Lineare Algebra I

Prof. Dr. Uwe Jannsen **Wintersemester 2005/06**

§0 Inhalte und Fragestellungen

1) Lösungen von linearen Gleichungssystemen: Ist das lineare Gleichungssystem in 4 Variablen

$$\begin{array}{ccccccccc} x & + & 2y & + & 3z & + & 4w & = & 1 \\ 2x & + & 3y & + & 4z & + & 5w & = & 1 \\ 3x & + & 4y & + & 5z & + & 6w & = & 1 \\ 4x & + & 5y & + & 6z & + & 5w & = & 1 \end{array}$$

lösbar? Wenn ja, wieviele Lösungen hat es? Welche Struktur hat die Lösungsmenge?

Allgemeiner betrachte das Gleichungssystem mit n Variablen und m Gleichungen (m und n beliebig!)

$$\begin{array}{ccccc} a_{11}x_1 + a_{12}x_2 + & \dots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + & \dots & +a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots \\ a_{m1}x_1 + & \dots & +a_{mn}x_n & = & b_m \end{array}$$

Formale Vereinfachung: schreibe

$$Ax = b$$

mit einer $(m \times n)$-Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

und Vektoren der Längen (Dimension) n bzw. m

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \; b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

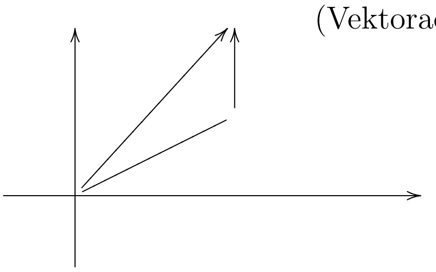
2) Formale (und praktische) Sprache: Vektorräume und lineare Abbildungen zwischen ihnen. V, W Vektorräume über einem festen Grundkörper K (z.B. $K = \mathbb{R}$ oder $K = \mathbb{C}$).

Beispiel: $\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{R}\}$ ist ein reeller Vektorraum (Vektorraum über $\mathbb{R}$). Manchmal schreiben wir die Vektoren auch als Spalten wie oben

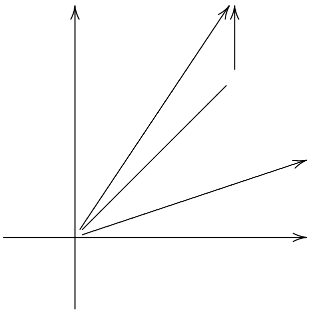
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Anschauliche Bilder für $n = 2$ und $n = 3$:
$n = 2$: Ebene

(Vektoraddition)



$n = 3$: Raum



Für höherdimensionale Räume gibt es teilweise auch physikalische Bedeutung und Anwendung:

$\mathbb{R}^4$	Raum-Zeit	(Relativitätstheorie)
$\mathbb{R}^6$	Ort-Impuls	(Klassische Mechanik)

Aber in Rechnungen (z.B. bei Optimierungsproblemen mit vielen Variablen, z.B. den Materialien einer chemischen Produktion) braucht man oft beliebig hohe Dimensionen. Das macht nichts - die Rechnerregeln sind immer diesselben, z.B.

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)\,.$$

Zurück zu den Gleichungssytemen: Matrizen können als lineare Abbildungen

$$\varphi : V \to W$$

zwischen Vektorräumen gedeutet werden. Die Matrix A oben liefert

$$\begin{array}{rcl} \mathbb{R}^n & \to & \mathbb{R}^m \\ x & \mapsto & Ax \end{array}$$

Die Lösbarkeit des Gleichungssystems wird zurückgeführt auf die Fragen: Ist A injektiv, surjektiv, bijektiv?

Außerdem werden wir einen **Dimensions**begriff entwickeln, mit Hilfe des Begriffs einer **Basis**, und die **Rangformel** benutzen

$$\begin{array}{rcl} \dim V & = & \dim\ker\varphi + \operatorname{rg}\,\varphi \\ n & = & \dim\ker A + \operatorname{rg}\,A\,. \end{array}$$

§1 Mengen

Aussagen und logische Symbole

Eine Aussage kann entweder wahr (W) oder falsch (F) sein. Seien $A, B, \ldots$ Aussagen.

$A \Leftrightarrow B$ ist wahr genau dann, wenn A und B beide wahr sind oder beide falsch sind.
$A \wedge B$ ist wahr genau dann, wenn A und B wahr sind.
$A \vee B$ ist wahr genau dann, wenn A oder B wahr sind. (kein ausschließendes "entweder oder"!)
$\neg A$ ist wahr genau dann, wenn A falsch ist.
$A \Rightarrow B$ ist wahr genau dann, wenn A falsch ist oder B wahr ist.

Wahrheitstafel

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
W	W	F	W	W	W	W
W	F	F	F	W	F	F
F	W	W	F	W	W	F
F	F	W	F	F	W	W

Corollar 1.1 (a) $\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B)$
(b) $\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B)$
(c) $(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$
$\Leftrightarrow (\neg B \Rightarrow \neg A)$
(d) $(A \Leftrightarrow B) \Leftrightarrow (A \Rightarrow B \wedge B \Rightarrow A)$

Beispiele 1.2 Betrachte die folgende Aussagen für eine natürliche Zahl n:

A	n ist gerade
B	n ist ungerade
C	n ist negativ

Dann gilt:

$A \vee B$	ist wahr
$A \wedge B$	ist falsch
C	ist falsch
$(A \vee B) \vee C$	ist wahr

Lemma 1.3 (a) $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$
(b) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$
(c) $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$
(d) $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$

Mengen

Wir halten uns nur an die folgende "naive" Definition, ohne grundlegendere Fragen der Mengenlehre zu diskutieren.

Def. 1.4 (Georg Cantor, 1845-1918) Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die Elemente der Menge genannt werden – zu einem Ganzen

Die Zusammenfassung wird durch Mengenklammern { } bezeichnet.

Beispiele 1.5
$\mathbb{N} = \{1, 2, 3, 4, 5, \ldots\}$ ist die Menge der natürlichen Zahlen.
$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\}$ ist die Menge der ganzen Zahlen.
$\mathbb{Q} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\}$ ist die Menge der rationalen Zahlen.
$\mathbb{R} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\}$ ist die Menge der reellen Zahlen.
$\mathbb{C} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, 4, \ldots\}$ ist die Menge der komplexen Zahlen.

Beispiele 1.6
$\{1, 1, 2\} = \{1, 2\}$
$\{1, \{1\}, 2\} \neq \{1, 2\}$
$\{\mathbb{N}, \mathbb{Z}\} \neq \mathbb{Z}$
$\{\{x\}\} \neq \{x\}$

Bez. 1.7 (a) $x \in M$ heißt x ist Element der Menge M (oder x liegt in M).
(b) $x \notin M \quad :\Leftrightarrow \quad \neg(x \in M)$

Beispiele 1.8: (a) $-1 \in \mathbb{Z}$, $-1 \notin \mathbb{N}$
(b) $\mathbb{N} \in \{\mathbb{N}, \mathbb{Z}\}$
(c) $1 \in \{1, \{1\}, 2\}, \{1\} \in \{1, \{1\}, 2\}$

Definition 1.9 Die leere Menge $\emptyset$ ist die Menge, die kein Element enthält.

Definition 1.10 Seien $M, N, \ldots$ Mengen.
(a) Zwei Mengen sind gleich, wenn sie diesselben Elemente enthalten, also

$$M = N \Leftrightarrow (x \in M \Leftrightarrow x \in N)$$

(b) $M \subseteq N$(M enthalten in N) wenn jedes Element von M auch in N liegt, also

$$M \subseteq N \Leftrightarrow (x \in M \Rightarrow x \in N)$$

(c) Der Durchschnitt zweier Mengen wird definiert durch

$$M \cap N = \{x \mid x \in M \wedge x \in N\}$$

(d) Die Vereinigung zweier Mengen wird definiert durch

$$M \cup N = \{x \mid x \in M \vee x \in N\}$$

(e) Das Komplement von N in M wird definiert als

$$M \setminus N = \{x \in M \mid x \notin N\}$$

(wir verlangen nicht $N \subseteq M$!)

Bild: ...

Hier haben wir, wie oft später, Mengen durch Aussagen definiert: Ist $A(x)$ eine Aussage die von einer Variable x abhängt (oft x in einer Menge N laufend), so können wir bilden

$$M = \{x \mid A(x)\},$$

die Menge derjeningen x für die $A(x)$ wahr ist. Betrachtet man dies nur für die Elemente von N, so schreibt man

$$M = \{x \in N \mid A(x)\}.$$

Beispiele 1.11 (a) $\mathbb{N} \subseteq \mathbb{Z}$, aber $\mathbb{N} \neq \mathbb{Z}$.
(b) $\mathbb{N} \in \{\mathbb{N}, \mathbb{Z}\}$ und $\{\mathbb{N}\} \subseteq \{\mathbb{N}, \mathbb{Z}\}$, aber $\mathbb{N} \nsubseteq \{\mathbb{N}, \mathbb{Z}\}$.
(c) $\mathbb{N} \cap \mathbb{Z} = \mathbb{N}$, $\mathbb{N} \cup \mathbb{Z} = \mathbb{Z}$, $\mathbb{Z} \smallsetminus \mathbb{N} = \{0, -1, -2, -3, \dots\}$
(d) Für jede Menge M gilt

$$\emptyset \subseteq M,\ M \subseteq M,\ M \cup \emptyset = M,\ M \cap \emptyset = \emptyset.$$

Insbesondere ist also die leere Menge Teilmenge jeder Menge!

Def. 1.12 Die Potenzmenge $\mathfrak{P}(M)$ einer Menge M ist die Menge ihrer Teilmengen:

$$\mathfrak{P}(M) = \{N \mid N \subseteq M\}.$$

Beispiele 1.13 (a) $\mathfrak{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$.
(b) $\{\mathbb{N}, \mathbb{Z}\} \subseteq \mathfrak{P}(\mathbb{Z})$.

Lemma 1.14 Für Mengen L, M, N gilt

(a)	$(L \cap M) \cap N$	$=$	$L \cap (M \cap N)$
(b)	$(L \cup M) \cup N$	$=$	$L \cup (M \cup N)$
(c)	$L \cap (M \cup N)$	$=$	$(L \cap M) \cup (L \cap N)$
(d)	$L \cup (M \cap N)$	$=$	$(L \cup M) \cap (L \cup N)$

Dies folgt sofort aus 1.3! (Warum?)

Bez. 1.15 (All- und Existenzquantoren)
Sei M eine Menge und $A(x)$ eine Aussage über $x (\in M)$
$\underset{x\in M}{\forall}\ A(x)$ heißt: Für alle $x \in M$ gilt $A(x)$.
$\underset{x\in M}{\exists}\ A(x)$ heißt: Es existiert ein $x \in M$, so dass $A(x)$ gilt.
$\underset{x\in M}{\exists!}\ A(x)$ heißt: Es existiert genau ein $x \in M$, so dass $A(x)$ gilt.
(Manche Bücher schreiben $\bigwedge\limits_{x\in M}$ für $\underset{x\in M}{\forall}$ bzw. $\bigvee\limits_{x\in M}$ für $\underset{x\in M}{\exists}$ bzw. $\overset{1}{\bigvee\limits_{x\in M}}$ oder $\underset{x\in M}{\exists^1}$ für $\underset{x\in M}{\exists!}$).
Wir schreiben auch oft $\forall\, x \in M : A(x)$, oder $\exists\, x \in M : A(x)$, usw.

Beispiele 1.16 (a) $\underset{x\in\mathbb{Z}}{\forall}\ x > 0$ ist falsch
(b) $\underset{x\in\mathbb{Z}}{\exists}\ x > 0$ ist richtig.

(c) $\underset{x\in\mathbb{Z}}{\exists!} \quad x > 0$ ist falsch.

(d) $\underset{x\in\mathbb{Z}}{\forall} \quad (x > 0 \Rightarrow x \in \mathbb{N})$ ist richtig.

(e) $\underset{x\in\mathbb{R}}{\forall} \quad (x > 0 \Rightarrow x \in \mathbb{N})$ ist falsch.

Es gilt die folgende wichtige Regel für die Negation:

Regel 1.17 (a) $\neg(\underset{x\in M}{\forall} A(x)) \Leftrightarrow \underset{x\in M}{\exists} \neg A(x)$

(b) $\neg(\underset{x\in M}{\exists} A(x)) \Leftrightarrow \underset{x\in M}{\forall} \neg A(x)$.

Beispiele 1.18 (a) Das Gegenteil von "Alle Hörer der LA I haben die Vorlesung verstanden" ist "Es gibt einen Hörer der LA I, der die Vorlesung nicht verstanden hat".

(b) Die Aussage 1.16 (e) war falsch. Die Negation ist

$$\begin{aligned} & \underset{x\in\mathbb{R}}{\exists} \neg(x > 0 \Rightarrow x \in \mathbb{N}) \\ \Leftrightarrow \quad & \underset{x\in\mathbb{R}}{\exists} x > 0 \wedge x \notin \mathbb{N} \end{aligned}$$

und ist richtig (wähle $x = 0,5$).

(c) Das Gegenteil von $M \subseteq N$ ist $\neg(\underset{x\in M}{\forall} x \in N)$, und dies ist äquivalent zu $\underset{x\in M}{\exists} x \notin N$.

§2 Abbildungen

Def. 2.1 Seien M und N Mengen. Eine Abbildung

$$f : M \to N$$

von M nach N ist eine Zuordnungsvorschrift, die jedem Element $x \in M$ genau ein $y \in N$ zuordnet. Dieses y wird dann mit $f(x)$ bezeichnet und heißt Bild von x unter f. M heißt Definitionsbereich oder Quelle der Abbildung und N heißt Wertebereich oder Ziel der Abbildung.

M und N sind hierbei fest vorgegeben! Insbesondere gilt

Def. 2.2 Zwei Abbildungen

$$f : M \to N, \quad \tilde{f} : \tilde{M} \to \tilde{N}$$

heißen gleich, wenn gilt:

$$M = \tilde{M} \wedge N = \tilde{N} \wedge (\underset{x\in M}{\forall} f(x) = \tilde{f}(x))\,.$$

Beispiele 2.3 (a)

$$\begin{array}{rrcl} f: & \{\text{Menschen}\} & \to & \{\text{Menschen}\} \\ & x & \mapsto & \text{Tochter von } x \end{array}$$

ist aus zwei Gründen keine Abbildung: nicht jeder Mensch hat eine Tochter (dann kann ich kein Element zuordnen), und es gibt Menschen, die mehr als eine Tochter haben (dann ist die Zuordnung nicht eindeutig).

(b) Nach 2.2 sind die Abbildungen

$$\begin{array}{rrcl} f: & \mathbb{N} & \to & \mathbb{N} \\ & x & \mapsto & x^2 \\ g: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & x^2 \end{array}$$

verschieden; es gibt also keine “Abbildung $f(x) = x^2$”.

Def. 2.4 Sei $f : M \to N$ eine Abbildung.
(a) Sei $x \in M$ und $y \in N$. Dann heißt x **Urbild von** y (unter f), wenn

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

(b) f heißt **injektiv** (und **Injektion**), wenn jedes $y \in N$ höchstens ein Urbild hat, wenn also gilt

$$\underset{x\in M}{\forall}\ \underset{z\in M}{\forall} \quad (f(x) = f(z) \Rightarrow x = z)\,.$$

$(\Leftrightarrow \underset{x\in M}{\forall}\ \underset{z\in M}{\forall} \quad (x \neq z \Rightarrow f(x) \neq f(y)))$

(c) f heißt **surjektiv** (und **Surjektion**), wenn jedes $y \in N$ (mindestens) ein Urbild hat, wenn also gilt

$$\underset{y\in N}{\forall}\ \underset{x\in M}{\exists} \quad f(x) = y\,.$$

(d) f heißt **bijektiv** (und **Bijektion**), wenn f injektiv und surjektiv ist, d.h., jedes $y \in N$ genau ein Urbild hat, wenn also gilt

$$\underset{y\in N}{\forall}\ \underset{x\in M}{\exists!} \quad f(x) = y\,.$$

In diesem Fall erhält man eine wohldefinierte Abbildung

$$\begin{array}{rrcl} f^{-1}: & N & \to & M \\ & y & \mapsto & x \text{ mit } f(x) = y \\ & & & \text{d.h., das Urbild von } y \end{array}$$

f^{-1} heißt inverse Abbildung zu f oder die Umkehrabbildung von f.

Beispiele 2.5
(a) $\begin{array}{rrcl} f: & \{a,b,c\} & \to & \{1,2\} \\ & a & \mapsto & 1 \\ & b & \mapsto & 2 \\ & c & \mapsto & 1 \end{array}$

ist surjektiv, aber nicht injektiv.

(b) $\begin{array}{rrcl} f: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & x^3 \end{array}$ ist bijektiv.

(c) $f: \begin{array}{ccc} \mathbb{N} & \to & \mathbb{N} \\ x & \mapsto & x^3 \end{array}$ ist
injektiv, aber nicht surjektiv.

Definition 2.6 Sei $f: M \to N$ eine Abbildung.
(a) Für $U \subseteq M$ heißt

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

das **Bild von** U unter f.

(b) Für $V \subseteq M$ heißt

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

das **Urbild von** V unter f.

Wir schreiben auch $f^{-1}(y)$ statt $f^{-1}(\{y\})$ für ein $y \in N$.

Beachte: In (b) setzen wir nicht voraus, dass f bijektiv ist und eine Umkehrabbildung f^{-1} existiert. Ist f aber bijektiv, mit Umkehrabbildung $g = f^{-1}$, so ist für $V \subseteq N$
$f^{-1}(V)$ (nach Def. (b))
$= g(V)$ (nach Def. (a)).

Bemerkung 2.7 Für eine Abbildung $f: M \to N$ gilt

$$f \text{ surjektiv } \Leftrightarrow f(M) = N\,.$$

Beispiele 2.8 Wir betrachten die Beispiele von 2.5
(a) Für $f: \begin{array}{ccc} \{a,b,c\} & \to & \{1,2\} \\ a & \mapsto & 1 \\ b & \mapsto & 2 \\ c & \mapsto & 1 \end{array}$
ist $f(\{a,b\}) = \{1,2\}$, $f(\{a,c\}) = \{1\}$, $f^{-1}(1) = \{a,c\}$, $f^{-1}(2) = \{b\}$.

(b) Für $f: \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$ und reelle Zahlen $a < b$ ist $f(]a,b[) =]a^3, b^3[$, und $f^{-1}(]a,b[) =]\sqrt[3]{a}, \sqrt[3]{b}[$
(Hierbei ist $]a,b[\ : \ = \{x \in \mathbb{R} \mid a < x < b\}$ das offene Intervall zwischen a und b.

(c) Für $f: \begin{array}{ccc} \mathbb{N} & \to & \mathbb{N} \\ x & \mapsto & x^3 \end{array}$ ist
$f^{-1}(\{1,2,3,\ldots,10\}) = \{1,2\}$.

Satz 2.9 Sei $f: M \to N$ eine Abbildung.

(a) Sind $U_1, U_2 \subseteq M$, so gilt

$$U_1 \subseteq U_2 \Rightarrow f(U_1) \subseteq f(U_2)\,.$$

(b) Sind $V_1, V_2 \subseteq M$, so gilt

$$V_1 \subseteq V_2 \Rightarrow f^{-1}(V_1) \subseteq f^{-1}(V_2)\,.$$

(c) Für alle $U \subseteq M$ gilt

$$f^{-1}(f(U)) \supseteq U$$

(d) f injektiv $\Leftrightarrow \underset{U \subseteq M}{\forall} \quad f^{-1}(f(U)) = U$

(e) Für alle $V \subseteq N$ gilt

$$f(f^{-1}(V)) \subseteq V$$

(f) f ist surjektiv $\Leftrightarrow \underset{V \subseteq N}{\forall} \quad f(f^{-1}(V)) = V$.

Beweis (a) Sei $U_1 \subseteq U_2$
$y \in f(U_1) \Rightarrow \underset{x \in U_1}{\exists} y = f(x) \underset{x \in U_2}{\Rightarrow} \underset{x \in U_2}{\exists} y = f(x) \Rightarrow y \in f(U_2)$

(b) Sei $V_1 \subseteq V_2$
$x \in f^{-1}(V_1) \Rightarrow f(x) \in V_1 \Rightarrow f(x) \in V_2 \Rightarrow x \in f^{-1}(V_2)$.

(c) $x \in U \Rightarrow f(x) \in f(U) \Rightarrow x \in f^{-1}(f(U))$

(d) z.z. f injektiv $\Leftrightarrow \underset{U \subseteq M}{\forall} f^{-1}(f(U)) \subseteq U$

"$\Rightarrow$" Sei f injektiv und $U \subseteq M$. Dann gilt: $x \in f^{-1}(f(U)) \overset{\text{Def.}}{\Rightarrow} f(x) \in f(U) \overset{\text{Def.}}{\Rightarrow} \underset{\tilde{x} \in U}{\exists} f(x) = f(\tilde{x}) \underset{f \text{ inj.}}{\Rightarrow} x = \tilde{x} \Rightarrow x \in U$

"$\Leftarrow$" Wir zeigen: f nicht injektiv $\Rightarrow \underset{V \subseteq M}{\exists} f^{-1}(f(U)) \subsetneq U$. Ist f nicht injektiv, so gibt es $x, \tilde{x} \in M$ mit $x \neq \tilde{x}$ aber $f(x) = f(\tilde{x})$. Sei $U = \{x\}$. Dann ist $f^{-1}(f(U)) = f^{-1}(f(x)) \supseteq \{x, \tilde{x}\}$ und $\tilde{x} \notin U$. (geht auch direkt)

(e) $y \in f(f^{-1}(V)) \overset{\text{Def.}}{\Rightarrow} \underset{x \in f^{-1}(V)}{\exists} f(x) = y$. Weiter gilt: $x \in f^{-1}(V) \overset{\text{Def.}}{\Leftrightarrow} f(x) \in V$. Zusammen folgt $y \in V$.

(f) "$\Rightarrow$" Sei f surjektiv und $V \subseteq N$. $y \in V \underset{f \text{ surj}}{\Rightarrow} \underset{x \in M}{\exists} y = f(x)$, und nach Def. folgt $x \in f^{-1}(V)$ und damit $y \in f(f^{-1}(V))$. Es gilt also $V \subseteq ff^{-1}(V)$.

"$\Leftarrow$" Sei $y \in N$. $\{y\} \subseteq f(f^{-1}(y)) \Rightarrow \underset{x \in f^{-1}(y)}{\exists} y = f(x) \Rightarrow y$ hat ein Urbild.

Satz 2.10 Sei $f : M \to N$ eine Abbildung, seien $U_1, U_2 \subseteq M$ und $V_1, V_2 \subseteq N$. Dann gilt

(a)	$f^{-1}(V_1 \cup V_2)$	$=$	$f^{-1}(V_1) \cup f^{-1}(V_2)$
(b)	$f^{-1}(V_1 \cap V_2)$	$=$	$f^{-1}(V_1) \cap f^{-1}(V_2)$
(c)	$f(U_1 \cup U_2)$	$=$	$f(U_1) \cup f(U_2)$
(d)	$f(U_1 \cap U_2)$	$\subseteq$	$f(U_1) \cap f(U_2)$ (im Allgemeinen gilt keine Gleichheit).

Beweis (a): $x \in f^{-1}(V_1 \cup V_2)$
$\Leftrightarrow f(x) \in V_1 \cup V_2$
$\Leftrightarrow f(x) \in V_1 \vee f(x) \in V_2$
$\Leftrightarrow x \in f^{-1}(V_1) \vee x \in f^{-1}(V_2)$
$\Leftrightarrow x \in f^{-1}(V_1) \cup f(V_2)$

(b): analog

(c): $y \in f(U_1 \cup U_2)$
$\Leftrightarrow \underset{x \in U_1 \cup U_2}{\exists} y = f(x)$
$\Leftrightarrow \underset{x \in U_1}{\exists} y = f(x) \vee \underset{x \in U_2}{\exists} y = f(x)$
$\Leftrightarrow y \in f(V_1) \cup f(U_2)$

(d): Übungsaufgabe!

Def. 2.11 Seien $g : L \to M$ und $f : M \to N$ Abbildungen. Dann ist die Komposition (oder die Hintereinanderausführung)

$$f \circ g : L \to N$$

von f und g definiert durch

$$(f \circ g)(x) = f(g(x)) \,.$$

Manchmal schreiben wir nur fg.

Beispiele 2.12 (a) Für jede Menge M hat man immer die identische Abbildung (oder Identität)

$$\begin{array}{rccc} \mathrm{id}_M & : & M & \to & M \\ & & x & \mapsto & x \,. \end{array}$$

Diese ist offenbar bijektiv. Ist $f : M \to N$ eine Abbildung, so ist

$$f \circ \mathrm{id}_M = f, \quad \mathrm{id}_N \circ f = f \,.$$

(b) Ist $f : M \to N$ eine bijektive Abb., so ist f^{-1} bijektiv, und es gilt

$$f^{-1} \circ f = \mathrm{id}_M, \quad f \circ f^{-1} = \mathrm{id}_N \,.$$

(c) Für $\begin{array}{l} g : \mathbb{R} \to \mathbb{R} \\ \quad x \mapsto x^2 \end{array}$ und $\begin{array}{l} f : \mathbb{R} \to \mathbb{R} \\ \quad x \mapsto x^3 \end{array}$ ist $\begin{array}{l} f \circ g = g \circ f : \mathbb{R} \to \mathbb{R}. \\ \qquad\qquad\qquad x \mapsto x^6 \end{array}$

(d) Für $\begin{array}{l} g : \mathbb{R} \to \mathbb{R} \\ \quad x \mapsto x^2 \end{array}$ und $\begin{array}{l} f : \mathbb{R} \to \mathbb{R} \\ \quad x \mapsto e^x \end{array}$ ist $f \circ g(x) = e^{x^2}$ und $g \circ f = e^{2x}$, also $f \circ g \neq g \circ f$.

.

(e) Sei $\begin{array}{l} f : \mathbb{Z} \to \mathbb{R} \\ \quad x \mapsto x \end{array}$ und $\begin{array}{l} g : \mathbb{R} \to \mathbb{Z} \\ \quad x \mapsto [x] \end{array}$
wobei $[x] :=$ größte ganze Zahl, die kleiner oder gleich x ist. Dann ist $g \circ f = \mathrm{id}_{\mathbb{Z}} : \mathbb{Z} \to \mathbb{Z}$ und $\begin{array}{l} f \circ g : \mathbb{R} \to \mathbb{R} \\ \qquad x \mapsto [x]. \end{array}$

Satz 2.13 Seien $g : L \to M$ und $f : M \to N$ Abbildungen.

Sind f und g $\left\{ \begin{array}{l} \text{injektiv} \\ \text{surjektiv} \\ \text{bijektiv} \end{array} \right\}$, so ist auch $f \circ g$ $\left\{ \begin{array}{l} \text{injektiv} \\ \text{surjektiv} \\ \text{bijektiv} \end{array} \right\}$

Beweis 1) Seien f und g injektiv.
Für $x, x' \in L$ gilt: $(f \circ g)(x) = (f \circ g)(x') \Leftrightarrow f(g(x)) = f(g(x')) \underset{f \text{ inj.}}{\Rightarrow} g(x) = g(x') \underset{g \text{ inj.}}{\Rightarrow} x = x'$. Also ist $f \circ g$ injektiv.

2) Seien f und g surjektiv. Dann ist

$$g(L) = M \quad \text{und} \quad f(M) = N\,,$$

also $(f \circ g)(L) = f(g(L)) = f(M) = N$, d.h., $f \circ g$ ist surjektiv.

3) Der dritte Fall folgt aus den ersten beiden.

Satz 2.14 Seien $f : M \to N$ und $g : N \to M$ Abbildungen. Ist $g \circ f = \mathrm{id}_M$, so ist f injektiv und g surjektiv.

Beweis Sei $g \circ f = \mathrm{id}_M$. Sind $x, x' \in M$ mit $f(x) = f(x')$, so ist auch $x = g(f(x)) = g(f(x')) = x'$. Also ist f injektiv.

Ist $x \in M$, so ist $x = g(f(x))$, x hat also das Urbild $f(x)$ unter g. Also ist g surjektiv.

Corollar 2.15 Eine Abbildung $f : M \to N$ ist genau dann bijektiv, wenn es eine Abbildung $g : N \to M$ gibt mit

$$g \circ f = \mathrm{id}_M \quad \text{und} \quad f \circ g = \mathrm{id}_N\,.$$

Beweis Die eine Richtung folgt aus 2.12(b): Ist f bijektiv, so leistet $g = f^{-1}$ das Verlangte. Die andere Richtung folgt aus 2.14.

Beispiel 2.16 Sei M eine Menge. Für $U \subseteq M$ heißt

$$C_M(U) := M \smallsetminus U$$

das Komplement (oder die Komplementärmenge) von U in M. Die Abbildung

$$\begin{array}{rrcl} C_M : & \mathfrak{P}(M) & \to & \mathfrak{P}(M) \\ & M & \mapsto & C_M(U) \end{array}$$

ist bijektiv, denn es ist $C_M \circ C_M = \mathrm{id}_{\mathfrak{P}(M)} : C_M(C_M(U) = M \smallsetminus (M \smallsetminus U) = U \quad \forall\, U \subseteq M$:

$$\begin{array}{rcl} (x \in M \wedge x \notin (M \smallsetminus U)) & \Leftrightarrow & (x \in M \wedge \neg(x \in M \wedge x \notin U)) \\ & \Leftrightarrow & x \in M \wedge (x \notin M \vee x \in U) \\ & \Leftrightarrow & (x \in M \wedge x \notin M) \vee (x \in M \wedge x \in U) \\ & \Leftrightarrow & x \in U \end{array}$$

Def. 2.17 Sei $f : M \to N$ eine Abbildung und $U \subseteq M$. Dann heißt die Abbildung

$$\begin{array}{rrcl} f_{|U} : & U & \to & N \\ & x & \mapsto & f(x) \end{array}$$

die Einschränkung von f auf U.

§3 Gruppen

Wie faßt man manthematisch Begriffe wie Addition oder Multiplikation (z.B. von natürlichen Zahlen oder reellen Zahlen, oder von Matrizen...)?

Def. 3.1 Eine (innere) Verknüpfung auf einer Menge G ist eine Abbildung

$$\mu : G \times G \to G\,.$$

Man wählt dafür auch oft Symbole wie $\circ, \cdot, +$ und schreibt zum Beispiel $g_1 \circ g_2$ statt $\mu(g_1, g_2)$, also für das Element, welches dem Paar (g_1, g_2) zugeordnet ist.

Def. 3.2 Eine Gruppe ist eine Menge G mit einer Verknüpfung

$$\circ : G \times G \to G$$

für die die folgenden Eigenschaften gelten

(i) (Assoziativität) Für alle $x, y, z \in G$ gilt

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

(ii) (Existenz eines neutralen Elementes) Es gibt ein $e \in G$ mit

$$e \circ x = x = x \circ e \quad \text{für alle } x \in G\,.$$

(iii) (Existenz von Inversen) Sei $x \in G$. Dann gibt es ein $y \in G$ mit

$$y \circ x = e = x \circ y\,.$$

Definition 3.3 Eine Gruppe $(G, \circ)$ heißt kommutativ, wenn gilt

$$\underset{x,y \in G}{\forall} \quad x \circ y = y \circ x\,.$$

Lemma 3.4 Sei $(G, \circ)$ eine Gruppe.

(a) Das neutrale Element ist eindeutig.

(b) Für jedes $x \in G$ ist das inverse Element eindeutig bestimmt; wir bezeichnen es mit x^{-1}.

Beweis (a): Seien e und e' neutrale Elemente. Dann ist

$$e \overset{*}{=} e \circ e' \overset{**}{=} e'$$

($*$, da e' neutral, $**$, da e neutral).

(b): Seien y und y' inverse Elemente zu $x \in G$. Dann ist

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

Beispiele 3.5 (a) Die Verknüpfung $+$ auf $\mathbb{N}$ ist assoziativ, aber es gibt kein neutrales Element, und keine Inversen. Auf $\mathbb{N}_0$ gibt es das neutrale Element 0, aber zu $n \in \mathbb{N}$ kein Inverses.

(b) $(\mathbb{Z}+)$ ist eine kommutative Gruppe. Das neutrale Element ist 0, das Inverse von x ist $-x$.

(c) Die Verknüpfung $\cdot$ auf $\mathbb{Z}$ ist assoziativ und kommutativ, die 1 ist ein links- und rechtsneutrales Element, aber für $x \neq 1$ hat x kein Inverses bezüglich der Multiplikation.

(d) Sei $\mathbb{Q}$ die Menge der rationalen Zahlen. Die Menge $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$ ist bezüglich der Multiplikation eine Gruppe.

(e) Die Verknüpfung

$$\begin{aligned} - : \mathbb{Z} \times \mathbb{Z} &\rightarrow \mathbb{Z} \\ (a,b) &\mapsto a - b \end{aligned}$$

ist weder assoziativ noch kommutativ; es gibt kein neutrales Element.

(f) Bei endlichen Gruppen wird die Verknüpfung manchmal durch eine **Gruppentafel** (oder Verknüpfungstafel, oder Multiplikationstafel) gegeben; in Zeile x und Spalte y trägt man $x \circ y$ ein. Zum Beispiel erhält man eine Gruppe mit 4 Elementen durch

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Dies ist die sogenannte **Kleinsche Vierergruppe**. Sie ist kommutativ (dies sieht man an der Symmetrie bezüglich der Diagonalen $\diagdown$). In einer Gruppentafel müssen in jeder Zeile und jeder Spalte alle Gruppenelemente vorkommen (also jedes Element genau einmal) (warum?).

(g) Sei G eine Gruppe mit 2 Elementen. Dann muß die Gruppentafel wie folgt aussehen (e ist das neutrale Element, und a ist das zweite Element)

	e	a
e	e	a
a	a	e

Bemerkung 3.6 (a) Das neutrale Element einer Gruppe nennt man auch das Einselement und bezeichnet es manchmal mit 1.

(b) Bei einer kommutativen Gruppe bezeichnet man die Verknüpfung oft mit $+$; dann bezeichnet man das neutrale Element mit 0 und das Inverse von x mit $-x$ (und spricht von Nullelement bzw. dem Negativen von x).

(c) Oft lassen wir das Zeichen $\circ$ für die Verknüpfung einfach weg und schreiben xy für die Verknüpfung $x \circ y$ von x mit y.

(d) Aus dem Assoziativgesetz folgt, dass man in einer Gruppe beliebig klammern kann; z.B. ist

$$(xy)(zt) = (x(yz))t\,.$$

(formaler Beweis für beliebig viele Faktoren: Übungsaufgabe!). Deshalb können wir auch Klammern ganz weglassen, d.h., xyz schreiben, bzw.

$$xyzt$$

für den obigen Ausdruck.

(e) Ist $(G, +)$ eine kommutative Gruppe, so kann man in einer Verknüpfung von meherern Elementen diese Elemente beliebig vertauschen. Also ist z.B.

$$x + y + z = z + y + x$$

(formaler Beweis für beliebig viele Elemente: Übungsaufgabe!)

Lemma 3.7 Sei G eine Gruppe. Dann gilt für alle Elemente $a, b \in G$:

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

Beweis Vorbemerkung: Ist in einer Gruppe $ba = e$ (e das neutrale Element der Gruppe), so gilt $b = a^{-1}$, wie durch Multiplikation mit a^{-1} von rechts folgt. Damit zeigen wir:

(a): $aa^{-1} = e \Rightarrow a$ Inverses von a^{-1}.

(b): $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}eb = b^{-1}b = e$, also ist $b^{-1}a^{-1}$ das Inverse von ab.

Lemma/Def. 3.8 Für jede natürliche Zahl n sei S_n die Menge der bijektiven Abbildungen

$$\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}\,.$$

Diese bildet mit der Verknüpfung

$$\begin{array}{rcll} S_n \times S_n & \to & S_n & \\ (\sigma, \tau) & \mapsto & \sigma \circ \tau & \text{(Komposition von Abbildungen)} \end{array}$$

eine Gruppe, und heißt die symmetrische Gruppe von Grad n. Die Elemente heißen Permutationen der Zahlen $1, \ldots, n$.

Beweis dass S_n eine Gruppe ist: Für **jede** Menge M ist die Menge $Bij(M, M)$ der bijektiven Abbildungen $f : M \to M$ eine Gruppe bezüglich der Komposition

$$Bij(M, M) \times Bij(M, M) \to Bij(M, M)$$

$$(f, g) \mapsto f \circ g\,,$$

denn: Die Verknüpfung ist nach 2.13 wohldefiniert ($f \circ g$ ist wieser bijektiv), und das Assoziativgesetz gilt: Es ist $(f \circ g) \circ h = f \circ (g \circ h)$ (Gleicheit von Abbildungen), denn für alle $x \in M$ gilt $((f \circ g) \circ h)(x) = (f \circ g)(h(x)) = f(g(h(x)))$, und $(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x)))$. Ein neutrales Element für die Komposition ist id_M, da

$$f \circ \mathrm{id}_M = f = \mathrm{id}_M \circ f \,,$$

und inverses Element zu f ist die Umkehrabbildung f^{-1} (siehe 2.12 (a) und (b)).

Bezeichnung 3.9 Für eine Permutation (=Vertauschung) σ schreiben wir auch σ in der Form

$$\begin{pmatrix} 1 \; 2 & \dots & n \\ \sigma(1)\sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

(unter i steht das Bild $\sigma(i)$ unter σ). Z. B. ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

die Abbildung

$$\begin{aligned} 1 &\mapsto 2 \\ 2 &\mapsto 3 \\ 3 &\mapsto 1 \,. \end{aligned}$$

Def. 3.10 Die Ordnung einer Gruppe G ist ihre Mächtigkeit und wird mit $|G|$ oder $(G : 1)$ bezeichnet (sie ist ∞ oder in $\mathbb{N}$).

Satz 3.11 Sei $n \in \mathbb{N}$. Die Ordnung der symmetrischen Gruppe S_n ist $n!\,(= 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n)$.

Beweis Die Darstellung 3.9 zeigt, dass die Permutation in S_n gerade den verschiedenen Anordnungen der Zahlen $1, \dots, n$ entsprechen. Bekanntlich gibt es hiervon $n!$ viele. (Ausführlicher Beweis mit vollständiger Induktion).

Def. 3.12 Sei $(G, \circ)$ eine Gruppe. Eine Teilmenge $U \subseteq G$ heißt Untergruppe von G, wenn U mit der Einschränkung der Verknüpfung von G wieder eine Gruppe ist.

Beispiele 3.13 (a) $\mathbb{Z}$ ist eine Untergruppe von $(\mathbb{R}, +)$.

(b) Für jedes $m \in \mathbb{Z}$ ist $m\mathbb{Z} := \{m \cdot x \mid x \in \mathbb{Z}\}$ eine Untergruppe von $(\mathbb{Z}, +)$.

Lemma 3.14 Sei G eine Gruppe, und sei $U \subseteq G$ eine nicht-leere Teilmenge. Dann sind äquivalent

(a) U ist Untergruppe

(b) Es gelten die drei Bedingungen

(i) $\underset{x,y \in U}{\forall} \; x, y \in U$ (Abgeschlossenheit unter der Verknüpfung)

(ii) $e \in U$, wenn e das neutrale Element von G ist.

(iii) $\underset{x \in U}{\forall} \; x^{-1} \in U$.

Beweis (a) $\Rightarrow$ (b): Sei U Untergruppe. Dann gilt (b)(i) nach Definition. Weiter gibt es ein $e' \in U$ mit $e'x = x = xe'$ für alle $x \in U$. Dann folgt $e' = e'e = e' \cdot e' \cdot (e')^{-1} = e' \cdot (e')^{-1} = e$, also (b)(ii). Schließlich gibt es für jedes $x \in U$ ein $y \in U$ mit $yx = e = xy$. Aus der Eindeutigkeit des Inversen in G folgt $x^{-1} = y \in U$, also (b)(iii).

(b) $\Rightarrow$ (a) ist klar.

Lemma/Definition 3.15 Seien $G_1, \ldots, G_n$ Gruppen. Dann ist das Produkt

$$\prod_{i=1}^{n} G_i = \{(g_1, \ldots, g_n) \mid g_i \in G_i\}$$

eine Gruppe bezüglich der Verknüpfung

$$(g_1, \ldots, g_n) \cdot (h_1, \ldots, h_n) := (g_1 h_1, g_2, h_2, \ldots, g_n h_n)\,.$$

Beweis Die Assoziativität folgt sofort aus den Assoziativgesetzen für die Gruppen G_i. Das Element $(e_1, \ldots, e_n)$ ist ein neutrales Element, wenn $e_i \in G_i$ das Einselement von G_i ist. Schließlich ist das Inverse von $(g_1, \ldots, g_n)$ offenbar durch $(g_1^{-1}, \ldots, g_n^{-1})$ gegeben.

