1 Script der LA-Vorlesung 2016/17 von Th.Zink

1.1 Vektorräume

Definition 1 *Es sei K ein Körper. Ein K-Vektorraum ist eine abelsche Gruppe $(V,+)$ zusammen mit einer Abbildung*

$$* : K \times V \to V$$

so dass folgende Eigenschaften erfüllt sind:

(1) $1_K * v = v, \quad v \in V.$

(2) Für $\lambda, \mu \in K$ und $v \in V$ gilt:

$$\lambda * (\mu * v) = (\lambda\mu) * v$$

(3) Für $\lambda, \mu \in K$ und $v, w \in V$ gilt:

$$\lambda * (v + w) = (\lambda * v) + (\lambda * w), \quad (\lambda + \mu) * v = (\lambda * v) + (\mu * v),$$

Die Elemente eines K-Vektorraums nennen wir Vektoren.

Wir bezeichnen das neutrale Element von $(V,+)$ mit $\mathbf{0}$. Wenn $v \in V$, so bezeichnet $(-v)$ das inverse Element in der Gruppe $(V,+)$. Wenn $\lambda \in K$, so bezeichnet $(-\lambda)$ das inverse Element in der Gruppe $(K,+)$. Aus der Definition folgt:

$$0_K * v = \mathbf{0}, \quad \lambda * \mathbf{0} = \mathbf{0} \quad (-1_K) * v = (-v).$$

Konvention: Wir schreiben zur Vereinfachung: $\lambda v := \lambda * v$, $1 = 1_K$, $0 = 0_K$.

Beispiele für Vektorräume sind:

Der Vektorraum der Parallelverschiebungen des Anschauungsraumes.

$M(m \times n, K)$ der Vektorraum der Matrizen mit m Zeilen und n Spalten mit Einträgen aus K. Als eine Art "Modellraum" wird der Raum der *Spaltenvektoren* benutzt:

$$K^n = M(n \times 1, K)$$

Es sei $n \in \mathbb{N}$ eine natürliche Zahl. Es sei $v_1, \ldots, v_n$ eine endliche Folge von Vektoren in einem Vektorraum V. Eine *Linearkombination* dieser Vektoren ist ein Vektor der Form:

$$\lambda_1 * v_1 + \ldots + \lambda_n * v_n,$$

wobei $\lambda_i \in K$.

Die Menge aller Linearkombinationen der Vektoren $v_1, \ldots, v_n$ ist ein Unterraum des Vektorraums V. Er heißt die *lineare Hülle* der Vektoren $v_1, \ldots, v_n$ und wird mit $\mathcal{L}(v_1, \ldots, v_n)$ bezeichnet. Wir sagen auch $\mathcal{L}(v_1, \ldots, v_n)$ ist der Raum, welcher von der Vektoren $v_1, \ldots, v_n$ *aufgespannt* wird.

Es sei $W \subset V$ ein Unterraum, so dass $v_1, \ldots, v_n \in W$. Dann gilt $\mathcal{L}(v_1, \ldots, v_n) \subset W$.

Wir sagen $v_1, \ldots, v_n$ ist ein *Erzeugendensystem* von V, wenn

$$\mathcal{L}(v_1, \ldots, v_n) = V. \tag{1}$$

Wenn es eine endliche Folge von Vektoren mit der Eigenschaft (1) gibt, so heißt der Vektorraum V *endlich erzeugt*.

Eine Folge $\lambda_1, \ldots, \lambda_n$ von Elementen aus K heißt eine *Relation* zwischen den Vektoren $v_1, \ldots, v_n$, wenn

$$\lambda_1 * v_1 + \ldots + \lambda_n * v_n = \mathbf{0}. \tag{2}$$

Wenn alle $\lambda_i = 0$, so heißt die Relation trivial.

Wenn es zwischen den Vektoren $v_1, \ldots, v_n$ eine nichttriviale Relation gibt, so heißen diese Vektoren *linear abhängig* Wenn es nur die triviale Relation gibt, nennt man die Vektoren $v_1, \ldots, v_n$ *linear unabhängig*

Wir wollen auch den Fall $n = 0$ zu lassen, wo die Folge der Vektoren leer ist. Dann definieren wir die lineare Hülle als den Nullvektorraum:

$$\mathcal{L}(\emptyset) = \{\mathbf{0}\} \subset V.$$

Wir legen fest, dass die leere Folge $\emptyset$ linear unabhängig ist.

Eine beliebige Teilmenge T eines Vektorraums V heißt linear unabhängig, wenn jede beliebige Folge $v_1, \ldots, v_n$ paarweise verschiedener Elemente aus T linear unabhängig ist.

Es sei eine nichttriviale Relation (2) gegeben. Dann gibt es einen Index i, so dass $\lambda_i \neq 0$. Daraus ergibt sich

$$\mathcal{L}(v_1, \ldots, \widehat{v_i}, \ldots, v_n) = \mathcal{L}(v_1, \ldots, v_n). \tag{3}$$

Hier bedeutet $\widehat{v_i}$, dass man v_i aus der Folge $v_1, \ldots, v_n$ streichen soll.

Wenn umgekehrt Gleichung (3) für ein i, $1 \leq i \leq n$ erfüllt ist, so sind die Vektoren $v_1, \ldots, v_n$ linear abhängig sind.

Wenn die Vektoren $v_1, \ldots, v_{n-1}$ linear unabhängig sind und wenn

$$v_n \notin \mathcal{L}(v_1, \ldots, v_{n-1}),$$

so sind auch die Vektoren $v_1, \ldots, v_{n-1}, v_n$ linear unabhängig.

Definition 2 *Es sei V ein Vektorraum. Wir nennen eine Folge von Vektoren $v_1, \ldots, v_n$ eine Basis von V, wenn die Vektoren linear unabhängig sind und wenn $\mathcal{L}(v_1, \ldots, v_n) = V$.*

Wenn $V = \{\mathbf{0}\}$ ist, so sagen wir, dass die Folge $\emptyset$ eine Basis von V ist.

Konvention: Es sei $v \in V$ und $\lambda \in K$. Dann werden wir in Zukunft schreiben:

$$\lambda v := \lambda * v.$$

Beispiel: Der Vektorraum der Spaltenvektoren K^n hat die Standardbasis $e_1, \ldots, e_n$. Hier ist e_i der Vektor

$$e_i = \begin{pmatrix} 0 \\ \ldots \\ 0 \\ 1 \\ 0 \\ \ldots \\ 0 \end{pmatrix}$$

Die 1 steht in der i-ten Zeile. In den anderen Zeilen steht eine Null. Wir nennen e_i den i-ten *Standardvektor*

Es sei $v_1, \ldots, v_n$ eine Basis von V. Dann lässt sich jeder Vektor $v \in V$ als Linearkombination schreiben:

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n, \quad \text{wo } \lambda_i \in K. \tag{4}$$

Die λ_i sind durch v eindeutig bestimmt. Sie heißen die *Koordinaten* des Vektors v bezüglich der Basis $v_1, \ldots, v_n$.

Satz 3 *Es sei V ein Vektorraum. Es sei $v_1, \ldots, v_n$ eine Folge von Vektoren von V, so dass $\mathcal{L}(v_1, \ldots, v_n) = V$. Dann kann man aus den Vektoren $v_1, \ldots, v_n$ eine Basis von V auswählen.*

Beweis: In der Tat, wenn $v_1, \ldots, v_n$ linear abhängig sind, so findet man einen Index i, so dass (3) gilt. *Q.E.D.*

Bemerkung: Man kann auch von einer Basis von V sprechen, wenn V nicht endlich erzeugt ist. Das ist eine Teilmenge $B \subset V$, die linear unabhängig ist und deren lineare Hülle V ist. Letzeres bedeutet, dass jedes Element $v \in V$ eine Linearkombination von Elementen aus B ist. Es sei $E \subset V$ eine Teilmenge. Die lineare Hülle $\mathcal{L}(E)$ ist die Menge aller Vektoren von V, die sich als Linearkombinationen von Vektoren aus E schreiben lassen. $\mathcal{L}(E)$ ist ein Untervektorraum von V. Wir zitieren hier die folgende Verallgemeinerung des letzten Satzes:

Korollar 4 *Es sei V ein Vektorraum. Es sei $E \subset V$ eine Teilmenge, so dass $\mathcal{L}(E) = V$.*

Dann gibt es eine Teilmenge $B \subset E$, die eine Basis von V ist. Insbesondere hat jeder Vektorraum eine Basis.

Man kann das leicht auf den letzten Satz zurückführen, wenn es eine abzählbare Teilmenge von V gibt, die V erzeugt.

Satz 5 *(Der Austauschsatz) Es sei $v_1, \ldots, v_d$ eine Basis eines Vektorraumes V. Es seien $x_1, \ldots, x_r$ linear unabhängige Vektoren.*

Dann ist $r \leq d$. Man kann aus der Menge $\{v_1, \ldots, v_d\}$ $d - r$ Vektoren $v_{i_{r+1}}, \ldots, v_{i_d}$ so auswählen, dass

$$x_1, \ldots, x_r, v_{i_{r+1}}, \ldots, v_{i_d}$$

eine Basis von V ist.

Zum Beweis formulieren wir:

Lemma 6 *Es sei $v_1, \ldots, v_d$ eine Basis eines Vektorraumes V. Es sei $s \in \mathbb{Z}$, so dass $0 \leq s \leq d$. Es sei $x \in V$ ein Vektor, so dass die Vektoren $x, v_1, \ldots, v_s$ linear unabhängig sind.*

Dann gibt es ein $i \in \mathbb{N}$, $d \geq i > s$ mit folgender Eigenschaft: Wenn man in der Folge $v_1, \ldots, v_d$ den Vektor v_i durch den Vektor x ersetzt, so erhält man wieder eine Basis von V:

$$v_1, \ldots, v_s, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_d \tag{5}$$

Beweis: Wir drücken x als Linearkombination von $v_1, \ldots, v_d$ aus:

$$x = \lambda_1 v_1 + \ldots + \lambda_d v_d. \tag{6}$$

Es gibt einen Index $i > s$, so dass $\lambda_i \neq 0$. In der Tat, sonst wären die Vektoren $x, v_1, \ldots, v_s$ linear abhängig.

Nach (6) ist v_i eine Linearkombination der Vektoren (5). Daraus folgt, dass die Vektoren (5) den Vektorraum V erzeugen.

Da die Darstellung (6) eindeutig ist und $\lambda_i \neq 0$ folgt, dass

$$x \notin \mathcal{L}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_d).$$

Also sind die Vektoren (5) linear unabhängig. *Q.E.D.*

Beweis von Satz 5:

Induktion nach r. Der Fall $r = 1$ ist nach Lemma 6 (Fall $s = 0$) klar.

Induktionsschritt: Man wende das Lemma auf die Basis

$$x_1, \ldots, x_{r-1}, v_{i_r}, \ldots, v_{i_d},$$

den Vektor $x = x_r$ und $s = r - 1$ an. *Q.E.D.*

Korollar 7 *(Basisergänzungssatz) Es sei V ein endlich erzeugter Vektorraum. Es sei $V = \mathcal{L}(w_1, \ldots, w_m)$. Es seien $x_1, \ldots, x_r \in V$ linear unabhängige Vektoren. Dann kann man Vektoren $x_{r+1}, \ldots, x_d \in \{w_1, \ldots, w_m\}$ finden, so dass $x_1, \ldots, x_d$ eine Basis von V ist.*

Beweis: Nach Satz 3 kann man annehmen, dass $w_1, \ldots, w_m$ eine Basis von V ist. Dann folgt die Behauptung aus dem Austauschsatz 5. *Q.E.D.*

Aus dem Austauschsatz sieht man, dass zwei Basen eines endlich erzeugten Vektorraums stets gleichviele Elemente enthalten:

Definition 8 *Es sei V ein endlich erzeugter Vektorraum. Die Anzahl der Elemente in einer Basis von V nennt man die Dimension von V.*

Korollar 9 *Es sei V ein endlich erzeugter Vektorraum der Dimension d. Dann sind d linear unabängige Vektoren von V stets eine Basis von V.*

Eine beliebige Folge von $d + 1$ Vektoren in V ist stets linear abhängig.

Es sei $W \subset V$ ein Unterraum. Wir nennen $W' \subset V$ einen *Komplementärraum* von W, wenn

$$W + W' = V \quad \text{und} \quad W \cap W' = \{0\}.$$

Satz 10 *Es sei V ein endlich erzeugter Vektorraum. Es sei $W \subset V$ ein Unterraum. Dann ist W endlich erzeugt.*

Wenn $\dim W = \dim V$, so gilt $W = V$.

Zu jedem Unterraum $W \subset V$ gibt es einen Komplementärraum.

Wenn W und W' komplementäre Unteräume in V sind, so gilt

$$\dim W + \dim W' = \dim V$$

Bemerkung: Mit Hilfe des Auswahlaxioms der Mengenlehre kann man beweisen, dass ein Komplementärraum eines Unterraumes auch dann existiert, wenn V nicht endlich erzeugt ist. Wir nehmen diese Tatsache hier zur Kenntnis.

Beispiel: Vektorräume von Funktionen. Es sei V eine K-Vektorraum und es sei I eine Menge. Die Menge aller Abbildungen (=Funktionen)

$$f : I \to V. \tag{7}$$

bezeichnen wir mit V^I oder auch mit $\mathrm{Abb}(I, V)$.

Man definiert die Summe von zwei Funktionen $f, g \in V^I$:

$$(f + g)(i) = f(i) + g(i), \quad i \in I.$$

Dadurch wird V^I eine abelsche Gruppe.

Die Multiplikation einer Funktion f mit einem $\lambda \in K$ ist folgendermaßen definiert:

$$(\lambda f)(i) = \lambda(f(i)).$$

Dadurch wird $V^I = \mathrm{Abb}(I, V)$ ein Vektorraum.

Man definiert $V^{(I)} \subset V^I$ als die Teilmenge aller Funktionen $f : I \to V$, für die eine endliche Teilmenge $S \subset I$ existiert, so dass

$$f(i) = 0, \quad i \notin S.$$

(S kann von f abhängen.) Dann ist $V^{(I)}$ ein Untervektorraum von V^I.

Übung: Man finde eine Basis von $V^{(I)}$.

Es sei $K = \mathbb{R}$, $V = \mathbb{R}$ als $\mathbb{R}$-Vektorraum und es sei $I = (a, b)$ ein Intervall (offen oder abgeschlossen) in $\mathbb{R}$. Dann ist die Menge $\mathbb{R}^I$ aller reellwertigen Funktionen auf (a, b) ein Vektorraum. In der Analysis zeigt man, dass die stetigen bzw. differenzierbaren Funktionen einen Unterraum bilden.

1.2 Lineare Abbildungen

Isomorphismus
Automorphismus

Definition 11 *Eine* lineare Abbildung $f : V \to W$ *von* K*-Vektorräumen ist eine Abbildung, die folgende Eigenschaften hat:*

(1) $f(v_1 + v_2) = f(v_1) + f(v_2), \quad v_1, v_2 \in V.$

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

$$f(\lambda v) = \lambda f(v).$$

Eine lineare Abbildung wird auch *Homomorphismus* genannt. Wenn $V = W$, so nennt man $f : V \to V$ auch einen *Endomorphismus*

Die Menge aller linearen Abbildungen von V nach W bilden einen Untervektorraum $\mathrm{Hom}_K(V, W) \subset \mathrm{Abb}(V, W)$ (vgl. (7)). Wir schreiben auch $\mathrm{End}_K(V) = \mathrm{Hom}_K(V, V)$ für den Vektorraum der Endomorphismen.

Der *Kern* von f ist ein Unterraum von V: $\mathrm{Ker}\, f = f^{-1}(\mathbf{0})$

Lemma 12 *Ein Homomorphismus von* K*-Vektorräumen* $f : V \to W$ *ist genau dann injektiv, wenn* $\mathrm{Ker}\, f = \{\mathbf{0}\}$.

Das *Bild* von f ist ein Unterraum von W: $\mathrm{Im}\, f = f(V)$

Das *Kompositum* $g \circ f$ linearer Abbildungen f und g:

$$f : V \to W \quad g : W \to U,$$

ist die lineare Abbildung $g \circ f : V \to U$, so dass $(g \circ f)(v) = g(f(v))$ für alle $v \in V$.

Eine lineare Abbildung $f : V \to W$ heißt *Isomorphismus* , wenn eine lineare Abbildung $g : W \to V$ existiert, so dass

$$g \circ f = \mathrm{id}_V, \quad f \circ g = \mathrm{id}_W\,.$$

Bemerkung: Eigenschaften und Konstruktionen die in dem Vektorraum V gelten kann man mit Hilfe der Isomorphismus f in den Vektorraum W tranportieren. Deshalb haben isomorphe Vektorräume für die Theorie der Vektorräume die gleichen Eigenschaften. Man überzeuge sich z.B., dass V genau dann endlich erzeugt ist, wenn W endlich erzeugt ist, und dass $\dim V = \dim W$.

Wir nennen einen Isomorphismus $f : V \to V$ auch einen *Automorphismus* von V.

Eine lineare Abbildung $f : V \to W$ ist genau dann injektiv, wenn $\operatorname{Ker} f = 0$. Sie ist genau dann surjektiv, wenn $\operatorname{Im} f = W$. Schließlich ist f genau dann bijektiv, wenn f ein Isomorphismus ist. Faktorraum

Ein *Diagramm* von Abbildungen von Vektorräumen

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle k} \\ T & \xrightarrow{g} & U \end{array} \tag{8}$$

heißt kommutativ, wenn $k \circ f = g \circ h$

Es sei $U \subset V$ ein Unterraum eines Vektorraumes V. Ein *Faktorraum* von V modulo U ist eine surjektive lineare Abbildung von Vektoräumen $\alpha : V \to W$, so dass $\operatorname{Ker} \alpha = U$. Ein Faktorraum existiert stets. Für Faktorraum wird synonym das Wort Quotientenraum verwendet.

Die Existenz eines Faktorraum (W, α) beweist man genau so wie für abelsche Gruppen (Satz 126).

Man kann die Existenz eines Faktorraumes auch mit Hilfe des Komplementärraumes zeigen: Man wählt einen Komplementärraum $U' \subset V$ zu dem Unterraum U. Jedes Element $v \in V$ hat eine eindeutige Darstellung

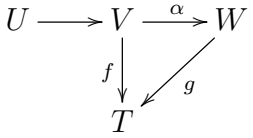
$$v = u + u', \quad u \in U,\ u' \in U'.$$

Man definiert eine lineare Abbildung $\alpha' : V \to U'$ durch die Gleichung

$$\alpha'(v) = u'.$$

Dann ist $\alpha' : V \to U'$ eine Faktorraum von V modulo U.

Satz 13 *Es sei $\alpha : V \to W$ ein Faktorraum modulo U. Es sei $f : V \to T$ eine lineare Abbildung, so dass $U \subset \operatorname{Ker} f$. Dann gibt es eine eindeutig bestimmte lineare Abbildung $g : W \to T$, so dass folgendes Diagramm kommutativ ist.*



Beweis: Es sei $w \in W$. Man wählt einen Vektor $v \in \pi^{-1}(w)$ und definiert:

$$g(w) = f(v).$$

Da $U \subset \operatorname{Ker} f$ ist diese Definition unabhängig von der Wahl von v. Also ist $g : W \to T$ ein wohldefinierte Abbildung von Mengen. Man zeigt leicht, dass g linear ist. $Q.E.D.$ Projektion

Es seien $\pi : V \to W$ und $\pi' : V \to W'$ zwei Faktorräume modulo U. Dann gibt es nach dem letzten Satz einen eindeutig bestimmten Isomorphismus $g : W \to W'$, so dass $g \circ \pi = \pi'$. Deshalb ist es nicht wichtig, welchen Faktorraum man wählt. Wir bezeichen einen beliebig gewählten Faktorraum oft mit $V \to V/U$.

Satz 14 Dimensionsformel *Es sei U ein Unterraum eines endlich erzeugten Vektorraumes V. Es sei $\alpha : V \to W$ ein Faktorraum modulo U, d.h. $\alpha(V) = W$ und $U = \operatorname{Ker} \alpha$. Dann gilt:*

$$\dim V = \dim U + \dim W.$$

Beweis: Aus Satz 10 folgt leicht, dass U und W wieder endlich erzeugt sind. Man wählt eine Basis $u_1, \ldots, u_n$ von U und eine Basis $w_1, \ldots, w_m$ von W. Weil α surjektiv ist, gibt es Elemente $\tilde{w}_1, \ldots, \tilde{w}_m \in V$, so dass $\alpha(\tilde{w}_i) = w_i$ für $i = 1, \ldots, m$. Man verifiziert leicht, dass die folgenden Elemente eine Basis von V sind:

$$u_1, \ldots, u_n, \tilde{w}_1, \ldots \tilde{w}_m.$$

Aus der Formel (13) kann man bequem alle übrigen Dimensionsformeln herleiten. $Q.E.D.$

Man kann die Dimensionsformel für eine beliebige lineare Abbildung $f : V \to W$ endlich erzeugter Vektorräume formulieren.

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V. \tag{9}$$

In der Tat, $f : V \to \operatorname{Im} f$ ist eine Surjektion mit dem Kern $\operatorname{Ker} f$. Deshalb können wir den letzten Satz anwenden.

Es seien V_1 und V_2 zwei komplementäre Unterräume eines Vektorraumes V. Dann gibt es genau eine lineare Abbildung

$$\pi : V \to V,$$

so dass $\pi(v_1) = v_1$ für alle $v_1 \in V_1$ und $\pi(v_2) = 0$ für alle $v_2 \in V_2$. Wir nennen π die *Projektion* auf V_1 längs V_2. Es gilt $\pi^2 = \pi$. Geometrischer kann man die Projektion auch so definieren:

$$\pi(v) = (v + V_2) \cap V_1$$

Anschaulich ist das der Durchschnitt der Parallele durch v zu V_2 mit dem Raum V_1. In der Geometrie ist das eine Parallelprojektion.

direkte Summe

innere direkte Summe

Es seien $V_1, \ldots, V_n$ Vektorräume. Dann definiert man auf der Produktmenge

$$V_1 \times \ldots \times V_n.$$

die Struktur eines Vektorraums. Die Addition zweier Elemente $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_n)$, wobei $v_i, w_i \in V$, ist wie folgt definiert

$$(v_1, \ldots, v_n) + (w_1, \ldots, w_n) := (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n).$$

Die Multiplikation mit einem Skalar $\lambda \in K$ ist nach Definition:

$$\lambda(v_1, v_2, \ldots, v_n) := (\lambda v_1, \lambda v_2, \ldots \lambda v_n)$$

Diesen Vektorraum nennt man *direkte Summe* der Vektorräume $V_1, \ldots, V_n$:

$$V_1 \oplus \ldots \oplus V_n.$$

Wenn $V_1 = \ldots = V_n = K$, so erhalten wir $K^n = K \oplus \ldots \oplus K$.

Es seien $V_1, \ldots, V_n$ Unterräume eines Vektorraums V. Wir betrachten die lineare Abbildung:

$$\begin{array}{ccc} V_1 \oplus \ldots \oplus V_n & \longrightarrow & V \\ (v_1, v_2, \ldots, v_n) & \mapsto & v_1 + v_2 + \ldots + v_n. \end{array} \tag{10}$$

Wenn (10) ein Isomorphismus ist, so sagen wir, dass V die *innere direkte Summe* seiner Unterräume $V_1, \ldots, V_n$ ist. Wir erlauben uns in diesem Fall die etwas unpräzise Schreibweise

$$V = V_1 \oplus \ldots \oplus V_n,$$

obwohl es sich hier nur um einen Isomorphismus handelt.

Die Aussage, dass zwei Unterräume U und W von einem Vektorraum V komplementär sind, schreibt sich dann auch so:

$$V = U \oplus W.$$

Satz 15 *Es seien $V_1, \ldots, V_n$ endlich erzeugte K-Vektorräume. Dann gilt*

$$\dim(V_1 \oplus \ldots \oplus V_n) = \dim V_1 + \dim V_2 + \ldots + \dim V_n.$$

Definition 16 *Es seien $f : V \to W$ und $g : U \to V$ zwei lineare Abbildungen von Vektorräumen. Wir nennen die Sequenz*

$$U \xrightarrow{g} V \xrightarrow{f} W \tag{11}$$

exakt in V, wenn $\operatorname{Im} g = \operatorname{Ker} f$. Insbesondere gilt dann $g \circ f = 0$.

Wenn $U = 0$, so ist (11) genau dann exakt, wenn f injektiv ist. Wenn $W = 0$, so ist (11) genau dann exakt, wenn g surjektiv ist.

Eine kurze exakte Sequenz ist eine Folge von Homomorphismen

$$0 \to U \xrightarrow{g} V \xrightarrow{f} W \to 0, \tag{12}$$

die an den Stellen U, V, und W exakt ist. Insbesondere ist dann g injektiv und f surjektiv.

Es sei V ein endlich erzeugter Vektorraum. Dann sind in der kurzen exakten Sequenz (12) die Vektorräume U und W auch endlich erzeugt und es gilt die Formel:

$$\dim U + \dim W = \dim V. \tag{13}$$

Das folgt aus Satz 14.

Beispiel: Es seien V_1 und V_2 Vektorräume. Dann hat man eine kurze exakte Sequenz

$$\begin{array}{ccccccc} 0 \to & V_1 & \to & V_1 \oplus V_2 & \to & V_2 & \to 0. \\ & v_1 & \mapsto & (v_1, 0)\ ;\ (v_1, v_2) & \mapsto & v_2 & \end{array}$$

Satz 17 Dimensionsformel für Unterräume *Es seien W_1 und W_2 Unterräume eines endlich erzeugten Vektorraumes V. Dann gilt:*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$

Beweis: Man hat eine kurze exakte Sequenz:

$$\begin{array}{ccccccc} 0 \to & W_1 \cap W_2 & \to & W_1 \oplus W_2 & \to & W_1 + W_2 & \to 0. \\ & w & \mapsto & (w, -w)\ ;\ (w_1, w_2) & \mapsto & w_1 + w_2 & \end{array}$$

Q.E.D.

Satz 18 *Es seien V und W Vektorräume über K. Es sei $B \subset V$ eine Basis von V. Die Abbildung* duale Basis

$$\mathrm{Hom}_K(V, W) \to \mathrm{Abb}(B, W),$$

die man durch Einschränkung einer linearen Abbildung $V \to W$ auf die Menge $B \subset V$ erhält, ist ein Isomorphismus von Vektorräumen. (Hier ist $\mathrm{Abb}(V, W)$ *der unter (7) definierte Vektorraum.)*

Beweis: Es sei $f : B \to V$ eine Abbildung (von Mengen). Einen Vektor $v \in V$ kann man auf eindeutige Weise als Linearkombination schreiben

$$v = \lambda_1 b_1 + \ldots + \lambda_t b_t,$$

wobei $b_1, \ldots b_t \in B$ verschiedene Elemente sind und $\lambda_i \in K$.

Man definiert eine Abbildung $\phi : V \to W$ durch

$$\phi(v) = \lambda_1 f(b_1) + \ldots + \lambda_t f(b_t).$$

Man sieht, dass ϕ eine lineare Abbildung ist.

Wenn f die Einschränkung einer linearen Abbildung $\psi : V \to W$ ist, so muss notwendig $\psi = \phi$ gelten. Das zeigt die Behauptung. *Q.E.D.*

Für einen endlich erzeugten Vektorraum formulieren wir diesen Satz etwas konkreter: Es sei $v_1, \ldots, v_n$ eine Basis von V. Dann gibt es zu n beliebigen Vektoren $w_1, \ldots, w_n \in W$ genau eine lineare Abbildung $f : V \to W$, so dass $f(v_i) = w_i$.

Spezialfälle: 1) Mit den Bezeichnungen des Satzes sei V ein endlich erzeugter Vektorraum und es sei $W = K$. Den Vektorraum $\mathrm{Hom}_K(V, K)$ nennt man den *dualen Vektorraum* zu V. Man schreibt $V^* = \mathrm{Hom}_K(V, K)$. Dann gibt es Elemente $\hat{v}_1, \ldots, \hat{v}_n \in V^*$, so dass $\hat{v}_i(v_j) = \delta_{ij}$ (Kroneckersymbol) für alle $i, j = 1, \ldots, n$. Die linearen Abbildungen $\hat{v}_1, \ldots, \hat{v}_n$ sind eine Basis von V^*. Diese Basis nennt man die *duale Basis* zu $v_1, \ldots v_n$. Also gilt

$$\dim V = \dim V^* \tag{14}$$

Es sei $v \in V$. Seine Koordinaten (4) bezüglich der Basis $v_1, \ldots, v_n$ sind die Elemente $\hat{v}_i(v) \in K$, d.h.

$$v = \sum_{i=1}^{n} \hat{v}_i(v) v_i.$$

Die duale Basis besteht also aus den Koordinatenfunktionen. Koordinatensyster

2) Wir bleiben in der Situation von Satz 18. Es sei $w \in W$. Wir betrachten die lineare Abbildung $\tilde{w} : K \to W,\ \lambda \mapsto \lambda w$. Das Bild ist also der eindimensionale Untervektorraum von W, der von w erzeugt wird. Man kann f mit Hilfe der dualen Basis schreiben

$$f = \sum_{i=1}^{n} \tilde{w}_i \circ \hat{v}_i, \tag{15}$$

wobei $w_i = f(v_i)$. Hier steht auf der rechten Seite, das Kompositum von Abbildungen.

3) Wir betrachten den Vektorraum $V = K^n$ mit der Standardbasis $e_1, \ldots, e_n$. Es sei $w_1, \ldots, w_n$ eine Folge von Vektoren aus W. Dann gibt es genau eine lineare Abbildung

$$b : K^n \to W \tag{16}$$

so dass $b(e_i) = w_i$. Die Abbildung b ist genau dann ein Isomorphismus von Vektorräumen, wenn $w_1, \ldots, w_n$ eine Basis von W ist.

Es sei $w_1, \ldots, w_n$ eine Basis von W. Es sei b der entsprechende Isomorphismus (16). Es sei $w \in W$ ein beliebiger Vektor. Den Spaltenvektor $\underline{x} = b^{-1}(w) \in K^n$ sind die Koordinaten des Vektors w bezüglich der Basis $w_1, \ldots, w_n$. Es gilt:

$$w = x_1 w_1 + \ldots + x_n w_n,$$

wobei die x_i die Einträge von $\underline{x}$ sind. Wir nennen einen Isomorphismus $b : K^n \to W$ ein *Koordinatensystem* von W. Wegen (16) ist die Wahl eines Koordinatensystems dasgleiche wie die Wahl einer Basis.

1.3 Matrizen

Es sei K ein Körper. Wir bezeichen mit

$$M(m \times n, K)$$

die Menge aller Matrizen mit m Zeilen und n Spalten und Koeffizienten aus K. Es sei $A \in M(m \times n, K)$. Es sei $a_{ij} \in K$, der Eintrag, welcher in der i-ten Zeile und j-ten Spalte von A steht. Dann schreiben wir $A = (a_{ij})$.

Stufenform

Man kann Matrizen addieren und multiplizieren:

$$\begin{array}{ccc} M(m\times n,K)\times M(m\times n,K) & \to & M(m\times n,K) \\ (A_1,A_2) & \mapsto & A_1+A_2. \end{array}$$

$$\begin{array}{ccc} M(\ell\times m,K)\times M(m\times n,K) & \to & M(\ell\times n,K) \\ (B,A) & \mapsto & B\cdot A. \end{array}$$

Wir bezeichnen mit $E_n \in M(n\times n,K)$ die Matrix, deren n Spalten die Standardvektoren $e_1,\dots,e_n \in K^n$ sind. Für jede Matrix $A\in M(m\times n,K)$:

$$E_m\cdot A = A, \quad A\cdot E_n = A.$$

Es sei $A\in M(m\times n,K)$ eine Matrix mit m Zeilen und n Spalten. Die lineare Hülle der n Spalten von von A in ein Unterraum von K^m. Die Dimension dieses Unterraumes nennt man den Rang von A. Genauso gut kann man sagen, dass der Rang von A die maximale Anzahl von linear unabhängigen Spalten der Matrix A ist.

Das homogene Gleichungssystem zu der Matrix A ist

$$A\underline{x}=0, \quad \underline{x}=\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \in K^n. \tag{17}$$

Wir bezeichnen mit $L_A\in K^n$ die Menge aller Lösungen von (17). Das ist ein Untervektorraum. Wir stellen uns die Aufgabe eine Basis von L_A zu finden.

Man bemerkt, dass eine Relation $x_1,\dots,x_n$ zwischen den Spalten der Matrix A das Gleiche ist wie eine Lösung von (17).

Wir betrachten zuerst spezielle Gleichungssysteme. Eine Matrix $C\in M(m\times n,K)$ hat *Stufenform* , wenn sie die folgende Gestalt besitzt:

$$C=\begin{pmatrix}
0 & \dots & 0 & c_{i_1} & * & & & & & \dots & & & & & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & c_{i_2} & * & & & \dots & & & * \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & c_{i_3} & * & & \dots & & * \\
 & & & & & & & & \dots & & & & & & & \\
0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & c_{i_r} & * & \dots & * \\
0 & & & & & & & & \dots & & & & & & & & 0 \\
 & & & & & & & & \dots & & & & & & & & \\
0 & & & & & & & & \dots & & & & & & & & 0
\end{pmatrix}$$

An den mit $*$ belegten Stellen können beliebige Einträge aus K stehen. Die Einträge c_{i_k} in der k-ten Zeile und i_k-ten Spalte sind für $k = 1, \dots r$ sind von 0 verschieden. Die Positionen (k, i_k) nennt man Pivots.

Scherung

Zeilenoperation

Die Zahl r ist der Rang der Matrix C.

Die Teilmenge $\{i_1, \dots, i_r\} \subset \{1, 2, \dots n\}$ nennt man die abhängigen Indexe. Die übrigen Zahlen in $\{1, 2, \dots, n\}$ nennt man die unabhängigen Indexe.

Zu beliebig vorgegebenen $x_\ell \in K$, wo $\ell = 1, \dots, n$ aber $\ell \neq i_1, \dots, \ell \neq i_r$, existieren eindeutig bestimmte $x_{i_1}, \dots, x_{i_r} \in K$, so dass:

$$C \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \tag{18}$$

Wenn man also für die unabhängigen Indexe ℓ beliebige Werte $x_\ell \in K$ vorgibt, so kann man sie eindeutig zu einer Lösung $(x_1, \dots, x_n)$ des Gleichungssystems (18) ergänzen.

Die letzte Aussage kann man wie folgt formulieren: Es sei $L_C \subset K^n$ die Menge aller Lösungen des Gleichungssystems (18). Es sei $j_1, \dots, j_{n-r}$ die Menge der unabhängigen Indexe. Dann ist die lineare Abbildung

$$\begin{array}{ccc} L_C & \to & K^{n-r} \\ (x_1, \dots, x_n) & \mapsto & (x_{j_1}, \dots, x_{j_{n-r}}) \end{array}$$

ein Isomorphismus von Vektorräumen. Anders ausgedrückt ist L_C ein Komplementärraum von $\mathcal{L}(e_{i_1}, \dots, e_{i_r})$, wobei $\{i_1, \dots, i_r\}$ die abhängigen Indexe sind.

Insbesondere finden wir:

$$\dim L_C = n - \operatorname{Rang} C. \tag{19}$$

Auf eine Matrix A eine *Scherung* anwenden heißt: man addiert zu einer gegebenen Zeile das Vielfache einer anderen Zeile.

Satz 19 *Jede Matrix läßt sich durch Scherungen auf Stufenform bringen.*

Auf eine Matrix A eine *Zeilenoperation* anwenden heisst, eine der folgenden Operationen ausführen:

1. eine Scherung,

2. die Multiplikation einer Zeile mit einem von Null verschiedenen Element aus K,

3. die Vertauschung zweier Zeilen.

Es sei A' eine Matrix, die aus A durch eine Zeilenoperation hervorgegangen ist. Dann gelten die folgenden trivialen aber wichtigen Beziehungen zwischen den Matrizen A und A':

Die homogenen Gleichungssysteme

$$A\underline{x} = 0 \quad \text{und} \quad A'\underline{x} = 0$$

haben die gleichen Lösungen.

Wenn eine Relation zwischen den Spalten von A besteht, so besteht die gleiche Relation zwischen den Spalten von A' und umgekehrt. Daraus folgt, dass A und A' den gleichen Rang haben.

Da man A durch Zeilenoperationen auf Stufenform bringen kann, ergibt sie aus (19):

$$\dim L_A = n - \operatorname{Rang} A. \tag{20}$$

Der Unterraum von K^n, der von den Zeilen von A erzeugt wird, stimmt mit dem Unterraum überein, der von den Zeilen von A' erzeugt wird. Die Dimension von diesem Unterraum nennen wir den *Zeilenrang* von A. Also haben A und A' auch den gleichen Zeilenrang.

Mit Zeilenoperationen kann man sogar Stufennormalform erreichen:

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

In den Spalten der Pivots gibt es einmal den Eintrag 1. Die übrigen Einträge sind 0.

Die Stufennormalform ist durch die Matrix A eindeutig bestimmt, d.h. unabhängig davon durch welche Zeilenoperationen, man diese Form erreicht.

Lösung eines Gleichungssystems der Form:

$$A\underline{x} = \underline{b}.$$

Man führt das auf den Fall $\underline{b} = 0$ zurück, indem man die Relationen zwischen den Spalten der Matrix (A, b) ausrechnet.

Satz 20 *Der Zeilenrang einer Matrix ist gleich ihrem Rang.*

Beweis: Wenn man auf eine Matrix A eine Zeilenoperation anwendet, so ändern sich Rang und Zeilenrang nicht. Also kann man annehmen, dass A eine Stufenmatrix mit r Pivots ist. Man sieht direkt, dass sowohl der Rang als auch der Zeilenrang einer solchen Matrix r ist. $Q.E.D.$

Man kann die Zeilenoperationen als Matrixmultiplikation interpretieren. Es sei $\lambda \in K$. Es seien $i, j, m \in \mathbb{N}$, so dass $i \leq m$ und $j \leq m$. Es gibt eine Matrix $E_{ij}(\lambda) \in M(m \times m, K)$ mit der folgenden Eigenschaft.

Es sei $n \in \mathbb{N}$ und $A \in M(m \times n, K)$ eine Matrix. Es sei A' die Matrix, die aus A entsteht, indem man zur i-ten Zeile von A das λ-fache der j-ten Zeile addiert. Dann gilt

$$A' = E_{ij}(\lambda)A.$$

Man findet $E_{ij}(\lambda)$, in dem für A die Einheitsmatrix E_m nimmt. Wenn man auf E_m die Scherung anwendet, so bekommt man $E_{ij}(\lambda)$.

1.4 Die Matrix einer linearen Abbildung

Eine Matrix $A \in M(m \times n, K)$ definiert eine lineare Abbildung $\phi_A : K^n \to K^m$:

$$\phi_A(\underline{x}) = A\underline{x}, \quad \underline{x} \in K^n.$$

Hier steht rechts die Matrixmultiplikation. Die Spalten von A sind die Bilder der Standardvektoren $e_i \in K^n$ bei ϕ_A. Nach Satz 18 ist die Zuordnung $A \mapsto \phi_A$ ein Isomorphismus von Vektorräumen:

$$M(m \times n, K) \to \mathrm{Hom}(K^n, K^m).$$

(In Satz 18 haben wir den Pfeil in die umgekehrte Richtung betrachtet, der einer linearen Abbildung $\phi : K^n \to K^m$ die Matrix mit den Spalten $\phi(e_i)$, $i = 1, \ldots, n$ zuordnet.)

Wir schreiben oft $A = \phi_A$. Dem Kompositum von linearen Abbildungen entspricht bei dieser Identifikation die Multiplikation von Matrizen, und die Addition von linearen Abbildungen entspricht der Addition von Matrizen. invertierbar

$\operatorname{Im}\phi_A$ wird von den Spalten von A erzeugt. Daher gilt $\dim \operatorname{Im}\phi_A = \operatorname{Rang} A$ und dann nach der Dimensionsformel

$$\dim \operatorname{Ker}\phi_A = n - \operatorname{Rang} A.$$

Offensichtlich ist $\operatorname{Ker}\phi_A$ gleich der Lösunsgsmenge L_A des linearen Gleichungssystems $A\underline{x} = 0$. Also ist die Dimensionsformel identisch mit (19).

Es sei $n = m$. Man nennt A *invertierbar* , wenn ϕ ein Isomorphimus ist. In Matrizen formuliert bedeutet das, dass eine Matrix $B \in M(n \times n, K)$ existiert, so dass

$$AB = E_n, \quad BA = E_n.$$

Die Matrix A ist genau dann invertierbar, wenn $\operatorname{Rang} A = n$. In der Tat, wenn die letzte Gleichung gilt ist ϕ surjektiv und dann nach der Dimensionsformel auch injektiv. Wir bezeichen die Menge der invertierbaren Matrizen mit $GL_n(K) \subset M(n \times n, K)$.

Es sei V ein Vektorraum mit einer Basis $v_1, \ldots, v_n$ und W ein Vektorraum mit einer Basis $w_1, \ldots, w_m$. Wir betrachten die entsprechenden Koordinatensysteme $b : K^n \to V$ und $c : K^m \to W$ (vgl. (16)).

Wenn $f : V \to W$ eine lineare Abbildung ist, so gibt es genau eine Matrix $A : K^n \to K^m$, so dass folgendes Diagramm kommutativ ist.

$$\begin{array}{ccc} K^n & \xrightarrow{A} & K^m \\ {\scriptstyle b}\downarrow & & \downarrow{\scriptstyle c} \\ V & \xrightarrow{f} & W \end{array} \tag{21}$$

Explizit erhält man die Matrix $A = (a_{ij})$, indem man die Bilder $f(v_j)$ als Linearkombination der Vektoren w_i schreibt:

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad j = 1, \ldots, n \tag{22}$$

Es sei $\hat{w}_1, \ldots, \hat{w}_m \in \operatorname{Hom}_K(W, K)$ die duale Basis zu $w_1, \ldots, w_m$. Dann kann man die letzte Formel auch schreiben:

$$a_{ij} = \hat{w}_i(f(v_j)), \quad i = 1, \ldots, m \quad \text{und } j = 1, \ldots, n. \tag{23}$$

Die j-te Spalte der Matrix (a_{ij}) besteht also aus den Koordinaten des Vektors $f(v_j)$ bezüglich der Basis $w_1, \ldots, w_m$.

Wenn $\underline{x} \in K^n$ die Koordinaten eines Vektors $v \in V$ bezüglich b sind, und $\underline{y}$ die Koordinaten von $f(v)$ bezüglich c, so gilt:

$$\underline{y} = A\underline{x} \tag{24}$$

Wir nennen A die Matrix von f bezüglich der Koordinatensysteme b und c bzw. bezüglich der Basen $v_1, \ldots, v_n$ und $w_1, \ldots, w_m$.

