Lineare Algebra

Claus Diem

Universität Leipzig
WS 2009/10 und SS 2010

Kapitel 1

Grundlegende Strukturen

1.1 Vorbemerkungen

Ziel dieses Abschnitts ist, einige logische Grundlagen zu klären.

Wir beginnen mit einigen mathematischen Aussagen $A, B, C, \ldots$. Zum Beispiel könnten dies diese Aussagen sein:

- 2 *ist gerade.*
- *Jede durch 4 teilbare natürliche Zahl ist durch 2 teilbar.*
- *Jede durch 2 teilbare natürliche Zahl ist durch 4 teilbar.*
- *Für je drei ganze Zahlen* x, y, z *mit* $x^3 + y^3 = z^3$ *gilt:* $x \cdot y \cdot z = 0$.
- *Für je drei ganze Zahlen* x, y, z *und jede natürliche Zahl* $n \geq 3$ *mit* $x^n + y^n = z^n$ *gilt:* $x \cdot y \cdot z = 0$.
- *Jede gerade natürliche Zahl* ≥ 4 *ist eine Summe von zwei Primzahlen.*

All dies sind sinnvolle mathematische Aussagen, die entweder wahr oder falsch sind.

Wir können nun diese oder andere sinnvolle (wahre oder falsche) Aussagen verwenden, um komplexere Aussagen zu betrachten. Ein Beispiel ist

$$A \textit{ und } B\,,$$

was man mit

$$A \wedge B$$

abkürzt. Dies ist eine Aussage, die genau dann wahr ist, wenn sowohl A als auch B wahr sind. Dies kann man mittels einer *Wahrheitstabelle* ausdrücken.

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

Eine oft benutzte Form ist auch

$A \wedge B$	w	f
w	w	f
f	f	f

.

Andere "Operatoren", die wir auf Aussagen anwenden können, sind *nicht*, *oder* und *entweder oder*, mit den folgenden offensichtlichen Wahrheitstabellen. (Das Wort *oder* wird in der Mathematik immer als *und/oder* benutzt.)

$\neg A$	w	f
	f	w

$A \vee B$	w	f
w	w	w
f	w	f

$A \dot{\vee} B$	w	f
w	f	w
f	w	f

Implikationen

Wir betrachten nun den Aussagesatz A *impliziert* B, abgekürzt $A \to B$. Wir legen fest, dass dieser Aussagesatz die folgende Bedeutung hat: *Es ist ausgeschlossen, dass A richtig ist und B falsch.* Hiermit ist $A \longrightarrow B$ wieder eine mathematische Aussage, welche eine Umformulierung der Aussage $\neg(A \wedge \neg B)$ ist.

Wiederum können wir in Abhängigkeit von A und B betrachten, ob die Aussage wahr oder falsch ist. Wir haben die folgende Wahrheitstabelle:

$A \to B$	w	f
w	w	f
f	w	w

Insbesondere ist $A \to B$ automatisch wahr, wenn A falsch ist.

Wir können nun die fünfte Beispiel-Aussage oben wie folgt umformulieren:
Für je drei ganze Zahlen x, y, z und jede natürliche Zahl $n \geq 3$ gilt:

$$x^n + y^n = z^n \longrightarrow x \cdot y \cdot z = 0$$

Die Aussage A *impliziert* B wird auch mit

Wenn A [gilt], dann [gilt] B

umschrieben. Z.B.

Für je drei ganze Zahlen x, y, z und jede natürliche Zahl $n \geq 3$ gilt:

$$\textit{Wenn } x^n + y^n = z^n, \textit{ dann gilt } x \cdot y \cdot z = 0 \ .$$

Ich möchte darauf hinweisen, dass die Bedeutung von A *impliziert* B bzw. *Wenn* A, *dann [gilt]* B nicht unbedingt dem allgemeinen Sprachgebrauch entspricht. Insbesondere könnte man geneigt sein, Aussagen der Form A *impliziert* B weder als wahr oder falsch sondern einfach als *unsinnig* zu betrachten, falls es keinen (offensichtlichen) engen Zusammenhang zwischen A und B gibt.

Mögliche Beispiele hierfür sind die folgenden beiden wahren Aussagen über ganze Zahlen:

$$3 > 4 \longrightarrow 100 < 0$$

$$3 = 2 + 1 \longrightarrow 3^2 + 4^2 = 5^2$$

Ich erwähne noch den Operator *genau dann wenn*, der durch die folgende Wahrheitstabelle definiert ist.

$A \leftrightarrow B$	w	f
w	w	f
f	f	w

Die Aussage $A \longleftrightarrow B$ liest man auch so: A *ist äquivalent zu* B.

Komplexere Zusammensetzungen

Die Operatoren *und*, *oder*, ... kann man selbstverständlich mehrfach anwenden. Man sollte Klammern setzen, um die Interpretation eines Aussagesatzes genau festzulegen.

Einige Beispiele:

Seien A, B, C drei (sinnvolle) mathematische Aussagen. Dann sind die folgenden beiden Aussagen äquivalent:

- $\neg(A \wedge B)$
- $\neg A \vee \neg B$

genauso:

- $\neg(A \vee B)$
- $\neg A \wedge \neg B$

(Diese beiden Äquivalenzen sind unter dem Namen *De Morgan'sche Gesetze* bekannt.)

Es sind auch äquivalent:

- $A \to B$
- $\neg(A \wedge \neg B)$
- $\neg A \vee B$
- $\neg B \to \neg A$

sowie:

- $A \wedge (B \vee C)$
- $(A \wedge B) \vee (A \wedge C)$

sowie:

- $A \vee (B \wedge C)$
- $(A \vee B) \wedge (A \vee C)$

und:

- $A \to (B \to C)$
- $(A \wedge B) \to C$

Dies kann man z.B. leicht mittels Wahrheitstabellen einsehen.

Beweisschemata

Nehmen wir an, wir wollen beweisen dass B wahr ist. Falls wir wissen, dass A wahr ist und $A \to B$ wahr ist, folgt dass B wahr ist. Dies kann man formal so beschreiben:

$$\frac{\begin{array}{l} A \\ A \to B \end{array}}{B}$$

Dies ist ein Beispiel eines *direkten Beweises*.

Wir nehmen nun an, dass wir wissen, dass A wahr ist und $\neg B \to \neg A$ gilt. In diesem Fall können wir auch schließen, dass B wahr ist.

$$\frac{\begin{array}{l} A \\ \neg B \to \neg A \end{array}}{B}$$

Dies ist ein Beispiel eines *Beweises durch Widerspruch*. Das ist natürlich sehr eng mit dem vorherigen Beweisschema verwandt, denn $\neg B \to \neg A$ ist ja äquivalent zu $A \longrightarrow B$. Man kann auch einen Widerspruch herbeiführen, indem man aus einer Aussage ihr Gegenteil ableitet. Dann muss die Aussage offensichtlich falsch sein.

$$\frac{A \longrightarrow \neg A}{\neg A}$$

Noch einige Bemerkungen:

- Statt $A \wedge B$ schreibt man oft A, B.
- Statt des Implikationspfeils $\to$ wird oft ein Doppelpfeil $\Longrightarrow$ geschrieben.
- Ein Beweis der Form *Es gilt A, und es gilt $A \to B$, folglich gilt also auch B.* wird oft (insbesondere in Vorlesungen) in der Form *Es gilt A.* $\Longrightarrow B$ abgekürzt.

- Die Aussagen $A \leftrightarrow B$ und $B \leftrightarrow C$ (und folglich auch $A \leftrightarrow C$) werden oft zu $A \leftrightarrow B \leftrightarrow C$ zusammengefasst (und statt $\leftrightarrow$ wird meist $\Longleftrightarrow$ geschrieben).

Es sei noch angemerkt, dass Aussagesätze je nach Kontext sinnvolle mathematische Aussagen ergeben können oder auch nicht. Es kann auch sein, dass ein Satz je nach Kontext entweder sinnlos, wahr oder falsch ist.

Betrachten Sie als Beispiel den folgenden Aussagesatz:

x ist gerade.

Dieser Satz ist so nicht interpretierbar, weil nicht klar ist, was x ist. Wenn wir vorher x als die ganze Zahl 2 definieren (d.h. x und 2 bezeichnen nun dasselbe mathematische Objekt), erhalten wir eine wahre Aussage, wenn wir vorher x als 3 definieren, erhalten wir eine falsche Aussage. Wenn wir aber x als die rationale Zahl 3/2 definieren, erhalten wir wieder einen nicht-interpretierbaren Satz (weil die Eigenschaft *gerade* nur für ganze Zahlen definiert ist).

1.2 Mengen

Was ist eine Menge? Wir stellen uns auf den folgenden Standpunkt Georg Cantors (1845-1918), der als Begründer der Mengenlehre gelten kann:

Eine Menge ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.

Aus der Schule kennen Sie die Mengen der *natürlichen Zahlen*, der *ganzen Zahlen*, der *rationalen Zahlen* oder der *reellen Zahlen* (wenn auch auf intuitiven Niveau). Die Bezeichnungen sind $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$. Wir legen hier fest, dass 0 keine natürliche Zahl ist und definieren $\mathbb{N}_0$ als die Menge der ganzen Zahlen ≥ 0.

Ich setze im Folgenden die natürlichen, die ganzen und die rationalen Zahlen als offensichtlich voraus. Man kann aber auch hierzu noch einiges sagen. Die Frage, was genau die reellen Zahlen sind, wird in der Analysis Vorlesung behandelt und ist für diese Vorlesung nicht so wichtig. Mengen werden oft durch (möglicherweise unvollständige) Aufzählung ihrer Elemente beschrieben. Beispiele sind

- $\emptyset = \{\}$, die leere Menge
- $\mathbb{N} = \{1, 2, 3, \ldots\}$
- $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$

- $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

Ein anderes Beispiel ist: Sei $a \in \mathbb{N}$, sei $X := \{1, 2, \ldots, a\}$. (D.h. X ist definiert als die Menge $\{1, 2, \ldots, a\}$.) Nun ist X eine Menge mit a Elementen.

Die Schreibweise $X = \{x_1, \ldots, x_n\}$ bedeutet hingegen nicht, dass alle x_i verschieden sind; X hat *höchstens* n aber nicht notwendigerweise genau n Elemente.

Beispielsweise bedeutet die Aussage $X = \{a, b, c\}$, dass X eine Menge mit höchstens drei Elementen ist, welche mit a, b und c bezeichnet werden.

Teilmengen

Definition Seien A und X Mengen. Wenn nun jedes Element von A in X enthalten ist, nennen wir A eine *Teilmenge* von X, und wir nennen X eine *Obermenge* von A. Wir sagen dann auch, dass X die Menge A *umfasst*. Wir schreiben dann $A \subseteq X$. Wenn A eine Teilmenge von X ist und es ein Element von X gibt, das nicht in A enthalten ist, nennen wir A eine *echte Teilmenge* und schreiben $A \subsetneq X$.

Bemerkung In Analogie zu den Relationen "kleiner-gleich" und "kleiner" für Zahlen wäre es folgerichtig, statt $A \subsetneq X$ einfach nur $Y \subset X$ zu schreiben. Allerdings schreiben viele Autoren $A \subset X$, wenn sie $A \subseteq X$ meinen. Ich vermeide die Bezeichnung $A \subset X$ ganz.

Definition Seien A und B Mengen. Dann besteht die *Vereinigung* $A \cup B$ aus allen Elementen, in A oder in B vorkommen. Der *Durchschnitt* $A \cap B$ besteht aus allen Elementen, die sowohl in A als auch in B vorkommen. Die *Differenz* $A \backslash B$ oder $A - B$ besteht aus genau den Elementen aus A, die nicht in B vorkommen.[1]

Definition Sei $A \subseteq X$ eine Teilmenge. Dann ist

$$A^c := \{x \in X \mid x \notin A\}$$

das *Komplement* von A in X.

Definition Seien A und B zwei Teilmengen von X. Dann ist X eine *disjunkte Vereinigung* von A und B, wenn gilt:

$$\text{Für alle } x \in X: \; x \in A \;\dot\vee\; x \in B \; .$$

[1] Ich schreibe immer $A - B$, aber die Schreibweise $A \backslash B$ ist wohl üblicher.

In diesem Fall schreiben wir

$$X = A \dot\cup B \ .$$

Für $A \subseteq X$ gilt also offensichtlich $A \dot\cup A^c = X$. Die folgenden Aussagen für Teilmengen A und B von X folgen sofort aus den De Morgan'schen Gesetzen für Aussagen:

$$(A \cap B)^c = A^c \cup B^c \qquad (A \cup B)^c = A^c \cap B^c \ .$$

Auch die folgende Aussage ist offensichtlich:

$$A \subseteq B \longleftrightarrow B^c \subseteq A^c \ .$$

Quantoren

Mittels Mengen kann man mathematische Aussagen wesentlich einfacher formulieren. Beispielsweise ist das Folgende eine Umformulierung der fünften Aussage zu Beginn.

Für alle $x, y, z \in \mathbb{Z}$, *für alle* $n \in \mathbb{N} : (n \geq 3 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0$.

Das kann man elegant mittels des All-Quantors $\forall$ aufschreiben:[2]

$$\forall x, y, z \in \mathbb{Z}, \forall n \in \mathbb{N} : (n \geq 3 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0 \ .$$

Diese Aussage ist (nach Auskunft der Experten) wahr. (Aber der Beweis ist sehr lang und schwierig.) Anderseits ist die folgende Aussage falsch:

$$\forall x, y, z \in \mathbb{Z}, \forall n \in \mathbb{N} : (n \geq 2 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0 \ .$$

(Ein Gegenbeispiel ist $x = 3, y = 4, z = 5, n = 2$.) Die Existenz so eines Gegenbeispiels man man mittels des Existenz-Quantors $\exists$ ausdrücken:

$$\exists x, y, z \in \mathbb{Z}, \exists n \in \mathbb{N} : \neg((n \geq 2 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0)$$

bzw.

$$\exists x, y, z \in \mathbb{Z}, \exists n \in \mathbb{N} : (n \geq 2 \wedge x^n + y^n = z^n) \wedge x \cdot y \cdot z \neq 0 \ .$$

(Die Klammern kann man weglassen.)

[2]Den All-Quantor schreibe ich mit $\forall$, den Existenzquantor mit $\exists$. Eine ältere Schreibweise ist $\bigwedge$ für den All-Quantor und $\bigvee$ für den Existenzquantor. Da letzteres leicht mit dem All-Quantor $\forall$ verwechselt werden kann, bitte ich Sie, diese Schreibweise nicht zu benutzen.

Paradoxien und Axiome

Wenn man den Mengenbegriff allzu generös auslegt, stößt man leicht auf Paradoxien. Ein Beispiel ist die folgende sogenannte *Russelsche Antinomie*:

Sei $\mathcal{M}$ die Menge aller Mengen, die nicht ein Element von sich selbst sind. D.h.

$$\mathcal{M} := \{M \text{ Menge} \mid M \notin M\}$$

Für jede Menge M gilt also

$$M \in \mathcal{M} \longleftrightarrow M \notin M \ .$$

Somit gilt insbesondere $\mathcal{M} \in \mathcal{M} \longleftrightarrow \mathcal{M} \notin \mathcal{M}$. Das ist offensichtlich absurd.

Aufgrund solcher Paradoxien (das Fachwort ist *Antinomien*), sollte man aufpassen, wenn man Mengen bildet. Als Regeln (Axiome) sollte man sich merken:

- Jedes Objekt ist eine Menge. ("Es gibt nichts außer Mengen.")[3]
- Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente enthalten.
- Es gibt (genau) eine Menge ohne Elemente:[4] die leere Menge $\emptyset$
- Für zwei Objekte (Mengen) a, b gibt es (genau) eine Menge, die genau a und b enthält: $\{a, b\}$.
- Zu jeder Menge X gibt es die *Potenzmenge* $\mathcal{P}(X)$. Die Elemente von $\mathcal{P}(X)$ sind genau die Teilmengen von X.
- Man kann Vereinigungen von Mengen bilden.
- Wenn X eine Menge ist und E eine Eigenschaft,[5] die für Elemente aus X definiert ist (und jeweils wahr oder falsch sein kann), dann gibt es die Teilmenge

  $$\{x \in X \mid E \text{ trifft auf } x \ zu \ \}$$

  von X.
- Es gibt die Menge der natürlichen Zahlen.

[3] Diese Aussage macht natürlich nur innerhalb der Mathematik Sinn. Und auch innerhalb der Mathematik ist das noch nicht das letzte Wort ...

[4] Die Eindeutigkeit folgt aus der 2. Regel

[5] Was ist eine "Eigenschaft"? Eine schwierige Frage ...

Man setzt in der Regel ein paar weitere Axiome voraus. Auf jeden Fall wird noch ein Axiom hinzukommen, das ich später erwähne.

Wenn Sie das genauer interessiert, schauen Sie mal unter *Zermelo-Fraenkel-Mengenlehre* nach!

Tupel

Seien X und Y Mengen. Dann besteht die Menge $X \times Y$ aus *geordneten Paaren* (x, y) von Elementen $x \in X, y \in Y$. Solche geordneten Paare heißen auch *Tupel* (oder *Zweiertupel*).

Die Menge $X \times Y$ heißt *kartesisches Produkt* von X und Y. Etwas allgemeiner erhält man zu Mengen $X_1, \ldots, X_n$ das kartesische Produkt $X_1 \times \cdots \times X_n$, das aus den so genannten n-Tupeln $(x_1, \ldots, x_n)$ mit $x_i \in X_i$ besteht. Wenn $X_1 = X_2 = \cdots = X_n = X$, schreibt man auch X^n für das kartesische Produkt.

Das Prinzip der vollständigen Induktion

Das *Prinzip der vollständigen Induktion* ist eine Beweismethode für Aussagen, die für alle natürlichen Zahlen gelten. Es handelt sich also um Aussagen der Form

$$\forall n \in \mathbb{N} : A(n) \ ,$$

wobei für $n \in \mathbb{N}$ $A(n)$ eine Aussage über die Zahl n ist.

Aussagen dieser Form kann man wie folgt beweisen:

1. Man beweist, dass $A(1)$ gilt.
2. Man beweist für alle $n \in \mathbb{N}$, dass die Implikation $A(n) \longrightarrow A(n+1)$ gilt.

Wenn nun $n \in \mathbb{N}$ beliebig ist, haben wir die folgenden (bewiesenen) Aussagen:

$$A(1), A(1) \longrightarrow A(2), A(2) \longrightarrow A(3), \ldots, A(n-1) \longrightarrow A(n) \ .$$

Hieraus folgt $A(n)$.

Es gibt zwei oft benutzte Varianten der vollständigen Induktion:

- Man beginnt die Induktion nicht bei 1 sondern bei $n_0 \in \mathbb{Z}$ und zeigt $\forall n \in \mathbb{Z}, n \geq n_0 : A(n)$.
- Man setzt im Induktionsschritt nicht nur $A(n)$ sondern alle vorherigen Aussagen voraus. D.h. man zeigt $A(n_0) \wedge A(n_0+1) \wedge \cdots \wedge A(n) \longrightarrow A(n+1)$.

Wir geben zwei Beispiele für Beweise mittels "vollständiger Induktion".

Beispiel 1.1 Wir wollen beweisen:

$$\forall n \in \mathbb{N} : \ 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \ .$$

Die Aussage $A(n)$ ist also $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.

Beweis von $A(1)$
Es ist $1 = \frac{1 \cdot 2}{2}$.

Beweis der Implikation $A(n) \longrightarrow A(n+1)$
Es gelte $A(n)$, also $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.

Dann ist $1 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+2)(n+1)}{2}$. □

Beispiel 1.2 Wir wollen beweisen:

Für alle $n \in \mathbb{N}, m \in \mathbb{N}_0$ gibt es eindeutig bestimmte $p \in \mathbb{N}_0$ und $r \in \{0, \ldots, n-1\}$ mit

$$m = pn + r \ .$$

("Division mit Rest")

Hierzu *fixieren* wir $n \in \mathbb{N}$ und betrachten die folgende Aussage:[6]

$$\forall m \in \mathbb{N}_0 \ : \ \exists! \, (p, r) \in \mathbb{N}_0 \times \{0, \ldots, n-1\} : m = pn + r$$

Wir zeigen dies nun mittels vollständiger Induktion nach m.

Beweis von $A(0)$
Es ist $0 = 0 \cdot n + 0$, und dies ist offensichtlich die einzige Möglichkeit, 0 in der Form $pn + r$ mit $p \in \mathbb{N}_0, r \in \{0, \ldots, n-1\}$ zu schreiben.

Beweis der Implikation $A(m) \longrightarrow A(m+1)$
Wir setzen voraus: Es gibt eindeutig bestimmte $p \in \mathbb{N}_0$ und $r \in \{0, \ldots, n-1\}$ mit $m = pn + r$.

Zuerst zur *Existenz* der Darstellung von $m+1$. Seien $p_0 \in \mathbb{N}_0, r_0 \in \{0, \ldots, n-1\}$ mit

$$m = p_0 n + r_0 \ .$$

Dann ist also $m+1 = p_0 n + r + 1$. Es gibt nun zwei Fälle: Wenn $r < n-1$, dann ist $m + 1 = p_0 n + (r+1)$ eine Darstellung wie gewünscht. Wenn andererseits $r = n - 1$, dann ist $m + 1 = (p_0 + 1)n + 0$ eine Darstellung wie gewünscht.

[6] Das Ausrufezeichen hinter "$\exists$" deutet an, dass es genau ein Element mit der angegebenen Eigenschaft gibt.

Nun zur *Eindeutigkeit*. Seien $m+1 = p_1 n + r_1$ und $m+1 = p_2 n + r_2$ zwei Darstellungen mit $p_1, p_2 \in \mathbb{N}_0, r_1, r_2 \in \{0, \ldots, n-1\}$.

Wir machen eine Fallunterscheidung in vier Fälle.

$\underline{r_1 \geq 1, r_2 \geq 1}$
In diesem Fall ist $m = p_1 n + (r_1 - 1)$ und $m = p_2 n + (r_2 - 1)$ mit $r_1 - 1, r_2 - 1 \in \{0, \ldots, n-1\}$. Damit ist nach der Eindeutigkeit der Darstellung von m $p_1 = p_2$ und $r_1 = r_2$.

$\underline{r_1 = 0, r_2 \geq 1}$
In diesem Fall ist $m = (p_1 - 1)n + (n-1)$ und $m = p_2 n + (r_2 - 1)$ mit $p_1 - 1, p_2 \in \mathbb{N}_0, r_1 - 1 \in \{0, \ldots, n-1\}$. Nach der Eindeutigkeit der Darstellung von m kann dieser Fall nicht auftreten.

$\underline{r_1 \geq 1, r_2 = 0}$
Analog zum zweiten Fall kann dieser Fall nicht auftreten.

$\underline{r_1 = 0, r_2 = 0}$
In diesem Fall ist $m = (p_1 - 1)n + (n-1)$ und $m = (p_2 - 1)n + (n-1)$ mit mit $p_1 - 1, p_2 - 1 \in \mathbb{N}_0$. Damit ist nach der Eindeutigkeit der Darstellung von m $p_1 = p_2$. □

1.3 Abbildungen

Seien X, Y Mengen. Intuitiv ist eine *Abbildung* von X nach Y eine Vorschrift, die jedem Element aus X in eindeutiger Weise ein Element aus Y zuordnet. Wenn f eine Abbildung von X nach Y ist, schreibt man $f : X \longrightarrow Y$, X heißt dann *Definitionsbereich* und Y heißt *Bildbereich* oder *Wertebereich*.

Zwei Abbildungen $f : X \longrightarrow Y, g : X \longrightarrow Y$ sind genau dann gleich, wenn für alle $x \in X$ $f(x) = g(x)$ ist.

Ein Beispiel ist

$$f : \mathbb{Z} \longrightarrow \mathbb{Z}\ ,\ x \mapsto 2x\ .$$

Hier gilt also $f(x) = 2x$ für alle $x \in \mathbb{Z}$.

Wenn allgemein $f : X \longrightarrow Y$ eine Abbildung ist, schreibt man $f(x)$ für den *Wert* von f an x, d.h. für dasjenige Element aus Y, welches x zugeordnet ist, bzw. auf welches x abgebildet wird.

Aus der Schule kennen Sie den Begriff der *Funktion*. "Funktion" und "Abbildung" kann man synonym benutzen, allerdings spricht man in der Regel eher dann von Funktionen, wenn der Wertebereich aus Zahlen besteht.

Die Abbildungen von X nach Y bilden auch eine Menge. Wir definieren:

Definition Die Menge der Abbildungen von X nach Y wird mit $\mathrm{Abb}(X, Y)$ bezeichnet.

Einige Beispiele:

Beispiel 1.3 Sei $X = \{a, b\}$, $Y = \{c, d\}$ mit $a \neq b$ und $c \neq d$. Dann besteht $\mathrm{Abb}(X, Y)$ aus den folgenden vier Elementen: $(a \mapsto c, b \mapsto c), (a \mapsto c, b \mapsto d), (a \mapsto d, b \mapsto c), (a \mapsto d, b \mapsto d)$.

Beispiel 1.4 Sei $X = \{x\}$. Dann besteht $\mathrm{Abb}(X, Y)$ genau aus den Abbildungen $x \mapsto a$ mit $a \in Y$.

Beispiel 1.5 Sei andererseits $Y = \{y\}$. Dann besteht $\mathrm{Abb}(X, Y)$ aus genau einem Element.

Eine kleine Spitzfindigkeit ist das folgende Beispiel:

Beispiel 1.6 Sei Y eine beliebige Menge. Natürlich können wir dann *jedem* Element der *leeren Menge* ein Element von Y zuordnen. Somit besteht $\mathrm{Abb}(\emptyset, Y)$ aus genau einem Element. Andererseits ist $\mathrm{Abb}(X, \emptyset)$ leer, wenn $X \neq \emptyset$.

Definition Die *identische Abbildung* id_X auf einer Menge X ist durch $x \mapsto x$ gegeben.

Definition Eine Abbildung $f : X \longrightarrow Y$ heißt

- *injektiv*, wenn für alle $x, x' \in X$ gilt: $f(x) = f(x') \longrightarrow x = x'$
- *surjektiv*, wenn für alle $y \in Y$ gilt: $\exists x \in X : f(x) = y$
- *bijektiv*, wenn sie injektiv und surjektiv ist.

Notation Wenn $f : X \longrightarrow Y$ injektiv ist, schreibt man auch $X \hookrightarrow Y$. Wenn $f : X \longrightarrow Y$ surjektiv ist, schreibt man auch $X \twoheadrightarrow Y$.

Einige offensichtliche Bemerkungen:

- f ist genau dann injektiv, wenn für alle $x, x' \in X$ gilt: $x \neq x' \longrightarrow f(x) \neq f(x')$.
- f ist genau dann bijektiv, wenn es zu jedem $y \in Y$ genau ein $x \in X$ mit $f(x) = y$ gibt. Dies kann man auch so beschreiben: $\forall y \in Y \exists! x \in X : f(x) = y$.

- Wenn f bijektiv ist, dann kann man wie folgt eine Abbildung $g : Y \longrightarrow X$ definieren: Jedem $y \in Y$ wird das eindeutig bestimmte $x \in X$ mit $f(x) = y$ zugeordnet. Diese Abbildung erfüllt $g \circ f = \mathrm{id}_X, f \circ g = \mathrm{id}_Y$. Sie heißt die *Umkehrabbildung* zu f und wird mit $f^{-1} : Y \longrightarrow X$ bezeichnet.

Ich erinnere noch daran, dass man Abbildungen verknüpfen kann: Gegeben zwei Abbildungen $f : X \longrightarrow Y, g : Y \longrightarrow Z$, hat man die Abbildung

$$g \circ f : X \longrightarrow Z,\ x \mapsto g(f(x)) \ .$$

Diese Definition kann man auch so ausdrücken: "Das Diagramm

$$\begin{array}{ccc} X & & \\ {\scriptstyle f}\downarrow & \searrow^{g\circ f} & \\ Y & \xrightarrow[g]{} & Z \end{array}$$

kommutiert."

Allgemein ist ein *kommutatives Diagramm* ein Diagramm von Mengen und Abbildungen, so dass gilt: Wenn immer man von einer Menge zu einer anderen "auf mehreren Wegen gehen kann", erhält man dieselbe Abbildung.

Beispiel 1.7 Seien X, Y, Z, W Mengen, und sei $f : X \longrightarrow Y, g : Y \longrightarrow W, h : X \longrightarrow Z, k : Z \longrightarrow W$. Die Aussage, dass das Diagramm

$$\begin{array}{ccc} X & \xrightarrow{h} & Z \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle k} \\ Y & \xrightarrow[g]{} & W \end{array}$$

kommutiert, heißt, dass $g \circ f = k \circ h : X \longrightarrow W$.

Definition Sei $U \subseteq X$ eine Teilmenge. Durch "Einschränkung" erhalten wir dann eine Abbildung

$$f|_U : U \longrightarrow Y\ ,\ x \mapsto f(x) \ .$$

Die Menge

$$f(U) := \{y \in Y \mid \exists x \in U : f(x) = y\}$$

heißt *Bild* von U unter f. Die Menge $f(X)$ wird auch mit $\mathrm{Bild}(f)$ bezeichnet. Es ist also $f(U) = \mathrm{Bild}(f|_U)$.

Eine weniger formale aber oft vorkommende Beschreibung ist

$$f(U) = \{f(x) \mid x \in U\} \ .$$

Bemerkung Beachten Sie den Unterschied zwischen dem *Bildbereich* (=Wertebereich) von f und dem *Bild* von f!

Definition Sei nun $V \subseteq Y$ eine Teilmenge. Die Menge

$$f^{-1}(V) := \{x \in X \mid f(x) \in V\}$$

heißt die *Urbildmenge* von V unter f.

Familien

Wir diskutieren noch eine weitere *Sichtweise* auf Abbildungen.

Seien zwei beliebige Mengen X und I gegeben.

Für jedes $i \in I$ sei genau ein Element aus X gegeben, welches wir mit x_i bezeichnen. Wir erhalten somit eine *Familie* von Elementen von X, die wir mit $(x_i)_{i\in I}$ bezeichnen können. Die Menge I heißt hierbei auch *Indexmenge*. Wenn $I = \{1, \ldots, n\}$, erhalten wir so die n-Tupel von Elementen aus X, d.h. die Elemente $(x_1, \ldots, x_n) \in X^n$.

So eine Familie $(x_i)_{i\in I}$ ist nichts weiter als eine Abbildung (d.h. Zuordnung): Jedem $i \in I$ wird genau das Element x_i zugeordnet, oder formaler: Die Familie $(x_i)_{i\in I}$ ist *per Definition* identisch mit der Abbildung $I \longrightarrow X, i \mapsto x_i$.

Es gibt also keinen inhaltlichen Unterschied zwischen einer Familie und einer Abbildung. Es ist eine Frage der Sichtweise bzw. der Notation.

Beispiel 1.8 Man kann die Menge X^n als die Menge der Abbildungen $\{1, \ldots, n\} \longrightarrow X$ *definieren*. In diesem Sinne definieren wir für jede beliebige Menge I:

$$X^I := \mathrm{Abb}(I, X) \; .$$

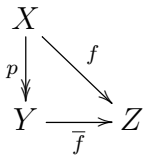
Beispiel 1.9 Eine *Folge* ist eine Familie von reellen Zahlen über der Indexmenge $\mathbb{N}$, d.h. ein Element $(a_n)_{n\in\mathbb{N}}$ mit $a_n \in \mathbb{R}$. Mit anderen Worten: Eine Folge ist per Definition eine Abbildung $\mathbb{N} \longrightarrow \mathbb{R}$. Die Menge der Folgen ist also $\mathbb{R}^{\mathbb{N}}$.

Beispiel 1.10 Seien $X_1, \ldots, X_n$ Mengen. Oben haben wir von Tupeln $(x_1, \ldots, x_n)$ (mit $x_i \in X_i$) sowie vom kartesischen Produkt $X_1 \times \cdots \times X_n$ gesprochen. Dieses Produkt kann man mittels Familien (d.h. mittels Abbildungen) so definieren: Zuerst definiert man X^n für eine beliebige Menge X. Wir definieren nun $X_1 \times \cdots \times X_n$ als die Menge der Familien $(x_1, \ldots, x_n) \in (X_1 \cup \cdots \cup X_n)^n$ mit $x_i \in X_i$.

Wohldefiniertheit

Es kommt häufig vor, dass die folgende Situation gegeben ist:

Gegeben sind drei Mengen X, Y, Z, eine *surjektive* Abbildung $p : X \twoheadrightarrow Y$ sowie eine weitere Abbildung $f : X \longrightarrow Z$. Man fragt sich nun, ob es eine Abbildung $\overline{f} : Y \longrightarrow Z$ mit $\overline{f} \circ p = f : X \longrightarrow Z$ gibt. Mit anderen Worten: Wir wollen, dass das Diagramm



kommutiert. Wiederum mit anderen Worten: Wir wollen, dass für alle $x \in X$ $\overline{f}(p(x)) = f(x)$ gilt. Da p surjektiv ist, ist $\overline{f}$ hierdurch – wenn es existiert – eindeutig bestimmt. Wir erhalten auch sofort eine *notwendige Bedingung* damit $\overline{f}$ existieren kann: Es muss gelten:

$$\forall x, x' \in X : p(x) = p(x') \longrightarrow f(x) = f(x') \tag{1.1}$$

Denn, wenn $\overline{f}$ existiert und $x, x' \in X$ mit $p(x) = p(x')$ gegeben sind, dann ist $f(x) = \overline{f}(p(x)) = \overline{f}(p(x')) = f(x')$. Wenn andererseits (1.1) gilt, dann können wir mittels

$$y \mapsto f(x) \text{ für } \textit{irgendein } x \in X \text{ mit } p(x) = y$$

eine Abbildung $\overline{f} : Y \longrightarrow Z$ definieren. Der entscheidende Punkt ist, dass $\overline{f}(y)$ nun nicht von der Wahl von x abhängt, man also eine *eindeutige Zuordnung*, eben eine Abbildung erhält. Die Unabhängigkeit von der Wahl von x nennt man *Wohldefiniertheit*. Wir fassen dies zusammen:

Aussage 1.11 *Seien X, Y, Z drei Mengen, $p : X \longrightarrow Y$ eine surjektive Abbildung, $f : X \longrightarrow Z$ irgendeine Abbildung. Dann gibt es höchstens eine Abbildung $\overline{f} : Y \longrightarrow Z$ mit $\overline{f} \circ p = f$. So eine Abbildung $\overline{f}$ gibt es genau dann, wenn die Bedingung (1.1) erfüllt ist.*

Aussage 1.12 *Seien die Notationen wie oben. Dann existiert $\overline{f}$ genau dann und ist injektiv, wenn gilt:*

$$\forall x, x' \in X : p(x) = p(x') \longleftrightarrow f(x) = f(x') \tag{1.2}$$

Der Beweis ist leicht.

Kardinalität

Definition Eine Menge X heißt *endlich*, wenn es eine natürliche Zahl n und eine Bijektion $\{1, \ldots, n\} \longrightarrow X$ gibt, andernfalls *unendlich*. Eine Menge heißt *abzählbar*, wenn es eine Surjektion $\mathbb{N} \longrightarrow X$ gibt.

Bemerkung Beachten Sie, dass eine "abzählbare Menge" endlich sein kann!

Man kann (mittels vollständiger Induktion) zeigen:

Lemma 1.13 *Sei X eine endliche Menge, und seien n und m zwei natürliche Zahlen, so dass es Bijektionen $\{1, \ldots, n\} \longrightarrow X$ und $\{1, \ldots, m\} \longrightarrow X$ gibt. Dann ist $n = m$.*

Damit können wir definieren:

Definition Wenn es eine Bijektion $\{1, \ldots, n\} \longrightarrow X$ gibt, dann heißt n die *Kardinalität* von X und wird mit $\#X$ oder $|X|$ bezeichnet. Wenn X unendlich ist, so schreibt man $\#X = |X| = \infty$.

1.4 Relationen

Sei X eine Menge. Eine *Relation* auf X ist intuitiv eine Eigenschaft, die für je zwei Elemente aus X gelten kann oder nicht. Wenn die Eigenschaft für $(x, y) \in X \times X$ gilt, dann sagt man auch, dass x *in Relation zu y steht* (bzgl. der gegebenen Eigenschaft).

Einfache Beispiele für die ganzen Zahlen sind die Beziehungen $<, \leq, >, \geq$ und natürlich auch $=$.

Ein anderes Beispiel wiederum für $\mathbb{Z}$ ist: "$x - y$ ist gerade" (wobei $x, y \in \mathbb{Z}$). Hier stehen also je zwei gerade Zahlen zueinander in Relation und je zwei ungerade Zahlen stehen zueinander in Relation. Hingegen stehen jeweils eine gerade und eine ungerade Zahl (oder umgekehrt) nicht zueinander in Relation.

Eine formale Definition einer Relation erhält man, indem man von der oben erwähnten Eigenschaft zu der Teilmenge von $X \times X$ übergeht, die durch die Eigenschaft definiert wird.

Definition Eine Relation auf einer Menge X ist eine Teilmenge von $X \times X$.

Sei nun R eine Relation, und seien $x, y \in X$. Dann sagen wir, dass x (bezüglich R) *in Relation zu y steht*, wenn $(x, y) \in R$. In diesem Fall schreiben

wir $x \sim_R y$.[7] (Andere Autoren schreiben auch xRy.)

Andernfalls schreiben wir $x \nsim_R y$.

Oft wird $\sim_R$ durch ein anderes Symbol wie z.B. die obigen $(<, \leq, \ldots)$ oder auch einfach $\sim$ ersetzt.

Beispiel 1.14 Die (übliche) Relation $\geq$ auf $\mathbb{Z}$ ist durch $x \geq y :\Longleftrightarrow x - y \in \mathbb{N}_0$ definiert. Mit anderen Worten: Sie ist durch die Menge

$$\{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid x - y \in \mathbb{N}_0\}$$

gegeben.

Äquivalenzrelationen

Sei X eine Menge und R eine Relation auf X.

Definition Die Relation R heißt *Äquivalenzrelation*, falls gilt:

$$\begin{aligned}
&(R) \quad \forall x \in X : x \sim_R x \\
&(S) \quad \forall x, y \in X : x \sim_R y \longleftrightarrow y \sim_R x \\
&(T) \quad \forall x, y, z \in X : x \sim_R y \wedge y \sim_R z \longrightarrow x \sim_R z
\end{aligned}$$

Die Bedingung (R) heißt *Reflexivität*, die Bedingung (S) heiß *Symmetrie*, und die Bedingung (T) heißt *Transitivität*.

Wenn R eine Äquivalenzrelation ist, schreibt man meist $x \sim y$ anstatt $x \sim_R y$. Man sagt dann auch, dass $\sim$ eine Äquivalenzrelation ist. Aufgrund der Symmetrie sagt man auch, dass x und y *zueinander in Relation stehen*, wenn $x \sim_R y$.

Beispiel 1.15 Zu Beginn dieses Abschnitts haben wir die folgende Relation auf $\mathbb{Z}$ erwähnt:

$$x \sim y :\Longleftrightarrow x - y \text{ ist gerade } .$$

Die Eigenschaften (R), (S), (T) sind offensichtlich, es handelt sich also um eine Äquivalenzrelation.

Man kann dieses Beispiel leicht wie folgt verallgemeinern: Sei n eine natürliche Zahl > 1, und seien $x, y \in \mathbb{Z}$. Dann heißen x und y *kongruent* zueinander modulo n, wenn $x - y$ durch n teilbar ist (d.h. falls ein $a \in \mathbb{Z}$ mit

[7] Immer wenn ich in einer *Definition* schreibe *"Wenn ..., dann sagen wir ..."*, meine ich *"Genau dann wenn ..."*.

$y = x + an$ existiert). Wenn x und y (modulo n) kongruent zueinander sind, schreibt man

$$x \equiv y \mod n \ .$$

Diese "Kongruenz modulo n" ist offensichtlich auch eine Äquivalenzrelation.

Beispiel 1.16 Sei X eine beliebige Menge. Dann ist "=" eine Äquivalenzrelation.

Sei nun eine Äquivalenzrelation $\sim$ gegeben.

Definition Zu $x \in X$ definieren wir

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

die *Äquivalenzklasse* zu x. Wir schreiben auch $[x]$, wenn die Relation offensichtlich ist.

Aussage 1.17 *Seien $x, y \in X$ (und somit $[x]_\sim$ und $[y]_\sim$ zwei Äquivalenzklassen). Dann gilt: Entweder ist $[x]_\sim = [y]_\sim$ oder es ist $[x]_\sim \cap [y]_\sim = \emptyset$.*

Beweis. Es sei zunächst $[x]_\sim = [y]_\sim$. Dann ist $x \in [x]_\sim = [y]_\sim$ und somit $[x]_\sim \cap [y]_\sim \neq \emptyset$.

Wir müssen nun zeigen:

$$[x]_\sim \cap [y]_\sim \neq \emptyset \longrightarrow [x]_\sim = [y]_\sim \ . \tag{1.3}$$

Bevor wir diese Implikation beweisen, legen wir noch eine Notation fest:

Wenn $x_1, \ldots, x_n \in X$ gegeben sind und $x_1 \sim x_2, x_2 \sim x_3, \ldots, x_{n-1} \sim x_n$, dann stehen nach der Transitivität alle x_i zueinander in Relation. Dies deuten wir mit $x_1 \sim x_2 \sim \cdots \sim x_n$ an.

Wir zeigen nun die Implikation. Sei also $[x]_\sim \cap [y]_\sim \neq \emptyset$. Dann gibt es ein $z \in [x]_\sim \cap [y]_\sim$. Es gilt also $x \sim z$, $y \sim z$. Nach der Symmetrie gilt $z \sim y$. Somit gilt $x \sim z \sim y$.

Sei nun $x' \in [x]_\sim$ beliebig. Dann ist

$$y \sim x \sim x' \ .$$

Somit ist also $x' \in [y]_\sim$.

Soeben haben wir gezeigt, dass $[x]_\sim \subseteq [y]_\sim$. Analog zeigt man $[y]_\sim \subseteq [x]_\sim$. Damit sind die beiden Mengen gleich. □

Alle Äquivalenzklassen sind Teilmengen von X, und somit sind sie *Elemente* der Potenzmenge $\mathcal{P}(X)$. Wir können somit die *Menge der Äquivalenzklassen* (in X bzgl. $\sim$) betrachten.

Notation Die Menge der Äquivalenzklassen bezeichnen wir mit $X_{/\sim}$.

Wir haben also

$$X_{/\sim} = \{M \in \mathcal{P}(X) \mid \exists x \in X : M = [x]_\sim\}$$

oder etwas weniger formal

$$X_{/\sim} = \{[x]_\sim \mid x \in X\} .$$

Beispiel 1.18 Wir kommen auf die in Beispiel 1.15 diskutierte Äquivalenzrelation "Kongruenz modulo n" zurück. Für $x \in \mathbb{Z}$ bezeichnen wir die Äquivalenzklasse "modulo n" mit $[x]_n$. Es ist also

$$[x]_n = \{x + an | a \in \mathbb{Z}\} .$$

Wir haben n Äquivalenzklassen, nämlich $[0]_n, [1]_n, \ldots, [n-1]_n$.

Im letzten Abschnitt haben wir das Konzept der Wohldefiniertheit einer Abbildung diskutiert. Wir wenden dies nun auf Äquivalenzrelationen an.

Wir nehmen an, dass wir eine Abbildung $f : X \longrightarrow Y$ gegeben haben. Wir fragen uns, ob es eine Abbildung $\overline{f} : X_{/\sim} \longrightarrow Y$ mit $\overline{f}([x]_\sim) = f(x)$ für alle $x \in X$ gibt. Hierzu betrachten wir die surjektive Abbildung $p : X \longrightarrow X_{/\sim}$, $x \mapsto [x]_\sim$. Aussage 1.11 liefert nun:

Aussage 1.19 *Sei $f : X \longrightarrow Y$ gegeben. Dann gibt es genau dann eine Abbildung $\overline{f} : X_{/\sim} \longrightarrow Y$ mit $\overline{f}([x]_\sim) = f(x)$ für alle $x \in X$, wenn gilt:*

$$\forall x, x' \in X : x \sim x' \longrightarrow f(x) = f(x')$$

(und in diesem Fall ist $\overline{f}$ eindeutig bestimmt).

Definition Eine *Partition* einer Menge X ist eine Teilmenge $\mathcal{M}$ von $\mathcal{P}(X)$ (d.h. eine Menge von Teilmengen von X) mit der folgenden Eigenschaften:

- Alle Mengen in $\mathcal{M}$ sind nicht-leer.
- Für alle $x \in X$ existiert genau eine Menge $M \in \mathcal{M}$ mit $x \in M$.

Wenn $\mathcal{M}$ eine Menge von Teilmengen von X mit der zweiten obigen Eigenschaft ist, sagt man auch, dass X die *disjunkte Vereinigung* der Mengen in $\mathcal{M}$ ist. (Dies ist eine Verallgemeinerung der entsprechenden Definition in Abschnitt 1.2.) Hierfür schreibt man:

$$X = \dot{\bigcup}_{M \in \mathcal{M}} M$$

Aus der Reflexivität und Lemma 1.17 folgt:

Aussage 1.20 *Sei $\sim$ eine Äquivalenzrelation auf X. Dann bildet die Menge der Äquivalenzklassen $X_{/\sim}$ eine Partition von X.*

Umgekehrt kann man auch jeder Partition eine Äquivalenzrelation zuordnen. Sei hierzu $\mathcal{M}$ eine Partition von X. Zuerst ordnen wir jedem x die eindeutig bestimmte Menge $M \in \mathcal{M}$ mit $x \in M$ zu. Diese bezeichnen wir mit $[x]_{\mathcal{M}}$. Dann definieren wir wie folgt eine Relation:

$$x \sim y :\Longleftrightarrow [x]_{\mathcal{M}} = [y]_{\mathcal{M}}$$

Die Eigenschaften (R), (S), (T) folgen sofort, und $[x]_{\mathcal{M}}$ ist nun die Äquivalenzklasse zu x.

Definition Sei $\mathcal{M}$ eine Partition von X. Ein *Repräsentantensystem* zu $\mathcal{M}$ ist eine Teilmenge A von X mit der folgenden Eigenschaft: Für alle $M \in \mathcal{M}$ existiert genau ein $a \in A$ mit $[a]_{\mathcal{M}} = M$.

Ein Repräsentantensystem einer Äquivalenzrelation ist per Definition ein Repräsentantensystem der zugehörigen Partition. D.h. ein Repräsentantensystem zu einer Äquivalenzrelation $\sim$ auf X ist eine Teilmenge A von X mit:

$$\forall x \in X \, \exists! \, a \in A : x \sim a \; .$$

Beispiel 1.21 Wir betrachten wieder die Relation "Kongruenz modulo n" auf $\mathbb{Z}$. Ein Repräsentantensystem dieser Relation ist z.B. die Menge $\{0, 1, \ldots, n-1\}$.

Das Auswahlaxiom (Diskussion)

Betrachten wir nun die folgende Aussage.

Jede Partition einer Menge hat ein Repräsentantensystem.

Ein Beweis dieser Aussage ist scheinbar sehr leicht: Gegeben eine Partition $\mathcal{M}$ auf X wählt man aus jeder Menge $M \in \mathcal{M}$ ein (beliebiges) Element $m \in M$ aus. Nehmen wir an, wir haben so eine Auswahl getroffen. Dann haben wir also eine Zuordnung (d.h. Abbildung) $f : \mathcal{M} \longrightarrow X$ mit $f(M) \in M$ für alle $M \in \mathcal{M}$.

Aber gibt es so eine Abbildung immer? Man würde gerne die Existenz so einer Abbildung mittels eines geeignet präzisierten Axiomensystems der Mengenlehre beweisen, wobei die Axiome in etwa so wie die oben angegebenen sein sollen. Eine überraschende Erkenntnis der Mengenlehre ist, dass dies nicht möglich ist. Man sollte die obige Aussage als ein zusätzliches Axiom der Mengenlehre akzeptieren. Es ist das sogenannte *Auswahlaxiom*.

Es sei noch bemerkt, dass die Tatsache, dass das Auswahlaxiom nicht so naheliegend wie die anderen Axiome der Mengenlehre ist, auch Kritiker auf den Plan ruft, die die Verwendung des Auswahlaxioms ablehnen. Diese Kritiker nehmen jedoch eine Außenseiterrolle in der Gemeinschaft der Mathematiker ein.

Ordnungsrelationen

Sei wiederum X eine Menge und R eine Relation auf X.

Definition Die Relation R heißt *Ordnungsrelation*, falls gilt:

$$\begin{aligned}
&(R) \quad \forall x \in X : x \sim_R x \\
&(A) \quad \forall x, y \in X : x \sim_R y \wedge y \sim_R x \longrightarrow x = y \\
&(T) \quad \forall x, y, z \in X : x \sim_R y \wedge y \sim_R z \longrightarrow x \sim_R z
\end{aligned}$$

Eine Ordnungsrelation heißt *lineare Relation* oder *vollständige Relation*, falls gilt:

$$(L) \quad \forall x, y \in X : x \sim_R y \vee y \sim_R x \ .$$

Die Bedingungen (R) und (T) sind die schon bekannte Reflexivität und Transitivität, die Bedingung (A) heißt *Antisymmetrie*. Wenn R eine lineare Relation auf X ist, heißt X (bezüglich R) *linear geordnet*.

Beispiel 1.22 Die Relationen kleiner-gleich bzw. größer-gleich auf $\mathbb{Z}, \mathbb{Q}$ oder $\mathbb{R}$ sind lineare Relationen.[8]

Beispiel 1.23 Sei X eine beliebige Menge. Dann ist "$\subseteq$" eine Ordnungsrelation auf der Potenzmenge $\mathcal{P}(X)$. Diese Relation ist aber nicht linear, wenn X mehr als ein Element besitzt. (Wenn x und y zwei verschiedene Elemente sind, gilt weder $\{x\} \subseteq \{y\}$ noch $\{y\} \subseteq \{x\}$.)

Lemma 1.24 *Sei X eine Menge, und sei $\leq$ eine Ordnungsrelation auf X. Dann gibt es höchstens ein $x \in X$ mit $\forall y \in X : y \leq x$.*

Beweis. Seien x_1, x_2 zwei solche Elemente. Dann ist insbesondere $x_2 \leq x_1$ und $x_1 \leq x_2$. Damit ist $x_1 = x_2$. □

[8] *Vorsicht.* Die übliche Relation $\leq$ auf $\mathbb{Q}$ ist linear, und dies nennt man wie schon gesagt auch vollständig. In der Analysis wird das Wort "vollständig" aber anders benutzt. Und mit der Definition aus der Analysis ist die Relation $\leq$ auf $\mathbb{Q}$ nicht vollständig!

Definition Sei eine Ordnungsrelation $\leq$ auf der Menge X gegeben. Ein Element $x \in X$ wie im letzten Lemma heißt *größtes Element* von X.

Ein Element $x \in X$, so dass

$$\forall y \in X : x \leq y \longrightarrow y = x$$

heißt ein *maximales Element* von X.

Wenn X ein größtes Element hat, dann ist dies (offensichtlich) auch ein maximales Element, und auch das einzige maximale Element.

In vielen wichtigen Beispielen gibt es jedoch mehrere maximale Elemente und kein größtes Element. Hier ist ein instruktives Beispiel.

Beispiel 1.25 Sei $X := \{1, 2, \ldots, 100\}$, und sei Y die Teilmenge von $\mathcal{P}(X)$, die aus den Teilmengen von X mit höchstens 10 Elementen besteht. Wir betrachten die partielle Ordnung "$\subseteq$" auf Y. Nun ist jede Teilmenge mit genau 10 Elementen ein maximales Element von Y, und es gibt kein größtes Element (es gibt keine Teilmenge von X mit höchstens 10 Elementen, die alle Teilmengen mit höchstens 10 Elementen umfasst).

Bemerkung Analog zu den obigen Definitionen kann man *kleinste Elemente* und *minimale Elemente* definieren.

Relationen zwischen zwei Mengen

Man kann den Begriff der Relation erweitern und allgemeiner Relationen zwischen zwei Mengen betrachten. Seien dazu zwei Mengen X und Y gegeben.

Definition Eine *Relation zwischen X und Y* ist eine Teilmenge von $X \times Y$.

Notation Wenn R eine Relation zwischen X und Y ist und $x \in X, y \in Y$, dann schreiben wir $x \sim_R y$ falls $(x, y) \in R$.

Wir werden diese Verallgemeinerung nicht weiter verfolgen und bemerken nur, wie man mittels dieses Begriffs der Relation definieren kann, was eine Abbildung ist. Wir erinnern uns an die intuitive Beschreibung von Abbildungen von X nach Y: jedem Element von X wird genau ein Element von Y zuordnet. Wir erhalten somit die folgende formale Definition.

Definition Eine *Abbildung* $f : X \longrightarrow Y$ ist eine Relation zwischen X und Y, so dass für jedes $x \in X$ genau ein $y \in Y$ mit $x \sim_f y$ existiert.

Bemerkung Aus der Schule kennen Sie den Begriff des *Graphen* einer Funktion. Dies kann man wie folgt für beliebige Abbildungen definieren: Sei $f : X \longrightarrow Y$ eine Abbildung. Dann ist der *Graph* von f die Menge $\{(x, y) \in X \times Y \mid y = f(x)\}$. Nach der obigen Definition sind allerdings eine Abbildung und ihr Graph identisch! Es ist jedoch üblich und sinnvoll, zwischen einer Abbildung und ihrem Graphen zu unterscheiden und z.B. den Graphen einer Abbildung f mit Γ_f zu bezeichen. (Dann ist also $f(x) = y \longleftrightarrow x \sim_f y \longleftrightarrow (x, y) \in \Gamma_f$.)

Diskussion In Beispiel 1.10 haben wir diskutiert, wie man Tupel mittels Abbildungen definieren kann, und oben haben wir Abbildungen mittels Tupel definiert. Diese zirkuläre Definition sollte natürlich aufgehoben werden. Der übliche Weg ist, rein mengentheoretisch zu definieren, was unter einem Zweiertupel (x, y) für $x, y \in X \times Y$ zu verstehen ist. Man definiert z.B.: $(x, y) := \{x, \{x, y\}\}$. Hier ist man aber noch nicht ganz fertig, denn man will ja nicht nur Zweiertupel sondern auch $X \times Y$, die Menge der Zweiertupel, definieren. Wie könnte das gehen?

Eine alternative Möglichkeit wäre, den Abbildungsbegriff axiomatisch vorauszusetzen und die Mengenlehre darauf aufzubauen.

1.5 Halbgruppen, Monoide und Gruppen

Sei im Folgenden X eine Menge.

Definition Eine *Abbildung* $X \times X \longrightarrow X$ heißt eine *Verknüpfung* auf X.

Beispiele für Mengen mit Verknüpfungen sind die Zahlbereiche $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ jeweils mit der Addition und der Multiplikation. Die Subtraktion ist eine Verknüpfung auf $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, nicht aber auf $\mathbb{N}_0$. Die Division ist eine Verknüpfung auf $\mathbb{Q} - \{0\}$ sowie $\mathbb{R} - \{0\}$, nicht jedoch auf $\mathbb{Z} - \{0\}$, $\mathbb{Q}$ oder $\mathbb{R}$.

An diesen Beispielen fällt auf: Man schreibt z.B. $2 + 3 = 5$ und nicht $+(2, 3) = 5$. Eine analoge Schreibweise ist ganz allgemein bei Verknüpfungen üblich. Übliche Symbole für allgemeine Verknüpfungen sind "$\circ$", "$\cdot$", "$*$".

Verknüpfungen auf endlichen Mengen können auch durch eine *Verknüpfungstabelle* angegeben werden.

Beispiel 1.26 Die folgende Tabelle definiert eine Verknüpfung auf der 5-

elementigen Menge $\{1, 2, 3, 4, 5\}$.

$\circ$	1	2	3	4	5
1	1	2	3	4	5
2	2	4	1	5	3
3	3	5	4	2	1
4	4	1	5	3	2
5	5	3	2	1	4

Zum Beispiel ist $2 \circ 3 = 1$ und $3 \circ 2 = 5$.

Definition Sei $\circ : X \times X \longrightarrow X$ eine Verknüpfung. Dann heißt $\circ$

- *assoziativ*, falls $\forall x, y, z \in X : x \circ (y \circ z) = (x \circ y) \circ z$.
- *kommutativ*, falls $\forall x, y \in X : x \circ y = y \circ x$.

Ein Element $e \in X$ heißt *neutrales Element*, wenn gilt: $\forall x \in X : x \circ e = e \circ x = x$.

Lemma 1.27 *Jede Verknüpfung hat höchstens ein neutrales Element.*

Beweis. Seien $e, e' \in X$ neutrale Elemente.[9] Dann gilt

$$e' = e \circ e' = e \ .$$

Bei der ersten Gleichung haben wir benutzt, dass e ein neutrales Element ist, und bei der zweiten Gleichung haben wir benutzt, dass e' ein neutrales Element ist. □

Definition Sei $\circ : X \times X \longrightarrow X$ eine Verknüpfung mit einem neutralen Element e, und sei $x \in X$. Ein Element $y \in X$ mit

- $y \circ x = e$ heißt ein *Links-Inverses* zu x
- $x \circ y = e$ heißt ein *Rechts-Inverses* zu x
- $y \circ x = e$ und $x \circ y = e$ heißt ein *(beidseites) Inverses* zu x.

[9] Mit dieser Formulierung meine ich, dass e und e' nicht notwendigerweise verschieden sein müssen.

Beispiel 1.28 Die Verknüpfung aus Beispiel 1.26 hat ein neutrales Element (die 1), und jedes Element hat (eindeutig bestimmte) Rechts- und Links-Inverse (die aber nicht identisch sind). Die Verknüpfung ist aber nicht assoziativ. Z.B. ist $(2 \circ 2) \circ 3 = 4 \circ 3 = 5$ und $2 \circ (2 \circ 3) = 2 \circ 1 = 2$. Sie ist nicht kommutativ, was man leicht daran sieht, dass die Tabelle nicht symmetrisch bzgl. Spiegelung an der Diagonalen (von oben links nach unten rechts) ist.

Beispiel 1.29 Sei X eine beliebige Menge. Wenn $f : X \longrightarrow X$ und $g : X \longrightarrow X$ zwei Abbildungen sind, dann können wir die Verknüpfung $f \circ g$ der beiden Abbildungen betrachten. Die Zuordnung $(f, g) \mapsto f \circ g$ ist eine Verknüpfung auf der Menge der Abbildungen $\mathrm{Abb}(X, X)$ im Sinne der obigen Definition. Diese Verknüpfung ist offensichtlich assoziativ. Außerdem gibt es ein neutrales Element, nämlich die identische Abbildung $\mathrm{id}_X : X \longrightarrow X,\ x \mapsto x$. Die Elemente mit beidseitigem Inversen sind genau die bijektiven Abbildungen.

Frage Welche Elemente in $\mathrm{Abb}(X, X)$ haben Links-, welche Rechts-Inverse? (Hierbei muss man das Auswahlaxiom benutzen.)

Im Folgenden betrachten wir ausschließlich assoziative Verknüpfungen. Dies bedeutet, dass Klammern grundsätzlich weggelassen werden können.

Definition

- Eine *Halbgruppe* ist eine Menge mit einer assoziativen Verknüpfung.
- Ein *Monoid* ist eine Halbgruppe mit einem neutralen Element.
- Eine *Gruppe* ist ein Monoid, so dass jedes Element ein Inverses hat.

Beispiele für Halbgruppen sind die bereits erwähnten Zahlbereiche $\mathbb{N}_0, \mathbb{N}$, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ jeweils mit der Addition oder der Multiplikation.

Die Zahlbereiche $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ bilden bezüglich der Addition auch Monoide (das neutrale Element ist die 0). Allerdings ist $\mathbb{N}$ bezüglich der Addition kein Monoid. Bezüglich der Multiplikation sind $\mathbb{N}_0, \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ Monoide (das neutrale Element ist die 1). Eine andere Beispielklasse ist in Beispiel 1.29 gegeben.

Beispiele für Gruppen sind die Zahlbereiche $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ bezüglich der Addition. Allerdings ist $\mathbb{N}_0$ bezüglich der Addition keine Gruppe. Bezüglich der Multiplikation sind $\mathbb{Q} - \{0\}$ und $\mathbb{R} - \{0\}$ Gruppen. Andererseits sind $\mathbb{Q}$ und $\mathbb{R}$ bezüglich der Multiplikation keine Gruppen (0 hat kein Inverses!).

Die ein-elementige Menge $\{e\}$ ist mit der Verknüpfung $e \circ e = e$ eine Gruppe (genannt die *triviale Gruppe*).

Ein (zugegebenerweise merkwürdiges) Beispiel für eine Halbgruppe ist die leere Menge.

Lemma 1.30 *Sei M ein Monoid. Zu $x \in M$ gibt es höchstens ein beidseitiges Inverses.*

Beweis. Seien y, z zwei beidseitige Inverse zu x. Dann ist

$$y = y \circ e = y \circ (x \circ z) = (y \circ x) \circ z = e \circ z = z \ .$$

□

Es gilt sogar die folgende stärkere Aussage:

Lemma 1.31 *Sei $x \in M$ invertierbar*[10] *mit Inversem y. Dann ist y das einzige Rechts- und das einzige Links-Inverse von x.*

Beweis. Sei z ein Rechts-Inverses zu x. Dann ergibt die obige Rechnung $y = z$. Die Aussage über Links-Inverse zeigt man analog. □

Notation Sei $x \in M$. Wenn x ein Inverses hat,[11] wird dies oft mit x^{-1} bezeichnet.

Lemma 1.32 *Sei M ein Monoid, und seien $x, y \in M$ invertierbar. Dann ist auch $x \circ y$ invertierbar, und es ist $(x \circ y)^{-1} = y^{-1} \circ x^{-1}$.*

Beweis. Seien $x, y \in M$. Dann ist $(y^{-1} \circ x^{-1}) \circ (x \circ y) = y^{-1} \circ (x^{-1} \circ x) \circ y = y^{-1} \circ y = e = x \circ x^{-1} = x \circ y \circ y^{-1} \circ x^{-1} = (x \circ y) \circ (y^{-1} \circ x^{-1})$. Damit ist per Definition $y^{-1} \circ x^{-1}$ ein Inverses (das Inverse) von $x \circ y$. □

Notation Wenn eine beliebige Verknüpfung $\circ$ auf einer Menge X gegeben ist, wird für $x \in X$ und $n \in \mathbb{N}$ das Element $\overbrace{x \circ \cdots \circ x}^{n \text{ mal}}$ mit x^n bezeichnet.

Sei nun M wieder ein Monoid. Wenn x ein Inverses hat, so ist (für $n \in \mathbb{N}$) $\overbrace{x^{-1} \circ \cdots \circ x^{-1}}^{n \text{ mal}} = (\overbrace{x \circ \cdots \circ x}^{n \text{ mal}})^{-1}$, was man leicht sieht (siehe auch das obige Lemma). Dieses Element wird mit x^{-n} bezeichnet. Man setzt $x^0 := e$. Mit dieser Notation gilt

$$x^n \circ x^m = x^{n+m} \text{ und } (x^n)^m = x^{nm} \ .$$

Das Verknüpfungssymbol selbst wird oft weggelassen. Also ist $xy = x \circ y$.

[10] Mit *invertierbar* meinen wir immer, dass es ein beidseitiges Inverses gibt.

[11] Mit einem *Inversen* meinen wir immer ein beidseitiges Inverses.

Definition Eine Halbgruppe resp. ein Monoid resp. eine Gruppe mit einer kommutativen Verknüpfung heißt *kommutative* oder *abelsche* Halbgruppe etc.

Notation Wenn eine abelsche Halbgruppe gegeben ist, benutzt man oft die folgende "additive Notation": Man benutzt das Symbol "+" für die Verknüpfung, und für ein Element x und $n \in \mathbb{N}$ setzt man $nx := \overbrace{x + \cdots + x}^{n \text{ mal}}$. Wenn Elemente $x_1, \ldots, x_r$ gegeben sind, setzt man $\sum_{i=1}^{r} x_i := x_1 + \cdots + x_r$.

Wenn ein abelsches Monoid M gegeben ist, schreibt man dann 0_M (oder 0) für das neutrale Element, und man setzt $0 \cdot x := 0_M$ für $x \in M$. Wenn x ein inverses Element hat, bezeichnet man dies mit $-x$.

Beachten Sie, dass die additive Notation der üblichen Notation bezüglich der Addition von Zahlen entspricht.

Beispiel 1.33 Sei Σ eine beliebige nicht-leere Menge. (Die Notation Σ deutet an, dass wir Σ als Alphabet betrachten, aber wie gesagt muss Σ nicht notwendigerweise endlich und auch nicht abzählbar sein.) Wir betrachten die Menge aller Tupel beliebiger Länge ≥ 1 von Elementen von Σ zusammen mit einem weiteren Element $\square$. [12] Diese Menge heißt die Menge der *Worte* in Σ und wird mit Σ^* bezeichnet, $\square$ wird *leeres Wort* genannt. Statt $(a_1, \ldots a_n)$ schreibt man $a_1 \cdots a_n$.

Durch "Hintereinanderschreiben" vw kann man zwei Wörtern v, w ein neues zuordnen (wobei $v\square := v, \square w := w$). Man erhält also eine Verknüpfung auf Σ^*. Mit dieser Verknüpfung ist Σ^* ein Monoid (mit neutralem Element $\square$). Wenn Σ mehr als ein Element enthält, ist dieses Monoid nicht abelsch.

Beispiel 1.34 Sei X eine beliebige Menge. Dann ist $\mathrm{Abb}(X, X)$ ein Monoid (siehe Beispiel 1.29). Wenn allerdings X mindestens zwei Elemente enthält[13] ist $\mathrm{Abb}(X, X)$ keine Gruppe und auch nicht abelsch.

Beweis. Seien $a, b \in X$ zwei verschiedene Elemente. Dann ist die Abbildung $f : x \mapsto a$ nicht injektiv (da insbesondere a und b auf a abgebildet werden), besitzt also kein Inverses (keine Umkehrabbildung). Damit ist X keine Gruppe. Sei $g : x \mapsto b$. Dann ist $g \circ f$ durch $x \mapsto b$ gegeben und $f \circ g$ durch $x \mapsto a$ gegeben. Damit ist X nicht abelsch. $\square$

[12] Da man für $n \geq 1$ die Menge der n-Tupel Σ^n als die Menge $\Sigma^{\{1,\ldots,n\}}$ auffassen kann (siehe Beispiel 1.10), ist es naheliegend, $\square$ als die eindeutige Abbildung $\emptyset \longrightarrow \Sigma$ zu definieren. Dies ist dann das "leere Tupel".

[13] Mit dieser Redewendung meine ich "mindestens zwei *verschiedene* Elemente".

Definition Sei wiederum X eine beliebige Menge. Eine *Permutation* auf X ist eine bijektive Abbildung $X \longrightarrow X$. Die Menge der Permutationen von X wird mit $\mathrm{Perm}(X)$ oder mit $S(X)$ bezeichnet. Für $n \in \mathbb{N}$ ist $S_n := S(\{1, \ldots, n\})$ die Menge der Permutationen auf $\{1, \ldots, n\}$.

Beispiel 1.35 $\mathrm{Perm}(X)$ ist eine Gruppe. Diese Gruppe ist genau dann abelsch, wenn X höchstens zwei Elemente besitzt.

Diese Gruppe heißt die *symmetrische Gruppe* auf X; die Gruppe S_n heißt die *symmetrische Gruppe auf n Elementen*.

Beweis. Da jede bijektive Abbildung eine Umkehrabbildung hat, ist $\mathrm{Perm}(X)$ eine Gruppe.

Wenn X kein oder ein Element besitzt, besteht $\mathrm{Perm}(X)$ nur aus der identischen Abbildung. Wenn X zwei Elemente a, b besitzt, besteht $\mathrm{Perm}(X)$ aus id_X und τ mit $\tau : a \mapsto b, b \mapsto a$. Damit ist $\mathrm{Perm}(X)$ kommutativ.

Seien nun a, b, c drei verschiedene Elemente von X. Betrachte die Abbildungen

$$f : a \mapsto b,\ b \mapsto a,\ x \mapsto x \text{ für } x \neq a, b$$

sowie

$$g : a \mapsto b,\ b \mapsto c,\ c \mapsto a,\ x \mapsto x \text{ für } x \neq a, b, c\ .$$

Dann ist $(g \circ f)(a) = g(b) = c$, $(f \circ g)(a) = f(b) = a$, also insbesondere $f \circ g \neq g \circ f$. □

Produkte

Seien X, Y zwei Halbgruppen. Wir definieren wie folgt eine Verknüpfung auf $X \times Y$:

$$(x, y) \circ (x', y') := (x \circ x', y \circ y')\ .$$

Diese "komponentenweise definierte" Verknüpfung ist offensichtlich assoziativ. Damit ist auch $X \times Y$ eine Halbgruppe.

Wenn X und Y Monoide mit neutralen Elementen e_X, e_Y sind, dann ist auch $X \times Y$ ein Monoid mit neutralem Element (e_X, e_Y).

Wenn nun x und y invertierbar sind, ist offensichtlich auch (x, y) invertierbar mit Inversem (x^{-1}, y^{-1}). Insbesondere ist $X \times Y$ eine Gruppe, wenn X und Y Gruppen sind.

Außerdem erhält man abelsche Halbgruppen, abelsche Monoide oder abelsche Gruppen, wenn X und Y abelsche Halbgruppen, abelsche Monoide oder abelsche Gruppen sind.

Selbstverständlich gelten diese Aussagen auch für mehr als zwei Faktoren X, Y. Damit ist also insbesondere X^n (für $n \in \mathbb{N}$) in natürlicher Weise eine

Halbgruppe, ein Monoid oder eine Gruppe, wenn X eine Halbgruppe, ein Monoid oder eine Gruppe ist. Zum Beispiel sind $\mathbb{Z}^n, \mathbb{Q}^n$ und $\mathbb{R}^n$ mit der "komponentenweisen" Addition abelsche Gruppen.

Sei nun X weiterhin eine Halbgruppe, und sei I eine Menge. Auch auf X^I können wir in natürlicher Weise eine Verknüpfung definieren. Wir erinnern daran, dass X^I aus den Abbildungen $I \longrightarrow X$ besteht, und diese Abbildungen werden oft in der Form $(x_i)_{i\in I}$ geschrieben. Wir folgen dieser Schreibweise. Wir haben die folgende Verknüpfung auf X^I:

Gegeben $(x_i)_{i\in I}, (x'_i)_{i\in I}$, definieren wir

$$(x_i)_{i\in I} \circ (x'_i)_{i\in I} := (x_i \circ x'_i)_{i\in I} \ .$$

Damit ist X^I wiederum eine Halbgruppe. Wenn X ein Monoid mit neutralem Element e ist, dann ist X^I ein Monoid mit neutralem Element $(e)_{i\in I}$ (dies ist die Abbildung, die jedem $i \in I$ das neutrale Element e von X zuordnet). Die Halbgruppe X^I ist eine Gruppe, wenn X eine Gruppe ist.

Unterstrukturen

Definition Sei X eine Menge mit einer Verknüpfung "$\circ$", und sei Y eine Teilmenge von X. Dann heißt Y *abgeschlossen* bezüglich "$\circ$", wenn gilt:

$$\forall y, y' \in Y : y \circ y' \in Y$$

In diesem Fall definiert $\circ$ durch Einschränkung auf $Y \times Y$ eine Verknüpfung auf Y, man spricht auch von der *induzierten Verknüpfung*.

Beachten Sie: Wenn die Verknüpfung auf X assoziativ (resp. kommutativ) ist, so ist auch die induzierte Verknüpfung assoziativ (resp. kommutativ). Wir haben somit:

- Sei H eine Halbgruppe, und sei $U \subseteq H$ abgeschlossen (bezüglich der Verknüpfung auf H). Dann ist U mit der induzierten Verknüpfung eine Halbgruppe; man spricht von einer *Unterhalbgruppe*.

- Sei M ein Monoid mit neutralem Element e, und sei $U \subseteq M$ abgeschlossen mit $e \in U$. Dann ist U mit der induzierten Verknüpfung ein Monoid mit neutralem Element e; man spricht von einem *Untermonoid*.

- Sei G eine Gruppe mit neutralem Element e, und sei $U \subseteq G$ abgeschlossen mit $e \in U$, so dass für jedes $x \in U$ auch das inverse Element x^{-1} in U liegt. Dann ist U mit der induzierten Verknüpfung eine Gruppe mit neutralem Element e; man spricht von einer *Untergruppe*.

Beispiel 1.36 Sei $(G, +)$ eine additiv geschriebene abelsche Gruppe, und seien $g_1, \ldots, g_r \in G$. Dann ist

$$\langle g_1, \ldots, g_r \rangle := \{z_1 g_1 + \cdots + z_r g_r \mid z_i \in \mathbb{Z}\}$$

eine Untergruppe von G (nachrechnen!). Es gilt: Wenn $U \subseteq G$ irgendeine Untergruppe mit $g_1, \ldots, g_r \in U$ ist, dann ist $\langle g_1, \ldots, g_r \rangle \subseteq U$ (warum?). $\langle g_1, \ldots, g_r \rangle$ ist also die *kleinste* Untergruppe von G, die $g_1, \ldots, g_r$ enthält (siehe den Unterabschnitt über Ordnungsrelationen im vorherigen Abschnitt).

Die Untergruppe $\langle g_1, \ldots, g_r \rangle$ von G heißt die von $g_1, \ldots, g_r$ *erzeugte* Untergruppe bzw. das *Erzeugnis* von $g_1, \ldots, g_r$.

Das kann man noch etwas abändern: Sei G wie zuvor und $S \subseteq G$ eine beliebige Teilmenge (die auch unendlich groß sein kann.) Dann ist

$$\begin{aligned}
&\langle S \rangle := \\
&\{g \in G \mid \exists k \in \mathbb{N}_0 \, \exists g_1, \ldots, g_k \in S, \exists z_1, \ldots, z_k \in \mathbb{Z} : g = z_1 g_1 + \cdots + z_k g_k\} = \\
&\{g \in G \mid \exists k \in \mathbb{N}_0 \, \exists g_1, \ldots, g_k \in S, \exists z_1, \ldots, z_k \in \{1, -1\} : g = z_1 g_1 + \cdots + z_k g_k\}
\end{aligned}$$

die kleinste Untergruppe von G, die die Menge S umfasst.[14] Man spricht wieder von der von S *erzeugten* Untergruppe.

Frage Natürlich kann man eine abelsche Gruppe auch "multiplikativ" schreiben. Sei $(G, \circ)$ so eine Gruppe. Wie lautet dann das Erzeugnis von $g_1, \ldots, g_k \in G$? Wie lautet z.B. das Erzeugnis von $2, 3, 5 \in \mathbb{Q}^*$? Sind die Untergruppen $\langle 2 \rangle$ und $\langle -2 \rangle$ von $\mathbb{Q}^*$ identisch?

Beispiel 1.37 Sei M ein Monoid, und sei $G \subseteq M$ die Menge der invertierbaren Elemente von M. Dann ist (nach Lemma 1.32) G abgeschlossen und somit ein Untermonoid. Es ist per Definition auch eine Gruppe, genannt die *Gruppe der invertierbaren Elemente* von M.

Beispiel 1.38 Die Gruppe der invertierbaren Elemente von $(\mathbb{Z}, \cdot)$ ist $\{1, -1\}$.

