{1, \ldots, n\}$, sodass $(w_{i_1}, \ldots, w_{i_\ell})$ eine Basis von V ist.*
Insbesondere hat jeder endlich-erzeugte Vektorraum V über K eine Basis.

Beweis:

(1) Wir beweisen die Aussage durch Induktion nach der Anzahl k der Vektoren in $\{w_1, \ldots, w_n\}$, die nicht in $\langle u_1, \ldots, u_m \rangle$ enthalten sind.
Wenn $k = 0$ ist, dann ist $\langle u_1, \ldots, u_m \rangle = \langle w_1, \ldots, w_n \rangle = V$, also ist $(u_1, \ldots, u_m)$ bereits eine Basis.
Wenn $k > 0$ ist, dann gibt es einen Vektor w_i, der nicht Element von $\langle u_1, \ldots, u_m \rangle$ ist. Nach Satz 91 ist das $m+1$-Tupel $(u_1, \ldots, u_m, w_i)$ linear unabhängig. Die Anzahl der Vektoren in $\{w_1, \ldots, w_n\}$, die nicht in $\langle u_1, \ldots, u_m, w_i \rangle$ enthalten sind, ist kleiner als k. Daher folgt aus der Induktionsannahme, dass $(u_1, \ldots, u_m, w_i)$ (und damit auch $(u_1, \ldots, u_m)$) durch Vektoren aus $\{w_1, \ldots, w_n\}$ zu einer Basis von V ergänzt werden kann.
(2) Weil $V \neq \{0\}$ ist, ist mindestens ein Element von $\{w_1, \ldots, w_n\}$ nicht 0. Sei also $w_i \neq 0$. Dann ist das 1-Tupel (w_i) linear unabhängig, nach (1) kann es durch Elemente von $\{w_1, \ldots, w_n\}$ zu einer Basis von V ergänzt werden. Diese Basis ist die gesuchte.

Satz 94 : *Sei V ein endlich-erzeugter Vektorraum über K. Dann enthalten je zwei Basen von V gleich viele Vektoren.*

Beweis: Sei $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V und $(w_1, \ldots, w_n)$ in V linear unabhängig. Wir zeigen zuerst, dass $n \leq m$ sein muss.
Weil $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V ist, können wir w_j als Linearkombination von $(v_1, \ldots, v_m)$ schreiben:

$$w_j = \sum_{i=1}^{m} A_{ij} v_i \,, 1 \leq j \leq n.$$

Wäre $m < n$, hätte die Matrix $A := (A_{ij})_{1\le i\le m, 1\le j\le n}$ mehr Spalten als Zeilen, also gäbe es nach Satz 88 eine n-Spalte $x \neq 0$ mit $Ax = 0$. Daher wäre

$$\sum_{j=1}^{n} x_j w_j = \sum_j x_j (\sum_i A_{ij} v_i) = \sum_i (\sum_j A_{ij} x_j) v_i = \sum_i 0 \cdot v_i = 0\,,$$

im Widerspruch zu $x \neq 0$ und $(w_1, \dots, w_n)$ linear unabhängig.
Wenn nun sowohl $(v_1, \dots, v_m)$ als auch $(w_1, \dots, w_n)$ eine Basis ist, ist einerseits $n \le m$, weil $(v_1, \dots, v_m)$ ein Erzeugendensystem von V und $(w_1, \dots, w_n)$ linear unabhängig ist, und andererseits $m \le n$, weil auch $(w_1, \dots, w_n)$ ein Erzeugendensystem von V und $(v_1, \dots, v_m)$ linear unabhängig ist.

Definition 95 : Sei $\{0\} \neq V$ ein endlich-erzeugter Vektorraum über K. Die Zahl

$$\dim_K(V) := \text{Anzahl der Elemente einer Basis von } V$$

(ist nach Satz 94 wohldefiniert und) heißt die *Dimension* von V.
Für den Nullraum $\{0\}$ definieren wir: die leere Menge ist eine Basis von $\{0\}$ und $\dim_K(V) = 0$.
Sprechweisen: Wenn $\dim_K(V) = n$ ist, dann nennt man V *n-dimensional*. Ein endlich-erzeugter Vektorraum heißt *endlich-dimensional*. Ein nicht endlich-erzeugter Vektorraum W heißt *unendlich-dimensional*.

Die Dimension eines Vektorraums ist also die kleinste Anzahl von Zahlen (oder, allgemeiner, von Elementen von K), die für die vollständige Beschreibung eines Vektors in V nötig sind.

Beispiel 96 : Nach Satz 79 ist
$\dim_K(K^n) = n$ und $\dim_K(K^{m\times n}) = m \cdot n$.

Satz 97 : *Sei V ein Vektorraum über K der Dimension n und $(v_1, \dots, v_n) \in V^n$. Dann sind die folgenden Aussagen äquivalent:*

(1) *$(v_1, \dots, v_n)$ ist eine Basis von V.*
(2) *$(v_1, \dots, v_n)$ ist ein Erzeugendensystem von V.*
(3) *$(v_1, \dots, v_n)$ ist linear unabhängig.*

Beweis: Aus (1) folgen immer (2) und (3).
(2) $\Rightarrow$ (1): Nach Satz 93 kann aus $(v_1, \dots, v_n)$ eine Basis ausgesondert werden, die wegen $\dim_K(V) = n$ wieder aus n Vektoren besteht. Somit muss $(v_1, \dots, v_n)$ eine Basis sein.
(3) $\Rightarrow$ (1): analog.

Definition 98: Sei V ein Vektorraum über K, $w \in V$ und $(v_1, \dots, v_n)$ eine Basis von V. Das eindeutig bestimmte n-Tupel $(c_1, \dots, c_n)$ von Skalaren mit

$$w = \sum_{i=1}^{n} c_i v_i$$

heißt das *Koordinaten-n-Tupel* des Vektors w bezüglich der Basis $(v_1, \dots, v_n)$. Das Element $c_i \in K$ heißt die *Koordinate* von w beim Basisvektor v_i (oder die i-te Koordinate von w bezüglich der Basis $(v_1, \dots, v_n)$). Die Spalte

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in K^{n \times 1}$$

heißt *Koordinatenspalte* des Vektors w bezüglich der Basis $(v_1, \dots, v_n)$.

Satz 99: *Sei V ein Vektorraum über K und $(v_1, \dots, v_n)$ eine Basis von V. Seien $w_1, \dots, w_m \in V$ und $T_{-1}, \dots, T_{-m}$ deren Koordinatenspalten bezüglich $(v_1, \dots, v_n)$.*

(1) *Seien $c_1, \dots, c_n \in K$. Die Koordinatenspalte von $\sum_{i=1}^{m} c_i w_i$ bezüglich $(v_1, \dots, v_n)$ ist $\sum_{i=1}^{m} c_i T_{-i}$. („Die Koordinatenspalte der Linearkombination ist die Linearkombination der Koordinatenspalten“).*
(2) *Das m-Tupel $(w_1, \dots, w_m)$ ist genau dann ein Erzeugendensystem von V bzw. linear unabhängig in V bzw. eine Basis von V, wenn das entsprechende für das m-Tupel $(T_{-1}, \dots, T_{-m})$ seiner Koordinatenspalten gilt.*

Beweis: (1) ist leicht nachzuprüfen, (2) folgt aus (1).

Aus Satz 99 folgt, dass man das Rechnen in *jedem* endlich-dimensionalen Vektorraum durch die Wahl einer Basis dieses Vektorraums auf das (komponentenweise) Rechnen mit n-Tupeln zurückführen kann.

Definition 100: Der *Rang* einer Matrix $A \in K^{m \times n}$ ist die Dimension des Spaltenraums von A (siehe Definition 81). Schreibweise: $\mathrm{rg}(A)$.

Satz 101: *Es sei $A \in K^{m \times n}$, $P \in \mathrm{GL}_m(K)$, $1 \le \ell \le n$ und $1 \le i_1 < i_2 < \ldots < i_\ell \le n$. Dann gilt:*

(1) *Das ℓ-Tupel der Spalten $(A_{-i_1}, \dots, A_{-i_\ell})$ von A ist genau dann linear unabhängig, wenn das n-Tupel der Spalten $((PA)_{-i_1} = PA_{-i_1}, \dots, (PA)_{-i_\ell} = PA_{-i_\ell})$ von PA linear unabhängig ist.*

(2) *Das ℓ-Tupel der Spalten $(A_{-i_1},\ldots,A_{-i_\ell})$ von A ist genau dann ein Erzeugendensystem des Spaltenraums von A, wenn das n-Tupel der Spalten $((PA)_{-i_1},\ldots,(PA)_{-i_\ell})$ von PA ein Erzeugendensystem des Spaltenraums von PA ist.*

(3) *Das ℓ-Tupel der Spalten $(A_{-i_1},\ldots,A_{-i_\ell})$ ist genau dann eine Basis des Spaltenraums von A, wenn das n-Tupel der Spalten $((PA)_{-i_1},\ldots,(PA)_{-i_\ell})$ eine Basis des Spaltenraums von PA ist. Insbesondere sind der Rang von A und von PA gleich.*

Beweis:

(1) Seien $c_1,\ldots,c_\ell \in K$. Weil P invertierbar ist, ist

$$\sum_{j=1}^{\ell} c_j(PA)_{-i_j} = P(\sum_{j=1}^{\ell} c_jA_{-i_j}) = 0$$

genau dann, wenn

$$\sum_{j=1}^{\ell} c_jA_{-i_j} = 0$$

ist.

(2) Wir nehmen an, dass $(A_{-i_1},\ldots,A_{-i_\ell})$ ein Erzeugendensystem des Spaltenraums von A ist. Sei $S := \sum_{j=1}^{n} d_j(PA)_{-j} = P\cdot\sum_{j=1}^{n} d_jA_{-j}$ ein Element des Spaltenraums von PA. Dann gibt es Elemente $b_1,\ldots,b_\ell \in K$ so, dass $\sum_{j=1}^{n} d_jA_{-j} = \sum_{j=1}^{\ell} b_jA_{-i_j}$ ist. Dann ist

$$S = P\cdot\sum_{j=1}^{\ell} b_jA_{-i_j} = \sum_{j=1}^{\ell} b_j(PA)_{-i_j}$$

eine Linearkombination von $((PA)_{-i_1},\ldots,(PA)_{-i_\ell})$. Die andere Implikation wird analog gezeigt.

(3) Folgt aus (1) und (2).

Satz 102 : *Seien $A \in K^{m\times n}$ und $P \in \mathrm{GL}_m(K)$ so, dass PA Stufenform hat. Sei r die Anzahl der Pivots in PA. Dann ist $r = \mathrm{rg}(A)$ und die Dimension von $L(A,0)$ ist $n-\mathrm{rg}(A)$.*
In Worten: Die Dimension des Lösungsraums eines homogenen Systems linearer Gleichungen ist gleich der Anzahl der Unbekannten minus dem Rang der Matrix der Koeffizienten dieses Systems.

Beweis: Nach Satz 83 ist $n-r$ die Dimension von $L(PA,0)$. Da PA Stufenform hat, ist leicht nachzuprüfen, dass $\mathrm{rg}(PA) = r$ ist. Weiters ist $L(PA,0) = L(A,0)$. Also folgen die Behauptungen aus Satz 101.

Satz 103: *Sei $A \in K^{m\times n}$.*

(1) *Mit dem folgenden Verfahren kann* $\mathrm{rg}(A)$ *berechnet und aus dem n-Tupel der Spalten von A eine Basis des Spaltenraums von A ausgesondert werden:*

Bringe A durch elementare Umformungen auf Stufenform PA. Sei r die Anzahl der Pivots von PA, und die Pivots seien in den Spalten mit Indizes $p_1,\ldots,p_r$. Dann ist $\mathrm{rg}(A) = r$ *und* $(A_{-p_1},\ldots,A_{-p_r})$ *ist eine Basis des Spaltenraums von A.*

(2) *Wenn die Spalten von A linear unabhängig sind, kann $(A_{-1},\ldots,A_{-n})$ mit dem folgenden Verfahren zu einer Basis von $K^{m\times 1}$ ergänzt werden:*

Bringe $(A \mid I_m)$ durch elementare Umformungen wie in Satz 86 auf Stufenform $Q(A \mid I_m)$. Dann stehen die Pivots in den Spalten mit Indizes

$$1,2,\ldots,n,n+r_1,\ldots,n+r_{m-n}\,,$$

und $(A_{-1},\ldots,A_{-n},e_{r_1},\ldots,e_{r_{m-n}})$ ist eine Basis von $K^{m\times 1}$. Dabei sind die Spalten e_j, $j = r_1,\ldots,r_{m-n}$, Standardspalten in $K^{m\times 1}$.

Beweis:

(1) Das r-Tupel $(PA_{-p_1},\ldots,PA_{-p_r})$ von Spalten ist eine Basis des Spaltenraums von PA, nach Satz 101 ist daher $(A_{-p_1},\ldots,A_{-p_r})$ eine Basis des Spaltenraums von A.

(2) Der Spaltenraum von $(A \mid I_m)$ ist $K^{m\times 1}$. Da die Spalten von A linear unabhängig sind, stehen in den ersten Spalten n von $Q(A \mid I_m)$ Pivots. Wende (1) auf die Matrix $(A \mid I_m)$ an.

§7. Fragen

Seien K ein Körper, V ein Vektorraum über K und $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren in V.

1. Erzeugendensystem
   Welche der folgenden Aussagen sind wahr?
   (a) Das n-Tupel $(v_1,...,v_n)$ von Vektoren ist ein Erzeugendensystem von V genau dann, wenn jedes Element von V eine Linearkombination von $v_1,...,v_n$ ist.
   (b) Das n-Tupel $(v_1,...,v_n)$ von Vektoren ist ein Erzeugendensystem von V genau dann, wenn sich jedes Element $v \in V$ auf genau eine Weise als Linearkombination von $v_1,...,v_n$ darstellen lässt.
   (c) Das n-Tupel $(v_1,...,v_n)$ von Vektoren ist ein Erzeugendensystem von V genau dann, wenn sich jedes Element $v \in V$ auf höchstens eine Weise als Linearkombination von $v_1,...,v_n$ darstellen lässt.

2. Lineare Unabhängigkeit
   Welche der folgenden Aussagen sind wahr?
   (a) Ist das Tripel von Vektoren (v_1, v_2, v_3) linear unabhängig, dann ist das 4-Tupel (v_1, v_2, v_3, v_4) linear abhängig.
   (b) Besitzt ein Vektor v verschiedene Darstellungen als Linearkombination der Vektoren v_1, v_2, v_3, v_4, dann sind diese Vektoren linear abhängig.
   (c) Ein n-Tupel $(v_1,...,v_n)$ von Vektoren heißt linear unabhängig, wenn für alle $c_1, \ldots, c_n \in K$ aus $\sum_{i=1}^n c_i \cdot v_i = 0$ folgt, dass $c_i = 0$, für $1 \leq i \leq n$, ist.

3. Basis eines Vektorraumes
   Welche der folgenden Aussagen sind wahr?
   (a) Eine Basis von V ist ein linear abhängiges Erzeugendensystem von V.
   (b) Nimmt man von einer Basis einen Vektor weg, so hat man immer noch ein Erzeugendensystem.
   (c) Sind alle Vektoren eines Vektorraums Linearkombinationen der Vektoren $v_1, \ldots, v_n$, und sind diese Vektoren linear unabhängig, dann ist das n-Tupel $(v_1,...,v_n)$ eine Basis.

4. Welche der folgenden Aussagen sind falsch?
   (a) Es sei V der von $(1,1,1)$ und $(2,1,3)$ erzeugte Untervektorraum von $\mathbb{Q}^3$, dann ist $(-4,-1,-7) \in V$.
   (b) Die Vektoren $(2,1)$ und$(4,2)$ bilden eine Basis des Vektorraums $\mathbb{Q}^2$.

(c) Die Vektoren $(3,0,2),(5,2,0),(7,-2,1),(8,0,0) \in \mathbb{Q}^3$ sind linear unabhängig.

In den folgenden drei Aufgaben ist jeweils ein System linearer Gleichungen in Form einer Matrix $(A|b)$ angegeben. Bestimmen Sie welche der drei Aussagen zutrifft und geben Sie die Lösungsmenge $L(A,b)$ an.

5. Lösungsmenge eines Systems linearer Gleichungen

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & -8 \\ 0 & 1 & 0 & -19 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

(a) Das System linearer Gleichungen $(A|b)$ hat genau eine Lösung.
(b) Das System linearer Gleichungen $(A|b)$ hat keine Lösung.
(c) Das System linearer Gleichungen $(A|b)$ hat mehr als eine Lösung.
$L(A,b) = ?$

6.

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

(a) Das System linearer Gleichungen $(A|b)$ hat genau eine Lösung.
(b) Das System linearer Gleichungen $(A|b)$ hat keine Lösung.
(c) Das System linearer Gleichungen $(A|b)$ hat mehr als eine Lösung.
$L(A,b) = ?$

7.

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 0 & 3 & 4 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & -7 \end{array}\right)$$

(a) Das System linearer Gleichungen $(A|b)$ hat genau eine Lösung.
(b) Das System linearer Gleichungen $(A|b)$ hat keine Lösung.
(c) Das System linearer Gleichungen $(A|b)$ hat mehr als eine Lösung.
$L(A,b) = ?$

8. Welche der folgenden Aussagen sind falsch?
(a) Ein System linearer Gleichungen mit gleich vielen Gleichungen wie Unbekannten hat immer genau eine Lösung.
(b) Ein homogenes System linearer Gleichungen hat immer mindestens eine Lösung.
(c) Ein System linearer Gleichungen über $\mathbb{Q}$ hat keine, genau eine oder unendlich viele Lösungen.

Gauss Algorithmus

Berechnen Sie in den folgenden zwei Aufgaben mit dem Gauss-Algorithmus eine Spalte $x \in \mathbb{Q}^{3\times 1}$ so, dass $A \cdot x = b$ ist und eine Basis des Vektorraums $L(A,0) = \{y \in \mathbb{Q}^{3\times 1} | A \cdot y = 0\}$.

9.
$$A := \begin{pmatrix} 0 & 2 & 1 \\ 0 & 1 & 5 \\ 0 & 3 & 2 \end{pmatrix} \in \mathbb{Q}^{3\times 3}, \quad b := \begin{pmatrix} 4 \\ 11 \\ 7 \end{pmatrix} \in \mathbb{Q}^{3\times 1}$$

10.
$$A := \begin{pmatrix} 0 & 3 & 1 \\ 3 & 2 & 5 \\ 1 & 4 & 2 \end{pmatrix} \in \mathbb{Q}^{3\times 3}, \quad b := \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \in \mathbb{Q}^{3\times 1}$$

11. Berechnen der inversen Matrix

Berechnen Sie die inverse Matrix von $\begin{pmatrix} 1 & 0 & 2 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$.

12. Begründen Sie:

Wendet man auf die Einheitsmatrix I_n dieselben elementaren Zeilenumformungen an, welche die Matrix A in die Einheitsmatrix I_n überführen, so erhält man die inverse Matrix von A, d.h.
$(A|I_n) \longrightarrow (I_n|A^{-1})$.

13. Vektorraum

Welche der folgenden Aussagen sind wahr?

(a) Ein Vektorraum ist eine Menge von Pfeilen.

(b) Die Lösungsmenge eines Systems von homogenen linearen Gleichungen (mit der komponentenweisen Addition und Skalarmultiplikation) ist ein Vektorraum.

(c) Sei $K^{n\times 1}$ der Vektorraum aller Spalten mit n Zeilen. Dann ist die Lösungsmenge eines Systems von linearen Gleichungen in n Unbekannten ein Untervektorraum von $K^{n\times 1}$.

14. Dimension eines Vektorraums

Verschiedene Basen eines endlich erzeugten Vektorraums V über K enthalten gleich viele Elemente.

(a) Die Aussage ist falsch.

(b) Die Aussage ist wahr.

15. Welche der folgenden Aussagen sind wahr?

Die Dimension des Vektorraums aller Matrizen mit 4 Zeilen und 5 Spalten, deren erste Zeile 0 ist, ist

(a) 20
(b) 4
(c) 15
(d) höchstens 3
(e) höchstens 5

16. Rang einer Matrix
 Welche der folgenden Aussagen sind wahr?
 (a) Die Anzahl der Spalten einer quadratischen Matrix ist gleich dem Rang der Matrix.
 (b) Die Anzahl der linear unabhängigen Spalten einer Matrix ist gleich dem Rang der Matrix.
17. Der Rang der Matrix $\begin{pmatrix} 1 & 2 & 4 \\ 3 & 5 & 7 \\ 6 & 8 & 9 \end{pmatrix}$ ist
 (a) 9
 (b) 6
 (c) 3
 (d) 2
 (e) 1
18. Der Rang der Matrix $\begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \\ 4 & 5 & 6 \end{pmatrix}$ ist
 (a) 9
 (b) 6
 (c) 3
 (d) 2
 (e) 1

19. Gib die Koordinaten von $\begin{pmatrix} 7 \\ 10 \\ 11 \end{pmatrix} \in \mathbb{Q}^3$ bezüglich der Basis

$$(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix})$$

 an.

20. Auswahl einer Basis aus einem Erzeugendensystem
 Es seien $s_1, \ldots, s_k \in K^{n \times 1}$ Spalten und V der von $s_1, \ldots, s_k$ erzeugte Vektorraum. Um aus $s_1, \ldots, s_k$ eine Basis von V auszuwählen, bringt man die Matrix $(s_1|s_2|\ldots|s_k)$ in Stufenform.
 Seien $i_1, \ldots, i_d$ die Indizes der Spalten mit Pivots.
 Welche der folgenden Aussagen sind wahr?
 (a) Das d-Tupel $(s_1, \ldots, s_d)$ ist eine Basis von V.

(b) Die Spalten s_j, deren Index j ein Element von $\{1,...,k\} \setminus \{i_1,...,i_d\}$ ist, bilden eine Basis von V.
(c) Das d-Tupel $(s_{i_1},...,s_{i_d})$ ist eine Basis von V.

21. Ergänzen einer linear unabhängigen Menge von Spalten zu einer Basis
Es seien $s_1,...,s_k \in K^{n\times 1}$ linear unabhängige Spalten und $e_1,...,e_n$ die Standardspalten in $K^{n\times 1}$.
Welche der folgenden Aussagen sind wahr?
(a) Die Spalten $s_1,...,s_k,e_{k+1},...,e_n$ bilden eine Basis von $K^{n\times 1}$.
(b) Man bringt die Matrix $(s_1|s_2|...|s_k|e_1|...|e_n)$ auf Stufenform. Seien $1,2,...,k,k+i_1,...,k+i_d$ die Indizes der Spalten mit Pivots, dann bilden die Spalten $s_1,...,s_k,e_{i_1},..,e_{i_d}$ eine Basis von $K^{n\times 1}$.
(c) Man bringt die Matrix $(s_1|s_2|...|s_k|e_1|...|e_n)$ auf Stufenform. Seien $1,2,...,k,k+i_1,...,k+i_d$ die Indizes der Spalten mit Pivots und $\{j_1,...,j_{n-d}\} = \{1,...,n\} \setminus \{i_1,...,i_d\}$. Dann bilden die Spalten $s_1,...,s_k,e_{j_1},..,e_{j_{n-d}}$ eine Basis von $K^{n\times 1}$.

