Lineare Algebra 2
Sommersemester 2019

Prof. Dr. Annette Huber-Klawitter

Fassung vom 20. Juli 2019

Dies ist ein Vorlesungsskript und kein Lehrbuch.
Mit Fehlern muss gerechnet werden!

Math. Institut
Ernst-Zermelo-Str. 1
79104 Freiburg

0761-203-5560
annette.huber@math.uni-freiburg.de

Inhaltsverzeichnis

Einleitung

Wintersemester

Im ersten Teil der Vorlesung ging es zunächst um Vektorräume und lineare Abbildungen. Konkret bedeutete dies das Rechnen mit Matrizen und das Lösen von linearen Gleichungssystemen. Determinanten waren ein wesentliches Hilfsmittel.

Das zweite Thema waren Endomorphismen, also lineare Selbstabbildungen $f : V \to V$. Wir haben den Begriff des Eigenwertes kennengelernt und gesehen, dass man Eigenwerte als Nullstellen des characteristischen Polynoms bestimmt.

Nebenbei haben wir mathematische Grundbegriffe (z.B. Gruppen) und Beweismethoden(vollständige Induktion, indirekter Beweis) kennengelernt.

Gauß-Algorithmus!

Sommersemester

Wir behandeln zwei große Themenblöcke:

(i) Bilineare Abbildungen und Skalarprodukte

(ii) Diagonalisierbarkeit von Endormorphismen und Normalformen

Ein Skalarprodukt weist einem Paar von Vektoren v, w in einem $\mathbb{R}$-Vektorraum V eine reelle Zahl, also einen Skalar zu. Prototyp ist das *Standardskalarprodukt* auf dem $\mathbb{R}^n$:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = x_1y_1 + x_2y_2 + \ldots x_ny_n.$$

Wir benutzen es, um Längen von Vektoren und Winkel zwischen Vektoren zu definieren. Damit wird die gesamte euklidische Geometrie zu einem Teilgebiet der linearen Algebra. Drei- und vierdimensionale euklidische Räume sind die natürliche Grundlage der Physik. Wir studieren dann vor allem diejenigen linearen Abbildungen und Endomorphismen, die auf die eine oder andere Art mit dem Skalarprodukt verträglich sind. Unser erstes großes Ziel ist der *Spektralsatz* (auch Satz zur *Hauptachsentransformation*), der für gewisse Endomorphismen die Existenz einer Basis aus Eigenvektoren garantiert.

Damit sind wir dann bei unserem zweiten Thema: Wir studieren (wieder über beliebigen Körpern) mögliche Normalformen von Matrizen zu Endomorphismen. Der schönste Fall ist, wenn V eine Basis aus Eigenvektoren hat. Dann kann die darstellende Matrix als Diagonalmatrix gewählt werden. Leider ist dies nicht immer der Fall. Wir werden dafür ein genaues Kriterium kennenlernen. Dabei knüpfen wir wieder an die Theorie des charakteristischen Polynoms aus dem Wintersemester an. Höhepunkt wird die Jordansche Normalform sein.

Zum Abschluss greifen wir noch einmal bilineare Abbildungen auf (ein erstes Beispiel war das Skalarprodukt), diesmal über beliebigen Körpern. Wir beschreiben sie durch Matrizen, ähnlich zum Fall von linearen Abbildungen. Dies führt auf den Begriff des Tensorproduktes und des äußeren Produktes. Letzteres steckt hinter Formeln aus der mehrdimensionalen Integrationstheorie und wirft ein neues Licht auf die Determinante, wie wir sie im ersten Semester kennengelernt haben.

Kein Gauß-Algorithmus!

Literatur

Das Material, das wir behandeln werden, ist in allen Standardlehrbüchern zur linearen Algebra enthalten. Die Reihenfolge variiert, auch der Grad der Allgemeinheit einiger Sätze. Mögliche Referenzen:

(i) Bosch: Lineare Algebra

(ii) Fischer: Lineare Algebra

(iii) Lang: Algebra (*Enthält viel mehr als den Stoff der linearen Algebra!*)

(iv) Lorenz: Lineare Algebra 2

Vorsicht mit Wikipedia u.ä.: Oft sehr gut und sehr hilfreich, kann aber auch lückenhaft/falsch sein.

Kapitel 10

Euklidische und unitäre Vektorräume

In diesem Kapitel schränken wir uns ein auf die Grundkörper $\mathbb{R}$ und $\mathbb{C}$. Wir schreiben abkürzen $\mathbb{K}$ für einen von beiden.
Wir erinnern uns $\mathbb{C} = \{x + iy | x, y \in \mathbb{R}\}$. Dies wird zu einem Körper mit der Rechenregel $i^2 = -1$. Wir benutzen die *komplexe Konjugation* $(x + iy) \mapsto (x - iy)$. Wir schreiben $z \mapsto \overline{z}$. Diese Abbildung ist verträglich mit Addition und Multiplikation.

Definition 10.1. *Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung*

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

heißt Bilinearform *($\mathbb{K} = \mathbb{R}$) bzw.* Sesquilinearform *($\mathbb{K} = \mathbb{C}$), wenn gilt:*

(i) $s(v, \lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 s(v, w_1) + \lambda_2 s(v, w_2)$ *für alle* $\lambda_1, \lambda_2 \in \mathbb{K}$, $v, w_1, w_2 \in V$.

(ii) $s(\mu_1 v_1 + \mu_2 v_2, w) = \overline{\mu}_1 s(v_1, w) + \overline{\mu}_2 s(v_2, w)$ *für alle* $\mu_1, \mu_2 \in \mathbb{K}$, $v_1, v_2, w \in V$.

Die Form heißt symmetrisch *bzw.* hermitesch*, falls zusätzlich gilt*

(iii) $s(v, w) = \overline{s(w, v)}$ *für alle* $v, w \in V$.

Die symmetrische/hermitsche Form heißt Skalarprodukt*, wenn sie zusätzlich* positiv definit *ist, d.h.:*

(iv) $s(v, v) > 0$ *für* $v \in V$ *mit* $v \neq 0$.

Das Paar (V, s) heißt euklidischer *bwz.* unitärer *Vektorraum. Wir schreiben meist $\langle v, w \rangle$ für $s(v, w)$.*

Bemerkung. (i) Damit Bedingung (iv) überhaupt formuliert werden kann, muss der Wert reell sein. Im unitären Fall folgt dies aus (iii):

$$s(v,v) = \overline{s(v,v)} \Rightarrow s(v,v) \in \mathbb{R}.$$

Dies ist auch der Grund, warum wir uns in diesem Körper auf reelle und komplexe Vektorräume einschränken. Den Begriff des Skalarproduktes gibt es über $\mathbb{F}_2$ nicht!

(ii) Im unitären Fall spricht man oft auch von einem *hermiteschen* Produkt.

(iii) Es gilt $s(0,0) = 0$ wegen (i) oder (ii). Oft formuliert man daher (iii) auch als: $s(v,v,) \geq 0$ mit $s(v,v) = 0$ genau für $v = 0$.

(iv) Es gibt auch die Begriffe *positiv semi-definit* ($\geq$ in (iv)), *negativ definit* ($<$ in (iv)), *negativ semi-definit* ($\leq$) und schließlich *indefinit* (keines davon, also manchmal $>$, manchmal $<$).

(v) Bedingung (ii) heißt *semi-linear* (bezüglich der komplexen Konjugation). In der Literatur wird das Skalarprodukt mal mit Semi-Linearität in der ersten, mal in der zweiten Variablen definiert. Die beiden Versionen sind natürlich äquivalent, aber die Formeln ändern sich.

Beispiel. Sei $V = \mathbb{R}^n$, $v, w \in V$. Wir setzen

$$s(v,w) = v^t w$$

oder ganz explizit:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = x_1 y_1 + x_2 y_2 + \dots x_n y_n.$$

Nach den Rechenregeln der Matrixmulitplikation ist die Abbildung linear in beiden Argumenten. An der expliziten Formel sieht man, dass die Abbildung symmetrisch ist. Im Fall $v = w$ erhalten wir eine Summe von Quadraten, die nur verschwindet, wenn alle Koordinaten verschwinden.
Im Fall $V = \mathbb{C}^n$ und $\mathbb{K} = \mathbb{C}$ setzen wir

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

oder explizit

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \overline{x}_1 y_1 + \overline{x}_2 y_2 + \dots \overline{x}_n y_n.$$

Sesquilinearität und hermitesch sind klar. Im Fall $v = w$ erhält man $\sum |x|_i^2$. Diese verschwindet genau, wenn alle Koordinaten verschwinden.

Beispiel. Sei $I = [a, b] \subset \mathbb{R}$ beschränktes, abgeschlossenes Intervall. Sei $V = \{f : I \to \mathbb{K} | f \text{ stetig}\}$. Wir definieren

$$\langle f, g \rangle = \int_a^b \overline{f(t)} g(t) \mathrm{d}t.$$

Das Riemanintegral existiert, da f, g stetig ist und I beschränkt. Die Abbildung ist offensicht bi- bzw. sesquilinear und symmetrisch bwz. hermitesch. Wir betrachten

$$\langle f, f \rangle = \int_a^b |f(t)|^2 \mathrm{d}t.$$

Der Integrand ist größer gleich 0. Wenn das Integral verschwindet, so muss nach dem Mittelwertsatz der Integralrechnung $|f(t)| = 0$ für alle $t \in I$ gelten, also $f = 0$ als Element von V.

Beispiel. Sei $V = \mathbb{R}^2$. Wir setzen

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle = 2x_1 y_1 + x_2 y_2.$$

Dies ist ebenfalls ein Skalarprodukt.

Lemma 10.2. *Sei V ein euklidischer bzw. unitärer Vektorraum und $U \subset V$ ein Untervektorraum. Dann definiert die Einschränkung von $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf U.*

Beweis: All Eigenschaften gelten für Element von U, weil sie für Elemente von V gelten. □

Definition 10.3. *Sei V euklidischer bzw. unitärer Vektorraum mit Basis $B = (v_1, \dots, v_n)$. Dann heißt*

$$M_B(s) = \left(s(v_i, v_j)_{i,j=1}^n \right)$$

darstellende Matrix *von s bezüglich B.*

Beispiel. Das Skalarprodukt $\langle (a_1, a_2)^t, (b_1, b_2)^t \rangle = 2a_1 b_1 + a_2 b_2$ hat bezüglich der Standardbasis auf $\mathbb{R}^2$ die darstellende Matrix

$$M_B(s) = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

Satz 10.4. *Das Skalarprodukt ist eindeutig bestimmt durch die darstellende Matrix. Für $v = \sum_{i=1}^n a_i v_i$ und $w = \sum_{j=1}^n b_j v_j$ gilt*

$$s(v, w) = \sum_{i,j=1}^n \overline{a}_i s(v_i, v_j) b_j$$

oder mit $a = (a_1, \dots, a_n)^t$, $b = (b_1, \dots, b_n)^t$

$$s(v, w) = \overline{a}^t M_B(s) b.$$

Die darstellende Matrix erfüllt

$$M_B(s)^t = \overline{M_B(s)}.$$

Sei $B' = (w_1, \dots, w_n)$ *eine weitere Basis von* V *mit Basiswechselmatrix* $M^B_{B'}(\mathrm{id})$. *Dann gilt*

$$M_B(s) = (\overline{M^B_{B'}(\mathrm{id})})^t M_{B'}(s) M^B_{B'}(\mathrm{id}).$$

Beweis: Wir setzen ein und benutzen die Sesequilinearität. Wir berechnen die Transponierte mittels $s(v_j, v_i) = \overline{s(v_i, v_j)}$.
Nach Definition von $M^B_{B'}(\mathrm{id}) = (t_{ij})$ gilt

$$v_j = \sum_{i=1}^{n} t_{ij} w_i.$$

Es folgt

$$s(v_a, v_b) = s(\sum_i t_{ia} w_i, \sum_j t_{jb} w_j) = \sum_{i,j} \overline{t_{ia}} s(w_i, w_j) t_{jb}.$$

Das ist genau der Eintrag von $\overline{M^B_{B'}(\mathrm{id})}^t M_B(s) M^B_{B'}(\mathrm{id})$ an der Stelle ab. □

Definition 10.5. *Sei* $M \in M_n(\mathbb{R})$ *bzw.* $M \in M_n(\mathbb{C})$. *Die Matrix heißt* symmetrisch *bzw.* hermitesch, *wenn*

$$M^t = \overline{M}.$$

Sie heißt positiv definit, *wenn*

$$s : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}, \qquad (x, y) \mapsto \overline{x}^t M y$$

positiv definit ist.

Die darstellenden Matrizen von Skalarprodukten sind also symmetrisch bzw. hermitesch und positiv definit. Es ist nicht offensichtlich, ob eine Matrix positiv definit ist, es gibt aber Kriterien.

Längen und Winkel

Skalarprodukte sind wichtig, weil sie Längen von Vektoren und Abstände von Punkten definieren.

Definition 10.6. *Sei* V *ein* $\mathbb{K}$*-Vektorraum. Eine Abbildung*

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

heißt Norm, *wenn für alle* $v, w \in V$, $\lambda \in \mathbb{K}$ *gilt* *(i)* $\|\lambda v\| = |\lambda|\|v\|$

(ii) $\|v + w\| \leq \|v\| + \|w\|$ *(Dreiecksungleichung)*

(iii) $\|v\| \geq 0$ *und* $\|v\| = 0$ *genau dann, wenn* $v = 0$.

Eine Abbildung

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

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

(i) $d(x, y) = d(y, x)$ *(Symmetrie)*

(ii) $d(x, z) \leq d(x, y) + d(y, z)$ *(Dreiecksungleichung)*

(iii) $d(x, y) \geq 0$ *und* $d(x, y) = 0$ *genau dann, wenn* $x = y$.

Die Zahl $d(x, y)$ *heißt auch* Abstand *von* x, y.

Lemma 10.7. *Sei* $\|\cdot\|$ *eine Norm auf dem* $\mathbb{K}$*-Vektorraum* V*. Dann definiert*

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

eine Metrik.

Beweis: Für alle $x, y \in V$ gilt

$$d(x, y) = \|x - y\| = \|(-1)(y - x)\| = |-1|\|y - x\| = d(y, x).$$

Wir überprüfen die Dreiecksungleichung. Für alle $x, y, z \in V$ gilt

$$d(x, z) = \|x - y\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z).$$

Und schließlich ist $d(x, y) = \|x - y\| \geq 0$. Hierbei ist $d(x, y) = 0$ genau dann, wenn $\|x - y\| = 0$, also $x - y = 0$. □

Beispiel. Nicht jede Metrik kommt von einer Norm her: z.B. die triviale Metrik

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

Sobald man eine Metrik hat, kann man den Begriff des Grenzwertes definieren und Analysis betreiben. Dies gilt z.B. für den $\mathbb{R}^n$ oder $\mathbb{C}^n$.

Lemma 10.8. *Sei* V *ein euklidischer bzw. unitärer Vektorraum. Dann ist*

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

eine Norm auf V*. Für alle* $v, w \in V$ *gilt die* Cauchy-Schwarzsche Ungleichung

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

Beweis: Da $\langle v,v\rangle$ reell und nicht-negativ ist, ist die Quadratwurzel als nicht-negative reelle Zahl wohldefiniert.
Wir rechnen die Eigenschaften nach: für alle $v \in V$ und $\lambda \in \mathbb{K}$ folgt

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

Es ist

$$\|v\|^2 = \langle v,v\rangle \geq 0$$

und $\|v\|^2 = 0$ nur für $v = 0$.
Für die Dreiecksungleichung müssen wir verifizieren:

$$\|v+w\| \leq \|v\| + \|w\|$$

oder (da beide Seiten nicht-negativ sind) äquivalent:

$$\langle v+w, v+w\rangle \leq \left(\sqrt{\langle v,v\rangle} + \sqrt{\langle w,w\rangle}\right)^2$$

bzw.

$$\langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle \leq \langle v,v\rangle + 2\sqrt{\langle v,v\rangle\langle w,w\rangle} + \langle w,w\rangle$$

Dies ist wegen $\langle w,v\rangle = \overline{\langle v,w\rangle}$ äquivalent zu:

$$2\mathrm{Re}\langle v,w\rangle \leq 2\|v\|\|w\|.$$

Für jede komplexe Zahl gilt $\mathrm{Re} z \leq |z|$, also genügt es, die Cauchy-Schwarzsche Ungleichung zu zeigen. Diese verifizieren wir nun. Für $w = 0$ ist nichts zu zeigen. Sei also $w \neq 0$. Es gilt für beliebiges, noch zu wählendes $\lambda \in \mathbb{K}$:

$$0 \leq \langle v - \lambda w, v - \lambda w\rangle = \langle v,v\rangle - \lambda\langle v,w\rangle - \overline{\lambda}\overline{\langle v,w\rangle} + \overline{\lambda}\lambda\langle w,w\rangle.$$

In dieser Ungleichung setzen wir $\lambda = \frac{\overline{\langle v,w\rangle}}{\langle w,w\rangle}$ (wohldefiniert, da $\|w\| \neq 0$) und erhalten

$$\begin{aligned} 0 \leq & \|v\|^2 - \frac{\overline{\langle v,w\rangle}}{\|w\|^2}\langle v,w\rangle - \frac{\langle v,w\rangle}{\|w\|^2}\overline{\langle v,w\rangle} + \frac{\overline{\langle v,w\rangle}\langle v,w\rangle}{\|w\|^4}\|w\|^2 \\ = & \|v\|^2 - \frac{|\langle v,w\rangle|^2}{\|w\|^2} \end{aligned}$$

Nach Multiplikation mit $\|w\|^2$ ist das die Behauptung. □

Definition 10.9. *Sei V ein euklidischer oder unitärer Vektorraum. Zwei Vektoren $x, y \in V$ heißen* orthogonal*, falls $\langle x,y\rangle = 0$. Wir schreiben $x \perp y$. Für eine Teilmenge $M \subset V$ setzen wir das* orthogonale Komplement

$$M^\perp = \{v \in V | m \perp v \text{ für alle } m \in M\}.$$

Sind $M, N \subset V$ Teilmengen, so schreiben wir $M \perp N$, wenn $m \perp n$ für alle $m \in M$, $n \in N$.
Für $x, y \in V \setminus \{0\}$ definieren wir durch

$$\cos(\alpha) = \frac{\langle x, y\rangle}{\|x\|\|y\|} \qquad\qquad 0 \leq \alpha \leq \pi$$

den Öffnungswinkel *zwischen x, y.*

Bemerkung. Wegen der Cauchy-Schwarzschen Ungleichung liegt der Bruch im Intervall $[-1, 1]$, also ist der Cosinus wohldefiniert.

Hier ist jetzt die Stelle, an der man alle bekannten und weniger bekannten Sätze der euklidischen Geometrie der Ebene, also der Schulgeometrie, formulieren und beweisen könnte. Leider reicht unsere Zeit nicht dafür.

Orthonormalbasen

Definition 10.10. *Sei V ein euklidischer bzw. unitärer Vektorraum. Ein Tupel $(v_1, \dots, v_n)$ von Vektoren in V heißt* orthonormal*, falls $v_i \perp v_j$ für alle $i \neq j$ und $\|v_i\| = 1$ für alle i. Es heißt* Orthonormalbasis*, wenn es zusätzlich eine Basis ist.*

Die darstellende Matrix des Skalarproduktes bezüglich einer Orthonormalbasis ist also einfach die Einheitsmatrix.

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

Bemerkung. Wir verzichten hier auf den unendlich dimensionalen Fall, der in der Praxis keine Rolle spielt. Im Fall von z.B. Vektorräumen von Funktionen will man unendliche Summen oder sogar Integrale betrachten. Dies ist nicht mehr Gegenstand der linearen Algebra.

Lemma 10.11. *Ein Orthonormalsystem ist linear unabhängig. Für $v \in \langle v_1, \dots, v_n\rangle$ gilt*

$$v = \sum_{i=1}^{n} \langle v_i, v\rangle v_i.$$

Beweis: Seien $v_1, \dots, v_n$ orthonormal, $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, und $v = \sum_{i=1}^{n} \lambda_i v_i$. Dann folgt für $1 \leq j \leq n$

$$\langle v_j, v\rangle = \langle v_j, \sum_{i=1}^{n} \lambda_i v_i\rangle = \sum_{i=1}^{n} \lambda_i \langle v_j, v_i\rangle = \lambda_j$$

nach Voraussetzung an das System. Dies beweist die Formel. Ist speziell $v = 0$, so erhalten wir

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

d.h. die Vektoren sind linear unabhängig. $\square$

Satz 10.12 (Gram-Schmidtsches Orthonormalisierungsverfahren)**.** *Sei V ein euklidischer bzw. unitärer Vektorraum. Seien $v_1, \dots, v_r \in V$ linear unabhängig. Dann ist durch die Rekursionsgleichungen*

$$\begin{aligned} \tilde{v}_1 &= \frac{v_1}{\|v_1\|} \\ w_{k+1} &= v_{k+1} - \sum_{i=1}^{k} \langle \tilde{v}_i, v_{k+1} \rangle \tilde{v}_i \\ \tilde{v}_{k+1} &= \frac{w_{k+1}}{\|w_{k+1}\|} \end{aligned}$$

ein Orthonormalsystem $(\tilde{v}_1, \dots, \tilde{v}_r)$ definiert mit

$$\langle v_1, \dots, v_r \rangle = \langle \tilde{v}_1, \dots, \tilde{v}_r \rangle.$$

Beispiel. $V = \mathbb{R}^2$ mit dem Skalarprodukt $\left\langle \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\rangle = 2a_1b_1 + a_2b_2$. Die Vektoren e_1 und e_2 sind linear unabhängig. Es ist

$$\langle e_1, e_1 \rangle = 2 \Rightarrow \tilde{e}_1 = \frac{1}{\sqrt{2}} e_1$$

Daher

$$w_2 = e_2 - \langle \tilde{e}_1, e_2 \rangle \tilde{e}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = e_2$$

Dieser Vektor ist bereits normiert, also $\tilde{e}_2 = e_2$.

Beweis: Wir argumentieren mit Induktion nach r. Für $r = 1$ ist nach Voraussetzung $v_1 \neq 0$ und daher $\tilde{v}_1$ wohldefiniert. Es gilt

$$\|\tilde{v}_1\| = \|\frac{1}{\|v_1\|} v_1\| = \frac{1}{|\|v_1\||} \|v_1\| = 1$$

wie gewünscht. Weiter gilt für den aufgespannten Vektorraum $\langle v_1 \rangle = \langle \frac{1}{\|v_1\|} v_1 \rangle$. Sei nun $r \geq 1$, so dass die Behauptung gilt. Wir betrachten w_{r+1}.

Behauptung. $w_{r+1} \perp \tilde{v}_j$ *für* $j = 1, \dots, r$.

Wir rechnen:

$$\begin{aligned} \langle \tilde{v}_j, w_{r+1} \rangle &= \langle \tilde{v}_j, v_{r+1}, \rangle - \left\langle \tilde{v}_j, \sum_{i=1}^{r} \langle \tilde{v}_i, v_{r+1} \rangle \tilde{v}_i \right\rangle \\ &= \langle \tilde{v}_j, v_{r+1}, \rangle - \sum_{i=1}^{r} \langle \tilde{v}_i, v_{r+1} \rangle \langle \tilde{v}_j, \tilde{v}_i \rangle \\ &= \langle v_{r+1}, \tilde{v}_j \rangle - \langle v_{r+r}, \tilde{v}_j \rangle = 0 \end{aligned}$$

Wen $v_{r+1} \notin \langle v_1, \dots, v_r \rangle = \langle \tilde{v}_1, \dots, \tilde{v}_r \rangle$ gilt $w_{r+1} \neq 0$ und $\tilde{v}_{r+1}$ ist wohldefiniert und normiert.

Behauptung. $\langle v_1, \ldots, v_{r+r} \rangle = \langle \tilde{v}_1, \ldots, \tilde{v}_{r+1} \rangle$.

Es genügt zu zeigen $v_{r+1} \in \langle \tilde{v}_1, \ldots, \tilde{v}_{r+1} \rangle$. Es gilt

$$v_{r+1} = w_{r+1} + \sum_{i=1}^{r} \langle \tilde{v}_i, v_{r+1} \rangle \tilde{v}_i = \|w_{r+1}\| \tilde{v}_{r+1} + \sum_{i=1}^{r} \langle v_{r+1}, \tilde{v}_i \rangle \tilde{v}_i.$$

□

Korollar 10.13. *Jeder endlich-dimensionale euklidische bzw. unitäre Vektorraum hat eine Orthonormalbasis.*

Beweis: Wir wenden das Verfahren auf eine Basis an. □

Lemma 10.14. *Sei $M \subset V$ eine Untermenge. Dann ist $M^\perp$ ein Untervektorraum. Es gilt $M \cap M^\perp \subset \{0\}$.*

Beweis: Seien $v_1, v_2 \in M^\perp$, $\lambda_1, \lambda_2 \in \mathbb{K}$. Zu überprüfen ist $\lambda_1 v_1 + \lambda_2 v_2 \in M^\perp$. Für $m \in M$ gilt

$$\langle m, \lambda_1 v_1 + \lambda_2 v_2 \rangle = \lambda_1 \langle v_1, m \rangle + \lambda_2 \langle v_2, m \rangle = 0 + 0 = 0.$$

Also liegt die Linearkombination in $M^\perp$.