§4 Ringe und Körper

Definition 4.1 Ein Ring ist eine Menge R mit zwei Verknüpfungen

$$\begin{array}{rccc} + : & R \times R & \to & R \\ \cdot : & R \times R & \to & R \end{array}$$

für die gilt:

(i) $(R, +)$ ist eine kommutative Gruppe.

(ii) Die zweite Verknüpfung $\cdot$ ist assoziativ.

(iii) (Distributivgesetze) Für alle $x, y, z \in R$ gilt:

$$\begin{array}{rcl} (x+y) \cdot z & = & x \cdot z + y \cdot z \\ x \cdot (y+z) & = & x \cdot y + x \cdot z\,. \end{array}$$

Bemerkungen 4.2 Wir nennen $+$ die **Addition** des Ringes R, und schreibn 0 für das neutrale Element von $(R, +)$, und $-x$ für das inverse Element von x in $(R, +)$. Die zweite Verknüpfung $\cdot$ heißt die **Multiplikation** von R. Eigentlich hätten wir oben schreiben müssen

$$(y+z) \cdot z = (x \cdot z) + (y \cdot z)\,,$$

aber wir vereinbaren (wie bei Rechnungen mit reellen Zahlen), dass Multiplikationen vor Addition ausgeführt werden. Weiter schreiben wir $x - y$ für $x + (-y)$, und xy für $x \cdot y$.

Definition 4.3 (a) Ein Ring $(R, +, \cdot)$ heißt **kommutativ**, wenn die Multiplikation kommutativ ist, d.h., wenn gilt

$$x \cdot y = y \cdot x \quad \text{für alle } x, y \in R\,.$$

(b) Ein Ring R heißt **Ring mit Eins**, wenn es ein neutrales Element 1 bezüglich der Multiplikation gibt, d.h., wenn es ein Element $1 \in R$ gibt mit

$$1 \cdot x = x = x \cdot 1 \quad \text{für alle } x \in R\,.$$

(c) Ein Ring R heißt trivial, wenn er nur aus der 0 besteht.

(d) Ein Körper ist ein nicht-trivialer kommutativer Ring, für den $(K \setminus \{0\}, \cdot)$ eine Gruppe ist. Schreibe auch $\frac{x}{y}$ für $xy^{-1} = y^{-1}x$, falls $x \in K$, $y \in K \setminus \{0\}$.

Beispiele 4.4 (a) $\mathbb{Z}$ mit der üblichen Addition und Multiplikation ist ein kommutativer Ring mit Eins. $\mathbb{Z}$ ist kein Körper, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist (2 hat z.B. kein Inverses).

(b) $\mathbb{Q}$ und $\mathbb{R}$ sind Körper.

(c) Bei der Matrizenrechnung werden wir später nicht-kommutative Ringe kennenlernen.

Lemma 4.5 In einem Ring mit Eins ist die 1 eindeutig.

Beweis Seien $1, 1' \in R$ mit

$$1 \cdot x = x \cdot 1 = x = 1' \cdot x = 1' \cdot x \quad \forall\, x \in R\,.$$

Es folgt

$$(1 - 1') \cdot x = 1 \cdot x - 1' \cdot x = 0 \quad \forall\, x \in R$$

Für $x = 1$ folgt $1 - 1' = 0$, also $1 = 1'$.

Proposition 4.6 Sei R ein Ring.

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

(b) Für alle $x, y \in R$ gilt $(-x) \cdot y = -xy = x \cdot (-y)$. Insbesondere gilt $(-1) \cdot x = -x = x \cdot (-1)$.

(c) Für alle $x, y \in R$ gilt $(-x) \cdot (-y) = x \cdot y$.

Beweis (a): Aus dem Distributivgesetz folgt

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

und

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

Es folgt die Behauptung, da $(R, +)$ eine Gruppe ist (auf beiden Seiten $-(x \cdot x)$ addieren).

(b): Es ist zu zeigen

$$x \cdot y + (-x) \cdot y = 0$$

und
$$x \cdot y + x \cdot (-y) = 0$$
(siehe Vorbemerkung im Beweis von Lemma 3.7). Aber nach dem Distributivgesetz ist

$$x \cdot y + (-x) \cdot y = (x - x) \cdot y = 0 \cdot y \overset{\text{(a)}}{=} 0$$

und
$$x \cdot y + x \cdot (-y) = x \cdot (y - y) = x \cdot 0 \overset{\text{(a)}}{=} 0\,.$$

Lemma 4.7 In jedem Körper K gilt für $x, y \in K$:

(a) $x \cdot y = 0 \Rightarrow x = 0 \ \vee \ y = 0$.

(b) $x^2 = 1 \Rightarrow x = 1 \ \vee \ x = -1$. (Hierbei ist $x^2 = x \cdot x$).

Beweis (a): Dies ist äquivalent zur Implikation

$$\neg(x = 0 \vee y = 0) \ \Rightarrow \ \neg(x \cdot y = 0)\,,$$

also zur Implikation
$$x \neq 0 \ \wedge \ y \neq 0 \ \Rightarrow \ x \cdot y \neq 0\,.$$
Dies gilt aber, da $K \setminus \{0\}$ abgeschlossen unter der Multiplikation ist.

(b): Es ist $1 \cdot 1 = 1$ und $(-1) \cdot (-1) \overset{4.6(c)}{=} 1 \cdot 1 = 1$. Umgekehrt gelte $x^2 = 1$. Dann ist

$$\begin{aligned} (x+1)(x-1) &= x^2 + x(-1) + 1x + 1(-1) \\ &= x^2 - x + x - 1 = 0 \end{aligned}$$

Nach (a) ist also $x + 1 = 0$ oder $x - 1 = 0$, also $x = -1$ oder $x = 1$.

Definition 4.8 Ein Teilkörper eines Körpers K ist eine Teilmenge $L \subseteq K$, die durch die Einschränkungen der Verknüpfungen $+$ und $\cdot$ von K wieder ein Körper ist. (Analog definiert man Unterringe von Ringen).

Definition 4.9 Der Körper $\mathbb{C}$ der komplexen Zahlen ist wie folgt definiert:

1. Definition (Physiker): $\mathbb{C}$ besteht aus allen Ausdrücken der Form $a + bi$ mit $a, b \in \mathbb{R}$, und die Ringoperationen sind definiert durch

$$\begin{aligned} (a + bi) + (c + di) &= (a + c) + (b + d)i \\ (a + bi) \cdot (c + di) &= (ac - bd) + (ad + bc)i\,. \end{aligned}$$

Identifiziere die reelle Zahl a mit $a + 0i$ und schreibe i für $0 + 1i$.

Beobachtung: Dann ist $i^2 = -1$ und $a + bi = a + b \cdot i$, wobei $+$ und $\cdot$ die Ringoperationen sind.

2. Definition (Mathematiker) Definiere auf der kommutativen Gruppe $(\mathbb{R}^2, +)$ eine Multiplikation $\cdot$ durch
$$(a, b) \cdot (c, d) = (ac - bd,\, ad + bc)\,.$$

Der so erhaltene Ring ist der Körper $\mathbb{C}$ der komplexen Zahlen. Identifiziere $\mathbb{R}$ mit einem Teilkörper von $\mathbb{C}$ durch $a \mapsto (a, 0)$, und schreibe i für $(0, 1)$.

Beobachtung: Es ist $i^2 = -1$, und jedes Element läßt sich schreiben als $a + b \cdot i$, wobei $+$ und $\cdot$ die Ringoperationen sind.

Wir können also in jedem Fall die Elemente von $\mathbb{C}$ in der Form $z = a + bi$ mit eindeutig bestimmten $a, b \in \mathbb{R}$ schreiben, wobei $i^2 = -1$, und $\mathbb{R}$ ist der Teilkörper von $\mathbb{C}$ der Zahlen z mit $b = 0$.

Beweis dass wir einen Körper erhalten:

1) Dass $(\mathbb{C}, +)$ eine Gruppe ist, ist klar (insbesondere bei der 2. Definition).

2) Man rechnet sofort nach, dass $\cdot$ assoziativ und kommutativ ist, und dass die Distributivgesetze gelten.

3) $1 = (1, 0) = 1 + 0 \cdot i$ ist ein Einselement bezüglich der Multiplikation.

4) $\mathbb{C} \setminus \{0\}$ ist ein Körper: Beobachtung: Es ist

$$(a + bi)(a - bi) = a^2 - (bi)^2 = a^2 + b^2 \,.$$

Ist nun $a + bi \neq 0$, so ist $a \neq 0$ oder $b \neq 0$, also $a^2 + b^2 > 0$ in $\mathbb{R}$, also $a^2 + b^2 \neq 0$. Dann besitzt $a^2 + b^2$ das Inverse $\frac{1}{a^2+b^2}$ in $\mathbb{R}$, und wir haben

$$1 = (a + bi)(a - bi)\frac{1}{a^2 + b^2} \,,$$

d.h., $a + bi$ hat das Inverse

$$(a - bi)\frac{1}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i \,.$$

Definition 4.10 Für eine komplexe Zahl

$$z = a + bi$$

heißt $\mathrm{Re}(z) := a$ der **Realteil** von z und $\mathrm{Im}(z) := b$ der **Imaginärteil** von z. Die Zahl

$$\overline{z} = a - bi$$

heißt das **komplex Konjugierte** von z, und

$$|z| = \sqrt{a^2 + b^2}$$

heißt der Betrag von z.

Oben haben wir gesehen:

$$z \cdot \overline{z} = a^2 + b^2 = |z|^2 \,.$$

Weiter gilt für $z_1, z_2 \in \mathbb{C}$

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2| \,,$$

wie man leicht nachrechnet (Zeige $|z_1 \cdot z_2|^2 = |z_1|^2 \cdot |z_2|^2$).

§5 Vektorräume

Die abelschen Gruppen $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3$ können als Mengen von Vektoren visualisiert werden, und man kann diese Vektoren mit Skalaren multiplizieren. Die mathematische Axiomatisierung ist:

Definition 5.1 Sei K ein Körper. Ein Vektorraum über K (oder K-Vektorraum) ist eine kommutative Gruppe $(V, +)$ zusammen mit einer **äußeren** Verknüpfung

$$\begin{array}{rrcl} \cdot : & K \times V & \to & V \\ & (\lambda, v) & \mapsto & \lambda \cdot v \end{array}$$

(genannt Skalarmultiplikation) mit folgenden Eigenschaften: Für alle $\lambda, \lambda_1, \lambda_2 \in K$ und $v, v_1, v_2 \in V$ gilt

(i) $(\lambda_1 + \lambda_2) \cdot v = \lambda_1 \cdot v + \lambda_2 \cdot v$ (1. Distributivgesetz)

(ii) $\lambda \cdot (v_1 + v_2) = \lambda \cdot v_1 + \lambda \cdot v_2$ (2. Distributivgesetz)

(iii) $\lambda_1 \cdot (\lambda_2 \cdot v) = (\lambda_1 \cdot \lambda_2) \cdot v$ ("Assoziativität")

(iv) $1 \cdot v = v$ (Neutralität der Eins).

Man beachte: es gibt zwei verschiedene Additionen $+$ und zwei verschiedene Multiplikationen $\cdot$: $+$ in K und in V, $\cdot$ in K und $\lambda \cdot v$ für $\lambda \in K$ und $v \in V$.

Generell lassen wir wieder einige Klammern weg, nach dem Motto "$\cdot$ geht vor $+$". Die Eelemente aus K werden oft Skalare genannt, die Elemente aus V Vektoren.

Beispiele 5.2 Sei K ein Körper.

(a) Für $n \in \mathbb{N}$ machen wir

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

immer wie folgt zu einem K-Vektorraum:

• Die Addition ist die von $+$ in K auf dem kartesischen Produkt induzierte (vergl. 3.15), d.h., wir setzen

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1,\, x_2 + y_2, \ldots, x_n + y_n)\,.$$

• Die Skalarmultiplikation ist definiert durch

$$\lambda \cdot (x_1, \ldots, x_n) := (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)\,.$$

(b) Sei M eine Menge. Dann machen wir die Menge

$$\mathrm{Abb}(M, \mathbb{R}) := \{f : M \to \mathbb{R}\}$$

zu einem $\mathbb{R}$-Vektorraum wie folgt:

- Für $f, g \in \mathrm{Abb}(M, \mathbb{R})$ definieren wir $f + g \in \mathrm{Abb}(M, \mathbb{R})$ durch

$$(f+g)(m) := f(m) + g(m)\,.$$

- Für $\lambda \in \mathbb{R}$ und $f \in \mathrm{Abb}(M, \mathbb{R})$ definieren wir $\lambda \cdot f \in \mathrm{Abb}(M, \mathbb{R})$ durch

$$(\lambda \cdot f)(m) := \lambda \cdot f(m)\,.$$

(c) Sei L ein Körper und $K \subseteq L$ ein Teilkörper. Dann ist L mittels der Verknüpfungen $+$ und $\cdot$ von L ein K-Vektorraum. Insbesondere ist $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum (da $\mathbb{R}$ ein Teilkörper von $\mathbb{C}$ ist).

Lemma 5.3 Sei K ein Körper und V ein K-Vektorraum.

(a) Für alle $v \in V$ gilt $0 \cdot v = 0$

(b) Für alle $\alpha \in K$ gilt $\alpha \cdot 0 = 0$.

(c) Für alle $v \in V$ gilt $(-1) \cdot v = -v$.

Beweis (a): Für alle $v \in V$ gilt

$$0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v\,,$$

also $0 = 0 \cdot v$, da $(V, +)$ eine Gruppe ist.

(b): $\alpha \cdot 0 = \alpha(0+0) = \alpha \cdot 0 + \alpha \cdot 0 \Rightarrow 0 = \alpha \cdot 0$

(c): $v + (-1)v = 1 \cdot v + (-1) \cdot v = (1-1) \cdot v = 0 \cdot v \overset{\text{(a)}}{=} 0$

Definition 5.4 Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper K. Eine Teilmenge $W \subseteq V$ heißt Untervektorraum (oder Unterraum), wenn die Einschränkungen von $+$ auf $\cdot$ auf W wieder die Struktur eines K-Vektorraums definieren.

Satz 5.5 (Unterraum-Kriterium) Sei V ein K-Vektorraum und $W \subseteq V$ nicht-leer. Dann sind äquivalent:

(a) Für alle $v, w \in W$ ist $v + w \in W$ und für alle $v \in W$ und alle $\alpha \in K$ ist $\alpha \cdot v \in W$.

(b) Für alle $v, w \in W$ und alle $\alpha, \beta \in K$ ist $\alpha v + \beta w \in W$.

(c) W ist K-Unterraum von V.

Beweis (c) $\Rightarrow$ (b) ist trivial
(b) $\Rightarrow$ (a): Setze $(\alpha, \beta) = (1, 1)$ und $(\alpha, \beta) = (\alpha, 0)$
(a) $\Rightarrow$ (c): Nach (a) ist W abgeschlossen bezüglich der Verknüpfungen $+$ und $\cdot$. Nach Voraussetzung existiert ein $v \in W$; dann folgt nach 5.3: $0 = 0 \cdot v \in W$. Weiter ist für jedes $v \in W$ auch $-v = (-1) \cdot v \in W$. Also ist W auch eine Untergruppe von $(V, +)$.

Beispiele 5.6 Sei K ein Körper.

(a) Für $m, n \in \mathbb{N}$ mit $m \leq n$ ist

$$W = \{(x_1, \ldots, x_n) \in K^n \mid x_i = 0 \quad \forall\, i > n\}$$

ein Unterraum von K^n.

(b) Betrachte die folgenden Untermengen von $\mathbb{R}^2$:

$$\begin{array}{rcl} W_1 & = & \{(x,y) \in \mathbb{R}^2 \quad \mid \quad 3x + 4y = 0\} \\ W_2 & = & \{(x,y) \in \mathbb{R}^2 \quad \mid \quad 3x + 4y = 5\}\,. \end{array}$$

W_1 ist ein Untervektorraum, W_2 nicht (z.B. $0 = (0,0) \notin W_2$)

(c) $\mathbb{Q}^2 \subseteq \mathbb{R}^2$ ist kein $\mathbb{R}$-Untervektorraum.

(d) Sei M eine Menge, und sei $m \in M$. Die Menge

$$\{f : M \to \mathbb{R} \mid f(m) = 0\}$$

ist ein Untervektorraum von Abb $(M, \mathbb{R})$.

(e) In der Analysis wird gezeigt: die Menge $C([a,b], \mathbb{R})$ der stetigen reellwertigen Funktionen auf dem Intervall $[a,b]$ ist ein Unterraum des $\mathbb{R}$-Vektorraums $\mathrm{Abb}([a,b]), \mathbb{R})$ aller Abbildungen.

Satz 5.7 Sei V ein K-Vektorraum, und seien $W_1, \ldots, W_n$ Unterräume von V. Dann ist

$$\bigcap_{i=1}^{n} W_i$$

wieder ein Unterraum von V.

Beweis

$x, y \in \bigcap\limits_{i=1}^{n} W_i \Rightarrow \quad \underset{i \in \{1,\ldots,n\}}{\forall} \quad x \in W_i \ \wedge \ y \in W_i$

$\Rightarrow \underset{i \in \{1,\ldots,n\}}{\forall} \quad x + y \in W_i \quad$ (da alle W_i Unterräume sind)

$\Rightarrow x + y \in \bigcap\limits_{i=1}^{n} W_i$

$x \in \bigcap\limits_{i=1}^{n} W_i, \lambda \in K \Rightarrow \underset{i \in \{1,\ldots,n\}}{\forall} \quad x \in W_i$

$\Rightarrow \underset{i \in \{1,\ldots,n\}}{\forall} \quad \lambda x \in W_i \quad$ (da alle W_i Unterräume sind)

$\Rightarrow \lambda x \in \bigcup\limits_{i=1}^{n} W_i.$

§6 Gruppenhomomorphismen

Definition 6.1 Seien G und H Gruppen. Eine Abbildung

$$\varphi : G \to H$$

heißt (Gruppen)homomorphismus, wenn sie verträglich mit den Gruppenstrukturen ist, d.h., wenn gilt

$$\varphi(g_1 \cdot g_2) = \varphi(g_1) \cdot \varphi(g_2) \quad \text{für alle } g_1, g_2 \in G\,.$$

φ heißt Monomorphismus (bzw. Epimorphismus bzw. Isomorphismus), wenn φ injektiv (bzw. surjektiv, bzw. bijektiv) ist.
Zwei Gruppen G und H heißen isomorph, wenn es einen Gruppenisomophismus zwischen ihnen gibt.

Bedeutung: Zwei isomorphe Gruppen G und H können für die Zwecke der Gruppentheorie miteinander identifiziert werden; insbesondere haben sie bis auf Permutation der Element die gleiche Gruppentafeln.

Beispiele 6.2 (a) Sei G eine Gruppe. Für jedes $g \in G$ haben wir eine Abbildung

$$\begin{array}{rrcl} L_g\colon & G & \to & G \\ & h & \mapsto & g \cdot h \end{array}$$

(genannt die Linkstranslation mit g). L_g ist bijektiv, denn ist $L_g(h) = L_g(h')$, also $gh = gh'$, so folgt durch Multiplikation mit g^{-1} von links $h = h'$; also ist L_g injektiv. L_g ist auch surjektiv, denn für $h \in G$ ist $g^{-1}h$ Urbild von h unter L_g: $L_g(g^{-1}h) = g(g^{-1}h) = (gg^{-1})h = eh = h$.

Die Menge $S_G = Bij(G,G) = \{f : G \to G \mid f \text{ bijektiv}\}$ ist eine Gruppe unter der Verknüpfung

$$\begin{array}{rcll} S_G \times S_G & \to & S_G & \\ (f, f') & \mapsto & f \circ f' & \text{(Komposition von Abbildungen)} \end{array}$$

(siehe 3.8). Weiter haben wir einen Monomorphismus

$$\begin{array}{rrcl} \varphi\colon & G & \to & S_G \\ & g & \mapsto & L_g \end{array}$$

Denn: φ ist Homomorphismus:
z.z.: Für $g, g' \in G$ ist $\varphi(gg') = \varphi(g) \circ \varphi(g'), d.h., L_{gg'} = L_g \circ L'_g$.

Hierzu ist zu zeigen: Für alle $h \in G$ ist

$$L_{gg}(h) = (L_g \circ L_{g'})(h)\,.$$

Es ist aber

$$L_{gg'}(h) = (gg')h$$

und

$$(L_g \circ L_{g'})(h) = L_g(L_{g'}(h)) = L_g(g'h) = g(g'h)\,.$$

Die Gleichheit dieser Terme gilt nach dem Assoziativgesetz.

φ ist injektiv:
$\varphi(g) = \varphi(g') \Rightarrow L_g = L_{g'} \Rightarrow\ \forall\, h \in G : L_g(h) = L_{g'}(h) \Rightarrow\ \forall\, h \in G : gh = g'h \Rightarrow g = g'$
(setze $h = e$, das neutrale Element von G)

Lemma 6.3 Sei $\varphi : G \to G'$ ein Gruppenhomomorphismus. Die neutralen Elemente von G und G' seien e bzw. e'. Dann gilt

(a) $\varphi(e) = e'$.

(b) Für alle $x \in G$ ist $\varphi(x^{-1}) = \varphi(x)^{-1}$.

Beweis (a): Es ist

$$\varphi(e) = \varphi(ee) \overset{\text{Homom.}}{=} \varphi(e)\varphi(e)$$

Durch Multiplikation mit $\varphi(e)^{-1}$ folgt

$$e' = \varphi(e)\,.$$

(b) Es ist

$$\varphi(x)\varphi(x^{-1}) \overset{\text{Homom.}}{=} \varphi(xx^{-1}) = \varphi(e) \overset{\text{(a)}}{=} e'$$

Durch Multiplikation mit $\varphi(x)^{-1}$ von links folgt

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

Satz 6.4 Sei $\varphi : G \to G'$ ein Gruppenhomomorphismus.

(a) Für jede Untergruppe $U \subseteq G$ ist $\varphi(U)$ eine Untergruppe von G'.

(b) Für jede Untergruppe $U' \subseteq G'$ ist $\varphi^{-1}(U')$ eine Untergruppe von G.

Beweis (a): Seien $x', y' \in \varphi(U) \Rightarrow \underset{x,y\in U}{\exists} \quad \varphi(x) = x' \wedge \varphi(y) = y'$. Dann ist

$$x' \cdot y' = \varphi(x) \cdot \varphi(y) \overset{\varphi\,\text{Homom.}}{=} \varphi(x \cdot y) \in \varphi(U)\,,$$

da $x \cdot y \in U$ (U Untergruppe). Weiter seien e und e' die neutralen Elemente von G bzw. G'. Dann ist $e \in U$, also $e' \overset{6.3(a)}{=} \varphi(e) \in \varphi(U)$. Schließlich ist für $x' = \varphi(x) \in \varphi(U)$ wie oben

$$(x')^{-1} = \varphi(x)^{-1} \overset{6.3(b)}{=} \varphi(x^{-1}) \in \varphi(U)\,,$$

weil $x^{-1} \in U$. Nach 3.14 ist $\varphi(U)$ also Untergruppe.

(b): Seien $x, y \in \varphi^{-1}(U')$ also $\varphi(x), \varphi(y) \in U'$. Dann ist

$$\varphi(x \cdot y) \overset{\varphi\,\text{Homom.}}{=} \varphi(x) \cdot \varphi(y) \in U' \quad (U' \text{ Untergruppe}),$$

also $x \cdot y \in \varphi^{-1}(U')$. Weiter ist $\varphi(e) = e' \in U'$, also $e \in \varphi^{-1}(U')$, und

$$\varphi(x^{-1}) \overset{6.3(b)}{=} \varphi(x)^{-1} \in U' \quad (\text{da } \varphi(x) \in U')\,,$$

also $x^{-1} \in \varphi^{-1}(U')$. Nach 3.14 ist also $\varphi^{-1}(U')$ Untergruppe.

Corollar/Definition 6.5 Sei $\varphi : G \to G'$ ein Gruppenhomomorphismus.

(a) Dann ist

$$\mathrm{im}\,(\varphi) := \varphi(G) \ = \{g' \in G' \mid \underset{g\in G}{\exists}\, \varphi(g) = g'\}$$

eine Untergruppe von G' und heißt das Bild von φ.

(b) Ist e' das neutrale Element von G', so ist

$$\ker(\varphi) := \varphi^{-1}(e') \ = \{g \in G \mid \varphi(g) = e'\}$$

eine Untergruppe von G und heißt der Kern von φ.

Dies sind Spezialfälle von 6.4, da $G \subseteq G$ und $\{e'\} \subseteq G'$ Untergruppen sind.

Satz 6.6 Ein Gruppenhomomorphismus $\varphi : G \to G'$ ist genau dann injektiv (also ein Monomorphismus), wenn $\ker(\varphi)$ trivial ist, d.h., $\ker(\varphi) = \{e\}$ (e das neutrale Element von G).

Beweis Sei e' das neutrale Element von G'. Sei φ injektiv. Es ist nach 6.3 (a) jedenfalls $\varphi(e) = e'$, also $e \in \ker(\varphi)$. Ist nun $x \in \ker(\varphi)$, also auch $\varphi(x) = e'$, so gilt $x = e$ wegen der Injektivität von φ, also folgt $\ker\varphi = \{e\}$.

Sei umgekehrt $\ker(\varphi) = \{e\}$. Seien $x, y \in G$ mit $\varphi(x) = \varphi(y)$. Dann ist

$$\varphi(xy^{-1}) = \varphi(x)\varphi(y)^{-1} = e\,,$$

also $xy^{-1} \in \ker(\varphi)$, also $xy^{-1} = e$, also $x = y$. Daher ist φ injektiv.

Lemma 6.7 Seien $\varphi : G \to G'$ und $\psi : G' \to G''$ Gruppenhomomorphismen. Dann ist die Komposition

$$\psi \circ \varphi : G \to G''$$

ein Gruppenhomomorphismus.

Beweis Für $x, y \in G$ ist

$$\begin{array}{rcll} (\psi\circ\varphi)(xy) & = & \psi(\varphi(xy)) & \text{(Def.)} \\ & = & \psi(\varphi(x)\varphi(y)) & (\varphi \text{ Homom.}) \\ & = & \psi(\varphi(x)\psi(\varphi(y)) & (\psi \text{ Homom.}) \\ & = & (\psi\circ\varphi)(x)\cdot(\psi\circ\varphi)(y) & \end{array}$$

Lemma 6.8 Sei $\varphi : G \to G'$ ein Gruppenhomomorphismus. Ist φ bijektiv (also ein Isomorphismus), so ist die inverse Abbildung φ^{-1} wieder ein Homomorphismus (also ein Isomorphismus).

Beweis: Seien $x', y' \in G'$. Wir haben zu zeigen

Beh. $\varphi^{-1}(x'y') = \varphi^{-1}(x')\varphi^{-1}(y')$.

Beweis Sei $\varphi^{-1}(x') = x$ und $\varphi^{-1}(y') = y$. Dies bedeutet: x und y sind die eindeutig bestimmten Elemente in G mit $\varphi(x) = x'$ und $\varphi(y) = y'$. Es gilt

$$\varphi(xy) = \varphi(x)\varphi(y) = x'y'\,,$$

und xy ist natürlich eindeutig mit dieser Eingenschaft. Daher ist

$$\varphi^{-1}(x'y') = xy = \varphi^{-1}(x')\varphi^{-1}(y') \, .$$

§7 Lineare Abbildungen

Def. 7.1 Sei K ein Körper. Eine Abbildung

$$\varphi : V \to V'$$

zwischen K-Vektorräumen heißt (K)-lineare Abbildung (oder ($K-$)Vektorraum-Homomorphismus), wenn gilt:

(i) Für alle $v_1, v_2 \in V$ gilt

$$\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)$$

(d.h., φ ist Gruppenhomomorphismus $(V, +) \to (V', +)$)

(ii) Für alle $\lambda \in K$ und alle $v \in V$ gilt

$$\varphi(\lambda v) = \lambda \varphi(v) \, .$$

Ist φ injektiv (bzw. surjektiv, bzw. bijektiv), so heißt φ Vektorraum-Monomorphismus (bzw. -Epimorphismus, bzw. -Isomorphismus). Ist $V = V'$, so heißt φ Endomorphismus. Ein Isomorphismus $\varphi : V \to V$ heißt (Vektorraum-)Automorphismus.

Bemerkung 7.2 Da φ insbseondere Gruppenhomomorphismus $(V, +) \to (V', +)$ ist, gilt nach 6.3

$$\begin{aligned} \varphi(0) &= 0 \\ \varphi(-x) &= -\varphi(x), \text{ für alle } x \in V \end{aligned}$$

Beispiele 7.3 Sei K ein Körper.

(a) Seien V und V' K-Vektorräume. Die Nullabbildung

$$\begin{array}{rrcl} 0: & V & \to & V' \\ & v & \mapsto & 0 \end{array}$$

ist eine lineare Abbildung (auch Nullmorphismus genannt).

(b) Die Identität $\mathrm{id}_V : V \to V$ auf einem K-Vektorraum ist ein Automorphismus von V.

(c) Sei $n \in \mathbb{N}$ Für jedes $i \in \{1, \ldots, n\}$ ist die i-te Projektion

$$\begin{array}{rrcl} p_i : & K^n & \to & K \\ & (x_1, \ldots, x_n) & \mapsto & x_i \end{array}$$

eine surjektive lineare Abbildung.

(d) Sei V ein K-Vektorraum. Für jedes $\alpha \in K$ ist dann die Multiplikation mit α

$$\begin{array}{rccc} M_\alpha : & V & \to & V \\ & x & \mapsto & \alpha \cdot x \end{array}$$

ein Endomorphismus von V, denn:

$$\begin{array}{rcl} M_\alpha(x+y) & = & \alpha(x+y) = \alpha x + \alpha y = M_\alpha(x) + M_\alpha(y) \\ M_\alpha(\beta x) & = & \alpha\beta x = \beta\alpha x = \beta M_\alpha(x)\,. \end{array}$$

M_α heißt auch Streckung mit dem Skalar α.

(e) Die Abbildung

$$\begin{array}{rccc} \varphi : & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & 3x+4y \end{array}$$

ist $\mathbb{R}$-linear.

Proposition 7.4 Sei K ein Körper, und seien V, V' und V'' K-Vektorräume.

(a) Sind $\varphi : V \to V'$ und $\psi : V' \to V''$ lineare Abbildungen, so auch $\psi \circ \varphi : V \to V''$.

(b) Ist $\varphi : V \to V'$ ein Vektorraum-Isomorphismus, so auch die inverse Abbildung φ^{-1}.

Beweis: Wir haben dies schon für Gruppenhomomorphismen gesehen (6.7 bzw. 6.8). Wir haben also nur noch die Skalarmultiplikation zu untersuchen:

(a): Für $v \in V$ und $\lambda \in K$ gilt

$$\begin{array}{rcl} (\psi \circ \varphi)(\lambda v) & = & \psi(\varphi(\lambda v)) \\ & \overset{\varphi \text{ lim.}}{=} & \psi(\lambda\varphi(v)) \\ & \overset{\psi \text{ lim.}}{=} & \lambda(\psi(\varphi(v))) = \lambda(\psi \circ \varphi)(v)\,. \end{array}$$

(b) z.z. $\varphi^{-1}(\lambda v') = \lambda\varphi^{-1}(v')$ (für $\lambda \in K$, $v' \in V'$)

$$\begin{array}{rcl} v = \varphi^{-1}(v') & \Leftrightarrow & \varphi(v) = v' \\ & \overset{\varphi \text{ lim.}}{\Rightarrow} & \varphi(\lambda v) = \lambda\varphi(v) = \lambda v' \\ & \Rightarrow & \varphi^{-1}(\lambda v') = \lambda v = \lambda\varphi^{-1}(v') \end{array}$$

Lemma 7.5 Sei $\varphi : V \to V'$ eine lineare Abbildung zwischen K-Vektorräumen.

(a) Für jeden Untervektorraum $W \subseteq V$ ist $\varphi(W) \subseteq V'$ ein Untervektorraum.

(b) Für jeden Untervektorraum $W' \subseteq V'$ ist $\varphi^{-1}(W')$ ein Untervektorraum von V.

(c) Insbesondere ist der Kern von φ,

$$\ker(\varphi) = \{x \in V \mid \varphi(x) = 0\}$$

ein Untervektorraum von V, und das Bild von φ,

$$\mathrm{im}(\varphi) := \{y \in V' \mid \underset{x \in V}{\exists}\ \varphi(x) = y\}$$

ein Untervektorraum von V'.

Beweis Nach dem entsprechenden Sätzen für Gruppen (6.4 und 6.5) müssen wir nur noch die Skalarmultiplikation betrachten.

(a): Sei $v \in \varphi^{-1}(W')$ und $\lambda \in K$

$\Rightarrow \quad \varphi(v) \in W' \overset{W'\,UVR}{\Rightarrow} \lambda\varphi(v) \in W'$

$\Rightarrow \quad \varphi(\lambda v) \overset{\varphi\,\text{linear}}{=} \lambda\varphi(v) \in W'$

$\Rightarrow \quad \lambda v \in \varphi^{-1}(W')$

(b): Sei $v' \in \varphi(W)$ und $\lambda \in K$

$\Rightarrow \quad \exists\, v \in W$ mit $\varphi(v) = v'$

$\Rightarrow \quad \lambda v \in W$ mit $\varphi(\lambda v) \overset{\varphi\,\text{linear}}{=} \lambda\varphi(v) = \lambda v'$ da W Untervektorraum (UVR)

$\Rightarrow \quad \lambda v \in \varphi(W)$

(c) $\ker(\varphi) = \varphi^{-1}(\{0\})$ und $\{0\} \subseteq V'$ ist ein Untervektorraum; $\operatorname{im}(\varphi) = \varphi(V)$ und V ist UVR (Untervektorraum).

Beispiele 7.6 (a) Betrachte die lineare Abbildung

$$\begin{array}{rrcl} \varphi: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & 3x+4y \end{array}$$

(Beispiel 7.3 (e)). Es ist

$$\ker(\varphi) = \{(x,y) \in \mathbb{R}^2 \mid 3x+4y=0\}$$

Dies ist der Untervektorraum $W_1 \subseteq \mathbb{R}^2$ von Beispiel 5.6 (b).

(b) Betrachte das lineare Gleichungssystem (in reellen Zahlen)

$$\begin{array}{rcrcrcl} 3x & + & 4y & + & 5z & = & 0 \\ -x & + & 2y & + & z & = & 0 \end{array}$$

Die Lösungsmenge

$$L = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; \begin{array}{rcrcrcl} 3x & + & 4y & + & 5z & = & 0 \\ -x & + & 2y & + & z & = & 0 \end{array} \right\}$$

ist der Kern der linearen Abbildung

$$\begin{array}{rrcl} \varphi: & \mathbb{R}^3 & \to & \mathbb{R}^2 \\ & (x,y,z) & \mapsto & (3x+4y+5z, -x+2y+z) \end{array}$$

und daher ein Untervektorraum von $\mathbb{R}^3$.

§8 Erzeugendensysteme und Dimension von Vektorräumen

Bild

Wie können wir die "Dimension" eines Vektorraums definieren – z.B. die Dimension des Vektorraums L aus Beispiel 7.6(b)?

Sei K ein Körper (z.B. $\mathbb{R}$ oder $\mathbb{C}$).

Satz/Definition 8.1 Sei V ein K-Vektorraum und $M \subseteq V$ eine Teilmenge. Dann ist

$$\begin{aligned} <M>_K := \{v \in V \mid \exists\, n \in \mathbb{N} \text{ und } v_1, \ldots, v_n \in M \text{ und } \alpha_1, \ldots, \alpha_n \in K \\ \text{mit } v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n\} \end{aligned}$$

der kleinste Unterraum von V, der M enthält und heißt die **lineare Hülle** (oder der Spann) von M. Wir sagen auch, dass $<M>_K$ von M aufgespannt wird (oder dass $<M>_K$ der von M erzeugte Unterraum ist). Setze $<\emptyset>= \{0\}$.

Beweis der Behauptung: 1): $<M>_K$ ist ein Unterraum, denn für $v, v' \in <M>_K$ gibt es $v_1, \ldots, v_n \in M$ und $\alpha_1, \ldots, \alpha_n \in K$ mit $v = \sum\limits_{i=1}^{n} \alpha_i v_i$, sowie $v'_1, \ldots, v'_m \in M$ und $\alpha'_1, \ldots \alpha'_m \in k$ mit $v' = \sum\limits_{j=1}^{n} \alpha'_j v'_j$. Daher ist

$$v + v' = \alpha_1 v_1 + \ldots \alpha_n v_n + \alpha'_1 v'_1 + \ldots + \alpha'_m v'_m$$

Element von $<M>_K$. Weiter ist für $\alpha \in K$

$$\alpha \cdot v = \alpha \cdot (\alpha_1 v_1 + \ldots + \alpha_n v_n) = (\alpha\alpha_1) v_1 + \ldots + (\alpha\alpha_n) v_n$$

Element von $<M>_K$. Nach dem Unterraumkriterium ist also $<M>_K$ ein Unterraum von V.

2): $M \subseteq <M>_K$, denn für $v \in M$ ist $v = 1 \cdot v \in <M>_K$ nach Definition.

3): Ist $W \subseteq V$ ein Unterraum mit $M \subseteq W$, so gilt für $v_1, \ldots, v_n \in M$ und $\alpha_1, \ldots, \alpha_n \in K$ auch $v_i \in W$ für alle $i = 1, \ldots, n$, also $\alpha_i v_i \in W$ für alle i und schließlich $\sum\limits_{i=1}^{n} \alpha_i v_i \in W$, da W Unterraum ist. Also ist $<M>_K \subseteq W$.

Für $M = \{v_1, \ldots, v_n\}$ schreiben wir auch $<v_1, \ldots, v_n>_K$ statt $<\{v_1, \ldots, v_n\}>_K$. Weiter schreiben wir oft $<M>$ für $<M>_K$, wenn der Körper K vorgegeben ist.

Definition 8.2 Seien $v_1, \ldots, v_n$ Vektoren in einem K-Vektorraum V und seien $\alpha_1, \ldots \alpha_n \in K$. Dann heißt

$$\sum_{i=1}^{n} \alpha_i v_i = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

eine **Linearkombination** der Vektoren $v_1, \ldots, v_n$.

Bemerkung 8.3 (a) $<M>_K$ besteht also aus allen Linearkombinationen von Vektoren in M.

(b) Ist $M = \{v_1, \ldots, v_n\}$ endlich, so ist

$$<M>_K = \{\sum_{i=1}^{n} \alpha_i v_i \mid \alpha_i \in K \ \ \forall\, i = 1, \ldots, n\}$$

Ist nämlich $J \subseteq \{1, \dots, n\}$ endlich und sind $\alpha_i \in K$ für $i \in J$, so ist mit

$$\alpha_i' := \begin{cases} \alpha_i & i \in J \\ 0 & i \in \{1, \dots, n\} \setminus J \end{cases}$$

$$\sum_{i \in J} \alpha_i v_i = \sum_{i=1}^n \alpha_i' v_i \,.$$

Definition 8.4 Eine Teilmenge M eines K-Vektorraums V heißt **Erzeugendensystem** von V, wenn $V = <M>_K$. V heißt **endlich erzeugt**, wenn es ein endliches Erzeugendensystem gibt.

Definition 8.5 Für $i = 1, \dots n$ ist der **i-te Einheitsvektor** $e_i \in K^n$ definiert als

$$e_i := (0, \dots, 0, 1, 0, \dots, 0)\,, \text{ wobei 1 an der } i\text{-ten Stelle steht}\,.$$

Es ist also $e_i = (\delta_{ji})_{j=1,\dots,n}$, wobei

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

das Kronecker-Symbol ist.

Beispiele 8.6 (a) $\{e_1, e_2, e_2\}$ ist ein Erzeugendensysten von $\mathbb{R}^3$.

(b) $\{e_1, e_2\}$ ist ein Erzeugendensystem der $x-y$-Ebene $U = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\} \subseteq \mathbb{R}^3$.

Lemma 8.7 Sei $\varphi : V \to V'$ ein lineare Abbildung von K-Vektorräumen, und sei $M \subseteq V$ eine Teilmenge. Dann ist

$$\varphi(<M>_K) = <\varphi(M)>_K \;\;.$$

Insbesonder gilt: Ist M ein Erzeugendensystem von V, so ist $\varphi(M)$ ein Erzeugendensystem von im (φ).

