LINEARE ALGEBRA UND ANALYTISCHE GEOMETRIE II (TU MÜNCHEN, SOMMERSEMESTER 2010)

BERNHARD HANKE

Die Referenz [Fischer] bezieht sich auf das Buch

Gerd Fischer: *Lineare Algebra*, Vieweg Studium, 15. Auflage.

Weiterhin nehmen wir des öfteren Bezug auf die Vorlesung *Lineare Algebra und Analytische Geometrie I* aus dem Wintersemester. Ein Skript ist im Internet verfügbar.

20.4.10

1. Eigenwerte, Diagonalisierung, Trigonalisierung

Es seien K ein Körper, V und W endlichdimensionale K-Vektorräume, $n := \dim V$, $m := \dim W$ sowie $f : V \to W$ eine K-lineare Abbildung. In der Vorlesung lineare Algebra 1 haben wir gesehen (Satz 4.17), dass nach geschickter Wahl von Basen $\mathcal{B}$ von V und $\mathcal{C}$ von W die darstellende Matrix $M^{\mathcal{B}}_{\mathcal{C}}(f) \in K^{m\times n}$ von f von der einfachen Gestalt

$$M^{\mathcal{B}}_{\mathcal{C}}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

ist, wobei $r :=$ rang(f).

Falls $V = W$, also $f \in \mathrm{End}(V)$, so ist die Forderung naheliegend, nicht zwei Basen (für Quelle und Ziel von f), sondern nur eine Basis von V zu variieren. In anderen Worten: Können wir eine Basis $\mathcal{B}$ von V finden, so dass $M^{\mathcal{B}}_{\mathcal{B}}(f) \in K^{n\times n}$, die *darstellende Matrix des Endomorphismus f bezüglich der Basis $\mathcal{B}$* von besonders einfacher Gestalt ist? Diese Frage ist nicht mehr einfach mit Satz 4.17 aus LinAlg 1 zu beantworten.

Es sei $\mathcal{C}$ eine beliebige Basis von V und $A := M^{\mathcal{C}}_{\mathcal{C}}(f) \in K^{n\times n}$. Bezeichnet $S := T^{C}_{B}$ die Matrix des Koordinatenwechsels von der Basis $\mathcal{C}$ in eine andere Basis $\mathcal{B}$ von V, so gilt (vgl. Prop. 4.10 aus LinAlg 1)

$$M^{\mathcal{B}}_{\mathcal{B}}(f) = S \cdot A \cdot S^{-1}$$

Wir können unser Problemstellung also folgendermaßen variieren: Gegeben sei eine Matrix $A \in K^{n\times n}$. Man finde eine Matrix $S \in \mathrm{GL}(n, K)$, so das SAS^{-1} eine besonders einfache Gestalt hat. In diesem Zusammenhang ist die folgende Sprechweise nützlich.

Definition. *Zwei Matrizen $A, B \in K^{n\times n}$ heißen* ähnlich*, falls es eine Matrix $S \in \mathrm{GL}(n, K)$ gibt mit $B = SAS^{-1}$.*

Dies definiert eine Äquivalenzrelation auf $K^{n\times n}$.

Definition. *Es sei $A \in K^{n\times n}$.*

- *Wir nennen A in* Diagonalform *oder eine* Diagonalmatrix*, falls*
$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$
wobei $\lambda_1, \ldots, \lambda_n \in K$. Die Matrix A hat also von Null verschiedene Einträge höchstens auf der Diagonalen.
- *Wir nennen A eine* obere Dreiecksmatrix*, falls*
$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$
mit $\lambda_1, \ldots, \lambda_n \in K$ und $*$ bedeutet, dass die Einträge oberhalb der Diagonalen beliebige Elemente aus K sind.
- *Wir nennen A* diagonalisierbar*, bzw.* trigonalisierbar*, falls A ähnlich zu einer Diagonalmatrix, bzw. zu einer oberen Dreiecksmatrix ist.*
- *Wir nennen einen Endomorphismus $f \in \mathrm{End}(V)$* diagonalisierbar*, bzw.* trigonalisierbar*, falls es eine Basis von V gibt, bezüglich der die darstellende Matrix von f eine Diagonalmatrix, bzw. eine obere Dreiecksmatrix ist.*

Gemäß dieser Definition ist also eine Matrix $A \in K^{n\times n}$ genau dann diagonalisierbar (trigonalisierbar), wenn die durch A beschriebene lineare Abbildung $K^n \to K^n$ (d.h. der entsprechende Endomorphismus in $\mathrm{End}(K^n)$) diagonalisierbar (trigonalisierbar) ist.

Jede Matrix $A \in K^{1\times 1}$ ist automatisch in Diagonalgestalt. Ist $A = (\lambda)$ so entspricht A der linearen Abbildung $K \to K$, $x \mapsto \lambda \cdot x$. Ist allgemeiner V ein eindimensionaler K-Vektorraum und $f : V \to V$ eine lineare Abbildung, so gibt es genau ein $\lambda \in K$, so dass $f(v) = \lambda v$ für alle $v \in V$. Insbesondere ist die darstellende Matrix von f bezüglich jeder Basis (v) von V eine (1×1)-Matrix von der Gestalt (λ). Diagonalmatrizen beschreiben also die einfachste mögliche Verallgemeinerung von Endomorphismen eindimensionaler Vektorräume auf höherdimensionale Räume.

Es sei $f \in \mathrm{End}(V)$ ein Endomorphismus, $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und A die darstellende Matrix von f bezüglich dieser Basis. Ist dann
$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$
in Diagonalgestalt, so gilt
$$f(v_i) = \lambda_i v_i$$
für alle $i = 1, \ldots, n$. Wollen wir umgekehrt zu einem Endomorphismus $f \in \mathrm{End}(V)$ eine Basis finden, bezüglich der f in Diagonalgestalt ist, so müssen wir also Vektoren $v \in V$, $v \neq 0$, suchen, so dass $f(v) = \lambda v$ für ein $\lambda \in K$.

Definition. *Es sei V ein beliebiger (nicht unbedingt endlichdimensionaler) K-Vektorraum und $f \in \mathrm{End}(V)$. Wir nennen $\lambda \in K$ einen* Eigenwert *von f, falls es ein $v \in V$ gibt mit $v \neq 0$ und $f(v) = \lambda v$. In diesem Fall heißt v ein* Eigenvektor *zum Eigenwert λ.*

In dieser Definition müssen wir $v \neq 0$ fordern, da die Gleichung $f(0) = \lambda \cdot 0$ für alle $\lambda \in K$ trivialerweise erfüllt ist. $0 \in K$ kann aber durchaus als Eigenwert auftreten. Für einen entsprechenden Eigenvektor v gilt dann $f(v) = 0$. Auf Eigenvektoren wirkt f in besonders einfacher Weise: Durch eine Streckung um einen gewissen Faktor.

Wir besprechen die Beispiele auf S. 223 f. in [Fischer].

22.4.10

Aus unserer obigen Diskussion folgt:

Proposition 1.1. *Es sei V endlichdimensional und $f \in \mathrm{End}(V)$. Dann ist f genau dann diagonalisierbar, wenn f eine Basis bestehend aus Eigenvektoren von f besitzt.*

Lemma 1.2. *Es sei V ein K-Vektorraum und $(v_i)_{i \in I}$ eine Familie von Eigenvektoren von f zu paarweise verschiedenen Eigenwerten. Dann ist diese Familie linear unabhängig.*

Beweis. Zum Testen der linearen Unabhängigkeit genügt es, endliche Teilfamilien der gegebenen Familie zu betrachten. Dieser Fall wird in [Fischer] auf Seite 226 diskutiert. □

Inbesondere kann ein n-dimensionaler Vektorraum höchstens n verschiedene Eigenwerte haben.

Eine weitere Folgerung ist:

Proposition 1.3. *Es sei V ein n-dimensionaler Vektorraum und $f \in \mathrm{End}(V)$ ein Endomorphismus mit n paarweise verschiedenen Eigenwerten. Dann ist f diagonalisierbar.*

Beweis. Es seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von f und $v_1, \ldots, v_n$ zugehörige Eigenvektoren. Dann ist nach Lemma 1.2 die Familie $(v_1, \ldots, v_n)$ linear unabhängig, wegen $\dim V = n$ also eine Basis von V. Also hat V eine Basis aus Eigenvektoren von f. □

Definition. *Ist V ein K-Vektorraum und $\lambda \in K$, so setzen wir*

$$\mathrm{Eig(f;\lambda) := \{v \in V \mid f(v) = \lambda v\} \subset V}\,.$$

Dies ist der zu λ gehörende Eigenraum.

Nach Definition ist also

$$\mathrm{Eig(f;\lambda) = ker(f - \lambda \cdot id_V)}$$

und daher ist $\mathrm{Eig(f;\lambda)}$ ein Untervektorraum von V. Weiterhin folgt direkt aus den Definitionen:

- $\mathrm{Eig(f;0) = ker(f)}$.

- $\lambda \in K$ ist Eigenwert von $f \Leftrightarrow \mathrm{Eig(f; \lambda)} \neq 0 \Leftrightarrow \dim \mathrm{Eig(f; \lambda)} > 0$. (Da wir die Dimension nur für endlichdimensionale Vektorräume definiert haben, nehmen wir für die letzte Äquivalenz $\dim V < \infty$ an.)
- Ist $\lambda \in K$ Eigenwert von f, so ist die Menge $\mathrm{Eig(f; \lambda)} \setminus \{0\}$ die Menge der zu λ gehörenden Eigenvektoren.

Wir erhalten somit auch:

Proposition 1.4. *Es sei V endlichdimensional und $f \in \mathrm{End}(V)$. Es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von f. Dann ist f genau dann diagonalisierbar, falls*

$$\dim \mathrm{Eig(f; \lambda_1)} + \ldots + \dim \mathrm{Eig(f; \lambda_k)} = \dim \mathrm{V} \ .$$

Beweis. Es sei $\dim V = n$. Falls f diagonalisierbar ist, hat V eine Basis $(v_1, \dots, v_n)$ von Eigenvektoren von f. Für jeden Eigenwert λ von V bilden die Vektoren aus dieser Basis zum Eigenwert λ eine Basis von $\mathrm{Eig(f; \lambda)}$. Da jeder Vektor aus der Basis $(v_1, \dots, v_n)$ zu genau einem Eigenwert gehört, folgt

$$\dim \mathrm{Eig(f; \lambda_1)} + \ldots + \dim \mathrm{Eig(f; \lambda_k)} = \dim \mathrm{V} \ .$$

Ist umgekehrt diese Gleichung erfüllt, fassen wir die Basen der Eigenräume $\mathrm{Eig(f; \lambda_1)}, \ldots, \mathrm{Eig(f; \lambda_k)}$ zu einer Familie von Vektoren in V zusammen. Diese Familie ist nach Lemma 1.2 linear unabhängig und hat nach Annahme die Länge n. Sie bildet also eine Basis von V aus Eigenvektoren. □

Es stellt sich die Frage, wie wir die Eigenwerte eines Endomorphismus bestimmen können. Wir untersuchen diese Frage hier nur für endlichdimensionale Vektorräume. Es sei also ab jetzt $\dim V = n$ und $f \in \mathrm{End}(V)$.

Nach Definition ist $\lambda \in K$ genau dann ein Eigenwert von f, falls

$$\ker(f - \lambda \cdot \mathrm{id}_V) \neq 0 \ .$$

Kombinieren wir dies mit der Dimensionsformel $\dim \ker(f - \lambda \, \mathrm{id}_V) + \dim \mathrm{im}\,(f - \lambda \, \mathrm{id}_V) = n$, so erhalten wir

Proposition 1.5. *Es ist $\lambda \in K$ genau dann ein Eigenwert von f, falls*

$$\mathrm{rang(f - \lambda \, id_V) < n}$$

Es sei nun $F \in \mathrm{End}(V)$ ein beliebiger Endomorphismus. Wir erinnern an die Definition der Determinante von F: Sei dazu $\mathcal{B}$ eine Basis von V. Wir setzen dann

$$\det(F) := \det(M^{\mathcal{B}}_{\mathcal{B}}(F)) \in K$$

Auf dem Tutoriumsblatt 12 aus dem Wintersemester wurde in Aufgabe 2 gezeigt, dass diese Definition nicht von der Auswahl von $\mathcal{B}$ abhängt. Weiterhin haben wir nach Prop. 6.9. aus LinAlg 1 die Äquivalenz

$$\mathrm{rangF < \dim V \Leftrightarrow \det(F) = 0} \, .$$

Indem wir nun $F := f - \lambda \cdot \mathrm{id}_V$ setzen, erhalten wir

Proposition 1.6. *Die folgenden Aussagen sind äquivalent:*

- *$\lambda \in K$ ist ein Eigenwert von f.*
- $\det(f - \lambda\, \mathrm{id}_V) = 0$.

Ist A die darstellende Matrix von f bezüglich einer (beliebigen) Basis von V, so ist

$$\det(f - \lambda\, \mathrm{id}_V) = \det(A - \lambda \cdot E_n)$$

mit der $(n \times n)$-Einheitsmatrix E_n. Diese Formel können wir benutzen, um die Eigenwerte von f in systematischer Weise zu bestimmen. Dazu ersetzen wir in der Matrix $A - \lambda \cdot E_n$ das Element $\lambda \in K$ durch eine Unbestimmte X und erhalten die Matrix $A - X \cdot E_n$ mit Koeffizienten im Polynomring $K[X]$. Die Determinante dieser Matrix ist nach der Leibnizformel ein Polynom in $K[X]$. Genauer erhalten wir nach Aufgabe 1 auf Übungsblatt 13 aus dem Wintersemester:

$$\det(A - X \cdot E_n) = (-1)^n X^n + (-1)^{n-1} \mathrm{spur(A)X^{n-1}} + \ldots + \det \mathrm{A}\,,$$

insbesondere ist $\det(A - X \cdot E_n)$ ein Polynom vom Grad n.

Definition. *Das Polynom* $\det(A - X \cdot E_n) \in K[X]$ *heißt das* charakteristische Polyon *von f und wird mit $P_f(X)$ bezeichnet.*

Oft (z.B. in [Fischer] wird die Unbestimmte auch mit t bezeichnet. Dies wollen wir ab jetzt auch so halten. Wir müssten dann also eher $P_f(t)$ schreiben. Wir schreiben allerdings der Einfachheit halber ab jetzt nur noch P_f.

Wir müssen an dieser Stelle allerdings noch zeigen, dass P_f nicht davon abhängt, bezüglich welcher Basis von V wir rechnen. Arbeiten wir aber mit einer anderen Basis von V und bezeichnet A' die darstellende Matrix von f bezüglich dieser neuen Basis, so gibt es eine Matrix $S \in \mathrm{GL}(n, K)$ mit $A' = SAS^{-1}$ und wir erhalten

$$\begin{aligned}\det(A' - t \cdot E_n) &= \det(SAS^{-1} - t \cdot E_n) = \det(SAS^{-1} - t \cdot (SE_nS^{-1}))\\ &= \det(S \cdot (A - tE_n) \cdot S^{-1}) = \det(A - tE_n)\end{aligned}$$

also ändert sich das Polynom $\det(A - tE_n) \in K[t]$ in der Tat nicht, wenn wir A durch die darstellende Matrix von f bezüglich einer anderen Basis von V ersetzen.

Aus unseren Betrachtungen folgt:

Proposition 1.7. *Es sei V ein endlichdimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Dann sind die folgenden Aussagen äquivalent:*

- *$\lambda \in K$ ist ein Eigenwert von f.*
- *λ ist Nullstelle des charakteristischen Polynoms $P_f \in K[t]$ von f.*

Beispiel. Wir betrachten die Matrix

$$A := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

die eine Drehung der Ebene gegen den Uhrzeigersinn um den Winkel $\pi/2$ beschreibt. Wir erhalten

$$P_A = 1 + t^2 \in \mathbb{R}[t] \,.$$

Über $\mathbb{R}$ hat dieses Polynom keine Nullstellen. Fassen wir jedoch A als Element von $\mathbb{C}^{2\times 2}$ auf, so hat das zugehörige charakterische Polyon $P_A(t) = 1+t^2 \in \mathbb{C}[t]$ die Nullstellen i und $-i$. Insbesondere ist dann A diagonalisierbar und ähnlich zur Diagonalmatrix

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} .$$

Dieses Beispiel zeigt sehr gut, wie die Diagonalisierbarkeit einer Matrix vom zugrunde liegenden Körper abhängen kann.

Ist nun $\lambda \in K$ ein Eigenwert von $A \in K^{n\times n}$, so kann man eine Basis des zugehörigen Eigenraumes bestimmen, indem man eine Basis von $\ker(A - \lambda E_n)$ berechnet. Dies macht man mit den Methoden aus dem Wintersemester (indem man zum Beispiel zunächst $A-\lambda E_n$ auf Zeilenstufenform bringt).

Erste Beispiele dazu finden sich in [Fischer] auf Seite 231 f.

27.4.10

Ist allgemeiner λ ein Eigenwert eines Endomorphismus $f \in \mathrm{End}(V)$ (wie immer ist hier $\dim V = n$), so wählt man zunächst eine Basis $\mathcal{B}$ von V, berechnet die darstellende Matrix A von f bezüglich dieser Basis und bestimmt anschließend eine Basis von $\ker(A - \lambda E_n) \subset K^n$. Die Vektoren in dieser Basis sind die Koordinatenvektoren (bezüglich der Basis $\mathcal{B}$) der Basisvektoren von $\mathrm{Eig(f;}\,\lambda)$.

Proposition 1.8.

- *Es sei $f \in \mathrm{End}(V)$ diagonalisierbar. Dann zerfällt P_f in Linearfaktoren.*
- *Das charakteristische Polynom P_f zerfalle in Linearfaktoren, wobei jede Nullstelle mit Vielfachheit 1 auftritt. Dann ist f diagonalisierbar.*

Beweis. Die erste Aussage folgt daraus, dass $\det(A-tE_n)$ in Linearfaktoren zerfällt, falls A eine Diagonalmatrix ist. Für die zweite Aussage beachte man, dass $\deg P_f = \dim V = n$, so dass unter der gegebenen Voraussetzung n paarweise verschiedene Eigenwerte von f existieren. Nach Proposition 1.3 ist dann f diagonalisierbar.

□

Es kann allerdings sein, dass f diagonalisierbar ist, aber Nullstellen von P_f in höherer Vielfachheit auftreten, z.B. ist für den (diagonalen) Endomorphismus $f : K^n \to K^n$, $x \mapsto \lambda x$ das charakteristische Polyonom $P_f = (\lambda - t)^n$. Wir halten zunächst fest:

Proposition 1.9. *Für alle $\lambda \in K$ gilt*

$$\dim \mathrm{Eig}(\mathrm{f}; \lambda) \leq \mu(\mathrm{P_f}; \lambda)$$

(die rechte Seite bezeichnet die Vielfachheit der Nullstelle λ des Polynoms P_f, vgl. S. 77 in LinAlg I).

Der Beweis findet sich in [Fischer] auf Seite 234 f.

Hier muss nicht immer Gleichheit auftreten, wie das Beispiel des *Jordanblocks* der Größe n zum Eigenwert λ

$$J(\lambda, n) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \in K^{n \times n}$$

zeigt. Wir erhalten für das charakteristische Polynom

$$P_{J(\lambda,n)} = (\lambda - t)^n \,,$$

jedoch ist

$$\dim \mathrm{Eig}(\mathrm{J}(\lambda, \mathrm{n}); \lambda) = 1 \,.$$

In diesem Zusammenhang kann man die Diagonalisierbarkeit von f auch folgendermaßen charakterisieren:

Satz 1.10. *Die folgenden Aussagen sind äquivalent:*

- *f ist diagonalisierbar.*
- *P_f zerfällt in Linearfaktoren und für alle Nullstellen λ von f gilt $\dim \mathrm{Eig}(\mathrm{f}; \lambda) = \mu(\mathrm{P_f}; \lambda)$.*
- *Sind $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Eigenwerte von f, so gilt*

  $$V = \mathrm{Eig}(\mathrm{f}; \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(\mathrm{f}; \lambda_k) \,.$$

  (Für die Diskussion direkter Familien von Unterverktorräumen siehe S. 36 f. im Skript zu Lin Alg I, insbesondere Proposition 3.14).

Der Beweis findet sich in [Fischer], S. 235 f.

Wir haben damit ein praktisches Verfahren zur Hand, wie wir bestimmen, ob f diagonalisierbar ist: In einem ersten Schritt prüft man, ob P_f in Linearfaktoren zerfällt (hierfür existiert kein allgemeiner Algorithmus; aber in vielen praktischen Fällen kann man trotzdem entscheiden, ob P_f in Linearfaktoren zerfällt oder nicht). Falls dies nicht der Fall ist, so ist f sicher nicht diagonalisierbar.

Zerfällt P_f in Linearfaktoren, so berechnet man für alle Nullstellen λ von P_f die Dimension von $\mathrm{Eig}(f; \lambda)$. Gilt für alle Nullstellen λ die Gleichheit $\dim \mathrm{Eig}(\mathrm{f}; \lambda) = \mu(\mathrm{P_f}; \lambda)$, so ist f diagonalisierbar.

Wir illustrieren dies an Hand von Beispiel 4.3.4 auf S. 236 in [Fischer].

Es fragt sich, was man im allgemeinen über den Endomorphismus f sagen kann, falls P_f in Linearfaktoren zerfällt. Ein erstes wichtiges Resultat in diese Richtung ist:

Satz 1.11. *Es sei $f \in \mathrm{End}(V)$. Dann sind äquivalent:*

- *P_f zerfällt in Linearfaktoren.*
- *f ist trigonalisierbar.*

Mit dem Fundamentalsatz der Algebra folgt also

Korollar 1.12. *Jeder Endomorphismus eines endlichdimensionalen komplexen Vektorraumes ist trigonalisierbar.*

29.4.10

Bevor wir das Theorem beweisen, holen wir noch eine Definition nach.

Definition. *Es sei V ein endlichdimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Es sei $\lambda \in K$ eine Nullstelle von P_f (d.h. ein Eigenwert von f).*

- *Die Ordnung der Nullstelle λ von P_f, also die Zahl $\mu(P_f;\lambda)$, heißt* algebraische Vielfachheit *von λ.*
- *Die Dimension* $\dim \mathrm{Eig(f;}\lambda)$ *heißt* geometrische Vielfachheit *von λ.*

In Proposition 1.9 haben wir gezeigt, dass die geometrische Vielfachheit eines Eigenwertes immer kleiner oder gleich dessen algebraischer Vielfachheit ist.

Wir kommen nun zum Beweis von Satz 1.11.

Beweis. Da das charakteristische Polynom einer oberen Dreiecksmatrix in Linearfaktoren zerfällt, ist eine Richtung einfach.

Die andere Richtung beweisen wir so, dass sich gleich ein algorithmisches Verfahren zur Trigonalisierung ergibt.

Wir setzen $W_1 := V$. Ohne Beschränkung der Allgemeinheit (durch Wahl einer Basis von V) können wir annehmen, dass $W_1 = V = K^n$. Insbesondere ist dann f durch eine Matrix $A_1 \in K^{n\times n}$ gegeben.

Es zerfalle nun P_f in Linearfaktoren. Es sei λ_1 ein Eigenwert von A_1 und $v_1 \in K^n$ ein zugehöriger Eigenvektor. Wir haben dann

$$P_{A_1} = (\lambda_1 - t)\cdot P',$$

wobei $P' \in K[t]$ ebenfalls in Linearfaktoren zerfällt.

Wir ergänzen nun die Familie (v_1) zu einer Basis $\mathcal{B}_2$ von K^n. Eine bequeme Möglichkeit ist $\mathcal{B}_2 = (v_1, e_1, \ldots, \hat{e}_{i_1}, \ldots, e_n)$, wobei i_1 so gewählt wurde, dass die i_1-te Komponente von v_1 ungleich 0 ist und das Hütchen bedeutet, dass die entsprechende Komponente ausgelassen wird ($(e_1, \ldots, e_n)$ ist die kanonische Basis von K^n).

Bezüglich der Basis $\mathcal{B}_2$ hat A_1 die Gestalt

$$S_1 A_1 S_1^{-1} = \begin{pmatrix} \lambda_1 & * \\ 0 & \\ \vdots & A_2 \\ 0 & \end{pmatrix}.$$

mit einer $((n-1)\times(n-1))$-Matrix A_2. Dabei ist $S_1 = T^{\mathcal{B}_1}_{\mathcal{B}_2}$ die Matrix der Koordinatentransformation von der Basis $\mathcal{B}_1 = (e_1, \ldots, e_n)$ in die Basis $\mathcal{B}_2$.

Wir bezeichnen mit $W_2 \subset W_1$ den von den Vektoren $(e_1, \ldots, \hat{e}_{i_1}, \ldots, e_n)$ aufgespannten Untervektorraum von W_1. Die Matrix A_2 definiert einen Endomorphismus $W_2 \to W_2$. Indem wir

$$P_{A_1} = \det \begin{pmatrix} \lambda_1 - t & & * & \\ 0 & & & \\ \vdots & & A_2 - tE_{n-1} & \\ 0 & & & \end{pmatrix}.$$

nach der ersten Spalte entwickeln, sehen wir, dass

$$P_{A_1} = (\lambda_1 - t) \cdot P_{A_2}$$

und da sich das Polynom P_{A_1} auf genau eine Weise durch $(\lambda_1 - t)$ mit Rest dividieren lässt, erhalten wir

$$P_{A_2} = P'\,.$$

Insbesondere zerfällt P_{A_2} in Linearfaktoren. Es sei λ_2 ein Eigenwert von A_2 und $v_2 \in W_2 \subset W_1$ ein zugehöriger Eigenvektor. Wir wählen nun ein $i_2 \neq i_1$, so dass die i_2-te Komponenten von v_2 (in der Basis $(e_1, \ldots, e_n)$) ungleich 0 ist (die i_1-te Komponenten ist automatisch gleich 0, da $v_2 \in W_2$).

Somit ist $\mathcal{B}_3 = (v_1, v_2, e_1, \ldots, \hat{e}_{i_1}, \ldots, \hat{e}_{i_2}, \ldots, e_n)$ wieder eine Basis von V und die darstellende Matrix von A_1 hat bezüglich $\mathcal{B}_3$ die Gestalt

$$S_2 A_1 S_2^{-1} = \begin{pmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \\ \vdots & \vdots & A_3 \\ 0 & 0 & \end{pmatrix}.$$

wobei $S_2 = T^{\mathcal{B}_1}_{\mathcal{B}_3}$ und $A_3 \in K^{(n-2)\times(n-2)}$.

Wir setzen nun $W_3 := \operatorname{span}(e_1, \ldots, \hat{e}_{i_1}, \ldots, \hat{e}_{i_2}, \ldots, e_n)$ und setzen das Verfahren induktiv fort.