Sei $m \in M$ mit $m \in M^\perp$. Dann ist m orthogonal zu allen Elementen von M, also auch zu sich selbst. Es folgt

$$\langle m, m \rangle = 0 \Rightarrow m = 0.$$

□

Satz 10.15. *Sei V ein endlich-dimensionaler euklidischer bwz. unitärer Vektorraum, $U \subset V$ ein Untervektorraum. Dann gilt $V = U \oplus U^\perp$.*

Beweis: Wir wissen bereits, dass $U^\perp$ ein Untervektorraum und $U \cap U^\perp = 0$. Zu zeigen bleibt, dass die beiden zusammen ganz V erzeugen. Hierfür wählen wir eine Basis $v_1, \ldots, v_k$ von U, die wir durch $v_{k+1}, \ldots, v_n$ zu einer Basis von V ergänzen. Das Gram-Schmidt-Verfahren liefert dann eine Orthonormalbasis $\tilde{v}_1, \ldots, \tilde{v}_n$, so dass gleichzeitig $\tilde{v}_1, \ldots, \tilde{v}_k$ eine Basis von U ist. Die Vektoren $\tilde{v}_{k+1}, \ldots, \tilde{v}_n$ liegen in $U^\perp$. Damit ist alles gezeigt. □

Orthogonale und unitäre Abbildungen

Definition 10.16. *Eine lineare Abbildung $f : V \to W$ zwischen euklidischen bzw. unitären Vektorräumen heißt* orthogonal *bzw.* unitär*, wenn*

$$\langle f(x), f(y) \rangle = \langle x, y \rangle \qquad \text{für alle } x, y \in V.$$

Falls f zusätzlich bijektiv ist, so heißt f Isomorphismus *(von euklidischen bzw. unitären Vektorräumen).*

Der Begriff *Isometrie* ist ebenfalls für orthogonale und unitäre Abbildungen üblich.

Beispiel. Sei $V = \mathbb{R}^2$ mit dem Standardskalarprodukt. Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ die Multiplikation mit der Drehmatrix $M_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. Dann gilt nach Definition

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

Wir rechnen nach:

$$\begin{aligned} M_\alpha^t M_\alpha &= \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\alpha + \sin^2\alpha & -\cos\alpha\sin\alpha + \sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha + \sin\alpha\cos\alpha & \sin^2\alpha + \cos^2\alpha \end{pmatrix} = E_2. \end{aligned}$$

Es gilt also:

$$\langle M_\alpha x, M_\alpha y\rangle = x^t y = \langle x, y\rangle.$$

Die Abbildung ist orthogonal.

Wir haben gerade gezeigt:

Korollar 10.17. *Sei V endlich-dimensionaler euklidischer bwz. unitärer Vektorraum. Dann ist V isomorph zu $\mathbb{K}^n$ mit dem Standardskalarprodukt.*

Beweis: Sei $v_1, \dots, v_n$ eine Orthonormalbasis von V. Wie immer definieren wir

$$f : \mathbb{K}^n \to V, \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto \sum_{i=1}^n a_i v_i.$$

Dies ist eine lineare Abbildung und bijektiv, da $v_1, \dots, v_n$ eine Basis ist. Wir überprüfen das Skalarprodukt. Sei $a = (a_1, \dots, a_n)^t$, $b = (b_1, \dots, b_n)^t$. Dann gilt

$$\langle f(a), f(b)\rangle_V = \sum_{i,j=1}^n \overline{a}_i b_j \langle v_i, v_j\rangle_V = \sum_{i=1}^n \overline{a}_i b_i = \langle a, b\rangle_{\mathbb{K}^n}$$

da $v_1, \dots, v_n$ eine Orthonormalbasis ist. □

Lemma 10.18. *Sei $f : V \to W$ orthogonal bzw. unitär.*

(i) $x \perp y \Rightarrow f(x) \perp f(y)$.

(ii) $\|f(v)\| = \|v\|$.

(iii) *f ist injektiv.*

(iv) *Falls $V = W$ und* $\dim V < \infty$, *so ist f ein Isomorphismus.*

(v) *Falls $V = W$ und λ ein Eigenwert, so gilt $|\lambda| = 1$.*

(vi) *Ist f ein Isomorphismus, so ist die Umkehrabbildung ebenfalls orthogonal bzw. unitär.*

(vii) *Die Komposition zweier Isometrien ist eine Isometrie.*

Beweis: Wir rechnen nach:

$$0 = \langle x, y \rangle = \langle f(x), f(y) \rangle$$

und

$$\|f(v)\| = \sqrt{\langle f(v), f(v) \rangle} = \sqrt{\langle v, v \rangle} = \|v\|.$$

Insbesondere

$$f(v) = 0 \Rightarrow \|f(v)\| = 0 \Rightarrow \|v\| = 0 \Rightarrow v = 0.$$

Ist zusätzlich $f : V \to V$ Endormorphismus eines endlich-dimensionalen Vektorraums, so ist f nach der Dimensionsformel auch surjektiv, also ein Isomorphismus.
Ist v Eigenvektor zum Eigenwert λ, so gilt

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

Da $\|v\| \neq 0$ folgt hieraus $|\lambda| = 1$.
Sei $f : V \to W$ ein Isomorphismus mit Umkehrabbildung g. Wir verizieren, dass g orthogonal bzw. unitär ist. Es gilt für alle $x, y \in W$

$$\langle g(x), g(y) \rangle_V = \langle f(g(x)), f(g(y)) \rangle_W = \langle x, y \rangle_W.$$

Seien $f : V \to W$ und $g : W \to U$ orthogonal bzw. unitär. Dann gilt für alle $x, y \in V$

$$\langle x, y \rangle_V = \langle f(x), f(y) \rangle_W = \langle g(f(x)), g(f(y)) \rangle_U,$$

d.h. $g \circ f$ ist ebenfalls orthogonal bzw. unitär. □

Korollar 10.19. *Die Menge der orthogonalen bzw. unitären Autormorphismen eines euklidischen bzw. unitären Vektorraums ist eine Gruppe.*

Definition 10.20. *Sei V euklidisch. Die* orthogonale Gruppe *$O(V)$ ist die Gruppe der orthogonalen Automorphismen von V. Speziell für $V = \mathbb{R}^n$ mit dem Standarskalarprodukt schreiben wir auch $O(n)$.*
Sei V unitär. Die unitäre Gruppe *$U(V)$ ist die Gruppe der unitären Automorphismen von V. Speziell für $\mathbb{C}^n$ mit dem Standardskalarprodukt schreiben wir auch $U(n)$.*

Ein Automorphismus des $\mathbb{R}^n$ ist stets durch Multiplikation mit einer Matrix $M \in M_n(\mathbb{R})$ gegeben, d.h. es ist $O(n) \subset \mathrm{GL}_n(\mathbb{R})$. Ebenso folgt $U(n) \subset \mathrm{Gl}_n(\mathbb{C})$.

Lemma 10.21. *Sei $M \in M_n(\mathbb{K})$. Es gilt $M \in O(n)$ bzw. $M \in U(n)$ genau dann, wenn*

$$\overline{M}^t M = E_n.$$

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

Beweis: Sei $x, y \in \mathbb{R}^n$. Dann gilt

$$\langle Mx, My\rangle = \overline{x}^t \overline{M}^t My.$$

Falls $\overline{M}^t M = E_n$, so ist die Abbildung orthogonal bzw. unitär. Ist umgekehrt die Abbildung orthogonal bzw. unitär, so setzen wir $x = e_i, y = e_j$. Dann ist Me_i die i-te Spalte von M. Die Bedingung

$$\langle Me_i, Me_j\rangle = \langle e_i, e_j\rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

besagt also, dass die Spalten eine Orthonormalbasis bilden. Der Eintrag von $\overline{M}^t M$ an der Stelle ij ist gerade $\langle Me_i, Me_j\rangle$, also folgt auch $\overline{M}^t M = E_n$. □

Definition 10.22. *Eine Matrix $M \in M_n(\mathbb{R})$ bzw $M \in M_n(\mathbb{C})$ mit $\overline{M}^t M = E_n$ heißt* orthogonal *bzw.* unitär.

Jede solche Matrix ist invertierbar mit $M^{-1} = \overline{M}^t$. Es sind genau die Elemente der orthogonalen bzw. unitären Gruppe.

Beispiel.

$$O(1) = \{\pm 1\}, \quad O(2) = \{M_\alpha | \alpha \in \mathbb{R}\} \times \{T^n | n \in \mathbb{Z}\}$$

wobei M_α wieder die Drehmatrix um den Winkel α ist und $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

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

Beweis: Für $n = 1$ haben wir es mit $M_1(\mathbb{K}) = \mathbb{K}$ zu tun, also mit $a \in \mathbb{K}$, das die Bedingung $|a|^2 = 1$ erfüllt. Für $\mathbb{K} = \mathbb{R}$ sind das gerade ± 1. Für $\mathbb{K} = \mathbb{C}$ erhalten wir die Einheitskreislinie.

Wir betrachten nun $O_2(\mathbb{R})$. Dies sind Matrizen

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$$

mit

$$a^2 + c^2 = b^2 + d^2 = 1, ab + cd = 0.$$

Die Transpositionsmatrix $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ liegt in $O(2)$ und hat die Determinante -1. Nach eventueller Multiplikation mit T können wir also erreichen, dass $\det(M) > 0$.

Insbesondere gilt $a, b, c, d \in [-1, 1]$. Sei $\alpha \in \mathbb{R}$ mit $a = \cos(\alpha)$. Dann ist $c = \pm\sin(\alpha)$. Indem wir eventuell α durch $-\alpha$ ersetzen, erreichen wir $c = -\sin(\alpha)$. Aus der Relation $ab + cd = 0$ schliessen wir

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

für ein $\lambda \in \mathbb{R}$. Es folgt

$$b^2 + d^2 = \lambda^2(\sin^2\alpha + \cos^2\alpha) = \lambda^2 = \lambda^2 = 1,$$

$\lambda = \pm 1$. Für die Determinante erhalten wir

$$\det M = ad - bc = \lambda \cos^2\alpha + \lambda \sin^2\alpha = \lambda > 0$$

also $\lambda = 1$. $\square$

Bemerkung. Sei $M \in U(n)$. Wir werden noch sehen, dass $\mathbb{C}^n$ eine Basis aus Eigenvektoren von M hat. Als Produkt der Eigenwerte hat die Determinante den komplexen Betrag 1. Jede orthogonale Matrix ist auch unitär. Also ist ihre Determinante reell mit Betrag 1, also ± 1. Wir haben diese Eigenschaften im obigen Beispiel konkret gesehen.

Bemerkung. Sei V euklidisch bzw. unitär mit Orthonormalbasis $v_1, \dots, v_n$. Der Isomorphismus aus Korollar 10.17 induziert einen Gruppenisomorphismus

$$O(V) \to O(n).$$

Kapitel 11

Der Spektralsatz

Wir arbeiten weiterhin über $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

Selbstadjungierte Abbildungen

Definition 11.1. *Sei V ein euklidischer bzw. unitärer Vektorraum. Ein Endomorphismus $f : V \to V$ heißt* selbstadjungiert, *wenn für alle $x, y \in V$ gilt*

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

Beispiel. Sei $V = \mathbb{K}^n$ mit dem Standardskalarprodukt, f gegeben durch die Multiplikation mit der Matrix $M \in M_n(\mathbb{K})$. Die Bedingung lautet dann

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

für alle x, y. Wenn M symmetrisch bzw. hermitesch ist, also $\overline{M}^t = M$, so ist die Abbildung selbstadjungiert.

Beispiel. In der Quantenmechanik ist der Zustandsraum ein unendlich-dimensionaler unitärer Vektorraum, z.B.

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

Physikalische Größen wie Ort und Impuls sind selbstadjungierte Endormorphismen. (Das ist ein bisschen gelogen. Die Wahrheit erfahren Sie vielleicht in Physik.) Die Messwerte sind Eigenwerte dieser Endomorphismen.

Lemma 11.2. *Sei V endlich-dimensionaler euklidischer bzw. unitärer Vektorraum, $f : V \to V$ ein Endomorphismus. Dann ist f genau dann selbstadjuniert, wenn die darstellende Matrix von f bezüglich einer Orthonormalbasis symmetrisch bzw. hermitesch ist.*

Beweis: Die Wahl der Orthonormalbasis definiert einen Isomorphismus von euklidischen bzw. unitären Vektorräumen $V \cong \mathbb{K}^n$. Sei M die darstellende Matrix von f in der Orthonormalbasis, also der auf $\mathbb{K}^n$ induzierte Endormorphismus. Wir haben gerade gesehen, dass die Abbildung selbstadjuniert ist, wenn $\overline{M}^t = M$. Sei umgekehrt die Abbildung selbstadjungiert. Wir setzen $x = e_i$, $y = e_j$. Dann ist Me_j die j-te Spalte von M und $e_i^t M e_j$ der i-te Eintrag dieser Spalte, also m_{ij}. Genauso ist $e_i^t \overline{M}^t e_j$ der ij-te Eintrag von $\overline{M}^t$, also $\overline{m}_{ji}$. Die Matrix ist symmetrisch bzw. hermitesch. □

Unser nächstes großes Ziel ist das folgende wichtige Resultat:

Theorem 11.3 (Spektralsatz)**.** *Sei V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum, $f : V \to V$ ein selbstadjungierter Endomorphismus. Dann besizt V eine Orthonormalbasis aus Eigenvektoren von f. Alle Eigenwerte sind reell.*

Wir tasten uns in mehreren Schritten an den Beweis heran.

Lemma 11.4. *Sei V unitärer Vektorraum, $f : V \to V$ ein selbstadjungierter Endomorphismus, λ ein Eigenwert von f. Dann ist λ reell.*

Beweis: Sei v Eigenvektor zum Eigenwert λ. Wir rechnen nach:

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

Wegen $\langle v, v\rangle \neq 0$ folgt

$$\overline{\lambda} = \lambda.$$

□

Lemma 11.5. *Sei V endlich-dimensionaler euklidischer Vektorraum, $f : V \to V$ selbstadjunigert. Dann besitzt f einen Eigenwert.*

Beweis: Sei M darstellende Matrix von f bezüglich einer Orthonormalbasis von V. Dann ist M symmetrisch über $\mathbb{R}$, also auch hermitesch über $\mathbb{C}$. Über $\mathbb{C}$ hat das charakteristische Polynom eine Nullstelle λ. Nach Lemma 11.4 ist dieser Eigenwert reell. Diese reelle Nullstelle des charakteristischen Polynoms ist der gesucht Eigenwert des ursprünglichen f. □

Lemma 11.6. *Sei V euklidischer oder unitärer Vektorraum, $f : V \to V$ selbstadjungiert. Sei v ein Eigenvektor von f. Wir setzen $U = v^\perp$. Dann gilt $f(U) \subset U$.*

Beweis: Sei v Eigenvektor zum Eigenwert λ. Sei $x \in U$. Wir überprüfen $f(x) \in U = v^\perp$. Es gilt

$$\langle v, f(x)\rangle = \langle f(v), x\rangle = \langle \lambda v, x\rangle = \overline{\lambda}\langle v, x\rangle = 0$$

denn $x \perp v$. □

Beweis des Spektralsatzes. Sei V endlich-dimensionaler euklidischer oder unitärer Vektorraum, $f : V \to V$ selbstadjunigert. Wir argumentieren mit vollständiger Induktion über $\dim(V)$.

Falls $\dim(V) = 0$, so ist nichts zu zeigen. Sei die Behauptung wahr für alle Vektorräume der Dimension $n-1$. Sei $\dim V = n$. Sei v ein Eigenvektor. Falls $\mathbb{K} = \mathbb{C}$, so existiert dieser als Nullstelle des charakteristischen Polynoms. Falls $\mathbb{K} = \mathbb{R}$, so existiert er nach Lemma 11.5. Wir setzen

$$v_1 = \frac{1}{\|v\|} v.$$

Dieser Vektor hat Länge 1.

Sei $U = v_1^\perp$. Dann gilt

$$V \cong \langle v_1 \rangle \oplus U.$$

Daher ist $\dim(U) = n-1$. Nach Lemma 11.6 induziert f einen Endomorphismus $f|_U : U \to U$. Dieser ist ebenfalls selbstadjunigert bezüglich des Skalarproduktes, das man durch Einschränken von V erhält. Nach Induktionsvoraussetzung hat U eine Orthonormalbasis $v_2, \dots, v_n$ aus Eigenvektoren von $f|_U$. Diese stehen senkrecht auf v_1. Das System $v_1, \dots, v_n$ ist die gesuchte Orthonormalbasis aus Eigenvektoren von f.

Wir haben bereits in Lemma 11.4 gezeigt, dass die Eigenwerte reell sind. □

Dieselben Ideen erlauben auch den Beweis eines allgemeineren Satzes. Dafür müssen wir ausholen.

Normale Abbildungen

Definition 11.7. *Seien V, W euklidische bzw. unitäre Vektorräume. Seien $f : V \to W$ und $f^* : W \to V$ $\mathbb{K}$-linear. Dann heißt f^** adjungiert *zu f, wenn für alle $v \in V$ und $w \in W$ gilt*

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

Beispiel. Ist $V = W$ und f selbstadjungiert, so ist $f^* = f$.

Bemerkung. Ist f^* adjungiert zu f, so auch f zu f^*, denn

$$\langle v, f^*(w) \rangle_V = \overline{\langle f^*(w), v \rangle} = \overline{\langle w, f(v) \rangle} = \langle f(v), w \rangle.$$

Lemma 11.8. *Die Abbildung f^* ist eindeutig bestimmt durch f.*

Beweis: Seien g, g' beide adjungiert zu f. Dann gilt für alle $w \in W$, $v \in V$

$$\langle (g - g')(w), v \rangle = \langle w - w, f(v) \rangle = 0,$$

also $(g - g')(w) \in V^\perp = 0$. □

Beispiel. Seien $V = \mathbb{K}^n$ und $W = \mathbb{K}^m$ mit Standardskalarprodukt, f gegeben durch Multiplikation mit der Matrix M. Dann ist f^* gegeben durch Multiplikation mit der Matrix $\overline{M}^t$, denn

$$\langle \overline{M}^t v, w\rangle = \overline{(\overline{M}^t v)}^t w = \overline{v}^t M w = \langle v, Mw\rangle.$$

Definition 11.9. *Sei $M \in M_{n\times m}(\mathbb{K})$. Dann heißt*

$$M^* := \overline{M}^t$$

adjungierte Matrix *von M.*

Diese Formel ist gleichzeitig der Existenzbeweis für die adjungierte Abbildung.

Korollar 11.10. *Seien V und W euklidische bwz. unitäre Vektorräume mit Orthonormalbasen $(v_1, \dots, v_n)$ und $(w_1, \dots, w_m)$. Sei $f : V \to W$ linear mit darstellender Matrix M bezüglich dieser Basen.*

Dann existiert die adjungierte Abbildung f^* und hat darstellende Matrix M^* bezüglich dieser Basen.

Wir kehren nun zurück zu Endomorphismen.

Definition 11.11. *Sei V ein endlich-dimensionaler unitärer Vektorraum, $f : V \to V$ ein Endomorphismus. Dann heißt f* normal*, wenn f mit f^* vertauscht, also*

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

Beispiel.
(i) Wenn f selbstadjungiert ist, also $f = f^*$, so ist f auch normal.

(ii) Wenn f unitär ist, so erfüllt die darstellende Matrix bezüglich einer Orthonormalbasis $M^*M = E_n$, also $M^* = M^{-1}$. Jede invertierbare Matrix vertauscht mit ihrer Inversen, also ist f normal.

Theorem 11.12. *Sei V endlich-dimensionaler unitärer Vektorraum, $f : V \to V$ ein normaler Endomorphismus. Dann besitzt V eine Orthonormalbasis aus gemeinsamen Eigenvektoren von f und f^*.*

Ist v Eigenvektor von f zum Eigenwert λ, dann ist v Eigenvektor von f^* zum Eigenwert $\overline{\lambda}$.

Wir zerlegen den Beweis wieder in mehrere Lemmas.

Lemma 11.13. *Sei V endlichdimensionaler unitärer Vektorraum und $f : V \to V$ normal. Sei λ ein Eigenwert von f. Dann ist $\overline{\lambda}$ ein Eigenwert von f^* und es existiert ein gemeinsamer Eigenvektor.*

Beweis: Sei V_λ der Eigenraum von f zum Eigenwert λ, also

$$V_\lambda = \{v \in V | f(v) = \lambda v\}.$$

Behauptung. $f^*(V_\lambda) \subset V_\lambda$.

Sei $v \in V_\lambda$. Wir betrachten $f^*(v)$. Es gilt

$$f(f^*(v)) = f^*(f(v)) = f^*(\lambda v) = \lambda f^*(v)$$

also liegt auch $f^*(v)$ in V_λ.
Wir betrachten nun $f^*|_{V_\lambda} : V_\lambda \to V_\lambda$. Das charakteristische Polynom dieses Endomorphismus hat eine Nullstelle μ, also gibt es einen Eigenvektor $v \in V_\lambda$ von f^*. Dieser ist gemeinsamer Eigenvektor von f und f^*. Weiter gilt

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

Wegen $\|v\| \neq$ folgt $\overline{\mu} = \lambda$, also $\mu = \overline{\lambda}$. □

Lemma 11.14. *Sei V unitärer Vektorraum, $f : V \to V$ normal, v ein Eigenvektor von f^*. Wir setzen $U = v^\perp$. Dann gilt $f(U) \subset U$.*

Beweis: Sei $x \in U$. Wir betrachten $f(x)$.

$$\langle v, f(x)\rangle = \langle f^*(v), x\rangle = \langle \mu v, x\rangle = \overline{\mu} v, x\rangle = 0.$$

□

Beweis des Spektralsatzes für normale Endomorphismen. Sei V endlich-dimensionaler unitärer Vektorraum, $f : V \to V$ normal. Wie im selbst-adjungierten Fall argumentieren wir mit vollständiger Induktion nach $\dim(V)$. Für $\dim(V) = 0$ ist nichts zu zeigen.
Sei die Aussage wahr für alle Vektorräume der Dimension $n-1$. Sei $\dim(V) = n$. Nach Lemma 11.13 existiert ein gemeinsamer Eigenvektor v_1 von f und f^*. Ohne Einschränkung erfüllt er $\|v_1\| = 1$. Sei $U = v_1^\perp$. Nach Lemma 11.14 gilt $f(U) \subset U$. Wegen $V = \langle v_1\rangle \oplus U$ gilt $\dim(U) = n - 1$. Wir wenden die Induktionsvoraussetzung an auf $f|_U$. Wegen $(f|_U)^* = f^*|_U$ ist dieser Endomorphismus normal. Also existiert eine Basis $v_2, \dots, v_n$ von U, die aus gemeinsamen Eigenvektoren von f und f^* besteht.
Insbesondere gilt $V_\lambda(f) = V_{\overline{\lambda}}(f^*)$, also sind alle Eigenvektoren gemeinsame Eigenvektoren. □

Diese Version des Spektralsatzes gilt insbesondere für unitäre Endormorphismen. Das Beispiel der Drehmatrizen zeigt, dass er für *reelle* normale Endormorphismen nicht gilt.

Korollar 11.15. *Sei V endlich-dimensionaler euklidischer bzw. unitärer Vektorraum, $f : V \to V$ orthogonal bzw. unitär. Dann gilt $|\det(f)| = 1$ bzw. $\det(f) = \pm 1$.*

Beweis: Sei M die darstellende Matrix von f in einer Orthonormalbasis, also M orthogonal bzw. unitär.

Für $\mathbb{K} = \mathbb{C}$ wenden wir den Spektralsatz an und sehen, dass es eine Basis aus Eigenvektoren gibt. Die Eigenwerte $\lambda_1, \dots, \lambda_n$ haben Betrag 1. Wegen

$$\det(f) = \det(M) = \prod_{i=1}^{n} \lambda_i$$

folgt $|\det(f)| = 1$.

Für $\mathbb{K} = \mathbb{R}$ bemerken wir, dass die orthogonale Matrix auch unitär ist, also $|\det(f)| = |\det(M)| = 1$. Da diesmal die Determinante reell ist, folgt $\det(f) = \pm 1$. $\square$

Hauptachsentransformation

Der Spektralsatz kann auch ganz anders interpretiert werden. Wir sind weiter in der Situation $\mathbb{K} = \mathbb{R}, \mathbb{C}$, aber jetzt ohne Skalarprodukt. Wir erinnern uns, dass die darstellende Matrix einer bilinearen bzw. sesquilinearen symmetrischen bzw. hermiteschen Abbildung symmetrisch bzw. hermitesch als Matrix ist.

Aus dem Spektralsatz wissen wir, dass es für diese Matrix M eine Basis aus Eigenvektoren gibt, also eine Orthonormalbasis $v_1, \dots, v_n \in \mathbb{K}^n$ mit $Mv_i = \lambda_i v_i$. Für die Sesquilinearform bedeutet dies

$$\overline{v}_i M v_j = \overline{v}_i(\lambda_j v_j) = \lambda_i \delta_{ij}.$$

Die Basiswechselmatrix ist orthogonal bzw. unitär. Damit haben wir gezeigt:

Korollar 11.16. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, $s : V \times V \to \mathbb{K}$ symmetrische Bilinearform bzw. hermitesche Sesquilinearform. Sei M darstellende Matrix von s bezüglich einer Basis. Dann existiert ein Basiswechselmatrix S, so dass die neue darstellende Matrix S^*MS eine Diagonalmatrix ist.*

Die Einträge der Diagonalmatrix sind also die Eigenwerte von M.