Beispiel 1.39 Sei X eine Menge. Dann ist die Gruppe der invertierbaren Elemente von $\mathrm{Abb}(X, X)$ gleich $\mathrm{Perm}(X)$.

Aussage 1.40 (Untergruppenkriterium) *Sei G eine Gruppe, und sei $U \subseteq G$ eine Teilmenge. Dann ist U (mit der induzierten Verknüpfung) genau dann eine Untergruppe von G, wenn gilt:*

- *U ist nicht-leer*

[14] Für $k = 0$ erhält man die "leere Summe", und die ist per Definition 0.

- $\forall x, y \in U : x \circ y^{-1} \in U$.

Beweis. Wenn U eine Untergruppe ist, gelten die Eigenschaften offensichtlich. Seien also nun die beiden Eigenschaften erfüllt. Da U nicht-leer ist, gibt es ein $x_0 \in U$. Damit ist nach der zweiten Eigenschaft $e = x_0 \circ x_0^{-1} \in U$. Für $x \in U$ beliebig ist dann auch $x^{-1} = e \circ x^{-1} \in U$. Seien nun $x, y \in U$ beliebig. Dann ist $y^{-1} \in U$, wie wir gerade gesehen haben. Somit ist auch $x \circ y = x \circ (y^{-1})^{-1} \in U$. Dies ist die Abgeschlossenheit. Die beiden anderen Eigenschaften wurden zuvor gezeigt. □

Faktorgruppen

Sei nun $(G, +)$ eine *abelsche* Gruppe, und sei U eine Untergruppe. Wir führen wie folgt eine Relation auf G ein:

$$x \sim_U y :\Longleftrightarrow x - y \in U \ .$$

Man sieht leicht, dass dies eine Äquivalenzrelation ist; die Äquivalenzklasse zu $x \in G$ bezeichnen wir mit $[x]_U$. Damit gilt also: Für $x \in G$ ist

$$[x]_U = \{x + u \mid u \in U\} \ .$$

Diese Menge wird auch mit $x + U$ bezeichnet.

Wir wollen auf $G_{/\sim_U}$ eine Verknüpfung definieren durch

$$[x]_U + [y]_U := [x + y]_U \ .$$

Hierzu müssen wir die Wohldefiniertheit nachweisen, d.h. wir müssen nachweisen, dass gilt:

$$\forall x, y, x', y' \in G : x \sim_U x' \wedge y \sim_U y' \longrightarrow x + y \sim_U x' + y'$$

Seien also $x, y, x', y' \in G$ mit $x \sim_U x'$ und $y \sim_U y'$. Dann ist also $x - x' \in U$ und $y - y' \in U$. Nun ist $(x + y) - (x' + y') = (x - x') + (y - y') \in U$, d.h. $x + y \sim_U x' + y'$. □

Man rechnet leicht nach, dass die Verknüpfung auf $G_{/\sim_U}$ assoziativ und abelsch ist. Damit ist also $G_{/\sim_U}$ mit der soeben definierten Verknüpfung eine Halbgruppe. Außerdem ist $[0_G]_U = U$ ein neutrales Element von $G_{/\sim_U}$, und für $x \in G$ ist $[-x]_U = -x + U$ ein inverses Element von $[x]_U = x + U$.

Damit ist $G_{/\sim_U}$ sogar eine abelsche Gruppe.

Definition Die soeben definierte Gruppe wird mit G/U bezeichnet und heißt die *Faktorgruppe* von G *nach* U (oder G *modulo* U), die Verknüpfung heißt wiederum die *induzierte Verknüpfung*.

Zur Verdeutlichung: Es ist $0_{G/U} = [0_G]_U$ und $-[x]_U = [-x]_U$ für alle $x \in G$.

Beispiel 1.41 Sei n eine natürliche Zahl > 1, und sei $n\mathbb{Z} := \{z \in \mathbb{Z} \mid n \text{ teilt } z\} = \{na \mid a \in \mathbb{Z}\}$. Dann ist $n\mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$. Die Äquivalenzrelation $\sim_{n\mathbb{Z}}$ ist (per Definition) durch

$$x \sim_{n\mathbb{Z}} y \longleftrightarrow x - y \in n\mathbb{Z} \longleftrightarrow n \text{ teilt } x - y \longleftrightarrow x \equiv y \mod n$$

gegeben. Mit anderen Worten, es ist die Relation "Kongruenz modulo n", die wir in den Beispielen 1.15 und 1.18 diskutiert haben. Ich erinnere daran, dass wir für $x \in \mathbb{Z}$ die Äquivalenzklasse mit $[x]_n$ bezeichnen, und es ist $[x]_n = \{x + na \mid a \in \mathbb{Z}\}$. Die Gruppe $\mathbb{Z}/n\mathbb{Z}$ besteht folglich aus den n Elementen $[0]_n, [1]_n, \ldots, [n-1]_n$.

Beispiel 1.42 Sei $(a_1, a_2) \in \mathbb{R}^2$ mit $(a_1, a_2) \neq (0, 0)$. Sei

$$U := \{(ra_1, ra_2) \mid r \in \mathbb{R}\} .$$

Geometrisch ist dies eine Gerade, die durch den Ursprung geht. Für $(x_1, x_2) \in \mathbb{R}^2$ ist $[(x_1, x_2)]_U = (x_1, x_2) + U$ die Parallele von U durch (x_1, x_2). Die Faktorgruppe $\mathbb{R}^2/U$ besteht somit aus allen Parallelen von U in $\mathbb{R}^2$.

1.6 Ringe und Körper

Definition Ein *Ring* ist eine Menge R mit zwei Verknüpfungen "+" und "$\cdot$", so dass

- $(R, +)$ eine abelsche Gruppe ist,
- $(R, \cdot)$ ein Monoid ist,
- die *Distributivgesetze* gelten, d.h.

$$\forall a, b, c \in R : (a + b)c = ac + bc , \quad c(a + b) = ca + cb .$$

Ein Ring heiß *kommutativ*, wenn die Multiplikation eine kommutative Verknüpfung ist.

Notation Bei den Distributivgesetzen haben wir die übliche Rechenregel "mal vor plus" benutzt und die Klammern auf der rechten Seite der Gleichung entsprechend weglassen. So verfahren wir auch im Folgenden.

Bemerkung Beachten Sie, dass sich das Adjektiv *kommutativ* hier nur auf die Multiplikation bezieht; die Addition eines Rings ist per Definition immer kommutativ.

Beispiel 1.43 Die ganzen, die rationalen sowie die reellen Zahlen bilden jeweils kommutative Ringe mit den Verknüpfungen "+" und "$\cdot$". Die neutralen Elemente sind 0 und 1.

Beispiele für nicht-kommutative Ringe werden wir später kennenlernen.

Notation Das neutrale Element bezüglich "+" wird mit 0 (oder genauer mit 0_R) bezeichnet, und das neutrale Element bezüglich "$\cdot$" wird mit 1 (oder genauer mit 1_R) bezeichnet. Es gilt also ("wie üblich") $0+r = r+0 = r$ und $1 \cdot r = r \cdot 1 = r$ für alle $r \in R$.

Aussage 1.44 *Sei R ein Ring.*

a) Für alle $r \in R$ gilt $0 \cdot r = 0$.

b) Für alle $r \in R$ gilt $(-1) \cdot r = -r$.

c) Wenn $0 = 1$ (in R), dann ist $R = \{0\}$.

Beweis.
zu a) Sei $r \in R$ beliebig. Dann ist $0 \cdot r = (0+0) \cdot r = 0 \cdot r + 0 \cdot r$. Hieraus folgt (mit Subtraktion von $0 \cdot r$) $0 = 0 \cdot r$.
zu b) Sei wiederum $r \in R$ beliebig. Dann ist $0 = 0 \cdot r = (1-1) \cdot r = 1 \cdot r + (-1) \cdot r = r + (-1) \cdot r$. Daraus folgt die Behauptung.
zu c) Sei $0 = 1$ und sei $r \in R$ beliebig. Dann ist $r = 1 \cdot r = 0 \cdot r = 0$. (Andererseits ist die Menge $\{0\}$ mit den Verknüpfungen $0 + 0 = 0$ und $0 \cdot 0 = 0$ ein Ring.) □

Definition Ein *Körper* ist ein kommutativer Ring mit $0 \neq 1$, in dem jedes Element $\neq 0$ ein Inverses bezüglich der Multiplikation hat.

Beispiel 1.45 Beispiele für Körper sind die rationalen und die reellen Zahlen.

Notation Körper werden oft mit K bezeichnet.

Definition Sei R ein Ring. Die Gruppe der invertierbaren Elemente von $(R, \cdot)$ wird mit R^* bezeichnet (siehe Beispiel 1.37).

Aussage 1.46 *Sei R ein Ring mit $0 \neq 1$. Dann ist $0 \notin R^*$. Wenn K ein Körper ist, dann ist $K^* = K - \{0\}$ (mit der induzierten Verknüpfung).*

Beweis. Sei also R ein Ring mit $0 \neq 1$. Angenommen $0 \in R^*$. Dann gibt es ein $r \in R$ mit $0 \cdot r = 1$. Dies ist ein Widerspruch, da nach Aussage 1.40 $0 \cdot r = 0$ ist.

Sei nun K ein Körper. Soeben haben wir gesehen, dass $K^* \subseteq K - \{0\}$. Es gilt aber auch $K - \{0\} \subseteq K^*$ nach Definition eines Körpers. □

Die folgende Definition ist analog zur Definition der Unterstrukturen in Abschnitt 1.5.

Bemerkung / Definition Sei R ein Ring, und sei $U \subseteq R$ abgeschlossen bezüglich Addition und Multiplikation, so dass U eine Untergruppe von R bezüglich der Addition und ein Untermonoid von R bezüglich der Multiplikation ist (insbesondere ist also $0, 1 \in U$). Dann ist U mit den induzierten Verknüpfungen ein Ring; man spricht von einem *Unterring*.

Sei nun R ein Körper. Dann haben insbesondere alle Elemente von $U - \{0\}$ ein multiplikatives Inverses in K. Falls diese Inversen alle in U liegen, ist U mit den induzierten Verknüpfungen ein Körper; man spricht von einem *Unterkörper* oder einem *Teilkörper*.

Restklassenringe und Primkörper

Es sei n eine natürliche Zahl ≥ 2. Wir betrachten die Faktorgruppe $\mathbb{Z}/n\mathbb{Z}$. Auf dieser Gruppe wollen wir auch eine Multiplikation definieren, und zwar so:

$$[a]_n \cdot [b]_n := [ab]_n$$

für $a, b \in \mathbb{Z}$. Wir müssen zeigen, dass dies wohldefiniert ist. Seien dazu $a, a', b, b' \in \mathbb{Z}$ mit $[a]_n = [a']_n$ und $[b]_n = [b']_n$. Dann haben wir also $a \equiv a' \mod n$ und $b \equiv b' \mod n$. Oder mit anderen Worten: $n|a-a'$ und $n|b-b'$. Nun gilt $ab \equiv a'b' \mod n$, denn: $ab-a'b' = ab-a'b+a'b-a'b' = (a-a')b+a'(b-b')$, und dies ist durch n teilbar.

Man sieht leicht, dass die Multiplikation wieder assoziativ und kommutativ ist und dass $[1]_n$ ein neutrales Element (und somit das neutrale Element) ist. Außerdem gelten die Distributivgesetze. Somit ist also $\mathbb{Z}/n\mathbb{Z}$ ein kommutativer Ring, genannt *Restklassenring modulo n*.

Erinnern Sie sich: Eine *Primzahl* ist eine natürliche Zahl, die nur durch sich selbst und 1 teilbar ist. Wir setzen nun das folgende Lemma voraus:

Lemma 1.47 *Sei n eine natürliche Zahl. Dann sind äquivalent:*

- *n ist eine Primzahl*
- $\forall a, b \in \mathbb{N} : n|ab \longrightarrow n|a \vee n|b$

Es ist offensichtlich, dass der zweite Punkt den ersten impliziert. (Denn: Es gelte die Bedingung im zweiten Punkt. Sei nun $a \in \mathbb{N}$ mit $a|n$ und $a \neq 1$. Dann gibt es ein $b \in \mathbb{N}$ ($b < n$) mit $n = ab$, insbesondere also $n|ab$. Damit gilt $n|a$ oder $n|b$. Die zweite Bedingung ist unmöglich. Und damit gilt $n|a$ und $a|n$, also $a = n$.)

Die Umkehrung ist ein wenig schwieriger. Wir kommen darauf in zweiten Semester zurück.

Aussage 1.48 *Der Ring $\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Körper, wenn n eine Primzahl ist.*

Beweis. Sei zunächst n keine Primzahl. Dann haben wir $a, b \in \mathbb{N}$ mit $a, b \neq 1$ und $ab = n$. Dann gilt in $\mathbb{Z}/n\mathbb{Z}$: $[a]_n \cdot [b]_n = [ab]_n = [0]_n$ aber $[a]_n, [b]_n \neq [0]_n$. Damit ist $\mathbb{Z}/n\mathbb{Z}$ kein Körper.

Sei umgekehrt n eine Primzahl. Sei $a \in \{1, \ldots, n-1\}$. Wir wollen zeigen, dass $[a]_n$ invertierbar ist. Wir betrachten hierzu die Abbildung $\mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}, [b]_n \mapsto [a]_n \cdot [b]_n = [ab]_n$. Ich behaupte, dass dies eine bijektive Abbildung ist. Nehmen wir für den Moment an, dass dies schon gezeigt ist. Dann gibt es insbesondere ein $b \in \mathbb{Z}$ mit $[a]_n \cdot [b]_n = [ab]_n = [1]_n$. Somit ist also $[a]_n$ invertierbar.

Um die Bijektivität der Abbildung zu zeigen, zeigen wir, dass die Abbildung injektiv ist. Sie ist dann auch bijektiv, da sie von einer endlichen Menge auf sich selbst geht. Seien nun $[b]_n, [b']_n \in \mathbb{Z}/n\mathbb{Z}$ mit $[ab]_n = [ab']_n$. Dann ist $[a(b-b')]_n = [0]_n$, d.h. $n|a(b-b')$. Nach Voraussetzung ist n kein Teiler von a, und somit gilt nach dem obigen Lemma $n|b-b'$, also $b \equiv b' \mod n$ bzw. $[b]_n = [b']_n$. □

Definition Für eine Primzahl p setzen wir $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ und nennen diesen Körper den *Primkörper* mit p Elementen.

Schließlich fixieren wir noch eine Notation:

Notation Sei R ein Ring und $z \in \mathbb{Z}$. Dann schreiben wir für $z \cdot 1_R$ auch z_R oder auch nur z. Hier muss man allerdings wirklich aufpassen, da z_R eben ein Element von R ist und R nicht $\mathbb{Z}$ liegen muss.

Beispiel 1.49 Wir betrachten den Ring $R := \mathbb{Z}/3\mathbb{Z}$. Wir haben nun allgemein für alle ganzen Zahlen z: $z_R = z \cdot [1]_3 = [z]_3$. Somit haben wir also z.B.

$$3_R = 0_R \quad , \quad 4_R = 1_R \quad , \quad 5_R = 2_R \quad , \quad -1_R = 2_R \, .$$

Indem wir den Index "R" weglassen, können wir das auch so schreiben:

$$3 = 0 \quad , \quad 4 = 1 \quad , \quad 5 = 2 \quad , \quad -1 = 2$$

Natürlich sind diese Identitäten in $\mathbb{Z}$ falsch, aber mit unseren Notationen sind sie im Restklassenring $\mathbb{Z}/3\mathbb{Z}$ richtig.

Diskussion Hier ist vielleicht ein allgemeiner Kommentar angebracht: Dass das Symbol "1" die erste natürlich Zahl bezeichnet, das Symbol "2" die zweite usw., ist nur eine Konvention ohne weitere Bedeutung für die Mathematik. Wenn wir festlegen, dass "3" die dritte natürlich Zahl bedeute und "0" das neutrale Element in $\mathbb{Z}$ bzgl. der Addition, dann gilt natürlich $3 \neq 0$. Wenn wir aber festlegen, dass "3" das Element $3 \cdot 1_R$ mit $R = \mathbb{Z}/3\mathbb{Z}$ bedeute und 0 das neutrale Element in diesem Ring, dann gilt eben $3 = 0$.

1.7 Die ganzen und die rationalen Zahlen

Bisher sind wir stillschweigend davon ausgegangen, dass die ganzen und die rationalen Zahlen existieren und einige offensichtliche Eigenschaften haben. Hier wollen wir nun zeigen, wie man explizit definieren kann, was der Ring der ganzen Zahlen und der Körper der rationalen Zahlen ist, bzw. wie man eine explizite Konstruktion der ganzen Zahlen und der rationalen Zahlen angeben kann.

Wir stellen uns auf den Standpunkt, dass die Menge $\mathbb{N}_0$ der natürlichen Zahlen einschließlich der Null existiert (das war sowieso ein Axiom), und die Addition und Multiplikation in dieser Menge die üblichen Gesetze wir Assoziativität, Kommutativität und Distributivität erfüllen.

Nun zuerst zu den ganzen Zahlen. Es gibt zwei naheliegende Möglichkeiten, von den natürlichen Zahlen zu den ganzen Zahlen zu gelangen.

Die erste Möglichkeit ist, $\mathbb{Z}$ als die Vereinigung von positiven, negativen Zahlen und der Null zu definieren. Man fixiert eine Menge M, die bijektiv zu $\mathbb{N}$ ist aber von $\mathbb{N}$ verschieden ist. Wir haben also eine Bijektion $m : \mathbb{N} \longrightarrow M$.

Dann setzt man $Z := \mathbb{N}_0 \cup M$; dies ist nun eine disjunkte Vereinigung, und für jedes Element $z \in Z$ gilt nun: Entweder es ist $z \in \mathbb{N}_0$ oder es gibt ein (eindeutig bestimmtes) $n \in \mathbb{N}$ mit $z = m(n)$.

Nun kann man die Operationen "+" und "$\cdot$" per Fallunterscheidung auf von $\mathbb{N}_0$ auf Z ausdehnen. Für "+" sieht das dann so aus: Wir definieren $0 + z := z$ und $z + 0 := z$ für alle $z \in Z$, sowie für $a, b \in \mathbb{N}$:

$$\begin{array}{lll} a + b & := a + b & \\ a + m(b) & := a - b & \text{wenn } a \geq b \\ a + m(b) & := m(b - a) & \text{wenn } b > a \\ m(a) + b & := b - a & \text{wenn } b \geq a \\ m(a) + b & := m(a - b) & \text{wenn } a > b \end{array}$$

Man beachte, dass sich die Operationen "+" und "-" auf der rechten Seite auf die schon bekannten natürlichen Zahlen beziehen.

Jetzt muss man noch die Multiplikation definieren und dann nachweisen, dass Assoziativität, Kommutativität und Distributivität immer noch gelten. Der Ring $\mathbb{Z}$ ist dann per Definition $(Z, +, \cdot)$.

Wir schildern nun eine andere Möglichkeit, die auf einer allgemeinen Methode beruht, die auch in anderem Kontext Anwendung findet.

Die Idee ist, dass sich jede ganze Zahl in der Form $a - b$ mit $a, b \in \mathbb{N}_0$ schreiben lässt, aber diese Darstellung ist nicht eindeutig. Genauer: Es gilt für $a, c, b, d \in \mathbb{N}_0$: $a - b = c - d \longleftrightarrow a + d = b + c$. Diese Überlegung nimmt man nun zum Anlass, die ganzen Zahlen mittels Äquivalenzklassen von Tupeln in $\mathbb{N}_0 \times \mathbb{N}_0$ zu definieren.

Man definiert eine Äquivalenzrelation auf $\mathbb{N}_0 \times \mathbb{N}_0$ wie folgt: $(a, b) \sim (c, d) :\Longleftrightarrow a + d = b + c$, und man setzt $Z := (\mathbb{N}_0 \times \mathbb{N}_0)_{/\sim}$. Wir wollen nun eine Verknüpfung "+" ("Addition") auf $\mathbb{Z}$ definieren durch

$$[(a, b)] + [(c, d)] := [(a + c, b + d)] \; .$$

Hierzu müssen wir nachprüfen, dass dies *wohldefiniert* ist. Sei hierzu $(a, b) \sim (a', b')$ und $(c, d) \sim (c', d')$ mit $a, a', b, b', c, c', d, d' \in \mathbb{N}_0$. Dann ist also $a + b' = a' + b$ und $c + d' = c' + d$. Somit ist

$$(a + c) + (b' + d') = (a + b') + (c + d') = (a' + b) + (c' + d) = (a' + c') + (b + d)$$

und somit $(a + c, b + d) \sim (a' + c', b' + d')$. □

Um die Multiplikation zu definieren, erinnern wir uns, dass $(a - b) \cdot (c - d) = ac + bd - (ad + bc)$. In diesen Sinne wollen wir eine weitere Verknüpfung "$\cdot$" ("Multiplikation") auf Z definieren durch

$$[(a, b)] \cdot [(c, d)] := [(ac + bd, ad + bc)] \; .$$

Wieder rechnet man nach, dass dies wohldefiniert ist. Dann muss man noch zeigen, dass die Assoziativ-, Kommutativ- und Distributivgesetze gelten.

Der Ring $\mathbb{Z}$ ist nun per Definition wieder $(Z, +, \cdot)$.

Wir kommen nun zu den rationalen Zahlen. Nun gehen wir von den ganzen Zahlen $\mathbb{Z}$ aus. Diesmal ist der Ausgangspunkt, dass sich jede rationale Zahl als Bruch $\frac{a}{b}$ mit $a \in \mathbb{Z}$ und $b \in \mathbb{N}$ schreiben lässt, und es gilt für $a, c \in \mathbb{Z}$ und $b, d \in \mathbb{N}$: $\frac{a}{b} = \frac{c}{d} \longleftrightarrow ad = cb$.

Somit definieren wir eine Äquivalenzrelation auf $\mathbb{Z}\times\mathbb{N}$: $(a, b) \sim (c, d) :\Longleftrightarrow ad = cb$, und wir setzen $Q := (\mathbb{Z} \times \mathbb{N})_{/\sim}$.

Wir wollen nun zwei Verknüpfungen "+" und "$\cdot$" ("Addition" und "Multiplikation") auf Q wie folgt definieren:

$$[(a, b)] + [(c, d)] := [(ad + bc, bd)] \ , \quad [(a, b)] \cdot [(c, d)] := [(ac, bd)]$$

Beachten Sie, dass diese Operationen die üblichen Operation "+" und "$\cdot$" auf den rationalen Zahlen nachempfinden, wenn man Äquivalenzklassen als Brüche auffasst.

Wiederum zeigt man, dass diese Operationen wohldefiniert sind und die Assoziativ-, Kommutativ- und Distributivgesetze gelten.

Nun kann man noch $\frac{a}{b} := [(a, b)]$ setzen und erhält die rationalen Zahlen in bekannter Darstellung. Der Körper der rationalen Zahlen ist dann $\mathbb{Q} := (Q, +, \cdot)$.

1.8 Morphismen

Oftmals will man "Rechnungen" von einem Monoid, einer Gruppe, einem Ring usw. in ein anderes Objekt "der gleichen Art" transferieren. Hierzu kann man *Homomorphismen* (strukturerhaltende Abbildungen) benutzen. Eine andere Frage, die hiermit in engem Zusammenhang steht, ist, wann man zwei mathematische Strukturen als "strukturgleich" ansehen sollte.

Homomorphismen von Halbgruppen, Monoiden und Gruppen

Definition

- Seien H und H' zwei Halbgruppen. Ein *Homomorphismus von Halbgruppen* von H nach H' ist eine Abbildung $\varphi : H \longrightarrow H'$ mit $\varphi(a \circ_H b) = \varphi(a) \circ_{H'} \varphi(b)$ für alle $a, b \in H$.
- Seien M und M' Monoide mit neutralen Elementen e und e'. Ein *Homomorphismus von Monoiden* von M nach M' ist eine Abbildung

$\varphi : M \longrightarrow M'$ mit $\varphi(a \circ_M b) = \varphi(a) \circ_{M'} \varphi(b)$ für alle $a, b \in M$ und $\varphi(e) = e'$.

- Seien G und G' Gruppen. Dann sind G und G' insbesondere Monoide, und ein Homomorphismus von Gruppen von G nach G' ist ein Homomorphismus von G nach G' als Monoide.

Beispiel 1.50 Die Abbildung $\mathbb{N}_0 \longrightarrow \mathbb{N}_0,\ x \mapsto 2x$ ist ein Homomorphismus von Monoiden von $(\mathbb{N}_0, +)$ nach $(\mathbb{N}_0, +)$. Ebenso ist die "Null-Abbildung" $\mathbb{N}_0 \longrightarrow \mathbb{N}_0,\ x \mapsto 0$ ein Homomorphismus von Monoiden von $(\mathbb{N}_0, +)$ nach $(\mathbb{N}_0, +)$. Es ist auch ein Homomorphismus von Halbgruppen von $(\mathbb{N}_0, \cdot)$ nach $(\mathbb{N}_0, \cdot)$, aber es ist kein Homomorphismus von Monoiden von $(\mathbb{N}_0, \cdot)$ nach $(\mathbb{N}_0, \cdot)$.

Beispiel 1.51 Sei H eine Halbgruppe (resp. ein Monoid, resp. eine Gruppe) und $U \subseteq H$ eine Unterhalbgruppe (resp. ein Untermonoid, resp. eine Untergruppe). Dann ist die Inklusion $U \hookrightarrow H$ ein Homomorphismus von Halbgruppen (resp. von Monoiden, resp. von Gruppen).

Beispiel 1.52 Sei G eine (additiv geschriebene) abelsche Gruppe, und sei $U \subseteq G$ eine Untergruppe. In Abschnitt 1.5 haben wir die Faktorgruppe G/U eingeführt. Ich wiederhole, dass die Verknüpfung "+" auf G/U $[a]_U + [b]_U = [a+b]_U$ für alle $a, b \in G$ erfüllt. Damit ist die Abbildung $G \twoheadrightarrow G/U,\ a \mapsto [a]_U$ ein Homomorphismus von Gruppen.

Homomorphismen kann man verknüpfen, und man erhält wieder einen Homomorphismus:

Aussage 1.53 *Seien A, A', A'' Halbgruppen (resp. Monoide), und seien $\varphi : A \longrightarrow A'$ und $\psi : A' \longrightarrow A''$ Homomorphismen von Halbgruppen (resp. Monoiden). Dann ist $\psi \circ \varphi : A \longrightarrow A''$ ein Homomorphismus von Halbgruppen (resp. Monoiden).*

Beweis. Seien $a, b \in A$ beliebig. Dann ist $(\psi \circ \varphi)(a \circ_A b) = \psi(\varphi(a \circ_A b)) = \psi(\varphi(a) \circ_{A'} \varphi(b)) = \psi(\varphi(a)) \circ_{A''} \psi(\varphi(b)) = (\psi \circ \varphi)(a) \circ_{A''} (\psi \circ \varphi)(b)$.

Wenn es sich um Monoide handelt, gilt zusätzlich $(\psi \circ \varphi)(e) = \psi(\varphi(e)) = \psi(e') = e''$, wobei e, e' und e'' jeweils die neutralen Elemente in A, A' und A'' sind. □

Bemerkung Diese Aussage gilt selbstverständlich auch für Gruppen. Denn Homomorphismen von Gruppen sind ja per Definition Homomorphismen von Monoiden.

Lemma 1.54 *Seien M und M' Monoide und sei $\varphi : M \longrightarrow M'$ eine Abbildung mit $\varphi(a \circ_M b) = \varphi(a) \circ_{M'} \varphi(b)$ für alle $a, b \in M$, so dass $\varphi(e)$ ein Rechts- oder ein Links-Inverses hat. Dann ist $\varphi(e) = e'$, und folglich ist φ ein Homomorphismus von Monoiden.*

Beweis. Es ist $\varphi(e) = \varphi(e \circ_M e) = \varphi(e) \circ_{M'} \varphi(e)$. Multiplikation mit einem Rechts- bzw. Links-Inversen (von Rechts bzw. Links) liefert $e' = \varphi(e)$. □

Da in einer Gruppe jedes Element invertierbar ist, ergibt sich:

Aussage 1.55 *Seien G und G' Gruppen und sei $\varphi : G \longrightarrow G'$ mit $\varphi(a \circ_G b) = \varphi(a) \circ_{G'} \varphi(b)$ für alle $a, b \in G$ (d.h. φ ist ein Homomorphismus von Halbgruppen von G nach G'). Dann ist φ ein Homomorphismus von Gruppen.*

Aussage 1.56 *Seien M und M' Monoide, $\varphi : M \longrightarrow M'$ ein Homomorphismus von Monoiden, und sei $a \in M$ invertierbar. Dann ist $\varphi(a)$ invertierbar, und es ist $\varphi(a)^{-1} = \varphi(a^{-1})$.*

Beweis Es ist $e' = \varphi(e) = \varphi(a \circ_M a^{-1}) = \varphi(a) \circ_{M'} \varphi(a^{-1})$ sowie $e' = \varphi(e) = \varphi(a^{-1} \circ_M a) = \varphi(a^{-1}) \circ_{M'} \varphi(a)$. Damit folgt die Behauptung. □

Aussage 1.57 *Seien M und M' Monoide, $\varphi : M \longrightarrow M'$ ein Homomorphismus von Monoiden. Dann ist* Bild(φ) *ein Untermonoid von M'. Wenn M eine Gruppe ist, dann ist auch* Bild(φ) *eine Gruppe.*

Beweis. Bild(φ) ist offensichtlich abgeschlossen. Da es auch e' enthält, ist es ein Untermonid von M'. Die zweite Behauptung folgt aus der obigen Aussage. □

Bemerkung Sei $\iota : M \longrightarrow M'$ ein injektiver Homomorphismus von Monoiden. Oftmals "identifiziert" man dann M mit seinem Bild in M'. Dies bedeutet, dass man nicht zwischen $a \in M$ und $\iota(a) \in M'$ unterscheidet. Ein typisches Beispiel hierfür ist das folgende: Im vorherigen Abschnitt haben wir die Menge $Z := (\mathbb{N}_0 \times \mathbb{N}_0)_{/\sim}$ definiert; die Elemente in dieser Menge sind per Definition die ganzen Zahlen. Wir haben einen injektiven Homomorphismus $\iota : \mathbb{N}_0 \longrightarrow Z, a \mapsto [(a, 0)]_\sim$. Dies ist ein Homomorphismus von Monoiden bezüglich der Addition und der Multiplikation. Sicher macht es Sinn, natürliche Zahlen mit ihrem Bild (der entsprechenden ganzen Zahl) zu identifizieren.

Man muss aber aufpassen, mittels dieser "Identifizierungen" keine unsinnigen "Identitäten" abzuleiten. Ein Beispiel hierzu: Die injektive Abbildung $\iota : \mathbb{Z} \longrightarrow \mathbb{Z}\,,\ x \mapsto 2x$ ist ein Homomorphismus von Monoiden bezüglich der

Addition. Wenn wir nun x mit $\iota(x)$ "identifizieren", "erhalten" wir $x = 2x$ für alle $x \in \mathbb{Z}$!

Sei $\varphi : G \longrightarrow G'$ ein Homomorphismus von Gruppen.

Definition Der *Kern von* φ ist $\mathrm{Kern}(\varphi) := \{a \in G \mid \varphi(a) = e'\}$.

Aussage 1.58 $\mathrm{Kern}(\varphi)$ *ist eine Untergruppe von* G.

Beweis. Offensichtlich ist $e \in \mathrm{Kern}(\varphi)$, da $\varphi(e) = e'$.

Sei $a \in \mathrm{Kern}(\varphi)$. Dann ist $\varphi(a^{-1}) = \varphi(a)^{-1} = e'$ nach Aussage 1.56. Damit ist auch $a^{-1} \in \mathrm{Kern}(\varphi)$. Seien nun $a, b \in \mathrm{Kern}(\varphi)$. Dann ist $\varphi(a \circ_G b) = \varphi(a) \circ_{G'} \varphi(b) = e' \circ_{G'} e' = e'$. Damit ist auch $a \circ_G b \in \mathrm{Kern}(\varphi)$. □

Aussage 1.59 *Der Homomorphismus* φ *ist genau dann injektiv, wenn* $\mathrm{Kern}(\varphi) = \{e\}$.

Beweis. Offensichtlich ist $e \in \mathrm{Kern}(\varphi)$. Wenn nun φ injektiv ist, gilt insbesondere $\#\,\mathrm{Kern}(\varphi) = \#\varphi^{-1}(\{e'\}) \leq 1$, also $\mathrm{Kern}(\varphi) = \{e\}$.

Sei andererseits $\mathrm{Kern}(\varphi) = \{e\}$, und seien $a, b \in G$ mit $\varphi(a) = \varphi(b)$. Dann ist also $\varphi(a \circ_G b^{-1}) = \varphi(a) \circ_{G'} \varphi(b^{-1}) = \varphi(a) \circ_{G'} \varphi(b)^{-1} = e'$, also $a \circ_G b^{-1} \in \mathrm{Kern}(\varphi)$ und somit $a \circ_G b^{-1} = e$. Damit ist $a = b$. □

Homomorphismen von Ringen und Körpern

Definition

- Seien R und R' Ringe. Ein *Homomorphismus von Ringen* von R nach R' ist eine Abbildung $\varphi : R \longrightarrow R'$ mit $\varphi(a +_R b) = \varphi(a) +_{R'} \varphi(b)$ und $\varphi(a \cdot_R b) = \varphi(a) \cdot_{R'} \varphi(b)$ für alle $a, b \in R$ sowie $\varphi(1_R) = 1_{R'}$.

- Seien K und K' Körper. Dann sind K und K' insbesondere Ringe. Ein *Homomorphismus von Körpern* von K nach K' ist ein Homomorphismus von Ringen von K nach K'.

Bemerkung Ein Homomorphismus von Ringen von R nach R' ist also eine Abbildung $R \longrightarrow R'$, welche ein Homomorphismus der abelschen Gruppe $(R, +)$ sowie des Monoids $(R, \cdot)$ ist.

Beispiel 1.60 Analog zu Beispiel 1.51 gilt: Sei R ein Ring und $U \subseteq R$ ein Unterring. Dann ist die Inklusion $U \hookrightarrow R$ ein Homomorphismus von Ringen.

Beispiel 1.61 Sei n eine natürliche Zahl. Dann ist die "kanonische" Abbildung $\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}$ (gegeben durch $z \mapsto [z]_n$) ein Homomorphismus von Ringen.

Wiederum gilt:

Aussage 1.62 *Seien A, A', A'' Ringe (resp. Körper), und seien $\varphi : A \longrightarrow A'$ und $\psi : A' \longrightarrow A''$ Homomorphismen von Ringen (resp. Körpern). Dann ist $\psi \circ \varphi : A \longrightarrow A''$ ein Homomorphismus von Ringen (resp. Körpern).*

Der *Beweis* ist analog zum Beweis von Aussage 1.53.

Aussage 1.63 *Ein Homomorphismus von Körpern ist immer injektiv.*

Beweis. Sei $\varphi : K \longrightarrow L$ ein Homomorphismus von Körpern. Es genügt zu zeigen, dass $\text{Kern}(\varphi) = \{0\}$ gilt. Mit anderen Worten: Wir müssen zeigen, dass für alle $a \in K - \{0\} = K^*$ $\varphi(a) \neq 0$ gilt. Sei hierfür $a \in K^*$. Dann gilt $a \cdot a^{-1} = 1$. Somit gilt $\varphi(a \cdot a^{-1}) = \varphi(a) \cdot \varphi(a^{-1}) = \varphi(a) \cdot \varphi(a)^{-1} = 1$. Also ist $\varphi(a) \in L^* = L - \{0\}$. □

Isomorphismen, Endomorphismen und Automorphismen

Definition Sei $\varphi : A \longrightarrow A'$ ein Homomorphismus von Halbgruppen (resp. Monoiden, resp. Gruppen, resp. Ringen, resp. Körpern). Dann heißt φ *Isomorphismus* wenn es einen Homomorphismus $\psi : A' \longrightarrow A$ mit $\psi \circ \varphi = \text{id}_A, \varphi \circ \psi = \text{id}_{A'}$ von Halbgruppen (resp. Monoiden, resp. Gruppen, resp. Ringen, resp. Körpern) gibt. Wenn es einen Isomorphismus $A \longrightarrow A'$ von Halbgruppen (resp. Monoiden, resp. Gruppen, resp. Ringen. resp. Körpern) gibt, heißen A und A' *isomorph* (als Halbgruppen (resp. Monoide, resp. Gruppen, resp. Ringe, resp. Körper)).

Aussage 1.64 *Ein Homomorphismus (von Halbgruppen, Monoiden, Gruppen, Ringen oder Körpern) ist genau dann ein Isomorphismus, wenn er bijektiv ist.*

Beweis. Ein Isomorphismus ist offensichtlich bijektiv (es gibt eine Umkehrabbildung). Sei andererseits $\varphi : A \longrightarrow A'$ ein bijektiver Homomorphismus, und sei $\varphi^{-1} : A' \longrightarrow A$ die Umkehrabbildung. Zu zeigen ist nun, dass φ^{-1} auch ein Homomorphismus ist.

Wir betrachten zuerst den Fall von Halbgruppen. Seien $a', b' \in A'$ mit $a' = \varphi(a), b' = \varphi(b)$. Dann ist $\varphi^{-1}(a' \circ_{A'} b') = \varphi^{-1}(\varphi(a) \circ_{A'} \varphi(b)) = \varphi^{-1}(\varphi(a \circ_A b)) = a \circ_A b = \varphi^{-1}(a') \circ_A \varphi^{-1}(b')$, was zu zeigen war.

Sei nun φ ein Homomorphismus von Monoiden. Dann ist φ^{-1} ein Homomorphismus von Halbgruppen, und es gilt $\varphi^{-1}(e') = \varphi^{-1}(\varphi(e)) = e$, d.h. φ^{-1} ist ein Homomorphismus von Monoiden.

Für Homomorphismen von Gruppen ist nun nichts mehr zu zeigen, und die Aussagen für Ringe und Körper folgen sofort aus den obigen Argumenten (angewandt auf Addition und Multiplikation). □

Definition Ein Homomorphismus $A \longrightarrow A$ heißt *Endomorphismus* von A. Ein Endomorphismus, der zusätzlich ein Isomorphismus ist, heißt *Automorphismus*.

Definition Seien wie oben A und A' beides Halbgruppen, Monoide, Gruppen, Ringe oder Körper. Dann wird die Menge der Homomorphismen von A nach A' mit $\mathrm{Hom}(A, A')$ und die Menge der Isomorphismen von A nach A' mit $\mathrm{Iso}(A, A')$ bezeichnet. Sei nun A eine Halbgruppe, ein Monoid, eine Gruppe, ein Ring oder ein Körper. Dann wird die Menge der Endomorphismen von A mit $\mathrm{End}(A)$ und die Menge der Automorphismen mit $\mathrm{Aut}(A)$ bezeichnet.

Bemerkung Wenn A, A' ein Monoide sind, sind sie auch eine Halbgruppen. Nun ist nicht notwendigerweise jeder Homomorphismus von A nach A' als Halbgruppen ein Homomorphismus von A nach A' als Monoide. Beispiel: Die Abbildung $\mathbb{N}_0 \longrightarrow \mathbb{N}_0, x \mapsto 0$ ist ein Endomorphismus der Halbgruppe $(\mathbb{N}_0, \cdot)$ aber kein Endomorphismus des Monoids $(\mathbb{N}_0, \cdot)$. Insbesondere sollte man also aufpassen, wenn man die Notation $\mathrm{Hom}(A, A')$ etc. benutzt und explizit angeben, ob man A, A' als Monoid oder als Halbgruppe "betrachtet".

Ein ähnliches Problem tritt bei Ringen auf: Wenn man die Multiplikation "vergißt", "wird" jeder Ring eine abelsche Gruppe. Wenn R ein Ring ist, ist aber nicht notwendigerweise jeder Homomorphismus der abelschen Gruppe $(R, +)$ auch ein Endomorphismus von R als Ring.

Aussage 1.65 *Sei A eine Halbgruppe, ein Monoid, eine Gruppe, ein Ring oder ein Körper. Dann ist* $\mathrm{End}(A)$ *bezüglich der normalen Verknüpfung von Abbildungen ein Monoid, und* $\mathrm{Aut}(A)$ *ist die Gruppe der invertierbaren Elemente in* $\mathrm{End}(A)$ *(siehe Beispiel 1.37).*

Bemerkung Beachten Sie, dass diese Aussage analog zu den Aussagen, dass $\mathrm{Abb}(X, X)$ ein Monoid und $\mathrm{Perm}(X)$ eine Gruppe ist, ist (siehe Beispiel 1.39).

Seien G und G' (additiv geschriebene) abelsche Gruppen. In Abschnitt 1.5 (Unterabschnitt über Produkte) haben wir eine Verknüpfung auf $\mathrm{Abb}(G, G')$ definiert. Da G' eine abelsche Gruppe ist, ist $\mathrm{Abb}(G, G')$ mit dieser Verknüpfung auch eine abelsche Gruppe. Ich wiederhole, dass die Verknüpfung wie folgt definiert ist: Für $\varphi, \psi \in \mathrm{Abb}(G, G')$ ist $\varphi + \psi : G \longrightarrow G'$ durch

$$(\varphi + \psi)(a) := \varphi(a) + \psi(a)$$

definiert. (Diese Definition ist identisch zu der in Abschnitt 1.5, nur die Notation ist anders). Beachten Sie, dass wir bisher nur die Gruppenstruktur von G' und nicht die von G ausgenutzt haben.

Offensichtlich ist $\mathrm{Hom}(G, G')$ eine Teilmenge von $\mathrm{Abb}(G, G')$ und sogar eine Untergruppe (nachrechnen!). Damit ist $\mathrm{Hom}(G, G')$ also auch eine abelsche Gruppe.

Insbesondere ist also $\mathrm{End}(G)$ eine abelsche Gruppe mittels der soeben definierten Addition von Homomorphismen. Wir wissen auch schon, dass $\mathrm{End}(G)$ ein Monoid bezüglich der Verknüpfung von Abbildungen ist. Außerdem gelten die Distributivgesetze (für $\varphi, \chi, \psi \in \mathrm{End}(G)$)

$$\varphi \circ (\chi + \psi) = \varphi \circ \chi + \varphi \circ \psi \qquad (\chi + \psi) \circ \varphi = \chi \circ \varphi + \psi \circ \varphi \; .$$

(Nachrechnen!) Damit ist $(\mathrm{End}(G), +, \circ)$ ein Ring, genannt der *Endomorphismenring* von G. Beachten Sie, dass dieser Ring nicht notwendigerweise kommutativ ist.

Strukturtransport

Sei nun A eine Halbgruppe, sei X eine Menge, und sei $f : A \longrightarrow X$ eine *bijektive* Abbildung. Wir definieren wie folgt auf X eine Verknüpfung $*$: Für $x, y \in X$ definieren wir

$$x * y := f(f^{-1}(x) \circ f^{-1}(y)) \; .$$

Damit gilt für alle $a, b \in A$:

$$f(a) * f(b) = f(f^{-1}(f(a)) \circ f^{-1}(f(b))) = f(a \circ b) \; . \tag{1.4}$$

Man sieht leicht, dass die Verknüpfung $*$ auf X assoziativ ist (nachrechnen!), d.h. X ist mit $*$ eine Halbgruppe. Aus (1.4) folgt nun, dass f ein Homomorphismus ist. Da f auch bijektiv ist, ist f somit ein Isomorphismus, und A und $(X, *)$ sind isomorph als Halbgruppen.

Man sieht nun leicht: Wenn A ein Monoid ist, dann ist auch $(X, *)$ ein Monoid (mit neutralem Element $f(e)$), und wenn A eine Gruppe ist, so auch $(X, *)$.

Analoge Aussagen gelten, wenn A ein Ring oder ein Körper ist.

Die soeben angewandte Methode, eine Verknüpfung auf X zu definieren und f zu einem Isomorphismus zu machen, nennt man *Strukturtransport*.

Beispiel 1.66 Sei $n > 1$. Die Zahlen $0, 1, \ldots, n-1$ bilden ein Repräsentantensystem bezüglich der Äquivalenzrelation "Kongruenz modulo n". Wir erhalten somit mittels "Strukturtransport" zwei Verknüpfungen "$\oplus_n$" und "$\odot_n$" auf $\{0, 1, \ldots, n\}$, so dass für alle $a, b \in \{0, 1, \ldots, n-1\}$ gilt: $[a \oplus_n b]_n = [a]_n + [b]_n = [a+b]_n$ und $[a \odot_n b]_n = [a]_n \cdot [b]_n = [a \cdot b]_n$. Da $\mathbb{Z}/n\mathbb{Z}$ ein Ring ist, ist somit $(\{0, 1, \ldots, n-1\}, \oplus_n, \odot_n)$ auch ein Ring.

Wahrscheinlich kennen Sie die Verknüpfungen "$\oplus_n$" und "$\odot_n$" aus der Schule: Sie beschreiben die Arithmetik "modulo n": $a \oplus_n b$ ist der Rest von $a + b$ bei der Division mit n, analog ist $a \odot_n b$ der Rest von $a \cdot b$ bei der Division mit n.

Sei für eine natürliche Zahl $n > 1$ und $a \in \mathbb{Z}$ $\mathrm{mod}(a, n)$ die kleinste Zahl in $\mathbb{N}_0$, die kongruent zu a modulo n ist. Dann ist für $a \in \mathbb{Z}$ $\mathrm{mod}(a, n)$ der Repräsentant von $[a]_n$ in $0, 1, \ldots, n-1$. Es ist also $a \oplus_n b = \mathrm{mod}(a + b, n)$ und $a \odot_n b = \mathrm{mod}(a \cdot b, n)$.

Kategorien (Diskussion)

Oben haben wir einer Vielzahl verschiedener Arten mathematischer Objekte entsprechende Homomorphismen zugeordnet. Dabei fällt auf, dass es in diesem Kontext einige Aussagen gibt, die gelten, egal ob man nun von Homomorphismen von Halbgruppen, Monoiden, Gruppen, Ringen oder Körpern redet. Beispiele hierfür sind Aussagen 1.53 und 1.62. Diese kann man wie folgt knapp zusammenfassen:

Seien A, A', A'' "Objekte von gleichem Typ", und seien $\varphi : A \longrightarrow A'$ und $A' \longrightarrow A''$ Homomorphismen dieser Objekte. Dann ist auch $\psi \circ \varphi : A \longrightarrow A''$ ein Homomorphismus dieser Objekte.

Dies ist natürlich keine mathematisch rigorose Aussage, da nicht klar ist, was "Objekte von gleichem Typ" und "Homomorphismen dieser Objekte" sein sollen.

Dies kann man jedoch präzise machen, indem man den Begriff einer *Kategorie* einführt. Der Begriff der Kategorie ist ziemlich abstrakt, ich versuche eine intuitive Annäherung zu geben.

Betrachten Sie als Beispiel alle Mengen zusammen mit allen Abbildungen. Wie wir zu Beginn gesehen haben, macht es keinen Sinn, von der *Menge aller Menge* zu reden. Um trotzdem alle Mengen zusammenfassen zu können, führt man den Begriff einer *Klasse* ein und spricht z.B. von der *Klasse* aller

Mengen. Die Elemente der Klasse (in diesem Fall die Mengen) heißen nun *Objekte*.

Eine Kategorie besteht nun aus einer Klasse zusammen mit Folgendem: Zu je zwei Objekten A, A' der Klasse gibt es eine Menge $\mathrm{Mor}(A, A')$ von sogenannten *Morphismen*. Diese Morphismen schreibt man suggestiv wie Abbildungen (d.h. wenn $\varphi \in \mathrm{Mor}(A, A')$, dann schreibt man $\varphi : A \longrightarrow A'$), obwohl es nicht notwendigerweise Abbildungen sein müssen.

Dabei sollen gewisse Eigenschaften gelten, die für Abbildungen offensichtlich sind. Z.B. soll man Morphismen "hintereinanderschalten" können.

Wir haben nun schon einige Kategorien gesehen, nämlich die Kategorien der Mengen, der Halbgruppen, der Monoide, der Gruppen, der Ringe und der Körper. Im Fall der Mengen sind die Morphismen per Definition die Abbildungen, und ansonsten sind die Morphismen die oben definierten Homomorphismen. Außerdem macht es noch Sinn, von den Kategorien der abelschen Halbgruppen, der abelschen Monoide, der abelschen Gruppen und der kommutativen Ringe zu sprechen.

1.9 Polynome

Sei im Folgenden R ein kommutativer Ring.

Intuitiv ist ein Polynom über R in der Unbestimmten X ein "formaler Ausdruck" der Form $\sum_{i=0}^{d} a_i X^i = a_d X^d + a_{d-1} X^{d-1} \cdots + a_1 X + a_0$ mit $a_i \in R$.

Solche "formalen Ausdrücke" kann man dann addieren und multiplizieren, wobei die üblichen Rechenregeln (Assoziativität, Distributivität, Kommutativität) gelten und X als Unbestimmte aufgefasst wird.

Beispielsweise ist $(1 + X + X^2) \cdot (1 - X) = 1 - X^3$.

Aber inwiefern kann ein "formaler Ausdruck" der angegebenen Form ein wohldefiniertes mathematisches Objekt sein? Sicher sollte man zwischen den mathematischen Objekten und seiner symbolischen Darstellung unterscheiden. Wie kann man also definieren, was Polynome über R sind?

Sie erinnern sich, wie wir bei den ganzen und den rationalen Zahlen vorgegangen sind: Wir haben nicht definiert, was eine ganze oder einer rationale Zahl "ist", sondern wir haben die Menge der ganzen bzw. der rationalen Zahlen definiert, und eine ganze bzw. rationale Zahl ist dann ein Element aus dieser Menge. So gehen wir auch hier vor.

Wir haben die folgende Wunschliste: Wir wollen einen kommutativen Ring, genannt $R[X]$, finden, der ein bestimmtes Element X enthält, so dass

- R ein Unterring von $R[X]$ ist

- sich jedes Element $p \in R[X]$ $(p \neq 0)$ in eindeutiger Weise in der Form $p = \sum_{i=0}^{d} a_i X^i$ mit $a_i \in R$ und $a_d \neq 0$ schreiben läßt.

Beachten Sie, dass die zweite Eigenschaft besagt, dass jedes Polynom eindeutig durch das Tupel $(a_0, \ldots, a_d)$ der Koeffizienten gegeben ist. Da Polynome beliebig lang werden können, liegt es nahe nicht Tupel sondern "Folgen" mit Werten in R zu betrachten,[15] allerdings mit zwei Modifikationen: Erstens sollte das erste Folgenglied nicht a_1 sondern a_0 heißen. Zweitens sollten wir nur solche Folgen betrachten, für die es nur endlich viele Folgenglieder gibt, die $\neq 0$ sind. (Weil Polynome auch nur endliche viele Terme enthalten.)

Definition und einfache Eigenschaften

Wir beginnen mit der Menge $R^{\mathbb{N}_0}$. Die Elemente hierin sind die Abbildungen $\mathbb{N}_0 \longrightarrow R$ bzw. die Familien von Elementen aus R über der Indexmenge $\mathbb{N}_0$. Diese Familien schreiben wir wie üblich in der Form $a = (a_n)_{n\in\mathbb{N}_0}$. Nun ist $R^{\mathbb{N}_0}$ mit der "komponentenweisen Verknüpfung" eine abelsche Gruppe (siehe den Unterabschnitt über Produkte in Abschnitt 1.5).

Ferner definieren wir für $r \in R$ und $a = (a_n)_{n\in\mathbb{N}_0} \in R^{\mathbb{N}_0}$: $r{\cdot}a := (r{\cdot}a_n)_{n\in\mathbb{N}_0}$.

Für $i, j \in \mathbb{N}_0$ definieren wir das so genannte *Kronecker-Delta* durch

$$\delta_{i,j} := \begin{cases} 1 \text{ falls } i = j \\ 0 \text{ falls } i \neq j \end{cases} ,$$

und wir definieren $e_j := (\delta_{i,j})_{i\in\mathbb{N}_0} \in R^{\mathbb{N}_0}$. Wir betrachten nun die folgende Teilmenge von $R^{\mathbb{N}_0}$:

$$P := \{a = (a_n)_{n\in\mathbb{N}_0} \in R^{\mathbb{N}_0} \mid \text{ es gibt nur endlich viele } n \in \mathbb{N}_0 \text{ mit } a_n \neq 0\}$$

Dies ist eine Untergruppe von $R^{\mathbb{N}_0}$. (Warum?) Auf dieser Gruppe P wollen wir eine Multiplikation definieren und ein Element $X \in P$ identifizieren, so dass P zusammen mit X die gewünschten Eigenschaften hat.

Beachten Sie: Wenn $a = (a_n)_{n\in\mathbb{N}} \in P$, dann gibt es ein $d \in \mathbb{N}_0$, so dass $a_n = 0$ für $n > d$. Mit so einem d gilt dann

$$a = \sum_{n=0}^{d} a_n e_n \ . \tag{1.5}$$

Dies schreiben wir auch in der Form

$$a = \sum_{n\in\mathbb{N}_0} a_n e_n \ , \tag{1.6}$$

[15] In Beispiel 1.9 haben wir Folgen als Abbildungen $\mathbb{N} \longrightarrow \mathbb{R}$ definiert. Mit einer *Folge mit Werten in einem Ring R* meinen wir eine Abbildung $\mathbb{N} \longrightarrow R$.

wobei zu beachten ist, dass in dieser Summe immer nur endlich viele Terme $\neq 0$ sind.

Wir definieren nun eine Multiplikation auf P durch

$$(a_n)_{n\in\mathbb{N}_0} \cdot (b_n)_{n\in\mathbb{N}_0} := (\sum_{i=0}^{n} a_i b_{n-i})_{n\in\mathbb{N}_0} \ . \tag{1.7}$$

Beachten Sie, dass das Ergebnis wieder in P liegt, da nur endlich viele a_n und b_n von 0 verschieden sind.

Beachten Sie weiter: Wenn wir (1.5) in (1.7) einsetzen, erhalten wir

$$(\sum_{n\in\mathbb{N}_0} a_n e_n) \cdot (\sum_{n\in\mathbb{N}_0} b_n e_n) = \sum_{n\in\mathbb{N}_0} (\sum_{i=0}^{n} a_i b_{n-i}) e_n \tag{1.8}$$

$$= \sum_{k\in\mathbb{N}_0} \sum_{\ell\in\mathbb{N}_0} a_k b_\ell \, e_{k+\ell} \tag{1.9}$$

Insbesondere ist

$$e_k \cdot e_\ell = e_{k+\ell} \ , \tag{1.10}$$

für alle $i, j \in \mathbb{N}$, und hieraus folgt insbesondere

$$e_1^n = e_n$$

für alle $n \in \mathbb{N}$. Außerdem sieht man, dass e_0 ein neutrales Element bezüglich der Multiplikation ist. Wie üblich setzen wir also $1 = 1_P := e_0$.

Wir setzen nun $X := e_1$. Dann ist also $e_n = X^n$. Wenn wir dies in (1.6) einsetzen, erhalten wir

$$a = \sum_{n\in\mathbb{N}_0} a_n X^n$$

für alle $a \in P$.

Seien nun $a, b \in P$. Dann ist

$$a \cdot b = (\sum_{n\in\mathbb{N}_0} a_n X^n) \cdot (\sum_{n\in\mathbb{N}_0} b_n X^n) =$$

$$\sum_{n\in\mathbb{N}_0} (\sum_{i=0}^{n} a_i b_{n-i}) X^n = \sum_{k\in\mathbb{N}_0} \sum_{\ell\in\mathbb{N}_0} a_k b_k X^{k+\ell}$$

nach (1.8). Man sieht leicht, dass die Distributivgesetze gelten. Damit ist P ein Ring. Außerdem ist $\iota : R \hookrightarrow P, r \mapsto re_0$ ein injektiver Ringhomomorphismus.

Damit ist P zusammen mit X ein Ring wie gewünscht; wir setzen also $R[X] := P$. □

Sei nun $p = p(X) = \sum_{i=0}^{d} a_i X^i \in R[X]$ ein Polynom. Sei $r \in R$. Dann können wir r in p "einsetzen" bzw. – was dasselbe ist – p an r "auswerten". Wir erhalten

$$p(r) := \sum_{i=0}^{d} a_i r^i \in R \; .$$

Wenn wir nun die Polynome variieren, erhalten wir eine Abbildung

$$R[X] \longrightarrow R \; , \; p(X) = \sum_{i=0}^{d} a_i X^i \mapsto p(r) = \sum_{i=0}^{d} a_i r^i \; .$$

Man kann zeigen, dass diese Abbildung ein Ringhomomorphismus ist (Übungsaufgabe).

Ferner können wir ein Polynom $p(X) \in R[X]$ fixieren und die Ringelemente variieren. Wir erhalten dann die *Polynomfunktion* zu $p(X)$:

$$R \longrightarrow R \; , \; r \mapsto p(r) \; .$$

Aus der Schule kennen Sie die Polynomfunktionen $\mathbb{R} \longrightarrow \mathbb{R}$ (die wahrscheinlich "Polynome" genannt wurden).

Wenn wir nun auch noch die sowohl die Polynome als auch die Ringelemente variieren, erhalten wir eine Abbildung

$$R[X] \longrightarrow \mathrm{Abb}(R, R) \; , \; p(X) \mapsto (r \mapsto p(r)) \; .$$

Wenn z.B. $R = \mathbb{R}$ (aber auch $R = \mathbb{Z}$), ist diese Abbildung injektiv, d.h. ein Polynom ist durch die entsprechende Polynomfunktion eindeutig bestimmt (ein Beweis hiervon ist später recht einfach). Dies ist aber *nicht* immer der Fall. Sei z.B. $R = \mathbb{Z}/2\mathbb{Z} = \mathbb{F}_2$, und sei $p_1(X) := 0, p_2(X) := X^2 + X$.

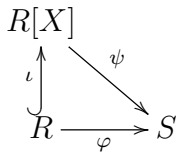
Nun ist sowohl die Polynomfunktion zu $p_1(X)$ als auch die Polynomfunktion zu $p_2(X)$ die Null-Abbildung, aber es ist $p_1(X) \neq p_2(X)$.

Begründung: Es ist klar, dass die Polynomfunktion zu $p_1(X)$ durch $r \mapsto 0$ gegeben ist. Zu $p_2(X)$: Der Körper $\mathbb{F}_2$ hat nur zwei Elemente, $0 = [0]_2$ und $1 = [1]_2$, und es ist $p_2(0) = 0 + 0 = 0$ und $p_2(1) = 1 + 1 = 0$.

Sei weiterhin $p(X) = \sum_{i=0}^{d} a_i X^i \in R[X]$, sei S ein kommutativer Ring, und sei $\varphi : R \longrightarrow S$ ein Homomorphismus von Ringen und $s \in S$. Wir definieren $p(s) := \sum_{i=0}^{d} \varphi(a_i) s^i \in S$. Man kann zeigen:

Aussage 1.67 *Die Abbildung $\psi : R[X] \longrightarrow S$, $p(X) \mapsto p(s)$ ist ein Homomorphismus von Ringen, und es ist der einzige Homomorphismus*

$\psi : R[X] \longrightarrow S$ für den $\psi(X) = s$ gilt und das Diagramm



kommutativ ist.

(Übungsaufgabe)

Definition Sei $p(X) = \sum_{i=0}^{d} a_i X^i$ mit $a_d \neq 0$. Dann heißt d der *Grad* von $p(X)$, Bezeichnung $\mathrm{Grad}(p(X))$. Der Grad des Nullpolynoms wird mit $-\infty$ definiert.

Ein Polynom $p(X) = \sum_{i=0}^{d} a_i X^i$ mit $a_d = 1$ heißt *normiert*.

Polynome über Körpern

Sei ab nun $R = K$ ein Körper.

Definition Seien $a(X), b(X) \in K[X]$. Dann sagen wir, dass $a(X)$ das Polynom $b(X)$ *teilt* und schreiben $a(X)|b(X)$, wenn es ein Polynom $c(X) \in K[X]$ mit $b(X) = a(X) \cdot c(X)$ gibt.

Definition Ein Polynom $p(X) \in K[X]$ mit $\mathrm{Grad}(p(X)) \geq 1$ heißt *irreduzibel*, wenn es nicht in der Form $p(X) = p_1(X) \cdot p_2(X)$ mit $p_1(X), p_2(X) \in K[X]$ und $\mathrm{Grad}(p_1(X)) \geq 1$ und $\mathrm{Grad}(p_2(X)) \geq 1$ geschrieben werden kann.

Aus der Schule kennen Sie die *Polynomdivision*, die große Ähnlichkeit mit der Division mit Rest von ganzen Zahlen hat. Auf Grundlage der Polynomdivision kann man beweisen:

Aussage 1.68 *Seien $a(X), b(X) \in K[X]$ zwei Polynome mit $b(X) \neq 0$. Dann gibt es eindeutig bestimmte Polynome $p(X), r(X) \in K[X]$ mit $\mathrm{Grad}(r(X)) < \mathrm{Grad}(b(X))$ und*

$$a(X) = p(X) \cdot b(X) + r(X) \; .$$

Mittels dieser Aussage kann man viele Aussagen über den Ring $\mathbb{Z}$ auf den Ring $K[X]$ "übertragen" (wobei die normierten Polynome die Rolle der natürlichen Zahlen und die irreduziblen Polynome die Rolle der Primzahlen einnehmen). Wir kommen hierauf ausführlich im zweiten Semester zurück.

Kapitel 2

Grundlagen der Linearen Algebra

2.1 Der Standardvektorraum

Wir fixieren für das gesamte Kapitel einen Körper K. Die Elemente von K werden auch *Skalare* genannt.

Spaltenvektoren

Die Elemente aus K^n (die "n-Tupel") werden von nun ab als "Spaltenvektoren" geschrieben, beispielsweise so:

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Dementsprechend nennen wir die Elemente von K^n auch *Vektoren*.

Man definiert eine Addition von Vektoren in K^n komponentenweise:

$$+ : K^n \times K^n \longrightarrow K^n , \quad (\underline{x}, \underline{y}) \mapsto \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Nun ist $(K, +)$ eine abelsche Gruppe (siehe Unterabschnitt über Produkte in Abschnitt 1.5).

Außerdem definiert man eine so genannte *Skalarmultiplikation*:

$$\cdot : K \times K^n \longrightarrow K^n , \quad (a, \underline{x}) \mapsto \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}$$

Die Menge K^n mit den soeben definierten Verknüpfungen heißt der *n-dimensionale Standardvektorraum*. (Bis jetzt ist das nur ein Name, der sich aber bald erklären wird.)

Definition Wir setzen $\underline{0} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$; dies ist der *Nullvektor*. Außerdem setzen wir

$$\underline{e}_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i - \text{te Zeile} ;$$

dies ist der i-te *Standardvektor*.

Ferner setzen wir $K^0 := \{0\}$ (mit der "trivialen" Addition $0 + 0 = 0$ und der "trivialen" Skalarmultiplikation $a \cdot 0 = 0$ für $a \in K$).

Bemerkung Die Addition ist (wie schon gesagt) eine Verknüpfung auf K^n. Die Skalarmultiplikation ist hingegen keine Verknüpfung auf K^n.

Lineare Abbildungen und Matrizen

Seien bis zum Ende des Kapitels $r, m, n \in \mathbb{N}_0$.

Definition Eine *lineare Abbildung* von K^n nach K^m ist eine Abbildung $\varphi : K^n \longrightarrow K^m$, so dass

- $\forall \underline{x}, \underline{y} \in K^n : \varphi(\underline{x} + \underline{y}) = \varphi(\underline{x}) + \varphi(\underline{y})$
- $\forall a \in K \, \forall \underline{x} \in K^n : \varphi(a\underline{x}) = a\varphi(\underline{x})$.

Bemerkungen

- Eine lineare Abbildung von K^n nach K^m ist also insbesondere ein Gruppenhomomorhismus von $(K^n, +)$ nach $(K^m, +)$.
- Wenn $\varphi : K^r \longrightarrow K^n$ und $\psi : K^n \longrightarrow K^m$ lineare Abbildungen sind, dann ist auch $\psi \circ \varphi : K^r \longrightarrow K^m$ eine lineare Abbildung.
- Wenn $\varphi : K^n \longrightarrow K^m$ eine bijektive lineare Abbildung ist, dann ist auch $\varphi^{-1} : K^m \longrightarrow K^n$ eine lineare Abbildung.[1]

Definition Eine $m \times n$-*Matrix* über K ist eine Familie von Elementen aus K über der Indexmenge $\{1, \ldots, m\} \times \{1, \ldots, n\}$.

Eine $m \times n$-Matrix ist also ein Element aus $K^{\{1,\ldots,m\}\times\{1,\ldots,n\}}$ bzw. eine Abbildung $\{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow K$.

Eine Matrix schreibt man oft auch als rechteckiges Schema mit "Klammern darum":

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{2,2} & \cdots & a_{m,n} \end{pmatrix}$$

wobei $a_{i,j} \in K$. Eine andere übliche Schreibweise, die wir benutzen, ist $((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n}$ oder kürzer $((a_{i,j}))_{i,j}$. Matrizen werden meist mit großen Buchstaben bezeichnet.

Wir schreiben $K^{m\times n}$ für $K^{\{1,\ldots,m\}\times\{1,\ldots,n\}}$, die Menge der $m \times n$-Matrizen. Andere oft benutzte Schreibweisen sind $\mathcal{M}_{m,n}(K)$ oder $\mathcal{M}_{m\times n}(K)$. Beachten Sie, dass auch $m = 0$ oder $n = 0$ sein kann. In diesem Fall hat die Matrix allerdings keinen Eintrag ("leere Matrix").

Oftmals geben wir uns Spaltenvektoren $\underline{a}_1, \ldots, \underline{a}_n$ vor und betrachten die Matrix, deren j-te Spalte gleich $\underline{a}_j$ ist. Wir schreiben dann $A = (\underline{a}_1 | \cdots | \underline{a}_n)$.

Definition Für $A = (\underline{a}_1 | \cdots | \underline{a}_n) \in K^{m\times n}$ und $\underline{x} \in K^n$ definieren wir:

$$A \cdot \underline{x} := \sum_{j=1}^{n} x_j \underline{a}_j$$

Damit ist also $A \cdot \underline{x} = \begin{pmatrix} \sum_{j=1}^n a_{1,j} x_j \\ \vdots \\ \sum_{j=1}^n a_{m,j} x_j \end{pmatrix}$.

[1] Wir werden in Aussage 2.20 sehen, dass dann auch $n = m$ gilt.

Beispiel 2.1 Sei $A \in K^{m\times n}$. Dann ist die Abbildung

$$\Lambda_A : K^n \longrightarrow K^m \,, \underline{x} \mapsto A\underline{x}$$

linear.

Wir haben somit jeder $m \times n$-Matrix eine lineare Abbildung von K^n nach K^m zugeordnet. Umgekehrt ordnen wir wie folgt jeder linearen Abbildung von K^n nach K^m eine Matrix in $K^{m\times n}$ zu:

Definition Sei $\varphi : K^n \longrightarrow K^m$ eine lineare Abbildung. Dann ist die zu φ *assoziierte Matrix* $M(\varphi)$ wie folgt definiert:

$$M(\varphi) := (\varphi(\underline{e}_1)|\cdots|\varphi(\underline{e}_n))$$

Hier ist $\underline{e}_j$ der j-te Standardvektor.

Aussage 2.2 *Die Zuordnungen $\varphi \mapsto M(\varphi)$ und $A \mapsto \Lambda_A$ definieren zueinander inverse Bijektionen zwischen der Menge der linearen Abbildungen von K^n nach K^m und der Menge der $m \times n$-Matrizen $K^{m\times n}$.*

Mit anderen Worten: Es gilt

$$\Lambda_{M(\varphi)} = \varphi \;\; \textit{und} \;\; M(\Lambda_A) = A$$

für alle linearen Abbildungen $\varphi : K^m \longrightarrow K^n$ und alle $m \times n$-Matrizen A.

Beweis. Sei zunächst A eine $m \times n$-Matrix, und sei $A = (\underline{a}_1|\cdots|\underline{a}_n)$. Dann ist $M(\Lambda_A) = (\Lambda_A(\underline{e}_1)|\cdots|\Lambda_A(\underline{e}_n)) = (A\underline{e}_1|\cdots|A\underline{e}_n) = (\underline{a}_1|\cdots|\underline{a}_n) = A$, was zu zeigen war.

Sei nun $\varphi : K^n \longrightarrow K^m$ eine lineare Abbildung. Zu zeigen ist, dass für alle $\underline{x} \in K^n$ gilt: $\Lambda_{M(\varphi)}(\underline{x}) = \varphi(\underline{x})$.

Sei hierfür $\underline{x} \in K^n$ beliebig. Dann gilt $\Lambda_{M(\varphi)}(\underline{x}) = M(\varphi) \cdot \underline{x} = (\varphi(\underline{e}_1)|\cdots|\varphi(\underline{e}_n)) \cdot \underline{x} = \sum_{j=1}^n x_j\, \varphi(\underline{e}_j) = \varphi(\sum_{j=1}^n x_j\, \underline{e}_j) = \varphi(\underline{x})$. □

Beispiele 2.3 Wir geben nun noch einige spezielle lineare Abbildungen $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$ mittels der entsprechenden assoziierten Matrizen an.

- $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ mit $a > 0$ – Streckung um a
- $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ – Spiegelung an der "y-Achse"
- $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ – Punktspiegelung an $\underline{0}$

- $\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$ mit $\varphi \in \mathbb{R}$ –

  Drehung mit Bogenmaß φ (entspricht $\varphi \cdot \frac{180^\circ}{\pi}$).

Da lineare Abbildungen auch Gruppenhomomorhismen sind, können wir den Begriffs des *Kerns* auch auf lineare Abbildungen anwenden. Für eine lineare Abbildung $\varphi : K^n \longrightarrow K^m$ haben wir also

$$\text{Kern}(\varphi) = \{\underline{x} \in K^n \,|\, \varphi(\underline{x}) = \underline{0}\} \; .$$

Wir definieren nun für jede Matrix $A \in K^{m \times n}$

$$\text{Kern}(A) := \text{Kern}(\Lambda_A),$$

d.h.

$$\begin{aligned} \text{Kern}(A) = \; & \{\underline{x} \in K^n \,|\, A\underline{x} = \underline{0}\} \\ & \{\underline{x} \in K^n \,|\, \forall i = 1, \ldots, n : \textstyle\sum_{j=1}^{n} x_j a_{i,j} = 0\} \end{aligned}$$

Der Kern von A ist somit die Lösungsmenge des Gleichungssystems

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & 0 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & 0 \\ \vdots \qquad\qquad\quad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & 0 \end{array}$$

mit "Unbestimmten" $X_1, \ldots, X_n$. So ein Gleichungsystem heißt *homogenes lineares Gleichungssystem.*

Ferner definieren wir

$$\text{Bild}(A) := \text{Bild}(\Lambda_A)$$

d.h.

$$\text{Bild}(A) = \{\underline{b} \in K^m \mid \exists \underline{x} \in K^n : \underline{b} = A \cdot \underline{x}\} \; .$$

Das Bild von A ist somit gleich der Menge von Vektoren $\underline{b} \in K^m$, für die das *lineare Gleichungssystem*

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & b_1 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & b_2 \\ \vdots \qquad\qquad\quad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & b_m \end{array}$$

lösbar ist.

Wie man lineare Gleichungssysteme explizit löst, behandeln wir später.

Matrizenaddition und -multiplikation

Wir definieren eine Addition auf der Menge der $m \times n$-Matrizen $K^{m\times n}$ "komponentenweise":

$$+ : K^{m\times n} \times K^{m\times n} \longrightarrow K^{m\times n} ,$$
$$(((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots n}, ((b_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n}) \mapsto ((a_{i,j} + b_{i,j}))_{i=1,\ldots,m,j=1,\ldots n}$$

(Diese Definition ist ein Spezialfall der in 1.5, Unterabschnitt über Produkte, definierten Verknüpfung.)

Mit dieser Addition bildet die Menge der $m \times n$-Matrizen eine abelsche Gruppe.

Beachten Sie, dass man auch die linearen Abbildungen von K^n nach K^m addieren kann (siehe wiederum den erwähnten Unterabschnitt über Produkte oder die Diskussion auf Seite 45): Wenn $\varphi, \psi : K^n \longrightarrow K^m$ lineare Abbildungen sind, dann ist auch die Abbildung $\varphi + \psi : K^n \longrightarrow K^m$ (gegeben durch $\underline{x} \mapsto \varphi(\underline{x}) + \psi(\underline{x})$) linear. Auch die "Nullabbildung" ($\underline{x} \mapsto \underline{0}$) ist linear. Damit bilden die linearen Abbildungen von K^n nach K^m eine abelsche Gruppe.

Wir haben nun

$$M(\varphi + \psi) = M(\varphi) + M(\psi) \quad \text{und} \quad \Lambda_{A+B} = \Lambda_A + \Lambda_B \tag{2.1}$$

für alle linearen Abbildungen $\varphi, \psi : K^n \longrightarrow K^m$ und alle Matrizen $A, B \in K^{m\times n}$.

Mit anderen Worten: Die Zuordungen $\varphi \mapsto M(\varphi)$ und $A \mapsto \Lambda_A$ sind Isomorphismen zwischen der abelschen Gruppe der linearen Abbildungen von K^n nach K^m und der abelschen Gruppe $K^{m\times n}$.

Wir wollen nun eine *Multiplikation* $\cdot : K^{m\times n} \times K^{n\times r} \longrightarrow K^{m\times r}$ definieren, welche der Verknüpfung linearer Abbildungen entspricht. Mit anderen Worten: Wir wollen, dass für alle $A \in K^{m\times n}$ und alle $B \in K^{n\times r}$ gilt:

$$\Lambda_{A\cdot B} = \Lambda_A \circ \Lambda_B , \tag{2.2}$$

bzw.

$$A \cdot B = M(\Lambda_A \circ \Lambda_B) . \tag{2.3}$$

Wir *definieren* die Multiplikation also mittels (2.3). Wir wollen nun wissen, wie man explizit zwei Matrizen multipliziert. Seien dazu $A \in K^{m\times n}$ und $B \in K^{n\times r}$ mit $B = (\underline{b}_1 | \cdots | \underline{b}_r)$, und sei $C := A \cdot B$ mit $C = (\underline{c}_1 | \cdots | \underline{c}_r) = ((c_{i,j}))_{i=1,\ldots,m,j=1,\ldots,r}$. Dann ist

$$\underline{c}_j = C \cdot \underline{e}_j = \Lambda_C(\underline{e}_j) = (\Lambda_A \circ \Lambda_B)(\underline{e}_j) = \Lambda_A(\Lambda_B(\underline{e}_j)) = \Lambda_A(\underline{b}_j) = A \cdot \underline{b}_j .$$

Eine Umformulierung hiervon ist:

$$c_{i,j} = \sum_{\ell=1}^{n} a_{i,\ell} b_{\ell,j}$$

Mit anderen Worten:

$$A \cdot B = (A\underline{b}_1 | \cdots | A\underline{b}_r) \; . \tag{2.4}$$

bzw.

$$((a_{i,j}))_{i,j} \cdot ((b_{i,j}))_{i,j} = ((\sum_{\ell=1}^{n} a_{i,\ell}\, b_{\ell,j}))_{i,j} \; .$$

Insbesondere sieht man, dass die Multiplikation einer Matrix mit einem Spaltenvektor ein Spezialfall der Multiplikation zweier Matrizen ist (setze $r = 1$ in Formel (2.4)).