Nach insgesamt $n-1$ Schritten erhalten wir eine Basis $(v_1, \ldots, v_n)$ von K^n, bezüglich der A_1 obere Dreiecksgestalt hat. $\square$

Wir illustrieren dieses Verfahren an Hand des Beispiels in [Fischer], Abschnitt 4.4.5. auf S. 246.

Die Diagonalisierung und Trigonalisierung von Endomorphismen hat eine wichtige Anwendung in der Theorie der gewöhnlichen Differentialgleichungen. Wir skizzieren hier nur die Methode ohne die entsprechende Theorie systematisch zu entwickeln. Näheres hierzu findet sich zum Beispiel in [O. Forster, Analysis 2, Vieweg Verlag, Kapitel 13 und 14].

Ziel ist es, Lösungen ϕ eines Systems von homogenen linearen Differentialgleichungen mit konstanten Koeffizienten der Gestalt

$$y' = Ay$$

zu finden. Dabei ist $A \in \mathbb{C}^{n\times n}$ eine quadratische Matrix mit komplexen Einträgen und $\phi = (\phi_1, \ldots, \phi_n)$ ist ein Tupel von komplexwertigen differenzierbaren Funktionen $\mathbb{R} \to \mathbb{C}$ mit $\phi'(t) = A\phi(t)$ für alle $t \in \mathbb{R}$. Wählen wir ein $t_0 \in \mathbb{R}$ und einen „Anfangswert" $c = (c_1, \ldots, c_n) \in \mathbb{C}^n$, so folgt aus der Theorie der gewöhnlichen Differentialgleichungen, dass es genau eine Lösung $\phi : \mathbb{R} \to \mathbb{C}^n$ der Gleichung $y' = Ay$ gibt, die der Anfangsbedingung $\phi(t_0) = c$ genügt. Diese Lösung ist unendlich oft differenzierbar.

Aus diesen Betrachtungen folgt, dass die Menge L der Lösungen der Differentialgleichung $y' = Ay$ ein komplexer Untervektorraum der Dimension n des Vektorraumes aller unendlich oft differenzierbaren Funktionen $\mathbb{R} \to \mathbb{C}^n$ ist. Ein Isomorphismus $L \cong \mathbb{C}^n$ ist dadurch gegeben, dass wir einer Lösung $\phi \in L$ den Wert $\phi(t_0) \in \mathbb{C}^n$ zuordnen (die Linearität ist klar; die Bijektivität folgt aus dem oben zitierten Existenz- und Eindeutigkeitssatz).

Eine Basis des Lösungsraumes bezeichnet man auch als *Lösungsfundamentalsystem* des Differentialgleichungssystems $y' = Ay$. Ist $(\phi^{(1)}, \ldots, \phi^{(n)})$ ein Lösungsfundamentalsystem, so folgt aus der allgemeinen Theorie, dass die Familie $(\phi^{(1)}(t), \ldots, \phi^{(n)}(t))$ aus Vektoren im $\mathbb{C}^n$ für alle $t \in \mathbb{R}$ linear unabhängig und somit eine Basis des $\mathbb{C}^n$ ist (nicht nur für $t = t_0$).

29.4.10

Der Einfachheit halber wählen wir im folgenden immer $t_0 := 0$.

Beispiel. Die eindeutig bestimmte Lösung der Differentialgleichung $y' = \lambda y$ mit Anfangswert $c \in \mathbb{C}$ zur Zeit $t = 0$ und $\lambda \in \mathbb{C}$ ist gegeben durch $\phi(t) = ce^{\lambda t}$. Damit ist $e^{\lambda t}$ ein Lösungsfundamentalsystem der Differentialgleichung $y' = \lambda y$.

Allgemeiner sei

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in \mathbb{C}^{n\times n}$$

eine Diagonalmatrix. Dann bilden die Funktionen $\phi^{(i)}(t) = (0, \ldots, 0, e^{\lambda_i t}, 0, \ldots, 0)$, $i = 1, \ldots, n$, ein Lösungsfundamentalsystem des Systems $y' = Ay$. Wir sagen in diesem Fall, das System ist *vollständig entkoppelt*.

In unserem Zusammenhang ist nun folgende Beobachtung wichtig.

Proposition 1.13. *Es sei $S \in \mathrm{GL}(n;\mathbb{C})$. Dann ist $\phi : \mathbb{R} \to \mathbb{C}^n$ genau dann Lösung der obigen Differentialgleichung, wenn die Funktion $S \cdot \phi : \mathbb{R} \to \mathbb{C}^n$ das Differentialgleichungssystem $y' = (SAS^{-1})y$ löst.*

Beweis. Die Gleichung $\phi'(t) = A\phi(t)$ ist gleichbedeutend mit

$$S\phi'(t) = SA\phi(t) = (SAS^{-1}) \cdot (S\phi(t)) \,.$$

Daher folgt die Behauptung aus der Gleichung $(S \cdot \phi)'(t) = S \cdot \phi'(t)$, die aus der Tatsache folgt, dass S konstante Einträge hat. $\square$

Ist also $A \in \mathbb{C}^{n \times n}$ diagonalisierbar und $c \in \mathbb{C}^n$, so bestimmen wir zunächst ein $S \in \mathrm{GL}(n;\mathbb{C})$, so dass $B := SAS^{-1}$ Diagonalgestalt hat und bestimmen die Lösung $\psi : \mathbb{R} \to \mathbb{C}^n$ des Differentialgleichungssystems $z' = Bz$ mit Anfangswert Sc. Dann ist $S^{-1} \cdot \psi$ eine Lösung des Differentialgleichungssystems $y' = Ay$ mit Anfangswert c.

Im allgemeinen ist jedoch A nicht diagonalisierbar, sondern nur trigonalisierbar. Ist nun $B \in \mathbb{C}^{n \times n}$ in oberer Dreiecksgestalt, so müssen wir also das Differentialgleichungssystem

$$\begin{aligned} z_1' &= b_{11} z_1 + \ldots + b_{1n} z_n \\ z_2' &= b_{22} z_2 + \ldots + b_{2n} z_n \\ \vdots &\quad \vdots \\ z_n' &= b_{nn} z_n \end{aligned}$$

lösen (mit einer passenden Anfangsbedingung). Dies kann man von unten nach oben durchführen, indem man die Lösung

$$\phi_n(t) = c e^{b_{nn} t}$$

in die vorletzte Gleichung einsetzt und dann die entstehende inhomogene Gleichung

$$z_{n-1}' - b_{n-1\,n-1} z_{n-1} = b_{n-1\,n} \phi_n$$

löst. Diese Lösung ist bei gegebener Anfangsbedigung ebenfalls eindeutig und kann explizit ermittelt werden. Dies macht man z.B. durch „Variation der Konstanten“, indem man zunächst die entsprechende homogene Gleichung

$$\psi' - b_{n-1\,n-1} \psi = 0$$

löst und dann für ψ_{n-1} den Ansatz

$$\psi_{n-1}(t) = \psi(t) \cdot u(t)$$

macht. Es ergibt sich für u die Lösung

$$u(t) = \int_0^t \psi(t)^{-1} b_{n-1\,n} \phi_n(t) dt + \text{const.}$$

Auf diese Weise arbeitet man das Differentialgleichungssystems von unten nach oben ab.

Genauere Informationen zur hier benötigten Theorie der gewöhnlichen Differentialgleichungen findet man zum Beispiel in [O. Forster, Analysis 2, Kapitel 13 und 14].

Wir illustrieren diese Methode am Beispiel der Differentialgleichung einer gedämpften Schwingung wie in [Fischer, S. 238].

Es handelt sich um eine lineare Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten

$$f'' + 2\mu f' + \omega^2 f = 0$$

wobei 2μ (mit $\mu \geq 0$) den Dämpfungsfaktor bezeichnet und ω^2 die Federkonstante. Dabei ist $f : \mathbb{R} \to \mathbb{C}$. Die Anfangsbedingungen sind gegeben durch $f(0) = \alpha$ und $f'(0) = \beta$ mit $\alpha, \beta \in \mathbb{R}$.

Diese Gleichung zweiter Ordnung schreiben wir in ein System von Gleichungen erster Ordnung

$$y' = Ay$$

um, wobei $y = (y_1, y_2) : \mathbb{R} \to \mathbb{C}^2$ und

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}$$

und $y_1(0) = \alpha$ und $y_2(0) = \beta$. Insbesondere gilt also $f = y_1$ und $f' = y_2$. Das charakteristische Polynom von A lautet

$$P_A(X) = X^2 + 2\mu X + \omega^2$$

und dieses hat die Nullstellen

$$X_{1/2} = -\mu \pm \sqrt{\mu^2 - \omega^2}\,.$$

Es ergeben sich folgende Fälle:

- $\mu > \omega$ (starke Dämpfung). In diesem Fall sind die Nullstellen verschieden und reelle negative Zahlen. Insbesondere ist A diagonalisierbar.
- $\mu = \omega$ (aperiodischer Grenzfall). Die beiden Nullstellen stimmen überein. In diesem Fall muss die Diagonalisierbarkeit von A noch eigens untersucht werden.
- $\mu < \omega$ (periodischer Fall). Die beiden Nullstellen sind rein komplex und verschieden. Die Matrix ist also wieder diagonalisierbar.

Im ersten Fall erhalten wir mit der Transformationsmatrix

$$S^{-1} = \begin{pmatrix} 1 & 1 \\ X_1 & X_2 \end{pmatrix}$$

dass

$$SAS^{-1} = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}$$

und somit hat die Gleichung $y' = Ay$ das Fundamentalsystem von Lösungen

$$\begin{aligned} \phi^{(1)} &= S^{-1} \begin{pmatrix} e^{X_1 t} \\ 0 \end{pmatrix} = \begin{pmatrix} e^{X_1 t} \\ X_1 e^{X_1 t} \end{pmatrix} \\ \phi^{(2)} &= S^{-1} \begin{pmatrix} 0 \\ e^{X_2 t} \end{pmatrix} = \begin{pmatrix} e^{X_2 t} \\ X_2 e^{X_2 t} \end{pmatrix} \end{aligned}$$

Da in der Gleichung zweiter Ordnung $y = y_1$ gilt, haben wir also als allgemeine Lösung der ursprünglichen Gleichung

$$f(t) = c_1 e^{X_1 t} + c_2 e^{X_2 t}\,,$$

wobei die Konstanten $c_1, c_2 \in \mathbb{C}$ so zu bestimmen sind, dass die Anfangsbedingungen erfüllt sind. Wir sehen insbesondere, dass für reelle c_1 und c_2 die Lösung ϕ für alle t reell ist.

6.5.10

2. Der Satz von Cayley-Hamilton und die Jordansche Normalform

Wir wollen die theoretische Untersuchung der Struktur von Endomorphismen endlichdimensionaler Vektorräume noch weiter ausführen. Insbesondere werden wir sehen, wie sich die Darstellung trigonalisierbarer Endomorphismen weiter verbessern lässt. Hierzu werden wir auch einige ringtheoretische Eigenschaften der Polynomringe $K[X]$ über Körpern K diskutieren und anwenden.

Ein theoretisch sehr wichtiges Zwischenresultat ist der Satz von Cayley-Hamilton.

Es sei V ein K-Vektorraum und $f \in \operatorname{End}(V)$. Wir betrachten die Teilmenge

$$K[f] := \{a_0 \operatorname{id}_V + a_1 f + \ldots + a_n f^n \mid n \in \mathbb{N}, a_0, \ldots, a_n \in K\} \subset \operatorname{End}(V)$$

Dies ist ein Unterring mit 1 von $\operatorname{End}(V)$ (d.h. $\operatorname{id}_V \in K[f]$ und $K[f]$ ist abgeschlossen unter den Ringverknüpfungen in $\operatorname{End}(V)$). Obwohl aber $\operatorname{End}(V)$ für $\dim V \geq 2$ nicht kommutativ ist, ist $K[f]$ immer ein kommutativer Ring. Diese Tatsache wird später nützlich, wenn wir Matrizen mit Einträgen in $K[f]$ betrachten.

Wir erhalten einen Auswertungshomomorphismus (t bezeichnet nach wie vor eine Unbestimmte)

$$\phi_f : K[t] \to K[f]\,, \sum_{i=0}^{n} a_i t^i \mapsto \sum_{i=0}^{n} a_i f^i\,.$$

Dabei setzen wir $f^0 := \operatorname{id}_V$. Diese Abbildung ist ein Ringhomomorphismus.

Wir formulieren nun das folgende wichtige theoretische Resultat.

Satz 2.1 (Satz von Cayley-Hamilton)**.** *Es sei* $\dim V < \infty$ *und* P_f *das charakteristische Polynon von* f. *Dann gilt* $P_f(f) = 0$. *Das heißt, setzen wir den Endomorphismus in sein eigenes charakteristisches Polynom ein, so erhalten wir die Nullabbildung* $V \to V$.

Beweis. Durch Wahl einer Basis von V können wir uns auf den Fall beschränken, dass $f = A \in K^{n \times n}$ mit einem $n \in \mathbb{N}$. Wir müssen dann $P_A(A) = 0$ zeigen.

Wir betrachten nun $(n \times n)$-Matrizen mit Einträgen in den kommutativen Ringen mit $K[t]$ und $K[A]$. Dabei beachte man, dass jedes Element λ aus K auch als Element von $K[t]$ (konstantes Polynom λ) und von $K[A]$ (lineare Abbildung $\lambda \cdot E_n$) aufgefasst werden kann.

Wir richten uns nun im weiteren Beweis nach der eleganten Darstellung in [Fischer], S. 252. □

11.5.10

Beispiel. Ist V endlichdimensional und $f \in \mathrm{End}(V)$ diagonalisierbar mit Eigenwerten $\lambda_1, \ldots, \lambda_k$, so erhalten wir eine direkte Summenzerlegung

$$V = \mathrm{Eig}(\mathrm{f}, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(\mathrm{f}, \lambda_\mathrm{k})\,.$$

Für $i = 1, \ldots, k$ ist

$$\mathrm{Eig}(\mathrm{f}, \lambda_\mathrm{i}) = \ker(\mathrm{f} - \lambda_\mathrm{i} \cdot \mathrm{id}_\mathrm{V})\,.$$

Setzen wir also

$$\Phi := (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_k)$$

so ist für alle $i = 1, \ldots, k$ die Abbildung $\Phi(f)$ gegeben durch

$$(f - \lambda_1 \cdot \mathrm{id}_V) \circ \ldots \circ \widehat{(f - \lambda_i \cdot \mathrm{id}_V)} \circ \ldots \circ (f - \lambda_k \cdot \mathrm{id}_V) \circ (f - \lambda_i \cdot \mathrm{id}_V)$$

da $K[f]$ kommutativ ist. Insbesondere ist für alle $i = 1, \ldots, k$ der Eigenraum $\mathrm{Eig}(\mathrm{f}, \lambda_\mathrm{i})$ im Kern von $\Phi(f)$ enthalten und wir erhalten

$$\Phi(f) = 0\,.$$

Das charakteristische Polynom von f lautet

$$P_f = (\lambda_1 - t)^{\mu_1} \cdot \ldots \cdot (\lambda_k - t)^{\mu_k}\,,$$

wobei $\mu_1, \ldots, \mu_k$ die algebraischen Vielfachheiten der Eigenwerte $\lambda_1, \ldots, \lambda_k$ sind. Also folgt wie oben die Aussage des Satzes von Cayley-Hamilton: $P_f(f) = 0$.

Es sei nun $A := J(\lambda, n)$ ein Jordanblock der Größe n zum Eigenwert λ (vgl. S. 7). Es gilt dann

$$P_A = (\lambda - t)^n$$

und wir haben

$$(\lambda \cdot E_n - A)^i \neq 0$$

für alle $0 \leq i < n$. Diese Beispiel zeigt , dass man im Satz von Cayley-Hamilton im allgemeinen die Vielfachheiten der Nullstellen von P_f berücksichtigen muss.

Diese Beobachtung legt es nahe, neben den Eigenräumen noch die sogannten *verallgemeinerten Eigenräume* eines Endomorphismus $f : V \to V$ zu betrachten (wie immer ist V ein endlichdimensionaler K-Vektorraum).

Definition. *Es sei $\lambda \in K$ ein Eigenwert von f mit algebraischer Vielfachheit $\mu := \mu(P_f, \lambda)$. Wir setzen*

$$\overline{\mathrm{Eig}}(f, \lambda) := \ker(f - \lambda \cdot \mathrm{id}_V)^\mu \subset V\,.$$

Dies ist der verallgemeinerte Eigenraum *zum Eigenwert λ.*

Dies ist offensichtlich ein f-invarianter Untervektorraum von V und es gilt $\mathrm{Eig}(\mathrm{f}, \lambda) \subset \overline{\mathrm{Eig}}(\mathrm{f}, \lambda)$.

Entscheidend ist nun die folgende Aussage.

Satz 2.2. *Es zerfalle P_f in Linearfaktoren,*

$$P_f = (t-\lambda_1)^{\mu_1} \cdot \ldots \cdot (t-\lambda_k)^{\mu_k}\,.$$

Dann gilt

$$V = \overline{\mathrm{Eig}}(f,\lambda_1) \oplus \ldots \oplus \overline{\mathrm{Eig}}(f,\lambda_k)\,.$$

Dies zeigt, dass wir für beliebige trigonalisierbare Endomorphismen immer eine direkte Summenzerlegung in verallgemeinerte Eigenräume haben. Dies ist im allgemeinen nicht richtig, wenn wir mit den gewöhnlichen Eigenräumen arbeiten, vgl. Theorem 1.10 (es sei denn, der gegebenen Endomorphismus ist diagonalisierbar).

Für den Beweis von Proposition 2.2 entwickeln wir etwas Teilbarkeitstheorie in Polynomringen. Dies ist ohnehin ein wichtiges Kapitel der klassischen Algebra.

Es bezeichne fortan R einen kommutativen Ring mit 1.

Definition. *Eine Teilmenge $I \subset R$ heißt* Ideal*, falls I eine additive Untergruppe von R ist und für alle $r \in R$ und $x \in I$ gilt, dass $r \cdot x \in I$.*

Nach Definition ist jedes Ideal auch ein Unterring von R (möglicherweise nicht mit 1).

Es ist aber nicht unbedingt jeder Unterring auch ein Ideal: Die Diagonalmatrizen in $\mathbb{R}^{2\times 2}$ bilden einen Unterring von $\mathrm{End}(\mathbb{R}^2)$, jedoch kein Ideal.

Ist $S \subset R$ eine Teilmenge, so ist die Menge

$$\{r_1s_1 + \ldots + r_ks_k \mid k \geq 0, s_1,\ldots,s_k \in S, r_1,\ldots,r_k \in R\}$$

ein Ideal in R. Dieses heißt das von S *erzeugte* Ideal und wird mit (S) bezeichnet. Dieses ist das kleinste Ideal in R, das die Menge S enthält, d.h. $S \subset (S)$ (dies gilt wegen $1 \in R$) und ist $I \subset R$ ein beliebiges Ideal mit $S \subset I$, so gilt $(S) \subset I$.

Definition. *Wir nennen $I \subset R$ ein* Hauptideal*, falls I von einem Element $a \in R$ erzeugt wird. Wir nennen R einen* Hauptidealring*, wenn jedes Ideal in R ein Hauptideal ist.*

Ist $I = (a)$, so gilt nach Definition $x \in I$ genau dann, wenn $a|x$.

Proposition 2.3. *Es sei K ein Körper. Dann ist $K[X]$ ein Hauptidealring.*

Beweis. Wir benutzen Division mit Rest in $K[X]$ (vgl. Satz 7.4. aus LinAlg I).

Es sei $I \subset K[X]$ ein Ideal. Falls $I = 0 := \{0\}$, so sind wir fertig. Andernfalls wählen wir ein Polynom $f \in I$ ungleich 0 und von minimalem Grad ≥ 0. Wir behaupten, dass $I = (f)$. Sei dazu $F \in I$ beliebig. Division mit Rest durch f führt auf eine Gleichung

$$F = g \cdot f + r$$

mit $g \in K[X]$ und $\deg r < \deg f$. Da $F, f \in I$, gilt auch $r = F - kf \in I$. Nach Wahl von f folgt $r = 0$. Also gilt $F \in (f)$. $\square$

Da wir auch in $\mathbb{Z}$ Division mit Rest durchführen können (vgl. Aufgabe 1 auf Blatt 4 der LinAlg I), ist $\mathbb{Z}$ ebenfalls ein Hauptidealring.

Definition. *Es seien $a_1, \ldots, a_k \in R$. Wir nennen $c \in R$ einen* größten gemeinsamen Teiler *der Elemente $a_1, \ldots, a_k$, falls $c|a_1, \ldots, c|a_k$ und falls jedes Element $d \in R$ mit $d|a_1, \ldots, d|a_k$ auch $d|c$ erfüllt. In diesem Fall schreiben wir $c = \mathrm{ggT}(a_1, \ldots, a_k)$.*

Wir nennen die Elemente $a_1, \ldots, a_k$ teilerfremd*, falls $1 = \mathrm{ggT}(a_1, \ldots, a_k)$.*

Man beachte, dass ein größter gemeinsamer Teiler in der Regel nicht eindeutig ist. Beispielsweise gilt in $\mathbb{Z}$ sowohl $2 = \mathrm{ggT}(6, 8)$ als auch $-2 = \mathrm{ggT}(6, 8)$.

Definition. *Ein Element $x \in R$ heißt* (multiplikative) Einheit*, falls es ein $y \in R$ gibt mit $xy = 1$.*

Ist R nullteilerfrei und sind e und e' zwei größte gemeinsame Teiler von $a_1, \ldots, a_k$, so gibt es eine Einheit $u \in R$ mit $e' = ue$ (siehe das Tutoriumsblatt). In diesem Fall ist also der größte gemeinsame Teiler jedenfalls bis auf Multiplikation mit einer Einheit eindeutig.

Proposition 2.4. *Es sei R ein Hauptidealring, und es seien $a_1, \ldots, a_k \in R$. Dann existiert $c = \mathrm{ggT}(a_1, \ldots, a_k)$ und es gilt*

$$(a_1, \ldots, a_k) = (c)\,.$$

Beweis. Da R ein Hauptidealring ist, gibt es ein $c \in R$ mit $(a_1, \ldots, a_k) = (c)$. Wir behaupten, dass $c = \mathrm{ggT}(a_1, \ldots, a_k)$. Da $a_1, \ldots, a_k \in (c)$, gilt $c|a_1, \ldots, c|a_k$. Ist $d|a_1, \ldots, d|a_k$, so gilt $(a_1, \ldots, a_k) \subset (d)$, somit (wegen $(a_1, \ldots, a_k) = (c)$) auch $(c) \subset (d)$ und damit gilt $d|c$. □

Konkret bedeutet diese Aussage: Ist R ein Hauptidealring und sind $a_1, \ldots, a_k \in R$, so existieren Elemente $b_1, \ldots, b_k \in R$ mit

$$b_1 a_1 + \ldots + b_k a_k = \mathrm{ggT}(a_1, \ldots, a_k)$$

Für den Ring $\mathbb{Z}$ ist das eine bekannte Aussage aus der elementaren Zahlentheorie.

Sind $f, g \in K[X]$ Polynome, die nicht beide gleich 0 sind, so bestimmt man $\mathrm{ggT}(f, g)$ mit dem *Euklidischen Algorithmus*. Ohne Einschränkung der Allgemeinheit sei $f \neq 0$ (falls $f = g = 0$, so ist $0 = \mathrm{ggT}(f, g)$). Wir setzen $r_0 := f$. Division mit Rest führt auf

$$g = q_1 r_0 + r_1$$

mit $\deg r_1 < \deg r_0$. Durch erneute Division mit Rest haben wir

$$\begin{aligned} r_0 &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \end{aligned}$$

Da immer $\deg r_{i+1} < \deg r_i$, erhalten wir nach endlich vielen Schritten

$$r_{n-1} = q_{n+1} r_n + 0$$

wobei $r_n \neq 0$.

Proposition 2.5. *Es gilt* $r_n = \mathrm{ggT}(\mathrm{f}, \mathrm{g})$.

Beweis. Es gilt $r_n \neq 0$ und $r_n | r_{n-1}$ aufgrund der letzten Gleichung. Verfolgen wir die Divisionen mit Rest von unten nach oben, sehen wir $r_k | r_{k-1}$ für $k = n, n-1, \dots, 1$. Die erste Gleichung zeigt dann $r_n | r_0$ und $r_n | g$. Somit ist r_n ein Teiler von f und g. Ist andererseits d ein beliebiger Teiler von f und g, so verfolgen wir die Divisionen mit Rest von oben nach unten und sehen $d | r_k$ für $k = 1, \dots, n$. □

Haben wir von Null verschiedene Polynome $f_1, \dots, f_k \in K[X]$ gegeben, so bestimmt man auf diese Weise zunächst $F_2 := \mathrm{ggT}(\mathrm{f_1}, \mathrm{f_2})$, anschließend $F_3 := \mathrm{ggT}(\mathrm{F_2}, \mathrm{f_3})$ und so fort. Es gilt dann $F_k = \mathrm{ggT}(\mathrm{f_1}, \dots, \mathrm{f_k})$.

Ganz ähnlich bestimmt man $\mathrm{ggT}(\mathrm{a_1}, \dots, \mathrm{a_k})$, falls $a_1, \dots, a_k \in \mathbb{Z}$.

20.5.10

In diesem Abschnitt sei immer R ein kommutativer Ring mit 1, der zudem nullteilerfrei ist.

Definition.

- *Wir nennen* $x \in R$ irreduzibel, *falls* $x \neq 0$ *und* x *keine Einheit ist und in jeder Darstellung* $x = ab$ *als Produkt zweier Elemente in* R *mindestens eines der Elemente* a *oder* b *eine Einheit ist.*
- *Wir sagen, ein Element* $x \in R$, $x \neq 0$, *besitzt eine* Faktorisierung in irreduzible Elemente, *wenn es eine Darstellung*
$$x = u \prod_{i=1}^{r} p_i$$
gibt, wobei $u \in R$ *eine Einheit ist und* $p_1, \dots, p_r \in R$ *irreduzibel sind. Hier ist* $r = 0$ *erlaubt, so dass nach Definition auch jede Einheit in* R *eine Faktorisierung in irreduzible Elemente besitzt.*
- *Wir nennen diese Faktorisierung* eindeutig, *wenn folgendes gilt: Ist*
$$x = u' \prod_{i=1}^{s} p'_i$$
eine zweite Darstellung mit einer Einheit u' *und irreduziblen Elementen* $p'_1, \dots, p'_s \in R$, *so gilt* $r = s$ *und es gibt Einheiten* $u_1, \dots, u_r \in R$, *so dass bis auf Permutation der Elemente* $p_1, \dots, p_r$ *gilt:* $p'_i = u_i p_i$.
- *Der Ring* R *heißt* faktoriell, *wenn jedes Element* $x \in R$ *mit* $x \neq 0$ *eine eindeutige Faktorisierung in irreduzible Elemente besitzt.*

Die für uns wichtige Aussage zu Polynomringen lautet nun:

Proposition 2.6. *Ist* K *ein Körper, so ist der Ring* $K[X]$ *faktoriell.*

Bevor wir dies zeigen, führen wir noch einen anderen Begriff ein, der von den ganzen Zahlen her wohlbekannt ist.

