Joachim Gräter
Elke Rosenberger

Lineare Algebra

Potsdam, März 2004

Prof. Dr. J. Gräter, E. Rosenberger
Universität Potsdam, Institut für Mathematik
Am Neuen Palais 10, 14469 Potsdam

Die neueste Version dieses Skriptes ist erhältlich unter
http://users.math.uni-potsdam.de/~graeter/

Inhaltsverzeichnis

KAPITEL 1

Grundlagen

1. Logische Zeichen

Sind $\mathcal{A}$ und $\mathcal{B}$ Aussagen, so können auf folgende Weise neue Aussagen gebildet werden:

$\mathcal{A} \wedge \mathcal{B}$	$\mathcal{A}$ und $\mathcal{B}$
$\mathcal{A} \vee \mathcal{B}$	$\mathcal{A}$ oder $\mathcal{B}$
$\neg\mathcal{A}$	nicht $\mathcal{A}$
$\mathcal{A} \Rightarrow \mathcal{B}$	Aus $\mathcal{A}$ folgt $\mathcal{B}$
$\mathcal{A} \Leftrightarrow \mathcal{B}$	$\mathcal{A}$ genau dann, wenn $\mathcal{B}$

Ist $\mathcal{A}$ wahr, so ist $\neg\mathcal{A}$ falsch und umgekehrt. Die Wahrheitswerte (w = wahr und f = falsch) der anderen Aussagen ergeben sich aus den Wahrheitswerten von $\mathcal{A}$ und $\mathcal{B}$ gemäß der folgenden Tabellen:

$\wedge$	w	f
w	w	f
f	f	f

$\vee$	w	f
w	w	w
f	w	f

$\Rightarrow$	w	f
w	w	f
f	w	w

$\Leftrightarrow$	w	f
w	w	f
f	f	w

Die logischen Zeichen $\wedge$ und $\vee$ nennt man Junktoren. Eine Aussage der Form $\mathcal{A} \Rightarrow \mathcal{B}$ bezeichnet man als Implikation, und man sagt: $\mathcal{A}$ *impliziert* $\mathcal{B}$. Gilt $\mathcal{A} \Leftrightarrow \mathcal{B}$, so sagt man, die beiden Aussagen $\mathcal{A}$ und $\mathcal{B}$ sind äquivalent oder gleichwertig.

Ist A eine Menge und $\mathcal{A}(x)$ eine für alle Elemente x von A definierte Aussage, dann bedeutet:

$\forall x \in A : \mathcal{A}(x)$	Für alle Elemente x der Menge A gilt die Aussage $\mathcal{A}(x)$.
$\exists x \in A : \mathcal{A}(x)$	Es gibt ein Element x der Menge A, für das die Aussage $\mathcal{A}(x)$ gilt.

Dabei heißen $\forall$ Allquantor und $\exists$ Existenzquantor. $\forall x \in A : \mathcal{A}(x)$ und $\exists x \in A : \mathcal{A}(x)$ sind selbst auch wieder Aussagen.

Für die Negation von Aussagen gelten die folgenden Regeln:

Negationen von Aussagen:

$$\neg(\mathcal{A} \wedge \mathcal{B}) \Leftrightarrow \neg\mathcal{A} \vee \neg\mathcal{B}, \qquad \neg\forall x \in A : \mathcal{A}(x) \Leftrightarrow \exists x \in A : \neg\mathcal{A}(x),$$
$$\neg(\mathcal{A} \vee \mathcal{B}) \Leftrightarrow \neg\mathcal{A} \wedge \neg\mathcal{B}, \qquad \neg\exists x \in A : \mathcal{A}(x) \Leftrightarrow \forall x \in A : \neg\mathcal{A}(x).$$

Beispiel.

1. Die Aussage ,,3 ist kleiner als 6" ist wahr.

2. Die Aussage ,,Jede gerade natürliche Zahl größer als 2 ist Summe von zwei Primzahlen" ist wahr oder falsch. Allerdings ist bis heute nicht bekannt, welcher der beiden Wahrheitswerte zutrifft.

3. Die Aussagen ,,p ist ein Primteiler von 6" und ,,$p = 2 \vee p = 3$" sind äquivalent.

4. Ist $\mathbb{Q}$ die Menge der rationalen Zahlen, so ist die Aussage

$$\forall q \in \mathbb{Q}\ \exists r \in \mathbb{Q} : r + r = q$$

wahr, denn sie besagt, daß jede rationale Zahl durch 2 teilbar ist. Falsch dagegen ist die Aussage

$$\forall q \in \mathbb{Q}\ \exists r \in \mathbb{Q} : r \cdot r = q,$$

denn sie besagt, daß jede rationale Zahl ein Quadrat ist. Wahr ist daher die Negation

$$\exists q \in \mathbb{Q}\ \forall r \in \mathbb{Q} : r \cdot r \neq q,$$

die besagt, daß es eine rationale Zahl gibt, die kein Quadrat in $\mathbb{Q}$ ist. Um dieses zu zeigen, braucht lediglich eine rationale Zahl angegeben zu werden, die in $\mathbb{Q}$ kein Quadrat ist, zum Beispiel 2.

Beweismethoden im Zusammenhang mit logischer Äquivalenz:

1. Der Widerspruchsbeweis beruht auf der Tatsache, daß $\mathcal{A} \Rightarrow \mathcal{B}$ und $\neg(\mathcal{A} \wedge \neg\mathcal{B})$ äquivalent sind (das kann z.B. leicht mit Hilfe der oben angegebenen Tabellen überprüft werden). Statt $\mathcal{A} \Rightarrow \mathcal{B}$ zu beweisen zeigt man, daß $\mathcal{A} \wedge \neg\mathcal{B}$ falsch ist. Es wird also $\mathcal{A} \wedge \neg\mathcal{B}$ zum Widerspruch geführt.

2. Die Kontraposition beruht auf der Tatsache, daß $\mathcal{A} \Rightarrow \mathcal{B}$ und $\neg\mathcal{B} \Rightarrow \neg\mathcal{A}$ äquivalent sind (auch das kann z.B. leicht mit Hilfe der oben angegebenen Tabellen überprüft werden). Statt $\mathcal{A} \Rightarrow \mathcal{B}$ zu beweisen zeigt man, daß $\neg\mathcal{B} \Rightarrow \neg\mathcal{A}$ wahr ist.

Beispiele hierzu werden in den folgenden Abschnitten behandelt.

2. Mengen

Ist M eine Menge und x ein Element von M, so schreibt man $x \in M$, anderenfalls $x \notin M$. Beispiele für Mengen sind:

$\emptyset$ die leere Menge,
$\mathbb{N} = \{1, 2, \ldots\}$ die Menge der natürlichen Zahlen,
$\mathbb{Z}$ die Menge der ganzen Zahlen,
$\mathbb{Q}$ die Menge der rationalen Zahlen,
$\mathbb{R}$ die Menge der reellen Zahlen,
$\mathbb{C}$ die Menge der komplexen Zahlen.

Die leere Menge ist dadurch charakterisiert, daß sie keine Elemente hat. Eine Menge A heißt Teilmenge einer Menge M, wenn jedes Element von A auch Element von M ist. Man schreibt dann $A \subseteq M$:

$$A \subseteq M \Leftrightarrow \forall x \in A : x \in M.$$

Gilt $A \subseteq M$ und $A \neq M$, so heißt A echte Teilmenge von M, geschrieben: $A \subset M$. Für jede Menge M gilt $M \subseteq M$ und $\emptyset \subseteq M$.

Ist für alle Elemente x aus M die Aussage $\mathcal{A}(x)$ definiert, so ist

$$A := \{x \in M | \mathcal{A}(x)\}$$

eine Teilmenge von M, und jedes x aus M ist genau dann in A, wenn $\mathcal{A}(x)$ wahr ist.

Beispiel. $A = \{n \in \mathbb{N} \mid n \text{ ist ein Quadrat und } n \leq 10\} = \{1, 4, 9\}$.

Definition 2.1 *Sind A und B Teilmengen der Menge M, so heißen*

$A \cup B := \{x \in M \mid x \in A \vee x \in B\}$ *Vereinigung,*
$A \cap B := \{x \in M \mid x \in A \wedge x \in B\}$ *Durchschnitt und*
$A \setminus B := \{x \in M \mid x \in A \wedge x \notin B\}$ *Differenz von A und B.*
$\mathcal{C}A := M \setminus A$ *heißt Komplement von A.*

Satz 2.2 *Sind A und B Teilmengen der Menge M, so gilt*

i) $\mathcal{C}(A \cup B) = \mathcal{C}A \cap \mathcal{C}B$.
ii) $\mathcal{C}(A \cap B) = \mathcal{C}A \cup \mathcal{C}B$.

Beweis. i) Für jedes x aus M gilt:
$x \in \mathcal{C}A \cap \mathcal{C}B \Longleftrightarrow x \in \mathcal{C}A \wedge x \in \mathcal{C}B \Longleftrightarrow x \notin A \wedge x \notin B \Longleftrightarrow$
$\neg(x \in A \vee x \in B) \Longleftrightarrow \neg(x \in A \cup B) \Longleftrightarrow x \in \mathcal{C}(A \cup B)$.
Für jedes $x \in M$ folgt somit $x \in \mathcal{C}A \cap \mathcal{C}B \Longleftrightarrow x \in \mathcal{C}(A \cup B)$, d.h. $\mathcal{C}A \cap \mathcal{C}B = \mathcal{C}(A \cup B)$.
ii) beweist man analog.

$\square$

Definition 2.3 *I sei eine Indexmenge und für jedes $i \in I$ sei A_i eine Teilmenge der Menge M. Dann heißen*

$$\bigcup_{i \in I} A_i := \{x \in M \mid \exists i \in I : x \in A_i\} \quad \textit{Vereinigung und}$$

$$\bigcap_{i \in I} A_i := \{x \in M \mid \forall i \in I : x \in A_i\} \quad \textit{Durchschnitt}$$

der $A_i, i \in I$.

Beispiel.
Sei $I = \mathbb{N}$ sowie $A_i = \{x \in \mathbb{Q} \mid 0 < x < \frac{1}{i}\}$ und $B_i = \{x \in \mathbb{Q} \mid 0 \leq x < \frac{1}{i}\}$.
Dann gilt $\bigcap\limits_{i \in \mathbb{N}} A_i = \emptyset$ und $\bigcap\limits_{i \in \mathbb{N}} B_i = \{0\}$.

A, B seien Mengen und $a \in A, b \in B$. Dann heißt (a, b) geordnetes Paar. Sind $a, a' \in A$ und $b, b' \in B$, so definiert man

$$(a, b) = (a', b') :\Longleftrightarrow a = a' \wedge b = b',$$

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

$A \times B$ ist eine Menge und heißt Produktmenge von A und B. Entsprechend definiert man

$$A_1 \times \ldots \times A_n = \{(a_1, \ldots, a_n) | a_1 \in A_1, \ldots, a_n \in A_n\}$$

als Produkt der Mengen $A_1, \ldots, A_n$. Die Elemente $(a_1, \ldots, a_n)$ von $A_1 \times \ldots \times A_n$ heißen n-Tupel .

$$A^n := A \times \ldots \times A = \{(a_1, \ldots, a_n) \mid a_1, \ldots, a_n \in A\}.$$

Definition 2.4 *Ist A eine Menge, so heißt die Menge aller Teilmengen von A die Potenzmenge von A, geschrieben: $\wp(A)$.*

Beispiel.

1. $A = \emptyset$. $\wp(A) = \{\emptyset\}$.
2. $A = \{a\}$. $\wp(A) = \{\emptyset, \{a\}\}$.
3. $A = \{a, b\}$. $\wp(A) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.
4. $A = \{a, b, c\}$. $\wp(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$.

Definition 2.5 *Ist A eine Menge, so bezeichnet $|A|$ die Anzahl der Elemente von A. Hat A unendlich viele Elemente, so schreibt man $|A| = \infty$.*

Satz 2.6 *Ist A eine endliche Menge mit n Elementen, so hat $\wp(A)$ genau 2^n Elemente:*

$$|A| = n \Rightarrow |\wp(A)| = 2^n.$$

Bevor wir Satz 2.6 durch vollständige Induktion beweisen, soll diese Beweismethode kurz vorgestellt werden.

Beweis durch vollständige Induktion

Ist $\mathcal{A}(n)$ für alle $n \in \mathbb{Z},\ n \geq n_0$ eine Aussage, so ist $\mathcal{A}(n)$ für alle $n \in \mathbb{Z},\ n \geq n_0$ wahr, wenn folgendes gilt:

(i) $\mathcal{A}(n_0)$ (Induktionsanfang).
(ii) $\forall n \geq n_0 : (\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1))$ (Induktionsschluß).

Beispiel.
$\mathcal{A}(n) : 1 + 2 + \ldots + n = \frac{1}{2}n(n+1), n \in \mathbb{N}$.
Behauptung: Für alle $n \in \mathbb{N}$ gilt $\mathcal{A}(n)$.
Induktionsanfang: $\mathcal{A}(1) : 1 = \frac{1}{2} \cdot 1 \cdot (1+1)$.
Induktionsschluß: $1 + 2 + \ldots + n + n + 1 \underset{\uparrow \mathcal{A}(n)}{=} \frac{1}{2}n(n+1) + n + 1 = \frac{1}{2}(n+1)(n+1+1)$.

Beweis von Satz 2.6: $\mathcal{A}(n) : |A| = n \Rightarrow |\wp(A)| = 2^n, n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
Induktionsanfang: $\mathcal{A}(0)$.
$|A| = 0 \Rightarrow A = \emptyset \Rightarrow \wp(A) = \{\emptyset\} \Rightarrow |\wp(A)| = 1 = 2^0$.
Induktionsschluß: $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$.
$|A| = n+1$, etwa $A = \{a_1, a_2, \ldots, a_n, a_{n+1}\}$. Definiere $A' := \{a_1, \ldots, a_n\}$, also $|A'| = n$.
Nach Induktionsvoraussetzung gilt $|\wp(A')| = 2^n$, etwa $\wp(A') = \{M_1, M_2, \ldots, M_{2^n}\}$.
Die M_i sind genau die Teilmengen von A, die a_{n+1} nicht enthalten. Es folgt

$$\wp(A) = \{M_1, M_2, \ldots, M_{2^n}, M_1 \cup \{a_{n+1}\}, M_2 \cup \{a_{n+1}\}, \ldots, M_{2^n} \cup \{a_{n+1}\}\},$$

d.h. $|\wp(A)| = 2 \cdot 2^n = 2^{n+1}$.

$\square$

3. Abbildungen

Sind A und B Mengen, so heißt eine Vorschrift f, die jedem a aus A genau ein b aus B zuordnet, eine Abbildung oder Funktion von A nach B, geschrieben

$$f : A \longrightarrow B.$$

Wird a aus A das Element b aus B zugeordnet, so schreibt man

$$a \longmapsto b \qquad \text{oder} \qquad f(a) = b.$$

A heißt Definitionsbereich von f, und B heißt Bildbereich von f.
Zwei Funktionen $f : A \to B,\quad f' : A' \to B'$ heißen gleich, wenn $A = A', B = B'$ und $f(a) = f'(a)$ für alle $a \in A$ gilt.

Beispiel.

1. $f : \mathbb{Z} \to \mathbb{Z}, x \mapsto 1 + x^2, \quad g : \mathbb{Z} \to \mathbb{N}, x \mapsto 1 + x^2$. Da f und g verschiedene Bildbereiche haben, sind f und g verschieden.

2. $f : \mathbb{Z} \to \mathbb{Z}, x \mapsto \frac{1}{2}(1 - (-1)^x)$.

   $$g : \mathbb{Z} \to \mathbb{Z}, x \mapsto \begin{cases} 0 & falls \ \ x \quad gerade \\ 1 & \quad ungerade. \end{cases}$$

   f und g sind gleich.

Definition 3.1 *Ist $f : A \to B$ eine Abbildung, so heißt*

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

das Bild von f, und f heißt surjektiv, wenn $f(A) = B$, d.h., wenn es zu jedem $b \in B$ ein $a \in A$ mit $f(a) = b$ gibt.

Beispiel.
$f : \mathbb{R} \longrightarrow \mathbb{R}, \ x \longmapsto x^2$. Dann ist $f(\mathbb{R}) = \mathbb{R}_0^+ := \{x \in \mathbb{R} | x \geq 0\}$ das Bild von f, und f ist nicht surjektiv. Hier wurde die Eigenschaft von $\mathbb{R}$ benutzt, daß ein $a \in \mathbb{R}$ genau dann ein Quadrat in $\mathbb{R}$ ist, wenn $a \geq 0$.

$g : \mathbb{R} \longrightarrow \mathbb{R}, \ x \mapsto x^3$. Dann ist $g(\mathbb{R}) = \mathbb{R}$ das Bild von g, d.h., g ist surjektiv. Für die Berechnung von $g(\mathbb{R})$ wurde ausgenutzt, daß jede Gleichung $x^3 = a$ mit $a \in \mathbb{R}$ in $\mathbb{R}$ lösbar ist.

Definition 3.2 *Ist $f : A \to B$ eine Abbildung, so heißt f injektiv, wenn für alle $a, a' \in A$ gilt:*

$$f(a) = f(a') \Longrightarrow a = a'.$$

Beispiel.
$f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ist nicht injektiv, da $f(1) = f(-1)$, aber $1 \neq -1$.

$g : \mathbb{R} \to \mathbb{R}, x \mapsto x^3$ ist injektiv:
Zunächst gilt für alle $a \in \mathbb{R} : \ \ a^3 = 1 \Rightarrow a = 1$, denn es gilt $a^3 \leq 0$ falls $a \leq 0$ und $1 > a > a^2 > a^3$ falls $0 < a < 1$ sowie $1 < a < a^2 < a^3$ falls $1 < a$.

Wir zeigen nun, daß g injektiv ist:
Gilt $x^3 = y^3$ und $x = 0$, dann folgt $x^3 = 0$, also $y^3 = 0$ und $y = 0 = x$.
Gilt $x^3 = y^3$ und $x \neq 0$, dann folgt $1 = \left(\frac{y}{x}\right)^3$, also $\frac{y}{x} = 1$, d.h. $x = y$.

Für den Nachweis der Injektivität von g wurden wesentlich die Ordnungseigenschaften von $\mathbb{R}$ ausgenutzt.

Definition 3.3 *Ist $f : A \to B$ eine Abbildung, so heißt f bijektiv, wenn f injektiv und surjektiv ist.*

Zum Beispiel ist $g : \mathbb{R} \longrightarrow \mathbb{R}, x \longmapsto x^3$ bijektiv.

Die Hintereinanderschaltung (Komposition) von Abbildungen.

Sind
$$f : A \longrightarrow B \qquad \text{und} \quad g : B \longrightarrow C$$
Abbildungen, so definiert man
$$g \circ f : A \longrightarrow C, a \longmapsto g(f(a)).$$
$g \circ f$ heißt Komposition von g und f, und $g \circ f$ ist auch definiert, wenn $f(A)$ *nur* im Defintionsbereich von g enthalten ist.

Beispiel.
$f : \mathbb{Z} \longrightarrow \mathbb{N}_0, z \longmapsto z^2.$
$g : \mathbb{N}_0 \longrightarrow \mathbb{Z}, n \mapsto 2n - 9,$
$g \circ f : \mathbb{Z} \longrightarrow \mathbb{Z}, z \longmapsto 2z^2 - 9,$
$f \circ g : \mathbb{N}_0 \longrightarrow \mathbb{N}_0, n \longmapsto 4n^2 - 36n + 81.$
Im allgemeinen gilt $g \circ f \neq f \circ g$.

Satz 3.4 *Sind $f : A \to B, g : B \to C$ Abbildungen, so gilt*
i) f, g surjektiv $\Longrightarrow g \circ f$ surjektiv.
ii) f, g injektiv $\Longrightarrow g \circ f$ injektiv.
iii) f, g bijektiv $\Longrightarrow g \circ f$ bijektiv.

Beweis. i) Zu jedem $c \in C$ gibt es $b \in B$ mit $g(b) = c$, da g surjektiv. Zu b gibt es $a \in A$ mit $f(a) = b$, da f surjektiv. Also $(g \circ f)(a) = g(f(a)) = g(b) = c$.
ii) Gilt $(g \circ f)(a) = (g \circ f)(a')$, so folgt $g(f(a)) = g(f(a'))$ und $f(a) = f(a')$, da g injektiv ist. Da f injektiv ist, folgt $a = a'$.
iii) folgt aus i) und ii).

$\square$

Satz 3.5 *$f : A \to B$ ist bijektiv $\Longleftrightarrow$ Es gibt $g : B \to A$ mit $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$.*

Bemerkung. Für jede Menge C ist $\mathrm{id}_C : C \longrightarrow C, c \longmapsto c$ die identische Abbildung von C.

Beweis $'' \Rightarrow''$: Da f surjektiv ist, existiert zu jedem $b \in B$ ein $a \in A$ mit $f(a) = b$, und da f injektiv ist, ist a eindeutig bestimmt. $g : B \to A$ sei nun die Funktion, die jedem $b = f(a)$ genau dieses a zuordnet.
$(g \circ f)(a) = a$ folgt sofort nach Definition von g, d.h. $g \circ f = \mathrm{id}_A$.
Jedes $b \in B$ läßt sich in der Form $b = f(a)$ schreiben, also $(f \circ g)(b) = (f \circ g)(f(a)) = f(g(f(a)) = f(a) = b$, d.h. $f \circ g = \mathrm{id}_B$.
$'' \Leftarrow''$: Für jedes $b \in B$ gilt $b = \mathrm{id}_B(b) = f(g(b)) \in f(A)$, d.h. f ist surjektiv.
Für $a, a' \in A$ gilt: $f(a) = f(a') \Rightarrow g(f(a)) = g(f(a')) \Rightarrow a = a'$, d.h. f ist injektiv.

$\square$

4. Aufgaben

A 4.1 Zeigen Sie, daß die Aussagen $\mathcal{A} \Rightarrow \mathcal{B}$ und $\neg(\mathcal{A} \wedge \neg\mathcal{B})$ bzw. $\mathcal{A} \Rightarrow \mathcal{B}$ und $\neg\mathcal{B} \Rightarrow \neg\mathcal{A}$ für beliebige Aussagen $\mathcal{A}, \mathcal{B}$ äquivalent sind. Erläutern Sie dieses Ergebnis.

A 4.2 Formalisieren Sie die folgende Aussage: *Ist das Produkt zweier rationaler Zahlen negativ, so ist genau einer dieser Faktoren negativ.*

A 4.3 Formalisieren Sie folgende Aussage: *Zu beliebigen reellen Zahlen a und b existiert eine natürliche Zahl n, so daß b kleiner ist als na.* Begründen Sie, daß die Aussage falsch ist, und bilden Sie ihre Negation.

A 4.4 Sind A und B Mengen, so definiert man $A\Delta B := (A \setminus B) \cup (B \setminus A)$. Zeigen Sie, daß für beliebige Teilmengen A, B und C einer Menge M gilt: $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.

A 4.5 Im folgenden sind A und B Teilmengen der Menge M mit $A \subseteq B$. Zeigen Sie, daß für jede Teilmenge X von M gilt: $X = (X \cap (M \setminus A)) \cup (X \cap B)$.

A 4.6 Es seien $f : A \to B$ und $g : B \to C$ zwei Abbildungen. Zeigen Sie:
(a) Wenn g injektiv ist und $g \circ f : A \to C$ surjektiv, dann ist f surjektiv.
(b) Wenn f surjektiv ist und $g \circ f$ injektiv, dann ist g injektiv.

A 4.7 Gegeben sind die Mengen N und M sowie die Abbildung $f : N \longrightarrow M$. Zeigen Sie, daß für beliebige Mengen $A, B \subseteq N$ gilt:
(a) $f(A \cup B) = f(A) \cup f(B)$.
(b) $f(A \cap B) \subseteq f(A) \cap f(B)$.
(c) Ist f injektiv, so gilt in (b) das Gleichheitszeichen.

A 4.8 (a) Zeigen Sie $2n + 1 < 2^n$ für alle $n \in \mathbb{N}, n \geq 3$.
(b) Für welche natürlichen Zahlen n gilt $n^2 < 2^n$? Benutzen Sie vollständige Induktion.

A 4.9 Beweisen oder widerlegen Sie (benutzen Sie gegebenenfalls vollständige Induktion):
(a) Für alle natürlichen Zahlen $n, n \geq 3$ gilt $(n + 1)! < n^n$.
(b) Für alle natürlichen Zahlen n gilt

$$\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \ldots + \frac{n}{n+1} = \frac{5n^2 - 9n + 5}{n(n+1)}.$$

(c) Für alle natürlichen Zahlen n gilt

$$\frac{1^2}{1 \cdot 3} + \frac{2^2}{3 \cdot 5} + \frac{3^2}{5 \cdot 7} + \ldots + \frac{n^2}{(2n-1)(2n+1)} = \frac{n(n+1)}{2(2n+1)}.$$

A 4.10 Für $n \in \mathbb{N}_0$ definiert man $n!$ rekursiv durch $0! := 1$ und $(n + 1)! := n! \cdot (n + 1)$. Zeigen Sie, daß für alle $a, b \in \mathbb{R}$ und alle $n \in \mathbb{Z}, n \geq 0$ gilt

$$(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \cdot b^{n-k},$$

wobei $\binom{n}{k} := \frac{n!}{k!(n-k)!}$ für alle $n, k \in \mathbb{Z}, 0 \leq k \leq n$.

KAPITEL 2

Gruppen, Ringe, Körper

1. Gruppen

Ist A eine Menge, so heißt eine Abbildung

$$A \times A \longrightarrow A, \quad (a, a') \longmapsto a \circ a'$$

Verknüpfung auf A. Eine Verknüpfung heißt...

... assoziativ, wenn für alle $a, b, c \in A$ gilt:

$$a \circ (b \circ c) = (a \circ b) \circ c.$$

... kommutativ, wenn für alle $a, b \in A$ gilt:

$$a \circ b = b \circ a.$$

Beispiel.

1. $A = \mathbb{R}$. Die Addition und Multiplikation sind assoziative und kommutative Verknüpfungen auf $\mathbb{R}$. Die Subtraktion $a \circ b = a - b$ ist weder assoziativ noch kommutativ:

$$0 \circ (0 \circ 1) = 0 - (0 - 1) = 1 \neq -1 = (0 - 0) - 1 = (0 \circ 0) \circ 1.$$

$$1 \circ 0 = 1 - 0 = 1 \neq -1 = 0 - 1 = 0 \circ 1.$$

2. Ist M eine Menge und A die Menge der Abbildungen $f : M \to M$, so ist die Komposition eine Verknüpfung auf A. Die Komposition ist sogar eine assoziative Verknüpfung, denn für $f, g, h : M \to M$ gilt:
   $f \circ (g \circ h) : M \to M, m \mapsto f(g(h(m)))$.
   $(f \circ g) \circ h : M \to M, m \mapsto f(g(h(m)))$.

Bemerkung. Grundsätzlich sind Ausdrücke der Form $a_1 \circ a_2 \circ \ldots \circ a_n$ nicht definiert, da stets nur zwei Elemente miteinander verknüpft werden können, z.B.:

$$(a_1 \circ a_2) \circ (a_3 \circ (a_4 \circ a_5)) \ldots$$

Es gilt jedoch

Satz 1.1 *Bei einer assoziativen Verknüpfung ist die Verknüpfung von n Elementen unter Beibehaltung der Reihenfolge unabhängig von der Klammerung.*

Beweis. Für jedes $n \in \mathbb{N}$ ist $\mathcal{A}(n)$ die Aussage: Jedes Produkt der Elemente $a_1, \ldots, a_n$ ist unter Beibehaltung der Reihenfolge gleich dem Ausdruck

$$(\ldots((a_1 \circ a_2) \circ a_3) \ldots) \circ a_n.$$

$\mathcal{A}(1) : a_1 = a_1.$
$\mathcal{A}(2) : a_1 \circ a_2 = a_1 \circ a_2.$
$\mathcal{A}(3) : (a_1 \circ a_2) \circ a_3 = (a_1 \circ a_2) \circ a_3 \wedge a_1 \circ (a_2 \circ a_3) = (a_1 \circ a_2) \circ a_3.$
$n \geq 3$: Sei $P = (a_1 \ldots a_k) \circ (a_{k+1} \ldots a_{n+1})$ die Verknüpfung von $n+1$ Elementen. Gilt $k = n$, so folgt nach Induktionsvoraussetzung:

$$P = (\ldots((a_1 \circ a_2) \circ a_3) \ldots) \circ a_{n+1}.$$

Gilt $k < n$, so folgt nach Induktionsvoraussetzung:

$$\begin{aligned} P &= (a_1 \ldots a_k) \circ ((a_{k+1} \ldots a_n) \circ a_{n+1}) \\ &= ((a_1 \ldots a_k) \circ (a_{k+1} \ldots a_n)) \circ a_{n+1} \\ &= (\ldots((a_1 \circ a_2) \circ a_3) \ldots) \circ a_{n+1}. \end{aligned}$$

□

Bemerkung. Bei einer assoziativen Verknüpfung können die Klammern weggelassen werden.

Definition 1.2 *Ist A eine Menge mit einer Verknüpfung "$\circ$", so heißt $e \in A$ neutrales Element oder Einselement , wenn $e \circ a = a \circ e = a$ für alle $a \in A$ gilt.*

Bemerkung. Gibt es ein neutrales Element e, so ist es eindeutig bestimmt: Ist e' auch ein neutrales Element, so gilt $e' = e' \circ e = e$.

Beispiel.

1. $A = \mathbb{Q}$. Dann ist 0 das neutrale Element der Addition und 1 das neutrale Element der Multiplikation.

2. Ist M eine Menge und A die Menge der Abbildungen $f : M \to M$, so ist

   $$\mathrm{id}_M : M \longrightarrow M, m \longmapsto m$$

   das neutrale Element bezüglich der Komposition, denn für alle $f \in A$ und $m \in M$ gilt $(f \circ \mathrm{id}_M)(m) = f(\mathrm{id}_M(m)) = f(m)$ und $(\mathrm{id}_M \circ f)(m) = \mathrm{id}_M(f(m)) = f(m)$, d.h. $f \circ \mathrm{id}_M = \mathrm{id}_M \circ f = f$.

Bemerkung. Eine Menge $A \neq \emptyset$ mit einer assoziativen Verknüpfung heißt Halbgruppe. Eine Halbgruppe mit Einselement heißt Monoid.

Definition 1.3 *Es sei A eine Menge mit einer Verknüpfung "$\circ$" und einem neutralen Element e. Ist $a \in A$, so heißt $b \in A$ inverses Element oder Inverses von a, wenn $a \circ b = b \circ a = e$ gilt. Existiert zu a ein Inverses, so heißt a invertierbar.*

Beispiel. Bezüglich der Addition hat $\mathbb{Z}$ das neutrale Element 0, und jedes $a \in \mathbb{Z}$ ist invertierbar. Bezüglich der Multiplikation ist 1 das neutrale Element, und nur 1 und -1 sind invertierbar.

Satz 1.4 *Ist A eine Menge mit einer assoziativen Verknüpfung $''\circ''$ und einem neutralen Element e, so gibt es zu jedem invertierbaren $a \in A$ genau ein $b \in A$ mit $a \circ b = e$ ($b \circ a = e$).*

Beweis. Ist a' ein Inverses von a, so gilt $b = e \circ b = (a' \circ a) \circ b = a' \circ (a \circ b) = a' \circ e = a'$. Die zweite Behauptung folgt entsprechend.

□

Bemerkung. Das unter den Voraussetzungen von Satz 1.4 eindeutig bestimmte Inverse von a wird mit a^{-1} bezeichnet.

Definition 1.5 *Ist G eine Menge und $''\circ''$ eine Verknüpfung auf G, so heißt G Gruppe (bezüglich $''\circ''$), wenn gilt:*
i) $''\circ''$ ist assoziativ.
ii) Es existiert ein neutrales Element $e \in G$.
iii) Jedes $g \in G$ ist invertierbar.
Ist weiterhin $''\circ''$ kommutativ, so heißt G abelsche Gruppe.

Beispiel.

1. $\mathbb{Z}$ ist bezüglich $''+''$ eine abelsche Gruppe.

2. $\mathbb{R}$ ist bezüglich $''\cdot''$ keine Gruppe, denn 0 ist nicht invertierbar. $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ ist bezüglich $''\cdot''$ eine abelsche Gruppe.

Satz 1.6 *Ist M eine Menge mit einer assoziativen Verknüpfung $''\circ''$ und einem neutralen Element e, so ist die Menge G der invertierbaren Elemente von M bezüglich $''\circ''$ eine Gruppe.*

Beweis. o) Zunächst zeigen wir: $\forall a, b \in G : a \circ b \in G$. Sind a, b aus G, so folgt $a^{-1}, b^{-1} \in M$. Wegen

$$(a \circ b) \circ (b^{-1} \circ a^{-1}) = a \circ (b \circ b^{-1}) \circ a^{-1} = a \circ e \circ a^{-1} = a \circ a^{-1} = e$$

und

$$(b^{-1} \circ a^{-1}) \circ (a \circ b) = b^{-1} \circ (a^{-1} \circ a) \circ b = b^{-1} \circ e \circ b = b^{-1} \circ b = e$$

ist $a \circ b$ invertierbar in M, also $a \circ b \in G$.
i) $a \circ (b \circ c) = (a \circ b) \circ c$ gilt sogar für alle $a, b, c \in M$.
ii) Wegen $e \circ e = e$ gilt $e \in G$, und e ist damit das neutrale Element von G.
iii) Für alle $a \in G$ gilt $a \circ a^{-1} = a^{-1} \circ a = e$. Also ist a^{-1} in M invertierbar, d.h. $a^{-1} \in G$, und a^{-1} ist das Inverse von a.

□

Bemerkung. Für invertierbare a, b wurde sogar $(a^{-1})^{-1} = a$ sowie $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$ gezeigt.

Beispiel. M sei eine nichtleere Menge und

$$\mathrm{Abb}(M) = \{f | f : M \to M\}$$

die Menge aller Abbildungen $f : M \to M$. Die Komposition ist eine Verknüpfung auf $\mathrm{Abb}(M)$, die assoziativ ist, und $\mathrm{id}_M : M \to M, m \mapsto m$ ist das neutrale Element. Wegen Satz 1.6 ist die Menge der invertierbaren Abbildungen $f : M \to M$ bezgl. der Komposition eine Gruppe, die mit $\mathbf{S_M}$ bezeichnet wird. Gemäß Definition ist $f \in \mathrm{Abb}(M)$ genau dann aus $\mathbf{S_M}$, wenn es ein $g : M \to M$ mit $f \circ g = \mathrm{id}_M$ und $g \circ f = \mathrm{id}_M$ gibt. Wegen Satz 3.5 aus Kapitel 1 folgt somit

$$\mathbf{S_M} = \{f | f : M \to M \text{ bijektiv }\}.$$

$\mathbf{S_M}$ heißt symmetrische Gruppe von M. Ist speziell $M = \{1, \ldots, n\}$, so schreibt man $\mathbf{S_n}$ statt $\mathbf{S_M}$. Die Elemente von $\mathbf{S_M}$ heißen Permutationen.

Darstellung von Permutationen. Für $\pi \in \mathbf{S_n}$, also $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ bijektiv, verwendet man auch die Bezeichnung

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

Ist zum Beispiel $n = 3$ und $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, so gilt $\pi(1) = 3, \pi(2) = 1, \pi(3) = 2$ und $\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$.

Da es genau $n!$ Möglichkeiten gibt, die Zahlen $1, 2, \ldots, n$ anzuordnen, folgt unter Benutzung obiger Darstellung

Satz 1.7 *Für alle $n \in \mathbb{N}$ gilt $|\mathbf{S_n}| = n!$.*

Definition 1.8 *Eine Permutation $\pi \in \mathbf{S_n}$ heißt Zykel der Länge r, wenn es paarweise verschiedene $k_1, \ldots, k_r$ aus $\{1, \ldots, n\}$ gibt, so daß $\pi(k_1) = k_2, \pi(k_2) = k_3, \ldots, \pi(k_r) = k_1$ und $\pi(k) = k$ für $k \neq k_i, i = 1, \ldots, r$ gilt.*

Bemerkung.

1. Zykeln der Länge 2 heißen Transpositionen.
2. Ist π ein Zykel wie in Definition 1.8, so schreibt man $\pi = (k_1 k_2 \ldots k_r)$, und es gilt dann $\pi = (k_2 k_3 \ldots k_r k_1) = (k_3 k_4 \ldots k_r k_1 k_2) = \ldots$

Beispiel. Für

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 5 & 1 & 7 & 6 & 4 \end{pmatrix} \in \mathbf{S_7}$$

gilt $\pi(1) = 3, \pi(3) = 5, \pi(5) = 7, \pi(7) = 4, \pi(4) = 1$ und $\pi(2) = 2, \pi(6) = 6$. Also folgt $\pi = (13574)$. Kein Zykel dagegen ist

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}.$$

Bemerkung.

1. Die Identität von $\{1, \ldots, n\}$ (also das neutrale Element von $\mathbf{S_n}$) wird auch mit (1) bezeichnet.

2. Jede Permutation läßt sich als Produkt von (elementfremden) Zykeln schreiben, z. B. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 2 & 4 & 1 & 5 \end{pmatrix} = (165) \circ (23) \circ (4)$.

Definition 1.9 *Für π aus $\mathbf{S_n}$ heißt*

$$\mathrm{sgn}\pi := \prod_{j<i} \frac{\pi(i) - \pi(j)}{i - j}$$

Signum von π.

Bemerkung.

1. Es gilt stets $\mathrm{sgn}\pi \in \{1, -1\}$, da im Zähler und Nenner bis auf das Vorzeichen und die Reihenfolge dieselben Faktoren vorkommen.

2. Gilt $a \neq b$, dann folgt $\mathrm{sgn}(ab) = -1$.
   Sei dazu o.B.d.A. $a = 1$ und $b = 2$, also $\pi = (12)$. Dann folgt

$$\mathrm{sgn}(12) = \prod_{j<i} \frac{\pi(i) - \pi(j)}{i - j} = \frac{\pi(2) - \pi(1)}{2 - 1} \cdot \prod_{3 \leq i} \frac{\pi(i) - \pi(1)}{i - 1} \frac{\pi(i) - \pi(2)}{i - 2} =$$
$$= \frac{1 - 2}{2 - 1} = -1.$$

Satz 1.10 *Für alle $\pi, \sigma \in \mathbf{S_n}$ gilt:*

$$\mathrm{sgn}(\pi \circ \sigma) = \mathrm{sgn}\pi \cdot \mathrm{sgn}\sigma.$$

Beweis.

$$\mathrm{sgn}(\pi \circ \sigma) = \prod_{j<i} \frac{\pi \circ \sigma(i) - \pi \circ \sigma(j)}{i - j} = \prod_{j<i} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(i) - \sigma(j)} \cdot \prod_{j<i} \frac{\sigma(i) - \sigma(j)}{i - j} = \mathrm{sgn}\pi \cdot \mathrm{sgn}\sigma.$$

□

Anwendung von Satz 1.10
Ist $\pi \in \mathbf{S_n}$ ein Zykel, etwa $\pi = (a_1 \ldots a_k)$, so folgt $\pi = (a_1 a_k) \circ (a_1 a_{k-1}) \circ \ldots \circ (a_1 a_3) \circ (a_1 a_2)$, und wegen Satz 1.10 gilt dann

$$\mathrm{sgn}(a_1 \ldots a_k) = (-1)^{k-1}.$$

Da jede Permutation Produkt von Zykeln ist, ist damit jede Permutation π Produkt von Transpositionen:

$$\pi = \tau_1 \circ \ldots \circ \tau_r,\ \tau_i \ \text{Transposition.}$$

Satz 1.10 liefert $\mathrm{sgn}\pi = (-1)^r$.

$$\pi \text{ heißt gerade } :\Longleftrightarrow \mathrm{sgn}\pi = 1.$$

$$\pi \text{ heißt ungerade } :\Longleftrightarrow \mathrm{sgn}\pi = -1.$$

π ist also genau dann gerade, wenn bei jeder Darstellung von π als Produkt von Transpositionen die Anzahl der Transpositionen gerade ist.

Beispiel. $\pi = (12387) \circ (6537)$. Dann folgt $\mathrm{sgn}\pi = (-1)^{4+3} = -1$. Somit ist π ungerade. Zum Beispiel gilt $\pi = (17) \circ (18) \circ (13) \circ (12) \circ (67) \circ (63) \circ (65)$.

Bemerkung. Zur Vereinfachung wird im folgenden das Verknüpfungssymbol $\circ$ zwischen den Permutationen weggelassen.

2. Ringe und Körper

Definition 2.1 *Ist R eine Menge mit den Verknüpfungen "$+$" und "$\cdot$", so heißt R Ring (bezüglich "$+$" und "$\cdot$"), wenn gilt:*
i) R ist bezüglich "$+$" eine abelsche Gruppe.
ii) "$\cdot$" ist assoziativ.
iii) Es gelten die Distributivgesetze, d.h., für alle $a, b, c \in R$ gilt

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c),\ (a + b) \cdot c = (a \cdot c) + (b \cdot c).$$

Ist weiterhin "$\cdot$" kommutativ, so heißt R kommutativer Ring .
Hat ein Ring R ein neutrales Element bezgl. der Multiplikation, so heißt R Ring mit Eins, und dann heißt weiterhin $a \in R$ invertierbar oder Einheit, wenn a bezgl. "$\cdot$" invertierbar ist.

Bemerkung.

1. Punktrechnung geht vor Strichrechnung, d.h., um Produkte werden keine Klammern gesetzt.

2. Das neutrale Element eines Ringes bezüglich "$+$" wird mit 0 bezeichnet.

3. Ist R ein Ring mit Eins, so wird das neutrale Element der Multiplikation mit 1 bezeichnet. In diesem Falle ist $\mathrm{E}(R) = \{a \in R \mid a \text{ Einheit in } R\}$ eine Gruppe bezüglich der Multiplikation, die Einheitengruppe von R genannt wird (vgl. Satz 1.6). Für jedes $a \in \mathrm{E}(R)$ ist a^{-1} das Inverse von a bezüglich der Multiplikation. Zum Beispiel ist $\mathbb{Z}$ bezüglich der gewöhnlichen Addition und Multiplikation ein Ring mit Eins und $\mathrm{E}(\mathbb{Z}) = \{1, -1\}$.

4. Für alle a aus R bezeichnet $-a$ das Inverse von a bezüglich "+". Für alle a, b aus R definiert man weiterhin $a - b := a + (-b)$.

5. Für alle a, b, c aus R gilt:
$a \cdot 0 = 0 \cdot a = 0$; $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$; $(-a) \cdot (-b) = a \cdot b$;
$a \cdot (b - c) = a \cdot b - a \cdot c$; $(a - b) \cdot c = a \cdot c - b \cdot c$.