KAPITEL 3

Vektorrechnung und Geometrie

In der Vorlesung „Analysis 1“ lernen Sie den Körper $(\mathbb{R},+,\cdot)$ der reellen Zahlen und die Ordnungsrelation $\leq$ auf $\mathbb{R}$ kennen. Ein Vektorraum über dem Körper der reellen Zahlen heißt kurz *reeller Vektorraum*.

§1. Rechnen mit Punkten

In diesem Abschnitt wird gezeigt, wie wir die Zeichenebene und den Anschauungsraum nach Wahl eines „Nullpunktes“ in „natürlicher Weise“ als Vektorraum betrachten können.

Wir nehmen an, wir haben ein „beliebig großes“ Zeichenblatt E und die folgenden Zeichengeräte:

- einen „beliebig fein gespitzten“ Bleistift,
- ein „beliebig langes“ Lineal und
- ein Dreieck.

Wir betrachten das Zeichenblatt als Menge von „Punkten“ und wählen einen davon aus. Diesen ausgewählten Punkt nennen wir *Nullpunkt* und bezeichnen ihn mit $0 \in E$.

Wir nehmen an, dass mit Lineal und Bleistift durch je zwei Punkte eine „Gerade“ gezeichnet werden kann und dass mit Lineal, Dreieck und Bleistift jede Gerade in jeden Punkt „parallelverschoben“ werden kann.

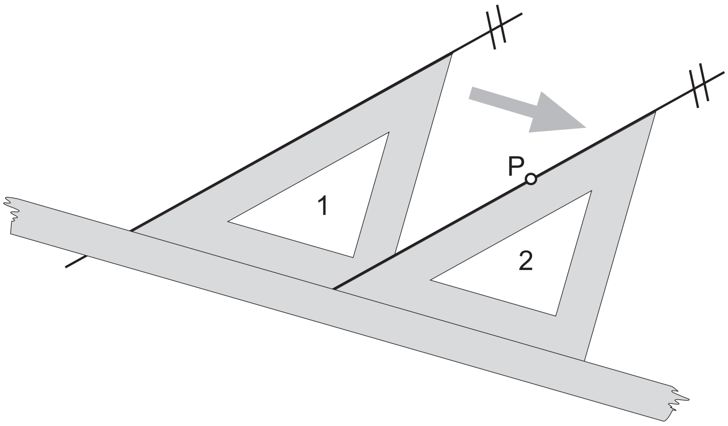


ABBILDUNG 1. Parallelverschieben

Je zwei Punkten $A,B \in E$ können wir wie folgt einen dritten Punkt, den wir mit $A+B$ bezeichnen, zuordnen:

- Falls $0, A$ und B nicht auf einer Geraden liegen:
  Zeichne eine Gerade durch 0 und B und verschiebe sie in den Punkt A. Zeichne eine Gerade durch 0 und A und verschiebe sie in den Punkt B.
  Dann sei $A+B$ der „Schnittpunkt“ dieser zwei Geraden.

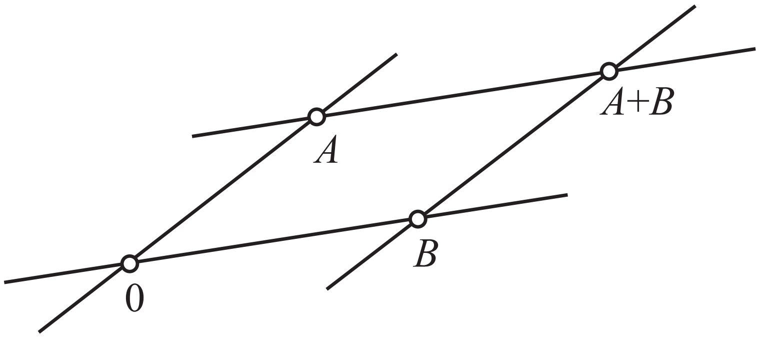


ABBILDUNG 2. Addieren von Punkten in allgemeiner Lage

- Falls $0, A$ und B auf einer Geraden liegen:
  Wähle einen Punkt $H \in E$, der nicht auf dieser Geraden liegt.
  Konstruiere wie oben die Punkte $A+H$ und $(A+H)+B$.
  Verschiebe die Gerade durch 0 und H in den Punkt $(A+H)+B$.
  Dann sei $A+B$ der Schnittpunkt dieser Geraden mit der Geraden durch 0 und A.

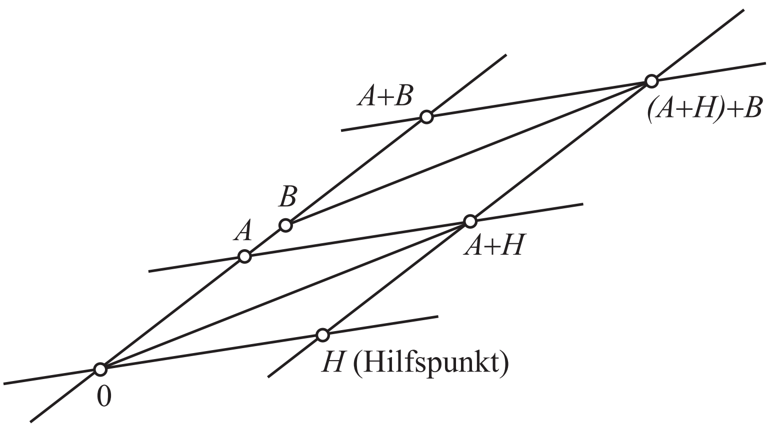


ABBILDUNG 3. Addieren von Punkten in spezieller Lage

Wir nehmen nun zusätzlich an, dass das Lineal mit einer Skala versehen ist, aus der jede reelle Zahl abgelesen werden kann. Dann kann jeder reellen Zahl c und jedem Punkt $A \in E$ ein weiterer Punkt, den wir mit $c \cdot A$ bezeichnen, wie folgt zugeordnet werden:

- Zeichne mit dem Lineal eine Gerade durch 0, die den Punkt A nicht enthält.
- Lege das Lineal so, dass die Zahl 0 über dem Punkt 0 liegt und zeichne dann die Punkte P bzw. Q, über denen die Zahlen 1 bzw. c liegen, auf dieser Geraden ein.
- Verschiebe die Gerade durch A und P in den Punkt Q.
- Dann sei $c \cdot A$ der Schnittpunkt dieser Geraden mit der Geraden durch 0 und A.

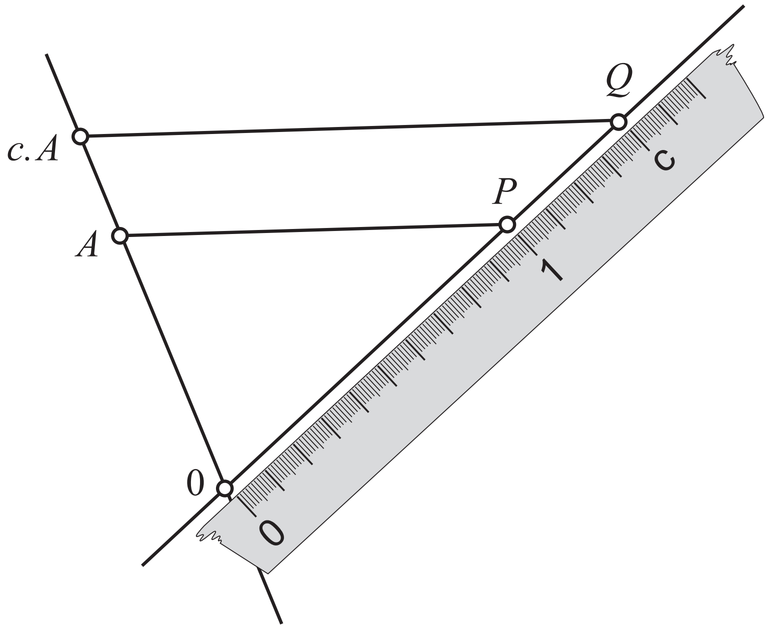


ABBILDUNG 4. Multiplikation von Punkten mit Zahlen

Wir nehmen an, dass die so definierten Rechenoperationen + („Addition von Punkten“) und $\cdot$ („Skalarmultiplikation von reellen Zahlen mit Punkten“) die Rechenregeln eines Vektorraums erfüllen. Dann ist die Zeichenebene E ein Vektorraum und ihre Punkte sind Vektoren.

Zum zeichnerischen Subtrahieren zweier Punkte A und B beachte man, dass $(A - B) + B = A$ ist. Wir beschränken uns auf den Fall, dass 0, A und B nicht auf einer Geraden liegen.
Zeichne die Gerade durch 0 und B und verschiebe sie in den Punkt A. Zeichne die Gerade durch A und B und verschiebe sie in den Punkt 0. Dann ist der „Schnittpunkt“ dieser zwei Geraden jener Punkt, den man zu B addieren muss, um A zu bekommen, also $A - B$.

Alternativ könnte $A - B$ $(= A + (-B))$ auch durch Addition von A und $-B$ erhalten werden.

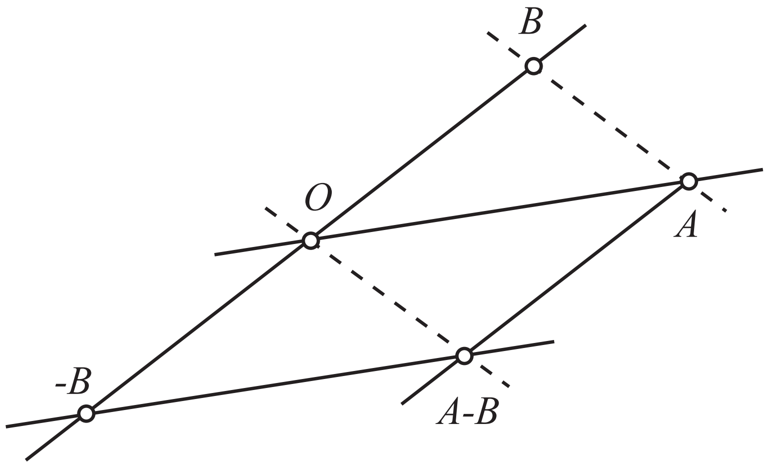


ABBILDUNG 5. Differenz zweier Punkte

Es sei A ein von 0 verschiedener Punkt. Die Gerade durch 0 und A ist (nach Definition der Skalarmultiplikation) die Menge aller skalaren Vielfachen von A. Wenn die Punkte $0, A$ und B nicht auf einer Geraden liegen, dann ist das Punktepaar (A, B) eine $\mathbb{R}$-Basis des Vektorraums E. Wenn a bzw. b die Koordinaten eines Punktes P bezüglich dieser Basis sind, dann erhält man $a \cdot A$ (bzw. $b \cdot B$) als Schnittpunkt der Geraden $\mathbb{R} \cdot A$ mit der Geraden, die man erhält, indem man die Gerade durch 0 und B in den Punkt P verschiebt (bzw. als Schnittpunkt der Geraden $\mathbb{R} \cdot B$ mit der Geraden, die man erhält, indem man die Gerade durch 0 und A in den Punkt P verschiebt). Der Punkt P wird eindeutig durch das Zahlenpaar (a, b) beschrieben.

Die Wahl eines Nullpunktes 0 in der Ebene und von zwei Punkten A, B so, dass $0, A$ und B nicht auf einer Geraden liegen, nennt man auch *Wahl eines Koordinatensystems*. Man kann diese Wahl auch dadurch treffen, dass man ein Paar von Geraden, die genau einen Punkt gemeinsam haben, und auf jeder Geraden einen Punkt, der nicht der Schnittpunkt ist, wählt. Der Schnittpunkt ist dann der Nullpunkt und das Paar der auf den Geraden gewählten Punkte ist die Basis.

Analog kann der Anschauungsraum nach Wahl eines Nullpunktes 0 in natürlicher Weise als Vektorraum betrachtet werden. Wir nehmen dazu an, dass drei Punkte $0, A, B$ jeweils in einer „Ebene“ liegen und definieren dann $A + B$ und $c \cdot A$ wie oben. Wenn die Punkte $0, A, B$ und C nicht in einer Ebene liegen, dann ist das Punktetripel (A, B, C) eine $\mathbb{R}$-Basis dieses Vektorraums.

§2. Affine Unterräume

Definition 104: Sei V ein Vektorraum über einem Körper K. Eine Teilmenge Z von V heißt *affiner Unterraum* von V, wenn ein Vektor $p \in V$ und ein Untervektorraum U von V existieren, sodass

$$Z = p + U := \{p + u \mid u \in U\}$$

ist. In diesem Fall heißt der Vektor p ein *Aufpunkt* von Z und U der zu Z *parallele Untervektorraum.*

Satz 105 : *Sei $Z = p + U$ ein affiner Unterraum von V. Dann gilt:*

(1) *$U = \{w - z \mid w, z \in Z\} =: Z - Z$, insbesondere ist der zu Z parallele Untervektorraum eindeutig durch Z bestimmt.*
(2) *Für alle $q \in Z$ gilt $Z = q + U$, und somit kann jedes Element von Z als Aufpunkt gewählt werden.*
(3) *Z ist ein Untervektorraum von V genau dann, wenn $0_V \in Z$ ist.*

Beweis:

(1) Für $u \in U$ ist $u = (p+u) - (p+0) \in Z - Z$, also $U \subset Z - Z$. Für $w, z \in Z$ gibt es umgekehrt $u, v \in U$ mit $w = p + u$ und $z = p + v$, woraus $w - z = (p+u) - (p+v) = u - v \in U$ folgt. Somit ist auch $Z - Z \subset U$ und insgesamt $U = Z - Z$.

(2) Zu $q \in Z$ existiert $v \in U$ mit $q = p + v$. Für $w = p + u \in Z$ ist dann $w = q + (p - q + u) \in q + U$, also $Z \subset q + U$. Für $u \in U$ ist umgekehrt $q + u = p + (v + u) \in Z$, also auch $q + U \subset Z$.

(3) Wenn $0_V \in Z$ ist, dann ist $Z = 0_V + U = U$ ein Untervektorraum. Wenn umgekehrt Z ein Untervektorraum ist, dann muss nach Definition Z den Nullvektor enthalten.

Satz 106 : *Sei (A, b) ein System linearer Gleichungen über K und $z \in K^{n \times 1}$ eine Lösung. Dann ist die Lösungsmenge $L(A, b)$ ein affiner Unterraum von $K^{n \times 1}$ mit Aufpunkt z und parallelem Untervektorraum $L(A, 0)$, d.h.*

$$L(A, b) = z + L(A, 0).$$

Weiters ist $L(A, b)$ ein Untervektorraum von $K^{n \times 1}$ genau dann, wenn $b = 0$ ist.

Beweis: Nach Satz 65 ist $L(A, 0)$ ein Untervektorraum von $K^{n \times 1}$, und nach Satz 64 somit $L(A, b) = z + L(A, 0)$ ein affiner Unterraum von $K^{n \times 1}$. Nach Satz 105 ist $L(A, b)$ ein Untervektorraum genau dann, wenn $L(A, b)$ die Spalte 0 enthält, d.h. $A \cdot 0 = b$ ist.

Definition 107 : Sei V ein Vektorraum über einem Körper K der Dimension n und

$$Z = p + U$$

ein affiner Unterraum von V mit Aufpunkt p und parallelem Untervektorraum U. Dann heißt

$$\dim_K(Z) := \dim_K(U)$$

die *Dimension* von Z. Im Fall $\dim_K(Z) = 0$ heißt Y ein *Punkt*,
im Fall $\dim_K(Z) = 1$ eine *Gerade*, im Fall $\dim_K(Z) = 2$ eine *Ebene* und im Fall $\dim_K(Z) = n-1$ eine *Hyperebene*.

Beispiel 108 : Es sei E die Zeichenebene, die wir nach Wahl eines Nullpunktes $0 \in E$ als Vektorraum betrachten. Es sei P ein von 0 verschiedener Punkt. Die „Gerade“ durch 0 und P ist die Menge aller skalaren Vielfachen von P, also ist sie ein eindimensionaler Untervektorraum von E. Die „Gerade“ durch zwei Punkte P und Q erhält man, indem man die „Gerade“ durch 0 und $Q-P$ in den Punkt P verschiebt, also ist sie gleich dem eindimensionalen affinen Unterraum

$$P + \mathbb{R} \cdot (Q-P) = \{P + c \cdot (Q-P) \mid c \in \mathbb{R}\} .$$

Der „intuitive Begriff“ „Gerade“ stimmt also mit dem oben definierten Begriff Gerade überein.

Definition 109 : Sei V ein Vektorraum über K und Z ein affiner Unterraum von V.

(1) Wenn ein Aufpunkt $p \in Z$ und eine Basis $(v_1, \ldots, v_k)$ des parallelen Untervektorraums von Z gegeben sind, dann lässt sich jedes Element $z \in Z$ schreiben als

$$z = p + \sum_{i=1}^{k} c_i v_i$$

mit eindeutig bestimmten Elementen $c_1, \ldots, c_k \in K$, und man sagt, dass Z in *Parameterform* gegeben ist.

(2) Sei $V = K^{n\times 1}$, $b \in K^{m\times 1}$ und $A \in K^{m\times n}$ so, dass

$$Z = L(A, b)$$

ist. Dann sagt man, dass Z durch das System linearer Gleichungen (A, b) in *impliziter Form* gegeben ist.

Die Parameterform eines affinen Unterraums hat den Vorteil, dass beliebig viele Elemente von Z angeschrieben werden können. Hingegen lässt sich die Frage, ob eine Spalte $y \in K^{n\times 1}$ in Z enthalten ist, leichter beantworten, wenn Z in impliziter Form gegeben ist, und zwar durch Multiplikation von A mit y.

Definition 110 : Es seien V ein Vektorraum über einem Körper K und $Z_1 = p_1 + U_1$, $Z_2 = p_2 + U_2$ affine Unterräume von V mit Aufpunkten p_1, p_2 und parallelen Untervektorräumen U_1, U_2.

Die affinen Unterräume Z_1 und Z_2 heißen *parallel*, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ ist.

Beispiel 111 : Ein affiner Unterraum und sein paralleler Untervektorraum sind parallel. Jeder Punkt von V ist zu jedem affinen Unterraum von V parallel.

Beispiel 112 : Sei $A \in K^{m\times n}$ eine Matrix und $b, b' \in K^{m\times 1}$ Spalten. Wenn die Lösungsmengen der Systeme linearer Gleichungen (A, b) und (A, b') nicht leer sind, dann sind sie parallel.
Insbesondere: Wenn a_1, a_2, b, b' Elemente von K sind mit $a_1 \neq 0$ oder $a_2 \neq 0$, dann sind die Mengen

$$\{(x,y) \in K^2 \,|\, a_1x + a_2y = b\} \quad \text{und} \quad \{(x,y) \in K^2 \,|\, a_1x + a_2y = b'\}$$

parallele Geraden.

§3. Skalarprodukte

In diesem Abschnitt sei V ein reeller Vektorraum.

Wir werden die folgende Eigenschaft der reellen Zahlen verwenden: Zu jeder reellen Zahl $a \geq 0$ gibt es genau eine reelle Zahl $b \geq 0$ mit $b^2 = a$. Schreibweise: $b = \sqrt{a}$. Ist $0 \leq a < b$, dann ist auch $\sqrt{a} < \sqrt{b}$.

Definition 113 : Ein *Skalarprodukt* auf V ist eine Funktion

$$\langle -,- \rangle : V \times V \longrightarrow \mathbb{R} \ , \quad (v,w) \longmapsto \langle v,w \rangle \ ,$$

mit den folgenden Eigenschaften:
Für alle $c \in \mathbb{R}$, $u, v, w \in V$ gilt

(1) $\langle v,w \rangle = \langle w,v \rangle$
(„$\langle -,- \rangle$ ist *symmetrisch*“)
(2) $\langle u, c(v+w) \rangle = c\langle u,v \rangle + c\langle u,w \rangle$
(„$\langle -,- \rangle$ ist *linear in der zweiten Komponente*“)
(3) Für $v \neq 0$ ist $\langle v,v \rangle$ eine positive reelle Zahl
(„$\langle -,- \rangle$ ist *positiv definit*“).

Aus (1) und (2) folgt:

$$\langle c(u+v), w \rangle = c\langle u,w \rangle + c\langle v,w \rangle \ .$$

(„$\langle -,- \rangle$ ist auch *linear in der ersten Komponente*“, also zusammengefasst: „$\langle -,- \rangle$ ist *bilinear*“)

Ein endlichdimensionaler reeller Vektorraum mit einem Skalarprodukt heißt *euklidischer Raum*.

Satz 114 : *Es seien $\langle -,- \rangle$ ein Skalarprodukt auf V, $(v_1,\dots,v_n)$ eine Basis von V und $c_1,\dots,c_n,d_1,\dots,d_n$ reelle Zahlen. Dann ist*

$$\langle \sum_{i=1}^{n} c_i v_i, \sum_{j=1}^{n} d_j v_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} c_i d_j \langle v_i, v_j \rangle ,$$

insbesondere ist das Skalarprodukt $\langle -,- \rangle$ durch seine Gram'sche Matrix bezüglich der Basis $(v_1,\dots,v_n)$

$$(\langle v_i, v_j \rangle)_{1 \le i,j \le n}$$

eindeutig bestimmt.

Beweis: Induktion über n.

Definition 115 : Ist $\langle -,- \rangle$ ein Skalarprodukt auf V, dann heißen die Funktionen

$$d : V \times V \longrightarrow \mathbb{R} \ , \quad (v,w) \longmapsto ||v-w|| := \sqrt{\langle v-w, v-w \rangle} \ ,$$

bzw.

$$||\cdot|| : V \longrightarrow \mathbb{R} \ , \quad v \longmapsto ||v|| := \sqrt{\langle v,v \rangle} \ ,$$

die von $\langle -,- \rangle$ induzierte *Metrik* bzw. *Norm* auf V. Die Zahl $d(v,w)$ heißt *Abstand* zwischen v und w. Die Zahl $||v|| = d(v,0)$ heißt *Abstand zwischen v und* 0, *Norm*, *Betrag* oder *Länge* von v.

Zwei Vektoren v,w stehen *zueinander senkrecht* oder *orthogonal*, wenn $\langle v,w \rangle = 0$ ist. Schreibweise: $v \perp w$.

Beispiel 116 : Die Funktion

$$\langle -,- \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \ , \quad ((a_1,\dots,a_n),(b_1,\dots,b_n)) \longmapsto \sum_{i=1}^{n} a_i b_i \ ,$$

ist ein Skalarprodukt auf $\mathbb{R}^n$ und heißt *Standardskalarprodukt* auf $\mathbb{R}^n$. Für die Standardbasis $(e_1,\dots,e_n)$ von $\mathbb{R}^n$ gilt

$$\langle e_i, e_j \rangle = \delta_{ij}, \ \|e_i\| = 1 \text{ und } \|e_i - e_j\| = \sqrt{2}(1-\delta_{ij}), \text{ für } 1 \le i,j \le n.$$

Satz 117 : *Es seien V mit $\langle -,- \rangle$ ein reeller Vektorraum mit Skalarprodukt und v,w Vektoren in V. Dann ist*

$$|\langle v,w \rangle| \le \|v\| \cdot \|w\|$$

*(„*Cauchy-Schwarz'sche Ungleichung*“).*

Weiters sind die Zahlen $|\langle v,w\rangle|$ und $\|v\|\cdot\|w\|$ genau dann gleich, wenn v und w linear abhängig sind.