Definition. *Es sei R ein kommutativer Ring mit 1. Ein Element $p \in R$ heißt* Primelement*, falls folgendes gilt: Es ist $p \neq 0$ und p ist keine Einheit und sind $a, b \in R$ mit $p|(ab)$, so gilt $p|a$ oder $p|b$.*

Proposition 2.7. *Ist $p \in R$ ein Primelement, so ist p irreduzibel.*

Beweis. Angenommen $p = ab$. Dann gilt $p|a$ oder $p|b$. Sei ohne Beschränkung der Allgemeinheit $p|a$. Es gibt also ein $q \in R$ mit $a = pq$. Wir erhalten

$$a = abq \quad \Longrightarrow \quad a(1 - bq) = 0.$$

Wäre $a = 0$, so wäre auch $p = ab = 0$. Aber p ist prim und somit $p \neq 0$ per Definition. Aus der Nullteilerfreiheit von R folgt also $bq = 1$ und b ist somit eine Einheit. □

Die umgekehrte Impliktation ist nicht in jedem Ring richtig. Es gilt aber:

Proposition 2.8. *Es sei R ein Hauptidealring. Dann ist jedes irreduzible Element $x \in R$ ein Primelement.*

Beweis. Es gelte $x|(ab)$. Im Hauptidealring R existiert $c = \mathrm{ggT}(x, a)$. Es gibt also ein $d \in R$ mit $x = cd$. Da x irreduzibel ist, muss entweder c oder d eine Einheit sein. Ist d eine Einheit, so folgt $x|a$ aus $c|a$. Ist c eine Einheit so sind x und a teilerfremd. Nach Proposition 2.4 gibt es $\alpha, \beta \in R$ mit

$$\alpha a + \beta x = 1\,.$$

Wegen $x|(ab)$, teilt x also auch $\alpha ab = b - \beta xb$. Da x offensichtlich βxb teilt, gilt somit auch $x|b$. □

Insbesondere stimmen also die Begriffe „irreduzibel“ und „prim“ in den Ringen $\mathbb{Z}$ und $K[X]$ überein und in $\mathbb{Z}$ sind die Primelemente genau die bekannten Primzahlen (und ihre Negativen).

Wir können nun zeigen, dass wir in $K[X]$ eine „eindeutige Zerlegung in Primfaktoren“ haben.

Proposition 2.9. *Der Polynomring $K[X]$ ist faktoriell.*

Beweis. Es sei $f \in K[X]$ gegeben. Wir beweisen die Existenz einer Faktorisierung in irreduzible Faktoren per Induktion über den Grad $\deg f$. Jedes Polynom vom Grad 0 ist eine Einheit und besitzt daher eine solche Faktorisierung. Haben wir eine Zerlegung

$$f = f_1 \cdot f_2$$

wobei weder f_1 noch f_2 eine Einheit ist, so gilt $\deg f_1 < \deg f$ und $\deg f_2 < \deg f$. Per Induktion besitzen f_1 und f_2 Faktorisierungen in irreduzible Faktoren. Dies gilt somit auch für f.

Nun zur Eindeutigkeit. Es seien $f_1, \ldots, f_r, f'_1, \ldots, f'_s \in K[X]$ irreduzibel und

$$u f_1 \cdot \ldots \cdot f_r = u' f'_1 \cdot \ldots \cdot f'_s$$

mit Einheiten $u, u' \in R$. Wir führen Induktion nach r. Falls $r = 0$, haben wir $u = u' f_1' \cdot \ldots \cdot f_s'$. Dann sind $f_1', \ldots, f_s'$ Einheiten und es gilt $s = 0$. Damit ist die gewünschte Eindeutigkeit gezeigt.

Nun zum Induktionsschritt. Da f_1' irreduzibel ist, ist f_1' prim nach Proposition 2.8. Daher existiert ein $i = 1, \ldots, r$ mit $f_1' | f_i$. Durch Umordnen der Faktoren könne wir $i = 1$ annehmen. Da f_1 irreduzibel und f_1' keine Einheit ist, gibt es also eine Einheit $u_1 \in K[X]$ mit $f_1 = u_1 f_1'$. Nach Kürzen obiger Gleichung durch f_1' erhalten wir die Gleichung

$$u_1 \cdot u \cdot f_2 \cdot \ldots \cdot f_r = u' f_2' \cdot \ldots \cdot f_s' .$$

Nach Induktionsannahme folgt daraus $r = s$ und die irreduziblen Faktoren stimmen auf beiden Seiten bis auf die Reihenfolge und Multiplikation mit Einheiten überein. □

Ebenso zeigt man, dass $\mathbb{Z}$ faktoriell ist. Dies ist der bekannte Satz über die eindeutige Primfaktorzerlegung.

Proposition 2.10. *Es seien $\lambda_1, \ldots, \lambda_k \in K$ paarweise verschieden und es seien $\mu_1, \ldots, \mu_k$ positive ganze Zahlen. Für $i = 1, \ldots, k$ setzen wir*

$$\Phi_i := \prod_{1 \le j \le k, i \ne j}^{k} (X - \lambda_j)^{\mu_j} \in K[X] .$$

Dann sind die Polynome $\Phi_1, \ldots, \Phi_k$ teilerfremd.

Beweis. Falls $\Phi_1, \ldots, \Phi_k$ nicht teilerfremd sind, existiert ein normierter, irreduzibler gemeinsamer Teiler ϕ dieser Polynome.

Da ϕ auch ein Primelement ist und die irreduziblen Faktoren der Polynome $\Phi_1, \ldots, \Phi_k$ alle von der Gestalt $X - \lambda_i$ sind, gibt es dann ein $i \in I$ mit $\phi = X - \lambda_i$. Ist $j \ne i$, so sind aber die Polynome $X - \lambda_i$ und $X - \lambda_j$ teilerfremd (siehe Tutorium 4b). Insbesondere kann $\phi = X - \lambda_i$ kein Teiler von Φ_i sein. Widerspruch. □

Wir können nun den Satz über die Zerlegung in verallgemeinerte Eigenräume beweisen.

Satz 2.11. *Es sei V ein endlichdimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$ ein Endomorphismus, so dass P_f in Linearfaktoren zerfällt (dies ist z.B. immer der Fall, wenn $K = \mathbb{C}$). Dann gilt*

$$V = \overline{\mathrm{Eig}}(f, \lambda_1) \oplus \ldots \oplus \overline{\mathrm{Eig}}(f, \lambda_k) .$$

Beweis. Wir schreiben

$$P_f(X) = \pm (X - \lambda_1)^{\mu_1} \cdot \ldots \cdot (X - \lambda_k)^{\mu_k} .$$

und definieren $\Phi_i \in K[X]$ wie oben. Dann gilt

$$\mathrm{im}\,(\Phi_i(f)) \subset \overline{\mathrm{Eig}}(f; \lambda_i)$$

nach dem Satz von Cayley-Hamilton (!) und wegen

$$P_f = \pm (X - \lambda_i)^{\mu_i} \cdot \Phi_i .$$

Da nach Proposition 2.10 $\Phi_1, \ldots, \Phi_k$ teilerfremd sind, existieren Polynome $\xi_1, \ldots, \xi_k \in K[X]$ mit

$$\Phi_1 \xi_1 + \ldots + \Phi_k \xi_k = 1 \in K[X] .$$

Einsetzen der Abbildung f in die Unbekannte X liefert

$$\phi_1(f) \circ \xi_1(f) + \ldots + \phi_k(f) \circ \xi_k(f) = \mathrm{id}_V$$

und somit

$$v = \phi_1(f)(\xi_1(f)(v)) + \ldots + \phi_k(f)(\xi_k(f)(v))$$

für alle $v \in V$. Daher gilt

$$V = \overline{\mathrm{Eig}}(f, \lambda_1) + \ldots + \overline{\mathrm{Eig}}(f, \lambda_k) .$$

Es bleibt zu zeigen, dass die Summe direkt ist. Sei dazu $j \in \{1, \ldots, k\}$ und

$$v \in \overline{\mathrm{Eig}}(f, \lambda_j) \cap \sum_{i \neq j} \overline{\mathrm{Eig}}(f, \lambda_i) .$$

Daraus folgt $(f - \lambda_j \, \mathrm{id}_V)^{\mu_j}(v) = 0$ und $\Phi_j(f)(v) = 0$. Die Polynome $(X - \lambda_j)^{\mu_j}$ und Φ_j sind teilerfremd, also gibt es Polynome $\alpha, \beta \in K[X]$ mit

$$\alpha \cdot (X - \lambda_j)^{\mu_j} + \beta \cdot \Phi_j = 1$$

Setzen wir hier für X wieder die Abbildung f ein, so erhalten wir

$$(\alpha \cdot \mathrm{id}_V) \circ (f - \lambda_j \, \mathrm{id}_V)^{\mu_j} + \beta \cdot \Phi_j(f) = \mathrm{id}_V$$

und dies impliziert nach Einsetzen des Vektors v die Gleichung $0 = v$ wie gewünscht. $\square$

Für $i = 1, \ldots, k$ setzen wir nun

$$N_i := f - \lambda_i \cdot \mathrm{id} : \overline{\mathrm{Eig}}(f, \lambda_i) \to \overline{\mathrm{Eig}}(f, \lambda_i)$$

Die Abbildung N_i ist nilpotent und kommutiert (offensichtlich) mit der Abbildung $\lambda_i \cdot \mathrm{id} : \overline{\mathrm{Eig}}(f; \lambda_i) \to \overline{\mathrm{Eig}}(f; \lambda_i)$.

Der nächste Satz stellt nun den ersten wichtigen Schritt zur Jordanschen Normalform dar.

Im Beweis dieses Satzes benutzen wir folgende Schreibweise: Sind $U_1, \ldots, U_k$ und $V_1, \ldots, V_k$ Vektorräume über K, sowie $f_i : U_i \to V_i$ lineare Abbildungen für $i = 1, \ldots, k$, so definieren wir die lineare Abbildung

$$f_1 \oplus \ldots \oplus f_k : U_1 \oplus \ldots \oplus U_k \to V_1 \oplus \ldots \oplus V_k$$

durch $(u_1, \ldots, u_k) \mapsto (f_1(u_1), \ldots, f_k(u_k))$.

Satz 2.12 (Jordan-Chevalley-Zerlegung)**.** *Es sei V ein endlichdimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$ und es zerfalle P_f in Linearfaktoren. Dann gibt es einen diagonalisierbaren Endomorphismus $D : V \to V$ sowie einen nilpotenten Endomorphismus $N : V \to V$ mit*

- *$f = D + N$.*
- *$D \circ N = N \circ D$.*

Beweis. Auf $V = \overline{\mathrm{Eig}}(f;\lambda_1) \oplus \ldots \oplus \overline{\mathrm{Eig}}(f;\lambda_k)$ definieren wir

$$D := \lambda_1 \cdot \mathrm{id}_{\overline{\mathrm{Eig}}(f;\lambda_1)} \oplus \ldots \oplus \lambda_k \cdot \mathrm{id}_{\overline{\mathrm{Eig}}(f;\lambda_k)}$$

und

$$N := N_1 \oplus \ldots \oplus N_k$$

Dann sind die behaupteten Gleichungen erfüllt. □

27.5.10

Die praktische Durchführung der Jordan-Chevalley-Zerlegung ist einfach: Es sei A eine darstellende Matrix von f und es sei λ eine Nullstelle von $P_A = \det(A - t \cdot E_n) \in K[t]$ mit Vielfachheit μ. Man berechnet zunächst $\overline{\mathrm{Eig}}(A;\lambda) := \ker(A - \lambda \cdot E_n)^\mu \subset K^n$ durch Lösen des entsprechenden linearen Gleichungssystems. Diese Rechnung führt man für alle Eigenwerte von A durch und bestimmt so Basen für alle verallgemeinerten Eigenräume.

Wählen wir nun eine Basis von K^n der Gestalt

$$\mathcal{B} = (v_1^{(1)}, \ldots, v_{n_1}^{(1)}, \ldots, v_1^{(k)}, \ldots, v_{n_k}^{(k)})$$

wobei $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von A, $n_i = \dim \overline{\mathrm{Eig}}(A;\lambda_i)$ und $(v_1^{(i)}, \ldots, v_{n_i}^{(i)})$ eine Basis von $\overline{\mathrm{Eig}}(A;\lambda_i)$ ist, so hat die darstellende Matrix von A bezüglich $\mathcal{B}$ die Gestalt

$$\begin{pmatrix} \lambda_1 \cdot E_{n_1} + N_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \cdot E_{n_k} + N_k \end{pmatrix}$$

mit nilpotenten Matrizen $N_i \in K^{n_i \times n_i}$.

Bevor wir fortfahren, machen wir noch folgende Bemerkung zur Faktorialität von $K[X]$: Ist K ein Körper, so sind die Einheiten in $K[X]$ genau die Elemente in $K \setminus \{0\}$, aufgefasst als konstante Polynome in $K[X]$ (warum?). Wir können daher die Faktorialität von $K[X]$ folgendermaßen präzisieren. Jedes Polynom $f \in K[X]$, $f \neq 0$, besitzt eine Zerlegung

$$f = u \cdot f_1 \cdot \ldots \cdot f_r$$

wobei $u \in K \setminus \{0\}$ und $f_1, \ldots, f_r$ normierte Polynome in $K[X]$ vom Grad mindestens 1 sind (normiert bedeutet, dass die höchste Potenz von X in f mit Koeffizient 1 auftritt). Haben wir eine zweite Zerlegung dieser Art

$$f = u' \cdot f_1' \cdot \ldots \cdot f_s'$$

so gilt $u = u'$, $s = r$ und $f_i = f_i'$ für alle $i = 1, \ldots, r$ (eventuell nach Permutation der Faktoren $f_1, \ldots, f_s$).

Ist nun $f \in \mathrm{End}(V)$ ein Endomorphismus eines endlichdimensionalen K-Vektorraumes V, so ist die höchste Potenz von X in $P_f(X)$ von der Form $\pm X^{\dim V}$. Also gilt

$$P_f(X) = \pm f_1 \cdot \ldots \cdot f_r$$

wobei $f_1, \dots, f_r \in K[X]$ irreduzible normierte Polynome sind, die bis auf die Reihenfolge eindeutig bestimmt sind. Weiterhin tritt das positive Vorzeichen genau dann auf, wenn $\dim_K V$ gerade ist.

Wir können nun die Dimensionen der verallgemeinerten Eigenräume bestimmen:

Proposition 2.13. *Es sei V ein endlichdimensionaler K-Vektorraum, $f \in \mathrm{End}(V)$ und es zerfalle P_f in Linearfaktoren. Für alle $i = 1, \dots, k$ gilt dann:*

$$\dim \overline{\mathrm{Eig}}(f, \lambda_i) = \mu_i$$

Die Dimension des verallgemeinerten Eigenraumes zum Eigenwert λ_i ist also genau die Vielfachheit der entsprechenden Nullstelle λ_i im charakteristischen Polynom P_f.

Beweis. Da das charakteristische Polynom einer Matrix in Diagonal-Blockgestalt das Produkt der charateristischen Polynome der Blöcke ist, haben wir

$$P_f = P_{\lambda_1 \cdot E_{n_1} + N_1} \cdot \ldots \cdot P_{\lambda_k \cdot E_{n_k} + N_k} = \pm (X - \lambda_1)^{\mu_1} \cdot \ldots \cdot (X - \lambda_k)^{\mu_k} .$$

Wir berechnen nun

$$P_{\lambda_1 \cdot E_{n_i} + N_i} = \det(N_i - (X - \lambda_i) \cdot E_{n_i}) = P_{N_i}(X - \lambda_i) = \pm (X - \lambda_i)^{n_i}.$$

Dabei nutzen wir aus, dass das charakteristische Polynom der nilpotenten $n_i \times n_i$-Matrix N_i gegeben ist durch $\pm X^{n_i}$ (wegen der Normalform für nilpotente Matrizen, siehe Aufgabe 1 von Blatt 2).

Wir haben somit die Gleichung

$$\pm (X - \lambda_1)^{\mu_1} \cdot \ldots \cdot (X - \lambda_k)^{\mu_k} = \pm (X - \lambda_1)^{n_1} \cdot \ldots \cdot (X - \lambda_k)^{n_k}.$$

Da jedes Polynom der Form $X - \lambda_i$ irreduzibel ist, folgt aus unserer Vorbemerkung zur eindeutigen Primfaktorzerlegung im faktoriellen Polynomring $K[X]$, dass $\mu_i = n_i$ für alle $i = 1, \dots, k$. □

Wir betrachten nun wieder eine darstellende Matrix A von f bezüglich einer Basis $\mathcal{B}$ wie sie auf Seite 21 konstruiert wurde und für $i = 1, \dots, k$ den Unterraum $\overline{\mathrm{Eig}}(A, \lambda_i) \subset K^n$. Wir müssen untersuchen, wie wir durch Wahl einer geschickten Basis (die dann die Basis $(v_1^{(i)}, \dots, v_{n_i}^{(i)})$ von $\overline{\mathrm{Eig}}(A, \lambda_i)$ ersetzt) erreichen können, dass der nilpotente Operator N_i eine besonders einfache Gestalt hat.

Wir betrachten als Vorbereitung dazu Jordanblöcke. Bezüglich der Standardbasis $(e_1, \dots, e_n)$ des K^n beschreibt die nilpotente Matrix $J(0, n)$ diejenige Abbildung , die durch

$$e_1 \mapsto 0, \;\; e_k \mapsto e_{k-1} \text{ für } 2 \leq k \leq n$$

gegeben ist. Die Idee ist nun, auch für N solche „Ketten„ von Basisvektoren in K^n zu finden und diese Ketten zu einer Basis von K^n zu kombinieren. Die entsprechende Konstruktion erfordert etwas Fingerspitzengefühl.

Es sei $p \geq 1$ die minimale natürliche Zahl mit $N^p = 0$. Für $i = 0, \ldots, p$ setzen wir

$$V_i := \ker N^i \subset K^n \,.$$

Insbesondere ist

$$0 = V_0 \subset V_1 \ldots \subset V_p = K^n \,.$$

Außerdem haben wir für alle $i = 1, \ldots, p$ die Gleichung

$$N^{-1}(V_{i-1}) = V_i \,.$$

Es sei nun U_p ein Komplement von V_{p-1} in V_p, also

$$V_p = V_{p-1} \oplus U_p \,.$$

(Man vgl. die Definition auf S. 38 im Skript zu LinAlg I; wir erinnern insbesondere, dass dieses Komplement in der Regel nicht eindeutig ist). Es gelten nach Konstruktion die folgenden Aussagen:

- $N|_{U_p}$ ist injektiv, falls $p > 1$ (wegen $U_p \cap V_1 \subset U_p \cap V_{p-1} = 0$).
- $N(U_p) \cap V_{p-2} = 0$, falls $p > 1$ (denn ist $v \in U_p$, so gilt $N^i(v) \neq 0$ für alle $i \leq p-1$).
- $N(U_p) \subset V_{p-1}$ (denn es gilt $N(V_p) \subset V_{p-1}$).

Wir wählen nun ein Komplement U'_{p-1} von $N(U_p) \oplus V_{p-2}$ in V_{p-1} und setzen $U_{p-1} := N(U_p) \oplus U'_{p-1}$. Wir erhalten die folgenden Aussagen:

- U_{p-1} ist ein Komplement von V_{p-2} in V_{p-1}.
- $N(U_p) \subset U_{p-1}$.

Dieses Verfahren setzen wir fort und erhalten Untervektorräume $U_1, \ldots, U_p \subset K^n$ mit den folgenden Eigenschaften für alle $i = 1, \ldots, p$.

- $V_i = V_{i-1} \oplus U_i$ für alle $i = 1, \ldots, p$. Insbesondere ist $U_1 = V_1$.
- $N(U_i) \subset U_{i-1}$ für alle $i = 2, \ldots, p$.
- $N|_{U_i}$ ist injektiv für alle $i = 2, \ldots, p$.

Nach Konstruktion der U_i gilt

$$K^n = V_p = V_{p-1} \oplus U_p = (V_{p-2} \oplus U_{p-1}) \oplus U_p = U_1 \oplus \ldots \oplus U_p \,.$$

Es sei nun

$$(v_1^{(p)}, \ldots, v_{d_p}^{(p)})$$

eine Basis von U_p. Falls $p = 1$, sind wir fertig. Ansonsten ist $N|_{U_p}$ injektiv und wir erhalten linear unabhängige Vektoren

$$v_1^{(p-1)} := N(v_1^{(p)}),\ \ldots,\ v_{d_p}^{(p-1)} := N(v_{d_p}^{(p)})$$

in U_{p-1}, die wir zu einer Basis

$$(v_1^{(p-1)}, \ldots, v_{d_p+d_{p-1}}^{(p-1)})$$

von U_{p-1} ergänzen. Insbesondere ist also $d_{p-1} = \dim U_{p-1} - \dim U_p$. Falls $p = 2$, sind wir jetzt fertig. Ansonsten wenden wir N auf diese Familie an und erhalten eine linear unabhängig Familie

$$(v_1^{(p-2)}, \ldots, v_{d_p+d_{p-1}}^{(p-2)})$$

in U_{p-2}, die wir zu einer Basis

$$(v_1^{(p-2)}, \dots, v_{d_p+d_{p-1}+d_{p-2}}^{(p-2)})$$

von U_{p-2} ergänzen. Entsprechend konstruieren wir Basen

$$(v_1^{(i)}, \dots, v_{d_p+\dots+d_i}^{(i)})$$

von U_i für alle $i = p, \dots, 1$.

Wir ordnen nun die so erhaltenen Basen der U_i in folgendem Schema

$$\begin{array}{l|llllllllll}
U_p & v_1^{(p)} & \dots & v_{d_p}^{(p)} & & & & & & \\
U_{p-1} & v_1^{(p-1)} & \dots & v_{d_p}^{(p-1)} & v_{d_p+1}^{(p-1)} & \dots & v_{d_p+d_{p-1}}^{(p-1)} & & & \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & & & \\
U_1 & v_1^{(1)} & \dots & v_{d_1}^{(1)} & v_{d_p+1}^{(1)} & \dots & v_{d_p+d_{p-1}}^{(1)} & v_{d_p+\dots d_2+1}^{(1)} & \dots & v_{d_p+\dots+d_1}^{(1)}
\end{array}$$

Die Vektoren in einer Zeile bilden jeweils eine Basis des angegebenen Unterraumes U_i. Die Vektoren in der untersten Zeile bilden also eine Basis von $U_1 = V_1 = \ker N$. Da

$$K^n = U_1 \oplus \dots \oplus U_p$$

ist die Vereinigung dieser Basen eine Basis von K^n. Gehen wir in diesem Schema von einem Vektor ausgehend eine Zeile tiefer, so entspricht dies genau der Anwendung von N.

Daher bildet jede Spalte (von unten nach oben) in diesem Schema die Basis eines N-invarianten Unterraumes $W \subset K^n$. Die lineare Unabhängigkeit folgt dabei aus der eben angegebenen direkten Summenzerlegung für K^n oder auch aus Aufgabe 4b auf Übungsblatt 8 aus LinAlg I. Weiterhin ist $N|_W$ ist bezüglich dieser Basis durch einen Jordanblock der Form $J(0, \nu)$ gegeben, wobei ν die Länge der betrachteten Spalte bezeichnet.

1.6.10

Bemerkung 2.14. *Ergänzend zur Vorlesung habe ich ein explizites Beispiel für die Berechnung der Normalform einer nilpotenten Matrix eingefügt.*

In letzter Zeit habe ich (versehentlich) die Schreibweise $\mathrm{Mat}(n \times n, K)$ *für den Ring der* $(n \times n)$*-Matrizen über* K *verwendet, die jedoch mit der in LinAlg 1 eingeführten Schreibweise* $\mathrm{Mat}(n, K)$ *kollidiert. Ich möchte wieder zu dieser Notation zurückkehren.*

Auf Seite 7 habe ich den Jordanblock der Größe n *zum Eigenwert* λ *mit* $J(\lambda, n)$ *bezeichnet. Gelegentlich habe ich der Vorlesung hier die Rollen von* λ *und* n *vertauscht. Ich hoffe, in den folgenden Notizen ist alles korrekt.*

Beim Studium der folgenden Ausführungen sollte man zunächst die Existenz der Jordanschen Normalform und ihre Konstruktion im Auge behalten. Das genaue Verständnis der Eindeutigkeit ist beim ersten Durchlesen (noch) nicht entscheidend.

Aus den Betrachtungen der letzten Vorlesung folgt:

Satz 2.15 (Normalform nilpotenter Matrizen)**.** *Es sei* $N \in \mathrm{Mat}(n, K)$ *nilpotent. Dann ist* N *ähnlich zu einer Matrix der Form*

$$\begin{pmatrix} J(0,n_1) & 0 & 0 & 0 \\ 0 & J(0,n_2) & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & J(0,n_r) \end{pmatrix}.$$

Die Anzahl r *der Jordanblöcke ist durch* $\dim(\ker N)$ *gegeben und die Größen* $n_1, \ldots, n_r$ *der Jordanblöcke sind ebenfalls eindeutig durch* N *festgelegt (bis auf die Reihenfolge der* n_j*).*

(Zur Definition der Jordanblöcke $J(\lambda, n)$ siehe S. 7)

Beweis. Die erste Aussage haben wir gerade mit Hilfe des Schemas auf Seite 24 bewiesen.

Die anderen Aussagen sind schwerer zu zeigen. Wir argumentieren wie folgt: Es sei eine Matrix J in obiger Form gegeben (d.h. mit Jordanblöcken zum Eigenwert 0 auf der Diagonalen), die ähnlich zu N ist. Wir schreiben die Vektoren einer Basis des zum j-ten Jordanblock $J(0, n_j)$, $1 \leq j \leq r$, gehörigen Untervektorraumes des K^n in die j-te Spalte eines Schemas, wobei wir die Basisvektoren von unten nach oben anschreiben. Dies liefert ein Schema ähnlich dem auf S. 24, wobei aber die Längen der Spalten nicht unbedingt nach rechts hin abnehmen müssen (diese Längen sind ja durch die n_j gegeben).

