Lineare Algebra II

Sommersemester 2022

Albert-Ludwigs-Universität Freiburg
Heike Mildenberger

Fassung vom quatorze juillet 2022

Warnung, das Skript kann Fehler enthalten.

Inhaltsverzeichnis

Kapitel 1

Die Jordan'sche Normalform

Quellen: [?], [?], [?], [?], [?], [?], [?] [?], [?], [?].

1.1 Zwei Matrizen-Muster

Definition 1.1. Es sei V ein n-dimensionaler VR. Sei $f \in \mathrm{End}(V)$. Ein Vektor v heißt Eigenvektor zum Eigenwert λ,, falls $v \neq 0$ und $f(v) = \lambda v$.

Wenn λ ein Eigenwert ist, dann heißt $V_\lambda = \{v \in V \,:\, f(v) = \lambda v\}$ der Eigenraum zu λ.

Wenn λ kein Eigenwert ist, kann man V_λ auch wie oben definieren, und erhält dann $V_\lambda = \{0\}$.

Definition 1.2. Eine n-n-Matrix der Form

$$A = \begin{pmatrix} \alpha_{1,1} & 0 & \cdots & 0 \\ 0 & \alpha_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{n,n} \end{pmatrix}$$

heißt Diagonalmatrix.

Definition 1.3. Es sei $\lambda \in K$. Eine n-n-Matrix J_λ^n der Form

$$J_\lambda^n = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & & 0 & \lambda \end{pmatrix}$$

heißt Jordanblock[1] der Dimension n zum Eigenwert λ oder kurz J_λ^n oder auch λ-Block.

1.2 Diagonalisierbarkeit und Trigonalisierbarkeit

Unsere ersten Aussagen und Definitionen lassen sich für beliebige Körper treffen.

Lemma 1.4. *Sei K ein Körper $P \in K[X]$ und $P(\lambda) = 0$. Dann teilt $(X - \lambda)$ das Polynom P.*

Beweis: Im Polynomring gibt es die Division mit Rest (Übung). Es sei

$$P = (X - \lambda) \cdot P_1 + P_2.$$

Es gilt $\deg(P_2) < \deg(X - \lambda)$, also $\deg(P_2) = 0$. Es ist $P(\lambda) = P_2(\lambda) = 0$, also ist P_2 das Nullpolynom. Nach Induktionsvoraussetzung gilt $(X-\lambda)$ teilt P_2. Der erste Summand wird durch $(X - \lambda)$ geteilt wird, teilt $(X - \lambda)$ auch ganz P. □

Definition 1.5. Set $P \in k[X]$, $P(\lambda) = 0$. Die Nullstellenvielfachheit oder Multiplizität von λ in P ist die maximale Zahl $m = \mathrm{ord}_\lambda(P)$, so dass $(X - \lambda)^m$ das Polynom P teilt.

Satz 1.6. *Es seien $\lambda_1, \dots, \lambda_m$ alle Nullstellen von P, die λ_i seien paarweise verschieden. Dann ist*

$$P = \prod_{i=1}^{m} (X - \lambda_i)^{\mathrm{ord}_{\lambda_i}(P)} \cdot Q$$

mit einem Q, das keine Nullstellen hat. Falls K algebraisch abgeschlossen ist, so ist $\sum_{i=1}^m \mathrm{ord}_{\lambda_i}(P) = \deg(P)$.

Beweis: Da $(X - \lambda_1)^{\mathrm{ord}_{\lambda_1}(P)}$ das Polynom P teilt, ist

$$P = (X - \lambda_1)^{\mathrm{ord}_{\lambda_1}(P)} \cdot P_1.$$

Nun teilt $(X - \lambda_2)^{\mathrm{ord}_{\lambda_2}(P)}$ das Polynom P_1, da $(X - \lambda_2)^{\mathrm{ord}_{\lambda_2}(P)}$ das Polynom P teilt, aber $(X - \lambda_2)$ nicht $(X - \lambda_1)^{\mathrm{ord}_{\lambda_1}(P)}$ teilt. (Zwischenrechnung!)

Wir fahren bis m fort.

Falls K algebraisch abgeschlossen ist, gilt $\deg(Q) = 0$.

□

Definition 1.7. Es seien V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$.

(a) f heißt diagonalisierbar, falls es eine Basis von V gibt, bezüglich der f durch eine Diagonalmatrix dargestellt wird.

(b) f heißt trigonalisierbar, wenn es eine Basis von V gibt, bezüglich der f durch eine obere Dreiecksmatrix dargestellt ist.

[1] Marie Ennemond Camille Jordan, genannt Camille Jordan, 5.1.1838 – 21.1.1922

1.2.1 Trigonalisierbarkeit

Satz 1.8. *Wenn das charakteristische Polynom von f in Linearfaktoren zerfällt, dann ist f trigonalisierbar.*

Wir stellen den Beweis zurück, denn er wird leicht aus der Jordan'schen Normalform ablesbar sein.

Definition 1.9 (Algebraische Vielfachheit und geometrische Vielfachheit)**.** Es ein von nun an K algebraisch abgeschlossen.

(1) Das charakteristische Polynom χ_f von f zerfällt in Linearfaktoren

$$\chi_f(X) = \prod_{i=1}^{r}(X - \lambda_i)^{\mu_i}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_r$. Die Ordnung von λ_i in χ_f, also $\mu_i = \mathrm{ord}_{\lambda_i}(\chi_f) \geq 1$, heißt algebraische Vielfachheit (oder Multiplizität) des Eigenwerts λ_i.

(2) Die Dimension des Eigenraums $V_{\lambda_i} = \{v \in V : f(v) = \lambda_i v\}$ heißt geometrische Vielfachheit von λ_i.

(3) Später in Beobachtung 1.43 werden wir sehen, dass es ein in der Teilbarkeitsrelation kleinstes Polynom $P \in K[X]$ gibt, so dass $P = \prod_{i=1}^{r}(X - \lambda_i)^{\nu(\lambda_i)}$ und $P(f) = 0 \in \mathrm{End}(V)$. Die Zahl $\nu(\lambda_i)$ heißt Nilpotenzindex oder Nilpotenzgrad von λ_i. Es gilt $\nu(\lambda_i) \leq \mu_i$ und P teilt χ_f.

1.2.2 Diagonalisierbarkeit

Satz 1.10. *Es sei V ein endlichdimensionaler k- Vektorraus $f \in \mathrm{End}(V)$. Es seien λ_1, $\ldots\lambda_m$ die Eigenwerte mit algebraischer Vielfachheit μ_i und geometrischer Vielfachheit $\varrho(\lambda_i)$. Es sind äquivalent*

(a) Das charakteristische Polynom zerfällt in Linearfaktoren und $\varrho(\lambda_i) = \mu_i$ für all $i = 1, \ldots m$.

(b) Der Vektorraum hat ein Basis aus Eigenvektoren von f.

Umformulierung

Satz 1.11. *Genau dann, wenn es eine Basis aus Eigenvektoren gibt, ist f diagonalisierbar.*

Satz 1.10 und Satz 1.11 folgen aus der Jordan'schen Normalform 1.27 und werden daher hier nicht bewiesen.

Definition 1.12. Es seien V_i, $1 \leq i \leq k$, Unterräume von V.

(1) Das Tupel $(V_1, \dots, V_k)$ heißt transversal, falls

$$\forall v_1 \in V_1 \dots \forall v_k \in V_k \, (\sum_{i=1}^{k} v_i = 0 \to v_1 = \dots = v_k = 0_V).$$

Man sagt hierzu auch V_1, ..., V_k sind als Unterräume linear unabhängig.

(2) Das Tupel $(V_1, \dots, V_k)$ heißt komplementär, falls es transversal ist und $V_1 + \dots + V_r = V$ ist.

Wir wissen schon aus LA 1, dass Eigenvektoren zu verschiedenen Eigenvektoren linear unabhängig sind. Dieses können wir nun noch einmal formulieren.

Lemma 1.13 (Transversalität der Eigenräume)**.** *Wenn λ_i, $1 \leq i \leq n$, paarweise verschieden sind, dann sind die V_{λ_i} transversal, d.h.*

$$\forall v_i \in V_{\lambda_i} (\sum_{i=1}^{n} v_i = 0 \to v_1 = \dots = v_n = 0_V).$$

Beweis: Wir führen Induktion über n. Für $n = 1$ ist das Lemma klar. Nun sei $n > 1$. Es sei $v_i \in V_i$ und $\sum_{i=1}^n v_i = 0$. Dann ist $f(0) = 0 = f(\sum_{i=1}^n v_i) = \sum_{i=1}^n \lambda_i v_i$. Aus $\sum_{i=1}^n v_i = 0$ und $\sum_{i=1}^n \lambda_i v_i = 0$ ergibt sich durch Abziehen von λ_n mal der ersten Gleichung von der zweiten Gleichung

$$\sum_{i=0}^{n-1} (\lambda_i - \lambda_n) v_i = 0.$$

Nach der Induktionsvoraussetzung ist $(\lambda_i - \lambda_n) v_i = 0$ für $i = 1, \dots, n-1$. Da $\lambda_i - \lambda_n \neq 0$, ist $v_i = 0$ für $i = 1, \dots, n-1$. Einsetzen in die Voraussetzung ergibt dann, dass auch $v_n = 0$. □

Die Summe $V_{\lambda_1} + \dots + V_{\lambda_n}$ ist also direkt. Doch spannt sie ganz V auf? Im Allgemeinen nicht.

Definition 1.14. Es sei $f \in \mathrm{End}(V)$. Wir definieren für $k \in \mathbb{N}$, $f^0 = \mathrm{id}_V$ und $f^{k+1} = f^k \circ f$. f^k heißt die k-fache Komposition von f oder die k-malige Hintereinanderausführung von f.

1.3 Eigenräume und Haupträume

Definition 1.15. Es sei K ein Körper, und V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$, $\lambda \in K$. Dann heißt

$$H_\lambda = \{v \in V : \text{es gibt ein } k \leq n, \text{ sodass } (f - \lambda \mathrm{id})^k(v) = 0\}$$

der Hauptraum von f zum Eigenwert λ.

Lemma 1.16 (Erste Eigenschaften von H_λ)**.** *Es seien K ein Körper, und V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$, $\lambda \in K$.*

(1) Genau dann, wenn λ ein Eigenwert von f ist, ist H_λ nicht der Nullraum. In diesem Fall ist $V_\lambda \subseteq H_\lambda$.

(2) Es gilt

$$\ker(f - \lambda\,\mathrm{id}) \subseteq \ker((f - \lambda\,\mathrm{id})^2) \subseteq \cdots \subseteq H_\lambda.$$

Die an einem Anfangsstück echt aufsteigende Folge von Unterräumen wird nach spätestens n Schritten stabil. Das Anfangsstück kann leer sein.

(3) Es gilt $f[H_\lambda] \subseteq H_\lambda$, oder in Worten: H_λ ist f-invariant.

Beweis: (1) und (2). Falls für $s \geq 1$, $(f - \lambda\,\mathrm{id})^s(v) = 0$, so ist für jedes $g \in \mathrm{End}(V)$, $g \circ (f - \lambda\,\mathrm{id})^s(v) = 0$, also insbesondere für $g = (f - \lambda\,\mathrm{id})^r$ für ein $r \in N$.

(3) Es sei $v \in H_\lambda$ via $(f - \lambda\,\mathrm{id})^s(v) = 0$, $s \geq 1$. Dann $\lambda v \in H_\lambda$ und $f(v) - \lambda v \in H_\lambda$, da

$$(f - \lambda\,\mathrm{id})^{s-1}(f(v) - \lambda v) = (f - \lambda\,\mathrm{id})^s(v) = 0.$$

Also ist $f(v) = f(v) - \lambda v + \lambda v \in H_\lambda$. □

Satz 1.17 ("Das Kästchenlemma)**.** *Es sei V ein endlichdimensionaler K-Vektorraum und es sei $f \in \mathrm{End}(V)$ und es sei $U \subseteq V$ ein f-invarianter Unterraum, d.h., $f[U] \subseteq U$. Dann ist $\tilde{f}: V/U \to V/U$ durch*

$$\tilde{f}(v + U) = f(v) + U$$

wohldefiniert, und es gilt

$$\chi_f = \chi_{f \upharpoonright U} \cdot \chi_{\tilde{f}}.$$

Beweis: Sei $v + U = v' + U$. Dann ist $v - v' \in U$ und daher $\tilde{f}(v - v' + U) = 0$. Wir nehmen eine angeordnete Basis $\vec{B}$ von U und ergänzen B zu einer angeordneten Basis $\vec{C} = \vec{B}^\frown\vec{D}$ von V. Die Länge von $\vec{B}$ sei r und die Länge von $\vec{D}$ sei $k = n - r$. Es sei $\vec{E} = (d_1 + U, \ldots d_k + U)$. Dann gibt es Matrizen

$$M = \mathrm{Mat}^{\vec{B}}_{\vec{B}}(f \upharpoonright U)$$

und

$$N = \mathrm{Mat}^{\vec{E}}_{\vec{E}}(\tilde{f}),$$

so dass

$$\mathrm{Mat}^{\vec{C}}_{\vec{C}}(f) = \begin{pmatrix} M & K \\ 0 & N \end{pmatrix} \tag{1.1}$$

Man rechnet nach, dass für $j = 1, \dots, k$,

$$\begin{aligned} f(d_j) &= \sum_{i=1}^{r} k_{i,j} b_i + \sum_{i=1}^{k} n_{i,j} d_i \\ \tilde{f}(d_j + U) &= \sum_{i=1}^{k} n_{i,j} (d_i + U), \end{aligned} \tag{1.2}$$

da die mit der K-Matrix gebildete erste Summation auf der rechten Seite in der ersten Zeile in (1.2) in U liegt.

Unter den Bedingungen des Kästchenlemmas gilt nun also (nach der Leinizformel zum Beispiel)

$$\det(f) = \det(M) \cdot \det(N).$$

Dann ist auch $\chi_f = \chi_{f\restriction U} \cdot \chi_{\tilde{f}}$, denn U ist auch unter $f - X\,\mathrm{id}$ invariant ist für jedes $X \in K$. $\square$

Übung 1.18. Wie vereinfacht sich die Kästchenmatrix in Zeile (1.1), falls zusätzlich auch $\vec{D}$ einen f-invarianten Unterraum aufspannt?

Satz 1.19 (Transversalität)**.** *Es seien K ein algebraisch abgeschlossener Körper, und V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$, $\lambda_i \in K$, $i = 1, \dots, r$ die paarweise verschiedenen Eigenwerte von f. Dann sind die Haupträume $H_{\lambda_1}, \dots, H_{\lambda_r}$ transversal.*

Beweis: Wir zeigen induktiv über $s \leq r$:

Falls $v_i \in H_{\lambda_i}$, $1 \leq i \leq s$ und $\sum_{i=1}^{s} v_i = 0$, so ist $v_1 = \dots = v_s = 0$.

Für $s = 1$ ist nicht zu zeigen. Insbesondere besteht V gerade aus dem Hauptraum H_{λ_1}.

Im Induktionsschritt von s auf $s + 1$ seien $v_i \in H_{\lambda_i}$, $1 \leq i \leq s + 1$ und

$$\sum_{i=1}^{s+1} v_i = 0. \tag{1.3}$$

Dann ist

$$(f - \lambda_{s+1}\,\mathrm{id})^{\nu(\lambda_{s+1})}\Big(\sum_{i=1}^{s+1} v_i\Big) = \sum_{i=1}^{s} (f - \lambda_{s+1}\,\mathrm{id})^{\nu(\lambda_{s+1})}(v_i) = 0.$$

Nach der Invarianz (Lemma 1.16(3)) ist $(f - \lambda_{s+1}\,\mathrm{id})^{\nu(\lambda_{s+1})}(v_i) \in H_{\lambda_i}$ für $i = 1, \dots, s$. Wir haben also in der rechten Seite eine Summe von s Summanden aus paarweise verschiedenen H_{λ_i}, die 0 ist. Nach Induktionsvoraussetzung ist für $i = 0, \dots s$, $w_i = (f - \lambda_{s+1}\,\mathrm{id})^k(v_i) = 0$.

Behauptung:

Dann ist für $i = 1, \dots, s$ die Abbildung

$$g = (f - \lambda_{s+1}\,\mathrm{id})^{\nu(\lambda_{s+1})}$$

injektiv auf H_{λ_i}.

Beweis der Injektivität:
Wir verfahren indirekt und nehmen für einen Widerspruch an, dass $0 \neq v \in \ker(g) \cap H_{\lambda_i}$. Es sei N minimal, so dass $(f - \lambda_i \operatorname{id})^N(v) = 0$. Dann ist $v' = (f - \lambda_i \operatorname{id})^{N-1}(v) \neq 0$ ein Eigenvektor von f zum Eigenwert λ_i. Durch Vertauschen erhalten wir

$$\begin{aligned} g(v') &= (f - \lambda_{s+1} \operatorname{id})^{\nu(\lambda_{s+1})}(f - \lambda_i \operatorname{id})^{N-1}(v) \\ &= (f - \lambda_i \operatorname{id})^{N-1}(f - \lambda_{s+1} \operatorname{id})^{\nu(\lambda_{s+1})}(v) = 0. \end{aligned}$$

Nun ist v' ein f-Eigenvektor zum Eigenwert λ_i, und daher haben wir

$$g(v') = (\lambda_i - \lambda_{s+1})^{\nu(\lambda_{s+1})} \cdot v' = 0$$

Dann ist $v' = 0$. Widerspruch.

Da nun g auf H_{λ_i} injektiv ist, impliziert $w_i = 0$ auch v_i, $i = 0, \dots, s$ gleich Null. Nach Gleichung (1.3) ist auch $v_{s+1} = 0$. □

Satz 1.20 (Dimension von H_λ und die algebraische Vielfachheit)**.** *Es seien K ein algebraisch abgeschlossener Körper, und V ein n-dimensionaler K-Vektorraum, $f \in \operatorname{End}(V)$, $\lambda \in K$ ein Eigenwert von f. Das charakteristische Polynom von $f \restriction H_\lambda : H_\lambda \to H_\lambda$ ist $(X - \lambda)^{\operatorname{ord}_\lambda(\chi_f)}$.*

Beweis: Wir wählen eine Basis $v_1, \dots, v_m$ von H_λ und ergänzen diese zu einer Basis von V, sei G ein Komplementärraum. Da H_λ f-invariant ist, gilt nach dem Kästchenlemma $\chi_f = \chi_{f \restriction H_\lambda} \cdot \chi_{f \restriction G}$.

Nach dem vorigen Satz (Satz 1.19) sind die H_μ zu verschiedenen μ transversal. Daher $\chi_{f \restriction H_\lambda}$ die Form

$$\chi_{f \restriction H_\lambda}(X) = (\lambda - X)^{\mu(\lambda)}$$

mit einem $\mu(\lambda) = \dim(H_\lambda) \leq \operatorname{ord}_\lambda(\chi_f)$, da $\chi_{f \restriction H_\lambda}$ das Polynom χ_f teilt. Andrerseits ist $\mu(\lambda) \geq \operatorname{ord}_\lambda(\chi_f)$, da $(X - \lambda)$ das Polynom $\chi_{f \restriction G}$ nicht teilt und $\chi_f = \prod_{i=1}^r (X - \lambda_i)^{\operatorname{ord}_{\lambda_i}(\chi_f)}$. □

Satz 1.21 (Komplementarität)**.** *Es seien K ein algebraisch abgeschlossener Körper, und V ein n-dimensionaler K-Vektorraum, $f \in \operatorname{End}(V)$, $\lambda_i \in K$, $i = 1, \dots, k$ die paarweise verschiedenen Eigenwerte von f. Dann ist die natürliche Abbildung*

$$\begin{aligned} g : H_{\lambda_1} \oplus \cdots \oplus H_{\lambda_k} &\to V \\ (v_1, \dots, v_k) &\mapsto \sum_{i=1}^k v_i \end{aligned}$$

ein Isomorphismus von $\bigoplus_{i=1}^k H_{\lambda_i}$ *auf* V*. Es gilt also* $V = H_1 + \cdots + H_k$*, und die Summen sind direkt.*

Beweis: Nach Satz 1.19 ist die natürliche Abbildung g injektiv und die Summe ist direkt. Es ist noch zu zeigen, dass die Summe ganz V ist.

Da $n = \sum_{i=1}^k \mathrm{ord}_{\lambda_i}(\chi_f)$ und $n = \dim(V)$ und nach dem vorigen Satz (i.e., Satz 1.20) ist $\dim(H_{\lambda_i}) = \mathrm{ord}_{\lambda_i}(\chi_f)$ ist die Abbildung g auch surjektiv. □

1.4 Analyse auf einem Hauptraum und der Satz von der Normalform

Wie immer sei V ein endlichdimensionaler K-Vektorraum. Wir betrachten nun für $f \in \mathrm{End}(V)$, λ Eigenwert von f die Abbildung $f \restriction H_\lambda$ genauer. Wir suchen Basen von H_λ, bezüglich denen $f \restriction H_\lambda$ eine besonders einfache Darstellung hat.

Definition 1.22. Es sei V ein K-Vektorraum und $f \in \mathrm{End}(V)$. Die Abbildung f heißt nilpotent, falls es ein $n \in N$ gibt, so dass f^n die Nullabbildung ist, also jedes $v \in V$ annihiliert, auf 0_V abbildet. Das kleinste solche n heißt Nilpotenzgrad von f.

Wir betrachten nun $h = f \restriction H_\lambda$ für einen Eigenwert λ. Dann ist nach Definition von H_λ die Abbildung

$$g = h - \lambda\,\mathrm{id}_{H_\lambda}$$

eine nilpotente Abbildung.

Definition 1.23. (1) Es sei V ein endlichdimensionaler Vektorraum und $f \in \mathrm{End}(V)$. Eine f-zyklische (angeordnete) Basis von V ist eine Basis der Form

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

Der Vektor v heißt der Hauptvektor der f-zyklischen Basis.

(2) Ein Unterraum W von V heißt f-zyklisch, wenn er f-invariant ist und $f \restriction W$-zyklisch ist.

(3) Die f-Ordnung eines beliebigen $x \in V$ ist das kleinste m, so dass $f^m(x) = 0$, falls dieses existiert. (Dieses ist kleiner oder gleich dem Nilpotenzgrad, falls f nilpotent ist.)

