LINEARE ALGEBRA

Vorlesung an der
Universität Rostock

Prof. Dr. R. Knörr

Winter– und Sommersemester 2004/05

Inhalt

Das Griechische Alphabet

Name	klein	groß
Alpha	α	A
Beta	β	B
Gamma	γ	Γ
Delta	δ	Δ
Epsilon	ε, ϵ	E
Zeta	ζ	Z
Eta	η	H
Theta	ϑ, θ	Θ
Iota	ι	I
Kappa	κ	K
Lambda	λ	Λ
My	μ	M
Ny	ν	N
Xi	ξ	Ξ
Omikron	o	O
Pi	π	Π
Rho	ρ	P
Sigma	σ	Σ
Tau	τ	T
Ypsilon	υ	Υ
Phi	φ	Φ
Chi	χ	X
Psi	ψ	Ψ
Omega	ω	Ω

1 Mengen und Abbildungen

1.1 Bemerkung: *Aussagen, Quantoren*

Eine Aufgabe der Mathematik ist es, aus streng definierten Voraussetzungen mit Hilfe richtiger Schlüsse interessante Aussagen herzuleiten. Eine Aussage ist ein Satz, d.h. hat mindestens Subjekt und Prädikat. Eine Aussage ist wahr (w) oder falsch (f), aber nicht beides. Aus gegebenen Aussagen lassen sich neue Aussagen durch logische Verknüpfung gewinnen. Die wichtigsten dieser Verknüpfungen sind die Negation:

A	nicht A ($\neg A$)
w	f
f	w

sowie die Konjunktion, Disjunktion, Implikation, und Äquivalenz:

A	B	A und B ($A \wedge B$)	A oder B ($A \vee B$)	aus A folgt B ($A \Rightarrow B$)	A und B sind äquivalent ($A \Leftrightarrow B$)
w	w	w	w	w	w
w	f	f	w	f	f
f	w	f	w	w	f
f	f	f	f	w	w

Wenn eine Aussage $A(x)$ eine Variable x enthält, dann kann man diese quantifizieren mit dem Allquantor: für alle x gilt $A(x)$, symbolisch

$$\forall x\ A(x),$$

oder mit dem Existenzquantor: es gibt ein x, so daß $A(x)$ gilt, symbolisch

$$\exists x\ A(x).$$

Eine Verschärfung ist: es gibt genau ein x, so daß $A(x)$ gilt, symbolisch

$$\exists!\, x\ A(x).$$

Bei der Negation einer quantifizierten Aussage werden Allquantor und Existenzquantor vertauscht:

$$\neg\,(\forall x\ A(x)) \Leftrightarrow \exists x\ \neg A(x),$$

d.h. die Aussage „$A(x)$ trifft für alle x zu“ ist genau dann falsch, wenn es wenigstens ein x gibt, für welches die Aussage „Nicht $A(x)$“ richtig ist.

Vorsicht: Wenn mehrere Quantoren auftreten, kommt es auf deren Reihenfolge an. Vergleiche z.B. die Aussage

„Für jede natürliche Zahl n gibt es eine größere natürliche Zahl m“,

symbolisch $\forall n \in \mathbb{N}\ \exists m \in \mathbb{N}\ m > n$, mit der Aussage

„Es gibt eine natürliche Zahl m, welche größer als alle natürlichen Zahlen n ist“,

symbolisch $\exists m \in \mathbb{N}\ \forall n \in \mathbb{N}\ m > n$.

1.2 Vollständige Induktion:

Sei $A(n)$ eine Behauptung über natürliche Zahlen n. Wenn $A(0)$ gilt und $A(n) \Rightarrow A(n+1)$, dann gilt $A(n)$ für alle natürlichen Zahlen.

1.3 „Definition“: *naive Definition der Menge*

Eine Menge ist eine Zusammenfassung wohl unterschiedener Objekte unserer Anschauung oder unseres Denkens. Diese Objekte nennt man Elemente der Menge. m ist ein Element von M wird geschrieben als $m \in M$, die Negation ist $m \notin M$. Standardbezeichnungen sind $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ für die Mengen der natürlichen, ganzen, rationalen, reellen, komplexen Zahlen. Wenn A eine Menge ist, und $E(x)$ eine Aussage, welche für ein Element x aus A zutreffen kann, so bezeichnet $B = \{x \in A \mid E(x)\}$ die Menge der Elemente x aus A, für welche $E(x)$ richtig ist. B ist eine Teilmenge von A im Sinne von:

1.4 Definition: *Teilmenge, leere Menge*

Seien A und B Mengen. Man nennt A eine Teilmenge von B und schreibt $A \subseteq B$, falls $a \in A \Rightarrow a \in B$. Wenn $A \subseteq B$ und $B \subseteq A$, dann gilt $A = B$.

Eine wichtige Teilmenge jeder Menge ist die leere Menge $A = \emptyset$, welche keine Elemente enthält.

1.5 Konstruktion von Mengen:

Seien A und B Mengen. Dann lassen sich neue Mengen wie folgt konstruieren:

Vereinigung:	$A \cup B = \{x \mid x \in A \vee x \in B\}$
	$\bigcup_{i \in I} A_i = \{x \mid \exists i \in I : x \in A_i\}$
Schnitt:	$A \cap B = \{x \mid x \in A \wedge x \in B\}$
	$\bigcap_{i \in I} A_i = \{x \mid \forall i \in I : x \in A_i\}$
Differenz:	$A \setminus B = \{x \in A \wedge x \notin B\}$
symmetrische Differenz:	$A \dot{+} B = \{x \mid x \in A \setminus B \vee x \in B \setminus A\}$
Produkt:	$A \times B = \{(a,b) \mid a \in A \wedge b \in B\}$,
	dabei ist das geordnete Paar $(a,b) = \{\{a\}, \{a,b\}\}$
Potenzmenge:	$\wp(A) = 2^A = \{M \mid M \subseteq A\}$

1.6 Definition: *Abbildung; Produkt, Identität, Einschränkung*

Sind A und B Mengen, so ist eine Abbildung $f : A \to B$ ein Paar (T, B), wobei $T \subseteq A \times B$ eine Teilmenge mit der folgenden Eigenschaft ist: $\forall a \in A \;\; \exists! \, b \in B \;\; : \;\; (a,b) \in T$. Dieses b wird dann oft $f(a)$ (oder auch af oder a^f) geschrieben.

Zwei Abbildungen $f : A \to B$ und $g : X \to Y$ sind also gleich, wenn $A = X$ und $B = Y$ und $\forall a \in A \; : f(a) = g(a)$.

Wenn $f : A \to B$ und $g : B \to C$ Abbildungen sind, so ist $g \circ f : A \to C$, definiert durch $(g \circ f)(a) = g(f(a))$, ebenfalls eine Abbildung, genannt Produkt von f und g. Das Produkt wird oft einfach gf geschrieben. Im Allgemeinen gilt $g \circ f \neq f \circ g$, auch wenn $A = B = C$ ist.

Spezialfälle:

1. Die durch $f(a) = a \; \forall a \in A$ definierte Abbildung heißt die Identität von A, $f = \mathrm{id}_A$.
2. Wenn $X \subseteq A$ und $f : A \to B$ eine Abbildung ist, dann kann man durch $f_{|X}(x) = f(x) \; \forall x \in X$ eine Abbildung $f_{|X} : X \to B$ definieren, die man die Einschränkung von f auf X nennt.

1.7 Lemma:

Sei $f : A \to B$. Dann ist $f = \mathrm{id}_B \circ f = f \circ \mathrm{id}_A$.

Beweis: ist trivial.

1.8 Lemma:

Die Multiplikation von Abbildungen ist assoziativ, d.h. für Abbildungen $f : A \to B$, $g : B \to C$ und $h : C \to D$ gilt $(hg)f = h(gf)$.

Beweis: Beides sind Abbildungen $A \to D$. Für jedes $a \in A$ gilt

$$\begin{aligned} ((hg)f)(a) = (hg)(f(a)) = \quad & h(g(f(a))) \\ & \shortparallel \\ (h(gf))(a) = h((gf)(a)) = \quad & h(g(f(a))). \end{aligned}$$

1.9 Definition: *Umkehrabbildung (Inverse)*

Sind $f : A \to B$ und $g : B \to A$ Abbildungen mit $gf = \mathrm{id}_A$ und $fg = \mathrm{id}_B$, so heißt g die Inverse oder Umkehrabbildung zu f.

1.10 Lemma:

Es gibt höchstens eine Inverse zu einer Abbildung.

Beweis: Seien $g, h : B \to A$ Umkehrabbildungen zu $f : A \to B$. Zu zeigen: $g = h$. Es ist

$$g = g \circ \mathrm{id}_B = g(fh) = (gf)h = \mathrm{id}_A \circ h = h.$$

1.11 Definition: *injektiv, surjektiv, bijektiv*

Sei $f : A \to B$ eine Abbildung. Man nennt f

injektiv,	wenn $a, a' \in A \wedge a \neq a' \Rightarrow f(a) \neq f(a')$.
surjektiv,	wenn $\forall b \in B \; \exists a \in A \; : f(a) = b$.
bijektiv,	wenn f injektiv und surjektiv ist.

(Bijektive Abbildungen werden auch Bijektionen genannt.)

1.12 Bemerkung/Definition: *Bild einer Abbildung*

Schreibt man $f(A) = \{f(a) \,|\, a \in A\}$, dann gilt f surjektiv $\Leftrightarrow f(A) = B$. Für $f(A)$ schreibt man auch $\mathrm{Im}\,(f)$ und nennt dies das Bild von f.

1.13 Satz: *Existenz der Inversen*

Ist $f : A \to B$ eine Abbildung, so existiert die Umkehrabbildung g genau dann, wenn f bijektiv ist. Dann ist auch g bijektiv.

Beweis:

„$\Rightarrow$“: Sei g die Umkehrabbildung zu f.
f ist injektiv. Denn sei $a, a' \in A$ mit $f(a) = f(a')$. Dann $a = \mathrm{id}_A(a) = (gf)(a) = g(f(a)) = g(f(a')) = a'$.
f ist surjektiv. Denn für $b \in B$ ist $g(b) \in A$ und $f(g(b)) = (fg)(b) = \mathrm{id}_B(b) = b$.

„$\Leftarrow$“: Sei $b \in B$. Da f surjektiv ist, existiert ein $a \in A$ mit $f(a) = b$. Es gibt aber auch nur ein solches a, da f injektiv ist. Definiere $g(b) = a$, falls $f(a) = b$. Dann ist $(fg)(b) = f(g(b)) = f(a) = b$, d.h. $fg = \mathrm{id}_B$, und $(gf)(a) = g(f(a)) = g(b) = a$, d.h. $gf = \mathrm{id}_A$.

Da dann auch zu g eine Umkehrabbildung existiert (nämlich f), muß auch g bijektiv sein.

1.14 Lemma: *Vererbung*

Seien $f : A \to B$ und $g : B \to C$ Abbildungen. Wenn beide Abbildungen injektiv (surjektiv, bijektiv) sind, so gilt das auch für ihr Produkt gf.

Beweis: Übungsaufgabe.

1.15 Lemma:

Sei $f : A \to A$ eine injektive Abbildung und $f(A) \subseteq X \subseteq A$. Dann gibt es eine bijektive Abbildung $g : A \to X$.

Beweis: Sei $Y_0 := A \setminus X$, $Y_1 := f(Y_0)$, $Y_2 := f(Y_1)$ ... und $Y = \bigcup_{n \geq 0} Y_n$ (also ist $Y \subseteq A$). Definiere $g : A \to X$ durch

$$g(a) = \begin{cases} f(a) & \text{falls } a \in Y \\ a & \text{falls } a \notin Y \end{cases}$$

Es ist $g(a) \in X$ für alle $a \in A$ (d.h. g ist wirklich Abbildung von A nach X), denn:
Falls $a \in Y$, dann $g(a) = f(a) \in f(A) \subseteq X$. Falls $a \notin Y$, dann $a \notin Y_0 = A \setminus X$, also $a \in X$. Wegen $g(a) = a$ folgt wieder $g(a) \in X$.

g ist injektiv:
Seien $a_1, a_2 \in A$ mit $g(a_1) = g(a_2)$, z.z. $a_1 = a_2$. Das ist klar, falls beide $a_1, a_2 \notin Y$, denn dann $a_1 = g(a_1) = g(a_2) = a_2$, oder falls beide $a_1, a_2 \in Y$, denn dann $f(a_1) = g(a_1) = g(a_2) = f(a_2)$, also $a_1 = a_2$ (da f injektiv ist; dies wird nur hier benutzt). Wir zeigen, dass der Fall $a_1 \in Y$, $a_2 \notin Y$ (oder umgekehrt) nicht eintritt: sonst wäre $f(a_1) =$

$g(a_1) = g(a_2) = a_2$. Aber $a_1 \in Y_n$ für ein n, also ist $a_2 = f(a_1) \in f(Y_n) = Y_{n+1} \subseteq Y$, ein Widerspruch.

g ist surjektiv:
Sei $x \in X$. Gesucht ist ein Element $a \in A$ mit $g(a) = x$. Wenn $x \notin Y$, dann $g(x) = x$. Wenn $x \in Y$, dann existiert ein $n \geq 0$ mit $x \in Y_n$. Wegen $x \in X$ ist $x \notin A \setminus X = Y_0$, also $n > 0$. Dann ist $x \in Y_n = f(Y_{n-1})$, also existiert $a \in Y_{n-1}$ mit $x = f(a) = g(a)$.

1.16 Satz: *(Schröder–Bernstein)*

Seien $s : A \to B$ und $t : B \to A$ injektive Abbildungen. Dann existiert eine Bijektion $r : A \to B$.

Beweis: Es ist $f := t\,s : A \to A$ injektiv nach 1.14. Da $s(A) \subseteq B$, ist $f(A) = t(s(A)) \subseteq t(B) \subseteq A$. Nach 1.15 (mit $X = t(B)$) existiert eine Bijektion $g : A \to t(B)$. Sei $t_1 : B \to t(B)$ definiert durch $t_1(b) = t(b)$. Dann ist t_1 bijektiv (injektiv, da t injektiv, surjektiv, da $t_1(B) = t(B)$). Nach 1.13 hat t_1 eine Umkehrabbildung $u : t(B) \to B$, und u ist auch bijektiv. Nach 1.14 ist dann auch $ug : A \to B$ bijektiv.

1.17 Definition: *Relation, Äquivalenzrelation, Ordnungsrelation*

Sei A eine Menge. Eine Relation R auf A ist eine Teilmenge $R \subseteq A \times A$. Man schreibt meistens $aR\,b$ für $(a, b) \in R$.

Eine Relation R heißt

reflexiv,	falls $aR\,a\ \forall a \in A$,
symmetrisch,	falls $aR\,b \Rightarrow bR\,a$,
transitiv,	falls $aR\,b \wedge bR\,c \Rightarrow aR\,c$.
antisymmetrisch,	falls $aR\,b \wedge bR\,a \Rightarrow a = b$.

Eine Relation R, welche reflexiv, symmetrisch und transitiv ist, heißt Äquivalenzrelation. Eine Relation R, welche reflexiv, transitiv und antisymmetrisch ist, heißt Ordnungsrelation.

1.18 Beispiele, Bemerkung, Definition:

(i) $A = \mathbb{Z}, n \in \mathbb{N}$. Man nennt „$a$ kongruent zu b modulo n“ (für $a, b \in \mathbb{Z}$), falls n ein Teiler von $a - b$ ist, geschrieben als $a \equiv b \bmod n$. Kongruenz ist eine Äquivalenzrelation.

(ii) Sei X eine beliebige Menge, $A = \wp(X)$. Dann ist „$\subseteq$“eine Ordnungsrelation auf A. (Zur Erinnerung: Die Elemente von A sind Teilmengen von X.)

(iii) Wenn X in (ii) mehr als ein Element enthält, dann gibt es Teilmengen $A, B \subseteq X$, welche nicht miteinander vergleichbar sind, d.h. es gilt weder $A \subseteq B$ noch $B \subseteq A$.

(iv) Dagegen sind je zwei Elemente in $\mathbb{Z}$ unter der natürlichen Ordnung miteinander vergleichbar. Solche Mengen nennt man vollständig geordnet.

1.19 Definition: *Äquivalenzklasse*

Sei A eine Menge und $\sim$ eine Äquivalenzrelation auf A. Für $a \in A$ nennt man $[a] = [a]_\sim = \{x \in A \mid x \sim a\}$ die Äquivalenzklasse von a.

1.20 Bemerkung:

Sei $\sim$ eine Äquivalenzrelation auf A und $a, b \in A$. Es gelten

(1) $a \in [a]$

(2) $[a] = [b] \Leftrightarrow a \sim b$

(3) $[a] \cap [b] \neq \emptyset \Rightarrow [a] = [b]$

Beweis:

(1) gilt, da $a \sim a$.

(2) „$\Rightarrow$“: Nach (1) ist $a \in [a] = [b] = \{x \in A \mid x \sim b\}$, daher $a \sim b$.

„$\Leftarrow$“: Aus Symmetriegründen genügt z.z. $[a] \subseteq [b]$. Sei $c \in [a]$. Dann $c \sim a$ und $a \sim b$, also $c \sim b$, d.h. $c \in [b]$.

(3) Sei $c \in [a] \cap [b]$. Dann $c \sim a$, also $[a] = [c]$ nach (2). Ebenso folgt $[b] = [c]$, also $[a] = [b]$.

1.21 Bemerkung:

Typischerweise treten Äquivalenzrelationen im Zusammenhang mit Abbildungen auf: Wenn $f : A \to M$ eine Abbildung ist, dann wird durch $a \sim b \Leftrightarrow f(a) = f(b)$ eine Äquivalenzrelation auf A definiert.

2 Gruppen, Ringe, Körper

2.1 Definition: *Verknüpfung*

Eine Verknüpfung auf einer Menge M ist eine Abbildung $M \times M \to M$.

2.2 Beispiele:

(i) $M = \mathbb{N}$, Verknüpfung ist die Addition, $(a, b) \mapsto a + b$

(ii) $M = \mathbb{R}_{>0}$, Verknüpfung ist die Multiplikation, $(a, b) \mapsto ab$

(iii) X sei eine Menge, $M = \{$ alle bijektiven Abbildungen $X \to X\}$, Verknüpfung ist die Multiplikation von Abbildungen (1.6), $(f, g) \mapsto f \circ g$

(iv) $M = \mathbb{Z}$, $\mathbb{Q}$, oder $\mathbb{R}$, Verknüpfung ist die Addition oder die Multiplikation

(v) $M = \mathbb{Z}$, Verknüpfung ist die Subtraktion, $(a, b) \mapsto a - b$

(vi) $M = \mathbb{N}$, $(a, b) \mapsto a - b$ ist *keine* Verknüpfung auf $\mathbb{N}$, weil z.B. $1 - 2 \notin \mathbb{N}$

2.3 Definition: *Gruppe*

Eine Gruppe ist eine Menge G mit einer Verknüpfung $(x, y) \mapsto xy$, welche folgenden Axiomen genügt:

(1) (Assoziativität) $(xy)z = x(yz)\ \forall x, y, z \in G$

(2) (Neutrales Element) $\exists e \in G$ mit $ex = xe = x\ \forall x \in G$

(3) (Inverses Element) $\forall x \in G\ \exists y \in G$ mit $xy = yx = e$

G heißt kommutativ (oder abelsch), falls zusätzlich $xy = yx\ \forall x, y \in G$ (Kommutativität) gilt.

2.4 Beispiel:

Welche der in 2.2 gegebenen Beispiele ist eine Gruppe?

(i) ist keine Gruppe: zwar ist die Addition assoziativ, es gibt auch ein neutrales Element in $\mathbb{N}$, nämlich 0, aber im Allgemeinen kein Inverses.

(ii) ist eine Gruppe: neutrales Element ist 1, Inverses zu x ist $\frac{1}{x}$ (> 0, da $x > 0$), sogar abelsch.

(iii) ist eine Gruppe: assoziativ nach 1.8, neutrales Element ist id_X nach 1.7 (offenbar ist id_X bijektiv), inverses Element ist die Umkehrabbildung (existiert und ist auch bijektiv, also in M, nach 1.13). Diese Gruppe ist nicht abelsch, wenn X mindestens drei Elemente hat: Seien a, b, c drei verschiedene Elemente in X. Sei $f(a) = b$, $f(b) = a$, und $f(x) = x$ falls $x \in X, x \neq a, b$. Sei $g(a) = c$, $g(c) = a$, und $g(x) = x$ falls $x \in X, x \neq a, c$. Dann sind $f, g : X \to X$ bijektive Abbildungen, also $f, g \in M$. Es ist $(fg)(a) = f(g(a)) = f(c) = c$, aber $(gf)(a) = g(f(a)) = g(b) = b$, also $fg \neq gf$.

(iv) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ sind bzgl. der Addition abelsche Gruppen, neutrales Element ist 0, Inverses zu x ist $-x$.

Bezüglich der Multiplikation sind diese keine Gruppen: zwar ist die Multiplikation assoziativ, und es gibt ein neutrales Element (nämlich 1), aber es gibt nicht immer Inverse: in $\mathbb{Z}$ haben nur 1 und -1 Inverse bzgl. der Multiplikation (diese Elemente sind jeweils zu sich selbst invers), in $\mathbb{Q}$ und $\mathbb{R}$ haben alle Elemente außer Null ein Inverses. Daher sind $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$ und $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ Gruppen bzgl. der Multiplikation.

(v) ist keine Gruppe: die Verknüpfung ist nicht assoziativ, z.B. $(1-2)-3 = -4 \neq 2 = 1-(2-3)$.

2.5 Lemma:

Sei G eine (multiplikativ geschriebene) Gruppe (also $(x,y) \mapsto xy$). Dann gelten:

(1) Es gibt nur ein neutrales Element e in G.

(2) Zu jedem $x \in G$ gibt es nur ein Inverses, geschrieben als x^{-1}.

(3) Wenn $xy = xz$ oder $yx = zx$, dann $y = z$ (Kürzungsregel). Insbesondere:

- $xy = y \Leftrightarrow x = e$, und analog $yx = y \Leftrightarrow x = e$
- $xy = e \Leftrightarrow x = y^{-1}$
- $(x^{-1})^{-1} = x$

(4) $(xy)^{-1} = y^{-1}x^{-1}$

Beweis:

(1) Sei auch e' ein neutrales Element, d.h. $e'x = xe' = x \; \forall x \in G$. Dann $e = ee' = e'$.

(2) Vergleiche Beweis von 1.10

(3) Multiplikation mit x^{-1} von rechts bzw. links.

(4) $(y^{-1}x^{-1})(xy) = y^{-1}(x^{-1}x)y = y^{-1}ey = e$. Die Behauptung folgt jetzt aus (3).

2.6 Bemerkung: *zur Schreibweise*

Die additive Schreibweise $(a,b) \mapsto a+b$ wird nur für abelsche Gruppen verwendet. Das neutrale Element wird dann "Null" genannt und "0" geschrieben. Das inverse Element zu a wird $-a$ geschrieben. Für $a + (-b)$ schreibt man $a - b$.

Die multiplikative Schreibweise $(a,b) \mapsto ab$ wird sowohl für abelsche Gruppen (z.B. $\mathbb{Q}^*$) als auch für nicht–abelsche Gruppen verwendet. Das neutrale Element wird dann "Eins" genannt und "1" geschrieben. Das inverse Element zu a wird a^{-1} geschrieben. Bei abelschen Gruppen schreibt man manchmal $\frac{a}{b}$ für $ab^{-1} = b^{-1}a$.

2.7 Definition: *symmetrische Gruppe, Fixpunkt, Transposition*

Sei X eine Menge. Die Menge S_X aller bijektiven Abbildungen von X auf X heißt die symmetrische Gruppe (auf X). Die Elemente dieser Gruppe heißen auch Permutationen von X. Wenn $X = \{1, \ldots, n\}$, dann schreibt man auch S_n für S_X.

Sei nun $s \in S_X$.

(1) Man nennt x einen Fixpunkt von s, falls $s(x) = x$.

(2) Mann nennt s eine Transposition, falls $s(x) = x$ für alle $x \in X$ mit genau zwei Ausnahmen.

2.8 Satz:

Jedes $s \in S_n$ ist ein Produkt von Transpositionen.

Beweis: Induktion über m = Anzahl der Nicht–Fixpunkte von s.

Induktionsverankerung: $m = 0$. Dann ist jeder Punkt Fixpunkt, also $s = 1$ = Produkt von 0 Transpositionen (ein leeres Produkt). (Wenn es überhaupt Transpositionen gibt, also wenn $n > 1$, dann kann man auch $1 = t^2$ für eine beliebige Transposition $t \in S_n$ schreiben.)
Induktionsschritt: Sei die Behauptung für $k \leq m$ schon bewiesen, und sei $m + 1$ die Anzahl der Nicht–Fixpunkte von s. Da $s \neq 1$, gibt es ein $a \in \{1, \ldots, n\}$ mit $s(a) \neq a$. Sei $b = s(a)$. Definiert man t durch $t(a) = b$, $t(b) = a$ und $t(i) = i$, falls $i \neq a, b$, dann ist t eine Transposition.
Wenn j ein Fixpunkt von s ist, dann ist $j \neq a$, da $s(a) \neq a$, und $j \neq b$, denn sonst wäre $s(b) = b = s(a)$, also - da s injektiv - $a = b$, ein Widerspruch. Daher ist $st(j) = s(j) = j$, d.h. j ist auch ein Fixpunkt von st. Es gilt auch $st(b) = s(a) = b$, also hat st hat noch den zusätzlichen Fixpunkt b.
Mit anderen Worten: st hat mehr Fixpunkte als s, also weniger Nicht–Fixpunkte, d.h. höchstens m Nicht–Fixpunkte. Nach Induktionsvoraussetzung ist st ein Produkt von Transpositionen, etwa $st = t_1 t_2 \cdots t_r$. Dann ist $s = s1 = st^2 = t_1 t_2 \cdots t_r t$ ebenfalls ein Produkt von Transpositionen.

2.9 Bemerkung/Definition: *Signum, gerade Permutation, ungerade Permutation*

Sei $s \in S_n$ und $T = \{i, j\}$ eine zwei–elementige Teilmenge von $[n] := \{1, \ldots, n\}$, dann ist auch $s(T) = \{s(i), s(j)\}$ eine zwei–elementige Teilmenge und die Abbildung $T \mapsto s(T)$ ist eine Bijektion auf der Menge der zwei–elementigen Teilmengen von $[n]$. Daher ist

$$\operatorname{sign}(s) := \prod_{\substack{T=\{i,j\}\subseteq[n] \\ |T|=2}} \frac{s(i) - s(j)}{i - j} = \pm 1,$$

denn im Zähler wie im Nenner treten bis auf's Vorzeichen die gleichen Zahlen auf. Man nennt $\operatorname{sign}(s)$ das Signum von s.

Wenn $\operatorname{sign}(s) = 1$, so heißt s gerade Permutation, andernfalls ungerade Permutation.

Wenn auch $t \in S_n$, dann gilt

$$\operatorname{sign}(s) = \prod_{\substack{T=\{i,j\}\subseteq[n] \\ |T|=2}} \frac{st(i) - st(j)}{t(i) - t(j)},$$

denn dabei wird nur die Reihenfolge der Faktoren geändert, da mit T auch $t(T)$ alle zwei–elementigen Teilmengen von $[n]$ durchläuft.

2.10 Satz:

Für alle $s, t \in S_n$ gilt

$$\operatorname{sign}(s \circ t) = \operatorname{sign}(s) \cdot \operatorname{sign}(t).$$

Beweis:

$$\begin{aligned}\operatorname{sign}(s) \cdot \operatorname{sign}(t) &= \prod_T \frac{st(i) - st(j)}{t(i) - t(j)} \cdot \prod_T \frac{t(i) - t(j)}{i - j} \\ &= \prod_T \frac{st(i) - st(j)}{i - j} = \operatorname{sign}(st)\end{aligned}$$

2.11 Lemma:

Für eine Transposition t ist $\operatorname{sign}(t) = -1$.

Beweis: Sei t die Transposition, welche a und b vertauscht und alle anderen Elemente festhält. Wenn T eine zwei–elementige Teilmenge von $[n]$ ist, dann gibt es vier Möglichkeiten:

(1) $T = \{a, b\}$
(2) $T \cap \{a, b\} = \{a\}$ d.h. $T = \{a, i\}, i \neq a, b$
(3) $T \cap \{a, b\} = \{b\}$ d.h. $T = \{b, i\}, i \neq a, b$
(4) $T \cap \{a, b\} = \emptyset$

Daher ist

$$\begin{aligned}\operatorname{sign}(t) &= \frac{t(a) - t(b)}{a - b} \cdot \prod_{i \neq a,b} \frac{t(a) - t(i)}{a - i} \cdot \prod_{i \neq a,b} \frac{t(b) - t(i)}{b - i} \cdot \prod_{i,j \neq a,b} \frac{t(i) - t(j)}{i - j} \\ &= \frac{b - a}{a - b} \cdot \prod_{i \neq a,b} \frac{b - i}{a - i} \cdot \frac{a - i}{b - i} \cdot \prod_{i,j \neq a,b} \frac{i - j}{i - j} = -1\end{aligned}$$

2.12 Satz:

Sei $s \in S_n$ als Produkt von Transpositionen geschrieben, etwa $s = \prod_{i=1}^{r} t_i$. Dann ist $\operatorname{sign}(s) = (-1)^r$.

Beweis:

$$\begin{aligned}\operatorname{sign}(s) &= \prod_{i=1}^{r} \operatorname{sign}(t_i) \quad \text{nach 2.10 (und trivialer Induktion)} \\ &= \prod_{i=1}^{r} (-1) \qquad \text{nach 2.11} \\ &= (-1)^r.\end{aligned}$$

2.13 Korollar:

Sei $t_1 t_2 \cdots t_r = t'_1 t'_2 \cdots t'_s$ mit Transpositionen t_i, t'_j. Dann ist $r \equiv s \mod 2$.

2.14 Definition: *Homo–, Epi–, Mono–, Isomorphismus*

Seien G und H Gruppen. Ein Homomorphismus α ist eine Abbildung $\alpha : G \to H$ mit $\alpha(xy) = \alpha(x)\alpha(y)\ \forall x, y \in G$.

Ist der Homomorphismus α surjektiv, so heißt er Epimorphismus, ist er injektiv, so heißt er Monomorphismus, ist er bijektiv, so heißt er Isomorphismus.

2.15 Beispiele:

(i) id_G ist Isomorphismus.

(ii) Für beliebige Gruppen G und H ist $\varepsilon : G \to H$ definiert durch $\varepsilon(g) = 1_H\ \forall g \in G$ ein Homomorphismus.

(iii) Für $G = H = \mathbb{Z}$ (mit der Addition) und $t \in \mathbb{Z}$ ist $\mu_t(z) = tz$ ein Monomorphismus, falls $t \neq 0$, und ein Epimorphismus (also sogar ein Isomorphismus), falls $t = \pm 1$.

(iv) Für $G = S_n$ und $H = \{+1, -1\}$ ist $\mathrm{sign} : S_n \to \{+1, -1\}$ ein Epimorphismus, falls $n \geq 2$.

2.16 Definition: *Ring, Körper*

Eine Menge R mit zwei Verknüpfungen Addition $(+)$ und Multiplikation $(\cdot)$ heißt Ring, falls folgende Axiome gelten:

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

(2) $(R, \cdot)$ ist assoziativ, d.h. $(ab)c = a(bc)$ für alle $a, b, c \in R$.

(3) $(a + b)c = ac + bc$ und $c(a + b) = ca + cb$ für alle $a, b, c \in R$. (Diese Eigenschaft nennt man Distributivität.)

Wenn zusätzlich

(4) $ab = ba$ für alle $a, b \in R$

gilt, so heißt R kommutativer Ring.

Wenn zusätzlich

(5) $\exists e \in R$ mit $ea = ae = a$ für alle $a \in R$

gilt, so heißt R Ring mit Eins.

Ein Körper ist ein Ring K, so daß $K \setminus \{0\}$ bzgl. der Multiplikation eine abelsche Gruppe ist („0“ ist dabei das neutrale Element für die Addition).

2.17 Bemerkung:

(i) Ein Körper ist also ein kommutativer Ring mit $1 \neq 0$ derart, daß jedes Element $\neq 0$ ein multiplikatives Inverse hat. Ein Körper hat mindestens zwei Elemente.

(ii) Wenn R ein Ring mit Eins ist, dann ist diese eindeutig bestimmt (siehe Beweis von 2.5 (2)), man bezeichnet sie mit 1.

2.18 Beispiel:

(i) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sind Ringe, die letzten drei sogar Körper. Diese Ringe sind alle kommutativ. Beispiele für nicht–kommutative Ringe werden wir später kennenlernen.

(ii) $R[x] = \{$Polynome in einer Unbestimmten mit Koeffizienten aus einem Ring R (mit Eins)$\}$ mit der üblichen Addition und Multiplikation von Polynomen ist ein Ring (mit Eins).

2.19 Lemma:

Sei R ein Ring. Dann gelten:

(1) $a \cdot 0 = 0 \cdot a = 0$ für alle $a \in R$ ('0' ist - wie immer - das neutrale Element bzgl. der Addition)

(2) $(-a) \cdot b = -(a \cdot b) = a \cdot (-b)$ für alle $a, b \in R$

(3) $(-a) \cdot (-b) = a \cdot b$ für alle $a, b \in R$

Falls R ein Körper ist, gilt auch die Umkehrung von (1), d.h.

(4) $a \cdot b = 0 \Leftrightarrow a = 0 \vee b = 0$

Beweis:

(1) Es ist $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$, also $a \cdot 0 = 0$ nach 2.5 (3). Analog zeigt man $0 \cdot a = 0$.

(2) Es ist $0 = 0 \cdot b = (a + (-a)) \cdot b = a \cdot b + (-a) \cdot b$, also $(-a) \cdot b = -(a \cdot b)$ nach 2.5 (3). Analog zeigt man $a \cdot (-b) = -(a \cdot b)$.

(3) Es ist $(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b))$ nach (2), also $(-a) \cdot (-b) = a \cdot b$ nach 2.5 (3).

(4) Sei $a \cdot b = 0$, zu zeigen ist $a = 0$ oder $b = 0$. Wenn $a \neq 0$, dann hat a ein Inverses a^{-1} bzgl. der Multiplikation, da R nach Voraussetzung ein Körper ist. Dann gilt $0 = a^{-1} \cdot 0 = a^{-1} \cdot (a \cdot b) = (a^{-1} \cdot a) \cdot b = 1 \cdot b = b$.

3 Vektorräume

3.1 Definition: *Vektorraum*

Sei $\mathbb{K}$ ein Körper. Ein Vektorraum V über $\mathbb{K}$ ist eine abelsche Gruppe $(V,+)$ zusammen mit einer Abbildung $\mathbb{K}\times V\to V$, geschrieben $(k,v)\mapsto k\cdot v$ für $(k,v)\in\mathbb{K}\times V$, mit folgenden Axiomen:

(1) (Assoziativität) $(k_1k_2)\cdot v = k_1\cdot(k_2\cdot v)\ \forall k_1,k_2\in\mathbb{K}\ \forall v\in V$

(2) (Einselement) $1_{\mathbb{K}}\cdot v = v\ \forall v\in V$

(3) (Distributivität) $(k_1+k_2)\cdot v = k_1\cdot v + k_2\cdot v\ \forall k_1,k_2\in\mathbb{K}\ \forall v\in V$ und $k\cdot(v_1+v_2) = k\cdot v_1 + k\cdot v_2\ \forall k\in\mathbb{K}\ \forall v_1,v_2\in V$

3.2 Bezeichnung/Bemerkung:

Die Elemente in $\mathbb{K}$ nennt man Skalare. Ein Element aus V heißt Vektor. Wenn man einen Vektor mit einem Skalar multipliziert, erhält man also wieder einen Vektor. Wir haben in diesem Produkt den Skalar links vor den Vektor geschrieben und sollten daher genau genommen von einem Linksvektorraum sprechen. Rechtsvektorräume sind analog definiert. „Vektorraum“ heißt in dieser Vorlesung immer „Linksvektorraum“. Für „Sei V ein Linksvektorraum über $\mathbb{K}$“ sagen wir kürzer „Sei V ein $\mathbb{K}$–Vektorraum“ oder noch kürzer „Sei V ein $\mathbb{K}$–VR“. Das gibt nur einen Sinn, falls $\mathbb{K}$ ein Körper ist.

Warnung: Die Addition in V wird mit dem gleichen Symbol „+“ bezeichnet wie die Addition in $\mathbb{K}$. Die neutralen Elemente in diesen beiden abelschen Gruppen werden beide „Null“ genannt (sind aber voneinander zu unterscheiden).

Die Multiplikation von zwei Skalaren gibt wieder einen Skalar (das ist einfach die Multiplikation in $\mathbb{K}$), die Multiplikation von einem Skalar mit einem Vektor ergibt einen Vektor. Die Multiplikation von zwei Vektoren ist gar nicht definiert.

3.3 Beispiel:

(i) $\mathbb{K}=\mathbb{R}$, $V=\mathbb{R}\times\mathbb{R}=\{(a,b)\,|\,a,b\in\mathbb{R}\}$ (mit der „komponentenweisen“ Verknüpfung)

(ii) $\mathbb{K}=\mathbb{R}$, $V=\{f:\mathbb{R}\to\mathbb{R}\}$ die Menge aller reellen Funktionen

(iii) $\mathbb{K}=\mathbb{R}$, $V=\{f:\mathbb{R}\to\mathbb{R}\,|\,f \text{ ist stetig}\}$

(iv) $\mathbb{K}=\mathbb{R}$, $V=\mathbb{R}[x]$

(v) $\mathbb{K}$ beliebiger Körper, $V=\mathbb{K}^n=\{(a_1,\dots,a_n)\,|\,a_i\in\mathbb{K}\}$

(vi) $\mathbb{K}=\mathbb{R}$, $V=\{\text{alle reellen Lösungen der Gleichung } 2x-y-3z=0\}$, z.B. $x=y=z=0$, oder $x=1,\ y=-1,\ z=1$, oder $x=2,\ y=1,\ z=1$

(vii) Punkte der Ebene, wobei ein Punkt als 'Koordinatenursprung' ausgezeichnet ist.

Die fehlenden Verknüpfungen zu erraten, sollte leicht sein.

3.4 Lemma:

Sei V ein $\mathbb{K}$–VR und $v \in V$, $k \in \mathbb{K}$. Dann gelten:

(1) $kv = 0 \Leftrightarrow k = 0 \vee v = 0$

(2) $(-1)v = -v$

Beweis:

(1)„$\Leftarrow$“: Sei $k = 0$. Dann ist $0v = (0+0)v = 0v + 0v$, also $0v = 0$ nach 2.5. Sei $v = 0$. Dann ist $k0 = k(0+0) = k0 + k0$, also $k0 = 0$, wieder nach 2.5.

„$\Rightarrow$“: Fertig, falls $k = 0$. Wenn $k \neq 0$, dann existiert das Inverse k^{-1} zu k in $\mathbb{K}$. Dann ist $0 = k^{-1}0 = k^{-1}(kv) = (k^{-1}k)v = 1v = v$.

(2) Es ist $0 = 0v = (1 + (-1))v = 1v + (-1)v = v + (-1)v$, also $(-1)v = -v$ nach 2.5.

3.5 Definition: *Unterraum*

Sei V ein $\mathbb{K}$–VR. Eine Teilmenge U von V heißt ein Unterraum von V, falls U selbst ein $\mathbb{K}$–VR ist mit der Addition und skalaren Multiplikation wie in V (geschrieben $U \leq V$).

3.6 Beispiel:

(i) $V = \mathbb{R}^3$, $U = \{(a,b,c) \in \mathbb{R}^3 \mid b = 0\}$

(ii) U in 3.3 (iii) ist ein Unterraum von V in 3.3 (ii).

(iii) V in 3.3 (vi) ist ein Unterraum von $\mathbb{R}^3$.

(iv) In 3.3 (iv) ist für jedes $d \in \mathbb{N}$ die Menge $U_d = \{p \in \mathbb{K}[x] \mid \ deg(p) < d\}$ ein Unterraum von $\mathbb{K}[x]$. (Der Grad eines Polynoms $p = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ ist definiert als $deg(p) = n$, falls $a_n \neq 0$. Der Grad des Nullpolynoms wird als $-\infty$ definiert.)

(v) $\{0\}$ und V sind Unterräume von V. Diese beiden werden manchmal die trivialen Unterräume genannt. $\{0\}$ heißt Nullraum.

(vi) Die Geraden durch den Nullpunkt sind Unterräume der Ebene in 3.3(vii).

3.7 Satz: *Unterraumkriterium*

Sei V ein $\mathbb{K}$–VR und U eine Teilmenge von V. Genau dann ist U ein Unterraum von V, wenn

(1) (Nullelement) $0_V \in U$

(2) (abgeschlossen unter „+“) $a, b \in U \Rightarrow a + b \in U$

(3) (abgeschlossen unter Multiplikation mit Skalaren) $k \in \mathbb{K},\ a \in U \Rightarrow ka \in U$

Beweis:

„$\Rightarrow$“: Sei $U \leq V$. Dann ist $(U, +)$ eine abelsche Gruppe, also ist $a + b \in U$ für $a, b \in U$, d.h. (2) gilt. Außerdem enthält U ein neutrales Element 0_U. Da $0_U + 0_U = 0_U$, folgt nach 2.5 (1) $0_V = 0_U \in U$, d.h. (1) gilt. Da schließlich U ein $\mathbb{K}$–VR ist, muß $ka \in U$ sein für alle $a \in U, k \in \mathbb{K}$, d.h. (3) gilt.

„⇐“: Umgekehrt seien (1)–(3) vorausgesetzt. Zu zeigen ist $U \leq V$.

Zunächst ist $(U, +)$ eine abelsche Gruppe, denn nach (2) ist „+“ eine Verknüpfung auf U, diese ist assoziativ und kommutativ, da dies sogar auf ganz V gilt. Nach (1) gibt es ein neutrales Element. Nach (3) ist für $a \in U$ auch $(-1)a \in U$, d.h. $-a \in U$ nach 3.4 (2), also hat jedes Element von U ein Inverses in U bzgl. „+“.

Nach (3) ist die Multiplikation mit Skalaren eine Abbildung $\mathbb{K} \times U \to U$. Die VR–Axiome gelten dann trivialerweise.

3.8 Definition: *Summen*

Für Teilmengen A und B des $\mathbb{K}$–VR V ist die Summe definiert durch

$$A + B := \{a + b \mid a \in A, b \in B\} \ .$$

Wenn $B = \{b\}$, schreibt man dafür einfach $A + b$. Wenn nun $U_i \leq V$, $i \in I$, so sei die Summe der U_i definiert durch

$$\sum_{i \in I} U_i := \left\{ \sum_{i \in I} u_i \mid u_i \in U_i, \text{ fast alle } u_i = 0 \right\} .$$

„Fast alle“ bedeutet „alle bis auf höchstens endlich viele Ausnahmen“.

3.9 Satz:

Schnitte und Summen von Unterräumen sind wieder Unterräume.

Beweis: Seien U_i, $i \in I$, Unterräume von V. Zu zeigen ist:

(a) $\bigcap_{i \in I} U_i \leq V$

(b) $\sum_{i \in I} U_i \leq V$.

Dazu verwenden wir das Unterraumkriterium.

(a) (1) $0 \in \bigcap_{i \in I} U_i$, da $0 \in U_i$ für alle $i \in I$.

(2) Wenn $a, b \in \bigcap_{i \in I} U_i$, dann $a, b \in U_i$ für alle $i \in I$, also auch $a + b \in U_i$ für alle $i \in I$ und daher $a + b \in \bigcap_{i \in I} U_i$.

(3) Wenn $k \in \mathbb{K}$, $a \in \bigcap_{i \in I} U_i$, dann $a \in U_i$ für alle $i \in I$, also auch $ka \in U_i$ für alle $i \in I$, d.h. $ka \in \bigcap_{i \in I} U_i$.

(b) (1) $0 = \sum_{i \in I} 0 \ \in \sum_{i \in I} U_i$, da $0 \in U_i$.

(2) Wenn $a = \sum_{i \in I} u_i$ und $b = \sum_{i \in I} u_i'$ in $\sum_{i \in I} U_i$, fast alle $u_i, u_i' = 0$, dann ist $u_i + u_i' = 0$ für fast alle i und $u_i + u_i' \in U_i$ für alle i, also ist $a + b = \sum_{i \in I} (u_i + u_i') \in \sum_{i \in I} U_i$.

(3) Wenn $a = \sum_{i \in I} u_i \in \sum_{i \in I} U_i$, fast alle $u_i = 0$, und $k \in \mathbb{K}$, dann ist $ka = \sum_{i \in I} ku_i \ \in \sum_{i \in I} U_i$, denn $ku_i \in U_i$ und $ku_i = 0$ für fast alle i.

3.10 Definition: *Erzeugnis*

Sei T eine Teilmenge des $\mathbb{K}$–VR V, dann heißt

$$\langle T\rangle := \bigcap_{U\leq V,\, T\subseteq U} U$$

der von T erzeugte Unterraum von V oder einfach das Erzeugnis von T.

3.11 Bemerkung:

Man schneidet also alle Unterräume von V, welche T enthalten (solche gibt es, z.B. V selbst). Der Schnitt enthält immer noch T und ist ein Unterraum nach 3.9. Wenn W ein beliebiger Unterraum mit $T \subseteq W$ ist, dann ist W „am Schnitt beteiligt“, also $W \geq \langle T\rangle$. Daher ist das Erzeugnis von T der kleinste Unterraum von V, welcher T enthält. Insbesondere ist $\langle\emptyset\rangle = \{0\}$.

3.12 Definition: *Erzeugenden–System*

Sei T eine Teilmenge des $\mathbb{K}$–VR V. Man nennt T ein Erzeugenden–System (von V), falls $\langle T\rangle = V$. Man nennt V endlich erzeugt, falls V ein endliches Erzeugenden–System besitzt.

3.13 Beispiel:

(i) $\langle V\rangle = V$, jeder VR hat also ein Erzeugenden–System.

(ii) $\langle\emptyset\rangle = \langle\{0\}\rangle = \{0\}$, der Nullraum.

(iii) $\langle\{(1,0),\ (0,1)\}\rangle = \mathbb{K}^2$, wie man leicht sieht. Entsprechend hat $\mathbb{K}^n$ ein Erzeugenden–System mit n Elementen.

(iv) Dagegen ist der Vektorraum $\mathbb{K}[x]$ nicht endlich erzeugt, da je endlich viele Polynome in einem echten Unterraum U_d wie in 3.6, (iv) enthalten sind. Ein naheliegendes Erzeugenden–System für diesen Vektorraum ist die Menge $\{1 = x^0,\ x = x^1,\ x^2,\ x^3,\ \ldots\}$.

4 Basis und Dimension

Sei V ein Vektorraum über dem Körper $\mathbb{K}$.

4.1 Definition: *Linearkombination, Koeffizient*

Sei X eine Teilmenge von V. Ein Vektor v der Form

$$v = \sum_{x \in X} k_x x$$

(höchstens endlich viele $k_x \neq 0$) heißt eine Linearkombination von X. In dieser Summe heißt k_x der Koeffizient von x.

4.2 Beispiel:

(i) $V = \mathbb{R}^3,\ X = \{(1,2,0),\ (1,-1,0),\ (2,2,0)\}$.
Nicht alle Vektoren von $\mathbb{R}^3$ sind Linearkombinationen von X. Der Vektor

$$(a,b,0) = (-a-b)\cdot(1,2,0) + (-b)\cdot(1,-1,0) + (a+b)\cdot(2,2,0)$$

ist Linearkombination von X. Es gilt auch

$$(a,b,0) = \frac{a+b}{3}\cdot(1,2,0) + \frac{2a-b}{3}\cdot(1,-1,0) + 0\cdot(2,2,0)\ ,$$

man kann also u.U. denselben Vektor auf verschiedene Weisen als Linearkombination von X schreiben.

(ii) 0 ist Linearkombination von X für jedes X (z.B. ganz einfach alle $k_x = 0$).

(iii) Jedes $x \in X$ ist Linearkombination von X ($k_x = 1$, $k_y = 0$ für $x \neq y \in X$).

4.3 Satz:

Für jede Teilmenge X von V ist $U := \{\text{alle Linearkombinationen von } X\}$ ein Unterraum von V, nämlich $U = \langle X \rangle$.

Beweis: Mit dem Unterraumkriterium:

(1) $0 \in U$ nach 4.2 (ii).

(2) Wenn $u = \sum_{x\in X} k_x x$ und $u' = \sum_{x\in X} k'_x x$ in U sind, dann ist auch $u+u' = \sum_{x\in X}(k_x + k'_x)x$ in U, denn $k_x + k'_x \in \mathbb{K}$ und $k_x + k'_x = 0$ für fast alle $x \in X$.

(3) Für u wie oben und $k \in \mathbb{K}$ ist $ku = \sum_{x\in X}(kk_x)x$ mit $kk_x \in \mathbb{K}$ und $kk_x = 0$ für fast alle $x \in X$, also $ku \in U$.

Daher ist U ein Unterraum von V. Außerdem ist $X \subseteq U$ nach 4.2 (iii). Da $\langle X \rangle$ der kleinste Unterraum von V ist, welcher X enthält (siehe 3.11), folgt $\langle X \rangle \leq U$. Andererseits ist $X \subseteq \langle X \rangle$ und da $\langle X \rangle$ ein Vektorraum ist, liegen auch alle Linearkombinationen von X in $\langle X \rangle$, d.h. $U \leq \langle X \rangle$. Also ist $U = \langle X \rangle$.

4.4 Korollar:

Genau dann ist eine Teilmenge E von V ein Erzeugenden–System von V, wenn jeder Vektor aus V sich als Linearkombination von E schreiben läßt.

4.5 Definition: *lineare Unabhängigkeit, Basis*

(1) Eine Teilmenge X von V heißt linear unabhängig (l.u.), wenn

$$0 = \sum_{x \in X} k_x x,\ k_x \in \mathbb{K} \Rightarrow k_x = 0 \text{ für alle } x \in X,$$

d.h. wenn es nur eine einzige Möglichkeit gibt, 0 als Linearkombination von X zu schreiben (nämlich alle Koeffizienten $= 0$). Andernfalls heißt X linear abhängig (l.a.).

(2) Eine Teilmenge B von V heißt eine Basis von V, falls B ein linear unabhängiges Erzeugenden–System ist.

4.6 Beispiel/Bemerkung/Bezeichnung:

(i) Sei $V = \mathbb{K}^n$ und sei $e_1 := (1, 0, \ldots, 0)$, $e_2 := (0, 1, 0, \ldots, 0)$ und allgemein

$$e_i := (0, \ldots, 0, \underset{\substack{\uparrow \\ i}}{1}, 0, \ldots, 0)$$

für $i = 1, \ldots, n$. Dann ist $\sum_{i=1}^n k_i e_i = (k_1, k_2, \ldots, k_n)$. Also läßt sich jeder Vektor aus $\mathbb{K}^n$ als Linearkombination von $B := \{e_1, \ldots, e_n\}$ schreiben, d.h. B ist ein Erzeugenden–System. Wenn $\sum_{i=1}^n k_i e_i = 0 = (0, \ldots, 0)$, dann sind alle Koeffizienten $k_i = 0$, d.h. B ist linear unabhängig. Also ist B eine Basis von $\mathbb{K}^n$. Man nennt B oft die Standardbasis von $\mathbb{K}^n$. Der Vektor e_i heißt der i-te Einheitsvektor von $\mathbb{K}^n$.

(ii) Falls $k_x \in \mathbb{K}$ existieren mit $\sum_{x \in X} k_x x = 0$, fast alle $k_x = 0$, aber wenigstens ein $k_x \neq 0$, dann ist X nicht linear unabhängig, sondern linear abhängig. Insbesondere: Wenn $0 \in X$, dann ist X linear abhängig. Aber auch die Menge X aus 4.2 (i) ist linear abhängig, denn z.B. gilt

$$0 = \frac{4}{3}(1, 2, 0) + \frac{2}{3}(1, -1, 0) + (-1)(2, 2, 0).$$

(iii) Die leere Menge ist linear unabhängig.

(iv) Im allgemeinen gibt es mehr als eine Basis für einen Vektorraum.
Zum Beispiel ist $B := \{(1, 1, 0),\ (1, 0, 1),\ (0, 1, 1)\}$ eine Basis von $\mathbb{R}^3$. Denn wenn

$$a \cdot (1, 1, 0) + b \cdot (1, 0, 1) + c \cdot (0, 1, 1) = 0 = (0, 0, 0),$$

dann ist

$$\begin{aligned} a + b &= 0 \\ a + c &= 0 \\ b + c &= 0, \end{aligned}$$

also $b = c = -a$, $2a = 0$, $a = 0 = b = c$. Also ist B linear unabhängig. Außerdem gilt

$$\begin{aligned}&\frac{1}{2}(a+b-c)\cdot(1,1,0)+\frac{1}{2}(a-b+c)\cdot(1,0,1)+\frac{1}{2}(-a+b+c)\cdot(0,1,1)\\ =&\frac{1}{2}\cdot(a+b-c+a-b+c, a+b-c-a+b+c, a-b+c-a+b+c)\\ =&\frac{1}{2}\cdot(2a,2b,2c)=(a,b,c).\end{aligned}$$

Also ist jedes $(a,b,c)\in\mathbb{R}^3$ eine Linearkombination von B, d.h. B ist auch Erzeugenden–System.

Diese Basis hat drei Elemente, genauso wie die Standardbasis von $\mathbb{R}^3$. Das ist kein Zufall!

(v) Eine Basis von $\mathbb{K}[x]$ (iv)) ist die Menge $B = \{x^n \mid n\in\mathbb{N}\}$ (siehe 3.13).

(vi) $X = \{v\}$ ist genau dann linear unabhängig, wenn $v \neq 0$.

(vii) Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.

4.7 Lemma:

Sei B eine Basis von V. Dann ist jedes $v\in V$ auf genau eine Weise Linearkombination von B.

Beweis: ist leichte Übungsaufgabe.

4.8 Lemma:

Sei L linear unabhängig und $v\in V$. Dann sind folgende Aussagen äquivalent:

(1) $v \notin \langle L\rangle$

(2) $v \notin L$ und $L\cup\{v\}$ ist linear unabhängig

Beweis:

(1)$\Rightarrow$(2) Da $L \subseteq \langle L\rangle$, ist $v \notin L$. Sei $kv + \sum_{x\in L} k_x x = 0$. Zu zeigen ist $k = k_x = 0$ für alle x. Wäre $k \neq 0$, dann multipliziere mit k^{-1}. Dies ergibt $v = \sum_{x\in L}(-k^{-1}k_x)x$, d.h. v ist Linearkombination von L, also $v \in \langle L\rangle$ (4.3), Widerspruch. Also ist $k = 0$. Daher ist $\sum_{x\in L} k_x x = 0$. Da L linear unabhängig ist, müssen alle $k_x = 0$ sein.

(2)$\Rightarrow$(1) Angenommen $v \in \langle L\rangle$. Dann ist $v = \sum_{x\in L} k_x x$, also $1v + \sum_{x\in L}(-k_x)x = 0$. Aber der Koeffizient von v in dieser Linearkombination ist $1 \neq 0$, da $v \notin L$. Dies ist ein Widerspruch zur linearen Unabhängigkeit von $L\cup\{v\}$.

4.9 Lemma: *Austausch–Lemma von Steinitz*

Sei $L \subseteq V$ linear unabhängig und E ein Erzeugenden–System von V, sei $t \in L$. Dann existiert ein $s\in E$ derart, daß $s \notin L' = L \setminus \{t\}$ und $L'\cup\{s\}$ linear unabhängig ist. (t wird gegen $s\in E$ ausgetauscht.)

Beweis: Als Teilmenge von L ist L' linear unabhängig (siehe 4.6 (vii)), $t \notin L'$ und $L = L'\cup\{t\}$ ist linear unabhängig. Nach 4.8 folgt $t \notin \langle L'\rangle$. Da E ein Erzeugenden–System ist, läßt sich t als Linearkombination von E schreiben, etwa $t = \sum_{y\in E} k_y y$. Wären alle

$y \in \langle L' \rangle$, dann auch diese Linearkombination, da $\langle L' \rangle$ ein Unterraum ist, d.h. $t \in \langle L' \rangle$, Widerspruch. Also gibt es ein $s \in E$ mit $s \notin \langle L' \rangle$. Nach 4.8 folgt: $s \notin L'$ und $L' \cup \{s\}$ ist linear unabhängig.

4.10 Satz: *Austausch–Satz von Steinitz*

Sei $L \subseteq V$ linear unabhängig und E ein Erzeugenden–System von V. Weiter sei T eine endliche Teilmenge von L mit n Elementen; wir schreiben $L' = L \setminus T$. Dann existiert eine ebenfalls n-elementige Teilmenge $S \subseteq E$ derart, daß $L' \cap S = \emptyset$ und $L' \cup S$ linear unabhängig ist. (T wird gegen S ausgetauscht.)

Beweis: Induktion über n. Der Fall $n = 0$ ist trivial. Sei $n > 0$ und $T = T' \dot\cup \{t\}$; dann ist also $L = L' \cup T' \cup \{t\}$. Nach dem Austausch–Lemma gibt es ein $s \in E$ derart, daß $s \notin L' \cup T'$ und $L \cup T' \cup \{s\}$ linear unabhängig ist.
Da T' nur $n-1$ Elemente hat, finden wir nach Induktionsvoraussetzung eine Teilmenge S' von E mit $n-1$ Elementen derart, daß $(L' \cup \{s\}) \cap S' = \emptyset$ und $(L' \cup \{s\}) \cup S'$ linear unabhängig ist. Daher hat $S = S' \cup \{s\}$ die gewünschten Eigenschaften.

4.11 Korollar:

Wenn V ein Erzeugenden–System mit n Elementen hat und $L \subseteq V$ linear unabhängig ist, dann hat L höchstens n Elemente.

Beweis: Andernfalls hätte L eine $(n+1)-$elementige Teilmenge T. Nach 4.10 findet man zu dieser eine $(n+1)-$elementige Teilmenge $S \subseteq E$, ein Widerspruch, da E nur n Elemente hat.