1.5 Basiswechsel

Es seien $b : K^n \to V$ und $b' : K^n \to V$ zwei Koordinatensysteme des Vektorraumes V. Dann ist die Abbildung

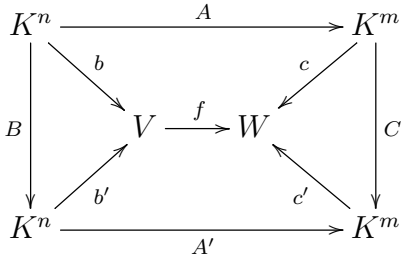
$$B = {b'}^{-1} b : K^n \to K^n$$

eine invertierbare Matrix. Das ist die *Übergangsmatrix* von dem Koordinatensystem b zu dem Koordinatensystem b'. Wenn $\underline{x} \in K^n$ die Koordinaten eines Vektors $v \in V$ in dem Koordinatensystem b sind, so sind

$$\underline{x}' = B\underline{x} \tag{25}$$

die Koordinaten desselben Vektors in dem Koordinatensystem b'.

Es sei $f : V \to W$ ein Homomorphismus. Es sei A die Matrix von f bezüglich der Koordinatensysteme b und c wie unter (21). Es seien $b' : K^n \to V$ und $c' : K^m \to W$ weitere Basen, und es sei A' die Matrix von f in diesen Basen. Es sei C die Übergangsmatrix vom Koordinatensystem c zum Koordinatensystem c' (wie unter (25)). Dann erhalten wir ein kommutatives Diagramm:



Daraus folgt, dass $CA = A'B$. Die Matrix von f in den neuen Koordinatensystemen b' und c' berechnet man also aus A wie folgt:

$$A' = CAB^{-1}.$$

Diese Formel bekommt man ganz bequem auch so: Es sei $v \in V$ und es sei $w = f(v)$. Es seien $\underline{x}$ bzw. $\underline{x'}$ die Koordinaten von v bezüglich b bzw. b' und es seien $\underline{y}$ bzw. $\underline{y'}$ die Koordinaten von w bezüglich c bzw. c'. Dann gilt: Spur

$$\underline{y'} = C\underline{y} = CA\underline{x} = CAB^{-1}\underline{x'}.$$

Im Fall $V = W$ will man die Basen häufig so wählen, dass $b = c$ bzw. $b' = c'$. Dann erhält man die Umrechnungsformel:

$$A' = BAB^{-1} \tag{26}$$

Man nennt zwei quadratische Marizen $A, A' \in M(n \times n, K)$ *ähnlich*, wenn es eine invertierbare Matrix $B \in M(n \times n, K)$ gibt, so dass (26) gilt.

Man kann sich fragen, ob ein Koordinatensystem b von V existiert, in dem die Matrix des Endomorphismus f möglichtst einfache Gestalt hat (Problem der Normalform siehe Abschnitt Jordansche Normalform). Äquivalent ist das Problem zu einer quadratischen Matrix A eine ähnliche von einfacher Gestalt zu finden.

Es sei $A = (a_{ij})$ eine $n \times n$-Matrix. Dann definiert man die *Spur* von A:

$$\text{Spur } A = \sum_{i=1}^{n} a_{ii}$$

Lemma 21 *Es seien m, n natürliche Zahlen. Es seien $B \in M(m \times n, K)$ und $C \in M(n \times m, K)$. Dann gilt:*

$$\text{Spur } BC = \text{Spur } CB$$

Der Beweis ist eine Übung in Matrixmultiplikation. Nach diesem Lemma haben ähnliche Matrizen die gleiche Spur. Deshalb ist die folgende Definition sinnvoll:

Definition 22 *Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraumes. Es sei A die Matrix von f bezüglich einer beliebig gewählten Basis von V. Dann definieren wir:*

$$\text{Spur } f = \text{Spur } A$$

Die Spur definiert eine lineare Abbildung von Vektorräumen:

$$\text{Spur} : \text{End}(V) \to K, \tag{27}$$

so dass für $f, g \in \operatorname{End} V$ gilt: $\operatorname{Spur}(f \circ g) = \operatorname{Spur}(g \circ f)$.

Es sei $\phi : V \to K$ eine lineare Abbildung und es sei $v \in V$. Wir bezeichnen mit $\tilde{v} : K \to V$ die lineare Abbildung, so dass $\tilde{v}(1) = v$. Dann gilt:

$$\operatorname{Spur}(\tilde{v} \circ \phi) = \phi(v) \tag{28}$$

In der Tat, man kann annehmen, dass $V = K^n$. Dann ist ϕ eine Zeilenmatrix. Die Abbildung $\tilde{v}$ und der Vektor v sind die gleiche Spaltenmatrix. Die Formel (28) ist dann ein einfacher Spezialfall von Lemma 21.

Es sei $v_1, \ldots, v_n$ eine Basis von V. Es sei $\hat{v}_1, \ldots, \hat{v}_n$ die duale Basis des dualen Vektorraumes. Es sei $f : V \to V$ ein Endomorphismus. Dann ergibt sich aus (28) und aus (15), dass

$$\operatorname{Spur} f = \sum_{i=1}^{i=n} \hat{v}_i(f(v_i)).$$

Diese Formel ergibt sich auch aus (22). Wir sehen, dass die Spur (27) als lineare Abbildung durch die Formel (28) eindeutig bestimmt ist.

1.6 Matrix eines Endomorphismus

Definition 23 *Es sei (V, f) eine endlich erzeugter Vektorraum V, der mit einem Endomorphismus $f : V \to V$ versehen ist. Es sei (W, h) von der gleichen Art. Wir sagen, dass (V, f) zu (W, h) ähnlich ist, wenn ein Isomorphismus $\pi : W \to V$ existiert, so dass das folgendes Diagramm*

$$\begin{array}{ccc} W & \xrightarrow{\pi} & V \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow{\pi} & V \end{array} \tag{29}$$

kommutativ ist.

Die Änlichkeit ist eine Äquivalenzrelation.

In dem Speziallfall, wo $W = K^d$ ist $\pi : K^d \to V$ ein Koordinatensystem von V. Dann ist $h = A$ eine quadratische Matrix. Wenn $f : V \to V$ zu der Matrix A ähnlich ist, bedeutet das also, dass A die Matrix von f in einem geeigneten Koordinatensystem ist.

Wir wollen Beziehungen zwischen den Eigenschaften von f und den zu f ähnlichen Matrizen untersuchen. Die wichtigste Beziehung ist die "Jordansche Normalform", die wir später herleiten. Dieser Abschnitt ist eine Einführung in das Thema.

Endomorphismus, diagonalisierbar
Endomorphismus, trigonalisierbar
Unterraum, invarianter

Definition 24 *Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeuten Vektorraums. Wir nennen f diagonalisierbar, wenn f ähnlich zu einer Diagonalmatrix ist und trigonalisierbar, wenn f ähnlich zu einer oberen Dreiecksmatrix ist.*

Zu Untersuchung dieser Begriffe brauchen wir die folgende Definition.

Definition 25 *Es sei $f : V \to V$ ein Endomorphismus eines Vektorraumes. Wir nennen einen Unterraum $U \subset V$ invariant bezüglich f, wenn $f(U) \subset U$.*

Es sei $P : K^n \to K^n$ eine quadratische Matrix. Es sei $U = \mathcal{L}(e_1, \dots e_r)$. Dann ist U genau dann ein invarianter Unterraum von K^n, wenn P eine Blockmatrix vom folgenden Typ ist:

$$P = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \tag{30}$$

wobei $A \in M(r \times r, K)$ und $B \in M(n-r \times n-r, K)$. In der Tat, die ersten r Spalten der Matrix müssen in U liegen. Wir sagen in diesem Fall, dass P eine obere Blockmatrix ist.

Es sei jetzt P so, dass $\mathcal{L}(e_{n-r+1}, \dots, e_n) \subset K^n$ invariant bei P ist. Dann ist P eine untere Blockmatrix:

$$P = \begin{pmatrix} B & 0 \\ C & A \end{pmatrix}, \tag{31}$$

wobei $A \in M(r \times r, K)$ und $B \in M(n-r \times n-r, K)$.

Wenn die komplementären Unterräume $\mathcal{L}(e_1, \dots e_r)$ und $\mathcal{L}(e_{n-r+1}, \dots, e_n)$ von K^n beide invariant bezüglich P sind, so ist P eine Diagonalblockmatrix

$$P = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

Es sei $n = \dim V$. Es sei $U \subset V$ ein invarianter Unterraum der Dimension r von $f : V \to V$. Dann wählen wir eine Basis $v_1, \dots, v_r$ von U und

ergänzen sie zu einer Basis $v_1, \ldots, v_n$ von V. Wir betrachten dass zugehörige Koordinatensystem $b : K^n \to V$, wo $b(e_i) = v_i$. Dann ergibt sich die Matrix von f aus dem Diagramm

Fahne
Fahne, zerf"allenden Basis

$$\begin{array}{ccc} K^n & \xrightarrow{P} & K^n \\ {\scriptstyle b}\downarrow & & \downarrow{\scriptstyle b} \\ V & \xrightarrow{f} & V. \end{array}$$

Da $b^{-1}(U) = \mathcal{L}(e_1, \ldots, e_r)$, lässt P den Unterraum $\mathcal{L}(e_1, \ldots, e_r) \subset K^n$ invariant. Daher ist P eine obere Blockmatrix von der Form (30). Man kann auch eine Basis $w_1, \ldots, w_n$ von V nehmen, so dass $w_{n-r+1}, \ldots, w_n$ eine Basis von U ist. Es sei c das Koordinatensystem von V, so dass $c(e_i) = w_i$. Dann ist die Matrix von f in dem Koordinatensystem c eine untere Blockmatrix von der Form (31).

Bemerkung: Wir stellen fest, dass es zu einem Endomorphismus $f : V \to V$ genau dann einen invarianten Unterraum $U \subset V$ der Dimension r gibt, wenn die Matrix von f in einer geeigneten Basis eine Blockmatrix von der Form (30) (bzw. von der Form (31)) ist.

Wenn V die Summe zweier f-invarianter komplementärer Räume ist

$$V = U \oplus W,$$

so wählen wir eine Basis $v_1, \ldots v_r$ von U und eine Basis $v_{r+1}, \ldots, v_n$ von W. Dann ist die Matrix von f in der Basis $v_1, \ldots, v_n$ von V eine Diagonalblockmatrix.

Wir können das verallgemeinern.

Definition 26 *Es sei V ein Vektorraum. Eine Fahne der Länge d ist eine aufsteigende Kette von Unterräumen V_i*

$$0 = V_0 \subset V_1 \subset V_2 \subset \ldots \subset V_{d-1} \subset V_d = V. \tag{32}$$

Wenn alle Vektorräume V_i verschieden sind, nennen wir die Fahne echt.

Es sei V von der Dimension n. Dann gibt es eine Basis $v_1, \ldots, v_n$ und ganze Zahlen

$$0 \leq i_1 \leq i_2 \ldots \leq i_{d-1} \leq i_d = n$$

so dass $v_1, \ldots, v_{i_t}$ eine Basis von V_t , für $i = 1, \ldots, e$ ist. Wir sprechen von einer *zerfällenden Basis* der Fahne.

Wir betrachten in K^n die Fahne

Fahne, voll

$$0 = \mathcal{L}(e_1, \ldots, e_{i_1}) \subset \ldots \subset \mathcal{L}(e_1, \ldots, e_{i_{d-1}}) \subset \mathcal{L}(e_1, \ldots, e_{i_d}) = K^n \qquad (33)$$

Es sei $b : K^n \to V$ ein Koordinatensystem, so dass $b(e_l) = v_l$ eine zerfällende Basis der Fahne (32) ist. Dann bildet b die Fahne (33) auf die Fahne (32) ab.

Wir wollen voraussetzen, dass die Fahne (33) echt ist. Es sei $P : K^n \to K^n$ eine Matrix, die die Fahne (33) invariant lässt. Dann ist P eine Blockmatrix der Form

$$P = \begin{pmatrix} A_{11} & * & * & \ldots & * \\ 0 & A_{22} & * & \ldots & * \\ 0 & 0 & A_{33} & \ldots & * \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & A_{dd} \end{pmatrix}, \qquad (34)$$

Hier ist $A_{tt} \in M((i_t - i_{t-1}) \times (i_t - i_{t-1}), K)$, für $t = 1, \ldots, d$ und oberhalb dieser Blöcke stehen beliebige Elemente $*$.

Es sei $f : V \to V$ ein Endomorphismus, der eine echte Fahne (32) invariant lässt, d.h. $f(V_t) \subset V_t$. Dann hat die Matrix von f in einer zerfällenden Basis die Form (34). Wenn man die Reihenfolge der Elemente in der zerfällenden Basis umkehrt: $v_n, v_{n-1}, \ldots, v_1$, so würde man anstelle von (34) eine untere Blockmatrix bekommen.

Man nennt eine Fahne (32) *voll* , wenn $\dim V_t = \dim V_{t-1} + 1$ für $t = 1, \ldots, d$. Dann ist $f = n = \dim V$. In diesem Fall ist die Matrix von f in einer zerfällenden Basis der Fahne eine obere Dreiecksmatrix. Das sieht man, weil die oberen Dreiecksmatrizen gerade diejenigen sind, welche die volle Fahne

$$0 = \mathcal{L}(e_1) \subset \mathcal{L}(e_1, e_2) \subset \ldots \subset \mathcal{L}(e_1, \ldots, e_{n-1}) \subset \mathcal{L}(e_1, \ldots, e_n) = K^n \qquad (35)$$

invariant lassen. Wir nennen das die Standardfahne im K^n. Wir haben bewiesen:

Satz 27 *Ein Endomorphismus $f : V \to V$ eines endlich erzeugten Vektorraums ist genau dann trigonalisierbar, wenn es eine volle Fahne von V gibt, die f-invariant ist.*

Übung: Mit den Bezeichnungen des Satzes, sei $U \subset V$ ein f-invarianter Unterraum. Es sei $V \to W$ ein Faktorraum modulo U. Dann induziert f lineare Abbildungen $g : U \to U$ und $h : W \to W$. Man zeige, dass f trigonalisierbar ist, wenn g und h trigonalisierbar sind.

Satz 28 *Ein Endomorphismus $f : V \to V$ eines endlich erzeugten Vektorraums ist genau dann diagonalisierbar, wenn es eine direkte Zerlegung in eindimensionale f-invariante Unterräume gibt*

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_n.$$

Beweis: Angenommen es gibt eine solche direkte Zerlegung. Dann wählen wir $v_i \in W_i$, so dass $v_i \neq 0$ für $i = 1, \ldots n$. Da W_i f-invariant und eindimensional ist, gibt es $\lambda_i \in K$, so dass $f(v_i) = \lambda_i v_i$. Damit ist die Matrix von f in der Basis $v_1, \ldots, v_n$ eine Diagonalmatrix mit der Diagonale $\lambda_1, \ldots \lambda_n$.

Wenn umgekehrt f in dem Koordinatensystem $b : K^n \to V$ Diagonalgestalt hat, so kann man $W_i = \mathcal{L}(b(e_i))$ nehmen, wobei e_i der i-te Standardvektor ist.

Korollar 29 *Es sei $f : V \to V$ diagonalisierbar. Es sei $U \subset V$ ein f-invarianter Unterraum. Dann ist die Einschränkung $f : U \to U$ diagonalisierbar.*

Beweis: Wir benutzen die Bezeichnungen des Satzes. Es sei $\dim U = r$ und $\dim V = n$. Nach dem Austauschsatz finden wir Indexe $i_{r+1}, \ldots, i_n$, so dass wir eine direkte Zerlegung in Unterräume haben:

$$V = U \oplus W_{i_{r+1}} \oplus \ldots \oplus W_{i_n}.$$

Nach Umnumerierung dürfen wir $i_{r+1} = r + 1, \ldots, i_n = n$ annehmen. Wir betrachten die Projektion

$$\pi : V \to W_1 \oplus \ldots \oplus W_r.$$

längs des Unterraumes $W_{r+1} \oplus \ldots \oplus W_n$. Da Endomorphismus f jedes W_i invariant läßt, erhalten wir als Einschränkung von f einen Endomorphismus:

$$f' : W_1 \oplus \ldots \oplus W_r \to W_1 \oplus \ldots \oplus W_r.$$

Es gilt offensichtlich $\pi \circ f = f' \circ \pi$. Wenn man π auf U einschränkt erhält man einen Isomorphismus *Endomorphismus, nilpotenter*

$$\pi_{|U} : U \to W_1 \oplus \ldots \oplus W_r.$$

Dadurch ist $(U, f_{|U})$ isomorph zu $(W_1 \oplus \ldots \oplus W_r, f')$. Aber der letzte Endomorphismus ist offensichtlich diagonalisierbar. $Q.E.D.$

Übung: Wie findet man eine Zerlegung von U in 1-dimensionale invariante Unterräume?

1.7 Die Fittingzerlegung

Satz 30 *Es sei V ein Vektorraum der Dimension n. Es sei $f : V \to V$ ein Endomorphismus. Dann sind die folgenden Bedingungen äquivalent:*

1. *Es gibt eine Zahl $m \in \mathbb{N}$, so dass $f^m = 0$.*
2. *Es gibt eine Fahne*

   $$0 = V_0 \subset V_1 \subset V_2 \subset \ldots \subset V_{d-1} \subset V_d = V,$$

   so dass $f(V_t) \subset V_{t-1}$, für $t = 1, \ldots, d$.
3. *Es gibt eine Basis von V in der die Matrix von f eine obere Dreiecksmatrix mit Nullen auf der Diagonale ist.*
4. *Es gilt: $f^n = 0$.*

Man nennt einen Endomorphismus mit diesen Eigenschaften *nilpotent*

Beweis: Wenn f die erste Bedingung erfüllt, so kann man folgende Fahne nehmen:

$$0 = f^m(V) \subset f^{m-1}(V) \subset \ldots f(V) \subset V = V.$$

Diese Fahne hat die Eigenschaft 2.

Wenn eine Fahne unter 2. existiert, so folgt

$$f^d(V) = f^d(V_d) \subset f^{d-1}(V_{d-1}) \subset \ldots \subset f(V_1) \subset V_0 = 0. \tag{36}$$

Also sind die ersten beiden Eigenschaften äquivalent.

Es sei die Eigenschaft 2 erfüllt. Wir können annehmen, dass die Fahne echt ist, wenn wir evtl. einige der V_t entfernen. Wir betrachten die Matrix

von f in einer zerfällenden Basis. Die Vektorräume V_t sind invariant, da $f(V_t) \subset V_{t-1} \subset V_t$. Also ist die Matrix von f eine obere Blockmatrix (34). Aus $f(V_t) \subset V_{t-1}$ folgt, dass alle Blöcke $A_{tt} = 0$ sind. Damit ist die Matrix erst recht eine obere Dreiecksmatrix mit Nullen auf der Diagonalen.

Eine obere Dreiecksmatrix P mit Nullen auf der Diagonalen erfüllt die Eigenschaft 2 bezüglich der Standardfahne (35). Deshalb folgt aus der Eigenschaft 3 die Eigenschaft 2 und nach (36) die Eigenschaft 4. *Q.E.D.*

Bemerkung: Die Eigenschaften 1 und 2 sind auch äquivalent, wenn V nicht endlich erzeugt ist.

Man kann den letzten Satz noch präzisieren:

Korollar 31 *(Jordansche Normalform für nilpotente Endomorphismen) Es sei $f : V \to V$ wie im letzten Satz. Dann gibt es Vektoren $v_1, \ldots, v_t \in V$ und natürliche Zahlen $\ell_1, \ldots, \ell_t$, so dass*

$$f^{\ell_1}(v_1) = 0, \ldots, f^{\ell_t}(v_t) = 0,$$

und so dass die folgenden Vektoren eine Basis von V sind:

$$\begin{array}{c} v_1, f(v_1), \ldots f^{\ell_1 - 1}(v_1), \\ v_2, f(v_2), \ldots f^{\ell_2 - 1}(v_2), \\ \cdots \\ v_t, f(v_t), \ldots f^{\ell_t - 1}(v_t). \end{array} \tag{37}$$

Beweis: Man wähle ℓ, so dass $f^\ell(V) = 0$. Wir setzen $W_i = \operatorname{Ker} f^i$ für $i > 0$ und $W_0 = 0$. Wir betrachten die Fahne

$$V = W_\ell \supset W_{\ell-1} \supset \ldots \supset W_1 \supset W_0.$$

Man sieht, dass der Endomorphismus f für jedes $i \geq 2$ eine Injektion induziert:

$$f : W_i/W_{i-1} \to W_{i-1}/W_{i-2}. \tag{38}$$

Man wählt eine Basis $\bar{B}_\ell \subset W_\ell/W_{\ell-1}$. Es $B_\ell \subset W_\ell$ ein System von Repräsentanten für $\bar{B}_\ell$.

Das Bild $f(\bar{B}_\ell)$ von $\bar{B}_\ell$ bei der Injektion (38) für $i = \ell$ ist wieder eine linear unabhägige Menge und wir können sie zu einer Basis $f(\bar{B}_\ell) \cup \bar{B}_{\ell-1}$ von $W_{\ell-1}/W_{\ell-2}$ ergänzen.

Als nächstes betrachten wir die Injektion (38) für $i = \ell - 1$. Durch Basisergänzung erhalten wir eine Basis von $W_{\ell-2}/W_{\ell-3}$ von der Form

$$f^2(\bar{B}_\ell) \cup f(\bar{B}_{\ell-1}) \cup \bar{B}_{\ell-2}.$$

So fährt man fort. Man wählt für die Mengen $\bar{B}_i \subset W_i/W_{i-1}$ Repräsentantensysteme $B_i \subset W_i$.

Dann betrachtet man die folgenden Mengen

$$\begin{array}{ll} B_\ell & \subset W_\ell \\ f(B_\ell) \cup B_{\ell-1} & \subset W_{\ell-1} \\ \quad \cdots & \\ f^{\ell-i}(B_\ell) \cup \ldots \cup f(B_{i+1}) \cup B_i & \subset W_i \\ \quad \cdots & \\ . & \end{array}$$

Die Teilmenge von W_i in der $\ell - i + 1$-ten Zeile gibt uns nach Konstruktion eine Basis von W_i/W_{i-1}. Also geben alle Teilmengen zusammen eine Basis von V. Diese Basis hat die Form wie sie das Korollar fordert. *Q.E.D.*

Übung: Es sei $g : V \to V$ ein Endomorphismus, so dass $g \circ f = f \circ g$. Dann sind $\operatorname{Ker} g$ und $\operatorname{Im} g$ beides f-invariante Unterräume.

Korollar 32 *Es seien $f : V \to V$ und $g : V \to V$ Endomorphismen, so dass $g \circ f = f \circ g$.*

Wenn f und g nilpotent sind, so auch $f + g$.

Beweis: Weil f und g vertauschbar sind, folgt $g(\operatorname{Im} f^i) \subset \operatorname{Im} f^i$ für alle $i \in \mathbb{N}$. Wir betrachten die Fahne:

$$0 = \operatorname{Im} f^{m+1} \subset \operatorname{Im} f^m \subset \ldots \subset \operatorname{Im} f^2 \subset \operatorname{Im} f \subset V.$$

Alle Vektorräume in dieser Fahne sind invariant bei $f + g$. Wir betrachten die Faktorräume $W_i = \operatorname{Im} f^i / \operatorname{Im} f^{i+1}$ für $i \geq 0$. Dann induziert $f + g$ auf jedem dieser Faktorräume einen Endomorphismus. Es genügt zu zeigen, dass

$$f + g : W_i \to W_i$$

für alle i nilpotent ist. Da f auf W_i den Nullmorphismus induziert, ist das klar. *Q.E.D.*

Übung: Einen anderen Beweis erhält man, wenn man bedenkt, dass für $(f + g)^{2m}$ die binomische Formel gilt, wenn f und g vertauschbar sind.

Wie findet man invariante Unterräume eines Endomorphismus $f : V \to V$? Offensichtlich sind $\operatorname{Ker} f$ und $\operatorname{Im} f$ invariante Unterräume. Hier ist eine weitere Antwort:

Satz 33 *(Fittingzerlegung) Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraums über K. Dann existiert eine eindeutig bestimmte Zerlegung in f-invariante komplementäre Unterräume* *Fittingzerlegung*

$$V = V_{nil} + V_{bij}, \tag{39}$$

so dass $f : V_{nil} \to V_{nil}$ nilpotent ist und $f : V_{bij} \to V_{bij}$ ein Isomorphismus ist.

Ein Vektor $v \in V$ liegt genau dann in V_{nil}, wenn es eine natürliche Zahl m gibt, so dass $f^m(v) = 0$.

Beweis: Man hat eine aufsteigende Kette von Unteräumen von V:

$$0 \subset \operatorname{Ker} f \subset \cdots \subset \operatorname{Ker} f^m \subset \operatorname{Ker} f^{m+1} \subset \ldots$$

Da die Dimensionen dieser Unterräume durch $\dim V$ beschränkt sind, findet man eine Zahl t, so dass

$$\operatorname{Ker} f^t = \operatorname{Ker} f^{t+1} = \operatorname{Ker} f^{t+2} = \ldots$$

Andererseits hat man eine absteigende Folge von Unterräumen von V:

$$V \supset \operatorname{Im} f \supset \operatorname{Im} f^2 \supset \ldots$$

Die Dimensionen dieser Unterräume erreichen irgendwann einer minimalen Wert. Daher findet man eine Zahl s, so dass

$$\operatorname{Im} f^s = \operatorname{Im} f^{s+1} = \operatorname{Im} f^{s+2} \ldots$$

Es sei $u = \max(s, t)$. Wir setzen $V_{nil} = \operatorname{Ker} f^u, V_{bij} = \operatorname{Im} f^u$.

Behauptung: $\operatorname{Ker} f^u \cap \operatorname{Im} f^u = 0$.

In der Tat, es sei $x \in \operatorname{Ker} f^u \cap \operatorname{Im} f^u$. Dann gibt es ein $y \in V$, so dass $x = f^u(y)$. Andererseits gilt $f^u(x) = 0$. Wir finden

$$0 = f^u(x) = f^{2u}(y) \Rightarrow y \in \operatorname{Ker} f^{2u} = \operatorname{Ker} f^u \Rightarrow f^u(y) = 0 \Rightarrow x = 0.$$

Das zeigt die Behauptung.

Behauptung: $\operatorname{Ker} f^u + \operatorname{Im} f^u = V$.

Es sei $v \in V$. Da $\operatorname{Im} f^u = \operatorname{Im} f^{2u}$ findet man ein $w \in V$, so dass $f^{2u}(w) = f^u(v)$, d.h. $f^u(v - f^u(w)) = 0$. Also gilt $v - f^u w \in \operatorname{Ker} f^u$. Wir finden für v die Darstellung:

$$v = f^u(w) + (v - f^u(w)) \in \operatorname{Im} f^u + \operatorname{Ker} f^u$$

Offenbar ist f auf dem ersten Summanden $\operatorname{Ker} f^u$ nilpotent und auf dem zweiten Summanden $\operatorname{Im} f^u$ surjektiv und damit ein Isomorphismus. Damit ist die Existenz der Fittingzerlegung bewiesen.

Wir zeigen die Eindeutigkeit. Es sei $W = W_{bij} + W_{nil}$ eine weitere Zerlegung in komplementäre f-invariante Unterräume wie im Satz. Dann gibt es eine Zahl m, so dass $f^m(W_{nil}) = 0$ und folglich gilt $W_{nil} \subset V_{nil}$. Andererseits sei $v \in V_{nil} = \operatorname{Ker} f^u$. Der Vektor v hat eine Zerlegung

$$v = w_{bij} + w_{nil}, \quad w_{bij} \in W_{bij},\ w_{nil} \in W_{nil}.$$

Wir folgern, dass

$$0 = f^u(v) = f^u(w_{bij}) + f^u(w_{nil}).$$

Der erste Summand auf der rechten Seite ist in W_{bij} und der zweite Summand in W_{nil}. Da beide Räume komplementär sind, sind beide Summanden gleich 0. Weil f auf W_{bij} ein Isomorphismus ist folgt $w_{bij} = 0$. Wir finden $v \in W_{nil}$. Damit haben wir auch gezeigt, dass $V_{nil} \subset W_{nil}$. Die Räume sind also gleich $V_{nil} = W_{nil}$.

Ein Element aus W_{bij} liegt im Bild jeder Abbildung f^m. Daher folgt $W_{bij} \subset V_{bij}$.

Es sei umgekehrt $v \in V_{bij}$. Wir wählen wieder m, so dass $f^m(W_{nil}) = 0$. Wir finden $w \in V$, so dass $f^m(w) = v$. Wir schreiben $w = w_{bij} + w_{nil}$. Wenn wir f^m auf diese Gleichung anwenden, so folgt $v = f^m(w) = f^m(w_{bij}) \in W_{bij}$. Also $V_{bij} \subset W_{bij}$ und damit sind die Vektorräume gleich. Damit ist auch die Eindeutigkeit der Zerlegung gezeigt. *Q.E.D.*

Man kann die letzte Behauptung auch aus den Dimensionsformeln erhalten. Wegen der ersten Behauptung folgt aus den Dimensionsformeln:

$$\dim(\operatorname{Ker} f^u + \operatorname{Im} f^u) = \dim \operatorname{Ker} f^u + \dim \operatorname{Im} f^u = \dim V.$$

Also sind $\operatorname{Ker} f^u$ und $\operatorname{Im} f^u$ komplementäre Räume.

Die Fittingzerlegung ist kanonisch. Das bedeutet folgendes: Es sei T ein weiterer Vektorraum und $g : T \to T$ ein Endomorphismus. Es sei $T = T_{nil} +$

T_{bij} die Fittingzerlegung bzgl. g. Es sei $\alpha : V \to T$ ein Homomorphismus, so dass folgendes Diagramm kommutativ ist:

Eigenwert
Eigenvektor
Hauptraum
verallgemeinerten Eigenraum

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\alpha} \\ T & \xrightarrow{g} & T \end{array} \tag{40}$$

Dann gilt $\alpha(V_{nil}) \subset T_{nil}$ und $\alpha(V_{bij}) \subset T_{bij}$.

In der Tat, es ist klar, dass $\alpha(\operatorname{Ker} f^m) \subset \operatorname{Ker} g^m$ und $\alpha(\operatorname{Im} f^m) \subset \operatorname{Im} g^m$ für alle $m \in \mathbb{N}$.

Die Eigenschaft kanonisch zu sein, hat zwei Folgerungen:

1) Wenn $T \subset V$ ein invarianter Unterraum ist und $f(T) = T$, dann gilt $T \subset V_{bij}$.

2) Es sei $h : V \to V$ ein Endomorphismus, der mit f kommutiert, d.h. $f \circ h = h \circ f$. Dann sind V_{nil} und V_{bij} invariante Unterräume bezüglich h.

Definition 34 *Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraumes über einem Körper K. Ein Element $\lambda \in K$ heißt ein* Eigenwert *, wenn $f - \lambda \operatorname{id}_V$ kein Isomorphismus ist.*

Dann gibt es einen Vektor $v \in V$, $v \neq 0$, so dass $f(v) = \lambda v$. Man nennt v einen *Eigenvektor* zum Eigenwert λ. Der Unterraum $\mathcal{L}(v)$ von V ist f-invarianter Unterraum von V.

Für einen Eigenwert λ betrachten wir die Fittingzerlegung bezüglich des Endomorphismus $f - \lambda \operatorname{id}_V$:

$$V = V_{nil} \oplus V_{bij}$$

Es ist klar, dass V_{nil} und V_{bij} auch f-invariante Unterräume sind, z.B. weil die Endomorphismen $f - \lambda \operatorname{id}_V$ und f vertauschbar sind. Die Einschränkung $f_{|V_{bij}}$ hat nicht den Eigenwert λ, da $f - \lambda \operatorname{id}_V$ auf V_{bij} einen Isomorphismus induziert. Man setzt $V(\lambda) = V_{nil}$ und nennt das den *Hauptraum* oder deb *verallgemeinerten Eigenraum* zum Eigenwert λ. Wenn λ kein Eigenwert von f ist, so setzen wir $V(\lambda) = 0$. Es gilt, dass $V(\lambda) = \operatorname{Ker}(f - \lambda \operatorname{id}_V)^{\dim V}$.

Lemma 35 *Es sei $g : T \to T$ ein nilpotenter Endomorphismus eines Vektorraumes. Dann ist für alle $\mu \in K$, $\mu \neq 0$ der Endomorphismus $g - \mu \operatorname{id}_T : T \to T$ ein Isomorphismus. (T braucht nicht endlich erzeugt zu sein.)*

Zerlegung in Hauptr""aume

Beweis: Man schreibt $g - \mu\,\mathrm{id}_T = -\mu(\mathrm{id}_T - \mu^{-1}g)$. Wir betrachten den nilpotenten Endomorphismus $\alpha = \mu^{-1}g$. Nach Voraussetzung gibt es eine Zahl n, so dass $\alpha^n = 0$. Es genügt zu beweisen, dass $\mathrm{id}_T - \alpha : T \to T$ ein Isomorphismus ist.

Wir zeigen, dass

$$\mathrm{id}_T + \alpha + \alpha^2 + \ldots + \alpha^{n-1} : T \to T$$

invers zu $\mathrm{id}_T - \alpha$ ist. In der Tat:

$$\begin{array}{ll} & (\mathrm{id}_T - \alpha) \circ (\mathrm{id}_T + \alpha + \alpha^2 + \ldots + \alpha^{n-1}) = \\ \mathrm{id}_T & + \alpha + \alpha^2 + \ldots + \alpha^{n-1} \\ & - \alpha - \alpha^2 - \ldots - \alpha^{n-1} - \alpha^n = \\ & = \mathrm{id}_T - \alpha^n = \mathrm{id}_T\,. \end{array}$$

Für das Kompositum

$$(\mathrm{id}_T + \alpha + \alpha^2 + \ldots + \alpha^{n-1}) \circ (\mathrm{id}_T - \alpha)$$

erhält man auf die gleiche Weise id_T. *Q.E.D.*

Satz 36 *Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraumes über einem Körper K. Dann gibt es nur endlich viele verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_r$ von f. Es seien $V(\lambda_1), \ldots, V(\lambda_r)$ die Haupträume.*

Dann gibt es einen f-invarianten Unterraum W von V, so dass die Einschränkung von f auf W keine Eigenwerte besitzt und so dass V eine innere direkte Summe ist:

$$V = V(\lambda_1) \oplus \ldots \oplus V(\lambda_r) \oplus W. \tag{41}$$

Die Zerlegung (41) ist kanonisch (vgl (40)).

Beweis: Wenn f keinen Eigenwert besitzt, so ist der Satz mit $V = W$ bewiesen.

Es sei λ_1 ein Eigenwert von f. Man nimmt die Fittingzerlegung von $f - \lambda_1\,\mathrm{id}_V$:

$$V = V(\lambda_1) \oplus U. \tag{42}$$

Dann gilt $\dim U < \dim V$. Wir führen eine Induktion nach $\dim V$ durch. Wenn $\dim V = 0$ so ist der Satz trivial. Wir wenden die Induktionsvoraussetzung auf die Einschränkung von $f_{|U} : U \to U$ an. Es seien $\lambda_2, \ldots, \lambda_r$ die

Eigenwerte von $f_{|U}$. Da $f_{|U} - \lambda_1 \mathrm{id}_U$ ein Isomorphismus ist, kann λ_1 unter den Eigenwerten von $f_{|U}$ nicht vorkommen. Nach Induktionsvoraussetzung können wir schreiben:

$$U = U(\lambda_2) \oplus \ldots \oplus U(\lambda_r) \oplus W$$

Wir setzen das für U in die Zerlegung (42) ein. Wenn wir zeigen können, dass $U(\lambda_i) = V(\lambda_i)$ für $i = 2, \ldots r$, so ist der Satz bewisen.

Es sei $i \geq 2$. Nach dem letzten Lemma ist die Einschränkung von $f - \lambda_i \mathrm{id}_V$ auf jeden der Unterräume $U(\lambda_j)$ $j \neq 2$, und $V(\lambda_1)$ ein Isomorphismus. Die Einschränkung auf W ist auch ein Isomorphismus, da $f_{|W}$ keine Eigenwerte besitzt. Es sei

$$Q = V(\lambda_1) \oplus U(\lambda_2) \oplus \ldots \widehat{U(\lambda_i)} \ldots \oplus U(\lambda_r) \oplus W. \tag{43}$$

Das Dach deutet an, dass wir den Summanden $U(\lambda_i)$ ausgelassen haben. $f - \lambda_i \mathrm{id}_V$ ist eingeschränkt auf jeden Summanden auf der rechten Seite von (43) ein Isomorphismus und folglich ist auch die Einschränkung auf den Unterraum $Q \subset V$ ein Isomophismus. Wir sehen, dass

$$V = U(\lambda_i) \oplus Q$$

die Fittingzerlegung des Endomorphismus $f - \lambda_i \mathrm{id}_V$ ist. Damit ist $U(\lambda_i) = V(\lambda_i)$ bewiesen. *Q.E.D.*

Als nächstes wollen wir Eigenwerte finden. Wir haben schon zur Fittingzerlegung von $f : V \to V$ bemerkt, dass es gut ist Endomorphismen zu kennen, die mit f vertauschen.

Wir betrachten ein Polynom

$$\mathbf{p}(X) = a_d X^d + a_{n-1} X^{n-1} + \ldots + a_1 X + a_0, \quad a_i \in K.$$

Man kann f in das Polynom einsetzen und erhält den Endomorphismus

$$\mathbf{p}(f) = a_d f^d + a_{n-1} f^{n-1} + \ldots + a_1 f + a_0 \mathrm{id}_V .$$

Das wichtigste ist hier, dass man mit der Potenz $f^t = f \circ \ldots \circ f$ das Kompositum von Abbildungen meint. Es ist klar, dass $\mathbf{p}(f)$ mit f vertauschbar ist. Allgemeiner ist es nicht schwer die folgenden Rechenregeln einzusehen. Es sei $\mathbf{q}$ ein weiteres Polynom. Dann gilt:

$$\begin{aligned} \mathbf{p}(f) + \mathbf{q}(f) &= (\mathbf{p} + \mathbf{q})(f) \\ \mathbf{p}(f) \circ \mathbf{q}(f) &= (\mathbf{pq})(f). \end{aligned}$$

Hier steht auf der rechten Seite die Summe und das Produkt von Polynomen. Auf der linken Seite steht die Summe und das Kompositum von linearen Abbildungen.

Damit können wir Eigenwerte berechnen. Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraums. Wir wählen einen Vektor $v \in V$, $v \neq 0$. Dann gibt es eine größte natürliche Zahl d, so dass die Vektoren

$$v, f(v), \ldots, f^{d-1}(v). \tag{44}$$

linear unabhängig sind. Also findet man eine Relation

$$f^d(v) + a_{d-1}f^{d-1}(v) + \ldots + a_1 f(v) + a_0 v = 0, \tag{45}$$

Aus dieser Relation folgert man, dass

$$f^d(v) \in \mathcal{L}(v, f(v), \ldots, f^{d-1}(v)).$$

Also ist $\mathcal{L}(v, f(v), \ldots, f^{d-1}(v))$ ein f-invarianter Unterraum von V, der den Vektor v enthält. Im allgemeinen muss man natürlich damit rechnen, dass dieser Unterraum gleich V ist. Wir betrachten jetzt das Polynom

$$\mathbf{p}(X) = X^d + a_{d-1}X^{d-1} + \ldots + a_1 X + a_0.$$

Dann gilt nach (45), dass

$$\mathbf{p}(f)(v) = 0 \tag{46}$$

Wir werden zeigen, dass die Nullstellen von $\mathbf{p}(X)$ Eigenwerte von f sind.

Dazu machen wir eine allgemeine Bemerkung über Polynome: Es sei λ eine Nullstelle von $\mathbf{p}(X)$. Dann gibt es eine Faktorisierung

$$\mathbf{p}(X) = (X - \lambda) \cdot \mathbf{q}(X), \tag{47}$$

wobei $\mathbf{q}(X) = X^{d-1} + b_{d-2}X^{d-2} + \ldots b_1 X + b_0$ und die $b_i \in K$ geeignete Koeffizienten sind.

Wir nehmen an, dass λ eine Nullstelle von $\mathbf{p}$ ist und beweisen, dass es dann eine Faktorisierung (47) gibt. Zunächst bemerken wir, dass die Gleichung (47) äquivalent mit dem folgenden Gleichungssystem ist.

$$a_n = b_{n-1},\ a_{n-1} = -\lambda b_{n-1} + b_{n-2},\ \ldots,\ a_1 = \lambda b_1 + b_0,\ a_0 = -\lambda b_0.$$

Aus den ersten n dieser $n+1$ Gleichungen eliminiere man $b_{n-1}, \ldots, b_0$ und zeige, dass dann die letzte Gleichung $a_0 = -\lambda b_0$ erfüllt ist, weil $\mathbf{p}(\lambda) = 0$.

Jetzt können wir zeigen, dass λ ein Eigenwert von f ist. Wir setzen f in die Gleichung (47) ein und finden

$$\mathbf{p}(f) = (f - \lambda \,\mathrm{id}_V) \circ \mathbf{q}(f).$$

Wir wenden diesen Endomorphismus auf v an, und finden nach (46)

$$0 = (f - \lambda \,\mathrm{id}_V)(\mathbf{q}(f)(v)). \tag{48}$$

Es sei

$$w := (\mathbf{q}(f)(v)) = f^{d-1}(v) + b_{d-2}f^{d-2}(v) + \ldots b_1 f(v) + b_0 v.$$

Wegen der linearen Unabhängigkeit der Vektoren (44) ist das ein von 0 verschiedener Vektor. Aus der Gleichung (48) folgt, dass w ein Eigenvektor ist mit dem Eigenwert λ ist. Wir haben bewiesen, dass jede Nullstelle von $\mathbf{p}$ ein Eigenwert von f ist.

Übung: Es sei $U = \mathcal{L}(v, f(v), \ldots, f^{d-1}(v))$. Dann gilt für jeden Vektor $u \in U$, dass $\mathbf{p}(f)(u) = 0$. Wenn u ein Eigenvektor mit dem Eigenwert μ ist, so folgt, dass μ eine Nullstelle von $\mathbf{p}$ ist.

Definition 37 *Man nennt einen Körper K algebraisch abgeschlossen, wenn für jedes Polynom $\mathbf{p}(X) = X^d + a_{d-1}X^{d-1} + \ldots + a_1 X + a_0$ mit $d \geq 1$ und $a_i \in K$ ein $\lambda \in K$ existiert, so dass $\mathbf{p}(\lambda) = 0$, d.h. $\mathbf{p}$ hat eine Nullstelle.*

Wir zitieren ohne Beweis einen berühmten Satz:

Satz 38 *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

Es sei K ein algebraisch abgeschlossener Körper und $\mathbf{p}(X) = a_d X^d + a_{d-1}X^{d-1} + \ldots + a_1 X + a_0$ mit Koeffizienten in K und $a_d \neq 0$. Dann gibt es $\lambda_i \in K$, für $i = 1, \ldots, d$, so dass

$$\mathbf{p}(X) = a_d \prod_{i=1}^{d} (X - \lambda_i). \tag{49}$$

Wenn für ein Polynom $\mathbf{p}(X)$ über einem beliebigen Körper K eine solche Darstellung existiert, so sagen wir, dass $\mathbf{p}(X)$ in lineare Faktoren zerfällt.

Satz 39 *Es sei K ein algebraisch abgeschlossener Körper. Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraums $V \neq 0$.*

Dann hat f einen Eigenwert.