Beispiel 1.24. Es sei $n \geq 2$. Die n-n-Matrix J_0^n der Form

$$J_0^n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix}$$

definiert $f_{J_0^n} = f \in \mathrm{End}(K^n)$. Dieses f ist nilpotent vom Nilpotenzgrad n und hat die zyklische Basis

$$(f^{n-1}(e_n) = e_1, \ \ldots, \ f(e_n) = e_{n-1}, \ e_n).$$

Ist umgekehrt f nilpotent mit der zyklischen Basis

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

so hat f bezüglich dieser die Darstellung J_0^n.

Lemma 1.25. *Es sei $f \in \mathrm{End}(V)$ nilpotent, $v \neq 0$, $d \geq 1$ und $f^{d-1}(v) \neq 0$ und $f^d(v) = 0$. Dann ist v der Hauptvektor des zyklischen Teilraumes $U(v)$, der von*

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

aufgespannt wird. Die Vektoren $f^{d-1}(v), \ldots, f(v), v$ sind linear unabhängig. Es gilt $U(v) \cap \ker(f) = \{\beta f^{d-1}(v) \,:\, \beta \in K\}$.

Beweis: Es seien $\alpha_i \in K$ und

$$\sum_{i=0}^{d-1} \alpha_i f^i(v) = 0.$$

Auf diese Gleichung wenden wir f^j an für $j = 1, \ldots, d-1$ an und erhalten

$$\sum_{i=0}^{d-2} \alpha_i f^{i+1}(v) = 0$$

$$\sum_{i=0}^{d-3} \alpha_i f^{i+2}(v) = 0$$

$$\vdots$$

$$\alpha_0 f^{d-1}(v) = 0$$

Rückwärtsarbeiten durch alle d Gleichungen ergibt $\alpha_0 = \cdots = \alpha_{d-1} = 0$.

Falls $w = \sum_{i=0}^{d-1} \beta_i f^i(v) \in \ker(f)$, so ist wegen der linearen Unabhängigkeit der $f^i(v)$ höchstens $\beta_{d-1} \neq 0$.

□

Satz 1.26 (Existenz zyklischer Basen). *Es sei V' ein endlichdimensionaler K-Vektorraum und es sei $g \in \operatorname{End}(V')$ nilpotent. Dann gibt es g-zyklische komplementäre Teilräume V'_1, …, V'_r von V'. Dabei sind r und die Dimensionen der V'_i bis auf die Reihenfolge eindeutig.*

Im Beweis des Satzes 1.27 wird Satz 1.26 für jedes $i = 1, \dots, k$ auf $V' = H_{\lambda_i}$ und $g = g_i = f \upharpoonright H_{\lambda_i} - \lambda_i \operatorname{id}_{H_{\lambda_i}}$ angewendet.

Beweis: Es sei $d \geq 1$ minimal mit $g^d = 0$. Wir beweisen die Aussage induktiv über d. Für $d = 1$ ist g die Nullabbildung. Die g-zyklischen Unterräume sind in diesem Fall alle eindimensional. Wir nehmen eine Basis $\vec{B} = (v_1, \dots v_n)$ von V' und setzen $V'_i = \operatorname{span}(\{v_i\})$. g hat also als Darstellung beüglich $\vec{B}$ die Nullmatrix.

Schritt von d auf $d + 1$:
Es sei $g[V] = W$. Dieser Teilraum W ist g-stabil und hat Nilpotenzgrad d. Es gilt nach Induktionsvoraussetzung also

$$W = W_1 \oplus \cdots \oplus W_s \tag{1.4}$$

für eine Zerlegung in g-zyklische Teilräume $W_j = \operatorname{span}(g^{d_j-1}(w_j), \dots, w_j)$ mit Hauptvektoren w_j und $d_j \leq d$ mit mindestens einem $j \in \{1, \dots s\}$ mit $d_j = d$ und $g^{d-1}(w_j) \neq 0$.

Nach Induktionsvoraussetzung hat für $j = 1, \dots, s$ die Teilabbildung $g \upharpoonright W_j$ auf dem g-zyklischen Unterraum W_j bezüglich der zyklischen Basis $(g^{d_j-1}(w_j), \dots, w_j)$ die Darstellung

$$J_0^{d_j} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix}$$

Sei $w_j = g(v_j)$. Wir nehmen also irgendein Urbild v_j von w_j. Dieses v_j ist nicht eindeutig. Wir setzen nun für $j = 1, \dots, s$,

$$U_j = \operatorname{span}(g^{d_j}(v_j) = g^{d_j-1}(w_j), \dots, w_j = g(v_j), v_j).$$

Der Nilpotenzgrad erhöht sich durch

$$g^{d_j}(v_j) = g^{d_j-1}(w_j) \neq 0$$

also um 1, und es gilt $g^{d_j+1}[U_j] = \{0_V\}$ und d_j+1 ist der genaue Exponent. Die Teilabbildung $g \upharpoonright U_j$ hat in diesem Fall auf dem g-zyklischen Unterraum U_j bezüglich der zyklischen Basis

$(g^{d_j-1}(w_j), \dots, w_j, v_j)$ die Darstellung

$$J_0^{d_j+1} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & \cdots & & 0 & 0 & 0 \end{pmatrix}$$

Wir zeigen als nächstes: Es gibt einen Unterraum X vom $\ker(g)$ mit

$$V' = X \oplus U_1 \oplus \cdots \oplus U_s, \tag{1.5}$$

Die Transversalität der U_j folgt aus (1.4) und einer Anwendung von g auf $\sum_{j=0}^s v_j = 0$.

Die U_j spannen zusammen mit $\ker(g)$ ganz V' auf: Sei $v \in V'$. Dann ist

$$g(v) = w = \sum_{j=1}^{s} w_j'$$

mit geeigneten $w_j' \in W_j$, denn $W = W_1 \oplus \cdots \oplus W_s$.

Die w_j' sind Linearkombinationen aus $g(v_j) = w_j$, $\dots g^{d_j}(v_j) = g^{d_j-1}(w_j)$, denn der Unterraum W_j hat ja die zyklische Basis $(g^{d_j-1}(w_j), \dots, w_j)$. Wir haben also eine Linearkombination

$$w_j' = \sum_{\iota=0}^{d_j-1} \alpha_{j,\iota} g^\iota(w_j) = \sum_{\iota=0}^{d_j-1} \alpha_{j,\iota} g^{\iota+1}(v_j).$$

Dann ist

$$v - \sum_{j=1}^{s} \sum_{\iota=1}^{d_j} \alpha_{j,\iota} g^\iota(v_j) \in \ker(g).$$

Wir können also jeden Vektor aus V' durch $U_1 \oplus \cdots \oplus U_s$ und einen Vektor aus dem Kern von g kombinieren. Wir wählen aus $\ker(g)$ eine Komplementärraum X zu $U_1 \oplus \cdots \oplus U_s$ aus. Damit ist die Zwischenbehauptung (1.5) gezeigt.

Der Unterraum X des Kerns wiederum lässt sich als Summe von $\dim(X)$-vielen eindimensionalen g-zyklischen Unterräumen schreiben. Wir erhalten also aus (1.5) eine Darstellung

$$V' = V_1' \oplus V_2' \oplus \cdots \oplus V_r'$$

und die U_1, …, U_s sind unter den V_i'. Hierbei ist

$$\dim(\ker(g)) = \dim(\ker(g^1)) - \dim(\ker(g^0)) = r$$

und für $1 \le j \le d$

$$\dim(\ker(g^j)) - \dim(\ker(g^{j-1})) = \text{ Anzahl der } V_i' \text{ mit } \dim(V_i') \ge j;$$

und demnach für $1 \le j \le d$

$$\begin{aligned}\dim(\ker(g^j)) - \dim(\ker(g^{j-1})) - \big(\dim(\ker(g^{j+1})) - \dim(\ker(g^j))\big)\\ = \text{ Anzahl der } V_i' \text{ mit } \dim(V_i') = j.\end{aligned}$$

□

Satz 1.27 (Jordan'sche Normalform)**.** *Es sei K ein algebraisch abgeschlossener Körper und V ein endlichdimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$. Dann gibt es eine angeordnete Basis von V, bezüglich der f die Darstellung Eine n-n-Matrix der Form*

$$J = \begin{pmatrix} J_{\kappa_1}^{n_1} & & & \\ & J_{\kappa_2}^{n_2} & & \\ & & \ddots & \\ & & & J_{\kappa_r}^{n_r} \end{pmatrix}$$

Hierbei sind die κ_ℓ, $\ell = 1, \dots, r$ und die n_ℓ eindeutig bis auf die Reihenfolge. Wiederholungen unter den κ_ℓ sind gestattet. (Alle κ_ℓ mit gleichem Wert λ gehören also zum Hauptraum H_λ.) In den freien Feldern stehen lauter Nullen. Für jeden vorkommenden Eigenwert λ gilt:

(1) Die algebraische Vielfachheit von λ ist die Summe der Größen aller κ_ℓ-Blöcke mit $\kappa_\ell = \lambda$.

(2) Der Nilpotenzgrad zum Eigenwert λ ist die Größe des größten κ_ℓ-Blocks für $\kappa_\ell = \lambda$.

(3) Die Dimension des Eigenraums zum Eigenwert λ ist die Anzahl der κ_ℓ-Blöcke, für die $\kappa_\ell = \lambda$ ist.

Beweis: Erster Schritt: Zerlegung in Haupträume.
Es seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von f. (Dies ist nun eine ganz andere Nummerierung als in der Matrix! Daher haben wir jene oben κ_ℓ genannt.) Dann gilt nach den Sätzen 1.19, 1.20 und 1.21

$$V = \bigoplus_{i=1}^{k} H_{\lambda_i}.$$

Die H_{λ_i} sind f-invariant. Die Darstellung ist eindeutig bis auf die Reihenfolge. Nun halten wir ein $i \in \{1, \dots, k\}$ fest. Wir setzen

$$g_i = f \restriction H_{\lambda_i} - \lambda_i \, \mathrm{id}_{H_{\lambda_i}}.$$

Zweiter Schritt: Zerlegung eines Hauptraums in g_i-zyklische Unterräume.

Dann ist nach Konstruktion H_{λ_i} auch g_i-invariant und g_i ist ein nilpotenter Endomorphismus von H_{λ_i}. Daher gilt nach Satz 1.26 angewandt auf g_i,

$$H_{\lambda_i} = \bigoplus_{j=1}^{r_i} V_{i,j}$$

mit g_i-zyklischen Unterräumen $V_{i,j}$, wobei nach Satz 1.26 der Raum $V_{i,j}$ durch

$$((f - \lambda_i \operatorname{id})^{d_{i,j}-1}(v_{i,j}), \dots v_{i,j})$$

aufgespannt sei und $g_i \upharpoonright V_{i,j}$ den Nilpotenzgrad $d_{i,j}$ hat. Für jedes $j \in \{1, \dots, r_i\}$ haben $g_i \upharpoonright V_{i,j}$ bzw. $f \upharpoonright V_{i,j}$ hat bezüglich dieser Matrix die Darstellung

$$J_0^{d_{i,j}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & & 0 & 0 \end{pmatrix} \qquad J_{\lambda_i}^{d_{i,j}} = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_i & 1 \\ 0 & \cdots & & 0 & \lambda_i \end{pmatrix}$$

Wiederum ist die Darstellung eindeutig bis auf die Reihenfolge. $r_i = \dim(V_{\lambda_i})$ ist die geometrische Vielfachheit von λ_i. Die Matrix $J_{\lambda_i}^{d_{i,j}}$ heißt Jordanblock oder λ_i-Block.

Dritter Schritt: Zurück zu einem ganzen Hauptraum.
Nun setzen wir die angeordneten Basen der $V_{i,1}, \dots, V_{i,r_i}$ hintereinander und erhalten bezüglich dieser angeordneten Basis von H_{λ_i} die Darstellung von $f \upharpoonright H_{\lambda_i}$ der Form

$$J_{\lambda_i}^{d_{i,1},\dots,d_{i,r_i}} := \begin{pmatrix} J_{\lambda_i}^{d_{i,1}} & & & \\ & J_{\lambda_i}^{d_{i,2}} & & \\ & & \ddots & \\ & & & J_{\lambda_i}^{d_{i,r_i}} \end{pmatrix}$$

hat. In den freien Feldern stehen lauter Nullen. Hierbei heißt $J_{\lambda_i}^{d_{i,1},\dots,d_{i,r_i}}$ "Jordankästchen zum Eigenwert λ_i" oder "Jordan-Matrix zum Eigenwert λ_i" oder "Jordan-Matrix für f auf dem Hauptraum H_{λ_i}". Es gilt für $1 \le i \le k$:

- Die algebraische Vielfachheit $\mu(\lambda_i)$ von λ_i ist gleich $\sum_{j=1}^{r_i} d_{i,j}$ und dies ist $\dim(H_{\lambda_i})$,
- der Nilpotenzgrad $\nu(\lambda_i)$ von g_i ist das Maximum der $d_{i,j}$, $j = 1, \dots, r_i$,

– die Dimension des Eigenraums V_{λ_i} ist r_i, d.h., die Anzahl der λ_i-Blöcke. Diese heißt auch die geometrische Vielfachheit von λ_i.

Vierter Schritt: Zusammensetzung der Haupträume zu ganz V.
Nun betrachten wir wieder alle $i \in \{1, \dots, k\}$ zusammen. Aus der Zusammensetzung der $f \restriction H_{\lambda_i}$ auf den f-invarianten Haupträumen H_{λ_i}, $i = 1, \dots, k$, erhält man eine Matrix wie in der Aussage des Satzes. □

Mit Satz 1.27 sind auch Satz 1.8 und Satz 1.11 bewiesen.

1.5 Folgerungen und Anwendungen

Beobachtung 1.28. *Jeder Jordanblock J^n_λ und jedes Jordan-Kästchen $J^{d_1,\dots,d_r}_\lambda$ mit $\sum_{i=1}^r d_i = n$ haben die Form*

$$\lambda E_n + N$$

mit einer nilpotenten n-n-Matrix. Wir multiplizieren die beiden Summanden und erhalten

$$\lambda E_n \cdot N = N \cdot \lambda E_n. \tag{1.6}$$

Jede Matrix in Jordan-Normalform lässt sich daher als

$$J = M_s + M_n$$

schreiben mit einer Diagonalmatrix M_s und einer nilpotenten Matrix M_n. Da jeder Hauptraum f-invariant ist, impliziert die Vertauschbarkeit (1.6) *folgendes Kommutieren:*

$$M_s \cdot M_n = M_n \cdot M_s.$$

Diese Beobachtung gilt nicht nur auf jedem Hauptraum einzeln, sondern auch für die Gesamtdarstellung J. Dies folgt aus einem Kästchenlemma für die Matrizenmultiplikation der folgenden Form und der Induktion über die Anzahl der verschiedenen Haupträume von f.

Lemma 1.29 (Kästchenlemma für die Matrizenmultiplikation)**.** *Es sei $n = r + k$, und es seien für $i = 1, 2$, $M_i \in M_{r,r}(K)$ und $N_i \in M_{k,k}(K)$.*

$$\begin{pmatrix} M_1 & 0 \\ 0 & N_1 \end{pmatrix} \cdot \begin{pmatrix} M_2 & 0 \\ 0 & N_2 \end{pmatrix} = \begin{pmatrix} M_1 \cdot M_2 & 0 \\ 0 & N_1 \cdot N_2 \end{pmatrix}. \tag{1.7}$$

Satz 1.30 (Additive Jordan-Zerlegung, auch Jordan-Chevalley-Zerlegung genannt)**.** *Es sei K algebraisch abgeschlossen und $M \in M_n(K)$. Dann gibt es eine Zerlegung*

$$M = M_s + M_n,$$

in der M_s diagonalisierbar und M_n nilpotent ist und

$$M_s M_n = M_n M_s$$

gilt. Die Zerlegung ist durch die drei Eigenschaften eindeutig bestimmt.

Definition 1.31. Für diagonalisierbare Matrizen ist auch der Begriff halbeinfach (auf Englisch semi-simple) üblich. Daher wählten wir den Namen M_s.

Beweis: des Satzes. Es seien S eine reguläre Basistransformationsmatrix und $J = S^{-1}MS$ eine Jordan-Normalform. Dann existiert nach der Beobachtung für J eine Zerlegung $J = M_s + M_n$. Die Transformation in umgekehrte Richtung ergibt

$$M = SJS^{-1} = SM_sS^{-1} + SM_nS^{-1}.$$

Der erste Summand ist diagonalisierbar und der zweite ist nilpotent und sie kommutieren.

Wir zeigen nun die Eindeutigkeit.

Dazu zeigen wir erst ein Lemma

Lemma 1.32. *Sei $f = f_s + f_n$ eine additive Jordan-Zerlegung. Dann kommutieren f_n und f_s mit allen Endomorphismen, die mit f kommutieren.*

Beweis: Sei $g \in \mathrm{End}(V)$ und $gf = fg$. Dann ist jeder Hauptraum H_λ von f auch g-invariant. Dies sieht man wie folgt: Es sei $x \in H_\lambda$. Wir zeigen $g(x) \in H_\lambda$.

Es sei $(f(x) - \lambda \cdot x)^n = 0$, $n = 1$. Dann ist $f(x) = \lambda x$ und $g(f(x)) = \lambda g(x) = f(g(x))$, also auch $g(x) \in V_\lambda \subseteq H_\lambda$.

Schritt von n auf $n+1$:
Es sei $(f(x) - \lambda \cdot x))^{n+1} = 0$. Dann ist $(f - \lambda\,\mathrm{id})^n(f(x) - \lambda \cdot x) = 0$, $f(x) - \lambda x \in H_\lambda$ mit Exponent n. Nach der Induktionssvoraussetzung für n ist dann

$$g(f(x) - \lambda x) = g(f(x)) - \lambda g(x) \overset{vertauscht}{=} f(g(x)) - \lambda g(x)) = (f - \lambda\,\mathrm{id})(g(x)) \in H_\lambda.$$

Also ist $g(x) \in H_\lambda$.

Auf H_λ ist $f \restriction H_\lambda = \lambda\,\mathrm{id}_{H_\lambda} + f'_n$, $f'_n \in \mathrm{End}(H_\lambda)$, f'_n nilpotent. Dann vertauscht $g \restriction H_\lambda$ mit $f \restriction H_\lambda$ und auch mit $\lambda \cdot \mathrm{id}_{H_\lambda}$ also auch mit f'_n.

Zusammenbetrachten aller Haupträume ergibt, dass g mit f_n vertauscht. Also vertauscht g auch mit f_s. □

Fortsetzung des Eindeutigkeitsbeweises: Sei nun also $M = M_s + M_n = D + N$, D diagonalisierbar und N nilpotent und $DN = ND$. Es gilt dann

$$DM = D(D+N) = D^2 + DN = D^2 + ND = (D+N)D = MD$$

Nach dem Lemma, angewandt auf $f = f_M$[2] und $g = f_D$, ist nun auch

$$M_s D = D M_s.$$

Da beide diagonalisierbar sind und vertauschen, gibt es eine gemeinsame Basis aus Eigenvektoren. (Übung, dies geht wie $n = 1$ im Lemma 1.32) In der Darstellung bezüglich so einer Basis ist dann auch $M_s - D$ eine Diagonalmatrix.

Wiederum ist

$$NM = N(D+N) = ND + N^2 = DN + N^2 = (D+N)N = MN$$

und somit kommutieren auch M_n und N. Daher ist $N - M_n$ nilpotent. Dies sieht man wie folgt: Es seinen $N^\nu = 0$ und $M_n^\mu = 0$. Dann ist

$$(N - M_n)^{\mu+\nu} \overset{vertauscht}{=} \sum_{j=0}^{\mu+\nu} \binom{\nu+\mu}{j} N^j M_n^{\mu+\nu-j} = 0.$$

Hierbei schaut man für $j \le \nu$ auf den hinteren Faktor und für $j > \nu$ auf den vorderen in dem jeweiligen Summanden und sieht, dass alle Summanden schon $0 \in \mathrm{End}(V)$ sind. Die binomische Formel darf man anwenden, weil die beiden Endomorphismen sich vertauschen. Es gilt $\binom{\nu+\mu}{j} = \frac{(\mu+\nu)!}{j!\cdot(\mu+\nu-j)!}$.

Da $M = M_s + M_n = D + N$ ist

$$M_s - D = N - M_n$$

nilpotent und diagonalisierbar und somit die Nullmatrix.

□

Definition 1.33. Eine Matrix $M \in M_n(K)$ heißt unipotent, falls $M - E_n$ nilpotent ist.

Beobachtung 1.34. *Wenn M invertierbar ist, so ist in $M = M_s + M_n$ der Summand M_s ebenfalls invertierbar, und es gilt*

$$M = M_s(E_n + M_s^{-1} M_n) = M_s M_u$$

und M_u ist unipotent und $M_s M_u = M_u M_s$.

[2] Erinnerung $f_M \in \mathrm{End}(K^n)$ wird in den Standardbasen durch Multiplikation mit M beschrieben.

Satz 1.35 (Multiplikative Jordan-Zerlegung). *Es sei K algebraisch abgeschlossen und $M \in M_n(K)$ invertierbar. Dann gibt es eine Zerlegung*

$$M = M_s M_u,$$

in der M_s diagonalisierbar und M_u invertierbar und unipotent ist und

$$M_u M_n = M_n M_u$$

gilt. Die Zerlegung ist durch die drei Eigenschaften eindeutig bestimmt.

Beweis: Es ist $M_u = E_n + M_s^{-1} M_n$. Die Matrix $M_s^{-1} M_n$ ist nilpotent. Sei $M = DU$ eine Faktorisierung. Dann ist $M = D + N$ mit $N = D(U - E_n)$. Aus der Eindeutigkeit der additiven Zerlegung folgt $D = M_s$ und damit auch $U = M_u$. □

Definition 1.36. Es sei $f \in \mathrm{End}(V)$. Wir setzen $f^0 = \mathrm{id}_V$ und $f^{n+1} = f \circ f^n$. Dann definieren wir für $P = \sum_{i=1}^n a_i X^i \in k[X]$ den Endomorphismus $P(f) = \sum_{i=0}^n a_i f^i$.