Definition 2.2 *Ist K ein kommutativer Ring mit Eins, so heißt K Körper, wenn $1 \neq 0$ und jedes $a \neq 0$ eine Einheit ist.*

Beispiel. Bezüglich der gewöhnlichen Addition und Multiplikation sind $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ kommutative Ringe mit Eins. $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ sind sogar Körper, $\mathbb{Z}$ aber nicht.

Im folgenden werden weitere Beispiele für Ringe und Körper vorgestellt, die sich von den bisher angegebenen Beispielen in vieler Hinsicht wesentlich unterscheiden.

Die Ringe $\mathbb{Z}_n, n \in \mathbb{N}$

Sind a und b ganze Zahlen, so schreibt man $a \mid b$, wenn b von a geteilt wird, wenn es also eine ganze Zahl k mit $b = k \cdot a$ gibt. Für $a, b \in \mathbb{Z}$ definiert man nun

$$a \sim b \;:=\; n \mid (a - b).$$

" $\sim$" ist eine Äquivalenzrelation:
$a \sim a$, da $n \mid (a - a)$.
$a \sim b \Rightarrow n \mid (a - b) \Rightarrow n \mid (b - a) \Rightarrow b \sim a$.
$a \sim b \wedge b \sim c \Rightarrow n \mid (a - b) \wedge n \mid (b - c) \Rightarrow n \mid (a - b + b - c) \Rightarrow n \mid (a - c) \Rightarrow a \sim c$.
Für die Äquivalenzklasse $\overline{a}$, in der a ist, gilt somit

$$\overline{a} = \{b \in \mathbb{Z} \mid a \sim b\} = \{a + k \cdot n \mid k \in \mathbb{Z}\}.$$

Offenbar sind $\overline{0}, \overline{1}, \ldots, \overline{n-1}$ sämtliche "Restklassen modulo n". Die Menge aller Äquivalenzklassen bezeichnet man mit $\mathbb{Z}_n$, also

$$\mathbb{Z}_n := \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}.$$

$a \sim b$ kann man auch so formulieren: a und b lassen bei Division durch n denselben Rest.

Auf der Menge $\mathbb{Z}_n$ läßt sich nun folgendermaßen eine Addition und eine Multiplikation erklären:
Addition in $\mathbb{Z}_n : \overline{a} + \overline{b} = \overline{a + b}$.
Multiplikation in $\mathbb{Z}_n$: $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$.

Da die Verknüpfungen der Restklassen offenbar mit den jeweiligen Repräsentanten zusammenhängen, muß gezeigt werden, daß sie von den Repräsentanten unabhängig sind (Wohldefiniertheit):
Gilt $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$, so folgt $a \sim a'$ und $b \sim b'$, also $n \mid (a - a')$ und $n \mid (b - b')$. Damit ist n ein Teiler von $a + b - (a' + b')$, d.h. $a + b \sim a' + b'$ und $\overline{a + b} = \overline{a' + b'}$.

Weiterhin folgt $n \mid (a-a')b$ aus $n \mid (a-a')$ sowie $n \mid (b-b')a'$ aus $n \mid (b-b')$. Daraus ergibt sich schließlich $n \mid (ab - a'b')$, also $ab \sim a'b'$.

Man rechnet leicht nach, daß $\mathbb{Z}_n$ ein kommutativer Ring mit dem Einselement $\overline{1}$ ist.

Satz 2.3 *$\mathbb{Z}_n$ ist genau dann ein Körper, wenn n eine Primzahl ist.*

Beweis. Sei zunächst $n = p$ eine Primzahl und $\overline{a} \in \mathbb{Z}_p, \overline{a} \neq \overline{0}$. Zu zeigen ist, daß $\overline{a}$ eine Einheit in $\mathbb{Z}_p$ ist. Wegen $\overline{a} \neq \overline{0}$ ist p kein Teiler von a. Da p eine Primzahl ist, sind a und p teilerfremd, und es existieren $x, y \in \mathbb{Z}$ mit $ax + py = 1$, also $\overline{a} \cdot \overline{x} = \overline{1}$.
Wegen $\overline{1} \neq \overline{0}$ ist schließlich gezeigt, daß $\mathbb{Z}_p$ ein Körper ist.

Sei nun andererseits n keine Primzahl. Gilt $n = 1$, so folgt $\overline{0} = \overline{1}$, und $\mathbb{Z}_n$ ist kein Körper. Sei daher im folgenden $n > 1$. Dann gibt es $a, b \in \mathbb{N}$ mit $n = a \cdot b$ und $1 < a, b < n$, also $\overline{a}, \overline{b} \neq \overline{0}$ und $\overline{a} \cdot \overline{b} = \overline{0} \notin \mathrm{E}(\mathbb{Z}_n)$. Folglich ist $\overline{a}$ oder $\overline{b}$ keine Einheit und $\mathbb{Z}_n$ kein Körper.

$\square$

Satz 2.4 *In einem Körper K gilt folgende Alternative:*

i) $\underbrace{1 + \ldots + 1}_{n-\text{mal}} \neq 0$ *für alle* $n \in \mathbb{N}$. *ii)* $\underbrace{1 + \ldots + 1}_{p-\text{mal}} = 0$ *für genau eine Primzahl* p.

Beweis. Wir nehmen an, daß i) nicht gilt, und zeigen, daß dann ii) gilt.
Gilt i) nicht, so existiert ein $n \in \mathbb{N}$ mit $\underbrace{1 + \ldots + 1}_{n-\text{mal}} = 0$. Wähle $n \in \mathbb{N}$ mit dieser Eigenschaft minimal. Wir zeigen, daß n eine Primzahl ist. Sei $n = a \cdot b$ mit $a, b \in \mathbb{N}$ und $1 < a, b < n$. Dann folgt

$$\underbrace{(1 + \ldots + 1)}_{a-\text{mal}} \cdot \underbrace{(1 + \ldots + 1)}_{b-\text{mal}} = 0.$$

Da K ein Körper ist, muß der erste Faktor oder der zweite Faktor 0 sein, im Widerspruch zu $1 < a, b < n$ und der Minimalität von n.
Wir definieren $p := n$. Sei nun q ebenfalls eine Primzahl mit $\underbrace{(1 + \ldots + 1)}_{q-\text{mal}} = 0$. Dann folgt $q \geq p$ wegen der Minimalität von p, und es gibt $a \in \mathbb{N}, b \in \mathbb{N}_0, 0 \leq b < p$ mit $q = a \cdot p + b$.

$$\underbrace{(1 + \ldots + 1)}_{a-\text{mal}} \cdot \underbrace{(1 + \ldots + 1)}_{p-\text{mal}} + \underbrace{(1 + \ldots + 1)}_{b-\text{mal}} = 0.$$

Es folgt $\underbrace{1 + \ldots + 1}_{b-\text{mal}} = 0$ und wegen der Minimalität von p sogar $b = 0$. Also gilt $q = a \cdot p$, d.h. $q = p$, da q prim.

$\square$

Ist nun K ein Körper, für den die Aussage i) aus Satz 2.4 gilt, so definiert man $\chi(K) = 0$. Gilt dagegen Aussage ii) aus Satz 2.4, so definiert man $\chi(K) = p$. Man nennt $\chi(K)$ die Charakteristik von K. So gilt zum Beispiel

$$\chi(\mathbb{Q}) = \chi(\mathbb{R}) = \chi(\mathbb{C}) = 0 \quad \text{und} \quad \chi(\mathbb{Z}_p) = p.$$

Beispiel. Für $n = 7$ gilt $\overline{1}+\overline{1}+\overline{1}+\overline{1}+\overline{1}+\overline{1}+\overline{1} = \overline{0}$ und z.B. $\overline{4}+\overline{3} = \overline{7} = \overline{0}$, also $\overline{4} = -\overline{3}$. Diese Gleichung erhält man auch wegen $\overline{4} = \{4 + k \cdot 7 \mid k \in \mathbb{Z}\} = \{-3 + k \cdot 7 \mid k \in \mathbb{Z}\} = \overline{-3}$. Bezüglich der Multiplikation ist z.B. $\overline{3} \cdot \overline{5} = \overline{15} = \overline{1}$, also $\overline{3}^{-1} = \overline{5}$.

3. Aufgaben

A 3.1 Prüfen Sie, ob folgende Verknüpfungen $\circ$ auf $\mathbb{R}$ assoziativ bzw. kommutativ sind:

(a) $a \circ b := ab^2$ (b) $a \circ b := a + b + ab$ (c) $a \circ b := b$
(d) $a \circ b := (a + b)^2$ (e) $a \circ b := a + b - 1$.

A 3.2 Wieviele Verknüpfungen gibt es auf einer n-elementigen Menge? Wieviele davon sind kommutativ?

A 3.3 Stellen Sie die Permutation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 9 & 5 & 6 & 8 & 4 & 1 & 3 & 7 \end{pmatrix}$$

durch Zykeln dar. Berechnen Sie $f \circ g$, $g \circ f$ und $f^{-1} \circ g$ mit $g = (143)(29) \in \mathbf{S_9}$ sowie $\operatorname{sgn} f$.

A 3.4 Zeigen Sie, daß sich jede Permutation $\sigma \in \mathbf{S_n}$ bis auf die Reihenfolge eindeutig als Produkt von elementfremden Zykeln schreiben läßt.

A 3.5 Es sei $L = \mathbb{Q} \times \mathbb{Q}$. Zeigen Sie, daß L bezüglich der Verknüpfungen

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

kein Körper ist. Welches Axiom ist verletzt?

A 3.6 Es sei $L = \mathbb{Q} \times \mathbb{Q}$. Zeigen Sie, daß L bezüglich der Verknüpfungen

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b) \cdot (c, d) = (ac + 2bd, bc + ad)$$

ein Körper ist. (Hinweis: Benutzen Sie, daß in $\mathbb{Q}$ kein Element x mit $x^2 = 2$ existiert.)

A 3.7 Berechnen Sie das multiplikative Inverse von $\overline{13}$ in $\mathbb{Z}_{89}$.

A 3.8 Berechnen Sie die Einheitengruppe von $\mathbb{Z}_{45}$.

KAPITEL 3

Vektorräume

1. Grundlagen

Definition 1.1 *Sei* $(V, +)$ *eine abelsche Gruppe und* K *ein Körper.* V *heißt* K*-Vektorraum, wenn es eine Abbildung*

$$K \times V \longrightarrow V, \quad (k, v) \longmapsto k \cdot v$$

gibt, so daß für alle $k, l \in K;\ v, w \in V$ *gilt:*
i) $(k \cdot l) \cdot v = k \cdot (l \cdot v)$.
ii) $(k + l) \cdot v = k \cdot v + l \cdot v$.
iii) $k \cdot (v + w) = k \cdot v + k \cdot w$.
iv) $1 \cdot v = v$.

Bemerkung.

1. Die Addition in V und die Addition in K werden mit $''+''$ bezeichnet. Entsprechendes gilt für die Multiplikation und $''\cdot''$. Oft schreibt man kv statt $k \cdot v$ für $k \in K$ und $v \in V$.
2. Die Elemente eines Vektorraumes heißen Vektoren. Übliche Notationen für Vektoren sind $\vec{v}$ und $\underline{v}$.
3. Das neutrale Element in $(V, +)$ heißt Nullvektor und wird mit $\mathcal{O}$ bezeichnet.
4. K heißt Skalarenkörper, die Elemente von K heißen Skalare.
5. V heißt reeller (komplexer) Vektorraum, falls $K = \mathbb{R}$ $(K = \mathbb{C})$.

Beispiel. In den folgenden Beispielen ist K ein Körper mit der Addition $''+''$ und der Multiplikation $''\cdot''$.

1. Der Körper K selbst ist ein K-Vektorraum bezüglich der Verknüpfungen in K.
   $\mathbb{R}$ und $\mathbb{C}$ sind $\mathbb{Q}$-Vektorräume mit der üblichen Addition und Multiplikation.
   $\mathbb{C}$ ist ein reeller Vektorraum.

2. Die Menge $V = K^n = K \times K \times \ldots \times K$ der n-Tupel ist bezüglich der Verknüpfung

$$''+'' : (k_1, \ldots, k_n) + (l_1, \ldots, l_n) := (k_1 + l_1, \ldots, k_n + l_n)$$

eine abelsche Gruppe. Wird die Multiplikation mit Skalaren definiert durch

$$''\cdot'' : k \cdot (k_1, \ldots, k_n) := (k \cdot k_1, \ldots, k \cdot k_n),$$

dann ist V ein K-Vektorraum.

3. Sei $[a, b] \subseteq \mathbb{R}$ ein Intervall und $\mathcal{F} := \{f \mid f : [a, b] \to \mathbb{R}\}$.
Für $f, g \in \mathcal{F}$, $k \in \mathbb{R}$ definiert man

$$\begin{aligned} f + g : [a, b] &\longrightarrow \mathbb{R}, \quad x \longmapsto f(x) + g(x) \quad \text{und} \\ k \cdot f : [a, b] &\longrightarrow \mathbb{R}, \quad x \longmapsto k \cdot f(x)\,. \end{aligned}$$

Bezüglich dieser Verknüpfungen ist $\mathcal{F}$ ein reeller Vektorraum.

Einfache Eigenschaften. Ist V ein K-Vektorraum, dann gilt für alle $v \in V, k \in K$:

1. $0 \cdot v = \mathcal{O}$.
2. $k \cdot \mathcal{O} = \mathcal{O}$.
3. $k \cdot v = \mathcal{O} \implies (k = 0 \ \vee \ v = \mathcal{O})$.
4. $(-1) \cdot v = -v\,.$

Definition 1.2 *Ist V ein K-Vektorraum und $U \subseteq V$, dann heißt U Unterraum von V, wenn U bezüglich der für V definierten Addition und Multiplikation ein K-Vektorraum ist, geschrieben $U \subseteq\!| \, V$.*

Satz 1.3 *Sei V ein K-Vektorraum und $U \subseteq V$. Genau dann ist U ein Unterraum von V, wenn gilt*

i) $U \neq \emptyset$.

ii) $\forall u, v \in U \ : \ u + v \in U$.

iii) $\forall u \in U \ \forall k \in K \ : \ k \cdot u \in U$.

Beweis. $'' \Rightarrow''$: Ergibt sich direkt aus Definition 1.2.
$'' \Leftarrow''$: Zunächst wird gezeigt, daß aus i) - iii) folgt, daß $(U, +)$ eine abelsche Gruppe ist. Die Assoziativität und Kommutativität von $''+''$ gelten, da $(V, +)$ eine abelsche Gruppe ist. Mit iii) gilt $(-1) \cdot u = -u \in U$ für alle $u \in U$. Mit u sind also auch $-u$ und $u + (-u) = \mathcal{O} \in U$, das heißt, U besitzt ein neutrales Element und jedes $u \in U$ ist invertierbar. Da U mit iii) abgeschlossen ist bezüglich der Skalarenmultiplikation und die restlichen Axiome sich direkt von V auf U übertragen, folgt die Behauptung.

$\square$

Beispiel.

1. Jeder K-Vektorraum V hat die *trivialen* Unterräume $\{\mathcal{O}\}$ und V. Man nennt $\{\mathcal{O}\}$ auch den Nullraum von V. Für jedes $v \in V$ ist weiterhin $K \cdot v := \{k \cdot v \mid k \in K\}$ ein Unterraum von V, denn

   i) $1 \cdot v \in K \cdot v \quad \Longrightarrow \quad K \cdot v \neq \emptyset.$

   ii) $k \cdot v,\ l \cdot v \in K \cdot v \quad \Longrightarrow \quad k \cdot v + l \cdot v = (k + l) \cdot v \in K \cdot v.$

   iii) $k \in K,\ l \cdot v \in K \cdot v \quad \Longrightarrow \quad k \cdot (l \cdot v) = (k \cdot l) \cdot v \in K \cdot v.$

   Mit dieser Bezeichnung gilt $\{\mathcal{O}\} = K \cdot \mathcal{O}$.

2. Für $n \in \mathbb{N}_0$ sei $C^n([a, b]) := \{f : [a, b] \to \mathbb{R} \mid f$ ist n-mal stetig differenzierbar$\}$. Dann gilt
$$\ldots C^n([a, b]) \subseteq\mid \ldots \subseteq\mid C^1([a, b]) \subseteq\mid C^0([a, b]) \subseteq\mid \mathcal{F}.$$

3. Sei $\mathrm{Pol}([a, b])$ die Menge aller Polynomabbildungen $f : [a, b] \longrightarrow \mathbb{R}$, das heißt, jedes $f \in \mathrm{Pol}([a, b])$ läßt sich in der Form
$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0, \quad a_i \in \mathbb{R},\ n \in \mathbb{N}_0$$
   darstellen. Dann gilt
$$\mathrm{Pol}([a, b]) \subseteq\mid \ldots \subseteq\mid C^n([a, b]) \subseteq\mid \ldots \subseteq\mid C^0([a, b]) \subseteq\mid \mathcal{F}.$$

Korollar 1.4 *Sei V ein K-Vektorraum, I eine Indexmenge und $U_i \subseteq\mid V$ für alle $i \in I$. Dann gilt auch $\bigcap_{i \in I} U_i \subseteq\mid V$.*

Beweis. Es wird gezeigt, daß sich die Eigenschaften i)-iii) aus Satz 1.3 von den einzelnen Unterräumen U_i auf deren Schnittmenge übertragen.
Da für alle $i \in I$ die Mengen U_i Unterräume von V sind, ist $\mathcal{O} \in U_i$ für alle $i \in I$ und damit auch $\mathcal{O} \in \bigcap_{i \in I} U_i \neq \emptyset$. Die Eigenschaft i) ist damit erfüllt.
Seien nun $u, v \in \bigcap_{i \in I} U_i$, also $u, v \in U_i$ für alle $i \in I$. Dann gilt $k \cdot u \in U_i$ und $u + v \in U_i$ für alle $k \in K$ und alle $i \in I$, also auch $k \cdot u,\ u + v \in \bigcap_{i \in I} U_i$. Die Eigenschaften ii) und iii) sind also auch erfüllt.

□

Definition 1.5 *Ist V ein K-Vektorraum sowie $n \in \mathbb{N}$, $v_1, \ldots, v_n \in V$ und $k_1, \ldots, k_n \in K$, dann heißt $k_1 v_1 + \ldots + k_n v_n$ Linearkombination der Vektoren $v_1, \ldots, v_n$. Ist $M \subseteq V$ nicht leer, so heißt $v \in V$ Linearkombination von M, wenn v eine Linearkombination endlich vieler $v_1, \ldots, v_n \in M$ ist.*

Korollar 1.6 *Sei V ein K-Vektorraum, $M \subseteq V$ und $M \neq \emptyset$. Dann ist die Menge aller Linearkombinationen von M ein Unterraum von V, der M umfaßt.*

Beispiel.

1. Sei K ein Körper und $V = K^2$. Ist $M = \{(1,0),(0,1)\}$, so läßt sich jedes $v \in V$ als Linearkombination von M darstellen:

$$v = (v_1, v_2) = v_1 \cdot (1,0) + v_2 \cdot (0,1).$$

2. Sei $V = \mathrm{Pol}([a,b])$ wie in Beispiel 3 nach Satz 1.3 definiert. Für $i \in \mathbb{N}_0$ bezeichne

$$p_i \,:\, [a,b] \to \mathbb{R}, \quad x \mapsto x^i$$

das Monom i-ten Grades, z.B. $p_0 : [a,b] \to \mathbb{R},\ x \mapsto 1$ und $p_1 : [a,b] \to \mathbb{R},\ x \mapsto x$.
Ist $M := \{p_0, p_1, \ldots\} = \{p_i \mid i \in \mathbb{N}_0\}$, dann gilt:
Jedes Polynom $f \in \mathrm{Pol}([a,b])$ ist eine Linearkombination von M, denn

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

läßt sich darstellen in der Form

$$f = a_n p_n + a_{n-1}p_{n-1} + \ldots + a_1 p_1 + a_0 p_0.$$

Definition 1.7 *V sei ein K-Vektorraum und $M \subseteq V$. Dann bezeichnet $[M]$ den Durchschnitt aller Unterräume von V, die M enthalten.*

Bemerkung.

1. Mit Korollar 1.4 gilt $[M] \subseteq| V$.

2. Ist speziell $M = \{v_1, \ldots, v_n\}$, dann schreibt man auch $[v_1, \ldots, v_n]$ statt $[M]$.

3. $[M]$ ist der kleinste Unterraum von V, der M umfaßt, z.B. $[\emptyset] = \{\mathcal{O}\}$.

Satz 1.8 *V sei ein K-Vektorraum, $M \subseteq V$ und $M \neq \emptyset$. Dann ist $[M]$ die Menge aller Linearkombinationen von M.*

Beweis. $''\supseteq''$: $[M]$ enthält nach Definition 1.7 jede Linearkombination von M.
$''\subseteq''$: Die Menge aller Linearkombinationen von M ist ein Unterraum von V, der M umfaßt, und $[M]$ ist der kleinste Unterraum mit dieser Eigenschaft (Bemerkung 3 vor Satz 1.8).

$\square$

Definition 1.9 *Sei V ein K-Vektorraum und $M \subseteq V$.*
i) V wird von M erzeugt (aufgespannt) $:\Longleftrightarrow V = [M]$.
M heißt dann Erzeugendensystem von V.
ii) V heißt endlich erzeugt $:\Longleftrightarrow$ V hat ein endliches Erzeugendensystem.

Beispiel. Sei $V = K^n$ und seien n Vektoren aus V gegeben durch

$$e_1 := (1, 0, \ldots, 0),\ e_2 := (0, 1, 0, \ldots, 0), \ldots,\ e_n := (0, \ldots, 0, 1).$$

Dann wird V von $M := \{e_1, \ldots, e_n\}$ erzeugt, denn ein beliebiger Vektor $v \in V$ läßt sich darstellen in der Form

$$\begin{aligned} v = (a_1, \ldots, a_n) &= (a_1, 0, \ldots, 0) + (0, a_2, 0, \ldots, 0) + \ldots + (0, \ldots, 0, a_n) \\ &= a_1 e_1 + \ldots + a_n e_n\,, \end{aligned}$$

also als Linearkombination von M. Mit Satz 1.8 folgt damit $V \subseteq [M]$ und wegen $[M] \subseteq V$ gilt $V = [M]$. Da M endlich ist, ist V endlich erzeugt.

Definition 1.10 *Sei V ein K-Vektorraum und seien $U_1, \ldots, U_n$ Unterräume von V. Dann heißt $U_1 + \ldots + U_n := \{u_1 + \ldots + u_n \mid u_i \in U_i,\ i = 1, \ldots, n\}$ Summenraum von $U_1, \ldots, U_n$.*

Bemerkung.

1. $U_1 + \ldots + U_n \subseteq\!| V$.
2. $U_1 + \ldots + U_n = [U_1 \cup \ldots \cup U_n]$.

Beispiel. V sei ein K-Vektorraum und $\{v_1, \ldots, v_n\}$ ein Erzeugendensystem von V. Dann gilt

$$\begin{aligned} V &= \{k_1 v_1 + \ldots + k_n v_n \mid k_i \in K,\ i = 1, \ldots, n\} \\ &= \{u_1 + \ldots + u_n \mid u_i \in K \cdot v_i,\ i = 1, \ldots, n\} \\ &= K \cdot v_1 + \ldots + K \cdot v_n. \end{aligned}$$

Ist konkret $V = K^n$ und sind $e_1, \ldots, e_n$ wie im Beispiel zu Definition 1.9, dann ist $K^n = K \cdot e_1 + \ldots + K \cdot e_n$. Jedes $v \in K^n$ läßt sich sogar eindeutig als Linearkombination von $e_1, \ldots, e_n$ schreiben. Insbesondere gilt

$$\mathcal{O} = k_1 e_1 + \ldots + k_n e_n = (k_1, \ldots, k_n) \implies k_1 = \ldots = k_n = 0\,.$$

Definition 1.11 *Sei V ein K-Vektorraum und $v_1, \ldots, v_n \in V$. Die Vektoren $v_1, \ldots, v_n$ heißen linear unabhängig, wenn gilt:*

$$\forall k_1, \ldots, k_n \in K : (\mathcal{O} = k_1 v_1 + \ldots + k_n v_n \iff k_1 = \ldots = k_n = 0).$$

Anderenfalls heißen $v_1, \ldots, v_n$ linear abhängig.

Beispiel. Sei $V = \mathbb{R}^3$.

1. Wir betrachten die Vektoren $v_1 = (2, 1, 0)$, $v_2 = (0, 1, 3)$, $v_3 = (0, 0, 2)$. Um die Bedingung für lineare Unabhängigkeit zu überprüfen, schreiben wir den Nullvektor als Linearkombination dieser drei Vektoren.

   $$\mathcal{O} = (0, 0, 0) = k_1 \cdot (2, 1, 0) + k_2 \cdot (0, 1, 3) + k_3 \cdot (0, 0, 2) = (2k_1, k_1 + k_2, 3k_2 + 2k_3)\,.$$

   Es folgt

   $$2k_1 = k_1 + k_2 = 3k_2 + 2k_3 = 0\,,$$

   also $k_1 = k_2 = k_3 = 0$. Damit ist gezeigt, daß v_1, v_2 und v_3 linear unabhängig sind.

2. Die Vektoren $v_1 = (1,1,2)$, $v_2 = (2,3,-1)$, $v_3 = (-1,-2,3)$ sind linear abhängig, denn
$$(-1)\cdot v_1 + v_2 + v_3 = (-1+2-1, -1+3-2, -2-1+3) = \mathcal{O},$$
das heißt, der Nullvektor läßt sich nichttrivial als Linearkombination von v_1, v_2 und v_3 darstellen.

Definition 1.12 *Sei V ein K-Vektorraum und $M \subseteq V$. Dann heißt M linear unabhängig, wenn je endlich viele verschiedene Vektoren aus M linear unabhängig sind. Anderenfalls heißt M linear abhängig.*

Beispiel.

1. $\emptyset$ ist linear unabhängig.

2. Sei $K = \mathbb{Q}$ und $V = \mathbb{R}$; wir betrachten also den Vektorraum $\mathbb{R}$ über dem Körper $\mathbb{Q}$.
Behauptung: $M = \{1, \pi, \pi^2, \ldots\} = \{\pi^i \mid i \in \mathbb{N}_0\}$ ist linear unabhängig.
Seien $n \in \mathbb{N}$, $k_1, \ldots, k_n \in \mathbb{Q}, i_1, \ldots, i_n \in \mathbb{N}_0$, so daß
$$k_n \pi^{i_n} + \ldots + k_1 \pi^{i_1} = 0, \quad i_1 < \ldots < i_n.$$
Da π transzendent ist, folgt $k_n = \ldots = k_1 = 0$. M ist also linear unabhängig.

Bemerkung.

1. Sind $v_1, \ldots, v_n \in V$, so heißt $\mathcal{O} = k_1 \cdot v_1 + \ldots + k_n \cdot v_n$ triviale Darstellung von $\mathcal{O}$ durch $v_1, \ldots, v_n$ genau dann, wenn $k_1 = \ldots = k_n = 0$. Mit dieser Bezeichnung gilt also:
$v_1, \ldots, v_n$ sind genau dann linear unabhängig, wenn sich $\mathcal{O}$ durch $v_1, \ldots, v_n$ nur trivial darstellen läßt.

2. Sind die Vektoren $v_1, \ldots, v_n \in V$ linear unabhängig, so gilt $v_i \neq \mathcal{O}$ für $i = 1, \ldots, n$, denn wäre z.B. $v_i = \mathcal{O}$, dann wäre durch
$$0 \cdot v_1 + \ldots + 0 \cdot v_{i-1} + 1 \cdot v_i + 0 \cdot v_{i+1} + \ldots + 0 \cdot v_n = \mathcal{O}$$
eine nichttriviale Darstellung von $\mathcal{O}$ gegeben.

3. Sind $v_1, \ldots, v_n$ linear unabhängig, so sind sie auch verschieden.

Satz 1.13 *Sei V ein K-Vektorraum, $M \subseteq V$ und $M \neq \emptyset$. Dann sind die folgenden Aussagen äquivalent.*

i) M ist linear unabhängig.

ii) Jedes $v \in [M]$ läßt sich eindeutig als Linearkombination von M schreiben.

iii) Für alle $v \in M$ gilt $v \notin [M \setminus \{v\}]$.

Beweis. i)⇒ii): Wir beweisen die Behauptung durch Kontraposition. Sei $v \in [M]$ und die Darstellung von v als Linearkombination von M nicht eindeutig. Dann gibt es $k_1, \ldots, k_n \in K$ und $l_1, \ldots, l_n \in K$, mit $k_i \neq l_i$ für mindestens ein $i \in \{1, \ldots, n\}$, so daß

$$v = k_1 \cdot v_1 + \ldots + k_n \cdot v_n = l_1 \cdot v_1 + \ldots + l_n \cdot v_n \,.$$

Dabei sind $v_1, \ldots, v_n \in M$ verschieden. Es ergibt sich

$$\mathcal{O} = (k_1 - l_1) \cdot v_1 + \ldots + (k_n - l_n) \cdot v_n.$$

Wegen $k_i - l_i \neq 0$ für mindestens ein i, ist dieses eine nichttriviale Darstellung von $\mathcal{O}$, d.h., $v_1, \ldots, v_n$ sind linear abhängig.
ii)⇒iii): Wir beweisen die Behauptung durch Kontraposition. Sei $v \in M$ und $v \in [M \setminus \{v\}]$.
1. Fall: $M \setminus \{v\} = \emptyset$. Dann ist $v = \mathcal{O}$, da $[\emptyset] = \{\mathcal{O}\}$. Für v lassen sich nun verschiedene Darstellungen als Linearkombination von M angeben: $v = 1 \cdot v = 0 \cdot v$.
2. Fall: $M \setminus \{v\} \neq \emptyset$. Wegen Satz 1.8 gibt es $v_1, \ldots, v_n \in M \setminus \{v\}$ und $k_1, \ldots, k_n \in K$ mit

$$v = k_1 \cdot v_1 + \ldots + k_n \cdot v_n.$$

Mit $v = 1 \cdot v$ erhält man also zwei Darstellungen von v als Linearkombination von M.
iii)⇒i): Auch diesen Beweis führen wir durch Kontraposition. Ist M linear abhängig, dann gibt es verschiedene $v_1, \ldots, v_n \in M$ und $k_1, \ldots, k_n \in K$, nicht alle Null, mit

$$\mathcal{O} = k_1 \cdot v_1 + \ldots + k_n \cdot v_n \,.$$

Sei o.B.d.A. $k_1 \neq 0$. Dann folgt

$$\begin{aligned} -k_1 v_1 &= k_2 v_2 + \ldots + k_n v_n \,, \quad \text{also} \\ v_1 &= -\frac{k_2}{k_1} \cdot v_2 - \ldots - \frac{k_n}{k_1} \cdot v_n \in [M \setminus \{v_1\}] \,. \end{aligned}$$

□

Definition 1.14 *Ist V ein K-Vektorraum und $B \subseteq V$ eine Teilmenge von V, so heißt B Basis von V, wenn V durch B erzeugt wird und B linear unabhängig ist.*

Beispiel.

1. Sei $V = K^n$ und $B = \{e_1, \ldots, e_n\}$ mit $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, wobei die 1 an der i-ten Position steht. Dann ist B eine Basis von V und heißt kanonische Basis von K^n.

2. $\emptyset$ ist eine Basis von $\{\mathcal{O}\}$.

Satz 1.15 *Ist V ein K-Vektorraum und $B \subseteq V$, so sind die folgenden Aussagen äquivalent.*

i) B ist eine Basis von V.

ii) B ist eine minimale Teilmenge von V, die V erzeugt.

iii) B ist eine maximale linear unabhängige Teilmenge von V.

Beweis. Ist $V = \{\mathcal{O}\}$, dann sind i), ii), iii) jeweils äquivalent zu $B = \emptyset$. Sei also $V \neq \{\mathcal{O}\}$.
i)⇒ii): Ist B eine Basis von V, dann wird V nach Definition von B erzeugt. Es bleibt zu zeigen, daß B mit dieser Eigenschaft minimal ist. Wir nehmen an, es gäbe $M \subset B$ mit $[M] = V$. Dann existiert $v \in B \setminus M$, und es gilt $v \in V = [M] \subseteq [B \setminus \{v\}]$. Mit Satz 1.13 folgt, daß B nicht linear unabhängig ist, und damit ein Widerspruch zu i).
ii)⇒iii): Da $V \neq \{\mathcal{O}\}$, folgt $B \neq \emptyset$. Wir zeigen zunächst die lineare Unabhängigkeit von B mittels eines Widerspruchsbeweises.
Annahme: B ist nicht linear unabhängig. Dann existiert $v \in B$, so daß $v \in [B \setminus \{v\}]$, also $[B \setminus \{v\}] = [B] = V$ (vgl. Aufgabe 3.8). Somit ist B als Erzeugendensystem nicht minimal, was ein Widerspruch zu ii) ist.
Es bleibt zu zeigen, daß B eine maximale linear unabhängige Menge ist. Sei $v \in V \setminus B$. Dann folgt $v \in [B]$ wegen ii) und mit Satz 1.13 die lineare Abhängigkeit von $B \cup \{v\}$.
iii)⇒i): Da B nach Voraussetzung linear unabhängig ist, bleibt nach Definition 1.14 nur $[B] = V$ zu zeigen. Sei $v \in V$. Gilt $v \in B$, dann folgt $v \in [B]$ trivial. Sei also $v \notin B$. Dann ist wegen iii) die Menge $B \cup \{v\}$ linear abhängig, also $v \in [B]$ (vgl. Aufgabe 3.8).

□

Satz 1.16 *Ist V ein K-Vektorraum und $V = [v_1, \ldots, v_n]$ mit $v_1, \ldots, v_n \in V$, dann existiert eine Basis B von V mit $B \subseteq \{v_1, \ldots, v_n\}$.*

Beweis. Die Potenzmenge $\wp(\{v_1, \ldots, v_n\})$ ist endlich. Also gibt es $B \in \wp(\{v_1, \ldots, v_n\})$ minimal mit der Eigenschaft $[B] = V$. Wegen Satz 1.15 folgt, daß B eine Basis ist.

□

Bemerkung. Mit Hilfe transfiniter Methoden der Mengenlehre kann gezeigt werden, daß jede linear unabhängige Menge eines Vektorraums zu einer Basis erweitert werden kann. Jeder Vektorraum hat damit eine Basis.

Satz 1.17 *Sei V ein K-Vektorraum und $\{v_1, \ldots, v_n\}$ eine Basis von V. Dann ist für $u \in V$ mit*

$$u = k_1 \cdot v_1 + \ldots + k_n \cdot v_n\,, \quad k_1, \ldots, k_n \in K \tag{1}$$

und $k_i \neq 0$ auch

$$\{v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_n\}$$

eine Basis von V.

Beweis. Sei in der Darstellung (1) o.B.d.A. $k_1 \neq 0$. Dann gilt

$$\begin{aligned} k_1 \cdot v_1 &= u - k_2 \cdot v_2 - \ldots - k_n \cdot v_n\,, \quad \text{also} \\ v_1 &= \frac{1}{k_1} \cdot u - \frac{k_2}{k_1} \cdot v_2 - \ldots - \frac{k_n}{k_1} \cdot v_n. \end{aligned}$$

Damit gilt $v_1 \in [u, v_2, \ldots, v_n]$, und wegen Aufgabe 3.8 folgt

$$[u, v_2, \ldots, v_n] = [u, v_1, \ldots, v_n] = V.$$

Es bleibt die lineare Unabhängigkeit von $\{u, v_2, \ldots, v_n\}$ zu zeigen. Nach Voraussetzung sind $v_2, \ldots, v_n$ linear unabhängig.
Widerspruchsannahme: $u, v_2, \ldots, v_n$ sind linear abhängig. Dann gilt $u \in [v_2, \ldots, v_n]$ und damit $[v_2, \ldots, v_n] = [u, v_2, \ldots, v_n] = V$ (vgl. Aufgabe 3.8). Das ist ein Widerspruch, da $\{v_1, \ldots, v_n\}$ nach Voraussetzung eine Basis und mit Satz 1.15 also ein minimales Erzeugendensystem von V ist.

□

Satz 1.18 (Austauschsatz von Steinitz) *Sei V ein K-Vektorraum und $\{v_1, \ldots, v_n\}$ eine Basis von V. Sind $u_1, \ldots, u_m \in V$ linear unabhängig, so gilt $m \leq n$, und bei geeigneter Indizierung von $v_1, \ldots, v_n$ ist $\{u_1, \ldots, u_m, v_{m+1}, \ldots, v_n\}$ auch eine Basis von V.*

Beweis. Wir beweisen den Austauschsatz durch vollständige Induktion nach m.
Induktionsanfang: $m = 0$. Nach Voraussetzung ist $\{v_1, \ldots, v_n\}$ eine Basis von V.
Induktionsschluß: Seien $u_1, \ldots, u_m, u_{m+1}$ linear unabhängig. Dann gilt nach Induktionsvoraussetzung $m \leq n$, und $\{u_1, \ldots, u_m, v_{m+1}, \ldots, v_n\}$ ist bei geeigneter Indizierung eine Basis von V. Wäre $m = n$, dann wäre $\{u_1, \ldots, u_m\}$ eine Basis von V und damit eine maximale linear unabhängige Menge, was im Widerspruch zu der Annahme steht, daß $u_1, \ldots, u_{m+1}$ linear unabhängig sind. Es muß also $m < n$ gelten, das heißt $m + 1 \leq n$. Nach Induktionsvoraussetzung können wir schreiben

$$u_{m+1} = k_1 \cdot u_1 + \ldots + k_m \cdot u_m + k_{m+1} \cdot v_{m+1} + \ldots + k_n \cdot v_n \, .$$

Nehmen wir nun an, es gelte $k_{m+1} = \ldots = k_n = 0$. Dann wäre durch

$$\mathcal{O} = k_1 \cdot u_1 + \ldots + k_m \cdot u_m - u_{m+1}$$

eine nichttriviale Darstellung von $\mathcal{O}$ gegeben, was im Widerspruch zur geforderten linearen Unabhängigkeit dieser Vektoren steht. Wir können also o.B.d.A. $k_{m+1} \neq 0$ annehmen. Dann folgt mit Satz 1.17, daß $\{u_1, \ldots, u_m, u_{m+1}, v_{m+2}, \ldots, v_n\}$ eine Basis von V ist.

□

Korollar 1.19 *Sei V ein K-Vektorraum. Dann gilt genau eine der beiden folgenden Aussagen.*

i) Jede Basis von V hat unendlich viele Elemente.

ii) Es gibt ein $n \in \mathbb{N}_0$, so daß jede Basis von V genau n Elemente hat.

Beweis. Wir nehmen an, i) gelte nicht. Dann besitzt V eine Basis B mit $n := |B| < \infty$. Für jede weitere Basis B' von V folgt mit Satz 1.18 sowohl $|B'| \leq |B|$ als auch $|B| \leq |B'|$.

□

Definition 1.20 *V sei ein K-Vektorraum. Hat V eine endliche Basis B, so heißt $n := |B| < \infty$ Dimension von V, geschrieben $\dim_K V = n$. Anderenfalls hat V die Dimension ∞, geschrieben $\dim_K V = \infty$.*

Bemerkung.

1. Wegen Korollar 1.19 ist der Begriff *Dimension* wohldefiniert.
2. V ist endlich erzeugt genau dann, wenn $\dim_K V < \infty$.

Beispiel.

1. $\dim_K\{\mathcal{O}\} = 0; \quad \dim_K K^n = n.$
2. $\dim_{\mathbb{R}} \mathbb{C} = 2$, denn $\{1, i\}$ ist eine Basis von $\mathbb{C}$ über $\mathbb{R}$.
3. $\dim_{\mathbb{Q}} \mathbb{R} = \infty$.

Korollar 1.21 *Sei V ein K-Vektorraum mit* $\dim_K V = n < \infty$.

i) Ist $M \subseteq V$ linear unabhängig, so ist M genau dann eine Basis von V, wenn $|M| = n$.

ii) Ist U ein Unterraum von V, dann gilt $\dim_K U \leq \dim_K V$ *und $U = V$ genau dann, wenn* $\dim_K U = \dim_K V$.

Beweis.

i) $'' \Rightarrow''$: Gilt wegen Korollar 1.19 und der Definition von $\dim_K V$.
$'' \Leftarrow''$: Ergänzt man M zu einer Basis B, so ist $M \subseteq B$. Da aber $n = |M| \leq |B| = n$, folgt $M = B$.

ii) Sei B eine Basis von U. Dann folgt mit Satz 1.18 aus der linearen Unabhängigkeit von B, daß $|B| \leq n$. Also gilt $\dim_K U \leq \dim_K V$. Weiterhin ist $\dim_K U = \dim_K V$ genau dann, wenn $|B| = n$. Das ist nach i) äquivalent dazu, daß B eine Basis auch von V ist, also $U = V$.

□

Satz 1.22 (Dimensionssatz) *Sei W ein K-Vektorraum und seien U, V Unterräume von W mit* $\dim_K U$, $\dim_K V < \infty$. *Dann gilt*

$$\dim_K U + \dim_K V = \dim_K(U + V) + \dim_K(U \cap V).$$

Beweis. Gilt $U = \{\mathcal{O}\}$ oder $V = \{\mathcal{O}\}$, so folgt die Behauptung offensichtlich. Sei also $U, V \neq \{\mathcal{O}\}$. Wir bezeichnen mit $\{w_1, \ldots, w_r\}$ eine Basis von $U \cap V$, die wegen Satz 1.18 zu einer Basis $\{w_1, \ldots, w_r, v_1, \ldots, v_s\}$ von V und zu einer Basis $\{w_1, \ldots, w_r, u_1, \ldots, u_t\}$ von U ergänzt werden kann, und wollen zeigen, daß dann $B = \{w_1, \ldots, w_r, v_1, \ldots, v_s, u_1, \ldots, u_t\}$ eine Basis von $U + V$ ist. Sei $u + v \in U + V$, wobei u als Linearkombination der Vektoren $w_1, \ldots, w_r, u_1, \ldots, u_t$ und v entsprechend als Linearkombination von $w_1, \ldots, w_r, v_1, \ldots, v_s$ dargestellt werden kann. Also ist $u + v \in [B]$. Es bleibt zu zeigen, daß B linear unabhängig ist. Seien hierfür $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s, \gamma_1, \ldots, \gamma_t \in K$, so daß

$$\alpha_1 \cdot w_1 + \ldots + \alpha_r \cdot w_r + \beta_1 \cdot v_1 + \ldots + \beta_s \cdot v_s + \gamma_1 \cdot u_1 + \ldots + \gamma_t \cdot u_t = \mathcal{O}\,, \quad \text{also}$$
$$\alpha_1 \cdot w_1 + \ldots + \alpha_r \cdot w_r + \beta_1 \cdot v_1 + \ldots + \beta_s \cdot v_s = -\gamma_1 \cdot u_1 - \ldots - \gamma_t \cdot u_t \in U \cap V\,. \qquad (2)$$

Es gibt $\sigma_1, \ldots, \sigma_r \in K$ mit

$$\begin{aligned} -\gamma_1 \cdot u_1 - \ldots - \gamma_t \cdot u_t &= \sigma_1 \cdot w_1 + \ldots + \sigma_r \cdot w_r \,, \quad \text{also} \\ \mathcal{O} &= \gamma_1 \cdot u_1 + \ldots + \gamma_t \cdot u_t + \sigma_1 \cdot w_1 + \ldots + \sigma_r \cdot w_r \,. \end{aligned}$$

Da die Vektoren $u_1, \ldots, w_r$ eine Basis von U bilden, sind sie linear unabhängig, und damit folgt

$$\gamma_1 = \ldots = \gamma_t = \sigma_1 = \ldots = \sigma_r = 0 \,.$$

Da auch die Vektoren $w_1, \ldots, v_s$ linear unabhängig sind, folgt mit (2)

$$\alpha_1 = \ldots = \alpha_r = \beta_1 = \ldots = \beta_s = 0 \,.$$

Insgesamt ergibt sich also

$$\dim_K U + \dim_K V = (r+t) + (r+s) = (r+s+t) + r = \dim_K(U+V) + \dim_K(U \cap V) \,.$$

$\square$

2. Lineare Gleichungssysteme

Ist K ein Körper, dann heißt

$$\begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ & & \vdots & & & & \vdots \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array} \tag{3}$$

mit $a_{ij}, b_i \in K$ lineares Gleichungssystem (LGS) mit m Gleichungen und n Unbekannten. Gesucht ist die Lösungsmenge

$$\mathbb{L} := \{(x_1, \ldots, x_n) \in K^n \mid x_1, \ldots, x_n \;\; \text{lösen} \quad (3)\}.$$

(3) heißt homogenes LGS, falls $b_1 = \ldots = b_m = 0$.