Beweis: Wenn $v=0$ oder $w=0$ ist, dann ist $|\langle v,w\rangle| = 0 = \|v\|\cdot\|w\|$.
Sei nun $v\neq 0$ und $w\neq 0$. Wenn v und w linear abhängig sind, gibt es ein $0\neq c\in\mathbb{R}$ mit $w=c\cdot v$. Daher ist

$$|\langle v,w\rangle| = |c|\langle v,v\rangle = |c|\cdot\|v\|^2 = \|v\|\cdot\|w\| \ .$$

Wenn v und w linear unabhängig sind, dann ist

$$0\neq w-(\langle v,w\rangle/\langle v,v\rangle)v$$

und

$$0 < \langle w-(\langle v,w\rangle/\langle v,v\rangle)v, w-(\langle v,w\rangle/\langle v,v\rangle)v\rangle = \langle w,w\rangle - \langle v,w\rangle^2/\langle v,v\rangle \ .$$

Daher ist

$$\langle v,w\rangle^2 < \|v\|^2\cdot\|w\|^2 \ .$$

Beispiel 118 : Für $V=\mathbb{R}^n$ mit dem Standardskalarprodukt und $a,b\in\mathbb{R}^n$ ergibt sich

$$|\sum_{i=1}^n a_ib_i| \leq \sqrt{\sum_{i=1}^n a_i^2}\sqrt{\sum_{i=1}^n b_i^2}$$

als Cauchy-Schwarz'sche Ungleichung.

Satz 119 : *Es seien V mit $\langle -,-\rangle$ ein reeller Vektorraum mit Skalarprodukt, $c\in\mathbb{R}$ und v,w Vektoren in V. Dann gilt:*

(1) $\|v+w\|\leq\|v\|+\|w\|$ *(* „Dreiecksungleichung“ *).*
(2) *Wenn v und w zueinander orthogonal sind, dann ist*

$$\|v-w\|^2 = \|v\|^2+\|w\|^2$$

(„Satz von Pythagoras“ *).*
(3) *Wenn $\|v\|=\|w\|$ ist, dann stehen $v+w$ und $v-w$ zueinander senkrecht (* „Satz von Thales“ *).*
(4) $\|v+w\|^2-\|v-w\|^2 = 4\langle v,w\rangle$.
Insbesondere: „Eine Parallelogramm ist genau dann ein Rechteck, wenn seine Diagonalen gleich sind.“

Beweis:

(1) $\|v+w\|^2 = \langle v+w,v+w\rangle = \langle v,v\rangle+\langle v,w\rangle+\langle w,v\rangle+\langle w,w\rangle \leq$
$\leq$ (Satz 117) $\|v\|^2+\|w\|^2+2\|v\|\cdot\|w\| = (\|v\|+\|w\|)^2$.
(2) $\|v-w\|^2 = \langle v,v\rangle - 0 + \langle w,w\rangle = \|v\|^2+\|w\|^2$.
(3) $\langle v+w,v-w\rangle = \langle v,v\rangle-\langle w,w\rangle = \|v\|^2-\|w\|^2 = 0$.

(4) $\|v+w\|^2 - \|v-w\|^2 = \langle v,v\rangle + 2\langle v,w\rangle + \langle w,w\rangle -$
$- (\langle v,v\rangle - 2\langle v,w\rangle + \langle w,w\rangle) = 4\langle v,w\rangle$.

Definition 120 : Es sei $u \in V$ und $v \in V,\ v \neq 0$. Mit $\mathbb{R}_{\geq 0}$ bezeichnen wir die Menge aller nicht-negativen reellen Zahlen. Die Menge

$$\mathbb{R}_{\geq 0}v := \{c \cdot v \mid c \in \mathbb{R},\ c \geq 0\}$$

heißt *Richtung in* V. Die Menge

$$u + \mathbb{R}_{\geq 0}v := \{u + c \cdot v \mid c \in \mathbb{R},\ c \geq 0\}$$

heißt *Halbgerade* oder *Strahl in* V *mit Anfangspunkt* u *und Richtung* v.

Sei V ein Vektorraum mit Skalarprodukt. Zu jeder nicht-negativen Zahl $a \in \mathbb{R}$ und jeder Richtung $H := \{cv \mid c \in \mathbb{R},\ c \geq 0\}$ gibt es genau ein Element $w \in H$ mit $\|w\| = a$, und zwar $a\|v\|^{-1} \cdot v$. Also ist jeder Vektor in einem Vektorraum mit Skalarprodukt durch Richtung und Betrag eindeutig bestimmt. Die Sprechweise „ein Vektor hat Betrag und Richtung“ ist daher dann sinnvoll, wenn von Elementen eines Vektorraums mit Skalarprodukt gesprochen wird.

§4. Orthonormalbasen

Sei $\langle -,- \rangle$ ein Skalarprodukt auf einem reellen Vektorraum V.

Definition 121 : Ein n-Tupel $(v_1,\dots,v_n)$ in V heißt *orthonormal* bezüglich $\langle -,- \rangle$, wenn für alle $1 \leq i,j \leq n$

$$\langle v_i, v_j \rangle = \delta_{ij}$$

ist. Ein n-Tupel $(v_1,\dots,v_n)$ in V heißt *Orthonormalbasis* (kurz: *ON-Basis*) von V bezüglich $\langle -,- \rangle$, wenn sie eine Basis von V und orthonormal bezüglich $\langle -,- \rangle$ ist.

Beispiel 122 : Die Standardbasis von $\mathbb{R}^n$ ist eine Orthonormalbasis bezüglich des Standardskalarproduktes.

Beispiel 123 : Eine Basis $(v_1,\dots,v_n)$ von V ist genau dann eine Orthonormalbasis bezüglich $\langle -,- \rangle$, wenn die Gram'sche Matrix

$$(\langle v_i, v_j \rangle)_{1 \leq i,j \leq n}$$

gleich der Einheitsmatrix I_n ist.

Nach Satz 114 bedeutet also „auf einem reellen Vektorraum ein Skalarprodukt wählen“ dasselbe, wie „von einer Basis dieses Vektorraums festlegen, dass sie eine ON-Basis sein soll“. Das nennt man auch “ein rechtwinkeliges Koordinatensystem wählen“.

Satz 124 : *Ein orthonormales n-Tupel ist linear unabhängig. Insbesondere: Wenn V endlich-dimensional ist, dann ist jedes orthonormale n-Tupel mit* $\dim_{\mathbb{R}}(V)$ *Elementen eine ON-Basis von V.*

Beweis: Sei $(v_1,\dots,v_n)$ ein orthonormales n-Tupel und $c_1,\dots c_n \in \mathbb{R}$. Wenn $\sum_{i=1}^n c_i v_i = 0$ ist, dann ist für $1 \le j \le n$ auch

$$0 = \langle v_j, \sum_{i=1}^n c_i v_i \rangle = \sum_{i=1}^n c_i \langle v_j, v_i \rangle = c_j \ .$$

Satz 125 : *Es sei $w \in V$ und $\underline{v} := (v_1,\dots,v_n)$ eine ON-Basis von V. Dann ist*

$$w = \sum_{i=1}^n \langle v_i, w \rangle v_i \ .$$

„Die Koordinate von w bei v_i ist das Skalarprodukt von v_i mit w“.
(Koordinaten von Vektoren bezüglich ON-Basen sind also leicht zu berechnen!)

Beweis: Sei $w = \sum_{j=1}^n c_j v_j$. Für $1 \le i \le n$ ist dann

$$\langle v_i, w \rangle = \langle v_i, \sum_{j=1}^n c_j v_j \rangle = \sum_{j=1}^n c_j \langle v_i, v_j \rangle = \sum_{j=1}^n c_j \delta_{ij} = c_i \ .$$

Beispiel 126 : Es sei $V := \mathbb{R}^2$ und $\langle -,- \rangle$ das Standardskalarprodukt. Dann ist

$$(v_1 := (\frac{3}{5}, \frac{4}{5}), v_2 := (\frac{4}{5}, -\frac{3}{5}))$$

eine ON-Basis von V. Dann ist

$$(1,1) = \langle v_1, (1,1) \rangle v_1 + \langle v_2, (1,1) \rangle v_2 = \frac{7}{5} v_1 + \frac{1}{5} v_2 \ .$$

Satz 127 : *Es seien $(v_1,\dots,v_n)$ eine ON-Basis von V, $u_1,\dots,u_n \in V$ und $S_{-1},\dots,S_{-n}$ die Koordinatenspalten von $u_1,\dots,u_n$ bezüglich der ON-Basis $(v_1,\dots,v_n)$. Dann ist*

$$S_{ij} = \langle v_i, u_j \rangle \,,$$

$1 \le i, j \le n$ *und die folgenden Aussagen sind äquivalent:*

(1) *Das n-Tupel* $(u_1,\dots,u_n)$ *ist eine ON-Basis von* V.
(2) *Die Spalten* $(S_{-1},\dots,S_{-n})$ *bilden eine ON-Basis von* $\mathbb{R}^{n\times 1}$ *mit dem Standardskalarprodukt.*

Beweis: Mit Satz 125 folgt

$$\langle u_i,u_j\rangle = \langle \sum_{k=1}^{n} S_{ki}v_k, \sum_{\ell=1}^{n} S_{\ell j}v_\ell\rangle = \sum_{k,\ell} S_{ki}S_{\ell j}\langle v_k,v_\ell\rangle =$$
$$= \sum_{k,\ell} S_{ki}S_{\ell j}\delta_{k\ell} = \sum_{k} S_{ki}S_{kj} = \langle S_{-i},S_{-j}\rangle\ ,$$

damit ist die Behauptung leicht nachzuprüfen.

Sobald man also eine ON-Basis in einem euklidischen Raum V gewählt hat, wird das Berechnen des Skalarprodukts von zwei Vektoren in V zum Berechnen des Standardskalarprodukts ihrer Koordinatenspalten vereinfacht! Aber: Gibt es immer eine ON-Basis?

Satz 128 : *Es sei* $(w_1,\dots,w_n)$ *eine Basis von* V. *Mit dem folgenden Verfahren („Schmidt'sches Orthonormalisierungsverfahren“) kann eine ON-Basis* $(v_1,\dots,v_n)$ *von* V *berechnet werden:*

- $u_1 := w_1$
- *Für* $2\le j\le n$ *sei*

$$u_j := w_j - \sum_{i=1}^{j-1} (\langle u_i,w_j\rangle/\langle u_i,u_i\rangle)u_i$$

- *Für* $1\le j\le n$ *sei*

$$v_j := \|u_j\|^{-1}u_j\,.$$

Insbesondere: Jeder euklidische Raum hat eine ON-Basis. Für alle j *ist* ${}_{\mathbb{R}}\langle v_1,\dots,v_j\rangle = {}_{\mathbb{R}}\langle w_1,\dots,w_j\rangle$.

Beweis: Nach Definition ist $\langle v_i,v_i\rangle = 1$, $1\le i\le n$. Es genügt also zu zeigen, dass für alle $1\le k<\ell\le n$ die Vektoren u_k und u_ℓ zueinander senkrecht stehen. Das kann einfach nachgerechnet werden:

$$\langle u_k,u_\ell\rangle = \langle u_k, w_\ell - \sum_{i=1}^{\ell-1}(\langle u_i,w_\ell\rangle/\langle u_i,u_i\rangle)u_i\rangle =$$

$$= \langle u_k,w_\ell\rangle - \sum_{i=1}^{\ell-1}(\langle u_i,w_\ell\rangle/\langle u_i,u_i\rangle)\delta_{ki}\langle u_k,u_k\rangle = \langle u_k,w_\ell\rangle - \langle u_k,w_\ell\rangle = 0\ .$$

Nach Satz 124 ist dann $(v_1,\dots,v_n)$ linear unabhängig, wegen $\dim_{\mathbb{R}}(V)=n$ folgt mit Satz 97, dass $(v_1,\dots,v_n)$ eine Basis ist.

Beispiel 129: Es sei V der von $w_1 := (1,0,1)$ und $w_2 := (1,2,3)$ erzeugte Untervektorraum von $\mathbb{R}^3$. Wir betrachten V mit der Einschränkung des Standardskalarproduktes (von $\mathbb{R}^3$) als euklidischen Raum.

Es ist $\|w_1\| = \sqrt{2}$, $\|w_2\| = \sqrt{14}$ und $\langle w_1, w_2\rangle = 4$, also ist (w_1, w_2) keine ON-Basis von V.

Mit den Bezeichnungen von Satz 128 erhalten wir

- $u_1 := w_1$,
- $u_2 := w_2 - (\langle u_1, w_2\rangle / \langle u_1, u_1\rangle) u_1 = (1,2,3) - \frac{4}{2}(1,0,1) = (-1,2,1)$
- $v_1 := \frac{1}{2}\sqrt{2}(1,0,1)$, $v_2 := \frac{1}{6}\sqrt{6}(-1,2,1)$.

(v_1, v_2) ist eine ON-Basis von V.

§5. Der Fußpunkt des Lotes

Sei $\langle -,-\rangle$ ein Skalarprodukt auf einem reellen Vektorraum V.

Definition 130: Es sei U ein endlich-dimensionaler Untervektorraum von V. Dann ist

$$U^\perp := \{v \in V \mid \text{ für alle } u \in U \text{ ist } \langle v, u\rangle = 0\}$$

ein Untervektorraum von V und heißt *das orthogonale Komplement von U in V*.

Satz 131: *Es sei U ein endlich-dimensionaler Untervektorraum von V, $v \in V$ und $(u_1, \ldots, u_n)$ eine ON-Basis von U.*

(1) *Der Vektor*
$$p_U(v) := \sum_{i=1}^{n} \langle u_i, v\rangle u_i \in U$$
hängt nicht von der Wahl der ON-Basis $(u_1, \ldots, u_n)$ ab und heißt Fußpunkt des Lotes von v auf U.

(2) *Jeder Vektor $v \in V$ lässt sich eindeutig als Summe eines Vektors in U und eines Vektors in $U^\perp$ schreiben, und zwar*
$$v = p_U(v) + (v - p_U(v)),$$
wobei $p_U(v) \in U$ und $(v - p_U(v)) \in U^\perp$ ist. Insbesondere ist $U \cap U^\perp = \{0\}$ und $V = U + U^\perp := \{y + y' \mid y \in U, y' \in U^\perp\}$.

(3) *Für $v \in V$ und $y \in U$ mit $p_U(v) \neq y$ ist*
$$\|v - y\| > \|v - p_U(v)\|,$$
das heißt: $p_U(v)$ ist der eindeutig bestimmte Vektor in U, der von v den kleinsten Abstand hat.

Beweis:

(1) Es sei $(w_1,\ldots,w_n)$ eine ON-Basis von U. Es ist zu zeigen, dass $p_U(v) = \sum_{k=1}^n \langle w_k, v\rangle w_k$ ist. (Dann hängt $p_U(v)$ nicht von der Wahl der ON-Basis in U ab). Nach Satz 125 ist

$$u_i = \sum_{j=1}^n \langle w_j, u_i\rangle w_j\,,\ 1 \le i \le n\,.$$

Daher ist

$$p_U(v) = \sum_{i=1}^n \langle u_i, v\rangle u_i = \sum_{i=1}^n \langle \sum_{j=1}^n \langle w_j, u_i\rangle w_j, v\rangle \sum_{k=1}^n \langle w_k, u_i\rangle w_k =$$

$$= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \langle w_j, u_i\rangle \langle w_j, v\rangle \langle w_k, u_i\rangle w_k =$$

$$= \sum_{k=1}^n (\sum_{j=1}^n \langle w_j, v\rangle \sum_{i=1}^n (\langle w_j, u_i\rangle \langle w_k, u_i\rangle)) w_k = (\text{ cf. Satz 127 })$$

$$= \sum_{k=1}^n (\sum_{j=1}^n \langle w_j, v\rangle \langle w_j, w_k\rangle) w_k = \sum_{k=1}^n \langle w_k, v\rangle w_k\,.$$

(2) Für $y \in U$ ist $\langle v, y\rangle = \langle p_U(v), y\rangle$, also

$$\langle v - p_U(v), y\rangle = \langle v, y\rangle - \langle p_U(v), y\rangle = 0\,.$$

Daher ist $v - p_U(v) \in U^\perp$ und $v = p_U(v) + (v - p_U(v)) \in U + U^\perp$. Wenn $y \neq 0$ ist, dann ist $0 < \langle y, y\rangle$, also $y \notin U^\perp$. Daher ist $U \cap U^\perp = \{0\}$.

(3) Für $y \in U$ mit $p_U(v) \neq y$ ist $0 \neq p_U(v) - y \in U$ und $v - p_U(v) \in U^\perp$. Deshalb ist

$$\|v - y\|^2 = \|(v - p_U(v)) + (p_U(v) - y)\|^2 =$$

$$= \|v - p_U(v)\|^2 + \|p_U(v) - y\|^2 > \|v - p_U(v)\|^2\,.$$

Beispiel 132: Es sei $0 \neq u \in V$ und U die Gerade $\mathbb{R}u$. Dann ist $\|u\|^{-1}u$ eine ON-Basis von U. Der Fußpunkt des Lotes von $v \in V$ auf die Gerade U ist

$$p_{\mathbb{R}u}(v) = \langle \|u\|^{-1}u, v\rangle \|u\|^{-1}u = (\langle u, v\rangle / \langle u, u\rangle) u\,.$$

Beispiel 133: Satz 131 ermöglicht eine geometrische Interpretation des Schmidt'schen Orthonormalisierungsverfahrens (Satz 128):
Seien $u_1,\ldots,u_{j-1}$ schon berechnet. Dann ist $(v_1,\ldots,v_{j-1})$ eine Orthonormalbasis des von $v_1,\ldots,v_{j-1}$ erzeugten Untervektorraums V_{j-1}. Der Vektor

$$\sum_{i=1}^{j-1} (\langle u_i, w_j\rangle / \langle u_i, u_i\rangle) u_i$$

ist dann der Fußpunkt des Lotes von w_j auf V_{j-1} und u_j der Fußpunkt des Lotes von w_j auf $V_{j-1}^{\perp}$. Wegen $w_j \notin V_{j-1}$ ist $u_j \neq 0$.

Definition 134 : Es sei Z ein endlichdimensionaler affiner Unterraum von V mit Aufpunkt z und parallelem Untervektorraum U. Der Vektor

$$p_Z(v) := z + p_U(v - z)$$

heißt *Fußpunkt des Lotes* von v auf den affinen Unterraum Z.
Die Zahl $\|v - p_Z(v)\|$ heißt *Abstand des Punktes v vom affinen Unterraum Z*.

Satz 135 : *Es seien Z und Z' endlichdimensionale affine Unterräume von V mit Aufpunkten z, z' und parallelen Untervektorräumen U, U'. Seien $u \in U$, $u' \in U'$ so, dass $-u + u'$ der Fußpunkt des Lotes von $z - z'$ auf $U + U'$ ist. Dann ist $v := z + u \in Z$, $v' := z' + u' \in Z'$ und für alle $w \in Z$, $w' \in Z'$ ist*

$$\|w - w'\| \geq \|v - v'\| \ .$$

Die Zahl $\|v - v'\|$ heißt Abstand der affinen Unterräume Z und Z' *und ist gleich dem Abstand des Punktes $z - z'$ vom Untervektorraum $U + U'$.*

Beweis: Für alle $x \in U$, $x' \in U'$ ist
$\|(z+x) - (z'+x')\| = \|(z-z') - (x'-x)\| \geq \|(z-z') - (-u+u')\| = \|v-v'\|$.

§6. Winkel

Sei V mit $\langle -, - \rangle$ ein euklidischer Raum.

Definition 136 : Es sei V zweidimensional, $w \in V$ und $r \in \mathbb{R}_{\geq 0} := \{t \in \mathbb{R} \mid t \geq 0\}$. Die Menge

$$\{v \in V \mid \|v - w\| = r\}$$

heißt *Kreis mit Mittelpunkt w und Radius r*. Der Kreis mit Mittelpunkt w und Radius 1 heißt *Einheitskreis um u*.

Es seien $u, v, w \in V$ mit $u \neq 0$, $v \neq 0$. Wir möchten die Lage der zwei Halbgeraden $w + \mathbb{R}_{\geq 0}u$ und $w + \mathbb{R}_{\geq 0}v$ zueinander durch eine Zahl beschreiben, welche der „Länge des (kürzeren) Bogens“ zwischen $w + \|u\|^{-1}u$ und $w + \|v\|^{-1}v$ auf dem Einheitskreis um w entsprechen soll.

Diese Zahl soll nur von u und v, aber nicht von w abhängen. Wir nehmen also im weiteren an, dass $w = 0$ ist. Sei

$$u' := \|u\|^{-1}u \quad \text{und} \quad v' := \|v\|^{-1}v \ .$$

Der Fußpunkt des Lotes $p_{\mathbb{R}u}(v)$ von v' auf die Gerade $\mathbb{R}u$ ist dann

$$\langle u', v' \rangle u'$$

(siehe Beispiel 132). Die Zahl $|\langle u', v' \rangle|$ ist der Abstand zwischen 0 und $p_{\mathbb{R}u}(v')$, kann also leichter gemessen werden als die „Länge des (kürzeren) Bogens" zwischen u' und v'. Nach Satz 117 ist $-1 \leq \langle u', v' \rangle \leq 1$.

Wenn $\alpha \in [0, \pi]$ die „Länge des (kürzeren) Bogens" zwischen u' und v' ist, setzen wir

$$\cos(\alpha) := \langle u', v' \rangle = \langle u, v \rangle / (\|u\| \cdot \|v\|)$$

(Sprechweise: Cosinus von α). In der Analysis zeigt man, dass es zu jeder Zahl $z \in [-1, 1]$ genau ein $\alpha \in [0, \pi]$ mit $\cos(\alpha) = z$ gibt. Anders formuliert: die Funktion

$$\cos : [0, \pi] \longrightarrow [-1, 1], \ \alpha \longmapsto \cos(\alpha),$$

ist bijektiv. Für $\alpha \in [0, \pi]$ bezeichnen wir mit $\sin(\alpha)$ (Sprechweise: Sinus von α) den Abstand von v' zur Geraden $\mathbb{R}u$. Aus dem Satz von Pythagoras folgt $\sin(\alpha)^2 + \cos(\alpha)^2 = 1$.

Definition 137 : Es seien $u, v, w \in V$, $u \neq 0$ und $v \neq 0$. Die eindeutig bestimmte Zahl $\alpha \in [0, \pi]$ mit

$$\cos(\alpha) = \langle u, v \rangle / (\|u\| \cdot \|v\|)$$

heißt *Winkel zwischen den Halbgeraden $w + \mathbb{R}_{\geq 0} u$ und $w + \mathbb{R}_{\geq 0} v$* oder kurz *Winkel zwischen u und v*.