4.12 Voraussetzung/Bemerkung: *Endlichkeit*

Man kann eine Variante des Steinitz'schen Austausch–Satzes formulieren (und beweisen!, allerdings mit mehr Aufwand), die auch für unendliches T gilt. Wir haben es uns leicht gemacht und wollen dies auch im Folgenden tun:
Wir betrachten nur noch Vektorräume, für welche es eine **endliche Maximalzahl** linear unabhängiger Elemente gibt. Es soll also zu dem Vektorraum V eine natürliche Zahl $d = d(V)$ geben derart, daß eine linear unabhängige Teilmenge $L \subseteq V$ höchstens d Elemente hat. Indem man ein solches d möglichst klein wählt, kann man annehmen, daß es in V wirklich eine linear unabhängige Teilmenge mit d Elementen gibt.
Die eben gemachte Annahme ist nach 4.11 jedenfalls für endlich–erzeugte Vektorräume erfüllt. Andererseits werden dadurch manche Vektorräume (z.B. $\mathbb{K}[x]$, vergleiche 4.6(v)) von der Betrachtung ausgeschlossen, die einen Mathematiker durchaus interessieren können.

4.13 Definition: *Dimension*

Sei V ein Vektorraum. Wenn es in V eine linear unabhängige Teilmenge mit d Elementen, aber keine solche Teilmenge mit $d+1$ Elementen gibt, dann heißt d die Dimension von V, geschrieben $d = \dim\ V = \dim_{\mathbb{K}} V$.

4.14 Satz:

Sei $d = \dim\ V$ und $B \subseteq V$. Dann sind folgende Aussagen äquivalent:

(1) B hat d Elemente und ist linear unabhängig.

(2) B ist eine Basis von V.

(3) B hat d Elemente und ist ein Erzeugenden–System von V.

Beweis:

(1)⇒(2) Es ist nur fraglich, ob $\langle B\rangle = V$. Andernfalls gibt es $v \in V \setminus \langle B\rangle$. Nach 4.8 ist dann $B \cup \{v\}$ eine linear unabhängige Menge mit $d+1$ Elementen, ein Widerspruch zu $d = \dim\ V$.

(2)⇒(3) Es ist nur fraglich, ob B genau d Elemente hat. Mehr kann es nicht haben, weil es linear unabhängig ist (als Basis). Es kann aber auch nicht weniger haben, denn es gibt ja wenigstens eine linear unabhängige Menge mit d Elementen. Nach 4.11 hat dann jedes Erzeugenden–System, insbesondere also B, mindestens d Elemente.

(3)⇒(1) Es ist nur fraglich, ob B linear unabhängig ist. Jedenfalls gibt es eine linear unabhängige Teilmenge L von V mit genau d Elementen. Jetzt kann man den Austausch–Satz mit $T = L$ und $E = B$ verwenden und findet eine d-elementige Teilmenge S von B (also $S = B$) mit $(L \setminus T) \cup S$ linear unabhängig, also B linear unabhängig.

4.15 Beispiel:

(i) $\dim\ \mathbb{K}^n = n$, denn die Standardbasis hat n Elemente.

(ii) In 4.6 (iv) haben wir gesehen, daß $\{(1,1,0),\ (1,0,1),\ (0,1,1)\}$ eine Basis von $\mathbb{R}^3$ ist. Dabei haben wir kontrolliert, daß diese Menge linear unabhängig und ein Erzeugenden–System von $\mathbb{R}^3$ ist. Nach 4.14 haben wir uns mehr Arbeit als nötig gemacht.

(iii) $\dim V = 0 \Leftrightarrow V = \{0\}$ ist der Nullraum.

4.16 Korollar:

Jeder Vektorraum hat eine Basis.

Beweis: Es gibt nach Voraussetzung eine linear unabhängige Teilmenge von V mit $\dim V$ Elementen. Diese ist eine Basis nach dem vorigen Satz.

4.17 Lemma:

Sei $X \subseteq V$ und E ein Erzeugenden–System von V. Wenn jedes $e \in E$ eine Linearkombination von X ist, dann ist auch X ein Erzeugenden–System von V.

Beweis: Nach Voraussetzung ist $E \subseteq \langle X\rangle$. Daher ist $V = \langle E\rangle \leq \langle X\rangle \leq V$.

4.18 Satz:

Sei $L \subseteq V$ linear unabhängig und E ein Erzeugenden–System von V. Dann existiert ein $T \subseteq E$ derart, daß $L \cap T = \emptyset$ und $L \cup T$ eine Basis von V ist.

Beweis: Wir wählen aus E eine möglichst große Teilmenge T derart, daß

$$L \cap T = \emptyset$$

und

$$L \cup T \text{ linear unabhängig.}$$

Das geht, denn jedenfalls erfüllt $T = \emptyset$ beide Bedingungen. Wenn dieses T vergrößert werden kann, fügt man ein geeignetes Element hinzu. Dies wiederholt man, bis keine Vergrößerung von T mehr möglich ist, ohne eine der Bedingungen zu verletzen. Da nach Voraussetzung $L \cup T$ höchstens $d(V)$ Elemente hat, endet dieser Prozeß.
Ein so gewonnenes T tut das gewüschte: Es ist nur zu zeigen, daß $L \cup T$ ein Erzeugenden–System ist. Nach 4.17 genügt es, $e \in \langle L \cup T \rangle$ für jedes $e \in E$ zu kontrollieren. Wenn dies für ein e nicht gilt, dann ist nach 4.8 sowohl $e \notin L \cup T$ als auch $L \cup T \cup \{e\}$ linear unabhängig. Aber dann kann man T durch Zufügen von e vergrößern, ein Widerspruch.

4.19 Korollar:

Es gelten:

(1) Jede linear unabhängige Teilmenge von V kann zu einer Basis ergänzt werden.

(2) Jedes Erzeugenden–System von V kann auf eine Basis reduziert werden.

Beweis:

(1) Verwende $E = V$, um die linear unabhängige Menge L zu einer Basis zu ergänzen.

(2) Ergänze die linear unabhängige Menge $\emptyset$ mit einer Teilmenge des Erzeugenden–System zu einer Basis.

4.20 Korollar:

Sei $\dim V = d$. Dann gelten:

(1) Jede linear unabhängige Teilmenge von V hat höchstens d Elemente.

(2) Jedes Erzeugenden–System von V hat mindestens d Elemente.

(3) Sei $U \leq V$, dann ist $\dim U \leq \dim V$, und $U = V \Leftrightarrow \dim U = \dim V$.

Beweis: Jede Basis von V hat d Elemente (nach 4.14).

(1) Nach 4.19 (1) läßt sich jede linear unabhängige Teilmenge zu einer Basis (mit d Elementen) ergänzen.

(2) Nach 4.19 (2) läßt sich jedes Erzeugenden–System zu einer Basis (mit d Elementen) abmagern.

(3) Eine Basis B von U ist eine linear unabhängige Teilmenge von V, nach (1) ist also $\dim U \leq d$. Wenn $\dim U = d$, dann hat B ebenfalls d Elemente. Nach 4.14 ist B schon eine Basis von V, also $U = \langle B \rangle = V$. Die Umkehrung ist trivial.

5 Faktorraum und lineare Abbildungen

Wieder sei V ein Vektorraum über dem Körper $\mathbb{K}$.

5.1 Definition: *Nebenklasse*

Sei $U \leq V$. Für jedes $x \in V$ nennt man die Menge

$$x + U := \{x + u \mid u \in U\}$$

eine Neben– oder Restklasse (von U in V); x heißt (Nebenklassen–) Vertreter von $x + U$.

5.2 Beispiel:

Sei U eine Gerade durch den Nullpunkt in der Ebene. Die Nebenklassen von U sind die zu U parallelen Geraden.

5.3 Lemma:

Sei $U \leq V$, $x, y \in V$. Es gelten:

(1) $U = 0 + U$ ist eine Nebenklasse.

(2) Äquivalent sind:
   (i) $(x + U) \cap (y + U) \neq \emptyset$
   (ii) $x - y \in U$
   (iii) $x + U = y + U$

(3) $x + U = U \Leftrightarrow x \in U$

(4) $(x + U) + (y + U) = (x + y) + U$, die Summe von zwei Nebenklassen ist also wieder eine Nebenklasse.

Beweis:

(1) trivial

(2) (i)⇒(ii): Sei $x + u_1 = y + u_2 \in (x + U) \cap (y + U)$, $u_1, u_2 \in U$. Dann ist $x - y = u_2 - u_1 \in U$.
(ii)⇒(iii): Sei $u \in U$. Dann ist $x + u = y + (x - y) + u \in y + U$, da $x - y, u \in U$, also $x - y + u \in U$. Also ist $x + U \subseteq y + U$. Genauso folgt die umgekehrte Inklusion, da aus $x - y \in U$ auch $y - x = -(x - y) \in U$ folgt.
(iii)⇒(i): $x = x + 0 \in x + U$, da $0 \in U$. Also ist $x + U \neq \emptyset$ und die Behauptung folgt.

(3) Es gilt $x + U = U$ genau dann, wenn $x + U = 0 + U$ (nach (1)), also genau dann, wenn $x - 0 \in U$ (nach (2)).

(4) Zur Erinnerung: $A + B = \{a + b \mid a \in A, b \in B\}$ für $A, B \subseteq V$ (siehe Definition 3.8). Sei $v \in (x + U) + (y + U)$. Dann ist $v = (x + u_1) + (y + u_2)$ mit $u_1, u_2 \in U$ geeignet. Also $v = (x + y) + (u_1 + u_2) \in (x + y) + U$, daher ist $(x + U) + (y + U) \subseteq (x + y) + U$. Wenn $w \in (x + y) + U$, etwa $w = x + y + u$, $u \in U$, dann $w = (x + 0) + (y + u) \in (x + U) + (y + U)$, also $(x + y) + U \subseteq (x + U) + (y + U)$.

5.4 Lemma:

Sei $U \leq V$, $x, y \in V$ und $k \in \mathbb{K}$. Wenn $x + U = y + U$, dann $kx + U = ky + U$.

Beweis: Wenn $x+U = y+U$, dann $x-y \in U$, nach 5.3 (2). Es folgt $kx-ky = k(x-y) \in U$, da U ein Unterraum ist. Also $kx + U = ky + U$, wieder nach 5.3 (2).

5.5 Satz/Definition: *Faktorraum*

Die Menge der Nebenklassen des Unterraumes U in V bildet mit der Addition

$$(x + U) + (y + U) = (x + y) + U,\ x, y \in V$$

und der skalaren Multiplikation

$$k(x + U) = kx + U,\ k \in \mathbb{K}, x \in V$$

einen Vektorraum über $\mathbb{K}$. Diesen Vektorraum bezeichnet man mit V/U (gesprochen V modulo U) und nennt ihn den Faktorraum von V nach U.

Beweis: V/U ist abelsche Gruppe:

- Assoziativität und Kommutativität sind klar, da sie für Elemente von V gelten.
- Neutrales Element ist U.
- Das Negative zu $x + U$ ist $(-x) + U$.

Die Wohldefiniertheit der skalaren Multiplikation ist gerade 5.4. Die Vektorraum–Axiome für V/U sind triviale Folgerungen aus den Axiomen für V.

5.6 Definition: *lineare Abbildung, Homomorphismus, Mono–, Epi–, Iso–, Endo–, Automorphismus, Bild, Kern*

(1) Seien V und W Vektorräume über $\mathbb{K}$. Eine Abbildung $\alpha : V \to W$ heißt lineare Abbildung oder (Vektorraum–)Homomorphismus, wenn

$$\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2) \text{ für alle } v_1, v_2 \in V$$

und

$$\alpha(kv) = k\alpha(v) \text{ für alle } v \in V, k \in \mathbb{K}.$$

Eine injektive (surjektive, bijektive) lineare Abbildung heißt Monomorphismus (Epimorphismus, Isomorphismus). Zwei Vektorräume V, W heißen isomorph ($V \cong W$), falls ein Isomorphismus $\alpha : V \to W$ existiert. Wenn $V = W$, dann nennt man die Homomorphismen auch Endomorphismen, und die Isomorphismen auch Automorphismen.

(2) Sei $\alpha : V \to W$ eine lineare Abbildung. Man nennt

$$\mathrm{Im}\,(\alpha) = \{\alpha(v) \mid v \in V\} \text{ das Bild von } \alpha$$

und

$$\mathrm{Ker}\,(\alpha) = \{v \in V \mid \alpha(v) = 0\} \text{ den Kern von } \alpha.$$

5.7 Beispiel/Bemerkung/Definition:

(i) Seien V und W beliebige $\mathbb{K}$–Vektorräume. Die Abbildung $v \mapsto 0_W$ für alle $v \in V$ ist linear. Diese Abbildung heißt die Nullabbildung.

(ii) Die Identität $\mathrm{id}_V : V \to V$ ist linear.

(iii) Sei $U \leq V$. Die Abbildung $\kappa : V \to V/U$, welche durch $\kappa(v) = v + U$ definiert wird, ist linear und surjektiv. Diese Abbildung heißt der kanonische Epimorphismus. Es ist $v \in \mathrm{Ker}\,(\kappa) \Leftrightarrow v + U = \kappa(v) = 0_{V/U} = U \Leftrightarrow v \in U$. Also ist $\mathrm{Ker}\,(\kappa) = U$.

(iv) Die Menge $\mathrm{Hom}\,(V, W)$ der linearen Abbildungen von V in W bildet bzgl. der Addition

$$(\alpha + \beta)(v) = \alpha(v) + \beta(v),\ \alpha, \beta \in \mathrm{Hom}\,(V, W)\,, v \in V$$

eine abelsche Gruppe; das Negative $-\alpha$ zu $\alpha \in \mathrm{Hom}\,(V, W)$ ist definiert durch $(-\alpha)(v) = -\alpha(v)$, und das neutrale Element ist die Nullabbildung. Für $\alpha \in \mathrm{Hom}\,(V, W)$ und $k \in \mathbb{K}$ definiert man $k\alpha$ durch

$$(k\alpha)(v) = k\alpha(v).$$

Damit ist $\mathrm{Hom}\,(V, W)$ ein $\mathbb{K}$–Vektorraum.

(v) Wenn $\alpha : U \to V$ und $\beta : V \to W$ linear sind, dann ist auch $\beta\alpha : U \to W$ linear. Diese Multiplikation von Abbildungen verträgt sich mit der in (iv) definierten Addition: es gelten die beiden Distributivitätsgesetze $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$ und $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$, sofern die Verknüpfungen definiert sind.

(vi) Insbesondere bildet die Menge $\mathrm{End}\,(V) = \mathrm{Hom}\,(V, V)$ der Endomorphismen von V mit den oben definierten Verknüpfungen einen Ring mit Eins ($= \mathrm{id}_V$). Dieser Ring heißt Endomorphismen–Ring von V.

(vii) „Isomorphie“ ist eine Äquivalenzrelation:

(1) $\mathrm{id}_V : V \to V$ ist ein Isomorphismus (Reflexivität).

(2) Wenn $\alpha : U \to V$ und $\beta : V \to W$ Isomorphismen sind, dann ist auch $\beta\alpha : U \to W$ ein Isomorphismus (Transitivität).

(3) Wenn $\alpha : V \to W$ ein Isomorphismus ist, dann gilt dies auch für die Umkehrabbildung $\alpha^{-1} : W \to V$. Der Beweis dazu ist eine Übungsaufgabe. Dies zeigt die Symmetrie.

5.8 Lemma:

Sei $\alpha : V \to W$ eine lineare Abbildung. Es gelten:

(1) $\alpha(0_V) = 0_W$

(2) $\alpha(-v) = -\alpha(v)$.

Beweis:

(1) $\alpha(0) = \alpha(0 \cdot 0) = 0\alpha(0) = 0$

(2) $\alpha(-v) = \alpha((-1) \cdot v) = (-1)\alpha(v) = -\alpha(v)$ (vergleiche 3.4)

5.9 Lemma:

Sei $\alpha : V \to W$ linear, dann:

(1) $\mathrm{Im}\,(\alpha) \leq W$

(2) $\mathrm{Ker}\,(\alpha) \leq V$.

Beweis: Verwende jeweils das Unterraumkriterium 3.7.

(1) Es ist $0_W = \alpha(0_V) \in \mathrm{Im}\,(\alpha)$.
Wenn $w_1, w_2 \in \mathrm{Im}\,(\alpha)$, etwa $w_1 = \alpha(v_1), w_2 = \alpha(v_2)$ mit $v_1, v_2 \in V$, dann ist $w_1 + w_2 = \alpha(v_1) + \alpha(v_2) = \alpha(v_1 + v_2) \in \mathrm{Im}\,(\alpha)$.
Wenn $w = \alpha(v) \in \mathrm{Im}\,(\alpha)$ und $k \in \mathbb{K}$, dann ist $kw = k\alpha(v) = \alpha(kv) \in \mathrm{Im}\,(\alpha)$.

(2) Es ist $\alpha(0_V) = 0_W$, also $0_V \in \mathrm{Ker}\,(\alpha)$.
Wenn $v_1, v_2 \in \mathrm{Ker}\,(\alpha)$ dann ist $\alpha(v_1 + v_2) = \alpha(v_1) + \alpha(v_2) = 0_W + 0_W = 0_W$, also $v_1 + v_2 \in \mathrm{Ker}\,(\alpha)$.
Wenn $v \in \mathrm{Ker}\,(\alpha)$ und $k \in \mathbb{K}$, dann ist $\alpha(kv) = k\alpha(v) = k0_W = 0_W$, also $kv \in \mathrm{Ker}\,(\alpha)$.

5.10 Satz:

Sei $\alpha : V \to W$ linear.

(1) Genau dann ist α ein Epimorphismus, wenn $\mathrm{Im}\,(\alpha) = W$.

(2) Genau dann ist α ein Monomorphismus, wenn $\mathrm{Ker}\,(\alpha) = \{0\}$.

Beweis:

(1) trivial

(2) „$\Rightarrow$“: Sei α ein Monomorphismus, und sei $v \in \mathrm{Ker}\,(\alpha)$. Dann ist $\alpha(v) = 0_W = \alpha(0_V)$. Da α injektiv ist, folgt $v = 0_V$. Also ist $\mathrm{Ker}\,(\alpha) = \{0_V\}$.
„$\Leftarrow$“: Zu zeigen: α ist injektiv. Sei $\alpha(v_1) = \alpha(v_2)$. Dann $0_W = \alpha(v_1) - \alpha(v_2) = \alpha(v_1 - v_2)$, d.h. $v_1 - v_2 \in \mathrm{Ker}\,(\alpha) = \{0_V\}$. Also ist $v_1 = v_2$.

5.11 Satz: *Erster Isomorphiesatz*

Sei $\alpha : V \to W$ linear. Dann ist $\mathrm{Im}\,(\alpha) \cong V/\mathrm{Ker}\,(\alpha)$.

Beweis: Definiere $\bar{\alpha} : V/\mathrm{Ker}\,(\alpha) \to \mathrm{Im}\,(\alpha)$ durch $\bar{\alpha}(v + \mathrm{Ker}\,(\alpha)) = \alpha(v)$. Dann ist $\bar{\alpha}$ wohldefiniert; denn wenn $v + \mathrm{Ker}\,(\alpha) = v' + \mathrm{Ker}\,(\alpha)$, dann ist $v - v' \in \mathrm{Ker}\,(\alpha)$ nach 5.3 (2), also $0 = \alpha(v - v') = \alpha(v) - \alpha(v')$ und daher $\alpha(v) = \alpha(v')$. Die Linearität von $\bar{\alpha}$ folgt sofort aus der Linearität von α und der Definition von Addition und skalarer Multiplikation in $V/\mathrm{Ker}\,(\alpha)$. Offenbar ist $\bar{\alpha}$ surjektiv. Wenn $v + \mathrm{Ker}\,(\alpha) \in \mathrm{Ker}\,(\bar{\alpha})$, dann $0 = \bar{\alpha}(v + \mathrm{Ker}\,(\alpha)) = \alpha(v)$, also $v \in \mathrm{Ker}\,(\alpha)$. Nach 5.3 (3) ist dann $v + \mathrm{Ker}\,(\alpha) = \mathrm{Ker}\,(\alpha)$, und dies ist gerade das Nullelement in $V/\mathrm{Ker}\,(\alpha)$. Wir haben gezeigt: Der Kern von $\bar{\alpha}$ besteht nur aus dem Nullelement. Nach 5.10 ist $\bar{\alpha}$ ein Monomorphismus.

5.12 Satz:

Sei $\alpha : V \to W$ linear. Ergänze eine Basis X von $\mathrm{Ker}\,(\alpha)$ zu einer Basis $B = X \dot\cup Y$ von V. Dann ist $\alpha_{|Y}$ injektiv, und $\alpha(Y)$ ist eine Basis von $\mathrm{Im}\,(\alpha)$.

Beweis: Sei $y_1, y_2 \in Y$. Wenn $\alpha(y_1) = \alpha(y_2)$, dann ist $y_1 - y_2 \in \mathrm{Ker}\,(\alpha)$, daher existieren $k_x \in \mathbb{K}$ mit $y_1 - y_2 - \sum_{x\in X} k_x x = 0$. Wäre $y_1 \neq y_2$, so wäre B nicht linear unabhängig. Also ist $\alpha_{|Y}$ injektiv.

Angenommen $0 = \sum_{y\in Y} k_y \alpha(y) = \alpha\left(\sum_{y\in Y} k_y y\right)$, also $\sum_{y\in Y} k_y y \in \mathrm{Ker}\,(\alpha)$. Dann ist $\sum_{y\in Y} k_y y = \sum_{x\in X} k_x x$ für geeignete $k_x \in \mathbb{K}$, also $0 = \sum_{y\in Y} k_y y - \sum_{x\in X} k_x x$. Weil B linear unabhängig ist, sind alle $k_y = 0$. Also ist $\alpha(Y)$ linear unabhängig.

Wenn $w \in \mathrm{Im}\,(\alpha)$, dann $w = \alpha(v)$ für ein $v \in V$. Da $v = \sum_{x\in X} k_x x + \sum_{y\in Y} k_y y$ für geeignete $k_x, k_y \in \mathbb{K}$, ist $\alpha(v) = \sum_{x\in X} k_x \alpha(x) + \sum_{y\in Y} k_y \alpha(y) = \sum_{y\in Y} k_y \alpha(y)$. Daher ist $\alpha(Y)$ auch ein Erzeugenden–System von $\mathrm{Im}\,(\alpha)$.

5.13 Korollar:

Sei $\alpha : V \to W$ linear. Dann ist

$$\dim V = \dim \mathrm{Ker}\,(\alpha) + \dim \mathrm{Im}\,(\alpha)\,.$$

Beweis: Es ist mit den Bezeichnungen des Satzes 5.12 $\dim V = |B| = |X| + |Y|$, da $B = X \dot\cup Y$, wobei $|B|$, ... die Anzahl der Elemente in B, ... bezeichnet. Da X eine Basis von $\mathrm{Ker}\,(\alpha)$ ist, gilt $|X| = \dim \mathrm{Ker}\,(\alpha)$. Da $\alpha_{|Y} : Y \to \alpha(Y)$ bijektiv und $\alpha(Y)$ eine Basis von $\mathrm{Im}\,(\alpha)$ ist, gilt $|Y| = |\alpha(Y)| = \dim \mathrm{Im}\,(\alpha)$. Die Behauptung folgt.

5.14 Korollar:

Sei $U \leq V$. Dann ist

$$\dim V = \dim U + \dim V/U$$

Beweis: Der kanonische Epimorphismus $\kappa : V \to V/U$ hat $\mathrm{Ker}\,(\kappa) = U$; vergleiche 5.7.

5.15 Definition: *Rang einer linearen Abbildung*

Sei $\alpha : V \to W$ linear. Man nennt $\mathrm{Rg}\,(\alpha) = \dim(\mathrm{Im}\,(\alpha))$ den Rang von α.

5.16 Korollar:

Seien V, W $\mathbb{K}$–VR mit $\dim V = \dim W = n$ und sei $\alpha : V \to W$ linear. Dann sind äquivalent:

(1) $\exists \beta : W \to V$ mit $\beta\alpha = \mathrm{id}_V$.

(2) α ist injektiv.

(3) $\mathrm{Rg}\,(\alpha) = n$

(4) α ist surjektiv.

(5) α ist ein Isomorphismus.

Außerdem gilt: β ist durch (1) eindeutig bestimmt, nämlich $\beta = \alpha^{-1}$. Insbesondere ist β auch linear und $\alpha\beta = \mathrm{id}_W$.

Beweis:

(1)⇒(2) Wenn $\alpha(v_1) = \alpha(v_2)$, dann $v_1 = \beta\alpha(v_1) = \beta\alpha(v_2) = v_2$, d.h. α ist injektiv.

(2)⇒(3) Nach 5.10 ist $\mathrm{Ker}\,(\alpha) = \{0\}$, also auch $\dim \mathrm{Ker}\,(\alpha) = 0$. Daher gilt mit 5.13 $n = \dim V = \dim \mathrm{Ker}\,(\alpha) + \dim \mathrm{Im}\,(\alpha) = \dim \mathrm{Im}\,(\alpha) = \mathrm{Rg}\,(\alpha)$.

(3)⇒(4) Es ist $\dim \mathrm{Im}\,(\alpha) = \mathrm{Rg}\,(\alpha) = n = \dim W$, also $\mathrm{Im}\,(\alpha) = W$ nach 4.20 (3), d.h. α ist surjektiv.

(4)⇒(5) Wegen $\mathrm{Im}\,(\alpha) = W$ gilt $\dim(\mathrm{Ker}\,(\alpha)) = \dim V - \dim W = 0$, also $\mathrm{Ker}\,(\alpha) = \{0\}$, d.h. α ist injektiv. Also ist α ein Isomorphismus.

(5)⇒(1) Wähle $\beta = \alpha^{-1}$.

Außerdem: Wenn $\beta\alpha = \mathrm{id}_V = \alpha^{-1}\alpha$, dann ist $\beta = \beta\alpha\alpha^{-1} = \alpha^{-1}\alpha\alpha^{-1} = \alpha^{-1}$. Daß α^{-1} linear ist, ist in 5.7 (7) bemerkt. Es gilt $\alpha\alpha^{-1} = \mathrm{id}_W$ nach Definition.

5.17 Satz:

Sei $B \subseteq V$. Äquivalent sind:

(1) B ist eine Basis von V.

(2) Für jeden $\mathbb{K}$–VR W und jede Abbildung $a : B \to W$ gibt es genau eine lineare Abbildung $\alpha : V \to W$ mit $\alpha_{|B} = a$.

(3) Für jede Abbildung $a : B \to \mathbb{K}$ gibt es genau eine lineare Abbildung $\alpha : V \to \mathbb{K}$ mit $\alpha_{|B} = a$.

Beweis:

(1)⇒(2) Wenn $v \in V$, dann ist nach 4.7 $v = \sum_{b\in B} k_b b$ mit eindeutig bestimmten Skalaren $k_b \in \mathbb{K}$. Setze dann $\alpha(v) = \sum_{b\in B} k_b \alpha(b) \in W$. Dann ist α wohldefiniert, linear, $\alpha_{|B} = a$, und offenbar ist α die einzige Abbildung mit diesen Eigenschaften.

(2)⇒(3) trivial

(3)⇒(1) B ist linear unabhängig: Sei $\sum_{b\in B} k_b b = 0$ und $b_0 \in B$ fest. Definiere $a_0 : B \to \mathbb{K}$ durch $a_0(b_0) = 1$ und $a_0(b) = 0$ für alle $b \neq b_0$. Sei $\alpha_0 : V \to \mathbb{K}$ linear mit $\alpha_{0|B} = a_0$. Dann ist $0 = \alpha_0(0) = \sum_{b\in B} k_b a_0(b) = k_{b_0}$. Da dies für jedes $b_0 \in B$ gilt, sind alle $k_b = 0$, d.h. B ist linear unabhängig.

Es läßt sich B zu einer Basis $\tilde{B} = B \dot\cup X$ von V ergänzen (nach 4.19 (1)). Angenommen $x \in X$. Sei $a_0 : \tilde{B} \to \mathbb{K}$ definiert durch $a_0(y) = 0$ für alle $y \in \tilde{B}$ und sei $a_1 : \tilde{B} \to \mathbb{K}$ definiert durch $a_1(x) = 1$ und $a_1(y) = 0$ für alle $y \neq x$, $y \in \tilde{B}$. Seien α_0 und α_1 die entsprechenden linearen Abbildungen $V \to \mathbb{K}$ (diese existieren nach (1)⇒(2)). Dann ist $\alpha_{0|B} = \alpha_{1|B}$, aber $\alpha_0 \neq \alpha_1$. Dies widerspricht der Eindeutigkeit. Also ist $X = \emptyset$ und $\tilde{B} = B$ eine Basis von V.

5.18 Definition: *lineare Fortsetzung*

Die Abbildung α in 5.17 (2) heißt die lineare Fortsetzung von a auf V.

5.19 Korollar:

Sei B eine Basis von V und seien $\alpha, \beta : V \to W$ linear. Dann gilt:

$$\alpha = \beta \Leftrightarrow \alpha_{|B} = \beta_{|B}.$$

5.20 Satz:

Seien V, W $\mathbb{K}$–VR. Dann gilt:

$$V \cong W \Leftrightarrow \dim V = \dim W.$$

Beweis:

„⇒“: Sei $\alpha : V \to W$ ein Isomorphismus. Es ist $\operatorname{Im}(\alpha) = W$ und $\operatorname{Ker}(\alpha) = \{0\}$ nach 5.10. Also ist mit 5.13 $\dim V = \dim \operatorname{Ker}(\alpha) + \dim \operatorname{Im}(\alpha) = 0 + \dim W = \dim W$.

„⇐“: Sei B eine Basis von V und C eine Basis von W. Dann ist nach Voraussetzung $|B| = |C|$, also existiert eine bijektive Abbildung $a : B \to C$. Sei $b : C \to B$ die Umkehrabbildung zu a, α die lineare Fortsetzung von a und β die lineare Fortsetzung von b. Da id_V die lineare Fortsetzung von id_B ist, folgt $\beta\alpha = \mathrm{id}_V$, und ebenso $\alpha\beta = \mathrm{id}_W$. Also ist α ein Isomorphismus.

5.21 Korollar:

Sei V ein endlich dimensionaler $\mathbb{K}$–VR, etwa $\dim V = n$. Dann ist $V \cong \mathbb{K}^n$.

Beweis: klar nach 5.20

5.22 Bemerkung:

Sei $\alpha : V \to W$ linear und $w \in W$. Setze $X = \{v \in V \mid \alpha(v) = w\}$. Dann gelten:

(i) Falls $\langle w, \operatorname{Im}(\alpha)\rangle \gneq \operatorname{Im}(\alpha)$, dann $X = \emptyset$.

(ii) Falls $\langle w, \operatorname{Im}(\alpha)\rangle = \operatorname{Im}(\alpha)$, dann $X = v + \operatorname{Ker}(\alpha)$, wobei $v \in V$ ein beliebiges Element mit $\alpha(v) = w$ ist.

Beweis: $X \neq \emptyset \Leftrightarrow \exists v \in V : \alpha(v) = w \Leftrightarrow w \in \operatorname{Im}(\alpha) \Leftrightarrow \langle w, \operatorname{Im}(\alpha)\rangle = \operatorname{Im}(\alpha)$. Wenn $v \in X$, dann $v_1 \in X \Leftrightarrow \alpha(v_1) = w = \alpha(v) \Leftrightarrow \alpha(v_1 - v) = 0 \Leftrightarrow v_1 - v \in \operatorname{Ker}(\alpha) \Leftrightarrow v_1 \in v + \operatorname{Ker}(\alpha)$ (siehe 5.3), also $X = v + \operatorname{Ker}(\alpha)$.

5.23 Satz: *Zweiter Isomorphiesatz*

Seien U und W Unterräume von V. Dann gilt:

$$(U + W)/U \cong W/(U \cap W)$$

Beweis: Vorbemerkung: $U \leq U + W \leq V$, also ist $(U + W)/U$ sinnvoll, ebenso $U \cap W \leq W \leq V$, also ist $W/(U \cap W)$ sinnvoll.

Zum eigentlichen Beweis: Sei $\varepsilon : W \to (U+W)/U$ definiert durch $\varepsilon(w) = w + U$. Dann ist ε linear, sogar Epimorphismus, denn jedes $x \in (U+W)/U$ ist von der Form $x = (u+w) + U = w + u + U = w + U = \varepsilon(w)$ mit $u \in U, w \in W$. Es ist $w \in \mathrm{Ker}\,(\varepsilon) \Leftrightarrow U = 0_{(U+W)/U} = \varepsilon(w) = w + U \Leftrightarrow w \in U \cap W$ nach 5.3. Daher ist $\mathrm{Ker}\,(\varepsilon) = U \cap W$. Nach dem Ersten Isomorphiesatz (5.11) gilt

$$(U+W)/U = \mathrm{Im}\,(\varepsilon) \cong W/\mathrm{Ker}\,(\varepsilon) = W/(U \cap W).$$

5.24 Korollar:

Seien U und W Unterräume von V. Dann gilt:

$$\dim U + \dim W = \dim(U+W) + \dim(U \cap W).$$

Beweis:

$$\begin{aligned}
\dim(U+W) \;+\; \dim(U \cap W) &= \dim(U+W)/U \;+\; \dim U \;+\; \dim(U \cap W), \text{ nach 5.14,} \\
&= \dim W/(U \cap W) \;+\; \dim(U \cap W) \;+\; \dim U, \\
&\qquad\qquad \text{nach 5.23 und 5.20,} \\
&= \dim W \;+\; \dim U, \text{ wieder nach 5.14.}
\end{aligned}$$

6 Dualer Raum und duale Abbildung

6.1 Definition: *Dualer Raum*

Sei V ein $\mathbb{K}$–VR. Man nennt

$$V^* = \{f : V \to \mathbb{K} \mid f \text{ ist linear}\}$$

den zu V dualen Raum.

6.2 Bemerkung:

Durch $(f+g)(v) = f(v) + g(v)$ und $(kf)(v) = k(f(v))$ ist V^* wieder ein $\mathbb{K}$–VR. Das rechtfertigt die Bezeichnung. (In der Tat ist $V^* = \mathrm{Hom}\,(V, \mathbb{K})$ ein Spezialfall von 5.7.)

6.3 Satz/Definition:

Sei $\dim V = n < \infty$ und sei $B = \{b_1, \ldots, b_n\}$ eine Basis von V. Definiere $b_i^* \in V^*$ durch

$$b_i^*(b_j) = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

(vergleiche 5.17). Dann ist $B^* = \{b_1^*, \ldots, b_n^*\}$ eine Basis von V^*, genannt die duale Basis zu B.

Beweis: $\{b_1^*, \ldots, b_n^*\}$ ist linear unabhängig, denn wenn $\sum_i k_i b_i^* = 0$ (die Nullabbildung), dann $0 = \left(\sum_i k_i b_i^*\right)(b_j) = \sum_i k_i b_i^*(b_j) = k_j$ für alle $j = 1, \ldots, n$.

Sei $f : V \to \mathbb{K}$ linear und $k_i = f(b_i)$. Dann ist $\left(\sum_i k_i b_i^*\right)(b_j) = k_j = f(b_j)$. Also stimmen die beiden linearen Abbildungen f und $\left(\sum_i k_i b_i^*\right)$ auf B überein. Nach 5.19 ist $f = \sum_i k_i b_i^*$. Also bilden die b_i^*'s ein Erzeugenden–System von V^*.

6.4 Korollar:

Wenn $\dim V < \infty$, dann $V \cong V^*$.

Beweis: Nach 6.3 ist $\dim V = \dim V^*$. Die Behauptung folgt daher aus 5.20.

6.5 Definition: *Duale Abbildung*

Sei $\alpha : V \to W$ linear. Die Abbildung $\alpha^* : W^* \to V^*$, definiert durch

$$\alpha^*(f) = f\alpha$$

für $f \in W^*$, heißt die zu α duale Abbildung.

6.6 Bemerkung:

(i) Wenn $f \in W^*$, dann ist $W \xrightarrow{f} \mathbb{K}$ linear. Nach 5.7 ist dann auch $V \xrightarrow{\alpha} W \xrightarrow{f} \mathbb{K}$ linear, d.h. $f\alpha \in V^*$.

(ii) α^* ist linear, denn $\alpha^*(f+g) = (f+g)\alpha = f\alpha + g\alpha = \alpha^*(f) + \alpha^*(g)$ und $\alpha^*(kf) = (kf)\alpha = k(f\alpha) = k\alpha^*(f)$ für alle $f, g \in W^*, k \in \mathbb{K}$.

(iii) Wenn $\beta : U \to V$ und $\alpha : V \to W$ linear sind, dann ist $\alpha\beta : U \to W$ ebenfalls linear, und es gilt $(\alpha\beta)^* = \beta^*\alpha^*$.

6.7 Satz:

Sei $\alpha : V \to W$ linear mit $\operatorname{Rg}(\alpha) < \infty$. Dann ist $\operatorname{Rg}(\alpha^*) = \operatorname{Rg}(\alpha)$.

Beweis: Sei $\operatorname{Rg}(\alpha) = n = \dim \operatorname{Im}(\alpha)$ und $B_0 = \{b_1, \ldots, b_n\}$ eine Basis von $\operatorname{Im}(\alpha) \leq W$. Ergänze B_0 zu einer Basis B von W. Definiere $f_i \in W^*$ für $i = 1, \ldots, n$ durch

$$f_i(b) = \begin{cases} 1 & \text{falls } b_i = b \\ 0 & \text{falls } b_i \neq b \in B \end{cases}$$

(vergleiche 5.17).

Behauptung: Die Menge $B_0^* = \{\alpha^*(f_i) \mid i = 1, \ldots, n\}$ besteht aus n Elementen und ist eine Basis von $\operatorname{Im}(\alpha^*)$.

Es folgt dann $\operatorname{Rg}(\alpha^*) = \dim \operatorname{Im}(\alpha^*) = n = \dim \operatorname{Im}(\alpha) = \operatorname{Rg}(\alpha)$.

B_0^* ist linear unabhängig: Sei $0 = \sum_{i=1}^n k_i\alpha^*(f_i) = \sum_{i=1}^n k_i f_i \alpha$ die Nullabbildung. Wähle $v_j \in V$ mit $\alpha(v_j) = b_j$, $j = 1, \ldots, n$. Dann ist $0 = 0(v_j) = \sum_{i=1}^n k_i f_i \alpha(v_j) = \sum_{i=1}^n k_i f_i(b_j) = k_j$ für alle $j = 1, \ldots, n$. Also ist $|B_0^*| = n$ und B_0^* ist linear unabhängig.

B_0^* ist ein Erzeugenden–System von $\operatorname{Im}(\alpha^*)$: Sei $g \in \operatorname{Im}(\alpha^*)$, etwa $g = \alpha^*(f) = f\alpha$ für ein $f \in W^*$. Sei $k_i = f(b_i)$, $i = 1, \ldots, n$. Dann ist

$$g = \sum_{i=1}^n k_i \alpha^*(f_i) \ ,$$

denn: Beide Seiten sind Elemente von V^*, d.h. Abbildungen $V \to \mathbb{K}$. Um die Gleichheit zu zeigen, muß man beide Abbildungen auf ein beliebiges Element $v \in V$ anwenden. Wenn $v \in V$, dann ist $\alpha(v) \in \operatorname{Im}(\alpha)$, also $\alpha(v) = \sum_{j=1}^n k'_j b_j$ für geeignete $k'_j \in \mathbb{K}$. Daher ist

$$g(v) = (f\alpha)(v) = f\left(\sum_{j=1}^n k'_j b_j\right) = \sum_{j=1}^n k'_j k_j$$

und

$$\left(\sum_{i=1}^n k_i \alpha^*(f_i)\right)(v) = \sum_{i=1}^n k_i f_i \alpha(v) = \sum_{i=1}^n k_i f_i \sum_{j=1}^n k'_j b_j = \sum_{i,j=1}^n k_i k'_j f_i(b_j) = \sum_{j=1}^n k'_j k_j.$$

7 Matrizen und lineare Gleichungssysteme

Sei $\mathbb{K}$ ein Körper.

7.1 Definition: *Matrix*

Seien $n, m \in \mathbb{N}$. Eine $n \times m$–Matrix über $\mathbb{K}$ ist ein $n \times m$–Tupel von Elementen aus $\mathbb{K}$, angeordnet in einem Rechteck von n Zeilen und m Spalten.

7.2 Beispiel/Bezeichnung/Bemerkung:

Sei

$$A = \begin{pmatrix} 1 & 2 & -\pi & \sqrt{3} \\ -1 & 0 & 2 & 10 \\ 11 & 11 & 0 & 0 \end{pmatrix}.$$

Dies ist eine 3×4–Matrix über $\mathbb{R}$. A hat also 3 Zeilen und 4 Spalten. Die Länge der Zeilen ist die Anzahl der Spalten (hier 4), ebenso ist die Spaltenlänge gleich der Zeilenzahl (hier 3). Jede Spalte kann also als Element von $\mathbb{R}^3$ aufgefaßt werden. Allgemein: Die m Spalten einer $n \times m$–Matrix A über $\mathbb{K}$ sind Elemente aus $\mathbb{K}^n$. Der von den Spalten erzeugte Unterraum von $\mathbb{K}^n$ heißt Spaltenraum von A. Entsprechend ist jede Zeile ein Element aus $\mathbb{K}^m$, und der von den Zeilen erzeugte Unterraum von $\mathbb{K}^m$ heißt Zeilenraum von A. Die Einträge von A werden oft mit $a_{i,j}$ bezeichnet, wobei $i = 1, \ldots, n$, $j = 1, \ldots, m$, und $a_{i,j}$ das Element in der i–ten Zeile und j–ten Spalte ist. Anstelle von A schreibt man dann auch oft $(a_{i,j})$. Im Beispiel ist $a_{2,2} = a_{3,3} = a_{3,4} = 0$, $a_{1,3} = -\pi$, $a_{3,1} = 11$.

7.3 Bemerkung/Definition: *Addition und skalare Multiplikation von Matrizen*

Seien A, B Matrizen der gleichen Größe über $\mathbb{K}$. Dann ist die Summe $A + B$ die Matrix, welche durch komponentenweise Addition entsteht, d.h. wenn $A = (a_{i,j})$, $B = (b_{i,j})$, dann ist $A + B = (c_{i,j})$ mit $c_{i,j} = a_{i,j} + b_{i,j}$ für alle i, j. Für $k \in \mathbb{K}$ ist das skalare Produkt kA die Matrix, welche man durch Multiplikation aller Einträge von A mit k erhält.

Mit dieser Addition und Multiplikation mit Skalaren bildet die Menge aller $n \times m$–Matrizen über $\mathbb{K}$ einen $\mathbb{K}$–Vektorraum der Dimension nm.

7.4 Bemerkung/Definition: *Multiplikation von Matrizen*

Sei A eine $n \times m$– und B eine $m \times \ell$–Matrix über $\mathbb{K}$. Dann ist das Produkt $C = AB$ die $n \times \ell$–Matrix, deren Einträge $c_{i,j}$ definiert sind durch

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

Das Produkt AB zweier Matrizen ist also nur dann definiert, wenn die Spaltenzahl von A gleich der Zeilenzahl von B ist. In diesem Fall ist die Zeilenzahl von AB gleich der Zeilenzahl von A und die Spaltenzahl von AB gleich der Spaltenzahl von B.

7.5 Beispiel/Definition: *Nullmatrix, quadratische Matrix, Einheitsmatrix*

(i) Sei A wie 7.2 und

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

Dann ist B eine 4×2–Matrix. Da A eine 3×4–Matrix ist, ist AB definiert und eine 3×2–Matrix, nämlich

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 & -\pi & \sqrt{3} \\ -1 & 0 & 2 & 10 \\ 11 & 11 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 2 & 3 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 0 + 2 \cdot 0 + (-\pi) \cdot 2 + \sqrt{3} \cdot (-1) & 1 \cdot 1 + 2 \cdot 0 + (-\pi) \cdot 3 + \sqrt{3} \cdot 0 \\ (-1) \cdot 0 + 0 \cdot 0 + 2 \cdot 2 + 10 \cdot (-1) & (-1) \cdot 1 + 0 \cdot 0 + 2 \cdot 3 + 10 \cdot 0 \\ 11 \cdot 0 + 11 \cdot 0 + 0 \cdot 2 + 0 \cdot (-1) & 11 \cdot 1 + 11 \cdot 0 + 0 \cdot 3 + 0 \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} -2\pi - \sqrt{3} & 1 - 3\pi \\ -6 & 5 \\ 0 & 11 \end{pmatrix}. \end{aligned}$$

BA ist nicht definiert, da Spaltenzahl von $B = 2 \neq 3 =$ Zeilenzahl von A.

(ii) Sei A eine $n \times m$– und B eine $k \times \ell$–Matrix. Wenn AB und BA definiert sind, dann ist $m = k$ und $\ell = n$. In diesem Fall ist AB eine $n \times n$–Matrix und BA eine $m \times m$–Matrix, also sind AB und BA im Allgemeinen von verschiedener Größe. Sogar wenn $n = m$ ist, brauchen AB und BA nicht gleich zu sein, zum Beispiel:

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

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \text{ die } 2 \times 2\text{–}\underline{\text{Nullmatrix}}.$$

(iii) Eine Matrix mit ebenso vielen Zeilen wie Spalten heißt <u>quadratisch</u>, zum Beispiel:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ die } 3 \times 3\text{–}\underline{\text{Einheitsmatrix}}.$$

7.6 Wichtiges Beispiel/Definition: *Matrix zu einer linearen Abbildung*

Sei $\alpha : V \to W$ linear. Sei $\{a_1, \ldots, a_m\}$ eine Basis von V und $\{b_1, \ldots, b_n\}$ eine Basis von W, also $\dim V = m < \infty$ und $\dim W = n < \infty$. Für jedes $j = 1, \ldots, m$ ist $\alpha(a_j) \in W$, kann also eindeutig als Linearkombination der b_i's geschrieben werden:

$$\alpha(a_j) = \sum_{i=1}^{n} k_{ij} b_i \text{ mit } k_{ij} \in \mathbb{K}.$$

Die Matrix $A = (k_{ij})$ ist eine $n \times m$–Matrix. Man nennt A die Matrix von α bezüglich $\{a_1, \ldots, a_m\}$ und $\{b_1, \ldots, b_n\}$. Die j–te Spalte von A erhält man also, indem man das Bild des j–ten Basisvektors von V unter α als Linearkombination der b_i's schreibt. Die Matrix A hängt also ab von α, aber auch von der speziellen Wahl der Basen in V und W und sogar von der Reihenfolge der Basisvektoren.

7.7 Bemerkung:

Sei $\alpha : V \to W$ und $A = (k_{ij})$ wie in 7.6. Wenn

$$v = \sum_{j=1}^{m} x_j a_j \in V,\ x_j \in \mathbb{K},$$

dann ist

$$\alpha(v) = \sum_{j=1}^{m} x_j \alpha(a_j) = \sum_{j=1}^{m} x_j \sum_{i=1}^{n} k_{ij} b_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} k_{ij} x_j \right) b_i,$$

d.h. $\sum_{j=1}^{m} k_{ij} x_j$ ist der Koeffizient von b_i in $\alpha(v)$. Man erhält diese Koeffizienten also gerade durch Multiplikation $A \cdot x$, wobei x der Spaltenvektor

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

ist.

7.8 Beispiel:

Sei $V = \mathbb{R}^2$, $W = \mathbb{R}^3$, und sei α definiert durch $\alpha(a, b) = (3a - b, 2a + 5b, -a)$.

(i) Basis von V und W sei jeweils die Standardbasis in der natürlichen Reihenfolge. Dann ist $\alpha(1, 0) = (3, 2, -1)$, $\alpha(0, 1) = (-1, 5, 0)$ und

$$A = \begin{pmatrix} 3 & -1 \\ 2 & 5 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 3 & -1 \\ 2 & 5 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3a - b \\ 2a + 5b \\ -a \end{pmatrix}.$$

(ii) Basis von V sei $\{a_1, a_2\}$ mit $a_1 = (1, 2)$, $a_2 = (1, 1)$. (Dies ist wirklich eine Basis, denn aus $0 = k_1 a_1 + k_2 a_2 = (k_1 + k_2, 2k_1 + k_2)$ folgt $k_2 = -k_1$ und dann $0 = 2k_1 + k_2 = k_1 = k_2$). Basis von W sei wieder die Standardbasis. Dann ist $\alpha(a_1) = (3 - 2, 2 + 10, -1)$, $\alpha(a_2) = (3 - 1, 2 + 5, -1)$, $(a, b) = (b - a)(1, 2) + (2a - b)(1, 1)$ und

$$A = \begin{pmatrix} 1 & 2 \\ 12 & 7 \\ -1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 12 & 7 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} b - a \\ 2a - b \end{pmatrix} = \begin{pmatrix} 3a - b \\ 2a + 5b \\ -a \end{pmatrix}.$$

7.9 Beispiel:

Sei V der Vektorraum der Polynome vom Grad kleiner als $n \in \mathbb{N}$ in $\mathbb{K}[x]$, $W = \mathbb{K}^n$, und seien $a_1, \ldots, a_n$ verschiedene Elemente von $\mathbb{K}$. Wir definieren $\alpha : V \to W$ durch $\alpha(p) = (p(a_1), p(a_2), \ldots, p(a_n))$. Es ist leicht zu sehen, dass α linear ist. Nun kann ein Polynom $\neq 0$ vom Grad d höchstens d verschiedene Nullstellen haben (wir werden das später im Abschnitt über Polynomringe beweisen). Daraus folgt, dass α injektiv ist: Wenn $p \in V$ und $0 = \alpha(p) = (p(a_1), p(a_2), \ldots, p(a_n))$, dann hat p mindestens n Nullstellen. Da der Grad von p kleiner ist, muss $p = 0$ sein. Daher ist $\operatorname{Ker}(\alpha) = 0$. Nun ist $\dim V = \dim W = n$; nach 5.16 folgt, dass α ein Isomorphismus ist. Anders formuliert: Zu beliebigen $k_1, \ldots, k_n \in \mathbb{K}$ gibt es genau ein Polynom $p \in V$ mit $p(a_i) = k_i$ für alle i. Es ist auch nicht schwer, dieses Polynom (das man das Interpolationspolynom nennt) direkt anzugeben, d.h. die Umkehrabbildung zu α zu konstruieren:
Für festes i sei

$$f_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - a_j}{a_i - a_j} \ .$$

Offenbar ist $f_i(a_i) = 1$ und $f_i(a_j) = 0$, wenn $j \neq i$. Setzt man

$$p = \sum_{i=1}^{n} k_i f_i \ ,$$

so folgt $p(a_j) = k_j$ für alle j, d.h. $\alpha(p) = (k_1, \ldots, k_n)$. Durch Angabe des Interpolationspolynoms wird also die Surjektivität von α direkt gezeigt; die Injektivität folgt wieder aus 5.16.

Wir berechnen jetzt die Matrix zu α, wobei wir in V die Basis $\{1 = x^0, x, x^2, \ldots, x^{n-1}\}$ wählen und in $W = \mathbb{K}^n$ die Standardbasis $\{e_1, \ldots, e_n\}$. Offenbar ist $\alpha(x^j) = (a_1^j, \ldots, a_n^j)$. Daher ist

$$A = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix} .$$

die Matrix von α bezüglich dieser Basen. Eine Matrix dieser Form (oder auch die transponierte Matrix; siehe 7.14 unten) nennt man eine Vandermonde'sche Matrix.

Wir wiederholen die Berechnung der Matrix zu α, wobei wir für $W = \mathbb{K}^n$ die Standardbasis beibehalten, aber in V eine andere Basis wählen, nämlich

$$\{b_0 = 1, b_1 = x - a_1, b_2 = (x - a_1)(x - a_2), \ldots, b_{n-1} = (x - a_1)(x - a_2) \cdots (x - a_{n-1})\}.$$

Da $b_j = x^j +$(Terme vom Grad kleiner j), sind die b_j's linear unabhängig, bilden also wirklich eine Basis von V. Es ist

$$\begin{array}{rcccccc}
\alpha(b_0) & = (1, & 1, & 1, & \dots \ \dots, & 1 &) \\
\alpha(b_1) & = (0, & a_2 - a_1, & a_3 - a_1, & \dots \ \dots, & a_n - a_1 &) \\
\alpha(b_2) & = (0, & 0, & (a_3 - a_1)(a_3 - a_2), & \dots \ \dots, & (a_n - a_1)(a_n - a_2) &) \\
\vdots & \vdots & \vdots & & \ddots & \vdots & \\
\vdots & \vdots & \vdots & & \quad \ddots & \vdots & \\
\alpha(b_{n-1}) & = (0, & 0, & 0, & \dots \ \dots, & (a_n - a_1)(a_n - a_2) \cdots (a_n - a_{n-1}) &)
\end{array}$$

Deshalb ist

$$B = \begin{pmatrix}
1 & & & & \\
1 & a_2 - a_1 & & & \mathbf{0} \\
1 & a_3 - a_1 & (a_3 - a_1)(a_3 - a_2) & & \\
\vdots & \vdots & \vdots & \ddots & \\
1 & a_n - a_1 & (a_n - a_1)(a_n - a_2) & \cdots & (a_n - a_1) \cdots (a_n - a_{n-1})
\end{pmatrix}$$

(nur Nullen oberhalb der Diagonalen) die Matrix von α bezüglich dieser Basen. Eine Matrix dieser Form nennt man eine untere Dreiecksmatrix.

7.10 Satz:

Sei $\dim V = m$ und $\dim W = n$, beide endlich. Wähle in V und W jeweils eine Basis, etwa $\{a_1, \dots, a_m\}$ und $\{b_1, \dots, b_n\}$. Zu jedem $\alpha : V \to W$ sei $A(\alpha)$ die zugehörige Matrix (bzgl. dieser Basen). Dann ist die Abbildung $\alpha \mapsto A(\alpha)$ ein Isomorphismus von $\mathrm{Hom}\,(V, W)$ auf den Vektorraum der $n \times m$–Matrizen.

Beweis: Wenn $\alpha, \beta : V \to W$ linear, dann ist $(\alpha + \beta)(a_j) = \alpha(a_j) + \beta(a_j)$, also ist die j–te Spalte von $A(\alpha + \beta)$ die Summe der j–ten Spalten von $A(\alpha)$ und $A(\beta)$, d.h. $A(\alpha + \beta) = A(\alpha) + A(\beta)$. Ebenso sieht man $A(k\alpha) = kA(\alpha)$. Also ist A linear. Wenn (k_{ij}) irgendeine $n \times m$–Matrix ist, setze $w_j = \sum_{i=1}^{n} k_{ij} b_i$. Nach 5.17 existiert genau eine lineare Abbildung $\alpha : V \to W$ mit $\alpha(a_j) = w_j$ für $j = 1, \dots, m$. Also existiert genau ein $\alpha \in \mathrm{Hom}\,(V, W)$ mit $A(\alpha) = (k_{ij})$. Daher ist A bijektiv.

7.11 Korollar:

Wenn $\dim V,\ \dim W < \infty$, dann ist

$$\dim \mathrm{Hom}\,(V, W) = \dim V \cdot \dim W.$$

Beweis: Wenn $\dim V = m$ und $\dim W = n$, dann ist $\mathrm{Hom}\,(V, W) \cong \{n \times m\text{–Matrizen}\}$ nach 7.10. Es folgt $\dim \mathrm{Hom}\,(V, W) = nm = \dim V \cdot \dim W$.

7.12 Satz:

Sei $X = \{x_1, \dots, x_\ell\}$ eine Basis von U, $Y = \{y_1, \dots, y_m\}$ eine Basis von V und $Z = \{z_1, \dots, z_n\}$ eine Basis von W. Sei $\alpha : V \to W$ linear mit der Matrix $A = (a_{ij})$ bzgl. Y

und Z, und sei $\beta : U \to V$ linear mit der Matrix $B = (b_{ij})$ bzgl. X und Y. Dann ist AB die Matrix von $\alpha\beta$ bzgl. X und Z.

Beweis: Nach Voraussetzung ist

$$\beta(x_h) = \sum_{j=1}^{m} b_{jh} y_j, \ h = 1, \ldots, \ell$$

und

$$\alpha(y_j) = \sum_{i=1}^{n} a_{ij} z_i, \ j = 1, \ldots, m \ .$$

Also ist

$$(\alpha\beta)(x_h) = \sum_{j=1}^{m} b_{jh} \alpha(y_j) = \sum_{j=1}^{m} b_{jh} \sum_{i=1}^{n} a_{ij} z_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij} b_{jh} \right) z_i, \ h = 1, \ldots, \ell \ .$$

Aber $\sum_{j=1}^{m} a_{ij} b_{jh}$ ist gerade der Eintrag von AB in der i–ten Zeile und der h–ten Spalte nach der Definition der Matrizenmultiplikation. (Da A eine $n \times m$– und B eine $m \times \ell$–Matrix ist, ist AB definiert und eine $n \times \ell$–Matrix).

7.13 Korollar:

Wenn definiert, dann sind die Addition und Multiplikation von Matrizen assoziativ und distributiv.

Beweis: Das gilt für lineare Abbildungen.

7.14 Definition: *transponierte Matrix*

Sei A eine $n \times m$–Matrix über $\mathbb{K}$. Die $m \times n$–Matrix A^t, die man erhält, indem man Zeilen und Spalten vertauscht, heißt die zu A transponierte Matrix. Formal: Wenn $A = (a_{ij})$, dann ist $A^t = (b_{ji})$ mit $b_{ji} = a_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$.

7.15 Beispiel/Bemerkung:

(i) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 7 \end{pmatrix}$. Dann $A^t = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 7 \end{pmatrix}$.

(ii) $(A^t)^t = A$

7.16 Satz:

Sei $X = \{x_1, \ldots, x_m\}$ eine Basis von V und $Y = \{y_1, \ldots, y_n\}$ eine Basis von W. Wenn $\alpha : V \to W$ linear und A die Matrix von α bzgl. X und Y ist, dann ist A^t die Matrix von α^* bzgl. Y^* und X^* (siehe 6.3 zur Definition von X^*).