Genauer ist V die innere direkte Summe seiner Haupträume

$$V = V(\lambda_1) \oplus \ldots \oplus V(\lambda_r).$$

Beweis: Dass f einen Eigenwert besitzt folgt, da $\mathbf{p}$ aus (46) eine Nullstelle besitzt. Der zweite Teil folgt aus Satz 36 , da es keinen von 0 verschiedenen f-invarianten Vektorraum W geben kann, wo die Einschränkung von f keine Eigenwerte besitzt. *Q.E.D.*

Korollar 40 *Der letzte Satz bleibt richtig, wenn wir die Voraussetzung K algebraisch abgeschlossen durch die Voraussetzung ersetzen, dass ein zerfallendes Polynom $\mathbf{p}$ existiert, so dass $\mathbf{p}(f) = 0$. Darüber hinaus ist dann jeder Eigenwert von f eine Nullstelle von $\mathbf{p}$.*

Beweis: Wir zeigen zuerst, dass f einen Eigenwert besitzt. Dazu nehmen wir an, dass das Polynom $\mathbf{p}$ die Zerlegung (49) besitzt. Wir können annehmen, dass λ_d kein Eigenwert ist. Dann ist $f - \lambda_d \operatorname{id}_V : V \to V$ ein Isomorphismus. Aus der Gleichung

$$0 = a_d(f - \lambda_1 \operatorname{id}_V) \circ \ldots \circ (f - \lambda_{d-1} \operatorname{id}_V) \circ (f - \lambda_d \operatorname{id}_V),$$

folgt dann, dass

$$0 = a_d(f - \lambda_1 \operatorname{id}_V) \circ \ldots \circ (f - \lambda_{d-1} \operatorname{id}_V).$$

Damit haben wir ein zerfallendes Polynom

$$\mathbf{q}(X) = a_d \prod_{i=1}^{d-1} (X - \lambda_i).$$

von kleinerem Grad, so dass $\mathbf{q}(f) = 0$. Mit Hilfe von Induktion schließen wir, dass f einen Eigenwert hat.

Es ist klar, dass dann auch die Einschränkung von f auf jeden invarianten Unterraum einen Eigenwert besitzt. Deshalb muss in der Zerlegung (41) der Vektorraum $W = 0$ sein.

Dass jeder Eigenwert von f eine Nullstelle von p ist, sei eine Übungaufgabe. *Q.E.D.*

Satz 41 *Es sei K ein Körper. Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraums. Es sei $P \in K[T]$ ein von 0 verschiedenes Polynom, das in lineare Faktoren zerfällt und so dass $P(f) = 0$. Dann ist f trigonalisierbar.*

Beweis:: Nach Satz 39 kann man annehmen, dass $V = V(\lambda)$. Man schreibt $f = \lambda \operatorname{id}_V + (f - \lambda \operatorname{id}_V)$. Da hier der erste Summand in jeder Basis eine Diagonalmatrix ist, genügt es den zweiten Summanden zu trigonalisieren. Da $(f - \lambda \operatorname{id}_V)$ nilpotent ist, ist die Trigonalisierung nach Satz 30 möglich. *Q.E.D.*
Bemerkung: Wenn $K = \mathbb{C}$ ist, so existiert stets ein Polynom P mit den genannten Eigenschaften. Folglich ist für $K = \mathbb{C}$ jeder Endomorphismus diagonalisierbar.

Satz 42 *Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraums über einem Körper K.*

Der Endomorphismus f ist genau dann diagonalisierbar, wenn ein Polynom

$$\mathbf{p}(X) = \prod_{i=1}^{d} (X - \lambda_i), \quad \lambda_i \in K, \tag{50}$$

existiert, so dass sämtliche λ_i verschieden sind und

$$\mathbf{p}(f) = 0.$$

Beweis: Angenommen $\mathbf{p}(X)$ existiert. Nach Korollar 40 können wir $V = V(\lambda)$ annehmen. Wir wissen, dass

$$f - \lambda_i \operatorname{id}_V$$

für $\lambda_i \neq \lambda$ ein Isomorphismus ist. Wenn wir aus dem Produkt (50) diese Faktoren $X - \lambda_i$ streichen, bleibt die Gleichung $\mathbf{p}(f) = 0$ erhalten. Da λ der einzige Eigenwert ist, muss dann

$$\mathbf{p} = (X - \lambda)$$

gelten. Dann ist $f = \lambda \operatorname{id}_V$ offensichtlich diagonalisierbar.

Wenn umgekehrt f diagonalisierbar ist, so ist die Existenz eines Polynoms $\mathbf{p}$ offensichtlich. *Q.E.D.*

Beispiel: Es sei V ein endlich erzeugter $\mathbb{C}$-Vektorraum. Es sei $f : V \to V$ ein Endomorphismus. Es existiere eine Zahl $t \in \mathbb{N}$, so dass $f^t = \mathrm{id}_V$. Dann ist f diagonalisierbar. In der Tat, das Polynom $X^t - 1$ hat keine mehrfachen Nullstellen.

Wir betrachten nochmal die Situation (44): Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten K-Vektorraums. Es sei $v \in V$, wo $v \neq 0$. Dann ist $U = \mathcal{L}(v, f(v), \ldots, f^{d-1}(v))$ ein f-invarianter Unterraum von V. Die Nullstellen der Polynoms $\mathbf{p}(X)$ sind genau die Eigenwerte des Endomorphismus $f_{|U} : U \to U$, der durch Einschränkung von f entsteht. Um weitere Eigenwerte von f zu finden benutzen wir einen Faktorraum $\sigma : V \to W$ modulo U. Da U f-invariant ist, induziert f eine lineare Abbildung $h : W \to W$.

Satz 43 *Es sei ein kommutatives Diagramm von Vektorräumen mit exakten Zeilen gegeben:*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U & \xrightarrow{\varkappa} & V & \xrightarrow{\sigma} & W & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & U & \xrightarrow{\varkappa} & V & \xrightarrow{\sigma} & W & \longrightarrow & 0 \end{array}$$

Für $\lambda \in K$ seien $U(\lambda)$, bzw. $V(\lambda)$, bzw. $W(\lambda)$ der Hauptraum von g, bzw. f, bzw. h. Dann induzieren die Abbildungen $\varkappa$ und σ eine kurze exakte Sequenz der Haupträume:

$$0 \to U(\lambda) \to V(\lambda) \to W(\lambda) \to 0 \tag{51}$$

Beweis: Das obige Diagramm bleibt kommutativ, wenn wir f, g, h durch $f - \lambda \,\mathrm{id}_V, g - \lambda \,\mathrm{id}_U, h - \lambda \,\mathrm{id}_W$ ersetzen. Wir betrachten die Fittingzerlegungen der drei letzten Endomorphismen:

$$U = U(\lambda) \oplus \breve{U}, \quad V = V(\lambda) \oplus \breve{V}, \quad W = W(\lambda) \oplus \breve{W}.$$

Da die Fittingzerlegung kanonisch ist, folgt $\varkappa(U(\lambda)) \subset V(\lambda)$ und $\varkappa(\breve{U}) \subset \breve{V}$. Das entsprechende gilt für σ.

Die Injektivität von $U(\lambda) \to V(\lambda)$ ist klar, da $U(\lambda) \subset U$ und $V(\lambda) \subset V$. Ähnlich leicht sieht man, dass der Kern von $V(\lambda) \to W(\lambda)$ gleich dem Bild von $U(\lambda)$ ist. In der Tat: Es sei $v \in V(\lambda)$ aus dem Kern. Dann gilt $v = \varkappa(u)$ für ein $u \in U$. Es sei $u = u_1 + u_2$ die Fittingzerlegung, wo $u_1 \in U(\lambda)$ und $u_2 \in \breve{U}$. Dann ist $v = \varkappa(u_1) + \varkappa(u_2)$ die Fittingzerlegung von $v \in V(\lambda)$. Daher folgt $\varkappa(u_2) = 0$ und $v = \varkappa(u_1)$ was wir zeigen wollten.

Schließlich zeigen wir die Surjektivität von $V(\lambda) \to W(\lambda)$. Es sei $w_1 \in W(\lambda)$. Es existiert ein $v \in V$, so dass $\sigma(v) = w_1$. Der Vektor v zerlegt sich $v = v_1 + v_2$, wo $v_1 \in V(\lambda)$ und $v_2 \in \breve{V}$. Wenn wir σ anwenden, erhalten wir

$$\sigma(v_1) + \sigma(v_2) = w_1 \in W(\lambda).$$

Da die Fittingzerlegung kanonisch ist, steht links die Fittingzerlegung von w_1. Also gilt $\sigma(v_2) = 0$ und $\sigma(v_1) = w_1$. Das wollten wir beweisen. *Q.E.D.*

Satz 44 *(Jordanzerlegung) Es sei K ein Körper. Es sei $f : V \to V$ ein trigonalisierbarer Endomorphismus eines endlich erzeugten Vektorraumes. Dann gibt es eine eindeutig bestimmte Zerlegung*

$$f = f_s + f_n, \quad f_s, f_n \in \mathrm{End}(V),$$

mit den folgenden Eigenschaften

1. *$f_s : V \to V$ ist diagonalisierbar.*
2. *$f_n : V \to V$ ist nilpotent.*
3. *Es gilt $f_s \circ f_n = f_n \circ f_s$.*

Wir bemerken, dass für $K = \mathbb{C}$ jeder Endomorphismus trigonalisierbar ist und folglich eine Jordanzerlegung $f = f_s. + f_n$ besitzt.

Beweis: Wir betrachten die Zerlegung in Haupträume $V(\lambda)$ bezüglich f. Da f_s und f_n nach Eigenschaft 3 mit f vertauschbar sein sollen, müssen sie die Haupträume invariant lassen. Deshalb genügt es den Satz in dem Fall zu beweisen, wo $V = V(\lambda)$.

Angenommen wir hätten eine Jordanzerlegung. Dann gilt:

$$f_s - \lambda \,\mathrm{id}_V = (f - \lambda \,\mathrm{id}_V) - f_n.$$

Auf der rechten Seite steht die Differenz von zwei nilpotenten Endomorphismen, die vertauschbar sind. Also ist die rechte Seite nilpotent. Andererseits ist die linke Seite der Gleichung diagonalisierbar. Also ist die rechte Seite ähnlich zu einer nilpotenten Diagonalmatrix. Aber davon gibt es nur die 0-Matrix. Deshalb muss $f_s = \lambda \,\mathrm{id}_V$ und $f_n = (f - \lambda \,\mathrm{id}_V)$ gelten. Aber die Zerlegung

$$f = \lambda \,\mathrm{id}_V + (f - \lambda \,\mathrm{id}_V)$$

erfüllt offenbar die Eigenschaften einer Jordanzerlegung. *Q.E.D.*

multilinear

Satz 45 *Es sei K ein algebraisch abgeschlossener Körper. Es sei V ein endlich erzeugter Vektorraum und $f : V \to V$ ein Endomorphismus. Dann gibt es eine Zerlegung in f-invariante Unterräume $V_i \subset V$, $i = 1, \ldots, t$*

$$V = \oplus_{i=1}^{t} V_i,$$

so dass für alle i die Matrix von $f_{|V_i}$ in einer geeigneten Basis die folgende Gestalt hat:

$$\begin{pmatrix} \lambda_i & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_i & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \lambda_i & 1 & \ldots & 0 & 0 \\ & & & \ldots & & & \\ 0 & 0 & 0 & 0 & \ldots & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & \lambda_i \end{pmatrix}.$$

Beweis: Man kann sich auf den Fall beschränken, dass es nur einen einzigen Hauptraum gibt. $V = V(\lambda)$. Dann ist $f - \lambda \, \mathrm{id}_V)$ ein nilpotenter Endomorphismus von V. Daher folgt der Satz aus Korollar 31. *Q.E.D.*

1.8 Determinanten

Es seien $V_1, \ldots, V_m, U$ Vektorräume. Wir betrachten Abbildungen:

$$\Phi : V_1 \times \ldots \times V_m \to U.$$

Man nennt Φ *multilinear* , wenn für alle $i = 1, \ldots, m$ und für alle fest gewählten $v_j \in V_j$ für $j \neq i$, die Funktion

$$f(v_i) = \Phi(v_1, \ldots, v_i, \ldots, v_m), \quad v_i \in V_i$$

eine lineare Funktion $V_i \to U$ ist. Wenn $U = K$ spricht man von einer Multilinearform. Wenn $m = 2$ so spricht man von einer bilinearen Abbildung.

Wir betrachten den Vektorraum $\mathbb{R}^2$ über dem Körper $\mathbb{R}$. Wir machen eine heuristische Betrachtung: Es seien $x, v, w \in \mathbb{R}^2$. Es sei

$$Q(v, w) = \{x + \lambda v + \mu w \mid 0 \leq \lambda \leq 1, \ 0 \leq \lambda \leq 1\},$$

das Parallelogramm, welche von v und w in x aufgespannt wird. Seine Fläche $D(v, w)$ ist von x unabhängig. Wir zählen $D(v, w)$ positiv, wenn der

Drehwinkel $0 < \sphericalangle(v,w) < 180^o$ und negativ, wenn $180^o < \sphericalangle(v,w) < 360^o$. Dann ist anschaulich klar, dass $D(v,w)$ folgende Eigenschaften hat:

$$\begin{array}{ll} 1) & D(v_1+v_2,w) = D(v_1,w) + D(v_2,w), \quad D(\lambda v, w) = \lambda D(v,w). \\ 2) & D(v, w_1+w_2) = D(v,w_1) + D(v,w_2), \quad D(v, \mu w) = \mu D(v,w). \\ 3) & D(v,w) = -D(w,v) \\ 4) & D(v,v) = 0 \end{array}$$

Die ersten beiden Bedingungen besagen, dass der orientierte $D(v,w)$ Flächeninhalt eine bilineare Abbildung ist.

$$D : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}.$$

Wir wollen multlineare Abbildungen Φ mit $V = V_1 = \ldots = V_m$ betrachten. Wir nennen eine multilineare Abbildung

$$\Phi : V \times \ldots \times V \to U \tag{52}$$

alternierend, wenn für alle Folgen von Vektoren $w_1, \ldots, w_m \in V$ so dass $w_i = w_j$ für ein Paar von Indexen $i \neq j$:

$$\Phi(w_1, \ldots, w_m) = 0. \tag{53}$$

Wir bezeichnen die Menge der alternierenden multilinearen Abbildungen in m Argumenten mit $\mathrm{Alt}^m(V,U)$.

Lemma 46 *Es seien V und U Vektorräume und es sei $\vartheta \in \mathrm{Alt}^m(V,U)$.*

Es seien $i, j \in [1,m]$, so dass $i \neq j$ und es sei $\lambda \in K$. Dann gilt für beliebige Vektoren $v_1, \ldots, v_m \in V$, dass

$$\vartheta(v_1, \ldots, v_i, \ldots, v_j, \ldots v_m) = \vartheta(v_1, \ldots, v_i + \lambda v_j, \ldots, v_j, \ldots v_m).$$

Man kann also zu einem Argument der Funktion ϑ das Vielfache eines anderen Arguments addieren, ohne dass sich der Wert von ϑ ändert (Invarianz bei Scherungen).

Wenn $v_1, \ldots, v_m$ linear abhängig sind, so gilt

$$\vartheta(v_1, \ldots, v_m) = 0.$$

Insbesondere gilt für $m > \dim V$, dass $\mathrm{Alt}^m(V,U) = 0$.

Lemma 47 *Es sei $\Phi : V \times V \to U$ eine alternierende bilineare Abbildung. Dann gilt für beliebige Vektoren $w_1, w_2 \in V$, dass*

$$\Phi(w_1, w_2) = -\Phi(w_2, w_1). \tag{54}$$

Es sei K ein Körper, dessen Charakteristik nicht 2 ist. Es sei $\Phi : V \times V \to K$ eine Bilinearform, die (54) erfüllt. Dann ist Φ alternierend.

Beweis: Aus der Bilinearität von f folgt die Gleichung

$$\Phi(w_1 + w_2, w_1 + w_2) = \Phi(w_1, w_1) + \Phi(w_1, w_2) + \Phi(w_2, w_1) + \Phi(w_2, w_2).$$

Wegen (53) erhält man die gewünschte Gleichung (54). Wenn umgekehrt (54) erfüllt ist, so folgt

$$\Phi(w, w) = -\Phi(w, w).$$

Da $2 \neq 0$ vorausgesetzt ist, erhalten wir $\Phi(w, w) = 0$. *Q.E.D.*

Im allgemeinen sei $\Phi \in \mathrm{Alt}^m(V, U)$. Wir lassen alle Argumente von Φ bis auf zwei verschiedene i und j fest. Dann ergibt sich aus Lemma 47

$$\Phi(\ldots, v_i, \ldots, v_j, \ldots) = -\Phi(\ldots, v_j, \ldots, v_i, \ldots).$$

Wir folgern, dass für eine beliebige Permutation aus $\sigma \in \mathcal{S}_m$

$$\Phi(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(m)}) = (\operatorname{sgn}\sigma)\Phi(v_1, v_2, \ldots, v_m). \tag{55}$$

Wenn die Charakteristik von K nicht 2 ist, so ist eine multilineare Abbildung Φ, die (55) erfüllt nach Lemma 47 alternierend.

Beispiel: Der orientierte Flächeninhalt ist eine alternierende Bilinearform $D \in \mathrm{Alt}^2(\mathbb{R}^2, \mathbb{R})$. Es seien e_1 und e_2 die Standardvektoren. Es sei $v = x_1e_1 + x_2e_2$ und $w = y_1e_1 + y_2e_2$. Da D alternierend ist finden wir

$$\begin{aligned}
&D(x_1e_1 + x_2e_2, y_1e_1 + y_2e_2) = x_1D(e_1, y_1e_1 + y_2e_2) + x_2D(e_2, y_1e_1 + y_2e_2) = \\
&x_1y_1D(e_1, e_1) + x_1y_2D(e_1e_2) + x_2y_1D(e_2, e_1) + x_2y_2D(e_2, e_2) = \\
&x_1y_2D(e_1e_2) + x_2y_1D(e_2, e_1) = \\
&(x_1y_2 - x_2y_1)D(e_1, e_2).
\end{aligned}$$

Die Zahl $D(e_1, e_2)$ ist die Fläche eines Quadrats mit der Kantenlänge 1. Wir geben ihr die Maßeinheit 1. Wir definieren die Determinante:

$$\det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} := D(v, w) = (x_1y_2 - x_2y_1).$$

Das ist der orientierte Flächeninhalt des Parallelogramms, welches von den Spalten der Matrix aufgespannt wird. Die Formel folgt formal aus der Tatsache, dass D multilinear und alternierend ist.

Dieser Sachverhalt verallgemeinert sich auf höhere Dimensionen:

Satz 48 *Es sei V ein K-Vektorraum der Dimension n. Es gibt eine von 0 verschiedene alternierende Multilinearform in n Argumenten $\vartheta \in \mathrm{Alt}^n(V, K)$. Jede andere alternierende Multilinearform dieser Art hat die Form $c\vartheta$, wobei $c \in K$.*

Wir verschieben der Beweis auf den nächsten Abschnitt (Satz 71). Die Funktion ϑ nennen wir eine *Determinantenfunktion* von V.

Korollar 49 *Es seien $v_1, \ldots, v_n \in V$ Vektoren. Sie bilden genau dann eine Basis von V, wenn $\vartheta(v_1, \ldots, v_n) \neq 0$.*

Beweis: Wir haben gesehen, dass $\vartheta(v_1, \ldots, v_n) = 0$, wenn die Vektoren $v_1, \ldots, v_n$ linear abhängig sind (Lemma 46).

Wir nehmen an, dass die Vektoren $v_1, \ldots, v_n$ linear unabhängig sind. Wir führen die Annahme $\vartheta(v_1, \ldots, v_n) = 0$ zum Widerspruch. Wir dürfen nach Induktion voraussetzen, dass das Korollar für Vektorräume W gilt, so dass $\dim W < n = \dim V$. Wir betrachten auf $W = \mathcal{L}(v_2, \ldots, v_n)$ die alternierende Multilinearform

$$\tau(w_2, \ldots, w_n) = \vartheta(v_1, w_2, \ldots, w_n),$$

Wenn τ nicht identisch 0 wäre, so wäre τ eine Determinantenfunktion auf W. Dann hätten wir nach Induktionsvoraussetzung $0 \neq \tau(v_2, \ldots, v_n)$. Das widerspräche unser Annahme $\vartheta(v_1, \ldots, v_n) = 0$. Also ist τ identisch 0.

Unsere Annahme bleibt richtig, wenn wir die Basiselemente $v_1, \ldots, v_n$ umnumerieren. Deshalb finden wir das für alle i, $1 \leq i \leq n$ und für beliebige Vektoren $w_2, \ldots, w_n \in V$

$$\vartheta(v_i, w_2, \ldots, w_n) = 0.$$

Aber dann ist ϑ nach Satz 18 identisch 0. Das widerspricht der Voraussetzung, dass ϑ eine Determinantenfunktion ist. *Q.E.D.*

Wir berechnen ϑ in dem Fall $V = K^3$. Es sei $A = (a_{ij}) \in M(3 \times 3, K)$. Die drei Spalten von A seien $s_1, s_2, s_3 \in K^3$. Dann gilt:

$$s_j = \sum_{i=1}^{3} a_{ij} e_i.$$

Wir wählen und eine Determinantenfunktion ϑ. Weil ϑ im ersten Argument linear ist finden wir:

$$\vartheta(s_1, s_2, s_3) = a_{11}\vartheta(e_1, s_2, s_3) + a_{21}\vartheta(e_2, s_2, s_3) + a_{31}\vartheta(e_3, s_2, s_3).$$

Wir berechnen jeden der Summanden, in dem wir die Linearität im zweiten Argument ausnutzen

$$\begin{aligned} = \quad & a_{11}a_{12}\vartheta(e_1, e_1, s_3) + a_{11}a_{22}\vartheta(e_1, e_2, s_3) + a_{11}a_{32}\vartheta(e_1, e_3, s_3) \\ & + a_{21}a_{12}\vartheta(e_2, e_1, s_3) + a_{21}a_{22}\vartheta(e_2, e_2, s_3) + a_{21}a_{32}\vartheta(e_2, e_3, s_3) \\ & + a_{31}a_{12}\vartheta(e_3, e_1, s_3) + a_{31}a_{22}\vartheta(e_3, e_2, s_3) + a_{31}a_{32}\vartheta(e_3, e_3, s_3). \end{aligned} \tag{56}$$

Da ϑ alternierend ist, ist der 1., 5. und 9. Summand auf der rechten Seite 0. Wir berechnen den 2. Summanden weiter:

$$\begin{aligned} & a_{11}a_{22}\vartheta(e_1, e_2, s_3) = \\ & \quad a_{11}a_{22}a_{13}\vartheta(e_1, e_2, e_1) + a_{11}a_{22}a_{23}\vartheta(e_1, e_2, e_2) + a_{11}a_{22}a_{33}\vartheta(e_1, e_2, e_3) \\ & = a_{11}a_{22}a_{33}\vartheta(e_1, e_2, e_3) \end{aligned}$$

Nur der letzte Summand spielt eine Rolle, die anderen werden null, da die Argumente e_1 und e_2 schon vorkommen. Wenn wir auf die gleiche Weise mit den übrigen Summand in (56) verfahren, so ergibt sich:

$$\begin{aligned} = \quad & a_{11}a_{22}a_{33}\vartheta(e_1, e_2, e_3) + a_{11}a_{32}a_{23}\vartheta(e_1, e_3, e_2) + a_{21}a_{12}a_{33}\vartheta(e_2, e_1, e_3) \\ & + a_{21}a_{32}a_{13}\vartheta(e_2, e_3, e_1) + a_{31}a_{12}a_{23}\vartheta(e_3, e_1, e_2) + a_{31}a_{22}a_{13}\vartheta(e_3, e_2, e_1). \end{aligned}$$

Aus der Gleichung (55) ergibt sich schließlich

$$\begin{aligned} = \quad & a_{11}a_{22}a_{33}\vartheta(e_1, e_2, e_3) - a_{11}a_{32}a_{23}\vartheta(e_1, e_2, e_3) - a_{21}a_{12}a_{33}\vartheta(e_1, e_2, e_3) \\ & + a_{21}a_{32}a_{13}\vartheta(e_1, e_2, e_3) + a_{31}a_{12}a_{23}\vartheta(e_1, e_2, e_3) - a_{31}a_{22}a_{13}\vartheta(e_1, e_2, e_3). \end{aligned}$$

Da die Funktion ϑ nur bis auf Multiplikation bis auf eine von 0 verschiedene Konstante festliegt, können wir ϑ so wählen, dass

$$\vartheta(e_1, e_2, e_3) = 1.$$

Für eine Matrix $A \in M(3 \times 3, K)$ mit den Spalten s_1, s_2, s_3 definieren wir

$$\begin{aligned}&\det A = \vartheta(s_1, s_2, s_3)\\ &= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13}.\end{aligned}$$

Es ist leicht nachzuweisen, dass die so definierte Funktion ϑ multilinear und alternierend ist. Wir werden das später in einem dafür besser geeigneten Kalkül tun und dadurch den Satz 48 beweisen.

Für $V = K^n$ wollen wir eine Determinantenfunktion ϑ so wählen, dass $\vartheta(e_1, e_2, \ldots, e_n) = 1$. Wenn $A \in M(n \times n, K)$ eine Matrix mit den Spalten $s_1, \ldots, s_n$ ist, so definieren wir die Determinante von A:

$$\det A = \vartheta(s_1, \ldots, s_n). \tag{57}$$

Nach dem Korollar 49 ist $\det A = 0$, genau dann wenn Rang $A < n$.

Man kann Determinanten mit Hilfe von Scherungen ausrechnen. Es sei $(v_1, \ldots, v_n)$ eine Folge von Vektoren in einem Vektorraum V. Eine Scherung auf diese Folge anwenden, heißt den Vektor an der Stelle i durch $v_i + \lambda v_j$ ersetzen, wobei $\lambda \in K$ und $j \neq i$ beliebig gewählt werden können. Man erhält dann die Folge:

$$(v_1, \ldots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \ldots, v_n).$$

für ein festes $\lambda \in K$, $\lambda \neq 0$ hat man die Sequenz von Scherungen

$$(v_1, v_2) \mapsto (v_1 + \lambda v_2, v_2) \mapsto (v_1 + \lambda v_2, -\lambda^{-1} v_1) \mapsto (\lambda v_2, -\lambda^{-1} v_1).$$

Insbesondere kann man folgende Operationen durch eine Folge von Scherungen erhalten:

$$(v_1, v_2) \mapsto (-v_2, v_1), \quad (v_1, v_2) \mapsto (\lambda v_1, \lambda^{-1} v_2)$$

Es seien $v_1, \ldots, v_n \in K^n$ linear unabhängige Vektoren.

Eine Spaltenscherung auf eine Matrix A anwenden, heisst zu einer Spalte von A das Vielfache einer anderen Spalte zu addieren. Wenn A' aus A durch eine Spaltenscherung hervorgeht, so geht die tranponierte Matrix ${}^tA'$ aus ${}^tA$ durch eine Zeilenscherung hervor, wie wir sie oben benutzt haben. Deshalb gelten für Spaltenscherungen (und Spaltenoperationen) ähnliche Aussagen wie wir sie für Zeilenscherungen (und Zeilenoperationen) kennen. Man kann zum Beispiel jede Matrix durch Spaltenscherungen auf Stufenform bringen und damit insbesondere auf obere (oder untere) Dreiecksform.

Die Determinante von A ändert sich nach Lemma 46 nicht, wenn man eine Spaltenscherung auf A anwendet.

Eine obere Dreiecksmatrix

$$P = \begin{pmatrix} \lambda_1 & * & * & \dots & * \\ 0 & \lambda_2 & * & \dots & * \\ 0 & 0 & \lambda_3 & \dots & * \\ & & \dots & & \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

mit von 0 verschiedenen Elementen auf der Hauptdiagonale kann man durch Scherungen in die Diagonalmatrix

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ & & \dots & & \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

überführen. Die Diagonalelemente bleiben unverändert. Also gilt:

$$\det P = \det D = (\prod_{i=1}^{n} \lambda_i) \det E_n = \prod_{i=1}^{n} \lambda_i. \tag{58}$$

Diese Gleichung bleibt nach Satz 48 richtig, wenn eins der $\lambda_i = 0$ ist, da dann Rang $P \leq n-1$. Die gleiche Formel gilt mit der gleichen Begründung für eine untere Dreiecksmatrix. Da man eine beliebige Matrix durch Scherungen auf Dreiecksform bringen kann, haben wir eine Methode zur Berechnung einer Determinante.

Es sei $f : V \to V$ ein Endomorphismus. Dann gibt es ein Element $\det f \in K$, so dass:

$$\vartheta(f(v_1), \dots, f(v_n)) = (\det f)\vartheta(v_1, \dots, v_n), \tag{59}$$

für beliebige $v_1, \dots, v_n \in V$. Das folgt aus Satz 48, weil die linke Seite der obigen Gleichung multilinear und alternierend in der Argumenten $v_1, \dots, v_n$ ist. Man nennt $\det f$, die *Determinante* von f. Die Definition von $\det f$ ist unabhängig von der Wahl von ϑ.

Satz 50 *Der Endomorphismus $f : V \to V$ ist genau dann ein Isomorphismus, wenn* $\det f \neq 0$.

Das folgt aus dem Korollar zu Satz 48.

Bemerkung: Es sei $A \in M(n\times n, K)$. Wir haben $\det A$ definiert (57). Es sei $f : K^n \to K^n$ die lineare Abbildung $f(\underline{x}) = A\underline{x}$. Dann gilt $\det A = \det f$. Die neue und die alte Definition sind also kompatibel.

Satz 51 . *Es sei*

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle\alpha}\uparrow & & \uparrow{\scriptstyle\alpha} \\ W & \xrightarrow[g]{} & W \end{array}$$

ein kommutatives Diagramm endlich erzeugter Vektorräume. Es sei α ein Isomorphismus.

Dann gilt:

$$\det(f) = \det(g).$$

Beweis: Es sei $n = \dim V = \dim W$. Man wähle eine Determinantenfunktion ϑ auf V. Man definiere $\hat{\vartheta} \in \mathrm{Alt}^n(W, K)$ wie folgt:

$$\hat{\vartheta}(w_1, \ldots, w_n) = \vartheta(\alpha(w_1), \ldots, \alpha(w_n)), \quad w_i \in W.$$

Das ist eine Determinantenfunktion auf W.

Man findet:

$$\begin{array}{rcl} \hat{\vartheta}(g(w_1), \ldots, g(w_n)) & = & \vartheta(\alpha \circ g(w_1), \ldots, \alpha \circ g(w_n)) = \\ \vartheta(f \circ \alpha(w_1), \ldots, f \circ \alpha(w_n)) & = & \det(f)\vartheta(\alpha(w_1), \ldots, \alpha(w_n)) = \\ \det(f)\hat{\vartheta}(w_1, \ldots, w_n). & & \end{array}$$

Die vorletzte Gleichung ist die Definition von $\det(f)$. Andererseits besagt die Definition von $\det(g)$:

$$\hat{\vartheta}(g(w_1), \ldots, g(w_n)) = \det(g)\hat{\vartheta}(w_1, \ldots, w_n).$$

Wir sehen, dass $\det(f) = \det(g)$. *Q.E.D.*

Korollar 52 *Es sei $f : V \to V$ ein Endomorphismus. Dann können wir ein Koordinatensystem wählen $\alpha : K^n \to V$. Es sei A die Matrix von f in diesem Koordinatensystem. Dann gilt*

$$\det f = \det A$$

Das ist klar.

Übung: Oft kann man eine Determinante berechnen, indem man das Koordinatensystem geeignet wählt: Es sei $C \in M(n \times n, K)$ eine Matrix vom Rang 1. Man beweise, dass

$$\det(E_n + C) = 1 + \operatorname{Spur} C.$$

Satz 53 *(Multiplikationssatz) Es seien f und g Endomorphismen eines endlich erzeugten Vektorraumes V.*

Dann gilt:

$$\det(f \circ g) = (\det f)(\det g).$$

Beweis: Es seien $v_1, \ldots, v_n \in V$ beliebige Vektoren. Dann finden wir nach (59).

$$\vartheta(f(g(v_1)), \ldots, f(g(v_n))) = (\det f)\vartheta(g(v_1), \ldots, g(v_n)) = (\det f)(\det g)\vartheta(v_1, \ldots, v_n).$$

Nach (59) für $f \circ g$ folgt aus dieser Gleichung die Behauptung. *Q.E.D.*

Bemerkung: Es sei $V = \mathbb{R}^n$. Es seien $x, v_1, \ldots, v_n \in \mathbb{R}^n$. Unter einem Parallelkörper $\mathcal{P}$ versteht man die Menge

$$\mathcal{P} = \{x + \lambda_1 v_1 + \ldots + \lambda_n v_n \mid \lambda_i \in \mathbb{R},\ 0 \leq \lambda \leq 1,\ \text{wo } i = 1, \ldots n\}.$$

Man definiert das Volumen eines Parallelkörpers als den Absolutbetrag:

$$\operatorname{Vol}(\mathcal{P}) = |\vartheta(v_1, \ldots, v_n)|.$$

Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ ein Endomorphismus. Offensichtlich ist $f(\mathcal{P})$ der Parallelkörper zu den Vektoren $f(x), f(v_1), \ldots, f(v_n)$. Daher gilt

$$\operatorname{Vol}(f(\mathcal{P})) = |\det f| \operatorname{Vol}(\mathcal{P}).$$

In der Analysis hat man den Begriff einer messbaren Menge $A \subset \mathbb{R}^n$ zur Verfügung. Dann gilt allgemeiner:

$$\operatorname{Vol}(f(A)) = |\det f| \operatorname{Vol}(A).$$

Korollar 54 *Es sei $A \in M(n \times n, K)$. Dann gilt* $\det A = \det {}^tA$.

Beweis: Es sei $E_{ij}(\lambda)$ eine Elementarmatrix. Sie geht durch eine Spaltenscherung aus der Einheitsmatrix hervor und daher gilt $\det E_{ij}(\lambda) = 1$. Da eine Zeilenscherung A' von A durch die Multiplikation mit einer elementaren Matrix von links entsteht, erhalten wir aus Satz 53, dass $\det A' = \det A$. Da sich die Matrizen ${}^tA$ und ${}^tA'$ um eine Spaltenscherung unterscheiden gilt auch $\det {}^tA = \det {}^tA'$. Nach fortgesetzter Anwendung von Zeilenscherungen, kann man daher annehmen, dass A eine obere Dreiecksmatrix ist. Aber dann erhalten wir die Behauptung aus der Formel (58). *Q.E.D.*

Aus diesem Korollar folgt, dass $\det A$ auch eine multilineare alternierende Funktion in den Zeilen der Matrix A ist.

Satz 55 *Es sei $f : V \to V$ ein Endomorphismus eines endlich erzeugten Vektorraums. Es sei $U \subset V$ eine f-invarianter Unterraum (d.h. $f(U) \subset U$). Wir bezeichnen die Einschränkung von f auf U mit $g : U \to U$. Es sei $\pi : V \to W$ eine Surjektion mit dem Kern U.*

Dann gibt es einen Endomorphismus $h : W \to W$, so dass folgendes Diagramm kommutativ ist

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U & \longrightarrow & V & \xrightarrow{\pi} & W & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & U & \longrightarrow & V & \xrightarrow{\pi} & W & \longrightarrow & 0. \end{array}$$

Es gilt:

$$\det f = (\det g)(\det h).$$

Beweis: Wir wählen eine Basis $b_1, \dots b_m$ von U und ergänzen sie zu einer Basis $b_1, \dots, b_n$ von V. Dann ist $\pi(b_{m+1}), \dots, \pi(b_n)$ eine Basis von W.

In der Basis $b_1, \dots, b_n$ hat die Matrix von f die folgenden Form:

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

wobei $A \in M(m \times m, K)$ und $B \in M(n-m \times n-m, K)$. Die Matrix A ist die Matrix von g in der Basis $b_1, \dots, b_m$ und die Matrix B ist die Matrix von h in der Basis $\pi(b_{m+1}), \dots, \pi(b_n)$. Die Aussage des Satzes ist daher äquivalent mit der Gleichung

$$\det P = \det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = (\det A)(\det B). \tag{60}$$

Zum Beweis betrachten wir zunächst den Fall, wo A und B obere Dreiecksmatrizen sind. Dann erhält man (60) direkt aus (58).

Den allgemeinen Fall kann man durch Spaltenscherungen darauf zurückführen. Zunächst wendet man auf die Blockmatrix P nur Spaltenscherungen an, an denen nur die ersten m Spalten beteiligt sind. Man erhält dann eine Blockmatrix der Form

$$P' = \begin{pmatrix} A' & C \\ 0 & B \end{pmatrix},$$

Da die Determinante bei Scherungen ihren Wert nicht ändert, findet man $\det P' = \det P$ und $\det A' = \det A$. Also genügt es die Formel (60) für P' zu beweisen. Man kann erreichen, dass A' eine obere Dreiecksmatrix ist.

Dann wendet man auf P' Scherungen an, an denen nur die letzten $n - m$ Spalten beteiligt sind. Man erhält dadurch eine Blockmatrix der Form

$$P'' = \begin{pmatrix} A' & C' \\ 0 & B' \end{pmatrix}.$$

Wieder kann man erreichen, dass B' eine obere Dreiecksform hat. Also gilt (60) für P''. Da $\det P'' = \det P$, $\det A' = \det A$ und $\det B = \det B'$ folgt die Formel (60) auch für P. *Q.E.D.*

Übung: Es sei ϑ die Determinantenfunktion auf V, so dass $\vartheta(b_1, \ldots, b_n) = 1$. Es sei σ die Determinantenfunktion auf U, so dass $\sigma(b_1, \ldots, b_m) = 1$. Es sei τ die Determinantenfunktion auf W, so dass $\tau((\pi(b_{m+1}), \ldots, \pi(b_n)) = 1$. Dann gilt beliebige Vektoren $v_1, \ldots, v_n \in V$, so dass $v_1, \ldots, v_m \in U$ die folgende Gleichung

$$\vartheta(v_1, \ldots, v_n) = \sigma(v_1, \ldots, v_m)\tau(\pi(v_{m+1}), \ldots \pi(v_n)).$$

Definition 56 *Es sei (V, f) eine endlich erzeugter Vektorraum V, der mit einem Endomorphismus versehen ist. Es sei (W, h) von der gleichen Art. Wir sagen, dass (V, f) zu (W, h) ähnlich ist, wenn ein Isomorphismus $\pi : V \to W$ existiert, so dass folgendes Diagramm kommutativ ist.*

$$\begin{array}{ccc} V & \xrightarrow{\pi} & W \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle h} \\ V & \xrightarrow{\pi} & W \end{array}$$

Die Änlichkeit ist eine Äquivalenzrelation. Ähnliche Endomorphismen haben nach dem Korollar die gleiche Determinante.

Beispiel: Wenn wir quadratische Matrizen als Endomorphismen auffassen, können wir von ähnlich reden. Nach Definition 56 sind zwei Matrizen $A, B \in M(n \times n, K)$ ähnlich, wenn es ein kommutatives Diagramm gibt:

$$\begin{array}{ccc} K^n & \xrightarrow{C} & K^n \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle B} \\ K^n & \xrightarrow{C} & K^n, \end{array}$$

wobei die Matrix C einen Isomorphismus definiert. Also sind A und B genau dann ähnlich, wenn es eine invertierbare Matrix $C \in M(n \times n, K)$ gibt, so dass

$$BC = CA, \quad \text{oder } B = CAC^{-1}.$$

Nach dem Multiplikationsatz 53 folgt

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

Da $\det C \neq 0$, folgt $\det B = \det A$. Das wissen wir bereits nach Satz 51.

Es sei $A \in M(n \times n, K)$ eine Matrix mit den Einträgen a_{kl}. Wir fixieren Indexe i und j, wobei $1 \leq i \leq n$. Wir streichen aus der Matrix A, die i-te Zeile und die j-te Spalte und erhalten eine Matrix $A_{ij} \in M((n-1) \times (n-1), K)$.

Satz 57 *(Entwicklungssatz) Es sei ein Index j fixiert. Dann gilt die Entwicklungsformel nach der j-ten Spalte:*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}. \tag{61}$$

Es sei i fixiert. Dann gilt die Entwicklungsformel nach der i-ten Zeile:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}. \tag{62}$$

Beweis: Wir zeigen zuerst die erste Formel des Satzes für $j = 1$. Dazu betrachten wir zunächst den Fall, wo die erste Spalte der Matrix A der Standardvektor e_1 ist. Dann gilt nach (60), dass

$$\det A = \det A_{11}. \tag{63}$$

Das ist (61) für dieses spezielle A.

Als nächtes betrachten wir den Fall, wo die erste Spalte der Matrix A der Standardvektor e_i für ein $i > 1$ ist. Wir vertauschen die i-te Zeile von A mit der 1-sten Zeile und erhalten eine Matrix B. Da det auch alternierend in den Zeilen einer Matrix ist, finden wir $\det B = (-1)^{i-1} \det A$. Man sieht, dass $A_{i1} = B_{11}$. Also gilt unter Benutzung von (63)

$$\det A = (-1)^{i-1} \det B = (-1)^{i-1} \det B_{11} = (-1)^{i-1} \det A_{i1} = (-1)^{i+1} \det A_{i1}.$$

Damit haben wir (61) für $j = 1$ bewiesen, wenn die erste Spalte von A ein Standardvektor e_i ist.

Wir betrachten jetzt den Fall $j = 1$ mit beliebiger erster Spalte von A. Es seien $s_1, \ldots, s_n$ die Spalten der Matrix A. Dann gilt

$$\det A = \det(s_1, \ldots, s_n) = \sum_{i=1}^{n} a_{i1} \det(e_i, s_2, \ldots s_n).$$

Da wir $\det(e_i, s_2, \ldots s_n) = (-1)^{i+1} A_{i1}$ bereits bewiesen haben, erhalten wir die Entwicklungsformel (61) für $j = 1$.

Wenn j beliebig ist, so schieben wir die j-te Spalte in die erste Spalte, indem wir sie jeweils mit vorhergehenden Spalte vertauschen. Dadurch führen wir die Formel (61) auf den Fall $j = 1$ zurück.

Die Formel (62) folgt aus Korollar (54). *Q.E.D.*

Definition 58 *Es sei $A \in M(n \times n, K)$ eine Matrix. Es sei $A_{ij} \in M((n-1) \times (n-1), K)$ die Matrix, die aus A durch Streichung der i-ten Zeile und j-ten Spalte entsteht. Wir definieren*

$$c_{ji} = (-1)^{i+j} \det A_{ij} \quad \textit{für } i, j = 1, \ldots, n.$$

Die Adjunkte zu A ist die Matrix C mit dem Eintrag c_{ji} in der j-ten Zeile und i-ten Spalte.