Satz 138 : *(„Cosinussatz") Es seien $v, w \in V$ mit $v \neq 0$, $w \neq 0$ und α der Winkel zwischen v und w. Dann ist*

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\| \cdot \|w\| \cdot \cos(\alpha) .$$

Beweis: $\|v - w\|^2 = \langle v - w, v - w \rangle = \|v\|^2 + \|w\|^2 - 2\langle v, w \rangle =$
$= \|v\|^2 + \|w\|^2 - 2\|v\| \cdot \|w\| \cdot \cos(\alpha)$.

KAPITEL 4

Permutationen, Determinanten und Eigenwerte

§1. Hintereinanderausführung von Funktionen

In diesem Abschnitt seien M,N,P und Q Mengen, die nicht leer sind.

Definition 139: Seien $f : M \to N$ und $g : P \to Q$ Funktionen so, dass für alle $m \in M$ die Bilder $f(m)$ Elemente von P sind (das ist zum Beispiel der Fall, wenn $N = P$ ist). Dann heißt die Funktion

$$g \circ f : M \to Q, m \mapsto g(f(m)),$$

die *Hintereinanderausführung* oder *Zusammensetzung* von f und g (sprich „g nach f"). Oft wird statt $g \circ f$ nur gf geschrieben.

Beispiel 140: Die Hintereinanderausführung von

$$f : \mathbb{N} \to \mathbb{N}\,, z \mapsto 2z+1\,,$$

und

$$g : \mathbb{Z} \to \mathbb{Z}\,, z \mapsto 3z-7\,,$$

ist

$$gf : \mathbb{N} \to \mathbb{Z}\,, z \mapsto 6z-4\,.$$

Die Hintereinanderausführung fg ist nicht definiert, weil zum Beispiel $g(0) = -7$ keine natürliche Zahl ist.

Beispiel 141: Die Hintereinanderausführung von

$$f : \mathbb{Q} \to \mathbb{Q}\,, z \mapsto z^2\,,$$

und

$$g : \mathbb{Q} \to \mathbb{Q}\,, z \mapsto z+1\,,$$

ist

$$gf : \mathbb{Q} \to \mathbb{Q}\,, z \mapsto z^2+1\,.$$

Die Hintereinanderausführung von g und f ist

$$fg : \mathbb{Q} \to \mathbb{Q}\,, z \mapsto (z+1)^2\,.$$

Es ist $(fg)(3) = 16$ und $(gf)(3) = 10$, also ist $fg \neq gf$.

Satz 142 : *Seien $f : M \to N$, $g : P \to Q$ und $h : R \to S$ Funktionen so, dass für alle $m \in M$ die Bilder $f(m)$ Elemente von P sind und für alle $p \in P$ die Bilder $g(p)$ Elemente von R sind. Dann gilt*

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

d.h. bei mehrfacher Hintereinanderausführung von Funktionen kommt es nicht auf die Reihenfolge an (die Hintereinanderausführung von Funktionen ist assoziativ*, auf das Setzen von Klammern kann verzichtet werden).*

Beweis: Sowohl $h \circ (g \circ f)$ als auch $(h \circ g) \circ f$ sind Funktionen von M nach S. Für jedes $m \in M$ ist

$$\begin{aligned}(h \circ (g \circ f))(m) &= h((g \circ f)(m)) = h(g(f(m))) = (h \circ g)(f(m)) \\ &= ((h \circ g) \circ f)(m).\end{aligned}$$

Definition 143 : Die Funktion

$$\mathrm{Id}_M : M \to M\,, m \mapsto m\,,$$

heißt die *identische Funktion* oder *Identität* auf M.

Beispiel 144 : Für alle Funktionen $f : M \longrightarrow N$ ist

$$\mathrm{Id}_N \circ f = f = f \circ \mathrm{Id}_M\ .$$

Definition 145 : Sei $f : M \to N$ eine bijektive Funktion. Dann heißt die (ebenfalls bijektive) Funktion

$$f^{-1} : N \to M, n \mapsto \text{Urbild von } n \text{ bezüglich } f,$$

die zu f *inverse* Funktion oder die *Umkehrfunktion* von f.

Beispiel 146 : Die zu

$$f : \{1,2,3\} \longrightarrow \{1,2,3\},\ 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1,$$

inverse Funktion ist

$$f^{-1} : \{1,2,3\} \longrightarrow \{1,2,3\},\ 1 \mapsto 3, 2 \mapsto 1, 3 \mapsto 2.$$

Die zu

$$g : \mathbb{Q} \longrightarrow \mathbb{Q}\,,\ z \longmapsto 3z+1,$$

inverse Funktion ist

$$g^{-1} : \mathbb{Q} \longrightarrow \mathbb{Q}\,,\ z \longmapsto \frac{1}{3}(z-1).$$

Satz 147 : *Sei $f : M \to N$ eine Funktion.*

(1) *Ist f bijektiv, dann gilt $f \circ f^{-1} = \mathrm{Id}_N$ und $f^{-1} \circ f = \mathrm{Id}_M$.*
(2) *Ist $g : N \to M$ eine Funktion mit $f \circ g = \mathrm{Id}_N$ und $g \circ f = \mathrm{Id}_M$, dann ist f bijektiv und $g = f^{-1}$.*

Beweis: Übung.

Satz 148 : *Seien $f : M \to N$ und $g : N \to P$ bijektive Funktionen. Dann ist auch $g \circ f$ bijektiv und es gilt*

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

Beweis: Wegen

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \mathrm{Id}_N \circ g^{-1} = \mathrm{Id}_P$$

und

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \mathrm{Id}_N \circ f = \mathrm{Id}_M$$

folgt die Behauptung aus Satz 147, (2).

Satz 149 : *Es sei*

$$S(M) := \{s \mid s : M \to M \text{ bijektiv}\}.$$

(1) *$S(M)$ mit der Hintereinanderausführung von Funktionen ist eine Gruppe mit neutralem Element* Id_M *und heißt die* symmetrische Gruppe *von M.*
(2) *Wenn M eine endliche Menge mit n Elementen ist, dann ist auch $S(M)$ endlich und hat $n! := n \cdot (n-1) \cdots 2 \cdot 1$ Elemente.*

Beweis:

(1) Folgt aus Satz 142 und Satz 148.
(2) Übung

§2. Translationen

In diesem Abschnitt sei V ein Vektorraum über einem Körper K.

Definition 150 : Es sei $v \in V$. Die Funktion

$$t_v : V \longrightarrow V\,,\; x \longmapsto x+v\,,$$

heißt *Translation* oder *Verschiebung um v* in V.
Sei

$$T(V) := \{t_v \mid v \in V\}$$

die Menge aller Translationen in V.

Satz 151 :

(1) *Jede Translation ist bijektiv, die Umkehrfunktion von t_v ist $t_{(-v)}$. Die Hintereinanderausführung zweier Translationen ist wieder eine Translation, für $v,w \in V$ ist*

$$t_v \circ t_w = t_{v+w} = t_w \circ t_v\,.$$

(2) *Mit*

$$T(V) \times T(V) \longrightarrow T(V)\,,\; (s,t) \longmapsto s \circ t\,,$$

und

$$K \times T(V) \longrightarrow T(V)\,,\; (c,t_v) \longmapsto t_{cv}\,,$$

ist $T(V)$ ein Vektorraum („Translationen sind Vektoren“).

Beweis: nachprüfen.

Definition 152 : Sei M eine Menge. Wir bezeichnen ein Paar $(a,e) \in M \times M$ von Elementen von M als *Pfeil in M mit Anfangspunkt a und Endpunkt e*.

Ein Pfeil enthält mehr Information als eine Menge von zwei Punkten (die durch eine Strecke, die diese Punkte verbindet, dargestellt werden kann). Die „Zusatzinformation“ ist die Reihenfolge der Punkte: es gibt einen ersten („Schaft“ des Pfeils) und einen zweiten Punkt („Spitze“ des Pfeils).

Ein Paar reeller Zahlen kann entweder als Punkt in der Zeichenebene oder als Pfeil in $\mathbb{R}$ zeichnerisch dargestellt werden. Ein Pfeil im Vektorraum $\mathbb{R}^2$ ist ein Element des vierdimensionalen Vektorraumes $\mathbb{R}^2 \times \mathbb{R}^2$. Elemente eines vierdimensionalen Vektorraumes können also zeichnerisch durch Pfeile in der Ebene dargestellt werden.

Sei $0 \neq v \in V$. Der Graph der Translation t_v ist die Menge

$$\{(y, y+v) \mid y \in V\} \subseteq V \times V\ .$$

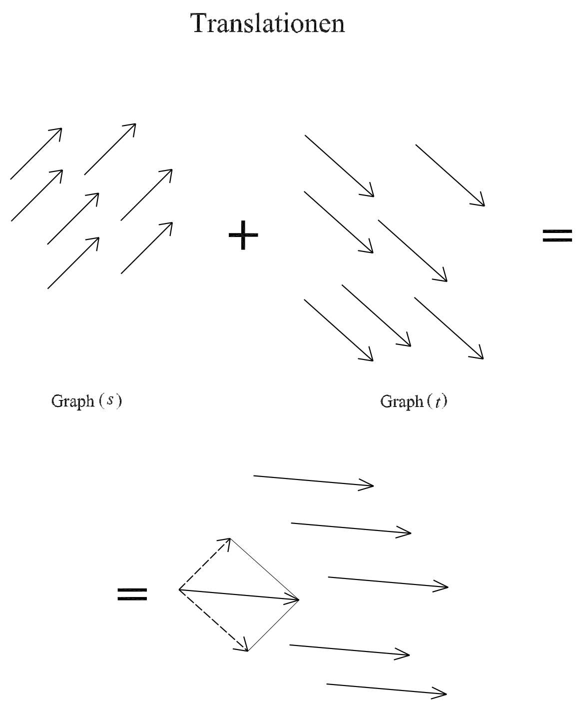


ABBILDUNG 1. Addition von (Graphen von) Translationen

Die Gerade durch die Punkte $x \in V$ und $x+v$ ist $x+\mathbb{R}v$, also sind für alle $y,z \in V$ die Geraden durch y und $y+v$ bzw. durch z und $z+v$ parallel.

§3. Permutationen

Definition 153 : Eine bijektive Funktion von $\{1,2,\dots,n\}$ nach $\{1,2,\dots,n\}$ heißt *Permutation* der Zahlen $1,2,\dots,n$. Wir bezeichnen mit S_n die Menge aller Permutationen der Zahlen $1,2,\dots,n$. S_n zusammen mit der Hintereinanderausführung von Funktionen heißt die *Permutationsgruppe* vom Grad n. Eine Permutation

$$\sigma : \{1,2,\dots,n\} \to \{1,2,\dots,n\}$$

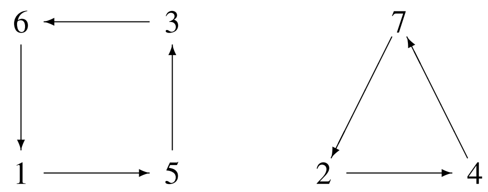
schreibt man oft als $2 \times n$-Matrix

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}.$$

S_n hat genau $n! = n(n-1)\cdots 2\cdot 1$ Elemente. Graphisch kann man eine Permutation darstellen, indem man die Zahlen $1,2,\dots,n$ anschreibt und für $1 \le i \le n$ einen Pfeil von i nach $\sigma(i)$ zeichnet. Zum Beispiel hat

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 4 & 6 & 7 & 3 & 1 & 2 \end{pmatrix}$$

die folgende Darstellung:



Da $\sigma : \{1,2,\dots,n\} \to \{1,2,\dots,n\}$ bijektiv ist, ist jede Zahl Anfangs- und Endpunkt von genau einem Pfeil.

Definition 154 : Sei $\ell \in \mathbb{N}$, $\ell \geq 2$, und seien $j_1,\dots,j_\ell$ paarweise verschiedene Elemente von $\{1,\dots,n\}$. Dann heißt die Funktion

$$\tau : \{1,2,\dots,n\} \to \{1,2,\dots,n\}\,,\, i \mapsto \begin{cases} i & \text{falls } i \notin \{j_1,\dots,j_\ell\}, \\ j_{k+1} & \text{falls } i = j_k \text{ mit } k < \ell, \\ j_1 & \text{falls } i = j_\ell, \end{cases}$$

ein *Zykel der Länge* ℓ und wird mit

$$(j_1, j_2, \dots, j_\ell) \quad \text{oder} \quad (j_1\ j_2\ \dots\ j_\ell)$$

bezeichnet. Ein Zykel der Länge 2 heißt eine *Transposition* oder *Vertauschung* von j_1 und j_2. Zwei Zykel $(i_1,\dots,i_k)$ und $(j_1,\dots,j_\ell)$ heißen *disjunkt*, wenn

$$\{i_1,\dots,i_k\} \cap \{j_1,\dots,j_\ell\} = \emptyset$$

ist.

Für $n = 5$ ist zum Beispiel

$$(2\ 4\ 1) = (1\ 2\ 4) = (4\ 1\ 2) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 1 & 5 \end{pmatrix}.$$

Satz 155 : *Es seien* $(i_1,\dots,i_k)$ *und* $(j_1,\dots,j_\ell)$ *zwei disjunkte Zykel.*
Es gilt:

$$(i_1,\dots,i_k)(j_1,\dots,j_\ell) = (j_1,\dots,j_\ell)(i_1,\dots,i_k)$$

und

$$(j_1, j_2, \dots, j_\ell)^{-1} = (j_\ell, j_{\ell-1}, \dots, j_1)\,,$$

insbesondere ist

$$(j_1, j_2)^{-1} = (j_2, j_1) = (j_1, j_2)\,.$$

Beweis: Der Zykel $(i_1,\dots,i_k)$ vertauscht nur Elemente in $\{i_1,\dots,i_k\}$, der Zykel $(j_1,\dots,j_\ell)$ nur Elemente in $\{j_1,\dots,j_\ell\}$. Daher spielt die Reihenfolge keine Rolle. Man prüft leicht nach, dass $(j_\ell, j_{\ell-1},\dots,j_1)$ die Umkehrabbildung von $(j_1,j_2,\dots,j_\ell)$ ist.

Definition 156 : Sei M eine Menge und $f : M \longrightarrow M$ eine Funktion. Ein Element $x \in M$ heißt ein *Fixpunkt* von f, wenn $f(x) = x$ ist.

Satz 157 : *Jede Permutation $\sigma \in S_n$ ist Produkt paarweise disjunkter Zyklen $\rho_1,\dots,\rho_m$, die eindeutig bis auf die Reihenfolge sind. Die Darstellung*

$$\sigma = \rho_1 \dots \rho_m$$

heißt die Zyklenzerlegung *von σ. Jene Elemente $i \in \{1,2,\dots,n\}$, die nicht in den Zyklen vorkommen, sind die Fixpunkte von σ.*

Die Zyklenzerlegung einer Permutation $\sigma \in S_n$ kann wie folgt berechnet werden:
Man startet mit $J := \{1,\dots,n\}$ und wiederholt die folgende Prozedur, bis J leer ist: Wähle ein $j \in J$ und berechne

$$j, \sigma(j), \sigma(\sigma(j)), \sigma(\sigma(\sigma(j))), \dots$$

solange, bis wieder j kommt. Dann ist entweder j ein Fixpunkt oder $(j, \sigma(j), \dots)$ ein Zykel von σ. Streiche diese Elemente aus der Menge J.

Beweis: Nachprüfen.

Definition 158 : Sei $\sigma \in S_n$ eine Permutation mit p Fixpunkten und m Zyklen. Dann heißt die Zahl

$$\operatorname{sign}(\sigma) := (-1)^{n-p-m}$$

das *Vorzeichen* oder *Signum* von σ.

Beispiel 159 : Die Zyklenzerlegung der Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 8 & 6 & 4 & 3 & 1 & 2 & 7 \end{pmatrix}$$

ist $\sigma = (1\ 5\ 3\ 6)(2\ 8\ 7)$, sie hat einen Fixpunkt und ihr Vorzeichen ist $(-1)^{8-1-2} = -1$.

Beispiel 160 : Sei $\mathrm{Id}_n \in S_n$ die durch $\mathrm{Id}_n(i) = i$, $1 \le i \le n$ definierte Permutation. Da Id_n keine Zyklen und n Fixpunkte hat, ist

$$\operatorname{sign}(\mathrm{Id}_n) = 1\,.$$

Jede Transposition $\tau \in S_n$ hat einen Zykel und $n-2$ Fixpunkte, daher ist

$$\operatorname{sign}(\tau) = (-1)^{n-(n-2)-1} = -1\,.$$

Satz 161 : *Jede Permutation $\sigma \in S_n$ ist Produkt von Transpositionen $\tau_1, \dots, \tau_r \in S_n$, die im Allgemeinen nicht eindeutig bestimmt sind. Es gilt aber*

$$\operatorname{sign}(\sigma) = (-1)^r = \begin{cases} 1 & \textit{falls } r \textit{ gerade ist,} \\ -1 & \textit{falls } r \textit{ ungerade ist.} \end{cases}$$

Eine Permutation σ heißt gerade *bzw.* ungerade*, wenn* $\operatorname{sign}(\sigma) = 1$ *bzw.* -1 *ist. Zum Beispiel ist jede Transposition ungerade.*

Beweis: Sei $\sigma = \rho_1 \dots \rho_m$ die Zyklenzerlegung von σ. Wegen

$$(i_1, \dots, i_k) = (i_1, i_2)(i_2, i_3) \dots (i_{k-1}, i_k)$$

kann man jeden Zykel ρ_i als Produkt von Transpositionen schreiben und daher auch σ. Das Beispiel

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (12)(23) = (13)(12)$$

zeigt, dass diese Darstellung im Allgemeinen nicht eindeutig ist.

Die Aussage $\operatorname{sign}(\sigma) = (-1)^r$ beweisen wir durch Induktion nach r. Für $r = 0$ ist $\sigma = \mathrm{Id}_n$ und $\operatorname{sign}(\sigma) = 1$. Für $r = 1$ ist σ eine Transposition, also ist $\operatorname{sign}(\sigma) = -1$.

Sei nun $r \geq 2$, $\tau_1 = (i, j)$ und $\rho := \tau_2 \dots \tau_r$. Dann ist zu zeigen, dass

$$\operatorname{sign}((i, j)\rho) = -\operatorname{sign}(\rho)$$

ist. Dazu untersuchen wir, wie sich die Transposition (i, j) auf die Zyklenzerlegung von ρ auswirkt, und unterscheiden 2 Fälle:

(1) Die Elemente i und j liegen im gleichen Zykel $(j_1, \dots, j_\ell)$ von ρ, wobei wir dann $j_1 = i$ und $j_k = j$ mit $2 \leq k \leq \ell$ annehmen können, d.h.

$$(j_1, \dots, j_\ell) = (i = j_1, \dots, j_{k-1}, j = j_k, \dots, j_\ell)\,.$$

Für $k = 2$ ist dann i ein Fixpunkt von $(i, j)\rho$, und
für $k \geq 3$ ist $(i, j_2, \dots, j_{k-1})$ ein Zykel von $(i, j)\rho$.
Für $k < \ell$ ist $(j, j_{k+1}, \dots, j_\ell)$ ein Zykel von $(i, j)\rho$, und
für $k = \ell$ ist j ein Fixpunkt von $(i, j)\rho$.
Ansonsten bleibt bei der Zyklenzerlegung von ρ alles gleich, sodass die Summe aus der Zahl der Fixpunkte und der Zahl der Zykel insgesamt um 1 steigt und sich damit das Vorzeichen ändert.

(2) Im anderen Fall ist i oder j Fixpunkt von ρ oder i, j liegen in verschiedenen Zyklen $(i_1, i_2, \dots, i_k)$ bzw. $(j_1, j_2, \dots, j_\ell)$ von ρ, wobei wir $i_1 = i$ bzw. $j_1 = j$ annehmen können. Dann ist

$$(i = i_1, \dots, i_k, j = j_1, \dots, j_\ell)$$

ein Zykel von $(i,j)\rho$, sodass insgesamt die Summe aus der Zahl der Fixpunkte und der Zahl der Zykel um 1 fällt und wiederum sich das Vorzeichen ändert.

Satz 162: *Für Permutationen $\sigma,\tau \in S_n$ gilt*

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma)\cdot\operatorname{sign}(\tau)\,.$$

Insbesondere ist $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$.

Beweis: Nach Satz 161 gibt es Transpositionen $\sigma_1,\ldots,\sigma_r$ und $\tau_1,\ldots,\tau_s$ mit $\sigma = \sigma_1\ldots\sigma_r$ bzw. $\tau = \tau_1\ldots\tau_s$. Dann ist $\sigma\tau = \sigma_1\ldots\sigma_r\tau_1\ldots\tau_s$, also $\operatorname{sign}(\sigma\tau) = (-1)^{r+s} = (-1)^r(-1)^s = \operatorname{sign}(\sigma)\operatorname{sign}(\tau)$. Wegen $\sigma^{-1}\cdot\sigma = \mathrm{Id}_n$ ist

$$\operatorname{sign}(\sigma^{-1})\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1}\cdot\sigma) = \operatorname{sign}(\mathrm{Id}_n) = 1\ ,$$

also $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$.

§4. Polynomfunktionen

In diesem Abschnitt sei K ein Körper.

Definition 163: Seien $n \in \mathbb{N}$ und $a_0,a_1,\ldots,a_n \in K$. Dann heißt die Funktion

$$f : K \to K\,, z \mapsto a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n = \sum_{i=0}^{n} a_i z^i\,,$$

eine *Polynomfunktion* von K nach K. Die Elemente $a_0,\ldots,a_n$ heißen *Koeffizienten von f*.

Mit Polynomfunktionen sind mehrere grundlegende Aufgaben verbunden:

- *Auswerten* einer Polynomfunktion f mit Koeffizienten $a_0,\ldots,a_n$ in einem Element c von K: Berechne das Bild

$$f(c) = \sum_{i=0}^{n} a_i c^i$$

  von c unter der Polynomfunktion f. Es ist klar, dass dieses Element von K immer durch Ausführen von Additionen und Multiplikationen in K berechnet werden kann. Darin liegt die „rechnerische Bedeutung“ der Polynomfunktionen.
- *Interpolation* durch eine Polynomfunktion: Gegeben sind eine endliche Teilmenge E von K und eine Funktion $g : E \longrightarrow K$. Gesucht ist eine Polynomfunktion f von K nach K so, dass für alle $z \in E$ gilt: $f(z) = g(z)$.

- *Überprüfen der Gleichheit* von zwei Polynomfunktionen: Zwei Polynomfunktionen seien durch ihre Koeffizienten gegeben. Wie kann man feststellen, ob diese zwei Funktionen gleich sind? Die Antwort ist nicht so leicht: zum Beispiel sind die Polynomfunktionen

$$f : \mathbb{Z}_2 \longrightarrow \mathbb{Z}_2\,,\; z \longmapsto z\,,\text{und } g : \mathbb{Z}_2 \longrightarrow \mathbb{Z}_2\,,\; z \longmapsto z^2\,,$$

  gleich.
- *Berechnen der Nullstellen* einer Polynomfunktion f: Finde alle Elemente $c \in K$ mit der Eigenschaft, dass $f(c) = 0$ ist. Einfacher zu beantwortende Fragen sind: Gibt es solche Elemente? Wenn ja, wieviele?