Beweis: Sei $A = (a_{ij})$, d.h. $\alpha(x_\ell) = \sum_{i=1}^{n} a_{i\ell} y_i$, $\ell = 1, \ldots, m$. Zu zeigen: Wenn $\alpha^*(y_j^*) = \sum_{i=1}^{m} b_{ij} x_i^*$, $j = 1, \ldots, n$, dann ist $b_{ij} = a_{ji}$. Es ist für jedes $\ell = 1, \ldots, m$:

$$b_{\ell j} = \sum_{i=1}^{m} b_{ij} x_i^*(x_\ell) = \left(\sum_{i=1}^{m} b_{ij} x_i^* \right)(x_\ell) = \left(\alpha^*(y_j^*) \right)(x_\ell)$$
$$= y_j^* \alpha(x_\ell) = y_j^* \sum_{i=1}^{n} a_{i\ell} y_i = \sum_{i=1}^{n} a_{i\ell} y_j^*(y_i) = a_{j\ell}.$$

7.17 Korollar:

Seien A, B Matrizen so, daß AB definiert ist. Dann ist auch $B^t A^t$ definiert und es gilt

$$(AB)^t = B^t A^t.$$

Beweis: Sei A die Matrix zu α und B die Matrix zu β (bzgl. fester Basen). Dann ist AB die Matrix zu $\alpha\beta$ (nach 7.12), und $(AB)^t$ die Matrix zu $(\alpha\beta)^*$ bzgl. der dualen Basen (7.16). Außerdem (wieder nach 7.16) ist B^t die Matrix zu β^* und A^t die Matrix zu α^*, also (nach 7.12) $B^t A^t$ die Matrix zu $\beta^*\alpha^*$ (immer bzgl. der dualen Basen). Da nach 6.6 $\beta^*\alpha^* = (\alpha\beta)^*$, folgt $B^t A^t = (AB)^t$.

7.18 Definition: *Spalten– und Zeilenrang*

Der Spaltenrang einer Matrix ist die Dimension des Spaltenraumes von A. Entsprechend ist der Zeilenrang definiert.

7.19 Satz:

Sei $\alpha : V \to W$ linear und sei A die Matrix von α (bzgl. fester Basen von V und W). Dann ist der Spaltenrang von A gleich dem Rang von α.

Beweis: Sei $\{x_1, \ldots, x_m\}$ Basis von V und $\{y_1, \ldots, y_n\}$ Basis von W derart, daß $A = (a_{ij})$ die Matrix von α bzgl. dieser Basen ist, d.h. $\alpha(x_j) = \sum_{i=1}^{n} a_{ij} y_i$. Sei $\sigma : \mathbb{K}^n \to W$ definiert durch $\sigma(k_1, \ldots, k_n) = \sum_{i=1}^{n} k_i y_i$. Dann ist σ ein Isomorphismus, und wenn

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

die j–te Spalte von A ist, dann ist $\sigma(a_j) = \sum_{i=1}^{n} a_{ij} y_i = \alpha(x_j)$. Daher ist die Einschränkung von σ auf den Spaltenraum $\langle a_j \,|\, j = 1, \ldots, m \rangle$ von A ein Isomorphismus auf $\langle \alpha(x_j) \,|\, j = 1, \ldots, m \rangle = \operatorname{Im}(\alpha)$. Insbesondere ist also der Spaltenrang von A gleich $\dim \operatorname{Im}(\alpha) = \operatorname{Rg}(\alpha)$.

7.20 Korollar/Definition:

Für jede Matrix A ist der Spaltenrang von A gleich dem Zeilenrang von A. Man spricht daher einfach vom Rang von A und schreibt $\operatorname{Rg}(A)$.

Beweis: Sei A die Matrix von $\alpha : \mathbb{K}^m \to \mathbb{K}^n$ (siehe 7.10). Dann ist

$$\begin{aligned}\text{Spaltenrang von } A &= \operatorname{Rg}(\alpha) \text{ nach 7.19}\\ &= \operatorname{Rg}(\alpha^*) \text{ nach 6.7}\\ &= \text{ Spaltenrang von } A^t \text{ nach 7.16 und 7.19}\\ &= \text{ Zeilenrang von } A\ ,\end{aligned}$$

da die Spalten von A^t die Zeilen von A sind.

7.21 Korollar:

Für jede Matrix A gilt $\operatorname{Rg}(A) = \operatorname{Rg}(A^t)$.

Beweis: Folgt aus 7.20.

7.22 Definition: *elementare Zeilentransformationen*

Die folgenden Umformungen einer Matrix heißen elementare Zeilentransformationen:

Typ I : Addition eines skalaren Vielfachen einer Zeile zu einer anderen Zeile.

Typ II : Vertauschen zweier Zeilen.

Typ III : Multiplikation einer Zeile mit einem Skalar $\neq 0$.

7.23 Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 & 4\\ 0 & 5 & 0 & 6\\ -1 & 3 & 0 & 0 \end{pmatrix} \xrightarrow{1.Z.+(-3)\cdot 3.Z.} \begin{pmatrix} 4 & -7 & 3 & 4\\ 0 & 5 & 0 & 6\\ -1 & 3 & 0 & 0 \end{pmatrix} \xrightarrow{2.Z.\leftrightarrow 3.Z.} \begin{pmatrix} 4 & -7 & 3 & 4\\ -1 & 3 & 0 & 0\\ 0 & 5 & 0 & 6 \end{pmatrix}$$

$$\xrightarrow{1.Z.\cdot\frac{1}{4}} \begin{pmatrix} 1 & -7/4 & 3/4 & 1\\ -1 & 3 & 0 & 0\\ 0 & 5 & 0 & 6 \end{pmatrix}$$

7.24 Bemerkung:

(i) Jede elementare Zeilentransformation läßt sich durch eine elementare Zeilentransformation wieder rückgängig machen.

(ii) Der Zeilenraum einer Matrix ändert sich bei einer elementaren Zeilentransformation nicht. Insbesondere ändert sich der Rang der Matrix nicht.

(iii) Die elementaren Zeilentransformationen entsprechen Multiplikationen von links mit geeigneten Matrizen, den so genannten Elementarmatrizen. Diese erhält man, indem man die Zeilentransformation auf die Einheitsmatrix der entsprechenden Größe anwendet. (Siehe Übungsaufgabe).

7.25 Definition: *Stufenform, strenge Stufenform, Treppenform*

Eine Matrix hat Stufenform (Treppenform), wenn gilt:

(1) Alle Nullzeilen (wenn es welche gibt) stehen am Ende,

(2) Der erste Eintrag $\neq 0$ in der $(i+1)$–ten Zeile steht weiter rechts als der in der i–ten Zeile, falls die $(i+1)$–te Zeile keine Nullzeile ist.

Wenn zusätzlich gilt:

(3) Der erste Eintrag $\neq 0$ in jeder Nicht–Nullzeile ist 1 (genannt eine „führende Eins“),

(4) Wenn eine Spalte eine führende Eins enthält, dann sind alle anderen Einträge in dieser Spalte $= 0$,

dann sprechen wir von strenger Stufenform (Treppennormalform).

7.26 Beispiel:

(i) Die Matrix

$$\begin{pmatrix} 3 & 1 & 4 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 4 & 6 \end{pmatrix}$$

hat keine Stufenform. Vertausche 2. und 3. Zeile. Die neue Matrix hat Stufenform.

(ii) Subtrahiere in der Matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 2 & 1 \end{pmatrix}$$

die 1. Zeile und das doppelte der 2. Zeile von der 3. Zeile. Die neue Matrix hat strenge Stufenform.

7.27 Bemerkung:

Wenn die Matrix A Stufenform hat, dann ist der Rang von A gleich der Anzahl der Nicht–Nullzeilen von A.

Beweis: Offenbar sind diese Zeilen linear unabhängig. Andererseits bilden sie ein Erzeugenden–System des Zeilenraumes von A. Also ist ihre Anzahl gleich der Dimension des Zeilenraumes, d.h. gleich dem Rang von A.

7.28 Satz:

Jede Matrix A kann durch elementare Zeilentransformationen vom Typ I und II (bzw. Typ I, II und III) auf Stufenform (bzw. strenge Stufenform) gebracht werden.

Beweis: Induktion über die Anzahl der Spalten.

1. Fall: Die erste Spalte von A besteht nur aus Nullen, dann ist

$$A = \begin{pmatrix} 0 & \\ \vdots & B \\ 0 & \end{pmatrix},$$

wobei die Matrix B eine Spalte weniger als A hat. Fertig per Induktion.

2. Fall: Es gibt einen Eintrag $\neq 0$ in der ersten Spalte von A. Durch geeignetes Vertauschen von Zeilen läßt sich erreichen, daß dieser Eintrag in der ersten Zeile steht (Typ II). Anschließend addiert man geeignete Vielfache der 1. Zeile zu den übrigen Zeilen derart, daß der erste Eintrag in den übrigen Zeilen 0 wird (Typ I). Die Matrix hat dann folgende Form:

$$\begin{pmatrix} k_1 & k_2 & \cdots & k_m \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}, \text{ mit } k_1 \neq 0.$$

Per Induktion läßt sich B auf Stufenform bringen. Dann hat die ganze Matrix Stufenform.

Wenn B schon strenge Stufenform hat und in der i–ten Zeile und j–ten Spalte eine führende Eins von B steht, dann addiere das $(-k_j)$–fache der i–ten Zeile zur 1. Zeile. Die neue 1. Zeile hat 0 in allen Spalten, welche eine führende Eins von B enthalten. Schließlich multipliziere die 1. Zeile mit k_1^{-1}. Das Ergebnis ist eine Matrix in strenger Stufenform.

7.29 Beispiel:

(i)

$$\begin{pmatrix} 2 & 1 & 1 & 3 & 5 \\ 4 & 2 & 1 & 1 & 2 \\ -2 & 1 & 0 & 0 & 0 \\ -2 & 2 & 1 & 3 & 5 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 1 & 1 & 3 & 5 \\ 0 & 0 & -1 & -5 & -8 \\ 0 & 2 & 1 & 3 & 5 \\ 0 & 3 & 2 & 6 & 10 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 1 & 1 & 3 & 5 \\ 0 & 2 & 1 & 3 & 5 \\ 0 & 0 & -1 & -5 & -8 \\ 0 & 3 & 2 & 6 & 10 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 2 & 1 & 1 & 3 & 5 \\ 0 & 2 & 1 & 3 & 5 \\ 0 & 0 & -1 & -5 & -8 \\ 0 & 0 & 1/2 & 3/2 & 5/2 \end{pmatrix} \longrightarrow \begin{pmatrix} 2 & 1 & 1 & 3 & 5 \\ 0 & 2 & 1 & 3 & 5 \\ 0 & 0 & -1 & -5 & -8 \\ 0 & 0 & 0 & -1 & -3/2 \end{pmatrix}$$

(ii) Die führenden Einsen sind fett gedruckt.

$$\begin{pmatrix} 2 & 1 & 1 & 3 & 5 \\ 4 & 2 & 1 & 1 & 2 \\ -2 & 1 & 0 & 0 & 0 \\ -2 & 2 & 1 & 3 & 5 \end{pmatrix} \longrightarrow \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 3 & 5 \\ 0 & 4 & 1 & 1 & 2 \\ 0 & \mathbf{1} & 1 & 3 & 5 \end{pmatrix} \longrightarrow \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 1 & 3 & 5 \\ 0 & 0 & -1 & -3 & -5 \\ 0 & 0 & -3 & -11 & -18 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 1 & 3 & 5 \\ 0 & 0 & -1 & -3 & -5 \\ 0 & 0 & 0 & -2 & -3 \end{pmatrix} \longrightarrow \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 3 & 5 \\ 0 & 0 & 0 & \mathbf{1} & 3/2 \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 1/2 \\ 0 & 0 & 0 & \mathbf{1} & 3/2 \end{pmatrix}$$

(iii)

$$\begin{pmatrix} \mathbf{1} & 0 & 1 & 2 \\ 2 & 1 & 3 & 4 \\ 3 & 2 & 5 & 1 \\ 4 & 1 & 5 & 8 \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{1} & 1 & 0 \\ 0 & 2 & 2 & -5 \\ 0 & \mathbf{1} & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{1} & 1 & 0 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{1} & 1 & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} \mathbf{1} & 0 & 1 & 0 \\ 0 & \mathbf{1} & 1 & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

7.30 Definition: *lineares Gleichungssystem, Koeffizientenmatrix, Konstantenvektor, erweiterte Matrix, Lösung, (in–)homogenes Gleichungssystem*

Ein lineares Gleichungssystem (LGS) über $\mathbb{K}$ mit n Gleichungen und m Unbestimmten ist ein System

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned} \qquad (*)$$

wobei alle $a_{ij}, b_i \in \mathbb{K}$ sind. Man nennt die Matrix $A = (a_{ij})$ die Koeffizientenmatrix des Systems und den Spaltenvektor $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ den Konstantenvektor. Die $n \times (m+1)$–Matrix $\tilde{A} = (A, b)$, die man erhält, indem man den Konstantenvektor als letzte Spalte an die Koeffizientenmatrix anhängt, heißt die erweiterte Koeffizientenmatrix. Schreibt man schließlich $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, dann läßt sich das System auch schreiben als $Ax = b$. Jedes $x \in \mathbb{K}^m$, für welches diese Gleichung gilt, heißt eine Lösung des Systems. Falls $b = 0$, nennt man das Gleichungssystem homogen, andernfalls inhomogen.

7.31 Fragen:

Gibt es überhaupt Lösungen? Genauer: Wann gibt es Lösungen? (Lösbarkeitskriterien). Wenn es welche gibt, wie findet man eine oder sogar alle?

7.32 Bemerkung/Definition: *triviale Lösung, Lösungsmenge*

(i) Ein homogenes LGS ist immer lösbar, nämlich $x = 0$ tut's. Dies ist die triviale Lösung eines homogenen Systems.

(ii) Es ist sehr nützlich umzudenken: Für jedes $v \in \mathbb{K}^m$ ist $Av \in \mathbb{K}^n$, und die Abbildung $\alpha : v \mapsto Av$ ist eine lineare Abbildung $\mathbb{K}^m \to \mathbb{K}^n$. Die Matrix von α bzgl. der Standardbasis ist gerade A. Daher ist $\mathrm{Im}\,(\alpha)$ der Spaltenraum von A. Die Lösungsmenge X des Gleichungssystems besteht also aus allen $x \in \mathbb{K}^m$ mit der Eigenschaft $\alpha(x) = b$.

7.33 Satz:

(1) Das LGS $(*)$ ist genau dann lösbar, wenn der Rang der Koeffizientenmatrix gleich dem Rang der erweiterten Koeffizientenmatrix ist.

(2) Wenn x_0 eine Lösung von $(*)$ ist, dann erhält man alle Lösungen als $x = x_0 + y$, wobei y eine beliebige Lösung des homogenen Systems $Ax = 0$ ist

Beweis:

(1) $(*)$ ist lösbar $\Leftrightarrow \exists x \in \mathbb{K}^m : \alpha(x) = b \Leftrightarrow b \in \operatorname{Im}(\alpha) \Leftrightarrow \langle b, \operatorname{Im}(\alpha)\rangle = \operatorname{Im}(\alpha) \Leftrightarrow \dim \operatorname{Im}(\alpha) = \dim\langle b, \operatorname{Im}(\alpha)\rangle$.
Die Behauptung folgt, da

$$\begin{aligned} \dim \operatorname{Im}(\alpha) &= \dim \text{ (Spaltenraum von } A) \\ &= \operatorname{Rg}(A) \\ \text{und} \qquad \dim \langle b, \operatorname{Im}(\alpha)\rangle &= \dim \text{ (Spaltenraum von } \tilde{A}) \\ &= \operatorname{Rg}(\tilde{A}) \ . \end{aligned}$$

(2) folgt aus 5.22 (ii).

7.34 Bemerkung:

Das Problem ist also:

(i) Die Ränge von A und $\tilde{A}$ berechnen.

Falls diese gleich sind,

(ii) eine spezielle Lösung von $(*)$ und

(iii) eine Basis von $\operatorname{Ker}(\alpha) = \{y \mid Ay = 0\}$ zu finden (damit hat man dann alle Elemente aus $\operatorname{Ker}(\alpha)$ als beliebige Linearkombination der Basisvektoren).

Dies geht alles gleichzeitig. Der Schlüssel liegt in den elementaren Zeilentransformationen. Es gilt nämlich, wie man leicht sieht: Wendet man eine elementare Zeilentransformation auf die erweiterte Matrix an, so hat das der neuen Matrix entsprechende lineare Gleichungssystem dieselben Lösungen wie das ursprüngliche. Nach 7.28 kann man die erweiterte Matrix durch elementare Zeilentransformationen auf Treppennormalform bringen (ohne die Lösungsmenge zu verändern). Dann läßt sich aber alles ablesen.

7.35 Verfahren zum Lösen des linearen Gleichungssystems $Ax = b$:

1. Schreibe die erweiterte Koeffizientenmatrix $\tilde{A}$ auf.
   Bemerkung: Achte auf die Reihenfolge der Unbestimmten. Bei homogenen Systemen schreibt man nur die Koeffizientenmatrix auf.
2. Forme $\tilde{A}$ mittels elementarer Zeilentransformationen zu einer Matrix $\tilde{A}_1 = (A_1, b_1)$ um derart, daß $\tilde{A}_1$ Treppennormalform hat.

Bemerkung: Falls nach einigen Umformungen eine Matrix (A', b') erreicht wird, für die offensichtlich $\mathrm{Rg}(A') < \mathrm{Rg}(A', b')$ gilt, kann man aufhören. Dann gibt es keine Lösung.

3. Sei $\mathrm{Rg}(A_1) = \mathrm{Rg}(A_1, b_1)$. Dann ergänze (A_1, b_1) mit Nullzeilen derart, daß A_1 zu einer quadratischen Matrix C wird und die führenden Einsen auf der Diagonalen stehen. Anschließend ersetze die Nullen auf der Diagonalen von C durch -1. Sei die neue Matrix $\tilde{D} = (D, d)$.
4. Eine Lösung des linearen Gleichungssystems ist d. Die Spalten von D mit -1 auf der Diagonalen bilden eine Basis des Lösungsraumes des zugehörigen homogenen Systems. Also ist die allgemeine Lösung des Systems: d+(beliebige Linearkombination dieser Spalten).

7.36 Beispiel:

(i)

$$\begin{aligned} 6x_2 + 15x_4 + 18x_5 - x_6 + 30x_7 &= -3x_1 - 9x_3 \\ 2x_1 + 4x_2 + 6x_3 + x_4 + x_6 + 7x_7 + 9 &= 6x_5 \\ 4 + x_1 + 2x_2 + 3x_3 + x_4 + x_6 + 6x_7 &= 2x_5 \\ 4 + x_1 + 2x_2 + 3x_3 + x_4 + 2x_6 + 9x_7 &= 2x_5 \end{aligned}$$

$$\tilde{A} = \begin{pmatrix} 3 & 6 & 9 & 15 & 18 & -1 & 30 & 0 \\ 2 & 4 & 6 & 1 & -6 & 1 & 7 & -9 \\ 1 & 2 & 3 & 1 & -2 & 1 & 6 & -4 \\ 1 & 2 & 3 & 1 & -2 & 2 & 9 & -4 \end{pmatrix} \longrightarrow \tilde{A}_1 = \begin{pmatrix} 1 & 2 & 3 & 0 & -4 & 0 & 1 & -5 \\ 0 & 0 & 0 & 1 & 2 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 \end{pmatrix}$$

(Nullzeile fortgelassen)

$$\longrightarrow \tilde{D} = \begin{pmatrix} 1 & 2 & 3 & 0 & -4 & 0 & 1 & -5 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}$$

Allgemeine Lösung:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} -4 \\ 0 \\ 0 \\ 2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 0 \\ 3 \\ -1 \end{pmatrix}$$

(ii)

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ 2x_2 + 2x_3 + 2x_4 &= -2 \\ x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ x_1 + x_2 + 2x_3 + 3x_4 &= 2 \end{aligned}$$

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

$$\longrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

An der letzten Zeile erkennt man, daß es keine Lösungen gibt.

7.37 Definition: *invertierbare Matrix, Inverse*

Eine $n \times n$–Matrix A heißt <u>invertierbar</u> (oder <u>regulär</u>), wenn eine $n \times n$–Matrix B existiert mit $AB = BA = E_n$ = die $n \times n$–Einheitsmatrix.

7.38 Bemerkung:

(i) A ist invertierbar genau dann, wenn A die Matrix eines Isomorphismus ist.

(ii) Wenn A invertierbar ist und $AB = BA = E_n$, dann ist B eindeutig bestimmt. Man nennt dann $B = A^{-1}$ die zu A <u>inverse Matrix</u>.

(iii) Wenn A und B invertierbar und von gleicher Größe sind, dann ist AB invertierbar, und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.

(iv) $(A^{-1})^{-1} = A$.

7.39 Satz:

Sei A eine $n \times n$–Matrix. Äquivalent sind:

(1) A ist regulär.

(2) $\exists$ $n \times n$–Matrix B mit $BA = E_n$

(3) $\mathrm{Rg}(A) = n$

(4) A ist Produkt von Elementarmatrizen.

Außerdem: Wenn (2) gilt, ist auch $AB = E_n$ und $B = A^{-1}$.

<u>*Beweis:*</u> A ist Matrix einer linearen Abbildung $\alpha : \mathbb{K}^n \to \mathbb{K}^n$ (bzgl. der Standardbasis). Beachtet man 7.19, so folgen die Äquivalenz von (1)–(3) und der Zusatz aus 5.16.

(3)⇒(4) Durch elementare Zeilentransformationen läßt sich A zu einer Matrix T in Treppennormalform umformen. Nach 7.24 (ii) gilt $\mathrm{Rg}\,(T) = \mathrm{Rg}\,(A) = n$. Da T Treppennormalform hat, folgt $T = E_n$. Also gibt es Elementarmatrizen $X_1, \ldots, X_k$ mit $X_k \cdot \ldots \cdot X_1 \cdot A = E$ (vergleiche 7.24 (iii)). Die Elementarmatrizen sind invertierbar, ihre Inversen sind ebenfalls Elementarmatrizen. Daher ist $A = X_1^{-1} \cdot \ldots \cdot X_k^{-1}$ wie behauptet.

(4)⇒(1) Als Produkt invertierbarer Matrizen ist auch A invertierbar.

7.40 Berechnung von A^{-1}:

Forme (A, E) durch elementare Zeilentransformationen zu (B, C) um, derart, daß B Treppennormalform hat. Genau dann ist A invertierbar, wenn $B = E$. Dann ist $C = A^{-1}$.

Beweis: Der Umformung entsprechen Multiplikationen mit Elementarmatrizen. Insgesamt werden also A und E beide mit einer Matrix C multipliziert (C = das Produkt dieser Elementarmatrizen). Wenn $B = CA = E$, dann $C = A^{-1}$.

7.41 Definition: *Matrix des Basiswechsel*

Sei $\dim V = n$ und seien $X = \{x_1, \ldots, x_n\}$ und $X' = \{x'_1, \ldots, x'_n\}$ beides Basen von V. Sei $x'_j = \sum_{i=1}^{n} t_{ij} x_i$ mit $t_{ij} \in \mathbb{K}$ für alle i, j. Dann heißt die $n \times n$–Matrix $T = (t_{ij})$ die Matrix des Basiswechsel von X nach X'.

7.42 Bemerkung:

(i) Wenn T wie oben, dann ist T die Matrix von id_V bzgl. der Basen X' und X.

(ii) Jedes solches T ist invertierbar, T^{-1} ist die Matrix des Basiswechsels von X' nach X.

Beweis:

(i) ist klar.

(ii) Sei S die Matrix des Basiswechsel von X' nach X. Dann ist S die Matrix von id_V bzgl. X und X', also ist (nach 7.12) TS die Matrix von id_V bzgl. der Basen X und X, also $TS = E$.

7.43 Satz: *Änderung der Matrix einer linearen Abbildung bei Basiswechsel*

Sei $\alpha : V \to W$ eine lineare Abbildung und seien X und X' Basen von V und Y und Y' Basen von W. Sei T die Matrix des Basiswechsel von X nach X' und S die Matrix des Basiswechsel von Y nach Y'. Wenn A die Matrix von α bzgl. X und Y und A' die Matrix von α bzgl. X' und Y' ist, dann gilt $A' = S^{-1}AT$.

Beweis: Es ist T die Matrix von id_V bzgl. X' und X, A die Matrix von α bzgl. X und Y, S^{-1} die Matrix von id_W bzgl. Y und Y'. Daher ist (nach 7.12) $S^{-1}AT$ die Matrix von $\mathrm{id}_W \,\alpha\, \mathrm{id}_V = \alpha$ bzgl. X' und Y', also gleich A'.

7.44 Beispiel:

(i) Sei $\alpha : \mathbb{R}^3 \to \mathbb{R}^2$ definiert durch $\alpha(a,b,c) = (a+2b+3c, a-b+c)$. Sei $X = \{e_1, e_2, e_3\}$, $X' = \{(1,1,1),(2,1,0),(5,2,-3)\}$, $Y = \{e_1, e_2\}$, $Y' = \{(6,1),(4,1)\}$. Dann ist

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \end{pmatrix}, T = \begin{pmatrix} 1 & 2 & 5 \\ 1 & 1 & 2 \\ 1 & 0 & -3 \end{pmatrix}, S = \begin{pmatrix} 6 & 4 \\ 1 & 1 \end{pmatrix}.$$

Wir berechnen S^{-1} nach 7.40:

$$\begin{pmatrix} 6 & 4 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -2 & 1 & -6 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 1/2 & -2 \\ 0 & 1 & -1/2 & 3 \end{pmatrix}$$

Dann ist

$$\begin{aligned} A' = S^{-1}AT &= \begin{pmatrix} 1/2 & -2 \\ -1/2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 5 \\ 1 & 1 & 2 \\ 1 & 0 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 1/2 & -2 \\ -1/2 & 3 \end{pmatrix} \begin{pmatrix} 6 & 4 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \end{aligned}$$

Probe: $\alpha(1,1,1) = (6,1)$, $\alpha(2,1,0) = (4,1)$, $\alpha(5,2,-3) = (0,0)$.

(ii) Wir nehmen das Beispiel aus 7.9 wieder auf. Die Matrix S des Basiswechsels von $\{x^0, \ldots, x^{n-1}\}$ zu der dort definierten neuen Basis $\{b_0, \ldots, b_{n-1}\}$ erhält man, indem man die b_i's ausmultipliziert. Zum Beispiel ist

$$b_2 = (x-a_1)(x-a_2) = x^2 - (a_1+a_2)x + a_1a_2$$

und entsprechend die dritte Spalte der Transformationsmatrix

$$\begin{pmatrix} a_1a_2 \\ -(a_1+a_2) \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Offenbar ist S eine obere Dreiecksmatrix mit Einsen auf der Diagonalen. Sie sollten sich für (z.B.) $n = 4$ davon überzeugen, dass $AS = B$ gilt.

7.45 Bemerkung:

Gegeben sei $\alpha : V \to W$. Dann hängt die Matrix A von der Wahl der Basen in V und W ab. Kann man diese Basen "geschickt" wählen, so daß A eine möglichst einfache Form hat? Im Beispiel oben ist A' schöner als A. Dieses Problem wird uns noch beschäftigen.

8 Determinanten

8.1 Definition: *Determinantenfunktion*

Sei $\dim V = n$. Eine Abbildung $D : V^n \to \mathbb{K}$ heißt eine Determinantenfunktion (auf V), wenn folgendes gilt:

(D1) $D(v_1, \ldots, v_{i-1}, v_i + kv_i', v_{i+1}, \ldots, v_n) = D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)$
$+kD(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n)$
für alle $i = 1, \ldots, n$ und alle $k \in \mathbb{K}$, d.h. D ist linear in jeder Komponente.

(D2) $D(v_1, \ldots, v_n) = 0$, falls $v_i = v_j$ für $i \neq j$

(D3) D ist nicht die Nullabbildung.

8.2 Lemma:

Sei D eine Determinantenfunktion. Dann gilt (für $i < j$):

$$D(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) = -D(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n)$$

Beweis: Es gilt nach (D1), (D2)

$$\begin{aligned} 0 &= D(\ldots, \underset{\substack{\uparrow \\ i}}{v_i + v_j}, \ldots, \underset{\substack{\uparrow \\ j}}{v_i + v_j}, \ldots) \\ &= D(\ldots, v_i, \ldots, v_i, \ldots) + D(\ldots, v_i, \ldots, v_j, \ldots) + D(\ldots, v_j, \ldots, v_i, \ldots) \\ &\quad + D(\ldots, v_j, \ldots, v_j, \ldots) \\ &= 0 + D(\ldots, v_i, \ldots, v_j, \ldots) + D(\ldots, v_j, \ldots, v_i, \ldots) + 0, \end{aligned}$$

also $D(\ldots, v_i, \ldots, v_j, \ldots) = -D(\ldots, v_j, \ldots, v_i, \ldots)$.

8.3 Lemma:

Sei D eine Determinantenfunktion und $i \neq j$. Dann gilt:

$$D(\ldots, \underset{\substack{\uparrow \\ i}}{v_i}, \ldots, \underset{\substack{\uparrow \\ j}}{v_j + kv_i}, \ldots) = D(\ldots, v_i, \ldots, v_j, \ldots)$$

Beweis: Nach (D1), (D2) gilt

$$\begin{aligned} D(\ldots, v_i, \ldots, v_j + kv_i, \ldots) &= D(\ldots, v_i, \ldots, v_j, \ldots) + kD(\ldots, v_i, \ldots, v_i, \ldots) \\ &= D(\ldots, v_i, \ldots, v_j, \ldots) + 0. \end{aligned}$$

8.4 Lemma:

Sei D eine Determinantenfunktion. Wenn $\{v_1, \ldots, v_n\}$ linear abhängig ist, dann gilt

$$D(v_1, \ldots, v_n) = 0.$$

Beweis: Sei z.B. $v_n = \sum_{i=1}^{n-1} k_i v_i$ Linearkombination der übrigen v_i. Dann ist nach (D1) und (D2)

$$D(v_1, \ldots, v_n) = \sum_{i=1}^{n-1} k_i D(v_1, \ldots, v_i, \ldots, v_i) = 0.$$

8.5 Satz:

Sei $\{b_1, \ldots, b_n\}$ eine Basis von V und sei $k \in \mathbb{K}$. Dann gibt es genau eine Funktion D auf V^n mit (D1), (D2) und $D(b_1, \ldots, b_n) = k$. Für diese gilt: Wenn $v_i = \sum_{j=1}^{n} a_{ij} b_j$, $j = 1, \ldots, n$, dann ist

$$D(v_1, \ldots, v_n) = k \cdot \sum_{\sigma \in S_n} \operatorname{sign} \sigma \prod_{i=1}^{n} a_{i,\sigma(i)}. \qquad (*)$$

D ist Determinantenfunktion genau dann, wenn $k \neq 0$.

Beweis: Sei D eine Funktion mit (D1), (D2) und $D(b_1, \ldots, b_n) = k$ und seien v_i wie oben. Wegen der Linearität in jeder Komponente (D1) gilt

$$\begin{aligned} D(v_1, \ldots, v_n) &= D\left(\sum_{j=1}^{n} a_{1j} b_j, \sum_{j=1}^{n} a_{2j} b_j, \ldots, \sum_{j=1}^{n} a_{nj} b_j \right) \\ &= \sum_{f:\{1,\ldots,n\} \to \{1,\ldots,n\}} \prod_{i=1}^{n} a_{i,f(i)} D\left(b_{f(1)}, \ldots, b_{f(n)} \right). \end{aligned}$$

Wenn f nicht injektiv ist, dann ist $f(i) = f(j)$ für $i \neq j$, also $D\left(b_{f(1)}, \ldots, b_{f(n)}\right) = 0$ nach (D2). Also braucht man nur über die injektiven f's zu summieren. Diese sind bijektiv, also Permutationen auf $\{1, \ldots, n\}$, d.h. Elemente von S_n. Daher ist

$$D(v_1, \ldots, v_n) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i,\sigma(i)} D\left(b_{\sigma(1)}, \ldots, b_{\sigma(n)} \right).$$

Es genügt deshalb zu zeigen, daß

$$D\left(b_{\sigma(1)}, \ldots, b_{\sigma(n)} \right) = k \operatorname{sign} \sigma$$

für jedes $\sigma \in S_n$ gilt. σ läßt sich als Produkt von Transpositionen schreiben; $\sigma = \tau_1 \cdot \ldots \cdot \tau_r$. Induktion über r:
Sei $r = 0$. Dann ist $\sigma = \mathrm{id}$ und

$$D\left(b_{\sigma(1)}, \ldots, b_{\sigma(n)} \right) = D(b_1, \ldots, b_n) = k = k \operatorname{sign} \sigma.$$

Sei $r > 0$. Die Transposition τ_1 vertausche i und j. Setze $\sigma_0 = \tau_2 \cdot \ldots \cdot \tau_r$, also $\sigma = \tau_1\sigma_0$, und sei $\sigma_0(x) = i$ und $\sigma_0(y) = j$. Dann ist

$$\begin{aligned}
D\left(b_{\sigma(1)}, \ldots, b_{\sigma(n)}\right) &= D\big(b_{\tau_1\sigma_0(1)}, \ldots, \underset{\substack{\uparrow\\ x}}{b_{\tau_1\sigma_0(x)}}, \ldots, \underset{\substack{\uparrow\\ y}}{b_{\tau_1\sigma_0(y)}}, \ldots, b_{\tau_1\sigma_0(n)}\big) \\
&= D\left(b_{\sigma_0(1)}, \ldots, b_{\tau_1(i)}, \ldots, b_{\tau_1(j)}, \ldots, b_{\sigma_0(n)}\right) \\
&= D\left(b_{\sigma_0(1)}, \ldots, b_{\sigma_0(y)}, \ldots, b_{\sigma_0(x)}, \ldots, b_{\sigma_0(n)}\right) \\
&= -D\left(b_{\sigma_0(1)}, \ldots, b_{\sigma_0(x)}, \ldots, b_{\sigma_0(y)}, \ldots, b_{\sigma_0(n)}\right) \text{ nach 8.2,} \\
&= -(k \operatorname{sign}\sigma_0) \text{ nach Induktion,} \\
&= k \operatorname{sign}\tau_1 \operatorname{sign}\sigma_0 \text{ nach 2.11,} \\
&= k \operatorname{sign}(\tau_1\sigma_0) \text{ nach 2.10,} \\
&= k \operatorname{sign}\sigma.
\end{aligned}$$

Dies zeigt: Wenn D eine Funktion mit (D1), (D2) und $D(b_1, \ldots, b_n) = k$ ist, dann gilt $(*)$ für D, insbesondere ist D eindeutig bestimmt.

Es bleibt die Existenz einer solchen Funktion zu zeigen. Dazu zeigen wir: die durch $(*)$ definierte Funktion $D : V^n \to \mathbb{K}$ erfüllt (D1), (D2) und $D(b_1, \ldots, b_n) = k$. Der Fall $n = 1$ ist trivial, sei $n > 1$.
Daß D linear in jeder Komponente ist, zeigt eine triviale Rechnung.
Wenn $v_r = v_s$ für $r \neq s$, dann ist $a_{rj} = a_{sj}$ für alle $j = 1, \ldots, n$. Sei τ die Transposition, welche r und s vertauscht. Die Relation $\sigma \sim \sigma' \Leftrightarrow \sigma = \sigma' \vee \sigma = \sigma'\tau$ ist eine Äquivalenzrelation auf S_n. Wenn σ_μ ein Vertreter–System der Äquivalenzklassen durchläuft, dann sind die Elemente von S_n genau $\{\sigma_\mu, \sigma_\mu\tau \mid \mu = 1, \ldots, n!/2\}$. Also ist

$$\begin{aligned}
D(\ldots, v_r, \ldots, v_s, \ldots) &= k \cdot \sum_{\sigma \in S_n} \operatorname{sign}\sigma \prod_{i=1}^{n} a_{i,\sigma(i)} \\
&= k \cdot \left(\sum_{\mu=1}^{n!/2} \operatorname{sign}\sigma_\mu \prod_{i=1}^{n} a_{i,\sigma_\mu(i)} + \operatorname{sign}\sigma_\mu\tau \prod_{i=1}^{n} a_{i,\sigma_\mu\tau(i)} \right) \\
&= 0,
\end{aligned}$$

denn $\operatorname{sign}\sigma_\mu\tau = -\operatorname{sign}\sigma_\mu$ und

$$\begin{aligned}
\prod_{i=1}^{n} a_{i,\sigma_\mu\tau(i)} &= \left(\prod_{\substack{i=1 \\ i\neq r,s}}^{n} a_{i,\sigma_\mu(i)} \right) a_{r,\sigma_\mu(s)} a_{s,\sigma_\mu(r)} \\
&= \left(\prod_{\substack{i=1 \\ i\neq r,s}}^{n} a_{i,\sigma_\mu(i)} \right) a_{s,\sigma_\mu(s)} a_{r,\sigma_\mu(r)} = \prod_{i=1}^{n} a_{i,\sigma_\mu(i)}.
\end{aligned}$$

Daher ist (D2) für D erfüllt.
Wenn schließlich $b_i = \sum_{j=1}^{n} a_{ij} b_j$, dann ist $a_{ii} = 1$ und $a_{ij} = 0$ für $i \neq j$. Für $\sigma \in S_n$ ist also $\prod_{i=1}^{n} a_{i,\sigma(i)} = 0$, außer $\sigma(i) = i$ für alle $i = 1, \ldots, n$, d.h. $\sigma = \mathrm{id}$. In diesem Fall ist

$\prod_{i=1}^{n} a_{i,\sigma(i)} = 1$. Daher ist

$$D(b_1, \ldots, b_n) = k \cdot \sum_{\sigma \in S_n} \operatorname{sign} \sigma \prod_{i=1}^{n} a_{i,\sigma(i)} = k \operatorname{sign} \mathrm{id} = k.$$

Offenbar ist D nicht die Nullabbildung, erfüllt also auch (D3), genau dann, wenn $k \neq 0$.

8.6 Korollar/Definition: *Determinante*

Es gibt genau eine Abbildung det von der Menge der $n \times n$–Matrizen über $\mathbb{K}$ in $\mathbb{K}$ mit den folgenden Eigenschaften:

(1) det ist linear in allen Zeilen.

(2) $\det A = 0$, falls zwei Zeilen von A gleich sind.

(3) $\det E = 1$.

Man nennt $\det A$ die Determinante von A.

Beweis: Die n Zeilen $z_1, \ldots, z_n$ einer $n \times n$–Matrix A sind Vektoren von $\mathbb{K}^n$. Betrachte die Determinantenfunktion $D : \mathbb{K}^n \times \cdots \times \mathbb{K}^n \to \mathbb{K}$ mit $D(e_1, \ldots, e_n) = 1$. Setzt man $\det A = D(z_1, \ldots, z_n)$, so folgt die Behauptung.

8.7 Bemerkung:

(i) Wenn $A = (a_{ij})$, dann ist also $\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$.

(ii) Wenn A' durch elementare Zeilentransformation aus A hervorgeht, dann bestehen die folgenden Zusammenhänge zwischen $\det A$ und $\det A'$:

- bei Typ I (Addition eines skalaren Vielfachen einer Zeile zu einer anderen) gilt $\det A' = \det A$.

  Beweis: 8.3

- bei Typ II (Vertauschen zweier Zeilen) gilt $\det A' = -\det A$.

  Beweis: 8.2

- bei Typ III (Multiplikation einer Zeile mit einem Skalar $k \neq 0$) gilt $\det A' = k \cdot \det A$.

  Beweis: folgt direkt aus (D1)

8.8 Satz:

Sei A eine $n \times n$–Matrix. Genau dann ist A invertierbar, wenn $\det A \neq 0$.

Beweis: Sei A invertierbar. Dann ist $\operatorname{Rg}(A) = n$ nach 7.39, also bilden die n Zeilen $z_1, \ldots, z_n$ eine Basis von $\mathbb{K}^n$. Daher ist $\det A = D(z_1, \ldots, z_n) \neq 0$ nach 8.5. Wenn A nicht invertierbar, dann sind $z_1, \ldots, z_n$ linear abhängig und daher $\det A = D(z_1, \ldots, z_n) = 0$ nach 8.4.

8.9 Satz:

Seien A und B $n \times n$–Matrizen. Dann gilt

$$\det(AB) = \det(A)\det(B).$$

Beweis: 1. Fall: A ist nicht invertierbar.
Dann ist $\det A = 0$ (nach 8.8), also $\det A \det B = 0$. Wäre AB invertierbar, etwa $(AB)C = E = C(AB)$, dann wäre A invertierbar nach 7.39, Widerspruch. Also ist AB nicht invertierbar, d.h. $\det AB = 0$ nach 8.8, also gilt in diesem Fall die Behauptung.

2.Fall: A ist invertierbar.
Dann ist A Produkt von Elementarmatrizen, etwa $A = X_1 \cdot \ldots \cdot X_s$ (7.39). Induktion über s:
Sei $s = 0$. Dann ist $A = E$, $AB = B$, $\det A = 1$, also $\det AB = \det B = 1 \cdot \det B = \det A \det B$.
Sei $s = 1$. Dann ist $A = X_1$ eine Elementarmatrix. Die Behauptung folgt dann aus 8.7 (ii).
Sei $s > 1$. Dann ist $A = X_1 A_0$ mit $A_0 = X_2 \cdot \ldots \cdot X_s$ und es gilt

$$\begin{aligned} \det(AB) &= \det(X_1(A_0B)) = \det(X_1)\det(A_0B) \quad (s = 1) \\ &= \det(X_1)\det(A_0)\det(B) \quad \text{(Induktion)} \\ &= \det(X_1A_0)\det(B) = \det(A)\det(B). \end{aligned}$$

8.10 Korollar:

Wenn A invertierbar, dann gilt

$$\det(A^{-1}) = (\det A)^{-1}.$$

Beweis: Es ist $1 = \det E = \det(A^{-1}A) = \det(A^{-1})\det(A)$. Die Behauptung folgt.

8.11 Korollar:

Seien A und T $n \times n$–Matrizen und T invertierbar. Dann gilt

$$\det(T^{-1}AT) = \det(A).$$

Beweis: Es ist $\det(T^{-1}AT) = (\det T)^{-1} \det A \det T = \det A$.

8.12 Korollar:

Sei A eine $n \times n$–Matrix, dann ist

$$\det(A) = \det(A^t).$$

(Also sind zur Berechnung auch elementare Spaltentransformationen erlaubt).

Beweis: Wenn A nicht regulär ist, dann ist auch A^t nicht regulär, da beide den gleichen Rang haben nach 7.21. Dann also $\det A = 0 = \det A'$ nach 8.8. Wenn A regulär ist, dann ist $A = X_1 \cdot \ldots \cdot X_s$ mit Elementarmatrizen X_i nach 7.39. Daher ist $\det A = \det X_1 \cdot \ldots \cdot \det X_s$. Andererseits ist $A^t = X_s^t \cdot \ldots \cdot X_1^t$ nach 7.17. Also $\det A^t = \det X_s^t \cdot \ldots \cdot \det X_1^t$. Es genügt daher $\det X = \det X^t$ für Elementarmatrizen X zu zeigen. Dies ist eine Übungsaufgabe.

8.13 Bemerkung:

Was wir über die Zeilen bewiesen haben, gilt auch für die Spalten (nach 8.12), z.B.: Wenn d eine Abbildung von den $n\times n$–Matrizen in $\mathbb{K}$ ist mit: d ist linear in allen Spalten, $d(A) = 0$ falls A zwei gleiche Spalten hat, und $d(E) = 1$, dann ist $d(A) = \det A$.

8.14 Bezeichnung:

Sei A eine $n\times n$–Matrix und sei $1 \leq i, j \leq n$. Mit A_{ij} bezeichnen wir die $(n-1)\times(n-1)$–Matrix, welche durch Streichen der i–ten Zeile und j–ten Spalte aus A entsteht.

8.15 Satz: *Entwicklungssatz für Determinanten*

Sei $A = (a_{ij})$ eine $n\times n$–Matrix und sei ein i mit $1 \leq i \leq n$ fest gewählt. Dann gilt

(1)
$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ji} \det A_{ji}.$$

(„Entwicklung nach der i–ten Spalte“.)

(2)
$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}.$$

(„Entwicklung nach der i–ten Zeile“.)

Beweis: Es genügt, die erste Behauptung zu zeigen. Die zweite folgt dann mit 8.12 durch Transponieren, denn $(A^t)_{ji} = (A_{ij})^t$, wie man leicht sieht. Dazu definieren wir die Abbildung D von der Menge der $n\times n$–Matrizen über $\mathbb{K}$ in $\mathbb{K}$ durch

$$D(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ji} \det A_{ji}.$$

und zeigen

(1) D ist linear in allen Zeilen.

(2) $D(A) = 0$, falls zwei Zeilen von A gleich sind.

(3) $D(E) = 1$.

Nach 8.6 ist dann $D = \det$, und das ist die Behauptung. Die dritte Aussage ist offenkundig, und die erste eine leichte Übung. Wenn zwei Zeilen in A gleich sind, etwa die r–te und die s–te mit $r < s$, dann hat auch A_{ji} noch zwei gleiche Zeilen, solange $j \neq r, s$, und daher dann $\det A_{ji} = 0$. Also ist

$$D(A) = (-1)^{i+r} a_{ri} \det A_{ri} + (-1)^{i+s} a_{si} \det A_{si} \ .$$

Da die r–te und die s–te Zeile von A gleich sind, ist $a_{ri} = a_{si}$. Aber auch A_{ri} und A_{si} haben die gleichen Zeilen, nur in einer anderen Reihenfolge. Vertauscht man die r–te Zeile von A_{si} der Reihe nach mit den Zeilen von $r+1$ bis $s-1$, dann erhält man gerade A_{ri}. Durch

diese $(s-1)-(r+1)+1 = s-r-1$ vielen Vertauschungen ändert sich das Vorzeichen der Determinante nach 8.7 um den Faktor $(-1)^{s-r-1}$. Daher ist $\det A_{ri} = (-1)^{s-r-1} \det A_{si}$. Setzt man dies oben ein, so ergibt sich

$$\begin{aligned} D(A) &= (-1)^{i+r} a_{si} (-1)^{s-r-1} \det A_{si} + (-1)^{i+s} a_{si} \det A_{si} \\ &= (-1)^{i+s-1} a_{si} \det A_{si} \, [(-1)^0 + (-1)^1] \\ &= 0 \ , \end{aligned}$$

was zu zeigen war.

8.16 Korollar: *Determinante von Dreiecksmatrizen*

Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente.

Beweis: Entwickeln nach der 1. Zeile und Induktion.

8.17 Beispiel: *Vandermonde'schen Determinante*

Wir nehmen noch einmal das Beispiel der Vandermonde'schen Matrix auf; vergleiche 7.9. Dass die Determinante einer solchen Matrix A nicht 0 ist, folgt schon aus 8.8, denn A ist die Matrix eines Isomorphismus'. Man kann die Determinante aber auch leicht genau berechnen: Da $AS = B$ (mit der Notation von 7.9 und 7.44), folgt aus dem Produktsatz für Determinanten, dass $\det A \det S = \det B$. Nach dem vorangehenden Korollar ist $\det S = 1$ und

$$\det B = \prod_{1 \le i < j \le n} a_j - a_i \ .$$

Dies ist also auch die Determinante von A; man spricht von einer Vandermonde'schen Determinante.

8.18 Definition: *adjungierte Matrix*

Sei A eine $n \times n$–Matrix, und sei $B = (b_{ij})$ die $n \times n$–Matrix, welche durch

$$b_{ij} = (-1)^{i+j} \det A_{ji}$$

definiert ist. Dann heißt B die zu A adjungierte Matrix, $B = \mathrm{adj}A$.

8.19 Satz:

Sei A eine $n \times n$–Matrix. Dann ist

$$A(\mathrm{adj}A) = (\mathrm{adj}A)A = (\det A)E.$$

Insbesondere gilt $(\det A)A^{-1} = \mathrm{adj}A$, falls A regulär ist.

Beweis: Wir zeigen zuerst $A(\mathrm{adj}A) = (\det A)E$, d.h.

$$\sum_{k=1}^{n} a_{ik} (-1)^{k+j} \det A_{jk} = \begin{cases} \det A & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

Sei zunächst $i = j$. Dann ist $\sum_{k=1}^{n} a_{ik}(-1)^{k+i} \det A_{ik} = \det A$ nach 8.15. Sei nun $i \neq j$. Es bezeichne $\tilde{A}$ die Matrix, welche man erhält, indem man die j–te Zeile von A durch die i–te Zeile ersetzt. Dann hat $\tilde{A}$ zwei gleiche Zeilen, also ist

$$0 = \det \tilde{A} = \sum_{k=1}^{n} \tilde{a}_{jk}(-1)^{k+j} \det \tilde{A}_{jk}.$$

Aber $\tilde{a}_{jk} = a_{ik}$ und $\tilde{A}_{jk} = A_{jk}$, deshalb $0 = \sum_{k=1}^{n} a_{ik}(-1)^{k+j} \det A_{jk}$. Also ist $A(\mathrm{adj}A) = (\det A)E$.

Ersetzt man A durch A^t und beachtet $\det A^t = \det A$, so erhält man $A^t(\mathrm{adj}A^t) = (\det A)E$. Es ist eine leichte Übung, $\mathrm{adj}A^t = (\mathrm{adj}A)^t$ zu zeigen, also gilt $A^t(\mathrm{adj}A)^t = (\det A)E$. Durch Transponieren folgt $(\mathrm{adj}A)A = (\det A)E$ (siehe 7.17).

8.20 Korollar: *Cramer'sche Regel*

Sei A eine reguläre $n \times n$–Matrix und $Ax = b$ mit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Sei A_k die Matrix, welche man erhält, indem man die k–te Spalte von A durch b ersetzt. Dann ist

$$x_k = \frac{\det A_k}{\det A}.$$

Beweis: Es gilt durch Entwicklung nach der k–ten Spalte

$$\det A_k = \sum_{j=1}^{n} (-1)^{k+j} b_j \det A_{jk} \ .$$

Es ist $(\det A)x = (\det A)A^{-1}b = (\mathrm{adj}A)b$, also

$$\begin{aligned} (\det A)x_k &= \sum_{j=1}^{n} (\mathrm{adj}A)_{kj} b_j \\ &= \sum_{j=1}^{n} b_j (-1)^{k+j} \det A_{jk} \\ &= \det A_k \ , \end{aligned}$$

wobei die letzte Gleichheit durch Entwickeln nach der k–ten Spalte folgt. Daher ist $x_k = \frac{\det A_k}{\det A}$.

8.21 Zur Berechnung von Determinanten:

(1) n klein (d.h. höchstens 3)

(i) $n = 1$
Dann ist $\det(a) = a$.

(ii) $n = 2$
Dann ist

$$\begin{aligned}\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \operatorname{sign}(\mathrm{id}) a_{1,\mathrm{id}(1)} a_{2,\mathrm{id}(2)} + \operatorname{sign}(\tau) a_{1,\tau(1)} a_{2,\tau(2)} \\ &= a_{11}a_{22} - a_{12}a_{21},\end{aligned}$$

wobei $\tau : 1 \leftrightarrow 2$.

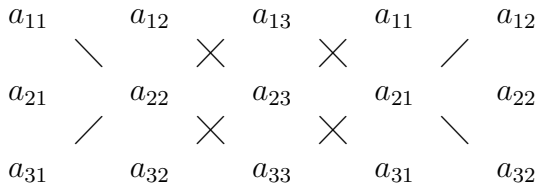
(iii) $n = 3$
Die Permutationen in S_3 sind folgende:

gerade	ungerade
id	$1 \leftrightarrow 2$
$1 \to 2 \to 3 \to 1$	$1 \leftrightarrow 3$
$1 \to 3 \to 2 \to 1$	$2 \leftrightarrow 3$

Also ist

$$\begin{aligned}\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =& a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ & - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.\end{aligned}$$

Eine einfache Merkhilfe ist die Regel von Sarrus:



(2) Spezielle Formen

(i) Eine Zeile (oder Spalte) enthält viele Nullen. Entwickle nach dieser. Das reduziert das Problem auf die Berechnung weniger $(n-1)\times(n-1)$–Determinanten.

(ii) Die Matrix

$$A = \begin{pmatrix} a_{11} & & 0 \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

ist untere Dreiecksmatrix. Dann ist

$$\det A = \prod_{i=1}^{n} a_{ii}$$

und analog für obere Dreiecksmatrizen. Das steht schon in 8.16.

(iii) Block–Dreiecksform
Sei

$$D = \begin{pmatrix} A & 0 \\ B & C \end{pmatrix}$$

mit quadratischen Matrizen A und C, nicht notwendig von der gleichen Größe. Dann ist

$$\det D = \det A \cdot \det C.$$

Beweis: Entwickeln nach der ersten Zeile und Induktion über die Größe von A. Übergang zur Transponierten zeigt, daß entsprechendes für obere Block–Dreiecksmatrizen gilt. Außerdem läßt sich diese Regel natürlich mehrfach anwenden, z.B.: Für

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & 0 \\ 0 & A_{32} & A_{33} \end{pmatrix}$$

mit quadratischen Untermatrizen A_{11}, A_{22}, A_{33} gilt

$$\det A = \det A_{11} \cdot \det A_{22} \cdot \det A_{33}.$$

(iv) Vorsicht! Einige nahe liegende „Vereinfachungen“ zur Berechnung von Determinanten sind leider im Allgemeinen falsch.
Betrachte z.B. für $n = m\ell$ eine Aufteilung der $n\times n$–Matrix A in ℓ^2 Untermatrizen M_{ij} vom Format $m\times m$:

$$A = \begin{pmatrix} M_{11} & \cdots & M_{1\ell} \\ \vdots & & \vdots \\ M_{\ell 1} & \cdots & M_{\ell\ell} \end{pmatrix}. \qquad (*)$$

Man kann die $\ell\times\ell$–Matrix

$$B = \begin{pmatrix} \det M_{11} & \cdots & \det M_{1\ell} \\ \vdots & & \vdots \\ \det M_{\ell 1} & \cdots & \det M_{\ell\ell} \end{pmatrix}.$$

bilden und anschließend deren Determinante berechnen; im Allgemeinen ist jedoch $\det B \neq \det A$.
Man kann auch A als $\ell\times\ell$–Matrix betrachten, deren Einträge nun aber keine Skalare, sondern $m\times m$–Matrizen sind, und die übliche Determinantenformel benutzen, um eine neue $m\times m$–Matrix

$$D = \sum_{\sigma\in S_\ell} \operatorname{sign}\sigma \prod_{i=1}^{\ell} M_{i,\sigma(i)}$$

zu berechnen; auch hier gilt aber im Allgemeinen $\det D \neq \det A$.
Beispiel:

$$A = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

mit

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

Dann ist $\det M_{11} = \det M_{22} = 0$, $\det M_{12} = \det M_{21} = 1$, also (mit den obigen Bezeichnungen)

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$D = M_{11}M_{22} - M_{12}M_{21} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & -1 \end{pmatrix},$$

somit $\det B = -1$, $\det D = 2$. Das richtige Ergebnis ist aber $\det A = 1$. In diesem Beispiel sind die M_{ij} alle Dreiecksmatrizen, außerdem kommutieren M_{11} und M_{22} sowie M_{12} und M_{21} jeweils miteinander.

(v) Wenn alle M_{ij} in $(*)$ obere Dreiecksmatrizen sind, d.h.

$$M_{ij} = \begin{pmatrix} a_{ij} & & & & \\ & b_{ij} & & * & \\ & & c_{ij} & & \\ & 0 & & \ddots & \\ & & & & z_{ij} \end{pmatrix},$$

dann ist

$$\det A = \det(a_{ij}) \cdot \det(b_{ij}) \cdot \ldots \cdot \det(z_{ij}).$$

Beweis: am Beispiel $\ell = 3$.

$$A = \left(\begin{array}{ccc|ccc|ccc} a_{11} & & * & a_{12} & & * & a_{13} & & * \\ & b_{11} & & & b_{12} & & & b_{13} & \\ & c_{11} & & & c_{12} & & & c_{13} & \\ 0 & & \ddots & 0 & & \ddots & 0 & & \ddots \\ \hline a_{21} & & * & a_{22} & & * & a_{23} & & * \\ & b_{21} & & & b_{22} & & & b_{23} & \\ & c_{21} & & & c_{22} & & & c_{23} & \\ 0 & & \ddots & 0 & & \ddots & 0 & & \ddots \\ \hline a_{31} & & * & a_{32} & & * & a_{33} & & * \\ & b_{31} & & & b_{32} & & & b_{33} & \\ & c_{31} & & & c_{32} & & & c_{33} & \\ 0 & & \ddots & 0 & & \ddots & 0 & & \ddots \end{array}\right)$$

Vertauscht man die mit a_{21} bzw. mit a_{31} beginnenden Zeilen nach oben an die zweite bzw. dritte Position und ebenso die mit a_{12} bzw. mit a_{13} beginnenden Spalten nach vorne, so hat man ebenso viele Zeilen– wie Spaltenvertauschungen

vorgenommen. Daher hat die neue Matrix

$$\left(\begin{array}{ccc|ccc|ccc|ccc}
a_{11} & a_{12} & a_{13} & & * & & & * & & & * & \\
a_{21} & a_{22} & a_{23} & & & & & & & & & \\
a_{31} & a_{32} & a_{33} & & & & & & & & & \\
\hline
 & & & b_{11} & & * & b_{12} & & * & b_{13} & & * \\
 & 0 & & & c_{11} & & & c_{12} & & & c_{13} & \\
 & & & 0 & & \ddots & 0 & & \ddots & 0 & & \ddots \\
\hline
 & & & b_{21} & & * & b_{22} & & * & b_{23} & & * \\
 & 0 & & & c_{21} & & & c_{22} & & & c_{23} & \\
 & & & 0 & & \ddots & 0 & & \ddots & 0 & & \ddots \\
\hline
 & & & b_{31} & & * & b_{32} & & * & b_{33} & & * \\
 & 0 & & & c_{31} & & & c_{32} & & & c_{33} & \\
 & & & 0 & & \ddots & 0 & & \ddots & 0 & & \ddots
\end{array}\right)$$

die gleiche Determinante wie A. Aber diese Matrix hat Block–Dreiecksform, und der untere Block hat (per Induktion) Determinante $\det(b_{ij}) \cdot \det(c_{ij}) \cdot \ldots$. Daher folgt die Behauptung aus (iii) oben.