Da die Verknüpfung zweier (linearer) Abbildungen assoziativ ist, ist die Matrizenmultiplikation auch automatisch assoziativ (warum?). Es gilt also für alle $A \in K^{m\times n}, B \in K^{n\times r}, C \in K^{r\times s} : A(BC) = (AB)C$.

Die zur identischen Abbildung $\mathrm{id}_{K^n} : K^n \longrightarrow K^n$ assoziierte Matrix ist die so genannte *Einheitsmatrix*

$$I_n := (\underline{e}_1 | \cdots | \underline{e}_n) = ((\delta_{i,j})_{i,j=1,\ldots,n} \; .$$

Für alle Matrizen $A \in K^{m\times n}$ gilt $AI_n = A$ und für alle Matrizen $A \in K^{n\times r}$ gilt $I_n A = A$. Insbesondere ist $(K^{n\times n}, \cdot)$ ein Monoid mit neutralem Element I_n.

Wir wissen schon, dass für die Addition und Verknüpfung linearer Abbildungen die Distributivgesetze gelten:

$$\varphi \circ (\chi + \psi) = \varphi \circ \chi + \varphi \circ \psi$$

für alle linearen Abbildungen $\varphi : K^n \longrightarrow K^m$ und $\chi, \psi : K^r \longrightarrow K^n$ sowie

$$(\chi + \psi) \circ \varphi = \chi \circ \varphi + \psi \circ \varphi$$

für alle linearen Abbildungen $\chi, \psi : K^n \longrightarrow K^m$ und $\varphi : K^r \longrightarrow K^n$. (Weil lineare Abbildungen Gruppenhomomorphismen sind, siehe S. 45.) Die analoge Aussage gilt nun auch für Matrizen:

$$A(B + C) = AB + AC \qquad (B + C)A = BA + CA \; ,$$

wenn immer die Multiplikation definiert ist. (Dies kann man direkt nachrechnen oder (2.3) anwenden.)

Insgesamt erhalten wir:

Aussage 2.4 *Die Menge der linearen Abbildungen $K^n \longrightarrow K^n$ bildet mit der Verknüpfung als Multiplikation einen Ring. Genauso bildet die Menge der $n \times n$-Matrizen $K^{n\times n}$ mit der soeben definierten Addition und Multiplikation einen Ring. Die Zuordnungen $\varphi \mapsto M(\varphi)$ sowie $A \mapsto \Lambda_A$ sind zueinander inverse Isomorphismen zwischen diesen Ringen.*

Bemerkungen

- Die Menge der linearen Abbildungen von K^n nach K^n ist ein Unterring des Rings der Endomorphismen der abelschen Gruppe $(K^n, +)$.
- Für $n \geq 2$ sind die obigen Ringe nicht kommutativ.

Wenn $\varphi : K^n \longrightarrow K^m$ eine lineare Abbildung ist und $c \in K$, dann ist auch die Abbildung $c\varphi : K^n \longrightarrow K^m$, die per Definition durch $\underline{x} \mapsto c\varphi(\underline{x})$ gegeben ist, linear. Es gilt

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

für alle $c \in K$ und alle linearen Abbildungen $\varphi : K^r \longrightarrow K^n$ und $\psi : K^n \longrightarrow K^m$.

In Analogie hierzu definieren wir nun noch eine *Skalarmultiplikation*:

$$K \times K^{m\times n} \longrightarrow K^{m\times n} \, , (c, A) \mapsto ((c \cdot a_{i,j}))_{i,j} \, .$$

Damit gilt

$$M(c\varphi) = cM(\varphi) \;\text{ und } \Lambda_{cA} = c\Lambda_A \, .$$

Außerdem gilt

$$(cA)B = c(AB) = A(cB)$$

für $c \in K, A \in K^{m\times n}$ und $B \in K^{n\times r}$.

2.2 Vektorräume und lineare Abbildungen

Vektorräume

Der Raum K^n mit den oben definierten Operationen ist ein Beispiel für einen *Vektorraum*:

Definition Ein *K-Vektorraum* (oder ein *Vektorraum über K*) ist eine abelsche Gruppe $(V, +)$ zusammen mit einer Abbildung (genannt *Skalarmultiplikation* (von K auf V))

$$\cdot : \; K \times V \longrightarrow V, (a, \mathfrak{v}) \mapsto a \cdot \mathfrak{v} \; ,$$

so dass

- $\forall a \in K \, \forall \mathfrak{v}, \mathfrak{w} \in V \; : a(\mathfrak{v} + \mathfrak{w}) = a\mathfrak{v} + a\mathfrak{w}$
- $\forall a, b \in K \, \forall \mathfrak{v} \in V \; : a(b\mathfrak{v}) = (ab)\mathfrak{v}$
- $\forall a, b \in K \, \forall \mathfrak{v} \in V \; : (a + b)\mathfrak{v} = a\mathfrak{v} + b\mathfrak{v}$
- $\forall \mathfrak{v} \in V \; : 1_K \cdot \mathfrak{v} = \mathfrak{v}$

Hier haben wir (wie üblich) die Multiplikationspunkte weggelassen und die Regel ("Mal vor Plus") angewandt.

Notation Das Nullelement bezeichnen wir mit $\mathfrak{o}$.

Sprachgebrauch Die Elemente eines Vektorraums heißen *Vektoren*, und die Elemente des Grundkörpers, wie schon gesagt, *Skalare*. Beachten Sie, dass der Begriff eines Vektorraums abstrakt ist; insbesondere sind die Elemente eines Vektorraums nicht notwendigerweise in irgendeinem anschaulichen Sinn Vektoren. Wir werden Vektoren aus "allgemeinen" Vektorräumen immer mit Frakturbuchstaben bezeichnen.

Bemerkung / Frage Es gilt immer $-\mathfrak{v} = (-1) \cdot \mathfrak{v}$. Warum?

Einige Beispiele:

Beispiele 2.5

- Für alle $n \in \mathbb{N}_0$ ist K^n mit der üblichen Addition und Skalarmultiplikation ein K-Vektorraum.
- Der Körper K ist mit seiner Addition und Multiplikation ein K-Vektorraum. ("Im Wesentlichen" ist $K = K^1$, siehe unten.)
- Der Raum $K^{m \times n}$ der $m \times n$-Matrizen ist mit der im letzten Kapitel definierten Addition und Skalarmultiplikation ein K-Vektorraum.
- Der Raum der Polynome $K[X]$ ist ein K-Vektorraum (mit der offensichtlichen Skalarmultiplikation).

- Sei M eine beliebige Menge. Dann ist $K^M (= \mathrm{Abb}(M, K))$ mit der komponentenweisen Addition (d.h. $(x_m)_{m\in M} + (y_m)_{m\in M} = (x_m + y_m)_{m\in M}$) sowie der Skalarmultiplikation

$$K \times K^M \longrightarrow K^M \ , \ (a, (x_m)_{m\in M}) \mapsto (a \cdot x_m)_{m\in M}$$

  ein K-Vektorraum.

Untervektorräume

Sei V ein K-Vektorraum und $U \subseteq V$ eine Teilmenge, die die $\mathfrak{o}$ enthält und abgeschlossen bezüglich der Addition und der Skalarmultiplikation ist. Dann ist U mit den induzierten Verknüpfungen auch ein K-Vektorraum; man spricht von einem K-*Untervektorraum* oder einem *Untervektorraum*. Untervektorräume nennt man auch *lineare Unterräume*, insbesondere in K^n.

Notation Wenn U ein Untervektorraum von V ist, schreibt man auch $U \leq V$.

Einige Beispiele hierzu:

Beispiele 2.6

- Für jeden Vektorraum V sind sowohl $\{\mathfrak{o}\}$ als auch V selbst Untervektorräume.
- In $\mathbb{R}^2$ sind alle Geraden durch den Ursprung Untervektorräume.
- In $\mathbb{R}^3$ sind alle Geraden durch den Ursprung und alle Ebenen durch den Ursprung Untervektorräume.
- Sei für eine beliebige Menge M

  $K^{(M)} := \{(a_m)_{m\in M} \in K^M \mid \text{Es gibt nur endlich viele } m \in M \text{ mit } a_m \neq 0\}$.[2]

  Dann ist $K^{(M)}$ ein K-Untervektorraum von K^M.

  Beachten Sie: Nach unserer Definition ist der K-Vektorraum $K[X]$ gleich $K^{(\mathbb{N}_0)}$.

- Die Menge der konvergenten Folgen bildet einen Untervektorraum im $\mathbb{R}$-Vektorraum $\mathbb{R}^{\mathbb{N}}$.

[2] Wenn Sie die Vorlesung Analysis I besuchen, kennen Sie sicher den Begriff "für fast alle". Dies heißt: bis auf endlich viele. Hiermit kann man die Bedingung auch so formulieren: Für fast alle $m \in M$ gilt $a_m = 0$.

- Die Menge der gegen Null konvergenten Folgen bildet einen Untervektorraum im $\mathbb{R}$-Vektorraum der konvergenten Folgen.
- Sei $I \subseteq \mathbb{R}$ eine beliebige Teilmenge. Dann bilden die stetigen Funktionen $I \longrightarrow \mathbb{R}$ einen Untervektorraum im $\mathbb{R}$-Vektorraum aller Funktionen $I \longrightarrow \mathbb{R}$. (Der letztere Raum ist $\mathbb{R}^I$.)

Lineare Abbildungen

Der Begriff der linearen Abbildung verallgemeinert sich sofort von Abbildungen zwischen Vektorräumen:

Definition Seien V und W K-Vektorräume. Eine *K-lineare Abbildung* (kurz: eine *lineare Abbildung*) von V nach W ist eine Abbildung $\varphi : V \longrightarrow W$, so dass

- $\forall \mathfrak{v}, \mathfrak{w} \in V : \varphi(\mathfrak{v} + \mathfrak{w}) = \varphi(\mathfrak{v}) + \varphi(\mathfrak{w})$
- $\forall a \in K \; \forall \mathfrak{v} \in V : \varphi(a\mathfrak{v}) = a\varphi(\mathfrak{v})$.

Die K-linearen Abbildungen sind genau die "strukturerhaltenden" Abbildungen zwischen K-Vektorräumen, genau wie z.B. die Gruppenhomomorphismen die strukturerhaltenden Abbildungen zwischen Gruppen sind.

Bemerkung Seien V, W K-Vektorräume. Dann ist eine Abbildung $\varphi : V \longrightarrow W$ genau dann linear, wenn für alle $a \in K$ und alle $\mathfrak{v}, \mathfrak{w} \in V$ $\varphi(a\mathfrak{v} + \mathfrak{w}) = a\varphi(\mathfrak{v}) + \varphi(\mathfrak{w})$ gilt.

Bemerkung Sei $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann haben wir, wie bei jedem Gruppenhomomorphismus, Kern(φ). Dies ist ein Untervektorraum von V. Außerdem haben wir Bild(φ), und dies ist ein Untervektorraum von W.

Lemma 2.7

a) Wenn $\varphi : V \longrightarrow W$ und $\psi : W \longrightarrow X$ lineare Abbildungen von K-Vektoräumen sind, dann auch $\psi \circ \varphi : V \longrightarrow X$.

b) Wenn $\varphi, \psi : V \longrightarrow W$ lineare Abbildungen von K-Vektorräumen sind, dann auch $\varphi + \psi : V \longrightarrow W$, $\mathfrak{v} \mapsto \varphi(\mathfrak{v}) + \psi(\mathfrak{v})$.

c) Wenn $a \in K$ und $\varphi : V \longrightarrow W$ eine lineare Abbildungen von K-Vektorräumen ist, dann auch $a\varphi : V \longrightarrow W$, $\mathfrak{v} \mapsto a\varphi(\mathfrak{v})$.

d) Wenn $\varphi : V \longrightarrow W$ eine bijektive lineare Abbildung von K-Vektorräumen ist, dann ist auch $\varphi^{-1} : W \longrightarrow V$ linear.

Definition Die K-linearen Abbildungen nennt man auch *Homomorphismen* von K-Vektorräumen. (Entsprechend der Philosophie von Abschnitt 1.8.

Dementsprechend bezeichnet man die Menge der K-linearen Abbildungen von V nach W mit $\mathrm{Hom}_K(V, W)$ (oder einfach mit $\mathrm{Hom}(V, W)$, wenn es klar ist, dass man "über dem Körper K" arbeitet).

Aussage 2.8 *Seien V, W zwei K-Vektorräume. Dann ist $\mathrm{Hom}_K(V, W)$ mit der in Lemma 2.7 b) definierten Addition und der in Lemma 2.7 c) definierten Skalarmultiplikation ein K-Vektorraum.*

Der *Beweis* dieser Aussage ist einfach.

Wir wenden nun auch die anderen Begriffe aus Abschnitt 1.8 auf lineare Abbildungen an:

- Ein *Isomorphismus* von V nach W (beides K-Vektorräume) ist eine lineare Abbildung $\varphi : V \longrightarrow W$, so dass es eine lineare Abbildung $\psi : W \longrightarrow V$ mit $\psi \circ \varphi = \mathrm{id}_V$ und $\varphi \circ \psi = \mathrm{id}_W$ existiert. (Nach Lemma 2.7 d) ist dies äquivalent dazu, dass φ eine bijektive lineare Abbildung ist.) Wenn es einen Isomorphismus zwischen V und W gibt, heißen V und W *isomorph*.

- Ein *Endomorphismus* eines K-Vektorraums V ist eine lineare Abbildung von V nach V.

- Ein *Automorphismus* eines K-Vektorraums V ist ein Isomorphismus von V nach V.

Die Bezeichnungen für die entsprechenden Mengen sind $\mathrm{Iso}_K(V, W)$, $\mathrm{End}_K(V)$ und $\mathrm{Aut}_K(V)$, wobei das K auch weggelassen werden kann. (Beachten Sie, dass die letzeren beiden Bezeichungen analog zu den Bezeichnungen in Abschnitt 1.8 sind.)

Es ist klar, dass $\mathrm{End}_K(V)$ mit der schon diskutierten Addition und der Verknüpfung als Multiplikation ein Ring ist, und $\mathrm{Aut}_K(V)$ ist mit der Verknüpfung eine Gruppe. Dabei ist $\mathrm{Aut}_K(V)$ gleich der Gruppe der (bzgl. der Multiplikation) invertierbaren Elemente von $\mathrm{End}_K(V)$, also $\mathrm{End}_K(V)^* = \mathrm{Aut}_K(V)$ (vergleiche die letzte Aussage mit Aussage 1.65).

Und nun noch einige Beispiele für Isomorphismen von K-Vektorräumen:

- Die Abbildung $K \longrightarrow K^1, x \mapsto (x)$ (wobei (x) das "1-Tupel" ist, das x enthält) ist ein Isomorphismus von K-Vektorräumen. (Wir identifizieren diese beiden Vektorräume.)

- Die Abbildung

  $$K^{m\times n} \longrightarrow K^{m\cdot n}, A = ((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n} \mapsto \underline{x} \text{ mit } x_{(j-1)\cdot m+i} := a_{i,j}$$

  für alle $i = 1, \ldots, m, j = 1, \ldots, n$ ist ein Isomorphismus von K-Vektorräumen.

- Die Abbildungen $\mathrm{Hom}_K(K^n, K^m) \longrightarrow K^{m\times n}, \varphi \mapsto M(\varphi)$ und $K^{n\times m} \longrightarrow \mathrm{Hom}_K(K^n, K^m), A \mapsto \Lambda_A$ sind zueinander inverse Isomorphismen von K-Vektorräumen.

2.3 Die komplexen Zahlen

In den reellen Zahlen ist die Gleichung $X^2 = -1$ nicht lösbar. Wir wollen einen "möglichst kleinen" Körpern finden, der die reellen Zahlen enthält, in dem die Gleichung lösbar ist.

Nehmen wir mal an, dass wir einen Körper M haben, der die reellen Zahlen enthält, in dem die Gleichung lösbar ist. Sei $\boldsymbol{i}$ ein Element aus M mit $\boldsymbol{i}^2 = -1$.

In M betrachten wir nun die Teilmenge

$$C := \{a + b\boldsymbol{i} \,|a, b \in \mathbb{R}\} .$$

Dann enthält C offensichtlich die Null (denn $0 = 0 + 0 \cdot \boldsymbol{i}$) und die 1 (denn $1 = 1 + 0 \cdot \boldsymbol{i}$). Außerdem gilt für alle $a, b, c, d \in \mathbb{R}$ (in M):

$$(a + b\boldsymbol{i}) + (c + d\boldsymbol{i}) = (a + c) + (b + d) \cdot \boldsymbol{i} \tag{2.5}$$

und

$$(a + b\boldsymbol{i}) \cdot (c + d\boldsymbol{i}) = (ac - bd) + (ad + bc) \cdot \boldsymbol{i} . \tag{2.6}$$

Wir sehen, dass C abgeschlossen bezüglich der Addition und der Multiplikation ist. Außerdem gilt

$$-(a + b\boldsymbol{i}) = -a + (-b) \cdot \boldsymbol{i} \in C$$

und falls $a + b\boldsymbol{i} \neq 0$:

$$(a + b\boldsymbol{i})^{-1} = \frac{a - b\boldsymbol{i}}{(a + b\boldsymbol{i})(a - b\boldsymbol{i})} = \frac{a}{a^2 + b^2} + \frac{-b}{a^2 + b^2}\boldsymbol{i} \in C$$

Somit ist C ein Unterkörper von M.

Wir haben nun eine Vorstellung davon, welche Eigenschaften der Körper der komplexen Zahlen $\mathbb{C}$ haben sollte: $\mathbb{R}$ ist ein Unterkörper von $\mathbb{C}$, es gibt ein Element $\boldsymbol{i} \in \mathbb{C}$ mit $\boldsymbol{i}^2 = -1$ und dabei gilt:

- Für alle $z \in \mathbb{C}$ existieren eindeutig bestimmte $a, b \in \mathbb{R}$: $z = a + b\boldsymbol{i}$
- Die Addition ist wie in (2.5) und die Multiplikation wie in (2.6) gegeben.

Nun *definieren* wir den Körper der komplexen Zahlen wie folgt: Die unterliegende Menge ist $\mathbb{R}^2$ (d.h. die Elemente sind Tupel (a, b) von reellen Zahlen). Wir haben wie üblich eine komponentenweise Addition und eine Skalarmultiplikation.

Wir definieren $\mathbf{1} := (1, 0), \boldsymbol{i} := (0, 1)$ und $\mathbf{0} := (0, 0)$. (Das sind fürs Erste nur Namen.) Dann gibt es also für alle $z \in \mathbb{C}$ eindeutig bestimmte $a, b \in \mathbb{R}$ mit $z = a \cdot \mathbf{1} + b \cdot \boldsymbol{i}$.

Wir definieren nun zusätzlich eine Multiplikation wie folgt: Für $a, b, c, d \in \mathbb{R}$ sei

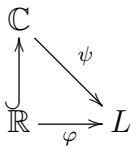
$$(a \cdot \mathbf{1} + b \cdot \boldsymbol{i}) \cdot (c \cdot \mathbf{1} + d \cdot \boldsymbol{i}) := (ac - bd) \cdot \mathbf{1} + (ad + bc) \cdot \boldsymbol{i}$$

Man kann nun überprüfen, dass $\mathbb{C}$ mit diesen Operationen und $\mathbf{0}$ als Nullelement und $\mathbf{1}$ als Einselement einen Körper bildet. Außerdem gilt $\boldsymbol{i}^2 = -\mathbf{1}$. Dies ist der *Körper der komplexen Zahlen*.

Wir haben den Homomorphismus von Körpern $\mathbb{R} \longrightarrow \mathbb{C}$, $a \mapsto a \cdot \mathbf{1}$. Wir "identifizieren" $\mathbb{R}$ mit seinem Bild in $\mathbb{C}$. Nun gibt es in der Tat für alle $z \in \mathbb{C}$ eindeutig bestimmte $a, b \in \mathbb{C}$: $z = a + b\boldsymbol{i}$.

Wir wollten ja, dass $\mathbb{C}$ in einer gewissen Hinsicht der "kleinste" Körper mit diesen Eigenschaften ist. Dies kann man wie folgt präzisieren:

Aussage 2.9 *Sei $\varphi : \mathbb{R} \longrightarrow L$ ein Homomorphismus von Körpern, und sei $\alpha \in L$ ein Element mit $\alpha^2 = -1_L$. Dann gibt es einen eindeutig bestimmten Homomorphismus von Körpern $\psi : \mathbb{C} \longrightarrow L$, so dass $\psi(\boldsymbol{i}) = \alpha$ gilt und das folgende Diagramm kommutativ ist:*



Beachten Sie hier, dass Homomorphismen von Körpern immer injektiv sind (siehe Aussage 1.63). Somit ist ψ ein Isomorphismus von $\mathbb{C}$ auf sein Bild in L; L enthält also einen zu $\mathbb{C}$ isomorphen Körper.

Bemerkung Sei L ein beliebiger Körper und $K \subseteq L$ ein Unterkörper. Dann ist L in "natürlicher Weise" ein K-Vektorraum, und zwar wie folgt: $(L, +)$ ist eine abelsche Gruppe, und wenn wir die Multiplikation $\cdot : L \times L \longrightarrow L$ auf $K \times L$ einschränken, erhalten wir eine Skalarmultiplikation von K auf L.

Dies kann man natürlich auch für $\mathbb{R} \subseteq \mathbb{C}$ betrachten. Die so erhaltene Skalarmultiplikation ist genau die Skalarmultiplikation von $\mathbb{R}$ auf $\mathbb{R}^2$.

Es gilt nun:

Fundamentalsatz der Algebra Jedes nicht-konstante Polynom $p(X) \in \mathbb{C}[X]$ hat eine Nullstelle (in $\mathbb{C}$).

Ein Beweis dieses Satzes würde uns zu weit wegführen. Wir geben noch das folgende einfache Korollar des Fundamentalsatzes an:

Korollar 2.10 *Jedes nicht-konstante Polynom $p(X) \in \mathbb{C}[X]$ zerfällt in Linearfaktoren, d.h. es gibt $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ (mit $d = \mathrm{Grad}(p(X))$ und $c \in \mathbb{C}$ mit*

$$p(X) = c \cdot (X - \alpha_1) \cdots (X - \alpha_d) \; .$$

In diesem Kontext ist die folgende Definition sinnvoll:

Definition Sei K ein Körper. Wenn nun jedes nicht-konstante Polynom in $K[X]$ eine Nullstelle (in K) besitzt, heißt K *algebraisch abgeschlossen.*

Offensichtlich zerfallen alle nicht-konstanten Polynome über algebraisch abgeschlossenen Körpern in Linearfaktoren. Man kann zeigen, dass man jeden Körper in einen algebraisch abgeschlossenen Körper einbetten kann. Das einzige Beispiel eines algebraisch abgeschlossenen Körpers, das für diese Vorlesung relevant ist, sind die komplexen Zahlen.

2.4 Endliche Systeme von Vektoren

Sei weiterhin K ein beliebiger Körper, und sei V ein K-Vektorraum.

Für das Folgende müssen wir den vielleicht offensichtlich erscheinenden Begriff eines "Systems von Vektoren" klären:

Definition Eine Familie von Vektoren $(\mathfrak{v}_i)_{i \in I}$ über einer beliebigen Indexmenge I nennen wir auch ein *System von Vektoren.*

Ich erinnere daran, dass ein Tupel $(\mathfrak{v}_1, \ldots, \mathfrak{v}_r)$ von Vektoren mit $r \in \mathbb{N}$ nichts anderes als eine Familie von Vektoren über der Indexmenge $\{1, \ldots, r\}$

ist. In diesem Sinne definieren wir das *leere Tupel* von Vektoren oder das *leere System* als die Familie über der leeren Menge. (Dies ist die eindeutig bestimmte Abbildung $\emptyset \longrightarrow V$.

Anstatt ein Tupel $(\mathfrak{v}_1, \ldots, \mathfrak{v}_r)$ anzugeben, kann man natürlich auch einfach die Elemente aufzählen: $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$. Wir sagen dann auch, dass $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ ein System von Vektoren ist oder wir sagen einfach, dass $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ Vektoren von V sind. Wenn nichts weiteres gesagt wird, ist $r = 0$ (das leere System) immer zugelassen.

Wie schon in Fußnote 14 in Kapitel 1 erwähnt, definieren wir die "leere Summe" (die Summe über das leere System) als $\mathfrak{o}$.

Es seien nun Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r \in V$ gegeben (mit $r \in \mathbb{N}_0$). Eine *Linearkombination* von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ ist ein Vektor $\mathfrak{v} \in V$, so dass es Skalare $a_1, \ldots, a_r \in K$ mit $\mathfrak{v} = a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r$ gibt.

Nun ist

$$\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_n \rangle_K := \{a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_n \mid a_1, \ldots, a_r \in K\}$$

offensichtlich ein Untervektorraum von V. Außerdem umfasst dieser Untervektorraum in jedem anderen Untervektorraum von V, der $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ enthält. Folglich ist $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K$ der kleinste Untervektorraum von V, der $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ enthält.

Definition Der Raum $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K$ heißt das *(lineare) Erzeugnis* oder die *lineare Hülle* von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ in V oder auch der *von* $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *aufgespannte / erzeugte Raum.*

Analog zu abelschen Gruppen kann man $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ auch durch eine (beliebige) Teilmenge von V ersetzen: Sei $S \subseteq V$. Dann ist wiederum

$$\langle S \rangle_K :=$$
$$\{\mathfrak{v} \in V \mid \exists k \in \mathbb{N}_0, \exists \mathfrak{v}_1, \ldots, \mathfrak{v}_k \in V, \exists a_1, \ldots, a_k \in K : \mathfrak{v} = a_1\mathfrak{v}_1 + \cdots + a_k\mathfrak{v}_k\}$$

der kleinste Untervektorraum von V, der die Menge S umfasst. Dieser Raum heißt das *Erzeugnis* von S in V.

Notation Wenn es vom Kontext her klar ist, dass man vom Erzeugnis in V als K-Vektorraum redet, schreibt man auch $\langle S \rangle$ statt $\langle S \rangle_K$.

Beispiele 2.11

- Das Erzeugnis von $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ in $\mathbb{R}^2$ ist $\{\begin{pmatrix} a \\ a \end{pmatrix} \mid a \in \mathbb{R}\}$ – die "Diagonale".

- Das Erzeugnis der leeren Menge im K-Vektorraum V ist $\{\mathfrak{o}\}$.
- Sei U ein Untervektorraum von V. Dann ist das Erzeugnis von U in V gleich U.
- Sei K ein beliebiger Körper und $a \in K, a \neq 0$. Wir betrachten K als K-Vektorraum (mit den offensichtlichen Operationen). Dann ist das Erzeugnis von a gleich K.
- Beispielsweise ist das Erzeugnis der 1 in $\mathbb{Q}$ als $\mathbb{Q}$-Vektorraum gleich $\mathbb{Q}$. Beachten Sie: Das Erzeugnis der 1 in der abelschen Gruppe $(\mathbb{Q}, +)$ ist hingegen $\mathbb{Z}$.

Definition Sei V ein K-Vektorraum.

- Seien $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ Vektoren ($r \in \mathbb{N}_0$). Wenn nun $V = \langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K$, sagen wir, dass V von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *erzeugt* wird. Das System $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ heißt dann ein *Erzeugendensystem* von V.
- Sei nun S eine Menge von Vektoren aus V. Wenn nun $V = \langle S \rangle_K$, sagen wir wiederum, dass V von S *erzeugt* wird; S heißt dann eine *erzeugende Menge* von V.
- Wenn V von einer endlichen Menge erzeugt wird (d.h. wenn es Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ gibt, die V erzeugen), nennt man V auch *endlich erzeugt*.

Im Folgenden diskutieren wir einige grundlegende Begriffe für endliche Systeme von Vektoren (genauer: für Tupel von Vektoren). Die Begriffe und ihre Beziehungen können auf beliebige Systeme erweitert werden. Dies kommt später.

Seien also Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ mit $r \in \mathbb{N}_0$ gegeben.

Definition Die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ bilden eine *Basis* von V, wenn gilt: Für alle $\mathfrak{v} \in V$ gibt es eindeutig bestimmte $a_1, \ldots, a_r \in K$ mit

$$\mathfrak{v} = a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r \ .$$

Bemerkung Beachten Sie den Unterschied zwischen einem Erzeugendensystem und einer Basis! In beiden Fällen kann man jeden Vektor als Linearkombination der Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ schreiben. Bei einer Basis sind die "Koeffizienten" $a_1, \ldots, a_r$ zusätzlich eindeutig bestimmt.

Beispiel 2.12

- Der "triviale Vektorraum" $\{\mathfrak{o}\}$ hat das leere System als Basis.
- Der Raum K^n hat $\underline{e}_1, \ldots, \underline{e}_n$ als Basis.

(Mit $n = 0$ ist der erste Punkt ein Spezialfall des zweiten.)

Definition Sei $\mathfrak{v} \in V$. Dann heißt $\mathfrak{v}$ *linear abhängig* von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$, wenn es $a_1, \ldots, a_r \in K$ mit

$$\mathfrak{v} = a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r$$

gibt. Wenn dies nicht der Fall ist, heißt $\mathfrak{v}$ *linear unabhängig* von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$.

Bemerkung Der Nullvektor ist von jedem System linear abhängig, auch vom leeren System.

Lemma 2.13 *Sei $\mathfrak{v} \in V$. Dann sind die folgenden Aussagen äquivalent:*

- *$\mathfrak{v}$ ist linear unabhängig von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$.*
- *$\mathfrak{v} \notin \langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K$.*
- *$\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K \subsetneq \langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{v} \rangle_K$.*

Beweis. Die ersten beiden Aussagen sind offensichtlich äquivalent, und die zweite impliziert die dritte.

Es ist zu zeigen, dass die dritte Aussage die zweite impliziert bzw. dass das Gegenteil der zweiten Aussage das Gegenteil der dritten Aussage impliziert.

Es gelte also $\mathfrak{v} \in \langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K$. Dann ist $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K$ ein linearer Unterraum, der $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{v}$ enthält, und es ist auch der kleinste solche lineare Unterraum. Denn: Sei V ein weiterer Unterraum mit dieser Eigenschaft. Dann enthält V auch $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$, und somit enthält V auch $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K$.

Der kleinste lineare Unterraum, der $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{v}$ enthält, ist jedoch $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{v} \rangle_K$. Damit ist $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r \rangle_K = \langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{v} \rangle_K$. □

Definition Die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ heißen *linear unabhängig*, wenn keiner der Vektoren von den anderen linear abhängig ist, mit anderen Worten, falls für alle $i = 1, \ldots, r$ gilt: $\mathfrak{v}_i$ ist linear unabhängig von $\mathfrak{v}_1, \ldots, \mathfrak{v}_{i-1}, \mathfrak{v}_{i+1}, \ldots, \mathfrak{v}_r$. Andernfalls heißen sie *linear abhängig*.

Lemma 2.14 *Die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ sind genau dann linear unabhängig, wenn gilt:*

$$\forall a_1, \ldots, a_r \in K : a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r = \mathfrak{o} \longrightarrow a_1 = \cdots = a_r = 0 \, .$$

Beweis. Wir zeigen, dass die Vektoren genau dann linear abhängig sind, wenn das Kriterium im Lemma nicht gilt.

Wenn die Vektoren linear abhängig sind, ist das Kriterium im Lemma offensichtlich falsch.

Sei nun das Kriterium im Lemma falsch. Dann gibt es $a_1, \ldots, a_r$, nicht alle $= 0$, so dass $a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r = 0$. Sei $a_i \neq 0$. Dann ist also

$$\mathfrak{v}_i = -\frac{a_1}{a_i}\mathfrak{v}_1 - \cdots - \frac{a_{i-1}}{a_i}\mathfrak{v}_{i-1} - \frac{a_{i+1}}{a_i}\mathfrak{v}_{i+1} - \cdots - \frac{a_r}{a_i}\mathfrak{v}_r \ ,$$

d.h. $\mathfrak{v}_i$ ist linear abhängig von den anderen Vektoren. □

Bemerkung Man sagt: "Das System $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ ist genau dann linear unabhängig, wenn sich der Nullvektor nur auf triviale Weise als Linearkombination der $\mathfrak{v}_i$ darstellen lässt."

Bemerkung Man sollte immer versuchen, das Kriterium im obigen Lemma anzuwenden, wenn man zeigen will, dass ein System von Vektoren linear unabhängig ist.

Definition

- Die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ bilden ein *maximales linear unabhängiges System*, wenn sie linear unabhängig sind und für alle Vektoren $\mathfrak{v}$ gilt: Das System $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{v}$ ist linear abhängig.
- Die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ bilden ein *minimales Erzeugendensystem* von V wenn sie ein Erzeugendensystem bilden und für $i = 1, \ldots, r$ gilt: $\mathfrak{v}_1, \ldots, \mathfrak{v}_{i-1}, \mathfrak{v}_{i+1}, \ldots, \mathfrak{v}_r$ ist kein Erzeugendensystem von V.

Übungsaufgabe Es gibt einen Zusammenhang zwischen den obigen Definitionen und den Begriffen "maximales Element" / "minimales Element" bei Ordnungsrelationen. Welche Menge und welche Ordnungsrelation sollte man betrachten, damit sich die obige Definition aus den allgemeinen Definitionen von Ordnungsrelationen ergeben?

Aussage 2.15 *Die folgenden Aussagen sind äquivalent:*

a) $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *ist eine Basis von* V*.*

b) $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *ist ein linear unabhängiges Erzeugendensystem von* V*.*

c) $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *ist ein maximales linear unabhängiges System von* V*.*

d) $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *ist ein minimales Erzeugendensystem von* V*.*

Beweis. Wir zeigen, dass alle Aussagen äquivalent zur Aussage b) sind. Zunächst zeigen wir, dass jede der Aussagen a), c), d) (für sich) die Aussage b) impliziert.

Es gelte a). Dann haben wir insbesondere ein Erzeugendensystem. Es ist zu zeigen, dass das System auch linear unabhängig ist. Seien dazu $a_1, \cdots, a_r \in K$ mit $a_1\mathfrak{v}_1 + \ldots + a_r\mathfrak{v}_r = \mathfrak{o}$. Wir haben auch $0 \cdot \mathfrak{v}_1 + \cdots + 0 \cdot \mathfrak{v}_r = \mathfrak{o}$. Hieraus folgt $a_1 = 0, \ldots, a_r = 0$ aufgrund der Eindeutigkeit der "Darstellung von $\mathfrak{o}$". Damit gilt b).

Es gelte nun c). Dann haben wir also ein linear unabhängiges System. Es ist zu zeigen, dass wir auch ein Erzeugendensystem haben. Sei dazu $\mathfrak{v} \in V$. Dann ist nach Voraussetzung das System $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{v}$ linear abhängig. Es gibt also $a_1, \ldots, a_r, a \in K$, nicht alle $= 0$, mit $a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r + a\mathfrak{v} = \mathfrak{o}$. Wenn nun $a = 0$ wäre, dann wären die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ linear abhängig, was aber nach Voraussetzung nicht der Fall ist. Also ist $a \neq 0$. Damit gilt $\mathfrak{v} = -\frac{a_1}{a}\mathfrak{v}_1 - \cdots - \frac{a_r}{a}\mathfrak{v}_r$. Damit gilt wiederum b).

Es gelte nun d). Dann haben wir also ein Erzeugendensystem. Es ist zu zeigen, dass wir auch ein linear unabhängiges System haben. Sei dazu $i = 1, \ldots, r$. Nach Voraussetzung ist $\mathfrak{v}_1, \ldots, \mathfrak{v}_{i-1}, \mathfrak{v}_{i+1}, \mathfrak{v}_r$ kein Erzeugendensystem und somit $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_{i-1}, \mathfrak{v}_{i+1}, \mathfrak{v}_r \rangle_K \subsetneq V$. Damit ist nach Lemma 2.13 $\mathfrak{v}_i$ linear unabhängig von den anderen Vektoren.

Es gelte nun b).

Wir zeigen zuerst a). Offensichtlich können wir jeden Vektor in der gewünschten Weise darstellen. Wir müssen die Eindeutigkeit der Darstellung zeigen. Sei also $\mathfrak{v} \in V$ und seien $a_1, \ldots, a_r \in K$ und $a'_1, \ldots, a'_r \in K$ mit $\mathfrak{v} = a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r = a'_1\mathfrak{v}_1 + \cdots + a'_r\mathfrak{v}_r$. Dann ist

$$\mathfrak{o} = (a_1 - a'_1)\mathfrak{v}_1 + \cdots + (a_r - a'_r)\mathfrak{v}_r \ .$$

Nach Voraussetzung ist nun $a_1 - a'_1 = \cdots = a_r - a'_r = 0$, d.h. $a_1 = a'_1, \ldots, a_r = a'_r$.

Nun zu c). Da $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ ein Erzeugendensystem von V ist, ist jeder Vektor in V linear abhängig von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$. Damit gilt c).

Zu d). Sei $i = 1, \ldots, r$. Nach Voraussetzung ist $\mathfrak{v}_i$ linear unabhängig von $\mathfrak{v}_1, \ldots, \mathfrak{v}_{i-1}, \mathfrak{v}_{x+1}, \mathfrak{v}_r$, also gilt $\mathfrak{v}_i \notin \langle \mathfrak{v}_1, \ldots, \mathfrak{v}_{i-1}, \mathfrak{v}_{i+1}, \ldots, \mathfrak{v}_r \rangle_K$. Damit ist insbesondere $\mathfrak{v}_1, \ldots, \mathfrak{v}_{i-1}, \mathfrak{v}_{x+1}, \ldots, \mathfrak{v}_r$ kein Erzeugendensystem von V. □

Hieraus folgt sofort:

Aussage 2.16 *Ein endlich erzeugter Vektorraum hat stets eine Basis.*

Beweisskizze. Wir starten mit einem Erzeugendensystem. Falls dieses System linear unabhängig ist, sind wir fertig. Andernfalls gibt es einen Vektor, der

von den anderen Vektoren linear abhängig ist. Wir nehmen so einen Vektor weg. Wir erhalten ein neues Erzeugendensystem. Wir wiederholen diese Prozedur, solange dies geht. Irgendwann erhalten wir so eine Basis.

(Können Sie dieses Argument zu einem formalen Beweis ausbauen?) □

Bemerkung Genauer kann man das Folgende zeigen: Sei $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ ein Erzeugendensystem. Dann gibt es ein $k \in \mathbb{N}_0$ und $i_1, \ldots, i_k \in \{1, \ldots, r\}$ mit $i_a < i_b$ für $a < b$, so dass $v_{i_1}, \ldots, v_{i_k}$ eine Basis ist. ("Aus einem Erzeugendensystem kann man eine Basis auswählen.")

Satz 2.1 (Basisergänzungssatz) *Sei $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ ein linear unabhängiges System und $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Basis. Dann gibt es Indices $i_1, \ldots, i_{n-r}$ mit $i_a < i_b$ für $a < b$, so dass $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{i_1}, \ldots, \mathfrak{b}_{i_{n-r}}$ eine Basis ist.*

Den Satz kann man salopp so ausdrücken: "Man kann $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ mit $n - r$ Vektoren aus der Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ zu einer Basis ergänzen."

Beweis. Die Idee des folgenden Beweises ist algorithmisch: Wir tauschen successive Vektoren aus der Basis gegen Vektoren aus dem linear unabhängigen System aus. Dies machen wir so, dass wir stets eine Basis haben. Ein etwas subtiler Punkt ist hierbei: Wie stellt man sicher, dass man einen Vektor, den man zuvor "eingetauscht" hat, nicht wieder austauscht?

Wir zeigen die Aussage für einen festen Vektorraum V per Induktion nach r. Der Induktionsanfang für $r = 0$ ist trivial. Zum Induktionsschritt von r nach $r + 1$:

Sei $\mathfrak{v}_1, \ldots, \mathfrak{v}_{r+1}$ ein linear unabhängiges System. Dann gibt es nach Induktionsvoraussetzung Indices $i_1, \ldots, i_{n-r}$ mit $i_a < i_b$ für $a < b$, so dass $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{i_1}, \ldots, \mathfrak{b}_{i_{n-r}}$ eine Basis ist. Mit "Umnummerieren" können wir "OE" (ohne Einschränkung) annehmen, dass die betrachtete Basis $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{r+1}, \ldots, \mathfrak{b}_n$ ist.

Nun gibt es eine eindeutige Linearkombination

$$\mathfrak{v}_{r+1} = a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r + a_{r+1}\mathfrak{b}_{r+1} + \cdots + a_n\mathfrak{b}_n$$

mit $a_1, \ldots, a_n \in K$. Da die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_{r+1}$ linear unabhängig sind, sind nicht alle $a_{r+1}, \ldots, a_n$ gleich 0. Sei $s \in \{r + 1, \ldots, n\}$, so dass $a_s \neq 0$. Wir nehmen OE an, dass $s = r + 1$ ist. Wir halten hier fest: Da die Linearkombination eindeutig ist (d.h. die Koeffizienten sind eindeutig), liegt $\mathfrak{v}_{r+1}$ nicht in $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{r+2}, \ldots, \mathfrak{b}_n \rangle_K$, d.h. $\mathfrak{v}_{r+1}$ ist linear unabhängig von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{r+2}, \ldots, \mathfrak{b}_n$.

Nun betrachten wir das System

$$\mathfrak{v}_1, \ldots, \mathfrak{v}_{r+1}, \mathfrak{b}_{r+2}, \ldots, \mathfrak{b}_n \ .$$

Wir behaupten, dass dies eine Basis ist.

Es ist ein Erzeugendensystem, denn: Der Raum $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_{r+1}, \mathfrak{b}_{r+2}, \ldots, \mathfrak{b}_n \rangle_K$ enthält den Vektor $\mathfrak{b}_{r+1}$ (obige Linearkombination umstellen und durch a_{r+1} teilen). Somit enthält $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_{r+1}, \mathfrak{b}_{r+2}, \ldots, \mathfrak{b}_n \rangle_K$ den Raum $\langle \mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{r+1}, \ldots, \mathfrak{b}_n \rangle_K = V$.

Es ist linear unabhängig, denn: Seien $c_1, \ldots, c_n \in K$ mit

$$c_1\mathfrak{v}_1 + \cdots + c_{r+1}\mathfrak{v}_{r+1} + c_{r+2}\mathfrak{b}_{r+2} + \cdots + c_n\mathfrak{b}_n = \mathfrak{o} \, .$$

Da $\mathfrak{v}_{r+1}$ linear unabhängig von $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{r+2}, \ldots, \mathfrak{b}_n$ ist, ist $c_{r+1} = 0$. Da die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r, \mathfrak{b}_{r+2}, \ldots, \mathfrak{b}_n$ linear unabhängig sind, sind dann alle Koeffizienten gleich 0. □

Aufgrund dieses Satzes erhält man sofort:

Satz 2.2 *Sei V ein endlich erzeugter Vektorraum. Dann haben je zwei Basen von V die gleiche Anzahl von Elementen.*

Aufgrund dieses Satzes können wir definieren:

Definition Die *Dimension* von V ist wie folgt gegeben: Falls V endlich erzeugt ist, ist sie die Anzahl der Elemente einer (jeder) Basis von V. Wenn V nicht endlich erzeugt ist, ist sie unendlich. Die Dimension von V bezeichnet man mit $\mathrm{Dim}_K(V)$ oder $\mathrm{Dim}(V)$.

Beispiel 2.17 Für $n \in \mathbb{N}_0$ hat der Raum K^n hat die Dimension n. (Das rechtfertigt den Namen "n-dimensionaler Standardvektorraum".)

Die folgende Aussage ist nun sehr leicht zu zeigen:

Aussage 2.18

a) Seien $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ linear unabhängig. Dann ist $r \leq \mathrm{Dim}(V)$. Ferner ist dann $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ genau dann eine Basis, wenn $r = \mathrm{Dim}(V)$.

b) Sei $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ ein Erzeugendensystem. Dann ist $r \geq \mathrm{Dim}(V)$. Ferner ist dann $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ genau dann eine Basis, wenn $r = \mathrm{Dim}(V)$.

Beachten Sie, dass diese Aussage auch für unendlich-dimensionale Vektorräume Sinn macht (wobei $x < \infty$ für alle $x \in \mathbb{N}_0$).

Sei nun W ein weiterer K-Vektorraum, und sei $\varphi : V \longrightarrow W$ eine lineare Abbildung.

Aussage 2.19

a) Wenn $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *linear unabhängig sind und* φ *injektiv ist, dann sind auch* $\varphi(\mathfrak{v}_1), \ldots, \varphi(\mathfrak{v}_r)$ *linear unabhängig.*

b) Wenn $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *ein Erzeugendensystem von* V *bilden und* φ *surjektiv ist, dann bilden* $\varphi(\mathfrak{v}_1), \ldots, \varphi(\mathfrak{v}_r)$ *ein Erzeugendensystem von* W*.*

c) Wenn $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ *eine Basis von* V *ist und* φ *ein Isomorphismus ist, dann ist* $\varphi(\mathfrak{v}_1), \ldots, \varphi(\mathfrak{v}_r)$ *eine Basis von* W*.*

Beweis.

a) Es seien $a_1, \ldots, a_r \in K$ mit $a_1\varphi(\mathfrak{v}_1) + \cdots + a_r\varphi(\mathfrak{v}_r) = \mathfrak{o}$. Dann ist $\varphi(a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r) = \mathfrak{o}$. Da φ injektiv ist, ist dann $a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r = \mathfrak{o}$. Und da $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ linear unabhängig sind, ist somit $a_1 = \cdots = a_r = 0$.

b) Es sei $\mathfrak{w} \in V$ beliebig. Dann gibt es nach Voraussetzung ein $\mathfrak{v} \in V$ mit $\mathfrak{w} = \varphi(\mathfrak{v})$. Und nun gibt es $a_1, \ldots, a_r \in K$ mit $\mathfrak{v} = a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r$. Hieraus folgt: $\mathfrak{w} = \varphi(a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r) = a_1\varphi(\mathfrak{v}_1) + \cdots a_r\varphi(\mathfrak{v}_r)$.

c) folgt sofort aus a) und b). □

Aus den beiden letzten Aussagen folgt sofort:

Aussage 2.20

a) Wenn φ *injektiv ist, dann ist* $\mathrm{Dim}(V) \leq \mathrm{Dim}(W)$*.*

b) Wenn φ *surjektiv ist, dann ist* $\mathrm{Dim}(V) \geq \mathrm{Dim}(W)$*.*

c) Wenn φ *ein Isomorphismus ist, dann ist* $\mathrm{Dim}(V) = \mathrm{Dim}(W)$*.*

Basen sind insbesondere deshalb interessant, weil man sie benutzen kann, um lineare Abbildungen zu definieren:

Aussage 2.21 *Sei* $\mathfrak{b}_1, \ldots, \mathfrak{b}_r \in V$ *eine Basis. Sei* W *ein weiterer* K*-Vektorraum, und seien* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r \in W$*. Dann gibt es genau eine lineare Abbildung* $\varphi : V \longrightarrow W$ *mit* $\varphi(\mathfrak{b}_i) = \mathfrak{x}_i$ *für alle* $i = 1, \ldots, r$*, und diese ist durch*

$$\varphi(a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r) = a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$$

für $a_1, \ldots, a_r \in K$ *gegeben. Dabei gilt* $\mathrm{Bild}(\varphi) = \langle \mathfrak{x}_1, \ldots, \mathfrak{x}_r \rangle_K$*. Ferner ist* φ

- *genau dann injektiv, wenn die Vektoren* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *linear unabhängig sind,*
- *genau dann surjektiv, wenn die Vektoren* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *ein Erzeugendensystem bilden,*

- *genau dann bijektiv, wenn die Vektoren* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *eine Basis bilden.*

Beweis. Wir zeigen zuerst die Eindeutigkeit. Sei $\mathfrak{v} \in V$. Dann gibt es (eindeutig bestimmte) $a_1, \ldots, a_r \in K$ mit $\mathfrak{v} = a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r$. Nun ist $\varphi(\mathfrak{v}) = \varphi(a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r) = a_1\varphi(\mathfrak{b}_1) + \cdots + a_r\varphi(\mathfrak{b}_r) = a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$.

Nun müssen wir noch nachweisen, dass es tatsächlich so eine lineare Abbildung gibt. Sei dazu wiederum $\mathfrak{v} \in V$ beliebig. Wie schon gesagt gibt es eindeutig bestimmte $a_1, \ldots, a_r \in K$ mit $\mathfrak{v} = a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r$. Wir setzen nun $\varphi(\mathfrak{v}) := a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$. Dies ist wohldefiniert, da die "Koeffizienten" $a_1, \ldots, a_r$ eindeutig sind.

Wir müssen noch die Linearität nachweisen.

Seien dazu $\mathfrak{v}, \mathfrak{w} \in V$ und $c \in K$. Sei $\mathfrak{v} = a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r$ und $\mathfrak{w} = b_1\mathfrak{b}_1 + \cdots + b_r\mathfrak{b}_r$.

Dann ist $\varphi(c \cdot \mathfrak{v} + \mathfrak{w}) = \varphi(ca_1\mathfrak{b}_1 + \cdots + ca_r\mathfrak{b}_r + b_1\mathfrak{b}_1 + \cdots + b_r\mathfrak{b}_r) = \varphi((ca_1 + b_1)\mathfrak{b}_1 + \cdots + (ca_r + b_r)\mathfrak{b}_r) \overset{\text{per Def.}}{=} (ca_1 + b_1)\mathfrak{x}_1 + \cdots + (ca_r + b_r)\mathfrak{x}_r = c(a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r) + (b_1\mathfrak{x}_1 + \cdots + b_r\mathfrak{x}_r) \overset{\text{per Def.}}{=} c\varphi(a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r) + \varphi(b_1\mathfrak{b}_1 + \cdots + b_r\mathfrak{b}_r) = c\varphi(\mathfrak{v}) + \varphi(\mathfrak{w})$.

Die Aussage über das Bild folgt sofort aus der Definition von φ. Die weiteren Aussagen sind leicht zu zeigen (und teilweise auch schon in Aussage 2.19 gezeigt worden. □

Als Spezialfall der obigen Aussage erhalten wir: Sei V ein K-Vektorraum, und seien $\mathfrak{v}_1, \ldots, \mathfrak{v}_k \in V$. Dann gibt es genau eine lineare Abbildung $\varphi : K^r \longrightarrow V$ mit $\varphi(\underline{e}_i) = \mathfrak{v}_i$ für $i = 1, \ldots, r$. Explizit ist diese Abbildung durch

$$\varphi\left(\begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}\right) = a_1\mathfrak{v}_1 + \cdots + a_r\mathfrak{v}_r$$

gegeben. Diese Abbildung ist genau dann injektiv, wenn die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$ linear unabhängig sind. Sie ist genau dann surjektiv, wenn die Vektoren ein Erzeugendensystem bilden und genau dann ein Isomorphismus, wenn die Vektoren eine Basis bilden.

Ein weiterer wichtiger Spezialfall ist wie folgt: Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_r \in V$ eine Basis von V. Dann haben wir eine eindeutig bestimmte lineare Abbildung $c : V \longrightarrow K^r$ mit $c(\mathfrak{b}_i) = \underline{e}_i$; diese Abbildung ist ein Isomorphismus und erfüllt

$$c(a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} .$$

Der Vektor $\underline{a} \in K^r$ heißt dann der *Koordinatenvektor* zum Vektor $a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r \in V$, und die Abbildung c heißt *Koordinatenabbildung*. (Achtung: Beide Begriffe beziehen sich auf eine feste Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_r$!) Es ist naheliegend, den Koordinatenvektor von $\mathfrak{v} \in V$ mit $\underline{v}$ zu bezeichnen.

2.5 Lineare Gleichungssysteme und der Gauß-Algorithmus

Ein *lineares Gleichungssystem (LGS)* über K ist ein Gleichungssystem der Form

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & b_1 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & b_2 \\ \vdots \qquad\qquad\quad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & b_m \end{array} \tag{2.7}$$

mit $a_{i,j} \in K$ und den "Unbestimmten" $X_1, \ldots, X_n$.

So einem linearen Gleichungssystem (2.7) ordnet man die so genannte *Koeffizientenmatrix* $A = ((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n} \in K^{m\times n}$ zu. Die so genannte *erweiterte Koeffizientenmatrix* ist dann die Matrix $(A|\underline{b}) \in K^{m\times(n+1)}$. Das LGS kann man dann auch so schreiben:

$$A \cdot \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$$

Wir wollen nun die Lösungsmenge so eines Gleichungssystems explizit bestimmen. Zunächst eine Wiederholung.

Definition Ein lineares Gleichungssystem der Form

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & 0 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & 0 \\ \vdots \qquad\qquad\quad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & 0 \end{array} \tag{2.8}$$

(d.h. mit "trivialer rechter Seite") heißt *homogenes lineares Gleichungssystem*. Ein lineares Gleichungssystem wie (2.7) heißt auch *inhomogenes lineares Gleichungssystem*.

Wenn man ein konkretes inhomogenes System (2.7) zugrundelegt und die rechte Seite durch Nullen ersetzt, erhält man das *zugehörige homogene lineare Gleichungssystem*.

Bemerkung Ein "inhomogenes System" kann auch eine "triviale rechte Seite" haben, d.h. ein "homogenes System" ist ein Spezialfall eines inhomogenen Systems.

Notation Gegeben ein LGS, wird die Lösungsmenge mit $\mathbb{L}$ bezeichnet. Die Lösungsmenge des zugehörigen homogenen LGS wird mit $\mathbb{L}_h$ bezeichnet.

Definition Sei $A \subseteq K^n$ eine Teilmenge und $\underline{x} \in K^n$. Dann definieren wir

$$\underline{x} + A := \{\underline{x} + \underline{a} \mid \underline{a} \in A\} .$$

Bemerkung Man kann (insbesondere für $K = \mathbb{R}$ und $n = 2$ oder $n = 3$) die Menge $\underline{x} + A$ als eine Parallelverschiebung von A interpretieren.

Aussage 2.22 *Sei ein LGS (2.7) gegeben. Wir nehmen an, dass das LGS lösbar ist und fixieren eine Lösung $\underline{x}_0$. Dann gilt*

$$\mathbb{L} = \underline{x}_0 + \mathbb{L}_h .$$

Der *Beweis* ist einfach.

Bemerkung Eine Lösung $\underline{x}_0$ wie in der Aussage heißt auch "spezielle Lösung". Man sagt: Man erhält alle Lösungen eines inhomogenen LGS, indem man eine spezielle Lösung sowie alle Lösungen des zugehörigen homogenen LGS berechnet.

Aussage 2.23 *Die Lösungsmenge eines homogenen LGS in n Variablen ist ein linearer Unterraum (ein Untervektorraum) von K^n.*

Der *Beweis* ist wiederum einfach.

Definition Sei A eine Teilmenge von K^n. Dann ist A ein *affiner Unterraum* von K^n, falls A leer ist oder es ein $\underline{x}_0 \in K^n$ und einen linearen Unterraum U von K^n mit $A = \underline{x}_0 + U$ gibt.

Lemma 2.24 *Seien $\underline{x}_0, \underline{y}_0 \in K^n$ und $U, V \subseteq K^n$ lineare Unterräume mit $\underline{x}_0 + U = \underline{y}_0 + V$. Dann ist $U = V$ und $\underline{x}_0 - \underline{y}_0 \in U$.*

Beweis. Es ist $(\underline{y}_0 - \underline{x}_0) + V = U$. Da $\underline{0} \in V$ ist $\underline{y}_0 - \underline{x}_0 \in U$ (und somit auch $\underline{x}_0 - \underline{y}_0 \in U$). Analog zeigt man, dass auch $\underline{x}_0 - \underline{y}_0 \in V$.

Sei nun $\underline{u} \in U$. Dann ist $\underline{x}_0 - \underline{y}_0 + \underline{u} \in V$. Somit ist auch $\underline{u} = \underline{x}_0 - \underline{y}_0 + \underline{u} - (\underline{x}_0 - \underline{y}_0) \in V$. Wir haben gesehen, dass $U \subseteq V$. Analog zeigt man, dass $V \subseteq U$. □

Definition Sei A ein nicht-leerer affiner Raum mit $A = \underline{x}_0 + U$. Dann heißt U der *zu A gehörige lineare Unterraum*. Dieser Unterraum wird mit U_A bezeichnet.

Beachten Sie, dass nach dem obigen Lemma dieser lineare Unterraum nur von A abhängt.

Bemerkung Die zwei letzten Definitionen und das Lemma machen auch in beliebigen Vektorräumen Sinn.

Der Zusammenhang mit linearen Gleichungssystemen ist durch die folgende Aussage gegeben.

Aussage 2.25 *Die Lösungsmenge $\mathbb{L}$ eines inhomogenen LGS ist stets ein affiner Unterraum. Falls $\mathbb{L}$ nicht-leer ist, ist die Lösungsmenge $\mathbb{L}_h$ des zugehörigen homogenen LGS gleich dem zu diesem affinen Unterraum gehörigen linearen Unterraum.*

Dies folgt aus Aussage 2.23 und Aussage 2.22. □

Der Gauß-Algorithmus

Gegeben sei also ein lineares Gleichungssystem

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & b_1 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & b_2 \\ \vdots \qquad\qquad\quad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & b_m \end{array} \tag{2.9}$$

über dem Körper K. Die Aufgabe besteht darin, zu entscheiden, ob das System lösbar ist, und gegebenenfalls eine Lösung $\underline{x}_0$ (genannt "spezielle Lösung") sowie eine Basis $\underline{x}_1, \ldots, \underline{x}_r$ des Lösungsraums des zugehörigen homogenen Systems zu finden. Der Lösungsraum des Systems ist dann der affine Raum

$$\underline{x}_0 + \langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K \ .$$

Wir sagen, dass zwei lineare Gleichungssysteme *äquivalent* sind, wenn ihre Lösungsmengen gleich sind. Der *vollständige Gauß-Algorithmus* (auch *Gauß-Jordan-Algorithmus* genannt) besteht nun darin, das System solange mittels bestimmter (einfacher) Operationen umzuformen, bis man eine "spezielle Lösung" sowie eine Basis des zugehörigen homogenen Systems ablesen kann.

Wir betrachten hierzu die folgenden drei Operationen.

(I) Multiplikation einer Gleichung mit einem Körperelement $\neq 0$.

(II) Vertauschen von zwei Gleichungen.
(III) Addition von c-mal Gleichung i zu Gleichung j (wobei $i \neq j$ und $c \in K$).

Zur Verdeutlichung: Die Operationen (I) und (III) sind konkret wie folgt gegeben:

(I) Sei $c \in K - \{0\}$ und $i = 1, \ldots, m$. Dann wird die i-te Gleichung aus (2.9) durch die Gleichung

$$ca_{i,1}X_1 + \cdots + ca_{i,n}X_n = cb_i$$

ersetzt.

(III) Sei $c \in K$ (c muss nicht $\neq 0$ sein, aber wenn es 0 ist, passiert nichts), und seien $i, j = 1, \ldots, m$ mit $i \neq j$. Dann wird die j-te Gleichung aus (2.9) durch die Gleichung

$$(a_{j,1} + ca_{i,1})X_1 + \cdots + (a_{j,n} + ca_{i,n})X_n = (b_j + cb_i)$$

ersetzt.

Lemma 2.26 *Operationen (I), (II), (III) überführen ein lineares Gleichungssystem in ein äquivalentes lineares Gleichungssystem.*

Beweis. Die Aussage ist offensichtlich für Operationen (I) und (II), wir betrachten also Operation (III).

Seien hierzu $i, j = 1, \ldots, k (i \neq j)$ und $c \in K$. Sei nun $\underline{x}$ eine Lösung von (2.9). Dann gilt insbesondere

$$a_{i,1}x_1 + \cdots + a_{i,n}x_n = b_i$$

und

$$a_{j,1}x_1 + \cdots + a_{j,n}x_n = b_j \ .$$

Hieraus folgt:

$$(a_{j,1} + ca_{i,1})x_1 + \cdots + (a_{j,n} + ca_{i,n})x_n = (b_j + cb_i) \ ,$$

und somit erfüllt $\underline{x}$ auch das umgeformte System.

Sei nun umgekehrt $\underline{x}$ eine Lösung des umgeformten Systems. Dann gilt insbesondere

$$a_{i,1}x_1 + \cdots + a_{i,n}x_n = b_i$$

und

$$(a_{j,1} + ca_{i,1})x_1 + \cdots + (a_{j,n} + ca_{i,n})x_n = (b_j + cb_i) \ .$$

Damit gilt auch

$$a_{j,1}x_1 + \cdots + a_{j,n}x_n = b_j \ .$$

Außerdem erfüllt $\underline{x}$ auch die Gleichungen $1, \ldots, j-1, j+1, \ldots, m$ des ursprünglichen Systems, weil diese nicht verändert werden. □

Definition Die Operationen (I), (II), (III) heißen *elementare Umformungen.*

Der Gauß-Algorithmus besteht nun darin, mittels der Operationen (I), (II), (III) ein System (2.9) in ein System in "Treppenform" umzuformen. Beim Gauß-Jordan-Algorithmus rechnet man noch ein wenig weiter, bis man ein System in einer bestimmten, sehr einfachen, "Treppenform" hat. Neben den elementaren Umformungen ist es noch zweckmäßig, Gleichungen der Form $0X_1 + \cdots + 0X_n = 0$ ("Nullzeilen") wegzulassen.

Ich verdeutliche die Algorithmen an einem Beispiel.

Beispiel 2.27 Die Lösungsmenge des folgenden Systems über $\mathbb{Q}$ sei gesucht.

$$\begin{array}{rrrrrr} X_1 & -X_2 & & +X_4 & +X_5 & = 1 \\ -X_1 & & +X_3 & -2X_4 & -X_5 & = 0 \\ & 2X_2 & +X_3 & -X_4 & +3X_5 & = 1 \\ -2X_1 & +3X_2 & +X_3 & -3X_4 & & = -1 \end{array}$$

Für die Rechnungen ist es zweckmäßig, das System in der folgenden symbolischen Form hinzuschreiben.

$$\begin{array}{rrrrr|r} X_1 & X_2 & X_3 & X_4 & X_5 & \\ 1 & -1 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -2 & -1 & 0 \\ 0 & 2 & 1 & -1 & 3 & 1 \\ -2 & 3 & 1 & -3 & 0 & -1 \end{array} \tag{2.10}$$

Die Variablen in der ersten Zeile werden wir im Folgenden auch weglassen. Wir wenden Operation (III) an: Wir addieren die erste Zeile zur zweiten, und wir addieren das 2-fache der ersten Zeile zur vierten. Wir erhalten:

$$\begin{array}{rrrrr|r} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & -1 & 0 & 1 \\ 0 & 2 & 1 & -1 & 3 & 1 \\ 0 & 1 & 1 & -1 & 2 & 1 \end{array}$$

Wir addieren nun $2\times$ die zweite Zeile zur dritten Zeile sowie die zweite Zeile zur vierten und erhalten:

$$\begin{array}{rrrrr|r} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & -1 & 0 & 1 \\ 0 & 0 & 3 & -3 & 3 & 3 \\ 0 & 0 & 2 & -2 & 2 & 2 \end{array}$$

Wir multiplizieren die zweite Zeile mit -1, die dritte mit $\frac{1}{3}$ und die vierte mit $\frac{1}{2}$ und erhalten:

$$\left.\begin{array}{ccccc|c} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right.$$

Nun steht in der vierten Zeile das Gleiche wie in der dritten, und wir können die vierte Zeile weglassen. (Mittels der elementaren Umformungen geht das so: Wir addieren $(-1)\times$ die dritte Zeile zur vierten. Dann erhalten wir eine "Null-Zeile", und diese können wir weglassen.) Wir erhalten:

$$\left.\begin{array}{ccccc|c} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right.$$

Dies ist ein lineares Gleichungssystem in so genannter *Treppenform* oder *(Zeilen-)Stufenform* (siehe unten für Definition).

Die Lösungsmenge kann nun durch "Auflösen" bestimmt werden. Hierzu geben wir uns einen belieben Vektor $\underline{x} \in \mathbb{Q}^n$ vor und setzen $\lambda := x_4, \mu := x_5$. Dann ist $\underline{x}$ genau dann eine Lösung, wenn gilt:

$$\begin{aligned} x_3 &= \lambda - \mu + 1 \\ x_2 &= x_3 - x_4 - 1 = (\lambda - \mu + 1) - \lambda - 1 = -\mu \\ x_1 &= x_2 - x_4 - x_5 + 1 = -\mu - \lambda - \mu + 1 = -\lambda - 2\mu + 1 \ . \end{aligned}$$

Damit ist die Lösungsmenge

$$\left\{ \begin{pmatrix} -\lambda - 2\mu + 1 \\ -\mu \\ \lambda - \mu + 1 \\ \lambda \\ \mu \end{pmatrix} | \lambda, \mu \in \mathbb{Q} \right\} =$$

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} -2 \\ -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \mid \lambda, \mu \in \mathbb{Q} \right\}$$

Der soeben durchgeführte Algorithmus heißt *Gauß-Algorithmus*.

Das "Auflöseverfahren" ist jedoch recht unübersichtlich und fehleranfällig. Besser ist es, noch ein wenig mit elementaren Umformungen weiterzurechnen. Das Ziel ist, dass auf allen "Treppenstufen" eine 1 steht (das sind hier

die Elemente mit Indices (1,1), (2,2), (3,3) und das ist hier schon der Fall), und dass über all diesen "Treppenstufen" nur Nullen stehen. Das bedeutet hier, dass die Elemente mit Indices $(1,2)$, $(1,3)$ und $(2,3)$ Null sein sollen.

Hierzu addieren wir zuerst die dritte Zeile zur zweiten und erhalten:

$$\left(\begin{array}{ccccc|c} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right)$$

Jetzt addieren wir die zweite zur ersten und erhalten:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right)$$

Dies ist ein System in so genannter *reduzierter (Zeilen-)Stufenform* oder *reduzierter Treppenform* (siehe unten für Definition).

Hieraus kann man direkt ablesen, dass $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ eine "spezielle Lösung" ist.

Gesucht ist nunnoch eine Basis des Lösungsraums des zugehörigen homogenen Systems:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right)$$

(Die Spalte rechts des Strichs kann man nun weglassen.)

Nun gibt es einen "Ablesetrick", der wie folgt funktioniert: Man fügt neue *Zeilen* ein, und zwar so: Die Zeilen enthalten nur Nullen und genau eine -1. Sie werden so eingefügt, dass man ein System mit gleich viel Spalten wie Zeilen erhält, das die folgenden Eigenschaften hat: Unter der Diagonalen stehen nur Nullen und jeder Eintrag auf der Diagonalen ist entweder eine 1 oder eine -1. In unserem Fall sieht das dann so aus:

$$\begin{array}{rccccc} & 1 & 0 & 0 & 1 & 2 \\ & 0 & 1 & 0 & 0 & 1 \\ & 0 & 0 & 1 & -1 & 1 \\ \text{eingefügt} \rightarrow & 0 & 0 & 0 & -1 & 0 \\ \text{eingefügt} \rightarrow & 0 & 0 & 0 & 0 & -1 \\ & & & & \uparrow & \uparrow \end{array}$$

(Wenn wir die Null-Zeile nicht gestrichen hätten, wäre diese nun weggefallen.)

Diejenigen *Spalten*, in denen eine -1 eingefügt wurde (mit ↑ gekennzeichnet), bilden nun eine Basis von $\mathbb{L}_h$. In unserem Fall sind dies die Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} .$$

Somit ist die Lösungsmenge des ursprünglichen Systems gleich

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \langle \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} \rangle_{\mathbb{Q}} .$$

Vergleichen Sie dies mit der zuerst berechneten Darstellung der Lösungsmenge!

Überlegen Sie sich auch, warum der "Ablesetrick" allgemein funktioniert!

Eine große Bitte. Der "Ablesetrick" (einfügen der "−1-Zeilen") ist eine grundsätzlich andere Operation als die vorangegangenen elementaren Umformungen. Deshalb sollte man das Einfügen auch entsprechend kenntlich machen, z.B. indem man schreibt "Ich wende nun den 'Ablesetrick' an." und / oder indem man die eingefügten "−1-Zeilen" mit Bleistift schreibt.

Für die weitere Beschreibung benutze ich Matrizen. Zunächst einige Definitionen:

Definition Eine Matrix in *(Zeilen-)Stufenform* bzw. in *Treppenform* ist eine Matrix der Gestalt

$$\begin{pmatrix} * & \circ & \cdots & \circ & \circ & \circ & \cdots & \circ & \circ & \circ & \cdots & \circ & \cdots & \circ & \circ & \cdots & \circ \\ & & & & * & \circ & \cdots & \circ & \circ & \circ & \cdots & \circ & \cdots & \circ & \circ & \cdots & \circ \\ & & & & & & & & * & \circ & \cdots & \circ & \cdots & \circ & \circ & \cdots & \circ \\ & & & & & & & & & & & & \ddots & \vdots & \vdots & \cdot & \vdots \\ & & & & & & & & & & & & & * & \circ & \cdots & \circ \\ & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & \end{pmatrix},$$

wobei die mit $*$ bezeichneten Einträge alle $\neq 0$ sind, die mit $\circ$ bezeichneten Einträge beliebig sind, und die Einträge ohne Bezeichnung (d.h. die Einträge unter der "Treppe" 0 sind.

Mit anderen Worten: So eine Matrix hat eine "Treppe", wobei jede Treppenstufe ein Eintrag $\neq 0$ ist. Jede Treppenstufe hat Höhe 1, und neben jeder Stufe dürfen beliebige Einträge stehen. Unter der Treppe stehen nur Nullen.

Beachten Sie, dass links 0-Spalten und unten 0-Zeilen stehen dürfen (aber nicht müssen) (diese Spalten bzw. Zeilen sind durch den freien Platz links und unten angedeutet.)

Definition Eine Matrix in *reduzierter (Zeilen-)Stufenform* bzw. in *reduzierter Treppenform* ist eine Matrix der Gestalt

$$\begin{pmatrix}
1 & \circ & \cdots & \circ & 0 & \circ & \cdots & \circ & 0 & \circ & \cdots & \circ & \cdots & 0 & \circ & \cdots & \circ \\
 & & & & 1 & \circ & \cdots & \circ & 0 & \circ & \cdots & \circ & \cdots & 0 & \circ & \cdots & \circ \\
 & & & & & & & & 1 & \circ & \cdots & \circ & \cdots & 0 & \circ & \cdots & \circ \\
 & & & & & & & & & & & & \ddots & \vdots & \vdots & . & \vdots \\
 & & & & & & & & & & & & & 1 & \circ & \cdots & \circ \\
 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & &
\end{pmatrix},$$

wobei die $\circ$'s wieder beliebige Einträge repräsentieren und unter der "Treppe" wiederum nur Nullen stehen. Mit anderen Worten: Eine Matrix in reduzierter (Zeilen-)Stufenform ist eine Matrix in (Zeilen-)Stufenform, so dass
- die "Treppenstufen" alle gleich 1 sind,
- über den "Treppenstufen" nur Nullen stehen.

(Diese Bedingungen beziehen sich wirklich nur auf die Stufen, nicht auf die Elemente, die rechts daneben stehen!)

Ich gebe nun eine formalisiertere Beschreibung des Gauß-Algorithmus zur Transformation einer Matrix in eine Matrix in (Zeilen-)Stufenform an. Ich wähle eine rekursive Beschreibung. Man würde den Algorithmus allerdings eher wohl mit Schleifen implementieren. In dem folgenden Algorithmus wird die Matrix ohne Kopien anzufertigen fortlaufend transformiert. (D.h. bei den rekursiven Aufrufen werden keine Kopien der Matrix (oder Teile der Matrix) angefertigt.)

Der Gauß-Algorithmus

Eingabe. Eine Matrix $A \in K^{m \times n}$ mit $m, n \in \mathbb{N}$.
Ausgabe. Eine Matrix $\tilde{A}$, die aus A durch elementare Zeilentransformationen hervorgeht und in (Zeilen-)Stufenform ist.

Wenn die erste Spalte eine Nullspalte ist,
  wenn die Matrix mehr als eine Spalte hat,
    wende den Algorithmus auf die Matrix an, die entsteht,
      indem man die erste Spalte streicht.
Ansonsten
  Wähle ein i, so dass $a_{i,1} \neq 0$.
  Vertausche die Zeilen 1 und i (Transformation (II)).
  (Für die Matrix gilt nun, dass $a_{1,1} \neq 0$.)
  Multipliziere eventuell die erste Zeile mit einer Konstanten (z.B. mit $a_{1,1}^{-1}$) (Transformation (I)).
  Für $i = 2, \ldots, m$: Addiere jeweils $-\frac{a_{i,1}}{a_{1,1}}$-mal die erste Zeile zur i-ten Zeile (Transformation (III)).
  Wenn die Matrix mehr als eine Zeile und mehr als eine Spalte hat,
    wende den Algorithmus auf die Matrix an, die entsteht,
      indem man die erste Zeile und die erste Spalte streicht.

Wenn die erste Spalte keine Nullspalte ist, sieht die Matrix nach dem vorletzten Schritt so aus:

$$\begin{pmatrix} * & \circ & \cdots & \circ \\ 0 & \circ & \cdots & \circ \\ \vdots & \vdots & \cdot & \vdots \\ 0 & \circ & \cdots & \circ \end{pmatrix} \tag{2.11}$$

(Mit $* \neq 0$ und $\circ$ beliebig.) Im letzten Schritt wird dann der Algorithmus mit der Matrix aufgerufen, die entsteht, wenn man in dieser Matrix die erste Zeile und die erste Spalte weglässt.

Sicher terminiert der Algorithmus, und der Algorithmus ist offensichtlich korrekt: Das Ergebnis des Algorithmus ist offensichtlich eine Matrix in (Zeilen-)Stufenform, die aus der ursprünglichen Matrix durch elementare Zeilentransformationen hervorgegangen ist.

Um eine Matrix in reduzierter (Zeilen-)Stufenform zu erhalten, geht man ausgehend von einer Matrix in (Zeilen-)Stufenform wie folgt vor:

Zuerst teilt man alle nicht-Nullzeilen durch ihren ersten Eintrag $\neq 0$. Dann sind also die Einträge auf allen Stufen gleich 1.

Dann "räumt" man mittels Operationen von Typ (III) die Einträge oberhalb der Stufen aus (man erzeugt Nullen). Dabei kann man die "Stufenspalten" in beliebiger Reihenfolge behandeln.

Das Ergebnis ist eine Matrix in reduzierter (Zeilen-)Stufenform. Der gesamte soeben beschriebene Algorithmus heißt *vollständiger Gauß-Algorithmus* oder *Gauß-Jordan-Algorithmus*.

Sei nun das LGS

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$$

gegeben. Um die Lösungsmenge zu bestimmen (genauer um eine spezielle Lösung und eine Basis des zugehörigen homogenen Systems zu bestimmen), kann man nun wie folgt vorgehen:

- Man wendet den Gauß-Algorithmus auf die Matrix $(A|\underline{b})$ an. Sei $(\tilde{A}|\tilde{b})$ das Ergebnis. Dann ist das System genau dann lösbar, wenn in der letzten Spalte keine Treppenstufe ist.

Wenn das System lösbar ist,
- erzeugt man eine Matrix in reduzierter (Zeilen-)Stufenform.
- liest man die Lösung mittels des "Ablesetricks" ab.

Bemerkung Oftmals wird zum Lösen linearer Gleichungssysteme noch eine weitere Operation erlaubt: Man vertauscht die Spalten der "linken Seite" des LGS und merkt sich dabei, welche Variable zu welcher Spalte gehört. (Die Variablen sollte man dann beim Rechnen "von Hand" über das System schreiben, so wie in (2.10).) Ich empfehle allerdings dringend, diese Operation nur in Ausnahmefällen anzuwenden, da sie sehr fehlerbehaftet ist. Dies gilt insbesondere für die Klausur!