Satz 164 : *Es sei $c \in K$ und f eine Polynomfunktion mit Koeffizienten $a_0,\dots,a_n \in K$. Mit dem folgenden Verfahren kann $f(c)$ mit höchstens n Additionen und höchstens n Multiplikationen in K berechnet werden:*

- *Setze $i := n$ und $w := a_n$.*
- *Solange $i \neq 0$ ist, ersetze i durch $i-1$ und dann w durch $w \cdot c + a_i$.*
- *Wenn $i = 0$ ist, dann ist $f(c) = w$.*

Beweis: $\sum_{i=0}^{n} a_i c^i = (\dots((a_n c + a_{n-1})c + a_{n-2})c + \dots + a_1)c + a_0$.

Satz 165 : *Sei M eine Menge und W ein Vektorraum über K. Dann ist die Menge*

$$\mathcal{F}(M,W) := \{f \mid f : M \to W\}$$

aller Funktionen von M nach W mit der punktweisen Addition

$$(f+g)(m) := f(m) + g(m) \quad \textit{für alle } m \in M$$

und der punktweisen Skalarmultiplikation

$$(cf)(m) := cf(m) \quad \textit{für alle } m \in M$$

ein Vektorraum über K und heißt Vektorraum der W-wertigen Funktionen auf M.

Für $M = \mathbb{N}$ erhält man als Spezialfall den Vektorraum der Folgen in W

$$\mathcal{F}(\mathbb{N},W) = \{(a_n)_{n\in\mathbb{N}} \mid \textit{für alle } n \in \mathbb{N} \textit{ ist } a_n \in W\}$$

mit der komponentenweisen Addition

$$(a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} := (a_n + b_n)_{n\in\mathbb{N}}$$

und der komponentenweisen Skalarmultiplikation

$$c(a_n)_{n\in\mathbb{N}} := (ca_n)_{n\in\mathbb{N}}\,.$$

Für $M = \{1,2,\dots,n\}$, wobei $n \in \mathbb{N}$ ist, erhält man als Spezialfall den Standard-Vektorraum von Satz 67.

Beweis: Übung.

Satz 166 : *Die Menge der Polynomfunktionen von K nach K ist ein Untervektorraum des K-Vektorraums $\mathcal{F}(K,K)$ aller Funktionen von K nach K.*

Beweis: Seien f und g Polynomfunktionen und $a_0,\ldots,a_n$ bzw. $b_0,\ldots,b_m$ ihre Koeffizienten. Ohne Einschränkung der Allgemeinheit sei $n \geq m$. Falls $n > m$ ist, setzen wir $b_{m+1} := 0,\ldots,b_n = 0$. Für alle $z \in K$ ist dann $f(z) = \sum_{i=0}^n a_i z^i$ und $g(z) = \sum_{i=1}^n b_i z^i$. Daher ist für alle $z \in K$

$$(f+g)(z) = f(z)+g(z) = \left(\sum_{i=0}^n a_i z^i\right) + \left(\sum_{i=0}^n b_i z^i\right) = \sum_{i=0}^n (a_i+b_i) z^i ,$$

also $f+g$ eine Polynomfunktion (mit den Koeffizienten $a_0+b_0,\ldots,a_n+b_n$). Analog zeigt man, dass cf (für $c \in K$) eine Polynomfunktion ist.

Satz 167 : *Es sei M eine Menge und $F := \mathcal{F}(M,K)$ die Menge aller Funktionen von M in den Körper K. Für $f,g \in F$ und $c \in K$ und $m \in M$ sei*

$$(fg)(m) := f(m)g(m) .$$

Mit den Rechenoperationen

$$F \times F \longrightarrow F \ , \ (f,g) \longmapsto f+g \ ,$$

und

$$F \times F \longrightarrow F \ , \ (f,g) \longmapsto fg \ ,$$

*(*punktweise Multiplikation*) ist F ein kommutativer Ring.*
Das Nullelement ist $0 : M \longrightarrow K$, $m \longmapsto 0_K$, *das Einselement ist* $1 : M \longrightarrow K$, $m \longmapsto 1_K$.

Beweis: Übung.

Definition 168 : Es sei R ein Ring und S eine nichtleere Teilmenge von R. Dann ist S ein *Unterring* von R, wenn $1 \in S$ und für alle $a,b \in S$ auch die Elemente

$$a+b, \ \ -a \ \text{ und } \ ab$$

in S enthalten sind.

Ein Unterring ist mit den auf diese Teilmenge eingeschränkten Rechenoperationen selbst ein Ring.

Satz 169 : *Die Menge der Polynomfunktionen von K nach K ist ein Unterring des Ringes $\mathcal{F}(K,K)$ aller Funktionen von K nach K.*

Beweis: Seien f und g Polynomfunktionen und $a_0,\dots,a_n$ bzw. $b_0,\dots,b_m$ ihre Koeffizienten. Für alle $z \in K$ ist dann $f(z) = \sum_{i=0}^n a_i z^i$ und $g(z) = \sum_{j=1}^m b_j z^j$. Daher ist für alle $z \in K$

$$(f\cdot g)(z) = f(z)g(z) = \left(\sum_{i=0}^n a_i z^i\right)\left(\sum_{j=0}^m b_j z^j\right) =$$

$$= \sum_{i=0}^n \sum_{j=0}^m a_i b_j z^{i+j} = \sum_{k=0}^{n+m}\left(\sum_{i=0}^k a_i b_{k-i}\right) z^k ,$$

also $f\cdot g$ eine Polynomfunktion. Die anderen Eigenschaften eines Unterringes folgen aus Satz 166.

§5. Determinanten

In diesem Abschnitt sei K ein Körper und n eine positive ganze Zahl.

Definition 170 : Für eine Matrix $A \in K^{n\times n}$ heißt

$$\det(A) := \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)1}A_{\sigma(2)2}\dots A_{\sigma(n)n} \in K$$

die *Determinante* von A.

Beispiel 171 : Im Fall $n = 1$ ist $S_n = \{\mathrm{Id}_1\}$ und $\det(A) = A_{11}$.
Im Fall $n = 2$ ist $S_n = \{\mathrm{Id}_2, (1,2)\}$ und

$$\det\begin{pmatrix} A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix} = A_{11}A_{22} - A_{12}A_{21}.$$

Im Fall $n = 3$ ist $S_n = \{\mathrm{Id}_3, (1,2,3), (1,3,2), (1,3), (2,3), (1,2)\}$ und

$$\det\begin{pmatrix} A_{11} & A_{12} & A_{13}\\ A_{21} & A_{22} & A_{23}\\ A_{31} & A_{32} & A_{33}\end{pmatrix} = A_{11}A_{22}A_{33} + A_{21}A_{32}A_{13} + A_{31}A_{12}A_{23} -$$

$$- A_{31}A_{22}A_{13} - A_{11}A_{32}A_{23} - A_{21}A_{12}A_{33}.$$

Für $n \geq 4$ hat S_n mindestens $4! = 24$ Elemente und die Berechnung der Determinante nach Definition ist zu aufwändig. Wir suchen daher ein Verfahren, mit dem man die Determinante „schnell“ berechnen kann.

Definition 172 : Seien m eine positive ganze Zahl und $A \in K^{m\times n}$. Dann heißt

$$A^T := (A_{ji})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \in K^{n\times m}$$

die *transponierte Matrix* von A.

Beispiel 173 :

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}$$

$$(3\ 2)^T = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Satz 174 : *Für $A \in K^{n\times n}$ ist*

$$\det(A^T) = \det(A)\,,$$

d.h. Transponieren ändert die Determinante nicht.

Beweis: Nach Definition ist

$$\det(A^T) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{j=1}^n (A^T)_{\sigma(j)j} = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{j=1}^n A_{j\sigma(j)}$$
$$= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^n A_{\sigma^{-1}(i)i},$$

wobei im Produkt $i := \sigma(j)$ gesetzt und umgeordnet wurde. Da die Funktion $S_n \to S_n\,,\ \sigma \mapsto \sigma^{-1}$, bijektiv ist, kann man $\tau := \sigma^{-1}$ setzen und erhält wegen $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$

$$\det(A^T) = \sum_{\tau\in S_n} \operatorname{sign}(\tau) \prod_{i=1}^n A_{\tau(i)i} = \det(A)\,.$$

Definition 175 : Sei $A \in K^{n\times n}$.

(1) A hat *obere Dreiecksform*, wenn $A_{ij} = 0$ für alle Indizes i, j mit $i > j$ ist. Dann hat A die Gestalt

$$\begin{pmatrix} * & * & \dots & * \\ 0 & * & & * \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & * \end{pmatrix},$$

wobei $*$ für beliebige Elemente von K steht.

(2) A hat *untere Dreiecksform*, wenn $A_{ij} = 0$ für alle Indizes i, j mit $i < j$ ist. Dann hat A die Gestalt

$$\begin{pmatrix} * & 0 & \dots & 0 \\ * & * & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ * & * & \dots & * \end{pmatrix}.$$

(3) A ist eine *Dreiecksmatrix*, wenn A obere oder untere Dreiecksform hat.
(4) A st eine *Diagonalmatrix*, wenn $A_{ij} = 0$ für alle Indizes i, j mit $i \neq j$ ist. Dann hat A hat die Gestalt

$$\begin{pmatrix} * & 0 & \dots & 0 \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & * \end{pmatrix}.$$

Beispiel 176 : Jede quadratische Matrix in Stufenform hat obere Dreiecksform. Daher kann jede quadratische Matrix durch elementare Umformungen in obere Dreiecksform übergeführt werden.

Satz 177 : *Sei $\sigma \in S_n$ mit $\sigma \neq \mathrm{Id}_n$. Dann gibt es eine Zahl $k \in \{1, 2, \dots, n\}$ mit $\sigma(k) > k$.*

Beweis: Wenn es keine derartige Zahl k gibt, dann ist $\sigma(i) \leq i$ für alle $i \in \{1, 2, \dots, n\}$, insbesondere $\sigma(1) = 1$, $\sigma(2) = 2$ usw., also $\sigma = \mathrm{Id}_n$ im Widerspruch zur Voraussetzung.

Satz 178 : *Die Determinante einer Dreiecksmatrix $A \in K^{n \times n}$ ist das Produkt ihrer Diagonalelemente, also*

$$\det(A) = A_{11}A_{22}\dots A_{nn}\,.$$

Insbesondere ist

$$\det(I_n) = 1\,.$$

Beweis: Sei A eine obere Dreiecksmatrix. Nach Satz 177 gibt zu jeder Permutation $\sigma \in S_n$, $\sigma \neq \mathrm{Id}_n$, eine Zahl $k \in \{1, 2, \dots, n\}$ mit $\sigma(k) > k$, also $A_{\sigma(k)k} = 0$. Daher verschwinden in

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{j=1}^{n} A_{\sigma(j)j}$$

alle Summanden bis auf den Summanden zur Identität

$$A_{11}A_{22}\dots A_{nn}\,.$$

Wenn B eine untere Dreiecksmatrix ist, dann ist B^T eine obere Dreiecksmatrix und nach Satz 174 gilt

$$\det(B) = \det(B^T) = (B^T)_{11} \dots (B^T)_{nn} = B_{11} \dots B_{nn}\,.$$

Satz 179 : *Für $\tau \in S_n$ und $A \in K^{n\times n}$ sei $\tau \cdot A$ die durch*

$$(\tau \cdot A)_{ij} := A_{i\tau(j)},\ 1 \le i, j \le n\,,$$

definierte Matrix in $K^{n\times n}$. („Die Matrix $\tau \cdot A$ erhält man aus A, indem man die Spalten von A mit τ^{-1} permutiert".) Dann ist

$$\det(\tau \cdot A) = \operatorname{sign}(\tau) \cdot \det(A)\,.$$

Beweis: Es ist

$$\det(\tau \cdot A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)\tau(1)} A_{\sigma(2)\tau(2)} \cdots A_{\sigma(n)\tau(n)}\,.$$

Wir setzen $\rho := \sigma\tau^{-1}$, dann ist $\sigma = \rho\tau$ und nach Satz 162 $\operatorname{sign}(\rho\tau) = \operatorname{sign}(\rho)\operatorname{sign}(\tau)$. Daher ist

$$\sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)\tau(1)} A_{\sigma(2)\tau(2)} \cdots A_{\sigma(n)\tau(n)} = \sum_{\rho \in S_n} \operatorname{sign}(\rho\tau) \prod_{i=1}^{n} A_{\rho(\tau(i))\tau(i)} =$$

$$= \operatorname{sign}(\tau) \sum_{\rho \in S_n} \operatorname{sign}(\rho) \prod_{j=1}^{n} A_{\rho(j)j} = \operatorname{sign}(\tau) \cdot \det(A)\,.$$

Satz 180 : *Wenn zwei Zeilen oder zwei Spalten einer Matrix gleich sind, dann ist ihre Determinante gleich* 0.

Beweis: Falls $1_K + 1_K \neq 0$ ist, folgt die Behauptung direkt aus Satz 179: Sei τ die Vertauschung der Indizes der zwei gleichen Zeilen. Dann ist $-\det(A) = \det(\tau \cdot A) = \det(A)$, daher muss $\det(A) = 0$ sein.

Ohne diese Voraussetzung wird der Beweis aufwendiger: Nach Satz 174 genügt es, die Aussage für Spalten zu beweisen. Sei $A \in K^{n\times n}$ eine Matrix und $1 \le i, j \le n$ so, dass $i \neq j$ und $A_{-i} = A_{-j}$ ist.

Sei $\tau := (i, j) \in S_n$ die Vertauschung von i und j. Nach Satz 162 und Satz 161 ist die Funktion

$$\{\rho \in S_n \mid \operatorname{sign}(\rho) = 1\} \to \{\sigma \in S_n \mid \operatorname{sign}(\sigma) = -1\}\,,\ \rho \mapsto \rho\tau\,,$$

wohldefiniert und bijektiv, ihre Umkehrfunktion ist

$$\{\sigma \in S_n \mid \operatorname{sign}(\sigma) = -1\} \to \{\rho \in S_n \mid \operatorname{sign}(\rho) = 1\}\,,\ \sigma \mapsto \sigma\tau\,.$$

Daher ist

$$\det(A) = \sum_{\sigma \in S_n, \operatorname{sign}(\sigma)=1} \operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(n)n} +$$

$$+ \sum_{\sigma \in S_n, \operatorname{sign}(\sigma)=1} \operatorname{sign}(\sigma\tau) A_{\sigma(\tau(1))1} A_{\sigma(\tau(2))2} \cdots A_{\sigma(\tau(n))n} =$$

$$= \sum_{\sigma \in S_n, \operatorname{sign}(\sigma)=1} (\operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(n)n} -$$

$$-\operatorname{sign}(\sigma)A_{\sigma(\tau(1))1}A_{\sigma(\tau(2))2}\dots A_{\sigma(\tau(n))n}) = 0\,,$$

weil nach Voraussetzung $A_{\sigma(\tau(1))1}\dots A_{\sigma(\tau(n))n} = A_{\sigma(1)1}\dots A_{\sigma(n)n}$ ist.

Satz 181 : *Seien A und B Matrizen in $K^{n\times n}$.*

(1) $\det(A\cdot B) = \det(A)\cdot\det(B) = \det(B\cdot A)$.
„Die Determinante des Produktes ist das Produkt der Determinanten."

(2) *Eine Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn ihre Determinante nicht* 0 *ist. In diesem Fall ist*

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

„Die Determinante der inversen Matrix ist zur Determinante der Matrix invers."

Beweis:

(1) Es ist

$$\det(AB) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(AB)_{\sigma(1)1}(AB)_{\sigma(2)2}\dots(AB)_{\sigma(n)n} =$$

$$= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)(\sum_{k_1=1}^{n} A_{\sigma(1)k_1}B_{k_1 1})\dots(\sum_{k_n=1}^{n} A_{\sigma(n)k_n}B_{k_n n}) =$$

$$= \sum_{k_1=1}^{n}\dots\sum_{k_n=1}^{n}(\sum_{\sigma\in S_n} \operatorname{sign}(\sigma)A_{\sigma(1)k_1}A_{\sigma(2)k_2}\dots A_{\sigma(n)k_n})B_{k_1 1}B_{k_2 2}\dots B_{k_n n}\,.$$

Seien $1 \le k_1,\dots,k_n \le n$ und

$$\tau : \{1,\dots,n\} \longrightarrow \{1,\dots,n\}\,,\ i \longmapsto k_i\,.$$

Es ist

$$\sum_{\sigma\in S_n} \operatorname{sign}(\sigma)A_{\sigma(1)k_1}A_{\sigma(2)k_2}\dots A_{\sigma(n)k_n}$$

die Determinante der Matrix, deren i-te Spalte die k_i-te Spalte von A ist, $1 \le i \le n$. Wir bezeichnen diese Matrix mit $A(k_1,\dots,k_n)$.
Seien $1 \le k_1,\dots,k_n \le n$ und

$$\tau : \{1,\dots,n\} \longrightarrow \{1,\dots,n\}\,,\ i \longmapsto k_i\,.$$

Wenn τ bijektiv, also eine Permutation ist, dann ist

$$\det(A(k_1,\dots,k_n)) = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma)A_{\sigma(1)\tau(1)}\dots A_{\sigma(n)\tau(n)} = \det(\tau\cdot A)\,,$$

also nach Satz 179 gleich $\operatorname{sign}(\tau)\cdot\det(A)$. Wenn τ nicht bijektiv ist, dann ist

$$\det(A(k_1,\dots,k_n)) = 0 \ \text{(nach Satz 180).}$$

Somit ist

$$\det(AB) = \sum_{k_1=1}^{n} \dots \sum_{k_n=1}^{n} \det(A(k_1,\dots,k_n))B_{k_1 1}B_{k_2 2}\dots B_{k_n n} =$$

$$= \sum_{\tau\in S_n} \operatorname{sign}(\tau)\det(A)B_{\tau(1)1}\dots B_{\tau(n)n} = \det(A)\cdot\det(B)\,.$$

(2) Wenn die Matrix A invertierbar ist, dann gibt es eine Matrix $A^{-1} \in K^{n\times n}$ mit $A\cdot A^{-1} = I_n$. Nach (1) ist dann

$$\det(A)\cdot\det(A^{-1}) = \det(I_n) = 1\,,$$

also kann die Zahl $\det(A)$ nicht 0 sein und $\det(A^{-1}) = \det(A)^{-1}$.
Wenn die Matrix A nicht invertierbar ist, dann gibt es nach Satz 89 eine invertierbare Matrix P so, dass die letzte Zeile von $P\cdot A$ die Nullzeile ist. Aus der Definition der Determinante folgt, dass dann $\det(P\cdot A) = 0$ ist. Da $\det(P)$ nicht 0 ist, folgt aus (1), dass $\det(A) = 0$ ist.

Da man Determinanten von Dreiecksmatrizen leicht ausrechnen kann, liegt die Frage nahe, wie sich die Determinante bei elementaren Zeilenumformungen der Matrix ändert.

Satz 182: *Sei $A \in K^{n\times n}$.*

(1) *Sei B die Matrix, die man erhält, indem man eine Zeile von A mit einem Element $c \in K$ multipliziert. Dann ist* $\det(B) = c\cdot\det(A)$. *Insbesondere ist* $\det(c\cdot A) = c^n\cdot\det(A)$.
(2) *Sei B die Matrix, die man erhält, indem man zwei Zeilen von A vertauscht. Dann ist* $\det(B) = -\det(A)$.
(3) *Sei B die Matrix, die man erhält, indem man ein skalares Vielfaches einer Zeile von A zu einer anderen Zeile von A addiert. Dann ist* $\det(B) = \det(A)$.
(4) *Die Aussagen (1) - (3) gelten analog für Spalten statt Zeilen.*
(5) *Die Determinante von A kann wie folgt berechnet werden: Forme A durch elementare Zeilenumformungen (oder Spaltenumformungen) vom Typ 1 (Addition eines Vielfachen einer Zeile zu einer anderen) und vom Typ 2 (Vertauschung zweier Zeilen) in eine Matrix B in Dreiecksform um. Sei k die Zahl der ausgeführten Zeilenvertauschungen. Dann ist*

$$\det(A) = (-1)^k B_{11}B_{22}\dots B_{nn}\,.$$

Beweis:

(1) folgt direkt aus der Definition.
(2) ist ein Spezialfall von Satz 179.

(3) Die Matrix B ist das Produkt von A mit einer Elementarmatrix vom Typ 1. Diese ist eine Dreiecksmatrix, deren Determinante 1 ist. Die Behauptung folgt daher aus Satz 181.
(4) folgt aus Satz 174.
(5) folgt aus (2) und (3).

§6. Orientierung, Volumen und Vektorprodukt

In diesem Abschnitt sei V ein reeller Vektorraum.

Definition 183: Es seien $(v_1,\ldots,v_n)$ und $(w_1,\ldots,w_n)$ Basen von V. Die Matrix $T \in K^{n\times n}$, deren Spalten $T_{-1},\ldots,T_{-n}$ die Koordinatenspalten von $w_1,\ldots,w_n$ bezüglich der Basis $(v_1,\ldots,v_n)$ sind, heißt *Transformationsmatrix* von $(v_1,\ldots,v_n)$ nach $(w_1,\ldots,w_n)$.
Die Basen $(v_1,\ldots,v_n)$ und $(w_1,\ldots,w_n)$ heißen *gleich orientiert*, wenn $\det(T) > 0$ ist, und *verschieden orientiert*, wenn $\det(T) < 0$ ist.
Wählt man eine Basis $(v_1,\ldots,v_n)$ von V aus, dann wird die Menge aller Basen in zwei disjunkte Teilmengen zerlegt: die Teilmenge aller gleich wie $(v_1,\ldots,v_n)$ orientierten Basen und die Teilmenge aller anderen Basen. Diese zwei Mengen heißen *Orientierungen* von V.
Durch die Wahl einer Basis von V wird eine Orientierung festgelegt. V zusammen mit einer Orientierung heißt *orientierter Vektorraum*. Die Basen in der gegebenen Orientierung heißen dann *positiv orientiert*, die anderen *negativ orientiert*.

Wird die Zeichenebene bzw. der physikalische Raum als reeller Vektorraum betrachtet, dann nennt man seine zwei Orientierungen „Orientierung im Uhrzeigersinn“ und „Orientierung gegen den Uhrzeigersinn“ bzw. „Orientierung nach der Linken-Hand-Regel“ und „Orientierung nach der Rechten-Hand-Regel“.

Beispiel 184: Die Standardbasis $(e_1,\ldots,e_{n-1},e_n)$ von $\mathbb{R}^n$ und die Basis $(e_1,\ldots,e_{n-1},-e_n)$ sind verschieden orientiert. Jede Basis von $\mathbb{R}^n$ ist also gleich orientiert wie genau eine dieser zwei Basen.

Sei V mit $\langle -,- \rangle$ ein n-dimensionaler euklidischer Raum.