Satz 11.17. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, $s : V \times V \to \mathbb{K}$ symmetrische Bilinearform bzw. hermitesche Sesquilinearform, M darstellende Matrix bezüglich einer Basis. Dann ist s genau dann ein Skalarprodukt, wenn alle Eigenwerte von M positiv sind.*

Beweis: Sei $v_1, \dots, v_n$ Basis aus Eigenvektoren für die lineare Abbildung f, die durch die Matrix M definiert wird. Dann gilt

$$s(v_i, v_i) = \lambda_i.$$

Wenn s ein Skalarprodukt ist, so ist diese Zahl positiv. Sind alle $\lambda_i > 0$, so gilt für beliebige $a = \sum a_i v_i \neq 0$

$$s(v, v) = \sum |a_i|^2 \lambda_i > 0.$$

$\square$

Bemerkung. Die darstellende Matrix eines Endomorphismus transformiert wie $S^{-1}MS$. Die darstellende Matrix einer Bilinearform dagegen als S^*MS. Dabei bleiben die Eigenwerte *nicht* erhalten. Im Spektralsatz erhält man eine Orthonormalbasis, d.h. die Basiswechselmatrix ist unitär und $S^* = S^{-1}$.

Theorem 11.18 (Trägheitssatz von Sylvester)**.** *Sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum, $s : V \times V \to \mathbb{R}$ eine symmetrische bilineare Abbildung. Dann gibt es eine Basis von V, so dass die darstellende Matrix von s die Form*

$$\begin{pmatrix} E_n & 0 & 0 \\ 0 & -E_m & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

hat. Die Zahlen n und m sind unabhängig von der Wahl der Basis.

Beweis: Sei $v_1, \dots, v_N$ eine Basis, so dass die darstellende Matrix von s eine Diagonalmatrix mit Einträgen $\lambda_1, \dots, \lambda_N$ ist. Wir sortieren so, dass die Einträge $\lambda_1, \dots, \lambda_n$ positiv sind, $\lambda_{n+1}, \dots, \lambda_{n+m}$ negativ und die restlichen 0. Für $i = 1, \dots, n+m$ gehen wir von v_i zu

$$\sqrt{|\lambda_i|}^{-1} v_i$$

über. In der neuen Basis hat die darstellende Matrix die behauptete Form.
Wir wenden uns der Eindeutigkeit zu. Wir setzen

$$V_1 = \langle v_1, \dots, v_n \rangle, V_{-1} = \langle v_{n+1}, \dots, v_{n+m} \rangle, V_0 = \langle v_{n+m+1}, \dots, v_n \rangle.$$

Auf V_1 definiert s ein Skalarprodukt, auf V_{-1} gilt dies für $-s$.
Sei $v'_1, \dots, v'_N$ eine zweite solche Basis und analog V'_1, V'_{-1}, V'_0.

Behauptung. $V_0 = \{v \in V | s(v, w) = 0 \text{ für alle } w \in V\}$

Man liest an der Matrix ab, dass die Element von v diese Bedingung erfüllen. Sei umgekehrt $v = v_1 + v_{-1} + v_0$ auf der rechten Seite. Dann gilt

$$0 = s(v, v_1) = s(v_1, v_1)$$

Da $s|_{V_1}$ ein Skalarprodukt ist, folgt $v_1 = 0$. Genauso sehen wir $V_{-1} = 0$, also $v = v_0 \in V_0$.
Die rechte Seite in dieser Beschreibung hängt nicht von den konkreten Basiswahlen ab, es gilt also $V_0 = V'_0$. Durch Vergleich der Dimensionen erhalten wir

$$n + m = n' + m'.$$

Sei $r = N - n - m = \dim V_0$.

Behauptung. $V_1 \cap (V'_{-1} \oplus V_0) = 0$

Sei $v \neq 0$ im Schnitt. Da $v \in V_1$ gilt $s(v,v) > 0$. Da $v \in V'_{-1} + V_0$ gilt $s(v,v) \leq 0$. Das ist ein Widerspruch.

Wir haben $n = \dim V_1$, $m = \dim V_{-1}$, $r = \dim V_0 = \dim V'_0$, $n' = \dim V'_1$, $m' = \dim V'_{-1}$. Wir lesen ab

$$n + m' + r = \dim V_1 \oplus V'_{-1} \oplus V_0 \leq \dim N = n + m + r \Rightarrow m' \leq m.$$

Genauso zeigt man $m \leq m'$, also $m' = m$. Dann gilt automatisch auch $m = m'$. □

Definition 11.19. *In der Situation des Trägheitssatzes heißt $n + m$ der* Rang *von s, n der* Index *und $n - m$* Signatur*.*

Eine reelle bilineare Abbildung ist also durch zwei dieser Zahlen eindeutig bestimmt. Oft gibt man die Signatur auch in der Form (n, m) an.

Beispiel. Spezielle und allgemeine Relativitätstheorie arbeiten mit dem *Ninkowski-Raum*: $\mathbb{R}^4$ und einer Bilinearform der Signatur $(3, 1)$.

Korollar 11.20. *Eine symmetrische Bilinearform der Dimension n ist genau dann ein Skalarprodukt, wenn sie Rang und Index n hat.*

Im Prinzip haben wir damit ein Verfahren, mit dem wir entscheiden können, ob s ein Skalarprodukt definiert.

Beispiel. Wir betrachten $M = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$. Das charakteristische Polynom ist

$$\det \begin{pmatrix} 2 - X & 2 \\ 2 & 2 - X \end{pmatrix} = (2 - X)^2 - 4 = X^2 - 4X = X(X - 4).$$

Die Eigenwerte sind 0 und 4, also definiert s kein Skalarprodukt.

Tatsächlich gibt es ein besseres Verfahren: mit quadratischer Ergänzung. Die Grundidee ist es, die quadratische Gleichung zu lösen, die durch die Bilinearform gelöst wird. Die Methode ist dieselbe wie in der Schule bei der Herleitung der Lösungsformel für quadratische Gleichungen in einer Variablen. Statt des allgemeinen Algorithmus stellen wir ein Beispiel vor.

Beispiel. Wir betrachten $V = \mathbb{R}^3$,

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

Wir erhalten für $x = (x_1, x_2, x_3)^t$

$$\begin{aligned} x^t M x &= 2x_1^2 + 2x_1x_2 + x_2^2 + 4x_1x_3 - 2x_2x_3 + x_3^2 \\ &= (\sqrt{2}x_1)^2 + \sqrt{2}(\sqrt{2}x_1)x_2 + x_2^2 + 2\sqrt{2}(\sqrt{2}x_1)x_3 - 2x_2x_3 + x_3^2 \end{aligned}$$

Wir rechnen mit dieser quadratischen Form weiter. Im ersten Schritt ersetzen wir $\sqrt{2}x_1$ durch x'_1. (Die Quadratwurzel existiert immer, denn $2 = s(e_1, e_1) > 0$. Falls $s(e_1, e_2) \leq 0$ ist s kein Skalarprodukt.) Dann erhält man eine neue quadratische Form

$$x'^2_1 + \sqrt{2}x'_1 x_2 + x_2^2 + 2\sqrt{2}x'_1 x_3 - 2x_2 x_3 + x_3^2$$

Im nächsten Schritt schreiben wir um zu

$$y_1 = x'_1 + \frac{1}{\sqrt{2}}x_2 + \sqrt{2}x_3$$

Die Linearkombination ist so gewählt, dass die gemischten Terme mit x_1 eliminert werden. Die neue Form ist

$$\begin{aligned} y_1^2 + \frac{1}{2}x_2^2 - 2x_2x_3 - x_3^2 \\ &= (x'^2_1 + \frac{1}{2}x_2^2 + 2x_3^2 + \sqrt{2}x'_1x_2 + 2\sqrt{2}x'_1x_3) + \frac{1}{2}x_2^2 - 2x_2x_3 - x_3^2 \\ &= x'^2_1 + x_2^2 + x_3^2 + \sqrt{2}x'_1x_2 + 2\sqrt{2}x'_1x_3 - 2x_2x_3 \end{aligned}$$

An dieser Stelle sieht man schon, dass es sich nicht um ein Skalarprodukt handelt, denn für $y_1 = 0$, $x_2 = 0$, $x_3 = 1$ erhalten wir den Wert -1. Aber streng im Algorithmus: Wir ersetzen $\sqrt{2}^{-1}x_2$ durch x'_2. Die Form wird dann zu

$$y_1^2 + x'^2_2 - \sqrt{2}x'_2x_3 - x_3^2.$$

Die nächste quadratische Ergänzung eliminiert die gemischten Terme mit x'_2. Wir setzen

$$y_2 = x'_2 - \frac{1}{\sqrt{2}}x_3$$

und erhalten als neue quadratische Form

$$y_1^2 + y_2^2 - \frac{3}{2}x_3^2.$$

Wir haben eine rein-quadratische Form erreicht. Sie ist nicht positiv-definit. Es handelt sich nur um lineare Koordinatenwechsel bzw. Basiswechsel auf $\mathbb{R}^3$. Auch die ursprüngliche Abbildung s ist kein Skalarprodukt.

Kapitel 12

Diagonalisierbarkeit und Trigonalisierbarkeit

Ab jetzt ist k wieder ein beliebiger Körper, nicht nur $\mathbb{R}$ oder $\mathbb{C}$. Die Vektorräume sind in der Regel endlich-dimensional.

Wir studieren wieder systematisch Eigenwerte, Eigenvektoren und Eigenräume. Aus LA 1 kennen wir das charakteristische Polynom.

Definition 12.1. *Sei V ein Vektorraum, $f : V \to V$. Der Endomorphismus heißt* diagonalisierbar*, wenn eine Basis aus Eigenvektoren existiert. Diese Basis heißt* Eigenbasis.

Eine Matrix $M \in M_n(k)$ heißt diagonisierbar*, wenn der zugehörige Endomorphismus von k^n diagonalisierbar ist.*

Bemerkung.

(i) Die darstellenden Matrix bezüglich einer Eigenbasis ist eine Diagonalmatrix, deren Einträge die Eigenwerte sind.

(ii) Eine Matrix $M \in M_n(k)$ ist diagonisierbar, wenn es $S \in \mathrm{Gl}_n(k)$ gibt mit $S^{-1}MS$ eine Diagonalmatrix.

(iii) Wenn f diagonalisierbar ist, so zerfällt das charakteristische Polynom in Linearfaktoren.

Wir kennen bereits einige Fälle, in denen der Endormorphismus automatisch diagonalisierbar ist:

(i) $k = \mathbb{R}$, V euklidisch, f selbst-adjungiert (Theorem 11.3)

(ii) $k = \mathbb{C}$, V unitär, f normal, z.b. unitär oder selbst-adjungiert (Theorem 11.12

(iii) k beliebig, $\dim V = n$ und χ_f hat n verschiedene Nullstellen (LA1 Korollar 7.19)

Wir wissen bereits, dass *nicht* jeder Endomorphismus diagonalisierbar ist.

(i) Falls k nicht algebraisch abgeschlossen ist, zerfällt das charakteristische Polynom i.a. nicht in Linearfaktoren. Ein Beispiel für $k = \mathbb{R}$ sind die Drehmatrizen.

(ii) Selbst wenn das charakteristische Polynom zerfällt, ist der Endomorphismus i.a. nicht diagonalisierbar. Ein Beispiel hierfür sind Matrizen der Form

$$\begin{pmatrix} \lambda & 1 & 0 & & \dots \\ 0 & \lambda & 1 & 0 & \dots \\ & & \vdots & & \\ 0 & & \dots & & \lambda \end{pmatrix}$$

Das charakteristische Polynom ist $(\lambda - X)^n$. Einziger Eigenvektor ist e_1, wie man leicht nachrechnet.

Eigenschaften des Polynomrings, Teil 2

Wir betrachten den Polynomring

$$k[X] = \left\{ \sum_{i=0}^{n} a_i X^i | n \in \mathbb{N}_0, a_i \in k \text{ für } i \geq 0 \right\}.$$

Definition 12.2. *Für $P \in k[X]$, $P \neq 0$, $\lambda \in k$ sei* $\operatorname{ord}_\lambda(P)$ *die maximale Zahl, so dass $(X - \lambda)^n$ ein Teiler von P ist. Sie heißt* Multiplizität *oder* Nullstellenvielfachheit *von λ.*

Wir wissen bereits aus LA 1 Korollar 7.7, dass ein Polynom höchstens so viele Nullstellen hat, wie sein Grad.

Satz 12.3. *Sei $P \in k[X]$, $\lambda_1, \dots, \lambda_r \in k$ die paarweise verschiedenen Nullstellen von P mit den Multiplizitäten $n_1, \dots, n_r \in \mathbb{N}$. Dann gilt*

$$P = (X - \lambda_1)^{n_1} \dots (X - \lambda_r)^{n_r} Q$$

mit $Q \in k[X]$ nullstellenfrei.

Wo ist das Problem? Nach Voraussetzung gilt

$$P = (X - \lambda_1)^{n_1} P_1, \quad P = (X - \lambda_2)^{n_2} P_2.$$

Warum ist $(X - \lambda_2)$ und sogar $(X - \lambda_2)^{n_2}$ ein Teiler von P_1?

Beweis: Der wesentliche Schritt ist:

Behauptung. *Seien $A, B \in k[X]$ und λ eine Nullstelle von AB. Dann ist $X - \lambda$ ein Teiler von A oder B (oder von beiden).*

Angenommen, $X - \lambda$ ist weder Teiler von A noch von B. Division mit Rest ergibt

$$A = A_1(X - \lambda) + A_2, B = A_2(X - \lambda) + B_2$$

wobei $\deg(A_2), \deg(B_2) < \deg(X - \lambda) = 1$. Also sind A_2 und B_2 konstant. Wir schreiben ab jetzt a_2 und b_2 für diese Konstanten. Einsetzen von λ ergibt

$$A(\lambda) = a_2, B(\lambda) = b_2$$

also sind $a_2, b_2 \neq 0$. Wir multiplizieren und erhalten

$$AB = A_1B_1(X - \lambda)^2 + A_1b_2(X - \lambda) + a_2B_1(X - \lambda) + a_2b_2.$$

Einsetzen von λ ergibt dann

$$AB(\lambda) = a_2b_2.$$

Das ist ungleich 0, da $a_2, b_2 \in k^*$. Dies ist ein Widerspruch. Die Behauptung ist bewiesen.

Der Rest des Argumentes ist vollständige Induktion nach dem Grad und den Multiplizitäten. (Übungsaufgabe oder Vorlesung Algebra und Zahlentheorie). □

Diagonalisierbarkeit

Definition 12.4. *Sei V endlich-dimensionaler k-Vektorraum, $f : V \to V$ ein Endomorphismus, λ ein Eigenwert.*

(i) Die algebraische Vielfachheit *von λ ist die Nullstellenordnung von λ im charakteristischen Polynom χ_f.*

(ii) Die geometrische Vielfachheit *von λ ist die Dimension des Eigenraumes $V_\lambda = \{v \in V | f(v) = \lambda v\}$.*

Satz 12.5. *Sei V endlich-dimensionaler Vektorraum, $f \in \mathrm{End}(V)$. Seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von f mit geometrischer Vielfachheit μ_i und algebraischer Vielfachheit ν_i. Dann gilt*

$$\mu_i \leq \nu_i \quad \text{für alle } i.$$

Es sind äquivalent:

(i) Das charakteristische Polynom zerfällt in Linearfaktoren und $\mu_i = \nu_i$ für alle $i = 1, \dots, n$.

(ii) Für V existiert eine Eigenbasis von f.

Beweis: Sei λ ein Eigenwert. Sei $v_1, \dots, v_m$ eine Basis von V_λ. Wir ergänzen durch $v_{m+1}, \dots, v_N$ zu einer Basis von V. Die darstellende Matrix von f hat in dieser Basis die Form

$$\begin{pmatrix} \lambda E_m & A \\ 0 & B \end{pmatrix}.$$

Wir lesen das charakteristische Polynom ab:

$$\chi_f = (\lambda - X)^m \chi_B.$$

Hieraus folgt $m = \mu(\lambda) \leq \nu(\lambda)$.
Sei nun $v_1, \ldots, v_N$ eine Eigenbasis. Dann ist die darstellende Matrix diagonal. Für jeden Eigenwert λ_i gibt es mindestens soviele Elemente in dieser Basis wie ν_i angibt, also $\dim V_{\lambda_i} \geq \nu_i$.
Für die Rückrichtung zerfalle das charakteristische Polynom in Linearfaktoren, also $N = \sum_{i=1}^n \nu_i$. Für $i = 1, \ldots, n$ wählen wir eine Basis $v_1^i, \ldots, v_{\nu_i}^i$ von V_{λ_i}.

Behauptung. *Das System der (v_j^i) für $i = 1, \ldots, n$, $j = 1, \ldots, \nu_i$ ist linear unabhängig.*

Sei

$$0 = \sum_{i,j} a_{ij} v_j^i.$$

Dann ist $x^i = \sum_j a_{ij} v_j^i \in V_{\lambda_i}$ und

$$0 = \sum_{i=1}^n x^i.$$

Nach LA 1 Satz 7.18 sind Elemente aus verschiedenen Eigenräumen linear unabhängi, wenn sie nicht Null sind. Es folgt also $x^i = 0$ für alle i. Aus der linearen Unabhängigkeit der v_j^i für festes i folgt dann $a_{ij} = 0$ für alle j und i. Damit ist die Zwischenbehauptung bewiesen.
Es folgt

$$N = \dim V \geq \dim \langle v_j^i | i, j \rangle = \sum_i \nu_i = \sum_i \mu_i = N.$$

Hieraus folgt, dass die linear unabhängige Familie auch ein Erzeugendensystem ist. Nach Konstruktion handelt es sich um eine Eigenbasis. □

Trigonalisierbarkeit

Sei nun k algebraisch abgeschlossen. Wir haben gesehen, dass nicht jede Matrix diagonalisierbar ist, wohl aber stets *mindestens einen* Eigenvektor hat. Daraus lässt sich mehr machen!

Satz 12.6. *Sei V endlich-dimensionaler Vektorraum über einem algebraisch abgeschlossenen Körper k, $f \in \mathrm{End}(V)$. Dann hat V eine Basis, so dass die darstellende Matrix eine obere Dreiecksmatrix ist*

$$M_B(f) = (a_{ij})_{i,j} \quad \text{mit } a_{ij} = 0 \text{ für } i > j.$$

Zum Beweis fassen wir die Eigenschaft konzeptioneller.

Definition 12.7. *Sei V ein Vektorraum. Eine* Fahne der Länge r *von V ist eine echt aufsteigende Folge*

$$V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_r$$

von Untervektorräumen. Die Fahne heißt vollständig, *falls $V_0 = 0$, $V_r = V$ und* $\dim V_{i+1} = \dim V_i + 1$.

Sei $f : V \to V$ ein Endomorphismus. Ein Unterraum W heißt f-invariant, *wenn $f(W) \subset W$. Eine Fahne heißt* f-invariant, *falls alle V_i f-invariant sind.*

Beispiel. (i) $0 \subset V$ ist eine (i.a. unvollständige) Fahne der Länge 1 (für $V \neq 0$). Sie ist immer f-invariant.

(ii) Sei $v_1, \dots, v_n$ eine Basis von V. Dann ist

$$0 \subsetneq \langle v_1 \rangle \subsetneq \langle v_1, v_2 \rangle \subsetneq \cdots \subsetneq \langle v_1, \dots, v_n \rangle = V$$

eine vollständige Fahne.

(iii) Für jedes $\lambda \in k$ ist der Eigenraum V_λ ein f-invarianter Unterraum.

Satz 12.8. *Sei V endlich-dimensionaler Vektorraum über einem algebraisch abgeschlossenen Körper, $f \in \mathrm{End}(V)$. Dann existiert eine vollständige f-invariante Fahne.*

Beweis: Wir argumentieren mit vollständiger Induktion nach $n = \dim V$. Für $n = 1$ ist nichts zu zeigen.

Sei nun $n > 1$ und die Aussage wahr für alle Vektorräume der Dimension $n - 1$. Da k algebraisch abgeschlossen ist, gibt es einen Eigenvektor $v_1 \in V$. Sei $W = \langle v_1 \rangle \subset V$. Dieser Unterraum ist f-invariant. Wir betrachten den Quotientenvektorraum

$$V' = V/W = \{v + W | v \in V\}.$$

Es gilt nach der Dimensionsformel $\dim V' = \dim V - \dim W = n - 1$. Wir betrachten

$$f' : V' \to V', \quad v + W \mapsto f(v) + W$$

Behauptung. *Die lineare Abbildung f' ist wohldefiniert.*

Wir können das mit Repräsentanten nachrechnen. Oder wir benutzen den Homomorphiesatz LA 1 Satz 4.14: Wir betrachten die Komposition

$$V \xrightarrow{f} V \xrightarrow{\pi} V/W$$

und überprüfen, dass $W = \langle v_1 \rangle$ im Kern von $\pi \circ f$ liegt. Es gilt

$$\pi(f(v_1)) = \pi(\lambda v_1) = \lambda \pi(v_1) = 0 + W.$$

Nach dem Homomorphiesatz faktorisiert $\pi \circ f$ eindeutig über f'.

Wir können nun die Induktionsvoraussetzung auf den Vektorraum V' und den Endomorphismus f' anwenden. Es existiert also eine vollständige Fahne

$$0 = V_1' \subset V_2' \subset \cdots \subset V_n' = V'.$$

Wir setzen für $i = 1, \ldots, n$

$$V_i = \pi^{-1}(V_i')$$

und $V_0 = 0$. Im Fall $i = 1$ erhalten wir gerade $V_1 = \mathrm{Ker}(\pi) = W$. Im Fall $i = n$ erhalten wir ganz V. Insgesamt

$$0 \subset V_1 \subset V_2 \subset \cdots \subset V_n = V.$$

Behauptung. $\dim V_i = i$.

Die Abbildungen $V_i \to V_i'$ sind jeweils surjektiv. Nach der Dimensionsformel gilt

$$\dim V_i = \dim \mathrm{Ker}(\pi) + \dim V_i' = 1 + (i - 1) = i,$$

nach Wahl der V_i' und wegen $\mathrm{Ker}(\pi) = W = \langle v_1 \rangle$. Es handelt sich also um eine vollständige Fahne.

Behauptung. $f(V_i) \subset V_i$ *für* $i = 0, \ldots, n$

Für $i = 0$ ist nichts zu zeigen. Sei $i > 0$, also $V_i = \pi^{-1}(V_i')$. Sei $x \in V_i$. Wir betrachten $f(x)$. Es gilt

$$\pi(f(x)) = f'\pi(x) \in V_i'$$

denn nach Induktionsvoraussetzung ist V_i' invariant unter f'. Jetzt ist $f(x) \in \pi^{-1}(V_i')$, wie gewünscht.
Dies beendet den Beweis. □

Bemerkung. Etwas genaueres Hinschauen zeigt: es genügt die Voraussetzung, das χ_f über k in Linearfaktoren zerfällt.

Beweis von Satz 12.6. Sei

$$0 = V_0 \subset V_1 \subset \cdots \subset V_n$$

eine vollständige f-invariante Fahne. Wir ergänzen rekursiv eine Basis $v_1, \ldots, v_i$ von V_i zu einer Basis $v_1, \ldots, v_{i+1}$ von V_{i+1}. Dann gilt

$$f(v_i) \in \langle v_1, \ldots, v_i \rangle.$$

Die darstellende Matrix hat die gewünschte Form. □

Beispiel. Sei $f : k^2 \to k^2$ die Multiplikation mit $\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$. Wir bestimmen einen Eigenvektor:

$$\chi_f = \det \begin{pmatrix} 1-X & 1 \\ -1 & 3-X \end{pmatrix} = (1-X)(3-X)+1 = X^2 - 4X + 4.$$

Der Eigenwert ist 2. Wir lösen die Eigenwertgleichung $f(v) - 2v = 0$.

$$0 = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x+y \\ -x+y \end{pmatrix}.$$

Wir erhalten $x = y$ und den Eigenvektor $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Die volle Fahne ist

$$0 \subset \langle v_1 \subset k^2 = \langle v_1, e_1 \rangle.$$

Die Bilder der Basisvektoren unter f sind

$$f(v_1) = 2v_1, f(e_1) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -1v_1 + 2e_1.$$

Die darstellende Matrix ist

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

Bemerkung. In oberen Dreiecksmatrizen können wir Determinante, Eigenwerte und charakteristisches Polynom leicht ablesen.

Der Satz von Cayley-Hamilton

Sei V ein Vektorraum. Dann ist $\mathrm{End}(V)$ ein Ring bezüglich der Addition und Komposition von Endomorphismen. Wir schreiben für $f \in \mathrm{End}(V)$ und $n \geq 0$

$$f^0 = \mathrm{id}, f^1 = f, f^{n+1} = f \circ f^n.$$

Ist $P \in k[X]$, also $P = \sum_{i=0}^n a_i X^i$, so können wir nicht Elemente aus k, sondern auch aus $\mathrm{End}(V)$ einsetzen

$$P(f) = \sum_{i=0}^n a_i f^i \in \mathrm{End}(V).$$

Die darstellende Matrix von $P(f)$ ist $P(M(f))$.