Beweis Die erste Aussage ist klar, da für $v_1, \dots, v_n \in M$ und $\alpha_1, \dots, \alpha_n \in K$ gilt

$$\varphi(\sum_{i=1}^n \alpha_i v_i) = \sum_{i=1}^n \alpha_i \varphi(v_i)\,.$$

Damit folgt auch die zweite Aussage.

Definition 8.8 Sei V ein K-Vektorraum. Dann ist die Dimension von V (Bezeichnung $\dim_K V$ oder einfach $\dim V$) definiert als die minimale Mächtigkeit eines Erzeugendensystems von V. Es ist also $\dim V \in [0, \infty]$, mit

$$\begin{array}{lcl} \dim V < \infty & \Leftrightarrow & V \text{ endlich erzeugt} \\ \dim V = 0 & \Leftrightarrow & V = 0\,. \end{array}$$

Beispiel 8.9 Es ist $\{e_1, \dots, e_n\}$ ein Erzeugendensystem von K^n (denn für $v = (x_1, \dots, x_n) \in K^n$ ist $v = \sum_{i=1}^n x_i e_i$). Also gilt $\dim K^n \leq n$. Gilt Gleichheit?

Lemma 8.10 Ist $\varphi : V \to V'$ eine surjektive lineare Abbildung, so ist $\dim V' \leq \dim V$.

Beweis Dies folgt sofort aus 8.7, da für die Mächtigkeiten $|M|$ und $|\varphi(M)|$ von M bzw. $\varphi(M)$ gilt: $|\varphi(M)| \leq |M|$.

Beobachtung 8.11 Sei V ein K-Vektorraum und $M \subseteq V$ ein Erzeugendensystem. Gibt es ein $v \in M$, so dass $M \setminus \{v\}$ immer noch ein Erzeugendensystem ist, so ist also $v \in <M \setminus \{v\}>$, d.h., es gibt $\{v_1, \ldots, v_n\} \subseteq M \setminus \{v\}$ und $\alpha_1, \ldots, \alpha_n \in K$ mit

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n \,.$$

Wir können dies auch schreiben als

$$v - \alpha_1 v_1 - \alpha_2 v_1 - \ldots - \alpha_n v_n = 0 \,.$$

Dies führt auf:

Definition 8.12 Eine Teilmenge $M \subseteq V$ heißt **linear abhängig**, wenn es Vektoren $v_1, \ldots, v_n \in M$ gibt, alle paarweise verschieden (d.h., $v_i \neq v_j$ für $i \neq j$), und $\alpha_1, \ldots, \alpha_n \in K$, nicht alle gleich 0, mit $\sum\limits_{i=1}^{n} \alpha_i v_i = 0$. Andernfalls heißt M **linear unabhängig** (oder frei).

! M ist also genau dann linear unabhängig, wenn gilt:
Sind $v_1, \ldots, v_n \in M$ paarweise verschiedene Vektoren und $\alpha_1, \ldots, \alpha_n \in K$ mit $\sum\limits_{i=1}^{n} \alpha_i v_i = 0$ so gilt $\alpha_i = 0 \;\; \forall\, i = 1, \ldots, n$.

Bemerkung 8.13 (a) $\emptyset \subseteq V$ ist immer linear unabhängig, $\{0\}$ ist linear abhängig (da 1.0=0).

(b) Endlich viele **Vektoren** $v_1, \ldots, v_n \in V$ heißen **linear unabhängig** (bzw. das Tupel $(v_1, \ldots, v_n)$ heißt linear unabhängig), wenn gilt

$(*)$ $\quad \sum\limits_{i=1}^{n} \alpha_i v_i = 0$ mit $\alpha_i \in K \Rightarrow \alpha_i = 0 \;\; \forall\, i = 1, \ldots, n\,.$

(c) Vektoren $v_1, \ldots, v_n \in V$ sind also genau dann linear unabhängig, wenn sie paarweise verschieden sind, und wenn die Menge $M = \{v_1, \ldots, v_n\}$ linear unabhängig ist. Sind nämlich $v_1, \ldots, v_n$ linear unabhängig, so ist für $i \neq j$ auch $v_i \neq v_j$, denn für $v_i = v_j$ wäre $v_i + (-1)v_j = 0$ eine nichttriviale Linearkombination zu null. Ist weiter $I \subseteq \{1, \ldots, n\}$ eine Teilmenge mit

$$\sum_{i \in I} \alpha_i v_i = 0 \quad \text{für } \alpha_i \in K\,,$$

so ist auch

$$\sum_{i=1}^{n} \alpha_i' v_i = 0 \quad \text{mit } \alpha_i' = \begin{cases} \alpha_i & i \in I \\ 0 & i \notin I\,. \end{cases}$$

Gilt nun $(*)$, so folgt $\alpha_i' = 0 \;\; \forall\, i = 1, \ldots, n$, also auch $\alpha_i = 0 \;\; \forall\, i \in I$. Damit ist M linear unabhängig. Die Rückrichtung ist klar.

Bemerkung 8.14 (a) In K^n sind die Einheitsvektoren $e_1, \ldots, e_n$ linear unabhängig:

$$0 = \sum_{i=1}^{n} \alpha_i e_i = (\alpha_1, \ldots, \alpha_n) \quad \overset{\text{Def.}}{\Rightarrow} \quad \alpha_1 = \ldots = \alpha_n = 0\,.$$

(b) In $\mathbb{R}^2$ sind e_1, e_2 und $v = (1,1)$ linear abhängig, denn es gilt

$$e_1 + e_2 - v = 0\,.$$

(c) In $\mathbb{R}^3$ betrachte die drei Vektoren

$$v_1 = (1,2,3), \quad v_2 = (4,5,6), \quad v_3 = (7,8,9)\,.$$

Sind diese linear unabhängig? Wir beobachten: $v_2 - v_1 = (3,3,3)$ und $v_3 - v_2 = (3,3,3)$. Es folgt $v_2 - v_1 = v_3 - v_2$, d.h.,

$$v_1 - 2v_2 + v_3 = 0\,.$$

Also sind die Vektoren linear abhängig!

Lemma 8.15 Sei V ein K-Vektorraum und $M \subseteq V$ eine linear unabhängige Teilmenge. Ist $v \in V$ mit $v \notin <M>_K$, so ist $M \mathbin{\dot\cup} \{v\}$ ebenfalls linear unabhängig.

Zur Bezeichnung: $A \mathbin{\dot\cup} B$ bezeichnet die **disjunkte Vereinigung**, also die Vereinigung zweier **disjunkter Mengen** ($A \cap B = \emptyset$). Beachte dass oben $v \notin M$.

Beweis Seien $v_1, \ldots, v_n$ paarweise verschiedene Vektoren aus M, und seien $\alpha, \alpha_1, \ldots, \alpha_n \in K$ mit

$$\alpha v + \alpha_1 v_1 + \ldots \alpha_n v_n = 0\,.$$

Ist $\alpha \neq 0$, so folgt

$$v = -\sum_{i=1}^{n} \tfrac{\alpha_i}{\alpha} v_i \in <M>_K \ .$$

Das ist ein Widerspruch zur Voraussetzung. Also ist $\alpha = 0$. Es folgt $\sum_{i=1}^{n} \alpha_i v_i = 0$, und damit auch $\alpha_1 = \ldots = \alpha_n = 0$, wegen der linearen Unabhängigkeit von M.

Satz/Definition 8.16 Sei V ein K-Vektorraum. Eine Teilmenge $B \subseteq V$ heißt **Basis** von V, wenn die folgenden äquivalenten Bedingungen gelten:

(a) B ist Erzeugendensystem und linear unabhängig.

(b) B ist minimales Erzeugendensytem.

(c) B ist maximale linear unabhängige Teilmenge.

Zur Erklärung: "minimal" in (b) heißt, dass keine echte Teilmenge $B' \subsetneq B$ Erzeugendensystem ist, und "maximal" in (c) heißt, dass keine echte größere Menge $B' \supsetneq B$ linear unabhängig ist.

Beweis der Äquivalenz:
(a) $\Rightarrow$ (b): Es gelte (a). Dann ist B ein Erzeugendensystem. Angenommen, B ist nicht minimal. Aus Beobachtung 8.11 folgt dann, dass B linear abhängig ist – Widerspruch!

(b) $\Rightarrow$ (c): Es gelte (b). Angenommen B ist linear abhängig. Dann gibt es paarweise verschiedene $v_1, \ldots, v_n \in B$, und $\alpha_1, \ldots, \alpha_n \in K$, nicht alle null, mit

$$\sum_{i=1}^{n} \alpha_i v_i = 0 \, .$$

Sei etwa $\alpha_j \neq 0$ für $j \in \{1, \ldots, n\}$. Dann ist also

$$v_j = -\sum_{\substack{i=1 \\ i \neq j}}^{n} \tfrac{\alpha_i}{\alpha_j} \cdot v_i \in <v_i \mid i \neq j> \, ,$$

und damit auch $B \smallsetminus \{v_j\}$ ein Erzeugendensystem (Beweis: selbst!), im Widerspruch zu (b). Also ist B linear unabhängig.

Angenommen, B ist nicht maximal linear unabhängig. Dann gibt es ein $v \in V$, mit $v \notin B$, so dass $B \cup \{v\}$ linear unabhängig ist. Dass ist ein Widerspruch dazu, dass B Erzeugendensystem ist (so dass $v \in <B>_K$).

(c) $\Rightarrow$ (a): Es gelte (c). Angenommen, B ist kein Erzeugendensystem. Dann gibt es ein $v \in V$ mit $v \notin <B>_K$, und nach Lemma 8.15 ist $B \cup \{v\}$ linear unabhängig – Widerspruch zur Maximalität von B!

Beispiele 8.17 (a) $\{e_1, \ldots, e_n\}$ ist eine Basis von K^n (da Erzeugendensystem nach 8.9 und linear unabhängig nach 8.14 (a)); diese heißt die **kanonische Basis** von K^n.

(b) $\{(1,1), (-1,1)\}$ ist eine Basis von $\mathbb{R}^2$.

(c) $\{(1,1)\}$, $\{(1,0), (1,1), (0,1)\}$ sind keine Basen von $\mathbb{R}^2$.

Wir beweisen nun zwei wichtige Resultate über Basen. Sei V ein K-Vektorraum.

Lemma 8.18 (Austauschlemma) Sei B eine Basis von V und E ein Erzeugendensystem von V. Sei $v \in B$ und $w \in E$ mit $w \notin <B \smallsetminus \{v\}>_K$. Dann ist $B \smallsetminus \{v\} \cup \{w\}$ wieder eine Basis von V.

Beweis Nach Lemma 8.15 ist $B \smallsetminus \{v\} \cup \{w\}$ linear unahängig. Aber $B \smallsetminus \{v\} \cup \{w\}$ ist auch ein Erzeugendensytem: Da B ein Erzeugendensystem ist, gibt es $v_1, \ldots, v_n \in B$ und $\alpha_1, \ldots, \alpha_n \in K$ mit $w = \sum\limits_{i=1}^{n} \alpha_i v_i$. Da $w \notin <B \smallsetminus \{v\}>_K$ gibt es ein $i_0 \in \{1, \ldots, n\}$ mit $v_{i_0} = v$ und $\alpha_{i_0} \neq 0$. Dann ist

$$v = v_{i_0} = \frac{1}{\alpha_{i_0}} w - \sum_{\substack{i=1 \\ i \neq i_0}}^{n} \tfrac{\alpha_i}{\alpha_{i_0}} v_i \, .$$

Ist nun $u \in V$, so gibt es paarweise verschiedene $w_1, \ldots, w_m \in B$ und $\beta_1, \ldots, \beta_m \in K$ mit $u = \sum\limits_{j=1}^{m} \beta_j w_j$.

Sind alle $w_j \neq v$, so ist offenbar $u \in <B \setminus \{v\}>_K$. Gibt es ein $j_0 \in \{1, \ldots, m\}$ mit $w_{j_0} = v$, so ist

$$u = \sum_{j \neq j_0} \beta_j w_j + \beta_{j_0} \left(\frac{1}{\alpha_{j_0}} w - \sum_{i \neq i_0} \frac{\alpha_i}{\alpha_{i_0}} v_i \right) \in <B \setminus \{v\} \cup \{w\}>_K$$

Satz 8.19 (Basisaustauschsatz von Steinitz) Sei V ein endlich-dimensionaler K-Vektorraum, sei $B = \{v_1, \ldots, v_n\}$ eine n-elementige Basis, und sei E ein Erzeugendensystem von V. Sei $m \in \mathbb{N}_0$ mit $m \leq n$ gegeben. Dann existieren m Elemente $w_1, \ldots, w_m \in E$, so dass

$$\{w_1, \ldots, w_m, \, v_{m+1}, \ldots, v_n\}$$

wieder eine (n-elementige) Basis von V ist.

(man kann also die ersten m Elemente durch Elemente aus E austauschen und erhält wieder eine Basis. Für $m = n$ erhalten wir $\{w_1, \ldots, w_n\}$).

Beweis durch vollständige Induktion über m ($\leq n$). Für $m = 0$ ist nichts zu zeigen.
Sei die Aussage für m bewiesen, so dass wir eine n-elementige Basis

$$B' = \{w_1, \ldots, w_m, \, v_{m+1}, \ldots, v_n\}$$

mit $w_1, \ldots, w_m \in E$ erhalten. Falls $m = n$, so ist nichts mehr zu zeigen. Sei $m < n$. Da B' eine Basis ist, also ein minimales Erzeugendensystem, ist $B' \setminus \{v_{m+1}\}$ kein Erzeugendensystem. Es gibt also ein $w \in E$ mit $w \notin < B' \setminus \{v_{m+1}\} >$ (denn sonst gilt $E \subseteq <B \setminus \{v_{m+1}\}>_K$ und damit $V = <E>_K \subseteq <B' \setminus \{v_{m+1}\}>$ – Widerspruch!). Nach dem Austauschlemma ist $B' \setminus \{v_{m+1}\} \cup \{w\} = \{w_1, \ldots, w_m, w, v_{m+2}, \ldots, v_n\}$ eine Basis (n-elementig da $w \notin B' \setminus \{v_{m+1}\}$), und wir können $w_{m+1} := w$ setzen.

Corollar 8.20 Sei V ein endlich erzeugter K-Vektorraum.

(a) Dann besitzt V eine endliche Basis, und für jede Basis B gilt $|B| = \dim V$. Insbesondere haben alle Basen die gleiche Mächtigkeit.

(b) Für jedes Erzeugendensystem E gilt $|E| \geq \dim V$.

(c) Für jede linear unabhängige Teilmenge B gilt $|F| \leq \dim V$.

Beweis Nach Definition gibt es ein endliches Erzeugendensystem E_0 mit $|E_0| = \dim V$. Dieses ist dann minimal, also eine Basis. Aus dem Steinitzschen Austauschsatz (mit $B = E_0$ und $m = n$ folgt:

Lemma 8.21 Sei V ein endlich-dimensionaler K-Vektorraum. Ist $E \subseteq V$ ein Erzeugendensystem, so gibt es eine endliche Teilmenge $E' \subseteq E$, die eine Basis von V ist.

Dies zeigt, dass jede Basis B von V endlich ist, denn nach 8.21 enthält B eine endliche Basis B' und wegen der Minimalität von B als Erzeugendensystem ist $B' = B$.
Aus dem Steinitzschen Austauschsatz folgt weiter: Ist B eine Basis und E ein Erzeugendensystem so gilt

$$(*) \qquad |B| \leq |E| \, .$$

Angewandt auf E_0 erhalten wir $|B| \leq \dim V$. Andererseits gilt, da B Erzeugendensystem ist, $|B| \geq \dim V$, und wir erhalten (a). Dann folgt (b) mit $(*)$.

(c) folgt aus (a) und:

Satz 8.22 (Basisergänzungssatz von Steinitz) Sei V ein endlich erzeugter K-Vektorraum. Ist $F \subseteq V$ eine linear unabhängige Menge und ist $E \subseteq V$ ein Erzeugendensystem von V, so gibt es endlich viele Vektoren $w_1, \ldots, w_m \in E$, so dass $F \cup \{w_1, \ldots, w_m\}$ eine Basis von V ist.

Beweis Zunächst ist E wegen 8.21 ohne Einschränkung endlich. Wir wählen die w_i nun induktiv. Ist F schon Erzeugendensystem, so sind wir fertig. Seien bereit $w_1, \ldots, w_n \in E$ gewählt, so dass $F \cup \{w_1, \ldots, w_n\}$ linear unabhängig ist. Wenn diese Menge auch erzeugend ist, so sind wir fertig. Andernfalls gibt es ein $w_{n+1} \in E$ mit $w_{n+1} \notin <F \cup \{w_1, \ldots, w_n\}>_K$. Nach Lemma 8.15 ist dann $F \cup \{w_1, \ldots, w_{n+1}\}$, linear unabhängig. Dieser Prozess muss bei $n = |E|$ abbrechen, da E erzeugend ist.

Fazit: Nach 8.22 kann man jede linear unabhängige Menge zu einer Basis vergrößern, und nach 8.21 kann man jedes Erzeugendensystem zu einer Basis verkleinern.

Corollar 8.23 Sei $\dim V = n < \infty$.

(a) Mehr als n Vektoren sind linear abhängig.

(b) Weniger als n Vektoren sind nicht erzeugend.

(c) Für n Vektoren $v_1, \ldots v_n$ ist äquivalent:

   (i) Die Vektoren bilden eine Basis.

   (ii) Die Vektoren sind linear unabhängig.

   (iii) Die Vektoren bilden ein Erzeugendensystem.

Beweis (a) Die maximale Mächtigkeit einer linear unabhängigen Menge ist nach 8.20 (c) gleich n.

(b) Die minimale Mächtigkeit eines Erzeugendensystems ist nach 8.20 (b) gleich n.

(c) (i) $\Rightarrow$ (ii), (iii) nach 8.16(a). Gilt umgekehrt (ii), so ist $\{v_1, \ldots, v_n\}$ nach (a) maximal linear unabhängig, also eine Basis nach 8.16 (c). Gilt (iii), so ist $\{v_1, \ldots, v_n\}$ nach (b) minimales Erzeugendensystem, also eine Basis nach 8.16 (b).

Definition 8.24 Sei V ein K-Vektorraum. Ein n-Tupel $(v_1, \ldots, v_n)$ von Vektoren $v_i \in V$ heißt eine **geordnete Basis** von V, wenn es linear unabhängig ist (8.13(b)) und erzeugend ($<v_1, \ldots, v_n>_K = V$).

Also ist $(v_1, \ldots, v_n)$ genau dann eine geordnete Basis, wenn die v_i paarweise verschieden sind und $\{v_1, \ldots, v_n\}$ eine Basis ist.

Wir haben dann die folgende Kennzeichnung:

Satz 8.25 $(v_1, \ldots, v_n)$ ist genau dann eine (geordnete) Basis von V, wenn jeder Vektor $v \in V$ eine Darstellung

$$(*) \qquad v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

mit eindeutig bestimmten $\alpha_1, \ldots, \alpha_n \in K$ hat.

Beweis Offenbar bilden $v_1, \ldots, v_n$ genau dann ein Erzeugendensytem, wenn es für jeden Vektor $v \in V$ eine solche Darstellung (mit vielleicht nicht eindeutigen α_i) gibt. Ist diese Darstellung eindeutig für $v = 0$, so ist offenbar $(v_1, \ldots, v_n)$ linear unabhängig. Ist umgekehrt $(v_1, \ldots, v_n)$ linear unabhängig, und gilt für einen Vektor $v \in V$

$$v = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i v_i \,,$$

so folgt $\sum\limits_{i=1}^{n} (\alpha_i - \beta_i) v_i = 0$, wegen der linearen Unabhängigkeit also $\alpha_i - \beta_i = 0$ für alle $i = 1, \ldots, n$, d.h., $\alpha_i = \beta_i$ für $i = 1, \ldots, n$. Also gilt die Eindeutigkeit der Darstellung $(*)$.

Lemma 8.26 Sei V ein K-Vektorraum und seien $v_1, \ldots, v_n \in V$ Vektoren. Dann ist die Abbildung

$$\begin{array}{rccl} \varphi : & K^n & \to & V \\ & (\alpha_1, \ldots, \alpha_n) & \mapsto & \sum\limits_{i=1}^{n} \alpha_i v_i \end{array}$$

linear. Weiter gilt

(a) $\mathrm{im}\,(\varphi) = <v_1, \ldots, v_n>_K$.

(b) φ surjektiv $\Leftrightarrow$ $(v_1, \ldots, v_n)$ Erzeugendensytem.

(c) φ injektiv $\Leftrightarrow$ $(v_1, \ldots, v_n)$ linear unabhängig.

(d) φ Isomorphismus $\Leftrightarrow$ $(v_1, \ldots, v_n)$ Basis.

Beweis Die Linearität von φ rechnet man sofort nach:
$\varphi((\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n)) = \varphi((\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n))$
$= (\alpha_1 + \beta_1) v_1 + \ldots + (\alpha_n + \beta_n) v_n$
$= \alpha_1 v_1 + \ldots + \alpha_n v_n + \beta_1 v_1 + \ldots + \beta_n v_n$
$= \varphi((\alpha_1, \ldots, \alpha_n)) + \varphi((\beta_1, \ldots, \beta_n))$. Weiter gilt für $\lambda \in K$
$\varphi(\lambda(\alpha_1, \ldots, \alpha_n)) = \varphi((\lambda\alpha_1, \ldots, \lambda\alpha_n))$
$= \lambda\alpha_1 v_1 + \ldots + \lambda\alpha_n v_n = \lambda(\alpha_1 v_1 + \ldots + \alpha_n v_n)$
$= \lambda\varphi((\alpha_1, \ldots, \alpha_n))$.

(a) folgt direkt aus der Definition der linearen Hülle, und (b) folgt direkt aus (a).

(c): φ injektiv $\Leftrightarrow$ $\ker(\varphi) = 0$. Letzteres ist aber offenbar äquivalent dazu, dass $(v_1, \ldots, v_n)$ linear unabhängig ist.

(d) folgt aus (b) und (c).

Definition 8.27 Die Abbildung φ aus 8.26 nennen wir auch $\varphi_{(v_1, \ldots, v_n)}$.

Bemerkung 8.28 (a) Sei $\psi : K^n \to V$ eine beliebige lineare Abbildung. Dann ist $\psi = \varphi_{(v_1,\dots,v_n)}$ für eindeutig bestimmtes Tupel $(v_1, \dots, v_n)$ von Vektoren in V. Setze nämlich $v_i := \psi(e_i)$, wobei e_i der i-te Einheitsvektor in K^n ist. Dann gilt $\psi = \varphi_{(v_1,\dots,v_n)}$:

$$\begin{aligned}
\psi((\alpha_1, \dots, \alpha_n)) &= \psi\left(\sum_{i=1}^{n} \alpha_1 e_i\right) \\
&= \sum_{i=1}^{n} \alpha_i \psi(e_i) \quad \text{(da } \psi \text{ linear)} \\
&= \sum_{i=1}^{n} \alpha_i v_i \\
&= \varphi_{(v_1,\dots,v_n)}((\alpha_1, \dots, \alpha_n)) \quad \text{(Definition).}
\end{aligned}$$

Natürlich gilt auch $\varphi_{(v_1,\dots,v_n)}(e_i) = v_i$; also ist $(v_1, \dots, v_n)$ eindeutig.

(b) Insbesondere folgt mit 8.26(d):
Es gibt eine Bijektion

$$\begin{array}{ccc}
\left\{ \begin{array}{c} \text{geordnete Basen} \\ (v_1, \dots, v_n) \text{ von } V \end{array} \right\} & \to & \left\{ \begin{array}{c} \text{Isomorphismen} \\ \varphi : K^n \xrightarrow{\sim} V \end{array} \right\} \\
(v_1, \dots, v_n) & \mapsto & \varphi_{(v_1,\dots,v_n)} \,.
\end{array}$$

Corollar 8.29 Jeder n-dimensionale K-Vektorraum V ist isomorph zu K^n (d.h., es gibt einen Isomorphismus $\varphi : K^n \xrightarrow{\sim} V$). .

Beweis: Es gibt eine Basis $(v_1, \dots, v_n)$ von V.

Beispiel 8.30 Sei

$$U = \left\{ (x, y, z) \in \mathbb{R}^3 \,\middle|\, \begin{array}{r} y + 3z = 0 \\ 4x - y + 5z = 0 \end{array} \right\}$$

Dies ist der Kern der linearen Abbildung

$$\begin{array}{cccc}
\varphi : & \mathbb{R}^3 & \to & \mathbb{R}^2 \\
& \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto & \begin{pmatrix} y + 3z \\ 4x - y + 5z \end{pmatrix}
\end{array}$$

Für welche v_1, v_2, v_3 ist $\varphi = \varphi_{(v_1,v_2,v_3)}$? Antwort: Für $v_1 = \varphi(e_1) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$, $v_2 = \varphi(e_2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $v_3 = \varphi(e_3) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$, nach der Konstruktion in 8.28. Tatsächlich ist

$$\begin{aligned}
\varphi\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) &= x \cdot \begin{pmatrix} 0 \\ 4 \end{pmatrix} + y \begin{pmatrix} 1 \\ -1 \end{pmatrix} + z \begin{pmatrix} 3 \\ 5 \end{pmatrix} \\
&= \varphi_{(v_1,v_2,v_3)}\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) \,.
\end{aligned}$$

(Hieraus folgt noch einmal, dass φ linear ist). Um eine Basis von U zu bestimmen, lösen wir das lineare Gleichungsystem

$$\begin{aligned} y + 3z &= 0 \\ 4x - y + 5z &= 0\,. \end{aligned}$$

Wir erhalten $y = -3z$ und $4x + 8z = 0$, also $x = -2z$. Wir können $z \in \mathbb{R}$ beliebig wählen und erhalten eine Lösung durch $y = -3z$, $x = -2z$. (Überprüfung durch Einsetzen in das Gleichungssystem!). Also ist

$$U = \left\{ z \cdot \begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix} \middle| \, z \in \mathbb{R} \right\},$$

d.h., U ist 1-dimensional mit Basis

$$w = \begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix}.$$

Wir erhalten einen Isomorphismus von Vektorräumen

$$\begin{aligned} \psi = \varphi_w : \quad \mathbb{R} &\xrightarrow{\sim} U \\ \lambda &\mapsto \lambda \cdot \begin{pmatrix} -2 \\ -3 \\ 1 \end{pmatrix}. \end{aligned}$$

Wir können U auch als Durchschnitt der beiden Vektorräume

$$U_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \middle| \, y + 3z = 0 \right\}$$

und

$$U_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \middle| \, 4x - y + 5z = 0 \right\}$$

beschreiben. Für U_1 ist zum Beispiel

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix}$$

eine Basis (warum?), und für U_2 ist

$$\begin{pmatrix} 0 \\ 5 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}$$

eine Basis (warum?). Also hat man einen Isomorphismus

$$\begin{aligned} \varphi : \quad \mathbb{R}^2 &\xrightarrow{\sim} U_1 \\ \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} &\mapsto \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix} \end{aligned}$$

("Parametrisierung" der Lösung des linearen Gleichungssystems

$$y + 3z = 0\,,$$

welches U_1 beschreibt). Entsprechend für U_2.

Wir schließen diesen Abschnitt mit der folgenden wichtigen Eigenschaft von Basen.

Satz 8.31 (universelle Eigenschaft einer Basis) Sei V ein endlich-dimensionaler K-Vektorraum, und sei $(v_1, \dots, v_n)$ eine Basis von V. Sei W ein weiterer K-Vektorraum. Zu jedem n-Tupel $(w_1, \dots, w_n)$ von Vektoren $w_i \in W$ gibt es dann eine eindeutig bestimmte lineare Abbildung

$$\psi = \varphi^{(v_1,\dots,v_n)}_{(w_1,\dots,w_n)} : V \to W$$

mit $\psi(v_i) = w_i$.

(Wir können also die Bilder auf einer Basis beliebig vorgeben. Dies verallgemeinert 8.26/27, welches der Spezialfall $V = K^n$ und $(v_1, \dots, v_n) = (e_1, \dots, e_n)$ von 8.31 ist. Tatsächlich ist $\varphi^{(e_1,\dots,e_n)}_{(w_1,\dots,w_n)}$ nach 8.31 dann $\varphi_{(w_1,\dots,w_n)}$ nach 8.27).

Beweis Eindeutigkeit: Gilt $\psi(v_i) = w_i$ für alle $i = 1, \dots, n$, so ist ψ hierdurch bestimmt, da $(v_1, \dots, v_n)$ ein Erzeugendensystem ist: Für $v \in V$ gilt $v = \sum\limits_{i=1}^{n} \alpha_i v_i$ mit $\alpha_1, \dots, \alpha_n \in K$, und es muss gelten

$$\psi(v) = \psi\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i \psi(v_i) = \sum_{i=1}^{n} \alpha_i w_i\,.$$

Existenz: Definiere ψ durch diese Formel! Wegen 8.25 ist die Darstellung

$$v = \sum_{i=1}^{n} \alpha_i v_i \quad (\alpha_i \in K)$$

eindeutig, d.h., die α_i sind eindeutig, und wir setzen mit diesen eindeutigen α_i

$$\psi(v) = \sum_{i=1}^{n} \alpha_i w_i\,.$$

Wir müssen noch zeigen, dass ψ linear ist. Dies kann man direkt aus der Definition nachrechnen; ein anderer Beweis ergibt sich so: Nach Definition ist das Diagramm

$$\begin{tikzcd}
V \arrow[rr, "\psi"] & & W \\
& K^n \arrow[ul, "\sim", "\varphi_{\underline{v}}=\varphi_{(v_1,\dots,v_n)}"'] \arrow[ur, "\sim"', "\varphi_{(w_1,\dots,w_n)}=\varphi_{\underline{w}}"] &
\end{tikzcd}$$

kommutativ, d.h., es ist

$$\psi \circ \varphi_{\underline{v}} = \varphi_{\underline{w}}\,.$$

Damit folgt, dass $\psi = \varphi_{\underline{w}} \circ \varphi_{\underline{v}}^{-1}$ linear ist.

§9 Dimensionsformeln

Sei K ein Körper.

Satz 9.1 Sei V ein endlich-dimensionaler K-Vektorraum. Dann ist jeder Untervektorraum $U \subseteq V$ auch endlich-dimensional, und es gilt

$$\dim U \leq \dim V\,.$$

Weiter ist $\dim U = \dim V$ genau dann wenn $U = V$.

Beweis Sei $F \subseteq U$ eine linear unabhängige Menge. Nach 8.20 gilt

$$|F| \leq \dim V\,.$$

Insbesondere ist F endlich. Also gibt es eine endliche (maximale lineare unabhängige Menge, und damit) Basis B' von U. Damit ist $\dim U = |B'| \leq \dim V$. Nach dem Basisergänzungssatz 8.22, angewandt auf B' und das Erzeugendensystem $E = V$, gibt es eine Basis $B \supseteq B'$ von V. Gilt nun $\dim U = \dim V$, so ist $|B'| = |B|$, also $B' = B$, also $U = V$. Die Umkehrung ist klar.

Beispiel 9.2 Wir können jetzt leichter sehen, dass der Untervektorraum

$$U_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, y + 3z = 0 \right\} \subseteq \mathbb{R}^3$$

aus Beispiel 8.30 die Basis

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -3 \\ 1 \end{pmatrix}$$

wie behauptet hat: Die Vektoren liegen in U_1 und sind linear unabhängig, also ist $\dim U_1 \geq 2$. Es ist $U_1 \neq \mathbb{R}^3$ (z.B. $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \notin U_1$), also ist nach 9.1 $\dim U_1 \neq 3$. Es folgt $\dim U_1 = 2$; damit bilden die 2 linear unabhängigen Vektoren eine Basis (siehe 8.23 (c)).

Satz/Definition 9.3 (Rangsatz) Sei $\varphi : V \to V'$ eine K-lineare Abbildung, wobei V endlich-dimensional ist. Dann ist $\mathrm{im}(\varphi)$ endlich-dimensional, und es ist

$$\dim \ker(\varphi) + \mathrm{rang}(\varphi) = \dim V\,,$$

wobei $\mathrm{rang}(\varphi) := \dim \mathrm{im}(\varphi)$ der **Rang von** φ ist.

Beweis Da V endlich erzeugt ist, gilt dies auch für $\mathrm{im}(\varphi)$ nach Lemma 8.7. Weiter ist $\ker(\varphi)$ endlich-dimensional nach 9.1. Sei $(v_1, \ldots, v_m)$ eine Basis von $\ker(\varphi)$, und sei $(w_1, \ldots, w_n)$ eine Basis von $\mathrm{im}(\varphi)$. Seien $\tilde{w}_1, \ldots, \tilde{w}_n$ Urbilder in V von $w_1, \ldots, w_n$, also $\varphi(\tilde{w}_i) = w_i$.

Behauptung: $(v_1, \ldots, v_m, \tilde{w}_1, \ldots, \tilde{w}_n)$ ist Basis von V. (Hieraus folgt dann $\dim V = m+n$, also die Behauptung des Satzes, da $m = \dim\ker(\varphi)$ und $n = \dim\,\mathrm{im}(\varphi)$).

Beweis: 1) Erzeugendensystem: Sei $v \in V$. Dann gibt es $\beta_1, \ldots, \beta_n \in K$ mit $\varphi(v) = \sum\limits_{j=1}^{n} \beta_j w_j$. Setze

$$\tilde{v} := \sum_{j=1}^{n} \beta_j \tilde{w}_j \,.$$

Dann gilt $\varphi(\tilde{v}) = \varphi(\sum\limits_{j=1}^{n} \beta_j \tilde{w}_j) = \sum\limits_{j=1}^{n} \beta_j \varphi(\tilde{w}_j) = \sum\limits_{j=1}^{n} \beta_j w_j = \varphi(v)$. Hieraus folgt $\varphi(v - \tilde{v}) = \varphi(v) - \varphi(\tilde{v}) = 0$, also $v - \tilde{v} \in \ker(\varphi)$. Also gibt es $\alpha_1, \ldots, \alpha_m \in K$ mit

$$v - \tilde{v} = \sum_{i=1}^{m} \alpha_i v_i \,.$$

Zusammen folgt

$$v = \sum_{i=1}^{m} \alpha_i v_i + \tilde{v} = \sum_{i=1}^{m} \alpha_i v_i + \sum_{j=1}^{n} \beta_j \tilde{w}_j \,,$$

also $v \in <v_1, \ldots, v_m, \tilde{w}_1, \ldots, \tilde{w}_n>_K$.

2) linear unabhängig:
Seien $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n \in K$ mit

$$\sum_{i=1}^{m} \alpha_i v_i + \sum_{j=1}^{n} \beta_j \tilde{w}_j = 0 \,.$$

Dann folgt

$$\begin{aligned} 0 = \varphi(0) &= \sum_{i=1}^{m} \alpha_i \varphi(v_i) + \sum_{j=1}^{n} \beta_j \varphi(\tilde{w}_j) \\ &= \sum_{j=1}^{n} \beta_j w_j \,, \end{aligned}$$

da $v_i \in \ker(\varphi)$ für alle $i = 1, \ldots, m$. Da $(w_1, \ldots, w_n)$ linear unabhängig ist, folgt $\beta_1 = \ldots = \beta_n = 0$. Also haben wir

$$\sum_{i=1}^{m} \alpha_i v_i = 0 \,,$$

woraus wegen der linearen Unabhängigkeit der v_i wiederum $\alpha_1 = \ldots = \alpha_m = 0$ folgt.

Beispiel 9.4 Wir können nun noch leichter sehen (ohne Angabe einer Basis!), dass für den Untervektorraum

$$U_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, y + 3z = 0 \right\}$$

(Beispiele 8.30, 9.2) $\dim U_1 = 2$ gilt: Es ist $U_1 = \ker(\varphi)$ für die lineare Abbildung

$$\begin{aligned} \varphi : \quad \mathbb{R}^3 &\to \mathbb{R} \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\mapsto y + 3z \end{aligned}$$

($\varphi = \varphi_{(0,1,3)}$ für die "Vektoren" $0, 1, 3 \in \mathbb{R}$).

Nach 9.1 kann dim im(φ) gleich 0 oder 1 sein (da $\dim \mathbb{R} = 1$). Da $\varphi \neq 0$, ist auch im(φ) $\neq 0$, also gilt dim im(φ) = 1. Es folgt mit dem Rangsatz

$$\dim U_1 = \dim \ker(\varphi) = 3 - 1 = 2\,.$$

Bemerkung 9.5 Sind V und W isomorphe K-Vektorräume, so ist $\dim V = \dim W$. Dies folgt aus 8.10, angewendet auf φ und φ^{-1}, falls $\varphi : V \xrightarrow{\sim} W$ ein Isomorphismus ist.

Lemma/Definition 9.6 Für K-Vektorräume $V_1, \ldots, V_n$ definiere den K-Vektorraum

$$\bigoplus_{i=1}^{n} V_i = V_1 \oplus \ldots \oplus V_n$$

(**direkte Summe** der Räume $V_1, \ldots, V_n$), wie folgt:

$$\bigoplus_{i=1}^{n} V_i = \prod_{i=1}^{n} V_i = \{(v_1, \ldots, v_n) \mid v_i \in V_i\}$$

als abelsche Gruppe (siehe 3.15), mit der Skalarmultiplikation

$$\lambda(v_1, \ldots, v_n) := (\lambda v_1, \ldots, \lambda v_n) \quad (\lambda \in K)\,.$$

Beweis, dass dies einen Vektorraum ergibt: selbst!

Satz 9.7 Sind $V_1, V_2, \ldots, V_n$ endlich-dimensionale K-Vektorräume, so auch $\bigoplus\limits_{i=1}^{n} V_i$, und es gilt

$$\dim \bigoplus_{i=1}^{n} V_i = \sum_{i=1}^{n} \dim V_i\,.$$

Beweis Sei zunächst $n = 2$. Ist $(v_1, \ldots, v_m)$ eine Basis von V_1 und $(w_1, \ldots, w_n)$ eine Basis von V_2, so sieht man leicht, dass

$$(v_1, 0), \ldots, (v_m, 0), (0, w_1), \ldots, (0, w_n)$$

eine Basis von $V_1 \oplus V_2$ bilden. Hieraus folgt

$$\dim V_1 \oplus V_2 = m + n = \dim V_1 + \dim V_2\,.$$

Den allgemeinen Fall beweisen wir nun per Induktion.

Sei $n > 2$ und die Aussage für $n - 1$ bewiesen. Wir haben einen Isomorphismus

$$\begin{array}{ccc} \bigoplus\limits_{i=1}^{n} V_i & \cong & \left(\bigoplus\limits_{i=1}^{n-1} V_i\right) \oplus V_n \\ (v_1, \ldots, v_n) & \mapsto & ((v_1, \ldots, v_{n+1}), v_n)\,. \end{array}$$

Damit schließen wir

$$\begin{aligned}
\dim \bigoplus_{i=1}^{n} V_i &\overset{9.3}{=} \dim\left(\bigoplus_{i=1}^{n-1} V_i \oplus V_n\right) \\
&= \dim \bigoplus_{i=1}^{n-1} V_i + \dim V_n \quad \text{(Fall } n=2\text{)} \\
&= \sum_{i=1}^{n-1} \dim V_i + \dim V_n \quad \text{(Induktions-Voraussetzung)} \\
&= \sum_{i=1}^{n} \dim V_i \,.
\end{aligned}$$

Satz 9.8 (Dimensionsformel für Unterräume) Seien U_1, U_2 Unterräume des endlich-dimensionalen K-Vektorraumes V. Dann ist

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim U_1 \cap U_2$$

(Hierbei ist $U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1,\ u_2 \in U_2\}$, vergl. Übungsaufgabe 22).

Beweis Definiere die lineare Abbildung

$$\begin{array}{rrcl}
\varphi: & U_1 \oplus U_2 & \to & V \\
& (u_1, u_2) & \mapsto & u_1 + u_2
\end{array}$$

(die Linearität von φ ist offensichtlich). Es ist $\mathrm{im}(\varphi) = U_1 + U_2$. Weiter ist

$$\ker(\varphi) = \{(u_1, u_2) \in U_1 \oplus U_2 \mid u_1 + u_2 = 0\}\,,$$

und wir erhalten einen Vektorraumisomorphismus

$$\begin{array}{rrcl}
\psi: & U_1 \cap U_2 & \overset{\sim}{\to} & \ker(\varphi) \\
& u & \mapsto & (u, -u)
\end{array}$$

Denn offenbar ist ψ wohldefiniert ($\psi(u) = (u, -u) \in \ker(\varphi)$ für $u \in U_1 \cap U_2$), linear und injektiv. Aber ψ ist auch surjektiv, denn für $(u_1, u_2) \in \ker(\varphi)$ ist $u_1 \in U_1$, $u_2 \in U_2$ und $u_1 + u_2 = 0$, also $u_1 = -u_2$. Damit ist $u_1 \in U_1 \cap U_2$ und $(u_1, u_2) = \psi(u_1)$. Es folgt

$\dim U_1 + \dim U_2$
$\overset{9.7}{=} \dim(U_1 \oplus U_2)$
$= \dim \ker(\varphi) + \dim\, \mathrm{im}(\varphi)$ (Rangsatz)
$\overset{9.5}{=} \dim U_1 \cap U_2 + \dim(U_1 + U_2)$, woraus die Behauptung folgt.