Satz 59 *Es sei $A \in M(n \times n, K)$ eine Matrix und $C \in M(n \times n, K)$ ihre Adjunkte. Dann gilt*

$$CA = (\det A) E_n = AC.$$

Beweis: Wir zeigen die erste Gleichung. Sie besagt, dass für alle Indexe k, j gilt, dass

$$\sum_{i=1}^{n} (-1)^{i+k} (\det A_{ik}) a_{ij} = \delta_{kj} \det A.$$

Das folgt für $k = j$ aus dem Entwicklungssatz nach der k-ten Spalte. Wenn $k \neq j$, so ersetzt man in der Matrix A die k-te Spalte s_k durch die Spalte s_j. Dann entwickelt man die erhaltene Matrix nach der k-ten Spalte in der jetzt die Elemente a_{ij} stehen:

$$0 = \vartheta(s_1, \ldots, s_j, \ldots, s_j, \ldots s_n) = \sum_{i=1}^{n} a_{ij}(-1)^{i+k} \det A_{ik}.$$

Anlog gewinnt man die zweite Gleichung aus der Entwicklungsatz nach einer Zeile oder man überlegt, das ${}^tC$ die Adjunkte zu ${}^tA$ ist. *Q.E.D.*

Es sei K ein Körper. Die Menge der Abbildungen von Mengen $\mathcal{F} = \mathrm{Abb}(K, K)$ ist eine K-Vektorraum. Die Multiplikation von Funktionen ist eine bilineare Abbildung

$$\mathcal{F} \times \mathcal{F} \to \mathcal{F}.$$

Dadurch wird $\mathcal{F}$ ein Ring.

Definition 60 *Es sei K ein unendlicher Körper. Eine Funktion $f \in \mathcal{F}$ heißt Polynomfunktion, wenn eine Zahl $m \in \mathbb{Z}_{\geq 0}$ und Elemente $a_m, \ldots, a_1, a_0 \in K$ existieren, so dass*

$$f(\lambda) = a_m\lambda^m + \ldots + a_1\lambda + a_0, \quad \lambda \in K. \tag{64}$$

Die Menge $\mathcal{P}$ der Polynomfunktionen sind ein K-Untervektorraum $\mathcal{P} \subset \mathcal{F}$. Das ist ein Unterring, da das Produkt von Polynomfunktionen wieder eine Polynomfunktion ist.

Die Funktionen

$$\mathbf{X}^s(\lambda) := \lambda^s, \quad s = 0, 1, 2, 3 \ldots,$$

sind eine Erzeugendensystem des Vektorraums $\mathcal{P}$. Hier soll man s zunächst als ein Index und nicht als Potenz auffassen. Es gilt aber

$$\mathbf{X}^s \cdot \mathbf{X}^t = \mathbf{X}^{s+t}.$$

Wir setzen

$$\mathcal{P}_m = \mathcal{L}(\mathbf{X}^0, \mathbf{X}^1, \ldots, \mathbf{X}^m)$$

Die folgende Übung zeigt, dass die Funktionen $\mathbf{X}^s$ linear unabhängig sind.

Übung: Es seien $\xi_0, \ldots, \xi_m \in K$ verschiedene Elemente. Man betrachte die Polynomfunktionen:

$$\delta_s(\lambda) = \prod_{t \neq s}(\lambda - \xi_t) / \prod_{t \neq s}(\xi_s - \xi_t), \quad \text{für } s = 0, \ldots, m$$

Hier durchläuft t die Zahlen $0 \leq t \leq m$, so dass $t \neq s$. Es ist klar, dass $\delta_s \in \mathcal{P}_m$. Man berechne diese Funktionen an den Stellen $\xi_0, \ldots, \xi_m$. Man folgere, dass die folgende lineare Abbildung

$$\begin{array}{ccc} \mathcal{P}_m & \to & K^{m+1}, \\ f & \mapsto & {}^t(f(\xi_0), f(\xi_1), \ldots, f(\xi_m)) \end{array} \tag{65}$$

eine Surjektion von Vektorräumen ist. Das ist dann ein Isomorphismus, weil $\dim_K \mathcal{P}_m \leq m + 1$. Daher ist $\mathbf{X}^0, \mathbf{X}^1, \ldots, \mathbf{X}^m$ eine Basis des Vektorraumes $\mathcal{P}_m$. Also kann ein von Null verschiedenes Polynom von Grad $\leq m$ nicht $m + 1$ verschiedene Nullstellen besitzen.

Es lohnt sich das explizit zu formulieren:

Satz 61 *Es seien f und g zwei Polynomfunktionen, die folgende Darstellungen besitzen:*

$$\begin{array}{ll} f(\lambda) & = a_m\lambda^m + \ldots + a_1\lambda + a_0, \\ g(\lambda) & = b_m\lambda^m + \ldots + b_1\lambda + b_0, \end{array}$$

wobei $m \in \mathbb{N}$ und $a_i, b_i \in K$ für $i = 1, \ldots, m$.

Es sei $f(\lambda) = g(\lambda)$ für $m + 1$ verschiedene Elemente $\lambda \in K$.

Dann gilt $a_t = b_t$ für $0 \leq t \leq m$.

Korollar 62 *Es seien $f, g \in \mathcal{P}$ zwei Polynomfunktionen. Wenn $f \cdot g = 0$, so gilt $f = 0$ oder $g = 0$.*

Beweis: Es sei $f \neq 0$ und $g \neq 0$. Dann gibt es nur endlich viele $\lambda \in K$, so dass $f(\lambda) = 0$ und nur endlich viele $\mu \in K$, so dass $g(\mu) = 0$. Da K unendlich ist, findet man ein $\tau \in K$ mit $f(\tau) \neq 0$ und $g(\tau) \neq 0$. Dann gilt $(f \cdot g)(\tau) := f(\tau)g(\tau) \neq 0$. *Q.E.D.*

Im Anhang haben wir für jeden Ring R den Ring der Polynome $R[X]$ definiert. Es sei K ein unendlicher Körper. Dann ist die Abbildung

$$\begin{array}{ccc} K[X] & \to & \mathcal{P} \\ a_mX^m + \ldots + a_1X + a_0 & \mapsto & a_m\mathbf{X}^m + \ldots + a_1\mathbf{X}^1 + a_0\mathbf{X}^0 \end{array}$$

ein Isomorphismus von Ringen. Deshalb brauchen wir über einem unendlichen Körper K kein Unterschied zwischen Polynomen und Polynomfunktionen zu machen. Insbesondere ist im weiteren $X = \mathbf{X}$. Grad

Wir können jedes von 0 verschiedene Polynom $f \in K[X]$ in der Form

$$f = a_m X^m + \ldots + a_1 X + a_0, \text{ wobei } a_m \neq 0$$

schreiben. Die Zahl $m \in \mathbb{Z}_{\geq 0}$ nennen wir den *Grad* des Polynoms und schreiben $m = \deg(f)$. Wir setzen $\deg(0) = -\infty$. Dann gelten für zwei Polynome $f, g \in K[X]$ die folgenden Regeln:

$$\begin{aligned} &\deg(fg) = \deg(f) + \deg(g) \\ &\deg(f+g) \leq max\{\deg(f), \deg(g)\}. \end{aligned} \tag{66}$$

Wenn die Grade von f und g verschieden sind, so ist die letzte Ungleichung eine Gleichheit. Die erste Gleichheit liefert einen anderen Beweis für Korollar 62.

Übung: Ein anderer Zugang zu dem Isomorphismus (65) ist die Vandermondsche Determinante.

$$V = \begin{pmatrix} 1 & \xi_0 & \xi_0^2 & \ldots & \xi_0^{m-1} & \xi_0^m \\ 1 & \xi_1 & \xi_1^2 & \ldots & \xi_1^{m-1} & \xi_1^m \\ 1 & \xi_2 & \xi_2^2 & \ldots & \xi_2^{m-1} & \xi_2^m \\ & & & \ldots & & \\ 1 & \xi_{m-1} & \xi_{m-1}^2 & \ldots & \xi_{m-1}^{m-1} & \xi_{m-1}^m \\ 1 & \xi_m & \xi_m^2 & \ldots & \xi_m^{m-1} & \xi_m^m \end{pmatrix}.$$

Die s-te Spalte dieser Matrix ist das Bild der Polynomfunktion $\mathbf{X}^s$ bei der Abbildung (65). Man beweise die Formel:

$$\det V = \prod_{i>j} (\xi_i - \xi_j), \quad i, j = 0, 1, \ldots, m.$$

Zum Beweis wendet man die folgenden Spaltenscherungen auf V an. Man multipliziert die vorletzte Spalte mit ξ_0 und subtrahiert sie von der letzten Spalte. Dann multipliziert die vorvorletzte Spalte mit ξ_0 und subtrahiert sie von der vorletzten Spalte. So fährt man fort. Am Schluß lautet die erste Zeile der Matrix $(1, 0, 0, \ldots, 0)$. Eine Induktion und der Entwicklungssatz

führen zum Ziel. Wenn die ξ_s alle verschieden sind, gilt $\det V \neq 0$. Das zeigt uns nochmal, dass (65) ein Isomorphismus ist.

Lemma 63 *Es seien $a_{ij}(\lambda)$, für $i, j = 1, \ldots, n$ Polynomfunktionen vom Grad ≤ 1. Dann ist*

$$f(\lambda) = \det(a_{ij}(\lambda))$$

eine Polynomfunktion vom Grad $\leq n$.

Übung: Man beweise das mit dem Entwicklungssatz.

Es sei V ein endlich erzeugter K-Vektorraum der Dimension n. Es sei $f : V \to V$ ein Endomorphismus. Man definiert das *charakteristische Polynom*:

$$\chi(f, \lambda) := \det(\lambda \operatorname{id}_V - f) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 \qquad (67)$$

Hier ist $\operatorname{id}_V : V \to V$ die identische Abbildung $\operatorname{id}_V(v) = v$. Die $a_i \in K$ sind Elemente, die nur von f abhängen. Man muss zeigen, dass die Determinante in (67) eine Polynomfunktion vom Grad $\leq n$ ist. Aber dazu kann man f durch einen ähnlichen Endomorphismus ersetzen. Man kann also annehmen, dass $f = A \in M(n \times n, K)$ eine Matrix ist. Dann folgt die Behauptung über den Grad aus der Übung. Wir zeigen, dass $a_n = 1$. Es sei $\lambda \neq 0$ und es sei $\mu\lambda = 1$. Wenn wir die Gleichung (67) mit μ^n multiplizieren finden wir

$$\det(\operatorname{id}_V - \mu f) = a_n + a_{n-1}\mu + \ldots + a_1\mu^{n-1} + a_0\mu^n.$$

Das sind zwei Polynomfunktionen, die für $\mu \neq 0$ übereinstimmen, und daher nach Korollar 61 auch für $\mu = 0$. Wenn man $\mu = 0$ einsetzt findet man $a_n = 1$. Für die Koeffizienten von (67) gilt:

$$a_0 = (-1)^n \det f, \quad a_{n-1} = -\operatorname{Spur} f, \quad a_n = 1.$$

Übung: Man beweise diese Formeln. Für die mittlere Formel kann man den Entwicklungssatz benutzen. Man beachte, dass man Determinanten, die Polynomfunktionen vom Grad $< n - 1$ sind, vernachlässigen kann.

Satz 64 *Es sei ein kommutatives Diagramm von endlich erzeugten Vektorräumen mit exakten Zeilen:*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & U & \xrightarrow{\varkappa} & V & \xrightarrow{\sigma} & W & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & U & \xrightarrow{\varkappa} & V & \xrightarrow{\sigma} & W & \longrightarrow & 0 \end{array}$$

Dann gilt für die charakteristischen Polynome:

$$\chi(f,\lambda) = \chi(g,\lambda)\chi(h,\lambda).$$

Das folgt aus Satz 55.

Satz 65 *Es sei V ein endlich erzeugter K-Vektorraum der Dimension n. Es sei $f : V \to V$ ein Endomorphismus.*

Die Nullstellen seines charakteristischen Polynoms $\chi(f,\lambda)$ sind die Eigenwerte von f.

Es sei μ ein Eigenwert von f. Dann kann man schreiben

$$\chi(f,\lambda) = (\lambda-\mu)^m g(\lambda), \quad m \in \mathbb{N},$$

wobei $g(\mu) \neq 0$. Die Zahl m ist die Dimension des Hauptraumes $V(\mu)$.

Beweis: Ein $\mu \in K$ ist genau dann ein Eigenwert, wenn der Kern des Endomorphismus $\mu\,\mathrm{id}_V - f : V \to V$ nicht null ist. Nach der Dimensionsformel ist das genau dann der Fall, wenn $\mu\,\mathrm{id}_V - f$ kein Isomorphismus ist, also wenn $\det(\mu\,\mathrm{id}_V - f) = 0$.

Wir betrachten die Fittingzerlegung von $\mu\,\mathrm{id}_V - f$

$$V = V(\mu) \oplus W.$$

Das ist eine Zerlegung in f-invariante Unterräume, so dass die Einschränkung von f auf W das Element μ nicht als Eigenwert hat. Wir finden nach Satz 64:

$$\chi(f,\lambda) = \chi(f_{|V(\mu)},\lambda)\chi(f_{|W},\lambda),$$

wobei $\chi(f_{|W},\mu) \neq 0$. Andererseits können wir eine Basis von $V(\mu)$ finden, in der die Matrix von $f_{|V(\mu)}$ obere Dreiecksform hat (Satz 30, 3):

$$P = \begin{pmatrix} \mu & * & * & \dots & * \\ 0 & \mu & * & \dots & * \\ 0 & 0 & \mu & \dots & * \\ & & \dots & & \\ 0 & 0 & 0 & \dots & \mu \end{pmatrix},$$

Also gilt $\chi(f_{|V(\mu)},\lambda) = \det(\lambda E_m - P) = (\lambda-\mu)^m$, wobei $m = \dim V(\mu)$. *Q.E.D.*

Satz 66 Cayley-Hamilton *Es sei V ein Vektorraum der Dimension n. Es sei $f : V \to V$ ein Endomorphismus mit dem charakteristischen Polynom:*

$$\chi(f, \lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0.$$

Dann gilt im Vektorraum $\mathrm{Hom}_K(V, V)$ *die Relation :*

$$f^n + a_{n-1}f^{n-1} + \ldots + a_1 f + a_0 \,\mathrm{id}_V = 0$$

Beweis: Wir wählen eine Basis des des Vektorraumes V. Wir können dann $\mathrm{End}\, V$ mit dem Ring der Matrizen $M(n \times n, K)$ identifizieren. Es sei $A \in M(n \times n, K)$ die Matrix von f. Es gilt:

$$\det(\lambda E_n - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0, \quad \lambda \in K.$$

Es sei $C(\lambda)$ die Adjunkte der Matrix $(\lambda E_n - A)$. Dann gilt nach Satz 59:

$$(\lambda E_n - A)C(\lambda) = \chi(\lambda)E_n. \tag{68}$$

Nach Lemma 63 findet man Matrizen $B_0, B_1, \ldots, B_{n-1} \in M(n \times n, K)$, so dass

$$C(\lambda) = B_0 + B_1\lambda + \ldots + B_{n-1}\lambda^{n-1}.$$

Wenn man das in (68) einsetzt, erhält man:

$$(\lambda E_n - A)(B_0 + B_1\lambda + \ldots + B_{n-1}\lambda^{n-1}) = \chi(\lambda)E_n.$$

Wir multiplizieren die linke Seite aus und ordnen nach Potenzen von λ:

$$-AB_0 + (B_0 - AB_1)\lambda + \ldots + (B_{n-2} - AB_{n-1})\lambda^{n-1} + B_{n-1}\lambda^n = \chi(\lambda)E_n.$$

Durch Vergleich der Koeffizienten bei λ^i (Satz 61), ergibt sich das Gleichungssystem:

(0) $-AB_0 = a_0E_n$

(1) $B_0 - AB_1 = a_1E_n$

(2) $B_1 - AB_2 = a_2E_n$

...

(n-1) $B_{n-2} - AB_{n-1} = a_{n-1}E_n$ zyklisch

(n) $B_{n-1} = E_n$

Man multipliziert die Gleichung mit der Nummer (i) von links mit der Matrix A^i. Wenn man dann alle Gleichungen addiert findet man die gewünschte Gleichung

$$0 = a_0E_n + a_1A + a_2A^2 + \ldots a_{n-1}A^{n-1} + A^n.$$

Q.E.D.

Beispiel: Es sei V ein endlich erzeugter Vektorraum. Wir nennen $f : V \to V$ *zyklisch* , wenn ein Vektor $v \in V$ existiert, so dass V von den folgenden Vektoren erzeugt wird

$$v, f(v), \ldots, f^i(v), \ldots.$$

Dann gibt es eine größte natürliche Zahl n, so dass die Vektoren

$$v, f(v), \ldots, f^{n-1}(v). \tag{69}$$

linear unabhängig sind. Also findet man eine Relation

$$f^n(v) + b_{n-1}f^{n-1}(v) + \ldots + b_1f(v) + b_0v = 0, \tag{70}$$

Aus dieser Relation folgert man, dass

$$v, f(v), \ldots, f^{n-1}(v)$$

eine Basis des Vektorraums V ist. Andererseits folgt aus dem Satz von Cayley-Hamilton mit den dort gewählten Bezeichnungen

$$f^n(v) + a_{n-1}f^{n-1}(v) + \ldots + a_1f(v) + a_0v = 0.$$

Durch Koeffizientenvergleich findet man $a_i = b_i$.

Übung: Wir können das auch ohne Cayley-Hamilton sehen. Aus der Gleichung 70 erhalten wir die Matrix von f bezüglich der Basis 69:

$$\text{Matrix}(f) = \begin{pmatrix} 0 & 0 & \ldots & 0 & -b_0 \\ 1 & 0 & \ldots & 0 & -b_1 \\ 0 & 1 & \ldots & 0 & -b_2 \\ & & & \ldots & \\ 0 & 0 & \ldots & 1 & -b_{n-1} \end{pmatrix}. \tag{71}$$

Mit dem Entwicklungssatz nach der ersten Zeile berechnet man leicht

$$\det(\lambda E_n - \text{Matrix}(f)) = \lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_1\lambda + b_0.$$

Also sind die b_i aus 70 die Koeffizienten des charakteristischen Polynoms.

Da f ähnlich zu Matrix(f) ist, folgern wir, dass zwei zyklische Endomorphismen mit dem gleichen charakteristischen Polynom zueinander ähnlich sind.

Wenn umgekehrt ein Polynom

$$\lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_1\lambda + b_0$$

gegeben ist, so nennen wir die Matrix (71) seine Begleitmatrix. Die Begleitmatrix $B : K^n \to K^n$ definiert einen zyklischen Endomorphimsus.

Übung: Es sei $\lambda \in K$. Man beweise, dass die $n \times n$-Matrix

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 & 0 \\ & & & \ldots & & & \\ 0 & 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}.$$

einen zyklischen Endomorphismus das K^n definiert. Sie ist also ähnlich zu ihrer Begleitmatrix. Wie sieht die Begleitmatrix aus?

1.9 Multilineare Abbildungen, Existenz der Determinante

Es seien $V_1, \ldots, V_m, U$ Vektorräume über dem Körper K. Wir bezeichnen mit Mult($V_1 \times V_2 \times \ldots \times V_m, U$) die Menge aller multilinearen Abbildungen

$$\Phi : V_1 \times \ldots \times V_m \to U.$$

Das ist ein Vektorraum nämlich ein Untervektorraum des Vektorraumes aller Funktionen auf der Menge $V_1 \times \ldots \times V_m$ mit Werten in U.

Es seien $\phi_i \in \mathrm{Hom}(V_i, K)$. Dann definiert man das *Tensorprodukt* : Das ist die multilineare Abbildung Tensorprodukt

$$\phi_1 \otimes \ldots \otimes \phi_m : V_1 \times \ldots \times V_m \to K,$$

die durch die Formel:

$$\phi_1 \otimes \ldots \otimes \phi_m(v_1, \ldots, v_m) = \phi_1(v_1) \cdot \ldots \cdot \phi_m(v_m) \in K$$

definiert ist.

Beispiel: Es sei $U = K^\ell$. Es sei $\hat{e}_i : K^\ell \to K$, $1 \leq i \leq \ell$ die lineare Abbildung, die einem Spaltenvektor $\underline{x} \in K^\ell$ seinen i-ten x_i Eintrag zuordnet. Entsprechend definieren wir auf $V = K^m$ für $1 \leq j \leq m$ die lineare Abbildung $\hat{f}_j$ und auf $W = K^n$ für $1 \leq k \leq n$ die linearen Abbildung $\hat{g}_k$, für $1 \leq k \leq n$. Wir fixieren die Indexe i, j, k. Dann ist das Tensorprodukt die Funktion:

$$\hat{e}_i \otimes \hat{f}_j \otimes \hat{g}_k(\underline{x}, \underline{y}, \underline{z}) = x_i y_j z_k, \quad \underline{x} \in K^\ell,\ \underline{y} \in K^m,\ \underline{z} \in K^n$$

Es seien $a^{ijk} \in K$ (hier stehen obere Indexe, nicht Potenzen!) für $1 \leq i \leq \ell$, $1 \leq j \leq m$, $1 \leq k \leq n$ beliebige Körperelemente. Wir definieren

$$\mathbf{t} = \sum_{i,j,k} a^{ijk} \hat{e}_i \otimes \hat{f}_j \otimes \hat{g}_i \in \mathrm{Mult}(K^\ell \times K^m \times K^n, K).$$

Manchmal nennt man das Datum a^{ijk} einen Tensor. Man könnte auch von der (mehrdimensionalen) Matrix der multilinearen Abbildung $\mathbf{t}$ sprechen. Man schreibt oft etwas nachlässiger

$$\mathbf{t} = \sum_{i,j,k} a^{ijk} x_i y_j z_k.$$

Satz 67 *Es sei $B_i \subset V_i$ eine Basis des Vektorraumes V_i für $i = 1, \ldots, m$. Dann ist die kanonische Abbildung*

$$\mathrm{Mult}(V_1 \times \ldots \times V_m, U) \to \mathrm{Abb}(B_1 \times \ldots \times B_m, U), \tag{72}$$

die einer multilinearen Abbildung ihre Einschränkung auf die Menge $B_1 \times \ldots \times B_m$ zuordnet, ein Isomorphismus von Vektorräumen.

Beweis: Wir führen den Beweis durch Induktion nach m. Wir haben das für $m = 1$ bereits gezeigt (Satz 18). Es sei $\mu : B_1 \times \ldots \times B_m \to U$ eine Abbildung von Mengen. Wir werden zeigen, dass μ im Bild der Abbildung (72) liegt. Für ein Element $b \in B_m$ definieren wir eine Abbildung $\mu_b : B_1 \times \ldots \times B_{m-1} \to U$:

$$\mu_b(b_1, \ldots, b_{m-1}) = \mu(b_1, \ldots, b_{m-1}, b).$$

Dann gibt es nach Induktionvoraussetzung multilineare Abbildungen $\alpha_b : V_1 \times \ldots \times V_{m-1} \to U$, deren Einschränkung auf $B_1 \times \ldots \times B_{m-1}$ mit μ_b übereinstimmt.

Wir definieren eine multilineare Abbildung $\alpha : V_1 \times \ldots \times V_m \to U$ wie folgt. Ein Element $v_m \in V_m$ hat eine eindeutige Darstellung

$$v_m = \sum_{b \in B_m} \lambda_b b, \quad \lambda_b \in B_m.$$

Es sei $v_1 \in V_1, \ldots, v_m \in V_m$. Dann setzen wir

$$\alpha(v_1, \ldots, v_m) = \sum_{b \in B_m} \lambda_b \alpha_b(v_1, \ldots, v_{m-1}).$$

Man überzeugt sich davon, dass α eine multilineare Abbildung ist, deren Einschränkung auf $B_1 \times \ldots \times B_m$ mit μ übereinstimmt.

Das gleiche Argument zeigt auch, dass die Abbildung (72) injektiv ist. *Q.E.D.*

Korollar 68 *Die V_i seien endlich erzeugte Vektorräume. Es sei $B_i \subset V_i$ eine Basis für $i = 1, \ldots, m$. Es sei $\hat{B}_i$ die duale Basis, des dualen Vektorraums V_i^*.*

Dann ist die Menge aller Multilinearformen

$$\hat{b}_1 \otimes \ldots \otimes \hat{b}_m, \quad \hat{b}_i \in \hat{B}_i$$

eine Basis des Vektorraumes $\mathrm{Mult}(V_1 \times \ldots \times V_m, K)$.

Beweis: Es genügt zu beweisen, dass die Einschränkung der angegebenen Multilinearformen auf die Menge $B_1 \times \ldots \times B_m$ eine Basis des Vektorraumes $\mathrm{Abb}(B_1 \times \ldots \times B_m, K)$ ist. Das ist klar, weil die Funktion

$$\hat{b}_1 \otimes \ldots \otimes \hat{b}_m : B_1 \times \ldots \times B_m \to K,$$

in dem Element $(b_1, \ldots, b_m)$ den Wert 1 annimmt, aber in allen übrigen Elementen der Menge $B_1 \times \ldots \times B_m$ den Wert 0. *Q.E.D.*

Wir nehmen weiter an, dass die Vektorräume $V_1, \ldots, V_m$ endlich erzeugt sind. Man kann eine Umkehrabbildung zu der Einschränkungsabbildung (72) hinschreiben:

$$\begin{array}{ccc} \mathrm{Abb}(B_1 \times \ldots \times B_m, U) & \tilde{\to} & \mathrm{Mult}(V_1 \times \ldots \times V_m, U) \\ f & \mapsto & \sum_{b_1, \ldots b_m \in B_1 \times \ldots \times B_m} f(b_1, \ldots b_m) \hat{b}_1 \otimes \ldots \otimes \hat{b}_m. \end{array} \tag{73}$$

(Wenn die V_i nicht endlich erzeugt wären, wäre diese Formel interpretationsbedürftig.)

Wir kommen jetzt auf die alternierenden multilinearen Abbildungen (52) zurück.

Korollar 69 *Es sei V ein endlich erzeugter Vektorraum der Dimension n. Es sei U ein Verktorraum. Dann gilt für $m > n$*

$$\mathrm{Alt}^m(V, U) = \{0\}.$$

Es sei $\Phi \in \mathrm{Alt}^n(V, U)$. Wenn Vektoren $v_1, \ldots, v_n \in V$ linear abhängig sind, so gilt

$$\Phi(v_1, \ldots, v_n) = 0. \tag{74}$$

Wenn es linear unabhängige Vektoren $v_1, \ldots, v_n$ gibt, für die (74) gilt, so ist $\Phi = 0$.

Beweis: Es sei B eine Basis von V. Wir wollen zeigen, dass jedes $\Phi \in \mathrm{Alt}^m(V, U)$ null ist. Nach Satz 67 genügt es zu zeigen, dass die Einschränkung von Φ auf $B \times \ldots \times B$ (hier stehen m Faktoren) null ist. Das ist richtig, da eine Folge von m Elementen aus B notwendig zwei gleiche Elemente enthält.

Die zweite Ausssage folgt aus der ersten, wenn man Φ auf den Unterraum $\mathcal{L}(v_1, \ldots, v_n)$ einschränkt.

Die letzte Aussage sieht man, wenn man $B = \{v_1, \ldots, v_n\}$ nimmt. *Q.E.D.*

Wir zeigen jetzt, wie man mit Hilfe von Satz 67 multilineare alternierende Abbildungen

$$\Phi : V \times \ldots \times V \to K, \tag{75}$$

konstruieren kann.

Es sei $f : B \times \ldots \times B \to K$ die Einschränkung von Φ. Weil Φ alternierend ist, ergeben sich für f die folgenden beiden Eigenschaften.

Es seien $i_1, \ldots, i_m \in [1, d]$. Wenn zwei dieser Indexe gleich sind, so muss gelten

$$f(b_{i_1}, \ldots, b_{i_m}) = 0, \quad \text{falls } i_s = i_t \text{ für } s \neq t. \tag{76}$$

Wenn $i_1, \ldots, i_m$ beliebig sind, so muss für jede Permutation $\sigma \in \mathcal{S}_m$ gelten

$$f(b_{i_{\sigma(1)}}, \ldots, b_{i_{\sigma(m)}}) = (\operatorname{sgn}\sigma) f(b_{i_1}, \ldots, b_{i_m}). \tag{77}$$

Umgekehrt sei $f : B \times \ldots \times B \to K$ eine Funktion, die die Gleichungen (76) und (77) erfüllt. Dann gibt es nach Satz 67 eine multilineare Abbildung $\Phi \in \operatorname{Mult}(V \times \ldots \times V, K)$, deren Einschränkung auf $B \times \ldots \times B$ gleich f ist.

Lemma 70 *Die multilineare Abbildung Φ ist alternierend.*

Beweis: Es sei $\sigma \in \mathcal{S}_m$. Dann definieren wir ein multilineare Abbildung $\Phi^\sigma \in \operatorname{Mult}(V \times \ldots \times V, K)$:

$$\Phi^\sigma(v_1, \ldots, v_m) := (\operatorname{sgn}\sigma)\Phi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}).$$

Nach (77) stimmen die beiden Einschränkungen von Φ und Φ^σ auf $B \times \ldots \times B$ überein. Daher gilt $\Phi = \Phi^\sigma$ nach Satz 67. Also erfüllt Φ die Gleichung

$$\Phi(v_1, \ldots, v_m) := (\operatorname{sgn}\sigma)\Phi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}),$$

für beliebige Vektoren $v_1, \ldots . v_m \in V$. Wenn die Charakteristik von K nicht 2 ist, zeigt das bereits, dass Φ alternierend ist (Lemma 47). Sonst kann man die folgende Übungsaufgabe benutzen. *Q.E.D.*

Übung: Es sei $f \in \operatorname{Mult}(V^2, K)$, so dass $f(w_1, w_2) = -f(w_2, w_1)$ für alle $w_1, w_2 \in V$. Wenn eine Basis $b_1, \ldots b_d$ von V existiert, so dass $f(b_i, b_i) = 0$ für $i = 1, \ldots d$, so ist f alternierend.

Wenn $m > d$ so muss eine Funktion auf $B^m = B \times \ldots \times B$, die (76) erfüllt, null sein, da in B nur d Elemente zu Verfügung stehen. Daher folgt aus Satz 67, dass

$$\operatorname{Alt}^m(V, K) = 0 \quad \text{für } m > \dim V. \tag{78}$$

Das wissen wir bereits nach Lemma 46.

Satz 71 *Es sei V ein Vektorraum der Dimension d, Dann gilt*

$$\dim \mathrm{Alt}^d(V,K) = 1.$$

Wenn $b_1,\ldots,b_d$ eine Basis von V ist, so gibt es eine eindeutig bestimmte Multilinearform $\vartheta_B \in \mathrm{Alt}^d(V,K)$, so dass

$$\vartheta_B(b_1,\ldots,b_d) = 1.$$

Wir nennen ϑ_B die Determinantenfunktion bezüglich der Basis B.

Beweis: Wir konstruieren zuerst die Funktion ϑ_B. Dazu definieren eine Funktion

$$\hat{\vartheta} : B^d \to K.$$

auf der Menge $B^d = B \times B \times \ldots \times B$. Wir setzen $\hat{\vartheta}(b_{i_1},\ldots,b_{i_d}) = 0$, wenn zwei der Argumente b_{i_k} gleich sind. Wenn alle Argumente verschieden sind, so ist $\sigma(k) = i_k$ für $k = 1,\ldots,d$ eine Permutation $\sigma \in \mathcal{S}_d$. Dann setzen wir $\hat{\vartheta}(b_{i_1},\ldots,b_{i_d}) = \operatorname{sgn}\sigma$. Nach Satz 67 können wir $\hat{\vartheta}$ zu einer multilinearen Abbildung

$$\vartheta_B \in \mathrm{Mult}^d(V,K)$$

fortsetzen. Nach Lemma 70 ist $\vartheta_B \in \mathrm{Alt}^d(V,K)$.

Es sei $\Phi \in \mathrm{Alt}^d(V,K)$. Wir setzen $\lambda = \Phi(b_1,\ldots,b_d)$. Aus der Definition von alternierend (53) und aus (55) erhalten wir

$$\Phi(b_{i_1},\ldots,b_{i_d}) = \lambda\vartheta_B(b_{i_1},\ldots,b_{i_d}).$$

Dann folgt aus Satz 67, dass $\Phi = \lambda\vartheta_B$. *Q.E.D.*

Als Folgerung aus dem Beweis von Satz 71 erhält man:

Korollar 72 *Es sei V ein Vektorraum der Dimension d. Es sei $B = \{b_1,\ldots,b_d\}$ eine Basis von V. Es sei $\hat{b}_1,\ldots,\hat{b}_d \in V^*$ die duale Basis zu .*

Dann gilt die Leibnizformel:

$$\vartheta_B = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\hat{b}_{\sigma(1)} \otimes \ldots \otimes \hat{b}_{\sigma(d)}. \tag{79}$$

Beweis: In der Tat, die Multilinearformen auf beiden Seiten der Gleichung nehmen die gleichen Werte an für alle Argumente $(b_{i_1},\ldots,b_{i_d}) \in B^d$. *Q.E.D.*

Wir betrachten den Fall, wo $V = K^d$ und wo $\hat{b}_1 = \mathbf{e}_1,\ldots,\hat{b}_d = \mathbf{e}_d$ die Standardbasis ist. In diesem Fall schreiben wir $\det = \vartheta_B$.

Die duale Basis besteht aus den Koordinatenfunktionen $\hat{\mathbf{e}}_i$ für $i = 1, \ldots, d$:

$$\hat{\mathbf{e}}_i(\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_d \end{pmatrix}) = x_i, \quad \text{wo} \quad \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_d \end{pmatrix} \in K^d.$$

Es sei $A = (a_{ij}) \in M(d \times d, K)$. Es seien $s_1, s_2, \ldots, s_d$ die Spalten der Matrix A. Man definiert:

$$\det A = \det(s_1, \ldots, s_d).$$

Es gilt:

$$\hat{\mathbf{e}}_i(s_j) = a_{ij}.$$

Also sagt die Leibnizformel:

$$\begin{aligned} \det A &= \textstyle\sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \cdot \hat{\mathbf{e}}_{\sigma(1)}(s_1) \cdot \hat{\mathbf{e}}_{\sigma(2)}(s_2) \cdot \ldots \cdot \hat{\mathbf{e}}_{\sigma(d)}(s_d) \\ &= \textstyle\sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \ldots \cdot a_{\sigma(d)d} \end{aligned}$$

Proposition 73 *Es sei $A \in M(d \times d, K)$ eine Matrix und es sei ${}^tA$ die transponierte Matrix. Dann gilt*

$$\det A = \det {}^tA.$$

Beweis: Es sei $A = (a_{ij})$. Nach der Leibnizformel gilt

$$\det {}^tA = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{d\sigma(d)}.$$

Wenn i die Zahlen $1, 2, \ldots, d$ durchläuft, so durchläuft $\sigma^{-1}(i)$ ebenfalls die Zahlen $1, 2, \ldots, d$. Nur die Reihenfolge kann eine andere sein. Daher gilt

$$\prod_{i=1}^{d} a_{i\sigma(i)} = \prod_{i=1}^{d} a_{\sigma^{-1}(i),\sigma(\sigma^{-1}(i))} = \prod_{i=1}^{d} a_{\sigma^{-1}(i),i}.$$

Also kann man die Leibnizformel für ${}^tA$ auch schreiben

$$\det {}^tA == \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma^{-1}(1)1} \cdot a_{\sigma^{-1}(2)2} \cdot \ldots \cdot a_{\sigma^{-1}(d)d}.$$

Man schreibt $\tau = \sigma^{-1}$ und bemerkt, dass mit σ auch τ genau einmal die Elemente von $\mathcal{S}_n$ durchlaüft und dass $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau)$. Dann gilt

$$\det {}^tA == \sum_{\tau \in \mathcal{S}_n} \operatorname{sgn}(\tau) \cdot a_{\tau(1)1} \cdot a_{\tau(2)2} \cdot \ldots \cdot a_{\tau(d)d}.$$

Der rechte Ausdruck in der letzten Gleichung ist nach Definition $\det A$. *Q.E.D.*

Aus dem Satz erhalten wir, dass

$$\dim \operatorname{Alt}^d(V, K) = 1 \quad \text{für } d = \dim V.$$

Übung: Es sei $1 \leq m \leq d$

$$\dim \operatorname{Alt}^m(V, K) = \left(\begin{array}{c} d \\ m \end{array} \right).$$

Übung: Es sei V ein dreidimensionaler K-Vektorraum. Es sei $\Phi \in \operatorname{Alt}^2(V, K)$. Man beweise, dass es einen Vektor $x \in V$ gibt, so dass

$$\Phi(u, v) = \det(x, u, v).$$

Wir haben bereits zum Satz 48 Folgendes erklärt.

Korollar 74 *Es seien $w_1, \ldots, w_d \in V$ Vektoren. Sie sind genau dann linear abhängig, wenn*

$$\vartheta_B(w_1, \ldots, w_d) = 0.$$

Beweis: Es sei $C = \{w_1, \ldots, w_d\}$ eine Basis. Dann gilt $\vartheta_C(w_1, \ldots, w_d) = 1$. Da $\vartheta_C = \mu \vartheta_B$, für ein $\mu \in K$. Daraus folgt $\vartheta_B(w_1, \ldots, w_d) \neq 0$. Wenn $w_1, \ldots, w_d$ linear abhängig sind, muss die Einschränkung von ϑ_B auf den Unterraum $\mathcal{L}(w_1, \ldots, w_d) \subset V$ identisch 0 sein, da $\dim \mathcal{L}(w_1, \ldots, w_d) < d$ (nach (78)). *Q.E.D.*

1.10 Polynomfunktionen

Es sei K ein unendlicher Körper. Es sei V ein K-Vektorraum. Wir betrachten die (mengentheoretischen) Abbildungen $\operatorname{Abb}(V, K)$. Das ist ein K-Vektorraum. Es sei seien $\ell_1, \ldots, \ell_d \in \operatorname{Hom}(V, K) \subset \operatorname{Abb}(V, K)$ lineare Funktionen. Dann definieren eine Funktion $m \in \operatorname{Abb}(V, K)$:

$$m(v) := \ell_1(v) \cdot \ldots \cdot \ell_d(v) \in K. \tag{80}$$

Hier steht auf der rechten Seite ein Produkt in K. Wir schreiben für diese Funktion einfach $\ell_1 \cdot \ldots \cdot \ell_d$.

Der Unterraumvektorraum $\mathcal{P}(V, K) \subset \mathrm{Abb}(V, K)$, der von allen Funktionen $\ell_1 \cdot \ldots \cdot \ell_d$ für beliebiges d und beliebige $\ell_1, \ldots, \ell_d$ und von den konstanten Funktionen erzeugt wird, heißt der Raum der *Polynomfunktionen* auf V. Es ist klar, dass $\mathcal{P}(V, K)$ ein kommutativer Ring ist. Das Einselement ist die konstante Funktion mit dem Wert $1 \in K$.

Beispiel: Es sei $V = K$. Eine Funktion $P : K \to K$ ist eine Polynomfunktion, gdw. eine Zahl $n \in \mathbb{N}$ existiert und Elemente $a_0, a_1, \ldots, a_n \in K$, so dass für alle $\lambda \in K$

$$P(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0. \tag{81}$$

Das ist die Definition (60). Mit der dort verwendeten Bezeichnung gilt $\mathcal{P} = \mathcal{P}(K, K)$.

Wir geben jetzt eine typische Anwendung von Polynomfunktionen in einer Variablen. Es seien $f : V \to V$ und $g : V \to V$ zwei Endomorphismen eines Vektorraums V. Dann sind die charakteristischen Polynome der Endomorphismen $f \circ g$ und $g \circ f$ gleich.

In der Tat, wir nehmen zunächst an, dass f ein Isomorphismus ist. Dann gilt

$$\begin{aligned} \chi(f \circ g, \lambda) &:= \det(\lambda \,\mathrm{id}_V - f \circ g) = \det(f^{-1}(\lambda \,\mathrm{id}_V - f \circ g) \circ f) \\ &= \det(\lambda \,\mathrm{id}_V - g \circ f) = \chi(g \circ f, \lambda). \end{aligned}$$

Jetzt betrachten wir den allgemeinen Fall. Es sei $D \subset K$ die Teilmenge aller $\mu \in K$, die kein Eigenwert von f sind. Es sei $\lambda \in K$ fest. Dann gilt die Gleichung

$$\chi((f - \mu \,\mathrm{id}_V) \circ g, \lambda) = \chi(g \circ (f - \mu \,\mathrm{id}_V), \lambda)$$

für alle $\mu \in D$. Beide Seiten der Gleichung sind nach Lemma 63 Polynomfunktionen in $\mu \in K$. Daher stimmen beide Seiten nach dem Satz 61 für alle $\mu \in K$ überein. Also insbesondere auch für $\mu = 0$.

Es sei V ein K-Vektorraum. Wenn wir $d > 0$ festhalten, so nennen wir den Unteraum von $\mathcal{P}(V, K)$ der von allen $\ell_1 \cdot \ldots \cdot \ell_d$ für beliebige $\ell_1, \ldots, \ell_d$ erzeugt wird, der Raum der homogenen Polynomfunktionen $\mathcal{P}(V, K)(d)$ vom Grad

d. Die konstanten Funktionen auf V sind nach Definition die homogenen Polynomfunktionen vom Grad 0: $\mathcal{P}(V,K)(0) = K$.

Es sei $F \in \mathcal{P}(V,K)(d)$. Dann gilt für alle $\lambda \in K$:

$$F(\lambda v) = \lambda^d F(v), \quad v \in V. \tag{82}$$

Lemma 75 *$\mathcal{P}(V,K)$ ist die direkte Summe seiner Unterräume $\mathcal{P}(V,K)(d)$, d.h.*

$$\mathcal{P}(V,K) = \oplus_{d=0}^{\infty} \mathcal{P}(V,K)(d).$$

Insbesondere ist eine Polynomfunktion genau dann homogen vom Grad d. wenn die Gleichung (82) erfüllt ist.

Beweis: Wir müssen zeigen, dass sich jede Polynomfunktion F eindeutig als Summe von homogenen Polynomfunktionen schreiben lässt:

$$F = \sum_{d=0}^{m} F_d.$$

Hier ist m eine natürliche Zahl und $F_d \in \mathcal{P}(V,K)(d)$.

Die Existenz einer solchen Darstellung ist nach Definition klar. Es bleibt zu zeigen, dass alle Funktionen F_d null sind, wenn $F = 0$.

Wir fixieren einen beliebigen Vektor $v \in V$. Für jedes $\lambda \in K$ gilt $F(\lambda v) = 0$. Wir erhalten

$$0 = \sum_{d=0}^{m} \lambda^d F_d(v).$$

Die rechte Seite ist eine Polynomfunktion in einer Variablen λ. Daher sind alle Koeffizienten null: $F_d(v) = 0$. *Q.E.D.*

Satz 76 *Es sei V ein endlich erzeugter Vektorraum und es sei $v_1, \ldots, v_n$ eine Basis von V. Es sei $\breve{v}_1, \ldots, \breve{v}_n$ die duale Basis.*

Dann sind die Funktionen $\breve{v}_1^{t_1} \cdot \ldots \cdot \breve{v}_n^{t_n}$, wo $\underline{t} \in \mathbb{Z}_{\geq 0}^n$ eine Basis des Vektorraums $\mathcal{P}(V,K)$.