Definition 185: Seien $w_1,\ldots,w_n \in V$. Die Menge

$$P(w_1,\ldots,w_n) := \{c_1w_1+\ldots+c_nw_n \mid 0 \le c_i \le 1,\ c_i \in \mathbb{R}\}$$

heißt das *von $w_1,\ldots,w_n$ erzeugte Parallelotop*. Wenn $n = 2$ ist, dann heißt ein Parallelotop *Parallelogramm*.
Es sei $(v_1,\ldots,v_n)$ eine ON-Basis von V und $S \in \mathbb{R}^{n\times n}$ die Matrix, deren i-te

Spalte die Koordinatenspalte von $w_i \in V$ bezüglich $(v_1,\ldots,v_n)$ ist, $1 \leq i \leq n$. Die Zahl

$$vol(P(w_1,\ldots,w_n)) := |\det(S)|$$

heißt das *Volumen* von $P(w_1,\ldots,w_n)$.

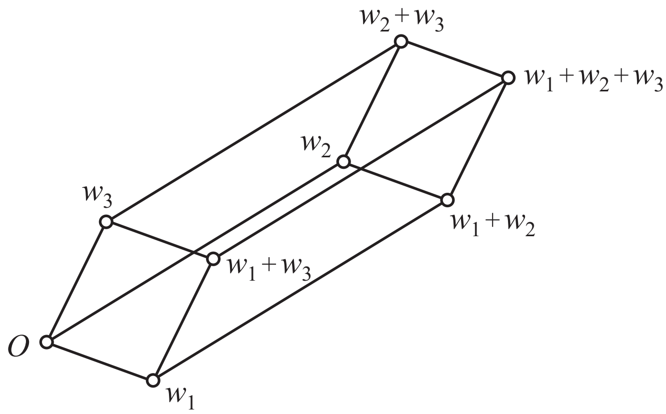


ABBILDUNG 2. Parallelotop $P(w_1,w_2,w_3)$

Satz 186 : *Es seien $(v_1,\ldots,v_n)$ eine ON-Basis von V und $w_1,\ldots,w_n$ Vektoren in V. Dann ist*

$$vol(P(w_1,\ldots,w_n)) = \sqrt{\det((\langle w_i,w_j\rangle)_{1\leq i,j\leq n})}\,.$$

Insbesondere: Das Volumen eines Parallelotops hängt nicht von der Wahl der ON-Basis $(v_1,\ldots,v_n)$ ab. Wenn $(w_1,\ldots,w_n)$ eine ON-Basis von V ist, dann ist $vol(P(w_1,\ldots,w_n)) = 1$.

Beweis: Sei S die Matrix, deren Spalten die Koordinatenspalten von $w_1,\ldots,w_n$ bezüglich $(v_1,\ldots,v_n)$ sind. Für $1 \leq i,j \leq n$ gilt nach Satz 127

$$\langle w_i,w_j\rangle = \langle S_{-i},S_{-j}\rangle = (S^T\cdot S)_{ij}\ .$$

Daher ist

$$\det((\langle w_i,w_j\rangle)_{1\leq i,j\leq n}) = \det(S^T\cdot S) = \det(S)^2 = vol(P(w_1,\ldots,w_n))^2\ .$$

Satz 187 : *Es seien $u,w \in V$, $u \neq 0$, $w \neq 0$ und α der Winkel zwischen u und w. Dann ist*

$$vol(P(u,w)) = \|u\|\cdot\|w\|\cdot\sin(\alpha)\ .$$

Beweis: Nach Satz 186 ist

$$vol(P(u,w))^2 = \det\begin{pmatrix} \langle u,u\rangle & \langle u,w\rangle \\ \langle u,w\rangle & \langle w,w\rangle \end{pmatrix} = \|u\|^2\cdot\|w\|^2 - \langle u,w\rangle^2 =$$

$$= \|u\|^2\cdot\|w\|^2\cdot(1-\cos^2(\alpha)) = (\|u\|\cdot\|w\|\cdot\sin(\alpha))^2 \ .$$

Definition 188 : V sei ein dreidimensionaler orientierter euklidischer Raum. Für $u,w\in V$ sei $u\times w$ der eindeutig bestimmte Vektor in V mit den drei Eigenschaften

- $\|u\times w\| = vol(P(u,w))$,
- $u\times w$ und u stehen zueinander senkrecht, $u\times w$ und w stehen zueinander senkrecht,
- wenn $u\times w\neq 0$ ist, dann ist $(u,w,u\times w)$ eine positiv orientierte Basis von V.

Dieser Vektor heißt das *Vektorprodukt (oder Kreuzprodukt) von u und w*. Sprechweise: „u Kreuz w“.

Satz 189 : *V sei ein dreidimensionaler euklidischer Raum, der durch eine ON-Basis (v_1,v_2,v_3) orientiert ist.*

(1) $v_1\times v_2 = v_3\,,\ v_1\times v_3 = -v_2\,,\ v_2\times v_3 = v_1\,.$
(2) *Ist (u,w) eine ON-Basis des von u und w erzeugten Untervektorraums, dann ist $(u,w,u\times w)$ eine wie (v_1,v_2,v_3) orientierte ON-Basis von V.*
(3) *Für $a_i,b_i\in\mathbb{R},\ 1\le i\le 3$, ist*

$$(\sum_{i=1}^{3} a_iv_i)\times(\sum_{i=1}^{3} b_iv_i) =$$

$$= \det\begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix} v_1 - \det\begin{pmatrix} a_1 & a_3 \\ b_1 & b_3 \end{pmatrix} v_2 + \det\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} v_3 \ .$$

(4) *Für $u,u',w,w'\in V,\ c,d\in\mathbb{R}$ ist*

$$c(u+u')\times w = c(u\times w)+c(u'\times w)\,,$$
$$u\times(d(w+w')) = d(u\times w)+d(u\times w')$$
$$\text{und}\ \ u\times w = -w\times u\ .$$

(5) *Für $x,y,z\in V$ ist $x\times(y\times z) = \langle x,z\rangle y - \langle x,y\rangle z\,.$*
(6) *Für $x,y,z\in V$ ist*

$$(x\times y)\times z + (z\times x)\times y + (y\times z)\times x = 0\ .$$

(Beachte: im Allgemeinen ist $(x\times y)\times z\neq x\times(y\times z)$).

Beweis:

(1) und (2) folgen aus der Definition.

(3) Man rechnet nach, dass

$$\det\begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix} v_1 - \det\begin{pmatrix} a_1 & a_3 \\ b_1 & b_3 \end{pmatrix} v_2 + \det\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} v_3$$

die in der Definition von $(\sum_{i=1}^{3} a_i v_i) \times (\sum_{i=1}^{3} b_i v_i)$ geforderten Eigenschaften hat.

(4) Kann mit (3) nachgerechnet werden.

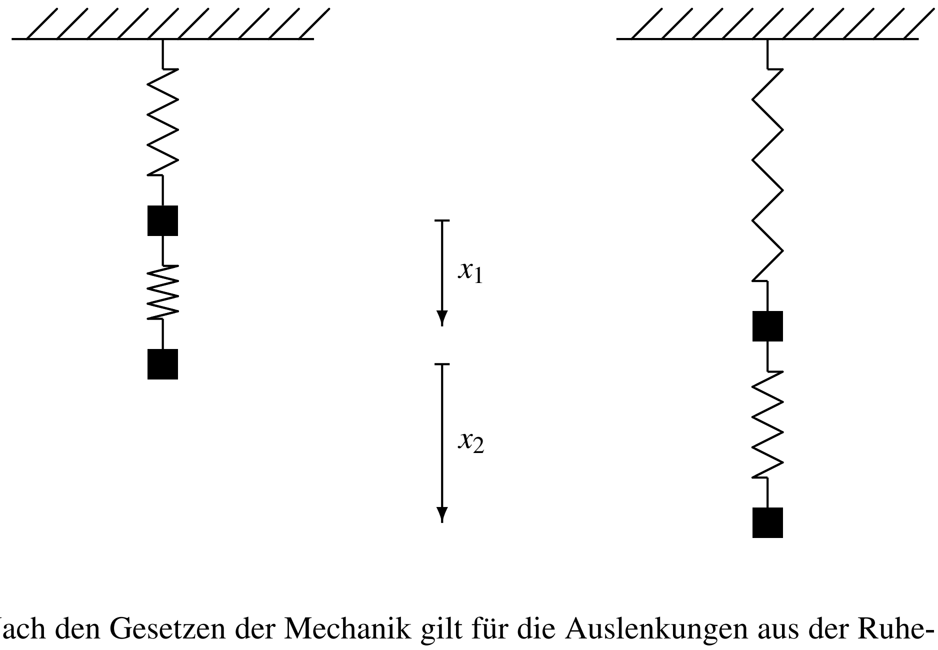
(5) Wenn $\{i,j,k\} = \{1,2,3\}$ ist , dann ist $v_i \times (v_j \times v_k) = 0$. Wenn $i \neq k$ bzw. $i \neq j$, dann ist $v_i \times (v_i \times v_k) = -v_k$ und $v_i \times (v_j \times v_i) = v_j$.
Sei $x = \sum_{i=1}^{3} a_i v_i$, $y = \sum_{j=1}^{3} b_j v_j$ und $z = \sum_{k=1}^{3} c_k v_k$. Dann ist
$x \times (y \times z) = \sum_{i,j,k} a_i b_j c_k (v_i \times (v_j \times v_k)) =$
$= -\sum_{i,k} a_i b_i c_k v_k + \sum_{i,j} a_i b_j c_i v_j = \langle x,z \rangle y - \langle x,y \rangle z$.

(6) Nach (5) ist $x \times (y \times z) + z \times (x \times y) + y \times (z \times x) =$
$= \langle x,z \rangle y - \langle x,y \rangle z + \langle z,y \rangle x - \langle z,x \rangle y + \langle y,x \rangle z - \langle y,z \rangle x = 0$.

§7. Eigenwerte und Eigenvektoren

In diesem Abschnitt sei K ein Körper und n eine positive ganze Zahl.

Beispiel 190 : Zwei Gewichte mit Masse m hängen hintereinander an zwei Federn mit Federkonstante k.



Nach den Gesetzen der Mechanik gilt für die Auslenkungen aus der Ruhelage x_1 bzw. x_2 des ersten bzw. zweiten Gewichts

$$\begin{array}{rcrcrcl} mx_1'' & + & kx_1 & - & k(x_2 - x_1) & = & 0 \\ mx_2'' & & & + & k(x_2 - x_1) & = & 0 \end{array} \quad ,$$

wobei $''$ die zweite Ableitung nach der Zeit bezeichnet. In Matrizenform umgeschrieben erhält man

$$m\begin{pmatrix} x_1'' \\ x_2'' \end{pmatrix} + \begin{pmatrix} 2k & -k \\ -k & k \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0\,.$$

Wir untersuchen nun die Frage, ob es eine Schwingung der Form

$$x_1(t) = a_1 \sin(\omega t)$$
$$x_2(t) = a_2 \sin(\omega t)$$

gibt, wobei a_1, a_2 die Amplituden sind und ω die Frequenz ist. In diesem Fall wäre

$$-m\omega^2 \sin(\omega t)\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \sin(\omega t)\begin{pmatrix} 2k & -k \\ -k & k \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0$$

für alle reellen Zahlen t, also

$$\begin{pmatrix} 2k & -k \\ -k & k \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = m\omega^2\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Daher suchen wir Spalten, die durch Multiplikation mit einer vorgegebenen Matrix in ein skalares Vielfaches übergehen (Fortsetzung in Beispiel 197).

Definition 191 : Sei $A \in K^{n\times n}$.

(1) Eine Spalte $u \in K^{n\times 1}$ heißt *Eigenvektor* von A, wenn $u \neq 0$ ist und ein Element $c \in K$ existiert mit

$$Au = cu\,.$$

(2) Ein Element $c \in K$ heißt *Eigenwert* von A, wenn es eine Spalte $u \in K^{n\times 1}$ gibt mit $u \neq 0$ und

$$Au = cu\,.$$

Eine solche Spalte u heißt *Eigenvektor* von A zum Eigenwert c.

(3) Für einen Eigenwert c von A ist

$$E(A,c) := \{y \in K^{n\times 1} \mid Ay = cy\} = L(cI_n - A, 0)$$

ein Untervektorraum von $K^{n\times 1}$, heißt der *Eigenraum* von A zum Eigenwert c, und besteht aus dem Nullvektor sowie allen Eigenvektoren von A zum Eigenwert c.

(4) Eine *Eigenbasis* von A ist eine Basis von $K^{n\times 1}$, deren Vektoren Eigenvektoren von A sind.

Beispiel 192 : Jede Spalte in $K^{n\times 1}$ ist ein Eigenvektor der Einheitsmatrix I_n zum Eigenwert 1. Jede Basis von $K^{n\times 1}$ ist eine Eigenbasis von I_n.

Satz 193: *Seien $A \in K^{n\times n}$, ℓ eine positive ganze Zahl und $u_1,\dots,u_\ell$ Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $c_1,\dots,c_\ell$. Dann ist das n-Tupel $(u_1,\dots,u_\ell)$ linear unabhängig. Insbesondere hat A höchstens n paarweise verschiedene Eigenwerte.*
Wenn $\ell = n$ ist, A also n paarweise verschiedene Eigenwerte hat, dann ist das n-Tupel $(u_1,\dots,u_n)$ von Eigenvektoren eine Eigenbasis von A.

Beweis: Wir zeigen die Behauptung durch Induktion nach ℓ. Für $\ell = 1$ folgt die Behauptung aus $u_1 \neq 0$. Sei nun $\ell \geq 2$ und die Behauptung gelte für $\ell - 1$ Eigenvektoren. Für $d_1,\dots,d_\ell \in K$ mit

$$0 = \sum_{i=1}^{\ell} d_i u_i \tag{1}$$

folgt

$$0 = A \cdot 0 = \sum_{i=1}^{\ell} d_i A u_i = \sum_{i=1}^{\ell} d_i c_i u_i \,.$$

Subtrahiert man davon das c_ℓ-fache von (1), so erhält man

$$0 = \sum_{i=1}^{\ell-1} d_i(c_i - c_\ell)u_i = 0\,.$$

Nach Induktionsannahme ist das $(\ell-1)$-Tupel $(u_1,\dots,u_{\ell-1})$ linear unabhängig und somit $d_i(c_i - c_\ell) = 0$ für $i = 1,\dots,\ell-1$. Da $c_1,\dots,c_\ell$ paarweise verschieden sind, ist auch $d_i = 0$ für $i = 1,\dots,\ell-1$. Aus (1) folgt $0 = d_\ell u_\ell$ und wegen $u_\ell \neq 0$ auch $d_\ell = 0$.

Satz 194: *Sei $A \in K^{n\times n}$ eine Matrix, die eine Eigenbasis hat. Sei $T \in K^{n\times n}$ eine Matrix, deren Spalten eine Eigenbasis von A bilden. Der Eigenwert von T_{-i} sei c_i, $1 \leq i \leq n$. Dann ist $T^{-1}AT$ eine Diagonalmatrix und*

$$(T^{-1}AT)_{ii} = c_i\,,\ 1 \leq i \leq n\,.$$

Beweis: Es seien e_i, $1 \leq i \leq n$, die Standardspalten in $K^{n\times 1}$. Für jede Matrix $B \in K^{n\times n}$ ist $B \cdot e_i = B_{-i}$, $1 \leq i \leq n$. Daher ist die i-te Spalte von $T^{-1}AT$

$$\begin{aligned}(T^{-1}AT)_{-i} &= (T^{-1}AT)e_i = (T^{-1}A)Te_i = \\ &= T^{-1}(AT_{-i}) = T^{-1}(c_iT_{-i}) = c_ie_i\,,\ 1 \leq i \leq n.\end{aligned}$$

Beispiel 195: Satz 194 kann zum Berechnen von „großen Potenzen“ von Matrizen A, die eine Eigenbasis haben, benutzt werden. Es ist nämlich

$$(T^{-1}AT)^k = (T^{-1}AT)(T^{-1}AT)\dots(T^{-1}AT) = T^{-1}A^kT\,,$$

daher

$$A^k = T(T^{-1}AT)^kT^{-1}\,.$$

Der Aufwand für die Berechnung der k-ten Potenz der Diagonalmatrix $T^{-1}AT$ ist im allgemeinen wesentlich geringer als der für die Berechnung von A^k.

Satz 196 : *Sei $A \in K^{n\times n}$. Dann ist $c \in K$ genau dann ein Eigenwert von A, wenn*

$$\det(cI_n - A) = 0$$

ist. Die Funktion

$$\chi_A : K \to K\,,\, z \mapsto \det(zI_n - A)\,,$$

ist eine Polynomfunktion und heißt das charakteristische Polynom *von A. („Die Eigenwerte einer Matrix sind die Nullstellen ihres charakteristischen Polynoms.“)*

Beweis: Es ist c genau dann ein Eigenwert von A, wenn ein Vektor $u \in K^{n\times 1}$ mit $u \neq 0$ existiert, sodass $Au = cu$ ist. Dies ist äquivalent dazu, dass das durch $cI_n - A$ gegebene homogene System linearer Gleichungen eine nicht-triviale Lösung hat. Das ist genau dann der Fall, wenn die Matrix $cI_n - A$ nicht invertierbar ist. Nach Satz 181 ist diese Matrix genau dann nicht invertierbar, wenn $\det(cI_n - A) = 0$ ist.

Satz 196 legt folgendes Verfahren nahe, die Eigenwerte und Eigenvektoren einer Matrix $A \in K^{n\times n}$ zu berechnen:

(1) Finde alle $c \in K$ mit $\det(cI_n - A) = 0$ („Berechne alle Nullstellen des charakteristischen Polynoms von A“).
(2) Bestimme für jeden Eigenwert c den Eigenraum $E(A,c)$ durch Lösen des homogenen Systems linearer Gleichungen $(cI_n - A, 0)$.

Beispiel 197 : Wir lösen nun das Eigenwertproblem aus Beispiel 190. Hier ist

$$A = \begin{pmatrix} 2k & -k \\ -k & k \end{pmatrix}.$$

Wegen

$$\det(cI_2 - A) = \det \begin{pmatrix} c-2k & k \\ k & c-k \end{pmatrix} = c^2 - 3kc + k^2$$

sind die Eigenwerte von A

$$c_1 = \frac{1}{2}(3+\sqrt{5})k \qquad \text{und} \qquad c_2 = \frac{1}{2}(3-\sqrt{5})k$$

(siehe Satz 200). Die zugehörigen Eigenräume sind

$$E(A,c_1) = \mathbb{R}\begin{pmatrix} 2 \\ 1-\sqrt{5} \end{pmatrix} \qquad \text{bzw.} \qquad E(A,c_2) = \mathbb{R}\begin{pmatrix} 2 \\ 1+\sqrt{5} \end{pmatrix}.$$

Beispiel 198 : Sei $A \in K^{n \times n}$ eine Dreiecksmatrix. Dann ist auch die Matrix $cI_n - A$ eine Dreiecksmatrix, und nach Satz 178 ist

$$\chi_A(z) = (z - A_{11}) \dots (z - A_{nn}) .$$

Daher sind die Diagonalelemente einer Matrix in Dreiecksform ihre Eigenwerte.
Zum Beispiel sind 2 und 3 die Eigenwerte der Matrix

$$\begin{pmatrix} 2 & 4 & 6 & 8 \\ 0 & 3 & 5 & 7 \\ 0 & 0 & 2 & 9 \\ 0 & 0 & 0 & 3 \end{pmatrix} .$$

§8. Quadratische Funktionen und komplexe Zahlen

In diesem Kapitel sei K ein Körper.

Definition 199 : Sei f eine Funktion von K nach K. Ein Element $y \in K$ ist genau dann eine *Nullstelle von* f , wenn $f(y) = 0$ ist.
Seien $a_0, a_1, a_2 \in K$. Die Polynomfunktion mit Koeffizienten a_0, a_1, a_2 heißt genau dann eine *quadratische Funktion* von K nach K, wenn $a_2 \neq 0$ ist.
Schreiben wir x für die identische Funktion

$$x : K \to K , z \mapsto z$$

und a_0 für die Funktion, die jedem Element von K das Element a_0 zuordnet, dann ist $a_2x^2 + a_1x + a_0$ die quadratische Funktion mit Koeffizienten a_0, a_1, a_2.

Satz 200 : *Seien $p, q \in K$ und $f := x^2 + px + q$ die quadratische Funktion mit den Koeffizienten $1, p, q$. Wir nehmen an, dass in K gilt: $2 := 1_K + 1_K \neq 0$. (Insbesondere ist K nicht der Körper $\mathbb{Z}_2$).*
Die quadratische Funktion f hat genau dann eine Nullstelle in K, wenn es in K ein Element z mit

$$z^2 = (\frac{p}{2})^2 - q$$

gibt. In diesem Fall sind

$$-\frac{p}{2} + z \quad und \quad -\frac{p}{2} - z$$

die (einzigen) Nullstellen von f.

Beweis: Für $y \in K$ ist

$$y^2 + py + q = (y + \frac{p}{2})^2 - (\frac{p}{2})^2 + q .$$

Es ist daher $f(y) = 0$ genau dann, wenn

$$(y+\frac{p}{2})^2 = (\frac{p}{2})^2 - q =: d$$

ist. Wenn ein Element $z \in K$ mit $z^2 = d$ existiert, dann hat auch $-z$ diese Eigenschaft. Ist u ein Element mit $u^2 = d$, dann folgt aus $u^2 = z^2$, dass $0 = u^2 - z^2 = (u-z)\cdot(u+z)$ ist, also $u = z$ oder $u = -z$. Daher ist $f(y) = 0$ genau dann, wenn $y = -\frac{p}{2} + z$ oder $y = -\frac{p}{2} - z$ ist.

Beispiel 201 : Es ist $f := x^2 + 4x - 12 = (x+2)^2 - 4 - 12 = (x+2)^2 - 16$. Wegen $(-4)^2 = 16 = 4^2$ sind $-2+4 = 2$ und $-2-4 = -6$ die Nullstellen der quadratischen Funktion f.
Es gibt keine reelle Zahl, deren Quadrat -1 ist. Daher hat die quadratische Funktion $x^2 + 1 : \mathbb{R} \longrightarrow \mathbb{R}$ keine Nullstellen.

Wir lernen nun einen Körper kennen, in dem jede nicht-konstante Polynomfunktion mindestens eine Nullstelle hat: den Körper der komplexen Zahlen. Der Beweis für die Existenz dieser Nullstellen wird in der Analysis geführt (weil dabei Eigenschaften der reellen Zahlen verwendet werden).

Satz 202 : *Sei $\mathbb{R}$ der Körper der reellen Zahlen. Die Menge $\mathbb{R} \times \mathbb{R}$ mit der Addition*

$$(a_1,a_2) + (b_1,b_2) := (a_1+b_1, a_2+b_2)$$

und der Multiplikation

$$(a_1,a_2)\cdot(b_1,b_2) := (a_1b_1 - a_2b_2, a_1b_2 + a_2b_1)$$

ist ein Körper mit Nullelement $(0,0)$ und Einselement $(1,0)$.
Dieser Körper heißt Körper der komplexen Zahlen *und wird mit $\mathbb{C}$ bezeichnet.*

Das zu $(a_1,a_2) \in \mathbb{C} \setminus \{(0,0)\}$ inverse Element ist

$$(a_1,a_2)^{-1} = (\frac{a_1}{a_1^2+a_2^2}, \frac{-a_2}{a_1^2+a_2^2}).$$

Es ist $(0,1)^2 = (-1,0) = -(1,0)$, also sind $(0,1)$ und $(0,-1)$ die Nullstellen von $x^2+1 : \mathbb{C} \longrightarrow \mathbb{C}\,, t \longmapsto t^2+1$.