Diskussion Das Wort "Algorithmus" besagt, dass zu jedem Zeitpunkt genau festgelegt sein muss, welche Operation als nächstes ausgeführt wird. In dem oben beschriebenen Gauß-Algorithmus sind allerdings wesentliche Unbestimmtheiten vorhanden:[3]
- Es ist nicht festgelegt, welche Zeile man "nach oben holt".

[3]Immer wenn man einen Algorithmus in Pseudocode angibt, handelt man sich gewisse Unbestimmtheiten ein, oder anders ausgedrückt, man läßt gewisse Freiheiten bei der Implementierung. Somit ist die Abgrenzung, ob man nun einen Algorithmus oder nur ein Algorihmenschema hat, etwas ungenau.

- Es ist nicht festgelegt, unter welchen Bedingungen man eine Transformation (I) durchführen soll, und wenn, mit welcher Konstante.
Somit handelt es sich streng genommen nicht um einen Algorithmus sondern eher um ein *Algorithmenschema*.

Man erhält einen Algorithmus, indem man eine konkrete Regel vorgibt, welche Zeile ausgewählt werden soll und nach dem Vertauschen die erste Zeile mit $a_{1,1}^{-1}$ multipliziert.

Die Auswahl so einer Zeile heißt *Pivotwahl*, und dasjenige Element so einer Zeile, das für die Elimination benutzt wird (in der obigen Darstellung das erste Element), wird *Piviotelement* genannt. Eine Möglichkeit für die Pivotwahl (in der obigen Darstellung des Gauß-Algorithmus) ist, das kleinste i mit $a_{1,i} \neq 0$ zu bestimmen und dann die 1-ste und die i-te Zeile zu vertauschen. (Dies bezieht sich wirklich nur auf die obige rekursive Darstellung des Algorithmus. Natürlich wählt man keine Zeile derjenigen Zeilen aus, die man schon behandelt hat.)

Für Rechnungen von Hand wird man jedoch eher eine Zeile mit einem "möglichst einfach aussehenden" ersten Eintrag nach oben holen.

Für approximative Rechnungen mit Fließkommazahlen (die reelle Zahlen darstellen) mittels eines Computers sollte man darauf achten, dass nur möglichst kleine Rundungsfehler auftreten. Hierzu ist es geschickt, ein Pivotelement mit möglichst großem Absolutbetrag zu wählen. (Dies hat etwas damit zu tun, dass man durch das Pivotelement teilen muss).

Wenn man ein LGS lösen will und auch das Vertauschen von Spalten zulässt, ist es also optimal, das betragmäßig größte Element der gesamten (restlichen) Matrix zu wählen und entsprechend Zeilen und Spalten (!) umzunummerieren (nicht umzukopieren!) (kein Problem auf dem Computer).

Komplexität Unter der *(Zeit)-Komplexität* eines Algorithmus versteht man eine Angabe der notwendigen Rechenschritte des Algorithmus. Dies erfordert eigentlich noch ein wenig Theorie. Insbesondere müsste man zunächst ein geeignetes formales *Rechenmodell* definieren. Ich gebe hier eine kurze Darstellung der Komplexität des Gauß-Algorithmus.

Sei eine $m \times n$-Matrix über K gegeben.

Offensichtlich benötigen wir höchstens $m-1$ Operationen vom Typ (II) und höchstens m Operationen vom Typ (I). Wir benötigen höchstens $(m-1)+(m-2)+\cdots+3+2+1 = \frac{m\cdot(m-1)}{2}$ Operationen vom Typ (III), um eine Matrix in (Zeilen)-Stufenform zu erzeugen. Danach benötigen wir noch höchstens $\frac{m\cdot(m-1)}{2}$ Operationen vom Typ (III), um eine Matrix in reduzierter (Zeilen-)Stufenform zu erzeugen.

Für jede dieser Operationen benötigen wir höchstens n Körperoperatio-

nen.

Damit erhalten wir:

Aussage 2.28 *Es gibt eine Konstante $C \in \mathbb{R}_{>0}$, so dass das Folgende gilt: Gegeben eine $m \times n$-Matrix $A \in K^{m \times n}$, kann man in $\leq C \cdot m^2 \cdot n$ Körperoperationen eine reduzierte (Zeilen-)Stufenform von A berechnen. Insbesondere gibt es eine Konstante $C' \in \mathbb{R}_{>0}$, so dass man die Lösungsmenge eines lineares Gleichungssystem mit n Unbestimmten und m Gleichungen in $\leq C' \cdot m^2 \cdot n$ Körperoperationen bestimmen kann.*

Grob gesagt hat das Lösungsverfahren mittels des Gauß-Algorithmus "kubische" Komplexität.

Ich erinnere noch einmal daran, dass "Bestimmen der Lösungsmenge" bedeutet, eine "spezielle Lösung" und eine Basis der Lösungsmenge des zugehörigen homogenen Systems anzugeben.

Bemerkung Durch die Angabe der Komplexität in Körperoperationen wird wenig über das Problem ausgesagt, wie man nun möglichst schnell und möglichst exakt die Lösungsmenge eines Gleichungssystems über den reellen Zahlen mit Fließkommazahlen approximativ berechnet. Dieses Problem haben wir oben kurz angedeutet. Hierzu und zu verwandten Fragen gibt es eine umfangreiche Theorie, die in die *Numerische Mathematik* fällt.

Anwendungen

Ich gebe jetzt noch Anwendungen des Gauß-Algorithmus auf die Frage, ob ein System von Vektoren *linear unabhängig* bzw. ein *Erzeugendensystem* ist.

Es seien $\underline{a}_1, \ldots, \underline{a}_r \in K^n$ gegeben. Wir betrachten nun die $n \times r$-Matrix $A = (\underline{a}_1 | \cdots | \underline{a}_r)$, die entsteht, wenn man die Vektoren $\underline{a}_1, \ldots \underline{a}_r$ als Spalten einer Matrix auffasst.

Nun sind die Vektoren $\underline{a}_1, \ldots, \underline{a}_r$ genau dann linear unabhängig, wenn das System

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix} = \underline{0} \tag{2.12}$$

ausschließlich die "triviale" Lösung $\underline{0}$ (und keine weitere Lösung) hat. Ob dies der Fall ist, kann man nun mit dem Gauß-Algorithmus überprüfen.

Sei nun $\tilde{A}$ eine Matrix in (Zeilen-)Stufenform, die aus A mittels elementarer Zeilenoperationen entsteht. Dann ist also das System (2.12) äquivalent

zum System

$$\tilde{A} \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix} = \underline{0} \,. \tag{2.13}$$

Dieses System hat genau dann eine "nicht-triviale" Lösung (eine Lösung $\neq \underline{0}$), wenn es Spalten von $\tilde{A}$ gibt, die (in $\tilde{A}$) keine "Treppenstufen" sind.

Mit den obigen Überlegungen können wir auch auf die folgende Frage eine algorithmische Antwort geben:

Gebeben ein System $\underline{a}_1, \ldots, \underline{a}_r \in K^n$, finde ein "Teilsystem", das eine Basis von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$ ist!

(Ein "Teilsystem" ist wie folgt definiert. Gegeben Seien $1 \leq i_1 < i_2 < \ldots < i_k \leq r$ für ein $k \leq r$. Dann ist $\underline{a}_{i_1}, \ldots, \underline{a}_{i_k}$ ein Teilsystem von $\underline{a}_1, \ldots, \underline{a}_r$.)

Hierzu definieren wir die Matrix A wie oben und berechnen mittels elementarer Zeilenumformungen eine Matrix $\tilde{A}$. Dann bilden diejenigen Spalten aus A (!), für die in $\tilde{A}$ eine Stufe steht, eine Basis von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$.

(Denn: Die "Nichtstufenspalten" von $\tilde{A}$ sind von den "Stufenspalten" von $\tilde{A}$ linear abhängig. Und das gilt auch für A, weil die Lösungsmenge eines LGS unter elementaren Zeilentransformationen eben nicht verändert wird. Außerdem sind die Stufenspalten von $\tilde{A}$ linear unabhängig, und das Gleiche gilt dann auch für die Stufenspalten von A.)

Wir kommen nun zu der Frage, wie man entscheiden kann, ob die $\underline{a}_1, \ldots, \underline{a}_r$ ein Erzeugendensystem von K^n bilden. Wieder betrachten wir die Matrix A, die definiert ist wie oben, sowie eine Matrix $\tilde{A}$ in (Zeilen-)Stufenform, die aus A durch elementare Transformationen hervorgeht.

Offensichtlich ist $\underline{a}_1, \ldots, \underline{a}_r$ genau dann ein Erzeugendensystem von K^n, wenn für alle $\underline{b} \in K^n$ das LGS

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix} = \underline{b} \tag{2.14}$$

lösbar ist. Diese Eigenschaft ist auch invariant unter elementaren Transformationen (warum?) und folglich äquivalent zu der Bedingung, dass das System (2.13) für alle $\underline{b} \in K^n$ lösbar ist. Dies wiederum ist dazu äquivalent, dass $\tilde{A}$ keine Nullzeile enthält.

Die Frage, ob wir eine *Basis* gegeben haben, kann nach dem Obigen wie folgt entschieden werden: Die folgenden Aussagen sind äquivalent:

- $\underline{a}_1, \ldots, \underline{a}_r$ ist eine Basis.
- Es gilt $r = n$ und in jeder Spalte von $\tilde{A}$ ist eine Stufe.

Sei nun $\tilde{A}$ eine Matrix, die aus A durch elementare Transformationen hervorgeht und in reduzierter Zeilenstufenform ist. Dann haben wir die Äquivalenzen:

- $\underline{a}_1, \ldots, \underline{a}_r$ ist eine Basis.
- $\tilde{A} = I_n$.

Bemerkung Wir sehen, dass die Matrix $\tilde{A}$ in reduzierter Zeilenstufenform eindeutig ist, wenn $\underline{a}_1, \ldots, \underline{a}_r$ eine Basis ist. Allgemeiner kann man zeigen, dass sich jede Matrix mittels elementarer Zeilenumformungen zu genau einer Matrix in reduzierter (Zeilen-)Stufenform umformen lässt. Diese Matrix heißt dann auch die *Zeilennormalform* von A.

2.6 Matrizenmultiplikation und Gauß-Algorithmus

Mittels der Matrizenmultiplikation kann man den Gauß-Algorithmus reinterpretieren:

Wir betrachten die drei Arten von elementaren Zeilenoperationen auf $m \times n$-Matrizen.

Jede der drei elementaren Zeilentransformationen entspricht einer Multiplikation mit einer bestimmten invertierbaren Matrix in $K^{m\times m}$ von links. Wir betrachten hierzu die drei elementaren Zeilentransformationen.

(I) Multiplikation der i-ten Zeile mit $c \in K^*$. Dies entspricht der Multiplikation mit der Matrix

$$(\underline{e}_1 \,|\, \cdots \,|\, \underline{e}_{i-1} |\, c\underline{e}_i \,|\, \underline{e}_{i+1} \,|\, \cdots \,|\, \underline{e}_n) \,.$$

(II) Vertauschen der i-ten und der j-ten Zeile. Sei $i < j$. Dann entspricht dies der Multiplikation mit der Matrix

$$(\underline{e}_1 \,|\, \cdots \,|\, \underline{e}_{i-1} \,|\, \underline{e}_j \,|\, \underline{e}_{i+1} \,|\, \cdots \,|\, \underline{e}_{j-1} \,|\, \underline{e}_i \,|\, \underline{e}_{j+1} \,|\, \cdots \,|\, \underline{e}_n) \,.$$

(III) Addition von c-mal Zeile i zu Zeile j (mit $i \neq j$). Dies entspricht der Multiplikation mit der Matrix

$$(\underline{e}_1 \,|\, \cdots \,|\, \underline{e}_{i-1} \,|\, \underline{e}_i + c\underline{e}_j \,|\, \underline{e}_{i+1} \,|\, \cdots \,|\, \underline{e}_n) \,.$$

Definition Die obigen Matrizen heißen *Elementarmatrizen.*

Bemerkung Gegeben eine elementare Zeilentransformation, erhält man die entsprechende Elementarmatrix, indem man die Transformation auf die Einheitsmatrix anwendet.

Bemerkung Beachten Sie die die unintuitive Rolle der Indices i und j in (III)!

Bemerkung /Frage Die Elementarmatrizen sind invertierbar, und die inversen Matrizen sind auch Elementarmatrizen. Wie lauten die inversen Matrizen?

Mittels des Gauß-Algorithmus kann man eine Matrix in eine Matrix in reduzierter (Zeilen-)stufenform transformieren. Dies kann man nun wie folgt ausdrücken:

Aussage 2.29 *Sei $A \in K^{m\times n}$. Dann gibt es Elementarmatrizen $E_1, \ldots, E_k \in K^{m\times m}$, so dass $E_k \cdots E_1 A$ eine Matrix in reduzierter (Zeilen-) stufenform ist.*

So eine Matrix $E_k \cdots E_1$ kann man auch leicht algorithmisch ausrechnen. Beachten Sie, dass

$$E_k \cdots E_1 (A \mid I_m) = (E_k \cdots E_1 A \mid E_k \cdots E_1)$$

gilt.

Wir können demnach so vorgehen: Wir gehen von der Matrix $(A|I_m)$ aus und wenden elementare Zeilentransformationen an, bis wir eine Matrix der Form $(\tilde{A}|M)$ erhalten, wobei $\tilde{A}$ in reduzierter (Zeilen-)stufenform ist. Dann ist M ein Produkt elementarer Matrizen, und es ist $MA = \tilde{A}$.

Wir können nun auch die Aussage, dass elementare Operationen die Lösungsmenge eines LGS nicht ändern, neu beweisen:

Seien $A \in K^{m\times n}$ und $\underline{b} \in K^n$. Sei $M \in K^{m\times m}$ eine Elementarmatrix oder allgemeiner eine invertierbare Matrix. Dann ist offenbar das LGS

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$$

äquivalent zum LGS

$$MA \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = M\underline{b} \, .$$

Die erweiterte Koeffizientenmatrix des ersten LGS ist $(A|\underline{b})$, und die erweiterte Koeffizientenmatrix des zweiten LGS ist $(MA|M\underline{b}) = M(A|\underline{b})$. Man sieht, dass die Multiplikation der erweiterten Koeffizientenmatrix mit M das erste LGS in das äquivalente zweite LGS überführt.

Lemma 2.30 *Sei $A \in K^{m\times n}$, und seien $M \in K^{m\times m}, N \in K^{n\times n}$ invertierbar. Dann sind die Spalten von A genau dann linear unabhängig (resp. ein Erzeugendensystem von K^m, resp. eine Basis von K^m), wenn die Spalten von MAN linear unabhängig (resp. ein Erzeugendensystem von K^m, resp. eine Basis von K^m) sind.*

Beweis. Die folgenden Aussagen sind äquivalent:

- Die Spalten von A sind linear unabhängig.
- Die Abbildung Λ_A ist injektiv.
- Die Abbildung $\Lambda_M \circ \Lambda_A \circ \Lambda_N = \Lambda_{MAN}$ ist injektiv.
- Die Spalten von MAN sind linear unabhängig.

Ebenso sind äquivalent:

- Die Spalten von A bilden ein Erzeugendensystem.
- Die Abbildung Λ_A ist surjektiv.
- Die Abbildung $\Lambda_M \circ \Lambda_A \circ \Lambda_N = \Lambda_{MAN}$ ist surjektiv.
- Die Spalten von MAN bilden ein Erzeugendensystem.

□

Hieraus folgt insbesondere:

Aussage 2.31 *Sei $M \in K^{m\times m}$ invertierbar, und sei $\tilde{A} = MA$ in (Zeilen-) Stufenform. Dann gilt:*

- *Die Spalten von A sind genau dann linear unabhängig, wenn $\tilde{A}$ nur "Stufenspalten" hat.*

- *Die Spalten von A bilden genau dann ein Erzeugendensystem, wenn $\tilde{A}$ keine Nullzeilen hat.*
- *Die Spalten von A bilden genau dann eine Basis, wenn $m = n$ und jeder Eintrag auf der Diagonalen von $\tilde{A} \neq 0$ ist. (Wenn $\tilde{A}$ in reduzierter (Zeilen-)Stufenform ist, ist dies äquivalent zu $\tilde{A} = I_m$.)*

Wenn man dies nun mit Aussage 2.29 verbindet, folgt:

Satz 2.3 *Sei $A \in K^{m\times n}$. Sei $M \in K^{m\times m}$ invertierbar, so dass $\tilde{A} := MA$ in reduzierter (Zeilen-)Stufenform ist. Dann sind äquivalent:*

a) Die Spalten von A bilden eine Basis von K^m.

b) Λ_A ist ein Isomorphismus.

c) Λ_A ist ein Automorphismus von K^n.

d) $m = n$ und $A \in K^{n\times n}$ ist invertierbar.

e) Es ist $\tilde{A} = I_m$.

f) Es gibt Elementarmatrizen $E_1, \ldots, E_k$, so dass $E_k \cdots E_1 A = I_m$.

g) Es gibt Elementarmatrizen $E_1, \ldots, E_k$ mit $A = E_k \cdots E_1$.

Beweis. Aussagen a) und b) sind offensichtlich äquivalent.

Aussage c) impliziert a). Es gelte a). Dann ist $n = m$, weil jede Basis von K^m m Elemente hat. Somit ist Λ_A ein Automorphismus.

Aussage c) ist offensichtlich äquivalent zu Aussage d).

Aussagen a) und e) sind äquivalent zueinander nach Aussage 2.31.

Wir wissen nun, dass Aussagen a), b), c), d), e) äquivalent zueinander sind.

Offensichtlich sind Aussagen f) und g) äquivalent zueinander, und ferner impliziert Aussage g) Aussage d).

Außerdem impliziert Aussage a) Aussage f). Denn nach Aussage 2.29 gibt es Elementarmatrizen $E_1, \ldots, E_k$, so dass $E_k \cdots E_1 A$ in reduzierter (Zeilen-)Stufenform ist. Dies bedeutet nach Aussage 2.31, das $E_k \cdots E_1 A = I_m$ gilt, also Aussage f). □

Gegeben eine Matrix $A \in K^{n\times n}$, kann man das Verfahren zu Aussage 2.29 benutzen, um zu entscheiden, ob A invertierbar ist und ggf. die inverse Matrix berechnen:

Man formt mittels elementarer Zeilentransformationen die Matrix $(A|I_n)$ um, bis man eine Matrix $(\tilde{A}|M)$ mit $\tilde{A}$ in reduzierter Treppenform erhält.

Wenn nun $\tilde{A} = I_n$, ist $M = A^{-1}$. Ansonsten ist A nicht invertierbar. Natürlich kann man den Algorithmus abbrechen, falls man während der Rechnung eine Nullzeile erhält. In diesem Fall ist A nicht invertierbar.

Transponieren

Gegeben eine Matrix $A \in K^{m\times n}$ definieren wir die *transponierte Matrix* A^t durch

$$A^t := ((a_{j,i}))_{i=1,\dots,n,j=1,\dots,m} \in K^{n\times m} .$$

Die transponierte Matrix zu

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & . & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

ist also

$$\begin{pmatrix} a_{1,1} & \cdots & a_{m,1} \\ \vdots & . & \vdots \\ a_{1,n} & \cdots & a_{m,n} \end{pmatrix} .$$

Beachten Sie, dass

$$(A+B)^t = A^t + B^t \quad \text{und} \quad (A^t)^t = A$$

für alle $A, B \in K^{m\times n}$.

Lemma 2.32 *Für $A \in K^{m\times n}$ und $B \in K^{n\times r}$ gilt*

$$(AB)^t = B^t A^t .$$

(Das Produkt auf der rechten Seite ist definiert, da $B^t \in K^{r\times n}$ und $A^t \in K^{n\times m}$.)

Beweis. Der Eintrag an der Stelle (i,j) von $(AB)^t$ ist gleich dem Eintrag an der Stelle (j,i) von AB, also $\sum_{\ell=1}^{n} a_{j,\ell} b_{\ell,i}$.

Andererseits ist der Eintrag an der Stelle (i,j) von $B^t A^t$ per Definition gleich $\sum_{\ell=1}^{n} b_{\ell,i} a_{j,\ell}$.

Damit sind die beiden Einträge gleich. □

Wir erhalten insbesondere:

Aussage 2.33 *Eine Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn A^t invertierbar ist. In diesem Fall ist $(A^t)^{-1} = (A^{-1})^t$.*

Beweis. Sei A invertierbar. Dann gilt $A^t(A^{-1})^t = (A^{-1}A)^t = I_n^t = I_n$ und $(A^{-1})^t A^t = (AA^{-1})^t = I_n^t = I_n$. Damit ist per Definition A^t invertierbar, und $(A^{-1})^t$ ist das Inverse von A^t.

Die Rückrichtung folgt auch, da $(A^t)^t = A$ ist. □

Spaltentransformationen

Analog zu elementaren Zeilentransformationen kann man eine Matrix auch mittels *elementarer Spaltentransformationen* umformen.

Diese Umformungen sind:

(I) Multiplikation einer Spalte mit einem Skalar $\neq 0$.
(II) Vertauschen von zwei Spalten.
(III) Addition des c-fachen der i-ten Spalte zur j-ten Spalte (für ein $c \in K$ und $i \neq j$).

Jede dieser drei Umformungen kann man erhalten, indem man die Matrix von *rechts* mit einer bestimmten invertierbaren Matrix multipliziert.

Frage Welche Matrizen sind dies genau?

Es gibt einen Zusammenhang zwischen Zeilentransformationen und Spaltentransformationen, der sich durch das Transponieren ergibt: Anstatt eine Spaltentransformation durchzuführen, kann man auch die Matrix transponieren, die "entsprechende" Zeilentransformation durchführen und dann wieder transponieren.

Beispiel 2.34 Sei $A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Wenn wir 2-mal die erste Spalte von der zweiten abziehen, erhalten wir $\begin{pmatrix} 1 & 0 & 3 \\ 4 & -3 & 6 \end{pmatrix}$. Wenn wir andererseits A transponieren, dann 2-mal die erste Zeile von der zweiten abziehen und dann wieder transponieren, erhalten wir der Reihe nach

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & 4 \\ 0 & -3 \\ 3 & 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 3 \\ 4 & -3 & 6 \end{pmatrix}.$$

Beachten Sie auch: Die Spaltentransformation entspricht der Multiplikation mit der Matrix

$$M := \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

von rechts, die Zeilentransformation entspricht der Multiplikation mit der Matrix

$$M^t = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

von links. (Man kann das Beispiel auch durch die Gleichung $AM = (M^t A^t)^t$ beschreiben; diese folgt aus Lemma 2.32.)

Analog zur Definition einer Matrix in (reduzierter) Zeilenstufenform definiert man, was eine Matrix in (reduzierter) Spaltenstufenform ist. Wir machen es uns einfach und definieren:

Definition Eine Matrix A ist in *Spaltenstufenform* (resp. in *reduzierter Spaltenstufenform*), wenn A^t in Zeilenstufenform (resp. in reduzierter Zeilenstufenform) ist.

Beachten Sie, dass die Stufen einer Matrix in Spaltenstufenform die *Breite* 1 haben.

Wir kommen nun zu einer *Anwendung*.

Seien $\underline{a}_1, \ldots, \underline{a}_r \in K^n$. Wir wollen eine Basis von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$ bestimmen. Ein Verfahren hierfür haben wir bereits in Abschnitt 2.5 kennen gelernt. Mit dem dort beschriebenen Verfahren kann man eine Basis finden, die ein Teilsystem von $\underline{a}_1, \ldots, \underline{a}_r$ ist.

Mit dem hier beschriebenen Verfahren findet man hingegen eine besonders "schöne" Basis, die man dann auch leicht zu einer Basis von K^n ergänzen kann.

Wir betrachten die Matrix $(\underline{a}_1 | \cdots | \underline{a}_r)$. Nehmen wir an, wir haben eine elementare *Spalten*transformation durchgeführt und haben nun die Matrix $(\tilde{\underline{a}}_1 | \cdots | \tilde{\underline{a}}_r)$. Dann liegen (offensichtlich) alle $\tilde{\underline{a}}_i$ in $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$. Umgekehrt gilt $\underline{a}_i \in \langle \tilde{\underline{a}}_1, \ldots, \tilde{\underline{a}}_r \rangle_K$, da die Transformation (mittels einer anderen elementaren Spaltentransformation) rückgängig gemacht werden kann. Damit gilt:

$$\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K = \langle \tilde{\underline{a}}_1, \ldots, \tilde{\underline{a}}_r \rangle_K \tag{2.15}$$

Mit anderen Worten, der von den Spalten aufgespannte lineare Unterraum ist invariant unter elementaren Spaltentransformationen.

Das Verfahren ist nun wie folgt: Wir transformieren die Matrix $(\underline{a}_1 | \cdots | \underline{a}_r)$ mittels elementarer Spaltentransformationen in eine Matrix $\tilde{A}$ in Spaltenstufenform. Die nicht-Nullspalten von $\tilde{A}$ sind dann offensichtlich linear unabhängig und ein Erzeugendensystem des von den Spalten aufgespannten Raums. Damit bilden diese eine Basis von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$.

Eine besonders "schöne Basis" erhält man, indem man bis zu einer Matrix in reduzierter Spaltenstufenform weiterrechnet.

Wenn man so eine "Stufenbasis" von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$ berechnet hat, kann man sie in offensichtlicher Weise zu einer Basis K^n ergänzen: Man nimmt Einheitsvektoren $\underline{e}_i$ hinzu, so dass man eine quadratische "Diagonalmatrix" hat.

Die Vektoren $\underline{a}_1, \ldots, \underline{a}_r$ sind genau dann linear unabhängig, wenn der ganze Prozess keine Nullspalten erzeugt. Sei dies der Fall, und seien $\underline{e}_{i_1}, \ldots, \underline{e}_{i_{n-r}}$ die hinzugenommenen Vektoren. Dann ist also $\tilde{\underline{a}}_1, \ldots, \tilde{\underline{a}}_r, \underline{e}_{i_1}, \ldots, \underline{e}_{i_{n-r}}$ eine

Basis von K^n. Nun ist auch $\underline{a}_1, \ldots, \underline{a}_r, \underline{e}_{i_1}, \ldots, \underline{e}_{i_{n-r}}$ eine Basis von K^n. Der Beweis hiervon ist leicht.

Bemerkung Aufgrund des Zusammenhangs zwischen Spalten- und Zeilentransformationen unter Transponieren kann man auch so vorgehen: Man transponiert die Vektoren $\underline{a}_1, \ldots, \underline{a}_r$ und schreibt diese als *Zeilen* in eine Matrix. Dann führt man den Gauß-Algorithmus durch (elementare Zeilenumformungen). Man betrachtet nun die nicht-Nullzeilen. Wenn man diese transponiert, erhält man eine gesuchte Basis von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$.

Bemerkung Die Invarianz (2.15) kann man mittels Matrizenmultiplikation auch so zeigen: Wie gesagt korrespondiert jede elementare Spaltentransformation zu einer Multiplikation mit einer bestimmten invertierbaren Matrix von rechts.

Sei $A \in K^{n\times r}$ die Ausgangsmatrix und M eine beliebige invertierbare $r \times r$-Matrix. Dann ist

$$\text{Bild}(\Lambda_{AM}) = \text{Bild}(\Lambda_A \circ \Lambda_M) = \text{Bild}(\Lambda_A) \ ,$$

da $\Lambda_M : K^r \longrightarrow K^r$ surjektiv (sogar bijektiv) ist. Es ist aber $\text{Bild}(\Lambda_A)$ der von den Spalten von A aufgespannte Raum und $\text{Bild}(\Lambda_{AM})$ der von den Spalten von AM aufgespannte Raum.

Der Rang

Definition Sei A eine Matrix über K. Die Dimension des von den Spalten von A erzeugten Vektorraums heißt der *Spaltenrang* von A. Die Dimension des von den Zeilen von A erzeugten Vektorraums heißt der *Zeilenrang* von A.

Bemerkung Der Zeilenrang von A ist gleich dem Spaltenrang von A^t (und ebenso, wenn man Zeilenrang und Spaltenrang vertauscht).

Ziel ist nun der Beweis des folgenden Satzes.

Satz 2.4 *Sei A eine Matrix über K. Dann ist der Spaltenrang von A gleich dem Zeilenrang von A.*

Lemma 2.35 *Sei $A \in K^{m\times n}$ eine beliebige Matrix, und seien $M \in K^{m\times m}$ und $N \in K^{n\times n}$ invertierbare Matrizen. Dann ist der Spaltenrang von A gleich dem Spaltenrang von MAN, und der Zeilenrang von A ist gleich dem Zeilenrang von MAN.*

Beweis. Wir zeigen zunächst die Aussage für den Spaltenrang. Der Spaltenrang von A (resp. MAN) ist per Definition gleich der Dimension von $\text{Bild}(\Lambda_A)$ (resp. $\text{Bild}(\Lambda_{MAN})$). Nun ist $\text{Bild}(\Lambda_{MAN}) = \text{Bild}(\Lambda_M \circ \Lambda_A \circ \Lambda_N) = \text{Bild}(\Lambda_M \circ \Lambda_A)$, da Λ_N surjektiv (sogar bijektiv) ist. Außerdem ist

$$\text{Dim}(\text{Bild}(\Lambda_M \circ \Lambda_A)) = \text{Dim}(\text{Bild}(\Lambda_A)) \ ,$$

da $(\Lambda_M)|_{\text{Bild}(\Lambda_A)} : \text{Bild}(\Lambda_A) \longrightarrow \text{Bild}(\Lambda_M \circ \Lambda_A)$ ein Isomorphismus ist (siehe Aussage 2.20).

Die Aussage über den Zeilenrang folgt, indem man zu den transponierten Matrizen übergeht und die schon bewiesenen Aussagen über den Spaltenrang anwendet. Genauer ist der Zeilenrang von A gleich dem Spaltenrang von A^t gleich dem Spaltenrang von $N^t A^t M^t = (MAN)^t$, gleich dem Zeilenrang von MAN. □

Lemma 2.36 *Sei $\tilde{A}$ in Zeilenstufenform, und sei r die Anzahl der Nicht-Nullzeilen. Dann gilt: Der Zeilenrang von $\tilde{A}$ ist gleich dem Spaltenrang von $\tilde{A}$ ist gleich r.*

Beweis. Die Nicht-Nullzeilen sind offensichtlich linear unabhängig, d.h. sie bilden eine Basis für den von ihnen aufgespannten Raum.

Andererseits sind die Stufenspalten auch offensichtlich linear unabhängig, und die anderen Spalten sind von diesen Spalten linear abhängig. Damit bilden die Stufenspalten auch eine Basis in dem von ihnen aufgespannten Raum.

Beachten Sie, dass die Anzahl der Stufenspalten gleich der Anzahl der Nicht-Nullzeilen ist. □

Aus diesen beiden Lemmata folgt der Satz. Denn: Sei $M \in K^{m \times n}$ invertierbar, so dass MA in Zeilenstufenform ist. Dann ist der Spaltenrang von A gleich dem Spaltenrang von MA gleich dem Zeilenrang von MA gleich dem Zeilenrang von A. □

Definition Da der Zeilenrang immer gleich dem Spaltenrang ist, spricht man einfach vom *Rang* von A. Bezeichnung: $\text{Rang}(A)$.

2.7 Faktorräume und Dimensionsformeln

Sei V ein K-Vektorraum, und sei $U \subseteq V$ ein Untervektorraum. Dann ist U insbesondere eine Untergruppe von V, und wir haben die Faktorgruppe V/U mit der induzierten Addition (also $[\mathfrak{v}]_U + [\mathfrak{w}]_U = [\mathfrak{v} + \mathfrak{w}]_U$ für alle $\mathfrak{v}, \mathfrak{w} \in V$).

Ich behaupte, dass die Skalarmultiplikatition von K auf V eine Skalarmultiplikation

$$\cdot : K \times V/U \longrightarrow V/U \ (a, [\mathfrak{v}]_U) \longrightarrow [a\mathfrak{v}]_U$$

auf V/U induziert. Hierzu müssen wir die "Wohldefiniertheit" überprüfen.

Seien dazu $a \in K$ und $\mathfrak{v}, \mathfrak{w} \in V$ mit $[\mathfrak{v}]_U = [\mathfrak{w}]_U$, d.h. $\mathfrak{v} \sim_U \mathfrak{w}$. Dann ist $a\mathfrak{v} - a\mathfrak{w} = a(\mathfrak{v} - \mathfrak{w}) \in U$, also $a\mathfrak{v} \sim_U a\mathfrak{w}$, d.h. $[a\mathfrak{v}]_U = [a\mathfrak{w}]_U$, was zu zeigen war.

Man erhält nun leicht:

Aussage 2.37 *V/U ist mit der soeben definierten Skalarmultiplikation ein K-Vektorraum.*

Aussage 2.38 *Sei V endlich erzeugt und $U \subseteq V$ ein Untervektorraum. Dann gilt* $\mathrm{Dim}(V) = \mathrm{Dim}(V/U) + \mathrm{Dim}(U)$.

Beweis. Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_r$ eine Basis von U. Wir ergänzen diese Basis zu einer Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_s$ von V (siehe Aussage 2.1). Ich behaupte nun, dass $[\mathfrak{b}_{r+1}]_U, \ldots, [\mathfrak{b}_s]_U$ eine Basis von V/U ist. (Dann ist $s = \mathrm{Dim}(V)$, $r = \mathrm{Dim}(U)$ und $s - r = \mathrm{Dim}(V/U)$, und hieraus folgt die Behauptung.)

Offensichtlich bilden $[\mathfrak{b}_1]_U, \ldots, [\mathfrak{b}_s]_U$ ein Erzeugendensystem von V/U. Da aber $[\mathfrak{b}_1]_U = \cdots = [\mathfrak{b}_r]_U = [\mathfrak{o}]_U = \mathfrak{o}_{V/U}$ ist, bilden auch $[\mathfrak{b}_{r+1}]_U, \ldots, [\mathfrak{b}_s]_U$ ein Erzeugendensystem.

Wir zeigen nun die lineare Unabhängigkeit. Seien dazu $a_{r+1}, \ldots, a_s \in K$ mit $a_{r+1}[\mathfrak{b}_{r+1}]_U + \cdots + a_s[\mathfrak{b}_s]_U = \mathfrak{o}$. Dann ist also $a_{r+1}\mathfrak{b}_{r+1} + \cdots + a_s\mathfrak{b}_s \in U$, d.h. es gibt $a_1, \ldots, a_r \in K$ mit $a_{r+1}\mathfrak{b}_{r+1} + \cdots + a_s\mathfrak{b}_s = a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r$, d.h. $-a_1\mathfrak{b}_1 - \cdots - a_r\mathfrak{b}_r + a_{r+1}\mathfrak{b}_{r+1} + \cdots + a_s\mathfrak{b}_s = \mathfrak{o}$. Da $\mathfrak{b}_1, \ldots, \mathfrak{b}_s$ eine Basis von V bilden, folgt $a_1 = \cdots = a_s = 0$. □

Sei nun auch W ein K-Vektorraum und $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann induziert φ einen Isomorphismus von Vektorräumen

$$\overline{\varphi} : V/\operatorname{Kern}(\varphi) \tilde{\longrightarrow} \operatorname{Bild}(\varphi) \ , [\mathfrak{x}]_{\operatorname{Kern}(\varphi)} \mapsto \varphi(\mathfrak{x}) \ . \tag{2.16}$$

Um zu überprüfen, dass die Abbildung wohldefiniert und injektiv ist, geben wir uns zwei beliebige Vektoren $\mathfrak{v}, \mathfrak{w} \in V$ vor. Dann gilt:

$$\mathfrak{v} \sim_{\operatorname{Kern}(\varphi)} \mathfrak{w} \longleftrightarrow \mathfrak{v} - \mathfrak{w} \in \operatorname{Kern}(\varphi) \longleftrightarrow \varphi(\mathfrak{v} - \mathfrak{w}) = \mathfrak{o} \longleftrightarrow \varphi(\mathfrak{v}) = \varphi(\mathfrak{w}) \ ,$$

was zu zeigen war.

Es ist leicht zu sehen, dass φ eine lineare Abbildung ist, also ist es ein Isomorphismus.

Satz 2.5 *Sei V endlich erzeugt, und sei $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann gilt* $\mathrm{Dim}(V) = \mathrm{Dim}(\mathrm{Kern}(\varphi)) + \mathrm{Dim}(\mathrm{Bild}(\varphi))$.

Dies folgt aus Aussage 2.38 und dem Isomorphismus (2.16).

Korollar 2.39 *Sei $A \in K^{m\times n}$. Dann gilt* $\mathrm{Rang}(A) + \mathrm{Dim}(\mathrm{Kern}(A)) = n$.

Definition Seien $U_1, U_2 \subseteq V$ Untervektorräume. Dann definieren wir die *Summe* von U_1 und U_2 als

$$U_1 + U_2 := \{\mathfrak{x}_1 + \mathfrak{x}_2 \mid \mathfrak{x}_1 \in U_1\,, \mathfrak{x}_2 \in U_2\}\ .$$

Dies ist der kleinste Untervektorraum von V, der die Mengen U_1 und U_2 umfasst. (überprüfen!).

Bemerkung Wenn $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ ein Erzeugendensystem von U_1 bilden und $\mathfrak{y}_1, \ldots, \mathfrak{y}_s$ ein Erzeugendensystem von U_2 bilden, dann bilden $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$, $\mathfrak{y}_1, \ldots, \mathfrak{y}_s$ ein Erzeugendensystem von $U_1 + U_2$.

Satz 2.6 *Sei V endlich erzeugt, und seien $U_1, U_2 \subseteq V$ Untervektorräume. Dann gilt* $\mathrm{Dim}(U_1 + U_2) + \mathrm{Dim}(U_1 \cap U_2) = \mathrm{Dim}(U_1) + \mathrm{Dim}(U_2)$.

Beweis 1. Wir betrachten die lineare Abbildung

$$U_1 \hookrightarrow U_1 + U_2 \longrightarrow (U_1 + U_2)/U_2\ ,\ \mathfrak{u} \mapsto [\mathfrak{u}]_{U_2}\ .$$

Man sieht leicht, dass diese Abbildung surjektiv ist und den Kern $U_1 \cap U_2$ hat. Somit haben wir nach (2.16) einen Isomorphismus

$$U_1/(U_1 \cap U_2) \longrightarrow (U_1 + U_2)/U_2\ .$$

Wenn wir hierauf Aussage 2.38 anwenden, erhalten wir:

$$\mathrm{Dim}(U_1) - \mathrm{Dim}(U_1 \cap U_2) = \mathrm{Dim}(U_1 + U_2) - \mathrm{Dim}(U_2)$$

□

Beweis 2. Sei $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ eine Basis von $U_1 \cap U_2$. Wir ergänzen diese Basis zu einer Basis $\mathfrak{x}_1, \ldots, \mathfrak{x}_s$ von U_1 sowie zu einer Basis $\mathfrak{x}_1, \ldots, \mathfrak{x}_r, \mathfrak{y}_1, \ldots, \mathfrak{y}_t$ von U_2.

Ich behaupte, dass dann $\mathfrak{x}_1, \ldots, \mathfrak{x}_s, \mathfrak{y}_1, \ldots, \mathfrak{y}_t$ eine Basis von $U_1 + U_2$ ist. (Dann ist $s+t = \mathrm{Dim}(U_1+U_2)$, $s = \mathrm{Dim}(U_1)$ und $t = \mathrm{Dim}(U_2) - \mathrm{Dim}(U_1 \cap U_2)$, und hieraus folgt die Behauptung.)

Es ist offensichtlich, dass es sich um ein Erzeugendensystem handelt.

Seien also $a_1, \ldots, a_s, b_1, \ldots, b_t \in K$ mit $a_1\mathfrak{x}_1 + \cdots + a_s\mathfrak{x}_s + b_1\mathfrak{y}_1 + \cdots + b_t\mathfrak{y}_t = \mathfrak{o}$. Dann ist also $a_1\mathfrak{x}_1 + \cdots + a_s\mathfrak{x}_s = -(b_1\mathfrak{y}_1 + \cdots + b_t\mathfrak{y}_t) \in U_1 \cap U_2$. Somit gibt es $c_1, \ldots, c_r \in K$ mit $c_1\mathfrak{x}_1 + \cdots + c_r\mathfrak{x}_r = a_1\mathfrak{x}_1 + \cdots + a_s\mathfrak{x}_s$ bzw. $(a_1 - c_1)\mathfrak{x}_1 + \cdots + (a_r - c_r)\mathfrak{x}_r + a_{r+1}\mathfrak{x}_{r+1} + \cdots + a_s\mathfrak{x}_s = \mathfrak{o}$. Hieraus folgt aufgrund der linearen Unabhängigkeit von $\mathfrak{x}_1, \ldots, \mathfrak{x}_s$: $c_1 = a_1, \ldots, c_r = a_r$ und $a_{r+1} = \cdots = a_s = 0$. Wir haben also $a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r + b_1\mathfrak{y}_1 + \cdots + b_t\mathfrak{y}_t = \mathfrak{o}$. Hieraus folgt nun $a_1 = \cdots = a_r = b_1 = \cdots = b_t = 0$ aufgrund der linearen Unabhängigkeit von $\mathfrak{x}_1, \ldots, \mathfrak{x}_r, \mathfrak{y}_1, \ldots, \mathfrak{y}_t$. □

Bemerkung Wir sagen, dass die *Summe von U_1 und U_2 direkt* ist, falls $U_1 \cap U_2 = \{0\}$. Man sieht leicht, dass dies äquivalent hierzu ist, dass es für jedes $\mathfrak{v} \in U_1 + U_2$ *eindeutig bestimmte* Vektoren $\mathfrak{u}_1, \mathfrak{u}_2$ mit $\mathfrak{v} = \mathfrak{u}_1 + \mathfrak{u}_2$ gibt. Wenn dies der Fall ist, bezeichnet man die Summe von U_1 und U_2 auch mit $U_1 \oplus U_2$.

Sei nun wie oben V endlich erzeugt. Dann folgt aus Satz 2.6: Die Summe von U_1 und U_2 ist genau dann direkt, wenn $\mathrm{Dim}(U_1 + U_2) = \mathrm{Dim}(U_1) + \mathrm{Dim}(U_2)$ ist.

2.8 Abbildungsmatrizen und Basiswechsel

Sei V ein endlich erzeugter K-Vektorraum.

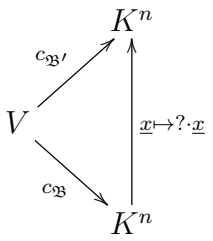
Sei nun $\mathfrak{B} := (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$ eine Basis von V. Wir haben dann die Koordinatenabbildung $c : V \longrightarrow K^n$, die eindeutig durch $\mathfrak{b}_i \mapsto \underline{e}_i$ gegeben ist. Da wir die Basis variieren werden, schreiben wir $c_{\mathfrak{B}}$ für diese Koordinatenabbildung.

Wir wählen nun eine zweite Basis $\mathfrak{B}' = (\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n)$ von V. Dann können wir $\mathfrak{b}'_j = \sum_{i=1}^n s_{i,j}\mathfrak{b}_i$ mit eindeutig bestimmten $s_{i,j} \in K$ schreiben.

Beachten Sie, dass der Vektor $\begin{pmatrix} s_{1,j} \\ \vdots \\ s_{n,j} \end{pmatrix}$ (die j-te Spalte von S) genau der Koordinatenvektor von $\mathfrak{b}'_j$ bzgl. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ ist. Dementsprechend nennen wir die Matrix $S = ((s_{i,j}))_{i,j}$ *Koordinatenmatrix von $\mathfrak{B}'$ bezüglich $\mathfrak{B}$*.

Wir wollen nun die Koordinatenvektoren von Vektoren von V bzgl. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ in Koordinatenvektoren bzgl. $\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n$ umrechnen. D.h. gegeben $\underline{x} \in K^n$ wollen wir $c_{\mathfrak{B}'} \circ c_{\mathfrak{B}}^{-1}(\underline{x})$ berechnen. Beachten Sie, dass $c_{\mathfrak{B}'} \circ c_{\mathfrak{B}}^{-1} : K^n \longrightarrow K^n$

eine lineare Abbildung ist.



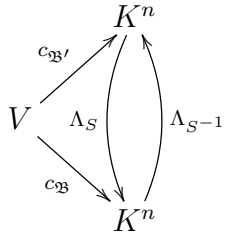
Wie muss die Matrix ? lauten, damit das Diagramm kommutativ ist?

Bevor wir hierzu kommen, betrachten wir ein einfacheres Problem: Wir rechnen die Koordinatenvektoren von $\mathfrak{x} \in V$ bzgl. $\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n$ in Koordinatenvektoren bzgl. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ um.

Dies bedeutet: Wir bestimmen die Matrix zur linearen Abbildung $c_{\mathfrak{B}} \circ c_{\mathfrak{B}'}^{-1}$: $K^n \longrightarrow K^n$. Wir wissen, dass $c_{\mathfrak{B}} \circ c_{\mathfrak{B}'}^{-1}(\underline{e}_j) = c_{\mathfrak{B}}(\mathfrak{b}'_j) = \begin{pmatrix} s_{1,j} \\ \vdots \\ s_{m,j} \end{pmatrix}$ ist. Und das bedeutet: Die gesuchte Matrix ist S.

Mit anderen Worten: Wenn $\underline{x}$ der Koordinatenvektor von $\mathfrak{x}$ bzgl. $\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n$ ist, dann ist $S\underline{x}$ der Koordinatenvektor von $\mathfrak{x}$ bzgl. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$.

Nun können wir auch die ursprüngliche Frage beantworten: Die Abbildung $c_{\mathfrak{B}'} \circ c_{\mathfrak{B}}^{-1} : K^n \longrightarrow K^n$ ist durch $\underline{x} \mapsto S^{-1}\underline{x}$ gegeben. Also: Wenn $\underline{x}$ der Koordinatenvektor von $\mathfrak{x}$ bzgl. $\mathfrak{B}$ ist, dann ist $S^{-1}\underline{x}$ der Koordinatenvektor von $\mathfrak{x}$ bzgl. $\mathfrak{B}'$.



Definition Die Matrix S^{-1} heißt *Transformationsmatrix des Basiswechsels von* $\mathfrak{B}$ *nach* $\mathfrak{B}'$. Sie wird mit $T_{\mathfrak{B}'}^{\mathfrak{B}}$ bezeichnet.

Wir haben also die Formeln

$$T_{\mathfrak{B}'}^{\mathfrak{B}} = S^{-1} \quad , \quad T_{\mathfrak{B}}^{\mathfrak{B}'} = S \; .$$

Bemerkung Die Verwendung des Begriffs "Transformationsmatrix des Basiswechsels $\mathfrak{B}$ nach $\mathfrak{B}'$" und die Notation $T_{\mathfrak{B}'}^{\mathfrak{B}}$ folgt dem Buch "Lineare Algebra" von G. Fischer und ist auch ansonsten üblich. Auch die weiteren Notationen folgen dem Buch von Fischer.

Den Begriff "Koordinatenmatrix von $\mathfrak{B}'$ bezüglich $\mathfrak{B}$" habe ich mir ausgedacht. Er sollte aber allgemein verständlich sein, schließlich ist S ja gerade die Matrix der Koordinatenvektoren der Elemente von $\mathfrak{B}'$ bezüglich $\mathfrak{B}$.

Manche Autoren nennen S die "Matrix des Basiswechsels von $\mathfrak{B}$ nach $\mathfrak{B}'$" (z.B. S. Bosch in seinem Buch "Lineare Algebra" (S.116)). Das ist etwas ungünstig, da man dies leicht mit "Transformationsmatrix des Basiswechsels" verwechseln kann.

Der Begriff "Transformationsmatrix des Basiswechsels von $\mathfrak{B}$ nach $\mathfrak{B}'$" und die Bezeichnung $T^{\mathfrak{B}}_{\mathfrak{B}'}$ sollte aber wirklich nur für S^{-1} (und nicht für S) verwandt werden.

Sei nun W ein zweiter endlich erzeugter K-Vektorraum, sei $\mathfrak{C} := (\mathfrak{c}_1, \ldots, \mathfrak{c}_m)$ eine Basis von W, und sei $\varphi : V \longrightarrow W$ eine lineare Abbildung.

Nun gibt es eine eindeutig bestimmte Matrix $M \in K^{m\times n}$, so dass das folgende Diagramm kommutiert.

$$\begin{array}{ccc} V & \xrightarrow{\quad\varphi\quad} & W \\ {\scriptstyle c_{\mathfrak{B}}}\downarrow & & \downarrow{\scriptstyle c_{\mathfrak{C}}} \\ K^n & \xrightarrow[\underline{x}\mapsto M\cdot\underline{x}]{} & K^m \end{array}$$

Dies bedeutet, dass für alle $\mathfrak{x} \in V$ $M \cdot c_{\mathfrak{B}}(\mathfrak{x}) = c_{\mathfrak{C}}(\varphi(\mathfrak{x}))$ gilt.

Wie lautet M?

Die j-te Spalte von M ist $M \cdot \underline{e}_j = c_{\mathfrak{C}}(\varphi(c^{-1}_{\mathfrak{B}}(\underline{e}_j)) = c_{\mathfrak{C}}(\varphi(\mathfrak{b}_j))$, dies ist der Koordinatenvektor von $\varphi(\mathfrak{b}_j)$ bezüglich $\mathfrak{C}$.

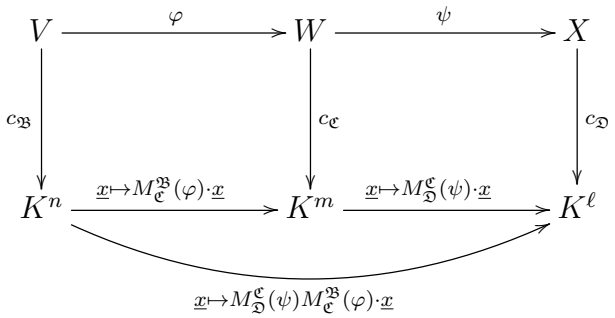
Definition Die soeben definierte $m \times n$-Matrix M heißt *Abbildungsmatrix* von φ bezüglich der Basen $\mathfrak{B}$ und $\mathfrak{C}$. Sie wird mit $M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)$ bezeichnet.

Beachten Sie: Die Abbildungsmatrix von id_V bezüglich $\mathfrak{B}'$ und $\mathfrak{B}$ ist gleich der Koordinatenmatrix von $\mathfrak{B}'$ bezüglich $\mathfrak{B}$. Außerdem ist die Abbildungsmatrix von id_V bezüglich $\mathfrak{B}$ und $\mathfrak{B}'$ gleich der Transformationsmatrix des Basiswechsels von $\mathfrak{B}$ nach $\mathfrak{B}'$. Mit den obigen Notationen haben wir:

$$M^{\mathfrak{B}'}_{\mathfrak{B}}(\mathrm{id}_V) = S \qquad M^{\mathfrak{B}}_{\mathfrak{B}'}(\mathrm{id}_V) = S^{-1} = T^{\mathfrak{B}}_{\mathfrak{B}'} \tag{2.17}$$

Sei nun X noch ein endlich erzeugter K-Vektorraum mit Basis $\mathfrak{D} := (\mathfrak{d}_1, \ldots, \mathfrak{d}_\ell)$, und sei $\psi : W \longrightarrow X$ eine lineare Abbildung. Dann sind das rechte und das linke "Kästchen" sowie der untere Teil des folgenden Dia-

gramms kommutativ:



Dies bedeutet, dass das gesamte Diagramm kommutativ ist. Insbesondere sieht man, dass die Abbildungsmatrix von $\psi \circ \varphi : V \longrightarrow X$ bzgl. $\mathfrak{B}$ und $\mathfrak{C}$ gleich $M^{\mathfrak{C}}_{\mathfrak{D}}(\psi)M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)$ ist. Also

$$M^{\mathfrak{B}}_{\mathfrak{D}}(\psi \circ \varphi) = M^{\mathfrak{C}}_{\mathfrak{D}}(\psi)\, M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)\ . \tag{2.18}$$

Eselsbrücke: "Man kann $\mathfrak{C}$ kürzen".

Wir betrachten nun, wie sich Abbildungsmatrizen unter Basiswechsel transformieren.

Seien dazu $\mathfrak{B}, \mathfrak{B}'$ Basen von V und $\mathfrak{C}, \mathfrak{C}'$ Basen von W. Sei S die Koordinatenmatrix von $\mathfrak{B}'$ bezüglich $\mathfrak{B}$ und T die Koordinatenmatrix von $\mathfrak{C}'$ bezüglich $\mathfrak{C}$. Dann ist nach (2.18) und (2.17)

$$M^{\mathfrak{B}'}_{\mathfrak{C}'}(\varphi) = M^{\mathfrak{C}}_{\mathfrak{C}'}(\mathrm{id}_W)M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)M^{\mathfrak{B}'}_{\mathfrak{B}}(\mathrm{id}_V) = T^{\mathfrak{C}}_{\mathfrak{C}'}M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)T^{\mathfrak{B}'}_{\mathfrak{B}} = T^{-1}M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)S\ . \tag{2.19}$$

Wir betrachten noch zwei Spezialfälle:

Sei $\varphi : V \longrightarrow V$ ein Endomorphismus. Die Abbildungsmatrix $M^{\mathfrak{B}}_{\mathfrak{B}}(\varphi)$ heißt dann die *Abbildungsmatrix von φ bezüglich* der Basis $\mathfrak{B}$ und wird auch mit $M_{\mathfrak{B}}(\varphi)$ bezeichnet. Wenn nun $\mathfrak{B}'$ eine weitere Basis ist und S (wie oben) die Koodinatenmatrix von $\mathfrak{B}'$ bezüglich $\mathfrak{B}$ ist, dann ist

$$M_{\mathfrak{B}'}(\varphi) = S^{-1}\, M_{\mathfrak{B}}(\varphi)\, S\ . \tag{2.20}$$

Wir geben uns eine Matrix $A \in K^{m\times n}$ vor und betrachten die Abbildung $\Lambda_A : K^n \longrightarrow K^m$. Die Abbildungsmatrix dieser Abbildung bzgl. den Standardbasen von K^n und K^m ist natürlich A.

Seien nun $\mathfrak{B} := (\underline{b}_1, \ldots, \underline{b}_n)$ und $\mathfrak{C} := (\underline{c}_1, \ldots, \underline{c}_m)$ Basen von K^n bzw. K^m, und seien B und C die Matrizen, die man erhält, wenn man die Basisvektoren jeweils als Spalten einer Matrix auffasst. Die Koordinatenmatrix von $\mathfrak{B}$ bezüglich der Standardbasis ist demnach B, und die Koordinatenmatrix von $\mathfrak{C}$ bezüglich der Standardbasis ist C.

Demnach ist die Abbildungsmatrix von Λ_A bzgl. $\mathfrak{B}$ und $\mathfrak{C}$ gleich $C^{-1}AB$. Mit anderen Worten: Wenn M diese Abbildungsmatrix ist, ist $CM = AB$. Überlegen Sie sich anhand der Definition der Abbildungsmatrix, warum dies richtig ist!

2.9 Der Dualraum

Linearformen

Sei V ein K-Vektorraum. Beachten Sie, dass K in offensichtlicher Weise auch ein K-Vektorraum ist. (Wenn man K als K-Vektorraum betrachtet, schreibt man auch K^1.)

Definition Eine *Linearform* auf V ist eine lineare Abbildung von V nach K.

Beispiel 2.40 Seien $i, n \in \mathbb{N}$ mit $i \leq n$. Dann ist die "Projektion auf die i-te Komponente" $p_i : K^n \longrightarrow K\ ,\ \underline{x} \mapsto x_i$ eine Linearform auf K^n.

Für zwei K-Vektorräume V und W ist auch die Menge $\mathrm{Hom}_K(V, W)$ der linearen Abbildungen von V nach W "in offensichtlicher Weise" ein K-Vektorraum. Somit ist auch $\mathrm{Hom}_K(V, K)$, die Menge der Linearformen auf V, "in offensichtlicher Weise" ein K-Vektorraum.

Definition Wir setzen $V^* := \mathrm{Hom}_K(V, K)$ und nennen diesen Raum den *Dualraum* von V.

Beispiel 2.41 Wir studieren den Dualraum von K^n. Eine lineare Abbildung von K^n nach K^m wird durch eine $m \times n$-Matrix gegeben. Somit wird also eine Linearform auf K^n durch eine $1 \times n$-Matrix gegeben. Solche Matrizen nennt man auch *Zeilenvektoren*. Nochmal etwas formaler: Wir haben einen Isomorphismus von K-Vektorräumen

$$K^{1\times n} \longrightarrow (K^n)^*\ ,\ (a_1\ a_2\ \cdots\ a_n) \mapsto (\underline{x} \mapsto \sum_{i=1}^{n} a_i x_i) = \sum_{i=1}^{n} a_i p_i\ .$$

Unter dem obigen Isomorphismus entspricht dann der Zeilenvektor $e_i^t = (\delta_{i,j})_{j=1,\ldots,n}$ der Linearform p_i.

Wir sehen, dass die Linearformen $p_1, \ldots, p_n$ eine Basis von $(K^n)^*$ bilden.

Wir setzen nun $X_i := p_i$ für alle $i = 1, \ldots, n$. Dann kann man also jede Linearform auf K^n in der From $a_1X_1 + \cdots a_nX_n$ mit eindeutig bestimmten

Koeffizienten $a_1, \ldots, a_n$ schreiben. Es gibt hier einen Zusammenhang zu homogenen Linearen Gleichungssystemen. Man kann definieren: Eine *homogene lineare Gleichung in n Unbestimmten über K* ist *per Definition* eine Linearform auf K^n. Wenn nun die Linearform $a_1X_1 + \cdots a_nX_n$ gegeben ist, sagt man auch "die Gleichung $a_1X_1 + \cdots a_nX_n = 0$" ist gegeben. Aber das "$= 0$" hat keine formale Bedeutung – es ist eher eine Aufforderung. (Man kann es auch weglassen; es ist vielleicht sogar besser, es wegzulassen, weil es so aussieht als wäre die Linearform gleich 0.)

Bemerkung Es ist entscheidend, dass man hier die Symbolik beachtet: Der Raum K^n ist per Definition der Raum der *Spaltenvektoren* der Länge n über K. Der Dualraum $(K^n)^*$ ist kanonisch isomorph zum Raum der *Zeilenvektoren* über K. Natürlich ist der Raum der Spaltenvektoren der Länge n auch isomorph zum Raum der Zeilenvektoren der Länge n (via transponieren). Man sollte jedoch strikt zwischen Spalten- und Zeilenvektoren unterscheiden.

Und noch etwas: Sowohl Spalten- als auch Zeilenvektoren sind Spezialfälle von Matrizen. Deshalb schreibe ich zwischen den Einträgen von Zeilenvektoren auch keine Kommata.

Wir betrachten nun Linearformen auf beliebigen endlich-dimensionalen Vektorräumen.

Sei also V ein endlich-dimensionaler Vektorraum mit Basis $\mathfrak{B} := (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$. Sei zunächst W ein weiterer K-Vektorraum. Wir wissen schon: Zu vorgegebenen Vektoren $\mathfrak{x}_1, \ldots, \mathfrak{x}_n \in W$ gibt es genau eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\varphi(\mathfrak{b}_i) = \mathfrak{x}_i$ für $i = 1, \ldots, n$.

Dies wenden wir nun auf $W = K$ an. Wir erhalten: Zu $a_1, \ldots, a_n \in K$ gibt es genau eine Linearform α auf V mit $\alpha(\mathfrak{b}_i) = a_i$ für alle $i = 1, \ldots, n$.

Dies können wir nun mittels Abbildungsmatrizen umformulieren: Die Abbildungsmatrix von α bezüglich der Basen $\mathfrak{B}$ einerseits und $1 \in K$ anderseits ist der Zeilenvektor $(a_1\ a_2\ \cdots\ a_n) \in K^{1\times n}$.

Definition Sei α eine Linearform auf V. Dann nennen wir die Abbildungsmatrix von α bezüglich $\mathfrak{B}$ einerseits und 1 andererseits auch die Abbildungsmatrix von α bezüglich $\mathfrak{B}$. Wir bezeichnen diese Matrix (diesen Zeilenvektor) mit $m_{\mathfrak{B}}(\alpha)$.

Wir erhalten:

Lemma 2.42 *Wir haben einen Isomorphismus von K-Vektorräumen*

$$K^{1\times n} \longrightarrow V^*\ ,\ (a_1\ a_2\ \cdots\ a_n) \mapsto \alpha$$

mit $\alpha(\mathfrak{b}_i) = a_i$ *für alle* $i = 1, \ldots, n$. *Die Umkehrabbildung ist*

$$V^* \longrightarrow K^{1\times n}\ ,\ \alpha \mapsto m_{\mathfrak{B}}(\alpha) = (\alpha(\mathfrak{b}_1)\ \ \alpha(\mathfrak{b}_2)\ \cdots\ \alpha(\mathfrak{b}_n))\ .$$

Besonders wichtig sind (wie auf K^n) die Linearformen, die den Zeilenvektoren $(\delta_{i,j})_{j=1,\ldots,n}$ entsprechen. Diese bezeichnen wir mit $\mathfrak{b}_i^*$. Es gilt also

$$\mathfrak{b}_i^*(\mathfrak{b}_j) = \delta_{i,j}$$

für $i, j = 1, \ldots, n$. Wir erhalten:

Aussage 2.43 *Die Linearformen* $\mathfrak{b}_1^*, \ldots, \mathfrak{b}_n^*$ *bilden eine Basis von* V^*.

Definition Die Basis $\mathfrak{B}^* := (\mathfrak{b}_1^*, \ldots, \mathfrak{b}_n^*)$ von V^* heißt *Dualbasis* zu $\mathfrak{B} := (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$.

Wir kommen nochmal auf den Isomorphismus von $K^{1\times n}$ nach V^* zurück: Sei $\alpha \in V^*$ mit $\alpha(\mathfrak{b}_i) = a_i$ für alle i. Dann ist Abbildungsmatrix von α bezüglich $\mathfrak{B}$ gleich dem Zeilenvektor $(a_1\ \cdots\ a_n)$. Als Element von V^* lässt sich α aber auch "in der Basis $\mathfrak{b}_1^*, \ldots, \mathfrak{b}_n^*$ entwickeln". Wir haben

$$\alpha = \sum_{i=1}^{n} \alpha(\mathfrak{b}_i)\, \mathfrak{b}_i^* = \sum_{i=1}^{n} a_i \mathfrak{b}_i^*$$

(denn beide linearen Abbildungen stimmen auf den Basisvektoren $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V überein). Somit ist der *Koordinatenvektor* von α bezüglich der Basis $\mathfrak{b}_1^*, \ldots, \mathfrak{b}_n^*$ gleich dem Spaltenvektor $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, was ja gerade das Transponierte der Abbildungsmatrix von α bezüglich $\mathfrak{B}$ ist. In Formeln:

$$c_{\mathfrak{B}^*}(\alpha) = m_{\mathfrak{B}}(\alpha)^t$$

Wir sehen hier, dass transponieren natürlicherweise auftritt, wenn man mit Linearformen rechnet. Ob man Linearfaktoren durch Zeilen- oder durch Spaltenvektoren angibt, ist eine Frage des Standpunkts: Wenn man von V und einer Basis in V ausgeht, sind Zeilenvektoren natürlich. Man kann aber auch "vergessen", dass es auch V gibt und "nur" V^* mit einer festen Basis betrachten. Dann sind Spaltenvektoren natürlich. (Meiner Ansicht nach sollte der letztere Standpunkt aber eher vermieden werden.)

Es gibt einen engen Zusammenhang zwischen der Dualbasis $\mathfrak{B}^*$ und der durch $\mathfrak{B}$ definierten Koordinatenabbildung $c_{\mathfrak{B}} : V \longrightarrow K^n$:

Es ist $c_{\mathfrak{B}}(\mathfrak{b}_j) = \underline{e}_j$ für alle $j = 1, \ldots, n$. Hieraus folgt

$$(p_i \circ c_{\mathfrak{B}})(\mathfrak{b}_j) = p_i(c_{\mathfrak{B}}(\mathfrak{b}_j)) = \delta_{i,j}$$

für alle $i, j = 1, \ldots, n$. Und dies bedeutet:

$$\mathfrak{b}_i^* = p_i \circ c_{\mathfrak{B}} \ .$$

Oder mit anderen Worten: Für $\mathfrak{v} \in V$ gilt

$$c_{\mathfrak{B}}(\mathfrak{v}) = \begin{pmatrix} \mathfrak{b}_1^*(\mathfrak{v}) \\ \vdots \\ \mathfrak{b}_n^*(\mathfrak{v}) \end{pmatrix} \ .$$

Das kann man auch so ausdrücken:

$$\mathfrak{v} = \sum_{i=1}^{n} \mathfrak{b}_i^*(\mathfrak{v})\, \mathfrak{b}_i$$

Diese Überlegungen kann man auch "umdrehen", aber hierfür benötigen wir noch etwas Theorie.

Der Bidualraum

Sei nun V ein beliebiger K-Vektorraum. Den Dualraum von V^* nennt man *Bidualraum* von V. Man setzt $V^{**} := (V^*)^*$.

Sei nun $\mathfrak{v} \in V$. Dann haben wir die Abbildung

$$V^* \longrightarrow K \ , \ \alpha \mapsto \alpha(\mathfrak{v}) \ .$$

Eine einfache Rechnung zeigt, dass diese Abbildung eine Linearform auf V^*, also ein Element von V^{**} ist.

Die von $\mathfrak{v}$ definierte Linearform auf V^* bezeichnen wir mit $\Phi(\mathfrak{v})$. Wir haben nun also die Abbildung

$$V \longrightarrow V^{**} \ , \ \mathfrak{v} \mapsto \Phi(\mathfrak{v}) \ .$$

Man sieht leicht, dass diese Abbildung linear ist.

Lemma 2.44 *Sei V endlich-dimensional. Dann ist $\Phi : V \longrightarrow V^{**}$ ein Isomorphismus von K-Vektorräumen.*

Beweis. Wir wählen nun eine Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V. Nun ist $\Phi(\mathfrak{b}_1), \ldots, \Phi(\mathfrak{b}_n)$ die Dualbasis zu $\mathfrak{b}_1^*, \ldots, \mathfrak{b}_n^*$ (denn $\Phi(\mathfrak{b}_j)(\mathfrak{b}_i^*) = \mathfrak{b}_i^*(\mathfrak{b}_j) = \delta_{i,j}$).

Dies zeigt, dass wir einen Isomorphismus haben. □

Definition Der Isomorphismus $\Phi : V \longrightarrow V^{**}$ heißt die *kanonische Abbildung* auf den Dualraum.

Bemerkung Wir haben auch einen Isomorphismus von V nach V^*, nämlich $c_{\mathfrak{B}^*}^{-1} \circ c_{\mathfrak{B}} : V \longrightarrow V^*$ – dieser Isomorphismus bildet $\mathfrak{b}_i$ auf $\mathfrak{b}_i^*$ ab. Aber dieser Isomorphismus hängt von der Basis $\mathfrak{B}$ ab und ist somit nicht "kanonisch".

Es ist übrigens entscheidend, dass V endlich-dimensional ist: Ohne diese Voraussetzung ist Φ immer injektiv und genau dann ein Isomorphismus, wenn V endlich-dimensional ist. Dies werden wir später beweisen.

Koordinatensysteme

Sei nun V wieder endlich-dimensional. Aus Lemma 2.44 folgt:

Lemma 2.45 *Sei $\beta_1, \ldots, \beta_n$ eine Basis von V^*. Dann gibt es genau ein System von Vektoren $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V mit $\beta_i(\mathfrak{b}_j) = \delta_{i,j}$ für alle $i, j = 1, \ldots, n$. Dies ist eine Basis von V.*

Beweis. Die Bedingung ist äquivalent zu $\Phi(\mathfrak{b}_j)(\beta_i) = \delta_{i,j}$ für alle i, j. Dies bedeutet, dass $\Phi(\mathfrak{b}_1), \ldots, \Phi(\mathfrak{b}_n)$ die Dualbasis zu $\beta_1, \ldots, \beta_n$ ist. Die Aussage folgt nun, weil Φ ein Isomorphismus ist. □

Definition Eine Basis von V^* heißt *Koordinatensystem* von V. Ein Isomorphismus von V nach K^n heißt *Koordinatenabbildung* auf V.

Beispiel 2.46 $X_1, \ldots, X_n$ ist ein Koordinatensystem auf K^n.

Bemerkung Die Linearformen auf V nennt man auch *Koordinatenfunktionen*. Ein Koordinatensystem ist also eine Basis von Koordinatenfunktionen. Wenn α eine Koordinatenfunktion und $\mathfrak{x}$ ein Vektor ist, nennt man $\alpha(\mathfrak{x})$ auch die Koordinate von $\mathfrak{x}$ unter α. In der Literatur wird allerdings nicht immer streng zwischen Koordinatenfunktionen und Koordinaten unterschieden. Diese Ungenauigkeit folgt aber einem allgemeinen Prinzip: Man unterscheidet nicht immer zwischen einer Funktion und ihrem Wert an einer Stelle. Z.B. schreibt man ja auch $\sin(x)$ anstatt $\mathbb{R} \longrightarrow \mathbb{R}\ ,\ x \mapsto \sin(x)$.

Die Begriffe werden auch nicht ganz einheitlich benutzt. Z.B. wird auch ein Isomorphismus $V \longrightarrow K^n$ oder auch ein Isomorphismus $K^n \longrightarrow V$ oder eine Basis von V als Koordinatensystem von V bezeichnet. Wir wollen das aber nicht machen.

Wir können nun von einem der folgenden drei Objekte in offensichtlicher Weise jeweils zum anderen geraten:

- eine Basis von V
- eine Koordinatenabbildung auf V
- ein Koordinatensystem von V

Wenn zum Beispiel eine Basis $\mathfrak{B}$ gegeben ist, haben wir die entsprechende Koordinatenabbildung $c_{\mathfrak{B}}$ und das Koordinatensystem $\mathfrak{B}^*$.

Überlegen Sie sich, wie man von einer Koordinatenabbildung oder einem Koordinatensystem ausgehen kann!

Bemerkung Es ist üblich, anstatt einer Basis ein Koordinatensystem zu fixieren. So ein Koordinatensystem bezeichnet man dann oft mit $X_1, \ldots, X_n$ oder so ähnlich. Hier sollte man allerdings etwas aufpassen: Wenn $X_1, \ldots, X_n$ ein Koodinatensystem auf V ist und $c : V \longrightarrow K^n$ die entsprechende Koordinatenabbildung ist, gilt (per Definition) $X_i = p_i \circ c$, wobei (wie oben) p_i die Projektion auf die i-te Komponente ist. Nun bezeichnen wir p_i auch mit X_i und erhalten somit $X_i = X_i \circ c$, wobei X_i auf der linken Seite eine Linearform auf V und auf der rechten Seite eine Linearform auf K^n ist. Dies *kann* Sinn machen, nämlich dann, wenn man sich auf den Standpunkt stellt, dass man V mittels c mit K^n "identifiziert". Hier muss man allerdings aufpassen, denn das Identifizieren kann auch zu Problemen führen. Insbesondere muss man natürlich eine andere Notation verwenden, wenn man ein Koordinatensystem auf K^n betrachtet, das verschieden vom Standardsystem ist.

Basiswechsel und Wechsel des Koordinatensystems

Seien Basen $\mathfrak{B} = (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$ und $\mathfrak{B}' = (\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n)$ von V gegeben. Sei $X_1, \ldots, X_n$ das durch $\mathfrak{B}$ definierte Koordinatensystem auf V und $X'_1, \ldots, X'_n$ das durch $\mathfrak{B}'$ definierte Koordinatensystem auf V. Wie hängen nun diese beiden Koordinatensysteme zusammen?

Sei $T := T^{\mathfrak{B}}_{\mathfrak{B}'}$ die Transformationsmatrix des Basiswechsels von $\mathfrak{B}$ nach $\mathfrak{B}'$. (Die Koordinatenmatrix von $\mathfrak{B}'$ bzgl. $\mathfrak{B}$ ist T^{-1}.)

Sei $\mathfrak{x} \in V$ mit Koordinatenvektoren $\underline{x}$ bezüglich $\mathfrak{B}$ und $\underline{x}'$ bezüglich $\mathfrak{B}'$. Dann kann man die Koordinatenvektoren ineinander umrechnen: Es ist $\underline{x}' = T\underline{x}$, also $x'_i = \sum_{j=1}^n t_{i,j} x_j$. Wenn man nun $\mathfrak{x}$ variiert, erhält man die entsprechende Identität für die Koordinatensysteme: Es ist $x_i = X_i(\mathfrak{x})$ und $x'_i = X'_i(\mathfrak{x})$ und somit

$$X'_i(\mathfrak{x}) = \sum_{j=1}^{n} t_{i,j} X_j(\mathfrak{x}) \ .$$

Da dies für alle $\mathfrak{x} \in V$ gilt, ist also

$$X_i' = \sum_{j=1}^{n} t_{i,j} X_j \,. \tag{2.21}$$

Merkregel: Wenn man die üblichen Transformationen mit "unbestimmten Koordinatenvektoren" ausführt, erhält man genau die richtigen Transformationsformeln für Koordinatenfunktionen.

Man sieht auch: Die Abbildungsmatrix von X_i' bezüglich $\mathfrak{B}$ ist gleich der i-ten Zeile von T (und der Koordinatenvektor von X_i' bezüglich der Basis $X_1, \ldots, X_n$ von V^* ist das Transponierte dieses Vektors).

Die duale Abbildung

Seien nun V und W beliebige K-Vektorräume und sei $\varphi : V \longrightarrow W$ eine lineare Abbildung. Beachten Sie: Wenn α eine Linearform auf W ist, dann ist $\alpha \circ \varphi$ eine Linearform auf V.

Wir betrachten nun die Abbildung

$$\varphi^* : W^* \longrightarrow V^* \,,\ \alpha \mapsto \alpha \circ \varphi \,.$$

Man sieht leicht, dass φ^* linear ist.

Definition Die Abbildung $\varphi^* : W^* \longrightarrow V^*$ ist die zu φ *duale* lineare Abbildung.

Offensichtlich ist $\mathrm{id}_V^* = \mathrm{id}_{V^*}$.

Sei nun X ein weiter K-Vektorraum und $\psi : W \longrightarrow X$ eine weitere lineare Abbildung. Für $\alpha \in X^*$ ist gilt nun

$$(\varphi^* \circ \psi^*)(\alpha) = \varphi^*(\psi^*(\alpha)) = \psi^*(\alpha) \circ \varphi = \alpha \circ \psi \circ \varphi = (\psi \circ \varphi)^*(\alpha)$$

und somit:

$$\varphi^* \circ \psi^* = (\psi \circ \varphi)^* \tag{2.22}$$

Hieraus folgt auch: Wenn $\varphi : V \longrightarrow W$ invertierbar ist, dann ist auch φ^* invertierbar und $(\varphi^*)^{-1} = (\varphi^{-1})^*$.