Satz/Definition 9.9 Seien V und W zwei K-Vektorräume. Die Menge

$$\mathrm{Hom}_K(V, W) = \{\varphi : V \to W \mid \varphi \text{ linear}\}$$

der linearen Abbildungen von V nach W wird wie folgt zu einem K-Vektorraum: Für $f, g \in \mathrm{Hom}_K(V, W)$ und $\lambda \in K$ definiere $f + g : V \to W$ durch

$$(f + g)(v) := f(v) + g(v) \qquad (v \in V)\,.$$

und $\lambda f : V \to W$ durch

$$(\lambda f)(v) := \lambda \cdot f(v) \qquad (v \in V)\,.$$

Beweis der Behauptung: Wir haben nur zu zeigen, dass die Verknüpfungen wohldefiniert sind, d.h., dass die Abbildungen $f + g$ und λf wieder linear sind, d.h., in $\mathrm{Hom}_K(V, W)$ liegen. Dann folgt mit dem Unterraum-Kriterium, dass $\mathrm{Hom}_K(V, W)$ ein Unterraum des K-Vektorraums $\mathrm{Abb}(V, W)$ *aller* Abbildungen $f : V \to W$ ist (siehe Beispiel 5.2(b)). Es gilt aber

$(f + g)(v_1 + v_2) = f(v_1 + v_2) + g(v_1 + v_2)$ (Definition)
$= f(v_1) + f(v_2) + g(v_1) + g(v_2)$ (f, g linear)
$= f(v_1) + g(v_1) + f(v_2) + g(v_2)$ ($(W, +)$ kommutativ)
$= (f + g)(v_1) + (f + g)(v_2)$ (Definition).

Weiter ist für $v \in V$ und $\lambda \in K$

$(f + g)(\lambda v) = f(\lambda v) + g(\lambda v)$ (Definition)
$= \lambda \cdot f(v) + \lambda \cdot g(v)$ (f, g linear)
$= \lambda \cdot (f(v) + g(v)) = \lambda \cdot (f + g)(v)$.

Also ist $f + g$ linear. Ähnlich zeigt man die Linearität von λf.

Seien nun V und W endlich-dimensional. Ist dann auch $\mathrm{Hom}_K(V, W)$ endlich-dimensional? Wenn ja, was ist die Dimension? Das werden wir später mit Matrizenrechnung beantworten. Jetzt betrachten wir den Spezialfall $W = K$ (d.h., $W = K^n$ mit $n = 1$).

Definition 9.10 Sei V ein K-Vektorraum. Der Vektorraum

$$V^* := \mathrm{Hom}_K(V, K)$$

heißt der **Dualraum** von V. Seine Elemente, also die linearen Abbildungen

$$\varphi : V \to K$$

heißen die **linearen Funktionale** auf V.

Beispiele 9.11 (a) Die Abbildung

$$\begin{array}{rccc} \varphi : & \mathbb{R}^3 & \to & \mathbb{R} \\ & \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto & y + 3z \end{array}$$

(siehe Beispiel 9.4) ist ein lineares Funktional auf $\mathbb{R}^3$.

(b) Für jedes feste $i \in \{1, \dots, n\}$ ist die i-te Projektion

$$\begin{array}{rccc} p_i : & \mathbb{R}^n & \to & \mathbb{R} \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & x_i \end{array}$$

ein lineares Funktional auf $\mathbb{R}^n$.

(c) Allgemeiner seien $a_1, \ldots, a_n \in \mathbb{R}$. Dann ist die Abbildung

$$\begin{array}{rccl} \varphi : & \mathbb{R}^n & \to & \mathbb{R} \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & a_1x_1 + \ldots + a_nx_n \end{array}$$

ein lineares Funktional auf $\mathbb{R}^n$ (genauso: für einen beliebigen Körper K an Stelle von $\mathbb{R}$).

(d) Seien $a, b \in \mathbb{R}$, $a < b$. Für jedes $x_0 \in [a, b]$ ist die Auswertungsabbildung

$$\begin{array}{ccc} C([a,b], \mathbb{R}) & \to & \mathbb{R} \\ f & \mapsto & f(x_0) \end{array}$$

ein lineares Funktional auf dem $\mathbb{R}$-Vektorraum $C([a, b], \mathbb{R})$ der stetige Funktionen auf dem Intervall $[a, b]$.

Satz 9.12 Sei V ein n-dimensionaler K-Vektorraum, $n < \infty$. Dann ist der Dualraum V^* auch ein n-dimensionaler K-Vektorraum.

Beweis Sei $(v_1, \ldots, v_n)$ eine Basis von V. Dann definieren wir eine Basis $(v_1^*, \ldots, v_n^*)$ von V^* wie folgt: Nach der universellen Eigenschaft von Basen (8.31) gibt es zu jedem n-Tupel $(\alpha_1, \ldots, \alpha_n) \in K^n$ genau eine lineare Abbildung

$$\varphi = \varphi^{(v_1,\ldots,v_n)}_{(\alpha_1,\ldots,\alpha_n)} : V \to K$$

mit

$$\varphi(v_i) = \alpha_i \quad (i = 1, \ldots, n)\,.$$

Wir setzen nun $v_i^* := \varphi^{(v_1,\ldots,v_n)}_{e_i}$, wobei e_i der i-te Einheitsvektor in K^n ist. Explizit ist also das lineare Funktional

$$v_i^* : V \to K$$

dadurch bestimmt, dass gilt

(9.12.1) $$v_i^*(v_j) = \delta_{ij} \quad \text{(Kronecker-Symbol)}$$

Behauptung $(v_1^*, \ldots, v_n^*)$ ist eine Basis von V^*.

Beweis 1) Erzeugendensystem:
Sei $\varphi : V \to K$ eine lineare Abbildung. Nach der Vorbemerkung gibt es dann ein n-Tupel $(\alpha_1, \ldots, \alpha_n)$ mit $\varphi(v_j) = \alpha_j$. Dann ist aber

(9.12.2) $$\varphi = \sum_{i=1}^{n} \alpha_i v_i^*\,,$$

denn es gilt für jedes $j \in \{1, \ldots, n\}$

$$\varphi(v_j) = \alpha_j$$

und
$$\left(\sum_{i=1}^{n} \alpha_i v_i^*\right)(v_j) = \sum_{i=1}^{n} \alpha_i v_i^*(v_j) = \sum_{i=1}^{n} \alpha_i \delta_{ij} = \alpha_j \,.$$

φ und $\sum\limits_{i=1}^{n} \alpha_i v_i^*$ stimmen also auf der Basis $v_1, \ldots, v_n$ überein, sind also gleich nach 8.31.

2) linear unabhängig:
Seien $\alpha_1, \ldots, \alpha_n \in K$ mit
$$\sum_{i=1}^{n} \alpha_i v_i^* = 0 \text{ in } V^* \,.$$

Für jedes $j \in \{1, \ldots, n\}$ gilt dann
$$0 = \left(\sum_{i=1}^{n} \alpha_i v_i^*\right)(v_j) = \sum_{i=1}^{n} \alpha_i \delta_{ij} = \alpha_j \,,$$

was zu zeigen war.

Definition 9.13 Sei $(v_1, \ldots, v_n)$ eine Basis von V. Dann heißt die in Beweis von 9.12 konstruierte Basis $(v_1^*, \ldots, v_n^*)$ von V^* – die also durch
$$v_i^*(v_j) = \delta_{ij}$$
bestimmt ist – die **Dualbasis** zu $(v_1, \ldots, v_n)$.

Beispiele 9.14 (a) Betrachte die kanonische Basis $(e_1, \ldots, e_n)$ von K^n. Wir bestimmen explizit die Dualbasis $(e_1^*, \ldots, e_n^*)$ dazu. Sei i fest. Nach Definition gilt $e_i^*(e_j) = \delta_{ij}$, also
$$e_i^*((x_1, \ldots, x_n)) = e_i^*\left(\sum_{j=1}^{n} x_i e_j\right) = \sum_{j=1}^{n} x_j e_i^*(e_j) = \sum_{j=1}^{n} x_j \delta_{ij} = x_i \,.$$

Also ist
$$e_i^* = p_i : K^n \to K$$
die i-te Projektion (oder Projektion auf die i-te Komponente).

(b) Betrachte das lineare Funktional
$$\begin{array}{rccc} \varphi: & \mathbb{R}^3 & \to & \mathbb{R} \\ & \begin{pmatrix} x \\ y \\ z \end{pmatrix} & \mapsto & y + 3z \end{array}$$

aus 9.11 (a). Nach Satz 9.12 muss φ Linearkombination der e_i^* sein – tatsächlich ist
$$\varphi = e_2^* + 3e_3^* \,,$$

denn es ist
$$e_2^*\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = p_2\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = y, \; e_3^*\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = p_3\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = z \,.$$

(c) Allgemeiner seien $\alpha_1, \ldots, \alpha_n \in K$ und φ das lineare Funktional

$$\begin{array}{rccl} \varphi: & K^n & \to & K \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \mapsto & \alpha_1 x_1 + \ldots + \alpha_n x_n \end{array} \tag{9.14.1}$$

(siehe 9.11.(c) für $K = \mathbb{R}$). Dann ist

$$\varphi = \sum_{i=1}^{n} \alpha_i e_i^* \qquad \left(= \sum_{i=1}^{n} \alpha_i p_i \right)$$

– und umgekehrt. Nach Satz 9.12 ist also jedes lineare Funktional auf K^n von der Form (9.14.1), für geeignete (eindeutige!) $\alpha_1, \ldots, \alpha_n \in K$.

§10 Lineare Abbildungen und Matrizen

Matrizen entstehen zunächst, wenn man lineare Gleichungssysteme wie

$$\begin{array}{l} 3x + 4y + 5z = 0 \\ 6x + 7y + 8z = 1 \end{array} \tag{10.0}$$

in mehreren Unbekannten betrachtet.

Definition 10.1 Sei K ein Körper. Ein (inhomogenes) lineares Gleichungssystem über K mit m Gleichungen und n Unbekannten ist ein Gleichungssystem

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + \ldots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \ldots + & a_{2n}x_n & = & b_2 \\ & & & \vdots & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \ldots + & a_{mn}x_n & = & b_n \end{array} \tag{10.11}$$

mit $a_{ij} \in K$ für alle $i = 1, \ldots, m$ und alle $j = 1, \ldots, n$ und $b_i \in K$ für alle $i = 1, \ldots, m$. Das Gleichungssystem heißt homogen, wenn $b_1 = b_2 = \ldots = b_m = 0$. Eine Lösung des Gleichungssystems ist ein $(x_1, \ldots, x_n) \in K^n$, welches diese Gleichungen erfüllt.

Definition 10.2 Sei K ein Körper (oder ein Ring), und seien $m, n \in \mathbb{N}$.

(a) Eine $(m \times n)$-Matrix über K ist eine Familie

$$A = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

von $m \cdot n$ Zahlen a_{ij} in K. Wir schreiben eine $(m \times n)$-Matrix in der Form

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

Sei $M(m \times n, K)$ die Menge der $(m \times n)$-Matrizen über K.

(b) Für $j \in \{1, \ldots, n\}$ heißt

$$S_j(A) := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = (a_{ij})_{i=1,\ldots,m} \in K^m$$

der j-te Spaltenvektor von A, und für $i \in \{1, \ldots, m\}$ heißt

$$Z_i(A) = (a_{i1}, a_{i2}, \ldots, a_{in}) = (a_{ij})_{j=1,\ldots,n} \in K^n$$

der i-te Zeilenvektor von A.

$$\begin{array}{rc} \begin{array}{r} \\ \\ \\ i - \text{te Zeile } \rightarrow \\ \\ \\ \end{array} & \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1j} & \ldots & a_{1n} \\ a_{21} & a_{22} & & \vdots & & a_{2n} \\ \vdots & & & \vdots & & \vdots \\ a_{i1} & \ldots & \ldots & a_{ij} & \ldots & a_{in} \\ \vdots & & & \vdots & & \vdots \\ a_{m1} & \ldots & \ldots & a_{mj} & \ldots & a_{mn} \end{pmatrix} \\ & \begin{array}{c} \uparrow \\ j - \text{te Spalte} \end{array} \end{array}$$

(c) Für $x = (x_1, \ldots, x_n) \in K^n$ sei Ax ("Anwendung von A auf x") der Vektor $y = (y_1, \ldots, y_m) \in K^m$ mit

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \, .$$

In Zukunft schreiben wir Vektoren in K^r als Spaltenvektoren. Dann sagt diese Definition

$$Ax = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ & \vdots & \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 & + & a_{12}x_2 & + \ldots + & a_{1n}x_n \\ a_{21}x_1 & + & a_{22}x_2 & + \ldots + & a_{2n}x_n \\ & & \vdots & & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n \end{pmatrix} \in K^m$$

Bemerkung 10.3 Definiert man das (Standard-) Skalarprodukt in K^n durch

$$x \cdot y = \sum_{j=1}^{n} x_i y_j \text{ für } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K, \text{ so ist also}$$

$$(Ax)_i = Z_i(A) \cdot x \, ,$$

also die i-te Komponente von Ax gleich dem Skalarprodukt des i-ten Zeilenvektors von A mit x.

Ist also A eine $(m \times n)$-Matrix und $b \in K^m$ ein Vektor, so ist

$$Ax = b$$

ein lineares Gleichungssystem mit m Gleichungen und n Unbekannten. Umgekehrt kann jedes solches Gleichungssystem so geschrieben werden.

Beispiel 10.4 Die Matrix zum Gleichungssystem (10.0)

$$\begin{array}{rcl} 3x + 4y + 5z &=& 0 \\ 6x + 7y + 8z &=& 1 \end{array}$$

ist die (2×3)-Matrix

$$A = \begin{pmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}.$$

Mit dem Vektor

$$b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

schreibt sich (10.0) als

$$Ax = b\,,$$

ausgeschrieben

$$\begin{pmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Es ist

$$\begin{pmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 - 4 + 10 \\ 6 - 7 + 16 \end{pmatrix} = \begin{pmatrix} 9 \\ 15 \end{pmatrix},$$

also ist $(1, -1, 2)$ keine Lösung des linearen Gleichungssystems (10.0).

Wir kommen nun zum Zusammenhang mit linearen Abbildungen.

Satz 10.5 Sei K ein Körper. Ist $A \in M(m \times n, K)$, so ist die Abbildung

$$\begin{array}{rccc} \varphi_A : & K^n & \to & K^m \\ & x & \mapsto & Ax \end{array}$$

eine lineare Abbildung. Ist umgekehrt

$$\varphi : K^n \to K^m$$

eine lineare Abbildung, so gibt es genau eine Matrix $A \in M(m \times n, K)$ mit $\varphi = \varphi_A$; wir schreiben $A = M(\varphi)$.

Beweis Für die erste Behauptung müssen wir zeigen

$$\begin{array}{rcll} A(x + y) &=& Ax + Ay & \forall x, y \in K^n \\ A(\lambda \cdot x) &=& \lambda \cdot Ax & \forall \lambda \in K,\ \forall x \in K^n \end{array}$$

Dies ist klar; zum Beispiel ist

$$\begin{aligned}
(A(x+y))_i &= \sum_{j=1}^n a_{ij}(x+y)_j = \sum_{j=1}^n a_{ij}(x_j+y_j) \\
&= \sum_{j=1}^n a_{ij}x_j + \sum_{j=1}^n a_{ij}y_j = (Ax)_i + (Ay)_i \\
&= (Ax+Ay)_i\,.
\end{aligned}$$

Sei nun $\varphi : K^n \to K^m$ linear. Wie muß A aussehen, damit $\varphi = \varphi_A$ ist?

Antwort: Sei $(e_1, \dots, e_n)$ die Standardbasis des K^n. Dann ist

$$\begin{aligned}
Ae_k &= \left(\sum_{j=1}^n a_{ij}(e_k)_j\right)_{i=1,\dots,m} \\
&= \left(\sum_{j=1}^n a_{ij}\delta_{kj}\right)_{i=1,\dots,m} = (a_{ik})_{i=1,\dots,m} \\
&= S_k(A) = k\text{-ter Spaltenvektor von } A\,.
\end{aligned}$$

Es muß also gelten

$$S_k(A) = \varphi(e_k)$$

Sei A die Matrix mit diesen Spalten, also

$$A = \begin{pmatrix} | & | & & | \\ \varphi(e_1) & \varphi(e_2) & \dots & \varphi(e_n) \\ | & | & & | \end{pmatrix}$$

(die $\varphi(e_j)$ als Spaltenvektoren geschrieben).

Behauptung $\varphi = \varphi_A$.

Beweis $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^n x_i e_j \in K^n$

$\Rightarrow \varphi(x) = \sum_{j=1}^n x_j \varphi(e_j) = \sum_{j=1}^n x_j S_j(A)$

$\Rightarrow \varphi(x)_i = \sum_{j=1}^n x_j S_j(A)_i = \sum_{j=1}^n x_j a_{ij} = (Ax)_i = \varphi_A(x)_i$.

q.e.d.

Bemerkungen 10.6 (a) Wir haben im Beweis gesehen: Ist $A \in M(m \times n, K)$, so ist die j-te Spalte von A

$$S_j(A) = Ae_j$$

also das Bild des j-ten Einheitsvektors (unter der Abbildung $\varphi_A : x \mapsto Ax$).

(b) Wir können die Matrix $A = M(\varphi)$ auch so beschreiben: Es ist

$$a_{ij} = \varphi(e_j)_i\,,$$

d.h., A ist bestimmt durch

(10.6.1)
$$\boxed{\varphi(e_j^{(n)}) = \sum_{i=1}^{m} a_{ij} e_i^{(m)}}$$

wobei $(e_1^{(n)}, \ldots, e_n^{(n)})$ bzw. $(e_1^{(m)}, \ldots, e_m^{(m)})$ die kanonischen Basen von K^n bzw. K^m sind. Man beachte den Unterschied zur Formel

(10.6.2)
$$\boxed{(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j}$$

in der die Summation über den *ersten* Index ist.

Nach Satz 10.5 haben wir eine Bijektion

$$\begin{array}{rcl} M(m \times n, K) & \to & \mathrm{Hom}_K(K^n, K^m) \\ A & \mapsto & \varphi_A\,. \end{array}$$

mit Umkehrabbildung $\varphi \mapsto M(\varphi)$. Genauer gilt:

Satz/Definition 10.7 Mache $M(m \times n, K)$ zu einem K-Vektorraum durch

$$\begin{array}{rcl} (a_{ij}) + (b_{ij}) & := & (a_{ij} + b_{ij}) \\ \lambda \cdot (a_{ij}) & := & (\lambda a_{ij}) \end{array}$$

für $A = (a_{ij})$, $B = (b_{ij}) \in M(m \times n, K)$ und $\lambda \in K$. Dann ist die Abbildung

$$\begin{array}{rcl} M(m \times n, K) & \to & \mathrm{Hom}_K(K^n, K^m) \\ A & \mapsto & \varphi_A \end{array}$$

ein Vektorraum-Isomorphismus, also auch die Umkehrabbildung $\varphi \mapsto M(\varphi)$.

Beweis Man sieht leicht, dass die angegebenen Verknüpfungen $M(m \times n, K)$ zu einem K-Vektorraum machen, und wir wissen schon, dass die Abbildung $A \mapsto \varphi_A$ bijektiv ist. Weiter ist offenbar

$$\begin{array}{rcl} \varphi_{A+B}(x) & = & (A+B)x = Ax + Bx \\ & = & \varphi_A(x) + \varphi_B(x) = (\varphi_A + \varphi_B)(x) \end{array}$$

und

$$\begin{array}{rcccl} \varphi_{\lambda A}(x) & = & (\lambda A)x & = & \lambda \cdot (Ax) \\ & = & \lambda \cdot \varphi_A(x) & = & (\lambda \cdot \varphi_A)(x) \end{array}$$

Dies zeigt, dass die Matrizenaddition und -Skalarmultiplikation gerade der Addition und Skalarmultiplikation in $\mathrm{Hom}_K(K^n, K^m)$ entspricht, d.h., dass die Abbildung $A \mapsto \varphi_A$ linear ist. Es folgen die restlichen Behauptungen.

Lemma 10.8 Eine Basis von $M(m \times n, K)$ ist gegeben durch die Matrizen

$$E_{rs} := (\delta_{ri}\delta_{sj})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} ,$$

für $r \in \{1, \dots, m\}$ und $s \in \{1, \dots, n\}$. Ausführlich geschrieben

$$E_{rs} = \begin{array}{c} \begin{pmatrix} 0 & \dots & 0 & \dots & \dots & 0 \\ \vdots & & \vdots & & & \vdots \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & 1 & \dots & \dots & 0 \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & 0 & \dots & \dots & 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow r - \text{te Zeile} \\ \\ \\ \end{array} \\ \uparrow \\ s - \text{te Spalte} \end{array}$$

Insbesondere ist $\dim \operatorname{Hom}_K(K^n, K^m) = \dim M(m \times n, K) = m \cdot n$.

Beweis Dies ist klar: es ist gerade $(a_{ij}) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \cdot E_{ij}$.

Was entspricht der Komposition von linearen Abbildungen auf der Seite von Matrizen? Antwort: das Matrizenprodukt:

Definition 10.9 Seien $A = (a_{ij}) \in M(m \times n, K)$ und $B = (b_{ij}) \in M(n \times r, K)$. Dann ist das Matrizenprodukt $A \cdot B \in M(m \times r, K)$ definiert als die Matrix $C = (c_{ik})$ mit

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

($= Z_i(A) \cdot S_k(B) =$ Skalarprodukt von i-ter Zeile von A mit j-ter Spalte von B).

$$i\text{-te Zeile} \rightarrow m\left\{ \overbrace{\begin{pmatrix} \\ a_{i1} & \dots & \dots & a_{in} \\ \\ \end{pmatrix}}^{n} \right. \begin{array}{c} \left. \overbrace{\begin{pmatrix} & b_{1k} & \\ & \vdots & \\ & & \\ & \vdots & \\ & b_{nk} & \end{pmatrix}}^{r} \right\} n \\ \uparrow \\ k\text{-te Spalte} \end{array} = m \left\{ \begin{array}{c} \overbrace{\begin{pmatrix} & & \\ & & c_{ik} \\ & & \end{pmatrix}}^{r} \leftarrow i\text{-te Zeile} \\ \uparrow \\ k\text{-te Spalte} \end{array} \right.$$

Also: $(m \times n)$-Matrix $\cdot (n \times r)$-Matrix $= (m \times r)$-Matrix.

Beispiel 10.10 Es ist

$$\begin{array}{ccccc} A & \cdot & B & & C \\ \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix} & & \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 3 \end{pmatrix} & = & \begin{pmatrix} 2 & 1 & 5 \\ 4 & 3 & 9 \\ 6 & 5 & 13 \\ 8 & 7 & 17 \end{pmatrix} \\ 4 \times 2 & & 2 \times 3 & & 4 \times 3 . \end{array}$$

Manchmal ist es nützlich, die Matrizen so anzuordnen:

$$\begin{array}{ccc} & \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 3 \end{pmatrix} & = B \\ A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix} & \begin{pmatrix} & | & \\ - & 3 & - \\ & | & \\ & | & \end{pmatrix} & = C\,. \end{array}$$

Bemerkung 10.11 Sind $w_1, \ldots, w_n$ die Spalten von B, so sind $Aw_1, \ldots, Aw_n$ die Spalten von C:

$$A \cdot \begin{pmatrix} | & | & & | \\ w_1 & w_2 & \ldots & w_n \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ Aw_1 & Aw_2 & \ldots & Aw_n \\ | & | & & | \end{pmatrix}\,.$$

Satz 10.12 Das Matrizenprodukt entspricht der Komposition von linearen Abbildungen, d.h., es ist

(10.10.1) $$\varphi_A \circ \varphi_B = \varphi_{A \cdot B}\,.$$

Mit anderen Worten: Für $x \in K^r$ gilt

(10.10.2) $$A(Bx) = (A \cdot B)x\,.$$

Beweis Die Aussagen sind äquivalent, denn für $x \in K^r$ gilt

$$\begin{array}{rrcl} & (\varphi_A \circ \varphi_B)(x) & = & \varphi_A(\varphi_B(x)) = A(Bx) \\ \text{und} & \varphi_{A \cdot B}(x) & = & (A \cdot B)x\,. \end{array}$$

Die Gleichheit (10.10.2) folgt nun so: Sei $i \in \{1, \ldots, m\}$; mit $C = A \cdot B$ gilt dann

$$\begin{aligned} (A(Bx))_i &= \sum_{j=1}^n a_{ij}(Bx)_j = \sum_{j=1}^n a_{ij} \sum_{k=1}^r b_{jk}x_k = \sum_{j=1}^n \sum_{k=1}^r a_{ij}b_{jk}x_k \\ &= \sum_{k=1}^r \sum_{j=1}^n a_{ij}b_{jk}x_k = \sum_{k=1}^r \left(\sum_{j=1}^n a_{ij}b_{jk} \right) x_k = \sum_{k=1}^r c_{ik}x_k = (Cx)_i\,. \end{aligned}$$

Lemma 10.13 Für Matrizen

$$\begin{array}{rcl} A, A' & \in & M(m \times n, K) \\ B, B' & \in & M(n \times r, K) \\ C & \in & M(r \times s, K) \end{array}$$

und $\lambda \in K$ gilt

(0)	(Kommutativität)	$A + A' = A' + A$.
(i)	(Distributivität)	$A \cdot (B + B') = A \cdot B + A \cdot B'$
		$(A + A) \cdot B = A \cdot B + A' \cdot B$.
(ii)	(Assoziativität)	$A \cdot (B \cdot C) = (A \cdot B) \cdot C$.
(iii)	(Linearität)	$(\lambda A) \cdot B = \lambda(A \cdot B) = A \cdot (\lambda B)$.

Beweis: durch einfaches Nachrechnen: Zum Beispiel ist

$$\begin{aligned}
\text{(ii)}: \ (A\cdot(B\cdot C))_{i\ell} &= \sum_{j=1}^{n} a_{ij}(B\cdot C)_{j\ell} \\
&= \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{r} b_{jk}c_{k\ell} \\
&= \sum_{k=1}^{r} \left(\sum_{j=1}^{n} a_{ij}b_{jk} \right) c_{k\ell} \\
&= \sum_{k=1}^{r} (A\cdot B)_{ik}c_{k\ell} = ((A\cdot B)\cdot C)
\end{aligned}$$

Ein anderer (eleganterer) Beweis ergibt sich aus 10.7 und 10.12, indem man das Lemma 10.17 (siehe unten) auf $V = K^n$, $W = K^m$, $U = K^r$ und $T = K^s$ anwendet.

Corollar 10.14 $M_n(K) := M(n \times n, K)$ ist mit $+$ und $\cdot$ ein Ring, der **Ring der quadratischen** $(n\times n)$**-Matrizen**.

Lemma/Definition 10.15 Der Ring $M_n(K)$ hat ein Einselement, nämlich die n**-te Einheitsmatrix**

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

d.h., $E_n = (\delta_{ij})$. Es ist also

$$E_n \cdot A = A = A \cdot E_n$$

für alle $A \in M_n(K)$.

Beweis: direktes Nachrechnen, oder: Es ist $E_n = M(\mathrm{id}_{K^n})$, d.h., $E_n x = x$ für alle $x \in K^n$, und Satz 10.7 sagt: $Ax = A'x$ für alle $x \in K^n \Rightarrow A = A'$.

Bemerkung 10.16 Allgemeiner gilt für $A \in M(m\times n, K)$

$$E_m \cdot A = A = A \cdot E_n\,.$$

Die entsprechenden Sätze für beliebige Vektorräume lauten:

Lemma 10.17 Für k-lineare Abbildungen

$$\begin{aligned}
\varphi, \varphi' &: V \to W \\
\psi, \psi' &: U \to V \\
\rho &: T \to U \qquad \text{und } \lambda \in K \text{ gilt:}
\end{aligned}$$

(0)	(Kommutativität)	$\varphi + \varphi' = \varphi' + \varphi$.
(i)	(Distributivität)	$\varphi\circ(\psi+\psi') = \varphi\circ\psi + \varphi\circ\psi$
		$(\varphi+\varphi')\circ\psi = \varphi\circ\psi + \varphi'\circ\psi'$.
(ii)	(Assoziativgesetz)	$\varphi\circ(\psi\circ\rho) = (\varphi\circ\psi)\circ\rho$.
(iii)	(Linearität)	$(\lambda\varphi)\circ\psi = \lambda\cdot(\varphi\circ\psi) = \varphi\circ(\lambda\cdot\psi)$.

Beweis (0) wissen wir schon: $\mathrm{Hom}_K(V,W)$ ist kommutative Gruppe bezüglich $+$.

(i): Für alle $u \in U$ ist

$$\begin{aligned}
(\varphi \circ (\psi + \psi'))(u) &= \varphi((\psi + \psi')(u)) \\
&= \varphi(\psi(u) + \psi'(u)) \\
&= \varphi(\psi(u)) + \varphi(\psi'(u)) \quad (\varphi \text{ linear}) \\
&= (\varphi \circ \psi)(u) + (\varphi \circ \psi')(u) \\
&= (\varphi \circ \psi + \varphi \circ \psi')(u)\,.
\end{aligned}$$

Der zweite Fall ist analog.

(ii) gilt allgemein für Verknüpfungen von Abbildungen.

(iii) selbst!

Corollar 10.18 Für jeden K-Vektorraum V ist

$$\mathrm{End}_K(V) := \mathrm{Hom}_K(V,V)$$

mit $+$ und $\circ$ ein Ring mit Eins – der **Endomorphismenring** von V.

Für endlich-dimensionale Vektorräume gilt der folgende bemerkenswerte Satz:

Satz 10.19 Seien V, W K-Vektorräume mit $\dim V = \dim W = n < \infty$. Für eine lineare Abbildung

$$\varphi : V \to W$$

sind äquivalent:

(i) Es gibt eine lineare Abbildung $\psi : W \to V$ mit $\psi \circ \varphi = \mathrm{id}_V$.

(ii) Es gibt eine lineare Abbildung $\psi' : W \to V$ mit $\varphi \circ \psi' = \mathrm{id}_W$.

(iii) φ ist injektiv.

(iv) φ ist surjektiv.

(v) φ ist ein Isomorphismus.

Beweis Es gilt (i) $\Rightarrow$ (iii) und (ii) $\Rightarrow$ (iv) (siehe Satz 2.14). Weiter gilt nach dem Rangsatz

$$\ker(\varphi) = 0 \Leftrightarrow \dim \mathrm{im}(\varphi) = \dim V = \dim W \Leftrightarrow \mathrm{im}(\varphi) = W\,,$$

also (iii) $\Leftrightarrow$ (iv) $\Leftrightarrow$ (v).
Aus (v) folgen aber auch (i) und (ii), mit $\psi = \psi' = \varphi^{-1}$.

Angewendet auf $V = W = K^n$ ergibt sich sofort das Folgende:

Lemma/Definition 10.20 Für eine quadratische Matrix $A \in M_n(K)$ sind äquivalent:

(i) Es gibt ein $B \in M_n(K)$ mit $B \cdot A = E_n$.

(ii) Es gibt ein $B' \in M_n(K)$ mit $A \cdot B' = E_n$.

(iii) φ_A ist injektiv.

(iv) φ_A ist surjektiv.

(v) φ_A ist ein Isomorphismus.

Gelten diese Bedingungen, so heißt die Matrix A **regulär** oder **invertierbar**, es ist $B = B'$ und wird mit A^{-1} bezeichnet; A^{-1} heißt die **inverse Matrix** zu A.

Wir verbinden nun beliebige endlich-dimensionale Vektorräume mit Matrizen:

Definition 10.21 Sei K ein Körper. Seien V und V' K-Vektorräume mit Basen $B = (v_1, \ldots, v_n)$ bzw. $B' = (v'_1, \ldots, v'_m)$ (es ist also $\dim V = n$ und $\dim V' = m$). Für eine lineare Abbildung

$$\varphi : V \to V'$$

definiere dann die Matrix

$$M_B^{B'}(\varphi) = (a_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} \in M(m \times n, K)$$

durch die Beziehung

$$\boxed{\varphi(v_j) = \sum_{i=1}^{m} a_{ij} v'_i \qquad , j = 1, \ldots, n\,.}$$

Dies ist sinnvoll: Da B' eine Basis ist, kann jedes $\varphi(v_j)$ als Linearkombination der v'_i dargestellt werden, mit eindeutigen Koeffizienten a_{ij}. $M_B^{B'}(\varphi)$ heißt **die Matrix von φ bezüglich der Basen B und B'.**

Satz 10.22 Die Abbildung

$$\begin{array}{rrcl} M_B^{B'} : & \mathrm{Hom}_K(V, V') & \to & M(m \times n, K) \\ & \varphi & \mapsto & M_B^{B'}(\varphi) \end{array}$$

ist ein Vektorraum-Isomorphismus. Insbesondere ist $\dim \mathrm{Hom}_K(V, V') = m \cdot n$.

Beweis 1) $M_B^{B'}$ ist linear, d.h.,

$$\begin{array}{rcl} M_B^{B'}(\varphi + \psi) & = & M_B^{B'}(\varphi) + M_B^{B'}(\psi) \\ M_B^{B'}(\lambda\varphi) & = & \lambda \cdot M_B^{B'}(\varphi)\,. \end{array}$$

Dies ist klar.

2) Die Umkehrabbildung ergibt sich aus der universellen Eigenschaft von Basen (8.31):

Gegeben $A = (a_{ij}) \in M(m \times n, K)$, so gibt es genau eine lineare Abbildung $\varphi : V \to V'$ mit

$$\varphi(v_j) = w_j := \sum_{i=1}^{m} a_{ij} v'_i\,.$$

Bemerkungen 10.23 (a) Wir können $M_B^{B'}(\varphi) = (a_{ij})$ auch anders interpretieren: Wir haben ein kommutatives Diagramm

$$\begin{array}{ccccc} v_j & & V \xrightarrow{\ \varphi\ } V' & & v'_i \\ \uparrow & & \varphi_B \uparrow \wr \qquad \wr \uparrow \varphi_{B'} & & \uparrow \\ e_j^{(n)} & & K^n \xrightarrow[M_B^{B'}(\varphi)]{} K^m & & e_i^{(m)}\,, \end{array}$$

wenn wir $\mathrm{Hom}_K(K^n, K^m)$ mit $M(m \times n, K)$ identifizieren. Ganz genau ist

$$M_B^{B'}(\varphi) = M(\varphi_{B'}^{-1} \circ \varphi \circ \varphi_B)\,.$$

Denn:

$$\begin{aligned} (\varphi_{B'}^{-1} \circ \varphi \circ \varphi_B)(e_j^{(n)}) = \varphi_{B'}^{-1}(\varphi(\varphi_B(e_j^{(n)}))) &= \varphi_{B'}^{-1}(\varphi(v_j)) \\ &= \varphi_{B'}^{-1}(\sum_{i=1}^m a_{ij} v'_i) \\ &= \sum_{i=1}^m a_{ij} e_i^{(m)} \end{aligned}$$

Hieraus folgt die Behauptung, nach Bemerkung 10.6 (b): $M(\varphi_B^{-1} \circ \varphi \circ \varphi_B) = (a_{ij}) \Leftrightarrow (\varphi_B^{-1} \circ \varphi \circ \varphi_B)(e_j^{(n)}) = \sum\limits_{i=1}^m a_{ij} e_i^{(m)}$.

(b) Offenbar ist für $\varphi \in \mathrm{Hom}_K(K^n, K^m)$ gerade $M(\varphi) = M_{E^{(n)}}^{E^{(m)}}(\varphi)$, wobei $E^{(r)}$ die Standardbasis von K^r ist; d.h., die vorigen Überlegungen über Matrizen sind nur Spezialfälle, für die kanonischen Basen von K^n und K^m.

Beispiele 10.24 (a) Betrachte die Standardbasis (e_1, e_2) zu $\mathbb{R}^2$ und die zwei Vektoren

$$w_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},\ w_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

in $\mathbb{R}^3$. Wir wissen (8.31): es gibt genau eine lineare Abbildung

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^3$$

mit $\varphi(e_1) = w_1$ und $\varphi(e_2) = w_2$. Nach der Konstruktion in 10.5 ist dies φ_A für die Matrix

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

deren Spalten die Bilder w_1 und w_2 von e_1 und e_2 sind (Bemerkung 10.6 (a): die j-te Spalte von A ist das Bild von e_j).

(b) Jetzt betrachte die Basis $B' = (v_1, v_2)$ von $\mathbb{R}^2$, mit $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Berechne die Matrix

$$A' = M_{B'}^{E^{(3)}}(\varphi)\,.$$

Es ist

$$\begin{aligned}
\varphi(v_1) &= \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} = 0e_1 + 2e_2 + 2e_3, \\
\varphi(v_2) &= \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 2 \\ 2 \\ 4 \end{pmatrix} = 2e_1 + 2e_2 + 4e_3 \,.
\end{aligned}$$

Also ist

$$A' = \begin{pmatrix} 0 & 2 \\ 2 & 2 \\ 2 & 4 \end{pmatrix} .$$

Lemma 10.25 Ist V'' ein weiterer K-Vektorraum, mit Basis $B'' = (v_1'', \ldots, v_r'')$ (also $\dim V'' = r$), so ist für eine lineare Abbildung $\psi : V' \to V''$

$$M_B^{B''}(\psi \circ \varphi) = M_{B'}^{B''}(\psi) \cdot M_B^{B'}(\varphi) \,.$$

Beweis: Dies folgt leicht aus der Definition 10.21, mit denselben Schlüssen wie für 10.12. Oder es folgt so aus 10.12:

$$\begin{aligned}
M_B^{B''}(\psi \circ \varphi) &\overset{10.23\,(a)}{=} M(\varphi_{B''}^{-1} \circ \psi \circ \varphi \circ \varphi_B) \\
&= M(\varphi_{B''}^{-1} \circ \psi \circ \varphi_{B'} \circ \varphi_{B'}^{-1} \circ \varphi \circ \varphi_B) \\
&\overset{10.12}{=} M(\varphi_{B''}^{-1} \circ \psi \circ \varphi_{B'}) \cdot M(\varphi_{B'} \circ \varphi \circ \varphi_B) \\
&\overset{10.23\,(a)}{=} M_{B'}^{B''}(\psi) \cdot M_B^{B'}(\varphi) \,.
\end{aligned}$$

Was ist, wenn wir von einer Basis B zu einer anderen Basis B' übergehen? Wie verändern sich die Matrixdarstellungen einer linearen Abbildung?

Definition 10.26 Sei V ein endlich-dimensionaler K-Vektorraum mit Basen $B = (v_1, \ldots, v_n)$ und $B' = (v_1', \ldots, v_n')$. Definiere die **Transformationsmatrix** (oder **Basiswechselmatrix**) von B nach B', $M_{B'}^B \in M(n \times n, K)$, als die Matrix (c_{ij}) mit

$$v_j' = \sum_{i=1}^{n} c_{ij} v_i \quad , j = 1, \ldots, n \,.$$

Sie beschreibt also, wie die neue Basis B' durch die alte Basis ausgedrückt wird. Mit der Definition 10.21 ist gerade

$$M_{B'}^B = M_{B'}^B(\mathrm{id}_V) \,.$$

Lemma 10.27 $M_{B'}^B$ ist invertierbar; es ist

$$(M_{B'}^B)^{-1} = M_B^{B'} \,.$$

Beweis Nach 10.25 ist

$$\begin{aligned}
M_{B'}^B \cdot M_B^{B'} &= M_{B'}^B(\mathrm{id}_V) \cdot M_B^{B'}(\mathrm{id}_V) \\
&= M_B^B(\mathrm{id}_V) = M_B^B = E_n \,.
\end{aligned}$$

Beispiel 10.28 Betrachte die Standardbasis $E^{(2)} = (e_1, e_2)$ von $\mathbb{R}^2$, sowie die Basis $B' = (v_1, v_2)$ mit

$$v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \; v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Es ist

$$\begin{aligned} v_1 &= e_1 - e_2 \\ v_2 &= e_1 + e_2 , \end{aligned}$$

also ist

$$M_{B'}^{E^{(2)}} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Umgekehrt ist, wie man sofort nachrechnet,

$$\begin{aligned} e_1 &= \tfrac{1}{2}v_1 + \tfrac{1}{2}v_2 \\ e_2 &= -\tfrac{1}{2}v_1 + \tfrac{1}{2}v_2 \quad , \end{aligned}$$

also

$$M_B^{B'} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Wir rechnen nach

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 .$$

Satz 10.29 Sei $\varphi : V \to W$ eine lineare Abbildung von K-Vekorräumen. Seien $B = (v_1, \ldots, v_n)$ und $B' = \{v'_1, \ldots, v'_n\}$ Basen von V, und seien $C = \{w_1, \ldots, w_n\}$ und $C' = \{w'_1, \ldots, w'_n\}$ Basen von W. Dann ist

$$M_{B'}^{C'}(\varphi) = M_C^{C'} \circ M_B^C(\varphi) \circ M_{B'}^B$$

Beweis Nach 10.25 ist

$$M_C^{C'}(\mathrm{id}_W) \cdot M_B^C(\varphi) \cdot M_{B'}^B(\mathrm{id}_V) = M_B^{C'}(\varphi) \cdot M_{B'}^B(\mathrm{id}_V) = M_{B'}^{C'}(\varphi) .$$

Beispiel 10.30 Betrachte die lineare Abbildung $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ aus Beispiel 10.24 (a). Die zugehörige Matrix ist

$$M(\varphi) = M_{E^{(2)}}^{E^{(3)}}(\varphi) = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix} = A .$$

Für die Basis $B' = (v_1, v_2)$ aus 10.24 (b) ist, unter Benutzung von 10.29 und Beispiel 10.28

$$M_{B'}^{E^{(3)}}(\varphi) = M_{E^{(2)}}^{E^{(3)}}(\varphi) \cdot M_{B'}^{E^{(2)}} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 2 & 2 \\ 2 & 4 \end{pmatrix} = A'$$

was mit 10.24 (b) übereinstimmt.