(vi) Natürlich gilt (v) auch, wenn alle M_{ij} in $(*)$ untere Dreiecksmatrizen sind. Das Beispiel in (iv) zeigt, daß (v) im Allgemeinen nicht mehr gültig bleibt, wenn sowohl obere als auch untere Dreiecksmatrizen auftreten.

(vii) Schließlich sei vermerkt, daß mit den Bezeichnungen von (iv) tatsächlich $\det D = \det A$ gilt, wenn die M_{ij} alle miteinander kommutieren. (Dies läßt sich natürlich im Allgemeinen nicht einfach an der Form erkennen.)
Beispiel: Die vier Matrizen

$$M_{11} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix},\ M_{12} = \begin{pmatrix} -1 & 2 \\ -2 & 1 \end{pmatrix},\ M_{21} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix},\ M_{22} = \begin{pmatrix} 4 & -3 \\ 3 & 1 \end{pmatrix}$$

kommutieren miteinander. Daher läßt sich für

$$A = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

die Determinante berechnen als $\det A = \det D$, wobei

$$D = M_{11}M_{22} - M_{12}M_{21} = \begin{pmatrix} 3 & -2 \\ 2 & 1 \end{pmatrix},\ \text{also } \det A = 7.$$

Beweis: wird hier nur angedeutet: Über einem genügend großen Körper (z.B. $\mathbb{C}$) gibt es eine invertierbare $m \times m$ Matrix S derart, daß alle Matrizen $S^{-1}M_{ij}S$ obere Dreiecksform haben. Verwende dann (v).

(3) Allgemeiner Fall
Verwende elementare Zeilentransformationen vom Typ I (d.h. Addition des Vielfachen einer Zeile zu einer anderen). Dabei ändert sich die Determinante nicht (nach 8.3). Die Matrix kann mit solchen Operationen auf obere (oder untere) Dreiecksform gebracht werden. Dann kann man die Determinante als Produkt der Diagonal–Elemente ablesen.

8.22 Definition: *Ähnlichkeit von Matrizen*

Zwei $n \times n$–Matrizen A und B heißen ähnlich, wenn es eine reguläre $n \times n$–Matrix T gibt mit $B = T^{-1}AT$.

8.23 Bemerkung/Definition: *Determinante einer linearen Abbildung*

(i) Ähnlichkeit ist eine Äquivalenzrelation.

(ii) Ähnliche Matrizen haben die gleiche Determinante (8.11).

(iii) Wenn $\alpha \in \mathrm{End}\,(V)$ und A (bzw. A') die Matrix von α bzgl. einer Basis B (bzw. B') von V ist, dann sind A und A' ähnlich (siehe 7.43). Nach (ii) ist also $\det A = \det A'$, d.h. unabhängig von der Wahl der Basis. Man kann daher einfach von $\det \alpha$, der Determinante von α, sprechen.

9 Eigenwerte und Eigenvektoren

Sei V ein n–dimensionaler Vektorraum über dem Körper $\mathbb{K}$.

9.1 Definition: *Eigenwert, Eigenvektor*

Sei α ein Endomorphismus von V und sei $\lambda \in \mathbb{K}$. Man nennt λ einen Eigenwert (EW) von α, wenn es einen Vektor $0 \neq v \in V$ gibt mit $\alpha(v) = \lambda v$. Jedes solche v heißt ein Eigenvektor (EV) von α zum Eigenwert λ.

Entsprechend sind Eigenwerte und Eigenvektoren zu einer $n \times n$–Matrix definiert.

9.2 Satz:

Sei α ein Endomorphismus von V und sei $\lambda \in \mathbb{K}$. Äquivalent sind:

(1) λ ist Eigenwert von α.

(2) $\operatorname{Ker}(\lambda - \alpha) \neq \{0\}$

(3) $\det(\lambda - \alpha) = 0$

(In (2) und (3) ist λ als $\lambda \operatorname{id}_V$ zu lesen.)

Zusatz: In diesem Fall sind die Eigenvektoren von α zum Eigenwert λ genau die Vektoren $\neq 0$ in $\operatorname{Ker}(\lambda - \alpha)$. Man nennt $\operatorname{Ker}(\lambda - \alpha)$ den Eigenraum von α zum Eigenwert λ. (Als Kern einer linearen Abbildung ist dies natürlich ein Unterraum.)

Beweis: Es gilt $\alpha(v) = \lambda v$ genau dann, wenn $(\lambda - \alpha)(v) = 0$, also genau dann, wenn $v \in \operatorname{Ker}(\lambda - \alpha)$. Dies ist die Äquivalenz von (1) und (2). Weiter gilt (2) genau dann, wenn $\lambda - \alpha$ nicht injektiv ist (nach 5.10), also genau dann, wenn $\lambda - \alpha$ nicht invertierbar ist (nach 5.16), also genau dann, wenn $\det(\lambda - \alpha) = 0$ (nach 8.8).

9.3 Definition: *charakteristisches Polynom*

(1) Sei A eine $n \times n$–Matrix. Das charakteristische Polynom von A ($\operatorname{char\,pol} A$) ist das Polynom

$$\det(xE - A) \in \mathbb{K}[x].$$

(2) Sei α ein Endomorphismus von V. Das charakteristische Polynom von α ist $\operatorname{char\,pol} \alpha = \operatorname{char\,pol} A$, wobei A die Matrix von α bzgl. einer Basis von V ist.

9.4 Bemerkung:

(1) $\operatorname{char\,pol} A$ ist ein normiertes (d.h. der führende Koeffizient ist 1) Polynom vom Grad n. Das absolute Glied ist gerade $(-1)^n \det A$.

(2) Wie in 8.23 sieht man, daß ähnliche Matrizen das gleiche charakteristische Polynom haben. Daher ist $\operatorname{char\,pol} \alpha$ unabhängig von der Wahl der Basis.

9.5 Definition: *Nullstelle*

Sei $p \in \mathbb{K}[x]$ ein Polynom. Eine Nullstelle (in $\mathbb{K}$) von p ist ein $\lambda \in \mathbb{K}$ mit $p(\lambda) = 0$.

9.6 Satz:

Sei α ein Endomorphismus von V und $\lambda \in \mathbb{K}$. Genau dann ist λ ein Eigenwert von α, wenn λ eine Nullstelle von $\operatorname{char pol} \alpha$ ist.

Beweis: Sei $\lambda \in \mathbb{K}$ und $p(x)$ das charakteristische Polynom von α. Dann ist $p(x) = \det(x - \alpha)$, also $p(\lambda) = \det(\lambda - \alpha)$. Die Behauptung folgt aus 9.2.

9.7 Definition: *Diagonalmatrix, diagonalisierbar*

Die $n \times n$–Matrix $A = (a_{ij})$ heißt Diagonalmatrix, wenn $a_{ij} = 0$ für $i \neq j$. Man nennt A diagonalisierbar, wenn A zu einer Diagonalmatrix ähnlich ist. Ein Endomorphismus α von V heißt diagonalisierbar, wenn es eine Basis von V derart gibt, daß die Matrix von α bzgl. dieser Basis eine Diagonalmatrix ist.

9.8 Bemerkung:

Genau dann ist α diagonalisierbar, wenn es eine Basis von V gibt, welche aus Eigenvektoren von α besteht. Die verschiedenen Einträge auf der Diagonalen sind dann genau die Eigenwerte von α.

9.9 Lemma:

Seien $v_1, \ldots, v_s$ Eigenvektoren von α zu verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_s$. Dann ist $\sum_{i=1}^{s} v_i \neq 0$.

Beweis: Induktion über s: Klar für $s = 1$, da per Definition 0 kein Eigenvektor ist. Sei $s > 1$ und $w_i = (\lambda_s - \lambda_i)v_i$ für $i = 1, \ldots, s-1$. Dann ist w_i ein Eigenvektor zum Eigenwert λ_i. Daher ist

$$(\lambda_s - \alpha)\left(\sum_{i=1}^{s} v_i\right) = \sum_{i=1}^{s}(\lambda_s - \lambda_i)v_i = \sum_{i=1}^{s-1} w_i \neq 0$$

per Induktion. Also ist $\sum_{i=1}^{s} v_i \neq 0$.

9.10 Satz:

Sei α ein Endomorphismus von V und seien $\lambda_1, \ldots, \lambda_s$ die verschiedenen Eigenwerte von α (in $\mathbb{K}$). Für $j = 1, \ldots, s$ sei $d_j = \dim \operatorname{Ker}(\lambda_j - \alpha)$. Dann gilt:

(1) $\sum_{j=1}^{s} d_j \leq n$

(2) Genau dann ist α diagonalisierbar, wenn $\sum_{j=1}^{s} d_j = n$.

Beweis:

(1) Für $j = 1, \ldots, s$ sei U_j der Eigenraum zum Eigenwert λ_j, also $d_j = \dim U_j$, und sei $\{b_{ji} \,|\, i = 1, \ldots, d_j\}$ eine Basis von U_j. Dann ist $\{b_{ji} \,|\, j = 1, \ldots, s,\ i = 1, \ldots, d_j\}$ linear unabhängig, denn wenn

$$0 = \sum_{j,i} k_{ji} b_{ji} = \sum_{j=1}^{s} v_j, \quad \text{wobei } v_j = \sum_i k_{ji} b_{ji},$$

und (für festes j) nicht alle $k_{ji} = 0$, dann ist $v_j \neq 0$, also ein Eigenvektor zu λ_j ; dies widerspricht 9.9. Nach 4.14 ist dann

$$n \geq |\{b_{ji} \,|\, j = 1, \ldots, s,\ i = 1, \ldots, d_j\}| = \sum_{j=1}^{s} d_j.$$

(2) Wenn „=“, dann sind die b_{ji} eine Basis (wieder nach 4.14). Da sie alle Eigenvektoren von α sind, ist α nach 9.8 diagonalisierbar. Umgekehrt sei α diagonalisierbar. Nachdem man eventuell die Basisvektoren umordnet, hat dann α die Matrix

$$\begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_s & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & \lambda_s \end{pmatrix},$$

wobei λ_i genau d_i–mal in der Hauptdiagonale auftritt. Also ist $\sum_{j=1}^{s} d_j = n$.

9.11 Korollar:

Sei A eine $n \times n$–Matrix mit n verschiedenen Eigenwerten. Dann ist A diagonalisierbar.

9.12 Bemerkung:

Wenn A diagonalisierbar ist, dann zerfällt das charakteristische Polynom in Linearfaktoren. Die Umkehrung gilt nicht!

9.13 Definition: *direkte Summe, Projektion*

Sei V ein VR, und seien $U_1, \ldots, U_s$ Unterräume. Man nennt V die direkte Summe der U_i, und schreibt

$$V = \sum_{i=1}^{s} \oplus\, U_i,$$

wenn jeder Vektor $v \in V$ sich eindeutig schreiben läßt als $v = \sum_{i=1}^{s} u_i$ mit $u_i \in U_i$.

Man sagt, die Summe von Unterräumen U_i, $i = 1, \ldots, s$, ist direkt, falls $\sum_{i=1}^{s} U_i = \sum_{i=1}^{s} \oplus U_i$.

Wenn $V = \sum_{i=1}^{s} \oplus U_i$, sei $\pi_i : V \to U_i$ definiert durch

$$\pi_i \left(\sum_{j=1}^{s} u_j \right) = u_i.$$

Man nennt π_i die Projektion von V auf U_i

9.14 Bemerkung:

(i) Wenn $V = \sum_{i=1}^{s} \oplus U_i$ und $\pi_i : V \to U_i$ die Projektion, dann ist π_i ein Epimorphismus.

(ii) Die Summe von Eigenräumen zu verschiedenen Eigenwerten (eines Endomorphismus' α) ist direkt.

(iii) α ist diagonalisierbar genau dann, wenn $V = \sum_{i=1}^{s} \oplus U_i$, wobei U_i der Eigenraum zum Eigenwert λ_i ist, und $\lambda_1, \ldots, \lambda_s$ die verschiedenen Eigenwerte von α sind.

9.15 Satz:

Seien α und β diagonalisierbare Endomorphismen von V, welche miteinander kommutieren (d.h. $\alpha\beta = \beta\alpha$). Dann sind α und β gemeinsam diagonalisierbar (d.h. es gibt eine Basis, die aus Eigenvektoren von α und β besteht).

Beweis: Seien $\lambda_1, \ldots, \lambda_s$ die Eigenwerte von α und $U_1, \ldots, U_s$ die zugehörigen Eigenräume. Wenn $u_i \in U_i$, dann ist auch $u_i' = \beta(u_i) \in U_i$, denn $\alpha(u_i') = \alpha\beta(u_i) = \beta\alpha(u_i) = \beta(\lambda_i u_i) = \lambda_i(\beta(u_i)) = \lambda_i u_i'$. Nach 9.14 ist $V = \sum_{i=1}^{s} \oplus U_i$, weil α diagonalisierbar ist. Sei $\{b_1, \ldots, b_n\}$ eine Basis von V, welche aus Eigenvektoren von β besteht (existiert, da β diagonalisierbar ist), etwa $\beta(b_j) = \mu_j b_j$. Jedes b_j läßt sich eindeutig schreiben als $b_j = \sum_{i=1}^{s} u_{ij}$ mit $u_{ij} \in U_i$. Dann ist

$$\sum_{i=1}^{s} \beta(u_{ij}) = \beta(b_j) = \mu_j b_j = \sum_{i=1}^{s} \mu_j u_{ij}.$$

Da $u_{ij}' = \beta(u_{ij}) \in U_i$ für jedes i (siehe oben) und $\mu_j u_{ij} \in U_i$, folgt aus der Eindeutigkeit $\beta(u_{ij}) = \mu_j u_{ij}$. Wenn also $u_{ij} \neq 0$, dann ist u_{ij} Eigenvektor von α (zum Eigenwert λ_i) und auch Eigenwert von β (zum Eigenwert μ_j). Die u_{ij}'s, welche $\neq 0$ sind, bilden also ein Erzeugendensystem aus gemeinsamen Eigenvektoren, aus welchem man nach 4.19 eine Basis auswählen kann.

10 Orthogonale Räume

10.1 Bemerkung/Definition: *komplexe Zahlen*

Auf $\mathbb{C} = \mathbb{R}^2 = \{(a,b) \mid a,b \in \mathbb{R}\}$ definiert man Addition komponentenweise, d.h.

$$(a,b) + (a',b') = (a+a', b+b')$$

und Multiplikation durch

$$(a,b) \cdot (a',b') = (aa' - bb', ab' + ba') \ .$$

Mit diesen beiden Verknüpfungen bildet $\mathbb{C}$ einen Körper, genannt der Körper der komplexen Zahlen.

Die Abbildung $a \mapsto (a,0)$ von $\mathbb{R} \to \mathbb{C}$ ist injektiv und ein Körperhomomorphismus, d.h. sie verträgt sich mit Addition und Multiplikation. Man kann also $\mathbb{R}$ als Unterkörper von $\mathbb{C}$ betrachten.

Setzt man $i = (0,1)$, dann gilt $i^2 = (-1,0) = -1$. Außerdem bilden $i = (0,1)$ und $1 = (1,0)$ eine $\mathbb{R}$–Basis von $\mathbb{C}$. Jede komplexe Zahl z hat also eine eindeutige Darstellung $z = a + bi$ mit $a,b \in \mathbb{R}$.

Die Abbildung $z = a + bi \mapsto \bar{z} = a - bi$ ist ein Körperautomorphismus von $\mathbb{C}$. Man nennt $\bar{z}$ das Konjugierte zu z, die Abbildung Konjugation. Es gilt $\bar{\bar{z}} = z$ und $\bar{z} = z \Leftrightarrow z \in \mathbb{R}$ ($\mathbb{R}$ ist der „Fixkörper“ der Konjugation). Außerdem ist $\bar{z}z \in \mathbb{R}_{\geq 0}$ und $\bar{z}z = 0 \Leftrightarrow z = 0$. Oft nützlich ist $z^{-1} = \frac{\bar{z}}{z\bar{z}}$ (für $z \neq 0$), z.B. $(1+2i)^{-1} = \frac{1}{5} - \frac{2}{5}i$.

10.2 Satz:

$\mathbb{C}$ ist algebraisch abgeschlossen, d.h. jedes nicht–konstante Polynom in $\mathbb{C}[x]$ hat eine Nullstelle in $\mathbb{C}$.

(ohne Beweis)

10.3 Definition: *Bilinearform, symmetrische Bilinearform*

Sei V ein $\mathbb{K}$–VR. Eine Abbildung $(\ ,\ ) : V \times V \to \mathbb{K}$ heißt eine Bilinearform von V, wenn folgendes gilt:

$\forall u, u_1, u_2, v, v_1, v_2 \in V \ \forall k \in \mathbb{K}$

(1) $(u_1 + u_2, v) = (u_1, v) + (u_2, v)$

(2) $(u, v_1 + v_2) = (u, v_1) + (u, v_2)$

(3) $(ku, v) = k(u, v) = (u, kv)$

Man nennt die Bilinearform symmetrisch, falls zusätzlich gilt:

(4) $(u,v) = (v,u) \ \forall u, v \in V$.

10.4 Beispiel:

Sei $V = \mathbb{K}^n$. Für $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ aus V wird durch $(x, y) = \sum_{i=1}^{n} x_i y_i$ eine symmetrische Bilinearform auf V definiert. Betrachtet man x und y als Spaltenvektoren, dann ist $(x, y) = x^t y$.

10.5 Bemerkung/Definition:

(i) Wenn $(\ ,\ )$ eine Bilinearform auf V, dann ist $(0, v) = (u, 0) = 0$ für alle $u, v \in V$.

(ii) Wenn eine Bilinearform auf V gegeben ist, dann ist durch Einschränken auch auf jedem Unterraum $U \leq V$ eine Bilinearform gegeben.

(iii) Wenn $(\ ,\ )$ eine Bilinearform auf V und $B = \{b_1, \ldots, b_n\}$ eine Basis, dann nennt man

$$G = \begin{pmatrix} (b_1, b_1) & \cdots & (b_1, b_n) \\ \vdots & & \vdots \\ (b_n, b_1) & \cdots & (b_n, b_n) \end{pmatrix},$$

die Gram'sche Matrix der Bilinearform (bzgl. der Basis B).

(iv) Wie bei Abbildungen genügt die Kenntnis von G, um den Wert (u, v) für beliebige Vektoren zu berechnen: Wenn

$$u = \sum_{i=1}^{n} x_i b_i \quad \text{und} \quad v = \sum_{i=1}^{n} y_i b_i \ ,$$

dann ist

$$(u, v) = x^t G y \ .$$

(v) Offenbar ist $(\ ,\ )$ symmetrisch genau dann, wenn G symmetrisch.

(vi) Geht man zu einer anderen Basis $B' = \{b'_1, \ldots, b'_n\}$ über, dann ist die Gram'sche Matrix G' von $(\ ,\ )$ bzgl. B' gegeben durch

$$G' = T^t G T \ ,$$

wobei T die Matrix des Basiswechsels von B nach B' ist (vergl. 7.41).

(vii) Wenn α ein Endomorphismus von V ist, dann kann man eine neue Bilinearform $<,>$ durch $< u, v >= (u, \alpha(v))$ definieren. Die Gram'sche Matrix dieser Bilinearform ist GA, wenn A die Matrix von α ist (alles bzgl. B).
Ebenso kann man durch $<< u, v >>= (\alpha(u), v)$ eine dritte Bilinearform definieren, deren Gram'sche Matrix dann $A^t G$ ist.

10.6 Definition: *ausgeartet,* ${}^{\perp}U$, $U^{\perp}$

Sei $(\ ,\ )$ eine Bilinearform auf V.

(1) V heißt nicht ausgeartet, falls $(u, v) = 0\ \forall v \Rightarrow u = 0$.

(2) Wenn U ein Unterraum von V ist, dann setzt man ${}^{\perp}U = \{v \in V \mid (v, U) = 0\}$ und $U^{\perp} = \{v \in V \mid (U, v) = 0\}$.

10.7 Bemerkung:

(i) ${}^{\perp}U$ und $U^{\perp}$ sind Unterräume von V.

(ii) V ist genau dann nicht ausgeartet, wenn ${}^{\perp}V = \{0\}$.

10.8 Satz:

Sei V ein nicht ausgearteter Vektorraum der Dimension n, und sei $U \leq V$.

(1) $V^{\perp} = \{0\}$ (also ist „nicht ausgeartet“ links–rechts–symmetrisch)

(2) $\dim U^{\perp} = \dim {}^{\perp}U = n - \dim U$

(3) ${}^{\perp}(U^{\perp}) = ({}^{\perp}U)^{\perp} = U$

(4) Äquivalent sind:

(i) U ist nicht ausgeartet.

(ii) ${}^{\perp}U \cap U = \{0\}$

(iii) ${}^{\perp}U \oplus U = V$

(iv) ${}^{\perp}U + U = V$

(v) ${}^{\perp}U$ ist nicht ausgeartet.

(ii')–(v') wie (ii)–(v) mit „$U^{\perp}$“ statt „${}^{\perp}U$“

Beweis:

(1) Für $u \in V$ sei $(u,-) : V \to \mathbb{K}$ die Abbildung $(u,-) : v \mapsto (u,v)$. Dann ist $(u,-)$ eine lineare Abbildung, also $(u,-) \in V^*$. Sei $\lambda : V \to V^*$ definiert durch $\lambda(u) = (u,-)$. Dann ist λ ebenfalls linear, und wenn $u \in \mathrm{Ker}\,(\lambda)$, dann ist $(u,-)$ die Nullabbildung, d.h. $(u,v) = 0$ für alle $v \in V$, also $u = 0$, weil V nicht ausgeartet ist. Also ist λ injektiv. Nach 6.3 ist $\dim V^* = \dim V$, also ist nach 5.16 λ auch surjektiv. Sei $0 \neq v \in V$. Dann existiert eine Basis $v = b_1, b_2, \ldots, b_n$ von V. Sei $b_1^*, \ldots, b_n^*$ die duale Basis von V^*. Dann ist $b_1^*(v) = 1$. Weil λ surjektiv, gibt es ein $u \in V$ mit $\lambda(u) = b_1^*$, d.h. $0 \neq 1 = b_1^*(v) = \lambda(u)(v) = (u,v)$. Also ist $v \notin V^{\perp}$. Daher ist $V^{\perp} = \{0\}$.

(2) Sei $m = \dim U$, $\{b_1, \ldots, b_m\}$ eine Basis von U. Ergänze diese zu einer Basis $\{b_1, \ldots, b_m, b_{m+1}, \ldots, b_n\}$ von V. Betrachte die Abbildung $\alpha : v \mapsto ((v,b_1), \ldots, (v,b_n)) \in \mathbb{K}^n$. α ist eine lineare Abbildung $V \to \mathbb{K}^n$. Wenn $v \in \mathrm{Ker}\,(\alpha)$, dann ist $(v,b_i) = 0$ für alle $i = 1, \ldots, n$. Wenn $w \in V$ beliebig, $w = \sum_{i=1}^n k_i b_i$, dann ist $(v,w) = \sum_{i=1}^n k_i (v,b_i) = 0$; also ist $v \in {}^{\perp}V = \{0\}$ (da V nicht ausgeartet), d.h. α ist injektiv. Nach 5.16 ist α auch surjektiv. Daher ist auch $\beta : V \to \mathbb{K}^m$, definiert durch $\beta(v) = ((v,b_1), \ldots, (v,b_m))$ surjektiv. Nach 5.13 ist $n = \dim V = \dim \mathrm{Ker}\,(\beta) + \dim \mathrm{Im}\,(\beta) = \dim \mathrm{Ker}\,(\beta) + m$, d.h. $\dim \mathrm{Ker}\,(\beta) = n - m = n - \dim U$. Aber $v \in \mathrm{Ker}\,(\beta) \Leftrightarrow (v,b_i) = 0,\ i = 1, \ldots, m \Leftrightarrow (v,U) = 0 \Leftrightarrow v \in {}^{\perp}U$, d.h. ${}^{\perp}U = \mathrm{Ker}\,(\beta)$. Wir haben gezeigt: $\dim {}^{\perp}U = n - \dim U$. Ebenso zeigt man $\dim U^{\perp} = n - \dim U$, da nach (1) auch $V^{\perp} = \{0\}$.

(3) Sei $u \in U$, $w \in U^{\perp}$. Dann ist $(u,w) = 0$, also $(u, U^{\perp}) = 0$ und daher $u \in {}^{\perp}(U^{\perp})$. Also ist $U \leq {}^{\perp}(U^{\perp})$. Da nach (2) $\dim {}^{\perp}(U^{\perp}) = n - \dim U^{\perp} = n - (n - \dim U) = \dim U$, folgt $U = {}^{\perp}(U^{\perp})$. Ebenso $U = ({}^{\perp}U)^{\perp}$.

(4) Nach 5.24 ist $\dim(U + {}^\perp U) = \dim U + \dim {}^\perp U - \dim(U \cap {}^\perp U)$, also nach (2) $\dim(U + {}^\perp U) = \dim V - \dim(U \cap {}^\perp U)$. Also

$$\begin{aligned} U + {}^\perp U = V &\Leftrightarrow \dim(U + {}^\perp U) = \dim V \text{ (nach 4.20)} \\ &\Leftrightarrow \dim(U \cap {}^\perp U) = 0 \\ &\Leftrightarrow U \cap {}^\perp U = \{0\} \\ &\Leftrightarrow (u \in U,\, (u, U) = 0 \Rightarrow u = 0) \\ &\Leftrightarrow U \text{ ist nicht ausgeartet .} \end{aligned}$$

Das zeigt die Äquivalenz von (i)–(iv). Wegen der Symmetrie von „nicht ausgeartet“ (siehe (1)), ist (i) äquivalent zu (ii')–(iv'), insbesondere gilt: U ist nicht ausgeartet $\Leftrightarrow U \cap U^\perp = \{0\}$. Verwendet man die letzte Äquivalenz für ${}^\perp U$, so folgt:

$$\begin{aligned} {}^\perp U \text{ ist nicht ausgeartet } &\Leftrightarrow {}^\perp U \cap ({}^\perp U)^\perp = \{0\} \\ &\Leftrightarrow {}^\perp U \cap U = \{0\} \text{ (nach (3)).} \end{aligned}$$

Also ist (v) äquivalent zu (ii). Analog ist (v') äquivalent zu (ii').

10.9 Bemerkung:

(i) Es ist $(\ ,\ )$ nicht ausgeartet genau dann, wenn die Gram'sche Matrix G von $(\ ,\ )$ regulär ist, denn $x^t G y = 0$ für alle x genau dann, wenn $Gy = 0$.

(ii) Wenn dies der Fall ist, dann gibt es zu gegebener Matrix A genau eine Matrix B mit $B^t G = GA$, wie man durch Rechtsmultiplikation mit G^{-1} sieht. Daher gibt es dann zu gegebenem Endomorphismus α von V genau einen Endomorphismus β mit $(\beta(u), v) = (u, \alpha(v))$ für alle u, v.

(iii) Wenn $(\ ,\ )$ ausgeartet ist, stimmt (ii) nicht:
Sei $V = \mathbb{R}^2$; für $x = (x_1, x_2),\ y = (y_1, y_2) \in V$ definiere $(x, y) = x_1 y_1 \in \mathbb{R}$ und $\alpha(x) = (x_2, 0)$. Seien $e_1 = (1, 0),\ e_2 = (0, 1)$ die Vektoren der Standardbasis. Dann ist $(x, e_2) = 0$ für jeden Vektor x und $\alpha(e_2) = e_1$. Es folgt

$$(e_1, \alpha(e_2)) = (e_1, e_1) = 1 \neq 0 = (\beta(e_1), e_2) \ ,$$

wie man auch β wählt.

10.10 Definition: *orthogonal, orthonormal*

Sei $(\ ,\ )$ eine symmetrische Bilinearform auf V. Zwei Vektoren $u, v \in V$ heißen orthogonal (senkrecht), falls $(u, v) = 0$.

Eine Basis $\{b_i \mid i \in I\}$, für welche $(b_i, b_j) = 0$ für $i \neq j$, heißt Orthogonalbasis. Falls zusätzlich $(b_i, b_i) = 1$ für alle i, so spricht man von einer Orthonormalbasis.

10.11 Beispiel:

In 10.4 bildet die Standardbasis eine Orthonormalbasis.

10.12 Bemerkung:

Die Gram'sche Matrix G von $(\ ,\ )$ bzgl. einer Basis B ist genau dann eine Diagonalmatrix, wenn B eine Orthogonalbasis ist, und G ist die Einheitsmatrix genau dann, wenn B eine Orthonormalbasis ist.

10.13 Lemma:

Sei $1+1 \neq 0$ in $\mathbb{K}$ (man sagt dann $\operatorname{char}\mathbb{K} \neq 2$) und V ein $\mathbb{K}$–VR mit einer nicht–trivialen symmetrischen Bilinearform. Dann existiert ein $v \in V$ mit $(v,v) \neq 0$.

Beweis: Es gibt u und w mit $(u,w) \neq 0$. Wenn $(u,u) \neq 0$ oder $(w,w) \neq 0$, sind wir fertig. Sonst setze $v = u + w$. Dann ist $(v,v) = (u,u) + (u,w) + (w,u) + (w,w) = 2(u,w) \neq 0$.

10.14 Satz:

Sei $\operatorname{char}\mathbb{K} \neq 2$ und V ein endlich dimensionaler $\mathbb{K}$–VR mit einer symmetrischen Bilinearform. Dann hat V eine Orthogonalbasis.

Beweis: Wenn die Bilinearform trivial ist, dann ist jede Basis eine Orthogonalbasis. Andernfalls existiert nach 10.13 ein $v \in V$ mit $(v,v) \neq 0$. Dann ist $U = \mathbb{K}v$ nicht ausgeartet, daher $U \cap U^\perp = \{0\}$. Also ist $\dim U^\perp < \dim V$. Per Induktion über die Dimension hat $U^\perp$ eine Orthogonalbasis $\{b_1, \ldots, b_r\}$. Dann ist $B = \{v = b_0, b_1, \ldots, b_r\}$ eine Orthogonalbasis von V: Offenkundig sind verschiedene Vektoren aus B zueinander orthogonal. Wenn $0 = \sum_{i=0}^r k_i b_i$, dann ist $k_0 b_0 = -\sum_{i=1}^r k_i b_i \in U \cap U^\perp = \{0\}$. Daher ist $k_0 = 0 = k_1 = \cdots = k_r$, weil $\{b_1, \ldots, b_r\}$ linear unabhängig ist. Also ist B linear unabhängig. Wenn $w \in V$ beliebig, setze $k = \frac{(b_0,w)}{(b_0,b_0)}$. Dann ist $w - kb_0 \in U^\perp$, also eine Linearkombination von $\{b_1, \ldots, b_r\}$. Daher ist w eine Linearkombination von B.

10.15 Definition: *Skalarprodukt, euklidischer Raum, Vektorlänge*

Sei V ein reeller VR. Eine symmetrische Bilinearform $(\ ,\ )$ auf V heißt Skalarprodukt, wenn $(v,v) > 0$ für alle $0 \neq v \in V$ („positiv definit"). Ein reeller VR mit einem Skalarprodukt heißt ein euklidischer Raum. Man definiert dann $|v| = \sqrt{(v,v)}$ als die Länge von v.

10.16 Beispiel:

10.4 mit $\mathbb{K} = \mathbb{R}$

10.17 Korollar:

Sei V ein endlich dimensionaler euklidischer Raum. Dann hat V eine Orthonormalbasis.

Beweis: Sei $\{b_1, \ldots, b_n\}$ eine Orthogonalbasis. Setze $e_i = \frac{1}{|b_i|} b_i$. Dann ist

$$(e_i, e_j) = \frac{1}{|b_i|}\frac{1}{|b_j|}(b_i, b_j) = \begin{cases} \frac{1}{|b_i|^2}(b_i,b_i) = 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

10.18 Bemerkung:

10.17 gilt auch noch für abzählbar–unendlich dimensionale euklidische Räume.

10.19 Satz:

Sei V ein euklidischer Raum und $v, w \in V$. Dann gelten:

(1) $(v,w)^2 \leq (v,v)(w,w)$ (Schwarz'sche Ungleichung),
wobei „=“ genau dann, wenn v und w linear abhängig

(2) $|rv| = |r|\,|v|$ für $r \in \mathbb{R}$

(3) $|v+w| \leq |v| + |w|$ (Dreiecks–Ungleichung),
wobei „=“ genau dann, wenn $v = 0$ oder $w = rv$ mit $r \geq 0$ ist.

Beweis:

(1) Wenn $v = 0$ oder $w = 0$, dann ist die Aussage trivial. Sei also $v, w \neq 0$. Für jedes $r \in \mathbb{R}$ gilt

$$0 \leq (v - rw, v - rw) = (v, v) + r^2(w, w) - 2r(v, w) \ .$$

Insbesondere für $r = \frac{(v,w)}{(w,w)}$:

$$0 \leq (v, v) + \frac{(v, w)^2}{(w, w)^2}(w, w) - 2\frac{(v, w)}{(w, w)}(v, w) = (v, v) - \frac{(v, w)^2}{(w, w)} \ .$$

Also ist $(v, w)^2 \leq (v, v)(w, w)$. Es gilt sogar „<“, außer wenn $v - rw = 0$, d.h. wenn $v = rw$, v und w linear abhängig. Umgekehrt: Falls $v = rw$, dann ist $(v, w)^2 = (rw, w)^2 = r^2(w, w)^2 = (rw, rw)(w, w) = (v, v)(w, w)$.

(2) ist trivial

(3) Nach (1) ist $(v, w) \leq |(v, w)| = \sqrt{(v, w)^2} \leq \sqrt{(v, v)(w, w)} = |v|\,|w|$. (Dabei gilt Gleichheit an der zweiten Stelle nur für v und w linear abhängig. Wenn dies der Fall, aber $v \neq 0$, dann ist $w = rv$. Aus Gleichheit an der ersten Stelle folgt dann $r \geq 0$.) Daher

$$\begin{aligned} |v + w|^2 = (v + w, v + w) &= (v, v) + (w, w) + 2(v, w) \\ &\leq |v|^2 + |w|^2 + 2|v|\,|w| = (|v| + |w|)^2 \,, \end{aligned}$$

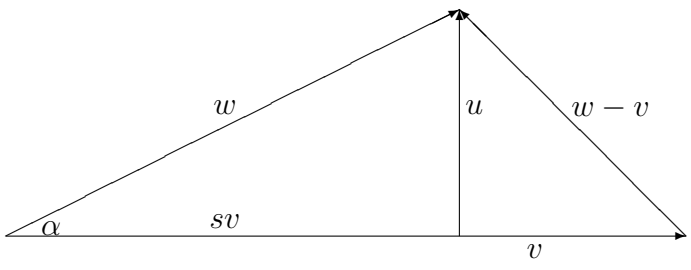
also $|v + w| \leq |v| + |w|$ mit Gleichheit nur im angegebenen Fall.

10.20 Bemerkung/Definition: *Kosinus*

Nach der Schwarz'schen Ungleichung gilt für Vektoren $v, w \neq 0$ aus einem euklidischen Raum, dass $-1 \leq r = \frac{(v,w)}{|v||w|} \leq 1$ ist. Daher gibt es einen Winkel $\alpha = \alpha(v, w)$ mit $\cos\alpha = r$, und durch die zusätzliche Annahme $0 \leq \alpha \leq \pi$ ist α eindeutig bestimmt. Man nennt α den von v und w eingeschlossenen Winkel. Insbesondere ist also $\alpha = \pi/2$ genau dann, wenn $\cos\alpha = 0$, d.h. $(v, w) = 0$, also v orthogonal zu w.

10.21 Bemerkung: *Kosinussatz*

(1) Seien v, w, α wie oben. Setzt man $s = \frac{|w|}{|v|} cos\ \alpha$ und $u = w - sv$, dann ist $w = sv + u$ und $(u, v) = (w, v) - s(v, v) = (w, v) - |v||w| cos\ \alpha = 0$, also u orthogonal zu v. Außerdem ist $\frac{s|v|}{|w|} = cos\ \alpha$ das Verhältnis von Ankathete zu Hypothenuse.



(2) Es ist $|w-v|^2 = |v|^2 + |w|^2 - 2(v,w) = |v|^2 + |w|^2 - 2|v||w|cos\ \alpha$. Das ist der Kosinussatz. Damit läßt sich aus zwei Seiten und dem eingeschlossenen Winkel die Länge der dritten Dreiecks-Seite berechnen. Für den Fall, dass v und w orthogonal sind, geht der Kosinussatz in den Satz des Pythagoras über.

10.22 Lemma:

A sei eine symmetrische, reelle $n \times n$–Matrix. Dann sind alle Eigenwerte von A reell.

Beweis: Nach 10.2 hat A jedenfalls einen Eigenwert $\lambda \in \mathbb{C}$, zu zeigen: $\lambda \in \mathbb{R}$, d.h. $\lambda = \bar{\lambda}$. Sei $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ ein Eigenvektor zu λ, d.h. $\lambda z = Az$. Durch Konjugieren und Transponieren folgt $\bar{z}^t\bar{\lambda} = \bar{z}^t\bar{A}^t = \bar{z}^tA$, weil A reell und symmetrisch. Also ist $\bar{z}^t\bar{\lambda}z = \bar{z}^tAz = \bar{z}^t\lambda z$. Da $\bar{z}^tz \neq 0$, weil $z \neq 0$, folgt $\lambda = \bar{\lambda}$.

10.23 Satz:

Sei V ein endlich dimensionaler euklidischer Raum mit Skalarprodukt $(\ ,\ )$, und sei $\langle\ ,\ \rangle$ eine zweite symmetrische Bilinearform auf V. Dann existiert eine Orthonormalbasis bzgl. $(\ ,\ )$, welche zugleich Orthogonalbasis bzgl. $\langle\ ,\ \rangle$ ist.

Beweis: Sei $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis bzgl. $(\ ,\ )$ (existiert nach 10.17), und sei $a_{ij} = \langle e_i, e_j\rangle$. Dann ist $a_{ij} = a_{ji}$ (da $\langle\ ,\ \rangle$ symmetrisch). Also ist $A = (a_{ij})$ eine symmetrische Matrix. Sei λ ein (reeller!) Eigenwert von A und $\begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$ ein Eigenvektor in $\mathbb{R}^n$, also $\sum_{j=1}^n a_{ij}k_j = \lambda k_i$. Setze $b = \sum_{j=1}^n k_je_j$. Dann ist $b \neq 0$ und

$$\lambda(e_i, b) = \lambda k_i = \sum_{j=1}^n a_{ij}k_j = \sum_{j=1}^n \langle e_i, e_j\rangle k_j = \langle e_i, \sum_{j=1}^n k_je_j\rangle = \langle e_i, b\rangle.$$

Daher gilt für $v = \sum_{i=1}^n r_ie_i \in V$

$$\lambda(v, b) = \lambda \sum_{i=1}^n r_i(e_i, b) = \sum_{i=1}^n r_i\langle e_i, b\rangle = \langle v, b\rangle.$$

Insbesondere: Wenn $(v, b) = 0$, dann auch $\langle v, b\rangle = 0$. Sei $U = b^\perp$ bzgl. $(\ ,\ )$. Setze $b_1 = \frac{b}{|b|}$ und sei $\{b_2, \ldots, b_n\}$ eine Orthonormalbasis von U bzgl. $(\ ,\ )$, welche zugleich Orthogonalbasis bzgl. $\langle\ ,\ \rangle$ ist (existiert per Induktion über die Dimension). Dann ist $\{b_1, \ldots, b_n\}$ wie gewünscht.

10.24 Definition: *orthogonale Abbildung und Matrix*

(1) Sei V ein euklidischer Raum mit Skalarprodukt $(\ ,\ )$ und α ein Endomorphismus von V. Man nennt α eine orthogonale Abbildung, wenn

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

(2) Eine reelle $n \times n$–Matrix A heißt orthogonal, wenn $A^tA = E$.

10.25 Bemerkung:

(i) A orthogonal $\Leftrightarrow$ A^t orthogonal

(ii) Genau dann ist A orthogonal, wenn die Spalten von A eine Orthonormalbasis von $\mathbb{R}^n$ bzgl. des üblichen Skalarprodukts bilden. (Ebenso für die Zeilen von A.)

(iii) Produkte orthogonaler Matrizen sind orthogonal.

10.26 Lemma:

Sei V ein endlich dimensionaler euklidischer Raum, und seien α und β Endomorphismen von V mit Matrizen A, B bzgl. einer Orthonormalbasis $\{e_1, \ldots, e_n\}$. Genau dann ist $(\alpha v, w) = (v, \beta w)$ für alle $v, w \in V$, wenn $B = A^t$.

Beweis: Sei $A = (a_{ij})$, $B = (b_{ij})$. Dann gilt

$$\begin{aligned}(\alpha v, w) = (v, \beta w)\, \forall v, w \in V &\Leftrightarrow (\alpha e_i, e_j) = (e_i, \beta e_j)\ \forall i, j \\ &\quad (\text{„}\Rightarrow\text{“ trivial, „}\Leftarrow\text{“ Bilinearität}) \\ &\Leftrightarrow a_{ji} = \left(\sum_k a_{ki} e_k, e_j\right) = \left(e_i, \sum_k b_{kj} e_k\right) = b_{ij}\ \forall i, j \\ &\Leftrightarrow B = A^t.\end{aligned}$$

10.27 Satz:

Sei V ein endlich dimensionaler euklidischer Raum, und sei α ein Endomorphismus von V. Äquivalent sind:

(1) α ist eine orthogonale Abbildung.

(2) $|\alpha v| = |v|$ für alle $v \in V$ (d.h. α ist 'längentreu')

(3) Wenn $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis von V ist, dann auch $\{\alpha e_1, \ldots, \alpha e_n\}$.

(4) Die Matrix von α bzgl. einer beliebigen Orthonormalbasis ist orthogonal.

(5) Es gibt eine Orthonormalbasis, so daß die Matrix von α bzgl. dieser Basis orthogonal ist.

Beweis:

(1)$\Rightarrow$(2) $(\alpha v, \alpha v) = (v, v) \Rightarrow |\alpha v| = |v|$

(2)$\Rightarrow$(3) α ist injektiv, denn wenn $\alpha v = 0$, dann ist $0 = |\alpha v| = |v|$, d.h. $v = 0$. Also ist α ein Automorphismus von V. Daher ist $\{\alpha e_1, \ldots, \alpha e_n\}$ wieder eine Basis. Bleibt zu zeigen, daß

$$(\alpha e_i, \alpha e_j) = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst.} \end{cases}$$

Wenn $i = j$, dann ist $(\alpha e_i, \alpha e_j) = |\alpha e_i|^2 = |e_i|^2 = (e_i, e_i) = 1$.
Wenn $i \neq j$, dann ist

$$\begin{aligned} 2 &= (e_i, e_i) + (e_j, e_j) + 2(e_i, e_j) = (e_i + e_j, e_i + e_j) \\ &= |e_i + e_j|^2 = |\alpha(e_i + e_j)|^2 \text{ (nach (2))} \\ &= (\alpha e_i + \alpha e_j, \alpha e_i + \alpha e_j) = (\alpha e_i, \alpha e_i) + (\alpha e_j, \alpha e_j) + 2(\alpha e_i, \alpha e_j) \\ &= 2 + 2(\alpha e_i, \alpha e_j), \end{aligned}$$

also $(\alpha e_i, \alpha e_j) = 0$.

(3)⇒(4) Sei $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis, und sei A die Matrix von α bzgl. dieser Basis. Sei β die Abbildung mit der Matrix A^t. Dann gilt nach Lemma 10.26 $(\alpha v, w) = (v, \beta w)$ für alle $v, w \in V$. Insbesondere für $v = e_i$, $w = \alpha e_j$:

$$\begin{aligned} (e_i, \beta\alpha e_j) = (\alpha e_i, \alpha e_j) &= \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} \\ &= (e_i, e_j) \end{aligned}$$

Daher ist $(e_i, \beta\alpha e_j - e_j) = 0$ für alle i, j. Es folgt $(v, \beta\alpha e_j - e_j) = 0$ für alle j und alle $v \in V$, also $e_j = \beta\alpha e_j$ für alle j. Dann ist aber $\beta\alpha = \mathrm{id}$, also $A^t A = E$ und daher A orthogonal.

(4)⇒(5) Klar mit 10.17.

(5)⇒(1) Sei $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis derart, daß die Matrix A von α bzgl. dieser Basis orthogonal ist. Sei β die Abbildung mit Matrix A^t, also $\beta\alpha = \mathrm{id}$. Dann gilt für alle $v, w \in V$: $(v, w) = (v, \beta\alpha w) = (\alpha v, \alpha w)$, nach 10.26. Also ist α orthogonal.

10.28 Korollar:

Sei V ein endlich dimensionaler euklidischer Raum und $\alpha : V \to V$ orthogonal. Dann ist $\det \alpha = \pm 1$, insbesondere ist α ein Automorphismus. Außerdem sind ± 1 die einzig möglichen Eigenwerte von α in $\mathbb{R}$.

Beweis: Es ist $1 = \det E = \det(A^t A) = \det A^t \det A = (\det A)^2$, also $\det A = \pm 1$. Wenn v Eigenvektor zum Eigenwert $\lambda \in \mathbb{R}$ von α, dann ist $(v, v) = (\alpha v, \alpha v) = (\lambda v, \lambda v) = \lambda^2 (v, v)$. Da $(v, v) \neq 0$, folgt $\lambda^2 = 1$, also $\lambda = \pm 1$.

10.29 Korollar:

Sei V ein endlich dimensionaler euklidischer Raum, und sei α ein Endomorphismus von V. Wenn α bzgl. einer Orthonormalbasis eine symmetrische Matrix hat, dann gilt dies bzgl. jeder Orthonormalbasis.

Beweis: Sei $\{e_1, \ldots, e_n\}$ eine Orthonormalbasis bzgl. welcher die Matrix A von α symmetrisch ist. Sei $\{b_1, \ldots, b_n\}$ eine zweite Orthonormalbasis, und sei T die Transformationsmatrix. Dann ist T die Matrix der orthogonalen Abbildung $e_i \mapsto b_i$ bezüglich $\{e_1, \ldots, e_n\}$, also orthogonal nach 10.27, d.h $T^{-1} = T^t$, außerdem ist $T^{-1}AT$ die Matrix von α bzgl. $\{b_1, \ldots, b_n\}$ nach 7.43. Dann ist $(T^{-1}AT)^t = T^t A^t (T^{-1})^t = T^{-1} A^t (T^t)^t = T^{-1}AT$, also $T^{-1}AT$ symmetrisch.

10.30 Satz: *Hauptachsen-Theorem*

Sei A eine symmetrische reelle $n \times n$–Matrix. Dann existiert eine orthogonale Matrix T derart, daß T^tAT eine Diagonalmatrix ist.

Beweis: Definiere auf $\mathbb{R}^n$ das übliche Skalarprodukt $(v, w) = v^t w$ und eine weitere Bilinearform $\langle\ ,\ \rangle$ durch $\langle v, w\rangle = v^t Aw$ ($\langle\ ,\ \rangle$ ist symmetrisch, da A symmetrisch). Nach 10.23 existiert eine Orthonormalbasis $\{b_1, \ldots, b_n\}$ bzgl. $(\ ,\ )$, welche zugleich eine Orthogonalbasis bzgl. $\langle\ ,\ \rangle$ ist. Sei T die Transformationsmatrix, d.h. $Te_i = b_i$. Nach 10.27 ist T orthogonal. Außerdem gilt für $i \neq j$: $0 = \langle b_i, b_j\rangle = b_i^t Ab_j$. Aber $b_i = Te_i$, also $b_i^t = e_i^t T^t$, und daher $0 = e_i^t(T^tAT)e_j$. Für eine beliebige Matrix $S = (s_{ij})$ gilt jedoch $e_i^t Se_j = s_{ij}$. Folglich ist für $i \neq j$ der Eintrag in der i–ten Zeile und j–ten Spalte von T^tAT gleich 0, d.h. T^tAT ist eine Diagonalmatrix.

10.31 Korollar:

Sei A eine reelle symmetrische $n \times n$–Matrix. Seien $\lambda_1, \ldots, \lambda_s$ die verschiedenen Nullstellen des charakteristischen Polynoms von A in $\mathbb{C}$. Dann sind alle λ_i reell. Wenn U_i der Eigenraum zum Eigenwert λ_i, dann ist $\mathbb{R}^n = U_1 \oplus \cdots \oplus U_s$, wobei U_i und U_j für $i \neq j$ orthogonal sind.

Beweis: Daß die Eigenwerte reell sind, steht schon in 10.22. Da A diagonalisierbar ist (nach 10.30), ist $\mathbb{R}^n$ die direkte Summe der Eigenräume (nach 9.14). Wenn $v \in U_i$, $w \in U_j$ für $i \neq j$, dann gilt

$$\begin{aligned}\lambda_i(v, w) = (\lambda_i v, w) &= (Av, w)\\ &= (v, Aw) \ \text{ (nach 10.26)}\\ &= (v, \lambda_j w) = \lambda_j(v, w),\end{aligned}$$

also $(\lambda_i - \lambda_j)(v, w) = 0$. Da $\lambda_i \neq \lambda_j$, folgt $(v, w) = 0$. Also sind U_i und U_j orthogonal.

10.32 Bemerkung:

Beschreibung der orthogonalen Abbildungen α von $\mathbb{R}^n$ (mit dem üblichen Skalarprodukt) für $n \leq 3$.

(i) $n = 1$
Nach 10.27 ist $|\alpha(e_1)| = |e_1| = 1$, also $\alpha(e_1) = \pm e_1$. Es gibt also nur zwei orthogonale Abbildungen, nämlich $\alpha = \mathrm{id}$ und $\alpha = -\mathrm{id}$.

(ii) $n = 2$
Wegen $|\alpha(e_1)| = 1$ ist $\alpha(e_1)$ wieder auf dem Einheitskreis. Sei φ der Winkel zwischen e_1 und $\alpha(e_1)$ (im positiven Umlaufsinn), also

$$\alpha(e_1) = \begin{pmatrix}\cos\varphi\\ \sin\varphi\end{pmatrix}.$$

Da $\alpha(e_2)$ senkrecht zu $\alpha(e_1)$ und $|\alpha(e_2)| = 1$, ist

$$\alpha(e_2) = \begin{pmatrix}-\sin\varphi\\ \cos\varphi\end{pmatrix} \ \text{ oder } \ \alpha(e_2) = -\begin{pmatrix}-\sin\varphi\\ \cos\varphi\end{pmatrix}.$$

Die Matrix von α bzgl. $\{e_1, e_2\}$ ist also

$$D_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \text{ oder } S_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}.$$

Dabei hat D_φ die Determinante 1 und ist gerade eine Drehung um den Winkel φ. Es ist $D_0 = E$ und $D_\pi = -E$; wenn $\varphi \neq 0, \pi$, dann hat D_φ keine Eigenwerte in $\mathbb{R}$. Aus $D_\varphi D_\psi = D_{\varphi+\psi}$ erhält man die Additionstheoreme für sin und cos.
Es ist $\det S_\varphi = -1$ und $\operatorname{char\,pol} S_\varphi = (x - \cos\varphi)(x + \cos\varphi) - \sin^2\varphi = x^2 - 1 = (x+1)(x-1)$. Also hat S_φ die Eigenwerte ± 1. Aus den Additionstheoremen für $\frac{\varphi}{2} = \varphi - \frac{\varphi}{2}$ erhält man

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \begin{pmatrix} \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} \end{pmatrix},$$

also ist

$$s = \begin{pmatrix} \cos\frac{\varphi}{2} \\ \sin\frac{\varphi}{2} \end{pmatrix}$$

ein Eigenvektor zum Eigenwert 1. Ein zu s senkrechter Vektor t ist dann Eigenvektor zum Eigenwert -1. Also ist S_φ die Spiegelung an der Geraden $\mathbb{R}s$.

(iii) $n = 3$

(A) 1 ist Eigenwert.
Sei U der Eigenraum zum Eigenwert 1.

(Aa) $\dim U = 3$. Dann ist $\alpha = \mathrm{id}$.

(Ab) $\dim U = 2$. Dann ist α die Spiegelung an der Ebene U (und -1 ist ebenfalls Eigenwert von α).

(Ac) $\dim U = 1$. Dann ist α die Drehung um die Achse U mit einem geeigneten Winkel $\varphi \neq 0$.

(B) 1 ist kein Eigenwert.
Da das charakteristische Polynom den Grad 3 hat, gibt es eine reelle Nullstelle. Diese ist dann -1 nach 10.28. Dann ist 1 ein Eigenwert von $-\alpha$. Also ist $-\alpha$ in (A) beschrieben.
Der Fall (Ab) tritt für $-\alpha$ nicht ein, da sonst -1 Eigenwert von $-\alpha$ ist, also 1 Eigenwert von α entgegen der Annahme.

10.33 Schmidt'sches Orthonormalisierungsverfahren:

Sei $\{b_1, b_2, \ldots\}$ eine höchstens abzählbare, linear unabhängige Menge im euklidischen Vektorraum V. Dann gibt es eine orthonormale Menge $\{e_1, e_2, \ldots\}$ mit $\langle b_1, \ldots, b_n\rangle = \langle e_1, \ldots, e_n\rangle$ für alle $n \leq |\{b_1, b_2, \ldots\}|$.

Beweis: Die e_i werden induktiv definiert. Sei $e_1 = \frac{b_1}{|b_1|}$. Dann ist $|e_1| = 1$ und $\langle e_1\rangle = \langle b_1\rangle$. Seien $e_1, \ldots, e_n$ schon definiert mit $|e_i| = 1$, $(e_i, e_j) = 0$ für $i \neq j$, und $\langle e_1, \ldots, e_n\rangle = \langle b_1, \ldots, b_n\rangle$. Sei

$$a_{n+1} = b_{n+1} - \sum_{i=1}^{n} (b_{n+1}, e_i) e_i.$$

Dann gilt $\langle a_{n+1}, e_1, \ldots, e_n\rangle = \langle b_{n+1}, e_1, \ldots, e_n\rangle = \langle b_{n+1}, b_1, \ldots, b_n\rangle \gneq \langle b_1, \ldots, b_n\rangle = \langle e_1, \ldots, e_n\rangle$, da $b_{n+1} \notin \langle b_1, \ldots, b_n\rangle$ wegen der linearen Unabhängigkeit der b_i. Insbesondere ist $a_{n+1} \neq 0$. Aber für $i \leq n$ gilt $(a_{n+1}, e_i) = (b_{n+1}, e_i) - (b_{n+1}, e_i) = 0$. Setze $e_{n+1} = \frac{a_{n+1}}{|a_{n+1}|}$. Dann ist $\{e_1, \ldots, e_{n+1}\}$ eine orthonormale Menge und $\langle e_1, \ldots, e_{n+1}\rangle = \langle e_1, \ldots, e_n, a_{n+1}\rangle = \langle b_1, \ldots, b_{n+1}\rangle$.

10.34 Algorithmus: *zur Diagonalisierung symmetrischer reeller Matrizen mit orthogonalen Transformationsmatrizen*

Gegeben: symmetrische reelle $n \times n$–Matrix A

Gesucht: orthogonale Matrix T mit $T^t A T$ diagonal (Existenz klar nach Hauptachsen Theorem)

Schritt 1: Berechne das charakteristische Polynom $p(x)$ zu A.

Schritt 2: Finde die Nullstellen von $p(x)$, d.h. die Eigenwerte von A.

Schritt 3: Zu jedem Eigenwert λ berechne eine Basis des zugehörigen Eigenraumes (durch Lösen des homogenen linearen Gleichungssystems $(A - \lambda E)x = 0$).

Schritt 4: Wende für jedes λ das Schmidt'sche Orthonormalisierungsverfahren auf die in Schritt 3 gefundenen Basisvektoren an.

Schritt 5: Die so gefundene Orthonormalbasis bildet die Spalten von T.

10.35 Beispiel:

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

Dann ist

$$\begin{aligned} \text{char pol}A &= \det \begin{pmatrix} x-2 & -1 & -1 \\ -1 & x-2 & 1 \\ -1 & 1 & x-2 \end{pmatrix} \\ &= (x-2)^3 + 1 + 1 - 3(x-2) \\ &= x^3 - 6x^2 + 9x \\ &= x(x-3)^2 \end{aligned}$$

A hat also die Eigenwerte 0 und 3. Wir bestimmen die dazugehörigen Eigenräume U_0 und U_3.

Bestimmung von U_0:

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

Also ist $b_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$ eine Basis von U_0.

Bestimmung von U_3:

$$\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Also bilden $b_2 = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$ und $b_3 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$ eine Basis von U_3.

Anwendung des Schmidt'schen Orthonormalisierungsverfahren auf $\{b_1\}$ und $\{b_2, b_3\}$ liefert

$$e_1 = \frac{b_1}{|b_1|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

und

$$e_2 = \frac{b_2}{|b_2|} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \ ,$$

$$a_3 = b_3 - (b_3, e_2)e_2 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \ ,$$

$$e_3 = \frac{a_3}{|a_3|} = -\frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} .$$

Also ist

$$T = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ -1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ -1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix}.$$

Zur Probe kann man

$$T^tT = E \text{ und } T^tAT = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

verifizieren.

10.36 Anwendung:

Gegeben sei

$$p(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 - 2x_2x_3 + 4x_1 + 3x_2 + x_3 + \frac{1}{6}.$$

Wir wollen die Punktmenge $\left\{ x = \left(\begin{smallmatrix} x_1 \\ x_2 \\ x_3 \end{smallmatrix}\right) \in \mathbb{R}^3 \,|\, p(x) = 0 \right\}$ beschreiben.