Aussage 2.47 *Sei weiterhin $\varphi : V \longrightarrow W$ linear. Dann haben wir ein kommutatives Diagramm*

$$\begin{array}{ccc} V & \xrightarrow{\ \varphi\ } & W \\ {\scriptstyle \Phi_V}\downarrow & & \downarrow{\scriptstyle \Phi_W} \\ V^{**} & \xrightarrow{\ \varphi^{**}\ } & W^{**} \end{array} \,,$$

wobei die vertikalen Pfeile die kanonischen Abbildungen $V \longrightarrow V^{**}$, $W \longrightarrow W^{**}$ *bedeuten.*

Beweis. Wir müssen zeigen: $\forall \mathfrak{v} \in V : \Phi_W(\varphi(\mathfrak{v})) = \varphi^{**}(\Phi_V(\mathfrak{v}))$.

Die Elemente aus W^{**} sind Linearformen auf W^*. Wir müssen somit zeigen: $\mathfrak{v} \in V, \forall \alpha \in W^* : (\Phi_W(\varphi(\mathfrak{v})))(\alpha) = (\varphi^{**}(\Phi_V(\mathfrak{v})))(\alpha)$.

Seien also $\mathfrak{v} \in V$ und $\alpha \in W^*$. Dann ist $(\Phi_W(\varphi(\mathfrak{v})))(\alpha) = \alpha(\varphi(\mathfrak{v}))$ und $(\varphi^{**}(\Phi_V(\mathfrak{v})))(\alpha) = (\Phi_V(\mathfrak{v}) \circ \varphi^*)(\alpha) = (\Phi_V(\mathfrak{v}))(\varphi^*(\alpha)) = (\Phi_V(\mathfrak{v}))(\alpha \circ \varphi) = (\alpha \circ \varphi)(\mathfrak{v}) = \alpha(\varphi(\mathfrak{v}))$. □

Die obige Aussage kann man so zusammenfassen: Wenn man endlich-dimensionale K-Vektorräume betrachtet, sind die kanonischen Isomorphismen kompatibel mit den linearen Abbildungen zwischen Vektorräumen und den entsprechenden bidualen linearen Abbildungen. Aus diesem Grund macht es Sinn, einen endlich-dimensionalen Vektorraum V mittels Φ mit seinem Dual zu "identifizieren".

Seien nun V, W endlich-dimensional. Wenn wir nun diese Identifizierungen vornehmen, haben wir die schöne Identität

$$\varphi = \varphi^{**} .$$

Hieraus folgt insbesondere:

Aussage 2.48 *Seien* V *und* W *endlich-dimensional. Dann ist die Abbildung* $\mathrm{Hom}_K(V, W) \longrightarrow \mathrm{Hom}_K(W^*, V^*)$ $\varphi \mapsto \varphi^*$ *ein Isomorphismus von Vektorräumen.*

Der Annullator

Sei weiterhin V ein K-Vektorrraum.

Definition Sei $S \subseteq V$. Dann ist $S^\circ := \{\alpha \in V^* \mid \forall \mathfrak{v} \in S : \alpha(\mathfrak{v}) = 0\}$ der *Annullator* (oder *Annullatorraum*) von S.

Bemerkung Sei V endlich-dimensional und sei $S \subseteq V^*$. Unter der Identifikation von V mit V^{**} ist dann $S^\circ := \{\mathfrak{v} \in V \mid \forall \alpha \in S : \alpha(\mathfrak{v}) = 0\}$.

Bemerkung Wenn $S \subseteq K^n$ ist, ist S° die Menge der homogenen Gleichungen, die auf S "verschwinden" (identisch Null sind).

Man kann aber auch mit $S \subseteq (K^n)^*$, d.h. mit einer Menge von homogenen Gleichungen anfangen. Unter der Identifikation von V mit V^{**} ist dann S° die Lösungsmenge des von S definierten homogenen LGS.

Das folgende Lemma ist leicht:

Lemma 2.49

a) Sei $S \subseteq V$. Dann ist S° ein Untervektorraum von V^*.

b) Seien $S_1 \subseteq S_2 \subseteq V$. Dann ist $S_2^\circ \leq S_1^\circ$.

c) Sei $S \subseteq V$. Dann ist $S^\circ = \langle S \rangle^\circ$.

Aussage 2.50 *Sei nun V endlich-dimensional, und sei U ein Untervektorraum von V. Dann gilt:*

a) $\mathrm{Dim}(U^\circ) = \mathrm{Dim}(V) - \mathrm{Dim}(U)$

b) Unter der Identifikation von V mit V^{**} ist $U^{\circ\circ} = U$.

Beweis.

a) Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_k$ eine Basis von U. Wir ergänzen diese Basis zu einer Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V. Wir behaupten, dass nun $\mathfrak{b}_{k+1}^*, \ldots, \mathfrak{b}_n^*$ eine Basis von U° ist. Klar ist das System linear unabhängig, wir müssen zeigen, dass es ein Erzeugendensystem ist. Sei hierzu $\alpha \in U^\circ$. Dann ist

$$\alpha = \sum_{i=1}^{n} \alpha(\mathfrak{b}_i)\mathfrak{b}_i^* \ .$$

Nun ist $\alpha(\mathfrak{b}_i) = 0$ für $i = 1, \ldots, k$, also

$$\alpha = \sum_{i=k+1}^{n} \alpha(\mathfrak{b}_i)\mathfrak{b}_i^* \ .$$

b) Offensichtlich ist $U \subseteq U^{\circ\circ}$. Nach a) ist ferner $\mathrm{Dim}(U) = \mathrm{Dim}(U^{\circ\circ})$. □

Bemerkung Die Identität $U = U^{\circ\circ}$ hat die folgende Interpretation in K^n: Sei U ein Untervektorraum von K^n. Dann ist $G := U^\circ$ die Menge der homogenen Gleichungen, die auf U verschwinden. Nun gilt also $G^\circ = U$, d.h. U ist die Lösungsmenge von G. Also: Man kann zwischen linearen Unterräumen von K^n und linearen Unterräumen von Gleichungen auf K^n "hin- und hergehen".

Aussage 2.51 *Sei $\varphi : V \longrightarrow W$ linear.*

a) Es ist $\mathrm{Kern}(\varphi^*) = \mathrm{Bild}(\varphi)^\circ$.

b) Sei W endlich-dimensional. Dann ist $\mathrm{Dim}(\mathrm{Bild}(\varphi)) = \mathrm{Dim}(\mathrm{Bild}(\varphi^*))$.

Beweis. a) Sei $\alpha \in W^*$. Dann gilt: $\alpha \in \mathrm{Kern}(\varphi^*) \longleftrightarrow \alpha \circ \varphi = 0 \longleftrightarrow \forall \mathfrak{v} \in V : \alpha(\varphi(\mathfrak{v})) = \mathfrak{o} \longleftrightarrow \alpha \in (\varphi(V))^\circ = \mathrm{Bild}(\varphi)^\circ$.

b) Es ist $\mathrm{Dim}(\mathrm{Bild}(\varphi^*)) = \mathrm{Dim}(W^*) - \mathrm{Dim}(\mathrm{Kern}(\varphi^*)) = \mathrm{Dim}(W) - \mathrm{Dim}((\mathrm{Bild}(\varphi))^\circ) = \mathrm{Dim}(\mathrm{Bild}(\varphi))$. □

Abbildungsmatrizen und die duale Abbildung

Seien nun V und W endlich-dimensionale K-Vektorräume. Wir fixieren Basen $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V und $\mathfrak{c}_1, \ldots, \mathfrak{c}_m$ von W. Seien $X_1, \ldots, X_n$ und $Y_1, \ldots, Y_m$ die entsprechenden Koordinatensysteme.

Nun gibt es zu beliebigen Vektoren $\mathfrak{x}_1, \ldots, \mathfrak{x}_n \in W$ genau eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\varphi(\mathfrak{b}_j) = \mathfrak{x}_j$ für alle j.

Ebenso gibt es für beliebige Linearformen $\alpha_1, \ldots, \alpha_m$ auf V genau eine lineare Abbildung von W^* nach V^*, die Y_i auf α_i abbildet. Nach Aussage 2.48 bedeutet dies, dass es genau eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\varphi^*(Y_i) = \alpha_i$ (d.h. mit $Y_i \circ \varphi = \alpha_i$) für alle i gibt. Kurz: Anstatt anzugeben, wie eine Basis von V abgebildet wird, kann man auch angeben, wie ein Koordinatensystem von W abgebildet wird.

Sei nun $M = ((m_{i,j}))_{i,j} := M_{\mathfrak{C}}^{\mathfrak{B}}(\varphi)$. Sei $\mathfrak{x} \in V$ mit Koordinatenvektor $\underline{x} \in K^n$, und sei $\underline{y}$ der Koordinatenvektor von $\varphi(\mathfrak{x}) \in W$ (bzgl. der angegebenen Basen). Dann ist $\underline{y} = M\underline{x}$. Nun ist $y_i = Y_i(\varphi(\mathfrak{x}))$ und $x_i = X_i(\mathfrak{x})$ und somit

$$Y_i(\varphi(\mathfrak{x})) = \sum_{j=1}^{n} m_{i,j} X_j(\mathfrak{x}) \ .$$

Da $\mathfrak{x}$ beliebig ist, gilt somit

$$\varphi^*(Y_i) = Y_i \circ \varphi = \sum_{j=1}^{n} m_{i,j} X_j \ .$$

(Mit $\varphi = \mathrm{id}$ erhalten wir wieder (2.21).)

Wir betrachten nun die Abbildungsmatrix von φ^* bzgl. $Y_1, \ldots, Y_m$ und $X_1, \ldots, X_n$. Nach der obigen Formel ist die i-te Spalte dieser Matrix gleich der i-ten Zeile von M. Also ist

$$M_{\mathfrak{C}}^{\mathfrak{B}}(\varphi)^t = M_{\mathfrak{B}^*}^{\mathfrak{C}^*}(\varphi^*) \ .$$

Diese Formel kann man sich auch wie folgt überlegen: Sei α eine Linearform auf W. Dann haben wir die Identität

$$m_{\mathfrak{B}}(\varphi^*(\alpha)) = m_{\mathfrak{B}}(\alpha \circ \varphi) = m_{\mathfrak{C}}(\alpha) \cdot M_{\mathfrak{C}}^{\mathfrak{B}}(\varphi)$$

Kurz: $M_{\mathfrak{C}}^{\mathfrak{B}}(\varphi)$ operiert auf Abbildungsmatrizen von Linearformen auf W "von rechts" (und auf Koordinatenvektoren von V von links). Durch transponieren erhält man:

$$c_{\mathfrak{B}^*}(\varphi^*(\alpha)) = (M_{\mathfrak{C}}^{\mathfrak{B}}(\varphi))^t \cdot c_{\mathfrak{C}^*}(\alpha)$$

Dies bedeutet gerade, dass $M_{\mathfrak{C}}^{\mathfrak{B}}(\varphi)^t = M_{\mathfrak{B}^*}^{\mathfrak{C}^*}(\varphi^*)$ ist.

Mit dieser Formel und Aussage 2.51 kann man auch einen "theoretischeren" Beweis von "Spaltenrang=Zeilenrang" geben. Sei hierfür $A \in K^{m\times n}$. Dann ist also A^t gleich der Abbildungsmatrix von A^* bezüglich der Standarddualbasen.

Nun ist per Definition der Spaltenrang von A gleich $\mathrm{Dim}(\mathrm{Bild}(\Lambda_A))$. Ferner ist der Zeilenrang von A gleich $\mathrm{Dim}(\mathrm{Bild}(\Lambda_{A^t})) = \mathrm{Dim}(\mathrm{Bild}(\Lambda_A^*))$. Somit ist nach Aussage 2.51 der Spaltenrang von A gleich dem Zeilenrang von A.

2.10 Determinanten

Zur Motivation der Definition der Determinante nehmen wir uns die folgende Aufgabe vor:

Gegeben $\underline{x}_1, \ldots, \underline{x}_n \in \mathbb{R}^n$ wollen wir definieren, was das *Volumen* des Spats

$$\{c_1\underline{x}_1 + \cdots + c_n\underline{x}_n \,|\, 0 \leq c_j \leq 1 \text{ für alle } j = 1, \ldots, n\}$$

ist. Hierzu suchen wir eine Abbildung $\mathrm{Vol} : (\mathbb{R}^n)^n \longrightarrow \mathbb{R}_{\geq 0}$, so dass für alle $\underline{x}_1, \ldots, \underline{x}_n \in \mathbb{R}^n$ $\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_n) \in \mathbb{R}$ unserer intuitiven Vorstellung des Volumens des Spats entspricht.

Wir stellen die folgenden naheliegenden Forderungen (wobei $\underline{x}_1, \ldots, \underline{x}_n$ beliebige Vektoren aus $\mathbb{R}^n$ sind, $c \in \mathbb{R}$ und $i, j = 1, \ldots, n$ mit $i \neq j$ ist).

Vol1 $\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_{j-1}, c\underline{x}_j, \underline{x}_{j+1}, \ldots, \underline{x}_n) = |c| \cdot \mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_j, \underline{x}_{j+1}, \ldots, \underline{x}_n)$

Vol2 $\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}_j + \underline{x}_i, \underline{x}_{j+1}, \ldots, \underline{x}_n) = \mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_n)$

Vol3 $\mathrm{Vol}(\underline{e}_1, \ldots, \underline{e}_n) = 1$

Eine andere Motivation der Determinante kann man über beliebigen Körpern formulieren: Wir suchen eine Abbildung, die jeder Matriz ein Skalar zuordnet, so dass die Matriz genau dann invertierbar ist, wenn dieses Skalar $\neq 0$ ist. Außerdem soll die Abbildung noch einige weitere "angenehme Eigenschaften" bzgl. elementaren Spaltentransformationen haben.

Sei also K ein Körper und $n \in \mathbb{N}$. Wir suchen eine Abbildung $\mathrm{Det} : K^{n\times n} \longrightarrow K$, die die folgenden Eigenschaften hat:

Det1 Sei $A \in K^{n\times n}$, und sei A' eine Matrix, die aus A durch Multiplikation einer Spalte mit einem Skalar c hervorgeht. Dann gilt $\mathrm{Det}(A') = c\,\mathrm{Det}(A)$.

Det2 Sei $A \in K^{n\times n}$, und sei A' die Matrix, die aus A durch Addition von Spalte i zu Spalte j (mit $i \neq j$) hervorgeht. Dann gilt $\mathrm{Det}(A) = \mathrm{Det}(A')$.

Det3 $\mathrm{Det}(I_n) = 1$.

(Beachten Sie, dass aber für $c = 0$ die Transformation in Det1 keine elementare Spaltentransformation ist.)

So eine Abbildung $K^{n\times n} \longrightarrow K$ heißt eine *Determinantenabbildung*. Wir werden sehen, dass es genau eine Determinantenabbildung $K^{n\times n} \longrightarrow K$ gibt. Wir werden auch sehen, dass diese Determinantenabbildung die Eigenschaft hat, dass $\mathrm{Det}(A) = 0$ genau dann wenn $\mathrm{Rang}(A) < n$, was die ursprüngliche Motivation war.

Ferner werden wir dann auch sehen, dass es genau eine Abbildung $\mathrm{Vol} : (\mathbb{R}^n)^n \longrightarrow \mathbb{R}$ mit den Eigenschaften Vol1, Vol2 und Vol3 gibt (Aussage 2.62).

Notation Wenn eine Abbildung $f : K^{m\times n} \longrightarrow K$ gegeben ist, erhält man mittels $(\underline{x}_1, \ldots, \underline{x}_n) \mapsto f(\underline{x}_1|\cdots|\underline{x}_n)$ eine Abbildung $(K^m)^n \longrightarrow K$. Wir bezeichnen diese Abbildung wieder mit f, d.h. wir setzen $f(\underline{x}_1, \ldots, \underline{x}_n) := f(\underline{x}_1|\ldots|\underline{x}_n)$. Umgekehrt identifizieren wir Abbildungen $(K^m)^n \longrightarrow K$ mit Abbildungen $K^{m\times n} \longrightarrow K$.

Eindeutigkeit und Existenz einer Determinantenabbildung

Wir fixieren nun eine Determinantenabbildung $d : K^{n\times n} \longrightarrow K$ und leiten einige Eigenschaften her. Mittels dieser Eigenschaften werden wir dann insbesondere zeigen, dass es höchstens eine Determintenabbildung $K^{n\times n} \longrightarrow K$ gibt. Danach werden wir eine dieser Eigenschaften als Ansatz für eine Definition benutzen und zeigen, dass die so definierte Abbildung wirklich eine Determinatenabbildung ist.

Es ist leicht zu beschreiben, wie sich $d(A)$ (mit $A \in K^{n\times n}$) unter elementaren Spaltentransformationen angewandt auf A ändert. Wir kennen schon das Transformationsverhalten unter Multiplikation einer Spalte mit einem Skalar $c \neq 0$. Außerdem haben wir:

Lemma 2.52 *Sei $A \in K^{n\times n}$ und sei A' eine Matrix, die aus A durch Anwendung einer elementaren Spaltentransformation hervorgeht. Dann gilt:*

a) Wenn A' aus A durch Vertauschen von zwei Spalten hervorgeht, gilt $d(A') = -d(A)$.

b) Wenn A' aus A durch Addition des c-fachen einer Spalte zu einer anderen Spalte hervorgeht, gilt $d(A') = d(A)$.

Beweis. Wir zeigen zuerst die zweite Aussage. Sei A' die Matrix, die man aus A durch Addition des c-fachen von Spalte i zu Spalte j erhält. Für $c = 0$ ist nichts zu zeigen, sei also $c \neq 0$. Wir haben

$$\begin{aligned} c\,d(A) &\overset{\text{Det1}}{=} d(\underline{a}_1, \ldots, \underline{a}_{i-1}, c\underline{a}_i, \underline{a}_{i+1}, \ldots, \underline{a}_{j-1}) \\ &\overset{\text{Det2}}{=} d(\underline{a}_1, \ldots, \underline{a}_{i-1}, c\underline{a}_i, \underline{a}_{i+1}, \ldots, \underline{a}_{j-1}, \underline{a}_j + c\underline{a}_i, \underline{a}_{j+1}, \underline{a}_n) \\ &\overset{\text{Det1}}{=} c\,d(A')\ , \end{aligned}$$

und hieraus folgt die Behauptung. (Wir haben hier implizit angenommen, dass $i < j$, aber das hat nur notationelle Gründe.)

Die erste Behauptung folgt, da das Vertauschen von zwei Spalten durch wiederholte Addition und Subtraktion einer Spalte, gefolgt mit Multiplikation einer Spalte mit -1, dargestellt werden kann (Mit anderen Worten: Man braucht Umformung (II) im Gauß-Algorithmus eigentlich gar nicht.)

In der Tat, die Vertauschung von Spalten i und j kann man so realisieren:
- Man addiert Spalte i zu Spalte j. (Dann steht in Spalte j $\underline{a}_i + \underline{a}_j$.)
- Man subtahiert Spalte j von Spalte i. (Dann steht in Spalte i $\underline{a}_i - (\underline{a}_i + \underline{a}_j) = -\underline{a}_j$.)
- Man addiert Spalte i zu Spalte j. (Dann steht in Spalte j $\underline{a}_i$.)
- Man multipliziert Spalte i mit -1. □

Bemerkung Zusammen mit der Eigenschaft $d(I_n) = 1$ folgt aus dem obigen Lemma insbesondere, dass $d(E)$ für eine Elementarmatrix E durch Axiome Det1, Det2, Det3 eindeutig festgelegt ist und dass immer $d(E) \neq 0$ gilt.

Wenn man beachtet, dass eine elementare Spaltentransformationen zur Multiplikation mit einer Elementarmatrix von rechts korrespondiert, erhält man aus dem obigen Lemma sofort:

Lemma 2.53 *Sei $A \in K^{n\times n}$ eine beliebige Matrix, und sei $E \in K^{n\times n}$ eine Elementarmatrix. Dann gilt $d(AE) = d(A) \cdot d(E)$.*

Per Induktion nach k folgt:

Lemma 2.54 *Sei $A \in K^{n\times n}$ beliebig, und seien $E_1, \ldots, E_k$ $n \times n$-Elementarmatrizen. Dann gilt $d(AE_1 \cdots E_k) = d(A) \cdot d(E_1) \cdots d(E_k)$.*

Aussage 2.55 *Es gibt höchstens eine Determinantenabbildung $d : K^{n\times n} \longrightarrow K$, und diese erfüllt $d(A) = 0$ genau dann wenn* $\mathrm{Rang}(A) < n$.

Beweis. Sei nach wie vor $d : K^{n\times n} \longrightarrow K$ eine Determinantenabbildung, und sei $A \in K^{n\times n}$.

Sei zunächst $\text{Rang}(A) = n$, also A invertierbar. Dann gibt es Elementarmatrizen $E_1, \ldots, E_k$ mit $A = E_1 \cdots E_k$ (siehe Satz 2.3). Dann gilt nach dem obigen Lemma $d(A) = d(E_1) \cdots d(E_k)$. Nun ist die Determinante einer Elementarmatrix eindeutig festgelegt und $\neq 0$ (s.o.). Damit ist auch die Determinante von A eindeutig festgelegt und $\neq 0$.

Sei nun $\text{Rang}(A) < n$. Dann gibt es also eine Matrix $\tilde{A}$, die eine Nullspalte enthält (z.B. eine Matrix in Spaltenstufenform) sowie Elementarmatrixen $E_1, \ldots, E_k$, so dass $A = \tilde{A}E_1 \cdots E_k$. Nach Det1 ist $d(\tilde{A}) = 0$ und somit $d(A) = d(\tilde{A}) \cdot d(E_1) \cdots d(E_k) = 0$. □

Die Beweismethode hat einige recht einfache Konsequenzen:

Aussage 2.56 *Für $A \in K^{n\times n}$ gilt $d(A) = d(A^t)$.*

Beweis. Offensichtlich ist für eine Elementarmatrix E $d(E) = d(E^t)$. Wenn $\text{Rang}(A) = n$ ist, gibt es Elementarmatrizen $E_1, \ldots, E_k$ mit $A = E_1 \cdots E_k$. Damit gilt $d(A^t) = d(E_k^t \cdots E_1^t) = d(E_k^t) \cdots d(E_1^t) = d(E_1) \cdots d(E_k) = d(A)$.

Wenn $\text{Rang}(A) < n$, ist auch $\text{Rang}(A^t) = \text{Rang}(A) < n$. Damit gilt $d(A^t) = 0 = d(A)$. □

Bemerkung Sei wiederum $A \in K^{n\times n}$. Aufgrund der obigen Aussage hat die Determinantenfunktion d die folgenden Eigenschaften bezüglich Transformation von A mittels elementarer Zeilenoperationen:

- Sei A' eine Matrix, die aus A durch Multiplikation einer Zeile mit einem Skalar c hervorgeht. Dann gilt $d(A') = c\, d(A)$.
- Sei A' eine Matrix, die aus A durch Vertauschen von zwei Zeilen hervorgeht. Dann gilt $d(A') = -d(A)$.
- Sei A' die Matrix, die aus A durch Addition des c-fachen einer Zeile zu einer anderen Zeile hervorgeht. Dann gilt $d(A') = d(A)$.

Aussage 2.57 (Multiplikativität) *Für $A, B \in K^{n\times n}$ gilt $d(AB) = d(A) \cdot d(B)$.*

Beweis. Sei zunächst $\text{Rang}(A) = \text{Rang}(B) = n$. Dann gibt es Elementarmatrizen $E_1, \ldots, E_k$ und $E_{k+1}, \ldots, E_\ell$ mit $A = E_1 \cdots E_k$ und $B = E_{k+1} \cdots E_\ell$. Damit ist $d(AB) = d(E_1 \cdots E_\ell) = d(E_1) \cdots d(E_\ell) = d(A)\, d(B)$.

Sei nun der Rang (mindestes) einer der beiden Matrizen $< n$. Wie das folgende Lemma zeigt, ist dann auch $\text{Rang}(AB) < n$. Damit gilt $d(AB) = 0 = d(A)\, d(B)$. □

Lemma 2.58 *Sei $A \in K^{m\times n}$ und $B \in K^{n\times r}$. Dann ist* $\text{Rang}(AB) \leq \min\{\text{Rang}(A), \text{Rang}(B)\}$.

Beweis. Wir benutzen die Definition des Rangs als Spaltenrang.

Der Spaltenraum von AB ist gleich $\text{Bild}(\Lambda_{AB}) = \Lambda_{AB}(K^r) = \Lambda_A(\Lambda_B(K^r))$. Die Dimension dieses Raumes ist $\leq \min\{\text{Dim}(\Lambda_A(K^n)), \text{Dim}(\Lambda_B(K^r))\} = \min\{\text{Rang}(A), \text{Rang}(B)\}$ (siehe Aussage 2.20). □

Eine Folgerung von Aussage 2.57 ist wiederum:

Aussage 2.59 *Die Determinantenabbildung d ist ein Homomorphismus von Monoiden von $(K^{n\times n}, \cdot)$ nach $(K, \cdot)$. Insbesondere ist $d : (K^{n\times n})^* \longrightarrow K^*$ ein Homomorphismus von Gruppen.*

Beweis. Die erste Aussage ist eine Zusammenfassung der obigen Aussage und des Axioms Det3: $d(I_n) = 1$. Die zweite Aussage folgt sofort aus der ersten.□

Aussage 2.60 (Multilinearität) *Es gilt für alle $j = 1, \ldots, n$ und alle $\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}_{j+1}, \ldots, \underline{x}_n \in K^n$: Die Abbildung*

$$d : K^n \longrightarrow K \, , \underline{x} \mapsto d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}, \underline{x}_{j+1}, \ldots, \underline{x}_n)$$

ist linear.

Beweis. Wir müssen nur zeigen, dass stets $d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}+\underline{y}, \underline{x}_{j+1}, \ldots, \underline{x}_n) = d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}, \underline{x}_{j+1}, \ldots, \underline{x}_n) + d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{y}, \underline{x}_{j+1}, \ldots, \underline{x}_n)$ gilt.

Wir wissen, dass $n+1$ Vektoren in K^n immer linear abhängig sind. Wir haben also $c, d \in K$ und $c_1, \ldots, c_{j-1}, c_{j+1}, \ldots, c_n \in K$, nicht alle $= 0$, mit

$$c\underline{x} + d\underline{y} + \sum_{k\neq j} c_k \underline{x}_k = 0 \, .$$

Wenn $c = d = 0$ ist, sind die Vektoren $\underline{x}_k$ $(k \neq j)$ linear abhängig. Damit haben alle drei Matrizen, die hier betrachtet werden, Rang $< n$. Somit gilt die Behauptung.

Sei also $c \neq 0$ oder $d \neq 0$. Wie können o.E. annehmen, dass $d \neq 0$ und sogar $d = -1$. Dann ist also $\underline{y} = c\underline{x} + \sum_{k\neq j} c_k \underline{x}_k$, und wir haben nun

$$\begin{aligned}
&d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{y}, \underline{x}_{j+1}, \ldots, \underline{x}_n) = \\
&d(\underline{x}_1, \ldots, \underline{x}_{j-1}, c\underline{x} + \textstyle\sum_{k\neq j} c_k \underline{x}_k, \underline{x}_{j+1}, \ldots, \underline{x}_n) = \\
&d(\underline{x}_1, \ldots, \underline{x}_{j-1}, c\underline{x}, \underline{x}_{j+1}, \ldots, \underline{x}_n) = \\
&c \cdot d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}, \underline{x}_{j+1}, \ldots, \underline{x}_n) \, .
\end{aligned}$$

Nach demselben Argument ist

$$\begin{aligned}
&d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x} + \underline{y}, \underline{x}_{j+1}, \ldots, \underline{x}_n) = \\
&d(\underline{x}_1, \ldots, \underline{x}_{j-1}, (1+c)\underline{x} + \textstyle\sum_{k\neq j} c_k \underline{x}_k, \underline{x}_{j+1}, \ldots, \underline{x}_n) = \\
&(1+c) \cdot d(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}, \underline{x}_{j+1}, \ldots, \underline{x}_n) \, .
\end{aligned}$$

Aus der Verbindung dieser beiden Identitäten folgt die Behauptung. □

Sei nun $n \geq 2$.

Notation Sei für $A \in K^{n\times n}$ und $i,j = 1,\dots n$ $A_{i,j}$ diejenige $n-1\times n-1$-Matrix, die entsteht, wenn man in A die i-te Zeile und die j-te Spalte streicht.

Betrachten wir die Abbildung $K^{(n-1)\times(n-1)} \longrightarrow K\,, A \mapsto d(\begin{pmatrix} 1 & \\ & A \end{pmatrix})$. Man sieht leicht, dass dies eine Determinantenabbildung auf $K^{(n-1)\times(n-1)}$ ist. Wir bezeichnen diese Abbildung mit d_{n-1}, und die wir setzen $d_n := d$.

Aussage 2.61 (Laplacescher Entwicklungssatz) *Sei $A \in K^{n\times n}$. Dann gilt für alle $j = 1,\dots,n$:*

$$d_n(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j}\, d_{n-1}(A_{i,j}) \qquad \textit{(Entwicklung nach Spalte } j\textit{)}$$

Analog gilt für alle $i = 1,\dots,n$:

$$d_n(A) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{i,j}\, d_{n-1}(A_{i,j}) \qquad \textit{(Entwicklung nach Zeile } i\textit{)}$$

Beweis. Die beiden Aussagen sind äquivalent da $d_n(A) = d_n(A^t)$ und $d_{n-1}(A) = d_{n-1}(A^t)$. Wir zeigen die Entwicklung nach Spalte j.

Aufgrund der Multilinearität ist

$$\begin{aligned} d(A) = &\ d(\underline{a}_1,\dots,\underline{a}_{j-1},\textstyle\sum_{i=1}^{n} a_{i,j}\,\underline{e}_i,\underline{a}_{j+1},\dots,\underline{a}_n) \\ = &\ \textstyle\sum_{i=1}^{n} a_{i,j}\,(\underline{a}_1,\dots,\underline{a}_{j-1},\underline{e}_i,\underline{a}_{j+1},\dots,\underline{a}_n)\,. \end{aligned}$$

Es reicht also zu zeigen, dass

$$d_n(\underline{a}_1,\dots,\underline{a}_{j-1},\underline{e}_i,\underline{a}_{j+1},\dots,\underline{a}_n) = (-1)^{i+j} d_n(A_{i,j})\,.$$

Sei zunächst

$$A = (\underline{e}_1|a_2|\cdots|\underline{a}_n) = \begin{pmatrix} 1 & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,n} \end{pmatrix}\,.$$

Dann ist $d_n(A) = d_n(\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,n} \end{pmatrix})$ aufgrund von Spaltentransformationen, und letzteres ist per Definition gleich $d_{n-1}(A_{1,1})$. Somit ist die Formel für solche Matrizen richtig.

Sei nun $A = (\underline{e}_i|a_2|\cdots|\underline{a}_n)$ eine Matrix wie oben, nur das die Eins in in der i-ten Zeile steht. Wieder "räumen wir zunächst die Einträge rechts der 1 aus" (Addition von Vielfachen der ersten Spalte zu anderen Spalten). Wir führen hintereinander die Zeilentransformationen "Vertauschen der Zeile i mit Zeile $i-1$", "Vertauschen der Zeile $i-1$ mit Zeile $i-2$", ..., "Vertauschen der Zeile 2 mit Zeile 1" durch und nennen das Ergebnis A'. Dann ist $A'_{1,1} = A_{i,1}$ und $A' = \begin{pmatrix} 1 & 0 \\ 0 & A_{i,1} \end{pmatrix}$. Damit ist $d_n(A) = (-1)^{i-1}\, d_n(A') = (-1)^{i+1}\, d_{n-1}(A_{i,1})$. Die Formel ist also wiederum richtig.

Sei nun $A = (\underline{a}_1|\cdots|\underline{a}_{j-1}|\underline{e}_i|\underline{a}_{j+1}|\cdots|\underline{a}_n)$. Wir gehen analog vor und "räumen zuerst die Einträge rechts und links vom Eintrag mit Index (i,j) aus". Dann vertauschen wir der Reihe nach die Spalte j mit der Spalte $j-1$, ..., die Spalte 2 mit der Spalte 1. Dann vertauschen wir noch die Zeilen wie oben. Wir erhalten die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & A_{i,j} \end{pmatrix}$. Wir haben nun $d_n(A) = (-1)^{i-1+j-1}\, d_{n-1}(A_{i,j}) = (-1)^{i+j}\, d_{n-1}(A_{i,j})$, abermals die richtige Formel. □

Wir kommen nun zum Hauptresultat über Determinantenabbildungen.

Satz 2.7 *Sei K ein Körper und $n \in \mathbb{N}$. Dann gibt es genau eine Determinantenabbildung $K^{n\times n} \longrightarrow K$.*

Beweis. Die Eindeutigkeit haben wir schon gezeigt.

Motiviert durch die obige Entwicklung nach Zeilen (die ja für jede Determinantenabbildung gelten muss) *definieren* wir induktiv für jeden Körper K:

$$\mathrm{Det}_1(a) := a \text{ für } a \in K \quad , \quad \mathrm{Det}_n(A) := \sum_{j=1}^{n} (-1)^{1+j} a_{1,j}\, \mathrm{Det}_{n-1}(A_{1,j})$$

für $n \in \mathbb{N}$ und $A \in K^{n\times n}$. Ich behaupte, dass für alle $n \in \mathbb{N}$ Det_n eine Determinantenabbildung ist.

Wir zeigen dies nach Induktion über n. Der Induktionsanfang $n = 1$ ist trivial. Wir setzen also voraus, dass die Behauptung für n richtig ist und zeigen die Behauptung für $n+1$.

Beachten Sie, dass wir dann insbesondere alle oben bewiesenen Eigenschaften für Determinantenabbildungen für Det_n benutzen dürfen.

Offensichtlich ist $\mathrm{Det}_{n+1}(I_{n+1}) = 1 \cdot \mathrm{Det}(I_n) = 1$.

Sei $A \in K^{(n+1)\times(n+1)}$, $j = 1, \ldots, n+1$ und A' diejenige Matrix, die aus A durch Multiplikation der j-ten Spalte mit $c \in K$ hervorgeht. Dann ist $\mathrm{Det}_n(A'_{1,k}) = c\, \mathrm{Det}_n(A_{1,k})$ für alle $k \neq j$. Damit gilt $\mathrm{Det}_{n+1}(A') = \sum_{j=1}^{n+1} (-1)^{1+j}\, c\, a_{1,j}\, \mathrm{Det}_n(A_{1,j}) = c \cdot \mathrm{Det}_{n+1}(A)$.

Sei nun $A \in K^{(n+1)\times(n+1)}$, und seien $i, j = 1, \ldots, n+1$ mit $i \neq j$. Sei A' diejenige Matrix, die aus A durch Addition der i-ten Spalte zur j-ten Spalte hervorgeht. Für $k \neq i, j$ entsteht dann $A'_{1,k}$ auch aus $A_{1,k}$, indem eine Spalte zu einer anderen addiert wird. Wenn wir nun anwenden, dass Det_n eine Determinantenabbildung ist, erhalten wir, dass $\mathrm{Det}_n(A'_{1,k}) = \mathrm{Det}_n(A_{1,k})$ für alle $k \neq i, j$. Außerdem ist dies offensichtlich auch für $k = j$ richtig, denn diese Spalte wird ja gerade gestrichen.

Wir untersuchen nun $\mathrm{Det}_n(A'_{1,i})$. Es ist $\mathrm{Det}_n(A'_{1,i}) = \mathrm{Det}_n(A_{1,i}) + \mathrm{Det}_n(B)$, wobei B aus A hervorgeht, indem man zuerst die Spalten i und j vertauscht und dann die Spalte i streicht. Also hat B bis auf Reihenfolge dieselben Spalten wie die Matrix $A_{1,j}$. Genauer geht B aus $A_{1,j}$ durch $|j-i|-1$ Spaltenvertauschungen hervor. Damit ist also $\mathrm{Det}_n(A'_{1,i}) = \mathrm{Det}_n(A_{1,i}) + (-1)^{|j-i|-1}\,\mathrm{Det}_n(A_{1,j}) = \mathrm{Det}_n(A_{1,i}) + (-1)^{j-i+1}\,\mathrm{Det}_n(A_{1,j})$.

Es folgt:

$$\begin{aligned}
& \mathrm{Det}_{n+1}(A') - \mathrm{Det}_{n+1}(A) \\
= \;& (-1)^{1+i} a_{1,i}\,\mathrm{Det}_n(A'_{1,i}) + (-1)^{1+j} a'_{1,j}\,\mathrm{Det}_n(A_{1,j}) \\
& -(-1)^{1+i} a_{1,i}\,\mathrm{Det}_n(A_{1,i}) - (-1)^{1+j} a_{1,j}\,\mathrm{Det}_n(A_{1,j}) \\
= \;& (-1)^{1+i} a_{1,i}\,\mathrm{Det}_n(A_{1,i}) + (-1)^{1+i}(-1)^{j-i+1} a_{1,i}\,\mathrm{Det}_n(A_{1,j}) \\
& \quad +(-1)^{1+j} a_{1,j}\,\mathrm{Det}_n(A_{1,j}) + (-1)^{1+j} a_{1,i}\,\mathrm{Det}_n(A_{1,j}) \\
& -(-1)^{1+i} a_{1,i}\,\mathrm{Det}_n(A_{1,i}) - (-1)^{1+j} a_{1,j}\,\mathrm{Det}_n(A_{1,j}) \\
= \;& 0
\end{aligned}$$

□

Notation Im Folgenden schreiben wir Det statt Det_n. Eine andere übliche Schreibweise für $\mathrm{Det}\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & . & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$ ist $\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & . & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}$.

Hier noch die zu Beginn dieses Abschnitts versprochene Aussage:

Aussage 2.62 *Es gibt genau eine Abbildung* $\mathrm{Vol} : (\mathbb{R}^n)^n \longrightarrow \mathbb{R}$*, die Vol1, Vol2 und Vol3 erfüllt.*

Beweis. Die Existenz ist klar: Die Funktion $(\mathbb{R}^n)^n \longrightarrow \mathbb{R}$, $(\underline{x}_1, \ldots, \underline{x}_n) \mapsto |\,\mathrm{Det}(\underline{x}_1, \ldots, \underline{x}_n)|$ erfüllt die Eigenschaften.

Sei anderseits Vol eine Funktion mit den angegebenen Eigenschaften. Sei

$$\epsilon : (\mathbb{R}^n)^n \longrightarrow \mathbb{R} \ , \ (\underline{x}_1, \ldots, \underline{x}_n) \mapsto \begin{cases} 1 \ , & \text{falls } \mathrm{Det}(\underline{x}_1, \ldots, \underline{x}_n) \geq 0 \\ -1 \ , & \text{falls } \mathrm{Det}(\underline{x}_1, \ldots, \underline{x}_n) < 0 \end{cases}$$

Dann ist die Funktion

$$(\mathbb{R}^n)^n \longrightarrow \mathbb{R} \ , \ (\underline{x}_1, \ldots, \underline{x}_n) \mapsto \epsilon(\underline{x}_1, \ldots, \underline{x}_n) \cdot \mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_n)$$

eine Determinantenfunktion. Also ist $\epsilon \cdot \mathrm{Vol} = \mathrm{Det}$ bzw. $\mathrm{Vol} = \epsilon \cdot \mathrm{Det}$. □

Die Formel von Leibniz

Man kann die Spalten- / Zeilnentwicklungen der Determinante verwenden, um eine nicht-rekursive Formel herzuleiten. Es ergibt sich:

$\underline{n = 2}$. Es ist $\mathrm{Det}(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,2}$.

$\underline{n = 3}$. Es ist $\mathrm{Det}(A) = a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{2,1} \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} + a_{3,1} \begin{vmatrix} a_{1,2} & a_{1,2} \\ a_{2,2} & a_{2,3} \end{vmatrix} =$
$a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3}$.
Dies ist die so genannte *Regel von Sarrus*.

Für beliebiges n benötigen wir zunächst eine Definition:

Definition / Bemerkung Wir ordnen jeder Permutation $\pi \in S_n$ die entsprechende *Permutationsmatrix* $M_\pi := (\underline{e}_{\pi(1)}, \ldots, \underline{e}_{\pi(n)}) \in K^{n \times n}$ zu. Man sieht leicht, dass diese Matrix immer invertierbar ist, und dass die Abbildung $S_n \longrightarrow (K^{n \times n})^*$ ein Gruppenhomomorphismus ist. Wir betrachten nun Permutationsmatrizen über $\mathbb{Q}$ und definieren nun das *Signum* von $\pi \in S_n$ wie folgt: $\mathrm{sign}(\pi) := \mathrm{Det}(M_\pi) \in \mathbb{Q}^*$. Da auch $\mathrm{Det} : (\mathbb{Q}^{n \times n})^* \longrightarrow \mathbb{Q}^*$ ein Gruppenhomomorphismus ist, ist also $\mathrm{sign} : S_n \longrightarrow \mathbb{Q}^*$ ein Gruppenhomomorphismus.

Eine *Transposition* ist per Definition eine Permutation, die genau zwei Elemente vertauscht und die anderen Elemente fest lässt. Man sieht leicht, dass jede Permutation ein Produkt von Transpositionen ist (= aus Transpositionen durch Verknüpfung hervorgeht). (Beweisen Sie dies per Induktion!)

Wenn π eine Transposition ist, ist $\mathrm{Det}(\pi) = -1$ per Definition. Damit gilt: Sei $\pi = \pi_1 \cdots \pi_k$, wobei die π_i Transpositionen sind. Dann ist $\mathrm{sign}(\pi) = (-1)^k$. Insbesondere ist also $\mathrm{sign}(\pi) = \pm 1$.

Man sieht auch: Wenn $\pi_1 \cdots \pi_k = \sigma_1 \cdots \sigma_\ell$ mit Transpositionen π_i und σ_i, dann sind entweder k und ℓ beide gerade oder beide sind ungerade.

Wir haben nun:

Aussage 2.63 (Formel von Leibniz) *Sei $A \in K^{n\times n}$. Dann ist*

$$\mathrm{Det}(A) = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \ .$$

Beweis. Es ist aufgrund der Multilinearität der Determinantenfunktion

$$\mathrm{Det}(A) = \mathrm{Det}(\sum_{i=1}^{n} a_{i,1}\, \underline{e}_i, \ldots, \sum_{i=1}^{n} a_{i,n}\, \underline{e}_i) =$$
$$\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_n=1}^{n} a_{i_1,1}\cdots a_{i_n,n} \cdot \mathrm{Det}(\underline{e}_{i_1}, \ldots, \underline{e}_{i_n}) =$$
$$\sum_{\underline{i}} a_{i_1,1}\cdots a_{i_n,n} \cdot \mathrm{Det}(\underline{e}_{i_1}, \ldots, \underline{e}_{i_n}) \ ,$$

wobei $\underline{i}$ alle Tupel in $\{1,\ldots,n\}^n$, d.h. alle Abbildungen von $\{1,\ldots,n\}$ nach $\{1,\ldots,n\}$ durchläuft. (Man nennt $\underline{i}$ dann einen *Multiindex* auf der Indexmenge $\{1,\ldots,n\}$ mit Werten in $\{1,\ldots,n\}$.) Wenn nun $\underline{i}$ keine Bijektion ist, ist $\mathrm{Det}(\underline{e}_{i_1}, \ldots, \underline{e}_{i_n}) = 0$. Wir haben also

$$\begin{aligned}\mathrm{Det}(A) &= \sum_{\sigma\in S_n} a_{\sigma(1),1}\cdots a_{\sigma(n),n} \cdot \mathrm{Det}(\underline{e}_{\sigma(1)}, \ldots, \underline{e}_{\sigma(n)}) \\ &= \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) \cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} \ .\end{aligned}$$

Wenn man nun beachtet, dass $\mathrm{Det}(A) = \mathrm{Det}(A^t)$, erhält man die gewünschte Formel. □

Bemerkung Diese Formel hat ihren Wert in theoretischen Betrachtungen. Für die algorithmische Berechnung der Determinante ist sie aber katastrophal: Mittels dieser Formel benötigt man $(n-1)\cdot n!$ Mltiplikationen in K und $n!-1$ Additionen in K. Wenn man mit elementaren Spalten- und Zeilenumformungen rechnet, benötigt man nur $\leq C\cdot n^3$ "Körperoperationen" für eine Konstante $C > 0$, genau wie beim Gauß-Algorithmus.

Die adjunkte Matrix und die Cramersche Regel

Definition Für $A \in K^{n\times n}$ definieren wir die *adjunkte Matrix* als

$$A^{\#} := ((-1)^{i+j}\,(\mathrm{Det}(A_{j,i})))_{i,j=1,\ldots n} \ .$$

(Beachten Sie die Rolle der Indices!)

Sei nun $A \in K^{n\times n}$ und $\underline{b} \in K^n$, und sei $\underline{c} := A^{\#}\underline{b}$. Dann ist

$$c_i = \sum_{j=1}^{n} (-1)^{i+j} \operatorname{Det}(A_{j,i})\, b_j = \operatorname{Det}(\underline{a}_1, \ldots, \underline{a}_{i-1}, \underline{b}, \underline{a}_{i+1}, \cdots, \underline{a}_n)\ . \tag{2.23}$$

aufgrund der Formel für die Entwicklung nach Spalten. Also: Man erhält c_i, indem man *die i-te Spalte von A durch $\underline{b}$ ersetzt und die Determinante dieser Matrix bildet.* (Beachten Sie wieder die unkonventionelle Rolle des Index i!)

Somit ist insbesondere $A^{\#}\underline{a}_j = \operatorname{Det}(A)\,\underline{e}_j$. (Wenn man die i-te Spalte von A durch $\underline{a}_j$ ersetzt und dann die Determinante bildet, erhält man $\operatorname{Det}(A)\,\delta_{i,j}$.)

Damit gilt

$$A^{\#}A = \operatorname{Det}(A)\, I_n\ . \tag{2.24}$$

Da A beliebig war, gilt auch $(A^t)^{\#}A^t = \operatorname{Det}(A^t)\, I_n = \operatorname{Det}(A)\, I_n$. Nun ist $(A^t)^{\#} = ((-1)^{i+j}\,(\operatorname{Det}(A_{i,j})))_{i,j=1,\ldots n} = (A^{\#})^t$.

Damit folgt:

$$(A\,A^{\#})^t = (A^{\#})^t A^t = (A^t)^{\#}A^t = \operatorname{Det}(A)\, I_n$$

Wenn man nochmal transponiert, erhält man

$$A\,A^{\#} = \operatorname{Det}(A)\, I_n\,.$$

Wenn A invertierbar ist, kann man dies natürlich auch durch

$$A^{-1} = \frac{1}{\operatorname{Det}(A)} \cdot A^{\#}$$

ausdrücken.

Wir geben uns nun eine invertierbare Matrix $A \in K^{n\times n}$ und einen Vektor $\underline{b} \in K^n$ vor. Dann wissen wir, dass das LGS $A \cdot \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$ genau eine Lösung hat; sei diese $\underline{x}$. Wenn wir beide Seiten von (2.24) von rechts mit $\underline{x}$ multiplizieren erhalten wir $A^{\#}\,\underline{b} = \operatorname{Det}(A)\,\underline{x}$, also

$$\underline{x} = \frac{1}{\operatorname{Det}(A)} \cdot A^{\#}\,\underline{b}\,.$$

Wenn wir (2.23) beachten, erhalten wir:

$$x_i = \frac{\operatorname{Det}(\underline{a}_i, \cdots, \underline{a}_{i-1}, \underline{b}, \underline{a}_{i+1}, \cdots, \underline{a}_n)}{\operatorname{Det}(A)} \tag{2.25}$$

Dies ist die so genannte *Cramersche Regel.*

Algorithmisch ist diese Formel aber kein Fortschritt gegenüber dem Gauß-Algorithmus.

Die Determinante eines Endomorphismus

Wir geben uns jetzt einen endlichen erzeugten K-Vektorraum V und einen Endomorphismus $\varphi : V \longrightarrow V$ vor. Wir wollen die *Determinante von* φ definieren.

Hierzu wählen wir uns (irgendwie) eine Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V und betrachten die Determinante der Abbildungsmatrix $M_{\mathfrak{B}}(\varphi)$ von φ bzgl. $\mathfrak{B}$.

Ich behaupte, dass diese Determinante nicht von der Wahl der Basis abhängt.

Sei hierzu $\mathfrak{B}'$ eine andere Basis von V, und sei S die Koordinatenmatrix von $\mathfrak{B}'$ bezüglich $\mathfrak{B}$. Dann gilt nach (2.20) und Aussage 2.59:

$$\begin{aligned}\mathrm{Det}(M_{\mathfrak{B}'}(\varphi)) &= \mathrm{Det}(S^{-1}\, M_{\mathfrak{B}}(\varphi)\, S) = \\ &\mathrm{Det}(S^{-1}) \cdot \mathrm{Det}(M_{\mathfrak{B}}(\varphi)) \cdot \mathrm{Det}(S) = \mathrm{Det}(M_{\mathfrak{B}}(\varphi))\end{aligned}$$

Damit können wir definieren:

Definition Die *Determinante* von φ ist die Determinante der Abbildungsmatrix von φ bezüglich irgendeiner Basis von V. Bezeichnung: $\mathrm{Det}(\varphi)$.

2.11 Eigenwerte und Diagonalisierbarkeit

Eigenwerte und Diagonalisierbarkeit von Eindomorphismen

Wir fixieren einen endlich erzeugten K-Vektorraum V der Dimension n. Sei ferner $\varphi : V \longrightarrow V$ ein Endomorphismus von V.

Es liegt nun nahe, zu fragen, ob es eine Basis von V gibt, so dass die Abbildungmatrix von φ bezüglich dieser Basis besonders "einfach" oder "schön" ist. Besonders einfach sind die Matrizen der Form $a\, I_n$ mit $a \in K$, aber wenn φ bezüglich irgendeiner Basis die Abbildungsmatrix $a\, I_n$ hat, dann ist $\varphi = a \cdot \mathrm{id}_V$. Nach diesen Matrizen sind die Diagonalmatrizen besonders einfach. Dies motiviert:

Definition Der Endomorphismus φ ist *diagonalisierbar*, falls eine Basis $\mathfrak{B}$ von V existiert, so dass die Abbildungsmatrix von φ bzgl. der Basis $\mathfrak{B}$ eine Diagonalmatrix ist.

Somit ist φ genau dann diagonalisierbar, wenn es eine Bais $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V sowie Skalare $a_1, \ldots, a_n \in K$ mit $\varphi(\mathfrak{b}_i) = a_i \cdot \varphi(\mathfrak{b}_i)$ für alle $i = 1, \ldots, n$ gibt.

Definition Ein *Eigenwert* von φ ist ein Skalar $a \in K$, so dass es einen Vektor $\mathfrak{v} \in V$ mit $\mathfrak{v} \neq \mathfrak{o}$ und $\varphi(\mathfrak{v}) = a \cdot \mathfrak{v}$ gibt.

Wenn nun a ein Eigenwert von φ ist, dann ist jeder Vektor $\mathfrak{v} \neq \mathfrak{o}$ mit $\varphi(\mathfrak{v}) = a \cdot \mathfrak{v}$ ein *Eigenvektor* von φ zum Eigenwert a.

Somit ist φ also genau dann diagonalisierbar, wenn eine Basis von V aus Eigenvektoren von φ existiert.

Aussage 2.64 *Seien $\mathfrak{v}_1, \ldots, \mathfrak{v}_k$ Eigenvektoren von φ zu paarweise verschiedenen Eigenwerten. Dann sind die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_k$ linear unahbhängig.*

Beweis. Wir beweisen die Aussage per Induktion nach k. Der Induktionsanfang $k = 1$ (oder auch $k = 0$) ist trivial.

Wir schließen nun von k auf $k + 1$:

Seien also $\mathfrak{v}_1, \ldots, \mathfrak{v}_{k+1}$ Eigenvektoren von φ zu den Eigenwerten $a_1, \ldots, a_{k+1}$, welche paarweise verschieden sind. Seien $c_1, \ldots, c_{k+1} \in K$ beliebig mit $c_1\mathfrak{v}_1 + \cdots + c_{k+1}\mathfrak{v}_{k+1} = \mathfrak{o}$. Wenn wir hierauf φ anwenden, erhalten wir:

$$\mathfrak{o} = c_1\varphi(\mathfrak{v}_1) + \cdots + c_{k+1}\varphi(\mathfrak{v}_{k+1}) = c_1a_1\mathfrak{v}_1 + \cdots + c_{k+1}a_{k+1}\mathfrak{v}_{k+1}$$

Andererseits ist

$$\mathfrak{o} = a_{k+1}c_1\mathfrak{v}_1 + \cdots + a_{k+1}c_{k+1}\mathfrak{v}_{k+1} \ .$$

Dies ergibt:

$$(a_1 - a_{k+1})\, c_1\mathfrak{v}_1 + \cdots + (a_k - a_{k+1})\, c_k\mathfrak{v}_k = \mathfrak{o}$$

Da nach Induktionsvoraussetzung $\mathfrak{v}_1, \ldots, \mathfrak{v}_k$ linear unabhängig sind, folgt $(a_1 - a_{k+1})\, c_1 = 0, \ldots, (a_k - a_{k+1})\, c_k = 0$. Da die Eigenwerte $a_1, \ldots, a_k$ paarweise verschieden sind, ist somit $c_1 = \cdots = c_k = 0$. Aufgrund der ursprünglichen Gleichung ist dann auch $c_{k+1} = 0$. □

Hieraus folgt:

Aussage 2.65 *Der Endomorphismus φ hat höchstens n Eigenwerte. Wenn φ n verschiedene Eigenwerte hat, dann ist φ diagonalisierbar.*

Achtung Es kann sein, dass φ weniger als n Eigenwerte hat und trotzdem diagonalisierbar ist. Zum Beispiel ist $a \cdot \mathrm{id}_V$ für ein $a \in K$ immer diagonalisierbar.

Aussage 2.66 *Sei $a \in K$. Dann ist die Menge*

$$\{\mathfrak{v} \in V \mid \varphi(\mathfrak{v}) = a \cdot \mathfrak{v}\}$$

ein Untervektorraum von V.

Beweis. Offensichtlich ist $\mathfrak{o} \in V$. Seien nun $\mathfrak{v}, \mathfrak{w} \in V$ und $c \in K$ mit $\varphi(\mathfrak{v}) = a \cdot \mathfrak{v}, \varphi(\mathfrak{w}) = a \cdot \mathfrak{w}$. Dann ist $\varphi(c \cdot \mathfrak{v} + \mathfrak{w}) = c \cdot \varphi(\mathfrak{v}) + \varphi(\mathfrak{w}) = c \cdot a \cdot \mathfrak{v} + a \cdot \mathfrak{w} = a \cdot (c \cdot \mathfrak{v} + \mathfrak{w})$. □

Definition Sei a ein Eigenwert von φ. Dann ist

$$E_a := \{\mathfrak{v} \in V \mid \varphi(\mathfrak{v}) = a \cdot \mathfrak{v}\}$$

der *Eigenraum* zum Eigenwert a.

Der Eigenraum E_a besteht also aus $\mathfrak{o}$ und allen Eigenvektoren zu a.

Bemerkung Es macht Sinn, die obige Definition von E_a auf alle $a \in K$ auszuweiten. Dann ist E_a also genau dann nicht-trivial (d.h. $\neq \{\mathfrak{o}\}$), wenn a ein Eigenwert ist.

Definition und Bemerkung Seien $U_1, \ldots, U_k$ Untervektorräume von V. Dann ist die *Summe* von $U_1, \ldots, U_k$

$$U_1 + \cdots + U_k := \{\mathfrak{u}_1 + \cdots + \mathfrak{u}_k \mid \mathfrak{u}_1 \in U_1, \ldots, \mathfrak{u}_k \in U_k\} .$$

Man sieht leicht, dass dies das Erzeugnis der Teilmenge $U_1 \cup \cdots \cup U_k$ von V ist, und dass $\mathrm{Dim}(U_1 + \cdots + U_k) \leq \mathrm{Dim}(U_1) + \cdots + \mathrm{Dim}(U_k)$ gilt.

Das folgende Lemma ist nicht schwer:

Lemma 2.67 *Seien wieder $U_1, \ldots, U_k$ Untervektorräume von V. Wir fixieren für jedes $i = 1, \ldots, k$ eine Basis $\mathfrak{b}_{i,1}, \ldots, \mathfrak{b}_{i,\mathrm{Dim}(U_i)}$ von U_i. Dann sind die folgenden Aussagen äquivalent:*

a) Für alle $\mathfrak{v} \in U_1 + \cdots + U_k$ gibt es eindeutig bestimmte *$\mathfrak{u}_1 \in U_1, \ldots, \mathfrak{u}_k \in U_k$ mit $\mathfrak{v} = \mathfrak{u}_1 + \cdots + \mathfrak{u}_k$ gibt.*

b) Für alle $i = 1, \ldots, k$ ist $U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_k) = \{\mathfrak{o}\}$

c) Für alle $\mathfrak{u}_1 \in U_1, \ldots, \mathfrak{u}_k \in U_k$ mit $\mathfrak{u}_1 \neq \mathfrak{o}, \ldots \mathfrak{u}_k \neq \mathfrak{o}$ sind die Vektoren $\mathfrak{u}_1, \ldots, \mathfrak{u}_k$ linear unabhängig.

d) $\mathfrak{b}_{1,1}, \ldots, \mathfrak{b}_{1,\mathrm{Dim}(U_1)}, \ldots, \mathfrak{b}_{k,1}, \ldots, \mathfrak{b}_{k,\mathrm{Dim}(U_k)}$ ist eine Basis von $U_1 + \cdots + U_k$.

e) $\mathrm{Dim}(U_1 + \cdots + U_k) = \mathrm{Dim}(U_1) + \cdots + \mathrm{Dim}(U_k)$.

(Die ersten drei Aussagen sind auch in nicht endlich erzeugten Vektorräumen äquivalent.)

Definition Wir sagen nun, dass die Summe *direkt* ist, falls die Bedingungen im obigen Lemma erfüllt sind.

Aussage 2.68 *Die Eigenräume von φ bilden eine direkte Summe in V. Ferner ist V genau dann eine (direkte) Summe der Eigenräume von φ, wenn φ diagonalisierbar ist.*

Beweis. Nach Äquivalenz c) im obigen Lemma ist die erste Aussage eine Umformulierung von Aussage 2.64.

Sei für die zweite Aussage zunächst V die Summe der Eigenräume. Dann hat nach Äquivalenz d) im obigen Lemma V eine Basis aus Eigenvektoren, also ist φ diagonalisierbar.

Sei nun φ diagonalisiserbar. Wir fixieren eine Basis aus Eigenvektoren und sortieren diese nach verschiedenen Eigenwerten: Seien $a_1, \ldots, a_k$ die verschiedenen Eigenwerte und sei j_i die Anzahl der Basiselemente zum Eigenwert a_i. Dann ist also $j_1 + \cdots + j_k = n$ und außerdem $j_i \leq \mathrm{Dim}(E_{a_i})$ für alle $i = 1, \ldots, k$ nach der Definition von j_i. Ferner ist $\mathrm{Dim}(E_{a_1}) + \cdots + \mathrm{Dim}(E_{a_k}) \leq n$, weil die Summe der Eigenräume direkt ist. Hieraus folgt $j_i = \mathrm{Dim}(E_{a_i})$ für alle $i = 1, \ldots, k$. □

Eigenwerte und Diagonalisierbarkeit von Matrizen

Definition Seien $A, B \in K^{n\times n}$. Dann sind A und B *ähnlich*, falls es eine invertierbare Matrix $S \in K^{n\times n}$ gibt, so dass $S^{-1}AS = B$.

Beachten Sie: Wenn $\varphi : V \longrightarrow V$ ein Endomorphismus ist, dann sind Abbildungsmatrizen von φ zu verschiedenen Basen von V ähnlich.

Man kann dies auch direkt auf Matrizen anwenden: Sei $A \in K^{n\times n}$. Nun ist eine Matrix $S \in K^{n\times n}$ genau dann invertierbar, wenn die Spalten von S eine Basis bilden. Wenn nun so eine Matrix $S = (\underline{s}_1 | \cdots | \underline{s}_n)$ gegeben ist, dann ist $S^{-1}AS$ die Abbildungsmatrix von Λ_A bzgl. $\underline{s}_1, \ldots, \underline{s}_n$.

Somit sind zwei Matrizen $A, B \in K^{n\times n}$ genau dann ähnlich, wenn es eine Basis von K^n gibt, so dass B die Abbildungsmatrix von Λ_A bzgl. dieser Basis ist.

Hieraus folgt:

Bemerkung und Definition Sei A eine $n \times n$-Matrix über K. Dann sind äquivalent:

a) Λ_A ist diagonalisierbar.

b) A ist ähnlich zu einer Diagonalmatrix.

Falls dies der Fall ist, sagen wir, dass A *diagonalisierbar* ist.

Die Begriffe des Eigenwerte, Eigenvektors und Eigenraums übertragen sich nun sofort auf A. Konrekt heißt das:

Definition Sei $A \in K^{n\times n}$.

- Sei $a \in K$ und $\underline{v} \neq \underline{0}$. Dann ist $\underline{v}$ ein *Eigenvektor* zum *Eigenwert* a, falls $A\underline{v} = a \cdot \underline{v}$.
- Sei a ein Eigenwert von A. Dann ist der *Eigenraum* von a

$$E_a := \{\underline{v} \in K^{n\times n} \mid A\underline{v} = a\underline{v}\} .$$

Ein Skalar a ist also genau dann ein Eigenwert von A, wenn $\mathrm{Rang}(A - aI_n) < n$ ist. Die obigen Aussagen über Endomorphismen übertragen sich dann auch sofort auf Matrizen.

Man sieht auch: Sei $\underline{s}_1, \ldots, \underline{s}_n$ eine Basis von K^n und sei $S := (\underline{s}_1, \ldots, \underline{s}_n)$. Dann ist $S^{-1}AS$ genau dann eine Diagonalmatrix, wenn S eine Basis von Eigenvektoren ist. Denn: Sei $B := S^{-1}AS$. Dann ist $\underline{b}_i$ der Koordinatenvektor von $A\underline{s}_i$ bzgl. der Basis $\underline{s}_1, \ldots, \underline{s}_n$. Damit ist $\underline{b}_i$ genau dann ein Vielfaches von $\underline{e}_i$, wenn $\underline{s}_i$ ein Eigenvektor ist. (Alternativ kann man auch die Identität $SB = AS$ betrachten.)

Das charakteristische Polynom

Sei weiterhin A eine $n \times n$-Matrix über K. Wir haben eben gesehen: Ein Skalar $a \in K$ ist genau dann ein Eigenwert von A, wenn $\mathrm{Rang}(A - aI_n) < n$. Dies bedeutet natürlich, dass $\mathrm{Det}(A - aI_n) = 0$.

Es ist somit naheliegend, das Polynom $\mathrm{Det}(A - TI_n) \in K[T]$ zu betrachten und nach den Nullstellen dieses Polynoms zu suchen, um die Eigenwerte zu bestimmen.

Es gibt hier allerdings ein Problem: Was soll eigentlich $\mathrm{Det}(A - T \cdot I_n) \in K[T]$ sein? Schließlich haben wir nur Determinanten von Matrizen über Körpern definiert.

Es gibt zwei Möglichkeiten, Determinanten von Matrizen in $K[T]$ zu definieren:

1. Möglichkeit Der Polynomring $K[T]$ ist ein kommutativer Ring. Sei R allgemeiner ein kommultativer Ring. Dann bilden die $n \times n$-Matrizen über R "wie gewohnt" einen Ring. Wir definieren nun für eine Matrix $M \in R^{n\times n}$:

$$\mathrm{Det}(M) := \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) \cdot m_{1,\sigma(1)} \cdots m_{n,\sigma(n)} .$$

Offensichtlich verallgemeinert dies unsere bisherige Definition von Determinanten über Körpern.

2. Möglichkeit Wir erinnern uns, dass man mittels eines formalen Prozesses den Körper $\mathbb{Q}$ aus dem Körper $\mathbb{Z}$ erhalten kann (siehe §1.7). Analog kann man auch Körper definieren, dessen Elemente Brüche von Polynome in $K[T]$ sind. Diesen Körper bezeichnet man mit $K(T)$, die Elemente nennt man *rationale Funktionen* über K. (Das ist nur ein Name; es sind keine Funktionen.) Die Elemente von $K(T)$ haben also die Form $\frac{a(T)}{b(T)}$, wobei $a(T), b(T) \in K[T]$ mit $g(T) \neq 0$ sind. Hierbei ist (per Definition) für $a(T), b(T), c(T), d(T) \in K[t]$: $\frac{a(T)}{b(T)} = \frac{c(T)}{d(T)}$ genau dann, wenn $a(T)d(T) = b(T)c(T)$.

Nun haben wir für alle Matrizen über $K(T)$ eine Determinante, also insbesondere auch für Matrizen über $K[T]$.

Da wir ja die Leibniz-Formel über beliebigen Körpern bewiesen haben, stimmen diese beiden Definitionen der Determinante einer Matrix über $K[T]$ überein.

Definition Das *charakteristische Polynom* von A ist

$$\chi_A = \chi_A(T) := \mathrm{Det}(A - T \cdot I_n) \in K[T] \ .$$

Explizit ist also

$$\begin{aligned} &\mathrm{Det}(A - T \cdot I_n) \\ &= \textstyle\sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \cdot (a_{1,\sigma(1)} - T \cdot \delta_{1,\sigma(1)}) \cdots (a_{n,\sigma(n)} - T \cdot \delta_{n,\sigma(n)}) \ . \end{aligned} \tag{2.26}$$

Diese Formel hat einige einfache Konsequenzen:

Lemma 2.69

a) Es ist $\chi_A = (-1)^n \cdot T^n + c_{n-1}T^{n-1} + \cdots + c_1 T + c_0$ *mit gewissen Koeffizienten* $c_0, \ldots, c_{n-1} \in K$. *Hierbei ist* $c_0 = \mathrm{Det}(A)$ *und* $c_{n-1} = (-1)^{n-1} \cdot (a_{1,1} + \cdots + a_{n,n}) =: \mathrm{Spur}(A)$.

b) Für $a \in K$ *ist* $\chi_A(a) = \mathrm{Det}(A - a \cdot I_n)$.

Beweis. Die erste Aussage erhält man, wenn man die obige Formel "nach T entwickelt". Die zweite Aussage erhält man, wenn man a "einsetzt" (T durch a ersetzt). □

Somit haben wir:

Aussage 2.70 *Die Eigenwerte von A sind genau die Nullstellen von χ_A.*

Ferner:

Aussage 2.71 *Ähnliche Matrizen haben dasselbe charakteristische Polynom.*

Beweis. Seien A und A' ähnlich. Dann existiert also eine invertierbare Matrix $S \in K^{n\times n}$ mit $A' = S^{-1}AS$. Nun ist

$$\begin{aligned}\chi_{A'} &= \mathrm{Det}(A' - T \cdot I_n) = \mathrm{Det}(S^{-1}AS - T \cdot I_n)\\ &= \mathrm{Det}(S^{-1} \cdot (A - T \cdot I_n) \cdot S) = \mathrm{Det}(S^{-1}) \cdot \mathrm{Det}(A - T \cdot I_n) \cdot \mathrm{Det}(S)\\ &= \chi_A \,.\end{aligned}$$

□

Bemerkung Hier einige Tipps zum Berechnen des charakteristischen Polynoms: Bei 2×2 und 3×3-Matrizen benutzt man am besten "Entwicklung". Bei einer größeren Matrix mit einige Nullen kann man versuchen, geschickt zu entwickeln. Ansonsten kann man den Gauß-Algorithmus benutzen. Beachten Sie: Wie oben beschrieben, berechnen wir die Determinante einer Matrix über dem Körper $K(T)$. Es ist nun kein Problem, z.B. durch $T - a$ für $a \in K$ zu teilen, denn dies ist auch nur ein ganz bestimmtes Körperelement in $K(T)$. Man muss hier also keine Fallunterscheidung machen, ob T gleich a sei oder nicht. (T ist per Definition eben nicht a.)

Beispiel 2.72 Das charakteristische Polynom einer Diagonalmatrix mit Diagonaleinträgen $d_1, \ldots, d_n$ ist $\prod_{i=1}^{n}(d_i - T)$. Die Nullstellen hiervon sind $d_1, \ldots, d_n$, und das sind offensichtlich die Eigenwerte der Matrix.

Sei allgemeiner $A = ((a_{i,j}))_{i,j} \in K^{n\times n}$ eine obere Dreiecksmatrix (d.h. für alle $i > j$ ist $a_{i,j} = 0$). Dann ist $\chi_A = \prod_{i=1}^{n}(a_{i,i} - T)$. Die Eigenwerte sind demnach $a_{1,1}, \ldots, a_{n,n}$. (Dies kann man wieder leicht direkt sehen.)

Wir können nun zurück zu Endomorphismen gehen. Wir definieren:

Definition Sei $\varphi : V \longrightarrow V$ ein Endomorphismus. Sei A die Abbildungsmatrix von φ bezüglich irgendeiner Basis von V. Dann ist das charakteristische Polynom von φ $\chi_\varphi := \mathrm{Det}(A - T \cdot I_n)$. Dies ist wohldefiniert, denn zwei Abbildungsmatrizen desselben Endomorphismus sind ähnlich und deren charakteristischen Polynome sind gleich nach Aussage 2.71.

Trigonalisierbarkeit und der Satz von Cayley-Hamilton

Definition Sei $U \leq V$ ein Untervektorraum und $\varphi : V \longrightarrow V$ ein Endomorphismus. Dann heißt U *φ-invariant*, falls $\varphi(U) \subseteq U$.

Wenn nun U φ-invariant ist, dann haben wir den Endomorphismus φ_U von U. Ferner haben wir den Endomorphismus $\overline{\varphi} : V/U \longrightarrow V/U$ $[\mathfrak{v}]_U \mapsto [\varphi(\mathfrak{v})]_U$. (Dies ist wohldefiniert, weil U φ-invariant ist.)

Das folgende Lemma ist offensichtlich:

Lemma 2.73 *Sei $\varphi : V \longrightarrow V$ ein Endomorphismus und $U \leq V$ φ-invariant. Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Basis von V, wobei $\mathfrak{b}_1, \ldots, \mathfrak{b}_k$ eine Basis von U ist. Somit ist also $[\mathfrak{b}_{k+1}]_U, \ldots, [\mathfrak{b}_n]_U$ eine Basis von V/U.*

Sei A die Abbildungsmatrix von φ_U bezüglich $\mathfrak{b}_1, \ldots, \mathfrak{b}_k$ und D die Abbildungsmatrix von $\overline{\varphi}$ bezüglich $[\mathfrak{b}_{k+1}]_U, \ldots, [\mathfrak{b}_n]_U$. Dann hat die Abbildungsmatrix von φ bezüglich $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ die Form

$$\begin{pmatrix} A & B \\ O & D \end{pmatrix}$$

mit einer $k \times n-k$-Matrix B.

Die Determinante einer Blockmatrix $\begin{pmatrix} A & B \\ O & D \end{pmatrix}$ ist $\mathrm{Det}(A) \cdot \mathrm{Det}(D)$ (Übungsaufgabe).

Wenn man dies auf die Matrix $\begin{pmatrix} A - TI_k & B \\ O & D - TI_{(n-k)\times(n-k)} \end{pmatrix}$ anwendet, erhält man:

$$\chi_\varphi = \chi_{\varphi_U} \cdot \chi_{\overline{\varphi}} \,. \tag{2.27}$$

Wir benötigen nun eine einfache Aussage über Polynome:

Aussage 2.74 *Sei $f \in K[T], f \neq 0$. Dann gibt es paarweise verschiedene $a_1, \ldots, a_k \in K$ ($k \in \mathbb{N}_0$) sowie $e_1, \ldots, e_k \in \mathbb{N}$ und ein Polynom $g \in K[T]$ ohne Nullstellen in K mit*

$$f(T) = (\prod_{i=1}^{k} (T - a_i)^{e_i}) \cdot g(T) \,.$$

Hierbei gilt für alle $i = 1, \ldots, k$: $e_i := \max\{e \in \mathbb{N} \mid (T-a)^e \text{ teilt } f(T)\}$.

Beweis. Die Darstellung folgt aus der folgenden Aussage per Induktion: Sei $f \in K[T], f \neq 0$ sowie $a \in K$ mit $f(a) = 0$. Dann teilt $T - a$ das Polynom f. Dies folgt aus der Polynomdivision.

Wir fixieren nun ein $i = 1, \ldots, k$ und setzen $m := \max\{e \in \mathbb{N} \mid (T - a)^e$ teilt $f(T)\}$. Dann gibt es also ein $h \in K[T]$ mit $f(T) = (T-a)^m \cdot h(T)$ und $h(a) \neq 0$. Ferner haben wir nach der obigen Gleichung ein Polynom $\ell(T) \in K[T]$ mit $f(T) = (T - a_1)^{e_i} \cdot \ell(T)$ und $\ell(a) \neq 0$. Wenn man dies von einander subtrahiert und $(T - a_i)^{e_i}$ ausklammert, folgt: $h(T) = (T - a_i)^{m-e_i} \cdot \ell(T)$. Wenn man a einsetzt, folgt $e_i = m$. □

Aussage 2.75 *Sei $\varphi : V \longrightarrow V$ ein Endomorphismus. Seien $a_1, \ldots, a_k \in K$ die Eigenwerte von φ. Sei*

$$\chi_\varphi = (T - a_1)^{n_1} \cdots (T - a_k)^{n_k} \cdot g(T) \ ,$$

wobei $n_i \in \mathbb{N}$ ist und $g(T)$ keine Nullstellen in K hat. Dann gilt $\mathrm{Dim}(E_{a_i}) \leq n_i$. Ferner ist der Endomorphismus genau dann diagonalisierbar, wenn χ_φ in Linearfaktoren zerfällt und für alle $i = 1, \ldots, k$ $\mathrm{Dim}(E_{a_i}) = n_i$ gilt.

Beweis. Der Eigenraum E_{a_i} ist φ-invariant. Wie oben ausgeführt teilt $\chi_{\varphi|_{E_{a_i}}}$ das Polynom χ_φ. Es ist aber $\chi_{\varphi|_{E_{a_i}}} = (a_i - T)^{\mathrm{Dim}(E_{a_i})}$. Dies zeigt die Ungleichung.

Es ist $\sum_i n_i \leq n$. Die angegebene Bedingung ist nun äquivalent zu $\sum_i \mathrm{Dim}(E_{a_i}) = n$. Und dies bedeutet gerade, dass φ diagonalisierbar ist. □

Wir haben noch eine schöne Anwendung:

Definition Sei $\varphi : V \longrightarrow V$ ein Endomorphismus. Dann heißt φ *trigonalisierbar*, wenn es eine Basis von V gibt, so dass die Abbildungsmatrix von φ bzgl. dieser Basis eine obere Dreiecksmatrix ist. (Analog heißt eine $n \times n$-Matrix über K trigonalisierbar, wenn sie ähnlich zu einer oberen Dreiecksmatrix ist.)

Satz 2.8 (Trigonalisierbarkeit) *Sei $\varphi : V \longrightarrow V$ ein Endomorphismus. Dann ist φ genau dann trigonalisierbar, wenn das charakteristische Polynom von φ in Linearfaktoren zerfällt.*

Beweis. Da das charakteristische Polynom einer oberen Dreiecksmatrix in Linearfaktoren zerfällt, ist die Notwendigkeit der Bedingung klar.

Wir zeigen die Rückrichtung per Induktion über die Dimension.