Lemma 1.37. *Für $P, Q \in k[X]$ is $Q(P(f)) = (Q(P))(f)$.*

Beweis: Induktiv über den Grad von Q. Man kann sich der Additivität wegen auf Monome Q beschränken. Wir führen Induktion über m. Anfang: $m = 0$. Wenn Q das konstante Polynom ist, ist nicht zu zeigen. Sei nun $Q = X^{m+1} = X \cdot Q_0$. Es sei $P = \sum_{i=0}^n a_i X^i$. Dann ist $Q(P(f)) = (\sum_{i=0}^n a_i f^i)^{m+1} = (\sum_{i=0}^n a_i f^i)(\sum_{i=0}^n a_i f^i)^m \overset{Ind.Vor.}{=} P(f) \circ (Q_0(P))(f) = (Q(P))(f)$. □

Satz 1.38. *Es gibt ein Polynom $P \in K[X]$, so dass $M_s = P(M)$.*

Beweis: Wir nehmen M in der Jordan-Normalform. Es seien $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von M. Wir setzen für $i = 1, \ldots, k$,

$$Q_i = \prod_{j \neq i} \frac{X - \lambda_j}{\lambda_i - \lambda_j}$$

Wir setzen nun für X das Jordankästchen $J_{\lambda_i}^{d_{i,1},\ldots,d_{i,d_{r_i}}}$ ein, das $f \restriction H_{\lambda_i}$ darstellt, und erhalten $Q_i(f) \restriction H_{\lambda_i}$ ist eine obere Dreiecksmatrix. Wir nennen unten $(Q_i(f) \restriction H_{\lambda_i})^{\nu_i} = J_i$, ν_i wird unten bestimmt oder gewählt.

Für $j \neq i$ ist $Q_i(f) \restriction H_{\lambda_j}$ nilpotent und daher gibt es ein ν_i[3], so dass

$$(Q_i(M))^{\nu_i} = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & J_i & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$$

für eine Dreiecksmatrix J_i mit Diagonaleinträgen 1. Auch in den oberen Feldern in J_i oberhalb dessen Nebendiagonale stehen möglicherweise keine Nullen. Es gibt ein Polynom, so dass $P_i(J_i) = J_i^{-1}$ (Übung, dies zeigt man zum Beispiel mit dem Satz von Cayley–Hamilton 1.40. Der Satz, den wir gerade beweisen, wird nirgendwo verwendet, also auch nicht im Beweis des Satzes von Cayley–Hamilton. Wir erzeugen also hier keinen Zirkelschluss.) Wir setzen $R_i = XP_i$. Dann ist $R_i(0) = 0$ und

$$R_i(J_i) = (X \cdot P_i)(J_i) = J_i \cdot P_i(J_i) = J_i J_i^{-1} = \mathrm{id}_{H_{\lambda_i}}.$$

Dies heißt

$$R_i((Q_i(M))^{\nu_i}) = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & \mathrm{id} \restriction H_{\lambda_i} & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$$

Dann ist $M_s = P(M)$ für

$$P = \sum_{i=1}^{k} \lambda_i \cdot R_i((Q_i(X))^{\nu_i})$$

□

Bemerkung 1.39. Aus Satz 1.38 erhält man einen komplizierteren alternativen Beweis des Lemmas 1.32.

[3] Man kann sich überlegen, dass $\nu_i = \max\{\nu(\lambda_j) : j \neq i\}$ ausreicht.

Eine Anwendung der additiven Zerlegung

Wir betrachten für $(y_1, \dots, y_n)^T \in \mathbb{R}^n$ die Differentialgleichung

$$(y_1', \dots, y_n')^T = A(y_1, \dots, y_n)^T,$$

in einem Argment x, also $y = y(x)$, $y\colon \mathbb{R} \to \mathbb{R}^n$, y differenzierbar. Für $n = 1$ ist die Lösung $y = \exp(ax)$ mit $A = (a)$. Für größere n bildet man

$$A = S^{-1}(D + N)S$$

mit einem nilpotenten N und einer Diagonalmatrix D. Dann ist

$$\exp(xA) = (\sum_{n=0}^{\infty} \frac{1}{n!}(x^n A^n) = \exp(x(S^{-1}JS) = S^{-1}\exp(xJ)S = S^{-1}\exp(xD)\cdot\exp(xN)S$$

Nun ist $\exp(xD)$ eine Diagonalmatrix mit Einträgen $e^{\lambda_i \cdot x}$ und $\exp(xN)$ eine abbrechende Summe und somit ein Element von $M_{n,n}(K[x])$.

1.6 Der Satz von Cayley–Hamilton

Satz 1.40 (Der Satz von Cayley–Hamilton)**.** *Es sei K ein Körper und V ein endlichdimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$. Dann ist $\chi_f(f)$ die Nullabbildung.*

Beweis: Falls K nicht algebraisch abgeschlossen ist, gehen wir zum algebraischen Abschluss $\bar{K}$ von K über (dieser wird in Algebra und Zahlentheorie hergeleitet) und erweitern V zu einem $\bar{K}$-Vektorraum und erweitern f entsprechend zu $\bar{f}$. Die Basen und die Matrizendarstellungen ändern sich dabei nicht. Es ist wichtig, dass die Matrix über K ist. Dies geht auf genau eine Weise[4] und es gilt $\chi_f = \chi_{\bar{f}} \in K[X] \subseteq \bar{K}[X]$. Wenn nun $\chi_{\bar{f}}(\bar{f}) = 0 \in \mathrm{End}(\bar{V})$, so ist $\chi_{\bar{f}}(f) = 0 \in \mathrm{End}(V)$.

Wir rechnen daher von nun an mit einem algebraisch abgeschlossenem K. Es seien λ_1, …, λ_k die Eigenwerte von f und es seien H_{λ_i} die zugehörigen Haupträume. Diese sind alle f-invariant. Für jedes $i = 1, \dots, k$ zerlegt sich nach Satz 1.26 der Eigenraum H_{λ_i} in zyklische Unterräume $V_{i,1}$, …, $V_{i,\varrho}$, …, V_{i,r_i} (hierbei ist $r_i = \dim(V_{\lambda_i})$), und $g_i \upharpoonright V_{i,\varrho}$ hat den Nilpotenzgrad $d_{i,\varrho}$ für $\varrho = 1, \dots, r_i$. Wie im Satz 1.27 gezeigt, ist für den Nilpotenzgrad

$$\nu(\lambda_i) = \max\{d_{i,\varrho} : \varrho = 1, \dots, r_i\}$$

die Abbildung

$$(f \upharpoonright H_{\lambda_i} - \lambda_i \,\mathrm{id}_{H_{\lambda_i}})^{\nu(\lambda_i)}$$

[4] Genaueres gibt es dazu im Kapitel über Tensoren

die Nullabbildung auf H_{λ_i}, also die Null in $\mathrm{End}(H_{\lambda_i})$. Dann ist auch

$$(f \upharpoonright H_{\lambda_i} - \lambda_i \,\mathrm{id}_{H_{\lambda_i}})^{\mu(\lambda_i)}$$

der Nullendomorphismus im Raum $\mathrm{End}(H_{\lambda_i})$, da jeweils der Nilpotenzgrad $\nu(\lambda_i)$ kleiner oder gleich der algebraischen Multiplizität $\mu(\lambda_i) = \mathrm{ord}_{\lambda_i}(\chi_f)$ ist.

Ingesamt haben wir also für $X = f$,

$$\prod_{i=1}^{k} (X - \lambda_i \,\mathrm{id}_V)^{\nu(\lambda_i)} = 0 \in \mathrm{End}(V).$$

Dies sieht man zum Beispiel so: Für $v \in H_{\lambda_j}$ nützt man

$$\prod_{i=1}^{k} (f - \lambda_i \,\mathrm{id}_V)^{\nu(\lambda_i)} = \prod_{i=1, i \neq j}^{k} (f - \lambda_i \,\mathrm{id}_V)^{\nu(\lambda_i)} (f - \lambda_j \,\mathrm{id}_V)^{\nu(\lambda_j)}$$

und wendet diese Abbildung auf v an. Dann ist schon die Anwendung des am weitesten rechts stehenden Faktors Null.[5] □

Definition 1.41. Das Minimalpolynom von f ist das von der Teilbarkeit her kleinste Polynom P, so dass $P(f) = 0$, d.h. $P(f) = 0$ und für jedes Q mit $Q(f) = 0$ gilt: P teilt Q.

Die Existenz eines in einem solch starken Sinne kleinsten Polynoms folgt daraus, das der Polynomring Teilen mit Rest gestattet:

Lemma 1.42. *Falls die Polynome P und Q beide f annullieren, so auch annulliert auch der größte gemeinsame Teiler von P und Q den Endomorphismus f, denn dieser größte gemeinsame Teiler R hat die Form $R = P_1 P + Q_1 Q$.*

Beobachtung 1.43. *Mit den Bezeichnungen von oben ist*

$$\prod_{i=1}^{k} (X - \lambda_i \,\mathrm{id}_V)^{\nu(\lambda_i)}$$

das Minimalpolynom von f. Das Minimalpolynom teilt das charakteristische Polynom.

[5] Ein anderer Beweis wird mittels einer vollständigen Fahne geführt. Dies kommt vielleicht in eine Übung.

Kapitel 2

Skalarprodukte, euklidische und unitäre Vektorräume

2.1 Definitionen und Beispiele

In diesem Kapitel wird unter Körper immer $\mathbb{R}$ oder $\mathbb{C}$ verstanden. Wir werden die lineare Ordnung $<$ auf $\mathbb{R}$ wesentlich in unseren Definitionen benutzen. Wir schreiben $\mathbb{K}$ für einen dieser Körper. Die komplexe Konjugation ist die Abbildung

$$(x + iy) = \alpha \mapsto \bar{\alpha} := (x - iy) \tag{2.1}$$

Diese Abbildung ist linear und es gilt $\bar{x}\bar{y} = \overline{xy}$. Man sagt hierzu auch, dass die Konjugation mit der Addition und mit der Multiplikation verträglich ist. Man schreibt für $c = x + iy \in \mathbb{C}$ mit reellen x, y, $x = \mathrm{Re}(c)$, $y = \mathrm{Im}(c)$ und nennt

$$|c| = \sqrt{x^2 + y^2}$$

den Betrag von c. Es ist $|c| = \sqrt{\bar{c}c}$.

Definition 2.1. Es sei V ein $\mathbb{K}$-Vektorraum.

(1) Eine Abbildung

$$s\colon V \times V \to \mathbb{K}$$

heißt Bilinearform (für $\mathbb{K} = \mathbb{R}$) bzw. Sequilinearform (für $\mathbb{K} = \mathbb{C}$), falls s im zweiten Argument linear und im ersten Argument semilinear ist, d.h. für alle $\alpha, \beta \in \mathbb{K}$ und $v, w, x \in V$,

$$\begin{aligned} s(\alpha v + \beta w, x) &= \bar{\alpha} s(v, x) + \bar{\beta} s(w, x), \\ s(v, \alpha x + \beta y) &= \alpha s(v, x) + \beta s(v, y). \end{aligned}$$

(2) Die Form s heißt symmetrisch bzw. hermitesch, falls gilt

$$s(v,w) = \overline{s(w,v)}.$$

(3) Die symmetrische/hermitesche Bilinearform heißt Skalarpodukt, falls zusätzlich gilt

$$\text{für } v \neq 0 \text{ ist } s(v,v) > 0.$$

Diese Eigenschaft nennt man “s ist positiv definit”.

(4) Das Paar (V,s) heißt euklidischer (im Falle $\mathbb{K} = \mathbb{R}$) bzw. unitärer (im Falle $\mathbb{K} = \mathbb{C}$) Vektorraum.

Bemerkung 2.2.

(1) Es gilt $s(v,v) = \overline{s(v,v)} \in \mathbb{R}$, also ist $s(v,v)$ gleich Null oder größer Null oder kleiner Null.

(2) Das Adjektiv “hermitesch” wird auch im Sinne von unitär gebraucht.

(3) Man kann auch nur $s(v,v) \geq 0$ fordern und nennt dies dann positiv semidefinit. Analog kann man negativ definit, negativ semidefinit und indefinit definieren.

Beispiel 2.3.

(1) Das Standardskalarprodukt auf dem $\mathbb{K}^n$ ist

$$s(x,y) = \bar{x}^T y = \sum_{i=1}^{n} \bar{x}_i \cdot y_i.$$

(2) Es sei $a < b \in \mathbb{R}$. Das Riemannintegral-Skalarprodukt auf $V = \{f\colon [a,b] \to \mathbb{K} : f \text{ stetig}\}$ ist gegeben durch

$$s(f,g) = \int_a^b \bar{f}(t) g(t) dt.$$

(3) Auf dem $\mathbb{C}^2$ ist

$$s(x,y) = 2\bar{x}_1 y_1 + \bar{x}_2 y_2$$

auch ein Skalarprodukt.

Lemma 2.4. *Es sei (V,s) ein euklidischer/ unitärer Vektorraum und es sei U ein Untervektorraum von V. Dann ist (U,s) (— genauer $(U, s \restriction U \times U)$ —) auch ein euklidischer/unitärer Vektorraum.*

Definition 2.5. Es sei V ein euklidischer/unitärer Vektorraum mit angeordneter Basis $\vec{B} = (b_1, \dots, b_n)$, $n \in \mathbb{N} \setminus \{0\}$. Dann heißt

$$\mathrm{Mat}_{\vec{B}}(s) := (s(b_i, b_j)_{1 \leq i,j \leq n})$$

die darstellende Matrix von s bezüglich $\vec{B}$.

Definition 2.6. Es seien V und W Vektorräume mit angeordneten Basen $\vec{B} = (b_1, \dots, b_n)$, $n \in \mathbb{N} \setminus \{0\}$, $\vec{C} = (c_1, \dots, c_m)$, $m \in \mathbb{N} \setminus \{0\}$ und

$$s \colon V \times W \to K$$

sei bilinear oder sesquilinear. Dann heißt

$$\mathrm{Mat}^{\vec{B},\vec{C}}(s) := (s(b_i, c_j)_{1 \leq i \leq n, 1 \leq j \leq m})$$

die darstellende Matrix von s bezüglich $\vec{B}$, $\vec{C}$. Falls $V = W$, nimmt man zweimal dieselbe Basis und schreibt dann $\mathrm{Mat}^{\vec{B}}(s)$.

Satz 2.7 (Transformationsformel für Bilinearformen)**.** *Es seien $\vec{B}$, $\vec{C}$, s wie oben.*

(1) Die Bilinearform s ist durch ihre darstellende Matrix bestimmt.

(2) Falls s symmetrisch oder hermitesch ist und $V = W$ mit Basis $\vec{B}$ ist, gilt $\mathrm{Mat}^{\vec{B}}(s)^T = \overline{\mathrm{Mat}^{\vec{B}}(s)}$.

(3) Sei $\vec{D} = (d_1, \dots, d_n)$ eine weitere Basis von V, und es gelte $b_j = \sum_{i=1}^n t_{i,j} d_i$, also $\mathrm{Mat}_{\vec{B}}^{\vec{D}}(\mathrm{id}) = (t_{i,j})$. Zusätzlich sei $\vec{F} = (f_1, \dots, f_m)$ eine weitere Basis von W, und es gelte für $j = 1, \dots, m$, $c_j = \sum_{i=1}^m u_{i,j} f_i$, also $\mathrm{Mat}_{\vec{C}}^{\vec{F}}(\mathrm{id}) = (u_{i,j})$.
Dann gilt

$$\mathrm{Mat}^{\vec{B},\vec{C}}(s) = (\overline{\mathrm{Mat}_{\vec{B}}^{\vec{D}}}(\mathrm{id}))^T \cdot \mathrm{Mat}^{\vec{D},\vec{F}}(s) \cdot \mathrm{Mat}_{\vec{C}}^{\vec{F}}(\mathrm{id}) \tag{2.2}$$

Beweis: (3) Es folgt für $i = 1, \dots, n$, $j = 1, \dots, m$,

$$s(b_i, c_j) = s(\sum_{k=1}^n t_{k,i} d_k, \sum_{\ell=1}^m u_{\ell,j} f_\ell) = \sum_{k,\ell} \bar{t}_{k,i} s(d_k, f_\ell) u_{\ell,j}.$$

Dies ist genau die (i,j)-Koordinate der Matrix

$$\mathrm{Mat}^{\vec{B},\vec{C}}(s) = (\overline{\mathrm{Mat}_{\vec{B}}^{\vec{D}}}(\mathrm{id}))^T \cdot \mathrm{Mat}^{\vec{D},\vec{F}}(s) \cdot \mathrm{Mat}_{\vec{C}}^{\vec{F}}(\mathrm{id})$$

Hier ist noch ein Diagramm. Es seien $\Phi_{\vec{B}}(v) = (\alpha_i)_i \in k^n$ für $v = \sum_{i=1}^n \alpha_i b_i$ usf. die Koordinatisierungen bezüglich der angeordneten Basen. Teil (3) des Satzes sagt gerade, dass das rot gezeichnete Dreieck kommutiert.

Dann ist das folgende kommutative Diagramm[1] eine alternative Beschreibung des Sachverhalts von Satz 2.7 (3).

$$
\begin{tikzcd}
 & k^n \times k^m \arrow[dd, "\mathrm{Mat}^{\vec{D}}_{\vec{B}}(\mathrm{id})\times\mathrm{Mat}^{\vec{F}}_{\vec{C}}(\mathrm{id})"] \arrow[dr, "\mathrm{Mat}^{\vec{B},\vec{C}}(s)"] & \\
V \times W \arrow[ur, "\Phi_{\vec{B}}\times\Phi_{\vec{C}}"'] \arrow[dr, "\Phi_{\vec{D}}\times\Phi_{\vec{F}}"] \arrow[rr, bend left=90, "s"] & & k \\
 & k^n \times k^m \arrow[ur, "\mathrm{Mat}^{\vec{D},\vec{F}}(s)"'] &
\end{tikzcd}
$$

Für die Beschriftung des senkrechten Pfeils haben wir die Spaltenform genommen. Es sei $x \in k^n$. Wenn $y = \mathrm{Mat}^{\vec{D}}_{\vec{B}}(\mathrm{id}) \cdot x$, so ist $\bar{y}^T = \bar{x}^T \overline{\mathrm{Mat}^{\vec{D}}_{\vec{B}}(\mathrm{id})}^T$, und letzteres ist in der Formel (2.2) gefragt. □

Definition 2.8. Sei $M \in M_n(\mathbb{K})$.

(1) M heißt symmetrisch/ hermitesch, wenn $M^T = \bar{M}$.

(2) M heißt positiv definit, wenn die bilineare Abbildung von $V \times V$ nach $\mathbb{K}$, die durch

$$(x, y) \mapsto \bar{x}^T M y$$

gegeben ist, positiv definit ist.

Es gibt Kriterien, wann eine Matrix positiv definit ist. Alle Haupt-Untermatrizen müssen positiv definit sein.

[1] Wir danken den Entwicklern des Pakets xypictures.

2.1.1 Längen und Winkel

Definition 2.9. Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$\|\cdot\|: V \to \mathbb{R}$$

heißt Norm, wenn für alle $v, w \in V$ und $\lambda \in \mathbb{K}$ folgendes gilt:

(1) $\|\lambda v\| = |\lambda| \cdot \|v\|$.

(2) $\|v + w\| \leq \|v\| + \|w\|$.

(3) $\|v\| \geq 0$ und $\|v\| = 0$ gdw $v = 0$.

Definition 2.10. Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$d: V \times V \to \mathbb{R}$$

heißt Metrik, wenn für alle $x, y, z \in V$ folgendes gilt:

(1) $d(x,y) = d(y,x)$.

(2) $d(x,z) \leq d(x,y) + d(y,z)$, Dreiecksungleichung.

(3) $d(x,y) \geq 0$ und $d(x,y) = 0$ gdw $x = y$.

Lemma 2.11. *Es sei $\|\cdot\|$ eine Norm auf V. Dann definiert*

$$(x,y) \mapsto \|x - y\|$$

eine Metrik.

Beispiel 2.12. Die Metrik

$$d(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

wird nicht durch eine Norm definiert.

Lemma 2.13. *Es sei V ein euklidischer/unitärer Vektorraum. Dann wird durch*

$$\|v\| = \sqrt{\langle v, v \rangle}$$

eine Norm definiert. Die Cauchy-Schwarz'sche Ungleichung vergleicht die Norm mit dem Betrag des Skalarprodukts:

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|.$$

Beweis: Wir zeigen zuerst $||\lambda v|| = |\lambda| \cdot ||v||$. Hier nutzt man

$$\lambda\bar{\lambda} = (\mathrm{Re}(\lambda) + i\mathrm{Im}(\lambda)) \cdot (\mathrm{Re}(\lambda) - i\mathrm{Im}(\lambda)) = (\mathrm{Re}(\lambda))^2 + (\mathrm{Im}(\lambda))^2 = |\lambda|^2.$$

Die Dreieckungleichung ist äquivalent to

$$2\mathrm{Re}(\langle v, w\rangle) \le 2||v|| \cdot ||w||.$$

Dies wird aus der Cauchy-Schwarz-Ungleichung folgen, die wir nun ohne Benutzung der Dreiecksungleichung zeigen. Falls $w = 0$, ist die Cauchy-Schwarz-Ungleichung richtig. Wir nehmen nun $w \neq 0$ and und erhalten mit der binomischen Formel

$$0 \le \langle ||w||^2 v - \overline{\langle v, w\rangle} w, ||w||^2 v - \overline{\langle v, w\rangle} w\rangle = ||w||^4 ||v||^2 - ||w||^2 |\langle v, w\rangle|^2.$$

Nun dividieren wir durch $||w||^2$ und sind fertig.

□

Definition 2.14. Es sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer oder ein unitärer Vektorraum. Zwei Vektoren v, w heißen orthogonal , wenn $\langle v, w\rangle = 0$. Man schreibt hierfür auch $v \perp w$. Für eine Teilmenge M von V definieren wir

$$M^\perp = \{v \in V \ : \ \forall m \in M, \langle m, v\rangle = 0\}.$$

Bemerkung 2.15. Es gilt Da $M^\perp = (\mathrm{span}(M))^\perp$ und $\mathrm{span}(M) + M^\perp = V$ und $\mathrm{span}(M) \cap M^\perp = \{0\}$. Daher wird $M^\perp$ auch das orthogonale Komplement von M genannt, denn $M^\perp$ ist ein Komplementärraum zum Unterraum $\mathrm{span}(M)$.

Definition 2.16. Es seien M, N Teilmengen von $(V, \langle\cdot,\cdot\rangle)$. Dann schreiben wir $M \perp N$, falls für jedes $m \in M$ und jedes $n \in N$ gilt $m \perp n$. (Man sieht leicht, dass dann $\mathrm{span}(M) \perp \mathrm{span}(N)$ folgt.)

Definition 2.17. Es seien $x, y \in V \setminus \{0\}$, V mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Der Öffnungswinkel α ist definiert durch

$$\cos(\alpha) = \frac{\langle x, y\rangle}{||x|| \cdot ||y||} \text{ und } \alpha \in [0, \pi].$$

Der Bruch liegt in [-1,1] und daher ist $\cos(\alpha)$ wohldefiniert.