Mit A und T wie in 10.35, $b = \left(\begin{smallmatrix} 4 \\ 3 \\ 1 \end{smallmatrix}\right)$ und $c = \frac{1}{6}$ kann man $p(x)$ schreiben als $x^tAx + b^tx + c$. Setzt man $y = T^tx = \left(\begin{smallmatrix} y_1 \\ y_2 \\ y_3 \end{smallmatrix}\right)$, dann ist $x = Ty$, also

$$p(x) = y^tT^tATy + (T^tb)^ty + c = q(y).$$

Es ist

$$y^tT^tATy = (y_1, y_2, y_3) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = 3y_2^2 + 3y_3^2$$

und

$$T^tb = \begin{pmatrix} 0 \\ -7/\sqrt{2} \\ -3/\sqrt{6} \end{pmatrix}.$$

Daher ist

$$\begin{aligned} q(y) &= 3y_2^2 + 3y_3^2 - \frac{7}{\sqrt{2}}y_2 - \frac{3}{\sqrt{6}}y_3 + \frac{1}{6} \\ &= 3\left(y_2 - \frac{7}{6\sqrt{2}}\right)^2 + 3\left(y_3 - \frac{1}{2\sqrt{6}}\right)^2 + \frac{1}{6} - 3\frac{7^2}{6^2 \cdot 2} - 3\frac{1^2}{2^2 \cdot 6} \\ &= 3\left(y_2 - \frac{7}{6\sqrt{2}}\right)^2 + 3\left(y_3 - \frac{1}{2\sqrt{6}}\right)^2 - 2, \end{aligned}$$

also $0 = p(x) = q(y)$ genau dann, wenn

$$\left(y_2 - \frac{7}{6\sqrt{2}}\right)^2 + \left(y_3 - \frac{1}{2\sqrt{6}}\right)^2 = \frac{2}{3}.$$

Dies ist ein Kreiszylinder, dessen Achse parallel zur y_1–Achse ist und durch den Punkt

$$\begin{pmatrix} 0 \\ 7/(6\sqrt{2}) \\ 1/(2\sqrt{6}) \end{pmatrix}$$

geht. Sein Radius ist $\sqrt{\frac{2}{3}}$. In den ursprünglichen Koordinaten ist die Achse parallel zu

$$\mathbb{R} \cdot T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \mathbb{R} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

und geht durch den Punkt

$$T \begin{pmatrix} 0 \\ 7/(6\sqrt{2}) \\ 1/(2\sqrt{6}) \end{pmatrix} = -\frac{1}{6} \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}.$$

10.37 Bemerkung:

Wie hier im Beispiel läßt sich die Nullstellen-Menge eines beliebigen, höchstens quadratischen Polynoms in $\mathbb{R}^n$ beschreiben. Sei also

$$p(x_1, \cdots, x_n) \;=\; \sum_{i=1}^{n} a_{ii} x_i^2 \;+\; 2 \sum_{i<j} a_{ij} x_i x_j \;+\; 2 \sum_{i=1}^{n} b_i x_i \;+\; c \;=\; x^t A x \;+\; 2 b^t x \;+\; c,$$

wobei $x^t = (x_1, \cdots, x_n)$, $b^t = (b_1, \cdots, b_n)$, c eine Konstante und A die (symmetrische) Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & . & . & . & a_{1,n} \\ a_{1,2} & a_{2,2} & a_{2,3} & . & . & . & a_{2,n} \\ a_{1,3} & a_{2,3} & a_{3,3} & . & . & . & a_{3,n} \\ . & . & . & . & & & . \\ . & . & . & & . & & . \\ . & . & . & & & . & . \\ a_{1,n} & a_{2,n} & a_{3,n} & & & & a_{n,n} \end{pmatrix}$$

ist. Wenn $A = 0$, dann ist p höchstens linear, und die Nullstellen-Menge ergibt sich als Lösungsmenge eines linearen Gleichungssystems. Diesen Fall ist schon ausführlich diskutiert worden und wird hier beiseite gelassen. Wir nehmen also $A \neq 0$ an. Für eine geeignete orthogonale Matrix T und Diagonalmatrix $D = (d_{ij})$ gilt nach 10.30 $A = T^t D T$, also (mit $y = Tx$ und $f = Tb$)

$$p(x) = x^t T^t D T x \;+\; 2 b^t T^t T x \;+\; c = y^t D y \;+\; 2 f^t y \;+\; c \;=\; q(y)$$

(der Übergang von x zu y entspricht einer Drehung des Koordinatensystems; das Polynom q hat keine gemischten Terme $y_i y_j$ mit $i \neq j$ mehr). Wenn $d_{ii} \neq 0$, kann man

$$d_{ii} y_i^2 + 2 f_i y_i = d_{ii} (y_i + \frac{f_i}{d_{ii}})^2 - \frac{f_i^2}{d_{ii}}$$

benutzen, um durch die Substitution $z_i = y_i + \frac{f_i}{d_{ii}}$ ein neues Polynom

$$r(z) = z^t D z + g^t z + k$$

zu gewinnen, welches zu keinem i sowohl einen quadratischen als auch einen linearen Term enthält (dies entspricht einer Verschiebung des Koordinatensystems). Wenn es jetzt überhaupt noch lineare Terme gibt, dann kann man sogar erreichen, dass es nur einen solchen gibt, indem man wieder dreht: wenn o.B.d.A. $d_{ii} \neq 0 \Leftrightarrow i \leq r$ und $0 \neq g = (0, \cdots, 0, g_{r+1}, \cdots, g_n)$ dann ist $g_0 = \frac{g}{|g|}$ ein Vektor der Länge 1, welcher senkrecht zu den Einheitsvektoren $e_1, \cdots, e_r$ steht; daher gibt es eine Orthonormalbasis $\{e_1, \cdots, e_r, g_0, \cdots\}$. Nochmaliger Übergang zu neuen Koordinaten liefert also schließlich ein Polynom

$$s(u) \;=\; \sum_{i=1}^{r} d_{ii} u_i^2 \;+ h u_{r+1} \;+\; l \;,$$

wobei $r < n$, falls $h \neq 0$. Außerdem kann man dann $h < 0$ und $l = 0$ durch die Substitution $u'_{r+1} = \pm u_{r+1} + \frac{l}{h}$ erreichen.
Wenn dagegen kein lineare Term mehr auftritt, aber der konstante Term ungleich 0 ist, kann man durch geeignete Multiplikation erreichen, dass die Konstante gleich -1 ist. Ist der konstante Term dagegen 0, dann läßt sich durch Multiplikation mit ± 1 erreichen, dass die Anzahl der positiven d_{ii} mindestens so groß wie die der negativen ist. Solche Multiplikationen mit einem Skalar $\neq 0$ ändern zwar das Polynom, aber nicht seine Nullstellen-Menge.

10.38 Bemerkung: *Klassifikation der Hyperflächen zweiter Ordnung in* $\mathbb{R}^2$ *und* $\mathbb{R}^3$

Alle auftretenden reellen Zahlen $a, b, \cdots$ seien positiv. Sei zunächst $p(x, y)$ ein quadratisches Polynom in zwei Unbestimmten. Die folgenden Fälle sind möglich; in Klammern steht immer eine geometrische Beschreibung der Nullstellen-Menge:

(A) $rg(D) = 2$

- (A1) $p(x,y) = \;\; ax^2 + by^2 - 1$ (Ellipse)
- (A2) $p(x,y) = \;\; ax^2 - by^2 - 1$ (Hyperbel)
- (A3) $p(x,y) = -ax^2 - by^2 - 1$ (leere Menge)
- (A4) $p(x,y) = \;\; ax^2 + by^2$ (Nullpunkt)
- (A5) $p(x,y) = \;\; ax^2 - by^2$ (zwei Geraden, die sich im Nullpunkt schneiden)

(B) $rg(D) = 1$

- (B1) $p(x,y) = \;\; ax^2 - 1$ (zwei parallele Geraden)
- (B2) $p(x,y) = -ax^2 - 1$ (leere Menge)
- (B3) $p(x,y) = \;\; ax^2$ (y-Achse)
- (B4) $p(x,y) = \;\; ax^2 - by$ (Parabel)

Die analoge Liste für ein quadratisches Polynom $p(x, y, z)$ in drei Unbestimmten:

(A) $rg(D) = 3$

(A1) $p(x,y,z) = \quad ax^2 + by^2 + cy^2 - 1$ (Ellipsoid)
(A2) $p(x,y,z) = \quad ax^2 + by^2 - cz^2 - 1$ (einschaliges Hyperboloid)
(A3) $p(x,y,z) = \quad ax^2 - by^2 - cz^2 - 1$ (zweischaliges Hyperboloid)
(A4) $p(x,y,z) = -ax^2 - by^2 - cy^2 - 1$ (leere Menge)
(A5) $p(x,y,z) = \quad ax^2 + by^2 + cz^2$ (Nullpunkt)
(A6) $p(x,y,z) = \quad ax^2 + by^2 - cz^2$ (elliptischer Kegel)

(B) $rg(D) = 2$

(B1) $p(x,y,z) = \quad ax^2 + by^2 - 1$ (elliptischer Zylinder)

(B2) $p(x,y,z) = \quad ax^2 - by^2 - 1$ (hyperbolischer Zylinder)

(B3) $p(x,y,z) = -ax^2 - by^2 - 1$ (leere Menge)

(B4) $p(x,y,z) = \quad ax^2 + by^2$ (z-Achse)

(B5) $p(x,y,z) = \quad ax^2 - by^2$ (zwei sich in der z-Achse schneidende Ebenen)

(B6) $p(x,y,z) = \quad ax^2 + by^2 - cz$ (elliptisches Paraboloid)

(B7) $p(x,y,z) = \quad ax^2 - by^2 - cz$ (hyperbolisches Paraboloid)

(C) $rg(D) = 1$

(C1) $p(x,y,z) = \quad ax^2 - 1$ (zwei parallele Ebenen)

(C2) $p(x,y,z) = -ax^2 - 1$ (leere Menge)

(C3) $p(x,y,z) = \quad ax^2$ ((y,z)-Ebene)

(C4) $p(x,y,z) = \quad ax^2 - by$ (parabolischer Zylinder)

10.39 Bemerkung: *Kegelschnitte*

Ellipse, Parabel und Hyperbel sind Kegelschnitte, d.h. sie entstehen als Schnitt eines Doppelkegels mit einer Ebene. Das läßt sich wieder leicht mit dem Hauptachsen-Theorem zeigen: Sei $p\,(x,y)$ ein quadratisches Polynom in zwei Unbestimmten und Koeffizienten aus $\mathbb{R}$. Definiert man

$$f(x,y,z) = z^2 p\,(\frac{x}{z}\,,\,\frac{y}{z}) \quad ,$$

dann ist $f(x,y,z)$ ein Polynom in drei Unbestimmten, in dem alle Terme vom Grad 2 sind; außerdem ist $p\,(x,y) = f(x,y,1)$. (Zum Beispiel führt $p\,(x,y) = 3x^2 - 2y^2 - 3x + 4$ zu $f(x,y,z) = 3x^2 - 2y^2 - 3xz + 4z^2$.) Wie oben kann man

$$f(x,y,z) = (x,y,z)A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

mit einer symmetrischen, reellen 3×3-Matrix A schreiben. Nach dem Hauptachsen-Theorem darf man nach Übergang zu neuen Unbestimmten x_1, y_1, z_1 annehmen, dass A eine Diagonalmatrix ist. Wir nehmen jetzt zusätzlich an, dass $\mathrm{Rg}\,(A) = 3$ (Sie sollten sich auch überlegen, was in den anderen Fällen geschieht!); dann sind die drei Einträge auf der Diagonalen alle ungleich 0. Wenn sie das gleiche Vorzeichen haben, dann ist $(x,y,z) = (0,0,0)$ die einzige Nullstelle von f, also hat p keine Nullstelle. Im anderen Fall kann man die Variablen so umbenennen, dass man eine Nullstelle genau dann erhält, wenn

$$z_1^2 = (\frac{x_1}{a})^2 + (\frac{y_1}{b})^2$$

gilt. Das beschreibt einen elliptischen Doppelkegel. Die Nullstellen von p ergeben sich als Schnitt dieses Doppelkegels mit der Ebene $z = 1$.

11 Hauptidealringe

11.1 Definition: *Links–, Rechts–, zweiseitiges Ideal*

Sei R ein Ring (siehe 2.16), und sei I eine Untergruppe von $(R,+)$. Man nennt I ein Linksideal (von R), falls $rx \in I$ für alle $r \in R$, $x \in I$. Entsprechend ($xr \in I$ für alle $r \in R$, $x \in I$) sind Rechtsideale definiert. Falls I sowohl Rechts– als auch Linksideal ist, heißt I ein (zweiseitiges) Ideal von R (Schreibweise: $I \triangleleft R$).

11.2 Beispiel/Bemerkung:

(i) $\{0\}$ (das Nullideal) und R sind Ideale für jeden Ring R.

(ii) $R = \mathbb{Z}$, $I = \{\text{alle geraden Zahlen}\}$ ist ein Ideal.

(iii) Wenn R kommutativ ist, fallen die Begriffe Linksideal, Rechtsideal und Ideal zusammen.

(iv) Die Menge R aller stetigen Funktionen von $\mathbb{R}$ in $\mathbb{R}$ bildet mit den Verknüpfungen

$$(f+g)(x) = f(x) + g(x)$$
$$(f \cdot g)(x) = f(x) \cdot g(x)$$

einen (kommutativen) Ring. Sei X eine Teilmenge von $\mathbb{R}$. Dann ist

$$N_X = \{f \in R \mid f(x) = 0 \ \forall x \in X\}$$

ein Ideal von R. Es ist

$$X = \emptyset \Leftrightarrow N_X = R,$$
$$X \text{ dicht in } \mathbb{R} \Leftrightarrow N_X = \{0\}.$$

(v) Wenn R ein Ring mit Eins ist und I ein einseitiges Ideal, dann gilt

$$I = R \Leftrightarrow 1 \in I.$$

Beweis:

„$\Rightarrow$“: trivial

„$\Leftarrow$“: Sei $r \in R$. Es ist $r = r \cdot 1 = 1 \cdot r \in I$, da I Links– oder Rechtsideal ist. Also ist $R \subseteq I$. Immer gilt $I \subseteq R$, also $R = I$.

(vi) Schnitte und Summen von Linksidealen sind wieder Linksideale.

(vii) Sei $a \in R$. Es ist $Ra = \{ra \mid r \in R\}$ ein Linksideal von R, genannt das von a erzeugte Linkshauptideal.

11.3 Definition: *Hauptideal, Hauptidealring*

(1) Sei R ein kommutativer Ring, und sei I ein Ideal von R. Man nennt I ein Hauptideal, wenn ein $a \in R$ existiert mit $I = Ra$.

(2) Ein Hauptidealring ist ein kommutativer, nullteilerfreier (d.h. für alle $a, b \in R$ gilt: $ab = 0 \Rightarrow a = 0$ oder $b = 0$) Ring mit Eins, in welchem jedes Ideal ein Hauptideal ist.

11.4 Lemma:

$\mathbb{Z}$ ist ein Hauptidealring.

Beweis: $\mathbb{Z}$ ist kommutativ, nullteilerfrei, mit Eins. Sei I ein Ideal von $\mathbb{Z}$, zu zeigen: I ist ein Hauptideal. Falls $I = \{0\}$, dann ist $I = \mathbb{Z} \cdot 0$, fertig. Sei also $I \neq \{0\}$ und $0 \neq x \in I$. Dann ist auch $-x \in I$ (da I Untergruppe von $(\mathbb{Z}, +)$). Entweder x oder $-x$ ist positiv, also aus $\mathbb{N}$. Daher ist $I \cap \mathbb{N} \neq \emptyset$. Sei m das kleinste Element in $I \cap \mathbb{N}$ ($\Rightarrow m \neq 0$).

Behauptung: $I = \mathbb{Z}m$

Beweis: Da $m \in I$, ist auch $zm \in I$ für alle $z \in \mathbb{Z}$, also $\mathbb{Z}m \subseteq I$. Umgekehrt sei $a \in I$. Division mit Rest ergibt $a = tm + r$ mit $t, r \in \mathbb{Z}$ und $0 \leq r < m$. Dann ist $r = a - tm \in I$, denn $a \in I$ und $tm \in \mathbb{Z}m \subseteq I$ und I ist Untergruppe. Da $r < m$ und m minimal in $I \cap \mathbb{N}$, ist $r \notin I \cap \mathbb{N}$, also $r \notin \mathbb{N}$. Daher ist $r = 0$, $a = tm \in \mathbb{Z}m$, also $I \subseteq \mathbb{Z}m$.

11.5 Definition: *Grad eines Polynoms*

Sei $\mathbb{K}$ ein Körper und $p = a_n x^n + \ldots + a_1 x + a_0 \in \mathbb{K}[x]$. Wenn $a_n \neq 0$, dann heißt $n = \deg p$ der Grad von p. Wenn p das Nullpolynom ist, setzt man $\deg p = -\infty$.

11.6 Lemma: *Gradformel*

Sei $p, q \in \mathbb{K}[x]$. Dann gelten

(1) $\deg(pq) = \deg(p) + \deg(q)$

(2) $\deg(p + q) \leq \max(\deg p, \deg q)$. Wenn $\deg p \neq \deg q$, gilt Gleichheit.

Beweis: ist trivial.

11.7 Lemma: *Division mit Rest in* $\mathbb{K}[x]$

Sei $\mathbb{K}$ ein Körper und seien $p, q \in \mathbb{K}[x]$ mit $q \neq 0$. Dann existieren $t, r \in \mathbb{K}[x]$ mit $p = tq + r$ und $\deg r < \deg q$. Die Polynome t und r sind dadurch eindeutig bestimmt.

Beweis:

Existenz: Induktion über $\deg p$
Wenn $\deg p < \deg q$, wähle $t = 0$, $r = p$. Sei also $p = a_n x^n + \ldots + a_0$, $q = b_m x^m + \ldots + b_0$, mit $a_n, b_m \neq 0$ und $n \geq m$. Sei $t_1 = a_n b_m^{-1} x^{n-m} \in \mathbb{K}[x]$. Dann ist $t_1 q = a_n x^n +$ Terme kleineren Grades. Setzt man $p_1 = p - t_1 q$, dann ist $\deg p_1 < \deg p$. Per Induktion gibt es $t_2, r \in \mathbb{K}[x]$ mit $p_1 = t_2 q + r$ und $\deg r < \deg q$. Also ist $p = p_1 + t_1 q = t_2 q + r + t_1 q = (t_1 + t_2)q + r$; fertig mit $t = t_1 + t_2$.

Eindeutigkeit: Sei auch $p = t'q + r'$ mit $\deg r' < \deg q$. Dann ist $tq + r = t'q + r'$, also $(t - t')q = r' - r$. Wäre $t \neq t'$, dann $\deg(t - t') \geq 0$, also $\deg(r' - r) = \deg((t' - t)q) = \deg(t' - t) + \deg q \geq \deg q$ nach 11.6 (1). Nach 11.6 (2) ist aber $\deg(r' - r) \leq \max(\deg r, \deg r') < \deg q$, ein Widerspruch. Also ist $t = t'$ und $r' - r = (t - t')q = 0$, d.h. $r = r'$.

11.8 Satz:

$\mathbb{K}[x]$ ist ein Hauptidealring für jeden Körper $\mathbb{K}$.

Beweis: $\mathbb{K}[x]$ ist kommutativ mit Eins.

Nullteilerfrei: Wenn $p, q \in \mathbb{K}[x]$ beide $\neq 0$, dann ist $\deg p,\ \deg q \geq 0$, also $\deg(pq) = \deg p + \deg q \geq 0$ und daher $pq \neq 0$.

Sei I ein Ideal von $\mathbb{K}[x]$, zu zeigen: I ist Hauptideal. Dies ist klar, falls $I = \{0\}$. Andernfalls wähle in I ein Polynom $q \neq 0$ vom kleinsten Grad. Klar ist $\mathbb{K}[x]q \subseteq I$, weil I ein Ideal ist. Umgekehrt sei $p \in I$. Dann ist $p = tq + r$ mit geeigneten $t, r \in \mathbb{K}[x]$ so, daß $\deg r < \deg q$ (nach 11.7). Da $r = p - tq \in I$, erzwingt die Wahl von q, daß $r = 0$, also $p = tq \in \mathbb{K}[x]q$. Daher ist $I = \mathbb{K}[x]q$ ein Hauptideal.

11.9 Definition: *Einheit, Inverse, Vielfaches, irreduzibel, teilerfremd, Primideal, Primelement, maximales Ideal*

Sei R ein kommutativer, nullteilerfreier Ring mit Eins.

(1) $u \in R$ heißt Einheit, falls ein $v \in R$ existiert mit $uv = 1$. (v ist dann eindeutig bestimmt und heißt das Inverse von u, $v = u^{-1}$).

(2) Seien $a, b \in R$. Man sagt a teilt b oder b ist ein Vielfaches von a und schreibt $a \mid b$, falls ein $c \in R$ existiert mit $ac = b$.

(3) Wenn p keine Einheit ist, aber aus $p = ab$ stets folgt, daß a oder b eine Einheit ist, dann heißt p irreduzibel.

(4) Seien $a_1, \ldots, a_n \in R$. Man nennt diese Elemente teilerfremd, falls gilt:

$$u \mid a_i \ \forall i = 1, \ldots, n \Rightarrow u \text{ ist Einheit.}$$

(5) Ein Ideal $P \neq R$ heißt Primideal, falls gilt:

$$ab \in P \Rightarrow a \in P \text{ oder } b \in P.$$

(6) Ein Element $0 \neq p \in R$ heißt Primelement, wenn Rp ein Primideal ist.

(7) Ein Ideal M heißt maximal, falls $\{0\} \subseteq M \subsetneq R$ und kein Ideal I existiert mit $M \subsetneq I \subsetneq R$.

11.10 Bemerkung:

(i) In einem Körper sind alle Elemente $\neq 0$ Einheiten; Primelemente gibt es nicht. In $\mathbb{Z}$ sind nur ± 1 Einheiten; Primelemente sind die Primzahlen und ihre Negativen. In $\mathbb{K}[x]$ sind die Einheiten genau die konstanten Polynome $\neq 0$; Primelemente sind die irreduziblen Polynome, wie wir gleich zeigen werden.

(ii) $a|b$ genau dann, wenn $Ra \geq Rb$. Insbesondere ist $Ra = Rb$ genau dann, wenn es eine Einheit $u \in R$ gibt mit $b = ua$.

(iii) Ein Primelement p ist stets irreduzibel.

Beweis: Sei $p = ab$. Dann ist $ab \in Rp$, und da dies ein Primideal ist, folgt o.B.d.A. $b \in Rp$, also $b = cp$ für ein $c \in R$. Daher ist $p = acp$, d.h. $(1 - ac)p = 0$ und daher $1 = ac$. Also ist a eine Einheit. Weil $Rp \neq R$, ist p keine Einheit.

(iv) Die Umkehrung gilt i.A. nicht. Beispiel: Die Menge $R = \{a + bi\sqrt{5} \mid a, b \in \mathbb{Z}\}$ ist ein Unterring mit Eins von $\mathbb{C}$, daher kommutativ und nullteilerfrei. Das Element $2 + i\sqrt{5} \in R$ ist irreduzibel, aber kein Primelement. (Übungsaufgabe)

(v) In einem Hauptidealring braucht man irreduzible Elemente und Primelemente nicht zu unterscheiden. Dies ist Teil des nächsten Satzes.

11.11 Satz:

Sei R ein Hauptidealring und $0 \neq p \in R$.

(1) Jede nichtleere Menge von Idealen enthält maximale Elemente. Insbesondere ist jedes Ideal $\neq R$ in einem maximalen Ideal enthalten.

(2) Die folgenden Aussagen sind äquivalent:

(i) p ist Primelement.

(ii) p ist irreduzibel.

(iii) Rp ist maximales Ideal.

(iv) Rp ist Primideal.

Beweis:

(1) Um diese Aussage in voller Allgemeinheit zu beweisen, benötigt man das Zorn'sche Lemma, welches wir nicht zur Verfügung haben. Daher beschränken wir uns beim Beweis auf die beiden Spezialfälle $R = \mathbb{Z}$ und $R = \mathbb{K}[x]$. Sei $\emptyset \neq \mathcal{M}$ eine Menge von Idealen. Wenn $\mathcal{M}$ nur das Nullideal enthält, ist nichts zu zeigen. Wir nehmen also das Gegenteil an. Jedes Ideal $\{0\} \neq I \in \mathcal{M}$ ist von einem Element $0 \neq a = a(I)$ erzeugt. Wählt man $I \in \mathcal{M}$ so, dass $0 < |a(I)|$ (für $R = \mathbb{Z}$) b.z.w. $0 \leq \deg(a(I))$ (für $R = \mathbb{K}[x]$) minimal wird, dann ist I maximal in $\mathcal{M}$, denn ein größeres Ideal wird nach 11.10, (ii) von einem echten Teiler von $a(I)$ erzeugt.
Die zweite Aussage folgt aus der ersten: sei $I \neq R$ ein Ideal; ein maximales Element in der (nicht-leeren) Menge $\mathcal{M} = \{A \mid I \leq A \lhd R,\ A \neq R\}$ ist dann ein maximales Ideal von R, welches I enthält.

(2) (i)⇒(ii): steht schon in 11.10, (iii).
(ii)⇒(iii): Weil p keine Einheit ist, ist $Rp \neq R$. Sei $Rp \lneq A$ für ein Ideal A. Zu zeigen: $A = R$. Da A ein Hauptideal ist, folgt $A = Ra$ für ein a. Wegen $p \in A$ gibt es ein b mit $p = ab$. Die Irreduzibilität von p erzwingt, daß a oder b eine Einheit ist. Wäre b eine Einheit, dann $a = pb^{-1} \in Rp$, also $A = Ra \leq Rp$, Widerspruch. Also ist a eine Einheit und daher $A = Ra = R$.
(iii)⇒(iv): Sei $ab \in Rp$. Angenommen, weder a noch b liegen in Rp. Dann sind die Ideale $Ra + Rp$ und $Rb + Rp$ beide echt größer als Rp, also gleich R wegen der Maximalität von Rp. Daher gibt es $r, s, t, u \in R$ mit $ra + sp = 1 = tb + up$. Dann folgt $1 = (ra + sp)(tb + up) = rtab + p(rua + stb + sup) \in Rp$, also $Rp = R$, Widerspruch.
(iv)⇒(i): ist trivial.

11.12 Bemerkung:

Außer den von den Primelementen erzeugten Idealen hat ein Hauptidealring noch genau ein weiteres Primideal, nämlich $\{0\}$ (da der Ring nullteilerfrei ist). Dieses Primideal ist genau dann ein maximales Ideal, wenn der Ring ein Körper ist.

11.13 Satz: *Primfaktorzerlegung in Hauptidealringen*

Sei R ein Hauptidealring und $\{Rp_i \mid i \in I\}$ die Menge der maximalen Ideale $\neq \{0\}$. Dann hat jedes Element $0 \neq a \in R$ eine (bis auf die Reihenfolge) eindeutige Faktorisierung

$$a = u \prod_{i \in I} p_i^{n_i}$$

mit $n_i \in \mathbb{N}_0$, fast alle $n_i = 0$, und einer Einheit u.

Beweis:

Existenz: Es sei

$$\mathcal{M} = \left\{ Rb \mid a = b \prod_{i \in I} p_i^{n_i}, \text{ fast alle } n_i = 0 \right\}.$$

Dann ist $\mathcal{M} \neq \emptyset$, denn $Ra \in \mathcal{M}$. Also gibt es ein maximales Element Rc in $\mathcal{M}$ nach 11.11 (1).

Wenn $Rc \neq R$, dann liegt Rc in einem maximalen Ideal M von R. Es ist $M \neq \{0\}$, denn sonst ist $Rc = \{0\}$, also $c = 0$, $a = 0$, Widerspruch. Also ist $M = Rp_j$ für ein $j \in I$. Dann ist $Rc \subseteq Rp_j$, also $c = dp_j$ mit geeignetem d. Wegen

$$a = c \prod_{i \in I} p_i^{n_i} = dp_j \prod_{i \in I} p_i^{n_i}$$

folgt $Rd \in \mathcal{M}$. Es ist aber $c = p_j d \in Rd$, also $Rc \subseteq Rd$, und daher $Rc = Rd$, wegen der Maximalität von Rc in $\mathcal{M}$. Es folgt, daß p_j eine Einheit ist, d.h. $Rp_j = R$. Aber Rp_j ist ein maximales Ideal, Widerspruch.

Also ist $Rc = R$, d.h. c ist Einheit und $a = c \prod_{i \in I} p_i^{n_i}$ eine Faktorisierung wie gewollt.

Eindeutigkeit: Angenommen, es ist

$$u \prod_{i \in I} p_i^{a_i} = v \prod_{i \in I} p_i^{b_i}$$

mit $a_1 > b_1$. Dann ist

$$up_1^{a_i - b_1} \prod_{i \neq 1} p_i^{a_i} = v \prod_{i \neq 1} p_i^{b_i} \in Rp_1.$$

Da Rp_1 ein Primideal ist, muß mindestens einer der Faktoren von $v \prod_{i \neq 1} p_i^{b_i}$ in Rp_1 liegen. v tut's nicht, weil v eine Einheit ist. Es ist aber auch $p_j \notin Rp_1$, denn sonst ist $Rp_j \subseteq Rp_1$, also $Rp_j = Rp_1$, wegen der Maximalität von Rp_j und Rp_1 (11.11), Widerspruch. Also ist $a_i = b_i$ für alle i und dann auch $u = v$.

11.14 Korollar:

Sei $0 \neq f \in \mathbb{K}[x]$ ein Polynom vom Grad n. Dann hat f höchstens n Nullstellen in $\mathbb{K}$.

Beweis: Wenn $a_1, \ldots, a_m \in \mathbb{K}$ Nullstellen von f sind, dann sind $x - a_1, \ldots, x - a_m$ verschiedene irreduzible Teiler von f. Daher ist auch das Produkt dieser m linearen Polynome ein Teiler von f. Die Behauptung folgt.

11.15 Lemma:

Seien $a_1, \ldots, a_n$ teilerfremde Elemente aus dem Hauptidealring R. Dann gibt es $b_1, \ldots, b_n \in R$ mit $1 = \sum_{i=1}^{n} b_i a_i$.

Beweis: Sei $I = \sum_{i=1}^{n} Ra_i$. Dann ist I ein Ideal, also $I = Ra$ für ein a. Da $a_i \in I$, gibt es ein c_i mit $a_i = c_i a$, also $a | a_i$ für $i = 1, \ldots, n$. Nach Voraussetzung ist dann a eine Einheit, also $I = R$. Insbesondere ist $1 \in I$, also gibt es b_i mit der gewünschten Eigenschaft.

11.16 Bemerkung/Definition: *größter gemeinsamer Teiler*

Seien $a_1, \ldots, a_n$ Elemente aus dem Hauptidealring R. Dann ist $I = \sum_{i=1}^{n} Ra_i$ ein Ideal, also $I = Rg$ für ein $g \in R$, welches bis auf Einheiten eindeutig bestimmt ist. Man nennt g den größten gemeinsamen Teiler (ggT) von $a_1, \ldots, a_n$.

11.17 Bemerkung:

(i) Nach Definition von g gibt es $b_1, \ldots, b_n \in R$ mit $g = \sum_{i=1}^{n} b_i a_i$. Dies ist eine Verallgemeinerung von 11.15.

(ii) Weil $Rg \geq Ra_i$, ist $g | a_i$ für jedes i nach 11.10 (ii), d.h. g ist gemeinsamer Teiler der a_i. Wenn auch $t | a_i$ für jedes i, dann ist $Ra_i \leq Rt$, also $Rg = \sum_{i=1}^{n} Ra_i \leq Rt$, und daher $t | g$. Daher ist g der 'größte' gemeinsame Teiler.

11.18 Bemerkung/Definition: *kleinstes gemeinsames Vielfaches*

Seien $a_1, \ldots, a_n$ Elemente aus dem Hauptidealring R. Dann ist $I = \bigcap_{i=1}^{n} Ra_i$ ein Ideal, also $I = Rk$ für ein $k \in R$, welches bis auf Einheiten eindeutig bestimmt ist. Man nennt k das kleinste gemeinsame Vielfache (kgV) von $a_1, \ldots, a_n$.

11.19 Bemerkung:

Weil $Rk \leq Ra_i$, ist $a_i | k$ für jedes i nach 11.10 (ii), d.h. k ist gemeinsames Vielfaches der a_i. Wenn auch $a_i | l$ für jedes i, dann ist $Ra_i \geq Rl$, also $Rk = \bigcap_{i=1}^{n} Ra_i \geq Rl$, und daher $k | l$. Daher ist k das 'kleinste' gemeinsame Vielfache.

11.20 Bemerkung:

Sei $0 \neq a_i$ für $i = 1, \ldots, n$ und sei $a_i = \prod_j p_j^{e_{ij}}$ die Faktorisierung in Primelemente (Einheiten können ignoriert werden). Für jedes j sei $u_j = \min(e_{1j}, \ldots, e_{nj})$ und $o_j = \max(e_{1j}, \ldots, e_{nj})$. Dann ist $g = \prod_j p_j^{u_j}$ der größte gemeinsame Teiler und $k = \prod_j p_j^{o_j}$ das

kleinste gemeinsame Vielfache der a_i. Wenn insbesondere $n = 2$, dann folgt hieraus und aus $\min(x, y) + \max(x, y) = x + y$, dass $k \cdot g = a_1 \cdot a_2$.

11.21 Bemerkung: *Euklidischer Algorithmus*

Um den größte gemeinsame Teiler zweier Polynome $a \neq 0$ und $b \neq 0$ in $\mathbb{K}[x]$ zu berechnen, ist es nicht nötig, diese in irreduzible Faktoren zu zerlegen. Wir dürfen $\deg(a) \geq \deg(b)$ annehmen. Durch wiederholte Division mit Rest erzeugen wir dann drei Folgen von Polynomen $a_0, a_1, \ldots, a_n, a_{n+1} = 0$, $s_0, s_1, \ldots, s_n = s$ und $t_0, t_1, \ldots, t_n = t$ mit $a_i = s_i a + t_i b$ für $i = 0, \ldots, n$ und $ggT(a, b) = ggT(a_i, a_{i+1}) = a_n = s \cdot a + t \cdot b$. Dazu setzen wir $a_0 = a$, $s_0 = 1$, $t_0 = 0$, $a_1 = b$, $s_1 = 0$ und $t_1 = 1$, so dass also $a_i = s_i a + t_i b$ für $i = 0$, 1 und $ggT(a, b) = ggT(a_i, a_{i+1})$ für $i = 0$ gelten. Solange nun $a_i \neq 0$ ist, definieren wir a_{i+1} durch Division mit Rest:

$$a_{i-1} = q_i a_i + a_{i+1} \quad .$$

Aus dieser Gleichung folgt schon, dass $a_i\mathbb{K}[x] + a_{i+1}\mathbb{K}[x] = a_{i-1}\mathbb{K}[x] + a_i\mathbb{K}[x]$, also

$$ggT(a_i, a_{i+1}) = ggT(a_{i-1}, a_i) = ggT(a, b) \quad ,$$

wobei die zweite Gleichheit per Induktion gilt. Setzt man $s_{i+1} = s_{i-1} - q_i s_i$ und $t_{i+1} = t_{i-1} - q_i t_i$, dann ist – wieder per Induktion –

$$a_{i+1} = a_{i-1} - q_i a_i = (s_{i-1}a + t_{i-1}b) - q_i(s_i a + s_i b) = s_{i+1}a + t_{i+1}b \quad .$$

Da die Grade der Polynome a_i fallen, geht die Division schließlich auf. Es ist dann $a_{n+1} = 0$, also $ggT(a, b) = ggT(a_n, a_{n+1}) = a_n$.
Wenn man nur am $ggT(a, b)$ interessiert ist, nicht an seiner Darstellung durch a und b, braucht man offenbar nur die a_i's zu berechnen, nicht die s_i's und t_i's.
Ganz analog kann natürlich in $\mathbb{Z}$ verfahren werden.

12 Normalformen von Matrizen I: Die kanonische rationale

Im ganzen Paragraphen ist $\mathbb{K}$ ein beliebiger Körper, V ein endlich–dimensionaler $\mathbb{K}$–VR, etwa $\dim V = n$, und α ein Endomorphismus von V.

12.1 Bemerkung/Definition: *Minimalpolynom von α*

Für jedes $f \in \mathbb{K}[x]$ ist $f(\alpha)$ ein Endomorphismus von V; es ist leicht zu sehen, dass $I = \{f \in \mathbb{K}[x] \mid f(\alpha) = 0\}$ ein Ideal von $\mathbb{K}[x]$ ist. Dieses Ideal ist nicht das Nullideal: da $\dim \mathrm{End}\,(V) = n^2$ (siehe 7.11), können die Elemente $\mathrm{id}_V = \alpha^0, \alpha, \alpha^2, \cdots, \alpha^{n^2}$ nicht linear unabhängig sein. Also gibt es $k_0, \cdots, k_{n^2} \in \mathbb{K}$, nicht alle $= 0$, mit $\sum_{i=0}^{n^2} k_i \alpha^i = 0$. Damit ist ein Polynom $\neq 0$ in I gefunden. Nach den Ergebnissen des vorigen Paragraphen ist $I = m\mathbb{K}[x]$ für ein Polynom $m \neq 0$, welches als normiert angenommen werden kann, und dann durch α eindeutig bestimmt ist. Dieses m heißt das Minimalpolynom von α.
Wenn A eine quadratische Matrix ist, definiert man entsprechend das Minimalpolynom von A.

12.2 Bemerkung:

(i) Minimalpolynom und charakteristisches Polynom von α müssen unterschieden werden! Zu den Beziehungen zwischen beiden folgt unten genaueres.

(ii) Wenn A die Matrix von α bezüglich einer beliebigen Basis ist, dann ist das Minimalpolynom von A zugleich das Minimalpolynom von α. Insbesondere haben ähnliche Matrizen das gleiche Minimalpolynom.

(iii) Das Argument in der Definition zeigt, dass es ein Polynom $\neq 0$ in I vom Grad $\leq n^2$ gibt. Also ist insbesondere $\deg m \leq n^2$. Wir werden zeigen, dass sogar $\deg m \leq n$ gilt.

(iv) Nach Definition von m gilt für beliebiges $f \in \mathbb{K}[x]$, dass $f(\alpha) = 0$ genau dann, wenn $m|f$.

12.3 Bezeichnung:

Im ganzen Paragraphen ist m das Minimalpolynom von α. Um Klammern zu sparen, schreiben wir einfach fv für $(f(\alpha))(v)$, wenn $v \in V$ und $f \in \mathbb{K}[x]$. Entsprechend sind $\mathrm{Ker}\,(f) = \mathrm{Ker}\,(f(\alpha))$ und $\mathrm{Im}\,(f) = \mathrm{Im}\,(f(\alpha))$.

12.4 Definition: *invarianter Unterraum*

Ein Unterraum $U \leq V$ mit $\alpha(u) \in U$ für alle $u \in U$ heißt invarianter Unterraum von V.

12.5 Bemerkung/Definition:

(i) Beispiele für invariante Unterräume sind $\mathrm{Im}\,(f)$ und $\mathrm{Ker}\,(f)$ für beliebiges $f \in \mathbb{K}[x]$. Außerdem sind Summen und Schnitte von invarianten Unterräumen invariant.

(ii) Wählt man ein beliebiges $v \in V$ und betrachtet dann den Unterraum U, der von $\{v, \alpha(v), \alpha^2(v), \cdots\}$ erzeugt wird, so erhält man einen invarianten Unterraum. Solche invarianten Unterräumen heißen zyklisch.

(iii) Wenn U ein invarianter Unterraum ist, dann kann man α auf U einschränken und erhält einen Endomorphismus $\alpha_{|U}$ von U. Dieser hat ebenfalls ein Minimalpolynom, etwa m_U. Da $m(\alpha_{|U}) = 0$ ist, folgt $m_U | m$ aus 12.2 (iv).

12.6 Lemma:

Wenn $f \cdot g = m$ mit f und g normiert, dann ist $U := \mathrm{Ker}\,(g) \geq \mathrm{Im}\,(f)$, und $m_U = g$.

Beweis: Es ist $gfV = mV = 0$, also $\mathrm{Im}\,(f) = fV \leq \mathrm{Ker}\,(g) = U$. Insbesondere ist $m_U fV = 0$, also $m | f \cdot m_U$ und daher $g | m_U$. Andererseits ist $g(\alpha_{|U}) = 0$ nach Definition von U, also $m_U | g$. Da beide Polynome normiert sind, folgt Gleichheit.

12.7 Satz: *zyklische Räume*

Wenn V ein zyklischer Raum ist, etwa $V =< v, \alpha(v), \alpha^2(v), \cdots >$, dann gelten:

(i) $n = \dim V = \deg m$, und $\{v, \alpha(v), \ldots, \alpha^{n-1}(v)\}$ ist eine Basis von V. Außerdem ist m das normierte Polynom kleinsten Grades mit $mv = 0$.

(ii) Bezüglich der Basis in (i) hat α die Matrix

$$A = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -k_0 \\ 1 & \ddots & & \vdots & \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \cdots & 1 & 0 & \\ 0 & \cdots & 0 & 1 & -k_{n-1} \end{pmatrix},$$

wobei

$$m(x) = x^n + \sum_{j=0}^{n-1} k_j x^j$$

das Minimalpolynom ist.

(iii) m ist zugleich das charakteristische Polynom von α.

(iv) Wenn $m = f \cdot g$ eine Faktorisierung von m mit normiertem f und g ist, dann ist $\mathrm{Im}\,(f) = \mathrm{Ker}\,(g)$ ein invarianter Unterraum U von V mit $m_U = g$. Der Unterraum U ist ebenfalls zyklisch, und es gilt $\dim U = \deg g$.

(v) Die Abbildung $g \mapsto \mathrm{Ker}\,(g)$ ist eine Bijektion zwischen den normierten Teilern von m und den invarianten Unterräumen von V.

Beweis: (i) Sei $r \in \mathbb{N}$ so gewählt, dass $\{v, \alpha(v), \ldots, \alpha^{r-1}(v)\}$ linear unabhängig, aber $\{v, \alpha(v), \ldots, \alpha^{r-1}(v), \alpha^r(v)\}$ linear abhängig ist. Offenbar ist dann $\alpha^r(v)$ eine Linearkombination von $\{v, \alpha(v), \ldots, \alpha^{r-1}(v)\}$ (vergleiche 4.8). Also gibt es $k_i \in \mathbb{K}$ mit

$$0 = \alpha^r(v) + \sum_{i=0}^{r-1} k_i \alpha^i(v) = pv$$

mit

$$p(x) = x^r + \sum_{i=0}^{r-1} k_i x^i \ ,$$

und nach Wahl von r ist p das normierte Polynom kleinsten Grades, welches v annulliert. Es annulliert dann aber auch alle $\alpha^j(v)$, also auch deren Erzeugnis, d.h. ganz V. Daher ist $p = m$, insbesondere $\deg m = r$.
Per Induktion sieht man leicht, dass sich mit $\alpha^r(v)$ auch jedes $\alpha^s(v)$ für $s \geq r$ als Linearkombination von $\{v, \alpha(v), \ldots, \alpha^{r-1}(v)\}$ schreiben läßt . Diese Menge ist also eine Basis von V, insbesondere ist $r = n$.
(ii) Das Bild eines der Basisvektoren $\alpha^j(v)$ unter α ist $\alpha^{j+1}(v)$, also der nächste Basisvektor außer im Fall $j = n - 1$. Dann ist

$$\alpha^n(v) = -\sum_{i=0}^{n-1} k_i \alpha^i(v) \ .$$

Daher hat A die angegebene Form.
(iii) Es ist

$$xE - A = \begin{pmatrix} x & \cdots & \cdots & 0 & k_0 \\ -1 & \ddots & & \vdots & \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \cdots & -1 & x & k_{n-2} \\ 0 & \cdots & 0 & -1 & x + k_{n-1} \end{pmatrix}$$

Die Untermatrix, welche durch Streichen der ersten Zeile und ersten Spalte entsteht, hat die gleiche Form, also (per Induktion) die Determinante

$$x^{n-1} + \sum_{i=1}^{n-1} k_i x^{i-1} \ .$$

Durch Entwickeln nach der ersten Zeile (siehe 8.15) erhält man daher

$$\det(xE - A) = x(x^{n-1} + \sum_{i=1}^{n-1} k_i x^{i-1} \ ) + (-1)^{n+1} k_0 (-1)^{n-1} = m(x) \ .$$

Das ist die Behauptung.
(iv) In 12.6 stehen schon viele der Aussagen. Wir zeigen, dass $\mathrm{Ker}\,(g) \leq \mathrm{Im}\,(f)$. Dazu sei $u \in \mathrm{Ker}\,(g)$. Weil V zyklisch ist, gibt es ein Polynom h mit $u = hv$, also $u \in \mathrm{Im}\,(h)$. Dann ist $0 = gu = ghv$, also nach (i) $m = gf$ ein Teiler von gh und daher $f|h$. Folglich ist $\mathrm{Im}\,(f) \geq \mathrm{Im}\,(h)$ und insbesondere $u \in \mathrm{Im}\,(f)$.
Es ist klar, dass $\{fv, \alpha(fv), \ldots\}$ ein Erzeugenden-System von $fV = U$ ist; also ist U zyklisch. Nach (i) ist dann $\dim U = \deg m_U = \deg g$.
(v) Wenn g und h normierte Teiler von m sind mit $\mathrm{Ker}\,(g) = \mathrm{Ker}\,(h) = U$, dann ist $g = m_U = h$ nach 12.6. Also ist die Abbildung $g \mapsto \mathrm{Ker}\,(g)$ injektiv. Sie ist auch surjektiv: Dazu sei U ein beliebiger invarianter Unterraum. Setzt man $I = \{h \in \mathbb{K}[x] \mid hv \in U\}$, dann ist leicht zu kontrollieren, dass I ein Ideal von $\mathbb{K}[x]$ ist, welches m enthält. Also ist $I = \mathbb{K}[x]f$ für ein geeignetes normiertes $f \in \mathbb{K}[x]$ mit $f|m$, etwa $m = f \cdot g$. Weil V zyklisch ist, ist $U = Iv = f\mathbb{K}[x]v = fV = \mathrm{Im}\,(f) = \mathrm{Ker}\,(g)$, wobei die letzte Gleichheit aus (iv) folgt.

12.8 Lemma:

Seien $\lambda_1, \ldots, \lambda_t \in V^*$. Dann

(i)
$$\dim(\bigcap_{i=1}^{t} \mathrm{Ker}\,(\lambda_i)\ ) \geq n - t\ .$$

(ii) Wenn $\mu \in V^*$ eine Linearkombination von $\{\lambda_1, \ldots, \lambda_t\}$, dann

$$\mathrm{Ker}\,(\mu) \cap \bigcap_{i=1}^{t} \mathrm{Ker}\,(\lambda_i) = \bigcap_{i=1}^{t} \mathrm{Ker}\,(\lambda_i)\ .$$

Beweis: (i) Sei $\lambda : V \to \mathbb{K}^t$ definiert durch $\lambda(v) = (\lambda_1(v), \ldots, \lambda_t(v))$. Dann hat $\mathrm{Im}\,(\lambda)$ als Unterraum von $\mathbb{K}^t$ höchstens die Dimension t. Weil

$$\mathrm{Ker}\,(\lambda) = \bigcap_{i=1}^{t} \mathrm{Ker}\,(\lambda_i)\ ,$$

folgt die Behauptung aus 5.13.
(ii) Ein $v \in \mathrm{Ker}\,(\lambda)$ wird auch von jeder Linearkombination der λ_i's annulliert, liegt also auch in $\mathrm{Ker}\,(\mu)$. Daraus folgt die Behauptung.

12.9 Satz:

Wenn $m = p^r$ eine Potenz eines irreduziblen Polynoms p ist, dann gilt:

(i) Es gibt einen zyklischen invarianten Unterraum Z von V mit $m_Z = m$.

(ii) Zu jedem Z wie in (i) gibt es einen invarianten Unterraum W mit $V = Z \oplus W$.

Beweis: (i) Sonst wäre zu jedem invarianten zyklischen Unterraum U das Minimalpolynom m_U ein Teiler von p^{r-1}. Aber dann wäre $p^{r-1}(\alpha) = 0$ entgegen der Minimalität von m.
(ii) Nach 12.7 (iv) und (v) enthält Z genau einen kleinsten invarianten Unterraum $\neq 0$, nämlich $S = p^{r-1}Z$. Sei $\lambda \in V^*$ so gewählt, dass $\lambda_{|S} \neq 0$. Für jedes $i = 0, 1, \ldots$ ist $\lambda_i := \lambda\alpha^i \in V^*$. Sei

$$W = \bigcap_{i=0}^{\infty} \mathrm{Ker}\,(\lambda_i)\ .$$

Dann ist W invariant, denn wenn $w \in W$, dann ist $\lambda_i\alpha(w) = \lambda_{i+1}(w) = 0$ für alle i, also $\alpha(w) \in W$. Setzt man $t = \deg m$, so ist α^t und jede höhere Potenz von α eine Linearkombination von $\{\alpha^0, \ldots, \alpha^{t-1}\}$, folglich sind $\lambda_t, \lambda_{t+1}, \ldots$ Linearkombinationen von $\{\lambda_0, \ldots \lambda_{t-1}\}$. Aus 12.8 folgt jetzt, dass

$$W = \bigcap_{i=0}^{t-1} \mathrm{Ker}\,(\lambda_i)$$

und dass $\dim W \geq n - t$.
Als Schnitt von invarianten Unterräumen ist $Z \cap W$ invariant. Daher ist $Z \cap W = 0$, denn

sonst wäre $S \leq Z \cap W \leq W \leq \mathrm{Ker}\,(\lambda)$ im Widerspruch zur Wahl von λ.
Weil Z zyklisch mit Minimalpolynom m ist, gilt $\dim Z = \deg m = t$ nach 12.7, (i). Folglich ist nach 5.24

$$n = \dim(V) \geq \dim(Z+W) = \dim(Z) + \dim(W) \geq t + (n-t) = n\ .$$

Es folgt $Z \oplus W = V$, wie behauptet.

12.10 Satz:

Wenn $m = m_1 \cdot m_2 \cdot \cdots \cdot m_r$ eine Faktorisierung in paarweise teilerfremde, normierte Faktoren ist (d.h. für $i \neq j$ sind m_i und m_j teilerfremd), und wenn $U_i = \mathrm{Ker}\,(m_i)$ für $i = 1, \cdots, r$, dann gilt:

(i) Jedes U_i ist invariant.

(ii) m_i ist das Minimalpolynom von $\alpha_{|U_i}$.

(iii) $V = U_1 \oplus \cdots \oplus U_r$.

Beweis: (i) steht schon in 12.5 und (ii) in 12.6.
(iii) Induktion über r. Der Fall $r = 1$ ist trivial, denn $\mathrm{Ker}\,(m) = V$.

$r = 2$. Nach 11.15 gibt es Polynome f und g mit $fm_1 + gm_2 = 1$. Für jedes $v \in V$ gilt daher $v = 1v = gm_2v + fm_1v \in U_1 + U_2$, denn $gm_2v \in \mathrm{Im}\,(m_2) \leq \mathrm{Ker}\,(m_1) = U_1$ nach 12.6 und ebenso $fm_1v \in U_2$. Wenn $v \in U_1 \cap U_2$, dann ist $v = 1v = gm_2v + fm_1v = 0$; damit ist $V = U_1 \oplus U_2$ gezeigt.

Für $r > 2$ setze $\widetilde{m}_1 = m_1 \cdot m_2 \cdot \cdots \cdot m_{r-1}$ und $\widetilde{U}_1 = \mathrm{Ker}\,(\widetilde{m}_1)$. Verwende den Fall $r = 2$ für $m = \widetilde{m}_1 \cdot m_r$ sowie Induktion für $\alpha_{|\widetilde{U}_1}$.

12.11 Satz: *kanonische rationale Form*

Es gibt eine Basis von V bzgl. derer die Matrix A von α die folgende Form hat:

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_t \end{pmatrix}$$

mit $n_i \times n_i$–Matrizen A_i der Form

$$A_i = \begin{pmatrix} 0 & \cdots & \cdots & 0 & -k_0^{(i)} \\ 1 & \ddots & & \vdots & \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \cdots & 1 & 0 & \\ 0 & \cdots & 0 & 1 & -k_{n_i-1}^{(i)} \end{pmatrix},$$

wobei

$$x^{n_i} + \sum_{j=0}^{n_i-1} k_j^{(i)} x^j = p_i(x)^{s_i}$$

mit $s_i \in \mathbb{N}$ und $p_i \in \mathbb{K}[x]$ irreduzibel ist. (Dabei gilt natürlich: Wenn $d_i = \deg p_i$, dann $n_i = d_i s_i$ und $\sum_{i=1}^{t} n_i = \dim V$.)

Beweis: Faktorisiere das Minimalpolynom in Potenzen von normierten irreduziblen Faktoren wie in 11.13 :

$$m = \prod_{j=1}^{r} p_j^{e_j} \ .$$

Mit $U_j = \operatorname{Ker}\left(p_j^{e_j}\right)$ ist dann $V = U_1 \oplus \cdots \oplus U_r$ nach 12.10, wobei $p_j^{e_j}$ das Minimalpolynom von $\alpha_{|U_j}$ ist. Nach 12.9 (und trivialer Induktion über die Dimension) ist jedes U_j eine direkte Summe von zyklischen invarianten Unterräumen. Für jeden von diesen kann man eine Basis wie in 12.7 (i) wählen. Die Vereinigung der Basen all dieser zyklischen Unterräume ist dann eine Basis von V, bezüglich derer die Matrix von α die behauptete Form hat.

12.12 Korollar:

Es gibt eine Zerlegung $V = Z_1 \oplus \cdots \oplus Z_t$ derart, dass jedes Z_i ein zyklischer invarianter Unterraum und jedes Minimalpolynom m_i von $\alpha_{|Z_i}$ Potenz eines irreduziblen Polynoms ist.

Beweis: Folgt aus dem Beweis von 12.11.

12.13 Definition: *kanonische rationale Form*

Man sagt, daß die Matrix A kanonische rationale Form hat, wenn A wie in 12.11 ist.

12.14 Lemma:

Sei $A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_t \end{pmatrix}$ in kanonisch rationaler Form. Dann gelten:

(1) Für jedes $f \in \mathbb{K}[x]$ ist $f(A) = \begin{pmatrix} f(A_1) & & \\ & \ddots & \\ & & f(A_t) \end{pmatrix}$.

(2) Wenn m_i das Minimalpolynom von A_i ist, dann ist m das kleinste gemeinsame Vielfache der m_i.

Beweis:

(1) ist trivial.
(2) Nach (1) ist $f(A) = 0$ genau dann, wenn $f(A_i) = 0$ für jedes i. Nach 12.2 (iv) ist dies äquivalent zu $m_i | f$ für jedes i, d.h. f ist gemeinsames Vielfaches aller m_i. Daraus folgt die Behauptung.

12.15 Korollar: *Cayley-Hamilton*

(i) Das Minimalpolynom ist ein Teiler des charakteristischen Polynoms.

(ii) α ist Nullstelle seines charakteristischen Polynoms.

(iii) Jeder irreduzible Teiler der charakteristischen Polynoms ist auch Teiler des Minimalpolynoms.

Beweis: Sei A eine Matrix zu α in kanonischer rationaler Form und m_i das Minimalpolynom von A_i wie oben. Dann ist m das kleinste gemeinsame Vielfache der m_i nach 12.14. Für jedes i ist m_i zugleich das charakteristische Polynom von A_i nach 12.7 (iii). Daher ist das charakteristische Polynom von A gleich dem Produkt der m_i. Daraus folgen alle Behauptungen.

12.16 Satz:

Jede quadratische Matrix B ist ähnlich zu einer Matrix $A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_t \end{pmatrix}$ in kanonisch rationaler Form. Bis auf die Reihenfolge der Blöcke A_i ist A durch B eindeutig bestimmt. Sei $p \in \mathbb{K}[x]$ irreduzibel und $0 < k \in \mathbb{N}$; dann läßt sich die Anzahl $z = z(p,k)$ der Blöcke A_i mit Minimalpolynom $m_i = p^k$ aus der Formel

$$z \cdot \deg p = \operatorname{Rg}\left(p^{k-1}(B)\right) + \operatorname{Rg}\left(p^{k+1}(B)\right) - 2 \cdot \operatorname{Rg}\left(p^k(B)\right)$$

bestimmen.

Beweis: Die erste Aussage ist nur eine Umformulierung von 12.11. Die behauptete Eindeutigkeit von A folgt aus der Formel für z ; es genügt also, diese zu beweisen. Weil ähnliche Matrizen gleiche Ränge haben (das folgt aus 7.19), darf man für den Beweis B durch A ersetzen. Aus 12.14 folgt, dass für $f \in \mathbb{K}[x]$ stets $\operatorname{Rg}(f(A)) = \sum_{i=1}^{t} \operatorname{Rg}(f(A_i))$ gilt. Daher genügt es, einen Block A_i zu betrachten und

$$r_i := \operatorname{Rg}\left(p^{k-1}(A_i)\right) + \operatorname{Rg}\left(p^{k+1}(A_i)\right) - 2 \cdot \operatorname{Rg}\left(p^k(A_i)\right) = \begin{cases} \deg p & \text{falls } m_i = p^k \\ 0 & \text{sonst} \end{cases}$$

zu zeigen.
Sei also $m_i = q^s$ mit einem irreduziblen Polynom q.
Wenn $p \neq q$, dann sind p^k und m_i teilerfremd. Nach 11.15 gibt es $f, g \in \mathbb{K}[x]$ mit $1 = f \cdot p^k + g \cdot m_i$. Durch Einsetzen von A_i folgt

$$E = f(A_i) \cdot p^k(A_i) + g(A_i) \cdot m_i(A_i) = f(A_i) \cdot p^k(A_i) \ ,$$

denn $m_i(A_i) = 0$. Hierbei ist E die Einheitsmatrix passender Größe. Insbesondere ist $p^k(A_i)$ invertierbar, hat also vollen Rang (7.39). Mit derselben Begründung gilt dies auch für $p^{k-1}(A_i)$ und $p^{k+1}(A_i)$. Da die drei Matrizen den gleichen Rang haben, ist $r_i = 0$.
Sei also $p = q$. Wenn $s < k$, dann ist $s \leq k-1$ und daher $p^{k-1}(A_i) = 0$ und erst recht $p^k(A_i) = p^{k+1}(A_i) = 0$. Auch in diesem Fall haben also alle beteiligten Matrizen den gleichen Rang (diesmal 0), und daher ist $r_i = 0$.
Wenn $s = k$, dann ist wieder $p^k(A_i) = p^{k+1}(A_i) = 0$. Dagegen ist $\operatorname{Rg}\left(p^{k-1}(A_i)\right) = \deg p$ nach 12.7 (iv) (die Faktorisierung ist hier $m_i = p^k = p^{k-1} \cdot p$; außerdem sollten Sie sich daran erinnern, dass der Rang einer Abbildung die Dimension des Bildes ist). In diesem Fall ist also $r_i = \deg p$.
Mit 12.7 (iv) wird auch im letzten Fall $s > k$ argumentiert: Wenn etwa $s = k + t$, dann ist $m_i = p^s = p^k \cdot p^t$ eine Faktorisierung, also $\operatorname{Rg}\left(p^k(A_i)\right) = \deg p^t = t \cdot \deg p$. Genauso ist $\operatorname{Rg}\left(p^{k+1}(A_i)\right) = (t-1) \cdot \deg p$ und $\operatorname{Rg}\left(p^{k-1}(A_i)\right) = (t+1) \cdot \deg p$. Daraus ergibt sich wieder $r_i = 0$.