Beispiel.

1. In dem Gleichungssystem

$$\begin{array}{ccc} x_1 & = & 0 \\ x_1 & = & 1 \end{array}$$

   ist die Zahl der Gleichungen $m = 2$ und die Zahl der Unbekannten $n = 1$. Die Lösungsmenge ist gegeben durch $\mathbb{L} = \emptyset$; das LGS ist also nicht lösbar.

2. Betrachten wir

$$x_1 = 1$$

   als LGS mit $n = 1$ und $m = 1$, so ist die Lösungsmenge $\mathbb{L} = \{1\}$; das System ist eindeutig lösbar.

3. In dem LGS
$$x_1 + x_2 = 0$$
ist $n = 2$, $m = 1$ und die Lösungsmenge $\mathbb{L} = \{(k, -k) \mid k \in K\}$. Dieses lineare Gleichungssystem ist also lösbar, aber nicht eindeutig.

Formale Schreibweise. Ein lineares Gleichungssystem kann formal durch eine Matrixgleichung beschrieben werden. Die Koeffizienten des LGS bilden dabei eine (m, n)-Matrix

$$A := \begin{array}{c} \downarrow \ \text{Spalten} \ \downarrow \\ \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \end{array} \begin{array}{l} \\ \leftarrow \\ \vdots \ \text{Zeilen} \\ \leftarrow \end{array}$$

die man Koeffizientenmatrix des LGS nennt. Die Menge aller (m, n)-Matrizen über K bezeichnet man mit $K_{m,n}$. Es gilt also $A \in K_{m,n}$. Die Matrix

$$A_{\text{erw}} := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

heißt erweiterte Koeffizientenmatrix des LGS. Im folgenden sei

$$A \neq \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix},$$

die Koeffizientenmatrix ist also nicht die Nullmatrix.

Definiert man

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + & \dots & + a_{1n}x_n \\ & \vdots & \\ a_{m1}x_1 + & \dots & + a_{mn}x_n \end{pmatrix}$$

und führt man die formale Bezeichnung

$$(x_1, \dots, x_n)^t := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

ein, so kann man das allgemeine lineare Gleichungssystem (3) in der Form

$$A \cdot x^t = b^t \tag{4}$$

schreiben. Man bezeichnet (4) oft auch als inhomogenes LGS und

$$A \cdot x^t = \mathcal{O}^t \tag{5}$$

als zugehöriges homogenes LGS. Die Lösungsmenge des inhomogenen LGS ist dann

$$\mathbb{L} = \{x \in K^n \mid A \cdot x^t = b^t\},$$

die des homogenen ist gegeben durch

$$\mathbb{L}_h = \{x \in K^n \mid A \cdot x^t = \mathcal{O}^t\}.$$

Beispiel. Für das LGS

$$\begin{array}{rcl} 2x_1 + 3x_2 + x_3 & = & 5 \\ 2x_2 + x_1 & = & 7 \end{array}$$

mit $n = 3$ und $m = 2$ ist die Koeffizientenmatrix gegeben durch

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

und die Inhomogenität durch $b = (5, 7)$. Die erweiterte Koeffizientenmatrix ist damit

$$A_{\text{erw}} = \left(\begin{array}{ccc|c} 2 & 3 & 1 & 5 \\ 1 & 2 & 0 & 7 \end{array}\right).$$

Das LGS hat also in der Matrixschreibweise die Form

$$\begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \end{pmatrix},$$

und das zugehörige homogene LGS ist gegeben durch

$$\begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Satz 2.1 *Ist $A \in K_{m,n}$ eine (m,n)-Matrix und $A \cdot x^t = \mathcal{O}^t$ ein homogenes LGS, dann ist $\mathbb{L} = \{x \in K^n \mid A \cdot x^t = \mathcal{O}^t\}$ ein Unterraum von K^n.*

Beweis. Wir beweisen Satz 2.1, indem wir für $\mathbb{L}$ die Eigenschaften i)-iii) aus Satz 1.3 nachprüfen. Für Lösungen des linearen Gleichungssystems muß für alle $i = 1, \ldots, m$ gelten

$$\sum_{j=1}^{n} a_{ij} x_j = 0.$$

i) $\mathbb{L}$ ist nicht leer, da der Nullvektor die Gleichungen löst. Es ist also $(0,\ldots,0) \in \mathbb{L}$.

ii) Sind $x = (x_1,\ldots,x_n)$, $y = (y_1,\ldots,y_n) \in \mathbb{L}$, dann folgt für alle $i = 1,\ldots,m$

$$\sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} a_{ij}y_j = 0\,, \quad \text{also} \quad \sum_{j=1}^{n} a_{ij}(x_j + y_j) = 0,$$

d.h. $x + y \in \mathbb{L}$. Damit ist $\mathbb{L}$ abgeschlossen bezüglich der Addition.

iii) Ist $x = (x_1,\ldots,x_n) \in \mathbb{L}$ und $k \in K$, dann folgt für alle $i = 1,\ldots,m$

$$\sum_{j=1}^{n} a_{ij}(k \cdot x_j) = k \cdot \sum_{j=1}^{n} a_{ij}x_j = k \cdot 0 = 0,$$

d.h. $k{\cdot}x \in \mathbb{L}$. Damit ist $\mathbb{L}$ auch bezüglich der Multiplikation mit Skalaren abgeschlossen.

□

Bemerkung. Zur Beschreibung von $\mathbb{L}$ braucht man somit nur eine Basis von $\mathbb{L}$.

Satz 2.2 *Ist $A \in K_{m,n}$ und $A \cdot x^t = b^t$ ein LGS, dann gilt für jedes $y \in \mathbb{L}$*

$$\mathbb{L} = y + \mathbb{L}_h = \{y + x \mid x \in \mathbb{L}_h\}.$$

Beweis. Seien $x, y, z \in K^n$ mit $x = (x_1,\ldots,x_n)$, $y = (y_1,\ldots,y_n)$ und $z = (z_1,\ldots,z_n)$.
$''\supseteq''$: Gilt $y \in \mathbb{L}$ und $x \in \mathbb{L}_h$, dann folgt für alle $i = 1,\ldots,m$

$$\sum_{j=1}^{n} a_{ij}(y_j + x_j) = \sum_{j=1}^{n} a_{ij}y_j + \sum_{j=1}^{n} a_{ij}x_j = b_i + 0 = b_i$$

und damit $y + x \in \mathbb{L}$.
$''\subseteq''$: Gilt $z \in \mathbb{L}$ und $y \in \mathbb{L}$, so folgt für alle $i = 1,\ldots,m$

$$\sum_{j=1}^{n} a_{ij}(z_j - y_j) = \sum_{j=1}^{n} a_{ij}z_j - \sum_{j=1}^{n} a_{ij}y_j = b_i - b_i = 0$$

und damit $z - y \in \mathbb{L}_h$. Es ergibt sich also für ein beliebiges Element z von $\mathbb{L}$ die Darstellung

$$z = y + (z - y) \in y + \mathbb{L}_h.$$

□

Bemerkung. Man sagt, daß die Gesamtlösung eines LGS eine spezielle Lösung des LGS plus der Gesamtlösung des zugehörigen homogenen Systems ist.

Die Eliminationsmethode zur Lösung eines LGS.
Sei die erweiterte Koeffizientenmatrix

$$A_{\text{erw}} = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right)$$

gegeben. Dann können die folgenden elementaren Umformungen zur Berechnung der Lösungsmenge durchgeführt werden.

Elementare Umformungen:

1. Vertauschen von Zeilen.
2. Vertauschen der ersten n Spalten. (Diese Umformung führt zu einer Umnumerierung der Unbekannten!)
3. Addieren/Subtrahieren einer Zeile zu/von einer anderen.
4. Multiplizieren einer Zeile mit einem $k \in K\backslash\{0\}$.

Jede Lösung des ursprünglichen Systems ist auch Lösung des neuen Systems, das durch eine elementare Umformung aus dem alten entstanden ist. Da jede elementare Umformung durch eine Umformung vom selben Typ rückgängig gemacht werden kann, verändert sich die Lösungsmenge durch elementare Umformungen nicht. Anstelle des ursprünglichen LGS genügt es, das umgeformte LGS zu lösen.
Lösungsprinzip. Man überlegt sich leicht, daß man nach einer endlichen Anzahl von elementaren Umformungen ein LGS mit der folgenden erweiterten Matrix erhalten kann:

$$A_{\text{erw}} = \left(\begin{array}{ccccccc|c}
1 & 0 & \dots & 0 & c_{1(r+1)} & \dots & c_{1n} & b'_1 \\
0 & 1 & & 0 & c_{2(r+1)} & \dots & c_{2n} & \vdots \\
\vdots & & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \dots & 0 & 1 & c_{r(r+1)} & \dots & c_{rn} & b'_r \\
0 & \dots & \dots & 0 & 0 & \dots & 0 & b'_{r+1} \\
0 & & \ddots & 0 & 0 & \ddots & 0 & \vdots \\
0 & \dots & \dots & 0 & 0 & \dots & 0 & b'_m
\end{array}\right). \tag{6}$$

Eine Koeffizientenmatrix, die diese spezielle Form hat, bezeichnet man als Koeffizientenmatrix in Normalform.

Beispiel. Die erweiterte Koeffizientenmatrix sei gegeben durch

$$A_{\text{erw}} = \left(\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 0 & 1 \end{array}\right)$$

Vertauschen der ersten beiden Zeilen $((Z1) \leftrightarrow (Z2))$:

$$\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \end{array}$$

Subtrahieren der ersten Zeile von der dritten $((Z3) - (Z1))$:

$$\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{array}$$

Subtrahieren der zweiten Zeile von der ersten $((Z1) - (Z2))$:

$$\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{array}$$

Vertauschen der dritten und vierten Spalte $((S3) \leftrightarrow (S4))$:

$$\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 \end{array}$$

Addieren der dritten Zeile zu der zweiten $((Z2) + (Z3))$:

$$\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 \end{array}$$

Multiplizieren der dritten Zeile mit -1 $((-1) \cdot (Z3))$:

$$\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{array}$$

Die erweiterte Koeffizientenmatrix ist nun in Normalform. Die Lösungsmenge setzt sich zusammen aus einer speziellen Lösung des inhomogenen Gleichungssystems und der allgemeinen Lösung des homogenen LGS. Das homogene LGS ist demnach durch

$$x_1 + x_4 = 0 \tag{7}$$
$$x_2 = 0 \tag{8}$$
$$x_3 = 0 \tag{9}$$

gegeben und hat die allgemeine Lösungsmenge $\mathbb{L}_h = \{k(1,\, 0,\, 0,\, -1) \mid k \in K\}$. Der Vektor $(1,\, 0,\, 1,\, 0)$ ist eine spezielle Lösung des inhomogenen Gleichungssystems. Um die richtige Lösungsmenge zu erhalten, muß nun noch die dritte Umformung $((S3) \leftrightarrow (S4))$ rückgängig gemacht werden. Damit ergibt sich als Lösungsmenge

$$\mathbb{L} = \{(1,\, 0,\, 0,\, 1) + k(1,\, 0,\, -1,\, 0) \mid k \in K\}\,.$$

Wir kommen zurück auf das allgemeine lineare Gleichungssystem mit der Normalform (6). Offenbar ist die Lösungsmenge genau dann nicht leer, wenn $b'_{r+1} = \ldots = b'_m = 0$. Um Lösungsvektoren für dieses allgemeine System in Normalform zu bekommen, betrachten wir die Vektoren $v_1, \ldots, v_{n-r} \in K^n$, die gegeben sind durch

$$v_i := (c_{1(r+i)}, \ldots,\, c_{r(r+i)},\, 0, \ldots, 0, -1, 0, \ldots,\, 0), \quad 1 \leq i \leq n-r, \tag{10}$$

wobei in v_i die -1 an der $(r+i)-$ten Stelle steht.

Lemma 2.3 *Ist die erweiterte Koeffizientenmatrix eines LGS in der Normalform (6) gegeben und sind* $v_1, \ldots, v_{n-r} \in K^n$ *wie in (10), dann bilden* $v_1, \ldots, v_{n-r}$ *eine Basis des Lösungsraumes* $\mathbb{L}_h$ *des zugehörigen homogenen LGS.*

Beweis. Die Vektoren $v_1, \ldots, v_{n-r}$ sind nach Konstruktion linear unabhängig. Zu zeigen ist also $\mathbb{L}_h = [v_1, \ldots, v_{n-r}]$. Zunächst überprüft man leicht durch Einsetzen, daß $v_1, \ldots, v_{n-r}$ tatsächlich Lösungen des homogenen LGS sind, d.h. $[v_1, \ldots, v_{n-r}] \subseteq \mathbb{L}_h$. Um die umgekehrte Inklusion zu beweisen sei $x \in \mathbb{L}_h$. Dann gibt es $\lambda_1, \ldots, \lambda_{n-r} \in K$ mit

$$x + \lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r} = (k_1, \ldots, k_r, 0, \ldots, 0) \in \mathbb{L}_h$$

(setze $\lambda_i = x_{r+i}$). Multipliziert man die Koeffizientenmatrix mit $(k_1, \ldots, k_r, 0, \ldots, 0) \in \mathbb{L}_h$, so folgt $k_1 = \ldots = k_r = 0$, und wir erhalten

$$x + \lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r} = \mathcal{O},$$

d.h., x ist eine Linearkombination von $v_1, \ldots, v_{n-r}$.

□

Für $\mathbb{L} \neq \emptyset$, also $b'_{r+1} = \ldots = b'_m = 0$, ist $(b'_1, \ldots, b'_r, 0, \ldots, 0)$ eine spezielle Lösung des inhomogenen LGS, und damit folgt

$$\boxed{\mathbb{L} = \{(b'_1, \ldots, b'_r, 0, \ldots, 0) + \lambda_1 v_1 + \ldots + \lambda_{n-r} v_{n-r} \mid \lambda_1, \ldots, \lambda_{n-r} \in K\}. \qquad (11)}$$

Beispiel. Die erweiterte Koeffizientenmatrix sei gegeben durch

$$\left(\begin{array}{cccc|c} 1 & 1 & 3 & 1 & 0 \\ 2 & -1 & 0 & 5 & -3 \\ 1 & -2 & -3 & 4 & -3 \\ 0 & 1 & 2 & -1 & 1 \end{array}\right),$$

und sie soll mit Hilfe elementarer Umformungen auf Normalform gebracht werden.
$(Z2) - 2 \cdot (Z1)$ und $(Z3) - (Z1)$:

$$\begin{array}{cccc|c} 1 & 1 & 3 & 1 & 0 \\ 0 & -3 & -6 & 3 & -3 \\ 0 & -3 & -6 & 3 & -3 \\ 0 & 1 & 2 & -1 & 1 \end{array}$$

$(Z3) - (Z2)$ und $(Z4) + \frac{1}{3} \cdot (Z2)$, dann $(-\frac{1}{3}) \cdot (Z2)$:

$$\begin{array}{cccc|c} 1 & 1 & 3 & 1 & 0 \\ 0 & 1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$

$(Z1) - (Z2)$:

$$\begin{array}{cccc|c} 1 & 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}$$

Dieses LGS ist also lösbar, und nach Formel (11) gilt für die Lösungsmenge

$$\mathbb{L} = \{(-1,1,0,0) + \lambda_1 \cdot (1,2,-1,0) + \lambda_2 \cdot (2,-1,0,-1) \mid \lambda_1, \lambda_2 \in K\}\,.$$

Definition 2.4 *Ist $A \in K_{m,n}$, so heißen die Vektoren des K^n, die durch die einzelnen Zeilen von A gegeben sind, Zeilenvektoren von A. Der Zeilenraum von A ist der Unterraum des K^n, der von den Zeilenvektoren von A aufgespannt wird. Die Dimension des Zeilenraumes von A ist die Maximalzahl linear unabhängiger Zeilenvektoren von A und heißt (Zeilen-)Rang der Matrix A, geschrieben* $\mathrm{Rg}A$.

Bemerkung.

1. Entsprechend definiert man den Spaltenrang als Maximalzahl linear unabhängiger Spaltenvektoren. Wir zeigen später, daß Zeilenrang und Spaltenrang gleich sind.

2. Entsteht die Matrix B aus der Matrix A durch elementare Umformungen, so überprüft man leicht, daß $\mathrm{Rg}A = \mathrm{Rg}B$ gilt.

Satz 2.5 *Gegeben sind $A \in K_{m,n}$ und $b \in K^m$ sowie das LGS*

$$Ax^t = b^t. \tag{12}$$

i) Das LGS (12) ist genau dann lösbar, wenn $\mathrm{Rg}A = \mathrm{Rg}A_{\mathrm{erw}}$.

ii) Ist (12) lösbar, so gilt $\dim_K \mathbb{L}_h = n - \mathrm{Rg}A$.

iii) Ist (12) lösbar, so existiert genau dann eine eindeutige Lösung, wenn $\mathrm{Rg}A = n$.

Beweis. Wegen der vorangegangenen Überlegungen kann man annehmen, daß die erweiterte Koeffizientenmatrix des LGS Normalform hat:

$$\left(\begin{array}{cccccccc|c} \multicolumn{4}{c}{\overbrace{\hphantom{1 \quad 0 \quad \dots \quad 0}}^{\mathrm{Rg}A}} & \multicolumn{3}{c}{\overbrace{\hphantom{0 \quad \dots \quad 0}}^{n-\mathrm{Rg}A}} & & \\ 1 & 0 & \dots & 0 & & & & & b'_1 \\ 0 & 1 & \dots & 0 & & * & & & \vdots \\ \vdots & & \ddots & & & & & & \vdots \\ 0 & 0 & \dots & 1 & & & & & b'_r \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & & b'_{r+1} \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & & b'_m \end{array}\right).$$

Hieran lassen sich aber die Behauptungen des Satzes direkt ablesen.

□

3. Aufgaben

A 3.1 Gegeben sind die Teilmengen

$$U_1 = \{(x, y, z) \mid x + y = z\},\ U_2 = \{(x, y, z) \mid xy = 0\},$$

$$U_3 = \{(x, y, z) \mid x = 0\},\ U_4 = \{(x, y, z) \mid z \neq 0\}$$

des $\mathbb{Z}_5^3$. Welche U_i sind Unterräume des $\mathbb{Z}_5^3$?

A 3.2 Es sei $\{U_i \mid i \in I\}, I \neq \emptyset$, eine Familie von Unterräumen eines Vektorraumes V mit folgender Eigenschaft:

$$\forall i, j \in I\ \exists k \in I : U_i, U_j \subseteq U_k.$$

Zeigen Sie, daß dann auch

$$U := \bigcup_{i \in I} U_i$$

ein Unterraum von V ist.

A 3.3 Es sei V ein Vektorraum über einem Körper K. Sind $U_1, \ldots, U_n$ Unterräume von V mit $V = U_1 + \ldots + U_n$, so heißt die Summe direkt (geschrieben $V = U_1 \oplus \ldots \oplus U_n$), wenn sich jedes $v \in V$ *eindeutig* als Linearkombination $v = u_1 + \ldots + u_n$ mit $u_i \in U_i, i = 1, \ldots, n$ darstellen läßt. Zeigen Sie, daß für von $\mathcal{O}$ verschiedene Vektoren $v_1, \ldots, v_n$, die V aufspannen, folgende Aussagen äquivalent sind:
(i) $v_1, \ldots, v_n$ sind linear unabhängig über K.
(ii) $V = Kv_1 \oplus \ldots \oplus Kv_n$.

A 3.4 Zeigen Sie für Unterräume $U_1, \ldots, U_n$ eines Vektorraumes V die Äquivalenz folgender Aussagen:
(i) $V = U_1 \oplus \ldots \oplus U_n$.
(ii) $V = U_1 + \ldots + U_n$ und $U_k \cap (U_1 + \ldots + U_{k-1} + U_{k+1} + \ldots + U_n) = \{\mathcal{O}\}$ für $k = 1, \ldots, n$.

A 3.5 Es sei V ein Vektorraum mit der Basis $\{v_1, \ldots, v_n\}$ über einem Körper K mit $\chi(K) \neq 2$. Für welche n bilden die Vektoren $u_1 = v_1 + v_2, u_2 = v_2 + v_3, \ldots, u_n = v_n + v_1$ ebenfalls eine Basis von V? Was gilt, wenn $\chi(K) = 2$ ist?

A 3.6 Berechnen Sie in Abhängigkeit von $a \in \mathbb{R}$ die Dimension des von den Vektoren v_1, v_2, v_3 mit $v_1 = (1, 1, a),\ v_2 = (1, a, 1),\ v_3 = (a, 1, 1)$ aufgespannten Unterraums des $\mathbb{R}^3$.

A 3.7 Sei $U = [v_1, v_2, v_3, v_4]$ ein Unterraum des $\mathbb{R}^4$ mit $v_1 = (1, -7, 1, 3), v_2 = (-3, -4, 2, 6)$, $v_3 = (-2, -11, 3, 9)$, $v_4 = (-4, 3, 1, 3)$. Ermitteln Sie hieraus eine Basis von U und stellen Sie die übrigen Erzeugenden von U als Linearkombination der Basisvektoren dar.

A 3.8 Es sei $M \subseteq V$ eine Teilmenge des Vektorraums V und $v \in M$. Zeigen Sie:
(a) Ist $M \setminus \{v\}$ linear unabhängig und M linear abhängig, dann ist $v \in [M \setminus \{v\}]$.
(b) $v \in [M \setminus \{v\}] \Longrightarrow [M \setminus \{v\}] = [M]$.

A 3.9 Es sei V ein K-Vektorraum, $1 < \dim_K V < \infty$, und U, W Unterräume von V. Welche Aussagen sind äquivalent?
(i) $V = U \oplus W$.
(ii) $\dim_K U + \dim_K W = \dim_K V$.
(iii) $\dim_K U + \dim_K W \geq \dim_K V$ und $U \cap W = \{\mathcal{O}\}$.

A 3.10 Es sei V ein K-Vektorraum und $v_1, v_2, v_3 \in V$ mit $v_1 \neq \mathcal{O}$. Es gelte

$$\dim_K[v_1, v_2, v_3] = 2 \quad \text{und} \quad \dim_K[v_1, v_2] = \dim_K[v_1, v_3].$$

Zeigen Sie $[v_1, v_2] = [v_1, v_3]$ und berechnen Sie $\dim_K[v_1, v_2]$.

A 3.11 Im 4-dimensionalen Vektorraum $\mathbb{R}^4$ sei der Unterraum U von den Vektoren $u_1 = (1, 0, 2, 0)$, $u_2 = (1, 4, 6, 4)$, $u_3 = (2, -2, -1, 0)$, $u_4 = (3, 0, 3, 2)$ und der Unterraum V von den Vektoren $v_1 = (0, 2, 2, 3)$, $v_2 = (-1, 9, 3, 15)$, $v_3 = (-1, 5, 3, 7)$ aufgespannt. Berechnen Sie für die Unterräume $U, V, U + V$ und $U \cap V$ jeweils die Dimension und eine Basis.

A 3.12 Es sei T ein Vektorraum über $\mathbb{R}$ und $\{e_1, e_2, e_3, e_4\}$ eine Basis von T. Ferner sei $u_1 = e_1 + e_2$, $u_2 = e_1 + e_3$, $u_3 = e_3 - e_1$, $v_1 = e_1 + 2e_2$, $v_2 = e_2 + e_4$, $v_3 = e_1 - 2e_4$ sowie $U = [u_1, u_2, u_3], V = [v_1, v_2, v_3]$. Berechnen Sie für die Unterräume $U, V, U + V$ und $U \cap V$ jeweils die Dimension und eine Basis.

A 3.13 Es sei $n \in \mathbb{N} \setminus \{1\}$ sowie $m = 2^n - 1$ und H eine (n, m)-Matrix über $\mathbb{Z}_2$, deren Spalten paarweise verschieden und ungleich $\mathcal{O}^t$ sind. Die Menge $C = \{x \in \mathbb{Z}_2^m \mid Hx^t = \mathcal{O}^t\}$ heißt *Hammingcode*, ihre Elemente *Codewörter*.
(a) Berechnen Sie die Anzahl der Codewörter, d.h. $\mid C \mid$.
(b) Zeigen Sie, daß sich zwei verschiedene Codewörter an mindestens 3 Stellen unterscheiden.

A 3.14 Für welche $a \in \mathbb{Z}_{13}$ besitzt das lineare Gleichungssystem

$$\begin{aligned} ax_1 + 2x_2 + 2x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \\ 2x_1 + x_2 + ax_3 &= 0 \end{aligned}$$

über $\mathbb{Z}_3$ nicht-triviale Lösungen?

A 3.15 Gegeben ist das folgende lineare Gleichungssystem über dem beliebigen Körper K:

$$\begin{array}{rcrcrcrcr} x_1 & - & 2x_2 & - & 5x_3 & - & 2x_4 & = & a+1 \\ -x_1 & + & x_2 & + & 4x_3 & + & 3x_4 & = & -2a-1 \\ -2x_1 & - & x_2 & + & 6x_3 & + & 3x_4 & = & -a \\ -3x_1 & & & + & 11x_3 & + & 6x_4 & = & -2. \end{array}$$

(a) Bringen Sie das LGS durch elementare Umformungen auf Normalform.
(b) Zeigen Sie, daß das LGS im Falle $K = \mathbb{Q}$ für jedes $a \in \mathbb{Q}$ eindeutig lösbar ist.
(c) Erörtern Sie die Lösbarkeit des LGS für die Fälle $\chi(K) = 2$ und $\chi(K) = 3$ und geben Sie jeweils die Lösungsmengen an.

KAPITEL 4

Lineare Abbildungen

1. Grundlagen

Definition 1.1 *V und W seien K-Vektorräume. Eine Funktion $\varphi : V \longrightarrow W$ heißt lineare Abbildung (Homomorphismus), wenn für alle $v, v' \in V$ und für alle $k \in K$ gilt*

i) $\varphi(v + v') = \varphi(v) + \varphi(v')$.

ii) $\varphi(k \cdot v) = k \cdot \varphi(v)$.

Die Menge aller linearen Abbildungen $\varphi : V \longrightarrow W$ wird mit $\mathrm{Hom}(V, W)$ bezeichnet.

Bemerkung. Sei $\varphi \in \mathrm{Hom}(V, W)$.

1. Mit i) ist $\varphi(\mathcal{O}) = \varphi(\mathcal{O} + \mathcal{O}) = \varphi(\mathcal{O}) + \varphi(\mathcal{O})$, und damit gilt $\varphi(\mathcal{O}) = \mathcal{O}$, wobei $\mathcal{O}$ den Nullvektor sowohl in V als auch in W bezeichnet.

2. Mit vollständiger Induktion folgt aus den Linearitätsbedingungen i) und ii) in Definition 1.1 für beliebige $n \in \mathbb{N}$
$$\varphi(k_1 \cdot v_1 + \ldots + k_n \cdot v_n) = k_1 \cdot \varphi(v_1) + \ldots + k_n \cdot \varphi(v_n).$$

3. Wegen Bemerkungen 1 und 2 folgt aus der linearen Unabhängigkeit der Vektoren $\varphi(v_1), \ldots, \varphi(v_n)$ in W die lineare Unabhängigkeit der Vektoren $v_1, \ldots, v_n$ in V:
$$\mathcal{O} = k_1 v_1 + \ldots + k_n v_n \Longrightarrow \mathcal{O} = k_1 \varphi(v_1) + \ldots + k_n \varphi(v_n) \Longrightarrow 0 = k_1 = \ldots = k_n.$$

4. Ist φ bijektiv, so heißt φ Isomorphismus. Dann ist auch $\varphi^{-1} : W \longrightarrow V$ ein Isomorphismus. Ein Isomorphismus $\varphi : V \longrightarrow V$ heißt Automorphismus von V. Die Automorphismen von V bilden bezüglich der Komposition eine Gruppe, die mit $\mathrm{Aut}(V)$ bezeichnet wird.

Beispiel. Die folgenden Abbildungen $\varphi : V \longrightarrow V$ sind Homomorphismen.

1. Die Nullabbildung, die für alle $v \in V$ durch $\varphi(v) = \mathcal{O}$ gegeben ist.

2. Die Identität, die für alle $v \in V$ durch $\varphi(v) = v$ gegeben ist. Sie ist sogar ein Isomorphismus.

3. Die Multiplikation mit $k \in K$, also $\varphi(v) = k \cdot v$ für alle $v \in V$. Für $k = 0$ ist φ dann die Nullabbildung, für $k = 1$ die Identität.

Satz 1.2 *V und W seien K-Vektorräume und B sei eine Basis von V. Ist jedem $v \in B$ ein $v' \in W$ zugeordnet, so gibt es genau eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\varphi(v) = v'$ für alle $v \in B$.*

Beweis. Existenz: Da B eine Basis ist, läßt sich nach Satz 1.13 in Kapitel 3 jedes $v \in V$ eindeutig als Linearkombination

$$v = k_1 \cdot v_1 + \ldots + k_n \cdot v_n$$

darstellen, wobei $v_1, \ldots, v_n \in B$ verschieden sind und $k_1, \ldots, k_n \in K$. Definieren wir die Abbildung $\varphi : V \longrightarrow V$ durch

$$\varphi(v) := k_1 \cdot v_1' + \ldots + k_n \cdot v_n' \,, \tag{13}$$

so ist φ linear: Für $k \in K$ folgt

$$k \cdot v = kk_1 \cdot v_1 + \ldots + kk_n \cdot v_n,$$

also

$$\begin{aligned} \varphi(k \cdot v) &= kk_1 \cdot v_1' + \ldots + kk_n \cdot v_n' \\ &= k(k_1 \cdot v_1' + \ldots + k_n \cdot v_n') = k \cdot \varphi(v), \end{aligned}$$

und analog ergibt sich für alle $u, v \in V$

$$\varphi(v + u) = \varphi(v) + \varphi(u) \,.$$

Daß für φ auch die Beziehung $\varphi(v) = v'$ für alle Basiselemente $v \in B$ gilt, folgt direkt aus der Darstellung $v = 1 \cdot v$ und $\varphi(v) = 1 \cdot v' = v'$.
Eindeutigkeit: Sei $\psi \in \mathrm{Hom}(V, W)$ mit $\psi(v) = v'$ für alle $v \in B$. Dann gilt

$$\begin{aligned} \psi(k_1 \cdot v_1 + \ldots + k_n \cdot v_n) &= k_1 \cdot \psi(v_1) + \ldots + k_n \cdot \psi(v_n) \\ &= k_1 \cdot v_1' + \ldots + k_n \cdot v_n' \,, \end{aligned}$$

d.h. $\varphi = \psi$. □

Beispiel. Seien $V = \mathbb{R}^3$ mit der kanonischen Basis $B = \{e_1, e_2, e_3\}$ (vgl. Beispiel 1 nach Definition 1.14 aus Kapitel 3) und $W = \mathbb{R}^2$ gegeben. Wählen wir in Analogie zu Satz 1.2

$$e_1' = (1, 2) \,, \quad e_2' = (-1, 1) \,, \quad e_3' = (-2, -1) \,,$$

dann ist für alle $(x, y, z) \in \mathbb{R}^3$:

$$\begin{aligned} \varphi((x, y, z)) &= \varphi(x \cdot e_1 + y \cdot e_2 + z \cdot e_3) = x \cdot (1, 2) + y \cdot (-1, 1) + z \cdot (-2, -1) = \\ &= (x - y - 2z, 2x + y - z) \,. \end{aligned}$$

Satz 1.3 *V und W seien K-Vektorräume und $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann gilt*

i) $M \subseteq V \Longrightarrow \varphi([M]) = [\varphi(M)]$.

ii) $U \subseteq\mid V \Longrightarrow \varphi(U) \subseteq\mid W$.

iii) $U \subseteq\mid V \Longrightarrow \dim_K \varphi(U) \leq \dim_K U$.

Beweis.

i) Ist $M = \emptyset$, so ist auch $\varphi(M) = \emptyset$ und $\varphi([\emptyset]) = \varphi(\{\mathcal{O}\}) = \{\mathcal{O}\} = [\emptyset]$.
Sei also $M \neq \emptyset$. Dann existiert für jedes $w \in \varphi([M])$ ein $v \in [M]$ mit $\varphi(v) = w$. Wegen Satz 1.8 aus Kapitel 3 gibt es $u_1, \ldots, u_n \in M$ und $k_1, \ldots, k_n \in K$, so daß $v = k_1u_1 + \ldots + k_nu_n$, also

$$\begin{aligned} w &= \varphi(k_1 \cdot u_1 + \ldots + k_n \cdot u_n) \\ &= k_1 \cdot \varphi(u_1) + \ldots + k_n \cdot \varphi(u_n) \in [\varphi(M)] \,. \end{aligned}$$

Damit gilt $\varphi([M]) \subseteq [\varphi(M)]$. Ist andererseits $w \in [\varphi(M)]$, so gibt es $u_1, \ldots, u_n \in M$ und $k_1, \ldots, k_n \in K$ mit $w = k_1\varphi(u_1) + \ldots + k_n\varphi(u_n)$, d.h. $w = \varphi(k_1u_1 + \ldots + k_nu_n) \in \varphi([M])$. Damit gilt auch $[\varphi(M)] \subseteq \varphi([M])$.

ii) Ist $U \subseteq\mid V$, so gilt $U = [U]$. Mit i) folgt daraus $\varphi(U) = [\varphi(U)] \subseteq\mid W$.

iii) Sind $\varphi(v_1), \ldots, \varphi(v_n)$ $(n \in \mathbb{N})$ linear unabhängig, so sind wegen Bemerkung 3 nach Definition 1.1 auch $v_1, \ldots, v_n$ linear unabhängig. Gilt also $\dim_K \varphi(U) = \infty$, so folgt $\dim_K U = \infty$ unmittelbar. Ist aber $\{\varphi(u_1), \ldots, \varphi(u_n)\}$ eine Basis von $\varphi(U)$, so kann $\{u_1, \ldots, u_n\}$ zu einer Basis von U ergänzt werden, das heißt $\dim_K \varphi(U) \leq \dim_K U$.

□

Bemerkung. Ist $\varphi : V \longrightarrow W$ linear, so ist $\mathrm{Bild}\varphi := \varphi(V)$ ein Unterraum von W.

Definition 1.4 *V und W seien K-Vektorräume und $\varphi : V \longrightarrow W$ sei linear. Dann heißt*

$$\mathrm{Rg}\varphi := \dim_K \mathrm{Bild}\varphi$$

Rang von φ.

Satz 1.5 *Sind V und W endlich dimensionale K-Vektorräume und ist $\varphi : V \longrightarrow W$ linear, dann sind die folgenden Aussagen äquivalent.*

i) $\varphi : V \longrightarrow W$ *ist surjektiv.*

ii) Für jede Basis B von V gilt $[\varphi(B)] = W$.

iii) $\mathrm{Rg}\varphi = \dim_K W$.

Beweis. i)⇒ii): Aus der Surjektivität von φ folgt $W = \varphi(V)$. Da B eine Basis von V ist, gilt $\varphi(V) = \varphi([B]) = [\varphi(B)]$ wegen Satz 1.3.
ii)⇒iii): Nach Definition 1.4 ist $\mathrm{Rg}\varphi = \dim_K \varphi(V) = \dim_K \varphi([B])$, da B eine Basis von V ist. Mit Satz 1.3 folgt $\dim_K \varphi([B]) = \dim_K[\varphi(B)] = \dim_K W$.
iii)⇒i): Aus $\dim_K \varphi(V) = \mathrm{Rg}\varphi = \dim_K W$ folgt $\varphi(V) = W$ mit Korollar 1.21 aus Kapitel 3, also die Behauptung.

□

Satz 1.6 *V und W seien K-Vektorräume und $\varphi : V \longrightarrow W$ sei linear. Ist U ein Unterraum von W, dann ist*

$$\varphi^{-}(U) := \{v \in V \mid \varphi(v) \in U\}$$

ein Unterraum von V.

Beweis. Wir benutzen Satz 1.3 aus Kapitel 3.

i) Wegen $\varphi(\mathcal{O}) = \mathcal{O} \in U$ folgt $\mathcal{O} \in \varphi^{-}(U)$, also $\varphi^{-}(U) \neq \emptyset$.

ii) Für alle $u, v \in \varphi^{-}(U)$ gilt $\varphi(u), \varphi(v) \in U$. Da U ein Unterraum von W ist, folgt $\varphi(u+v) = \varphi(u) + \varphi(v) \in U$, also $u + v \in \varphi^{-}(U)$.

iii) Für $v \in \varphi^{-}(U)$ und $k \in K$ gilt $\varphi(k \cdot v) = k \cdot \varphi(v) \in U$, da $\varphi(v) \in U$, das heißt $k \cdot v \in \varphi^{-}(U)$.

□

Bemerkung. Insbesondere ist $\varphi^{-}(\{\mathcal{O}\})$ ein Unterraum von V.

Definition 1.7 *Sind V und W zwei K-Vektorräume und ist $\varphi : V \longrightarrow W$ linear, dann heißt*

$$\mathrm{Kern}\varphi := \varphi^{-}(\{\mathcal{O}\})$$

Kern von φ, und

$$\mathrm{Def}\varphi := \dim_K \mathrm{Kern}\varphi$$

heißt Defekt von φ.

Satz 1.8 *V und W seien K-Vektorräume und $\varphi : V \longrightarrow W$ sei linear. φ ist genau dann injektiv, wenn* $\mathrm{Kern}\varphi = \{\mathcal{O}\}$.

Beweis. "⇒": Nach Definition 1.7 ist $\mathrm{Kern}\varphi = \{v \in V \mid \varphi(v) = \mathcal{O}\}$. Da φ injektiv ist, hat $\mathcal{O}$ genau ein Urbild, d. h. $\mathrm{Kern}\varphi = \{\mathcal{O}\}$.
"⇐": Seien $u, v \in V$ mit $\varphi(u) = \varphi(v)$. Dann ist $\varphi(u - v) = \varphi(u) - \varphi(v) = \mathcal{O}$, und damit $u - v \in \mathrm{Kern}\varphi = \{\mathcal{O}\}$, also $u - v = \mathcal{O}$. Es folgt $u = v$.

□

Beispiel.

1. Sei $B = \{e_1, e_2, e_3\}$ die kanonische Basis von $\mathbb{R}^3$ und $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ linear mit

$$\varphi(e_1) = (1,1)\,, \quad \varphi(e_2) = (-1,0)\,, \quad \varphi(e_3) = (-2,-1)\,.$$

Dann ist

$$\varphi((x,y,z)) = \varphi(x \cdot e_1 + y \cdot e_2 + z \cdot e_3) = (x - y - 2z, x - z)\,.$$

Zunächst wollen wir den Kern von φ berechnen. Genau die Vektoren $(x,y,z) \in \mathbb{R}^3$ werden durch φ auf den Nullvektor abgebildet, für die $\varphi((x,y,z)) = (0,0)$ gilt. Also muß das folgende LGS

$$\begin{aligned} x - y - 2z &= 0 \\ x - z &= 0 \end{aligned}$$

gelöst werden. Die erweiterte Koeffizientenmatrix (vergleiche Kapitel 3.2)

$$A_{\text{erw}} = \left(\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right)$$

läßt sich durch elementare Umformungen folgendermaßen vereinfachen:

$$\left(\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right) \quad \leadsto \quad \left(\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right) \quad \leadsto \quad \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right).$$

Mit Lemma 2.3 und Gleichung (11) aus Kapitel 3 ist $\mathbb{L} = \{\lambda(-1,1,-1) \mid \lambda \in \mathbb{R}\}$, also

$$\text{Kern}\varphi = \{\lambda(-1,1,-1) \mid \lambda \in \mathbb{R}\}$$

und $\text{Def}\varphi = 1$. Um nun das Bild von φ zu berechnen genügt es, eine Basis von $\varphi(\mathbb{R}^3)$ zu ermitteln. Aus

$$\varphi(-1,0,-1) = (1,0) \quad \text{und} \quad \varphi(0,2,-1) = (0,1)$$

folgt $\varphi(\mathbb{R}^3) = \mathbb{R}^2$, also $\text{Rg}\varphi = 2$.