Wir beobachten nun, dass für alle $i = 1, \ldots, p$ (p ist wieder die minimale Zahl mit $N^p = 0$) der Untervektorraum $V_i = \ker N^i \subset K^n$ von allen Vektoren aufgespannt wird, die in diesem Schema in den ersten i Zeilen stehen (von unten gezählt). Für alle $1 \leq j \leq i$ ist nun die Anzahl der Vektoren in der j-ten Zeile des Schemas (von unten gezählt) gleich der Anzahl der Jordanblöcke der Größe mindestens j in J. Daher ist für alle $i = 1, \ldots, p$

$$\dim V_i = \sharp\{J(0, n_j) \mid n_j \geq 1\} + \ldots + \sharp\{J(0, n_j) \mid n_j \geq i\}.$$

Aufgrund dieser Beziehung ist die Anzahl der Jordanblöcke der Größe mindestens 1 (also aller Jordanblöcke) gegeben durch $\dim V_1$, die Anzahl der Jordanblöcke der Größe mindestens 2 gegeben durch $\dim V_2 - \dim V_1$ und allgemeiner die Anzahl der Jordanblöcke der Größe mindestens i (mit $1 \leq i \leq p$) durch $\dim V_i - \dim V_{i-1}$. Wenn wir für alle $i = 1, \ldots, p$ wissen, wie viele Jordanblöcke der Größe mindestens i existieren, können wir (da die maximale Größe gleich p ist) die Zahlen $n_1, \ldots, n_r$ berechnen. Damit sind die Zahlen $n_1, \ldots, n_r$ durch die Dimensionen der V_i und damit durch N eindeutig festgelegt. □

In dem Schema oben auf Seite 24 treten für $i = 1, \ldots p$ genau d_i viele Spalten der Länge i auf und damit gibt es genau d_i Jordanblöcke der Größe i.

Beispiel. Das ab Seite 22 beschriebene Verfahren zur Berechnung der Normalform nilpotenter Matrizen illustrieren wir an der Matrix

$$A := \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{C})$$

von Seite 270 in [Fischer]. Das charakteristische Polynom lautet

$$P_A(t) = -(t-2)^3 \, .$$

Somit ist 2 der einzige Eigenwert, und dieser hat algebraische Vielfachheit 3. Nach der Struktur der Jordan-Chevalley-Zerlegung (bzw. direkt nach dem Satz von Cayley-Hamilton) ist die Matrix $N := A - 2 \cdot E_2$ nilpotent und $N^3 = 0$. Somit gilt (in der Schreibweise von S. 22) $V_3 = \mathbb{C}^3$. Wir haben

$$V_1 = \ker N = \mathrm{span} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

somit ist die geometrische Vielfachheit des Eigenwertes λ gleich 1. Weiterhin haben wir

$$V_2 = \ker N^2 = \mathrm{span} \left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right) .$$

Und $p = 3$ ist die kleinste Zahl mit $N^p = 0$. Ein Komplement U_3 von V_2 in V_3 wird beispielsweise von

$$v_1^{(3)} := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

aufgespannt (denn V_2 und dieser Vektor spannen $\mathbb{C}^3$ auf). Wir setzen also

$$U_3 := \mathrm{span} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

Wir bestimmen nun ein Komplement U_2' von $\mathrm{span}(N(v_1^{(3)})) \oplus V_1$ in V_2. Da jedoch $\dim V_2 = 2$ ist $U_2' = 0$ und wir erhalten

$$U_2 := \mathrm{span}(v_1^{(2)}) \, ,$$

wobei $v_1^{(2)} := N(v_1^{(3)}) = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$. Da $\dim V_1 = 1$ haben wir weiterhin

$$U_1 := \mathrm{span}(v_1^{(1)})$$

mit $v_1^{(1)} := N(v_1^{(2)}) = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix}$. Das auf Seite 24 oben angegebene Schema besteht also aus genau einer Spalte der Höhe 3 und bezüglich der Basis

$(v_1^{(1)}, v_1^{(2)}, v_1^{(3)})$ wird A durch den Jordanblock

$$J(2,3) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

dargestellt. Die Matrix S der Koordinatentransformation in die neue Basis ist gegeben durch

$$S^{-1} = \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix}$$

also

$$S = \frac{1}{4} \begin{pmatrix} -1 & -3 & 0 \\ 2 & 2 & 0 \\ 0 & 4 & 4 \end{pmatrix} .$$

Insbesondere ist $SAS^{-1} = J(2,3)$.

Kombinieren wir die Normalform für nilpotente Matrizen mit der Jordan-Chevalley-Zerlegung, so erhalten wir das folgende wichtige Ergebnis der klassischen linearen Algebra:

Satz 2.16 (Existenz und Eindeutigkeit der Jordanschen Normalform)**.** *Es sei $f \in \mathrm{End}(V)$ ein Endomorphismus eines endlich-dimensionalen K-Vektorraumes V. Das charakteristische Polynom P_f zerfalle in Linearfaktoren. Dann wird f bezüglich einer geeigneten Basis von V durch eine Matrix der Form*

$$\begin{pmatrix} J(\lambda_1, m_1) & 0 & 0 & 0 \\ 0 & J(\lambda_2, m_2) & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & J(\lambda_q, m_q) \end{pmatrix}$$

dargestellt.

Dabei sind $m_1, \ldots, m_q$ positive natürliche Zahlen mit $m_1 + \ldots + m_q = \dim V$ und $\lambda_1, \ldots, \lambda_q$ sind (nicht unbedingt paarweise verschiedene) Eigenwerte von f.

Zu jedem Eigenwert $\lambda \in K$ treten genau $\dim \mathrm{Eig}(\mathrm{f}, \lambda)$ *viele Jordanblöcke zu diesem Eigenwert auf. Die Größen der auftretenden Jordanblöcke zum Eigenwert λ hängen (bis auf die Reihenfolge) nur von f ab, die Summe dieser Größen ist gleich der algebraischen Vielfachheit von λ in P_f und diese stimmt mit $\dim_K \overline{Eig}(f, \lambda)$ überein.*

Beweis. Wir können eine Basis so wählen (siehe S. 21), dass die entsprechende darstellende Matrix von f von der Gestalt

$$A = \begin{pmatrix} \lambda_1 \cdot E_{n_1} + N_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \cdot E_{n_k} + N_k \end{pmatrix}$$

mit nilpotenten Matrizen $N_i \in K^{n_i \times n_i}$ ist. Dabei ist $n_1 + \ldots + n_k = \dim V$ und $\lambda_1, \ldots, \lambda_k$ sind die paarweise verschiedenen Eigenwerte von f.

Wir finden nun für alle $i = 1, \ldots, k$ Matrizen $S_i \in \mathrm{GL}(m_i; K)$, so dass $S_i N_i S_i^{-1}$ in einer Normalform vorliegt, wie sie in Theorem 2.15 angegeben wurde. Da

$$S_i(\lambda_i \cdot E_{n_i} + N_i)S_i^{-1} = \lambda_i \cdot E_{n_i} + S_i N_i S_i^{-1}$$

eine Blockdiagonalmatrix mit Jordankästchen zum Eigenwert λ_i ist, ist

$$\begin{pmatrix} S_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & S_k \end{pmatrix} A \begin{pmatrix} S_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & S_k \end{pmatrix}^{-1}$$

von einer Gestalt wie sie in Theorem 2.16 behauptet wird.

Die Eindeutigkeitssaussage folgt daraus, dass für $i = 1, \ldots, k$ die Matrix N_i die darstellende Matrix der Abbildung

$$(f - \lambda_i \cdot \mathrm{id}_V)|_{\overline{Eig}(f,\lambda_i)}$$

ist (und somit nur von f abhängt) und aus der Eindeutigkeitsaussage in Theorem 2.15.

$\square$

Aufgrund der Eindeutigkeitsaussage nennen wir die Paare $(\lambda_1, m_1), \ldots, (\lambda_q, m_q)$ auch die *Jordan-Invarianten* des Endomorphismus f. Diese hängen (bis auf die Reihenfolge) nur von f ab. Die Jordansche Normalform eines Endomorphismus f ist in gewisser Hinsicht die „einfachste“ Möglichkeit, f durch eine Matrix zu beschreiben.

Wir fassen die Konstruktion einer Jordanschen Normalform zu einer Matrix $A \in \mathrm{Mat}(n, K)$, bei dem P_A in Linearfaktoren zerfällt (was zum Beispiel für $K = \mathbb{C}$ immer der Fall ist), nochmal zusammen:

- Man bestimmt zunächst die Eigenwerte $\lambda_1, \ldots, \lambda_k$ von A. Es sei μ_i die algebraische Vielfachheit zum Eigenwert λ_i für $i = 1, \ldots, k$.
- Man berechne Basen von $\overline{Eig}(A, \lambda_i) = \ker(A - \lambda_i E_n)^{\mu_i}$ für $i = 1, \ldots, k$ und stelle $(A - \lambda_i E_n)|_{\overline{Eig}(A,\lambda_i)}$ bezüglich dieser Basen durch nilpotente Matrizen N_i dar.
- Für N_i bestimme man eine Matrix S_i, so dass $S_i N_i S_i^{-1}$ eine nilpotente Matrix in Normalform wie in Theorem 2.15 ist.
- Aus diesen Daten berechne man eine Matrix $S \in \mathrm{GL}(n, K)$, so dass SAS^{-1} eine Matrix in Jordanscher Normalform ist.

Ist eine Matrix in Jordanscher Normalform gegeben, so lassen sich ihre Potenzen besonders bequem berechnen. Dies veranschaulichen wir einerseits auf dem Übungsblatt und andererseits durch folgende Bemerkung zur Lösungstheorie von Systemen homogener Differentialgleichungen erster Ordnung.

Wir nehmen Bezug auf die Exponentialabbildung für Matrizen

$$\exp : \mathrm{Mat}(n, \mathbb{C}) \to \mathrm{Mat}(n, \mathbb{C})$$

definiert durch
$$\exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!}\,,$$
vgl. Blatt 4, Aufgabe 4. Wir betrachten weiterhin die Abbildung
$$\phi : \mathbb{R} \to \mathrm{Mat}(n, \mathbb{C})\,,\; t \mapsto \exp(t \cdot A)$$

Proposition 2.17. *Die Abbildung*
$$\phi : \mathbb{R} \to \mathrm{Mat}(n; \mathbb{C}) = \mathbb{R}^{(2n)^2}$$
ist differenzierbar. Wir erhalten für die Ableitung
$$\phi'(t) = A \cdot \phi(t)$$
für alle $t \in \mathbb{R}$.

Beweis. Für alle $t \in \mathbb{R}$ existiert (wir betrachten hier Konvergenz in $\mathrm{Mat}(n; \mathbb{C})$ wie üblich mit der Operatornorm)
$$\lim_{h\to 0, h\neq 0} \frac{\phi(t+h) - \phi(t)}{h} = \lim_{h\to 0, h\neq 0} \frac{\exp(tA)\cdot\exp(hA) - \exp(tA)}{h}$$
da die Matrizen tA und hA vertauschen (vgl. Blatt 5, Aufgabe 4). Der letzte Ausdruck ist gleich
$$\lim_{h\to 0, h\neq 0} \frac{\exp(hA) - 1}{h} \cdot \exp(tA) = A \cdot \exp(tA)\,,$$
denn
$$\lim_{h\to 0, h\neq 0} \frac{\exp(hA) - 1}{h} = A + \lim_{h\to 0} \sum_{k=2}^{\infty} \frac{A^k h^{k-1}}{k!}$$
und
$$\| \sum_{k=2}^{\infty} \frac{A^k h^{k-1}}{k!} \| \leq \sum_{k=2}^{\infty} \frac{\|A\|^k \cdot h^{k-1}}{k!}$$
konvergiert für $h \to 0$ gegen 0, da die rechte Seite eine Potenzreihe in h mit Konvergenzradius ∞ darstellt und somit stetig in $h = 0$ ist. $\square$

Diese Tatsache erlaubt es uns, Systeme homogener Differentialgleichungen der Form $y' = Ay$ mit Anfangsbedinung $y(0) = c_0 \in \mathbb{C}^n$ bequem zu lösen.

Proposition 2.18. *Die (nach der Theorie der gewöhnlichen Differentialgleichungen eindeutig bestimmte) Lösung $\phi : \mathbb{R} \to \mathbb{C}^n$ dieser Differentialgleichung ist gegeben durch*
$$\phi(t) = \exp(tA) \cdot c_0\,.$$

Beweis. Wir haben $\phi'(t) = A\cdot\exp(tA)\cdot c_0$ und $\phi(0) = \exp(0\cdot A)\cdot c_0 = c_0$. $\square$

Wir erhalten also eine ganz ähnliche Lösungsformel wie wir sie aus der Analysis einer Veränderlichen kennen.

Natürlich stellt sich nun die Frage, wie wir $\exp(tA)$ berechnen können. Dies geht besonders leicht, wenn A in Jordanscher Normalform vorliegt.

Proposition 2.19. *Es gilt*

$$\exp(t \cdot J(\lambda, n)) = \exp(t\lambda) \cdot \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{n-1}}{(n-1)!)} \\ 0 & 1 & t & \dots & \frac{t^{n-2}}{(n-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}$$

Der Beweis erfolgt durch eine direkte Rechnung (siehe Tutorium).

Da nun

$$\exp\left(t \cdot \begin{pmatrix} J(\lambda_1, m_1) & 0 & 0 & 0 \\ 0 & J(\lambda_2, m_2) & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & J(\lambda_q, m_q) \end{pmatrix}\right)$$

durch die Matrix

$$\begin{pmatrix} \exp(tJ(\lambda_1, m_1)) & 0 & 0 & 0 \\ 0 & \exp(tJ(\lambda_2, m_2)) & 0 & 0 \\ & & \ddots & \\ 0 & 0 & 0 & \exp(tJ(\lambda_q, m_q)) \end{pmatrix}$$

gegeben ist, lässt sich damit die Exponentialfunktion auf Matrizen in Jordanscher Normalform leicht auswerten.

Damit ergibt sich eine neue Perspektive auf die ab Seite 9 unten diskutierte Lösungstheorie von Systemen $y' = Ay$ linearer Differentialgleichungen mit Anfangsbedingung $y(0) = c_0$: Durch geeigneten Basiswechsel bringe man A zunächst auf Jordansche Normalform, d.h. man bestimme $S \in \mathrm{GL}(n; \mathbb{C})$ so, dass $J := SAS^{-1}$ in Jordanscher Normalform vorliegt. Man bestimme nun die Lösung ψ des Differentialgleichungssystems $y' = Jy$ mit Anfangswert Sc_0 nach der in Proposition 2.18 gegebenen Formel. Dann löst $\phi := S^{-1} \cdot \psi$ das ursprüngliche Differentialgleichungssystem mit Anfangsbedingung $\phi(0) = c_0$ (vgl. Proposition 1.13). Wir haben also die Lösung derartiger Differentialgleichungssysteme vollständig auf ein Problem der linearen Algebra zurückgeführt.

8.6.10

3. Euklidische und unitäre Vektorräume

Neben der Gegenseiten Lage von Graden, Ebenen usw. spielen Längen- und Winkelmessungen in der Geometrie eine fundamentale Rolle. In diesem Abschnitt diskutieren wir die dazu relevanten Strukturen.

Bevor wir die formale Diskussion beginnen, erinnern wir an die Definition des Skalarproduktes im $\mathbb{R}^n$ (vgl. auch Abschnitt 5.1. in [Fischer]): Für $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ definieren wir das *Skalarprodukt* von x und y als

$$\langle x, y \rangle := x_1 y_1 + \dots + x_n y_n \, .$$

Dies erlaubt die Definition der *Länge* des Vektors x durch

$$\|x\| := \sqrt{\langle x, x\rangle}$$

sowie des *Winkels* ϕ zwischen den Vektoren x und y (falls $x \neq 0 \neq y$) durch die Gleichung

$$\cos\phi = \frac{\langle x, y\rangle}{\|x\| \cdot \|y\|}$$

und die Bedingung $\phi \in [0, \pi]$. Die Definition der Länge von x ist dabei durch den Satz des Pythagoras motiviert. Bei der Definition des Winkels benutzen wir die *Cauchy-Schwarzsche Ungleichung* (die wir später beweisen)

$$\langle x, y\rangle\| \leq \|x\| \cdot \|y\|\,.$$

Die obige Definition des Winkels stimmt mit der aus der Analysis bekannten Definition über die Bogenlänge überein: Für $x = (\cos\alpha, \sin\alpha) \in \mathbb{R}^2$ und $y = (\cos\beta, \sin\beta) \in \mathbb{R}^2$ mit $0 \leq \alpha \leq \beta \leq \pi$ impliziert die obige Gleichung, dass

$$\cos\phi = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\beta - \alpha)$$

und somit $\phi = \beta - \alpha$. Dies ist nach Betrachtungen aus der Analysis genau die Bogenlänge des Kreisbogens, der x mit y gegen den Uhrzeigersinn verbindet.

Wir entwickeln nun den formalen Rahmen für diese Überlegungen.

Definition. *Es sei K ein Körper und V ein K-Vektorraum. Eine* Bilinearform *auf V ist eine Abbildung*

$$\gamma : V \times V \to K\,,$$

so dass für alle $v \in V$ die Abbildung

$$\gamma(v, -) : V \to K\,, w \mapsto \gamma(v, w)$$

und für alle $w \in V$ die Abbildung

$$\gamma(-, w) : V \to K\,, v \mapsto \gamma(v, w)$$

K-linear sind. Die Abbildung γ ist also linear in jedem Argument.

Ist V ein K-Vektorraum mit einer Bilinearform γ und ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V, so definieren wir die *darstellende Matrix* $M_{\mathcal{B}}(\gamma) \in K^{n\times n}$ von γ bezüglich $\mathcal{B}$ durch

$$M_{\mathcal{B}}(\gamma)_{ij} := \gamma(b_i, b_j)\,.$$

Diese legt die Form γ eindeutig fest. Dies ist ein Korollar des folgenden Resultates.

Proposition 3.1. *Es sei $A \in K^{n\times n}$ die darstellende Matrix einer Bilinearform γ auf einem n-dimensionalen K-Vektorraum V bezüglich einer Basis $\mathcal{B} = (b_1, \ldots, b_n)$. Sind $v, w \in V$ mit Koordinatenvektoren*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\,,\; y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

so gilt

$$\gamma(v,w) = x^T A y\,.$$

Beweis. Da die linke Seite linear in v und w und die rechte Seite linear in x und y sind und außerdem x linear von v und y linear von w abhängt, genügt es den Fall $v = b_i$, $x = e_i$, $w = b_j$, $y = e_j$ zu behandeln, wobei $(e_1, \ldots, e_n)$ die kanonische Basis des K^n bezeichnet. Aber für diesen Fall folgt die Gleichung direkt aus der Definition von A. □

Umgekehrt definiert für jede Matrix $A \in K^{n\times n}$ die Abbildung

$$(v,w) \mapsto x^T A y$$

eine Bilinearform auf V, wobei x und y die Koordinatenvektoren (als Spaltenvektoren) von v und w bezüglich $\mathcal{B}$ sind.

Korollar 3.2. *Sind γ und γ' zwei Bilinearformen mit $M_{\mathcal{B}}(\gamma) = M_{\mathcal{B}}(\gamma')$, so gilt $\gamma = \gamma'$.*

Wir erkennen an dieser Rechnung auch, wie sich die darstellende Matrix einer Bilinearform bei Basiswechsel transformiert:

Proposition 3.3. *Es sei $\mathcal{C}$ eine weitere Basis von V und $T^{\mathcal{B}}_{\mathcal{C}}$ die Matrix der Koordinatentransformation von der Basis $\mathcal{B}$ in die Basis $\mathcal{C}$. Dann gilt*

$$M_{\mathcal{B}}(\gamma) = (T^{\mathcal{B}}_{\mathcal{C}})^T M_{\mathcal{C}}(\gamma) T^{\mathcal{B}}_{\mathcal{C}}\,.$$

Beweis. Wir setzen der Einfachheit halber $S := T^{\mathcal{B}}_{\mathcal{C}}$. Die Koordinaten von v und w bezüglich $\mathcal{C}$ sind durch Sx und Sy gegeben. Somit haben wir

$$x^T M_{\mathcal{B}}(\gamma) y = (Sx)^T M_{\mathcal{C}}(\gamma') S y = x^T (S^T M_{\mathcal{C}}(\gamma') S) y$$

für alle $x, y \in K^n$. Daraus folgt die Behauptung (hier beachte man, dass für eine Matrix $A \in K^{n\times n}$ die ij-te Komponente von A durch $(e_i)^T A e_j$ gegeben ist). □

10.6.10

An dieser Stelle ist ein Vergleich mit dem Verhalten von darstellenden Matrizen bei einem Basiswechsel angebracht. Es sei also $f : V \to V$ ein Endomorphismus und $M_{\mathcal{B}}(f) : K^n \to K^n$ die darstellende Matrix von f bezüglich einer Basis $\mathcal{B}$ von V. Ist dann $\mathcal{C}$ eine weitere Basis von V und $S = T^{\mathcal{B}}_{\mathcal{C}}$ die Matrix der Koordinatentransformation von der Basis $\mathcal{B}$ in die Basis $\mathcal{C}$, so haben wir

$$M_{\mathcal{B}}(f) = S^{-1} M_{\mathcal{C}}(f) S\,,$$

d.h. die Matrix S^{-1} ist durch S^T zu ersetzen, wenn wir die darstellenden Matrizen von Bilinearformen bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$ vergleichen.

Dieser Sachverhalt lässt sich wie folgt erklären. Es sei $\gamma : V \times V \to K$ eine Bilinearform. Wir erhalten daraus eine lineare Abbildung

$$\Phi : V \to V^*\,,\; w \mapsto \gamma(-, w)$$

von V in den Dualraum $V^* = \mathrm{Hom}(V, K)$. Es sei nun wieder $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V (wir nehmen an, dass V endlichdimensional ist). Wir bezeichnen die duale Basis von V^* mit $\mathcal{B}^*$. Es gilt dann die Gleichung

$$M_{\mathcal{B}}(\gamma) = M^{\mathcal{B}}_{\mathcal{B}^*}(\Phi)$$

d.h. die darstellende Matrix der Bilinearform γ bezüglich der Basis $\mathcal{B}$ ist genau die darstellende Matrix der linearen Abbildung Φ bezüglich der Basis $\mathcal{B}$ und der dazu dualen Basis $\mathcal{B}^*$ von V^*. Dies folgt aus der Gleichung

$$(M^{\mathcal{B}}_{\mathcal{B}^*}(\Phi))_{ij} = \Phi(b_j)(b_i) = \gamma(b_i, b_j)\,.$$

Wählen wir nun eine weitere Basis $\mathcal{C}$ von V und setzen wir $S := T^{\mathcal{B}}_{\mathcal{C}}$, so haben wir

$$M^{\mathcal{B}}_{\mathcal{B}^*}(\Phi) = T^{\mathcal{C}^*}_{\mathcal{B}^*} M^{\mathcal{C}}_{\mathcal{C}^*}(\Phi) T^{\mathcal{B}}_{\mathcal{C}}$$

Beachten wir nun

$$T^{\mathcal{C}^*}_{\mathcal{B}^*} = ((T^{\mathcal{C}}_{\mathcal{B}})^T)^{-1} = (T^{\mathcal{B}}_{\mathcal{C}})^T\,,$$

(vgl. Aufgabe 3 auf Blatt 10 aus LinAlg 1), so erhalten wir genau die obige Formel für den Übergang von $M_{\mathcal{B}}(\gamma)$ nach $M_{\mathcal{C}}(\gamma)$.

Definition. *Wir nennen eine Bilinearform $\gamma : V \times V \to K$* symmetrisch*, falls $\gamma(v, w) = \gamma(w, v)$ für alle $v, w \in V$.*

Ist V endlichdimensional und $\mathcal{B}$ eine Basis von V, so ist nach Proposition 3.1 γ genau dann symmetrisch, wenn die darstellende Matrix $M_{\mathcal{B}}(\gamma)$ symmetrisch ist, d.h. $M_{\mathcal{B}}(\gamma) = (M_{\mathcal{B}}(\gamma))^T$.

Definition. *Es sei nun $K = \mathbb{R}$. Wir nennen eine symmetrische Bilinearform $\gamma : V \to V$* positiv definit*, falls $\gamma(v, v) > 0$ für alle $v \in V$ mit $v \neq 0$. Eine symmetrische, positiv definite Bilinearform auf einem reellen Vektorraum V heißt* (euklidisches) Skalarprodukt*. Skalarprodukte werden in der Regel mit spitzen Klammern bezeichnet, d.h. wir schreiben $\langle v, w\rangle$ statt $\gamma(v, w)$. Ein Paar $(V, \langle -, -\rangle)$ bestehend aus einem reellen Vektorraum und einem Skalarpdodukt heißt* euklidischer Vektorraum*.*

Das Standardbeispiel eines euklidischen Vektorraumes ist der reelle Vektorraum $\mathbb{R}^n$ zusammen mit dem (bereits oben erwähnten) *kanonischen Skalarprodukt* gegeben durch

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle := \sum_{i=1}^{n} x_i y_i\,.$$

Bevor wir die Euklidischen Vektorräume genauer untersuchen und auch einige Beispiele diskutieren, betrachten wir noch die Entsprechung für komplexe Vektorräume. Die grundlegende Beobachtung ist, dass der Absolutbetrag einer komplexen Zahl $x \in \mathbb{C}$ nicht durch $\sqrt{x^2}$, sondern durch $\sqrt{x \cdot \overline{x}}$ gegeben ist. Entsprechend definiert man das kanonische Skalarprodukt auf $\mathbb{C}^n$ durch

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n)\rangle = x_1\overline{y_1} + \ldots + x_n\overline{y_n}\,.$$

Dann ist nach wie vor $\langle x, x\rangle \geq 0$ und die übliche Länge von x ist gegeben durch

$$\|x\| = \sqrt{\langle x, x\rangle}\,.$$

Dies motiviert die folgende Definition

Definition. *Es sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung*

$$\gamma : V \times V \to \mathbb{C}$$

heißt Sesquilinearform, *falls γ linear im ersten Argument, jedoch konjugiert-linear im zweiten Argument ist, d.h. für alle $v \in V$ und $w_1, w_2 \in V$, $\lambda_1, \lambda_2 \in \mathbb{C}$ gilt*

$$\gamma(v, \lambda_1 w_1 + \lambda_2 w_2) = \overline{\lambda_1}\gamma(v, w_1) + \overline{\lambda_2}\gamma(v, w_2)\,.$$

Eine Sesquilinearform γ heißt Hermitesch, *falls*

$$\gamma(v, w) = \overline{\gamma(w, v)}$$

für alle $v, w \in V$ (dies ersetzt die Eigenschaft der Symmetrie auf reellen Vektorräumen). Weiter nennen wir eine Hermitesche Sesquilinearform γ ein (unitäres) Skalarprodukt *auf V, falls γ zusätzlich positiv definit ist, d.h.*

$$\gamma(v, v) > 0$$

für alle $v \in V$, $v \neq 0$. Hier beachte man, dass $\gamma(v, v) \in \mathbb{R}$ für alle $v \in V$, denn

$$\gamma(v, v) = \overline{\gamma(v, v)}\,.$$

Ist $\gamma : V \times V \to \mathbb{C}$ eine Sesquilinearform auf einem komplexen Vektorraum V und ist $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V, so definieren wir wie früher bei Bilinearformen die *darstellende Matrix* durch

$$M_{\mathcal{B}}(\gamma)_{ij} := \gamma(b_i, b_j)\,.$$

Sind $v, w \in V$ mit Koordinatenvektoren x und y (geschrieben als Spaltenvektoren im $\mathbb{C}^n$), so haben wir nun auf Grund der konjugierten Linearität im zweiten Eintrag

$$\gamma(v, w) = x^T M_{\mathcal{B}}(\gamma)\overline{y}\,,$$

wobei rechts alle Komponenten von y konjugiert werden. Die Form γ ist insbesondere genau dann Hermitesch, falls

$$M_{\mathcal{B}}(\gamma)^T = \overline{M_{\mathcal{B}}(\gamma)}\,.$$

Es sei nun $(V, \langle -, -\rangle)$ ein euklidischer oder unitäter Vektorraum. Für v in V setzen wir

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

Ist $(V, \langle -, -\rangle)$ ein euklidischer Vektorraum und sind $v, w \in V$ mit $v \neq 0 \neq w$, so definieren wir wie oben den *Winkel* ϕ zwischen v und w durch die Gleichung

$$\cos\phi = \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}$$

und die Bedingung $\phi \in [0, \pi]$. Diese Definition kann nicht direkt auf den unitären Fall übertragen werden (da wir es dann auch mit imaginären Winkeln zu tun bekämen).