2.1.2 Orthonormalbasen

Definition 2.18. Es sei (V, s) ein endlichdimensionaler euklidischer bzw. unitärer Vektorraum. Eine Menge $B = \{b_1, \dots, b_n\}$ heißt orthonormal, wenn jedes b_i die s-Norm 1 hat und für $i \neq j$ im Sinn von s, $b_i \perp b_j$.

Die Menge B heißt Orthonormalbasis von V, wenn sie orthonormal und zusätzlich eine Basis von V ist.

Bemerkung 2.19. Die darstellende Matrix eines Skalarprodukts ist in einer Darstellung bezüglich einer angeordneten Orthonormalbasis gerade die Einheitsmatrix.

Im $\mathbb{R}^n$ ist die Standardbasis eine Orthonormalbasis.

Lemma 2.20. *Es sei $\vec{B} = (b_1, \dots, b_n)$ orthonormal in $(V, \langle\cdot,\cdot\rangle)$. Dann ist B linear unabhängig. Für jedes $x \in \operatorname{span}(B)$ gilt*

$$x = \sum_{i=1}^{n} \langle b_i, x\rangle b_i.$$

Beweis: Es sei $x \in \operatorname{span}(B)$, also $x = \sum_{i=1}^{n} \alpha_i b_i$. Dann ist $\langle b_j, x\rangle = \alpha_j$. Dies beweist die Formel. Setzt man nun speziell $x = 0$ ein, so erhält man für $j = 1, \dots, n$, $\alpha_j = 0$. Die Vektoren $b_1, \dots, b_n$ sind also linear unabhängig. □

Satz 2.21 (Gram-Schmidt'sches Orthonormalisierungsverfahren)**.** *Es sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer bzw. ein unitäre Vektorraum, und es seien $b_1, \dots, b_n$ linear unabhängig. Wir definieren durch Rekursion über $1 \leq i \leq n$ die folgenden Vektoren*

$$\begin{aligned}
\tilde{b}_1 &= \frac{b_1}{\|b_1\|}, \\
c_{i+1} &= b_{i+1} - \sum_{j=1}^{i} \langle \tilde{b}_j, b_{i+1}\rangle \tilde{b}_j, \\
\tilde{b}_{i+1} &= \frac{c_{i+1}}{\|c_{i+1}\|}.
\end{aligned}$$

Dann bilden die so definierten Vektoren $\tilde{b}_1, \dots, \tilde{b}_n$ ein Orthonormalsystem mit selbem Erzeugnis wie $b_1, \dots, b_n$.

Beweis: Induktiv über $1 \leq i \leq n$ zeigen wir die entsprechende Behauptung für $b_1, \dots, b_i$.

Für $i = 1$ ist nichts zu zeigen. Seien $\tilde{b}_1, \dots \tilde{b}_i$ gewählt und b_{i+1} gegeben. Da $b_{i+1} \notin \operatorname{span}(b_1, \dots, b_i)$ ist b_{i+1} auch nicht im $\operatorname{span}(\tilde{b}_1, \dots, \tilde{b}_i)$. Dann ist $c_{i+1} \neq 0$. Durch Einsetzen in die Gleichung für c_{i+1} sieht man $c_{i+1} \perp \tilde{b}_j$ für $j = 1, \dots, i$. Wie im Beweis des Steinitz'schen Satzes sieht nan $\operatorname{span}(c_{i+1}, \tilde{b}_1, \dots, \tilde{b}_i) = \operatorname{span}(b_1, \dots, b_i, b_{i+1})$. Nun normiert man c_{i+1} noch und erhält dann $\tilde{b}_{i+1}$ mit $\operatorname{span}(\tilde{b}_{i+1}, \tilde{b}_1, \dots, \tilde{b}_i) = \operatorname{span}(b_1, \dots, b_i, b_{i+1})$.

□

Durch Anwendung des Verfahrens auf eine Basis findet man also eine Orthormalbasis.

Korollar 2.22. *Jeder endlich-dimensionale euklidische Vektorraum/ unitäre Vektorraum hat eine Orthonormalbasis.*

Satz 2.23. *Es sei V ein endlichdimensionaler euklidischer/unitärer Vektorraum und es sei $M \subseteq V$. Dann ist* $\operatorname{span}(M) \oplus M^\perp = V$.

Beweis: Wir nehmen eine Orthonormalbasis $b_1, \dots, b_k$ von $\operatorname{span}(M)$ und ergänzen diese zu einer Orthonormalbasis $b_1, \dots, b_k, b_{k+1}, \dots, b_n$ von V. Dann ist $M^\perp = \operatorname{span}(b_{k+1}, \dots, b_n)$. □

2.1.3 Hilberträume und Orthonormalbasen in Hilberträumen

Das Wort „Orthonormalbasis " hat leider zwei verschiedene Belegungen. Zur Vermeidung von Missverständnissen nennen wir noch den anderen Gebrauch.

Definition 2.24. Ein Hilbertraum ist ein Vektorraum über dem Körper der reellen oder komplexen Zahlen, versehen mit einem Skalarprodukt, der vollständig bezüglich der vom Skalarprodukt induzierten Norm ist. Dies heißt, jede Cauchy-Folge von Vektoren konvergiert. Eine Cauchy-Folge ist wie in der Analysis definiert: $(v_n)_n$, so dass für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$\forall n, m \geq n_0, \|v_n - v_m\| < \varepsilon.$$

Definition 2.25. Eine Teilmenge $\{b_i : i \in I\}$ eines Hilbertraums heißt Orthonormalsystem, falls für alle $i, j \in I$, $\langle b_i, b_j \rangle = \delta_{i,j}$.

Ein Orthonormalsystem B, das dicht liegt in H, heißt Orthomormalbasis (im Hilbertraumsinn) von H. Die Eigenschaft „B liegt dicht in H" ist definiert durch:

$$\begin{aligned}&\forall h \in H, \forall \varepsilon > 0, \exists n \in \mathbb{N},\\ &\exists \{b_{i_1}, \dots, b_{i_n}\} \subseteq I, \exists \alpha_i \in \mathbb{K}, \|\sum_{i=1}^{n} \alpha_i b_i - h\| < \varepsilon.\end{aligned}$$

Die endlich-dimensionalen Hilberträume sind (modulo Isometrie, s.u.) gerade die euklidischen und die unitären Räume $\mathbb{R}^n$ und $\mathbb{C}^n$. Der Polynomraum mit dem Standardskalarprodukt $\langle x^i, x^j \rangle = \delta_{i,j}$ ist nicht vollständig. Im Vektorraumsinn sind die Hilberträume immer mindestens $2^{\aleph_0}$-dimensional. Viele Rechnungen werden jedoch mit einem abzählbaren dichten Orthonormalsystem vorgenommen. Wenn es so eines gibt, nennt man den Hilbertraum separabel.

2.1.4 Orthogonale und unitäre Abbildungen

Definition 2.26. Es seien V und W euklidische/unitäre Vektorräume und es bezeichne $\langle\cdot,\cdot\rangle$ das Skalarprodukt auf V und auch auf W. Ein Homomorphismus $f\colon V\to W$ heißt orthogonal/ unitär , falls für alle $x,y\in V$,

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

Falls f zusätzlich surjektiv ist, heißt f Isometrie oder Isomorphismus von euklidischen/unitären Vektorräumen.

Bemerkung 2.27. Für jeden orthogonalen/unitären Homomorphismus $f\colon V\to W$ gilt für jedes $v\in V$, $\|f(v)\|_W = \|v\|_V$. Auch Winkel werden erhalten. Aus Längen- und Winkeltreue folgt wiederum die Orthogonalität/Unitarität.

Beispiel 2.28. Die Drehung des $\mathbb{R}^2$ um α. Es sei

$$M_\alpha = \begin{pmatrix}\cos(\alpha) & \sin(\alpha)\\ -\sin(\alpha) & \cos(\alpha)\end{pmatrix}.$$

Dann ist $M_\alpha^T\cdot M_\alpha = E_2$. Wir haben

$$\langle M_\alpha x, M_\alpha y\rangle = (M_\alpha x)^T M_\alpha y = x^T M_\alpha^T M_\alpha y = x^T y = \langle x,y\rangle.$$

Hier ist noch ein Korollar aus dem Korollar 2.22

Korollar 2.29. *Es sei V ein n-dimensionaler euklidischer/unitärer Vektorraum. Dann ist V isometrisch zum $\mathbb{K}^n$ mit dem Standardskalarprodukt.*

Beweis: Es sei $\vec{B} = (b_1,\dots b_n)$ eine angeordnete Orthonormalbasis von V. Dann setzen wir $f\colon\mathbb{K}^n\to V$ an durch

$$f(\sum_{i=1}^n \alpha_i e_i) = \sum_{i=1}^n \alpha_i b_i.$$

Man rechnet leicht nach, dass f eine Isometrie ist. □

Lemma 2.30. *Es sei $f\colon V\to W$ orthogonal / unitär.*

(1) Dann erhält f die Orthogonalität und die Norm und ist injektiv und, falls $\dim(V)=\dim(W)$ *endlich ist, so ist f ein Isomorphismus.*

(2) Falls $V=W$ und λ ein Eigenwert von f ist, so ist $|\lambda|=1$.

(3) Falls f ein Isomorphismus ist, so ist f^{-1} auch orthogonal/unitär.

(4) Die Komposition zweier orthogonaler/unitärer Homomophismen ist orthogonal/unitär.

Korollar 2.31. *Die Menge der orthogonalen/unitären Automorphismen eines euklidischen Vektorraums ist ist eine Gruppe.*

Definition 2.32.

(1) Es sei V euklidisch. Die orthogonale Gruppe von V is die Gruppe $O(V)$ der orthogonalen Isomorphismen von V. $O(\mathbb{R}^n)$ heißt auch $O(n)$.

(2) Es sei V unitär. Die unitäre Gruppe von V is die Gruppe $U(V)$ der unitären Isopmorphismen von V. $U(\mathbb{C}^n)$ heißt auch $U(n)$.

Aus dem Lemma 2.30 (1) folgt, dass $O(n) \subseteq \mathrm{GL}_n(\mathbb{R})$ und $U(n) \subseteq \mathrm{GL}_n(\mathbb{C})$.

Erinnerung $f_M(x) = Mx$ ist die lineare Abbildung von $\mathbb{K}^n$ in $\mathbb{K}^n$, die durch die Multiplikation von M vermittelt wird.

Lemma 2.33. *Es sei $M \in M_n(\mathbb{K})$. Dann ist f_M orthogonal/unitär genau dann, wenn*

$$\bar{M}^T M = E_n.$$

Dies ist genau dann der Fall, wenn die Spalten von M eine Orthonormalbasis von $\mathbb{R}^n$ bezüglich des Standardskalarprodukts bilden.

Übung 2.34. Man überlege sich, wie man Orthogonalität/Unitarität bei nicht notwendig quadratischen Matrizen durch Transponierte charakterisiert.

Definition 2.35. Es sei $M \in M_{n,n}(\mathbb{R})$ und $\bar{M}^T M = E_n$. Dann heißt M orthogonal/unitär.

Ein quadratische Matrix M is also orthogonal/unitär genau dann, wenn M bezüglich der Standardbasis des $\mathbb{K}^n$ eine orthogonale/unitäre Abbildung darstellt.

Beispiel 2.36. (1) $O(1) = \{1, -1\}$.

(2) $O(2) = \{M_\alpha : \alpha \in \mathbb{R}\} \cup \{N_\alpha : \alpha \in \mathbb{R}\}$, mit $N_\alpha = \begin{pmatrix} \sin(\alpha) & \cos(\alpha) \\ \cos(\alpha) & -\sin(\alpha) \end{pmatrix}$.

(3) $U(1) = \{z \in \mathbb{C} : |z| = 1\} = \{e^{i\alpha} : 0 \leq \alpha < 2\pi\}$.

Beweis: (1) Es sei $M = (a) \in O(1)$. Dann ist $|a| = 1$, also $s = -1$ oder $a = 1$. (2) Im Falle von $\mathbb{K} = \mathbb{C}$ folgt aus $|a| = 1$, dass $a = e^{i\alpha}$ für ein $\alpha \in [0, 2\pi)$.

(2) Sei $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(2)$. Dann ist $\|Me_1\|^2 = a^2 + c^2 = 1$ und $\|Me_2\|^2 = b^2 + d^2 = 1$ und $Me_1 \perp Me_2$, also $ab + cd = 0$. Wir setzen $a = \cos(\alpha)$. Dann ist $c = \pm\sin(\alpha)$. Wir nehmen obdA den positiven Fall. Aus der Gleichung $ab + cd = 0$ folgt dann, dass

$$(b, d) = \lambda(\sin(\alpha), \cos(\alpha))$$

für ein $\lambda \in \mathbb{R}$ mit $|\lambda|^2 = 1$. Also $\lambda = \pm 1$. Wir erhalten gerade die M_α und die N_α. Übrigens is $N_\alpha = M_\alpha \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = M_{\pi/2+\alpha} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
(3) $\overline{e^{i\alpha}} = \cos(\alpha) - i\sin(\alpha) = e^{-i\alpha} = \frac{1}{e^{i\alpha}}$. Es ist also $\bar{a} = \overline{e^{i\alpha}} = (e^{i\alpha})^{-1} = \frac{1}{a}$. Die Einermatrix der konjugierten Zahl $(\bar{a})$ ist die inverse Matrix der unitären 1×1-Matrix (a). □

Übung 2.37. Gibt es eine nichttiviale (d.h., nicht einelementige) echte Untergruppe on $O(2)$? Denken Sie an die Orientierung.

2.2 Der Spektralsatz

2.2.1 Selbstadjungierte Abbildungen

Definition 2.38. Es sei V ein euklidischer/unitärer Vektrorraum. Ein Endomorphismus f von V heißt selbstadjungiert, falls für alle $x, y \in V$,

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

Beispiel 2.39. Es sei $V = \mathbb{K}^n$ mit dem Standardskalarprodukt und es sei $f = f_M$ durch die Matrix M definiert und selbstadjungiert. Dann ist

$$\bar{M}^T = M,$$

denn

$$\bar{x}^T My = \langle x, My \rangle = \langle Mx, y \rangle = \overline{Mx}^T y = \bar{x}^T \bar{M}^T y$$

für alle $x, y \in \mathbb{K}^n$. Nun lässt man x, y über e_i, e_j rangieren und sieht $\bar{M}^T = M$. Die Matrix M ist also symmetrisch/hermitesch.

Beispiel 2.40. Ein Beispiel für einen Zustandsraum in der Quantenmechanik.

$$V = \{g : \mathbb{R}^3 \to \mathbb{C} \, : \, \int_{\mathbb{R}^3} |g|^2 < \infty\}.$$

Das Beipiel 2.39 lässt sich auf beliebige endlichdimensionale V übertragen wir folgt:

Lemma 2.41. *Es sei V ein endlich-dimensionaler euklidischer/unitärer Vektorraum und es sei $f : V \to V$ ein Endomorphismus. Dann ist f genau dann selbstadjungiert, wenn die darstellende Matrix von f bezüglich einer Orthonormalbasis von V eine symmetrische/hermitesche Matrix ist.*

Beweis: Sei $b_1, \dots, b_n$ eine angeordnete Orthonormalbasis von V und sei $f \in \operatorname{End}(V)$ und $M = \operatorname{Mat}_{\vec{B}}^{\vec{B}}(f)$, also [2] für jedes j, $f(b_j) = \sum_i m_{i,j} b_i$. Diese Formel ergibt nun, weil $\vec{B}$ eine Orthonormalbasis ist, ist für jedes i, j,

$$\langle b_i, f(b_j) \rangle = {b_i}^T M b_j = m_{i,j}.$$

Nun ist $\langle f(b_i), b_j \rangle = \langle b_i, f(b_j) \rangle$ gdw $\bar{m}_{j,i} = m_{i,j}$. □

Satz 2.42 (Der Spektralsatz)**.** *Es sei V ein endlich-dimensionaler euklidischer/unitärer Vektorraum und es sei $f \colon V \to V$ ein selbstadjungierter Endomorphismus. Dann hat V eine Orthonormalbasis aus Eigenvektoren von f und alle Eigenwerte von f sind reell.*

Beweis: Wir zeigen zuerst das letztere. Sei $v \in V$ ein Eigenvektor zum Eigenwert λ, also $f(v) = \lambda v$. Dann ist

$$\langle v, f(v) \rangle = \lambda |v|^2 = \langle f(v), v \rangle = \langle \lambda v, v \rangle = \bar{\lambda} |v|^2.$$

Also ist $\lambda = \bar{\lambda} \in \mathbb{R}$.

Nun hat das charakteristische Polynom von f immer eine Nullstelle in $\mathbb{C}$. Da aber $\lambda \in \mathbb{R}$, gibt es auch eine Nullstelle in $\mathbb{R}$. Auch im euklidischen Fall hat f also einen Eigenwert $\lambda \in R$ und einen Eigenvektor v zu λ.

Wir stellen fest, dass $U = \{v\}^\perp$ ein f-invarianter Unterraum ist: Es sei $u \perp v$. Dann ist

$$\langle f(u), v \rangle = \langle u, f(v) \rangle = \langle u, \lambda v \rangle = \lambda \langle u, v \rangle = 0,$$

also auch $f(u) \perp v$.

Nun beweisen wir den Satz induktiv über die Dimension. Für die Dimension 1 nehmen wir einen normierten Eigenvektor v zum reellen Eigenwert λ und haben so die Orthonormalbasis $\{v\}$ und die Darstellung $\operatorname{Mat}_{\vec{B}}^{\vec{B}}(f) = (\lambda)$.

Im Schritt von n auf $n+1$ suchen wir zuerst einen Eigenvektor v_{n+1} zu einem reellen Eignewert λ_{n+1}.

Dann verwenden wir das Kästchenlemma (Satz 1.17) auf den invarianten Unterraum $\operatorname{span}(\{v_{n+1}\})$ und den f-invarianten Unterraum $U = \{v_{n+1}\}^\perp$ an und erhalten bezüglich irgendeiner Orthonormalbasis $\vec{B}$ von U eine Darstellung von f bezüglich $\vec{B}{}^\frown v_{n+1}$ der Form

$$\begin{pmatrix} M & 0 \\ 0 & \lambda_{n+1} \end{pmatrix},$$

und es gilt $M^T = \bar{M}$. Nun können wir auf dem n-dimensionalen Vektorraum U auf die Abbildung, die durch M beschrieben wird, also auf $f \restriction U$, die Induktionsvoraussetzung

[2] nach der definierenden Gleichung für $\operatorname{Mat}_{\vec{B}}^{\vec{B}}(f)$ aus LA 1

anwenden und finden eine angeordnete Orthonormalbasis $(v_1, \dots, v_n)$ von U, so dass $f \upharpoonright U$ bezüglich dieser Diagononalgestalt mit reellen Einträgen auf der Diagonale hat. Bezüglich $(v_1, \dots, v_n, v_{n+1})$ hat f also die Darstellung

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \lambda_{n+1} \end{pmatrix}.$$

□

2.2.2 Adjungierte Abbildungen

Achtung, nun wird ein Zeichen recycled. Wir nehmen f^* nun nicht im dualen Sinne zu f sondern im folgenden Sinn.

Definition 2.43. Es sei $f: V \to W$ eine lineare Abbildung zwischen zwei euklidischen/unitären Vektorräumen V und W, und das jeweilige Skalarprodukt wird mit $\langle \cdot, \cdot \rangle$ bezeichnet. [3]

$f^*: W \to V$ heißt zu f adjungiert, genauer "adjungiert bezüglich $\langle \cdot, \cdot \rangle_V$ und $\langle \cdot, \cdot \rangle_W$", falls für alls $v \in V$, $w \in W$,

$$\langle w, f(v) \rangle = \langle f^*(w), v \rangle.$$

Bemerkung 2.44. Falls f selbstadjungiert ist, so ist $f^* = f$. Ist f^* adjungiert to f, so ist auch f adjungiert zu f^*.

Lemma 2.45. *f^* ist durch die definierende Bedingung eindeutig bestimmt.*

Beweis: Aus $\langle x, v \rangle = 0$ für alle $v \in V$ folgt $x = 0$. Nun setzt man $x = (f^*)_1(w) - (f^*)_2(w)$ für zwei adjungierte Abbildungen $(f^*)_1, (f^*)_2$. □

Lemma 2.46. *Wenn die beiden Vektorräume endlich-dimensional ist, so existiert f^*.*

Beweis: Es seien $\vec{B} = (b_i \, : \, i = 1, \dots, n)$ und $\vec{C} = (c_i \, : \, i = 1, \dots, m)$ ONB. Für $j = 1, \dots, m$ ist

$$f^*(c_j) = \sum_{i=1}^{n} \langle f^*(c_j), b_i \rangle_V \cdot b_i = \sum_{i=1}^{n} \langle c_j, f(b_i) \rangle_W \cdot b_i.$$

□

Übung 2.47. Man kann die Existenz auch herleiten, wenn nur einer der beiden Räume endlichdimensional ist.

[3] Wer möchte, kann $\langle \cdot, \cdot \rangle_V$ schreiben und selbes für W

Beweis: Skizze: Falls V endlich-dimensional und $\vec{B} = (b_1, \ldots, b_n)$ ist, funktioniert die Formel für beliebiges $j \in J$, $\vec{C} = (c_j : j \in J)$. Falls W endlich-dimensional ist, ist $V/\ker(f)$ endlich-dimensional und hat eine Basis $b'_1 + \ker(f)$, $\ldots$, $b'_m + \ker(f)$. Gram-Schmidt auf irgendwelche Repräsentanten, z.B. auf die b'_1, $\ldots$, b'_m, liefert dann eine ONB eines m-dimensionalen Teilraums T von V, für den das Bild von $f \restriction T$ schon der Bildraum B ist. Man ergänzt den Bildraum um einen orthogonalen Komplementärraum B'. Auf $f \restriction T$ und $f^* \restriction B$ kann man die Formel anwenden. Auf allen Vektoren v aus einem orthogonalen Komplementärraum von T (so einen gibt es), ist $f(v) = 0$. Auf allen Vektoren $w \in B'$ ist $f^*(w) = 0$. □

Beispiel 2.48. Wenn $V = \mathbb{K}^n$ mit dem Standardskalarprodukt und Orthonormalbasis $\vec{B}$ und ebenso $W = \mathbb{K}^m$ und Orthonormalbasis $\vec{C}$, und $f\colon V \to W$ mit

$$\mathrm{Mat}^{\vec{C}}_{\vec{B}}(f) = M.$$

Dann ist

$$\mathrm{Mat}^{\vec{B}}_{\vec{C}}(f^*) = \bar{M}^T.$$

Definition 2.49. Es sei $M \in M_{n\times n}(\mathbb{K})$. Dann heißt

$$M^* = \bar{M}^T$$

die zu M adjungierte Matrix.