2. Wir betrachten $V = C^1([a,b]) = \{f : [a,b] \longrightarrow \mathbb{R} \mid f \text{ ist stetig differenzierbar}\}$ und $W = C^0([a,b]) = \{f : [a,b] \longrightarrow \mathbb{R} \mid f \text{ ist stetig}\}$ (vergleiche Kapitel 3, Beispiel 2 nach Satz 1.3). Dann ist die Abbildung

$$\varphi : V \longrightarrow W\,, \quad f \longmapsto f + f'$$

linear, denn für alle $f, g \in V$ und $k \in \mathbb{R}$ gilt

$$\begin{aligned} \varphi(f+g) &= f + g + (f+g)' = f + f' + g + g' = \varphi(f) + \varphi(g) \quad \text{und} \\ \varphi(k \cdot f) &= k \cdot f + (k \cdot f)' = k \cdot f + k \cdot f' = k \cdot (f + f') = k \cdot \varphi(f)\,. \end{aligned}$$

Zunächst ermitteln wir den Kern von φ. Er besteht aus denjenigen Elementen f von V, für die $\varphi(f)$ die Nullabbildung ist. Es muß also $f + f' = \mathcal{O}$ und damit $f(x) + f'(x) = 0$

für alle $x \in [a,b]$ gelten. Aus der Theorie der gewöhnlichen Differentialgleichungen ergibt sich

$$\text{Kern}\varphi = \{ke^{-x} \mid k \in \mathbb{R}\}, \quad \text{also Def}\varphi = 1.$$

Um nun das Bild von φ zu berechnen, stellen wir zunächst fest, daß mit $f \in C^0([a,b])$ auch die Funktion $\exp \cdot f : [a,b] \longrightarrow \mathbb{R}, x \longmapsto e^x f(x)$ im Intervall $[a,b]$ stetig ist und somit eine Stammfunktion besitzt:

$$F(x) = \int_a^x e^t f(t) dt.$$

Es gilt $F \in C^1([a,b])$ mit $F'(x) = e^x f(x)$. Definieren wir die Funktion $g \in C^1([a,b])$ durch $g(x) := e^{-x} F(x)$, so ist

$$\varphi(g) = e^{-x}F(x) - e^{-x}F(x) + e^{-x}e^x f(x) = f(x).$$

Jedes $f \in C^0([a,b])$ hat also ein Urbild bezüglich φ, das heißt

$$\text{Bild}\varphi = C^0([a,b]) \quad \text{und} \quad \text{Rg}\varphi = \infty.$$

Lemma 1.9 *V und W seien K-Vektorräume und $\varphi : V \longrightarrow W$ sei linear. Ist B eine Basis von* Kernφ *und B' eine Basis von V mit $B \subseteq B'$, so gilt:*

i) Sind $v_1, \ldots, v_n \in B' \setminus B$ verschieden, dann sind die Vektoren $\varphi(v_1), \ldots, \varphi(v_n)$ linear unabhängig.

ii) $\varphi(B' \setminus B)$ ist eine Basis von $\varphi(V)$.

Beweis.

i) Seien $k_1, \ldots, k_n \in K$, so daß $k_1\varphi(v_1) + \ldots + k_n\varphi(v_n) = \mathcal{O}$, also $\varphi(k_1v_1 + \ldots + k_nv_n) = \mathcal{O}$. Dann folgt $k_1v_1 + \ldots + k_nv_n \in \text{Kern}\varphi$, und es gibt $w_1, \ldots, w_m \in B$ und $l_1, \ldots, l_m \in K$ mit
$$k_1v_1 + \ldots + k_nv_n = l_1w_1 + \ldots + l_mw_m\,.$$
Da $v_1, \ldots, v_n, w_1, \ldots, w_m \in B'$ linear unabhängig sind, folgt $k_1 = \ldots = k_n = l_1 = \ldots = l_m = 0$.

ii) Der Fall $B' = \emptyset$ ist klar. Wir nehmen also $B' \neq \emptyset$ an. Mit i) folgt, daß $\varphi(B' \setminus B)$ linear unabhängig ist. Es bleibt zu zeigen, daß diese Menge $\varphi(V)$ erzeugt. Dazu sei $\varphi(v) \in \varphi(V)$. Dann gibt es $v_1, \ldots, v_n \in B'$ und $k_1, \ldots, k_n \in K$ mit $v = k_1v_1 + \ldots + k_nv_n$. O.B.d.A. gelte $v_1, \ldots, v_i \in B, v_{i+1}, \ldots, v_n \in B' \setminus B$, und es ergibt sich
$$\begin{aligned} \varphi(v) &= k_1\varphi(v_1) + \ldots + k_i\varphi(v_i) + k_{i+1}\varphi(v_{i+1}) + \ldots + k_n\varphi(v_n) = \\ &= k_{i+1}\varphi(v_{i+1}) + \ldots + k_n\varphi(v_n) \in [\varphi(B' \setminus B)]. \end{aligned}$$

□

Satz 1.10 *V und W seien K-Vektorräume mit $\dim_K V < \infty$ und $\varphi : V \longrightarrow W$ sei linear. Dann gilt*

$$\dim_K V = \text{Rg}\varphi + \text{Def}\varphi\,.$$

Beweis. Mit Satz 1.18 aus Kapitel 3 kann eine Basis B von Kernφ zu einer Basis B' von V ergänzt werden. Wegen Lemma 1.9 gilt $|B' \setminus B| = |\varphi(B' \setminus B)|$, und es folgt

$$\dim_K V = |B'| = |B' \setminus B| + |B| = \mathrm{Rg}\varphi + \mathrm{Def}\varphi.$$

□

Satz 1.11 *Es seien V und W endlich-dimensionale K-Vektorräume und $\varphi : V \longrightarrow W$ sei linear. Dann sind folgende Aussagen äquivalent.*

i) φ ist ein Isomorphismus.

ii) Ist B eine Basis von V, so gilt $|B| = |\varphi(B)|$ und $\varphi(B)$ ist eine Basis von W.

iii) $\dim_K V = \dim_K W = \mathrm{Rg}\varphi$.

Beweis. i)⇒ii): Aus der Injektivität von φ folgt $|B| = |\varphi(B)|$ und Kern$\varphi = \{\mathcal{O}\}$. Damit ist $\emptyset$ eine Basis von Kernφ, und wegen Lemma 1.9 ist demnach $\varphi(B \setminus \emptyset) = \varphi(B)$ eine Basis von $\varphi(V)$. Da φ surjektiv ist, gilt $\varphi(V) = W$.
ii)⇒iii): Sei B eine Basis von V. Dann ist $\varphi(B)$ eine Basis von W, und mit Satz 1.3 ergibt sich $W = [\varphi(B)] = \varphi([B]) = \varphi(V)$. Daraus folgt

$$\mathrm{Rg}\varphi = \dim_K \varphi(V) = \dim_K W = |\varphi(B)| = |B| = \dim_K V\,.$$

iii)⇒i): Wegen Satz 1.10 gilt Def$\varphi = 0$, also Kern$\varphi = \{\mathcal{O}\}$, und wegen Satz 1.8 ist damit φ injektiv. Wegen Satz 1.5 ist φ auch surjektiv.

□

Definition 1.12 *V und W seien K-Vektorräume. V und W heißen isomorph (geschrieben $V \cong W$), wenn es einen Isomorphismus $\varphi : V \longrightarrow W$ gibt.*

Bemerkung. Durch $'' \cong ''$ wird eine Äquivalenzrelation erklärt.

Korollar 1.13 *Sind V und W endlich-dimensionale K-Vektorräume, so gilt $V \cong W$ genau dann, wenn* $\dim_K V = \dim_K W$.

Beweis. $''\Rightarrow''$: Folgt direkt aus Satz 1.11.
$''\Leftarrow''$: Sei $\{v_1, \ldots, v_n\}$ eine Basis von V und sei $\{u_1, \ldots, u_n\}$ eine Basis von W. Dann folgt aus Satz 1.2, daß eine lineare Abbildung $\varphi : V \longrightarrow W$ existiert mit $\varphi(v_i) = u_i$ für alle $i = 1, \ldots, n$. Wegen Satz 1.3 ist dann

$$\varphi(V) = \varphi([v_1, \ldots, v_n]) = [\varphi(v_1), \ldots, \varphi(v_n)] = [u_1, \ldots, u_n],$$

also Rg$\varphi = n$. Aus Satz 1.11 ergibt sich damit, daß φ ein Isomorphismus ist, d.h. $V \cong W$.

□

Bemerkung. Ist V ein K-Vektorraum mit $\dim_K V = n < \infty$, so gilt insbesondere $V \cong K^n$. Jeder vermittelnde Isomorphismus kann dabei als Koordinatensystem gedeutet werden.

Koordinaten und Koordinatensysteme.
Sei $\dim_K V = n$ und $B := \{v_1, \ldots, v_n\}$ eine Basis von V. Dann folgt mit Satz 1.2, daß es genau eine lineare Abbildung φ_B mit

$$\varphi_B : V \longrightarrow K^n \quad \text{und} \quad \varphi_B(v_1) = e_1, \ldots, \varphi_B(v_n) = e_n$$

gibt. Dabei ist $\{e_1, \ldots, e_n\}$ die kanonische Basis des K^n und φ_B ein Isomorphismus, den man als Koordinatensystem bezeichnet. Für jedes $x \in V$ heißt $\varphi_B(x) = (x_1, \ldots, x_n)$ Koordinatenvektor von x bezüglich B, die x_i heißen Koordinaten von x bezüglich B, die sich auch folgendermaßen deuten lassen: Ist $x \in V$, so gibt es $x_1, \ldots, x_n \in K$ mit

$$x = x_1 v_1 + \ldots + x_n v_n.$$

Wegen Satz 1.13 aus Kapitel 3 sind $x_1, \ldots, x_n$ eindeutig bestimmt, und es gilt

$$\begin{aligned} \varphi_B(x) &= x_1 \varphi_B(v_1) + \ldots + x_n \varphi_B(v_n) = \\ &= x_1 e_1 + \ldots + x_n e_n = (x_1, \ldots, x_n). \end{aligned}$$

Beispiel. Wir betrachten den reellen Vektorraum $V = \mathbb{R}^3$ mit der Basis $B = \{v_1, v_2, v_3\}$, wobei $v_1 = (1, 2, 1)$, $v_2 = (-1, 1, 1)$ und $v_3 = (-1, 2, 1)$. Gesucht sind die Koordinaten x_1, x_2, x_3 des Vektors $x = (-4, 2, 2)$ bezüglich B. Es gilt

$$\begin{aligned} x &= x_1 \cdot v_1 + x_2 \cdot v_2 + x_3 \cdot v_3\,, \text{ d.h.} \\ (-4, 2, 2) &= x_1 \cdot (1, 2, 1) + x_2 \cdot (-1, 1, 1) + x_3 \cdot (-1, 2, 1)\,. \end{aligned}$$

Die Koordinaten x_1, x_2, x_3 bilden also die eindeutige Lösung des linearen Gleichungssystems

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

Es folgt $x_1 = -1, x_2 = 2, x_3 = 1$, und damit ist $(-1, 2, 1)$ der Koordinatenvektor von x bezüglich der Basis B.

2. Die Vektorräume Hom(V, W) und $K_{m,n}$

Im folgenden sind V, W zwei K-Vektorräume sowie $\mathrm{Hom}(V, W) = \{f : V \longrightarrow W \mid f \text{ linear}\}$ wie im vorangehenden Abschnitt beschrieben. Für $f, g \in \mathrm{Hom}(V, W)$ und $k \in K$ definieren wir

$$\begin{aligned} f + g &: V \longrightarrow W\,, \qquad v \longmapsto f(v) + g(v), && (14) \\ k \cdot f &: V \longrightarrow W\,, \qquad v \longmapsto k \cdot f(v)\,. && (15) \end{aligned}$$

Dann gilt $f + g \in \mathrm{Hom}(V, W)$, denn für alle $v, v' \in V$ ist

$$\begin{aligned} (f+g)(v+v') &= f(v+v') + g(v+v') \\ &= f(v) + f(v') + g(v) + g(v') \\ &= (f+g)(v) + (f+g)(v'), \end{aligned}$$

und für alle $v \in V, k \in K$ folgt

$$\begin{aligned}(f+g)(k \cdot v) &= f(k \cdot v) + g(k \cdot v) = k \cdot f(v) + k \cdot g(v) \\ &= k \cdot (f(v) + g(v)) = k \cdot (f+g)(v)\,.\end{aligned}$$

Entsprechend beweist man $kf \in \mathrm{Hom}(V, W)$.

Satz 2.1 *Die Menge* $\mathrm{Hom}(V, W)$ *ist bezüglich der durch (14) und (15) definierten Verknüpfungen ein* K*-Vektorraum.*

Beweis. Das neutrale Element der Addition ist die Nullabbildung

$$\mathcal{O} : V \longrightarrow W\,; \quad v \longmapsto \mathcal{O}.$$

Es gilt $\mathcal{O} \in \mathrm{Hom}(V, W)$. Bei dieser Schreibweise bezeichnet $\mathcal{O}$ einerseits den Nullvektor von W und andererseits die Nullabbildung.
Das additive Inverse von $f \in \mathrm{Hom}(V, W)$ ist

$$-f : V \longrightarrow W\,; \quad v \longmapsto -f(v)\,,$$

da

$$f + (-f) : V \longrightarrow W\,; \quad v \longmapsto f(v) - f(v) = \mathcal{O}\,,$$

und es gilt $-f \in \mathrm{Hom}(V, W)$. Alle anderen Vektorraumeigenschaften lassen sich jetzt leicht nachprüfen.

$\square$

Definieren wir nun weiterhin auf der Menge $K_{m,n} = \{A \mid A \text{ ist eine } (m,n)\text{-Matrix über } K\}$ die Addition

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix} \tag{16}$$

sowie die Multiplikation mit Skalaren

$$k \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} := \begin{pmatrix} ka_{11} & \dots & ka_{1n} \\ \vdots & & \vdots \\ ka_{m1} & \dots & ka_{mn} \end{pmatrix}, \tag{17}$$

so folgt unmittelbar

Satz 2.2 *Die Menge* $K_{m,n}$ *ist bezüglich der durch (16) und (17) gegebenen Verknüpfungen ein* K*-Vektorraum. Es gilt* $\dim_K K_{m,n} = m \cdot n$.

Wir wollen im folgenden den Zusammenhang zwischen dem Vektorraum $\mathrm{Hom}(V,W)$ der linearen Abbildungen und dem Vektorraum $K_{m,n}$ der (m,n)-Matrizen darstellen, wobei $\dim_K V = n$ und $\dim_K W = m$ gilt. Dazu sei $B_V = \{v_1,\dots,v_n\}$ eine K-Basis von V und $B_W = \{w_1,\dots,w_m\}$ eine K-Basis von W. Ist $\varphi \in \mathrm{Hom}(V,W)$ mit

$$\varphi(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_m\,, \quad j = 1,\dots,n\,,$$

so definieren wir $A_\varphi \in K_{m,n}$ durch

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

A_φ heißt Matrixdarstellung von φ bezüglich B_V und B_W. In der j-ten Spalte von A_φ stehen also die Koordinaten von $\varphi(v_j)$ bezüglich der Basis $\{w_1,\dots,w_m\}$ von W.

Satz 2.3 *Sind V und W endlich-dimensionale K-Vektorräume mit den Basen B_V bzw. B_W, dann ist die Abbildung*

$$\Theta_{B_V,B_W} : \mathrm{Hom}(V,W) \longrightarrow K_{m,n}\,, \quad \varphi \longmapsto A_\varphi\,,$$

ein Isomorphismus.

Beweis. Sei $B_V = \{v_1,\dots,v_n\}$ und $B_W = \{w_1,\dots,w_m\}$.
Θ_{B_V,B_W} ist linear: Sind $\varphi,\psi \in \mathrm{Hom}(V,W)$ mit $A_\varphi = (a_{ij})$ und $A_\psi = (b_{ij})$, dann gilt nach Definition $\varphi(v_j) = \sum_{i=1}^m a_{ij}w_i$ und $\psi(v_j) = \sum_{i=1}^m b_{ij}w_i$ für $j = 1,\dots,n$. Es folgt

$$(\varphi+\psi)(v_j) = \varphi(v_j) + \psi(v_j) = \sum_{i=1}^m (a_{ij}w_i + b_{ij}w_i) = \sum_{i=1}^m (a_{ij}+b_{ij})w_i$$

und damit

$$A_{\varphi+\psi} = (a_{ij}+b_{ij}) = (a_{ij}) + (b_{ij}) = A_\varphi + A_\psi.$$

Für $k \in K$ ist weiterhin

$$(k\cdot\varphi)(v_j) = k\cdot(\varphi(v_j)) = \sum_{i=1}^m k a_{ij} w_i\,,$$

also

$$A_{k\varphi} = (ka_{ij}) = k(a_{ij}) = k\cdot A_\varphi.$$

Θ_{B_V,B_W} ist injektiv: Wegen Satz 1.8 ist $\mathrm{Kern}\Theta_{B_V,B_W} = \{\mathcal{O}\}$ zu zeigen, wobei die Inklusion $\{\mathcal{O}\} \subseteq \mathrm{Kern}\Theta_{B_V,B_W}$ trivial ist. Sei $\varphi \in \mathrm{Kern}\Theta_{B_V,B_W}$, d.h.

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

Dann gilt für alle $j = 1, \ldots, n$

$$\varphi(v_j) = 0 \cdot w_1 + \ldots + 0 \cdot w_m = \mathcal{O},$$

und mit Satz 1.2 folgt $\varphi = \mathcal{O}$, d.h. $\mathrm{Kern}\Theta_{B_V,B_W} \subseteq \{\mathcal{O}\}$.
Θ_{B_V,B_W} ist surjektiv: Sei $A = (a_{ij}) \in K_{m,n}$ gegeben. Dann folgt aus Satz 1.2, daß es ein $\varphi \in \mathrm{Hom}(V, W)$ mit

$$\varphi(v_j) = a_{1j}w_1 + \ldots + a_{mj}w_m$$

für $j = 1, \ldots, n$ gibt. Somit gilt $A_\varphi = (a_{ij})$. □

Korollar 2.4 *Sind V und W zwei K-Vektorräume mit $\dim_K V = n$ und $\dim_K W = m$, dann gilt*

$$\dim_K \mathrm{Hom}(V, W) = m \cdot n\,.$$

Beweis. Die Behauptung folgt sofort aus den Sätzen 2.2 und 2.3. □

Wir wollen nun noch untersuchen, wie sich die Matrix der Hintereinanderausführung zweier linearer Abbildungen aus den beiden Matrizen der einzelnen Abbildungen ermitteln läßt. Dazu betrachten wir die K-Vektorräume U, V, W und die linearen Abbildungen $\varphi : V \longrightarrow W$, $\psi : W \longrightarrow U$. Die Komposition $\psi \circ \varphi : V \longrightarrow U$ ist ebenfalls linear, denn für alle $v, v' \in V$ und $k \in K$ gilt

$$\begin{aligned} \psi \circ \varphi(v + v') &= \psi(\varphi(v + v')) = \psi(\varphi(v) + \varphi(v')) \\ &= \psi(\varphi(v)) + \psi(\varphi(v')) = \psi \circ \varphi(v) + \psi \circ \varphi(v') \end{aligned}$$

und

$$\psi \circ \varphi(k \cdot v) = \psi(\varphi(k \cdot v)) = \psi(k \cdot \varphi(v)) = k \cdot \psi(\varphi(v)) = k \cdot (\psi \circ \varphi(v))\,.$$

Seien $B_V = \{v_1, \ldots, v_n\}$, $B_W = \{w_1, \ldots, w_m\}$ und $B_U = \{u_1, \ldots, u_l\}$ Basen von V, W und U sowie $A_\varphi = (a_{ij}) \in K_{m,n}$ und $A_\psi = (b_{ij}) \in K_{l,m}$ die Matrixdarstellungen der Abbildungen φ und ψ bezüglich dieser Basen. Ist dann weiterhin $A_{\psi\circ\varphi} = (c_{ij}) \in K_{l,n}$ die Matrixdarstellung von $\psi \circ \varphi$ bezüglich B_V und B_U, so gilt einerseits $\psi \circ \varphi(v_j) = \sum_{i=1}^{l} c_{ij}u_i$ und andererseits

$$\begin{aligned} \psi \circ \varphi(v_j) &= \psi(\varphi(v_j)) = \psi(\sum_{k=1}^{m} a_{kj}w_k) \\ &= \sum_{k=1}^{m} a_{kj}\psi(w_k) = \sum_{k=1}^{m} a_{kj} \sum_{i=1}^{l} b_{ik}u_i \\ &= \sum_{i=1}^{l}(\sum_{k=1}^{m} b_{ik}a_{kj})u_i\,. \end{aligned}$$

Daraus ergibt sich

$$c_{ij} = \sum_{k=1}^{m} b_{ik}a_{kj}\,.$$

Definition 2.5 *Sind $(a_{ij}) \in K_{m,n}$ und $(b_{ij}) \in K_{l,m}$ zwei Matrizen, dann heißt die Matrix $(c_{ij}) \in K_{l,n}$ mit*

$$c_{ij} = \sum_{k=1}^{m} b_{ik} a_{kj}$$

das Produkt von (b_{ij}) und (a_{ij}), geschrieben $(c_{ij}) = (b_{ij}) \cdot (a_{ij})$.

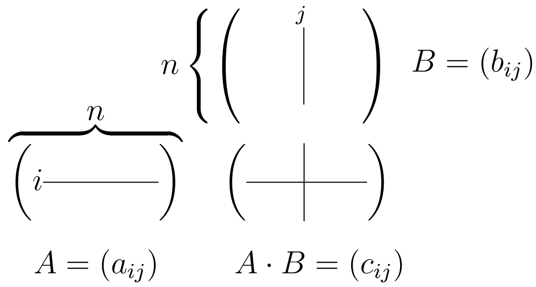
Bemerkung.

1. Mit den Bezeichnungen wie oben gilt

$$\boxed{A_{\psi\circ\varphi} = A_\psi \cdot A_\varphi \,.}$$

2. Sind $A \in K_{m,n}$ und $B \in K_{r,s}$ Matrizen, so ist $A \cdot B$ genau dann definiert, wenn $n = r$, wenn also die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist.

Rechenschema.



Das Element (c_{ij}) ergibt sich durch "Multiplikation" der i-ten Zeile von A mit der j-ten Spalte von B.

Deutung der Matrixdarstellung.
Ist V ein K-Vektorraum mit der Basis $B = \{v_1, \ldots, v_n\}$ und dem Koordinatensystem φ_B und W entsprechend mit der Basis $B' = \{w_1, \ldots, w_m\}$ und dem Koordinatensystem $\varphi_{B'}$, dann haben wir für eine lineare Abbildung $\varphi : V \longrightarrow W$ das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ \varphi_B \downarrow & & \downarrow \varphi_{B'} \\ K^n & \xrightarrow{?} & K^m \end{array} \qquad\qquad \begin{array}{ccc} x & \xmapsto{\varphi} & \varphi(x) = y \\ \varphi_B \downarrow & & \downarrow \varphi_{B'} \\ (x_1, \ldots, x_n) & \xmapsto{?} & (y_1, \ldots, y_m) \end{array}$$

wobei zunächst offen bleibt, wie man $(y_1, \ldots, y_m)$ direkt aus $(x_1, \ldots, x_n)$ berechnen kann. Bezüglich der Basis B hat x die Darstellung

$$x = x_1 v_1 + \ldots + x_n v_n,$$

und daraus ergibt sich für $y = \varphi(x)$:

$$\begin{aligned} y &= \varphi(x) = x_1\varphi(v_1) + \ldots + x_n\varphi(v_n) \\ &= \sum_{j=1}^{n} x_j\varphi(v_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}w_i \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j \right) w_i = \sum_{i=1}^{m} y_i w_i \,. \end{aligned}$$

Also gilt $y_i = \sum_{j=1}^{n} a_{ij}x_j$ und in Matrixschreibweise

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + & \ldots & +a_{1n}x_n \\ & \vdots & \\ a_{m1}x_1 + & \ldots & +a_{mn}x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} .$$

In den Koordinaten stellt sich eine lineare Abbildung also als Multiplikation mit einer Matrix dar:

$$\boxed{A_\varphi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}} \quad \leftarrow \text{Koordinatenvektor von } \varphi(x) \text{ bezüglich } B'$$

$\uparrow$

Koordinatenvektor von x bezüglich B

Das Produkt von A_φ und $(x_1, \ldots, x_n)^t$ ist das Matrizenprodukt gemäß Definition 2.5.

Beispiel. Es sei $V = \mathbb{R}^2$ mit der Basis $B = \{v_1, v_2\}$ wobei $v_1 = (1,1)$, $v_2 = (1,-1)$, und für $W = \mathbb{R}^3$ betrachten wir die Basis $B' = \{w_1, w_2, w_3\}$ mit $w_1 = (1,1,0), w_2 = (1,0,1)$ und $w_3 = (0,1,1)$. Die Bilder der Elemente von B unter der linearen Abbildung φ seien

$$\varphi((1,1)) = (2,2,2) \quad \text{und} \quad \varphi((1,-1)) = (0,1,-1).$$

Wir berechnen A_φ bezüglich B und B'. Es gilt

$$\begin{aligned} \varphi((1,1)) &= w_1 + w_2 + w_3 \\ \varphi((1,-1)) &= w_1 - w_2, \end{aligned}$$

also folgt

$$A_\varphi = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{pmatrix} .$$

Der Vektor $x = (2,0) = v_1 + v_2$ hat bezüglich B den Koordinatenvektor $(x_1, x_2) = (1,1)$. Wegen

$$A_\varphi \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

ist $(2,0,1)$ der Koordinatenvektor von $\varphi((2,0))$ bezüglich B', und $\varphi(x)$ ergibt sich durch

$$\varphi((2,0)) = 2w_1 + w_3 = (2,3,1)\,.$$

Anwendung. Es sei $\varphi : V \longrightarrow W$ linear, wobei V und W endliche Dimension haben, $\dim_K V = n$. Dann gilt

$$\begin{aligned} \mathrm{Rg}\varphi &= \dim_K[\varphi(v_1),\ldots,\varphi(v_n)] \\ &= \dim_K[\varphi_{B'}\varphi(v_1),\ldots,\varphi_{B'}\varphi(v_n)]\,. \end{aligned}$$

Da $\varphi_{B'}\varphi(v_1),\ldots,\varphi_{B'}\varphi(v_n)$ die Spaltenvektoren von A_φ sind, ist $\mathrm{Rg}\varphi$ der Spaltenrang von A_φ. Zusammen mit Satz 2.5 aus Kapitel 3 ergibt sich daraus

$$\begin{aligned} \mathrm{Def}\varphi = \dim_K \mathrm{Kern}\varphi &= \dim_K\{v \in V \mid \varphi(v) = \mathcal{O}\} \\ &= \dim_K\{x \in K^n \mid A_\varphi x^t = \mathcal{O}^t\} \\ &= n - \text{ Zeilenrang von } A_\varphi\,. \end{aligned}$$

Wegen Satz 1.10 können wir nun schließen

$$n = \mathrm{Rg}\varphi + \mathrm{Def}\varphi = \text{Spaltenrang von} A_\varphi + (n - \text{ Zeilenrang von } A_\varphi)\,.$$

Damit folgt

Korollar 2.6 *Für jede Matrix $A \in K_{m,n}$ stimmen Spaltenrang und Zeilenrang überein.*

Korollar 2.7 *Mit den Bezeichnungen wie oben gilt* $\mathrm{Rg}\varphi = \mathrm{Rg}A_\varphi$.

Satz 2.8 *Sind V und W zwei n-dimensionale K-Vektorräume und ist $\varphi : V \longrightarrow W$ linear, dann gilt*

i) *φ ist genau dann injektiv, wenn φ surjektiv ist.*

ii) *φ ist genau dann ein Isomorphismus, wenn es ein $\psi \in \mathrm{Hom}(W,V)$ mit $\psi \circ \varphi = \mathrm{id}_V$ gibt.*

iii) *φ ist genau dann ein Isomorphismus, wenn es ein $\psi \in \mathrm{Hom}(W,V)$ mit $\varphi \circ \psi = \mathrm{id}_W$ gibt.*

Beweis

i) Wegen Satz 1.8 ist φ genau dann injektiv, wenn $\mathrm{Kern}\varphi = \{\mathcal{O}\}$ und damit $\mathrm{Def}\varphi = 0$ gilt. Nach Satz 1.10 ist das wiederum äquivalent zu $\mathrm{Rg}\varphi = n$, und mit Satz 1.5 ist dieses genau dann der Fall, wenn φ surjektiv ist.

ii) $''\Rightarrow''$: Ist φ ein Isomorphismus, dann ist wegen Bemerkung 4 nach Definition 1.1 auch φ^{-1} ein Isomorphismus, und es gilt $\varphi^{-1} \circ \varphi = \mathrm{id}_V$.
$''\Leftarrow''$: Wegen $\psi \circ \varphi = \mathrm{id}_V$ ist $\psi \circ \varphi$ insbesondere injektiv. Damit ist auch φ injektiv und wegen i) ein Isomorphismus.

iii) $'' \Rightarrow''$: Für φ^{-1} gilt $\varphi \circ \varphi^{-1} = \mathrm{id}_W$.
$'' \Leftarrow''$: Wegen $\varphi \circ \psi = \mathrm{id}_W$ folgt mit ii), daß ψ ein Isomorphismus ist. Damit ist auch $\varphi = \psi^{-1}$ ein Isomorphismus.

□

Definition 2.9 *Die Matrix*

$$E_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & \dots & 0 & 1 \end{pmatrix} \in K_{n,n}$$

heißt (n-reihige) Einheitsmatrix. Eine Matrix $A \in K_{n,n}$ heißt invertierbar, wenn eine Matrix $B \in K_{n,n}$ existiert mit $A \cdot B = B \cdot A = E_n$.

Bemerkung.

1. Ist V ein n-dimensionaler K-Vektorraum mit der Basis B und $\varphi : V \longrightarrow V$ eine lineare Abbildung mit der Matrixdarstellung A_φ bezüglich B, dann gilt $\varphi = \mathrm{id}_V$ genau dann, wenn $A_\varphi = E_n$.
2. Ist W ebenfalls ein n-dimensionaler K-Vektorraum und B' eine Basis von W sowie $\varphi : V \longrightarrow W$ linear mit der Matrixdarstellung A_φ bezüglich B und B', so ist φ genau dann ein Isomorphismus, wenn A_φ invertierbar ist.

Es gilt nun Satz 2.8 in folgender Matrixdarstellung:

Satz 2.10 *Für $A \in K_{n,n}$ sind die folgenden Aussagen äquivalent.*

i) *A ist invertierbar.*

ii) $\mathrm{Rg} A = n$.

iii) *Es existiert eine Matrix $B \in K_{n,n}$ mit $B \cdot A = E_n$.*

iv) *Es existiert eine Matrix $B \in K_{n,n}$ mit $A \cdot B = E_n$.*

Definition 2.11 *Ist $n \in \mathbb{N}$ und K ein Körper, so heißt*

$$\mathrm{GL}(n; K) := \{A \in K_{n,n} \mid A \textit{ ist invertierbar}\}$$

allgemeine lineare Gruppe.

Bemerkung. $\mathrm{GL}(n; K)$ ist eine Gruppe bezüglich des Matrizenproduktes.

Satz 2.12 *Für $A \in K_{n,n}$ sind die folgenden Aussagen äquivalent:*

i) Für jedes $b \in K^n$ ist $Ax^t = b^t$ eindeutig lösbar.

ii) Für jedes $b \in K^n$ ist $Ax^t = b^t$ lösbar.

iii) $Ax^t = \mathcal{O}^t$ hat nur die triviale Lösung.

iv) Es gibt ein $b \in K^n$, so daß $Ax^t = b^t$ eindeutig lösbar ist.

v) A ist invertierbar.

Beweis. i)⇒iii): offensichtlich.
iii)⇒iv): Wähle $b = \mathcal{O}$.
iv)⇒v): Satz 2.5, Kapitel 3.
v)⇒ii): Da A invertierbar ist, folgt $x^t = A^{-1}b^t$.
ii)⇒i): A sei die Matrixdarstellung der linearen Abbildung $\varphi : K^n \longrightarrow K^n$ bezüglich einer geeigneten Basis von K^n. Dann ist ii) äquivalent dazu, daß φ surjektiv ist, und i) dazu, daß φ bijektiv ist. Wegen Satz 2.8 folgt die Behauptung.

□

Die Darstellung linearer Abbildungen durch Matrizen soll nun im Zusammenhang mit Koordinatentransformationen und Basiswechsel untersucht werden. Dazu seien im K-Vektorraum V die Basen $B = \{v_1, \ldots, v_n\}$ und $B' = \{v_1', \ldots, v_n'\}$ mit den zugehörigen Koordinatensystemen φ_B und $\varphi_B{}'$ gegeben:

$$\begin{tikzcd} & V \arrow[dl, "\varphi_B"'] \arrow[dr, "\varphi_{B'}"] & \\ K^n \arrow[rr] & & K^n \end{tikzcd}$$

Dann ist $\varphi_{B'} \circ \varphi_B^{-1} : K^n \longrightarrow K^n$ ein Isomorphismus, und $T_{B,B'}$ bezeichne die zugehörige Matrixdarstellung bezüglich der kanonischen Basen. Somit stehen in der j-ten Spalte von $T_{B,B'}$ die Koordinaten von $\varphi_{B'} \circ \varphi_B^{-1}(e_j)$ bezüglich $\{e_1, \ldots, e_n\}$, also die Koordinaten von $\varphi_{B'}(v_j)$ bezüglich $\{e_1, \ldots, e_n\}$:

$$\varphi_{B'}(v_j) = t_{1j}e_1 + \ldots + t_{nj}e_n.$$

Wendet man auf diese Gleichung $\varphi_{B'}^{-1}$ an, so sieht man, daß in der j-ten Spalte von $T_{B,B'}$ die Koordinaten von v_j bezüglich $\{v_1', \ldots, v_n'\}$ stehen. Mit T statt $T_{B,B'}$ gilt dann

$$\boxed{T = (t_{ij}) \text{ und } v_j = t_{1j}{v_1}' + \ldots + t_{nj}{v_n}' .}$$

Ist also $(x_1, \ldots, x_n)$ der Koordinatenvektor von x bezüglich B und $(x_1', \ldots, x_n')$ der Koordinatenvektor von x bezüglich B', so folgt

$$\boxed{T \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} {x_1}' \\ \vdots \\ {x_n}' \end{pmatrix} .}$$

T heißt Transformationsmatrix des Basiswechsels $B \longrightarrow B'$, und T ist invertierbar. Offenbar ist T^{-1} die Transformationsmatrix des Basiswechsels $B' \longrightarrow B$.

Beispiel. Wir betrachten für $V = \mathbb{R}^2$ die Basen B und B' mit $B = \{(3,-1), (-1,-3)\}$ und $B' = \{(1,1), (1,-1)\}$. Wegen

$$\begin{array}{rcl} (3,-1) & = & (1,1) + 2 \cdot (1,-1) \quad \text{und} \\ (-1,-3) & = & (-2) \cdot (1,1) + (1,-1) \end{array}$$

folgt

$$T = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}.$$

Der Vektor $x = (7,-9)$ besitzt die beiden Darstellungen

$$\begin{array}{rcl} x & = & 3 \cdot (3,-1) + 2 \cdot (-1,-3) \\ & = & (-1) \cdot (1,1) + 8 \cdot (1,-1), \end{array}$$

d.h., $(3,2)$ ist der Koordinatenvektor von x bezüglich B und $(-1,8)$ der Koordinatenvektor von x bezüglich B'. Es gilt

$$T \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 8 \end{pmatrix}.$$

Sind nun V und W zwei K-Vektorräume mit den Dimensionen n bzw. m und ist $\varphi : V \longrightarrow W$ linear, so betrachten wir die beiden Matrixdarstellungen A_φ und A'_φ von φ bezüglich der Basen B_V und B_W bzw. B'_V und B'_W:

$$\begin{array}{ccccc} & K^n & \xrightarrow{A'_\varphi} & K^m & \\ B'_V & \uparrow & & \uparrow & B'_W \\ & V & \xrightarrow{\varphi} & W & \\ B_V & \downarrow & & \downarrow & B_W \\ & K^n & \xrightarrow{A_\varphi} & K^m & \end{array}.$$

Sind weiterhin x_i die Koordinaten von $x \in V$ bezüglich B_V und x'_i diejenigen bezüglich B'_V und entsprechend y_i die Koordinaten von $\varphi(x) \in W$ bezüglich B_W und y'_i die Koordinaten bezüglich B'_W, dann gilt

$$A_\varphi \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad \text{und} \quad A'_\varphi \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \begin{pmatrix} y'_1 \\ \vdots \\ y'_m \end{pmatrix}.$$

Bezeichnet T die Transformationsmatrix des Basiswechsels $B_V \longrightarrow B'_V$ und S die Transformationsmatrix des Basiswechsels $B_W \longrightarrow B'_W$, dann folgt

$$\boxed{A_\varphi = S^{-1} {A'}_\varphi T\,.}$$

Definition 2.13 *Sind $A, B \in K_{m,n}$, dann heißen A und B äquivalent (geschrieben $A \sim B$), wenn $S \in \mathrm{GL}(m; K)$ und $T \in \mathrm{GL}(n; K)$ existieren, so daß*

$$A = S^{-1}BT.$$

Bemerkung. Es seien $A, B \in K_{m,n}$.

1. $A \sim B$ gilt genau dann, wenn eine lineare Abbildung $\varphi : K^n \longrightarrow K^m$ existiert, die bezüglich geeigneter Basen die Matrixdarstellungen A und B hat.

2. $A \sim B$ gilt genau dann, wenn $\mathrm{Rg}A = \mathrm{Rg}B$.

3. Determinanten

Wir betrachten die Matrix $A = (a_{ij}) \in K_{n,n}$, also

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Dann heißt $a_i := (a_{i1}, \dots, a_{in})$ der i-te Zeilenvektor von A, und man kann A formal durch

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

darstellen.

Definition 3.1 *Für $n \in \mathbb{N}$ heißt eine Abbildung*

$$\det : K_{n,n} \longrightarrow K\,, \quad (a_{ij}) \longmapsto |a_{ij}|$$

Determinante, wenn für alle $(a_{ij}) \in K_{n,n}$, $k \in K$ gilt

i) $\det \begin{pmatrix} \vdots \\ a'_j + a''_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a'_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a''_j \\ \vdots \end{pmatrix}$ *für alle $j \in \{1, \dots, n\}$.*

ii) $\det \begin{pmatrix} \vdots \\ k \cdot a_j \\ \vdots \end{pmatrix} = k \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \end{pmatrix}$ *für alle $j \in \{1, \dots, n\}$.*

iii) Hat A zwei gleiche Zeilen, so ist $\det A = 0$.

iv) $\det E_n = 1$.

Bemerkung. Die Eigenschaften i) und ii) bedeuten, daß die Abbildung det linear in jeder Zeile ist.

Einfache Eigenschaften.

1. Aus ii) folgt $\det(k \cdot A) = k^n \det A$.
2. Aus $a_i = (0, \dots, 0)$ folgt $\det A = 0$, denn mit der Eigenschaft i) gilt
$$\det \begin{pmatrix} & \vdots & \\ 0 & \dots & 0 \\ & \vdots & \end{pmatrix} = \det \begin{pmatrix} & \vdots & \\ 0 & \dots & 0 \\ & \vdots & \end{pmatrix} + \det \begin{pmatrix} & \vdots & \\ 0 & \dots & 0 \\ & \vdots & \end{pmatrix}.$$
3. Das Vertauschen zweier Zeilen ändert das Vorzeichen, das heißt
$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix},$$
denn mit den Eigenschaften i) - iii) aus Definition 3.1 gilt
$$\begin{aligned} \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} &= \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \end{pmatrix} \\ &= \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \end{pmatrix} = 0. \end{aligned}$$
4. Es gilt
$$\det \begin{pmatrix} \vdots \\ a_i + k a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix} + k \det \begin{pmatrix} \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \end{pmatrix};$$

das Addieren des Vielfachen einer Zeile zu einer anderen Zeile ändert den Wert der Determinante also nicht.

5. Sei $\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ eine Permutation (siehe Kapitel 2.1) und $\{e_1, \ldots, e_n\}$ die kanonische Basis des K^n. Dann gilt

$$\det \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = \mathrm{sgn}\sigma\,.$$

Um das einzusehen, schreiben wir $\sigma = \tau_1 \ldots \tau_\nu$ als Produkt von Transpositionen und verwenden $\mathrm{sgn}\sigma = (-1)^\nu$. Aus Eigenschaft 3 (siehe oben) folgt dann die Behauptung.

Satz 3.2 *Ist die Abbildung* $\det : K_{n,n} \longrightarrow K$ *eine Determinante, so gilt*

$$\det(a_{ij}) = \sum_{\sigma \in \mathbf{S_n}} \mathrm{sgn}\sigma \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}\,. \tag{18}$$

Beweis. Mit der Darstellung

$$a_i = (a_{i1}, \ldots, a_{in}) = a_{i1}e_1 + \ldots + a_{in}e_n$$

folgt aus den Eigenschaften i) und ii) der Determinante (Definition 3.1)

$$\det(a_{ij}) = \det \begin{pmatrix} a_{11}e_1 + \ldots + a_{1n}e_n \\ \vdots \\ a_{n1}e_1 + \ldots + a_{nn}e_n \end{pmatrix} = \sum_{k_1,\ldots,k_n=1}^{n} a_{1k_1} \cdot \ldots \cdot a_{nk_n} \det \begin{pmatrix} e_{k_1} \\ \vdots \\ e_{k_n} \end{pmatrix}.$$

Mit Eigenschaft iii) der Determinante läßt sich die Summe reduzieren zu

$$\det(a_{ij}) = \sum_{\substack{k_1,\ldots,k_n=1 \\ \text{verschieden}}}^{n} a_{1k_1} \cdot \ldots \cdot a_{nk_n} \det \begin{pmatrix} e_{k_1} \\ \vdots \\ e_{k_n} \end{pmatrix} = \sum_{\sigma \in \mathbf{S_n}} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \det \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix},$$

und mit der Eigenschaft 5 folgt die Behauptung des Satzes.

□

Bemerkung. Gleichung (18) bezeichnet man als Leibnizformel für Determinanten. Insbesondere besagt Satz 3.2, daß es für jedes $n \in \mathbb{N}$ höchstens eine Determinante gibt.

Korollar 3.3 *Hat eine Matrix* $A \in K_{n,n}$ *die spezielle Gestalt*

$$A = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ * & a_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ * & & * & a_{nn} \end{pmatrix},$$

d.h., ist A *eine untere Dreiecksmatrix, so gilt* $\det A = a_{11} \cdot \ldots \cdot a_{nn}$.

Definition 3.4 *Ist* $A = (a_{ij}) \in K_{m,n}$, *so heißt die Matrix* $A^t \in K_{n,m}$, $A^t = (b_{ij})$ *mit* $b_{ij} = a_{ji}$ *die zu* A *transponierte Matrix.*

Bemerkung.

1. $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \Longrightarrow A^t = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}.$

2. $(A^t)^t = A$.

3. Die Zeilen (bzw. Spalten) von A entsprechen den Spalten (bzw. Zeilen) von A^t.

Korollar 3.5 *Für jede Matrix* $A \in K_{n,n}$ *gilt* $\det A = \det A^t$.

Beweis. Mit Satz 3.2 und $\text{sgn}\sigma = \text{sgn}\sigma^{-1}$ folgt

$$\begin{aligned} \det A &= \sum_{\sigma \in \mathbf{S_n}} \text{sgn}\sigma \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathbf{S_n}} \text{sgn}\sigma \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma \in \mathbf{S_n}} \text{sgn}\sigma^{-1} \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\ &= \sum_{\tau \in \mathbf{S_n}} \text{sgn}\tau \cdot a_{\tau(1)1} \cdot \ldots \cdot a_{\tau(n)n} = \det A^t\,. \end{aligned}$$

□

Bemerkung. Die Funktion det ist damit multilinear in den Spalten und die Eigenschaften 2-5 nach Definition 3.1 gelten entsprechend für die Spalten von A.