Beweis: Man betrachtet den Isomorphismus $K^n \to V$, der die Standardbasis e_i auf die Basis v_i abbildet. Dadurch kann man sich auf den Fall $V = K^n$ und $v_i = e_i$ beschränken. Es sei $\hat{e}_i$ die duale Basis zu Standardbasis. Dann gilt:

$$\hat{e}_i(\underline{x}) = x_i, \quad \underline{x} \in K^n.$$

Es ist üblich für die Funktion $\hat{e}_i : K^n \to K$ einfach x_i zu schreiben. Eine beliebige lineare Funktion $\ell : K^n \to K$ hat dann die Form

$$\ell(\underline{x}) = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n.$$

Wenn man ein Monom (80) ausmultipliziert erhält man dass man eine Polynomfunktion $P : K^n \to K$ in der folgenden Form schreiben kann

$$P(\underline{x}) = \sum_{\underline{t} \in \mathbb{Z}^n_{\geq 0}} a_{\underline{t}} x_1^{t_1} x_2^{t_2} \cdot \ldots \cdot x_n^{t_n}, \quad \text{wo } \underline{x} = (x_1, \ldots, x_n) \in K^n.$$

Hier bezeichnet $\mathbb{Z}^n_{\geq 0} \subset \mathbb{Q}^n$ die Menge aller Vektoren $\underline{t} = (t_1, \ldots, t_n)$ mit ganzzahligen nicht negativen Komponenten t_i. Die $a_{\underline{t}} \in K$ sind Elemente, so dass $a_{\underline{t}} = 0$ bis auf endlich viele Ausnahmen $\underline{t} \in \mathbb{Z}^n_{\geq 0}$ gilt. Der Satz behauptet, dass die Funktionen $x_1^{t_1} x_2^{t_2} \cdot \ldots \cdot x_n^{t_n} \in \mathcal{P}(K^n, K)$ linear unabhängig sind. Wir beweisen eine etwas schärfere Aussage:

Satz 77 *Es seien Elemente $a_{\underline{t}}, b_{\underline{t}} \in K$ für alle $\underline{t} \in \mathbb{Z}^n_{\geq 0}$ gegeben, die nur für endlich viele $\underline{t}$ verschieden von 0 sind. Es seien $P, Q \in \mathcal{P}(K^n, K)$ die folgenden Polynomfunktionen:*

$$P(\underline{x}) = \textstyle\sum_{\underline{t} \in \mathbb{Z}^n_{\geq 0}} a_{\underline{t}} x_1^{t_1} x_2^{t_2} \cdot \ldots \cdot x_n^{t_n},$$
$$Q(\underline{x}) = \textstyle\sum_{\underline{t} \in \mathbb{Z}^n_{\geq 0}} b_{\underline{t}} x_1^{t_1} x_2^{t_2} \cdot \ldots \cdot x_n^{t_n}$$

Angenommen es gibt eine unendliche Menge $D \subset K$, so dass für alle $\underline{x} \in D^n$

$$P(\underline{x}) = Q(\underline{x}).$$

Dann gilt $a_{\underline{t}} = b_{\underline{t}}$ für alle $\underline{t} \in \mathbb{Z}^n_{\geq 0}$.

Beweis: In dem man die Differenz $P - Q$ betrachtet, sieht man, dass man $b_{\underline{t}} = 0$ annehmen kann. Wir müssen $a_{\underline{t}} = 0$ zeigen.

Wir fixieren eine ganze Zahl $j > 0$ und setzen:

$$P_j(x_2, \ldots, x_n) = \sum_{t_2, \ldots, t_n} a_{j, t_2, \ldots, t_n} x_2^{t_2} \cdot \ldots \cdot x_n^{t_n},$$

Es gibt eine Zahl m, so dass $a_{j,t_2,\ldots,t_n} = 0$ für $j > m$. Dann gilt:

$$P(\underline{x}) = x_1^m P_m(x_2, \ldots, x_n) + x_1^{m-1} P_{m-1}(x_2, \ldots, x_n) + \ldots + P_0(x_2, \ldots, x_n). \tag{83}$$

Wir fixieren eine beliebigen Vektor $(x_2, \ldots, x_n) \in D^{n-1}$ und fassen die rechte Seite von (83) als Polynom in der Variablen x_1 auf. Aus Satz 61 erhalten wir, dass $P_j(x_2, \ldots, x_n) = 0$ für alle $(x_2, \ldots x_n) \in D^{n-1}$. Durch Induktion über n erhalten wir $a_{j,t_2,\ldots,t_n} = 0$. *Q.E.D.*

Ein wichtiges Beispiel einer Polynomfunktion ist die Determinante $\det : M(n \times n, K) \to K$, als Funktion auf dem Vektorraum $M(n \times n, K)$. Man kann mit der Leibnizformel (79) beweisen, dass det eine Polynomfunktion ist. In der Tat, es sei $X = (x_{ij}) \in M(n \times n, K)$. Dann besagt (79), dass

$$\det X = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) x_{1\sigma(1)} x_{2\sigma(2)} \cdot \ldots \cdot x_{n\sigma(n)}. \tag{84}$$

Hier steht offenbar eine Polynomfunktion. Alternativ kann man den Entwicklungssatz benutzen, um durch Induktion nach n zu zeigen, dass det eine Polynomfunktion ist. Da $\det(\lambda X) = \lambda^n \det X$, ist det homogen vom Grad n.

Abstrakter können wir formulieren: Es sei V ein Vektorraum der Dimension n. Dann ist

$$\det : \operatorname{End}(V) \to K \tag{85}$$

eine homogene Polynomfunktion von Grad n auf dem Vektorraum $\operatorname{End}(V)$.

Satz 78 *Es sei $K \subset L$ ein Teilkörper. Es sei K unendlich. Es seien $A, A' \in M(n \times n, K)$ Matrizen. Es sei $C \in M(n \times n, L)$ eine invertierbare Matrix, so dass $A = C^{-1}A'C$, d.h. die Matrizen A und A' sind über dem Körper L ähnlich.*

Dann gibt es eine invertierbare Matrix $B \in M(n \times n, K)$, so dass $A = B^{-1}A'B$.

Beweis: Wir betrachten das homogene lineare Gleichungssystem

$$XA - A'X = 0, \quad X \in M(n \times n, L). \tag{86}$$

Die Einträge von X sind die Unbekannten. Es sei $V \subset M(n \times n, L)$ der Untervektorraum der Lösungen. Da die Koeffizienten des Gleichungssystems in K liegen, gibt es Matrizen $B_1, \ldots, B_d \in M(n \times n, K)$, die eine Basis von $V \cap M(n \times n, K)$ sind. Das ist auch eine Basis des L-Vektorraumes V. Wir suchen ein $B \in V \cap M(n \times n, K)$, das invertierbar ist. Deshalb genügt es zu zeigen, dass die folgende Funktion auf K^d nicht identisch 0 ist.

$$F(t_1, \ldots, t_d) = \det(t_1 B_1 + \ldots + t_d B_d).$$

Wir werden weiter unten begründen, dass F eine Polynomfunktion ist.

Da C eine Lösung des Gleichungssystems (86) ist, können wir schreiben $C = \sum_{j=1}^{d} \lambda_i B_i$, wo $\lambda_i \in L$. Es gilt $F(\lambda_1, \ldots, \lambda_d) = \det C \neq 0$. Daher kann nicht das Polynom F nicht identisch 0 sein.

Der Satz ist bewiesen, wenn wir zeigen, dass F eine Polynomfunktion ist. Dazu schreiben wir F als Kompositum von zwei Abbildungen:

$$\begin{array}{ccccc} K^d & \to & M(n \times n, K) & \stackrel{\det}{\longrightarrow} & K. \\ (t_1, \ldots, t_d) & \mapsto & t_1 B_1 + \ldots + t_d B_d & & \end{array}$$

Die erste Abbildung ist linear und die zweite ist eine Polynomfunktion (85). Nach folgender Übung ist F eine Polynomfunktion. *Q.E.D.*

Übung: Es sei $f : W \to V$ eine lineare Abbildung von endlich erzeugten Vektorräumen. Es sei $P : V \to K$ eine Polynomfunktion. Dann ist $P \circ f : W \to K$ eine Polynomfunktion.

Satz 79 *Es seien $P, Q \in \mathcal{P}(V, K)$ Polynomfunktionen. Es sei $P \cdot Q = 0$. Dann gilt $P = 0$ oder $Q = 0$.*

Beweis: Wir nehmen an, dass $P \neq 0$. Wir können annehmen, dass $V = K^n$. Wir führen der Beweis durch Induktion nach n. Der Fall $n = 1$ ist das Korollar 62.

Wie im Beweis von Satz 77 schreiben wir:

$$\begin{aligned} P(\underline{x}) &= x_1^m P_m(x_2, \ldots, x_n) + x_1^{m-1} P_{m-1}(x_2, \ldots, x_n) + \ldots + P_0(x_2, \ldots, x_n), \\ Q(\underline{x}) &= x_1^m Q_m(x_2, \ldots, x_n) + x_1^{m-1} Q_{m-1}(x_2, \ldots, x_n) + \ldots + Q_0(x_2, \ldots, x_n). \end{aligned}$$

Mindestens eine der Polynomfunktionen P_i muss von 0 verschieden sein. Wir halten $(x_2, \ldots, x_n)$ fest. Wenn $P_i(x_2, \ldots, x_n) \neq 0$, so folgt aus dem Fall einer Variablen (Korollar 62), dass $Q_j(x_2, \ldots, x_n) = 0$ für $j = 0, 1, \ldots, m$.

Also gilt $P_i Q_j = 0$ für $j = 0, 1, \ldots, m$. Nach Induktionsvoraussetzung gilt daher $Q_j = 0$ und folglich $Q = 0$. *Q.E.D.*

Anwendung: Es seien m, n natürliche Zahlen. Es sei $A \in M(m \times n, K)$ und es sei $B \in M(n \times m, K)$. Dann gilt

$$\det(E_m + AB) = \det(E_n + BA). \tag{87}$$

Um das einzusehen nehmen wir o.B.d.A. an, dass $m < n$. Wir deuten A und B als lineare Abbildungen:

$$K^n \stackrel{A}{\to} K^m \stackrel{B}{\to} K^n.$$

Wir nehmen zunächt an, dass Rang $B = m$, d.h. die lineare Abbildung B ist injektiv. Zum Beweis der Formel können wir auf K^m und K^n jeweils eine beliebige Basistransformation durchführen. Deshalb können wir annehmen, dass $B(e_i) = e_i$ für $i = 1, \ldots, m$. Jetzt kann man die Behauptung sehr einfach nachprüfen.

Wir betrachten jetzt den Fall, wo B beliebig ist. Es sei B' die quadratische Matrix aus den ersten m Zeilen von B. Wir können $\det B'$ als eine Polynomfunktion auf dem Vektorraum $M(n \times m, K)$ auffassen. Die Gleichung (87) gilt für $\det B' \neq 0$. Also gilt

$$\det B'(\det(E_m + AB) - \det(E_n + BA)) = 0.$$

Da $\det(E_m + AB) - \det(E_n + BA)$ ebenfalls eine Polynomfunktion in $B \in M(n \times m, K)$ ist, folgt nach Satz 79, dass diese Polynomfunktion 0 ist. Das beweist (87).

Es sei K ein Körper der Charakteristik 0. Es sei V ein Vektorraum der Dimension n. Es sei $P : V \to K$ eine Polynomfunktion die homogen vom Grad d ist. Dann kann man schreiben

$$P = \sum_{i=1}^{M} a^{(i)} \ell_1^{(i)} \cdot \ldots \cdot \ell_d^{(i)}.$$

Man betrachte den Tensor (= Multilinearform)

$$T = \sum_{i=1}^{M} a^{(i)} \ell_1^{(i)} \otimes \ldots \otimes \ell_d^{(i)}.$$

Dann ist $T \in \mathrm{Mult}(V \times \ldots \times V, K)$ eine multilineare Abbildung in d Argumenten, so dass

$$P(v) = T(v, \ldots, v).$$

Man macht aus T eine symmetrischen Tensor S:

$$S(v_1, \ldots v_d) = (1/d!) \sum_{\sigma \in \mathcal{S}_d} T(v_{\sigma(1)}, \ldots, v_{\sigma(d)}).$$

Dann gilt die Gleichung

$$P(v) = S(v, \ldots, v) \tag{88}$$

Wir zeigen jetzt, dass S durch P eindeutig bestimmt ist. Das folgt aus dem folgenden Lemma. Wir nennen S die Polarisierung der homogenen Polynomfunktion P.

Lemma 80 *Es sei*

$$S : V \times \ldots \times V \to K$$

eine symmetrische Bilinearform in d Argumenten, so dass $S(v, v, \ldots, v) = 0$ *für alle* $v \in V$*. Dann gilt* $S = 0$*.*

Beweis: Es seien v, w beliebige Vektoren. Man betrachte die Polynomfunktion $f(\lambda) = S(v + \lambda w, \ldots, v + \lambda w)$. Sie ist identisch 0. Daher ist der Koeffizient von λ gleich 0. Also gilt $S(v, \ldots, v, w) = 0$ für alle $v, w \in V$. Nach Induktion folgt, dass $S(v_1, \ldots, v_{d-1}, w) = 0$ für beliebige $v_1, \ldots, v_{d-1} \in V$. *Q.E.D.*

Zusammenfassend haben wir bewiesen:

Satz 81 *Es sei* K *ein Körper der Charakteristik* 0*. Es sei* $Sym^d(V, K)$ *der Vektorraum aller symmetrischen Multilinearformen in d Argumenten. Es sei* $S \in \mathrm{Sym}^d(V, K)$*. Dann ist* $P(v) = S(v, \ldots, v)$ *eine homogene Polynomfunktion vom Grad d. Die Abbildung* $S \mapsto P$ *ist ein Isomorphismus von Vektorräumen:*

$$\mathrm{Sym}^d(V, K) \to \mathcal{P}(V, K)(d).$$

Man nennt S *die Polarisierung des homogenen Polynoms* P*.*

1.11 Bilineare Abbildungen und Dualität

Es sei V ein endlich erzeugter Vektorraum. Es sei $V^* = \mathrm{Hom}_K(V, K)$ der duale Vektorraum. Es sei $v_1, \ldots, v_n$ eine Basis von V. Nach Satz 18 gibt es eine Basis $\hat{v}_1, \ldots, \hat{v}_n \in V^*$, so dass

$$\hat{v}_i(v_j) = \delta_{ij}.$$

Wir haben das die duale Basis genannt. Insbesondere folgt $\dim V = \dim V^*$. Für jeden Vektor $v \in V$ gilt:

$$v = \sum_{i=1}^{n} \hat{v}_i(v) v_i. \tag{89}$$

In der Tat, man schreibt

$$v = \sum_{i=1}^{n} \lambda_i v_i, \quad \text{wo } \lambda_i \in K.$$

Wenn man auf diese Gleichung $\hat{v}_i$ anwendet, findet man $\hat{v}_i(v) = \lambda_i$. Wir haben $\hat{v}_i$ deshalb die Koordinatenfunktionen relativ zur Basis $v_1, \ldots, v_n$ genannt. Gramsche Matrix

Wir betrachten bilineare Abbildungen, das sind multilineare Abbildungen in zwei Argumenten:

$$\Phi : V \times W \to U.$$

Wir bezeichnen den K-Vektorraum der bilinearen Abbildungen mit $\mathrm{Bil}(V \times W, U)$. Wenn $U = K$, so nennen wir Φ eine Bilinearform. Wir wollen eine Bilinearform mit B bezeichnen.

$$B : V \times W \to K. \tag{90}$$

Es sei V endlich erzeugt mit der Basis $v_1, \ldots, v_n$ und es sei W endlich erzeugt mit der Basis $w_1, \ldots, w_m$, Es seien $\hat{v}_1, \ldots, \hat{v}_n \in V^*$ und $\hat{w}_1, \ldots, \hat{w}_m \in W^*$ die dualen Basen. Dann können wir B als Tensor schreiben (73)

$$B = \sum_{i=1}^{n} \sum_{j=1}^{m} B(v_i, w_j) \hat{v}_i \otimes \hat{w}_j.$$

Wir nennen $G := (B(v_i, w_j)) \in M(n \times m, K)$ die *Gramsche Matrix* der Bilinearform B. Nach Korollar 68 erhalten wir einen Isomorphismus von Vektorräumen:

$$\mathrm{Bil}(V \times W, K) \xrightarrow{\sim} M(n \times m, K),$$

der einer Bilinearform ihre Gramsche Matrix zuordnet.

Es sei $V = K^n$ und $W = K^m$. Es sei G die Gramsche Matrix von B bezüglich der Standardbasen. Dann gilt

$$B(\underline{x}, \underline{y}) = {}^t\underline{x} G \underline{y}.$$

Übung: Man beweise dies durch Anwendung von Satz 67. Die rechte Seite der Gleichung ist eine Bilinearform.

Jeder Bilinearform (90) entsprechen zwei lineare Abbilddungen:

$$\alpha : V \to W^* \tag{91}$$

$$\beta : W \to V^* \tag{92}$$

Die erste Abbildung (91) ordnet einem $v \in V$ folgendes Element $\ell_v \in W^*$ zu: $\ell_v(w) = B(v, w)$. Die zweite Abbildung (92) ordnet einem $w \in W$ folgendes Element $\ell_w \in V^*$ zu: $\ell_w(v) = B(v, w)$.

Standardskalarpro
Evaluation

Umgekehrt kann man aus der Abbildung (91) oder (92) die Bilinearform B zurückgewinnen:

$$B(v,w) = \alpha(v)(w), \quad B(v,w) = \beta(w)(v).$$

Allgemeiner gilt der folgende Sachverhalt: **Übung:** Es sei $\Phi : V \times W \to U$ eine bilineare Abbildungen von Vektorräumen. Für festes $v \in V$ definiert Φ eine lineare Abbildung $\Phi_v : W \to U$, wo $\Phi_v(w) = \Phi(v,w)$. Es gibt einen Iomorphismus von Vektorräumen (Formel von Cartan:)

$$\begin{array}{rccc} \mathbf{T}: & \mathrm{Bil}(V \times W, U) & \cong & \mathrm{Hom}(V, \mathrm{Hom}(W,U)). \\ & \Phi & \mapsto & (v \mapsto \Phi_v) \end{array}$$

Definition 82 *Man nennt die Bilinearform B nichtausgeartet in V, wenn ein beliebiger Vektor $v \in V$, der folgende Eigenschaft erfüllt, gleich 0 ist.*

Für jeden Vektor $w \in W$ gilt

$$B(v,w) = 0.$$

Entsprechend definiert man, dass B nichtausgeartet in W ist. Wenn wir einfach sagen, dass B nichtausgeartet ist, so meinen wir, dass B in V und W nicht ausgeartet ist.

B ist nichtausgeartet in V, wenn die Abbildung (91) injektiv ist und B nichtausgeartet in W, wenn (92) injektiv ist.

Beispiel: Es sei $V = W = K^n$. Für $\underline{x}, \underline{y} \in K^n$ definieren wir das *Standardskalarprodukt*

$$B(\underline{x}, \underline{y}) = {}^t\underline{x}\underline{y} = \sum_{i=1}^{n} x_i y_i.$$

Man sieht sofort, dass B in beiden Argumenten nicht ausgeartet ist.

Das wichtigste Beispiel einer nicht ausgearteten Bilinearform ist die *Evaluation*: Es sei V Vektorraum

$$\begin{array}{rccc} \mathcal{E}: & V \times V^* & \to & K \\ & (v, \ell) & \mapsto & \ell(v) \end{array} \tag{93}$$

Wenn für ein $\ell \in V^*$ gilt, dass $\ell(v) = 0$ für alle v, so folgt $\ell = 0$. Also ist $\mathcal{E}$ im zweiten Argument nicht ausgeartet.

Um zu zeigen, dass $\mathcal{E}$ im ersten Argument nicht ausgeartet ist, muss man zeigen: Es sei $v \in V$, $v \neq 0$. Dann existiert ein $\ell \in V^*$, so dass $\ell(v) \neq 0$. In der Tat, es sei U ein Komplementärraum zu Kv, also $V = Kv \oplus U$. Dann liefert die Projektion $V \to Kv$ längs U das gewünschte ℓ.

Die der Evaluation entsprechende Abbildung (92) ist die Identität: $\mathrm{id}_{V^*} : V^* \to V^*$.

Mit Hilfe der Evaluation können wir die Definition so schreiben

$$\mathcal{E}_V(v, \beta(w)) = B(v, w) = \mathcal{E}_W(w, \alpha(v)). \tag{94}$$

Satz 83 *Es seien V und W Vektorräume, von denen wenigstens einer endlich erzeugt ist. Es sei $B : V \times W \to K$ nichtausgeartet in V und in W.*

Dann gilt $\dim V = \dim W$. Wenn $\ell_2 : W \to K$ eine lineare Abbildung ist, so gibt es genau einen Vektor $v \in V$ mit der Eigenschaft:

$$\ell_2(w) = B(v, w), \text{ für alle } w \in W.$$

Wenn $\ell_1 : V \to K$ eine lineare Abbildung ist, so gibt es genau einen Vektor $w \in W$ mit der Eigenschaft:

$$\ell_1(v) = B(v, w), \text{ für alle } v \in V.$$

Beweis: Man kann sich auf den Fall beschränken, wo V endlich erzeugt ist. Da $\beta : W \to V^*$ nach Voraussetzung injektiv ist, ist dann auch W endlich erzeugt und es gilt

$$\dim W \leq \dim V^* = \dim V.$$

Genauso sieht man mit Hilfe der Abbildung (91), dass

$$\dim V \leq \dim W.$$

Also gilt $\dim V = \dim W$. Daraus folgt, dass die beiden injektiven linearen Abbildungen (91) und (92) Isomorphismen sein müssen. Der Satz behauptet, dass diese Abbildungen surjektiv sind. *Q.E.D.*

Bemerkung: Es seien $B : V \times W \to K$ und $B' : V' \times W' \to K$ zwei Bilinearformen. Ein Isomorphismus von B nach B' ist nach Definition

ein Paar (μ, ν) von Isomorphismen von Vektorräumen: $\mu : V \to V'$ und $\nu : W \to W'$, so dass

Nullraum

$$B(v, w) = B'(\mu(v), \nu(w)).$$

(μ, ν) ist ein Isomorphismus in der Kategorie der Bilinearformen. Heuristisch sind isomorphe Bilinearformen in allen relevanten Eigenschaften gleich.

Es sei $B : V \times W \to K$ eine nicht ausgeartete Bilinearform endlich erzeugter Vektorräume, d.h. wie im Satz 83. Dann hat man einen Isomorphismus von Bilinearformen $B \to \mathcal{E}_V$. Mit den Bezeichnungen (94) ist der Isomorphismus durch das Paar (id_V, β) gegeben. Daher kann man sich auf die Evaluation beschränken, wenn man etwas über B beweisen will.

Es sei $B : V \times W \to K$ eine Bilinearform. Wir definieren

$$V_0 = \{v \in V \mid B(v, w) = 0 \text{ für alle } w \in W\}.$$

$V_0 \subset V$ ist ein Untervektorraum und wir nennen ihn den *Nullraum* bzgl. B. Entsprechend definieren wir den Nullraum $W_0 \subset W$:

$$W_0 = \{w \in W \mid B(v, w) = 0 \text{ für alle } v \in V\}.$$

Wir betrachten Faktorräume V/V_0 und W/W_0. Dann induziert B eine nichtausgeartete Bilinearform

$$\bar{B} : V/V_0 \times W/W_0 \to K$$

Satz 84 *(Dimensionsformel) Es seien V und W endlich erzeugte Vektorräume. Es sei*

$$B : V \times W \to K$$

eine Bilinearform. Wir bezeichenen mit $V_0 \in V$ und $W_0 \in W$ die Nullräume. Dann gilt

$$\dim V - \dim V_0 = \dim W - \dim W_0.$$

Beweis: Das folgt aus Satz 83, weil $\bar{B}$ nicht ausgeartet ist. *Q.E.D.*

Korollar 85 *Es sei V ein endlicher erzeugter Vektorraum. Es sei $B : V \times W \to K$ eine Bilinearform. Dann sind folgende Bedingungen äquivalent:*

(i) B ist nichtausgeartet in V.

(ii) Die Abbildung (92) $W \to V^*$ ist surjektiv.

Die zweite Bedingung besagt, dass zu jeder linearen Abbildung $\ell : V \to K$, ein Vektor $w \in W$ existiert, so dass

$$\ell(v) = B(v, w), \text{ für alle } v \in V.$$

Beweis: Angenommen (i) ist erfüllt. Dann wendet man Satz 83 nichtausgeartete Bilinearform

$$\bar{B} : V \times W/W_0 \to K$$

an und erhält (ii).

Umkehrt sei (ii) erfüllt. Es sei $v \in V$, so dass $v \neq 0$. Bei der Definition der Evaluation haben wir gesehen, dass ein $\ell \in V^*$ existiert, so dass $\ell(v) = 0$. Nach Voraussetzung finden wir ein $w \in W$, so dass für alle $v' \in V$

$$\ell(v') = B(v', w).$$

Insbesondere gilt also $B(v, w) \neq 0$. *Q.E.D.*

Es sei $f : V \to W$ ein Homomorphismus von Vektorräumen. Man definiert den dualen Homomorphismus

$$f^* : W^* \to V^*,$$

wie folgt: Es sei $\phi \in W^*$. Dann setzt man

$$f^*(\phi) = \phi \circ f \in V^*.$$

Man nennt f^* den dualen Homomorphismus. Wir drücken das mit Hilfe der Evaluationen $\mathcal{E}_V$ und $\mathcal{E}_W$ aus:

$$\mathcal{E}_V(v, f^*(\phi)) = \mathcal{E}_W(f(v), \phi), \quad v \in V,\ \phi \in W^*. \tag{95}$$

Wenn $h : V \to W$ ein weiterer Homomorphismus ist, so gilt

$$(f + h)^* = f^* + h^*, \quad (\lambda f)^* = \lambda f^* \text{ wo } \lambda \in K.$$

Es sei $v_1, \ldots, v_n$ eine Basis von V und $\hat{v}_1, \ldots, \hat{v}_n$ die duale Basis von V^*. Es sei $w_1, \ldots, w_m$ eine Basis von W und $\hat{w}_1, \ldots, \hat{w}_m$ die duale Basis von W^*.

Die Matrizen von f und f^* seien durch folgende Gleichungen definiert

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \quad f^*(\hat{w}_l) = \sum_{k=1}^{n} b_{kl} \hat{v}_k.$$

Wir finden

$$a_{ij} = \mathcal{E}_W(f(v_j), \hat{w}_i) = \mathcal{E}_V(v_j, f^*(\hat{w}_i)) = b_{ji}.$$

Also ist die Matrix von f^* in den dualen Basen die transponierte der Matrix von f.

Es sei $g : U \to V$ ein weiterer Homomorphismus von Vektorräumen. Dann gilt:

$$(f \circ g)^* = g^* \circ f^*. \tag{96}$$

In Matrizen entspricht das der folgenden Eigenschaft der transponierten Matrix

$${}^t(AB) = {}^tB \ {}^tA.$$

Präziser kann man Folgendes formulieren. Es sei S_n das Standardskalarprodukt auf K^n und S_m das Standardskalarprodukt auf K^m. Es sei $A : K^n \to K^m$, wo $A \in M(m \times n, K)$. Dann gilt

$$S_m(A\underline{x}, \underline{y}) = S_n(\underline{x}, {}^tA\underline{y}), \quad \underline{x} \in K^n, \ \underline{y} \in K^m.$$

Das ist natürlich klar.

Es sei V ein Vektorraum. Die Evaluation (93) definiert nach (91) eine lineare Abbildung

$$\kappa_V : V \to V^{**}.$$

Wir schreiben $\kappa_V(v) = \tilde{v}$ für einen Vektor $v \in V$. Explizit ist $\tilde{v}$ so definiert:

$$\tilde{v}(\phi) = \phi(v), \quad \text{für alle } \phi \in V^*.$$

Satz 86 *Es sei V ein endlich erzeugter Vektorraum. Die Abbildung:*

$$\kappa_V : V \to V^{**}, \quad \kappa_V(v) = \tilde{v}$$

ist ein Isomorphismus von Vektorräumen.

Beweis: Da $\dim V = \dim V^{**}$ genügt es zu zeigen, dass die Abbildung κ injektiv ist. Das folgt aus (89). *Q.E.D.*

Korollar 87 *Es sei $f : V \to W$ ein Homomorphismus von Vektorräumen. Dann ist folgendes Diagramm kommutativ:*

$$\begin{array}{ccc} V & \xrightarrow{\kappa_V} & V^{**} \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f^{**}} \\ W & \xrightarrow{\kappa_W} & W^{**} \end{array} \tag{97}$$

Beweis: Wir verwenden die Bezeichnungen aus Satz 86. Es sei $\ell \in W^*$. Dann gilt nach Definition von f^* für alle $v \in V$

$$f^*(\ell)(v) = \ell(f(v)).$$

Es sei $v \in V$ und es sei $\tilde{v} = \kappa_V(v)$. Wir müssen beweisen

$$f^{**}(\tilde{v})(\ell) = \widetilde{f(v)}(\ell).$$

In der Tat:

$$f^{**}(\tilde{v})(\ell) = \tilde{v}(f^*(\ell)) = f^*(\ell)(v) = \ell(f(v)) = \widetilde{f(v)}(\ell).$$

Die erste Gleichung ist die Definition von $f^{**} := (f^*)^*$. Die zweite Gleichung ist die Definition von $\tilde{v}$. Die dritte Gleichung ist die Definition von f^*. Die letzte Gleichung ist die Definition von $\widetilde{f(v)}$. *Q.E.D.*

Ein alternativer Beweis für das letzte Korollar geht so. Wir wählen eine Basis $v_1, \ldots, v_n$ von V und eine Basis $w_1, \ldots, w_n$ von W. Die duale Basis zu $\hat{v}_1, \ldots, \hat{v}_n \in V^*$ in V^{**} ist die Basis $\tilde{v}_1, \ldots \tilde{v}_n$. Entsprechendes gilt für W^* und W^{**}. In diesen Basen sind die Matrizen von κ_V und κ_W die Einheitsmatrizen. Die Matrix von f sei A. Wir haben gesehen, dass ${}^tA$ die Matrix von f^* ist. Genauso ist die Matrix von $f^{**} := (f^*)^*$ gleich ${}^t({}^tA) = A$. Damit ist die Kommutativität von (97) in den gewählten Koordinaten offensichtlich:

$$AE_n = E_m A.$$

Das folgende Lemma ist eine Reformulierung der Tatsache, dass es Komplementärräume gibt.

Lemma 88 *Es sei $f : V \to W$ ein Homomorphismus von Vektorräumen. Dann gilt*

1. *Die Abbildung f ist genau dann surjektiv, wenn ein Homomorphismus $s : W \to V$ existiert, so dass $f \circ s = \mathrm{id}_W$.*

2. *Die Abbildung f ist genau dann injektiv, wenn ein Homomorphismus $t : W \to V$ existiert, so dass $t \circ f = \mathrm{id}_V$.*

Beweis: Wenn s existiert, so sieht man leicht, dass f surjektiv ist und wenn t existiert, dass f injektiv ist.

Es sei f surjektiv. Es sei $L \subset V$ ein Komplementärraum zu $\mathrm{Ker}\, f$. Dann ist die Abbildung $f : L \to W$ immer noch surjektiv und ausser dem injektiv. Dann ist f bijektiv und deshalb gibt es eine Umkehrabbildung $g : W \to L$. Das Kompositum von g mit der Inklusion $L \subset V$ ist die gesuchte Abbildung $s : W \to V$.

Es sei jetzt f injektiv. Dann ist f ein Isomorphismus von V auf das Bild $f(V)$. Daher kann man annehmen, dass $V = f(V)$ und f die Inklusion $f(V) \subset V$ ist.

Jetzt ist also V ein Unterraum von W und $f(v) = v$ für $v \in V$. Es sei $M \in W$ ein Komplementärraum von V und es sei $t : W \to V$ die Projektion längs M. Dann gilt $t(v) = v$ und dann $t \circ f(v) = v$ für alle $v \in V$. Das wollten wir beweisen. *Q.E.D.*

Satz 89 *Es sei*

$$0 \longrightarrow U \xrightarrow{g} V \xrightarrow{f} W \longrightarrow 0$$

eine kurze exakte Sequenz von Vektorräumen.

Dann ist

$$0 \longrightarrow W^* \xrightarrow{f^*} V^* \xrightarrow{g^*} U^* \longrightarrow 0$$

eine kurze exakte Sequenz.

Beweis: Da f surjektiv ist, findet man eine Abbildung $s : W \to V$, so dass $f \circ s = \mathrm{id}_W$. Nach (96) findet man daraus $s^* \circ f^* = \mathrm{id}_{W^*} = (\mathrm{id}_W)^*$. Deshalb ist die Abbildung f^* injektiv. Genauso folgert man aus Lemma 88, dass g^* surjektiv ist.

Man muss noch die Exaktheit der Sequenz in V^* beweisen. Nach (96) findet man $g^* \circ f^* = (f \circ g)^* = 0^* = 0$. Deshalb bleibt nur noch folgendes zu zeigen: Es sei $\phi \in V^*$ ein Element, so dass $g^*(\phi) = 0$. Dann gibt es ein $\psi \in W^*$ mit $f^*(\psi) = \phi$. Das Element ϕ ist eine lineare Abbildung $\phi : V \to K$. Nach Voraussetzung ist $\phi \circ g = 0$. Aber dann existiert nach

Satz 13 ein Homomorphismus $\psi : W \to K$, so dass folgendes Diagramm kommutativ ist:

orthogonale Komplement

$$\begin{array}{ccccc} U & \xrightarrow{g} & V & \xrightarrow{f} & W \\ & & \downarrow{\scriptstyle \phi} & \swarrow{\scriptstyle \psi} & \\ & & K & & \end{array}$$

Q.E.D.

Korollar 90 *Es sei*

$$U \xrightarrow{g} V \xrightarrow{f} W \tag{98}$$

eine Sequenz von Vektorräumen, die in V exakt ist.

Dann ist

$$W^* \xrightarrow{f^*} V^* \xrightarrow{g^*} U^*$$

in V^* exakt.

Wenn f eine Injektion ist, so ist f^* surjektiv. Wenn g eine Surjektion ist, so ist g^* injektiv.

Es sei $B : V \times W \to K$ eine Bilinearform. Es sei $S \subset V$ ein Unterraum. Dann definiert man einen Unterraum $S^\perp \subset W$:

$$S^\perp = \{w \in W \mid B(v, w) = 0 \text{ für } \quad \text{alle } v \in S\}$$

$S^\perp$ nennt man das *orthogonale Komplement* zu S. Analog definiert man für einen Unterraum $T \subset W$ das orthogonale Komplement $T^\perp \subset V$.

Man kann die Definition des orthogonalen Komplements auch so formulieren: Es sei $B_S : S \times W \to K$ die Einschränkung von B. Dann ist $S^\perp \in W$ der Nullraum von B_S bezüglich des zweiten Arguments. Genauso erhält man $T^\perp \in V$ als Nullraum der Einschränkung $B_T : V \times T \to K$.

Es seien S_1, S_2 Unterräume von V. Dann gilt

$$\begin{aligned} S_1 \subset S_2 \quad &\Rightarrow \quad S_2^\perp \subset S_1^\perp \\ (S_1 + S_2)^\perp &= S_1^\perp \cap S_2^\perp. \end{aligned} \tag{99}$$

Die erste Implikation ist klar. Wir beweisen die letzte Gleichung. Die Inklusion $\subset$ folgt aus der ersten Implikation. Es sei umgekehrt $w \in S_1^\perp \cap S_2^\perp$. Dann gilt für jeden Vektor $s = s_1 + s_2 \in S_1 + S_2$, dass

$$B(s, w) = B(s_1, w) + B(s_2, w) = 0 + 0 = 0,$$

d.h. $w \in (S_1 + S_2)^\perp$.

Satz 91 *Es seien V und W endlich erzeugte Vektorraum. Es sei $B : V \times W \to K$ eine Bilinearform, die in V nicht ausgeartet ist. Es seien $S, S_1, S_2 \subset V$ Unterräume. Dann gilt:*

$$\dim S + \dim S^\perp = \dim W, \quad (S^\perp)^\perp = S. \tag{100}$$

$$S_1^\perp + S_2^\perp = (S_1 \cap S_2)^\perp \tag{101}$$

Beweis: Die Dimensionsformel folgt, wenn man Satz 84 auf die Einschränkung $B_S : S \times W \to K$ anwendet.

Man sieht leicht, dass $S \subset (S^\perp)^\perp$. Um die Gleichheit zu beweisen, kann man W durch W/W_0 ersetzen. Man kann also annehmen, dass B auch in W nichtausgeartet ist. Dann gilt $\dim V = \dim W$. Wenn wir die Rollen von V und W vertauschen folgt aus der schon bewiesenen Dimensionsformel

$$\dim(S^\perp)^\perp + \dim S^\perp = \dim V = \dim W.$$

Also gilt: $\dim S = \dim(S^\perp)^\perp$ und da wir eine Inklusion haben auch $S = (S^\perp)^\perp$.

Auch für den Beweis von 101 kann man annehmen, dass B in W nicht augeartet ist. W nicht ausgeartet ist. Dann ist (100) auch für Unterräume $T \subset W$ gültig: $T = (T^\perp)^\perp$.

Wir setzen $T_1 = S_1^\perp$ und $T_2 = S_2^\perp$. Nach (99) gilt:

$$(T_1 + T_2)^\perp = T_1^\perp \cap T_2^\perp.$$

Wir nehmen das orthogonale Komplement auf beiden Seiten der Gleichung.

$$S_1^\perp + S_2^\perp = (T_1 + T_2) = (T_1^\perp \cap T_2^\perp)^\perp = (S_1 \cap S_2)^\perp.$$

Q.E.D.

Bemerkung: Aus dem Beweis folgt, dass die Formel $\dim S + \dim S^\perp = \dim W$ unter der schwächeren Voraussetzung gilt, dass die Einschränkung $B_S : S \times W \to K$ in S nicht ausgeartet ist.

Es sei $u_1, \ldots, u_t$ ein Erzeugendensystem von $S \subset V$. Es sei $w_1, \ldots, w_m$ ein Basis von W. Ein Vektor

$$w = \sum_{j=1}^{m} \xi_j w_j, \quad \xi_j \in K,$$

liegt genau dann im orthogonalen Komplement von S, wenn die ξ_j dem folgenden Gleichungssystem genügen:

$$\sum_{j=1}^{m} B(u_i, w_j)\xi_j = 0, \quad \text{für } i = 1, \ldots, t.$$

Der Rang der Koeffizientenmatrix ist $m - \dim S^\perp = \dim S$. Das gibt uns die Möglichkeit, dass orthogonale Komplement in konkreten Fällen zu berechnen.

Satz 92 *Es sei $f : V \to W$ eine lineare Abbildung von Vektorräumen. Dann gilt:*

$$(\operatorname{Ker} f)^\perp = \operatorname{Im} f^*, \quad \operatorname{Ker} f^* = (\operatorname{Im} f)^\perp. \tag{102}$$

Hier sind die orthogonalen Komplemente bezüglich der Evaluation $\mathcal{E}_V$ bzw. $\mathcal{E}_W$ genommen.

Beweis: Zum Beweis der ersten Gleichung verwenden wir die exakte Sequenz

$$0 \to \operatorname{Ker} f \xrightarrow{\iota} V \xrightarrow{f} W.$$

Genauer meinen wir, dass diese Sequenz an den Stellen $\operatorname{Ker} f$ und V exakt ist.

Wenn wir dualisieren, erhalten wir die exakte Sequenz

$$W^* \xrightarrow{f^*} V^* \xrightarrow{\iota^*} (\operatorname{Ker} f)^* \to 0.$$

Der Kern von ι^* ist nach Definition $(\operatorname{Ker} f)^\perp$. Die Exaktheit sagt, dass dieser Kern das Bild $\operatorname{Im} f^*$ ist.

Zum Beweis der zweiten Gleichung betrachten wir die exakte Sequenz

$$V \xrightarrow{f} W \to W/\operatorname{Im} f \to 0.$$

Dann können wir die zweite Gleichung aus der exakten Sequenz ablesen, die durch Dualisierung entsteht:

$$0 \to (W/\operatorname{Im} f)^* \to W^* \xrightarrow{f^*} V^*.$$

Q.E.D.

Anwendung: Es seien $B_1 : V \times V' \to K$ und $B_2 : W \times W' \to K$ nicht ausgeartete Bilinearformen endliche erzeugter Vektorräume. Sie definieren Isomorphismen $V' \cong V^*$ und $W' \cong W^*$. Deshalb kann man so tun als wäre V' der duale Raum zu V und W' der zu W. Es sei $f : V \to W$ ein Homomorphismus. Der duale Homomorphismus $f' : W' \to V'$ ist durch

$$B_1(v, f'(w')) = B_2(f(v), w'), \quad v \in V,\ w' \in W,$$

gegeben (vgl.(95)). Wenn man f' vermittels der Isomorphismen $V' \cong V^*$ und $W' \cong W^*$ zu einem Morphismus $W^* \to V^*$ tranportiert, erhält man f^*. Daher gelten die Gleichungen:

$$(\operatorname{Ker} f)^\perp = \operatorname{Im} f', \quad \operatorname{Ker} f' = (\operatorname{Im} f)^\perp. \tag{103}$$

Wir nehmen jetzt an, dass $V = V' = K^n$ und dass B_1 das Standardskalarprodukt ist. Genauso sei $W = W' = K^m$ und B_2 das Standardskalarprodukt. Dann ist $f : K^n \to K^m$ die Multiplikation mit einer Matrix $A \in M(m \times n, K)$. Die Abbildung $f' : K^m \to K^n$ ist durch die duale Matrix ${}^tA$ gegeben:

$$B_1(\underline{x},\ {}^tA\underline{y}) = B_2(A\underline{x}, \underline{y}), \quad \underline{x} \in K^n,\ \underline{y} \in K^m.$$

Es gilt nach (103), dass

$$(\operatorname{Ker}\ {}^tA)^\perp = \operatorname{Im} A, \tag{104}$$

wobei das orthogonale Komplement bezüglich B_2 genommen ist. Man sieht, dass

$$\operatorname{Ker}\ {}^tA = \{\underline{z} \mid\ {}^t\underline{z}A = 0\}.$$

Also gilt:

Ein Gleichungssystem

$$A\underline{x} = \underline{b}, \quad \underline{b} \in K^m$$

hat genau dann eine Lösung $\underline{x}$, wenn $\underline{b}$ die folgende Eigenschaft hat:

Für alle $\underline{z} \in K^m$, so dass ${}^t\underline{z}A = 0$, gilt ${}^t\underline{z}\,\underline{b} = 0$.

Man diese Tatsache auch so einsehen: Wenn eine Lösung $\underline{x}$ existiert, so gilt ${}^t\underline{z}\,\underline{b} = 0$. Umgekehrt sei $\underline{b}$ mit der geforderten Eigenschaft gegeben. Es sei $S \subset K^m$ der Raum, welcher von den Spalten der Matrix A aufgespannt

wird. Die Gleichung ${}^t zA = 0$ ist äquivalent mit $\underline{z} \in S^\perp$ (bezüglich B_2. Die für $\underline{b}$ geforderte Eigenschaft besagt also $\underline{b} \in (S^\perp)^\perp = S$. Also liegt $\underline{b}$ in der linearen Hülle der Spalten von A. Aber das besagt gerade, dass $A\underline{x} = \underline{b}$ eine Lösung hat.