Die Formel liefert im Endlichdimensionalen einen Existenzbeweis. Im Unendlichdimensionalen wird es wie bei den Dualräumen und beim Dualisieren der Homomorphismen (in der Vorlesung Linear Algebra 1) wieder trickreich. Wir beschränken uns hier auf endlichdimensionale Vektorräume. [4]

Korollar 2.50. *Es seien V und W euklidische/unitäre Vektorräume mit Orthonormalbasen $(b_1, \ldots, b_n)$ und $(c_1, \ldots, c_m)$. Es sei f linear mit darstellung M bezüglich dieser beiden angeordneten Basen. Dann hat die adjungierte Abbildung f^* als Darstellung bezüglich $(c_1, \ldots, c_m)$ und $(b_1, \ldots, b_n)$ die adjungierte Matrix $M^* = \bar{M}^T$.*

2.2.3 Normale Endomorphismen

Wir betrachten nun wieder $V = W$. Vorsicht, nun brauchen wir $\mathbb{K} = \mathbb{C}$. Ohne Nullstellen geht es nicht.

[4]Etliche Beweise, die induktiv über die Dimension des Vektorraums geführt werden, haben im Unendlichdimensionalen schlichtweg keine Analoga.

Definition 2.51. Es sei V ein endlich-dimensionaler unitärer Vektorraum und es sei $f \in \mathrm{End}(V)$. Dann heißt f normal, wenn

$$f^* \circ f = f \circ f^*.$$

Beispiel 2.52.

(1) Jeder selbstadjungierte Endomorphismus ist normal.

(2) Jeder unitäre Endomorphismus ist normal.

Satz 2.53 (Spektralsatz für normale Endomorphismen)**.** *Es sei V ein endlich-dimensionaler unitärer Vektorraum und es sei $f \in \mathrm{End}(V)$ normal. Dann hat V eine angeordnete Orthonormalbasis aus gemeinsamen Eigenvektoren von f und von f^*. Wenn $f(v) = \lambda v$, so $f^*(v) = \bar{\lambda} v$.*

Beweis: Diesmal arbeiten wir in $\mathbb{C}$. Es sei $n \geq 1$ die Dimension von V. Dann hat f einen Eigenvektor v zu einem Eigenwert λ. Wir zeigen, dass v auch Eigenvektor von f^* ist:

Es sei $V_\lambda = \{v \in V \,:\, f(v) = \lambda v\}$. Dann ist für $w \in V_\lambda$

$$f(f^*(w)) = f^*(f(w)) = f^*(\lambda w) = \lambda f^*(w).$$

Also ist auch $f^*(w) \in V_\lambda$. Der Eigenraum V_λ ist f^*-invariant. Die Abbildung $f^* \restriction V_\lambda \to V_\lambda$ hat einen Eigenvektor $v \in V_\lambda$ zu einem f^*-Eigenwert μ. Dann ist

$$\bar{\mu}\langle v, v\rangle = \langle \mu v, v\rangle = \langle f^*(v), v\rangle = \langle v, f(v)\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle,$$

also $\bar{\mu} = \lambda$, da $\langle v, v\rangle \neq \varnothing$.

Wie oben haben wir, dass $U = \{v\}^\perp$ sowohl f-invariant als auch f^*-invariant ist. Wir nehmen $\|v\| = 1$ an. Nach dem Kästchenlemma (in der Form, dass sowohl $\mathrm{span}(\{v\})$ also auch U beide f-invariant sind) hat f bezüglich einer angeordneten Orthonormalbasis $\vec{B}$ von U und v daran angehängt die Darstellung

$$\begin{pmatrix} M & 0 \\ 0 & \lambda \end{pmatrix}$$

und f^* die Darstellung

$$\begin{pmatrix} N & 0 \\ 0 & \bar{\lambda} \end{pmatrix}$$

und $MN = NM$.

Wir haben, dass U unter f invariant ist. Dies sieht man wie folgt: Es sei $\langle v, u\rangle = 0$. Dann ist

$$\langle v, f(u)\rangle = \langle f^*(v), u\rangle = \lambda\langle v, u\rangle = 0.$$

Ebenso sieht man, dass U unter f^* invariant ist und $(f \upharpoonright U)^* = f^* \upharpoonright U$.

Wir können daher auf $f \upharpoonright U$ und $f^* \upharpoonright U$ die und den niedriger-dimensionalen Vektorraum U die Induktionsvoraussetzung anwenden und erhalten für U eine Orthonormal-Basis $\vec{C}$ aus gemeinsamen Eigenvektoren für f und für f^*. Insgesamt haben wir bezüglich $\vec{C}\frown v$ die Darstellung von f bzw f^*

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & 0 \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \qquad \begin{pmatrix} \overline{\lambda_1} & 0 & \dots & 0 \\ 0 & \overline{\lambda_2} & 0 & 0 \\ \vdots & & \ddots & \\ 0 & \dots & 0 & \overline{\lambda_n} \end{pmatrix}.$$

□

Korollar 2.54. *Es sei V ein endlichdimensionaler euklidischer bzw unitärer Vektorraum und es ei $f \in \mathrm{End}(V)$ orthogonal/ unitär. Dann gilt* $\det(f) = \pm 1$ / $|\det(f)| = 1$.

Beweis: Für $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$ wenden wir den Spektralsatz 2.42 an und sehen, dass es eine Basis aus Eigenvektoren gibt. Jeder Eigenwert hat den Betrag 1, also hat $\det(f) = \prod \lambda_i$ den Betrag 1. Es ist $|a \cdot b| = |a| \cdot |b|$. Nun spezialisieren wir auf $\mathbb{R}$ und wissen, dass es eine Basis aus Eigenvektoren zu reellen Eigenwerten gibt, da f sowohl selbstadjungiert als auch unitär ($f \circ f^* = 1$) ist. □

2.2.4 Hauptachsentransformation

Wiederholung:

Definition 2.55. Es sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum

$$\varphi: V \times V \to \mathbb{K}$$

heißt bilinear/sesquilinear oder Bilinearform/Sesquilinearform wenn φ in der zweiten Koordinate linear und in der ersten sesquilinear ist, d.h.

$$\varphi(\lambda v, w) = \bar{\lambda}\varphi(v, w)$$

im Falle $\mathbb{K} = \mathbb{C}$.

Eine Bilinearform/ Sequilinearform φ heißt symmetrisch/hermitesch, falls

$$\varphi(v, w) = \overline{\varphi(w, v)}.$$

Es sei $\vec{B}$ eine angeordnete Basis von V.

$$M = \mathrm{Mat}_{\vec{B},\vec{B}}(\varphi) := (\varphi(b_i, b_j))_{1\le i,j\le n}$$

heißt die darstellende Matrix von φ bezüglich $\vec{B}$.

Dies ist eine symmetrische/ hermitesche Matrix.

Für Basiswechsel mit einer regulären Matrix S^{-1} gilt nach Satz 2.7

$$\mathrm{Mat}_{\vec{B}',\vec{B}'}(\varphi) := \bar{S}^T MS = S^* MS$$

Nun können wir Satz 6.9 aus Lineare Algebra 1 noch einmal beweisen

Satz 2.56 (Satz 6.9 aus LA I)**.** *Es sei V eine endlichdimensionaler euklidischer/unitärer $\mathbb{K}$-Vektorraum und es sei $s: V \times V \to \mathbb{K}$ eine symmetrische Bilinearform/ hermitesche Sesquilinearform. Es sei M die Darstellung von s bezüglich irgendeiner angeordneten Orthonormalbasis $\vec{B}$, d.h., $s(b_i, b_j) = b_i^T M b_j$. Dann gibt es eine orthogonale/unitäre Basiswechselmatrix S, und eine neue Basis $\vec{C}$, so dass $N = \mathrm{Mat}_{\vec{C},\vec{C}}(s)$*

$$N = S^* MS = S^{-1} MS$$

Diagonalgestalt hat.

Beweis: Aus dem Spektralsatz 2.42 wissen wir, dass es für M und f_M (Erinnerung $f_M(x) = Mx$ für $x \in k^n$) eine (bzgl des Skalarprodukts auf V) Orthonormalbasis $(c_1, \dots, c_n)$ aus Eigenvektoren gibt, also $Mc_i = \lambda_i c_i$. Dann sei S die reguläre Matrix $\mathrm{Mat}^{\vec{B}}_{\vec{C}}(\mathrm{id})$. Da sowohl $\vec{B}$ und $\vec{C}$ Orthonormalbasen sind, ist S orthogonal/unitär. Dann gilt nach dem Transformationssatz aus LA I für die Diagonalmatrix $N = S^{-1}MS = \mathrm{Mat}^{\vec{C}}_{\vec{C}}(f_M)$. Nun haben wir $S^* = S^{-1}$ und daher folgt nach Satz 2.7 die Gleichung $N = \mathrm{Mat}_{\vec{C},\vec{C}}(s)$, denn $M = \mathrm{Mat}_{\vec{B},\vec{B}}(s)$ und N erfüllen die Gleichung $N = S^* MS$. □

Satz 2.57 (Satz 6.10 aus LA I)**.** *Es sei V eine endlichdimensionaler euklidischer/unitärer $\mathbb{K}$-Vektorraum und es sei $\varphi: V \times V \to \mathbb{K}$ eine symmetrische Bilinearform/ hermitesche Sesquilinearform. Es sei M die Darstellung von φ bezüglich irgendeiner angeordneten Basis $\vec{B}$. Dann ist φ positiv definit, genau dann, wenn alle Eigenwerte von M positiv (> 0) sind.*

Beweis: Wir können die Basis so wählen, dass M Diagonalgestalt hat. Dann ist

$$\bar{b}_i M b_j = \lambda_j \delta_{i,j}.$$

Sei nun φ nun positiv definit. Dann ist φ ein Skalarprodukt und λ_i ist die φ-Länge von b_i und also positiv.

Seien umgekehrt alle λ_i positiv. Dann ist M positiv definit und also ein Skalarprodukt. □

Wir erinnern an den Satz von Sylvester.

Satz 2.58 (Der Trägheitssatz von Sylvester[5]). *Zu jeder reellen symmetrischen n-n-Matrix S gibt es eindeutig bestimmte Zahlen p und $q \in \mathbb{N}$ und ein nicht notwendig eindeutiges $M \in \mathrm{GL}(n, K)$, so dass*

$$S = M^\top \begin{pmatrix} 1_{M_{p,p}(K)} & 0 & 0 \\ 0 & -1_{M_{q,q}(K)} & 0 \\ 0 & 0 & 0 \end{pmatrix} M. \tag{2.3}$$

$p + q$ heißt Rang, p wird der positive Trägheitsindex, q der negative Trägheitsindex, $p - q$ wird manchmal der Trägheitsindex oder auch die Signatur genannt.

Beweis: Existenz: Wir nehmen nach Satz 2.56 eine reguläre orthogonale Matrix W, so dass $D = W^\top SW$ eine Diagonalmatrix ist. Wir nehmen W', so dass $w'_{i,j} = 0$ für $i \neq j$ und $w'_{i,i} = \frac{1}{\sqrt{|d_{i,i}|}}$, falls $d_{i,i} \neq 0$, und $w'_{i,i} = 1$ falls $d_{i,i} = 0$. Dann ist $W'^\top DW'$ eine Diagonalmatrix, in der auf der Diagonalen höchstens 1, 0, und -1 vorkommt. Durch Permutation der angeordneten Basis erreicht man nun noch die normierte Diagonalgestalt wie in der Mitte der rechten Seite der Gleichung (2.3).

Wir nehmen an, dass D_1 und D_2 zwei unterschiedliche diagonale normierte Darstellungen für S sind: $D_i = W_i^\top SW_i$ beginne mit p_i Einsen auf der Diagonalen, und dann q_i minus Einsen, und $p_1 < p_2$. Der Rang von D_i ist $p_i + q_i = r$, unabhängig von i, da beide S darstellen. Sei $\varphi_i(x) = x^\top D_i x$. Die φ_i sind wegen der Normierung auf die Längen 1 (der Multiplikation mit W' von oben) jetzt nicht notwendig Darstellungen von φ.

Nun seien $y = (\eta_1, \dots, \eta_n)^\top$ die Koordinaten von $v \in K^n$ nach dem Basiswechsel mit W_1, $x = (\xi_1, \dots, \xi_n)^\top$ nach dem Basiswechsel mit W_2. Sei $U = W_2(W_1^{-1}) \restriction \{1, \dots, r\}^2$ eine reguläre r-r-Matrix. Dann ist

$$S = W_1^\top D_1 W_1 = W_2^\top D_2 W_2$$

und daher

$$y^\top D_1 y = y^\top (W_1^{-1})^\top W_2^\top D_2 W_2 W_1^{-1} y = (Uy)^\top D_2 Uy = \varphi_2(Uy).$$

Wir haben somit

$$\varphi_1(y) = \varphi_2(Uy).$$

Nun seien $x = (\xi_1, \dots, x_n)^\top$ and $y = (\eta_1, \dots, \eta_n)^\top$. Wir setzen nun ein, dass wir die D_i kennen:

$$\begin{aligned} \varphi_1(y) = \varphi_2(Uy) = \varphi_2(x) &= \eta_1^2 + \dots + \eta_{p_1}^2 - \eta_{p_1+1}^2 - \dots - \eta_r^2 \\ &= \xi_1 + \dots + \xi_{p_2}^2 - \xi_{p_2+1}^2 - \dots - \xi_r^2. \end{aligned} \tag{2.4}$$

[5]James Joseph Sylvester, 1814 – 1897

Nun hat das lineare Gleichungssystem $\eta_1 = \cdots = \eta_{p_1} = 0 = \xi_{p_2+1} = \cdots = \xi_r$, in dem die Werte von $(\eta_1, \ldots, \eta_r)$ gesucht sind, eine nicht triviale Lösung $(\eta_1, \ldots, \eta_r)$, denn sein Rang ist $r - (p_2 - p_1) < r$. Für diese $(\eta_1, \ldots, \eta_r)$ wäre aber nach Gleichung (2.4) (in die die $p_1 + q_2$ Nullen aus dem lineares Gleichungssystem eingesetzt werden)

$$-\eta_{p_1+1}^2 - \cdots - \eta_r^2 = \xi_1^2 + \cdots + \xi_{p_2}^2,$$

also doch $\eta_i = 0$ für alle i. □

Beispiel 2.59. Der Minkowski-Raum ist ein vierdimensionaler reeller Vektorraum, auf dem eine symmetrische nicht definite Bilinearform der Form

$$\langle x, y \rangle = -x_0 y_0 + x_1 y_1 + x_2 y_2 + x_3 y_3$$

definiert ist. Hierbei ist die Koordinate $x_0 = c \cdot t$ ebenfalls reell definiert: sie geht mit Hilfe der Lichtgeschwindigkeit c aus der Zeitkoordinate t hervor.

Kapitel 3

Das Tensorprodukt

3.1 Produkte von Räumen und von Vektoren

Definition 3.1. Es seien V, W k-Vektorräume. Das Tensorprodukt $V \otimes W$ ist ein neuer Vektorraum T, der aus den Elementen

$$T = V \bigotimes W = \operatorname{span}\{v \otimes w : v \in V, w \in W\}$$

besteht. Hierbei ist wie immer $\operatorname{span}(M) = \{\sum_{i=1}^{n} \alpha_i m_i : n \in N, m_i \in M, \alpha_i \in k \text{ für } i = 1, \ldots, n\}$ und die leere Summe wird als 0_M interpretiert. Für die Abbildung

$$V \times W \to V \bigotimes W; \qquad (v, w) \mapsto v \otimes w$$

gelten die folgenden Rechenregeln

(1) Für alle $v_1, v_2 \in V$, $\alpha_1, \alpha_2 \in k$, $w \in W$ ist

$$(\alpha_1 v_1 + \alpha_2 v_2) \otimes w = \alpha_1 (v_1 \otimes w) + \alpha_2 (v_2 \otimes w).$$

(2) Für alle $v \in V$, $\alpha_1, \alpha_2 \in k$, $w_1, w_2 \in W$ ist

$$v \otimes (\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 (v \otimes w_1) + \alpha_2 (v \otimes w_2).$$

Vorsicht: Die Abbildung $(v, w) \mapsto v \otimes w$ ist nicht surjektiv, falls $\dim(V) > 1$ und $\dim(W) > 1$. Die Vektoren $v \otimes w$ heißen reine Tensoren. Die Linearkombinationen mit mindestens zwei Summanden, die nicht vereinfacht werden können, heißen zusammengesetzte Tensoren.

Bemerkung 3.2. Aus den obigen Regeln folgt zum Beispiel $\alpha(v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w)$.

Das Tensorprodukt ist durch eine (sogenannte) universelle Eigenschaft eindeutig charakterisiert.

Lemma 3.3. *Ein k-Vektorraum T zusammen mit einer bilinearen Abbildung $\otimes : U \times V \to T$ ist genau dann ein Tensorprodukt der beiden k-Vektorräume V und W, wenn $\otimes$ und T die folgende universelle Eigenschaft haben: Jede bilineare Abbildung $s: V \times W \to Y$ (mit einem weiteren k-Vektorraum Y) lässt sich als Komposition von $\otimes$ mit einer eindeutig bestimmten linearen Abbildung $f: T \to Y$ schreiben.*

$$\begin{tikzcd}
V \times W \arrow[r, "\otimes"] \arrow[dr, "s"'] & T = V \bigotimes W \arrow[d, dashed, "f"] \\
& Y
\end{tikzcd}$$

Beweis: $\Leftarrow$ $\otimes$, T vom Lemma sind ein Tensorprodukt: Man zeigt, dass $\otimes$ und T die Rechenregeln 1) und 2) repektieren. Dies folgt aus der Bilinearität von s, die durch $f \circ \otimes$ treu widergegeben werden muss: $s(\alpha_1 v_1 + \alpha_2 v_2, w) = f((\alpha_1 v_1 + \alpha_1 v_2) \otimes w)$ und andererseits $s(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 s(v_1, w) + \alpha_2 s(v_2, w) = \alpha_1 f(v_1 \otimes w) + \alpha_2 f(v \otimes w) = f(\alpha_1 v_1, \otimes w + \alpha_2 v_2 \otimes w)$. Wenn nun das Argument $\alpha_1 v_1, \otimes w + \alpha_2 v_2 \otimes w = (\alpha_1 v_1 + \alpha_1 v_2) \otimes w)$ ist, hat f auf beiden gleichen Argumenten auch denselben Wert.

Warum muss das Argument in beiden Darstellungen gleich sein? Da man den kleinsmöglichen Raum T definiert, denn man fordert ja dass f eindeutig ist. Dies sieht man wie folgt:

Wenn wir

$$\alpha_1 v_1 \otimes w + \alpha_2 v_2 \otimes w \neq (\alpha_1 v_1 + \alpha_1 v_2) \otimes w$$

annehmen, dann ist

$$d = \alpha_1 v_1 \otimes w + \alpha_2 v_2 \otimes w - (\alpha_1 v_1 + \alpha_1 v_2) \otimes w \neq 0_T$$

und es gibt ein lineares $g: T \to Y$, das $g(d) \neq 0_Y$ und $g(v \otimes w) = 0_Y$ hat. Man kann zum Beispiel $Y = T$ selbst nehmen. Dieses g wird zu keiner Bilinearform gebraucht, um $s(v, w)$ als $f(v \otimes w)$ to schreiben. Man kann g zu jeder Lösung f von $\forall v \in V \forall w \in W s(v, w) = f(v \otimes w)$ auf allen Elementen von T hinzuaddieren und erhält eine neue Lösung. Dies widerspricht der Eindeutigkeit von f.

Wenn $\alpha_1 v_1 \otimes w + \alpha_2 v_2 \otimes w = v' \otimes w'$, $w' \neq w$ für irgendein v', dann gibt es ein bilineares s', so dass $s(v', w') = 0$ für alle v' und $s'(\alpha_1 v_1 + \alpha_2 v_2, w) \neq 0$, für das es nun kein f gibt.

Wenn $\alpha_1 v_1 \otimes w + \alpha_2 v_2 \otimes w = v' \otimes w$, $v' \neq \alpha_1 v_1 + \alpha_2 v_2$, dann gibt es ein bilineares s'', so dass $s''(v', w) = 0$ und $s'(\alpha_1 v_1 + \alpha_2 v_2, w) \neq 0$, für das es nun kein f gibt.

$\Rightarrow$ Jedes Tensorprodukt hat die universelle Eigenschaft. Man setzt $f(v \otimes w) = s(v, w)$. f is wohldefiniert und linear, da s bilinar ist. $\square$

Wie zeigt man die Existenz und die Eindeutigkeit (bis auf Isomorphie) des Tensorprodukts? Durch eine Konstruktion.

Definition 3.4. Sei $\{a_i : i \in I\}$ eine Basis von V. Dann heißt $(a_i)_{i \in I}$ indizierte Basis von V. Mengentheoretisch kann man $(a_i)_{i \in I}$ mit dem Graphen

$$\{(i, a_i) : i \in I\}$$

identifizieren. Man kennt also nicht nur die Elemente der Basis, sondern weiß bei jedem Basisvektor, welchen Index er trägt. Falls I eine lineare Ordnung $(I, <)$ trägt, spricht man auch von einer angeordneten Basis $(a_i)_{i \in I}$.

Beispiel: Es ist also für $a_0 \neq a_1$ die indizierte Basis $\{(0, a_0), (1, a_1)\}$ verschieden von $\{(0, a_1), (1, a_0)\}$.

Satz 3.5. *Je zwei Vektorräume V und W haben ein Tensorprodukt $\otimes V \times W \to T$. Der Raum T ist eindeutig bestimmt in folgendem Sinn: Wenn $\otimes' \colon V \times W \to T'$ ein zweites Tensorprodukt ist, gibt es genau einen Isomorphismus $t \colon T \to T'$, der das Diagramm*

$$\begin{tikzcd}
V \times W \arrow[r, "\otimes"] \arrow[dr, "\otimes'"'] & T \arrow[d, "t"] \\
 & T'
\end{tikzcd}$$

kommutativ macht.

Beweis: Im Falle, dass V oder W der Nullraum ist, ist auch T der Nullraum.