Korollar 3.6 *Hat eine Matrix* $A \in K_{n,n}$ *die spezielle Gestalt*

$$A = \begin{pmatrix} a_{11} & * & \dots & * \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & * \\ 0 & & 0 & a_{nn} \end{pmatrix},$$

d.h., ist A *eine obere Dreiecksmatrix, so gilt* $\det A = a_{11} \cdot \ldots \cdot a_{nn}$.

Korollar 3.7 *Für* $A \in K_{n,n}$ *ist genau dann* $\text{Rg}A = n$, *wenn* $\det A \neq 0$ *gilt.*

Beweis. Ist eine Matrix $B \in K_{n,n}$ aus A durch eine elementare Umformung entstanden, so gilt genau dann $\det A = 0$, wenn $\det B = 0$. Da dann weiterhin $\mathrm{Rg}A = \mathrm{Rg}B$ folgt, können wir also annehmen, daß A in Normalform gegeben ist:

$$\begin{pmatrix} \overbrace{\begin{matrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1 \end{matrix}}^{\mathrm{Rg}A} & \overbrace{\begin{matrix} \quad & * & \quad \end{matrix}}^{n-\mathrm{Rg}A} \\ & \\ 0 & 0 \end{pmatrix}.$$

Somit ist $\mathrm{Rg}A = n$ genau dann, wenn $\det A \neq 0$. □

Bemerkung. Eine Matrix $A \in K_{n,n}$ heißt regulär, wenn $\mathrm{Rg}A = n$. Mit Korollar 3.7 ist das äquivalent dazu, daß $\det A \neq 0$. Ist A nicht regulär, so heißt A singulär.

Satz 3.8 *Für alle $A, B \in K_{n,n}$ gilt* $\det A \cdot B = \det A \cdot \det B$.

Beweis. Mit den Bezeichnungen $A = (a_{ij})$, $B = (b_{ij})$ sowie $A \cdot B =: C = (c_{ij})$ gilt $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ nach Definition 2.5. Ist weiterhin $b_i = (b_{i1}, \dots, b_{in})$ der i-te Zeilenvektor von B, dann folgt

$$\begin{aligned} (c_{i1}, \dots, c_{in}) &= (\sum_{k=1}^{n} a_{ik} b_{k1}, \dots, \sum_{k=1}^{n} a_{ik} b_{kn}) \\ &= \sum_{k=1}^{n} (a_{ik} b_{k1}, \dots, a_{ik} b_{kn}) = \sum_{k=1}^{n} a_{ik} b_k \,. \end{aligned}$$

Damit ergibt sich für die Determinante auf Grund der Multilinearität

$$\det C = \sum_{i_1, \dots, i_n = 1}^{n} a_{1 i_1} \cdot \ldots \cdot a_{n i_n} \det \begin{pmatrix} b_{i_1} \\ \vdots \\ b_{i_n} \end{pmatrix}$$

und wegen iii) aus Definition 3.1

$$\det C = \sum_{\sigma \in \mathbf{S_n}} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \det \begin{pmatrix} b_{\sigma(1)} \\ \vdots \\ b_{\sigma(n)} \end{pmatrix}.$$

Eigenschaft 3 liefert schließlich

$$\det C = \sum_{\sigma \in \mathbf{S_n}} \mathrm{sgn}\sigma \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \det B = \det A \cdot \det B \,.$$

□

Korollar 3.9 *Ist eine Matrix $A \in K_{n,n}$ regulär, so gilt* $\det A^{-1} = (\det A)^{-1}$.

Satz 3.10 *Zu jedem $n \in \mathbb{N}$ existiert genau eine Determinante* $\det : K_{n,n} \longrightarrow K$.

Beweis. Die Eindeutigkeit der Determinante folgt aus Satz 3.2. Die Existenz soll durch vollständige Induktion nach n gezeigt werden.
Für $n = 1$ ist $\det : K_{1,1} \longrightarrow K,\ (k) \longmapsto k$ offensichtlich eine Determinante.
Als Induktionsvoraussetzung nehmen wir an, daß es bereits die Determinante für $n-1$ gibt, wobei $n > 1$ ist. Für den Induktionsschluß definieren wir zunächst die Adjunkte

$$A_{ij} := (-1)^{i+j} \begin{array}{c} j-\text{te Spalte} \\ \left|\begin{array}{cc|cc} a_{11} & \dots & \dots & a_{1n} \\ \vdots & & & \vdots \\ \hline \vdots & & & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{array}\right| \end{array} \ i - \text{te Zeile}\ ,$$

die bis auf das Vorzeichen die Determinante der Matrix bezeichnet, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Da die entstandene Matrix eine $(n-1, n-1)$-Matrix ist, existiert ihre Determinante nach Voraussetzung. Sei nun $j \in \{1, \dots, n\}$ fest gewählt; dann setzen wir

$$\det(a_{ij}) := \sum_{i=1}^{n} a_{ij} A_{ij}\,. \tag{19}$$

Es bleibt zu prüfen, daß die so definierte Abbildung die Eigenschaften der Determinante hat:

i) Die k-te Zeile von $A = (a_{ij})$ sei durch $a_k = a'_k + a''_k$ gegeben. Wir betrachten nun die beiden (n,n)-Matrizen $A' = (a'_{ij})$ und $A'' = (a''_{ij})$, wobei A' und A'' bis auf die k-ten Zeilen mit A übereinstimmen und a'_k die k-te Zeile von A' sowie a''_k die k-te Zeile von A'' ist. Zu zeigen bleibt $\det A = \det A' + \det A''$. Nach (19) gilt

$$\det(a_{ij}) = \sum_{\substack{i=1\\ i\neq k}}^{n} a_{ij}A_{ij} + a_{kj}A_{kj}\,.$$

Für $k \neq i$ ist nach Induktionsvoraussetzung $A_{ij} = A'_{ij} + A''_{ij}$, und da A, A' und A'' bis auf die k-ten Zeilen übereinstimmen, folgt $A_{kj} = A'_{kj} = A''_{kj}$ sowie $a_{ij} = a'_{ij} = a''_{ij}$ für $k \neq i$. Wegen $a_{kj} = a'_{kj} + a''_{kj}$ ergibt sich schließlich

$$\det(a_{ij}) = \sum_{\substack{i=1\\ i\neq k}}^{n} (a'_{ij}A'_{ij} + a''_{ij}A''_{ij}) + a'_{kj}A'_{kj} + a''_{kj}A''_{kj} = \det A' + \det A''\,.$$

ii) Analog zu i).

iii) Sei $a_\nu = a_\mu$ mit $\nu \neq \mu$ und wir nehmen desweiteren zunächst an, daß a_ν und a_μ benachbart sind mit $\mu = \nu + 1$. Dann gilt

$$\begin{aligned} \det(a_{ij}) &= \sum_{\substack{i=1\\ i\neq\nu,\nu+1}}^{n} a_{ij}A_{ij} + a_{\nu j}A_{\nu j} + a_{(\nu+1)j}A_{(\nu+1)j} \\ &= a_{\nu j}A_{\nu j} + a_{\nu j}(-A_{\nu j}) = 0\,, \end{aligned}$$

da nach Induktionsvoraussetzung $A_{ij} = 0$ für $i \neq \nu, \nu + 1$ gilt. Wie im Beweis zu Eigenschaft 3 der Determinante folgt somit, daß das Vertauschen benachbarter Zeilen nur das Vorzeichen ändert. Daraus ergibt sich, daß das Vertauschen zweier beliebiger Zeilen auch nur das Vorzeichen ändert, und es folgt insgesamt $\det A = 0$ bei Gleichheit zweier beliebiger Zeilen.

iv) Für $(a_{ij}) = E_n$ gilt

$$\det(a_{ij}) = \sum_{i=1}^{n} a_{ij} A_{ij} = A_{jj} = (-1)^{j+j} \det E_{n-1} = 1 \, .$$

$\square$

Bemerkung. Die Darstellung (19) für die Determinante einer Matrix $A = (a_{ij})$ heißt Laplacescher Entwicklungssatz. Wegen der Eindeutigkeit von det ist die Darstellung (19) unabhängig von der Spalte j, nach der ″entwickelt″ wird. Wegen Korollar 3.5 kann man auch nach einer Zeile entwickeln, und für die Entwicklung nach der i-ten Zeile ergibt sich

$$\det(a_{ij}) = \sum_{j=1}^{n} a_{ij} A_{ij} \, .$$

Beispiel.

1. Für $n = 2$ ist die Determinante einer Matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ nach Satz 3.2 gegeben durch
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \, .$$

2. Für $n = 3$ gilt
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} \, .$$
Dieses ist die Regel von Sarrus; zum besseren Merken kann man die ersten beiden Spalten der Matrix noch einmal rechts neben die dritte Spalte schreiben und dann immer schräg nach oben und schräg nach unten multiplizieren, wie unten angedeutet. Dabei bedeuten alle nach unten gerichteten Pfeile ein positives und alle nach oben gerichteten Pfeile ein negatives Vorzeichen:
$$\begin{array}{ccccccccc}
 & & & & - & & - & & - \\
a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\
 & \searrow & & \searrow\!\!\!\!\!\!\nearrow & & \searrow\!\!\!\!\!\!\nearrow & & \nearrow & \\
a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\
 & \nearrow & & \searrow\!\!\!\!\!\!\nearrow & & \searrow\!\!\!\!\!\!\nearrow & & \searrow & \\
a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32} \\
 & & & & + & & + & & +
\end{array}$$
(diese Regel gilt nur für n=3).

3. Durch elementare Umformungen (vgl. Eigenschaften 1,3 und 4 nach Definition 3.1) bringt man die Matrix auf Dreiecksgestalt und verwendet dann Korollar 3.6:

$$\begin{vmatrix} 1 & 2 & 1 & 1 \\ -1 & -3 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 & 1 \\ 0 & -1 & 3 & 2 \\ 0 & -2 & 0 & 1 \\ 0 & 3 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 & 1 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & -6 & -3 \\ 0 & 0 & 12 & 8 \end{vmatrix}$$
$$= \begin{vmatrix} 1 & 2 & 1 & 1 \\ 0 & -1 & 3 & 2 \\ 0 & 0 & -6 & -3 \\ 0 & 0 & 0 & 2 \end{vmatrix} = 1 \cdot (-1) \cdot (-6) \cdot 2 = 12.$$

4. Beim Entwickeln nach einzelnen Zeilen oder Spalten (Laplacescher Entwicklungssatz) kann man zur Vereinfachung das Vorzeichenschema

$$\begin{matrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & & & \end{matrix}$$

benutzen. Im folgenden Beispiel entwickeln wir nach der 2. Spalte und verwenden somit die Vorzeichenfolge $-, +, -, +$:

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 1 & 1 \\ -1 & -3 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & 2 & 1 \end{vmatrix} &= -2 \cdot \begin{vmatrix} -1 & 2 & 1 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{vmatrix} + (-3) \cdot \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{vmatrix} - 0 + \begin{vmatrix} 1 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{vmatrix} \\ &= 0 + (-3)(-1) \cdot \begin{vmatrix} 1 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{vmatrix} \\ &= 4 \cdot (4 + 1 - 1 - (2 + 1 - 2)) = 4 \cdot 3 = 12\,. \end{aligned}$$

Satz 3.11 *Für jede Matrix $A = (a_{ij}) \in K_{n,n}$ gilt*

$$\sum_{k=1}^{n} a_{ik} A_{jk} = \left\{ \begin{array}{ccc} \det A & & j = i. \\ & \textit{für} & \\ 0 & & j \neq i\,. \end{array} \right.$$

Beweis. Für $j = i$ ist der Satz gerade der Laplacesche Entwicklungssatz (Entwicklung nach der i-ten Zeile). Es bleibt also der Fall $j \neq i$ zu untersuchen. Sei $B \in K_{n,n}$ mit den Zeilenvektoren $b_1, \ldots, b_n$, wobei $b_\nu = a_\nu$ für $\nu \neq j$ und $b_j = a_i$, also

$$B = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} \begin{array}{l} \\ \leftarrow \quad i - \text{te Zeile} \\ \\ \leftarrow \quad j - \text{te Zeile} \\ \\ \end{array}$$

Dann folgt $\det B = 0$ mit iii) aus Definition 3.1. Andererseits kann man die Determinante von B durch Entwickeln nach der j-ten Zeile berechnen, also

$$0 = \det B = \sum_{k=1}^{n} b_{jk} B_{jk} = \sum_{k=1}^{n} a_{ik} A_{jk} \, .$$

□

Definition 3.12 *Für $A = (a_{ij}) \in K_{n,n}$ ist $\tilde{A} := (\tilde{a}_{ij}) \in K_{n,n}$ mit $\tilde{a}_{ij} = A_{ij}$.*

Satz 3.13 *Ist $A = (a_{ij}) \in K_{n,n}$, so gilt $A \cdot \tilde{A}^t = \tilde{A}^t \cdot A = (\det A) \cdot E_n$.*

Beweis. Nach Definition 3.12 und wegen Satz 3.11 ist

$$A \cdot \tilde{A}^t = (\sum_{k=1}^{n} a_{ik} \tilde{a}_{jk}) = (\sum_{k=1}^{n} a_{ik} A_{jk}) = (\det A) \cdot E_n.$$

□

Korollar 3.14 *Ist $A \in K_{n,n}$ invertierbar, so gilt $A^{-1} = \frac{1}{\det A} \tilde{A}^t$.*

Beispiel. Für $A = \begin{pmatrix} 1 & 2 & 1 \\ -2 & 1 & -1 \\ 1 & 1 & 2 \end{pmatrix}$ gilt $\det A = 2 - 2 - 2 - (1 - 1 - 8) = 6$ und

$$\tilde{A} = \begin{pmatrix} 3 & 3 & -3 \\ -3 & 1 & 1 \\ -3 & -1 & 5 \end{pmatrix}, \quad \text{also} \quad A^{-1} = \frac{1}{6} \begin{pmatrix} 3 & -3 & -3 \\ 3 & 1 & -1 \\ -3 & 1 & 5 \end{pmatrix}.$$

Korollar 3.15 *Ist $A = (a_{ij}) \in \mathbb{Q}_{n,n}$ regulär mit $a_{ij} \in \mathbb{Z}$, dann sind die Komponenten von A^{-1} genau dann auch aus $\mathbb{Z}$, wenn* $\det A = \pm 1$.

Beweis. Gilt $\det A = \pm 1$, so sind die Komponenten von A^{-1} ganzzahlig wegen Korollar 3.14. Andererseits folgt aus $1 = \det E_n = \det AA^{-1} = \det A \det A^{-1}$ und $\det A, \det A^{-1} \in \mathbb{Z}$, daß $\det A^{-1} = \pm 1$.

□

Bemerkung. Definition 3.1 kann sinngemäß auch zur Einführung von Determinanten über kommutativen Ringen mit 1 verwendet werden. Viele der bewiesenen Sätze gelten dann auch für Matrizen über kommutativen Ringen mit 1, so zum Beispiel Satz 3.13.

Anwendung. Ist V ein n-dimensionaler $\mathbb{Q}-$ bzw. $\mathbb{R}-$Vektorraum mit der Basis $\{v_1, \ldots, v_n\}$, dann heißt

$$\mathfrak{h} := \mathbb{Z} \cdot v_1 + \ldots + \mathbb{Z} \cdot v_n$$

n-dimensionales Gitter. Es soll nun untersucht werden, unter welchen Bedingungen für eine andere Basis $\{w_1, \ldots, w_n\} \subseteq \mathfrak{h}$ von V

$$\mathbb{Z} \cdot w_1 + \ldots + \mathbb{Z} \cdot w_n = \mathfrak{h}$$

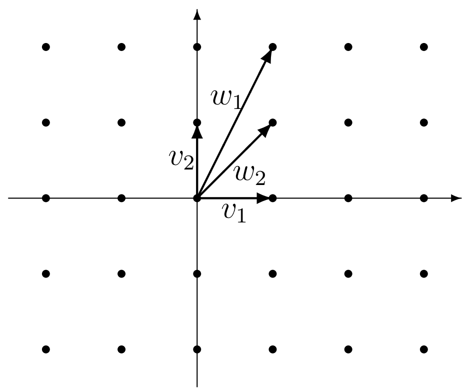
gilt, wann also $\{w_1, \ldots, w_n\}$ das ganze Gitter erzeugt. Wegen $\mathbb{Z} \cdot w_1 + \ldots + \mathbb{Z} \cdot w_n \subseteq \mathfrak{h}$ hat jedes w_j die Darstellung

$$w_j = a_{1j}v_1 + \ldots + a_{nj}v_n\,,$$

wobei $a_{ij} \in \mathbb{Z}$. Damit ist (a_{ij}) die Transformationsmatrix des Basiswechsels $\{w_1, \ldots, w_n\} \to \{v_1, \ldots, v_n\}$ und $(b_{ij}) = (a_{ij})^{-1}$ die Transformationsmatrix des umgekehrten Wechsels. Wegen $v_j = b_{1j}w_1 + \ldots + b_{nj}w_n$ gilt also

$$\mathbb{Z} \cdot w_1 + \ldots + \mathbb{Z} \cdot w_n = \mathfrak{h} \quad \Longleftrightarrow \quad b_{ij} \in \mathbb{Z} \quad \Longleftrightarrow \quad \det(a_{ij}) = \pm 1.$$

Beispiel. Sei $V = \mathbb{R}^2$ mit der Basis $\{v_1, v_2\}$ und sei durch $w_1 = v_1 + 2v_2$ und $w_2 = v_1 + v_2$ eine weitere Basis gegeben.



Dann ist $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ die zugehörige Transformationsmatrix, und es gilt $\det A = -1$. Also ist die Transformationsmatrix des umgekehrten Wechsels durch $A^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$ gegeben. Es gilt $v_1 = -w_1 + 2w_2$ und $v_2 = w_1 - w_2$.

Cramersche Regel. Ist ein lineares Gleichungssystem mit quadratischer Koeffizientenmatrix eindeutig lösbar, so kann man mit Hilfe der Determinante die Lösung explizit angeben (Cramersche Regel):

Für $A \in K_{n,n}$ und $b \in K^n$ sei das LGS

$$Ax^t = b^t \tag{20}$$

gegeben. Die Matrix A_i entstehe aus A, indem man die i-te Spalte von A durch b^t ersetzt. Ist $\det A \neq 0$ und $(x_1, \ldots, x_n) \in K^n$ die eindeutige Lösung von (20), so gilt

$$x_i = \frac{\det A_i}{\det A}\,, \quad i = 1, \ldots, n. \tag{21}$$

Beweis. Da A invertierbar ist, folgt

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

wobei $A^{-1} = \frac{1}{\det A}\tilde{A}^t$ wegen Korollar 3.14 und

$$\tilde{A}^t = \begin{pmatrix} A_{11} & \dots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \dots & A_{nn} \end{pmatrix}.$$

Damit ist

$$x_i = \frac{1}{\det A} \sum_{k=1}^{n} A_{ki} b_k.$$

Wegen

$$A_i = \begin{pmatrix} a_{11} & \dots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \dots \\ a_{21} & & a_{2(i-1)} & b_2 & a_{2(i+1)} & \\ \vdots & & \vdots & \vdots & \vdots & \\ a_{n1} & \dots & a_{n(i-1)} & b_n & a_{n(i+1)} & \dots \end{pmatrix}$$

folgt auf Grund des Laplaceschen Entwicklungssatzes $\det A_i = \sum_{k=1}^n b_k A_{ki}$, also die Behauptung. □

Beispiel. Es sei das LGS

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 - x_2 + x_3 &= -1 \\ x_1 - x_2 - x_3 &= 0 \end{aligned}$$

gegeben. Die Determinante der Koeffizientenmatrix ist also

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & 0 \end{vmatrix} = 2 \cdot 2 = 4,$$

und mit der Cramerschen Regel ergibt sich

$$\begin{aligned} x_1 &= \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ 0 & -1 & -1 \end{vmatrix} = \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & -1 & -1 \end{vmatrix} = \frac{1}{2}, \\ x_2 &= \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{vmatrix} = \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & -1 \end{vmatrix} = 1, \\ x_3 &= \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 0 \end{vmatrix} = \frac{1}{4} \begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & 0 \\ 1 & -1 & 0 \end{vmatrix} = -\frac{1}{2}. \end{aligned}$$

4. Aufgaben

A 4.1 Es sei $\{v_1, \ldots, v_n\}$ eine Basis des reellen Vektorraums V und $\varphi : V \longrightarrow V$ die durch

$$\varphi(v_1) := v_2,\ \varphi(v_2) := v_3,\ \ldots\ ,\ \varphi(v_{n-1}) := v_n,\ \varphi(v_n) := v_1$$

definierte lineare Abbildung. Für welche Dimensionen n gibt es einen Vektor $v \neq \mathcal{O}$ mit $\varphi(v) = -v$?

A 4.2 Es seien V_1 und V_2 zwei K-Vektorräume sowie $\varphi : V_1 \to V_2$ eine surjektive, lineare Abbildung. Weiterhin sei $\mathcal{U}_2$ die Menge aller Unterräume von V_2 und $\mathcal{U}_1$ die Menge aller Unterräume von V_1, die Kernφ enthalten.
Zeigen Sie, daß dann

$$\Phi : \mathcal{U}_1 \longrightarrow \mathcal{U}_2, \quad U \longmapsto \varphi(U)$$

eine bijektive Abbildung ist.

A 4.3 Es sei V ein K-Vektorraum mit der Basis $\{e_1, e_2, e_3\}$ und $\varphi : V \to V$ die durch

$$\varphi(e_1) := e_1 + e_2 + e_3, \quad \varphi(e_2) := e_1 - 2e_2 + e_3, \quad \varphi(e_3) := e_1 + 3e_2 + e_3$$

definierte lineare Abbildung. Berechnen Sie Rgφ und Defφ.

A 4.4 Es sei $V = \mathbb{R}^3$ und $\{e_1, e_2, e_3\}$ die kanonische Basis des $\mathbb{R}^3$. Für fest gewähltes $a \in \mathbb{R}$ sei $\varphi_a : V \to V$ die durch

$$\varphi_a(e_1) := (1, a^2, a),\ \varphi_a(e_2) := (a, 1, a^2),\ \varphi_a(e_3) := (a^2, a, 1)$$

definierte lineare Abbildung. Berechnen Sie Rgφ_a und Defφ_a.

A 4.5 Es sei $A \in K_{m,n},\ B \in K_{n,k},\ C \in K_{k,l}$. Zeigen Sie

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

A 4.6 Berechnen Sie $|\ \mathrm{GL}(3; \mathbb{Z}_2)\ |$.

A 4.7 (a) Invertieren Sie die reelle Matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 5 & 4 & 3 & 0 \\ -1 & 1 & -2 & 3 \\ 3 & 6 & 1 & -1 \end{pmatrix}.$$

(b) Für welche $a \in \mathbb{R}$ ist die reelle Matrix

$$B = \begin{pmatrix} a & a^2 & a^3 \\ 1 & a & -a \\ 0 & 1 & -1 \end{pmatrix}$$

invertierbar? Berechnen Sie hierfür B^{-1}.

A 4.8 Es sei V ein Vektorraum über dem Körper K und $\mathcal{B} = \{v_1, v_2, v_3\}$ eine Basis von V sowie

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

die Matrixdarstellung von $\varphi \in \mathrm{Hom}(V, V)$ bezüglich der Basis $\mathcal{B}$. Zeigen Sie, daß die Vektoren

$$u_1 = v_1 + v_2, \quad u_2 = 2v_1 + v_2, \quad u_3 = 4v_1 + 2v_2 + v_3$$

eine Basis von V bilden, und berechnen Sie die zugehörige Matrixdarstellung A'_φ von φ.

A 4.9 Für $a, b \in K$ ist die (n, n)-Matrix A gegeben mit

$$A = \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{pmatrix}.$$

Zeigen Sie durch vollständige Induktion $\det A = (a-b)^{n-1}(a + (n-1)b)$. Für welche a, b ist A invertierbar?

A 4.10 Es sei $B \in K_{n,n}, C \in K_{n,m}, D \in K_{m,m}$ sowie

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}.$$

Zeigen Sie $\det A = \det B \cdot \det D$ unter Verwendung von Satz 3.2.

A 4.11 Es seien $x_1, \ldots, x_n \in K$. Zeigen Sie

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{vmatrix} = \prod_{k>j} (x_k - x_j).$$

A 4.12 Zeigen Sie: Sind $a_1, \ldots, a_n \in \mathbb{R}$ paarweise verschieden und $b_1, \ldots, b_n \in \mathbb{R}$ beliebig, so gibt es genau ein reelles Polynom f mit $f(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$ und $f(a_1) = b_1, \ldots, f(a_n) = b_n$.

A 4.13 Geben Sie alle reellen Polynome f mit $f(x) = x^4 + ax^3 + bx^2 + cx + d$ an, für die gilt:
(a) $f(1) = 0, f(2) = 7, f(-1) = 4$.
(b) $f(1) = 1, f(0) = 0, f(2) = 2, f(-1) = -1$.

A 4.14 Gegeben ist die reelle Matrix

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

Zeigen Sie, daß A für jedes $a \in \mathbb{R}$ regulär ist, und berechnen Sie mit Hilfe der Adjunktenformel (Korollar 3.14) die inverse Matrix A^{-1}.

KAPITEL 5

Der Endomorphismenring End(V)

1. Polynome

Ist K ein Körper, so bezeichnet man den formalen Ausdruck

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots = \sum_{n=0}^{\infty} a_n x^n$$

mit $a_n \in K$ und $a_n \neq 0$ für nur endlich viele n als Polynom über K mit den Koeffizienten $a_0, a_1, a_2, \ldots$. Gilt $a_i = 0$ für alle $i = 0, 1, 2, \ldots$, so heißt $f(x)$ Nullpolynom, geschrieben $f(x) = 0$.

Definition 1.1 *Seien $f(x) = \sum_{n=0}^{\infty} a_n x^n$ und $g(x) = \sum_{n=0}^{\infty} b_n x^n$ zwei Polynome.*

i) $f(x) = g(x) :\Longleftrightarrow a_n = b_n$ *für alle* $n \in \mathbb{N}_0$.

ii) $\mathrm{grad} f(x) := -\infty$, *falls* $f(x) = 0$, *und* $\mathrm{grad} f(x) := \max\{n \in \mathbb{N}_0 \mid a_n \neq 0\}$ *sonst.*

iii) Mit $K[x]$ bezeichnet man die Menge aller Polynome über K.

iv)
$$f(x) + g(x) = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n := \sum_{n=0}^{\infty} (a_n + b_n) x^n =: (f+g)(x).$$

v)
$$f(x) \cdot g(x) = \sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} b_n x^n := \sum_{n=0}^{\infty} c_n x^n =: (f \cdot g)(x),$$

wobei $c_n := \sum_{k=0}^{n} a_k b_{n-k}$.

vi) Sei $k \in K$.
$$k \cdot f(x) = k \cdot \sum_{n=0}^{\infty} a_n x^n := \sum_{n=0}^{\infty} (k \cdot a_n) x^n.$$

Bemerkung.

1. $\text{grad} f(x)$ heißt Grad von $f(x)$.
2. Die Multiplikation von Polynomen ist wohldefiniert, da nur endlich viele der Summen $\sum_{k=0}^{n} a_k b_{n-k}$ von 0 verschieden sind.
3. Bezüglich der in Definition 1.1 beschriebenen Verknüpfungen ist $K[x]$ ein kommutativer Ring mit 1 und ein K-Vektorraum mit $\dim_K K[x] = \infty$.
4. Man schreibt oft die Summanden $0x^n$ nicht mit. Dann ist 0 das Nullpolynom mit dem Grad $-\infty$, und $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ mit $a_n \neq 0$ ist ein Polynom vom Grad n.
5. Ein Polynom n-ten Grades $f(x)$ heißt normiert, wenn $a_n = 1$.
6. Es gilt $\text{grad}(f(x) \cdot g(x)) = \text{grad} f(x) + \text{grad} g(x)$, denn für $f(x), g(x) \neq 0$ mit $f(x) = a_n x^n + \ldots + a_0$ und $g(x) = b_m x^m + \ldots + b_0$ gilt
$$f(x) \cdot g(x) = a_n b_m x^{n+m} + (a_n b_{m-1} + a_{n-1} b_m) x^{n+m-1} + \ldots + (a_1 b_0 + a_0 b_1) x + a_0.$$

Satz 1.2 (Division mit Rest) *Seien $f(x), g(x) \in K[x]$ und es gelte $g(x) \neq 0$. Dann gibt es $q(x), r(x) \in K[x]$ mit $f(x) = q(x) \cdot g(x) + r(x)$ und $\text{grad} r(x) < \text{grad} g(x)$.*

Beweis. Gilt $\text{grad} g(x) > \text{grad} f(x)$, dann folgt mit $q(x) = 0$ und $r(x) = f(x)$ wegen $f(x) = 0 \cdot g(x) + f(x)$ die Behauptung. Sei also $\text{grad} g(x) \leq \text{grad} f(x)$ und damit insbesondere $\text{grad} f(x) \neq -\infty$. Wir beweisen den Satz nun durch Induktion nach $n := \text{grad} f(x) \geq 0$.
$n = 0$: Dann sind $g(x) = b_0$ und $f(x) = a_0$ konstant und $f(x) = (a_0 \cdot b_0^{-1}) \cdot g(x) + 0$.
$n > 0$: Mit der Darstellung

$$\begin{aligned} f(x) &= a_n x^n + \ldots + a_1 x + a_0 \quad \text{und} \\ g(x) &= b_m x^m + \ldots + b_1 x + b_0 \quad \text{wobei } m \leq n,\ b_m \neq 0 \end{aligned}$$

definieren wir $h(x) := f(x) - a_n b_m^{-1} x^{n-m} \cdot g(x)$, also

$$h(x) = a_n x^n + \ldots + a_0 - a_n x^n - a_n b_m^{-1} b_{m-1} x^{n-1} - \ldots - a_n b_m^{-1} b_0 x^{n-m}.$$

Es folgt $\text{grad} h(x) < n$ und mit der Induktionsvoraussetzung ist $h(x) = q'(x) \cdot g(x) + r(x)$ mit $q'(x), r(x) \in K[x]$, wobei $\text{grad} r(x) < \text{grad} g(x)$. Nach Konstruktion ergibt sich

$$f(x) = h(x) + a_n b_m^{-1} x^{n-m} g(x) = (q'(x) + a_n b_m^{-1} x^{n-m}) g(x) + r(x).$$

□

Beispiel. Wir betrachten über dem Körper $K = \mathbb{Q}$ die Polynome $f(x) = x^3 + 2x^2 + x + 1$ und $g(x) = x^2 - 2x + 1$. Dann gilt

$$\begin{array}{l} x^3 + 2x^2 + x + 1 : x^2 - 2x + 1 = x + 4 \\ \underline{x^3 - 2x^2 + x} \\ \qquad 4x^2 + 1 \\ \qquad \underline{4x^2 - 8x + 4} \\ \qquad\qquad 8x - 3 \end{array}$$

Wir erhalten also die Darstellung $x^3 + 2x^2 + x + 1 = (x + 4) \cdot (x^2 - 2x + 1) + 8x - 3$.

Definition 1.3 *Ein Polynom $f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \in K[x]$ hat an der Stelle $a \in K$ den Wert*

$$f(a) := a_n a^n + a_{n-1} a^{n-1} + \ldots + a_1 a + a_0.$$

a heißt Nullstelle von $f(x)$, wenn $f(a) = 0$.

Bemerkung.

1. Für $f(x), g(x) \in K[x], a \in K$ gilt

$$(f + g)(a) = f(a) + g(a) \quad \text{und} \quad (f \cdot g)(a) = f(a) \cdot g(a).$$

2. Gilt $f(x) = a_0 \in K$, so folgt $f(a) = a_0$ für alle $a \in K$.

Korollar 1.4 *Sei $f(x) \in K[x]$ und $a \in K$. Genau dann ist a eine Nullstelle von $f(x)$, wenn ein $g(x) \in K[x]$ existiert mit $f(x) = g(x)(x - a)$, wenn also $x - a$ das Polynom $f(x)$ teilt.*

Beweis. Es gibt $q(x), r(x) \in K[x]$ mit $\text{grad} r(x) < \text{grad}(x - a) = 1$ und $f(x) = q(x) \cdot (x - a) + r$. Einsetzen liefert $f(a) = q(a)(a - a) + r = r$.

$\square$

Durch vollständige Induktion folgt aus Korollar 1.4

Korollar 1.5 *Sei $f(x) \in K[x]$ mit $\text{grad} f(x) = n$ und seien $a_1, \ldots, a_l \in K$ verschieden. $a_1, \ldots, a_l$ sind genau dann Nullstellen von $f(x)$, wenn ein $g(x) \in K[x]$ existiert, so daß $f(x) = g(x) \cdot (x - a_1) \cdot \ldots \cdot (x - a_l)$. Insbesondere hat $f(x)$ also höchstens n verschiedene Nullstellen.*

Ist K ein Körper und $f(x) \in K[x]$, dann sagt man, daß $f(x)$ über K vollständig in Linearfaktoren zerfällt, wenn es ein $a \in K, a \neq 0$ und verschiedene $a_1, \ldots, a_k \in K$ sowie $n_1, \ldots, n_k \in \mathbb{N}$ gibt, so daß

$$f(x) = a \cdot (x - a_1)^{n_1} \cdot \ldots \cdot (x - a_k)^{n_k}$$

gilt. $a_1, \ldots, a_k$ sind dann genau die verschiedenen Nullstellen von $f(x)$, und $n_1, \ldots, n_k$ sind eindeutig bestimmt. a_i heißt n_i-fache Nullstelle von $f(x)$.

Der Fundamentalsatz der Algebra besagt, daß jedes Polynom $f(x) \in \mathbb{C}[x]$ über $\mathbb{C}$ vollständig in Linearfaktoren zerfällt. Ist nun speziell $f(x) \in \mathbb{R}[x]$, wobei $f(x) = x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$ ($a_i \in \mathbb{R}$), dann zerfällt $f(x)$ zwar über $\mathbb{C}$, i.a. aber nicht über $\mathbb{R}$ vollständig in Linearfaktoren. In diesem Falle existieren $\alpha_1, \ldots, \alpha_l \in \mathbb{R}$ und quadratische, normierte Polynome $q_1, \ldots, q_k \in \mathbb{R}[x]$ mit

$$f(x) = (x - \alpha_1) \cdot \ldots \cdot (x - \alpha_l) \cdot q_1(x) \cdot \ldots \cdot q_k(x) \,,$$

wobei kein $q_i(x)$ eine reelle Nullstelle hat. Ist eine komplexe Zahl $\alpha = a+bi \in \mathbb{C}$ mit $a, b \in \mathbb{R}$ Nullstelle von $f(x)$, dann ist auch $\bar{\alpha} = a - bi$ eine Nullstelle von $f(x)$. Ein $q_i(x)$ aus der obigen Darstellung hat dann die Form

$$q_i(x) = (x-\alpha)(x-\bar{\alpha}) = x^2 - 2ax + a^2 + b^2.$$

Beispiel. $f(x) = x^4 - 1 = (x^2-1)(x^2+1) = (x+1)(x-1)(x^2+1)$. Also hat $f(x)$ die Nullstellen $-1, 1, i, -i \in \mathbb{C}$.

Anwendung in der linearen Algebra.

Ist V ein n-dimensionaler K-Vektorraum, so bezeichnet

$$\mathrm{End}(V) := \{f : V \longrightarrow V \mid f \text{ linear}\}$$

die Menge aller Endomorphismen von V.
$\mathrm{End}(V)$ ist ein K-Vektorraum und ein Ring mit 1 bezüglich der Verknüpfungen $''+''$ und $''\circ''$, die durch

$$\begin{array}{rcl} f+g: & V \longrightarrow V & ; \quad v \longmapsto f(v)+g(v), \\ f\circ g: & V \longrightarrow V & ; \quad v \longmapsto f(g(v)), \\ k\cdot f: & V \longrightarrow V & ; \quad v \mapsto k\cdot f(v) \end{array}$$

gegeben sind. Wir prüfen nur beispielhaft die Distributivität nach:

$$\begin{array}{rcl} (f\circ(g+h))(v) & = & f((g+h)(v)) = f(g(v)+h(v)) = f(g(v)) + f(h(v))\,. \\ (f\circ g + f\circ h)(v) & = & f\circ g(v) + f\circ h(v) = f(g(v)) + f(h(v)). \end{array}$$

Das Einselement von $\mathrm{End}(V)$ ist die Identität von V: $1 = \mathrm{id}_V \in \mathrm{End}(V)$.
$\mathrm{End}(V)$ heißt Endomorphismenring ($\sim$ algebra) des Vektorraums V, und wegen Korollar 2.4 aus Kapitel 4 gilt $\dim_K \mathrm{End}(V) = n^2$.

$K_{n,n}$ ist bezüglich der Matrizenaddition und des Matrizenproduktes ebenfalls ein Ring mit 1 und ein K-Vektorraum mit $\dim_K K_{n,n} = n^2$. Bezüglich einer festen Basis von V ist eine Darstellung der Endomorphismen durch Matrizen gegeben:

$$\Theta : \mathrm{End}(V) \longrightarrow K_{n,n}, \quad \varphi \longmapsto A_\varphi.$$

Diese Zuordnung ist bijektiv und für alle $\varphi, \psi \in \mathrm{End}(V), k \in K$ gilt:

$$\begin{array}{rcl} A_{\varphi\circ\psi} & = & A_\varphi \cdot A_\psi \\ A_{\mathrm{id}} & = & E_n \\ A_{\varphi+\psi} & = & A_\varphi + A_\psi \\ A_{k\varphi} & = & k\cdot A_\varphi. \end{array}$$

Die ersten drei Gleichungen besagen, daß Θ ein Ringhomomorphismus, die letzten beiden, daß Θ ein Vektorraumhomomorphismus ist. Insbesondere sind die beiden Gruppen $\mathrm{Aut}(V)$

und $\mathrm{GL}(n;K)$ als Einheitengruppen der Ringe $\mathrm{End}(V)$ und $K_{n,n}$ isomorph.
Ist $f(x) \in K[x]$ mit $f(x) = a_n x^n + \ldots + a_1 x + a_0$ und $\varphi \in \mathrm{End}(V)$, so definiert man

$$f(\varphi) := a_n \cdot \varphi^n + \ldots + a_1 \cdot \varphi + a_0 \cdot \mathrm{id} \in \mathrm{End}(V),$$

wobei $\varphi^i = \underbrace{\varphi \circ \ldots \circ \varphi}_{i-\mathrm{mal}}$ die i-fache Hintereinanderausführung von φ bezeichnet. Damit gilt für $f(x), g(x) \in K[x]$

$$(f+g)(\varphi) = f(\varphi) + g(\varphi) \quad \text{und} \quad (f \cdot g)(\varphi) = f(\varphi) \circ g(\varphi).$$

φ heißt Nullstelle von $f(x)$, wenn $f(\varphi)$ die Nullabbildung ist.
Entsprechend definieren wir für $A \in K_{n,n}$ und $f(x) = a_n x^n + \ldots + a_1 x + a_0 \in K[x]$

$$f(A) := a_n A^n + \ldots + a_1 A + a_0 E_n \in K_{n,n},$$

und es gilt

$$(f+g)(A) = f(A) + g(A) \quad \text{und} \quad (f \cdot g)(A) = f(A) \cdot g(A).$$

A heißt Nullstelle von $f(x)$, wenn $f(A)$ die Nullmatrix ist.

Beispiel. Ist $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \in \mathbb{Q}_{2,2}$ und $f(x) = x^2 - 2x - 1 \in \mathbb{Q}[x]$, dann ist

$$f(A) = A^2 - 2A - E_2 = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} - \begin{pmatrix} 2 & 4 \\ 2 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

A ist also eine Nullstelle von $f(x)$. Für die assoziierte lineare Abbildung

$$\varphi : \mathbb{Q}^2 \longrightarrow \mathbb{Q}^2 \quad \text{mit} \quad e_1 \mapsto (1,1),\ e_2 \mapsto (2,1)$$

gilt $f(\varphi) = \varphi^2 - 2\varphi - \mathrm{id} = \mathcal{O}$, das heißt, φ ist ebenfalls eine Nullstelle von $f(x)$.

Satz 1.6 *Sei V ein K-Vektorraum mit $\dim_K V = n$. Dann existiert zu jedem $\varphi \in \mathrm{End}(V)$ ein $f(x) \in K[x], f(x) \neq 0$ mit $\mathrm{grad} f(x) \leq n^2$ und $f(\varphi) = \mathcal{O}$.*

Beweis. Wegen $\dim_K \mathrm{End}(V) = n^2$ sind $\mathrm{id}, \varphi, \varphi^2, \ldots, \varphi^{n^2}$ linear abhängig. Damit existieren $a_0, a_1, \ldots, a_{n^2}$, nicht alle 0, mit

$$a_{n^2} \cdot \varphi^{n^2} + \ldots + a_1 \cdot \varphi + a_0 \cdot \mathrm{id} = \mathcal{O},$$

und $f(x) := a_{n^2} x^{n^2} + \ldots + a_1 x + a_0$ erfüllt die Bedingungen des Satzes.

□

Entsprechend gilt

Satz 1.7 *Zu jedem $A \in K_{n,n}$ existiert ein $f(x) \in K[x], f(x) \neq 0$ mit $\mathrm{grad} f(x) \leq n^2$ und $f(A) = \mathcal{O}$.*

Bemerkung. Ist $\varphi \in \mathrm{End}(V)$ und $\dim_K V < \infty$, dann gibt es wegen Satz 1.6 ein $f(x) \in K[x], f(x) \neq 0$ mit den Eigenschaften

1. $f(\varphi) = \mathcal{O}$.
2. $f(x)$ ist normiert.
3. $f(x)$ hat mit 1) und 2) minimalen Grad.

Satz 1.8 *φ und $f(x)$ seien wie in der vorangehenden Bemerkung und $g(x) \in K[x]$ mit $g(\varphi) = \mathcal{O}$. Dann gibt es ein $h(x) \in K[x]$ mit $g(x) = f(x) \cdot h(x)$.*

Beweis. Wegen Satz 1.2 existieren $q(x), r(x) \in K[x]$ mit $g(x) = q(x) \cdot f(x) + r(x)$ und $\mathrm{grad} r(x) < \mathrm{grad} f(x)$. Daraus folgt $\mathcal{O} = g(\varphi) = q(\varphi) \circ f(\varphi) + r(\varphi)$, also $r(\varphi) = \mathcal{O}$, da $f(\varphi) = \mathcal{O}$. Wegen Punkt 3 der voranstehenden Bemerkung (gegebenenfalls muß $r(x)$ zunächst normiert werden) gilt $r(x) = 0$.