Gleichwohl geben wir im euklidischen und unitären Fall die folgende Definition:

Definition. *Es sei V ein euklidischer oder unitäter Vektorraum. Zwei Vektoren $v, w \in W$ heißen* orthogonal*, falls*

$$\langle v, w \rangle = 0\,.$$

Wir schreiben in diesem Fall $v \perp w$. Wir nennen eine Familie $(v_i)_{i \in I}$ von Vektoren in V orthogonal*, falls $v_i \perp v_j$ für alle $i \neq j$. Wir nennen diese Familie* orthonormal*, falls sie orthogonal ist und zusätzlich $\|v_i\| = 1$ für alle $i \in I$ erfüllt. Eine orthonormale Basis von V nennt man* Orthonormalbasis*.*

Es ist nun ein Leichtes, den Satz des Pythagoras zu beweisen:

Proposition 3.4. *Es sei V ein euklidischer oder unitärer Vektorraum und $v, w \in V$ mit $v \perp w$. Dann gilt*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2\,.$$

Der Beweis ist eine direkte Rechnung. Als Folgerung haben wir

Proposition 3.5 (Cauchy-Schwarzsche Ungleichung)**.** *Es sei V ein euklidischer oder unitärer Vektorraum. Für alle $v, w \in V$ gilt dann*

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

und Gleichheit tritt genau dann ein, wenn v und w linear abhängig sind.

Beweis. Falls $w = 0$, so sind beide Aussagen klar (die Familie (v, w) dann auf jeden Fall linear abhängig). Es sei nun $w \neq 0$. Wir setzen dann

$$v' := v - \frac{\langle v, w \rangle \cdot w}{\|w\|^2}\,.$$

Es gilt $v' \perp w$, denn

$$\langle v', w \rangle = \langle v, w \rangle - \frac{\langle v, w \rangle \cdot \|w\|^2}{\|w\|^2} = 0\,.$$

Damit steht auch v' senkrecht auf $\frac{\langle v, w \rangle}{\|w\|^2} \cdot w$ (da dies ein Vielfaches von w ist). Aus dem Satz des Pythagoras ergibt sich:

$$\|v'\|^2 + \|\frac{\langle v, w \rangle \cdot w}{\|w\|^2}\|^2 = \|v\|^2\,,$$

also

$$\|v'\|^2 + \frac{|\langle v, w \rangle|^2}{\|w\|^2} = \|v\|^2\,.$$

Da $\|v'\|^2 \geq 0$, folgt daraus

$$\frac{|\langle v, w \rangle|^2}{\|w\|^2} \leq \|v\|^2$$

und daraus nach Multiplikation mit $\|w\|^2$ und Wurzelziehen die Cauchy-Schwarzsche Ungleichung.

Es tritt genau dann Gleichheit ein, wenn $\|v'\|^2 = 0$, also $v' = 0$ und dies ist äquivalent zu

$$v = \frac{\langle v, w\rangle}{\|w\|^2} \cdot w\,.$$

Dies ist gleichbedeutend dazu, dass $v = \lambda w$ mit einem λ aus dem Grundkörper ($\mathbb{R}$ oder $\mathbb{C}$). Die eine Richtung dieser Implikation ist klar. Für die andere beachten wir, dass $v = \lambda w$ (und $w \neq 0$) die Gleichung

$$\lambda = \frac{\langle v, w\rangle}{\|w\|^2}$$

impliziert wie man durch Nehmen des Skalarproduktes mit w auf beiden Seiten sieht. □

Diese Ungleichung hat die folgende wichtige Konsequenz.

Korollar 3.6. *Es sei V ein euklidischer oder unitärer Vektorraum. Dann erhält man mit der Setzung*

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

eine Norm auf V.

Beweis. Die Eigenschaften

- $\|v\| = 0 \Leftrightarrow v = 0$ und
- $\|\lambda v\| = |\lambda| \cdot \|v\|$ für alle λ aus dem Grundkörper

folgen direkt aus der Definition. Für die Dreiecksungleichung seien $v, w \in V$. Wir haben dann

$$\|v + w\|^2 = \langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle\,.$$

Da

$$\langle v, w\rangle + \langle w, v\rangle = 2\mathrm{Re}\langle v, w\rangle$$

(im unitären und im euklidischen Fall) und

$$\mathrm{Re}\langle v, w\rangle \leq |\langle v, w\rangle| \leq \|v\| \cdot \|w\|$$

(durch Anwenden der Cauchy-Schwarzschen Ungleichung) erhalten wir also

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

woraus durch Wurzelziehen die Behauptung folgt. □

Da jeder normierte (reelle oder komplexe) Vektorraum durch die Setzung

$$d(v, w) := \|v - w\|$$

ein metrischer Raum wird, sind also euklidische und unitäre Vektorräume automatisch auch metrische Räume.

15.6.10

Es sei V ein euklidischer oder unitärer Vektorraum. Ist $(v_i)_{i\in I}$ eine orthogonale Familie und sind alle $v_i \neq 0$, so ist die Familie $(\frac{v_i}{\|v_i\|})_{i\in I}$ orthonormal. Zum Beweis siehe [Fischer], S. 295.

Darüberhinaus gilt für orthogonale Familien folgende wichtige Tatsache.

Proposition 3.7. *Es sei V ein euklidischer oder unitärer Vektorraum, $(v_i)_{i\in I}$ eine orthogonale Familie und $v_i \neq 0$ für alle $i \in I$. Dann ist die Familie (v_i) linear unabhängig.*

Beweis. Wir betrachten eine Linearkombination

$$\sum_{i\in I} \lambda_i v_i = 0\,.$$

Ist nun $j \in I$, so nehmen wir das Skalarprodukt von beiden Seiten mit v_j und erhalten

$$\lambda_j \langle v_j, v_j\rangle = 0\,,$$

da $v_j \perp v_i$, falls $j \neq i$. Wegen $v_j \neq 0$ folgt daraus mit der positiven Definitheit des Skalarproduktes $\lambda_j = 0$. □

Im Zusammenhang mit diesem Beweis halten wir noch die folgende Beobachtung fest: Es sei $\mathcal{B} := (v_i)$ eine Orthonormalbasis eines euklidischen oder unitären Vektorraumes V. Ist nun $v \in V$, und setzen wir

$$\lambda_i := \langle v, v_i\rangle$$

so haben wir

$$v = \sum_{i\in I} \lambda_i v_i$$

Wir können also die Koordinaten von v bezüglich der Orthonormalbasis $\mathcal{B}$ recht einfach bestimmen.

Wir erweitern das Konzept der Orthogonalität auf Untervektorräume eines euklidischen oder unitären Vektorraumes wie folgt:

Definition. *Wir nennen zwei Untervektorräume $U, W \subset V$* orthogonal, *falls $u \perp v$ für alle $u \in U$ und $w \in W$. Wir schreiben dann $U \perp W$.*

Ist $U \subset V$ ein Untervektorraum, so definieren wir sein orthogonales Komplement *durch*

$$U^{\perp} := \{v \in V \mid v \perp u \text{ für alle } u \in U\}\,.$$

Dies ist ein Untervektorraum von V.

Ein Vektor v heißt orthogonal *zu $U \subset V$, falls $U \perp \operatorname{span}(v)$.*

Offensichtlich ist $U^{\perp}$ ebenfalls ein Untervektorraum von V und es gilt $U \perp U^{\perp}$.

Wir behandeln nun die Existenz von Orthonormalbasen.

Satz 3.8. *Jeder endlichdimensionale euklidische oder unitäre Vektorraum besitzt eine Orthonormalbasis.*

Beweis. Wir wenden das *Gram-Schmidtsche Orthonormalisierungsverfahren* an und machen Induktion nach $n := \dim V$.

Falls $\dim V = 1$ wählen wir ein $v \in V$ mit $v \neq 0$. Dann ist $(\frac{v}{\|v\|})$ eine Orthonormalbasis.

Falls $\dim V = n+1$ wählen wir einen n-dimensionalen Untervektorraum $W \subset V$ und eine Orthonormalbasis $(w_1, \ldots, w_n)$ von W. Wir ergänzen diese zu einer Basis $(w_1, \ldots, w_n, v)$ von V. Wir setzen nun

$$v' := v - \sum_{i=1}^{n} \langle v, w_i \rangle w_i \,.$$

Dann gilt $v' \perp W$, denn für alle $j = 1, \ldots, n$ ist

$$\langle v', w_j \rangle = \langle v, w_j \rangle - \sum_{i=1}^{n} \langle v, w_i \rangle \cdot \langle w_i, w_j \rangle = 0$$

wobei wir ausnutzen, dass $(w_1, \ldots, w_n)$ eine orthonormale Familie ist. Außerdem haben wir $v' \notin W$, denn $v \notin W$.

Somit ist also

$$(w_1, \ldots, w_n, \frac{v'}{\|v'\|})$$

eine Orthonormalbasis von V. □

Korollar 3.9. *Es sei V ein euklidischer oder unitärer Vektorraum und $W \subset V$ ein endlichdimensionaler Untervektorraum. Ist $v \in V$, so existiert genau ein $w \in W$ mit $v - w \perp W$. Wir nennen w die* orthogonale Projektion *von v auf W.*

Beweis. Es sei $(w_1, \ldots, w_n)$ eine Orthonormalbasis von W. Wir setzen

$$w := \sum_{i=1}^{n} \langle v, w_i \rangle w_i \,.$$

Wie im vorhergenden Beweis folgt $(v - w) \perp W$.

Ist nun $w' \in W$ ein weiterer Vektor mit $(v - w') \perp W$, so haben wir

$$w - w' = \big((v - w') - (v - w)\big) \perp W \,,$$

woraus wegen $w - w' \in W$ und der positiven Definitheit des Skalarproduktes folgt, dass $w - w' = 0$. □

Wir erhalten mit Hilfe dieses Korollars eine Abbildung

$$\mathrm{pr}_W^{\perp} : V \to W \subset V$$

indem wir jedes $v \in V$ auf seine orthogonale Projektion in W schicken. Diese Abbildung ist linear und schränkt sich auf die Identitiät in W ein. Damit ist $\mathrm{pr}_W^{\perp}$ ein Projektion im Sinne von Aufgabe 3 auf Übungsblatt 8 aus LinAlg 1. Die Abbildung $\mathrm{pr}_W^{\perp}$ heißt die *orthogonale Projektion* auf den Untervektorraum W.

Korollar 3.10. *Es seien V ein euklidischer oder unitärer Vektorraum und $W \subset V$ ein endlichdimensionaler Teilraum. Dann gilt*

$$V = W \oplus W^{\perp}\,.$$

Beweis. Es gilt $W \cap W^{\perp} = 0$, da das Skalarprodukt positiv definit ist.

Es sei nun $v \in V$ beliebig. Dann ist $v = \mathrm{pr}^{\perp}_W(v) + (v - \mathrm{pr}^{\perp}_W(v))$. Wegen $v - \mathrm{pr}^{\perp}_W(v) \perp W$ gilt somit $V = W + W^{\perp}$ □

Wir betonen, dass diese Aussage im allgemeinen nicht richtig ist, falls W nicht endlichdimensional ist (vgl. Aufgabe 4 auf Übungsblatt 7).

Wir besprechen nun noch eine geometrische Anwendung unserer Überlegung auf die Messung k-dimensionaler Volumina in euklidischen Räumen. Diese wird später in der Analysis bei der Oberflächenmessung wichtig.

Es sei dazu V ein euklidischer Vektorraum und $(v_1, \ldots, v_k)$ eine endliche Familie von Vektoren in V. Wir setzen

$$\mathrm{Gram}(v_1, \ldots, v_k) := \det \begin{pmatrix} \langle v_1, v_1\rangle & \ldots & \langle v_1, v_k\rangle \\ \vdots & & \vdots \\ \langle v_k, v_1\rangle & \ldots & \langle v_k, v_k\rangle \end{pmatrix} \in \mathbb{R}\,.$$

Dies ist die *Gramsche Determinante* von $(v_1, \ldots, v_k)$.

Proposition 3.11. *Es gilt* $\mathrm{Gram}(v_1, \ldots, v_k) \geq 0$*, wobei Gleichheit genau dann eintritt, falls* $(v_1, \ldots, v_k)$ *linear abhängig sind.*

Beweis. Wir wandeln das Gramsche Orthonormalisierungsverfahren zum *Gramschen Orthogonalisierungsverfahren* ab und konstruieren für $i = 1, \ldots k$ aus $(v_1, \ldots, v_k)$ schrittweise orthogonale Familien (die nicht unbedingt aus Vektoren der Länge 1 bestehen) $(\widetilde{v}_1, \ldots, \widetilde{v}_i)$ mit $\mathrm{span}(v_1, \ldots, v_i) = \mathrm{span}(\widetilde{v}_1, \ldots, \widetilde{v}_i)$ wie folgt.

Wir setzen $\widetilde{v}_1 := v_1$.

Ist $1 \leq i < k$ und sind $\widetilde{v}_1, \ldots, \widetilde{v}_i$ schon konstruiert, so setzen wir

$$\widetilde{v}_{i+1} := v_{i+1} - \sum_{j=1,\ldots,k \text{ und } \widetilde{v}_j \neq 0}^{i} \frac{\langle v_{i+1}, \widetilde{v}_j\rangle}{\|\widetilde{v}_j\|^2} \cdot \widetilde{v}_j$$

Man beachte, dass wir nur über $\widetilde{v}_j \neq 0$ summieren, so dass keine Division durch 0 auftritt.

Dann gilt

- $(\widetilde{v}_1, \ldots, \widetilde{v}_{i+1})$ ist orthogonal. Es ist dazu nur noch zu zeigen, dass $\widetilde{v}_{i+1} \perp \widetilde{v}_m$ für alle $1 \leq m \leq i$. Dies folgt aber aus obiger Definition von $\widetilde{v}_{i+1}$ mit einer leichten Rechnung, da nach Induktionsannahme $(\widetilde{v}_1, \ldots, \widetilde{v}_i)$ orthogonal ist.
- $\mathrm{span}(v_1, \ldots, v_{i+1}) = \mathrm{span}(\widetilde{v}_1, \ldots, \widetilde{v}_{i+1})$. Dies folgt zusammen mit der Induktionsvoraussetzung ebenfalls aus der Definition von $\widetilde{v}_{i+1}$.

Schrittweise erhalten wir somit eine orthogonale Familie $(\widetilde{v}_1, \ldots, \widetilde{v}_k)$.

Diese Familie enthält genau dann keinen Nullvektor, wenn $(v_1, \ldots, v_k)$ linear unabhängig war. Denn in diesem und nur in diesem Fall ist

$$k = \dim \operatorname{span}(v_1, \ldots, v_k) = \dim \operatorname{span}(\widetilde{v}_1, \ldots, \widetilde{v}_k)\,,$$

also die orthogonale Familie $(\widetilde{v}_1, \ldots, \widetilde{v}_k)$ linear unabhängig. Dies ist aber nach Proposition 3.7 äquivalent dazu, dass sie keinen Nullvektor enthält.

Da nun für alle $i = 1, \ldots, k-1$ der Vektor $\widetilde{v}_{i+1}$ aus v_i durch Addition von Vielfachen von $\widetilde{v}_1, \ldots, \widetilde{v}_i$ hervorgeht, haben wir

$$\operatorname{Gram}(\widetilde{v}_1, \ldots, \widetilde{v}_i, v_{i+1}, v_{i+2}, \ldots, v_k) = \operatorname{Gram}(\widetilde{v}_1, \ldots, \widetilde{v}_i, \widetilde{v}_{i+1}, v_{i+2}, \ldots, v_k)$$

für alle $i = 1, \ldots, k-1$, denn die Determinante einer Matrix ändert sich nicht, wenn man Vielfache einer Zeile zu einer anderen Zeile oder Vielfache einer Spalte zu einer anderen Spalte addiert (man setzte direkt die Definition der Gramschen Determinante ein).

Insbesondere (für $i = k$) ist also

$$\operatorname{Gram}(v_1, \ldots, v_k) = \operatorname{Gram}(\widetilde{v}_1, \ldots, \widetilde{v}_k)\,.$$

Es gilt aber

$$\operatorname{Gram}(\widetilde{v}_1, \ldots, \widetilde{v}_k) = \det \begin{pmatrix} \|\widetilde{v}_1\|^2 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & \|\widetilde{v}_k\|^2 \end{pmatrix}$$

Und somit ist die Gramsche Determinante nicht-negativ und außerdem genau dann gleich 0, wenn $(\widetilde{v}_1, \ldots, \widetilde{v}_n)$ einen Nullvektor enthält, d.h. falls $(v_1, \ldots, v_n)$ linear abhängig ist. $\square$

Aufgrund dieses Ergebnisses können wir das Volumen des von der Familie $(v_1, \ldots, v_k)$ aufgespannten *Spates* (Parallelotopes)

$$\operatorname{Spat}(v_1, \ldots, v_k) := \{t_1 v_1 + \ldots + t_k v_k \mid 0 \leq t_i \leq 1 \text{ für } i = 1, \ldots, k\}$$

wie folgt definieren.

Definition. *Es sei V ein euklidischer Vektorraum und $(v_1, \ldots, v_k)$ eine endliche Familie von Vektoren in V. Wir definieren das k-dimensionale Volumen des von dieser Familie aufgespannten Spates als*

$$\operatorname{Vol}(\operatorname{Spat}(v_1, \ldots, v_k)) := \sqrt{\operatorname{Gram}(v_1, \ldots, v_k)}\,.$$

Diese Definition stimmt mit unserer intuitiven Vorstellung eines Volumens überein: In den Schritten des eben durchgeführten Orthogonalisierungsprozesses ändert sich das Volumen nach dem Cavalierischen Prinzip nicht („Parallelogramme mit gleicher Grundlinie und gleicher Höhe haben gleiches Volumen") und für eine orthogonale Familie $(v_1, \ldots, v_k)$ haben wir

$$\operatorname{Gram}(v_1, \ldots, v_k) = \|v_1\|^2 \cdot \ldots \cdot \|v_k\|^2$$

so dass sich nach dem Wurzelziehen genau das Produkt der Seitenlängen eines k-dimensionalen Quaders ergibt.

17.6.10

Wir illustrieren den eben besprochenen Orthogonalisierungsprozess an Hand der orthogonalen Polynome aus der Analysis. Wir betrachten den $\mathbb{R}$-Vektorraum $\mathbb{R}[x]$ zusammen mit dem Skalarprodukt

$$\langle f, g\rangle := \int_{-1}^{1} f(t)g(t)dt\,.$$

Wir betrachten die Familie $(1, x, x^2, \ldots)$ und wenden darauf das obige Orthogonalisierungsverfahren an. Ist $V_n := \operatorname{span}(1, \ldots, x^n)$ und betrachten wir V_{n-1} als Unterraum von V_n, so haben wir

$$\dim \operatorname{span}(1, \ldots, x^{n-1})^{\perp} = 1\,.$$

Somit ist das im Orthogonalisierungsprozess im n-ten Schitt entstehende Element $P_n \in V_n$ bis auf Multiplikation mit einem Skalar in $\mathbb{R}$ eindeutig bestimmt.

Wir behaupten, dass man (bis auf ein skalares Vielfaches) für $n = 0, 1, 2, \ldots$ als P_n das n-te *Legendre-Polynom* P_n nehmen kann, das durch

$$P_0(x) = 1\,, P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n \text{ für } n \geq 1$$

definiert ist.

Wir machen dazu Induktion nach n. Der Fall $n = 0$ ist klar. Da $P_n \neq 0$ und $P_n \in \operatorname{span}(1, x, \ldots, x^n)$ (als n-te Ableitung eines Polynoms vom Grad $2n$) bleibt noch zu zeigen, dass

$$\langle x^k, P_n\rangle = 0$$

für alle $0 \leq k < n$. Dies zeigt man mittels einer direkten Rechnung durch partielle Integration.

Die Legendre-Polynome erfüllen noch die Normierungsbedingung $P_n(1) = 1$. Durch diese Bedingung (und die Eigenschaft, im orthogonalen Komplement von V_{n-1} in V_n zu liegen) sind sie eindeutig bestimmt.

Nimmt man auf $\mathbb{R}[x]$ geeignete andere Skalarprodukte, erhält man durch Orthogonalisierung der Familie $(1, x, x^2, \ldots)$ ganz analog die Chebycheff und Hermite-Polynome.

Ein anderes wichtiges Beispiel zum Begriff der Orthonormalität führt in die Fourieranalysis.

Wir betrachten den komplexen Vektorraum

$$V := \{f : [0, 2\pi] \to \mathbb{C} \mid f \text{ stetig } \}$$

mit dem unitären Skalarprodukt

$$\langle f, g\rangle := \frac{1}{2\pi}\int_{0}^{2\pi} f(t)\overline{g(t)}dt\,.$$

Die Familie $(e_n)_{n\in\mathbb{Z}}$ mit $e_n(x) := e^{inx}$ ist orthonormal. Den von dieser Familie aufgespannten Untervektorraum $P \subset V$ nennt man den Raum der *trigonometrischen Polynome*. Die Fourieranalysis beschäftigt sich mit der

Frage, ob und wie gut beliebige Elemente in V (oder auch in allgemeineren Funktionenräumen) durch Elemente in P approximiert werden können.

Die aus der Fourieranalysis bekannte Herangehensweise können wir geometrisch wie folgt deuten.

Wir betrachten die Unterräume

$$P_n := \mathrm{span}(e_{-n}, \ldots, e_n) \subset P \, .$$

Für ein beliebiges $f \in V$, betrachten wir die orthogonale Projektion $S_n(f) := \mathrm{pr}^{\perp}_{P_n}(f) \in P_n$. Diese nennt man die n-te *Fourierapproximation* von f.

Wir können nach dem Beweis von Korollar 3.9 diese Projektion durch die Formel

$$S_n(f) = \sum_{k=-n}^{n} c_k e_k$$

mit

$$c_k = \langle f, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt$$

dem k-ten *Fourierkoeffizienten* (für $k = -n, \ldots, n$) von f berechnen.

Das obige Skalarprodukt auf V induziert eine Norm $\| - \|_2$ auf V. Falls eine Folge $(f_n)_{n \in \mathbb{N}}$ in V bezüglich dieser Norm gegen $f \in V$ konvergiert, spricht man von *Konvergenz im quadratischen Mittel*. Ein Ergebnis der Fouriertheorie besagt, dass für alle $f \in V$ die Folge $S_n(f)$ im quadratischen Mittel gegen f konvergiert.

22.6.10

Es besteht ein enger Zusammenhang zwischen Bilinearformen und Dualität, auf den wir nun eingehen.

Definition. *Es sei V ein K-Vektorraum. Eine Bilinearform $\gamma : V \times V \to K$ heißt* nicht ausgeartet*, falls die K-lineare Abbildung*

$$\Phi : V \to V^* \, , w \mapsto \gamma(-, w)$$

(vgl. Seite 32 unten) injektiv ist. Mit anderen Worten: Für alle $0 \neq w \in V$ existiert ein $v \in V$ mit $\gamma(v, w) \neq 0$.

Ein ähnlicher Begriff existiert für Sesquilinearformen auf $\mathbb{C}$-Vektorräumen V. Man beachte dabei, dass für alle $w \in V$ die Abbildung

$$v \mapsto \gamma(v, w)$$

$\mathbb{C}$-linear ist (da Sesquilinearformen nach unserer Konvention linear im ersten Argument sind) und somit ein Element in V^* definiert. Allerdings ist die in obiger Definition betrachtete Abbildung $w \mapsto \gamma(-, w)$ nicht mehr $\mathbb{C}$-linear, sondern konjugiert-linear. Diese Besonderheit lässt sich im Zusammenhang mit Sesquilinearformen nicht vermeiden.

Wir werden im folgenden auf den Fall von Sesquilinearformen auf komplexen Vektorräumen nicht immer gesondert eingehen.

Euklidische und unitäre Skalarprodukte sind immer nicht ausgeartet.

Es ist leicht zu sehen, dass eine Bilinearform $\gamma : V \times V \to K$ auf einem endlichdimensionalen Vektorraum V genau dann nicht-ausgeartet ist, wenn die darstellende Matrix $M_{\mathcal{B}}(\gamma)$ bezüglich einer (und damit bezüglich jeder) Basis von V injektiv und somit invertierbar ist. Daraus folgt:

Proposition 3.12. *Es sei V ein endlichdimensionaler K-Vektorraum und $\gamma : V \times V \to K$ eine nicht-ausgeartete Bilinearform. Dann sind die Abbildungen*

$$\Phi : V \to V^* \, , w \mapsto \gamma(-, w)$$

und

$$\Psi : V \to V^* \, , v \mapsto \gamma(v, -)$$

Isomorphismen. Mit anderen Worten: Für jede Linearform $f : V \to K$ gibt es genau ein $w \in V$ mit $f(v) = \gamma(v, w)$ für alle $v \in V$ und genau ein $v \in V$ mit $f(w) = \gamma(v, w)$ für alle $w \in V$.

Diese Tatsache wird später im Kontext der Hilberträume zum Rieszschen Darstellungssatz erweitert.