12.17 Beispiel:

Sei

$$B = \begin{pmatrix} 2 & -6 & 10 & 4 & -4 & & \\ 17 & 24 & -26 & 22 & 2 & & \\ -1 & 6 & -8 & -2 & 2 & & \\ & & & -3 & 3 & & \\ & & & -27 & 15 & & \\ & & & -24 & 12 & -12 & -12 \\ & & & 27 & -15 & 27 & 24 \end{pmatrix}.$$

Es ist $c_B = (x-6)^7$ das charakteristische Polynom von B; dieses enthält also nur einen irreduziblen Faktor, nämlich $p(x) = x - 6$. Durch elementare Zeilentransformationen erhalten wir die Treppennormalform der Matrix $p(B) = B - 6E$:

$$X = \begin{pmatrix} 1 & 0 & 4/5 & 0 & 0 & 6/5 & 4/5 \\ & 1 & -11/5 & 0 & 0 & 6/5 & 4/5 \\ & & & 1 & 0 & -3/2 & -1 \\ & & & & 1 & -9/2 & -3 \end{pmatrix}$$

(Nullzeilen fortgelassen). Also hat $p(B)$ den Rang 4. Es ist

$$X \cdot p(B) = \frac{6}{5} \begin{pmatrix} -4 & -1 & -1 & -4 & 0 & 0 & 0 \\ 16 & 4 & 4 & 16 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

also ist

$$Y = \begin{pmatrix} 1 & 1/4 & 1/4 & 1 & 0 & 0 & 0 \end{pmatrix}$$

die Treppennormalform der Matrix $p^2(B)$. Der Rang ist jetzt 1. Es ist $Y \cdot p(B) = 0$. Folglich ist p^3 das Minimalpolynom von B. Außerdem gilt für

$$z = z(k) = \frac{1}{\deg p}\left[\operatorname{Rg}\left(p^{k-1}(B)\right) + \operatorname{Rg}\left(p^{k+1}(B)\right) - 2 \cdot \operatorname{Rg}\left(p^k(B)\right)\right],$$

dass

k	$\operatorname{Rg}\left(p^k(B)\right)$	z
0	7	–
1	4	$7+1-2\cdot 4 = 0$
2	1	$4+0-2\cdot 1 = 2$
3	0	$1+0-2\cdot 0 = 1$
≥ 4	0	$0+0-2\cdot 0 = 0$

Die kanonische rationale Form von B ist also

$$A = \begin{pmatrix} 0 & 0 & 6^3 & & & & \\ 1 & 0 & -3\cdot 6^2 & & & & \\ 0 & 1 & 3\cdot 6 & & & & \\ & & & 0 & -6^2 & & \\ & & & 1 & 2\cdot 6 & & \\ & & & & & 0 & -6^2 \\ & & & & & 1 & 2\cdot 6 \end{pmatrix}.$$

12.18 Bemerkung:

(i) Im obigen Beispiel ist die Faktorisierung des charakteristischen Polynoms in irreduzible Faktoren einfach. Im allgemeinen liegt da aber eine große Schwierigkeit, obwohl nach Satz 11.13 diese Faktorisierung immer möglich ist. Ein Patentrezept, wie man das macht, kann ich Ihnen nicht anbieten.

(ii) Noch in einer zweiten Hinsicht ist das Beispiel nicht typisch: Im allgemeinen enthält das charakteristische Polynom natürlich mehrere irreduzible Faktoren p_1, $p_2, \dots$. Man muss dann für jedes j die Ränge von $p_j^1(B), p_j^2(B), \dots$ berechnen, und zwar solange, bis zum ersten Mal $\mathrm{Rg}\left(p_j^k(B)\right) = \mathrm{Rg}\left(p_j^{k+1}(B)\right)$ ist. Für größere Potenzen ändert sich der Rang nicht mehr. Dieses k ist dann der Exponent von p_j im Minimalpolynom.

(iii) Mit der im Beispiel gezeigten Methode findet man zu B die kanonische rationale Form $A = T^{-1} \cdot B \cdot T$. Wenn auch eine Matrix T mit dieser Eigenschaft gesucht wird, hilft ein genauer Blick auf den Beweis von 12.9. Wir zeigen das Vorgehen am gleichen Beispiel:

12.19 Beispiel: *Fortsetzung*

Es ist

$$p(B)^2 = \begin{pmatrix} -96 & -24 & -24 & -96 & & & \\ 264 & 66 & 66 & 264 & & & \\ 120 & 30 & 30 & 120 & & & \\ & & & 0 & 0 & & \\ & & & 0 & 0 & & \\ & & & & & 0 & 0 \\ & & & & & 0 & 0 \end{pmatrix}.$$

Weil die ersten vier Spalten nicht 0 sind, kann man als einen ersten Basisvektor z.B. jeden der Vektoren $e_1, \dots, e_4$ nehmen. Wir entscheiden uns für $b_1 = e_2$. Dann liegen $b_2 = B \cdot b_1$ (d.h. die zweite Spalte von B) und $b_3 = B^2 \cdot b_1$ (d.h. die zweite Spalte von B^2) fest.
In der Matrix sind die ersten drei Einträge in der zweiten Spalte nicht 0 ; also kann man für λ z.B. die Projektion auf eine dieser Komponenten nehmen. Wir entscheiden uns für die dritte Komponente. Den Schnitt

$$\bigcap_{i=0}^{\infty} \mathrm{Ker}\left(\lambda \cdot B^i\right) \;=\; \bigcap_{i=0}^{2} \mathrm{Ker}\left(\lambda \cdot B^i\right)$$

erhält man dann als Lösungsmenge des homogenen linearen Gleichungssystems, dessen Koeffizientenmatrix aus den dritten Zeilen von B^0, B^1 und B^2 besteht, also

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 6 & -8 & -2 & 2 & 0 & 0 \\ 108 & 102 & -102 & 96 & 24 & 0 & 0 \end{pmatrix} \cdot x = 0 \,.$$

Eine Basis des Lösungsraumes sind die Vektoren

$$b_4 = \begin{pmatrix} 26 \\ -4 \\ 0 \\ -25 \\ 0 \\ 0 \\ 0 \end{pmatrix}, b_5 = \begin{pmatrix} -2 \\ 8 \\ 0 \\ 0 \\ -25 \\ 0 \\ 0 \end{pmatrix}, b_6 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, b_7 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

Zusammen mit b_1, b_2, b_3 bilden sie die Spalten einer Matrix

$$T_1 = \begin{pmatrix} 0 & -6 & -96 & 26 & -2 & 0 & 0 \\ 1 & 24 & 318 & -4 & 8 & 0 & 0 \\ 0 & 6 & 102 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -25 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Damit ist

$$B_1 := T_1^{-1} \cdot B \cdot T_1 = \begin{pmatrix} 0 & 0 & 216 & & & & \\ 1 & 0 & -108 & & & & \\ 0 & 1 & 18 & & & & \\ & & & -3 & 3 & 0 & 0 \\ & & & -27 & 15 & 0 & 0 \\ & & & 600 & -300 & -12 & -12 \\ & & & -675 & 375 & 27 & 24 \end{pmatrix} = \begin{pmatrix} A_1 & \\ & C_1 \end{pmatrix} ,$$

d.h. die Matrix B_1 enthält schon den richtigen 3×3–Block links oben auf der Diagonalen. Um C_1 zu verändern, wird das Verfahren für den Unterraum $< b_4, b_5, b_6, b_7 >$ wiederholt: Weil $(C_1 - 6) \cdot b_4 \neq 0$, aber $(C_1 - 6)^2 = 0$, kann man $c_4 = b_4$ wählen und hat dann $c_5 = C_1 \cdot c_4$, d.h die erste Spalte von C_1. Wählt man für λ die Projektion auf b_4, so ist die Koeffizientenmatrix des zugehörigen homogenen Systems gerade

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

(erste Zeilen von C_1^0 und C_1^1). Der Lösungsraum hat offenbar Basis $\{b_6, b_7\}$. Entsprechend findet man als nächste Transformationsmatrix

$$T_2 = \begin{pmatrix} 1 & -3 & 0 & 0 \\ 0 & -27 & 0 & 0 \\ 0 & 600 & 1 & 0 \\ 0 & -675 & 0 & 1 \end{pmatrix}$$

und damit

$$C_2 := T_2^{-1} \cdot C_1 \cdot T_2 \begin{pmatrix} 0 & -36 & 0 & 0 \\ 1 & 12 & 0 & 0 \\ 0 & 0 & -12 & -12 \\ 0 & 0 & 27 & 24 \end{pmatrix} .$$

Jetzt muss nur noch der 2×2–Block rechts unten auf die richtige Form gebracht werden. Dies ist offenbar mit $T_3 = \begin{pmatrix} 1 & -12 \\ 0 & 27 \end{pmatrix}$ zu erreichen.
Insgesamt findet man als Matrix des Basiswechsels

$$T = T_1 \cdot \begin{pmatrix} E_3 & \\ & T_2 \end{pmatrix} \cdot \begin{pmatrix} E_5 & \\ & T_3 \end{pmatrix} = \begin{pmatrix} 0 & -6 & -96 & 26 & -24 & 0 & 0 \\ 1 & 24 & 318 & -4 & -204 & 0 & 0 \\ 0 & 6 & 102 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -25 & 75 & 0 & 0 \\ 0 & 0 & 0 & 0 & 675 & 0 & 0 \\ 0 & 0 & 0 & 0 & 600 & 1 & -12 \\ 0 & 0 & 0 & 0 & -675 & 0 & 27 \end{pmatrix}$$

und kann $T^{-1} \cdot B \cdot T = A$ (oder einfacher $B \cdot T = T \cdot A$ und T regulär) kontrollieren.

13 Normalformen von Matrizen II: Die Jordan'sche Form

Sei weiterhin V ein n–dimensionaler VR über dem Körper $\mathbb{K}$ und α ein Endomorphismus von V mit Minimalpolynom m. Wir setzen voraus, dass alle irreduziblen Faktoren des charakteristischen Polynoms von α linear sind.

13.1 Bemerkung:

Diese Voraussetzung ist für jedes α erfüllt, wenn $\mathbb{K}$ algebraisch abgeschlossen ist, denn wenn $p \in \mathbb{K}[x]$ nicht konstant ist, dann hat p nach Voraussetzung eine Nullstelle, etwa λ. Aber dann ist $x - \lambda$ ein Teiler von p. Insbesondere gilt: $p = x - \lambda$ ist linear, wenn p irreduzibel und normiert ist.

13.2 Satz/Definition: *Jordan'sche Normalform*

Bzüglich einer geeigneten Basis von V hat die Matrix A von α Jordan'sche Normalform, d.h.

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_t \end{pmatrix}$$

mit „Jordan–Kästchen" A_i der Form

$$A_i = \begin{pmatrix} \lambda_i & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda_i \end{pmatrix}.$$

Bis auf die Reihenfolge der A_i ist die Jordan'sche Normalform durch α eindeutig bestimmt.

Beweis: Existenz einer geeigneten Basis: Nach 12.12 ist $V = Z_1 \oplus \cdots \oplus Z_t$ mit zyklischen invarianten Unterräumen, deren Minimalpolynome m_i Potenzen von irreduziblen Teilern (nach 12.15) des charakteristischen Polynoms von α sind. Es genügt, für jedes Z_i eine Basis anzugeben derart, dass die Matrix von $\alpha_{|Z_i}$ ein Jordan-Kästchen ist. Also darf man annehmen, dass $V = Z_1$ zyklisch ist und dass $m = (x - \lambda)^n$ für geeignetes $\lambda \in \mathbb{K}$ gilt. Daher gibt es eine Basis $v,\ \alpha(v), \ldots, \alpha^{n-1}(v)$. Dann ist leicht zu sehen, dass auch $b_1 = v,\ b_2 = (\alpha - \lambda)(v), \ldots, b_n = (\alpha - \lambda)^{n-1}(v)$ eine Basis ist. Aus $(\alpha - \lambda)(b_k) = b_{k+1}$ (für $k < n$) bzw. $(\alpha - \lambda)(b_n) = (\alpha - \lambda)^n(v) = 0$ folgt $\alpha(b_k) = b_{k+1} + \lambda b_k$ (für $k < n$) bzw. $\alpha(b_n) = \lambda b_n$. Bezüglich dieser Basis hat α also die angegebene Form.
Die Eindeutigkeit ergibt sich aus 13.3 unten.

13.3 Bemerkung:

Für jeden Teiler $x - \lambda$ des charakteristischen Polynoms und für jedes s ist die Anzahl $z = z(\lambda, s)$ der Jordan–Kästchen

$$\begin{pmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix}$$

der Größe $s \times s$ gegeben durch

$$z = \operatorname{Rg}(\alpha - \lambda)^{s-1} + \operatorname{Rg}(\alpha - \lambda)^{s+1} - 2 \cdot \operatorname{Rg}(\alpha - \lambda)^{s} \ .$$

Beweis: Das ist genau die Formel aus 12.16, da $\deg p = 1$.

13.4 Bemerkung:

(i) Um die Jordan'sche Normalform zu finden, muß man die Ränge von $(\alpha - \lambda)^s$ für jeden Eigenwert λ von α und für jedes s berechnen.

(ii) Mit 13.3 findet man zwar ziemlich leicht die Jordan'sche Normalform (falls man die Nullstellen des charakteristischen Polynoms findet), aber man hat damit noch nicht eine geeignete Basis (bzw. Transformationsmatrix) gefunden. Dazu sind im wesentlichen die Schritte wie in 12.19 durchzuführen.

(iii) Wenn die kanonische rationale Form schon gegeben ist, dann ist eine Transformationsmatrix leicht zu finden. Es genügt, dies blockweise zu tun. Man muss dann also die Matrix des Basiswechsels von $\{\alpha^j(v) \mid j = 0, \ldots, s-1\}$ zur Basis $\{(\alpha-\lambda)^j(v) \mid j = 0, \ldots, s-1\}$ angeben, d.h. die neuen Basisvektoren durch die alten ausdrücken; dabei hilft der Binomische Lehrsatz:

$$(\alpha - \lambda)^j(v) = \sum_{\mu=0}^{j} \binom{j}{\mu} (-\lambda)^{j-\mu} \alpha^{\mu}(v).$$

13.5 Beispiel:

Ein 4×4–Matrix mit Minimalpolynom $(x - \lambda)^4$ hat die kanonische rationale Form

$$R = \begin{pmatrix} 0 & 0 & 0 & -\lambda^4 \\ 1 & 0 & 0 & 4\lambda^3 \\ 0 & 1 & 0 & -6\lambda^2 \\ 0 & 0 & 1 & 4\lambda \end{pmatrix}$$

und die Jordan'sche Normalform

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

Die Transformationsmatrix lautet

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

(Probe: $TJ = RT$)

14 Näherungslösungen reeller linearer Gleichungssysteme

14.1 Bemerkung:

Gegeben sei das reelle lineare Gleichungssystem $Ax = b$ mit einer $n \times m$-Matrix A und $b \in \mathbb{R}^n$. Auch wenn dieses System unlösbar ist, kann man nach Näherungslösungen fragen, d.h. nach solchen $x \in \mathbb{R}^m$, für die Ax 'dicht' an b liegt. Damit soll gemeint sein, dass der Abstand $||Ax - b||$ möglichst klein ist. Das führt zu:

14.2 Definition: *Beste Näherung*

Sei A eine reelle $n \times m$-Matrix und $b \in \mathbb{R}^n$. Man nennt $u \in \mathbb{R}^m$ eine beste Näherungslösung des linearen Gleichungssystems $Ax = b$, wenn $||Au - b|| \leq ||Av - b||$ für alle $v \in \mathbb{R}^m$ ist.

14.3 Beispiel:

Sei

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} , \qquad b = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} .$$

Wählt man $v = (-3, 3)$, so erhält man

$$Av - b = \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix} ,$$

also $||Av - b|| = \sqrt{4 + 1 + 4} = 3$. Wählt man dagegen $u = (2/3, 1/3)$, so erhält man

$$Au - b = \begin{pmatrix} 1/3 \\ -2/3 \\ 1/3 \end{pmatrix} ,$$

also $||Au - b|| = \sqrt{1/9 + 4/9 + 1/9} = 1/3\sqrt{6} < 1$. Daher ist v sicher keine beste Näherungslösung. Dagegen wird sich zeigen, dass u eine beste Näherungslösung ist (und sogar die einzige solche).

14.4 Bemerkung:

Wenn das Gleichungssystem lösbar ist, dann sind natürlich die Lösungen u die besten Näherungslösungen, denn für diese und keine anderen Vektoren gilt $||Au - b|| = 0$.

14.5 Lemma:

Sei A eine reelle $n \times m$-Matrix. Die Unterräume $\mathrm{Im}\,(A)$ und $\mathrm{Ker}\,(A^t)$ von $\mathbb{R}^n$ enthalten nur den Nullvektor gemeinsam.

Beweis: Sei $v = Au \in \mathrm{Ker}\,(A^t) \cap \mathrm{Im}\,(A)$. Dann ist $0 = A^t v = A^t \cdot Au$. Erst recht ist $0 = u^t A^t \cdot Au = (Au)^t \cdot Au = v^t v$. Da das Skalarprodukt auf $\mathbb{R}^n$ positiv definit ist, folgt $v = 0$ wie behauptet.

14.6 Satz:

Sei A eine reelle $n \times m$-Matrix und $b \in \mathbb{R}^n$. Dann gilt:

(i) Das Gleichungssystem $A^t \cdot Ax = A^t b$ ist immer lösbar.

(ii) Die Lösungen dieses Gleichungssystems sind genau die besten Näherungslösungen von $Ax = b$.

Beweis:

Behauptung 1: $\operatorname{Im}(A^t \cdot A) = \operatorname{Im}(A^t)$.
Bew.: Sei $u \in \operatorname{Ker}(A^t \cdot A)$ und $v = Au$. Dann ist $v \in \operatorname{Im}(A) \cap \operatorname{Ker}(A^t) = 0$ (nach dem Lemma), also $u \in \operatorname{Ker}(A)$. Wir haben gezeigt, dass $\operatorname{Ker}(A^t \cdot A) \leq \operatorname{Ker}(A)$. Daher ist $\dim \operatorname{Ker}(A^t \cdot A) \leq \dim \operatorname{Ker}(A)$. Aus 5.13 folgt

$$\dim \operatorname{Im}\left(A^t \cdot A\right) \geq \dim \operatorname{Im}(A) = \operatorname{Rg}(A) = \operatorname{Rg}\left(A^t\right) = \dim \operatorname{Im}\left(A^t\right)$$

(vergleiche 7.21). Da offensichtlich $\operatorname{Im}(A^t \cdot A) \leq \operatorname{Im}(A^t)$, folgt die Gleichheit aus 4.20.
Behauptung 2: Das Gleichungssystem $A^t \cdot Ax = A^t b$ ist lösbar.
Bew.: Da $\operatorname{Im}(A^t) = \operatorname{Im}(A^t \cdot A)$ nach (1) und da $A^t b \in \operatorname{Im}(A^t)$, gibt es u mit $(A^t \cdot A)u = A^t b$. Das ist die Behauptung.
Behauptung 3: Jede Lösung von $A^t \cdot Ax = A^t b$ ist eine beste Näherungslösung von $Ax = b$.
Bew.: Wir müssen zeigen, dass $||Au - b|| \leq ||Av - b||$ für alle $v \in \mathbb{R}^m$, wenn

$$(*) \qquad A^t \cdot Au = A^t b \; .$$

Es genügt offenbar, $||Av - b||^2 - ||Au - b||^2 \geq 0$ zu zeigen. Dazu schreiben wir $v = u + w$, also $Av - b = Au - b + Aw$. Dann ist

$$\begin{aligned} ||Av - b||^2 - ||Au - b||^2 &= (Au - b + Aw)^t (Au - b + Aw) - (Au - b)^t (Au - b) \\ &= 2(Aw)^t (Au - b) + ||Aw||^2 \\ &= 2w^t (A^t \cdot Au - A^t b) + ||Aw||^2 \\ &= ||Aw||^2 \\ &\geq 0 \; . \end{aligned}$$

Behauptung 4: Jede beste Näherungslösung von $Ax = b$ ist eine Lösung von $A^t \cdot Ax = A^t b$.
Bew.: Sei v eine beste Näherungslösung von $Ax = b$. Nach (2) existiert eine Lösung u von $A^t \cdot Ax = A^t b$ und diese ist nach (3) ebenfalls eine beste Näherungslösung von $Ax = b$; daher ist $||Au - b|| = ||Av - b||$. Schreibt man wieder $v = u + w$, dann ist also $0 = ||Av - b||^2 - ||Au - b||^2 = ||Aw||^2$, wie gerade im Beweis von (3) berechnet. Das geht aber nur, wenn $Aw = 0$, also erst recht $A^t \cdot Aw = 0$. Folglich ist $A^t \cdot Av = A^t \cdot Au + A^t \cdot Aw = A^t \cdot Au = A^t b$, also v eine Lösung von $A^t \cdot Ax = A^t b$.

14.7 Bemerkung:

Da $A^t \cdot A$ eine quadratische und sogar symmetrische Matrix ist, darf man sich also bei der Lösung reeller linearer Gleichungssysteme auf solche mit quadratischer und symmetrischer Koeffizienten-Matrix S beschränken! Für solche gibt es nach dem Hauptachsen-Theorem eine orthogonale Matrix T derart, dass $T^t \cdot S \cdot T = D$ eine Diagonalmatrix ist. Durch Multiplikation mit T^t wird aus $Sx = b$ das äquivalente Gleichungssystem $T^t b = T^t \cdot Sx = T^t \cdot S \cdot T \cdot T^t x = D \cdot T^t x$, welches sehr leicht zu lösen ist, da D diagonal ist.

14.8 Beispiel:

Bevor ein Bäcker in den Urlaub fährt, möchte er seine Vorräte möglichst verbrauchen. Eine Inventur ergibt die folgenden Bestände (Wert in Euro)

Mehl	Eier	Butter	Milch	Gewürze
120	170	250	63	30

,

aus denen drei verschiedene Kuchensorten K_1, K_2, K_3 hergestellt werden sollen. Den Wert der Zutaten für diese Kuchen kann man der folgenden Tabelle entnehmen:

	K_1	K_2	K_3
Mehl	0.5	0.5	0.6
Eier	0.6	0.9	0.9
Butter	1.0	1.5	1.0
Milch	0.3	0.2	0.3
Gewürze	0.1	0.1	0.2

Wenn x_i die Anzahl der Kuchen vom Typ $i = 1, 2, 3$ ist, dann ergibt sich daraus ein lineares Gleichungssystem, nämlich

$$\begin{pmatrix} 0.5 & 0.5 & 0.6 \\ 0.6 & 0.9 & 0.9 \\ 1.0 & 1.5 & 1.0 \\ 0.3 & 0.2 & 0.3 \\ 0.1 & 0.1 & 0.2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 120 \\ 170 \\ 250 \\ 63 \\ 30 \end{pmatrix}$$

Dieses Gleichungssystem ist nicht lösbar, da der Rang der Koeffizienten-Matrix 3, der Rang der erweiterten Matrix aber 4 ist. Also multipliziert man mit der Transponierten der Koeffizienten-Matrix und erhält

$$\begin{pmatrix} 1.71 & 2.36 & 1.95 \\ 2.36 & 3.36 & 2.69 \\ 1.95 & 2.69 & 2.30 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0.5 & 0.6 & 1.0 & 0.3 & 0.1 \\ 0.5 & 0.9 & 1.5 & 0.2 & 0.1 \\ 0.6 & 0.9 & 1.0 & 0.3 & 0.2 \end{pmatrix} \begin{pmatrix} 0.5 & 0.5 & 0.6 \\ 0.6 & 0.9 & 0.9 \\ 1.0 & 1.5 & 1.0 \\ 0.3 & 0.2 & 0.3 \\ 0.1 & 0.1 & 0.2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5 & 0.6 & 1.0 & 0.3 & 0.1 \\ 0.5 & 0.9 & 1.5 & 0.2 & 0.1 \\ 0.6 & 0.9 & 1.0 & 0.3 & 0.2 \end{pmatrix} \begin{pmatrix} 120 \\ 170 \\ 250 \\ 63 \\ 30 \end{pmatrix} = \begin{pmatrix} 433.9 \\ 603.6 \\ 499.9 \end{pmatrix}$$

Die Determinante der (quadratischen, symmetrischen) Koeffizienten-Matrix ist nicht 0, also hat dieses Gleichungssystem genau eine Lösung, die dann die einzige beste Näherung für das ursprüngliche Problem ist. Gerundet ergibt sich $(x_1, x_2, x_3) = (111.6, 47.1, 67.6)$. Jedenfalls kann der Bäcker also 111 Kuchen der ersten, 47 Kuchen der zweiten, und 67 Kuchen der dritten Sorte backen. Wie man kontrolliert, reichen die restlichen Zutaten

nicht mehr für einen weiteren Kuchen nach einem der drei Rezepte.
Es sollte angemerkt werden, dass wir (oder der Bäcker) hier Glück gehabt haben, weil alle Komponenten der besten Näherungslösung positiv sind. Was könnte der Bäcker mit der Information anfangen, dass er am besten -10 Kuchen der Sorte K_2 backen sollte? Die Bedingung $x_i \geq 0$ ergibt sich aus der Problemstellung, wir haben sie aber bei Berechnung nicht berücksichtigt und eben nur Glück, dass sie von der Lösung erfüllt wird. Der richtige Ansatz ist $Ax \leq b,\ x \geq 0$ und x ganzzahlig. Das (optimale) Lösen von linearen Ungleichungssystemen ist Gegenstand der 'Linearen Optimierung', auf die in dieser Vorlesung nicht eingegangen wird.

15 Multilineare Abbildungen und Tensorprodukt

15.1 Definition: *n–fach linear*

Seien $V_1, V_2, \ldots, V_n$ und W alles Vektorräume über demselben Körper $\mathbb{K}$. Eine Abbildung $\varphi : V_1 \times V_2 \times \cdots \times V_n \to W$ heißt n–fach linear, wenn sie linear in jeder Komponente ist, d.h. wenn für jedes $i = 1, \ldots, n$, jedes $k \in \mathbb{K}$ und alle Vektoren $v_j \in V_j$, $j \neq i$, $v_i, v_i' \in V_i$ gilt

$$\varphi(v_1, \ldots, v_{i-1}, v_i + kv_i', v_{i+1}, \ldots, v_n) \\ = \varphi(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + k\varphi(v_1, \ldots, v_{i-1}, v_i', v_{i+1}, \ldots, v_n).$$

15.2 Beispiel:

(i) Die Determinantenabbildung det ist n–fach linear (für $V_1 = V_2 = \cdots = V_n = \mathbb{K}^n$, $W = \mathbb{K}$).

(ii) Bilinearformen sind 2–fach linear (für $V_1 = V_2$, $W = \mathbb{K}$).

(iii) Wenn $\varphi : V_1 \times V_2 \times \cdots \times V_n \to W$ n–fach linear und $\alpha : W \to U$ linear, dann ist $\alpha\varphi : V_1 \times V_2 \times \cdots \times V_n \to U$ n–fach linear.

15.3 Bemerkung:

Wenn φ und ψ wie in 15.1 sind, dann ist auch $\varphi + \psi$, definiert durch

$$(\varphi + \psi)(v_1, \ldots, v_n) = \varphi(v_1, \ldots, v_n) + \psi(v_1, \ldots, v_n),$$

eine n–fach lineare Abbildung $V_1 \times \cdots \times V_n \to W$. Ebenso ist für $k \in \mathbb{K}$ durch

$$(k\varphi)(v_1, \ldots, v_n) = k\varphi(v_1, \ldots, v_n)$$

eine n–fach lineare Abbildung definiert. Mit dieser Addition und skalaren Multiplikation bilden die n–fach linearen Abbildungen von $V_1 \times \cdots \times V_n$ in W einen Vektorraum, den wir mit $\mathcal{L}(V_1, \ldots, V_n; W)$ bezeichnen.

15.4 Satz:

Für jedes i ist die Abbildung

$$\sigma : \mathrm{Hom}\,(V_i, \mathcal{L}(V_1, \ldots, V_{i-1}, V_{i+1}, \ldots, V_n; W)) = \mathrm{H} \to \mathcal{L}(V_1, \ldots, V_n; W),$$

welche für $\alpha \in \mathrm{H}$ durch

$$\alpha^\sigma(v_1, \ldots, v_n) = \alpha(v_i)(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$$

definiert wird, ein Isomorphismus.

Beweis: Sicher ist α^σ eine Abbildung $V_1 \times \cdots \times V_n \to W$.

α^σ ist n–fach linear: In der j–ten Komponente ($j \neq i$), weil $\alpha(v_i) \in \mathcal{L}(V_1, \ldots, V_{i-1}, V_{i+1}, \ldots, V_n; W)$; in der i–ten Komponente, weil α linear und nach Definition der Vektorraum–Struktur auf $\mathcal{L}$.

Also ist σ wirklich eine Abbildung $\mathrm{H} \to \mathcal{L}(V_1, \ldots, V_n; W)$. Nach Definition der Vektorraum–Strukturen auf $\mathcal{L}$ und H ist σ linear.

σ ist bijektiv: Für $\varphi \in \mathcal{L}(V_1, \ldots, V_n; W)$ sei

$$\varphi^\tau : V_i \to \mathcal{L}(V_1, \ldots, V_{i-1}, V_{i+1}, \ldots, V_n; W)$$

definiert durch

$$\varphi^\tau(v_i)(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) = \varphi(v_1, \ldots, v_n).$$

Dann ist $\varphi^\tau \in \mathrm{H}$ und

$$\begin{aligned} \alpha^{\sigma\tau}(v_i)(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) &= \alpha^\sigma(v_1, \ldots, v_n) \\ &= \alpha(v_i)(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n), \end{aligned}$$

also $\alpha^{\sigma\tau} = \alpha$, $\sigma\tau = \mathrm{id}$, und

$$\varphi^{\tau\sigma}(v_1, \ldots, v_n) = \varphi^\tau(v_i)(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) = \varphi(v_1, \ldots, v_n),$$

d.h. $\varphi^{\tau\sigma} = \varphi$, $\tau\sigma = \mathrm{id}$.

15.5 Korollar:

Wenn die Vektorräume V_i, W alle endlich–dimensional sind, dann gilt

$$\dim \mathcal{L}(V_1, \ldots, V_n; W) = \dim V_1 \cdot \dim V_2 \cdot \ldots \cdot \dim V_n \cdot \dim W.$$

Beweis: durch Induktion über n.

Für $n = 1$ hat der VR $\mathcal{L}(V_1; W) = \mathrm{Hom}\,(V_1, W)$ die Dimension $\dim V_1 \cdot \dim W$. Für $n > 1$ ist nach 15.4 und Induktionsvoraussetzung

$$\begin{aligned} \dim \mathcal{L}(V_1, \ldots, V_n; W) &= \dim \mathrm{Hom}\,(V_1, \mathcal{L}(V_2, \ldots, V_n; W)) \\ &= \dim V_1 \cdot \dim \mathcal{L}(V_2, \ldots, V_n; W) \\ &= \dim V_1 \cdot \dim V_2 \cdot \ldots \cdot \dim V_n \cdot \dim W. \end{aligned}$$

15.6 Bemerkung:

(i) Wir wollen jetzt aus $V_1, \ldots, V_n$ einen neuen Vektorraum (genannt das Tensorprodukt, geschrieben $V_1 \otimes V_2 \otimes \cdots \otimes V_n$) konstruieren derart, daß für jedes W die n–fach linearen Abbildungen $V_1 \times V_2 \times \cdots \times V_n \to W$ bijektiv den linearen Abbildungen $V_1 \otimes V_2 \otimes \cdots \otimes V_n \to W$ entsprechen. Dazu folgende Vorbemerkung:

(ii) Sei M eine beliebige Menge, $\mathbb{K}$ ein Körper. Die Menge

$$F_M = \{f : M \to \mathbb{K} \mid f(m) = 0 \text{ für fast alle } m \in M\}$$

bildet mit der üblichen Addition und Multiplikation einen Vektorraum. Eine Basis von F_M ist $\{d_m \mid m \in M\}$, wobei $d_m : M \to \mathbb{K}$ definiert ist durch

$$d_m(x) = \begin{cases} 1 & \text{falls } x = m \\ 0 & \text{sonst.} \end{cases}$$

Die Basis ist also in Bijektion zu M. Schreibt man $[m]$ für d_m, dann besteht also F_M aus allen (endlichen) Linearkombinationen

$$\sum_{m \in M} k_m [m]$$

mit $k_m \in \mathbb{K}$.

15.7 Definition: *Tensorprodukt*

Seien $V_1, \dots, V_n$ gegeben. Sei U der Unterraum von $F_{V_1 \times \dots \times V_n} = F$, welcher von allen Elementen der Form

$$[(v_1, \dots, v_{i-1}, v_i + v_i', v_{i+1}, \dots, v_n)] \\ - [(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n)] - [(v_1, \dots, v_{i-1}, v_i', v_{i+1}, \dots, v_n)]$$

und

$$[(v_1, \dots, v_{i-1}, kv_i, v_{i+1}, \dots, v_n)] - k[(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n)]$$

erzeugt wird. Der Faktorraum F/U wird das Tensorprodukt von $V_1, \dots, V_n$ genannt und als

$$F/U = V_1 \otimes \dots \otimes V_n = \bigotimes_{i=1}^{n} V_i$$

geschrieben. Wenn $(v_1, \dots, v_n) \in V_1 \times \dots \times V_n$, dann ist $[(v_1, \dots, v_n)] \in F$, also

$$T(v_1, \dots, v_n) := [(v_1, \dots, v_n)] + U \in F/U = V_1 \otimes \dots \otimes V_n.$$

Zur Abkürzung schreibt man $T(v_1, \dots, v_n) = v_1 \otimes \dots \otimes v_n$. Also ist T eine Abbildung von $V_1 \times \dots \times V_n$ in $V_1 \otimes \dots \otimes V_n$.

15.8 Lemma:

T ist n–fach linear.

Beweis: Es ist

$$\begin{aligned} T(\dots, v_i + v_i', \dots) &= [(\dots, v_i + v_i', \dots)] + U \\ &= [(\dots, v_i, \dots)] + [(\dots, v_i', \dots)] + [(\dots, v_i + v_i', \dots)] \\ &\quad - [(\dots, v_i, \dots)] - [(\dots, v_i', \dots)] + U \\ &= [(\dots, v_i, \dots)] + [(\dots, v_i', \dots)] + U, \\ &\quad \text{weil } [(\dots, v_i + v_i', \dots)] - [(\dots, v_i, \dots)] - [(\dots, v_i', \dots)] \in U, \\ &= ([(\dots, v_i, \dots)] + U) + ([(\dots, v_i', \dots)] + U) \\ &= T(\dots, v_i, \dots) + T(\dots, v_i', \dots). \end{aligned}$$

Ebenso für Skalare.

15.9 Bemerkung:

(i) Jedes Element von $V_1 \otimes \cdots \otimes V_n$ ist eine endliche Summe von Elementen der Form $v_1 \otimes \cdots \otimes v_n$.

Beweis: Da die Elemente $[v_1, \ldots, v_n]$ eine Basis von F bilden, sind die $v_1 \otimes \cdots \otimes v_n = [v_1, \ldots, v_n] + U$ ein Erzeugendensystem von $V_1 \otimes \cdots \otimes V_n$. Jedes Element aus $V_1 \otimes \cdots \otimes V_n$ ist also eine endliche Linearkombination von Elementen dieses Typs. Wegen

$$k(v_1 \otimes \cdots \otimes v_n) = kT(v_1, \ldots, v_n) = T(kv_1, \ldots, v_n) = (kv_1) \otimes \cdots \otimes v_n$$

folgt die Behauptung.

(ii) Beim Rechnen in $V_1 \otimes \cdots \otimes V_n$ gilt

$$(v_1 + v_1') \otimes v_2 \otimes \cdots \otimes v_n = v_1 \otimes v_2 \otimes \cdots \otimes v_n + v_1' \otimes v_2 \otimes \cdots \otimes v_n$$

und

$$(kv_1) \otimes v_2 \otimes \cdots \otimes v_n = k(v_1 \otimes v_2 \otimes \cdots \otimes v_n).$$

Ebenso für alle anderen Komponenten. Insbesondere darf man Skalare schieben:

$$v_1 \otimes \cdots \otimes kv_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_n = v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes kv_j \otimes \cdots \otimes v_n$$

für alle i, j.

15.10 Satz:

Seien $V_1, \ldots, V_n$ und W Vektorräume über $\mathbb{K}$. Die Abbildung $\alpha \mapsto \alpha T$ ist ein Isomorphismus von $\mathrm{Hom}\,(V_1 \otimes \cdots \otimes V_n, W)$ auf $\mathcal{L}(V_1, \ldots, V_n; W)$.

Beweis: Da $T : V_1 \times \cdots \times V_n \to V_1 \otimes \cdots \otimes V_n$ n–fach linear ist (15.8) und $\alpha : V_1 \otimes \cdots \otimes V_n \to W$ linear, ist $\alpha T : V_1 \times \cdots \times V_n \to W$ n–fach linear nach 15.2, d.h. $\alpha T \in \mathcal{L}(V_1, \ldots, V_n; W)$.

Offenkundig ist $\alpha \mapsto \alpha T$ eine lineare Abbildung. Wenn $\alpha \neq 0$, dann gibt es ein Element $x \in V_1 \otimes \cdots \otimes V_n$ mit $\alpha(x) \neq 0$. Da nach 15.9 das Element x eine Summe von Elementen der Form $v_1 \otimes \cdots \otimes v_n$ ist, gibt es also ein solches Element $v_1 \otimes \cdots \otimes v_n$ mit $\alpha(v_1 \otimes \cdots \otimes v_n) = \alpha T(v_1, \ldots, v_n) \neq 0$. Also ist dann $\alpha T \neq 0$, d.h. $\alpha \mapsto \alpha T$ ist injektiv.

Die Abbildung ist auch surjektiv: Sei $\varphi \in \mathcal{L}(V_1, \ldots, V_n; W)$. (Gesucht ist $\alpha \in \mathrm{Hom}(V_1 \otimes \cdots \otimes V_n, W)$ mit $\alpha T = \varphi$.) Dann ist durch $\hat{\varphi}([(v_1, \ldots, v_n)]) = \varphi(v_1, \ldots, v_n)$ eine Abbbildung auf der Basis von F in W definiert. Diese läßt sich nach 5.16 eindeutig zu einer linearen Abbildung $\tilde{\varphi} : F \to W$ fortsetzen. Es ist

$$\begin{aligned}\tilde{\varphi}([(\ldots, v_i + v_i', \ldots)] - [(\ldots, v_i, \ldots)] - [(\ldots, v_i', \ldots)]) \\ &= \varphi(\ldots, v_i + v_i', \ldots) - \varphi(\ldots, v_i, \ldots) - \varphi(\ldots, v_i', \ldots), \\ &\quad \text{nach Definition von } \tilde{\varphi}, \\ &= 0, \text{ da } \varphi \text{ } n\text{–fach linear.}\end{aligned}$$

Ebenso für $[(\ldots, kv_i, \ldots)] - k[(\ldots, v_i, \ldots)]$. Also gilt $U \leq \mathrm{Ker}\,(\tilde{\varphi})$. Daher ist durch $\alpha(f + U) = \tilde{\varphi}(f)$ eine lineare Abbildung $\alpha : F/U = V_1 \otimes \cdots \otimes V_n \to W$ definiert, und es gilt

$$\alpha T(v_1, \ldots, v_n) = \alpha([(v_1, \ldots, v_n)] + U) = \tilde{\varphi}([(v_1, \ldots, v_n)]) = \varphi(v_1, \ldots, v_n),$$

d.h. $\alpha T = \varphi$.

15.11 Korollar:

Zu jedem Vektorraum W und jeder n–fach linearen Abbildung $\varphi : V_1 \times \cdots \times V_n \to W$ existiert genau eine lineare Abbildung $\alpha : V_1 \otimes \cdots \otimes V_n \to W$ mit $\varphi = \alpha T$.

Der nächste Satz zeigt, daß dies das Paar $(V_1 \otimes \cdots \otimes V_n, T)$ charakterisiert.

15.12 Satz:

Sei U ein Vektorraum und $S : V_1 \times \cdots \times V_n \to U$ eine n–fach lineare Abbildung mit der Eigenschaft:

> Zu jedem Vektorraum W und zu jeder n–fach linearen Abbildung $\varphi : V_1 \times \cdots \times V_n \to W$ existiert genau eine lineare Abbildung $\alpha : U \to W$ mit $\varphi = \alpha S$.

Dann gilt: Es gibt genau einen Homomorphismus $\sigma : V_1 \otimes \cdots \otimes V_n \to U$ mit $S = \sigma T$, und dieses σ ist ein Isomorphismus.

Beweis: Zunächst sei $W = V_1 \otimes \cdots \otimes V_n$ und $\varphi = T : V_1 \times \cdots \times V_n \to V_1 \otimes \cdots \otimes V_n$ gewählt (φ ist n–fach linear nach 15.8). Nach Voraussetzung gibt es dann genau eine lineare Abbildung $\tau : U \to V_1 \otimes \cdots \otimes V_n$ mit $T = \tau S$. Weil $S : V_1 \times \cdots \times V_n \to U$ n–fach linear ist, gibt es nach 15.11 (mit $W = U$, $\varphi = S$) genau eine lineare Abbildung $\sigma : V_1 \otimes \cdots \otimes V_n \to U$ mit $S = \sigma T$.

Dann ist $\mathrm{id}T = T = \tau S = \tau\sigma T$. Wegen 15.11 (Eindeutigkeit) ist also $\tau\sigma = \mathrm{id}_{V_1 \otimes \cdots \otimes V_n}$. Ebenso ist $\mathrm{id}S = S = \sigma T = \sigma\tau S$, also nach Voraussetzung an (U, S) (Eindeutigkeit) $\sigma\tau = \mathrm{id}_U$. Daher ist σ ein Isomorphismus (und $\tau = \sigma^{-1}$).

15.13 Satz:

Wenn $\dim V_i$ endlich für $i = 1, \ldots, n$, dann gilt:

$$\dim \left(\bigotimes_{i=1}^{n} V_i \right) = \prod_{i=1}^{n} \dim V_i.$$

Beweis: Es ist für den dualen Raum

$$\begin{aligned} \left(\bigotimes_{i=1}^{n} V_i \right)^* &= \mathrm{Hom}\,(V_1 \otimes \cdots \otimes V_n, \mathbb{K}) \\ &\cong \mathcal{L}(V_1, \ldots, V_n; \mathbb{K}), \text{ nach 15.10.} \end{aligned}$$

Nach 15.5 ist $\dim \mathcal{L}(V_1, \ldots, V_n; \mathbb{K}) = \prod_{i=1}^{n} \dim V_i$. Also ist dies die Dimension von $(\otimes_{i=1}^{n} V_i)^*$. Da $\dim(\otimes_{i=1}^{n} V_i)$ endlich, gilt nach 6.3

$$\dim \left(\bigotimes_{i=1}^{n} V_i \right) = \dim \left(\bigotimes_{i=1}^{n} V_i \right)^* = \prod_{i=1}^{n} \dim V_i.$$

15.14 Satz:

Das Tensorprodukt ist assoziativ, d.h. für Vektorräume V_1, V_2, V_3 gilt

$$(V_1 \otimes V_2) \otimes V_3 \cong V_1 \otimes V_2 \otimes V_3 \cong V_1 \otimes (V_2 \otimes V_3)$$

(kanonisch).

Beweis: Sei $S = V_1 \times V_2 \times V_3 \to (V_1 \otimes V_2) \otimes V_3$ definiert durch $S(v_1, v_2, v_3) = (v_1 \otimes v_2) \otimes v_3$. Dann ist S 3–fach linear (triviale Verifikation). Wenn W ein beliebiger Vektorraum ist, und $\varphi : V_1 \times V_2 \times V_3 \to W$ 3–fach linear, dann sei $\psi_{v_3} : V_1 \times V_2 \to W$ für festes $v_3 \in V_3$ definiert durch $\psi_{v_3}(v_1, v_2) = \varphi(v_1, v_2, v_3)$. Dann ist ψ_{v_3} bilinear, also gibt es ein lineares $\alpha_{v_3} : V_1 \otimes V_2 \to W$ mit $\alpha_{v_3}(v_1 \otimes v_2) = \varphi(v_1, v_2, v_3)$. Die Abbildung $(V_1 \otimes V_2) \times V_3 \to W$, welche durch $(x, v_3) \mapsto \alpha_{v_3}(x)$ für $x \in V_1 \otimes V_2$ definiert wird, ist bilinear, definiert also eine lineare Abbildung $\alpha : (V_1 \otimes V_2) \otimes V_3 \to W$ mit $\alpha((v_1 \otimes v_2) \otimes v_3) = \alpha_{v_3}(v_1 \otimes v_2) = \varphi(v_1, v_2, v_3)$. Also ist $\alpha S = \varphi$, und offenbar ist α die einzige lineare Abbildung mit dieser Eigenschaft. Wir haben gezeigt, dass S und $U = (V_1 \otimes V_2) \otimes V_3$ die Voraussetzungen von Satz 15.12 erfüllen. Daher gibt es (genau einen) Isomorphismus $\sigma : V_1 \otimes V_2 \otimes V_3 \to (V_1 \otimes V_2) \otimes V_3$ mit $S = \sigma T$, d.h.

$$\sigma(v_1 \otimes v_2 \otimes v_3) = \sigma T(v_1, v_2, v_3) = S(v_1, v_2, v_3) = (v_1 \otimes v_2) \otimes v_3.$$

Analog sieht man $V_1 \otimes V_2 \otimes V_3 \cong V_1 \otimes (V_2 \otimes V_3)$.

15.15 Lemma:

Seien $x_1, \ldots, x_n \in X$ und $y_1, \ldots, y_n \in Y$, die y_i linear unabhängig (X und Y VR'e über $\mathbb{K}$). Wenn $\sum_{i=1}^n x_i \otimes y_i = 0$, dann alle $x_i = 0$.

Beweis: Sei $\lambda : X \to \mathbb{K}$ eine beliebige lineare Abbildung, und sei $\varepsilon_i : Y \to \mathbb{K}$ eine lineare Abbildung mit $\varepsilon_i(y_i) = 1$, $\varepsilon_i(y_j) = 0$ für alle $j \neq i$. Dann ist $\beta_i(x, y) = \lambda(x)\varepsilon_i(y)$ eine bilineare Abbildung $X \times Y \to \mathbb{K}$. Also gibt es einen Homomorphismus $\alpha_i : X \otimes Y \to \mathbb{K}$ mit $\beta_i = \alpha_i T$, wobei $T : X \times Y \to X \otimes Y$ die kanonische Abbildung ist (d.h. $T(x, y) = x \otimes y$). Dann ist

$$\begin{aligned}
\lambda(x_i) &= \sum_{j=1}^n \lambda(x_j)\varepsilon_i(y_j) = \sum_{j=1}^n \beta_i(x_j, y_j) \\
&= \sum_{j=1}^n \alpha_i T(x_j, y_j) = \sum_{j=1}^n \alpha_i(x_j \otimes y_j) \\
&= \alpha_i \left(\sum_{j=1}^n x_j \otimes y_j \right) = 0,
\end{aligned}$$

da $\sum_{j=1}^n x_j \otimes y_j = 0$ nach Voraussetzung. Also gilt $\lambda(x_i) = 0$ für jedes $\lambda : X \to \mathbb{K}$. Daher ist $x_i = 0$ (für alle i).

15.16 Satz:

Seien V und W $\mathbb{K}$–Vektorräume. Wenn $\{a_i \mid i \in I\}$ eine Basis von V und $\{b_j \mid j \in J\}$ eine Basis von W, dann ist $\{a_i \otimes b_j \mid i \in I, j \in J\}$ eine Basis von $V \otimes W$.

Beweis: $\{a_i \otimes b_j \mid i \in I, j \in J\}$ ist linear unabhängig: Wenn

$$0 = \sum_{i,j} k_{ij}(a_i \otimes b_j) = \sum_j \left(\sum_i k_{ij} a_i \right) \otimes b_j,$$

dann ist $\sum_i k_{ij} a_i = 0$ für jedes i nach Lemma 15.15. Da die a_i's linear unabhängig sind, folgt $k_{ij} = 0$ für alle i, j.

$\{a_i \otimes b_j \mid i \in I, j \in J\}$ ist Erzeugenden–System: Wenn $v \in V$, etwa $v = \sum_{i \in I} k_i a_i$, und $w \in W$, etwa $w = \sum_{j \in J} l_j b_j$, dann ist

$$v \otimes w = \left(\sum_i k_i a_i \right) \otimes \left(\sum_j l_j b_j \right) = \sum_{i,j} k_i l_j (a_i \otimes b_j)$$

eine Linearkombination der $a_i \otimes b_j$'s. Nach 15.9 (i) folgt die Behauptung.

15.17 Bemerkung:

(i) Der vorige Satz läßt sich leicht auf Tensorprodukte mit mehr als zwei Faktoren verallgemeinern.

(ii) Man erhält dann einen anderen Beweis für 15.13.

15.18 Bemerkung:

Seien $X_1, \ldots, X_n$ und $Y_1, \ldots, Y_n$ $\mathbb{K}$–Vektorräume. Dann kann man die folgenden weiteren Räume konstruieren: $X_1 \otimes \cdots \otimes X_n$, $Y_1 \otimes \cdots \otimes Y_n$ und $\mathrm{Hom}(X_1 \otimes \cdots \otimes X_n, Y_1 \otimes \ldots \otimes Y_n) = H_1$. Ebenso kann man zuerst die Homomorphismenräume bilden: $\mathrm{Hom}(X_1, Y_1)$, $\ldots, \mathrm{Hom}(X_n, Y_n)$, und dann das Tensorprodukt $\mathrm{Hom}(X_1, Y_1) \otimes \cdots \otimes \mathrm{Hom}(X_n, Y_n) = H_2$. Was ist der Zusammenhang zwischen H_1 und H_2?

Sei $\alpha_i \in \mathrm{Hom}(X_i, Y_i)$, $i = 1, \ldots, n$, und sei $x_i \in X_i$. Durch

$$x_1 \otimes \cdots \otimes x_n \mapsto \alpha_1(x_1) \otimes \cdots \otimes \alpha_n(x_n)$$

ist eine lineare Abbildung

$$\varphi(\alpha_1, \ldots, \alpha_n) : X_1 \otimes \cdots \otimes X_n \to Y_1 \otimes \cdots \otimes Y_n$$

bestimmt. Also ist φ eine Abbildung von $\mathrm{Hom}(X_1, Y_1) \times \cdots \times \mathrm{Hom}(X_n, Y_n)$ in H_1. Man rechnet leicht nach, daß φ n–fach linear ist. Also existiert nach 15.11 genau eine lineare Abbildung

$$\gamma : H_2 = \mathrm{Hom}(X_1, Y_1) \otimes \cdots \otimes \mathrm{Hom}(X_n, Y_n) \to H_1$$

mit der Eigenschaft

$$\gamma(\alpha_1 \otimes \cdots \otimes \alpha_n)(x_1 \otimes \cdots \otimes x_n) = \alpha_1(x_1) \otimes \cdots \otimes \alpha_n(x_n).$$

15.19 Satz:

Die lineare Abbildung γ aus 15.18 ist injektiv.

Wenn die Räume X_i alle endliche Dimensionen haben, dann ist γ ein Isomorphismus.

Beweis: Jedes Element ξ aus $\mathrm{Hom}\,(X_1, Y_1) \otimes \cdots \otimes \mathrm{Hom}\,(X_n, Y_n)$ läßt sich darstellen als

$$\xi = \sum_{i=1}^{m} \omega_i \otimes \alpha_i$$

mit $\omega_i \in \mathrm{Hom}\,(X_1, Y_1) \otimes \cdots \otimes \mathrm{Hom}\,(X_{n-1}, Y_{n-1})$ und $\alpha_i \in \mathrm{Hom}\,(X_n, Y_n)$, $i = 1, \ldots, m$. Weiter kann man annehmen, daß $\omega_1, \ldots, \omega_n$ linear unabhängig sind, denn wenn etwa $\omega_m = \sum_{i=1}^{m-1} k_i \omega_i$, dann $\omega_m \otimes \alpha_m = \sum_{i=1}^{m-1} \omega_i \otimes k_i \alpha_m$, also

$$\sum_{i=1}^{m} \omega_i \otimes \alpha_i = \sum_{i=1}^{m-1} \omega_i \otimes \alpha_i + \sum_{i=1}^{m-1} \omega_i \otimes k_i \alpha_m = \sum_{i=1}^{m-1} \omega_i \otimes (\alpha_i + k_i \alpha_m).$$

Sei nun $\xi \in \mathrm{Ker}\,(\gamma)$. Wenn $\xi \neq 0$, dann o.B.d.A. $\alpha_1 \neq 0$. Sei $x_n \in X_n$ mit $\alpha_1(x_n) \neq 0$. Nach geeigneter Umnumerierung kann man annehmen, daß für ein k mit $1 \leq k \leq m$ die ersten k Vektoren $\alpha_1(x_n), \ldots, \alpha_k(x_n)$ linear unabhängig sind, während $\alpha_j(x_n)$ für $k < j \leq m$ eine Linearkombination dieser Vektoren ist, etwa $\alpha_j(x_n) = \sum_{i=1}^{k} c_{ji} \alpha_i(x_n)$ mit $c_{ji} \in \mathbb{K}$. Seien nun $x_1 \in X_1, \ldots, x_{n-1} \in X_{n-1}$ beliebig. Dann gilt

$$\begin{aligned}
0 &= \gamma(\xi)(x_1 \otimes \cdots \otimes x_n) \\
&= \textstyle\sum_{i=1}^{m} \gamma(\omega_i \otimes \alpha_i)(x_1 \otimes \cdots \otimes x_n) \\
&= \textstyle\sum_{i=1}^{m} \gamma'(\omega_i)(x_1 \otimes \cdots \otimes x_{n-1}) \otimes \alpha_i(x_n), \\
&\quad \text{wobei } \gamma' \text{ definiert ist wie } \gamma \text{ für } n-1 \text{ Vektorräume,} \\
&= \textstyle\sum_{i=1}^{k} \gamma'(\omega_i)(x_1 \otimes \cdots \otimes x_{n-1}) \otimes \alpha_i(x_n) + \sum_{j=k+1}^{m} \gamma'(\omega_j)(x_1 \otimes \cdots \otimes x_{n-1}) \otimes \sum_{i=1}^{k} c_{ji} \alpha_i(x_n) \\
&= \textstyle\sum_{i=1}^{k} \left(\gamma'(\omega_i) + \sum_{j=k+1}^{m} c_{ji} \gamma'(\omega_j) \right) (x_1 \otimes \cdots \otimes x_{n-1}) \otimes \alpha_i(x_n) \ .
\end{aligned}$$

Da die $\alpha_i(x_n)$'s linear unabhängig sind, folgt nach Lemma 15.15

$$\left(\gamma'(\omega_i) + \sum_{j=k+1}^{m} c_{ji} \gamma'(\omega_j) \right) (x_1 \otimes \cdots \otimes x_{n-1}) = 0$$

für $i = 1, \ldots, k$. Dies gilt für alle $x_1 \in X_1, \ldots, x_{n-1} \in X_{n-1}$. Also ist

$$0 = \gamma'(\omega_i) + \sum_{j=k+1}^{m} c_{ji} \gamma'(\omega_j) = \gamma' \left(\omega_i + \sum_{j=k+1}^{m} c_{ji} \omega_j \right).$$

Per Induktion über n ist γ' injektiv, d.h. $\omega_i + \sum_{j=k+1}^{m} c_{ji} \omega_j = 0$ für $i = 1, \ldots, k$. Dies widerspricht der linearen Unabhängigkeit der ω_i's, also ist $\xi = 0$ und daher γ injektiv.

Nun sei vorausgesetzt, dass alle X_i endliche Dimensionen haben. Für jedes i wählen wir Basen $B(X, i)$ und $B(Y, i)$ von X_i und Y_i. Nach 15.16 und 15.17 ist dann $B_X = \{x_1 \otimes \cdots \otimes x_n \mid x_i \in B(X, i))\}$ eine Basis von $X_1 \otimes \cdots \otimes X_n$, und analog findet man eine Basis $B_Y = \{y_1 \otimes \cdots \otimes y_n \mid y_i \in B(Y, i))\}$ von $Y_1 \otimes \cdots \otimes Y_n$. Sei nun $\alpha_{x,y}$ die lineare Abbildung, die einen festen Basisvektor $x \in B_X$ auf $y \in B_y$ abbildet, und alle anderen Elemente in

B_X auf 0. Es ist dann leicht zu sehen, dass $\alpha_{x,y} \in \operatorname{Im}(\gamma)$.
Wenn $\alpha : X_1 \otimes \cdots \otimes X_n \to Y_1 \otimes \cdots \otimes Y_n$ eine beliebige lineare Abbildung ist, dann lässt sich α als endliche Summe von Abbildungen schreiben, die jeweils alle bis auf einen Basisvektor von B_X auf 0 abbilden, da B_X endlich ist. Das Bild dieses Basisvektors ist dann eine (endliche!) Linearkombination von B_Y. Daher ist α eine Linearkombination der $\alpha_{x,y}$'s, also auch in $\operatorname{Im}(\gamma)$. Daher ist γ surjektiv.