Beweis: Übung.

Definition 203 : Für $z := (a_1,a_2) \in \mathbb{C}$ wird $a_1 + a_2 i \in \mathbb{C}$ geschrieben, $\mathrm{Re}(z) := a_1$ bzw. $\mathrm{Im}(z) := a_2$ heißt *Realteil* bzw. *Imaginärteil* von z .
Statt $0 + a_2 i$ bzw. $a_1 + 0i$ schreibt man einfach $a_2 i$ bzw. a_1. Die komplexe Zahl $\overline{a_1 + a_2 i} := a_1 - a_2 i$ heißt die *zu $a_1 + a_2 i$ konjugierte komplexe Zahl.*

Üblicherweise identifiziert man eine reelle Zahl a mit der komplexen Zahl $(a,0) = a+0i$, weil hier reelle und komplexe Addition bzw. Multiplikation übereinstimmen:

$$(a_1,0)+(b_1,0) = (a_1+b_1,0) \quad \text{und} \quad (a_1,0)\cdot(b_1,0) = (a_1\cdot b_1,0).$$

Daher kann $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ aufgefasst werden.

Für jede komplexe Zahl z ist

$$z\cdot\overline{z} = \mathrm{Re}(z)^2+\mathrm{Im}(z)^2$$

eine (nicht negative) reelle Zahl. Daher ist

$$z^{-1} = (\mathrm{Re}(z)^2+\mathrm{Im}(z)^2)^{-1}\overline{z}.$$

Beispiel 204 : Die Eigenwerte der Matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

sind die Nullstellen von x^2+1, also i und $-i$. Eine Eigenbasis dieser Matrix ist

$$(\begin{pmatrix} 1 \\ -i \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix}).$$

§9. Fragen

1. Welche der folgenden Aussagen sind wahr?
   (a) Sei $N = \{1,2,3,4,5\}$.
   Eine bijektive Funktion von N nach N nennt man eine Permutation der Zahlen $1,...,5$.
   (b) Die Umkehrfunktion der Permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$ ist $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix}$.

2. Gegeben sei die Permutation in Tabellenform
$$s := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 5 & 7 & 3 & 6 & 8 & 4 & 2 \end{pmatrix}.$$
   Welche der folgenden Aussagen sind wahr?
   (a) Die Zahl 1 ist ein Fixpunkt von s.
   (b) Die Permutation s hat genau drei disjunkte Zykel.
   (c) Das Vorzeichen von s ist 1.
   (d) Die Permutation s ist ungerade.

3. Die Determinante von $A \in K^{n\times n}$ kann wie folgt definiert werden.
   (a)
$$\det(A) := \sum_{\sigma\in S_n} sign(\sigma)(A_{\sigma(1)\sigma(n)}A_{\sigma(2)\sigma(n-1)}...A_{\sigma(n)\sigma(1)})$$
   (b)
$$\det(A) := \sum_{\sigma\in S_n} sign(\sigma)A_{1\sigma(1)}A_{2\sigma(2)}...A_{n\sigma(n)}$$
   (c)
$$\det(A) := \sum_{\sigma\in S_n} sign(\sigma)(A_{\sigma(1)1} + A_{\sigma(2)2} + ... + A_{\sigma(n)n})$$
   (d)
$$\det(A) := \sum_{\sigma\in S_n} sign(\sigma)A_{\sigma(1)1}A_{\sigma(2)2}...A_{\sigma(n)n}$$

4. Sei $A \in K^{n\times n}$.
   Welche der folgende Aussagen sind wahr?
   (a) $\det(A^T) = -\det(A)$
   (b) Wenn A eine Dreiecksmatrix ist, dann ist $\det(A) = A_{11}A_{22}...A_{nn}$.
   (c) Die Determinante der Matrix $\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 1 & 3 \end{pmatrix}$ ist Null.

5. Es sei $A := \begin{pmatrix} 1 & 0 & 2 & 4 \\ 2 & 1 & 0 & 3 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

   Welche der folgenden Aussagen sind wahr?
   (a) $\det(A) = 9$
   (b) $\det(A) =$

$$= A_{11}A_{22}A_{33}A_{44} + A_{12}A_{23}A_{34}A_{41} + A_{13}A_{24}A_{31}A_{42} + A_{14}A_{21}A_{32}A_{43} -$$
$$-A_{14}A_{23}A_{32}A_{41} - A_{11}A_{24}A_{33}A_{42} - A_{12}A_{21}A_{34}A_{43} - A_{13}A_{22}A_{31}A_{44}$$

6. Welche der folgenden Aussagen sind für alle $A \in K^{3\times 3}$ wahr ?
   Sei B die Matrix, die man aus A durch Multiplikation der ersten Zeile von A mit 3 erhält.
   (a) $\det(B) = \det(A)^3$
   (b) $\det(B) = 3 \cdot \det(A)$
   (c) $\det(B) = \det(A) + 3$

7. Welche der folgenden Aussagen sind für alle $A \in K^{3\times 3}$ wahr ?
   (a) $\det(3A) = 3\det(A)$
   (b) $\det(3A) = 3^3\det(A)$
   (c) $\det(3A) = \frac{1}{3}\det(A)$

8. Welche der folgenden Aussagen sind für alle $A, B \in K^{n\times n}$ wahr ?
   (a) $\det(A \cdot B) = \det(A) + \det(B)$
   (b) $\det(A + B) = \det(A) + \det(B)$
   (c) $\det(A \cdot B) = \det(B) \cdot \det(A)$

9. Sei A die Matrix $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{pmatrix}$.
   Welche der folgenden Aussagen sind wahr?
   (a) Die Matrix A hat genau drei Eigenwerte, und zwar 0, 1 und 2.
   (b) Die Eigenwerte erhält man, indem man die Nullstellen des charakteristischen Polynoms von A berechnet.
   (c) $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ ist ein Eigenvektor von A mit Eigenwert 2.
   (d) Der Eigenraum von A zum Eigenwert 2 ist die Lösungsmenge des Systems linearer Gleichungen
   $\begin{pmatrix} 1 & -2 & -1 \\ -2 & 2 & 2 \\ 1 & -2 & -1 \end{pmatrix} x = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

KAPITEL 5

Lineare Funktionen

In diesem Kapitel sei K ein Körper und n eine positive ganze Zahl.

§1. Lineare Funktionen

In diesem Abschnitt seien V und W Vektorräume über einem Körper K.

Definition 205 : Eine Funktion $f : V \to W$ heißt *K-linear*, wenn die folgenden zwei Bedingungen erfüllt sind:

(1) Für alle $v, w \in V$ ist $f(v+w) = f(v) + f(w)$.
(„Das Bild der Summe ist die Summe der Bilder.“)

(2) Für alle $c \in K$ und für alle $v \in V$ ist $f(cv) = cf(v)$.
(„Das Bild des c-fachen ist das c-fache des Bildes.“)

Beispiel 206 : Die *Nullfunktion*

$$0 : V \to W \,,\, v \mapsto 0_W \,,$$

und die Identität

$$\mathrm{Id}_V : V \to V \,,\, v \mapsto v \,,$$

sind K-linear.

Beispiel 207 : V sei ein Untervektorraum des K-Vektorraums $\mathcal{F}(K,K)$ aller Funktionen von K nach K und $t_0, \ldots, t_n$ seien paarweise verschiedene Elemente von K. Dann ist die „Auswertungsfunktion“

$$a : V \longrightarrow K^{n+1}, f \longmapsto (f(t_0), f(t_1), \ldots, f(t_n)),$$

linear.

Beispiel 208 : Für jede Matrix $A \in K^{m \times n}$ ist die Funktion

$$K^{n\times 1} \to K^{m\times 1} \,, x \mapsto Ax \,,$$

K-linear. Später werden wir sehen, dass jede lineare Funktion vom Vektorraum aller m-Spalten in den Vektorraum aller n-Spalten durch Multiplikation mit einer Matrix gegeben ist.

Beispiel 209 : Ein Kaufhaus bietet n Waren an. Kauft jemand a_i Einheiten der i-ten Ware, $1 \leq i \leq n$, so muss er $p(a_1, \ldots, a_n)$ Euro zahlen. Die Funktion

$$p : \mathbb{Q}^n \longrightarrow \mathbb{Q}\,,\ (a_1, \ldots, a_n) \longmapsto p(a_1, \ldots, a_n)$$

ist genau dann linear, wenn es keinen Mengenrabatt, keine Sonderaktionen („nimm drei, zahl' zwei") oder ähnliches gibt, also:

(1) Nimmt man bei einem Einkauf a_i und bei einem anderen Einkauf b_i Einheiten der i-ten Ware, $1 \leq i \leq n$, dann bezahlt man in Summe dasselbe, wie wenn man alles bei einem Einkauf genommen hätte ($p(a_1, \ldots, a_n) + p(b_1, \ldots, b_n) = p(a_1 + b_1, \ldots, a_n + b_n)$).
(2) Kauft man von jeder Ware das c-fache, dann muss man c-mal soviel zahlen ($p(ca_1, \ldots, ca_n) = c \cdot p(a_1, \ldots, a_n)$).

Satz 210 : *Seien $f : V \to W$ eine lineare Funktion, $v_1, \ldots, v_n \in V$ und $c_1, \ldots, c_n \in K$. Dann ist*

$$f(\sum_{i=1}^{n} c_i v_i) = \sum_{i=1}^{n} c_i f(v_i)\,.$$

(„Das Bild der Linearkombination ist die Linearkombination der Bilder.") Speziell ist

$$f(0_V) = 0_W$$

und, für alle $v \in V$,

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

Beweis: Induktion nach n.

Satz 211 : *Seien U ein Vektorraum über K und $f : U \to V$ sowie $g : V \to W$ lineare Funktionen. Dann gilt:*

(1) *Die Zusammensetzung $g \circ f : U \to W\,,\ u \mapsto g(f(u))$, ist K-linear.*
(2) *Wenn f bijektiv ist, dann ist die Umkehrfunktion $f^{-1} : V \to U$ auch K-linear.*

(„Die Zusammensetzung von linearen Funktionen ist linear. Die Umkehrfunktion einer bijektiven linearen Funktion ist linear.")

Beweis:

(1) Für $v, w \in U$ und $c \in K$ ist
$(g \circ f)(v + w) = g(f(v + w)) = g(f(v) + f(w)) =$
$= g(f(v)) + g(f(w)) = (g \circ f)(v) + (g \circ f)(w)$,
$(g \circ f)(cv) = g(f(cv)) = g(cf(v)) = cg(f(v)) = c(g \circ f)(v)$,
(2) Für $v, w \in V$ und $c \in K$ ist
$f(f^{-1}(v) + f^{-1}(w)) = (f \circ f^{-1})(v) + (f \circ f^{-1})(w) = v + w$,

also $f^{-1}(v+w) = f^{-1}(v) + f^{-1}(w)$, und
$f(cf^{-1}(v)) = c(f \circ f^{-1})(v) = cv$, also $f^{-1}(cv) = cf^{-1}(v)$.

Definition 212 : Eine lineare und bijektive Funktion von V nach W heißt *Isomorphismus* von Vektorräumen. V und W heißen isomorph, wenn es einen Isomorphismus von V nach W (oder von W nach V) gibt. Schreibweise: $V \cong W$.

Satz 213 :

(1) *Jeder Vektorraum V über K der Dimension n ist zum Standard-Vektorraum K^n isomorph.*
Nach Wahl einer Basis $(v_1, \ldots, v_n)$ erhält man einen Isomorphismus durch

$$V \longrightarrow K^n \,,\, u \mapsto \textit{Koordinaten-n-Tupel von } u \textit{ bezüglich } (v_1, \ldots, v_n)\,.$$

(2) *Zwei endlich-dimensionale Vektorräume über K sind genau dann isomorph, wenn ihre Dimensionen gleich sind.*

Beweis:

(1) Weil $(v_1, \ldots, v_n)$ eine Basis ist, ist die Funktion

$$K^n \to V \,,\, (c_1, \ldots, c_n) \mapsto \sum_{i=1}^{n} c_i v_i \,,$$

bijektiv. Nach Satz 99 ist sie auch linear. Nach Satz 211 ist dann die Koordinaten-Funktion als Umkehrfunktion linear und bijektiv, also ein Isomorphismus.

(2) Seien V, W endlich-dimensionale Vektorräume über K. Wenn V und W isomorph sind, dann gibt es eine bijektive lineare Funktion $f : V \to W$. Sei $(v_1, \ldots, v_n)$ eine Basis von V. Mit Satz 210 prüft man leicht nach, dass dann $(f(v_1), \ldots, f(v_n))$ eine Basis von W ist, also ist $\dim_K(V) = \dim_K(W)$.
Sei umgekehrt $\dim_K(V) = \dim_K(W) =: n$. Dann gibt es nach (1) Isomorphismen $f : V \to K^n$ und $g : W \to K^n$. Nach Satz 211 ist dann auch $g^{-1} \circ f : V \to W$ ein Isomorphismus.

§2. Die Matrix einer linearen Funktion

In diesem Abschnitt sei V ein Vektorraum über K mit Dimension $n \in \mathbb{N}$, $\underline{v} := (v_1, \ldots, v_n)$ eine Basis von V, W ein Vektorraum über K mit Dimension $m \in \mathbb{N}$ und $\underline{w} := (w_1, \ldots, w_m)$ eine Basis von W.

Satz 214 : *Seien $u_1, \ldots, u_n \in W$. Dann gibt es genau eine lineare Funktion $f : V \to W$ mit*

$$f(v_i) = u_i \, , 1 \leq i \leq n \; .$$

Somit kann eine lineare Funktion zwischen Vektorräumen durch (beliebige) Vorgabe der Bilder einer Basis eindeutig definiert werden.

Beweis: Wenn eine derartige Funktion f existiert, dann ist für einen Vektor $\sum_{i=1}^n c_i v_i \in V$

$$f(\sum_{i=1}^n c_i v_i) = \sum_{i=1}^n c_i f(v_i) = \sum_{i=1}^n c_i u_i \, ,$$

was die Eindeutigkeit der Funktion beweist.
Um die Existenz einer solchen Funktion zu zeigen, definieren wir eine Funktion $f : V \to W$ durch

$$f(x) := \sum_{i=1}^n c_i u_i \, ,$$

wobei $c_1, \ldots, c_n \in K$ die Koordinaten von $x \in V$ bezüglich der Basis $\underline{v}$ sind. Dann ist f K-linear, weil für $x = \sum_{i=1}^n c_i v_i$, $y = \sum_{i=1}^n d_i v_i$ und $t \in K$ wegen $x + y = \sum_{i=1}^n (c_i + d_i) v_i$ und $tx = \sum_{i=1}^n (tc_i) v_i$

$$f(x+y) = \sum_{i=1}^n (c_i + d_i) u_i = \sum_{i=1}^n c_i u_i + \sum_{i=1}^n d_i u_i = f(x) + f(y)$$

sowie

$$f(tx) = \sum_{i=1}^n t c_i u_i = t \left(\sum_{i=1}^n c_i u_i \right) = t f(x)$$

ist. Schließlich gilt wegen $v_j = \sum_{i=1}^n \delta_{ij} v_i$ für alle $1 \leq j \leq n$:

$$f(v_j) = \sum_{i=1}^n \delta_{ij} u_i = u_j \, .$$

Definition 215 : Seien $f : V \to W$ eine K-lineare Funktion und $A_{1j}, \ldots, A_{mj}$ die Koordinaten von $f(v_j)$ bezüglich $\underline{w}$, d.h.

$$f(v_j) = \sum_{i=1}^m A_{ij} w_i \; , 1 \leq j \leq n \, .$$

Dann heißt

$$M(f, \underline{v}, \underline{w}) := (A_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in K^{m \times n}$$

die *Matrix von f bezüglich der Basen $\underline{v}$ und $\underline{w}$*.
Die j-te Spalte von $M(f, \underline{v}, \underline{w})$ ist also die Koordinatenspalte von $f(v_j)$ bezüglich $\underline{w}$.

Im Spezialfall $V = W$ und $\underline{v} = \underline{w}$ schreibt man statt $M(f, \underline{v}, \underline{v})$ kurz $M(f, \underline{v})$.

Beispiel 216 : Für $1 \leq j \leq n$ ist

$$\mathrm{Id}_V(v_j) = v_j = \sum_{i=1}^{n} \delta_{ij} v_i$$

und somit

$$M(\mathrm{Id}_V, \underline{v}) = I_n \,.$$

Beispiel 217 : Sei $A \in K^{m\times n}$. Dann ist die Matrix der linearen Abbildung

$$K^{n\times 1} \to K^{m\times 1}, x \mapsto Ax,$$

bezüglich der Standardbasen $(e_1, \ldots, e_n)$ von $K^{n\times 1}$ und $(e'_1, \ldots, e'_m)$ von $K^{m\times 1}$ gleich A, weil für $1 \leq j \leq n$

$$Ae_j = A_{-j} = \sum_{i=1}^{m} A_{ij} e'_i$$

ist.
Wenn $f : K^{n\times 1} \to K^{m\times 1}$ eine lineare Funktion ist und $A := M(f, \underline{e}, \underline{e}')$, dann ist

$$f(x) = f(\sum_{i=1}^{n} x_i e_i) = \sum_{i=1}^{n} x_i f(e_i) = \sum_{i=1}^{n} x_i A_{-i} = Ax \,.$$

Daher ist jede lineare Funktion von $K^{n\times 1}$ nach $K^{m\times 1}$ durch Multiplikation mit einer eindeutig bestimmten Matrix $A \in K^{m\times n}$ gegeben. Oft identifiziert man eine $m \times n$-Matrix mit dieser linearen Funktion.

Beispiel 218 : Wir betrachten $\mathbb{C} = \mathbb{R}^2$ als zweidimensionalen reellen Vektorraum mit der Basis $(1, i)$. Es sei $f : \mathbb{R}^2 = \mathbb{C} \longrightarrow \mathbb{R}^2 = \mathbb{C}$ die Funktion, die jeder komplexen Zahl z ihr Produkt mit $1+2i$ zuordnet. Man prüft leicht nach, dass diese Funktion $\mathbb{R}$-linear ist. Es ist $f(1) = 1+2i$ und $f(i) = i-2$. Daher ist

$$M(f, (1, i)) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

die Matrix von f bezüglich der Standardbasis.

Satz 219 : *Seien U ein Vektorraum über K der Dimensionen ℓ mit Basis $\underline{u}$ und $f : V \to W$ sowie $g : W \to U$ K-lineare Funktionen. Dann ist*

$$M(g \circ f, \underline{v}, \underline{u}) = M(g, \underline{w}, \underline{u}) \cdot M(f, \underline{v}, \underline{w}) \,,$$

d.h. der Zusammensetzung von linearen Funktionen entspricht die Multiplikation der zugehörigen Matrizen.

Beweis: Sei $A := M(f,\underline{v},\underline{w}) \in K^{m\times n}$ und $B := M(g,\underline{w},\underline{u}) \in K^{\ell\times m}$. Dann ist für $1 \leq j \leq n$

$$
\begin{aligned}
(g\circ f)(v_j) = g(f(v_j)) &= g(\sum_{i=1}^{m} A_{ij}w_i) = \sum_{i=1}^{m} A_{ij}g(w_i) \\
&= \sum_{i=1}^{m} A_{ij} \sum_{k=1}^{\ell} B_{ki}u_k = \sum_{k=1}^{\ell} \left(\sum_{i=1}^{m} B_{ki}A_{ij} \right) u_k \\
&= \sum_{k=1}^{\ell} (BA)_{kj}u_k\,,
\end{aligned}
$$

also $BA = M(g\circ f,\underline{v},\underline{u})$.

Satz 220 : *Die lineare Funktion $f : V \to W$ ist genau dann ein Isomorphismus, wenn die Matrix $M(f,\underline{v},\underline{w})$ invertierbar ist. In diesem Fall ist*

$$M(f^{-1},\underline{w},\underline{v}) = M(f,\underline{v},\underline{w})^{-1}\,.$$

Beweis: Wenn f ein Isomorphismus ist, dann ist nach Satz 211 die Umkehrfunktion f^{-1} linear und nach Satz 213 ist $n = m$. Aus $f^{-1}\circ f = \mathrm{Id}_V$ und $f\circ f^{-1} = \mathrm{Id}_W$ folgt nach Satz 219

$$M(f^{-1},\underline{w},\underline{v})\cdot M(f,\underline{v},\underline{w}) = M(f^{-1}\circ f,\underline{v},\underline{v}) = M(\mathrm{Id}_V,\underline{v},\underline{v}) = I_n$$

und

$$M(f,\underline{v},\underline{w})\cdot M(f^{-1},\underline{w},\underline{v}) = M(f\circ f^{-1},\underline{w},\underline{w}) = M(\mathrm{Id}_W,\underline{w},\underline{w}) = I_m\,.$$

Wenn umgekehrt $A := M(f,\underline{v},\underline{w}) \in K^{m\times n}$ invertierbar ist, dann ist $m = n$ und es existiert $B := A^{-1} \in K^{n\times n}$. Definiert man nach Satz 214 eine lineare Funktion $g : W \to V$ durch

$$g(w_j) := \sum_{i=1}^{n} B_{ij}v_i$$

für $1 \leq j \leq n$, so folgt nach Satz 219

$$M(g\circ f,\underline{v},\underline{v}) = M(g,\underline{w},\underline{v})\cdot M(f,\underline{v},\underline{w}) = BA = I_n = M(\mathrm{Id}_V,\underline{v},\underline{v})$$

und

$$M(f\circ g,\underline{w},\underline{w}) = M(f,\underline{v},\underline{w})\cdot M(g,\underline{w},\underline{v}) = AB = I_n = M(\mathrm{Id}_W,\underline{w},\underline{w})\,.$$

Daher ist $g\circ f = \mathrm{Id}_V$ und $f\circ g = \mathrm{Id}_W$ nach Satz 214, also f ein Isomorphismus und $g = f^{-1}$.

Nach Wahl von Basen im Definitionsbereich und im Bildbereich einer linearen Funktion ist diese eindeutig durch ihre Matrix bestimmt. Anstatt mit linearen Funktionen kann mit den entsprechenden Matrizen gerechnet werden. Der Zusammensetzung von linearen Abbildungen entspricht die

Matrizenmultiplikation, dem Berechnen der Umkehrabbildung entspricht das Invertieren von Matrizen.

Wir können nun die Begriffe Eigenvektor, Eigenwert und Eigenraum auch für lineare Funktionen (statt für Matrizen) definieren:

Definition 221 : Sei V ein Vektorraum über K und $f : V \to V$ eine lineare Funktion.

(1) Ein Vektor $u \in V$ heißt *Eigenvektor* von f, wenn $u \neq 0$ ist und eine Zahl $c \in K$ existiert mit

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

In diesem Fall ist c eindeutig bestimmt und heißt der *Eigenwert* von f zum Eigenvektor u.