Beispiel. Sei $f \in \mathrm{End}(k^2)$ die Multiplikation mit der Matrix

$$M = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \lambda E_2 + N$$

mit $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Beachte, dass $N^2 = 0$. Dann gilt

$$M^2 = (\lambda E_2 + N)^2 = \lambda^2 E_2 + 2\lambda N = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}$$

$$M^n = \left((\lambda E_2 + N)^n = \sum_{i=0}^{n} \binom{n}{i} \lambda^i N^{n-i} \right) = \lambda^n E_2 + n\lambda^{n-1} N = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}$$

Diese Matrizen sind linear abhängig! Es gilt einfach

$$(M - \lambda E_2)^2 = N^2 = 0$$

d.h. M ist Nullstelle von

$$P = X^2 - 2\lambda X + \lambda^2 = (X - \lambda)^2.$$

Dies ist das charakteristische Polyom von M.

Theorem 12.9 (Satz von Cayley-Hamilton). *Sei k Körper, V endlich-dimensionaler k-Vektorraum, $f \in \mathrm{End}(V)$. Dann gilt*

$$\chi_f(f) = 0 \in \mathrm{End}(V).$$

Beweis: Sei zunächst k algebraisch abgeschlossen. Wir wählen eine Basis $v_1, \dots, v_n$, so dass die darstellende Matrix obere Dreieckstgestalt hat mit Diagonale $(\lambda_1, \dots, \lambda_n)$. Es gilt

$$\chi_f = (\lambda_1 - X)(\lambda_2 - X) \dots (\lambda_n - X).$$

Wir müssen also

$$\chi_f(f) = (\lambda_1 \,\mathrm{id} - f)(\lambda_2 - f) \dots (\lambda_n \,\mathrm{id} - f)$$

berechnen. Wir wollen zeigen, dass $\chi_f(f)(V) = 0$, d.h. alle Vektoren werden auf 0 abgebildet.
Sei

$$\Phi_i = (\lambda_1 \,\mathrm{id} - f)(\lambda_2 - f) \dots (\lambda_i \,\mathrm{id} - f)$$

Sei $V_i = \langle v_1, \dots, v_i \rangle$ die vollständige f-invariante Fahne, die zu unserer Basis gehört.

Behauptung. $\Phi_i(V_i) = 0$

Wir argumentieren mit vollständiger Induktion nach i. Für $i = 0$ ist nichts zu zeigen. Sei nun $i > 0$ und die Aussage wahr für Vektorräume der Dimension $i - 1$. Wir betrachten einen beliebigen Vektor $v \in V_i$. Es gilt

$$v = a_1 v_1 + \dots a_i v_i$$

mit $a_i \in k$. Es gilt

$$\begin{aligned}\Phi_i(v) &= \Phi_{i-1} \circ (\lambda_i \operatorname{id} - f)(v) \\ &= \Phi_{i-1}(\lambda_i a_1 v_1 + \lambda_i a_2 v_2 + \dots \lambda_i a_i v_i - a_1 f(v_1) - a_2 f(v_2) - \dots - a_i f(v_i))\end{aligned}$$

Fast alle Summanden liegen in V_{i-1}, da V_{i-1} stabil unter f. Einzige eventuelle Ausnahmen ist $\lambda_i a_i v_i - a_i f(v_i)$. Nach Definition gilt

$$f(v_i) = \lambda_i v_i + v' \text{ mit } v' \in V_{i-1}.$$

Also ist ebenfalls

$$\lambda_i a_i v_i - a_i f(v_i) = v' \in V_{i-1}.$$

Also ist das Argument von Φ_{i-1} in V_{i-1}. Nach Induktionsvoraussetzung gilt $\Phi_{i-1}(V_{i-1}) = 0$, also auch $\Phi_i(v) = 0$. Dies beweist die Behauptung. Für $i = n$ erhalten wir die Aussage des Theorems.

Sei nun k ein beliebiger Körper. Sei M darstellende Matrix bezüglich einer beliebigen Basis von V. Zu zeigen ist $\chi_f(M) = 0$. Jeder Körper k ist ein einem algebraische abgeschlossenen Körper $\overline{k}$ enthalten (Vorlesung Algebra und Zahlentheorie). Es gilt $M_n(k) \subset M_n(\overline{k})$, daher genügt es zu zeigen, dass $\chi_f(M) = 0$ in $M_n(\overline{k})$. Das Polynom χ_f ist das charakteristische Polynom der linearen Abbildung $\overline{f} : \overline{k}^n \to \overline{k}^n$, die durch Multiplikation mit M gegeben ist. Nach dem ersten Fall folgt $\chi_f(\overline{f}) = \chi_f(M) = 0$. □

Wir können jetzt die Matrizen λE_2 und $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ auseinander halten: beide sind Nullstellen des charakteristischen Polynoms $(\lambda - X)^2$. Die erste aber sogar schon von $(\lambda - X)$.

Definition 12.10. *Sei V endlich-dimensionaler Vektorraum, $f \in \operatorname{End}(V)$. Das* Minimalpolynom *von f*

$$\mu_f \in k[X]$$

ist das normierte Polynom kleinsten Grades mit $\mu_f(f) = 0$.

Ein Polynom heißt *normiert*, wenn der höchste Koeffizient 1 ist, also

$$P = X^n + a_{n-1}X^n + \dots + a_0.$$

Normierte Polynome sind immer ungleich 0.

Der Satz von Cayley-Hamilton garantiert die Existenz eines solchen Polynoms. Genauer:

Lemma 12.11. *Sei $P \in k[X]$ ein Polynom mit $P(f) = 0$. Dann ist μ_f ein Teiler von P. Das Minimalpolynom μ_f ist eindeutig bestimmt.*

Beweis: Polynomdivison mit Rest ergibt

$$P = Q\mu_f + R \text{ mit } \deg R < \deg \mu_f$$

Einsetzen von f ergibt

$$0 = P(f) = Q(f)\mu_f(f) + R(f) = R(f).$$

Sei mit $\deg(R) = m$. Angenommen, $R \neq 0$, also $m \geq 0$.

$$R = b_m X^m + \dots b_m.$$

Dann ist $b_m \neq 0$, also $b_m^{-1}R$ ein normiertes Polynom mit Nullstelle f. Es hat einen kleineren Grad als μ_f. Dies ist ein Widerspruch zur Wahl von μ_f, also gilt $R = 0$. Dies beweist die Teilbarkeitsaussage.
Seien μ_f und μ'_f zwei Minimalpolynome. Dann gilt

$$\mu_f = A\mu'_f, \quad \mu'_f = B\mu_f = BA\mu_f.$$

Hieraus folgt, dass $\deg BA = 0$, also B und A beide konstant. Wir vergleichen den höchsten Koeffizienten. Er ist für μ_f und für μ'_f jeweils 1, also ist $A = B = 1$. Die beiden Polynome sind gleich. □

Satz 12.12. *Sei V endlich-dimensionaler k-Vektorraum, $f \in \mathrm{End}(V)$. Dann ist das Minimalpolynom μ_f ein Teiler des charakteristischen Polynoms χ_f. Die Nullstellen stimmen überein.*

Beweis: Nach dem Satz von Cayley-Hamilton ist $\chi_f(f) = 0$. Nach Lemma 12.11 folgt $\chi_f = Q\mu_f$ für ein Polynom Q. Insbesondere ist jede Nullstelle von μ_f eine Nullstelle von χ_f.
Sei umgekehrt λ eine Nullstelle von χ_f, also ein Eigenwert. Sei v der zugehörige Eigenvektor. Wir berechnen $\mu_f(f)(v)$. Sei dafür $\mu_f = X^d + b_{d-1}X^{d-1} + \dots + b_0$. Es folgt

$$\begin{aligned}
\mu_f(f)(v) &= (f^d + b_{d-1}f^{d-1} + \dots + b_0\,\mathrm{id})(v)\\
&= f^d(v) + b_{d-1}f^{d-1}(v) + \dots + b_0 v\\
&= \lambda^d v + b_{d-1}\lambda^{d-1}v + \dots b_0 v\\
&= \mu_f(\lambda)v.
\end{aligned}$$

denn $f^i(v) = \lambda^i v$ (Induktion). Es ist $\mu_f(f) = 0 \in \mathrm{End}(V)$, also $\mu_f(f)(v) = 0$. Wir haben daher

$$0 = \mu_f(\lambda)v$$

für ein $v \neq 0$. Hieraus folgt $\mu_f(\lambda) = 0$. Wie behauptet ist λ auch eine Nullstelle von μ_f. □

Beispiel. Sei $M = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ für $a, b \in k$. Dann ist

$$\chi_M = (a - X)(b - X) = X^2 - (a + b)X + ab.$$

Einsetzen von M ergibt

$$\chi_M(M) = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix} - (a+b)\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} + ab\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} a^2 - a^2 - ab + ab & 0 \\ 0 & b^2 - ab - b^2 + ab \end{pmatrix} = 0.$$

Die möglichen Teiler von χ_M sind $(X-a)$ und $(X-b)$. Wir setzen M ein

$$(X-a)(M) = \begin{pmatrix} a-a & 0 \\ 0 & b-a \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & b-a \end{pmatrix}.$$

Dies ist 0 für $a = b$! Es gilt also

$$\mu_M = \begin{cases} X - a & a = b \\ \chi_M & a \neq b \end{cases}.$$

Beispiel. Für $M = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ ist $\chi_M = (X-\lambda)^2$. In diesem Fall ist M nicht Nullstelle von $X - \lambda$, also $\mu_M = \chi_M$.

Satz 12.13. *Sei V endlich-dimensionaler Vektorraum, $f \in \mathrm{End}(V)$. Dann ist f genau dann diagonalisierbar, wenn μ_f in Linearfaktoren zerfällt und die Nullstellen Multiplizität 1 haben.*

Beweis: Sei $v_1, \dots, v_n$ eine Eigenbasis. Dann zerfallen χ_f und sein Teiler μ_f in Linearfaktoren. Sei

$$\mu_f = (X - \lambda_1)^{\nu_1} \dots (X - \lambda_d)^{\nu_d}$$

mit paarweise verschiedenen $\lambda_1, \dots, \lambda_d$. Setze

$$\mu'_f = (X - \lambda_1) \dots (X - \lambda_d).$$

Wir setzen v_i in $\mu'_f(f)$ ein. Dies ist ein Eigenvektor. Sei $\lambda_{j(i)}$ der zugehörige Eigenwert. Es folgt wie im letzten Beweis

$$\mu'_f(f)(v_i) = \mu'_f(\lambda_{j(i)})v_i = 0.$$

Dies gilt für alle v_i in einer Basis, also $\mu'_f(f)(V) = 0$. Mit anderen Worten, $\mu'_f(f) = 0$. Nach Definition des Minimalpolynoms folgt $\mu_f = \mu'_f$.
Den Beweis der Rückrichtung verschieben wir auf später. Er fällt als Nebenprodukt der Hauptraumzerlegung ab.

□

Kapitel 13

Die Jordansche Normalform

Wir formulieren gleich zu Anfang unser Ziel.

Definition 13.1. *Sei k Körper, $\lambda \in k$, $n \geq 1$. Die Matrix*

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots & & & 0 & \lambda \end{pmatrix}$$

heißt Jordan-Block der Länge n zum Eigenwert λ..

Theorem 13.2 (Jordansche Normalform)**.** *Sei k algebraisch abgeschlossen, V endlich dimensionaler Vektorraum $f \in \mathrm{End}(V)$. Dann gibt es eine Basis von V, so dass die darstellende Matrix die Form*

$$\begin{pmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \dots & \\ & & & J_{n_m}(\lambda_m) \end{pmatrix}$$

hat. Hierbei sind die λ_i und n_i eindeutig bis auf Reihenfolge.

Bemerkung. (i) Der Endormorphismus ist diagonalisierbar, wenn alle Jordanblöcke die Größe 1 haben.

(ii) An der Jordanschen Normalform liest man leicht das charakteristische Polynom ab, aber auch das Minimalpolynom ab. Es ist

$$\chi_{J_n(\lambda)} = \mu_{J_n(\lambda)} = (\lambda - X)^n$$

und für M in Jordan-Normalform

$$\chi_M = \prod_{i=1}^{n} \chi_{J_{n_i}(\lambda_i)} = \prod (\lambda_i - X)^{n_i}, \mu_M = \mathrm{kgv}(\lambda_i - X)^{n_i}.$$

Insbesondere lesen wir das Diagonalisierbarkeitskriterium 12.13 aus dem letzten Kapitel ab. Die algebraische Vielfachheit von λ erhalten wir durch Aufsummieren der n_i mit $\lambda_i = \lambda$. Die geometrische Vielfachheit ist die Anzahl der Jordanblöcke zum Eigenwert λ. Dies liefert das erste Diagonalisierbarkeitskriterium 12.5.

Der Beweis wird uns eine Weile beschäftigen. Wir gehen in zwei Schritten vor:

(i) Zerlegung von V in die Anteile, die zu einem λ gehören.

(ii) Auf einem solchen Raum betrachten wir $f - \lambda\,\mathrm{id}$. Diese Abbildung wird nilpotent sein, d.h. wir müssen darstellende Matrizen für nilpotente Abbildungen verstehen.

Der Schlüssel sind also die Abbildungen $f - \lambda\,\mathrm{id}$ und deren Potenzen.

Hauptraumzerlegung

Definition 13.3. *Sei k Körper, V ein k-Vektorraum, $\lambda \in k$. Dann heißt*

$$H_\lambda = \{v \in V | \textit{es gibt } n \geq 1 \textit{ mit } (f-\lambda)^n(v) = 0\}$$

Hauptraum *oder* verallgemeinerter Eigenraum *zum Eigenwert λ.*

Bemerkung. Es gilt $V_\lambda \subset H_\lambda$ (mit $n = 1$). Wenn $H_\lambda \neq 0$, so ist λ ein Eigenwert, denn für minimales n mit $(f-\lambda)^n(v) = 0$ folgt $(f-\lambda)^{n-1}(v) \in V_\lambda$.

Lemma 13.4. *H_λ ist ein Untervektorraum. Ist* $\dim V < \infty$*, so gibt es $n \geq 1$ mit $H_\lambda = \mathrm{Ker}(f-\lambda)^n$.*

Beweis: Nach Definition ist

$$H_\lambda = \bigcup_{n=1}^{\infty} \mathrm{Ker}(f - \lambda\,\mathrm{id})^n).$$

Die Kerne sind ineinander enthalten:

$$\mathrm{Ker}(f - \lambda\,\mathrm{id}) \subset \mathrm{Ker}(f - \lambda\,\mathrm{id})^2 \subset \dots.$$

Jedes $\mathrm{Ker}(f-\lambda)^n$ ist abgeschlossen unter skalarer Multiplikation, also auch die Vereinigung. Für $v_i \in \mathrm{Ker}(f-\lambda)^{n_i}$ mit $i = 1, 2$ sind beide enthalten in $\mathrm{Ker}(f-\lambda)^n$ für $n \geq n_1, n_2$, also auch ihre Summe.
Ist $\dim V < \infty$, so muss die Kette von Untervektorräumen stabil werden. □

Lemma 13.5. *Einziger Eigenwert von $f|_{H_\lambda}$ ist λ.*

Beweis: Sei $v \in H_\lambda$ Eigenvektor zum Eigenwert μ. Dann gilt rekursiv

$$(f - \lambda\,\mathrm{id})^n(v) = (\mu - \lambda)^n v.$$

Für genügend großes n ist also $(\mu - \lambda)^n = 0$, und dann auch $\mu - \lambda = 0$. □

Lemma 13.6. *H_λ ist f-invariant.*

Beweis: Sei $v \in H_\lambda$, $(f - \lambda)^n(v) = 0$. Wir betrachten jetzt $f(v)$. Wir wissen, dass $f(v) - \lambda v \in H_\lambda$, denn dieses Element wird von $(f - \lambda)^{n-1}$ annihiliert. Nach Voraussetzung ist $v \in H_\lambda$. Dies ist ein Vektorraum, also liegen auch λv und $(f(v) - \lambda v) + \lambda v$ in H_λ. □

Satz 13.7. *Sei k algebraisch abgeschlossen, V endlich-dimensionaler Vektorraum, $f \in \mathrm{End}(V)$ und λ ein Eigenwert. Dann ist H_λ ein f-invarianter Teilraum mit Dimension gleich der algebraischen Vielfachheit $\nu(\lambda)$ von λ. Das charakteristische Polynom von $f|_{H_\lambda}$ ist $(\lambda - X)^{\nu(\lambda)}$.*

Beweis: Wir wählen eine Basis $v_1, \dots, v_m$ von H_λ und ergänzen zu einer Basis von V. Da H_λ invariant unter f ist, erhalten wir eine darstellende Matrix der Form

$$\begin{pmatrix} A & * \\ 0 & B \end{pmatrix}$$

mit $A \in M_m(k)$, $B \in M_{\dim V - m}(k)$. Es gilt

$$\chi_f = \chi_A \chi_B.$$

Einziger Eigenwert von A ist λ, also muss gelten $\chi_A = (\lambda - X)^m$. Angenommen, λ ist eine Nullstelle von χ_B. Dies ist charakteristische Polynom des von f auf V/H_λ induzierten Endormorphismus. Zu dieser Nullstelle gehört ein Eigenvektor $\overline{v}$. Sei $v \in V$ ein Urbild. Dann gilt

$$(f - \lambda\,\mathrm{id})(v) \in H_\lambda$$

also gibt es $n \geq 1$ mit

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

Also gilt $v \in H_\lambda$. Dies ist ein Widersproch zu $\overline{v} \neq 0$. Damit haben wir gezeigt, dass λ keine Nullstelle von χ_B ist. Die Nullstellenvielfachheit von λ in χ_f ist genau m. □

Satz 13.8 (Hauptraum-Zerlegung)**.** *Sei k algebraisch abgeschlossen, V endlich-dimensionaler k-Vektorraum, $f \in \mathrm{End}(V)$. Seien $\lambda_1, \dots, \lambda_r$ die Eigenwerte von f. Dann ist die natürliche Abbildung*

$$H_{\lambda_1} \oplus H_{\lambda_2} \oplus \cdots \oplus H_{\lambda_r} \to V$$

ein Isomorphismus.

Beweis: Nach Satz 13.7 gilt

$$\dim H_{\lambda_1} + \cdots + \dim H_{\lambda_r} = \nu(\lambda_1) + \ldots \nu(\lambda)_r) = \deg \chi_f = \dim V.$$

Es genügt daher die Injektivität der natürlichen Abbildung zu zeigen.
Seien $v_i \in H_{\lambda_i}$ für $i = 1, \ldots, r$ mit $v_1 + \cdots + v_r = 0$.

Behauptung. *Dann gilt $v_i = 0$ für $i = 1, \ldots, r$.*

Sei s die Anzahl der $v_i \neq 0$. Ist $s = 1$, so gilt die Behauptung automatisch. Sei also jetzt $s > 1$. Wir wählen $v_1, \ldots, v_r$ so, dass s minimal. Nach Umnummerieren ist $v_1, \ldots, v_s \neq 0$, $v_{s+1}, \ldots, v_r = 0$. Sei $g = (f - \lambda_1)^{\nu(\lambda_1)}$. Alle H_{λ_i} sind g-invariant, weil sie f-invariant sind. Es gilt

$$0 = g(v_1 + \cdots + v_s) = g(v_1) + g(v_2) + \cdots + g(v_s), \quad g(v_i) \in H_{\lambda_i}.$$

Dies ist wieder eine Relation der Art, die wir betrachten. Wir betrachten $g(v_1)$. Wegen $v_1 \in H_{\lambda_1}$ gilt $(f - \lambda_1 \operatorname{id})^{\nu(\lambda_i)}(v_1) = 0$, also $g(v_1) = 0$. Die Relation hat weniger als s Einträge. Nach Wahl von s bedeutet das

$$g(v_2) = g(v_3) = \cdots = g(v_s) = 0.$$

Behauptung. *g ist injektiv auf H_λ mit $\lambda \neq \lambda_1$.*

Sei $v \in \operatorname{Ker}(g) \cap H_\lambda$. Sei N minimal mit $(f - \lambda \operatorname{id})^N(v) = 0$. Dann ist $v' = (f - \lambda \operatorname{id})^{N-1}$ ein Eigenvektor zum Eigenwert λ und

$$g(v') = (f - \lambda_1 \operatorname{id})^{\nu(\lambda_1)}(f - \lambda \operatorname{id})^{N-1}(v) = (f - \lambda \operatorname{id})^{N-1}(f - \lambda_1)^{\nu(\lambda_1)}(v) = 0.$$

Für einen Eigenvektor ist aber

$$= g(v') = (f - \lambda_1)^{\nu(\lambda_1)}(v') = (\lambda - \lambda_1)^{\nu(\lambda_1)} v' \Rightarrow \lambda = \lambda'.$$

Damit die Aussage über die Injektivität bewiesen. Es folgt

$$v_2 = \cdots = v_r = 0,$$

wie zu beweisen war. □

Wir können uns jetzt also auf Haupträume konzentrieren. Sei also $V = H_\lambda$. Wir betrachten $g = f - \lambda \operatorname{id}$. Diese Abbildung ist jetzt nilpotent.

Nilpotente Endormorphismen

Wir wollen die Jordansche Normalform für nilpotente Endormorphismen herleiten. Die Jordanblöcke haben die Form

$$J_n(0) = \begin{pmatrix} 0 & 1 & 0 & & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \\ 0 & \ldots & & & 0 & 0 \end{pmatrix}$$

Alle Einträge sind 0, nur die erste Nebendiagonale hat die Einträge 1. Als lineare Abbildung ist dies

$$e_1 \mapsto 0, e_2 \mapsto e_1, \ldots, e_n \mapsto e_{n-1}.$$

Definition 13.9. *Sei V ein endlich-dimensionaler Vektorraum, $f \in \mathrm{End}(V)$. Eine* f-zyklische Basis *von V ist eine Basis der Form*

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

Der Vektor v heißt dann Hauptvektor. *Der Vektorraum heißt* f-zyklisch, *wenn eine f-zyklische Basis existiert. Ein Untervektorraum $W \subset V$ heißt* f-zyklisch, *wenn er f-invariant und $f|_W$-zyklisch. Die f-Ordnung von $v' \in V$ ist die kleinste Zahl $i \in \mathbb{N}_0$, so dass $f^i(v') = 0$.*

Beispiel. (i) $f = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Dann ist e_2 ein Hauptvektor, da $f(e_2) = e_1$, also $f(e_2), e_2$ eine Basis.

(ii) Sei $g = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Wieder ist e_2 ein Hauptvektor, da $g(e_2) = e_1 + e_2$ und wieder $g(e_2), e_2$ eine Basis.

(iii) Sei $h = \lambda E_2$. Dann gibt es keinen Hauptvektor. Der Vektorraum ist nicht h-zyklisch. Jeder eindimensionale Unterraum ist zyklisch.

Bemerkung. Hat $f : V \to V$ als darstellende Matrix den Jordanblock $J_n(0)$, so haben wir eine f-zyklische Basis gefunden. Ist umgekehrt f nilpotent und $f^{n-1}(v), \ldots, v$ eine f-zyklische Basis, so ist die darstellende Matrix der Jordanblock $J_n(0)$.

Im nilpotenten Fall gibt es viele zyklische Unterräume.

Lemma 13.10. *Sei f nilpotent, $v \in V$ mit $v \neq 0$. Dann ist v Hauptvektor des zyklischen Teilraums*

$$U(v) = \langle v, f(v), f^2(v), \ldots \rangle.$$

Es gilt $U(v) \cap \mathrm{Ker}(f) = f^{n-1}(v)$, wobei n die f Ordnung von v ist.

Beweis: Sei n minimal mit $f^n(v) = 0$. Wir zeigen, dass

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

linear unabhängig sind. Dies ist dann die gesuchte zyklische Basis. Seien $a_0, \ldots, a_{n-1} \in k$ mit

$$a_0 v + a_1 f(v) + \ldots a_{n-1} f^{n-1}(v) = 0.$$

Durch Anwenden der Potenzen von f erhalten wir die neuen Relationen:

$$\begin{aligned} a_0 f(v) + a_1 f^2(v) + \cdots + a_{n-2} f^{n-1}(v) &= 0 \\ a_0 f^2(v) + a_1 f^3(v) + \ldots a_{n-3} f^{n-1}(v) &= 0 \\ &\ldots \\ a_0 f^{n-1}(v) &= 0 \end{aligned}$$

Aus der letzten Relation folgt $a_0 = 0$, dann rückwärts $a_1 = 0$ bis $a_{n-1} = 0$.
Sei $w = b_0 v + b_1 f(v) + \cdots + b_{n-1} f^{n-1}(v)$ im Kern von f, also

$$0 = f(w) = b_0 f(v) + b_1 f^2(v) + \cdots + b_{n-2} f^{n-1}(v) + 0.$$

Wegen der linearen Unabhängigkeit der $f^i(v)$ folgt $b_0 = b_1 = \ldots b_{n-2} = 0$. □

Satz 13.11 (Existenz von zyklischen Basen)**.** *Sei V endlich-dimensionaler Vektorraum, $f \in \mathrm{End}(V)$ nilpotent. Dann gibt es f-zyklische Teilräume $V_1, \ldots, V_r$ so dass die natürliche Abbildung*

$$V_1 \oplus \cdots \oplus V_r \to V$$

ein Isomorphismus ist. Dabei ist r eindeutig. Die Dimensionen der V_i sind eindeutig bis auf Reihenfolge.