Beweis. Es sei $\mathcal{B}$ eine Basis von V und $\mathcal{B}^*$ die dazu duale Basis von V^*. Es sei $A = M_{\mathcal{B}}(\gamma)$ die darstellende Matrix von γ bezüglich $\mathcal{B}$. Diese ist nach Voraussetzung ein Isomorphismus.

Die Abbildung ϕ ist bezüglich der Basen $\mathcal{B}$ und $\mathcal{B}^*$ durch die lineare Abbildung

$$K^n \to K^n \, , \, y \mapsto Ay$$

gegeben und damit ein Isomorphismus.

Ganz analog ist die Abbildung ψ bezüglich dieser Basen durch die lineare Abbildung

$$K^n \to K^n \, , x \mapsto (x^T A)^T = A^T x$$

gegeben. Da A ein Isomorphismus ist, gilt dies auch für A^T und somit auch für ψ. $\square$

Eine wichtige nicht-ausgeartete Bilinearform auf dem $\mathbb{R}^4$ ist die *Minkowski-Form*

$$\gamma\left(\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}\right) \mapsto x_0x'_0 - x_1x'_1 - x_2x'_2 - x_3x'_3 \in \mathbb{R}$$

Diese spielt eine fundamentale Rolle in der speziellen und allgemeinen Relativitätstheorie. Zusammen mit dieser Form nennt man $\mathbb{R}^4$ die *vierdimensionale Raumzeit*. Dabei ist x_0 die Zeitkoordinate und (x_1, x_2, x_3) sind Raumkoordinaten.

Die Form γ ist jedoch kein Skalarprodukt, denn sie ist nicht positiv definit.

Die Punkte mit Abstand 0 zum Ursprung bilden den sogenannten *Lichtkegel*

$$\{x = (x_0, x_1, x_2, x_3) \mid \gamma(x, x) = 0\} = \{(x_0, x_1, x_2, x_3) \mid \|(x_1, x_2, x_3)\| = \|x_0\|\} \, .$$

Dieser beschreibt die möglichen Raum-Zeit-Koordinaten eines Lichtteilchens, das im Ursprung des Koordinatensystems startet (falls man die Lichtgeschwindigkeit gleich 1 setzt).

Wir werden später nicht-ausgeartete symmetrische Bilinearformen auf endlich-dimensionalen reellen Vektorräumen mit Hilfe des Sylvesterschen Trägheitssatzes und der Hauptachsentransformation genauer studieren.

Zurück zu unseren allgemeinen Betrachtungen. Ist V ein endlichdimensionaler K-Vektorraum und $\psi : V \to V^*$ ein Isomorphismus, so erhalten wir durch

$$\gamma(v,w) := \psi(v)(w)$$

eine nicht-ausgeartete Bilinearform auf V und die Gleichung $v \mapsto \gamma(v,-)$ liefert wieder genau die Abbildung ψ. Mit Proposition 3.12 folgt damit:

Proposition 3.13. *Die Isomorphismen $V \to V^*$ entsprechen genau den nicht-ausgearteten Bilinearformen $V \times V \to K$. Dabei ordnen wir einem Isomorphismus $\psi : V \to V^*$ die Bilinearform $(v,w) \mapsto \psi(v)(w)$ zu.*

Diese Aussage verdeutlicht nocheinmal, dass es keinen kanonischen Isomorphismus $V \to V^*$ gibt, denn dieser würde einer kanonischen nicht-ausgearteten Bilinearform $V \times V \to K$ entsprechen. Diese gibt es aber nicht.

Wir wissen, dass wir nach Wahl einer Basis $\mathcal{B} = (b_1, \ldots, b_n)$ von V einen Isomorphismus $\psi : V \to V^*$ erhalten, der durch $b_i \mapsto b_i^*$ gegeben ist. Übersetzen wir diesen in eine Bilinearform $\gamma : V \times V \to K$ wie oben, so ist diese durch

$$\gamma(b_i, b_j) = \delta_{ij}$$

gegeben. Insbesondere ist diese Bilinearform symmetrisch.

Wir behandeln diese Entsprechung jetzt konkreter im Kontext der euklidischen Vektorräume.

Es sei also $(V, \langle -,- \rangle)$ ein endlichdimensionaler euklidischer Vektorraum. Wir erhalten wie oben einen induzierten Isomorphismus

$$\psi : V \to V^*, v \mapsto \langle v, - \rangle .$$

Wir können auf diese Weise Untervektorräume von V und von V^* miteinander identifizieren. Zum Beispiel haben wir

Proposition 3.14. *Es sei $W \subset V$ ein Untervektorraum. Dann gilt*

$$\psi(W^\perp) = W^0 \, (= \{f \in V^* \mid f|_W = 0\}) .$$

D.h. in einem euklidischen Vektorraum kann der Annulator von W in V^* mit dem orthogonalen Komplement von W in V identifiziert werden.

Ist also $0 \neq v \in V$, so ist der Kern der durch diesen Vektor gegebenen Linearform auf V genau das orthogonale Komplement von v in V.

Beweis. Die Inklusion $\psi(W^\perp) \subset W^0$ folgt aus der Konstruktion von ψ. Diese Inklusion muss aber eine Gleichheit sein, denn $\dim \psi(W^\perp) = \dim W^\perp = \dim V - \dim W = \dim W^0$. $\square$

Wir diskutieren in diesem Licht nun noch einmal eine Fragestellung, die wir bereits am Ende von Kapitel 5 aus LinAlg1 betrachtet haben. Gegeben sei ein Untervektorraum $W \subset \mathbb{R}^n$. Wir suchen eine Matrix $A \in \mathbb{R}^{m\times n}$, so dass $\ker A = W$.

Mit den Begriffen dieses Kapitels lässt sich diese Aufgabe wir folgt lösen. Es sei $(v_1, \dots, v_m)$ eine Basis von $W^\perp$. Schreiben wir diese Vektoren in die Zeilen einer $m \times n$-Matrix A, so gilt

$$\ker A = W\,.$$

Die Inklusion $W \subset \ker A$ folgt direkt aus der Konstruktion von A und aus $(W^\perp)^\perp = W$. Diese Inklusion muss wieder eine Gleichheit sein, da $\dim\ker A = \dim V - \mathrm{rang(A) = dim\,V - m = dim\,W}$.

Wir vergleichen dies mit dem früheren Verfahren, das auf dem Begriff der Dualität basierte (siehe Skript zur LinAlg I, S. 60). Wir wählen dazu eine Basis $(b_1, \dots, b_k)$ von W und schreiben diese als Spalten einer Matrix $B \in K^{n\times k}$. Nach Definition des Standardskalarproduktes auf $\mathbb{R}^n$ gilt dann

$$\ker B^T = W^\perp\,.$$

Ist also $(v_1, \dots, v_m)$ eine Basis von $\ker B^T = W^\perp$ und schreiben wir diese als Zeilen einer Matrix A (wie im Wintersemester), so erhalten wir genau das gleiche Resultat wie oben.

Damit haben wir eine geometrische Veranschaulichung des auf Dualitätsbetrachtungen basierenden Verfahrens aus dem Wintersemester.

24.6.10

4. Orthogonale und unitäre Endomorphismen, die klassischen Matrixgruppen

Definition. *Es seien V und W Euklidische, bzw. unitäre Vektorräume. Eine ($\mathbb{R}$, bzw. $\mathbb{C}$-lineare Abbildung)*

$$f : V \to W$$

heißt orthogonal*, bzw.* unitär*, falls*

$$\langle f(v), f(w)\rangle_W = \langle v, w\rangle_V$$

für alle $v, w \in V$.

Orthogonale und unitäre Abbildungen sind offensichtlich längenerhaltend (insbesondere injektiv) und bilden orthogonale Familien von Vektoren wieder auf solche ab. Orthogonale Abbildungen sind zudem winkelerhaltend. Umgekehrt gilt:

Proposition 4.1. *Es sei $f : V \to W$ linear und längenerhaltend. Dann ist f orthogonal, bzw. unitär.*

Beweis. Der Beweis folgt mit Hilfe der sogenannte *Polarisierung*: In einem Euklidischen Vektorraum V gilt für alle $v, w \in V$ die Gleichung

$$\langle v, w\rangle = \frac{1}{2}(\|v+w\|^2 - \|v\|^2 - \|w\|^2)$$

und in unitären Vektorräumen V gilt für alle $v, w \in V$ die Gleichung

$$\langle v, w\rangle = \frac{1}{4}(\|v+w\|^2 - \|v-w\|^2 + i\|v+iw\|^2 - i\|v-iw\|^2)\,.$$

Im Euklidischen Fall folgt also mit Hilfe der Linearität von f

$$\begin{aligned}\langle f(v), f(w)\rangle &= \frac{1}{2}(\|f(v)+f(w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2)\\ &= \frac{1}{2}(\|f(v+w)\|^2 - \|f(v)\|^2 - \|f(w)\|^2)\\ &= \frac{1}{2}(\|v+w\|^2 - \|v\|^2 - \|w\|^2) = \langle v, w\rangle\,.\end{aligned}$$

Im unitären Fall argumentiert man analog.

□

Sind V und W Euklidische Vektorräume und ist $f : V \to W$ eine (nicht unbedingt lineare) Abbildung, die $f(0) = 0$ erfüllt und zudem abstandserhaltend ist, d.h. $\|f(v) - f(w)\| = \|v - w\|$ für alle $v, w \in V$, so ist f bereits linear und somit eine Isometrie. Ein Beweis erfolgt in den Übungen.

Definition. *Eine Matrix $A \in \mathbb{R}^{n\times m}$, bzw. in $\mathbb{C}^{n\times m}$ heißt* orthogonal, *bzw.* unitär, *falls sie bezüglich der Standardskalarprodukte eine orthogonale, bzw. unitäre Abbildung beschreibt.*

Dies ist offensichtlich gleichbedeutend damit, dass

$$(Ax)^T(Ay) = x^T A^T A y = x^T y$$

für alle $x, y \in \mathbb{R}^m$ also

$$A^T A = E_m\,,$$

d.h. die Spalten von A bilden eine orthonormale Familie von Vektoren im $\mathbb{R}^n$ (bestehend aus m Vektoren). Im unitären Fall erhält man entsprechend die Charakterisierung

$$A^T \overline{A} = E_m\,.$$

Der Zusammenhang von orthogonalen (unitären) Abbildungen zu orthogonalen Matrizen ist wie folgt.

Proposition 4.2. *Es seien V und W endlichdimenisonale Euklidische, bzw. unitäre Vektorräume. Eine Abbildung $f : V \to W$ ist genau dann orthogonal (unitär), wenn folgendes gilt: Es seien $\mathcal{B}$ und $\mathcal{C}$ Orthonormalbasen von V und W. Dann ist die darstellende Matrix $M^{\mathcal{B}}_{\mathcal{C}}(f)$ orthogonal (unitär).*

Beweis. Wir behandeln nur den euklidischen Fall. Der unitäre Fall ist analog. Es sei $\dim V = m$, $\dim W = n$.

Zunächst beobachten wir: Sind $v, w \in V$ mit Koordinatenvektoren $x, y \in \mathbb{R}^m$ bezüglich der Basis $\mathcal{B}$, so gilt

$$\langle v, w \rangle_V = \langle x, y \rangle_{\mathbb{R}^m}$$

wobei links das Skalarprodukt in V genommen wird und rechts das kanonische Skalarprodukt in $\mathbb{R}^m$. Eine entsprechende Aussage gilt für Vektoren in W mit Koordinatenvektoren in $\mathbb{R}^n$ bezüglich der Basis $\mathcal{C}$.

Aus dieser allgemeinen Tatsache folgt leicht die Behauptung. □

Wir behandeln nun den besonders wichtigen Spezialfall von orthogonalen, bzw. unitären Endomorphismen.

Proposition 4.3. *Es sei V ein endlichdimensionaler euklidischer oder unitärer Vektorraum. Es sei $f : V \to V$ ein orthogonaler (unitärer) Endomorphismus. Dann ist f ein Isomorphismus und f^{-1} ist ebenfalls orthogonal (unitär). Alle Eigenwerte haben den Betrag* 1. *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beweis. Wir wissen bereits, dass f injektiv ist. Damit ist f ein Isomorphismus, denn $\dim V < \infty$.

Dass f^{-1} ebenfalls orthogonal (unitär) ist, sieht man wie folgt: Es seien $v, w \in V$. Dann gilt

$$\langle f^{-1}(v), f^{-1}(w) \rangle = \langle f(f^{-1}(v)), f(f^{-1}(w)) \rangle = \langle v, w \rangle \,,$$

da f orthogonal (unitär) ist. Liest man diese Gleichung von rechts nach links, folgt die Behauptung.

Es sei nun λ ein Eigenwert. Wir wählen einen Eigenvektor v. Dann ist

$$\langle v, v \rangle = \langle f(v), f(v) \rangle = \lambda \overline{\lambda} \langle v, v \rangle$$

woraus nach Division durch $\langle v, v \rangle$ folgt (beachte, dass nach Definition $v \neq 0$), das $|\lambda| = 1$.

Es seien nun λ, μ Eigenwerte mit Eigenvektoren v, w. Falls $\langle v, w \rangle \neq 0$, sieht man mit einer Rechnung ähnlich wie eben, dass $\lambda \overline{\mu} = 1$, woraus wegen $|\mu| = 1$ (somit $\mu^{-1} = \overline{\mu}$)) folgt, dass $\lambda = \mu$. □

Ist $A \in \mathbb{R}^{n \times n}$ orthogonal, so gilt auf Grund der Gleichung $A^T A = E_n$ die Beziehung $A^T = A^{-1}$. Insbesondere gilt auch $AA^T = E_n$. Entsprechend ist für unitäre Matrizen $\overline{A} A^T = E_n$. Damit sind äquivalent:

- A ist orthogonal (unitär).
- Die Spalten von A bilden eine Orthonormalbasis.
- Die Zeilen von A bilden eine Orthonormalbasis.
- $A^{-1} = A^T$ (im Euklidischen Fall), $A^{-1} = \overline{A}^T$ (im unitären Fall).

Wir betrachten nun die Teilmengen

$$\mathrm{O}(n) := \{A \in \mathrm{Mat}(n, \mathbb{R}) \mid A \text{ orthogonal}\} \subset \mathbb{R}^{n \times n}$$

und
$$\mathrm{U}(n) := \{A \in \mathrm{Mat}(n, \mathbb{C}) \mid A \text{ unitär}\} \subset \mathbb{C}^{n\times n}$$
Diese Teilmengen sind Untergruppen von $\mathrm{GL}(n, \mathbb{R})$, bzw. $\mathrm{GL}(n, \mathbb{C})$ und heißen die *orthogonale*, bzw. *unitäre* Gruppe der Ordnung n.

29.6.10

Für die Diskussion der speziellen orthogonalen Gruppen brauchen wir noch den Begriff der Orientierung endlichdimensionaler reeller Vektorräume.

Definition. *Es sei V ein endlichdimensionaler reeller Vektorraum. Wir nennen zwei (wie üblich geordnete) Basen $\mathcal{B}$ und $\mathcal{C}$ von V* gleich orientiert, *falls für die Matrix der Koordinatentransformation von der Basis $\mathcal{B}$ in die Basis $\mathcal{C}$ die folgende gilt:*
$$\det T^{\mathcal{B}}_{\mathcal{C}} > 0\,.$$
Dies definiert eine Äquivalenzrelation auf der Menge der Basen von V mit genau zwei Äquivalenzklassen. Diese Äquivalenzklassen heißen Orientierungen *von V.*

Wir nennen einen bijektiven Endomorphismus $f : V \to V$ orientierungserhaltend, *falls $\det f > 0$. Ist also $\mathcal{B}$ eine Basis von V und $f : V \to V$ orientierungserhaltend, so bildet f diese Basis auf eine gleich orientierte Basis ab.*

Für den Fall $V = \mathbb{R}^n$, bestimmt die kanonische Basis $(e_1, \ldots, e_n)$ die Standardorientierung *oder* positive Orientierung. *Dazu entgegengesetzt orientierte Basen heißen* negativ orientiert.

Wir bemerken, dass es für allgemeine Vektorräume keine Standardorientierung gibt, da es auch keine ausgezeichneten Basen gibt.

Wir setzen nun
$$\mathrm{SO}(n) = \{A \in \mathrm{O}(n) \mid \det A = 1\}\,.$$
Dies ist die *spezielle orthogonale Gruppe* der Ordnung n und enthält genau die orientierungserhaltenden, orthogonalen Endomorphismen $\mathbb{R}^n \to \mathbb{R}^n$.

Analog definieren wir
$$\mathrm{SU}(n) = \{A \in \mathrm{U}(n) \mid \det A = 1\}$$
die *spezielle unitäre Gruppe* der Ordnung n. (Hier ist eine Interpretation mit Hilfe von Orientierungen leider nicht mehr möglich).

Dass diese Teilmengen wirklich Untergruppen sind, ist leicht einzusehen.

Bemerkung 4.4. *Für alle $A \in \mathrm{O}(n)$ und $A \in \mathrm{U}(n)$ gilt $|\det(A)| = 1$. (Dies folgt aus dem Determinantenmultiplikationssatz und aus der Tatsache, dass für orthogonale (unitäre) Matrizen die Gleichung $A^{-1} = A^T$ $(A^{-1} = \overline{A}^T)$ gilt.*

Wir betrachten nun diese Gruppen genauer. Die Klassifikation der Elemente $A \in \mathrm{O}(2)$ diskutieren wir wie in [Fischer], S. 305 f. Wir erhalten genau die Drehungen und Spiegelungen im $\mathbb{R}^2$.

Für die unitären Gruppen haben wir das folgende zentrale Ergebnis.

Satz 4.5. *Es sei V ein endlichdimensionaler unitärer Vektorraum und $f : V \to V$ ein unitärer Endomorphismus. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von f.*

Korollar 4.6. *Ist $A \in \mathrm{U}(n)$, so existiert ein $S \in \mathrm{U}(n)$, so dass*

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

wobei $\lambda_1, \ldots, \lambda_n$ komplexe Zahlen vom Betrag 1 sind. (Als Spalten der Matrix S nimmt man eine Orthonormalbasis aus Eigenvektoren von A.)

Da $S^{-1} = \overline{S}^T$ können wir statt der linken Seite auch schreiben

$$\overline{S}^T AS\,.$$

Wir beweisen Satz 4.5 per Induktion nach $\dim V$ wie in [Fischer], S. 307 f.

Eine Orthonormalbasis aus Eigenvektoren von f bestimmt man, indem man zu den verschiedenen Eigenräumen von f unabhängig voneinander Orthonormalbasen berechnet und diese zu einer Basis von V zusammensetzt (wir erinnern, dass nach Proposition 4.3 Eigenvektoren zu verschiedenen Eigenwerten ohnehin orthogonal zueinander sind).

1.7.10

Die Entsprechung von Satz 4.5 für orthogonale Endomorphismen ist komplizierter, da nicht immer das charakteristische Polynom in Linearfaktoren zerfällt.

Das Ergebnis lautet wie folgt:

Satz 4.7. *Es sei V ein endlichdimensionaler Euklidischer Vektorraum und $f : V \to V$ ein orthogonaler Endomorphismus. Dann existiert eine Orthonormalbasis von V, bezüglich der f durch eine Matrix der Form*

$$\begin{pmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & A_1 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & A_k \end{pmatrix}$$

dargestellt wird. Dabei sind $A_1, \ldots, A_k \in SO(2)$ Drehmatrizen der Form

$$A_i = \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

mit $\theta_i \in (0, \pi) \cup (\pi, 2\pi)$.

Für den Beweis ist die folgende Aussage wichtig (diese ersetzt gewissermaßen die Tatsache, dass nicht jeder orthogonale Endomorphismus einen reellen Eigenwert besitzt).

Lemma 4.8. *Es sei $A : \mathbb{R}^n \to \mathbb{R}^n$ ein orthogonaler Endomorphismus. Dann existiert ein Untervektorraum $W \subset \mathbb{R}^n$ der Dimension 1 oder 2 mit $f(W) = W$.*

Beweis. Falls A einen reellen Eigenwert hat, so wählen wir einen zugehörigen Eigenvektor $w \in \mathbb{R}^n$ und setzen $W := \mathbb{R} \cdot w \subset \mathbb{R}$. Wir nehmen nun an, A hat keinen reellen Eigenwert.

Wir fassen A als lineare Abbildung $\mathbb{C}^n \to \mathbb{C}^n$ auf. Diese Abbildung hat einen Eigenwert $\lambda \in \mathbb{C}$. Es sei $0 \neq z \in \mathbb{C}^n$ ein zugehöriger Eigenvektor. Wir bezeichnen mit $\overline{z} \in \mathbb{C}^n$ den Vektor, der durch komponentenweise Konjugation entsteht. Wir haben dann

$$A\overline{z} = \overline{Az} = \overline{\lambda z} = \overline{\lambda}\,\overline{z}\,,$$

wobei wir in der ersten Gleichung benutzen, dass A nur reelle Komponenten hat. (Da z und $\overline{z}$ zu verschiedenen Eigenwerten gehören, sind diese Vektoren linear unabhängig über $\mathbb{C}$.)

Wir setzen nun $z = x + iy$, wobei $x, y \in \mathbb{R}^n$ und $\lambda = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$. Dann gilt

$$\begin{aligned} Ax &= \frac{1}{2}(Az + A\overline{z}) = \mathrm{Re}(\lambda z) = \alpha x - \beta y \\ Ay &= \frac{1}{2i}(Az - A\overline{z}) = \mathrm{Im}(\lambda z) = \beta x + \alpha y \end{aligned}$$

und somit ist $W := \mathrm{span}_{\mathbb{R}}(x, y) \subset \mathbb{R}^n$ invariant unter A.

Weiterhin gilt $\dim_{\mathbb{R}}(W) = 2$, denn wegen $z \neq 0$ wäre sonst $\dim W = 1$ und somit hätte A einen reellen Eigenwert, im Widerspruch zur Annahme.

Da f ein Isomorphismus, also injektiv ist, und $f(W) \subset W$, haben wir $f(W) = W$ wie gewünscht. □

Mit diesem Lemma erfolgt der Beweis von Theorem 4.7 ganz analog zum unitären Fall (vgl. [Fischer], S. 309): Ist $\dim V = 1$, so ist die Aussage offensichtlich richtig. Es sei nun $\dim V = n$. Wir wählen ein $W \subset V$ der Dimension 1 oder 2 mit $f(W) = W$ (dieser Unterraum existiert nach dem Lemma) und setzen $V' := W^{\perp}$. Dann ist $f(V') = V'$ und wir können die Induktionsvoraussetzung auf $f|_{V'} : V' \to V'$ anwenden und eine entsprechende Orthonormalbasis finden. Nach der Klassifikation der Elemente in $\mathrm{O}(1)$ und $\mathrm{O}(2)$ hat W ebenfalls eine Orthonormalbasis der geforderten Art. Zusammen mit der Orthonormalbasis von V' erhält man eine Orthonormalbasis von V wie gewünscht.

Wir betrachten nun die Gruppe

$$\mathrm{SU}(2) = \{A \in \mathrm{Mat}(2, \mathbb{C}) \mid \overline{A}^T A = E_n, \det A = 1\}$$

etwas genauer.

Wir haben die Darstellung

$$\mathrm{SU}(2) = \{\begin{pmatrix} w & -\overline{z} \\ z & \overline{w} \end{pmatrix} \mid w, z \in \mathbb{C}, |w|^2 + |z|^2 = 1\}$$

wie man direkt nachrechnet (vgl. das Tutorium) und damit ist

$$\mathrm{SU}(2) \to \mathbb{R}^4\,, \begin{pmatrix} w & -\overline{z} \\ z & \overline{w} \end{pmatrix} \mapsto (\mathrm{Re}(w), \mathrm{Im}(w), -\mathrm{Re}(z), \mathrm{Im}(z))$$

eine injektive Abbildung mit Bild $S^3 := \{x \in \mathbb{R}^4 \mid \|x\| = 1\}$. (Das Minuszeichen ist durch den weiter unten erläuterten Zusammenhang zu den Quaternionen motiviert).

Wir fassen nun $\mathbb{C}^{2\times 2}$ als reellen Vektorraum auf und setzen

$$\mathbb{H} := \mathrm{span}_{\mathbb{R}}\,\mathrm{SU}(2) \subset \mathbb{C}^{2\times 2}\,.$$

Man kann zeigen (siehe Tutorium), dass

$$\mathbb{H} = \mathbb{R}\cdot\mathrm{SU}(2) := \{\lambda\cdot A \mid \lambda \in \mathbb{R}, A \in \mathrm{SU}(2)\}\,.$$

Dies impliziert auch, dass zum Beispiel die Matrizen

$$\eta_0 := E_2\,,\ \eta_1 := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\,,\ \eta_2 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\,,\ \eta_3 := \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\,,$$

eine Basis des reellen Vektorraumes $\mathbb{H}$ bilden: Diese Elemente sind in der Tat linear unabhängig über $\mathbb{R}$ und geeignete reelle Linearkombinationen dieser Matrizen ergeben alle Matrizen in SU(2) und somit auch in $\mathbb{H}$, weil ja $\mathbb{H} = \mathbb{R}\cdot\mathrm{SU}(2)$.

Der entsprechende $\mathbb{R}$-lineare Isomorphismus $\mathbb{R}^4 \to \mathbb{H}$, $(x_0, x_1, x_2, x_3) \mapsto \sum_{i=0}^3 x_i\eta_i$ ist gegeben durch

$$(x_0, x_1, x_2, x_3) \mapsto \begin{pmatrix} x_0 + ix_1 & x_2 + ix_3 \\ -x_2 + ix_3 & x_0 - ix_1 \end{pmatrix}$$

und die Umkehrung durch

$$\lambda\cdot\begin{pmatrix} w & -\overline{z} \\ z & \overline{w} \end{pmatrix} \mapsto \lambda\cdot(\mathrm{Re}(w), \mathrm{Im}(w), -\mathrm{Re}(z), \mathrm{Im}(z))\,,$$

wobei $\lambda \in \mathbb{R}$ und $w, z \in \mathbb{C}$ mit $|w|^2 + |z|^2 = 1$.

6.7.10

Da $\mathbb{H} = \mathbb{R}\cdot\mathrm{SU}(2) \subset \mathrm{Mat}(2,\mathbb{C})$ abgeschlossen unter der Matrixmultiplikation ist, induziert diese auf $\mathbb{H}$ eine $\mathbb{R}$-bilineare Abbildung

$$\mathbb{H}\times\mathbb{H} \to \mathbb{H}\,.$$

Diese ist jedoch nicht kommutativ (wie wir gleich sehen werden).