15.20 Bemerkung:

Im Allgemeinen ist γ nicht surjektiv. Als Beispiel betrachte man zwei Vektorräume V_1 und V_2 mit $\dim V_1 = \dim V_2 = |\mathbb{N}|$, und $W_1 = W_2 = \mathbb{K}$. Sei $\{a_1, a_2, \dots\}$ eine Basis von V_1 und $\{b_1, b_2, \dots\}$ eine von V_2. Dann ist $\{a_i \otimes b_j \mid i, j = 1, 2, \dots\}$ eine Basis von $V_1 \otimes V_2$, und durch

$$\varepsilon(a_i \otimes b_j) = \begin{cases} 1 \otimes 1 & \text{wenn } i = j \\ 0 & \text{sonst} \end{cases}$$

wird eine lineare Abbildung $\varepsilon : V_1 \otimes V_2 \to W_1 \otimes W_2$ definiert. Es ist $\varepsilon \notin \operatorname{Im}(\gamma)$. Andernfalls gibt es $\alpha_1, \dots, \alpha_r : V_1 \to W_1$ und $\beta_1, \dots, \beta_r : V_2 \to W_2$ mit

$$\varepsilon = \gamma \left(\sum_{t=1}^{r} \alpha_t \otimes \beta_t \right).$$

Setzt man $A_i = (\alpha_1(a_i), \alpha_2(a_i), \dots, \alpha_r(a_i))$ und $B_i = (\beta_1(b_i), \beta_2(b_i), \dots, \beta_r(b_i))$, dann sind $A_i, B_j \in \mathbb{K}^r$ und

$$\begin{aligned} \varepsilon(a_i \otimes b_j) &= \gamma \left(\sum_{t=1}^{r} \alpha_t \otimes \beta_t \right) (a_i \otimes b_j) \\ &= \sum_{t=1}^{r} \alpha_t(a_i) \otimes \beta_t(b_j) \\ &= \sum_{t=1}^{r} \alpha_t(a_i)\beta_t(b_j) \otimes 1 \\ &= (A_i, B_j) \otimes 1, \end{aligned}$$

wobei $(\ ,\ ) : \mathbb{K}^r \times \mathbb{K}^r \to \mathbb{K}$ das übliche Skalarprodukt ist. Also ist $(A_i, B_j) = \delta_{ij}$. Aber alle Vektoren A_i sind aus $\mathbb{K}^r$; die unendlich vielen A_i's können also nicht linear unabhängig sein. Daher gibt es eine Linearkombination $A_n = \sum_{s=1}^{n-1} k_s A_s$. Es folgt der Widerspruch

$$1 = (A_n, B_n) = \sum_{s=1}^{n-1} k_s (A_s, B_n) = 0.$$

15.21 Bemerkung:

Seien wieder $\alpha_i : X_i \to Y_i$ lineare Abbildungen und

$$\alpha = \gamma(\alpha_1 \otimes \cdots \otimes \alpha_n) : X_1 \otimes \cdots \otimes X_n \to Y_1 \otimes \cdots \otimes Y_n$$

wie oben. Was ist der Kern von α? Er sollte etwas zu tun haben mit $\operatorname{Ker}(\alpha_i)$ für $i = 1, \dots, n$. Zur Beschreibung dient:

15.22 Bezeichnung:

Sei U_i ein Unterraum von X_i. Man setzt

$$[U_1, \ldots, U_n] = \langle x_1 \otimes \cdots \otimes x_n \in X_1 \otimes \cdots \otimes X_n \mid \exists\, i : x_i \in U_i \rangle.$$

Also ist $[U_1, \ldots, U_n]$ ein Unterraum von $X_1 \otimes \cdots \otimes X_n$.

15.23 Satz:

$\operatorname{Ker}(\alpha) = [\operatorname{Ker}(\alpha_1), \ldots, \operatorname{Ker}(\alpha_n)] =: K$

Beweis: Sei $x_1 \otimes \cdots \otimes x_n \in X_1 \otimes \cdots \otimes X_n$ und $x_i \in \operatorname{Ker}(\alpha_i)$ für ein festes i. Dann ist $\alpha(x_1 \otimes \cdots \otimes x_n) = \alpha_1(x_1) \otimes \cdots \otimes \alpha_n(x_n) = 0$, da $\alpha_i(x_i) = 0$. Also liegen die erzeugenden Elemente von K in $\operatorname{Ker}(\alpha)$ und daher ist $K \leq \operatorname{Ker}(\alpha)$. Man kann dann

$$\bar{\alpha} : (X_1 \otimes \cdots \otimes X_n)/K \to \operatorname{Im}(\alpha_1) \otimes \cdots \otimes \operatorname{Im}(\alpha_n)$$

durch

$$\bar{\alpha}(x_1 \otimes \cdots \otimes x_n + K) = \alpha_1(x_1) \otimes \cdots \otimes \alpha_n(x_n)$$

definieren.

Die andere Inklusion ist schwieriger: Sei $y_i \in \operatorname{Im}(\alpha_i)$ für $i = 1, \ldots, n$, etwa $y_i = \alpha_i(x_i)$. Wenn auch $y_i = \alpha_i(x_i')$, dann ist $x_i - x_i' \in \operatorname{Ker}(\alpha_i)$, also

$$\begin{aligned} x_1 \otimes \cdots \otimes x_n - x_1' \otimes \cdots \otimes x_n' = {} & (x_1 - x_1') \otimes x_2 \otimes \cdots \otimes x_n \\ & + x_1' \otimes (x_2 - x_2') \otimes x_3 \otimes \cdots \otimes x_n \\ & + x_1' \otimes x_2' \otimes (x_3 - x_3') \otimes x_4 \otimes \cdots \otimes x_n \\ & \cdots \\ & + x_1' \otimes x_2' \otimes \cdots \otimes x_{n-1}' \otimes (x_n - x_n') \\ \in {} & K. \end{aligned}$$

Definiere

$$t : \operatorname{Im}(\alpha_1) \times \cdots \times \operatorname{Im}(\alpha_n) \to (X_1 \otimes \cdots \otimes X_n)/K$$

durch

$$t(y_1, \ldots, y_n) = x_1 \otimes \cdots \otimes x_n + K, \text{ falls } y_i = \alpha(x_i).$$

Diese Abbildung ist wohldefiniert und n–fach linear, also gibt es genau einen Homomorphismus

$$\tau : \operatorname{Im}(\alpha_1) \otimes \cdots \otimes \operatorname{Im}(\alpha_n) \to (X_1 \otimes \cdots \otimes X_n)/K$$

mit $\tau(y_1 \otimes \cdots \otimes y_n) = t(y_1, \ldots, y_n)$. Dann ist $\tau\bar{\alpha} : (X_1 \otimes \cdots \otimes X_n)/K \to (X_1 \otimes \cdots \otimes X_n)/K$ die Identität. Also ist $\{0\} = \operatorname{Ker}(\bar{\alpha}) = \operatorname{Ker}(\alpha)/K$, also $\operatorname{Ker}(\alpha) = K$.

15.24 Korollar:

Sei $U_i \leq V_i$, $i = 1, \ldots, n$. Dann ist

$$V_1/U_1 \otimes \cdots \otimes V_n/U_n \cong (V_1 \otimes \cdots \otimes V_n)/[U_1, \ldots, U_n].$$

Beweis: Betrachte den kanonischen Epimorphismus $\kappa_i : V_i \to V_i/U_i$. Dann ist $\gamma(\kappa_1 \otimes \cdots \otimes \kappa_n)$ ein Epimorphismus von $V_1 \otimes \cdots \otimes V_n$ auf $V_1/U_1 \otimes \cdots \otimes V_n/U_n$, und nach 15.23 ist $\mathrm{Ker}\,(\gamma(k_1 \otimes \cdots \otimes \kappa_n)) = [\mathrm{Ker}\,(\kappa_1), \ldots, \mathrm{Ker}\,(\kappa_n)] = [U_1, \ldots, U_n]$. Die Behauptung folgt aus dem ersten Isomorphiesatz 5.11.

16 Alternierende Abbildungen und äusseres Produkt

16.1 Definition: *alternierende Abbildung*

Seien V und W $\mathbb{K}$–Vektorräume und $\alpha : V^n \to W$ n–fach linear ($n \geq 2$). Man nennt α eine alternierende Abbildung, wenn für jedes n–Tupel $(v_1, \ldots, v_n) \in V^n$ mit zwei gleichen v_i's gilt

$$\alpha(v_1, \ldots, v_n) = 0.$$

16.2 Beispiel:

(i) Die Determinantenabbildung det ist alternierend (mit $V = \mathbb{K}^n$ und $W = \mathbb{K}$).

(ii) Die Nullabbildung ist alternierend.

(iii) Seien α und $\beta : V^n \to W$ alternierend. Dann ist $\alpha + \beta$ alternierend und ebenso $k\alpha$ für $k \in \mathbb{K}$. Die Menge $A_n(V, W)$ der alternierenden Abbildungen bildet damit einen Vektorraum.

(iv) Wenn $\alpha : V^n \to W$ alternierend und $\lambda : W \to U$ linear, dann ist $\lambda\alpha : V^n \to U$ alternierend.

16.3 Lemma:

Sei $\alpha : V^n \to W$ alternierend. Wenn $v_1, \ldots, v_n \in V$ linear abhängig sind, dann ist $\alpha(v_1, \ldots, v_n) = 0$.

Beweis: Sei $0 = \sum_{i=1}^{n} k_i v_i$ und o.B.d.A. $k_1 \neq 0$. Dann ist

$$0 = \alpha\left(\sum_{i=1}^{n} k_i v_i, v_2, \ldots, v_n\right) = \sum_{i=1}^{n} k_i \alpha(v_i, v_2, \ldots, v_n) = k_1 \alpha(v_1, v_2, \ldots, v_n),$$

also $\alpha(v_1, \ldots, v_n) = 0$.

16.4 Satz:

Seien V und W $\mathbb{K}$–Vektorräume. Sei $B = \{b_i \mid i \in I\}$ eine Basis von V und sei X die Menge aller n–elementigen Teilmengen von B. Zu jedem $x = \{b_{i_1}, \ldots, b_{i_n}\} \in X$ wähle man eine Anordnung $b_{i_1} < \cdots < b_{i_n}$ und ein Element $w_x \in W$. Dann gibt es genau eine alternierende Abbildung $\alpha : V^n \to W$ mit $\alpha(b_{i_1}, \ldots, b_{i_n}) = w_x$.

Beweis: Eindeutigkeit:

Da α n–fach linear ist, ist α schon bestimmt durch $\alpha(a_1, \ldots, a_n)$ für jedes n–Tupel von Basiselementen. Dies ist gleich Null, wenn nicht alle a_i's verschieden sind (da α alternierend). Wenn sie alle verschieden sind, dann ist $\{a_1, \ldots, a_n\} = x \in X$. Daher gibt es eine

Permutation $\pi \in S_n$ mit $a_j = b_{i_{\pi(j)}}$, $j = 1, \dots, n$. Da α alternierend ist, gilt

$$\begin{aligned} 0 &= \alpha(\dots, v_i \underset{i}{+} v_j, \dots, v_i \underset{j}{+} v_j, \dots) \\ &= \alpha(\dots, v_i, \dots, v_i, \dots) + \alpha(\dots, v_j, \dots, v_j, \dots) + \alpha(\dots, v_i, \dots, v_j, \dots) \\ &\quad + \alpha(\dots, v_j, \dots, v_i, \dots) \\ &= \alpha(\dots, v_i, \dots, v_j, \dots) + \alpha(\dots, v_j, \dots, v_i, \dots), \end{aligned}$$

also

$$\alpha(\dots, v_i, \dots, v_j, \dots) = -\alpha(\dots, v_j, \dots, v_i, \dots),$$

d.h. das Vorzeichen des Wertes von α ändert sich, wenn man zwei der v_i's vertauscht. Da π sich als Produkt von k Transpositionen schreiben läßt, ändert sich das Vorzeichen von π genau k–mal, also

$$\alpha(a_1, \dots, a_n) = \alpha(b_{i_{\pi(1)}}, \dots, b_{i_{\pi(n)}}) = (-1)^k \alpha(b_{i_1}, \dots, b_{i_n}) = \operatorname{sign}\pi \cdot w_x.$$

Existenz:

Es genügt, α auf n–Tupeln von Basiselementen zu definieren und dann n–fach linear fortzusetzen. Auf einem solchen n–Tupel $(a_1, \dots, a_n)$ definiert man α wie oben, d.h.

$$\alpha(a_1, \dots, a_n) = \begin{cases} 0 & \text{wenn nicht alle } a_i \text{ verschieden} \\ \operatorname{sign}\pi \cdot w_x & \text{wenn } \{a_1, \dots, a_n\} = x \in X \text{ und } a_j = b_{i_{\pi(j)}},\ j = 1, \dots, n, \\ & \pi \in S_n. \end{cases}$$

Dann ist α alternierend, denn wenn $v_i = v_j = \sum_{\lambda=1}^{\ell} k_\lambda b_\lambda$, dann ist

$$\begin{aligned} &\alpha(v_1, \dots, v_i, \dots, v_j, \dots, v_n) \\ &\quad = \sum_{\lambda,\mu=1}^{\ell} k_\lambda k_\mu \alpha(v_1, \dots, b_\lambda, \dots, b_\mu, \dots, v_n) \\ &\quad = \sum_{\substack{\lambda,\mu=1 \\ \lambda \neq \mu}}^{\ell} k_\lambda k_\mu \alpha(v_1, \dots, b_\lambda, \dots, b_\mu, \dots, v_n) \\ &\quad = \sum_{1 \le \lambda < \mu \le \ell} k_\lambda k_\mu \big(\alpha(v_1, \dots, b_\lambda, \dots, b_\mu, \dots, v_n) + \alpha(v_1, \dots, b_\mu, \dots, b_\lambda, \dots, v_n)\big) \\ &\quad = \sum_{1 \le \lambda < \mu \le \ell} k_\lambda k_\mu \big(\alpha(v_1, \dots, b_\lambda, \dots, b_\mu, \dots, v_n) - \alpha(v_1, \dots, b_\lambda, \dots, b_\mu, \dots, v_n)\big) \\ &\quad = 0. \end{aligned}$$

16.5 Korollar:

Seien V und W endlich–dimensional mit Basen $B = \{b_1, \dots, b_d\}$ von V und $\{w_1, \dots, w_r\}$ von W. Sei $x = \{b_{i_1}, \dots, b_{i_n}\}$ eine feste n–elementige Teilemenge von B mit $i_1 < \dots < i_n$ und sei $1 \le \varrho \le r$ fest. Sei $\alpha_{x,\varrho}$ diejenige alternierende Abbildung $V^n \to W$, für welche $\alpha_{x,\varrho}(b_{i_1}, \dots, b_{i_n}) = w_\varrho$ und $\alpha_{x,\varrho}(a_1, \dots, a_n) = 0$ für jede Teilmenge $\{a_1, \dots, a_n\} \neq x$, $\{a_1, \dots, a_n\} \subseteq B$ gilt (nach 16.4 gibt es genau eine solche alternierende Abbildung).

Dann bilden die $\alpha_{x,\varrho}$'s eine Basis von $A_n(V,W)$.

Beweis: Lineare Unabhängigkeit:

Sei $\sum_{x,\varrho} k_{x,\varrho}\alpha_{x,\varrho} = 0$. Dann ist für jedes feste $y = \{b_{i_1}, \ldots, b_{i_n}\} \subseteq B$

$$0 = \sum_{x,\varrho} k_{x,\varrho}\alpha_{x,\varrho}(b_{i_1}, \ldots, b_{i_n}) = \sum_{\varrho} k_{y,\varrho} w_\varrho.$$

Da die w_ϱ's linear unabhängig sind, ist $k_{y,\varrho} = 0$ für alle ϱ, y.

Erzeugenden–System:

Sei $\alpha \in A_n(V,W)$ und sei $x = \{b_{i_1}, \ldots, b_{i_n}\}$ fest. Dann ist

$$\begin{aligned}\alpha(b_{i_1}, \ldots, b_{i_n}) &= w_x \in W \\ &= \sum_{\varrho} k_{x,\varrho} w_\varrho \\ &= \sum_{\varrho} k_{x,\varrho}\alpha_{x,\varrho}(b_{i_1}, \ldots, b_{i_n}) \\ &= \sum_{\varrho,y} k_{y,\varrho}\alpha_{y,\varrho}(b_{i_1}, \ldots, b_{i_n}).\end{aligned}$$

Also stimmen α und $\sum_{\varrho,y} k_{y,\varrho}\alpha_{y,\varrho}$ auf allen $(b_{i_1}, \ldots, b_{i_n})$ überein. Da beides alternierende Abbildungen sind, müssen sie wegen der Eindeutigkeit (Satz 16.4) gleich sein. Daher ist α eine Linearkombination der $\alpha_{y,\varrho}$'s.

16.6 Korollar:

Seien V und W wie in 16.5. Dann ist

$$\dim A_n(V,W) = \binom{d}{n} r.$$

Beweis: Es gibt $\binom{d}{n}$ n–elementige Teilmengen x von $B = \{b_1, \ldots, b_d\}$, also $\binom{d}{n} r$ Abbildungen $\alpha_{x,\varrho}$, $x \subseteq B$, $|x| = n$, $\varrho = 1, \ldots, r$. Die Behauptung folgt aus 16.5.

16.7 Bemerkung:

Wir wollen jetzt zu gegebenem V und n einen neuen Vektorraum $V^{\wedge n}$, genannt das n–fache äußere Produkt von V, und eine alternierende Abbildung $A : V^n \to V^{\wedge n}$ konstruieren derart, daß für jeden Vektorraum W die Abbildung

$$\psi : \mathrm{Hom}\,(V^{\wedge n}, W) \to A_n(V,W), \quad \psi(\lambda) = \lambda A$$

ein Isomorphismus ist.

16.8 Definition: *n–fache äußeres Produkt*

Sei

$$V^{\otimes n} = \underbrace{V \otimes \cdots \otimes V}_{n}$$

und sei U der Unterraum von $V^{\otimes n}$, welcher von allen Elementen $v_1 \otimes \cdots \otimes v_n$ mit zwei gleichen $v_i = v_j$ erzeugt wird. Der Faktorraum

$$V^{\wedge n} = V^{\otimes n}/U$$

wird das n–fache äußere Produkt von V genannt. Für das Bild von $v_1 \otimes \cdots \otimes v_n$ unter der kanonischen Abbildung $\kappa : V^{\otimes n} \to V^{\otimes n}/U$ schreibt man auch $v_1 \wedge \cdots \wedge v_n$, d.h.

$$v_1 \wedge \cdots \wedge v_n = v_1 \otimes \cdots \otimes v_n + U \in V^{\wedge n}.$$

Sei $A = \kappa T$, also

$$A(v_1, \ldots, v_n) = \kappa T(v_1, \ldots, v_n) = \kappa(v_1 \otimes \cdots \otimes v_n) = v_1 \otimes \cdots \otimes v_n + U = v_1 \wedge \cdots \wedge v_n.$$

16.9 Lemma:

Die Abbildung $A : V^n \to V^{\wedge n}$ ist alternierend.

Beweis: Da T n–fach linear (15.8) und κ linear ist, ist A n–fach linear (15.2). Wenn $(v_1, \ldots, v_n) \in V^n$ mit zwei gleichen $v_i = v_j$, dann ist $T(v_1, \ldots, v_n) = v_1 \otimes \cdots \otimes v_n \in U$, also $A(v_1, \ldots, v_n) = \kappa T(v_1, \ldots, v_n) = 0$, da $U = \ker \kappa$. Daher ist A alternierend.

16.10 Satz:

Seien V und W $\mathbb{K}$–Vektorräume. Die Abbildung

$$\lambda \;\mapsto\; \lambda A \;\text{ ist ein Isomorphismus }\; \mathrm{Hom}\,(V^{\wedge n}, W) \;\to\; A_n(V, W).$$

Beweis: Da A alternierend ist (16.9), ist λA alternierend (16.2), d.h. $\lambda A \in A_n(V, W)$. Daß die Abbildung linear ist, ist trivial.

Injektiv:

Sei $\lambda \neq 0$, zu zeigen ist $\lambda A \neq 0$. Da $\lambda \neq 0$, gibt es ein Element x in $V^{\wedge n} = V^{\otimes n}/U$ mit $\lambda(x) \neq 0$. Da x eine endliche Summe von Elementen der Form $v_1 \otimes \cdots \otimes v_n + U$ ist, gibt es $(v_1, \ldots, v_n) \in V^n$ mit $0 \neq \lambda(v_1 \otimes \cdots \otimes v_n + U) = \lambda A(v_1, \ldots, v_n)$. Daher ist $\lambda A \neq 0$.

Surjektiv:

Sei $\alpha : V^n \to W$ eine alternierende Abbildung. Gesucht ist $\lambda : V^{\wedge n} \to W$ linear mit $\alpha = \lambda A$. Da α insbesondere n–fach linear ist, gibt es es nach 15.10 eine lineare Abbildung $\mu : V^{\otimes n} \to W$ mit $\alpha = \mu T$. Sei $v_1 \otimes \cdots \otimes v_n \in V^{\otimes n}$ mit zwei gleichen $v_i = v_j$. Dann ist

$$\mu(v_1 \otimes \cdots \otimes v_n) = \mu T(v_1, \ldots, v_n) = \alpha(v_1, \ldots, v_n) = 0,$$

da α alternierend ist. Also ist dann $v_1 \otimes \cdots \otimes v_n \in \mathrm{Ker}\,(\mu)$. Da diese Elemente den Unterraum U erzeugen, folgt $U \leq \mathrm{Ker}\,(\mu)$. Also ist durch $\lambda(x + U) = \mu(x)$ für $x \in V^{\otimes n}$ eine lineare Abbildung $\lambda : V^{\wedge n} \to W$ wohldefiniert, und es gilt $\alpha = \mu T = \lambda \kappa T = \lambda A$.

16.11 Korollar:

Wenn V die endliche Dimension d hat, dann ist

$$\dim V^{\wedge n} = \binom{d}{n}.$$

Wenn $\{b_1, \ldots, b_d\}$ eine Basis von V ist, dann ist

$$\{b_{i_1} \wedge \cdots \wedge b_{i_n} \mid 1 \leq i_1 < \cdots < i_n \leq d\}$$

eine Basis von $V^{\wedge n}$.

Beweis: Verwende Satz 16.10 mit $W = \mathbb{K}$. Dann ist

$$(V^{\wedge n})^* = \mathrm{Hom}\,(V^{\wedge n}, \mathbb{K}) \cong A_n(V, \mathbb{K}).$$

Nach 16.6 ist $\dim A_n(V, \mathbb{K}) = \binom{d}{n}$. Also hat $(V^{\wedge n})^*$ endliche Dimension $\binom{d}{n}$. Daher hat auch $V^{\wedge n}$ endliche Dimension $\binom{d}{n}$, da $\dim W = \dim W^*$ für $\dim W < \infty$. Die Aussage über die Basis folgt dann, da die angegebene Menge nach 15.17 ein Erzeugenden System ist.

16.12 Definition: *antisymmetrische Abbildung*

Eine n–fach lineare Abbildung $\alpha : V^n \to W$ heißt antisymmetrisch, wenn

$$\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

für alle $v_1, \ldots, v_n \in V$ gilt.

16.13 Bemerkung:

(i) Jede alternierende Abbildung ist antisymmetrisch.

(ii) Wenn α antisymmetrisch ist, dann gilt

$$\alpha(v_{\pi(1)}, \ldots, v_{\pi(n)}) = \mathrm{sign}\pi \cdot \alpha(v_1, \ldots, v_n)$$

für alle $\pi \in S_n$, $v_1, \ldots, v_n \in V$.

(Vergleiche den Beweis von 16.4)

16.14 Definition:

Sei $\varphi : V^n \to W$ eine n–fach lineare Abbildung. Dann definiert man $\varphi_a : V^n \to W$ durch

$$\varphi_a(v_1, \ldots, v_n) = \sum_{\pi \in S_n} \mathrm{sign}\pi \cdot \varphi(v_{\pi(1)}, \ldots, v_{\pi(n)}).$$

16.15 Satz:

Die Abbildung φ_a ist n–fach linear und alternierend (also insbesondere antisymmetrisch). Ist umgekehrt $\alpha : V^n \to W$ alternierend, dann existiert eine n–fach lineare Abbildung $\varphi : V^n \to W$ mit $\alpha = \varphi_a$.

Beweis: φ_a ist n–fach linear:

$$\begin{aligned}\varphi_a(v_1, \ldots, kv_i, \ldots, v_n) &= \sum_{\pi \in S_n} \operatorname{sign}\pi \cdot \varphi\left(v_{\pi(1)}, \ldots, kv_{\pi(j_\pi)}, \ldots, v_{\pi(n)}\right), \\ &\qquad \text{wobei } \pi(j_\pi) = i,\ j_\pi = \pi^{-1}(i), \\ &= k \sum_{\pi \in S_n} \operatorname{sign}\pi \cdot \varphi\left(v_{\pi(1)}, \ldots, v_{\pi(n)}\right) \\ &= k\varphi_a(v_1, \ldots, v_n).\end{aligned}$$

Ebenso für Summen.

φ_a ist alternierend:

Sei $v_i = v_j$, $i \neq j$, und sei τ die Transposition (ij). Dann ist $S_n = A_n \,\dot\cup\, \tau A_n$ und

$$\varphi_a(v_1, \ldots, v_n) = \sum_{\pi \in A_n} \operatorname{sign}\pi \cdot \varphi\left(v_{\pi(1)}, \ldots, v_{\pi(n)}\right) - \sum_{\pi \in A_n} \operatorname{sign}\pi \cdot \varphi\left(v_{\tau\pi(1)}, \ldots, v_{\tau\pi(n)}\right).$$

Für festes $\pi \in A_n$ sei $\pi(k) = i$ und $\pi(\ell) = j$. Dann ist

$$\begin{aligned}&\varphi\left(v_{\pi(1)}, \ldots, v_{\pi(n)}\right) - \varphi\left(v_{\tau\pi(1)}, \ldots, v_{\tau\pi(n)}\right) \\ &\quad = \varphi\left(v_{\pi(1)}, \ldots, v_i, \ldots, v_j, \ldots, v_{\pi(n)}\right) - \varphi\left(v_{\tau\pi(1)}, \ldots, v_j, \ldots, v_i, \ldots, v_{\tau\pi(n)}\right) \\ &\quad = 0,\end{aligned}$$

weil $v_i = v_j$. Daher ist $\varphi_a(v_1, \ldots, v_n) = 0$.

Sei schließlich $\alpha : V^n \to W$ eine beliebige alternierende Abbildung. Sei B eine Basis von V. Für jede n–elementige Teilmenge x von B sei eine Anordnung gewählt; $x = \{b_{i_1} < \cdots < b_{i_n}\}$. Sei nun

$$\varphi_0(b_{j_1}, \ldots, b_{j_n}) = \begin{cases} \alpha(b_{j_1}, \ldots, b_{j_n}) & \text{falls die } b_{j_r} \text{ alle verschieden sind und in der} \\ & \text{richtigen Reihenfolge vorkommen} \\ 0 & \text{sonst.} \end{cases}$$

Dann läßt sich φ_0 eindeutig zu einer n–fach linearen Abbildung $\varphi : V^n \to W$ fortsetzen. Es ist $\varphi_a = \alpha$, denn:

$$\begin{aligned}\varphi_a(b_{i_1}, \ldots, b_{i_n}) &= \sum_{\pi \in S_n} \operatorname{sign}\pi \cdot \varphi\left(b_{\pi(i_1)}, \ldots, b_{\pi(i_n)}\right) \\ &= \sum_{\pi \in S_n} \operatorname{sign}\pi \cdot \varphi_0\left(b_{\pi(i_1)}, \ldots, b_{\pi(i_n)}\right) \\ &= \alpha(b_{i_1}, \ldots, b_{i_n}).\end{aligned}$$

Also stimmen φ_a und α auf allen $(b_{i_1}, \ldots, b_{i_n})$ überein. Da beide alternierend sind, sind sie nach 16.4 gleich.

17 Affine und projektive Ebenen

In G.Pickert, Projektive Ebenen, Springer 1955, findet man zu den folgenden Kapiteln viele weitere Ergebnisse.

17.1 Definition: *Inzidenzstruktur*

Seien zwei Mengen $\mathcal{P}$ (die 'Punktmenge') und $\mathcal{G}$ (die 'Geradenmenge') sowie eine Teilmenge $I \subseteq \mathcal{P} \times \mathcal{G}$ gegeben. Wenn $(P, \gamma) \in I$ (P ein Punkt, γ eine Gerade), dann sagen wir, dass P auf γ liegt, oder auch, dass γ durch P geht. Man nennt das Tripel $(\mathcal{P}, \mathcal{G}, I)$ eine Inzidenzstruktur, wenn es durch zwei verschiedene Punkte höchstens eine Gerade gibt. Punkte $P_1, \ldots, P_n$ heißen kollinear, wenn sie alle auf einer Geraden liegen.

17.2 Bemerkung:

Zwei verschieden Geraden $\gamma \neq \delta$ können daher höchstens einen Punkt P gemeinsam haben. Wenn ein solcher Punkt existiert, heißt er der Schnittpunkt $\gamma \cap \delta$ der Geraden; man sagt auch, dass die Geraden sich in P schneiden.

17.3 Definition: *Affine und Projektive Ebenen*

Eine Inzidenzstruktur heißt affine Ebene, wenn die folgenden Axiome erfüllt sind:

(i) Durch je zwei verschiedene Punkte gibt es eine Gerade.

(ii) Wenn P ein Punkt ist, der nicht auf der Geraden γ liegt, dann gibt es genau eine Gerade, welche durch P geht und γ nicht schneidet.

(iii) Es gibt drei nicht kollineare Punkte.

Eine Inzidenzstruktur heißt projektive Ebene, wenn die folgenden Axiome erfüllt sind:

(i) Durch je zwei verschiedene Punkte gibt es eine Gerade.

(ii) Je zwei verschiedene Geraden haben einen Schnittpunkt.

(iii) Es gibt vier Punkte, von denen keine drei kollineare sind.

17.4 Bezeichnung/Bemerkung: *Parallelen*

(i) Durch je zwei verschiedene Punkte P, Q einer projektiven oder affinen Ebene gibt es also genau eine Gerade, die wir die Verbindungsgerade PQ nennen. Wenn diese Bezeichnung auftaucht, ist also immer zu kontrollieren, ob wirklich $P \neq Q$. Ebenso ergibt $\gamma \cap \delta$ nur dann einen Sinn, wenn die Geraden γ und δ im projektiven Fall verschieden und im affinen Fall nicht parallel (s.u.) sind. Die Verifikation dieser Bedingungen wird manchmal stillschweigend Ihnen überlassen!

(ii) Die Eigenschaft, das keine drei Punkte einer Punktmenge kollinear sind, wird manchmal auch so formuliert, dass die Punkte 'in allgemeiner Lage' sind.

(iii) Die Axiome für eine projektiven Ebene $\mathcal{E}$ lassen sich auch so ausdrücken:

- Durch je zwei verschiedene Punkte geht genau eine Gerade.
- Zwei verschiedene Geraden schneiden sich in genau einem Punkt.
- Es gibt vier Punkte in allgemeiner Lage.

Die zu $\mathcal{E}$ duale Ebene $\mathcal{E}^* = (\mathcal{P}^*, \mathcal{G}^*, I^*)$ hat als Punkte die Geraden von $\mathcal{E}$, als Geraden die Punkte von $\mathcal{E}$ und die Inzidenz ist durch $(\gamma, P) \in I^* \Leftrightarrow (P, \gamma) \in I$ definiert. Die beiden ersten Axiome sind dann für $\mathcal{E}^*$ klar. Das dritte gilt, da es in $\mathcal{E}$ vier Geraden gibt, von denen keine drei durch einen Punkt gehen.
Vertauscht man in einem Satz über projektive Ebenen 'Punkt', und 'Gerade', so erhält man also wieder einen (den dualen) Satz. Zum Beispiel werden wir gleich zeigen, dass auf jeder Geraden einer projektiven Ebene wenigstens drei Punkte liegen. Dualisieren ergibt, dass durch jeden Punkt wenigstens drei Geraden gehen.

(iv) Die in einer affinen Ebene nach (ii) der Definition existierende Gerade nennt man die Parallele zu γ durch P. Wenn P auf γ liegt, versteht man darunter einfach γ selbst.

(v) Geraden γ und δ heißen parallel, wenn $\gamma = \delta$ oder γ und δ sich nicht schneiden. Man schreibt dann $\gamma \,||\, \delta$.

(vi) Die Parallelität ist eine Äquivalenz-Relation auf den Geraden einer affinen Ebene: Symmetrie und Reflexivität sind klar. Sei $\alpha \,||\, \beta$ und $\beta \,||\, \gamma$; zu zeigen ist $\alpha \,||\, \gamma$. Wenn $\alpha = \beta$, dann gilt dies offenbar; wir nehmen daher $\alpha \neq \beta$ an. Wenn α und γ keinen Punkt gemeinsam haben, sind sie parallel. Im anderen Fall sei P ein Punkt, welcher sowohl auf α als auch auf γ liegt. Dann liegt P nicht auf β, weil α und β parallel, aber verschieden sind. Daher sind α und γ Parallelen zu β durch P. Wegen der Eindeutigkeit ist $\alpha = \gamma$, also wieder $\alpha \,||\, \gamma$. Dies zeigt die Transitivität. Die Menge aller zueinander parallelen Geraden nennt man auch eine Parallelenschar.

(vii) Das dritte Axiom ist für die projektiven Ebenen etwas stärker formuliert als für die affinen, aber auch die affinen Ebenen erfüllen diese stärkere Forderung. Denn seien A, B, C drei nicht kollineare Punkte. Sei γ die Parallele zu AB durch C und β die Parallele zu AC durch B. Dann ist $\gamma \neq AB$ (denn C liegt auf γ, aber nach Voraussetzung nicht auf AB) und ebenso $\beta \neq AC$. Außerdem sind γ und β nicht parallel, denn sonst wäre auch $(AB) \,||\, (AC)$ (wie wir gerade unter (vi) gesehen haben), also $AB = AC$, was wieder nicht geht, da A, B, C nicht kollinear sind. Daher haben γ und β einen Schnittpunkt S. Keine drei der Punkte A, B, C, S sind kollinear: Für A, B, C ist dies klar. Wären A, B, S kollinear, also S auf AB, dann hätten γ und AB den Punkt S gemeinsam. Da diese Geraden parallel sind, wären sie gleich, ein Widerspruch. Analog zeigt man, dass A, C, S nicht kollinear sind. Insbesondere ist S verschieden von A, B, C. Wäre S auf der Geraden BC, dann wäre $\gamma = CS = CB$, und B wäre Schnittpunkt von γ und AB, Widerspruch.

17.5 Beispiel: *Affine und projektive Ebenen*

Sei K ein Schiefkörper, d.h. die Multiplikation in K braucht nicht kommutativ zu sein; alle anderen Körperaxiome gelten in K. Es ist leicht zu kontrollieren, dass die elementaren Sätze über Vektorräume (wie üblich definiert) immer noch gelten, da ihre Beweise keinen Gebrauch von der Kommutativität der Multiplikation machen. Man muss sich allerdings entscheiden, ob man Rechts- oder Links-Vektorräume betrachtet.

(i) Zu einem zweidimensionalen K-Vektorraum V konstruiert man eine affine Ebene $\mathcal{A} = \mathcal{A}(V)$ wie folgt: Die Punkte von $\mathcal{A}$ sind die Vektoren, die Geraden die sämtlichen Nebenklassen $v + U$, wobei v alle Vektoren und U alle eindimensionalen Unterräume von V durchläuft. Die Inzidenz ist durch $a\, I\, (v + U) \Leftrightarrow a \in (v + U) \Leftrightarrow a + U = v + U$ definiert (vergleiche 5.3). Durch zwei verschiedene Punkte a, b geht dann genau eine Gerade, nämlich $a + K(b - a)$. Wenn $a + U_1$ und $b + U_2$ zwei verschiedene Geraden sind, dann kann $U_1 = U_2$ sein. In diesem Fall schneiden sich die Geraden nicht. Wenn dagegen $U_1 \neq U_2$, dann ist $V = U_1 + U_2$, da $\dim V = 2$. Folglich gibt es $u_1 \in U_1,\ u_2 \in U_2$ mit $u_1 - u_2 = b - a$, also $a + u_1 = b + u_2 \in (a + U_1) \cap (b + U_2)$. Schließlich sind die drei Punkte $0, a, b$ nicht kollinear, wenn a, b linear unabhängig sind. Daher ist $\mathcal{A}$ eine affine Ebene.

(ii) Zu einem dreidimensionalen K-Vektorraum W konstruiert man eine projektive Ebene $\mathcal{E} = \mathcal{E}(W)$ wie folgt: Die Punkte von $\mathcal{E}$ sind alle eindimensionalen Unterräume, die Geraden alle zweidimensionalen Unterräume von W. Wenn U ein Punkt, V eine Gerade ist, dann ist die Inzidenz durch $U\, I\, V \Leftrightarrow U \leq V$ definiert. Durch zwei verschiedene Punkte U_1, U_2 geht dann genau eine Gerade, nämlich $U_1 + U_2$. Wenn V_1, V_2 zwei verschiedene Geraden sind, dann ist $\dim (V_1 \cap V_2) = 1$ nach 5.24, also ist $V_1 \cap V_2$ der Schnittpunkt der Geraden. Wenn b_1, b_2, b_3 drei linear unabhängige Vektoren sind, dann erzeugen je drei der vier Unterräume Kb_1, Kb_2, Kb_3, $K(b_1 + b_2 + b_3)$ schon den ganzen Raum, sind also nicht kollinear. Daher liefert $\mathcal{E}$ ein Beispiel für eine projektive Ebene.

17.6 Bemerkung:

(i) Aus jeder projektiven Ebene $\mathcal{E}$ läßt sich durch Weglassen einer Geraden γ und aller Punkte auf dieser Geraden eine affine Ebene $\mathcal{E}_\gamma$ gewinnen: Wenn α eine Gerade und P ein Punkt von $\mathcal{E}_\gamma$ sind, und P nicht auf α liegt, dann findet man die Parallele zu α durch P, indem man die Verbindungsgerade β (in $\mathcal{E}$) von P mit $\gamma \cap \alpha$ bildet. In $\mathcal{E}_\gamma$ sind α und β dann parallel. Jede andere Gerade durch P schneidet α in einem Punkt, der nicht auf γ liegt, also noch zu $\mathcal{E}_\gamma$ gehört. Damit ist das zweite Axiom kontrolliert. Wir müssen noch drei nicht kollineare Punkte in $\mathcal{E}_\gamma$ finden. Nach Voraussetzung gibt es in $\mathcal{E}$ vier Punkte P_1, P_2, P_3, P_4, von denen keine drei kollinear sind. Wenn wenigstens drei dieser Punkte nicht auf γ liegen, dann sind wir fertig. Wenn dagegen etwa P_3, P_4 auf γ liegen, dann kann man zu P_1, P_2 als dritten Punkt den Schnittpunkt von P_1P_3 und P_2P_4 hinzunehmen, denn dieser liegt weder auf γ noch auf P_1P_2.
Wir werden gleich sehen, dass man *alle* affinen Ebenen auf diese Weise aus projektiven Ebenen erhält.

(ii) Die Konstruktion einer affinen Ebene aus einer projektiven Ebene durch Weglassen einer Geraden läßt sich umkehren. Für jede Äquivalenzklasse paralleler Geraden in einer affinen Ebene $\mathcal{A}$ fügt man einen (uneigentlichen oder unendlichen) Punkt hinzu, durch den genau die Geraden dieser Parallelenschar gehen; diese schneiden sich in der so erweiterten Ebene also alle in diesem uneigentlichen Punkt (das ist der Sinn der etwas mystischen Aussage, dass sich zwei Parallelen 'im Unendlichen' schneiden). Außerdem fügt man noch *genau eine* Gerade, die uneigentliche Gerade, hinzu, auf der genau die uneigentlichen Punkte liegen. Es ist klar, dass die so gewonnene Inzidenzstruktur $\overline{\mathcal{A}}$ die beiden ersten Axiome für eine projektive Ebene erfüllt. Das dritte gilt sogar schon in $\mathcal{A}$, wie wir in 17.4, (vii) gesehen haben.
Wenn man die uneigentliche Gerade aus $\overline{\mathcal{A}}$ wieder entfernt, so erhält man natürlich $\mathcal{A}$

zurück; also enstehen alle affinen Ebenen aus projektiven.

(iii) Während sich aus einer projektiven Ebene viele affine Ebenen gewinnen lassen, da ja die wegzulassende Gerade beliebig gewählt werden kann, ist die Konstruktion der projektiven Ebene aus einer gegebenen affinen Ebene völlig eindeutig. Wir sprechen auch von der zugehörigen projektiven Ebene.

(iv) Im Beispiel oben sei V ein zweidimensionaler Unterraum des dreidimensionalen Raums $W = V + Kb$. Zu jedem Punkt der affinen Ebene $\mathcal{A} = \mathcal{A}(V)$, also zu jedem Vektor $v \in V$, bilden wir den eindimensionalen Unterraum $K(v+b) \leq W$. Dies ist ein Punkt der projektiven Ebene $\mathcal{E} = \mathcal{E}(W)$. Wir erhalten so eine Bijektion zwischen den Punkten von $\mathcal{A}$ und den Punkten von $\mathcal{E}$, welche nicht auf V liegen. Zu jeder Geraden von $\mathcal{A}$, also zu jeder Nebenklasse $v+U$ mit einem eindimensionalen $U \leq V$, bilden wir den zweidimensionalen Unterraum $K(v+b)+U \leq W$. Dies ist eine Gerade von $\mathcal{E}$. Wir erhalten so eine Bijektion zwischen den Geraden von $\mathcal{A}$ und den Geraden $\neq V$ von $\mathcal{E}$. Als uneigentliche Gerade nehmen wir noch V selbst dazu. Offenbar ist $V \cap (K(v+b)+U) = U$, also schneiden sich alle Parallelen $v+U$ von $\mathcal{A}$ im uneigentlichen Punkt U von $\mathcal{E}$.
Damit ist gezeigt, dass $\overline{\mathcal{A}(V)} \cong \mathcal{E}(W)$. (Die genaue Definition von Isomorphie folgt unten.)

(v) Diese Diskussion ist in $W = R^3$ gut zu veranschaulichen: Als V wählen wir zum Beispiel die (x,y)-Ebene. Dazu parallel, etwa im Abstand 1, betrachten wir eine zweite Ebene $H = \{(x,y,1) \mid x,y \in \mathbb{R}\} = (0,0,1) + V$. Jede Gerade U durch den Nullpunkt, die nicht in der (x,y)-Ebene liegt, schneidet H in genau einem Punkt $(x_U, y_U, 1)$; fällt man von diesem Punkt das Lot auf V (d.h. ersetzt man 1 durch 0), so erhält also zu U einen Punkt $v_U \in V$. Jeder zweidimensionale Unterraum $X \neq V$ von W (d.h X ist eine andere Ebene durch den Nullpunkt) schneidet H in einer Geraden, die wir wieder in die (x,y)-Ebene projizieren können. (Die Gerade, welche man dabei erhält, braucht aber nicht durch den Nullpunkt zu gehen.)

(vi) In einer projektiven Ebene entfällt bei zwei Geraden die lästige Unterscheidung, ob sie parallel sind oder nicht. Außerdem gibt die Dualität ein bequemes Hilfsmittel bei Beweisen. Daher sind projektive Ebenen 'einfacher' als affine. Sie sind allerdings weniger anschaulich. Wir werden daher hauptsächlich mit affinen Ebenen arbeiten und uns bei Bedarf erinnern, dass diese in einer etwas größeren projektiven Ebene enthalten sind.

17.7 Satz:

Sei $\mathcal{E}$ eine projektive Ebene.

(1) Auf jeder Geraden liegen mindestens drei Punkte.

(2) Durch jeden Punkt gehen mindestens drei Geraden.

(3) Zu je zwei Geraden gibt es einen Punkt, der auf keiner der beiden liegt.

(4) Zu je zwei Punkten gibt es eine Gerade, die durch keinen der beiden geht.

(5) Sei P ein Punkt und γ eine Gerade, welche nicht durch P geht. Dann ist $\sigma : \alpha \mapsto \alpha \cap \gamma$ eine Bijektion zwischen den Geraden durch P und den Punkten auf γ.

(6) Seien $\alpha \neq \beta$ Geraden. Dann gibt es eine Bijektion zwischen den Punkten von α und den Punkten von β, bei der $\alpha \cap \beta$ fest bleibt.

(7) Seien $P \neq Q$ Punkte. Dann gibt es eine Bijektion zwischen den Geraden durch P und den Geraden durch Q, bei der PQ fest bleibt.

Beweis:

(1) Sei γ eine Gerade und A, B, C, D vier Punkte in allgemeiner Lage. Höchstens zwei dieser Punkte können auf γ liegen; o.E. liegt A nicht auf γ. Da die Geraden AB, AC, AD verschieden sind, haben sie verschiedene Schnittpunkte mit γ.

(2) Dual.

(3) Wir können annehmen, dass α und β verschiedene Geraden sind. Auf jeder dieser Geraden gibt es dann nach (1) außer $\alpha \cap \beta$ noch weitere Punkte, etwa P auf α und Q auf β. Wieder nach (1) liegt auf PQ ein dritter Punkt, der weder auf α noch auf β liegt.

(4) Dual.

(5) Wenn S ein Punkt auf γ ist, dann ist $\alpha = PS$ eine Gerade durch P. Dies definiert die Umkehrabbildung zu σ.

(6) Wähle einen Punkt P, der weder auf α noch auf β liegt; nach (3) ist dies möglich. Die Abbildung $Q \mapsto PQ \cap \beta$ für Q auf α ist nach (5) eine Bijektion.

(7) Dual.

17.8 Korollar:

Wenn $\mathcal{E}$ eine projektive Ebene ist und durch den Punkt P genau $n+1$ Geraden gehen ($n \in \mathbb{N}$), dann gehen durch jeden Punkt genau $n+1$ Geraden, und auf jeder Geraden liegen genau $n+1$ Punkte. Die Anzahl der Geraden wie auch die Anzahl der Punkte ist dann $n^2 + n + 1$.

Beweis: Dass die Anzahl der Geraden durch einen Punkt unabhängig von der Wahl des Punktes und gleich der Anzahl der Punkte auf jeder Geraden ist, steht im vorigen Satz (7), (5), (6). Auf jeder der $n+1$ Geraden durch P liegen also außer P noch n weitere Punkte, und jeder Punkt $Q \neq P$ liegt auf genau einer dieser Geraden, nämlich auf PQ. Also gibt es insgesamt $n(n+1)+1$ Punkte. Dual ist dies auch die Anzahl der Geraden.

17.9 Korollar:

Sei $\mathcal{A}$ eine affine Ebene.

(1) Auf jeder Geraden liegen mindestens zwei Punkte.

(2) Durch jeden Punkt gehen mindestens drei Geraden.

(3) Zu jeder Geraden gibt es einen Punkt, der nicht auf der Geraden liegt.

(4) Zu je zwei Punkten gibt es eine Gerade, die durch keinen der beiden geht.

(5) Sei P ein Punkt und γ eine Gerade, welche nicht durch P geht. Dann ist $\sigma : \alpha \mapsto \alpha \cap \gamma$ eine Bijektion zwischen den Geraden durch P, welche nicht zu γ parallel sind, und den Punkten auf γ.

(6) Seien α und β Geraden. Dann gibt es eine Bijektion zwischen den Parallelen zu α und den Punkten auf β.

Beweis: Nach 17.6 darf man annehmen, dass $\mathcal{A}$ aus einer projektiven Ebene durch Weglassen einer Geraden und der Punkte dieser Geraden hervorgeht. Die Behauptungen (1)–(3)

und (5) folgen.
(4) Seien o.E. $P \neq Q$ die beiden Punkte. Die Parallele zu PQ durch einen Punkt, der nicht auf PQ liegt, tut das verlangte.
(6) Wenn α und β nicht parallel sind, dann liefern die Schnittpunkte der Parallelen zu α mit β die gewünschte Bijektion. Wenn $\alpha \,||\, \beta$, dann sei P ein Punkt auf α. Durch diesen gibt es nach (2) noch zwei weitere Geraden, etwa γ und δ. Dann gibt es Bijektionen $\{\text{Parallelen zu } \alpha\} \to \{\text{Punkte auf } \gamma\} \to \{\text{Parallelen zu } \delta\} \to \{\text{Punkte auf } \beta\}$.

17.10 Definition: *Isomorphismus, Kollineation*

Seien $\mathcal{E} = (\mathcal{P}, \mathcal{G}, I)$ und $\mathcal{E}' = (\mathcal{P}', \mathcal{G}', I')$ zwei Inzidenzstrukturen.

(1) Wenn $\alpha : \mathcal{P} \to \mathcal{P}'$ und $\beta : \mathcal{G} \to \mathcal{G}'$ zwei Bijektionen sind und

$$PI\gamma \Leftrightarrow (P\alpha)I'(\gamma\beta)$$

für alle $P \in \mathcal{P}$ und alle $\gamma \in \mathcal{G}$ gilt dann heißt das Paar (α, β) ein Isomorphismus von $\mathcal{E}$ nach $\mathcal{E}'$. Man nennt $\mathcal{E}$ und $\mathcal{E}'$ isomorph und schreibt $\mathcal{E} \cong \mathcal{E}'$, wenn es einen Isomorphismus von $\mathcal{E}$ nach $\mathcal{E}'$ gibt.

(2) Eine bijektive Abbildung $\kappa : \mathcal{P} \to \mathcal{P}'$ heißt eine Kollineation von $\mathcal{E}$ nach $\mathcal{E}'$, wenn

$$P, Q, R \in \mathcal{P} \text{ kollinear} \quad \Rightarrow \quad P\kappa, Q\kappa, R\kappa \text{ kollinear.}$$

17.11 Bemerkung:

(1) Es ist klar, dass 'Isomorphie' reflexiv, symmetrisch und transitiv ist.

(2) Ebenfalls klar ist, dass die Punktabbildung α in einem Isomorphismus (α, β) eine Kollineation ist.

(3) Der nächste Satz zeigt, dass die Umkehrung auch gilt, jedenfalls für die uns hier interessierenden Inzidenzstrukturen.

17.12 Satz: *Isomorphismus = Kollineation*

Seien $\mathcal{E} = (\mathcal{P}, \mathcal{G}, I)$ und $\mathcal{E}' = (\mathcal{P}', \mathcal{G}', I')$ zwei projektive oder zwei affine Ebenen und sei $\kappa : \mathcal{P} \to \mathcal{P}'$ eine Kollineation. Dann gibt es genau eine Bijektion $\kappa^* : \mathcal{G} \to \mathcal{G}'$ derart, dass (κ, κ^*) ein Isomorphismus ist.

Beweis: Sei γ eine Gerade von $\mathcal{E}$. Wähle zwei Punkte $P \neq Q$ auf γ (vergleiche 17.9). Da κ injektiv ist, gibt es genau eine Verbindungsgerade $\gamma\kappa^* := (P\kappa)(Q\kappa)$ in $\mathcal{E}'$. Dadurch ist κ^* wohldefiniert, denn wenn auch $P_1 \neq Q_1$ auf γ, dann ist $(P\kappa)(Q\kappa) = (P_1\kappa)(Q_1\kappa)$, da $(P\kappa), (Q\kappa), (P_1\kappa), (Q_1\kappa)$ kollinear sind nach Voraussetzung an κ.
Es ist klar, dass bei dieser Definition von κ^* dann

$$PI\gamma \Rightarrow (P\kappa)I'(\gamma\kappa^*)$$

gilt und dass κ^* dadurch eindeutig bestimmt ist.
Wenn $\gamma' \in \mathcal{G}'$, wähle zwei Punkte $P' \neq Q'$ auf γ'. Weil κ surjektiv ist, gibt es $P, Q \in \mathcal{P}$ mit $P\kappa = P'$ und $Q\kappa = Q'$. Dann ist $(PQ)\kappa^* = \gamma'$, also ist κ^* surjektiv.

κ^* ist auch injektiv. Dazu seien $\gamma \neq \delta$ Geraden von $\mathcal{E}$. Wir nehmen $\gamma\kappa^* = \delta\kappa^*$ an und suchen einen Widerspruch.
Sei zunächst $\mathcal{E}$ eine projektive Ebene. Setze $S = \gamma \cap \delta$. Wenn P ein Punkt auf γ oder auf δ ist, dann liegt $P\kappa$ auf $\gamma\kappa^*$. Wenn P weder auf γ noch auf δ liegt, dann wähle einen Punkt $T \neq S$ auf γ. Die Gerade PT schneidet dann δ in einem Punkt $U \neq T$, also ist $PT = UT$ und daher $(P\kappa)(T\kappa) = (U\kappa)(T\kappa) = \gamma\kappa^*$, denn die beiden Punkte $U\kappa$ und $T\kappa$ liegen auf $\gamma\kappa^*$. Also liegt wieder $P\kappa$ auf $\gamma\kappa^*$. Daher liegen alle Punkte $P\kappa$ (und damit alle Punkte von $\mathcal{E}'$) auf einer Geraden, ein Widerspruch.
Sei jetzt $\mathcal{E}$ affin. Wenn γ und δ nicht parallel sind, dann schneidet jede Gerade α von $\mathcal{E}$ wenigstens eine der beiden Geraden. Aber dann schneidet $\alpha\kappa^*$ auch $\gamma\kappa^*$. Da κ^* surjektiv ist, schneidet $\gamma\kappa^*$ alle Geraden von $\mathcal{E}'$, was in einer affinen Ebene nicht sein kann. Wenn dagegen $\gamma \,||\, \delta$ und P ein Punkt, dann gibt es eine Gerade α durch P, welche γ und δ schneidet, etwa in T und U. Dann ist $\alpha\kappa^* = (TU)\kappa^* = \gamma\kappa^*$, also $P\kappa$ auf $\gamma\kappa^*$ für jeden Punkt P, ein Widerspruch.
Schließlich seien $P \in \mathcal{P}$ und $\gamma \in \mathcal{G}$. Wenn $(P\kappa)\, I'\, (\gamma\kappa^*)$, dann wähle noch einen zweiten Punkt auf $\gamma\kappa^*$, etwa $Q\kappa$. Dann ist $\gamma\kappa^* = (PQ)\kappa^*$, also $\gamma = PQ$, da κ^* injektiv ist. Insbesondere gilt $PI\gamma$. Damit ist (κ, κ^*) ein Isomorphismus.

17.13 Bemerkung:

Nach dem letzten Satz ist ein Isomorphismus (α, β) zwischen projektiven oder affinen Ebenen schon durch α eindeutig bestimmt. Wir werden daher auch die Abbildung zwischen den Geraden einfach mit α bezeichnen.

17.14 Satz: *Koordinaten*

Sei $\mathcal{E}$ eine projektive Ebene und O, X, Y, E vier Punkte in allgemeiner Lage. Wir nennen $\xi = OX$ die x-Achse, $\eta = OY$ die y-Achse und $\delta = OE$ die Diagonale. Weiter sei $\mathcal{A}$ die affine Ebene, welche durch Weglassen von XY entsteht. Die Menge K der Punkte $\neq Y$ auf η nennen wir den Koordinatenbereich. Daher sind $0 := O$ und $1 := \eta \cap EX$ in K. Die übrigen Elemente von K bezeichnen wir mit kleinen lateinischen Buchstaben. Für $a, b \in K$ definieren wir einen Punkt mit den Koordinaten (a, b), nämlich $P(a, b) = (aX \cap \delta)Y \cap bX$, d.h. wir schneiden die Parallele zu ξ durch a mit δ und ziehen durch den Schnittpunkt die Parallele zu η. Dann ist $P(a, b)$ der Schnittpunkt dieser Geraden mit der Parallelen zu ξ durch b.
Für $u, v \in K$ sei $\gamma_{u,v}$ die Gerade, welche durch v geht und zu $OP(1, u)$ parallel ist. Dann:

(1) Die Abbildung $K^2 \ni (a, b) \mapsto P(a, b)$ ist eine Bijektion von K^2 auf die Punkte von $\mathcal{A}$.

(2) Die Abbildung $K^2 \ni (u, v) \mapsto \gamma_{u,v}$ ist eine Bijektion von K^2 auf diejenigen Geraden von $\mathcal{A}$, welche nicht parallel zur y-Achse η sind.

(3) Der Schnittpunkt der Geraden $\gamma_{u,v}$ mit der Parallelen zu η durch $P(x, x)$ sei S. Dann hängt die y-Koordinate y_S von S ab von u, x und v, also $y_S = T(u, x, v)$ für eine Abbildung $T : K^3 \to K$. Der Punkt $P(x, y)$ liegt genau dann auf der Geraden $\gamma_{u,v}$, wenn $y = T(u, x, v)$ gilt. Außerdem hat diese Abbildung T die folgenden Eigenschaften:

(i) $T(u, 1, 0) = u$ für alle u.

(ii) $T(1, x, 0) = x$ für alle x.

(iii) $T(0,x,v) = v$ für alle x, v.

(iv) $T(u,0,v) = v$ für alle u, v.