(2) Für einen Eigenwert c von f ist

$$E(f,c) := \{x \in V \mid f(x) = cx\}$$

ein Untervektorraum von V, heißt der *Eigenraum* von f zum Eigenwert c, und besteht aus dem Nullvektor sowie allen Eigenvektoren von f zum Eigenwert c.

Satz 222 : *Sei V ein n-dimensionaler Vektorraum über K, $f : V \longrightarrow V$ eine lineare Funktion, $\underline{v}$ eine Basis von V und $A := M(f,\underline{v})$ die Matrix von f bezüglich $\underline{v}$.*
Die Spalte $y \in K^{n\times 1}$ ist genau dann ein Eigenvektor von A zum Eigenwert $c \in K$, wenn der Vektor in V mit Koordinatenspalte y bezüglich $\underline{v}$ ein Eigenvektor von f zum Eigenwert $c \in K$ ist.

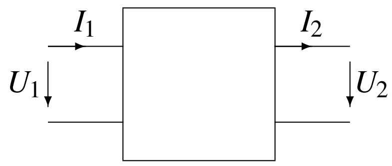
Beweis: Übung.

Satz 222 legt folgendes Verfahren nahe, die Eigenwerte und Eigenvektoren einer linearen Funktion $f : V \to V$ zu berechnen, falls V endlichdimensional ist:

(1) Wähle eine Basis $\underline{v} := (v_1, \ldots, v_n)$ von V und bestimme die Matrix A von f bezüglich $\underline{v}$.
(2) Finde alle $c \in K$ mit $\det(cI_n - A) = 0$. Diese Zahlen sind die Eigenwerte von A und daher auch von f.
(3) Bestimme für alle Eigenwerte c eine Basis des Eigenraums $E(A,c)$ von A zum Eigenwert c. Die Vektoren in V, deren Koordinatenspalten die Elemente dieser Basis sind, bilden eine Basis des Eigenraums $E(f,c)$ von f zum Eigenwert c.

§3. Lineare Funktionen und Vierpole

In diesem Abschnitt wenden wir die Überlegungen von §2 auf eine Fragestellung der Elektrotechnik an.
Unter einem *Vierpol* versteht man ein elektrisches Schaltelement mit je zwei Anschlüssen am Eingang und am Ausgang.



Ein Vierpol wird mathematisch durch die *Vierpolfunktion*

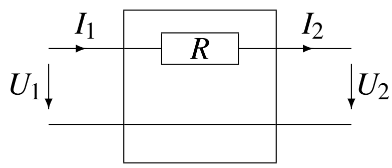
$$\begin{array}{ccc} \mathbb{R}^{2\times 1} & \longrightarrow & \mathbb{R}^{2\times 1} \\ \begin{pmatrix} U_1 \\ I_1 \end{pmatrix} & \longmapsto & \begin{pmatrix} U_2 \\ I_2 \end{pmatrix} \end{array}$$

modelliert, die dem Zahlenpaar (U_1, I_1):=(Spannung am Eingang in Volt, Stromstärke am Eingang in Ampère) das Zahlenpaar (U_2, I_2):=(Spannung am Ausgang in Volt, Stromstärke am Ausgang in Ampère) zuordnet. Wenn diese Funktion linear ist (das bedeutet unter anderem, dass sich Ausgangsspannung und Ausgangsstrom verdoppeln, wenn man Eingangsspannung und Eingangsstrom verdoppelt), können wir sie durch eine Matrix
$K \in \mathbb{R}^{2\times 2}$, wir nennen sie *Kettenmatrix*, beschreiben:

$$\begin{pmatrix} U_2 \\ I_2 \end{pmatrix} = K \cdot \begin{pmatrix} U_1 \\ I_1 \end{pmatrix}.$$

Wir betrachten zwei einfache Vierpole:

1. Sei



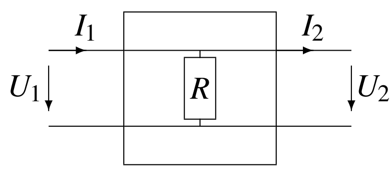
ein Vierpol mit einem Widerstand von R Ohm, der in Serie geschaltet ist. Dann ist die dazugehörige Funktion

$$\begin{array}{rcccccc} f: & \mathbb{R}^{2\times 1} & \longrightarrow & \mathbb{R}^{2\times 1} & & & \\ & \begin{pmatrix} U_1 \\ I_1 \end{pmatrix} & \longmapsto & \begin{pmatrix} U_2 \\ I_2 \end{pmatrix} & = & \begin{pmatrix} 1 & -R \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} U_1 \\ I_1 \end{pmatrix} \end{array}$$

linear.

Die Kettenmatrix dieses Vierpols ist $K_S := \begin{pmatrix} 1 & -R \\ 0 & 1 \end{pmatrix}$.

2. Sei



ein Vierpol mit einem Widerstand von R Ohm, der parallel geschaltet ist. Dann ist die dazugehörige Funktion

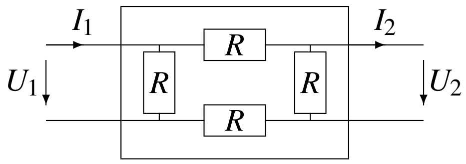
$$\begin{array}{rccccc} g: & \mathbb{R}^{2\times 1} & \longrightarrow & \mathbb{R}^{2\times 1} & & \\ & \begin{pmatrix} U_1 \\ I_1 \end{pmatrix} & \longmapsto & \begin{pmatrix} U_2 \\ I_2 \end{pmatrix} & = & \begin{pmatrix} 1 & 0 \\ -\frac{1}{R} & 1 \end{pmatrix} \begin{pmatrix} U_1 \\ I_1 \end{pmatrix} \end{array}$$

linear.

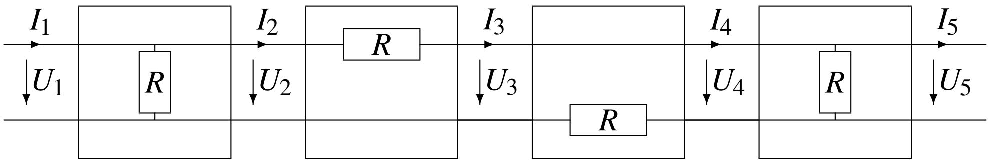
Die Kettenmatrix dieses Vierpols ist $K_P := \begin{pmatrix} 1 & 0 \\ -\frac{1}{R} & 1 \end{pmatrix}$.

Werden Vierpole mit den Vierpolfunktionen h und k und Kettenmatrizen L und M hintereinandergeschaltet (die Eingangsspannung bzw. die Eingangsstromstärke des zweiten Vierpols ist die Ausgangsspannung bzw. Ausgangsstromstärke des ersten), dann erhält man wieder einen Vierpol, seine Vierpolfunktion ist die Hintereinanderausführung $k \circ h$ der Funktionen h und k. Die Kettenmatrix des neuen Vierpols ist daher das Produkt $M \cdot L$ der Kettenmatrizen M und L.

Wie erhält man nun die Vierpolfunktion des folgenden Vierpols?



Dieser Vierpol entsteht durch Hinereinanderschaltung von vier Vierpolen, deren Vierpolfunktionen wir bereits kennen:



Die Vierpolfunktion davon ist $g \circ f \circ f \circ g$ (f und g wie oben), die entsprechende Kettenmatrix ist daher

$$K_P \cdot K_S \cdot K_S \cdot K_P = \begin{pmatrix} 3 & -2R \\ -\frac{4}{R} & 3 \end{pmatrix}.$$

§4. Fragen

1. Welche der folgenden Funktionen sind linear?
   (a) $f : \mathbb{R} \longrightarrow \mathbb{R},\, t \longmapsto 2t+1$
   (b) Wir betrachten die Ebene E nach Wahl eines Nullpunktes als Vektorraum. $g : E \longrightarrow E,\, P \longmapsto 2 \cdot P$
   (c) Wir wählen einen vom Nullpunkt verschiedenen Punkt $Q \in E$. $h : E \longrightarrow E,\, P \longmapsto P + Q$
   (d) $a : \mathcal{F}(\mathbb{R}, \mathbb{R}) \longrightarrow \mathbb{R},\, b \longmapsto b(2)$

2. Welche der folgenden Funktionen sind linear?
   (a)
$$\begin{array}{ccc} \mathbb{Q}^{2\times 1} & \longrightarrow & \mathbb{Q}^{2\times 1} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} x+y \\ 3x-4y \end{pmatrix} \end{array}$$
   (b)
$$\begin{array}{ccc} \mathbb{Q}^{2\times 1} & \longrightarrow & \mathbb{Q}^{2\times 1} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} 3x+2y-2 \\ x+4 \end{pmatrix} \end{array}$$
   (c)
$$\begin{array}{ccc} \mathbb{Q}^{2\times 1} & \longrightarrow & \mathbb{Q}^{2\times 1} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} x^2+2y \\ xy \end{pmatrix} \end{array}$$

In den folgenden drei Beispielen sind die lineare Funktion
$$\begin{array}{rcl} f : \mathbb{Q}^2 & \longrightarrow & \mathbb{Q}^2 \\ (x,y) & \longmapsto & (-2x+3y, x-y), \end{array}$$
eine Basis des Definitionsbereichs $\underline{v}$ und eine Basis des Bildbereichs $\underline{w}$ gegeben.

3. Sei $\underline{v} = ((1,-1),(1,1))$ und $\underline{w} = ((1,0),(0,1))$. Welche der folgenden Matrizen ist die Matrix von f bezüglich der Basen $\underline{v}$ und $\underline{w}$?
   (a) $\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}$
   (b) $\begin{pmatrix} -5 & 1 \\ 2 & 0 \end{pmatrix}$
   (c) $\begin{pmatrix} 7 & -10 \\ -3 & 4 \end{pmatrix}$
   (d) $\begin{pmatrix} 1 & -2 \\ 0 & -\frac{1}{2} \end{pmatrix}$

4. Sei $\underline{v} = ((1,0),(0,1))$ und $\underline{w} = ((1,0),(0,1))$. Welche der folgenden Matrizen ist die Matrix von f bezüglich der Basen $\underline{v}$ und $\underline{w}$?
   (a) $\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}$
   (b) $\begin{pmatrix} -5 & 1 \\ 2 & 0 \end{pmatrix}$
   (c) $\begin{pmatrix} 7 & -10 \\ -3 & 4 \end{pmatrix}$
   (d) $\begin{pmatrix} 1 & -2 \\ 0 & -\frac{1}{2} \end{pmatrix}$

5. Sei $\underline{v} = ((1,0),(0,1))$ und $\underline{w} = ((-2,1),(2,-2))$. Welche der folgenden Matrizen ist die Matrix von f bezüglich der Basen $\underline{v}$ und $\underline{w}$?
   (a) $\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}$
   (b) $\begin{pmatrix} -5 & 1 \\ 2 & 0 \end{pmatrix}$
   (c) $\begin{pmatrix} 7 & -10 \\ -3 & 4 \end{pmatrix}$
   (d) $\begin{pmatrix} 1 & -2 \\ 0 & -\frac{1}{2} \end{pmatrix}$

6. Seien K ein Körper, V,W Vektorräume über K, f eine lineare Funktion von V nach W und $\underline{v}$ bzw. $\underline{w}$ eine Basis von V bzw. W. Der Eintrag in der i-ten Zeile und j-ten Spalte der Matrix von f bezüglich $\underline{v}$ und $\underline{w}$ ist
   (a) die Koordinate bei w_j von $f(v_i)$
   (b) die Koordinate bei $f(v_i)$ von w_j
   (c) die Koordinate bei w_i von $f(v_j)$
   (d) die i-te Komponente des Bildes unter f des j-ten Standardbasisvektors.

7. Sei f,g lineare Funktionen von V nach V, $\underline{v}$ eine Basis von V und A bzw. B die Matrix von f bzw. g bezüglich $\underline{v}$.
   Welche der folgenden Aussagen sind wahr?
   (a) Die Matrix von $f \circ g$ bezüglich $\underline{v}$ ist $B \cdot A$.
   (b) Wenn f und g invertierbar sind, dann ist die Matrix von $(f \circ g)^{-1}$ bezüglich $\underline{v}$ gleich $B^{-1} \cdot A^{-1}$.

KAPITEL 6

Antworten

§1. Mengen, Funktionen, Zahlen und Rechenregeln

	1	2	3	4	5	6	7
a						X	
b	X		X				X
c	X	X	X				
d		X					

1. Siehe Definition 7 und Satz 11.

2. Siehe Definition 2, Definition 4 und §4.
   (a) Falsch, denn $B \cap C = \{\}$.
   (b) Die Aussage $2 \in A$ ist wahr. Die Aussage $2 \notin B$ ist falsch. Daher ist $(2 \in A) \wedge (2 \notin B)$ falsch.
   (c) Die Aussage $C \subseteq A$ ist wahr. Die Aussage $B \not\subseteq A$ ist falsch. Daher ist $(C \subseteq A) \vee (B \not\subseteq A)$ wahr.
   (d) Wahr, denn $B \cup C = \{1,2,3,5,6,7,9,10\} = A$.

3. (a) Diese Aussage ist nur dann wahr, wenn $m = 1$ ist. Wenn $m > 1$ ist, könnten zum Beispiel A_1 falsch und $A_2, \ldots, A_n$ wahr sein.
   (b) Wahr. Wenn A_1 falsch ist, ist die Aussage „wenn A_1 wahr ist, dann auch A_2" immer wahr. Siehe §4 von Kapitel 0.
   (c) Wahr. Wenn nicht alle Aussagen $A_1, \ldots, A_n$ wahr sind, dann muss eine der Voraussetzungen (1) oder (2) in Satz 15 falsch sein.

4. Siehe Satz 17.
   (a) Der ganzzahlige Quotient ist -2 und der Rest ist 2.
   (b) Der ganzzahlige Quotient ist -3 und der Rest ist 3.
   (c) Der ganzzahlige Quotient ist 3 und der Rest ist 3.

5. (a) $x = 111$ und $y = 11001$
   (b) $0,0100$ (siehe Satz 23).
   (c) $0,2800$ (siehe Satz 23).

6. (a) Wahr, siehe Definition 28.
   (b) Falsch. Zum Beispiel: in $(\mathbb{Z}, +, \cdot)$ ist $1 \cdot 0 = 2 \cdot 0$, aber $1 \neq 2$. Es gibt Ringe, in denen auch für $c \neq 0$ Elemente a, b existieren mit $a \neq b$ und $a \cdot c = b \cdot c$ (siehe Kapitel 1).

(c) Falsch. Siehe Definition 32.

7. Siehe Satz 42.
   (a) Diese Aussage ist nur dann wahr, wenn $n = 1$ ist.
   (b) Wahr, weil
   $$\sum_{i=1}^{n} a_i = \sum_{k=1}^{n} a_k$$
   und
   $$\sum_{j=1}^{n} b_j = \sum_{\ell=1}^{n} b_\ell$$
   ist.

§2. Matrizenrechnung

	1	2	3	4
a		X		X
b	X			X
c			X	
d				

1. (a) Nur die Addition von Matrizen und die Multiplikation einer Matrix mit einer Zahl erfolgen komponentenweise (siehe Definition 45), nicht aber die Multiplikation zweier Matrizen (siehe Definition 49).
   (b) Siehe Definition 49.
   (c) Siehe Definition 49.

2. (a) Die Matrizenmultiplikation ist im Allgemeinen nicht kommutativ (siehe Satz 55).
   (b) Siehe Satz 51.
   (c) Siehe Satz 54.

3. Siehe Definition 59 und Satz 60.

4. (a) Siehe Satz 61.
   (b) Siehe Satz 31.
   (c) Siehe Satz 31.

§3. Systeme linearer Gleichungen

	1	2	3	4	5	6	7	8	9	10	11	12
a	X				X			X				
b		X		X			X					
c		X	X	X		X						

	13	14	15	16	17	18	19	20	21
a									
b	X	X		X					X
c			X		X			X	
d						X			
e									

1. Siehe Definitionen 74 und 75.

2. Siehe Definition 75.

3. Siehe Definition 75.

4. (a) Es ist $(-4,-1,-7) = 2\cdot(1,1,1)+(-3)\cdot(2,1,3)$.
   (b) Wegen $2\cdot(2,1)-(4,2) = (0,0)$ sind die Vektoren $(2,1)$ und $(4,2)$ linear abhängig.
   (c) Nach Satz 88 hat das durch $\begin{pmatrix} 3 & 5 & 7 & 8 \\ 0 & 2 & -2 & 0 \\ 2 & 0 & 1 & 0 \end{pmatrix}$ gegebene System homogener linearer Gleichungen eine nicht-triviale Lösung.

5. $L(A,b) = \{\begin{pmatrix} -8 \\ -19 \\ 0 \end{pmatrix}\}$
   Siehe Satz 83.

6. $L(A,b) = \{\begin{pmatrix} \frac{1}{3} \\ -2 \\ 0 \end{pmatrix} + t\cdot\begin{pmatrix} -\frac{2}{3} \\ 0 \\ 1 \end{pmatrix} \mid \quad t \in \mathbb{Q}\}$
   Siehe Satz 83.

7. $L(A,b) = \{\}$
   Siehe Satz 83.

8. (a) Falsch, denn zum Beispiel hat $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot x = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ unendlich viele Lösungen in $\mathbb{Q}^{2\times 1}$.

   (b) Richtig, denn es gibt für ein homogenes System linearer Gleichungen immer zumindest eine Lösung, und zwar die Nullspalte.

   (c) Siehe Sätze 83 und 87.

9. $L(A,b) = \{ \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mid \quad t \in \mathbb{Q} \}$
   Siehe Sätze 86 und 87.

10. $L(A,b) = \{ \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \}$
    Siehe Sätze 86 und 87.

11. $\begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & 0 \\ -\frac{1}{3} & -\frac{1}{3} & 1 \\ \frac{2}{3} & -\frac{1}{3} & 0 \end{pmatrix}$
    Siehe Satz 89.

12. Sind $P_1, \ldots, P_k$ Elementarmatrizen $(P_1(P_2(\ldots(P_kA)\ldots))) = I_n$, dann ist $P_1, \ldots P_k$ die inverse Matrix von A. Aus
    $$(P_1(P_2(\ldots(P_kI_n)\ldots))) = P_1, \ldots P_k = A^{-1}$$
    folgt die Behauptung.

13. (a) Falsch, denn eine Menge von Pfeilen allein bildet keinen Vektorraum. Erst durch die Rechenoperationen Addition und Skalarmultiplikation wird ein Vektorraum definiert.
    (b) Richtig, siehe Satz 73.
    (c) Falsch. Die Aussage gilt nicht für alle Systeme linearer Gleichungen, sondern nur für homogene. Denn nur bei homogenen Systemen linearer Gleichungen ist die Summe zweier Lösungen wieder eine Lösung und jedes Vielfache einer Lösung auch wieder eine Lösung.

14. Siehe Satz 94.

15. (c) ist wahr. „Zum Anschreiben einer solchen Matrix können genau 15 Zahlen frei gewählt werden.“

16. (a) trifft nur zu, wenn die Spalten linear unabhängig sind.
    (b) Richtig, siehe Definition 100.

17. Der Rang der Matrix ist 3.

18. Der Rang der Matrix ist 2.

19. Die Koordinaten des Vektors bezüglich der gegebenen Basis sind $(4,1,2)$. Um sie zu berechnen, muss das durch $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 1 & 3 \end{pmatrix}$ und $\begin{pmatrix} 7 \\ 10 \\ 11 \end{pmatrix}$ gegebene System linearer Gleichungen gelöst werden.

20. Satz 103 beschreibt ein Verfahren zur Auswahl einer Basis aus einem EZS.
    (a) Aus der Stufenform können Sie ablesen, welche Dimension dieser Vektorraum hat. Sie können aber nicht daraus schließen, dass dann die ersten d Spalten linear unabhängig und somit eine Basis sind.
    (b) Falsch.
    (c) Richtig.

21. Satz 103 beschreibt ein Verfahren, mit dem linear unabhängige Vektoren zu einer Basis ergänzt werden können.
    (a) Falsch. Jede der Standardspalten $e_{k+1},\ldots,e_n$ könnte von den Spalten $s_1,\ldots,s_k$ linear abhängig sein.
    (b) Richtig.
    (c) Falsch.

§4. Permutationen, Determinanten und Eigenwerte

	1	2	3	4	5	6	7	8	9
a	X	X			X				
b			X	X		X	X		X
c				X				X	X
d		X	X						

1. (a) Richtig, siehe Definition 153.
   (b) Falsch.
   Die Umkehrfunktion ist $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}$.

2. (a) Richtig, da 1 auf 1 abgebildet wird. Siehe Definition 156.
   (b) Falsch. Die Permutation besitzt nur zwei Zykel und zwar $(2,5,6,8)$ und $(3,7,4)$.
   (c) Falsch, das Vorzeichen von s ist -1. Siehe Definition 158.
   (d) Richtig, da das Vorzeichen von s gleich -1 ist. Siehe Definition 158.

3. (b) und (d) sind wahr, siehe Definition 170 und Satz 174.

4. (a) Falsch, siehe Satz 174.
   (b) Richtig, siehe Satz 178.
   (c) Richtig, weil die ersten zwei Spalten gleich sind (siehe Satz 180).

5. (a) Richtig. Siehe Satz 182.
   (b) Falsch. Siehe Definition 170.

6. Siehe Satz 182 (1).

7. Siehe Satz 182 (1).

8. (a) Falsch, siehe Satz 181.
   (b) Falsch.
   (c) Richtig, siehe Satz 181.

9. (a) Falsch. Die Eigenwerte der Matrix sind 0 und 2.
   Denn es ist $\det(cI_3 - A) = c^3 - 4c^2 + 4c = c(c-2)^2$.
   Also ist $\det(cI_3 - A)$ genau dann 0, wenn c gleich 0 oder 2 ist.
   (b) Richtig, siehe Satz 196.
   (c) Richtig, denn $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & -2 \\ -1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.
   (d) Falsch. Der Eigenraum zum Eigenwert 2 ist der Lösungsmenge des homogenen Systems linearer Gleichungen
   $(2I_n - A)x = 0$.

§5. Lineare Funktionen

	1	2	3	4	5	6	7
a		X		X			
b	X		X				X
c						X	
d	X				X		

1. (a) Nicht linear, zum Beispiel weil $f(0) \neq 0$ ist.
   (b) Linear (nachprüfen)
   (c) Nicht linear, zum Beispiel weil $h(0) \neq 0$ ist.
   (d) Linear (nachprüfen)

2. (a) Die Funktion ist nach Definition 205 linear.
   Ein anderes Argument: Wegen

   $$\begin{pmatrix} x+y \\ 3x-4y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

   ist die Funktion durch die die Multiplikation der Spalten in $\mathbb{Q}^{2\times 1}$ mit der Matrix $\begin{pmatrix} 1 & 1 \\ 3 & -4 \end{pmatrix}$ gegeben, also linear.

   (b) und (c) erfüllen die Bedingungen aus Definition 205 nicht und sind somit auch nicht linear.

3. Siehe Definition 215.

4. Siehe Definition 215.

5. Siehe Definition 215.

6. Siehe Definition 215.

7. (a) Falsch, die Matrix von $f \circ g$ ist $A \cdot B$, siehe Satz 219.
   (b) Richtig, siehe Satz 220.