Übung: Es seien $L \subset K^m$ und $R \subset K^n$ Unterräume. Es gelte

$$n - \dim R = m - \dim L.$$

Man beweise, dass es eine Matrix A gibt, so dass

$$L = \{\underline{x} \in K^m \mid {}^t\underline{x}A = 0\}, \quad R = \{\underline{y} \in K^n \mid A\underline{y} = 0\}.$$

Wir betrachten jetzt den Basiswechsel. Es sei

$$B : V \times W \to K.$$

Wir wählen ein Koordinatensystem $\gamma : K^m \to V$ und schreiben $\underline{x}(v) = \gamma^{-1}(v)$ für Koordinaten eines Vektors $v \in V$. Wir wählen ein Koordinatensystem $\delta : K^n \to W$ und schreiben $\underline{y}(w) = \delta^{-1}(w)$ für $w \in W$.

Dann gilt:

$$B(v, w) = {}^t\underline{x}(v) G \underline{y}(w),$$

wobei G die Gramsche Matrix ist, bezüglich der Basen, die durch die Koordinatensytem γ und δ definiert werden.

Wir wählen weitere Koordinatensysteme $\gamma' : K^m \to V$ und $\delta' : K^n \to W$. Dann können wir bezüglich dieser Koordinatensysteme schreiben:

$$B(v, w) = {}^t\underline{x'}(v) G' \underline{y'}(w)$$

Es seien $C \in M(m \times m, K)$ und $D \in M(n \times n, K)$ die Übergangsmatrizen, d.h. $C\underline{x'}(v) = \underline{x}(v)$ und $D\underline{y'}(w) = \underline{y}(w)$ Dann finden wir:

$$B(v, w) = {}^t\underline{x}(v) G \underline{y}(w) = {}^t\underline{x'}(v)\ {}^t C G D \underline{y'}(w).$$

Daraus ergibt sich, dass die beiden Gramschen Matrizen G und G' von B bezüglich verschiedener Koordinatensysteme wie folgt zusammenhängen:

$$G' = {}^t C G D. \tag{105}$$

Beispiel: Es sei $V = K^m$ und $W = K^n$. Es sei $B : V \times W \to K$ und es sei G ihre Gramsche Matrix.

Der Nullraum W_0 ist die Menge der Lösungen des Gleichungssystems

kongruent

$$G\underline{y} = 0, \quad y \in K^n.$$

Also gilt Spaltenrang $G = n - \dim W_0 = \dim W - \dim W_0$. Der Nullraum V_0 ist die Menge der Lösungen des Gleichungssystems

$${}^t\underline{x}G = 0 \text{ oder } {}^tG\underline{x} = 0, \quad \underline{x} \in K^m.$$

Also gilt: Spaltenrang ${}^tG = m - \dim V_0 = \dim V - \dim V_0$. Weil die bilineare Abbildung

$$B : V/V_0 \times W/W_0 \to K$$

in beiden Argumenten nichtausgeartet ist, gilt $\dim V - \dim V_0 = \dim W - \dim W_0$. Wir folgern die bereits bewiesene Tatsache:

$$\text{Spaltenrang}\, G = \text{Spaltenrang}\ {}^tG.$$

Wir betrachten Bilinearformen

$$B : V \times V \to K$$

auf einem Vektorraum V der Dimension n. Wenn wir ein Koordinatensystem $\underline{x}(v) \in K^n$ auf V wählen, so gilt:

$$B(v_1, v_2) = {}^t\underline{x}(v_1)G\underline{x}(v_2), \quad v_1, v_2 \in V.$$

Hier ist $G \in M(n \times n, K)$ die Gramsche Matrix. Wenn $\underline{x'}(v)$ ein weiteres Koordinatensystem von V ist und G' die Gramsche Matrix von B bezüglich dieses Koordinatensystems, so gilt

$$G' = {}^tDGD, \tag{106}$$

wobei $D \in GL_n(K)$ die Übergangsmatrix ist, d.h. $D\underline{x'} = \underline{x}$. Wir nennen zwei Matrizen $G, G' \in M(n \times n, K)$ *kongruent* , wenn es eine Matrix $D \in \mathrm{GL}_n(K)$ gibt, so dass die Gleichung (106) gilt. Kongruent ist eine Äquivalenzrelation.

Verschiedene Basen auf V führen also zu kongruenten Gramschen Matrizen von B.

Definition 93 *Wir nennen zwei Bilinearformen $B : V \times V \to K$ und $A : U \times U \to K$ kongruent, wenn es einen Isomorphimsus von Vektorräumen $\alpha : U \to V$ gibt, so dass*

$$A(u_1, u_2) = B(\alpha(u_1), \alpha(u_2)), \quad \textit{für alle } u_1, u_2 \in U.$$

Angenommen wir haben auf U und V beliebige Basen gewählt. Dann sind B und A genau dann kongruent, wenn ihre Gramschen Matrizen kongruent sind.

1.12 Quadratische Formen

Nullraum einer symmetrischen Bilinearform

Isometrie

Definition 94 *Ein Bilinearform $B : V \times V \to K$ heißt symmetrisch, wenn für alle $v_1, v_2 \in V$*

$$B(v_1, v_2) = B(v_2, v_1).$$

Das ist der Fall, wenn eine Gramsche Matrix S von B symmetrisch ist, d.h.

$${}^tS = S.$$

Es sei $V_0 \subset V$ der Nullraum von B bezüglich des ersten Arguments. Da B symmetrisch ist, ist das auch der Nullraum bezüglich des zweiten Arguments. Deshalb können wir einfach vom *Nullraum einer symmetrischen Bilinearform* B sprechen.

Zu einer symmetrischen Bilinearform B betrachtet man die Funktion $Q(v) = B(v, v)$ (*quadratische Form*). Wenn $2 \neq 0$ in K, so erhält man aus B wieder Q zurück:

$$B(v, w) = \frac{1}{2}(Q(v + w) - Q(v) - Q(w)) \tag{107}$$

$Q : V \to K$ ist keine lineare Abbildung. Es gilt:

$$Q(\lambda v) = \lambda^2 Q(v), \text{ für alle } \lambda \in K.$$

Bemerkung: Es sei V ein endlich erzeugter Vektorraum. Wenn man B als Tensor schreibt (siehe (73)) sieht man, dass Q eine Polynomfunktion ist.

Es sei K ein Körper der Charakteristik 0. Dann ist Q die Polarisierung von B im Sinne von Satz 81.

Wir bezeichen eine symmetrische Bilinearform in diesem Abschnitt mit (V, B). Eine Morphismus $\alpha : (V_1, B_1) \to (V_2, B_2)$ ist nach Definition eine lineare Abbildung von Vektorräumen $\alpha : V_1 \to V_2$, so dass

$$B_2(\alpha(v_1), \alpha(w_1)) = B_1(v_1, w_1), \quad \text{für alle } v_1, w_1 \in V_1.$$

Mit diesen Morphismen bilden die symmetrischen Bilinearformen eine Kategorie. Wir haben zwei quadratische Formen kongruent genannt, wenn sie in dieser Kategorie isomorph sind. Einen Isomorphismus von symmetrischen Bilinearformen nennen wir eine *Isometrie* .

Es seien Q_1 bzw. Q_2 die quadratischen Formen zu B_1 und B_2. Eine lineare Abbildung $\alpha : V_1 \to V_2$ ist genau dann ein Morphismus $(V_1, B_1) \to (V_2, B_2)$, wenn

$$Q_2(\alpha(v_1)) = Q_1(v_1), \quad \text{für alle } v_1 \in V_1.$$

Das folgt aus (107).

Definition 95 *Es sei (V, B) eine symmetrische Bilinearform. Es sei V endlich erzeugt. Wir nennen die Zahl $\dim V - \dim V_0$ den Rang von (V, B).*

Unter der Dimension von (V, B) verstehen wir $\dim V$.

Es sei S eine Gramsche Matrix zu (V, B). Man sieht, dass der Rang von (V, B) gleich dem Rang der Matrix S ist.

Eine quadratische Form B heißt nichtausgeartet, wenn $V_0 = 0$. Das ist genau dann der Fall wenn die Gramsche Matrix invertierbar ist. In diesem Fall gilt für einen beliebigen Unterraum $U \subset V$ nach Satz 91 die Dimensionsformel

$$\dim U + \dim U^{\perp} = \dim V.$$

Beispiel: Das Standardskalarprodukt auf K^n ist die symmetrische Bilinearform $B : K^n \times K^n \to K$, so dass die Gramsche Matrix in der Standardbasis die Einheitsmatrix E_n ist.

Die quadratische Form Q zu B ist die Funktion

$$Q(\underline{x}) = x_1^2 + x_2^2 + \ldots + x_n^2.$$

Es sei $2K \neq 0$. Es sei $A \in M(n \times n)$ eine Matrix. Dann ist die Funktion $F : K^n \to K$, so dass $F(\underline{x}) = {}^t\underline{x}A\underline{x}$, die quadratische Form einer symmetrischen Bilinearform. In der Tat,

$$S = (1/2)(A + {}^tA)$$

ist die Gramsche Matrix dieser Bilinearform.

Es sei $B : V \times V \to K$ eine symmetrische Bilinearform. Wir sagen, dass ein Unterraum $S \subset V$ nicht ausgeartet ist, wenn die Einschränkung von B auf $S \times S$ nicht ausgeartet ist. Wir sprechen auch einfach von der Einschränkung von B auf S.

Lemma 96 *Es sei $2K \neq 0$. Es sei $B : V \times V \to K$ eine symmetrische Bilinearform.*

orthogonale direkte Summe

Es sei $S \subset V$ ein nicht ausgearteter endlich erzeugter Unterraum. Dann gilt

$$V = S \oplus S^{\perp}.$$

Es sei allgemeiner $V = W_1 \oplus W_2$ die direkte Summe zweier Unterräume. Wir nennen dies eine *orthogonale direkte Summe* , wenn $W_1 \subset W_2^{\perp}$ oder äquivalent dazu, wenn $W_2 \subset W_1^{\perp}$.

Beweis: Es sei B' die Einschränkung von B auf S. Dann ist $S \cap S^{\perp} \subset S$ der Nullraum von B'. Daher gilt nach Voraussetzung, dass

$$S \cap S^{\perp} = 0.$$

Es sei $v \in V$. Wir definieren $\ell \in S^*$, durch $\ell(s) = B(v, s)$. Nach Satz 83 finden wir ein $t \in S$, so dass $\ell(s) = B'(t, s) := B(t, s)$ für alle $s \in S$. Also gilt $v - t \in S^{\perp}$. Wir finden die Zerlegung

$$v = t + (v - t), \quad t \in S,\ (v - t) \in S^{\perp}.$$

Q.E.D.

Es sei V endlich erzeugt. Es sei $v_1, \ldots, v_m$ eine Basis von S und $v_{m+1}, \ldots, v_n$ eine Basis von $S^{\perp}$. Dann ist nach dem Lemma $v_1, \ldots, v_n$ eine Basis von V. Da $B(v_i, v_j) = 0$, wenn $i \leq m$ und $j > m$ hat die Gramsche Matrix von B in der Basis $v_1, \ldots, v_n$ die Gestalt einer Blockmatrix:

$$\begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix}, \quad A \in M(m \times m, K),\ C \in M((n-m) \times (n-m), K).$$

Es sei $B : V \times V \to K$ eine symmetrische Bilinearform. Zwei Vektoren $v_1, v_2 \in V$ heißen *orthogonal*, wenn $B(v_1, v_2) = 0$. Eine Basis $v_1, \ldots, v_n$ von V heißt orthogonal, wenn $B(v_i, v_j) = 0$ für $i \neq j$. Wenn ausserdem $B(v_i, v_i) = 1$, für $i = 1, \ldots, n$, heißt sie *orthonormal*. Eine Basis ist orthogonal, wenn die Gramsche Matrix von B in dieser Basis diagonal ist. Es sei $\underline{x}(v)$ das Koordinatensystem zu einer orthogonalen Basis. Die quadratische Form Q sieht dann folgendermaßen aus

$$Q(v) = \lambda_1 x_1(v)^2 + \lambda_2 x_2(v)^2 + \ldots + \lambda_n x_n(v)^2.$$

Hier sind $\lambda_i \in K$.

Satz 97 *Es sei V ein endlichdimensionaler Vektorraum über einem Körper K, so dass $2K \neq 0$. Es sei $B : V \times V \to K$ eine symmetrische Bilinearform. Dann gibt es eine orthogonale Basis von V bezüglich B.*

Beweis: Wir führen den Beweis durch Induktion nach $\dim V$. Wenn $B = 0$ so ist nichts zu beweisen. Wenn $B \neq 0$, so folgt aus (107), dass $Q \neq 0$. Also gibt es ein $v \in V$ mit $B(v, v) \neq 0$. Es sei $L = \mathcal{L}(v)$. Die Einschränkung von B auf L ist nicht ausgeartet. Daher folgt aus dem letzten Lemma:

$$V = L \oplus L^{\perp}.$$

Nach Induktionsvoraussetzung gibt es eine orthogonale Basis $v_2, \ldots, v_n \in L^{\perp}$. Dann ist $v, v_2, \ldots, v_n$ die gesuchte orthogonale Basis von V. *Q.E.D.*

Beispiel: Wir betrachten die quadratische Form auf $V = \mathbb{Q}^2$:

$$Q(\underline{x}) = x_1^2 - 6x_1x_2 + 10x_2^2.$$

Die assozierte symmetrische Bilinearform ist $B : \mathbb{Q}^2 \times \mathbb{Q}^2 \to \mathbb{Q}$ ist gegeben durch:

$$B(e_1, e_1) = 1,\ B(e_2, e_2) = 10,\ B(e_1, e_2) = -3. \tag{108}$$

Wir wollen Q diagonalisieren. Vom algebraischen Standpunkt heißt das, die quadratische Ergänzung zu finden:

$$Q(\underline{x}) = (x_1 - 3x_2)^2 + x_2^2.$$

Man wählt neue Koordinaten:

$$y_1 = x_1 - 3x_2, \quad y_2 = x_2.$$

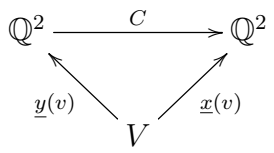
Oder äquivalent dazu

$$x_1 = y_1 + 3y_2, \quad y_2 = x_2. \tag{109}$$

In den Koordinatensystem $\underline{y}$ hat Q Diagonalgestalt:

$$y_1^2 + y_2^2. \tag{110}$$

Wir ermitteln jetzt welcher Basis von $V = \mathbb{Q}^2$ das Koordinatensystem $\underline{y}$ entspricht: Dazu betrachten wir das kommutative Diagramm



Hier ist $\underline{x}(v)$ die identische Abbildung und C die Matrix:

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

In der Tat, die Gleichung (109) besagt gerade $\underline{x} = C\underline{y}$.

Wegen (110) entsprechen die Standardvektoren in den $\underline{y}$-Koordinaten einer orthogonalen Basis von (V, B). In den $\underline{x}$ Koordinaten sind diese Standardvektoren die Spalten der Matrix C.

$$e_1,\ 3e_1 + e_2.$$

Das können wir überprüfen:

$$B(e_1, e_1 + 3e_2) = B(e_1, e_1) + 3B(e_1, e_2) = 0.$$

Das folgt aus (108).

Statt die quadratische Ergänzung zu benutzen, können wir auch geometrisch argumentieren. Da die Einschränkung von B auf $\mathcal{L}(e_1)$ nicht ausgeartet ist, haben wir eine direkte Summe

$$V = \mathcal{L}(e_1) \oplus \mathcal{L}(e_1)^\perp.$$

Daher können wir schreiben $e_2 = \lambda e_1 + u$, wo $\lambda \in \mathbb{Q}$ und $u \in \mathcal{L}(e_1)^\perp$. Wir finden

$$B(e_2, e_1) = \lambda B(e_1, e_1) + B(u, e_1) = \lambda B(e_1, e_1).$$

Mit (108) erhalten wir $\lambda = -3$. Also ist $u = 3e_1 + e_2$ orthogonal zu e_1. Das ist das gleiche Ergebnis, welches wir mit Hilfe der quadratischen Ergänzung schon gefunden haben.

Beispiel: Mit den Voraussetzungen des Satzes sei $K = \mathbb{C}$. Dann findet man eine orthogonale Basis $v_1, \ldots, v_n$ von V, so dass $B(v_i, v_i) = 1$ oder $= 0$. In der Tat: Es sei $B(v_i, v_i) = \mu_i \neq 0$. Man findet eine eine komplexe Zahl λ_i, so dass $\lambda_i^2 = \mu_i$. Dann gilt $B(\frac{1}{\lambda_i}v_i, \frac{1}{\lambda_i}v_i) = 1$. Man braucht also nur diese v_i durch $\frac{1}{\lambda_i}v_i$ zu ersetzen und erhält die gewünschte Basis.

Die Gramsche Matrix hat in dieser Basis die Form:

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Hier ist E_r die Einheitsmatrix und r der Rang von B.

Satz 98 *Es sei $K = \mathbb{C}$ der Körper der komplexen Zahlen. Zwei symmetrische Bilinearformen (V_1, B_1) und (V_2, B_2), so dass V_1 und V_2 endlich erzeugt sind, sind genau dann kongruent, wenn sie die gleiche Dimension und den gleichen Rang haben.*

Das ist klar.

Für $K = \mathbb{R}$ sind die Verhältnisse gerinfügig komplizierter. Es sei (V, B) eine symmetrische Bilinearform, so dass V endlich erzeugt ist. Es sei $v_1, \ldots, v_n$ eine orthogonale Basis von V. Wir können sie so wählen, dass

$$B(v_i, v_i) = \begin{cases} 1 \\ -1 \\ 0 \end{cases} .$$

In der Tat, es sei $B(v_i, v_i) = \mu_i \neq 0$. Es sei $\lambda_i = \sqrt{|\mu_i|}$. Wir ersetzen v_i durch $\frac{1}{\lambda_i} v_i$.

Definition 99 *Es sei (V, B) eine symmetrische Bilinearform über dem Körper $\mathbb{R}$. Man nennt (V, B) positiv definit, wenn für jeden Vektor $v \in V$, $v \neq 0$, gilt dass*

$$B(v, v) > 0.$$

Man nennt (V, B) negativ definit, wenn $(V, -B)$ positiv definit ist. Wir bemerken, dass eine positiv definite symmetrische Bilinearform nicht ausgeartet ist, denn $B(v, v) > 0$ besagt, dass v nicht im Nullraum liegt.

Es sei $G \in M(n \times n, K)$ eine symmetrische Matrix, d.h. $G = {}^tG$. Wir nennen G positiv definit, wenn die Bilinearform

$$S(\underline{x}, \underline{y}) = {}^t\underline{x} G \underline{y}$$

auf dem K^n positiv definit ist.

Es sei (V, B) eine symmetrische Bilinearform über $\mathbb{R}$ und es sei G die Gramsche Matrix bezüglich eine Basis. Dann ist B positiv definit, genau dann wenn G positiv definit ist.

Es sei (V, B) eine symmetrische Bilinearform. Wir nennen einen Unterraum $W \subset V$ positiv definit, wenn die Einschränkung von B auf W positiv definit ist. Wir nennen eine Vektor $v \in V$ positiv, wenn $B(v, v) > 0$.

Das Standardskalarprodukt ist ein Beispiel für eine positiv definite symmetrische Bilinearform.

Satz 100 *(von Sylvester) Es sei V ein Vektorraum der Dimension n über $\mathbb{R}$. Es sei $B : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gibt es nichtnegative ganze Zahlen r und s, $r + s \leq n$ und eine orthogonale Basis von V:*

$$v_1, \ldots, v_r, v_{r+1}, \ldots, v_{r+s}, v_{r+s+1}, \ldots, v_n, \tag{111}$$

so dass

$$\begin{aligned} B(v_i, v_i) &= 1, \quad \textit{für} \quad 1 \leq i \leq r, \\ B(v_i, v_i) &= -1, \quad \textit{für} \quad r+1 \leq i \leq r+s, \\ B(v_i, v_i) &= 0, \quad \textit{für} \quad r+s+1 \leq i \leq n. \end{aligned}$$

Für jede andere Basis dieser Art sind die Zahlen r und s die gleichen.

Wir nennen das Paar (r, s) nichtnegativer ganzer Zahlen die Signatur der symmetrischen Bilinearform (V, B).

Beweis: Die Existenz einer solchen Basis haben wir eben bewiesen.

Man nennt einen Unterraum $W \subset V$ einen maximalen positiv definiten Unterraum, wenn W positiv definit ist und wenn es keinen positiv definiten Unterraum von V gibt, der W echt enthält.

Man sieht leicht, dass ein positiv definiter Unterraum $W \subset V$ genau dann maximal ist, wenn in seinem orthogonalen Komplement $W^\perp \subset V$ kein Vektor v existiert, so dass $B(v, v) > 0$.

In der Tat, angenommen im orthogonalen Komplement von W gibt es keinen Vektor v, so dass $B(v, v) > 0$. Es sei W' ein positiv definiter Unterraum, so dass $W \subset W'$. Da W nicht ausgeartet ist, kann man nach Lemma 96 schreiben:

$$W' = L \oplus W, \quad \text{wo } L \subset W^\perp.$$

Für jeden Vektor $v \in L$, $v \neq 0$ gilt $B(v, v) > 0$. Daher gilt nach Annahme über W, dass $L = 0$. Also ist W maximal. Die Umkehrung ist jetzt ebenfalls klar.

Wenn man eine Basis $v_1, \ldots v_n$ wie im Satz hat, so ist $\mathcal{L}(v_1, \ldots, v_r)$ ein maximaler positiv definiter Unterraum, denn in seinem orthogonalen Komplement $\mathcal{L}(v_{r+1}, \ldots, v_n)$ liegt kein positiver Vektor.

Es seien W_1 und W_2 zwei maximale positiv definite Unterräume von V. Wir werden zeigen, dass $\dim W_1 = \dim W_2$. Nach dem was wir eben gesagt haben, folgt dass r unabhängig von der Wahl der Basis (111) ist. Da $r + s$ der Rang von B ist, ist diese Zahl ebenfalls unabhängig von der Basis (111). Damit wäre also der Satz von Sylvester bewiesen.

Um $\dim W_1 = \dim W_2$ zu zeigen, betrachten wir die Einschränkung von B:

$$B' : W_1 \times W_2 \to \mathbb{R}.$$

Nach Satz 83 genügt es zu zeigen, dass B' nicht ausgeartet ist. Aus Symmetriegründen reicht es, dies für das Argument W_1 zu zeigen. Angenommen es gäbe ein $w_1 \in W_1$, so dass $w_1 \neq 0$ und $B(w_1, w_2) = 0$ für alle $w_2 \in W_2$. Dann wäre $w_1 \in W_2^\perp$ ein positiver Vektor. Das widerspricht der Maximalität von W_2. *Q.E.D.*

Korollar 101 *Es sei (V, B) ein endlich erzeugter Vektorraum mit einer symmetrischen Bilinearform B. Es sei (s, t) die Signatur von (V, B).*

Es sei $W \subset V$ ein Unterraum. Wir bezeichen mit (s_W, t_W) die Signatur der Einschränkung von B auf W. Dann gilt:

$$s_W \leq s, \quad t_W \leq t.$$

Beweis: Es sei $W_0 \subset W$ der Nullraum von B auf W. Es sei W' ein Komplement: $W = W_0 \oplus W'$. Wenn man eine Basis von W' wie im Satz 100 nimmt und sie zu einer Basis von W ergänzt, so sieht man, dass W und W' die gleiche Signatur haben.

Also kann man annehmen, dass die Einschränkung von B auf W nichtausgeartet ist. Es sei $W^\perp$ das orthogonale Komplement von W. Dann gilt:

$$V = W \oplus W^\perp.$$

Die Behauptung folgt, wenn man Satz 100 auf W und $W^\perp$ anwendet. *Q.E.D.*

Übung: Es sei (V, B) eine symmetrische Bilinearform über K, so dass V endlich erzeugt ist. Ein Unterraum $W \subset V$ heißt isotrop, wenn die Einschränkung von B auf $W \times W$ null ist.

Man beweise, dass zwei maximale isotrope Unterräume die gleiche Dimension haben. (Hinweis: Es sei W' ein zweiter maximaler isotroper Unterraum. Man betrachte die Einschränkung $B_r : W \times W' \to K$ von B. Man zeige, dass $W \cap W'$ der Nullraum bezüglich beider Argumente ist.)

Es sei $B : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform über den reellen Zahlen $\mathbb{R}$. Es sei V endlich erzeugt. Wir wählen zwei verschiedene Basen $v_1, \ldots, v_n$ und $v'_1, \ldots, v'_n$ von V. Wir bezeichen mit $x(v)$ bzw. $x'(v)$ die Koordinaten eines Vektors bezüglich dieser Basen. Es seien S und S' die Gramschen Matrizen von B bezüglich dieser Basen. Es sei $D \in \mathrm{GL}_n(K)$ die

Übergangsmatrix: $Dx'(v) = x(v)$. Dann gilt nach (106), dass $S' = {}^tDSD$. Wenn wir den Multiplikationssatz anwenden und $\det D = \det {}^tD$ beachten, so finden wir

Hauptminoren

$$\det S' = \det S(detD)^2. \tag{112}$$

Insbesondere ist das Vorzeichen von $\det S$ unabhängig von der Basis, in der wir die Gramsche Matrix von B berechnet haben.

Wir erhalten: Es sei (V, B) nichtausgeartet von der Signatur (r, s). Dann gilt für jede Gramsche Matrix S von (V, B), dass

$$\det S = (-1)^s\epsilon, \quad \epsilon \in \mathbb{R},\ \epsilon > 0. \tag{113}$$

Insbesondere ist $\det S$ positiv, wenn B positiv definit ist.

Es sei

$$S = \begin{pmatrix} s_{11} & \dots & s_{1n} \\ & \dots & \\ s_{n1} & \dots & s_{nn} \end{pmatrix}$$

Wir betrachten für $1 \leq r \leq n$ die folgenden Matrizen:

$$\begin{pmatrix} s_{11} & \dots & s_{1r} \\ & \dots & \\ s_{r1} & \dots & s_{rr} \end{pmatrix}$$

Wir nennen sie die *Hauptminoren* von S.

Satz 102 *Es sei $S \in M(n\times n, \mathbb{R})$ eine symmetrische Matrix. Wir betrachten auf $\mathbb{R}^n$ die symmetrische Bilinearform:*

$$B(\underline{x}, \underline{y}) = {}^t\underline{x}S\underline{y}$$

Dann ist $(\mathbb{R}^n, B)$ genau dann positiv definit, wenn die Determinanten aller Hauptminoren von S positiv sind.

Beweis: Wir nehmen an, dass die Determinanten aller Hauptminoren positiv sind und zeigen, dass dann B positiv definit ist.

Es sei $W = \mathcal{L}(e_1, \dots, e_{n-1}) \subset \mathbb{R}^n$. Wir können mit Hilfe von Induktion annehmen, dass die Einschränkung von B auf W positiv definit ist. Die Signatur von W ist dann $(n-1, 0)$. Es sei (r, s) die Signatur von $(\mathbb{R}^n, B)$. Da $\det S \neq 0$ ist B nichtausgeartet und folglich $r + s = n$ Nach dem Korollar 101 ist nur $s = 1$ oder $s = 0$ möglich. Aber im Fall $s = 1$ wäre nach (113) $\det S < 0$. Also ist $s = 0$ und B positiv definit. *Q.E.D.*

(Man folgere aus dem Hauptminorenkriterium die Ungleichung von Cauchy-Schwarz, siehe unten.)

Es sei $B : V \times V \to K$ eine symmetrische nichtausgeartete Bilinearform auf einem endlichdimensionalen K-Vektorraum. Es sei $f : V \to V$ ein Endomorphismus. Dann gibt es genau einen Endomorphismus $f^* : V \to V$, so dass für beliebige $v, w \in V$:

$$B(f(v), w) = B(v, f^*(w))$$

Satz 103 *Wenn f und f^* adjungdiert sind so gilt:*

$$\begin{array}{l} (f^*)^* = f, \\ (\mathrm{Ker}\, f)^\perp = \mathrm{Im}\, f^*, \quad \mathrm{Ker}\, f = (\mathrm{Im}\, f^*)^\perp, \\ (\mathrm{Im}\, f)^\perp = \mathrm{Ker}\, f^*, \quad \mathrm{Im}\, f = (\mathrm{Ker}\, f^*)^\perp. \end{array}$$

1.13 Euklidische und unitäre Vektorräume

Definition 104 *Ein euklidischer Vektorraum (V, B) ist ein endlich erzeugter Vektorraum V über dem Körper der reellen Zahlen $\mathbb{R}$, der mit einer symmetrischen, positiv definiten Bilinearform $B : V \times V \to \mathbb{R}$ versehen ist. Wir nennen B eine euklidische Metrik.*

Es sei (V, B) ein euklidischer Vektorraum. Wenn keine Verwechslungen zu befürchten sind, schreiben wir

$$< v_1, v_2 >:= B(v_1, v_2),$$

und erwähnen B gar nicht mehr explizit. Wir schreiben auch

$$< v_1, v_2 >_V =< v_1, v_2 >,$$

wenn mehrere euklidische Vektorräume im Spiel sind.

Das wichtigste Beispiel eines euklidischen Vektorraums ist $\mathbb{R}^n$ mit dem Standardskalarprodukt:

$$< \underline{x}, \underline{y} >:= {}^t\underline{x}\underline{y} \quad \text{wo } \underline{x}, \underline{y} \in \mathbb{R}^n.$$

Wir werden einfach von dem euklidischen Vektorraum $\mathbb{R}^n$ reden, wenn wir das Standardsklarprodukt meinen.

L“”ange
Schmidtorthogona

Wir definieren die *Länge* eines Vektors $v \in V$ als

$$|v| = \sqrt{< v, v >}. \tag{114}$$

Für zwei Vektoren $v, w \in V$ gilt:

$$| < v, w > | \leq |v||w| \tag{115}$$

In der Tat: Es seien v und w linear unabhängig. Es sei $U = \mathcal{L}(v, w)$. Da die Einschränkung der euklidischen Metrik auf U positiv definit ist, ist die Determinante der Gramschen Matrix bezüglich der Basis v, w positiv. Also gilt (115) mit “echt kleiner”:

$$\det \begin{pmatrix} < v, v > & < v, w > \\ < w, v > & < w, w > \end{pmatrix} > 0.$$

Wenn v, w linear abhängig sind, so sind beide Seiten von (115) gleich.

Aus (115) erhält man die Dreiecksungleichung:

$$|v + w| \leq |v| + |w|.$$

Man definiert den Winkel $\alpha = \angle(v, w)$ zwischen zwei Vektoren $v, w \in V \setminus \{0\}$ als die Gradzahl zwischen 0^o und 180^o, so dass

$$\cos \alpha = \frac{< v, w >}{|v||w|}. \tag{116}$$

Man beachte, dass nach (115) die rechte Seite dieser Gleichung im Intervall $[-1, 1]$ liegt.

Nach dem Satz von Sylvester besitzt ein Euklidischer Vektorraum (V, B) eine orthonormale Basis. Orthogonale Basen kann mit der *Schmidtorthogonalisierung* konstruieren. Es sei $v_1, \ldots, v_n$ eine beliebige Basis von V. Man konstruiert schrittweise Vektoren $w_1, \ldots, w_r$, die orthogonal sind, und so dass

$$\mathcal{L}(v_1, \ldots, v_r) = \mathcal{L}(w_1, \ldots, w_r). \tag{117}$$

Für $r = 1$ wählt man $v_1 = w_1$. Angenommen $w_1, \ldots, w_r$ wären bereits konstruiert.

Dann setzt man:

$$w_{r+1} = v_{r+1} + \sum_{i=1}^{r} \lambda_i w_i, \tag{118}$$

und bestimmt die $\lambda_i \in \mathbb{R}$, so dass $B(w_{r+1}, w_i) = 0$ für $i \leq r$. Da die w_i für $i \leq r$ schon orthogonal sind, findet man die λ_i aus den Gleichungen $0 = B(w_{r+1}, w_i) = B(v_{r+1}, w_i) + \lambda_i B(w_i, w_i)$. Wir sagen, dass $w_1, \ldots, w_n$ aus $v_1, \ldots, v_n$ durch Schmidtorthogonalisierung entsteht.

Am Schluss kann man noch w_i durch $\breve{w}_i = w_i/(\sqrt{B(w_i, w_i)})$ ersetzen. Dann erhählt man eine orthonormale Basis $\breve{w}_1, \ldots, \breve{w}_n$.

Wenn B nicht positiv definit ist, kann das Verfahren versagen, da $B(w_i, w_i) = 0$ möglich ist.

Die Schmidtorthogonalisierung hat folgende Interpretation mit Hilfe von Matrizen. Es sei $s \in V$ ein Vektor. Es seien $\underline{x}(s)$ seine Koordinaten bezüglich der Basis $v_1, \ldots, v_n$ und es seien $\underline{y}(s)$ seine Koordinaten bezüglich der Basis $w_1, \ldots, w_n$. Es sei $U = (u_{ij}) \in GL_n(\mathbb{R})$ die Übergangsmatrix, d.h. $\underline{y}(s) = U\underline{x}(s)$. Anders ausgedrückt sind die u_{ij} durch die folgenden Gleichungen bestimmt

$$v_j = \sum_{i=1}^{n} u_{ij} w_i, \quad j = 1, \ldots, n.$$

Diese Gleichung besagt, dass die y-Koordinaten des Vektors v_j gerade die Einträge der j-ten Spalte der Matrix U sind. Ein Vergleich mit den Gleichungen (117) zeigt, dass U eine obere Dreiecksmatrix ist, auf deren Diagonale Einsen stehen.

Es sei G die Gramsche Matrix von B bezüglich der Basis $v_1, \ldots, v_n$. Es sei H die Gramsche Matrix von B bezüglich der Basis $w_1, \ldots, w_n$. Wegen $B(w_i, w_j) = 0$ für $i \neq j$ ist Da dies eine orthogonale Basis ist H eine Diagonalmatrix. Da B positiv definit ist, sind die Einträge auf der Diagonale von B positiv. Es gilt nach Definition

$$B(s,t) = {}^t\underline{y}(s) H \underline{y}(t) = {}^t\underline{x}(s)\, {}^tUHU\underline{x}(t).$$

Folglich gilt $G = {}^tUHU$.

Offensichtlich gibt es eine Diagonalmatrix D mit $D^2 = H$. Wenn also $G \in M(n \times n, \mathbb{R})$ eine symmetrische positiv definite Matrix ist, so existiert eine obere Dreiecksmatrix $T \in GL_n(\mathbb{R})$, so dass $G = {}^tTT$. In der Tat $T = DU$ erfüllt das.

Übung: Man beweise, dass $\det U = 1$. Insbesondere gilt $\det G = \det H$. Man beweise, dass $B(w_{r+1}, w_{r+1}) \leq B(v_{r+1}, v_{r+1})$. Man folgere, dass

$$\det G \leq \prod_{i=1}^{n} B(v_i, v_i).$$

Es sei V ein euklidischer Vektorraum und $W \subset V$ ein Unterraum. Dann hat man nach Satz 97 eine orthogonle Zerlegung:

$$V = W \oplus W^{\perp}.$$

Die Projektion $p_W : V \to V$ von V auf W längs $W^{\perp}$ nennen wir die orthogonale Projektion von V auf W. Es gilt:

$$p_W(v) \in W, \quad v - p_W(v) \in W^{\perp}.$$

Man kann die Projektion p_W so berechnen: Man verschafft sich eine orthonormale Basis $u_1, \ldots, u_m$ von W. Man macht den Ansatz

$$p_W(v) = \sum_{i=1}^{m} \lambda_i u_i, \quad \lambda_i \in \mathbb{R}.$$

Die Koeffizienten gewinnt man aus $\lambda_i =< p_W(v), u_i >=< v, u_i >$.

Definition 105 *Es seien V und W euklidische Vektorräume. Ein Isomorphismus $\phi : V \to W$ von Vektorräumen heißt eine Isometrie, wenn*

$$< \phi(v_1), \phi(v_2) >=< v_1, v_2 > \quad v_1, v_2 \in V. \tag{119}$$

Wenn $\dim V = \dim W$ ist, so folgt aus (119) automatisch, dass ϕ ein Isomorphismus sein muss.

Es sei $v_1, \ldots, v_n$ eine orthonormale Basis von V. Wenn ϕ eine Isometrie ist, so ist $\phi(v_1), \ldots, \phi(v_n)$ eine orthonormale Basis von W. Wenn $w_1, \ldots, w_n$ eine orthonormale Basis von W ist, so ist die eindeutig bestimmte lineare Abbildung $\psi : V \to W$, mit $\psi(v_i) = w_i$ eine Isometrie.

Wenn also V und W euklidische Vektorräume der gleichen Dimension sind, so gibt es eine Isometrie $\phi : V \to W$.

Übung: Es sei V ein euklidischer Vektorraum der Dimension n. Es seien

$$\{0\} = V_0 \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = V$$
$$\{0\} = V_0' \subset V_1' \subset \ldots \subset V_{n-1}' \subset V_n' = V,$$

zwei volle Fahnen von V, d.h. $\dim V_i = \dim V_i' = i$.

Man beweise, dass eine Isometrie $\phi : V \to V$ existiert, so dass $\phi(V_i) = V_i'$.

(Hinweis: Gram-Schmidt.)

Ein *kartesisches Koordinatensystem* auf einem euklidischen Vektorraum V ist eine Isometrie $\gamma : \mathbb{R}^n \to V$. Die Bilder der Standardvektoren $\gamma(e_1), \ldots, \gamma(e_n)$ sind eine orthonormale Basis von V. Auf diese Weise entsprechen sich orthonormale Basen und kartesische Koordinatensysteme von V bijektiv. Ein euklidischer Vektorraum besitzt also stets ein kartesisches Koordinatensystem.

kartesisches Koordinatensyst
Drehung
orthogonale Matrix

$$\gamma : \mathbb{R}^n \to V. \tag{120}$$

Eine Isometrie $\phi : V \to V$ heißt *Drehung* , wenn $\det U = 1$.

Eine Matrix $U \in \mathrm{GL}_n(\mathbb{R})$ definiert genau dann eine Isometrie $\mathbb{R}^n \to \mathbb{R}^n$, wenn

$$< U\underline{x}, U\underline{y} >=< \underline{x}, \underline{y} >, \quad \text{für alle } \underline{x}, \underline{y} \in \mathbb{R}^n.$$

Diese Gleichung besagt, dass ${}^tUU = E_n$. Eine solche Matrix U heißt *orthogonale Matrix* . Eine Matrix $U \in \mathrm{GL}_n(\mathbb{R})$ ist genau dann orthogonal, wenn ihre Spalten die Länge 1 haben und paarweise orthogonal sind. Man kann also orthogonale Matrizen finden, in dem man das Gram-Schmidt-Verfahren anwendet.

Übung: Man finde eine orthogonale Matrix $U \in M(3\times 3, \mathbb{R})$, deren erste Spalte der folgende Vektor ist:

$$\begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$$

Es sei

$$U = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$$

eine orthogonale Matrix. Da die Länge der Spaltenvektoren 1 ist, folgt $a^2 + b^2 = 1$ und $c^2 + d^2 = 1$. Es gilt

$$\begin{pmatrix} c \\ d \end{pmatrix} \in \mathcal{L}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right)^{\perp} = \mathcal{L}\left(\begin{pmatrix} -b \\ a \end{pmatrix}\right).$$

Wenn man berücksichtigt, dass die Längen der auftretenden Vektoren 1 sind, so erhält man bei festen reellen Zahlen a und b mit $a^2 + b^2 = 1$ die folgenden Möglichkeiten für U:

$$U = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{oder} \quad U = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$$

Die zweiten Spalten dieser beiden Matrizen sind genau die beiden Vektoren der Länge 1, die zu ${}^t(a,b)$ orthogonal sind.

metrisch kongruent

Im ersten Fall gilt $\det U = 1$ und im zweiten Fall $\det U = -1$. Da $a^2 + b^2 = 1$ die Punkte auf (a,b) dem Einheitskreis beschreibt gibt es genau einen Winkel α, so dass $0^o \leq \alpha < 360^o$, so dass $a = \cos\alpha$ und $b = \sin\alpha$. Man definiert die Drehung $D(\alpha) : \mathbb{R}^2 \to \mathbb{R}^2$ um den Winkel α als die lineare Abbildung, die durch die Matrix

$$D(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}. \tag{121}$$

Aus den Additionstheoremen für die Winkelfunktionen ergibt sich, dass

$$D(\alpha) \circ D(\beta) = D(\alpha + \beta).$$

Im zweiten Fall gilt $U^2 = 1$. Daher hat U zwei Eigenvektoren

$$\begin{pmatrix} b \\ 1-a \end{pmatrix}, \quad \begin{pmatrix} 1-a \\ -b \end{pmatrix}$$

zu den Eigenwerten 1 bzw. -1. Wenn man diese Vektoren auf die Länge 1 normiert, erhält man eine orthonormale Basis von $\mathbb{R}^2$ in der

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

die Matrix von U ist.

1.14 Hauptachsentransformation und Spektralsatz

Auch in diesem Abschnitt geht es fast ausschließlich um $\mathbb{R}$-Vektorräume.

Es seien V und W zwei euklidische Vektorräume. Es sei $B_1 : V \times V \to \mathbb{R}$ und $B_2 : W \times W \to \mathbb{R}$ zwei Bilinearformen. Wir nehmen nicht an, dass B_i die positiv definiten Bilinearformen sind, die auf V und W vorgegeben sind, sondern die B_i sind ganz beliebig. Wir nennen B_1 und B_2 *metrisch kongruent* wenn eine Isometrie $\phi : V \to W$ von euklidischen Vektorräumen existiert, so dass

$$B_2(\phi(v_1), \phi(v_2)) = B_1(v_1, v_2), \quad \text{für alle } v_1, v_2 \in V. \tag{122}$$

Außerdem gilt für ϕ auch (119), weil ϕ eine Isometrie ist.

selbstadjundiert

Zwei quadratische Formen Q_1 auf V und Q_2 auf W heißen metrisch kongruent, wenn eine Isometrie $\phi : V \to W$ von euklidischen Vektorräumen existiert, so dass

$$Q_2(\phi(v)) = Q_1(v), \quad \text{für alle } v \in V.$$

Es ist natürlich das Gleiche zu sagen, dass die assoziierten symmetrischen Bilinearformen metrisch kongruent sind. Z.B. sind die quadratischen Formen auf dem $\mathbb{R}^2$:

$$Q_1(\underline{x}) = x_1^2 + x_2^2, \quad Q_2(\underline{x}) = \frac{x_1^2}{4} + \frac{x_2^2}{9},$$

zwar kongruent, aber nicht metrisch kongruent. Der Satz von Sylvester liefert die Klassifikation von quadratischen Formen bis auf Kongruenz (d.h. man vergisst die euklidische Struktur). Wir fragen jetzt nach der Klassifikation der quadratischen Formen bis auf metrische Kongruenz.