Der Vektorraum $\mathbb{H}$, den wir wie oben durch Wahl der Basis $(\eta_0, \eta_1, \eta_2, \eta_3)$ mit $\mathbb{R}^4$ identifizieren, erbt von $\mathbb{R}^4$ das kanonische Skalarprodukt und wird damit ein euklidischer Vektorraum. Die induzierte Norm $\|-\|$ erfüllt die Gleichung

$$\|xy\| = \|x\|\cdot\|y\|$$

für alle $x, y \in \mathbb{H}$.

Proposition 4.9. *$\mathbb{H}$ ist bezüglich der eben eingeführten Multiplikation ein (nicht-kommutativer) Ring mit $1 = \eta_0$. Jedes Element $x \in \mathbb{H} \setminus \{0\}$ besitzt ein multiplikatives Inverses. Damit ist $\mathbb{H} \setminus \{0\}$ bezüglich der Multiplikation eine Gruppe und $\mathbb{H}$ ein* Schiefköper *(d.h. ein Körper, in dem die Multiplikation nicht notwendigerweise kommutativ ist).*

Beweis. Falls $x \in \mathbb{H} \setminus \{0\}$, so ist $\frac{x}{\|x\|} \in \mathrm{SU}(2)$, somit gibt es ein $y \in \mathrm{SU}(2)$ mit $\frac{x}{\|x\|} \cdot y = E_2 = 1$. Also ist $\frac{y}{\|x\|}$ ein Links- und Rechtsinverses von x. □

Die oben eingeführten Basiselemente von $\mathbb{H}$ kürzen wir wie folgt ab:

$$I := \eta_1, J := \eta_2, K := \eta_3$$

und erhalten die gewohnten Relationen

$$I^2 = J^2 = K^2 = -1, IJ = -JI = K, JK = -KJ = I, KI = -IK = J\,.$$

(An der Stelle ist die spezielle Wahl der Basisvektoren η_1, η_2, η_3 wichtig). Jedes Quaternion hat somit eine eindeutige Darstellung als $\alpha + \beta I + \gamma J + \delta K$ mit $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ und die Multiplikation solcher Elemente genügt den gerade angegebenen Regeln (reelle Koeffizienten vertauschen mit I, J und K). Dies zeigt auch, dass die Abbildung

$$\mathbb{C} \to \mathbb{H}\,, \alpha + \beta i \mapsto \alpha + \beta I$$

ein injektiver, normerhaltender Ringhomomorphismus ist. Damit fassen wir $\mathbb{C}$ als Unterring von $\mathbb{H}$ auf.

Ist $x := \alpha + \beta I + \gamma J + \delta K \in \mathbb{H}$, so nennen wir $\mathrm{Re}(x) := \alpha$ den *Realteil* und $\mathrm{Im}(x) := \beta I + \gamma J + \delta K$ den *Imaginärteil* von x. Wie im Komplexen setzen wir

$$\overline{x} := \mathrm{Re}(x) - \mathrm{Im}(x)\,.$$

Die Norm von x berechnet sich dann nach der aus dem Komplexen bekannten Formel

$$\|x\|^2 = x\overline{x}\,.$$

Ab sofort schreiben wir in $\mathbb{H}$ statt $\|x\|$ nur noch $|x|$, so wie bei den komplexen Zahlen auch.

Unsere Definition der Quaternionen über eine Matrixgruppe wirkt ungewöhnlich, kann aber ganz analog auch für die Definition der komplexen Zahlen (aus den reellen Zahlen) herangezogen werden. Diese Analogie stellen wir nun heraus. Wir haben

$$\mathrm{SO}(2) = \{ \begin{pmatrix} w & -z \\ z & w \end{pmatrix} \mid w, z \in \mathbb{R}, w^2 + z^2 = 1\}$$

und somit können wir $\mathrm{SO}(2)$ vermöge der Abbildung $\begin{pmatrix} w & -z \\ z & w \end{pmatrix} \mapsto (w, z)$ mit $S^1 \subset \mathbb{R}^{2\times 2}$ identifizieren. Wir setzen

$$\mathbb{C} := \mathrm{span}_{\mathbb{R}}\, \mathrm{SO}(2) \subset \mathbb{R}^{2\times 2}\,.$$

Dies ist ein zweidimensionaler, reeller Vektorraum mit Basis $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ und außerdem ein Unterring von $\mathrm{Mat}(2, \mathbb{R})$ mit Einselement $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Nennen wir den zweiten Basisvektor i, so erhalten wir genau die komplexen Zahlen mit den gewohnten Rechenregeln (Multiplikation mit i entspricht ja genau einer Drehung um $\pi/2$ gegen den Uhrzeigersinn in der Ebene).

Es stellt sich nun ganz allgemein die Frage, für welche n eine $\mathbb{R}$-bilineare Abbildung („Multiplikation“)

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n\,, (v, w) \mapsto v \cdot w$$

existiert, so dass für alle $0 \neq a \in \mathbb{R}^n$ und alle $b \in \mathbb{R}^n$ die Gleichungen $xa = b$ und $ay = b$ eindeutig lösbar sind.

Es ist ein erstaunliches Resultat von J. F. Adams aus den 60er Jahren des letzten Jahrhunderts, dass dies nur im Fall $n = 1, 2, 4, 8$ möglich ist. Dieser Beweis benutzt tiefe Methoden der Topologie, wie zum Beispiel K-Theorie. Beispiele in diesen Dimensionen sind durch die reellen Zahlen, durch die komplexen Zahlen, durch die Quaternionen und durch die sogenannten Cayley-Zahlen gegeben. Die Quaternionen wurden 1843 von William Rowan Hamilton (1805-1865) entdeckt, nachdem er jahrelang vergeblich versucht hatte, auf $\mathbb{R}^3$ eine entsprechende Multiplikation zu definieren. Heute kann man leicht zeigen, dass es auf $\mathbb{R}^3$ so eine Multiplikation nicht geben kann. Genauere Informationen zu den verschiedenen Zahlsystemen (insbesondere auch zu den Quaternionen und den Cayley-Zahlen) findet man in dem schönen Buch: Ebbinghaus et al., *Zahlen*, Springer-Verlag.

Wir behandeln zum Abschluss noch die sogenannten infinitesimale Erzeuger der klassischen Matrixgruppen.

Definition. *Es sei $G := \mathrm{GL}(n, \mathbb{R})$ oder $G := \mathrm{GL}(n; \mathbb{C})$. Eine* Einparametergruppe *in G ist eine differenzierbare Abbildung*

$$\phi : \mathbb{R} \to G$$

die außerdem ein Gruppenhomomorphismus $(\mathbb{R}, +, 0) \to (G, \cdot, E_n)$ ist. Für die Differenzierbarkeit fassen wir G als offene Teilmenge von $\mathbb{R}^{n^2}$, bzw. von $\mathbb{C}^{n^2} = \mathbb{R}^{2n^2}$ auf. (D.h. die einzelnen Komponentenfunktionen $t \mapsto (\phi(t))_{ij}$, $1 \leq i, j \leq n$ sind differenzierbar).

Proposition 4.10. *Für alle $A \in \mathrm{Mat}(n, \mathbb{R})$, bzw. $A \in \mathrm{Mat}(n, \mathbb{C})$ definiert*

$$\phi_A : \mathbb{R} \to G\,, t \mapsto \exp(t \cdot A)$$

eine Einparametergruppe in $\mathrm{GL}(n, \mathbb{R})$, bzw. in $\mathrm{GL}(n, \mathbb{C})$.

Wir interessieren uns nun für die Frage, für welche A die Abbildung ϕ Werte in den bisher behandelten Matrixgruppen annimmt. Wir haben zum Beispiel:

Proposition 4.11. *Es gilt* $\mathrm{im}\, \phi_A \in \mathrm{O}(n)$ *genau dann, wenn* $A^T = -A$.

Beweis. Angenommen, $\mathrm{im}\, \phi_A \subset \mathrm{O}(n)$. Durch Ableiten der Gleichung $\phi_A(t) \cdot \phi_A(t)^T = E_n$ für alle t erhalten wir (wegen $\exp(tA)^T = \exp(tA^T)$, wie direkt aus der Definition folgt)

$$A \exp(tA) \exp(tA^T) + \exp(tA) A^T \exp(tA^T) = 0$$

und daher, da A mit $\exp(tA)$ vertauscht,

$$\exp(tA)(A + A^T) \exp(tA^T) = 0\,.$$

Da $\exp(tA)$ und $\exp(tA^T)$ invertierbar sind, folgt daraus $A + A^T = 0$.

Umgekehrt folgt wegen $AA^T = A(-A) = -AA = A^T A$ im Falle $A + A^T = 0$

$$E_n = \exp(t(A + A^T)) = \exp(tA) \exp(tA)^T$$

für alle $t \in \mathbb{R}$. □

8.7.10

Da für alle $A \in \mathbb{R}^{n\times n}$ und für alle $A \in \mathbb{C}^{n\times n}$

$$\frac{d}{dt} \phi_A(t)|_{t=0} = A\,,$$

nennt man A den *infinitesimalen Erzeuger* der Einparametergruppe ϕ_A. Daher heißt

$$\mathfrak{o}(n) := \{A \in \mathbb{R}^{n\times n} \mid A^T = -A\}\,,$$

also der (reelle) Vektorraum der schiefsymmetrischen $(n \times n)$-Matrizen, auch der *Vektorraum der infinitesimalen Erzeuger* von $\mathrm{O}(n)$.

Wegen

$$\det(\exp(tA)) = \exp(t \cdot \mathrm{spur(A)})$$

gilt außerdem $\det \phi_A(t) = 1$ für alle t genau dann, wenn $\mathrm{spur(A)} = 0$. Daher heißt

$$\mathfrak{so}(n) := \{A \in \mathbb{R}^{n\times n} \mid A^T = -A,\ \mathrm{spur\,A} = 0\}$$

der *Vektorraum der infinitesimalen Erzeuger von* $\mathrm{SO}(n)$. Dieser stimmt mit $\mathfrak{o}(n)$ überein, wie aus jeder der beiden folgenden Beobachtungen folgt:

- die Eigenschaft $A^T = -A$ impliziert, dass auf der Diagonalen von A nur Nullen stehen,
- eine stetige Familie von Drehungen, die mit der Identität startet, besteht nur aus orientierungserhaltenden Drehungen (die Determinante kann wegen der Stetigkeit nicht von $+1$ nach -1 springen).

Entsprechend erhält man für die Vektorräume der inifinitesimalen Erzeuger (diese betrachten wir immer als Vektorräume über $\mathbb{R}$) der anderen Matrixgruppen:

- Für $\mathrm{U}(n)$ haben wir $\mathfrak{u}(n) := \{A \in \mathbb{C}^{n\times n} \mid \overline{A}^T = -A\}$ ($\mathbb{R}$-Vektorraum der *schief-hermiteschen* Matrizen - diese Matrizen haben auf der Diagonalen nur rein-imaginäre Einträge, insbesondere ist $\mathfrak{u}(n)$ nur abgeschlossen unter Multiplikation mit reellen Skalaren).

- Für $\mathrm{SU}(n)$ haben wir $\mathfrak{su}(n) := \{A \in \mathfrak{u}(n) \mid \mathrm{spur\,A} = 0\}$ (Vektorraum der spurfreien schief-hermiteschen Matrizen).
- Für $\mathrm{GL}(n, \mathbb{R})$ haben wir $\mathfrak{gl}(n; \mathbb{R}) := \mathbb{R}^{n \times n}$.
- Für $\mathrm{GL}(n, \mathbb{C})$ haben wir $\mathfrak{gl}(n; \mathbb{C}) := \mathbb{C}^{n \times n}$.
- Für $\mathrm{SL}(n; \mathbb{R}) := \{A \in \mathrm{GL}(n; \mathbb{R}) \mid \det A = 1\}$ haben wir $\mathfrak{sl}(n; \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid \mathrm{spur\,A} = 0\}$
- Für $\mathrm{SL}(n; \mathbb{C})$ haben wir $\mathfrak{sl}(n; \mathbb{C}) := \{A \in \mathbb{C}^{n \times n} \mid \mathrm{spur\,A} = 0\}$

Nach Konstruktion definiert exp jeweils eine Abbildung auf dem Vektorraum der infinitesimalen Erzeuger einer Matrixgruppe in diese Matrixgruppe. Diese Abbildung ist aber in der Regel nur dann ein Gruppenhomomorphismus, wenn man sie auf eindimensionale $\mathbb{R}$-Untervektorräume einschränkt.

Auf dem Vektorraum der infinitesimalen Erzeuger einer Matrixgruppe G existiert eine weitere Struktur, die der *Lie-Klammer*, die für Matrizen A, B durch den Kommutator $AB - BA$ gegeben ist. Diese reflektiert die Nichtkommutativität der entsprechenden Matrixgruppe. Dieser Zusammenhang wird in der Theorie der Liegruppen und Liealgebren genauer studiert und soll hier nicht weiter vertieft werden.

5. Selbstadjungierte Endomorphismen, Spektralsatz, Hauptachsentransformation

Definition. *Es sei $(V, \langle -, - \rangle)$ ein euklidischer oder unitärer Vektorraum. Ein Endomorphismus $f : V \to V$ heißt* selbstadjungiert*, wenn*

$$\langle v, f(w) \rangle = \langle f(v), w \rangle$$

für alle $v, w \in V$.

Die Bedeutung dieser Eigenschaft wird klar, wenn man wieder den Zusammenhang zu den dualen Vektorräumen heranzieht: Es sei V endlichdimensional. Die Abbildung

$$\Phi : V \to V^*, w \mapsto \langle -, w \rangle$$

ist dann ein ($\mathbb{R}$-linearer) Isomorphismus (vgl. S. 42), da das Skalarprodukt auf V nicht ausgeartet ist. Andererseits induziert jeder Endomorphismus $f : V \to V$ eine duale Abbildung $f^* : V^* \to V^*$ gegeben durch $\omega \mapsto \omega \circ f$.

Proposition 5.1. *Es sei* $\dim V < \infty$. *Der Endomorphismus $f : V \to V$ ist genau dann selbstadjungiert, wenn das Diagramm*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\cong"', "\Phi"] & V \arrow[d, "\cong"', "\Phi"] \\ V^* \arrow[r, "f^*"] & V^* \end{tikzcd}$$

kommutiert, d.h. f und f^* beschreiben (nachdem man V und V^* mit Hilfe von ϕ identifiziert) genau die gleichen Abbildungen.

Beweis. Wir müssen überlegen, wann die Gleichung

$$\Phi \circ f(w) = f^* \circ \Phi(w) \in V^*$$

für alle $w \in V$ gilt. Dies ist (für alle $w \in V$) eine Gleichung von Abbildungen in V^* und diese stimmen genau dann überein, wenn man nach Einsetzen beliebiger Vektoren $v \in V$ links und rechts das gleiche Resultat erhält.

Es sei $v \in V$. Werten wir die linke Seite auf v aus, so erhalten wir

$$\langle v, f(w) \rangle$$

und die rechte Seite ausgewertet auf v ergibt

$$\langle f(v), w \rangle .$$

Daraus folgt in der Tat, dass das obige Diagramm genau dann kommutiert, wenn f selbstadjungiert ist. □

Selbstadjungierte Abbildungen spielen eine zentrale Rolle in vielen Anwendungen in der Mathematik und Physik.

Proposition 5.2. *Es sei V ein endlichdimensionaler euklidischer oder unitärer Vektorraum und $f \in \mathrm{End}(V)$. Dann ist f genau dann selbstadjungiert, wenn folgendes gilt: Es sei $\mathcal{B}$ eine Orthonormalbasis von V und $A = M_{\mathcal{B}}(f)$ die darstellende Matrix von f bezüglich $\mathcal{B}$. Dann ist*

$$\overline{A}^T = A$$

d.h. A ist hermitesch*. Im Euklidischen Fall bedeutet das $A^T = A$, d.h. A ist* symmetrisch*.*

Zum Beweis siehe [Fischer], S. 312. Enscheidend ist die einfache Darstellung von Euklidischen oder unitären Skalarprodukten bezüglich Orthonormalbasen (vgl. auch den Beweis von Proposition 4.2).

Lemma 5.3. *Es sei $f : V \to V$ selbstadjungiert. Dann sind alle Eigenwerte von f reell und Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Zum Beweis siehe [Fischer], S. 312 bzw. 313 f.

Wir beweisen nun das folgende wichtige Resultat:

Satz 5.4 (Spektralsatz für selbstadjungierte Operatoren)**.** *Es sei* $\dim V < \infty$*. Ist $f : V \to V$ selbstadjungiert, dann besitzt V eine Orthonormalbasis bestehend aus Eigenvektoren von f .*

Nehmen wir hier $V = \mathbb{C}^n$, bzw. $V = \mathbb{R}^n$, so erhalten wir

Korollar 5.5. *Es sei $A \in \mathbb{C}^{n\times n}$ hermitesch, bzw. $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann existiert eine unitäre Matrix $S \in \mathrm{U}(n)$, bzw. orthogonale Matrix $S \in \mathrm{O}(n)$, so dass die Matrix $\overline{S}^T AS$ eine Diagonalmatrix ist, die auf der Diagonalen nur reelle Einträge hat.*

(Als Spalten von S nimmt man wieder die Vektoren einer Orthonormalbasis bestehend aus Eigenvektoren von A. In anderen Worten: Ist $\mathcal{B}$ so eine Orthonormalbasis und ist $\mathcal{E}$ die Standardbasis von $\mathbb{R}^n$, bzw. $\mathbb{C}^n$, so ist $S = T^{\mathcal{B}}_{\mathcal{E}}$.)

Die Besonderheiten von Satz 5.4 sind:

- Selbstadjungierte Endomorphismen sind (anders als orthogonale Endomorphismen) auch im Euklidischen Fall diagonalisierbar.
- Selbstadjungierte Endomorphismen besitzen auch im unitären Fall nur reelle Eigenwerte.

Wir beweisen Satz 5.4 wie in [Fischer], S. 313.

13.7.10

Der Name Spektralsatz rührt daher, dass man die Menge der Eigenwerte eines Endomorphismus $V \to V$, wobei V ein endlichdimensionaler reeller oder komplexer Vektorraum ist, auch häufig als *Spektrum* dieses Endomorphismus bezeichnet. Diese Bezeichnung ist durch die Physik motiviert (siehe unten).

Es sei V ein endlichdimensionaler Euklidischer oder unitärer Vektorraum und $W \subset V$ ein Untervektorraum. Die orthogonale Projektion $p^{\perp}_W : V \to W \subset V$ ist dann ein selbstadjungierter Endomorphismus. Ist umgekehrt $p : V \to V$ eine selbstadjungierte Projektion, dann existiert ein Untervektorraum $W \subset V$ mit $p = \mathrm{pr}^{\perp}_W$. Beide Aussagen werden in den Übungen gezeigt.

Es seien nun $\lambda_1, \dots, \lambda_k$ die (paarweise verschiedenen) reellen Eigenwerte eines selbstadjungierten Endomorphismus $f : V \to V$. Setzen wir $W_i := \mathrm{Eig(f;\lambda_i)}$, so haben wir eine Summenzerlegung

$$f = \sum_{i=1}^{k} \lambda_i \cdot \mathrm{pr}^{\perp}_{W_i}$$

von f als Linearkombination von selbstadjungierten (somit orthogonalen) Projektionen. Diese Zerlegung nennen man *Spektralzerlegung* von f. Diese Aussage wird in der Funktionalanalysis auf bestimmte Endomorphismen zwischen unendlichdimensionalen Vektorräumen verallgemeinert.

Wir betrachten als illustratives Beispiel den komplexen Vektorraum $V := \{f : [0, 2\pi] \to \mathbb{C} \mid f \text{ unendlich oft differenzierbar}, f(0) = f(2\pi)\}$ zusammen mit dem unitären Skalarprodukt

$$\langle f, g\rangle := \int_0^{2\pi} f(x)\overline{g(x)}dx$$

Der Operator

$$D : f \mapsto i\frac{df}{dx}$$

ist selbstadjungiert, denn

$$\langle Df, g\rangle = \int_0^{2\pi} (Df)(x)\cdot\overline{g(x)}dx = \big[if(x)\cdot\overline{g(x)}\big]_0^{2\pi} - i\int_0^{2\pi} f(x)\frac{d}{dx}\overline{g(x)}dx = \langle f, Dg\rangle\,.$$

Nach dem ersten Teil von Satz 5.4 sind daher alle Eigenwerte von D reell (dass man auch eine direkte Summenzerlegung von V in Eigenräume hat, stimmt nicht mehr ganz. Die korrekte Aussage lernt man in der Funktionalanalysis).

Die Eigenwerte von D erhält man durch Lösen der Differentialgleichung $Df = -i\lambda f$ mit der Randbedingung $f(0) = f(2\pi)$. Die Lösungen dieser Gleichung mit der gegebenen Randbedingung sind gegeben durch

$$f(x) = A \cdot \exp(-i\lambda x)$$

mit $A \in \mathbb{C}$ und $\lambda \in \mathbb{Z}$. Das Spektrum von D ist also gleich $\mathbb{Z}$.

Auch der sogenannte Schrödingeroperator in der Quantenmechanik ist ein selbstadjungierter Operator. Das Spektrum dieses Operators hat direkte physikalische Bedeutung (als mögliche Messwerte eines quantenmechanischen Systems). Die zugehörige „Spektraltheorie" ist heutzutage ein wichtiger Zweig der mathematischen Physik.

Wir wenden nun die Diagonalisierbarkeit symmetrischer reeller Matrizen auf die Klassifikation symmetrischer Bilinearformen an. Es sei V ein endlichdimensionaler K-Vektorraum und $\gamma : V \times V \to K$ eine Bilinearform. Sind $\mathcal{B}$ und $\mathcal{C}$ Basen von V und ist $S = T^{\mathcal{C}}_{\mathcal{B}}$ die Matrix der Koordinatentransformation von der Basis $\mathcal{C}$ in die Basis $\mathcal{B}$, so haben wir nach Proposition 3.3

$$M_{\mathcal{C}}(\gamma) = S^T M_{\mathcal{B}}(\gamma) S\,.$$

Da hier S^T auftritt und nicht S^{-1}, kann es durchaus vorkommen, dass $M_{\mathcal{B}}(\gamma)$ und $M_{\mathcal{C}}(\gamma)$ verschiedene Eigenwerte haben: Bezüglich der Basis $\frac{1}{2}$ von $\mathbb{R}$ ist das Standardskalarprodukt durch die Matrix $(\frac{1}{4})$ gegeben! Daher ist es nicht sinnvoll, von Eigenwerten der Bilinearform γ zu sprechen.

Trotzdem ist die Eigenwerttheorie für die Klassifikation von Bilinearformen von Nutzen, denn es gilt $S^{-1} = S^T$, falls $K = \mathbb{R}$ und falls $S \in \mathrm{O}(n)$. Für symmetrische Bilinearformen über $\mathbb{R}$ folgt zum Beispiel aus dem Spektralsatz:

Satz 5.6. *Es sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $\gamma : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann existiert eine Basis $\mathcal{B}$ von V, so dass $M_{\mathcal{B}}(\gamma)$ eine Diagonalmatrix ist.*

Beweis. Es sei zunächst $A \in \mathbb{R}^{n\times n}$ die darstellende Matrix von γ bezüglich einer beliebigen Basis $\mathcal{C}$ von V. Dann ist A symmetrisch. Es gibt also eine orthogonale Matrix $S \in \mathrm{O}(n)$, so dass $S^T A S$ eine Diagonalmatrix ist. Da $S \in \mathrm{GL}(n;\mathbb{R})$, gibt es genau eine Basis $\mathcal{B}$ von V, so dass $S = T^{\mathcal{B}}_{\mathcal{C}}$ (die Spalten von S sind die Koordinaten der Basisvektoren in $\mathcal{B}$ bezüglich der Basis $\mathcal{C}$). Nach Konstruktion ist dann $\mathcal{B}$ die gesuchte Basis. □

Falls $V = \mathbb{R}^n$, so können wir (nach dem Spektralsatz) die Basis $\mathcal{B}$ sogar als Orthonormalbasis (für das Standard-Skalarprodukt auf $\mathbb{R}^n$) wählen. Die von den Vektoren einer solchen Basis aufgespannten eindimensionalen Unterräume des $\mathbb{R}^n$ bezeichnet man in diesem Zusammenhang auch als *Hauptachsen* für die gegebene Bilinearform γ.

15.7.10

Diese Bezeichnung wird klar, wenn wir den Zusammenhang zu quadratischen Formen herausarbeiten.

Definition. *Eine* quadratische Form vom Rang *n über einem Körper K ist ein Polynom* $Q \in K[X_1, \ldots, X_n]$ *der Form*

$$Q = \sum_{1 \leq i,j \leq n} \alpha_{ij} X_i X_j$$

mit $\alpha_{ij} \in K$ *für alle* $i, j \in \{1, \ldots, n\}$*. Man sagt auch, Q ist ein* homogenes Polynom vom Grad 2*.*

Ist Q eine quadratische Form, so definiert Q eine Abbildung $\phi_Q : K^n \to K$, gegeben durch

$$(x_1, \ldots, x_n) \mapsto Q(x_1, \ldots, x_n)\,,$$

also durch Einsetzen der Körperelemente $x_1, \ldots, x_n$ in die Unbestimmten $X_1, \ldots, X_n$. Fassen wir die auftretenden Koeffizienten zu einer Matrix

$$A := (\alpha_{ij})_{1 \leq i,j \leq n} \in K^{n \times n}$$

zusammen, so schreibt sich ϕ_Q auch als

$$\phi_Q(x) = x^T A x$$

wobei

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K\,.$$

Ist $\gamma : K^n \times K^n \to \mathbb{R}$ die durch A beschriebene Bilinearform, also

$$\gamma(x, y) = x^T A y$$

für Spaltenvektoren $x, y \in K^n$, so haben wir

$$\phi_Q(x) = \gamma(x, x)\,.$$

Können wir Q, also die Matrix A und somit die Bilinearform γ aus dieser Funktion ϕ_Q zurückgewinnen?

Wir nehmen im folgenden an, dass $2 \neq 0$ in K gilt (dies ist z.B. für $K = \mathbb{R}$ oder $K = \mathbb{C}$, nicht jedoch für $K = \mathbb{Z}/2\mathbb{Z}$ erfüllt.)

Die Funktion ϕ_Q ändert sich nicht, wenn man in Q jeden Koeffizient α_{ij} durch

$$\frac{\alpha_{ij} + \alpha_{ji}}{2}$$

ersetzt. Die so erhaltende quadratische Form $Q \in K[X_1, \ldots, X_n]$ ist *symmetrisch*, d.h. die dann erhaltene Matrix $A = (\alpha_{ij}) \in K^{n \times n}$ (und die entsprechende Bilinearform) sind symmetrisch. Im diesem Falle können die Einträge von A durch die schon früher betrachtete *Polarisierungsformel*

$$\alpha_{ij} = e_i^T A e_j = \frac{1}{2}(\phi_Q(e_i + e_j) - \phi_Q(e_i) - \phi_Q(e_j))$$

zurückgewonnen werden, wobei $e_1, \ldots, e_n \in K^n$ die Standard-Basisvektoren sind.