Der Induktionsanfang ist trivial. Es zerfalle nun das charakteristische Polynom von φ in Linearfaktoren. Dann hat χ_φ insbesondere eine Nullstelle und somit hat φ einen Eigenwert. Sei a so ein Eigenwert. Dann ist der Raum E_a φ-invariant; sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_k$ eine Basis von E_a. Nun ist $\overline{\varphi} \in \mathrm{End}_K(V/E_a)$ trigonalisierbar. Sei $[\mathfrak{b}_{k+1}]_U, \ldots, [\mathfrak{b}_n]_U$ eine Basis von V/E_a mit $\mathfrak{b}_{k+1}, \ldots, \mathfrak{b}_n \in V$, so dass $\overline{\varphi}$ bzgl. dieser Basis die Abbildungsmatrix B hat, welche eine obere Dreicksmatrix ist.. Dann ist $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Basis von V und nach dem obigen Lemma ist die Abbildungsmatrix bzgl. dieser Basis auch eine obere Dreicksmatrix. □

Korollar 2.76 *Sei K algebraisch abgeschlossen, z.B. $K = \mathbb{C}$. Dann ist jede Matrix über K trigonalisierbar.*

Wenn $f(T) = \sum_{i=0}^{n} a_i T^i \in K[T]$ ein beliebiges Polynom über K ist, kann man φ in f "einsetzen". Man erhält dann den Endomorphismus $\sum_{i=0}^{n} a_i \varphi^i$. Hier ist (wie immer) $\varphi^0 := \mathrm{id}_V$. Man sieht sofort: Die Abbildung

$$K[T] \longrightarrow \mathrm{End}_K(V)\ ,\ f(T) \mapsto f(\varphi) \tag{2.28}$$

ist ein Ringhomomorphismus und ein Homomorphismus von K-Vektorräumen.

Der K-Vektorraum $\mathrm{End}_K(V)$ hat die Dimension n^2. Somit sind also die Potenzen φ^i für $i = 0, \ldots, n^2$ linear abhängig. Es gibt also ein Polynom $f(T) \in K[T]$ mit $f \neq 0$ und $f(\varphi) = 0_V$.

Bemerkenswerterweise hat auch das charakteristische Polynom diese Eigenschaft. Dies ist vielleicht der berühmteste Satz der Linearen Algebra:

Satz 2.9 (Satz von Cayley-Hamilton) *Es ist $\chi_\varphi(\varphi) = 0_V$, die Nullabbildung.*

Zum Beweis. Wir zeigen den Satz zunächst unter der Voraussetzung, dass χ_φ in Linearfaktoren zerfällt. Am Ende des Beweises gehen wir kurz darauf ein, wie man dann auch den Satz auch in voller Allgemeinheit erhalten kann. Wir werden später den Satz mit einer anderen Methode beweisen.

Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Basis von V, so dass die Abbildungsmatrix von φ bzgl. dieser Basis eine obere Dreiecksmatrix ist. Sei A die entsprechende Abbildungsmatrix.

Nun ist also $\chi_\varphi = (a_{1,1} - T) \cdots (a_{n,n} - T)$. Sei $U_i := \langle \mathfrak{b}_1, \ldots, \mathfrak{b}_n \rangle$ und $\chi_i := (a_{1,1} - T) \cdots (a_{i,i} - T)$. Beachten Sie, dass alle Räume U_i φ-invariant sind. Damit is also $\chi_n = \chi_\varphi$. Ich behaupte, dass für alle $i = 1, \ldots, n$ $\chi_i(\varphi)|_{U_i} = 0_{U_i}$ ist.

Der Beweis erfolgt per Induktion über i:

Der Induktionsanfang $i = 1$ ist trivial.

Zum Induktionsschritt von i auf $i + 1$: Wir müssen zeigen:

$$\forall j = 1, \ldots, i+1 : (\chi_{i+1}(\varphi))(\mathfrak{b}_j) = \mathfrak{o}\ .$$

Sei $j = 1, \ldots, i + 1$. Dann ist $(\varphi - a_{i+1,a+1}\,\mathrm{id})(\mathfrak{b}_i) \in U_i$. (Für $j \leq i$ ist das so, weil U_i φ-invarant ist und f ür $j = i$ folgt das aus der Struktur der Abbildungsmatrix A.) Somit ist $\chi_{i+1}(\varphi) = (\chi_i(\varphi) \circ (\varphi - a_{i+1,a+1}\,\mathrm{id}))(\mathfrak{b}_i) = \mathfrak{o}$ nach Induktionsvoraussetzung.

Zum zum allgemeinen Fall: Zunächst kann man die Aussage auch für Matrizen definieren. Sie lautet dann wie folgt: Sei $A \in K^{n \times n}$, so dass χ_A in $K[T]$ vollständig in Linearfaktoren zerfällt. Dann ist $\chi_A(A) = O$.

Nun gibt es für jeden Körper K und jedes Polynom $f \in K[T]$ einen Körper L, der K enthält, so dass f aufgefasst als Polynom in $L[T]$ in Linearfaktoren zerfällt (**Lücke!**).

Sei nun $A \in K^{n\times n}$ eine beliebige Matriz. Dann gibt es also einen Körper L, der K enthält, so dass χ_A in $L[T]$ in Linearfaktoren zerfällt. Wir können auch A als Matrix in $L^{n\times n}$ auffassen und erhalten $\chi_A(A) = O$.

Da dies nun für alle Matrizen über K gilt, gilt auch für alle Endomorphismen φ endlich-dimensionaler K-Vektorräume: $\chi_\varphi(\varphi) = 0$. □

Kapitel 3

Bilinearformen, Euklidische und unitäre Vektorräume

3.1 Bilinearformen

Grundlegende Definitionen

Aus der Schule kennen Sie das *Standardskalarprodukt* auf dem $\mathbb{R}^n$: Für $\underline{x}, \underline{y} \in \mathbb{R}^n$ setzt man

$$\underline{x} \bullet \underline{y} := \sum_{i=1}^{n} x_i y_i = \underline{x}^t \underline{y} \ .$$

Nun ist für jedes feste $\underline{x} \in \mathbb{R}^n$ die Abbildung

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n \ , \ \underline{y} \mapsto \underline{x} \bullet \underline{y}$$

linear. Analog ist die Situation, wenn man "in der zweiten Komponente" $\underline{y}$ fixiert und $\underline{x}$ variiert. Eine Abbildung auf $\mathbb{R}^n \times \mathbb{R}^n$ mit diesen Eigenschaften heißt *Bilinearform*. Dies können wir gleich allgemeiner für beliebige Vektorräume definieren.

Wie im letzten Kapitel fixieren wir einen Körper K.

Definition Sei V ein K-Vektorraum. Eine *Bilinearform* auf V ist eine Abbildung $\beta : V \times V \longrightarrow K$, so dass gilt:

- Für alle $\mathfrak{x} \in V$ ist die Abbildung $V \longrightarrow K \ , \mathfrak{y} \mapsto \beta(\mathfrak{x}, \mathfrak{y})$ linear.
- Für alle $\mathfrak{y} \in V$ ist die Abbildung $V \longrightarrow K \ , \mathfrak{x} \mapsto \beta(\mathfrak{x}, \mathfrak{y})$ linear.

Wir fixieren einige Notationen:

Notation Bilinearformen bezeichnet man oft mit $\langle ., . \rangle$. Die Punkte deuten an, dass man hier "was einsetzen" kann. Analoge Notationen mit "Platzhaltern" sind auch ansonsten üblich. Zum Beispiel: Wenn β eine Bilinearform auf V ist und $\mathfrak{x} \in V$ ein fester Vektor ist, bezeichnet $\beta(\mathfrak{x}, .)$ die Linearform $V \longrightarrow K\ ,\ \mathfrak{y} \mapsto \beta(\mathfrak{x}, \mathfrak{y})$.

Übrigens ist es eher unüblich, Bilinearformen mit einem "Multiplikationspunkt" zu bezeichnen.

Wir verallgemeinern die obige Definition nun nochmal:

Definition Seien V, W zwei K-Vektorräume. Eine *bilineare Abbildung* oder eine *Bilinearform* zwischen V und W ist eine Abbildung $\beta : V \times W \longrightarrow K$, so dass für jedes feste $\mathfrak{v} \in V$ und jedes feste $\mathfrak{w} \in W$ die Abbildungen

$$\beta(\mathfrak{v}, .) : W \longrightarrow K \quad \text{und} \quad \beta(., \mathfrak{w}) : V \longrightarrow K$$

linear sind.

Bemerkung und Definition Man sieht leicht, dass die Bilinearformen zwischen V und W einen K-Vektorraum bilden. Diesen Vektorraum bezeichnen wir mit $\mathrm{Bil}_K(V, W)$.

Beispiel 3.1 Ein große Beispielklasse für eine Bilinearformen kennen Sie schon: Sei V ein K-Vektorraum und V^* der Dualraum von V. Dann ist die Abbildung

$$V^* \times V\ ,\ (\alpha, \mathfrak{v}) \mapsto \alpha(\mathfrak{v})$$

bilinear. Diese Abbildung heißt *Dualprodukt* oder *kanonische Bilinearform* zu V. Wir bezeichnen sie mit $\langle ., . \rangle_{\mathrm{kan}}$.

Diese Sichtweise auf die duale Abbildung mittels des Dualprodukts ist insbesondere aufgrund der folgenden Überlegung passend: Wir haben auch ein Dualprodukt für V^*. Wenn nun $\Phi : V \longrightarrow V^{**}$ die kanonische Abbildung ist, dann ist für $\mathfrak{v} \in V$ und $\alpha \in V^*$:

$$\langle \alpha, \mathfrak{v} \rangle_{\mathrm{kan},V} = \langle \Phi(\mathfrak{v}), \alpha \rangle_{\mathrm{kan},V^*}$$

Wenn nun V endlich-dimensional ist und wir V mittels Φ mit V^{**} identifizieren, haben wir also

$$\langle \alpha, \mathfrak{v} \rangle_{\mathrm{kan},V} = \langle \mathfrak{v}, \alpha \rangle_{\mathrm{kan},V^*}\ .$$

Bemerkung Das Dualprodukt wird oft als eine Abbildung von $V \times V^*$ nach K definiert. Die obige Definition erscheint mir aber "natürlicher". Denn mit der obigen Notation gilt insbesondere das Folgende (und noch weitere "schöne" Eigenschaften):

Wir identifizieren $(K^n)^*$ mit $K^{1\times n}$. Dann ist das Dualprodukt für K^n durch

$$K^{1\times n} \times K^n \ , \ (\underline{a}, \underline{x}) \mapsto \underline{a} \cdot \underline{x}$$

gegeben.

Diese Bemerkung können wir zu einer Beispielklasse ausbauen:

Beispiel 3.2 Sei $B \in K^{m\times n}$. Dann ist die Abbildung

$$K^{1\times m} \times K^n \longrightarrow K \ (\underline{a}, \underline{x}) \mapsto \underline{a}B\underline{x}$$

eine Bilinearform zwischen $K^{1\times m}$ und K^n.

Einer Bilinearform $\beta : V \times W \longrightarrow K$ kann man zwei Abbildungen

$$\beta_1 : V \longrightarrow W^* \ , \ \mathfrak{v} \mapsto \beta(\mathfrak{v}, .) = (\mathfrak{w} \mapsto \beta(\mathfrak{v}, \mathfrak{w}))$$

und

$$\beta_2 : W \longrightarrow V^* \ , \ \mathfrak{w} \mapsto \beta(., \mathfrak{w}) = (\mathfrak{v} \mapsto \beta(\mathfrak{v}, \mathfrak{w}))$$

zuordnen. Diese Abbildungen sind offensichtlich linear.

Definition Die Bilinearform β ist *nicht-ausgeartet in der ersten Variablen* oder *links-nicht-ausgeartet*, falls β_1 injektiv ist.

Da die Abbildung β_1 linear ist, bedeutet dies, dass $\mathrm{Kern}(\beta_1) = \{\mathfrak{o}_V\}$ ist.
Konkret ist

$$\mathrm{Kern}(\beta_1) = \{\mathfrak{v} \in V \mid \forall \mathfrak{w} \in W : \beta(\mathfrak{v}, \mathfrak{w}) = 0\} \ .$$

Dies nennt man auch den *Links-Kern* von β.

Somit ist β genau dann links-nicht-ausgeartet, wenn

$$\forall \mathfrak{v} \in V : (\forall \mathfrak{w} \in W : \beta(\mathfrak{v}, \mathfrak{w}) = 0 \longrightarrow \mathfrak{v} = 0)$$

gilt.

Analog definiert man nicht-Ausgeartetheit bzgl. der zweiten Variablen und den Rechts-Kern von β. Die Aussagen gelten dann auch analog.

Beispiel 3.3 Wir betrachten wieder das Dualprodukt zu einem Vektorraum V. Dieses ist offensichtlich links-nicht-ausgeartet. Ferner ist es offensichtlich genau dann rechts-nicht ausgeartet, wenn die kanonische Abbildung $\Phi : V \longrightarrow V^{**}$ injektiv ist. Wir wissen, dass dies der Fall ist, wenn V endlich-dimensional ist. Man kann dies auch allgemein zeigen, aber das kommt – wie schon bei der Definition von Φ geschrieben – später.

Wir halten das folgende Lemma fest.

Lemma 3.4 *Seien V und W endlich-dimensional und β beidseitig nicht-ausgeartet. Dann gilt:*

- $\mathrm{Dim}(V) = \mathrm{Dim}(W)$
- *β_1 und β_2 sind Isomorphismen.*

Beweis. Da β_1 und β_2 injektiv sind, ist $\mathrm{Dim}(V) \leq \mathrm{Dim}(W^*) = \mathrm{Dim}(W)$ und $\mathrm{Dim}(W) \leq \mathrm{Dim}(V) = \mathrm{Dim}(V^*)$. Dies zeigt die Aussage. □

Beispiel 3.5 Wir fahren mit Beispiel 3.2 fort: Für $\underline{a} \in K^{1\times m}$ ist $\beta_1(\underline{a})$ gleich der durch den Zeilenvektor $\underline{a}B$ gegebenen Linearform. Der Links-Kern der Bilinearform besteht ist somit $\{\underline{a} \in K^{1\times m} \mid \underline{a}B = \underline{0}^t\}$. Dies ist der so genannte *Links-Kern* von B.

Analog ist der Rechts-Kern von β gleich dem (normalen) Kern von B.

Man sieht hier explizit: Wenn β links-nicht-ausgeartet ist, ist $\mathrm{Rang}(B) = m$, wenn β rechts-nicht-ausgeartet ist, ist $\mathrm{Rang}(B) = n$. Und wenn β beidseitig nicht ausgeartet ist, ist $m = n$ und B ist invertierbar.

Wir haben soeben jeder Bilinearform zwischen V und W eine lineare Abbildung von V nach W^* zugeordnet. Diesen Prozess kann man auch umkehren:

Sei $\varphi : V \longrightarrow W^*$ eine lineare Abbildung. Dann ist die Abbildung

$$V \times W\ ,\ (\mathfrak{v}, \mathfrak{w}) \mapsto \langle \varphi(\mathfrak{v}), \mathfrak{w} \rangle_{\mathrm{kan}}$$

eine Bilinearform.

Analog kann man verfahren, wenn eine lineare Abbildung von W nach V gegeben ist.

Wir haben nun:

Lemma 3.6

a) Die Abbildung

$$\mathrm{Bil}_K(V,W) \longrightarrow \mathrm{Hom}_K(V,W^*)\ ,\ \beta \mapsto \beta_1$$

ist ein Isomorphismus von Vektorräumen. Die Umkehrabbildung ist

$$\mathrm{Hom}_K(V,W^*) \longrightarrow \mathrm{Bil}_K(V,W)\ , \varphi \mapsto \langle \varphi(.),.\rangle_K = ((\mathfrak{v},\mathfrak{w}) \mapsto \langle \varphi(\mathfrak{v}),\mathfrak{w}\rangle_{\mathrm{kan},W})\ .$$

b) Die Abbildung

$$\mathrm{Bil}_K(V,W) \longrightarrow \mathrm{Hom}_K(W,V^*)\ ,\ \beta \mapsto \beta_2$$

ist ein Isomorphismus von Vektorräumen. Die Umkehrabbildung ist

$$\mathrm{Hom}_K(W,V^*) \longrightarrow \mathrm{Bil}_K(V,W)\ , \varphi \mapsto ((\mathfrak{v},\mathfrak{w}) \mapsto \langle \varphi(\mathfrak{w}),\mathfrak{v}\rangle_{\mathrm{kan},V}).$$

(Beachten Sie hier die Vertauschung der Reihenfolge von $\mathfrak{v}$ und $\mathfrak{w}$!)

Beweis. Wir zeigen nur die erste Aussage – der Beweis der zweiten Aussage geht analog.

Sei β eine Bilinearform. Dann ist für $\mathfrak{v} \in V, \mathfrak{w} \in W$: $\langle \beta_1(\mathfrak{v}),\mathfrak{w}\rangle_{\mathrm{kan}} = \beta_1(\mathfrak{v})(\mathfrak{w}) = \beta(\mathfrak{v},\mathfrak{w})$.

Sei nun $\varphi : V \longrightarrow W^*$ linear. Sei β die durch φ definierte lineare Abbildung. Dann ist also für alle $\mathfrak{v} \in V, \mathfrak{w} \in W : \beta(\mathfrak{v},\mathfrak{w}) = \langle \varphi(\mathfrak{v}),\mathfrak{w}\rangle_{\mathrm{kan}}$. Somit ist für alle $\mathfrak{v} \in V$: $\beta(\mathfrak{v},.) = \varphi(\mathfrak{v})$. □

Bemerkung Wenn V und W endlich-dimensional sind, ist somit $\mathrm{Dim}(\mathrm{Bil}_K(V,W)) = \mathrm{Dim}(V) \cdot \mathrm{Dim}(W)$.

Die adjungierte Abbildung

Seien nun V und W wieder beliebige K-Vektorräume und β eine Bilinearform zwischen V und W.

Definition Seien $\varphi \in \mathrm{End}_K(V)$ und $\psi \in \mathrm{End}_K(W)$ mit

$$\beta(\varphi(\mathfrak{v}),\mathfrak{w}) = \beta(\mathfrak{v},\psi(\mathfrak{w}))$$

für alle $\mathfrak{v} \in V, \mathfrak{w} \in W$. Dann heißt ψ *rechts-adjungiert* zu φ und φ heißt *links-adjungiert* zu ψ. Die beiden Abbildungen heißen dann *adjungiert* zueinander. (Diese Begriffe beziehen sich natürlich auf die fest vorgegebene Bilinearform β.)

Beispiel 3.7 Wir betrachten das Dualprodukt zu V: Sei $\varphi \in \mathrm{End}_K(V)$ und $\varphi^* \in \mathrm{End}_K(V^*)$ die zugehörige duale Abbildung. Dann ist per Definition

$$\langle \varphi^*(\alpha), \mathfrak{v} \rangle_{\mathrm{kan}} = \langle \alpha, \varphi(\mathfrak{v}) \rangle_{\mathrm{kan}} \tag{3.1}$$

für alle $\alpha \in V^*$ und $\mathfrak{v} \in V$. Somit sind φ^* und φ adjungiert zueinander.

In diesen Beispiel ist φ^* auch der einzige Endomorphismus, der links-adjungiert zu φ ist. (Denn für $\alpha \in V^*$ ist (3.1) ja gerade die Definition von $\varphi^*(\alpha)$.) Dies kann man verallgemeinern.

Wir gehen wieder von einer Bilinearfrom β auf $V \times W$ aus. Beachten Sie zunächst: Wenn $\varphi \in \mathrm{End}_K(V)$ ist, dann ist $\gamma := \beta(\varphi(.), .)$ auch eine Bilinearform auf $V \times W$. Ferner ist $\gamma_1 = \beta_1 \circ \varphi$. Wenn nun β links-nicht-ausgeartet ist, dann ist also β_1 injektiv und φ durch γ_1 (d.h. durch γ) eindeutig bestimmt. Wir erhalten:

Aussage 3.8 *Sei β links-nicht-ausgeartet und γ eine weitere Bilinearform zwischen V und W. Dann gibt es höchstens einen Endomorphismus $\varphi : V \longrightarrow V$ mit $\gamma = \beta(\varphi(.), .)$, d.h. mit*

$$\gamma(\mathfrak{v}, \mathfrak{w}) = \beta(\varphi(\mathfrak{v}), \mathfrak{w})$$

für alle $\mathfrak{v} \in V$ und $\mathfrak{w} \in W$.

Wenn nun $\psi \in \mathrm{End}_K(W)$ ist, dann ist $\beta(., \psi(.))$ eine Bilinearform zwischen V und W. Wenn wir auf β und diese Bilinarform das obige Lemma anwenden, erhalten wir:

Aussage 3.9 *Sei β links-nicht-ausgeartet und $\psi \in \mathrm{End}_K(W)$. Dann gibt es höchstens einen Endomorphismus $\varphi \in \mathrm{End}_K(V)$, der links-adjungiert zu ψ ist.*

Notation Sei weiterhin β links-nicht-ausgeartet und $\psi \in \mathrm{End}_K(W)$. Wenn nun ψ einen links-adjungierten Endomorphismus bzgl. β hat, sprechen wir von der *links-Adjungierten* von ψ und bezeichnen diesen Endomorphismus mit ψ^*.

Wir betonen, dass diese Notation nur Sinn macht, wenn die Bilinearform fixiert wurde. Eine Möglichkeit für so eine Bilinearform ist das Dualprodukt zu einem Vektorraum, und in diesem Fall stimmt die Notation mit der Bezeichnung der dualen Abbildung überein.

Die beiden obigen Aussagen und die Notation machen natürlich auch Sinn, wenn man “rechts” und “links” vertauscht.

Wenn nun V und W endlich-dimensional sind und β beidseitig nicht-ausgeartet sind, existieren die adjungierten Endomorphismen immer:

Aussage 3.10 *Sei β eine beidseitig nicht-ausgeartete Bilinearform zwischen V und W. (Es ist dann $\mathrm{Dim}(V) = \mathrm{Dim}(W)$.) Sei nun γ eine weitere Bilinearform zwischen V und W. Dann gibt es genau einen Endomorphismus φ von V mit*

$$\gamma(\mathfrak{v}, \mathfrak{w}) = \beta(\varphi(\mathfrak{v})), \mathfrak{w})$$

für alle $\mathfrak{v} \in V$ und alle $\mathfrak{w} \in W$. Analog gibt es genau einen Endomorphismus ψ von W mit

$$\gamma(\mathfrak{v}, \mathfrak{w}) = \beta(\mathfrak{v}, \psi(\mathfrak{w}))$$

für alle $\mathfrak{v} \in V$ und alle $\mathfrak{w} \in W$.

Beweis. Wir zeigen nur die erste Aussage.

Wie schon gesagt erfüllt φ die angegebene Bedingung genau dann, wenn $\gamma_1 = \beta_1 \circ \varphi$ gilt. Da nun β beidseitig nicht-ausgeartet ist und V und W endlich-dimensional sind, ist β_1 nach Lemma 3.4 ein Isomorphismus. Somit gibt es genau ein solches φ, nämlich $\varphi = \beta_1^{-1} \circ \gamma_1$. □

Hieraus folgt sofort:

Aussage 3.11 *Seien V und W endlich-dimensional und sei β nicht-ausgeartet. Dann hat jeder Endomorphismus von V eine (eindeutig bestimmte) rechts-Adjungierte und jeder Endomorphismus von W eine (eindeutig bestimmte) links-Adjungierte.*

Der endlich-dimensionale Fall

Wenn die Vektorräume endlich-dimensional sind, kann man Bilinearformen genau wie lineare Abbildungen mit Matrizen darstellen.

Seien also ab sofort V und W endlich-dimensional. Sei $\mathfrak{B} = (\mathfrak{b}_1, \ldots, \mathfrak{b}_m)$ eine Basis von V und $\mathfrak{C} = (\mathfrak{c}_1, \ldots, \mathfrak{c}_n)$ eine Basis von W. Sei ferner β eine Bilinearform zwischen V und W.

Wir definieren die *Matrix* von β bezüglich $\mathfrak{B}$ und $\mathfrak{C}$ wie folgt:

$$B := ((\beta(\mathfrak{b}_i, \mathfrak{c}_j)))_{i,j} \in K^{m \times n} \tag{3.2}$$

Seien nun $\mathfrak{x} \in V$ und $\mathfrak{y} \in W$ mit Koordinatenvektoren $\underline{x}, \underline{y}$ bezüglich der angegebenen Basen. Dann ist

$$\beta(\mathfrak{x}, \mathfrak{y}) = \beta(\sum_i x_i \mathfrak{b}_i, \sum_j y_i \mathfrak{c}_j) = \sum_{i,j} x_i \beta(\mathfrak{b}_i, \mathfrak{c}_j) y_j = \sum_{i,j} x_i b_{i,j} y_j = \underline{x}^t B \underline{y} \, .$$

Man sieht hier: Unter der Koordinatenabbildung $c_{\mathfrak{C}} : W \longrightarrow K^n$ entspricht der Rechts-Kern von β (das ist der Kern von β_2) genau dem Kern

von B. Analoge Aussagen gelten bez. der Links-Kerne. Ferner ist β genau dann beidseitig nicht-ausgeartet, wenn B quadratisch und invertierbar ist.

Den Zusammenhang zwischen der Bilinearform und ihrer Matrix bzgl. der vorgegebenen Basen kann man auch schön mit Dualbasen ausdrücken: Sei $X_1, \ldots, X_m$ die Dualbasis zu $\mathfrak{B}$ und $Y_1, \ldots, Y_n$ die Dualbasis zu $\mathfrak{C}$. Seien $p_V : V \times W \longrightarrow V, p_W : V \times W \longrightarrow W$ die Projektionen. Dann sind $p_V^*(X_i) = X_i \circ p_V$ Linearformen auf $V \times W$. Diese sind explizit wie folgt gegeben: Sei $(\mathfrak{x}, \mathfrak{y}) \in V \times W$, und seien $\underline{x}, \underline{y}$ die entsprechenden Koordinatenvektoren bzgl. der angegebenen Basen. Dann bildet $X_i \circ p_V$ $(\mathfrak{x}, \mathfrak{y})$ auf x_i ab. Wir bezeichnen deshalb die Linearform $X_i \circ p_V$ wieder mit X_i. Analoge Überlegungen gelten für W und $Y_1, \ldots, Y_n$. Man sieht leicht, dass $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ ein Koordinatensystem von $V \times W$ ist.

Nun ist

$$\beta = \sum_{i,j} b_{i,j} X_i Y_j \ . \tag{3.3}$$

Beachten Sie hierbei das Folgende: X_i und Y_j sind Linearformen auf einem gemeinsamen Vektorraum. Wenn nun allgemein α_1, α_2 Linearformen auf einem K-Vektorraum X sind, dann ist $\alpha_1 \cdot \alpha_2 = \alpha_1 \alpha_2$ per Definition die Abbildung

$$X \longrightarrow K \ , \ \mathfrak{v} \mapsto \alpha_1(\mathfrak{v}) \cdot \alpha_2(\mathfrak{v}) \ .$$

Diese Definition entspricht der Definition vom Produkt zweier Funktionen in der Analysis.

Also: Linearformen kann man multiplizieren, aber man erhält dann in der Regel keine Linearform mehr.

Wie verändert ("transformiert") sich nun die Matrix der Bilinearform, wenn man die Basen wechselt?

Sei hierzu $\mathfrak{B}'$ eine weitere Basis von V und $\mathfrak{C}'$ eine weitere Basis von W. Sei S die Koordinatenmatrix von $\mathfrak{B}'$ bzgl. $\mathfrak{B}$ und T die Koordinatenmatrix von $\mathfrak{C}'$ bzgl. $\mathfrak{C}$. Sei B' die Matrix von β bzgl. den neuen Basen. Seien $\mathfrak{x} \in V, \mathfrak{y} \in W$ mit Koordinatenvektoren $\underline{x}, \underline{y}$ bzw. $\underline{x}', \underline{y}'$ bzgl. der angegebenen Basen. Dann ist also $\underline{x} = S\underline{x}'$ und $\underline{y} = T\underline{y}'$.

Nun ist einerseits

$$\beta(\mathfrak{x}, \mathfrak{y}) = (\underline{x}')^t B' \underline{y}'$$

und andererseits

$$\beta(\mathfrak{x}, \mathfrak{y}) = \underline{x}^t B \underline{y} = (S\underline{x}')^t B (T\underline{y}') = \underline{x}'^t S^t B T \underline{y}' \ .$$

Da die Vektoren beliebig waren, gilt somit:

$$B' = S^t B T \tag{3.4}$$

Wir gehen jetzt nochmal auf adjungierte Endomorphismen ein.

Seien $\varphi \in \mathrm{End}_K(V), \psi \in \mathrm{End}_K(W)$. Sei A die Abbildungsmatrix von φ (bzgl. $\mathfrak{B}$) und C die Abbildungsmatrix von ψ (bzgl. $\mathfrak{C}$). Dann ist die Matrix von $\beta(\varphi(.), .)$ gleich A^tB und die Matrix von $\beta(., \psi(.))$ gleich BC. Somit sind die beiden Endomorphismen genau dann adjungiert zueinander, wenn

$$A^tB = BC$$

gilt.

Sei nun β beidseitig nicht-ausgeartet. Dann ist also B quadratisch und invertierbar. Wenn nun φ (oder A) gegeben ist, gibt es genau eine Matrix C wie oben, nämlich $B^{-1}A^tB$. Wir sehen wieder, dass nun die rechts-Adjungierte immer existiert und eindeutig bestimmt ist. Analog ist die Situation bzgl. der links-Adjungierten.

3.2 Symmetrische Bilinearformen

Sei V ein K-Vektorraum und β eine Bilinearform auf V. Dies ist nun also eine bilineare Abbildung von $V \times V$ nach K.

Wir halten nochmal fest, dass wir nun die Homomorphismen $\beta_1 : V \longrightarrow V^*$, $\mathfrak{v} \mapsto \beta(\mathfrak{v}, .)$ und $\beta_2 : V \longrightarrow V^*$, $\mathfrak{v} \mapsto \beta(., \mathfrak{v})$ haben.

Besonders wichtig ist nun der Fall, dass β *symmetrisch* ist. Dies bedeutet, dass für alle $\mathfrak{x}, \mathfrak{y} \in V$

$$\beta(\mathfrak{x}, \mathfrak{y}) = \beta(\mathfrak{y}, \mathfrak{x})$$

ist. Dies ist übrigens äquivalent dazu, dass $\beta_1 = \beta_2$ ist.

Wenn V endlich erzeugt ist, kann man dies noch konkreter beschreiben: Sei V endlich erzeugt und $\mathfrak{B} := (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$ eine Basis von V. Dann ist die *Matrix* von β bzgl. dieser Basis gleich

$$B := ((\beta(\mathfrak{b}_i, \mathfrak{b}_j))_{i,j} \in K^{n \times n} .$$

Dies ist natürlich ein Spezialfall von (3.2), wenn man "rechts" und "links" die gleiche Basis $\mathfrak{B}$ wählt.

Es gilt nun das folgende einfache Lemma:

Lemma 3.12 *Die Bilinearform β ist genau dann symmetrisch, wenn $B^t = B$ gilt.*

Definition Eine Matrix $B \in K^{n\times n}$ mit $B^t = B$ heißt auch *symmetrisch*.

In diesem Abschnitt behandeln wir symmetrische Bilinearformen.

Eine symmetrische Bilinearform ist natürlich genau dann links-nicht-ausgeartet, wenn sie rechts-nicht-ausgeartet ist. Man sagt dann einfach, dass sie *nicht-ausgeartet* ist. Auch ist der Rechts-Kern gleich dem Links-Kern; man spricht vom *Kern* oder dem *Radikal* der Bilinearform.

Ein Beispiel für eine nicht-ausgeartete symmetrische Bilinearform ist das Standardskalarprodukt auf dem $\mathbb{R}^n$ (oder allgemeiner auf dem K^n).

An dieser Stelle ist es hilfreich, den Begriff der *Charakteristik* eines Körpers einzuführen:

Definition Die Charakteristik des Körpers K ($\mathrm{char}(K)$) ist wie folgt definiert: Wenn es eine Zahl $n \in \mathbb{N}$ mit $n \cdot 1_K = 0$ gibt, ist n die Charakteristik von K. Andernfalls ist die Charakteristik von K gleich 0.

Es ist offensichtlich, dass die Charakteristik entweder 0 oder eine Primzahl ist. Außerdem gilt: Sei $\mathrm{char}(K) > 0$. Dann ist für $n \in \mathbb{Z}$ genau dann $n\cdot 1_K = 0$, wenn $\mathrm{char}(K)|n$ gilt.

Die Theorie der symmetrischen Bilinearformen ist unterschiedlich, je nachdem ob die Charakteristik des Grundkörpers gleich 2 oder ungleich 2 ist. Wir behandeln hier nur den "allgemeinen Fall", dass die Charakteristik ungleich 2 ist.

Wir legen also fest:

Grundlegende Voraussetzung *Sei von nun an bis zum Ende dieses Abschnitts* $\mathrm{char}(K) \neq 2$.

Sei nun V ein endlich-dimensionaler K-Vektorraum und β eine symmetrische Bilinearform auf V.

Sei wieder $\mathfrak{B} := (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$ eine Basis von V und B die Matrix von β bzgl. dieser Basis.

Wir halten fest: Wenn $\mathfrak{B}'$ eine andere Basis mit Koordinatenwechselmatrix S ist und B' die entsprechende Matrix der Bilinearform ist, dann ist

$$B' = S^t B S \; . \tag{3.5}$$

Der Zusammenhang von B und B' motiviert die folgende Definition:

Definition Seien $A, B \in K^{n\times n}$. Dann ist A *kongruent* zu B, wenn es eine invertierbare Matrix $S \in K^{n\times n}$ gibt, so dass $A = S^tBS$ ist.

Offensichtlich definiert dies eine Äquivalenzrelation auf $K^{n\times n}$. Analog zu Endomorphismen wollen wir eine Basis finden, so dass die Matrix von β bzgl. dieser Basis besonders "schön" ist.

Es gilt nun der folgende Satz:

Satz 3.1 *Es gibt eine Basis von V bezüglich welcher die Matrix von β Diagonalgestalt hat.*

Dieser Satz ist offenbar äquivalent zu der folgenden Aussage:

Aussage 3.13 *Sei $A \in K^{n\times n}$ symmetrisch. Dann ist A kongruent zu einer Diagonalmatrix.*

Der *Beweis* beruht auf einem Algorithmus, der ähnlich zum Gauß-Algorithmus ist und auch recht elementar ist. Der Algorithmus liefert dann auch eine passende Matrix invertierbare Matrix S, so dass S^tAS Diagonalgestalt hat.

Ich beschreibe nun diesen Algorithmus, wobei selbstredend $\mathrm{char}(K) \neq 2$ vorausgesetzt wird.

Wir starten mit einer beliebigen symmetrischen Matrix $A \in K^{n\times n}$. Wir versuchen, mit "gleichgearteten" elementaren Zeilen- und Spaltentransformationen die Matrix in eine Matrix der Form

$$A' := \begin{pmatrix} a'_{1,1} & \underline{0}^t \\ \underline{0} & A'_{\mathrm{Rest}} \end{pmatrix}$$

zu transformieren. Hiernach ist dann A'_{Rest} automatisch wieder symmetrisch. (Wenn S irgendeine Matrix – also z.B. auch eine Elementarmatrix – ist, dann ist S^tAS wieder symmetrisch.) Danach fahren wir mit A'_{Rest} fort.

Wir müssen nur zeigen, wie man eine Matrix der Form A' erhält.

Es gibt vier Fälle:

1. Die erste Spalte und die erste Zeile von A enthalten nur Nulleinträge. Dann machen wir nichts.

2. Der Eintrag $a_{1,1}$ ist ungleich 0. Dann "eleminieren" wir wie beim Gauß-Algorithmus die Einträge rechts und unterhalb des ersten Eintrags. Hierbei verwenden wir "zueinander transponierte" Zeilen- und Spaltentransformationen.

3. Es ist $a_{1,1} = 0$, aber erste Spalte und die erste Zeile von A enthalten nicht nur Nulleinträge. Ferner gibt es ein i, so dass $a_{i,i} \neq 0$ ist. Dann vertauschen wir Spalten 1 und i und Zeilen 1 und i. Danach ist also der Eintrag mit Index $(1,1)$ gleich $a_{i,i}$. Wir fahren mit Fall 2 fort.

4. Es ist $a_{1,1} = 0$, aber die erste Spalte und die erste Zeile von A enthalten nicht nur Nulleinträge. Außerdem ist die Diagonale "trivial" (alle "Diagonaleinträge" sind gleich 0). Wir whlen ein i mit $a_{1,i} \neq 0$. Dann addieren wir die i-Spalte zur ersten Spalte und die i-te Zeile zur ersten Zeile. Danach ist also der Eintrag mit Index $(1,1)$ gleich $a_{1,i} + a_{i,1} = 2a_{1,i}$. Dies ist $\neq 0$, da wir $\mathrm{char}(K) \neq 2$ vorausgesetzt haben. Wir fahren wiederum mit Fall 2 fort.

Die Matrix S ist dann das Produkt der Elementarmatrizen, die zu den ausgeführten Spaltentransformationen gehören. Dieses Produkt kann man analog zur Berechnung der inversen Matrix mitberechnen. Hierzu schreibt man nun die Einheitsmatrix unter die Matrix A und wendet die Spaltentransformationen immer auf beide Matrizen gleichzeitig an. (Analog kann man auch S^t berechnen, indem man die Einheitsmatrix neben A schreibt und die Zeilenoperationen immer auf beide Matrizen anwendet.)

Übrigens sollte man hier die Assoziativität der Matrizenmultiplikation im Kopf behalten: Für beliebige Matrizen $B, C \in K^{n\times n}$ ist $(BA)C = B(AC)$. Dies kann man auch so ausdrücken: Es ist egal, ob man zuerst die durch B beschriebene Zeilentranformation und dann die duch C beschriebene Spaltentransformation anwendet oder ob man umgekehrt vorgeht. Und das gilt dann wohlgemerkt für beliebige Umformungen, nicht nur für elementare Umformungen.

Quadratische Formen

Sei β eine symmetrische Bilinearform auf einem Vektorraum V. Wir haben dann die Abbildung

$$q : V \longrightarrow K \ , \ \mathfrak{v} \mapsto \beta(\mathfrak{v}, \mathfrak{v}) \ .$$

Diese Abbildung heißt die zu β assoziierte *quadratische Form*. Eine wichtige Eigenschaft dieser Abbildung ist, dass

$$q(a\,\mathfrak{v}) = a^2 \cdot q(\mathfrak{v}) \ \text{ für alle } a \in K \text{ und } \mathfrak{v} \in V \tag{3.6}$$

gilt.

Interessanterweise kann man β aus q zurückgewinnen: Seien $\mathfrak{v}, \mathfrak{w} \in V$. Dann ist

$$q(\mathfrak{v} + \mathfrak{w}) = \beta(\mathfrak{v} + \mathfrak{w}, \mathfrak{v} + \mathfrak{w}) = q(\mathfrak{v}) + 2\beta(\mathfrak{v}, \mathfrak{w}) + q(\mathfrak{w}) \ .$$

Somit ist also

$$\beta(\mathfrak{v}, \mathfrak{w}) = \frac{1}{2} \cdot \big(q(\mathfrak{v} + \mathfrak{w}) - q(\mathfrak{v}) - q(\mathfrak{w})\big) \ . \tag{3.7}$$

Wieder haben wir benutzt, dass $\mathrm{char}(K) \neq 2$ ist.

Dies kann man zu einer Definition machen:

Definition Eine *quadratische Form* auf V ist eine Abbildung $V \longrightarrow K$ mit den folgenden Eigenschaften: Es gilt (3.6) und wenn man eine Abbildung $\beta : V \times V \longrightarrow K$ mittels (3.7) definiert, dann ist β eine (symmetrische) Bilinearform auf V.

Die Menge der quadratischen Formen auf V ist auch in offensichtlicher Weise ein K-Vektorraum. Per Definition ist nun die Zuordnung von symmetrischen Bilinearformen auf V zu quadratischen Formen auf V ein Homomorphismus von K-Vektorräumen. Es ist auch ein Isomorphismus, denn man per Definition kann man jeder quadratischen Form eine Bilinearform zuordnen, und wenn q eine quadratische Form ist und β die q zugeordnete Bilinearform ist, dann definiert β wieder die quadratische Form q, wie man leicht sieht.

Sei nun wieder V endlich-dimensional. Sei $\mathfrak{B}$ eine Basis von V und B die Matrix der Bilinearform β. Sei wie immer $X_1, \ldots, X_n$ die Dualbasis zu $\mathfrak{B}$. Dann ist die β zugeordnete quadratische Form die folgende Abbildung:

$$\sum_{i,j=1}^{n} b_{i,j} X_i X_j = \sum_{i=1}^{n} b_{i,i} X_i^2 + \sum_{i<j} 2 b_{i,j} X_i X_j$$

Ferner hat jede quadratische Form so eine Darstellung. Die Koeffizienten $b_{i,j}$ sind dabei wieder eindeutig bestimmt. Wenn man die obige abstrakte Definition von quadratischen Formen nicht mag, kann man quadratische Formen auf endlich-dimensionalen Vektorräumen auch als Abbildungen definieren, die man wie hier beschrieben darstellen kann.

Satz 3.1 kann man nun auch für quadratische Formen formulieren. Er lautet wie folgt:

Satz 3.2 *Sei V eindlich-dimensional und $q : V \longrightarrow K$ eine quadratische Form. Dann gibt es ein Koordinatensystem $X_1, \ldots, X_n$ auf V sowie*

$a_1, \ldots, a_n \in K$, so dass

$$q = \sum_{i=1}^{n} a_i X_i^2$$

ist.

Den Algorithmus, den wir für den Beweis von Satz 3.1 benutzt haben, kann man auch so modifizieren, dass man einen Algorithmus erhält, der direkt mit quadratischen Formen arbeitet. Schritt 2 entspricht dann einer quadratischen Ergänzung. Überlegen Sie sich das!

Symmetrische Bilinearformen über den reellen Zahlen

Sei nun V ein endlich erzeugter $\mathbb{R}$-Vektorraum und β eine symmetrische Bilinearform auf V.

Wir wissen schon: Es gibt eine Basis $\mathfrak{B}$ von V und reelle Zahlen $a_1, \ldots, a_n$, so dass die Matrix von β bzgl. $\mathfrak{B}$ die Diagonalmatrix

$$A = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$$

ist.

Wir nehmen nun an, dass die ersten s Diagonaleinträge positiv, die nächsten t Einträge negativ und die letzten Einträge 0 sind.

Sei nun T die $n \times n$-Diagonalmatrix, deren Diagonaleinträge wie folgt gegeben sind: $\sqrt{a_1}, \cdots \sqrt{a_s}, \sqrt{-a_{s+1}}, \ldots, \sqrt{-a_{s+t}}, 1, \ldots, 1$.[1]

Sei ferner D diejenige $n \times n$-Diagonalmatrix, deren Diagonaleinträge wie folgt geben sind: Die ersten s Einträge sind 1, die nächsten t Einträge sind -1, die restlichen Einträge sind 0.

Dann ist offensichtlich

$$A = TDT = T^t DT$$

Wir haben den folgenden Satz:

Satz 3.3 (Trägheitssatz von Sylvester) *Sei β eine symmetrische Bilinearform auf einem endlich erzeugten $\mathbb{R}$-Vektorraum. Dann gibt es eine Basis von V bezüglich welcher die Matrix von β eine Diagonalmatrix ist, wobei*

[1] Eigentlich müssten wir nun beweisen, dass man aus jeder positiven reellen Zahl die Wurzel ziehen kann ...

die Diagonaleinträge alle 1, -1 oder 0 sind. Die Anzahl der Einträge, welche gleich 1, -1 bzw. 0 sind, ist hierbei unabhängig von der gewählten Basis mit der angegebenen Eigenschaft.

Umformuliert in quadratische Formen bedeutet dies:

Satz 3.4 *Sei q eine quadratische Form auf einem endlich erzeugten $\mathbb{R}$-Vektorraum V. Dann gibt es ein Koordinatensystem $X_1, \ldots, X_n$ auf V sowie nicht-negative ganze Zahlen s und t, so dass*

$$q = X_1^2 + \cdots + X_s^2 - X_{s+1}^2 - \cdots - X_{s+t}^2 \, .$$

Die Zahlen s und t sind hierbei unabhängig vom Koordinatensystem mit der angegebenen Eigenschaft.

Hierzu müssen wir noch die Eindeutigkeit (Unabhängigkeit von der Wahl der Basis bzw. des Koordinatensystems) zeigen. Sei hierzu β eine Bilinearform und q die entsprechende quadratische Form.

Zunächst ist die Anzahl der Nullen auf der Diagonalen gleich der Dimension des Kerns von β. Somit ist die Anzahl der Nullen unabhängig von der Wahl der Basis. Oder mit anderen Worten: $s + t$ ist durch q eindeutig bestimmt.

Seien nun zwei Koordinatensysteme $X_1, \ldots, X_n$ und $X_1', \ldots, X_n'$ mit der angegebenen Eigenschaft gegeben; seien $\mathfrak{B}, \mathfrak{B}'$ die entsprechenden Basen. Seien die Basisvektoren wie oben angegeben geordnet und sei s wie oben und analog s'. Sei nun $V_> := \langle \mathfrak{b}_1, \ldots, \mathfrak{b}_s \rangle$ und $V_\leq := \langle \mathfrak{b}_{s+1}, \ldots, \mathfrak{b}_n \rangle$. Seien $V_>'$ und $V_\leq'$ analog definiert.

Dann gilt also für alle $\mathfrak{v} \in V_>$ mit $\mathfrak{v} \neq \mathfrak{o}$: $q(\mathfrak{v}) > 0$. (Wenn $a_1, \ldots, a_s \in \mathbb{R}$ sind, ist $q(a_1\mathfrak{b}_1 + \cdots + a_s\mathfrak{b}_s) = a_1^2 + \cdots + a_s^2$, und wenn eines der Skalare a_i von Null verschieden ist, dann ist die Summe der Quadrate > 0.)

Ferner gilt für alle $\mathfrak{v} \in V_\leq$: $q(\mathfrak{v}) \leq 0$. Analoge Aussagen gelten für $V_>'$ und $V_\leq'$.

Nun angenommen, es wäre $s \neq s'$. Sei OE z.B. $s > s'$. Dann ist also $\mathrm{Dim}(V_>) + \mathrm{Dim}(V_\leq') = s + (n - s') > n$. Damit haben die beiden Räume einen nicht-trivial Schnitt. Dies ist ein Widerspruch. □

Bemerkung Über den komplexen Zahlen ist die Situation noch besser, weil wir dann aus jeder Zahl die Quadratwurzel ziehen können. Wenn also q eine quadratische Form auf einem endlich erzeugten komplexen Vektorraum ist, gibt es ein $r \in \mathbb{N}_0$ und ein Koordintensystem $X_1, \ldots, X_n$, so dass

$$q = X_1^2 + \cdots + X_r^2$$

ist. Hierbei ist r durch q eindeutig bestimmt. Aber das ist ja fast schon langweilig ...

3.3 Skalarprodukte

Skalarprodukte sind spezielle Bilinearformen auf $\mathbb{R}$-Vektorräumen.

Wir fixieren einen $\mathbb{R}$-Vektorraum V.

Definition

- Eine symmetrische Bilinearform $\beta : V \times V \longrightarrow \mathbb{R}$ ist *positiv definit*, wenn für alle $\mathfrak{v} \in V$ mit $\mathfrak{v} \neq \mathfrak{o}$ $\beta(\mathfrak{v}, \mathfrak{v}) > 0$ gilt.
- Eine positiv definite symmetrische Bilinearform auf V heißt auch *Skalarprodukt*.

Skalarprodukte bezeichnen wir in der Regel mit $\langle ., . \rangle$.

Ein Beispiel für ein Skalarprodukt ist natürlich das Standardskalarprodukt auf dem $\mathbb{R}^n$. In der Analysis kommen auch oft Skalarprodukte auf nicht endlich erzeugten Vektorräumen vor. Hier ist ein Beispiel:

Beispiel 3.14 Sei $\mathcal{C}([0,1], \mathbb{R})$ der $\mathbb{R}$-Vektorraum der stetigen $\mathbb{R}$-wertigen Funktionen auf dem Intervall $[0,1]$. Da das Produkt jeder stetigen Funktion stetig ist, stetige Funktionen Riemann-integrierbar und das Riemann-Integral bzgl. fester Grenzen linear ist, kann man wie folgt eine symmetrische Bilinearform auf $\mathcal{C}([0,1], \mathbb{R})$ definieren:

$$\langle f, g \rangle := \int_0^1 f(t)g(t)dt$$

Man sieht leicht, dass die Bilinearform auch positiv definit also ein Skalarprodukt ist.

Sei nun also so ein Skalarprodukt $\langle ., . \rangle$ gegeben. Sei ferner q die assoziierte quadratische Form. Wir setzen für $\mathfrak{v} \in V$: $\| \mathfrak{v} \| := \sqrt{q(\mathfrak{v})} = \sqrt{\langle \mathfrak{v}, \mathfrak{v} \rangle}$. Diese Zahl nennen wir die *Norm* von $\mathfrak{v}$ bzgl. des gebenen Skalarprodukts.

Aussage 3.15 *Für $\mathfrak{v}, \mathfrak{w} \in V$ und $c \in K$ ist*

a) $\| \mathfrak{v} \| = 0 \longleftrightarrow \mathfrak{v} = \mathfrak{o}$

b) $\| c\,\mathfrak{v} \| = |c| \cdot \| \mathfrak{v} \|$

c) $\| \mathfrak{v} + \mathfrak{w} \| \leq \| \mathfrak{v} \| + \| \mathfrak{w} \|$ *("Dreiecksungleichung")* .

Die ersten beiden Aussagen sind offensichtlich, für den Beweis der Dreiecksungleichung führen wir zunächst einige Begriffe ein. Wir halten aber schon mal fest:

Definitionen Eine Abbildung von einem $\mathbb{R}$-Vektorraum X nach $\mathbb{R}_{\geq 0}$ mit den Eigenschaften in der obigen Aussage heißt *Norm* auf X. Ein Skalarprodukt induziert also eine Norm. Aber nicht jede Norm wird von einem Skalarprodukt induziert, wie man leicht sieht.

Die folgenden Definitionen beziehen sich auf das fest gewählte Skalarprodukt auf V.

Definition Vektoren $\mathfrak{v}, \mathfrak{w} \in V$ sind *orthogonal* zueinander, wenn $\langle \mathfrak{v}, \mathfrak{w} \rangle = 0$ ist. In diesem Fall schreiben wir auch $\mathfrak{v} \perp \mathfrak{w}$. Ein Vektor $\mathfrak{v} \in V$ ist *normiert*, wenn $\| \mathfrak{v} \| = 1$ ist.

Ein *Orthogonalsystem* von Vektoren in V ist ein System von Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_r$, welche ungleich dem Nullvektor und paarweise orthogonal zueinander sind. Ein *Orthonormalsystem* ist ein Orthogonalsystem, dessen Vektoren normiert sind.

Beachten Sie: Wenn $\mathfrak{v} \neq \mathfrak{o}$ ist, dann ist der Vektor $\frac{1}{\|\mathfrak{v}\|} \cdot \mathfrak{v}$ normiert. Somit kann man "durch skalieren" ein Orthogonalsystem immer in ein Orthonormalsystem überführen.

Lemma 3.16 *Ein Orthogonalsystem ist linear unabhängig.*

Beweis. Sei $\mathfrak{v}_1, \ldots, \mathfrak{v}_n$ so ein System, und seien $a_1, \ldots, a_n \in \mathbb{R}$ mit $\sum_i a_i \mathfrak{v}_i = \mathfrak{o}$. Dann ist für alle $j = 1, \ldots, n$: $0 = \sum_i a_i \langle \mathfrak{v}_i, \mathfrak{v}_j \rangle = a_j \| \mathfrak{v}_j \|^2$ und somit $a_j = 0$. □

Seien nun $\mathfrak{v}, \mathfrak{w} \in V$. Wenn nun $\mathfrak{v}$ und $\mathfrak{w}$ orthogonal zueinander sind, haben wir

$$\langle \mathfrak{v} + \mathfrak{w}, \mathfrak{v} + \mathfrak{w} \rangle = \langle \mathfrak{v}, \mathfrak{v} \rangle + \langle \mathfrak{w}, \mathfrak{w} \rangle \ ,$$

d.h.

$$\| \mathfrak{v} + \mathfrak{w} \|^2 = \| \mathfrak{v} \|^2 + \| \mathfrak{w} \|^2 \ . \tag{3.8}$$

Bemerkung Diese Gleichung wird manchmal "Satz von Pythagoras" genannt. Das halte ich allerdings für irreführend. Unter einem "Satz von Pythagoras" würde ich eher eine Aussage verstehen, dass unter gewissen "offensichtlichen geometrischen Axiomen" eine Gleichung wie oben gilt. Die Aussage, dass die Normabbildung von einem Skalarprodukt induziert sei kann man zwar auch als so ein Axiom auffassen, allerdings sind die Axiome des Skalarprodukts (im Vergleich mit anderen möglichen Axiomen) geometrisch eher unituitiv.

Wir setzen nun voraus, dass $\mathfrak{v} \neq \mathfrak{o}$ ist; weitere Voraussetzungen machen wir nicht. Dann hat $\tilde{\mathfrak{v}} := \frac{1}{\|\mathfrak{v}\|} \cdot \mathfrak{v}$ die Norm 1. Nun ist $\mathfrak{v}$ orthogonal zu $\mathfrak{w} -$

$\langle \tilde{\mathfrak{v}}, \mathfrak{w} \rangle \cdot \tilde{\mathfrak{v}} = \mathfrak{w} - \frac{\langle \mathfrak{v}, \mathfrak{w} \rangle}{\langle \mathfrak{v}, \mathfrak{v} \rangle} \cdot \tilde{\mathfrak{v}}$. Wir erhalten die folgende "Zerlegung" von $\mathfrak{w}$ in zueinander orthogonale Vektoren:

$$\mathfrak{w} = (\mathfrak{w} - \langle \tilde{\mathfrak{v}}, \mathfrak{w} \rangle \cdot \tilde{\mathfrak{v}}) + \langle \tilde{\mathfrak{v}}, \mathfrak{w} \rangle \cdot \tilde{\mathfrak{v}}$$

Wenn wir hierauf (3.8) anwenden, erhalten wir:

$$\langle \tilde{\mathfrak{v}}, \mathfrak{w} \rangle^2 \leq \| \mathfrak{w} \|^2$$

Wir haben nun:

Aussage 3.17 (Cauchy-Schwarzsche Ungleichung) *Es ist*

$$|\langle \mathfrak{v}, \mathfrak{w} \rangle| \leq \| \mathfrak{v} \| \cdot \| \mathfrak{w} \| \ .$$

In dieser Ungleichung gilt genau dann Gleichheit, wenn $\mathfrak{v}$ *und* $\mathfrak{w}$ *linear abhängig sind.*

Beweis. Wir haben die Ungleichung bewiesen, falls $\mathfrak{v} \neq \mathfrak{o}$ ist. Für $\mathfrak{v} = \mathfrak{o}$ ist sie offensichtlich erfüllt.

Für die zweite Aussage können wir $\mathfrak{v} \neq \mathfrak{o}$ voraussetzen. Dann sind die folgenden Aussagen äquivalent:

- $\mathfrak{v}$ und $\mathfrak{w}$ sind linear abhängig.
- $\mathfrak{w} = \langle \tilde{\mathfrak{v}}, \mathfrak{w} \rangle \cdot \tilde{\mathfrak{v}}$
- $|\langle \mathfrak{v}, \mathfrak{w} \rangle| = \| \mathfrak{v} \| \cdot \| \mathfrak{w} \|$

□

Wir kommen nun zum *Beweis der Dreiecksungleichung.*

Seien $\mathfrak{v}, \mathfrak{w} \in V$. Dann ist $\| \mathfrak{v} + \mathfrak{w} \|^2 = \langle \mathfrak{v} + \mathfrak{w}, \mathfrak{v} + \mathfrak{w} \rangle = \| \mathfrak{v} \|^2 + \| \mathfrak{w} \|^2 + 2\langle \mathfrak{v}, \mathfrak{w} \rangle \leq \| \mathfrak{v} \|^2 + 2 \| \mathfrak{v} \| \cdot \| \mathfrak{w} \| + \| \mathfrak{w} \|^2 = (\| \mathfrak{v} \| + \| \mathfrak{w} \|)^2$. □

Die Zahl $\frac{\langle \mathfrak{v}, \mathfrak{w} \rangle}{\|\mathfrak{v}\| \cdot \|\mathfrak{w}\|}$ ist also im Intervall $[-1, 1]$. Somit kann man definieren (wobei wir die Existenz der Cosinus Funktion voraussetzen):

Definition Der *Winkel* zwischen $\mathfrak{v}$ und $\mathfrak{w}$ ist die eindeutig bestimmte Zahl $\alpha \in [0, \pi]$ mit $\cos(\alpha) = \frac{\langle \mathfrak{v}, \mathfrak{w} \rangle}{\|\mathfrak{v}\| \cdot \|\mathfrak{w}\|}$.

Endlich erzeugte Vektorräume mit Skalarprodukt

Der Begriff der positiven Definitheit überträgt sich sofort auf Matrizen:

Definition Sei $B \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann ist B *positiv definit*, wenn die durch B definierte Bilinearform $\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, $(\underline{x}, \underline{y}) \mapsto \underline{x}^t B \underline{y}$ positiv definit ist.

Somit ist also B genau dann positiv definit, wenn für alle $\underline{x} \in \mathbb{R}^n$ mit $\underline{x} \neq \underline{0}$ die Zahl $\underline{x}^t B \underline{x}$ positiv ist.

Man sieht auch sofort, dass positive Definitheit "invariant unter Kongruenz von Matrizen" ist. D.h.: Wenn $B \in \mathbb{R}^{n\times n}$ symmetrisch und $S \in \mathbb{R}^{n\times n}$ invertierbar ist, dann ist B genau dann positiv definit, wenn $S^t B S$ positiv definit ist.

Beispiel 3.18 Sei S eine beliebige invertierbare reelle Matrix. Dann ist $S^t S$ positiv definit. Denn: Diese Matrix ist kongruent zur Einheitsmatrix, und die Einheitsmatrix ist offensichtlich positiv definit.

Wir haben auch das folgende offensichtliche Lemma:

Lemma 3.19 *Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum und β eine symmetrische Bilinearform auf V. Sei B die Matrix von β bezüglich irgendeiner Basis von V. Dann ist β genau dann positiv definit, wenn B positiv definit ist.*

Sei nun V ein endlich erzeugter $\mathbb{R}$-Vektorraum mit Dimension n und sei wie zuvor $\langle ., . \rangle$ ein Skalarprodukt auf V. Die folgende Definition bezieht sich wieder auf das fixierte Skalarprodukt.

Definition

- Eine Basis von V, die auch ein Orthogonalsystem ist, heißt *Orthogonalbasis*. Wenn sie ein Orthonormalsystem ist, heißt sie *Orthonormalbasis*.
- Ein Koordinatensystem von V, die durch eine Orthonormalbasis von V definiert wird, heißt *kartesisches Koordinatensystem*.

Somit ist eine Basis $\mathfrak{B}$ offensichtlich genau dann eine Orthogonalbasis, wenn die Matrix von $\langle ., . \rangle$ bzgl. $\mathfrak{B}$ eine Diagonalmatrix ist. Sie ist genau dann eine Orthonormalbasis, wenn die Matrix von $\langle ., . \rangle$ bzgl. $\mathfrak{B}$ die Einheitsmatrix ist.

Bezüglich des zweiten Punkts beachte man: Ein Koordinatensystem von V ist per Definition eine Basis von V^*. Ferner definiert jede Basis $\mathfrak{B}$ von V eine Basis von V^*: die Dualbasis zu $\mathfrak{B}$.

Wir betrachten kurz Orthogonal- und Orthonormalbasen in $\mathbb{R}^n$ bezüglich des Standardskalarprodukts. Sei $\underline{b}_1, \ldots, \underline{b}_n$ ein System von Vektoren in $\mathbb{R}^n$,

und sei B die durch dieses System definierte Matrix. Nun ist $\underline{b}_i^t \underline{b}_j$ der Eintrag (i,j) der Matrix B^tB. Somit ist das System genau dann orthogonal, wenn B^tB eine Diagonalmatrix ist, deren "Diagonaleinträge" alle positiv sind. Das System ist genau dann orthonormal, wenn $B^tB = I_n$ ist.

Man definiert nun:

Definition Sei $B \in \mathbb{R}^{n\times n}$. Dann ist B *orthogonal*, wenn $B^tB = I_n$ ist.

Warnung Per Definition ist das System von Vektoren $\underline{b}_1, \ldots, \underline{b}_n$ genau dann ortho*normal* (bzgl. des Standardskalarprodukts), wenn die Matrix $(\underline{b}_1 | \cdots | \underline{b}_n)$ ortho*gonal* ist.

Orthogonale Matrizen sind offensichtlich invertierbar und es ist

$$B^t = B^{-1} \ .$$

Beachten Sie, dass somit vom algorithmischen Gesichtspunkt aus das Invertieren einer orthogonalen Matrix trivial ist.

Wir kehren zum endlich erzeuten Vektorraum V und dem fixierten Skalarprodukt zurück:

Orthogonal- und Orthonormalbasen sind insbesondere aufgrund der folgenden Aussage praktisch:

Aussage 3.20

a) Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Orthonormalbasis von V und sei $\mathfrak{v} \in V$. Dann ist

$$\mathfrak{v} = \sum_{i=1}^{n} \langle \mathfrak{b}_i, \mathfrak{v} \rangle \cdot \mathfrak{b}_i \ .$$

b) Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Orthogonalbasis von V und sei $\mathfrak{v} \in V$. Dann ist

$$\mathfrak{v} = \sum_{i=1}^{n} \frac{\langle \mathfrak{b}_i, \mathfrak{v} \rangle}{\langle \mathfrak{b}_i, \mathfrak{b}_i \rangle} \cdot \mathfrak{b}_i \ .$$

Beweis. Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Orthonormalbasis. Dann gibt es $a_1, \ldots, a_n \in \mathbb{R}$ mit $\mathfrak{v} = \sum_{i=1}^n a_i \mathfrak{b}_i$. Nun ist $\langle \mathfrak{v}, \mathfrak{b}_j \rangle = \sum_{i=1}^n a_i \delta_{i,j} = a_j$.

Die zweite Aussage folgt sofort aus der ersten. □

Satz 3.5 *V hat eine Orthonormalbasis.*

Beweis. Das folgt sofort aus dem Trägheitssatz von Sylvester und der positiv-Definitheit. □

Man kann an dieser Stelle noch etwas mehr sagen: Sei $\mathfrak{B}$ eine beliebige Basis und sei B die Matrix des Skalarprodukts bzgl. $\mathfrak{B}$. Dann sind offensichtlich die Diagonaleinträge alle positiv. Wenn wir nun den Algorithmus für Aussage 3.13 auf diese Matrix anwenden, haben wir nach jedem Schritt wieder eine Matrix des Skalarprodukts bzgl. irgendeiner Basis von V. Und dann sind die Diagonaleinträge wieder alle positiv. Damit kommt in dem ganzen Algorithmus nur der 2.Fall vor. Zum Schluss kann man dann noch wie im Beweis des Trägheitssatzes von Sylvester die Diagonaleinträge zu Einsen machen.

Hieraus erhält man die folgenden beiden Sätze. Doch zunächst eine Definition:

Definition Sei S eine quadratische Dreiecksmatrix. Dann ist S *normiert*, falls alle Diagonaleinträge von S gleich 1 sind.

Bemerkung Manche Autoren nennen solche Matrizen auch strikte obere Dreiecksmatrizen. Das ist aber verwirrend, da auch Dreiecksmatrizen, deren Diaognaleinträge gleich 0 sind, strikte Dreiecksmatrizen heißen.

Übrigens bilden die normierten oberen Dreiecksmatrizen in $\mathbb{R}^{n\times n}$ eine Gruppe.

Satz 3.6 *Sei $\mathfrak{B}$ eine beliebige Basis von V.*

a) Es gibt eine eindeutig bestimmte Orthogonalbasis $\mathfrak{B}'$ von V, so dass die Koordinatenmatrix von $\mathfrak{B}'$ bzgl. $\mathfrak{B}$ eine normierte obere Dreiecksmatrix ist.

b) Es gibt eine eindeutig bestimmte Orthonormalbasis $\mathfrak{B}'$ von V, so dass die Koordinatenmatrix von $\mathfrak{B}'$ bzgl. $\mathfrak{B}$ eine obere Dreiecksmatrix mit positiven Diagonaleinträgen ist.

Satz 3.7 *Sei A eine symmetrische positiv definite Matrix.*

a) Es gibt eine eindeutig bestimmte normierte obere Dreiecksmatrix B, so dass B^tAB eine Diagonalmatrix ist. Diese Diagonalmatrix hat dann nur positive Einträge auf der Diagonalen.

b) Es gibt eine eindeutig bestimmte normierte obere Dreiecksmatrix B und eine eindeutig bestimmte Diagonalmatrix D, so dass $A = B^tDB$ ist. Hierbei hat D nur positive Einträge auf der Diagonalen.

c) Es gibt eine eindeutig bestimmte invertierbare obere Dreiecksmatrix C mit positiven Diagonaleinträgen, so dass $A = C^tC$ ist.

Beweis. Wir haben die Existenz der angegebenen Basis im ersten Satz bewiesen. Hieraus folgen alle anderen Existenzaussagen. Die Eindeutigkeit der Matrix C mit $A = C^tC$ kann man direkt nachrechnen. Hieraus folgt dann auch in allen anderen Fällen die Eindeutigkeit. □

Definition Eine wie in b) oder c) angegebene "Zerlegung" der Matrix A heißt *Cholesky-Zerlegung.*

Wir kommen noch zu einer Anwendung.

Wir betrachten nun den Vektorraum $\mathbb{R}^n$ mit dem Standardskalarprodukt. Sei eine beliebige Basis $\mathfrak{B} = (\underline{b}_1, \ldots, \underline{b}_n)$ von $\mathbb{R}^n$ gegeben. Nach Satz 3.6 a) gibt es eine eindeutig bestimmte Orthonormalbasis $\tilde{\mathfrak{B}} = (\tilde{\underline{b}}_1, \ldots, \tilde{\underline{b}}_n)$ von $\mathbb{R}^n$, so dass die Koordinatenmatrix von $\tilde{\mathfrak{B}}$ bzgl. $\mathfrak{B}$ eine normierte obere Dreiecksmatrix ist. Wenn wir nun $B := (\underline{b}_1 | \ldots | \underline{b}_n)$, $\tilde{B} := (\tilde{\underline{b}}_1 | \ldots | \tilde{\underline{b}}_n)$ setzen und S die angesprochene Koordinatenmatrix ist, dann ist $\tilde{B}$ orthogonal und $\tilde{B} = BS$ bzw. $B = \tilde{B}S^{-1}$.

Die folgende Aussage ist eine Umformulierung der obigen Überlegungen:

Aussage 3.21 (QR-Zerlegung) *Sei eine invertierbare Matrix $B \in \mathbb{R}^{n\times n}$ gegeben. Dann gibt es eine eindeutig bestimmte obere Dreiecksmatrix $R \in \mathbb{R}^{n\times n}$ mit positiven Diagonaleinträgen und eine orthogonale Matrix $Q \in \mathbb{R}^{n\times n}$, so dass $B = QR$ ist.*

Wenn B gegeben ist, kann man Q und R wie folgt berechnen:

Man wendet den Algorithmus zu Aussage 3.13 auf die Matrix B^tB an (das ist die Koordinatenmatrix des Standardskalarprodukts bzgl. B) und normalisiert dann dann die Diagonaleinträge. Man erhält eine obere Dreiecksmatrix mit positiven Diagonaleinträgen S, so dass $S^tB^tBS = I_n$ ist. Wenn man nun $Q := BS$ setzt, ist $Q^tQ = I_n$, d.h. Q ist orthogonal. Mit $R := S^{-1}$ ist $B = QR$.

Das Gram-Schmidt Verfahren

Es gibt noch einen anderen Zugang, zu diesem Themenkomplex, der seinen eigenen Wert hat:

Definition Sei U ein Untervektorraum von V (der hier nicht endlich erzeugt sein muss). Dann ist das *orthogonale Komplement* von U in V gleich

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

Es gibt einen engen Zusammenhang zwischen dem orthogonalen Komplement (einem Untervektorraum von V) und dem Annullator (einem Untervektorraum von V°):

Das Skalarprodukt auf dem endlich erzeugten Vektorraum V induziert einen Isomorphismus $\Psi : V \longrightarrow V^*$, $\mathfrak{v} \mapsto \langle \mathfrak{v}, . \rangle$. (Wenn wir das Skalarprodukt mit β bezeichnen würden, wäre dies β_1.)

Nun ist für $\mathfrak{x}, \mathfrak{y} \in V$ $\langle \mathfrak{x}, \mathfrak{y} \rangle = \langle \Psi(\mathfrak{x}), \mathfrak{y} \rangle_{\text{kan}}$. Somit ist

$$\Psi(U^\perp) = U^\circ \ .$$

Beispiel 3.22 Wir betrachten das Standardskalarprodukt auf dem Standardvektorraum K^n. Wenn man $(K^n)^*$ mit $K^{1\times n}$ identifiziert, ist Ψ genau der Isomorphismus $K^n \longrightarrow K^{1\times n}$, $\underline{x} \mapsto \underline{x}^t$.

Aussage 3.23 *Es ist $U \oplus U^\perp = V$.*

Beweis. Sei $\mathfrak{v} \in U \cap U^\perp$. Dann ist insbesondere $\langle \mathfrak{v}, \mathfrak{v} \rangle = 0$ und somit $\mathfrak{v} = \mathfrak{o}$. Damit ist die Summe direkt.

Nach Aussage 2.50 ist ferner $\text{Dim}(U) + \text{Dim}(U^\perp) = \text{Dim}(U) + \text{Dim}(U^\circ) = \text{Dim}(V)$. Damit ist die direkte Summe gleich dem gesamten Raum. □

Wir definieren nun wie folgt eine Abbildung $\pi_U : V \longrightarrow V$: Wir schreiben $\mathfrak{v} \in V$ in der Form $\mathfrak{u} + \mathfrak{w}$ mit $\mathfrak{u} \in U, \mathfrak{w} \in U^\perp$. Dann setzen wir $\pi_U(\mathfrak{v}) := \mathfrak{u}$. Man sieht leicht, dass dies eine lineare Abbildung ist, die $\pi_U \circ \pi_U = \pi_U$ erfüllt. Diese lineare Abbildung heißt die *Orthogonalprojektion* von V auf U. Wir haben also

$$\pi_U + \pi_{U^\perp} = \text{id}_V \ .$$

Wenn $\mathfrak{b}_1, \ldots, \mathfrak{b}_r$ eine Orthonormalbasis von U ist, dann ist

$$\pi_U(\mathfrak{v}) = \sum_{i=1}^{r} \langle \mathfrak{b}_i, \mathfrak{v} \rangle \cdot \mathfrak{b}_i \ .$$

Demnach ist dann

$$\pi_{U^\perp}(\mathfrak{v}) = \mathfrak{v} - \sum_{i=1}^{r} \langle \mathfrak{b}_i, \mathfrak{v} \rangle \cdot \mathfrak{b}_i \ .$$

Bemerkung Der Begriff der Projektion ist allgemeiner: Man nennt einen Endomorphismus φ eines Vektorraums eine *Projektion*, wenn $\varphi^2 = \varphi$ ist. Wenn V ein beliebiger Vektorraum über einem Körper K ist und $V = U \oplus W$ mit Untervektorräumen U und W von V ist, dann gibt es genau eine Projektion von V mit Bild U (man sagt: eine Projektion *auf* U) und Kern W. Diese ist durch $\mathfrak{u} + \mathfrak{w} \mapsto \mathfrak{u}$ für $\mathfrak{u} \in U$ und $\mathfrak{w} \in W$ gegeben.

Nach der letzten Aussage kann man jedes Orthonormalsystem in V zu einer Basis von V ergänzen. Dies kann man auch iterativ machen, und zwar mittels des so genannten *Gram-Schmidt Orthogonalisierungsverfahrens*:

Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Basis von V und sei $U_i := \langle \mathfrak{b}_1, \ldots, \mathfrak{b}_i \rangle$ für $i = 0, \ldots, n$ sowie

$$\tilde{\mathfrak{b}}_i := \pi_{(U_{i-1})^\perp}(\mathfrak{b}_i) \ .$$

Nun ist $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_i$ eine Orthogonalbasis von U_i.

Dies sieht man per Induktion über i wie folgt: Der Induktionsanfang $i = 0$ ist trivial. Zum Induktionsschritt: Per Induktionsannahme ist $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_i$ eine Orthogonalbasis von U_i. Nun liegt $\mathfrak{b}_{i+1}$ nicht in U_i und somit ist $\tilde{\mathfrak{b}}_{i+1} = \pi_{U_i^\perp}(\mathfrak{b}_{i+1}) \neq \mathfrak{o}$. Per Definition ist ferner $\tilde{\mathfrak{b}}_{i+1}$ orthogonal zu allen Vektoren zu U_i, insbesondere also zu $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_i$. Schließlich liegt $\pi_{U_i^\perp}(\mathfrak{b}_{i+1})$ in U_{i+1} denn $\mathfrak{b}_{i+1}$ und $\pi_{U_i}(\mathfrak{b}_{i+1})$ liegen in U_{i+1}, also auch deren Differenz. Folglich ist $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_{i+1}$ ein Orthogonalsystem in U_{i+1}. Da dieser Raum auch die Dimension $i + 1$ hat, ist es eine Orthogonalbasis von U_{i+1}. □

Wir haben die folgenden explizite Formeln:

$$\tilde{\mathfrak{b}}_i = \mathfrak{b}_i - \sum_{j=1}^{i-1} \frac{\langle \tilde{\mathfrak{b}}_j, \mathfrak{b}_i \rangle}{\langle \tilde{\mathfrak{b}}_j, \tilde{\mathfrak{b}}_j \rangle} \cdot \tilde{\mathfrak{b}}_j \tag{3.9}$$

Sei $\tilde{\tilde{\mathfrak{b}}}_i := \frac{1}{\|\mathfrak{b}_i\|} \cdot \mathfrak{b}_i$. Dann ist also $\tilde{\tilde{\mathfrak{b}}}_1, \ldots, \tilde{\tilde{\mathfrak{b}}}_n$ eine Orthonormalbasis von V und

$$\tilde{\mathfrak{b}}_i = \mathfrak{b}_i - \sum_{j=1}^{i-1} \langle \tilde{\mathfrak{b}}_j, \mathfrak{b}_i \rangle \cdot \tilde{\mathfrak{b}}_j \ .$$

Da $\tilde{\mathfrak{b}}_i$ eine Summe von $\mathfrak{b}_i$ und einem Element in U_{i-1} ist, ist die Koordinatenmatrix der Orthogonalbasis $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_n$ bzgl. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ auch eine normierte obere Dreiecksmatrix.

Wir können das nun zusammenfassen:

Aussage 3.24 *Die oben definierte Orthogonalbasis $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_n$ ist eine Orthogonalbasis von V, und es ist die einzige Orthogonalbasis von V, deren Koordinatenmatrix bzgl. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine normierte obere Dreiecksmatrix ist.*

Beweis. Wir zeigen noch die Eindeutigkeit. Sei also eine Basis $\mathfrak{c}_1, \ldots, \mathfrak{c}_n$ mit den angegebenen Eigenschaften gegeben. Dann ist $\mathfrak{c}_1, \ldots, \mathfrak{c}_i$ eine Orthogonalbasis von U_i. Somit liegen sowohl $\tilde{\mathfrak{b}}_i$ als auch $\mathfrak{c}_i$ in $U_i \cap U_{i-1}^\perp$. Dies ist aber das orthogonale Komplement von U_{i-1} in U_i. Dieses orthogonale Komplement ist 1-dimensional, also sind $\tilde{\mathfrak{b}}_i$ und $\mathfrak{c}_i$ Vielfache voneinander. Somit sind auch die entsprechenden Koordinatenvektoren Vielfache voneinander.