§11 Lineare Gleichungssysteme

Sei K ein Körper. Für eine Matrix $A \in M(m \times n, K)$ und einen Vektor $b \in K^m$ betrachte das zugehörige lineare Gleichungssystem

$$Ax = b$$

für $x \in K^n$.

Definition 11.1 $L(A, b) \subseteq K^n$ sei die Menge der Lösungen dieses Gleichungssystems, also

$$L(A, b) := \{x \in K^n \mid Ax = b\}.$$

Satz 11.2 (a) Die Menge $L(A, 0)$ der Lösungen des homogenen linearen Gleichungssystems

$$Ax = 0$$

ist ein Untervektorraum in K^n.

(b) Die Menge $L(A, b)$ der Lösungen des inhomogenen linearen Gleichungssystems

$$Ax = b$$

ist entweder leer oder von der Form

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

wenn v eine Lösung des linearen Gleichungssystems $Ax = b$ ist.

Beweis (a): *1. Beweis*: Seien $x, y \in L(A, 0)$ und $\lambda \in K$. Dann ist $Ax = 0$ und $Ay = 0$. Es folgt $A(x + y) = Ax + Ay = 0 + 0 = 0$ und $A(\lambda x) = \lambda \cdot Ax = \lambda \cdot 0 = 0$. Also sind $x + y, \lambda x \in L(A, 0)$.

2. Beweis: $L(A, 0) = \ker(\varphi_A)$.

(b): Sei $v \in L(A, b)$, also $Av = b$. Wir zeigen $L(A, b) = v + L(A, 0)$:

"$\supseteq$": $x \in L(A, 0) \Rightarrow Ax = 0 \Rightarrow A(v + x) = Av + Ax = b + 0 = b \Rightarrow v + x \in L(A, b)$
"$\subseteq$": $y \in L(A, b) \Rightarrow Ay = b \Rightarrow A(y - v) = Ay - Av = b - b = 0 \Rightarrow y - v \in L(A, 0) \Rightarrow y = v + (y - v) \in v + L(A, 0)$.

Bemerkung 11.3 (a) (Unschöne) Merkregel: " Die allgemeine Lösung des inhomogenen Systems ist gleich einer speziellen Lösung plus einer allgemeinen Lösung des homogenen Systems".

(b) Für $b \neq 0$ kann $L(A, b)$ leer sein! Beispiel: Sei $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$Ax = b \Leftrightarrow \begin{matrix} x_1 = 0 \\ 0 = 1 \end{matrix}$ nicht lösbar!
Dagegen enthält $L(A, 0)$ immer die 0.

Definition 11.4 Sei V ein K-Vektorraum. Ein **affiner Teilraum** von V ist eine Teilmenge von V von der Form

$$v + W := \{v + w \mid w \in W\}$$

wobei $v \in V$ und $W \subseteq V$ ein Untervektorraum ist.

Beispiel 11.5 In $\mathbb{R}^2$ betrachte den 1-dimensionalen Unterraum

$$G_0 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = -x \right\} = L((1,1),0) = \mathbb{R} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Dann ist

$$G = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + G_0 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \end{pmatrix} \middle| \lambda \in \mathbb{R} \right\} = L((1,1),2) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| x + y = 2 \right\}$$

ein affiner Teilraum von $\mathbb{R}^2$.

Bild

Lemma 11.6 Eine Teilmenge A eines K-Vektorraums V ist genau dann ein affiner Teilraum, wenn für ein $v_0 \in A$ die Menge

$$W_{v_0} := A - v_0 := \{v - v_o \mid v \in A\}$$

ein Untervektorraum von V ist. Dann gilt dies für jedes $v_0 \in A$, und der Unterraum W_{v_0} hängt nicht von v_0 ab. Er heißt der Unterraum $W(A)$ zu A.

Beweis A affiner Teilraum $\Rightarrow$ es existieren $v_0 \in V$ und ein Untervektorraum $W \subseteq V$ mit

$$A = v_0 + W.$$

Dann ist

$$W_{v_0} = A - v_0 = W$$

ein Unterraum von V.

Sei umgekehrt $v_0 \in A$ und $W_{v_0} = A - v_0$ ein Unterraum von V. Dann ist

$$A = v_0 + W_{v_0}$$

denn $v_0 + W_{v_0} = \{v_0 + x \mid x \in W_{v_0}\} = \{v_0 + v - v_0 \mid v \in A\} = A$.
Also ist A affiner Teilraum. Ist schließlich $v_0' \in A$, so ist $u = v_0' - v_0 \in W_{v_0}$ und damit $W_{v_0}' = A - v_0' = A - v_0 - u = W_{v_0} - u = W_{v_0}$.

Wir haben also gesehen:

$L(A,0)$ bildet einen Unterraum von K^n.
$L(A,b)$ ist leer, oder bildet einen affinen Teilraum von K^n (nämlich $v + L(A,0)$, wobei $v \in L(A,b)$ beliebig ist.)

Es ergeben sich folgende Fragen:

1) Wie berechnet man $\dim L(A,0)$?

2) Wie sieht man, ob $Ax = b$ lösbar ist?

3) Wenn $Ax = b$ lösbar ist – wie bestimmt man eine spezielle Lösung?

zu 1): Nach dem Rangsatz ist

$$\dim L(A,0) = \dim \ker(\varphi_A) = n - rg\,\varphi_A$$

Corollar 11.7 Für $A \in M(m \times n, K)$ ist

$$\dim L(A,0) = n - rg\,A\,,$$

wobei $rg\,A \;:=\; rg\,\varphi_A \quad (= \dim\; \mathrm{im}(\varphi_A))$ der Rang der Matrix A ist.

Satz 11.8 Für eine Matrix $A \in M(m \times n, K)$ definiere den **Spaltenrang** (bzw. **Zeilenrang**) als die maximale Anzahl linear unabhängiger Spalten (bzw. Zeilen) von A. Dann ist

$$rg\,A = \text{ Spaltenrang von } A$$

Es ist $\mathrm{rg}\,A \leq \min(m,n)$.

Beweis: Für $x \in K^n$ ist

$$Ax = \sum_{j=1}^{n} x_j S_j(A)\,,$$

wobei $S_j(A)$ der j-te Spaltenvektor ist. Daher ist

$$\mathrm{im}(\varphi_A) = \{Ax \mid x \in K^n\} = \left\{ \sum_{j=1}^{n} x_j S_j(A) \mid x_1, \ldots, x_n \in K \right\} =< S_1(A), \ldots, S_n(A) >_K \;.$$

Damit ist $S_1(A), \ldots, S_n(A)$ ein Erzeugendensystem von $\mathrm{im}(\varphi_A)$. Ist $M \subseteq \{S_1(A), \ldots, S_n(A)\}$ eine maximale linear unabhängige Teilmenge unter den Spalten, etwa $M = \{S_{j_1}, \ldots, S_{j_r}\}$, so ist M nach dem folgenden Lemma eine Basis. Es folgt $rg\,A = \dim\,(\mathrm{im}\,\varphi_A) = |M| = r =$ Spaltenrang.

Zur letzten Behauptung: Nach Definition ist $\mathrm{rg}\,A = \dim \mathrm{im}(A) \leq \dim K^m = m$. Andererseits ist der Spaltenrang $\leq n$, da es nur n Spalten gibt.

Lemma 11.9 Sei V ein K-Vektorraum, $E \subseteq V$ ein Erzeugendensystem, und $M \subseteq E$ eine in E maximale linear unabhängige Teilmenge. Dann ist M eine Basis von V.

Beweis: Z.z.: $< M >_K = V$. Angenommen, $< M >_K \neq V$. Dann gibt es ein $v \in E$ mit $v \notin <M>_K$ (denn sonst $E \subseteq <M>_K$, also auch $V = <E>_K \subseteq <M>_K$, im Widerspruch zur Voraussetzung.) Dann ist aber $M \cup \{u\}$ linear unabhängig nach Lemma 8.15 – Widerspruch!

Bemerkung: Wir werden später zeigen: Spaltenrang von A = Zeilenrang von A.

Definition 11.10 Für ein inhomogenes Gleichungssystem $Ax = b$ heißt die Matrix

$$(A \mid b),$$

die aus A durch Anfügen der Spalte b entsteht, die **erweiterte Matrix** des Gleichungssystems.

Satz 11.11 Das Gleichungssystem $Ax = b$ ist genau lösbar, wenn gilt

$$\text{rg } A = \text{rg } (A \mid b).$$

Beweis: Seien $S_1(A), \ldots, S_n(A)$ die Spalten von A. Ist $Ax = b$, mit $x \in K^n$, so ist

$$b \in \text{ im } (\varphi_A) = < S_1(A), \ldots, S_n(A) >_K,$$

also

$$\begin{aligned} \text{rg } (A \mid b) &= \dim < S_1(A), \ldots, S_n(A), b >_K \\ &= \dim < S_1(A), \ldots, S_n(A) >_K \quad = \text{ rg } A. \end{aligned}$$

Ist umgekehrt rg $(A|b) =$ rg A, so folgt $b \in <S_1(A), \ldots, S_n(A)>_K$. Also gibt es $x_1, \ldots, x_n \in K$ mit

$$b = \sum_{j=1}^{n} x_j S_j(A) = \sum_{j=1}^{n} x_j A e_j = A \left(\sum_{j=1}^{n} x_j e_j \right) = Ax$$

wobei

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

Beispiel 11.12 Im Beispiel 11.3 (b) ist

$$\text{rg } (A \mid b) = \text{ rg } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 2 \neq 1 = \text{ rg } \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \text{ rg } A,$$

also das Gleichungssystem unlösbar.

Wie berechnet man nun *rg A* und spezielle Lösungen?

Vorüberlegung: Betrachte zum Beispiel das lineare Gleichungssystem

$$\begin{aligned} 2x + y &= 1 \\ 3x + 4y &= 4 \end{aligned}$$

Was macht man üblicherweise? Man vereinfacht die Gleichungen:

$$\begin{array}{rcccl} & 2x & + & y & = 1 \\ \text{2. Gleichung - 4. (1. Gleichung):} & & - & 5x & = 0 \end{array}$$

Es folgt $x = 0$ und $y = 1$. In Matrizenschreibweise schreibt sich der obige Prozess so:

$$\begin{array}{rcc} & A & b \\ & \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} & \begin{pmatrix} 1 \\ 4 \end{pmatrix} \\ \text{2. Zeile - 4. (1. Zeile)} & \begin{pmatrix} 2 & 1 \\ -5 & 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{array}$$

Die Matrix wird einfacher (hat mehr Nullen). Ein etwas "automatischeres" Verfahren (denke an 1000×1000-Matrizen) ist wie folgt: Elimination der 1. Variablen x:

$$\begin{array}{rccccc} & 2x & + & y & = & 1 \\ \text{2. Gleichung} - \frac{3}{2}\cdot \text{(1. Gleichung)} & 0 & + & 2,5y & = & 2,5 \end{array}$$

woraus wieder $y = 1$ und $x = 0$ folgt. In Matrizen:

$$\begin{array}{rcc} & \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} & \begin{pmatrix} 1 \\ 4 \end{pmatrix} \\ \text{2. Zeile -}\frac{3}{2}\cdot \text{ (1. Zeile)} & \begin{pmatrix} 2 & 1 \\ 0 & 2,5 \end{pmatrix} & \begin{pmatrix} 1 \\ 2,5 \end{pmatrix} \end{array}$$

Dies führt auf das Gaußsche Eliminationsverfahren (siehe weiter unten). Zunächst motiviert die Vorüberlegung die folgende Definition.

Definition 11.13 Sei $A \in M(m \times n, K)$. Eine **elementare Zeilen-** (bzw. **Spalten-**) **Umformung** (oder **-Transformation**) ist eine der folgenden beiden Operationen:

(a) Multiplikation einer Zeile (bzw. Spalte) von A mit einem $\lambda \in K^\times = K \setminus \{0\}$,

(b) Addition einer Zeile (bzw. Spalte) von A zu einer anderen Zeile (bzw. Spalte).

Lemma 11.14 Durch Iteration dieser Operationen erhält man die folgenden weiteren Operationen:

(c) Addition einer mit einem $\lambda \in K$ multiplizierten Zeile (bzw. Spalte) zu einer anderen Zeile (bzw. Spalte).

(d) Vertauschung zweier Zeilen (bzw. Spalten).

Beweis: (für Zeilen, der Fall von Spalten ist analog) Für (c) sei ohne Einschränkung $\lambda \neq 0$. Dann erhalten wir (c) durch die Hintereinanderausführung folgender elementarer Umformungen:

$$\begin{array}{rl} A = \begin{pmatrix} Z_1(A) \\ \vdots \\ Z_m(A) \end{pmatrix} =: & (Z_1(A), \ldots, Z_m(A))' \\ \underset{\text{für } i\text{-te Zeile und } \lambda}{\overset{\text{(a)}}{\longmapsto}} & (Z_1(A), \ldots, \lambda Z_i(A), \ldots, Z_j(A), \ldots, Z_m(A))' \\ \underset{i\text{-te Zeile auf } j\text{-te Zeile}}{\overset{\text{(b)}}{\longmapsto}} & (Z_1(A), \ldots, \lambda Z_i(A), \ldots, Z_j(A) + \lambda Z_i(A), \ldots)' \\ \underset{\text{für } i\text{-te Zeile und } \lambda^{-1}}{\overset{\text{(a)}}{\longmapsto}} & (Z_1(A), \ldots, Z_i(A), \ldots, Z_j(A) + \lambda Z_i(A), \ldots, Z_m(A))' \end{array}$$

Weiter ergibt sich (d) durch die folgenden elementaren Umformungen:

$$\begin{aligned}
A = \quad & (Z_1(A), \ldots, Z_i(A), \ldots, Z_j(A), \ldots, Z_m(A))' \\
\underset{i\text{-te Zeile auf } j\text{-te Zeile}}{\overset{(b)}{\mapsto}} \quad & (Z_1(A), \ldots, Z_i(A), \ldots, Z_j(A) + Z_i(A), \ldots, Z_m(A))' \\
\underset{\text{für } i\text{-te Zeile und } -1}{\overset{(a)}{\mapsto}} \quad & (\ldots, -Z_i(A), \ldots, Z_j(A) + Z_i(A), \ldots)' \\
\underset{j\text{-te Zeile auf } i\text{-te Zeile}}{\overset{(b)}{\mapsto}} \quad & (\ldots, Z_j(A), \ldots, Z_j(A) + Z_i(A), \ldots)' \\
\underset{\text{für } i\text{-te Zeile und } -1}{\overset{(a)}{\mapsto}} \quad & (\ldots, -Z_j(A), \ldots, Z_j(A) + Z_i(A), \ldots,)' \\
\underset{i\text{-te Zeile auf } j\text{-te Zeile}}{\overset{(b)}{\mapsto}} \quad & (\ldots, Z_j(A), \ldots, Z_i(A), \ldots)' \,.
\end{aligned}$$

Definition 11.15 Alle durch Iteration von elementaren Umformungen (bzw. Zeilen-Umformungen bzw. Spalten-Umformungen) erhaltenen Operationen nennen wir **Umformungen** (bzw. **Zeilenumformungen** bzw. **Spaltenumformungen**.

Satz 11.16 Seien $A, B \in M(m \times n, K)$ und $b, c \in K^m$. Geht $(B \mid c)$ aus $(A \mid b)$ durch *Zeilenumformungen* hervor, so ist

$$L(A, b) = L(B, c)\,,$$

also die Lösungsmengen von $Ax = b$ und $Bx = c$ gleich.

Beweis: Es genügt, dies für *elementare* Zeilenumformungen zu beweisen.

Typ (a):
Wird die i-te Zeile

$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i \tag{11.16.1}$$

mit $\lambda \in K, \lambda \neq 0$, multipliziert, also durch

$$\lambda a_{i1}x_1 + \ldots + a_{in}x_n = \lambda b_i \tag{11.16.2}$$

ersetzt, so sind die Bedingungen äquivalent: Löst x (11.16.1), so offenbar auch (11.16.2), und die Umkehrung gilt auch, indem man mit λ^{-1} multipliziert.

Typ (b): Seien $k, \ell \in \{1, \ldots, m\}$, $k \neq \ell$, und es entstehe $(B \mid c)$ aus $(A \mid b)$, indem man die ℓ-te Zeile auf die k-te Zeile aufaddiert. Ist nun $x \in K^n$ eine Lösung von $Ax = b$, so gilt

$$a_{i1}x_1 + \ldots + a_{in}x_n = b_i$$

für alle $i \in \{1, \ldots, m\}$. Dann gilt für jedes $\lambda \in K$ auch

$$(a_{k1} + \lambda a_{\ell 1})x_1 + \ldots + (a_{kn} + \lambda a_{\ell n})x_n = b_k + \lambda b_\ell$$

also auch $Bx = c$ (Fall $\lambda = 1$). Es gilt also $L(A, b) \subseteq L(B, c)$. Da umgekehrt $(A \mid b)$ aus $(B \mid c)$ entsteht, indem von der k-ten Zeile von $(B \mid c)$ die ℓ-te Zeile von $(B \mid c)$ abgezogen wird (Fall $\lambda = -1$), so gilt auch $L(B, c) \subseteq L(A, b)$.

Satz 11.17 Bei einer Umformung einer Matrix ändern sich Spaltenrang und Zeilenrang nicht.

Beweis: Sei $A \in M(m \times n, K)$.

1) Es gehe B aus A durch Zeilenumformungen hervor. Nach Satz 11.16 gilt

$$\ker(\varphi_A) = L(A, 0) = L(B, 0) = \ker(\varphi_B)\,,$$

und mit dem Rangsatz gilt

$$\text{rg } A = n - \dim\ker(\varphi_A) = n - \dim\ker(\varphi_B) = \text{rg } B\,.$$

2) Nun gehe B aus A durch Spaltenumformungen hervor. Wir behaupten, dass ebenfalls der Rang gleich bleibt.
Es genügt, dies für elementare Spaltenumformungen zu zeigen.

Typ (a): Sei $\lambda \in K \setminus \{0\}$. Erhalten wir B, indem wir die j-te Spalte $S_j(A)$ mit λ multiplizieren, so ist

$$\begin{aligned} \text{im }(\varphi_A) &= \; < S_1(A), \ldots, S_n(A) >_K \; = \; < S_1(A), \ldots, \lambda S_j(A), \ldots, S_n(A) >_K \\ &= \; < S_1(B), \ldots, S_n(B) >_K \; = \; \text{im }(\varphi_B)\,, \end{aligned}$$

also rg A = rg B.

Typ (b): Seien $i, j \in \{1, \ldots, n\}$, $i \neq j$. Erhalten wir B, indem wir die j-te Spalte $S_j(A)$ auf die i-te Spalte $S_i(A)$ aufaddieren, so ist wegen

$$< S_1(A), \ldots, S_i(A) + S_j(A), \ldots, S_j(A), \ldots, S_n(A) >_K \; = \; < S_1(A), \ldots, S_n(A) >_K$$

ebenfalls rg B = rg A.

Bei Umformungen ändert sich also der Rang nicht, nach Satz 11.8 also auch nicht der *Spaltenrang*.

3) Für den *Zeilenrang* betrachten wir die folgende Konstruktion:

Definition 11.18 Für eine Matrix $A = (a_{ij}) \in M(m \times n, K)$ wird die **transponierte Matrix** definiert als

$$A^t = (a_{ji}) \in M(n \times m, K)\,.$$

Beispiel 11.19 Die transponierte Matrix zu

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in M(2 \times 3, K)$$

ist

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in M(3 \times 2, K)\,.$$

Die Zeilen (bzw. Spalten) von A werden also gerade zu den Spalten (bzw. Zeilen) von A^t.

Insbesondere ist also der Zeilenrang von A gleich dem Spaltenrang von A^t. Entsteht weiter B durch Umformung aus A, so entsteht B^t ebenfalls durch Umformung aus A^t. Damit folgt nun:

$$\begin{aligned} \text{Zeilenrang } (A) &= \text{Spaltenrang } (A^t) \\ &\overset{1),2)}{=} \text{Spaltenrang } (B^t) = \text{Zeilenrang } (B). \end{aligned}$$

Damit bleibt auch der Zeilenrang bei einer Umformung erhalten und Satz 11.17 ist bewiesen.

Wir kommen nun zum Haupthilfsmittel bei der konkreten Behandlung von linearen Gleichungssystemen.

Satz 11.20 (Gauß'sches Eliminationsverfahren) Sei $A \in M(m \times n, K)$. Dann lässt sich A durch Zeilenumformungen auf die folgende **Zeilenstufenform** bringen:

$$\tilde{A} = \begin{array}{c} \begin{array}{cccccccccccccccccc} & & & j_1 & & & & j_2 & & & & j_3 & & & j_k & & & \end{array} \\ \begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots & * & 0 & * & & 0 & * & \dots & * \\ \vdots & & \vdots & & & & & 1 & * & \dots & * & 0 & * & & \vdots & \vdots & & \vdots \\ & & & & & & & & & & & 1 & * & \dots & & & & \\ & & & & & & & & & & & & & & \vdots & \vdots & & \vdots \\ & & & & & & & & & & & & & & 0 & & & \\ & & & & & & & & & 0 & & & & & 1 & * & \dots & * \\ & & & & & & & & & & & & & & & & & \\ \vdots & & \vdots & & & & & & & & & & & & & & & \\ 0 & \dots & 0 & & & & & & & & & & & & & & & \end{pmatrix} \end{array}$$

wobei $*$ für irgendwelche Einträge steht.
Dabei gilt
(a) rg $A = k =$ Zeilenrang von A.

(b) Die Spaltenvektoren $S_{j_1}(A), \dots, S_{i_k}(A)$ bilden eine Basis von $\mathrm{im}(A) := \mathrm{im}\,(\varphi_A)$. (Achtung: die Spaltenvektoren von $\tilde{A}$ haben mit $\mathrm{im}(A)$ im Allgemeinen nichts zu tun!)

(c) $\tilde{A}$ ist eindeutig bestimmt.

Beweis Möglichkeit der Umformung: Sei $A = (a_{ij})$. Sind alle $a_{ij} = 0$, so ist die Behauptung richtig. Sei $S_{j_1}(A)$ die erste Spalte von A, die nicht null ist. Dann gibt es ein i mit $a_{ij_1} \neq 0$. Durch Zeilenvertauschung (Umformung von Typ (d)) lässt sich a_{ij_1} in die erste Zeile bringen, also können wir annehmen, dass $a_{1j_1} \neq 0$. Nach Multiplikation der 1. Zeile mit $\lambda = a_{1j_1}^{-1}$ können wir dann annehmen, dass $a_{1j_1} = 1$. Nun können wir die Spalte unter a_{1j_1} wie folgt "ausräumen":

Ist $i \geq 2$ und $a_{ij_1} = 0$, so ändern wir die i-te Zeile nicht; ist $a_{ij_1} \neq 0$, so subtrahieren von der i-ten Zeile die mit a_{ij_1} multiplizierte 1. Zeile. Hierdurch wird also $a_{ij_1} = 0$ für alle $i \geq 2$. Dann hat A die Gestalt

$$
\begin{array}{c}
\begin{array}{cccc} & j_1 & & \end{array} \\
A = \left(\begin{array}{cccc}
 & 1 & * \; \dots & * \\
 & 0 & & \\
 & \vdots & & \\
0 & & A_1 & \\
 & \vdots & & \\
 & 0 & &
\end{array}\right)
\end{array}
$$

Nun können wir die Teilmatrix A_1 genauso wie vorher A behandeln, wobei wir die Zeilentransformation für die gesamte Matrix A ausgeführt denken, wobei sich aber an den Stellen außerhalb von A_1 nichts ändert. Das heißt: Induktiv (z.B. Induktion über die Spaltenzahl) können wir annehmen, dass wir A_1 durch solche Zeilentransformationen auf Zeilenstufenform bringen können, wodurch dann auch A insgesamt Zeilenstufenform erhält, außer dass vielleicht in den Spalten über den $a_{\nu j_\nu} (\nu = 1, \dots, k)$ noch keine Nullen stehen. Diese Spalten können wir aber oberhalb von $a_{\nu j_\nu}$ wie vorher unterhalb $a_{\nu j_\nu}$ ausräumen, wodurch sich jeweils in den Spalten links davon nichts mehr ändert.

Wir zeigen nun die restlichen Aussagen:

(a): Es ist nach Satz 11.17

$$\text{rg } A = \text{ rg } \tilde{A}, \quad \text{Zeilenrang } (A) = \text{ Zeilenrang } (\tilde{A})\,.$$

Weiter gilt rg $\tilde{A} \geq k$, da die Spalten $S_{j_1}(\tilde{A}) = e_1,\ S_{j_2}(\tilde{A}) = e_2, \dots, S_{j_k}(\tilde{A}) = e_k$ linear unabhängig sind. Es ist aber auch rg $\tilde{A} \leq k$, da nur die ersten k Zeilen $\neq 0$ sind, so dass $\text{im}(\tilde{A}) \subseteq <e_1, \dots, e_k>_K \subseteq K^n$ ist. Schließlich ist offensichtlich der Zeilenrang von $\tilde{A}$ gleich k: Es gibt nur k Zeilen ungleich null, und diese sind linear unabhängig.

(b): Die Spaltenvektoren $S_{j_1}(\tilde{A}), \dots, S_{j_k}(\tilde{A})$ sind linear unabhängig, und die $(m \times k)$-Matrix $(S_{j_1}(\tilde{A}), \dots, S_{j_k}(\tilde{A}))$ geht aus der $(n \times k)$-Matrix $(S_{j_1}(A), \dots, S_{j_k}(A))$ durch Zeilentransformation hervor. Daher haben beide Matrizen denselben Rang, nämlich k, wie man an $S_{j_1}(\tilde{A}), \dots, S_{j_k}(\tilde{A})$ sieht. Also sind auch die Spalten $S_{j_1}(A), \dots, S_{j_k}(A)$ linear unabhängig, wegen dim im(A) = rg $A = k$ also auch eine Basis von $\text{im}(A) = < S_1(A), \dots, S_n(A) >_K$.

(c): Sei $U =< Z_1(A)^t, \dots, Z_m(A)^t >_K \subseteq K^n$ der von den Zeilen von A erzeugte Unterraum, und sei

$$\begin{array}{rcl} p_j : K^n & \to & K^j \\ (x_1, \dots, x_n) & \mapsto & (x_1, \dots, x_j) \end{array}$$

die Projektion auf die ersten j Koordinaten. Da auch $U =< Z_1(\tilde{A})^t, \dots, Z_m(\tilde{A})^t >_K$ ist, sind die Stufen $j_1, \dots, j_k$ genau an den Stellen $j \in \{1, \dots, n\}$, wo

$$\dim p_j(U) > \dim p_{j-1}(U)$$

(wobei wir $P_0(U) := 0$ setzen). Daher sind $j_1, \dots, j_k$ eindeutig bestimmt. Ist nun $\tilde{\tilde{A}}$ eine weitere Zeilenstufenmatrix, die aus A durch Zeilentransformation hervorgeht, so sind wegen $U =< Z_1(\tilde{A})^t, \dots, Z_m(\tilde{A})^t >_K =< Z_1(\tilde{\tilde{A}})^t, \dots, Z_m(\tilde{\tilde{A}})^t >_K$ auch die Stufen von $\tilde{\tilde{A}}$ an den Stellen $j_1, \dots, j_k$.

Insbesondere sind $\{Z_1(\tilde{\tilde{A}})^t, \ldots, Z_k(\tilde{\tilde{A}})^t\}$ und $\{Z_1(\tilde{\tilde{A}})^t, \ldots, Z_k(\tilde{\tilde{A}})^t\}$ beides Basen von U ($k =$ Zeilenrang $(A) =$ Zeilenrang $(\tilde{A}) =$ Zeilenrang $(\tilde{\tilde{A}})$; Zeilen von $\tilde{A}$ und $\tilde{\tilde{A}}$ sind 0 für $i > k$). Daher gibt es $\lambda_{\mu\nu} \in K$, $\mu, \nu \in \{1, \ldots, k\}$, mit

$$Z_\nu(\tilde{\tilde{A}})^t = \sum_{\mu=1}^{k} \lambda_{\mu\nu} Z_\mu(\tilde{A})^t\,. \tag{11.20.1}$$

Es folgt für $\alpha \in \{1, \ldots, k\}$

$$\delta_{\nu\alpha} = (Z_\nu(\tilde{\tilde{A}})^t)_{j_\alpha} \quad = \quad \sum_{\mu=1}^{k} \lambda_{\mu\nu}(Z_\mu(\tilde{\tilde{A}})^t)_{j_\alpha} = \sum_{\mu=1}^{k} \lambda_{\mu\nu}\delta_{\mu\alpha} = \lambda_{\alpha\nu}\,,$$

also $\lambda_{\mu\nu} = \delta_{\mu\nu}$ für alle μ, ν. Mit (11.20.1) folgt $Z_\nu(\tilde{\tilde{A}}) = Z_\nu(\tilde{A})$ für alle ν, also $\tilde{\tilde{A}} = \tilde{A}$.

Corollar 11.21 Für jede Matrix A gilt

$$\text{rg } A = \text{ Spaltenrang von } A = \text{ Zeilenrang von } A\,.$$

Dies folgt sofort aus 11.20 (a) (und 11.8).

Corollar 11.22 Sei $A \in M(m \times n, K)$. Dann lässt sich A durch Umformungen auf die Gestalt

$$D_r = \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & 0 & & \\ & & & 1 & & & & \\ & & & & 0 & & & \\ & & & & & 0 & & \\ & & 0 & & & & \ddots & \\ & & & & & & & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow r-\text{te Zeile} \\ \\ \\ \\ \\ \end{matrix}$$

bringen, also $D_r = (d_{ij})$ mit

$$d_{ij} = \begin{cases} \delta_{ij} & , \quad i, j \le r, \\ 0 & , \quad \text{sonst.} \end{cases}$$

Es ist dann $r = \text{rg } A$.

Beweis Nach dem Gaußschen Eliminationsverfahren lässt sich A durch Zeilenumformungen auf Zeilenstufenform bringen:

$$
\begin{array}{c}
\begin{array}{ccccccccccccccc}
 & & & j_1 & & j_2 & & & j_3 & & j_4 & & j_k & & \\
\end{array}\\
\left(\begin{array}{ccccccccccccccc}
0 & \dots & 0 & 1 & *\dots* & 0 & * & * & 0 & *\dots* & 0 & & * & \dots & * \\
 & & & & & 1 & *\dots* & & 0 & * & 0 & & \vdots & & \vdots \\
 & & & & & & & & 1 & *\dots* & 0 & & & & \\
 & & & & & & & & & & 1 & \dots & & & \\
 & & & & & & & & & & & & \vdots & & \vdots \\
 & & & & & & & & & & & & 0 & & \\
 & & & & & & & & & & & & 1 & * \quad \dots & * \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & &
\end{array}\right)
\end{array}
\begin{array}{l}
\\ \\ \\ \\ \\ \\ \leftarrow k\text{-te Zeile.} \\ \\ \\ \\ \\
\end{array}
$$

Durch Spaltenumformungen können wir in jeder der ersten k Zeilen die Stellen neben der ersten 1 ausräumen, d.h., zu null machen. Jetzt sind nur noch die Spalten $S_{j_1}, \dots, S_{j_k}$ ungleich null, und sind gleich den Einheitsvektoren $e_1, \dots, e_k$. Durch Spaltenvertauschungen können wir diese nun nacheinander in die Spalten 1 bis k bekommen, wodurch wir die Form D_r mit $r = k = \text{rg } A$ erhalten.

Beispiele 11.23 (a) Wir berechnen den Rang der 3×4-Matrix

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

durch Umformungen.

1) 3. Zeile -1. Zeile

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

2) 3. Zeile -2. Zeile

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

3) 1. Zeile $\cdot \frac{1}{3}$

$$\begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

4) Durch Ausräumen können wir jetzt die Zeilen rechts neben den Einsen ausräumen (das

geht immer, wenn man schon eine Form

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

erreicht hat). Explizit zum Beispiel:

(i) 2. Spalte - $\frac{2}{3}\cdot$ 1. Spalte

$$\begin{pmatrix} 1 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(ii) 3. Spalte - $\frac{1}{3}\cdot$ 1. Spalte, 4. Spalte - $\frac{2}{3}\cdot$ 1. Spalte

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

(iiii) 3. Spalte - 2. Spalte

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

Es ist also rg $A = 2$.

(b) Damit ist insbesondere rg $A < 3$ und A (d.h., φ_A) nicht surjektiv. Das Gleichungssystem

$$Ax = b$$

ist also nicht für alle $b \in \mathbb{R}^3$ lösbar. Wir untersuchen die Lösbarkeit für

$$b = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} .$$

Wir berechnen den Rang der erweiterten Matrix

$$(A \mid b) = \left(\begin{array}{cccc|c} 3 & 2 & 1 & 2 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 3 & 3 & 2 & 2 & 0 \end{array}\right) ,$$

indem wir zum Beisiel diesselben Schritte machen:

1') 3. Zeile - 1. Zeile

$$\left(\begin{array}{cccc|c} 3 & 2 & 1 & 2 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1 \end{array}\right)$$

2') 3. Zeile - 2. Zeile

$$\left(\begin{array}{cccc|c} 3 & 2 & 1 & 2 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -2 \end{array}\right)$$

3') 1. Zeile $\cdot \frac{1}{3}$

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

4') 3. Zeile $\cdot \left(-\frac{1}{2}\right)$

$$\left(\begin{array}{cccc|c} 1 & \frac{1}{3} & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

5') Vertauschen von 3. und 5. Spalte

$$\left(\begin{array}{cccc|c} 1 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right)$$

6') Ausräumen der Zeilen

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

Damit ist rg $(A \mid b) = 3 > 2 = \mathrm{rg} A$. Das Gleichungssystem ist also **nicht** lösbar.

(c) Jetzt lösen wir das Gleichungssystem

$$Ax = b$$

für

$$b = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

mit dem Gaußschen Eliminationsverfahren. Wir benutzen die *Zeilen*umformungen 1') - 4') von (b), weil wir dann schon wissen, was für A herauskommt – wir brauchen also nur noch b betrachten:

$$(A \mid b) = \left(\begin{array}{cccc|c} 3 & 2 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 3 & 3 & 2 & 2 & 5 \end{array}\right)$$

1) 3. Zeile - 1. Zeile

$$\left(\begin{array}{cccc|c} 3 & 2 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 0 & 3 \end{array}\right)$$

2) 3. Zeile - 2. Zeile

$$\left(\begin{array}{cccc|c} 3 & 2 & 1 & 2 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

3) 1. Zeile $\cdot\frac{1}{3}$

$$\left(\begin{array}{cccc|c} 1 & \frac{2}{3} & \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

4) 1. Zeile $-\frac{2}{3}\cdot$ 2. Zeile

$$\left(\begin{array}{cccc|c} 1 & 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{4}{3} \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{11.23.1}$$

Dies ist die Zeilenstufenform. Sie entspricht dem Gleichungssystem

$$\begin{array}{ccccccc} x_1 & & - & \frac{1}{3}x_3 & +\frac{2}{3}x_4 & = & -\frac{4}{3} \\ & x_2 & + & x_3 & & = & 3\,. \end{array}$$

Wir erhalten eine spezielle Lösung x, indem wir zum Beispiel $x_3 = x_4 = 0$ setzen; durch Auflösen der Gleichungen nach x_1 und x_2 erhalten wir dann $x_2 = 3$, $x_1 = -\frac{4}{3}$, also die spezeille Lösung

$$\begin{pmatrix} -\frac{4}{3} \\ 3 \\ 0 \\ 0 \end{pmatrix}.$$

Tatsächlich ist

$$A\begin{pmatrix} -\frac{4}{3} \\ 3 \\ 0 \\ 0 \end{pmatrix} = -\frac{4}{3}\begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} + 3\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} = b\,.$$

Um die gesamte Lösungsmenge $L(A,b)$ zu bestimmen, müssen wir noch den Lösungsraum der homogenen Gleichung $Ax = 0$ beschreiben. Wegen $A \in M(3 \times 4, \mathbb{R})$ und $\mathrm{rg}A = 2$ ist $\dim L(A,0) = 4 - 2 = 2$; der Lösungsraum ist also 2-dimensional. Wir können eine Basis bestimmen, indem wir an Stelle von (11.23.1) das homogene System

$$\left(\begin{array}{cccc|c} 1 & 0 & -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{11.23.2}$$

betrachten. Ausgeschrieben lautet es

$$\begin{array}{cccccc} x_1 & & - & \frac{1}{3}x_3 & +\frac{2}{3}x_4 & = 0 \\ & x_2 & + & x_3 & & = 0 \end{array}$$

Wir können x_3 und x_4 wählen, und nach dieser Wahl sind x_1 und x_2 bestimmt. Eine Basis erhalten wir, wenn wir für $\begin{pmatrix} x_3 \\ x_4 \end{pmatrix}$ einmal $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und einmal $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ nehmen. Dies ergibt die Lösungen

$$v_1 = \begin{pmatrix} \frac{1}{3} \\ -1 \\ 1 \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} -\frac{2}{3} \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

Es ist also

$$L(A,b) = \begin{pmatrix} -\frac{4}{3} \\ 3 \\ 0 \\ 0 \end{pmatrix} + < v_1, v_2 >_{\mathbb{R}} = \left\{ \begin{pmatrix} -\frac{4}{3} \\ 3 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} \frac{1}{3} \\ -1 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -\frac{2}{3} \\ 0 \\ 0 \\ 1 \end{pmatrix} \,\middle|\, \lambda, \mu \in \mathbb{R} \right\} .$$

§12 Konkrete Verfahren

12.1 Wir diskutieren noch einmal die Berechnung des Lösungsraums $L(A,b)$ eines linearen Gleichungssystems

$$Ax = b$$

mit dem Gauß'schen Eliminationsverfahren.

1) Bringe $(A \mid b)$ durch *Zeilen*transformation auf Zeilenstufenform

$$(\overline{A} \mid \overline{b}) = \begin{array}{c} \begin{array}{cccccccc} j_1 & & j_2 & & & j_k & & \end{array} \\ \left(\begin{array}{cccccccc|c} 1 & * & 0 & * & & 0 & & & \overline{b}_1 \\ & & 1 & * & \dots & \vdots & * & & \vdots \\ & & & & & \vdots & * & & \vdots \\ & & & & & 1 & & & \overline{b}_k \\ & & & & & & & & \overline{b}_{k+1} \\ & & & & & & & & \vdots \\ & & & & & & & & \overline{b}_m \end{array} \right) \end{array}$$

Wir wissen: $L(A,b) = L(\overline{A} \mid \overline{b})$.