Beweis: Es gilt $V = H_0$. Nach Lemma 13.4 gibt es $n \geq 1$ mit $f^n = 0$. Wir argumentieren mit vollständiger Induktion nach dem kleinesten n mit $f^n = 0$.
Für $n = 1$ ist f die Nullabbildung. Die zyklischen Unterräume sind 1-dimensional.
Sei nun $n \geq 2$. Wir setzen voraus, dass die Aussage für Abbildungen der Ordnung $n - 1$ gilt. Wir betrachten $W = f(V)$. Dieser Teilraum ist f-stabil und f hat Ordnung $n - 1$. Sei

$$W \cong W_1 \oplus \cdots \oplus W_s$$

eine Zerlegung in zyklische Teilräume und w_i Hauptvektor von W_i. Nach Definition gibt es $v_i \in V$ mit $f(v_i) = w_i$. Seien $V_1, \ldots, V_s$ die von ihnen erzeugten zyklischen Teilräume.

Behauptung. *Die natürliche Abbildung $V_1 \oplus V_2 \oplus \cdots \oplus V_s \to V$ ist injektiv.*

Sei $(x_1, \ldots, x_s)$ im Kern, also $x_i \in V_i$ für $i = 1, \ldots, s$,

$$x_1 + \cdots + x_s = 0.$$

Wir wenden f auf die Relation an und erhalten

$$0 + f(x_1) + \ldots f(x_n) = 0$$

mit $f(x_i) \in W_i$. Wir waren mit einer Zerlegung von $\mathrm{Im}(f)$ gestartet, also folgt $f(x_i) = 0$ für $i = 1, \ldots, s$.
Wir haben nun $x_i \in V_i \cap \mathrm{Ker}(f)$, also ist x_i ein Vielfaches des Basisvektors $f^{n_i - 1}(v_i)$, der nach Konstruktion in W_i liegt. Wieder wegen der Unabhängikeit der W_i folgt dann auch $x_i = 0$ für alle i.

Behauptung. $\mathrm{Ker}(f) \oplus V_1 \oplus \ldots V_s \to V$ *ist surjektiv.*

Sei $x \in V$, also $f(x) \in W$. Dann gibt es $y_i \in W_i$, so dass

$$f(v) = y_1 + \cdots + y_s.$$

Die Abbildung $f : V_i \to W_i$ ist surjektiv, also gibt es $x_i \in V_i$ mit $f(x_i) = y_i$. Die Differenz

$$x - \sum_{i=1}^{s} x_i$$

liegt im Kern von f.

Wir können nun die zyklischen Basen der V_i durch linear unabhängige Elemente des Kerns von f zu einer Basis von V ergänzen. Dies ist die gesuchte zyklische Basis.

Zu zeigen bleibt noch die Eindeutigkeitsaussage. Es ist

$$r = \dim \mathrm{Ker}(f)$$

unabhängig von der Zerlegung. Weiter ist

$$\dim \mathrm{Ker}(f^2) = \text{Anzahl der } V_i \text{ mit } \dim V_i \geq 2$$

oder allgemein

$$\dim \mathrm{Ker}(f^n) = \text{Anzahl der } V_i \text{ mit } \dim V_i \geq n.$$

Insbesondere sind die Dimensionen also unabhängig von der konkreten Zerlegung. □

Beweis von Theorem 13.2. Nach Satz 13.8 existiert eine Zerlegung

$$V \cong \bigoplus H_{\lambda_i}$$

in die Haupträume zu den verschiedenen Eigenwerten. Nach Satz 13.11 existiert eine Zerlegung von H_{λ_i} in $f - \lambda_i$ id-zyklische Teilräume. In dieser Basis ist die darstellende Matrix ein Jordan-Block.

Die Eindeutigkeit folgt, weil die Zerlegung in Haupträume unabhängig von Basiswahlen ist und aus der Eindeutigkeit in Satz 13.11. □

Bemerkung. Wir haben die Informationen des letzten Kapitels nicht benutzt. Die Sätze folgen umgekehrt aus der Jordan-Zerlegung.

Konsequenzen

Jeder Jordan-Block hat die Form

$$J_m(\lambda) = \lambda E_m + N$$

mit N nilpotent. Die beiden Summanden kommutieren. Allgemeiner erhalten wir für jede Matrix in Jordan-Normalform eine Zerlegung

$$M = M_s + M_n$$

mit M_s eine Diagonalmatrix und M_n nilpotent, die beiden Summanden kommutieren.

Satz 13.12 (Additive Jordan-Zerlegung)**.** *Sei k algebraisch abgeschlossen, $M \in M_N(k)$. Dann gibt es eine Zerlegung*

$$M = M_s + M_n$$

mit M_s diagonalisierbar, M_n nilpotent, so dass die beiden Faktoren kommutieren, also $M_sM_n = M_nM_s$. Die Zerlegung ist durch diese Eigenschaften eindeutig charakterisiert.

Für diagonalisierbare Matrizen ist auch der Begriff *halb-einfach* (englisch: *semi-simple*) üblich, daher die Notation.

Beweis: Sei $M' = S^{-1}MS$ in Jordan-Normalform. Wie oben erklärt existiert dann die Zerlegung $M' = M'_s + M'_n$ und die beiden Summanden kommutieren. Dann ist $M_s = S^{-1}M'_sS$ diagonalisierbar und $M_n = S^{-1}M'_nS$ nilpotent.
Wir werden noch zeigen, dass es ein Polynom $P \in k[X]$ gibt mit $M_s = P(M)$. Damit können wir auch die Eindeutigkeit herleiten. Sei $M = D + N$ eine weitere Zerlegung mit D diagonalisierbar, N nilpotent und $DN = ND$. Es gilt dann sofort

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

Wegen $M_s = P(M)$ folgt dann auch $M_SD = DM_s$. Beide sind diagonalisierbar und sie vertauschen, also existiert eine gemeinsame Basis von Eigenvektoren (Übungsaufgabe). In dieser Basis sehen wir, dass auch $M_s - D$ diagonalisierbar ist. Mit den selben Argumenten sehen wir, dass $M_n = M - P(M)$ und N vertauschen. Dann ist $N - M_n$ nilpotent. Also ist

$$M_s - D = N - M_n$$

sowohl nilpotent als auch diagonalisierbar. Dies ist nur möglich für die Nullmatrix. Dies beweist die Eindeutigkeit. □

Ist M invertierbar, dann sind alle Eigenwerte ungleich 0 und M_s ist ebenfalls invertierbar. Wir erhalten daher eine Faktorisierung

$$M = M_s(\mathrm{id} + M_s^{-1}M_n) = M_sM_u$$

Definition 13.13. *Eine Matrix M heißt* unipotent*, wenn $M - \mathrm{id}$ nilpotent ist.*

Satz 13.14 (multiplikative Jordan-Zerlegung)**.** *Sei $M \in \mathrm{Gl}_m(k)$. Dann gibt es eine Zerlegung*

$$M = M_sM_u$$

mit $M_s \in \mathrm{Gl}_m(k)$ diagonalisierbar, $M_u \in \mathrm{Gl}_m(k)$ unipotente und die beiden Faktoren kommutieren. Durch diese Eigenschaften ist die Zerlegung eindeutig bestimmt.

Beweis: Wie oben bereits gesagt setzen wir $M_u = \mathrm{id} + M_s^{-1} M_n$. Die Matrix $M_s^{-1} M_n$ ist nilpotent, da die beiden Faktoren kommutieren und M_n nilpotent ist. Dies zeigt die Existenz. Sei $M = DU$ eine weitere Faktorisierung. Dann ist $M = D + N$ mit $N = D(U - \mathrm{id})$. Aus der Eindeutigkeit der additiven Zerlegung erhalten wir $D = M_s$ und dann auch $U = M_u$. □

Lemma 13.15. *In der additiven Jordan-Zerlegung gilt $M_s = P(M)$ für ein geeignetes $P \in k[X]$.*

$$P = \sum_i \lambda_i P$$

Beweis: Wir können M in Jordan-Normalform annehmen. Seien $\lambda_1, \dots, \lambda_r$ die verschiedenen Eigenwerte von M. Für $i = 1, \dots, r$ sei Sei

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

Setzen wir einen Jordan-Block zum Eigenwert λ ein, so erhalten wir a priori eine obere Dreiecksmatrix mit Diagonaleintrag

$$\prod_{j \neq i} \left(\frac{\lambda - \lambda_j}{\lambda_i - \lambda_j} \right).$$

Für $\lambda = \lambda_i$ erhalten wir den Diagonaleintrag 1. Für $\lambda = \lambda_m$ mit $m \neq i$ verschwindet der Diagonaleintrag und das Bild ist nilpotent.Eine genügend hohe Potenz, z.B. $N = \dim V$ verschwindet. Daher ist

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

für eine obere Dreieckstmatrix J_i mit Diagonaleinträgen 1. Es gibt ein Polynom P_i mit $P_i(J_i) = J_i^{-1}$ (Übungsaufgabe). Setze $R_i = XP_i$. Es gilt $R_i(0) = 0$ und $R_i(J_i) = \mathrm{id}$. Dies bedeutet

$$R_i(Q_i(M)^N) = \begin{pmatrix} 0 & & \\ & \mathrm{id} & \\ & & 0 \end{pmatrix}$$

d.h. Projektion auf den verallgemeinerten Eigenraum zu λ_i. Wir setzen

$$P = \sum_{i=1}^{s} \lambda_i R_i(Q_i(X)^N).$$

□

Eine konkrete Anwendung der Jordanschen Normalform gibt es in der Theorie der linearen Differentialgleichungen. Im eindimensionalen hat die Gleichung

$$y' = ay \quad a \in \mathbb{C}$$

die Lösung $y = \exp(ax)$. Mehrdimensional erhält man $\exp(Ax)$ für $A \in M_n(\mathbb{C})$. Diese Exponentialfunktion ist definiert als

$$\exp(Ax) = \sum_{i=0}^{\infty} \frac{A^i x^i}{i!}.$$

Sei $A = S^{-1}(D+N)S$ mit $D = [\lambda_1, \dots, \lambda_n]$ eine Diagonalmatrix und N nilpotent und $ND = DN$. Wir berechnen

$$\exp(Ax) = S^{-1} \exp(Dx) \exp(Nx) S.$$

Hierbei ist $\exp(Dx) = [e^{\lambda_1 x}, \dots, e^{\lambda_n x}]$ und $\exp(Nx)$ ein Polynom in x mit Koeffizienten in $M_n(\mathbb{C})$.
Aus dieser Darstellung folgt auch die Konvergenz der Reihe und ihre analytischen Eigenschaften.

Nachtrag

Wir tragen ein besseres Eindeutigkeitsargument nach.

Lemma 13.16. *Sei $f = f_s + f_n$ eine additive Jordanzerlegung. Dann vertauschen f_s und f_n mit allen Endormorphismen, die mit f vertauschen.*

Beweis: Sei g ein Endomorphismus, der mit f vertauscht. Dann sind die Haupträume von f automatisch g-invariant. Wir schränken uns auf einen Hauptraum ein, also $f = \lambda\,\mathrm{id} + n$ mit n nilpotent. Dann vertauscht g mit f und $\lambda\,\mathrm{id}$, also auch mit n. □

Dieses Lemma kann im Beweis der Eindeutigkeit anstelle des komplizierten Lemma 13.15 benutzt werden.

Kapitel 14

Bilineare Abbildungen

Wir arbeiten über einem beliebigen Körper k.

Definition 14.1. *Seien V, W, U Vektorräume über k. Eine Abbildung*

$$s : V \times W \to U$$

heißt bilinear*, falls*

(i) $s(v, \cdot) : W \to U$ linear ist für alle $v \in V$;

(ii) $s(\cdot, w) : V \to U$ linear ist für alle $w \in W$.

Die Abbildung heißt Paarung*, falls $U = k$, also $s : V \times W \to k$. Sie heißt* Bilinearform*, falls $U = k$ und $V = W$, also $s : V \times V \to k$.*

Beispiel. Für $k = \mathbb{R}$ und $s : V \times V \to \mathbb{R}$ ein Skalarprodukt ist s eine Bilinearform. Dies ist falsch für $k = \mathbb{C}$, denn Skalarprodukte sind sesquilinear.

Beispiel. Wir erinnern uns an den *Dualraum* $V^* = \operatorname{Hom}_k(V, k)$. Wir erhalten

$$V \times V^* \to k, \quad (v, \alpha) \mapsto \alpha(v).$$

Dies ist eine Paarung.

Definition 14.2. *Eine Paarung $s : V \times W \to k$ heißt* nicht-ausgeartet *oder* perfekt*, wenn gilt*

(i) Ist $v \in V$ mit $s(v, w) = 0$ für alle $w \in W$, so folgt $v = 0$.

(ii) Ist $w \in W$ mit $s(v, w) = 0$ für alle $v \in V$, so folgt $w = 0$.

Eine Bilinearform $s : V \times V \to k$ heißt symmetrisch*, wenn $s(x, y) = s(y, x)$ für alle $x, y \in V$. Sie heißt* definit*, wenn $s(x, x) \neq 0$ für alle $x \in V$.*

Beispiel. Skalarprodukte sind symmetrisch und definit.

Bemerkung. Definite Bilinearformen sind nicht-ausgeartet (zu jedem $x \neq 0$ gibt es ein y mit $s(x,y) \neq 0$ und $s(y,x) \neq 0$, nämlich $y = x$.) Die Umkehrung gilt nicht.

Beispiel. Sei $V = \mathbb{R}^4$,

$$s(x,y) = x_1 y_y + x_2 y_2 - x_3 y_3 - x_4 y_4.$$

Diese symmetrische Bilinearform ist nicht definit, denn für $x = (1,0,0,1)^t$ erhalten wir $s(x,x) = 0$. Sie ist aber nicht-ausgeartet, denn für $(x_1, x_2, x_3, x_4)^t \neq 0$ ist $y = (x_1, x_2, x_3, -x_4)^t$ ein Vektor mit

$$s(x,y) = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

Dies ist wieder die Pseudo-Metrik der speziellen Relativitätstheorie.

Lemma 14.3. *Sei $s : V \times W \to k$ eine Paarung. Dann sind*

$$s_2 : V \to W^*, \quad v \mapsto [s(v,\cdot) : W \to k]$$

und

$$s_1 : W \to V^*, \quad w \mapsto [s(\cdot,w) : V \to k]$$

lineare Abbildungen. Die Abbildung s_2 ist genau dann injektiv, wenn die erste Bedingung zu nicht-ausgeartet gilt. Die Abbildunge s_1 ist genau dann injektiv, wenn die zweite Bedingung zu nicht-ausgeartet gilt.
Im Fall $V = W$ stimmen s_1 und s_2 genau dann überein, wenn s symmetrisch ist.

Beweis: Wir überprüfen Linearität von s_2. Seien $x, y \in V$, $\lambda, \mu \in k$. Dann ist

$$s_2(\lambda x + \mu y) : W \to k$$

nach Definition die lineare Abbildung

$$w \mapsto s(\lambda x + \mu y, w).$$

Andererseits ist

$$\lambda s_2(x) + \mu s_2(y) : W \to k$$

die lineare Abbildung

$$w \mapsto \lambda s(x,w) + \mu s(y,w).$$

Wegen der Linerarität von s im ersten Argument stimmen die beiden überein.
Injektivität von s_2 ist äquivalent zu $s_2(x) \neq 0$ für alle $x \neq 0$. Dies ist wiederum äquivalent zu: Für alle $x \neq 0$ gibt es $w \in W$ mit $s_2(x)(w) \neq 0$, d.h $s(x,w) \neq 0$. Dies ist genau die Bedingung für nicht-ausgeartet.
Das Argument für s_1 ist analog.
Sei nun $V = W$, also $s : V \times V \to k$. Sowohl s_1 als auch s_2 sind lineare Abbildungen $V \to V^*$.

Die Bedingung $s_1 = s_2$ bedeutet $s_1(x) = s_2(x)$ für alle $x \in V$, d.h.

$$s(x, \cdot) = s(\cdot, x) : V \to k$$

d.h. $s(x, y) = s(y, x)$ für alle $x \in V$ und alle $y \in V$. Dies ist äquivalent zur Symmetrie. $\square$

Lemma 14.4. *Seien V, W endlich-dimensionale Vektorräume, $s : V \times W \to k$ eine Paarung. Dann sind äquivalent:*

(i) s ist nicht-ausgeartet.

(ii) s_1 ist ein Isomorphismus.

(iii) s_2 ist ein Isomorphismus.

Beweis: Wir erinnern uns: $\dim V = \dim V^*$, falls der Vektorraum endlich-dimensional ist.
Sei s nicht-ausgeartet, also s_1 und s_2 injektiv. Dann gilt

$$\dim V \leq \dim W^* = \dim W, \dim W \leq \dim V^* = \dim V$$

also haben die beiden Vektorräume dieselbe Dimension. Nach der Dimensionsformel sind s_1 und s_2 dann auch surjektiv, also Isomorphismen.
Sei nun $s_1 : W \to V^*$ bijektiv. Dann gilt $\dim W = \dim V^* = \dim V$. Wir müssen zeigen, dass auch s_2 injektiv ist. Sei $v \in V$ mit $s_2(v) = 0$, also $s(v, w) = 0$ für alle $w \in W$. Falls $v \neq 0$, so gibt es $v^* \in V^*$ mit $v^*(v) = 1$ (ergänze v zu einer Basis, setze v^* beliebig auf die anderen Basiselemente fort). Sei $w \in W$ mit $s_1(w) = v^*$. Dann gilt

$$s(v, w) = s_1(w)(v) = v^*(v) = 1.$$

Dies ist ein Widerspruch, also gilt $v = 0$. $\square$

Korollar 14.5. *Sei V endlich-dimensional, $s : V \times V \to k$ bilinear. Dann sind äquivalent:*

(i) s ist nicht-ausgeartet.

(ii) s_1 ist injektiv.

(iii) s_2 ist injektiv.

Beweis: Wegen $\dim V = \dim V^*$ ist die die Injektivität von s_1 äquivalent zur Bijektivität, genauso für s_2. Das Korollar folgt dann direkt aus dem Lemma. $\square$

Matrizen

Wir schreiben alles mit Matrizen. Die Formeln kennen wir bereits aus der Theorie der Skalarprodukte.

Lemma 14.6. *Sei $A \in M_{m\times n}(k)$. Dann ist die Abbildung*

$$s : k^m \times k^n \to k, \quad (x, y) \mapsto x^t Ay$$

bilinear. Sie ist symmetrisch genau dann, wenn $n = m$ und $A = A^t$ (d.h. A symmetrisch).

Beweis: Bilinearität ist klar nach den Rechenregeln des Produktes von Matrizen. Wenn A symmetrisch ist, so ist die Abbildung symmetrisch.
Ist s symmetrisch, so setzen wir die Standardbasisvektoren $x = e_i$ und $y = e_j$ ein und erhalten

$$a_{ij} = e_i^t Ae_j = s(e_i, e_j) = s(e_j, e_i) = e_j^t Ae_i = a_{ji}.$$

□

Satz 14.7. *Sei $B = (v_1, \dots, v_m)$ eine Basis von V, $B' = (w_1, \dots, w_n)$ eine Basis von W, $s : V \times W \to k$ bilinear. Dann gibt es eine eindeutige Matrix $M^{B,B'}(s) \in M_{n\times m}(k)$ so dass gilt*

$$s(v, w) = (a_1, \dots, a_m)M^{B,B'}(s) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

für

$$v = \sum a_i v_i, w = \sum b_i w_i.$$

Im Fall $V = W$ ist $M^{B,B}(s)$ genau dann symmetrisch, wenn s es ist.

Die Matrix $M^{B,B'}(s)$ heißt *darstellende Matrix.*

Beweis: Wir setzen

$$M^{B,B'}(s) = (s(v_i, w_j))_{ij}.$$

(Dies ist einzige Möglichkeit wie Einsetzen von $v = v_i$ und $w = w_j$ zeigt.) Dann gilt

$$s(\sum a_i v_i, \sum b_j v_j) = \sum_{i,j} a_i b_j s(v_i, w_j) = (a_1, \dots, a_m)M^{B,B'}(s) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

wie angegeben.
Ist $V = W$ und s symmetrisch, so folgt $M^{B,B}(s)_{ij} = M^{B,B}(s)_{ji}$. Ebenso folgt die Umkehrung. □

Satz 14.8 (Transformationsformel). *Sei B die Basis von V, C die Basis von W. Sei $B' = (v'_1, \dots, v'_m)$ eine zweite Basis von V mit Basiswechselmatrix $S = M^B_{B'}(\mathrm{id})$. Sei $C' = (w'_1, \dots, w'_m)$ eine zweite Basis von W mit Basiswechselmatrix $T = M^C_{C'}(\mathrm{id})$. Dann gilt*

$$M^{BC}(s) = S^t M^{B'C'}(s) T.$$

Beweis: Einsetzen, rechnen. □

In der LA 1 Satz 5.4 haben wir die Transformationsformel für lineare Abbildungen behandelt. Sie lässt sich am einfachsten mit einem kommutativen Diagramm verstehen. Dasselbe gilt auch jetzt.

$$\begin{tikzcd}
 & k^m \times k^n \arrow[dr, "M^{BC}(s)"] \arrow[dd, "M^B_{B'}(\mathrm{id}) \times M^C_{C'}(\mathrm{id})"] & \\
V \times W \arrow[ur, "\Phi_B \times \Phi_C"] \arrow[rr] \arrow[dr, "\Phi_{B'} \times \Phi_{C'}"'] & & k \\
 & k^m \times k^n \arrow[ur, "M^{B'C'}(s)"'] &
\end{tikzcd}$$

Hält man ein Argument fest, so erhält man ein Diagramm wie in LA 1, nur dass jetzt der Zielvektorraum k ist mit der festen Basis $1 \in k$.

Bemerkung. Matrizen können lineare Abbildunge darstellen oder bilineare. Wir erhalten unterschiedliche Transformationsformeln, die durch die unterschiedliche Notation auch angedeutet wird. Physiker sprechen von $(1,1)$ bzw. $(2,0)$-Tensoren.

Lemma 14.9. *Sei A darstellende Matrix einer Paarung $s : V \times W \to k$ von endlich dimensionalen Vektorräumen bezüglich Basen B und C. Die Paarung ist genau dann nicht-ausgeartet, wenn $\dim V = \dim W$ und A invertierbar ist.*

Beweis: Wir wissen bereits, dass im nicht-ausgearteten Fall gilt $\dim V = \dim W$. Wir benutzen die Basen, um V und W mit $\mathbb{K}^n$ zu identifizieren und die Paarung, die durch A definiert wird.

$$s(x, y) = x^t A y.$$

Wenn A nicht invertierbar ist, so gibt es $y \neq 0$ mit $Ay = 0$. Für dieses y folgt dann $s(x, y) = 0$ für alle $x \in k^n$. Die Paarung ist ausgeartet.

Sei nun A invertierbar. Sei $x = (x_1, \dots, x_n)^t \in k^n$ mit $s(x, y) = 0$ für alle $y \in k^n$. Sei $y_i = A^{-1} e_i$. Das Tupel $(y_1, \dots, y_n)$ ist eine Basis, da A invertierbar ist. Es folgt

$$0 = x^t A y_i = x^t e_i = x_i$$

für alle i, also $x = 0$. □

Orthogonalbasen

Definition 14.10. *Seien (V, s) und (V', s') Vektorräume mit Bilinearform. Eine lineare Abbildung $f : V \to V'$ heißt* orthogonal, *falls*

$$s(x, y) = s'(f(x), f(y)) \quad \text{für alle } x, y \in V.$$

Satz 14.11. *Sei V endlich-dimensionaler Vektorraum, $s : V \times V \to k$ biliniear. Dann gibt es einen orthogonalen Isomorphismus $f : V \to k^n$, wobei k^n mit der Bilinearform $(x, y) \mapsto x^t Ay$ für eine eindeutige Matrix A versehen wird.*

Beweis: Das ist nur eine Umformulierung des Satzes über die darstellende Matrix. Ist B eine Basis, so ist der induzierte Isomorphismus $V \to k^n$ automatisch orthogonal, wenn A richtig gewählt wird. □

Bemerkung. Können wir A besonders einfach wählen? Im Fall von Skalarprodukten konnten wir mit Orthonormalbasen arbeiten. Dann wird die darstellende Matrix der Bilinearform die Einheitsmatrix. Wir werden so etwas Ähnliches allgemein zeigen.

Definition 14.12. *Sei $s : V \times V \to k$ eine symmetrische Bilinearform. Dann heißt*

$$q : V \to k, \quad x \mapsto s(x, x)$$

die zu s gehörige quadratische Form.

Bemerkung. Für $x \in V$, $\lambda \in k$ gilt

$$q(\lambda x) = s(\lambda x, \lambda x) = \lambda^2 s(x, x) = \lambda^2 q(x).$$

Die Abbildung ist homogen vom Grad 2.

Satz 14.13. *Sei k ein Körper der Charakteristik ungleich 2, d.h. $0 \neq 2$ in k. Dann ist eine symmetrische Bilinearform eindeutig durch die zugehörige quadratische Form bestimmt.*

Beweis: Sei $s : V \times V \to k$, $x, y \in V$. Dann gilt

$$\begin{aligned} q(x+y) - q(x) - q(y) &= s(x+y, x+y) - s(x, x) - s(y, y) \\ &= s(x, x) + s(x, y) + s(y, x) + s(y, y) - s(x, x) - x(y, y) = 2s(x, y). \end{aligned}$$

Wenn $2 \neq 0$, so können wir $s(x, y)$ berechnen. □

Theorem 14.14 (Existenz von Orthogonalbasen)**.** *Sei k ein Körper der Charakteristik ungleich 2, V endlich-dimensionaler Vektorraum, $s : V \times V \to k$ eine symmetrische Bilinearform. Dann gibt es eine* Orthogonalbasis $b_1, \ldots, b_n$ *von V, d.h. es gilt $s(b_i, b_j) = 0$ für $i \neq j$. Die darstellende Matrix von s bezüglich dieser Basis hat Diagonalgestalt.*

Beweis: Sei q die zu s gehörige quadratische Form. Wir argumentieren mit Induktion nach $\dim V$. Für $\dim V = 0$ ist nichts zu zeigen. Sei nun $\dim V > 0$. Wir nehmen an, dass die Behauptung gilt für alle Vektorräume der Dimension echt kleiner als $\dim V$.