Wir nehmen nun an, dass weder V noch W der Nullraum ist. Wir wählen indizierte ("angeordnete") Basen $(a_i)_{i \in I}$ von V und $(b_j)_{j \in J}$ von W. Für T nehmen wir einen Vektorraum mit einer durch $I \times J = \{(i, j) : i \in I, j \in J\}$ indizierten Basis $(c_{i,j})_{(i,j) \in I \times J}$. Schließlich definieren wir die bilineare Abbildung $\otimes \colon V \times W \to T$ durch

$$a_i \otimes b_j = c_{i,j}$$

und setzen sie bilinear fort, das heißt für alle Paare aus endlichen Linearkombinationen gilt:

$$\Big(\sum_{i \in I_0} \alpha_i a_i\Big) \otimes \Big(\sum_{j \in J_0} \beta_i b_i\Big) = \sum_{(i,j) \in I_0 \times J_0} \alpha_i \beta_j (a_i \otimes b_j) = \sum_{(i,j) \in I_0 \times J_0} \alpha_i \beta_j \, c_{i,j}.$$

Sei nun eine konkurierende Abbildung $\otimes' : V \times W \to T'$ gegeben, die auch die Tensorprodukt-Eigenschaften hat. Wenn $t : T \to T'$ linear ist, ist $t \circ \otimes$ bilinear und stimmt mit $\otimes'$ genau dann überein, wenn für alle i und j $t(a_i \otimes b_j) = t(c_{i,j}) = a_i \otimes' b_j$. Man kann aber Werte von t auf der Basis $(c_{i,j})_{(i,j)\in I\times J}$ beliebig vorschreiben und t ist dadurch eindeutig bestimmt. t ist ein Isomorphismus von T auf T'. □

Wir geben noch einen alternativen Beweis mit einer Quotientenkonstruktion [1]

Definition 3.6. Es seien V und W k-Vektorräume.

(1) Wir definieren
$$Z = \bigoplus_{(v,w)\in V\times W} k$$
Wir setzen (v,w) Abkürzung für as Element von Z, das an der Stelle (v,w) den Eintrag 1 hat und Null an allen anderen Stellen.

(2) Im Vektorraum Z definieren wir den Unterraum
$$\begin{aligned} R = \operatorname{span}(&\{(\alpha_1 v_1 + \alpha_2 v_2, w) - \alpha_1(v_1, w) - \alpha_2(v_2, w) \,:\, \alpha_1, \alpha_2 \in k, v_1, v_2 \in V, w \in W\} \cup \\ &\{(v, \beta_1 w_1 + \beta_2 w_2) - \beta_1(v, w_1) - \beta_2(v, w_2) \,:\, \beta_1, \beta_2 \in k, w_1, w_2 \in W, v \in V\}) \end{aligned}$$

(3) Dann setzen wir
$$V \bigotimes W = \bigoplus_{(v,w)\in V\times W} k/R$$

(4) Wir definieren
$$\theta : V \times W \to V \otimes W$$
durch $(v,w) \mapsto (v,w) + R =: v \otimes w$.

Übung 3.7. Rechnen Sie nach, dass das eben definierte Tensorprodukt $\bigotimes$ die unverselle Eigenschaft von Lemma 3.3 und dass die Operation $\otimes$ die Eigenschaften aus der Deifnition 3.1 hat. Die Bezeichnungen in Satz 3.5 stimmen also mit denen aus der Quotientendefinition 3.6 überein.

Bemerkung 3.8. Die Elemente von T heißen auch Tensoren. Ein Tensor $t \in T$ heißt rein oder Elementartensor , wenn $t = v \otimes w = (v,w) + R$ für Vektoren $v \in V$, $w \in W$ also t im Bild von θ ist. Das Bild von θ is kein Unterraum von T. Es gibt keine Regel zum Vereinfachen von $v_1 \otimes w_1 + v_1 \otimes w_2$. Wir sehen, dass die reinen Tensoren den Raum aufspannen. Die Menge der reinen Tensoren ist linear abhängig, falls $\dim(V) \cdot \dim(W) \neq 0$. Sie bleibt linear abhängig, auch wenn man Tensoren der Form $O_V \otimes w$ weglässt.

[1] Diese wird meistens als sehr abstrakt empfunden. Wenn jemand sie als besonders schön empfindet, möge er es mich bitte wissen lassen. :-)

Korollar 3.9. *Es sei $(a_i)_i$ eine indizierte Basis von V, I Menge.*

(1) Wenn (b_j) eine Basis von W ist und J so ist $(a_i \otimes b_j)_{(i,j)\in I\times J}$ eine Basis von T. (Man hat $I \times J = \varnothing$, falls eine der beiden Mengen leer ist. Die Aussage stimmt also auch in diesen Extremfällen.)

(2) Jedes Element von $V \otimes W$ lässt sich schreiben als endliche Summe von

$$a_i \otimes w_i$$

für eindeutig bestimmte $w_i \in W$.

Beweis: (1) Es sei $\sum \alpha_{i,j} a_i \otimes b_j = 0_T$. Das heißt, jede lineare Abbildung in einen weiteren k-Vektorraum U, bildet $\sum \alpha_{i,j} a_i \otimes b_j$ auf 0_U ab. Nun nehmen wir eine bilineare Abbildung, die genau $s(a_{i_0}, b_{j_0}) = 1$ hat und alle anderen Basis-Paare auf 0 abbildet. Nach der universellen Eigenschaft gibt es ein lineares $f\colon T \to U$, so dass

$$f(a_i \otimes a_j) = s(a_i, b_j) \text{ für alle } i \in I_0, j \in J_0.$$

Dann ist

$$0 = f(0_T) = f\Big(\sum_{i\in I_0, j\in J_0} \alpha_{i,j} a_i \otimes b_j\Big) = \sum \alpha_{i,j} f(a_i \otimes b_j) = \sum \alpha_{i,j} s(a_j, b_j) = \alpha_{i_0,j_0}.$$

(2) Wir haben

$$\sum_{i\in I_0, j\in J_0} \gamma_{i,j}\, a_i \otimes b_j = \sum_{i\in I_0} a_i \otimes \Big(\sum_{j\in J_0} \gamma_{i,j} b_j\Big) = \sum_{i\in I_0} a_i \otimes w_i$$

mit

$$w_i = \sum_{j\in J_0} \gamma_{i,j} b_j.$$

□

Beispiel 3.10. Wir nehmen $V = k^2$ mit Basis (e_1, e_2) und $W = k^3$ mit Basis (e_1, e_2, e_3) und $t = e_1 \otimes e_1 +_T e_2 \otimes e_2$. Dann ist t kein Elementartensor. Der Ansatz $t = (\alpha_1 e_1 + \alpha_2 e_2) \otimes (\beta_1 e_1 + \beta_2 e_2) = \alpha_1\beta_1 e_1 \otimes e_1 + \alpha_2\beta_1 e_2 \otimes e_1 + \alpha_1\beta_2 e_1 \otimes e_2 + \alpha_2\beta_2 e_2 \otimes e_2$ und somit $\alpha_1\beta_1 = 1$, $\alpha_1\beta_2 = 0$, $\alpha_2\beta_2 = 1$ und $\alpha_2\beta_1 = 0$. Also $\alpha_1 = 0$ oder $\beta_2 = 0$. Beides führt zum Widerspruch. Es gibt also keine solchen $\alpha_1, \beta_1, \alpha_2, \beta_2$

Korollar 3.11. *Aus 1) folgt, dass für endlichdimensionale Vektorräume*

$$\dim(V \otimes W) = \dim(V) \cdot \dim(W).$$

Korollar 3.12. *Aus 2) folgt, dass jedes Element von T eine Summe von reinen Tensoren ist. Nicht jeder Tensor ist rein. Man kann sich nämlich leicht überlegen, daß $c = \sum_{i,j} \gamma_{i,j} a_i \otimes b_j$ genau dann ein reiner Tensor ist, wenn die Matrix $A = (\gamma_{i,j})$ höchstens den Rang 1 hat. Der Rang von A, der auch im Unendlichdimensionalen existiert, hängt nur von c ab und nicht von der Wahl der Basis. Er heißt der Rang von c.*

Lemma 3.13. *Die kanonische Abbildung $K \otimes V \to V$, definiert durch*

$$\alpha \otimes v = \alpha v$$

ist ein Isomorphismus.

Beweis: Die Umkehrabbildung ist $v \mapsto 1 \otimes v$, weil $\alpha \otimes v = \alpha(1 \otimes v) = 1 \otimes (\alpha v)$. $\square$

Definition 3.14. (und Existenzbehauptung) Es seien $f: V \to V'$ und $g: W \to W'$ lineare Abbildungen. Weil

$$s(v, w) \mapsto f(v) \otimes g(w)$$

bilinear ist, gibt es wegen der universellen Eigenschaft des Tensorprodukts genau eine lineare Abbildung

$$f \otimes g: V \bigotimes W \to V' \bigotimes W'$$

mit

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

Hierdurch ist $\otimes$ auf $\mathrm{Hom}(V, V') \times \mathrm{Hom}(W, W')$ definiert.

Satz 3.15. *Die so definierte Abbildung*

$$\otimes: \mathrm{Hom}(V, V') \times \mathrm{Hom}(W, W') \to \mathrm{Hom}(V \bigotimes W, V' \bigotimes W')$$
$$(f, g) \mapsto f \otimes g$$

im Diagramm[2]

$$\begin{tikzcd}
V \times W \arrow[r, "{(f,g)}"] \arrow[d, "\theta"'] & V' \times W' \arrow[d, "\theta'"'] \\
V \bigotimes W \arrow[r, "f \otimes g"] & V' \bigotimes W'
\end{tikzcd}$$

wird auf diese Weise zu einem exakten linearen Bifunktor, d.h., es gelten folgende Bifunktor-Gesetze

[2] Hierbei sind θ und θ' die Abbildungen aus der Quotientenkonstruktion in Definition 3.6.

(1) $\mathrm{id}_V \otimes \mathrm{id}_W = \mathrm{id}_{V\otimes W}$.

(2) $(f' \circ f) \otimes (g' \circ g) = (f' \otimes g') \circ (f \otimes g)$.

(3) $(f + f') \otimes g = (f \otimes g) + (f' \otimes g)$.

(4) $(\alpha f) \otimes g = \alpha(f \otimes g)$.

(5) 3) und 4) gelten auch für das zweite Argument.

(6) Wir können also den Urbildraum der Abbildung $\otimes$ *aud Satz 3.15 als* $\mathrm{Hom}(V, V') \bigotimes \mathrm{Hom}(W, W')$ *benennen.*

(7) Der Funktor ist exakt, erster Teil: $\mathrm{Im}(f \otimes \mathrm{id}_W) \cong \mathrm{Im}(f) \bigotimes W$.

(8) Der Funktor ist exakt, zweiter Teil: $\ker(f \otimes \mathrm{id}_W) \cong \ker(f) \bigotimes W$.

Beweis: Um zu zeigen, dass die linearen Abbildungen, die jeweils auf den beiden Seiten der Gleichungen stehen, gleich sind, genügt es zu zeigen, dass diese Abbildungen auf den Erzeugenden $v \otimes w$ übereinstimmen. Das sind leichte Rechnungen.

(2) Es seien $f \in \mathrm{Hom}(V, V')$, $f' \in \mathrm{Hom}(V', V'')$, $g \in \mathrm{Hom}(W, W')$, $g' \in \mathrm{Hom}(W', W'')$. Dann ist für $v \in V$, $w \in W$,

$$\begin{aligned}
(f' \circ f) \otimes (g' \circ g)(v \otimes w) &= ((f' \circ f)(v)) \otimes ((g' \circ g)(w)) \\
&= (f'(f(v))) \otimes g'(g(w)) \\
&= (f' \otimes g')(f(v) \otimes g(w)) \\
&= (f' \otimes g')((f \otimes g)(v \otimes w)) \\
&= ((f' \otimes g') \otimes (f \otimes g))(v \otimes w).
\end{aligned}$$

(6) Nach (3), (4) und (5) gelten für $(\mathrm{Hom}(V, V'), +)$, $\mathrm{Hom}(W, W'), +)$ und $f \otimes g$ gerade die Identifikations-Gesetze aus Definition 3.1.

(7) und (8) Es sei ker(f) durch $b_1, \ldots, b_k$ aufgespannt und es seien $b_{k+1}, \ldots, b_n$ eine Ergänzung zu einer Basis von V. Dann spannen nach dem Noether'schen Isomomrphiesatz $f(b_{k+1}), \ldots, f(b_n)$ das Bild $f[V] = \mathrm{Im}(f)$ auf.

Die Abbildung $(f \otimes \mathrm{id}_W)(v, w) \mapsto (f(v)) \otimes w$ ist wohldefiniert und auf den (v, w) mit $v \in \mathrm{span}(b_{k+1}, \ldots, b_n)$ bijektiv.

Genau für $v \in \ker(f)$ ist für jedes w das Tensorprodukt $f \otimes \mathrm{id}_W(v, w) = 0$. Falls $f(v) \neq 0$, so ist für jedes $w \in W$, $f(v) \otimes w = 0$. Falls für ein $w \in W$, $f(v) \otimes w = 0_{V' \otimes W}$, so ist schon $f(v) = 0$.

□

Beispiel 3.16. Es sei V ein $\mathbb{R}$-Vektorraum. Dann ist $\mathbb{C} \bigotimes V$ ein $\mathbb{C}$-Vektorraum mit der Skalarmultiplikation für $\alpha, \lambda \in \mathbb{C}$,

$$\lambda(\alpha \otimes v) = \lambda\alpha \otimes v \tag{3.1}$$

Man nennt $\mathbb{C} \bigotimes V$ die Komplexifizierung von V. Es sei μ_λ die Multiplikation mit $\lambda \in \mathbb{C}$. Aus $\lambda(\alpha \otimes v) = (\mu_\lambda \otimes \mathrm{id}_V)(\alpha \otimes v)$ folgt, dass Gleichung (3.1) eine $\mathbb{C}$-Vektorraumstruktur $\mathbb{C} \bigotimes V = V_{\mathbb{C}}$ über V definiert. Jede $\mathbb{R}$-lineare Abbildung $f\colon V \to W$ induziert eine $\mathbb{C}$-lineare Abbildung $\mathrm{id}_{\mathbb{C}} \otimes f = f_{\mathbb{C}}\colon V_{\mathbb{C}} \to W_{\mathbb{C}}$. Die darstellenden Matrizen bleiben gleich.

Dies haben wir in unseren Beweisen über $\mathrm{End}(V)$ schon mehrfach benutzt and den Stellen, als wir über einem algebraisch abgeschlosssenen Körper arbeiten wollten. Ein Beispiel ist der Satz von Cayley–Hamilton 1.40.

Definition 3.17 (und Existenzbeweis)**.** Das Tensorprodukt einer endlichen Familie $V_1, \dots V_n$ von Vektorräumen ist ein Vektorraum

$$V_1 \bigotimes V_2 \bigotimes \cdots \bigotimes V_n$$

zusammen mit einer multilinearen Abbildung

$$\otimes\colon V_1 \times \cdots \times V_n \to V_1 \bigotimes V_2 \bigotimes \cdots \bigotimes V_n,$$

so dass sich jede multilineare Abbildung

$$\mu\colon V_1 \times \cdots \times V_n \to W$$

eindeutig in der Form

$$\mu = f \circ \otimes$$

für ein lineares $f : V_1 \bigotimes \cdots \bigotimes V_n \to W$ schreiben lässt.

Wir schreiben immer $v_1 \otimes \cdots \otimes v_n$ statt $\otimes(v_1, \dots, v_n)$. Die lineare Abbildung f ist dann gegeben durch $f(v_1 \otimes \cdots \otimes v_n) = \mu(v_1, \dots, v_n)$. Man sagt, dass μ durch $\otimes$ faktorisiert. Existenz und Eindeutigkeit zeigt man wie in Satz 3.5. Wenn für $j = 1, \dots, n$, die Folge $(b_{j,i})_{i \in I_j}$ (mit einer Menge I_j) eine indizierte Basis von V_j ist, so ist

$$(b_{1,i_1} \otimes \cdots \otimes b_{n,i_n})_{(i_1,\dots,i_n) \in I_1 \times \cdots \times I_n}$$

eine indizierte Basis von $V_1 \bigotimes \cdots \bigotimes V_n$. (Ende der Definition)

Satz 3.18.

(1) Durch $(v_1 \otimes \cdots \otimes v_m) \otimes (w_1 \otimes \cdots \otimes w_n) \mapsto (v_1 \otimes \dots v_m \otimes w_1 \otimes \cdots \otimes w_n)$ wird ein Isomorphismus

$$(V_1 \bigotimes \cdots \bigotimes V_m) \bigotimes (W_1 \bigotimes \cdots \bigotimes W_n) \to V_1 \bigotimes \cdots \bigotimes V_m \bigotimes W_1 \bigotimes \cdots \bigotimes W_n$$

definiert.

(2) *Die Räume $V \otimes W$ und $W \otimes V$ sind kanonisch isomorph. Der Isomorphismus wird durch $v \otimes w \mapsto w \otimes v$ gegeben.*

Definition 3.19. Sei $\bigotimes^0 V = k$ und (für positive p) $\bigotimes^p V$ das Tensorprodukt von p Kopien von V. Dann ist nach dem vorhergehenden Satz für alle p und q eine bilineare Abbildung

$$\otimes\colon \bigotimes^p V \times \bigotimes^q V \to \bigotimes^{p+q} V$$

definiert. Wenn p oder q Null sind, soll diese Abbildung Multiplikation mit Elementen von k sein.

Satz 3.20. *Die Abbildungen*

$$\otimes\colon \bigotimes^p V \times \bigotimes^q V \to \bigotimes^{p+q} V$$

machen

$$T(V) = \bigoplus_{n\in\mathbb{N}} \bigotimes^n V$$

machen $(T(V), +, \otimes, \cdot_{\text{skalar}}, 1_k)$ zu einer assoziativen Algebra mit Einselement, der Tensoralgebra von V.

3.2 Tensorprodukt, Verjüngung (contraction) und Dualität

Definition 3.21.

(1) Die durch die bilineare Abbildung $(\lambda, x) \mapsto \lambda(x)$ gegebene lineare Abbildung

$$V^* \times V \to K$$

heißt Verjüngung (contraction) .

Auch ihre Faktorisierungs-Abbildung auf $V^* \otimes V$ heißt Verjüngung.

Auch in der umgekehrten Reihenfolge kann man verjüngen: $V \times V^* \to K$, $(x, \lambda) \mapsto \lambda(x)$.

(2) Allgemeiner definiert man die Verjüngung über V_i und V_j falls $V_i = V_j^*$ durch

$$\Gamma^i_j = v\colon \bigotimes_{k=1}^p V_k \to \bigotimes_{k\neq i,j} V_k$$

durch [3]

$$v(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_j \otimes \cdots \otimes v_p) = v_i(v_j) \cdot v_1 \otimes \cdots \otimes \hat{v}_i \otimes \cdots \otimes \hat{v}_j \otimes \cdots \otimes v_p.$$

[3] Die Schreibweise bedeutet, dass im rechten Produkt die Faktoren v_i und v_j weggelassen sind.

Die Elemente von

$$V_q^p = \bigotimes^p V^* \bigotimes \bigotimes^q V$$

nennt man p-fach kontravariante und q-fach kovariante Tensoren (oder auch Tensoren der Stufe (p,q)). Tensorieren liefert eine lineare Abbildung

$$V_q^p \times V_{q'}^{p'} \to V_{q+q'}^{p+p'}.$$

Verjüngungen über der i-ten Kopie von V^* und der j-ten Kopie von V liefert eine Abbildung

$$\Gamma_j^i\colon V_p^q \to V_{p-1}^{q-1}.$$

(Dies entspricht der Schreibweise $v_i(v_{j+p})$ von oben, die Nummerierung startet in der Γ-Schreibweise bei den hinteren Faktoren neu mit 1.) Man schreibt mit $e^i = e_i^*$ zum Beispiel

$$\sum_{i_1,\dots,i_p,j_1,\dots,j_q} \alpha_{i_1,\dots,i_p}^{j_1,\dots j_q} e^{i_1} \otimes \dots \otimes e^{i_p} \otimes e_{j_1} \otimes \dots \otimes e_{j_q}.$$

(3) Die Verjüngung in der zweiten und in der dritten Koordinate

$$j\colon U \bigotimes V^* \bigotimes V \to U$$

liefert für jedes $a = a_1 \otimes a_2 \in U \otimes V^*$ eine lineare Abbildung

$$\begin{aligned} f_a\colon V &\to U \\ v &\mapsto j(a \otimes v) = a_2(v) \cdot a_1 \end{aligned}$$

Die Abbildung $a \mapsto f_a$ ist offensichtlich linear und es gilt $f_{a\otimes\lambda}(v) = \lambda f_a(v)$.

Satz 3.22. *Die Abbildung $a \mapsto f_a$ ist ein Isomorphismus zwischen $(U \otimes V^*)$ und dem Raum $L_e(V,U)$ der linearen Abbildungen in* $\mathrm{Hom}(V,U)$*, die endlichen Rang haben.*

Beweis: Die Abbildung bildet nach $L_e(V,U)$ ab: Sei $a = u_1 \otimes \lambda_1 + \dots + u_n \otimes \lambda_n \in U \otimes V^*$. Da die Abbildung linear ist, hat f_a endlichen Rang, denn für beliebiges $v \in V$ gilt

$$f_a(v) = \sum_{i=1}^n f_{u_i\otimes\lambda_i}(v) = \sum_{i=1}^n \lambda_i(v) u_i \in \mathrm{span}(u_1,\dots,u_n).$$

Die Abbildung ist injektiv: Es sei $a \in U \otimes V^*$, sodass $f_a(v) = 0$ ist für alle $v \in V$. Sei genauer $a = u_1 \otimes \lambda_1 + \dots + u_n \otimes \lambda_n$ für linear unabhängige $u_1,\dots,u_n$. Dann ist

$$f_a(v) = \sum_{i=1}^n \lambda_i(v) u_i = 0,$$

wegen der linearen Unabängigkeit der u_i ist also $\lambda_i(v) = 0$ für alle v und $i \leq n$, also ist $a = 0$.
Die Abbildung ist surjektiv: Sei $g \in L_e(V,U)$. Es sei $u_1, \ldots, u_n$ eine Basis von Bild(g) und $v_1, \ldots, v_n$ eine Basis von einem Komplementärraum von $\ker(g)$, so dass $g(v_i) = u_i$. Außerdem seien v_i^* die dualen Basisvektoren von $(\ker(g))^*$. Dann setzen wir

$$a = \sum_{i=1}^{n} u_i \otimes v_i^*$$

und erhalten für $v \in V$, $f_a(v) = g(v)$. Dies rechnet man für $v_1, \ldots, v_n$ durch $v_i^*(v_j) = \delta_{i,j}$ nach und für $v \in \ker(g)$ ist $f_a(v) = 0$. □

Definition 3.23 (Definition 4.28 aus LA I)**.** Das Paar V, W heißt duales Paar, wenn es eine Bilinearform $B{:}\, V \times W \to K$ gibt, so dass $\dim(V)$ endlich ist und $\dim(V) = \dim(W) = \text{Rang}(B)$.