Wir spezialisieren uns nun auf den Fall $K = \mathbb{R}$ und betrachten eine reelle symmtrische quadratische Form, gegeben durch eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$. Durch Hauptachsentransformation können wir die zugehörige quadratische Form Q in eine besonders einfache Gestalt bringen.

Wir wählen also eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^n$ bestehend aus Eigenvektoren von A. Es sei $S = T^{B}_{(e_1,\ldots,e_n)} \in \mathrm{O}(n)$ die entsprechende Transformationsmatrix, so dass

$$S^{-1}AS = S^T AS = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} .$$

Nach Proposition 3.3 gilt

$$D = M_{\mathcal{B}}(\gamma) \, ,$$

d.h. sind $v, w \in \mathbb{R}^n$ und x, y die Koordinatenvektoren bezüglich der Basis $\mathcal{B}$, so gilt

$$v^T A w = \sum_{i=1}^{n} \lambda_i x_i y_i$$

und insbesondere

$$\phi_Q(v) = \sum_{i=1}^{n} \lambda_i x_i^2 \, .$$

Durch die Hauptachsentransformation werden also die gemischten Terme in der quadratischen Form eliminiert. Den Übergang zu dem Orthonormalsystem $\mathcal{B}$ bezeichnet man als Hauptachsentransformation, denn die Lösungsmenge der Gleichung

$$\phi_Q(x) = 1$$

wird zum Beispiel im Fall, dass alle Eigenwerte λ_i positiv sind, durch ein Ellipsoid mit Hauptachsen $\mathbb{R}v_1, \ldots, \mathbb{R}v_n$ beschrieben, wobei $(v_1, \ldots, v_n) = \mathcal{B}$.

Wir diskutieren in diesem Zusammenhang das Beispiel auf S. 320 f. in [Fischer]. Das dort verwendete Verfahren kann allgemeiner dazu benutzt werden, *affine Quadriken*, d.h. Teilmenge des $\mathbb{R}^n$ der Form

$$\{x \in \mathbb{R}^n \mid x^T A x + \langle b, x\rangle + c = 0\}$$

zu klassifizieren, wobei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix ist und $b \in \mathbb{R}^n$. Der erste Schritt dieser Klassifikation besteht darin, durch eine Hauptachsentransformation A auf Diagonalgestalt zu bringen. Den linearen Term $\langle b, x\rangle$ versucht man anschließend, mittels quadratischer Ergänzung (d.h. Verschieben des Koordinatenursprungs) zu eliminieren. Anschließend kann man durch Fallunterscheidung die Klassifikation durchführen (ähnlich wie im eben betrachteten Beispiel). Dies führt auf klassische Gebilde wie Zylinder, Kegel, Sphären, Ellipsoide, Hyperboloide und so weiter. Eine Übersicht der dabei entstehenden Figuren im Raum (d.h. für $n = 3$) findet man zum Beispiel auf der Web-Seite http://www.mathematik.uni-stuttgart.de/studium/infomat/HM-Stroppel-Material/bsp-quadriken/

20.7.10

Wir wir bereits gesehen haben, hängen die Eigenwerte einer darstellenden Matrix einer reellen symmetrischen Bilinearform von der gewählten Basis des zu Grunde liegenden Vekorraumes ab. Die auftretenden Vorzeichen dieser Eigenwerte sind jedoch Invarianten der Bilinearform (d.h. von der gewählten Basis unabhängig). Auf diese Resultat gehen wir im Folgenden ein.

Satz 5.7 (Trägheitssatz von Sylvester)**.** *Es sei V ein endlichdimensionaler reeller Vektorraum und $\gamma : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Es seien $\mathcal{B}$ und $\mathcal{C}$ zwei Basen von V und $S := M_{\mathcal{B}}(\gamma)$ und $T := M_{\mathcal{C}}(\gamma)$ die entsprechenden darstellenden Matrizen. Es seien s_+ und s_- die Anzahlen der positiven, bzw. negativen Eigenwerte von S. Entsprechend definieren wir t_+ und t_-. Dann gilt*

$$s_+ = t_+\,,\ s_- = t_-\,.$$

Beweis. Wir können nach dem Spektralsatz neue Basen $\mathcal{G}$ und $\mathcal{H}$ von V finden, so dass die entsprechenden darstellenden Matrizen $X := M_{\mathcal{G}}(\gamma)$ und $Y := M_{\mathcal{H}}(\gamma)$ Diagonalmatrizen mit den gleichen Eigenwerten wie S und T sind. (Dazu diagonalisiere man die symmetrischen Matrizen S und T mittels Orthonormalbasen von $\mathbb{R}^n$.)

Es sei $V_{\mathcal{G}}^+ \subset V$ der Untervektorraum, der von allen Basisvektoren v in $\mathcal{G}$ mit $\gamma(v, v) > 0$ aufgespannt wird. Entsprechend sei $V_{\mathcal{G}}^- \subset V$ der Untervektorraum, der von allen Basisvektoren v in $\mathcal{G}$ mit $\gamma(v, v) < 0$ aufgespannt wird, den von den verbleibenden Basisvektoren v in $\mathcal{G}$ mit $\gamma(v, v) = 0$ aufgespannten Untervektorraum von V bezeichnen wir mit $V_{\mathcal{G}}^0$. Ensprechend definieren wir die Untervektorräume $V_{\mathcal{H}}^+, V_{\mathcal{H}}^-, V_{\mathcal{H}}^0$ von V.

Dann gilt

$$V_{\mathcal{G}}^0 = \{v \in V \mid \gamma(w, v) = 0 \text{ für alle } w \in V\} = V_{\mathcal{H}}^0\,,$$

wobei man den in der Mitte stehenden Untervektorraum von V (der nicht von der Wahl einer Basis von V abhängt) auch als *Ausartungsraum* von γ bezeichnet. Wir gehen auf die linke Gleichheit näher ein (die rechte zeigt man analog). Es sei $v \in V_{\mathcal{G}}^0$ und $y \in \mathbb{R}^n$ der zugehörige Koordinatenvektor bezüglich $\mathcal{G}$. Da X in Diagonalgestalt ist und alle Basisvektoren von $V_{\mathcal{G}}^0$ zum Eigenwert 0 gehören, gilt dann $x^T X y = 0$ für alle $x \in \mathbb{R}^n$. Somit ist $\gamma(w, v) = 0$ für alle $w \in V$. Sei umgekehrt $v \in V$ und $\gamma(w, v) = 0$ für alle $w \in V$. Es sei wieder $y \in \mathbb{R}^n$ der Koordinatenvektor von v bezüglich $\mathcal{G}$, also $v = \sum y_i v_i$. Angenommen, für einen Basisvektor $v_i \in \mathcal{G}$ mit $\gamma(v_i, v_i) \neq 0$ ist $y_i \neq 0$. Es sei $x \in \mathbb{R}^n$ der i-te kanonische Basisvektor. Dann gilt

$$x^T X y = \gamma(v_i, v_i) \cdot y_i \neq 0$$

und somit gibt es ein $w \in V$ (nämlich $w := v_i$) mit $\gamma(w, v) \neq 0$. Also liegt v nicht im Ausartungsraum von γ.

Da $V = V_{\mathcal{G}}^+ \oplus V_{\mathcal{G}}^- \oplus V_{\mathcal{G}}^0 = V_{\mathcal{H}}^+ \oplus V_{\mathcal{H}}^- \oplus V_{\mathcal{H}}^0$ und $s_+ = \dim V_{\mathcal{G}}^+$, $s_- = \dim V_{\mathcal{G}}^-$ und entsprechend $t_+ = \dim V_{\mathcal{H}}^+$, $t_- = \dim V_{\mathcal{H}}^-$, brauchen wir also nur noch $\dim V_{\mathcal{G}}^+ = \dim V_{\mathcal{H}}^+$ zu zeigen. Nach Definition der entsprechenden Räume gilt aber

$$V_{\mathcal{G}}^+ \cap (V_{\mathcal{H}}^- \oplus V_{\mathcal{H}}^0) = 0\,,$$

denn ist $v \in V$ im Schnitt dieser Räume, so gilt

- $\gamma(v, v) \geq 0$ und $\gamma(v, v) = 0$ genau dann, wenn $v = 0$ (wegen $v \in V_{\mathcal{G}}^+$)
- $\gamma(v, v) \leq 0$ (wegen $V \in V_{\mathcal{H}}^- \oplus V_{\mathcal{H}}^0$).

Daher gilt $\dim V_{\mathcal{G}}^+ \leq \dim V_{\mathcal{H}}^+$. Entsprechend zeigt man die umgekehrte Ungleichung. $\square$

Die Anzahl der positiven und negativen Eigenwerte, bzw. des Eigenwertes 0 einer darstellenden Matrix von γ hängt also nicht von der Wahl der Basis ab. Diese Anzahlen sind daher Invarianten der Bilinearform γ. Sie werden mit $r_+(\gamma)$, $r_-(\gamma)$ und $r_0(\gamma)$ bezeichnet. Die Zahlen $r_+(\gamma)$ und $r_-(\gamma)$ fasst man manchmal auch zur *Signatur* (r_+, r_-) von γ zusammen. Symmetrische Bilinearformen der Signatur $(1, 3)$ (manchmal auch $(+, -, -, -)$ geschrieben) auf vierdimensionalen reellen Vektorräumen werden in der Relativitätstheorie untersucht.

Korollar 5.8. *(Normalform für reelle symmetrische Bilinearformen) Es sei V ein endlichdimensionaler reeller Vektorraum und $\gamma : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform der Signatur (r_+, r_-) (folglich $r_0 = n - (r_+ + r_-)$). Dann existiert eine Basis $\mathcal{B}$ von V, so dass*

$$M_{\mathcal{B}}(\gamma) = \begin{pmatrix} E_{r_+} & & 0 \\ & E_{r_-} & \\ 0 & & \mathbf{0} \end{pmatrix}$$

wobei die fette $\mathbf{0}$ *unten rechts die Nullmatrix in* $\mathbb{R}^{r_0 \times r_0}$ *bezeichnet.*

Beweis. Wir wählen eine Basis $\mathcal{C} := (w_1, \ldots, w_n)$ von V bezüglich der die darstellende Matrix von γ in Diagonalform ist. Wir können annehmen, dass die ersten r_+ Eigenwerte positiv sind, die nächsten r_- Eigenwerte negativ und die verbleibenden r_0 Eigenwerte gleich 0. Wir definieren nun eine neue Basis $\mathcal{B} = (v_1, \ldots, v_n)$ durch

$$v_i := \frac{w_i}{\sqrt{|\gamma(v_i, v_i)|}}$$

falls $1 \leq i \leq r_+ + r_-$ und

$$v_i = w_i$$

sonst. Diese Basis hat die im Korollar behauptete Eigenschaft. $\square$

Für hermitesche Sesquilinearformen auf endlichdimensionalen unitären Vektorräumen gilt ebenfalls der Sylvestersche Trgheitssatz und der Satz über die Normalform solcher Sequilineaformen. (Die entsprechenden darstellenden Matrizen sind hermitesch und damit alle Eigenwerte reell). Die Beweise können anschließend (weitgehend) unverändert von oben übernommen werden.

6. Epilog: Das Tensorprodukt

Definition. *Es seien V, W und Z Vektorräume über K. Wir nennen eine Abbildung $f : V \times W \to Z$* K-bilinear*, falls für alle $w \in W$ die Abbildung $f(-, w) : V \to Z$ und für alle $v \in V$ die Abbildung $f(v, -) : W \to Z$ K-linear sind.*

Für bilineare Abbildungen gelten zunächst nicht so schöne Regeln wie für lineare Abbildungen: Beispielsweise ist $\mathrm{Bild}(f) \subset Z$ in der Regel kein Untervektorraum (vgl. die Beispiele auf Seite 350 in [Fischer]). Können wir aus f in möglichst kanonischer Weise eine lineare Abbildung nach Z „basteln"? Das Tensorprodukt beantwortet genau diese Frage.

Als Vorbereitung überlegen wir uns, welche Teilmengen von $V \times W$ die Rolle von „Basen" übernehmen, wenn es um bilineare Abbildungen geht.

Lemma 6.1. *Es seien $(v_i)_{i\in I}$ und $(w_j)_{j\in J}$ Basen von V und von W. Es sei $(z_{ij})_{i\in I, j\in J}$ eine beliebige Familie in Z. Dann gibt es genau eine bilineare Abbildung $f : V \times W \to Z$, so dass $f(v_i, w_j) = z_{ij}$ für alle $i \in I$ und $j \in J$.*

Beweis. Ist $v = \sum_{i\in I} \lambda_i v_i$ und $w = \sum_{j\in J} \mu_j w_j$, so setzen wir

$$f(v, w) := \sum_{i\in I, j\in J} \lambda_i \mu_j z_{ij}\,.$$

Diese Abbildung ist bilinear und sie ist die einzige bilineare Abbildung mit der geforderten Eigenschaft. $\square$

Die Idee des Tensorproduktes von V und W ist nun, die Menge der Paare $(v_i, w_j)_{i\in I, j\in J}$ zu einer Basis eines neuen Vektorraumes T zu erklären: Dann entsprechen nämlich (nach den Eigenschaften von Basen und nach dem Lemma 6.1) die bilinearen Abbildungen $V \times W \to Z$ genau den linearen Abbildungen $T \to Z$.

22.7.10

Satz 6.2. *Es gibt einen Vektorraum T über K zusammen mit einer bilinearen Abbildung $\phi : V \times W \to T$ mit der folgenden* universelle Eigenschaft*:*

Es sei $f : V \times W \to Z$ eine bilineare Abbildung. Dann existiert genau eine lineare Abbildung $F : T \to Z$, so dass das Diagramm

$$\begin{tikzcd} V \times W \arrow[d, "\phi"'] \arrow[dr, "f"] & \\ T \arrow[r, "F"] & Z \end{tikzcd}$$

kommutiert.

(Insofern ist die Abbildung $F : T \to Z$ die „beste“ Art, aus f eine lineare Abbildung nach Z zu konstruieren.)

Ist $\phi' : V \times W \to T'$ eine weitere bilineare Abbildung in einen Vektorraum T' mit dieser universellen Eigenschaft, so existiert ein kanonischer Isomorphismus $T \cong T'$, und dieser macht das Diagramm

$$\begin{tikzcd} & V \times W \arrow[dl, "\phi"'] \arrow[dr, "\phi'"] & \\ T \arrow[rr, "\cong"] & & T' \end{tikzcd}$$

kommutativ.

Beweis. Wir wählen Basen (v_i) von V und (w_j) von W und setzen

$$T := \{ \sum_{i \in I, j \in J} \lambda_{ij} \cdot (v_i, w_j) \mid \lambda_{ij} \in K, \text{ fast alle } \lambda_{ij} = 0 \} ,$$

wobei auf der rechten Seite formale Linearkombinationen der Paare (v_i, w_j) stehen. Dieser Vektorraum hat nach Lemma 6.1 die geforderte Eigenschaft.

(Eine formal saubere Definition von T geht vom K-Vektorraum der Abbildungen $I \times J \to K$ aus, siehe [Fischer], S. 353).

Ist $\phi' : V \times W \to T'$ eine weitere bilineare Abbildung mit dieser Eigenschaft, so existeren nach Voraussetzung eindeutig bestimmte lineare Abbildungen

$$\alpha : T \to T' , \ \beta : T' \to T$$

so dass die Diagramme

$$\begin{tikzcd} V \times W \arrow[d, "\phi"] \arrow[dr, "\phi'"] & \\ T \arrow[r, "\alpha"] & T' \end{tikzcd} \quad \text{und} \quad \begin{tikzcd} V \times W \arrow[d, "\phi'"] \arrow[dr, "\phi"] & \\ T' \arrow[r, "\beta"] & T \end{tikzcd}$$

kommutieren. Dann kommutieren auch die Diagramme

$$\begin{tikzcd} V \times W \arrow[d, "\phi"] \arrow[dr, "\phi"] & \\ T \arrow[r, "\beta\circ\alpha"] & T \end{tikzcd} \quad \text{und} \quad \begin{tikzcd} V \times W \arrow[d, "\phi'"] \arrow[dr, "\phi'"] & \\ T' \arrow[r, "\alpha\circ\beta"] & T' \end{tikzcd}$$

Und da diese Diagramme auch dann kommutieren, wenn wir in den unteren horizontalen Abbildungen die Identitäten einsetzen, folgt aus der Eindeutigkeitsaussage in der angenommenen universellen Eigenschaft von $\phi : V \times W \to T$ und $\phi' : V \times W \to T'$, dass $\beta \circ \alpha = \mathrm{id}_T$ und $\alpha \circ \beta = \mathrm{id}_{T'}$. Daher ist $\alpha : T \to T'$ der gesuchte Isomorphismus. $\square$

Die am Schluss dieses Beweises verwandte Schlussweise, die geschickt mit universellen Eigenschaften mathematischer Konstrukte hantiert, wird manchmal als „abstract nonsense“ bezeichnet. Dieser Begriff entstammt der Kategorientheorie, einer mathematichen Teidisziplin, in der die Konzepte von mathematischen Objekten und strukturerhaltenden Abbildungen abstrahiert werden.

Der in Satz 6.2 konstruierte Vektorraum T wird mit $V \otimes_K W$ bezeichnet und das *Tensorprodukt* von V und W genannt. Den Index K (den Grundkörper) lässt man oft weg. Es sei $\phi : V \times W \to V \otimes W$ die obige bilineare Abbildung. Sind $v \in V$ und $w \in W$, so schreiben wir statt $\phi(v, w) \in V \otimes W$ auch $v \otimes w$. Elemente in $V \otimes W$ dieser Form heißen *elementare Tensoren.*

Starten wir mit anderen Basen von V und W, so erhalten wir ein Modell T' des Tensorproduktes, das nach der letzten Aussage von Satz 6.2 kanonisch isomorph zu $T = V \otimes W$ ist. Insofern hängt die Konstruktion von $V \otimes W$ (bis auf kanonische Isomorphie) nicht von der Wahl von Basen von V und W ab .

Insbesondere gilt: Sind $(v_i)_{i\in I}$ und $(w_j)_{j\in J}$ beliebige Basen von V und W, so ist $(v_i \otimes w_j)_{(i,j)\in I\times J}$ eine Basis von $V \otimes W$. Insbesondere gilt $\dim(V \otimes W) = n \cdot m$, falls $\dim V = n$, $\dim W = m$ (im Gegensatz dazu ist $\dim(V \oplus W) = n + m$). Man beachte, dass beliebige Elemente in $V \otimes W$ nicht elementare Tensoren der Form $v \otimes w$ sind, sondern endliche Summen solcher elementaren Tensoren.

Wichtige Rechenregeln für das Tensorprodukt folgen direkt aus der Konstruktion (siehe [Fischer], S. 354). Wir behandeln die Beispiele auf Seite 355 f. in [Fischer].

Wir konstruieren nun wichtige kanonische Isomorphismen. Sind $\phi \in V^*$ und $\psi \in W^*$ Linearformen, so definiert

$$V \times W \to K\,,\ (v, w) \mapsto \phi(v) \cdot \psi(w)$$

eine Bilinearform, somit eine lineare Abbildung $V \otimes W \to K$ und damit ein Element in $(V \otimes W)^*$. Die so erhaltene Abbildung $V^* \times W^* \to (V \otimes W)^*$ ist bilinear und induziert damit eine lineare Abbildung

$$\alpha : V^* \otimes W^* \to (V \otimes W)^*$$

Ebenso ist die Abbildung

$$W \times V^* \to \mathrm{Hom}(V, W)\,,\ (w, \phi) \mapsto f \text{ mit } f(v) := \phi(v) \cdot w$$

bilinear und induziert damit eine lineare Abbildung

$$\beta : W \otimes V^* \to \mathrm{Hom}(V, W)\,.$$

Man beachte, dass die Definitionen von α und β nicht von der Wahl von Basen von V und W abhängen.

Satz 6.3. *Es seien V und W endlichdimensionale K-Vektorräume. Dann sind α und β Isomorphismen.*

Der Beweis erfolgt wie auf S. 357 in [Fischer].

Identifikation von linearen Abbildungen $V \to W$ mit Elementen in $W \otimes V^*$ ist die Grundlage für den besonders in der Physik gebräuchlichen Kalkül mit hoch- und tiefgestellten Indizes. Wir fixieren dazu Basen $(v_1, \ldots, v_n)$ von V

und $(w_1, \ldots, w_m)$ von W. Die dazu dualen Basen von W^* und V^* werden mit $(v_1^*, \ldots, v_n^*)$ und $(w_1^*, \ldots, w_m^*)$ bezeichnet.

Wir können somit Elemente $f \in \mathrm{Hom}(V, W) \cong W \otimes V^*$ nach der Basis $(w_i \otimes v_j^*)$ entwickeln, also

$$f = \sum_{i=1,\ldots,n,\, j=1,\ldots,m} \alpha_{ij} \cdot (w_i \otimes v_j^*) \,.$$

Wir führen nun die folgende Konvention für die Indizes bei den auftretenden Koeffizienten $\alpha_{ij} \in K$ ein: Indizes, die zu Vektorräumen gehören, stehen oben, Indizes, die zu Dualräumen gehören, stehen unten. Die Reihenfolge der Indizes entspricht der Reihenfolge der Vektor- (und Dualräume) in dem Tensorprodukt. Die (z.B. bei Matrizen üblichen) zusammenfassenden Klammern werden weggelassen. Wir schreiben also einfach

$$f = \alpha^i{}_j \,,$$

wobei der Ausdruck auf der rechten Seite das $(m \times n)$-Tupel der Koeffizienten in der obigen Linearentwicklung von f bezeichnet. Dies zeigt sehr schön den Zusammenhang zu Matrizen: Bezüglich der Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ wird f in der Tat durch die Matrix (α_{ij}) dargestellt.

Die gewohnte Multiplikation von Matrizen entspricht der sogenannten *Kontraktion* von Indizes. Wir betrachten dazu die bilineare Abbildung

$$V \times V^* \to k \,,\ (v, \phi) \mapsto \phi(v) \,.$$

Diese induziert eine lineare Abbildung

$$\tau : V \otimes V^* \to K \,.$$

Nach Wahl einer Basis $(v_1, \ldots, v_n)$ von V (und der dazu dualen Basis von V^*) ist diese Abbildung durch

$$\alpha^i{}_j \mapsto \sum_{i=1}^{n} \alpha^i{}_i \in K$$

gegeben. Damit ist $\tau : \mathrm{End}(V, V) \cong V \otimes V^* \to K$ eine basisunabhängige Definition der Spur $\mathrm{spur : End(V, V) \to K}$.

Führen wir die *Einsteinsche Summenkonvention* ein, nach der über Paare gleicher hoch- und tiefgesteller Indizes summiert werden muss, so schreibt sich τ einfach als

$$\alpha^i{}_j \mapsto \alpha^i{}_i \,.$$

Die Verkettung zweier linearer Abbildung $f = \alpha^i{}_j \in \mathrm{Hom}(V, W), g = \beta^k_l \in \mathrm{Hom}(W, Z)$ schreibt sich in diesem Kalkül als $g \circ f = \gamma^k_j \in Z \otimes V^*$ mit

$$\gamma^k_j = \beta^k_l \alpha^l{}_j \,.$$

Dies ist genau die Formel für die Multiplikation der beiden darstellenden Matrizen von f und g.

Diese Konstruktionen können iteriert werden: Sind K-Vektorräume $V_1, \ldots, V_k$ gegeben, so erhalten wir das Tensorprodukte $V_1 \otimes \ldots \otimes V_k$ zusammen mit einer K-multilinearen (d.h. in jeder Komponente K-linearen) Abbildung

$$\phi : V_1 \times \ldots \times V_k \to V_1 \otimes \ldots \otimes V_k$$

so dass jede K multilineare Abbildung

$$f : V_1 \times \ldots \times V_k \to Z$$

in einen K-Vektorraum Z eindeutig über eine lineare Abbildung

$$F : V_1 \otimes \ldots \otimes V_k \to Z$$

faktorisiert (d.h. $F \circ \phi = f$). Mehr Details dazu findet man in [Fischer], S 366 f.

Manchmal bezeichnet man Elemente von K als *Skalare* (oder *Tensoren nullter Stufe*), Elemente von V als *Vektoren* oder *Tensoren erster Stufe* (Elemente von V^* heißen manchmal auch *Kovektoren*), Elemente von $W \otimes V$, $V \otimes W$, $W^* \otimes V$, etc. als *Tensoren zweiter Stufe* und so weiter. Falls $\dim V < \infty$ und $\dim W < \infty$, sind Tensoren zweiter Stufe in $W \otimes V^*$ nichts Anderes als Elemente in $\mathrm{Hom}(V, W)$ und - nach der Wahl von Basen von V und W - nichts Anderes als Matrizen. Elemente eines Vektorraumes V werden (nach Wahl einer Basis) durch Tupel α^i gegeben. Insbesondere entspricht die Stufe eines Tensors einfach der Anzahl der auftretenden Indizes bei den entsprechenden Koeffiziententupeln.

Ist zum Beispiel $k = 4$ und wählen wir Basen von $V_1, \ldots, V_4$, so schreiben wir Elemente von $V_1 \otimes V_2^* \otimes V_3^* \otimes V_4$ (d.h. Tensoren vierte Stufe) als

$$\alpha^i{}_{jk}{}^l \,.$$

Wir erhalten wie oben Kontraktionsabbildungen: Falls zum Beispiel $V_1 = V_2 = V$, so haben wir die Kontraktion

$$\alpha^i{}_{jk}{}^l \mapsto \alpha^i{}_{ik}{}^l \,.$$

Obige Konvention der hoch- und tiefgestellten Indizes betrifft die **Koeffizienten** in Linearentwicklungen nach Basen (diese Koeffizienten sind Elemente des Grundköpers!). Die Vektoren in V und die Kovektoren in V^* werden umgekehrt indiziert: Familien in V notieren wir als $(v_i)_{i \in I}$, Familien in V^* als $(w^i)_{i \in I}$. Ist (v_i) eine Basis von V, so wird die dazu duale Basis von V^* mit (v^i) bezeichnet. Insbesondere gilt $v^i(v_j) = \delta_{ij}$. Diese Regel passt gut zur Einsteinschen Summenkonvention: Ist $v \in V$ und $v = \sum_{i \in I} \alpha_i v_i$ mit $\alpha_i \in K$ für alle i, so schreiben wir

$$v = \alpha^i = \alpha^i v_i \,.$$

Der hier vorgestellte Tensorkalkül ist sehr geschmeidig und hat mannigfache Anwendungen in der Differentialgeometrie und mathematischen Physik.

* * *