Da der i-te Eintrag beider Koordinatenvektoren bzgl. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ nach Voraussetzung aber 1 ist, sind beide Koordinatenvektoren gleich. Also sind auch die Vektoren gleich. □

Dies zeigt dann nochmal Satz 3.6 a) und damit auch alle anderen Aussagen in den Sätzen 3.6 und 3.7.

Für explizite Berechnungen kann man sowohl bei der Cholesky Zerlegung als auch bei der QR-Zerlegung sowohl den Algorithmus zu Aussage 3.13 als auch das Gram-Schmidt benutzen. In der Numerischen Mathematik (in der man approximativ mit Fließkommazahlen rechnet) gibt es noch andere Algorithmen.

Bemerkung Mittels des Gram-Schmidt Verfahrens können wir auch nochmal beweisen, dass $U \oplus U^\perp = V$ gilt (ohne den Dualraum zu benutzen): Sei eine Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V gegeben, wobei $\mathfrak{b}_1, \ldots, \mathfrak{b}_r$ eine Basis von U ist. Wenn man nun $\tilde{\mathfrak{b}}_i$ mittels der Formel (3.9) definiert, ist $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_n$ eine Orthogonalbasis von V und $\tilde{\mathfrak{b}}_1, \ldots, \tilde{\mathfrak{b}}_r$ eine Orthogonalbasis von U. Man sieht nun leicht, dass $\tilde{\mathfrak{b}}_{r+1}, \ldots, \tilde{\mathfrak{b}}_n$ eine Orthogonalbasis von $U^\perp$ ist.

3.4 Euklidische Vektorräume

Normierte Räume

Definition Ein *normierter Vektorraum* oder ein *normierter Raum* ist ein $\mathbb{R}$-Vektorraum V zusammen mit einer Norm $\| \cdot \| \colon V \longrightarrow \mathbb{R}$.

Notation Das Paar $(V, \| \cdot \|)$ wird auch mit V bezeichnet. Die Norm wird dann oftmals mit $\| \cdot \|_V$ bezeichnet. Wenn man diese Bezeichnungen wählt muss man natürlich aufpassen, ob man mit V nun den normierten Raum oder nur den zugrundeliegenden Vektorraum meint.

Beachten Sie, dass ein normierter Vektorraum immer eine feste Norm hat, genau wie eine Gruppe immer eine feste Verknüpfung hat.

Wie immer definieren wir gleich die entsprechenden Homomorphismen:

Definition Seien $(V, \| \cdot \|_V)$ und $(W, \| \cdot \|_W)$ normierte Räume. Ein *Homomorphismus von normierten Räumen* oder eine *Isometrie* von $(V, \| \cdot \|)$ nach $(W, \| \cdot \|)$ ist eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\| \varphi(\mathfrak{v}) \|_W = \| \mathfrak{v} \|_V$ für alle $\mathfrak{v} \in V$.

Entsprechend dem allgemeinen Begriff des Isomorphismus als eines umkehrbaren Homomorphismus kann man auch Isomorphismen von normierten Räumen definieren. Isomorphismen von normierten Räumen heißen auch *isometrische Isomorphismen*. Ein Homomorphismus von normierten Räumen ist dann genau ein Isomorphismus, wenn er bijektiv ist. Für einen normierten Raum $(V, \| \cdot \|)$ bilden die isometrischen Isomorphismen von $(V, \| \cdot \|)$ nach $(V, \| \cdot \|)$ eine Untergruppe der Gruppe der Automorphismen des Vektorraums V.

Man sieht sofort, dass Isometrien injektiv sind. Insbesondere ist eine Isometrie zwischen normierten Räumen derselben Dimension immer ein isometrischer Isomorphismus.

Sei nun $(V, \| \cdot \|)$ ein normierter Vektorraum. Dann induziert die Norm $\| \cdot \|$ eine so genannte *Metrik* oder *Abstandsfunktion*

$$d : V \times V \longrightarrow \mathbb{R} \geq 0 \ , \ (\mathfrak{v}, \mathfrak{w}) \mapsto \| \mathfrak{v} - \mathfrak{w} \| \ .$$

Diese erfüllt die folgenden Bedingungen:

Für alle $\mathfrak{x}, \mathfrak{y}, \mathfrak{z} \in V$ gilt:

- $d(\mathfrak{x}, \mathfrak{y}) = 0 \longleftrightarrow \mathfrak{x} = \mathfrak{y}$
- $d(\mathfrak{x}, \mathfrak{y}) = d(\mathfrak{y}, \mathfrak{x})$
- $d(\mathfrak{x}, \mathfrak{z}) \leq d(\mathfrak{x}, \mathfrak{y}) + d(\mathfrak{y}, \mathfrak{z})$

Allgemein ist eine *Metrik* auf einer Menge M eine Abbildung $d : M \times M \longrightarrow \mathbb{R}_{\geq 0}$ mit den obigen Eigenschaften. Ein *metrischer Raum* ist eine dann eine Menge M zusammen mit einer Metrik auf M. Wenn (M, d_M) und (N, d_N) zwei metrische Räume sind, ist eine *Isometrie* (oder ein Morphismus von metrischen Räumen) eine Abbildung $f : M \longrightarrow N$ mit $d_N(f(a), f(b)) = d_M(a, b)$ für alle $a, b \in M$.

Wenn Sie die Vorlesung Analysis hören, kennen Sie metrische Räume schon.

Zwei Aspekte sollten Sie dann beachten: Erstens sind Isometrien zwischen metrischen Räumen stetig. Zweitens übertragen sich alle Begriffe und Aussagen über metrische Räume direkt auf normierte Räume. Insbesondere können wir von stetigen Abbildungen zwischen normierten Räumen reden, und Isometrien zwischen normierten Räumen sind stetig.

Ich gehe hier nur auf einige wenige analytische Aspekte ein.

Wir kommen zurück zu normierten Vektorräumen. Aus der Dreiecksungleichung folgt sofort die so genannte *umgekehrte Dreiecksungleichung*. Für $\mathfrak{x}, \mathfrak{y} \in V$ ist

$$| \| \mathfrak{x} \| - \| \mathfrak{y} \| | \leq \| \mathfrak{x} + \mathfrak{y} \|$$

Hieraus ergibt sich sofort, dass die Norm eine stetige Abbildung ist.

Stetigkeit lässt sich nun einfach charakterisieren:

Aussage 3.25 *Seien V und W normierte Räume und $\varphi : V \longrightarrow W$ linear. Dann sind äquivalent:*

- *φ ist stetig.*
- *φ ist stetig in 0.*
- *Es gibt eine Konstante $C > 0$, so dass für alle $\mathfrak{v} \in V$ gilt: $\| \varphi(\mathfrak{v}) \|_W \leq C \cdot \| \mathfrak{v} \|_V$.*
- *Es gibt eine Konstante $C > 0$, so dass für alle $\mathfrak{v} \in V$ mit $\| \mathfrak{v} \| = 1$ gilt: $\| \varphi(\mathfrak{v}) \|_W \leq C$.*
- *Es gilt $\sup\{ \| \varphi(\mathfrak{v}) \|_W \mid \mathfrak{v} \in V \text{ mit } \| \mathfrak{v} \|_V = 1\} < \infty$.*

Definition Die Zahl $\| \varphi \| := \sup\{\| \varphi(\mathfrak{v}) \|_W \mid \mathfrak{v} \in V \text{ mit } \| \mathfrak{v} \|_V = 1\}$ heißt die *Norm* von φ.

Bemerkung Lineare Abbildungen werden auch lineare *Operatoren* genannt, die Norm einer Linearen Abbildung heißt dann auch *Operatornorm*. Die hier angeschnittenen Themen werden ausführlich in der *Funktionalanalysis* behandelt. Übrigens ist ein *Funktional* nichts anderes als eine Abbildung von einem K-Vektorraum nach K. (Bei der Verwendung des Wortes ist meist $K = \mathbb{R}$ oder $\mathbb{C}$.) Ein lineares Funktional ist also nichts anderes als eine Linearform.

Mittels des Lemmas sieht man, dass alle linearen Abbildungen zwischen $\mathbb{R}^m$ und $\mathbb{R}^n$ für beliebige m und n ausgestattet mit beliebigen Normen stetig sind. Und hieraus ergibt sich leicht:

Lemma 3.26 *Jede lineare Abbildung zwischen endlich-dimensionalen normierten Räumen ist stetig.*

Wir halten noch fest:

Lemma 3.27 *Jeder Eigenwert einer Isometrie zwischen normierten Räumen ist gleich 1 oder -1.*

Euklidische Räume

Definition Ein Euklidischer Raum ist ein $\mathbb{R}$-Vektorraum V zusammen mit einem Skalarprodukt auf V. Wie oben bezeichnet man das Tupel $(V, \langle ., . \rangle)$ oftmals einfach mit V.

Ein Euklidischer Raum definiert somit immer einen normierten Raum und damit insbesondere einen metrischen Raum. Man kann auch mittels (3.7) das Skalarprodukt aus der Norm zurückgewinnen.

Wir kommen direkt zu den entsprechenen Homomorphismen.

Definition Seien $(V, \langle ., . \rangle_V)$ und $(W, \langle ., . \rangle_W)$ Euklidische Vektorräume. Ein *Homomorphismus von Euklidischen Vektorräumen* von $(V, \langle ., . \rangle_V)$ nach $(W, \langle ., . \rangle_W)$ ist eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\langle \varphi(\mathfrak{x}), \varphi(\mathfrak{y}) \rangle_W = \langle \mathfrak{x}, \mathfrak{y} \rangle_V$ für alle $\mathfrak{x}, \mathfrak{y} \in V$.

Seien $(V, \langle ., . \rangle_V)$ und $(W, \langle ., . \rangle_W)$ Euklidische Vektorräume und sei $\varphi : V \longrightarrow W$ linear. Mittels (3.7) sieht man sofort: φ ist genau dann ein Homomorphismus von Euklidischen Räumen, wenn es ein Homomorphismus der entsprechenden normierten Räume ist. Eine solche lineare Abbildung nennt man wie schon gesagt eine *Isometrie*. Eine Isometrie zwischen Euklidischen Vektorräumen nennt man auch eine *orthogonale Abbildung*.

Wie schon bei normierten Räumen gesagt, bilden für einen Euklidischen Raum $(V, \langle ., . \rangle_V)$ die isometrischen Isomorphismen von $(V, \langle ., . \rangle_V)$ eine Untergruppe der Gruppe der Automorphismen von V als Vektorraum.

Ferner gilt:

Lemma 3.28 *Sei $V = (V, \langle ., . \rangle)$ ein Euklidischer Vektorraum und φ ein Endomorphismus des Vektorraums V. Dann ist φ genau dann ein isometrischer Isomorphismus, wenn φ bijektiv ist und φ^{-1} adjungiert zu φ ist.*

Beispiel 3.29 Wir betrachten den $\mathbb{R}^n$ mit den Standardskalarprodukt. Dann sind die orthogonalen Abbildungen (oder Isometrien) von $\mathbb{R}^n$ nach $\mathbb{R}^n$ (man kann auch sagen: die isometrischen Endomorphismen von $\mathbb{R}^n$) genau diejenigen linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^n$, welche durch orthogonale Matrizen definiert sind.

Isometrien auf einem endlich erzeugten Euklidischen Raum

Sei V ein endlich erzeugter Euklidischer Raum.

Lemma 3.30 *Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Orthonormalbasis von V. Sei W ein zweiter Euklidischer Raum und $\varphi : V \longrightarrow W$ linear. Dann ist φ genau dann eine Isometrie, wenn $\varphi(\mathfrak{b}_1), \ldots, \varphi(\mathfrak{b}_n)$ ein Orthonormalsystem in W ist.*

Der *Beweis* ist einfach.

Jede Basis $\mathfrak{B}$ von V induziert einen Isomorphismus von Vektorräumen $\mathbb{R}^n \longrightarrow V$, deren Umkehrabbildung die Koordinatenabbildung $c_{\mathfrak{B}} : V \longrightarrow \mathbb{R}^n$ ist. Wir haben nun nach dem obigen Lemma:

Lemma 3.31 *Sei $\mathfrak{B}$ eine Basis von V. Dann ist $\mathfrak{B}$ genau dann orthonormal, wenn die Isomorphismen $\mathbb{R}^n \longrightarrow V$ und $c_{\mathfrak{B}} : V \longrightarrow \mathbb{R}^n$ Isometrien sind.*

Wir studieren nun Isometrien auf dem festen Euklidischen Raum V. Sei $\mathfrak{B}$ eine Orthonormalbasis von V und $\varphi : V \longrightarrow V$ linear. Sei $A := M_{\mathfrak{B}}(V)$. Wir haben dann das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{c_{\mathfrak{B}}} & \mathbb{R}^n \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\Lambda_A} \\ V & \xrightarrow{c_{\mathfrak{B}}} & \mathbb{R}^n \end{array},$$

wobei die horizontalen Pfeile Isometrien sind. Somit ist φ genau dann eine Isometrie, wenn Λ_A eine ist. Dies ist aber genau dann der Fall, wenn A orthogonal ist.

Die orthogonalen Matrizen bilden eine Gruppe, und hier ist ein wenig Terminologie sinnvoll:

Definition Sei K ein beliebiger Körper und $n \in \mathbb{N}$.

- Die Gruppe der invertierbaren $n \times n$-Matrizen wird auch die *allgemeine lineare Gruppe (general linear group)* (vom Grad n über K) genannt und mit $\mathrm{GL}(n, K)$ bezeichnet.
- Die *spezielle lineare Gruppe* (vom Grad n über K) $\mathrm{SL}(n, K)$ ist die Untergruppe von $\mathrm{GL}(n, K)$ der Matrizen mit Determinante 1.
- Die *orthogonale Gruppe* (vom Grad n) $\mathrm{O}(n)$ ist die Untergruppe von $\mathrm{GL}(n, K)$ bestehend aus den orthogonalen $n \times n$-Matrien, also $\mathrm{O}(n) := \{A \in \mathbb{R}^{n\times n} \mid A^t A = I_n\}$.
- Die *spezielle orthogonale Gruppe* (vom Grad n) $\mathrm{SO}(n)$ ist die Gruppe der orthogonalen $n \times n$-Matrizen mit Determinante 1, also $\mathrm{SO}(n) := \mathrm{SL}(n, \mathbb{R}) \cap \mathrm{O}(n) = \{A \in \mathbb{R}^{n\times n} \mid A^t A = I_n \ , \ \mathrm{Det}(A) = 1\}$.

Beachten Sie, dass jede orthogonale Matrix Determinate 1 oder -1 hat.

Damit erhalten wir:

Aussage 3.32 *Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Orthonormalbasis von V. Dann induziert der Gruppenisomorphismus* $\mathrm{Aut}(V) \longrightarrow (\mathbb{R}^{n\times n})^* = \mathrm{GL}(n, K)$, $\varphi \mapsto M_{\mathfrak{B}}(\varphi)$ *einen Gruppenisomorphismus zwischen der Gruppe der Isometrien von V und der Gruppe der orthogonalen $n \times n$-Matrizen. (Hierbei sei* $\mathrm{Aut}(V)$ *die Gruppe der Automorphismen des Vektorraums V.)*

Wir wollen nun beweisen, dass es zu einer gegebenen Isometrie $\varphi : V \longrightarrow V$ eine Orthonormalbasis gibt bezüglich der die Abbildungsmatrix von φ besonders "schön" ist.

Wir beginnen mit einer Beschreibung der Gruppe $\mathrm{SO}(2)$.

Sei für $\alpha \in \mathbb{R}$ $D(\alpha) := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$. Die lineare Abbildung, die durch diese Matrix definiert wird, ist eine Drehung um den Winkel α. Dementsprechend heißt die Matrix *Drehmatrix* zum Winkel α. Man nennt sie auch "Drehkästchen". Diese Matrix liegt in $\mathrm{SO}(2)$, und es gilt:

Lemma 3.33 *Es ist* $\mathrm{SO}(2) = \{D(\alpha) \mid 0 \leq \alpha < 2\pi\}$.

Es ist intuitiv klar, dass sie für $\alpha \neq \pi \cdot \mathbb{Z}$ die Matrix $D(\alpha)$ keine (rellen) Eigenwerte hat. Und in der Tat ist die Diskriminante des charakteristischen Polynoms $T^2 - 2\cos(\alpha)T + 1$ gleich $\cos^2(\alpha) - 1$, und für $\alpha \neq \pi \cdot \mathbb{Z}$ ist dies < 0.

Nach den *Additionstheoremen* für Sinuns und Cosinus ist

$$D(\alpha) \cdot D(\beta) = D(\alpha + \beta)$$

für $\alpha, \beta \in \mathbb{R}$.

Somit ist die Abbildung $\mathbb{R} \longrightarrow \mathrm{SO}(2)$, $\alpha \mapsto D(\alpha)$ ein surjektiver Gruppenhomomorphismus. Der Kern ist $2\pi \cdot \mathbb{Z}$. Wir erhalten also einen Gruppenisomorphismus

$$\mathbb{R}/2\pi\mathbb{Z} \longrightarrow \mathrm{SO}(2) .$$

Wir kommen nun zur Gruppe $\mathrm{O}(2)$. Sei

$$S := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} .$$

Die entsprechende lineare Abbildung ist eine Spiegelung an der "x-Achse".

In $\mathrm{O}(2)$ liegen nun genau die Matrizen $D(\alpha)$ und $D(\alpha) \cdot S$ für $\alpha \in [0, 2\pi)$, mit anderen Worten:

$$\mathrm{O}(2) = \mathrm{SO}(2) \,\dot\cup\, \mathrm{SO}(2) \cdot S$$

Wir untersuchen die Matrizen $D(\alpha) \cdot S$ genauer. Es ist

$$D(\alpha) \cdot S = S \cdot D(-\alpha)$$

bzw.

$$D(\alpha) \cdot S \cdot D(\alpha) = S$$

für alle $\alpha \in \mathbb{R}$. Hiermit ist auch

$$D(\alpha) \cdot S = D(\frac{\alpha}{2}) \cdot S \cdot D(-\frac{\alpha}{2}) = D(-\frac{\alpha}{2})^{-1} \cdot S \cdot D(-\frac{\alpha}{2})$$

bzw.

$$D(\frac{\alpha}{2})^{-1} \cdot (D(\alpha) \cdot S) \cdot D(\frac{\alpha}{2}) = S \, .$$

Damit ist also $D(\alpha) \cdot S$ diagonalisierbar mit den Eigenwerten 1 und -1. Ferner sehen wir, dass die beiden Eigenräume orthogonal zueinander sind.

Das Ziel ist nun der Beweis des folgenden Satzes.

Satz 3.8 *Sei V ein endlich erzeugter Euklidischer Vektorraum und $\varphi : V \longrightarrow V$ eine Isometrie. Dann gibt es in V eine Orthonormalbasis $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ bezüglich welcher die Abbildungsmatrix von φ die folgende Gestalt hat:*

$$\begin{pmatrix} D(\alpha_1) & & & & & & & & \\ & \ddots & & & & & & & \\ & & D(\alpha_r) & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & -1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & -1 \end{pmatrix}$$

Hierbei sind die α_i in $(0, 2\pi) - \{\pi\}$.

Man kann den gleichen Sachverhalt auch rein matrizentheoretisch formulieren:

Satz 3.9 *Sei $A \in \mathrm{O}(n)$. Dann gibt es eine Matrix $S \in \mathrm{O}(n)$, so dass die Matrix $S^{-1}AS = S^tAS$ die im obigen Satz angegebene Form hat.*

Seien V und φ wie in Satz 3.8.

Wir betrachten zunächst den Fall, dass $\mathrm{Dim}(V) = 1$ oder $= 2$ ist. Im ersten Fall ist die Abbildungsmatrix bezüglich jeder Basis (jeden nicht-trivialen Vektors) gleich (1) oder gleich (-1). Im zweiten Fall gibt es zwei Möglichkeiten: Entweder ist $\mathrm{Det}(\varphi) = 1$. Dann ist φ bezüglich jeder Orthonormalbasis durch ein "Drehkästchen" gegeben. Oder es ist $\mathrm{Det}(\varphi) = -1$. Dann ist nach den obigen Betrachtungen φ bezüglich einer bestimmten Orthonormalbasis

durch die Matrix S gegeben (φ ist eine Spiegelung an einer bstimmten Achse). In jedem Fall gibt es eine Orthonormalbasis, so dass die Abbildungsmatrix die angegebene Form hat.

Zum Beweis des Satzes zeigen wir zwei Lemmata.

Lemma 3.34 *Sei U ein φ-invarianter Untervektorraum von V. Dann ist auch $U^\perp$ φ-invariant.*

Beweis. Es gilt nach Voraussetzung $\varphi(U) \subseteq U$, und da φ ein Isomorphismus und U endlich-dimensional ist, gilt dann auch $\varphi(U) = U$. Somit gilt auch $\varphi^{-1}(U) = U$, d.h. U ist auch φ^{-1}-invariant.

Sei nun $\mathfrak{v} \in U^\perp$. Dann gilt für alle $\mathfrak{u} \in U$: $\langle \mathfrak{u}, \varphi(\mathfrak{v})\rangle = \langle \varphi^{-1}(\mathfrak{u}), \mathfrak{v}\rangle = 0$. Dies bedeutet gerade, dass $\varphi(\mathfrak{v})$ in $U^\perp$ liegt. □

Wir wollen in V einen echten φ-invarianten Untervektorraum finden und die Zerlegung $V = U \oplus U^\perp$ betrachten.

Lemma 3.35 *Sei* $\mathrm{Dim}(V) > 0$. *Dann hat V einen 1- oder einen 2-dimensiona??-len φ-invarianten Untervektorraum.*

Dieses Lemma beweisen wir im nächsten Abschnitt.

Wir bemerken nur, dass das Lemma für ungerades n offensichtlich richtig ist. Denn: Dann hat das charakteristische Polynom ungeraden Grad, und jedes Polynom ungeraden Grades über den reellen Zahlen hat eine Nullstelle nach dem Zwischenwertsatz. Hiermit ist der Satz für $n = 3$ schon bewiesen.

Wir zeigen die Aussage per Induktion über die Vektorraumdimension. Der Induktionsanfang ist trivial. Zum Induktionsschritt: Sei $\mathrm{Dim}(V) > 0$ und $\varphi : V \longrightarrow V$ eine Isometrie. Dann hat V einen φ-invarianten Untervektorraum U von Dimension 1 oder 2. Wie oben gezeigt, gibt es in U eine Orthonormalbasis, so dass die Abbildungsmatrix von φ_U die behauptete Gestalt hat. Ferner ist $U^\perp$ auch φ-invariant nach Lemma 3.34, und $\varphi_{U^\perp}$ ist wieder eine Isometrie. Nach Induktionsvoraussetzung gibt es hierin auch eine Orthonormalbasis, so dass die Abbildungsmatrix von $\varphi_{U^\perp}$ die angegebene Gestalt hat. Beide Orthonormalbasen zusammen ergeben – bei geeigneter Anordnung der Vektoren – eine Orthonormalbasis von V, so dass die Abbildungsmatrix von φ die behauptete Gestalt hat. □

Man nennt eine orthogonale Matrix der Gestalt wie in Satz 3.8 (oder auch mit einer andere Anordnung der Drehkästchen, Einsen und "Minus Einsen") eine orthogonale Matrix in *Normalform*.

Sei nun A so eine Matrix, und seien e, f die Anzahlen der Einsen bzw. "Minus Einsen" auf der Diagonalen. Dann ist das charakteristische Polynom

von A gleich $\chi_A = \left(\prod_{i=1}^{r}(T^2 - 2\cos(\alpha)T + 1)\right) \cdot (T-1)^e \cdot (T+1)^f$. Nun ist die Faktorisierung eines normierten Polynoms über den reellen Zahlen in irreduzible normierte Faktoren "eindeutig bis auf Vertauschung". Man sieht dies leicht, indem man zu den komplexen Zahlen übergeht. Im nächsten Kapitel werden wir auch noch einen abstrakten Beweis führen, der über jedem Körper funktioniert. Es folgt, dass A durch sein charakteristisches Polynom "bis auf vertauschen der Drehkästchen" durch χ_A bestimmt ist.

Wenn nun A irgendeine Matrix in $\mathrm{O}(n)$ und $S \in (\mathbb{R}^{n\times n})^*$ ist, so dass $S^{-1}AS$ die gewünschte Gestalt hat, dann ist $\chi_{S^{-1}AS} = \chi_A$. Das heißt: "Bis auf Vertauschen der Drehkästchen oder der Einsen / 'Minus Einsen' " ist A zu genau einer Matrix in Normalform ähnlich.

Der Begriff der Normalform kommt übrigens häufiger vor (vgl. Zeilennormalform oder Jordansche Normalform in nächsten Kapitel). Man sollte stets sagen, auf was sich die "Normalform" bezieht. Hier sind es die orthogonalen Matrizen.

Selbstadjungierte Endomorphismen

Für diesen Abschnitt ist es hilfreich, wenn wir den Begriff der Orthogonalität auf beliebige symmetrische Bilinearformen ausdehnen: Wenn V ein Vektorraum, β eine symmetrische Bilinearform und $\mathfrak{v}, \mathfrak{w} \in V$ sind, heißen $\mathfrak{v}, \mathfrak{w}$ *orthogonal* bezüglich β, wenn $\beta(\mathfrak{v}, \mathfrak{w}) = 0$ gilt. Hiermit verallgemeinern sich sofort die Begriffe Orthogonalsystem und Orthogonalbasis. Wenn V endlich erzeugt ist, dann ist eine Basis von V genau dann eine Orthogonalbasis bzgl. β, wenn die Matrix von β eine Diagonalmatrix ist. Der Trägheitssatz von Sylvester lässt sich dann auch so formulieren: Zu jeder Bilinearform auf einem endlich erzeugten Vektorraum gibt es eine Orthogonalbasis.

Sei nun $V = (V, \langle ., . \rangle)$ ein beliebiger Euklidischer Raum. Wir werden nun weitere Bilinearformen auf V betrachten. Wir legen fest: "Orthogonalität" oder "Orthonormalität" bezieht sich – wenn nichts anderes gesagt wird – immer auf das fest gewählte Skalarprodukt des Euklidischen Raums. Ferner legen wir fest: Unter einem *Endomorphismus* von V verstehen wir einen Endomorphismus des Vektorraums V.

Sei nun φ ein Endomorphismus von V. Dann ist $\langle ., \varphi(.)\rangle : V \times V \longrightarrow \mathbb{R}$ eine Bilinearform zwischen V und V. Wir fragen uns, wann wir wieder eine symmetrische Bilinearform haben.

Zunächst eine Vorbemerkung: Da $\langle ., . \rangle$ nicht-ausgeartet ist, ist die Zuordnung $\mathrm{End}_{\mathbb{R}}(V) \longrightarrow \mathrm{Bil}_{\mathbb{R}}(V, V)$ $\varphi \mapsto \langle ., \varphi(.)\rangle$ injektiv. Deshalb hat jeder Endomorphismus von V höchstens einen rechts-adjunigierten und höchstens einen links-adjungierten Endomorphismus. Da $\langle ., . \rangle$ symmetrisch ist, sind diese – wenn sie exisitieren – gleich.

Lemma 3.36 *Die folgenden Aussagen sind äquivalent:*

a) $\langle ., \varphi(.)\rangle$ *ist symmetrisch.*

b) Für alle $\mathfrak{x}, \mathfrak{y} \in V$ *ist* $\langle \mathfrak{x}, \varphi(\mathfrak{y})\rangle = \langle \varphi(\mathfrak{x}), \mathfrak{y}\rangle$.

c) Der zu φ *adjungierte Endomorphismus* φ^* *existiert und ist gleich* φ.

Definition Ein Endomorphismus von V, der die Bedingungen im obigen Lemma erfüllt, heißt *selbstadjungiert*.

Die selbstadjunigerten Endomorphismen bilden offensichtlich einen Untervektorraum und einen Unterring in $\mathrm{End}_{\mathbb{R}}(V)$.

Sei nun V wieder endlich erzeugt und der Dimension n und $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ eine Orthonormalbasis von V.

Nach Aussage 3.10 ist die Abbildung $\mathrm{End}_{\mathbb{R}}(V) \longrightarrow \mathrm{Bil}_{\mathbb{R}}(V, V)$, $\varphi \mapsto \langle ., \varphi(.)\rangle$ ein Isomorphismus von Vektorräumen. Dieser Isomorphismus induziert nun eine Bijektion zwischen dem Vektorraum der selbstadjungierten Endomorphismen von V und dem Vektorraum der symmetrischen Bilinearformen auf V.

Mit Matrizen erhält man mit den Überlegungen am Ende von Abschnitt 3.1 das Folgende (wobei entscheidend ist, dass wir eine Orthonormalbasis zugrundelegen):

Sei $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$. Dann ist die Matrix der Bilinearform $\langle ., \varphi(.)\rangle$ gleich $M_{\mathfrak{B}}(\varphi)$. Somit ist also die Bilinearform genau dann symmetrisch, wenn die Matrix $M_{\mathfrak{B}}(\varphi)$ symmetrisch ist.

Ferner ist

$$M_{\mathfrak{B}}(\varphi^*) = M_{\mathfrak{B}}(\varphi)^t \ .$$

Somit ist φ also genau dann selbstadjungiert, wenn $M_{\mathfrak{B}}(\varphi)$ symmetrisch ist. Es passt also alles zusammen.

Überlegen Sie sich, dass auch alles zusammenpasst, wenn man einen Basiswechsel von einer Orthonormalbasis zu einer anderen Orthonormalbasis betrachtet!

Sei nun $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$ selbstadjungiert und $\beta := \langle ., \varphi(.)\rangle$ die entsprechende symmetrische Bilinearform. Sei ferner $B := M_{\mathfrak{B}}(\varphi)$. Dann ist ja wie gesagt B auch die Matrix der Bilinearform β.

Da B sowohl die Abbildungsmatrix von φ als auch die Matrix von β ist, sind äquivalent:

1. Die Basis $\mathfrak{B}$ besteht aus Eigenvektoren von φ.

2. Die Matrix B ist eine Diagonalmatrix.

3. Die Basis $\mathfrak{B}$ ist eine Orthogonalbasis von β.

Ich betone hier nochmal, dass $\mathfrak{B}$ nach Voraussetzung eine Orthonormalbasis bezüglich des festen Skalarprodukts $\langle ., . \rangle$ ist.

Es gilt nun der folgende bemerkenswerte Satz:

Satz 3.10 *Sei $\varphi \in \mathrm{End}_{\mathbb{R}}(V)$ selbstadjungiert. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von φ.*

Nach den obigen Bemerkungen zum Zusammenhang zwischen selbstadjungierten Endomorphismen und symmetrischen Bilinarformen auf V haben wir auch:

Satz 3.11 *Sei β eine symmetrische Bilinearform auf V. Dann gibt es eine Orthonormalbasis des Euklidischen Vektorraums V, die bezüglich β orthogonal ist.*

Beachten Sie hier nochmal, dass wir in diesem Satz zwei Bilinearformen betrachten, die schon von Anfang an gegebene positiv definite Bilinearform $\langle ., . \rangle$ des Euklidischen Vektorraums und die Bilinearform β. Die Behauptung ist, dass es eine Basis von V gibt, die orthonormal bezüglich $\langle ., . \rangle$ und orthogonal bezüglich β ist.

Wenn man dies auf das Standardskalarprodukt anwendet, erhält man:

Satz 3.12 *Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann gibt es eine orthogonale Matrix S, so dass $S^{-1}AS = S^tAS$ eine Diagonalmatrix ist.*

Man beachte hier, dass für eine Orthogonalmatrix S gilt: $S^{-1} = S^t$. Man kann den letzten Satz natürlich rein matrizentheoretisch betrachten. Aber mit der Theorie wird klar, dass hier eigentlich zwei verschiedene Sachverhalte in einem betrachtet werden, je nachdem ob man A als Matrix einer linearen Abbildung oder einer Bilinearform interpretiert.

Beweis von Satz 3.10 Der Beweis ist analog zum Beweis von Satz 3.8.

Sei φ ein selbstadjungierter Endomorphismus von V.

Lemma 3.37 *Sei U ein φ-invarianter Untevektorraum von V. Dann ist $U^\perp$ auch φ-invariant.*

Beweis. Sei U φ-invariant und sei $\mathfrak{v} \in U^\perp$ Wir wollen zeigen, dass $\varphi(\mathfrak{v}) \in U^\perp$ ist. Sei hierzu $\mathfrak{u} \in U$. Dann ist $\langle \varphi(\mathfrak{v}), \mathfrak{u} \rangle = \langle \mathfrak{v}, \varphi(\mathfrak{u}) \rangle = 0$. □

Ferner gilt:

Lemma 3.38 *φ hat einen Eigenwert.*

Dies zeigen wir wieder im nächsten Abschnitt.

Nun folgt der Satz wie Satz 3.8 per Induktion über die Dimension. □

3.5 Unitäre Vektorräume

Grundlegende Begriffe und Resultate

Wir betrachten in diesem Abschnitt Vektorräume über den komplexen Zahlen. Zunächst einige Begriffe zu den komplexen Zahlen.

Wir legen die folgende Konvention fest: Wenn wir eine komplexe Zahl z in der Form $z = r + is$ schreiben, dann werden r und s immer als reell angenommen. Also: r ist der so genannte *Realteil* und s der so genannte *Imaginärteil* der Zahl. Man schreibt $r = \mathrm{Re}(z), s = \mathrm{Im}(z)$.

Definition Sei $z = r + si$ eine komplexe Zahl. Dann ist die zu z *konjugierte* komplexe Zahl $\overline{z} := r - si$.

Die so genannte *komplexe Konjugation* $z \mapsto \overline{z}$ ist ein Körperautomorphismus von $\mathbb{C}$.

Wir haben für $z = r + is \in \mathbb{C}$ $\mathrm{Re}(z) = \frac{1}{2} \cdot (z + \overline{z})$ und $\mathrm{Im}(z) = \frac{1}{2i}(z - \overline{z}) = \frac{-i}{2}(z - \overline{z})$.

Definition Die *Norm* oder der *Absolutbetrag* einer komplexen Zahl $z = r + si$ ist $|z| := \sqrt{\overline{z}z} = \sqrt{r^2 + s^r}$.

Beachten Sie, dass dies genau die übliche 2-Norm ist, wenn wir den $\mathbb{R}$-Vektorraum $\mathbb{C}$ mit $\mathbb{R}^2$ identifizieren.

Die komplexe Konjugation induziert auch eine Abbildung $\mathbb{C}^n \longrightarrow \mathbb{C}^n$, $\underline{x} \mapsto \overline{\underline{x}} := \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix}$, die auch *komplexe Konjugation* genannt wird.

Im $\mathbb{C}^n$ haben wir nun auch ein *Standardskalarprodukt*. Dieses ist durch

$$\langle \underline{x}, \underline{y} \rangle := \sum_{i=1}^{n} \overline{x}_i y_i = \overline{\underline{x}}^t \underline{y}$$

gegeben. Für $\underline{x} \in \mathbb{C}^n$ ist also $\langle \underline{x}, \underline{x} \rangle = \sum_{i=1}^{n} |\underline{x}_i|^2$; dies liegt natürlich immer in $\mathbb{R}_{\geq 0}$.

Die komplexe Konjugation $\mathbb{C}^n \longrightarrow \mathbb{C}^n$ ist nicht $\mathbb{C}$-linear sondern nur $\mathbb{R}$-linear, und das Standardskalarprodukt ist auch nicht $\mathbb{C}$-linear in der ersten Variablen sondern nur $\mathbb{R}$-linear.

Beachten Sie hierbei und im Folgenden, dass jeder $\mathbb{C}$-Vektorraum durch Einschränkung der Skalarmultiplikation zu einem $\mathbb{R}$-Vektorraum wird. Wenn V und W $\mathbb{C}$-Vektorräume sind, dann verstehen wir unter einer $\mathbb{R}$-*linearen* Abbildung von V nach W eine Abbildung $\varphi : V \longrightarrow W$, die ein Homomorphismus ist, wenn wir V und W als $\mathbb{R}$-Vektorräume auffassen.

Das Standardskalarprodukt motiviert die folgenden Definitionen.

Definition Sei V ein komplexer Vektorraum.

a) Sei W ein weiterer komplexer Vektorräume. Eine $\mathbb{R}$-lineare Abbildung $\varphi : V \longrightarrow W$ ist *semilinear*, falls für alle $a \in \mathbb{C}$ und alle $\mathfrak{v} \in V$ gilt: $\varphi(a\mathfrak{v}) = \overline{a}\,\varphi(\mathfrak{v})$.

b) Eine *Sesquilinearform* auf V ist eine Abbildung $\beta : V \times V \longrightarrow \mathbb{C}$, die semilinear bezüglich der ersten Variablen und linear bezüglich der zweiten Variablen ist. Dies bedeutet: Für alle $\mathfrak{v} \in V$ ist die Abbildung $\beta(\mathfrak{v}, .)$ eine Linearform und die Abbildung $\beta(., \mathfrak{v})$ eine Semilinearform.

c) Eine Sesquilinearform β ist *hermitesch*, falls für alle $\mathfrak{v}, \mathfrak{w} \in V$ gilt: $\beta(\mathfrak{v}, \mathfrak{w}) = \overline{\beta(\mathfrak{w}, \mathfrak{v})}$. Eine hermitesche Sesquilinearform nennt man auch kurz eine *hermitesche Form*.

Für eine hermitesche Form $\langle ., . \rangle$ auf einem komplexen Vektorraum V und $\mathfrak{v} \in V$ ist somit $\langle \mathfrak{v}, \mathfrak{v} \rangle \in \mathbb{R}$. Wir definieren:

Definition Sei nun β eine hermitesche Sesquilinearform auf V. Dann ist β *positiv definit*, falls für alle $\mathfrak{v} \in V - \{\mathfrak{o}_V\}$ gilt: $\beta(\mathfrak{v}, \mathfrak{v}) > 0$. Eine positiv definite hermitesche Form nennt man auch ein *Skalarprodukt* auf V.

Notation Skalarprodukte bezeichnen wir wie zuvor mit $\langle ., . \rangle$.

Definition

- Ein *unitärer Raum* ist ein $\mathbb{C}$-Vektorraum V zusammen mit einem Skalarprodukt auf V.
- Seien V und W unitäre Räume. Dann ist ein *Homomorphismus von unitären Räumen* oder eine *unitäre Abbildung* von V nach W eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\langle \varphi(\mathfrak{x}), \varphi(\mathfrak{y}) \rangle_W = \langle \mathfrak{x}, \mathfrak{y} \rangle_V$ für alle $\mathfrak{x}, \mathfrak{y} \in V$.

Beispiel 3.39 Wir betrachten den $\mathbb{C}^n$ mit dem Standardskalarprodukt. Sei $A \in \mathbb{C}^{n\times n}$. Dann ist die Abbildung $\Lambda_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ genau dann unitär, wenn $\overline{A}^t A = I_n$ gilt.

Beispiel 3.40 Sei B eine Matrix in $\mathbb{C}^{n\times n}$. Dann ist die Abbildung $\mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$, $(\underline{x}, \underline{y}) \mapsto \overline{\underline{x}}^t B\underline{y}$ eine Sesquilinearform. Die Sesquilinearform ist genau dann hermitesch, wenn $\overline{B}^t = B$ ist.

Sei dies der Fall. Dann ist die Sesquilinearform genau dann positiv definit, wenn für alle $\underline{x} \in \mathbb{C}^n - \{\overline{0}\}$ $\overline{\underline{x}}^t B\underline{x} > 0$ gilt.

Diese Beispiele motivieren:

Definition

- Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist *unitär*, falls $\overline{A}^t A = I$ gilt. Dies ist natürlich äquivalent dazu, dass A^{-1} existiert und gleich $\overline{A}^t$ ist.
- Die *unitäre Gruppe* (vom Grad n) $U(n)$ ist die Gruppe aller unitären $n \times n$-Matrizen.

Definition Eine *hermitesche Matrix* ist eine quadratische Matrix B über $\mathbb{C}$ mit $\overline{B}^t = B$. Eine hermitesche Matrix B ist *positiv definit*, wenn für alle $\underline{x} \in \mathbb{C}^n - \{\underline{0}\}$ $\overline{\underline{x}}^t B\underline{x} > 0$ gilt.

Bemerkung Eine hermitesche Matrix hat nur reelle Zahlen auf der Diagonalen. Eine positiv definite hermitesche Matrix hat nur positive reelle Zahl auf der Diagonalen.

Sei nun $V = (V, \langle ., . \rangle)$ ein unitärer Vektorraum. Dann definieren wir für ein $\mathfrak{v} \in V$ die *Norm* von $\mathfrak{v}$ als $\| \mathfrak{v} \| := \sqrt{\langle \mathfrak{v}, \mathfrak{v} \rangle} \in \mathbb{R}_{\geq 0}$.

Die *Norm* $\| \cdot \| : V \longrightarrow \mathbb{R}_{\geq 0}$ erfüllt dann für $c \in \mathbb{C}$ und $\mathfrak{v}, \mathfrak{w} \in V$:

a) $\| \mathfrak{v}|| = 0 \longleftrightarrow \mathfrak{v} = \mathfrak{o}_V$

b) $\| c \cdot \mathfrak{v} \| = |c| \cdot \| \mathfrak{v} \|$

c) $\| \mathfrak{v} + \mathfrak{w} \| \leq \| \mathfrak{v} \| + \| \mathfrak{w} \|$

Diese Eigenschaften sind natürlich vollkommen analog zu den Eigenschaften einer Norm eines $\mathbb{R}$-Vektorraums.

Nun kann man wie gesagt jeden $\mathbb{C}$-Vektorraum durch Einschränkung der Skalarmultiplikation zu einem $\mathbb{R}$-Vektorraum machen. Und man sieht jetzt:

Wenn wir V zu einem $\mathbb{R}$-Vektorraum machen, wird die oben definierte Normabbildung zu einer Norm im Sinne von $\mathbb{R}$-Vektorräumen. Also wird dann V zu einem normierten Raum (über $\mathbb{R}$). Man kann recht leicht zeigen, dass für unitäre Räume V, W und lineare Abbildungen $\varphi : V \longrightarrow W$ das Folgende gilt: φ ist genau dann eine unitäre Abbildung, wenn φ eine Isometrie der durch V und W definierten normierten Räume (über $\mathbb{R}$) ist (Übungsaufgabe). Dies rechtfertigt, eine unitäre Abbildung auch eine *Isometrie* zu nennen.

Achtung: Ein nicht-trivialer unitärer Vektorraum wird durch Einschränkung der Skalarmultiplikation *nicht* zu einem Euklidischen Raum. Das Bild des Skalarprodukts ist nämlich immer noch der Körper der komplexen Zahlen. Es gilt allerdings: Wenn V ein unitärer Vektorraum ist, dann wird V durch Einschränkung der Skalarmultiplikation zusammmen mit der Abbildung $\mathrm{Re}(\langle .,. \rangle) : V \times V \longrightarrow \mathbb{R}\,, (\mathfrak{v}, \mathfrak{w}) \mapsto \mathrm{Re}(\langle \mathfrak{v}, \mathfrak{w} \rangle)$ ein Euklidischer Vektorraum.

Nun übertragen sich viele weitere Begriffe von Euklidischen Räumen auf unitäre Räume. Beispielsweise haben wir wieder die Begriffe der Orthogonalität, des Orthogonalsystems, des Orthonormalsystems, der Orthogonal- und der Orthonormalbasis und des orthogonalen Komplements.

Auch viele Aussagen über Euklidische Räume übertragen sich direkt auf unitäre Räume. Man sollte nur die folgenden allgemeinen Regeln beachten. *(Die Details sind natürlich bei jeder Anwendung zu überprüfen.)*

- Eine Sesquilinearform ist per Definition nicht linear sondern semilinear in der ersten Variablen.
- Wenn man zuvor einen Vektor oder eine Matrix transponieren muss, muss man jetzt transponieren und konjugieren.
- Insbesondere müssen orthogonale Matrizen durch unitäre Matrizen und symmetrische Matrizen durch hermitesche Matrizen ersetzt werden.
- Beim Bilden des Skalarprodukts kommt es nun auf die Reihenfolge der Vektoren an.

So übertragen sich beispielsweise so gut wie alle Aussagen aus Abschnitt 3.3 auf unitäre Räume. Die einzige Ausnahme ist, dass die Abbildung $\Psi : V \longrightarrow V^*\,,\ \mathfrak{v} \mapsto \langle \mathfrak{v}, . \rangle$ nun semilinear ist, aber das spielt keine Rolle. (Allgemein ist das Bild einer semilinearen Abbildung auch ein $\mathbb{C}$-Untervektorraum.)

Beispielsweise hat man in endlich-dimensionalen unitären Vektorräumen wieder das Gram-Schmidt Orthogonalisierungsverfahren und jeder solche Raum hat eine Orthonormalbasis.

Lineare Abbildungen auf endlich erzeugten unitären Räumen

Wir studieren Isometrien und selbstadjungiere Abbildungen auf einem endlich erzeugten unitären Raum und später eine gemeinsame Verallgemeinerung dieser beiden Arten linearer Abbildungen. Außerdem schließen wir noch die Lücken in den Beweisen der Sätze des vorherigen Abschnitts.

Sei im Folgenden V ein endlich erzeugter unitärer Raum.

Das folgende Lemma ist offensichtlich (und gilt nicht nur in endlich erzeugten unitären Räumen).

Lemma 3.41 *Jeder Eigenwert einer unitären Abbildung hat Absolutbetrag 1.*

Satz 3.13 *Sei φ ein Endomorphismus des $\mathbb{C}$-Vektorraums V. Dann ist φ genau dann eine unitäre Abbildung, wenn es eine Orthonormalbasis von V gibt bezüglich welcher die Abbildungsmatrix von φ eine Diagonalmatrix ist, wobei alle Diagonalelemente Absolutbetrag 1 haben.*

Oder bezüglich Matrizen:

Satz 3.14 *Sei $A \in \mathbb{C}^{n\times n}$. Dann ist A genau dann unitär, wenn es eine unitäre Matrix S gibt, so dass $S^{-1}AS = \overline{S}^t AS$ eine Diagonalmatrix ist, deren Diagonalelemente Absolutbetrag 1 haben.*

Beweis. Es ist klar, dass linaere Abbildungen, die wie angegeben dargestellt werden können, unitär sind.

Sei also φ unitär. Genau wie im Falle orthogonaler Abbildungen gilt: Wenn U ein φ-invarianter Untervektorraum von V ist, dann ist auch $U^\perp$ φ-invariant. Anders als im reellen hat aber nun jede lineare Abbildung einen Eigenwert. Und ein Eigenwert von φ hat den Absolutbetrag 1. Hiermit folgt die Behauptung per Induktion über die Dimension. □

Wir kommen zur Lücke im Beweis von Satz 3.8 (und Satz 3.9).

Hierzu kommen wir zunächst nochmal auf die Gruppe $\mathrm{SO}(2)$ zurück. Diese Gruppe ist offensichtlich eine Untergruppe von $U(2)$. (Allgemein ist $\mathrm{O}(n)$ eine Untergruppe von $U(n)$ für $n \in \mathbb{N}$.) Wir wissen schon, dass alle Matrizen in $U(2)$ diagonalisierbar sind, es stellt sich also die Frage, wie das hier explizit aussieht.

Wir setzen die folgende Formel voraus: Für $\alpha \in \mathbb{R}$ ist

$$\cos(\alpha) + \sin(\alpha)i = e^{\alpha i} \ .$$

Wenn Sie wollen, können Sie diese Formel als Definition von $e^{\alpha i}$ nehmen. Die Additionstheoreme für Sinus und Cosinus sind nun äquivalent zur Aussage, dass $e^{\alpha i} \cdot e^{\beta i} = e^{(\alpha+\beta)i}$ für alle $\alpha, \beta \in \mathbb{R}$.

Als Matrix über den komplexen Zahlen hat $D(\alpha)$ die Eigenwerte $e^{\alpha i}$ und $e^{-\alpha i}$, die beiden Eigenräume sind $\langle \begin{pmatrix} 1 \\ -i \end{pmatrix} \rangle$ bzw. $\langle \begin{pmatrix} 1 \\ i \end{pmatrix} \rangle$. Es ist also

$$\frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \cdot D(\alpha) \cdot \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} = \begin{pmatrix} e^{\alpha i} & \\ & e^{-\alpha i} \end{pmatrix}$$

beziehungsweise

$$D(\alpha) = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \cdot \begin{pmatrix} e^{\alpha i} & \\ & e^{-\alpha i} \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \ . \qquad (3.10)$$

Wir beweisen nun Lemma 3.35. Genauer beweisen wir das folgende explizitere Lemma. Lemma 3.35 folgt dann, indem man eine beliebige Orthonormalbasis wählt und die entsprechende Koordinatenmatrix betrachtet.

Lemma 3.42 *Sei $A \in \mathrm{O}(n)$ für $n \in \mathbb{N}$. Dann hat $\mathbb{R}^n$ einen 1- oder einen 2-dimensionalen φ-invarianten Untervektorraum.*

Beweis. A habe keinen Eigenwert. Wir betrachten nun A als komplexe Matrix. Dann hat A einen Eigenwert, sei a so ein Eigenwert. Da $|a| = 1$ und $a \neq 1, -1$ ist, gibt es ein (eindeutig bestimmtes) $\alpha \in (0, 2\pi) - \{\pi\}$ mit $a = e^{i\alpha}$.

Sei $\underline{b}$ ein Eigenvektor zu a. Dann ist $A\overline{\underline{b}} = \overline{A\underline{b}} = \overline{a\underline{b}} = \overline{a} \cdot \overline{b}$. Mit anderen Worten: $\overline{a}$ ist auch ein Eigenwert und $\overline{\underline{b}}$ ist ein entsprechender Eigenvektor. Sei $U := \langle \underline{b}, \overline{\underline{b}} \rangle_{\mathbb{C}}$.

Wir betrachten nun die Vektoren $\underline{b}_1 := \frac{1}{\sqrt{2}} \cdot (\underline{b} + \overline{\underline{b}})$ sowie $\underline{b}_2 := \frac{1}{\sqrt{2}} \cdot (i\underline{b} - i\overline{\underline{b}})$. Dies ist durch den Basiswechsel in (3.10) motiviert. Diese Vektoren bilden wieder eine Orthonormalbasis von U. Ferner liegen sie in $\mathbb{R}^n$. Nach (3.10) ist nun die Abbildungsmatrix von φ_U genau $D(\alpha)$. Hieraus sieht man, dass der $\mathbb{R}$-Vektorraum $\langle \underline{b}_1, \underline{b}_2 \rangle_{\mathbb{R}}$ φ-invariant ist.

Übrigens kann man auch leicht zeigen, dass $U \cap \mathbb{R}^n = \langle \underline{b}_1, \underline{b}_2 \rangle_{\mathbb{R}}$ ist. □

Bemerkung Mit den Ideen im obigen Lemma hätten wir auch Satz 3.8 direkt aus Satz 3.13 ableiten können. Auf jeden Fall liefert das Lemma eine Konstruktion, wie man von einer Basis wie in Satz 3.13 zu einer Basis wie in Satz 3.8 kommt und umgekehrt.

Wir kommen nun zu selbstadjungierten Endomorphismen des Vektorraums V. Die Definition ist wie im reellen Fall. Der Hauptsatz ist nun wie folgt:

Satz 3.15 *Sei φ ein Endomorphismus des $\mathbb{C}$-Vektorraums V. Dann ist φ genau dann selbstadjungiert, wenn es eine Orthonormalbasis von V gibt bezüglich welcher φ eine Diagonalmatrix mit reellen Diagonaleinträgen ist.*

Für Matrizen bedeutet dies:

Satz 3.16 *Sei $A \in \mathbb{C}^{n\times n}$. Dann ist A genau dann hermitesch, wenn es eine unitäre Matrix S gibt, so dass $S^{-1}AS = \overline{S}^t AS$ eine Diagonalmatrix mit reellen Diagonaleinträgen ist.*

Beweis. Wiederum ist klar, dass eine lineare Abbildung mit einer wie beschriebenen Abbildungsmatrix selbstadjungiert ist.

Sei also φ selbstadjungiert. Nun ist wiederum das orthogonale Komplement eines φ-invarianten Untervektorraums φ-invariant. Ferner hat φ als Endomorphismus eines komplexen endlich erzeugten Vektorraums einen Eigenwert. Sei a so ein Eigenwert und $\mathfrak{v}$ ein entsprechender Eigenvektor. Dann ist $a \cdot \langle \mathfrak{v}, \mathfrak{v}\rangle = \langle \mathfrak{v}, \varphi(\mathfrak{v})\rangle = \langle \varphi(\mathfrak{v}), \mathfrak{v}\rangle = \overline{\langle \mathfrak{v}, \varphi(\mathfrak{v})\rangle} = \overline{a \cdot \langle \mathfrak{v}, \mathfrak{v}\rangle} = \overline{a} \cdot \langle \mathfrak{v}, \mathfrak{v}\rangle$. Also ist $a = \overline{a}$ und somit a reell. □

Wir kommen noch zu Lemma 3.38. Analog zum Beweis von Lemma 3.35 (siehe Lemma 3.42) folgt das Lemma aus der folgenden expliziteren Aussage.

Lemma 3.43 *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann hat A einen reellen Eigenwert.*

Der *Beweis* hiervon wurde soeben schon gegeben.

Wir kommen nun zu einer gemeinsamen Verallgemeinerung von unitären und selbstadjungierten Abbildungen.

Definition Sei φ ein Endomorphismus des $\mathbb{C}$-Vektorraums V. Dann ist φ *normal*, falls $\varphi^* \circ \varphi = \varphi \circ \varphi^*$ gilt. (Wir brauchen hier nicht, dasss V endlich erzeugt sei.)

Selbstadjungierte Endomorphismen sind sicher normal. Unitäre Abbildungen aber auch. Diese erfüllen nämlich $\varphi^* = \varphi^{-1}$.

Wir haben nun:

Satz 3.17 *Sei φ ein Endomorphismus des $\mathbb{C}$-Vektorraums V. Dann ist φ genau dann normal, wenn es eine Orthonormalbasis von V gibt bezüglich der die Abbildungsmatrix von φ eine Diagonalmatrix ist.*

Zunächst wieder einige Lemmata.

Lemma 3.44 *Sei V ein endlich erzeuger unitärer Vektorraum und φ ein beliebiger Endomorphismus des -Vektorraums V.*

a) Sei U ein φ-invarianter Untervektorraum von V. Dann ist $U^\perp$ φ^*-invariant.

b) Sei U ein φ^*-invarianter Untervektorraum von V. Dann ist $U^\perp$ φ-invariant.

c) Sei U ein φ und φ^*-invarianter Untervektorraum von V. Dann ist die adjungierte Abbildung zu φ_U gerade $(\varphi^*)_U$, und die zu $\varphi_{U^\perp}$ adjungierte Abbildung ist gerade $(\varphi^*)_{U^\perp}$.

d) Sei φ normal und U ein φ und φ^*-invarianter Untervektorraum von V. Dann ist sowohl φ_U also auch $\varphi_{U^\perp}$ wieder normal.

Dies ist einfach.

Lemma 3.45 *Sei φ ein normaler Endomorphismus des $\mathbb{C}$-Vektorraums V und sei a ein Eigenwert von φ. Sei $E_{a,\varphi}$ der Eigenraum zum Eigenwert a. Dann ist $E_{a,\varphi}$ φ^*-invariant.*

Beweis. Sei $\mathfrak{v}$ ein Eigenvektor von φ zum Eigenwert a. Dann gilt $\varphi(\varphi^*(\mathfrak{v})) = \varphi^*(\varphi(\mathfrak{v})) = \varphi^*(a\mathfrak{v}) = a \cdot \varphi^*(\mathfrak{v})$. Somit ist $\varphi^*(\mathfrak{v}) \in E_{a,\varphi}$. □

Bemerkung Dieses Lemma gilt analog für zwei beliebige "kommutierende Endomorphismen" auf einem Vektorraum. Im obigen Kontext folgt auch noch mehr: Da wir V als endlich erzeugt vorausgesetzt haben, hat der Endomorphismus $\varphi^*_{E_{a,\varphi}}$ von $E_{a,\varphi}$ hat auch einen Eigenwert. Man kann nun mittels des Skalarprodukts leicht zeigen, dass so ein Eigenwert gleich $\overline{a}$ ist. Also: Es gibt einen Vektor $\mathfrak{v} \in V$, der sowohl ein Eigenvektor von φ zum Eigenwert a als auch ein Eigenvektor von φ^* zum Eigenwert $\overline{a}$ ist.

Aus dem obigen Lemma sieht man, dass V einen nicht-trivialen φ- und φ^*-invarianten Untervektorraum U enthält. Aus Lemma 3.44 d) folgt dann, dass φ_U und $\varphi_{U^\perp}$ wieder normal sind.

Mit diesen Bemerkungen folgt der Satz per Induktion über die Dimension. □

Die Notation der Mathematiker und der Physiker

Abschließend noch einige Bemerkungen zum Begriff des Skalarprodukts in der Mathematik und der Physik und zu speziellen Notationen in der Physik.

In der Linaren Algebra werden Sesquilinarformen normalerweise so definiert, dass sie linear in der 1.Variablen und semilinear in der 2.Variablen sind. Dies ist also umgekehrt zu der obigen Notation.

Skalarprodukte kommen auch in der Physik vor, besonders in der Quantenmechanik, und dort werden sie so definiert wie oben.

Mir erscheint die "Physikerkonvention" natürlicher. Ein Grund ist, dass dann beim Standardskalarprodukt links ein transponierter und konjugierter Vektor steht.

In der Quantenmechanik sind Skalarprodukte auf unendlich-dimensionalen $\mathbb{C}$-Vektorräumen und (partiell definierte) selbstadjungierte Operatoren grundlegend. Um den Kontext zu erläutern, gehe ich kurz auf einige physikalische Aspekte ein: Der Zustand eines Teilchensystems wird durch einen Vektor in einem "großen" unitären Vektorraum V beschrieben. Zu jeder Messgröße korrespondiert ein (partiell definierter) selbstadungierter Operator. Diese Operatoren sind leider oftmals noch nicht mal stetig, was ihre mathematische Beschreibung sehr schwierig macht.

Sei zunächst V ganz allgemein ein unitärer Vektorraum. Dann haben wir die semilineare Abbildung $\Psi : V \longrightarrow V^*\ ,\ \mathfrak{v} \mapsto \langle \mathfrak{v}, . \rangle$. Wir können auch aus Ψ in eindeutiger Weise wieder das Skalarprodukt zurückgewinnen. Es ist ja $\langle \mathfrak{v}, \mathfrak{w} \rangle = \langle \Psi(\mathfrak{v}), \mathfrak{w} \rangle_{\text{kan}}$.

Zurück zur Quantenmechanik. Hier sollte man das Skalarprodukt auf V am besten vergessen und nur Ψ im Kopf behalten. Man bezeichnet nun einen "Teilchenzustand", d.h. einen Vektor in V, mit $|\psi\rangle$. Das Bild von $|\psi\rangle$ unter Ψ bezeichnet man mit $\langle\psi|$. Dies ist eben eine Linearform, und man kann sie auf einen "Zustand" $|\psi'\rangle$ anwenden (Dualpodukt). Man erhält die komplexe Zahl $\langle\psi|\psi'\rangle$.

Sei nun H ein (partiell definierter) selbstadjungierter (oder wie man auch sagt hermitescher) Operator auf V. Dann ist ja $\langle H(\psi)|\psi'\rangle = \langle\psi|H(\psi')\rangle$, wann immer das Sinn macht. Um zu betonen, dass H eine neue Sesquilinearform definiert, schreibt man H auch in der Mitte, also $\langle\psi|H|\psi'\rangle$.

Was ich soeben erklärt habe, ist die so genannte "Bra-Ket" Notation der Quantenmechanik.

Kapitel 4

Euklidische Ringe und die Jordansche Normalform

4.1 Euklidische Ringe

Die Ringe der ganzen Zahlen, $\mathbb{Z}$, sowie Polynomringe über Körpern, $K[X]$, wobei K ein Körper ist, haben die folgenden Gemeinsamheiten: Erstens sind sie kommutativ und "nicht-trivial". Zweitens ist das Produkt zweier Elemente ungleich Null immer ungleich Null. Drittens, und das ist wirklich das Besondere, gibt es "Größenfunktion" oder – wie man allgemein sagt – *Gradfunktion* (den Absolutbetrag für $\mathbb{Z}$ und den Grad für $K[X]$). Bezüglich dieser Gradfunktion gibt es eine Division mit Rest und ist das Produkt zweier Element ist größer-gleich den Faktoren.

Einen Ring mit diesen Eigenschaften nennt man einen *Euklidischen Ring*. In diesem Abschnitt geht es hauptsächlich um solche Ringe.

Ideale und Teilbarkeit

Wir beginnen mit einer grundlegenden Betrachtung über Ringe:

Sei R ein Ring und U eine Untegruppe von $(R, +)$. Dann haben wir ja die Faktorgruppe R/U (mit der induzierten Addition $[a]_U + [b]_U = [a+b]_U$). Es stellt sich die folgende Frage: Wann gibt es eine Verknüpfung $\cdot : R/U \times R/U \longrightarrow R/U$ mit $[a]_U \cdot [b]_U = [a \cdot b]_U$ für alle $a, b \in R$? Man sagt: Wann *induziert* die Multiplikation auf R eine Verknüpfung auf R/U? Die Antwort liefert die folgende Aussage.

Aussage 4.1 *Seien R und U wie oben. Dann induziert die Multiplikation auf R genau dann eine Verknüpfung auf R/U, wenn*

$$\forall r \in R \, \forall a \in U : ra \in U \wedge ar \in U \; .$$

Falls dies der Fall ist, ist R/U mit der induzierten Addition und der Verknüpfung $\cdot$ als Multiplikation ein Ring, und die kanonische Abbildung $R \longrightarrow R/U\ , r \mapsto [r]_U$ ist ein Ringhomomorphismus.

Beweis. Es existiere zunächst die Verknüpfung. Dann gilt für $a \in U$ und $r \in R$: $[ar]_U = [a]_U \cdot [r]_U = [0]_U \cdot [r]_U = [0 \cdot r]_U = [0]_U$ und damit $ar \in U$. Analog zeigt man $ra \in U$.

Es gelte nun die angegebene Bedingung. Seien $a, a' \in R$ mit $a \sim_U a'$ und $b \sim_U b'$. Dann ist $ab - a'b' = ab - a'b + a'b - a'b' = (a - a') \cdot b + a'(b - b') \in U$, da nach Voraussetzung $(a - a') \cdot b$ und $a \cdot (b - b')$ in U liegen.

Die weiteren Aussagen sind einfach. □

Definition Eine Untergruppe I von R mit der Eigenschaft

$$\forall r \in R\, \forall a \in I : ra \in I \wedge ar \in I$$

heißt *Ideal* von R. Wenn I ein Ideal von R ist, heißt der Ring R/I *Restklassenring von R modulo I*.

Sei nun S ein weiterer Ring und $\varphi : R \longrightarrow S$ ein Ringhomomorphismus. Der *Kern* von φ ist per Definition der Kern von φ aufgefasst als Homomorphismus von abelschen Gruppen, d.h. $\mathrm{Kern}(\varphi) = \{r \in R \mid \varphi(r) = 0_S\}$. Man sieht sofort, dass $\mathrm{Kern}(\varphi)$ ein Ideal in R ist. Wir haben nun die induzierte Abbildung $\overline{\varphi} : R/\mathrm{Kern}(\varphi) \longrightarrow S$ mit $\overline{\varphi}([r]_U) = \varphi(r)$ für alle $r \in R$. Diese Abbildung ist injektiv und ein Homomorphismus von Ringen.

Wir beschäftigen uns im Folgenden nur mit kommutativen Ringen. Sei also R kommutativ.

Beispiel 4.2 Seien $a_1, \ldots, a_k \in R$. Dann ist die Menge

$$\{r_1 a_1 + \cdots r_k a_k \mid r_1, \ldots, r_k \in R\}$$

ein Ideal von R. Es ist (bzgl. der Inklusion) das kleinste Ideal von R, das die Elemente $a_1, \ldots, a_k$ enthält.

Dieses Ideal heißt das *von $a_1, \ldots, a_k$ erzeugte Ideal*. Es wird mit $(a_1, \ldots, a_k)$ bezeichnet.

Wenn wir nur ein Element nehmen, sagen wir $a \in R$, erhalten wir das so genannte *von a erzeugte Hauptideal* $(a) = a \cdot R$.

Konkret ist das von einer natürlichen Zahl n erzeugte Ideal in $\mathbb{Z}$ gleich $n \cdot \mathbb{Z}$, und der Restklassenring modulo n, $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/(n)$, ist ein Spezialfall eines Restklassenrings modulo eines Ideals.

Achtung: Das Tupel von Elementen $(a_1, \ldots, a_k)$ und das von $a_1, \ldots, a_k$ erzeugte Ideal werden genau gleich bezeichnet, sind aber unterschiedlich.

Definition Seien $a, b \in R$. Dann sagen wir, dass das Element a das Element b *teilt*, $a|b$, falls ein $c \in R$ mit $ca = b$ existiert. In diesem Fall sagen wir auch, dass a ein *Teiler* von b ist.

Bemerkung Für $a, b \in R$ gilt $a|b$ genau dann, wenn $b \in (a)$ d.h. wenn $(b) \subseteq (a)$.

Bemerkung Teilbarkeit ist eine reflexive und transitive Relation. Für $\mathbb{Z}, K[X]$ (K ein Körper) und viele weitere Ringe ist sie aber nicht antisymmetrisch und somit keine Ordnungsrelation.

Definition Ein Element aus R^* (der multiplikativen Gruppe des Rings) heißt auch eine *Einheit* von R.

Somit ist $a \in R$ genau dann eine Einheit, wenn $1 \in (a)$ ist, und dies ist äquivalent zu $(a) = (1)$, d.h. $(a) = R$.

Definition

- Ein *Nullteiler* in R ist ein Element $a \in R - \{0\}$, so dass ein Element $b \in R - \{0\}$ mit $ab = 0$ existiert.
- Ein kommutativer Ring ohne Nullteiler heißt *nullteilerfrei*.
- Ein *Integritätsring* bzw. *Integritätsbereich* ist ein kommutativer nullteilerfreier Ring R mit $0_R \neq 1_R$ (bzw. mit $R \neq \{0_R\}$).

Achtung: In jedem Ring teilt jedes Element die 0. Das folgt direkt aus der Definition von "teilt". Die Aussage "r teilt 0" sollte man *nicht* in "r ist ein Nullteiler" unformulieren. Das wäre falsch.

Bemerkung Ein Ring ist genau dann eine Integritätsbereich, wenn $R-\{0\}$ ein Untermonoid von $(R, \cdot)$ ist.

Lemma 4.3 *Sei R ein Integritätsbereich, und seien $a, b \in R$. Dann sind äquivalent:*

- *$a|b$ und $b|a$*
- *$(a) = (b)$*
- *$\exists \epsilon \in R^* : b = \epsilon a$.*

Beweis. Die Äquivalenz der ersten beiden Aussagen ist offensichtlich. Wir wollen, zeigen, dass die zweite und die dritte Aussage äquivalent sind.

Sei $(a) = (b)$. Dann gibt es $c, d \in R$: $b = ca, a = db$. Nun ist $dca = db = a$, also $(dc - 1)a = 0$. Da R nullteilerfrei ist, folgt $dc = 1$.

Die Rückrichtung ist einfach. □

Definition Wenn zwei Elemente a, b in einem Integritätsbereich die Bedingungen im obigen Lemma erfüllen, heißen sie *assoziiert zueinander*. Wir schreiben dann $a \sim b$. Wenn a ein Teiler von b ist aber die Elemente nicht assoziiert zueinander sind, sagen wir, dass a ein *echter Teiler* von b ist.

Offensichtlich ist Assoziiertheit eine Äquivalenzrelation auf dem Ring.

Beispiele 4.4 In $\mathbb{Z}$ sind zwei Elemente genau dann assoziiert zueinander, wenn sie denselben Absolutbetrag haben. Ein Repräsentantensystem der Äquivalenzrelation Assoziiertheit auf $\mathbb{Z} - \{0\}$ wird hier durch die natürlichen Zahlen (die Elemente in $\mathbb{N}$) gegeben.

In $K[X]$, K ein Körper, sind zwei Elemente (Polynome) f, g genau dann assoziiert zueinander, wenn es eine Konstante $c \in K^*$ mit $g = cf$ gibt. In $K[X] - \{0\}$ bilden die normierten Polynome ein Repräsentatensystem.

Bemerkung Wie schon gesagt ist Teilbarkeit nicht für jeden Ring eine Ordnungsrelation. Aber wenn man Teilbarkeit auf ein Repräsentatensystem der Relation Assoziiertheit einschränkt, erhält man eine Ordnungsrelation. Alternativ kann man zur durch "Assoziiertkeit" definierten Partition übergehen. Hierauf induziert dann Teilbarkeit eine Ordungsrelation. Man kann aber auch einfach so mit der Relation Teilbarkeit arbeiten.

Definition

- Sei R ein Integritätsbereich. Eine *Gradfunktion* auf R ist eine Abbildung $\delta : R \longrightarrow \mathbb{N}_0 \dot\cup \{-\infty\}$, so dass
  - für alle $a, b \in R$ mit $b \neq 0$ Elemente $q, r \in R$ mit $a = qb + r$ und $\delta(r) < \delta(b)$ existieren,
  - für alle $a, b \in R - \{0\}$ $\delta(a) \leq \delta(ab)$ gilt.
- Ein *Euklidischer Ring* ist ein Integritätsbereich, auf dem eine Gradfunktion existiert.

Die Bezeichnung δ deutet übrigens auf "degree" hin.

Achtung: Die Gradfunktion ist nicht Bestandteil der Struktur des Euklidischen Rings. Es ist also *nicht* so wie z.B. bei Euklidischen Vektorräumen, welche Tupel von Vektorräumen und Skalarprodukten sind. Verlangt wird nur, dass eine Gradfunktion existiert.

Die Standardbeispiele für Euklidische Ring sind, wie bereits erwähnt, $\mathbb{Z}$ und die Polynomringe $K[X]$, wobei K ein Körper ist. (Wie man die "Unbestimmte" bezeichnet, spielt natürlich keine Rolle.) Für $\mathbb{Z}$ kann man als Gradfunktion den Absolutbetrag nehmen und für $K[X]$ die normale Gradfunktion (mit $\text{Grad}(0) := -\infty$).

Es gibt übrigens abweichende Definitionen von "Gradfunktion". Die zweite Bedingung wird oft weggelassen (das ist eine starke Änderung), und der 0 wird manchmal kein Element zugeordnet (das ist eine geringfügige Änderung). Dementsprechend werden auch Euklidische Ringe unterschiedlich definiert.

Lemma 4.5 *Sei R ein Euklidischer Ring mit einer Gradfunktion δ, und seien $a, b \in R$ mit $a|b$. Dann sind äquivalent:*

a) $a \sim b$

b) $\delta(a) = \delta(b)$

Beweis. Wenn $a \sim b$ gilt, ist nach Definition $\delta(a) \leq \delta(b)$ und $\delta(b) \leq \delta(a)$.

Es gelte nun $\delta(a) = \delta(b)$ (und sowieso $a|b$). Sei $c \in R$ mit $ca = b$. Seien $q, r \in R$ mit $a = qb + r$ und $\delta(r) < \delta(b) = \delta(a)$. Dann ist also $a = qca + r$ und somit $(1 - qc) \cdot a = r$. Wenn nun $1 - qc \neq 0$ wäre, wäre $\delta(a) \leq \delta(r)$, ein Widerspruch. Also ist $1 - qc = 0$ bzw. $qc = 1$. Somit ist also c eine Einheit, d.h. $a \sim b$. □

Aussage 4.6 *Sei R ein Euklidischer Ring mit einer Gradfunktion δ, und sei I ein nicht-triviales Ideal in R. Sei $a \in I$, so dass $\delta(a) = \min\{\delta(b) \mid b \in I - \{0\}\}$. Dann ist $I = (a)$.*

Beweis. Da $a \in I$ gilt auch $(a) \subseteq I$. Sei nun $b \in I$. Dann gibt es $q, r \in R$ mit $b = qa + r$ und $\delta(r) < \delta(a)$. Es gilt nun $r = b - qa \in I$. Aus der Definition von a folgt $r = 0$ also $b = qa$. □

Wir haben soeben bewiesen, dass jedes Ideal in einem Euklidischen Ring ein Hauptideal ist.

Definition Ein *Hauptidealring* oder *Hauptidealbereich* ist ein Integritätsbereich, in dem jedes Ideal ein Hauptideal ist.

Somit ist also jeder Euklidische Ring ein Hauptidealring.

Aussage 4.7 *Sei R ein Hauptidealring, und seien $a_1, \ldots, a_k \in R$, nicht alle gleich 0. Sei ferner $a \in R$. Dann sind äquivalent:*

a) $(a_1, \ldots, a_k) = (a)$

b) $a|a_1, \ldots, a|a_k$ *und es existieren* $s_1, \ldots, s_k \in R$ *mit* $s_1a_1 + \cdots + s_ka_k = a$.

c) $a|a_1, \ldots, a|a_k$ *und für alle* $b \in R$ *mit* $b|a_1, \ldots, b|a_k$ *gilt* $b|a$.

Je zwei Elemente a, a' mit den obigen Eigenschaften sind zueinander assoziiert.

Wenn R ein Euklidischer Ring mit einer Gradfunktion δ ist, sind die obigen Bedingungen äquivalent zu: $a|a_1, \ldots, a|a_k$ *und* $\delta(a) = \max\{\delta(b) \mid b|a_1, \ldots, b|a_k\}$.

Beweis. Aussagen a) und b) sind offensichtlich äquivalent.

a) $\to$ *c)*. Es gelte a) und es sei b wie in c). Dann gilt $a_1, \ldots, a_k \in (b)$ und somit $(a) = (a_1, \ldots, a_k) \subseteq (b)$. Also ist $a \in (b)$ bzw. $b|a$.

c) $\to$ *a)*. Es gelte c). Dann gilt sicher $(a_1, \ldots, a_k) \subseteq (a)$. Es sei $b \in R$ mit $(a_1, \ldots, a_k) = (b)$. (Wir benutzen, dass R ein Hauptidealring ist.) Dann gilt also $b|a_1, \ldots, b|a_k$. Somit gilt $b|a$. Also gilt $(a) \subseteq (b) = (a_1, \ldots, a_k)$. Insgesamt gilt also $(a_1, \ldots, a_k) = (a)$.

Nach a) ist es offensichtlich, dass je zwei solche Elemente a, a' zueinander assoziiert sind.

Sei nun R Euklidisch mit Gradfunktion δ. Es gelten zunächst die ersten Bedingungen. Dann ist $\max\{\delta(b) \mid b|a_1, \ldots, b|a_k\} = \max\{\delta(b) \mid b|a\} = \delta(a)$.