Es sei V ein euklidischer Vektorraum. Es sei Q_1 eine quadratische Form auf V und B_1 die assozierte symmetrische Bilinearform. Dann gibt es eine eindeutig bestimmte lineare Abbildung $\alpha_1 : V \to V$ so dass

$$B_1(v_1, v_2) =< \alpha_1(v_1), v_2 > \quad \text{für alle } v_1, v_2 \in V. \tag{123}$$

Da B_1 symmetrisch ist, gilt die Gleichung:

$$< \alpha_1(v_1), v_2 >=< v_1, \alpha_1(v_2) > \quad \text{für alle } v_1, v_2 \in V. \tag{124}$$

Ein Endomorphismus α_1 eines Euklidischen Vektorraums V, der (124) erfüllt, heißt *selbstadjundiert* . Umgekehrt definiert ein selbstadjundierter Endomorphismus α_1 durch die Formel (123) eine symmetrische Bilinearform oder äquivalent dazu eine quadratische Form.

Es sei W ein weiterer Euklidischer Vektorraum. Es sei B_2 eine symmetrische Bilinearform auf W und es sei $\alpha_2 : W \to W$ der assozierte selbsadjundierte Endomorphismus. Es sei $\phi : V \to W$ eine Isometrie. Dann ist die Gleichung (122) genau dann erfüllt, wenn folgendes Diagramm kommutativ ist

$$\begin{array}{ccc} V & \xrightarrow{\alpha_1} & V \\ {\scriptstyle\phi}\downarrow & & {\scriptstyle\phi}\downarrow \\ W & \xrightarrow{\alpha_2} & W. \end{array} \tag{125}$$

Man kann sagen, dass α_1 und α_2 *metrisch ähnlich* sind. In der Tat, angenommen die Gleichung (122) gilt. Dann finden wir für beliebige Vektoren $v \in V$ und $w \in W$:

metrisch
“”ahnlich

$$< \phi \circ \alpha_1(v), w >=< \alpha_1(v), \phi^{-1}(w) >= B_1(v, \phi^{-1}(w)) = B_2(\phi(v), w) \\ =< \alpha_2 \circ \phi(v), w > .$$

Die erste Gleichung gilt, weil ϕ eine Isometrie ist, und die dritte Gleichung gilt nach (122). Da die Gleichungen für alle $w \in W$ gelten, folgt die Kommutativität von (125). Genauso sieht man, dass diese Kommutativität die Gleichung (122) impliziert. Die Behauptung, dass B_1 und B_2 metrisch kongruent sind, ist also äquivalent zu der Behauptung, dass die Endomorphismen α_1 und α_2 metrisch ähnlich sind.

Lemma 106 *Ein Unterraum $W \subset V$ ist genau dann α_1-invariant, wenn*

$$B_1(u, w) = 0, \quad \textit{für alle } u \in W^\perp,\ w \in W. \tag{126}$$

In diesem Fall ist auch $W^\perp$ invariant bei α_1.

Beweis: Es sei $\alpha_1(W) \subset W$. Dann finden wir

$$B_1(u, w) =< u, \alpha_1(w) >= 0.$$

Umgekehrt sei (126) erfüllt. Es sei $w \in W$. Dann gilt

$$< u, \alpha_1(w) >= B_1(u, w) = 0, \quad \text{für alle } u \in W^\perp.$$

Also ist $\alpha_1(w) \in (W^\perp)^\perp = W$ (Satz 91).

Schließlich zeigen wir, dass $\alpha_1(W^\perp) \subset W^\perp$. Das folgt aus dem Bewiesenen, wenn wir in (126) die Rollen von W und $W^\perp$ vertauschen. *Q.E.D.*

Ein Vektor $v \in V$, $v \neq 0$ ist also genau dann ein Eigenvektor von α_1, wenn

$$B_1(u, v) = 0, \text{für alle } u \in \mathcal{L}(v)^\perp.$$

Wenn λ der Eigenwert ist, so gilt

$$B_1(v, v) =< \alpha_1(v), v >= \lambda < v, v > .$$

Übung: Es sei w ein weiterer Eigenvektor von α_1 zu einem Eigenwert $\mu \neq \lambda$. Man beweise, dass

$$< v_1, v_2 >= 0, \quad \text{und} \quad B_1(v_1, v_2) = 0.$$

Korollar 107 *Es sei $v_1, \dots v_n$ eine orthogonale Basis des Euklidischen Vektorraums V. Dann sind die folgenden Bedingungen äquivalent:*

1. *Es gilt $B_1(v_i, v_j) = 0$ für alle $i \neq j$.*
2. *Alle Vektoren $v_1, \dots, v_n$ sind Eigenvektoren für α_1.*

Der folgende Satz heißt *Spektralsatz* , wenn man an α_1 denkt und er heißt die *Hauptachsentransformation* , wenn man an B_1 denkt.

Satz 108 *Es sei V ein Euklidischer Vektorraum. Es sei B_1 eine symmetrische Bilinearform auf V und es sei α_1 der assoziierte selbstadjungdierte Endomorphismus.*

Dann gibt es eine orthogonale Basis $v_1, \dots v_n$ des Euklidischen Vektorraum V, die die äquivalenten Bedingungen von Korollar 107 erfüllt.

Beweis: Wir beweisen das durch Induktion nach n. Es genügt zu zeigen, dass ein selbstadjungdierter Endomorphismus stets einen Eigenvektor besitzt.

In der Tat, es sei v_1 ein Eigenvektor von α_1. Dann gilt

$$B_1(v_1, \mathcal{L}(v_1)^\perp) = 0.$$

Durch die Einschränkung von $< , >$ auf $\mathcal{L}(v_1)^\perp$ ist $\mathcal{L}(v_1)^\perp$ ein Euklidischer Vektorraum der Dimension $n-1$. Also gibt es nach Induktionsvoraussetzung eine orthogonale Basis $v_2, \dots v_n$ von $\mathcal{L}(v_1)^\perp$, so dass $B(v_i, v_j) = 0$ für $i \neq j$ und $2 \leq i, j$. Also ist $v_1, \dots v_n$ die gewünschte Basis.

Es bleibt zu zeigen, dass es einen Eigenvektor gibt. Wir wählen eine Isometrie $\gamma : \mathbb{R}^n \to V$. Wir definieren die Matrix A durch das kommutative Diagramm

$$\begin{array}{ccc} V & \xrightarrow{\alpha_1} & V \\ {\scriptstyle\gamma}\uparrow & & {\scriptstyle\gamma}\uparrow \\ \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^n. \end{array}$$

Dann ist A selbstadjungdiert bezüglich des Standardskalarprodukts auf $\mathbb{R}^n$. Folglich ist $A \in M(n \times n, \mathbb{R})$ eine symmetrische Matrix, d.h. ${}^tA = A$. Wir wollen zeigen, dass es einen Eigenvektor $v \in \mathbb{R}^n$ von A gibt. Dazu fassen wir A als Matrix über den komplexen Zahlen auf: $A : \mathbb{C}^n \to \mathbb{C}^n$. Da das

charakteristische Polynom von A eine Nullstelle $\mu \in \mathbb{C}$ hat, gibt es einen Eigenvektor $\underline{z} \in \mathbb{C}^n$, so dass

$$A\underline{z} = \mu\underline{z}.$$

Wir multiplizieren diese Gleichung von links mit ${}^t\bar{\underline{z}}$. Der obere Querstrich bedeutet, dass wir die Einträge von $\underline{z}$ durch die komplex konjugierten Zahlen ersetzt haben. Wir erhalten:

$${}^t\bar{\underline{z}}A\underline{z} = \mu\ {}^t\bar{\underline{z}}\underline{z}. \tag{127}$$

Man sieht leicht, dass sowohl ${}^t\bar{\underline{z}}\underline{z}$ als auch ${}^t\bar{\underline{z}}A\underline{z}$ reelle Zahlen sind. In der Tat der letzte Ausdruck verändert sich nicht, wenn man in transponiert:

$${}^t\bar{\underline{z}}A\underline{z} = {}^t\underline{z}\,{}^tA\bar{\underline{z}} = {}^t\underline{z}\bar{A}\bar{\underline{z}}.$$

Die letzte Gleichung gilt, da $A = \bar{A}$ und $A = {}^tA$. Also ist die linke Seite von (127) gleich ihrem komplex Konjugierten und daher reell.

Dann ist aber μ eine reelle Zahl. Da $\det(\mu E_n - A) = 0$ muss es dann auch einen reellen Eigenvektor $v \in \mathbb{R}^n$ geben. $Q.E.D.$

1.15 Drehungen eines euklidischen Raumes

Satz 109 *Es sei V ein Euklidischer Vektorraum. Es sei $f : V \to V$ eine Drehung von V. Dann gibt es eine orthonormale Basis von V, in der die Matrix f die folgende Gestalt hat:*

$$\begin{pmatrix} 1 & & & & & & & & & \\ & 1 & & & & & & & & \\ & & \ldots & & & & & & & \\ & & & 1 & & & & & & \\ & & & & D(\alpha_1) & & & & & \\ & & & & & D(\alpha_2) & & & & \\ & & & & & & D(\alpha_3) & & & \\ & & & & & & & \ldots & & \\ & & & & & & & & D(\alpha_k) & \end{pmatrix} \tag{128}$$

Hier sind $\alpha_1, \ldots, \alpha_k$ Winkel aus dem offenen Intervall $(0, 360)$. Die Anzahl der Einsen in der Matrix und die Winkel bis auf ihre Reihenfolge, sind durch f eindeutig bestimmt. Die $D(\alpha_i) \in M(2 \times 2, \mathbb{R})$ sind die Matrizen aus (121). Bis auf die Einsen und die $D(\alpha_i)$ stehen in der Matrix Nullen.

Wir beweisen zunächst:

Lemma 110 *Es sei V ein endlich erzeugter $\mathbb{R}$-Vektorraum. Es sei $f : V \to V$ ein Endomorphismus. Dann gibt es einen f-invarianten Unterraum $W \subset V$, so dass $W \neq 0$ und $\dim W \leq 2$.*

Beweis: Wenn f einen Eigenvektor besitzt, so ist das Lemma klar. Wir setzen deshalb voraus, dass das charakteristische Polynom von f keine reellen Nullstellen besitzt.

Im allgemeinen können wir ein Koordinatensystem wählen und annehmen, dass $V = \mathbb{R}^n$. Dann ist f durch eine Matrix $A \in M(n \times n, \mathbb{R})$ gegeben.

Wir fassen wie im letzten Paragraphen A als Endomorphismus von $\mathbb{C}^n$ auf. Dann hat A einen Eigenvektor $\underline{z} \in \mathbb{C}^n$:

$$A\underline{z} = \mu\underline{z}, \quad \mu \in \mathbb{C}.$$

Nach der Bemerkung am Anfang ist μ keine reelle Zahl. Wenn wir auf die letzte Gleichung die komplexe Konjugation anwenden und berücksichtigen, dass $\bar{A} = A$, so erhalten wir

$$A\bar{\underline{z}} = \bar{\mu}\bar{\underline{z}}.$$

Da $\bar{\mu} \neq \mu$, sind die beiden Vektoren $\underline{z}$ und $\bar{\underline{z}}$ linear unabhängig über $\mathbb{C}$. Das folgt, weil die Vektoren zu verschiedenen Haupträumen von A gehören. Es ist aber eine leichte Übung das direkt zu verifizieren.

Es sei $\tilde{W} = \mathcal{L}_{\mathbb{C}}(\underline{z}, \bar{\underline{z}}) \subset \mathbb{C}^n$. Das ist eine zwei dimensionaler A-invarianter Unterraum. Eine weitere Basis von $\tilde{W}$ sind die Vektoren

$$w_1 := \underline{z} + \bar{\underline{z}}, \quad w_2 := \sqrt{-1}(\underline{z} - \bar{\underline{z}}).$$

Diese Vektoren liegen in $\mathbb{R}^n \subset \mathbb{C}^n$. Sie bilden eine Basis des A-invarianten reellen Vektorraums $W = \tilde{W} \cap \mathbb{R}^n$.

In der Tat, es sei $w \in W$. Da w_1, w_2 eine Basis von $\tilde{W}$ sind, können wir schreiben

$$w = \lambda_1 w_1 + \lambda_2 w_2, \quad \text{wo, } \lambda_i \in \mathbb{C}. \tag{129}$$

Alle Vektoren liegen in $\mathbb{C}^n$. Deshalb können wir auf die letzte Gleichung die komplexe Konjugation anwenden. Es gilt $\bar{w} = w$, $\bar{w}_1 = w_1$ $\bar{w}_2 = w_2$, denn diese Vektoren liegen in $\mathbb{R}^n$. Also folgt aus (129)

$$w = \bar{\lambda_1} w_1 + \bar{\lambda_2} w_2.$$

Aber dann muss $\bar{\lambda_i} = \lambda_i$ gelten, d.h. die Zahlen λ_1 und λ_2 sind reell.

Damit ist $W = \mathcal{L}_{\mathbb{R}}(w_1, w_2)$. Da W auch A-invariant ist, sind wir am Ziel. *Q.E.D.*

Jetzt kommen wir zum Beweis des Satzes:

Wir beweisen zunächst eine etwas andere Aussage:

(*) Es sei $f : V \to V$ eine Isometrie. Dann gibt es eine orthonormale Basis von V, in der die Matrix von f die Gestalt (128) hat, wobei wir die Einsen durch ± 1 ersetzen.

Wir beweisen das durch Induktion nach $\dim V$. Für den Induktionsschritt benutzen wir folgende Tatsache. Es sei $W \subset V$ ein f-invarianter Unterraum. Dann ist auch $W^{\perp}$ ein f-invarianter Unterraum. In der Tat, es sei $u \in W^{\perp}$. Wir müssen zeigen, dass

$$< f(u), w >= 0, \text{für alle } w \in W.$$

Es ist klar, dass die Einschränkung von f auf W ein Isomorphismus ist. Deshalb gilt $w = f(w')$ für ein $w' \in W$. Wir finden

$$< f(u), w >=< f(u), f(w') >=< u, w' >= 0,$$

da f eine Isometrie ist. Also gilt $f(W^{\perp}) \subset W^{\perp}$.

Angenommen f hat einen reellen Eigenwert λ, so gilt $\lambda = \pm 1$. Es sei nämlich $v \in V$ ein Eigenvektor:

$$f(v) = \lambda v.$$

Da die Längen der Vektoren auf beiden Seiten der Gleichung übereinstimmen, sieht man $|\lambda| = 1$.

Es sei $W = \mathcal{L}(v)$. Wir wenden die Induktionsvoraussetzung auf die Einschränkung von f auf $W^{\perp}$ an und haben die Aussage (*) bewiesen.

Wenn f keinen reellen Eigenvektor besitzt, so nehmen wir einen zweidimensionalen invarianten Unterraum $W \subset V$. Wir schreiben die Matrix der Einschränkung von f auf W in einer orthonormalen Basis von W. Da es keine reellen Eigenwerte gibt kommt nur eine Matrix der Form $D(\alpha)$ (vgl. (121) infrage.

Jetzt brauchen wir nur noch die Induktionsvoraussetzung auf das orthogonale Komplement von W anzuwenden und haben (*) vollständig bewiesen.

Jetzt sei f eine Drehung. Wenn wir die Determinante der in (*) beschriebenen Matrix nehmen muss 1 herauskommen. Also ist die Anzahl der -1 auf der Diagonale gerade. Wir fassen immer zwei zu einem Kästchen zusammen

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

Aber diese Matrix ist die Drehung $D(180^o)$. Damit ist bewiesen, dass eine othonormale Basis existiert, wie in Satz 109 behauptet wird.

Wir müssen noch zeigen, dass die α_j bis auf Permutation eindeutig bestimmt sind. Das folgt, wenn man das charakteristische Polynom der Matrix ausrechnet:

$$\chi(f,\lambda) = (\lambda-1)^\ell \prod_{j=1}^{k} (\lambda - e^{i\alpha_j})(\lambda - e^{-i\alpha_j})$$

Hier ist $i = \sqrt{-1}$ und ℓ ist die Anzahl der Einsen in (128). *Q.E.D.*

Es sei $A \in M(3\times 3, \mathbb{R})$ eine orthogonale Matrix, so dass $\det A = 1$. Nach unser Definition ist A eine Drehung des $\mathbb{R}^3$. Nach dem Satz 109 gibt es eine orthogonale Matrix $U \in M(3\times 3, \mathbb{R})$, so dass

$$^tUAU = U^{-1}AU = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}. \tag{130}$$

Die Matrix auf der rechten Seite ist also sowohl ähnlich als auch kongruent zu A. Es sei $A \neq E_3$. Wir sehen aus (130), dass der Eigenraum zum Eigenwert 1 von A die Dimension 1 hat. Man nennt den Eigenraum die Drehachse von A.

Übung: Es sein $A = (a_{ij})$. Wenn $A^2 \neq 1$, so wird die Drehachse von A von dem folgenden Vektor ezeugt:

$$\begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix}.$$

Hinweis: Unter Benutzung des Vektorprodukts zeigt man z.B. folgende Gleichung: $a_{32} = -a_{11}a_{23} + a_{13}a_{21}$.

Den Drehwinkel α, $0 \leq \alpha < 360^o$ nennen wir den Drehwinkel der Drehung A. Es ist leicht seinen Kosinus zu berechnen

$$\cos\alpha = \frac{1}{2}(\text{Spur}\, A - 1).$$

Diese Formel folgt unmittelbar aus (130).

1.16 Orientierung, Vektorprodukt

Es sei V ein Euklidischer Vektorraum der Dimension n. Es sei $\phi : V \to V$ eine Isometrie. Wenn wir auf V ein kartesisches Koordinatensystem wählen, ist die Matrix von ϕ eine orthogonale Matrix. Folglich gilt

$$\det \phi = \pm 1.$$

Definition 111 *Ein Orientierung auf V ist eine Determinantenfunktion $\vartheta \in \mathrm{Alt}^n(V, \mathbb{R})$, so dass für jede orthonormale Basis $v_1, \ldots, v_n$ von V gilt, dass*

$$\vartheta(v_1, \ldots, v_n) = \pm 1. \tag{131}$$

Wenn wir eine orthonormale Basis $w_1, \ldots, w_n$ wählen gibt es genau eine Determinantenfunktion ϑ, so dass $\vartheta(w_1, \ldots, w_n) = 1$. Jede andere orthonormale Basis von V hat die Form $\phi(w_1), \ldots, \phi(w_n)$, wobei ϕ eine Isometrie ist. Daher erfüllt ϑ die Bedingung (131). Wir sehen daraus, dass es auf einem Euklidischen Vektorraum genau zwei Orientierungen gibt. Drehungen erhalten die Orientierung aber Spieglungen nicht.

Wenn eine Orientierung gegeben ist, so nennen wir eine orthonormale Basis $v_1, \ldots, v_n$ positiv orientiert, wenn $\vartheta(v_1, \ldots, v_n) = 1$ und sonst negativ orientiert.

Der folgende Sachverhalt liegt auf der Hand, aber es lohnt sich ihn nochmal zu formulieren:

Satz 112 *Es sei V ein Euklidischer Vektorraum mit einer Orientierung ϑ. Es sei $v_1, \ldots, v_n$ eine positiv orientierte orthonormale Basis von V.*

Eine lineare Abbildung $f : V \to V$ ist genau dann eine Drehung, wenn $f(v_1), \ldots, f(v_n)$ eine positiv orientierte orthonormale Basis von V.

Es sei umgekehrt $w_1, \ldots, w_n$ eine positiv orientierte orthonormale Basis von V. Dann gibt es genau eine Drehung $f : V \to V$, so dass $f(v_i) = w_i$.

Es sei V ein dreidimensionaler Vektorraum. Wir wählen eine Orientierung ϑ auf V. Wir definieren eine bilineare Abbildung

$$\begin{array}{ccc} V \times V & \to & V \\ (v, w) & \mapsto & v \times w \end{array}$$

Der Vektor $v \times w \in V$ ist durch folgende Bedingung bestimmt:
Für alle Vektoren $x \in V$ gilt:

$$< v \times w, x >= \vartheta(v, w, x). \tag{132}$$

Die Existenz von $V \times W$ ergibt sich aus Satz 83. Man beweist leicht, dass $v \times w$ in den beiden Argumenten v und w linear ist. Es gilt

$$v \times w = -w \times v, \quad \text{und insbesondere } v \times v = 0.$$

Es sei v_1, v_2, v_3 eine positiv orientierte orthonormale Basis. Dann gilt

$$v_1 \times v_2 = v_3, \quad v_2 \times v_3 = v_1, \quad v_3 \times v_1 = v_2. \tag{133}$$

In der Tat, wir verifizieren die erste Gleichung. Nach (132) muss man beweisen, dass für alle $x \in V$:

$$< v_3, x >= \det(v_1, v_2, x).$$

Es genügt die Gleichung für $x = v_1$, $x = v_2$ und $x = v_3$ zu verifizieren. Das ist klar.

Übung: Man beweise, dass das Vektorprodukt nicht assoziativ ist.

Satz 113 *Es sei V ein euklidischer Vektorraum der Dimension 3. Es seien $v, w, z \in V$. Dann gilt*

$$(v \times w) \times z =< v, z > w - < w, z > v.$$

Beweis: Beide Seiten der Gleichung sind multilineare Abbildungen $V \times V \times V \to V$. Deshalb reicht es nach Satz 67, zu zeigen, dass beide Abbildungen auf einer Basis von V übereinstimmen. Wir wählen eine orthonormale Basis u_1, u_2, u_3 von V. Es reicht die Gleichung in dem Fall zu verifizieren, wo $v, w, z \in \{u_1, u_2, u_3\}$. In diesen Fällen folgt die behauptete Gleichung aus (133). *Q.E.D.*

Korollar 114 *Es gilt:*

$$< v \times w, v \times w >=< v, v >< w, w > - < v, w >^2 .$$

Beweis: Nach Definition gilt

$$< v \times w, v \times w >= \vartheta(v, w, v \times w) = \vartheta(v \times w, v, w) =< (v \times w) \times v, w > \\ =< v, v >< w, w > - < w, v >< v, w > .$$

Die letzte Gleichung folgt aus dem Satz. *Q.E.D.*

Es sei wieder V ein 3-dimensionaler Euklidischer Raum. Es sei ϑ eine Orientierung.

Es sei d ein ein Vektor der Länge 1. Wir betrachten eine Drehung $\phi : V \to V$ um die Achse d im positiven Drehsinn um der einen Winkel α. Der positive Drehsinn ist so definiert. Der Unterraum $W = \mathcal{L}(d)^{\perp}$ erbt eine Orientierung: Eine orthonormale Basis u_1, u_2 von W heißt positiv orientiert, wenn

$$\vartheta(d, u_1, u_2) = 1.$$

Die positive Drehung um d um den Winkel α lässt W invariant, und ist in der Basis u_1, u_2 durch die Matrix (121) gegeben. Die Matrix von ϕ in der Basis d, u_1, u_2 ist

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha. \end{pmatrix}.$$

Es sei $v \in V$. Die folgende Formel für $\phi(v)$ geht auf Euler zurück:

$$\phi(v) = \cos\alpha\, v + < d, v > (1 - \cos\alpha)d + \sin\alpha(d \times v). \tag{134}$$

Die rechte Seite der Gleichung ist eine lineare Abbildung in v. Daher braucht man die Gleichung nur für $v \in \{d, u_1, u_2\}$ zu verifizieren. Aber die Bilder der Vektoren d, u_1, u_2 bei ϕ können wir aus der obigen Matrix ablesen. Wir erhalten ganz leicht die Eulersche Formel.

Etwas geometrischer kann man so argumentieren. Wir betrachten die folgenden orthogonalen Zerlegungen

$$v = w + \lambda d, \quad \phi(v) =: v' = w' + \lambda d,$$

wobei $w' = \phi(w)$. Es gilt $\lambda =< v, d >=< v', d >$. Die Basis $d, w, d \times w$ ist eine positiv orientierte orthogonale Basis. Die Vektoren w und $d \times w$ haben zwar nicht die Länge 1, aber sie haben die gleiche Länge. Daher gilt:

$$w' = \phi(w) = \cos\alpha\, w + \sin\alpha\, (d \times w).$$

Daraus folgt:

$$v' = w' + \lambda d = \cos\alpha\,(v - \lambda d) + \sin\alpha\,(d \times (v - \lambda d)) + \lambda d.$$

Wenn man für $\lambda =< d, v >$ einsetzt, hat man die Eulersche Formel.

Wir betrachten den Fall, wo $V = \mathbb{R}^3$ mit dem Standardskalarprodukt. Man setzt $q_0 = \cos(\alpha/2)$ und $\underline{q} = \sin(\alpha/2)\underline{d}$. Es sei:

$$\underline{q} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$$

Man findet $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$. Aus der Formel (134) ergibt sich die Matrix von ϕ (*Eulersche Drehmatrix*):

$$\begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}$$

1.17 Aufstieg und Abstieg zu anderen Körpern

Es sei $\tau : K \to L$ ein Homomorphismus von Körpern. Nach Definition gilt $\tau(1) = 1$. Daraus folgt, dass τ injektiv ist. In der Tat, angenommen es wäre $\tau(k) = 0$, wo $k \in K$ und $k \neq 0$. Dann folgt $\tau(1) = \tau(kk^{-1}) = \tau(k)\tau(k^{-1}) = 0$. Das wäre ein Widerspruch.

Wenn man τ komponentenweise auf Spaltenvektoren anwendet, so erhält man für jede natürliche Zahl n eine Abbildung

$$\tau : K^n \to L^n. \tag{135}$$

Das ist ein injektiver Homomorphismus von abelschen Gruppen.

Lemma 115 *Es seien $v_1, \ldots, v_d \in K^n$ Vektoren. Dann sind $v_1, \ldots, v_d$ linear unabhängig über K genau dann wenn ihre Bilder $\tau(v_1), \ldots, \tau(v_d)$ linear unabhägig über L sind.*

Beweis: Wir setzen voraus, dass $v_1, \ldots, v_d$ linear unabhängig über K sind und zeigen, dass $\tau(v_1), \ldots, \tau(v_d)$ linear unabhängig über L sind. Alles andere ist klar.

Wir ergänzen $v_1, \ldots, v_d$ zu einer Basis $v_1, \ldots v_d, v_{d+1}, \ldots, v_n$ von K^n. Es genügt zu zeigen, dass $\tau(v_1), \ldots, \tau(v_n)$ eine Basis von L^n ist. Aber $\mathcal{L}_K(v_1, \ldots, v_n)$ enthält alle Standardvektoren e_i und daher enthält $\mathcal{L}_L(\tau(v_1), \ldots, \tau(v_n))$ alle Standardvektoren $\tau(e_i)$ von L^n. *Q.E.D.*

Einen anderen Beweis erhält man, indem man die Matrix $(v_1, \ldots, v_d)$ mit Zeilenoperationen auf Stufenform bringt. Man kann das Lemma auch so formulieren:

Korollar 116 *Es sei $A \in M(m \times n, K)$. Dann gilt*

$$\operatorname{Rang} A = \operatorname{Rang} \tau(A).$$

Wenn $A = (a_{ij})$, so ist nach Definition $\tau(A) = (\tau(a_{ij}))$.

Wir wollen im folgenden zur Vereinfachung der Notation annehmen, dass τ eine Inklusion ist $K \subset L$. Wir betrachten den Lösungsraum des Gleichungssystems $Ax = 0$, wo $A \in M(m \times n, K)$. Uns interessieren nicht nur Lösungen in K^n sondern Lösungen in L^n:

$$V = \{z \in L^n \mid Az = 0\}.$$

Nach dem Korollar hat der L-Vektorraum V eine Basis von Vektoren in K^n. Denn es gibt über K linear unabhängige Vektoren $\underline{x}_1, \ldots, \underline{x}_d \in K^n$, wobei $d = n - \operatorname{Rang} A = \dim V$, so dass $A\underline{x}_i = 0$.

Wir betrachten jetzt für τ die Inklusion $\mathbb{R}^n \subset \mathbb{C}^n$. Wir definieren eine Homomorphismus abelscher Gruppe $\iota : \mathbb{C}^n \to \mathbb{C}^n$:

$$\iota\left(\begin{pmatrix} z_1 \\ z_2 \\ \ldots \\ z_n \end{pmatrix}\right) = \begin{pmatrix} \bar{z_1} \\ \bar{z_2} \\ \ldots \\ \bar{z_n} \end{pmatrix}$$

Hier bezeichnet $\bar{z_i}$ die komplex konjugierte Zahl zu $z_i \in \mathbb{C}$. Der Homomorphismus ι ist nicht $\mathbb{C}$-linear, sondern es gilt:

$$\iota(\lambda v) = \bar{\lambda}\iota(v), \quad \lambda \in \mathcal{C},\ v \in \mathbb{C}^n.$$

Wir schreiben auch $\iota(v) = \bar{v}$ für $v \in \mathbb{C}^n$. Es gilt $v \in \mathbb{R}^n$ genau dann wenn $\iota(v) = v$.

Satz 117 *Es sei $V \subset \mathbb{C}^n$ ein $\mathbb{C}$-Untervektorraum, so dass $\iota(V) \subset V$.*

Dann hat V eine Basis aus Vektoren in $V \cap \mathbb{R}^n$.

Die Zuordnung $V \mapsto V_0 = V \cap \mathbb{R}^n$ ist ein Bijektion zwischen $\mathbb{C}$-Untervektorräumen V von $\mathbb{C}^n$ mit $\iota(V) \subset V$ und $\mathbb{R}$-Untervektorräumen V_0 von $\mathbb{R}^n$. Es gilt $V = \mathcal{L}_{\mathbb{C}}(V_0)$.

Beweis: Wir zeigen nur, dass V die gewünschte Basis hat. Der Rest ist dann einfach. Es sei $d = \dim V$. Wir finden Standardvektoren $e_{i_{d+1}}, \ldots, e_{i_n}$, so dass $\mathcal{L}_{\mathbb{C}}(e_{i_{d+1}}, \ldots, e_{i_n})$ ein Komplementärraum von V ist:

$$\mathbb{C}^n = V \oplus \mathcal{L}_{\mathbb{C}}(e_{i_{d+1}}, \ldots, e_{i_n}) \tag{136}$$

Es seien $e_{i_1}, \ldots, e_{i_d}$ die restlichen der n Standardvektoren. Dann finden wir eine Relationen

$$e_{i_k} = u_i + \sum_{l=d+1}^{n} \lambda_{lk} e_{i_l}, \qquad \text{wo } u_i \in V,\ \lambda_{lk} \in \mathbb{C},\ k = 1, \ldots d. \tag{137}$$

Wenn man auf diese Gleichungen ι anwendet, so folgt

$$e_{i_k} = \iota(u_i) + \sum_{l=d+1}^{n} \bar{\lambda}_{lk} e_{i_l}, \qquad \text{wo } \ k = 1, \ldots d. \tag{138}$$

Nach Voraussetzung gilt $\iota(u_i) \in V$. Wenn man die beiden Gleichungen voneinander abzieht, so erhält man

$$\sum_{l=d+1}^{n} (\lambda_{lk} - \bar{\lambda}_{lk}) e_{i_l} \in V, \quad \text{für } k = 1, \ldots, d.$$

Das impliziert nach (136), dass $\lambda_{lk} - \bar{\lambda}_{lk} = 0$ für alle $k = 1, \ldots, d$ und $l = d+1, \ldots, n$. Also sind alle Koeffizienten λ_{lk} reelle Zahlen. Daraus folgt

$$e_{i_k} - \sum_{l=d+1}^{n} \lambda_{lk} e_{i_l} = u_i \in V \cap \mathbb{R}^n.$$

Da diese Elemente über $\mathbb{C}$ linear unabhängig sind, bilden sie eine Basis von V. *Q.E.D.*

Bemerkung: Der gleiche Beweis funktioniert in einer viel allgemeineren Situation. Es seien L und E Körper. Es sei $\tau_i : L \to E$, wo $i \in I$, eine Menge von Körperhomomorphismen. Es sei

$$K = \{k \in L \mid \tau_i(k) = \tau_j(k), \text{ für alle } i, j \in I\}.$$

Dann ist K ein Teilköper von L.

Die τ_i setzen wir auf die Spaltenvektoren fort (135):

$$\tau_i : L^n \to E^n.$$

Dann gilt: Es sei $V \subset L^n$ ein K-Untervektorraum, so dass $\mathcal{L}_E(\tau_i(V)) = \mathcal{L}_E(\tau_j(V))$ für alle $i, j \in I$. Dann hat V eine Basis aus Vektoren in $V \cap K^n$.

Übung: Es sei K ein Körper der Charakteristik p. Dann hat man einen Homomorphismus von Körpern $\sigma : K \to K$, wo $\sigma(\lambda) = \lambda^p$.

Wir betrachten den Teilkörper $\mathbb{F}_p \subset K$:

$$\mathbb{F}_p = \{\mu \in K \mid \mu^p = \mu\}.$$

Wir fixieren eine natürliche Zahl n. Es sei $\underline{x} \in K^n$. Wir schreiben $\sigma(\underline{x}) \in K^n$ für den Vektor mit den Einträgen $\sigma(x_i) = x_i^p$.

Es sei $W \subset K^n$ ein Untervektorraum, so dass $\sigma(W) \subset W$. Man beweise, dass es eine Basis $v_1, \ldots, v_d$ von W gibt, so dass

$$v_1, \ldots, v_d \in \mathbb{F}_p^n \subset K^n.$$

1.18 Hermitesche Formen

Es sei V ein $\mathbb{C}$-Vektorraum (komplexer Vektorraum).

Definition 118 *Eine Funktion $H : V \times V \to \mathbb{C}$ heißt hermitesch, wenn sie $\mathbb{R}$-bilinear ist, und wenn gilt:*

$$\begin{aligned} H(\lambda v, w) &= \lambda H(v, w) \\ H(v, \lambda w) &= \bar{\lambda} H(v, w) \\ H(w, v) &= \overline{H(v, w)}, \end{aligned}$$

für alle $v, w \in V$ und $\lambda \in \mathbb{C}$.

selbstadjungdiert

Die Querstriche bezeichnen die komplexe Konjugation auf $\mathbb{C}$. Wichtig ist, dass $H(v,v) \in \mathbb{R}$. Insbesondere gibt es positiv definite hermitesche Formen. Der Satz von Sylvester, die Schmidt-Orthogonalisierung und die Dimensionsformeln für orthogonale Komplemente gelten.

Es sei V ein komplexer Vektorraum. Wir definieren einen komplexen Vektorraum $\bar{V}$ wie folgt. Wenn wir V und $\bar{V}$ nur als abelsche Gruppen ansehen, so gilt $V = \bar{V}$. Wir definieren die Multiplikation

$$\begin{array}{rccc} * : & \mathbb{C} \times \bar{V} & \to & \bar{V} \\ & \lambda \times v & \mapsto & \bar{\lambda} v. \end{array}$$

Mit anderen Worten gilt für die neue Multiplikation $*$:

$$\lambda * v = \bar{\lambda}(v).$$

Es sei W ein weiterer komplexer Vektorraum. Eine Abbildung $\phi : V \to W$ heißt antilinear, wenn

$$\begin{array}{c} \phi(v_1 + v_2) = \phi(v_1) + \phi(v_2) \\ \phi(\lambda v) = \bar{\lambda} v. \end{array}$$

Wenn man ϕ als eine Abbildung von $\bar{V} \to W$ auffasst, so ist es eine $\mathbb{C}$-lineare Abbildung. Genauso hat man eine $\mathbb{C}$-lineare Abbildung $V \to \bar{W}$.

Eine Hermitesche Form H definiert ein Bilinearform von $\mathbb{C}$-Vektorräumen

$$B : V \times \bar{V} \to \mathbb{C},$$

wobei $B(v,w) := H(v,w)$. In dem man von H zu B übergeht, können wir Tatsachen die wir für Bilinearformen kennen, bedenkenlos auch auf H anwenden.

Ein komplexer Vektorraum V mit einer positiv definiten hermiteschen Form H heißt ein *unitärer Vektorraum.*

Es sei (V,H) ein unitärer Vektorraum. Es sei $\phi : V \to V$ eine $\mathbb{C}$-lineare Abbildung. Dann gibt es genau eine lineare Abbildung $\tilde{(\phi)}$, so dass

$$< \phi(v_1), v_2 > = < v_1, \tilde{\phi}(v_2) > .$$

Man nennt $\tilde{\phi} : V \to V$ die zu ϕ adjungdierte lineare Abbildung. Man nennt ϕ *selbstadjungdiert* , wenn $\phi = \tilde{\phi}$.

Definition 119 *Eine lineare Abbildung $\phi : V \to V$ heißt normal, wenn*

$$\phi \circ \tilde{\phi} = \tilde{\phi} \circ \phi$$

gilt

Satz 120 *(zweite Form des Spektralsatzes) Es sei (V, H) ein unitärer Vektorraum. Es sei $\phi : V \to V$ eine normale lineare Abbildung.*

Dann besitzt (V, H) eine orthonormale Basis, die aus Eigenvektoren von ϕ besteht.

Beispiel: Es seien V, H wie im Satz. Eine *Isometrie* ist eine lineare Abbildung $f : V \to V$, so dass

$$H(f(v), f(w)) = H(v, w).$$

Eine Isometrie ist offensichtlich normal und ihre Eigenwerte sind komplexe Zahlen vom Absolutbetrag 1.

Satz 121 *(Cauchy-Schwarz) ([**F**] S.293, 275) Es sei (V, H) unitär (oder euklidisch). Dann gilt:*

$$|H(v, w)| \leq |H(v, v)|^{\frac{1}{2}} |H(w, w)|^{\frac{1}{2}}$$

1.19 Drehungen des Raumes und unitäre Matrizen

Es sei

$$S^2 = \{\underline{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}.$$

Es sei $M \subset \mathbb{R}^3$ der Untervektorraum, der durch $x_3 = 0$ definiert wird. Es sei $\underline{n} = (0, 0, 1) \in \mathbb{R}^3$.

Für einen Punkt $\underline{x} \in S^2$, wo $\underline{x} \neq \underline{n}$ definiert man $\pi(\underline{x})$ als den Durchschnitt der affinen Geraden $(1 - t)\underline{n} + t\underline{x}$, $t \in \mathbb{R}$ mit M. Es gilt

$$\pi(\underline{x}) = \frac{1}{1 - x_3}(x_1, x_2).$$

$\pi : S^2 \setminus \{\underline{n}\} \to M$ heißt stereographische Projektion.

Wir fassen die Punkte von M als komplexe Zahlen auf $\xi = x_1 + \mathbf{i}x_2 \in \mathbb{C}$. In diesem Sinne schreiben wir einen Punkt von $\mathbb{R}^3$ in der Form (ξ, x_3). Zu

der Menge $\mathbb{C}$ fügen wir ein weiteres Element hinzu, dass ∞ heißt. Dann schreiben wir die streographische Projektion

$$\pi : S^2 \to \mathbb{C} \cup \{\infty\}, \quad \pi(\underline{x}) = \tfrac{\xi}{1-x_3}, \text{ wo } x_3 \neq 1, \quad \pi(\underline{n}) = \infty. \tag{139}$$

Diese Abbildung ist eine Bijektion.

Man findet für die Umkehrabbildung ρ die folgende Formel. Es sei $z \in \mathbb{C}$ und $(\xi, x_3) = \pi^{-1}(z)$. Dann gilt:

$$\xi = \frac{2z}{|z|^2+1}, \quad x_3 = \frac{|z|^2-1}{|z|^2+1}. \tag{140}$$

Wir bezeichen mit $\mathbb{P}$, die Menge aller eindimensionalen Unterräume des $\mathbb{C}$-Vektorraums $\mathbb{C}^2$. Man nennt $\mathbb{P}$ die projektive Gerade über dem Körper $\mathbb{C}$.

Wir bezeichnen den eindimensionalen Unterraum von $\mathbb{C}^2$, der von dem Vektor

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \neq 0, \quad z_1, z_2 \in \mathbb{C}$$

erzeugt wird mit $(z_1 : z_2)$. Wenn $\lambda \in \mathbb{C}$, $\lambda \neq 0$, so gilt

$$(z_1 : z_2) = (\lambda z_1 : \lambda z_2).$$

Mit den Bezeichungen von (140) gilt beispielsweise, wenn $x_3 \neq \pm 1$

$$(z : 1) = (\xi : (1 - x_3)) = (\xi\bar{\xi} : (1 - x_3)\bar{\xi}) = ((1 + x_3) : \bar{\xi}).$$

Wir definieren die Bijektion $\mathbb{C} \cup \{\infty\} \to \mathbb{P}$ durch $z \mapsto (z : 1)$ und $\infty \mapsto (1 : 0)$.

Die Umkehrabbildung zur streographischen Projektion gibt uns dann

$$\rho : \mathbb{P} \to S^2. \tag{141}$$

Das Bild $\rho((z_1 : z_2)) = (\xi, x_3) \in S^2$ ist durch die folgende Formel gegeben:

$$\xi = \frac{2z_1\bar{z}_2}{|z_1|^2 + |z_2|^2}, \quad x_3 = \frac{|z_1|^2 - |z_2|^2}{|z_1|^2 + |z_2|^2}. \tag{142}$$

Hier bezeichnet $\bar{z}_2$ die zu z_2 komplex konjugierte Zahl. In dieser Formel spielt der Punkt ∞ keine Sonderrolle mehr.

Für zwei komplexe Zahlen α, β, so dass $|\alpha|^2 + |\beta|^2 = 1$ betrachten wir die Matrix

$$U = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in GL_2(\mathbb{C}).$$

Alle diese Matrizen bilden eine Untergruppe $SU_2 \subset GL_2(\mathbb{C})$.

Wir betrachten auf $\mathbb{C}^2$ die hermitesche Form $H(\underline{v}, \underline{w}) = {}^t\bar{\underline{v}}\underline{w}$. Dann gilt für $U \in SU_2$:

$$H(U(\underline{v}), U(\underline{w})) = H(\underline{v}, \underline{w}).$$

Es sei $L \subset \mathbb{C}^2$ ein eindimensionaler Unterraum, d.h. $L \in \mathbb{P}$. Wir bezeichnen mit $U(L)$ sein Bild bei der linearen Abbildungen $U : \mathbb{C}^2 \to \mathbb{C}^2$. Die Abbildung $\hat{U} : L \mapsto U(L)$ ist ein Automorphismus (=Permutation) der Menge $\mathbb{P}$. Wir erhalten einen Gruppenhomomorphismus

$$SU_2 \to \mathrm{Aut}(\mathbb{P}).$$

Statt von einem Gruppenhomomorphismus spricht man auch von einer Operation von SU_2 auf $\mathbb{P}$.

Wegen der Bijektion ρ (141) erhalten wir auch eine Operation von SU_2 auf S^2:

$$SU_2 \to \mathrm{Aut}(S^2), \quad U \mapsto T. \tag{143}$$

Dabei ist T durch die Kommutativität des folgenden Diagramms von Mengen definiert.

$$\begin{array}{ccc} \mathbb{P} & \xrightarrow{\rho} & S^2 \\ {\scriptstyle \hat{U}}\downarrow & & \downarrow{\scriptstyle T} \\ \mathbb{P} & \xrightarrow{\rho} & S^2 \end{array} \tag{144}$$

Lemma 122 *Es gibt eine lineare Abbildung $\tilde{T} : \mathbb{R}^3 \to \mathbb{R}^3$, deren Einschränkung auf S^2 gleich T ist.*

Wir werden das beweisen, indem wir T explizit ausrechen. Insbesondere gilt dann für einen Vektor $\underline{x} \in S^2 \subset \mathbb{R}^3$, dass $\tilde{T}(\underline{x}) \in S^2$. Also hat man die Implikation

$$|\underline{x}| = 1 \quad \Rightarrow \quad |\tilde{T}(\underline{x})| = 1.$$

Daraus folgt, dass die Matrix $\tilde{T}$ orthogonal sein muss: ${}^t\tilde{T}T = E_3$.