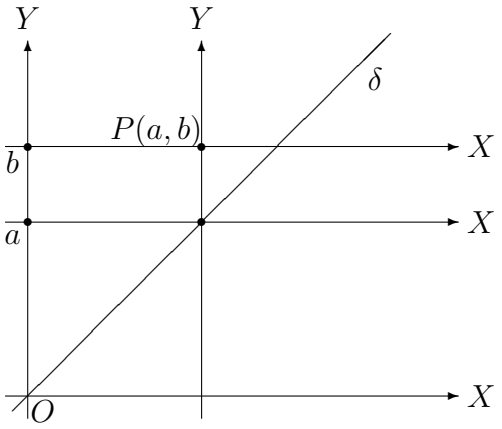
(v) Für alle u_1, u_2, v_1, v_2 gilt: wenn $u_1 \neq u_2$, dann gibt es genau ein x mit $T(u_1,x,v_1) = T(u_2,x,v_2)$.

(vi) Für alle x, y, u gibt es genau ein v mit $y = T(u,x,v)$.

(vii) Für alle x_1, x_2, y_1, y_2 gilt: wenn $x_1 \neq x_2$, dann gibt es u, v mit $T(u,x_1,v) = y_1$ und $T(u,x_2,v) = y_2$.

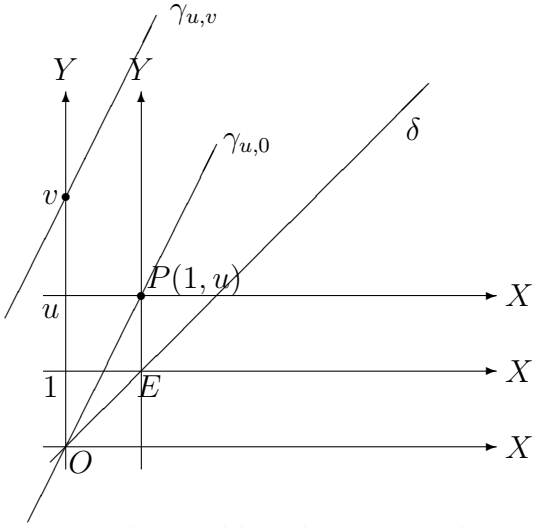
Beweis:

(1) Hier ist das Bild zur Konstruktion:



Gegeben ein Punkt P von $\mathcal{A}$; man findet y_P als Schnittpunkt der Parallele zu ξ durch P mit der y-Achse. Um x_P zu finden, schneidet man zunächst die Parallele durch P zu η mit δ und dann die Parallele zu ξ durch diesen Schnittpunkt mit der y-Achse. Es ist klar, dass $P \mapsto (x_P, y_P) \in K^2$ die Umkehrabbildung zu $K^2 \ni (a,b) \mapsto P(a,b)$ ist.
Offenbar sind die affinen Punkte auf der x-Achse gerade diejenigen mit $y_P = 0$. Ebenso lassen sich die Punkte auf der y-Achse durch $x_P = 0$ kennzeichnen und die auf der Diagonalen durch $x_P = y_P$. Insbesondere ist $O = P(0,0),\ E = P(1,1)$.

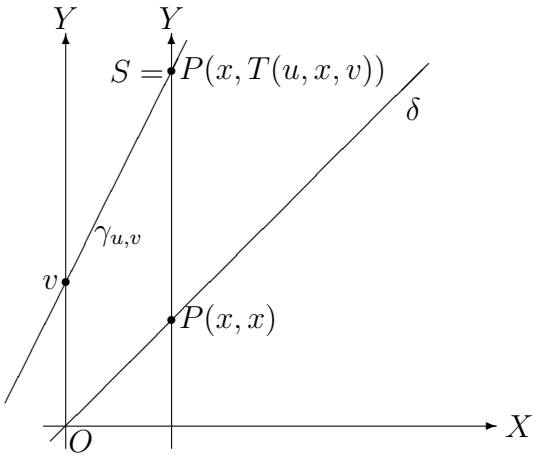
(2) Hier ist das Bild zur Konstruktion:



Zunächst ist klar, dass $\gamma_{u,v}$ nicht parallel zur y-Achse ist, denn sonst wäre auch $OP(1,u)$

parallel zu η. Da $OP(1,u)$ mit η den Punkt O gemeinsam hat, folgt Gleichheit. Dann läge aber $P(1,u)$ auf der y-Achse, ein Widerspruch, weil $1 \neq 0$.
Wenn γ eine nicht zu η parallele Gerade von $\mathcal{A}$ ist, dann ist $v_\gamma = \gamma \cap \eta$. Um u_γ zu finden, schneidet man die Parallele zu γ durch O mit der Parallelen zu η durch E. Die Koordinaten des Schnittpunkts sind $(1, u_\gamma)$. Es ist klar, dass $\gamma \mapsto (u_\gamma, v_\gamma) \in K^2$ die Umkehrabbildung zu $K^2 \ni (u,v) \mapsto \gamma_{u,v}$ ist.
Offenbar ist $\gamma_{u,v}$ parallel zur x-Achse genau dann, wenn $u = 0$, und parallel zur Diagonalen genau dann, wenn $u = 1$.

(3) Hier ist das Bild zur Konstruktion:



Die Charakterisierung der Punkte auf $\gamma_{u,v}$ ist klar. Zum Beweis der Eigenschaften von T:

(i) Die Gerade $\gamma_{u,0} = OP(1,u)$ schneidet sich mit der Parallelen zu y-Achse durch $P(1,1) = E$ in $P(1,u)$.

(ii) Es ist $\gamma_{1,0} = \delta$; darauf liegt der Punkt $P(x,x)$.

(iii) Es ist $\gamma_{0,v}$ die Parallele zur x-Achse durch v. Jeder Punkt auf dieser Geraden hat also v als y-Koordinate.

(iv) Der Schnittpunkt von $\gamma_{u,v}$ mit der y-Achse ist v.

(v) Weil $u_1 \neq u_2$, ist auch $OP(1,u_1) \neq OP(1,u_2)$; insbesondere sind die Geraden nicht parallel, denn sie haben ja den Punkt O gemeinsam. Also sind auch γ_{u_1,v_1} und γ_{u_2,v_2} nicht parallel. Sie schneiden sich daher in genau einem Punkt, dessen x-Koordinate dann die gewünschte Eigenschaft hat.

(vi) Durch den Punkt $P(x,y)$ gibt es genau eine Parallele zu $\gamma_{u,0}$. Deren Schnittpunkt mit der y-Achse ist das gesuchte v.

(vii) Weil $x_1 \neq x_2$, ist die Gerade durch $P(x_1,y_1)$ und $P(x_2,y_2)$ nicht parallel zur y-Achse. Sie ist daher nach (2) oben von der Form $\gamma_{u,v}$ für geeignetes (u,v).

17.15 Definition: *Ternärkörper*

Eine Menge K mit einer ternären Verknüpfung $T : K^3 \to K$ heißt Ternärkörper, wenn es in K zwei verschiedene Elemente 0 und 1 gibt, so dass gilt:

(i) $T(u,1,0) = u$ für alle u.

(ii) $T(1,x,0) = x$ für alle x.

(iii) $T(0, x, v) = v$ für alle x, v.

(iv) $T(u, 0, v) = v$ für alle u, v.

(v) Für alle u_1, u_2, v_1, v_2 gilt: wenn $u_1 \neq u_2$, dann gibt es genau ein x mit $T(u_1, x, v_1) = T(u_2, x, v_2)$.

(vi) Für alle x, y, u gibt es genau ein v mit $y = T(u, x, v)$.

(vii) Für alle x_1, x_2, y_1, y_2 gilt: wenn $x_1 \neq x_2$, dann gibt es u, v mit $T(u, x_1, v) = y_1$ und $T(u, x_2, v) = y_2$.

17.16 Beispiel/Bemerkung:

(i) Jeder Schiefkörper K liefert ein Beispiel für einen Ternärkörper, wenn man $T(u, x, v) = ux + v$ definiert.

(ii) Die Elemente 0 und 1 sind im Ternärkörper eindeutig bestimmt: Wenn auch 0' und 1' die geforderten Eigenschaften haben, dann gilt $1 = T(1, 1', 0) = 1'$ nach den Axiomen (i) und (ii). Nach den Axiomen (i) und (iii) ist $0' = T(0', 1, 0) = 0$.

(iii) Im letzten Satz haben wir gezeigt, wie sich zu einer projektiven Ebene ein Ternärkörper konstruieren lässt. Dieser hängt auch ab von der Wahl der Punkte O, E, X, Y.

(iv) Umgekehrt sei ein Ternärkörper K gegeben. Wir konstruieren dann eine affine Ebene $\mathcal{A} = \mathcal{A}(K^2)$ wie folgt: $\mathcal{P} = K^2$, d.h. die Punkte von $\mathcal{A}$ sind alle Paare (x, y) mit $x, y \in K$. Geraden sind bestimmte Teilmengen von K^2, nämlich für jedes $c \in K$ die Menge $\eta_c = \{(c, y) \mid y \in K\}$ und für jedes $(u, v) \in K^2$ die Menge $\gamma_{u,v} = \{(x, T(u, x, v)) \mid x \in K\}$. Die Inzidenz ist definiert durch: Ein Punkt P liegt auf einer Geraden γ, wenn $P \in \gamma$.
Es gibt dann drei nicht kollineare Punkte, z.B. kann man $(0, 0)$, $(0, 1)$ und $(1, 0)$ nehmen; denn weil $0 \neq 1$, liegen diese Punkte nicht auf einer Geraden η_c. Sie liegen auch nicht auf einem $\gamma_{u,v}$, weil dann zugleich $T(u, 0, v) = 0$ und $T(u, 0, v) = 1$ sein müsste.
Wenn $c \neq d$, dann haben η_c und η_d offenbar keinen Punkt gemeinsam, sind also parallel. Kein η_c ist parallel zu einem $\gamma_{u,v}$, denn $(c, y) \in \gamma_{u,v}$ genau dann, wenn $y = T(u, c, v)$. Also haben diese beiden Geraden genau einen Schnittpunkt. Wenn zwei verschiedene Geraden γ_{u_1,v_1} und γ_{u_2,v_2} gegeben sind, gibt es zwei Möglichkeiten: Wenn $u_1 = u_2 = u$, dann ist $v_1 \neq v_2$. In diesem Fall gibt es keinen Schnittpunkt, denn sonst wäre für ein (x, y) zugleich $T(u, x, v_1) = y$ und $T(u, x, v_2) = y$ im Widerspruch zu Axiom (vi). Die Geraden sind dann also parallel. Wenn dagegen $u_1 \neq u_2$, dann folgt aus Axiom (v), dass sich die Geraden in genau einem Punkt schneiden. Insgesamt ergibt sich, dass zwei verschiedene Geraden höchstens einen Punkt gemeinsam haben. Außerdem haben wir gezeigt, dass die η_c's alle zueinander parallel sind und ebenso die $\gamma_{u,v}$'s mit festem u, aber keine anderen Geraden. Da durch jeden Punkt ein η_c geht und zu gegebenem u nach Axiom (vi) auch ein $\gamma_{u,v}$, ist damit das zweite Axiom für affine Ebenen verifiziert.
Schließlich gibt es durch je zwei Punkte (x_1, y_1) und (x_2, y_2) eine Gerade: Wenn $x_1 = x_2 = c$, dann tut es η_c. Im anderen Fall folgt die Existenz einer geeigneten Geraden $\gamma_{u,v}$ aus Axiom (vii).
Damit ist gezeigt, dass $\mathcal{A}$ eine affine Ebene ist. Wie in 17.6 (iv) beschrieben, kann man $\mathcal{A}$ zu einer projektiven Ebene erweitern.
Aus 17.14 folgt, dass jede affine Ebene isomorph zu $\mathcal{A}(K^2)$ für einen geeigneten Ternärkörper ist.

(v) Diese Konstruktion führt auf das in 17.5 (i) betrachtete Beispiel $\mathcal{A}(V)$ mit dem Rechts-Vektorraum $V = K^2$, wenn der Ternärkörper K ein Schiefkörper ist: seien e_1, e_2 die Einheitsvektoren in K^2. Dann ist $\eta_c = e_1c + e_2K$ und $\gamma_{u,v} = e_2v + (1,u)K$, also sind die Geraden Nebenklassen von eindimensionalen Unterräume von K^2. Man bekommt sogar alle Nebenklassen: seien $v \in K^2$ und U ein eindimensionaler Unterraum, etwa $v = (v_1, v_2)$ und $U = (u_1, u_2)K$. Wenn $u_1 \neq 0$, dann ist $U = (1,u)K$ mit $u = u_2u_1^{-1}$ und $v + U = v - (1,u)v_1 + U = \gamma_{u,v_3}$ mit $v_3 = v_2 - uv_1$. Wenn $u_1 = 0$, dann ist $u_2 \neq 0$ und $U = e_2K$; außerdem ist $v + U = v - e_2v_2 + U = e_1v_1 + e_2K = \eta_{v_1}$.

(vi) Isomorphe Ternärkörper (mit der offenkundigen Definition) führen zu isomorphen affinen Ebenen, und dann sind natürlich auch die zugehörigen projektiven Ebenen isomorph. Umgekehrt sei $\alpha : \mathcal{E} \to \mathcal{E}'$ ein Isomorphismus von projektiven Ebenen. Wenn O, X, Y, E Punkte von $\mathcal{E}$ in allgemeiner Lage sind, dann sind auch $O\alpha, X\alpha, Y\alpha, E\alpha$ in allgemeiner Lage, und die zugehörigen Ternärkörper sind isomorph.

(vii) Algebraische Eigenschaften der Ternärkörper lassen sich im geometrische Eigenschaften der projektiven Ebene übersetzen und umgekehrt. Ein klassisches Beispiel dafür werden wir im nächsten Kapitel untersuchen.

(viii) Wir werden Punkte einer affinen Ebene künftig kurz durch ihre Koordinaten beschreiben; für $x, y \in K$ ist also (x, y) der Punkt mit diesen Koordinaten. Ebenso behalten wir die Bezeichnungen für die Geraden bei. Sinnvoll ist das aber erst nach Wahl der Punkte O, X, Y, E.

17.17 Lemma:

In einem Ternärkörper gilt $(a = 0$ oder $b = 1) \Leftrightarrow T(a,b,0) = a \Leftrightarrow T(b,a,0) = a$.

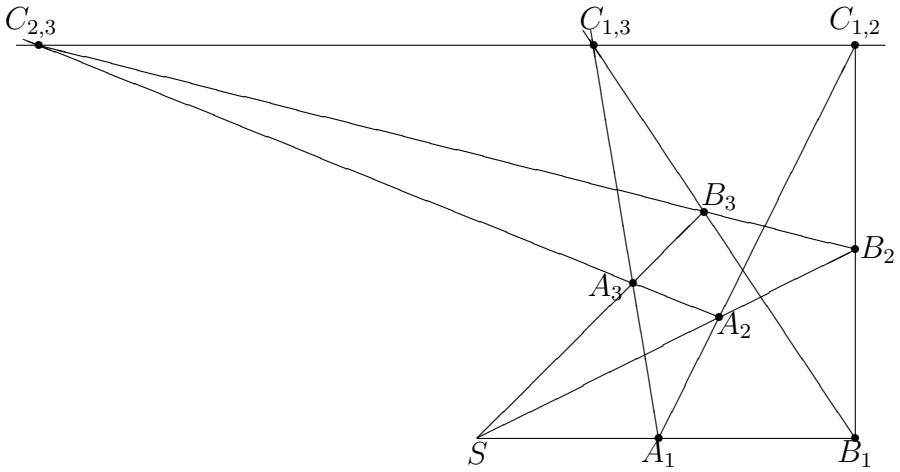
Beweis:

Wenn $b = 1$, dann ist $T(a,b,0) = T(a,1,0) = a$ nach 17.15 (i) und $T(b,a,0) = T(1,a,0) = a$ nach 17.15 (ii).
Wenn $a = 0$, dann ist $T(a,b,0) = T(0,b,0) = 0 = a$ nach 17.15 (iii) und $T(b,a,0) = T(b,0,0) = 0 = a$ nach 17.15 (iv).
Wenn $b \neq 1$, dann gibt es nur eine Gerade durch $(1,a)$ und (b,a), nämlich $\gamma_{0,a}$. Wenn $T(a,b,0) = a$, dann geht auch $\gamma_{a,0}$ durch beide Punkte, also $a = 0$.
Wenn $a \neq 0$, dann gibt es nur eine Gerade durch $(0,0)$ und (a,a), nämlich $\delta = \gamma_{1,0}$. Wenn $T(b,a,0) = a$, dann geht auch $\gamma_{b,0}$ durch beide Punkte, also $b = 1$.

18 Desargues'sche Ebenen

18.1 Definition: *Desargues'sche Ebenen*

Eine projektive Ebene heißt Desargues'sch, wenn das folgende gilt:
Zu je drei verschiedenen Geraden γ_1, γ_2, γ_3, welche sich in einem Punkt S schneiden, und je zwei weiteren Punkte $S \neq A_i \neq B_i \neq S$ auf γ_i sei $C_{1,2} = A_1A_2 \cap B_1B_2$ und entsprechend seien $C_{1,3}$ und $C_{2,3}$ definiert. Dann gibt es eine Gerade durch $C_{1,2}$, $C_{1,3}$ und $C_{2,3}$.



Eine affine Ebene nennt man Desargues'sch, wenn die zugehörige projektive Ebene dies ist.

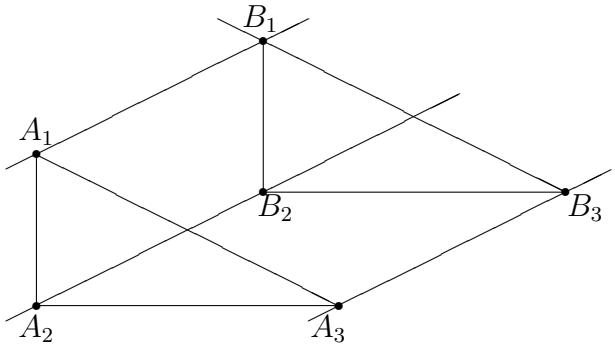
18.2 Bemerkung:

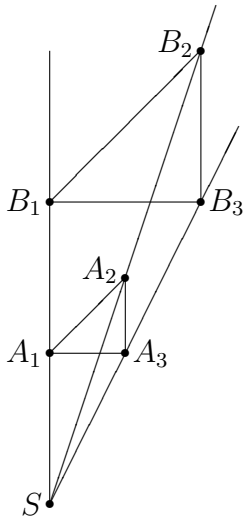
(i) Im Bild oben ist die Gerade durch $C_{1,2}$, $C_{1,3}$ und $C_{2,3}$ parallel zu γ_1. Das ist *kein* Teil der Definition (und ergibt in projektiven Ebenen keinen Sinn)!

(ii) Die Aussage, dass die drei Punkte $C_{1,2}$, $C_{1,3}$, $C_{2,3}$ auf einer Geraden liegen, ist natürlich dann trivial, wenn sie nicht alle verschieden sind. In diesem Fall, also wenn etwa $C_{1,2} = C_{1,3} = T$, dann ist $A_1A_2 = A_1T = A_1A_3$, also sind A_1, A_2, A_3 kollinear und ebenso B_1, B_2, B_3. Es fallen dann also alle $C_{i,j}$ zusammen.

(iii) Die zu einer Desargues'schen projektiven Ebene duale Ebene ist wieder Desargues'sch (Übungsaufgabe).

(iv) Die drei Geraden der Definition können in einer affinen Ebene drei verschiedene Parallelen sein, denn diese schneiden sich ja in einem uneigentlichen Punkt. Wenn die affine Ebene Desargues'sch ist und etwa $A_1A_2 \,||\, B_1B_2$ und $A_1A_3 \,||\, B_1B_3$, dann liegen die Schnittpunkte $C_{1,2}$, $C_{1,3}$ auf der uneigentlichen Geraden. Also liegt auch $C_{2,3}$ auf der uneigentlichen Geraden, d.h. auch $A_2A_3 \,||\, B_2B_3$. Dieser und ein ähnlicher Spezialfall sind hier veranschaulicht:





18.3 Lemma:

Seien $S =< s >$, A, B drei verschiedene eindimensionale Unterräume des zweidimensionalen Vektorraums V. Dann gibt es eine Basis $\{s, a\}$ von V mit $A =< a >$ und $B =< a + s >$.

Beweis: (für Links-Vektorräume) Sei $0 \neq a_0 \in A$, also $A =< a_0 >$. Weil $A \neq S$, sind a_0 und s linear unabhängig, bilden also eine Basis von V. Insbesondere läßt sich $0 \neq b_0 \in B$ als Linearkombination $b_0 = k_1 a_0 + k_2 s$ schreiben. Dabei sind die Koeffizienten $k_1, k_2 \in K$ beide von 0 verschieden, da sonst $B = A$ oder $B = S$ wäre. Setzt man $b = k_2^{-1} b_0$ und $a = k_2^{-1} k_1 a_0$, dann ist $A =< a >$, $B =< b >$ und $b = a + s$.

18.4 Beispiel: *für Desargues'sche Ebenen*

Sei W ein dreidimensionaler Vektorraum W über einem Schiefkörper. Dann ist $\mathcal{E}(W)$ (vergleiche 17.5) Desargues'sch: Sei S der gemeinsame Schnittpunkt der drei Geraden V_1, V_2, V_3 (dies sind also zweidimensionale Unterräume von W, die Punkte sind eindimensionale Unterräume). Wähle die Basen $\{s, a_i\}$ von V_i wie im Lemma. Dann sind $A_1A_2 =< a_1, a_2 >$ und $B_1B_2 =< a_1 + s, a_2 + s >$ die Verbindungsgeraden. Ihr Schnittpunkt ist $C_{1,2} =< a_1 - a_2 >$. Ebenso findet man $C_{1,3} =< a_1 - a_3 >$ und $C_{2,3} =< a_2 - a_3 >$. Da die drei Vektoren $a_1 - a_2$, $a_1 - a_3$ und $a_2 - a_3 = (a_1 - a_3) - (a_1 - a_2)$ linear abhängig sind, liegen sie in einem zweidimensionalen Unterraum von W; also liegen $C_{1,2}, C_{1,3}, C_{2,3}$ auf einer Geraden.
Im folgenden wollen wir zeigen, dass umgekehrt alle Desargues'schen projektiven Ebenen von dreidimensionalen Vektorräumen herrühren.

18.5 Satz: *Translationen*

Seien $P \neq Q$ zwei Punkte einer Desargues'schen affinen Ebene $\mathcal{A}$. Dann gibt es genau eine Kollineation $\tau = \tau(P, Q)$ von $\mathcal{A}$ auf $\mathcal{A}$ mit den folgenden Eigenschaften:

(1) $P\tau = Q$

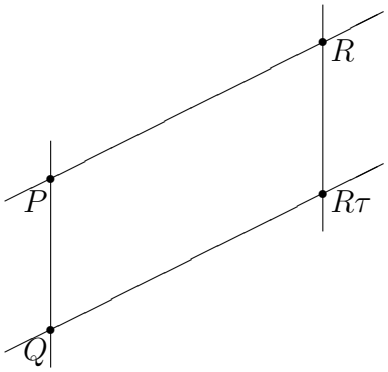
(2) $\gamma\tau \,||\, \gamma$ für alle Geraden γ

(3) $\gamma\tau = \gamma$, wenn $\gamma \,||\, (PQ)$

Man nennt τ die Translation von P nach Q; diese hat keine Fixpunkte.

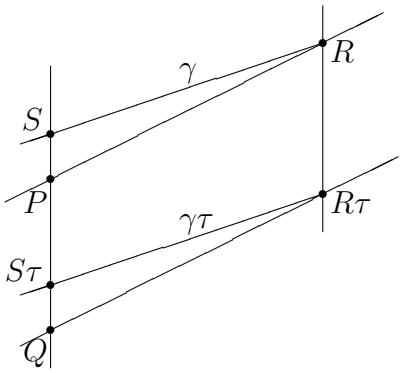
Beweis:

Eindeutigkeit:

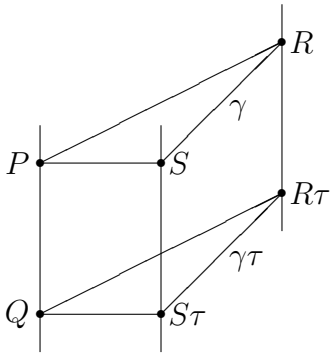


Sei R nicht auf PQ; weil τ eine Kollineation ist, muss $(PR)\tau = (P\tau)(R\tau) = Q(R\tau)$ sein; nach Bedingung (2) ist also $R\tau$ auf der Parallelen zu PR durch Q. Andererseits muss $R\tau$ auf der Parallelen zu PQ durch R liegen, denn diese bleibt nach (3) fest. Es ist klar, dass $R\tau \neq R$ und dass $R(R\tau)$ parallel zu PQ ist.

Wenn S auf PQ liegt, dann wähle einen Punkt R, der nicht auf dieser Geraden liegt. Die Verbindungsgerade $\gamma = RS$ wird dann durch τ auf die Parallele zu γ durch $R\tau$ abgebildet. Da PQ fest bleibt, ist $S\tau$ als Schnittpunkt von PQ und $\gamma\tau$ eindeutig bestimmt. Wieder ist klar, dass $S\tau \neq S$.

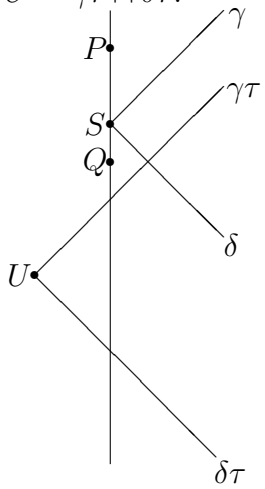


Wir haben also gezeigt, dass es höchstens eine Translation von P nach Q gibt und dass diese keine Fixpunkte haben kann.
Um die Existenz von τ zu zeigen, orientieren wir uns an der obigen Konstruktion; wieder wird zunächst der Fall betrachtet, dass R nicht auf PQ liegt. Wie man $R\tau$ findet, ist schon beschrieben. Wenn γ eine Gerade durch R ist, definiert man $\gamma\tau$ als die Parallele zu γ durch $R\tau$. Wir müssen zeigen, dass für jeden Punkt S auf γ dann $S\tau$ auf $\gamma\tau$ liegt. Das ist klar, wenn $\gamma \,||\, (PQ)$, denn dann $\gamma\tau = \gamma = S(S\tau)$. Sonst betrachten wir



Dabei ist o.E. $S \neq R$ und wieder vorausgesetzt, dass S nicht auf PQ liegt. Die drei Geraden PQ, $S(S\tau)$ und $R(R\tau)$ sind dann parallel und verschieden; nach Konstruktion ist auch $PR \,||\, Q(R\tau)$ und $PS \,||\, Q(S\tau)$. Da die Ebene Desargues'sch ist, folgt $RS \,||\, (R\tau)(S\tau)$.

Wir haben jetzt τ mit den gewünschten Eigenschaften auf allen Geraden definiert und auf allen Punkten, welche nicht auf PQ liegen. Auf diesen Punkten ist τ übrigens eine Bijektion; die Umkehrabbildung findet man, indem man die Rollen von P und Q vertauscht. Auch für einen Punkt S auf PQ liegt schon fest, was $S\tau$ ist: wähle $\gamma \neq PQ$ durch S und setze $S\tau = PQ \cap \gamma\tau$. Wir müssen noch zeigen, dass dies wohldefiniert, also nicht abhängig von der Wahl von γ ist. Dazu sei $\gamma \neq \delta \neq PQ$ eine weitere Gerade durch S und sei $U = \gamma\tau \cap \delta\tau$.



Wäre U nicht auf PQ, so wäre $U = V\tau$ für einen Punkt V, welcher nicht auf PQ liegt. Wenn γ_1 die Parallele zu γ durch V, dann ist $\gamma_1\tau$ die Parallele zu γ durch $V\tau = U$, also $\gamma_1\tau = \gamma\tau$ und daher $\gamma_1 = \gamma$. Also geht γ durch V; ebenso findet man, dass V auch auf δ liegt. Aber dann ist $V = S$, ein Widerspruch, denn S liegt auf PQ, aber V nicht. Damit ist gezeigt, dass $U = PQ \cap \gamma\tau = PQ \cap \delta\tau$ unabhängig von der Wahl von γ ist.

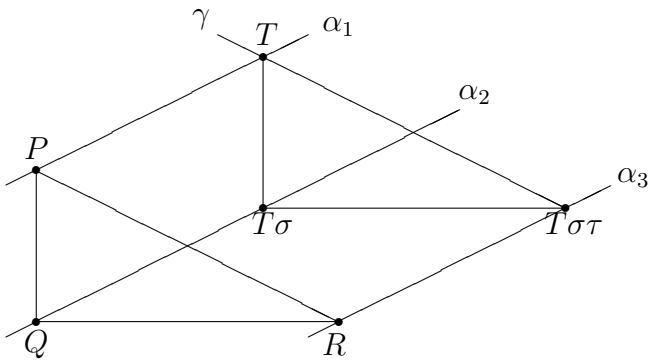
18.6 Korollar:

Mit den Bezeichnungen des Satzes gilt für jeden Punkt R, welcher nicht auf PQ liegt, dass $R\tau$ der Schnittpunkt der Parallelen zu PQ durch R mit der Parallelen zu PR durch Q ist.

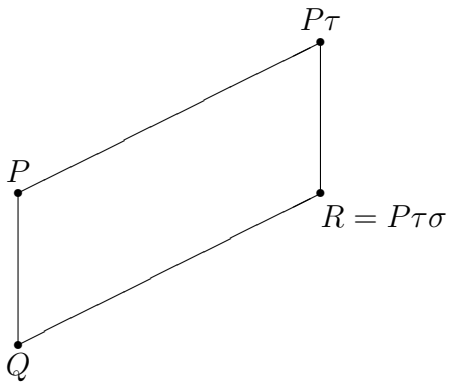
18.7 Satz: *Translationsgruppe*

Sei $\mathcal{A}$ eine Desargues'schen affinen Ebene. Die sämtlichen Translationen zusammen mit der Identität bilden eine abelsche Gruppe, die Translationsgruppe. (Die Gruppenverknüpfung ist die übliche Verknüpfung von Abbildungen.)

Beweis: Seien σ und τ Translationen. Wenn $P = P\sigma\tau$ für jeden Punkt P gilt, dann ist $\sigma\tau = id$. Andernfalls wähle P mit $P \neq P\sigma\tau$, etwa $P\sigma = Q$ und $Q\tau = R$. Wir zeigen, dass $\sigma\tau$ die Translation von P nach R ist. Dabei sind die beiden ersten Eigenschaften in 18.5 klar. Es bleibt zu kontrollieren, dass $\gamma\sigma\tau = \gamma$ für jede Gerade $\gamma \,||\, PR$ gilt. Wenn P, Q, R kollinear sind, dann lassen σ und τ beide jedes solche γ fest. Wir können also annehmen, dass P, Q, R nicht kollinear sind. Es genügt, einen Punkt T auf γ zu finden, für den auch $T\sigma\tau$ auf γ liegt, denn $\gamma\sigma\tau \,||\, \gamma$ und $\gamma\sigma\tau$ geht durch $T\sigma\tau$, also $\gamma\sigma\tau = \gamma$. Für $\gamma = PR$ kann man $T = P$ nehmen, also o.E. $\gamma \neq PR$. Wenn es nur zwei Parallelen in der Parallelenschar von PR gibt, ist notwendigerweise auch $\gamma\sigma\tau = \gamma$ und wir sind fertig. Im anderen Fall gibt es auf γ mindestens drei Punkte (vergleiche 17.9 (6)), also gibt es T auf γ, welches nicht auf PQ und nicht auf der Parallelen zu QR durch P liegt.



Nach Konstruktion sind $PQ \,||\, T(T\sigma)$, $\alpha_1 = PT \,||\, Q(T\sigma) = \alpha_2$, $QR \,||\, (T\sigma)(T\sigma\tau)$ und $Q(T\sigma) \,||\, R(T\sigma\tau) = \alpha_3$. Die drei Parallelen $\alpha_1, \alpha_2, \alpha_3$ sind verschieden: $\alpha_1 \neq \alpha_2$, weil sonst T, P, Q kollinear wären, entgegen der Wahl von T; auch $\alpha_1 \neq \alpha_3$, weil sonst T, P, R kollinear wären, also T auf PR und dann $\gamma = PR$; schließlich ist $\alpha_2 \neq \alpha_3$, denn sonst ist $\alpha_2 = QR$ und dann $\alpha_1 \,||\, QR$, wieder entgegen der Wahl von T. Außerdem sind $P \neq T$, $Q \neq T\sigma$, $R \neq T\sigma\tau$. Da die Ebene Desargues'sch ist, folgt $T(T\sigma\tau) \,||\, PR$. Da auch $\gamma \,||\, PR$ und T auf γ liegt, ist $\gamma = T(T\sigma\tau)$; insbesondere liegt $T\sigma\tau$ auf γ, was zu zeigen war.
Wir haben gezeigt, dass das Produkt von zwei Translationen entweder $\mathrm{id}_{\mathcal{A}}$ oder eine Translation ist. Sei σ die Translation von P nach Q und τ die Translation von Q nach P, dann ist $P = P\sigma\tau$ ein Fixpunkt. Nach 18.5 ist $\sigma\tau$ keine Translation und daher $\sigma\tau = \mathrm{id}_{\mathcal{A}}$. Das zeigt die Existenz von Inversen. Bekanntlich ist die Hintereinander-Ausführung von Abbildungen assoziativ. Die Translation bilden also eine Gruppe. Um zu zeigen, dass diese Gruppe abelsch ist, sei wieder $P\sigma = Q$ und $Q\tau = R$
Wir nehmen zunächst P, Q, R nicht kollinear an. Dann ist $\sigma\tau$ die Translation von P nach R, wie gerade gezeigt. Aber auch $P\tau\sigma = R$:



Da τ die Translation von Q nach R ist, erhält man $P\tau$, indem man die Parallele zu QR durch P mit der Parallelen zu PQ durch R schneidet. Wendet man auf diesen Punkt jetzt σ an, so erhält man offenbar R. Daher ist auch $\tau\sigma$ die Translation von P nach R, d.h. $\tau\sigma = \sigma\tau$.

Wenn P, Q, R kollinear sind, sei S ein Punkt, der nicht auf PQ liegt, σ_1 die Translation von P nach S und σ_2 die Translation von S nach Q. Dann ist $\sigma = \sigma_1\sigma_2$, also

$$\sigma\tau = (\sigma_1\sigma_2)\tau = \sigma_1(\sigma_2\tau) = \sigma_1(\tau\sigma_2) = (\sigma_1\tau)\sigma_2 = (\tau\sigma_1)\sigma_2 = \tau(\sigma_1\sigma_2) = \tau\sigma \ ,$$

denn σ_1 und σ_2 kommutieren mit τ, wie schon gezeigt. Daher kommutieren alle Translationen miteinander.

18.8 Definition:

Sei $(K, T, 0, 1)$ ein Ternärkörper. Wir definieren dann zwei binäre Verknüpfungen, genannt Addition und Multiplikation, auf K durch

$$a + b = T(1, a, b) \quad \text{und} \quad ab = T(a, b, 0) \ .$$

18.9 Bemerkung:

Nach den Ternärkörper-Axiomen verhalten sich 0 und 1 erwartungsgemäß, d.h. $a + 0 = 0 + a = a1 = 1a = a$ und $0a = a0 = 0$ für alle $a \in K$.

18.10 Voraussetzung: *für 18.11 bis 18.13*

O, X, Y, E sind vier Punkte in allgemeiner Lage in einer Desargues'schen projektiven Ebene $\mathcal{E}$. Der zugehörige Ternärkörper ist $(K, T, 0, 1)$, und die Punkte der affinen Ebene $\mathcal{A} = \mathcal{E} \setminus XY$ sind wieder durch ihre Koordinaten (x, y), die Geraden als $\gamma_{u,v}$ beziehungsweise als η_c (die Parallelen zur y-Achse) bezeichnet.

18.11 Satz:

Sei $0 \neq a \in K$ und σ beziehungsweise τ die Translationen von O nach $(0, a)$ beziehungsweise nach $(a, 0)$. Dann gelten:

(1) Für alle $x, y \in K$ ist $(x, y)\sigma = (x, y + a)$.

(2) $T(u, x, a) = ux + a$ für alle $u, x, a \in K$.

(3) $(K, +)$ ist eine abelsche Gruppe.

(4) Für alle $x, y \in K$ ist $(x, y)\tau = (x + a, y)$.

(5) Es gilt das Links-Distributivgesetz $u(x + a) = ux + ua$ für alle $u, x, a \in K$.

Beweis:

(1) Sei $0 \neq y \in K$. Die Parallele zu $O(y, y) = \delta = \gamma_{1,0}$ durch $(0, a)$ ist $\gamma_{1,a}$. Die Parallele zu $O(0, a) = \eta = \eta_0$ durch (y, y) ist η_y. Weil $(y, y)\sigma$ der Schnittpunkt dieser beiden Geraden ist nach 18.6, folgt $(y, y)\sigma = \eta_y \cap \gamma_{1,a} = (y, T(1, y, a)) = (y, y + a)$. Die Gerade $\gamma_{0,y}\sigma$ geht durch $(y, y)\sigma = (y, y + a)$ und ist parallel zu $\gamma_{0,y}$, also ist $\gamma_{0,y}\sigma = \gamma_{0,y+a}$. Aus $(x, y) = \eta_x \cap \gamma_{0,y}$ folgt jetzt $(x, y)\sigma = \eta_x\sigma \cap \gamma_{0,y}\sigma = \eta_x \cap \gamma_{0,y+a} = (x, y + a)$.

(2) Da $\gamma_{u,0}\sigma$ parallel zu $\gamma_{u,0}$ ist und durch $(0, a)$ geht, ist $\gamma_{u,0}\sigma = \gamma_{u,a}$. Da (x, ux) auf $\gamma_{u,0}$ liegt, liegt $(x, ux)\sigma = (x, ux + a)$ auf $\gamma_{u,a}$, also ist $T(u, x, a) = ux + a$.

(3) Diejenigen Elemente der Translationsgruppe (vergleiche 18.7), welche die y-Achse fest lassen, bilden eine Untergruppe A. Die Abbildung $A \ni \sigma \mapsto O\sigma$ ist eine Bijektion von A auf K. Weil $O\sigma_1 + O\sigma_2 = O\sigma_1\sigma_2$ (nach (1)) ist, folgt die Behauptung.

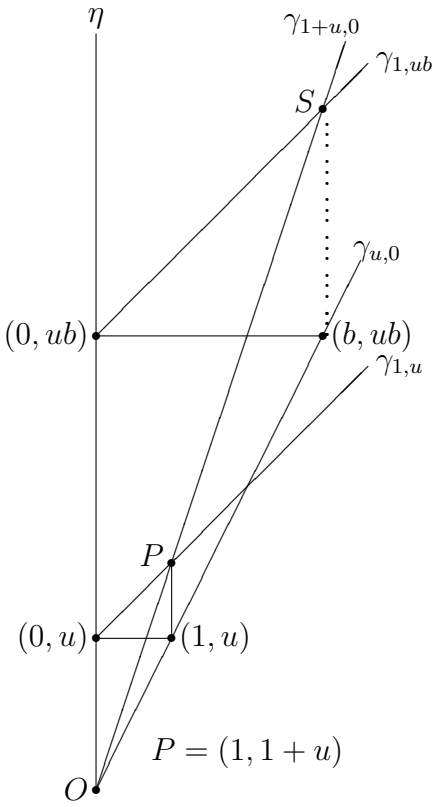
(4) Es ist $\gamma_{1,0}\tau = \gamma_{1,b}$ für ein geeignetes b, denn die beiden Geraden sind ja parallel. Weil O auf $\gamma_{1,0}$, ist $(a, 0) = O\tau$ auf $\gamma_{1,0}\tau = \gamma_{1,b}$ und daher $0 = T(1, a, b) = a + b$, d.h. $b = -a$. Weil τ die Parallelen zur x-Achse festläßt, ist $(x, x)\tau = (x_1, x)$ für ein x_1, und da dieser Punkt auf $\gamma_{1,-a}$ liegt, folgt $x = T(1, x_1, -a) = x_1 - a$, also $(x, x)\tau = (x + a, x)$. Daher ist $\eta_x\tau = \eta_{x+a}$ und damit schließlich $(x, y)\tau = (x + a, y)$.

(5) Es ist $(0, ua)$ auf $\gamma_{u,ua}$, also $(0, ua)\tau = (a, ua)$ auf $\gamma_{u,ua}\tau$. Offenkundig liegt (a, ua) auch auf $\gamma_{u,0}$. Da diese beiden Geraden parallel sind, folgt Gleichheit, d.h. $\gamma_{u,ua}\tau = \gamma_{u,0}$. Da $T(u, x, ua) = ux + ua$ nach (2), liegt $(x, ux + ua)$ auf $\gamma_{u,ua}$. Wendet man τ an, erhält man, dass $(x + a, ux + ua)$ auf $\gamma_{u,0}$ liegt, also $ux + ua = T(u, x + a, 0) = u(x + a)$.

18.12 Lemma:

Es gilt $(1 + u)b = b + ub$ für alle $u, b \in K$.

Beweis: Die Behauptung ist richtig, wenn $u = 0$ oder $b = 0$ oder $b = 1$. Wir nehmen daher $u \neq 0 \neq b \neq 1$ an und betrachten



Definiere $S = \gamma_{1+u,0} \cap \gamma_{1,ub}$. Auf den drei Geraden durch O liegen jeweils noch zwei verschiedene Punkte, denn $u \neq ub$ nach 17.17. Da die Gerade durch $(0, u)$ und $(1, u)$ parallel zur Geraden durch $(0, ub)$ und (b, ub) ist und ebenso die Gerade durch $(0, u)$ und P parallel zur Geraden durch $(0, ub)$ und S, ist auch die Gerade durch $(1, u)$ und P (und damit die y-Achse η) parallel zur Geraden durch (b, ub) und S. Daher ist b die x-Koordinate von S. Weil S auf $\gamma_{1,ub}$ und auf $\gamma_{1+u,0}$ liegt, ist die y-Koordinate von S also $b + ub = (1 + u)b$, wie behauptet.

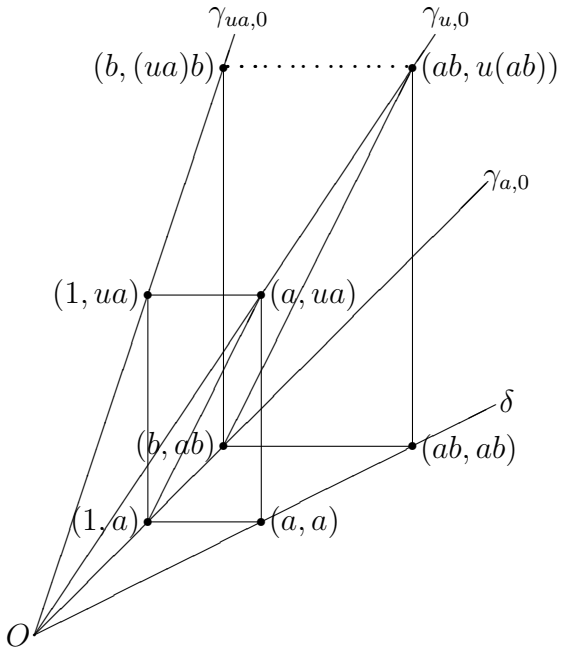
18.13 Satz:

$K^* = K \setminus \{0\}$ ist bezüglich der Multiplikation eine Gruppe mit 1 als neutralem Element.

Beweis: Wenn $u \neq 0 \neq a$, dann ist $ua \neq 0$, denn sonst läge $P = (a, ua)$ auf der x-Achse. Aber P liegt auch auf $\gamma_{u,0}$; daher $P = (0, 0)$, ein Widerspruch. Dass $u1 = u = 1u$, ist klar nach 18.9.

Die Existenz des multiplikativen Inversen kann man wie folgt zeigen: Verwendet man 17.15 (v) mit $u_1 = u \neq 0 = u_2$, $v_1 = 0$ und $v_2 = 1$, so erhält man ein x mit $ux = T(u, x, 0) = T(0, x, 1) = 1$. Daher gibt es zu u ein Rechtsinverses. Verwendet man 17.15 (vii) mit $x_1 = u \neq 0 = x_2$, $y_1 = 1$ und $y_2 = 0$, so erhält man s und r mit $T(s, u, r) = 1$ und $T(s, 0, r) = 0$. Aus der zweiten Gleichung folgt $r = 0$ und damit aus der ersten Gleichung $su = 1$. Daher gibt es zu u auch ein Linksinverses. Dass diese beiden einseitigen Inversen gleich sind, folgt dann aus der Assoziativität (die wir gleich zeigen werden): $s = s1 = s(ux) = (su)x = 1x = x$.

Wir müssen also noch $u(ab) = (ua)b$ für alle $u, a, b \in K^*$ zeigen; diese Gleichung gilt dann sogar in K, denn wenn einer der Faktoren 0 ist, dann auch das Produkt, egal wie geklammert. Die Behauptung ist klar, wenn einer der Faktoren 1 ist; wir nehmen daher das Gegenteil an. Nach 17.17 gilt dann $a \neq ab,\ ua \neq u$ und $ua \neq a$. Außerdem sind alle Produkte $\neq 0$.
Es folgt, dass (a, a) und (ab, ab) auf $\gamma_{1,0} = \delta$, beziehungsweise $(1, a)$ und (b, ab) auf $\gamma_{a,0}$, beziehungsweise (a, ua) und $(ab, u(ab))$ auf $\gamma_{u,0}$, beziehungsweise $(1, ua)$ und $(b, (ua)b)$ auf $\gamma_{ua,0}$ jeweils zwei voneinander und vom Schnittpunkt O dieser Geraden verschiedene Punkte sind.
Wenn $u \neq a$, dann sind $\delta,\ \gamma_{u,0}$ und $\gamma_{a,0}$ drei verschiedene Geraden durch O.



Weil

$$(1, a)(a, a) \,||\, (b, ab)(ab, ab) \quad \text{und} \quad (a, a)(a, ua) \,||\, (ab, ab)(ab, u(ab)) \ ,$$

ist auch

$$(1, a)(a, ua) \,||\, (b, ab)(ab, u(ab)) \ .$$

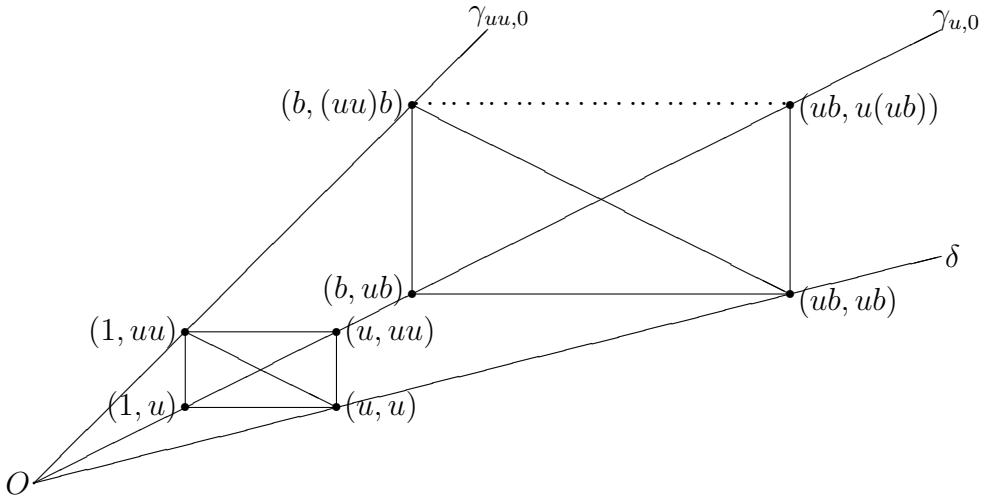
Ebenso sind $\gamma_{u,0}$, $\gamma_{a,0}$ und $\gamma_{ua,0}$ drei verschiedene Geraden durch O. Weil

$$(1, a)(a, ua) \,||\, (b, ab)(ab, u(ab)) \quad \text{und} \quad (1, a)(1, ua) \,||\, (b, ab)(b, (ua)b)) \ ,$$

ist auch

$$(1, ua)(a, ua) \,||\, (b, (ua)b)(ab, u(ab)) \ .$$

Aber das heißt gerade, dass diese beiden letzten Punkte die gleiche y-Koordinate haben, also $(ua)b = u(ab)$.
Wenn dagegen $u = a$, aber $uu \neq 1$, dann sind die drei Geraden $\delta,\ \gamma_{u,0}$ und $\gamma_{uu,0}$ verschieden. Auf ihnen liegen die Punkte (u, u) und (ub, ub), beziehungsweise $(1, u)$ und (u, uu) sowie (b, ub) und $(ub, u(ub))$, beziehungsweise $(1, uu)$ und $(b, (uu)b)$, die jeweils paarweise voneinander und alle von O verschieden sind (der Fall $ub = 1$ ist möglich).



Nun ist

$$(1,u)(1,uu) \,||\, (b,ub)(b,(uu)b) \quad \text{und} \quad (1,u)(u,u) \,||\, (b,ub)(ub,ub) \ ,$$

also auch

$$(u,u)(1,uu) \,||\, (ub,ub)(b,(uu)b) \ .$$

Dies zusammen mit

$$(u,u)(u,uu) \,||\, (ub,ub)(ub,u(ub))$$

impliziert

$$(1,uu)(u,uu) \,||\, (b,(uu)b)(ub,u(ub)) \ ,$$

also wieder die Behauptung.
Es bleibt der Fall $u \neq 1 = uu$. Man kann auch wieder geometrisch argumentieren; kürzer ist es, zu rechnen: $u(1+u) = u + uu = u + 1 = 1 + u$, denn das Links-Distributivgesetz gilt ja, und die Addition ist kommutativ. Nach 17.17 folgt $1 + u = 0$. Das Rechts-Distributivgesetz haben wir im Spezialfall in 18.12 gezeigt. Es folgt $0 = (1+u)b = b + ub$ und daher $ub = -b$. Daraus schließt man $u(ub) = -(-b) = b = (uu)b$.

18.14 Satz: *Koordinaten-Schiefkörper*

Seien O, X, Y, E vier Punkte in allgemeiner Lage in einer projektiven Ebene $\mathcal{E}$. Genau dann ist der Koordinaten-Ternärkörper $K = K(O, E, X, Y)$ ein Schiefkörper, wenn $\mathcal{E}$ eine Desargues'sche Ebene ist. In diesem Fall ist $\mathcal{E} \cong \mathcal{E}(W)$ für einen dreidimensionalen K-Vektorraum W.

Beweis: Sei $\mathcal{E}$ Desargues'sch; in 18.11 haben wir gezeigt, dass die ternäre Verknüpfung T in K sich durch zwei binäre Verknüpfungen '$+$' und '$\cdot$' ausdrücken läßt, nämlich $T(u,x,v) = ux + v$. Dort steht auch, dass $(K,+)$ eine abelsche Gruppe ist, und dass das Links-Distributivgesetz gilt. In 18.13 ist gezeigt, dass $(K^*,\cdot)$ eine Gruppe ist. Es fehlt noch das

Rechts-Distributivgesetz $(a+b)c = ac+bc$. Dies ist trivial, wenn $a = 0$. Andernfalls ist $a + b = a(1 + a^{-1}b)$, denn das Links-Distributivgesetz gilt ja. Daher

$$(a+b)c = [a(1+a^{-1}b)]c = a[(1+a^{-1}b)c] = a[c+a^{-1}bc] = ac+bc\ ,$$

wobei die Assoziativität der Multiplikation, der Spezialfall 18.12 der Rechts-Distributivität und noch einmal die Links-Distributivität benutzt wurden. Damit ist K ein Schiefkörper. Dass dann $\mathcal{E} \cong \mathcal{E}(W)$ gilt, folgt aus 17.16 (v) und daraus, dass die zugehörige projektive Ebene dann –bis auf Isomorphie– $\mathcal{E}(K^3)$ ist (vergleiche 17.6 (iv)).
Die Umkehrung steht schon in 18.4.

18.15 Beispiel: *Ternärkörper*

Sei $K = \mathbb{R}$ und

$$T(u,x,v) = \begin{cases} 2ux+v & \text{falls } u,x \text{ beide negativ} \\ ux+v & \text{sonst.} \end{cases}$$

Es ist eine Übungsaufgabe, die Ternärkörper-Axiome zu kontrollieren. Offenbar ist $T(1,a,b) = a+b$ die übliche Addition in $\mathbb{R}$. Dagegen weicht die neue Multiplikation $a * b = T(a,b,0)$ von der üblichen ab, wenn a und b negativ sind. Es ist $(-1)*(1+(-1)) = (-1)*0 = 0$, aber $(-1)*1+(-1)*(-1) = -1+2 = 1$, also gilt das Links-Distributivgesetz nicht. Insbesondere ist $(K,+,*)$ kein Schiefkörper und daher die entsprechende Ebene nicht Desargues'sch. (Es gibt noch viele andere Beispiele von Ternärkörpern, die keine Schiefkörper sind.)

Zum Schluss untersuchen wir, wann der Schiefkörper im vorigen Satz eine kommutative Multiplikation hat, also tatsächlich ein Körper ist.

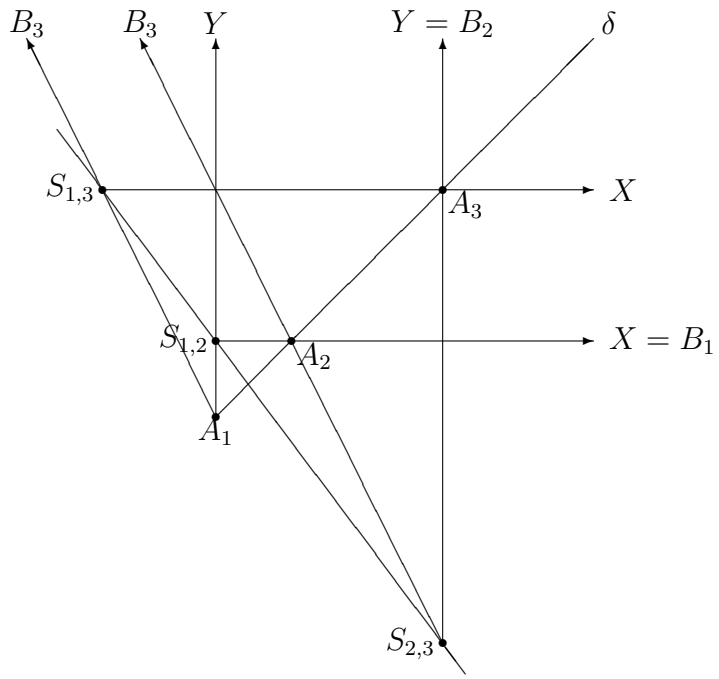
18.16 Definition: *Pappos'sche Ebenen*

Eine projektive Ebene heißt Pappos'sch, wenn sie die folgende Bedingung erfüllt: Zu je zwei verschiedenen Geraden $\gamma \neq \delta$ und je drei Punkten A_1, A_2, A_3 auf γ und B_1, B_2, B_3 auf δ, die jeweils voneinander und vom Schnittpunkt $\gamma \cap \delta$ verschieden sind, setze $S_{i,j} = A_iB_j \cap A_jB_i$ für $i \neq j$. Dann sind $S_{1,2}$, $S_{1,3}$ und $S_{2,3}$ kollinear.

18.17 Bemerkung:

(i) Wenn $\mathcal{E}$ Pappos'sch ist, dann auch die duale Ebene (Übungsaufgabe).

(ii) Unter der obigen Voraussetzung sind offenbar A_1, A_2, B_1, B_2 vier Punkte in allgemeiner Lage. Also kann man $A_1 = O$, $A_2 = E$, $B_1 = X$ und $B_2 = Y$ nehmen. Dann ist $A_3 = (v,v)$ ein weiterer Punkt auf der Diagonalen, und B_3 eine weitere Richtung, etwa u, außer 'waagerecht' in X-Richtung, 'senkrecht' in Y-Richtung, und 'diagonal', weil $B_3 \neq A_1A_2 \cap B_1B_2$. Das entsprechende Bild ist dann:



18.18 Satz:

Sei $\mathcal{E}$ eine Desargues'sche Ebene. Genau dann ist $\mathcal{E}$ Pappos'sch, wenn der Koordinaten-Schiefkörper K von $\mathcal{E}$ ein Körper ist.

Beweis:

Mit den Bezeichnungen der vorigen Bemerkung ist

$$\begin{array}{lllll} A_1B_2 & = & OY & = & \eta_0 \\ A_2B_1 & = & EX & = & \gamma_{0,1} \\ A_1B_3 & = & OB_3 & = & \gamma_{u,0} \qquad \text{für ein } u \neq 0,\, 1 \\ A_3B_1 & = & (v,v)X & = & \gamma_{0,v} \qquad \text{für ein } v \neq 0,\, 1 \\ A_2B_3 & = & EB_3 & = & \gamma_{u,1-u} \qquad \text{weil } u1 + (1-u) = 1 \\ A_3B_2 & = & (v,v)Y & = & \eta_v \qquad . \end{array}$$

Daher ist

$$\begin{array}{lllll} S_{1,2} & = & \eta_0 \cap \gamma_{0,1} & = & (0,1) \\ S_{1,3} & = & \gamma_{u,0} \cap \gamma_{0,v} & = & (u^{-1}v, v) \\ S_{2,3} & = & \gamma_{u,1-u} \cap \eta_v & = & (v, uv + 1 - u)\ . \end{array}$$

Genau dann sind diese drei Punkte kollinear, wenn ein $t \in K$ existiert mit
$(S_{1,3} - S_{1,2})t = S_{2,3} - S_{1,2}$, d.h. $(u^{-1}v, v-1)t = (v, uv-u)$.
Wenn dies der Fall ist, dann ist $u^{-1}vt = v$, also $vt = uv$, und $uv - u = (v-1)t = vt - t = uv - t$, d.h. $u = t$ und damit $uv = vu$.
Umgekehrt tut's $t = u$ natürlich, wenn u und v kommutieren.
Die Behauptung folgt, da u und v fast beliebig wählbar sind und die ausgeschlossenen Werte 0 und 1 ohnehin mit allen Elementen von K kommutieren.

18.19 Bemerkung:

Man kann rein algebraisch zeigen, dass jeder endliche Schiefkörper ein Körper ist. Endliche Desargues'sche Ebenen sind also Pappos'sch.