□

Korollar 1.9 *φ und $f(x)$ seien wie in der voranstehenden Bemerkung. Dann ist $f(x)$ durch φ eindeutig bestimmt.*

Beweis. Ist $\tilde{f}(x) \in K[x]$ auch ein Polynom mit den drei in der Bemerkung beschriebenen Eigenschaften, dann gibt es wegen Satz 1.8 zwei Polynome $h(x), \tilde{h}(x) \in K[x]$ mit $\tilde{f}(x) = f(x) \cdot h(x)$ und $f(x) = \tilde{f}(x) \cdot \tilde{h}(x)$, also $f(x) = f(x) \cdot h(x) \cdot \tilde{h}(x)$. Es folgt $h(x), \tilde{h}(x) \in K$. Da $f(x)$ und $\tilde{f}(x)$ normiert sind, ergibt sich $h(x) = \tilde{h}(x) = 1$, das heißt $f(x) = \tilde{f}(x)$.

□

Definition 1.10 *i) Sei V ein K-Vektorraum mit $\dim_K V < \infty$. Das durch $\varphi \in \mathrm{End}(V)$ eindeutig bestimmte normierte Polynom $f(x) \in K[x], f(x) \neq 0$ kleinsten Grades mit $f(\varphi) = \mathcal{O}$ heißt Minimalpolynom von φ, geschrieben $\mu_\varphi(x)$.*

ii) Das durch $A \in K_{n,n}$ eindeutig bestimmte normierte Polynom $f(x) \in K[x], f(x) \neq 0$ kleinsten Grades mit $f(A) = \mathcal{O}$ heißt Minimalpolynom von A, geschrieben $\mu_A(x)$.

Beispiel. Sei $K = \mathbb{Q}$. Dann ist $f(x) = x^2 - 2x - 1$ das Minimalpolynom von $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ und $f(x) = x - 1$ das Minimalpolynom von $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

2. Eigenwerte und Eigenvektoren

Definition 2.1 *Sei V ein K-Vektorraum und $\varphi \in \mathrm{End}(V)$. Dann heißt $c \in K$ Eigenwert von φ, wenn es ein $v \in V$, $v \neq \mathcal{O}$ gibt, so daß $\varphi(v) = c \cdot v$. Ist $c \in K$ ein Eigenwert von φ, so heißt jedes $v \in V$, $v \neq \mathcal{O}$ mit $\varphi(v) = c \cdot v$ Eigenvektor von φ zum Eigenwert c, und der von den Eigenvektoren aufgespannte Unterraum U_c von V heißt Eigenraum von φ zum Eigenwert c.*

Bemerkung.

1. Ist v ein Eigenvektor, so folgt $v \neq \mathcal{O}$.
2. Es gilt $U_c = \{v \in V \mid \varphi(v) = c \cdot v\}$.

Berechnung der Eigenwerte.
Es sei $\dim_K V = n$ und $A(= A_\varphi)$ die Matrixdarstellung von φ bezüglich einer fest gewählten Basis B. Ist c ein Eigenwert von φ und x der Koordinatenvektor eines Eigenvektors v von c, so gilt wegen $\varphi(v) = c \cdot v$ auch $Ax^t = c \cdot x^t$. Man nennt c dann auch Eigenwert von A und x Eigenvektor von A zum Eigenwert c. Das Problem, die Eigenwerte eines Endomorphismus′ zu berechnen, kann also zurückgeführt werden auf die Frage, für welche $c \in K$ das LGS

$$Ax^t = c \cdot x^t \tag{22}$$

eine nichttriviale Lösung besitzt. Dafür bringen wir (22) zunächst auf die Form

$$Ax^t - c \cdot x^t = \mathcal{O}^t\,, \quad \text{also} \quad (A - c \cdot E_n)x^t = \mathcal{O}^t\,.$$

Es gilt

$$A - c \cdot E_n = \begin{pmatrix} a_{11} - c & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - c & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - c \end{pmatrix}.$$

Wegen Satz 2.12 aus Kapitel 4 hat Gleichung (22) genau dann eine nichttriviale Lösung, wenn $|A - c \cdot E_n| = 0$. Es folgt also

$$\boxed{c \in K \text{ ist Eigenwert von } \varphi \quad \Longleftrightarrow \quad |A - c \cdot E_n| = 0\,.}$$

Für die Berechnung von $|A - c \cdot E_n|$ definieren wir $A - c \cdot E_n =: (b_{ij})$. Mit der Leibnizformel (Kapitel 4, Satz 3.2) ist

$$\begin{aligned} |A - c \cdot E_n| &= \sum_{\sigma \in \mathbf{S_n}} \mathrm{sgn}\sigma \cdot b_{1\sigma(1)} \cdot \ldots \cdot b_{n\sigma(n)} \\ &= b_{11} \cdot \ldots \cdot b_{nn} + \sum_{\substack{\sigma \in \mathbf{S_n} \\ \sigma \neq \mathrm{id}}} \mathrm{sgn}\sigma \underbrace{b_{1\sigma(1)} \cdot \ldots \cdot b_{n\sigma(n)}}_{c \text{ tritt höchstens } (n\text{-}2)\text{-mal auf}} \\ &= (a_{11} - c) \cdot \ldots \cdot (a_{nn} - c) + \sum_{\substack{\sigma \in \mathbf{S_n} \\ \sigma \neq \mathrm{id}}} \mathrm{sgn}\sigma \cdot b_{1\sigma(1)} \cdot \ldots \cdot b_{n\sigma(n)} \\ &= (-1)^n c^n + c_{n-1}c^{n-1} + c_{n-2}c^{n-2} + \ldots + c_1 c + c_0. \end{aligned}$$

Da in jedem Produkt $b_{1\sigma(1)} \cdot \ldots \cdot b_{n\sigma(n)}$ mit $\sigma \neq \mathrm{id}$ das c in höchstens $n-2$ Faktoren auftritt, folgt $c_{n-1} = (-1)^{n-1}(a_{11} + \ldots + a_{nn})$. Weiterhin gilt $c_0 = \det A$.

$$\chi_A(x) = (-1)^n x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$$

heißt charakteristisches Polynom von A, und es gilt $\mathrm{grad}\chi_A(x) = n$. Man nennt $\mathrm{Spur}A := a_{11} + \ldots + a_{nn}$ Spur von A.

Beispiel. Für $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ ist $A - c \cdot E_2 = \begin{pmatrix} 1-c & 2 \\ 2 & 1-c \end{pmatrix}$. Wegen

$$|A - cE_2| = (1-c)(1-c) - 4 = c^2 - 2c - 3$$

ist $\chi_A(x) = x^2 - 2x - 3$ das charakteristische Polynom von A, und es gilt $\mathrm{Spur}A = 2$ sowie $|A| = -3$. Die Eigenwerte von A sind -1 und 3, weil $\chi_A(x) = (x+1)(x-3)$.

Satz 2.2 *V sei ein n-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$. Sind A und A' zwei Matrixdarstellungen von φ bezüglich zweier Basen von V, so gilt $\chi_A = \chi_{A'}$.*

Beweis. Mit den Ausführungen vor Definition 2.13 aus Kapitel 4 gibt es ein $S \in \mathrm{GL}(n; K)$, so daß $A = S^{-1}A'S$. Wegen

$$c \cdot E_n = c \cdot S^{-1}S = S^{-1}(c \cdot E_n)S$$

folgt

$$\begin{aligned} |A - c \cdot E_n| &= |S^{-1}A'S - S^{-1}(c \cdot E_n)S| = |S^{-1}(A' - c \cdot E_n)S| \\ &= |S^{-1}| \cdot |A' - cE_n| \cdot |S| = |A' - cE_n| \,. \end{aligned}$$

□

Definition 2.3 *Es sei V ein n-dimensionaler K-Vektorraum, $\varphi \in \mathrm{End}(V)$ und A eine Matrixdarstellung von φ. Dann heißt*

i) $\chi_\varphi(x) := \chi_A(x)$ *charakteristisches Polynom von φ.*

ii) $\mathrm{Spur}\varphi := \mathrm{Spur}A$ *Spur von φ.*

iii) $\det \varphi := \det A$ *Determinante von φ.*

Bemerkung. Wegen Satz 2.2 sind $\chi_\varphi(x), \mathrm{Spur}\varphi$ und $\det \varphi$ wohldefiniert.

Korollar 2.4 *Ist V ein n-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$, so ist $c \in K$ genau dann ein Eigenwert von φ, wenn c eine Nullstelle von $\chi_\varphi(x)$ ist. Insbesondere hat φ also höchstens n verschiedene Eigenwerte.*

Satz 2.5 *Sei V ein K-Vektorraum und $\varphi \in \mathrm{End}(V)$. Sind $v_1, \ldots, v_n$ Eigenvektoren von φ zu den paarweise verschiedenen Eigenwerten $c_1, \ldots, c_n$, so sind $v_1, \ldots, v_n$ linear unabhängig.*

Beweis. Wir führen einen Widerspruchsbeweis und nehmen an, daß $v_1, \dots, v_n$ linear abhängig sind. Dann ist $n \geq 2$, und wir können sogar annehmen, daß jeweils $n-1$ der Vektoren linear unabhängig sind. Wegen Satz 1.13 aus Kapitel 3 läßt sich ein v_i als Linearkombination der anderen v_j schreiben. Es gelte also o.B.d.A.

$$v_n = k_1 v_1 + \ldots + k_{n-1} v_{n-1}\,.$$

Da φ linear ist, folgt

$$\varphi(v_n) = k_1 \varphi(v_1) + \ldots + k_{n-1}\varphi(v_{n-1})\,.$$

Wegen der Voraussetzungen des Satzes ist dann

$$c_n v_n = k_1 c_1 v_1 + \ldots + k_{n-1} c_{n-1} v_{n-1},$$

und mit der Darstellung von v_n ergibt sich

$$\mathcal{O} = k_1(c_1 - c_n)v_1 + \ldots + k_{n-1}(c_{n-1} - c_n)v_{n-1}\,.$$

Da wir angenommen haben, daß $v_1, \dots, v_{n-1}$ linear unabhängig sind, gilt $k_1(c_1 - c_n) = \ldots = k_{n-1}(c_{n-1} - c_n) = 0$, und da die Eigenwerte c_i nach Voraussetzung verschieden sind, folgt $k_1 = \ldots = k_{n-1} = 0$. Also ergibt sich $v_n = \mathcal{O}$, im Widerspruch zur Voraussetzung, daß v_n ein Eigenvektor ist.

□

Berechnung der Eigenvektoren.
Um die Eigenvektoren zu einem gegebenen Eigenwert c zu berechnen, muß das LGS

$$(A - c \cdot E_n)x^t = \mathcal{O}^t \tag{23}$$

gelöst werden. Die Dimension $\dim_K U_c$ des Eigenraumes $U_c = \{v \in V \mid \varphi(v) = c \cdot v\}$ ist dann durch die Dimension des Lösungsraumes von (23) gegeben.

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

$$\begin{aligned} |A - c \cdot E_3| &= \begin{vmatrix} 2-c & 1 & 1 \\ 1 & 2-c & 1 \\ 1 & -1 & 2-c \end{vmatrix} \\ &= (2-c)^3 + 1 - 1 - (2-c) + (2-c) - (2-c) \\ &= (2-c)((2-c)^2 - 1) \\ &= (2-c)(3-c)(1-c). \end{aligned}$$

Also sind $1, 2$ und 3 die Eigenwerte von A. Für den Eigenwert $c_1 = 1$ ergibt sich

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -2 & 0 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right).$$

Ein Eigenvektor von A zum Eigenwert c_1 ist also $v_1 = (1, 0, -1)$. Für den zweiten Eigenwert $c_2 = 2$ ergibt sich

$$\left(\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right),$$

und ein zugehöriger Eigenvektor ist $v_2 = (1, 1, -1)$. Entsprechend ergibt sich für $c_3 = 3$

$$\left(\begin{array}{ccc|c} -1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

und ein zugehöriger Eigenvektor ist $v_3 = (1, 1, 0)$.
Wegen Satz 2.5 sind die berechneten Eigenvektoren im Fall $\chi(K) \neq 2$ linear unabhängig und bilden eine Basis des K^3 ($\chi(K)$ bezeichnet die Charakteristik von K; vergleiche Kapitel 2.2). Wie man leicht sieht, sind v_1, v_2 und v_3 auch im Fall $\chi(K) = 2$ linear unabhängig.

Ist V ein n-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$, so ist eine Basis B von V gesucht, bezüglich der die Matrixdarstellung $A (= A_\varphi)$ von φ eine ″möglichst einfache″ Gestalt hat. Im folgenden wird nun gezeigt, daß in diesem Zusammenhang die Eigenvektoren von φ eine zentrale Rolle spielen.

Definition 2.6 *Eine Matrix $A = (a_{ij}) \in K_{n,n}$ heißt Diagonalmatrix, wenn $a_{ij} = 0$ für alle $i \neq j$ gilt, das heißt*

$$A = \begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}.$$

Satz 2.7 *Mit den Bezeichnungen wie oben gilt:*

i) A_φ ist genau dann eine Diagonalmatrix, wenn B nur aus Eigenvektoren besteht.

ii) Ist A_φ eine Diagonalmatrix, so sind die Diagonalelemente von A_φ genau die Eigenwerte von φ.

Beweis.

i) Ist $A_\varphi = (a_{ij})$ und $B = \{v_1, \ldots, v_n\}$, dann gilt gemäß der Definition von A_φ

$$\varphi(v_j) = a_{1j}v_1 + \ldots + a_{nj}v_n.$$

A_φ ist also genau dann eine Diagonalmatrix, wenn $\varphi(v_j) = a_{jj}v_j$ für alle $j \in \{1, \ldots, n\}$ gilt. Dieses ist aber gerade äquivalent dazu, daß v_j ein Eigenvektor von φ ist.

ii) Ist $A_\varphi = \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$ eine Diagonalmatrix, so gilt für das charakteristische Polynom $\chi_\varphi(x) = (a_{11} - x) \cdot \ldots \cdot (a_{nn} - x)$. Damit sind $a_{11}, \ldots, a_{nn}$ genau die Eigenwerte von φ.

□

Definition 2.8 *V sei ein endlich-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$. Dann heißt φ diagonalisierbar, wenn es eine Basis von V gibt, bezüglich der φ einer Diagonalmatrix zugeordnet ist.*

Korollar 2.9 *Zerfällt $\chi_\varphi(x)$ über K vollständig in paarweise verschiedene Linearfaktoren, so ist φ diagonalisierbar.*

Beweis. Da $\mathrm{grad}\chi_\varphi(x) = \dim_K V = n$, hat φ genau n paarweise verschiedene Eigenwerte. Wegen Satz 2.5 folgt die Behauptung.

□

Bemerkung. Die Umkehrung von Korollar 2.9 gilt nicht.

Satz 2.10 *Ist V ein K-Vektorraum mit $n = \dim_K V$, dann ist $\varphi \in \mathrm{End}(V)$ genau dann diagonalisierbar, wenn gilt*

i) $\chi_\varphi(x)$ zerfällt über K vollständig in Linearfaktoren.

ii) Ist c_i ein n_i-facher Eigenwert (d.h., ist c_i eine n_i-fache Nullstelle von $\chi_\varphi(x)$), so gilt $\dim_K U_{c_i} = n_i$.

Beweis. "$\Rightarrow$": Ist φ diagonalisierbar, so gilt bezüglich einer geeigneten Basis

$$A_\varphi = \begin{pmatrix} c_1 & & & \\ & c_2 & & 0 \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix},$$

und damit gilt für das charakteristische Polynom $\chi_\varphi(x) = (c_1 - x) \cdot \ldots \cdot (c_n - x)$. Sei c_1 ein m-facher Eigenwert, etwa $c_1 = \ldots = c_m$ und $c_1 \neq c_i$ für $i > m$. Da $\dim_K U_{c_1}$ gerade durch die Dimension des Lösungsraumes von $(A_\varphi - c_1 E_n)x^t = \mathcal{O}^t$ und damit durch $n - \mathrm{Rg}(A_\varphi - c_1 E_n)$ gegeben ist, folgt wegen

$$A_\varphi - c_1 E_n = \begin{pmatrix} 0 & & & & & 0 \\ & \ddots & & & & \\ & & 0 & & & \\ & & & c_{m+1} - c_1 & & \\ & & & & \ddots & \\ 0 & & & & & c_n - c_1 \end{pmatrix},$$

daß $\mathrm{Rg}(A_\varphi - c_1 E_n) = n - m$ und damit $\dim_K U_{c_1} = m$ gilt. Entsprechendes folgt für $c_i, i > 1$.
"$\Leftarrow$": Nach Voraussetzung ist $\chi_\varphi(x) = (c_1 - x)^{n_1} \cdot \ldots \cdot (c_l - x)^{n_l}$, wobei $c_1, \ldots, c_l$ verschieden sind. Sei $\{v_1^i, \ldots, v_{n_i}^i\}$ eine Basis von U_{c_i}. Wir zeigen, daß $\{v_j^i \mid i = 1, \ldots, l;\ j = 1, \ldots, n_i\}$ linear unabhängig ist. Wegen $n_1 + \ldots + n_l = n$ und Satz 2.7 folgt dann die Behauptung. Sei also

$$\sum_{i=1}^{l} \underbrace{\sum_{j=1}^{n_i} k_{ij} v_j^i}_{=:w_i} = \mathcal{O} .$$

Wegen $w_i \in U_{c_i}$ ist $w_i = \mathcal{O}$ oder w_i ein Eigenvektor zum Eigenwert c_i. Aus $w_1 + \ldots + w_l = \mathcal{O}$ folgt $w_1 = \ldots = w_l = \mathcal{O}$ mit Satz 2.5. Damit ist $\mathcal{O} = w_i = \sum_{j=1}^{n_i} k_{ij} v_j^i$. Weil $\{v_1^i, \ldots, v_{n_i}^i\}$ linear unabhängig ist, folgt $k_{i1} = \ldots = k_{in_i} = 0$ für alle i.

□

Bemerkung. Die Vielfachheit von c_i ist stets eine obere Schranke für $\dim_K U_{c_i}$.

Erinnerung (vergleiche Kapitel 4.2)

$$\begin{tikzcd}[row sep=small] & A_\varphi \text{ bzgl. } B \\ \varphi \arrow[ur] \arrow[dr] & A'_\varphi = S^{-1} A_\varphi S \\ & A'_\varphi \text{ bzgl. } B' \end{tikzcd}$$

Ist A_φ die Matrixdarstellung von φ bezüglich einer Basis B und A'_φ die Matrixdarstellung von φ bezüglich einer Basis B' von V, so gilt $A'_\varphi = S^{-1} A_\varphi S$, wobei S die Transformationsmatrix des Basiswechsels $B' \to B$ ist. Es gilt dann $S \in \mathrm{GL}(n; K)$. Ist also zum Beispiel φ durch A_φ bezüglich B gegeben, so ist φ genau dann diagonalisierbar, wenn ein $S \in \mathrm{GL}(n; K)$ existiert, so daß $S^{-1} A_\varphi S$ diagonal ist.

Definition 2.11 *Seien $A, B \in K_{n,n}$.*

i) A und B heißen ähnlich, wenn ein $S \in \mathrm{GL}(n; K)$ existiert, so daß $A = S^{-1} B S$ gilt.

ii) A heißt diagonalisierbar, wenn A zu einer Diagonalmatrix ähnlich ist.

Bemerkung.

1. $\varphi \in \mathrm{End}(V)$ ist genau dann diagonalisierbar, wenn A_φ diagonalisierbar ist.

2. Ähnliche Matrizen haben dasselbe charakteristische Polynom, dieselbe Determinante und dieselbe Spur (vgl. Satz 2.2).

3. Die Sätze 2.5 und 2.10 gelten entsprechend für Matrizen.

4. Ist $S^{-1} A S$ eine Diagonalmatrix, so ist die i-te Spalte von S ein Eigenvektor von A zum Eigenwert c_i und umgekehrt, denn

$$S^{-1} A S = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} \quad \Longleftrightarrow \quad A S = S \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} .$$

Beispiel. Sei $K = \mathbb{Q}$.

1. Für $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 2 & -1 \end{pmatrix}$ gilt $|A - cE_3| = \begin{vmatrix} -c & 1 & 0 \\ 0 & -c & 1 \\ 2 & 2 & -1-c \end{vmatrix}$. Mit Hilfe der Regel von Sarrus erhält man also
$$\chi_A(x) = x^2(-1-x) + 2 + 2x = -x^3 - x^2 + 2x + 2 = (2-x^2)(x+1)\,.$$
Da $\chi_A(x)$ über $\mathbb{Q}$ nicht in Linearfaktoren zerfällt, ist A über $\mathbb{Q}$ nicht diagonalisierbar. Über $\mathbb{R}$ zerfällt $\chi_A(x)$ aber in paarweise verschiedene Linearfaktoren, das heißt, über $\mathbb{R}$ ist A diagonalisierbar.

2. Die Matrix $A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ hat das charakteristische Polynom $\chi_A(x) = x^2 - 2x + 1 = (x-1)^2$. Also ist 1 ein doppelter Eigenwert von A.
Wegen $A - 1 \cdot E_2 = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}$ folgt $\dim U_1 = 2 - \mathrm{Rg}(A - 1 \cdot E_2) = 2 - 1 = 1 \neq 2$.
A ist also nicht diagonalisierbar (weder über $\mathbb{Q}$, noch über $\mathbb{R}$ oder einem anderen Körper).

3. Für die Matrix A wie im Beispiel nach Satz 2.5 gilt
$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix} \quad \text{und} \quad \begin{array}{c|c|c|c|} \text{EW:} & 1 & 2 & 3 \\ \hline \text{EV:} & (1,0,-1) & (1,1,-1) & (1,1,0) \end{array}\,,$$
wobei EW die Eigenwerte und EV zugehörige Eigenvektoren bezeichnen. Um A zu diagonalisieren, können wir die Eigenvektoren als Spaltenvektoren der Transformationsmatrix S wählen (vergleiche Bemerkung 4):
$$\mathrm{GL}(3;\mathbb{Q}) \ni S := \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{pmatrix} \quad \text{und} \quad S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Satz 2.12 *Sei V ein endlich-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$. Ein $c \in K$ ist genau dann ein Eigenwert von φ, wenn c eine Nullstelle des Minimalpolynoms $\mu_\varphi(x)$ ist.*

Beweis. "$\Rightarrow$": Ist v ein Eigenvektor von φ zu c, dann gilt
$$\mathcal{O} = \mu_\varphi(\varphi)(v) = \mu_\varphi(c) \cdot v\,,$$
also $\mu_\varphi(c) = 0$, da $v \neq \mathcal{O}$.
"$\Leftarrow$": Gilt $\mu_\varphi(c) = 0$, dann existiert nach Korollar 1.4 ein Polynom $q(x) \in K[x]$, so daß $\mu_\varphi(x) = (x-c) \cdot q(x)$. Somit folgt $\mu_\varphi(\varphi) = (\varphi - c \cdot \mathrm{id}) \circ q(\varphi)$. Es ist $\mathrm{grad} q(x) = \mathrm{grad} \mu_\varphi(x) - 1 < \mathrm{grad} \mu_\varphi(x)$, und $q(x)$ ist normiert. Da $\mu_\varphi(x)$ das Minimalpolynom von φ ist, folgt $q(\varphi) \neq \mathcal{O}$. Also gibt es ein $v \in V$ mit $w := q(\varphi)(v) \neq \mathcal{O}$ und $(\varphi - c \cdot \mathrm{id})(w) = (\varphi - c \cdot \mathrm{id}) q(\varphi)(v) = \mu_\varphi(\varphi)(v) = \mathcal{O}$. Damit ist $\varphi(w) = c \cdot w$ und w Eigenvektor von φ zum Eigenwert c.

$\square$

Bemerkung. Im allgemeinen sind das charakteristische Polynom $\chi_\varphi(x)$ und das Minimalpolynom $\mu_\varphi(x)$ verschieden.

Satz 2.13 *Ist V ein endlich-dimensionaler K-Vektorraum, dann ist ein Endomorphismus $\varphi \in \mathrm{End}(V)$ genau dann diagonalisierbar, wenn das Minimalpolynom $\mu_\varphi(x)$ über K vollständig in paarweise verschiedene Linearfaktoren zerfällt.*

Beweis. $''\Rightarrow''$: Wegen Satz 2.7 hat V eine Basis $B = \{v_1, \ldots, v_n\}$ von Eigenvektoren, etwa $\varphi(v_i) = c_i v_i$, $i = 1, \ldots, n$. O.B.d.A. seien $c_1, \ldots, c_r$ genau die verschiedenen Eigenwerte von φ. Wir zeigen, daß $\mu_\varphi(x)$ das Polynom

$$f(x) := (x - c_1) \cdot \ldots \cdot (x - c_r) \in K[x]$$

teilt. Für jedes $i = 1, \ldots, r$ gilt

$$f(x) = \left(\prod_{\substack{j=1 \\ j \neq i}}^{r} (x - c_j) \right) (x - c_i)$$

und damit

$$f(\varphi)(v) = \left(\prod_{\substack{j=1 \\ j \neq i}}^{r} (\varphi - c_j \cdot \mathrm{id}) \right) (\varphi - c_i \cdot \mathrm{id})(v) = \left(\prod_{\substack{j=1 \\ j \neq i}}^{r} (\varphi - c_j \cdot \mathrm{id}) \right) (\varphi(v) - c_i \cdot v) = \mathcal{O},$$

falls $v \in B$ zum Eigenwert c_i gehört. Da B eine Basis von V ist, folgt $f(\varphi) = \mathcal{O}$, das heißt, $f(\varphi)$ ist die Nullabbildung. Wegen Satz 1.8 teilt somit $\mu_\varphi(x)$ das Polynom $f(x)$.
$''\Leftarrow''$: Nach Voraussetzung ist $\mu_\varphi(x) = (x - c_1) \cdot \ldots \cdot (x - c_r)$ mit paarweise verschiedenen Nullstellen. Wir definieren

$$l_i(x) := \frac{(x - c_1) \cdot \ldots \cdot (x - c_{i-1}) \cdot (x - c_{i+1}) \cdot \ldots \cdot (x - c_r)}{(c_i - c_1) \cdot \ldots \cdot (c_i - c_{i-1}) \cdot (c_i - c_{i+1}) \cdot \ldots \cdot (c_i - c_r)} \quad \text{und}$$

$$c_{(i)} := (c_i - c_1) \cdot \ldots \cdot (c_i - c_{i-1}) \cdot (c_i - c_{i+1}) \cdot \ldots \cdot (c_i - c_r).$$

Dann gilt $l_i(c_i) = 1$ und $l_i(c_j) = 0$ für $i \neq j$. Definieren wir weiterhin $\rho(x) := l_1(x) + \ldots + l_r(x) - 1$, so ist

$$\rho(c_i) = l_1(c_i) + \ldots + l_r(c_i) - 1 = 0, \quad i = 1, \ldots, r.$$

$\rho(x)$ hat damit r verschiedene Nullstellen. Da aber $\mathrm{grad}\rho(x) < r$, folgt wegen Korollar 1.5, daß $\rho(x) = 0$ und damit $l_1(x) + \ldots + l_r(x) = 1$. Also gilt $l_1(\varphi) + \ldots + l_r(\varphi) = \mathrm{id}$, das heißt,

$$l_1(\varphi)(V) + \ldots + l_r(\varphi)(V) = V \,. \tag{24}$$

Im Hinblick auf Satz 2.7 bleibt zu zeigen, daß V eine Basis von Eigenvektoren hat. Da jedes $l_i(\varphi)(V)$ ein Unterraum von V ist, reicht es aus nachzuweisen, daß jedes $v \in l_i(\varphi)(V)$ mit $v \neq \mathcal{O}$ ein Eigenvektor von φ zum Eigenwert c_i ist. Sei also $v = l_i(\varphi)(w)$, $w \in V$. Dann gilt

$$(\varphi - c_i \cdot \mathrm{id})(v) = (\varphi - c_i \cdot \mathrm{id}) \circ l_i(\varphi)(w) = c_{(i)}^{-1} \cdot \mu_\varphi(\varphi)(w) = \mathcal{O}\,,$$

also $\varphi(v) - c_i \cdot v = \mathcal{O}$.

□

Beispiel. Die Matrix $A = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & -2 \\ -1 & 1 & 0 \end{pmatrix}$ ist keine Diagonalmatrix, also $\mu_A(x) \neq x - c$. Wegen

$$A^2 - 3A + 2E_3 = \begin{pmatrix} 4 & 0 & 0 \\ -3 & 7 & -6 \\ -3 & 3 & -2 \end{pmatrix} - 3 \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & -2 \\ -1 & 1 & 0 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

folgt $\mu_A(x) = x^2 - 3x + 2 = (x-1)(x-2)$ und damit die Diagonalisierbarkeit von A.

3. Der Satz von Cayley-Hamilton

Satz 3.1 *Ist V ein endlich-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$ mit dem Minimalpolynom $\mu_\varphi(x)$ und dem charakteristischen Polynom $\chi_\varphi(x)$, dann gilt*

i) $\mu_\varphi(x)$ teilt $\chi_\varphi(x)$.

ii) Es gibt ein $r \in \mathbb{N}$, so daß $\chi_\varphi(x)$ das Polynom $\mu_\varphi^r(x)$ teilt.

Korollar 3.2 *Es gilt $\chi_\varphi(\varphi) = \mathcal{O}$.*

Korollar 3.3 (Cayley-Hamilton) *Für jedes $A \in K_{n,n}$ gilt $\chi_A(A) = \mathcal{O}$.*

Beispiel. Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist das charakteristische Polynom $\chi_A(x) = x^2 - (a+d)x + ad - bc$, und es folgt

$$\begin{aligned} \chi_A(A) &= A^2 - (a+d)A + (ad-bc)E_2 \\ &= \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} - \begin{pmatrix} a^2+ad & ba+bd \\ ac+cd & ad+d^2 \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Beweis von Satz 3.1. Wir definieren rekursiv eine Folge von Unterräumen:

$$\begin{aligned} U_0 &:= \{\mathcal{O}\}; \text{ sei } v_1 \notin U_0\,. \\ U_1 &:= U_0 + [\{v_1, \varphi(v_1), \varphi^2(v_1), \ldots\}]; \text{ sei } v_2 \notin U_1. \\ U_2 &:= U_1 + [\{v_2, \varphi(v_2), \varphi^2(v_2), \ldots\}]; \text{ sei } v_3 \notin U_2. \\ &\vdots \\ U_i &:= U_{i-1} + [\{v_i, \varphi(v_i), \varphi^2(v_i), \ldots\}]; \text{ sei } v_{i+1} \notin U_i. \\ &\vdots \\ U_r &:= U_{r-1} + [\{v_r, \varphi(v_r), \varphi^2(v_r), \ldots\}] = V. \end{aligned}$$

Dann gilt

$$\begin{array}{rcl} U_1 & = & [\{\varphi^\nu(v_1) \mid \nu \in \mathbb{N}_0\}], \\ U_2 & = & [\{\varphi^\nu(v_1), \varphi^\nu(v_2) \mid \nu \in \mathbb{N}_0\}], \\ & \vdots & \\ U_i & = & [\{\varphi^\nu(v_j) \mid \nu \in \mathbb{N}_0;\ j = 1, \ldots, i\}], \\ & \vdots & \\ U_r & = & [\{\varphi^\nu(v_j) \mid \nu \in \mathbb{N}_0;\ j = 1, \ldots, r\}], \end{array}$$

und für $k \in \mathbb{N}_0$ ist

$$\varphi^k(U_i) = [\{\varphi^{\nu+k}(v_j) \mid \nu \in \mathbb{N}_0,\ j = 1, \ldots, i\}] \subseteq U_i\,.$$

Damit gilt für jedes $f(x) \in K[x]$

$$f(\varphi)(U_i) \subseteq U_i\,. \tag{25}$$

Wir betrachten nun zunächst U_1. Es gibt ein $m_1 \in \mathbb{N}$ mit

$$\begin{array}{rcl} v_1 & \neq & \mathcal{O}, \\ \varphi(v_1) & \notin & [v_1], \\ \varphi^2(v_1) & \notin & [v_1, \varphi(v_1)], \\ & \vdots & \\ \varphi^{m_1-1}(v_1) & \notin & [v_1, \varphi(v_1), \ldots, \varphi^{m_1-2}(v_1)], \\ \varphi^{m_1}(v_1) & \in & [v_1, \varphi(v_1), \ldots, \varphi^{m_1-1}(v_1)]. \end{array}$$

Wegen Aufgabe 3.8 aus Kapitel 3 ist $\{v_1, \varphi(v_1), \ldots, \varphi^{m_1-1}(v_1)\} =: B_1$ linear unabhängig. Durch vollständige Induktion zeigt man nun, daß $\varphi^l(v_1) \in [v_1, \varphi(v_1), \ldots, \varphi^{m_1-1}(v_1)]$ für alle $l \in \mathbb{N}$ gilt, und damit ist B_1 eine Basis von U_1. Sei $\varphi^{m_1}(v_1) = \sum_{\nu=0}^{m_1-1}(-a_{1\nu})\varphi^\nu(v_1)$, also

$$\varphi^{m_1}(v_1) + \sum_{\nu=0}^{m_1-1} a_{1\nu}\varphi^\nu(v_1) = \mathcal{O}\,.$$

Wir setzen $f_1(x) := x^{m_1} + \sum_{\nu=0}^{m_1-1} a_{1\nu}x^\nu \in K[x]$. Dann ist $f_1(x)$ normiert mit $f_1(\varphi)(v_1) = \mathcal{O}$ und bezüglich dieser Eigenschaften hat $f_1(x)$ minimalen Grad.
Entsprechend der vorangehenden Überlegungen hat U_{i-1} eine Basis der Form

$$B_{i-1} = \{\varphi^{\nu_j}(v_j) \mid j = 1, \ldots, i-1;\ \nu_j = 0, \ldots, m_j - 1\}.$$

Wir betrachten nun U_i. Es gibt ein $m_i \in \mathbb{N}$ mit

$$\begin{array}{rcl} v_i & \notin & U_{i-1}, \\ \varphi(v_i) & \notin & [B_{i-1} \cup \{v_i\}], \\ \varphi^2(v_i) & \notin & [B_{i-1} \cup \{v_i, \varphi(v_i)\}], \\ & \vdots & \\ \varphi^{m_i-1}(v_i) & \notin & [B_{i-1} \cup \{v_i, \varphi(v_i), \ldots, \varphi^{m_i-2}(v_i)\}], \\ \varphi^{m_i}(v_i) & \in & [B_{i-1} \cup \{v_i, \varphi(v_i), \ldots, \varphi^{m_i-1}(v_i)\}]. \end{array}$$

Wegen Aufgabe 3.8 aus Kapitel 3 ist $B_{i-1} \cup \{v_i, \varphi(v_i), \ldots, \varphi^{m_i-1}(v_i)\} =: B_i$ linear unabhängig, und wie oben ergibt sich, daß B_i eine Basis von U_i ist. Sei

$$\varphi^{m_i}(v_i) = u_{i-1} + \sum_{\nu=0}^{m_i-1} (-a_{i\nu})\varphi^{\nu}(v_i)\,,$$

wobei $u_{i-1} \in U_{i-1}$. Das Polynom $f_i(x) := x^{m_i} + \sum_{\nu=0}^{m_i-1} a_{i\nu}x^{\nu} \in K[x]$ ist normiert, es gilt $f_i(\varphi)(v_i) \in U_{i-1}$ und mit diesen Eigenschaften hat $f_i(x)$ minimalen Grad. Damit ist $B = \{\varphi^{\nu_j}(v_j) \mid j = 1, \ldots, r;\ \nu_j = 0, \ldots, m_j - 1\}$ eine Basis von V, und A sei die zugehörige Matrixdarstellung von φ, also

$$\left(\begin{array}{cccccc|cccccc|cc}
0 & 0 & \cdots & \cdots & 0 & -a_{10} & * & \cdots & & & \cdots & * & * & \cdots \\
1 & 0 & \cdots & & & -a_{11} & \vdots & & & & & \vdots & \vdots & \\
0 & 1 & & & & \vdots & \vdots & & & & & \vdots & & \\
\vdots & 0 & \ddots & & & \vdots & \vdots & & & & & \vdots & & \\
\vdots & \vdots & & \ddots & 0 & \vdots & \vdots & & & & & \vdots & \vdots & \\
0 & 0 & \cdots & 0 & 1 & -a_{1(m_1-1)} & * & \cdots & & & \cdots & * & * & \cdots \\ \hline
0 & 0 & \cdots & & 0 & 0 & 0 & \cdots & & \cdots & 0 & -a_{20} & * & \cdots \\
\vdots & \vdots & & & \vdots & \vdots & 1 & 0 & & & \vdots & -a_{21} & \vdots & \\
\vdots & & & & & \vdots & 0 & 1 & & & \vdots & \vdots & & \\
\vdots & & & & & \vdots & \vdots & 0 & \ddots & & \vdots & \vdots & & \\
\vdots & & & & & \vdots & \vdots & \vdots & & 1 & 0 & \vdots & \vdots & \\
0 & 0 & \cdots & & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -a_{2(m_2-1)} & * & \cdots \\ \hline
0 & 0 & \cdots & & 0 & 0 & 0 & \cdots & & \cdots & 0 & 0 & * & \cdots \\
\vdots & & & & & \vdots & \vdots & & & & \vdots & \vdots & \vdots &
\end{array}\right) = A.$$

In Blockform hat A die Darstellung

$$A = \begin{pmatrix} A_1 & & & * \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_r \end{pmatrix}, \text{ wobei } A_i = \begin{pmatrix} 0 & 0 & & 0 & -a_{i0} \\ 1 & 0 & & \vdots & -a_{i1} \\ 0 & 1 & & & \vdots \\ \vdots & 0 & \ddots & 0 & \vdots \\ & & & 1 & \vdots \end{pmatrix}.$$

Wegen Aufgabe 4.10 aus Kapitel 4 und Aufgabe 5.6 ergibt sich nun

1. $\chi_A(x) = \chi_{A_1}(x) \cdot \ldots \cdot \chi_{A_r}(x)$.

2. $\chi_{A_i}(x) = \pm f_i(x)$.

Damit folgt $\chi_\varphi(x) = \chi_A(x) = \pm f_1(x) \cdot \ldots \cdot f_r(x)$. Wir zeigen jetzt die Behauptungen i) und ii) des Satzes.

i) $\mu_\varphi(x)$ teilt $\chi_\varphi(x)$. Wegen Satz 1.8 genügt es, $\chi_\varphi(\varphi) = \mathcal{O}$ zu zeigen.
Behauptung: $f_i(\varphi)(U_i) \subseteq U_{i-1}$.
Da $f_i(x)x^\nu = x^\nu f_i(x)$ für alle $\nu \in \mathbb{N}_0$, ist auch $f_i(\varphi)\varphi^\nu = \varphi^\nu f_i(\varphi)$ für alle $\nu \in \mathbb{N}_0$, und wegen (25) und den Eigenschaften von $f_i(x)$ folgt

$$\begin{aligned} f_i(\varphi)(U_i) &= f_i(\varphi)(U_{i-1}) + f_i(\varphi)([\varphi^\nu(v_i) \mid \nu \in \mathbb{N}_0]) \subseteq U_{i-1} + [f_i(\varphi)\varphi^\nu(v_i) \mid \nu \in \mathbb{N}_0] \\ &= U_{i-1} + [\varphi^\nu f_i(\varphi)(v_i) \mid \nu \in \mathbb{N}_0] \subseteq U_{i-1} + U_{i-1} \subseteq U_{i-1}\,. \end{aligned}$$

Also ist mit $U_r = V$

$$\begin{aligned} \chi_\varphi(\varphi)(V) &= f_1(\varphi) \circ \ldots \circ f_r(\varphi)(U_r) \subseteq f_1(\varphi) \circ \ldots \circ f_{r-1}(\varphi)(U_{r-1}) \subseteq \ \ldots \ \subseteq \\ &\subseteq f_1(\varphi)(U_1) \subseteq U_0 = \{\mathcal{O}\}. \end{aligned}$$

ii) $\exists\, r \in \mathbb{N}:\ \chi_\varphi(x)$ teilt $\mu_\varphi^r(x)$.
Behauptung: $f_i(x)$ teilt $\mu_\varphi(x)$ (dann ist $\chi_\varphi(x)$ Teiler von $\mu_\varphi^r(x)$).
Es gilt $\mu_\varphi(x) = q_i(x) \cdot f_i(x) + r_i(x)$ mit $\operatorname{grad} r_i(x) < \operatorname{grad} f_i(x)$. Damit ist

$$\begin{aligned} \mathcal{O} = \mu_\varphi(\varphi)(v_i) &= q_i(\varphi) \circ f_i(\varphi)(v_i) + r_i(\varphi)(v_i) \\ &\in q_i(\varphi)(U_{i-1}) + r_i(\varphi)(v_i) \subseteq U_{i-1} + r_i(\varphi)(v_i). \end{aligned}$$

Daraus folgt $r_i(\varphi)(v_i) \in U_{i-1}$. Wegen der Minimalitätseigenschaft von $\operatorname{grad} f_i(x)$ und $\operatorname{grad} r_i(x) < \operatorname{grad} f_i(x)$ ergibt sich $r_i(x) = 0$.

□

Korollar 3.4 *Sei V ein K-Vektorraum mit $\dim_K V = n$ und sei $\varphi \in \operatorname{End}(V)$. Gibt es ein $v \in V$ mit $V = [v, \varphi(v), \ldots, \varphi^{n-1}(v)]$, so gilt $\chi_\varphi = \pm\mu_\varphi$.*

Beweis. Wähle $v_1 = v$ im obigen Beweis.

□

4. Die Jordansche Normalform

Sei V ein n-dimensionaler K-Vektorraum und $\varphi \in \operatorname{End}(V)$. Ist φ nicht diagonalisierbar, so ist A_φ bezüglich keiner Basis von V eine Diagonalmatrix. Im folgenden soll nun erörtert werden, wie "einfach" eine Matrixdarstellung A_φ sein kann und wie man erkennen kann, ob eine gegebene Matrix A eine Darstellung von φ bezüglich einer geeigneten Basis von V ist.

Satz 4.1 (Jordansche Normalform) *Sei V ein n-dimensionaler K-Vektorraum sowie $\varphi \in \operatorname{End}(V)$. Zerfällt $\chi_\varphi(x)$ über K vollständig in Linearfaktoren, so hat φ bezüglich einer geeigneten Basis die Darstellung*

$$A_\varphi = \begin{pmatrix} A_1 & & & \\ & A_2 & & 0 \\ 0 & & \ddots & \\ & & & A_k \end{pmatrix} \quad \textit{mit} \quad A_i = \begin{pmatrix} c_i & & & & \\ 1 & c_i & & 0 & \\ & 1 & \ddots & & \\ & 0 & \ddots & \ddots & \\ & & & 1 & c_i \end{pmatrix}.$$

Auf der Hauptdiagonalen von A_φ stehen die Eigenwerte von φ gemäß ihrer Vielfachheit.

Bemerkung. Gilt $A_i \in K_{n_i,n_i}$, so heißt A_i Jordanblock der Größe n_i zum Eigenwert c_i.