2) Lösbarkeit und spezielle Lösung: Ist eine der Zahlen $\overline{b}_{k+1}, \dots, \overline{b}_m \neq 0$, so ist das System *nicht* lösbar (für $\overline{A}x$ sind immer die letzten $m-k$ Komponenten null). Ist $\overline{b}_{k+1} = \dots = \overline{b}_m = 0$, so ist das System lösbar, zum Beispiel durch

$$\begin{array}{rcll} x_{j_\nu} & = & \overline{b}_\nu & , \quad \nu = 1, \dots, k, \\ x_j & = & 0 & , \quad j \notin \{j_1, \dots, j_k\} . \end{array}$$

Dies ist also eine *spezielle Lösung.*

3) Lösungsraum: Wir müssen noch $L(A,0)$ bestimmen, wobei $L(A,0) = L(\overline{A},0)$ ist (es ist $\overline{0} = 0$, da alle Zeilentransformationen wieder auf den Nullvektor führen).

Seien $r_1 < r_2 < \ldots < r_{n-k}$ die Indizes in $\{1,\ldots,n\} \smallsetminus \{j_1,\ldots,j_k\}$ (die Indizes der "nicht-speziellen" Spalten von $\overline{A}$), und sei

$$B = (S_{r_1}(\overline{A}), S_{r_2}(\overline{A}), \ldots, S_{r_{n-k}}(\overline{A}))$$

die Matrix, die aus $\overline{A}$ durch Streichen der "speziellen" Spalten $S_{j_1}(\overline{A}),\ldots,S_{j_k}(\overline{A})$ entsteht. Dann gilt die Äquivalenz

$$0 = \overline{A}x = \begin{pmatrix} E_k \\ \hline 0_{m-k,k} \end{pmatrix} \begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_k} \end{pmatrix} + B \begin{pmatrix} x_{r_1} \\ \vdots \\ x_{r_{n-k}} \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} x_{j_k} \\ \vdots \\ x_{j_k} \end{pmatrix} = -\overline{B} \begin{pmatrix} x_{r_1} \\ \vdots \\ x_{r_{n-k}} \end{pmatrix}$$

wobei $0_{m-k,k}$ für die $(m-k) \times k$-Nullmatrix steht, und $\overline{B}$ aus B durch Streichen der letzten $m-k$ Zeilen entsteht (die alle 0 sind).

Dies ist für beliebiges $\begin{pmatrix} x_{r_1} \\ \vdots \\ x_{r_{n-k}} \end{pmatrix} \in K^{n-k}$ lösbar. Eine Basis des Lösungsraums $L(A,0)$ erhält man, indem man hier z.B. die Basisvektoren $e_1,\ldots,e_{n-k}$ in K^{n-k} wählt.

Basisvektoren sind also $v_1,\ldots,v_{n-k}$; mit

$$(v_\nu)_j = \begin{cases} -(S_\nu(\overline{B}))_j & , \quad j = j_\alpha \quad (j \text{ "speziell"}) \\ \delta_{\nu,\beta} & , \quad j = r_\beta \quad (j \text{ "nicht speziell"})\,. \end{cases}$$

Etwas übersichtlicher wird dies nach einer Spaltenvertauschung ($\widehat{=}$ Umnummerierung der Variablen!) in $\overline{A}$, so dass wir die Gestalt

$$\overline{\overline{A}} = \left(\begin{array}{ccc|c} 1 & & & \\ & \ddots & & \overline{B} \\ & & 1 & \\ \hline & & 0 & \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 1\end{array}}\right\} k \\ \left.\vphantom{0}\right\} n-k \end{array}$$

erhalten. Eine Basis von $L(\overline{\overline{A}},0)$ ist dann

$$\begin{pmatrix} -S_1(\overline{B}) \\ \hline e_1 \end{pmatrix}, \ldots, \begin{pmatrix} -S_{n-k}(\overline{B}) \\ \hline e_{n-k} \end{pmatrix} \begin{array}{l} \}\, k \\ \}\, n-k \end{array} \quad .$$

Das Gauß'sche Eliminationsverfahren kann auch zum Invertieren von Matrizen benutzt werden:

Konstruktion 12.2 Sei $A \in M_n(K)$ eine reguläre (invertierbare) $n \times n$-Matrix. Um die inverse Matrix $B = A^{-1}$ von A zu bestimmen, müssen wir die Gleichung

$$A \cdot B = E$$

lösen ($E = E_n$ die $n \times n$-Einheitsmatrix).

Für die Spaltenvektoren $v_j = S_j(B)$ von B bedeutet dies wegen $A \cdot B = (Av_1, \ldots, Av_n)$ und $S_j(E) = e_j$ gerade

$$Av_j = e_j \qquad (j = 1, \ldots, n)\,. \tag{12.2.1}$$

Wir können also diese linearen Gleichungssysteme lösen (die Lösung ist eindeutig, vergleiche unten), und es ist dann $B = (v_1, \ldots, v_n)$, also die Matrix, die die Spalten $v_1, \ldots, v_n$ hat.

Umgekehrt ist $A \in M_n(K)$ genau dann regulär, wenn all die Gleichungen (12.2.1) lösbar sind.

Die Gleichungssysteme (12.2.1) können wir mit dem Eliminationsverfahren auf Lösbarkeit überprüfen beziehungsweise lösen. Dies können wir nach dem Schema 12.1 machen, und zwar für alle e_j gleichzeitig.

Beispiel 12.3 Wir bestimmen die Inverse der reellen Matrix

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix}\,.$$

Wir starten mit der "mehrfach erweiterten" Matrix $(A \mid e_1 \mid e_2 \mid e_3)$:

$$\begin{pmatrix} 1 & 1 & 2 & & 1 & 0 & 0 \\ 3 & 2 & 1 & & 0 & 1 & 0 \\ 2 & 1 & 1 & & 0 & 0 & 1 \end{pmatrix}\,.$$

Ausräumung der 1. Spalte (2. Zeile - $3 \cdot 1$. Zeile, 3. Zeile - $2 \cdot 1$. Zeile) liefert

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

2. Zeile $\cdot(-1)$ und 2. Spalte ausräumen:

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

3. Zeile $\cdot\frac{1}{2}$ und 3. Spalte ausräumen

$$\begin{pmatrix} 1 & 0 & 0 & & -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ 0 & 1 & 0 & & \frac{1}{2} & \frac{3}{2} & -\frac{5}{2} \\ 0 & 0 & 1 & & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}\,.$$

Damit ist

$$A^{-1} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & \frac{3}{2} & -\frac{5}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 & -1 & 3 \\ 1 & 3 & -5 \\ 1 & -1 & 1 \end{pmatrix}\,,$$

denn die Lösung eines Gleichungssystems

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ \vdots \\ b_n \end{pmatrix},$$

also $Ex = b$, ist

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

also $x = b$, wegen $Ex = x$.

Damit stehen rechts bereits die Spalten der Matrix $B = A^{-1}$, d.h., die gesamte Matrix $B = A^{-1}$. Nachrechnen (wichtig!):

$$A \cdot B = \frac{1}{2} \begin{pmatrix} 1 & 1 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & -1 & 3 \\ 1 & 3 & -5 \\ 1 & -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = E\,.$$

Wir notieren noch die folgenden zwei nützlichen Beobachtungen

Lemma 12.4 Sei $A \in M(m \times n, K)$ und $b \in K^m$. Sei das Gleichungssystem $Ax = b$ lösbar. Dann sind äquivalent

(i) Die Lösung ist eindeutig.

(ii) $L(A, 0) = \{0\}$.

(iii) $\operatorname{rg} A = n$.

Beweis Sei $v_0 \in K^m$ eine spezielle Lösung. Dann ist $L(A, b) = v_0 + L(A, 0)$ (siehe 11.2). Dies zeigt die Äquivalenz der ersten beiden Bedingungen. Weiter gilt $L(A, 0) = \{0\}$ genau dann wenn $\dim L(A, 0) = 0$, und dies ist nach den Rangsatz (bzw. 11.7) äquivalent zu $n - \operatorname{rg} A = 0$, also zu (iii).

Lemma 12.5 Für $A \in M(m \times n, K)$ sind äquivalent:

(i) $Ax = b$ ist für alle $b \in K^m$ lösbar.

(ii) $\operatorname{rg} A = m$.

Beweis: (i) $\Leftrightarrow$ A (d.h., φ_A) ist surjektiv $\Leftrightarrow$ $\operatorname{rg} A = \dim \operatorname{im}(\varphi_A) = \dim K^m = m$.

Schließlich ziehen wir noch eine Verbindung zwischen Matrizen und dem Thema von Abschnitt 8 (Erzeugendensysteme, lineare Unabhängigkeit, Basen).

Seien $v_1, \ldots, v_n$ Vektoren in K^m (hier sind m, n beliebig in $\mathbb{N}_0$). Bilde die Matrix

$$A = (v_1, \ldots, v_n) \in M(m \times n, K)$$

mit Spalten $v_1, \ldots, v_n$. Dann ist nach dem Beweis von 11.8 $\operatorname{im} A = <v_1, \ldots, v_n>_K$, also

$$\operatorname{rg} A = \dim < v_1, \ldots, v_n >_K \;.$$

Hieraus folgt

Lemma 12.6 Es gilt

(i) $(v_1, \ldots, v_n)$ Erzeugendensystem von $K^m \Leftrightarrow \operatorname{rg} A = m$.

(ii) $(v_1, \ldots, v_n)$ linear unabhängig $\Leftrightarrow \operatorname{rg} A = n$.

(iii) $(v_1, \ldots, v_n)$ Basis von $K^m \Leftrightarrow \operatorname{rg} A = m = n$.

Beweis (i) ist unmittelbar klar nach der Vorbemerkung, und (ii) folgt aus 11.8 ($\operatorname{rg} A =$ Spaltenrang (A)). (iii) folgt wiederum aus (i) und (ii).

Bemerkung 12.7 Das Gauß'sche Eliminationsverfahren (Teil 11.20 (b)) , angewandt auf A, kann benutzt werden, um unter $v_1, \ldots, v_n$ eine Basis von $W = <(v_1, \ldots, v_n)>_K$ zu finden. So können wir also für jeden Untervektorraum $W \subseteq K^m$ und jedes Erzeugendensystem $v_1, \ldots, v_n$ von W eine Basis von W unter den v_i finden.

§13 Die Determinante

Motivation: Betrachte 2×2-Matrizen

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

• Definiere $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.

Dann gilt, wie man leicht nachrechnet:

• A invertierbar $\Leftrightarrow \det A \neq 0$

• A invertierbar $\Rightarrow A^{-1} = \frac{1}{\det A} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

• $\det(A \cdot B) = \det A \cdot \det B$

Verallgemeinere dies auf beliebige (quadratische Matrizen). Bei 3×3-Matrizen:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a\,e\,i + \ldots + \ldots - \ldots - \ldots - c\,e\,g\,,$$

also 6 Terme. Bei 4×4-Matrizen: 24 Terme! Es ist besser, dies konzeptioneller zu betrachten.

Sei K ein Körper (z. B. $\mathbb{R}$ oder $\mathbb{C}$).

Satz 13.1 Es gibt genau eine Abbildung

$$\det : M_n(K) \to K$$

mit den folgenden Eigenschaften

(i) det ist linear in jeder Spalte, d.h., für jedes $j = 1, \ldots, n$ ist

$$\det(v_1, \ldots, \alpha v_j + \beta v_j', \ldots, v_n) = \alpha \det(v_1, \ldots, v_j, \ldots, v_n) + \beta \det(v_1, \ldots, v_j', \ldots, v_n)$$

(hierbei sind $v_1, \ldots, v_n, v_j'$ beliebige (Spalten-)Vektoren in K^n und $\alpha, \beta \in K$ beliebig);

(ii) Ist der (Spalten-)Rang von A kleiner als n, so ist $\det A = 0$;

(iii) $\det E = 1$.

Diese eindeutig bestimmte Abbildung det heißt die **Determinante**; $\det A$ heißt die **Determinante von** A.

Bemerkung: Für eine $n \times n$-Matrix A gilt nach Lemma 10.20: A invertierbar $\Leftrightarrow$ $\mathrm{rg} A = n$ (denn offenbar ist A (d.h., φ_A) genau dann surjektiv, wenn $\mathrm{rg} A = n$). Bedingung (ii) bedeutet also

(ii) Ist A nicht invertierbar, so ist $\det A = 0$.

Wir führen den Beweis von 13.1 in mehreren Schritten.

Zur Eindeutigkeit:

Lemma 13.2 Sei $\det : M_n(K) \to K$ eine Abbildung mit der Eigenschaft 13.1 (i). Die Eigenschaft 13.1 (ii) ist dann äquivalent zu:
(ii') $\det(v_1, \ldots, v_n) = 0$, falls $v_i = v_j$ für ein Paar (i, j) mit $i \neq j$.

Beweis (ii) $\Rightarrow$ (ii'): Ist $v_i = v_j$ für i, j mit $i \neq j$, so ist der Rang von $(v_1, \ldots, v_n)$ kleiner als n, nach (ii) ist also $\det(v_1, \ldots, v_n) = 0$.

(ii') $\Rightarrow$ (ii): Ist der Rang von $A = (v_1, \ldots, v_n)$ kleiner als n, so sind $v_1, \ldots, v_n$ linear abhängig. Es gibt dann also ein $i \in \{1, \ldots, n\}$ mit

$$v_i = \sum_{j \neq i} \alpha_j v_j$$

für $\alpha_j \in k \quad (j \in \{1, \ldots, n\} \smallsetminus \{i\})$.

Dann ist nach 13.1(i)

$$\begin{aligned} \det(v_1, \ldots, v_n) &= \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_j \det(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_n) \\ &= 0 \text{ nach (ii')} \end{aligned}$$

Definition 13.3 Sei V ein K-Vektorraum, und sei $m \in \mathbb{N}$.

(a) Eine Abbildung
$$\phi : V^m \to K$$
heißt (m-)**multilinear** (oder m-**linear**), wenn sie linear in jedem Argument ist, d.h., wenn für jedes $j \in \{1, \dots, m\}$ gilt

$$\begin{aligned}&\phi(v_1, \dots, v_{j-1}, \alpha v_j + \beta v_j', v_{j+1}, \dots, v_m)\\ &= \alpha\phi(v_1, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_m) + \beta\phi(v_1, \dots, v_{j-1}, v_j', v_{j+1}, \dots, v_m)\end{aligned}$$

für alle $v_1, \dots, v_{j-1}, v_j, v_j', v_{j+1}, \dots, v_m \in V$ und alle $\alpha, \beta \in K$
(äquivalent: Für jedes $j \in \{1, \dots, m\}$ und alle $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m \in V$ ist die Abbildung
$$\begin{array}{rcl} V & \to & K \\ v & \mapsto & \phi(v_1, \dots, v_{j-1}, v, v_{j+1}, \dots, v_m) \end{array}$$
linear). Man nennt ϕ auch eine **Multilinearform**, oder m-**lineare Form**, oder eine m-**Form** auf V.

(b) Eine m-lineare Abbildung $\phi : V^m \to K$ heißt **alternierend**, wenn gilt $\phi(v_1, \dots, v_m) = 0$, falls $v_i = v_j$ für ein Paar (i, j) mit $i \neq j$.

Bemerkung 13.4 (i) Eine 1-lineare Form $\phi : V \to K$ ist gerade eine Linearform auf V, also ein Element aus dem Dualraum V^* von V.

(ii) Eine 2-lineare Abbildung
$$\phi : V \times V \to K$$
heißt auch bilineare Abbildung oder **Bilinearform auf** V. Zum Beispiel ist das Skalarprodukt
$$\begin{array}{rcl} \phi : \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ (x, y) & \mapsto & < x, y >= \sum\limits_{i=1}^{n} x_i y_i \end{array}$$
eine Bilinearform auf $\mathbb{R}^n$.

Definition 13.5 Eine m-lineare Form $\phi : V^m \to K$ heißt **symmetrisch** (bzw. **anti-symmetrisch**), wenn für alle $i < j$ gilt

$$\begin{array}{rcl} \phi(v_1, \dots, v_i, \dots, v_j, \dots, v_n) & = & \phi(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\ (\text{bzw.} & = & -\phi(v_1, \dots, v_j, \dots, v_i, \dots, v_n)) \, . \end{array}$$

Bemerkung 13.6 (i) Das Skalarprodukt auf $\mathbb{R}^n$ ist symmetrisch.

(ii) Eine alternierende m-lineare Form ist anti-symmetrisch: Für $i < j$ ist

$$\begin{array}{rcl} 0 & = & \phi(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_n) \\ & = & \phi(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + \phi(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \end{array}$$

da ϕ m-linear und alternierend ist.

(iii) Ist $1+1=0$ in K (man sagt hierzu, dass die Charakteristik von K gleich 2 ist: char $K=2$), so sind die anti-symmetrischen Formen gleich den symmetrischen ($-1=+1!$), und nicht notwendig alternierend(Beispiel ?).

Ist char $K \neq 2$ (also $2 = 1+1 \neq 0$), so sind die alternierenden m-linearen Formen gleich den anti-symmetrischen:

$$\begin{array}{lll} \phi \text{ anti-symmetrisch} & \Rightarrow & 2\cdot\phi(v_1,\ldots,v_i,\ldots,v_i,\ldots,v_m)=0 \\ & \Rightarrow & \phi(v_1,\ldots,v_i,\ldots,v_i,\ldots,v_m)=0\ . \end{array}$$

Proposition 13.7 Sei V ein n-dimensionaler K-Vektorraum, und sei $b=(b_1,\ldots,b_n)$ eine Basis von V. Dann gibt es höchstens eine alternierende n-lineare Form

$$\phi : V^n \to K$$

mit $\phi(b_1,\ldots,b_n)=1$.

Beweis Seien ϕ,ϕ' zwei solche Formen, und seien $v_1,\ldots,v_n \in V$. Dann gibt es $a_{ij}\in K$ mit

$$v_i = \sum_{j=1}^n a_{ij}b_j, \quad i=1,\ldots,n\,,$$

und es ist

$$\begin{aligned} \phi(v_1,\ldots,v_n) &= \phi(\sum_{j_1=1}^n a_{1j_1}b_{j_1},\ldots,\sum_{j_n=1}^n a_{1j_n}b_{j_n}) \\ &= \sum_{(j_1,\ldots,j_n)\in\{0,\ldots,n\}^n} a_{1j_1}\ldots a_{nj_n}\phi(b_{j_1},\ldots,b_{j_n})\,, \end{aligned}$$

entsprechend für ϕ'. Also ist ϕ durch die Werte auf den Tupeln $(b_{j_1},\ldots,b_{j_n})$ bestimmt. Es ist aber $\phi(b_{j_1},\ldots,b_{j_n})=0$, falls dasselbe b_i zweimal vorkommt, da ϕ alternierend ist; dasselbe gilt für ϕ'. Wir können also annehmen, dass die $b_{j_1},\ldots,b_{j_n}$ paarweise verschieden sind. Da ϕ auch antisymmetrisch ist (13.6 (ii)), ist für $k<\ell$

$$\phi(b_{j_1},\ldots,b_{j_k},\ldots,b_{j_\ell},\ldots,b_{j_n}) = -\phi(b_{j_1},\ldots,b_{j_\ell},\ldots,b_{j_k},\ldots,b_{j_n})\,,$$

entsprechend für ϕ', und durch endlich viele solche Vertauschungen kann man das Tupel $(b_{j_1},\ldots,b_{j_n})$ auf die Gestalt $(b_1,\ldots,b_n)$ bringen. Damit sind ϕ und ϕ' durch ihren Wert auf $(b_1,\ldots,b_n)$ bestimmt, also gleich.

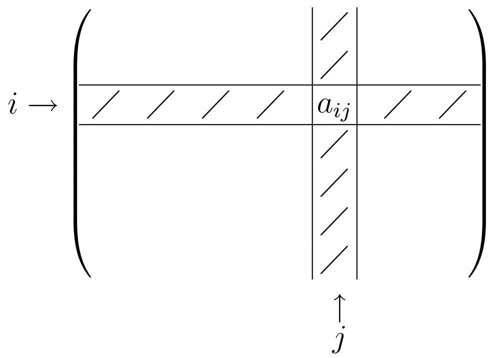
Hieraus folgt die Eindeutigkeit von ϕ in Satz 13.1: eine Abbildung $\phi: M_n(K)\to K$ mit 13.1 (i) und (ii) ist gerade eine alternierende n-Form auf K^n und nach 13.1 (iii) soll gelten $\phi(e_1,\ldots,e_n)=1$ für die Standardbasis $(e_1,\ldots,e_n)$ von K^n.

Existenz von ϕ:
Dies beweisen wir durch Induktion über n. Für $n=1$ ist nichts zu zeigen: es ist $\det(a)=a$.

Sei nun $n>1$.

Definition 13.8 Sei $A\in M_n(K)$. Für $i,j\in\{1,\ldots,n\}$ bezeichne $A_{ij}\in M_{n-1}(K)$ die

Matrix, die aus A durch Weglassen der i-ten Zeile und der j-ten Spalte entsteht.



Fixiere nun ein $i \in \{1, \ldots, n\}$. Dann definieren wir rekursiv

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} ,$$

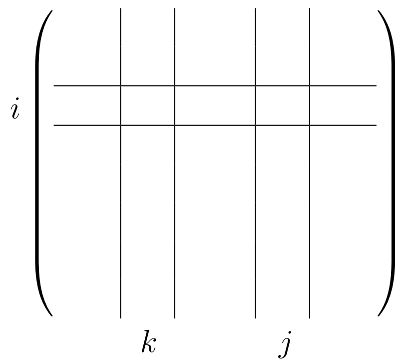
wobei $\det A_{ij}$ nach Induktionsannahme bereits definiert ist.

Wir zeigen nun, dass die entstehende Abbildung

$$\det : M_n(K) \rightarrow K$$

die Eigenschaften 13.1 (i)-(iii) hat.

(i): Wir zeigen, dass jeder Summand $(-1)^{i+j} a_{ij} \det A_{ij}$ linear in der k-ten Spalte ist (k beliebig). Für $k \neq j$ folgt, dies daraus, dass $\det A_{ij}$ linear in den Spalten ist und a_{ij} nicht von der k-ten Spalte abhängt:



Für $k = j$ hängt $\det A_{ij}$ nicht von der k-ten Spalte ab (diese wurde gerade weggelassen), aber $a_{ij} = a_{ik}$ (d.h., die Abbildung

$$\begin{array}{rcl} M_n(K) & \rightarrow & K \\ A & \mapsto & a_{ik} \end{array}$$

ist linear in der k-ten Spalte von A).

(ii) Wir zeigen die Eigenschaft (ii') aus 13.2. Sei die k-te Spalte v_k von A gleich der ℓ-ten Spalte v_ℓ von A, $k < \ell$. Dann ist

$$\det A = (-1)^{i+k} a_{ik} \det A_{ik} + (-1)^{i+\ell} a_{i\ell} \det A_{i\ell} ,$$

denn für $j \neq k, \ell$ ist $\det A_{ij} = 0$ nach Induktionsvoraussetzung, da A_{ij} zwei gleiche Spalten hat. Sei zunächst $\ell = k+1$ (benachbarte Spalten). Dann ist offenbar $A_{ik} = A_{i\ell}, a_{ik} = a_{i\ell}$, und damit

$$\begin{aligned} \det A &= (-1)^{i+k} a_{ik} \det A_{ik} + (-1)^{i+k+1} a_{ik} \det A_{ik} \\ &= 0\,. \end{aligned}$$

Hieraus ergibt sich weiter wie in Bemerkung 13.6 (ii), dass

$$\det(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n) = -\det(v_1, \ldots, v_{k+1}, v_k, \ldots, v_n)\,;$$

d.h., Vertauschung benachbarter Spalten führt zu Vorzeichenwechsel. Ist nun $k < \ell$ beliebig, so kann man durch fortlaufendes Vertauschen von benachbarten Spalten – wobei sich nur das Vorzeichen der Determinante ändern kann – erreichen, dass zwei benachbarte Spalten gleich sind; es ist also auch in diesem Fall $\det A = 0$.

(iii):
$$\begin{aligned} \det E &= \sum_{j=1}^{n} (-1)^{i+j} \delta_{ij} \det E_{ij} \\ &= (-1)^{2i} \det E_{ii} = 1 \end{aligned}$$

nach Induktionsvoraussetzung.

Damit ist die Determinante (mit den Eigenschaften von 13.1) definiert.

Definition 13.9 Die Berechnungsformel

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det A_{ij}$$

(für $i \in \{1, \ldots, n\}$ beliebig, aber fest) heißt die **Entwicklung** der Determinante **nach der i-ten Zeile** ("Laplace'sche Entwicklungsformel")

Das Schema der Vorzeichen $(-1)^{i+j}$ ist einfach; es sieht so aus

$$\begin{pmatrix} + & - & + & - & + & - & \\ - & + & - & + & & & \\ + & - & + & - & & & \\ & \vdots & & & & \ddots & \\ & \vdots & & & & & \end{pmatrix}.$$

Beispiele 13.10 (a) Für eine 2×2-Matrix

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

ergibt Entwicklung nach der 1. Zeile

$$\det A = (-1)^0 \cdot a \cdot d + (-1)^1 b \cdot c = ad - bc\,,$$

wie anfangs angegeben.

(b) (Regel von Sarrus) Für eine 3×3-Matrix

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

ergibt Entwicklung nach der 1. Zeile

$$\begin{aligned} \det A &= a \cdot (ei - fh) - b(di - fg) + c(dh - eg) \\ &= a \cdot e \cdot i + b \cdot f \cdot g + c \cdot d \cdot h - c \cdot e \cdot g - b \cdot d \cdot i - a \cdot f \cdot h \end{aligned}$$

Diese Formel läßt sich wie folgt verbildlichen

$$\left(\begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix}\right) \begin{matrix} a & b \\ d & e \\ g & h \end{matrix}$$

bzw.

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

(im Geiste immer die Diagonalen zu 3er-Reihen ergänzen)

Da es nicht darauf ankommt, kann man eine beliebige Zeile zur Entwicklung wählen; möglichst eine mit vielen Nullen:

Beispiel 13.11 Sei

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

Entwicklung nach der 3. Zeile:

$$\begin{aligned} \det A &= \det A_{31} - \det A_{34} = \det \begin{pmatrix} 2 & 3 & 4 \\ 3 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} - \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \\ 0 & 1 & 1 \end{pmatrix} \\ &= (3 + 12 - 8 - 2) - (3 + 12 - 8 - 2) = 0\,. \end{aligned}$$

($\underset{\text{s.u.}}{\Rightarrow} A$ ist nicht regulär)

Satz 13.12 Für eine Matrix $A = (a_{ij}) \in M_n(K)$ sei $A^t = (a_{ij})$ die transponierte Matrix. Dann gilt

$$\det A^t = \det A\,.$$

Beweis Betrachte die Abbildung

$$\begin{array}{rrcl} \widetilde{\det}: & M_n(K) & \to & K \\ & A & \mapsto & \det A^t\,. \end{array}$$

$\widetilde{\det}$ erfüllt 13.1 (iii), denn es ist $\widetilde{\det}\, E = \det E^t = \det E = 1$.

$\widetilde{\det}$ erfüllr 13.1 (ii): $\mathrm{rg} A < n \Rightarrow \mathrm{rg} A^t = \mathrm{rg} A < n \Rightarrow \widetilde{\det}\ A = \det A^t = 0$.

$\widetilde{\det}$ erfüllt 13.1 (i): die Spalten von A werden die Zeilen von A^t; wir haben also zu zeigen:

Lemma 13.13 det ist auch linear in den Zeilen.

Beweis: Für die Abhängigkeit von der k-ten Zeile betrachten wir die Formel (Entwicklung nach der k-ten Zeile)

$$\det A = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj}\,.$$

Hier hängt A_{kj} nicht von der k-ten Zeile ab (diese wurde gerade gestrichen) und $a_k j$ hängt offenbar linear von der k-ten Zeile ab, für alle j. Dies zeigt die Behauptung. □

Ende des Beweises von 13.12: Da $\widetilde{\det}$ 13.1.(i) – (iii) erfüllt, ist $\widetilde{\det} = \det$. □

Corollar 13.14 Für jedes $j \in \{1, \dots, n\}$ ist auch

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

(**Entwicklung nach der j-ten Spalte**).

Beweis Dies folgt durch Anwendung von 13.9 auf A^t, zusammen mit 13.12.

Satz 13.15 Sei $A \in M_n(K)$ eine obere Dreiecksmatrix, also

$$A = \begin{pmatrix} a_{11} & & & & \\ & a_{22} & & * & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & a_{nn} \end{pmatrix}$$

(hier steht $*$ für beliebige Einträge oberhalb der Diagonalen). Dann ist

$$\det A = a_{11} \cdot a_{22} \dots a_{nn} = \prod_{i=1}^{n} a_{ii}\,.$$

Entsprechendes gilt für untere Dreieckmatrizen

$$B = \begin{pmatrix} b_{11} & & & & \\ & \ddots & & 0 & \\ & & \ddots & & \\ & * & & \ddots & \\ & & & & b_{nn} \end{pmatrix},$$

auch hier ist $\det B = \prod_{i=1}^{n} b_{ii}$.

Beweis Durch Induktion über n, wobei der Fall $n = 1$ trivial ist.

Der Induktionsschnitt folgt durch Entwicklung nach der 1. Spalte: es ist

$$\begin{aligned}
\det A &= a_{11} \cdot \det \begin{pmatrix} a_{22} & & & \\ & \ddots & & * \\ & & \ddots & \\ & 0 & & \ddots \\ & & & & a_{nn} \end{pmatrix} \\
&= a_{11} \cdot \prod_{i=2}^{n} a_{ii} \quad \text{nach Induktionsvoraussetzung}
\end{aligned}$$

Der Fall von unteren Dreiecksmatrizen folgt analog (Entwicklung nach der 1. Zeile), oder durch Betrachten der Transponierten.

Satz 13.16 (Multiplikativität der Determinante) Für $A, B \in M_n(K)$ gilt

$$\det(A \cdot B) = \det A \cdot \det B\,.$$

Beweis Fixiere A und betrachte die Abbildung

$$\begin{aligned}
\phi : M_n(K) &\to K \\
B &\mapsto \det AB\,.
\end{aligned}$$

ϕ erfüllt 13.1 (i), d.h., ist linear in den Spalten:
Für

$$B = (\underbrace{v_1, \dots, v_n}_{\text{Spalten von } B}) \quad \text{ist} \quad AB = (\underbrace{Av_1, \dots, Av_n}_{\text{Spalten von } AB})\,.$$

Für feste $v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n$ ist also die Abbildung

$$\begin{aligned}
K^n &\to K \\
v &\mapsto \phi(v_1, \dots, v_{j-1}, v, v_{j+1}, \dots, v_n) \\
&\quad = \det(Av_1 \dots, Av_{j-1}, Av, Av_{j+1}, \dots, Av_n)
\end{aligned}$$

linear, als Komposition der linearen Abbildungen

$$\begin{array}{ccccc}
K^n & \to & K^n & \to & K \\
v & \mapsto & Av & & \\
& & w & \mapsto & \det(Av_1, \dots, Av_{j-1}, w, Av_{j+1}, \dots, Av_n)
\end{array}$$

(Linearität von det in der j-ten Spalte).

ϕ erfüllt auch 13.2 (ii'): ist $v_i = v_j$ für $i \neq j$, so ist $Av_i = Av_j$ und damit $\det(Av_1, \dots, Av_n) = 0$.

ϕ erfüllt im allgemeinen nicht 13.1 (iii), aber es ist

$$\phi(E) = \det(A)\,.$$

Ist nun $\det(A) = 0$, so ist

$$\det - \phi : M_n(K) \to K$$

eine Abbildung, die 13.1 (i) – (iii) erfüllt, also

$$\det - \phi = \det\,,$$

d.h., es ist ϕ die Nullabbildung und damit

$$\det(A \cdot B) \;=\; 0 \;=\; \det A \cdot \det B \qquad \forall\, B\,.$$

Ist $\det A \neq 0$, so ist

$$\frac{1}{\det A} \cdot \phi : M_n(K) \to K$$

eine Abbildung die 13.1 (i) – (iii) erfüllt, also

$$\frac{1}{\det A} \cdot \phi = \det\,,$$

also

$$\frac{1}{\det A} \cdot \det AB = \det B \qquad \forall\, B$$

q.e.d

Corollar 13.17 Eine Matrix $A \in M_n(K)$ ist genau dann regulär ($\Leftrightarrow A$ hat den maximalen Rang $n \Leftrightarrow A$ ist invertierbar), wenn $\det A \neq 0$.

Beweis Nach 13.1 (ii) gilt

$$A \text{ nicht regulär } \Rightarrow \det A = 0$$

Nach 13.16 gilt aber auch:

$$\begin{aligned} A \text{ regulär } &\Rightarrow\; A \text{ besitzt ein Inverses } A^{-1} \Rightarrow A \cdot A^{-1} = E \\ &\Rightarrow\; 1 = \det E = \det A \cdot A^{-1} = \det A \cdot \det(A^{-1})\,. \end{aligned}$$

Also muss gelten $\det A \neq 0$.

Corollar 13.18 Ist A regulär, so gilt

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

Dies folgt aus dem Beweis von 13.17.

Die Determinate erlaubt eine Berechnung der Inversen einer regulären Matrix:

Definition 13.19 Für $A \in M_n(K)$ definiere die **komplementäre Matrix** $\tilde{A} = (\widetilde{a_{ij}})$ durch

$$\widetilde{a_{ij}} := (-1)^{i+j} \det A_{ji}\,.$$

Beachte: hier steht A_{ji}, nicht A_{ij}!

Es ist also

$$\tilde{A} = \begin{pmatrix} \det A_{11} & -\det A_{21} & \det A_{31} & \\ -\det A_{12} & \det A_{22} & & \\ \det A_{13} & -\det A_{23} & \dots & \\ & & & \end{pmatrix}$$

Satz 13.20 Ist $A \in M_n(K)$, so ist

$$A \cdot \tilde{A} = \det A \cdot E\,.$$

Insbesondere gilt, falls A regulär ist (und damit $\det A \neq 0$)

$$A^{-1} = \frac{1}{\det A} \cdot \tilde{A}\,.$$

Beweis Sei

$$A \cdot \tilde{A} =: (c_{ij})$$

Dann ist

$$\begin{aligned} c_{ik} &= \sum_{j=1}^{n} a_{ij} \cdot \widetilde{a_{jk}} \\ &= \sum_{j=1}^{n} a_{ij} \cdot (-1)^{j+k} \det A_{kj}\,. \end{aligned}$$

Für $i = k$ ist (Entwicklung nach der i-ten Zeile)

$$c_{ii} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} = \det A\,.$$

Für $i \neq k$ behaupten wir, dass $c_{ik} = 0$ ist. Betrachte nämlich die Matrix $A' = (a'_{\mu\nu})$, die aus A entsteht, indem man die k-te Zeile durch die i-te Zeile von A ersetzt. Dann gilt (Entwicklung nach der k-ten Zeile, A' hat 2 gleiche Zeilen)

$$\begin{aligned} 0 = \det A' &= \sum_{j=1}^{n} \cdot (-1)^{k+j} a'_{kj} \det A'_{kj} \\ &= \sum_{j=1}^{n} \cdot (-1)^{k+j} a_{ij} \det A_{kj} \\ &= c_{ik}\,, \end{aligned}$$

denn die Matizen A_{kj} hängen nicht von der k-ten Zeile von A ab (diese wurde gestrichen). Insgesamt ergibt sich

$$A \cdot \tilde{A} = \begin{pmatrix} \det A & & & \\ & \ddots & 0 & \\ & & & \\ & 0 & \ddots & \\ & & & \det A \end{pmatrix} = \det A \cdot E\,.$$

Ist nun A regulär, so gilt $A \cdot \left(\frac{1}{\det A} \cdot \tilde{A}\right) = E$, also $A^{-1} = \frac{1}{\det A} \cdot \tilde{A}$. . q.e.d.

Beispiele 13.21 (a) Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(K)$ ist

$$\tilde{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} .$$

Ist $\det A = ad - bc \neq 0$, so ist also $A \in Gl_2(K)$ und

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

(b) Für

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

ist

$$\det A = 6 + 6 - 2 - 2 = 8$$

$$\tilde{A} = \begin{pmatrix} -2 & 4 & 2 \\ 2 & -8 & 2 \\ 2 & 4 & -2 \end{pmatrix}$$

Es ist

$$A \cdot \tilde{A} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} -2 & 4 & 2 \\ 2 & -8 & 2 \\ 2 & 4 & -2 \end{pmatrix} = \begin{pmatrix} -2+4+6 & 4-16+12 & 2+4-6 \\ -2+2 & 4+4 & 2-2 \\ -6+4+2 & 12-16+4 & 6+4-2 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} .$$

Die Determinante erlaubt auch das Lösen eines Gleichungssystems

$$(*) \qquad Ax = b$$

mit quadratischer Matrix $A \in M_n(K)$ (also, $x, b \in K^n$), falls A **invertierbar** ist. Wir erinnern uns, dass in diesem Fall das System $(*)$ eindeutig lösbar ist (für jeses $b \in K^n$).

Corollar 13.22 (Cramersche Regel) Sei $A \in M_n(K)$ invertierbar, und seien $v_1, \ldots, v_n$ die Spaltenvektoren von A. Dann wird für $b \in K^n$ das System

$$Ax = b$$

gelöst durch den Vektor

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{mit} \quad x_i = \frac{\det(v_1, \ldots, v_{i-1}, b, v_i, \ldots, v_n)}{\det A}$$

für $i = 1, \ldots, n$ (Beachte, dass $\det A = \det(v_1, \ldots v_n)$).

Beweis Es ist
$$x = A^{-1}b = \frac{1}{\det A}\tilde{A}b\,,$$
also
$$\begin{aligned} x_i &= \frac{1}{\det A}\cdot\sum_{j=1}^{n}\tilde{A}_{ij}bj \\ &= \frac{1}{\det A}\cdot\sum_{j=1}^{n}(-1)^{i+j}b_j\cdot\det A_{ji}\,, \end{aligned}$$
und $\sum\limits_{j=1}^{n}(-1)^{i+j}b_j\det A_{ji}$ ist die Entwicklung von $\det(v_1,\dots,v_{i-1},b,v_{i+1},\dots,v_n)$ nach der i-ten Spalte.

Beispiel 13.23 Wir lösen
$$\begin{pmatrix}1&2&3\\1&0&1\\3&2&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}5\\3\\7\end{pmatrix}.$$
Die Determinante der Matrix ist 8 (siehe 13.21 (b)), und damit folgt
$$\begin{aligned} x_1 &= \tfrac{1}{8}\det\begin{pmatrix}5&2&3\\3&0&1\\7&2&1\end{pmatrix} = \tfrac{1}{8}(14+18-6-10) = \tfrac{16}{2} = 2 \\ x_2 &= \tfrac{1}{8}\det\begin{pmatrix}1&5&3\\1&3&1\\3&7&1\end{pmatrix} = \tfrac{1}{8}(3+15+21-27) = \tfrac{1}{8}\,0 = 0 \end{aligned}$$
$$x_3 = \ldots = 1$$
also
$$x = \begin{pmatrix}2\\0\\1\end{pmatrix}$$
Nachrechnen!

Beispiel 13.24 (a) Manchmal werden 13.20 und 13.22 beide Cramersche Regeln genannt (erste und zweite); wichtiger ist 13.20.

(b) Manchmal wird die Determinante einer Matrix A auch mit $|A|$ bezeichnet, und für eine Matrix
$$\begin{pmatrix}a_{11}&\dots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\dots&a_{nn}\end{pmatrix}\text{ schreibt man }\begin{vmatrix}a_{11}&\dots&a_{1n}\\\vdots&&\vdots\\a_{n1}&\dots&a_{nn}\end{vmatrix}\text{ für die Determinante.}$$

Bemerkung 13.25 (Rechenregeln für Determinanten) Wir wissen schon

(a) $\det(v_1,\dots,\lambda v_i,\dots,v_n) = \lambda\cdot\det(v_1,\dots,v_i,\dots,v_n)$

(b) $\det(v_1, \ldots, v_i + v_i', \ldots, v_n) = \det(v_1, \ldots, v_i, \ldots, v_n) + \det(v_1, \ldots, v_i', \ldots, v_n)$

((a) + (b) $\Leftrightarrow$ Linearität in der i-ten Spalte), sowie

(c) $\det(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\det(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$

(vergleiche 13.6 (iii))

(a), (b) und (c) beschreiben insbesondere das Verhalten bei elementaren Spaltenumformungen vom Typ 11.13 (a), (b) und 11.14 (c). Für Typ 11.14 (d) gilt

(d) Für $i \neq j$ und $\lambda \in K$ ist

$$\det(v_1, \ldots, v_i + \lambda v_j, \ldots, v_n) = \det(v_1, \ldots, v_i, \ldots, v_n)$$

Denn sei ohne Einschränkung $i < j$; dann gilt wegen (a) und (b)

$$\begin{aligned} & \det(v_1, \ldots, v_i + \lambda v_j, \ldots, v_n) \\ = \;& \det(v_1, \ldots, v_n) + \lambda \det(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n) \\ = \;& \det(v_1, \ldots, v_n) \end{aligned}$$

da $\det(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n) = 0$ wegen 13.2.