Satz 3.24. *Wenn U und V endlichdimensional sind, bilden $U \otimes V$ und $U^* \otimes V^*$ ein duales Paar.*

Beweis: Eine bezeugende Bilinearform vom vollen Rang ist $((\lambda \otimes \mu), (u \otimes v)) = \lambda(u) \cdot \mu(v)$. □

Bemerkung 3.25. $(U \otimes V)^*$ kann also mit $U^* \otimes V^*$ identifiziert werden.

3.3 Multilineare Abbildungen

Definition 3.26. Es seien V, W K-Vektorräume.

(1) Eine Abbildung $f{:}\, V^n \to W$ heißt multilinear, wenn sie in jedem Argument K-linear ist.

(2) Eine multilineare Abbildung $f{:}\, V^n \to W$ heißt symmetrisch, wenn für jede Permutation $\sigma \in S_n$ gilt
$$f(v_1, \ldots, v_n) = f(v_{\sigma(1)}, \ldots v_{\sigma(n)}).$$

(3) Eine multilineare Abbildung $f{:}\, V^n \to W$ heißt alternierend, wenn für jedes Tupel $(v_1, \ldots, v_n)$ mit einer Wiederholung also $v_i = v_j$ für ein $i \neq j$ gilt, $f(v_1, \ldots, v_n) = 0$.

Beispiel 3.27. det : $(K^n)^n \to K$ ist eine alternierende Multilinearform. (Form = Lineare Abbildung mit Bildraum K)

Lemma 3.28. *Es genügt, die Symmetriebedingungen nur für Transpositionen zu verlangen. Für alternierende f gilt*

$$f(v_{\sigma(1)}, \dots v_{\sigma(n)}) = \operatorname{sign}(\sigma) f(v_1, \dots, v_n).$$

Beweis: Die erste Aussage folgt aus LAI 3.13. Die zweite Aussage ist gerade LAI 3.32. □

Definition 3.29. Es sei V ein K-Vektorraum, $n \geq 1$.

(1) Es sei $I_n \subseteq \bigotimes^n V$ der Unterraum von $\bigotimes^n V$, der von

$$\{v_1 \otimes \cdots \otimes v_n - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)} \,:\, v_1, \dots, v_n \in V, \sigma \in S_n\}$$

erzeugt wird. Die n-te symmetrische Potenz von V ist der Quotienten-Vektorraum

$$S^n(V) := \bigotimes^n V / I_n.$$

Die zugehörige Projektion

$$\begin{aligned} \pi_{S^n} \colon V^n &\to S^n(V) \\ (v_1, \dots, v_n) &\mapsto v_1 \otimes \cdots \otimes v_n + I_n =: v_1 \bullet \cdots \bullet v_n \end{aligned}$$

heißt das symmetrische Produkt.

(2) Es sei $J_n \subseteq \bigotimes^n V$ der Unterraum von $\bigotimes^n V$, der von

$$\{v_1 \otimes \cdots \otimes v_n \,:\, \exists i \neq j, v_i = v_j\}$$

erzeugt wird. Die n-te äußere Potenz (exterior product, wedge product) von V ist der Quotienten-Vektorraum

$$\bigwedge^n(V) := \bigotimes^n V / J_n.$$

Die zugehörige Projektion

$$\begin{aligned} \pi_{\bigwedge^n} \colon V^n &\to \bigwedge^n V \\ (v_1, \dots, v_n) &\mapsto v_1 \otimes \cdots \otimes v_n + J_n =: v_1 \wedge \cdots \wedge v_n \end{aligned}$$

heißt das Dachprodukt, äußeres Produkt, wedge product.

Übung 3.30. Das symmetrische Produkt ist symmetrisch. Das äußere Produkt ist alternierend.

Das Dachprodukt ist also antisymmetrisch:

$$e_1 \wedge e_2 + e_2 \wedge e_1 = e_1 \wedge e_2 + e_2 \wedge e_2 + e_1 \wedge e_1 + e_2 \wedge e_1 = (e_1 + e_2) \wedge e_2 + (e_1 + e_2) \wedge e_1 = (e_1 + e_2) \wedge (e_1 + e_2) = 0.$$

Die erste Gleichung folgt, da man modulo J_n addiert.

Satz 3.31 (Universelle Eigenschaften)**.**

(1) Es sei $f\colon V^n \to U$ eine symmetrische multilineare Abbildung. Dann gibt es eine eindeutig bestimmte Abbildung $\bar{f}$, so dass das folgende Diagramm kommutiert.

$$\begin{tikzcd} V^n \arrow[r, "\pi_{S^n}"] \arrow[dr, "f"'] & S^n V \arrow[d, dashed, "\bar{f}"] \\ & U \end{tikzcd}$$

(2) Es sei $f\colon V^n \to U$ eine alternierende multilineare Abbildung. Dann gibt es eine eindeutig bestimmte Abbildung $\bar{f}$, so dass das folgende Diagramm kommutiert.

$$\begin{tikzcd} V^n \arrow[r, "\pi_{\wedge^n}"] \arrow[dr, "f"'] & \bigwedge^n V \arrow[d, dashed, "\bar{f}"] \\ & U \end{tikzcd}$$

Beweis: (1) Da f multilinear ist existiert aufgrund der universellen Eigenschaft des Tensorprodukts genau eine lineare Abbildung $\tilde{f} : \bigotimes^n V \to U$, mit $f(v_1, \ldots, v_n) = \tilde{f}(v_1 \otimes \cdots \otimes v_n)$ für alle $v_1, \ldots, v_n \in V$. Wir definieren nun $\bar{f}$ durch

$$\bar{f}(v_1 \bullet \cdots \bullet v_n) = \tilde{f}(v_1 \otimes \cdots \otimes v_n).$$

Da f symmetrisch ist, ist $\bar{f}$ wohldefiniert: Seien $v_1 \bullet \cdots \bullet v_n = v'_1 \bullet \cdots \bullet v'_n$, d.h.

$$v_1 \otimes \cdots \otimes v_n - v'_1 \otimes \cdots \otimes v'_n \in I_n.$$

In anderen Worten existiert eine Permutation $\sigma \in S_n$ mit $v_i = v'_{\sigma(i)}$ für alle $i \leq n$. Also ist

$$\tilde{f}(v_1 \otimes \cdots \otimes v_n) = f(v_1, \ldots, v_n) = f(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \tilde{f}(v'_1 \otimes \cdots \otimes v'_n).$$

Schließlich haben wir

$$(\bar{f} \circ \pi_{S^n})(v_1, \ldots, v_n) = \bar{f}(v_1 \bullet \cdots \bullet v_n) = \tilde{f}(v_1 \otimes \cdots \otimes v_n) = f(v_1, \ldots, v_n).$$

(2) Genauso wie in (1). In diesem Fall ist die Vorschrift

$$\bar{f}(v_1 \wedge \cdots \wedge v_n) = \tilde{f}(v_1 \otimes \cdots \otimes v_n)$$

wohldefiniert weil f alternierend ist.

□

Die Konstruktion von symmetrischen und äußeren Potenzen ist verträglich mit linearen Abbildungen.

Lemma 3.32. *Es sei $f: V \to V'$ linear und es sei $n \geq 1$. Dann gibt es eindeutige lineare Abbildungen*

$$\begin{aligned} S^n(f): S^n(V) &\to S^n(V') \\ v_1 \bullet \cdots \bullet v_n &\mapsto f(v_0) \bullet \cdots \bullet f(v_n) \end{aligned}$$

$$\begin{aligned} \bigwedge^n(f): \bigwedge^n(V) &\to \bigwedge^n(V') \\ v_1 \wedge \cdots \wedge v_n &\mapsto f(v_0) \wedge \cdots \wedge f(v_n) \end{aligned}$$

Beweis: Die Abbildung $\pi_{S^n} \circ f^n$ ist multilinear und symmetrisch, die Abbildung $\pi_{\bigwedge^n} \circ f^n$ ist multilinear und alternierend. Die jeweiligen universellen Eigenschaften der symmetrischen bzw. äußeren Potenz ergeben dann jeweils die eindeutigen Abbildungen $S^n(f)$ und $\bigwedge^n(f)$.

□

Es sei nun V endlich-dimensional und $\dim(V) = n$. Sei außerdem $\{e_1, \ldots, e_n\}$ eine Basis von V. Für eine beliebige p-elementige Teilmenge $I = \{i_1, \ldots, i_p\}$ von $\{1, \ldots, n\}$ mit $i_1 < \cdots < i_p$ betrachten wir

$$e_I = e_{i_1} \wedge \cdots \wedge e_{i_p} \in \bigwedge^p V.$$

Satz 3.33. *Die Menge der e_I, I p-elementige Teilmenge von $\{1, \ldots, n\}$, ist eine Basis von $\bigwedge^p V$. Insbesondere hat $\bigwedge^p V$ die Dimension $\binom{n}{p}$.*

Beweis: Die e_I bilden ein Erzeugendensystem: Sei $v_1 \wedge \cdots \wedge v_p \in \bigwedge^p V$. Es reicht zu zeigen, dass sich Elemente der äußeren Potenz in dieser Form als Linearkombinaton der e_I schreiben lassen, denn jede endliche Summe von Linearkombinationen ist ebenfalls eine Linearkombination. Für $i \leq p$ schreiben wir den Vektor v_i als Linearkombination unserer Basis von V,

$$v_i = \sum_{j=1}^{n} \alpha_{j,i} e_j.$$

Da das Dachprodukt multilinear und alterneriend ist bekommen wir dann

$$\begin{aligned}
v_1 \wedge \cdots \wedge v_n &= \left(\sum_{j=1}^{n} \alpha_{j,1} e_j\right) \wedge \cdots \wedge \left(\sum_{j=1}^{n} \alpha_{j,p} e_j\right) \\
&= \sum_{j_1=1}^{n} \alpha_{j_1,1} \cdot \left(e_{j_1} \wedge \left(\sum_{j_2=1}^{n} \alpha_{j_2,2} e_{j_2}\right) \wedge \cdots \wedge \left(\sum_{j_p=1}^{n} \alpha_{j_p,p} e_{j_p}\right)\right) \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_p=1}^{n} \left(\alpha_{j_1,1} \cdot \cdots \cdot \alpha_{j_p,p} \cdot (e_{j_1} \wedge \cdots \wedge e_{j_p})\right) \\
&= \sum_{I} \alpha_I \cdot e_I.
\end{aligned}$$

Dabei läuft die letzte Summe über alle p-elementigen Teilmengen von $\{1, dots, n\}$ und es ist

$$\alpha_I = \sum_{\sigma \in S_p} \operatorname{sign}(\sigma) \prod_{i=1}^{p} \alpha_{j_{\sigma(i)},i}.$$

Die e_I sind linear unabhängig: Für festes $I = \{i_1, \dots, i_p\} \subseteq \{1, \dots, n\}$ mit $i_1 < \cdots < i_p$ gibt es nach Korollar 3.33 LA I genau eine p-Form $\mu_I : V^p \to K$ mit

$$\mu_I(e_{j_1}, \dots, e_{j_p}) = \prod_{k=1}^{p} e_{i_k}^*(e_{j_k}) = \begin{cases} 1 & I = J, \\ 0 & I \neq J. \end{cases}$$

Dabei ist $J = \{j_1, \dots, j_p\} \subseteq \{1, \dots, n\}$ mit $j_1 < \cdots < j_p$ beliebig. Aufgrund der universellen Eigenschaft der äußeren Potenz gibt es also eine eindeutige lineare Abbildung $e_I^* \in (\bigwedge^p V)^*$ mit

$$e_I^*(e_J) = \begin{cases} 1 & I = J, \\ 0 & I \neq J. \end{cases}$$

Die Situation wird durch das folgende Diagramm verdeutlicht:

$$\begin{tikzcd}
V^p \arrow[r, "\pi_{\wedge^p}"] \arrow[dr, "\mu_I"'] & \bigwedge^p V \arrow[d, dashed, "e_I^*"] \\
& K
\end{tikzcd}$$

Seien nun $I_1, \dots, I_k$ paarweise verschiedene p-elementige Teilmengen von $\{1, \dots, n\}$ und

$$\sum_{j=1}^{k} \alpha_j e_{I_j} = 0.$$

Wir müssen nun $\alpha_j = 0$ für alle $j \leq k$ zeigen um die lineare Unabhängigkeit der e_I zu beweisen. Für festes $j \leq k$ gilt aber tatsächlich schon

$$0 = e^*_{I_j}(0) = e^*_{I_j}\left(\sum_{l=1}^{k} \alpha_l \cdot e_{I_l}\right) = \sum_{l=1}^{k} \alpha_l \cdot e^*_{I_j}(e_{I_l}) = \alpha_l.$$

□

Korollar 3.34. *Im Spezialfall* $n = p$ *erhalten wir* $\dim(\bigwedge^n V) = 1$ *und für jedes* $f \in \mathrm{End}(V)$, $(\bigwedge^n f)(e_1 \wedge \dots \wedge e_n) = \det(f) \cdot e_1 \wedge \dots \wedge e_n$.

Beweis: Sei $A = (\alpha_{i,j})_{i,j \leq n}$ die darstellende Matrix von f zur Basis $\{e_1, \dots, e_n\}$, d.h. für $j \leq n$ ist

$$f(e_j) = \sum_{i=1}^{n} \alpha_{i,j} e_i.$$

Durch anwenden der Definition von $\bigwedge^n(f)$ und einer Rechnung wie im vorherigen Beweis erhalten wir

$$\left(\bigwedge^n(f)\right)(e_1 \wedge \dots \wedge e_n) = f(e_1) \wedge \dots \wedge f(e_n) = \alpha \cdot (e_1 \wedge \dots \wedge e_n),$$

wobei wir für α jetzt die Leibniz-Formel benutzen können:

$$\alpha = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} \alpha_{\sigma(i),i} = \det(A) = \det(f).$$

□

Kapitel 4

Affine Räume

Nun definieren wir affine Räume. Diese haben drei Komponenten: eine Trägermenge A, einen Vektorraum V und eine Abbildung $\vec{\ }: A \times A \to V$.

Definition 4.1. Gegeben seien eine Menge $A \neq \varnothing$, deren Elemente geometrisch als Punkte aufgefasst werden, ein Vektorraum V über einem Körper K und eine Abbildung von $A \times A$ nach V, die zwei Punkten P, $Q \in A$ einen Verbindungsvektor $\overrightarrow{PQ}$ in V zuordnet, so dass die folgenden beiden Regeln gelten:

(1) Dreiecksregel, Regel von Chasles. Für je drei Punkte $P, Q, R \in A$ gilt:

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR},$$

(2) für jeden Punkt $P \in A$ und jeden Vektor $v \in V$ gibt es einen eindeutig bestimmten Punkt $Q \in A$, so dass $v = \overrightarrow{PQ}$ (Abtragbarkeitsregel).

Das Tripel $(A, V, \to)$ heißt affiner Raum. Wenn klar ist, welcher Vektorraum V und welche Pfeilabbildung $\to$ zugrunde liegt, spricht man auch allein vom affinen Raum A. Bei dem Körper K handelt es sich oft um den Körper $\mathbb{R}$ der reellen Zahlen.

Aus (1) folgt $\overrightarrow{PP} = 0_V$ und aus der Eindeutigkeitsaussage in (2) folgt, dass $\overrightarrow{PQ} = 0_V$ schon $P = Q$ impliziert.

Beispiel 4.2.

(1) Affine Räume der Dimension 0, 1, 2 heißen Punkte, Geraden, Ebenen.

(2) Jeder Vektorraum kann als affiner Raum aufgefasst werden, man nimmt $\overrightarrow{PQ} = Q - P$. Falls $V = K^n$, so nennen wir V als affinen Raum aufgefasst $\mathbb{A}^n$.

(3) Es sei $M \in M_{m,n}(K)$, $b \in K^n$. Die Menge

$$A = \{x \in V \;:\; Ax = b\}$$

ist leer oder ein affiner Raum zum Vektorraum $\{x \;:\; Ax = 0\}$.

Jedem affinen Raum, also jeder Struktur $(A, V, \vec{\ })$, kann man ein "gemischtes" Plus zuordnen, das bei fester zweiter Komponente als Verschiebung fungiert.

Definition 4.3 (Translationen)**.** Im affinen Raum ist eine Abbildung + von $A \times V \to A$,

$$(P, v) \mapsto P + v,$$

dadurch definiert, dass $P + v = Q$ gerade der durch $\overrightarrow{PQ} = v$ eindeutig bestimmte Punkt Q ist. Für festgelegtes $v \in V$ heißt die zugehörige Abbildung $T_v \colon A \to A$, $P \mapsto P + v$ die Translation (Verschiebung) um den Vektor v.

Proposition 4.4. *Die Translationen sind stets Bijektionen. Sie bilden zusammen mit der Hintereinanderschaltung als Gruppenverknüpfung eine Untergruppe der Automorphismengruppe* $\mathrm{Aut}(A)$*. Es ist* $T_{0_V} = \mathrm{id}_A$ *und* $T_{-v} = (T_v)^{-1}$ *und* $T_v \circ T_w = T_{v+w}$*.*

Definition 4.5. Wenn P ein festgelegter Punkt aus A ist und U ein Untervektorraum von V, dann ist $B = P + U = \{P + u \;:\; u \in U\}$ ein affiner Unterraum von A, auch affiner Teilraum genannt.

Der zu einem affinen Teilraum gehörige Untervektorraum U ist eindeutig bestimmt.

Definition 4.6. Die Dimension eines affinen Raums A zu einem Vektorraum V über einem Körper K ist definiert als die Dimension des Vektorraums V. Oft ist es bequem, auch die leere Menge als affinen (Teil-)Raum anzusehen. Diesem leeren Teilraum wird dann die Dimension -1 zugeordnet.

Der affine Punktraum und der ihm zugeordnete Vektorraum: Aus A wird V eindeutig bestimmt, aus V und P wird A bestimmt. Wegen dieser engen Verwandtschaft wird manchmal auf eine rigide Unterscheidung zwischen dem affinen Punktraum einerseits und dem Vektorraum der Verschiebungsvektoren andererseits verzichtet.

Beispiel 4.7. Jeder Vektorraum kann als affiner Raum aufgefasst werden. Dadurch ist auch jeder Unterraum eines Vektorraums ein affiner Raum.

Die Lösungen eines inhomogenen linearen Gleichungssystems bilden einen affinen Raum über dem Vektorraum der Lösungen des zugehörigen homogenen Systems. Das gilt analog auch für Systeme linearer Differentialgleichungen.

4.1 Etwas affine Geometrie der Ebene

Es sei k ein Körper. Wir nennen $\mathbb{A}^2 = k^2$ die affine Ebene und die Elemente $P \in \mathbb{A}^2$ Punkte. Wir möchten zwischen den Punkten und den Vektoren unterscheiden. Für P, $Q \in \mathbb{A}^2$ schreiben wir $\overrightarrow{PQ} = Q - P \in k^2$ für den Verbindungsvektor.

Definition 4.8. Eine affine Gerade in der Ebene ist eine Teilmenge von $\mathbb{A}^2$ von der Form

$$L = P + V$$

mit einem Punkt $P \in \mathbb{A}^2$ und einem Unterraum $V \subseteq k^2$ der Dimension 1.

Lemma 4.9. *Es sei L eine affine Gerade, $L = P + V$. Dann gilt für jedes $Q \in L$, dass $L = Q + V$. Außerdem ist V für je zwei verschiedene P, $Q \in L$ durch $\overrightarrow{PQ}$ aufgespannt.*

Lemma 4.10. *Seien $P \neq Q \in \mathbb{A}^2$. Dann gibt es genau eine affine Gerade L, die P und Q enthält.*

Definition 4.11. Zwei affine Geraden heißen parallel, falls ihre Richtungsvektoren übereinstimmen.

Satz 4.12 (Parallelaxiom)**.** *Es sei L eine affine Gerade und $P \in \mathbb{A}^2$ ein Punkt. Dann gibt es genau eine zu L parallele affine Gerade durch P.*

Warum kann man hier ein Axiom beweisen? Dies liegt an den starken Voraussetzungen an den zugrundeliegenden Vektorraum. Es gibt auch geometrische Räume mit mehreren "Parallelen" und keine "Parallelen" in dem Sinn, dass eine von L unterschiedliche Parallele zu L die Gerade L nicht schneiden soll.

Beweis: Es sei $L_0 = P_0 + V$, $V \subseteq k^2$ ein Unterraum der Dimension 1. Wir setzen $L = P + V$. □

Lemma 4.13. *Zwei nicht parallele Geraden in der Ebene schneiden sich in genau einem Punkt.*

Beweis: Es sei für $i = 0, 1$ die gerade L_i durch $L_i = P_i + V_i$ gegeben. Da L_0 und L_1 nicht parallel sind, ist $V_0 \neq V_1$. Da V zweidimensional ist, ist daher V_1 ein Komplementärraum zu V_0. Es sei $V_i = \operatorname{span}(\{v_i\})$. Die Gleichung $\alpha v_0 + \beta v_1 = \overrightarrow{(P_0, P_1)}$ hat daher genau eine Lösung (α, β). Dann ist $P_0 + \overrightarrow{(P_0, P_1)} = P_1$, also $P_0 + \alpha v_0 + \beta v_1 = P_1$ und somit $P_0 + \alpha v_0 = P_1 - \beta v_1$. Dies ist der gesuchte Schnittpunkt. Er ist eindeutig, da $V_0 \cap V_1$ der Nullraum ist. □

Satz 4.14. *Es sei k ein Körper der Charakteristik ungleich 2,3. Dann gilt: Die Seitenhalbierenden eines echten Dreiecks (d.h., es gibt zwei linear unabhängige Seiten) schneiden sich in gerade einem Punkt.*

Beweis: Es seien A, B und C die Eckpunkte des Dreiecks im affinen Raum k^2. Es sei $a = \overrightarrow{(B,C)}$, $b = \overrightarrow{(C,A)}$, $c = \overrightarrow{(A,B)} = -a - b$.