Spezialfall: Ist φ diagonalisierbar, so ist $A_\varphi = \begin{pmatrix} c_1 & & & 0 \\ & c_2 & & \\ & & \ddots & \\ 0 & & & c_n \end{pmatrix}$, also $A_i = (c_i)$.

Satz 4.2 (Jordansche Normalform) *Zerfällt für $B \in K_{n,n}$ das charakteristische Polynom $\chi_B(x)$ über K vollständig in Linearfaktoren, so ist B ähnlich zu einer Matrix $A \in K_{n,n}$ mit*

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix}, \quad \textit{wobei } A_i \textit{ ein Jordanblock ist.}$$

Wir zeigen zunächst:

Lemma 4.3 *Satz 4.1 gilt, wenn alle Eigenwerte von φ gleich sind.*

Beweis. φ habe zunächst den n-fachen Eigenwert 0. Dann gilt $\chi_\varphi(x) = \pm x^n$, also $\varphi^n = \mathcal{O}$ wegen des Satzes von Cayley-Hamilton (Korollar 3.2). Definieren wir für jedes $v \in V, v \neq \mathcal{O}$:

$$\mathrm{ord} v := \min\{i \in \mathbb{N} \mid \varphi^i(v) = \mathcal{O}\} \leq n\,,$$

dann gilt $v, \varphi(v), \ldots, \varphi^{\mathrm{ord} v - 1}(v) \neq \mathcal{O}$ und $\varphi^{\mathrm{ord} v}(v) = \mathcal{O}$. Die Menge

$$B(v) = \{v, \varphi(v), \ldots, \varphi^{\mathrm{ord} v - 1}(v)\}$$

heißt Bahn von v.
Behauptung: Es gibt $v_1, \ldots, v_l \in V$, so daß $B := B(v_1) \dot\cup \ldots \dot\cup B(v_l)$ eine Basis von V ist, wobei $\dot\cup$ die disjunkte Vereinigung bezeichnet. Bezüglich B hat φ dann die Darstellung

$$\left(\begin{array}{c|c|c}
\begin{array}{ccccc} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ddots & \vdots & \vdots \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 1 & 0 \end{array} & \multicolumn{2}{c}{0} \\ \hline
0 & \begin{array}{ccccc} 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ddots & \vdots & \vdots \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 1 & 0 \end{array} & \\ \hline
 & & \ddots
\end{array}\right),$$

also die gewünschte Form.
Beweis der Behauptung: Jede Basis B von V hat die Form

$$B = B(v_1)\dot{\cup}\ldots\dot{\cup}B(v_l)\dot{\cup}\{a_1,\ldots,a_s\}$$

mit $\mathrm{ord}a_s \le \ldots \le \mathrm{ord}a_1 \le \mathrm{ord}v_i =: t_i$ für $i = 1,\ldots,l$. Wähle B so, daß zunächst s und dann mit diesem s auch $\mathrm{ord}a_1 + \ldots + \mathrm{ord}a_s$ minimal ist.
Annahme: $s \ge 1$. Wir definieren $\mathrm{ord}a_1 =: t$. Da s minimal gewählt war, sind die Vektoren der Mengen $B(v_1),\ldots,B(v_l),B(a_1)$ linear abhängig (sonst ließe sich $B(v_1)\dot{\cup}\ldots\dot{\cup}B(v_l)\dot{\cup}B(a_1)$ durch einige der $a_2,\ldots,a_s$ zu einer Basis von V ergänzen). Dann gilt

$$\mathcal{O} = (\sum_{j=1}^{l}\sum_{i=0}^{t_j-1} k_{ij}\varphi^i(v_j)) + k_0a_1 + \ldots + k_{t-1}\varphi^{t-1}(a_1), \tag{26}$$

und nicht alle k_{ij}, k_ν sind 0. Wäre $k_0 = \ldots = k_{t-1} = 0$, so wäre $k_{ij} = 0$, da $B(v_1)\dot{\cup}\ldots\dot{\cup}B(v_l)$ linear unabhängig ist.
Sei also ν minimal mit $k_\nu \ne 0$ für $0 \le \nu \le t-1$, also $1 \le t-\nu$. Wenden wir $\varphi^{t-\nu}$ auf (26) an, so ergibt sich

$$\mathcal{O} = \sum_{j=1}^{l}\sum_{i=0}^{t_j-1} k_{ij}\varphi^{i+t-\nu}(v_j) + \underbrace{k_\nu\varphi^{\nu+t-\nu}(a_1) + \ldots + k_{t-1}\varphi^{t-1+t-\nu}(a_1)}_{=\mathcal{O}}.$$

Es folgt $k_{ij} = 0$ für alle $j = 1,\ldots,l$ und i mit $i+t-\nu < t_j$, also $i < t_j - t + \nu$. Wenden wir nun $\varphi^{t-\nu-1}$ auf (26) an, so ist weiterhin

$$\mathcal{O} = \sum_{j=1}^{l}\sum_{i=0}^{t_j-1} k_{ij}\varphi^{i+t-\nu-1}(v_j) + k_\nu\varphi^{\nu+t-\nu-1}(a_1) = \sum_{j=1}^{l} k_{(t_j-t+\nu)j}\varphi^{t_j-1}(v_j) + k_\nu\varphi^{t-1}(a_1).$$

Definiere

$$a_1' := \underbrace{\sum_{j=1}^{l} k_{(t_j-t+\nu)j}\varphi^{t_j-t}(v_j)}_{\in[B(v_1)\cup\ldots\cup B(v_l)]} + k_\nu a_1.$$

Wegen $t \le t_j$ ist a_1' wohldefiniert, und $B(v_1)\dot{\cup}\ldots\dot{\cup}B(v_l)\dot{\cup}\{a_1', a_2,\ldots,a_s\}$ ist eine Basis von V wegen Satz 1.17, Kapitel 3. Es gilt aber

$$\varphi^{t-1}(a_1') = \sum_{j=1}^{l} k_{(t_j-t+\nu)j}\varphi^{t_j-1}(v_j) + k_\nu\varphi^{t-1}(a_1) = \mathcal{O},$$

also $\mathrm{ord}a_1' < \mathrm{ord}a_1$. Das widerspricht der Minimalität von $\mathrm{ord}a_1 + \ldots + \mathrm{ord}a_s$. Damit ist die Annahme "$s \ge 1$" falsch, und $B(v_1)\dot{\cup}\ldots\dot{\cup}B(v_l)$ ist eine Basis von V.

Sei nun der Eigenwert c von φ nicht notwendig 0. Es gilt $\chi_{\varphi-c\cdot\mathrm{id}}(x) = \chi_\varphi(x+c)$, und damit hat $\varphi - c\cdot\mathrm{id}$ den n-fachen Eigenwert 0. Es gibt also eine Basis, bezüglich der $\varphi - c\cdot\mathrm{id}$ die Darstellung

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & 0 \\ 0 & & \ddots & \\ & & & A_k \end{pmatrix} \quad \text{mit} \quad A_i = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$$

hat. Da $c \cdot E_n$ die Darstellung von $c \cdot \mathrm{id}$ ist, hat $\varphi = \varphi - c \cdot \mathrm{id} + c \cdot \mathrm{id}$ die Darstellung

$$\begin{pmatrix} A'_1 & & & 0 \\ & A'_2 & & \\ & & \ddots & \\ 0 & & & A'_k \end{pmatrix} \quad \text{mit} \quad A'_i = \begin{pmatrix} c & & & & 0 \\ 1 & c & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & c \end{pmatrix}.$$

□

Beweis des Satzes über die Jordansche Normalform.
Wir beweisen den Satz durch vollständige Induktion nach $\dim_K V$, wobei der Induktionsanfang $\dim_K V = 1$ offensichtlich ist. Sei nun c ein Eigenwert von φ. Dann ist

$$\mathrm{Kern}(\varphi - c \cdot \mathrm{id}) \subseteq \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^2 \subseteq \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^3 \subseteq \dots$$

eine aufsteigende Kette von Unterräumen. Da $\dim_K V < \infty$, gibt es ein $m \in \mathbb{N}$ mit

$$\mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m = \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^{m+r} \quad \text{für alle } r \in \mathbb{N}.$$

Aus Satz 1.10 in Kapitel 4 folgt

$$\dim_K \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m + \dim_K \mathrm{Bild}(\varphi - c \cdot \mathrm{id})^m = \dim_K V.$$

Behauptung: $\mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m \cap \mathrm{Bild}(\varphi - c \cdot \mathrm{id})^m = \{\mathcal{O}\}$.
Sei $v \in \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m \cap \mathrm{Bild}(\varphi - c \cdot \mathrm{id})^m$, also $v = (\varphi - c \cdot \mathrm{id})^m(w)$ für $w \in V$ und $\mathcal{O} = (\varphi - c \cdot \mathrm{id})^{2m}(w)$. Dann gilt $w \in \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^{2m} = \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m$, das heißt $v = \mathcal{O}$.
Es folgt also $V = \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m \oplus \mathrm{Bild}(\varphi - c \cdot \mathrm{id})^m = U \oplus W$, wobei $U = \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m$ und $W = \mathrm{Bild}(\varphi - c \cdot \mathrm{id})^m$.
$\varphi(U) \subseteq U$: Für alle $v \in U$ folgt $(\varphi - c \cdot \mathrm{id})^m(v) = \mathcal{O}$ und damit $\mathcal{O} = \varphi(\varphi - c \cdot \mathrm{id})^m(v) = (\varphi - c \cdot \mathrm{id})^m \varphi(v)$. Also gilt $\varphi(v) \in \mathrm{Kern}(\varphi - c \cdot \mathrm{id})^m = U$ für alle $v \in U$.
$\varphi(W) \subseteq W$: Für $v \in W$ gilt $v = (\varphi - c \cdot \mathrm{id})^m(w)$ für ein $w \in V$. Damit ist $\varphi(v) = \varphi(\varphi - c \cdot \mathrm{id})^m(w) = (\varphi - c \cdot \mathrm{id})^m(\varphi(w)) \in W$.
Wir definieren nun die Abbildungen

$$\begin{aligned} \phi : \quad U &\longrightarrow U\,; \quad u \longmapsto \varphi(u) \quad \text{und} \\ \psi : \quad W &\longrightarrow W\,; \quad w \longmapsto \varphi(w)\,. \end{aligned}$$

Sei B_U eine Basis von U und B_W eine Basis von W. Dann ist $B_U \cup B_W$ eine Basis von V, da $V = U \oplus W$. Bezeichnen wir mit A_ϕ die Matrixdarstellung von ϕ bezüglich B_U und mit A_ψ diejenige von ψ bezüglich B_W, so ist $\begin{pmatrix} A_\phi & 0 \\ 0 & A_\psi \end{pmatrix}$ die Matrixdarstellung von φ bezüglich $B_U \cup B_W$ und $\chi_\varphi(x) = \chi_\phi(x) \cdot \chi_\psi(x)$. Die Polynome $\chi_\phi(x)$ und $\chi_\psi(x)$ zerfallen über K vollständig in Linearfaktoren.
$\dim W < \dim V$: Zu zeigen ist $U \neq \{\mathcal{O}\}$.
Sei $v \in V$ Eigenvektor von φ zum Eigenwert c. Dann gilt $(\varphi - c \cdot \mathrm{id})(v) = \varphi(v) - c \cdot v = \mathcal{O}$, also $(\varphi - c \cdot \mathrm{id})^m(v) = \mathcal{O}$, das heißt $v \in U$. Wegen $v \neq \mathcal{O}$ folgt $U \neq \{\mathcal{O}\}$.

Nach Induktionsvoraussetzung kann nun B_W so gewählt werden, daßA_ψ die gewünschte Form hat. Zu zeigen bleibt, daß auch B_U so gewählt werden kann, daß A_ϕ die gewünschte Form hat. Wegen des Lemmas reicht es dazu aus nachzuweisen, daß alle Eigenwerte von ϕ gleich sind:
Sei $c' \in K$ ein Eigenwert von ϕ und $v \in U$ ein Eigenvektor von ϕ zum Eigenwert c'. Dann gilt

$$\mathcal{O} = (\varphi - c \cdot \mathrm{id})^m(v) = (\varphi - c \cdot \mathrm{id})^{m-1}((c' - c)(v)) = \ldots = (c' - c)^m v,$$

also $c' = c$.

□

Ergänzung. Die Darstellung von φ durch eine Matrix A mit Jordanblöcken ist bis auf die Reihenfolge der Jordanblöcke eindeutig. Wir zeigen dieses nur für die Jordanblöcke zum Eigenwert $c = 0$. Für $c \neq 0$ betrachte man $A - cE_n$. Dann ist die Anzahl der Jordanblöcke von A der Größe n_i zum Eigenwert c gleich der Anzahl der Jordanblöcke von $A - cE_n$ der Größe n_i zum Eigenwert 0. Die Behauptung folgt nun unmittelbar aus dem folgenden

Satz 4.4 *Die Zahl* $\mathrm{Rg}A^{\nu-1} - 2\mathrm{Rg}A^\nu + \mathrm{Rg}A^{\nu+1}$ *ist gleich der Anzahl der Jordanblöcke der Größe* ν *zum Eigenwert* 0.

Beweis. Für

$$A = \begin{pmatrix} A_1 & & & & & 0 \\ & \ddots & & & & \\ & & A_l & & & \\ 0 & & & B_1 & & \\ & & & & \ddots & \\ & & & & & B_m \end{pmatrix} \quad \text{gilt} \quad A^\nu = \begin{pmatrix} A_1^\nu & & & & & 0 \\ & \ddots & & & & \\ & & A_l^\nu & & & \\ 0 & & & B_1^\nu & & \\ & & & & \ddots & \\ & & & & & B_m^\nu \end{pmatrix},$$

wobei A_i die Jordanblöcke zu einem Eigenwert $\neq 0$ und B_i die Jordanblöcke zum Eigenwert 0 bezeichnen. Da A_i regulär ist, folgt $\mathrm{Rg}A_i = \mathrm{Rg}A_i^r$ für alle $r \in \mathbb{N}$. Weiterhin gilt

$$B_i = \begin{pmatrix} 0 & \ldots & & & 0 \\ 1 & 0 & \ldots & & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{pmatrix}, \; B_i^2 = \begin{pmatrix} 0 & \ldots & & & & 0 \\ 0 & 0 & \ldots & & & 0 \\ 1 & 0 & 0 & \ldots & & 0 \\ 0 & 1 & 0 & 0 & \ldots & 0 \\ & & \ddots & \ddots & \ddots & \\ 0 & & & 1 & 0 & 0 \end{pmatrix} \text{ und so weiter.}$$

B_i habe n_i Zeilen und Spalten. Dann ist $\mathrm{Rg}B_i^\nu = \begin{cases} n_i - \nu & 0 \leq \nu \leq n_i. \\ & \text{für} \\ 0 & n_i \leq \nu. \end{cases}$

Speziell gilt also

$$\boxed{\mathrm{Rg}B_i^\nu = 0 \quad \Longleftrightarrow \quad n_i \leq \nu.}$$

Es folgt

$$\mathrm{Rg}A^\nu = \mathrm{Rg}A_1 + \ldots + \mathrm{Rg}A_l + \mathrm{Rg}B_1^\nu + \ldots + \mathrm{Rg}B_m^\nu$$

und damit

$$\begin{aligned}
\mathrm{Rg}A^{\nu-1} - \mathrm{Rg}A^\nu &= \sum_{i=1}^{m}(\mathrm{Rg}B_i^{\nu-1} - \mathrm{Rg}B_i^\nu) \\
&= |\{i \mid \mathrm{Rg}B_i^{\nu-1} \neq \mathrm{Rg}B_i^\nu\}| \\
&= |\{i \mid \mathrm{Rg}B_i^{\nu-1} \neq 0\}| \\
&= |\{i \mid n_i > \nu - 1\}| \\
&= |\{i \mid n_i \geq \nu\}|.
\end{aligned}$$

Also gilt $\mathrm{Rg}A^\nu - \mathrm{Rg}A^{\nu+1} = |\{i \mid n_i > \nu\}|$ und

$$\mathrm{Rg}A^{\nu-1} - 2\mathrm{Rg}A^\nu + \mathrm{Rg}A^{\nu+1} = |\{i \mid n_i = \nu\}|\,.$$

$\square$

Beispiel. Für

$$A = \begin{pmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{pmatrix} \quad \text{ist} \quad A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & -1 & 1 \end{pmatrix} \quad \text{und} \quad A^3 = \mathcal{O}.$$

Damit ist $\mathrm{Rg}A^0 = 4, \mathrm{Rg}A = 2, \mathrm{Rg}A^2 = 1, \mathrm{Rg}A^3 = 0$, und A hat den 4-fachen Eigenwert 0. Die Anzahl der Blöcke der Größe 1 zum Eigenwert 0 ist also $4 - 2 \cdot 2 + 1 = 1$, die Anzahl der Blöcke der Größe 2 zum Eigenwert 0 ist $2 - 2 \cdot 1 + 0 = 0$, und die Anzahl der Blöcke der Größe 3 zum Eigenwert 0 ist $1 - 2 \cdot 0 + 0 = 1$. Als Jordansche Normalform von A erhalten wir somit

$$\left(\begin{array}{c|ccc} \boxed{0} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

5. Aufgaben

A 5.1 Überprüfen Sie, ob die reelle Matrix

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

diagonalisierbar ist.

A 5.2 Berechnen Sie alle $a \in \mathbb{R}$, für die die Matrix

$$A = \begin{pmatrix} 3 & a & 0 \\ a & 4 & a \\ 0 & a & 0 \end{pmatrix}$$

den Eigenwert 2 besitzt. Für welche a ist dann A auch diagonalisierbar? Geben Sie gegebenenfalls $S \in \mathrm{GL}(3;\mathbb{R})$ so an, daß $S^{-1}AS$ eine Diagonalmatrix ist.

A 5.3 V sei ein endlich-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$ mit $\varphi^2 = \varphi$. Zeigen Sie, daß $V = \mathrm{Kern}\varphi \oplus \mathrm{Bild}\varphi$ gilt und daß φ diagonalisierbar ist. Berechnen Sie die Eigenwerte von φ und bringen Sie die Eigenräume mit $\mathrm{Kern}\varphi$ sowie $\mathrm{Bild}\varphi$ in Zusammenhang.

A 5.4 Gegeben ist $A = (a_{ij}) \in K_{n,n}$ mit $n \geq 2$, so daß $a_{ii} = 0$ für $i = 1, \ldots, n$ und $a_{ij} = 1$ sonst. Für welche n ist A diagonalisierbar? (Beachten Sie die Charakteristik von K.) Geben Sie gegebenenfalls ein $S \in \mathrm{GL}(n;K)$ so an, daß $S^{-1}AS$ eine Diagonalmatrix ist.

A 5.5 $V \neq \{\mathcal{O}\}$ sei ein endlich-dimensionaler K-Vektorraum und $\varphi \in \mathrm{End}(V)$. Das charakteristische Polynom von φ habe die Zerlegung $\chi_\varphi(x) = f_1(x) \cdot \ldots \cdot f_l(x)$ in Polynome $f_i(x) \in K[x]$ und es sei $m := \max\{\mathrm{grad} f_i(x) \mid i = 1, \ldots, l\} \geq 1$. Zeigen Sie, daß es einen Unterraum $U \neq \{\mathcal{O}\}$ von V mit $\varphi(U) \subseteq U$ und $\dim_K U \leq m$ gibt. (Hinweis: Benutzen Sie Korollar 3.2)

A 5.6 Berechnen Sie für $a_0, \ldots, a_{n-1} \in K$ das charakteristische Polynom von $A_n \in K_{n,n}$ mit

$$A_n = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & & \vdots & -a_2 \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}.$$

A 5.7 Es sei $K = \mathbb{Q}$. Zeigen Sie $\chi_A(x) = \mu_A(x)$ mit Hilfe von Korollar 3.4 für

$$A = \begin{pmatrix} 0 & 0 & -1 & -1 \\ 5 & -4 & -4 & -4 \\ 1 & 0 & -1 & -1 \\ -7 & 5 & 3 & 4 \end{pmatrix}.$$

A 5.8 Gegeben ist $A = (a_{ij}) \in K_{n,n}$ mit $n \geq 2$ und $a_{ii} = 0$ für $i = 1, \ldots, n$ sowie $a_{ij} = 1$ für $i \neq j$ sonst. Zeigen Sie mit Hilfe von Satz 2.13, daß A genau dann diagonalisierbar ist, wenn n nicht von $\mathrm{char}K$ geteilt wird (vgl. Aufgabe 5.4).

A 5.9 Geben Sie für die reelle Matrix A die Jordansche Normalform an, wobei

$$A = \begin{pmatrix} -4 & 8 & 14 & -6 & -3 \\ 1 & -2 & -3 & 0 & 1 \\ -3 & 7 & 12 & -4 & -3 \\ -3 & 7 & 13 & -5 & -3 \\ -1 & 4 & 6 & -2 & -2 \end{pmatrix}.$$

KAPITEL 6

Euklidische Vektorräume

1. Das Skalarprodukt

Definition 1.1 *Für einen K-Vektorraum V heißt eine Abbildung*

$$\beta : V \times V \longrightarrow K\,; \quad (v,w) \longmapsto \beta(v,w)$$

Bilinearform, wenn für alle $v, v', w, w' \in V$ und $k \in K$ gilt

i) $\beta(v+v', w) = \beta(v,w) + \beta(v', w)$ *und* $\beta(k \cdot v, w) = k \cdot \beta(v,w)$.

ii) $\beta(v, w+w') = \beta(v,w) + \beta(v, w')$ *und* $\beta(v, k \cdot w) = k \cdot \beta(v,w)$.

Bemerkung.

1. Eine Bilinearform ist also linear in beiden Argumenten.
2. Ist β eine Bilinearform, so gilt $\beta(\mathcal{O}, v) = \beta(v, \mathcal{O}) = 0$ für alle $v \in V$.
3. β heißt symmetrisch, wenn für alle $v, w \in V$ gilt $\beta(v,w) = \beta(w,v)$.

Beispiel.

1. Ist $\beta(v,w) := 0$ für alle $v, w \in V$, so ist β eine symmetrische Bilinearform. β heißt dann trivial.
2. Für $V = K^n, A \in K_{n,n}$ und $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in K^n$ wird durch

$$\beta(x,y) := (x_1, \ldots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

eine Bilinearform definiert, die i.a. nicht symmetrisch ist. Ist speziell $A = E_n$, dann gilt $\beta(x,y) = x_1 y_1 + \ldots + x_n y_n$, und β ist symmetrisch.

Satz 1.2 *Sei V ein n-dimensionaler K-Vektorraum, $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $b_{ij} \in K$, $i, j = 1, \ldots, n$. Dann gibt es genau eine Bilinearform $\beta : V \times V \longrightarrow K$ mit $\beta(v_i, v_j) = b_{ij}$. Sie ist genau dann symmetrisch, wenn $b_{ij} = b_{ji}$ für alle $i, , j = 1, \ldots, n$.*

Beweis. Existenz: Da B eine Basis von V ist, gibt es für alle $x, y \in V$ eine eindeutige Darstellung

$$x = x_1 \cdot v_1 + \ldots + x_n \cdot v_n; \quad y = y_1 \cdot v_1 + \ldots + y_n \cdot v_n$$

mit $x_i, y_i \in K$. Die Abbildung $\beta : V \times V \longrightarrow K$ mit

$$\beta(x, y) := (x_1, \ldots, x_n) \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \ldots & b_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K$$

ist dann eine Bilinearform, für die gilt

$$\begin{aligned} \beta(v_i, v_j) &= (0, \ldots, 1, \ldots, 0) \begin{pmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \ldots & b_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \\ &= (b_{i1}, \ldots, b_{in}) \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = b_{ij} \,. \end{aligned}$$

Eindeutigkeit: x, x_i, y und y_i seien wie oben. Wegen der Bilinearität gilt für jedes β mit $\beta(v_i, v_j) = b_{ij}$

$$\begin{aligned} \beta(x, y) &= \beta(\sum_i x_i v_i, \sum_j y_j v_j) = \sum_i \sum_j x_i y_j \beta(v_i, v_j) \\ &= \sum_i \sum_j x_i y_j b_{ij} \,, \end{aligned}$$

das heißt, β ist eindeutig bestimmt.
Symmetrie: Ist β symmetrisch, dann ist $b_{ij} = \beta(v_i, v_j) = \beta(v_j, v_i) = b_{ji}$. Ist andererseits $b_{ij} = b_{ji}$, so folgt $\beta(v_i, v_j) = \beta(v_j, v_i)$ für alle $i, j = 1, \ldots, n$ und damit

$$\begin{aligned} \beta(x, y) = \sum_i \sum_j x_i y_j \beta(v_i, v_j) &= \sum_i \sum_j y_i x_j \beta(v_j, v_i) \\ &= \sum_i \sum_j y_i x_j \beta(v_i, v_j) = \beta(y, x). \end{aligned}$$

□

Definition 1.3 *Eine Matrix $A \in K_{n,n}$ heißt symmetrisch, wenn $A = A^t$, also $a_{ij} = a_{ji}$ für alle $i, j = 1, \ldots, n$ gilt.*

Wegen Satz 1.2 ist bei vorgegebener Basis $B = \{v_1, \ldots, v_n\}$ von V die Zuordnung

$$\beta \longmapsto \begin{pmatrix} \beta(v_1, v_1) & \ldots & \beta(v_1, v_n) \\ \vdots & & \vdots \\ \beta(v_n, v_1) & \ldots & \beta(v_n, v_n) \end{pmatrix} =: A_B(\beta)$$

eine Bijektion zwischen der Menge der Bilinearformen von V und $K_{n,n}$. Dabei heißt $A_B(\beta)$ Matrixdarstellung von β bezüglich B, und β ist genau dann symmetrisch, wenn $A_B(\beta)$ symmetrisch ist.

Deutung. Sind $(x_1, \ldots, x_n)$ bzw. $(y_1, \ldots, y_n)$ die Koordinatenvektoren von $x, y \in V$ bezüglich B, so kann $\beta(x, y)$ direkt mit Hilfe der Koordinaten und $A_B(\beta)$ berechnet werden:

$$\beta(x, y) = \sum_j \left(\sum_i x_i \beta(v_i, v_j) \right) y_j = (x_1, \ldots, x_n) A_B(\beta) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Definition 1.4 *Ist V ein reeller Vektorraum und $\beta : V \times V \longrightarrow \mathbb{R}$ symmetrisch und bilinear, dann heißt β Skalarprodukt, wenn $\beta(x, x) > 0$ für alle $x \in V \setminus \{\mathcal{O}\}$ gilt. Ist auf V ein Skalarprodukt definiert, so heißt V euklidischer Vektorraum.*

Bemerkung.

1. Nach Definition ist eine reelle, symmetrische Bilinearform β ein Skalarprodukt, wenn β positiv definit ist.

2. Ist β ein Skalarprodukt, so schreibt man auch $\langle x, y \rangle$ statt $\beta(x, y)$.

Beispiel.

1. Für $V = \mathbb{R}^n$ ist die Bilinearform $\langle \cdot , \cdot \rangle$ mit

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle := x_1 y_1 + \ldots + x_n y_n$$

ein Skalarprodukt, denn es gilt

$$\langle (x_1, \ldots, x_n), (x_1, \ldots, x_n) \rangle = x_1^2 + \ldots + x_n^2 > 0$$

für $x \neq \mathcal{O}$. Man nennt $\langle \cdot , \cdot \rangle$ das kanonische Skalarprodukt des $\mathbb{R}^n$.

2. Sei für $V = \mathbb{R}^2$ die Bilinearform $\langle \cdot , \cdot \rangle$ durch

$$\langle (x_1, x_2), (y_1, y_2) \rangle := (x_1, x_2) \begin{pmatrix} 4 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

gegeben. Dann ist $\langle (x_1, x_2), (x_1, x_2) \rangle = (2x_1 - x_2)^2 + 2x_2^2$, also $\langle \cdot , \cdot \rangle$ ein Skalarprodukt.

3. Ist für $V = \mathbb{R}^2$ die Bilinearform $\langle\,\cdot\,,\,\cdot\,\rangle$ durch
$$\langle (x_1, x_2), (y_1, y_2) \rangle = (x_1, x_2) \begin{pmatrix} 4 & -2 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
gegeben, so ist $\langle\,\cdot\,,\,\cdot\,\rangle$ wegen $\langle (0,1), (0,1) \rangle = -1$ kein Skalarprodukt.

4. Ist $V = C^0([a,b]) = \{f : [a,b] \to \mathbb{R} \mid f \ \text{ stetig}\}$, dann wird durch
$$\langle f, g \rangle := \int_a^b f \cdot g \, dx \tag{27}$$
eine symmetrische Bilinearform definiert. Es gilt
$$\langle f, f \rangle = \int_a^b f^2 \, dx \geq 0\,.$$
Da jedes $f \in V$ stetig ist, folgt weiterhin, daß $\langle f, f \rangle = 0$ genau dann gilt, wenn $f = \mathcal{O}$. Durch (27) wird also auf $C^0([a,b])$ ein Skalarprodukt definiert.

Definition 1.5 *Sei V ein euklidischer Vektorraum und $v \in V$. Dann heißt $||v|| := {}_+\sqrt{\langle v, v \rangle}$ Norm oder Länge von v. Weiterhin heißt v normiert, wenn $||v|| = 1$.*

Satz 1.6 *Ist V ein euklidischer Vektorraum, dann gilt für alle $v, w \in V$, $k \in \mathbb{R}$*

i) $|\langle v, w \rangle| \leq ||v|| \cdot ||w||$ *(Cauchy-Schwarzsche Ungleichung).*
Das Gleichheitszeichen gilt genau dann, wenn v und w linear abhängig sind.

ii) $||v|| \geq 0$ *und* $(||v|| = 0 \Leftrightarrow v = \mathcal{O})$.

iii) $||k \cdot v|| = |k| \cdot ||v||$.

iv) $||v + w|| \leq ||v|| + ||w||$ *(Dreiecksungleichung).*

v) $|\; ||v|| - ||w|| \;| \leq ||v - w||$.

Beweis.

i) Sei $w \neq \mathcal{O}$. Dann gilt mit der Definition $k := \frac{\langle v,w \rangle}{\langle w,w \rangle} \in \mathbb{R}$ wegen der Bilinearität und Symmetrie des Skalarproduktes
$$\begin{aligned} 0 &\leq \langle v - k \cdot w, v - k \cdot w \rangle = \langle v, v \rangle - k \cdot \langle v, w \rangle - k \cdot \langle w, v \rangle + k^2 \cdot \langle w, w \rangle \\ &= \langle v, v \rangle - 2\frac{\langle v, w \rangle^2}{\langle w, w \rangle} + \frac{\langle v, w \rangle^2}{\langle w, w \rangle} = \langle v, v \rangle - \frac{\langle v, w \rangle^2}{\langle w, w \rangle}\,, \end{aligned}$$
also
$$\langle v, w \rangle^2 \leq \langle v, v \rangle \cdot \langle w, w \rangle\,. \tag{28}$$
Gilt in (28) sogar Gleichheit, so folgt $v - k \cdot w = \mathcal{O}$ und damit $v = k \cdot w$, das heißt, v und w sind linear abhängig. Sind andererseits v und w linear abhängig, so gilt $v = c \cdot w$ für ein $c \in \mathbb{R}$ und damit Gleichheit in (28).

ii) Die erste Behauptung folgt aus der Definition der Norm, die zweite aus der positiven Definitheit des Skalarproduktes.

iii) $||k \cdot v|| = {}_{+}\sqrt{\langle k \cdot v, k \cdot v\rangle} = {}_{+}\sqrt{k^2 \langle v, v\rangle} = |k| \cdot ||v||$.

iv)
$$\begin{aligned} \langle v+w, v+w\rangle &= |\langle v,v\rangle + 2\langle v,w\rangle + \langle w,w\rangle| \\ &\leq ||v||^2 + 2|\langle v,w\rangle| + ||w||^2 \\ &\leq ||v||^2 + 2||v||\,||w|| + ||w||^2 = (||v|| + ||w||)^2. \end{aligned}$$

Dabei gilt die letzte Ungleichung wegen i) (Cauchy-Schwarzsche Ungleichung).

v) Es gilt wegen der Dreiecksungleichung $||v|| = ||v-w+w|| \leq ||v-w|| + ||w||$ und damit $||v|| - ||w|| \leq ||v-w||$. Analog ist $||w|| - ||v|| \leq ||w-v|| = ||v-w||$ und damit
$$-||v-w|| \leq ||v|| - ||w|| \leq ||v-w||\,.$$

□

Ist V ein euklidischer Vektorraum, dann gilt für alle $v, w \neq \mathcal{O}$ aus V wegen der Cauchy-Schwarzschen Ungleichung $|\langle v,w\rangle| \leq ||v|| \cdot ||w||$, also $\left|\frac{\langle v,w\rangle}{||v|| \cdot ||w||}\right| \leq 1$, das heißt
$$-1 \leq \frac{\langle v,w\rangle}{||v|| \cdot ||w||} \leq 1.$$

Bekanntlich ist die in der Analysis definierte Cosinusfunktion $\cos : \mathbb{R} \longrightarrow [-1,1]$ mit
$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$
auf dem Intervall $[0,\pi]$ bijektiv. Sind also $v, w \neq \mathcal{O}$ aus V, so gibt es genau ein $\alpha \in [0,\pi]$ mit $\cos\alpha = \frac{\langle v,w\rangle}{||v|| \cdot ||w||}$. Man nennt α Winkel zwischen v und w. Die Vektoren v und w heißen orthogonal (geschrieben $v \perp w$), wenn $\langle v,w\rangle = 0$, das heißt, der Winkel zwischen v und w ist $\frac{\pi}{2}$. Sind v und w orthogonal, so auch kv und $k'w$ mit $k, k' \in \mathbb{R}$.

Beispiel. Ist $V = C^0([-\pi,\pi])$, dann definieren wir für $f, g \in V$ wie im Beispiel 4 vor Definition 1.5
$$\langle f, g\rangle := \int_{-\pi}^{\pi} f \cdot g \, dx\,.$$

Für die Funktionen $f(x) = \sin x$ und $g(x) = \cos x$ gilt dann zum Beispiel
$$\langle f, g\rangle = \int_{-\pi}^{\pi} \sin x \cos x \, dx = 0\,.$$

Nach obiger Sprechweise sind damit sin und cos orthogonal.

Satz 1.7 *Sei V ein euklidischer Vektorraum und seien $v_1, \ldots, v_n \in V \setminus \{\mathcal{O}\}$ paarweise orthogonal. Dann sind $v_1, \ldots, v_n$ linear unabhängig.*

Beweis. Sind $k_1, \ldots, k_n \in \mathbb{R}$ mit $v := k_1 v_1 + \ldots + k_n v_n = \mathcal{O}$, dann gilt wegen der paarweisen Orthogonalität der Vektoren $v_1, \ldots, v_n$:

$$0 = \langle v_i, v \rangle = \langle v_i, k_1 v_1 + \ldots + k_n v_n \rangle = k_1 \langle v_i, v_1 \rangle + \ldots + k_n \langle v_i, v_n \rangle = k_i \langle v_i, v_i \rangle,$$

also $k_i = 0$, da $\langle v_i, v_i \rangle \neq 0$.

□

Definition 1.8 *Ist V ein euklidischer Vektorraum, dann heißt eine Basis B von V Orthonormalbasis von V, wenn gilt:*

i) Die Vektoren aus B sind paarweise orthogonal.

ii) Die Vektoren aus B sind normiert.

Beispiel. Bezüglich des kanonischen Skalarproduktes

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle := x_1 y_1 + \ldots + x_n y_n$$

ist die kanonische Basis des $\mathbb{R}^n$ eine Orthonormalbasis.

Satz 1.9 *Sei V ein euklidischer Vektorraum und $B = \{b_1, \ldots, b_n\}$ eine Orthonormalbasis von V. Sind $x, y \in V$ und $x_1, \ldots, x_n$ bzw. $y_1, \ldots, y_n$ die Koordinaten von x bzw. y bezüglich B, so gilt*

i) $\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n$.

ii) $x_i = \langle x, b_i \rangle$ für $i = 1, \ldots, n$.

Beweis. Da B eine Orthonormalbasis ist, gilt $\langle b_i, b_j \rangle = 0$ für $i \neq j$ und $\langle b_i, b_i \rangle = 1$.

i)

$$\begin{aligned} \langle x, y \rangle &= (x_1, \ldots, x_n) \begin{pmatrix} \langle b_1, b_1 \rangle & \ldots & \langle b_1, b_n \rangle \\ \vdots & & \vdots \\ \langle b_n, b_1 \rangle & \ldots & \langle b_n, b_n \rangle \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= (x_1, \ldots, x_n) E_n \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \ldots + x_n y_n \,. \end{aligned}$$

ii) Mit $x = x_1 b_1 + \ldots + x_n b_n$ folgt

$$\langle x, b_i \rangle = \langle x_1 b_1 + \ldots + x_n b_n, b_i \rangle = x_1 \langle b_1, b_i \rangle + \ldots + x_n \langle b_n, b_i \rangle = x_i \langle b_i, b_i \rangle = x_i \,.$$

□

Beispiel. Im $\mathbb{R}^3$ ist bezüglich des kanonischen Skalarproduktes die Menge $B = \{b_1, b_2, b_3\}$ mit $b_1 = \frac{1}{\sqrt{6}}(1,-1,2)$, $b_2 = \frac{1}{\sqrt{2}}(1,1,0)$ und $b_3 = \frac{1}{\sqrt{3}}(1,-1,-1)$ eine Orthonormalbasis. Der Vektor $v = (1,1,1)$ hat die Darstellung $v = \frac{2}{\sqrt{6}}b_1 + \frac{2}{\sqrt{2}}b_2 - \frac{1}{\sqrt{3}}b_3$.

Satz 1.10 *Ist V ein euklidischer Vektorraum und $B = \{a_i \mid i \in \mathbb{N}\}$ eine abzählbare Basis von V, dann existiert eine Orthonormalbasis $\{e_i \mid i \in \mathbb{N}\}$ von V, so daß für alle $n \in \mathbb{N}$ gilt* $[a_1, \ldots, a_n] = [e_1, \ldots, e_n]$.

Beweis. Wir definieren die e_i rekursiv.
$n = 1$: Da B eine Basis ist, gilt $a_1 \neq \mathcal{O}$, und wir konstruieren hieraus den ersten Vektor der Orthonormalbasis durch Normalisierung, also

$$\boxed{e_1 := \frac{1}{||a_1||}a_1 \ .}$$

Offenbar gilt $||e_1|| = 1$ und $[e_1] = [a_1]$.
$n \to n+1$: Es sei schon durch $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis von $[a_1, \ldots, a_n]$ gegeben. Sei $a_{n+1} \notin [a_1, \ldots, a_n]$. Mit der Definition

$$\boxed{e'_{n+1} := a_{n+1} - \sum_{i=1}^{n} \langle a_{n+1}, e_i\rangle \cdot e_i} \tag{29}$$

gilt für $j = 1, \ldots, n$

$$\langle e'_{n+1}, e_j\rangle = \langle a_{n+1}, e_j\rangle - \sum_{i=1}^{n} \langle a_{n+1}, e_i\rangle\langle e_i, e_j\rangle = \langle a_{n+1}, e_j\rangle - \langle a_{n+1}, e_j\rangle = 0\,.$$

Damit sind $e_1, \ldots, e_n, e'_{n+1}$ paarweise orthogonal. Wäre $e'_{n+1} = \mathcal{O}$, so wäre wegen (29) dann $a_{n+1} \in [e_1, \ldots, e_n] = [a_1, \ldots, a_n]$. Also gilt $e'_{n+1} \neq \mathcal{O}$ und $\|e'_{n+1}\| \neq 0$. Wir definieren

$$\boxed{e_{n+1} := \frac{1}{||e'_{n+1}||} \cdot e'_{n+1}\,.}$$

$e_1, \ldots, e_{n+1}$ sind somit paarweise orthogonal und normiert.
Wegen (29) ist $e_{n+1} \in [a_{n+1}, e_1, \ldots, e_n] = [a_{n+1}, a_1, \ldots, a_n]$ und damit $[e_1, \ldots, e_{n+1}] \subseteq [a_1, \ldots, a_{n+1}]$. Ebenfalls wegen (29) folgt $[a_1, \ldots, a_{n+1}] \subseteq [e_1, \ldots, e_{n+1}]$, also $[e_1, \ldots, e_{n+1}] = [a_1, \ldots, a_{n+1}]$. Aus Satz 1.7 folgt die lineare Unabhängigkeit von $\{e_i \mid i \in \mathbb{N}\}$. Da nach Voraussetzung und Konstruktion

$$V = \bigcup_{n=1}^{\infty} [a_1, \ldots, a_n] = \bigcup_{n=1}^{\infty} [e_1, \ldots, e_n]$$

gilt, ist $\{e_i \mid i \in \mathbb{N}\}$ eine Orthonormalbasis von V.

$\square$

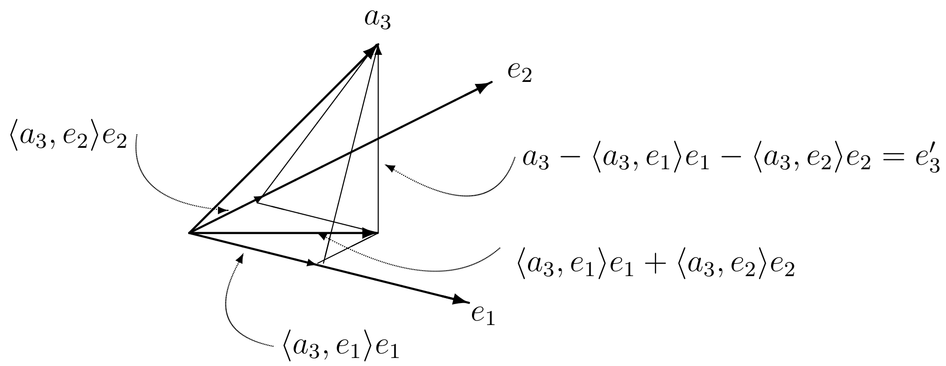
Bemerkung.

1. Das in diesem Beweis entwickelte Konstruktionsverfahren für die Vektoren e_i heißt Orthonormalisierungsverfahren von E. Schmidt.

2. Sind $e_1, \ldots, e_n$ normiert und paarweise orthogonal, so kann $\{e_1, \ldots, e_n\}$ mit dem Orthonormalisierungsverfahren zu einer Orthonormalbasis ergänzt werden.

3. Ist V endlich-dimensional, so hat V eine Orthonormalbasis, die mit dem obigen Verfahren konstruiert werden kann. Dabei braucht man keine Basis $\{a_i\}$ zu kennen. Im Rekursionsschritt muß lediglich a_{n+1} mit $a_{n+1} \notin [e_1, \ldots, e_n]$ gewählt werden.