Entsprechendes gilt für Zeilen (wegen $\det A = \det A^t$).

Dies erlaubt die Berechnung der Determinante durch Zeilen- und Spaltenumformungen. Z.B.:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{vmatrix} = 0 \quad \text{, und}$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 0 & 0 & -1 \end{vmatrix} = -1 \cdot \begin{vmatrix} 1 & 2 \\ 3 & 3 \end{vmatrix} = 3\,.$$

§14 Permutationen und Leibnizformel

Erinnerung: Für $n \in \mathbb{N}$ war die **symmetrische Gruppe** S_n die Menge aller Bijektionen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$; diese werden Permutationen genannt (und S_n auch die Permutationsgruppe). Die Ordung von S_n ist $n!$.

Definition 14.1 Eine **Transposition** ist eine Permutation $\tau \in S_n$, die zwei Zahlen $i \neq j$ vertauscht und alle anderen Zahlen festläßt; diese Abbildung wird mit $(i\,j)$ bezeichnet.

Offenbar gilt $(i\,j)^2 = \mathrm{id}$.

Satz 14.2 Es gibt genau einen Homomorphismus

$$\mathrm{sign}\; : S_n \to \{\pm 1\}$$

mit $\mathrm{sign}(\tau) = -1$ für jede Transposition (Die Verknüpfung auf $\{\pm 1\}$ ist die Multiplikation). Die Abbildung sign heißt das **Signum**; für $\sigma \in S_n$ heißt $\mathrm{sign}(\sigma)$ das **Signum von** σ.

Beweis:
Eindeutigkeit: folgt aus

Lemma 14.3 Jede Permutation $\sigma \in S_n$ ist Produkt von Transpositionen.

Beweis: Durch Induktion über n, wobei der Fall $n = 1$ trivial ist. Wir haben einen Monomorphismus

$$\begin{array}{rcl} \varphi : S_{n-1} & \hookrightarrow & S_n \\ \varphi & \mapsto & \tilde{\varphi} : \{1, \ldots, n\} \to \{1, \ldots, n\} \\ & & \tilde{\varphi}(i) = \left\{ \begin{array}{ccc} \varphi(i) & , & i \leq n-1 \\ n & , & i = n \end{array} \right. \end{array}$$

der Transpositionen auf Transpositionen abbildet. Nach Induktionsannahme ist die Behauptung richtig in S_{n-1}, also auch in $\mathrm{im}\, \varphi = \{\sigma \in S_n \mid \sigma(n) = n\}$. Sei nun $\sigma \in S_n$. Ist $\sigma(n) = n$, so sind wir fertig. Sonst sei $\sigma(n) = i < n$. Dann ist $\sigma' = (i\, n) \cdot \sigma \in \mathrm{im}\, \varphi$, also Produkt von Transpositionen, und dasselbe gilt für $(i\, n)\sigma' = (i\, n)^2 \sigma = \sigma$.

Existenz in 14.2: Setze

$$\mathrm{sign}(\sigma) = \prod_{i\, j} \frac{j-i}{\sigma(j) - \sigma(i)}$$

Für $j \neq i$ ist auch $\sigma(j) \neq \sigma(i)$, also ist die rechte Seite wohldefiniert.

Erklärung: Es ist $\mathrm{sign}(\sigma) = (-1)^m$ mit $m = m(\sigma) =$ Anzahl der (i, j) mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$.

Denn es ist

$$\begin{aligned} \prod_{i<j} (\sigma(j) - \sigma(i)) &= \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}} (\sigma(j) - \sigma(i)) \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} (\sigma(j) - \sigma(i)) \\ &= (-1)^m \prod_{i<j} |\sigma(j) - \sigma(i)| = (-1)^m \prod_{i<j} (j-i)\,, \end{aligned}$$

da $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ eine Bijektion ist. Wir zeigen nun, dass sign ein Homomorphismus ist.

$$\begin{aligned} \mathrm{sign}(\sigma\tau) &= \prod_{i<j} \frac{j-i}{\sigma\tau(j) - \sigma\tau(i)} = \prod_{i<j} \left(\frac{j-i}{\tau(j) - \tau(i)} \; \frac{\tau(j) - \tau(i)}{\sigma\tau(j) - \sigma\tau(i)} \right) \\ &= \prod_{i<j} \frac{j-i}{\tau(j) - \tau(i)} \cdot \prod_{\substack{i<j \\ \tau(i)<\tau(j)}} \frac{\tau(j) - \tau(i)}{\sigma\tau(j) - \sigma\tau(i)} \cdot \prod_{\substack{i<j \\ \tau(j)<\tau(i)}} \frac{\tau(i) - \tau(j)}{\sigma(\tau(i)) - \sigma(\tau(j))} \\ &= \prod_{i<j} \frac{j-i}{\tau(j) - \tau(i)} \cdot \prod_{i'<j'} \frac{j' - i'}{\sigma(j') - \sigma(i')} = \mathrm{sign}(\tau) \cdot \mathrm{sign}(\sigma)\,. \end{aligned}$$

Es bleibt noch zu zeigen, dass $\mathrm{sign}(\tau) = -1$, wenn τ eine Transposition ist. Sei $\tau = (k\,\ell)$, ohne Einschränkung $k < \ell$. Die Paare (i, j) mit $i < j$ aber $\tau(i) > \tau(j)$ sind dann gerade

$$\begin{array}{lll} (k,\ell),(k+1,\ell) & ,\ldots, & (\ell-1,\ell) \\ (k,k+1),(k,k+1) & ,\ldots, & (k,\ell-1)\,. \end{array}$$

Ihre Anzahl ist $m(\tau) = \ell - 1 - (k-1) + \ell - 1 - k = 2(\ell - k) - 1$. Daher ist $\mathrm{sign}(\tau) = (-1)^{m(\tau)} = -1$.

Definition/Lemma 14.4 (a) Eine Permutation $\sigma \in S_n$ heißt **gerade**, wenn $\mathrm{sign}(\sigma) = 1$ ist, und **ungerade**, wenn $\mathrm{sign}(\sigma) = -1$.

(b) Die Menge $A_n = \{\sigma \in S_n \mid \mathrm{sign}(\sigma) = 1\}$ der geraden Permutationen ist eine Untergruppe und heißt die alternierende Gruppe.

Dass A_n eine Untergruppe ist, folgt sofort daraus, dass sign ein Gruppenhomomorphismus ist.

Satz 14.5 (Leibniz-Determinantenformel) Für $A = (a_{ij}) \in M_n(K)$ ist

$$\det A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)}\,.$$

Beweis Wir zeigen, dass die durch die rechte Seite definierte Abbildung

$$\widetilde{\det} : M_n(K) \to K, \quad \widetilde{\det} A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \cdot a_{2\sigma(2)} \ldots a_{n\sigma(n)}$$

die Bedingungen 13.1 (i) – (iii) erfüllt.

13.1 (i): Jeder Summand $\mathrm{sign}(\tau) a_{1\sigma(1)} \cdot a_{2\sigma(2)} \ldots a_{n\sigma(n)}$ ist offenbar linear in der k-ten Spalte.

13.1 (ii'): Sei $j \neq k$ und $S_j(A) = S_k(A)$ für die j-te und die k-te Spalte von A. Sei $\tau = (j\,k)$ die Transposition, die j und k vertauscht. Dann ist die Abbildung

$$\begin{array}{rcl} f : A_N & \to & \{\text{ungerade Permutation}\} \\ \sigma & \mapsto & \tau\sigma \end{array}$$

eine Bijektion (mit Umkehrabbildung $\sigma' \mapsto \tau\sigma'$), und daher

$$\widetilde{\det} A = \sum_{\sigma \in A_n} \mathrm{sign}(\sigma) a_{1\sigma(1)} \ldots a_{n\sigma(n)} + \sum_{\sigma \in A_n} \mathrm{sign}(\tau\sigma) a_{1\tau\sigma(1)} \ldots a_{n\tau\sigma(n)}\,.$$

Es ist aber für $\sigma \in A_N$

$$\mathrm{sign}(\tau\sigma) = \mathrm{sign}(\tau) \cdot \mathrm{sign}(\sigma) = -\mathrm{sign}(\sigma)$$

und

$$a_{1\sigma(1)} \ldots a_{n\sigma(n)} = a_{1\tau\sigma(1)} \ldots a_{n\tau\sigma(n)}\,,$$

denn für $\sigma(\nu) \neq j, k$ ist $\tau\sigma(\nu) = \sigma(\nu)$, und für $\sigma(\nu') = j$, $\sigma(\nu'') = k$ ist $a_{\nu'\sigma(\nu')} = a_{\nu' j} = a_{\nu' k} = a_{\nu'\tau\sigma(\nu')}$ und ebenso $a_{\nu''\sigma(\nu'')} = a_{\nu''\tau\sigma(\nu'')}$.

Es ist also $\widetilde{\det} A = 0$.

13.1(iii): für $A = E$ ist

$$\begin{aligned}\widetilde{\det} E &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma)\delta_{1\sigma(1)} \cdot \delta_{n\sigma(n)} \\ &= 1\,,\end{aligned}$$

da nur der Summand für $\sigma = \mathrm{id}$ ungleich 0 ist.

Beispiele 14.6 (a) Für $n = 2$ ist $S_2 = \{\mathrm{id}, (12)\}$ und sign(id)$= 1$, $\operatorname{sign}((12)) = -1$. Die Leibniz-Formel zeigt also

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

wie bekannt.

(b) Für paarweise verschiedene Zahlen $k_1, \ldots, k_r \in \{1, \ldots, n\}$ definiere den **Zyklus der Länge** r

$$(k_1, \ldots, k_r) \in S_n$$

als die Permutation $\sigma \in S_n$ mit

$$\sigma(k_1) = k_2,\ \sigma(k_2) = k_3, \ldots, \sigma(k_{r-1}) = k_r,\ \sigma(k_r) = k_1$$

und $\sigma(i) = i$ für $i \notin \{k_1, \ldots, k_r\}$. Dann kann man leicht durch Induktion beweisen:

$$\operatorname{sign}(\sigma) = (-1)^{r-1}\,.$$

(c) Für $n = 3$ besteht die Gruppe S_3 aus den Elementen id, $\sigma_1 = (2\,3), \sigma_2 = (1\,3), \sigma_3 = (1\,2), \varphi = (1\,2\,3)$ und $\varphi^2 = (1\,3\,2)$. Es ist

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} a_{11}\,a_{22}\,a_{33} + a_{12}\,a_{23}\,a_{31} + a_{13}\,a_{21}\,a_{31} \\ -a_{11}\,a_{23}\,a_{32} - a_{13}\,a_{22}\,a_{31} - a_{12}\,a_{21}\,a_{33} \end{array}$$

Dies ist die Regel von Sarrus. Allgemein erhalten wir für die Determinante einer $n \times n$-Matrix eine Summe von $n!$ Produkten.

§15 Die Determinante eines Endomorphismus

Sei K ein Körper.

Lemma 15.1 Für $A, B \in M_n(K), B$ invertierbar, ist

$$\det(B^{-1}AB) = \det A\,.$$

Beweis: Nach 13.16 und 13.18 ist $\det(B^{-1}AB) = \det(B^{-1}) \cdot A \cdot \det B = (\det B)^{-1} \cdot \det A \cdot \det B = \det A$.

Bemerkung 15.2 Die Operation $A \mapsto B^{-1}AB$ heißt **Konjugation mit** B. Lemma 15.1 sagt also, dass die Determinante **Konjugations-invariant** ist.

Lemma 15.3 Sei $f : V \to V$ eine lineare Abbildung (also ein Endomorphismus von V). Dann ist $\det f$, die **Determinante von** f, wie folgt definiert: Sei $b = (b_1, \ldots, b_n)$ eine Basis von V, und sei $A = M_b^b(f) \in M_n(K)$ die Matrix von f bezüglich b [Erinnerung: $A = (a_{ij})$ mit $f(b_j) = \sum\limits_{i=1}^{n} a_{ij} b_i (j = 1, \ldots, n)$]. Dann setze

$$\det f := \det A\,.$$

Dies ist wohldefiniert.

Beweis der Wohldefiniertheit: Wir haben zu zeigen, dass dies unabhängig von der gewählten Basis ist. Die Matrix A ist durch das folgende kommutative Diagramm definiert

$$\begin{tikzcd}
V \arrow[r, "f"] & V & \sum\limits_{i=1}^{n} x_i b_i & b_i \\
K^n \arrow[u, "\varphi_b", "\wr"'] \arrow[r, "A"] & K^n \arrow[u, "\varphi_b"', "\wr"] & (x_1, \ldots x_n) \arrow[u, mapsto] & e_i \arrow[u, mapsto]
\end{tikzcd}$$

Ist nun $c = (c_1, \ldots, c_n)$ eine andere Basis von V und ist $B = M_c^b (= M_c^b(\mathrm{id}))$ die zugehörige Basiswechsel-Matrix [j-te Spalte von B=Darstellung von c_j in der Basis b, d.h., $c_j = \sum\limits_{i=1}^{n} b_{ij} b_i$], so zeigt das kommutative Diagramm

$$\begin{tikzcd}
 & V \arrow[r, "f"] & V & \\
K^n \arrow[ur, "\varphi_c"] \arrow[r, "B"] & K^n \arrow[u, "\varphi_b"'] \arrow[r, "A"] & K^n \arrow[u, "\varphi_b"] & K^n \arrow[ul, "\varphi_c"'] \arrow[l, "B"']
\end{tikzcd}$$

dass die Matrix $A' = M_c^c(f)$ von f bezüglich c gleich $B^{-1}AB$ ist. (Dies ist die Formel $M_c^c(f) = M_b^c M_b^b(f) M_c^b$ aus Satz 10.28, zusammen mit der Formel $M_b^c = (M_c^b)^{-1}$ aus Lemma 10.27). Mit Lemma 3.1 folgt nun $\det M_c^c(f) = \det(B^{-1}AB) = \det(A) = \det M_b^b(f)$.

Aus den Eigenschaften von Determinanten folgt sofort

Proposition 15.4 Es gilt

(a) $\det f \neq 0 \quad \Leftrightarrow \quad f$ Isomorphismus

(b) $\det f \circ g = \det f \cdot \det g$

(c) $\det \mathrm{id} = 1$.

Beispiel 15.5 Betrachte $\mathbb{C}$ als $\mathbb{R}$-Vektorraum und die Abbildung (komplexe Konjugation)

$$\begin{array}{lrcl} c: & \mathbb{C} & \to & \mathbb{C} \\ & \alpha = a + bi & \mapsto & \overline{\alpha} = a - bi \end{array}$$

c ist $\mathbb{R}$-linear (nicht $\mathbb{C}$-linear!), und es ist $\det c = -1$: $\mathbb{R}$-Basis von $\mathbb{C}$ ist $(1, i)$, und es ist $c(1) = 1$, $c(i) = -i$; die Matrix von c in dieser $\mathbb{R}$-Basis ist also

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

mit Determinante -1.

§16 Polynome

Sei K ein Körper.

Definition 16.1 (naive Definition) Ein **Polynom** über K (oder mit Koeffizienten in K) ist ein Ausdruck

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n ,$$

mit $n \in \mathbb{N}_0$ und $a_0, a_1, \ldots, a_n \in K$. Hierbei heißt a_i der i-te Koeffizient von $f(x)$.

Wir vereinbaren noch, dass wir einen Ausdruck $0x^i$ nicht mitschreiben müssen. Also ist z.B. $1 + x^2 = 1 + 0x + x^2 + 0x^3$.

Definition 16.2 Für zwei Polynome $f(x) = \sum\limits_{i=0}^{m} a_i x^i$ und $g(x) = \sum\limits_{j=0}^{n} b_j x^j$ werde die Summe und das Produkt wie folgt definiert

$$f(x) + g(x) := \sum_{j=0}^{\max(m,n)} (a_j + b_j) x^j$$

(wobei wir $a_j = 0$ für $j > m$ und $b_j = 0$ für $j > n$ setzen),

$$f(x) \cdot g(x) = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i \cdot b_j \right) x^k .$$

Die Addition ist also einfach komponentenweise, und die Multiplikation ist gerade so gemacht, dass gilt

Proposition 16.3 Die Menge $K[x]$ der Polynome über K wird mit der obigen Addition und Multiplikation ein kommutativer Ring mit Eins.

Beweis: selbst! Die Multiplikation folgt aus dem Distributivgesetz, also durch "Ausmultiplizieren".

Definition 16.4 Für ein Polynom

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

mit $a_n \neq 0$ heißt n der Grad von f (Bez.: $\deg(f)$) und a_n der Leitkoeffizient von f (Bez.: $a_{\max}(f)$). Für das Nullpolynom 0 setze $\deg(0) := -\infty$ und $a_{\max}(0) := 0$. Nenne f normiert, wenn $a_{\max}(f) = 1$.

Lemma 16.5 (a) $\deg(f \cdot g) = \deg(f) + \deg(g)$

(b) $\det(f + g) \leq \max(\deg(f), \deg(g))$

(c) $a_{\max}(f \cdot g) = a_{\max}(f) \cdot a_{\max}(g)$.

Beweis: selbst!

Satz 16.7 (Polynomdivision/Division und Rest) Seien $f, g \in K[x]$, $g \neq 0$. Dann existieren eindeutig bestimmte Polynome $q, r \in K[x]$ mit

$$f = q \cdot g + r, \quad \deg(r) < \deg(g)$$

(Bedeutung: f geteilt durch g ist q Rest r).

Beweis: *Existenz*: Induktion nach $\deg(f)$. Für $f = 0$ setze $q = r = 0$.

Sei nun $f \neq 0$. Für $\deg(f) < \deg(g)$ setze $q = 0$ und $r = f$. Für $\deg(f) \geq \deg(g)$ sei

$$\begin{aligned} g &= b_n x^n + \ldots \quad \text{(niedrigere Potenzen von } x) \\ f &= a_{n+k} x^{n+k} + \ldots \quad \text{(niedrigere Potenzen von } x) \end{aligned}$$

mit $n, k \in \mathbb{N}_0$, $b_n \neq 0$ und $a_{n+k} \neq 0$. Dann hat

$$f - \frac{a_{n+k}}{b_n} \cdot x^k \cdot g$$

kleineren Grad als f. Nach Induktionsvoraussetzung gibt es $q_1, r \in K[x]$ mit $\deg(r) < \deg(g)$ und

$$f - \frac{a_{n+k}}{b_n} x^k \cdot g = q_1 \cdot g + r_1 \,.$$

Also ist $f = q \cdot g + r$, mit

$$q = \frac{a_{n+k}}{b_n} x^k + q_1 \in K[x] \,.$$

Eindeutigkeit: Sei $f = q_1 \cdot g + r_1 = q_2 \cdot g + r_2$ mit $q_i, r_i \in K[x]$ und $\deg(r_i) < \deg(g) \quad (i = 1, 2)$. Dann ist

$$(q_1 - q_2) g = r_1 - r_2 \,.$$

Wegen $\deg(r_1 - r_2) < \deg(g)$ folgt $(q_1 - q_2) = 0$ (vergleiche 16.5(a)), also auch $r_1 - r_2 = 0$.

Beispiel 16.8 Wir teilen mit Rest wie beim Teilen von ganzen Zahlen:

$$\begin{array}{l} x^3 - 1 \;: x + 1 = x^2 - x + 1 \text{ Rest } -2 \\ \underline{x^2 + x^2} \\ \quad -x^2 - 1 \\ \quad \underline{-x^2 - x} \\ \qquad\quad x - 1 \end{array}$$

$$\begin{array}{r} x+1 \\ \hline -2 \end{array}$$

also $x^3 - 1 = (x^2 - x + 1)(x + 1) - 2$.

Definition 16.9 Für $f, g \in K[x]$ sagen wir g teilt f (in Zeichen $g \mid f$), wenn es ein $q \in K[x]$ gibt mit

$$f = q \cdot g\,.$$

(Für $g \neq 0$ bedeutet dies also, dass beim Teilen von f durch g Rest 0 herauskommt).

Proposition 16.10 Sei $f \in K[x]$. Ist $\lambda \in K$ mit $f(\lambda) = 0$, so gibt es ein $f_1 \in K[x]$ mit

$$f = f_1 \cdot (x - \lambda)$$

(d.h., $(x - \lambda)$ teilt f).

Beweis Sei

$$f = q \cdot (x - \lambda) + r$$

mit $q, r \in K[x]$ und $\deg(r) < \deg(x - \lambda) = 1$. Angenommen, $r \neq 0$. Dann ist $\deg(r) = 0$, also $r = a$ mit $a \in K, a \neq 0$ (ein konstantes Polynom). Es folgt

$$0 = f(\lambda) = q(\lambda) \cdot (\lambda - \lambda) + r = a\,.$$

Widerspruch! Also ist $r = 0$ und wir können $f_1 = q$ nehmen.

Definition 16.11 Sei $f \in K[x]$. Ein Element $\lambda \in K$ heißt Nullstelle (oder Wurzel) von f, wenn

$$f(\lambda) = 0\,.$$

Corollar 16.12 Sei $f \in K[x]$ von Grad $n, f \neq 0$. Dann hat f höchstens n Nullstellen in k.

Beweis Durch Induktion über n, wobei der Fall $n = 0$ trivial ist.

Hat f keine Nullstelle, so sind wir fertig.

Hat f eine Nullstelle λ, so ist nach 16.10

$$f = f_1 \cdot (x - \lambda)$$

mit einem Polynom $f_1 \in K[x]$. Dann ist $f_1 \neq 0$ und $\deg(f_1) = n - 1$ (16.5(a)); nach Induktionsvoraussetzung hat f_1 also höchstens $n-1$ Nullstellen. Ist weiter μ eine Nullstelle von f mit $\mu \neq \lambda$, so ist

$$0 = f(\mu) = f_1(\mu) \cdot (\mu - \lambda) \quad \text{mit } \mu - \lambda \neq 0\,,$$

also μ auch eine Nullstelle von f_1. Dies zeigt die Behauptung.

Die folgende Tatsache wird in der Analysis bewiesen; wir werden sie hier benutzen.

Satz 16.13 ("Fundamentalsatz der Algebra") Jedes nicht konstante Polynom $f \in \mathbb{C}[x]$ besitzt eine Nullstelle.

Bemerkung 16.14 (a) Körper mit dieser Eigenschaft heißen **algebraisch abgeschlossen**.

(b) $\mathbb{R}$ ist nicht algebraisch abgeschlossen: das Polynom $x^2 + 1$ hat keine Nullstelle in $\mathbb{R}$.

(c) Man kann zeigen ($\curvearrowright$ Algebra): Jeder Körper K ist in einem algebraisch abgeschlossenen Körper K' enthalten. Es gibt einen "kleinsten", den "algebraischen Abschluß", der bis auf Isomorphie eindeutig ist. Der algebraische Abschluß von $\mathbb{R}$ ist $\mathbb{C}$.

Durch sukzessives Anwenden von 16.10 erhalten wir:

Corollar 16.15 Sei $f \in \mathbb{C}[x]$ von Grad n. Dann ist

$$f = a \cdot \prod_{i=1}^{n} (x - \lambda_i)$$

mit $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, wobei die λ_i die Nullstellen von f sind.

Ein λ_i kann mehrfach auftreten. Genauer kann man jedes Polynom $f \in \mathbb{C}[x]$ bis auf die Reihenfolge eindeutig in Linearfaktoren zerlegen:

$$f = a \cdot \prod_{i=1}^{r} (x - \alpha_i)^{n_i} .$$

Hierbei sind $\alpha_1, \ldots, \alpha_r$ die verschiedenen Nullstellen von f, $n_1, \ldots, n_r$ heißen ihre **Vielfachheiten**, und a ist der führende Koeffizient von f.

Die Eindeutigkeit – bis auf die Reihenfolge – folgt aus der Polynomdivision; weiter ist $\sum\limits_{i=1}^{r} n_i = n := \deg(f)$.

Dasselbe gilt offenbar für jeden algebraisch abgeschlossenen Körper K an Stelle von $\mathbb{C}$.

§17 Eigenwerte

Sei K ein Körper.

Definition 17.1 Sei $\varphi : V \to V$ ein Endomorphismus eines K-Vektorraums V. Eine Zahl $\lambda \in K$ heißt **Eigenwert von** φ, wenn es einen Vektor $v \neq 0$ (!) in V gibt mit

$$\varphi(v) = \lambda v .$$

Ein solches v heißt **Eigenvektor von** φ zum Eigenwert λ.

Definition 17.1′ Entsprechend definiert man Eigenwerte und Eigenvektoren von quadratischen Matrizen $A \in M_n(K)$ (Spezialfall $V = K^n$) : $v \in K^n \setminus \{0\}$ heißt **Eigenvektor** von A zum **Eigenwert** λ, wenn

$$Av = \lambda v .$$

Beispiele 17.2 (a) $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$ hat keinen Eigenwert (in $\mathbb{R}$):

$$Av = \lambda \cdot v \qquad \text{für } v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \neq 0 \qquad \Rightarrow \qquad \begin{array}{rcl} -v_2 & = & \lambda v_1 \\ v_1 & = & \lambda v_2 \end{array}$$

Es folgt $v_2 = -\lambda^2 v_2$ und $v_1 = -\lambda^2 v_1$, also $\lambda^2 = -1$ (da $v \neq 0$). Dies gilt aber für kein $\lambda \in \mathbb{R}$.

(b) Betrachte die Abbildung

$$\begin{array}{rcl} D : C^\infty(\mathbb{R}) & \to & C^\infty(\mathbb{R})\,, \\ f & \mapsto & f' \end{array}$$

wobei $C^\infty(\mathbb{R})$ die Menge der unendlich oft differenzierbaren Funktionen auf $\mathbb{R}$ ist. Dann ist D linear, und für $\lambda \in \mathbb{R}$ ist $f(x) = e^{\lambda x}$ Eigenvektor von D zum Eigenwert λ.

Lemma 17.3 (a) Ein Skalar $\lambda \in K$ ist genau dann ein Eigenwert der linearen Abbildung $\varphi : V \to V$, wenn $\ker(\varphi - \lambda \cdot \; \mathrm{id}) \neq 0$.

(b) Ein Vektor $v \neq 0$ ist genau dann Eigenvektor von φ zum Eigenwert λ, wenn $v \in \ker(\varphi - \lambda \; \mathrm{id})$.

Beweis $\varphi(v) = \lambda v \Leftrightarrow \varphi(v) - \lambda v = 0 \Leftrightarrow (\varphi - \lambda \; \mathrm{id})(v) = 0$

Definition 17.4 Für $\lambda \in K$ heißt

$$V(\lambda) := V(\varphi, \lambda) := \ker(\varphi - \lambda \; \mathrm{id})$$

der **Eigenraum von φ zum Eigenwert** λ.

Nach 17.3 besteht dieser aus 0 und allen Eigenvektoren zum Eigenwert λ.

Wieder lässt sich dies auf Matrizen $A \in M_n(K)$ umschreiben:

Lemma 17.3′ (a) $\lambda \in K$ ist Eigenwert von $A \Leftrightarrow \ker(A - \lambda E) \neq 0$.

(b) $v \neq 0$ ist Eigenvektor zum Eigenwert $\lambda \Leftrightarrow (A - \lambda E)v = 0$.

Definition 17.4′ Der Eigenraum von A zum Eigenwert λ ist

$$V(\lambda) = \ker(A - \lambda E)\,.$$

Dies ist also der Lösungsraum des homogenen Gleichungssystems $(A - \lambda E)v = 0$.

Satz 17.5 Sei V ein endlich-dimensionaler K-Vektorraum, und sei $\varphi : V \to V$ ein Endomorphismus. Dann gilt für $\lambda \in K$

$$\lambda \text{ Eigenwert von } \varphi \quad \Leftrightarrow \quad \det(\varphi - \lambda \cdot \; \mathrm{id}) = 0$$

Beweis: Wir wissen nach 10.19 und 15.4(a):

$$\begin{aligned}\ker(\varphi - \lambda\,\mathrm{id}) \neq 0 \quad &\Leftrightarrow \quad \varphi - \lambda\,\mathrm{id} \text{ ist kein Isomorphismus}\\ &\Leftrightarrow \quad \det(\varphi - \lambda\cdot\ \mathrm{id}) = 0\,.\end{aligned}$$

Die Bedingung lautet für quadratische Matrizen $A \in M_n(K)$:

Satz 17.5′ λ Eigenwert von $A \Leftrightarrow \det(A - \lambda E) = 0$.

Definition 17.6 Sei $A \in M_n(K)$. Das (normierte) **charakteristische Polynom von** A wird definiert als

$$\chi_A(x) = \det(xE - A) \in K[x]\,.$$

Dies ist wie folgt gemeint: Wir können z.B. die Leibniz-Formel nehmen und für $A = (a_{ij})$ definieren

$$\det(xE - A) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma)(x\delta_{1\sigma(1)} - a_{1\sigma(1)})(x\delta_{2\sigma(2)} - a_{2\sigma(2)}) \cdots (x\delta_{n\sigma(n)} - a_{n\sigma(n)})$$

Dies ist ein wohlbestimmtes Polynom in $K[x]$. Genauer gilt

Satz 17.7 (a) $\chi_A(x)$ ist ein Polynom in $K[x]$ vom Grad n.
(b) $\chi_A(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ mit

$$\begin{aligned} a_n \quad &= \quad 1 \qquad (\text{d.h., } \chi_A(x) \text{ ist normiert})\,,\\ a_{n-1} \quad &= \quad -(a_{11} + a_{22} + \ldots + a_{nn}) = -\sum_{i=1}^{n} a_{ii}\,,\\ a_0 \quad &= \quad (-1)^n \det(A)\,. \end{aligned}$$

Beweis Jeder der Summanden

$$\mathrm{sign}(\sigma)(x\delta_{1\sigma(1)} - a_{1\sigma(1)}) \cdots (x\delta_{n\sigma(n)} - a_{n\sigma(n)})$$

ist ein Polynom vom Grad $\leq n$; also gilt dies auch für $\chi_A(x)$.
Für $\sigma \neq \mathrm{id}$ ist mindestens ein $\delta_{i\sigma(i)} = 0$, und damit der obige Term vom Grad $\leq n-1$. Es ist dann sogar für ein weiteres j auch $\delta_{j\sigma(j)} = 0$, denn wäre $\sigma(\nu) = \nu$ für $n-1$ der Zahlen $\{1, \ldots, n\}$, so auch für alle Zahlen $1, \ldots, n$, wegen der Bijektivität von σ. Es ist also

$$\sum_{\sigma \neq \mathrm{id}} \mathrm{sign}(\sigma)(x\delta_{1\sigma(1)} - a_{1\sigma(1)}) \cdots (x\delta_{n\sigma(n)} - a_{n\sigma(n)})$$

vom Grad $\leq n-2$. Der verbleibende Term für $\sigma = \mathrm{id}$ ist

$$(x - a_{11})(x - a_{22}) \cdots (x - a_{nn})\,.$$

Nach Ausmultiplizieren und Zusammenfassen nach Potenzen von x ist hier der Koeffizient vor x^n gleich 1, und der vor x^{n-1} gerade $-(a_{11} + a_{22} + \ldots + a_{nn})$; also gilt dies auch für das Gesamtpolynom. Schließlich gilt

$$a_0 = \chi_A(0) = \det(-A) = (-1)^n \cdot \det(A)\,.$$

Definition 17.8 Für $A = (a_{ij}) \in M_n(K)$ heißt $\sum\limits_{i=1}^{n} a_{ii}$ die **Spur** von A (Bez.: $tr(A)$ oder $Sp(A)$).

Bemerkung 17.9 Für die Bestimmung von $\chi_A(x)$ kann man die üblichen Determinantenregeln benutzen; d.h., man schreibt formal

$$xE - A = \begin{pmatrix} x - a_{11} & -a_{12} & -a_{13} & & -a_{1n} \\ -a_{21} & x - a_{22} & -a_{23} & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ -a_{11} & \dots & \dots & & x - a_{nn} \end{pmatrix}$$

und berechnet die Determinante in $K[x]$, nach den üblichen Determinaten-Rechenregeln (Entwicklung nach Zeilen oder Spalten, Verhalten bei Spalten- oder Zeilenumformungen...)

Offenbar gilt:

Lemma 17.10 Für $A \in M_n(K)$ und $\lambda \in K$ ist

$$\chi_A(\lambda) = \det(\lambda E - A) = (-1)^n \cdot \det(A - \lambda E)$$

Zusammen mit 17.5′ folgt:

Satz 17.11 Für $A \in M_n(K)$ sind die Eigenwerte von A gerade die Nullstellen des charakteristischen Polynoms $\chi_A(x)$, d.h., für $\lambda \in K$ gilt

$$\lambda \text{ Eigenwert von } A \quad \Leftrightarrow \quad \chi_A(\lambda) = 0$$

Corollar 17.12 Eine Matrix $A \in M_n(K)$ hat höchstens n verschiedene Eigenwerte.

Dies folgt aus 17.11 und 16.12.

Beispiel 17.13 Wir bestimmen die Eigenwerte von $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in M_2(\mathbb{R})$. Es ist

$$\begin{aligned} \chi_A(x) &= \det(xE - A) = \det \begin{pmatrix} x-1 & -2 \\ -2 & x-1 \end{pmatrix} \\ &= (x-1)^2 - 4 = x^2 - 2x - 3 \end{aligned}$$

Die Nullstellen sind

$$\chi_{1,2} = 1 \pm \sqrt{1+3} = 1 \pm 2\,,$$

also $\lambda_1 = 3$ und $\lambda_2 = -1$. Dies sind die Eigenwerte von A.

Der Eigenraum $V(3)$ zu $\lambda_1 = 3$ ist $\ker(\lambda_1 E - A)$, also die Lösungsmenge des homogenen Gleichungssystems

$$\begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

Der Lösungsraum ist $V(3) = \mathbb{R}\begin{pmatrix}1\\1\end{pmatrix}$. Entsprechend ist

$$V(-1) = \ker\begin{pmatrix}-2 & -2\\-2 & -2\end{pmatrix} = \mathbb{R}\begin{pmatrix}1\\-1\end{pmatrix}.$$

Definition 17.14 Wir folgen in der Normierung Bourbaki und der englisch-sprachigen Literatur. In vielen deutschen Büchern und in der Physik wird auch oft

$$\det(A - xE) = (-1)^n \det(xE - A)$$

als charakteristisches Polynom von A bezeichnet. Der Nachteil ist, dass dies nicht normiert ist.

Satz 17.15 Für $A, B \in M_n(K)$ mit invertierbarem B gilt

(a)
$$\chi_{B^{-1}AB}(x) = \chi_A(x)$$

(b)
$$Sp(B^{-1}AB) = Sp(A)\,.$$

(Das charakteristische Polynom und die Spur sind also Konjugations-invariant).

Beweis: (a): Es gilt

$$\chi_{B^{-1}AB}(x) \overset{\text{Def.}}{=} \det(xE - B^{-1}AB) \overset{(1)}{=} \det(B^{-1}(xE - A)B) \overset{(2)}{=} \det(xE - A) \overset{\text{Def.}}{=} \chi_A(x)\,.$$

Die Gleichung (1) folgt daraus, dass $B^{-1}xEB = xB^{-1}EB = xB^{-1}B = xE$. Wir können aber nicht unmittelbar Lemma 15.1 anwenden ($\det(B^{-1}CB) = \det(C)$), um die Gleichung (2) zu beweisen, da hier $C = xE - A$ Einträge in $K[x]$ hat und dies kein Körper ist. Dieses Problem können wir aber auf zwei Arten lösen.

1. Methode: Man kann $K[x]$ in einen Körper $K(x)$ einbetten, den sogenannten Quotientenkörper von $K[x]$ (dieser wird so gebildet wie $\mathbb{Q}$ aus $\mathbb{Z}$: $K[x]$ besteht aus allen "Brüchen" $\frac{p(x)}{q(x)}$ mit Polynomen $p(x), q(x) \in K[x], q(x)$ ungleich dem Nullpolynom 0, und man rechnet mit den üblichen Regeln der Bruchrechnung). Dann können wir im Körper $K(x)$ rechnen und Lemma 15.1 anwenden, so dass (2) folgt.

2. Methode: Für jedes $\lambda \in K$ gilt nach Einsetzen

$$\chi_{B^{-1}AB}(\lambda) = \det(B^{-1}(\lambda E - A)B) \overset{15.1}{=} \det(\lambda E - A) = \chi_A(\lambda)\,.$$

Das Polynom $f(x) = \chi_{B^{-1}AB}(x) - \chi_A(x)$ hat also alle $\lambda \in K$ als Nullstelle. Ist K unendlich, so folgt hieraus $\chi_{B^{-1}AB}(x) - \chi_A(x)$, da ein Polynom $f(x) \neq 0$ nur endlich viele Nullstellen hat (16.12). Für (zum Beispiel) $K = \mathbb{Q}, \mathbb{R}$ oder $\mathbb{C}$ sind wir also fertig. Für endliche Körper K kann man aber zeigen (Algebra), dass sie sich in einen unendlichen Körper K' einbetten lassen (z.B. den algebraischen Abschluss), und kann dann den obigen Schluss in K' benutzen.

(b) Dies folgt aus (a), da $-Sp(A)$ der $(n-1)$-te Koeffizient (a_{n-1}) von $\chi_A(x)$ ist (17.7(b)).

Wie in Paragraph 15 können wir nun das charakteristische Polynom und die Spur auf Endomorphismen ausdehnen.

Definition 17.16 Sei V ein endlich-dimensionaler K-Vektorraum, und sei $\varphi : V \to V$ ein Endomorphismus. Dann sind das **charakteristische Polynom** $\chi_\varphi(x)$ und die **Spur** $Sp(\varphi)$ von φ definiert durch

$$\begin{aligned} \chi_\varphi(x) &= \chi_A(x) \\ Sp(\varphi) &= Sp(A)\,, \end{aligned}$$

wobei $A = M_b^b(\varphi)$ die darstellende Matrix von φ bezüglich irgendeiner Basis $b = (b_1, \dots, b_n)$ von V ist.

Wie in 15.3 folgt aus der Konjugations-Invarianz von $\chi_A(x)$ und $Sp(A)$, dass dies nicht von der Wahl der Basis abhängt.

Aus 17.11 und 17.12 folgt:

Satz 17.17 (a) $\lambda \in K$ ist genau dann Eigenwert von φ, wenn λ Nullstelle von $\chi_\varphi(x)$ ist.

(b) Ist $\dim V = n$, so hat φ höchstens n Eigenwerte.

§18 Euklidische und unitäre Skalarprodukte

Definition 18.1 Sei V ein $\mathbb{R}$-Vektorraum. Eine symmetrische Bilinearform

$$\psi : V \times V \to \mathbb{R}$$

(siehe 13.4 (ii)) heißt

(a) **positiv definit**, wenn $\psi(v,v) > 0$ für alle $v \in V \setminus \{0\}$,

(b) **negativ definit**, wenn $\psi(v,v) < 0$ für alle $v \in V \setminus \{0\}$, und

(c) **indefinit**, wenn ψ weder positiv noch negativ definit ist.

Beispiele 18.2 (a) Das Standard-Skalarprodukt

$$(x,y) \mapsto x^t y = \sum_{i=1}^n x_i y_i$$

auf $\mathbb{R}^n$ ist eine positiv definite symmetrische Bilinearform, denn für $x \neq 0$ ist $x^t x = \sum\limits_{i=1}^n x_i^2 > 0$.

(b) Auf dem Raum $C([a,b])$ der stetigen Funktionen auf dem Intervall $[a,b]$ ist die Bilinearform

$$(f,g) \mapsto \int\limits_a^b f(t)g(t)dt$$

symmetrisch und positiv definit, denn für stetiges $f(t) \not\equiv 0$ ist $\int\limits_a^b f(t)^2 dt > 0$.

(c) In der Relativitätstheorie betrachtet man den 4-dimensionalen Raum $\mathbb{R}^4$ mit 3 Ortskoordinaten x_1, x_2, x_3 und einer Zeitkoordinate x_4, sowie der "Minkowski-Metrik"

$$\psi(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4 \,,$$

die indefinit ist.

Definition 18.3 Sei K ein Körper, V ein endlich-dimensionaler K-Vektorraum, und

$$\psi : V \times V \to K$$

eine symmetrische Bilinearform. Ist $b = (b_1, \dots, b_n)$ eine Basis von V, so heißt die Matrix

$$A = M_b(\psi) := (\, \psi(b_i, b_j)\,) \in M_n(K)$$

die **Fundamentalmatrix von ψ bezüglich** b.

Definition 18.4 Eine Matrix $A = (a_{ij}) \in M_n(K)$ heißt **symmetrisch**, wenn $A = A^t$, d.h., wenn $a_{ij} = a_{ji}$ für alle i, j.