1. Fall: $q(x) = 0$ für alle $x \in V$, also $s(x,y) = 0$ für alle $x, y \in V$. Jede Basis ist eine Orthogonalbasis.

2. Fall: Es gibt $x_0 \in V$ mit $q(x_0) \neq 0$. Wir betrachten das orthogonale Komplement

$$W = \{y \in V | s(x_0, y) = 0\}.$$

Wegen $x_0 \notin W$ gilt $\dim W < \dim V$. Die Einschränkung $s|_W$ ist eine symmetrische Bilinearform. Nach Induktionvoraussetzung gibt es eine Orthogonalbasis $x_1, \dots, x_n$ von W. Offensichtlich ist $s(x_i, x_j)$ für $i, j = 0, \dots, n$ mit $i \neq j$.

Behauptung. *$x_0, x_1, \dots, x_n$ ist eine Basis von V.*

Wir zeigen zunächst, dass das Tupel linear unabhängig ist. Sei

$$\sum_{i=0}^{n} a_i x_i = 0.$$

Dann folgt

$$0 = s(\sum a_i x_i, x_0) = \sum a_i s(x_i, x_0) = a_0 s(x_0, x_0).$$

Wegen $s(x_0, x_0) \neq 0$ folgt $a_0 = 0$. Nach Voraussetzung sind $x_1, \dots, x_n$ linear unabhängig, also auch $a_1 = a_2 = \dots = a_n = 0$.

Wir überprüfen, dass das Tupel ein Erzeugendensystem ist. Sei $x \in V$ beliebig. Wir betrachten

$$x' = x - \frac{s(x, x_0)}{s(x_0, x_0)} x_0.$$

Es gilt

$$s(x_0, x') = s(x_0, x) - \frac{s(x, x_0)}{s(x_0, x_0)} s(x_0, x_0) = 0.$$

Also ist $x' \in W$ und eine Linearkombination von $x_1, \dots, x_n$. Damit ist x eine Linearkombination von $x_0, x_1, \dots, x_n$. □

Bemerkung. (i) Der Beweis ist sehr ähnlich zum Gram-Schmidtschen Orthonormalisierungsverfahren. Es fehlt nur die Normierung auf $s(x_i, x_i) = 1$. Diese ist möglich, wenn der Wert eine Quadratzahl in k ist, also nicht immer.

(ii) Im Fall $k = \mathbb{R}$ erhalten wir wieder den Sylvesterschen Trägheitssatz. Indem wir die Basisvektoren skalieren, können wir annehmen, dass die darstellende Matrix eine Diagonalmatrix mit den einzigen Einträgen 1, -1 und 0 ist. In Beweis von Theorem 11.18 haben wir statt der Theorie der quadratischen Formen mit dem Spektralsatz für selbst-adjungierte Matrizen gearbeitet. Der Eindeutigkeitsteil des Theorems wird wie dort bewiesen.

(iii) Für $k = \mathbb{C}$ ist der Satz zwar ähnlich zu den Ergebnissen des Kapitels 11 zu selbst-adjungierten Matrizen, aber nicht identisch. Damals ging es um Matrizen mit $\overline{A}^t = A$, diesmal mit hermiteschenn Fall erhält man eine reelle Diagonalmatrix, diesmal eine komplexe.

Adjungierte Abbildungen

Definition 14.15. *Sei $s : V \times W \to k$ nicht-ausgeartete Paarung, $f : V \to V$ und $g : W \to W$ Endomorphismen. Dann heißen f und g* adjungiert*, falls*

$$s(f(v), w) = s(v, g(w)) \quad \text{für alle } v \in V,\ w \in W,$$

genauer f ist linksadjungiert *zu g und g* rechts-adjungiert *zu f. Ist $V = W$ und $f = g$, so heißt f* selbst-adjungiert.

Bemerkung. Das ist identisch zu unseren Definitionen euklidischen Fall, also Vektorräume mit Skalarprodukt über $\mathbb{R}$. Ist s symmetrische Bilinearform, so muss nicht zwischen rechts und links unterschieden werden.

Beispiel. Sei $s : V \times V^* \to k$ die kanonische Paarung und f ein Endomorphismus. Wir betrachten die duale Abbildung $f^* : V^* \to V^*$ (vergleiche LA1), also $f^*(\alpha) = \alpha \circ f$ für alle $\alpha \in V^*$. Dann sind f und f^* adjungiert:

$$s(f(v), \alpha) = \alpha(f(v)), s(v, f^*\alpha) = f^*(\alpha)(v) = \alpha \circ f(v).$$

Bemerkung. Man schreibt oft f^* für die Linksadjungierte zu f, g_* für die Rechstadjungierte zu g, also

$$s(f^*v, w) = s(v, f(w)), s(g(v), w) = s(v, g_*(w)).$$

Lemma 14.16. *Seien V, W endlich-dimensionale Vektorräume, $s : V \times W \to k$ eine nicht-ausgeartete Paarung, $f : V \to V$ ein Endomorphismus. Dann existiert ein eindeutiges Rechtsadjungiertes zu f. (Ebenso ein Linksadjungiertes für $g : W \to W$.)*

Beweis: Wir betrachten den Isomorphismus $s_1 : W \to V^*$. Wir erhalten

$$\begin{tikzcd} W \arrow[r, "s_1"] \arrow[d, dotted] & V^* \arrow[d, "f^*"] \\ W \arrow[r, "s_1"'] & V^* \end{tikzcd}$$

Wir definieren

$$g = s_1^{-1} \circ f^* \circ s_1.$$

Behauptung. *g ist rechts-adjungiert zu f, also $s(v, g(w)) = s(f(v), w)$.*

Wir schreiben $\tilde{s} : V \times V^* \to k$ für die kanonische Paarung. Dann ist $V \times W \xrightarrow{\mathrm{id} \times s_1} V \times V^*$ verträglich mit den Paarungen: Für alle $v \in V$, $w \in W$ gilt nach Definition
$$\tilde{s}(v, s_1(w)) = s_1(w)(v) = s(v, w)$$
und dann auch
$$s(v, s_1^{-1}\alpha) = \tilde{s}(v, \alpha).$$
Daher gilt auch für alle $v \in V$, $w \in W$
$$s(v, g(w)) = s(v, s_1^{-1} f^* s_1(w)) = \tilde{s}(v, f^* s_1(w)) = \tilde{s}(f(v), s_1(w)) = s(f(v), w).$$
Nun zur Eindeutigkeit: Seien g und g' beide rechtsadjungiert zu f. dann gilt für alle $v \in V$ und $w \in W$
$$s(v, g(w)) = s(f(v), w) = s(v, g'(w)) \Rightarrow s(v, g(w) - g'(w)).$$
Da s nicht-ausgeartet ist, folgt $g(w) - g'(w) = 0$ für alle $w \in W$, also $g = g'$. $\square$

Umgekehrt können wir Endomorphismen benutzen, um neue Bilinearformen zu definieren. Sei $s : V \times V \to k$ eine Bilinearform, $f \in \mathrm{End}(V)$ ein Endomorphismus. Dann ist
$$s' : V \times V \to k, (x, y) \mapsto s(f(x), y)$$
eine neue Bilinearform.

Lemma 14.17. *Sei s symmetrische, nicht-ausgeartete Bilinearform auf V, $f \in \mathrm{End}(V)$, s' wie oben. Dann ist s' genau dann symmetrisch, wenn f selbstadjungiert ist.*

Beweis: Es gilt für alle $x, y \in V$
$$s'(x, y) = s(f(x), y) = s(y, f(x)), s'(y, x) = s(f(y), x).$$
Ist s' symmetrisch, so sind die Ausdrücke gleich und f selbst-adjungiert. Ist f selbstadjungiert, so sind die Terme gleich. $\square$

Noch einmal Matrizen

Lemma 14.18. *Seien V und W endlich-dimensionale Vektorräume, $s : V \times W \to k$ eine nicht-ausgeartete Paarung. Sei $v_1, \dots, v_n$ eine Basis von V. Dann gibt es eine eindeutige Basis $w_1, \dots, w_n$ von W, so dass*
$$s(v_i, w_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Die so gefundene Basis heißt *duale Basis*. Wir schreiben auch $v_1^*, \dots, v_n^*$. Im Fall der kanonischen Paarung $V \times V^* \to k$ erhalten wir den Begriff aus der LA 1 zurück.

Beweis: **1. Argument:** Wir benutzen den Isomorphismus $s_1 : W \to V^*$, um uns auf den Fall der kanonischen Paarung zu reduzieren. In diesem Fall wurde die duale Basis in LA 1 konstruiert.

2. Argument: Wir wähle eine beliebige Basis $x_1, \dots, x_n$. Dann hat s bezüglich $v_1, \dots, v_n$ und $x_1, \dots, x_n$ eine darstellende Matrix A. Da s nicht-ausgeartet ist, ist sie invertierbar. Sei $A^{-1} = (b_{ij})_{ij}$. Dann löst $w_j = \sum_{i=1}^n b_{ij} x_i$ das Problem:

$$s(v_i, w_j) = \sum_{l=1}^n b_{lj} s(v_i, x_l) = \sum_{l=1}^n b_{lj} a_{il} = (AB)_{ij} = \delta_{ij}.$$

□

Beispiel. Ist s symmetrisch und $v_1, \dots, v_n$ eine Orthonormalbasis, so gilt $v_i^* = v_i$ für $i = 1, \dots, n$.

Lemma 14.19. *Seien V und W endlich-dimensional, $s : V \times W \to k$ eine nicht-ausgeartete Paarung. Seien B und B^* duale Basen B und B^*. Sei $f \in \mathrm{End}(V)$ adjungiert zu $f^* \in \mathrm{End}(W)$. Dann ist die darstellende Matrix $M_B^B(f)$ von f transponiert zur darstellenden Matrix $M_{B^*}^{B^*}(f^*)$ der adjungierten Abbildung bezüglich der dualen Basis.*

Beweis: Sei $B = (v_1, \dots, v_n)$ und $B^* = (v_1^*, \dots, v_n^*)$. Wir schreiben $C = M_B^B(f) = (c_{ij})$ und $D = M_{B^*}^{B^*}(f^*) = (d_{ij})$. Nach Definition ist

$$f(v_j) = \sum_{i=1}^n c_{ij} v_i \Rightarrow s(f(v_j), v_i^*) = c_{ij}$$

und

$$f^*(v_j^*) = \sum_{i=1}^n d_{ij} v_i^* \Rightarrow s(v_i, f^*(v_j)) = d_{ij}$$

und zusammen

$$c_{ij} = s(f(v_j), v_i^*) = s(v_j, f^*(v_i^*)) = d_{ji}.$$

□

Bemerkung. Wir haben bereits zwei Spezialfälle kennengelernt. Im Fall $V = V^*$ tauchte die transponierte Matrix in LA 1 als Matrix der dualen Abbildung auf. Im Fall $V = W$, $k = \mathbb{R}$ und einem Skalarprodukt war es die Matrix der adjungierten Abbildung. Das war also kein Zufall, sondern es ging beide Male um nicht-ausgeartete Paarungen.

Korollar 14.20. *Sei V endlich-dimensional, $s : V \times V \to k$ nicht-ausgeartete symmetrische Bilinearform und $f \in \mathrm{End}(V)$. Wir nehmen an, dass V eine selbst-duale Basis besitzt (also eine Orthonormalbasis). Dann ist f genau dann selbst-adjungiert, wenn die darstellende Matrix bezüglich einer Orthonormalbasis symmetrisch ist.*

Beweis: Wegen $B = B^*$ und nach dem Lemma gilt

$$M_B^B(f)^{\text{Spur}} = M_{B^*}^{B^*}(f^*) = M_B^B(f^*).$$

Ist $f = f^*$, so ist die Matrix symmetrisch. Ist die Matrix symmetrisch, so ist $f = f^*$. □

Schlussbemerkung

Bemerkung. Für nicht ausgeartete symmetrische Bilinearform gelten viele der Sätze über das Skalarprodukt auch. Eine große Ausnahme gibt es:

Ist $s : V \times V \to \mathbb{R}$ ein Skalarprodukt und $W \subset V$ ein Untervektorraum, dann ist $s|_W$ ebenfalls ein Skalarprodukt. Das ist i.a. falsch für nicht-ausgeartete Bilinearformen. Als Beispiel kann man im Fall $k = \mathbb{R}$, $V = \mathbb{R}^2$ die quadratische Form $q(x, y) = x^2 - y^2$ betrachten. Sie gehört zur Bilinearform mit darstellender Matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Eingeschränkt auf der Untervektorraum W, der von $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ aufgespannt wird, erhält man $q|_W = 0$.

Kapitel 15

Das Tensorprodukt

Wir arbeiten weiter über einem allgemeinen Körper k. Wir betrachten jetzt bilineare Abbildungen

$$V \times W \to U$$

für k-Vektorräume V, W, U.

Das *Tensorprodukt* zweiter Vektorräume ist ein neuer Vektorraum $V \otimes W$. Das *Tensorprodukt* zweier Vektoren $v \in V$ und $w \in W$ ist ein Element in $V \otimes W$. Die Abbildung

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

ist bilinear. Es gelten also die Rechenregeln:

(i) Für alle $v_1, v_2 \in V$, $a_1, a_2 \in k$, $w \in W$ gilt

$$(a_1 v_1 + a_2 v_2) \otimes w = a_1(v_1 \otimes w) + a_2(v_2 \otimes w).$$

(ii) Für alle $w_1, w_2 \in W$, $b_1, b_2 \in k$, $v \in V$ gilt

$$v \otimes (b_1 w_1 + b_2 w_2) = b_1(v \otimes w_1) + b_2(v \otimes w_2).$$

Mehr gibt es eigentlich nicht zu sagen. Die einzigen Rechenregeln sind diejenigen, die aus dem obigen folgen. So gilt z.B. für alle $a \in k$, $v \in V$, $w \in W$

$$a(v \otimes w) = (av) \otimes w = v \otimes (aw).$$

In der Literatur finden sich zwei Definitionen von $V \otimes W$:

(i) durch eine universelle Eigenschaft,

(ii) durch eine explizite Konstruktion.

Wie die meisten Mathematiker sehe ich in der ersten Version die “richtige” Charakterisierung mit der auch die Beweise leichter werden, aber die zweite ist vermutlich auf den ersten Blick zugänglicher. Beginnen wir also mit der Konstruktion.

Definition 15.1. *Seien V, W zwei k-Vektorräume. Wir definieren*

$$V \otimes W = \bigoplus_{(v,w)\in V\times W} k/R$$

wobei R der Untervektorraum ist, der von den Elementen

$$(a_1v_1 + a_2v_2, w) - a_1(v_1, w) - a_2(v_2, w)$$
$$(v, b_1w_1 + b_2w_2) - b_1(v, w_1) - b_2(v, w_2)$$

für $a_1, a_2, b_1, b_2 \in k$, $v, v_1, v_2 \imath V$, $w, w_1, w_2 \in W$ erzeugt wird. Hier schreiben wir abkuerzend (v, w) für das Element von $\bigoplus_{(v',w')\in V\times W} k$, das an der Stelle (v, w) den Eintrag 1 hat und 0 sonst.
Wir definieren

$$\theta : V \times W \to V \otimes W$$

als $(v, w) \mapsto (v, w) + U$.

Beispiel. Sei $V = W = k$. Dann erzwingen die Relationen, dass

$$a \otimes b = a(1 \otimes b) = ab(1 \otimes 1),$$

d.h. der Vektorraum $k \otimes k$ wird von $1 \otimes 1$ erzeugt. Gilt $1 \otimes 1 \neq 0$? Dies lässt sich leicht aus dem nächsten Satz folgern.

Bemerkung. Elemente der Form $v \otimes w$ heißen *Elementartensoren*. Im allgemeinen ist *nicht* jedes Element ein Elementartensor! Elementartensoren sind die einfach die Elemente im Bild von θ. Das Bild einer bilineare Abbildung ist i.a. *kein* Untervektorraum. Statt dessen ist jedes Element von der Form $\sum_{i=1}^n v_i \otimes w_i$ für $v_1, \dots, v_n \in V$, $w_1, \dots, w_n \in W$.
Es gibt *keine* Rechenregel zum Vereinfachen von $v_1 \otimes w_1 + v_2 \otimes w_2$ (es sei denn, v_i oder w_i habe spezielle Eigenschaften).

Theorem 15.2 (Universelle Eigenschaft)**.** *Sei k Körper, V, W Vektorräume über k*

(i) $\theta : V \times W \to V \otimes W$ is bilinear.

(ii) Sei U ein weiterer Vektorraum, $s : V \times W \to U$ bilinear. Dann gibt es eine eindeutige lineare Abbildung

$$\tilde{s} : V \otimes W \to U$$

so dass $s = \tilde{s} \circ \theta$.

(iii) Das Paar $(V \otimes W, \theta)$ ist durch diese Eigenschaft eindeutig bestimmt bis auf eindeutigen Isomorphismus.

Beweis: Die Bilinearität von θ folgt automatisch aus der Definition des Relationenraumes U.
Wir betrachten s. Die mengentheoretische Abbildung $V \times W \to U$ definiert eine k-lineare Abbildung

$$s' : \bigoplus_{(v,w)\in V\times W} k \to U$$

die den Basisvektor (v, w) auf $s(v, w)$ abbildet. Wir überprüfen, dass die Abbildung über $V \otimes W$ faktorisiert. Dazu müssen wir zeigen, dass die Erzeuger von R auf 0 abgebildet werden. Es gilt

$$\begin{aligned} &s'((a_1v_1 + a_2v_2, w) - a_1(v_1, w) - a_2(v_2, w)) \\ &\qquad = s(a_1v_1 + a_2v_2, w) - a_1s(v_1, w) - a_2s(v_w, w) = 0 \end{aligned}$$

da s bilinear ist. Ebenso rechnen wir für den zweiten Typ von Erzeugern. Damit faktorisiert s' über eine lineare Abbildung $\tilde{s} : V \otimes W \to U$. Die Relation $s = \tilde{s}\circ\theta$ gilt nach Konstruktion, da

$$\tilde{s} \circ \theta(v, w) = \tilde{s}(v \otimes w) = s'((v, w)) = s(v, w).$$

Die Abbildung $\tilde{s}$ ist durch diese Bedingung eindeutig bestimmt, das die Elemente im Bild von θ den Raum $V \otimes W$ erzeugen.
Sei (X, ϕ) ein weiterer Vektorraum, $\phi : V \times W \to X$ bilinear, so dass die universelle Eigenschaft erfüllt ist, d.h. jede belineare Abbildung in jedes U faktorisiert eindeutig über X. Wir wenden dies an auf $U = V \otimes W$ und die bilineare Abbildung θ. Wir erhalten also eine lineare Abbildung

$$f : X \to V \otimes W, f \circ \phi = \theta.$$

Umgekehrt erhalten wir aus der universelle Eigenschaft von $(V \otimes W, \theta)$ mit $U = X$, $s = \phi$ eine lineare Abbildung

$$g : V \otimes W \to X, g \circ \theta = \phi.$$

Wir betrachten $f \circ g : V \otimes W \to V \otimes W$. Die Abbildung ist linear. Außerdem gilt (Einsetzen)

$$f \circ g \circ \theta = f \circ \phi = \theta = \mathrm{id} \circ\theta.$$

Wir betrachten nun die universelle Eigenschaft für $(V \otimes W, \theta)$ mit $U = V \otimes W$, $s = \theta$. Die Eindeutigkeit von $\tilde{s}$ impliziert dass

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

Analog folgt auch $g \circ f = \mathrm{id}$. Damit sind die Abbildungen f und g zueinander inverse Isomorphismen. Sie sind verträglich mit θ und ϕ. Die Verträglichkeit mit θ und ϕ macht die Abbildungen eindeutig. □

Beispiel. Sei $V = W = k$, $U = k$. Sei $s : k \times k \to k$ die Multiplikation. Diese Abbildung ist bilinear. Also gibt es eine eindeutige lineare Abbildung

$$\tilde{s} : k \otimes k \to k$$

mit $\tilde{s}(a \otimes b) = s(a, b) = ab$. Insbesondere gilt $\tilde{s}(1 \otimes 1) = 1$. Daher ist $1 \otimes 1 \neq 0$ in $k \otimes k$. Der Vektorraum hat die Dimension 1.

Satz 15.3. *Sei $B = (b_i)_{i \in I}$ eine Familie in V, $C = (C_j)_{j \in J}$ eine Familie in in W. Wir betrachten die Familie $B \otimes C = (b_i \otimes c_j)_{(i,j) \in I \times J}$ in $V \otimes W$.*

(i) Sind B und C linear unabhängig, dann ist auch $B \otimes C$ linear unabhängig.

(ii) Sind B und C Erzeugendensysteme von V und W, so ist $B \otimes C$ ein Erzeugendensystem von $V \otimes W$.

(iii) Sind B und C Basen, so ist $B \otimes C$ eine Basis von $V \otimes W$.

Beweis: Seien B, C linear unabhängig. Sei

$$\sum_{i,j} a_{ij} b_i \otimes c_j = 0$$

für $a_{ij} \in k$, fast alle 0. Die Rechenregeln helfen uns nicht beim Vereinfachen. Statt dessen nutzen wir die universelle Eigenschaft. Sei $i_0 \in I$, $j_0 \in J$. Sei $f_{i_0} : V \to k$ mit

$$f_{i_0}(b_i) = \begin{cases} 1 & i = i_0 \\ 0 & i \neq i_0 \end{cases}$$

Sei $g_{j_0} : W \to k$ mit

$$g_{j_0}(c_j) = \begin{cases} 1 & j = j_0 \\ 0 & j \neq j_0 \end{cases}$$

Dann ist

$$V \times W \to k; (v, w) \mapsto f_{i_0}(v) g_{j_0}(w)$$

bilinear. Aus der universellen Eigenschaft erhalten wir eine eindeutige lineare Abbildung

$$f_{i_0} \otimes g_{j_0} : V \otimes W \to k.$$

Diese wenden wir auf unsere Relation an und erhalten

$$0 = f_{i_0} \otimes g_{j_0}(\sum_{i,j} a_{ij} b_i \otimes c_j) = \sum a_{ij} f_{i_0}(b_i) g_{j_0}(c_j) = a_{i_0 j_0}.$$

Also verschwindet jeder der Koeffizienten und die Tensoren sind linear unabhängig.

Seien nun B und C Erzeugendensysteme. Der Vektorraum $V \otimes W$ wird von den Elementartensoren $v \otimes w$ erzeugt, also genügt es, diese durch die Elemente von

$B \otimes C$ auszudrücken. Sei $v = \sum_i \beta_i b_i \in V$ und $w = \sum_j \gamma_j c_j \in W$ (jeweils fast alle Koeffizienten 0). Dann folgt

$$v \otimes w = (\sum_i \beta_i b_i) \otimes (\sum_j \gamma_j c_j) = \sum_{i,j} \beta_i \gamma_j (b_i \otimes c_j).$$

Sind B und C Basen, so ist $B \otimes C$ sowohl linear unabhängig als auch ein Erzeugendensystem, also eine Basis. □

Korollar 15.4. *Sei* $\dim V = n$, $\dim W = m$, *dann ist* $\dim V \otimes w = nm$.

Beweis: Wir zählen die Anzahl der Elemente in $B \otimes C$ ab, wenn $|B| = n$, $|C| = m$. □

Bemerkung. Benutzen wir die Basis B, um V mit k^n zu identifizieren und die Basis C, um W mit k^m zu identifizieren, so können wir $V \otimes W$ mit k^{nm} als Matrizen der Größe $n \times m$ auffassen. Ein Basiswechsel für V und W führt dann zu veränderten Einträgen in der Matrix von der Form

$$M \mapsto S^t M T$$

wobei S und T die Basiswechselmatrizen sind. Das erinnert an die Transformationsformel für Bilinearformen, ist aber tatsächlich dual dazu, denn

$$\mathrm{Bilin}_k(V, W; k) \cong \mathrm{Hom}_k(V \otimes W, k) = (V \otimes W)^*$$

nach der universellen Eigenschaft des Tensorproduktes.

Diese Art von Definition ist in der Physik üblich: Eine Größe ist ein Tensor, wenn sie sich bzgl. eines Koordiantenwechsels (d.h. eines Basiswechsels) auf eine bestimmte Art transformiert. Der Zugang der Mathematik ist koordinatenunabhängig.

Rechenregeln

Lemma 15.5 (Funktorialität)**.** *Seien* $f : V \to V'$ *und* $g : W \to W'$ *lineare Abbildungen. Dann gibt es genau eine lineare Abbildung*

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

so dass das folgende Diagramm kommutiert:

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

Mit anderen Worten

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

für alle $v \in V$, $w \in W$.

Beweis: Die Abbildung

$$V \times W \to V' \times W' \to V' \otimes W'$$

ist bilinear, faktorisiert also eindeutig über $V \otimes W$. Dies ist die gesuchte Abbildung. $\square$

Lemma 15.6. *Es gibt genau eine lineare Abbildung*

$$k \otimes V \to V$$

mit $a \otimes v \mapsto av$ und diese ist ein Isomorphismus.