4. Geometrische Deutung der Konstruktion zum Beispiel von e_3.



Beispiel. Im $\mathbb{R}^4$ soll bezüglich des kanonischen Skalarproduktes eine Orthonormalbasis von $[a_1, a_2, a_3]$ berechnet werden, wobei $a_1 = (1,0,2,2), a_2 = (5,0,1,1)$ und $a_3 = (6,1,6,0)$. Der erste Vektor e_1 ergibt sich durch die Normierung von a_1, also

$$e_1 = \frac{1}{3}(1,0,2,2)\,.$$

Um e_2 zu berechnen, ziehen wir zunächst von a_2 die Projektion auf $[e_1]$ ab. Der verbleibende Anteil ist orthogonal zu $[e_1]$:

$$e_2' = (5,0,1,1) - \frac{1}{3} \cdot 9 \cdot \frac{1}{3}(1,0,2,2) = (4,0,-1,-1).$$

Normieren wir diesen Vektor, so ergibt sich

$$e_2 = \frac{1}{\sqrt{18}}(4,0,-1,-1)\,.$$

Zur Berechnung von e_3 müssen zunächst von a_3 sowohl die Anteile in e_1- als auch in e_2-Richtung abgezogen werden, also

$$e_3' = (6,1,6,0) - 2(1,0,2,2) - (4,0,-1,-1) = (0,1,3,-3).$$

Nach der Normierung ergibt sich

$$e_3 = \frac{1}{\sqrt{19}}(0,1,3,-3).$$

Als nächstes untersuchen wir den Übergang zwischen verschiedenen Orthonormalbasen. Dazu sei V ein endlich-dimensionaler euklidischer Vektorraum mit der Orthonormalbasis $\{e_1, \ldots, e_n\} = B$. Sei $\{v_1, \ldots, v_n\} = B'$ eine zweite Basis von V und

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

die Transformationsmatrix von B' nach B, also

$$v_j = a_{1j} \cdot e_1 + \ldots + a_{nj} \cdot e_n\,, \quad j = 1, \ldots, n\,.$$

$a_{1j}, \ldots, a_{nj}$ sind also die Koordinaten von v_j bezüglich $\{e_1, \ldots, e_n\}$, das heißt

$$\langle v_i, v_j \rangle = (a_{1i}, \ldots, a_{ni}) \underbrace{\begin{pmatrix} \langle e_1, e_1 \rangle & \ldots & \langle e_1, e_n \rangle \\ \vdots & & \vdots \\ \langle e_n, e_1 \rangle & \ldots & \langle e_n, e_n \rangle \end{pmatrix}}_{E_n} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = (a_{1i}, \ldots, a_{ni}) \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}\,,$$

also

$$\begin{pmatrix} \langle v_1, v_1 \rangle & \ldots & \langle v_1, v_n \rangle \\ \vdots & & \vdots \\ \langle v_n, v_1 \rangle & \ldots & \langle v_n, v_n \rangle \end{pmatrix} = A^t A.$$

Die Basis $\{v_1, \ldots, v_n\}$ ist damit genau dann eine Orthonormalbasis, wenn $A^t A = E_n$.

Definition 1.11 *Eine Matrix $A \in \mathbb{R}_{n,n}$ heißt orthogonal, wenn $A^t A = E_n$.*

Bemerkung. Seien $A, B \in \mathbb{R}_{n,n}$.

1. Ist A orthogonal, dann ist A auch regulär, das heißt $A \in \mathrm{GL}(n; \mathbb{R})$.
2. Ist A orthogonal, dann gilt $A^{-1} = A^t$.
3. $\mathrm{O}(n) := \{A \in \mathrm{GL}(n; \mathbb{R}) \mid A^{-1} = A^t\}$, $n \in \mathbb{N}$.
4. Sind $A, B \in \mathrm{O}(n)$, dann gilt $(A \cdot B)^t = B^t \cdot A^t = B^{-1} \cdot A^{-1} = (A \cdot B)^{-1}$ und damit $A \cdot B \in \mathrm{O}(n)$.
5. Für $A \in \mathrm{O}(n)$ ist $(A^{-1})^t = (A^t)^t = A = (A^{-1})^{-1}$ und damit $A^{-1} \in \mathrm{O}(n)$.
6. $\mathrm{O}(n)$ ist eine Untergruppe von $\mathrm{GL}(n; \mathbb{R})$ und heißt orthogonale Gruppe.

Obige Überlegungen liefern nun

Satz 1.12 *Sei V ein endlich-dimensionaler euklidischer Vektorraum, B eine Orthonormalbasis von V und B' eine Basis von V. Genau dann ist B' eine Orthonormalbasis, wenn die Transformationsmatrix des Basiswechsels orthogonal ist.*

Korollar 1.13 *Sei V ein endlich-dimensionaler euklidischer Vektorraum, $B = \{e_1, \ldots, e_n\}$ eine Orthonormalbasis und $\varphi \in \mathrm{End}(V)$. Genau dann ist $\varphi(B)$ eine Orthonormalbasis, wenn die zu φ gehörende Matrixdarstellung A_φ orthogonal ist.*

Beweis. Mit $A_\varphi = (a_{ij})$ ist $\varphi(e_j) = a_{1j}e_1 + \ldots + a_{nj}e_n$ für $j = 1, \ldots, n$. Ist $\varphi(B)$ eine Orthonormalbasis von V, so ist A_φ auch die Transformationsmatrix des Basiswechsels $\varphi(B) \to B$, also A_φ orthogonal. Ist andererseits A_φ orthogonal, so ist A_φ regulär und damit $\varphi(B)$ eine Basis von V. Da A_φ dann auch wieder die Transformationsmatrix des Basiswechsels $\varphi(B) \longrightarrow B$ ist, ist $\varphi(B)$ eine Orthonormalbasis.

□

2. Orthogonale Endomorphismen

Definition 2.1 *Ist V ein euklidischer Vektorraum, dann heißt $\varphi \in \mathrm{End}(V)$ orthogonal, wenn für alle $v, w \in V$ gilt $\langle \varphi(v), \varphi(w) \rangle = \langle v, w \rangle$.*

Einfache Eigenschaften. Sei $\varphi \in \mathrm{End}(V)$ orthogonal.

1. $\|\varphi(v)\| = \|v\|$ für alle $v \in V$.
2. $v \perp w \Longrightarrow \varphi(v) \perp \varphi(w)$ für alle $v, w \in V$.
3. Ist $k \in \mathbb{R}$ ein Eigenwert von φ, so gilt $k \in \{1, -1\}$.
   Um das einzusehen, nehmen wir an, v sei ein zugehöriger Eigenvektor von φ. Dann gilt $\varphi(v) = kv$ und damit $\|\varphi(v)\| = \|kv\| = |k| \cdot \|v\|$. Wegen Eigenschaft 1 ist dann $||v|| = |k| \cdot \|v\|$, also $|k| = 1$.
4. φ ist injektiv.
5. Ist φ bijektiv, so ist φ^{-1} orthogonal.
6. Ist V endlich-dimensional, so ist φ bijektiv.
7. Sind die Endomorphismen φ und ψ orthogonal, dann ist es auch $\varphi \circ \psi$.

Beispiel.

1. Die Identität $\mathrm{id} : V \longrightarrow V$ ist orthogonal.
2. Der Endomorphismus $-\mathrm{id} : V \longrightarrow V\,; \quad v \longmapsto -v$ ist orthogonal.

Definition 2.2 *Sei V ein euklidischer Vektorraum. Dann heißt*

$$\mathrm{O}(V) := \{\varphi : V \longrightarrow V \mid \varphi \in \mathrm{Aut}(V) \text{ und } \varphi \text{ ist orthogonal}\}$$

Gruppe der orthogonalen Automorphismen von V.

Bemerkung.

1. $\mathrm{O}(V)$ ist eine Untergruppe von $\mathrm{Aut}(V)$.

2. Ist V endlich-dimensional, so gilt $\mathrm{O}(V) = \{\varphi \in \mathrm{End}(V) \mid \varphi \text{ ist orthogonal}\}$.

Satz 2.3 *Sei V ein euklidischer Vektorraum und $B = \{e_1, \ldots, e_n\}$ eine Orthonormalbasis von V. Ein $\varphi \in \mathrm{End}(V)$ ist genau dann orthogonal, wenn $\{\varphi(e_1), \ldots, \varphi(e_n)\}$ ebenfalls eine Orthonormalbasis ist.*

Beweis.
"$\Rightarrow$": Wegen Eigenschaft 6 ist $\{\varphi(e_1), \ldots, \varphi(e_n)\}$ eine Basis, deren Vektoren wegen Eigenschaft 1 auch normiert sind. Die Orthogonalität der Basiselemente folgt aus Eigenschaft 2.
"$\Leftarrow$": Sind $v, w \in V$ mit $v = k_1 e_1 + \ldots + k_n e_n$ und $w = l_1 e_1 + \ldots + l_n e_n$, dann gilt

$$\langle v, w \rangle = k_1 l_1 + \ldots + k_n l_n.$$

Andererseits ist aber $\varphi(v) = k_1\varphi(e_1) + \ldots + k_n\varphi(e_n)$ und $\varphi(w) = l_1\varphi(e_1) + \ldots + l_n\varphi(e_n)$, also

$$\langle \varphi(v), \varphi(w) \rangle = k_1 l_1 + \ldots + k_n l_n.$$

□

Korollar 2.4 *Ist V ein endlich-dimensionaler euklidischer Vektorraum und B eine Orthonormalbasis, dann ist $\varphi \in \mathrm{End}(V)$ genau dann orthogonal, wenn A_φ (bezüglich B) orthogonal ist.*

Beweis. Das Korollar folgt direkt aus Korollar 1.13 und Satz 2.3.

□

Im folgenden sollen die Matrizen $A \in \mathrm{O}(2)$ explizit berechnet und geometrisch gedeutet werden.

Für eine allgemeine Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist

$$A^t A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}.$$

Damit A orthogonal ist, muß nach Definition 1.11 also gelten

1. $a^2 + c^2 = 1$ und damit $a = \cos\alpha$, $c = \sin\alpha$ für ein $\alpha \in [0, 2\pi)$.

2. $b^2 + d^2 = 1$, also $d = \cos\beta$, $b = \sin\beta$ für ein $\beta \in [0, 2\pi)$.

3. $ab + cd = \cos\alpha\sin\beta + \cos\beta\sin\alpha = \sin(\alpha + \beta) = 0$ und damit $\alpha + \beta = k\pi$ mit $k \in \mathbb{Z}$.

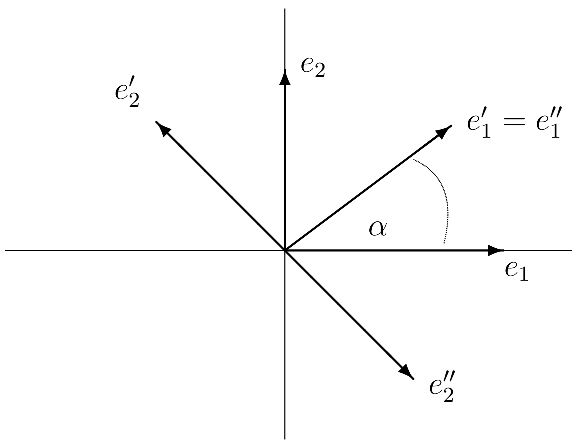
Ist k gerade, so gilt $d = \cos(-\alpha) = \cos\alpha$ und $b = \sin(-\alpha) = -\sin\alpha$. Die Matrix A hat dann die Gestalt

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Ist k ungerade, dann ist $d = -\cos\alpha$ sowie $b = \sin\alpha$ und es ergibt sich

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

Sei nun V ein 2-dimensionaler euklidischer Vektorraum mit der Orthonormalbasis $\{e_1, e_2\}$ sowie $\varphi \in \mathrm{O}(V)$ mit der zugehörigen Matrixdarstellung $A_\varphi = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. Dann folgt $\varphi(e_1) = \cos\alpha e_1 + \sin\alpha e_2 =: e_1'$ und $\varphi(e_2) = -\sin\alpha e_1 + \cos\alpha e_2 =: e_2'$.
Gilt aber $A_\varphi = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$, so ist $\varphi(e_1) =: e_1'' = e_1'$ und $\varphi(e_2) =: e_2'' = -e_2'$. Die beiden Matrizen beschreiben also eine Drehung bzw. eine Spiegelung.



Ob φ eine Drehung oder Spiegelung ist, kann bereits durch $\det\varphi$ erkannt werden:

$$\det\varphi = 1: \quad A_\varphi = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \qquad \text{und } \varphi \text{ ist eine Drehung.}$$

$$\det\varphi = -1: \quad A_\varphi = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \qquad \text{und } \varphi \text{ ist eine Spiegelung.}$$

Ist φ eine Spiegelung, so kann die Richtung der Spiegelachse mit Hilfe der Eigenvektoren berechnet werden. Sei dazu A_φ wie oben und $\alpha \neq 0, \pi$. Das charakteristische Polynom der Spiegelungsmatrix ist $\chi_\varphi(x) = x^2 - 1 = (x-1)(x+1)$. Ein Eigenvektor von A_φ zum Eigenwert 1 ergibt sich zu $(-\sin\alpha, \cos\alpha - 1)$, und der zugehörige normierte Vektor ist $(\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2})$. Ein

Eigenvektor von A_φ zum Eigenwert -1 ist $(-\sin\alpha, \cos\alpha+1)$, und der zugehörige normierte Vektor ergibt sich zu $(-\sin\frac{\alpha}{2}, \cos\frac{\alpha}{2})$. Mit

$$\tilde{e}_1 = \cos\frac{\alpha}{2}e_1 + \sin\frac{\alpha}{2}e_2 \quad \text{und} \quad \tilde{e}_2 = -\sin\frac{\alpha}{2}e_1 + \cos\frac{\alpha}{2}e_2$$

ist $\{\tilde{e}_1, \tilde{e}_2\}$ eine Orthonormalbasis von V sowie $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ die zugehörige Matrixdarstellung von φ. Damit liegt $\tilde{e}_1$ in Richtung der Spiegelachse.
Für $\alpha = 0$ ist $A_\varphi = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, und für $\alpha = \pi$ ist $A_\varphi = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

Definition 2.5 *Die Menge* $\mathrm{SO}(n) := \{A \in \mathrm{O}(n) \mid \det A = 1\}$ *heißt Gruppe der eigentlich orthogonalen Matrizen.*

Bemerkung.

1. Für jedes $A \in \mathrm{O}(n)$ gilt $\det A = \pm 1$, denn

$$1 = \det E_n = \det(A^t \cdot A) = \det A^t \cdot \det A = (\det A)^2\,.$$

2. $\mathrm{SO}(n)$ ist eine Untergruppe von $\mathrm{O}(n)$, denn
   (a) $E_n \in \mathrm{SO}(n) \quad \Longrightarrow \quad \mathrm{SO}(n) \neq \emptyset$.
   (b) $A \in \mathrm{SO}(n) \quad \Longrightarrow \quad \det A^{-1} = (\det A)^{-1} = 1 \quad \Longrightarrow \quad A^{-1} \in \mathrm{SO}(n)$.
   (c) $A, B \in \mathrm{SO}(n) \quad \Longrightarrow \quad \det(A \cdot B) = \det A \cdot \det B = 1 \quad \Longrightarrow \quad A \cdot B \in \mathrm{SO}(n)$.

Satz 2.6 *Sei V ein n-dimensionaler euklidischer Vektorraum und sei $\varphi \in \mathrm{End}(V)$ orthogonal. Dann gibt es eine Orthonormalbasis B von V, so daß gilt*

$$A_\varphi = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&& 0 & \\ &&&& \ddots &&&& \\ 0 &&&&& -1 &&& \\ &&&&&& A_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& A_k \end{pmatrix}$$

wobei $A_i = \begin{pmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{pmatrix} \in \mathrm{SO}(2)$ *mit* $\alpha_i \in (0, 2\pi)$, $\alpha_i \neq \pi$.

Bevor wir den Satz beweisen, führen wir einige Vorüberlegungen durch. Ist V ein n-dimensionaler euklidischer Vektorraum und U ein Unterraum von V, dann heißt

$$U^\perp := \{v \in V \mid \forall u \in U : \langle u, v\rangle = 0\}$$

orthogonales Komplement von U.

Bemerkung.

1. $U^\perp$ ist ein Unterraum von V.
2. $U \cap U^\perp = \{\mathcal{O}\}$, denn für $v \in U \cap U^\perp$ gilt $\langle v, v\rangle = 0$ und damit $v = \mathcal{O}$.
3. $U + U^\perp = V$. Um das zu sehen, wählen wir eine Orthonormalbasis B von U und ergänzen diese zu einer Orthonormalbasis $B \cup B'$ von V. Dann ist jedes $v \in B'$ orthogonal zu jedem $w \in B$ und damit auch zu jedem $w \in U$, das heißt $B' \subseteq U^\perp$. Es folgt $B \cup B' \subseteq U + U^\perp$ und damit $\dim(U + U^\perp) \geq |B \cup B'| = \dim V$, also $U + U^\perp = V$.
4. $V = U \oplus U^\perp$.

Folgerung.

1. Ist B_1 eine Orthonormalbasis von U und B_2 eine Orthonormalbasis von $U^\perp$, dann ist $B_1 \cup B_2$ eine Orthonormalbasis von V.
2. Ist φ orthogonal, dann gilt $\varphi(U)^\perp = \varphi(U^\perp)$.

Beweis von Satz 2.6.
Wir beweisen den Satz durch Induktion nach n.
$n = 1$: Sei $V = [e_1]$ mit $\|e_1\| = 1$ und $\varphi(e_1) = k \cdot e_1$ mit $|k| = 1$. Dann ist $A_\varphi = (1)$ oder $A_\varphi = (-1)$.
$n > 1$: Zunächst gilt $\chi_\varphi(x) = f_1(x) \cdot \ldots \cdot f_r(x)$ mit $f_i(x) \in \mathbb{R}[x]$ und $1 \leq \mathrm{grad} f_i(x) \leq 2$. Wegen Aufgabe 5.5 aus Kapitel 5 gibt es einen Unterraum U von V mit $1 \leq \dim U \leq 2$ und $\varphi(U) \subseteq U$. Da φ bijektiv ist, folgt $\varphi(U) = U$. Weiterhin gilt $\varphi(U^\perp) = \varphi(U)^\perp = U^\perp$. Definieren wir jeweils die orthogonalen Abbildungen eingeschränkt auf die beiden Unterräumen durch

$$\begin{aligned} \psi :\quad & U \longrightarrow U\,; & u \longmapsto \varphi(u) \quad \text{und} \\ \psi^\perp :\quad & U^\perp \longrightarrow U^\perp\,; & u \longmapsto \varphi(u)\,, \end{aligned}$$

dann gilt wegen $\dim U^\perp < \dim V$ und der Induktionsvoraussetzung, daß $U^\perp$ eine Orthonormalbasis B hat, bezüglich der $\psi^\perp$ die folgende Matrixdarstellung besitzt

$$A_{\psi^\perp} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & 0 & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & 0 & & & & A_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & A_l \end{pmatrix}.$$

1. Fall: $\dim U = 1$. Dann gilt $U = [e]$ mit $\|e\| = 1$ und $\varphi(e) = \pm 1 \cdot e$. Wegen obiger Folgerung ist $\{e\} \cup B$ eine Orthonormalbasis von V bezüglich der $A_\varphi = \begin{pmatrix} \pm 1 & 0 \\ 0 & A_{\psi^\perp} \end{pmatrix}$ gilt. Um letztendlich A_φ in der gewünschten Form zu haben, müssen die Vektoren aus $\{e\} \cup B$ gegebenenfalls umsortiert werden.
2. Fall: $\dim U = 2$. Wir behandeln die Drehung und Spiegelung getrennt.
$\det \psi = 1$: Sei B' eine Orthonormalbasis von U, so daß $A_\psi = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ mit $\alpha \in [0, 2\pi)$, also insbesondere $A_\psi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ falls $\alpha = 0$ und $A_\psi = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ falls $\alpha = \pi$. Dann ist $B' \cup B$ eine Orthonormalbasis von V, und A_φ hat eine der folgenden Formen

$$\begin{pmatrix} 1 & & \\ & 1 & 0 \\ & 0 & A_{\psi^\perp} \end{pmatrix}; \begin{pmatrix} -1 & & \\ & -1 & 0 \\ & 0 & A_{\psi^\perp} \end{pmatrix}; \begin{pmatrix} \cos\alpha & -\sin\alpha & \\ \sin\alpha & \cos\alpha & 0 \\ & 0 & A_{\psi^\perp} \end{pmatrix}, \alpha \in (0, 2\pi), \alpha \neq \pi.$$

Um A_ψ in der gewünschten Form zu haben, muß $B' \cup B$ gegebenenfalls umsortiert werden.
$\det \psi = -1$: Sei B' eine Orthonormalbasis von U mit $A_\psi = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Dann ist $B' \cup B$ eine Orthonormalbasis von V mit

$$A_\varphi = \begin{pmatrix} 1 & & \\ & -1 & 0 \\ & 0 & A_{\psi^\perp} \end{pmatrix}.$$

$B' \cup B$ muß gegebenenfalls wieder umsortiert werden.

□

Definition 2.7 *Sei V ein n-dimensionaler euklidischer Vektorraum und $\varphi \in \mathrm{End}(V)$ orthogonal.*

i) φ heißt eigentlich orthogonal oder Drehung, wenn $\det \varphi = 1$.

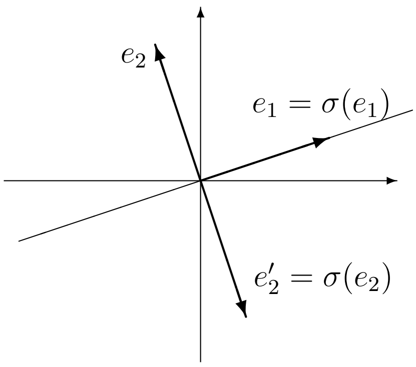
ii) φ heißt uneigentlich orthogonal, wenn $\det \varphi = -1$.

Bemerkung. Die eigentlich orthogonalen Automorphismen von V bilden eine Untergruppe von $\mathrm{O}(V)$, die mit $\mathrm{SO}(V)$ bezeichnet wird.

Spiegelungen:
Ein Unterraum H von V heißt Hyperebene, wenn $H \neq V$ und $V = H + [v]$ für $v \in V \setminus H$ gilt. Ist $\dim V = n$, so sind die Hyperebenen genau die Unterräume der Dimension $n - 1$. Sei nun H eine Hyperebene, $\{e_1, \ldots, e_{n-1}\}$ eine Orthonormalbasis von H und $\{e_1, \ldots, e_{n-1}, e_n\} =: B$ eine Orthonormalbasis von V. Es gibt ein $\sigma \in \mathrm{End}(V)$ mit $\sigma(e_1) = e_1, \ldots, \sigma(e_{n-1}) = e_{n-1}$

und $\sigma(e_n) = -e_n$. Bezüglich B hat σ die Matrixdarstellung $\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$, und damit ist σ uneigentlich orthogonal. Jedes $v \in V$ besitzt eine eindeutige Darstellung $v = h + k \cdot e_n$ mit $h \in H, k \in \mathbb{R}$, und damit ist $\sigma(v) = h - k \cdot e_n$. Die Abbildung σ heißt Spiegelung an der Hyperebene H.

Für $n = 2$ ist eine Hyperebene ein 1-dimensionaler Unterraum, also eine Gerade.



Drehungen und Spiegelungen im Falle $n = 2$.

Sei $B = \{e_1, e_2\}$ eine Orthonormalbasis von V und $\varphi, \psi \in \mathrm{End}(V)$ orthogonal.

Sind φ und ψ Drehungen, also $\det \varphi = \det \psi = 1$, dann haben sie Matrixdarstellungen der Gestalt $A_\varphi = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ und $A_\psi = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$. Für die Hintereinanderausführung $\varphi \circ \psi$ ergibt sich

$$A_{\varphi\circ\psi} = \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}$$

und für φ^{-1} im Falle $\alpha \neq 0$

$$A_{\varphi^{-1}} = \begin{pmatrix} \cos(2\pi-\alpha) & -\sin(2\pi-\alpha) \\ \sin(2\pi-\alpha) & \cos(2\pi-\alpha) \end{pmatrix}.$$

Die Drehungen bilden damit bezüglich der Hintereinanderschaltung eine abelsche Gruppe.
Wählt man nun eine andere Matrixdarstellung von φ zum Beispiel dadurch, daß man e_1 und e_2 vertauscht, so erhält man

$$A'_\varphi = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha' & -\sin\alpha' \\ \sin\alpha' & \cos\alpha' \end{pmatrix}$$

mit $\alpha' = 2\pi - \alpha$ für $\alpha \neq 0$, also $\alpha = 2\pi - \alpha'$. Dieses zeigt, daß der Winkel α im allgemeinen nicht eindeutig bestimmt ist. Um einer Drehung eindeutig einen Drehwinkel zuordnen zu können, muß V orientiert werden. Sind B und B' Orthonormalbasen von V und T die zugehörige Transformationsmatrix, dann ist T wegen Satz 1.12 orthogonal. Die Basen B

und B' heißen nun gleich bzw. entgegengesetzt orientiert, wenn $\det T = 1$ beziehungsweise $\det T = -1$. Definiert man die Basen einer der beiden Klassen von jeweils gleich orientierten Basen als positiv orientiert, so heißt V orientiert. Sind B und B' beide positiv orientiert, dann gilt für die zugehörigen Darstellungen A_φ und A'_φ von φ der Zusammenhang $A_\varphi = T^{-1}A'_\varphi T = A'_\varphi$, wobei T^{-1}, T und A'_φ jeweils eigentlich orthogonal sind. Der Drehwinkel $\alpha \in [0, 2\pi)$ ist dann eindeutig bestimmt.

Ist φ uneigentlich orthogonal, also $\det \varphi = -1$, dann ist $A_\varphi = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$. Für die Vektoren $v := (\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2})$ und $w := (-\sin\frac{\alpha}{2}, \cos\frac{\alpha}{2})$ gilt dann

$$A_\varphi v^t = \left(\cos\alpha\cos\frac{\alpha}{2} + \sin\alpha\sin\frac{\alpha}{2}, \sin\alpha\cos\frac{\alpha}{2} - \cos\alpha\sin\frac{\alpha}{2}\right)^t = \left(\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2}\right)^t.$$

v ist also ein Eigenvektor von A_φ zum Eigenwert 1. Entsprechend ergibt sich $A_\varphi w^t = -w^t$, also ist w ein Eigenvektor von A_φ zum Eigenwert -1.
Mit den Definitionen $e'_1 := \cos\frac{\alpha}{2}e_1 + \sin\frac{\alpha}{2}e_2$ und $e'_2 := -\sin\frac{\alpha}{2}e_1 + \cos\frac{\alpha}{2}e_2$ ist $\{e'_1, e'_2\}$ eine Orthonormalbasis, bezüglich der die Matrixdarstellung von φ durch $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ gegeben ist. φ ist damit eine Spiegelung an der "Hyperebene" $[e'_1]$.

Satz 2.8 *Jede Drehung in der euklidischen Ebene ist das Produkt von zwei Spiegelungen.*

Beweis.

$$\underbrace{\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}}_{\text{Drehung}} = \underbrace{\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}}_{\text{Spiegelung}} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\text{Spiegelung}}$$

□

Drehungen und Spiegelungen im Falle $n = 3$.
Sei $\varphi \in \mathrm{End}(V)$ orthogonal.
$\det\varphi = 1$: Es gibt eine Orthonormalbasis $B = \{e_1, e_2, e_3\}$ mit

$$A_\varphi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}.$$

Für $\alpha = 0$ ist dann $\varphi = \mathrm{id}$.
Für $\alpha \neq 0$ ist $[e_1]$ Eigenraum zum Eigenwert 1. Man nennt $[e_1]$ Drehachse und $[e_2, e_3]$ Drehebene. Die Drehebene ist offenbar orthogonal zur Drehachse. α heißt Drehwinkel und ist eindeutig bestimmt, wenn die Drehebene orientiert ist. Mit dem Zusammenhang $\mathrm{Spur}\varphi = 1 + 2\cos\alpha$ läßt sich dann der Drehwinkel durch $\cos\alpha = \frac{1}{2}(\mathrm{Spur}\varphi - 1)$ berechnen.

Beispiel. Bezüglich einer Orthonormalbasis sei $A_\varphi = \frac{1}{11}\begin{pmatrix} 2 & -6 & 9 \\ 6 & -7 & -6 \\ 9 & 6 & 2 \end{pmatrix}$. Dann gilt $\det\varphi = 1$, die Drehachse ist $[(1,0,1)]$ und die Drehebene $[(0,1,0),(1,0,-1)]$. Bezüglich einer geeigneten Orthonormalbasis hat φ dann die Darstellung:

$$\frac{1}{11}\begin{pmatrix} 11 & 0 & 0 \\ 0 & -7 & 6\sqrt{2} \\ 0 & -6\sqrt{2} & -7 \end{pmatrix}.$$

Es gilt $\mathrm{Spur}\varphi = \frac{1}{11}(2-7+2) = \frac{1}{11}(11-7-7) = -\frac{3}{11}$, und für den Drehwinkel α ergibt sich $\cos\alpha = \frac{1}{2}(-1-\frac{3}{11}) = -\frac{7}{11}$.

Aus Satz 2.8 folgt nun

Satz 2.9 *Im 3-dimensionalen euklidischen Raum ist jede Drehung das Produkt von zwei Spiegelungen und jeder uneigentlich orthogonale Endomorphismus das Produkt von drei Spiegelungen.*

Reelle, symmetrische Matrizen.
Für die Anwendungen in der Geometrie sind zum Beispiel diejenigen reellen Matrizen A wichtig, die sich durch orthogonale Matrizen diagonalisieren lassen, das heißt $U^tAU = D$ mit $U^tU = E_n$ und D diagonal. Gilt $U^tAU = D$, so ist A symmetrisch, denn mit $A = UDU^t$ folgt $A^t = (U^t)^tD^tU^t = UDU^t = A$. Wir zeigen nun umgekehrt, daß sich jede relle, symmetrische Matrix A auch durch eine orthogonale Matrix auf Diagonalform transformieren läßt. Sei also im folgenden $A \in \mathbb{R}_{n,n}$ symmetrisch. Für alle $x \in \mathbb{R}^n$ gilt dann $Ax^t = A^tx^t = (xA)^t$, das heißt, x ist genau dann ein Eigenvektor von A zum Eigenwert c, wenn $xA = cx$. Jeder Spalteneigenvektor von A ist also auch Zeileneigenvektor zum selben Eigenwert und umgekehrt. Sei nun weiterhin $V = \mathbb{R}^n$ mit dem kanonischen Skalarprodukt $\langle x,y\rangle = x\cdot y^t$ versehen. Dann gilt

$$\langle x, yA\rangle = x\cdot(yA)^t = xA\cdot y^t = \langle xA, y\rangle\,.$$

Satz 2.10 *Sei $A \in \mathbb{R}_{n,n}$ symmetrisch.*

i) Sind $c, c' \in \mathbb{R}$ verschiedene Eigenwerte von A mit zugehörigen Eigenvektoren v, v', so gilt $\langle v, v'\rangle = 0$.

ii) A ist diagonalisierbar.

Beweis.

i) Es gilt

$$\langle v, v'A\rangle = \langle v, c'v'\rangle = c'\langle v, v'\rangle$$

und andererseits

$$\langle v, v'A\rangle = \langle vA, v'\rangle = \langle cv, v'\rangle = c\langle v, v'\rangle\,.$$

Also folgt $(c'-c)\cdot\langle v, v'\rangle = 0$ und damit $\langle v, v'\rangle = 0$.

ii) Bezeichnet $\mu_A(x)$ das Minimalpolynom von A, dann genügt es wegen Satz 2.13 aus Kapitel 5 zu zeigen, daß $\mu_A(x)$ vollständig in paarweise verschiedene Linearfaktoren zerfällt. Sei $\mu_A(x) = f_1(x) \cdot \ldots \cdot f_r(x)$ mit $f_i(x) \in \mathbb{R}[x]$ normiert und $1 \leq \mathrm{grad} f_i(x) \leq 2$. Behauptung: Gilt $\mathrm{grad} f_i(x) = 2$, dann ist $f_i(x) = (x-\alpha)(x-\beta)$ mit verschiedenen $\alpha, \beta \in \mathbb{R}$.
O.B.d.A. sei $i = 1$ und $f_1(x) = (x-a)^2 + b$ mit $a, b \in \mathbb{R}$. Ist $b < 0$, so folgt $f_1(x) = (x - a - \sqrt{-b})(x - a + \sqrt{-b})$ und damit die Behauptung. Wir nehmen also $b \geq 0$ an und führen dies zum Widerspruch. Mit der Definition

$$g(x) := (x-a) f_2(x) \cdot \ldots \cdot f_r(x)$$

folgt $\mathrm{grad} g(x) < \mathrm{grad} \mu_A(x)$. Wir zeigen $g(A) = \mathcal{O}$ im Widerspruch zur Minimalität von $\mathrm{grad} \mu_A(x)$. Zu zeigen ist also $g(A) \cdot v^t = \mathcal{O}^t$ für alle $v \in V = \mathbb{R}^n$.
1.Fall: $f_2(A) \cdot \ldots \cdot f_r(A) \cdot v^t = \mathcal{O}^t$. Dann folgt

$$g(A) \cdot v^t = (A - a \cdot E_n) f_2(A) \cdot \ldots \cdot f_r(A) \cdot v^t = \mathcal{O}^t \, .$$

2. Fall: $f_2(A) \cdot \ldots \cdot f_r(A) \cdot v^t =: w^t \neq \mathcal{O}^t$. Dann gilt $f_1(A) \cdot w^t = \mu_A(A) \cdot v^t = \mathcal{O}^t$, also

$$\begin{aligned} 0 &= w \cdot f_1(A) \cdot w^t = w \cdot ((A - aE_n)^2 + bE_n) \cdot w^t \\ &= w \cdot (A - aE_n)(A - aE_n)^t \cdot w^t + bw \cdot w^t \\ &= \langle w(A - aE_n), w(A - aE_n) \rangle + b \langle w, w \rangle \, . \end{aligned}$$

Wegen $b \geq 0$ folgt insbesondere $w(A - aE_n) = \mathcal{O}$, das heißt

$$g(A) \cdot v^t = (A - aE_n) f_2(A) \cdot \ldots \cdot f_r(A) \cdot v^t = (A - aE_n) w^t = \mathcal{O}^t.$$

□

Satz 2.11 *Ist $A \in \mathbb{R}_{n,n}$ symmetrisch, dann gibt es ein $U \in \mathrm{O}(n)$, so daß $U^t A U$ eine Diagonalmatrix ist.*

Beweis. Wegen Satz 2.10 ist A diagonalisierbar. Seien $c_1, \ldots, c_k$ die paarweise verschiedenen Eigenwerte und $V = U_{c_1} \oplus \ldots \oplus U_{c_k}$, wobei U_{c_i} den Eigenraum zum Eigenwert c_i bezeichnet. B_i sei eine Orthonormalbasis von U_{c_i} bezüglich des kanonischen Skalarproduktes von $\mathbb{R}^n$. Dann ist $B = B_1 \cup \ldots \cup B_k$ eine Basis von $\mathbb{R}^n$ mit normierten Vektoren. Aus Satz 2.10 folgt, daß die Vektoren aus B paarweise orthogonal sind. Ist $B = \{v_1, \ldots, v_n\}$ und wählen wir für U die (n,n)-Matrix, die $v_1, \ldots, v_n$ als Spaltenvektoren hat, also $U = (v_1^t, \ldots, v_n^t)$, dann gilt $U^t U = (v_i v_j^t)$ wobei

$$v_i v_j^t = \begin{cases} 1 & \\ & \text{für} \\ 0 & \end{cases} \begin{matrix} i = j \\ \\ i \neq j \end{matrix} \, ,$$

also $U^t U = E_n$. Es folgt $U \in \mathrm{O}(n)$, und $U^t A U = U^{-1} A U$ ist eine Diagonalmatrix.

□

Anwendung. Sei $A \in \mathbb{R}_{n,n}$ symmetrisch. Wegen Satz 1.2 wird durch $\langle x, y \rangle := xAy^t$ auf dem $\mathbb{R}^n$ eine symmetrische Bilinearform definiert. Wir wollen nun nachprüfen, ob $\langle\ ,\ \rangle$ auch positiv definit und damit ein Skalarprodukt ist.
Behauptung: $\langle\ ,\ \rangle$ ist genau dann ein Skalarprodukt, wenn A nur positive Eigenwerte hat.
Beweis: "$\Rightarrow$": Sei c ein Eigenwert von A und v ein zugehöriger Eigenvektor, also insbesondere $v \neq \mathcal{O}$. Dann gilt

$$0 < \langle v, v \rangle = vAv^t = v(c \cdot v)^t = c \cdot vv^t = c \cdot (v_1^2 + \ldots + v_n^2)$$

und damit $c > 0$.
"$\Leftarrow$": Sei $v \in \mathbb{R}^n$, $v \neq \mathcal{O}$. Zu zeigen ist $\langle v, v \rangle = vAv^t > 0$. Wegen Satz 2.11 gibt es ein $U \in \mathrm{O}(n)$ mit $U^tAU = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}$, wobei $c_1, \ldots, c_n$ die (positiven) Eigenwerte von A sind. Wegen $v \neq \mathcal{O}$ ist auch $w := vU \neq \mathcal{O}$. Setzen wir $w = (w_1, \ldots, w_n)$, so folgt

$$vAv^t = v\ UU^t\ A\ UU^tv^t = w \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} w^t = c_1w_1^2 + \ldots + c_nw_n^2 > 0\,.$$

Also ist $\langle\ ,\ \rangle$ positiv definit.

Beispiel. Sei

$$A = \begin{pmatrix} 85 & 12 & -18 \\ 12 & 53 & -6 \\ -18 & -6 & 58 \end{pmatrix}.$$

Wird durch $\langle x, y \rangle := xAy^t$ auf dem $\mathbb{R}^3$ ein Skalarprodukt definiert?
Das charakteristische Polynom von A ist $\chi_A(x) = -x^3 + 196x^2 - 12005x + 235298$. Damit hat A die Eigenwerte $49, 49$ und 98, und $\langle\ ,\ \rangle$ ist tatsächlich ein Skalarprodukt.

Um zu zeigen, daß $\langle\ ,\ \rangle$ positiv definit ist, brauchen die Eigenwerte von A nicht explizit berechnet zu werden. Da $\chi_A(x)$ über $\mathbb{R}$ vollständig in Linearfaktoren zerfällt, kann bereits mit Hilfe der Koeffizienten von $\chi_A(x)$ die Anzahl der positiven Nullstellen von $\chi_A(x)$ ermittelt werden, zum Beispiel durch die kartesische Zeichenregel:

Sei $a_0, a_1, \ldots, a_n$ eine Folge reeller Zahlen. Zwischen a_i und a_j mit $i < j$ liegt ein Zeichenwechsel vor, wenn $a_i \cdot a_j < 0$ und die Folgeglieder zwischen a_i und a_j alle 0 sind.
Beispiel: Die Folge $5, 1, 0, -1, 2, 0, -3, 0, 0, -3$ hat 3 Zeichenwechsel.

Satz 2.12 *Ein Polynom $f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 \in \mathbb{R}[x]$ habe nur reelle Nullstellen. Dann ist die Anzahl der positiven Nullstellen gleich der Anzahl der Zeichenwechsel in $a_n, a_{n-1}, \ldots, a_0$.*

Ist $f(x)$ das charakteristische Polynom einer reellen, symmetrischen Matrix, so zerfällt $f(x)$ über $\mathbb{R}$ vollständig in Linearfaktoren. Damit ist die Voraussetzung von Satz 2.12 erfüllt. Im obigen konkreten Beispiel für A ergab sich $\chi_A(x) = -x^3 + 196x^2 - 12005x + 235298$. Da in der Folge $-1, 196, -12005, 235298$ insgesamt 3 Zeichenwechsel vorkommen, hat $\chi_A(x)$ genau 3 positive Nullstellen.

3. Aufgaben

A 3.1 V sei ein endlich-dimensionaler K-Vektorraum mit den Basen B und B', und die Abbildung $\beta : V \times V \to K$ sei bilinear. Welcher Zusammenhang besteht zwischen den Matrixdarstellungen von β bezüglich B und B'? (Hinweis: Benutzen Sie die Transformationsmatrix des zugehörigen Basiswechsels.)

A 3.2 V sei der Unterraum des $\mathbb{R}^4$, der von $\{a_1, a_2, a_3\}$ aufgespannt wird, wobei $a_1 = (2, -1, 2, 0), a_2 = (1, -3, 2, -2)$ und $a_3 = (-1, -3, 4, -1)$. Berechnen Sie bezüglich des kanonischen Skalarprodukts eine Orthonormalbasis von V und geben Sie den Lösungsweg an. Stellen Sie $a_1 - a_2 - a_3$ als Linearkombination der neuen Basisvektoren dar.

A 3.3 V sei der Unterraum von $C^0([0,1])$, der von $\{f_0, f_1, f_2, f_3\}$ mit $f_n(x) = x^n$ aufgespannt wird. Das Skalarprodukt sei gegeben durch

$$< f, g >= \int_0^1 f(x)g(x)dx.$$

Berechnen Sie eine Orthonormalbasis von V.

A 3.4 Gegeben ist die reelle Matrix

$$A = \begin{pmatrix} 2 & -2 & -1 \\ -2 & 5 & 2 \\ -1 & 2 & 2 \end{pmatrix}.$$

1. Zeigen Sie, daß durch $< x, y >= xAy^t$ auf $\mathbb{R}^3$ ein Skalarprodukt definiert wird.
2. Berechnen Sie eine Orthonormalbasis des $\mathbb{R}^3$ bezüglich dieses Skalarprodukts.
3. Berechnen Sie $U \in \mathrm{O}(3)$ so, daß U^tAU eine Diagonalmatrix ist.

A 3.5 Zeigen Sie, daß es zu jeder Matrix $A \in \mathrm{SO}(3)$ Winkel $\alpha, \beta, \gamma \in [0, 2\pi)$ gibt mit:

$$A = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Winkel α, β, γ heißen Eulersche Winkel zu A.

A 3.6 Berechnen Sie die Eulerschen Winkel für $A \in \mathrm{SO}(3)$ mit

$$A = \begin{pmatrix} \frac{1}{4}\sqrt{3} + \frac{1}{2} & \frac{1}{4}\sqrt{3} - \frac{1}{2} & -\frac{1}{4}\sqrt{2} \\ \frac{1}{4}\sqrt{3} - \frac{1}{2} & \frac{1}{4}\sqrt{3} + \frac{1}{2} & -\frac{1}{4}\sqrt{2} \\ \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2} & \frac{1}{2}\sqrt{3} \end{pmatrix}.$$

A 3.7 Zeigen Sie, daß es für die Matrix A aus Aufgabe 3.6 ein $U \in \mathrm{O}(3)$ gibt, so daß gilt

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

Index