Bemerkungen 18.5 (a) Die Fundamentalmatrix $M_b(\psi)$ ist symmetrisch und bestimmt ψ, denn es ist $\psi(b_i, b_j) = \psi(b_j, b_i)$, und

$$\psi\left(\sum_{i=1}^n x_i b_i \sum_{j=1}^n y_j b_j\right) = \sum_{i,j=1}^n x_i y_j \psi(b_i, b_j)$$

wegen der Bilinearität von ψ. Umgekehrt gibt es zu jeder symmetrischen Matrix $A = (a_{ij}) \in M_n(K)$ genau eine symmetrische Bilinearform ψ mit $M_b(\psi) = A$, nämlich $\psi : V \times V \to K$ mit

$$\psi\left(\sum_{i=1}^n x_i b_i, \sum_{j=1}^n y_j b_j\right) = \sum_{i,j=1}^n x_i a_{ij} x_j = x^t A y$$

für $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots y_n)$.

(b) Insbesondere ist jede symmetrische Bilinearform ψ auf K^n von der Form

$$\begin{array}{rcl} \psi_A : K^n \times K^n & \to & K \\ (x, y) & \mapsto & x^t A y \end{array}$$

mit einer eindeutig bestimmten symmetrischen Matrix $A \in M_n(K)$, nämlich $A = M_e(\psi) = (\, \psi(e_i, e_j)\,)$ für die Standardbasis $e = (e_1, \dots, e_n)$ von K^n. Beachte hierzu, dass $e^t A e_j = a_{ij}$ für $A = (a_{ij})$.

Definition 18.6 Sei $A \in M_n(\mathbb{R})$ symmetrisch. Dann heißt A **positiv definit** (bzw. **negativ definit** bzw. **indefinit**), wenn dies für die zugehörige symmetrische Bilinearform

$$\begin{array}{rcl} \psi_A : \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ (x, y) & \mapsto & x^t A y \end{array}$$

gilt, d.h., wenn $x^tAx > 0 \; \forall x \in \mathbb{R}^n \setminus \{0\}$ (bzw. $x^tAx < 0 \; \forall x \in \mathbb{R}^n \setminus \{0\}$, bzw. wenn es $x, y \in \mathbb{R}^n \setminus \{0\}$ gibt mit $x^tAx \leq 0$ und $y^tAy \geq 0$.

Wir kommen nun zu komplexen Vektorräumen.

Definition 18.7 Sei V ein $\mathbb{C}$-Vektorraum. Eine **hermitesche Form auf** V ist eine Abbildung

$$\psi : V \times V \to \mathbb{C}\,,$$

so dass für alle $u, v, w \in V$ und $\lambda, \mu \in \mathbb{C}$ gilt:

(i) $\psi(\lambda u + \mu v, w) = \lambda\psi(u, w) + \mu\psi(v, w)$,

(ii) $\psi(v, w) = \overline{\psi(w, u)}$.

Hierbei ist $\overline{\alpha}$ das komplex Konjugierte von $\alpha \in \mathbb{C}$.

Eine hermitesche Form ist also **linear** im 1. Argument und **anti-linear** im zweiten Argument:

$$\psi(w, \lambda u + \mu v) \overset{(ii)}{=} \overline{\psi(\lambda u + \mu v, w)} \overset{(i)}{=} \overline{\lambda}\,\overline{\psi(u, w)} + \overline{\mu}\,\overline{\psi(v, w)} \overset{(ii)}{=} \overline{\lambda}\psi(w, u) + \overline{\mu}\psi(w, v)\,.$$

Es folgt weiter aus (ii), dass $\psi(v, v) = \overline{\psi(v, v)}$, also dass $\psi(v, v) \in \mathbb{R}$ für $v \in V$. Deswegen macht die folgende Definition Sinn:

Definition 18.8 Eine hermitesche Form $\psi : V \times V \to \mathbb{C}$ heißt

(a) **positiv definit**, wenn $\psi(v, v) > 0$ für alle $v \in V \setminus \{0\}$,

(b) **negativ definit**, wenn $\psi(v, v) < 0$ für alle $v \in V \setminus \{0\}$, und

(c) **indefinit**, falls ψ weder positiv noch negativ definit ist.

Beispiele 18.9 (a) Die Abbildung

$$\begin{array}{rccl} <,>: & \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ & (x, y) & \mapsto & x^t\overline{y} = \sum\limits_{i=1}^{n} x_i\overline{y_i} \end{array}$$

ist eine positiv definite hermitesche Form auf $\mathbb{C}^n$ und heißt das Standard-Skalarprodukt auf $\mathbb{C}^n$.

(b) Auf dem Raum $C([a, b], \mathbb{C})$ der komplexwertigen stetigen Funktionen auf $[a, b]$ ist

$$\psi(x, y) := \int\limits_a^b f(t)\overline{g(t)}dt$$

eine postitiv definite hermitesche Form.

Analog zum reellen Fall definiert und beweist man:

Definition/Satz 18.10 (a) Für endlich-dimensionales V mit Basis $b = (b_1, \ldots, b_n)$ sei $A = M_b(\psi) = (\,\psi(b_i, b_j)\,) \in M_n(\mathbb{C})$. Dann ist die Matrix A **hermitesch**, d.h., es gilt

$$A^t = \overline{A}\,,$$

wobei für $A = (a_{ij})$ die Matrix $\overline{A} = (\overline{a_{ij}})$ die **komplex konjugierte** Matrix sei.

(b) Hierdurch erhält man eine Bijektion zwischen hermiteschen Formen auf V und hermiteschen Matrizen $A \in M_n(\mathbb{C})$. Insbesondere ist jede hermitesche Form auf $\mathbb{C}^n$ von der Form

$$\begin{array}{rrcl} \psi_A: & \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ & (x,y) & \mapsto & x^t A \overline{y} \end{array}$$

mit einer hermiteschen Matrix $A \in M_n(\mathbb{C})$.

Definition 18.11 (a) Ein **euklidischer Raum** ist ein reeller Vektorraum V mit einem **euklidischen Skalarprodukt**

$$\begin{array}{rrcl} <,>: & V \times V & \to & \mathbb{R} \\ & (v,w) & \mapsto & <v,w>, \end{array}$$

d.h., einer positiv definiten symmetrischen Bilinearform auf V.

(b) Ein **unitärer Raum** ist ein komplexer Vektorraum V mit einem **unitären Skalarprodukt**,

$$\begin{array}{rrcl} <,>: & V \times V & \to & \mathbb{C} \\ & (v,w) & \mapsto & <v,w>, \end{array}$$

d.h., einer positiv definiten hermiteschen Form auf V.

Wir verwenden hier eine vereinfachte Schreibweise, $<v,w>$ statt $\psi(v,w)$.

Sei $(V,<,>)$ ein euklidischer Raum und $\mathbb{K} = \mathbb{R}$, oder ein unitärer Raum und $\mathbb{K} = \mathbb{C}$.

Definition 18.12 (a) Für $v \in V$ heißt

$$||v|| := \sqrt{<v,v>}$$

die **Norm** von v (Hier nehmen wir die nicht-negative reelle Quadratwurzel; beachte, dass $<v,v>$ reell und ≥ 0 ist).

(b) Ein Vektor $v \in V$ heißt **normiert** (oder Einheitsvektor), wenn $||v|| = 1$.

(c) Zwei Vektoren $v, w \in V$ heißen **orthogonal** zueinander, wenn $<v,w>= 0$.

Definition 18.13 (a) Ein **Orthonormalsystem** in V ist eine Familie $(v_i)_{i \in I}$ von Vektoren in V mit

$$<v_i, v_j>= \begin{cases} 1 & , \quad i = j, \\ 0 & , \quad i \neq j \end{cases}$$

(also $<v_i, v_j>= \delta_{ij}$).

(b) Ein Ortonormalsystem $(v_i)_{i \in I}$ heißt **Orthonormalbasis** (ONB) von V, wenn $(v_i)_{i \in I}$ zusätzlich eine Basis von V ist.

Bemerkung 18.14 (a) Die Vektoren in einem Orthonormalsystem sind offenbar normiert.

(b) Ein Orthonormalsysten ist immer linear unabhängig: Sei

$$\sum_{\nu=1}^{r} \alpha_\nu v_{i_\nu} = 0$$

mit $\alpha_1, \dots, \alpha_r \in \mathbb{R}$ bzw. $\mathbb{C}$. Dann ist für jedes $\mu = 1 \dots, r$

$$0 =< \sum_{\nu=1}^{r} \alpha_\nu v_{i_\nu}, v_{i_\mu} >= \alpha_\mu \,.$$

Satz 18.15 (Gram-Schmidtsches Orthonormalisierungsverfahren) Sei $\dim V = n$ und $(w_1, \dots, w_n)$ eine Basis von V. Dann erhält man eine Orthonormalbasis $(v_1, \dots, v_n)$ durch das folgende rekursive Verfahren: Setze

$$v_1 = \frac{w_1}{||w_1||} \quad (:= \frac{1}{||w_1||} w_1)$$

und für $2 \le k \le n$:

$$\begin{aligned} v_k' &= w_k - \sum_{i=1}^{k-1} < w_k, v_i > v_i \\ v_k &= \frac{v_k}{||v_k||} \,. \end{aligned}$$

Beweis Es ist $w_1 \neq 0$, also $||w_1|| > 0$. Weiter ist

$$\left|\left| \frac{w_1}{||w_1||} \right|\right| = < \frac{w_1}{||w_1||}, \frac{w_1}{||w_1||} > = \frac{1}{||w_1||^2} < w_1, w_1 > = 1 \,,$$

also v_1 normiert. Sei nun $k > 1$ und bereits bewiesen, dass $(v_1, \dots, v_{k-1})$ ein Orthonormalsystem ist. Dann ist für $j = 1, \dots, k-1$

$$\begin{aligned} < v_k', v_j > &= < w_k, v_j > - < w_k, v_j > ||v_i||^2 \\ &= 0 \,. \end{aligned}$$

Weiter ist $v_k' \neq 0$, da $v_1, \dots, v_{k-1}$ Linearkombinationen von $w_1, \dots, w_{k-1}$ sind, und die Vektoren $w_1, \dots, w_k$ linear unabhängig sind . Nach dem selben Schluss wie oben ist dann $v_k = \frac{v_k'}{||v_k||}$ normiert. Es folgt, dass $(v_1, \dots, v_k)$ ein Orthonormalsystem ist, und damit linear unabhängig (8.14(b)). Für $k = n = \dim V$ muss $(v_1, \dots, v_n)$ dann auch eine Basis sein (8.23).

§19 Hauptachsentransformation/Spektralsätze

Sei $(V, <,>)$ ein euklidischer Raum und $\mathbb{K} = \mathbb{R}$, oder ein unitärer Raum und $\mathbb{K} = \mathbb{C}$.

Definition 19.1 Ein Endomorphismus $\varphi : V \to V$ heißt **selbstadjungiert** (bezüglich $<,>$), wenn für alle $v, w \in V$ gilt

$$< v, \varphi(w) > = < \varphi(v), w > \,.$$

Beispiele 19.2 Betrachte $\mathbb{R}^n$ mit dem Standard-Skalarprodukt $<,>$. Eine bilineare Abbildung

$$A : \mathbb{R}^n \to \mathbb{R}^n \qquad (A \in M_n(\mathbb{R}))$$

ist genau dann selbstadjungiert, wenn A symmetrisch ist. Denn es gilt

$$\begin{aligned} & < x, Ay > = < Ax, y > \quad \forall\, x, y \in \mathbb{R}^n \\ \Leftrightarrow \quad & x^t Ay = (Ax)^t y = x^t A^t y \quad \forall\, x, y \in \mathbb{R} \\ \Leftrightarrow \quad & A = A^t\,. \end{aligned}$$

Für das letzte "$\Rightarrow$" beachte, dass $e_i^t A e_j = a_{ij}$. Weiter haben wir benutzt, dass offenbar für einen beliebigen Körper K und Matrizen $A \in M(m \times n, K), B \in M(n, r, K)$ gilt

$$(AB)^t = B^t A^t\,.$$

(b) Genauso folgt: Für $\mathbb{C}^n$ mit dem Standard-Skalarprodukt $< x, y > = x^t \overline{y} = \sum\limits_{i=1}^{n} x_i \overline{y}_i$ ist

$$A : \mathbb{C}^n \to \mathbb{C}^n \qquad (A \in M_n(\mathbb{C}))$$

genau dann selbstadjungiert, wenn A hermitesch ist ($A = \overline{A}^t$). Beachte, dass $\overline{A}^t = \overline{A^t}$.

Satz 19.3 (Spektralsatz) Sei $\dim_{\mathbb{K}} V = n < \infty$ und $\varphi : V \to V$ eine selbstadjungierte lineare Abbildung. Dann gilt

(a) Es ist $\chi_\varphi(x) = \prod\limits_{i=1}^{n} (x - \lambda_i)$ mit reellen $\lambda_1, \dots, \lambda_n$ ("Alle Eigenwerte von φ sind reell").

(b) V beseitzt eine Basis $(b_1, \dots, b_n)$ aus Eigenvektoren für φ.

Wir führen den Beweis in mehreren Schritten.

Lemma 19.4 φ besitzt einen reellen Eigenwert λ.

Beweis: später!

Sei $U := V(\lambda)$ der Eigenraum von φ bezüglich λ, und sei

$$U^\perp := \{w \in V \mid < u, w > = 0 \quad \text{für alle } u \in U\}$$

das **orthogonale Komplement** von U.

Lemma 19.5 (a) $\varphi(U) \subseteq U$.

(b) $\varphi(U^\perp) \subseteq U^\perp$.

Beweis (a): $u \in U \Rightarrow \varphi(u) = \lambda u \in U$.

(b) $w \in U^\perp \Rightarrow < u, w > = 0\ \forall\, u \in U \Rightarrow < u, \varphi(w) > = < \varphi(u), w > = 0\ \forall\, u \in U$ (da $\varphi(u) \in U$ nach (a)) $\Rightarrow \varphi(w) \in U^\perp$.

Nach Satz 18.15 (Gram-Schmidt) besitzt U eine Orthonormalbasis $(v_1, \dots, v_r)$. Dies sind alles Eigenvektoren von φ, zum Eigenwert λ.

Nach 19.5 (b) induziert φ durch Einschränkung einen Endomorphismus

$$\begin{array}{rccc} \tilde{\varphi}: & U^\perp & \to & U^\perp \\ & v & \mapsto & \varphi(v)\,. \end{array}$$

Es ist $U \neq 0$ (nach 19.4) und daher $U^\perp \neq V$ (ist $u \neq 0$ in U, so ist $u \notin U^\perp$, da $< u, u > > 0$). Mit Induktion über $\dim V$ können wir also annehmen, dass $U^\perp$ eine Orthonormalbasis $(w_1, \dots, w_s)$ aus Eigenvektoren für $\tilde{\varphi}$ besitzt, mit reellen Eigenwerten; dies sind dann auch Eigenvektoren von φ. Der Spektralsatz 19.3 folgt also aus dem folgenden Lemma.

Lemma 19.6 $(v_1, \dots, v_r, w_1, \dots, w_s)$ ist eine Orthonormalbasis von V.

Beweis Offenbar bilden diese Vektoren ein Orthonormalsystem (wegen $w_j \in U^\perp$ ist $< v_i, w_j >= 0 \; \forall\, i, j$).

Es genügt also zu zeigen, dass diese Vektoren V erzeugen (die lineare Unabhängigkeit folgt aus 18.14 (b)). Sei aber $v \in V$. Dann ist

$$w := v - \sum_{i=1}^{r} < v, v_i > v_i \ \in U^\perp\,.$$

Denn es ist w orthogonal zu allen v_i (vergleiche den Beweis von 18.15), also zu allen Linearkombinationen der v_i, also zu allen $u \in U$. Damit ist w Linearkombination der w_j, also v Linearkombination von $v_1, \dots, v_r, w_1, \dots, w_s$.

Wir interpretieren dies für Matrizen:

Definition 19.7 Eine Matrix $A \in M_n(\mathbb{R})$ heißt **orthogonal**, wenn gilt

(19.7.1) $$A^t A = E\,,$$

d.h., A ist invertierbar und $A^{-1} = A^t$.

Bemerkung 19.8 Offenbar ist (19.7.1) äquivalent dazu, dass die Spalten $(v_1, \dots, v_n)$ von A eine Orthonormalbasis von $\mathbb{R}^n$ bilden (bezüglich des Standard-Skalarprodukts). Denn es ist

$$A^t A = \begin{pmatrix} - & v_1^t & - \\ & \vdots & \\ - & v_n^t & - \end{pmatrix} \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix} = (< v_i, v_j >)\,.$$

Satz 19.9 (Hauptachsentransformation) Sei $A \in M_n(\mathbb{R})$ symmetrisch. Dann gibt es eine orthogonale Matrix $T \in M_n(\mathbb{R})$, so dass gilt

$$T^{-1} A T = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Die λ_i sind die Eigenwerte von A, und es gilt $\chi_A(x) = \prod_{i=1}^{n} (x - \lambda_i)$ für das charakteristische Polynom von A.

Beweis Die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^n$ ist nach 19.2 selbstadjungiert (bezüglich des Standard-Skalarprodukts). Nach dem Spektralsatz 19.3 gibt es also eine Orthonormalbasis $(v_1, \dots, v_n)$ von $\mathbb{R}^n$ aus Eigenvektoren von A, und die zugehörigen Eigenwerte $\lambda_1, \dots, \lambda_n$ $(Av_i = \lambda_i v_i)$ sind reell. Sei

$$T = (v_1, \dots, v_n)$$

die Matrix mit den Spalten $v_1, \dots, v_n$. Dann ist T orthogonal (Bemerkung 19.8) und

$$T^t AT = T^t(Av_1, \dots, Av_n) = (v_1, \dots, v_n)^t(\lambda_1 v_1, \dots, \lambda v_n) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Wegen $T^t = T^{-1}$ folgt die erste Behauptung. Weiter ist

$$\chi_A(x) \overset{17.15}{=} \chi_{T^{-1}AT}(x) = \det \begin{pmatrix} x - \lambda_1 & & \\ & \ddots & \\ & & x - \lambda_n \end{pmatrix} = \prod_{i=1}^n (x - \lambda_i)$$

Beispiel 19.10 Betrachte die symmetrische Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in M_2(\mathbb{R}).$$

Nach Beispiel 17.13 sind die Eigenwerte von A die reellen Zahlen $\lambda_1 = 3$ und $\lambda_2 = -1$. Die zugehörigen Eigenräume sind eindimensional, und zwar $V(3) = \mathbb{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $V(-1) = \mathbb{R}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Die Eigenvektoren $v_1' = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in V(3)$ und $v_2' = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in V(-1)$ sind orthogonal zueinander, aber noch nicht normiert; die Eigenvektoren

$$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

bilden eine Orthonormalbasis. Die Matrix

$$T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

ist dann orthogonal, und es gilt

$$\begin{aligned} T^{-1}AT &= T^t AT = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 3 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 6 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \end{aligned}$$

Ganz entsprechend ergibt sich das Folgende für hermitesche Matrizen.

Definition 19.11 Eine Matrix $A \in M_n(\mathbb{C})$ heißt **unitär**, wenn gilt

$$\overline{A}^t A = E,$$

d.h., A ist invertierbar und $A^{-1} = \overline{A}^t$.

Bemerkung 19.12 Dies ist äquivalent dazu, dass die Spalten $(v_1, \ldots, v_n)$ von A eine Orthonormalbasis von $\mathbb{C}^n$ bilden (bezüglich des unitären Standard-Skalarproduktes $< x, y >= x^t \overline{y}$).

Satz 19.13 Sei $A \in M_n(\mathbb{C})$ eine hermitesche Matrix. Dann hat A reelle Eigenwerte $\lambda_1, \ldots, \lambda_n$ (das charakteristische Polynom ist $\chi_A(x) = \prod_{i=1}^{n}(x - \lambda_i)$), und es gibt eine unitäre Matrix $U \in M_n(\mathbb{C})$ mit

$$U^{-1}AU = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Die Beweise von 19.12 und 19.13 sind völlig analog zu denen von 19.8 und 19.9. Insbesondere finden wir die (im Allgemeinen nicht eindeutig bestimmte) Matrix U in 19.13, indem wir eine Orthonormalbasis $(v_1, \ldots, v_n)$ von $\mathbb{C}^n$ aus Eigenvektoren von A konstruieren und für U die Matrix mit den Spalten $v_1, \ldots, v_n$ nehmen.

Wir wenden uns nun dem Beweis von 19.4 zu. Sei $(V, <,>)$ ein euklidischer Raum und $\mathbb{K} = \mathbb{R}$, oder ein unitärer Raum und $\mathbb{K} = \mathbb{C}$.

Satz 19.14 Ist $\varphi : V \to V$ ein selbstadjungierter Endomorphismus, so sind alle Eigenwerte von φ reell.

Beweis Sei $v \in V \setminus \{0\}$ und $\lambda \in \mathbb{K}$ mit $\varphi(v) = \lambda v$. Dann ist

$$\begin{aligned} \lambda < v, v > &= < \lambda v, v > &&= < \varphi(v), v > \\ &= < v, \varphi(v) > && \text{(da } \varphi \text{ selbstadjungiert ist)} \\ &= < v, \lambda v > &&= \overline{\lambda} < v, v > . \end{aligned}$$

Es folgt $(\lambda - \overline{\lambda}) < v, v >= 0$, wegen $< v, v > \neq 0$ also $\lambda = \overline{\lambda}$, d.h., λ ist reell.

Wir kommen nun zum

Beweis von Lemma 19.4: Sei V endlich-dimensional und $\varphi : V \to V$ ein selbstadjungierter Endomorphismus.

1. Fall: Sei zunächst $(V, <,>)$ unitär, also insbesondere $\mathbb{K} = \mathbb{C}$. Für $V \neq 0$ ist das charakteristische Polynom $\chi_\varphi(x)$ vom Grad $n = \dim V > 0$, also nicht konstant. Nach 16.3 besitzt $\chi_A(x)$ also eine Nullstelle $\lambda \in \mathbb{C}$. Nach 17.13 ist λ ein Eigenwert von φ, und nach 19.14 ist λ reell. Damit ist 19.4 in diesem Fall bewiesen.

Für den euklidischen Fall benutzen wir das folgende Lemma.

Lemma 19.15 Ist $b = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V, so ist ein Endomorphismus $\varphi : V \to V$ genau dann selbstadjungiert, wenn die darstellende Matrix $A = M_b^b(\varphi)$ symmetrisch beziehungsweise hermitesch ist.

Beweis: Sei $A = (a_{ij})$; dann ist nach Definition $\varphi(b_j) = \sum\limits_{i=1}^{n} a_{ij} b_i$. Es folgt

$$(19.17.1) \qquad < \varphi(b_j), b_k > = \sum_{i=1}^{n} a_{ij} < b_i, b_k > = a_{kj} \, ,$$

wegen $< b_i, b_k > = \delta_{ik}$. Ist φ selbstadjungiert, so folgt für alle j, k

$$(19.17.2) \qquad < \varphi(b_j), b_k > = < b_j, \varphi(b_k) > = \overline{< \varphi(b_k), b_j >} \, ,$$

also

$$(19.17.3) \qquad a_{kj} = \overline{a_{jk}} \, ,$$

d.h., $A = \overline{A}^t (= A^t$ falls A reell ist). Umgekehrt folgt aus (19.17.3) wegen (19.17.1) auch (19.17.2) und damit

$$< \varphi(v), w > = < v, \varphi(w) >$$

für alle $v, w \in V$, da jedes $v \in V$ Linearkombination der b_i ist.

Bemerkung: Im Spezialfall $V = \mathbb{K}^n$, $<,> =$ Standard-Skalarprodukt ergeben sich wieder die Beispiele 19.2, da für die lineare Abbildung $A : \mathbb{K}^n \to \mathbb{K}$ die Matrix A die darstellende Matrix bezüglich der Standardbasis $e = (e_1, \ldots, e_n)$ ist, welche eine Orthonormalbasis für das Standard-Skalarprodukt ist.

2. Fall: Sei nun $(V, <,>)$ (endlich-dimensional und) euklidisch. Nach 18.15 besitzt V eine Orthonormalbasis $b = (b_1, \ldots, b_n)$, und nach 19.15 ist $A = M_b^b(\varphi) \in M_n(\mathbb{R})$ symmetrisch. Weiter ist nach Definition $\chi_\varphi(x) = \chi_A(x)$. Es genügt also zu zeigen, dass $\chi_A(x)$ eine reelle Nullstelle λ hat (diese ist dann ein Eigenwert von φ).

Fassen wir aber A als komplexe Matrix in $M_n(\mathbb{C})$ auf, so ist A hermitesch: $(A = A^t = \overline{A}^t)$, liefert also einen selbstadjungierten Endomorphismus $A : \mathbb{C}^n \to \mathbb{C}^n$. Nach dem ersten Fall besitzt also A einen reellen Eigenwert λ, und somit $\chi_A(x)$ die reelle Nullstelle λ. Damit ist Lemma 19.4 auch im euklidischen Fall bewiesen.

Zur tatsächlichen Bestimmung einer Orthonormalbasis aus Eigenvektoren sind die folgenden Betrachtungen nützlich. Sei $(V, <,>)$ wieder ein euklidischer Raum und $\mathbb{K} = \mathbb{R}$, oder ein unitärer Raum und $\mathbb{K} = \mathbb{C}$.

Definition 19.17 (a) (siehe oben) Für einen Unterraum $U \subseteq V$ heißt

$$U^\perp = \{w \in V \mid < u, w > = 0 \quad \text{für alle } u \in U\}$$

das **orthogonale Komplement** von U.

(b) Zwei Unterräume $U_1, U_2 \subseteq V$ heißen **orthogonal** zueinander, wenn

$$< u_1, u_2 > = 0 \quad \text{für alle } u_1 \in U_1, u_2 \in U_2 \, .$$

Bemerkung 19.18 (a) $U^\perp$ ist offenbar wieder ein Unterraum von V, denn für $w_1, w_2 \in U^\perp, \lambda, \mu \in \mathbb{K}$ und $u \in U$ gilt

$$< u, \lambda w_1 + \mu w_2 > = \overline{\lambda} < u, w_1 > + \overline{\mu} < u, w_2 > = 0 \, ,$$

also $\lambda w_1 + \mu w_2 \in U^\perp$.

(b) Unterräume U_1 und U_2 in V sind offenbar genau dann orthogonal zueinander, wenn gilt $U_2 \subseteq U_1^\perp$ (oder, äquivalent dazu, $U_1 \subseteq U_2^\perp$).

Satz 19.19 Sei $\varphi : V \to V$ ein selbstadjungierter Endomorphismus. Sind $\lambda \neq \mu$ zwei verschiedene Eigenwerte von φ, so sind die zugehörigen Eigenräume $V(\lambda)$ und $V(\mu)$ orthogonal zueinander.

Beweis Sei $v \in V(\lambda)$ und $w \in V(\mu)$. Dann ist

$$\begin{aligned} \lambda < v, w > &= < \lambda v, w > &&= < \varphi(v), w > \\ &= < v, \varphi(w) > &&= < v, \mu w) > = \mu < v, w > \end{aligned}$$

da μ reell ist (19.14). Es folgt $(\lambda - \mu) < v, w >= 0$, wegen $\lambda \neq \mu$ also $< v, w >= 0$.

Corollar 19.20 Sei V endlich-dimensional, und sei $\varphi : V \to V$ ein selbst-adjungierter Endomorphismus. Seien $\lambda_1, \ldots, \lambda_r$ die verschiedenen Eigenwerte von φ und sei $(v_1^{(i)}, \ldots, v_{n_i}^{(i)})$ eine Orthonormalbasis des Eigenraums $V(\lambda_i)$ von φ bezüglich $\lambda_i \quad (i = 1, \ldots, r)$ (Insbesondere gilt $n_i = \dim V(\lambda_i)$). Dann bilden die Vektoren

$$(v_1^{(1)}, \ldots, v_{n_1}^{(1)}, v_1^{(2)}, \ldots, v_{n_2}^{(2)}, \ldots, v_1^{(r)}, \ldots, v_{n_r}^{(r)})$$

eine Orthonormalbasis von V. Insbesondere ist $n = \dim V = \sum\limits_{i=1}^{r} n_i$.

Beweis Nach 19.19 bilden die Vektoren ein Orthonormalsystem in v. Sie bilden aber auch ein Erzeugendensystem von V, denn nach dem Spektralsatz 19.3 ist jeder Vektor $v \in V$ eine Linearkombination von Eigenvektoren $b_1, \ldots, b_n$ von φ, und jeder Vektor b_i liegt in einem Eigenraum, ist also Linearkombination von $v_1^{(1)}, \ldots, v_{n_r}^{(r)}$.

Beispiel 19.21 Betrachte die symmetrische Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 5 \end{pmatrix} \in M_2(\mathbb{R})$$

Es ist

$$\begin{aligned} \chi_A(x) &= \det \begin{pmatrix} x-1 & 0 & -1 \\ 0 & x-1 & -2 \\ -1 & -2 & x-5 \end{pmatrix} = (x-1) \begin{vmatrix} x-1 & -2 \\ -2 & x-5 \end{vmatrix} - 1 \begin{vmatrix} 0 & x-1 \\ -1 & -2 \end{vmatrix} \\ &= (x-1)((x-1)(x-5)-4) - (x-1) \\ &= (x-1)^2(x^2 - 6x + 5 - 4 - 1) = (x-1)(x-6)x\,. \end{aligned}$$

Also sind die Eigenwerte von A die Zahlen $1, 6$ und 0. Basisvektoren für die Eigenräume $V(1), V(6)$ und $V(0)$ sind

$$v_1' = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \text{ bzw. } v_2' = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} \text{ bzw. } v_3' = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix},$$

wie man leicht nachrechnet: $V(\lambda)$ ist der Lösungsraum des homogenen linearen Gleichungssystems $(\lambda E - A)x = 0$, also

$$\begin{pmatrix} \lambda-1 & 0 & -1 \\ 0 & \lambda-1 & -2 \\ -1 & -2 & \lambda-5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Für $\lambda = 1$ erhält man zum Beispiel

$$\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -2 \\ -1 & -2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

mit Lösungsraum $\mathbb{R}v_1'$. Die Vektoren v_1', v_2' und v_3' sind paarweise orthogonal zueinander, und die drei Vektoren

$$v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \; v_2 = \frac{1}{\sqrt{30}} \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \; v_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

bilden eine Orthonormalbasis aus Eigenvektoren für A.

Wir können Corollar 19.19 noch etwas konzeptioneller formulieren.

Nachtrag zu 18.12 (b) und 19.18 (b): Sind $v, w \in V$ orthogonal zueinander, so schreiben wir $v \perp w$. Sind Unterräume $U_1, U_2 \subseteq V$ orthogonal zueinander, so schreiben wir $U_1 \perp U_2$.

Definition 19.21 Seien $U_1, \ldots, U_r \subseteq V$ Unterräume. Wir sagen V ist die **orthogonale Summe** der U_i (Bez.: $V = U_1 \perp \ldots \perp U_r$), wenn gilt

(a) Die Räume sind paarweise orthogonal zueinander: $U_i \perp U_j$ für $i \neq j$.

(b) V ist die Summe der Unterräume U_i, d.h., $V = U_1 + \ldots + U := \{u_1 + \ldots + u_r \mid u_i \in U_i$ für alle $i = 1, \ldots, r\}$.

Lemma 19.22 Ist V die orthogonale Summe der Unterräume $U_1, \ldots, U_r$, so ist V auch die **direkte Summe** der U_i, d.h., die lineare Abbildung

$$\begin{aligned} \varphi : U_1 \oplus \ldots \oplus U_r &\rightarrow V \\ (u_1, \ldots, u_r) &\mapsto u_1 + \ldots + u_r \end{aligned}$$

ist ein Isomorphismus (Siehe 9.6 für die Definition von $U_1 \oplus \ldots \oplus U_n$). Insbesondere gilt

$$\dim V = \sum_{i=1}^{r} \dim U_i .$$

Beweis Nach 19.21 (b) ist φ surjektiv. Weiter zeigen $\ker(\varphi) = 0$: Sind $u_i \in U_i \quad (i = 1, \ldots, n)$ mit $u_1 + \ldots + u_r = 0$, so gilt für jedes $i \in \{1, \ldots, r\}$ wegen 19.21 (a)

$$0 = < u_i, u_1 + \ldots + u_r > = < u_i, u_i > ,$$

also $u_i = 0$. Damit ist φ auch injektiv. Die letzte Behauptung des Lemmas folgt mit 9.7.

Wir erhalten dann

Corollar 19.19′ Sei V endlich-dimensional, $\varphi : V \to V$ ein selbstadjungierter Endomorphismus und seien $\lambda_1, \dots, \lambda_r$ die verschiedenen Eigenwerte von φ. Dann ist

$$V = V(\lambda_1) \perp \dots \perp V(\lambda_r)\,.$$

Beweis Nach Satz 19.18 sind die $V(\lambda_i)$ paarweise orthogonal zueinander. Da V eine Orthonormalbasis aus Eigenvektoren von φ besitzt, erzeugen die $V(\lambda_i)$ auch V, d.h., $V = \sum\limits_{i=1}^{r} V(\lambda_i)$.

Hieraus folgt auch wieder die Aussage 19.19:

Lemma 19.23 Ist $V = U_1 \perp \dots \perp U_r$, und ist $(v_1^{(i)}, \dots, v_{n_i}^{(i)})$ eine Orthonormalbasis von $U_i \quad (i = 1, \dots, r)$, so ist

$$(v_1^{(1)}, \dots, v_{n_1}^{(1)}, v_1^{(2)}, \dots, v_{n_2}^{(2)}, \dots, v_1^{(r)}, \dots, v_{n_r}^{(r)})$$

eine Orthonormalbasis von V.

Beweis Offenbar ist dies ein Orthonormalsystem von V, und ebenso ein Erzeugendensystem von V, da $V = \sum\limits_{i=1}^{r} U_i$.

Wir notieren noch die folgende Beobachtung.

Lemma 19.24 Sei $U \subseteq V$ ein endlich-dimensionaler Unterraum. Dann ist

$$V = U \perp U^{\perp}\,.$$

Insbesondere gilt $\dim V = \dim U + \dim U^{\perp}$.

Beweis Nach Definition ist $U \perp U^{\perp}$. Wir zeigen noch $V = U + U^{\perp}$; die letzte Behauptung folgt dann aus 19.22. Sei $(u_1, \dots, u_m)$ eine Orthonormalbasis von U (existiert nach Gram-Schmidt), und sei $v \in V$. Dann ist

$$w := v - \sum_{i=1}^{m} < v, u_i > u_i \ \in U^{\perp}\,,$$

denn $< w, u_j > = < v, u_j > \; - \; < v, u_j > \; ||u_j||^2 = 0$ für alle $j = 1, \dots, m$. Es folgt

$$v = \sum_{i=1}^{m} < v, u_i > u_i + w \ \in U + U^{\perp}\,.$$

Beispiel 19.25 Betrachte $\mathbb{R}^3$ mit dem Standard-Skalarprodukt, und

$$U = < u_1 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} > \subseteq \mathbb{R}^3 .$$

Dieser Unterraum ist 2-dimensional, da die beiden Vektoren linear unabhängig sind. Es gilt

$$\begin{aligned} x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in U^\perp \quad &\Leftrightarrow \quad u_i^t x = 0 \text{ und } u_2^t x = 0 \\ &\Leftrightarrow \quad \begin{pmatrix} 1 & 2 & 5 \\ 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 . \end{aligned}$$

Dies ist ein homogenes lineares Gleichungssystem mit Lösungsraum

$$\mathbb{R} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = U^\perp .$$

Also ist $U^\perp$ 1-dimensional. (vergleiche Beispiel 19.20).

§20 Orthogonale und adjungierte Abbildungen

Seien $(V, <,>)$ und $(V', <,>)$ zwei euklidische (bzw. zwei unitäre) Räume. Schreiben wir nur $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$), so meinen wir das Standard-Skalarprodukt.

Definition 20.1 Eine lineare Abbildung $\varphi : V \to V'$ heißt **orthogonal** (bzw. **unitär**), wenn gilt

$$< \varphi(v), \varphi(w) >' = < v, w >$$

für alle $v, w \in V$.

Bemerkungen 20.2 (a) Es folgt, dass φ injektiv ist: $\varphi(v) = 0 \Rightarrow 0 =< \varphi(v), \varphi(v) > = < v, v > \Rightarrow v = 0$.

(b) Ist V endlich-dimensional und $b = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V, so ist

$$\begin{array}{rcccccl} \varphi_b & : & \mathbb{R}^n & \to & V & , & e_i \mapsto b_i \\ (\text{bzw. } \varphi_b & : & \mathbb{C}^n & \to & V & , & e_i \mapsto b_i) \end{array}$$

eine orthogonale Isomorphie (bzw. unitäre Isomorphie). V ist also orthogonal (bzw. unitär) isomorph zum Standard-euklidischen Raum $\mathbb{R}^n$ (bzw. Standard-unitären Raum $\mathbb{C}^n$).

Lemma/Definition 20.3 Die orthogonalen (bzw. unitären) Isomorphismen $\varphi : V \to V$ bilden eine Gruppe unter Komposition. Diese heißt die **orthogonale Gruppe** $O(V, <,>)$ (bzw. die **unitäre Gruppe** $U(V, <,>)$). Wenn das Skalarprodukt fest ist, lassen wir es auch in der Bezeichnung weg.

Beweis dass wir eine Gruppe erhalten: Es ist klar, dass die Komposition von orthogonalen Abbildungen wieder orthogonal ist, und dass das Inverse eines orthogonalen Isomorphismus wieder orthogonal ist. Dasselbe gilt im unitären Fall.

Beispiele 20.4 (a) Eine lineare Abbildung

$$T : \mathbb{R}^n \to \mathbb{R}^n \quad (T \in M_n(\mathbb{R}))$$

ist genau dann orthogonal, wenn T eine orthogonale Matrix ist:

$$\begin{aligned} & < Tx, Ty > = < x, y > \quad \forall x, y \in \mathbb{R}^n \\ \Leftrightarrow \quad & x^t T^t T y \text{ für } = x^t y = x^t E y \quad \forall x, y \in \mathbb{R}^n \\ \Leftrightarrow \quad & T^t T = E\,. \end{aligned}$$

Insbesondere ist T dann invertierbar. Die orthogonale Gruppe von $\mathbb{R}^n$ ist also

$$O(n) := \{T \in M_n(\mathbb{R}) \mid T^t T = E\}\,.$$

(b) Entsprechend ist eine lineare Abbildung

$$U : \mathbb{C}^n \to \mathbb{C}^n \quad (U \in M_n(\mathbb{C}))$$

genau dann unitär, wenn U eine unitäre Matrix ist, und die unitäre Gruppe von $\mathbb{C}^n$ ist

$$U(n) := \{U \in M_n(\mathbb{C}) \mid \overline{U}^t U = E\}$$

Definition 20.5 Sei $\varphi : V \to V'$ eine lineare Abbildung. Eine lineare Abbildung

$$\varphi^* : V' \to V$$

heißt **adjungiert zu** φ, wenn gilt

$$< \varphi(v), v' >' = < v, \varphi^*(v') >$$

für alle $v \in V, v' \in V'$.

Bemerkungen 20.6 Besitzt φ eine adjungierte Abbildung φ^*, so ist diese eindeutig: Ist $\tilde{\varphi}$ ebenfalls adjungiert zu φ, so ist für alle $v \in V, v' \in V'$

$$< v, \varphi^*(v') > = < \varphi(v), v' > = < v, \tilde{\varphi}(v') >\,,$$

also $< v, \varphi^*(v') - \tilde{\varphi}(v') > = 0$. Für $v = \varphi^*(v') - \tilde{\varphi}(v')$ folgt $\varphi^*(v') = \tilde{\varphi}(v')$, also $\varphi^* = \tilde{\varphi}$, da v' beliebig war.

Lemma 20.7 Sind V und V' endlich dimensional, so existiert immer eine adjungierte Abbildung.

Beweis Nach Bemerkung 20.2 (b) können wir annehmen, dass V gleich $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) ist und V' gleich $\mathbb{R}^m$ (bzw. $\mathbb{C}^m$). Dann ist φ durch eine $(m \times n)$-Matrix A gegeben, und φ^* durch die **adjungierte Matrix** $A^* := \overline{A}^t (= A^t$ im reellen Fall). Denn es ist

$$< Ax, y > = (Ax)^t \overline{y} = x^t A^t \overline{y} = x^t \overline{\overline{A}^t y} = < x, A^* y > \ .$$

Bemerkungen 20.8 Für einen Endomorphismus $\varphi : V \to V$ gilt also nach Definition:

$$\varphi \text{ selbstadjungiert } \Leftrightarrow \quad \varphi^* = \varphi\,.$$