Beweis: Die skalare Multiplikation $k \times V \to V$ ist bilinear, faktorisiert also eindeutig über $k \otimes V$. Ist B eine Basis von V, so ist $1 \otimes B$ eine Basis von $k \otimes V$. Sie wird on der linearen Abbildung auf B abgebildet. Daher erhalten wir einen Isomorphismus. $\square$

Satz 15.7. *Seien U, V, W Vektorräume über k.*

(i) Es gibt genau eine lineare Abbildung

$$\tau : V \otimes W \to W \otimes V$$

mit $\tau(v \otimes w) = w \otimes v$ für alle $v \in V$, $w \in W$. Sie ist ein Isomorphismus.

(ii) Es gibt genau eine lineare Abbildung

$$\alpha : U \otimes (V \otimes W) \to (U \otimes V) \otimes W$$

mit $\alpha(u \otimes (v \otimes w)) = (u \otimes v) \otimes w$ für alle $u \in U$, $v \in V$, $w \in W$. Sie ist ein Isomorphismus.

(iii) Es gibt genau eine lineare Abbildung

$$U \otimes (V \oplus W) \to (U \otimes V) \oplus (U \otimes W)$$

mit $u \otimes (v, w) \mapsto (u \otimes v, u \otimes w)$ für alle $u \in U$, $v \in V$, $w \in W$. Sie ein Isomorphismus.

Beweis: Die Abbildung $V \times W \to W \times V \to W \otimes V$ ist bilinear, faktorisiert also über τ wie gewünscht. Beim Vertauschen der Rollen von V und W erhält man die inverse Abbildung.

Die Behauptung mit drei Faktoren sieht man ähnlich, Übungsaufgabge.

Für die Distributivität betrachten wir

$$U \times (V \oplus V) \to U \otimes V \oplus U \otimes W, \quad (u, (v, w)) \mapsto (u \otimes v, u \otimes w).$$

Die Abbildung ist bilinear, faktorisiert also wie gewünscht. Für die inverse Abbildung betrachten wir die Inklusionen

$$i_1 : V \to V \oplus W, i_2 : W \to V \oplus W.$$

Die Funktorialität des Tensorproduktes induziert dann Abbildungen

$$\mathrm{id} \otimes i_1 : U \otimes V \to U \otimes U \otimes (V \oplus W), \mathrm{id} \otimes i_2 : U \otimes W \to U \otimes (V \oplus W).$$

Zusammen erhalten wir

$$(\mathrm{id} \otimes i_1, \mathrm{id} \otimes i_2) : U \otimes V \oplus U \otimes W \to U \otimes (V \oplus W).$$

Die beiden Abbildungen sind offensichtlich invers zueinander. □

Bemerkung. Die letzte Rechenregeln gilt auch für beliebige direkte Summen, also z.B. nach der Wahl einer Basis $W \cong \bigoplus_{i \in I} k$

$$V \otimes W \cong V \otimes (\bigoplus_{i \in I} k) \cong \bigoplus_{i \in I} (V \otimes k) \cong \bigoplus_{i \in I} V.$$

Wendet man diese Idee auch noch auf V an, so erhält man die Aussage über Basen zurück, die wir mit der Hand nachgerechnet haben.

Beispiel. Sei $V \cong k^n$, $W \cong k^m$, so folgt

$$V \otimes W \cong V \otimes k^m \cong (V \otimes k)^m \cong V^m \cong (k^n)^m = k^{nm}.$$

Beispiel. Sei V ein $\mathbb{R}$-Vektorraum. Der Körper $\mathbb{C}$ ist ein $\mathbb{R}$-Vektorraum, also können wir

$$V_{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{R}} V$$

definieren. Dies ist ein $\mathbb{R}$-Vektorraum der Dimension $2 \dim V$. Mit der skalaren Multiplikation

$$\mathbb{C} \times V_{\mathbb{C}} \to V_{\mathbb{C}}$$

induziert von der Multiplikation in $\mathbb{C}$ via

$$\mathbb{C} \times \mathbb{C} \to V \to \mathbb{C} \times V \to V_{\mathbb{C}}$$

wird $V_{\mathbb{C}}$ zu einem $\mathbb{C}$-Vektorraum. Ist $v_1, \dots, v_n$ eine $\mathbb{R}$-Basis von V, so ist $1 \otimes v_1, \dots, 1 \otimes v_n$ eine $\mathbb{C}$-Basis von $V_{\mathbb{C}}$. Jede $\mathbb{R}$-lineare Abbildung $f : V \to W$ induziert eine $\mathbb{C}$-lineare Abbildung $f_{\mathbb{C}} : V_{\mathbb{C}} \to W_{\mathbb{C}}$. Die darstellenden Matrizen bleiben einfach gleich.

Implizit haben wir das mehrfach benutzt, um Sätze über $f \in \mathrm{End}(V)$ zu beweisen. Statt einfach $f_{\mathbb{C}}$ zu betrachten, sind wir zur darstellenden Matrix übergegangen und haben dann argumentiert, dass eine reelle Matrix doch auch eine komplexe ist.

Der Beweis des Satzes von Cayley-Hamilton benutzte $V_{\overline{k}}$ für einen algebraischen Abschluss $\overline{k}$ von k.

Satz 15.8 (Exaktheit)**.** *Sei $f : V \to V'$ linear, W ein Vektorraum. Wir betrachten $f \otimes \mathrm{id} : V \otimes W \to V' \otimes W$. Es gilt*

$$\mathrm{Im}(f \otimes \mathrm{id}) \cong \mathrm{Im}(f) \otimes W,$$
$$\mathrm{Ker}(f \otimes \mathrm{id}) \cong (\mathrm{Ker}(f)) \otimes W.$$

Beweis: Da wir es mit Vektorräumen zu tun haben, können wir nach Basisergängssatz V und V' zerlegen als

$$V \cong K \oplus B, V' \cong B \oplus C$$

so dass die von f induzierte Abbildung

$$f' : K \oplus B \to B \oplus C$$

die Form $f'|_K = 0$, $f'|_B = \mathrm{id}_B$ hat. Hierin ist also K der Kern und B das Bild. Wegen der Funktorialität des Tensorprodukts und der Verträglichkeit mit direkten Summen, erhalten wir

$$f' \otimes \mathrm{id} : K \otimes W \oplus B \otimes W \to B \otimes W \oplus C \otimes W$$

mit

$$(f' \otimes \mathrm{id})|_{K \otimes W} = f'|_K \otimes \mathrm{id} = 0 \otimes \mathrm{id} = 0$$

und

$$(f' \otimes \mathrm{id})|_{B \otimes W} = f'_B \otimes \mathrm{id} = \mathrm{id}_B \otimes \mathrm{id}_W = \mathrm{id}\,.$$

Damit ist $K \otimes W$ der Kern und $B \otimes W$ das Bild. □

Bemerkung. Vieles aus der linearen Algebra funktioniert auch für Ringe statt Körpern, das Tensorprodukt und seine Rechenregeln gehören dazu, mit Ausnahme der Berechnung des Kerns im letzten Satz. Der Beweis benutzte die Existenz von Basen, genau das geht im allgemeinen schief.

Tensoren in der Physik

Definition 15.9. *Sei V ein endlich-dimensionaler Vektorraum. Ein (r,s)-Tensor ist ein Element von*

$$V^{(r,s)} := V^{\otimes r} \otimes (V^*)^{\otimes s}.$$

Er heißt r-fach kovariant und s-fach kontravariant.

Die Wahl einer Basis $e_1, \ldots, e_n$ auf V induziert Basen auf V^* (duale Basis $e^1, \ldots, e^n$) und auf allen $V^{(r,s)}$. Die Elemente haben die Form

$$\sum a^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} e_{i_1} \otimes \ldots \otimes e_{i_r} \otimes e^{j_1} \otimes \ldots \otimes e^{j_s}.$$

In Einsteinkonvention schreibt man

$$a^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} e_{i_1} \otimes \ldots \otimes e_{i_r} \otimes e^{j_1} \otimes \ldots \otimes e^{j_s}.$$

Zu Summieren ist über jeden Index, der oben und unten auftaucht. Ein Basiswechsel auf V induziert eine transformation der $(a^{i_1,\ldots,i_r}_{j_1,\ldots,j_s})$. Hierbei verhalten sich die ko- und die kontravarianten Indizes verschieden (Übungsaufgabe). Ein Basiswechsel bedeutet einen Koordinatenwechsel im Raum. Oft werden hierbei krummlinige Koordinaten benutz. Die korrekte mathematische Behandlung findet in der Theorie der Mannigfaltigkeiten statt.

Kapitel 16

Multilineare Abbildungen

Definition 16.1. *Seien V, U Vektorräume.*

(i) Eine Abbildung $f : V^n \to U$ heißt multilinear*, wenn sie k-linear in jedem Argument ist.*

(ii) Eine multilineare Abbildung heißt symmetrisch*, wenn*

$$f(v_1, \ldots, v_n) = f(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$$

für alle $\sigma \in S_n$. Im Fall $U = k$ heißt sie Multilinearform.

(iii) Eine multilineare Abbildung heißt alternierend*, wenn $f(v_1, \ldots, v_n) = 0$ für alle Tupel mit $v_i = v_j$ für ein Paar $i \neq j$.*

Beispiel. $\det : (k^n)^n \to k$ ist eine alternierende Multilinearform.

Beispiel. Sei $f : V^n \to k$ eine Multilinearform, $g : W^n \to k$ ebenfalls. Dann ist $f + g : (V \oplus W)^n \to k$ (definiert durch

$$((v_1, w_1), (v_2, w_2), \ldots, (v_n, w_n)) \mapsto f(v_1, \ldots, v_n) + g(w_1, \ldots, w_n)$$

ebenfalls multilinear. Sind f und g beide symmetrisch oder beide alternierend, dann auch $f + g$.

Bemerkung. Es genügt, die Symmetriebedingung nur für Transpositionen zu verlangen.

Lemma 16.2. *Sei f alternierend. Dann gilt*

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

Beweis: Es genügt, die Formel für Transpositionen zu überprüfen. Wir rechnen im Fall $n = 2$ (d.h. wir halten alle anderen Argumente einfach fest).

Behauptung. $f(v_1, v_2) = -f(v_2, v_1)$.

Es gilt

$$\begin{aligned}0 = f(v_1 + v_2, v_1 + v_2) = f(v_1, v_1) + f(v_1, v_2) + f(v_2, v_1) + f(v_2, v_2)\\ = f(v_1, v_2) + f(v_2, v_1).\end{aligned}$$

□

Dieselbe Rechnung haben wir bereits für Determinanten gemacht.

Nach der universellen Eigenschaft des Tensorproduktes und dess Assoziativität faktorisiert jede multilineare Abbildung $V^n \to U$ eindeutig über eine lineare Abbildung $V^{\otimes n} \to U$. Eine analoge Aussage wollen wir auch für symmetrische und alternierende Abbildungen.

Definition 16.3. *Sei V ein Vektorraum, $n \geq 1$. Die n-*te symmetrische Potenz *von V ist der Vektorraum*

$$S^n(V) = V^{\otimes V}/I_n$$

wobei $I_n \subset V^{\otimes}$ erzeugt wird von den Elementen $v_1 \otimes \ldots \otimes v_n - v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(n)}$ für alle $v_1, \ldots, v_n \in V$, $\sigma \in S_n$.

*Die n-*te äußere Potenz *von V ist der Vektorraum*

$$\Lambda^n(V) = V^{\otimes V}/J_n,$$

wobei $J_n \subset V^{\otimes n}$ erzeugt wird von Elemente $v_1 \otimes \ldots \otimes v_n$ mit $v_1, \ldots, v_n \in V$, so dass $v_i = v_j$ für ein Paar von Indizes mit $i \neq j$.

Satz 16.4 (Universelle Eigenschaft)**.** *Sei V ein Vektorraum, $n \geq 1$.*

(i) Die Komposition $V^n \to V^{\otimes n} \to S^n(V)$ ist symmetrisch. Sie ist universell, d.h. jede symmetrische multilineare Abbildung $V^n \to U$ fakorisiert eindeutig über eine lineare Abbildung $S^n(V) \to U$.

(ii) Die Komposition $V^n \to V^{\otimes n} \to \Lambda^n(V)$ ist alternierend. Sie ist universell, d.h. jede alternierende multilineare Abbildung $V^n \to U$ fakorisiert eindeutig über eine lineare Abbildung $\Lambda^n(V) \to U$.

Beweis: Nach Konstruktion ist $V^n \to S^n(V)$ symmetrisch. Sei $f : V^n \to U$ symmetrische mutilineare Abbildung. Nach der universellen Eigenschaft des Tensorproduktes erhalten wir eine lineare Abbildung $\tilde{f} : V^{\otimes n} \to U$. Da f symmetrisch ist, bildet $\tilde{f}$ die Erzeuger von I_n auf 0 ab. Also faktorisiert die Abbildung über $S^n(V)$.

Das Argument im alternierenden Fall ist dasselbe. □

Wir schreiben $v_1 \bullet \cdots \bullet v_n$ bzw. $v_1 \wedge \cdots \wedge v_n$ für das Bild von $v_1 \otimes \ldots \otimes v_n$ in $S^n(V)$ bzw. $\Lambda^n(V)$. Das Produkt $\bullet$ ist kommutativ, $\wedge$ ist alternierend ($v \wedge v = 0$), insbesondere anti-kommutativ (d.h. $v_1 \wedge v_2 = -v_2 \wedge v_1$).

Die Konstruktion von symmetrischen und äußeren Potenzen ist natürlich, d.h. verträglich mit linearen Abbildungen.

Lemma 16.5. *Sei $f : V \to V'$ linear, $n \geq 1$. Dann gibt es eindeutige lineare Abbildungen*

$$S^n(f) : S^n(V) \to S^n(V'), \quad \Lambda^n(f) : \Lambda^n(V) \to \Lambda^n(V')$$

mit

$$v_1 \bullet \cdots \bullet v_n \mapsto f(v_1) \bullet \cdots \bullet f(v_n), \quad v_1 \wedge \cdots \wedge v_n \mapsto f(v_1) \wedge \cdots \wedge f(v_n).$$

Beweis: Die Abbildung

$$V^n \xrightarrow{f^n} V'^n \to S^n(V')$$

ist multilinear und symmetrisch. Daher faktorisiert sie wie angegeben. Dasselbe Argument funktioniert für äußere Potenzen. □

Um die Konstruktionen besser zu verstehen, berechnen wir Basen. Sei ab jetzt $e_1, \ldots, e_N$ eine Basis von V. Dann wird $S^n(V)$ erzeugt von Elementen der Form $e_{i_1} \bullet \cdots \bullet e_{i_n}$. Dabei können wir die Faktoren vertauschen, so dass ohne Einschränkung $i_1 \leq i_2 \leq \cdots \leq i_n$.
Ebenso wird $\Lambda^n(V)$ erzeugt von den Elementen der Form $e_{i_1} \wedge \cdots \wedge e_{i_n}$ mit $i_1 < i_2 < \cdots < i_n$. Daher ist $\Lambda^n(V) = 0$ für $n > N$. Besonders interessant ist also der Fall $n = N$. Dieser Vektorraum wird erzeugt von $e_1 \wedge \cdots \wedge e_n$.

Lemma 16.6. *Sei V Vektorraum mit Basis $e_1, \ldots, e_n$. Dann ist $e_1 \wedge \cdots \wedge e_n \neq 0$.*

Beweis: Wir definieren eine multilineare Abbildung $D : V^n \to k$, indem wir ihren Wert auf allen Tupeln von Basisvektoren angeben. Wir bilden das Tupel auf 0 ab, wenn ein Vektor doppelt vorkommt. Anderfass ist das Tupel von der Form $(e_{\sigma(1)}, \ldots, e_{\sigma(n)})$. Dann bilden wir es auf $\mathrm{sgn}(\sigma)$ ab.

Behauptung. *D ist alternierend.*

Seien $v_1, \ldots, v_n \in V$ mit $v_i = v_j$ für ein Paar $i \neq j$. Wir müssen zeigen, dass $D(v_1, \ldots, v_n) = 0$. Zur Vereinfachung der Notation nehmen wir $i = 1, j = 2$ an. Der allgemeine Fall geht genauso. Wegen der Multilinearität genügt es, den Fall zu betrachten, dass $v_3, \ldots, v_n$ Basisvektoren sind. Sei also $v_3 = e_{\sigma(3)}, \ldots, v_n = e_{\sigma(n)}$. Hier ist σ nicht notwendig eine Permutation, sondern nur eine Abbildung $\{3, \ldots, n\} \to \{1, \ldots, n\}$. Sei $v = v_1 = v_2 = \sum_{s=1}^n a_s e_s$. Dann ist

$$D(v, v, e_{\sigma(3)}, e_{\sigma(4)}, \ldots, e_{\sigma(n)}) = \sum_{s,t} a_s a_t D(e_s, e_t, e_{\sigma(3)}, e_{\sigma(4)}, \ldots, e_{\sigma(n)})$$

Ist σ nicht injektiv, so ist jeder der Summanden 0 und die Bedingung ist erfüllt. Sei nun σ injektiv. Dann lässt σ zwei Werte x und y aus. In der Summe verschwinden alle Summanden, außer denen zu $s = x$, $t = y$ und $s = y$,$t = x$. Der Wert ist

$$a_x a_y D(e_x, e_y, e_{\sigma(3)}, \ldots, e_{\sigma(n)}) - a_y a_x D(e_y, e_x, e_{\sigma(3)}, \ldots, e_{\sigma(n)})$$

Diese unterscheiden sich um eine Transposition, also haben die Werte verschiedene Vorzeichen und das Ergebnis ist wieder 0.

Die zu D gehörende lineare Abbildung bildet $e_1 \wedge \cdots \wedge e_n$ auf 1 ab, also ist der Vektor nicht 0. □

Bemerkung. Wir hätten auch einfach

$$D : V^n \to (k^n)^n \xrightarrow{\det} k$$

benutzen können. Unser Argument kam ohne die Existenz der Determinante aus, bzw. wir haben die Determinante explizit konstruiert.

Satz 16.7. *Sei V ein Vektorraum der Dimension N, $n \geq 1$. Dann sind die $e_{i_1} \wedge \ldots e_{i_n}$ mit $1 \leq i_1 < \cdots < i_n \leq N$ eine Basis von $\Lambda^n(V)$. Es gilt*

$$\dim \Lambda^n(V) = \binom{N}{n}.$$

Beweis: Wir haben bereits geklärt, dass es sich um ein Erzeugendensystem handelt. Wir zeigen nun die lineare Unabhängigkeit. Wir betrachten

$$\sum_{i_1 < \cdots < i_n} a_{i_1,\ldots,i_n} e_{i_1} \wedge \cdots \wedge e_{i_n} = 0$$

mit Koeffizienten $a_{i_1,\ldots,i_n} \in k$. Wir wählen ein Tupel von Indizes, zur Vereinfachung der Notation $1 < 2 < \cdots < n$. (Für $n > N$ ist nichts zu zeigen.) Sei $W = V/\langle e_{n+1}, \ldots, e_N \rangle$ und $\pi : V \to W$ die Projektion. Wir nutzen die Natürlichkeit der äußeren Potenz aus. Es gilt

$$0 = \Lambda^n(\sum a_{i_1,\ldots,i_n} e_{i_1} \wedge \cdots \wedge e_{i_n}) = \sum a_{i_1,\ldots,i_n} \pi(e_{i_1}) \wedge \ldots \pi e_{i_n}.$$

Es gilt $\pi(e_j) = 0$ für $j > n$, daher tragen nur Summanden bei, bei denen alle Indizes kleiner gleich n sind. Davon gibt es nur einen einzigen, $i_1 = 1, \ldots, i_n = n$. Daher folgt

$$0 = a_{1,\ldots,n} \pi(e_1) \wedge \cdots \wedge \pi(e_n).$$

Nach dem Lemma ist $\pi(e_1) \wedge \cdots \wedge \pi(e_n) \neq 0$ (W hat Dimension n und dies sind genau die Basiselemente). Daher folgt

$$0 = a_{1,\ldots,n}.$$

Genauso sieht man das Verschwinden aller anderen $a_{i_1,\ldots,i_n}$.

Für die Dimensionsberechnung müssen wir die Anzahl der Basiselemente abzählen. Es ist die Anzahl der n-elementigen Mengen in einer Menge mit N Elementen. Das ist gerade $\binom{N}{n}$. □

Korollar 16.8. *Sei V ein Vektorraum der Dimension n, $f \in \mathrm{End}(V)$. Dann ist*

$$\Lambda^n(f) : \Lambda^n(V) \to \Lambda^n(V)$$

Multiplikation mit $\det(f)$.

Beweis: Sei $e_1, \dots, e_n$ eine Basis. Dann ist

$$\Lambda^n(f)(e_1 \wedge \dots e_n) = f(e_1) \wedge \dots f(e_n).$$

Wir müssen dies als Vielfaches von $e_1 \wedge \dots \wedge e_n$ ausdrücken. Dafür drücken wir $f(e_j)$ als Linearkombination der e_i aus, d.h. wir bestimmen die Einträge der darstellenden Matrix. Durch die Multilinearität entsteht eine große n-fache Summe. In dieser verschwinden aber alle Summanden, in denen ein Basisvektor doppelt vorkommt. Durch Vertauschen der Einträge erhalten wir lauter Vielfache von $e_1 \wedge \dots e_n$. Als Vorfaktor erhalten wir die Determinante der darstellenden Matrix laut Leibniz-Formel. □

Bemerkung. Alternativ können wir $\Lambda^n(f)$ als Definition der Determinante benutzen. Es ist

$$\Lambda^n(f) \in \mathrm{End}(\Lambda^n(V)) = k,$$

denn die einzigen Endomorphismen eines eindimensionalen Vektorraums sind die Vielfachen der Identität. Oft benutzt man daher auch die Notation

$$\det(V) = \Lambda^n(V)$$

für die höchste äußere Potenz.
In LA 1 hatten wir die Determinante naiv als Abbildung

$$M_n(k) \to k$$

definiert. Jetzt sehen wir, dass es um zwei kanonische Konstruktionen geht:

$$V^n \to \det(V)$$

wobei $\det(V)$ ein Vektorraum der Dimension 1 ist; und

$$\det(f) : \det(V) \to \det(f)$$

für einen Endormorphismus f von V.

Alle Eigenschaften der Determinante können aus der neuen Definition hergeleitet werden.

Satz 16.9 (Produktformel)**.** *Sei V ein endlich-dimensionaler Vektorraum, $f, g \in \mathrm{End}(V)$. Dann gilt*

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

Beweis: Wir betrachten

$$\Lambda^n(V) \xrightarrow{\Lambda^n(g)} \Lambda^n(V) \xrightarrow{\Lambda^n(f)} \Lambda^n(V).$$

Man sieht sofort, dass die Komposition $\Lambda^n(fg)$ ist. □

Besonders interessant ist das Zusammenspiel von Dualität und äußeren Potenzen.

Satz 16.10. *Sei V ein Vektorraum, $n \geq 0$. Dann gibt es eine eindeutige lineare Abbildung*

$$*: \Lambda^n(V) \otimes \det(V^*) \to \Lambda^{N-n}(V^*)$$

so dass für jede Basis $e_1, \dots, e_N$ gilt

$$e_1 \wedge \dots \wedge e_n \otimes e_1^* \wedge \dots e_N^* \mapsto e_{n+1}^* \wedge \dots e_N^*$$

Sie ist ein Isomorphismus.

Die Abbildung heißt *Hodge-*-Operator* oder **-Produkt.*

Beispiel. Für $N = 1$ gibt es nur die Fälle $n = 0, 1$. Die Aussagen sind dann

$$k \otimes V^* \cong V^*, V \otimes V^* \cong k.$$

Beweis: Wir fixieren eine Basis $e_1, \dots, e_N$. Für $\sigma \in S_n$ definieren wir

$$*e_{\sigma(1)} \wedge \dots e_{\sigma(n)} \otimes e_1^* \wedge \dots e_N^* = \mathrm{sgn}(\sigma) e_{\sigma(n+1)}^* \wedge \dots \wedge e_{\sigma(N)}^*.$$

Damit ist die Abbildung eindeutig festgelegt. Die Formel ist verträglich mit Vertauschen der Basivektoren. Beide Seiten haben dieselbe Dimension $\binom{N}{n}$ und die Abbildung ist eine Bijektion von Basen, also ein Isomorphismus.

Wir müssen also die Unabhängikeit von der Wahl einer Basis überprüfen. Sei $f_1, \dots, f_n$ eine zweite Basis. Jeder Basiswechsel lässt sich zerlegen in eine Folge von elementaren Transformationen, daher genügt es zwei Fälle zu betrachten:

$$f_i = \begin{cases} e_i & i \neq i_0 \\ ae_i & i = i_0 \end{cases}, \quad f_i = \begin{cases} e_i & i \neq i_0 \\ e_{i_0} + be_j & i = i_0 \end{cases}$$

wobei $a \in k^*$, $b \in k$, $j \neq i_0$. Zur Vereinfachung der Notation betrachten wir $i_0 = 1$, $j = 2$. Der allgemeine Fall geht genauso. Ohne Einschränkung ist sogar $b = 1$.