Zur Berechnung von T nehemen wir einen beliebigen Punkt $(\xi, x_3) \in S^2$. Seine Position im Diagramm (144) sei die rechte obere Ecke. Es gilt $\rho((\xi :$

$(1-x_3)))$, wie man aus (139) erkennt. Das Bild $(z_1' : z_2')$ von $(\xi, (1-x_3))$ bei U ergibt sich durch Matrixmultiplikation mit U:

$$\begin{aligned} z_1' &= \alpha\xi + \beta(1-x_3) \\ z_2' &= -\bar{\beta}\xi + \bar{\alpha}(1-x_3) \end{aligned} \tag{145}$$

Wir setzen $(\xi', x_3') = T(\xi, x_3)$. Wegen der Kommutativität des Diagramms (144) gilt $(\xi', x_3') = \rho(z_1', z_2')$. Wir berechnen das mit (142). Wir betrachten die Vektoren:

$$\underline{z} = \begin{pmatrix} \xi \\ 1-x_3 \end{pmatrix}, \quad \underline{z}' = \begin{pmatrix} z_1' \\ z_2' \end{pmatrix}, \quad \text{d.h. } \underline{z}' = U\underline{z}.$$

$$2(1-x_3) = x_1^2+x_2^2+(1-x_3)^2 = H(\underline{z},\underline{z}) = H(U\underline{z}, U\underline{z}) = H(\underline{z}', \underline{z}') = |z_1'|^2+|z_2'|^2.$$

Deshalb sieht die Formel (142) für $(z_1' : z_2')$ so aus:

$$\xi' = \frac{z_1'\bar{z}_2'}{1-x_3}, \quad x_3' = 1 - \frac{|z_2'|^2}{(1-x_3)}. \tag{146}$$

Wegen (145) können wir die letzte Gleichung schreiben:

$$\begin{aligned} x_3' &= 1 - \tfrac{1}{(1-x_3)}(\bar{\alpha}(1-x_3) - \bar{\beta}\xi)(\alpha(1-x_3) - \beta\bar{\xi}) \\ &= 1 - \tfrac{1}{(1-x_3)}(\bar{\alpha}\alpha(1-x_3)^2 - \bar{\beta}\alpha\xi(1-x_3) - \beta\bar{\alpha}\bar{\xi}(1-x_3) + \beta\bar{\beta}\xi\bar{\xi} \\ &= 1 - \tfrac{1}{(1-x_3)}(|\alpha|^2(1-x_3) - \bar{\beta}\alpha\xi - \beta\bar{\alpha}\bar{\xi} + |\beta|^2(1+x_3)) \\ &= 2\mathcal{R}(\bar{\beta}\alpha)x_1 - 2\mathcal{I}(\bar{\beta}\alpha)x_2 + (|\alpha|^2 - |\beta|^2)x_3. \end{aligned}$$

Hier bezeichnen $\mathcal{R}$ bzw. $\mathcal{I}$ den Realteil bzw. den Imaginärteil einer komplexen Zahl.

Die erste Gleichung von (146) können wir schreiben:

$$\begin{aligned} \xi' &= \tfrac{1}{(1-x_3)}(\alpha\xi + \beta(1-x_3))(-\beta\bar{\xi} + \alpha(1-x_3)) \\ &= \tfrac{1}{(1-x_3)}(\alpha^2\xi(1-x_3) - \beta^2\bar{\xi}(1-x_3) - \alpha\beta\xi\bar{\xi} + \alpha\beta(1-x_3)^2) \\ &= \alpha^2\xi - \beta^2\bar{\xi} - 2\alpha\beta x_3. \end{aligned}$$

Wir sehen, dass jeweils die letzten Ausdrücke für x_3' und ξ' linear in x_1, x_2, x_3. Damit ist das Lemma bewiesen. *Q.E.D.*

Wir bezeichnen Real- und Imaginärteile der komplexen Zahlen α, β und ξ' so:

leere Menge

$$\alpha = p + \mathbf{i}q, \quad \beta = s + \mathbf{i}t, \quad \xi' = x_1' + \mathbf{i}x_2'.$$

Dann bekommen die Gleichungen für x_3 und ξ' die folgenden Form:

$$\begin{array}{lrrr} x_1' = & (p^2 - q^2 - s^2 + t^2)x_1 & -2(pq + st)x_2 & +2(qt - ps)x_3 \\ x_2' = & 2(pq - st)x_1 & +(p^2 - q^2 + s^2 - t^2)x_2 & -2(pt + qs)x_3 \\ x_3' = & 2(ps + tq)x_1 & +(pt - qs)x_2 & +(p^2 + q^2 - s^2 - t^2)x_3 \end{array}$$

Das ist die lineare Transformation $\tilde{T}$. Wir wissen schon, dass die Matrix der Koeffizienten orthogonal ist. Man kann $\det \tilde{T} = 1$ nachprüfen, z.B. indem man die Matrix mit der Eulerschen Drehmatrix vergleicht. (siehe Skript)

Daraus ergibt sich, dass der Gruppehomomorphismus (143) über einen Isomorphismus von Gruppen

$$SU_2 \to SO_3$$

faktorisiert. Hier ist $SO_3 = \{A \in M(3 \times 3, \mathbb{R}) \mid {}^tAA = E_3, \det A = 1\}$.

2 Anhang

2.1 Klasseneinteilungen, Äquivalenzrelationen, Faktorgruppen

Es sei M eine Menge. Eine Teilmenge von M betrachtet man oft als eine Eigenschaft und umgekehrt eine Eigenschaft als Teilmenge. Es sei E eine Eigenschaft. Man schreibt

$$T = \{m \in M \mid m \text{ hat die Eigenschaft } E\}$$

Das ist die Teilmenge aller Elemente von M, die die Eigenschaft E besitzen. Die Menge, welche überhaupt keine Elemente enthält, wird mit $\emptyset$ bezeichnet und heißt die *leere Menge* . Man kann sie so definieren:

$$\emptyset = \{m \in M \mid m \neq m\}.$$

Es seien M und N Mengen. Es sei

$$f : M \to N$$

eine Abbildung. Es sei $A \subset M$ eine Teilmenge. Dann definieren wir ihr *Bild* $f(A) \subset N$:

$$f(A) = \{f(a) \mid a \in A\}$$

Es sei $B \subset N$ eine Teilmenge. Dann definieren wir ihr *Urbild* $f^{-1}(B) \subset M$:

$$f^{-1}(B) = \{m \in M \mid f(m) \in B\}.$$

Wenn $n \in N$, so bezeichnen wir mit $\{n\}$ die Teilmenge von N, die genau aus dem Element n besteht. Die Menge $f^{-1}(\{n\})$ nennen wir die *Faser* der Abbildung f über dem Element $n \in N$.

Wir nennen f *surjektiv* , wenn $f(M) = N$. Dann sind alle Fasern der Abbildung f nichtleer.

Es seien M und N zwei Mengen. Die *Produktmenge* $M \times N$ wird definiert als die Menge aller geordneten Paare (m, n), wo $m \in M$ und $n \in N$. Man schreibt: $M^2 = M \times M$.

Bild
Urbild
Faser
surjektiv
Produktmenge
Klasseneinteilung
Klasseneinteilung
Klassen
Repr“”asentanten
Repr“”asentantens

Definition 123 *Es sei M eine Menge. Eine Menge $\mathcal{A}$ von Teilmengen von M heißt eine* Klasseneinteilung *von M, wenn die folgenden Bedingungen erfüllt sind:*

1. *Alle Mengen $A \in \mathcal{A}$ sind nicht leer.*

2. *Zwei verschiedene Teilmengen $A_1, A_2 \in \mathcal{A}$ haben keine gemeinsamen Elemente:*
$$A_1 \cap A_2 = \emptyset, \quad \text{wenn } A_1 \neq A_2.$$

3. *Für jedes $m \in M$ gibt es eine Teilmenge $A \in \mathcal{A}$, so dass $m \in A$.*

Die Mengen $A \in \mathcal{A}$ nennen wir auch *Klassen* . Ein Element $a \in A$ nennen wir einen *Repräsentanten* der Klasse A. Eine Teilmenge $S \subset M$ heißt ein *Repräsentantensystem* für die Klasseneinteilung $\mathcal{A}$, wenn für jede Menge $A \in \mathcal{A}$ der Durchschnitt $S \cap A$ aus genau einem Element besteht.

Beispiel 1: Es sei M die Ebene. Wir fassen sie als eine Menge von Punkten auf. Es sei A_0 eine Gerade der Ebene M. Es sei $\mathcal{A}$ die Menge aller Geraden der Ebene, die zu A_0 parallel sind (einschließlich der Gerade A_0 selbst). Dann ist $\mathcal{A}$ eine Klasseneinteilung der Menge M.

Beispiel 2: Es sei $f : M \to N$ eine Abbildung. Es sei $\mathcal{A}$ die Menge aller nichtleeren Teilmengen von $A \subset M$, so dass $A = f^{-1}(\{n\})$ für ein $n \in N$.

Dann ist $\mathcal{A}$ eine Klasseneinteilung von M. Die Klassen sind die nichtleeren Fasern der Abbildung f.

klassifizierende Abbildung

Es sei umgekehrt $\mathcal{A}$ eine Klasseneinteilung von M. Dann gibt es zu jedem $m \in M$ genau eine Teilmenge $A \in \mathcal{A}$, so dass $m \in A$. Wir bezeichnen A mit $\pi(m)$. Dadurch erhalten wir eine surjektive Abbildung

$$\pi : M \to \mathcal{A}. \tag{147}$$

Die Fasern dieser Abbildung, sind genau die Teilmengen von M, die Elemente von $\mathcal{A}$ sind. Wir nennen π die klassifizierende Abbildung der Klasseneinteilung $\mathcal{A}$. Allgemeiner ist eine *klassifizierende Abbildung* eine Surjektion $f : M \to N$, deren Fasern genau die Elemente von $\mathcal{A}$ sind.

Wegen (147) gibt es zu jeder Klasseneinteilung eine klassifzierende Abbildung.

Beispiel: Es sei $M = \mathbb{Z}$. Es sei $m \in \mathbb{N}$ eine natürliche Zahl. Wenn man eine ganze Zahl g durch m teilt, so bleibt ein Rest $r \in \{0, 1, 2, \ldots, m-1\}$. Man schreibt das

$$g = qm + r, \text{ wo } q \in \mathbb{Z}.$$

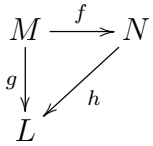
Man kann das als Abbildung auffassen

$$\begin{array}{rccc} \rho : & \mathbb{Z} & \to & \{0, 1, 2, \ldots, m-1\} \\ & g & \mapsto & r \end{array} \tag{148}$$

Die Fasern von ρ nennt man die Restklassen modulo m.

Proposition 124 *Es sei $f : M \to N$ eine surjektive Abbildung. Es sei $g : M \to L$ eine Abbildung, die auf den Fasern von f konstant ist, d.h. dass die Mengen $g(f^{-1}(\{n\}))$ für alle $n \in N$ genau aus einem Element bestehen.*

Dann gibt es eine eindeutig bestimmte Abbildung $h : N \to L$, so dass folgendes Diagramm kommutativ ist.



Beweis: Man bezeichnet das einzige Element in der Menge $g(f^{-1}(n))$ mit $h(n)$. Das ist die gewünschte Abbbildung. *Q.E.D.*

Es seien $m_1, m_2 \in M$. Man schreibt

$$m_1 \equiv m_2, \tag{149}$$

wenn m_1 und m_2 zu der gleichen Menge $A \in \mathcal{A}$ gehören.

Es sei in der Proposition 124 $f : M \to N$ eine klassifizierende Abbildung für die Klasseneinteilung $\mathcal{A}$. Dann kann man die Bedingung, dass g konstant auf den Fasern ist auch so ausdrücken: Es seien $m_1, m_2 \in M$. Dann gilt:

$$m_1 \equiv m_2 \quad \text{impliziert} \quad g(m_1) = g(m_2).$$

Es sei $(G, +)$ eine abelsche Gruppe. Es sei $H \subset G$ eine Untergruppe. Wir bezeichen mit $\mathcal{A}$ die Menge aller Teilmengen von G von der folgenden Form:

$$g + H = \{g + h \mid h \in H\}, \tag{150}$$

wobei $g \in G$ ein beliebig gewähltes Element sein kann. Die Mengen 150 heißen *Nebenklassen* von H in G. Diese Nebenklassen sind eine Klasseneinteilung $\mathcal{A}$ von G. Genauer gilt:

Lemma 125 *Für zwei Elemente $g_1, g_2 \in G$ sind die folgenden Bedingungen sind äquivalent:*

(i) $g_1 + H = g_2 + H$.

(ii) $(g_1 + H) \cap (g_2 + H) \neq \emptyset$.

(iii) $g_1 - g_2 \in H$.

Wenn diese äquivalenten Bedingungen erfüllt sind, schreiben wir entsprechend (149)

$$g_1 \equiv g_2 \quad (\mod H).$$

Satz 126 *Es sei $(G, +)$ eine abelsche Gruppe und es sei $H \subset G$ eine Untergruppe. Es sei $\mathcal{A}$ die Klasseneinteilung von G durch die Mengen (150), d.h. durch die Nebenklassen der Untergruppe H. Es sei*

$$\pi : G \to F$$

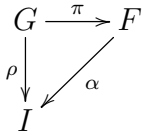
eine klassifizierende Abbildung für $\mathcal{A}$.

Dann gibt es genau eine Abbildung:

$$+_F : F \times F \to F \tag{151}$$

die auf F die Struktur einer abelschen Gruppe definiert, so dass π ein Homomorphismus von Gruppen ist.

Wenn $\rho : G \to I$ eine zweite klassifizierende Abbildung ist, so gibt es nach Proposition 124 genau eine Bijektion $\alpha : F \to I$, so dass folgendes Diagramm kommutativ ist



Bezüglich der Operationen $+_I$ und $+_F$ ist α ein Isomorphismus abelscher Gruppen.

Beweis: Es seien $\ell, g \in G$. Angenommen $+_F$ existiert. Weil dann π ein Homomorphismus ist, muss gelten:

$$\pi(\ell + g) = \pi(\ell) +_F \pi(g). \tag{152}$$

Das zwingt uns die Abbildung (151) folgendermaßen zu definieren:

Es seien $a, b \in F$. Dann wählen wir Elemente $\ell, g \in G$, so dass

$$\pi(\ell) = a, \quad \pi(g) = b.$$

Wir definieren

$$a +_F b = \pi(\ell + g).$$

Wir müssen beweisen, dass durch diese Gleichung ein Element von F, definiert wird, welches nur von a und b abhängt, aber nicht von der Wahl der Elemente ℓ und g. In der Tat, es seien $\ell', g' \in G$, so dass $\pi(\ell') = a, \quad \pi(g') = b$. Dann ist

$$\pi(\ell + g) = \pi(\ell' + g') \tag{153}$$

zu zeigen. Die Fasern die Abbildung π sind Mengen der Form (150). Daher gilt:

$$\ell + H = \ell' + H, \quad g + H = g' + H.$$

Deshalb können wir schreiben:

$$\ell = \ell' + x, \quad g = g' + y, \quad \text{wo } x, y \in H.$$

Daraus folgt:

$$\ell + g = \ell' + g' + (x + y).$$

Da das Element in der Klammer in H liegt, liegen die beiden Elemente $\ell + g$ und $\ell' + g'$ in der gleichen Menge unser Klasseneinteilung durch Nebenklassen. Sie werden also bei der klassifizierenden Abbildung π auf das gleichen Element abgebildet. Das beweist (153). Also haben wir eine Abbildung $+_F$ definiert, so dass (152) erfüllt ist.

Es ist jetzt einfach, die Gruppenaxiome für $+_F$ aus denen für $+$ abzuleiten. Wir überlassen das dem Leser. Die Existenz des Isomorphismus α ist klar. *Q.E.D.*

Bemerkung: Die Behauptung (153) können wir auch so formulieren: Für Elemente ℓ, g, ℓ', g' gilt die Implikation:

$$\ell \equiv \ell' \text{ und } g \equiv g' \quad \Rightarrow \quad \ell + g \equiv \ell' + g' \quad (\mod H).$$

Beispiel: Es sei $G = \mathbb{Z}$ die Gruppe der ganzen Zahlen bezüglich der Addition. Es sei $m \in \mathbb{N}$. Wir betrachten die Untergruppe

$$H = \{qm \mid q \in \mathbb{Z}\}.$$

Die Nebenklassen von H sind genau die Fasern der Abbildung ρ (148):

$$\rho : \mathbb{Z} \to \{0, 1, 2, \ldots, m-1\}.$$

Also erhält man auf der Menge $F := \{0, 1, 2, \ldots, m-1\}$ die Struktur einer abelschen Gruppe. Die Struktur ist durch die Gleichung

$$\rho(g_1) +_F \rho(g_2) = \rho(g_1 + g_2), \quad g_i \in \mathbb{Z},$$

Man kann sich fragen, ob es auf der Menge F auch eine Multiplikation $\cdot_F$ gibt, so dass eindeutig bestimmt.

$$\rho(g_1) \cdot_F \rho(g_2) = \rho(g_1 \cdot g_2), \quad g_i \in \mathbb{Z}.$$

Es seien $a, b \in F$ und es sei $\rho(g_1) = a$ und $\rho(g_2) = b$. Dann will man definieren:

$$a \cdot_F b = \rho(g_1 \cdot g_2)$$

Damit diese Definition sinnvoll ist, muss man wie im Beweis des letzten Satzes zeigen:

$$g_1 \equiv g_1' \text{ und } g_2 \equiv g_2' \quad \Rightarrow \quad g_1 \cdot g_2 \equiv g_1' \cdot g_2' \quad (\mod H). \tag{154}$$

Weil sich äquivalente Elemente durch einen Summanden aus $m\mathbb{Z}$ unterscheiden, kann man schreiben: $g_1 = g_1' + mu_1$ und $g_2 = g_2' + mu_2$, wobei $u_1, u_2 \in \mathbb{Z}$. Dann gilt

$$g_1 \cdot g_2 = g_1' \cdot g_2' + m(g_1' u_2 + u_1 g_2' + m u_1 u_2)$$

Das beweist die Behauptung (154).

Das F mit den Operationen $+_F$ und $\cdot_F$ ein Ring ist, ist eine Übungsaufgabe.

Beispiel: Es sei $G = \mathbb{R}^2$ mit der Addition

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \quad \text{wo } (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2.$$

Es sei $a \in \mathbb{R}$ fest gewählt. Es sei $H \subset \mathbb{R}^2$ die Gerade mit der Gleichung $y = ax$, d.h.

$$H = \{(x, y) \mid y = ax\}.$$

H ist eine Untergruppe von $\mathbb{R}^2$. Die Nebenklassen zu U sind die Parallelen zu der Geraden H. Sie haben die Gleichungen

$$y = ax + b.$$

Die Abbildung

$$\begin{array}{rccc} \pi: & \mathbb{R}^2 & \to & \mathbb{R} \\ & (x, y) & \mapsto & y - ax \end{array}$$

ist ein Homomorphismus mit dem Kern H. Also ist π eine Faktorgruppe von $\mathbb{R}^2$ modulo H.

2.2 Permutationen

Definition 127 *Es sei G eine Menge, die mit einer Abbildung $\circ : G \times G \to G$ versehen ist. Das Bild eines Paars $(\sigma, \tau) \in G \times G$ bei dieser Abbildung bezeichen wir mit $\sigma \circ \tau$.*

Wir nennen $(G, \circ)$ eine Gruppe, wenn folgende Eigenschaften erfüllt sind:

(1) (Assoziativgesetz) Für alle $\pi, \sigma, \tau \in G$ gilt:

$$(\pi \circ \sigma) \circ \tau = \pi \circ (\sigma \circ \tau).$$

(2) Es gibt ein Element $\iota \in G$, so dass für alle $\sigma \in G$

$$\iota \circ \sigma = \sigma, \quad \sigma \circ \iota = \sigma.$$

(3) Für jedes $\sigma \in G$ gibt es ein $\tau \in G$, so dass

$$\sigma \circ \tau = \iota, \quad \tau \circ \sigma = \iota.$$

Man nennt die Abbildung $\circ$ auch das Gruppengesetz und man sagt, dass $\sigma \circ \tau$ das Produkt der Faktoren σ und τ ist. Produkte von mehreren Elementen schreibt man einfach

$$\sigma_1 \circ \sigma_2 \circ \ldots \circ \sigma_n.$$

Hier ist $n \in \mathbb{N}$ die Anzahl der Faktoren. Wegen des Assoziativgesetzes (1) braucht man nicht anzugeben in welcher Reihenfolge man die Produkte bilden soll. Wenn alle Faktoren gleich σ sind, benutzt man die Potenzschreibweise:

$$\sigma^n := \sigma \circ \sigma \circ \ldots \circ \sigma.$$

Das Element ι ist durch die Eigenschaft (2) eindeutig bestimmt. Denn wenn ι' ebenfalls (2) erfüllt, so gilt

$$\iota' = \iota' \circ \iota = \iota.$$

Wir nennen ι das neutrale Element der Gruppe.

Das Element τ aus der 3. Eigenschaft ist durch σ eindeutig bestimmet. Denn wenn τ' ein weitres solches Element wäre, so folgt:

$$\tau' = \tau' \circ \iota = \tau' \circ (\sigma \circ \tau) = (\tau' \circ \sigma) \circ \tau = \iota \circ \tau = \tau.$$

Man nennt τ das inverse Element zu σ und bezeichnet es mit σ^{-1}. Man dehnt die Potenzschreibweise auf Exponenten aus $\mathbb{Z}$ aus:

$$\sigma^0 := \iota, \quad \sigma^{-n} := (\sigma^{-1})^n = (\sigma^n)^{-1} \text{ für } n \in \mathbb{N}.$$

Definition 128 *Es sei M eine Menge. Eine Permutation ist eine bijektive Abbildung $\sigma : M \to M$. Wir bezeichnen die Menge aller Permutationen mit* $\mathrm{Aut}(M)$.

Transposition

Es seien $\sigma, \pi \in \mathrm{Aut}(M)$. Das Kompositum dieser Abbildungen ist wieder eine Permutation:

$$(\pi \circ \sigma)(m) := \pi(\sigma(m)).$$

Bezüglich der Komposition von Abbildungen ist die Menge $\mathrm{Aut}(M)$ eine Gruppe. Die Permutation ι, deren Existenz unter 2. gefordert wird, ist die identische Abbildung

$$\iota := \mathrm{id}_M\,.$$

Es sei $\sigma \in \mathrm{Aut}(M)$. Da σ bijektiv ist, gibt es zu jedem $m \in M$ genau ein $n \in M$, so dass $\sigma(n) = m$. Durch

$$\tau(m) := n,$$

wird daher eine Permutation definiert, die 3. erfüllt.

Wir werden das Kompositum von Permutationen auch ohne das Zeichen $\circ$ schreiben:

$$\pi\sigma = \pi \circ \sigma..$$

Es seien $x, y \in M$ zwei verschiedene Elemente. Die *Transposition* $t_{(x,y)} \in \mathrm{Aut}(M)$ dieser beiden Elemente ist wie folgt definiert:

$$t_{(x,y)}(m) = \begin{cases} y, & \text{für } m = x \\ x, & \text{für } m = y \\ m, & \text{für } m \neq x, m \neq y. \end{cases}$$

Die Transposition $t_{(x,y)}$ vertauscht also die Elemente x und y, aber lässt alle übrigen Elemente von M fest. Wie bemerken, dass

$$t_{(x,y)}t_{(x,y)} = \mathrm{id}_M\,.$$

Die Menge der Permutationen der Menge $M = \{1, \ldots, n\}$ bezeichnen wir mit $\mathcal{S}_n := \mathrm{Aut}(M)$. Man nennt $\mathcal{S}_n$ eine symmetrische Gruppe. Eine Permutation $\sigma \in \mathcal{S}_n$ schreiben wir in Form einer Tabelle

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}. \tag{155}$$

Eine solche Tabelle kann man auch schreiben, wenn M eine beliebige endliche Menge ist. In der erste Zeile stehen die Elemente von M in einer beliebigen

Reihenfolge und in der zweiten Zeile steht $\sigma(m)$ unter dem Element m der ersten Zeile.

Die Tabelle der Permutation $\sigma t_{(x,y)}$ erhält man aus der Tabelle von σ, in dem man in der zweiten Zeile die Elemente $\sigma(x)$ und $\sigma(y)$ die Plätze tauschen lässt.

Lemma 129 *Jede Permutation einer endlichen Menge M ist Produkt von Transpositionen.*

Beweis: Man kann die Elemente die zweite Zeile einer Permutation durch fortgesetzte Vertauschung zweier Elemente so anordnen, dass sie mit der ersten Zeile übereinstimmt. D.h. wir finden Transpositionen $t_1, t_2, \ldots, t_\ell$, so dass

$$\sigma t_1 t_2 \ldots t_\ell = \mathrm{id}_M \;\Rightarrow\; \sigma = t_\ell \ldots t_2 t_1 \qquad Q.E.D.$$

Wir werden beweisen, dass $(-1)^\ell$ nur von der Permutation σ abhängt und nicht von den Elementen $t_1, \ldots, t_\ell$, die wir im Beweis des Lemmas gewählt haben. Man nennt $\mathrm{sgn}(\sigma) := (-1)^\ell$ das Signum (Vorzeichen) der Permutation σ.

Um die Unabhängigkeit von $(-1)^a$ zu beweisen, gehen wir von einer anderen Definition des Signums aus. Es sei M eine beliebige endliche Menge. Es sei P die Menge aller Teilmengen von M, die genau zwei Elemente enthalten. Es sei $p = \{m_1, m_2\} \in P$. Es sei $\sigma \in \mathrm{Aut}(M)$. Wir definieren $\sigma(p) := \{\sigma(m_1), \sigma(m_2)\}$. Die Zuordnung $p \mapsto \sigma(p)$ ist eine Permutation der Menge P.

Wir bezeichen mit $\vec{P}$ die Menge aller geordneten Paare (m_1, m_2) von zwei verschiedenen Elementen $m_1, m_2 \in M$.

Für ein geordnetes Paar $\vec{p} = (m_1, m_2)$ definieren wir $\sigma(\vec{p}) = (\sigma(m_1), \sigma(m_2))$.

Zu jeder Menge $p = \{m_1, m_2\} \in P$ gibt es genau zwei geordnete Paare (m_1, m_2) und (m_2, m_1) aus Elementen dieser Menge. Wir wählen zu jedem $p \in P$ eins dieser geordneten Paare aus.

Die Menge der ausgewählten geordneten Paare bezeichen wir mit P^+. Es seien $P^- = \vec{P} \setminus P^+$ die übrigen Paare. Wir benutzen die Schreibweise $(m_1, m_2) = -(m_2, m_1)$. Dann gilt

$$P^- = \{-\vec{p} \mid \vec{p} \in P^+\}.$$

Mit dieser fixierten Wahl von P^+ definieren wir:

Definition 130 *Es sei* $\sigma \in \mathrm{Aut}(M)$. *Das* Signum *von* σ *ist die Zahl* *Signum*

$$\mathrm{sgn}\,\sigma = (-1)^{\sharp(\sigma(P^+)\cap P^-)}.$$

Hier bezeichnet $\sigma(P^+)$ die Menge aller geordneten Paare $\sigma(\vec{p})$ mit $\vec{p} \in P^+$. Der Exponent $\sharp(\sigma(P^+) \cap P^-)$ ist die Anzahl der Elemente im Durchschnitt der beiden Mengen $\sigma(P^+)$ und P^-.

Für ein $p \in P$ bezeichnen wir mit $p^+ \in \vec{P}$ die gewählte Anordnung der Menge p. Wir definieren

$$\begin{aligned} \mathrm{sgn}_p(\sigma) = 1, &\quad \text{wenn} \quad \sigma(p^+) \in P^+ \\ \mathrm{sgn}_p(\sigma) = -1, &\quad \text{wenn} \quad \sigma(p^+) \in P^-. \end{aligned}$$

Mit dieser Definition können wir schreiben:

$$\mathrm{sgn}(\sigma) = \prod_{p\in P} \mathrm{sgn}_p(\sigma). \tag{156}$$

Lemma 131 *Die Zahl* $\mathrm{sgn}(\sigma)$ *ist unabhängig von der Wahl der Menge* P^+

Beweis: Es sei $P^{+'} \subset M \times M$ eine zweite Wahl von geordneten Paaren. Zu einer Menge $p \in P$ haben wir zwei evtl. verschiedene Anordnungen $p^+ \in P^+$ und $p^{+'} \in P^{+'}$. Wir bezeichnen das Signum, welches durch die zweite Wahl definiert ist mit sgn' bzw. sgn'_p.

Wir müssen beweisen, dass $\mathrm{sgn} = \mathrm{sgn}'$. Man kann o.B.d.A. annehmen, dass sich die geordneten Paare p^+ und $p^{+'}$ nur für ein einziges $p = s \in P$ verschieden sind.

Wir nehmen zuerst an, dass $\sigma(s) = s$. Dann gilt $\mathrm{sgn}_p = \mathrm{sgn}'_p$ für alle $p \in P$. Nach der Formel (156) gilt dann erst recht, dass $\mathrm{sgn} = \mathrm{sgn}'$.

Jetzt betrachten wir den Fall, wo $\sigma(s) \neq s$.

Es sei $p \in P$. Wir unterscheiden 3 Möglichkeiten für p:

Zuerst sei $\sigma(p) = q$, wobei p und q von s verschieden sind. Dann gilt $\mathrm{sgn}_p(\sigma) = \mathrm{sgn}'_p(\sigma)$.

Als zweites sei $\sigma(p) = s$. Dann ist $p \neq s$ und es gilt $p^+ = p^{+'}$. Da $s^+ = -s^{+'}$ erhalten wir, dass $\mathrm{sgn}_p(\sigma) = -\mathrm{sgn}'_p(\sigma)$.

Drittens sei $s = p$. Dann ist $\sigma(s) = q \neq s$. Wie im zweiten Fall sieht man, dass $\mathrm{sgn}_s(\sigma) = -\mathrm{sgn}'_s(\sigma)$. Daher ändern genau zwei Faktoren in (156) ihr Vorzeichen, wenn man von P^+ zu $P^{+'}$ übergeht.

Damit haben wir gezeigt, dass die Definition von $\mathrm{sgn}(\sigma)$ von der zusätzlichen Wahl von P^+ unabhängig ist. *Q.E.D.*

Satz 132 *Es sei M eine endliche Menge. Es seien $\sigma, \tau \in \mathrm{Aut}(M)$. Dann gilt:*

$$\mathrm{sgn}(\sigma\tau) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau)$$

Beweis: Wir wählen eine Menge P^+. Offenbar genügt es zu zeigen, dass

$$\mathrm{sgn}_p(\sigma\tau) = \mathrm{sgn}_{\tau(p)}(\sigma)\,\mathrm{sgn}_p(\tau). \tag{157}$$

Es sei $\tau(p) = q$. Wenn $\tau(p^+) = q^+ \in P^+$, ist (157) klar.

Es sei $\tau(p^+) = -q^+$. Wir müssen zeigen, dass

$$\mathrm{sgn}_p(\sigma\tau) = -\,\mathrm{sgn}_q(\sigma).$$

Aber das folgt, da $\sigma(-q^+) = -\sigma(q^+)$. *Q.E.D.*

Es sei M eine endliche Menge. Es seien m, n zwei verschiedene Elemente von M. Es sei $\sigma \in \mathrm{Aut}(M)$ die Transposition von m und n.

Wir wählen eine Menge von Paaren P^+ wie folgt:

$$(m, l) \in P^+ \text{ für alle } l \neq m$$
$$(n, l) \in P^+, \text{ für } l \notin \{m, n\}.$$

Die übrigen Paare von P^+, in denen weder m noch n vorkommt, wählen wir ganz beliebig. Dann gilt $\sigma((m, n)) = (n, m) \in P^-$. Für alle übrigen Paare $p \in P^+$ gilt offenbar $\sigma(p) \in P^+$. Also ist (n, m) das einzige Element in $\sigma(P^+) \cap P^-$. Wir finden $\mathrm{sgn}\,\sigma = -1$. Wir haben bewiesen:

Proposition 133 *Das Signum einer Transposition ist -1.*

Eine Permutation (155) . Wenn wir die untere Zeile durch Vertauschungen in die natürliche Reihenfolge bringen, erhalten wir eine Darstellung von (155) als Produkt von Transpositionen. Daher erhalten wir aus dem Satz 132, dass $\mathrm{sgn}\,\sigma = (-1)^a$.

Es sei M eine endliche Menge. Man nennt eine Permutation von M zyklisch, wenn es ein $m \in M$ gibt, so dass die Folge

$$\sigma^0(m), \sigma(m), \sigma^2(m), \ldots, \sigma^n(m), \ldots$$

sämtliche Elemente von M enthält.

Übung: Man berechne das Signum einer zyklischen Permutation.

Eine typische zyklische Permutation aus $\mathcal{S}_n$ ist

zyklische Gruppe

$$\zeta = \begin{pmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ 2 & 3 & 4 & \dots & n & 1 \end{pmatrix}. \tag{158}$$

Sämtliche Potenzen dieser Permutation ζ^i, $i \in \mathbb{Z}$ bilden eine Untergruppe $\mathcal{Z}_n$ von $\mathcal{S}_n$. Das ist die *zyklische Gruppe* der Ordnung n. Ordnung bedeutet, dass die Gruppe $\mathcal{Z}_n$ genau n Elemente enthält. Man zeigt leicht, dass

$$\zeta^{i\pm n} = \zeta^i, \quad \text{für alle } i \in \mathbb{Z}.$$

Man soll sich die Gruppe $\mathcal{Z}_n$ als zwei Zifferblätter vorstellen, die übereinanderliegen und die man gegeneinander dreht.

2.3 Polynome

Es sei A eine abelsche Gruppe und I eine Menge. Es sei

$$A^I = \mathrm{Abb}(I, A).$$

Es seien $f, g \in A^I$. Die Summe $f + g \in A^I$ dieser Funktionen ist wie folgt definiert:

$$(f+g)(i) := f(i) + g(i), \quad i \in I.$$

Dadurch wird A^I eine abelsche Gruppe. Man sagt, dass eine Funktion $f \in A^I$ endlichen Träger hat, wenn eine endliche Teilmenge $S \subset I$ existiert, so dass $f(i) = 0$ für $i \notin S$. Die Teilmenge der Funktionen mit endlichem Träger

$$A^{(I)} \subset A^I$$

ist eine abelsche Untergruppe.

Es sei $A = R$ eine Ring mit Einselement 1. Es sei $I = \mathbb{Z}_{\geq 0}$ die Menge der nicht negativen ganzen Zahlen. Wir definieren die Funktionen $X^m \in R^{(I)}$ für $m \in \mathbb{Z}_{\geq 0}$:

$$X^m(m) = 1, \quad X^m(n) = 0 \text{ für } n \neq m,\ n \in \mathbb{Z}_{\geq 0}.$$

Wenn $a \in R$ und $f \in R^{(I)}$, so definieren wir $af \in R^{(I)}$

$$(af)(m) := a(f(m)).$$

Dann kann man jede Funktion $f \in R^{(I)}$ schreiben

$$f = a_d X^d + a_{d-1} X^{d-1} + \ldots + a_1 X^1 + a_0 X^0, \quad a_i \in A. \tag{159}$$

Es gilt $f(m) = a_m$ oder $f(m) = 0$, wenn X^m in (159) nicht vorkommt.

Man definiert ein Produkt

$$\begin{array}{ccc} R^{(I)} \times R^{(I)} & \to & R^{(I)} \\ (f, g) & \mapsto & fg \end{array}$$

durch die Gleichung:

$$(fg)(m) = \sum_{u=0}^{m} f(u)g(m-u) \tag{160}$$

Das ist nicht das übliche Produkt von Funktionen. Dieses Produkt genügt der Regel:

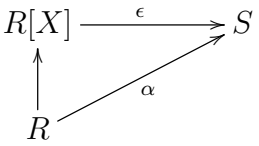
$$(aX^m)(bX^n) = abX^{m+n}.$$

Die abelsche Gruppe $R^{(I)}$ wird durch die Multiplikation (160) ein Ring, der mit $R[X]$ bezeichnet wird. Das ist der Polynomring. Man hat einen Ringhomomorphismus $R \to R[X]$, $a \mapsto aX^0$. Man definiert $a := aX^0 \in R[X]$ und betrachtet R als einen Unterring von $R[X]$. Das Element $1 = 1X^0$ ist ein Einselement von $R[X]$. Der Ring $R[X]$ ist kommutativ, wenn R kommutativ ist.

Es sei $R \stackrel{\alpha}{\to} S$ ein Homomorphismus von Ringen mit Einselement, d.h. $\alpha(1) = 1$. Es sei $s \in S$ ein Element, so dass $s\alpha(r) = \alpha(r)s$ für alle $r \in R$. Dann definiert man eine Abbildung

$$\begin{array}{rccc} \epsilon: & R[X] & \longrightarrow & S \\ a_dX^d + \ldots + a_1X^1 + a_0X^0 & & \mapsto & \alpha(a_d)s^d + \ldots + \alpha(a_1)s^1 + \alpha(a_0). \end{array}$$

Man beweist leicht, dass ϵ ein Ringhomomorphismus ist. Man hat ein kommutatives Diagramm von Ringhomomorphismen



2.4 Kategorien

Wichtige Eigenschaften ganz verschiedener Objekte, welche in der Algebra vorkommen (Gruppen, Ringe, Vektorräume usw.), kann man sehr vorteilhaft in der folgenden universellen Sprache formulieren.

Definition 134 *Eine Kategorie besteht aus folgenden Daten:*

1) *eine Menge $\mathcal{C}$, deren Elemente man die Objekte der Kategorie nennt,*

2) *eine Menge* $\mathrm{Hom}(A, B)$ *zu jedem Paar A, B von Objekten aus $\mathcal{C}$ (Die Elemente dieser Menge nennt man die Morphismen von A nach B),*

3) *Ein ausgezeichnetes Element* $\mathrm{id}_A \in \mathrm{Hom}(A, A)$ *für jedes Objekt $A \in \mathcal{C}$.*

4) *Für je 3 Objekte $A, B, C \in \mathcal{C}$ eine Abbildung*

$$\begin{array}{ccc} \mathrm{Hom}(A,B) \times \mathrm{Hom}(B,C) & \to & \mathrm{Hom}(A,C) \\ \phi \times \psi & \mapsto & \psi \circ \phi \end{array}$$

Die folgenden Eigenschaften sollen gelten:

- *Die Mengen* $\mathrm{Hom}(A, B)$ *sind disjunkt.*
- *Wenn* $\chi \in \mathrm{Hom}(C, D)$ *so gilt*

$$\chi \circ (\psi \circ \phi) = (\chi \circ \psi) \circ \phi.$$

- *Es gelten die Gleichungen:*

$$\phi \circ \mathrm{id}_A = \phi, \qquad \mathrm{id}_B \circ \phi = \phi.$$

Wenn mehrere Kategorien im Spiel sind, so schreibt man $\mathrm{Hom}_{\mathcal{C}}(A, B)$ anstelle von $\mathrm{Hom}(A, B)$.

Definition 135 *Einen Morphismus $\phi : A \to B$ der Kategorie $\mathcal{C}$ nennt man einen* Isomorphismus*, wenn ein Morphismus $\psi : B \to A$ existiert, so dass*

$$\psi \circ \phi = \mathrm{id}_A, \quad \phi \circ \psi = \mathrm{id}_B .$$

Definition 136 *Es seien $\mathcal{C}$ und $\mathcal{D}$ zwei Kategorien. Ein Funktor $F : \mathcal{C} \to \mathcal{D}$ besteht aus einer Menge von Abbildungen:*

1) *Eine Abbildung:*

$$\begin{array}{ccc} \mathcal{C} & \to & \mathcal{D}. \\ A & \mapsto & F(A) \end{array}$$

2) Für jedes Paar $A, B \in \mathcal{C}$ eine Abbildung:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathcal{C}}(A,B) & \to & \mathrm{Hom}_{\mathcal{D}}(F(A),F(B)). \\ \phi & \mapsto & F(\phi) \end{array}$$

Die folgenden Eigenschaften sind erfüllt:

$$F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$$

Wenn $A, B, C \in \mathcal{C}$, $\phi \in \mathrm{Hom}_{\mathcal{C}}(A,B)$ und $\psi \in \mathrm{Hom}_{\mathcal{C}}(B,C)$, so gilt:

$$F(\psi \circ \phi) = F(\psi) \circ F(\phi).$$

Index

Übergangsmatrix, 17
ähnlich, 18

Zeilenoperation, 14

alternierend, 38
aufgespannt, 1
Automorphismus, 6

Bild, 6, 120

Determinantenfunktion, 40
direkte Summe, 9
Drehung, 102
duale Basis, 11

Eigenvektor, 28
Eigenwert, 28
Eigenwert, Berechnung, 31
endlich erzeugt, 1
Endomorphismus, 6
Erzeugendensystem, 1
Evaluation, 77

Faktorraum, 7
Faser, 120

Grad, 51
Gramsche Matrix, 76

Hauptachsentransformation, 106
Hauptminoren, 98
Hauptraum, 29
Homomorphismen, 6

innere direkte Summe, 9
invertierbar, 16
Isometrie, 90
Isomorphismus, 6

kartesisches Koordinatensystem, 102
Kern, 6
Klassen, 121
Klasseneinteilung, 120
klassifizierende Abbildung, 121
Komplementärraum, 5
kongruent, 89
Koordinaten, 3
Koordinatensystem, 12

Länge, 99
leere Menge, 120
linear abhängig, 1
linear unabhängig, 1
lineare Abbildungen, 6
lineare Hülle, 1
Linearkombination, 1

metrisch ähnlich, 105
metrisch kongruent, 104
multilinear, 37

Nebenklassen, 122
nilpotent, 25
Nullraum, 79
Nullraum einer symmetrischen Bilinearform, 90

orthogonale direkte Summe, 91
orthogonale Komplement, 84
orthogonale Matrix, 102

Polynomfunktionen, 62
Produktmenge, 120
Projektion, 8

Relation, 1
Repräsentanten, 121
Repräsentantensystem, 121

Scherung, 14
Schmidtorthogonalisierung, 100
selbstadjundiert, 105
selbstadjungdiert, 119
Signum, 128
Spaltenvektoren, 1
Spektralsatz, 106
Spur, 19
Standardskalarprodukt, 77
Standardvektor, 2
Stufenform, 13
surjektiv, 120

Tensorprodukt, 57

Urbild, 120

verallgemeinerten Eigenraum, 29
voll, 23

Zeilenrang, 15
zerfällenden Basis, 22
zyklisch, 55
zyklische Gruppe, 130