Die Seitenmittelpunkte sind

$$P_a = B + \frac{1}{2}a, \quad P_b = C + \frac{1}{2}b, \quad P_c = B + \frac{1}{2}(a+b).$$

Die Gerade durch A und P_a, bzw. durch B und P_b, C und P_c ist

$$\begin{aligned}
L_a &= A + \operatorname{span}(\overrightarrow{(A,P_a)}) = A - \operatorname{span}(\frac{1}{2}a - b);\\
L_b &= B + \operatorname{span}(\overrightarrow{(B,P_b)}) = B + \operatorname{span}(a + \frac{1}{2}b);\\
L_c &= C + \operatorname{span}(\overrightarrow{(C,P_c)}) = C + \operatorname{span}(\frac{1}{2}(b-a)).
\end{aligned}$$

Nun sei $M = A - \frac{1}{3}a - \frac{2}{3}b$. Dann ist

$$\begin{aligned}
M = A - \frac{2}{3}\left(\frac{1}{2}a + b\right) =&\\
= B + a + b - \frac{1}{3}a - \frac{2}{3}b = B + \frac{2}{3}a + \frac{1}{3}b = B + \frac{2}{3}\left(a + \frac{1}{2}b\right) =&\\
= C + b - \frac{1}{3}a - \frac{2}{3}b = C = \frac{1}{3}a + \frac{1}{3}b = C + \frac{1}{3}\left(b - a\right).&
\end{aligned}$$

Da a aund b linear unabhängig sind, ist die Lösung M eindeutig. □

Für den Satz vom Schnittpunkt der Winkelhalbierenden würde man nun Winkel einführen und dadurch Gleichungen für die Winkelhalbierenden erhalten. Danach kann man in ähnlichem Stil weiterrechnen. Für den Schnittpunkt der Höhen und den Schnittpunkt der Mittelsenkrechten arbeitet man mit einer geeigneten Beschreibung des Senkrechtstehens.

4.2 Affine Teilräume

Definition 4.15. Es sei $(A, V, \rightarrow)$ ein affiner Raum. Ein affiner Teilraum ist ein affiner Raum der Form $(A', V', \rightarrow \restriction A' \times A')$ mit einer Teilmenge $A' \subseteq A$ und einem Untervektorraum V' von V.

Lemma 4.16. *(A', V') ist ein affiner Teilraum von (A, V), wenn V' ein Unterraum ist und für einen Punkt $P \in A \cap A'$ gilt, $A' = \{P + v : v \in V'\}$.*

Bespiele: In der affinen Ebene des k^2 ist jede affine Gerade ein affiner Unterraum. Gegenbeispiele: Im k^3 gibt es affine Ebenen und affine Geraden, die die jeweilige Ebene womöglich nur in einem Punkt durchstoßen.

Lemma 4.17 (Parameterdarstellung eines affinen Teilraums des $\mathbb{A}^n$)**.** *Es sei $(A, V) \subseteq \mathbb{A}^n$ (Erinnerung $\mathbb{A}^n = k^n$) ein affiner Teilraum der Dimension d und es sei $P \in A$. Dann ist $A = \{P + v : v \in V\}$ und es gibt Vektoren $v_1, \ldots, v_d \in k^n$, die eine Basis von V bilden. Ausführlich also $A = \{P + \sum_{i=1}^d \lambda_i v_i : \lambda_i \in k^n\}$.*

Lemma 4.18 (Implizite Gleichung für einen affinen Teilraum des $\mathbb{A}^n$)**.** *Es sei $(A, V) \subseteq \mathbb{A}^n$ (Erinnerung $\mathbb{A}^n = k^n$) ein affiner Teilraum der Dimension d. Dann ist gibt es eine Matrix $B \in M_{(n-d)\times n}(k)$ vom Rang $n - d$ und ein $b \in k^{n-d}$, so dass*

$$A = \{x \in k^n : Bx = b'\}.$$

Beweis: Es sei $P \in A$. Falls $A = V$, nehmen wir $b = 0$, andernfalls $b = P$. Es sei $v_1, \ldots, v_d$ eine angeordnete Basis von V. Wir setzen diese durch $v_{d+1}, \ldots, v_n$ zu einer Basis von k^n fort. Weiter sei $w_1, \ldots w_{n-d}$ eine Basis von k^{n-d}. Nun definieren wir $\beta\colon k^n \to k^{n-d}$ durch

$$\beta(v_i) = \begin{cases} 0 & \text{für } 1 \le i \le d, \\ w_{i-d} & \text{für } d < i \le n. \end{cases}$$

Sei nun B eine Matrix, die β in den Standardbasen darstellt. Wir setzen $b' = \beta(P)$. Dann ist

$$\ker(\beta) = V = \ker(f_B)$$

und

$$A = \{x \in k^n : Bx = b'\}.$$

□

Beispiel 4.19. Eine affine Hyperebene $A \subseteq \mathbb{A}^n$, $\dim(A) = n - 1$, ist gegeben durch eine Gleichung

$$A = \{(x_1, \ldots, x_n) \in k^n : \sum_{i=1}^n a_i x_i = b\}$$

mit einem $(a_1, \ldots, a_n) \neq 0_{k^n}$. Für $n = 2$ erhält man eine Gerade in der Ebene als

$$L = \{(x, y) \in k^2 : ax + by = c\},$$

mit $(s, b) \neq (0, 0)$. Eine Parameterdarstellung ist $L = \{(x, \frac{ax-c}{b}) : x \in k\}$.

Korollar 4.20. *Jeder affine Teilraum des $\mathbb{A}^n$ der Dimension d ist der Schnitt von $n - d$ Hyperebenenen.*

Beweis: In der impliziten Darstellung von Lemma 4.18 besteht die Matrix gerade aus $n-d$ linear unabhängigen Zeilen. Jede dieser Zeilen beschreibt eine Hyperebene. □

Definition 4.21. Es sei (A, V) ein affiner Raum. Zwei Teilräume (A_i, V_i), $i = 1, 2$, heißen parallel, falls $V_1 \subseteq V_2$ oder $V_2 \subseteq V_1$.

Lemma 4.22. *Es seien A_1, A_2 parallele Teilräume. Dann sind A_1, A_2 disjunkt oder es gilt $A_1 \subseteq A_2$ oder $A_2 \subseteq A_1$.*

Beweis: Falls $P \in A_1 \cap A_2$ und $V_1 \subseteq V_2$, so ist $A_1 \subseteq A_2$. □

Definition 4.23. Es seien (A_1, V_1) und (A_2, V_2) zwei affine Teilräume eines affinen Raums (A, v). Es sei $A_1 + A_2$ der von $A_1 \cup A_2$ aufgespannte affine Teilraum , also der kleinste affine Teilraum von A, der A_1 und A_2 als Teilmengen enthält.

Satz 4.24 (Dimensionsformel für affine Teilräume)**.** *Es seien (A_1, V_1) und (A_2, V_2) zwei nicht leere affine Teilräume eines affinen Raums (A, V). Dann gilt*

$$\dim(A_1 + A_2) = \dim(A_1) + \dim(A_2) + \begin{cases} -\dim(A_1 \cap A_2), & \textit{falls } A_1 \cap A_2 \neq \emptyset \\ -\dim(V_1 \cap V_2) + 1, & \textit{falls } A_1 \cap A_2 = \emptyset. \end{cases}$$

Beweis: Falls $P \in A_1 \cap A_2 \neq 0$, so ist $A_i = P + V_i$ und $A_1 \cap A_2 = P + V_1 \cap V_2$. Die Dimensionsformel für Unterräume lautet $\dim(V_0 + V_1) = \dim(V_0) + \dim(V_1) - \dim(V_0 \cap V_1)$. Nach Definition ist $\dim(A_i) = \dim(V_i)$.

Falls $A_1 \cap A_2 = \emptyset$, so sei $P_i \in A_i$ und $v = \overrightarrow{(P_0, P_1)}$. Der kleinste Vektorraum, der zu $A_1 + A_2$ gehört, wird durch $V_1 + V_2 + \operatorname{span}(\{v\})$ aufgespannt. Da $A_1 \cap A_2 = \emptyset$, ist $v \notin V_1 + V_2$. Also ist $\dim(A_1 + A_2) = \dim(V_1 + V_2) + 1$. □

Beispiel 4.25. Zwei Geraden L_1, L_2 im $\mathbb{A}^3$. Es gibt vier Fälle:

1. L_1 und L_2 sind windschief. Es sei also $L_1 \cap L_2 = \emptyset$. Dann ist $\dim(L_1 + L_2) = 3$.

2. $L_1 \neq L_2$, aber sie liegen in einer Ebene. Hierzu gehören zwei Konstellationen:

2a: L_1 und L_2 sind nicht parallel. Es gibt genau einen Punkt P in $L_1 \cap L_2$. Dann spannen $L_1 + L_2$ diese Ebene $P + V_1 + V_2$ auf.

2b: $L_1 \neq L_2$ und beide sind parallel mit $V_1 = V_2 = V$. Dann spannt $P_1 + \operatorname{span}(\overrightarrow{(P_1, P_2)}) + V$ einen zweidimensionalen affinen Raum auf, die Ebene, in der sie liegen.

3. $L_1 = L_2$. Dann spannt $L_1 + L_2 = L_1$ einen eindimensionalen affinen Raum auf.

4.3 Der projektive Raum

Definition 4.26. Es sei k ein Körper und V ein k-Vektorraum der Dimension $n + 1$. Der n-dimensionale projektive Raum $\mathbb{P}(V)$ über V, k ist die Menge der Geraden in V, die durch den Nullpunkt gehen. Speziell für $V = k^{n+1}$ schreiben wir

$$\mathbb{P}^n = \mathbb{P}(k^{n+1}),$$

und nennen diese Menge (mit ihrem unten beschriebenen Strukturmerkmalen) den n-dimensionalen projektiven Raum über k .

Bemerkung 4.27. In der algebraischen Geometrie schreibt man meist $\mathbb{P}^n(k)$ oder $\mathbb{P}^n_k$. In der Differentialgeometrie schreibt man für $k = \mathbb{R}$ den Raum $\mathbb{P}^n$ als $\mathbb{P}(P^n)$ und für $k = \mathbb{C}$ den Raum $\mathbb{P}^n$ als $\mathbb{C}(P^n)$. Man kann sich $\mathbb{P}(n)$ als n-dimensionale Einheitssphäre im k^{n+1} vorstellen.

Definition 4.28. Es seien $x, y \in k^{n+1} \setminus \{0\}$. Wir definieren $x \sim y$, falls es ein $\lambda \in k \setminus \{0\}$ gibt, so dass für alle $0 \le 1 \le n$, $x_i = \lambda y_i$.

Bemerkung 4.29. Die Relation $\sim$ ist eine Äquivalenzrelation auf $k^{n+1} \setminus \{0\}$.

Definition 4.30. Es sei $x \in k^{n+1} \setminus \{0\}$. Wir bezeichnen mit

$$[x_0 : x_1 : \cdots : x_n]$$

die Äquivalenzklasse von x. Dieses bestimmt nun eindeutig eine Gerade im k^{n+1} und ist daher ein Punkt im $\mathbb{P}^n$. $x, y \in k^{n+1} \setminus \{0\}$ sind also $\sim$-äquivalent, wenn sie dieselbe Gerade im k^{n+1} bestimmen.

Beispiel 4.31.

(1) Für $n = 0$ gibt es genau eine Gerade im k^1. Diese Gerade ist der einzigen Punkt im $\mathbb{P}^0$. Jeder nulldimensionale projektive Raum besteht aus einem einzigen Punkt.

(2) Für $n = 1$ sind die Elemente des $\mathbb{P}^1 = \mathbb{P}(k^2)$ bestimmt durch $[x_0 : x_1]$, $(x_0, x_1) \ne (0, 0)$. Es gibt zwei Fälle: Falls $x_1 \ne 0$, kann man normieren und erhält die Repräsentanten $[x_0 : 1]$ für $x_0 \in k \setminus \{0\}$. Falls $x_1 = 0$, erhält man einen weiteren Repäsentanten $[0 : 1]$. Dieser heißt "der unendlich ferne Punkt".

Definition 4.32.

(1) Es sei $0 \le i \le n$ und $U_i = \{[x_0 : x_1 : \cdots : x_n] \in \mathbb{P}^n \ : \ x_i \ne 0\}$. U_i heißt die i-te standardaffine Teilmenge.

(2) Die i-te standardaffine Karte ist die folgende Abbildung $\varphi_i : U_i \to k^n$,

$$[x_0 : x_1 : \cdots : x_n] \mapsto \left(\frac{x_0}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_n}{x_i} \right).$$

Lemma 4.33.

(1) φ_i *is bijektiv mit Umkehrabbildung* $\psi_i : k^n \to U_i$,

$$(y_1, y_2, \ldots, y_n) \mapsto [y_1 : \cdots : y_i : 1 : y_{i+1} : \cdots : y_n]$$

(2) $\mathbb{P}^n \setminus U_i = \mathbb{P}(V_i)$ *mit* $V_i = \{x \in k^{n+1} : x_i = 0\}$.

(3) *Es gilt* $\mathbb{P}^n = \bigcup_{i=0}^{n} U_i$.

Bemerkung 4.34. Für $k = \mathbb{R}$ oder $k = \mathbb{C}$ machen diese Karten und ihre Umkehrungen $\mathbb{P}^n$ zu einer Mannigfaltigkeit.

Es sei $U \subseteq V$ ein Untervektorraum. Jede Nullpunktgerade in U ist eine Nullpunktgerade in V, also $\mathbb{P}(U) \subseteq \mathbb{P}(V)$.

Definition 4.35.

(0) Ab jetzt fassen wir $\mathbb{A}^n$ via ψ_0 als Teilmenge des $\mathbb{P}^n$ auf. ψ_0 ist nicht surjektiv. Es fehlt die Differenz

$$H_\infty = \mathbb{P}^n \setminus \mathbb{A}^n = \mathbb{P}(\{0\} \times k^n), \tag{4.1}$$

die unendlich ferne Hyperebene H_∞.

(1) Für einen Unterraum U von V heißt $\mathbb{P}(U)$ projektiver Teilraum von $\mathbb{P}(V)$.

(2) Es habe V die Dimension $n + 1$. Dann definiert man $\mathbb{P}(V)$ hat die Dimension n. Projektive Teilräume der Dimensionen 0, 1, 2, bzw. $n - 1$ heißen Punkte, Geraden, Ebenen bzw. Hyperebenen.

(3) Es seien H_1, H_2 projektive Teilräume. Der von H_1 und H_2 aufgespannte projektive Teilraum ist der kleinste projektive Teilraum, der H_1 und H_2 enthält. Wir schreiben $H_1 + H_2$ für diesen Teilraum.

Satz 4.36 (Dimensionsformel für projektive Teilräume)**.** *Es seien H_1 und H_2 zwei projektive Teilräume eines affinen Raums $\mathbb{P}^n$. Dann gilt*

$$\dim(H_1 + H_2) = \dim(H_1) + \dim(H_2) - \dim(H_1 \cap H_2).$$

Beweis: Wenn $H_i = \mathbb{P}(U_i)$, so ist $H_1 + H_2 = \mathbb{P}(U_1 + U_2)$. $\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)$. □

Beispiel 4.37. Es sei $n = 2$, wir sind also im $\mathbb{P}(k^3)$. Es seien $H_1 \neq H_2$ projektive Geraden. $H_1 \cap H_2$ enthält dann genau einen Punkt. Dann ist $H_1 + H_2 = \mathbb{P}^2$.

Beachten Sie, dass der folgende Satz von der Setzung, dass die projektive Koordinate x_0 die zusätzliche ist, abhängt. Wir legen fest, wie k^n in $\mathbb{P}^n$ liegt, die Karte ψ_0 ist dabei ausgewählt.

Satz 4.38. *Es sei $A \subseteq \mathbb{A}^n$ ein affiner Teilraum des $\mathbb{A}^n$ mit dem Vektorraum $V \subseteq k^n$, also $A = \{P + v \;:\; v \in V\}$. Dann gibt es einen eindeutig bestimmten projektiven Unterraum $\bar{A} \subseteq \mathbb{P}(k^{n+1})$, so dass*

$$\bar{A} \cap \mathbb{A}^n = A.$$

Es gilt $\bar{A} \cap H_\infty = \mathbb{P}(V)$.

Beweis: Es sei A d-dimensional und $r = n - d$. Dann gibt es eine $r \times n$-Matrix M und einen Vektor $b \in k^r$, so dass

$$A = \{x \in k^n \;:\; Mx = b.\}$$

Die Juxtaposition $M' := (-b|M)$ ist eine $r \times (n+1)$-Matrix. Es gilt

$$\begin{pmatrix} -b\, M \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0 \;\leftrightarrow\; M \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = b. \tag{4.2}$$

Wir nehmen nun $x' = (x_0, \ldots, x_n) \in k^{n+1}$ und setzen

$$\bar{A} = \{[x'] \;:\; M'x' = 0\}.$$

Die behauptete Gleichheit $\bar{A} \cap \mathbb{A}^n = A$ folgt aus Gleichung (4.2). Es gilt

$$\bar{A} \cap H_\infty = \{[x'] \;:\; [0 : x_1 : \cdots : x_n] \;:\; M'x' = 0\} = \{[0 : x] \;:\; Mx = 0\} = \mathbb{P}(V).$$

$\bar{A}$ ist eindeutig, da ja nach Definition A mit ψ_0 in $\bar{A}$ eingebettet wird. Es sei $U \subseteq k^{n+1}$ ein Untervektorraum so dass $\mathbb{P}(U) \cap \mathbb{A}^n = A$. Aus $A \subseteq \bar{A} = \mathbb{P}(U)$ folgt $(1, x_1, \ldots, x_n)^t \in U$ für alle $(x_1, \ldots, x_n)^t$ mit $Mx = 0$. Somit enthält U alle $x' = (1, x_1, \ldots, x_n)^t \in k^{n+1}$ mit $M'x' = 0$ und $\bar{A} \subseteq \mathbb{P}(U)$.

Sei nun umgekehrt $(u_0, \ldots, u_n)^t \in U$ und $\mathbb{P}(U) \cap \mathbb{A}^n = A$. Zunächst sei $u_0 = 0$. Dann ist $u = [u_0 : u_1 : \cdots : u_n] = [1 : u_1/u_0 : \cdots : u_n/u_0] \in \mathbb{P}(U) \cap \mathbb{A}^n = A \subseteq \bar{A}$ und $M'u = 0$.

Falls $u_0 = 0$, wählen wir einen Hilfspunkt $x = (x_0, \ldots, x_n)^t \in U$, so dass $x_0 \neq 0$. So einen Punkt gibt es das $A \neq \varnothing$ nach Gleichung (4.2). Dann hat auch $[x + u]$ keine 0 in der ersten Koordinate. Wir können also den ersten Fall auf x und auf $x + u$ anwenden und erhalten $M'x = 0$, $M'(x + u) = 0$. Somit ist $M'u = 0$. Also $\mathbb{P}(U) \subseteq \bar{A}$.

□

Lemma 4.39. *Es seien L_1, $L_2 \subseteq \mathbb{A}^n$ parallele Geraden. Dann schneiden sich $\bar{L_1}$ und $\bar{L_2}$ auf der unendlich fernen Hyperebene. Jeder Punkt der unendlich fernen Hyperebene entspricht genau einer Menge von parallelen Geraden im $\mathbb{A}^n$.*

Beweis: Nun ist $L_i = P_i + V$ und $\dim(V) = 1$. Der Satz liefert nun $\bar{L_i} \cap H_\infty = V$ für $i = 1, 2$. □

Lemma 4.40. *Es seien $[x] = [x_0 : x_1 : x_2]$, $[y]$ und $[z] \in \mathbb{P}^2$. Dann sind äquivalent:*

(1) Die drei Punkte liegen auf einer Geraden. Man sagt dazu, die projektiven Punkte sind kolinear .

(2) Die Vektoren (x_0, x_1, x_2), (y_0, y_1, y_2) und (z_0, z_1, z_2) sind linear abhängig.

Beweis: Falls $x = (x_0, x_1, x_2)^t$, y, z linear abhängig sind, spannen sie einen zweidimensionalen Vektorraum auf, also eine projektive Gerade. Im anderen Fall spannen x, y, z eine projektive Ebene auf. □

Wir schreiben im folgenden Satz $[(x_0, x_1, x_2)]$ für $x/\sim = [x_0 : x_1 : x_2]$. Die Addition ist die Addition von k^3.

Satz 4.41 (Der Satz von Desargues)**.** *Es seien zwei Dreiecke ABC und $A'B'C'$ in der projektiven Ebene gegeben. Gehen die Verbindungsgeraden der entsprechenden Punkte durch einen gemeinsamen Schnittpunkt, so liegen die Schnittpunkte einander entsprechender Seiten auf einer Gerade.*

Beweis: Wir schreiben $A = [a]$ mit $a \in k^3 \setminus \{0\}$. Jeder Punkt auf der Geraden durch A und A' hat die Form $[\lambda_1 a + \mu_1 a'] = [\lambda_1' a + \mu_1'(a' - a)]$, analog für B und B' mit $[\lambda_2 b + \mu_2 b']$ und C und C' mit $[\lambda_3 c + \mu_3 c']$ mit $\lambda_i, \mu_i \in k$. Der gemeinsame Schnittpunkt hat die Form

$$S = [\lambda_1 a + \mu_1 a'] = [\lambda_2 b + \mu_2 b'] = [\lambda_3 c + \mu_3 c'].$$

Wir nehmen Vielfache, so dass die Gleichheit gilt. Dann fassen wir die drei Gleichungen anders auf als

$$\begin{aligned} \lambda_1 a - \lambda_2 b = -\mu_1 a' + \mu_2 b' &=: \quad r_1; \\ \lambda_1 a - \lambda_3 c = -\mu_1 a' + \mu_3 c' &=: \quad r_2; \\ \lambda_2 c - \lambda_3 c = -\mu_2 b' + \mu_3 c' &=: \quad r_3. \end{aligned}$$

Da das Dreieck echt ist, sind die $r_i \neq 0$. Es sei $R_i = [r_i]$. Dann ist

$$R_1 \in AB \cap A'B', R_2 \in AC \cap A'C', R_3 \in BC \cap B'C'.$$

Es gilt

$$r_1 - r_2 + r_3 = 0$$

die drei Punkte sind also kolinear. □

Bemerkung 4.42. Im axiomatischen Zugang zur projektiven Geometrie gilt der Satz von Desargues nicht automatisch, sondern nur dann, wenn die projektive Ebene von einem Schiefkörper herkommt. Addition und Multiplikation von Punkten auf einer Geraden kann man rein geometrisch definieren. Der Satz von Desargues ist das Assoziativgesetz der Multiplikation.