Sei also $f_1 = ae_1$. Dann gilt für die duale Basis

$$f_1^* = a^{-1} e_1^*, f_2^* = e_2^*, \dots, f_N^* = e_N^*.$$

Wir betrachten das Bild von

$$f_{\sigma(1)} \wedge \dots \wedge f_{\sigma(n)} \otimes f_1^* \wedge \dots \wedge f_N^*$$

unter der linearen Abbildung $*$. Es gibt zwei Fälle: $1 \in \{\sigma(1), \dots, \sigma(n)\}$ oder $1 \in \{\sigma(n+1), \dots, \sigma(N)\}$. Im ersten Fall gilt

$$f_{\sigma(1)} \wedge \dots \wedge f_{\sigma(n)} \otimes f_1^* \wedge \dots \wedge f_N^* = ae_{\sigma(1)} \wedge \dots \wedge e_{\sigma(n)} \otimes a^{-1} e_1^* \wedge \dots \wedge e_N^*$$

und das Bild ist

$$\mathrm{sgn}(\sigma) e_{\sigma(n+1)}^* \wedge \dots \wedge e_{\sigma(N)}^* = f_{\sigma(1)}^* \wedge \dots \wedge f_{\sigma(N)}^*.$$

Im zweiten Fall gilt

$$f_{\sigma(1)} \wedge \cdots \wedge f_{\sigma(n)} \otimes f_1^* \wedge \cdots \wedge f_N^* = e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} \otimes a^{-1} e_1^* \wedge \cdots \wedge e_N^*$$

und das Bild ist

$$\operatorname{sgn}(\sigma) a^{-1} e^*_{\sigma(n+1)} \wedge \cdots \wedge e^*_{\sigma(N)} = \operatorname{sgn}(\sigma) f^*_{\sigma(1)} \wedge \cdots \wedge f^*_{\sigma(N)}.$$

Wir betrachten nun $f_1 = e_1 + e_2$. Dann gilt für die duale Basis

$$f_1^* = e_1^*, f_2^* = e_2^* - e_1^*, f_3^* = e_3^*, \ldots, f_N^* = e_N^*$$

und daher

$$f_1^* \wedge \ldots f_N^* = e_1^* \wedge \cdots \wedge e_N^*.$$

Diesmal gibt es vier Fälle. Sei $1 = \sigma(i)$, $2 = \sigma(j)$.
1. Fall: $i, j \leq n$. Durch Vertauschen erreichen wir $i = 1$, $j = 2$. Ohne Einschränkung der Allgemeinheit rechnen wir nur im Fall $N = 2$. Es ist

$$f_1 \wedge f_2 \otimes f_1^* \wedge f_2^* = (e_1 + e_2) \wedge e_2 \otimes e_1^* \wedge e_2^* = e_1 \wedge e_2 \otimes e_1^* \wedge e_2^*$$

und daher

$$* f_1 \wedge f_2 \otimes f_1^* \wedge f_2^* = 1.$$

2. Fall: $i \leq n$, $j \geq n + 1$. Durch Vertauschen und Weglassen von Faktoren erreichen wir den wesentlichen Fall $N = 2$, $n = 1$, $i = 1$, $j = 2$, also

$$* f_1 \otimes f_1^* \wedge f_2^* = *(e_1 + e_2) \otimes e_1^* \wedge e_2^*) = e_1^* - e_2^* = f_2^*.$$

3. Fall: $i > n$, $j \leq n$. Durch Vertauschen und Weglassen von Faktoren erreichen wir den wesentlichen Fall $N = 2$, $n = 1$, $i = 2$, $j = 1$, also

$$* f_2 \otimes f_1^* \wedge f_2^* = * e_2 \otimes e_1^* \wedge e_2^* = -e_1^* = -f_1^*.$$

4. Fall: $i, j > n$. Der wesentliche Fall ist $n = 0$, $N = 1$, also

$$* f_1^* \wedge f_2^*) = *(e_1^* \wedge e_2^*) = e_1^* \wedge e_2^* = f_1^* \wedge f_2^* ..$$

In jedem Fall ist die Zuordnung verträglich mit dem Basiswechsel. □

Speziell für $n = N$ erhalten wir

$$\det(V) \otimes \det(V^*) \cong k \Rightarrow (\det(V))^* \cong \det(V^*).$$

Korollar 16.11. *Sei V endlich-dimensionaler Vektorraum, $f \in \operatorname{End}(V)$. Dann gilt*

$$\det(f) = \det(f^*).$$

Beweis: Wenn f kein Isomorphismus ist, so sind beide Seiten 0. Sei also f ein Isomorphismus, $e_1, \dots, e_N$ eine Basis. Dann ist $f(e_1), \dots, f(e_N)$ eine andere Basis mit dualer Basis $f^{*-1}(e_1^*), \dots, f^{*-1}(e_N)$.

$$f(e_1) \wedge \dots \wedge f(e_N) \otimes f^{*-1}(e_1^*) \wedge \dots \wedge f^{-1} * (e_N^*) \mapsto 1.$$

Andererseits ist die linke Seite auch

$$\det(f)\det(f^*)^{-1} e_1 \wedge \dots \wedge e_N \otimes e_1^* \wedge \dots \wedge e_N^* \mapsto \det(f)\det(f^*)^{-1}.$$

□

Das Kreuz-Produkt

Sei $k = \mathbb{R}$ und V endlich-dimensional. Die Erzeuger von $\det(V)$ unterscheiden sich um einen Faktor in $\mathbb{R}^*$. Dieser Faktor kann positiv oder negativ sein.

Definition 16.12. *Sei V ein reeller Vektorraum. Eine* Orientierung *auf V ist definiert als die Wahl einer Äquivalenzklasse von Erzeugern von* $\det(V)$*, die sich um einen positiven Faktor unterscheiden.*

Je zwei Basen von V definieren die gleiche Orientierung, wenn die Basiswechselmatrix positive Determinante hat.
Sei nun noch spezieller V euklidisch und orientiert. Je zwei Orthonormalbasen der gleichen Orientierung unterscheiden sich um eine Basiswechselmatrix in $SO(n)$. Sie definieren denselben Erzeuger von $\det(V)$, die wir benutzen, um $\det(V)$ und $\det(V^*)$ mit k zu identifizieren. Außderdem induziert das Skalarprodukt einen Isomorphismus $V \to V^*$.
Der Hodge-*-Operator definiert dann Isomorphismen

$$\Lambda^n(V) \cong \Lambda^n(V^*) \to \Lambda^{N-n}(V^*) \cong \Lambda^{N-n}(V).$$

Wir spezialisieren auf $N = 3$, $n = 2$ under halten

$$\Lambda^2(V) \cong V.$$

Insbesondere erhalten wir eine Multiplikation

$$* \circ \wedge : V \times V \to \Lambda^2(V) \to V^* \to V.$$

Sie heißt *Vektorprodukt* oder *Kreuzprodukt*.
Wir rechnen explizit aus. Sei e_1, e_2, e_3 eine orientiere Orthonormalbasis. Es gilt

$$\begin{aligned}
(e_1, e_2) &\mapsto e_1 \wedge e_2 \mapsto e_3^* \mapsto e_3 \\
(e_1, e_3) &\mapsto e_1 \wedge e_3 \mapsto -e_2^* \mapsto e_2 \\
(e_2, e_3) &\mapsto e_2 \wedge e_3 \mapsto e_1^* \mapsto e_1
\end{aligned}$$

Dann ist diese Abbildung explizit:

$$\begin{pmatrix} x_1 \\ x_2 \\ +x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$$

Die Rechenregeln folgen aus den Rechenregeln für $\wedge$, d.h. die Abbildung ist bilinear und alternierend.

Lemma 16.13. *Sei V orientierter Vektorraum mit Skalarprodukt s. Dann gilt für alle $x, y \in V$ gilt*

$$s(x, x \times y) = s(y, x \times y) = 0.$$

Beweis: Wir berechnen $s(x, x \times y)$. Sind x und y linear abhängig, so ist $x \times y = 0$, da $\times$ alternierend ist. Also sind sie ohne Einschränkung linear unabhängig. Es ist $y = y' + ax$ so y' orthogonal zu x (Gram-Schmidt). Da $\times$ alternierend ist, gilt $x \times y = x \times y'$. Sei also ohne Einschränkung y senkrecht auf x. Nach Skalieren erreichen wir, dass $e_1 = x, e_2 = y$ orthonormal sind. Es gibt einen eindeutigen normierten Vektor e_3, so dass e_1, e_2, e_3 eine orientiere Orthonormalbasis von V ist. Daher gilt $e_3 = x \times y$, insbesondere steht er auf e_1 senkrecht. □

Bemerkung. Das Kreuzprodukt spielt in den Formeln der Physik eine große Rolle, da es dort oft um $\mathbb{R}^3$ als orientierten euklidischen Vektorraum geht. In der Mathematik ist $n = 3$ nur ein Spezialfall. Das Kreuzprodukt verallgemeinert sich nicht, bzw. in der Form wie wir es oben diskutiert haben.

Differentialformen

Äußere Potenzen spielen eine große Rolle in der höherdimensionalen Differentiation und Integration. Hier ist nicht der Ort, diese Theorie vollständig zu entwickeln, aber wir wollen die Situation skizzieren.

Wir beginnen mit reellen differentierbaren Funktionen auf offenen Teilmengen $U \subset \mathbb{R}^n$. Einer Funktion $f : U \to \mathbb{R}$ ordnen wir eine Ableitung zu

$$f \mapsto \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right).$$

Diese Ableitung ist aber nicht unabhängig von der Wahl der Koordinaten. Statt dessen gilt die Kettenregel. Das ist schlecht, wenn wir von Teilmengen von $\mathbb{R}^n$ zu *Mannigfaltigkeiten* übergehen wollen. Eine Manngifaltigkeit ist ist ein topologischer Raum M, so dass es für jeden Punkt $P \in M$ eine Umgebung gibt, die mit einer offenen Teilmenge von $\mathbb{R}^n$ identifiziert wird. Mit anderen Worten: wir können Koordinaten einführen.

Beispiel. Sie Kugel $S^n = \{x \in \mathbb{R}^{n+1} | \|x\| = 1\}$ ist eine Untermannigfaltigkeit der Dimension n (Satz über implizite Funktionen).

Eine (lokal differenzierbare) Funktion $f : M \to \mathbb{R}$ hat keinen globalen Ableitungsvektor. Wir lösen das Problem, indem wir die *Differentialform*

$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n$$

definieren. Die Kettenregel sorgt dafür, dass dies ein wohldefinierter Ausdruck ist, unabhängig von der Kooridnatenwahl. Aber was ist eine Differentialform? Zu jedem Punkt P von M gehört der Vektorraum T_pM der Tangentialvektoren. Für eine Untermannigfaltigkeit sind das alle Tangentialrichtungen im umgebenden Raum. Sie organisieren sich in das *Tangentialbündel* $TM \to M$. Das Urbild von P ist genau T_PM. Alle Operationen mit Vektorräumen lassen sich auch mit Vektorbündeln definieren. Insbesondere gibt es $T^*M \to M$, das *Kotangentialbündel*. Das Urbild von $P \in M$ ist der duale Vektorraum von T_PM.

Definition 16.14. *Eine* Differentialform *auf M ist ein Abbildung $s : M \to T^*M$, so dass $s(P)$ in T_PM^* liegt.*

Die Wahl von Koordinaten idenfiziert eine Teilmenge von M mit einer offenen Teilmenge des $\mathbb{R}^n$, alle Tangentialräume mit $\mathbb{R}^n$ und das Tangentialbündel mit $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$. Die Ausdrücke $dx_1, \ldots, dx_n$ sind eine Basis des dualen Raums.

Die Ableitung einer Differentialform ist eine 2-Form. Diese ist ein Schnitt von $\Lambda^2(T^*M)$. Genauso geht es mit höheren Ableitungen weiter. Für $n = \dim M$ erhalten wir ein Geradenbündel $\Lambda^n(T^*M) \to M$, d.h. das Urbild von P ist ein eindimensionaler Vektorraum. Die Schnitte haben lokal die Form

$$f dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$$

für Funktionen f. Solche Ausdrücke kann man dann integrieren. Die äußeren Potenzen verwalten die Transformationsformeln genau richtig.

Der oben definierte Hodge-*-Operator wird dann in der Theorie der orientierten Riemannschen Mannigfaltigkeiten benutzt. Dann hat jedes T_pM die Struktur eines orientieren euklidischen Vektorraums.

Kapitel 17

Algebren

Definition 17.1. *Sei k ein Körper. Eine k-Algebra ist eine Menge A zusammen mit*

- *einer Addition: $+ : A \times A \to A$,*
- *einer Multiplikation: $\cdot : A \times A \to A$,*
- *einer Skalarmultiplikation: $\cdot : k \times A$,*

so dass, $(A, +, \cdot)$ ein Ring ist, $(A, +)$ mit der Skalarmultiplikation ein k-Vektorraum und

$$\lambda(ab) = (\lambda a)b = a(\lambda b) \quad \text{für alle } \lambda \in k,\ a, b \in A.$$

Die Algebra heißt kommutativ*, falls $ab = ba$ für alle $a, b \in A$. Sie heißt* Algebra mit Eins*, falls es ein Element $1 \in A$ gibt mit $1a = a1 = a$ für alle $a \in A$.*

Beispiel. (i) $M_n(k)$ mit der Addition und Multiplikation von Matrizen ist eine k-Algebra mit Eins (LA 1).

(ii) Sei V ein Vektorraum. Dann ist $\mathrm{End}(V)$ mit der Addition und Verknüpfung von Endomorphismen eine k-Algebra mit Eins.

(iii) $k[X]$ ist eine kommutative Algebra mit Eins.

(iv) Der Polynomring in n Variablen $k[X_1, \ldots, X_n]$ ist eine kommutative Algebra mit Eins.

(v) Sei M eine Menge $\mathrm{Abb}(M, k)$ mit der Addition und Multiplikation von Funktionen ist eine kommutative k-Algebra mit Eins (1 ist die konstante Funktion 1).

(vi) Sei U ein topologischer Raum, z.B., $U \subset \mathbb{R}^n$ offen und $C(U, \mathbb{R})$ die Menge der stetigen Funktionen. Dies ist eine kommutative $\mathbb{R}$-Algebra mit 1.

(vii) Sei $C_c(\mathbb{R}, \mathbb{R})$ die Menge der stetigen Funktionen mit kompaktem Träger, d.h. die Menge der $f : \mathbb{R} \to \mathbb{R}$ stetig, so dass es $N > 0$ gibt mit $f(x) = 0$ für alle $|x| > N$. Dies ist eine kommutative k-Algebra ohne Eins. (Das Einselement wäre die konstante Funktion 1, diese hat aber keinen kompakten Träger.)

(viii) In der Funktionalanalysis werden solche Algebren zusätzlich mit einer Norm versehen.

Wir wollen nun die Tensor-Algebra $T^*(V)$ für einen Vektorraum V einführen. Dafür sei $T^i(V) = V^{\otimes i}$ das i-fache Tensorprodukt von V mit sich selbst. Wegen der Assoziativität des Tensorproduktes kommt es auf die Klammerung nicht an. Für $i = 0$ setzen wir $T^0(V) = k$. Tensorieren defininiert Multiplikationen

$$\begin{gathered} T^n(V) \times T^m(V) \to T^{n+m}(V) \\ (v_1 \otimes \ldots \otimes v_n, w_1 \otimes \ldots \otimes w_m) \mapsto v_1 \otimes \ldots v_n \otimes w_1 \otimes \ldots \otimes w_m). \end{gathered}$$

Speziell für $n = 0$ oder $m = 0$ benutzen wir die skalare Multiplikation.

Definition 17.2. *Sei V ein k-Vektorraum.*

(i) Die Tensor-Algebra $T^*(V)$ *über V ist der Vektorraum*

$$\bigoplus_{i=0}^{\infty} T^i(V) \quad \textit{mit} \quad T^(V) = V^{\otimes i}$$

(hier ist $T^0(V) = k$) mit dem Tensorprodukt als Multiplikation.

(ii) Die symmetrische Algebra $S^*(V)$ *über V ist der Vektorraum*

$$\bigoplus_{i=0}^{\infty} S^i(V)$$

mit $\bullet$ als Multiplikation.

(iii) Die äußere Algebra $\Lambda^*(V)$ *über V ist der Vektorraum*

$$\bigoplus_{i=0}^{\dim V} \Lambda^i(V)$$

mit $\wedge$ als Multiplikation.

Lemma 17.3. *$T^*(V)$, $S^*(V)$ und $\Lambda^*(V)$ sind k-Algebren mit Eins.*

Beweis: Nach Definition ist $T^*(V)$ ein Vektorraum, insbesondere ist $(T^*(V), +)$ eine abelsche Gruppe. Wir überprüfen die Assoziativität der Multiplikation. Es genügt, dies für $v \in T^n(V)$, $w \in T^m(V)$, $u \in T^l(V)$ nachzurechnen. Dann folgt es aus der Assoziativitätsregel des Tensorproduktes. Die Verträglichkeit von +

und $\otimes$ folgt aus der Distributivität und Bilinearität des Tensorproduktes. Die Verträglichkeit von Multiplikation und skalarer Multiplikation folgt ebenfalls aus der Bilinearität von $\otimes$. Das Einselement ist $1 \in T^0(V)$.
Dieselben Argumente gelten auch für die beiden andere Fälle. □

Definition 17.4. *Ein* k-Algebrenhomomorphismus *ist eine k-lineare Abbildung $f : A \to B$ zwische k-Algebren, die zusätzlich ein Ringhomomorphismus ist.*
Ein Homomorphismus von Ringen mit Eins erfüllt zusätzlch $f(1) = 1$.

Beispiel. (i) Sei V ein k-Vektorraum der Dimension n. Die Wahl einer Basis B induziert einen Algebrenisomorphismus

$$\phi : \mathrm{End}(V) \to M_n(k) \quad f \mapsto M_B^B(f).$$

(ii) Sei $\alpha \in k$. Dann definiert die Auswertungabbildung

$$k[X] \to k \quad P \mapsto P(\alpha)$$

einen Algebrenhomomorphismus. Lassen wir α variieren, so erhalten wir

$$\mathrm{ev} : k[X] \to \mathrm{Abb}(k, k) \quad P \mapsto (\alpha \mapsto P(\alpha)).$$

Auch diese Abbildung ist ein Algebrenhomomorphismus. Für Körper mit unendlich vielen Elementen ist er injektiv (LA 1).

(iii) Sei V ein k-Vektorraum, $f \in \mathrm{End}(V)$. Dann definiert

$$k[X] \to \mathrm{End}(V) \quad P \mapsto P(f)$$

einen k-Algebrenhomomorphismus. Nach Definition gilt $1 \mapsto \mathrm{id}$. Er ist nicht injektiv. Nach dem Satz von Cayley-Hamilton gilt $\chi_f(f) = 0$.

(iv) Sei V ein k-Vektorraum. Dann sind die natürlichen Projektionen

$$T^*(V) \to S^*(V), \quad T^*(V) \to \Lambda^*(V)$$

Homomorphismen von Algebren mit Eins.

Satz 17.5 (Universelle Eigenschaft der Tensor-Algebra)**.** *Sei V ein k-Vektorraum. Jede lineare Abbildung*

$$f : V \to A$$

in eine k-Algebra A mit Eins setzt sich eindeutig fort zu einem Homomorphismus

$$F : T^*(V) \to A$$

von Algebren mit Eins.

Beweis: Die lineare Abbildung $f : V \to A$ definiert für $n \geq 1$ multilineare Abbildungen

$$f_n : V \times \cdots \times V \to A \quad (x_1, \dots, x_n) \mapsto f(x_1)f(x_2)\dots f(x_n).$$

Nach der universellen Eigenschaft des Tensorproduktes faktorisiert sie über eine eindeutige lineare Abbildung

$$F_n : V^{\otimes n} \to A.$$

Wir definieren $F_0 : k \to A$ als $\alpha \mapsto \alpha 1_A$. Wir definieren damit

$$F = (F_0, F_1, \dots) : \bigoplus V^{\otimes n} \to A.$$

Diese Abbildung ist nach Konstruktion linear, also verträglich mit $+$ und der skalaren Multiplikation. Die Verträglichkeit mit der Multiplikation ist leicht. Das Argument benutzt, dass $A \times A \to A$ ebenfalls bilinear ist. $\square$

Bemerkung. Ist $b_1, \dots, b_n$ eine Basis von V, so sind die Ausdrücke $b_{i_1} \otimes b_{i_2} \otimes b_{i_n}$ für $n \geq 0$, $i_j \in \{1, \dots, n\}$ eine Basis von $T^*(V)$. Der Algebrenhomomorphismus des Satzes bildet ihn ab auf $f(b_{i_1})f(b_{i_2})\dots f(b_{i_n})$, wie es die Verträglichkeit mit der Multiplikation verlangt.

Speziell für $n = 1$ und $f : V \to k[X]$ mit $b \mapsto X$ erhalten wir einen Algebrenisomorphismus

$$T^*(V) \to k[X].$$

Sei nun A kommutativ, wie vorher $V \to A$ linear. Dann liegt $v \otimes w - w \otimes v \in T^2(V)$ im Kern der induzierten Abbildung $T^*(V) \to A$.

Satz 17.6 (Universelle Eigenschaft)**.** *Die symmetrische Algebra ist kommutativ. Ist A eine kommutative Algebra mit Eins, $f : V \to A$ linear, so existiert ein eindeutiger Homomorphismus $F : S(V) \to A$ von Algebren mit Eins, der f fortsetzt.*

Beweis: Wir erinnern uns an die Definition:

$$S^n(V) = T^n(V)/I_n$$

wobei I_n als Vektorraum von Elementen der Form $v_1 \otimes \dots v_n - v_{\sigma(1)} \otimes \dots v_{\sigma(n)}$ für $v_1, \dots, v_n \in V$ und $\sigma \in S_n$ erzeugt wird.

Nun überprüfen wir die universelle Eigenschaft. Wir haben bereits einen Algebrenhomomorphismus $F : T^*(V) \to A$. Wir wollen zeigen, dass er über den Quotientenvektorraum $S^*(V)$ faktorisiert. Nach dem Homomorphiesatz müssen wir dafür überprüfen, dass $F(I_n) = 0$ für alle n. Es genügt die Erzeuger zu betrachten. Da A kommutativ ist, verschwinden deren Bilder in A. Wir erhalten wie gewünscht den Wert 0.

Die induzierte lineare Abbildung $S^*(V) \to A$ ist verträglich mit der Multiplikation, weil die Multiplikation auf $S^*(V)$ die Multiplikation von Restklassen ist. $\square$

Beispiel. Sei $e_1, \dots, e_n$ eine Basis von V. Wir definieren $V \to k[X_1, \dots, X_n]$ durch $e_i \mapsto X_i$. Aus der universellen Eigenschaft der symmetrischen Algebra erhalten wir einen Homomorphismus

$$S^*(V) \to k[X_1, \dots, X_n].$$

Man überprüft gradweise, dass es sich um einen Isomorphismus handelt. Die Basiselemente $e_{i_1} \bullet e_{i_2} \cdots \bullet e_{i_m}$ stehen in Bijektion zu den Monomen $X_{i_1} X_{i_2} \dots X_{i_m}$. Insbesondere ist $S^*(k) \cong k[X]$.

Es ließe sich noch mehr über die symmetrische Algebra sagen, aber tatsächlich sind wir mehr an der äußeren Algebra interessiert.

Dort gilt ebenfalls eine Vertauschungsregel:

Lemma 17.7. *Seien $x \in \Lambda^n(V)$, $y \in \Lambda^m(V)$. Dann gilt*

$$x \wedge y = (-1)^{nm} y \wedge x.$$

Beweis: Ohne Einschränkung ist $x = v_1 \wedge \cdots \wedge v_n$, $y = w_1 \wedge \dots w_m$. Wir tauschen w_1 mit v_{n-1}, dann mit v_{n-2} etc. Dies ergibt einen Faktor von $(-1)^n$. Dann wiederholen wir den Prozess für $w_2, \dots, w_m$. Also erhalten wir m mal das Vorzeichen $(-1)^n$, zusammen $(-1)^{nm}$. □

Definition 17.8. *Eine* graduierte Algebra *ist eine k-Algebra der Form*

$$A = \bigoplus_{i=0} \infty A_i$$

wobei für alle $x \in A_i$, $y \in A_j$ gilt $xy \in A_{i+j}$. Die Elemente von A_i heißen homogen vom Grad *i.*

Eine graduierte Algebra heißt graduiert-kommutativ*, wenn*

$$xy = (-1)^{nm} \quad \text{für alle } x \in A_n, y \in A_m.$$

Beispiel. $S^*(V)$ und $k[X_1, \dots, X_n]$ sind graduiert und kommutativ, aber nicht graduiert-kommutativ. $\Lambda^*(V)$ ist graduiert-kommutativ.

Satz 17.9 (Universelle Eigenschaft der äußeren Algebra)**.** *Sei $A = \bigoplus_{i=0}^{\infty} A_i$ eine graduiert-kommutative Algebra mit Eins in der $a^2 = 0$ für alle $a \in A_1$, $f : V \to A_1$ eine lineare Abbildung. Dann setzt sich f fort zu einem eindeutigen Homomorphismus von Algebren mit Eins*

$$F : \Lambda^*(V) \to A,$$

der die Graduiering respektiert, also $F(\Lambda^n(V)) \subset A_n$

Beweis: Genau wie für die symmetrische Algebra. Die Bedingung an A_1 garantiert, dass die Abbildung

$$A_1^n \to A_n$$

alternierend ist. □

Bemerkung. Wenn $0 \neq 2$ in k, dann die Bedingung an A_1 eine Konsequenz der Vorzeichenregel: $a^2 = (-1)^{1 \cdot 1} a^2 \Rightarrow 2a^2 = 0$.