Es gelte nun die angebene Bedingung. Dann ist zunächst $(a_1, \ldots, a_k) \subseteq (a)$. Sei $a' \in R$ mit $(a') = (a_1, \ldots, a_k)$. Dann gilt also $(a') \subseteq (a)$, d.h. $a' \in (a)$, d.h. $a|a'$.

Ferner ist, wie soeben gezeigt, $\max\{\delta(b) \mid b|a_1, \ldots, b|a_k\} = \delta(a')$, d.h. $\delta(a) = \delta(a')$. Mit Lemma 4.5 folgt: $(a) = (a') = (a_1, \ldots, a_k)$. □

Dies motiviert:

Definition Sei R ein Integritätsbereich und seien $a_1, \ldots, a_k \in R$, nicht alle gleich 0. Ein Element $g \in R$ mit $(g) = (a_1, \ldots, a_k)$ heißt ein *größter gemeinsamer Teiler (ggT)* von $a_1, \ldots, a_k$.

Bemerkung und Notation Wenn g ein ggT von $a_1, \ldots, a_k$ ist, dann sind genau die zu g assoziierten Elemente alle ggTs von $a_1, \ldots, a_k$. Es gibt also in der Regel nicht *den* ggT. Nichtsdestoweniger schreibt man: $\mathrm{ggT}(a_1, \ldots, a_k) = g$, wenn g *ein* ggT von $a_1, \ldots, a_k$ ist. Besser ist allerdings die uneindeutige Aussage $(g) = (a_1, \ldots, a_k)$.

Wir nehmen nun an, dass wir ein Repräsentantensystem der Relation Assoziiertheit ausgewählt haben. Dann gibt es zu jedem System $a_1, \ldots, a_k$ wie oben genau einen ggT in dem Repräsentantensystem, und oftmals meint man mit *dem* ggT dann denjenigen ggT in dem Repräsentatensystem (z.B. eine natürliche Zahl oder ein normiertes Polynom).

Auf dem Repräsentatensystem ist ja Teilbarkeit eine Äquivalenzrelation. Wenn nun $a_1, \ldots, a_k$ aus dem Repräsentantensystem sind und g der ggT von $a_1, \ldots, a_k$ aus dem Repräsentantenystem ist, dann ist g das größte Element, das $a_1, \ldots, a_k$ teilt im Sinne der allgemeinen Definition von "größtem Element".

Alternativ kann man die Definition von "größten Elementen" auf reflexive und transitive Relationen ausweiten. Dann ist ein ggT von $a_1, \ldots, a_k$ *ein* größtes Element, das $a_1, \ldots, a_k$ teilt.

Der Euklidische Algorithmus

Der *Euklidische Algorithmus* ist ein Algorithmus, um den ggT zweier Elemente in einem Euklidischen Ring explizit auszurechnen.

Sei hierzu R ein Euklidischer Ring mit einer Gradfunktion δ.

Die Grundidee ist die folgende: Seien $a, b \in R - \{0\}$. Wir betrachten nun die Division mit Rest:

$$a = qb + r$$

mit $\delta(r) < \delta(b)$. Nun gibt es zwei Fälle:

Entweder $r = 0$. Dann ist b ein Teiler von a und $\mathrm{ggT}(a, b) = b$.

Oder $r \neq 0$. Dann ist $(a, b) = (b, r)$, denn: $a \in (b, r)$ und $r \in (b, a)$ und somit $(a, b) \subseteq (b, r)$ und $(b, r) \subseteq (a, b)$. Somit ist $\mathrm{ggT}(a, b) = \mathrm{ggT}(b, r)$. (Man beachte die Symbolik: Wir behaupten nicht, dass es einen einzigen ggT gibt, aber wenn wir einen ggT von a, b haben, dann haben wir auch einen von b, r und umgekehrt.) Nun kann man a, b durch b, r ersetzen und das Vefahren iterieren. Da die Reste immer kleiner werden, gelangt man irgendwann zum ggT von a und b.

Sei nun g ein ggT von a und b. Dann gibt es wie oben bemerkt Elemente $c, d \in R$ mit $ca + db = g$. Mit dem so genannten *erweiterten Euklidischen Algorithmus* kann man solche Elemente ausrechnen.

Nehmen wir an, dass wir mit den Euklidischen Algorithmus das Folgende berechnet haben:

$$\begin{aligned} a &= p_1 b + r_1 \\ b &= p_2 r_1 + r_2 \\ r_1 &= p_3 r_2 + r_3 \\ &\vdots \\ r_{k-3} &= p_{k-1} r_{k-2} + r_{k-1} \\ r_{k-2} &= p_k r_{k-1} + 0 \end{aligned}$$

Dann ist also $\mathrm{ggT}(a,b) = \mathrm{ggT}(b,r_1) = \mathrm{ggT}(r_1,r_2) = \cdots = \mathrm{ggT}(r_{k-2},r_{k-1}) = r_{k-1}$.

Die Idee ist, von "unten nach oben" zurückzurechnen. Wir haben $\mathrm{ggT}(a,b) = r_{k-1} = r_{k-3} - p_{k-1} r_{k-2}$. Nun substituieren wir r_{k-2} mittels der Zeile "darüber". Wir erhalten eine Darstellung der Form $\mathrm{ggT}(a,b) = c_{k-4} r_{k-4} + d_{k-4} r_{k-3}$ mit gewissen Ringelementen c_{k-4}, d_{k-4}. Wenn wir so fortfahren, erhalten wir Darstellungen

$$\mathrm{ggT}(a,b) = c_i r_i + d_i r_{i+1}$$

mit gewissen $c_i, d_i \in \mathbb{Z}$ für alle $i \geq 1$. Ausgehend von

$$\mathrm{ggT}(a,b) = c_1 r_1 + d_1 r_2$$

liefert eine Substitution von r_2 mittels der zweiten Zeile eine Darstellung

$$\mathrm{ggT}(a,b) = c_0 b + d_0 r_1 \ ,$$

und eine weitere Substitution mittels der ersten Zeile liefert eine Darstellung

$$\mathrm{ggT}(a,b) = ca + db \ ,$$

wie gewünscht.

Faktorisierung

Definition Sei R ein Integritätsbereich und sei $r \in R$.

- r ist *irreduzibel*, falls r nicht 0 und keine Einheit ist und es gilt:

$$\forall s \in R : s | r \longrightarrow s \sim r \vee s \in R^*$$

- r ist *prim*, falls r nicht 0 und keine Einheit ist und es gilt:

$$\forall a, b \in R : r | ab \longrightarrow r | a \vee r | b$$

Bemerkung Eine *Primzahl* ist also per Definition ein irreduzibles Element im Ring $\mathbb{Z}$, das in $\mathbb{N}$ liegt. Wir werden sogleich sehen, dass Primzahlen auch prim im Sinne der obigen Definition sind.

Bemerkung Wenn r prim ist und r ein Produkt beliebig vieler Faktoren teilt, dann teilt r auch einen Faktor. Dies sieht man sofort aus der Definition.

Lemma 4.8 *Jedes Primelement in einem Integritätsbereich ist auch irreduzibel.*

Beweis. Sei R ein Integritätsbereich und $p \in R$ prim. Sei nun $s|p$ mit $s \notin R^*$. Dann gibt es also ein $a \in R$ mit $as = p$; wir fixieren so ein a. Es gilt nun $p|s$ oder $p|a$. Wir nehmen zuerst an, dass $p|a$ gilt. Dann gibt es ein $c \in R$ mit $cp = a$. Mit so einem c ist nun $scp = p$ und somit $sc = 1$. Hiermit ist s eine Einheit, ein Widerspruch.

Es gilt also $p|s$. Sei $d \in R$ mit $dp = s$. Dann ist $dap = p$, also $ad = 1$. Hiermit ist a eine Einheit und somit $s \sim p$. □

Satz 4.1 *Sei R ein Hauptidealring und $r \in R$. Dann ist r genau dann irreduzibel, wenn es prim ist.*

Beweis. Die eine Richtung haben wir schon gezeigt.

Sei nun r irreduzibel und seien $a, b \in R$ mit $r|ab$. Sei $s \in R$ mit $s = \mathrm{ggT}(r, a)$, d.h. $(s) = (r, a)$.

Wir nehmen an, dass $r \nmid a$ ist und wollen zeigen, dass dann $b|r$ gilt. Nun ist (r, a) echt größer als (r) und somit $s \notin (r)$ bzw. $r \nmid s$. Damit ist s ein echter Teiler von r, also ist s eine Einheit und $(r, a) = (1) = R$. Hieraus folgt: $b \in (b) = b \cdot (1) = b \cdot (r, a) = (br, ba) = (br, ab)$. Da sowohl br als auch ab Vielfache von r sind, folgt, dass jedes Element in (br, ab) ein Vielfaches von r ist. Somit ist auch b ein Vielfaches von r. □

Dieser Satz zeigt auch Lemma 1.47.

Wir wollen nun die eindeutige Primfaktorzerlegung in Euklidischen Ringen beweisen, also insbesondere in $\mathbb{Z}$ und $K[X]$, K ein Körper, beweisen.

Wir wählen bezüglich der Äquivalenzrelation Assoziiertheit in jeder Äquivalenzklasse primer Elemente genau eins aus. (Wir wählen ein Repräsentationsystem.) Sei $\mathbb{P}$ so eine Menge.

Die folgenden Beispiele sollte man im Kopf behalten: Im Ring $\mathbb{Z}$ ist es natürlich, die Primzahlen auszuwählen. Im Ring $K[X]$, K ein Körper, ist es natürlich, die normierten irreduziblen Polynome auszuwählen.

Es gilt nun:

Satz 4.2 *Sei R ein Euklidischer Ring und sei $\mathbb{P}$ eine Menge wie beschrieben. Sei nun $r \in R - \{0_R\}$. Dann gibt es ein $k \in \mathbb{N}_0$ und Elemente $p_1, \ldots, p_k \in \mathbb{P}, \epsilon \in R^*$ mit $r = \epsilon p_1 \cdots p_k$. So eine "Darstellung" ist eindeutig bis auf die Reihenfolge der p_i.*

Beweis. Sei δ eine Gradfunktion auf R. Beachten Sie, dass aus Lemma 4.5 folgt: Für $a, b \in R - \{0\}$, wobei a ein echter Teiler von b ist, ist $\delta(a) < \delta(b)$. Die Existenz folgt nun sofort per Induktion über den Grad.

Zur Eindeutigkeit: Es gelte $\epsilon p_1 \cdots p_k = \epsilon' p'_1 \cdots p'_{k'}$ mit $\epsilon, \epsilon' \in R^*$ und $p_i, p'_i \in R$. Da p_k prim ist, gibt es ein $i = 1, \ldots, k'$ mit $p_k | p'_i$. Hieraus folgt dann sofort $p_k = p'_i$. Ohne Einschränkung können wir annehmen, dass $i = k'$. Hiermit ist $\epsilon\, p_1 \ldots p_{k-1} = \epsilon'\, p_1 \cdots p_{k'-1}$.

Aus diesen Überlegungen folgt der Satz per Induktion nach k. □

Wir haben also bewiesen, dass es sowohl im Ring $\mathbb{Z}$ als auch in dem Ringen $K[X]$, K ein Körper, eindeutige Primfaktorzerlegung gibt.

Wir diskutieren noch Varianten der Notation der Primfaktorzerlegung. Seien hierzu R und $\mathbb{P}$ wie oben.

Man kann die eindeutige Primfaktorzerlegung auch so ausdrücken:

Zu $r \in R - \{0\}$ gibt es $\epsilon \in R^*, p_1, \ldots, p_k \in \mathbb{P}$, wobei die p_i paarweise verschieden sind, sowie $e_1, \ldots, e_k \in \mathbb{N}$, so dass $r = p_1^{e_1} \cdots p_k^{e_k}$. Wiederum ist die Darstellung eindeutig bis auf die Reihenfolge.

Eleganter ist die folgende Vorgehensweise: Sei für eine Menge X $\mathbb{N}_0^{(X)}$ die Menge der Abbildungen von X nach $\mathbb{N}_0$, die nur für endlich viele $x \in X$ ungleich Null sind. D.h.

$$\mathbb{N}_0^{(X)} = \{(n_x)_{x \in X} \mid \text{ es gibt nur endlich viele } x \in X \text{ mit } n_x \neq 0\} .$$

Dies ist ein Untermonoid des Monoids $\mathbb{N}_0^X$ mit der komponentenweisen Addition.

Sei nun $(n_p)_{p \in \mathbb{P}} \in \mathbb{N}_0^{(\mathbb{P})}$. Dann gilt ja für alle bis auf endlich viele $p \in \mathbb{P}$ $p^{n_p} = 1$. Wir definieren nun $\prod_{p \in \mathbb{P}} p^{n_p}$ als dasjenige Element in R, das man erhält, wenn man alle Elemente p^{n_p}, die nicht 1 sind, aufmultipliziert. (Also, $\mathbb{P}$ kann unendlich groß sein, aber es gibt immer nur endlich viele p^{n_p}, die ungleich 1 sind, und hierüber wird das Produkt gebildet.) Es gilt nun für $(n_p)_{p \in \mathbb{P}}, (n'_p)_{p \in \mathbb{P}} \in \mathbb{N}_0^{(\mathbb{P})}$:

$$\prod_{p \in \mathbb{P}} p^{n_p} \cdot \prod_{p \in \mathbb{P}} p^{n'_p} = \prod_{p \in \mathbb{P}} p^{n_p + n'_p}$$

Man erhält, dass die Abbildung $\mathbb{N}_0^{(\mathbb{P})} \longrightarrow R - \{0\}$, $(n_p)_{p \in \mathbb{P}} \mapsto \prod_{p \in \mathbb{P}} p^{n_p}$ ein Homomorphismus von Monoiden von $(\mathbb{N}_0^{(\mathbb{P})}, +)$ nach $(R - \{0\}, \cdot)$ ist.

Die Aussage des Satzes ist nun: Zu $r \in R - \{0\}$ gibt es eine eindeutig bestimmte Einheit ϵ und einen eindeutig bestimmtes "Tupel von Exponenten" $(e_p)_{p\in\mathbb{P}}$ mit

$$r = \epsilon \cdot \prod_{p\in\mathbb{P}} p^{e_p} \ .$$

Man kann dies auch so formulieren: Die Abbildung

$$R^* \times \mathbb{N}_0^{(\mathbb{P})} \longrightarrow R - \{0\} \ , \ (\epsilon, (n_p)_{p\in\mathbb{P}}) \mapsto \epsilon \cdot \prod_{p\in\mathbb{P}} p^{n_p}$$

ist ein Isomorphismus von Monoiden.

Mit dieser Darstellung haben wir übrigens auch für beliebige $(e_p)_{p\in\mathbb{P}} \in \mathbb{N}_0^{(\mathbb{P})}, (e'_p)_{p\in\mathbb{P}} \in \mathbb{N}_0^{(\mathbb{P})}$:

$$\mathrm{ggT}\,\big(\prod_{p\in\mathbb{P}} p^{e_p}, \prod_{p\in\mathbb{P}} p^{e'_p}\big) = \prod_{p\in\mathbb{P}} p^{\min\{e_p, e'_p\}} \ ,$$

wie man leicht sieht.

4.2 Die Jordansche Normalform

Wir kehren zur Linearen Algebra zurück. Sei K ein Körper, V ein endlich erzeugter K-Vektorraum der Dimension n und φ ein Endomorphismus von V.

Wie schon in Abschnitt 2.11 besprochen, können wir φ in Polynome in $K[T]$ einsetzen und erhalten wieder einen Endomorphismus von V. Wir erinnern uns: Die Abbildung in (2.28)

$$K[T] \longrightarrow \mathrm{End}_K(V) \ , \ f \mapsto f(\varphi)$$

ist ein Homomorphismus von K-Vektorräumen und von Ringen. Ferner ist die Abbildung nicht injektiv, weil $\mathrm{End}_K(V)$ die Dimension n^2 hat und $K[T]$ nicht endlich erzeugt ist.

Definition Das *Minimalpolynom* von φ, μ_φ, ist das eindeutig bestimmte normierte Polynom kleinsten Grades $f \in K[T]$ mit $f(\varphi) = 0$.

Anders ausgedrückt: Das Minimalpolynom von φ ist das eindeutig bestimte normierte Polynom kleinsten Grades im Kern des Homomorphismus $K[T] \longrightarrow \mathrm{End}_K(V)$. Dieser Kern ist ein Ideal in $K[T]$, und nach Aussage 4.6 wissen wir: Das Minimalpolynom erzeugt den Kern. Mit anderen Worten:

Lemma 4.9 *Sei $f \in K[T]$ mit $f(\varphi) = 0$. Dann teilt μ_φ das Polynom f.*

Definition Sei $S \subseteq V$. Das *φ-Erzeugnis* von S ist der Untervektorraum

$$\langle S \rangle_\varphi := \langle \{\varphi^i(\mathfrak{v}) \mid i \in \mathbb{N}_0, \mathfrak{v} \in S\} \rangle$$

von V.

Dies ist offensichtlich der kleinste φ-invariante Untervektorraum von V, der die Menge S umfasst.

Definitionen Der Vektorraum V ist *φ-zyklisch*, falls er das φ-Erzeugnis eines einzigen Vektors von V ist. Dies heißt also, dass $V = \langle \mathfrak{v} \rangle_\varphi$ für ein $\mathfrak{v} \in V$.

Das Hauptziel dieses Abschnitts ist wie folgt: Wir wollen versuchen, V in eine direkte Summe "möglichst kleiner" φ-invariante Untervektorräume zu zerlegen. Wir werden auch den Satz von Cayley-Hamilton erneut beweisen, diesmal ohne die Lücke im Beweis des Satzes in Abschnitt 2.11.

Wir setzen nun – fürs Erste – voraus, dass V das φ-Erzeugnis eines Vektors $\mathfrak{v} \in V$ ist.

Lemma 4.10 *Für $f \in K[T]$ sind äquivalent*

- $f(\varphi) = 0$
- $f(\varphi)(\mathfrak{v}) = \mathfrak{o}$.

Beweis. Wir machen eine Vorbemerkung: Für beliebige Polynome $f, g \in K[T]$ gilt $f(\varphi) \circ g(\varphi) = (f \cdot g)(\varphi) = (g \cdot f)(\varphi) = g(\varphi) \circ f(\varphi)$. Insbesondere ist also für $i \in \mathbb{N}_0$ $f(\varphi) \circ \varphi^i = \varphi^i \circ f(\varphi)$.

Es gelte die zweite Bedingung. Dann ist für $i \in \mathbb{N}_0$ $f(\varphi)(\varphi^i(\mathfrak{v})) = \varphi^i(f(\varphi)(\mathfrak{v})) = 0$. Da die Menge $\{\varphi^i(\mathfrak{v}) \mid i \in \mathbb{N}_0\}$ ein Erzeugendensystem von V bildet, gilt $f(\varphi) = 0$. □

Somit ist also μ_φ das eindeutig bestimmte normierte Polynom $f \in K[T]$ kleinsten Grades mit $f(\varphi)(\mathfrak{v}) = 0$.

Sei $\mu_\varphi = \sum_{i=0}^{n} a_i T^i$ mit $a_n = 1$. Nun ist $\mathfrak{v}, \varphi(\mathfrak{v}), \ldots, \varphi^{n-1}(\mathfrak{v})$ eine Basis von V.

Denn: Erstens sind die Vektoren linear unabhängig, weil ansonsten μ_φ nicht minimalen Grad unter allen Polynomen $p \in K[T]$ mit $p(\varphi)(\mathfrak{v}) = 0$ hätte.

Zweitens bilden die Vektoren ein Erzeugendensystem. Hierzu ist zu zeigen, dass jeder Vektor $\varphi^i(\mathfrak{v})$ eine Linearkombination dieser Vektoren ist. Nun gibt es Polynome $q, r \in K[T]$ mit $T^i = q \cdot \mu_\varphi + r$ und $\mathrm{Grad}(r) < \mathrm{Grad}(\mu_\varphi)$, d.h. $\mathrm{Grad}(r) < n$. Dann ist $\varphi^i(\mathfrak{v}) = r(\varphi)(\mathfrak{v}) \in \langle \mathfrak{v}, \varphi(\mathfrak{v}), \ldots, \varphi^{n-1}(\mathfrak{v}) \rangle$.

Die Abbildungsmatrix bezüglich dieser Basis ist

$$\begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & 0 & -a_1 \\ 0 & 1 & & 0 & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & & 1 & -a_{n-1} \end{pmatrix}.$$

Man rechnet leicht aus, dass das charakteristische Polynom der Matrix – und damit des Endomorphismus φ – gleich $(-1)^n \cdot \sum_{i=0}^{n} a_i T^i = (-1)^n \cdot \mu_\varphi$ ist.

Wir halten fest:

Aussage 4.11 *Wenn V das φ-Erzeugnis eines Vektors $\mathfrak{v} \in V$ ist, dann ist das Minimalpolynom von φ "bis auf Vorzeichen" gleich dem charakteristischen Polynom von φ und gleich dem normierten Polynom kleinsten Grades f mit $f(\varphi)(\mathfrak{v}) = 0$.*

Sei nun V wieder ein beliebiger endlich erzeugter K-Vektorraum mit einem Endomorphismus φ.

Wir beweisen nun nochmal den *Satz von Cayley-Hamilton*. Wir wollen also zeigen, dass $\chi_\varphi(\varphi) = 0$ ist. Dies kann man übrigens auch so ausdrücken:

$$\mu_\varphi \mid \chi_\varphi \,.$$

Sei hierzu $\mathfrak{v} \in V$ beliebig. Sei U das φ-Erzeugnis von $\mathfrak{v}$. Dann gilt

$$(-1)^n \cdot \mu_{\varphi_U} = \chi_{\varphi_U} \mid \chi_\varphi \,.$$

nach der obigen Aussage und (2.27). Somit gilt also $\chi_\varphi(\varphi)_U = \chi_\varphi(\varphi_U) = 0$ und insbesondere $\chi_\varphi(\mathfrak{v}) = \mathfrak{o}$. □

Definition Sei $p \in K[T]$ irreduzibel. Dann ist V ist *p-primär*, falls für jeden Vektor $\mathfrak{v} \in V$ ein $i \in \mathbb{N}$ mit $p(\varphi)^i(\mathfrak{v}) = 0$ existiert.

Lemma 4.12 *Sei $p \in K[T]$ normiert und irreduzibel. Die folgenden Aussagen sind äquivalent:*

a) V ist p-primär.

b) μ_φ ist eine p-Potenz.

c) χ_φ ist "bis auf das Vorzeichen" eine p-Potenz.

Beweis. Offensichtlich impliziert b) a), und nach dem Satz von Cayley-Hamilton impliziert c) b). (Man sieht auch leicht, dass a) b) impliziert, aber das benötigen wir gar nicht.)

Wir zeigen, dass a) c) impliziert. Wir zeigen die Implikation per Induktion über die Dimension von V. Der Induktionsanfang ist trivial. Sei nun V nicht-trivial. Dann gibt es einen von $\mathfrak{o}$ verschiedenen Vektor $\mathfrak{v} \in V$. Sei $\overline{\varphi} : V/\langle\mathfrak{v}\rangle_\varphi \longrightarrow V/\langle\mathfrak{v}\rangle_\varphi$ der induzierte Endomorhismus. Dann haben wir $\chi_\varphi = \chi_{\varphi_{\langle\mathfrak{v}\rangle_\varphi}} \cdot \chi_{\overline{\varphi}}$ nach (2.27). Nach Aussage (4.11) ist der erste Faktor "bis auf das Vorzeichen" eine p-Potenz und nach Induktionsvoraussetzung der zweite. □

Definition Sei $p \in K[T]$ irreduzibel. Der *p-primäre Anteil* von V ist

$$V(p) := \{\mathfrak{v} \in V \mid \exists i \in \mathbb{N}_0 : p(\varphi)^i(\mathfrak{v}) = \mathfrak{o}\} .$$

Wenn man weiß, was man damit meint, kann man auch $V(p) = \mathrm{Kern}(p(\varphi)^\infty)$ schreiben. Offensichtlich ist $V(p)$ φ-invariant.

Satz 4.3 *Sei* $\mathrm{Dim}(V) = n$ *und* $\chi_\varphi = (-1)^n \cdot \prod_{i=1}^k p_i^{n_i}(T) \in K[T]$ *mit paarweise verschiedenen normierten irreduziblen Polynomen* p_i *und* $n_i \in \mathbb{N}_0$*. Dann gilt:*

a) $V = V(p_1) \oplus \cdots \oplus V(p_k)$

b) $\chi_{\varphi_{V(p_i)}} = \pm p_i^{n_i}$ *für alle* i*. Insbesondere ist* $\mathrm{Dim}(V(p_i)) = n_i \cdot \mathrm{Grad}(p_i)$.

c) $\mu_\varphi = \mu_{\varphi_{V(p_1)}} \cdots \mu_{\varphi_{V(p_k)}}$.

d) Die Polynome χ_φ *und* μ_φ *haben dieselben Primteiler.*

Bemerkung Es ist kein Versehen, dass die Exponenten auch 0 sein dürfen. Dann gilt der Satz nämlich auch.

Wir zeigen zunächst ein Lemma:

Lemma 4.13 *Seien* $f, g \in K[T]$ *zwei Polynome mit* $\mathrm{ggT}(f, g) = 1$ *und* $\mu_\varphi | f \cdot g$*. Dann ist* $\mathrm{Kern}(f(\varphi)) = \mathrm{Bild}(g(\varphi))$*, und* $g(\varphi)_{\mathrm{Kern}(f(\varphi))}$ *ist ein Automorphismus von* $\mathrm{Kern}(f(\varphi))$.

Beweis. Es folgt direkt aus der Voraussetzung, dass das Bild im Kern enthalten ist. Offensichtlich ist auch $\mathrm{Kern}(f(\varphi))$ φ-invariant.

Seien nun $a, b \in K[T]$ mit $af + bg = 1$. Dann ist also $a(\varphi) \circ f(\varphi) + b(\varphi) \circ g(\varphi) = \mathrm{id}_V$. Hieraus folgt: $b(\varphi)_{\mathrm{Kern}(f(\varphi))} \circ g(\varphi)_{\mathrm{Kern}(f(\varphi))} = \mathrm{id}_{\mathrm{Kern}(f(\varphi))}$ und $g(\varphi)_{\mathrm{Kern}(f(\varphi))} \circ b(\varphi)_{\mathrm{Kern}(f(\varphi))} = \mathrm{id}_{\mathrm{Kern}(f(\varphi))}$. □

Beweis des Satzes

Zunächst eine Vorbemerkung zu a) und b). Sei $i = 1, \ldots, k$. Wir haben $\chi_{\varphi_{V(p_i)}} | \chi_\varphi$, und $\chi_{\varphi_{V(p_i)}}$ ist "bis auf Vorzeichen" eine p_i-Potenz. Somit gilt also $\chi_{\varphi_{V(p_i)}} | p^{n_i}$.

a). Sei $q_i := p_1^{n_1} \cdots p_{i-1}^{n_{i-1}} \cdot p_{i+1}^{n_{i+1}} \cdots p_k^{n_k}$ für $i = 1, \ldots, k$. Nun ist

$$V(p_i) = \mathrm{Bild}(q_i(\varphi)) \ ,$$

$q_i(\varphi)_{V(p_i)}$ ist ein Automorphismus von $V(p_i)$, und für alle $i \neq j$ gilt $q_i(\varphi)_{V(p_j)} = 0$.

Nun gilt $\mathrm{ggT}(q_1, \ldots, q_k) = 1$, es gibt also $s_1, \ldots, s_k \in K[T]$ mit $s_1 q_1 + \cdots + s_k q_k = 1$. Wir fixieren solche $s_1, \ldots, s_k$.

Sei nun $\mathfrak{v} \in V$. Dann gilt $\mathfrak{v} = \mathrm{id}(\mathfrak{v}) = (\sum_i s_i q_i)(\varphi)(\mathfrak{v}) = \sum_i s_i(\varphi)(q_i(\varphi)) \in V(p_1) + \cdots + V(p_k)$.

Wir müssen noch zeigen, dass die Summe direkt ist. Sei hierzu $(\mathfrak{v}_1, \ldots, \mathfrak{v}_k) \in (V(p_1), \ldots, V(p_k))$ mit $\mathfrak{v}_1 + \cdots + \mathfrak{v}_k = \mathfrak{o}$. Wenn wir hierauf $q_i(\varphi)$ anwenden, erhalten wir $q_i(\varphi)(\mathfrak{v}_i) = \mathfrak{o}$. Da $q_i(\varphi)_{V(p_i)}$ ein Automorphismus von $V(p_i)$ ist, folgt $\mathfrak{v}_i = \mathfrak{o}$.

b). Wir haben $\mathrm{Dim}(V_i) = \mathrm{Grad}(\chi_{\varphi_{V(P_i)}}) \leq n_i \cdot \mathrm{Grad}(p_i)$. Wenn nun für ein i $\chi_{\varphi_{V(p_i)}}$ ein echter Teiler von $p_i^{n_i}$ wäre, wäre die Dimension der (direkten) Summe der $V(p_i)$ kleiner als n, ein Widerspruch.

c) Wir überprüfen, dass $\mu_{\varphi_{V(p_1)}} \cdots \mu_{\varphi_{V(p_k)}}$ das Minimalpolynom von φ ist:

Offensichtlich verschwindet der Endomorphismus $\mu_{\varphi_{V(p_1)}} \cdots \mu_{\varphi_{V(p_k)}}$ auf allen $V(p_i)$. Damit verschweindet er auf ganz V. Sei nun $f \in K[T]$ ein Polynom, so dass $f(\varphi) = 0$ ist und sei $i = 1, \ldots, k$. Dann verschwindet $f(\varphi)$ auch auf $V(p_i)$ und somit gilt $\mu_{\varphi_{V(p_i)}} | f$. Da $\mu_{\varphi_{V(p_i)}}$ eine p_i-Potenz ist, gilt $\mu_{\varphi_{V(p_1)}} \cdots \mu_{\varphi_{V(p_k)}} | f$.

d) Dies folgt sofort aus c). □

Wenn wir jeweils in den Räumen $V(p_i)$ Basen wählen, erhalten wir eine Abbildungsmatrix der Form

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix} \tag{4.1}$$

mit $\chi_{A_i} = \pm p_i^{n_i}$.

Definition Wir nennen eine Matrix wie in (4.1) (ohne Bedingungen an die chakteristischen Polynome der Matrizen A_i) eine *Blockdiagonalmatrix* mit *Diagonalblöcken* $A_1, \ldots, A_k$.

Wir betrachten noch einen wichtigen Spezialfall p-primärer Untervektorräume:

Sei $a \in K$. Wenn nun a ein Eigenwert ist, heißt $V(T-a)$ auch *Hauptraum* zum Eigenwert a und wird mit H_a bezeichnet. Ganz allgemein setzen wir für $a \in K$: $H_a := V(T-a)$. Sei nun $(T-a)^e$ die maximale Potenz von $T-a$, die χ_φ teilt. Dann ist also $e = \mathrm{Dim}(H_a)$. Diese Zahl heißt die *algebraische Vielfachheit* von a. Die Dimension von E_a heißt die *geometrische Vielfachkeit* von a.

Wir wollen nun die p-primären Anteile von V in zyklische Untervektorräume zerlegen und damit auch Abbildungsmatrizen in Blockdiagonalgestalt mit "besonders schönen" Blöcken finden. Hierzu machen wir die folgende Voraussetzung:

Voraussetung Das charakteristische Polynom von φ zerfalle in Linearfaktoren.

Nach Punkt d) des obigen Satzes ist die Voraussetzung äquivalent zur scheinbar schwächeren Voraussetzung, dass das Minimalpolynom von φ in Linearfaktoren zerfalle.

Die normierten irreduziblen Teiler von χ_φ sind nun von der Form $T-a$ mit $a \in K$, und nach Satz 4.3 ist V eine direkte Summe der Haupträume von φ. (Man spricht hier von *Hauptraumzerlegung.*)

Wir nehmen nun an, dass χ_φ eine Potenz eines einzigen Linearfaktors ist. Sei also $\chi_\varphi = (T-a)^n$ und somit $V = H_a = V(T-a)$ von Dimension n. Damit ist insbesondere $(\varphi - a \cdot \mathrm{id})^n = 0$.

Der Hauptsatz ist jetzt:

Satz 4.4 *Sei $a \in K$ und sei $V = H_a$, d.h. V ist $(T-a)$-primär. Dann ist V eine direkte Summe von φ-zyklischen Untervektorräumen.*

Hierzu zunächst eine Definition.

Definition Ein Endomorphismus φ von V ist *nilpotent*, falls eine Potenz von φ gleich der Nullabbildung ist.

Dies bedeutet nichts anderes, als dass $V = H_0$ ist, d.h. V ist T-primär.

In unserem Fall ist also $\varphi' := \varphi - a \cdot \mathrm{id}$ nilpotent. Außerdem ist ein Untervektorraum von V genau dann φ-zyklisch, wenn er φ'-zyklisch ist.

Der Satz folgt also sofort aus der folgenden Aussage.

Aussage 4.14 *Sei φ nilpotent. Dann ist V eine direkte Summe von φ-zyklischen Untervektorräumen.*

Ich gebe zwei Beweise dieser Aussage an. Der zweite Beweis hat den Vorteil, dass er sich leicht in einen effizienten Algorithmus verwandeln lässt.

Sei φ nilpotent.

Beweis 1. (nach Bröcker: Lineare Algebra und Analytische Geometrie) Wir zeigen die Aussage per Induktion über die Dimension. Wir fixieren einen Vektor $\mathfrak{v} \in V$, dessen φ-zyklisches Erzeugnis maximale Dimension unter allen φ-zyklischen Untervektorräumen von V hat und setzen $e := \mathrm{Dim}(\langle \mathfrak{v} \rangle_\varphi)$.

Wir wollen zeigen: $\langle \mathfrak{v} \rangle_\varphi$ hat eine φ-invariantes Komplement, d.h. es gibt einen φ-invarianten Untervektorraum U von V mit

$$V = \langle \mathfrak{v} \rangle_\varphi \oplus U \ .$$

Wenn man dann die Induktionsvoraussetzung auf U anwendet, erhält man die Aussage.

Um die Existenz so eines Untervektorraums U zu zeigen, wählen wir einen φ-invarianten Untervektorraum U von V der maximal mit der Eigenschaft ist, dass $\langle \mathfrak{v} \rangle_\varphi \cap U = \{\mathfrak{o}\}$ ist. Wir müssen nun zeigen, dass $\langle \mathfrak{v} \rangle_\varphi + U = V$ gilt.

Hierzu nehmen wir an, dass die Aussage nicht gilt. Der folgende Beweis hat mehrere Schritte ist wohl der "technischste" der ganzen Vorlesung. Wir nehmen also die folgende Annahme an:

$$\exists \mathfrak{x} \in V : \mathfrak{x} \notin \langle \mathfrak{v} \rangle_\varphi \oplus U \ . \tag{4.2}$$

Sei $\mathfrak{x}$ so wie in (4.2). Es gibt nun ein $i \in \mathbb{N}$ mit $\varphi^i(\mathfrak{x}) \in \langle \mathfrak{v} \rangle_\varphi \oplus U$. Sei i_0 das minimale solche i und $\mathfrak{x}' := \varphi^{i_0 - 1}(\mathfrak{x})$. Dann gilt $\mathfrak{x}' \notin \langle \mathfrak{v} \rangle_\varphi \oplus U$ und $\varphi(\mathfrak{v}') \in \langle \mathfrak{v} \rangle_\varphi \oplus U$. Wir haben also die folgende Aussage:

$$\exists \mathfrak{x} \in V : \mathfrak{x} \notin \langle \mathfrak{v} \rangle_\varphi \oplus U \wedge \varphi(\mathfrak{x}) \in \langle \mathfrak{v} \rangle_\varphi \oplus U \ . \tag{4.3}$$

Wieder fixieren wir einen solchen Vektor $\mathfrak{x}$. Seien $\mathfrak{y} \in \langle \mathfrak{v} \rangle_\varphi, \mathfrak{u} \in U$ mit $\varphi(\mathfrak{x}) = \mathfrak{y} + \mathfrak{u}$. Nun gibt es $a \in K$ und $\mathfrak{z} \in \langle \mathfrak{v} \rangle_\varphi$ mit $\mathfrak{y} = a \cdot \mathfrak{v} + \varphi(\mathfrak{z})$. Wir fixieren so eine "Darstellung" und betrachten nun das Element $\mathfrak{x} - \mathfrak{z}$. Es ist $\mathfrak{x} - \mathfrak{z} \notin \langle \mathfrak{v} \rangle_\varphi + U$ und $\varphi(\mathfrak{x} - \mathfrak{z}) = a \cdot \mathfrak{v} + \mathfrak{u} \in \langle \mathfrak{v} \rangle + U$. Wir haben also:

$$\exists \mathfrak{x} \in V : \mathfrak{x} \notin \langle \mathfrak{v} \rangle_\varphi \oplus U \wedge \varphi(\mathfrak{x}) \in \langle \mathfrak{v} \rangle \oplus U \tag{4.4}$$

(Beachten Sie den Unterschied zwischen dieser Aussage und Aussage (4.3).) Und wieder fixieren wir so ein $\mathfrak{x}$. Seien $a \in K$ und $\mathfrak{u} \in U$ mit $\varphi(\mathfrak{x}) = a \cdot \mathfrak{v} + \mathfrak{u}$.

Wir wollen zeigen, dass $a = 0$ ist. Hierzu benutzen wir, dass $e = \mathrm{Dim}(\langle \mathfrak{v} \rangle_\varphi)$ maximal und $\langle \mathfrak{v} \rangle_\varphi \cap U = \{\mathfrak{o}\}$ ist.

Nach Voraussetzung ist $\mathfrak{o} = \varphi^e(\mathfrak{x}) = \varphi^{e-1}(\varphi(\mathfrak{x})) = \varphi^{e-1}(a \cdot \mathfrak{v}) + \varphi^e(\mathfrak{u}) = a \cdot \varphi^{e-1}(\mathfrak{v}) + \varphi^e(\mathfrak{u})$. Hieraus folgt $a \cdot \varphi^{e-1}(\mathfrak{v}) = \varphi^e(\mathfrak{u}) = \mathfrak{o}$ und, da $\varphi^{e-1}(\mathfrak{v}) \neq \mathfrak{o}$, $a = \mathfrak{o}$. Es ist also $\mathfrak{x} \in U$. Wir fassen zusammen:

$$\exists \mathfrak{x} \in V : \mathfrak{x} \notin \langle \mathfrak{v} \rangle_\varphi \oplus U \wedge \varphi(\mathfrak{x}) \in U \tag{4.5}$$

Sei $\mathfrak{x}$ so ein Vektor. Dann haben wir die direkte Summe $\langle \mathfrak{v} \rangle_\varphi \oplus U \oplus \langle \mathfrak{x} \rangle$ in V. Sei $U' := U \oplus \langle \mathfrak{x} \rangle$. Dann ist U' φ-invariant und $\langle \mathfrak{v} \rangle_\varphi \cap U' = \{\mathfrak{o}\}$. Somit ist U kein maximaler φ-invarianter Untervektorraum mit $\langle \mathfrak{v} \rangle_\varphi \cap U = \{\mathfrak{o}\}$, ein Widerspruch. □

Beweis 2. (nach Fischer: Lineare Algebra) Die Idee ist: Zunächst bestimmt man maximal viele linear unabhängige Vektoren $\mathfrak{v}$, die linear unabhängig zu allen Vektoren $\mathfrak{v}'$ mit $\varphi^{n-1}(\mathfrak{v}') = \mathfrak{o}$ sind. Dann bestimmt man maximal viele linear unabhängige Vektoren $\mathfrak{v}$ mit $\varphi^{n-1}(\mathfrak{v}) = \mathfrak{o}$, die linear unabhängig zu allen Vektoren $\mathfrak{v}'$ mit $\varphi^{n-2}(\mathfrak{v}') = \mathfrak{o}$ und linear unabhängig zu allen Bildern der vorher bestimmten Vektoren sind. Allgemein bestimmt man für $i = n, \ldots, 1$ in "absteigender Weise" maximal viele linear unabhängige Vektoren $\mathfrak{v}$ mit $\varphi^i(\mathfrak{v}) = \mathfrak{o}$, die linear unabhängig zu allen Vektoren $\mathfrak{v}'$ mit $\varphi^{i-1}(\mathfrak{v}') = \mathfrak{o}$ und linear unabhängig zu allen Vektoren in φ-Erzeugnis der vorherigen Vektoren sind.

Anstatt mit linear unabhängigen Systemen zu arbeiten, ist es übersichtlicher, mit Untervektorräumen zu arbeiten und dann später in diesen Untervektorräumen Basen zu wählen.

Wir beginnen mit einer *Vorbemerkung*:

Sei $i = 1, \ldots, n-1$ und sei W ein Untervektorraum von V mit $W \cap \mathrm{Kern}(\varphi^i) = \{\mathfrak{o}\}$. (D.h. W und $\mathrm{Kern}(\varphi^i)$ bilden eine direkte Summe in V.) Dann ist $\varphi|_W : W \longrightarrow V$ injektiv (d.h. diese Abbildung induziert einen Isomorphismus $W \longrightarrow \varphi(W)$), und es ist $\varphi(W) \cap \mathrm{Kern}(\varphi^{i-1}) = \{\mathfrak{o}\}$.

(Für die erste Aussage brauchen wir nur, dass $W \cap \mathrm{Kern}(\varphi) = \{\mathfrak{o}\}$ ist. Für die zweite Aussage sei $\mathfrak{v} \in \varphi(W) \cap \mathrm{Kern}(\varphi^{i-1})$. Dann ist $\mathfrak{v} = \varphi(\mathfrak{w})$ mit $\mathfrak{w} \in W$. Es ist nun $\varphi^i(\mathfrak{w}) = \mathfrak{o}$. Nach Voraussetzung folgt $\mathfrak{w} = \mathfrak{o}$, also auch $\mathfrak{v} = \varphi(\mathfrak{w}) = \mathfrak{o}$.)

Wir kommen nun zum eigentlichen Beweis.

Wir wählen zunächst ein Komplement W_n von $\mathrm{Kern}(\varphi^{n-1})$ in $V = \mathrm{Kern}(\varphi^n)$.[1] Mit anderen Worten: Wir wählen einen Untervektorraum W_n von V mit

$$W_n \oplus \mathrm{Kern}(\varphi^{n-1}) = V \ .$$

[1] Wenn V ein Vektorraum und U ein Untervektorraum von V ist, ist ein *Komplement* von U in V ein Untervektorraum W mit $U \oplus W = V$.

Nun ist $\varphi(W_n)$ ein Untervektorraum von $\mathrm{Kern}(\varphi^{n-1})$ mit $\varphi(W_{n-1}) \cap \mathrm{Kern}(\varphi^{n-2}) = \{\mathfrak{o}\}$ (nach der Vorbemerkung). Wir wählen ein Komplement Z_{n-1} von $\varphi(W_n) \oplus \mathrm{Kern}(\varphi^{n-2})$ in $\mathrm{Kern}(\varphi^{n-1})$:

$$Z_{n-1} \oplus \varphi(W_n) \oplus \mathrm{Kern}(\varphi^{n-2}) = \mathrm{Kern}(\varphi^{n-1})$$

Nun setzen wir $W_{n-1} := Z_{n-1} \oplus \varphi(W_n)$. Wie zuvor bilden W_{n-1} und $\mathrm{Kern}(\varphi^{n-2})$ eine direkte Summe in $\mathrm{Kern}(\varphi^{n-1})$. Wir wählen ein Komplement von $\varphi(W_{n-1}) \oplus \mathrm{Kern}(\varphi^{n-1})$ in $\mathrm{Kern}(\varphi^{n-2})$:

$$Z_{n-2} \oplus \varphi(W_{n-1}) \oplus \mathrm{Kern}(\varphi^{n-1}) = \mathrm{Kern}(\varphi^{n-2})$$

So fahren wir fort. Genauer machen wir die folgenden Definitionen und treffen die folgenden Auswahlen: Wir wählen W_n wie oben, und für $i := n-1, \ldots, 1$ in absteigender Weise wählen wir einen Untervektorraum Z_i in $\mathrm{Kern}(\varphi^i)$ mit

$$Z_i \oplus \varphi(W_{i+1}) \oplus \mathrm{Kern}(\varphi^{i-1}) = \mathrm{Kern}(\varphi^i)$$

und setzen

$$W_i := Z_i \oplus \varphi(W_{i+1}) \ .$$

Wir haben dann also

$$W_i \oplus \mathrm{Kern}(\varphi^{i-1}) = \mathrm{Kern}(\varphi^i)$$

und $\varphi(W_{i+1}) \subseteq W_i$.

Ferner setzen wir $Z_n := W_n$. Wir haben dann

$$\begin{aligned} V &= W_n \oplus \mathrm{Kern}(\varphi^{n-1}) \\ &= W_n \oplus W_{n-1} \oplus \mathrm{Kern}(\varphi^{n-2}) \\ &= \cdots \\ &= W_n \oplus \cdots \oplus W_i \oplus \mathrm{Kern}(\varphi^{i-1}) \\ &= \cdots \\ &= W_n \oplus W_{n-1} \oplus \cdots \oplus W_1 \ . \end{aligned}$$

Ich behaupte, dass ferner

$$W_i = Z_i \oplus \varphi(Z_{i+1}) \oplus \cdots \oplus \varphi^{n-i}(Z_n) = \bigoplus_{j=0}^{n-i} \varphi^j(Z_{i+j})$$

ist. Man sieht leicht, dass W_i eine Summe von $Z_i, \ldots, \varphi^{n-i}(Z_n)$ ist, wir müssen aber zeigen, dass die Summe direkt ist. Dies kann man per Induktion über i von $i = n$ bis $i = 1$ (absteigend) zeigen. Der Induktionsanfang ist trivial, denn per Definition ist $W_n = Z_n$. Zum Induktionsschritt von i nach $i-1$ (mit $i \geq 2$). Nach Induktionsvoraussetzung ist

$W_i = \bigoplus_{j=0}^{n-i} \varphi^j(Z_{i+j})$ und somit $\varphi(W_i) = \varphi(\bigoplus_{j=0}^{n-i} \varphi^j(Z_{i+j}))$. Nach Vorbemerkung ist $\varphi|_{W_i}$ injektiv, und dies impliziert, dass die direkte Summe unter φ "erhalten bleibt". Wir haben $\varphi(W_i) = \bigoplus_{j=1}^{n-i+1} \varphi^j(Z_{i+j})$ und somit $W_{i-1} = Z_{i-1} \oplus \varphi(W_i) = \bigoplus_{j=0}^{n-(i-1)} \varphi^j(Z_{(i-1)+j})$.

Wir haben somit

$$\begin{aligned} V &= W_1 \oplus W_2 \oplus \cdots \oplus W_n \\ &= \bigoplus_{i=1}^{n} \bigoplus_{j=0}^{n-i} \varphi^j(Z_{i+j}) = \bigoplus_{k=1}^{n} \bigoplus_{j=0}^{k-1} \varphi^j(Z_k) \ . \end{aligned}$$

Wir wählen nun in den Räumen Z_k Basen $\mathfrak{B}_k$. Da wie schon gesagt die Abbildungen $\varphi|_{W_i}$ für $i \geq 2$ injektiv sind, ist dann $\varphi^j(\mathfrak{B}_k)$ eine Basis von $\varphi^j(Z_k)$ für $j \leq k-1$.[2] Wir erhalten schließlich

$$V = \bigoplus_{k=1}^{n} \bigoplus_{j=0}^{k-1} \bigoplus_{\mathfrak{b} \in \mathfrak{B}_k} \langle \varphi^j(\mathfrak{b}) \rangle = \bigoplus_{k=1}^{n} \bigoplus_{\mathfrak{b} \in \mathfrak{B}_k} \bigoplus_{j=0}^{k-1} \langle \varphi^j(\mathfrak{b}) \rangle$$

und somit

$$V = \bigoplus_{k=1}^{n} \bigoplus_{\mathfrak{b} \in \mathfrak{B}_k} \langle \mathfrak{b} \rangle_\varphi \ .$$

□

Wie schon angekündigt suchen wir nun "schöne" Abbildungsmatrizen.

Wir betrachten hierzu zunächst den Spezialfall, dass $V = \langle \mathfrak{v} \rangle_\varphi$ und φ nilpotent ist. Mit $n = \mathrm{Dim}(V)$ ist dann $\mathfrak{v}, \varphi(\varphi), \ldots, \varphi^{n-1}(\varphi)$ eine Basis von V, und die entsprechende Abbildungsmatrix ist

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

Sei nun $V = \langle \mathfrak{v} \rangle_\varphi$ mit $\mathrm{Dim}(V) = n$ und $V = H_a$ mit $a \in K$. Dann ist $\varphi - a \cdot \mathrm{id}$ nilpotent. Bezüglich der Basis $\mathfrak{v}, (\varphi - a \cdot \mathrm{id})(\mathfrak{v}), \ldots, (\varphi - a\,\mathrm{id})^{n-1}(\mathfrak{v})$ hat also

[2] Wir haben hier φ auf ein System von Vektoren angewandt. Dies soll bedeuten, dass wir φ auf jeden einzelnen Vektor anwenden.

$\varphi - a \cdot \mathrm{id}$ die obige Abbildunsmatrix. Damit ist die Abbildungsmatrix von φ bzgl. dieser Basis gleich

$$\begin{pmatrix} a & & & & \\ 1 & a & & & \\ & 1 & \ddots & & \\ & & \ddots & a & \\ & & & 1 & a \end{pmatrix}.$$

Definition Sei $a \in K$ und $n \in \mathbb{N}$. Dann heißt die obige Matrix *Jordankästchen* zum Eigenwert a der Dimension n. Die Matrix wird mit $J(a, n)$ bezeichnet.

Mit Satz 4.3, Satz 4.4 und den obigen Überlegungen erhalten den folgenden Satz:

Satz 4.5 *Sei V ein endlich erzeugter K-Vektorraum und φ ein Endomorphismus von V. Das charakteristische Polynom von φ zerfalle in Linearfaktoren, $\chi_\varphi = (-1)^n \cdot \prod_{i=1}^{k} (T - a_i)^{n_i}$ mit paarweise verschiedenen $a_i \in K$ und $n_i \in \mathbb{N}$. Sei ferner $\mu_\varphi = \prod_{i=1}^{k} (T - a_i)^{m_i}$. Dann gibt es eine Basis von V, bezüglich welcher die Abbildungsmatrix von φ Blockdiagonalgestalt hat, wobei die Diagonalblöcke Jordankästchen zu den Eigenwerten $a_1, \ldots, a_k$ sind. Ferner ist die Anzahl der Kästchen $J(a_i, j)$ (mit $i = 1, \ldots, k, j \in \mathbb{N}$) in der Matrix durch φ eindeutig bestimmt, und es gilt für $i = 1, \ldots, k$:*

- *Die Anzahl der Kästchen zum Eigenwert a_i ist gleich der Dimension des Eigenraums E_{a_i} (der geometrischen Vielfachheit von a_i),*
- *die Summe der Dimensionen der Kästchen zum Eigenwert a_i ist gleich n_i (der algebraischen Vielfachheit von a_i),*
- *die Dimension des größten Kästchens zum Eigenwert a_i ist gleich m_i.*

Beweis. Wir müssen nur noch die Eindeutigkeit zeigen. Man sieht aber leicht, dass die folgende Formel gilt: Die Anzahl der Jordankästchen der Größe j zum Eigenwert a_i gleich

$$\begin{aligned} &\mathrm{Dim}((\varphi - a_i \,\mathrm{id})^j) - \mathrm{Dim}((\varphi - a_i \,\mathrm{id})^{j-1}) \\ &\qquad -(\mathrm{Dim}((\varphi - a_i \,\mathrm{id})^{j+1}) - \mathrm{Dim}((\varphi - a_i \,\mathrm{id})^j)) \\ &= 2\,\mathrm{Dim}((\varphi - a_i \,\mathrm{id})^j) - \mathrm{Dim}((\varphi - a_i \,\mathrm{id})^{j-1}) - \mathrm{Dim}((\varphi - a_i \,\mathrm{id})^{j+1}) \,. \end{aligned}$$

□

Bemerkung Wie schon bei Satz 2.8 ist die Voraussetzung, dass das charakteristische Polynom in Linearfaktoren zerfalle, immer erfüllt, wenn der Körper K algebraisch abgeschlossen ist, und dies ist insbesondere für $K = \mathbb{C}$ der Fall.

Korollar 4.15 *Eine Matrix ist genau dann diagonalisierbar, wenn ihr Minimalpolynom in paarweise verschiedene normierte Linearfaktoren zerfällt.*

Definition Eine Blockdiagonalmatrix, deren Diagonalblöcke Jordankästchen sind, heißt eine *Matrix in Jordannormalform.*

Insbesondere hat also jeder Endomorphismus eines endlich erzeugten komplexen Vektorraums eine Abbildungsmatrix in Jordannormalform, und diese ist bis auf Anordnung der Kästchen eindeutig durch den Endomorphismus bestimmt. Für Matrizen erhalten wir:

Satz 4.6 *Sei $A \in K^{n\times n}$, so dass χ_A in Linearfaktoren zerfällt. Dann ist A ähnlich zu einer Matrix in Jordannormalform, und so eine Matrix ist durch A "bis auf Vertauschen der Kästchen" eindeutig bestimmt.*

Man sprich dann auch von "der" Jordannormalform einer Matrix. Das ist nicht ganz richtig, besser wäre von "einer" Jordannormalform zu sprechen. Wenn man allerdings eine Totalordnung auf dem Körper fixiert, kann man eine Anordnung der Kästchen vorgeben, so dass man Eindeutigkeit erhält. Wenn man diese Bemerkungen beachtet, hat man auch:

Korollar 4.16 *Zwei Matrizen in $\mathbb{C}^{n\times n}$ sind genau dann ähnlich, wenn sie dieselbe Jordansche Normalform haben.*

Um "die" Jordannormalform einer Matrix A wie oben auszurechnen, muss man zuerst die Eigenwerte bestimmen. Seien diese $a_1, \ldots, a_k$. Dann kann man $\mathrm{Dim}(\mathrm{Kern}(A - a_i \cdot \mathrm{id})^j)$ berechnen und daraus die Struktur der Jordannormalform ableiten. Um eine geeignete Basis zu bestimmen, kann man sich an Beweis 2 von Satz 4.14 orientieren.

In Holz, Wille: Repetitorium der Linearen Algebra, Teil 2, wird detallierter beschrieben, wie man aus dem Beweis einen Algorithmus erhält. Ein ganz anderes Verfahren wird implizit im englischsprachigen Artikel zur Jordannormalform auf Wikipedia[3] beschrieben.

[3] am 4.7.2010

4.3 Beliebige Systeme von Vektoren

Wir kommen nun noch zu einem Themenbereich, den wir bisher vernachlässigt haben: beliebige, möglicherweise unendliche, Systeme von Vektoren in einem Vektorraum.

Dieser Abschnitt kann man Ergänzung zu Abschnitt 2.4 betrachtet werden und hat keinen Bezug zu dem beiden vorherigen Abschnitten dieses Kapitels.

Sei im Folgenden K ein Körper und V ein K-Vektorraum.

Wir starten mit einer Wiederholung. Sei I eine beliebige Menge. Dann ist ein *System* (oder eine *Familie*) von Vektoren in V über der "Indexmenge" I, sagen wir $(v_i)_{i\in I}$, nichts anderes als die Abbildung $I \longrightarrow V$. Explizit ist das System $(v_i)_{i\in I}$ per Definition identisch mit der Abbildung $I \longrightarrow V\ ,\ i \mapsto v_i$. Wenn I endlich ist, sprechen wir auch von einem *endlichen System* von Vektoren. Ein Tupel $(\mathfrak{v}_1, \ldots, \mathfrak{v}_n)$ von Vektoren ist nichts anderes als ein System von Vektoren über der Indexmenge $\{1, \ldots, n\}$.

Sei nun so ein System $(v_i)_{i\in I}$ gegeben und sei $J \subseteq I$. Dann haben wir das *Teilsystem* $(v_j)_{j\in J}$ des Systems $(v_i)_{i\in I}$. Dies ist nichts anderes als die Restriktion der Abbildung $I \longrightarrow V\ ,\ i \mapsto v_i$ zu J. Wir nennen dann $(v_i)_{i\in I}$ ein *Obersystem* von $(v_j)_{j\in J}$. Wenn $J \subsetneq I$ ist, sprechen wir von einem *echten* Teil- bzw. Obersystem. Wenn J endlich ist, sprechen wir von einem *endlichen Teilsystem*. Diese Definitionen verallgemeiner wir noch etwas: Wenn J eine Menge und $\iota : J \hookrightarrow I$ eine Inklusion ist, erhalten wir aus dem System $(\mathfrak{v}_i)_{i\in I}$ das System $(\mathfrak{v}_{\iota(j)})_{j\in J}$. Wir nennen dann $(\mathfrak{v}_{\iota(j)})_{j\in J}$ wieder ein *Teilsystem* von $(\mathfrak{v}_i)_{i\in I}$ und umgekehrt $(\mathfrak{v}_i)_{i\in I}$ ein *Obersystem* von $(\mathfrak{v}_{\iota(j)})_{j\in J}$. Dies bezieht sich dann immer auf die feste Inklusion $\iota : J \hookrightarrow I$. Wenn die Inklusion nicht surjektiv (d.h. nicht bijektiv) ist, sprechen wir wieder von einem echten *Teil-* bzw. *Obersystem*.

Wir wollen nun die Begriffe *Erzeugendensystem*, *linear unabhängiges System* und *Basis* auf beliebige Systeme anwenden. Die Begriffe übertragen sich sofort auf *endliche Systeme* von Vektoren. (Es macht keinen Unterschied, ob $I = \{1, \ldots, n\}$ oder eine beliebige *endliche* Menge ist. Zum Beispiel ist offensichtlich, was für *endliches* I und ein System $(\mathfrak{v}_i)_{i\in I}$ unter der Summe $\sum_{i\in I} \mathfrak{v}_i \in V$ gemeint ist. – Wir benutzen hier natürlich, dass die Operation $+$ auf V kommutativ ist, es bei der Summation also nicht auf die Reihenfolge ankommt.)

Wir hatten für eine beliebige Teilmenge $S \subseteq V$ schon das Erzeugnis $\langle S\rangle$ von S in V definiert. Dies ist der kleinste Untervektorraum von V, der die Menge S umfasst. Er besteht explizit aus allen Linearkombinationen jeweils endlich vieler Vektoren in S.

Entsprechend definieren wir für ein System $(\mathfrak{v}_i)_{i\in I}$:

Definition

a) Das *Erzeugnis* von $(\mathfrak{v}_i)_{i\in I}$ ist $\langle\{\mathfrak{v}_i \mid i \in I\}\rangle$. Das System $(\mathfrak{v}_i)_{i\in I}$ heißt *Erzeugendensystem* von V, wenn $\langle\{\mathfrak{v}_i \mid i \in I\}\rangle = V$ ist.

b) Das System $(\mathfrak{v}_i)_{i\in I}$ heißt *linear unabhängig*, wenn jedes endliche Teilsystem des Systems linear unabhängig ist.

Eleganter kann man diese Begriffe wie folgt ausdrücken: Sei wiederum I eine beliebige Menge. Sei nun $(\mathfrak{v}_i)_{i\in I}$ ein System von Vektoren, so dass nur endlich viele $i \in I$ mit $\mathfrak{v}_i \neq \mathfrak{o}$ existieren. Dann definieren wir $\sum_{i\in I} \mathfrak{v}_i \in V$ als die Summe über diese *endlich vielen* Vektoren $\neq \mathfrak{o}$.

Wir definieren nun für eine Menge I:

$$K^{(I)} := \{(a_i)_{i\in I} \in K^I \mid \text{Für alle bis auf endlich viele } i \in I \text{ ist } a_i = 0\}$$

Dies ist ein Untervektorraum von K^I. Dieser Raum wird nun die Rolle des Raums $K^n = K^{\{1,\ldots,n\}}$ einnehmen, den wir für n-Tupel, d.h. für Systeme mit Indexmenge $\{1,\ldots,n\}$, benutzt haben. Wir nennen diesen Raum den *Standardvektorraum* zur Indexmenge I.

Sei nun $(\mathfrak{v}_i)_{i\in I}$ ein *beliebiges* System von Vektoren mit Indexmenge I und $(a_i)_{i\in I} \in K^{(I)}$. Dann haben wir das System von Vektoren $(a_i\mathfrak{v}_i)_{i\in I}$, und es gilt natürlich $a_i\mathfrak{v}_i = \mathfrak{o}$ für alle bis auf endlich viele $i \in I$. Somit haben wir wieder die Summe $\sum_{i\in I} a_i\mathfrak{v}_i \in V$. Wir erhalten hiermit eine Abbildung

$$\varphi : K^{(I)} \longrightarrow V \ , \ (a_i)_{i\in I} \longrightarrow \sum_{i\in I} a_i\mathfrak{v}_i \ .$$

Man sieht leicht, dass diese Abbildung linear ist. Nun gilt offensichtlich:

- $(\mathfrak{v}_i)_{i\in I}$ ist genau dann ein Erzeugendensystem, wenn die Abbildung φ surjektiv ist.
- $(\mathfrak{v}_i)_{i\in I}$ ist genau dann linear unabhängig, wenn die Abbildung φ injektiv ist.

Wir definieren:

Definition Das System $(\mathfrak{v}_i)_{i\in I}$ ist eine *Basis* von V, wenn die Abbildung φ ein *Isomorphismus* ist.

Somit ist $(\mathfrak{v}_i)_{i\in I}$ also genau dann eine Basis, wenn es zu jedem $\mathfrak{v} \in V$ genau ein $(a_i)_{i\in I} \in K^{(I)}$ mit $\mathfrak{v} = \sum_{i\in I} a_i\mathfrak{v}_i$ gibt. Wenn nun $(\mathfrak{b}_i)_{i\in I}$ eine Basis ist und $\mathfrak{v} \in V$ mit $\mathfrak{v} = \sum_{i\in I} a_i\mathfrak{b}_i$ ist, sagen wir auch: "$\mathfrak{v}$ hat die Koordinaten

a_i ($i \in I$) bezüglich der Basis $(\mathfrak{b}_i)_{i\in I}$". Wie zuvor nennen wir die Abbildung $c := \varphi^{-1} : V \longrightarrow K^{(I)}$ dann *Koordinatenabbildung* zur vorgegebenen Basis.

Wir wollen nun Aussage 2.15 verallgemeinern. Unter einem *maximalen linear unabhängigen System* verstehen wir ein System $(\mathfrak{v}_i)_{i\in I}$, das kein echtes linear unabhängiges Obersystem hat. Unter einem *minimalen Erzeugendensystem* verstehen wir ein System $(\mathfrak{v}_i)_{i\in I}$, das kein echtes Teilsystem hat, welches auch ein Erzeugendensystem ist.

Aussage 4.17 *Die folgenden Aussagen sind äquivalent:*

a) $(\mathfrak{v}_i)_{i\in I}$ *ist eine Basis von* V*.*

b) $(\mathfrak{v}_i)_{i\in I}$ *ist ein linear unabhängiges Erzeugendensystem von* V*.*

c) $(\mathfrak{v}_i)_{i\in I}$ *ist ein maximales linear unabhängiges System von* V*.*

d) $(\mathfrak{v}_i)_{i\in I}$ *ist ein minimales Erzeugendensystem von* V*.*

Der *Beweis* ist analog zum Beweis von Aussage 2.15.

Der *Standardvektorraum* $K^{(I)}$ hat wiederum eine *Standardbasis*. Diese ist vollkommen analog zur Standardbasis in K^n definiert: Wir setzen für $i \in I$: $e_i := (\delta_{i,j})_{j\in I}$. Dann ist für $(a_i)_{i\in I} \in K^{(I)}$:

$$(a_i)_{i\in I} = \sum_{j\in I} a_j e_j \ .$$

Es ist hiermit offensichtlich, dass $(e_i)_{i\in I}$ eine Basis von $K^{(I)}$ ist. Wir nennen diese Basis die *Standardbasis* des $K^{(I)}$.[4]

Der folgende Satz verallgemeinert Aussage 2.21.

Aussage 4.18 *Sei* $(\mathfrak{b}_i)_{i\in I}$ *eine Basis von* V*. Sei* W *ein weiterer* K*-Vektorraum, und sei* $(\mathfrak{x}_i)_{i\in I}$ *ein System von Vektoren in* W*. Dann gibt es genau eine lineare Abbildung* $\varphi : V \longrightarrow W$ *mit* $\varphi(\mathfrak{b}_i) = \mathfrak{x}_i$ *für alle* $i \in I$*, und diese ist durch*

$$\varphi(\sum_{i\in I} a_i \mathfrak{b}_i) = \sum_{i\in I} a_i \mathfrak{x}_i$$

für $(a_i)_{i\in I} \in K^{(I)}$ *gegeben. Dabei gilt* $\mathrm{Bild}(\varphi) = \langle\{\mathfrak{x}_i \mid i \in I\}\rangle_K$*. Ferner ist* φ

- *genau dann injektiv, wenn das System* $(\mathfrak{x}_i)_{i\in I}$ *linear unabhängig ist,*

[4] Vergleichen Sie die Definition von $K^{(I)}$ und die Definition der Standardbasis mit der Definition des Polynomrings $R[X]$ über einem kommutativen Ring R in Abschnitt 1.9!

- *genau dann surjektiv, wenn das System* $(\mathfrak{x}_i)_{i\in I}$ *ein Erzeugendensystem von* W *ist,*
- *genau dann bijektiv, wenn das System* $(\mathfrak{x}_i)_{i\in I}$ *eine Basis von* W *ist.*

Als Spezialfall erhalten wir die folgende Aussage:

Aussage 4.19 *Sei* V *ein* K*-Vektorraum und* $(\mathfrak{v}_i)_{i\in I}$ *ein System von Vektoren in* V*. Dann ist die Abbildung* $K^{(I)} \longrightarrow V$, $(a_i)_{i\in I} \longrightarrow \sum_{i\in I} a_i\mathfrak{v}_i$ *die eindeutig bestimmte lineare Abbildung mit* $e_i \mapsto \mathfrak{v}_i$ *für alle* $i \in I$*.*

Anstatt Systeme von Vektoren zu betrachten, kann man auch *Teilmengen* von V betrachten. Sei V eine Teilmenge von V. Wir haben schon definiert, was das *Erzeugnis* von S ist und wann wir S eine *erzeugende Menge* von V nennen.

Ferner definieren wir: S ist *linear unabhängig*, wenn für alle Systeme $\mathfrak{v}_1, \ldots, \mathfrak{v}_n$ mit $\mathfrak{v}_i \neq \mathfrak{v}_j$ für $i \neq j$ (und n beliebig) gilt: Die Vektoren $\mathfrak{v}_1, \ldots, \mathfrak{v}_n$ sind linear unabhängig. Ferner ist S eine *Basis* oder *Basismenge* von V, wenn S ein linear unabhängiges Erzeugendensystem ist.

In gewisser Hinsicht kann man Teilmengen von V als Spezialfälle von Systemen von Vektoren in V betrachten: Ein System von Vektoren besteht aus einer Indexmenge I und einer Abbildung $I \longrightarrow V$. Wenn nun $S \subseteq V$ ist, können wir S auch als Indexmenge betrachten und als Abbildung die Inklusion $S \hookrightarrow V$ wählen. Wir erhalten dann also das System $(\mathfrak{v})_{\mathfrak{v}\in S}$. Damit finden alle Definitionen und Resultate über Systeme von Vektoren in V auch auf Teilmengen von V Anwendung. Natürlich sind die allgemeinen Definition für Systeme von Vektoren konsistent mit den obigen für Mengen von Vektoren.

Wenn nun $S \subseteq V$ endlich ist, haben wir die Summe $\sum_{\mathfrak{v}\in S} \mathfrak{v}$. Für beliebiges S haben wir wieder den Raum $K^{(S)}$, dessen Elemente Systeme $(a_\mathfrak{v})_{\mathfrak{v}\in S}$ mit $a_\mathfrak{v} = 0$ für alle bis auf endlich viele $\mathfrak{v} \in S$ sind. Für $(a_\mathfrak{v})_{\mathfrak{v}\in S} \in K^{(S)}$ haben wir dann die Summe

$$\sum_{\mathfrak{v}\in S} a_\mathfrak{v}\,\mathfrak{v} \in V \ .$$

Somit ist eine Teilmenge B von V genau dann eine Basis(menge) von V, wenn für alle $\mathfrak{v} \in V$ genau ein $(a_\mathfrak{b})_{\mathfrak{b}\in B} \in K^{(B)}$ mit $\mathfrak{v} = \sum_{\mathfrak{b}\in B} a_\mathfrak{b}\mathfrak{b}$ existiert.

Wir wollen noch den folgenden grundlegenden Satz beweisen.

Satz 4.7 *Sei* V *ein Vektorraum und* A *eine linear unabhängige Teilmenge in* V*. Dann hat* V *eine Basismenge, die* A *umfasst. Insbesondere hat jeder Vektorraum eine Basis.*

Der Beweis beruht wesentlich auf dem *Auswahlaxiom*. Ich wiederhole: Das Auswahlaxiom besagt: Jede Partition einer Menge hat ein Repräsentantensystem. Zusammen mit "elementaren" mengentheoretischen Aussagen impliziert das Auswahlaxiom das so genannte *Zornsche Lemma*. Hierzu zunächst einige Definitionen: Wenn X eine Menge und $\leq$ eine Ordnungsrelation auf X ist, heißt $(X, \leq)$ auch eine *geordnete Menge*. Man schreibt dann auch X anstelle von $(X, \leq)$. Wenn $\leq$ linear ist, heißt die geordnete Menge X *total geordnet*. Eine total geordnete Menge nennt man auch eine *Kette*. Um Missverständnisse zu vermeiden, nennt man eine geordnete Menge auch eine *partiell geordnete Menge*. Eine total geordnete Menge ist somit ein Spezialfall einer geordneten Menge, was das Gleiche wie eine partiell geordnete Menge ist. Wenn nun $X = (X, \leq)$ eine total geordnete Menge ist und Y eine Teilmenge von X ist, schränkt sich $\leq$ auf Y ein, und Y wird wieder eine geordnete Menge. Eine *obere Schranke* von Y in X ist dann ein Element $s \in X$ mit $y \leq s$ für alle $y \in Y$. Das Zornsche Lemma ist nun wie folgt:

Satz 4.8 (Zornsches Lemma) *Jede (partiell) geordnete nicht-leere Menge M, in der jede total geordnete Teilmenge eine obere Schranke in M hat, hat ein maximales Element.*

Wir beweisen dies nicht.

Wir zeigen nun Satz 4.7. Sei also V ein Vektorraum und A eine linear unabhängige Teilmenge von V. Sei $\mathcal{M}$ die Menge aller linear unabhänigen Teilmengen M von V mit $A \subseteq M$. Dies ist eine partiell geordnete Menge bezüglich der Inklusion. Eine Basismenge von V, die A umfasst, ist gerade ein maximales Element in $\mathcal{M}$. Um die Existenz einer Basis zu beweisen, müssen wir also zeigen, dass jede total geordnete Teilmenge $\mathcal{T}$ von $\mathcal{M}$ eine obere Schranke in $\mathcal{M}$ hat. Sei $\mathcal{T} \subseteq \mathcal{M}$ so eine Teilmenge. Sei nun

$$S := \bigcup_{M \in \mathcal{T}} M \ .$$

Wir behaupten, dass S eine obere Schranke von $\mathcal{T}$ in $\mathcal{M}$ ist. Offensichtlich gilt $M \subseteq S$ für alle $M \in \mathcal{T}$. Es stellt sich aber die Frage, ob S in $\mathcal{M}$ liegt, das heißt, ob S linear unabhängig ist. Seien hierzu $\mathfrak{v}_1, \ldots, \mathfrak{v}_n \in S$ mit $\mathfrak{v}_i \neq \mathfrak{v}_j$ für $i \neq j$. Nach Definition von S gibt es nun Mengen $M_1, \ldots, M_n \in \mathcal{T}$ mit $\mathfrak{v}_i \in M_i$. Da $\mathcal{T}$ aber total geordnet ist, können wir OE annehmen, dass $M_1 \subseteq M_2 \subseteq \cdots \subseteq M_n$ (ansonsten können wir umnummerieren und dies erreichen). Damit gilt $\mathfrak{v}_i \in M_n$ für alle i. Aber M_n ist per Definition linear unabhängig. Es folgt, dass das System $\mathfrak{v}_1, \ldots, \mathfrak{v}_n$ linear unabhängig ist. Somit liegt S in $\mathcal{M}$, ist also eine obere Schranke von $\mathcal{T}$.

Mit dem Lemma von Zorn folgt nun: Es gibt ein maximales Element in $\mathcal{M}$. Und dies besagt: Es gibt eine Basis von V, welche A umfasst. □

Wie schon in Fußnote 1 dieses Kapitels angesprochen ist ein *Komplement* eines Untervektorraums U eines Vektorraums V ein Untervektorraum W von V mit $V = U \oplus W$.

Wenn nun U ein Untervektorraum von V ist, hat nach Satz 4.7 U eine Basismenge B. Ferner hat – wieder nach Satz 4.7 – V eine Basismenge $D = B \dot\cup C$. Mit $W := \langle C \rangle$ gilt dann $V = U \oplus W$. Wir erhalten somit:

Korollar 4.20 *Jeder Untervektorraum in einem Vektorraum hat ein Komplement.*

Alternativ kann man dies übrigens auch direkt mittels des Zornschen Lemmas zeigen (Übungsaufgabe). Hieraus folgt insbesondere (Übungsaufgabe):

Aussage 4.21 *Die Abbildung $\Phi : V \longrightarrow V^{**}$ ist injektiv.*

Abschließend noch ein paar Wörter zum Dualraum V^*: Sei $(\mathfrak{b}_i)_{i\in I}$ eine Basis des Vektorraums V. Dann ist also V (mittels der Koordinatenabbildung zur Basis $(\mathfrak{b}_i)_{i\in I}$) isomorph zu $K^{(I)}$. Die Menge der Linearformen ist nun nach Aussage 4.18 in Bijektion zur Menge K^I (nicht $K^{(I)}$!): Wenn $(a_i)_{i\in I} \in K^I$ gegeben ist, ist die entsprechende Linearform definiert durch $\mathfrak{b}_i \mapsto a_i$ (und lineare Fortsetzung). Wir haben also einen Isomorphismus $K^I \longrightarrow V^*$. Insbesondere ist $(K^{(I)})^*$ selbst kanonisch isomorph zu K^I.

Nun kann der Raum K^I "viel größer" als der Raum $K^{(I)}$ sein. Beispielsweise ist $\mathbb{Q}^{(\mathbb{N})}$ abzählbar aber $\mathbb{Q}^{\mathbb{N}}$ überabzählbar (siehe "Diagonalargument" aus der Analysis).

Indem man in V eine Basis wählt, kann man sich auch davon überzeugen, dass $\Phi : V \longrightarrow V^{**}$ genau dann ein Isomorphismus ist, wenn V endlich erzeugt ist.

Literatur

Abschließend gebe ich noch einige Bücher zur Linearen Algebra an. Mein Favorit ist im Moment das Buch Lineare Algebra von Gerd Fischer. Daneben empfehle ich die Bücher von Theodor Bröcker, Siegfried Bosch, Falko Lorenz und Heiner Zieschang. Das Buch von Herrn Bröcker ist besonders für Studenten mit Nebenfach Physik geeignet. Zur Prüfungsvorbereitung empfehle ich besonders die Bücher "Repetitorium der Linearen Algebra" (Teile 1 und 2) von Wille sowie von Holz und Wille.

Inhaltsverzeichnis